Emanuele Pardini

Elements of Ergodic Theory

Contents

1

Invariant Measures and Recurrence

1.1 Invariant measures

In the rest of this chapter $(M, \mathcal{B}, \mu)$ will be a generic measure space and $T : M \to M$ a measurable transformation.

Definition 1.1.1: We say that T is *non-singular w.r.t.* μ or that μ is *non-singular w.r.t.* T if $\mu(B) = 0 \iff \mu(T^{-1}(B)) = 0$ for every $B \in \mathcal{B}$.

Definition 1.1.2: We say that μ *is invariant under* T (μ is T*-invariant*) or that T *preserves* μ, if for all $E \in \mathcal{B}$

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

Remark 1.1.3: If T preserves μ, then T is non-singular w.r.t. μ.

It is possible and useful to extend the definition to flows.

Definition 1.1.4: A *flow* on $(X, \mathcal{B})$ is a family of measurable transformations $T^t : M \to M$, $t \in \mathbb{R}$, s.t.

$$T^0 = \text{id} \quad , \quad T^{s+t} = T^s \circ T^t \quad \forall s, t \in \mathbb{R}$$

Remark 1.1.5: If $\{T^t\}_{t\in\mathbb{R}}$ is a flow, then each T^t is invertible with inverse T^{-t}.

Remark 1.1.6: Suppose $M \subset \mathbb{R}^d$ open. Given an ODE $\dot{\gamma} = X(\gamma)$ with enough regularity for the vector field X, then for each $x \in M$ there exists a solution γ_x starting at x at $t = 0$, then setting $T^t(x) = \gamma_x(t)$ for each $t \in \mathbb{R}$ and $x \in M$ we obtain the flow $\{T^t\}_{t\in\mathbb{R}}$.

Definition 1.1.7: We say that μ is *invariant* under a flow $\{T^t\}_{t\in\mathbb{R}}$ if is invariant under each one of the transformations T^t, $t \in \mathbb{R}$.

Proposition 1.1.8: *The measurable transformation T preserves μ if and only if*

$$\int_M \phi \, \mathrm{d}\mu = \int_M \phi \circ T \, \mathrm{d}\mu \tag{1.1}$$

for all $\phi \in L^1(\mu)$.

Proof. Let μ be invariant under T. For $B \in \mathcal{B}$ let $\mathbb{1}_B$ to be the indicator function of B, then

$$\int_M \mathbb{1}_B \, \mathrm{d}\mu = \mu(B) = \mu(T^{-1}(B)) = \int_M \mathbb{1}_{T^{-1}(B)} \, \mathrm{d}\mu = \int_M \mathbb{1}_B \circ T \, \mathrm{d}\mu.$$

So (1.1) holds for simple functions, next take a function $\phi \in L^1(M, \mu)$, it can be approximated by a sequence of simple functions $(\psi_n)_{n\in\mathbb{N}}$ s.t. $|\psi_n| \leq |\phi|$ for all $n \in \mathbb{N}$, then (1.1) for ϕ follows by dominated convergence.

The converse is trivial. □

1.2 Recurrence

Definition 1.2.1: A set $W \in \mathcal{B}$ is said to be *wandering* for T w,r,t, μ if for every $n, m \in \mathbb{N}$

$$\mu(T^{-n}(W) \cap T^{-m}(W)) = 0.$$

We call $\mathcal{W}(T)$ the family of all the wandering sets for T.

Definition 1.2.2: The sets

$$\mathcal{D}(T) = \bigcup_{W\in\mathcal{W}(T)} W$$
$$\mathcal{C}(T) = M \setminus \mathcal{D}(T)$$

are called respectively *dissipative part* and *conservative part* of T.

Definition 1.2.3: We say that

- T is *dissipative* w.r.t. μ if
$$\mu(M \setminus \mathcal{D}(T)) = 0;$$
- T is *conservative* w.r.t. μ if
$$\mu(M \setminus \mathcal{C}(T)) = 0.$$

Lemma 1.2.4: *Suppose T non-sngular w.r.t. μ. For every $E \in \mathcal{B}$ with $\mu(E) \neq 0$ holds that $\mu(E \cap W) = 0$ for every $W \in \mathcal{W}(T)$ if and only if for every $B \in \mathcal{B}$, $B \subset E$ holds*

$$\mu(\{x \in B \mid T^n(x) \in B \text{ infinitely often}\}) = \mu(B).$$

Proof. Assume first that $\mu(E \cap W) = 0$ for every wandering set W. Fix $B \subset \mathcal{B}$, $B \subset E$ and let

$$N = \{x \in B \mid T^k(x) \notin B \ \ \forall k \in \mathbb{N}_+\} = B \setminus \bigcup_{k\in\mathbb{N}_+} T^{-k}(B).$$

It suffices to show that

$$\mu\left(B \setminus \bigcap_{k\in\mathbb{N}} \bigcup_{n\geq k} T^{-n}(B)\right) = 0.$$

From the definition of N it follows that $T^{-n}(B) \cap T^{-m}(B) = \emptyset$ for every $n, m \in \mathbb{N}_+$, $n \neq m$, so N is wandering and by assumption

$$\mu(N) = \mu(N \cap E) = 0,$$

hence using the non-singularity of T it follows that

$$0 = \mu(T^{-k}(N)) = \mu\left(T^{-k}(B) \setminus \bigcup_{k \geq k+1} T^{-n}(B)\right)$$

so also

$$\mu\left(\bigcup_{n \geq k} T^{-n}(B) \setminus \bigcup_{n \geq k+1} T^{-n}(B)\right) = 0$$

and since $\bigcup_{n \geq k+1} T^{-n}(B) \subset \bigcup_{n \geq k} T^{-n}(B)$ we can rewrite the above equation to obtain

$$\mu\left(\bigcup_{n \geq k} T^{-n}(B)\right) = \mu\left(\bigcup_{n \geq k+1} T^{-n}(B)\right)$$

for every $k \geq 1$. As a consequence

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

for every $k \geq 1$, from this follows $\mu\left(B \setminus \bigcap_{k \in \mathbb{N}} \bigcup_{n \geq k} T^{-n}(B)\right) = 0$.

For the viceversa let W be a wandering set s.t. $\mu(E \cap W) \neq 0$. Set $B = E \cap W$, then $B \subset E$ so by assumption $\mu(\{x \in B \mid T^n(x) \in B \text{ infinitely often}\}) = \mu(B) \neq 0$, but B is wandering, hence there is no $A \subset B$, $A \in \mathcal{B}$, with positive measure s.t. every $x \in A$ returns to B. This gives a contradiction.

□

Theorem 1.2.5 (Halmos recurrence): *Suppose T non-singular w.r.t. μ. The transformation T is conservative w.r.t. μ if and only if for each $E \in \mathcal{B}$ with $\mu(E) \neq 0$ holds that for μ-a.e. $x \in E$ there exists infinitely many $n \in \mathbb{N}$ s.t. $T^n(x) \in E$.*

Proof. If T is conservative, then $\mu(W) = 0$ for every $W \in \mathcal{W}(T)$, hence by the previous Lemma (with $E = M$) we get that for every $B \in \mathcal{B}$ holds

$$\mu(\{x \in B \mid T^n(x) \in B \text{ infinitely often}\}) = \mu(B).$$

Viceversa if for every $B \in \mathcal{B}$ holds

$$\mu(\{x \in B \mid T^n(x) \in B \text{ infinitely often}\}) = \mu(B)$$

then, using again the previous Lemma, we obtain $\mu(W) = 0$ for every $W \in \mathcal{W}(T)$ and T is conservative. □

Theorem 1.2.6 (Poincaré recurrence): *Let μ be a T-invariant finite measure, then T is conservative w.r.t. μ. In particular if $E \in \mathcal{B}$ is s.t. $\mu(E) > 0$, then for μ-a.e. $x \in E$ there exists infinitely many $n \in \mathbb{N}$ s.t. $T^n(x) \in E$.*

Proof. Let $E \in \mathcal{B}$ with $\mu(E) \in (0, \infty)$. Consider the set

$$E_0 = \{x \in E \mid \nexists n \in \mathbb{N}_+ \text{ s.t. } T^n(x) \in E\};$$

let us prove that $\mu(E_0) = 0$. The family of sets $\{T^{-n}(E_0)\}_{n \in \mathbb{N}_+}$ is disjoint. Indeed suppose by contradiction that there exist $m > n \geq 1$ s.t. $T^{-n}(E_0) \cap T^{-m}(E_0) \neq \emptyset$. Then we can take a point $x \in T^{-n}(E_0) \cap T^{-m}(E_0)$ and if we denote $y = T^n(x)$, we have $y \in E_0$ and $T^{m-n}(y) = T^m(x) \in E_0$, but $m - n \geq 1$ hence we have a contradiction.

Since μ is T-invariant we have $\mu(T^{-n}(E_0) = \mu(E_0)$ for every $n \in \mathbb{N}_+$ and we can write

$$\mu\left(\bigcup_{n \in \mathbb{N}_+} T^{-n}(E_0)\right) = \sum_{n \in \mathbb{N}_+} \mu(T^{-n}(E_0)) = \sum_{n \in \mathbb{N}_+} \mu(E_0)$$

now the left-hand side is finite (since μ is finite), hence $\mu(E_0) = 0$ necessarly.

Now denote with F the measurable set of all the points $x \in E$ s.t. $|\{n \in \mathbb{N}_+ \mid T^n(x) \in E\}| < +\infty$. Since every $x \in F$ returnes to E only a finite number of times there exists some $k \in \mathbb{N}$ s.t. $T^k(x) \in E_0$. Hence

$$F \subset \bigcup_{k \in \mathbb{N}} T^{-k}(E_0)$$

and holds

$$\begin{aligned} \mu(F) &\leq \mu\left(\bigcup_{k \in \mathbb{N}} T^{-k}(E_0)\right) \\ &\leq \sum_{k \in \mathbb{N}} \mu(T^{-k}(E_0)) \\ &= \sum_{k \in \mathbb{N}} \mu(E_0) = 0 \end{aligned}$$

thus $\mu(F) = 0$ and we have the thesis. □

An anlogous for flows follows from the previous theorem.

Corollary 1.2.7: *Let μ be a finite invariant measure under a flow $\{T^t\}_{t \in \mathbb{R}}$. If $E \in \mathcal{B}$ is s.t. $\mu(E) > 0$, then for μ-a.e. $x \in E$ there exists a sequence of times $(t_j)_{j \in \mathbb{N}}$, $t_j \to +\infty$, s.t. $T^{t_j}(x) \in E$ for every $j \in \mathbb{N}$.*

Proof. Indeed μ is invariant under the transformation $T = T^1$, hence the thesis follows from the previous theorem applied to this T. □

Now we present a topological version of the Poincaré recurrence theorem.

Definition 1.2.8: A point $x \in M$ is *recurrent* for T if there exists a sequence $(n_j)_{j \in \mathbb{N}} \subset \mathbb{N}$, $n_j \to +\infty$, s.t. $T^{n_j}(x) \to x$.

Theorem 1.2.9 (Poincaré recurrence, topological version): *Suppose M to be a N2 topological space. Let μ be a T-invariant finite measure on M, then μ-a.e. $x \in M$ is recurrent for T.*

Proof. Let $\{U_k\}_{k \in \mathbb{N}}$ be a countable basis of open sets of the topology of M. For each $k \in \mathbb{N}$ denote

$$V_k = \{x \in U_k \mid \nexists n \in \mathbb{N}_+ \text{ s.t. } T^n(x) \in U_k\}$$

then, by Theorem 1.2.6, $\mu(V_k) = 0$ for every $k \in \mathbb{N}$. Hence if $V = \bigcup_{k\in\mathbb{N}} V_k$ we have $\mu(V) = 0$. To conclude let us prove that every $x \in M \setminus V$ is recurrent. Take $x \in M \setminus V$ and U any neighborhood of x. Take an U_k s.t. $U_k \subset U$ (such U_k exists from the definition of basis of open sets). Since $x \notin V$ holds also $x \notin V_k$, hence exists $n \in \mathbb{N}_+$ s.t. $T^n(x) \in U_k \subset U$. Since U is arbitrary the thesis follows. □

Remark 1.2.10: The Poincarè recurrence theorems are not true in general. For example consider $T : \mathbb{R} \longrightarrow \mathbb{R}$ be the function $T(x) = x + 1$, $x \in \mathbb{R}$. Obviously T preserve the Lebesgue measure on $\mathbb{R}$ (which is not finite) and no point $x \in \mathbb{R}$ is recurrent for T (easy to check). In particular, according to the Poincarè recurrence theorem in the topological version, T can not admit any finite invariant measure.

1.3 Kač theorem

Definition 1.3.1: Let μ be a σ-finite measure on M non-singular w.r.t. T. Consider $E \in \mathcal{B}$ with $\mu(E) \neq 0$, we define the *first-return time* function $\rho_E : E \longrightarrow \mathbb{N}_+ \cup \{\infty\}$ s.t. for every $x \in E$

$$\rho_E(x) = \min\{n \in \mathbb{N}_+ \mid T^n(x) \in E\}.$$

if $\{n \in \mathbb{N}_+ \mid T^n(x) \in E\} \neq \emptyset$, $\rho_E(x) = \infty$ otherwise (that is if x has no iterate in E). By Theorem 1.2.6 we have $\rho_E = \infty$ only on a zero measure set.

Remark 1.3.2: Let μ be a σ-finite measure on M non-singular w.r.t. T and conservative (for example μ an invariant finite measure, for Poincarè recurrence Theorem 1.2.6). Consider $E \in \mathcal{B}$ with $\mu(E) \neq 0$ and define

$$E_0 = \{x \in E \mid T^n(x) \notin E \ \ \forall n \in \mathbb{N}_+\}.$$

By Halmos recurrence Theorem 1.2.5 we have that $\mu(E_0) = 0$. In particular ρ_E is finite μ-a.e. on E.

Theorem 1.3.3 (Kač): *Let μ be a σ-finite invariant and conservative measure and $E \in \mathcal{B}$ with $\mu(E) \neq 0$, then ρ_E is finite μ-a.e. on E. Moreover if μ is invariant and finite, then $\rho_E \in L^1(\mu)$ and*

$$\int_E \rho_E \, \mathrm{d}\mu = \mu(M) - \mu(E_0^*)$$

where $E_0^* = \{x \in M \mid T^n(x) \notin E \ \forall n \in \mathbb{N}\}$.

Proof. The first part has been proven in the previous Remark. Suppose μ finite. Let E_0 be like in the previous Remark and E_0^* like in the statement. For every $n \in \mathbb{N}_+$ define also

$$\begin{aligned} E_n &= \{x \in E \mid T^i(x) \notin E \ \forall i = 1, \ldots, n-1,\ T^n(x) \in E\} \\ E_n^* &= \{x \in M \mid x \notin E,\ T^i(x) \notin E \ \forall i = 1, \ldots, n-1,\ T^n(x) \in E\}. \end{aligned}$$

Observe that $E_n = \rho_E^{-1}(\{n\})$ for every $n \in \mathbb{N}_+$. Moreover the family $\{E_n\}_{n\in\mathbb{N}} \cup \{E_n^*\}_{n\in\mathbb{N}}$ is a partition of M. Hence (using the previous remark)

$$\begin{aligned} +\infty > \mu(M) &= \sum_{n\in\mathbb{N}} (\mu(E_n) + \mu(E_n^*)) \\ &= \mu(E_0^*) + \sum_{n\in\mathbb{N}_+} (\mu(E_n) + \mu(E_n^*)) \\ &= \sum_{n\in\mathbb{N}_+} \mu(E_n) + \sum_{n\in\mathbb{N}} \mu(E_n^*) \end{aligned} \tag{1.2}$$

that, in particular, gives $\mu(E_n^*) \to 0$. Now let us prove that for every $n \in \mathbb{N}$ holds

$$T^{-1}(E_n^*) = E_{n+1}^* \cup E_{n+1}, \tag{1.3}$$

in fact $T(y) \in E_n^*$ implies that $T^i(T(y)) \notin E$ for $i = 1, \ldots, n-1$ and $T^n(T(y)) = T^{n+1}(y) \in E$ and this can be true only if and only if $y \in E_{n+1}^*$ or $y \in E_{n+1}$. Using (1.3) and the invariance of μ we obtain

$$\mu(E_n^*) = \mu(T^{-1}(E_n^*)) = \mu(E_{n+1}^*) + \mu(E_{n+1})$$

for every $n \in \mathbb{N}$ and using this relation compulsively we get

$$\mu(E_n^*) = \mu(E_m^*) + \sum_{i=n+1}^{m} \mu(E_i)$$

for every $m, n \in \mathbb{N}$, $m > n$. But we know that $\mu(E_m^*) \to 0$, so taking the limits for $m \to +\infty$ we find that

$$\mu(E_n^*) = \sum_{i \geq n+1} \mu(E_i).$$

Substituting the previous equation in (1.2) we obtain

$$\begin{aligned} \mu(M) - \mu(E_0^*) &= \sum_{n \in \mathbb{N}_+} \left(\sum_{i \geq n} \mu(E_i) \right) \\ &= \sum_{n \in \mathbb{N}_+} n\mu(E_n) = \int_E \rho_E \, \mathrm{d}\mu. \end{aligned}$$

$\square$

Corollary 1.3.4: *Let μ be a finite invariant measure and $E \in \mathcal{B}$ with $\mu(E) > 0$. Then $\rho_E \in L^1(\mu)$ and if we suppose that $\mu(E_0^*) = 0$ we have*

$$\frac{1}{\mu(E)} \int_E \rho_E \, \mathrm{d}\mu = \frac{\mu(M)}{\mu(E)}.$$

In words, the above equation means that the mean return time *to E (the left-hand side) is inversely proportional to the measure of E.*

1.4 Existence of invariant measures for continuous transformations

Let (M, d) be a compact metric space and $T : M \to M$ a continuous map. Like in the previous chapter $\mathcal{B}$ will be the Borel σ-algebra of M.

Definition 1.4.1: We define $\mathcal{M}(M)$ the family of all the signed finite Borel measures on M. Moreover $\mathcal{M}_1(M)$ will be the set of all the probability measures on M.

Remark 1.4.2: Obviously $\mathcal{M}_1(M) \subset \mathcal{M}(M)$. It is well known that if $\|\cdot\|_{\mathrm{TV}}$ is the total variation norm, the couple $\mathcal{M}(M)$, $\|\cdot\|_{\mathrm{TV}}$ form a Banach space, and with this structure, by the Riesz-Markov Theorem $\mathcal{M}(M)$ is linearly isometric to the topological dual of $C(M)$, that is $C(M)^*$. In particular we can use this isometry to put on $\mathcal{M}(M)$ the *weak* topology* induced by from the weak* topology of $C(M)^*$.

