CHAPTER 3

Review of Ergodic Theory

In this chapter we present a brief review of ergodic theory. Many of the results will be used in the sequel. For a more complete study of ergodic theory the reader is referred to the excellent texts [Petersen, 1983] or [Cornfeld, Fomin and Sinai, 1982].

3.1 Measure-Preserving Transformations

Let $(X, \mathfrak{B}, \mu)$ be a normalized measure space.

Definition 3.1.1. The transformation $\tau : X \to X$ is *measurable* if $\tau^{-1}(\mathfrak{B}) \subset \mathfrak{B}$, i.e., $B \in \mathfrak{B} \Rightarrow \tau^{-1}(B) \in \mathfrak{B}$, where $\tau^{-1}(B) \equiv \{x \in X : \tau(x) \in B\}$.

Definition 3.1.2. We say the measurable transformation $\tau : X \to X$ *preserves measure* μ or that μ is τ*-invariant* if $\mu(\tau^{-1}(B)) = \mu(B)$ for all $B \in \mathfrak{B}$.

Definition 3.1.3. Let $(X, \mathfrak{B}, \mu)$ be a normalized measure space and let $\tau : X \to X$ preserve μ. The quadruple $(X, \mathfrak{B}, \mu, \tau)$ is called a *dynamical system.*

In practice, it is usually difficult to check whether τ preserves a measure since one does not have explicit knowledge of all members of $\mathfrak{B}$. However, we often know a π-system $\mathcal{P}$ that generates $\mathfrak{B}$. For example, if X is the unit interval, then the family $\mathcal{P}$ of all intervals is a π-system. The following result is very useful in checking if a transformation is measure preserving.

Theorem 3.1.1. *Let $(X, \mathfrak{B}, \mu)$ be a normalized measure space and let $\tau : X \to X$ be measurable. Let $\mathcal{P}$ be a π-system (Definition 2.1.4) that generates $\mathfrak{B}$. If $\mu(\tau^{-1}A) = \mu(A)$ for any $A \in \mathcal{P}$, then τ is measure preserving (preserves μ).*

Proof. Let us define a new measure on $\mathfrak{B}$, $\eta(A) = \mu(\tau^{-1}(A))$ (see Problem 3.1.10). The measures μ and η agree on the π-system $\mathcal{P}$. By Theorem 2.1.2, $\mu = \eta$ on $\mathfrak{B}$. $\square$

The following theorem gives a necessary and sufficient condition for τ-invariance of μ.

Theorem 3.1.2. *Let $\tau : X \to X$ be a measurable transformation of $(X, \mathfrak{B}, \mu)$. Then τ is μ-preserving if and only if*

$$\int_X f(x) d\mu = \int_X f(\tau(x)) d\mu \tag{3.1.1}$$

for any $f \in \mathfrak{L}^\infty$. If X is compact and (3.1.1) holds for any continuous function f, then τ is μ-preserving.

Two examples of measure-preserving transformations are presented below. More examples can be found in the problem section at the end of the chapter.

Example 3.1.1. Let $X = [0,1], \mathfrak{B} =$ Borel σ-algebra of $[0,1]$ and $\lambda =$ Lebesgue measure on $[0,1]$. Let $\tau : X \to X$ be defined by $\tau(x) = rx$ (mod 1), where r is a positive integer greater than or equal to 2. Then τ is measure preserving.

Proof. Let $[a,b] \subset [0,1]$ be a subinterval of $[0,1]$. Its preimage $\tau^{-1}([a,b])$ consists of r disjoint intervals $I_1, \ldots, I_r$ and $\lambda(I_i) = \frac{1}{r}(b-a)$ for $i = 1, \ldots, r$. Thus, $\lambda(\tau^{-1}[a,b]) = \lambda([a,b])$. Since the family $\mathcal{P} = \{[a,b] \subset [0,1]\}$ is a π-system generating $\mathfrak{B}$, Theorem 3.1.1 implies that λ is τ-invariant. □

Example 3.1.2. Let $(X, \mathfrak{B}, \mu)$ be as in Example 3.1.1. Define $\tau : X \to X$ by $\tau(x) = x + \alpha$ (mod 1), where $\alpha > 0$. Then τ preserves Lebesgue measure.

Proof. As in Example 3.1.1 it is enough to show that $\lambda(\tau^{-1}[a,b]) = \lambda[a,b]$ for any subinterval $[a,b] \subset [0,1]$. The preimage $\tau^{-1}[a,b]$ consists of one or two disjoint intervals and $\lambda(\tau^{-1}[a,b]) = \lambda([a,b])$. A more natural way to view this example is to interpret it as a rotation of the circle. Then the τ-invariance of λ is obvious. □

The following theorem establishes the existence of invariant measures for an important class of transformations.

Theorem 3.1.3. *(Krylov–Bogoliubov Theorem) [Krylov and Bogoliubov, 1937] Let X be a compact metric space and let $\tau : X \to X$ be continuous. Then there exists a τ-invariant normalized measure on X.*

Proof. Let ν be a normalized measure on X. We consider a sequence

$$\mu_n = \frac{1}{n}(\nu + \tau_*\nu + \cdots + \tau_*^{n-1}\nu),$$

$n = 1, 2, \ldots$, where τ_* is the operator on the space of measures defined by $\tau_*\nu = \nu \circ \tau^{-1}$. By Corollary 2.2.3, the sequence $\{\mu_n\}_{n=1}^{\infty}$ is precompact in the weak topology of measures, i.e., it contains a weakly convergent subsequence $\{\mu_{n_k}\}_{k=1}^{\infty}$. Let μ be a limit point of this subsequence:

$$\mu_{n_k} \xrightarrow{\text{weak}} \mu$$

as $k \to +\infty$. We will prove that μ is τ-invariant, i.e., that μ is a fixed point of τ_*. To this end it is enough to show that for any continuous function $g : X \to \mathbb{R}$,

$$\int_X g d\mu = \int_X g \circ \tau d\mu. \tag{3.1.2}$$

We have

$$\begin{aligned}
|\mu(g) - \mu(g \circ \tau)| &= \lim_{k \to +\infty} |\mu_{n_k}(g) - \mu_{n_k}(g \circ \tau)| \\
&= \lim_{k \to +\infty} |\frac{1}{n_k}(\nu + \tau_*\nu + \cdots + \tau_*^{n_k - 1}\nu)(g) \\
&\qquad - \frac{1}{n_k}(\tau_*\nu + \tau_*^2\nu + \cdots + \tau_*^{n_k-1}\nu + \tau_*^{n_k}\nu)(g)| \\
&= \lim_{k \to +\infty} \frac{1}{n_k}|\nu(g) - \tau_*^{n_k}\nu(g)| \le \lim_{k \to +\infty} \frac{2 \sup |g|}{n_k} = 0,
\end{aligned}$$

and (3.1.2) is proved. Continuity of τ is necessary to claim that $\mu(g \circ \tau) = \lim_{k \to +\infty} \mu_{n_k}(g \circ \tau)$. See Problem 3.1.15 for a counter example. □

3.2 Recurrence and Ergodicity

Let $\tau : X \to X$ be a transformation. The nth iterate of τ is denoted by τ^n, i.e.,

$$\tau^n(x) = \tau \circ \ldots \circ \tau(x)$$

n times. In the study of dynamical systems, we are interested in properties of the *orbit* $\{\tau^n(x)\}_{n \ge 0}$. For example, in the recurrence of orbits of τ, i.e., the property that if the orbit starts in a specified set, it returns to that set infinitely many times. If τ is measure preserving, then we have a simply stated but powerful result regarding the recurrence of orbits.

Theorem 3.2.1. *(Poincaré Recurrence Theorem, 1899)*
Let τ be a measure-preserving transformation on a normalized measure space $(X, \mathfrak{B}, \mu)$. Let $E \in \mathfrak{B}$ such that $\mu(E) > 0$. Then almost all points of E return infinitely often to E under iterations of τ.

This theorem has interesting physical and philosophical implications.

Proof. Let A be a measurable set with $\mu(A) > 0$, and let us define the set B of points that never return to A, i.e., $B = \{x \in A : \tau^k(x) \notin A,\ k = 1, 2, \dots\}$. We will prove that

$$\tau^{-i}(B) \cap \tau^{-j}(B) = \emptyset,$$

for $i > j \geq 0$. If $x \in \tau^{-i}(B) \cap \tau^{-j}(B)$, then $\tau^j(x) \in B$ and $\tau^{i-j}(\tau^j(x)) = \tau^i(x) \in B$, which contradicts the definition of B. Hence, we have

$$\sum_{i=0}^{\infty} \mu(\tau^{-i}(B)) = \mu(\cup_{i=0}^{\infty} \tau^{-i}(B)) \leq \mu(X) = 1.$$

Since μ is τ-invariant, this implies that $\sum_{i=0}^{\infty} \mu(B) \leq 1$. Therefore, $\mu(B) = 0$. $\square$

Example 3.2.1. Poincaré's Recurrence Theorem has an interesting consequence for almost every number $x \in [0, 1]$. Let $\tau(x) = 10 \cdot x \pmod 1$ on $[0, 1]$. τ preserves Lebesgue measure λ and is closely related to the decimal expansion of numbers. For any $x \in [0, 1)$, we have

$$x = \frac{\varepsilon_1}{10} + \frac{\varepsilon_2}{10^2} + \frac{\varepsilon_3}{10^3} + \dots,$$

where $\varepsilon_i = [10 \cdot \tau^{i-1}(x)],\ i = 1, 2, \dots$, so $0 \leq \varepsilon_i \leq 9$. Then

$$\tau(x) = \frac{\varepsilon_2}{10} + \frac{\varepsilon_3}{10^2} + \frac{\varepsilon_4}{10^3} + \dots.$$

Let $A = [0, 0.\varepsilon_1\varepsilon_2\varepsilon_3 \dots \varepsilon_n]$. $\lambda(A) > 0$, so almost every point visits A infinitely many times. This means that the group of digits $\varepsilon_1\varepsilon_2\varepsilon_3 \dots \varepsilon_n$ repeats infinitely many times in the decimal expansion of almost every number $x \in [0, 1)$.

Let $(X, \mathfrak{B}, \mu)$ be a normalized measure space and let $\tau : X \to X$ be a measure-preserving transformation on $(X, \mathfrak{B}, \mu)$. If $\tau^{-1}B = B$ for some $B \in \mathfrak{B}$, then $\tau^{-1}(X \backslash B) = X \backslash B$ and the study of τ splits into two parts: $\tau\mid_B$ and $\tau\mid_{X \backslash B}$. It is useful to have a concept of *indecomposability* for measure-preserving transformations, so that if τ has this indecomposability property then the study of τ cannot be split into separate parts. This property is called *ergodicity*.

Definition 3.2.1. We call a measure-preserving transformation $\tau : (X, \mathfrak{B}, \mu) \to (X, \mathfrak{B}, \mu)$ *ergodic* if for any $B \in \mathfrak{B}$, such that $\tau^{-1}B = B$, $\mu(B) = 0$ or $\mu(X \setminus B) = 0$.

Since ergodicity is a property of the pair (τ, μ), we often say that (τ, μ) is ergodic.

Example 3.1.1 is ergodic (see Example 3.2.2). Example 3.1.2 is ergodic if and only if α is irrational (see Problem 3.2.1).

Example 3.2.2. We will prove the ergodicity of $\tau(x) = 2x \pmod 1$, $x \in [0, 1]$.

Let $A = \tau^{-1}(A)$ be an invariant set. Then, whenever $x_1 \in A$ and $\tau(x_1) = \tau(x_2)$, $x_2 \in A$ as well. Since $\tau([0, \frac{1}{2}]) = \tau([\frac{1}{2}, 1])$, we have

$$\lambda(A) = 2\lambda(A \cap [0, \frac{1}{2}]) = \frac{\lambda(A \cap [0, \frac{1}{2}])}{\lambda([0, \frac{1}{2}])}$$

or

$$\lambda(A \cap [0, \frac{1}{2}]) = \lambda(A) \cdot \lambda([0, \frac{1}{2}]). \tag{3.2.1}$$

Similarly, $\lambda(A \cap [\frac{1}{2}, 1]) = \lambda(A) \cdot \lambda([\frac{1}{2}, 1])$. For any $B \in \mathfrak{B}$, let $B_1 = \tau^{-1}(B) \cap [0, \frac{1}{2}]$ and $B_2 = \tau^{-1}(B) \cap [\frac{1}{2}, 1]$. Then

$$\lambda(A \cap \tau^{-1}(B)) = 2 \cdot \lambda(A \cap B_1) = 2 \cdot \lambda(A \cap B_2). \tag{3.2.2}$$

Using (3.2.1) and (3.2.2), we can show by induction that

$$\lambda(A \cap E) = \lambda(A) \cdot \lambda(E),$$

for any dyadic interval E (i.e., one with dyadic endpoints) and then for any union of dyadic intervals. The set A can be approximated arbitrarily closely by a union of dyadic intervals and we obtain, for any $\varepsilon > 0$,

$$|\lambda(A \cap A) - \lambda(A) \cdot \lambda(A)| < \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, $\lambda(A) = \lambda^2(A)$, which implies that $\lambda(A)$ equals 0 or 1. Thus, τ is ergodic.

