Course Notes

Ecole Polytechnique Fédérale de Lausanne

Mathematics Department

Ergodic theory and its applications to number theory

Author
Vincent Dumoncel

Professor
Florian Richter

July 15, 2023

Contents

1. Measure-preserving systems

In this first section, we present the basic objects of ergodic theory, and define our running examples for the entire course.

1.1 Definitions and examples

The first object to consider is what are called *measure-preserving transformation*. For this, fix a probability space $(X, \mathcal{A}, \mu)$. Recall that the *push-forward* of μ under any measurable map $T \colon X \longrightarrow X$ is the measure $T_*\mu$ on X defined by

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

for all $A \in \mathcal{A}$, where $T^{-1}A := \{x \in X : Tx \in A\}$.

Definition 1.1. The map $T \colon X \longrightarrow X$ is measure-preserving if $T_*\mu = \mu$.

Thus T is measure-preserving if $\mu(T^{-1}A) = \mu(A)$ for all $A \in \mathcal{A}$.

Definition 1.2. A measure-preserving system is a quadruple $(X, \mathcal{A}, \mu, T)$ where $(X, \mathcal{A}, \mu)$ is a probability space and $T \colon X \longrightarrow X$ is measure-preserving.

Here are some examples of measure-preserving systems.

Example 1.3. (i) Let $X = \{x\}$ be a singleton, with $\mathcal{A} = \{\emptyset, \{x\}\}$ and $\mu(\emptyset) = 0$, $\mu(\{x\}) = 1$. Let $T = \mathrm{Id}_X$. Then $(X, \mathcal{A}, \mu, T)$ is a measure-preserving system.

(ii) More generally, if $(X, \mathcal{A}, \mu)$ is an arbitrary probability space and $T = \mathrm{Id}_X$, then $(X, \mathcal{A}, \mu, T)$ is a measure-preserving system, called the *identity system*.

(iii) Consider an integer $m \geq 2$, and $X = \{0, 1, \ldots, m-1\}$, which can be identified with the cyclic group of order m. Let $\mathcal{A} = \mathcal{P}(X)$, $\mu(\{k\}) = \frac{1}{m}$ for all $k = 0, \ldots, m-1$. Define the transformation $T \colon X \longrightarrow X$ by

$$T(k) := k + 1 \mod m$$

for all $k = 0, \ldots, m-1$. Then T is clearly measure-preserving, and $(X, \mathcal{A}, \mu, T)$ is a measure-preserving system. It is called the *rotation on m points*.

(iv) Let $X = [0, 1]$, endowed with the Borel $\sigma-$algebra and λ the Lebesgue measure. Fix $\alpha \in \mathbb{R}$, and define the map

$$T(x) := x + \alpha \mod 1.$$

The translation invariance of λ implies that T preserves the measure. Alternatively, we can identify X with the compact group $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. The Lebesgue measure on $[0, 1]$ is identified with the Haar measure on $\mathbb{T}$, and T becomes the map

$$T(x) = x + \tilde{\alpha}$$

where $\tilde{\alpha} = x + \mathbb{Z} \in \mathbb{T}$. Lastly, $\mathbb{T}$ is isometrically isomorphic to the circle $\mathbb{S}^1 \subset \mathbb{C}$, viewed as a group under multiplication. Under this identification, one has

$$T(z) = z e^{2\pi i \alpha}.$$

This system is usually called a *circle rotation*.

(v) More generally, if X is a compact abelian group equipped with its Borel $\sigma-$algebra and the Haar measure m_X, then for any $\alpha \in X$ the map $T(x) = x + \alpha$ preserves m_X, so that $(X, \mathcal{A}, \mu, T)$ is a measure-preserving system, called a *group rotation*.

In the last example, the fact that T preserves the Haar measure follows from the next more general result.

Lemma 1.4. Let X be a compact abelian group, and $T\colon X \longrightarrow X$ be a surjective endomorphism. Then T preserves the Haar measure m_X.
In particular, $(X, \mathcal{B}(X), m_X, T)$ is a measure-preserving system.

Proof. Consider on X the measure $\mu := T_* m_X$. Fix $A \in \mathcal{B}(X)$, and $x \in X$. Since T is surjective, there is $y \in X$ such that $T(x) = y$, and we compute that

$$\mu(A + x) = m_X(T^{-1}(A + x)) = m_X(T^{-1}A + y) = m_X(T^{-1}A) = \mu(A).$$

Thus μ is invariant by left translation, and the uniqueness of the Haar measure on X forces $\mu = m_X$, which means T is measure-preserving. $\square$

Let us provide additional examples.

Example 1.5. (i) Take $(X, \mathcal{A}, \lambda)$ to be the unit interval $[0, 1]$ equipped with its Borel $\sigma-$algebra and the Lebesgue measure. The *doubling-map* is the transformation T of X defined as

$$T(x) := 2x \mod 1.$$

It preserves the Lebesgue measure, since for a closed interval $[a, b] \subset [0, 1]$, one has

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

and thus $\mu(T^{-1}[a, b]) = \mu([a, b])$. Closed intervals of this form generate the Borel $\sigma-$algebra on $[0, 1]$, hence T is measure-preserving.

(ii) Let $X = \{0,1\}^{\mathbb{N}}$ be the space of infinite sequences of 0 and 1. Give $\{0,1\}$ the discrete topology, so that X is a compact topological space. Let $\mathcal{B}(X)$ be its Borel $\sigma-$algebra. Define a probability measure μ_0 on $\{0,1\}$ by $\mu_0(\{0\}) = 1-p$, $\mu_0(\{1\}) = p$, $p \in (0,1)$, and let $\mu = \mu_0^{\mathbb{N}}$ be the product measure on X. Equivalently, μ is the unique measure on X satisfying

$$\mu(\{(x_n)_{n\geq 0} \in X : x_0 = a_0, \ldots, x_m = a_m\}) = \prod_{i=0}^{m} \mu_0(\{a_i\})$$

for every $m \in \mathbb{N}$ and $a_0, \ldots, a_m \in \{0,1\}$. Define the *left-shift* as

$$T((x_n)_{n\geq 0}) := (x_{n+1})_{n\geq 0}$$

for all $(x_n)_{n\geq 0} \in X$. The quadruple $(X, \mathcal{B}(X), \mu, T)$ is a measure-preserving system, called a *Bernoulli shift*.

(iii) Given two measure-preserving systems $(X, \mathcal{A}, \mu, T)$, $(Y, \mathcal{B}, \nu, S)$, we define their product to be the space $X \times Y$ with the product $\sigma-$algebra $\mathcal{A} \otimes \mathcal{B}$, the product measure $\mu \otimes \nu$, and the transformation $T \times S$ defined as

$$(T \times S)(x,y) := (Tx, Sy)$$

for all $(x,y) \in X \times Y$.

1.2 Recurrence

Here is the first important result in ergodic theory. It is known as the *Poincaré's recurrence theorem*.

Theorem 1.6. Let $(X, \mathcal{A}, \mu, T)$ be a measure preserving system, and let $A \in \mathcal{A}$ with $\mu(A) > 0$. Then there exists $n \in \mathbb{N}$ so that

$$\mu(A \cap T^{-n}A) > 0.$$

Proof. The sets $A, T^{-1}A, T^{-2}A, \ldots$ all have the same measure $\mu(A) > 0$, and all lie in X which has measure 1. Therefore we must have $\mu(T^{-i}A \cap T^{-j}A) > 0$ for some $i < j$. Setting $n := j - i$, we get

$$\mu(A \cap T^{-n}A) = \mu(T^{-i}(A \cap T^{-n}A)) = \mu(T^{-i}A \cap T^{-j}A) > 0$$

as claimed. □

As a consequence, we obtain the next corollary.

Corollary 1.7. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system, and $A \in \mathcal{A}$. Then for $\mu-$almost every $x \in A$, there exists $n \in \mathbb{N}$ so that $T^n x \in A$.

Proof. Consider the set $B := \{x \in A : \forall n \in \mathbb{N},\ T^n x \notin A\} \subset A$. First, observe that $B \in \mathcal{A}$ since we can write

$$B = \bigcap_{n \in \mathbb{N}} \{x \in A : T^n x \notin A\} = \bigcap_{n \in \mathbb{N}} (A \cap T^{-n}(A^c))$$

and $A^c \in \mathcal{A}$, so $T^{-n}(A^c) \in \mathcal{A}$ because T^n is measurable for all $n \geq 0$. A $\sigma-$algebra being closed under countable unions, we indeed have $B \in \mathcal{A}$. To prove the corollary, we thus have to prove that $\mu(B) = 0$. Suppose, towards a contradiction, that $\mu(B) > 0$. By Poincaré's recurrence theorem, there exists $k \in \mathbb{N}$ such that

$$\mu(B \cap T^{-k} B) > 0.$$

In particular, $B \cap T^{-k} B \neq \emptyset$, so pick $x \in B \cap T^{-k} B$. Then $x \in B$, so $T^n x \notin A$ for every $n \in \mathbb{N}$. On the other hand, $x \in T^{-k} B$, so $T^k x \in B$, in particular $T^k x \in A$. This is a contradiction, and therefore $\mu(B) = 0$. □

In words, this result says that almost every point of A has an orbit visiting A at least one time.

These two results can be significantly strengthened. For instance, if X is a compact metric space, every point returns arbitrary close to its initial position.

Proposition 1.8. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system, with X compact metric and $\mathcal{A} = \mathcal{B}(X)$ its Borel $\sigma-$algebra. Then, for $\mu-$a.e. $x \in X$, we have

$$\inf_{n \geq 1} \mathrm{d}(x, T^n x) = 0.$$

The points $x \in X$ with this property are called *recurrent*. By the previous proposition, $x \in X$ is recurrent if and only if there exists a subsequence $(T^{n_k} x)_{k \in \mathbb{N}}$ of $(T^n x)_{n \in \mathbb{N}}$ converging to x.

Proof. Let $A \subset X$ be the set of non-recurrent points. We will prove $\mu(A) = 0$. Since X is metric compact, it is second countable, and we fix $(B_n)_{n \in \mathbb{N}}$ a basis for its topology. Let $x \in A$. By definition, this means $\inf_{n \geq 1} \mathrm{d}(x, T^n x) > 0$, so there is $\varepsilon > 0$ such that $\mathrm{d}(x, T^n x) \geq \varepsilon$, for all $n \geq 1$. In particular, $T^n x \notin B(x, \varepsilon)$ for all $n \geq 1$. Now let $k \in \mathbb{N}$ be such that $x \in B_k \subset B(x, \varepsilon)$. By Corollary 1.7, we find a full measure subset $C_k \subset B_k$ such that for all $y \in C_k$, there is $m(k) \in \mathbb{N}$ with $T^{m(k)} y \in C_k$. Thus $x \in B_k \setminus C_k$, because if $x \in C_k$, then $T^{m(k)}x \in C_k \subset B(x,\varepsilon)$ for some $m(k) \in \mathbb{N}$, but we already proved the orbit of x avoids $B(x,\varepsilon)$. We then have

$$A \subset \bigcup_{k=0}^{\infty} B_k \setminus C_k$$

and since $\mu(B_k \setminus C_k) = 0$ for all $k \in \mathbb{N}$, this implies $\mu(A) = 0$, and we are done. □

Here is now a stronger version of Poincaré's recurrence theorem.

Theorem 1.9. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system, and $A \in \mathcal{A}$. For any $\varepsilon > 0$, there exists $n \in \mathbb{N}$ such that

$$\mu(A \cap T^{-n}A) \geq \mu(A)^2 - \varepsilon.$$

Proof. Fix $A \in \mathcal{A}$ with $\mu(A) > 0$. Denote, for $k \geq 1$, $A_k := T^{-k}A$. Since T is measure-preserving, $\mu(A_k) = \mu(A)$ for all $k \geq 1$, and we denote $\alpha := \mu(A)$. Let now $n \geq 1$, and set $f := \sum_{k=1}^{n} \mathbf{1}_{A_k}$. On the one hand, we compute that

$$\int_X f \, \mathrm{d}\mu = \sum_{k=1}^{n} \mu(A_k) = n\alpha$$

and on the other hand, we have

$$\int_X f^2 \, \mathrm{d}\mu = \int_X \left(\sum_{k=1}^{n} \mathbf{1}_{A_k} + 2 \sum_{1 \leq i < j \leq n} \mathbf{1}_{A_i} \mathbf{1}_{A_j} \right) \mathrm{d}\mu = n\alpha + 2 \sum_{1 \leq i < j \leq n} \mu(A_i \cap A_j).$$

To reach a contradiction, suppose there is $\varepsilon > 0$ such that $\mu(A \cap A_k) \leq \mu(A)^2 - \varepsilon$ for all $k \geq 1$. This implies in particular that $\mu(A_i \cap A_j) \leq \mu(A)^2 - \varepsilon$ for all $i \neq j$. Therefore, by the Cauchy-Schwartz inequality and the computations above, we obtain

$$n^2\alpha^2 = \left(\int_X f \, \mathrm{d}\mu \right)^2 \leq \int_X f^2 \, \mathrm{d}\mu \leq n\alpha + 2(n-1)n(\mu(A)^2 - \varepsilon)$$

for all $n \geq 1$. This is a contradiction, since this last inequality does not hold for n large enough, for instance for $n > \frac{\alpha}{\varepsilon}$. Thus, for any $\varepsilon > 0$, there must exists $n \in \mathbb{N}$ such that $\mu(A \cap T^{-n}A) \geq \mu(A)^2 - \varepsilon$, and the proof is complete. □

Poincaré's recurrence theorem, and its generalizations and applications, form a subfield of ergodic theory called the *theory of recurrence*. Broadly said, the goal is to understand how and when orbits of points go back to their initial position, if they do. Recurrence properties of dynamical systems can provide important informations about their long-run behaviour. We are then willing to introduce the following terminology.

Definition 1.10. A subset $R \subset \mathbb{N}$ is called a set of recurrence if for every measure-preserving system $(X, \mathcal{A}, \mu, T)$ and for every $A \in \mathcal{A}$ with $\mu(A) > 0$, there is $n \in R$ such that
$$\mu(A \cap T^{-n}A) > 0.$$

With this definition, Poincaré's recurrence theorem exactly says $\mathbb{N}$ is a set of recurrence. In fact, the same proof shows that for an infinite subset $E \subset \mathbb{N}$, the set
$$E - E := \{i - j : i, j \in E\}$$
is a set of recurrence. Finally, the same proof with the sequence $T^{-2}A, T^{-4}A, \ldots$ proves $2\mathbb{N}$ is a set of recurrence.

Note furthermore that any subset of $\mathbb{N}$ containing a set of recurrence is itself a set of recurrence.

On the other hand, the set $2\mathbb{N}+1$ of odd integers is *not* a set of recurrence. Consider for instance the rotation on 2 points system: $X = \{0, 1\}$, $\mathcal{A} = \mathcal{P}(X)$, $\mu(\{0\}) = \mu(\{1\}) = \frac{1}{2}$, and $T : X \longrightarrow X$ is given by $T(0) = 1$ and $T(1) = 0$. Then, for all $k \geq 0$, $T^{2k} = \mathrm{Id}_X$ and $T^{2k+1} = T$. If $A = \{0\}$, then for all $k \geq 0$, one has
$$\mu(A \cap T^{-(2k+1)}A) = \mu(A \cap T^{-1}A) = \mu(\{0\} \cap \{1\}) = \mu(\emptyset) = 0$$
proving that $2\mathbb{N} + 1 = \{2k + 1 : k \geq 0\}$ is not a set of recurrence.

Also, the set $\{2, 3, 5, \ldots\}$ of prime numbers is not a set of recurrence. Indeed, as above, considering the rotation on 4 points system and $A = \{0\}$ shows that $\{2\} \cup (2\mathbb{N}+1)$ is not a set of recurrence. Since the latter contains the set of prime numbers, the claim follows.

1.3 Ergodicity

The word *ergodic* is derived from Ludwig Boltzmann's hypothesis "ergodic hypothesis" in thermodynamics. In the language of measure-preserving systems, this hypothesis means the amount of time that the orbit $Tx, T^2x, T^3x, \ldots$ of a typical point $x \in X$ spends in A a measurable set should be proportional to the measure of that set. For instance, if a set has measure $\frac{1}{2}$, we expect that for half of all $n \in \mathbb{N}$, we have $T^n x \in A$. What we have just described is the conclusion of the so called *Birkhoff's Pointwise Ergodic Theorem*, which will be established in the sequel.

Although Boltzmann's hypothesis should occurs in many dynamical systems we care, it is not true it occurs in every system. We should then distinguish betweem the two. This motivates the next definition.

Definition 1.11. A measure-preserving system $(X, \mathcal{A}, \mu, T)$ is ergodic if

$$A = T^{-1}A \Longrightarrow \mu(A) \in \{0, 1\}.$$

A subset $A \in \mathcal{A}$ is called *invariant* if $A = T^{-1}A$, and *almost invariant* if $\mu(A \Delta T^{-1}A) = 0$. Similarly, a measurable function $f\colon X \longrightarrow \mathbb{C}$ is invariant if $f \circ T = f$, and almost everywhere invariant if $f(Tx) = f(x)$ for μ–almost every $x \in X$.

With this terminology, $(X, \mathcal{A}, \mu, T)$ is ergodic if it has no non-trivial invariant sets. Intuitively, this means the system behaves in a homogeneous manner, filling the entire space. It does not stay concentrated in a certain region of the space.

Example 1.12. (i) The trivial system is clearly ergodic.

(ii) Since any subset of the identity system is invariant, such a system is not ergodic.

(iii) The rotation on 2 points is ergodic, since the only invariant subsets in this case are $\emptyset$ and $X = \{0, 1\}$. More generally, a rotation on m points is ergodic.

Establishing the (non-)ergodicity of more complicated systems is much more involved, and will be done later. However, we can already observe the following.

Remark 1.13. Ergodicity is usually not preserved under direct products. For instance, if $X = \{0, 1\}$, $T(0) = 1$, $T(1) = 0$ is the rotation on 2 points, then the product system $X \times X$ has non-trivial invariant subsets, for instance $A = \{(0, 0), (1, 1)\}$. Clearly, $\mu(A) = \frac{1}{2}$.

Since measurable sets give naturally measurable functions, we are led naturally to the following equivalent characterizations of ergodicity.

Proposition 1.14. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system.
The following claims are equivalent.

(i) The system $(X, \mathcal{A}, \mu, T)$ is ergodic.

(ii) If $A \in \mathcal{A}$ is almost everywhere invariant, then $\mu(A) \in \{0, 1\}$.

(iii) If $f\colon X \longrightarrow \mathbb{C}$ is measurable and invariant, then f equals a constant almost everywhere.

(iv) If $f\colon X \longrightarrow \mathbb{C}$ is measurable and almost everywhere invariant, then f equals a constant almost everywhere.

Proof. (i)$\Longrightarrow$(ii) : Suppose $(X, \mathcal{A}, \mu, T)$ is ergodic, and let $A \in \mathcal{A}$ be almost everywhere invariant. Then the set

$$A' := \bigcap_{k=0}^{\infty} \bigcup_{i=k}^{\infty} T^{-i}A$$

is invariant and we have $\mu(A') = \mu(A)$. By ergodicity, we have $\mu(A') \in \{0, 1\}$, so $\mu(A) \in \{0, 1\}$ as well.

(ii) $\Longrightarrow$ (i) : Obvious, since an invariant set is in particular almost everywhere invariant.

(iii) $\Longrightarrow$ (iv) : Suppose $f: X \longrightarrow \mathbb{C}$ is measurable and almost everywhere invariant. This means the set $A_f := \{x \in X : f(Tx) = f(x)\}$ has full measure, and it implies that

$$A := \bigcap_{k=0}^{\infty} T^{-k} A_f$$

also has full measure, and is invariant. Thus the function $f'(x) = f(x)\mathbf{1}_A$ is measurable and invariant, and by (iii) we deduce that f' is almost everywhere constant. It follows that f is also almost everywhere constant.

(iv) $\Longrightarrow$ (iii) : It is immediate, as (ii) $\Longrightarrow$ (i).

(i) $\Longrightarrow$ (iii) : Fix a measurable and invariant function $f: X \longrightarrow \mathbb{C}$. Recall its *essential supremum* is defined as

$$\operatorname{ess\,sup}(f) := \inf\{\alpha \in \mathbb{R} : \mu(\{x \in X : f(x) > \alpha\}) = 0\}.$$

If $\alpha < \operatorname{ess\,sup}(f)$, then the set $A_\alpha := \{x \in X : f(x) < \alpha\}$ is invariant since f is invariant. By ergodicity, we get $\mu(A_\alpha) = 0$ or $\mu(A_\alpha) = 1$. However, this last case is excluded since $\alpha < \operatorname{ess\,sup}(f)$. Thus $\mu(A_\alpha) = 0$, and it follows that f is almost everywhere constant, equals to $\operatorname{ess\,sup}(f)$.