Lemma 1.4.3: *The space $C(M)$ with the uniform norm $\|\cdot\|_\infty$ is separable.*

Proof. Consider $F = \{\mathbb{1}\} \cup \{f_i\}_{i\in\mathbb{N}}$, where $\mathbb{1}$ is the function constant 1 and for every $i \in \mathbb{N}$ and every $x \in M$

$$f_i(x) = d(x, x_i)$$

with $\{x_i\}_{i\in\mathbb{N}}$ a countable dense subset of M (which is separable since it is a compact metric space). Then for $\mathbb{K} = \mathbb{R}, \mathbb{Q}$ consider

$$\mathbb{K}[F] = \left\{ \prod_{i=1}^{n} g_i^{n_i} \mid N \in \mathbb{N}_+,\ n_i \in \mathbb{N}_+,\ g_i \in F\ \forall i = 1, \dots, N \right\}.$$

Observe that both $\mathbb{Q}[F]$ and $\mathbb{R}[F]$ are algebras of functions contained in $C(M)$ and that $\mathbb{Q}[F]$ is dense in $\mathbb{R}[F]$ uniformly. In fact If

$$g = \sum_{(n_i)_{i=1}^k \in S} \lambda_{n_1,\dots,n_k} g_1^{n_1} \dots g_k^{n_k} \in \mathbb{R}[F]$$

with $S \subset \mathbb{N}^k$ finite, $k \in \mathbb{N}_+$, then for every $(n_i)_{i=1}^k \in S$ exists $(a^{(j)}_{n_1,\dots,n_k})_{j\in\mathbb{N}} \subset \mathbb{Q}$ s.t. $a^{(j)}_{n_1,\dots,n_k} \to \lambda_{n_1,\dots,n_k}$ when $j \to \infty$. Then if we define for each $j \in \mathbb{N}$

$$g^{(j)} = \sum_{(n_i)_{i=1}^k \in S} a^{(j)}_{n_1,\dots,n_k} g_1^{n_1} \dots g_k^{n_k} \in \mathbb{Q}[F]$$

we have that

$$\left\| g - g^{(j)} \right\|_\infty \le \left(\sum_{(n_i)_{i=1}^k \in S} \left\| g_1^{n_1} \dots g_k^{n_k} \right\|_\infty \right) \max_{(n_i)_{i=1}^k \in S} |\lambda n_1, \dots, n_k - a^{(j)}_{n_1,\dots,n_k}| \to 0$$

for $j \to \infty$. Therefore, to obtain the thesis we can show that $\mathbb{R}[F]$ is uniformly dense in $C(M)$ and to do this we use the Stone Theorem. The algebra of functions $\mathbb{R}[F]$ contains $\mathbb{1}$ and let us show that is separating. Let $x, y \in M,\ x \neq y$, since $\{x_i\}_{i\in\mathbb{N}}$ is dense there exists a x_m s.t.

$$d(x, x_m) \le \frac{d(x, y)}{3}$$

and can not happens $d(x, x_m) = d(y, x_m)$, otherwise

$$\begin{aligned} d(x, y) &\le d(x, x_m) + d(y, x_m) \\ &< \frac{2}{3} d(x, y) \end{aligned}$$

that is clearly an absurd. Hence f_m separetes x and y. □

Corollary 1.4.4: *The set $\mathcal{M}_1(M)$ is weak* sequentially compact in $\mathcal{M}(M)$.*

Proof. By the well known Banach-Alaoglu Theorem the unitary ball of $\mathcal{M}(M)$ is weak* sequentially compact and $\mathcal{M}_1(M)$ is contained in it. Moreover $\mathcal{M}_1(M)$ is the intersection of kernels of evaluation operators of functions in $C(M)$ so it is weak* closed, hence it is weak* sequentially compact. □

Definition 1.4.5: We define $\mathcal{M}_1^T(M)$ to be the set of all the T-invariant probability measures in $\mathcal{M}_1(M)$.

Theorem 1.4.6 (Krylov-Bogolyubov): *The set* $\mathcal{M}_1^T(M)$ *is a* non-empty*, weak* sequentially compact, convex subset of* $\mathcal{M}(M)$.

Proof. Step 1: There exists a T*-invariant borel probability measure defined on* M.

Fix $x_0 \in M$ and for every $n \in \mathbb{N}_+$ define the borel probability measure $\mu_n : \mathcal{B} \to [0,1]$ setting

$$\int_M \phi \, \mathrm{d}\mu_n = \frac{1}{n} \sum_{k=0}^{n-1} \phi(T^k(x_0))$$

for every $\phi \in C(M)$. In the prequel we have shown that $\mathcal{M}_1(M)$ is weak* sequentially compact, hence there exists s subsequence $(\mu_{n_k})_{k \in \mathbb{N}}$ and a $\mu \in \mathcal{M}_1(M)$ s.t. $\mu_{n_k} \xrightarrow{*} \mu$ for $k \to \infty$. It remains to prove that μ is T-invariant. Pick $\phi \in C(M)$

$$\begin{aligned} \int_M \phi \circ T \, \mathrm{d}\mu &= \lim_{k \to \infty} \frac{1}{n_k} \sum_{i=1}^{n_k} \phi(T^i(x_0)) \\ &= \lim_{k \to \infty} \frac{1}{n_k} \sum_{i=0}^{n_k - 1} \phi(T^i(x_0)) = \int_M \phi \, \mathrm{d}\mu \end{aligned}$$

so that $\mu \in \mathcal{M}_1^T(M)$.

Step 2: The set $\mathcal{M}_1^T(M)$ is convex, bounded (w.r.t. the total variation norm) and weak* closed (by definition, in fact $\mathcal{M}_1^T(M) = \{\mu \in \mathcal{M}(M) \mid \mu \geq 0, \ \mu(M) = 1, \ \mu = f_*\mu\}$ that is clearly an intersection of preimages of evaluation operators on $\mathcal{M}(M)$ of closed subsets of $[0,1]$), hence it is weak* sequentially compact since it is contained in $\mathcal{M}_1(M)$. □

2

Ergodic Theorems

2.1 Functional and Von Neumann ergodic theorems

Let $(X, \|\cdot\|)$ be a Banach space and $\mathrm{L}(X)$ be the space of all the bounded linear operators form X to X. Recall that a *projection* is a $P \in \mathrm{L}(X)$ s.t. $P^2 = P$. Recall also that if P is a projection then $X = X_0 \oplus X_1$ with $X_0 = \ker(P)$ and $X_1 = \mathrm{ran}(P) = \ker(I - P)$ where I is the identity operator of X.

Theorem 2.1.1 (Functional ergodic theorem): *Let $L \in \mathrm{L}(X)$. Assume that there exists a constant $c \geq 1$ s.t. $\|L^n\|_{op} \leq c$ $\forall n \in \mathbb{N}$. For $n \in \mathbb{N}$ define the bounded operators $S_n : X \to X$ by $S_n = \frac{1}{n}\sum_{k=0}^{n-1} L^k$. Then the following holds.*

(1) Let $x \in X$,then the sequence $(S_n x)_{n\in\mathbb{N}}$ converges if and only if it has a weakly convergent subsequence.

(2) The set $Z = \{x \in X \mid (S_n x)_{n\in\mathbb{N}} \text{ converge}\}$ is closed, T-invariant linear subspace of X and

$$Z = \ker(I - L) \oplus \overline{\mathrm{ran}(I - L)}.$$

Moreover if X is reflexive holds $Z = X$.

(3) Define the bounded operator $S : Z \to Z$ by $S(x+y) = x$ for $x \in \ker(I-L)$ and $y \in \overline{\mathrm{ran}(I - L)}$. Then $\lim_{n\to+\infty} S_n z = Sz$ $\forall z \in Z$ and $SL = LS = S^2 = S$ and $\|S\|_{op} \leq c$.

Proof. Step 1: Let $n \in \mathbb{N}$. Then $\|S_n\|_{op} \leq \frac{1+c}{n}$.

By the triangle inequality holds

$$\|S_n\|_{\mathrm{op}} \leq \frac{1}{n}\sum_{k=0}^{n-1}\|L^k\|_{\mathrm{op}} \leq c$$

for every $n \in \mathbb{N}$,. Moreover $S_n(I - L) = \frac{1}{n}(I - L^n)$ so that for every $n \in \mathbb{N}$

$$\|S_n(I - L)\|_{\mathrm{op}} = \frac{1}{n}\|I - L^n\|_{\mathrm{op}} \leq \frac{1+c}{n}.$$

Step 2: Le $x \in \ker(I - L)$, then $S_n x = x$ for every $n \in \mathbb{N}$ and $\|x\| \le c\,\|x + \xi - L\xi\|$ for all $\xi \in X$.

Since $Lx = x$, by induction follows $L^k x = x$ $\forall k \in \mathbb{N}$, hence $x = \frac{1}{n}\sum_{k=0}^{n-1} L^k x = S_n x$ $\forall n \in \mathbb{N}$. Moreover, by Step 1, holds that

$$\lim_{n\to\infty} \|S_n(\xi - L\xi)\|_{\mathrm{op}} \le \lim_{n\to\infty} \frac{1+c}{n}\,\|\xi\|$$

hence

$$\|x\| = \lim_{n\to\infty} \|x + S_n(\xi - L\xi)\| = \lim_{n\to\infty} \|S_n(x + \xi - L\xi)\| \le c\,\|x + \xi - L\xi\|$$

for every $\xi \in X$.

Step 3: If $x \in \ker(I - L)$ and $y \in \overline{\mathrm{ran}(I - L)}$ then $\|x\| \le c\,\|x + y\|$.

Choose a sequence $(\xi_n)_{n\in\mathbb{N}} \subset X$ s.t. $y = \lim_{n\to\infty}(\xi_n - L\xi_n)$. Then by Step 2 we have

$$\|x\| \le c\,\|x + \xi_n - \xi_n\|$$

for every $n \in \mathbb{N}$ and taking the limit for $n \to \infty$ we obtain $\|x\| \le x\,\|x + y\|$,

Step 4: $\ker(I - L) \cap \overline{\mathrm{ran}(I - L)} = \{0\}$ and the direct sum $Z = \ker(I - L) \oplus \overline{\mathrm{ran}(I - L)}$ is a closed subspace of X.

Let $x \in \ker(I - L) \cap \overline{\mathrm{ran}(I - L)}$ and define $y = -x$. Then by Step 3 holds

$$\|x\| \le c\|x + y\| = 0$$

hence $x = 0$. This prove the first part. Now we prove that Z is closed. Choose two sequences $(x_n)_{n\in\mathbb{N}} \subset \ker(I - L)$ and $(y_n)_{n\in\mathbb{N}} \subset \overline{\mathrm{ran}(I - L)}$ s.t. $z_n = z_n + y_n \to z$ for some $z \in X$. Then $(z_n)_{n\in\mathbb{N}}$ is a Cauchy sequence and for Step 3 also $(x_n)_{n\in\mathbb{N}}$. This implies that also $(y_n)_{n\in\mathbb{N}}$ is a Cauchy sequence and as a consequence defining $x = \lim_{n\to\infty} x_n \in \ker(I-L)$ and $y = \lim_{n\to\infty} y_n \in \overline{\mathrm{ran}(I - L)}$ we have $z = x + y \in Z$.

Step 5: If $z \in Z$ then $Tz \in Z$.

Let $z \in Z$, then $z = x + y$ with $x \in \ker(I - L)$ and $y \in \overline{\mathrm{ran}(I - L)}$. Choosing a sequence $(\xi_n)_{n\in\mathbb{N}} \subset X$ s.t. $y = \lim_{n\to\infty}(I - L)\xi_n$ we get

$$Ty = \lim_{n\to\infty} L(I - L)\xi_n = \lim_{n\to\infty} (I - L)L\xi_n \in \overline{\mathrm{ran}(I - L)}$$

hence $Tz = Tx + Ty = x + Ty \in Z$.

Step 6: If $x \in \ker(I - L)$ and $y \in \overline{\mathrm{ran}(I - L)}$ then $x = \lim_{n\to\infty} S_n(x + y)$.

By Step 1 the sequence $(S_n(I-L)\xi)_{n\in\mathbb{N}}$ converges to 0 for every $\xi \in X$. Hence, since $\|S_n\|_{\mathrm{op}} \le c$ by the uniform limitness principle follows that $\lim_{n\to\infty} S_n y = 0$ for all $y \in \overline{\mathrm{ran}(I - L)}$. Moreover $S_n x = x$ for every $n \in \mathbb{N}$ by Step 2, so that $x = \lim_{n\to\infty} S_n x = \lim_{n\to\infty} S_n(x + y)$.

Step 7: Let $x, z \in X$, t.f.a.e.

(1) $x \in \ker(I - L)$ and $z - x \in \overline{\mathrm{ran}(I - L)}$;

(2) $\lim_{n\to\infty} \|S_n z - x\| = 0$;

(3) $\exists (n_k)_{k\in\mathbb{N}} \subset \mathbb{N}$, $n_k \nearrow \infty$, s.t. $\lim_{i\to\infty}\langle x^*, S_{n_i} z\rangle = \langle x^*, x\rangle$ for every $x^* \in X^*$.

Taht (1) implies (2) follows directly from Step 6. Moreover (2) implies (3) is obvious (since strong convergence implies weak convergence). Let us prove that (3) implies (1). Fix $x^* \in X^*$, then $L^* x^* = x^* \circ L \in X^*$ and

$$\begin{aligned}\langle x^*, x - Lx\rangle &= \langle x^* - L^* x^*, x\rangle \\ &= \lim_{i\to\infty} \langle x^* - L^* x^*, S_{n_i} z\rangle \\ &= \lim_{i\to\infty} \langle x^*, (I - L) S_{n_i} z\rangle = 0\end{aligned}$$

where the last equality follows from Step 1. Hence $\|(I-L)x\| = 0$ (from Hahn-Banach theorem), so $x \in \ker(I-L)$. Now we prove that $z - x \in \overline{\operatorname{ran}(I-L)}$. Assume, by contradiction, that $z - x \notin \overline{\operatorname{ran}(I-L)}$, then by Hahn-Banach theorem we know that exist $x^* \in X^*$ s.t. $\langle x^*, z-x\rangle = 1$ and $\langle x^*, \xi - L\xi\rangle = 0$ for any $\xi \in X$. This implies

$$\langle x^*, T^k\xi - T^{k+1}\xi\rangle = 0$$

for every $k \in \mathbb{N}$ and every $\xi \in X$. Hence $\langle x^*, \xi\rangle = \langle x^*, T^k\xi\rangle$ for every $k \in \mathbb{N}$ and every $\xi \in X$, thus

$$\langle x^*, S_n z\rangle = \frac{1}{n}\sum_{k=0}^{n-1}\langle x^*, T^k z\rangle = \langle x^*, z\rangle$$

for every $n \in \mathbb{N}_+$. So by (3) we have

$$\langle x^*, z - x\rangle = \lim_{i\to\infty}\langle x^*, S_{n_i}z - x\rangle = 0$$

that is an equality that contradicts $\langle x^*, z-x\rangle = 1$.

Step 8: We conclude the proof.

The subspace Z is closed and L-invariant by Step 4 and Step 5. Moreover, by Step 7, for all $z \in X$ holds $z \in Z$ if and only if $(S_n z)_{n\in\mathbb{N}_+}$ converge in X if and only if $(S_n z)_{n\in\mathbb{N}_+}$ admits a subsequence weakly convergent. If X is reflexive, it is locally weakly sequentially compact so that $Z = X$ (since by Step 1 $(S_n z)_{n\in\mathbb{N}_+}$ is bounded in X for every $z \in X$). Now let $S : Z \to Z$ be like in the statement of the theorem. Then $\|S\|_{\mathrm{op}} \leq c$ by Step 3, the equation $Sz = \lim_{n\to\infty} S_n z$ follows from Step 6 and $S^2 = S$ by definition. Lastly $SL = LS = S$ follows from the fact that S commutes with $L_{|Z}$ and vanishes on $\overline{\operatorname{ran}(I-T)}$. □

Now let us fix a measurable space $(M, \mathcal{B})$, a measurable transformation $T : M \to M$ and a T-invariant probability measure μ on M.

Definition 2.1.2: We say that a function $\phi : M \to \mathbb{R}$ measurable is (T,μ)*-invariant* (or only *invariant*) if $\phi = \phi \circ T$ μ-a.e. on M. Moreover we say that a set $B \in \mathcal{B}$ is (T,μ)*-invariant* (or only *invariant*) if $\mathbb{1}_B$ is a (T,μ)-invariant function.

Definition 2.1.3: Fix $p \in (1,\infty)$. The operator $L_T : L^p(\mu) \to L^p(\mu)$ s.t. $L_T\phi = \phi \circ T$ is called *T-Koopman operator* on $L^p(\mu)$.

Remark 2.1.4: For $p \in (1,\infty)$ is well known that the space $L^p(\mu)$ is reflexive, so if wemconsider the T-Koopman operator L_T on $L^p(\mu)$ we have that $\|L_T\phi\|_{L^p(\mu)} = \|\phi\|_{L^p(\mu)}$ for every $\phi \in L^p(\mu)$, hence $\|L_T\|_{\mathrm{op}} = 1$. In particular all the requirements for the application of the previous theorem are fulfilled. Thus if we consider the projection $S : Z \to \ker(I - L_T)$ given by the functional ergodic theorem, we have $Z = L^p(\mu)$ and the space $\ker(I - L_T)$ is exactly the space of all the invariant functions in $L^p(\mu)$. In particular every function $\phi \in L^p(\mu)$ admits the *projection $\tilde{\phi} = S\phi$ to the space of all the invariant functions in $L^p(\mu)$.*

Theorem 2.1.5 (Von Neumann ergodic therorem): *Fix $p \in (1,\infty)$. Given $\phi \in L^p(\mu)$ let $\tilde{\phi} \in L^p(\mu)$ be the projection of ϕ to the linear subspace of all the invariant functions in $L^p(\mu)$. Then the sequence $S_n\phi = \frac{1}{n}\sum_{k=0}^{n-1}\phi \circ T^k$, $n \in \mathbb{N}_+$, converge to $\tilde{\phi}$ in $L^p(\mu)$ for $n \to \infty$.*

Proof. Let L_T be the T-Koopman operator on $L^p(\mu)$. Like we noticed in the previous remark L_T satisfies all the requirements of the functional ergodic theorem. Like we noticed in the previous observation the space $L^p(\mu)$ is reflexive, hence (by the functional ergodic theorem) the sequence $(S_n\phi)_{n\in\mathbb{N}_+}$ converges in $L^p(\mu)$ to $\tilde{\phi}$. □

2.2 Kingman subadditive ergodic theorem

Let us fix a measurable space $(M, \mathcal{B})$, a measurable transformation T and μ a T-invariant probability measure on M.