The symbol Δ denotes the symmetric difference of sets: $A \Delta B = (A \setminus B) \cup (B \setminus A)$.

Definition 3.2.2. Let $(X, \mathfrak{B}, \mu, \tau)$ be a dynamical system. A set $B \in \mathfrak{B}$ is called τ-invariant if $\tau^{-1}(B) = B$ and almost τ-invariant if $\mu(\tau^{-1}(B) \triangle B) = 0$. Similarly, a measurable function is called τ-invariant if $f \circ \tau = f$ and almost τ-invariant if $f \circ \tau = f$ μ-a.e.

Theorem 3.2.2. *The following three statements are equivalent for the transformation $\tau : (X, \mathfrak{B}, \mu) \to (X, \mathfrak{B}, \mu)$ preserving a normalized measure μ:*
(1) τ is ergodic.
(2) $\mu((\tau^{-1}B)\triangle B) = 0$, $B \in \mathfrak{B} \Rightarrow \mu(B) = 0$ or 1.
(3) For any $A, B \in \mathfrak{B}$ with $\mu(A) > 0$, $\mu(B) > 0$, there exists $n > 0$ such that $\mu((\tau^{-n}A) \cap B) > 0$.

We will prove Theorem 3.2.2 in a series of lemmas that are of independent interest.

Lemma 3.2.1. *If a normalized measure μ is τ-invariant and $\tau^{-1}B \subset B$, then there exists a set $B_1 \subset B$, $\mu(B \backslash B_1) = 0$ and $\tau^{-1}(B_1) = B_1$.*

Proof. We have $B \supset \tau^{-1}B \supset \tau^{-2}B \supset \dots$. Let $B_1 = \bigcap_{k=0}^{\infty} \tau^{-k}(B)$. Then, $B_1 \subset B$ and $\mu(B_1) = \lim_{k\to+\infty} \mu(\tau^{-k}(B)) = \mu(B)$. Also, $\tau^{-1}(B_1) = \bigcap_{k=1}^{\infty} \tau^{-k}(B) = B_1$. □

Lemma 3.2.2. *If a normalized measure μ is τ-invariant and $\mu(\tau^{-1}(B) \triangle B) = 0$, then there exists a set B_1 such that $\mu(B \triangle B_1) = 0$, and $\tau^{-1}(B_1) = B_1$.*

Proof. If $\mu(\tau^{-1}(B)\triangle B) = 0$, then $\mu(\tau^{-1}(B) \setminus B) = 0$. Let $B_2 = \bigcup_{k=0}^{\infty} \tau^{-k}(B)$. We can write

$$B_2 \setminus B = (\bigcup_{k=0}^{\infty} \tau^{-k}(B)) \setminus B = \bigcup_{k=0}^{\infty} (\tau^{-k-1}(B) \setminus \tau^{-k}(B)).$$

Since μ is τ-invariant,

$$\mu(B_2 \setminus B) \le \sum_{k=0}^{\infty} \mu(\tau^{-k-1}(B) \setminus \tau^{-k}(B)) = \sum_{k=0}^{\infty} \mu(\tau^{-k}(\tau^{-1}(B) \setminus B)) = 0.$$

Furthermore, $\tau^{-1}(B_2) = \bigcup_{k=1}^{\infty} \tau^{-k}(B) \subset B_2$. By Lemma 3.2.1, there exists $B_1 \subset B_2$, $\mu(B_2 \setminus B_1) = 0$ and such that $\tau^{-1}(B_1) = B_1$. We have $\mu(B_1 \setminus B) \le \mu(B_2 \setminus B) = 0$. Thus, $\mu(B\triangle B_1) = 0$. □

Lemma 3.2.3. *If a normalized τ-invariant measure μ is ergodic, then for any set B such that $\tau^{-1}(B) \subset B$, we have $\mu(B)$ equal to 0 or 1.*

Proof. Since $\tau^{-1}(B) \subset B$, we can find $B_1 \subset B$ such that $\tau^{-1}(B_1) = B_1$ and $\mu(B \setminus B_1) = 0$ (Lemma (3.2.1). By ergodicity, $\mu(B_1) = 0$ or 1. Since $\mu(B \setminus B_1) = 0$ we obtain $\mu(B) = 0$ or 1. □

Lemma 3.2.4. *If a normalized τ-invariant measure μ is ergodic and $\mu(A) > 0$, then* $\mu(\cup_{k=1}^{\infty}\tau^{-k}(A)) = 1$.

Proof. Let $B = \bigcup_{k=1}^{\infty} \tau^{-k}(A)$. Then $\tau^{-1}(B) \subset B$, and $\mu(B)$ equals 0 or 1 by Lemma 3.2.3. It cannot be 0 since $\mu(A) > 0$. $\square$

Ergodicity can also be characterized by means of functions:

Theorem 3.2.3. *Let $\tau : (X, \mathfrak{B}, \mu) \to (X, \mathfrak{B}, \mu)$ be measure preserving. Then the following statements are equivalent:*
(1) τ is ergodic.
(2) If f is measurable and $(f \circ \tau)(x) = f(x)$ a.e., then f is constant a.e.
(3) If $f \in \mathfrak{L}^2(\mu)$ and $(f \circ \tau)(x) = f(x)$ a.e., then f is constant a.e.

Proof. It follows from Theorem 3.2.2 and the denseness of characteristic functions both in the space of measurable functions and in $\mathfrak{L}^2(\mu)$. $\square$

Proposition 3.2.1. *Let X be a compact metric space and let μ be a Borel normalized measure on X, which gives positive measure to every non-empty open set. If $\tau : X \to X$ is continuous and ergodic with respect to μ, then*

$$\mu\{x : \{\tau^n x : n \geq 0\} \textit{ is dense in } X\}\} = 1$$

Proof. Let $\{U_n\}_{n=1}^{\infty}$ be a base for the topology of X. Then $\{\tau^n(x) : n \geq 0\}$ is dense in $X \Leftrightarrow x \in \cap_{n=1}^{\infty} \cup_{k=0}^{\infty} \tau^{-k} U_n$. This follows from the fact that $Y = \cup_{k=0}^{\infty}\tau^{-k}U_n$ is the set of points that go into U_n after k iterations of τ for some $k \geq 0$. Since denseness requires that x visits every U_n, we need $x \in \cap_{n=1}^{\infty} \cup_{k=0}^{\infty} \tau^{-k} U_n$.

Since μ is measure preserving, $\mu(\tau^{-1}Y) = \mu(Y)$. But

$$\tau^{-1}(\cup_{k=0}^{\infty}\tau^{-k}U_n) \subset \cup_{k=0}^{\infty}\tau^{-k}U_n.$$

Hence, $\mu(Y \triangle \tau^{-1}Y) = 0$. By ergodicity, we have $\mu(\cup_{k=0}^{\infty}\tau^{-k}U_n) = 0$ or 1. Since $\cup_{k=0}^{\infty}\tau^{-k}U_n$ is a non-empty open set (by continuity of τ), we have $\mu(\cup_{k=0}\tau^{-k}U_n) = 1$. Thus,

$$\mu(\cap_{n=1}^{\infty} \cup_{k=0}^{\infty} \tau^{-k} U_n) = 1.$$

$\square$

For ergodic transformations we have the following stronger version of the Poincaré Recurrence Theorem, known as Kac's Lemma [Kac, 1947]. Let A be a measurable set with $\mu(A) > 0$ and let us define, for $x \in A$,

$$n(x) = \min\{k \geq 1 : \ \tau^k(x) \in A\}.$$

$n(x)$ is the time of the first return of x to the set A.

Theorem 3.2.4. *(Kac's Lemma) If μ is τ-invariant and ergodic, and A is a measurable set with $\mu(A) > 0$, then*

$$\int_A n(x) d\mu(x) = 1. \tag{3.2.3}$$

In terms of the conditional measure $\mu_A(B) = \frac{\mu(A\cap B)}{\mu(A)}$, this can be written as

$$\int_A n(x) d\mu_A(x) = \frac{1}{\mu(A)},$$

which says that the expected time of return to a set A is $\frac{1}{\mu(A)}$.

Proof. Let us define $A_k = \{x \in A : \ n(x) = k\}$, $B_k = \{x \in X : \tau^k(x) \in A \text{ and } \tau^j(x) \notin A \text{, for } j = 1, \ldots, k-1\}$ and $C_k = B_k \backslash A_k$, $k = 1, 2, \ldots$. It is easy to see that $\{A_k\}_{k=1}^{\infty}$ are mutually disjoint subsets of A, $\{B_k\}_{k=1}^{\infty}$ are mutually disjoint subsets of X and $A_k \subset B_k$, for $k = 1, 2, \ldots$. Then

$$(X \backslash \cup_{k=1}^{\infty} B_k) \cap \tau^{-j}(A) = \emptyset, \text{ for } j = 1, 2, \ldots,$$

so $\mu(X \backslash \cup_{k=1}^{\infty} B_k) = 0$ by Lemma 3.2.4. Hence, $\mu(\cup_{k=1}^{\infty} B_k) = 1$. On the other hand, for any $k \geq 1$, we have $\tau^{-1}(C_k) = B_{k+1}$ and

$$\mu(B_k) = \mu(A_k) + \mu(C_k) = \mu(A_k) + \mu(\tau^{-1}(C_k)) = \mu(A_k) + \mu(B_{k+1}).$$

Thus, we can write

$$\begin{aligned} \mu(B_k) &= \mu(A_k) + \mu(B_{k+1}) \\ &= \mu(A_k) + \mu(A_{k+1}) + \mu(B_{k+2}) = \cdots = \sum_{i \geq k} \mu(A_i). \end{aligned}$$

This implies

$$\begin{aligned} 1 = \mu(\bigcup_{k=1}^{\infty} B_k) &= \sum_{k=1}^{\infty} \mu(B_k) = \sum_{k=1}^{\infty} \sum_{i \geq k} \mu(A_i) \\ &= \sum_{i=1}^{\infty} \sum_{k=1}^{i} \mu(A_i) = \sum_{i=1}^{\infty} i \cdot \mu(A_i) = \int_A n d\mu. \end{aligned}$$

□

Some details of this proof are further discussed in Problem 3.2.10.

Theorem 3.2.5. *If μ_1 and μ_2 are two different normalized τ-ergodic measures, then $\mu_1 \perp \mu_2$ (μ_1 and μ_2 are mutually singular).*

Proof. Let $\mu = \frac{\mu_1+\mu_2}{2}$. Since $\mu_1 << \mu$ and $\mu_2 << \mu$, there exist integrable functions f_1, f_2 such that $\mu_1 = f_1 \cdot \mu$ and $\mu_2 = f_2 \cdot \mu$, respectively. Let us define $A_1 = \{x : f_1(x) > 0\}$ and $A_2 = \{x : f_2(x) > 0\}$. We will prove that A_1 and A_2 are almost invariant (i.e., $\mu(A_i \Delta \tau^{-1}(A_i)) = 0$, $i = 1, 2$). We have $\mu_1(\tau^{-1}(A_1) \setminus A_1) = 0$ (since $\mu_1(A_1) = 1$) and $\mu_1(\tau^{-1}(A_1)) = \mu_1(A_1)$. Thus,

$$\mu_1(A_1 \setminus \tau^{-1}(A_1)) = \mu_1(A_1) + \mu_1(\tau^{-1}(A_1) \setminus A_1) - \mu_1(\tau^{-1}(A_1)) = 0.$$

Hence, $\mu_1(A_1 \triangle \tau^{-1}(A_1)) = 0$ and there exists a τ-invariant set $\tilde{A}_1$ such that $\mu_1(A_1 \triangle \tilde{A}_1) = 0$. Similarly, there exists a τ-invariant set $\tilde{A}_2$ such that $\mu_2(A_2 \triangle \tilde{A}_2) = 0$. Then, the set $A = \tilde{A}_1 \cap \tilde{A}_2$ is also τ-invariant. Since μ_1 is ergodic, $\mu_1(A)$ equals either 1 or 0. Similarly, $\mu_2(A)$ equals either 1 or 0. Let us consider all four possible pairs of values $\mu_1(A)$, $\mu_2(A)$. In three of the possible cases, 0 and 0, 0 and 1, 1 and 0, the measures μ_1, μ_2 are mutually singular. In the fourth case, $\mu_1(A) = \mu_2(A) = 1$ and we have

$$\mu_1 = f \cdot \mu_2,$$

where $f = \frac{f_1}{f_2}$ a.e. μ_2. Using Lemma 3.2.5 we obtain $\mu_1 = \mu_2$, which contradicts the assumption $\mu_1 \neq \mu_2$. □

Lemma 3.2.5. *If μ is a normalized τ-ergodic measure and $\mu_1 << \mu$ is a normalized τ-invariant measure, then $\mu_1 = \mu$.*

Proof. Since $\mu_1 << \mu$, there exists an integrable function $f \geq 0$, such that $\mu_1 = f \cdot \mu$. Since both μ_1 and μ are τ-invariant, for any measurable set B, we have

$$\int_B f d\mu = \mu_1(B) = \mu_1(\tau^{-1}(B)) = \int_{\tau^{-1}(B)} f d\mu = \int_B f \circ \tau d\mu.$$

Hence, $f = f \circ \tau$ μ-a.e. Since μ is ergodic, f is constant a.e. Since both measures are normalized, $f \equiv 1$, a.e., and $\mu_1 = \mu$. □

Let us consider the set $\mathfrak{M}_1$ of normalized τ-invariant measures. It is easy to show that $\mathfrak{M}_1$ is convex. We will prove that the extremal points of $\mathfrak{M}_1$ are precisely the τ-ergodic measures.