(iii) $\Longrightarrow$ (i) : Suppose $A \in \mathcal{A}$ satisfies $A = T^{-1}A$. Then $f := \mathbf{1}_A$ is measurable and invariant, so constant almost everywhere. This implies that $\mu(A) \in \{0, 1\}$, and the system is ergodic. □

Example 1.15. Consider $R_\alpha: \mathbb{T} \longrightarrow \mathbb{T}$ a rational rotation of the torus. Choose $n \in \mathbb{Z}$ such that $n\alpha \in \mathbb{Z}$. Then $f(x) := \mathrm{e}^{2\pi i n x}$ is invariant, since

$$f(R_\alpha x) = f(x + \alpha) = \mathrm{e}^{2\pi i n(x+\alpha)} = \mathrm{e}^{2\pi i n\alpha}\mathrm{e}^{2\pi i n x} = f(x).$$

However f is clearly non-constant. This proves this system is not ergodic.

Below are additional characterizations of ergodicity.

Proposition 1.16. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system. The following claims are equivalent.

(i) The system $(X, \mathcal{A}, \mu, T)$ is ergodic.

(ii) If $A \in \mathcal{A}$ satisfies $\mu(A) > 0$, then $\mu\left(\bigcup_{k=1}^{\infty} T^{-k} A\right) = 1$.

(iii) If $A, B \in \mathcal{A}$ satisfies $\mu(A)\mu(B) > 0$, then there exists $n \geq 1$ such that $\mu(T^{-n}A \cap B) > 0$.

Proof. (i) $\Longrightarrow$ (ii) : Let $A \in \mathcal{A}$. Observe that $\bigcup_{k=1}^{\infty} T^{-k}A$ is invariant, so it has either zero or full measure. If it has zero measure, then $\mu(A) = \mu(T^{-1}A) = 0$, which is excluded.

(ii) $\Longrightarrow$ (iii) : Towards a contradiction, suppose $\mu(T^{-n}A \cap B) = 0$ for all $n \geq 1$. Then, since $\mu(A), \mu(B) > 0$, it follows from (ii) that

$$0 < \mu(B) = \mu\left(\left(\bigcup_{n=1}^{\infty} T^{-n}A\right) \cap B\right) \leq \sum_{n=1}^{\infty} \mu(T^{-n}A \cap B) = 0$$

which is absurd. Thus there must exists $n \geq 1$ with $\mu(T^{-n}A \cap B) > 0$.

(iii) $\Longrightarrow$ (i) : We prove the contrapositive. Suppose the system is not ergodic, and let $A \in \mathcal{A}$ having $0 < \mu(A) < 1$. Consider $B := X \setminus A \in \mathcal{A}$, which also has $0 < \mu(B) < 1$. Hence $\mu(A)\mu(B) > 0$, although

$$\mu(T^{-n}A \cap B) = \mu(A \cap (X \setminus A)) = \mu(\emptyset) = 0$$

for all $n \geq 1$. This proves that (iii) does not hold, and concludes the proof. $\square$

2. The Von Neumann's Mean Ergodic Theorem

In this section, we establish a second major result, known as the *Von Neumann's mean ergodic theorem*. For a measure-preserving system $(X, \mathcal{A}, \mu, T)$, it describes for any measurable $f : X \longrightarrow \mathbb{C}$ a convergence in average for the sequence $(f \circ T^n)_{n\in\mathbb{N}}$ to a invariant function.

2.1 The Koopman operator

Definition 2.1. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system.
The linear operator U_T on $L^2(X, \mathcal{A}, \mu)$ defined by

$$\begin{aligned} U_T \colon L^2(X, \mathcal{A}, \mu) &\longrightarrow L^2(X, \mathcal{A}, \mu) \\ f &\longmapsto f \circ T \end{aligned}$$

is called the Koopman operator associated to the transformation T.

First of all, we shall check U_T is well-defined. This follows from T being measure-preserving, since if $f \in L^2(X, \mathcal{A}, \mu)$, one has

$$\|f \circ T\|_2^2 = \int_X |f \circ T|^2 \,\mathrm{d}\mu = \int_X |f|^2 \,\mathrm{d}(T_*\mu) = \int_X |f|^2 \,\mathrm{d}\mu = \|f\|_2^2 < \infty.$$

Moreover, U_T is clearly linear, and in fact by the computation above, we have $\|U_T f\|_2 = \|f\|_2$ for all $f \in L^2(X, \mathcal{A}, \mu)$, *i.e.* U_T is an isometry. In particular, $\|U_T\| = 1$, and U_T is continuous.

In fact, U_T preserves arbitrary scalar products.

Lemma 2.2. For all $f, g \in L^2(X, \mathcal{A}, \mu)$, it holds that $\langle U_T f, U_T g\rangle = \langle f, g\rangle$.

Proof. This is straightforward, since

$$\langle U_T f, U_T g\rangle = \int_X f(Tx)\overline{g(Tx)} \,\mathrm{d}\mu = \int_X f(x)\overline{g(x)} \,\mathrm{d}(T_*\mu) = \int_X f(x)\overline{g(x)} \,\mathrm{d}\mu = \langle f, g\rangle$$

by the change-of-variables formula. □

The next lemma is a fair reformulation of Proposition 1.14.

Lemma 2.3. A measure-preserving system $(X, \mathcal{A}, \mu, T)$ is ergodic if and only if 1 is a simple eigenvalue of U_T.

Proof. The eigenfunctions of U_T for the eigenvalue 1 are exactly the invariant functions, and they form a one dimensional eigenspace if and only if they are all constant almost everywhere. The lemma then follows from the equivalence (i) $\Longleftrightarrow$ (iv) in Proposition 1.14, and the fact that $L^2(X, \mathcal{A}, \mu)$ is dense in $L^0(X, \mathcal{A}, \mu)$. $\square$

This last observation allows to use Fourier analysis coming from the Hilbert structure of L^2 to study ergodicity.

Example 2.4. (i) Consider $R_\alpha \colon \mathbb{T} \longrightarrow \mathbb{T}$ an irrational rotation of the torus. Let $f \in L^2(\mathbb{T})$ be invariant. Writing its Fourier expansion

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

and using that $f(R_\alpha x) = f(x+\alpha) = f(x)$, we get $\sum_{n\in\mathbb{Z}} c_n e^{2\pi i n\alpha} e^{2\pi i n x} = \sum_{n\in\mathbb{Z}} c_n e^{2\pi i n x}$. Uniqueness of the Fourier coefficients, and the fact that $\alpha \notin \mathbb{Q}$, now implies that $c_n = 0$ for all $n \neq 0$, and therefore $f = c_0$ is constant. Thus R_α is ergodic. Combined with Example 1.15, we get that

$$R_\alpha \text{ is ergodic} \Longleftrightarrow \alpha \notin \mathbb{Q}.$$

(ii) The circle doubling map $T \colon \mathbb{T} \longrightarrow \mathbb{T}$, $T(x) = 2x \mod 1$ is ergodic. Again, let $f \in L^2(\mathbb{T})$ be written as

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

with the Parseval identity $\sum_{n\in\mathbb{Z}} |c_n|^2 = \|f\|_2^2 < \infty$. Then the fact that $f(Tx) = f(x)$ and the uniqueness of the Fourier coefficients yields to $c_{2n} = c_n$ for all $n \in \mathbb{Z}$. Thus, if there was $n \neq 0$ with $c_n \neq 0$, it would imply $\sum_{n\in\mathbb{Z}} |c_n|^2 = \infty$, which is impossible. Hence $c_n = 0$ for all $n \neq 0$, and $f = c_0$ is constant almost everywhere.

(iii) Let $X = \mathbb{T}$ equipped with its Borel $\sigma-$algebra and the Lebesgue measure. Consider R_α an irrational rotation, so that (X, R_α) is ergodic. The function f on $X \times X$ defined by $f(x, y) := x - y$ is invariant under $R_\alpha \times R_\alpha$, as proved by

$$f(R_\alpha x, R_\alpha y) = f(x+\alpha, y+\alpha) = (x+\alpha) - (y+\alpha) = x - y = f(x, y)$$

whereas f is not almost everywhere constant. This proves $(X \times X, R_\alpha \times R_\alpha)$ is not ergodic, providing another example that ergodicity is not preserved under direct products.

These examples can be widely generalized to compact abelian groups. Recall that a *character* of a locally compact group is a continuous homomorphism $\chi \colon X \longrightarrow \mathbb{S}^1$. Note that if $T \colon X \longrightarrow X$ is continuous and χ is a character, then $\chi \circ T \colon X \longrightarrow \mathbb{S}^1$ is also a character. We say that χ is said to be *trivial* if $\chi(x) = 1$ for all $x \in X$.

Proposition 2.5. Let X be a compact abelian group, and $T: X \longrightarrow X$ be a continuous surjective homomorphism. Then T is ergodic with respect to the Haar measure m_X if and only if the identity $\chi(T^n x) = \chi(x)$ for some character χ of X and some $n > 0$ implies that χ is trivial.

Proof. Suppose χ is a non-trivial character of X, which satisfies $\chi(T^n x) = \chi(x)$ for some $n > 0$ and all $x \in X$. Consider the function $f: X \longrightarrow \mathbb{C}$ defined by

$$f(x) := \chi(x) + \chi(Tx) + \cdots + \chi(T^{n-1}x).$$

Then we have that $f \circ T = f$, whereas f is not constant, as it is the sum of non-trivial distinct characters. Hence T is not ergodic.

Conversely, fix $f \in L^2(X, \mathcal{B}(X), m_X)$ an invariant function. It has a Fourier expansion

$$f = \sum_{\chi} c_\chi \chi$$

and additionally $\|f\|_2^2 = \sum_{\chi} |c_\chi|^2 < \infty$. Since f is invariant, we have $c_\chi = c_{\chi\circ T} = c_{\chi\circ T^2} = \ldots$, so for a fixed χ, either $c_\chi = 0$ or there are only finitely many distinct characters among $\chi, \chi \circ T, \chi \circ T^2, \ldots$. It follows there exists $p > q$ such that $\chi \circ T^p = \chi \circ T^q$, so $\chi \circ T^{p-q} = \chi$ (the map $\chi \longmapsto \chi \circ T$ is injective since T is surjective). By hypothesis, we conclude that χ is trivial, and thus f is constant. This proves T is ergodic. $\square$

The exact same proof allows one to prove that if G is a compact abelian group and $g \in G$, then the group rotation R_g defined as $R_g(h) = gh$ is ergodic with respect to the Haar measure if and only if the subgroup $\langle g \rangle$ is dense in G.

2.2 The Mean Ergodic Theorem

Let us take a closer look at invariant functions. For $(X, \mathcal{A}, \mu, T)$ a measure-preserving system, we denote

$$\mathcal{H}_{\text{inv}} := \{f \in L^2(X, \mathcal{A}, \mu) : U_T f = f\}$$

and its orthogonal $\mathcal{H}_{\text{erg}} := \mathcal{H}_{\text{inv}}^\perp$. Note that $\mathcal{H}_{\text{inv}} = (U_T - \mathrm{Id}_{\mathcal{H}})^{-1}(\{0\})$ is a closed subspace of $L^2(X, \mathcal{A}, \mu)$. In particular, it follows that

$$L^2(X, \mathcal{A}, \mu) = \mathcal{H}_{\text{inv}} \oplus \mathcal{H}_{\text{erg}}$$

and every $f \in L^2(X, \mathcal{A}, \mu)$ can be written uniquely as $f = f_{\text{inv}} + f_{\text{erg}}$ where f is an invariant function and f_{erg} is *orthogonal* to all invariant functions. Note that f_{inv} is exactly the orthogonal projection of f onto the subspace $\mathcal{H}_{\text{inv}}$ or, said differently, f_{inv} is the element of $\mathcal{H}_{\text{inv}}$ for which $\|f - f_{\text{inv}}\|_2$ is minimal.

It turns out the subspace $\mathcal{H}_{\text{erg}}$ can be described more precisely. This requires the following terminology.

Definition 2.6. A function $f \in L^2(X, \mathcal{A}, \mu)$ is a coboundary if there exists $g \in L^2(X, \mathcal{A}, \mu)$ such that $f = g - g \circ T$.

In the sequel, we will denote $C := \{f \in L^2(X, \mathcal{A}, \mu) : f \text{ is a coboundary}\}$. It is clearly a subspace of $L^2(X, \mathcal{A}, \mu)$, but it is not closed.

Theorem 2.7. We have $\overline{C} = \mathcal{H}_{\text{erg}}$.

Proof. To begin, observe it is enough to prove that $C^\perp = \mathcal{H}_{\text{inv}}$, and this implies $\overline{C}^\perp = \mathcal{H}_{\text{inv}}$, which in turn implies the conclusion by taking orthogonal on both sides. Fix then $f \in C^\perp$. Then, since f is orthogonal to every coboundary, we have

$$\langle f, f - U_T f \rangle = 0$$

so that $\langle f, U_T f \rangle = \langle f, U_T f \rangle + \langle f, f - U_T f \rangle = \langle f, f \rangle = \|f\|_2^2$. It then follows that

$$\begin{aligned}
\|f - U_T f\|^2 &= \|f\|_2^2 + \|U_T f\|^2 - 2\text{Re}\langle f, U_T f \rangle \\
&= 2\|f\|_2^2 - 2\text{Re}\|f\|_2^2 \\
&= 0
\end{aligned}$$

using that U_T has norm 1. Thus $f = U_T f$, and $f \in \mathcal{H}_{\text{inv}}$. For the converse, fix $f \in \mathcal{H}_{\text{inv}}$ and $g \in C$. We must show that $\langle f, g \rangle = 0$. Since g is a coboundary, write $g = h - U_T h$. Since U_T preserves the inner product, and f is invariant, we directly obtain

$$\langle f, g \rangle = \langle f, h - U_T h \rangle = \langle f, h \rangle - \langle U_T f, U_T h \rangle = \langle U_T f, U_T h \rangle - \langle U_T f, U_T h \rangle = 0$$

and thus $f \in C^\perp$. This finishes the proof. □

This allows us to already establish the Von Neumann's Mean Ergodic Theorem. Let us first state and prove the general case.

Theorem 2.8. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system.
For every $f \in L^2(X, \mathcal{A}, \mu)$, we have

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} U_T^n f = f_{\text{inv}}$$

with respect to $\| \cdot \|_2$.

Proof. Let $f \in L^2(X, \mathcal{A}, \mu)$, and write $f = f_{\text{inv}} + f_{\text{erg}}$. Then

$$\frac{1}{N}\sum_{n=0}^{N-1} U_T^n f = \frac{1}{N}\sum_{n=0}^{N-1} U_T^n f_{\text{inv}} + \frac{1}{N}\sum_{n=0}^{N-1} U_T^n f_{\text{erg}} = f_{\text{inv}} + \frac{1}{N}\sum_{n=0}^{N-1} U_T^n f_{\text{erg}}$$

so we will be done if we can show that $\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} U_T^n f_{\text{erg}} = 0$. By Theorem 2.7, we can assume that f_{erg} is a coboundary, and write $f_{\text{erg}} = h - U_T h$. With this, the last sum is telescopic, yielding

$$\frac{1}{N}\sum_{n=0}^{N-1} U_T^n f_{\text{erg}} = \frac{h - U_T^N h}{N}.$$

Now, we get $\left\|\frac{h-U_T^N h}{N}\right\|_2 \leq \frac{2\|h\|_2}{N}$ and this last quantity goes to 0 as $N \to \infty$. This concludes our proof. □

The quantity studied in this theorem is called an *ergodic average*, and will be now a central object for us. To shorten notations sometimes, we will denote

$$\mathrm{A}_N(f) := \frac{1}{N}\sum_{n=0}^{N-1} f \circ T^n$$

the $N-$th ergodic average of a measurable function $f : X \longrightarrow \mathbb{C}$.

2.3 Consequences of the Mean Ergodic Theorem

First of all, the Mean Ergodic Theorem takes a meaningful form if the system is ergodic.

Theorem 2.9. Let $(X, \mathcal{A}), \mu, T)$ be an ergodic measure-preserving system. For every $f \in L^2(X, \mathcal{A}, \mu)$, we have

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} U_T^n f = \int_X f \, \mathrm{d}\mu$$

with respect to $\|\cdot\|_2$.

Proof. By Theorem 2.8, we have

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} U_T^n f = f_{\text{inv}}$$

in L^2–norm. By ergodicity, f_{inv} must equals a constant almost everywhere, which we denote by c. By writing $f = f_{\text{inv}} + f_{\text{erg}}$ we get

$$\int_X f \,\mathrm{d}\mu = \int_X f_{\text{inv}} \,\mathrm{d}\mu + \int_X f_{\text{erg}} \,\mathrm{d}\mu = c + \int_X f_{\text{erg}} \,\mathrm{d}\mu$$

and we observe that $\int_X f_{\text{erg}} \,\mathrm{d}\mu = \langle f_{\text{erg}}, \mathbf{1}_X \rangle = 0$ since $\mathbf{1}_X$ is obviously invariant and f_{erg} is orthogonal to any invariant function. This shows the claim. □

Since an ergodic system has a certain homogeneity, the orbit of a subset $A \in \mathcal{A}$ should, in average, behaves indepently of any other subset $B \in \mathcal{A}$. This is confirmed by the next corollary.

Corollary 2.10. A measure-preserving system $(X, \mathcal{A}, \mu, T)$ is ergodic if and only if

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

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

Proof. Suppose first the system is not ergodic, and let $A \in \mathcal{A}$ be an invariant subset such that $0 < \mu(A) < 1$. Consider $B := X \setminus A$, which also has $0 < \mu(B) < 1$. Since $T^{-n}A = A$ for all $n \geq 0$, we directly get that

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

while $\mu(A)\mu(B) > 0$.

Conversely, suppose $(X, \mathcal{A}, \mu, T)$ is ergodic. First note that $\mathbf{1}_{T^{-n}A} = \mathbf{1}_A \circ T^n = U_T^n \mathbf{1}_A$, so we can write

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(T^{-n}A \cap B) &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X \mathbf{1}_{T^{-n}A} \mathbf{1}_B \,\mathrm{d}\mu \\
&= \lim_{N\to\infty} \int_X \left(\frac{1}{N} \sum_{n=0}^{N-1} U_T^n \mathbf{1}_A \right) \mathbf{1}_B \,\mathrm{d}\mu.
\end{aligned}$$

Using ergodicity, we may apply Theorem 2.9 to obtain that $\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} U_T^n \mathbf{1}_A = \mu(A)$ in L^2–norm. Since norm convergence implies weak convergence, it follows that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(T^{-n}A \cap B) = \int_X \left(\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} U_T^n \mathbf{1}_A \right) \mathbf{1}_B \,\mathrm{d}\mu$$

$$= \int_X \mu(A)\mathbf{1}_B \, \mathrm{d}\mu$$
$$= \mu(A)\mu(B)$$

and we are done. □

Since finite linear combinations of indicator functions are dense in $L^2(X, \mathcal{A}, \mu)$, we in fact proved that a measure-preserving system $(X, \mathcal{A}, \mu, T)$ is ergodic if and only if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \langle U_T^n f, g\rangle = \langle f, \mathbf{1}_X\rangle\langle \mathbf{1}_X, g\rangle$$

for all $f, g \in L^2(X, \mathcal{A}, \mu)$.

The next result is a stronger version of the Mean Ergodic Theorem, sometimes called the *Uniform Mean Ergodic Theorem*.

Theorem 2.11. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system. For every $f \in L^2(X, \mathcal{A}, \mu)$, we have

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} U_T^n f = f_{\mathrm{inv}}$$

with respect to $\|\cdot\|_2$.

Proof. We proceed exactly as in the proof of Theorem 2.8. Writing $f = f_{\mathrm{inv}} + f_{\mathrm{erg}}$, it is enough to prove that

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} U_T^n f_{\mathrm{erg}} = 0$$

in L^2–norm. Fix then $\varepsilon > 0$, and choose $h \in C$ a coboundary so that $\|f_{\mathrm{erg}} - h\|_2 < \frac{\varepsilon}{2}$. Write $h = g - U_T g$, and exactly as before we note that

$$\left\| \frac{1}{N-M} \sum_{n=M}^{N-1} U_T^n h \right\|_2 = \left\| \frac{1}{N-M} \sum_{n=M}^{N-1} U_T^n h - U_T^{n+1} h \right\|_2 = \frac{\|U_T^M h - U_T^N h}{N-M} \le \frac{2\|h\|_2}{N-M}$$

and this last quantity tends to 0 as $N - M \to \infty$. Therefore we have

$$\left\| \frac{1}{N-M} \sum_{n=M}^{N-1} U_T^n h \right\|_2 < \frac{\varepsilon}{2}$$

if $N - M$ is large enough. It thus follows that

$$\left\| \frac{1}{N-M} \sum_{n=M}^{N-1} U_T^n f_{\mathrm{erg}} \right\|_2 = \left\| \frac{1}{N-M} \sum_{n=M}^{N-1} (U_T^n f_{\mathrm{erg}} - U_T^n h) + \frac{1}{N-M} \sum_{n=M}^{N-1} U_T^n h \right\|_2$$

$$\begin{aligned}
&\leq \|f_{\mathrm{erg}} - h\|_2 + \left\| \frac{1}{N-M} \sum_{n=M}^{N-1} U_T^n h \right\|_2 \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\
&= \varepsilon
\end{aligned}$$

if $N - M$ is large enough. This concludes the proof. □

Now we use this new version to establish *Khintchine's recurrence theorem*. To that end, recall that a subset $S = \{s_0 < s_1 < \dots\}$ of $\mathbb{N}$ is *syndetic* if it has bounded gaps, in the sense that

$$\sup_{i \in \mathbb{N}} (s_{i+1} - s_i) < \infty.$$

In particular, if A contains an infinite syndetic set, then it is itself syndetic.