Definition 2.2.1: A sequence of functions $\phi_n : M \to \mathbb{R}$ is said to be (T, μ)-*subadditive* (or only *subadditive*) if for any $m, n \in \mathbb{N}_+$

$$\phi_{m+n} \leq \phi_m + \phi_n \circ T^m \quad \mu\text{-a.e. on } M.$$

Theorem 2.2.2 (Kingman subadditive ergodic theorem): *Let $\phi_n : M \to \mathbb{R}$, $n \in \mathbb{N}_+$, a subadditive sequence of measurable functions s.t. $\phi_1^+ \in L^1(\mu)$. Then the sequence $(\phi_n/n)_{n\in\mathbb{N}_+}$ converges μ-a.e. on M to some function $\phi : M \to [-\infty, +\infty)$ that is invariant under T. Moreover $\phi^+ \in L^1(\mu)$ and*

$$\int_M \phi \, \mathrm{d}\mu = \lim_{n\to\infty} \int_M \phi_n \, \mathrm{d}\mu = \inf_{n\in\mathbb{N}_+} \frac{1}{n} \int_M \phi_n \, \mathrm{d}\mu \in [-\infty, +\infty).$$

In order to prove this theorem we need some lemmas.

Definition 2.2.3: A sequence $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ is *subadditive* if fro every $m, n \in \mathbb{N}$ holds $a_{m+n} \leq a_m + a_n$.

Lemma 2.2.4: *If $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ is a subadditive sequence then*

$$\exists \lim_{n\to\infty} \frac{a_n}{n} = \inf_{n\in\mathbb{N}_+} \frac{a_n}{n}.$$

Proof. Given $d \in \mathbb{N}_+$, for all $n \in \mathbb{N}_+$ we have $n = p_n d + k$ with $0 \leq k \leq d$ and $p_n = \lfloor n/d \rfloor$. Then for subadditivity

$$\frac{a_n}{n} \leq \frac{1}{n}(p_n a_d + a_k)$$

and

$$\inf_{m\in\mathbb{N}_+} \frac{a_m}{m} \leq \frac{a_n}{n} \leq \frac{1}{n}(p_n a_d + a_k) \leq \frac{p_n d}{n} \frac{a_d}{d} + \frac{1}{n} \max_{1\leq k\leq d} a_k.$$

Hence for all $d \in \mathbb{N}_+$ we have

$$\inf_{m\in\mathbb{N}_+} \frac{a_m}{m} \leq \limsup_{n\to\infty} \frac{a_n}{n} \leq \frac{a_d}{d}$$

where we used that $p_n d/n \to 1$ and $(\max_{1\leq k\leq d} a_k)/n \to 0$ for $n \to \infty$. So taking the infimum for $d \in \mathbb{N}_+$ we obtain

$$\limsup_{n\to\infty} \frac{a_n}{n} = \inf_{n\in\mathbb{N}_+} \frac{a_n}{n}$$

then the thesis follows easily. □

Remark 2.2.5: Let $(\phi_n)_{n\in\mathbb{N}}$ be a sequence like in the statement of the Kingman ergodic theorem, then by subadditivity we have

$$\phi_n \leq \phi_1 + \phi_1 \circ T + \cdots + \phi_1 \circ T^{n-1}$$

and this relation remains valid als if we replace ϕ_1 and ϕ_n with ϕ_1^+ and ϕ_n^+. Hence the hypothesis $\phi_1^+ \in L^1(\mu)$ implies $\phi_n^+ \in L^1(\mu)$ for every $n \in \mathbb{N}_+$. Moreover if we set

$$a_n = \int_M \phi_n \, \mathrm{d}\mu$$

for every $n \in \mathbb{N}_+$, then $(a_n)_{n\in\mathbb{N}_+} \subset [-\infty, +\infty)$ is subadditive, indeed by subadditivity of $(\phi_n)_{n\in\mathbb{N}_+}$ we have

$$a_{n+m} = \int_M \phi_{n+m} \, d\mu \leq \int_M (\phi_n + \phi_m \circ T^n) \, d\mu = a_n + \int_M \phi_m \circ T^n \, d\mu = a_n + a_m$$

for every $n, m \in \mathbb{N}_+$ (where, in the last equality, we used that μ is T-invariant). Therefore by the previous lemma there exists the limit

$$L = \lim_{n\to\infty} \frac{a_n}{n} = \inf_{n\in\mathbb{N}_+} \frac{a_n}{n} \in [-\infty, +\infty).$$

Define $\phi_+, \phi_- : M \longrightarrow \mathbb{R}$ to be for all $x \in M$ $\phi_-(x) = \liminf_{n\to\infty} \frac{\phi_n(x)}{n}$ and $\phi_+(x) = \limsup_{n\to\infty} \frac{\phi_n(x)}{n}$ and observe that obviously $\phi_- \leq \phi_+$. Moreover $\phi_-^+ \in L^1(\mu)$, because using Fatou lemma and the T-invariance of μ, we get

$$\int_M \phi_-^+ \, d\mu = \int_M \liminf_{n\to\infty} \frac{\phi_n^+}{n} \, d\mu \leq \liminf_{n\to\infty} \frac{1}{n} \int_M \phi_n^+ \, d\mu \leq \int_M \phi_1^+ \, d\mu < +\infty.$$

The strategy of the proof of Kingman ergodic theorem is the following: we are going to prove that if ϕ_n is bounded from below for every $n \in \mathbb{N}_+$ then

$$\int_M \phi_- \, d\mu \geq L \geq \int_M \phi_+ \, d\mu.$$

Consequently $\phi_- = \phi_+ = \phi$ μ-a.e. on M and their integrals coincide with L. Moreover the two functions ϕ_+ and ϕ_- are invariants, indeed by subadditivity, for every $n \in \mathbb{N}$ holds $\phi_{n+1} \leq \phi_1 + \phi_n \circ T$, hence

$$\frac{n+1}{n} \frac{\phi_{n+1}}{n+1} \leq \frac{\phi_1}{n} + \frac{\phi_n \circ T}{n}$$

and taking the $\liminf_{n\to\infty}$ and the $\limsup_{n\to\infty}$ we obtain

$$\phi_\pm \leq \phi_\pm \circ T.$$

At the same time, using the T-invariance of μ, we get the inequality

$$\int_M \phi_\pm \, d\mu = \int_M \phi_\pm \circ T \, d\mu$$

that gives $\int_M [\phi_\pm \circ T - \phi_\pm] \, d\mu = 0$, but we saw that $\phi_\pm \circ T - \phi_\pm \geq 0$, so it follows that

$$\phi_\pm \circ T = \phi_\pm \quad \mu\text{-a.e. on } M.$$

Thus the Kingman ergodic theorem will be proved in the case of boundess from below, at the end we will remove this boundness assumption using a truncation argument.

In the following lemmas we will use the ojects and notations introduced in this remark.

Lemma 2.2.6: *Assume $\phi_- > -\infty$ at every point of M. Fix $\varepsilon > 0$, let us define for evry $k \in \mathbb{N}_+$*

$$E_k = \{x \in \mathcal{M} \mid \phi_j \leq j(\phi_-(x) + \varepsilon) \text{ for some } j \in \{1, \dots, k\}\}$$

and

$$\psi_k(x) = \begin{cases} \phi_-(x) + \varepsilon & x \in E_k \\ \phi_1(x) & x \notin E_k. \end{cases}$$

Then $E_k \subset E_{k+1}$ *for every* $k \in \mathbb{N}_+$*,* $M = \bigcup_{k\in\mathbb{N}_+} E_k$*,* $\phi_1(x) > \phi_-(x) + \varepsilon$ *for every* $x \in E_k^c$ *and* $\psi_k \searrow \phi_- + \varepsilon$ *on* M*. In particular*

$$\int_m \psi_k \,\mathrm{d}\mu \to \int_M \phi_- \,\mathrm{d}\mu + \varepsilon \text{ for } k \to \infty$$

Moreover holds the estimate

$$\phi_n(x) \leq \sum_{i=0}^{n-k-1} \psi_k(T^i(x)) + \sum_{i=n-k}^{n-1} \max\{\psi_k, \phi_1\}(T^i(x))$$

for every $n > k \geq 1$ *and* μ*-a.e.* $x \in M$*.*

Proof. Remember that $\phi_- = \liminf_{n\to\infty} \phi_n/n$, so if $x \in M$ there exists $n \in \mathbb{N}$ s.t.

$$\inf_{1\leq j\leq n} \frac{\phi_j}{j} \leq \phi_-(x) + \varepsilon,$$

hence $x \in E_n$ and for all $k \geq n$

$$\inf_{1\leq j\leq k} \frac{\phi_j}{j} \leq \phi_-(x) + \varepsilon$$

that is $x \in E_k$, hence $E_k \nearrow M$. If $x \in E_k^c$ then for every $j \in \{1, \ldots, \}$ holds

$$\phi_j(x) > j(\phi_-(x) + \varepsilon)$$

in particular $\phi_1(x) > \phi_-(x) + \varepsilon$ for every $x \in \bigcup_{k\in\mathbb{N}_+} E_k^c$. Hence we also have that $\psi_k \searrow \phi_- + \varepsilon$ on M and from the monotone convergence theorem (applied to the non-negative increasing sequence of functions $(\phi_1^+ - \psi_k)_{k\in\mathbb{N}_+}$) we obtain

$$\int_M \psi_k \,\mathrm{d}\mu \to \int_M (\phi_- + \varepsilon)\,\mathrm{d}\mu \text{ as } k \to \infty.$$

Now let us prove the estimate of the statement. Take $x \in M$ s.t. $\phi_-(x) = \phi_-(T^j(x))$ for all $j \in \mathbb{N}_+$ (remember that ϕ_- is invariant so $\phi_- = \phi_- \circ T^j$ μ-a.e. for every $j \in \mathbb{N}_+$, hence $\phi_- = \phi_- \circ T^j$ for every $j \in \mathbb{N}_+$ μ-a.e.). Now consider the sequence, possibly finite, of integers $m_0 \leq n_1 > m_1 \leq n_2 < m_2 \leq \ldots$ defined inductively as follows. Define $m_0 = 0$ and let n_j be the smallest integer greater or equal to m_{j-1} satisfying $T^{n_j}(x) \in E_k$, if such an integer exists, otherwise stop the process (and the sequence is finite). Then by the definition of E_k exists m_j s.t. $1 \leq m_j - n_j \leq k$ and

$$\phi_{m_j-n_j}(T^{n_j}(x)) \leq (m_j - n_j)(\phi_-(T^{n_j}(x) + \varepsilon) \tag{2.1}$$

This completes the definition of the sequence above. Now take $n \geq k$ and let $l \geq 0$ be the largest integer s.t. $m_l \leq n$. By subadditivity for every $j = 1, \ldots, l$ s.t. $m_{j-1} \neq n_j$ holds

$$\phi_{n_j-m_{j-1}}(T^{m_{j-1}}(x)) \leq \sum_{i=m_{j-1}}^{n_j-1} \phi_1(T^i(x))$$

and analogously

$$\phi_{n-m_l}(T^{m_l}(x)) \leq \sum_{i=m_l}^{n-1} \phi_1(T^i(x)).$$

Thus using the sublinearity compulsively: $n = (n - m_l) + m_l$ so

$$\phi_n \leq \phi_{m_l} + \phi_{n-m_l} \circ T^{m_l},$$

$m_l = (m_l - n_l) + n_l$ so

$$\phi_n \leq \phi_{n_l} + \phi_{m_l - n_l} \circ T^{n_l} + \phi_{n-m_l} \circ T^{m_l},$$

$n_l = (n_l - m_{l-1}) + m_{l-1}$ so

$$\phi_n \leq \phi_{m_l - 1} + \phi_{n_l - m_{l-1}} \circ T^{m_{l-1}} + \phi_{m_l - n_l} \circ T^{n_l} + \phi_{n-m_l} \circ T^{m_l}$$

and continuing in this way, at the end, we obtain

$$\phi_n(x) \leq \sum_{j=1}^{l} \phi_{n_j - m_{j-1}}(T^{m_{j-1}}(x)) + \phi_{n-m_l}(T^{M_l}(x)) + \sum_{j=1}^{l} \phi_{m_j - n_j}(T^{n_j}(x))$$

and using the estimates written before

$$\phi_n(x) \leq \sum_{i \in I} \phi_1(T^i(x)) + \sum_{j=1}^{l} \phi_{m_j - n_j}(T^{n_j}(x))$$

with $I = \left[\bigcup_{j=1}^{l} [m_{j-1}, n_j) \cup [m_l, n)\right] \cap \mathbb{N}$. Now observe that

$$\phi_1(T^i(x)) = \psi_k(T^i(x)) \quad \forall i \in \left[\bigcup_{j=1}^{l} [m_{j-1}, n_j) \cup [m_l, \min\{n_{l+1}, n\})\right] \cap \mathbb{N} \tag{2.2}$$

since $T^i(x) \in E_k^c$ for all those i. Moreover since ϕ_- is invariant it is constant in μ-almost all the orbits and $\psi_k \geq \phi_- + \varepsilon$, the relation (2.1) gives for every $j = 1, \ldots, l$

$$\begin{aligned}
\phi_{m_j - n_j}(T^{n_j}(x)) &\leq (m_j - n_j)(\phi_-(T^{n_j}(x) + \varepsilon) \\
&= \sum_{i=n_j}^{m_j - 1} (\phi_-(T^{n_j}(x)) + \varepsilon) \\
&= \sum_{i=n_j}^{m_j - 1} (\phi_-(T^i(x)) + \varepsilon) \\
&\leq \sum_{i=n_j}^{m_j - 1} \psi_k(T^i(x)).
\end{aligned}$$

In this way using (2.2) we get

$$\phi_n(x) \leq \sum_{i=0}^{\min\{n_{l+1}, n\} - 1} \psi_k(T^i(x)) + \sum_{i=n_{l+1}}^{n-1} \phi_1(T^i(x))$$

and since $n_{l+1}, n > n - k$ the thesis follows easily. □

Lemma 2.2.7: *It holds that* $\int_M \phi_- \, \mathrm{d}\mu = L$.

Proof. Suppose for a while that the sequence $(\phi_n/n)_{n \in \mathbb{N}_+}$ is uniformly bounded from below, that is: exists $K > 0$ s.t. $\phi_n/n \geq -K$ for every $n \in \mathbb{N}_+$. Using the Fatou Lemma to the sequence of non-negative functions $(\phi_n/n + K)_{n \in \mathbb{N}_+}$ we get

$$\int_M \phi_- \, \mathrm{d}\mu + K \leq \liminf_{n \to \infty} \int_M \phi_n/n \, \mathrm{d}\mu + K = L + K.$$

So $\phi_- \in L^1(\mu)$ and $\int_M \phi_- \,\mathrm{d}\mu \leq L$. To prove the opposite inequality observe that the previous Lemma implies

$$\frac{1}{n}\int_M \phi_n \,\mathrm{d}\mu \leq \frac{n-k}{n}\int_M \psi_k \,\mathrm{d}\mu + \frac{k}{n}\int_M \max\{\psi_k, \phi_1\}\,\mathrm{d}\mu$$

and observe now that $\max\{\psi_k, \phi_1\} \leq \max\{\phi_- + \varepsilon, \phi_1^+\} \in L^1(\mu)$. Hence taking the $\limsup_{n\to\infty}$ we obtain

$$L \leq \int_M \psi_k \,\mathrm{d}\mu \ \ \forall k \in \mathbb{N}_+$$

so taking the limit for $k \to \infty$ and using the monotone convergence theorem (applied to the sequence $(\max\{\phi_- + \varepsilon, \phi_1^+\} - \psi_k)_{n\in\mathbb{N}_+}$) we get

$$L \leq \lim_{n\to\infty}\int_M \psi_k \,\mathrm{d}\mu = \int_M \phi_- \,\mathrm{d}\mu + \varepsilon$$

and for $\varepsilon \to 0^+$ we obtain the thesis in the case of uniform boundness from below of $(\phi_n/n)_{n\in\mathbb{N}_+}$. We are left to remove this hypothesis. For $K > 0$ define

$$\phi_n^K = \max\{\phi_n, -kn\}, \ \ \phi_-^K = \max\{\phi_-, -K\}.$$

The sequence $(\phi_n^K)_{n\in\mathbb{N}_+}$ satisfies all the conditions of Kingman ergodic Theorem (it is subadditive and $(\phi_1^K)^+ \in L^1(\mu)$). Moreover it is clear that $\phi_-^K = \liminf_{n\to\infty}\phi_n^K/n$. So from the same argument used previously we get

$$\int_M \phi_n \,\mathrm{d}\mu = \inf_{K>0}\int_m \phi_n^K \,\mathrm{d}\mu.$$

By the monotone convergence theorem (applied to $(\phi_n^+ - \phi_n^K)_{K>0}$ and $(\phi_-^+ - \phi_-^K)_{K>0}$) we get

$$\int_M \phi_n \,\mathrm{d}\mu = \inf_{K>0}\int_M \phi_n^K \,\mathrm{d}\mu, \ \ \int_M \phi_- \,\mathrm{d}\mu = \inf_{K>0}\int_M \phi_-^K \,\mathrm{d}\mu$$

as a consequence we have

$$\begin{aligned}\int_M \phi_- \,\mathrm{d}\mu &= \inf_{K>0}\int_M \phi_-^K \,\mathrm{d}\mu \\ &= \inf_{K>0}\inf_{n\in\mathbb{N}_+}\int_M \phi_m^K \,\mathrm{d}\mu = \inf_{n\in\mathbb{N}_+}\frac{1}{n}\int_M \phi_n \,\mathrm{d}\mu = L.\end{aligned}$$

□

Lemma 2.2.8: *If $\psi \in L^1(\mu)$ then*

$$\frac{\psi \circ T^n}{n} \to 0 \ \ \mu\text{-a.e. on } M \text{ for } n \to \infty.$$