Theorem 3.2.6. *The extremal points of* $\mathfrak{M}_1$ *are ergodic measures.*

Proof. Let $\mu \in \mathfrak{M}_1$ be a τ-ergodic measure. Let us assume that $\mu = \alpha\mu_1 + \beta\mu_2$, where $\mu_1, \mu_2 \in \mathfrak{M}_1$, $\alpha, \beta \geq 0$, and $\alpha + \beta = 1$. Then $\mu_1 << \mu$ and $\mu_2 << \mu$ and by Lemma 3.2.5 we have $\mu_1 = \mu$, $\mu_2 = \mu$. Thus, any ergodic measure is an extremal point of $\mathfrak{M}_1$.

Now, let us assume that μ is an extremal point of $\mathfrak{M}_1$. We will show that μ is ergodic. If μ is not ergodic, then there exists an invariant set A with $0 < \mu(A) < 1$. Its complement A^c is also invariant and $0 < \mu(A^c) < 1$. Let us define $\mu_1 = \frac{1}{\mu(A)} \cdot \chi_A \cdot \mu$ and $\mu_2 = \frac{1}{\mu(A^c)} \cdot \chi_{A^c} \cdot \mu$. Both $\mu_1, \mu_2 \in \mathfrak{M}_1$ and we have

$$\mu = \mu(A) \cdot \mu_1 + \mu(A^c) \cdot \mu_2,$$

which contradicts the assumption that μ is an extremal point of $\mathfrak{M}_1$. □

If X is a compact metric space, then $\mathfrak{M}_1$ is compact in the weak topology. Let us denote by $Ex(\mathfrak{M}_1)$ the set of extremal points of $\mathfrak{M}_1$, i.e., the set of ergodic measures. By the Krein–Milman Theorem [Dunford and Schwartz, 1964], there exists a measure M on $Ex(\mathfrak{M}_1)$ such that for any $\mu \in \mathfrak{M}_1$, there exists a function f_μ on $Ex(\mathfrak{M}_1)$ such that

$$\mu = \int_{Ex(\mathfrak{M}_1)} f_\mu dM. \tag{3.2.4}$$

The representation (3.2.4) is called the ergodic decomposition of μ. In a simpler case, where $Ex(\mathfrak{M}_1)$ is countable, we have for any $\mu \in \mathfrak{M}_1$,

$$\mu = \sum_{i=1}^{\infty} \alpha_i \mu_i,$$

where $Ex(\mathfrak{M}_1) = \{\mu_i\}_{i=1}^{\infty}$, and $\alpha_i \in \mathbb{R}$, $i = 1, 2, \ldots$.

Example 3.2.3. Let $I = [0, 1]$ and $\tau(x) = x$ be the identity on I. Any measure μ is τ-invariant. The only ergodic measures are the Dirac measures $\delta_x, x \in I$. Thus, any measure can be represented as an integral over Dirac's measures. This example shows that ergodic measures in the ergodic decomposition of μ may have properties completely different from those of μ itself. For example, when we write an ergodic decomposition of an invariant measure absolutely continuous with respect to λ we cannot be sure that the components are also absolutely continuous with respect to λ.

Let X be a compact metric space and let $\tau : X \to X$ be measurable. For any $x \in X$, we denote by $\omega(x)$ a set of accumulation points of the orbit of $x : \{\tau^n(x)\}_{n=0}^{\infty}$.

Definition 3.2.3. A point $x \in X$ is called τ-*recurrent* if and only if $x \in \omega(x)$, i.e., there exists a strictly increasing sequence of positive integers $\{n_i\}_{i=1}^{\infty}$, such that

$$x = \lim_{i \to +\infty} \tau^{n_i}(x).$$

We denote by R_τ the set of all τ-recurrent points.

For example: every fixed or periodic point is recurrent and for an irrational rotation of the circle every point is recurrent.

The following theorem is a topological counterpart of the Poincaré Recurrence Theorem.

Theorem 3.2.7. *For any τ-invariant finite measure μ, $\mu(X \setminus R_\tau) = 0$, i.e., any τ-invariant measure is supported on the set of τ-recurrent points.*

Proof. Let $\{B_n\}_{n \geq 0}$ be a basis of open balls covering X with diameters tending to 0 as $n \to +\infty$ and such that $\bigcup_{n>N} B_n = X$, for any $N > 0$. Let μ be a τ-invariant measure. By the Poincaré Recurrence Theorem, we can find sets $\tilde{B}_n \subset B_n$ of points returning infinitely many times to B_n with $\mu(B_n \setminus \tilde{B}_n) = 0,\ n = 1, 2, \dots$. If $\mu(B_n) = 0$, then we define $\tilde{B}_n = \emptyset$. Let

$$\tilde{X} = \bigcap_{N \geq 0} \bigcup_{n \geq N} \tilde{B}_n.$$

Then we have

$$\begin{aligned}
\mu(X \setminus \tilde{X}) &= \mu\Big(\bigcup_{N \geq 0} \bigcup_{n \geq N} B_n \setminus \bigcap_{N \geq 0} \bigcup_{n \geq N} \tilde{B}_n\Big) \\
&= \mu\Big(\Big(\bigcup_{n \geq 0} \bigcup_{n \geq N} B_n\Big) \cap \Big(\bigcup_{N \geq 0} \bigcap_{n \geq N} (X \setminus \tilde{B}_n)\Big)\Big) \\
&\leq \mu\Big(\bigcup_{N \geq 0} \bigcup_{n \geq N} (B_n \setminus \tilde{B}_n)\Big) = 0.
\end{aligned}$$

Let $x \in \tilde{X}$. Then for any $\varepsilon > 0$ we can find a ball B_{n_0} with radius smaller than ε, such that $x \in \tilde{B}_{n_0}$ and therefore some image $\tau^{k_0}(x) \in B_{n_0}$ and $\rho(x, \tau^{k_0}) < 2\varepsilon$, where ρ is the metric on X. Since $\varepsilon > 0$ was arbitrary and k_0 can be chosen arbitrarily large, $x \in R_\tau$. Thus, we have $\tilde{X} \subset R_\tau$ and $\mu(X \setminus R_\tau) = 0$. $\square$

Another important set of points is the set of nonwandering points $-_{\tau}$.

Definition 3.2.4. A point $x \in X$ is called *nonwandering* if and only if, for any neighborhood U of x, there exists an $n > 1$ such that $U \cap \tau^{-n}(U) \neq \emptyset$. The set of all nonwandering points is denoted by $-_{\tau}$.

Theorem 3.2.8. *For a measurable transformation $\tau : X \to X$, $R_\tau \subset -_{\tau}$. Thus, for any finite τ-invariant measure μ,*

$$\mu(X \setminus -_{\tau}) = 0. \tag{3.2.5}$$

Proof. Let $x \in R_\tau$. Then there exists a sequence of positive integers $n_i \to +\infty$ such that $\rho(x, \tau^{n_i}(x)) \to 0$, as $i \to +\infty$. For any neighborhood U of x, we can find n_i such that $\tau^{n_i} x \in U$. Thus, $\tau^{-n_i}(U) \cap U \neq \emptyset$. This proves that $R_\tau \subset -_{\tau}$. (3.2.5) follows by Theorem 3.2.7. □

3.3 The Birkhoff Ergodic Theorem

Let $\tau : (X, \mathfrak{B}, \mu) \to (X, \mathfrak{B}, \mu)$ be measure preserving and $E \in \mathfrak{B}$. For $x \in X$, a question of physical interest is: With what frequency do the points of the orbit $\{x, \tau(x), \tau^2(x), ...\}$ occur in the set E?

Clearly, $\tau^i(x) \in E$ if and only if $\chi_E(\tau^i(x)) = 1$. Thus, the number of points of $\{x, \tau(x), ..., \tau^{n-1}(x)\}$ in E is equal to $\sum_{k=0}^{n-1} \chi_E(\tau^k(x))$, and the relative frequency of elements of $\{x, \tau(x), ..., \tau^{n-1}(x)\}$ in E equals $\frac{1}{n}\sum_{k=0}^{n-1} \chi_E(\tau^k(x))$.

The first major result in ergodic theory was proved in 1931 by G.D. Birkhoff [Birkhoff, 1931].

Theorem 3.3.1. *Suppose $\tau : (X, \mathfrak{B}, \mu) \to (X, \mathfrak{B}, \mu)$ is measure preserving, where $(X, \mathfrak{B}, \mu)$ is σ-finite, and $f \in \mathfrak{L}^1(\mu)$. Then there exists a function $f^* \in \mathfrak{L}^1(\mu)$ such that*

$$\frac{1}{n}\sum_{k=0}^{n-1} f(\tau^k(x)) \to f^*, \ \mu - a.e.$$

Furthermore, $f^* \circ \tau = f^*$ $\mu - a.e.$ and if $\mu(X) < \infty$, then $\int_X f^* d\mu = \int_X f d\mu$.

There are different proofs of the Birkhoff Ergodic Theorem, (see [Halmos, 1956], [Cornfeld, Fomin and Sinai 1982], [Rudolph, 1990], [Krengel, 1985]). Our presentation is based on [Randolph, 1968], which is closer to Birkhoff's original proof. It is our belief that this proof is more intuitive and can be easily grasped by novices in ergodic theory.

We present some lemmas first.

Definition 3.3.1. Given a real sequence $x_0, x_1, \dots, x_{n-1}$ of fixed length n, a term x_j is called a vit (very important term) if at least one of the sums

$$\begin{aligned}
&x_j \\
&x_j + x_{j+1} \\
&\vdots \\
&x_j + x_{j+1} + \cdots + x_{n-1}
\end{aligned}$$

is positive (i.e., strictly greater than 0).

Example 3.3.1. Let $n = 5$ and let the sequence be: $-1, 1, -\frac{1}{2}, -\frac{2}{3}, 1$. The vits are: $-1, 1, -\frac{2}{3}, 1$. $x_2 = -\frac{1}{2}$ is not a vit since $-\frac{1}{2} < 0, -\frac{1}{2} + (-\frac{2}{3}) < 0$ and $-\frac{1}{2} + (-\frac{2}{3}) + 1 < 0$.

Lemma 3.3.1. *In any finite sequence the sum of vits is greater than or equal to* 0. *(If there are no vits we assume that their sum is* 0*).*

Proof. We will use induction on the lengths of sequences. Let ν denote the sum of vits. For any sequence of length 1, the lemma holds since either $x_0 \geq 0$ and $\nu = x_0 \geq 0$ or $x_0 < 0$ in which case there are no vits and $\nu = 0$. Let $n \geq 2$ and let us assume that the lemma holds for any sequence of length $\leq n-1$. Take any sequence of length n,

$$x_0, x_1, x_2, \dots, x_{n-1}. \tag{3.3.1}$$

Form the sums

$$\begin{aligned}
s_0 &= x_0 \\
s_1 &= x_0 + x_1 \\
&\vdots \\
s_{n-1} &= x_0 + x_1 + \cdots + x_{n-1}.
\end{aligned}$$

We consider three cases:

Case 1: All $s_k \leq 0$. Then x_0 is not a vit of the sequence $x_0, x_1, x_2, \dots$. Hence all vits of (3.3.1), if there are any, are vits of the $(n-1)$-length

sequence $x_1, x_2, \dots, x_{n-1}$. Then $\nu \geq 0$ by the inductive assumption.
Case 2: $s_0 > 0$. Then $x_0 > 0, x_0$ is a vit of (3.3.1), and any other vit of (3.3.1) is also a vit of $x_1, x_2, \dots, x_{n-1}$. Hence $\nu \geq x_0 > 0$.
Case 3: $s_0 \leq 0,\ s_1 \leq 0, \dots, s_{k-1} \leq 0$ but $s_k > 0$ with $1 \leq k \leq n-1$. Then,

$$\begin{aligned}
&0 < s_k = s_{k-1} + x_k \\
&0 < s_k = s_{k-2} + (x_k + x_{k-1}) \\
&\vdots \\
&0 < s_k = s_0 + (x_k + x_{k-1} + \cdots + x_1)
\end{aligned}$$

Hence, $x_1 + x_2 + \cdots + x_{k-1} + x_k,\ x_2 + x_3 + \cdots + x_k, \dots, x_{k-1} + x_k$ are all positive and $x_1, x_2, \dots, x_{k-1}$ are all vits of (3.3.1). Also, x_0 is a vit since $0 < s_k = x_0 + (x_k + x_{k-1} + \cdots + x_1)$. The vits $x_0, x_1, x_2, \dots, x_{k-1}$ have $s_k > 0$ as their sum. If there are any vits of (3.3.1) other than these, then the others are also vits of the shorter sequence $x_{k+1}, \dots, x_{n-1}$, whose sum is greater than or equal to 0 by the inductive hypothesis. Hence, $\nu \geq 0$.

$\square$

For a function $f \in \mathfrak{L}^1(X, \mathfrak{B}, \mu)$ and a μ-preserving transformation τ, we define the ergodic averages

$$A_n(f) = \frac{1}{n} \sum_{i=0}^{n-1} f(\tau^i),$$

$n = 1, 2, \dots$.