Theorem 2.12. Let $(X, \mathcal{A}, \mu, T)$ be measure-preserving, and let $A \in \mathcal{A}$. For all $\varepsilon > 0$, the set

$$\{n \in \mathbb{N} : \mu(A \cap T^{-n}A) > \mu(A)^2 - \varepsilon\}$$

is syndetic.

Proof. First, we apply Theorem 2.11 to $\mathbf{1}_A$ to obtain that

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} \mathbf{1}_{T^{-n}A} = (\mathbf{1}_A)_{\mathrm{inv}}$$

in $\|\cdot\|_2$. Thus the convergence also holds weakly, and in particular we get

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} \mu(T^{-n}A \cap A) = \left\langle \lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} \mathbf{1}_{T^{-n}A}, \mathbf{1}_A \right\rangle = \langle (\mathbf{1}_A)_{\mathrm{inv}}, \mathbf{1}_A \rangle.$$

Writing $\mathbf{1}_A = (\mathbf{1}_A)_{\mathrm{inv}} + (\mathbf{1}_A)_{\mathrm{erg}}$, we also have

$$\langle (\mathbf{1}_A)_{\mathrm{inv}}, \mathbf{1}_A \rangle = \|(\mathbf{1}_A)_{\mathrm{inv}}\|_2^2 = \int_X (\mathbf{1}_A)_{\mathrm{inv}}^2 \,\mathrm{d}\mu$$

and by Cauchy-Schwarz it follows that

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} \mu(T^{-n}A \cap A) = \int_X (\mathbf{1}_A)_{\mathrm{inv}}^2 \,\mathrm{d}\mu \geq \left(\int_X (\mathbf{1}_A)_{\mathrm{inv}} \,\mathrm{d}\mu \right)^2 = \mu(A)^2.$$

Fix $\varepsilon > 0$, and towards a contradiction suppose the given set is not syndetic. This means there exists arbitrary large intervals of integers $[M, N)$ such that, for every n in this interval, we have $\mu(T^{-n}A \cap A) \leq \mu(A)^2 - \varepsilon$. This implies that

$$\frac{1}{N-M} \sum_{n=M}^{N-1} \mu(T^{-n}A \cap A) \leq \mu(A)^2 - \varepsilon < \mu(A)^2$$

and this contradicts the above inequality. Thus $\{n \in \mathbb{N} : \mu(A \cap T^{-n}A) > \mu(A)^2 - \varepsilon\}$ is syndetic, as claimed. $\square$

3. The Birkhoff's Pointwise Ergodic Theorem

In the previous section, we established a convergence in L^2 for ergodic averages of functions of a measure-preserving system. Another important type of convergence to consider in measure and probability theory is the convergence *almost everywhere*, or *almost surely*. We show now an analog of Von Neumann's theorem about almost sure convergence of ergodic averages associated to a measure-preserving system.

3.1 The Maximal Inequality

The proof of the Pointwise Ergodic Theorem hinges on a technical result called the *maximal inequality*.

For $f \in L^1(X, \mathcal{A}, \mu)$, let $S_0 := 0$ and $S_m(x) := \sum_{n=0}^{m-1} f(T^n x)$ for $m \geq 1$ and $x \in X$. Also, let $F_N(x) := \max_{0 \leq m \leq N} S_m(x)$, and $P := \{x \in X : F_N(x) > 0\}$.

Proposition 3.1. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system. With the above notations, it holds that

$$\int_P f \, \mathrm{d}\mu \geq 0.$$

Proof. Since $F_N(x) \geq S_m(x)$ for all $m = 0, \ldots, N$, we have

$$F_N(Tx) + f(x) \geq S_m(Tx) + f(x) = S_{m+1}(x)$$

and thus $F_N(Tx) + f(x) \geq \max_{1 \leq m \leq N} S_m(x)$ for all $x \in X$. For $x \in P$, we have

$$\max_{1 \leq m \leq N} S_m(x) = \max_{0 \leq m \leq N} S_m(x)$$

and hence $F_N(Tx) + f(x) \geq \max_{0 \leq m \leq N} S_m(x) = F_N(x)$ for all $x \in P$. It thus follows that

$$\begin{aligned}
\int_P f(x) \, \mathrm{d}\mu &\geq \int_P F_N(x) - F_N(Tx) \, \mathrm{d}\mu \\
&= \int_P F_N(x) \, \mathrm{d}\mu - \int_P F_N(Tx) \, \mathrm{d}\mu \\
&\geq \int_X F_N(x) \, \mathrm{d}\mu - \int_X F_N(Tx) \, \mathrm{d}\mu \\
&= \int_X F_N(x) \, \mathrm{d}\mu - \int_X F_N(x) \, \mathrm{d}(T_* \mu)
\end{aligned}$$

using the change of variables formula. As T is measure-preserving, the latter difference is 0. This concludes the proof. $\square$

In measure theory, Markov's inequality tells that whenever $(X, \mathcal{A}, \mu)$ is a measure space and $f : X \longrightarrow \mathbb{R}$ is measurable, it holds that

$$\mu(\{x \in X : |f(x)| \geq \varepsilon\}) \leq \frac{1}{\varepsilon} \int_X |f| \, \mathrm{d}\mu$$

for all $\varepsilon > 0$. Applying it with the ergodic average $\mathrm{A}_N(f)$ of f, using the triangle equality and the fact that T is measure-preserving, we get

$$\mu\left(\left\{x \in X : \left|\frac{1}{N} \sum_{n=0}^{N-1} f(T^n x)\right| \geq \varepsilon\right\}\right) \leq \frac{1}{\varepsilon} \int_X |f| \, \mathrm{d}\mu$$

for all $N \geq 1$ and $\varepsilon > 0$. The next result can therefore be seen as a uniform version of this inequality over $N \geq 1$. It is sometimes referred as the *Maximal Ergodic Theorem*.

Theorem 3.2. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system. For all $f \in L^1(X, \mathcal{A}, \mu)$ and $\varepsilon > 0$, we have

$$\mu\left(\left\{x \in X : \sup_{N \geq 1} \left|\frac{1}{N} \sum_{n=0}^{N-1} f(T^n x)\right| \geq \varepsilon\right\}\right) \leq \frac{1}{\varepsilon} \int_X |f| \, \mathrm{d}\mu.$$

Proof. By writing $f = f^+ - f^-$ into its positive and negative parts, we can assume without restriction that f is positive. By applying Proposition 3.1 to the function $f - \varepsilon$, we obtain

$$\int_{P_M} f(x) - \varepsilon \, \mathrm{d}\mu \geq 0$$

for all $M \geq 1$, where $P_M := \left\{x \in X \mid \sup_{1 \leq N \leq M} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x) > \varepsilon\right\}$. Observe that, if $P := \left\{x \in X \mid \sup_{N \geq 1} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x) > \varepsilon\right\}$, then $P = \bigcup_{M \geq 1} P_M$ and it follows that

$$\int_P f(x) - \varepsilon \, \mathrm{d}\mu \geq 0.$$

Hence $\mu(P) \leq \frac{1}{\varepsilon} \int_P f \, \mathrm{d}\mu \leq \frac{1}{\varepsilon} \int_P |f| \, \mathrm{d}\mu \leq \frac{1}{\varepsilon} \int_X |f| \, \mathrm{d}\mu$, and we are done. $\square$

3.2 The Pointwise Ergodic Theorem

We are now ready to formulate the Pointwise Ergodic Theorem, first in the general case of an arbitrary measure-preserving system.

Theorem 3.3. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system. For every $f \in L^2(X, \mathcal{A}, \mu)$, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x) = f_{\text{inv}}(x)$$

for $\mu-$almost every $x \in X$.

Proof. First, notice that if we prove the existence of a limit F of $(\mathrm{A}_N(f))_{N\geq 1}$ almost everywhere, then F must be f_{inv}. Indeed, since $(\mathrm{A}_N(f))_{N\geq 1}$ converges to f_{inv} in L^2-norm, it has a subsequence that converges almost everywhere to f_{inv}, but this subsequence also has to converge to F almost everywhere, yielding $F = f_{\text{inv}}$. We are then left to show the existence of the limit almost everywhere.

Let $\mathcal{L}$ be the subset of all real-valued functions $f \in L^2(X, \mathcal{A}, \mu)$ for which the limit

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x)$$

exists for $\mu-$almost every $x \in X$. Our goal is to prove that $\mathcal{L} = L^2(X, \mathcal{A}, \mu)$.
To begin, observe first that $\mathcal{L}$ is a subspace of $L^2(X, \mathcal{A}, \mu)$, and that $\mathcal{H}_{\text{inv}} \subset \mathcal{L}$. Thus, it is enough to show that $\mathcal{H}_{\text{erg}} \subset \mathcal{L}$, since we will then have

$$L^2(X, \mathcal{A}, \mu) = \mathcal{H}_{\text{inv}} \oplus \mathcal{H}_{\text{erg}} \subset \mathcal{L}$$

which implies the conclusion. Fix then $f \in \mathcal{H}_{\text{erg}}$, and $\varepsilon > 0$. By Theorem 2.7, we can pick h a coboundary with $\|f - h\|_2 \leq \varepsilon^2$. In particular, $\int_X |f - h| \,\mathrm{d}\mu \leq \varepsilon^2$, and applying Theorem 3.2 with the function $f - h$, it follows that

$$\mu\left(\left\{x \in X : \sup_{N\geq 1} \left|\frac{1}{N} \sum_{n=0}^{N-1} f(T^n x) - h(T^n x)\right| \geq \varepsilon\right\}\right) \leq \frac{1}{\varepsilon} \int_X |f - h| \,\mathrm{d}\mu \leq \varepsilon.$$

Since the lim sup of a sequence of real numbers is bounded by the supremum of this sequence, and since μ is increasing, we obtain also

$$\mu\left(\left\{x \in X : \limsup_{N\to\infty} \left|\frac{1}{N} \sum_{n=0}^{N-1} f(T^n x) - h(T^n x)\right| \geq \varepsilon\right\}\right) \leq \varepsilon.$$

Now, $h = g - g \circ T$ is a coboundary, so its ergodic average is telescopic, and gives $\lim_{N\to\infty} \mathrm{A}_N(h) = 0$ for μ–almost every $x \in X$. Thus

$$\mu\left(\left\{x \in X : \limsup_{N\to\infty} \left|\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x)\right| \geq \varepsilon\right\}\right) \leq \varepsilon.$$

As $\varepsilon > 0$ was arbitrary, this last inequality exactly says that the limit

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} f(T^n x)$$

exists and equals 0 for μ–almost every $x \in X$. Hence $f \in \mathcal{L}$, and this finishes the proof. $\square$

In the ergodic case, we know precisely the orthogonal projection of f onto $\mathcal{H}_{\text{inv}}$. This leads to the next corollary.

Corollary 3.4. Let $(X, \mathcal{A}, \mu, T)$ be an ergodic measure-preserving system. For every $f \in L^2(X, \mathcal{A}, \mu)$, it holds that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) = \int_X f \, \mathrm{d}\mu$$

for μ–almost every $x \in X$.

To end this subsection, we prove a pointwise ergodic theorem for non-integrable functions.

Theorem 3.5. Let $(X, \mathcal{A}, \mu, T)$ be an ergodic measure-preserving system and f a measurable function with $\int_X f \, \mathrm{d}\mu = \infty$. Then we have

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) = \infty$$

for μ–almost every $x \in X$.

Proof. We start by defining $\overline{f} : X \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ by

$$\overline{f}(x) := \liminf_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} f(T^n x)$$

and the measurable subset $A := \{x \in X : \overline{f}(x) < \infty\}$. First, $\overline{f}$ is invariant. Indeed, fix $x \in X$, and note that

$$\frac{N+1}{N}\left(\frac{1}{N+1}\sum_{n=0}^{N} f(T^n x)\right) = \frac{1}{N}\sum_{n=0}^{N-1} f(T^n(Tx)) + \frac{1}{N}f(x).$$

The left hand side has a subsequence converging to $\liminf_{N\to\infty} \mathrm{A}_{N+1}(f) = \overline{f}(x)$, so $\overline{f} \leq \overline{f} \circ T$. Doing the same for the right hand side, we get $\overline{f} \geq \overline{f} \circ T$, and hence $\overline{f}$ is invariant. Now we restrict $\overline{f}$ by setting $g := \overline{f}\mathbf{1}_A$. It is a measurable function since $\overline{f}$ and A is measurable. Moreover, we have

$$g \circ T(x) = \begin{cases} \overline{f}(Tx) & \text{if } \overline{f}(Tx) < \infty \\ 0 & \text{otherwise} \end{cases} = \begin{cases} \overline{f}(x) & \text{if } \overline{f}(x) < \infty \\ 0 & \text{otherwise} \end{cases} = g(x)$$

proving that g is also invariant. By ergodicity, g must be then constant almost everywhere. We denote this constant by c. It implies that either $\mu(A) = 0$, in which case we are done, or $\mu(A) = 1$, and $\overline{f}$ is constant on a full measure subset A' of A. Assume for a contradiction that this second case occurs. Define, for $m \geq 1$, the function

$$f_m(x) := \begin{cases} f(x) & \text{if } f(x) < m \\ 0 & \text{otherwise} \end{cases}$$

for all $x \in X$. The sequence $(f_m)_{m\geq 1}$ is in $L^\infty(X, \mathcal{A}, \mu)$, is increasing and converges pointwise to f. It follows from the monotone convergence theorem that also

$$\lim_{m\to\infty} \int_X f_m \, \mathrm{d}\mu = \int_X f \, \mathrm{d}\mu = \infty.$$

Lastly, we let

$$\overline{f_m} = \liminf_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} f_m \circ T^n$$

and we remark that $\overline{f_m} \leq \overline{f}$ for all $m \geq 1$ since $f_m \leq f$ for all $m \geq 1$. This implies that

$$\int_X \overline{f_m} \, \mathrm{d}\mu \leq \int_X \overline{f} \, \mathrm{d}\mu = \int_A \overline{f} \, \mathrm{d}\mu = \int_{A'} \overline{f} \, \mathrm{d}\mu = c$$

whence $c = \infty$, a contradiction. Hence $\mu(A) = 0$ and we are done. □

3.3 Normal numbers and the Borel's theorem

Recall that for a subset $A \subset \mathbb{N}$, its *density* is the number $d(A)$ defined as

$$d(A) := \lim_{N\to\infty} \frac{|A \cap \{1, \dots, N\}|}{N}.$$

In words, this corresponds to the proportion that A occupies in the integers. For instance, $2\mathbb{N}$ and $2\mathbb{N}+1$ have density $\frac{1}{2}$, and $10\mathbb{N}$ has density $\frac{1}{10}$. The set of prime numbers has density 0.

Let now $p \geq 2$. Any real number $x \in [0,1)$ has a *base p digit expansion*, of the form

$$x = \sum_{i=1}^{\infty} d_i p^{-i}, \ d_1, d_2, \dots \in \{0, \dots, p-1\}.$$

The numbers $d_1, d_2, \dots$ are called the *digits* of x in base p. The question is therefore to understand with which frequency each digit appears.

Definition 3.6. A number $x = \sum_{i=1}^{\infty} d_i p^{-i}$ is called normal in base p if for all $k \geq 1$ and $c_1, \dots, c_k \in \{0, \dots, p-1\}$, the set

$$\{n \in \mathbb{N} : d_{n+1} = c_1, \dots, d_{n+k} = c_k\}$$

has density p^{-k}.

Here is then the Borel's theorem on normal numbers.

Theorem 3.7. For any $p \geq 2$, almost every real number $x \in [0,1)$ is normal in base p.

Proof. Given $x \in [0,1)$, denote its digits in base p as $d_i(x)$, in such a way that

$$x = \sum_{i=1}^{\infty} d_i(x) p^{-i}.$$

Fix $k \geq 1$ and $c_1, \dots, c_k \in \{0, \dots, p-1\}$. Consider the set

$$C := \{x \in [0,1) \mid d_1(x) = c_1, \dots, d_k(x) = c_k\}.$$

One checks easily that the Lebesgue measure of C equals p^{-k}. On the other hand, the map $T\colon [0,1) \longrightarrow [0,1)$, $T(x) = px \mod 1$ is a Lebesgue measure-preserving ergodic transformation of $[0,1)$ (similarly to Example 2.4(ii)). We can thus apply the Pointwise Ergodic Theorem with $f = \mathbf{1}_C$ to get

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{1}_C(T^n x) = \int_0^1 1_C(x)\, \mathrm{d}x = p^{-k}.$$

for almost every $x \in [0,1)$. To conclude it suffices to notice that $\mathbf{1}_C(T^n x) = 1$ if and only if $d_{n+1}(x) = c_1, \dots, d_{n+k}(x) = c_k$, so that the limit on the left-hand side is precisely the natural density of $\{n \in \mathbb{N} : d_{n+1} = c_1, \dots, d_{n+k} = c_k\}$. $\square$

3.4 Continued fractions expansion

A *continued fraction* is an expression of the form

$$[a_0; a_1, a_2, \dots] = a_0 + \cfrac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots}}}$$

where $a_0 \in \mathbb{Z}$ and $a_1, a_2, \dots \in \mathbb{N}$. As we will see in this part, any continued fraction corresponds to a unique irrational number, and any irrational number has a unique continued fraction expansion.

A *finite truncation* of a continued fraction $[a_0; a_1, a_2, \dots]$ is the fraction

$$\frac{p_n}{q_n} = a_0 + \cfrac{1}{a_1 + \frac{1}{a_2 + \dots \frac{1}{a_{n-1} + \frac{1}{a_n}}}}$$

and is called the n–th convergent to $[a_0; a_1, a_2, \dots]$.

We first establish general properties of the convergents.

Proposition 3.8. Let $[a_0; a_1, a_2, \dots]$ be a continued fraction and denote $\alpha_n = \frac{p_n}{q_n}$ its n–th convergent. The following holds.

(i) For all $n \geq 1$, we have

$$p_{n+1} = a_{n+1} p_n + p_{n-1},$$
$$q_{n+1} = a_{n+1} q_n + q_{n-1},$$
$$p_{n+1} q_n - p_n q_{n+1} = (-1)^n.$$

(ii) The number $\alpha := \lim_{n \to \infty} \alpha_n$ exists, is irrational, and satisfies

$$|\alpha - \alpha_n| < \frac{1}{q_n q_{n+1}}.$$

(iii) One has $\alpha_0 < \alpha_2 < \alpha_4 < \cdots < \alpha < \cdots < \alpha_5 < \alpha_3 < \alpha_1$.

Proof. (i) We proceed by induction on n. Note first that $p_0 = a_0$, $q_0 = 1$, $p_1 = a_0 a_1 + 1$ and $q_1 = a_1$. So if we set $p_{-1} = 1$ and $q_{-1} = 0$ then the formulas hold for $n = 0$. Now suppose they have been verified for some $n \geq 0$. Define $\tilde{p} := a_{n+2} p_{n+1} + p_n$ and $\tilde{q} := a_{n+2} q_{n+1} + q_n$. Our goal is to prove that $(\tilde{p}, \tilde{q}) = (p_{n+2}, q_{n+2})$ and $\tilde{p} q_{n+1} - p_{n+1} \tilde{q} = (-1)^{n+1}$. First, one computes that

$$\tilde{p} q_{n+1} - p_{n+1} \tilde{q} = (a_{n+2} p_{n+1} + p_n) q_{n+1} - p_{n+1}(a_{n+2} q_{n+1} + q_n)$$

$$\begin{aligned}
&= p_n q_{n+1} - p_{n+1} q_n \\
&= -(-1)^n \\
&= (-1)^{n+1}
\end{aligned}$$

using the induction hypothesis. This proves already the third identity, and that $\tilde{p}$ and $\tilde{q}$ are coprime. Furthermore, observe that

$$\begin{aligned}
\frac{\tilde{p}}{\tilde{q}} &= \frac{a_{n+2}p_{n+1} + p_n}{a_{n+2}q_{n+1} + q_n} \\
&= \frac{a_{n+2}(a_{n+1}p_n + p_{n-1}) + p_n}{a_{n+2}(a_{n+1}q_n + q_{n-1}) + q_n} \\
&= \frac{(a_{n+1} + \frac{1}{a_{n+2}})p_n + p_{n-1}}{(a_{n+1} + \frac{1}{a_{n+2}})q_n + q_{n-1}} \\
&= \left[a_0; a_1, \ldots, a_{n+1} + \frac{1}{a_{n+2}}\right] \\
&= [a_0; a_1, \ldots, a_{n+1}, a_{n+2}] \\
&= \frac{p_{n+2}}{q_{n+2}}
\end{aligned}$$

and since $\tilde{p}$ and $\tilde{q}$ are coprime, it follows that $\tilde{p} = p_{n+2}$ and $\tilde{q} = q_{n+2}$, as wanted. This concludes the induction, and (i) is shown.