Proof. Fix $\varepsilon > 0$. Since μ is invariant we have that

$$\mu(|\phi \circ T^n| \geq n\varepsilon) = \mu(|\phi| \geq n\varepsilon) = \sum_{k\geq n}\mu(k \leq |\phi|/\varepsilon < k+1)$$

then adding over $n \in \mathbb{N}_+$ we obtain

$$\sum_{n\in\mathbb{N}_+}\mu(|\phi \circ T^n| \geq n\varepsilon) = \sum_{k\in\mathbb{N}_+} k\mu(k \leq |\phi|/\varepsilon < k+1) \leq \int_M \frac{|\phi|}{\varepsilon}\,\mathrm{d}\mu < +\infty$$

because $\phi \in L^1(\mu)$. Then for the Borel-Cantelli lemma we get

$$\mu(|\phi \circ T^n| \geq n\varepsilon \text{ for infinitely many } n \in \mathbb{N}_+) = 0$$

. Hence for μ-a.e. $x \in M$ there exists $P_x \in \mathbb{N}_+$ s.t. for every $n \geq P_x$ holds $|\phi(T^n(x))| < n\varepsilon$. Choosing $\varepsilon = 1/i$, $i \in \mathbb{N}_+$, we can find a $B_i \in \mathcal{B}$ s.t. $\mu(B_i) = 0$ and $|\phi \circ T^n| < n/i$ on B_i^c. Then, if $B = \bigcup_{i \in \mathbb{N}_+} B_i$, holds $\mu(B) = 0$ and

$$\lim_{n\to\infty} \frac{1}{n} |\phi \circ T^n| = 0 \text{ on } B^c.$$

□

Lemma 2.2.9: *For any $k \in \mathbb{N}_+$ holds*

$$\limsup_{n\to\infty} \frac{\phi_{kn}}{n} = k \limsup_{n\to\infty} \frac{\phi_n}{n}.$$

Proof. Obviously $\limsup_{n\to\infty} \phi_{kn}/n \leq k \limsup_{n\to\infty} \phi_n/n$ since $(\phi_{kn}/n)_{n\in\mathbb{N}_+}$ is a subsequence of $(\phi_n/n)_{n\in\mathbb{N}_+}$. To prove the opposite inequality let us write $n = kq_n + r_n$ with $r_n \in \{1, \ldots, k\}$. By subadditivity

$$\phi_n \leq \phi_{kq_n} + \phi_{r_n} \circ T^{kq_n} \leq \phi_{kq_n} + \psi \circ T^{kq_n} \tag{2.3}$$

with $\psi = \max\{\phi_1^+, \ldots, \phi_k^+\}$. Observe now that $n/q_n \to k$ for $n \to \infty$ (since $r_n \leq k$). Moreover by the fact that $\psi \in L^1(\mu)$, using the previous Lemma we get

$$\frac{\psi \circ T^n}{n} \to 0 \ \ \mu\text{-a.e. on } M \text{ for } n \to \infty.$$

Hence dividing by n in (2.3) and taking the $\limsup_{n\to\infty}$ we obtain

$$\begin{aligned}
\limsup_{n\to\infty} \frac{\phi_n}{n} &\leq \limsup_{n\to\infty} \frac{\phi_{kq_n}}{n} + \limsup_{n\to\infty} \frac{\psi \circ T^{kq_n}}{n} \\
&= \limsup_{n\to\infty} \frac{kq_n}{n} \frac{\phi_{kq_n}}{kq_n} + \underbrace{\limsup_{n\to\infty} \frac{kq_n}{n} \frac{\psi \circ T^{kq_n}}{kq_n}}_{=0} \\
&= \frac{1}{k} \limsup_{n\to\infty} \underbrace{\frac{kq_n}{n}}_{\to 1} \frac{\phi_{kq_n}}{q_n} \\
&= \frac{1}{k} \limsup_{q\to\infty} \frac{\phi_{kq}}{q}.
\end{aligned}$$

where we used that $\limsup_{n\to\infty} \frac{\phi_{kq_n}}{q_n} = \limsup_{q\to\infty} \frac{\phi_{kq}}{q}$ (since $q_n \to \infty$ for $n \to \infty$). □

Lemma 2.2.10: *Suppose that* $\inf_{x\in M} \phi_n(x) > -\infty$ *for every $n \in \mathbb{N}_+$. Then* $\int_M \phi_+ \, \mathrm{d}\mu \leq L$.

Proof. Fix $k \in \mathbb{N}_+$. For each $n \in \mathbb{N}_+$ consider $\theta_n = -\sum_{j=0}^{n-1} \phi_k \circ T^{jk}$. Observe that since μ is invariant holds

$$\int_M \theta_n \, \mathrm{d}\mu = -n \int_M \phi_k \, \mathrm{d}\mu \ \ \forall k \in \mathbb{N}_+. \tag{2.4}$$

Since the sequence $(\phi_n)_{n\in\mathbb{N}_+}$ is subadditive (note that $\phi_{n+m} \leq \phi_n + \phi_m \circ T^n$ implies $-\phi_m \circ T^n \leq \phi_n - \phi_{n+m}$) we have

$$\theta_n \leq \sum_{j=0}^{n-1} (\phi_{jk} - \phi_{(j+1)k}) = -\phi_{nk} \ \ \forall n \in \mathbb{N}_+.$$

Hence using the previous Lemma we obtain

$$\theta_- = \liminf_{n\to\infty} \frac{\theta_n}{n} \leq -\limsup_{n\to\infty} \frac{\phi_{nk}}{n} = -k \limsup \frac{\phi_n}{n} = -k\phi_+$$

so

$$\int_M \theta_- \, \mathrm{d}\mu \leq -k \int_m \phi_+ \, \mathrm{d}\mu. \tag{2.5}$$

Observe also that the sequence $(\theta_n)_{n\in\mathbb{N}_+}$ satisfies $\theta_{n+m} = \theta_m + \theta_n \circ T^{mk}$ for every $m, n \in \mathbb{N}_+$ (that in particular implies that the sequence of functions $(\theta_n)_{n\in\mathbb{N}_+}$ is subadditive w.r.t. T^k that obviously preserves μ). Since $\theta_1 = -\phi_k$ is bounded form above by $-\inf_{k\in\mathbb{N}_+} \phi_k$ we also have that θ_1^+ is bounded and in particular integrable. Hence we may use the Lemma 2.2.7, together with (2.4), to obtain

$$\int_M \theta_- \, \mathrm{d}\mu = \inf_{n\in\mathbb{N}_+} \int_M \frac{\theta_n}{n} \, \mathrm{d}\mu = -\int_M \phi_k \, \mathrm{d}\mu, \tag{2.6}$$

hence, putting together (2.5) and (2.6) we get that

$$\int_M \phi_+ \, \mathrm{d}\mu \leq \frac{1}{k} \int_M \phi_k \, \mathrm{d}\mu$$

and taking the $\inf_{k\in\mathbb{N}_+}$, we obtain the thesis. □

Proof of the Kingman subadditive ergodic Theorem 2.2.2. By the Lemmas 2.2.7 and 2.2.10 we get that if $inf_{x\in M} \phi_n(x) > -\infty$ for every $n \in \mathbb{N}_+$, then

$$\int_M \phi_- \, \mathrm{d}\mu = L \geq \int_M \phi_+ \, \mathrm{d}\mu \geq \int_M \phi_- \, \mathrm{d}\mu$$

that implies

$$\int_M \phi_+ \, \mathrm{d}\mu = \int_m \phi_- \, \mathrm{d}\mu = L.$$

Hence for what we said in the Remark 2.2.5 the theorem is proved in the assumption of boundness form below. In the general case consider for every $n \in \mathbb{N}_+$ and every $K > 0$

$$\phi_n^K = \max\{\phi_n, -Kn\}, \;\; \phi_-^K = \max\{\phi_-, -K\}, \;\; \phi_-^K = \max\{\phi_+, -K\}.$$

Then the previous argument applies to the sequence $(\phi_n^K)_{n\in\mathbb{N}_+}$ for each $K > 0$, hence $\phi_-^K = \phi_+^K$ μ-a.e. on M for every $K > 0$. But $\phi_+^K \to \phi_+$ and $\phi_-^K \to \phi_-$ for $K \to +\infty$, so $\phi_- = \phi_+$ μ-a.e. on M and the thesis follows like in the previous case. □

2.3 Birkhoff ergodic theorem

Again let us fix a measurable space $(M, \mathcal{B})$, a measurable transformation T and μ a T-invariant probability measure on M $\mathscr{A}$.

Lemma 2.3.1: *For any $\phi \in L^1(\mu)$ there exists $\tilde{\phi} \in L^1(\mu)$ invariant s.t.*

$$\frac{1}{n} \sum_{j=0}^{n-1} \phi \circ T^j \to \tilde{\phi} \;\; \textit{in } L^1(\mu) \textit{ for } n \to \infty.$$

Proof. By Von Neumann ergodic Theorem 2.1.5 we know that for $\psi \in L^\infty(\mu) \subset L^2(\mu)$ exists $\psi' \in L^2(\mu)$ s.t.

$$\frac{1}{n}\sum_{j=0}^{n-1} \psi \circ T^j \to \psi' \text{ in } L^2(\mu) \text{ for } n \to \infty.$$

We claim that $\psi' \in L^1(\mu)$. Indeed for every $n \in \mathbb{N}_+$ holds

$$\left\| \frac{1}{n}\sum_{j=0}^{n-1} \psi \circ T^j \right\|_{L^1(\mu)} \leq \|\psi\|_{L^\infty(\mu)}$$

and, since we can bound the $L^1(\mu)$-norm with the $L^2(\mu)$-norm (because μ is a probability), we also have

$$\frac{1}{n}\sum_{j=0}^{n-1} \psi \circ T^j \to \psi' \text{ in } L^1(\mu) \text{ for } n \to \infty$$

hence we get

$$\|\psi'\|_{L^1(\mu)} \leq \|\psi\|_{L^\infty(\mu)} < +\infty.$$

So the thesis holds for functions in $L^\infty(\mu)$. Now take $\phi \in L^1(\mu)$ and remember that for every $\varepsilon > 0$ there exists a function $\psi \in L^\infty(\mu)$ s.t. $\|\phi - \psi\|_{L^1(\mu)} \leq \varepsilon$. By averaging

$$\left\| \frac{1}{n}\sum_{j=0}^{n-1} \phi \circ T^j - \frac{1}{n}\sum_{j=0}^{n-1} \psi \circ T^j \right\|_{L^1(\mu)} \leq \varepsilon$$

since T preserves μ. Moreover by what we proved previously there exists $\psi' \in L^1(\mu)$ and $n_0 \in \mathbb{N}_+$ s.t.

$$\left\| \frac{1}{n}\sum_{j=0}^{n-1} \psi \circ T^j - \psi' \right\|_{L^1(\mu)} \leq \varepsilon \;\; \forall n \in \mathbb{N}_+, \;\; n \geq n_0.$$

As a consequence

$$\left\| \frac{1}{n}\sum_{j=0}^{n-1} \phi \circ T^j - \frac{1}{n'}\sum_{j=0}^{n'-1} \phi \circ T^j \right\|_{L^1(\mu)} \leq 4\varepsilon \;\; \forall n, n' \in \mathbb{N}_+, \;\; n, n' \geq n_0,$$

hence $(n^{-1}\sum_{j=0}^{n-1} \phi \circ T^j)_{n\in\mathbb{N}_+}$ is a Cauchy sequence in $L^1(\mu)$, so it converges in $L^1(\mu)$. Moreover for every $n \in \mathbb{N}_+$ holds

$$\left\| \left(\frac{1}{n}\sum_{j=0}^{n-1} \phi \circ T^j\right) \circ T - \frac{1}{n}\sum_{j=0}^{n-1} \phi \circ T^j \right\|_{L^1(\mu)} \leq \frac{2}{n}\left\| \sum_{j=0}^{n-1} \phi \circ T^j \right\|_{L^1(\mu)} = \frac{2}{n}\|\phi\|_{L^1(\mu)}$$

hence the limit function must be invariant. □

Theorem 2.3.2 (Birkhoff ergodic theorem): *Let $\phi \in L^1(\mu)$. There exists $\tilde{\phi} \in L^1(\mu)$ invariant, called* time average *of ϕ w.r.t. T (or only* time average *of ϕ), s.t.*

$$\frac{1}{n}\sum_{j=0}^{n-1} \phi \circ T^j \to \tilde{\phi} \;\; \mu\text{-a.e. on } M \text{ and in } L^1(\mu).$$

In particular holds the equality

$$\int_M \tilde{\phi}\, \mathrm{d}\mu = \int_M \phi\, \mathrm{d}\mu.$$

Proof. Define for avery $n \in \mathbb{N}_+$ the functions

$$\phi_n = \frac{1}{n}\sum_{j=0}^{n-1} \phi \circ T^j,$$

we have that $\phi_{n+m} = \phi_m + \phi_n \circ T^m$, in particular the sequence $(\phi_n)_{n\in\mathbb{N}_+}$ is subadditive and $\phi_1 = \phi \in L^1(\mu)$, so also $\phi_1^+ \in L^1(\mu)$. Hence the Kingman subadditive ergodic Theorem 2.2.2 applies and we obtain that $(\phi_n/n)_{n\in\mathbb{N}_+}$ converges to some $\tilde{\phi} : M \to \mathbb{R}$ measurable, invariant and with $\tilde{\phi}^+ \in L^1(\mu)$ s.t.

$$\int_M \tilde{\phi}\,\mathrm{d}\mu = \inf_{n\in\mathbb{N}_+} \frac{1}{n}\int_M \phi_n\,\mathrm{d}\mu = \int_M \phi\,\mathrm{d}\mu,$$

in particular $\tilde{\phi} \in L^1(\mu)$. The convergence is also in $L^1(\mu)$ for the previous Lemma. □

Corollary 2.3.3: *Let $E \in \mathcal{B}$. The function*

$$\tau(E,x) = \lim_{n\to\infty} \frac{1}{n}\#\left\{j \in \{0,1,\dots,n-1\} \mid T^j(x) \in E\right\},$$

called mean sojourn time *in , is well defined for μ-almost all $x \in M$. Moreover $\tau(E,\cdot) \in L^1(\mu)$ and*

$$\int_M \tau(E,x)\,\mathrm{d}\mu(x) = \mu(E).$$

Proof. The thesis follows from Birkhoff ergodic Theorem 2.3.2 observing that for every $x \in M$ holds

$$\lim_{n\to\infty} \frac{1}{n}\#\left\{j \in \{0,1,\dots,n-1\} \mid T^j(x) \in E\right\} = \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \mathbb{1}_E(T^j(x)).$$

□

Now we give an alternative proof of Von Neumann ergodic Theorem 2.1.5 that is indipendent by the functional ergodic Theorem 2.1.1 and uses only the Birkhoff ergodic Theorem 2.3.2.

Proof of Theorem 2.1.5 using Theorem 2.3.2. Consider any $p \in (1,\infty)$ and $\phi \in L^p(\mu)$. Let $\tilde{\phi}$ be the time average of ϕ, then $\tilde{\phi} \in L^p(\mu)$. Indeed

$$|\tilde{\phi}| \le \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} |\phi \circ T^j|$$

so holds

$$|\tilde{\phi}|^p \le \lim_{n\to\infty} \left[\frac{1}{n}\sum_{j=0}^{n-1} |\phi \circ T^j|\right]$$

and then using Fatou Lemma

$$\begin{aligned}\left[\int_M |\tilde{\phi}|^p\,\mathrm{d}\mu\right]^{1/p} &\le \liminf_{n\to\infty} \left[\int_M \left[\frac{1}{n}\sum_{j=0}^{n-1} |\phi \circ T^j|^p\right]\mathrm{d}\mu\right]^{1/p} \\ &= \liminf_{n\to\infty} \frac{1}{n}\left\|\sum_{j=0}^{n-1} |\phi \circ T^j|\right\|_{L^p(\mu)}\end{aligned}$$

hence using Minkowski inequality and the fact that μ is invariant, we get that

$$\begin{aligned}\|\tilde{\phi}\|_{L^p(\mu)} &= \left[\int_M |\tilde{\phi}|^p \,\mathrm{d}\mu\right]^{1/p} \\ &\leq \liminf_{n\to\infty} \frac{1}{n}\left\|\sum_{j=0}^{n-1} |\phi\circ T^j|\right\|_{L^p(\mu)} \leq \|\phi\|_{L^p(\mu)} < +\infty.\end{aligned}$$

Now let us show that

$$\frac{1}{n}\sum_{j=0}^{n-1} \phi\circ T^j \to \tilde{\phi} \ \text{ in } L^p(\mu).$$

Suppose first ϕ bounded, hence exists a $C > 0$ s.t. $|\phi| \leq C$, then $\left|\frac{1}{n}\sum_{j=0}^{n-1}\phi\circ T^j\right| \leq C$ for evry $n \in \mathbb{N}_+$ and $|\tilde{\phi}| \leq C$. Hence we may use the dominated convergence Theorem to obtain

$$\lim_{n\to\infty}\int_M \left|\frac{1}{n}\sum_{j=0}^{n-1}\phi\circ T^j - \tilde{\phi}\right|^p \mathrm{d}\mu = 0$$

that is the desired convergence. Now let us drop the boundness assumption in ϕ. Consider any sequence of bounded measurable functions $(\phi_k)_{k\in\mathbb{N}_+}$ converging to ϕ in $L^p(\mu)$ and denote $\tilde{\phi_k}$ the time average of ϕ_k for every $k \in \mathbb{N}_+$. Let us write

$$\left\|\frac{1}{n}\sum_{j=0}^{n-1}\phi\circ T^j - \tilde{\phi}\right\|_{L^p(\mu)} \leq \left\|\frac{1}{n}\sum_{j=0}^{n-1}(\phi-\phi_k)\circ T^j\right\|_{L^p(\mu)} + \left\|\frac{1}{n}\sum_{j=0}^{n-1}\phi_k\circ T^j - \tilde{\phi}_k\right\|_{L^p(\mu)} + \|\tilde{\phi}_k - \tilde{\phi}\|_{L^p(\mu)}$$

and notice that exists $k_0 \in \mathbb{N}_+$ s.t. $\|\phi-\phi_k\|_{L^p(\mu)} \leq \varepsilon/3$ for every $k \geq k_0$, that implies

$$\|(\phi-\phi_k)\circ T^j\|_{L^p(\mu)} \leq \varepsilon/3 \ \ \forall n\in\mathbb{N}_+ \ \ \forall k\in\mathbb{N}_+, \ \ k\geq k_0$$

since μ is invariant, thus

$$\left\|\frac{1}{n}\sum_{j=0}^{n-1}(\phi-\phi_k)\circ T^j\right\|_{L^p(\mu)} \leq \|\phi-\phi_k\|_{L^p(\mu)} \leq \frac{\varepsilon}{3} \ \ \forall n\in\mathbb{N}_+ \ \ \forall k\in\mathbb{N}_+, \ \ k\geq k_0.$$

Furthermore using Fatou Lemma in the previous inequality gives

$$\|\tilde{\phi}-\tilde{\phi}_k\|_{L^p(\mu)} \leq \|\phi-\phi_k\|_{L^p(\mu)} \leq \frac{\varepsilon}{3} \ \ \forall k\in\mathbb{N}_+, \ \ k\geq k_0$$

in fact $\widetilde{\phi-\phi_k} = \tilde{\phi}-\tilde{\phi}_k$. By the previous case for every $k\in\mathbb{N}_+$ there exists $n_0(k)\in\mathbb{N}_+$ s.t.

$$\left\|\frac{1}{n}\sum_{j=0}^{n-1}\phi_k\circ T^j - \tilde{\phi}_k\right\|_{L^p(\mu)} \leq \frac{\varepsilon}{3} \ \ \forall n\in\mathbb{N}_+, \ \ n\geq n_0(k)$$

hence

$$\left\|\frac{1}{n}\sum_{j=0}^{n-1}\phi\circ T^j - \tilde{\phi}\right\|_{L^p(\mu)} \leq \varepsilon \ \ \forall n\in\mathbb{N}_+, \ \ n\geq n_0(k_0).$$

□

2.4 Ergodic theorem for flows

As before we fix a measurable space $(M, \mathcal{B})$, a flow $\{T^t\}_{t\in\mathbb{R}}$ and μ an invariant probability measure for the flow on M.