Lemma 3.3.2. *(Maximal Ergodic Theorem) Let τ be a transformation preserving the measure μ, and $f \in \mathfrak{L}^1(X, \mathfrak{B}, \mu)$. The set*

$$M = \{x : \sup_{n \geq 1} A_n(f)(x) > 0\}$$

is measurable and

$$\int_M f d\mu \geq 0.$$

More generally, for any $\alpha \in \mathbb{R}$ and $M^\alpha = \{x : \sup_{n \geq 1} A_n(f)(x) > \alpha\}$, we have

$$\int_{M^\alpha} f d\mu \geq \alpha \mu(M^\alpha).$$

Proof. For $k = 1, 2, \dots$, let M_k be defined by

$$M_k = \{x : \sup_{1 \leq n \leq k} A_n(f)(x) > 0\}.$$

Obviously,

$$M_k = \{x : \sup_{1 \le n \le k} \sum_{i=0}^{n-1} f(\tau^i(x)) > 0\}.$$

Thus, $x \in M_k$ if and only if at least one of

$$\begin{aligned} &f(x) \\ &f(x) + f(\tau(x)) \\ &\vdots \\ &f(x) + f(\tau(x)) + \cdots + f(\tau^{k-1}(x)) \end{aligned} \tag{3.3.2}$$

is positive, i.e., if and only if $f(x)$ is a vit of the sequence $f(x)$, $f(\tau(x))$, $\ldots$, $f(\tau^{k-1}(x))$. Since all $f, f \circ \tau, \ldots, f \circ \tau^{k-1}$ are measurable, M_k is a measurable set. Moreover,

$$M_1 \subset M_2 \subset M_3 \subset \cdots \subset M_k \subset M_{k+1} \subset \ldots$$

and $M = \bigcup_{k=1}^{\infty} M_k$ so M is also a measurable set and

$$\int_M f d\mu = \lim_{k \to +\infty} \int_{M_k} f d\mu$$

by the Monotone Convergence Theorem applied to positive and negative parts of f, f^+ and f^-. The convergence of this sequence implies the convergence of its arithmetic means

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=1}^{n} \int_{M_k} f d\mu = \int_M f d\mu.$$

Therefore the first part of the lemma will be proved upon showing that

$$\sum_{k=1}^{n} \int_{M_k} f d\mu \ge 0, \ n = 1, 2, 3, \ldots. \tag{3.3.3}$$

The sum in (3.3.3) is equal to

$$\int_{M_n} f(x) d\mu + \int_{M_{n-1}} f(x) d\mu + \cdots + \int_{M_1} f(x) d\mu = \int_{\tau^{-1}(M_n)} f(\tau(x)) d\mu$$
$$+ \int_{\tau^{-1}(M_{n-1})} f(\tau(x)) d\mu + \cdots + \int_{\tau^{-(n-1)}(M_1)} f(\tau^{n-1}(x)) d\mu.$$

By setting $N_j = \tau^{-j}(M_{n-j}),\ j = 0, 1, \dots, n-1$, we can write

$$\sum_{k=1}^{n} \int_{M_k} f(x)d\mu = \sum_{j=0}^{n-1} \int_{N_j} f(\tau^j x)d\mu = \sum_{j=0}^{n-1} \int_X f(\tau^j x)\chi_{N_j}(x)d\mu$$
$$= \int_X \sum_{j=0}^{n-1} f(\tau^j x)\chi_{N_j}(x)d\mu. \tag{3.3.4}$$

We will establish (3.3.3) by showing that the integrand in (3.3.4) is nonnegative for all $x \in X$. The point $x \in M_k$ if and only if at least one of the sums in (3.3.2) is positive and $x \in M_{n-j},\ j = 1, \dots, (n-1)$ if and only if at least one of

$$\begin{aligned}
&f(x)\\
&f(x) + f(\tau(x))\\
&\vdots\\
&f(x) + f(\tau(x)) + \cdots + f(\tau^{n-j-1}(x))
\end{aligned}$$

is positive. Since $N_j = \tau^{-j}(M_{n-j}),\ x \in N_j$ if and only if $\tau^j(x) \in M_{n-j}$ and hence if and only if at least one of

$$\begin{aligned}
&f(\tau^j(x)\\
&f(\tau^j(x)) + f(\tau^{j+1}(x))\\
&\vdots\\
&f(\tau^j(x)) + f(\tau^{j+1}(x)) + \cdots + f(\tau^{n-1}(x))
\end{aligned}$$

is positive. Stated in terms of vits: $x \in N_j,\ j = 0, 1, \dots, n-1$, if and only if $f(\tau^j(x))$ is a vit of the sequence

$$f(x), f(\tau(x)), f(\tau^2(x)), \dots, f(\tau^{n-1}(x)). \tag{3.3.5}$$

Consider any $x \in X$ and any j among $0, 1, 2, \dots, n-1$. Either $x \notin N_j$, in which case $\chi_{N_j}(x) = 0$, or else $x \in N_j$ in which case $f(\tau^j(x))\chi_{N_j}(x) = f(\tau^j(x))$ is a vit of (3.3.5). Hence, for any $x \in X$, the sequence (3.3.5) has its sum of vits $\nu(x)$ given by

$$\nu(x) = \sum_{j=0}^{n-1} f(\tau^j(x))\chi_{N_j}(x),$$

which is nonnegative by virtue of Lemma 3.3.1. Since this sum is the integrand in (3.3.4), we have established (3.3.3) and proved the first part of the lemma. The second part of the lemma follows upon replacing f in the first part by $f - \alpha$. $\square$

Lemma 3.3.3. *Let τ be a transformation preserving the measure μ and let $f \in \mathfrak{L}^1(X, \mathfrak{B}, \mu)$. Let $\alpha \in \mathbb{R}$. If A is a measurable invariant set such that for each $x \in A$,*

$$(a) \quad \sup_{n\geq 1} A_n(f)(x) > \alpha, \quad \textit{then} \int_A f d\mu \geq \alpha\mu(A);$$

$$(b) \quad \inf_{n\geq 1} A_n(f)(x) < \alpha, \quad \textit{then} \int_A f d\mu \leq \alpha\mu(A).$$

Proof. First consider the special case $\alpha = 0$. Let $g = f \cdot \chi_A$ and

$$M = \{x : \sup_{n\geq 1} \frac{1}{n} \sum_{i=0}^{n-1} g(\tau^i(x)) > 0\}.$$

By Lemma 3.3.2, $\int_M g d\mu \geq 0$. The characteristic function χ_A is invariant, i.e., $\chi_A = \chi_A \circ \tau$. Thus, $g(\tau^i) = f(\tau^i)\chi_A$ so

$$g(\tau^i(x)) = \begin{cases} f(\tau^i(x)), & \text{a.e. in } A \\ 0, & \text{a.e. in } X \backslash A. \end{cases}$$

Hence,

$$\frac{1}{n} \sum_{i-0}^{n-1} g(\tau^i(x)) = \begin{cases} A_n(f)(x), & \text{a.e. in } A \\ 0, & \text{a.e. in } X \backslash A. \end{cases}$$

Therefore, $M \subset A$ differs from A by a set of measure zero and

$$0 \leq \int_M g d\mu = \int_A g d\mu = \int_A f d\mu,$$

which proves (a) under the special case $\alpha = 0$. In case α is not zero, then under the hypothesis in (a),

$$\sup_{n\geq 1} \frac{1}{n} \sum_{i=0}^{n-1} (f(\tau^i(x)) - \alpha) = \sup_{n\geq 1} (A_n(f)(x) - \alpha) > 0$$

and

$$0 \leq \int_A (f(x) - \alpha) d\mu = \int_A f d\mu - \alpha\mu(A)$$

by the above special case, with f replaced by $f - \alpha$. Hence (a) holds. Now apply (a) with f replaced by $-f$ and obtain (b). $\square$

Proof of Theorem 3.3.1. First we prove that the limit

$$f^*(x) = \lim_{n\to+\infty} A_n(f)(x)$$

exists μ-a.e. Let

$$A^-(x) = \liminf_{n\to+\infty} A_n(f)(x)$$

and

$$A^+(x) = \limsup_{n\to+\infty} A_n(f)(x).$$

Both A^- and A^+ are measurable functions. Also,

$$\begin{aligned} A^-(\tau(x)) &= \liminf_{n\to+\infty} A_n(f)(\tau(x)) \\ &= \liminf_{n\to+\infty}\Big(-\frac{f(\tau^n(x))}{n} + (\frac{n+1}{n})\frac{1}{n+1}\sum_{i=0}^{n} f(\tau^i(x))\Big) \\ &= \liminf_{n\to+\infty} A_{n+1}(f)(x) = A^-(x). \end{aligned}$$

Thus, A^- is τ-invariant and in the same way A^+ is τ-invariant. For any constants $a < b$, the set

$$A_{a,b} = \{x : A^-(x) < a < b < A^+(x)\}$$

is a measurable set. Also,

$$\begin{aligned} \tau^{-1}(A_{ab}) &= \{x : \tau(x) \in A_{ab}\} \\ &= \{x : A^-(\tau(x)) < a < b < A^+(\tau(x))\} = A_{ab} \end{aligned}$$

since A^-, A^+ are τ-invariant. Moreover,

$$A_{ab} \subset \{x : \inf_{n\ge 1} A_n(f)(x) < a < b < \sup_{n\ge 1} A_n(f)(x)\}.$$

Hence, by Lemma 3.3.3,

$$\int_{A_{ab}} f d\mu \le a\mu(A_{ab}) \le b\mu(A_{ab}) \le \int_{A_{ab}} f d\mu.$$

Consequently, $\mu(A_{ab}) = 0$, for all $a < b$. Thus, $\mu(A) = 0$, where

$$A = \{x : A^-(x) < A^+(x)\} = \bigcup_{a,b\le \mathcal{Q}} \{x : A^-(x) < a < b < A^+(x)\}.$$

Thus, $A^-(x) = A^+(x)$ μ-a.e. and $f^*(x) = \lim_{n\to+\infty} A_n(f)(x)$ exist μ-a.e. and is τ-invariant.

Now, we prove that $f^* \in \mathfrak{L}^1(X, \mathfrak{B}, \mu)$. We have $\|A_n(|f|)\|_1 = \||f|\|_1,\ f \in \mathfrak{L}^1(X, \mathfrak{B}, \mu)$ and by Fatou's Lemma,

$$\int_X \lim_{n\to+\infty} A_n(|f|)d\mu \leq \lim_{n\to+\infty} \int_X A_n(|f|)d\mu = \int_X |f|d\mu.$$

Since $|A_n(f)| \leq A_n(|f|),\ n = 1, 2, \ldots$, we have

$$\int_X |f^*|d\mu \leq \int_X |f|d\mu,$$

which proves that $f^* \in \mathfrak{L}^1(X, \mathfrak{B}, \mu)$. Furthermore, it proves that the operator $f \longmapsto f^*$ is a contraction on $\mathfrak{L}^1(X, \mathfrak{B}, \mu)$.

The last fact we have to prove is that $\int_X f^* d\mu = \int_X f d\mu$, assuming μ is a finite measure. We have

$$\int_X A_n(f)d\mu = \int_X f d\mu, \tag{3.3.6}$$

for $n = 1, 2, \ldots$. Since $f^* = \lim_{n\to+\infty} A_n(f)$ μ-a.e., for any bounded function f, Lebesgue Dominated Convergence Theorem and (3.3.6) imply

$$\int_X f^* d\mu = \int f d\mu,$$

for any bounded function. Let $f \in \mathfrak{L}^1(X, \mathfrak{B}, \mu)$ be arbitrary. For any $\epsilon > 0$, we can find a bounded function f_B such that $\|f - f_B\|_1 < \epsilon$. Then $\|(f - f_B)^*\|_1 < \epsilon$ and we have,

$$\begin{aligned}
|\int_X f^* d\mu - \int_X f d\mu| &= |\int_X f_B^* + (f - f_B)^* d\mu - \int_X f_B + (f - f_B) d\mu| \\
&\leq |\int_X f_B^* d\mu - \int_X f_B d\mu| + \|(f - f_B)^*\|_1 + \|(f - f_B)\|_1 \\
&\leq 0 + 2\epsilon.
\end{aligned}$$

Since $\epsilon > 0$ is arbitrary,

$$\int_X f^* d\mu = \int_X f d\mu.$$

□

Corollary 3.3.1. *If τ is ergodic, then f^* is constant $\mu - a.e.$ and if $\mu(X) < \infty$, then*

$$f^* = \frac{1}{\mu(X)} \int_X f d\mu \quad \text{a.e.}$$

Thus, if $\mu(X) = 1$ *and* τ *is ergodic, we have*

$$\frac{1}{n}\sum_{i=0}^{n-1} \chi_E(\tau^i(x)) \to \mu(E), \quad \mu - a.e.,$$

and thus the orbit of almost every point of X *occurs in the set* E *with asymptotic relative frequency* $\mu(E)$*.*

We define the *time average* of $f \in \mathfrak{L}^1(\mu)$ to be

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(\tau^i(x)),$$

and the *space average* of f to be

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

If τ is ergodic, Corollary 3.3.1 states that these averages are equal. The converse is also true, i.e., if the time average equals the space average, then τ is ergodic.