(ii) Note that $p_{n+1}q_n - p_n q_{n+1} = (-1)^n$ implies

$$\frac{p_{n+1}}{q_{n+1}} = \frac{p_n}{q_n} + \frac{(-1)^n}{q_n q_{n+1}}$$

and iterating this relation provides $\dfrac{p_n}{q_n} = a_0 + \displaystyle\sum_{j=0}^{n-1} \frac{(-1)^j}{q_j q_{j+1}}$, for all $n \geq 1$. Now from (i) we see that

$$1 = q_0 \leq q_1 < q_2 < q_3 < \ldots$$

since $a_n \geq 1$ for all $n \geq 1$. By induction, $q_n \geq 2^{\frac{n-2}{2}}$ for all $n \geq 1$. Thus the series

$$\sum_{j=0}^{\infty} \frac{(-1)^j}{q_j q_{j+1}}$$

is absolutely convergent, and $\alpha := \lim\limits_{n\to\infty} \dfrac{p_n}{q_n}$ is well-defined. It also satisfies

$$\left|\alpha - \frac{p_n}{q_n}\right| = \left|\sum_{j=n}^{\infty} \frac{(-1)^j}{q_j q_{j+1}}\right| < \frac{1}{q_n q_{n+1}}$$

as desired. We now prove α is irrational, by contradiction. Suppose $\alpha = \frac{a}{b}$ for some $a \in \mathbb{Z}$, $b \in \mathbb{N}$. The last inequality multiplied by $q_n b$ then implies $|q_n a - b p_n| < \frac{b}{q_{n+1}}$, which tends to 0 as $n \longrightarrow \infty$. Since $q_n a - bp_n$ is an integer, we must have $q_n a = bp_n$ for all n large enough, and hence $\frac{p_n}{q_n} = \frac{a}{b}$ for all n large enough. This contradicts the fact that $q_n \longrightarrow \infty$ as $n \longrightarrow \infty$.

(iii) follows from

$$\frac{p_n}{q_n} = a_0 + \sum_{j=0}^{n-1} \frac{(-1)^j}{q_j q_{j+1}}$$

and the fact that in this sum the terms are decreasing and have alternating signs. □

Now we introduce a dynamical approach to the theory of continued fractions.

Definition 3.9. The Gauss map is the map $T\colon [0,1) \longrightarrow [0,1)$ defined as

$$T(x) = \begin{cases} \frac{1}{x} \bmod 1 & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

and the Gauss measure μ is given by

$$\mu(B) = \frac{1}{\log 2} \int_B \frac{1}{1+x} \, \mathrm{d}x$$

for every Borel set $B \subset [0,1)$.

An observation that will be useful in the sequel is that μ and the Lebesgue measure on $[0,1)$ are absolutely continuous with respect to each other, *i.e.* they share the same null sets.

The first thing to show is that it is indeed a measure-preserving system.

Lemma 3.10. The map T preserves the Gauss measure μ.

Proof. It is enough to prove that $\mu(T^{-1}([0,s])) = \mu([0,s])$ since intervals of this form generate the Borel σ−algebra on $[0,1)$. Note that

$$T^{-1}([0,s]) = \{0\} \cup \left\{ x \in (0,1) : \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor \leq s \right\} = \bigcup_{k \in \mathbb{N}} \left[\frac{1}{k+s}, \frac{1}{k} \right]$$

is a disjoint union. Hence we have

$$\mu(T^{-1}([0,s])) = \frac{1}{\log 2} \sum_{k \in \mathbb{N}} \int_{\frac{1}{k+s}}^{\frac{1}{k}} \frac{1}{1+x} \, \mathrm{d}x$$

$$
\begin{aligned}
&= \frac{1}{\log 2}\sum_{k\in\mathbb{N}}\left(\log\left(1+\frac{1}{k}\right)-\log\left(1+\frac{1}{k+s}\right)\right)\\
&= \frac{1}{\log 2}\sum_{k\in\mathbb{N}}\left(\log\left(1+\frac{s}{k}\right)-\log\left(1+\frac{s}{k+1}\right)\right)\\
&= \frac{1}{\log 2}\sum_{k\in\mathbb{N}}\int_{\frac{s}{k+1}}^{\frac{s}{k}}\frac{1}{1+x}\,\mathrm{d}x\\
&= \mu([0,s])
\end{aligned}
$$

since also $[0,s]=\bigcup_{k\in\mathbb{N}}\left[\frac{s}{k+1},\frac{s}{k}\right]$ is a disjoint union. This proves the claim. □

For our dynamical purposes, we will take for granted the following fact.

Proposition 3.11. The map T is ergodic with respect to the Gauss measure.

With our previous notations, the continued fraction expansion

$$
x=\cfrac{1}{a_1+\cfrac{1}{a_2+\cfrac{1}{\ddots}}}
$$

of an irrational number $x\in[0,1)$ is denoted $[0;a_1,a_2,\dots]$. Observe that in this case $T(x)=\frac{1}{x}-a_1=[0;a_2,a_3,\dots]$, so T acts as the left shift on the continued fraction representation of a number.

Next, observe that if $x=[0;a_1,a_2,\dots]\in[0,1)$, then $a_1=k$ if and only if $x\in(\frac{1}{k+1},\frac{1}{k}]$. In other words, the continued fraction expansion of all numbers in $(\frac{1}{2},1]$ starts with $a_1=1$, those of numbers in $(\frac{1}{3},\frac{1}{2}]$ starts with $a_1=2$, and so on. We let then

$$
a(x)=\sum_{k\in\mathbb{N}}k\mathbf{1}_{(\frac{1}{k+1},\frac{1}{k}]}(x)
$$

the function that maps $x\in[0,1)$ to the first digit in its continued fraction expansion. Since the Gauss map T acts as the left shift on the continued fraction representation of a real number x, we can recover all digits of the expansion of x through its orbit $x,Tx,T^2x,\dots$. More precisely, for any irrational $x=[0;a_1,a_2,\dots]$, we have

$$
a_{n+1}=a(T^nx)
$$

for all $n\in\mathbb{N}\cup\{0\}$. This establishes a direct link between the dynamical behaviour of the Gauss map T and the continued fraction representation of irrationals in the $[0,1)$ interval. This link allows one to invoque Birkhoff's Pointwise Ergodic Theorem to get qualitative results about the continued fractions expansion of real numbers.

Theorem 3.12. The following hold for Lebesgue almost every $x = [0; a_1, a_2, \dots] \in (0,1)$.

(i) The digit k appears in the expansion $[a_0; a_1, a_2, \dots]$ of x with frequency

$$\frac{2\log(k+1) - \log(k) - \log(k+2)}{\log 2}.$$

(ii) $\displaystyle\lim_{N\to\infty} \frac{a_1 + \cdots + a_N}{N} = \infty.$

(iii) $\displaystyle\lim_{N\to\infty} (a_1 \dots a_N)^{\frac{1}{N}} = C$ where $\displaystyle C = \prod_{k=1}^{\infty} \left(\frac{(k+1)^2}{k(k+2)}\right)^{\frac{\log(k)}{\log 2}}.$

(iv) If $\frac{p_n}{q_n}$ are the convergents of x, then $\lim_{n\to\infty} \frac{1}{n} \log |x - \frac{p_n}{q_n}| = -\frac{\pi^2}{6\log(2)}$. In particular, $|x - \frac{p_n}{q_n}| = O(\mathrm{e}^{-\lambda n})$ for all $0 < \lambda < \frac{\pi^2}{6\log(2)}$.

As point (iv) is quite challenging to obtain, we will omit its proof.

Proof. (i) The frequency of the digit k in the expansion $[a_0; a_1, a_2, \dots]$ is given by

$$\lim_{N\to\infty} \frac{1}{N} |\{1 \le n \le N : a_n = k\}|.$$

As observed above, $a_n = k$ if and only if $T^{n-1}x \in (\frac{1}{k+1}, \frac{1}{k}]$. Therefore

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N} |\{1 \le n \le N : a_n = k\}| &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_{(\frac{1}{k+1}, \frac{1}{k}]}(T^{n-1}x) \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{1}_{(\frac{1}{k+1}, \frac{1}{k}]}(T^{n}x) \\
&= \int_{[0,1)} \mathbf{1}_{(\frac{1}{k+1}, \frac{1}{k}]} \, \mathrm{d}\mu(x) \\
&= \mu\left(\left(\frac{1}{k+1}, \frac{1}{k}\right]\right)
\end{aligned}$$

using Corollary 3.4 for the third equality, which then holds for μ–almost every $x \in [0,1)$. A direct computation yields to

$$\mu\left(\left(\frac{1}{k+1}, \frac{1}{k}\right]\right) = \frac{1}{\log 2} \int_{\frac{1}{k+1}}^{\frac{1}{k}} \frac{1}{1+x} \, \mathrm{d}x = \frac{2\log(k+1) - \log(k) - \log(k+2)}{\log 2}$$

and thus (i) holds for μ–almost every $x \in [0,1)$. Since μ and the Lebesgue measure share the same null sets, (i) holds also for Lebesgue almost every $x \in [0,1)$.

(ii) We have

$$\lim_{N\to\infty} \frac{a_1 + \cdots + a_N}{N} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} a(T^n x).$$

Define also $a_M(x) = \sum_{k=1}^{M} k\mathbf{1}_{(\frac{1}{k+1},\frac{1}{k}]}(x)$ and observe that a_M converges pointwise to a as $M \to \infty$. By Corollary 3.4, we have then

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} a_M(T^n x) = \int_{[0,1)} a_M \, \mathrm{d}\mu = \sum_{k=1}^{M} k\mu\left(\left(\frac{1}{k+1}, \frac{1}{k}\right]\right) = \frac{1}{\log 2} \sum_{k=1}^{M} k \log\left(\frac{(k+1)^2}{k(k+2)}\right)$$

for μ−almost every $x \in [0,1)$, so also for Lebesgue almost every $x \in [0,1)$. The claim follows by letting $M \to \infty$ and using that $\sum_{k=1}^{\infty} k \log\left(\frac{(k+1)^2}{k(k+2)}\right) = \infty$.

(iii) Define $f(x) = \log(a(x))$ and apply Corollary 3.4 to get

$$\lim_{N\to\infty} \log((a_1 \ldots a_N)^{\frac{1}{N}}) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x) = \int_{[0,1)} f \, \mathrm{d}\mu.$$

Note that for a fixed $x \in [0,1)$, $f(x)$ is in fact an integer, equals to k if $x \in (\frac{1}{k+1}, \frac{1}{k}]$, so that we can write

$$f(x) = \sum_{k=1}^{\infty} \log(k)\mathbf{1}_{(\frac{1}{k+1},\frac{1}{k}]}(x).$$

A straightforward calculation shows then that

$$\int_{[0,1)} f \, \mathrm{d}\mu = \sum_{k=1}^{\infty} \frac{\log(k)}{\log 2} \int_{\frac{1}{k+1}}^{\frac{1}{k}} \frac{1}{1+x} \, \mathrm{d}x = \log(C)$$

where $C := \prod_{k=1}^{\infty} \left(\frac{(k+1)^2}{k(k+2)}\right)^{\frac{\log(k)}{\log 2}}$. This proves (iii). □

3.5 Beatty sequences

This short subsection is devoted to another application of ergodic theory to number theory. Recall that the *floor function* $\lfloor \cdot \rfloor \colon \mathbb{R} \longrightarrow \mathbb{Z}$ takes any real number x to the integer $\lfloor x \rfloor$ such that $\lfloor x \rfloor < x \leq \lfloor x \rfloor + 1$.

Definition 3.13. A Beatty sequence is a sequence of integers of the form $(\lfloor n\alpha \rfloor)_{n\geq 1}$ where $\alpha > 1$ is an irrational number.

In the proof of the result below, we even do not need to apply Birkhoff's Theorem. It just suffices to translate the problem suitably into a dynamical setting.

Theorem 3.14. The complement of a Beatty sequence is a Beatty sequence.

Proof. Let $A := \{\lfloor n\alpha \rfloor \mid n \geq 1\}$ and $B := \{\lfloor n\beta \rfloor \mid n \geq 1\}$, where $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. Note that also β is irrational since α is irrational. We will show that A and B form a partition of $\mathbb{N}$. Denote by $R_{\alpha^{-1}}$ and $R_{\beta^{-1}}$ the toral rotations by α^{-1} and β^{-1} respectively. Observe that

$$\begin{aligned} m \in A &\iff \exists n \geq 1,\ m = \lfloor n\alpha \rfloor \\ &\iff n\alpha - 1 < m \leq n\alpha \\ &\iff n - \frac{1}{\alpha} < \frac{m}{\alpha} \leq n \\ &\iff \left\{\frac{m}{\alpha}\right\} \in \left(1 - \frac{1}{\alpha}, 1\right] \end{aligned}$$

which shows that $A = \left\{m \in \mathbb{N} \mid R^m_{\alpha^{-1}}(0) \in \left(1 - \frac{1}{\alpha}, 1\right]\right\}$. One obtains a similar description for B, with α replaced by β. Now for all $m \geq 1$ we have

$$\left\{\frac{m}{\alpha}\right\} + \left\{\frac{m}{\beta}\right\} = \frac{m}{\alpha} + \frac{m}{\beta} - \left\lfloor\frac{m}{\alpha}\right\rfloor - \left\lfloor\frac{m}{\beta}\right\rfloor = m - \left\lfloor\frac{m}{\alpha}\right\rfloor - \left\lfloor\frac{m}{\beta}\right\rfloor$$

so the left-hand side is an integer equals to either 0, 1 or 2, and we easily see it cannot be 0 or 2. Thus $\left\{\frac{m}{\alpha}\right\} + \left\{\frac{m}{\beta}\right\} = 1$, and we deduce then

$$m \in B \iff R^m_{\beta^{-1}}(0) \in \left(1 - \frac{1}{\beta}, 1\right] \iff R^m_{\alpha^{-1}}(0) \in \left[0, 1 - \frac{1}{\alpha}\right) \iff m \notin A$$

so that $B = \mathbb{N} \setminus A$, as announced. □

4. Classifying measure-preserving systems

In ergodic theory, protagonists are measure-preserving systems, and as any other structured object in mathematics it is natural to develop the tools to classify them. We want to understand when the dynamical behaviour of two measure-preserving systems are independent, correlated, or identical. We will essentially turn our attention to two types of systems, arising naturally everywhere, and that will play a crucial role in the next section.

4.1 Factors, extensions, isomorphisms

To reasonably classify some dynamical systems, we need to adopt several terminologies.

Definition 4.1. Let $(X, \mathcal{A}, \mu, T)$, $(Y, \mathcal{B}, \nu, S)$ be measure-preserving systems. A measurable map $\pi\colon X \longrightarrow Y$ is a factor map if

(i) $\pi(X)$ has full measure.

(ii) $\pi_*\mu = \nu$.

(iii) $\pi \circ T(x) = S \circ \pi(x)$ for μ−a.e. $x \in X$.

The first condition means that π is almost surjective, in the sense that the complement of its image has zero measure. When a map satisfies (iii) above, we say it *intertwines* the transformations T and S.

When there is a factor map from X to Y, we say that $(X, \mathcal{A}, \mu, T)$ is an *extension* of $(Y, \mathcal{B}, \nu, S)$, and that $(Y, \mathcal{B}, \nu, S)$ is a *factor* of $(X, \mathcal{A}, \mu, T)$.

We can now state a reasonable definition of isomorphic measure-preserving systems.

Definition 4.2. A factor map $\varphi\colon X \longrightarrow Y$ between two measure-preserving systems $(X, \mathcal{A}, \mu, T)$ and $(Y, \mathcal{B}, \nu, S)$ is called an isomorphism if there exists a factor map $\psi\colon Y \longrightarrow X$ such that $\psi \circ \varphi$ and $\varphi \circ \psi$ are almost everywhere the identities.

When this occurs, we say X and Y are *isomorphic*.

Example 4.3. Consider $X = Y = \mathbb{T}$ the torus, along with its Borel σ−algebra and μ the Lebesgue measure. Consider R_α a rotation, and T the doubling map. Then these two systems are *not* isomorphic. To reach a contradiction, suppose there is a factor map

$$\varphi\colon (X, \mathcal{B}(X), \mu, T) \longrightarrow (X, \mathcal{B}(X), \mu, R_\alpha)$$

invertible on a full measure subset $U \subset X$. Then $\varphi \circ T = R_\alpha \circ \varphi$ on U, and since R_α is invertible on U, T has to be invertible on U. Now, note that for any $x \in U$, $T(x + \frac{1}{2}) = 2x = T(x)$, and injectivity then forces $x + \frac{1}{2} \notin U$. But then $1 = m_X(U) \leq \frac{1}{2}$, a contradiction.

4.2 Kronecker systems

The first class of dynamical systems we study consists of those whose behaviour is, in some sense, predictable.

Definition 4.4. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system.
A non-zero function $f \in L^2(X, \mathcal{A}, \mu)$ is an eigenfunction for T if there exists $\lambda \in \mathbb{C}$ so that $f \circ T = \lambda f$ μ–almost everywhere.

In that case, we say λ is the eigenvalue associated to f. Observe that f is an eigenfunction if it is an eigenvector for the Koopman operator U_T on $L^2(X, \mathcal{A}, \mu)$.

The set of all eigenvalues of U_T is called the *point-spectrum* of T and is denoted $\sigma(T)$. An eigenvalue is *simple* if its eigenspace is one-dimensional.

Let us first look at some basic properties of eigenfunctions.

Lemma 4.5. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system.

(i) The set $\sigma(T)$ is a subgroup of $(\mathbb{S}^1, \cdot)$, where $\mathbb{S}^1 \subset \mathbb{C}$ is the unit circle.

(ii) If f, g are eigenfunctions with eigenvalues λ_f and λ_g and if $\lambda_f \neq \lambda_g$, then $\langle f, g\rangle = 0$.

(iii) If the system is ergodic, then for every eigenfunction f, $|f|$ is constant almost everywhere, and every eigenvalue is simple.

Proof. (i) To start, fix $\lambda \in \sigma(T)$, and let $f \in L^2(X, \mathcal{A}, \mu)$ be an eigenfunction of eigenvalue λ. Then, using Lemma 2.2, we have

$$\|f\|_2^2 = \langle f, f\rangle = \langle U_T f, U_T f\rangle = \langle \lambda f, \lambda f\rangle = |\lambda|^2 \langle f, f\rangle = |\lambda|^2 \|f\|_2^2.$$

Using that $\|f\| \neq 0$, we get $|\lambda| = 1$, hence $\lambda \in \mathbb{S}^1$. Also note that $\overline{f}$ is an eigenfunction with eigenvalue $\overline{\lambda} = \lambda^{-1}$, proving that $\lambda^{-1} \in \sigma(T)$. Lastly, if λ and μ are two eigenvalues with eigenfunctions f and g, then fg is an eigenfunction with eigenvalue $\lambda\mu$, proving $\lambda\mu \in \sigma(T)$. Thus $\sigma(T)$ is a subgroup of $\mathbb{S}^1$.

(ii) Again, since U_T preserves inner products, one has

$$\langle f, g\rangle = \langle U_T f, U_T g\rangle = \lambda_f \overline{\lambda_g}\langle f, g\rangle$$

and since $\lambda_f \neq \lambda_g$, the only way this equation holds is to have $\langle f, g\rangle = 0$.

(iii) Suppose $f \in L^2(X, \mathcal{A}, \mu)$ is an eigenfunction with eigenvalue λ. Then it follows that

$$|f| \circ T(x) = |f(Tx)| = |\lambda f(x)| = |f|(x)$$

for μ–almost every $x \in X$. Hence $|f|$ is almost everywhere invariant, and by ergodicity, $|f|$ is almost everywhere constant, as claimed. Finally, if f_1 and f_2 are two eigenfunctions with the same eigenvalue λ, then $\frac{f_1}{f_2}$ is invariant, and thus constant almost everywhere by ergodicity. This implies f_1 is a scalar multiple of f_2, and concludes the proof. □

Remarquably, the converse of (i) above is also true, namely every countable subgroup of $\mathbb{S}^1$ can be realized as the spectrum of an ergodic measure-preserving system.