Theorem 2.4.1 (Ergodic theorem for flows): *Let $\phi \in L^1(\mu)$. There exists $\tilde{\phi} \in L^1(\mu)$ invariant s.t.*

$$\frac{1}{\tau}\int_0^\tau \phi \circ T^t \,\mathrm{d}t \to \tilde{\phi} \;\; \mu\text{-a.e. on } M \text{ and in } L^1(\mu) \text{ for } \tau \to \infty.$$

Moreover if $\phi \in L^p(\mu)$ for some $p \in [1, \infty)$ then $\tilde{\phi} \in L^p(\mu)$ and the limit is also in $L^p(\mu)$.

Proof. We have

$$\int_0^\tau \phi \circ T^t \,\mathrm{d}t + \int_\tau^{\lfloor\tau\rfloor+1} \phi \circ T^t \,\mathrm{d}t = \sum_{k=0}^{\lfloor\tau\rfloor} \int_k^{k+1} \phi \circ T^t \,\mathrm{d}t$$

and

$$\begin{aligned}\int_k^{k+1} \phi \circ T^t \,\mathrm{d}t &= \int_0^1 \phi \circ T^{t+k} \,\mathrm{d}t \\ &= \int_0^1 \phi \circ T^t \circ T^k \,\mathrm{d}t = \Phi \circ T^k\end{aligned}$$

where $\Phi = \int_0^1 \phi \circ T^t \,\mathrm{d}t$. Now Φ is obviously measurable and if $\phi \in L^p(\mu)$ then $\Phi \in L^p(\mu)$ (by Minkowski's inequality in integral form). Moreover if $\phi \in L^\infty(\mu)$ then

$$\frac{1}{\tau}\int_\tau^{\lfloor\tau\rfloor+1} \phi \circ T^t \,\mathrm{d}t \to 0 \;\; \text{uniformly for } \tau \to \infty.$$

Further

$$\frac{1}{\tau}\int_0^\tau \phi \circ T^t \,\mathrm{d}t = \frac{\lfloor\tau\rfloor}{\tau}\frac{1}{\lfloor\tau\rfloor}\sum_{k=0}^{\lfloor\tau\rfloor} \Phi \circ (T^1)^k - \frac{1}{\tau}\int_\tau^{\lfloor\tau\rfloor+1} \phi \circ T^t \,\mathrm{d}t$$

This, with Birkhoff ergodic Theorem 2.3.2 and Von Neumann ergodic Theorem 2.1.5 (used with Φ), holds that if $\phi \in L^\infty(\mu)$, then $\frac{1}{\tau}\int_0^\tau \phi \circ T^t \,\mathrm{d}t$ converges μ-a.e. and in $L^p(\mu)$ for every $p \in [1, \infty)$. Now, Minkowski inequality in integral form and the invariance of μ yield that for every $\phi_1, \phi_2 \in L^p(\mu)$

$$\left\|\frac{1}{\tau}\int_0^\tau \phi_1 \circ T^t \,\mathrm{d}t - \frac{1}{\tau}\int_0^\tau \phi_2 \circ T^t \,\mathrm{d}t\right\|_{L^p(\mu)} \le \frac{1}{\tau}\int_0^\tau \|\phi_1 - \phi_2\|_{L^p(\mu)} \,\mathrm{d}t < \varepsilon$$

whenever $\|\phi_1 - \phi_2\|_{L^p(\mu)} < \varepsilon$ and since we can always approximate an $L^p(\mu)$ function with a sequence in $L^\infty(\mu)$ we obtain that $\frac{1}{\tau}\int_0^\tau \phi \circ T^t \,\mathrm{d}t$ converge in $L^p(\mu)$ if $\phi \in L^p(\mu)$ with $p \in [1, \infty)$. Moreover if $\phi \in L^1(\mu)$, using Lemma 2.2.8 we have

$$\begin{aligned}\left|\frac{1}{\tau}\int_\tau^{\lfloor\tau\rfloor+1} \phi \circ T^t \,\mathrm{d}t\right| &\le \frac{1}{\lfloor\tau\rfloor}\int_{\lfloor\tau\rfloor}^{\lfloor\tau\rfloor+1} |\phi| \circ T^t \,\mathrm{d}t \\ &= \frac{1}{\lfloor\tau\rfloor}\int_0^1 |\phi| \circ T^t \circ T^{\lfloor\tau\rfloor} \,\mathrm{d}t \\ &= \frac{1}{\lfloor\tau\rfloor}\psi \circ T^{\lfloor\tau\rfloor} \to 0 \;\; \mu\text{-a.e. for } \tau \to \infty\end{aligned}$$

where $\psi = \int_0^1 |\phi| \circ T^t \,\mathrm{d}t$, hence we have the complete thesis. □

3

Ergodicity and Mixing

In the rest of this chapter, if nothing is specified, (M, d) will always be a metric space, $\mathcal{B}$ the Borel σ-algebra of M and $T : M \to M$ a measurable transformation.

3.1 Ergodic systems

Definition 3.1.1: We say that the system (T, μ), where μ is a σ-finite Borel measure on M, is *ergodic* (or that μ *is ergodic w.r.t.* T or that T *is ergodic w.r.t.* μ) if for every T-invariant $A \in \mathcal{B}$ holds that $\min\{\mu(A), \mu(A^c)\} = 0$.

Ther are many useful ways to reformulate ergodicity.

Proposition 3.1.2: *Let μ be an invariant σ-finite Borel measure on M. The following are equivalent:*

(1) Every invariant $\psi \in L^1(\mu)$ is constant μ-a.e. on M.

(2) For every invariant $A \in \mathcal{B}$ holds that $\min\{\mu(A), \mu(A^c)\} = 0$.

Moreover if μ is a probability measure the above are equivalent also to the following:

(3) For every $B \in \mathcal{B}$ we have $\tau(B, x) = \mu(B)$ for μ-a.e. $x \in M$.

(4) For every $B \in \mathcal{B}$ the function $\tau(B, \cdot)$ is constant μ-a.e. on M.

(5) For every invariant $\psi \in L^1(\mu)$ holds $\psi(x) = \int_M \psi \, \mathrm{d}\mu$ for μ-a.e. $x \in M$.

(6) For every $\phi \in L^1(\mu)$ holds $\tilde{\phi}(x) = \int_M \phi \, \mathrm{d}\mu$ for μ-a.e. $x \in M$.

(7) For every $\phi \in L^1(\mu)$ the time average $\tilde{\phi}$ is constant μ-a.e. on M.

Proof. First we prove the first part with μ σ-finite.

- *(1) implies (2).* Let A be an invariant set. Obviously either A or A^c has finite measure, suppose without lost of generality that A has finite measure. Then $\mathbb{1}_A$ is an invariant integrable function, hence by hypothesis it is constant μ-a.e. on M. It follows that $\mu(A) = \mu(M)$, so that $\mu(A^c) = 0$.

- *(2) implies (1).* Let $\psi \in L^1(\mu)$ be invariant, then every level set $B_c = \{\psi \leq c\}$ is an invariant set. Hence by the hypothesis we have that $\min\{\mu(B_c), \mu(B_c^c)\} = 0$ for every $c \in \mathbb{R}$. Now observe that $c \mapsto \mu(B_c)$ is non-decreasing, hence there exists a $c_0 \in \mathbb{R}$ s.t. $\mu(B_c) = 0$ for every $c < c_0$ and $\mu(B_c) = \mu(M)$ for every $c \geq c_0$. Then $\psi = c_0$ μ-a.e. on M.

Now suppose μ a probability measure. It is obvious that (3) implies (4), that (6) implies (7) and that (5) implies (1). Since the time average is an invariant function, we easily see that (1) implies (4) and that (5) implies (6). Moreover, since the mean sojourn time is itself a time average (of the the respective indicator function), we have also that (6) implies (3) and that (7) implies (4).

Now suppose that (4) holds. Let A be an invariant set. Then $\tau(A, x) = \mathbb{1}_A(x)$ for μ-a.e. $x \in M$. Since $\tau(A, \cdot)$ is by hypothesis constant μ-a.e. on M it follows that $\min\{\mu(A), \mu(A^c)\} = 0$, that is (2). Hence (4) implies (2).

Finally note that, using the same argument used in the second point of the first part, one can see that (2) implies (5), indeed if $\psi \in L^1(\mu)$, then we obtain $\psi = c_0 \in \mathbb{R}$ μ-a.e. on M and since μ is a probability measure it follows that actually $\psi = c_0 = \int_M \psi \, d\mu$ μ-a.e. on M. □

Now fix a $\mu \in \mathcal{M}_1^T(M)$, we are going to see that in this case we can characterize the ergodicity property also using the T-Koopman operator L_T.

Proposition 3.1.3: *The following are equivalents:*

(1) (T, μ) *is ergodic.*

(2) For any $\phi \in L^p(\mu)$ *and* $\psi \in L^q(\mu)$*, with* p, q *conjugate exponents, holds*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \int_M (L_T^j \phi)\psi \, d\mu = \left(\int_M \phi \, d\mu\right)\left(\int_M \psi \, d\mu\right). \tag{3.1}$$

(3) For any $A, B \in \mathcal{B}$ *holds*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu(T^{-j}(A) \cap B) = \mu(A)\mu(B). \tag{3.2}$$

Proof. It is clear that (2) implies (3), simply taking $\phi = \mathbb{1}_A$ and $\psi = \mathbb{1}_B$. Next let us show that (3) implies (1). Let $A \in \mathcal{B}$ be invariant and set $B = A$ in (3), then

$$\mu(A) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu(T^{-j}(A) \cap A) = \mu(A)^2,$$

hence $\mu(A) \in \{0, 1\}$.

Now, to conclude the proof, is sufficient to prove that (1) implies (2). Consider $\phi \in L^p(\mu)$ and $\psi \in L^q(\mu)$, by (1) and the Birkhoof ergodic Theorem 2.3.2 we get that

$$\frac{1}{n} \sum_{j=0}^{n-1} L_T^j \phi \to \int_M \phi \, d\mu \quad \mu\text{-a.e. on } M.$$

Assume at first that $|\phi| \leq k$ for some $k \in \mathbb{N}_+$. Then for every $n \in \mathbb{N}$

$$\left| \left(\frac{1}{n} \sum_{j=0}^{n-1} L_T^j \phi \right) \psi \right| \leq k|\psi|.$$

Hence, since $k|\psi| \in L^1(\mu)$, we may use the dominated convergence theorem to get that

$$\int_M \left(\frac{1}{n}\sum_{j=0}^{n-1} L_T^j \phi\right) \psi \,\mathrm{d}\mu \to \left(\int_M \phi \,\mathrm{d}\mu\right)\left(\int_M \psi \,\mathrm{d}\mu\right).$$

This proves (2) when ϕ is bounded. Now we want to remove this assumption. Given any $\phi \in L^p(\mu)$ and $k \in \mathbb{N}_+$, define

$$\phi_k(x) = \begin{cases} k & \phi(x) > k \\ \phi(x) & \phi(x) \in [-k, k] \\ -k & \phi(x) < -k. \end{cases}$$

Fix $\varepsilon > 0$. Each ϕ_k is in bounded, hence the previous case gives us that for any $k \in \mathbb{N}_+$ there exists a $n_0(k) \in \mathbb{N}_+$ s.t.

$$\left|\int_M \left(\frac{1}{n}\sum_{j=0}^{n-1} L_T^j \phi_k\right) \psi \,\mathrm{d}\mu - \left(\int_M \phi_k \,\mathrm{d}\mu\right)\left(\int_M \psi \,\mathrm{d}\mu\right)\right| < \varepsilon$$

for $n \geq n_0(k)$. Next observe that for every $p \in [1, \infty]$ holds

$$\|\phi_k - \phi\|_{L^p(\mu)} \to 0 \text{ for } k \to \infty,$$

in fact for $p = \infty$ we have $\phi_k = \phi$ for $k > \|\phi\|_{L^\infty(\mu)}$ and for $p < \infty$ it follows from the dominated (by $2|\phi|^p \in L^1(\mu)$) convergence theorem. Hence, using the Hölder inequality, we get

$$\left|\left(\int_M (\phi_k - \phi) \,\mathrm{d}\mu\right)\left(\int_M \psi \,\mathrm{d}\mu\right)\right| \leq \|\phi_k - \phi\|_{L^p(\mu)} \left|\int_M \psi \,\mathrm{d}\mu\right| < \varepsilon$$

for $k \geq k_0$ for some $k_0 \in \mathbb{N}_+$. Similarly

$$\begin{aligned} \left|\int_M \frac{1}{n}\sum_{j=0}^{n-1} L_T^j(\phi_k - \phi)\psi \,\mathrm{d}\mu\right| &\leq \frac{1}{n}\sum_{j=0}^{n-1} \left|\int_M L_T^j(\phi_k - \phi)\psi \,\mathrm{d}\mu\right| \\ &\leq \frac{1}{n}\sum_{j=0}^{n-1} \left\|L_T^j(\phi_k - \phi)\right\|_{L^p(\mu)} \|\psi\|_{L^q(\mu)} \\ &= \|\phi_k - \phi\|_{L^p(\mu)} \|\psi\|_{L^q(\mu)} < \varepsilon \end{aligned}$$

for every $n \in \mathbb{N}_+$ and every $k \geq k_1$ for some $k_1 \in \mathbb{N}_+$. Hence taking $k \geq \max\{k_0, k_1\}$ and $n \geq n_0(k)$ we get

$$\left|\frac{1}{n}\sum_{j=0}^{n-1} \int_M (L_T^j \phi)\psi \,\mathrm{d}\mu - \left(\int_M \phi \,\mathrm{d}\mu\right)\left(\int_M \psi \,\mathrm{d}\mu\right)\right| < 3\varepsilon.$$

□

Corollary 3.1.4: *Suppose that the condition* (3.2) *holds for every $A, B \in \mathcal{S}$, with $\mathcal{S}$ a π-system that generates $\mathcal{B}$. Then (T, μ) is ergodic.*

Proof. It follows by the monotone class theorem. In fact, for each $B \in \mathcal{B}$, the family

$$\mathcal{F}_B = \{A \in \mathcal{B} \mid (3.2) \text{ holds with } A \text{ and } B\}$$

is a λ-system. To see this observe firstly that obviously $M \in \mathcal{F}_B$ and that if $A \in \mathcal{F}_B$, then

$$\begin{aligned}
\left|\frac{1}{n}\sum_{j=0}^{n-1}\mu(T^{-j}(A^c)\cap B)-\mu(A^c)\mu(B)\right| &= \left|\frac{1}{n}\sum_{j=0}^{n-1}\mu(T^{-j}(A)^c\cap B)-(1-\mu(A))\mu(B)\right| \\
&= \left|\frac{1}{n}\sum_{j=0}^{n-1}(\mu(B)-\mu(T^{-j}(A)\cap B))-(1-\mu(A))\mu(B)\right| \\
&= \left|\frac{1}{n}\sum_{j=0}^{n-1}\mu(T^{-j}(A)\cap B)-\mu(A)\mu(B)\right| \to 0
\end{aligned}$$

that is $A^c \in \mathcal{F}_B$. Now it remains to prove that if $(A_k)_{k\in\mathbb{N}} \subset \mathcal{F}_B$ and $A_k \nearrow A$, then $A \in B$. Let us write

$$\begin{aligned}
\left|\frac{1}{n}\sum_{j=0}^{n-1}\mu(T^{-j}(A)\cap B)-\mu(A)\mu(B)\right| \leq{}& \left|\frac{1}{n}\sum_{j=0}^{n-1}\mu(T^{-j}(A_k)\cap B)-\mu(A_k)\mu(B)\right| \\
&+\left|\frac{1}{n}\sum_{j=0}^{n-1}[\mu(T^{-j}(A)\cap B)-\mu(T^{-j}(A_k)\cap B)]\right| \\
&+|\mu(A_k)-\mu(A)|\mu(B)
\end{aligned}$$

and each of the three therms on the right-hand side is infitesimal for $k \to \infty$, hence $A \in \mathcal{F}_B$. So by the monotoone class theorem follows $\mathcal{F}_B = \mathcal{B}$ for every $B \in \mathcal{B}$, that is the thesis. □

Corollary 3.1.5: *Suppose that the condition* (3.1) *holds for every ϕ, ψ in dense subsets of $L^p(\mu)$ and $L^q(\mu)$ respecrively, for some $p, q \in [1,\infty]$. Then (T,μ) is ergodic.*

Proof. It is a simple approximation argument. □

3.2 Rokhlin disintegration and ergodic decomposition

Let us fix (M,d) a complete separable metric space, $\mathcal{B}$ the σ-algebra of its Borel sets, $T : M \to M$ a measurable transformation and μ a probability measure on M.