Example 3.3.2. Let $\tau(x) = 10 \cdot x \pmod 1$, $x \in [0,1]$. τ preserves Lebesgue measure λ and (τ, λ) is ergodic. Let $i = 0, 1, \ldots, 9$ and $A_i = [\frac{i}{10}, \frac{i+1}{10})$. By the Birkhoff Ergodic Theorem,

$$\frac{1}{n}\sum_{k=0}^{n-1} \chi_{A_i}(\tau^k x) \to \frac{1}{10},$$

for almost every $x \in [0,1]$. This proves the following famous result:

Theorem 3.3.2. *(Borel Normal Number Theorem) For almost every* $x \in [0,1]$ *(with respect to Lebesgue measure), the frequency of any digit in the decimal expansion of* x *is* $\frac{1}{10}$*, i.e., almost every* $x \in [0,1]$ *is a normal number.*

Since for any real number $x > 0$ we can find an $n \geq 1$ such that $x \cdot 10^{-n} \in [0,1]$, the Borel Normal Number Theorem holds for almost all real numbers. It remains an open question as to whether π is a normal number.

Corollary 3.3.2. *Let* $(X, \mathfrak{B}, \mu)$ *be a normalized measure space and let* $\tau : X \to X$ *be measure preserving. Then* τ *is ergodic if and only if* *for all* $A, B \in \mathfrak{B}$,

$$\frac{1}{n}\sum_{i=0}^{n-1}\mu(\tau^{-i}A\cap B)\to\mu(A)\mu(B).$$

Corollary 3.3.3. *$\mathfrak{L}^p$-Ergodic Theorem (J. von Neumann).*

Let $1 \le p < \infty$ and let τ be measure preserving on the normalized measure space $(X, \mathfrak{B}, \mu)$. If $f \in \mathfrak{L}^p(\mu)$, then there exists $f^* \in \mathfrak{L}^p(\mu)$ such that $f^*\circ\tau = f^*$ μ-a.e. and $\|\frac{1}{n}\sum_{i=0}^{n-1} f(\tau^i x) - f^*(x)\|_p \to 0$ as $n \to \infty$.

Proof. Let us fix $1 \le p < +\infty$ and $f \in \mathfrak{L}^p(X, \mathfrak{B}, \mu)$. Since $\|A_n(f)\|_p \le \|f\|_p$, we have by Fatou's Lemma,

$$\int_X |f^*|^p d\mu \le \liminf_{n\to+\infty}\int_X |A_n(f)|^p d\mu \le \int_X |f|^p d\mu.$$

Hence, the operator $L : \mathfrak{L}^p \to \mathfrak{L}^p$ defined by $L(f) = f^*$ is a contraction on $\mathfrak{L}^p(X, \mathfrak{B}, \mu)$. Since $\|f^* - A_n(f)\|_p^p = \int_X |f^* - A_n(f)|^p d\mu$ and by the Birkhoff Ergodic Theorem $A_n(f) \to f^*$ μ-a.e., $\|f^* - A_n(f)\|_p \to 0$, as $n \to +\infty$ for any bounded function $f \in \mathfrak{L}^p(X, \mathfrak{B}, \mu)$. Let $f \in \mathfrak{L}^p(X, \mathfrak{B}, \mu)$ be a function, not necessarily bounded. For any $\epsilon > 0$ we can find a bounded function $f_B \in \mathfrak{L}^p(X, \mathfrak{B}, \mu)$ such that $\|f - f_B\|_p < \epsilon$. Then, since L is a contraction on $\mathfrak{L}^p(X, \mathfrak{B}, \mu)$, we have

$$\begin{aligned}
&\|f^* - A_n(f)\|_p \\
&\qquad = \|f_B^* + (f - f_B)^* - A_n(f_B) - (A_n(f) - A_n(f_B))\|_p \\
&\qquad \le \|f_B^* - A_n(f_B)\|_p + \|A_n(f) - A_n(f_B)\|_p + \|(f - f_B)^*\|_p \\
&\qquad \le \|f_B^* - A_n(f_B)\|_p + 2\epsilon,
\end{aligned}$$

which can be made arbitrarily small. □

3.4 Mixing and Exactness

Recall that τ is ergodic if and only if for all $A, B \in \mathfrak{B}$,

$$\frac{1}{n}\sum_{i=0}^{n-1}\mu(\tau^{-i}A\cap B)\to\mu(A)\mu(B)\ \text{ as } n\to+\infty.$$

Definition 3.4.1. We say $\tau : (X, \mathfrak{B}, \mu) \to (X, \mathfrak{B}, \mu)$ is *weakly mixing* if for all $A, B \in \mathfrak{B}$,

$$\frac{1}{n}\sum_{i=0}^{n-1}|\mu(\tau^{-i}A\cap B)-\mu(A)\mu(B)|\to 0 \quad \text{as } n\to+\infty.$$

τ is *strongly mixing* if for all $A, B \in \mathfrak{B}$,

$$\mu(\tau^{-n}A \cap B) \to \mu(A)\mu(B) \quad \text{as } n \to +\infty.$$

τ is mixing of multiplicity $r \geq 1$, if for any $B, A_1, A_2, ..., A_r \in \mathfrak{B}$:

$$\mu(\tau^{-n_1}A_1 \cap \tau^{-n_2}A_2 \cap .... \cap \tau^{-n_r}A_r \cap B) \to \mu(A_1)\mu(A_2)...\mu(A_r)\mu(B)$$

as $n_1, n_2, ..., n_r \to +\infty$ and $|n_i - n_j| -> \infty, \quad i \neq j$.

Obviously τ mixing of multiplicity $r \Rightarrow \tau$ strongly mixing $\Rightarrow \tau$ weakly mixing $\Rightarrow \tau$ ergodic. Examples of τ ergodic but not weakly mixing and τ weakly but not strongly mixing are known. There are no known examples of τ strongly mixing but not mixing of multiplicity $r > 1$.

The following result shows it is sufficient to check the convergence properties on an algebra generating $\mathfrak{B}$.

Theorem 3.4.1. *If $\tau : X \to X$ is measure preserving and $\mathcal{P}$ is a π-system generating $\mathfrak{B}$, then*

(i) *τ is ergodic if and only if for all $A, B \in \mathcal{P}$,*

$$\frac{1}{n}\sum_{i=0}^{n-1} \mu(\tau^{-i}A \cap B) \to \mu(A)\mu(B)$$

as $n \to +\infty$.

(ii) *τ is weakly mixing if and only if for all $A, B \in \mathcal{P}$,*

$$\frac{1}{n}\sum_{i=0}^{n-1} |\mu(\tau^{i}A \cap B) - \mu(A)\mu(B)| \to 0 \text{ as } n \to +\infty.$$

(iii) *τ is strongly mixing if and only for all $A, B \in \mathcal{P}$,*

$$\mu(\tau^{-n}A \cap B) \to \mu(A)\mu(B) \text{ as } n \to +\infty.$$

Remark 3.4.1. The strong mixing of τ means that any set $B \in \mathfrak{B}$ under the action of τ, becomes asymptotically independent of a fixed set $A \in \mathfrak{B}$. The weak mixing of τ means that B becomes independent of A if we neglect a finite number of initial iterations. The ergodicity of τ means B becomes independent of A on the average.

We will now express the foregoing concepts in functional form. For that purpose, it is convenient to use the Koopman operator.

Definition 3.4.2. Let $\tau : (X, B, \mu) \to (X, B, \mu)$ be a measurable transformation. The operator $U_\tau : \mathfrak{L}^\infty \to \mathfrak{L}^\infty$ defined by

$$U_\tau f = f \circ \tau$$

is called the *Koopman operator.* It is easy to see that U_τ is well defined and that $\|U_\tau f\|_\infty \leq \|f\|_\infty$ for any $f \in \mathfrak{L}^\infty$. Usually the Koopman operator is defined as an operator on $\mathfrak{L}^2$, but for our purposes it is more convenient to define it on $\mathfrak{L}^\infty$.

Theorem 3.4.2. *Let $(X, \mathfrak{B}, \mu)$ be a normalized measure space and let $\tau : X \to X$ be measure preserving. Then*
(a) *τ is ergodic if and only if for all $f \in \mathfrak{L}^1, g \in \mathfrak{L}^\infty$*

$$\frac{1}{n}\sum_{k=0}^{n-1}\int_X f(U_\tau^k g)d\mu \to \int_X f d\mu \int_X g d\mu$$

as $n \to +\infty$.
(b) *τ is weakly mixing if and only if for all $f \in \mathfrak{L}^1, g \in \mathfrak{L}^\infty$,*

$$\frac{1}{n}\sum_{k=0}^{n-1}\left| \int_X f(U_\tau^k g)d\mu - \int_X f d\mu \int_X g d\mu \right| \longrightarrow 0 \quad \text{as } n \to +\infty.$$

(c) *τ is strongly mixing if and only if for all $f \in \mathfrak{L}^1, g \in \mathfrak{L}^\infty$,*

$$\int_X f(U_\tau^n g)d\mu \longrightarrow \int_X f d\mu \int_X g d\mu$$

as $n \to +\infty$.

There is a notion in ergodic theory that is even stronger than mixing. This is the property of exactness, which was introduced in [Rochlin, 1964].

Definition 3.4.3. Let $(X, \mathfrak{B}, \mu)$ be a normalized measure space and let $\tau : X \to X$ be measure preserving such that $\tau(A) \in \mathfrak{B}$ for each $A \in \mathfrak{B}$. If

$$\lim_{n\to\infty} \mu(\tau^n A) = 1$$

for every $A \in \mathfrak{B}, \mu(A) > 0$, then τ is *exact.*

It can be proved that exactness of τ implies that τ is strongly mixing. The converse is not true in general. Note that if τ is invertible, it cannot be exact, since

$$\mu(\tau A) = \mu(\tau^{-1}\tau A) = \mu(A),\ 0 < \mu(A) < 1,$$

and by induction $\mu(\tau^n A) = \mu(A)$ for all n.

Theorem 3.4.3. *Let $(X, \mathfrak{B}, \mu)$ be a normalized measure space and let $\tau : (X, \mathfrak{B}, \mu) \to (X, \mathfrak{B}, \mu)$ be measure preserving. Then τ is exact if and only if*

$$\mathfrak{B}^T = \bigcap_{n=0}^{\infty} \tau^{-n}(\mathfrak{B})$$

consists of the sets of μ-measure 0 or 1.

Proof. Let us assume that $A \in \mathfrak{B}^T$, $0 < \mu(A) < 1$ and let $B_n \in \mathfrak{B}$ be such that $A = \tau^{-n} B_n$, $n = 1, 2, \ldots$. Since τ preserves μ, we have $\mu(B_n) = \mu(A), n = 1, 2, \ldots$. We also have $\tau^n(A) = \tau^n(\tau^{-n} B_n) \subset B_n$. Hence, $\mu(\tau^n(A)) \leq \mu(A) < 1$ for $n = 1, 2, \ldots$, which contradicts the exactness of τ. Let $A \in \mathfrak{B}$ and $\mu(A) > 0$. If $\lim_{n\to+\infty} \mu(\tau^n A) < 1$, we may assume that for some $a < 1$, $\mu(\tau^n(A)) \leq a < 1$, $n = 1, 2, \ldots$. For any $n \geq 0$ we have $\tau^{-(n+1)}(\tau^{n+1} A) \supset \tau^{-n}(\tau^n A)$. Thus, the set $B = \bigcup_{n=0}^{\infty} \tau^{-n}(\tau^n A)$ belongs to $\mathfrak{B}^T$. Since $B \supset A$ and $\mu(B) \geq \mu(A) > 0$, $\mu(B) = 1$. On the other hand,

$$\mu(B) = \lim_{n\to+\infty} \mu(\tau^{-n}(\tau^n A)) = \lim_{n\to+\infty} \mu(\tau^n(A)) \leq a < 1.$$

□

3.5 The Spectrum of the Koopman Operator and the Ergodic Properties of τ

Let $(X, \mathfrak{B}, \tau, \mu)$ be a dynamical system with a finite measure μ. We recall that the Koopman operator $U_\tau = U : \mathfrak{L}^\infty(X, \mathfrak{B}, \mu) \to \mathfrak{L}^\infty(X, \mathfrak{B}, \mu)$ is defined by

$$Uf = f \circ \tau.$$

It is easy to see that $\|Uf\|_\infty \leq \|f\|_\infty$. Since constant functions are U-invariant, $\|U\|_\infty = 1$ and 1 is always an eigenvalue of U. Since U preserves integrals, all eigenvalues of U have modulus 1. We will study the relation between the spectrum of U and the ergodic properties of τ.

Lemma 3.5.1. *Let (τ, μ) be ergodic. A number η such that $\eta^k = 1$ is an eigenvalue of U if and only if there exist mutually disjoint sets $C_1, \ldots, C_k \in \mathfrak{B}$ of positive μ-measure such that $\tau^{-1}(C_i) = C_{i+1}$, $i = 1, \ldots, k-1$, and $\tau^{-1}(C_k) = C_1$.*

Proof. $\Rightarrow$ Let f be an eigenfunction of U corresponding to η. Then

$$U(f^k) = f^k \circ \tau = (f \circ \tau)^k = (Uf)^k = \eta^k f^k = f^k$$

and f^k is τ-invariant function. Since τ is ergodic, f^k is constant, which means that f attains at most k different values. We can find a complex number $a_0 \in \mathbb{C}$ such that $C_{a_0} = \{x : f(x) = a_0\}$ is of positive μ-measure. Let $C_1 = C_{a_0}$, and $C_{i+1} = \tau^{-i}(C_1),\ i = 1, \ldots, k-1$. We have $\tau^{-1}(C_k) = \{x : f(x) = \frac{a_0}{\eta^k} = a_0\} = C_1$.