Example 4.6. Let $(X, \mathcal{A}, \mu, T)$ be a circle rotation, *i.e.* $X = [0, 1]$ is equipped with its Borel σ–algebra $\mathcal{A}$, μ is the Lebesgue measure, and $T(x) = x + \alpha \mod 1$. The function $f(x) = \mathrm{e}^{2\pi i x}$ is an eigenfunction, since

$$U_T f(x) = f(Tx) = f(x + \alpha) = \mathrm{e}^{2\pi i(x+\alpha)} = \mathrm{e}^{2\pi i\alpha} f(x)$$

and the corresponding eigenvalue is $\mathrm{e}^{2\pi i\alpha}$. In fact, all eigenfunctions are of the form $x \longmapsto c\mathrm{e}^{2\pi i n x}$ for some $c \in \mathbb{C}$ and $n \in \mathbb{Z}$, and the eigenvalues are $\mathrm{e}^{2\pi i n\alpha}$. Indeed, fix $f \in L^2(X, \mathcal{A}, \mu)$ an arbitrary eigenfunction, and write its Fourier expansion as

$$f(x) = \sum_{n\in\mathbb{Z}} c_n \mathrm{e}^{2\pi i n x}.$$

Using that $f(x + \alpha) = \lambda f(x)$, and the uniqueness of the Fourier coefficients, it follows that

$$c_n \mathrm{e}^{2\pi i n\alpha} = \lambda c_n$$

for all $n \in \mathbb{Z}$. If there is two integers $n \neq m$ such that $c_n, c_m \neq 0$, then $\lambda = \mathrm{e}^{2\pi i n\alpha}$ and also $\lambda = \mathrm{e}^{2\pi i m\alpha}$, which is excluded. Thus $c_n \neq 0$ for exactly one integer $n \in \mathbb{Z}$, proving the claim.

This example can be generalized in a strong manner to arbitrary group rotations.

Proposition 4.7. Let $(X, \mathcal{A}, \mu, T)$ be a group rotation.
Then there exists a basis of $L^2(X, \mathcal{A}, \mu)$ consisting of eigenfunctions of T.

Proof. Recall that a character of X is a continuous homomorphism $\chi\colon X \longrightarrow \mathbb{S}^1$. By the Stone-Weierstrass theorem, finite linear combinations of characters are dense in L^2, and any two distinct characters are orthogonal. Moreover, such a character is clearly an eigenfunction, since

$$\chi(Tx) = \chi(x+\alpha) = \chi(\alpha)\chi(x)$$

and the corresponding eigenvalue is $\chi(\alpha)$. Thus we are done. □

Such measure-preserving systems are important, and therefore get their own name.

Definition 4.8. A measure-preserving system $(X, \mathcal{A}, \mu, T)$ has discrete spectrum if there exists a basis of $L^2(X, \mathcal{A}, \mu)$ consisting of eigenfunctions of T. Moreover, if $(X, \mathcal{A}, \mu, T)$ is ergodic it is called a Kronecker system.

So, with this terminology, Proposition 4.7 tells precisely that an ergodic group rotation is a Kronecker system. It turns out it is essentially the only way of constructing such systems.

Theorem 4.9. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system, with $(X, \mathcal{A}, \mu)$ being a standard probability space. Then $(X, \mathcal{A}, \mu, T)$ is a Kronecker system if and only if it is isomorphic to an ergodic group rotation.

A standard probability is any probability space which is measurably isomorphic to $(X, \mathcal{A}, \mu)$ where X is compact metric, $\mathcal{A}$ is the Borel $\sigma-$algebra, and μ is a Borel probability measure on X.

Proof. That every ergodic group rotation is a Kronecker system is the content of Proposition 4.7. Let us then prove the converse. Since $(X, \mathcal{A}, \mu)$ is isomorphic to a standard probability space, we can assume without loss of generality that X is compact metric, $\mathcal{A}$ is the Borel $\sigma-$algebra and μ is a Borel probability measure. Let $\chi_1, \chi_2, \dots$ be an orthonormal basis of $L^2(X, \mathcal{A}, \mu)$ consisting of eigenfunctions, and denote $\lambda_1, \lambda_2, \dots$ the corresponding eigenvalues. In view of Lemma 4.5(iii), we may assume that χ_n takes values in $\mathbb{S}^1$. Let then

$$\varphi\colon X \longrightarrow \mathbb{S}^{\mathbb{N}}$$

be the map defined as $\varphi(x) := (\chi_1(x), \chi_2(x), \dots)$. Also let $\alpha := (\lambda_1, \lambda_2, \dots) \in \mathbb{S}^{\mathbb{N}}$ and $Y := \overline{\{\alpha^n \mid n \in \mathbb{Z}\}}$. Then Y is a subgroup of $\mathbb{S}^{\mathbb{N}}$. Moreover this product is compact by Tychonoff's theorem, and Y is closed by definition, so Y is also compact. Consider then $\mathcal{B}$ the Borel $\sigma-$algebra on Y, ν its normalized Haar measure, and the transformation $S(y) := \alpha y$. The system $(Y, \mathcal{B}, \nu, S)$ is a group rotation, and we claim it is isomorphic to $(X, \mathcal{A}, \mu, T)$.

First, we prove that φ is an isomorphism between X and $(Y^*, \mathcal{B}^*, \nu^*, S)$, where $Y^* = \varphi(X)$, $\mathcal{B}^*$ is the Borel $\sigma-$algebra restricted to Y^*, and $\nu^* = \varphi_*\mu$. Note that by definition we have

$$\varphi(Tx) = (\chi_1(Tx), \chi_2(Tx), \dots) = (\lambda_1\chi_1(x), \lambda_2\chi_2(x), \dots) = \alpha\varphi(x) = S(\varphi(x))$$

so φ intertwines T and S. Moreover clearly its image has full measure in Y^*, and ν^* is the pushforward of μ under φ. Thus φ is a factor map. Also it is surjective by definition. Hence we only need to prove it is almost everywhere injective. Denote d the metric on X, and fix $\varepsilon > 0$. We are going to show there is a set of measure at least $1-\varepsilon$ on which φ is almost injective. Let $B_1, \dots, B_r$ be a finite collection of balls of radius at most ε that cover X. Such a cover exists because X is compact. Since $\chi_1, \chi_2, \dots$ form a basis of $L^2(X, \mathcal{A}, \mu)$, the subspace $\operatorname{span}\{\chi_1, \chi_2, \dots\}$ is dense in $L^2(X, \mathcal{A}, \mu)$, so we can find $f_1, \dots, f_r \in \operatorname{span}\{\chi_1, \chi_2, \dots\}$ so that

$$\|f_i - \mathbf{1}_{B_i}\|_2 \le \frac{\varepsilon}{2r}$$

for all $i = 1, \dots, r$. Let $\Delta_\varepsilon := \big\{x \in X : \max_{1\le i\le r} |f_i(x) - \mathbf{1}_{B_i}(x)| \ge \frac{1}{2}\big\}$. By Chebyshev's inequality, one has then

$$\begin{aligned}
\mu(\Delta_\varepsilon) &\le 2\int_X \max_{1\le i\le r} |f_i(x) - \mathbf{1}_{B_i}(x)| \,\mathrm{d}\mu(x) \\
&\le \sum_{i=1}^r 2\int_X |f_i(x) - \mathbf{1}_{B_i}(x)| \,\mathrm{d}\mu(x) \\
&\le 2r\frac{\varepsilon}{2r} = \varepsilon
\end{aligned}$$

using that $\|\cdot\|_1 \le \|\cdot\|_2$. Set then $\Omega_\varepsilon := X \setminus \Delta_\varepsilon$, and note that $\mu(\Omega_\varepsilon) \ge 1-\varepsilon$. Let $x, y \in \Omega_\varepsilon$ and suppose $\varphi(x) = \varphi(y)$. By definition of φ, this means $\chi_i(x) = \chi_i(y)$ for all $i \ge 1$, and in particular it implies $f_i(x) = f_i(y)$ for all $i = 1, \dots, r$. We therefore get

$$\begin{aligned}
|\mathbf{1}_{B_i}(x) - \mathbf{1}_{B_i}(y)| &= |\mathbf{1}_{B_i}(x) - f_i(x) + f_i(y) - \mathbf{1}_{B_i}(y)| \\
&\le |\mathbf{1}_{B_i}(x) - f_i(x)| + |f_i(y) - \mathbf{1}_{B_i}(y)| \\
&< \frac{1}{2} + \frac{1}{2} = 1
\end{aligned}$$

and so $\mathbf{1}_{B_i}(x) = \mathbf{1}_{B_i}(y)$ for all $i = 1, \dots, r$. We deduce that x and y are in a common ball, meaning $d(x,y) \le \varepsilon$. This proves that φ is almost injective on Ω_ε. Finally, consider

$$\Omega := \bigcap_{n\ge1}\bigcup_{k\ge n} \Omega_{\frac{1}{k}}$$

and note that by monotonicity Ω has full measure. Moreover if $x, y \in \Omega$ are so that $\varphi(x) = \varphi(y)$, then $d(x,y) \le \frac{1}{n}$ for all $n \ge 1$ by what we have shown above, so in fact $x = y$, and φ is injective on Ω. This concludes the proof that φ is an isomorphism.

It remains to prove now that $(Y^*, \mathcal{B}^*, \nu^*, S)$ and $(Y, \mathcal{B}, \nu, S)$ are isomorphic. Consider the quotient $\mathbb{S}^{\mathbb{N}}/Y$ and the natural projection $\pi\colon \mathbb{S}^{\mathbb{N}} \longrightarrow \mathbb{S}^{\mathbb{N}}/Y$. Consider ρ the push-forward of ν^* under π. If it not a point mass, then we find a subset C of the quotient with $0 < \rho(C) < 1$. The set $A = \pi^{-1}(C)$ is then an invariant subset of Y^* with $0 < \nu^*(C) < 1$, which is impossible because $(Y^*, \mathcal{B}^*, \nu^*, S)$ is ergodic. Hence ρ is a point mass, and we denote uY its support. It is now direct to prove that the map $\psi(y) = uy$ is an isomorphism from $(Y, \mathcal{B}, \nu, S)$ to $(Y^*, \mathcal{B}^*, \nu^*, S)$ and we leave the details to the interested reader. $\square$

4.3 Weakly mixing systems

On the other hand, we are now going to characterize some systems that tend to be unpredictable.

Definition 4.10. A measure-preserving system $(X, \mathcal{A}, \mu, T)$ is weak-mixing if the product system $(X \times X, \mathcal{A} \otimes \mathcal{A}, \mu \otimes \mu, T \times T)$ is ergodic.

The following proposition provides several equivalent characterizations of weak-mixing systems.

Theorem 4.11. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system. The following are equivalent.

(i) The system $(X, \mathcal{A}, \mu, T)$ is weak-mixing.

(ii) For any ergodic measure-preserving system $(Y, \mathcal{B}, \nu, S)$, the product $(X \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu, T \times S)$ is ergodic.

(iii) For any $A, B \in \mathcal{A}$, we have

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

(iv) For any $f, g \in L^2(X, \mathcal{A}, \mu)$, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} |\langle U_T^n f, g\rangle - \langle f, \mathbf{1}_X\rangle\langle \mathbf{1}_X, g\rangle| = 0.$$

(v) For any $A, B \in \mathcal{A}$, there exists $E \subset \mathbb{N}$ with zero density such that

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

Proof. (i) $\Longrightarrow$ (iv) : Suppose X is weak mixing. Fix $f, g \in L^2(X, \mathcal{A}, \mu)$. Replacing f by $f - \int_X f \, \mathrm{d}\mu$ if necessary, we can assume that $\displaystyle\int_X f \, \mathrm{d}\mu = 0$. By Cauchy-Schwartz inequality, we have

$$\begin{aligned}
\limsup_{N\to\infty} \left(\frac{1}{N}\sum_{n=1}^{N} |\langle U_T^n f, g\rangle|\right)^2 &\leq \limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} |\langle U_T^n f, g\rangle|^2 \\
&= \limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \langle U_{T\times T}^n (f\otimes \overline{f}), g\otimes \overline{g}\rangle \\
&= \langle \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} U_{T\times T}^n (f\otimes \overline{f}), g\otimes \overline{g}\rangle
\end{aligned}$$

and the last equality holds because the product of X with itself is ergodic, so Theorem 2.9 implies that $\displaystyle\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} U_{T\times T}^n (f\otimes\overline{f}) = \int_{X\times X} f\otimes\overline{f} \, \mathrm{d}\mu\otimes\mu$ in L^2–norm, and therefore also weakly, allowing the exchange of the limit and the inner product. Now

$$\int_{X\times X} f\otimes\overline{f} \, \mathrm{d}\mu\otimes\mu = \left(\int_X f \, \mathrm{d}\mu\right)^2 = 0$$

and thus (iv) holds.

(iv) $\Longrightarrow$ (iii) : apply (iv) with $f = \mathbf{1}_A$ and $g = \mathbf{1}_B$.

(iii) $\Longrightarrow$ (iv) : this follows from (iii) and the fact that finite linear combinations of indicator functions are dense in $L^2(X, \mathcal{A}, \mu)$.

(iii) $\Longrightarrow$ (v) : For $m \in \mathbb{N}$, let $A_m := \{n \in \mathbb{N} \mid |\mu(T^{-n}A\cap B) - \mu(A)\mu(B)| > \frac{1}{m}\}$. Then one has

$$\frac{|A_m \cap \{1, \ldots, N\}|}{N} \leq \frac{m}{N}\sum_{n=1}^{N} |\mu(T^{-n}A\cap B) - \mu(A)\mu(B)|$$

and letting $N \to \infty$ and using (iii), we get $\overline{d}(A_m) = 0$. Henceforth, for $m \in \mathbb{N}$, we can find $N_m \in \mathbb{N}$ so that

$$N \geq N_m \Longrightarrow \frac{|A_m \cap \{1, \ldots, N\}|}{N} \leq \frac{1}{m}.$$

Define then $E := \displaystyle\bigcup_{m=1}^{\infty} (A_m \cap [N_m+1, N_{m+1}))$. Since $A_m \subset A_{m+1}$, for $N \in \mathbb{N}$, we can choose $m \in \mathbb{N}$ so that $N \in [N_m+1, N_{m+1})$, and hence $E\cap\{1, \ldots, N\} \subset A_m \cap \{1, \ldots, N\}$. This provides $|E\cap\{1, \ldots, N\}| \leq \frac{N}{m}$, and thus $\overline{d}(E) = 0$ as well.

(v) $\Longrightarrow$ (iii) : It suffices to observe that

$$\frac{1}{N}\sum_{n=1}^{N} |\mu(T^{-n}A\cap B) - \mu(A)\mu(B)| = \frac{1}{N}\sum_{n\in E\cap\{1,\ldots,N\}} |\mu(T^{-n}A\cap B) - \mu(A)\mu(B)|$$

$$+\frac{1}{N}\sum_{n\in\{1,\ldots,N\}\setminus E}|\mu(T^{-n}A\cap B)-\mu(A)\mu(B)|$$

and that the first term is 0 since bounded by $\overline{d}(E)$, and the second goes to 0 as $N\to\infty$ by (v).

(iv) $\Longrightarrow$ (ii) : To come.

(ii) $\Longrightarrow$ (i) : It suffices to prove that $(X,\mathcal{A},\mu,T)$ is ergodic, since therefore it will follow from (ii) applied with $Y = X$ that the product of X with itself is ergodic, *i.e.* X is weak-mixing. Hence, for a contradiction, suppose that X is not ergodic, and let $A\in\mathcal{A}$ be a non-trivial invariant set. Consider $(Y,\mathcal{B},\nu,S)$ the ergodic one-point system. Then $A\times Y$ is a non-trivial invariant set of the product system $X\times Y$. But by (ii), $X\times Y$ is ergodic, so has no non-trivial invariant sets. Thus $(X,\mathcal{A},\mu,T)$ is ergodic as announced. □

This theorem has the following immediate consequences.

Corollary 4.12. Let $(X,\mathcal{A},\mu,T)$ be a weak mixing system. Then the following holds.
(i) The system $(X,\mathcal{A},\mu,T)$ is ergodic.
(ii) The $k-$fold product $(X^k,\mathcal{A}^k,\mu^{\otimes k},T^k)$ is weak-mixing.
(iii) All eigenfunctions of T are constant almost everywhere.

Point (iii) above is an opposite version of having discrete spectrum, and thus systems with this property are said to have *continuous spectrum*. Then, weak mixing systems have continuous spectrum.

Proof. (i) directly follows from point (iv) of the previous theorem, and Corollary 2.10.

(ii) By induction, it is enough to prove the case $k = 2$. X being weak mixing means $X\times X$ is ergodic. Therefore, by (ii) of Theorem 4.11, we deduce that $X\times X\times X$ is ergodic. Applying this one more time, it follows that $X\times X\times X\times X$ is ergodic, meaning that $X\times X$ is weak mixing.

(iii) Let $f\in L^2(X,\mathcal{A},\mu)$ be an eigenfunction of U_T, and $\lambda\in\mathbb{S}^1$ be the corresponding eigenvalue. Observe that, for all $(x,y)\in X\times X$, one has

$$U_{T\times T}(f\otimes\overline{f})(x,y) = f\otimes\overline{f}(Tx,Ty) = f(Tx)\overline{f}(Ty) = \lambda f(x)\overline{\lambda}\overline{f}(y) = f\otimes\overline{f}(x,y)$$

using that $\lambda\overline{\lambda} = |\lambda|^2 = 1$. Thus $f\otimes\overline{f}$ is invariant under $U_{T\times T}$, and since $(X\times X,\mathcal{A}\otimes\mathcal{A},\mu\otimes\mu,T\times T)$ is ergodic, Proposition 1.16 implies $f\otimes\overline{f}$ is constant almost everywhere. Hence f is also constant almost everywhere, proving the claim. □

It turns out that, in fact, having all eigenfunctions almost everywhere constant *characterize* weak mixing systems. This result will follow from the Jacobs-de Leeuw-Glicksberg decomposition established in the next section.

However, the converse of (i) in Corollary 4.12 is not true, *i.e.* an ergodic system is not necessarily weak mixing. Consider for instance an ergodic group rotation, such as $R_\alpha \colon \mathbb{T} \longrightarrow \mathbb{T}$ with $\alpha \notin \mathbb{Q}$.

The next proposition shows we can actually relax the condition (iv) in Theorem 4.11.

Proposition 4.13. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system Then it is weak mixing if and only if for any $f \in L^2(X, \mathcal{A}, \mu)$, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} |\langle U_T^n f, f\rangle - \langle f, \mathbf{1}_X\rangle\langle \mathbf{1}_X, f\rangle| = 0.$$

Proof. $\Longrightarrow$: If the system is weak mixing, it satisfies point (iv) of Theorem 4.11, which we may apply with $g = f$ to get the claim.
$\Longleftarrow$: This direction relies on the following two identities, usually called *polarization identities*, namely that

$$\begin{aligned} 4\langle U_T^n f, g\rangle &= \langle U_T^n(f+g), f+g\rangle + i\langle U_T^n(f+ig), f+ig\rangle \\ &\quad - \langle U_T^n(f-g), f-g\rangle - i\langle U_T^n(f-ig), f-ig\rangle \end{aligned}$$

and similarly

$$\begin{aligned} 4\langle f, 1\rangle\langle 1, g\rangle &= \langle f+g, 1\rangle\langle 1, f+g\rangle + i\langle f+ig, 1\rangle\langle 1, f+ig\rangle \\ &\quad - \langle f-g, 1\rangle\langle 1, f-g\rangle - i\langle f-ig, 1\rangle\langle 1, f-ig\rangle \end{aligned}$$

for all $f, g \in L^2(X, \mathcal{A}, \mu)$, where we denote by 1 the function $\mathbf{1}_X$. Using these equalities, one shows directly that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} |\langle U_T^n f, g\rangle - \langle f, \mathbf{1}_X\rangle\langle \mathbf{1}_X, g\rangle| = 0$$

for all $f, g \in L^2(X, \mathcal{A}, \mu)$, which implies that $(X, \mathcal{A}, \mu, T)$ is weak mixing by Theorem 4.11. □

4.4 Mixing systems

In this subsection, we provide a stronger version of weak mixing, namely *strong mixing*, or *mixing* systems.

Definition 4.14. A measure-preserving system $(X, \mathcal{A}, \mu, T)$ is mixing if

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

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

From this definition, it is clear that mixing systems are weak mixing, in particular ergodic.

As for weak-mixing systems, let us reformulate this notion in several manners.

Proposition 4.15. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system. The following are equivalent.

(i) The system $(X, \mathcal{A}, \mu, T)$ is mixing.

(ii) For any $f, g \in L^2(X, \mathcal{A}, \mu)$, we have

$$\lim_{n\to\infty} \langle U_T^n f, g\rangle = \langle f, \mathbf{1}_X\rangle\langle \mathbf{1}_X, g\rangle.$$

(iii) For any $f \in L^2(X, \mathcal{A}, \mu)$ with $\displaystyle\int_X f \,\mathrm{d}\mu = 0$, the sequence $(U_T^n f)_{n\geq 0}$ converges weakly to 0.