If $\mathcal{P}$ is a partition of measurable sets of M the function $\pi : M \to \mathcal{P}$ will always be the canonical projection that assign to each $x \in M$ the set $\mathcal{P}(x) = \pi(x) \in \mathcal{P}$ s.t. $x \in \pi(x)$. The map π endows $\mathcal{P}$ with a structure of probability space as follows:

- a subset $\mathcal{Q} \subset \mathcal{P}$ is measurable if and only if $\pi^{-1}(\mathcal{Q}) = \bigcup \mathcal{Q} \in \mathcal{B}$; it is not hard to prove that the family $\hat{\mathcal{B}}$ of such measurable subsets of $\mathcal{P}$ is a σ-algebra called *quotient σ-algebra*;
- then we define the *quotient measure* $\hat{\mu} : \hat{\mathcal{B}} \to [0,1]$ s.t. for every $\mathcal{Q} \in \hat{\mathcal{B}}$

$$\hat{\mu}(\mathcal{Q}) = \pi_*\mu(\mathcal{Q}) = \mu(\pi^{-1}(\mathcal{Q}))$$

In what follows, when we will deal with a partition P, $\hat{\mathcal{B}}$ and $\hat{\mu}$ will always be the quotient σ-algebra and the quotient measure respectively.

The statement of the main theorem of this section is the following.

Theorem 3.2.1 (Ergodic decomposition): *Let μ be invariant. There exists $M_0 \in \mathcal{B}$ with $\mu(M_0) = 1$, a partition of measurable sets $\mathcal{P}$ of M_0 and a family $\{\mu_P\}_{P\in\mathcal{P}}$ of probability measures on M s.t.*

(1) $\mu_P(P) = 1$ for $\hat{\mu}$-a.e. $P \in \mathcal{P}$;

(2) for every $E \in \mathcal{B}$ the map $P \mapsto \mu_P(E)$ is measurable w.r.t. $\hat{\mathcal{B}}$;

(3) μ_P is invariant and ergodic for $\hat{\mu}$-a.e. $P \in \mathcal{P}$;

(4) for every $E \in \mathcal{B}$ holds

$$\mu(E) = \int_{\mathcal{P}} \mu_P(E)\,\mathrm{d}\hat{\mu}(P).$$

In order to prove this theorem we are going to introduce an important result in measure theory, the Rokhlin disintegration.

Definition 3.2.2: A *disintegration* of μ w.r.t. a partition $\mathcal{P}$ of M is a family $\{\mu_P \mid P \in \mathcal{P}\}$ of probability measures on M s.t. for every $E \in \mathcal{B}$

(1) $\mu_P(P) = 1$ for $\hat{\mu}$-a.e. $P \in \mathcal{P}$;

(2) the map $\mathcal{P} \to \mathbb{R}$ s.t. $P \mapsto \mu_P(E)$ is measurable w.r.t. $\hat{\mathcal{B}}$;

(3) holds

$$\mu(E) = \int_{\mathcal{P}} \mu_P(E)\,\mathrm{d}\hat{\mu}(P).$$

The probability measures $\{\mu_P \mid P \in \mathcal{P}\}$ are called *conditional probabilities* for μ w.r.t. $\mathcal{P}$.

Lemma 3.2.3 (Uniqueness of the disintegration): *If $\{\mu_P\}_{P\in\mathcal{P}}$ and $\{\mu'_P\}_{P\in\mathcal{P}}$ are disintegrations of μ w.r.t. the same partition $\mathcal{P}$, then $\mu_P = \mu'_P$ for $\hat{\mu}$-a.e. $P \in \mathcal{P}$.*

Proof. (Optional) Let $\mathcal{U}$ be a countable basis of open sets of the topology of M and let $\mathcal{A}$ be the countable σ-algebra generated by $\mathcal{U}$. For each $A \in \mathcal{A}$ let us consider

$$\mathcal{Q}_A = \{P \in \mathcal{P} \mid \mu_P(A) > \mu'_P(A)\},$$
$$\mathcal{R}_A = \{P \in \mathcal{P} \mid \mu_P(A) < \mu'_P(A)\}.$$

Observe that if $P \in \mathcal{Q}_A$ then $P \subset \pi^{-1}(\mathcal{Q}_A) = \bigcup \mathcal{Q}_A$, hence using property (1) of the definition of disintegration we get

$$\mu_P(A \cap \pi^{-1}(\mathcal{Q}_A)) = \mu_P(A);$$

otherwise if $P \cap \pi^{-1}(\mathcal{Q}_A) = \emptyset$ we have $\mu_P(A \cap \pi^{-1}(\mathcal{Q}_A)) = 0$. THis fact remains obviously valid if we replace μ_P with μ'_P. Now using the property (3) of the definition of disinteration, we obtain

$$\mu(A \cap \pi^{-1}(\mathcal{Q}_A))) = \begin{cases} \int_{\mathcal{P}} \mu_P(A \cap \pi^{-1}(\mathcal{Q}_A))\,\mathrm{d}\hat{\mu}(P) = \int_{\mathcal{Q}_A} \mu_P(A)\,\mathrm{d}\hat{\mu}(P) \\ \int_{\mathcal{P}} \mu'_P(A \cap \pi^{-1}(\mathcal{Q}_A))\,\mathrm{d}\hat{\mu}(P) = \int_{\mathcal{Q}_A} \mu'_P(A)\,\mathrm{d}\hat{\mu}(P) \end{cases}$$

and since $\mu_P(A) > \mu'_P(A)$ for every $P \in \mathcal{Q}_A$, this implies $\hat{\mu}(\mathcal{Q}_A) = 0$ for every $A \in \mathcal{A}$. Analogously follows that $\hat{\mu}(\mathcal{R}_A) = 0$ for every $A \in \mathcal{A}$. Hence for $\mathcal{N} = \bigcup_{A\in\mathcal{A}}(\mathcal{Q}_A \cup \mathcal{R}_A)$ holds $\mu(\mathcal{N}) = 0$ and for every $P \in \mathcal{N}^c$ holds that $\mu_P(A) = \mu'_P(A)$ for every $A \in \mathcal{A}$, that implies $\mu_P = \mu'_P$ on the whole $\mathcal{B}$, since $\mathcal{A}$ generates $\mathcal{B}$. $\square$

If $\mathcal{P}, \mathcal{P}'$ are partitions we say that $\mathcal{P}$ is *coarser* than $\mathcal{P}'$ or that $\mathcal{P}'$ if *finer* then $\mathcal{P}$ and we write $\mathcal{P} \prec \mathcal{P}'$ if every element of $\mathcal{P}'$ is contaiced in some element of $\mathcal{P}$. Moreover if $n \in \mathbb{N}_+$ and $\mathcal{P}_1, \dots, \mathcal{P}_n$ are partitions we define

$$\bigvee_{j=1}^{n} \mathcal{P}_j = \left\{P_1 \cap \dots \cap P_n \mid P_j \in \mathcal{P}_j \ \ \forall j = 1, \dots, n\right\}$$

and if $\{\mathcal{P}_n\}_{n\in\mathbb{N}_+}$ is a sequence of partitions we define

$$\bigvee_{n\in\mathbb{N}_+} \mathcal{P}_n = \left\{\bigcap_{n\in\mathbb{N}_+} P_n \mid P_n \in \mathcal{P}_n \ \ \forall n \in \mathbb{N}_+\right\}.$$

Definition 3.2.4: We say that a partition of measurable sets $\mathcal{P}$ is a *measurable partition* if there exists $M_0 \in \mathcal{B}$ with $\mu(M_0) = 1$ and an increasing asequence of countable partitions

$$\mathcal{P}_1 \prec \mathcal{P}_2 \prec \dots \prec \mathcal{P}_n \prec \dots$$

s.t.

$$\mathcal{P} = \bigvee_{n\in\mathbb{N}_+} \mathcal{P}_n.$$

Remark 3.2.5: Every countable partition of measurable sets $\mathcal{P}$ of M is a measurable partition, indeed we can choose $M_0 = M$ and observe that

$$\mathcal{P} = \bigvee_{n\in\mathbb{N}_+} \mathcal{P}.$$

Remark 3.2.6: (Optional) Let us fix a measurable partition $\mathcal{P}$ and any increasing sequence $\mathcal{P}_1 \prec \mathcal{P}_2 \prec \dots \prec \mathcal{P}_n \prec \dots$ of countable partitions s.t. $\mathcal{P} = \bigvee_{n\in\mathbb{N}_+} \mathcal{P}_n$ restricted to some $M_0 \in \mathcal{B}$ with $\mu(M_0) = 1$ and since we can modify M_0 with zero measure sets, we can suppose that each $\mathcal{P}_n$ is generated by a countable family of measurable sets with positive measure. As said at the beginning of the section, for every $x \in M$ and every $n \in \mathbb{N}_+$, $\mathcal{P}_n(x)$ will be the element of $\mathcal{P}_n$ that contains x. For any $n \in \mathbb{N}_+$ and any $\psi \in L^1(\mu)$ let us consider the conditional expectation

$$\mathbf{E}[\psi \mid \mathcal{P}_n](x) = \frac{1}{\mu(\mathcal{P}_n(x))} \int_{\mathcal{P}_n(x)} \psi \, \mathrm{d}\mu.$$

The equality holds because a function is measurable with respect to $\mathcal{P}_n$ if and only if it is constant on every element of $\mathcal{P}_n$, so the right-hand side is measurable w.r.t. $\mathcal{P}_n$ and if $P' \in \mathcal{P}_n$ holds

$$\int_P \frac{1}{\mu(\mathcal{P}_n(x))} \int_{\mathcal{P}_n(x)} \psi \, \mathrm{d}\mu \, \mathrm{d}\mu(x) = \int_P \frac{1}{\mu(P)} \int_P \psi \, \mathrm{d}\mu \, \mathrm{d}\mu(x) = \int_P \psi \, \mathrm{d}\mu.$$

Observe that since $\mathcal{P}_n \prec \mathcal{P}_{n+1}$ for each $P \in \mathcal{P}_n$ there exists $\mathcal{Q}_P \subset \mathcal{P}_{n+1}$ s.t. $P = \bigcup \mathcal{Q}_P$. Hence if for every $n \in \mathbb{N}_+$ we denote with $\mathcal{F}_n$ the σ-algebra generated by $\mathcal{P}_n$ we have $\mathcal{P}_n \subset \mathcal{P}_{n+1}$ for every $n \in \mathbb{N}_+$, so $(\mathcal{F}_n)_{n\in\mathbb{N}_+}$ is a filtration. Moreover the sequence $(\mathbf{E}[\psi \mid \mathcal{P}_n])_{n\in\mathbb{N}_+}$ is a martingale with respect to the filtration $(\mathcal{F}_n)_{n\in\mathbb{N}_+}$ and it holds

$$|\mathbf{E}[\psi \mid \mathcal{P}_n]| \le |\psi| \in L^1(\mu)$$

so by the martingale convergence theorem follows that exists a $\psi_\infty \in L^1(\mu)$ s.t.

$$\mathbf{E}[\psi \mid \mathcal{P}_n] \to \psi_\infty \ \ \mu\text{-a.e. on } M \text{ and in } L^1(\mu),$$

hence, in particular

$$\int_M \psi \, \mathrm{d}\mu = \int_M \psi_\infty \, \mathrm{d}\mu.$$

We will call $M_\psi \in \mathcal{B}$ the full measure set where $\mathbf{E}[\psi \,|\, \mathcal{P}_n]$ converge to ψ_∞. Moreover we define $\psi_\infty = 0$ on $M \setminus M_\psi$

We are particularly interested in case of $\psi = \mathbb{1}_B$ with $B \in \mathcal{B}$. If $B \in \mathcal{B}$, we will use the notation $\mathcal{P}_B$ to refer to the subset of elements P of the partition $\mathcal{P}$ that intersect $M_{\mathbb{1}_B}$. Observe that

$$\hat{\mu}(\mathcal{P}_B) = \mu\left(\bigcup \mathcal{P}_B\right) \geq \mu(M_{\mathbb{1}_B}) = 1$$

so $\hat{\mu}(\mathcal{P}_B) = 1$, and that

$$(\mathbb{1}_B)_\infty(x) = \lim_{n\to\infty} \frac{\mu(B \cap \mathcal{P}_n(x))}{\mu(\mathcal{P}_n(x))} \quad \forall x \in M_{\mathbb{1}_B}.$$

Moreover, we define the function $E(B, \cdot) : \mathcal{P} \to \mathbb{R}$ by setting for all $P \in \mathcal{P}$

$$E(B, P) = \begin{cases} (\mathbb{1}_B)_\infty(x) \ \text{ for any } x \in M_{\mathbb{1}_B} \cap P & P \in \mathcal{P}_B \\ 0 & P \notin \mathcal{P}_B \end{cases}$$

(it is a well posed definition since a $(\mathbb{1}_B)_\infty$ is the limit of the functions $\mathbf{E}[\mathbb{1}_B \,|\, \mathcal{P}_n] = \mathbf{P}(B \,|\, \mathcal{P}_n)$, $n \in \mathbb{N}_+$ and for every $n \in \mathbb{N}$ the function $\mathbf{P}(B \,|\, \mathcal{P}_n)$ is constant on every element of the partition $\mathcal{P}_n$, but $\mathcal{P}_n$ is coarser than $\mathcal{P}$ so $\mathbf{P}(B \,|\, \mathcal{P}_n)$ is constant also on every element of $\mathcal{P}$). Note also that $(\mathbb{1}_B)_\infty = E(B, \cdot) \circ \pi$, hence $E(B, \cdot)$ is measurable with respect to $\hat{\mathcal{B}}$ and satisfies

$$\int_{\mathcal{P}} E(B, P) \, \mathrm{d}\hat{\mu}(P) = \int_{\mathcal{P}} E(B, P) \, \mathrm{d}(\pi_* \mu)(P) = \int_M (\mathbb{1}_B)_\infty \, \mathrm{d}\mu = \int_M (\mathbb{1}_B)_\infty \, \mathrm{d}\mu = \mu(B)$$

Theorem 3.2.7 (Rokhlin disintegration): *Let $\mathcal{P}$ be a measurable partition. Then the probability measure μ admits some disintegration w.r.t. $\mathcal{P}$.*

Proof. (Optional) Let $\mathcal{U} = \{U_k\}_{k\in\mathbb{N}}$ a countable basis of open sets for the topology of M (that is a separable metric space, hence N2) and let $\mathcal{A}$ be the ring generated by $\mathcal{U}$. Obviously $\mathcal{A}$ is countable and it generates $\mathcal{B}$. Recall the notations introduced in the previous Remark. Define

$$\mathcal{P}_* = \bigcap_{A\in\mathcal{A}} \mathcal{P}_A,$$

since $\hat{\mu}(\mathcal{P}_A) = 1$ for every $A \in \mathcal{A}$ and the intersection is countable, holds that $\hat{\mu}(\mathcal{P}_*) = 1$. For each $P \in \mathcal{P}_*$ define the function $\mu_P : \mathcal{A} \to [0, 1]$ s.t.

$$\mu_P(A) = E(A, P) \ \ \forall A \in \mathcal{A}.$$

In partiular $\mu_P(M) = E(M, P) = 1$ (by construction $E(M, \cdot) = 1$). It is also clear that $\mu_p(\emptyset) = 0$ and that μ_P is additive. Indeed if $A, B \in \mathcal{A}$ are s.t. $A \cap B = \emptyset$, we have

$$\begin{aligned} E(A \cup B, P) &= (\mathbb{1}_{A\cup B})_\infty(x) \\ &= \lim_{n\to\infty} \frac{\mu((A \cup B) \cap \mathcal{P}_n(x))}{\mu(\mathcal{P}_n(x))} \\ &= \lim_{n\to\infty} \frac{\mu(A \cap \mathcal{P}_n(x))}{\mu(\mathcal{P}_n(x))} + \lim_{n\to\infty} \frac{\mu(B \cap \mathcal{P}_n(x))}{\mu(\mathcal{P}_n(x))} \\ &= E(A, P) + E(B, P) \end{aligned}$$

where x is any element of $M_{\mathbb{1}_{A\cup B}} \cap P$ (that is nonempty because $P \in \mathcal{P}_*$). Furthermore μ_P is also σ-subadditive, in fact if $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$, called $A = \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$ we have

$$\begin{aligned} E(A,P) &= \lim_{n\to\infty} \frac{\mu(A\cap \mathcal{P}_n(x))}{\mu(\mathcal{P}_n(x))} \\ &\le \lim_{n\to\infty} \sum_{n\in\mathbb{N}} \frac{\mu(A_n\cap \mathcal{P}_n(x))}{\mu(\mathcal{P}_n(x))} \\ (\text{monotone conv.}) &= \sum_{n\in\mathbb{N}} \lim_{n\to\infty} \frac{\mu(A_n\cap \mathcal{P}_n(x))}{\mu(\mathcal{P}_n(x))} = \sum_{n\in\mathbb{N}} E(A_n,P) \end{aligned}$$

where x is any element of $M_{\mathbb{1}_A} \cap P$ (that is nonempty because $P \in \mathcal{P}_*$). Hence from a well known theorem from measure theory (Hahn-Carathéodory extension theorem), we can extend μ_P to a probability measure on $\mathcal{B}$ that we still denote μ_P. Now let us check that the family of probability measures $\{\mu_P\}_{P\in\mathcal{P}_*}$ is a disintegration.

Let $P \in \mathcal{P}_*$, for every $n \in \mathbb{N}_+$ let P_n be the element of $\mathcal{P}_n$ that contains P_n. Notice that if $A \in \mathcal{A}$ is s.t. $A \cap P_n = \emptyset$ for some $n \in \mathbb{N}_+$, then holds

$$\mu_P(A) = E(A,P) = \lim_{m\to\infty} \frac{\mu(A\cap P_m)}{\mu(P_m)} = 0, \tag{3.3}$$

since $P_m \subset P_n$ for every $m \ge n$. Fix $n \in \mathbb{N}_+$ and for every $s \in \mathbb{N}$ consider

$$P_n^s = \bigcup\{U_{j_0}^{a_0} \cap \cdots \cap U_{j_s}^{a_s} \mid a_i \in \{0,1\},\ \ j_i \in \mathbb{N}\ \forall i \in \{0,1,\dots,s\},\ \ [U_{j_0}^{a_0} \cap \cdots \cap U_{j_s}^{a_s}] \cap P_n \neq \emptyset\}$$

where $U^0 = U$ and $U^1 = U^c$ for every $U \in \mathcal{U}$. From (3.3) follows that $\mu_P(P_n^s) = 1$ for every $s \in \mathbb{N}$. Now observe that any open set that contains P_n is of the form $\bigcap_{i\in\mathbb{N}} U_{j_i}^{a_i}$, with $j_i \in \mathbb{N}$ and $a_i \in \{0,1\}$ for every $n \in \mathbb{N}_+$, it follows that $\mu(U) = 1$ for every open set U that contains P_n. Since every finite measure on a separable metric space is a Radon measure, it follows that $\mu_P(P_n) = 1$. Taking the limit for $n \to \infty$ we find that $\mu_P(P) = 1$ for every $P \in \mathcal{P}_*$. We have shown the condition (1) of the definition of disintegration.