$\Leftarrow$ Let $f = \sum_{i=1}^{k} \eta^{k-i} \chi_{C_i}$. Then $Uf = \sum_{i=1}^{k-1} \eta^{k-i} \chi_{C_{i+1}} + \chi_{C_1} = \eta f$.

□

Theorem 3.5.1. (τ, μ) *is ergodic* $\Leftrightarrow$ 1 *is a simple eigenvalue of* U.

Proof. We have proved that τ is ergodic $\Leftrightarrow$ any measurable τ-invariant function is constant. □

Theorem 3.5.2. *The following conditions are equivalent:*

(i) (τ, μ) *is weakly mixing;*
(ii) (τ, μ) *is ergodic and* 1 *is the only eigenvalue of* U*;*
(iii) every eigenfunction of U *is constant.*

Proof. (i)$\Rightarrow$(ii) Let $\eta \neq 1,\ |\eta| = 1$, be an eigenvalue of U, and let f be an eigenfunction corresponding to η. Let $g \equiv \bar{f}$. We have

$$\begin{aligned}
\lim_{n\to+\infty} \frac{1}{n} \sum_{k=1}^{n-1} & \left| \int_X f \circ \tau^k \bar{f} d\mu - \int_X f d\mu \int_X \bar{f} d\mu \right| \\
&= \lim_{n\to+\infty} \frac{1}{n} \sum_{k=1}^{n-1} \left| \eta^k \int_X |f|^2 d\mu - |\int_X f d\mu|^2 \right| \\
&= \int_X |f|^2 d\mu \cdot \lim_{n\to+\infty} \frac{1}{n} \sum_{k=1}^{n-1} |\eta^k - a| \neq 0,
\end{aligned}$$

where $a = |\int_X f d\mu|^2 / \int_X |f|^2 d\mu \leq 1$ (see Problem 3.5.1). This contradicts (i).

(ii)$\Rightarrow$(iii) The only eigenvalue of U is 1, so any eigenfunction of U is τ-invariant and hence constant.

(iii)$\Rightarrow$(ii) This part requires a deeper proof that uses the general spectral theorem [Dunford and Schwartz, 1964].

Both implications (i)$\Rightarrow$(ii) and (ii)$\Rightarrow$(iii) can be repeated for U extended to $\mathfrak{L}^2(X, \mathfrak{B}, \mu)$. Thus, we know that all eigenfunctions of U on

$\mathfrak{L}^2$ are constant. Let $f \in \mathfrak{L}^2$ be a nonconstant function with $\int_X f d\mu = 0$. Let $g \in \mathfrak{L}^2$. We will show that

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=1}^{n-1} \left| \int f \circ \tau^k \cdot \bar{g} d\mu \right|^2 = 0.$$

Let ν be a measure on the spectrum $\sigma(U)$ defined by

$$\nu(A) = \int (E(A)f)\bar{g} d\mu,$$

for all Borel subsets $A \subset \sigma(U)$, where $E(\cdot)$ is the spectral measure corresponding to the operator U. For any complex number $\eta \in \mathbb{C}$ with $|\eta| = 1$, we have

$$\begin{aligned} U(E(\{\eta\})f) &= \int_{\sigma(U)} z\chi_{\{\eta\}} dE(z) f \\ &= \eta \int_{\sigma(U)} \chi_{\{\eta\}} dE(z) f = \eta E(\{\eta\}) f. \end{aligned}$$

Thus, $E(\{\eta\})f$ is an eigenfunction of U and is constant. We have

$$\begin{aligned} 0 = \int_X E(\{\eta\}) f \cdot \bar{f} d\mu &= \int_X E(\{\eta\})^2 f \cdot \bar{f} d\mu \\ &= \int E(\{\eta\}) f \overline{E(\{\eta\}) f} d\mu, \end{aligned}$$

so $E(\{\eta\})f = 0$, for any $|\eta| = 1$. We have

$$\begin{aligned}
\frac{1}{n}\sum_{k=0}^{n-1}\left|\int_X f\circ\tau^k\cdot\bar{g}d\mu\right|^2 &= \frac{1}{n}\sum_{k=0}^{n-1}\left|\int_{\sigma(U)} z^k d(E(z)f,g)\right|^2\\
&= \frac{1}{n}\sum_{k=0}^{n-1}\left|\int_{\sigma(U)} z^k d\nu(z)\right|^2\\
&= \frac{1}{n}\sum_{k=0}^{n-1}\left(\int_{\sigma(U)} z^k d\nu(z)\right)\left(\int_{\sigma(U)} \bar{w}^k d\bar{\nu}(w)\right)\\
&= \frac{1}{n}\sum_{k=0}^{n-1}\iint_{\sigma(U)\times\sigma(U)} z^k\bar{w}^k d\nu(z)d\bar{\nu}(w)\\
&= \iint_{\sigma(U)\times\sigma(U)} \frac{1}{n}\sum_{k=0}^{n-1} z^k\bar{w}^k d\nu(z)d\bar{\nu}(w)\\
&= \iint_{\sigma(U)\times\sigma(U)} \frac{1}{n}\frac{1-(z\bar{w})^n}{1-z\bar{w}} d\nu(z)d\bar{\nu}(w).
\end{aligned}$$

The last equality holds because $1 - z\bar{w} = 0$ only for $z = w$ and the diagonal $\triangle = \{(zw) \in \sigma(U)\times\sigma(U) : z = w\}$ is of $(\nu\times\nu)$-measure 0, since the measure ν vanishes on points. Since $\frac{1}{n}\frac{1-(z\bar{w})^n}{1-z\bar{w}} \to 0$, $(\nu\times\nu)$-a.e., the Bounded Convergence Theorem yields

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{k=0}^{n-1}\left|\int_X f\circ\tau^k\cdot\bar{g}d\mu\right|^2 = 0.$$

This implies (see Problem 3.5.2) that

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{k=0}^{n-1}\left|\int_X f\circ\tau^k\cdot\bar{g}d\mu\right| = 0.$$

In general, we replace f by $f - \int_X f d\mu$ and obtain

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{k=0}^{n-1}\left|\int_X f\circ\tau^k\bar{g}d\mu - \int_X fd\mu\int_X\bar{g}d\mu\right| = 0,$$

which proves that τ is weakly mixing. $\square$

Definition 3.5.1. Let $\mathbb{J}\subset\mathbb{N}\cup\{0\}$. We define the density of $\mathbb{J}$ as

$$\limsup_{n\to+\infty}\frac{1}{n}\#(\mathbb{J}\cap\{0,1,\ldots,n-1\}).$$

Theorem 3.5.3. *Let $(X, \mathfrak{B}, \tau, \mu)$ be a dynamical system. The following conditions are equivalent:*

(i) τ *is weakly mixing;*
(ii) for any $A, B \in \mathfrak{B}$, *there exists a subset* $\mathbb{J} \subset \mathbb{N} \cup \{0\}$ *of density* 0 *such that*

$$\lim_{\substack{n\to+\infty \\ n\in\mathbb{N}\setminus\mathbb{J}}} \mu(\tau^{-n}(A)\cap B) = \mu(A)\mu(B);$$

(iii) $\tau \times \tau$ *is weakly mixing;*
(iv) $\tau \times T$ *is ergodic, for any ergodic system* $(Y, \mathfrak{A}, T, \nu)$;
(v) $\tau \times \tau$ *is ergodic.*

Proof. (i) $\Leftrightarrow$ (ii) follows from the definition of weak mixing and the Koopman–von Neumann Lemma (Problem 3.5.3).

(ii)$\Rightarrow$(iii) It is enough to show that for any $A, B, C, D \in \mathfrak{B}$,

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} |\mu(\tau^{-k}(A)\cap B)\cdot\mu(\tau^{-k}(C)\cap D) - \mu(A)\mu(B)\mu(C)\mu(D)| = 0. \quad (3.5.1)$$

By (ii), there exist sets of density 0, $\mathbb{J}_1$ and $\mathbb{J}_2$, such that

$$\lim_{\substack{n\to+\infty \\ n\in\mathbb{N}\setminus\mathbb{J}_1}} |\mu(\tau^{-n}(A)\cap B) - \mu(A)\mu(B)| = 0$$

and

$$\lim_{\substack{n\to+\infty \\ n\in\mathbb{N}\setminus\mathbb{J}_2}} |\mu(\tau^{-n}(C)\cap B) - \mu(C)\mu(D)| = 0.$$

The set $\mathbb{J} = \mathbb{J}_1 \cup \mathbb{J}_2$ is of density 0 and

$$\begin{aligned}
&\lim_{\substack{n\to+\infty \\ n\in\mathbb{N}\setminus\mathbb{J}}} |\mu(\tau^{-n}(A)\cap B)\mu(\tau^{-n}(C)\cap (D) - \mu(A)\mu(B)\mu(C)\mu(D)| \\
&\qquad \le \lim_{\substack{n\to+\infty \\ n\in\mathbb{N}\setminus\mathbb{J}_1}} |\mu(\tau^{-n}(A)\cap B) - \mu(A)\mu(B)|\cdot\mu(\tau^{-n}(C)\cap (D) \\
&\qquad\quad + \lim_{\substack{n\to+\infty \\ n\in\mathbb{N}\setminus\mathbb{J}_2}} \mu(A)\mu(B)|\mu(\tau^{-n}(C)\cap (D) - \mu(C)\mu(D)| = 0.
\end{aligned}$$

Now (3.5.1) follows by the Koopman–von Neumann Lemma.

(iii)$\Rightarrow$(iv) If $\tau\times\tau$ is weakly mixing, then so is τ itself. Let $(Y, \mathfrak{A}, T, \nu)$ be ergodic. To prove that $\tau \times T$ is ergodic on $X \times Y$, it is enough to show that, for any $A, B \in \mathfrak{B}$ and any $C, D \in \mathfrak{A}$,

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} \mu(\tau^{-k}(A)\cap B)\nu(T^{-k}(C)\cap D) = \mu(A)\mu(B)\nu(C)\nu(D).$$

We have

$$\begin{aligned}
&\lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} \mu(\tau^{-k}(A)\cap B)\nu(T^{-k}(C)\cap D)\\
&\quad= \lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1}\Big(\mu(A)\mu(B)\nu(T^{-k}(C)\cap D)\\
&\qquad\qquad + (\mu(\tau^{-k}(A)\cap B) - \mu(A)\mu(B))\nu(T^{-k}(C)\cap (D)\Big)\\
&\quad= \mu(A)\mu(B)\nu(C)\nu(D) + 0,
\end{aligned}$$

since by ergodicity of T, $\frac{1}{n}\sum_{k=0}^{n-1}\nu(T^{-k}(C)\cap(D) \to \nu(C)\nu(D)$, and by weak mixing of τ, $\frac{1}{n}\sum_{k=0}^{n-1}|\mu(\tau^{-k}(A)\cap B) - \mu(A)\mu(B)| \to 0$.

(iv)$\Rightarrow$(v) τ is ergodic since $\tau \times \mathrm{I}d$ is ergodic, where Id is the identity on the space consisting of a single point. Thus, $\tau\times\tau$ is ergodic.