Proof. (i) $\Longrightarrow$ (ii) follows from the fact that finite linear combinations of indicator functions are dense in $L^2(X, \mathcal{A}, \mu)$. Also (ii) $\Longrightarrow$ (i) and (ii) $\Longrightarrow$ (iii) are immediate.
(iii) $\Longrightarrow$ (ii) : Fix $f, g \in L^2(X, \mathcal{A}, \mu)$. We set $\tilde{f} := f - \displaystyle\int_X f \,\mathrm{d}\mu$, and applying (iii) with this new function, which integrates to 0, yields

$$\lim_{n\to\infty} \langle U_T^n \tilde{f}, g\rangle = \langle 0, g\rangle = 0$$

and on the other hand, using linearity of the inner product, we have

$$\langle U_T^n \tilde{f}, g\rangle = \langle U_T^n f - \langle f, \mathbf{1}_X\rangle, g\rangle = \langle U_T^n f, g\rangle - \langle f, \mathbf{1}_X\rangle\langle \mathbf{1}_X, g\rangle$$

and this implies (ii). □

Here are some examples of (non-)mixing systems.

Example 4.16. (i) Consider a rotation on 2 points, *i.e.* $X = \{0, 1\}$, $\mathcal{A} = \mathcal{P}(X)$, and $\mu(\{0\}) = \mu(\{1\}) = \frac{1}{2}$, $T(0) = 1$, $T(1) = 0$. Then $T^{2k} = \mathrm{Id}_X$, and $T^{2k+1} = T$ for all $k \geq 0$. For $A = B = \{0\}$, we have then

$$\mu(A \cap T^{-2k}A) = \mu(A) = \frac{1}{2},\ \mu(A \cap T^{-(2k+1)}A) = \mu(\{0\} \cap \{1\}) = 0$$

so the limit condition does not hold, and the system is not mixing.

Alternatively, as already seen, the product system $X \times X$ is not ergodic, so X is not weak mixing, and therefore not mixing either. This generalizes to rotations on an arbitrary number $m \geq 2$ of points.

(ii) The torus rotation $R_\alpha \colon \mathbb{T} \longrightarrow \mathbb{T}$ is not mixing, and this can be seen in several ways. First, if $\alpha \in \mathbb{Q}$, it is even not ergodic. Suppose then $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Then in fact R_α is not weak mixing, because for instance $R_\alpha \times R_\alpha$ is not ergodic (Example 2.4), or because it has a lot of non-constant eigenfunctions, namely all $x \longmapsto \mathrm{e}^{2\pi i n x}$, $n \in \mathbb{Z}$.

(iii) A Bernoulli shift is mixing. Using notations of Example 1.5, fix two cylinder sets

$$A := \{(x_n)_{n\geq 0} \in X \mid x_0 = a_0, \dots, x_m = a_m\},\ B := \{(x_n)_{n\geq 0} \in X \mid x_0 = b_0, \dots, x_k = b_k\}$$

where $a_0, \dots, a_m, b_0, \dots, b_k \in \{0, 1\}$. From the definition of the left shift $T \colon X \longrightarrow X$, we get that

$$T^{-j}A = \prod_{i=0}^{j-1}\{0,1\} \times \{a_0\} \times \cdots \times \{a_m\} \times \prod_{i \geq j+m+1} \{0,1\}$$

for all $j \geq 0$. Hence, for j large enough, we have

$$T^{-j}A \cap B = \{b_0\} \times \cdots \times \{b_k\} \times \cdots \times \{a_0\} \times \cdots \times \{a_m\} \times \prod \{0,1\}$$

and it follows that $\mu(T^{-j}A \cap B) = \mu(A)\mu(B)$ for j large enough. We therefore have the limit condition satisfied for cylinder sets, and since those generate the σ−algebra on X, the system is mixing.

We are going to formulate a more general version of Proposition 4.15.

Proposition 4.17. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system. Then it is mixing if and only if for any f, g in a dense subset of $L^2(X, \mathcal{A}, \mu)$, we have

$$\lim_{n\to\infty} \langle U_T^n f, g\rangle = \langle f, \mathbf{1}_X\rangle\langle \mathbf{1}_X, g\rangle.$$

Proof. If the system is mixing, the conclusion has already been established in 4.15, recalling that measurable simple functions are dense in $L^2(X, \mathcal{A}, \mu)$.

Conversely, denote by S the subset of $L^2(X, \mathcal{A}, \mu)$ for which the hypothesis holds. By Proposition 4.15, it is enough to prove that

$$\lim_{n\to\infty} \langle U_T^n f, g\rangle = 0$$

for all $f, g \in L^2(X, \mathcal{A}, \mu)$, under the assumption that $\langle f, \mathbf{1}_X\rangle = 0$. Note that the result is clear if $f = 0$ or $g = 0$. Hence we may assume $f, g \neq 0$ and additionally

$\|f\|_2 = \|g\|_2 = 1$. Fix then $\varepsilon > 0$. By density of S, we find $\tilde{f}, \tilde{g} \in S$ such that $\|\tilde{f}\|_2 = \|\tilde{g}\|_2 = 1$ and

$$\|f - \tilde{f}\|_2 < \varepsilon, \ \|g - \tilde{g}\|_2 < \varepsilon.$$

Thus, the triangle and the Cauchy-Schwartz inequalities, and the fact that U_T preserves inner products imply that

$$\begin{aligned}
|\langle U_T^n f, g\rangle| &= |\langle U_T(f - \tilde{f} + \tilde{f}), g\rangle| \\
&\leq |\langle U_T^n(f - \tilde{f}), g\rangle| + |\langle U_T^n \tilde{f}, g\rangle| \\
&\leq \|U_T(f - \tilde{f})\|_2 \|g\|_2 + |\langle U_T^n \tilde{f}, g - \tilde{g}\rangle| + |\langle U_T^n \tilde{f}, \tilde{g}\rangle| \\
&\leq \varepsilon + \varepsilon + |\langle U_T^n \tilde{f}, \tilde{g}\rangle| \\
&= 2\varepsilon + |\langle U_T^n \tilde{f}, \tilde{g}\rangle|
\end{aligned}$$

for all $n \geq 1$. Also observe that $\left| \int_X \tilde{f} \ \mathrm{d}\mu \right| = \left| \int_X \tilde{f} - f \ \mathrm{d}\mu \right| \leq \|f - \tilde{f}\|_1 \leq \|f - \tilde{f}\|_2 < \varepsilon$ and that the hypothesis implies

$$|\langle U_T^n \tilde{f}, \tilde{g}\rangle| - |\langle \tilde{f}, \mathbf{1}_X\rangle\langle \mathbf{1}_X, \tilde{g}\rangle| \leq |\langle U_T^n \tilde{f}, \tilde{g}\rangle - \langle \tilde{f}, \mathbf{1}_X\rangle\langle \mathbf{1}_X, \tilde{g}\rangle| < \varepsilon$$

if n is large enough. Hence one gets

$$|\langle U_T^n \tilde{f}, \tilde{g}\rangle| < \varepsilon + |\langle \tilde{f}, \mathbf{1}_X\rangle\langle \mathbf{1}_X, \tilde{g}\rangle| < 2\varepsilon$$

if n is large enough. With the previous computation, we finally have $|\langle U_T^n f, g\rangle| < 4\varepsilon$ for n sufficiently large. Hence $\langle U_T f, g\rangle$ converges to 0, and we are done. □

This result allows us to give another important example of mixing systems.

Proposition 4.18. Let X be a compact abelian group, $\mathcal{B}(X)$ its Borel σ−algebra, m_X the Haar measure, and $T \colon X \longrightarrow X$ a continuous ergodic automorphism. Then $(X, \mathcal{B}(X), m_X, T)$ is mixing.

Proof. First, recall from Lemma 1.4 that $(X, \mathcal{B}(X), m_X, T)$ is a measure-preserving system. To prove the claim, we will prove that

$$\lim_{n\to\infty} \langle U_T^n \chi, \psi\rangle = \langle \chi, \mathbf{1}_X\rangle\langle \mathbf{1}_X, \psi\rangle$$

for all $\chi, \psi \in \hat{X}$ the character group of X, which is a dense subset of $L^2(X)$. Fix then such $\chi, \psi \in \hat{X}$. Note that if $\chi \equiv 1$ is the trivial character, the result holds clearly. We may assume then that $\chi \neq 1$. Then, by the orthogonality of characters, we have $\langle \chi, 1\rangle = 0$, so we are left to prove that

$$\lim_{n\to\infty} \langle U_T^n \chi, \psi\rangle = 0.$$

Observe that there is at most one $k \in \mathbb{N}$ such that $U_T^k \chi = \psi$. Indeed, if there are $p > q$ such that $U_T^p \chi = \psi = U_T^q \chi$, then $\chi \circ T^p = \chi \circ T^q$, so $\chi \circ T^{p-q} = \chi$ (the map $\chi \longmapsto \chi \circ T$ is injective since T is surjective). By Proposition 2.5, which we may apply since T is ergodic, it implies that χ is the trivial character, which is excluded. Henceforth, the orthogonality of characters implies that

$$\langle U_T^n \chi, \psi \rangle = 0$$

if n is large enough, finishing the proof. □

To close this subsection, let us mention also a notion of mixing of order k.

Definition 4.19. A measure-preserving system $(X, \mathcal{A}, \mu, T)$ is mixing of order $k \geq 2$ if for every $A_1, \ldots, A_k \in \mathcal{A}$ and every $a_1, \ldots, a_k : \mathbb{N} \longrightarrow \mathbb{N}$ with $\lim_{n\to\infty} a_i(n) - a_j(n) = \infty$ for all $1 \leq i < j \leq n$, one has

$$\lim_{n\to\infty} \mu(T^{-a_1(n)} A_1 \cap \cdots \cap T^{-a_k(n)} A_k) = \mu(A_1) \ldots \mu(A_k).$$

Note that mixing of order $k = 2$ is the same as mixing. It is clear that being mixing of order k implies being mixing of order $k - 1$. It is a major open problem in ergodic theory to know whether the converse holds, even for $k = 3$.

5. The Jacobs-de Leeuw-Glicksberg decomposition

We are now going to describe one of the most profound phenomenon in ergodic theory, a dichotomy between eigenfunctions and weak mixing structure. As seen above, a Kronecker system is spanned entirely by its eigenfunctions, whereas a weak mixing system has no non-constant eigenfunctions. A typical dynamical system is generally a mixture of these two extreme, but it is always possible to separate a structured component, closed to a Kronecker system, and thus predictable, and on the other hand a mixing, a chaotic component. This result is known as the *Jacobs-de Leeuw-Glicksberg decomposition*.

5.1 The spectral theorem

The first tool we will need is from functional analysis, and is a version of the spectral theorem for unitary operators.

Theorem 5.1. Let $U\colon \mathcal{H} \longrightarrow \mathcal{H}$ be a unitary operator on a Hilbert space.
For every $f \in \mathcal{H}$, there exists a unique finite Borel measure μ_f on the torus $\mathbb{T}$ such that

$$\langle U^n f, f\rangle = \int_{\mathbb{T}} \mathrm{e}^{2\pi i n x}\,\mathrm{d}\mu_f(x).$$

In this context, the measure μ_f is called the *spectral measure* of f.

Example 5.2. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system. Suppose $f \in L^2(X)$ is an eigenfunction with eigenvalue $\mathrm{e}^{2\pi i \alpha}$. Then one has

$$\langle U_T^n f, f\rangle = \mathrm{e}^{2\pi i n\alpha}\|f\|_2^2 = \int_{\mathbb{T}} \mathrm{e}^{2\pi i n x}\,\mathrm{d}(\|f\|_2^2\delta_\alpha)(x)$$

and by uniqueness it follows that $\mu_f = \|f\|_2^2\delta_\alpha$.

5.2 Weak mixing functions

Definition 5.3. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system.
A function $f \in L^2(X, \mathcal{A}, \mu)$ is called weak mixing if for all $g \in L^2(X, \mathcal{A}, \mu)$, we have

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} |\langle U_T^n f, g\rangle| = 0.$$

Note that a weak mixing function automatically satisfies $\int_X f \, \mathrm{d}\mu = 0$. Moreover, in view of Theorem 4.11, a system is weak mixing if and only if every function with zero integral is a weak mixing function.

The next proposition can be seen as an analog of Proposition 4.13 for weak mixing functions.

Lemma 5.4. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system and let $f \in L^2(X, \mathcal{A}, \mu)$. If we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |\langle U_T^n f, f\rangle| = 0$$

then f is a weak mixing function.

Proof. Fix $g \in L^2(X, \mathcal{A}, \mu)$. Consider the product space $(X \times X, \mathcal{A} \otimes \mathcal{A}, \mu \otimes \mu, T \times T)$, and denote $F = f \otimes \overline{f}$, $G = g \otimes \overline{g}$. By using twice the Mean Ergodic Theorem with F, we have

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |\langle U_T^n f, g\rangle|^2 &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \langle U_{T\times T}^n F, G\rangle \\
&= \langle F_{\text{inv}}, G\rangle \\
&\leq \|F_{\text{inv}}\|_2 \|G\|_2 \\
&= \|G\|_2 \langle F_{\text{inv}}, F_{\text{inv}}\rangle^{\frac{1}{2}} \\
&= \|G\|_2 \langle F_{\text{inv}}, F\rangle^{\frac{1}{2}} \\
&= \|G\|_2 \left(\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \langle U_{T\times T}^n F, F\rangle \right)^{\frac{1}{2}} \\
&= \|G\|_2 \left(\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |\langle U_T^n f, f\rangle|^2 \right)^{\frac{1}{2}} \\
&= 0
\end{aligned}$$

hence $\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |\langle U_T^n f, g\rangle| = 0$, and f is weak mixing. □

Here is an example that illustrates the use of this lemma.

Example 5.5. Consider the space $X = \mathbb{T}^2$ with the product $\sigma-$algebra, the product of two Haar measures, and the transformation $T\colon X \longrightarrow X$ given by

$$T(x, y) = (x + \alpha, x + y)$$

for all $(x, y) \in X$. Let $f, g \colon X \longrightarrow \mathbb{C}$ given by $f(x, y) = \mathrm{e}^{2\pi i x}$ and $g(x, y) = \mathrm{e}^{2\pi i y}$. Then we compute that

$$U_T f(x, y) = f(T(x, y)) = \mathrm{e}^{2\pi i(x+\alpha)} = \mathrm{e}^{2\pi i\alpha}\mathrm{e}^{2\pi i x} = \mathrm{e}^{2\pi i\alpha} f(x, y)$$

so f is an eigenfunction with eigenvalue $\mathrm{e}^{2\pi i\alpha}$. On the other hand, by induction we get $U_T^n g = (\mathrm{e}^{2\pi i\alpha})^{\frac{n(n-1)}{2}} f^n g$ for all $n \geq 0$, and this yields to

$$\begin{aligned}
|\langle U_T^n g, g\rangle| &= \left|\int_X (\mathrm{e}^{2\pi i\alpha})^{\frac{n(n-1)}{2}} f^n g\overline{g} \,\mathrm{d}\mu\right| \\
&= \left|\int_X \mathrm{e}^{2\pi i n x} \,\mathrm{d}\mu\right| \\
&= \left|\int_0^1 \mathrm{e}^{2\pi i n x} \,\mathrm{d}x\right| \\
&= 0
\end{aligned}$$

and thus g is a weak mixing function.

Our goal is now to characterize weak mixing functions in terms of their spectral measure. We will make use of the following result, known as the *Wiener's lemma*.

Lemma 5.6. Let μ be a finite Borel measure on $\mathbb{T}$.
Then one has

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \left|\int_{\mathbb{T}} \mathrm{e}^{2\pi i n x} \,\mathrm{d}\mu(x)\right|^2 = \sum_{x\in\mathbb{T}} |\mu(\{x\})|^2.$$

Proof. Let μ' be the pushforward of μ under the map $x \longmapsto -x$, *i.e.* $\mu'(B) = \mu(-B)$ for all Borel sets B. Let also ν be the convolution of μ and μ', *i.e.*

$$\nu(B) = \mu * \mu'(B) = \int\int \mathbf{1}_B(x+y) \,\mathrm{d}\mu(x)\mathrm{d}\mu'(y)$$

for every Borel set B. Observe that $\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathrm{e}^{2\pi i n x} = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{otherwise} \end{cases}$, so the dominated convergence theorem implies

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_{\mathbb{T}} \mathrm{e}^{2\pi i n x} \,\mathrm{d}\nu(x) = \int_{\mathbb{T}} \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathrm{e}^{2\pi i n x} \,\mathrm{d}\nu(x) = \nu(\{0\}).$$

Now, on one hand, we have

$$\int_{\mathbb{T}} \mathrm{e}^{2\pi i n x} \,\mathrm{d}\nu(x) = \int\int \mathrm{e}^{2\pi i n(x+y)} \,\mathrm{d}\mu(x)\mathrm{d}\mu'(y)$$

$$
\begin{aligned}
&= \int \int e^{2\pi i n x} e^{2\pi i n y} \, d\mu(x) d\mu'(y) \\
&= \int \int e^{2\pi i n x} e^{-2\pi i n y} \, d\mu(x) d\mu(y) \\
&= \left| \int_{\mathbb{T}} e^{2\pi i n x} \, d\mu(x) \right|^2
\end{aligned}
$$

while on the other hand, it holds that

$$
\begin{aligned}
\nu(\{0\}) &= \int \int \mathbf{1}_{\{0\}}(x+y) \, d\mu(x) d\mu'(y) \\
&= \int \mu(\{-y\}) \, d\mu'(y) \\
&= \int \mu(\{y\}) \, d\mu(y) \\
&= \sum_{y \in \mathbb{T}} |\mu(\{y\})|^2
\end{aligned}
$$

and these two computations, together with previous limit, give the desired claim. □

We can now easily deduce the following. Recall before that a Borel measure μ is called *continuous* if it is non-atomic, *i.e.* it gives zero mass to singletons.

Proposition 5.7. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system. A function $f \in L^2(X, \mathcal{A}, \mu)$ is weak mixing if and only if μ_f is continuous.

Proof. By the previous results, we have

$$
\begin{aligned}
f \text{ is weak mixing} &\iff \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} |\langle U_T^n f, f \rangle| = 0 \\
&\iff \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} |\langle U_T^n f, f \rangle|^2 = 0 \\
&\iff \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \left| \int e^{2\pi i n x} \, d\mu_f(x) \right|^2 = 0 \\
&\iff \sum_{x \in \mathbb{T}} |\mu_f(\{x\})|^2 = 0 \\
&\iff \forall x \in \mathbb{T}, \ \mu_f(\{x\}) = 0
\end{aligned}
$$

where the first equivalence is Lemma 5.4, the second is a general fact from analysis, the third is the spectral theorem, and the fourth follows from Lemma 5.6. Thus we are done. □

5.3 The splitting $\mathcal{H}_{\mathrm{c}} \oplus \mathcal{H}_{\mathrm{wm}}$

For a measure-preserving system $(X, \mathcal{A}, \mu, T)$, let $\mathcal{H}_{\mathrm{c}}$ denote the closure of the subspace generated by the eigenfunctions, and denote $\mathcal{H}_{\mathrm{wm}}$ the subspace of weak mixing functions.

Theorem 5.8. We have $\mathcal{H}_{\mathrm{c}} \perp \mathcal{H}_{\mathrm{wm}}$ and $L^2(X, \mathcal{A}, \mu) = \mathcal{H}_{\mathrm{c}} \oplus \mathcal{H}_{\mathrm{wm}}$.