Now let us prove conditions (2) and (3) of the definition of disintegration. Consider the family

$$\mathcal{C} = \{E \in \mathcal{B} \mid \text{conditions (2) and (3) holds}\}.$$

By construction, given any $A \in \mathcal{A}$ the function $P \mapsto \mu_P(A) = E(A,P)$ is measurable and satisfies

$$\mu(A) = \int_{\mathcal{P}} E(A,P)\,\mathrm{d}\hat{\mu}(P) = \int_{\mathcal{P}_*} \mu_P(A)\,\mathrm{d}\hat{\mu}(P),$$

hence $\mathcal{A} \subset \mathcal{C}$. We claim that $\mathcal{C}$ is a λ-system. Indeed if $A, B \in \mathcal{C}$, $A \subset B$, then $E(B\setminus A, P) = E(B,P) - E(A,P)$ and

$$\mu(B\setminus A) = \mu(B) - \mu(A) = \int_{\mathcal{P}_*} (\mu_P(B) - \mu_P(A))\,\mathrm{d}\hat{\mu}(P) = \int_{\mathcal{P}_*} \mu_P(B\setminus A)\,\mathrm{d}\hat{\mu}(P)$$

for any $P \in \mathcal{P}_*$, so $B \setminus A \in \mathcal{C}$; moreover if A is the union of a family $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{C}$ s.t. $A_n \subset A_{n+1}$ for any $n \in \mathbb{N}$, then

$$P \mapsto \mu_P(A) = \sup_{n\to\infty} \mu_P(A_n)$$

is measurable and by the monotone converge theorem

$$\mu(A) = \sup_{n\in\mathbb{N}} \mu(A_n) = \sup_{n\in\mathbb{N}} \int_{\mathcal{P}_*} \mu_P(A_n)\,\mathrm{d}\hat{\mu}(P) = \int_{\mathcal{P}_*} \sup_{n\in\mathbb{N}} \mu_P(A_n)\,\mathrm{d}\hat{\mu}(P) = \int_{\mathcal{P}_*} \mu_P(A)\,\mathrm{d}\hat{\mu}(P),$$

so $A \in \mathcal{C}$. Hence $\mathcal{C}$ is a λ-system that contains the algebra $\mathcal{A}$ that generates $\mathcal{B}$ ad by the monotone class theorem follows that $\mathcal{C} = \mathcal{B}$. The proof is complete.

□

Now we are ready to prove the Ergodic decompodition Theorem 3.2.1.

Proof of Theorem 3.2.1. Let $\mathcal{U}$ be a countable basis of open sets of the topology of M and $\mathcal{A}$ be the algebra generated by $\mathcal{U}$. Observe that $\mathcal{A}$ is countable and generates $\mathcal{B}$. By the Birkhoff ergodic Theorem 2.3.2, for every $A \in \mathcal{A}$ there exists $M_A \in \mathcal{B}$ with $\mu(M_A) = 1$ s.t. the mean sjourn time $\tau(A, x)$ is well defined for every $x \in M_A$. Take $M_0 = \bigcap_{A \in \mathcal{A}} M_A$ and note that $\mu(M_0) = 1$, since the intersection is countable. Consider the partition $\mathcal{P}$ of M_0 defined by the following equivalent relation: for every $x, y \in M_0$ we have that x and y are equivalent if and only if $\tau(A, x) = \tau(A, y)$ for every $A \in \mathcal{A}$. Let us prove now that $\mathcal{P}$ is a measurable partition. Let us enumerate $\mathcal{A} = \{A_k\}_{k \in \mathbb{N}_+}$ and $\mathbb{Q} = \{q_k\}_{k \in \mathbb{N}_+}$. For every $n \in \mathbb{N}_+$ consider the partition $\mathcal{P}_n$ of M_0 defined as follows: we can equip each $x \in M_0$ with a function $\omega^x \in \{0, 1\}^{\{1, \dots, n\}^2}$ s.t. for any $i, j \in \{1, \dots, n\}$

$$\omega^x(i, j) = \begin{cases} 0 & \tau(A_i, x) \leq q_j \\ 1 & \tau(A_i, x) > q_j \end{cases}$$

then consider the equivalent relation s.t. for every $x, y \in M_0$ we have that x is equivalent to y if and only if $\omega^x = \omega^y$, then define $\mathcal{P}_n$ to be the partition induced by this relation. It is easy to see that $\mathcal{P}_n$ has at most 2^{n^2} elements. It is also clear that two elements of M_0, say x and y, are in the same element of the partition $\bigvee_{n \in \mathbb{N}_+} \mathcal{P}_n$ if and only if $\tau(A_i, x) = \tau(A_i, y)$ for every $i \in \mathbb{N}_+$, that gives us

$$\mathcal{P} = \bigvee_{n \in \mathbb{N}_+} \mathcal{P}_n$$

which implies that $\mathcal{P}$ is a measurable partition.

Hence by Theorem 3.2.7, there exists some disintegration $\{\mu_P\}_{P \in \mathcal{P}}$ of μ w.r.t. $\mathcal{P}$. The point (1),(2) and (4) of the thesis are parts of the definition of disintegration, let us prove (3). Consider the family of probability measures $\{T_* \mu_P\}_{P \in \mathcal{P}}$ and observe for every $P \in \mathcal{P}$ holds $T^{-1}(P) = P$, since for every set $A \in \mathcal{B}$ the mean sojourn time $\tau(A, \cdot)$ is constant on orbits, so in particular

$$T_* \mu_P(P) = \mu_P(T^{-1}(P)) = \mu_P(P) = 1.$$

Moreover, given any $E \in \mathcal{B}$ the function $P \mapsto T_* \mu_P(E) = \mu_P(T^{-1}(P))$ is measurable w.r.t. $\hat{\mathcal{B}}$ since it is a composition of measurable maps (recall the point (2)) and using the fact that T preserves μ we get

$$\mu(E) = \mu(T^{-1}(E)) = \int_{\mathcal{P}} \mu_P(T^{-1}(E)) \, \mathrm{d}\hat{\mu}(P) = \int_{\mathcal{P}} T_* \mu_P(E) \, \mathrm{d}\hat{\mu}(P).$$

So $\{T_* \mu_P\}_{P \in \mathcal{P}}$ is a disintegration of μ w.r.t. $\mathcal{P}$, hence by Proposition 3.2.3 follows that $T_* \mu_P = \mu_P$ for $\hat{\mu}$-a.e. $P \in \mathcal{P}$, that shows the invariance of μ_P for $\hat{\mu}$-a.e. $P \in \mathcal{P}$.

Since $\mu(M_0) = 1$ we have that $\mu_P(M_0 \cap P) = 1$ for $\hat{\mu}$-a.e. $P \in \mathcal{P}$, indeed

$$0 = 1 - \mu(M_0) = \int_{\mathcal{P}} [1 - \mu_P(M_0)] \, \mathrm{d}\hat{\mu}(P) = \int_{\mathcal{P}} \underbrace{[1 - \mu_P(M_0 \cap P)]}_{\geq 0} \, \mathrm{d}\hat{\mu}(P).$$

Hence it is enough to prvoe that for every $P \in \mathcal{P}$ and every $E \in \mathcal{B}$ we have that $\tau(E, x)$ is well defined for any $x \in M_0 \cap P$ and is constant on that set. Fix $P \in \mathcal{P}$ and denote

$$\mathcal{C}_P = \{E \in \mathcal{B} \mid \forall E \in \mathcal{B} \ \tau(E, x) \text{ is well defined } \forall x \in M_0 \cap P \text{ and is constant on } M_0 \cap P\}.$$

Let us prove that $\mathcal{C}_P$ is a λ-system. For $A, B \in \mathcal{C}_P$, $A \subset B$ we have that

$$\tau(B \setminus A, x) = \tau(B, x) - \tau(A, x)$$

is well defined for any $x \in M_0 \cap P$ and is constant on $M_0 \cap P$, so $B \setminus A \in \mathcal{C}_P$. In particular, since $M \in \mathcal{C}_P$, for every $E \in \mathcal{C}_P$ holds $E^c \in \mathcal{C}_P$. Moreover if $\{E_i\}_{i\in\mathbb{N}}$ are pairwise disjoint we have that

$$\tau\left(\bigcup_{i\in\mathbb{N}} E_i, x\right) = \sum_{i\in\mathbb{N}} \tau(E_i, x)$$

isis well defined for any $x \in M_0 \cap P$ and is constant on $M_0 \cap P$, so $\bigcup_{i\in\mathbb{N}} E_i \in \mathcal{C}_P$. Hence $\mathcal{C}_P$ is a λ-system that contains the algebra $\mathcal{A}$ that generates $\mathcal{B}$, then by the monotone class theorem follows $\mathcal{C}_P = \mathcal{B}$ for every $P \in \mathcal{P}$. □

3.3 Unique ergodicity and minimality

Throughout this section (M, d) will always be a compact metric space, $\mathcal{B}$ the Borel σ-algebra of $\mathcal{M}$ and a $T : M \to M$ a continuous transformation.

Definition 3.3.1: The transformation T is said to be *uniquely ergodic* if it admits exactly one invariant Borel probability measure.

Proposition 3.3.2: *If T is uniquely ergodic, the unique invariant Borel probability measure μ is necessarly ergodic.*

Proof. Suppose by contradition that there exists an invariant set $A \in \mathcal{B}$ with $\mu(A) \in (0, 1)$. Then the measure

$$\mu_A(E) = \frac{\mu(A \cap E)}{\mu(E)} \quad \forall E \in \mathcal{B}$$

would be itself an invariant probability measure, different from μ and that is contradiction. □

Theorem 3.3.3: *The following conditions are equivalent:*

(1) T is uniquely ergodic;

(2) T admits a unique ergodic probability measure;

(3) for every $\phi \in C(M)$ the sequence of time averages $n^{-1} \sum_{j=0}^{n-1} \phi \circ T^j$ converges everywhere on M to a constant;

(4) for every $\phi \in C(M)$ the sequence of time averages $n^{-1} \sum_{j=0}^{n-1} \phi \circ T^j$ converges uniformly on M to a constant.

In particular, if T is uniquely ergodic then for every $\phi \in C(M)$ the sequence of time averages $n^{-1} \sum_{j=0}^{n-1} \phi \circ T^j$ converges uniformly on M to the constant $\int_M \phi \, \mathrm{d}\mu$.

Proof. In the previous Proposition we proved that (1) implies (2) and from the ergodic decomposition Theorem 3.2.1 follows that any invariant probability measure is a convex combination of ergodic measures, hence if there exists only one ergodic probability measure, there is also only one invariant probability measure. It is also clear that (4) implies (3). Let us prove that (3) implies (2). Suppose that μ and ν are erogdic probability measures, then for every $\phi \in C(M)$ by the Birkhoff ergodic Theorem 2.3.2

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \phi(T^j(x)) = \begin{cases} \int_M \phi \, \mathrm{d}\mu & \mu\text{-a.e. on } M \\ \int_M \phi \, \mathrm{d}\nu & \nu\text{-a.e. on } M. \end{cases}$$

Hence, since (3) holds, we obtain

$$\int_M \phi \, \mathrm{d}\mu = \int_M \phi \, \mathrm{d}\nu \;\; \forall \phi \in C(M)$$

that implies $\mu = \nu$.

Now (2) implies

It remains to prove that (1) implies (4). Recall that from Krylov-Bogolyubov Theorem 1.4.6 T admits some invariant probability measure. Suppose that (4) does not hold, hence there exists some $\phi \in C(M)$ s.t. $n^{-1} \sum_{j=0}^{n-1} \phi \circ T^j$ does not converge uniformly to any constant. In particular it does not converge uniformly to $\int_M \phi \, \mathrm{d}\mu$. Hence there exists a $\varepsilon > 0$ s.t. for every $k \in \mathbb{N}_+$ we can find $n_k \geq k$ and $x_k \in M$ s.t.

$$\left| \frac{1}{n_k} \sum_{j=0}^{n_k - 1} \phi(T^j(x_k)) - \int_M \phi \, \mathrm{d}\mu \right| \geq \varepsilon.$$

Let us consider the sequence $(\nu_k)_{k \in \mathbb{N}_+} \subset \mathcal{M}_1(M)$ s.t. for every $k \in \mathbb{N}_+$

$$\nu_k = \frac{1}{n_k} \sum_{j=0}^{n_k - 1} \delta_{T^j(x_k)},$$

since $\mathcal{M}_1(M)$ is weak* sequentially compact we can suppose, up to consider a subsequence, that $(\nu_k)_{k \in \mathbb{N}_+}$ converges in the weak* topology to some $\nu \in \mathcal{M}_1(M)$. The limit measure ν is invariant, indeed

$$\begin{aligned} T_* \nu_k &= \frac{1}{n_k} \sum_{j=1}^{n_k} \delta_{T^j(x_k)} \\ &= \nu_k + \frac{1}{n_k} (\delta_{T^{n_k}(x_k)} - \delta_{x_k}) \xrightarrow{*} \nu \end{aligned}$$

and since T is continuous, for every $\psi \in C(M)$ we have $\psi \circ T \in C(M)$, hence

$$\int_M \psi \, \mathrm{d}(T_* \nu_k) = \int_M \psi \circ T \, \mathrm{d}\nu_k \to \int_M \psi \circ T \, \mathrm{d}\nu = \int_M \psi \, \mathrm{d}(T_* \nu)$$

for every $\psi \in C(M)$, that is $T_* \nu_k \xrightarrow{*} T_* \nu$. As a consequence $T_* \nu = \nu$.

The fact that $(\nu_k)_{k \in \mathbb{N}_+}$ converges to ν in the weak* topology implies that

$$\int_M \phi \, \mathrm{d}\nu = \lim_{k \to \infty} \int_M \phi \, \mathrm{d}\nu_k = \lim_{k \to \infty} \frac{1}{n_k} \sum_{j=0}^{n_k - 1} \phi(T^j(x_k))$$

so by the choice of $(n_k)_{k \in \mathbb{N}_+}$ we have

$$\left| \int_M \phi \, \mathrm{d}\nu - \int_M \phi \, \mathrm{d}\mu \right| \geq \varepsilon$$

and in particular $\nu \neq \mu$, so (1) does not hold. □

Definition 3.3.4: We say that a cleased invariant set $\Lambda \in \mathcal{B}$ is *minimal* if for every $x \in \Lambda$ holds

$$\overline{\{T^n(x)\}_{n \in \mathbb{N}}} = \Lambda.$$

Remark 3.3.5: Recall that the support of a measure μ on M is

$$\operatorname{supp}(\mu) = \{x \in M \mid \forall V \in \mathcal{I}_x \ \ \mu(V) > 0\}$$

where $\mathcal{I}_x$ is the family of all the neightborhoods of x.

Obviously $\operatorname{supp}(\mu)$ is a closed subset of M. Indeed if $x \notin \operatorname{supp}(\mu)$, then there exists an open $V_0 \in \mathcal{I}_x$ s.t. $\mu(V_0) = 0$ and then $V_0 \subset \operatorname{supp}(\mu)^c$.

Proposition 3.3.6: *The support of any invariant measure μ is s.t. $T(\operatorname{supp}(\mu)) \subset \operatorname{supp}(\mu)$.*

Proof. Let $x \in \operatorname{supp}(\mu)$, then take $V \in \mathcal{I}_{T(x)}$. By the continuity of T we get $T^{-1}(V) \in \mathcal{I}_x$, hence $\mu(T^{-1}(V)) > 0$, because $x \in \operatorname{supp}(\mu)$. Now using that μ is invariant we get

$$\mu(V) = \mu(T^{-1}(V)) > 0$$

for any $V \in \mathcal{I}_{T(x)}$, that is $T(x) \in \operatorname{supp}(\mu)$. □

Theorem 3.3.7: *It T is uniquely ergodic then the support of the unique invariant Borel probability measure μ is a minimal set.*

Proof. Suppose by contradiction that there exists $x \in \operatorname{supp}(\mu)$ s.t. the orbit $\{T^j(x)\}_{j\in\mathbb{N}}$ is not dense in $\operatorname{supp}(\mu)$. Then there exists a point $y \in \operatorname{supp}(\mu)$ and a $V \in \mathcal{I}_y$ s.t. $V \cap \{T^j(x)\}_{j\in\mathbb{N}} = \emptyset$. In particular we can find an open set U of M s.t. $\operatorname{supp}(\mu) \cap U \neq \emptyset$ and

$$U \cap \{T^j(x)\}_{j\in\mathbb{N}} = \emptyset. \tag{3.4}$$

Consieer the sequence of probability measures $(\nu_n)_{n\in\mathbb{N}_+}$ s.t.

$$\nu_n = \frac{1}{n}\sum_{j=0}^{n-1} \delta_{T^j(x)}$$

and take ν be any accumulation point of this sequence in weak* sequentially compact set $\mathcal{M}_1(M)$. Equation (3.4) gives us that $\nu_n(U) = 0$ for every $n \in \mathbb{N}_+$, hence $\nu(U) = 0$. Indeed we can find functions $(\psi_k)_{k\in\mathbb{N}_+} \subset C_c(M)$ taking values in $[0,1]$ s.t. $\operatorname{supp}(\psi_k) \subset U$ and $\psi_k \nearrow \mathbb{1}_U$, then $\psi_k \leq \mathbb{1}_U$ and in particular $\int_M \psi_k \,\mathrm{d}\nu_n = 0$ for any $k, n \in \mathbb{N}_+$, but by weak* convergence we know that $\int_M \psi_k \,\mathrm{d}\nu_n \to \int_M \psi_k \,\mathrm{d}\nu$ for every $k \in \mathbb{N}_+$, hence $\int_M \psi_k \,\mathrm{d}\nu = 0$ for every $k \in \mathbb{N}_+$ and by monotone convergence we deduce $\nu(U) = 0$. By the same reasoning used in the previous proof we deduce that ν is an invariant measure, hence $\mu = \nu$ by unique ergodicity of T, so $\mu(U) = 0$ that implies $U \subset \operatorname{supp}(\mu)^c$ that contradicts $U \cap \operatorname{supp}(\mu) \neq \emptyset$. □

The converse is false in general.