(v)$\Rightarrow$(i) For any $A, B \in \mathfrak{B}$, we have

$$\begin{aligned}
&\frac{1}{n}\sum_{k=0}^{n-1}(\mu(\tau^{-k}(A)\cap B) - \mu(A)\mu(B))^2\\
&\quad= \frac{1}{n}\sum_{k=0}^{n-1}\mu(\tau^{-k}(A)\cap B)^2 - 2\mu(A)\mu(B)\frac{1}{n}\sum_{k=0}^{n-1}\mu(\tau^{-k}(A)\cap B)\\
&\qquad + (\mu(A)\mu(B))^2\\
&\quad= \frac{1}{n}\sum_{k=0}^{n-1}(\mu\times\mu)((\tau\times\tau)^{-k}(A\times A)\cap(B\times B))\\
&\quad - 2\mu(A)\mu(B)\frac{1}{n}\sum_{k=0}^{n-1}(\mu\times\mu)((\tau\times\tau)^{-k}(A\times X)\cap(B\times X))\\
&\qquad + (\mu(A)\mu(B))^2.
\end{aligned}$$

Since $\tau\times\tau$ is ergodic, this converges to

$$\begin{aligned}
&(\mu\times\mu)(A\times A)(\mu\times\mu)(B\times B)\\
&\quad - 2\mu(A)\mu(B)(\mu\times\mu)(A\times X)(\mu\times\mu)(B\times X) + (\mu(A)\mu(B))^2 = 0,
\end{aligned}$$

as $n\to+\infty$. By Problem 3.5.2, τ is weakly mixing. $\square$

3.6 Basic Constructions of Ergodic Theory

Definition 3.6.1. Induced Transformations

Let $\tau : X \to X$ be a measurable transformation preserving a normalized measure μ. Let $A \in \mathfrak{B}$ and $\mu(A) > 0$. According to Kac's Lemma, the first return–time function $n = n_A$ is integrable and we can define a transformation

$$\tau_A(x) = \tau^{n(x)}(x), \quad x \in A.$$

The transformation $\tau_A : A \to A$ is called an *induced transformation* or the *first return transformation.*

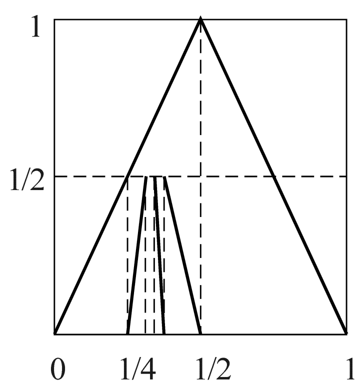

FIGURE 3.6.1

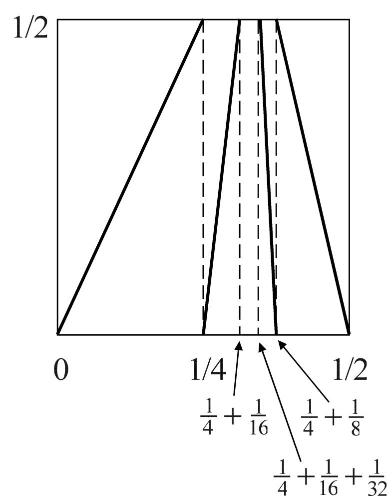

FIGURE 3.6.2

Example 3.6.1. Let τ be a tent transformation:

$$\tau(x) = \begin{cases} 2x & , \quad \text{for } x \in [0, \frac{1}{2}]; \\ 2 - 2x & , \quad \text{for } x \in (\frac{1}{2}, 1]; \end{cases}$$

and let $A = [0, \frac{1}{2}]$. We will construct the first return transformation τ_A. We use Figure 3.6.1 for the construction of τ_A while τ_A is shown in Figure 3.6.2. We have $n = 1$ on $[\frac{0}{4}, \frac{1}{4}]$, i.e., all points in $[0, \frac{1}{4}]$ return to A in one iteration of τ; $n = 2$ on $[\frac{3}{8}, \frac{4}{8}]$ $(\tau^2(\frac{4}{8}) = 0)$, i.e., points in $[\frac{3}{8}, \frac{4}{8}]$ return to A in two iterations of τ; $n = 3$ on $[\frac{4}{16}, \frac{5}{16}]$ and $\tau^3(\frac{4}{16}) = 0$; $n = 4$ on $[\frac{11}{32}, \frac{12}{32}]$ and $\tau^4(\frac{12}{32}) = 0$. In general, if $n = k$ on $[\frac{s}{2^{k+1}}, \frac{s+1}{2^{k+1}}]$ and k is odd, then $n = k + 2$ on $[\frac{4\cdot(s+1)}{4\cdot 2^{k+1}}, \frac{4(s+1)+1}{4\cdot 2^{k+1}}]$ and $\tau^{k+2}(\frac{4(s+1)}{4\cdot 2^{k+1}}) = 0$. If

$n = k$ on $[\frac{s}{2^{k+1}}, \frac{s+1}{2^{k+1}}]$ and k is even, then $n = k+2$ on $[\frac{4s-1}{4\cdot 2^{k+1}}, \frac{4s}{4\cdot 2^{k+1}}]$ and $\tau^{k+2}(\frac{4s}{4\cdot 2^{k+1}}) = 0$. This allows us to construct the consecutive branches of τ_A inductively.

Definition 3.6.2. Integral Transformations

Let $\tau : X \to X$ be a measurable transformation preserving the measure μ, and let $f : X \to \mathbb{N}$ be an integrable function. Let us define

$$X^f = \{(x,i) : \; x \in X, \; 1 \le i \le f(x)\}.$$

Then, we define an integral transformation $\tau^f : X^f \to X^f$ as follows:

$$\tau^f(x,i) = \begin{cases} (x, i+1), & \text{if } \; i+1 \le f(x), \\ (\tau(x), 1), & \text{if } \; i+1 > f(x). \end{cases}$$

Example 3.6.2. Let $\tau(x)$ be the tent transformation on $X = [0,1]$ and let $f(x) = 2 \cdot \chi_{[0,\frac{1}{2})} + 3 \cdot \chi_{[\frac{1}{2},1]}$. The integral transformation τ^f is shown in Figure 3.6.3. We use the notation $x_n = (\tau^f)^n(x), \; n = 1, 2, \dots$.

The invariant measures for a transformation τ and the induced transformation τ_A are closely related. We describe this relationship in the following two propositions.

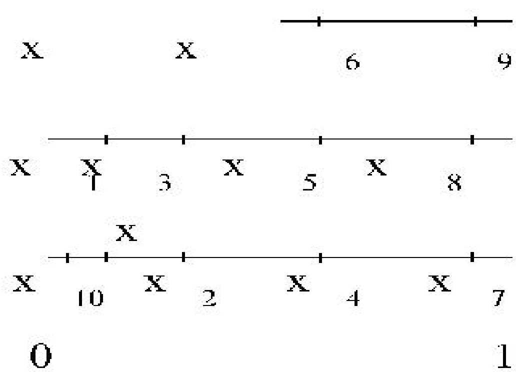

FIGURE 3.6.3

Proposition 3.6.1. *Let τ be a transformation preserving a normalized measure μ and let $\mu(A) > 0$. Then $\mu_{|A}$ is invariant under the induced transformation τ_A.*

Proof. Let $B \subset A$. For the measure μ, we have

$$\mu(B) = \mu(\tau^{-1}B) = \mu(\tau^{-1}B \cap A) + \mu(\tau^{-1}B \setminus A)$$

$$
\begin{aligned}
&= \mu(\tau^{-1}B \cap A) + \mu(\tau^{-2}B \setminus \tau^{-1}A) \\
&= \mu(\tau^{-1}B \cap A) + \mu((\tau^{-2}B \setminus \tau^{-1}A) \cap A) + \mu(\tau^{-2}B \setminus (A \cup \tau^{-1}A)) \\
&= \sum_{i=1}^{\infty} \mu(\tau^{-i}B \setminus \cup_{k=1}^{i-1}\tau^{-k}A) \cap A) + \lim_{i\to+\infty} \mu(\tau^{-i}B \setminus (\cup_{k=0}^{i-1}\tau^{-k}A)).
\end{aligned}
$$

We will show that

$$\lim_{i\to+\infty} \mu(\tau^{-i}B \setminus (\cup_{k=0}^{i-1}\tau^{-k}(A))) = 0. \tag{3.6.1}$$

Let $\bar{A} = \bigcup_{k=0}^{\infty} \tau^{-k}(A)$. We have $\mu(\bar{A}) = \lim_{i\to+\infty} \mu\left(\bigcup_{k=0}^{i-1} \tau^{-k}(A)\right)$. Since $\tau^{-i}B \setminus \bigcup_{k=0}^{i-1} \tau^{-k}(A) \subset \bar{A} \setminus \bigcup_{k=0}^{i-1} \tau^{-k}(A)$, (3.6.1) follows. Thus, we have

$$
\begin{aligned}
\mu(B) &= \sum_{i=1}^{\infty} \mu((\tau^{-i}B \setminus \cup_{k=1}^{i-1}\tau^{-k}(A)) \cap A) \\
&= \sum_{i=1}^{\infty} \mu(\tau^{-i}B \cap A_i) = \mu(\tau_A^{-1}(B)),
\end{aligned}
$$

where $A_i = \{x \in A : n(x) = i\}$. □

Proposition 3.6.2. *Let $\tau : X \to X$ be a measurable transformation and let $A \subset X$. Let the induced transformation $\tau_A : A \to A$ preserve the measure μ_A. Then τ preserves the measure μ, where*

$$\mu(B) = \sum_{k=1}^{\infty} \sum_{i=0}^{k-1} \mu_A(\tau^{-i}(B) \cap A_k),$$

$B \in \mathfrak{B}$ and $A_k = \{x \in A : n(x) = k\}$. In particular,

$$\mu(X) = \sum_{k=1}^{\infty} k \cdot \mu_A(A_k).$$

Proof. Let $B \in \mathfrak{B}$. We have

$$
\begin{aligned}
\mu(\tau^{-1}(B)) - \mu(B) &= \sum_{k=1}^{\infty} \mu_A(\tau^{-k}(B) \cap A_k) - \sum_{k=1}^{\infty} \mu_A(B \cap A_k) \\
&= \mu_A(\tau_A^{-1}(B)) - \mu_A(B) = 0.
\end{aligned}
$$

□

The idea for the definitions of measures μ_A in terms of μ and μ in terms of μ_A comes from the fact that constructions of induced and integral transformations are inverses of each other when τ is a 1-to-1 transformation. Unfortunately, this is not the case in general. The formula given in Proposition 3.6.2 is frequently used to construct an invariant measure for transformations we cannot deal with directly. For example, for transformations on an interval, that have critical points (i.e., points where $\tau'(x_0) = 0$), it can be used to prove the existence of an invariant measure that is absolutely continuous with respect to Lebesgue measure (acim). It turns out that if x_0 is a critical point of τ, then it is often possible to find a neighborhood U of x_0 such that the induced transformation τ_U is piecewise expanding (although with a countable number of branches). Then, we can prove the existence of an acim for τ_U and, by using Proposition 3.6.2, for τ itself. More details on this method can be found in [de Melo and van Strien, 1993].

Example 3.6.3. Let $\tau(x) = 4 \cdot x \cdot (1 - x)$, $x \in [0, 1]$. In Figures 3.6.4 and 3.6.5 we show τ_{U_1} and τ_{U_2}, where $U_1 = [0, \frac{1}{2}]$ and $U_2 = [\frac{1}{4}, \frac{1}{2}]$. τ_{U_1} is not piecewise expanding, and it can be seen in the picture that τ_{U_2} is piecewise expanding. This can be proved rigorously. Both τ_{U_1} and τ_{U_2} have countably many branches, but we can show only a finite number of them. What is left out are almost vertical lines accumulating densely.

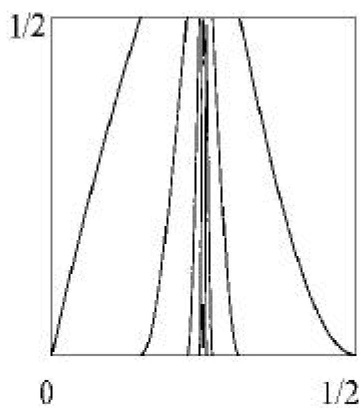

FIGURE 3.6.4

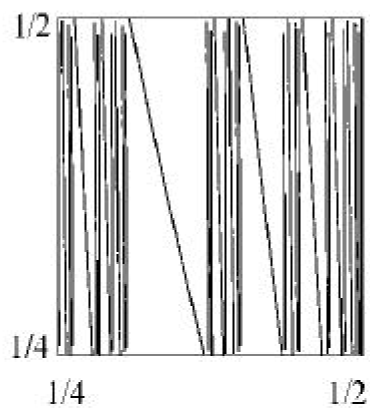

FIGURE 3.6.5

Proposition 3.6.3. *Let the transformation τ preserve measure μ, and let τ_A be an induced transformation on $A \subset X$, $\mu(A) > 0$. We assume that $\mu(X \setminus \bigcup_{n=0}^{\infty} \tau^{-n}(A)) = 0$. Then (τ, μ) is ergodic if and only if (τ_A, μ_A) is ergodic.*

Proof. First, let us assume that (τ, μ) is ergodic. Let $B \subset A$ be a τ_A-invariant set with $\mu_A(B) = \mu(B) > 0$. Then $B = \bigcup_{n=0}^{\infty} \tau_A^{-n}(B) = (\bigcup_{n=0}^{\infty} \tau^{-n} B) \cap A$. Since τ is ergodic, $\bigcup_{n=0}^{\infty} \tau^{-n} B$ is a set of full μ-measure in X. Thus, B is a set of full μ_A-measure in A, which proves the ergodicity of τ_A.

Let (τ_A, μ_A) be ergodic. Let $B \subset X$ be a τ-invariant set with $\mu(B) > 0$. Let $\bar{B} = \bigcup_{n=0}^{\infty} \tau^{-n}(B)$. Then $\tau^{-n}(\bar{B}) \subset \bar{B}$, for $n = 1, 2, \ldots$. If $\mu(\bar{B} \cap A) = 0$, then for $n = 1, 2, \ldots,$ $\mu(\bar{B} \cap \tau^{-n}(A)) = 0$, which is impossible since $\mu(X \setminus \bigcup_{n=0}^{\infty} \tau^{-n}(A)) = 0$. Thus, $C = \bar{B} \cap A$ is of positive measure. Since (τ_A, μ_A) is ergodic, we have $A = \bigcup_{n=0}^{\infty} \tau_A^{-n}(C) \subset \bar{B}$ and hence $\mu(X \setminus \bar{B}) = 0$. Thus, (τ, μ) is ergodic. □

Remark 3.6.1. Assume τ preserves an ergodic measure μ, $\mu(A) > 0$, and that τ_A preserves the measure $\mu_A = \mu_{|A}$. Then the construction of Proposition 3.6.2 applied to μ_A gives back the measure μ.