Proof. To start, fix $f \in \mathcal{H}_{\mathrm{wm}}$ and g an eigenfunction. Say that $U_T g = \lambda g$, $\lambda \in \mathbb{S}^1$. Then we compute

$$|\langle f, g\rangle| = \frac{1}{N}\sum_{n=1}^{N} |\langle U_T^n f, U_T^n g\rangle| = \frac{1}{N}\sum_{n=1}^{N} |\lambda^n||\langle U_T^n f, g\rangle| = \frac{1}{N}\sum_{n=1}^{N} |\langle U_T^n f, g\rangle|$$

and since this last quantity tends to 0 as $N \longrightarrow \infty$, we must have $|\langle f, g\rangle| = 0$. Hence $\langle f, g\rangle = 0$, and by linearity f is orthogonal to the subspace generated by eigenfunctions. Since the latter is dense in $\mathcal{H}_{\mathrm{c}}$, we deduce $\mathcal{H}_{\mathrm{wm}} \perp \mathcal{H}_{\mathrm{c}}$. To prove the second claim, it suffices to prove that $\mathcal{H}_{\mathrm{wm}} = \mathcal{H}_{\mathrm{c}}^{\perp}$, and since $\mathcal{H}_{\mathrm{wm}} \subset \mathcal{H}_{\mathrm{c}}^{\perp}$, it only remains to see the other inclusion. We will in fact prove that $f \notin \mathcal{H}_{\mathrm{wm}} \implies f \notin \mathcal{H}_{\mathrm{c}}^{\perp}$. Let then $f \notin \mathcal{H}_{\mathrm{wm}}$. By Proposition 5.7, μ_f is not continuous, so there exists $\alpha \in \mathbb{T}$ such that $\mu_f(\{\alpha\}) > 0$. Consider now the measure-preserving system $Y = \mathbb{T}$, $\mathcal{B}$ is the Borel $\sigma-$algebra, ν is the Haar measure, and $S(y) = y + \alpha$. Denote furthermore $g(y) = \mathrm{e}^{2\pi i y}$, and observe that $U_S g = \mathrm{e}^{2\pi i \alpha} g$. On the product system $(X \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu, T \times S)$, let $F = f \otimes \overline{g}$. By the Mean Ergodic Theorem, we have

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} U_{T\times S}^n F = F_{\mathrm{inv}}$$

in L^2-norm. Therefore, this convergence holds also weakly, and we can compute that

$$\begin{aligned}
\langle F_{\mathrm{inv}}, F\rangle &= \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \langle U_{T\times S}^n F, F\rangle \\
&= \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \langle (U_T^n f) \otimes (U_S^n \overline{g}), f \otimes \overline{g}\rangle \\
&= \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \langle U_T^n f, f\rangle \underbrace{\langle U_S^n \overline{g}, \overline{g}\rangle}_{=\mathrm{e}^{-2\pi i n\alpha}} \\
&= \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \mathrm{e}^{-2\pi i n\alpha} \langle U_T^n f, f\rangle
\end{aligned}$$

$$
\begin{aligned}
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathrm{e}^{-2\pi i n\alpha} \int \mathrm{e}^{2\pi i n x} \,\mathrm{d}\mu_f(x) \\
&= \int \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathrm{e}^{2\pi i n(x-\alpha)} \,\mathrm{d}\mu_f(x) \\
&= \mu_f(\{\alpha\}) > 0
\end{aligned}
$$

using the spectral theorem (Theorem 5.1) and the dominated convergence theorem. This shows that $F_{\mathrm{inv}} \neq 0$. Now observe that

$$U_{\mathrm{Id}\times S^{-1}} F = (U_{\mathrm{Id}} f) \otimes (U_{S^{-1}} \overline{g}) = f \otimes \mathrm{e}^{2\pi i\alpha} \overline{g} = \mathrm{e}^{2\pi i\alpha} F$$

and thus we also have

$$
\begin{aligned}
U_{\mathrm{Id}\times S^{-1}} F_{\mathrm{inv}} &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} U_{\mathrm{Id}\times S^{-1}} U_{T\times S}^n F \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} U_{T\times S}^n U_{\mathrm{Id}\times S^{-1}} F \\
&= \mathrm{e}^{2\pi i\alpha} \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} U_{T\times S}^n F \\
&= \mathrm{e}^{2\pi i\alpha} F_{\mathrm{inv}}.
\end{aligned}
$$

It then follows that

$$U_{T\times \mathrm{Id}} F_{\mathrm{inv}} = U_{\mathrm{Id}\times S^{-1}} U_{T\times S} F_{\mathrm{inv}} = U_{\mathrm{Id}\times S^{-1}} F_{\mathrm{inv}} = \mathrm{e}^{2\pi i\alpha} F_{\mathrm{inv}}$$

which means exactly that $F_{\mathrm{inv}}(Tx, y) = \mathrm{e}^{2\pi i\alpha} F_{\mathrm{inv}}(x,y)$. Let then $h_y(x) = F_{\mathrm{inv}}(x,y)$, which is therefore an eigenfunction for U_T with eigenvalue $\mathrm{e}^{2\pi i\alpha}$. Lastly, by Fubini's theorem one can write

$$
\begin{aligned}
0 < \langle F_{\mathrm{inv}}, F\rangle &= \int_{X\times Y} F_{\mathrm{inv}}(x,y)\overline{f}(x) g(y) \,\mathrm{d}(\mu\otimes\nu)(x,y) \\
&= \int_Y g(y) \left(\int_X h_y(x) \overline{f}(x) \,\mathrm{d}\mu(x) \right) \mathrm{d}\nu(y) \\
&= \int_Y g(y) \langle h_y, f\rangle \,\mathrm{d}\nu(y)
\end{aligned}
$$

so there must exist $y \in Y$ such that $h_y \neq 0$ and $\langle h_y, f\rangle > 0$. Thus f correlates with an eigenfunction of U_T, which means $f \notin \mathcal{H}_{\mathrm{c}}^{\perp}$, and finishes the proof. □

As promised, we now show having all eigenfunctions almost everywhere constant characterize weak mixing systems.

Theorem 5.9. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system.
It is weak mixing if and only all eigenfunctions are almost everywhere constant.

Proof. The direct implication has been seen in Corollary 4.12.

Conversely, suppose all eigenfunctions are constant almost everywhere. To show the system is weak mixing, it suffices to prove that every $f \in L^2(X, \mathcal{A}, \mu)$ with integral $\langle f, \mathbf{1}_X \rangle = 0$ is weak mixing. By Theorem 5.8, we write $f = f_{\mathrm{c}} + f_{\mathrm{wm}}$ with $f_c \in \mathcal{H}_{\mathrm{c}}$ and $f_{\mathrm{wm}} \in \mathcal{H}_{\mathrm{wm}}$. By hypothesis, f_{c} is almost everywhere constant, and we denote this constant by K. But then

$$0 = \int_X f \, \mathrm{d}\mu = \int_X f_{\mathrm{c}} \, \mathrm{d}\mu + \int_X f_{\mathrm{wm}} \, \mathrm{d}\mu = K$$

using the fact that a weak mixing function has integral equals to 0. Thus $f = f_{\mathrm{wm}}$ is weak mixing, as was to be shown. □

6. Modeling $\mathbb{N}$ through dynamical systems

In this part, we begin our preparation to the next section, and explain how integers relate deeply to ergodic theory. This relation is known as the *Furstenberg's correspondence principle*, and can therefore be used to prove results in number theory via ergodic theory.

6.1 The Bogolyubov-Krylov Theorem

Definition 6.1. Let X be a compact metric space, μ a Borel probability measure, and $T\colon X \longrightarrow X$ be continuous. A point $x \in X$ is called generic for μ if we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(T^n x) = \int_X f \, \mathrm{d}\mu$$

for all $f \in C(X)$. More generally, if $(I_k)_{k\in\mathbb{N}}$ is a sequence of intervals whose length tends to infinity, then $x \in X$ is generic for μ along $(I_k)_{k\in\mathbb{N}}$ if we have

$$\lim_{k\to\infty} \frac{1}{|I_k|} \sum_{n\in I_k} f(T^n x) = \int_X f \, \mathrm{d}\mu$$

for all $f \in C(X)$.

Remark 6.2. If x is generic for μ, then for all $f \in C(X)$, we have

$$\int_X f \, \mathrm{d}(T_*\mu) = \int_X f \circ T \, \mathrm{d}\mu = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f \circ T(T^n x) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(T^n x) = \int_X f \, \mathrm{d}\mu$$

so μ and $T_*\mu$ integrate continuous functions the same way, and this is enough to get that $T_*\mu = \mu$. The same argument applies for generic points along a sequence of intervals. Note that the second equality above crucially relies on the fact that T is continuous.

The existence of generic points is in fact always guaranteed in a compact space.

Proposition 6.3. Let X be a compact metric space, $T\colon X \longrightarrow X$ continuous. For any $x \in X$ and any sequence $(I_k)_{k\in\mathbb{N}}$, there is a subsequence $(I'_k)_{k\in\mathbb{N}}$ and a Borel probability measure μ such that x is generic for μ along $(I'_k)_{k\in\mathbb{N}}$.

Proof. Let $\mathcal{M}(X)$ denote the set of Borel probability measures on X. For $k \in \mathbb{N}$, let $\mu_k := \frac{1}{|I_k|}\sum_{n\in I_k} \delta_{T^n x} \in \mathcal{M}(X)$. Since $\mathcal{M}(X)$ is compact, there exists $\mu \in \mathcal{M}(X)$ and a subsequence $(\mu_{k_j})_{j\in\mathbb{N}}$ of $(\mu_k)_{k\in\mathbb{N}}$ converging to μ. Set then $I'_j := I_{k_j}$ for $j \in \mathbb{N}$. □

This yields to the Bogolyobov-Krylov theorem.

Theorem 6.4. Let X be a compact metric space.
For any continuous map $T\colon X \longrightarrow X$, there exists a Borel probability measure μ preserved by T.

Proof. Let $x \in X$ be an arbitrary point. By Proposition 6.3, we can find μ a Borel probability measure and a sequence of intervals whose length tends to infinity such that x is generic for μ along this sequence. Then T must preserves μ by Remark 6.2, which we may use since T is continuous. □

In the sequel, we will also make use of the following proposition.

Proposition 6.5. Let X be a compact metric space, $\mathcal{A}$ its Borel $\sigma-$algebra, μ a Borel probability measure, and $T\colon X \longrightarrow X$ a continuous measure-preserving transformation. If $(X, \mathcal{A}, \mu, T)$ is ergodic then every sequence of intervals $(I_k)_{k\in\mathbb{N}}$ has a subsequence $(I'_k)_{k\in\mathbb{N}}$ such that $\mu-$almost every point $x \in X$ is generic for μ along $(I'_k)_{k\in\mathbb{N}}$.

6.2 Furstenberg's correspondence principle

Recall that the upper density of a set $A \subset \mathbb{N}$ is defined as

$$\overline{d}(A) := \limsup_{N\to\infty} \frac{|A \cap \{1,\dots,N\}|}{N}.$$

For instance, $\overline{d}(2\mathbb{N}) = \overline{d}(2\mathbb{N}+1) = \frac{1}{2}$, while any finite set has upper density 0. The set of perfect squares has upper density $\frac{6}{\pi^2}$, while the subset of prime numbers has density 0.

Here the bridge between number and ergodic theory. It is called the Furstenberg's correspondence principle.

Theorem 6.6. For any $A \subset \mathbb{N}$, there exists a compact metric space X, a continuous map $T\colon X \longrightarrow X$, a Borel probability measure preserved by T, a point $x \in X$ generic for μ along a sequence $(I_k)_{k\in\mathbb{N}}$ and a clopen set $E \subset X$ so that $\mu(E) = \overline{d}(A)$ and $A = \{n \in \mathbb{N} : T^n x \in E\}$.

Proof. Suppose $A \subset \mathbb{N}$ is given. Take $X = \{0,1\}^{\mathbb{N}\cup\{0\}}$, $T\colon X \longrightarrow X$ the left shift, and $x = \mathbf{1}_A$. Let $(N_k)_{k\in\mathbb{N}}$ be such that

$$\overline{d}(A) = \lim_{k\to\infty} \frac{|A \cap \{1,\dots,N_k\}|}{N_k},$$

which is possible since the limsup of a sequence of real numbers is always an accumulation point of this sequence. Choose $I_k := \{1,\dots,N_k\}$. Finally, let

$$E := \{w \in X : w_0 = 1\}.$$

We now show all claims of the theorem holds. First, note that $\{0,1\}$ is compact metric, so that X is compact metric for the product topology, by Tychonoff's theorem. Secondly, T is continuous. We can now apply Proposition 6.3 and we find a subsequence $(I'_k)_{k\in\mathbb{N}}$ and a Borel probability measure μ such that x is generic for μ along $(I'_k)_{k\in\mathbb{N}}$. Then, by Remark 6.2, μ is preserved by T. Next, note that E is open in X, since we can write

$$E = \{0\} \times \prod_{n\geq 1}\{0,1\}$$

and $\{0\}$ is open in $\{0,1\}$. Similarly, $X \setminus E$ is open, so E is also closed in X. We have also that

$$\mathbf{1}_E(T^n x) = \mathbf{1}_E(T^n \mathbf{1}_A) = \mathbf{1}_E(\mathbf{1}_{A-n}) = \mathbf{1}_{A-n}(0) = \mathbf{1}_A(n)$$

proving that $A = \{n \in \mathbb{N} : T^n x \in A\}$. To conclude, we compute that

$$\mu(E) = \int_X \mathbf{1}_E \,\mathrm{d}\mu = \lim_{k\to\infty} \frac{1}{|I'_k|} \sum_{n\in I'_k} \mathbf{1}_E(T^n x) = \lim_{k\to\infty} \frac{1}{|I'_k|} \sum_{n\in I'_k} \mathbf{1}_A(n) = \lim_{k\to\infty} \frac{|A \cap I'_k|}{|I'_k|} = \overline{d}(A)$$

where the second equality follows from the fact that x is generic for μ along $(I'_k)_{k\in\mathbb{N}}$, and the fact that $\mathbf{1}_E$ is indeed continuous since E is clopen. This concludes the proof. □

We can in fact make two additional assumptions on the space X used to model the subset $A \subset \mathbb{N}$. These requires the following terminology.

Definition 6.7. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system, with X compact metric and $T\colon X \longrightarrow X$ continuous.
It has continuous eigenfunctions if every eigenfunction has a continuous representative in its $L^2(X, \mathcal{A}, \mu)$ equivalence class.

We will take for granted the following fact.

Lemma 6.8. Without restrictions, we can assume that the system stemming from Furstenberg's correspondence principle is ergodic and has continuous eigenfunctions.

7. Applications to number theory

In this section, we show how Furstenberg's correspondence principle can be applied to solve problems from number theory. The main part of this chapter is devoted to the proof of the so called *Erdös sumset conjecture*.

7.1 First applications

The first lemma we will need for this section is the following. We place ourselves in the context of Furstenberg's correspondence principle, *i.e.* Theorem 6.6.

Lemma 7.1. Let $A \subset \mathbb{N}$.
For any $n_1, \dots, n_k \in \mathbb{N}$, it holds that

$$\mu\left(\bigcap_{i=1}^{k} T^{-n_i} E\right) \leq \overline{d}\left(\bigcap_{i=1}^{k} (A - n_i)\right).$$

Proof. From the proof of Theorem 6.6, we have $X = \{0,1\}^{\mathbb{N}\cup\{0\}}$, $T : X \longrightarrow X$ the left shift, $x = \mathbf{1}_A$, $E = \{w \in X \mid w_0 = 1\}$, so that $A = \{n \in \mathbb{N} \mid T^n x \in E\}$. We also find a Borel probability measure μ such that x is generic for μ along a sequence $(I_k)_{k\in\mathbb{N}}$, and we have proved that

$$\mu(E) = \lim_{k\to\infty} \frac{1}{|I_k|} \sum_{n\in I_k} \mathbf{1}_E(T^n x) = \overline{d}(A).$$

The map T being continuous, and $E \subset X$ being clopen, also $T^{-n_1}E, \dots, T^{-n_k}E \subset X$ are clopen, so the intersection $\bigcap_{i=1}^{k} T^{-n_i} E$ is clopen, hence its indicator function is continuous, and since x is generic for μ along $(I_k)_{k\in\mathbb{N}}$ we get

$$\mu\left(\bigcap_{i=1}^{k} T^{-n_i} E\right) = \int_X \mathbf{1}_{\bigcap_{i=1}^{k} T^{-n_i} E} \, \mathrm{d}\mu = \lim_{k\to\infty} \frac{1}{|I_k|} \sum_{n\in I_k} \mathbf{1}_{\bigcap_{i=1}^{k} T^{-n_i} E}(T^n x).$$

Now note that $T^n x \in \bigcap_{i=1}^{k} T^{-n_i} E$ if and only if $T^{n+n_i} x \in E$ for all $i = 1, \dots, k$, which is equivalent to $n + n_i \in A$ for all $i = 1, \dots, k$, or also $n \in A - n_i$ for all $i = 1, \dots, k$, *i.e.*

$n \in \bigcap_{i=1}^{k}(A - n_i)$. Thus

$$\begin{aligned}
\mu\left(\bigcap_{i=1}^{k} T^{-n_i}E\right) &= \lim_{k\to\infty} \frac{1}{|I_k|} \sum_{n\in I_k} \mathbf{1}_{\bigcap_{i=1}^{k} T^{-n_i}E}(T^n x) \\
&= \lim_{k\to\infty} \frac{|\bigcap_{i=1}^{k}(A - n_i) \cap \{1,\dots,N_k\}|}{|\{1,\dots,N_k\}|} \\
&\le \limsup_{N\to\infty} \frac{|\bigcap_{i=1}^{k}(A - n_i) \cap \{1,\dots,N\}|}{|\{1,\dots,N\}|} \\
&= \overline{d}\left(\bigcap_{i=1}^{k}(A - n_i)\right)
\end{aligned}$$

and this proves the claim. □

We now give two applications of these results. Remember that a subset of $\mathbb{N}$ containing an infinite syndetic set is itself syndetic.

Proposition 7.2. Let $A \subset \mathbb{N}$ be a set of positive density.
Then $A - A := \{a - b \mid a, b \in A\}$ is syndetic.

Proof. Let's first apply Furstenberg's correspondence principle, to find a compact metric space X, a continuous transformation $T\colon X \longrightarrow X$, a clopen set $E \subset X$ such that $A = \{n \in \mathbb{N} \mid T^n x \in E\}$ and $\mu(E) = \overline{d}(A)$. By a weak version of Kintchine's theorem (Theorem 2.12), the set

$$R := \{n \in \mathbb{N} : \mu(E \cap T^{-n}E) > 0\}$$

is syndetic. Now if $S := \{n \in \mathbb{N} \mid \overline{d}(A \cap (A - n)) > 0\}$, then it follows from Lemma 7.1 that $R \subset S$, and thus S is also syndetic. Lastly, note that if $n \in S$, then $A \cap (A - n) \neq \emptyset$, and this implies $n \in A - A$. Hence $A - A$ contains S which is syndetic, so $A - A$ is syndetic. □

For the next result, we require a terminology.

Definition 7.3. A set $A \subset \mathbb{N}$ is called intersective if for any $B \subset \mathbb{N}$ of positive density, we have $A \cap (B - B) \neq \emptyset$.

The following proposition then gives us a bunch of intersective sets.

Proposition 7.4. A set of recurrence $R \subset \mathbb{N}$ is intersective.

Proof. Let $R \subset \mathbb{N}$ be a set of recurrence, and $A \subset \mathbb{N}$ having positive density. By Theorem 6.6, we find a measure-preserving system $(X, \mathcal{A}, \mu, T)$ and a clopen set $E \subset X$ such that $\mu(E) = \overline{d}(A)$ and $A = \{n \in \mathbb{N} : T^n x \in E\}$. In particular, $\mu(E) > 0$ since A has positive density, and R being a set of recurrence now implies there is $n \in R$ such that

$$\mu(E \cap T^{-n} E) > 0.$$

By Lemma 7.1, it follows that $\overline{d}(A \cap (A - n)) > 0$, which in turn implies $n \in A - A$. Hence $n \in R \cap (A - A)$, which is therefore not empty, and we proved R is intersective. □

7.2 The Szemerédi's theorem

Szemerédi's theorem is a qualitative result about arithmetic progressions. In order to prove it, we admit the next result, which can be viewed as a wide generalization of Poincaré's recurrence theorem. This is called the *Furstenberg's multiple recurrence theorem*.

Theorem 7.5. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system, and $A \in \mathcal{A}$ with $\mu(A) > 0$. For any $k \in \mathbb{N}$, there exists $n \in \mathbb{N}$ such that

$$\mu(A \cap T^{-n} A \cap T^{-2n} A \cap \cdots \cap T^{-kn} A) > 0.$$

Here is the goal of this subsection.

Theorem 7.6. Let $A \subset \mathbb{N}$ be a set of positive density.
Then A contains arbitrary long arithmetic progressions.

Proof. By the Furstenberg's correspondence principle, we find a measure-preserving system $(X, \mathcal{A}, \mu, T)$, a point $x \in X$ and a clopen set $E \subset X$ such that $\mu(E) = \overline{d}(A)$ and $A = \{n \in \mathbb{N} : T^n x \in E\}$. Now fix $k \in \mathbb{N}$. We will prove A contains an arithmetic progression of length k. Since $\mu(E) = \overline{d}(A) > 0$, Theorem 7.5 guarantees the existence of $n \in \mathbb{N}$ such that

$$\mu(E \cap T^{-n} E \cap \cdots \cap T^{-kn} E) > 0.$$

It then follows from Lemma 7.1 that

$$\overline{d}(A \cap (A - n) \cap \cdots \cap (A - kn)) > 0$$

and, in particular, $A \cap (A-n) \cap \cdots \cap (A-kn) \neq \emptyset$. Let then $m \in \bigcap_{j=0}^{k}(A-jn)$. Equivalently, $m, m+n, m+2n, \dots, m+kn \in A$, which shows A contains an arithmetic progression of length k, as wanted. □

7.3 The Erdös sumset conjecture

Let's now turn ourselves to the main result of this section, the Erdös sumset conjecture. It claims that any set $A \subset \mathbb{N}$ of positive density contains the sum of two infinite sets. Initially asked by Paul Erdös in the 70's, it was proved in 2019 by Joel Moreira, Florian Richter and Donald Robertson.

The first step towards the proof is to reformulate what it means for a set A to contain the sum of two infinite sets.

Lemma 7.7. Let $A \subset \mathbb{N}$.
There exists two infinite sets $B, C \subset \mathbb{N}$ such that $B + C \subset A$ if and only if there exists an increasing sequence $s_1 < s_2 < \dots$ of integers and $L \subset \mathbb{N}$ with

$$\mathbf{1}_L(n) = \lim_{i\to\infty} \mathbf{1}_A(n+s_i)$$

and such that the family $(L \cap (A - s_i))_{i\in\mathbb{N}}$ has the large intersection property, *i.e.* any finite subfamily intersects in an infinite set.