Theorem 3.3.8 (Furstenberg): *There exists some real-analytic diffeomorphism $T : \mathbb{T}^2 \to \mathbb{T}^2$ that is minimal, preserve the Lebesgue measure m on $\mathbb{T}^2$ but is not ergodic for m. In particular T is not uiquely ergodic.*

Proof. Not done. □

3.4 Haar measure on Lie groups

3.5 Mixing systems

In this section we always suppose μ to be T-invariant.

Definition 3.5.1: The *correlations sequence* of two measurable functions $\phi, \psi : M \to \mathbb{R}$ is $(C_n(\phi,\psi))_{n\in\mathbb{N}}$ s.t.

$$C_n(\phi,\psi) = \int_m (\phi \circ T^n)\psi \, \mathrm{d}\mu - \int_M \phi \, \mathrm{d}\mu \int_M \psi \, \mathrm{d}\mu \;\; \forall n \in \mathbb{N}.$$

Definition 3.5.2: We say that the system (T,μ) is *mixing* if

$$\lim_{n\to\infty} C_n(\mathbb{1}_A, \mathbb{1}_B) = \lim_{n\to\infty} \mu(T^{-n}(A) \cap B) - \mu(a)\mu(B) = 0$$

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

Proposition 3.5.3: *A mixing system (T,μ) is ergodic.*

Proof. By ontradiction suppose that there exists an invariant $A \in \mathcal{B}$ s.t. $\mu(A) \in (0,1)$, taking $B = A^c$ we get $T^{-n}(A) \cap B = \emptyset$ for evry $n \in \mathbb{N}$. Then

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

that is a contradiction with the mixing propery of (T,μ). □

Example 3.5.4 (Ergodicity is strictly weaker than mixing): Let $\theta \in \mathbb{R}$. As we have seen the rotation $R_\theta : S^1 \to S^1$ is ergodic with respect to the Lebesgue measure m on S^1. However (R_θ, m) is not mixing. In fact if $I, J \subset S^1$ are two sufficiently small intervals, then $R_\theta^{-n}(I) \cap J = \emptyset$ and thus $m(R_\theta^{-n}(I) \cap J) = 0$ frequently on $n \in \mathbb{N}$, while $m(A)m(B) \neq 0$.

Proposition 3.5.5: *Assume that*

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

holds for every A, B in a π-system $\mathcal{S}$ that generates $\mathcal{B}$. Then (T,μ) is mixing.

Proof. Consider the family

$$\mathcal{C} = \{A \in \mathcal{B} \mid \lim_{n\to\infty} \mu(T^{-n}(A) \cap B) = \mu(A)\mu(B) \;\; \forall A \in \mathcal{S}\}.$$

We know that $\mathcal{C} \supset \mathcal{S}$, we claim that $\mathcal{C}$ is a λ-system. Take $A_1, A_2 \in \mathcal{C}$, $A_1 \subset A_2$, then

$$\begin{aligned}
\lim_{n\to\infty} \mu(T^{-n}(A_2 \setminus A_1) \cap B) &= \lim_{n\to\infty} \mu((T^{-n}(A_2) \setminus T^{-n}(A_1)) \cap B) \\
&= \lim_{n\to\infty} \mu(T^{-n}(A_2) \cap B) - \lim_{n\to\infty} \mu(T^{-n}(A_1) \cap B) \\
&= (\mu(A_2) - \mu(A_1))\mu(B) = \mu(A_2 \setminus A_1)\mu(B)
\end{aligned}$$

for every $B \in \mathcal{S}$, hence $A_2 \setminus A_1 \in \mathcal{C}$. Moreover if $A = \bigcup_{k \in \mathbb{N}_+} A_k$ is the union of an increasing sequence of sets $A_1 \subset A_2 \subset \cdots \subset A_n \subset \ldots$ in $\mathcal{C}$, then

$$\begin{aligned}\mu(A)\mu(B) = \lim_{n\to\infty} \mu(A_k)\mu(B) &= \lim_{n\to\infty} \mu(T^{-n}(A_k) \cap B) \\ &= \lim_{n\to\infty} \mu((T^{-n}(A) \setminus T^{-n}(A \setminus A_k)) \cap B) \\ &= \lim_{n\to\infty} \mu(T^{-n}(A) \cap B) - \lim_{n\to\infty} \mu(T^{-n}(A \setminus A_k) \cap B) \\ &= \lim_{n\to\infty} \mu(T^{-n}(A) \cap B)\end{aligned}$$

since

$$\lim_{n\to\infty} \mu(T^{-n}(A \setminus A_k) \cap B) \leq \lim_{n\to\infty} \mu(T^{-n}(A \setminus A_k)) = \lim_{n\to\infty} \mu(A_k \setminus A) = 0$$

so $A \in \mathcal{C}$ and $\mathcal{C}$ is actually a λ-system. Hence by the monotone class theorem $\mathcal{C} = \mathcal{B}$. We proved that $\lim_{n\to\infty} \mu(T^{-n}(A) \cap B) = \mu(A)\mu(B)$ for every $A \in \mathcal{B}$ and every $B \in \mathcal{S}$, now by argument similar to the one we have just applied we can extend this equality to all the $A, B \in \mathcal{B}$. □

Now we present a topological version of the notion of mixing.

Definition 3.5.6: Let M to be a topological space and $\mathcal{B}$ its Borel σ-algebra. We say that the measurable transformation T is *topologically mixing* if for every open sets $U, V \subset M$ there exists $n_0 \in \mathbb{N}_+$ s.t. $T^{-n}(U) \cap V \neq \emptyset$ for every $n \geq n_0$.

Proposition 3.5.7: *If (T, μ) is mixing then the restriction $T_{|\mathrm{supp}(\mu)} : \mathrm{supp}(\mu) \to \mathrm{supp}(\mu)$ is topologically mixing.*

Proof. Let $A, B \subset \mathrm{supp}(\mu)$ be two nonempty open sets, then by the definiton of support of μ we have $\mu(A), M(B) > 0$. Hence, from the fact that (T, μ) is mixing, we have the existence of of $n_0 \in \mathbb{N}_+$ s.t.

$$\mu(T^{-n}(A) \cap B) > \frac{\mu(A)\mu(B)}{2} > 0$$

for every $n \geq n_0$. □

Corollary 3.5.8: *If (T, μ) is mixing and μ is positive on open sets, then T is topologically mixing.*

Proof. It follows form the previous Proposition. Indeed if μ is positive on open sets then $\mathrm{supp}(\mu) = M$. □

Proposition 3.5.9: *The following are equivalent:*

(1) (T, μ) is mixing;

(2) there exist $p, q \in [1, \infty]$ conjugate exponents s.t. $\lim_{n\to\infty} C_n(\phi, \psi) = 0$ for any $\phi \in L^p(\mu)$ and any $\psi \in L^q(\mu)$;

(3) the point (2) holds for ϕ in a dense subset of $L^p(\mu)$ and ψ in a dense subset of $L^q(\mu)$.

Proof. Obviously (2) implies (1). And (1) implies that $\lim_{n\to\infty} C_n(\phi, \psi) = 0$ for any simple functions ϕ, ψ, but simple functions are dense in $L^p(\mu)$ for every $p \in [1, \infty]$, so we have (3). It remains to show that (3) implies (2). Observe that if $\phi_1, \phi_2 \in L^p(\mu)$ and $\psi_1, \psi_2 \in L^q(\mu)$, by Hölder inequality we get

$$\begin{aligned}
&\left|\int_M (\phi_1 \circ T^n)\psi_1 \, \mathrm{d}\mu - \int_M (\phi_2 \circ T^n)\psi_2 \, \mathrm{d}\mu\right| \\
&\leq \left|\int_M (\phi_1 \circ T^n)\psi_1 \, \mathrm{d}\mu - \int_M (\phi_2 \circ T^n)\psi_1 \, \mathrm{d}\mu\right| + \left|\int_M (\phi_2 \circ T^n)\psi_1 \, \mathrm{d}\mu - \int_M (\phi_2 \circ T^n)\psi_2 \, \mathrm{d}\mu\right| \\
&\leq \|\phi_1 - \phi_2\|_{L^p(\mu)} \|\psi_1\|_{L^q(\mu)} + \|\phi_2\|_{L^p(\mu)} \|\psi_1 - \psi_2\|_{L^q(\mu)}
\end{aligned}$$

and similarly

$$\begin{aligned}
\left|\int_M \phi_1 \, \mathrm{d}\mu \int_M \psi_1 \, \mathrm{d}\mu - \int_M \phi_2 \, \mathrm{d}\mu \int_M \psi_2 \, \mathrm{d}\mu\right| &\leq \|\phi_1 - \phi_2\|_{L^1(\mu)} \|\psi_1\|_{L^1(\mu)} + \|\phi_2\|_{L^1(\mu)} \|\psi_1 - \psi_2\|_{L^1(\mu)} \\
&\leq \|\phi_1 - \phi_2\|_{L^p(\mu)} \|\psi_1\|_{L^q(\mu)} + \|\phi_2\|_{L^p(\mu)} \|\psi_1 - \psi_2\|_{L^q(\mu)}
\end{aligned}$$

and adding this inequalities together we obtain

$$|C_n(\phi_1, \psi_1) - C_n(\phi_2, \psi_2)| \leq 2\|\phi_1 - \phi_2\|_{L^p(\mu)} \|\psi_1\|_{L^q(\mu)} + 2\|\phi_2\|_{L^p(\mu)} \|\psi_1 - \psi_2\|_{L^q(\mu)}$$

for every $n \in \mathbb{N}_+$. Then taking any $\varepsilon > 0$ □

3.6 Weak-mixing systems

In this section we always suppose μ to be T-invariant.

Definition 3.6.1: The system (T, μ) is *weak-mixing* if for any $A, B \in \mathcal{B}$ holds

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} |C_j(\mathbb{1}_A, \mathbb{1}_B)| = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} |\mu(T^{-j}(A) \cap B) - \mu(A)\mu(B)| = 0.$$

Remark 3.6.2: Every mixing system is weak-mixing. It folllows from the Cesaro means theorem.

Proposition 3.6.3: *If (T, μ) is weak-mixing then it is ergodic.*

Proof. If $A \in \mathcal{B}$ is invariant, then $\mu(T^{-j}(A) \cap A^c) = 0$ for every $j \in \mathbb{N}$ and

$$0 = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} |C_j(\mathbb{1}_A, \mathbb{1}_{A^c})| = \mu(A)\mu(A^c)$$

hence $\mu(A) = 0$ or $\mu(A^c) = 0$. □

Proposition 3.6.4: *Assume that*

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

holds for every A, B in a π-system $\mathcal{S}$ that generates $\mathcal{B}$. Then (T, μ) is weak-mixing.

Proof. The proof is similar to the one of Proposition 3.5.5. □

Given the system (T, μ) we can consider the transformation $T_2 : M \times M \to M \times M$ that is measurable w.r.t. the product σ-algebra $\mathcal{B} \otimes \mathcal{B}$ and the product measure $\mu_2 = \mu \otimes \mu$. We can easily observe that T_2 preserves μ_2.

Proposition 3.6.5: *If (T_2, μ_2) is ergodic then (T, μ) is ergodic. Moreover the converse is false in general.*

Proof. If A is T-invariant then $A \times A$ is T_2-invariant and $\mu_2(A \times A) \in \{0, 1\}$. A counterexample for the converse is a rotation $T = R_\theta : S^1 \to S^1$ with $\theta \in \mathbb{R} \setminus \mathbb{Q}$ and d a distance invariant under rotations. Then any neighborhood of the diagonal $\{(x, y) \in S^1 \times S^1 \mid d(x, y) < r\}$ is T_2-invariant, but in heneral can exists an $r > 0$ s.t. $m(\{(x, y) \in S^1 \times S^1 \mid d(x, y) < r\}) \in (0, 1)$. □

Definition 3.6.6: We say that a set $E \subset \mathbb{N}$ has *zero density at infinity* if

$$\lim_{n\to\infty} n^{-1}\#(E \cap \{0, 1, \ldots, n-1\}) = 0.$$

Lemma 3.6.7: *Let $(a_n)_{n\in\mathbb{N}} \subset [0, +\infty)$ a sequence of non negative numbers. Then*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} a_j = 0 \iff \exists E \subset \mathbb{N} \text{ with zero density at infinity s.t. } \lim_{\substack{n\to\infty\\ n\notin E}} a_n = 0.$$

In particular

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} a_j = 0 \iff \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} a_j^2 = 0.$$

Proof. Suppose that the right-hand side of the equivalence holds. Take $\varepsilon > 0$ and $N \in \mathbb{N}_+$ s.t. $a_n < \varepsilon$ for every $n \geq N$, $n \notin E$. Define $A = sup_{n\notin E} a_n$, then for $n \geq N$ holds

$$\begin{aligned} 0 \leq \frac{1}{n}\sum_{j=0}^{n-1} a_j &= \frac{1}{n}\sum_{\substack{j=0\\ j\notin E}}^{n-1} a_j + \frac{1}{n}\sum_{\substack{j=0\\ j\in E}}^{n-1} a_j \\ &\leq \frac{NA + (n-N)\varepsilon}{n} + \frac{A}{n}\#(E \cap \{0, 1, \ldots, n-1\}) \to \varepsilon \end{aligned}$$

for $n \to \infty$, where we used that E ha zero density at infinity. Therefore $0 \leq \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} a_j \leq \varepsilon$ for every $\varepsilon > 0$, hence $\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} a_j = 0$.

Now suppose that the left-hand side of the equality holds. For $m \in \mathbb{N}_+$ let

$$E_m = \{n \in \mathbb{N} \mid a_n \geq m^{-1}\},$$

then $E_j \subset E_{j+1}$ for every $j \in \mathbb{N}_+$ and E_m ha zero density at infinity for every $m \in \mathbb{N}_+$. Indeed

$$\begin{aligned} 0 &= m \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} a_j \\ &\geq m \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} a_j \mathbb{1}_{E_m}(j) \\ &\geq \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \mathbb{1}_{E_m}(j) = \lim_{n\to\infty} \frac{1}{n}\#\{E_m \cap \{0, 1, \ldots, n-1\}\} \end{aligned}$$

for every $m \in \mathbb{N}_+$. Now consider $0 = N_0 < N_1 < \ldots$ s.t. $n^{-1}\#(E_m \cap \{0, 1, \ldots, n-1\}) < m^{-1}$ for every $n \geq N_{m-1}$ for every $m \in \mathbb{N}_+$. Define

$$E = \bigcup_{m \in \mathbb{N}_+} (E_m \cap \{N_{m-1}, \ldots, N_m - 1\}),$$

then if $m = m(n)$ is the maximal $m \in \mathbb{N}_+$ s.t. $N_{m-1} < n$ we get

$$\begin{aligned}
\frac{1}{n}\#(E \cap \{0, 1, \ldots, n-1\}) &\leq \frac{1}{n}\#(E_{m-1} \cap \{0, 1, \ldots, N_{m-1}\}) + \frac{1}{n}\#(E_{m-1} \cap \{N_{m-1}, \ldots, n-1\}) \\
&\leq \frac{1}{N_{m-1}}\#(E_{m-1} \cap \{0, 1, \ldots, N_{m-1}\}) + \frac{1}{n}\#(E_m \cap \{0, 1, \ldots, n-1\}) \\
\text{(for the choice of } N_{m-1}\text{)} &\leq \frac{1}{m-1} + \frac{1}{n}\#(E_m \cap \{0, 1, \ldots, n-1\}) \to 0
\end{aligned}$$

for $m \to \infty$. □

Proposition 3.6.8: *The following conditions are equivalent:*

(1) (T, μ) *is weak-mixing;*

(2) (T_2, μ_2) *is weak-mixing;*

(3) (T_2, μ_2) *is ergodic.*

Proof. Suppose that (1) holds. Take $A, B, C, D \in \mathcal{B}$, then

$$\begin{aligned}
&|\mu_2(T_2^{-j}(A \times B) \cap C \times D) - \mu_2(A \times B)\mu_2(C \times D)| \\
&\qquad = |\mu(T^{-j}(A) \cap C)\mu(T^{-j}(B) \cap D) - \mu(A)\mu(B)\mu(C)\mu(D)| \\
&\qquad \leq |\mu(T^{-j}(A) \cap C) - \mu(A)\mu(C)| + |\mu(T^{-j}(B) \cap D) - \mu(B)\mu(D)|
\end{aligned}$$

where in the last inequality we used the fact that for $a, b, c, d \in [0, 1]$ holds

$$|ab - cd| = |ab - cb + cb - cd| \leq b|a - c| + c|b - d| \leq |a - c| + |b - d|,$$

therefore, using (1), we get

$$\lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} [\mu_2((A \times B) \cap C \times D) - \mu_2(A \times B)\mu_2(C \times D)] = 0.$$

Hence, since $\{A \times B \mid A, B \in \mathcal{B}\}$ is a π-system that generates $\mathcal{B} \otimes \mathcal{B}$, using Proposition 3.6.4 we get that (T_2, μ_2) is weak-mixing. Therefore (1) implies (2).

We proved in a previous Proposition that (2) implies (3). Let us prove that (3) implies (1). Take $A, B \in \mathcal{B}$ and write

$$\begin{aligned}
&\frac{1}{n} \sum_{j=0}^{n-1} [\mu(T^{-j}(A) \cap B) - \mu(A)\mu(B)]^2 \\
&\qquad = \frac{1}{n} \sum_{j=0}^{n-1} [\mu(T^{-j}(A) \cap B)^2 - 2\mu(A)\mu(B)\mu(T^{-j}(A) \cap B) + \mu(A)^2\mu(B)^2] \\
&\qquad = \frac{1}{n} \sum_{j=0}^{n-1} [\mu_2(T_2^{-j}(A \times A) \cap B \times B) - \mu_2(A \times A)\mu_2(B \times B)] \\
&\qquad\qquad - 2\mu(A)\mu(B)\frac{1}{n} \sum_{j=0}^{n-1} [\mu(T^{-j}(A) \cap B) - \mu(A)\mu(B)].
\end{aligned}$$

Since (T_2, μ_2) is ergodic we obtain that also (T, μ) is ergodic and Proposition 3.1.3 (2) gives that the expression above converge to 0 as $n \longrightarrow \infty$. We proved that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} [\mu(T^{-j}(A) \cap B) - \mu(A)\mu(B)]^2 = 0$$

for any $A, B \in \mathcal{B}$ and using the previous Lemma we obtain that (T, μ) is weak-mixing. $\square$