Proof. To start, suppose $B + C \subset A$, and write $B = (b_i)_{i\in\mathbb{N}}$, $C = (c_j)_{j\in\mathbb{N}}$. By a diagonalization argument, there is a subsequence $(s_i)_{i\in\mathbb{N}}$ of $(b_i)_{i\in\mathbb{N}}$ such that $\lim_{i\to\infty} \mathbf{1}_A(n+s_i)$ exists for all $n \in \mathbb{N}$. Define $L \subset \mathbb{N}$ so that

$$\mathbf{1}_L(n) = \lim_{i\to\infty} \mathbf{1}_A(n+s_i).$$

Observe then that $C \subset L$, and since $C \subset A - s_i$ for all $i \in \mathbb{N}$, it follows that $C \subset L \cap (A - s_i)$ for all $i \in \mathbb{N}$. Hence $(L \cap (A - s_i))_{i\in\mathbb{N}}$ has the large intersection property.

Conversely, we are going to construct A and B inductively. To start, take $b_1 = s_1$ and c_1 any element in $L \cap (A - s_1) = L \cap (A - b_1)$. Suppose that $c_1, \dots, c_n \in L$, $b_1, \dots, b_n \in \{s_1, s_2, \dots\}$ have been found, and that they satisfy $c_j + b_i \in A$ for $i, j = 1, \dots n$. Let c_{n+1} be any element of

$$\bigcap_{i=1}^{n}(L \cap (A - b_i))$$

with $c_{n+1} > c_n$. Such an element exists because the above intersection is infinite by hypothesis. In particular, $c_{n+1} + b_i \in A$ for all $i = 1, \dots, n$. Now since $c_1, \dots, c_{n+1} \in L$, we obtain

$$\lim_{i \to \infty} \mathbf{1}_A(c_j + s_i) = \mathbf{1}_L(c_j) = 1$$

for all $j = 1, \dots, n+1$. This implies that $\mathbf{1}_A(c_j + s_i)$ becomes constant equal to 1 for i sufficiently large, *i.e.* $c_j + s_i \in A$ for i large enough. We can therefore find $b_{n+1} \in \{s_1, s_2, \dots\}$ so that $b_{n+1} > b_n$ and $c_j + b_{n+1} \in A$ for all $j = 1, \dots, n+1$. This achieves the inductive step, and our proof. □

Here are two remarks of importance for the proof of the main theorem.

Remark 7.8. (i) If $f = f_c + f_{wm}$ is the Jacobs-de Leeuw-Glicksberg decomposition of f, and if $f \geq 0$, then $f_c \geq 0$. In particular the orthogonal projection onto the subspace $\mathcal{H}_c$ preserves the order.

(ii) If f is weak mixing then $\displaystyle\lim_{k\to\infty} \frac{1}{|I_k|} \sum_{n \in I_k} |\langle U_T^n f, g\rangle| = 0$ for any increasing sequence of intervals $(I_k)_{k\in\mathbb{N}}$ and any $g \in L^2(X, \mathcal{A}, \mu)$. Indeed, Wiener's lemma still holds when the average is taken over a sequence $(I_k)_{k\in\mathbb{N}}$, so the limit equals 0 if and only if μ_f is continuous, which holds since f is weak mixing.

Theorem 7.9. Let X be a compact metric space, $\mathcal{A}$ the Borel $\sigma-$algebra, μ a Borel probability measure, and $T \colon X \longrightarrow X$ a continuous measure-preserving transformation. Let $x \in X$ be generic for μ along a sequence of intervals $(I_k)_{k\in\mathbb{N}}$. Suppose that $(X, \mathcal{A}, \mu, T)$ is ergodic and has continuous eigenfunctions. Then for any clopen set $E \subset X$ with $\mu(E) > 0$, there exists $y \in X$, an increasing sequence of integers $s_1 < s_2 < \dots$ and a Borel probability measure λ on $X \times X$ such that :

(i) The sequence $(T^{s_i}x)_{i\in\mathbb{N}}$ converges to y as $i \to \infty$.

(ii) The point (x, y) is generic for λ along a subsequence of $(I_k)_{k\in\mathbb{N}}$.

(iii) For any $k \in \mathbb{N}$, $\lambda((T^{-s_1}E \cap \cdots \cap T^{-s_k}E) \times E) > 0$.

Proof. To come. □

From there, Erdös sumset conjecture follows straightforwardly.

Theorem 7.10. Let $A \subset \mathbb{N}$ be a set of positive density.
Then A contains the sum of two infinite sets $B, C \subset \mathbb{N}$.

Proof. Let $A \subset \mathbb{N}$ be such that $\overline{d}(A) > 0$. By the Furstenberg's correspondence principle, namely Theorem 6.6, we find a compact metric space X, a Borel probability measure μ, a continuous transformation $T\colon X \longrightarrow X$ preserving μ, a point $x \in X$ generic for μ along a sequence $(I_k)_{k\in\mathbb{N}}$, such that $\mu(E) = \overline{d}(A)$ and $A = \{n \in \mathbb{N} : T^n x \in E\}$. Moreover, by Lemma 6.8, we can assume $(X, \mathcal{A}, \mu, T)$ is ergodic and has continuous eigenfunctions. Henceforth, we may apply Theorem 7.9, and we have a point $y \in X$, a sequence $s_1 < s_2 < \ldots$ of integers and a Borel probability measure λ on $X \times X$ such that the three points above hold. Let now $L := \{n \in \mathbb{N} : T^n y \in E\}$. One has then

$$\mathbf{1}_L(n) = \mathbf{1}_E(T^n y) = \mathbf{1}_E(\lim_{i\to\infty} T^{n+s_i}x) = \lim_{i\to\infty} \mathbf{1}_E(T^{n+s_i}x) = \lim_{i\to\infty} \mathbf{1}_A(n+s_i)$$

for all $n \in \mathbb{N}$, since T and $\mathbf{1}_E$ are continuous. Furthermore, note that

$$\begin{aligned} L \cap \bigcap_{i=1}^{k}(A - s_i) &= \{n \in \mathbb{N} : T^n y \in E\} \cap \{n \in \mathbb{N} : T^{n+s_1}x \in E, \ldots, T^{n+s_k}x \in E\} \\ &= \{n \in \mathbb{N} : (T \times T)^n(x,y) \in (T^{-s_1}E \cap \cdots \cap T^{-s_k}E) \times E\} \end{aligned}$$

and by combining (ii) and (iii) of Theorem 7.9, we see that the latter is infinite. Thus the family $(L \cap (A - s_i))_{i\in\mathbb{N}}$ has the large intersection property, and by Lemma 7.7, we conclude that A contains the sum $B + C$ of two infinite sets. □

8. Entropy

Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system. Knowing the position of a point $x \in X$ and $Tx, T^2x, \dots, T^{n-1}x$, how accurate are the predictions that one can make about the approximate position of $T^n x$? The answer depends crucially on the transformation T. If the transformation is in nature deterministic, the past trajectory determines the imminent future and so informations about $x, Tx, \dots, T^{n-1}x$ leads to probable predictions about $T^n x$. On the other hand, if T is very chaotic, the past orbit has little or no influence on the future, and does not allow one to make reasonable conjectures on T^n. The purpose of this section is to make these ideas rigorous and mathematically precise.

8.1 Entropy of a partition

A *partition* of a probability space $(X, \mathcal{A}, \mu)$ is a finite of countable infinite collection of pairwise disjoint measurable subsets of X whose union covers X. We will denote such a partition by $\xi = \{A_1, \dots, A_r\}$ or $\xi = \{A_1, A_2, \dots\}$. From the probabilistic point a view, a partition is a discrete random variable, taking values either in a finite set $\{1, \dots, r\}$ or in $\mathbb{N}$.

For a partition ξ, we denote $\sigma(\xi)$ the $\sigma-$algebra generated by ξ. The sets $A_1, A_2, \dots$ of the partition are called its *atoms*. For $x \in X$, we denote by $[x]_\xi$ the element of the partition that contains x. A partition η is called a *refinment* of ξ if each atom of ξ is a union of atoms of η. We then write $\xi \le \eta$. The *common refinment* of two partitions ξ and η is the *joint partition* $\xi \vee \eta$, which consists of all sets of the form $A \cap B$, $A \in \xi, B \in \eta$.

A partition $\xi = \{A_1, A_2, \dots\}$ of a probability space may be thought as giving the possible outcomes $1, 2, \dots$ of an experiment, with the probability of outcome i being $\mu(A_i)$. We can therefore associate to ξ a number $H(\xi)$ describing the amount of uncertainty about the outcome of the experience.

Definition 8.1. Let $(X, \mathcal{A}, \mu)$ be a probability space, and $\xi = \{A_1, A_2, \dots\}$ be a partition of X. The entropy of ξ, denoted $H(\xi)$, defined by

$$H(\xi) = -\sum_{i \ge 1} \mu(A_i) \log_2(\mu(A_i))$$

with the convention that $0 \log 0 = 0$.

For two measurable subsets $A, B \in \mathcal{A}$, the conditional measure of A with respect to B is

$$\mu(A|B) = \frac{\mu(A \cap B)}{\mu(B)}$$

and describes the probability of the event A happening assuming that B has occured. This idea readily transposes to partitions.

Definition 8.2. Let $(X, \mathcal{A}, \mu)$ be a probability space, with two partitions $\xi = \{A_1, A_2, \dots\}$ and $\eta = \{B_1, B_2, \dots\}$. The conditional entropy of ξ given η, denoted $H(\xi|\eta)$, is the number

$$H(\xi|\eta) = \sum_{j\geq 1} \mu(B_j)\Big(-\sum_{i\geq 1} \mu(A_i|B_j)\log_2(\mu(A_i|B_j))\Big).$$

This formula should be viewed as a weighted average of entropies of the partition ξ conditioned on the individual atoms of η.

As in probability theory, we will say ξ and η are *independent* if $\mu(A\cap B) = \mu(A)\mu(B)$ for all $A \in \xi, B \in \eta$.

Let us establish the first basic properties of entropy, that will be useful in the sequel.

Proposition 8.3. Let $(X, \mathcal{A}, \mu)$ be a probability space, with three partitions ξ, η and γ. Then the following holds.

(i) $H(\xi) \geq 0$, and $H(\xi) = 0$ if and only if $\mu(A) = 1$ for some atom $A \in \xi$.

(ii) If $|\xi| = r$, $H(\xi) \leq \log(r)$ and $H(\xi) = \log(r)$ if and only if $\mu(A_i) = \frac{1}{r}$ for all $A_i \in \xi$.

(iii) $H(\xi \vee \eta) = H(\eta) + H(\xi|\eta)$.

(iv) $H(\xi) \geq H(\xi|\eta)$, and moreover $H(\xi|\eta) \geq H(\xi|\eta \vee \gamma)$.

(v) $H(\xi|\{X\}) = H(\xi)$ and $H(\xi|\xi) = 0$.

(vi) ξ and η are independent if and only if $H(\xi \vee \eta) = H(\xi) + H(\eta)$ if and only if $H(\xi|\eta) = H(\xi)$.

(vii) If $T\colon X \longrightarrow X$ is measure-preserving, then $H(T^{-1}\xi) = H(\xi)$ and $H(T^{-1}\xi|T^{-1}\eta) = H(\xi|\eta)$.

Proof. (i) The function $x \longmapsto x\log(x)$ is negative on $[0, 1]$, so $\mu(A_i)\log_2(\mu(A_i) \leq 0$ for all $i \geq 1$. Hence $H(\xi) \geq 0$. If $\mu(A) = 1$ for some $A \in \xi$, then all other atoms of ξ has measure 0, yielding $H(\xi) = -\mu(A)\log_2(\mu(A)) = 0$. Conversely, if $\mu(A) < 1$ for all $A \in \xi$, then all terms in $H(\xi)$ give a non-trivial contribution, so $H(\xi) > 0$.

(ii) If $\varphi(x) = x\log(x)$, then one has

$$\frac{1}{r}\log\left(\frac{1}{r}\right) = \varphi\left(\frac{1}{r}\right) = \varphi\left(\sum_{i=1}^{r}\mu(A_i)\frac{1}{r}\right) \leq \sum_{i=1}^{r}\frac{1}{r}\varphi(\mu(A_i)) = \frac{1}{r}\sum_{i=1}^{r}\mu(A_i)\log(\mu(A_i))$$

by Jensen's inequality which applies since φ is convex. Thus $\log(r) \geq -\sum_{i=1}^{r}\mu(A_i)\log(\mu(A_i)) = H(\xi)$. Moreover Jensen's inequality is an equality if and only if $\mu(A_i) = \frac{1}{r}$ for all $A_i \in \xi$.

(iii) By definition, we have

$$\begin{aligned}
H(\xi|\eta) + H(\eta) &= \sum_{j\geq 1}\mu(B_j)\left(-\sum_{i\geq 1}\mu(A_i|B_j)\log_2(\mu(A_i|B_j))\right) - \sum_{j\geq 1}\mu(B_j)\log_2(\mu(B_j)) \\
&= \sum_{j\geq 1}\mu(B_j)\left(-\log_2(\mu(B_j)) - \sum_{i\geq 1}\mu(A_i|B_j)\log_2(\mu(A_i|B_j))\right) \\
&= -\sum_{j\geq 1}\left(\mu(B_j)\log_2(\mu(B_j)) + \sum_{i\geq 1}\mu(B_j)\mu(A_i|B_j)\log_2(\mu(A_i|B_j))\right)
\end{aligned}$$

and note that $\mu(B_j) = \sum_{i\geq 1}\mu(B_j|A_i)\mu(A_i) = \sum_{i\geq 1}\mu(A_i\cap B_j)$, for all $j \geq 1$. It thus follows that

$$\begin{aligned}
H(\xi|\eta) + H(\eta) &= -\sum_{j\geq 1}\left(\sum_{i\geq 1}\mu(A_i\cap B_j)\log_2(\mu(B_j)) + \sum_{i\geq 1}\mu(A_i\cap B_j)\log_2(\mu(A_i|B_j))\right) \\
&= -\sum_{j\geq 1}\sum_{i\geq 1}\mu(A_i\cap B_j)\log_2(\mu(A_i\cap B_j)) \\
&= H(\xi\vee\eta)
\end{aligned}$$

as claimed.

(iv) To come.

(v) The first point follows from $\mu(X) = 1$ and the second from $\mu(A_i|A_j) = 0$ if $i \neq j$.

(vi) We show that ξ and η are independent if and only if $H(\xi|\eta) = H(\xi)$.
$\Longrightarrow$: Suppose ξ and η are independent. Then one computes that

$$H(\xi|\eta) = \sum_{j\geq 1}\mu(B_j)\left(-\sum_{i\geq 1}\mu(A_i|B_j)\log_2(\mu(A_i|B_j))\right) = H(\xi)\sum_{j\geq 1}\mu(B_j) = H(\xi)$$

as claimed.

(vii) Obvious. $\square$

Let's now go back to ergodic theory.

Definition 8.4. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system. Its entropy with respect to a partition ξ is the number

$$h_\mu(T, \xi) = \lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right).$$

The entropy of the system is then $h_\mu(T) = \sup_\xi h_\mu(T, \xi)$.

To see $h_\mu(T, \xi)$ is well-defined, we appeal Fekete's lemma, which assures that if $(a_n)_{n\in\mathbb{N}}$ is a subadditive sequence, then $\lim_{n\to\infty} \frac{a_n}{n}$ exists and equals $\inf_{n\in\mathbb{N}} \frac{a_n}{n}$. In our case we denote then

$$a_n = H\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right)$$

for $n \in \mathbb{N}$. A combination of (iii) and (iv) in Proposition 8.3 shows that $H(\xi \vee \eta) \leq H(\xi) + H(\eta)$, so we have

$$\begin{aligned} a_{n+m} = H\left(\bigvee_{i=0}^{n+m-1} T^{-i}\xi\right) &\leq H\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right) + H\left(\bigvee_{i=n}^{n+m-1} T^{-i}\xi\right) \\ &= H\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right) + H\left(\bigvee_{i=0}^{m-1} T^{-i}\xi\right) \\ &= a_n + a_m \end{aligned}$$

where the third equality follows from Proposition 8.3(vii). Hence $h_\mu(T, \xi)$ is well-defined by Fekete's lemma.

We now prove a more convenient formula to use for computations.

Theorem 8.5. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system, and ξ be a partition of X. Then it holds that

$$h_\mu(T, \xi) = \lim_{n\to\infty} H\left(T^{-n}\xi \,\middle|\, \bigvee_{i=0}^{n-1} T^{-i}\xi\right).$$

Proof. The first thing to check is that the limit appearing in the statement indeed exists. Introduce then $w(n) := H\left(T^{-n}\xi \,\middle|\, \bigvee_{i=0}^{n-1} T^{-i}\xi\right)$. By points (iii) and (vii) of Proposition

8.3, one gets

$$\begin{aligned} w(n+1) &= H\left(T^{-(n+1)}\xi \,\middle|\, \bigvee_{i=0}^{n} T^{-i}\xi\right) \\ &\leq H\left(T^{-(n+1)}\xi \,\middle|\, \bigvee_{i=1}^{n} T^{-i}\xi\right) \\ &= H\left(T^{-n}\xi \,\middle|\, \bigvee_{i=0}^{n-1} T^{-i}\xi\right) \\ &= w(n) \end{aligned}$$

so $(w(n))_{n\in\mathbb{N}}$ is decreasing and bounded below, hence its limit exists, as was to be shown. Now we observe that

$$\begin{aligned} H\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right) &= H\left(T^{-(n-1)}\xi \vee \bigvee_{i=0}^{n-2} T^{-i}\xi\right) \\ &= H\left(\bigvee_{i=0}^{n-2} T^{-i}\xi\right) + H\left(T^{-(n-1)}\xi \,\middle|\, \bigvee_{i=0}^{n-2} T^{-i}\xi\right) \\ &= H\left(\bigvee_{i=0}^{n-2} T^{-i}\xi\right) + w(n-1) \\ &= \cdots = w(0) + \cdots + w(n-1) \end{aligned}$$

so we can conclude that $h_\mu(T,\xi) = \lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right) = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} w(k) = \lim_{n\to\infty} w(n)$ by Cesaro's theorem. □

We can therefore use this result to obtain a first explicit value for the entropy of a measure-preserving system.

Example 8.6. Consider a Bernoulli shift, *i.e.* $X = \{0,1\}^{\mathbb{N}\cup\{0\}}$, T is the left-shift, and $\mu = \nu^{\mathbb{N}\cup\{0\}}$ where $\nu(\{0\}) = \nu(\{1\}) = \frac{1}{2}$. Let ξ be the partition of X into two pieces $[0]_0$, $[1]_0$. The first one consists of all sequences beginning with 0 and the second one contains all sequences beginning by 1. Then $T^{-n}\xi = \{[0]_n, [1]_n\}$, and this partition is independent of $\bigvee_{i=0}^{n-1} T^{-i}\xi$, so Theorem 8.5 yields

$$h_\mu(T,\xi) = \lim_{n\to\infty} H(T^{-n}\xi) = H(\xi) = \log(2)$$

using Proposition 8.3(ii) for the last equality.

Note that in the previous example we just computed the entropy with respect to a particular partition, and we should do a similar computation for every countable partition of the space to get the entropy $h_\mu(T)$. The next theorem, that we will take for granted, tells us in fact it is not necessary if we are in a good situation.

Theorem 8.7. Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving system.
If ξ generates $\mathcal{A}$, meaning that $\sigma\left(\bigvee_{i=0}^{\infty} T^{-i}\xi\right) = \mathcal{A}$, then $h_\mu(T) = h_\mu(T, \xi)$.

Lastly, let us investigate how entropy behaves with respect to factors and extensions.

Proposition 8.8. Let $(Y, \mathcal{B}, \nu, S)$ be a factor of $(X, \mathcal{A}, \mu, T)$.
It holds that $h_\nu(S) \leq h_\mu(T)$.

Proof. Let ξ be a partition of Y, and denote $\pi\colon X \longrightarrow Y$ the factor map. Then $\pi^{-1}(\xi)$ is a partition of X, so $h_\mu(T, \pi^{-1}(\xi)) \leq h_\mu(T)$. On the other hand, one can check directly that

$$h_\mu(T, \pi^{-1}(\xi)) = h_\nu(S, \xi)$$

and it follows that $h_\nu(S) \leq h_\mu(T)$. □

The direct consequence of this proposition is that entropy is an invariant of measure-preserving system.

Corollary 8.9. If $(X, \mathcal{A}, \mu, T)$ and $(Y, \mathcal{B}, \nu, S)$ are isomorphic, then $h_\mu(T) = h_\nu(S)$.

Proof. If X and Y are isomorphic, there is two factor maps $\pi\colon X \longrightarrow Y$ and $\varphi\colon Y \longrightarrow X$, so Proposition 8.8 implies $h_\nu(S) \leq h_\mu(T)$ and $h_\mu(T) \leq h_\nu(S)$, giving the claim. □