Proof. Let ν be the measure obtained from μ_A via the construction of Proposition 3.6.2. Then ν is τ-invariant and $\nu = \mu$ on A. By the τ-invariance of ν and μ, $\nu = \mu$ on $\bigcup_{n=0}^{\infty} \tau^{-n} A$, which is a set of full measure. □

Definition 3.6.3. Natural Extension of a Transformation

Let $\tau : X \to X$ be a measurable transformation. We define a natural extension T_τ of τ as follows: Let

$$X_\tau = \{(x_0, x_1, x_2, \ldots) : x_n = \tau(x_{n+1}),\ x_n \in X,\ n = 0, 1, 2, \ldots\},$$

and let $T_\tau : X_\tau \to X_\tau$ be defined by

$$T_\tau((x_0, x_1, \ldots)) = (\tau(x_0), x_0, x_1, \ldots).$$

T_τ is 1-to-1 on X_τ. If τ preserves a measure μ, then we can define a measure $\overline{\mu}$ on X_τ by defining $\overline{\mu}$ on the cylinder sets

$$C(A_0, A_1, \ldots, A_k) = \{(x_0, x_1, \ldots) : x_0 \in A_0,\ x_1 \in A_1, \ldots, x_k \in A_k\}$$

as follows:

$$\overline{\mu}(C(A_0, A_1, \ldots, A_k)) = \mu(\tau^{-k}(A_0) \cap \tau^{-k+1}(A_1) \cap \cdots \cap A_k).$$

Proposition 3.6.4. *If τ preserves the measure μ, then T_τ preserves the measure $\overline{\mu}$. (τ, μ) is ergodic if and only if $(T_\tau, \overline{\mu})$ is ergodic. (τ, μ) is weakly mixing if and only if $(T_\tau, \overline{\mu})$ is weakly mixing.*

Proof. It is enough to check that T_τ preserves $\overline{\mu}$ on cylinder sets. We have

$$\begin{aligned} C(A_0, A_1, \ldots, A_k) \;=\; & C(A_0, \tau^{-1}(A_0) \cap A_1,\; \tau^{-2}(A_0) \cap A_2, \ldots, \\ & \tau^{-k}(A_0) \cap A_k,\;\; \tau^{-k-1}(A_0) \cap \tau^{-1}(A_k)). \end{aligned}$$

Therefore,

$$\begin{aligned} T_\tau^{-1}(C(A_0, A_1,\; \ldots,\; A_k)) \;=\; & C(\tau^{-1}(A_0) \cap A_1,\;\; \tau^{-2}(A_0) \cap A_2,\; \ldots, \\ & \tau^{-k}(A_0) \cap A_k, \tau^{-k-1}(A_0) \cap \tau^{-1}(A_k)). \end{aligned}$$

Since

$$\begin{aligned} & \mu(\tau^{-k}(A_0) \cap \tau^{-k+1}(A_1) \cap \;\cdots \cap A_k) \;= \\ & \quad \mu(\tau^{-k}(\tau^{-1}(A_0) \cap A_1) \cap \tau^{-k+1}(\tau^{-2}(A_0) \cap A_2), \ldots, \\ & \qquad \tau^{-1}(\tau^{-k}(A_0) \cap A_k) \cap \tau^{-k-1}(A_0) \cap \tau{-}1(A_k)), \end{aligned}$$

we have

$$\bar{\mu}(C(A_0, A_1, \ldots, A_k)) = \bar{\mu}(T_\tau^{-1}(C(A_0, A_1, \ldots, A_k)),$$

for any cylinder set.

Ergodicity If C is a τ-invariant subset of X, then $A = \{(x_0, x_1, \ldots) : x_i \in C,\; i = 0, 1, \ldots\}$ is a T_τ-invariant subset of X_τ, and $\mu(C) = \bar{\mu}(A)$. Thus, if τ is not ergodic, then T_τ is also not ergodic. This proves that the ergodicity of T_τ implies the ergodicity of τ.

Now, let us assume that (τ, μ) is ergodic. We will use von Neumann's Ergodic Theorem. For any $f \in \mathfrak{L}^1(X, \mu)$, we have

$$\frac{1}{n} \sum_{k=0}^{n-1} f(\tau^k(x)) \xrightarrow{\mathfrak{L}^1} \int_X f d\mu. \tag{3.6.2}$$

Let $F \in \mathfrak{L}^1(X_\tau, \bar{\mu})$ be of the form $F(\bar{x}) = f(x_{i_0})$, where $\bar{x} = (x_0, x_1, \ldots) \in X_\tau$ and $x_{i_0} \in X,\; i_0 \geq 0$. By (3.6.2), we have

$$\frac{1}{n} \sum_{k=0}^{n-1} F(T_\tau^k(\bar{x})) = \frac{1}{n} \sum_{k=0}^{n-1} f(\tau^k(x_{i_0})) \xrightarrow{\mathfrak{L}^1} \int_{X_\tau} F(\bar{x}) d\bar{\mu}. \tag{3.6.3}$$

For any $\bar{x} = (x_0, x_1, \ldots) \in X_\tau$, we have $x_i = \tau^{i_0 - i}(x_{i_0}),\; i = 0, 1, \ldots, i_0$. For any integrable function $G(\bar{x})$ on X_τ depending on a finite number of coordinates, $G(\bar{x}) = G(x_0, x_1, \ldots, x_{i_0})$, we can write $G(\bar{x}) =$

$G(\tau^{i_0}(x_{i_0}), \tau^{i_0-1}(x_{i_0}), \dots, x_{i_0})$. Thus, (3.6.3) holds for G. Since functions G of this form are dense in $\mathfrak{L}^1(X_\tau, \bar{\mu})$, (3.6.3) holds for any $F \in \mathfrak{L}^1(X_\tau, \bar{\mu})$ and therefore $(T_\tau, \bar{\mu})$ is ergodic.

Weak mixing Again, we will use the close relationship between $\mathfrak{L}^p(X, \mu)$ and $\mathfrak{L}^p(X_\tau, \overline{\mu})$, $p = 1, \infty$. If $(T_\tau, \overline{\mu})$ is weakly mixing, then for any $F \in \mathfrak{L}^1(X_\tau, \overline{\mu})$ and any $G \in \mathfrak{L}^\infty(X_\tau, \overline{\mu})$, we have

$$\frac{1}{n}\sum_{k=0}^{n-1} |F(\overline{x}) \cdot G(T_\tau^k(\overline{x})) - \int F d\overline{\mu} \cdot \int G d\overline{\mu}| \to 0$$

as $n \to +\infty$. In particular, it is true for $F(\overline{x}) = f(x_0)$ and $G(\overline{x}) = g(x_0)$. Thus, (τ, μ) is weakly mixing. If (τ, μ) is weakly mixing, we prove the weak mixing of $(T_\tau, \overline{\mu})$ in the same way we proved its ergodicity above.

□

Example 3.6.4. Let S be a compact metric space with measure ν on a Borel σ-algebra of subsets of S. Let $X = \prod_{k=0}^\infty S$ with the product σ-algebra and the product measure $\mu = \prod_{k=0}^\infty \nu$. Let $\tau : X \to X$ be the left shift on X, i.e.,

$$\tau((x_0, x_1, x_2, \dots)) = (x_1, x_2, \dots).$$

Then τ is noninvertible (in general). We will construct a natural extension of τ. By Definition 3.6.3, we define

$$X_\tau = \{\overline{x} = (y_0, y_1, \dots) : \ y_i = (x_0^{(i)}, x_1^{(i)}, \dots), \\ \tau(y_{i+1}) = y_i, \ y_i \in X, \ i = 0, 1, \dots\}.$$

By virtue of the condition $\tau(y_{i+1}) = y_i, \ i = 0, 1, \dots$, the sequences y_i are of the form

$$\begin{aligned} y_0 &= (x_0, x_1 \dots), \\ y_1 &= (x_{-1}, x_0, x_1, \dots), \\ y_2 &= (x_{-2}, x_{-1}, x_0, x_1, \dots), \\ &\vdots \\ y_n &= (x_{-n}, x_{-n+1}, \dots, x_{-1}, x_0, x_1, \dots). \end{aligned}$$

It is natural then to write the double sequence $\bar{x} = (y_0, y_1, \dots)$ as one two-sided sequence:

$$\bar{x} = (\dots, x_{-n}, x_{-n+1}, \dots, x_{-1}, \overset{\downarrow}{x}_0, x_1, \dots).$$

We interpret the transformation T_τ, defined by $T_\tau((y_0, y_1, y_2, \dots)) = (\tau(y_0), y_0, y_1, \dots)$, as the left shift on the space of two-sided sequences.

Definition 3.6.4. Skew Product

Let $(- , \mathfrak{A}, \sigma, \nu)$ be a dynamical system and let $(S, \mathfrak{B}, \tau_\omega, \mu_\omega)_{\omega \in}$ be a family of dynamical systems such that the function $\tau_\omega(x)$ is $\mathfrak{A} \times \mathfrak{B}$ measurable. A skew product of σ and $\{\tau_\omega\}_{\omega \in}$ is a transformation $T : - \times X \to - \times X$ defined by

$$T(\omega, x) = (\sigma(\omega), \tau_\omega(x)),$$

$\omega \in -$, $x \in X$.

Proposition 3.6.5. *If ν is σ-invariant and μ_ω is τ_ω-invariant for $\omega \in -$, then the measure on $\mathfrak{A} \times \mathfrak{B}$*

$$\mu(A \times B) = \int_A \mu_\omega(B) d\nu(\omega) \tag{3.6.4}$$

is T-invariant. If μ is a T-invariant measure and $\mathfrak{B}$ is countably generated, then there exists a σ-invariant measure ν on $\mathfrak{A}$ and a family of measures $\{\mu_\omega\}_{\omega \in}$ on $\mathfrak{B}$ such that μ_ω is τ_ω-invariant, and the representation (3.6.4) holds.

Proof. We have $T^{-1}(\omega, x) = \{(\sigma^{-1}(\omega), \tau_\omega^{-1}(x))\}$ and

$$\begin{aligned}
\mu(T^{-1}(A \times B)) &= \int_{\sigma^{-1}(A)} \mu_{\sigma(\omega)}(\tau_{\sigma(\omega)}^{-1}(B)) d\nu(\omega) \\
&= \int_{\sigma^{-1}(A)} \mu_{\sigma(\omega)}(B) d\nu(\omega) \\
&= \int_A \mu_\omega(B) d\nu(\omega) = \mu(A \times B).
\end{aligned}$$

If μ is T-invariant, then we define $\nu(A) = \mu(A \times X)$ for $A \in \mathfrak{A}$. Obviously ν is σ-invariant. For almost every $\omega \in -$, there exists a measure μ_ω such that (E.1) holds. For any measurable $A \in \mathfrak{A}$, we have

$$\int_A \mu_\omega(\tau_\omega^{-1}(B)) d\nu(\omega) = \int_A \mu_\omega(B) d\nu(\omega).$$

Thus, $\mu_\omega(\tau_\omega^{-1}(B)) = \mu_\omega(B)$, ν-a.e., for any $B \in \mathfrak{B}$. If $\mathfrak{B}$ is countably generated, then we can find a set $A_1 \in \mathfrak{A}$, $\nu(A_1) = 1$ such that μ_ω is τ_ω-invariant for $\omega \in A_1$. □

An important application of a skew product construction is the so–called *random transformation*. Let $- = \Sigma^+ = Y^{\{\mathbb{N} \cup 0\}}$, where Y is a compact space with Borel probability measure η. Let $\nu = \eta^{\{\mathbb{N} \cup (0)\}}$ be the product measure and $\sigma : - \to -$ be the shift to the left. Let $\{X, \mathfrak{B}, \lambda\}$ be a measure space and $\{\tau_y\}_{y \in Y}$ a family of transformations $\tau_y : X \to X$, such that $\tau_y(x)$ is a measurable function. A skew product T of σ and $\{\tau_y\}_{y \in Y}$ can be interpreted as a "random transformation" $\{\tau_y, \eta\}$, where the transformation τ_y is chosen according to the probability η. If Y is a finite space this model is especially simple: We have a finite number of transformations $\{\tau_i\}_{i=1}^k$ that act with probabilities $\{\eta_i\}_{i=1}^k$. If all the τ_i are nonsingular transformations (with respect to λ), we can write the Frobenius–Perron operator (see Chapter 4) of $T_\eta = \{\tau_i, \eta_i\}_{i=1}^k$. It is easy to check that

$$P_{T_\eta} = \sum_{i=1}^{k} \eta_i P_{\tau_i}.$$

If the transformations $\tau_i \in \mathcal{T}(I)$, a theory analogous to that of Lasota–Yorke (see Chapter 5 of this text) can be developed. See [Pelikan, 1984].

Isomorphism of dynamical systems.

It often happens that two dynamical systems that appear to be completely different behave essentially the same way. To formalize the notion of "essentially the same", we introduce the notion of isomorphism or conjugacy of dynamical systems. In this book we will use two notions of isomorphism.

Definition 3.6.5. Measure Theoretic Isomorphism
Let $(X, \mathfrak{B}_X, \mu, \tau)$ and $(Y, \mathfrak{B}_Y, \nu, T)$ be dynamical systems. We say that they are measure theoretically isomorphic (or conjugated) if there exist $\tilde{X} \subset X$, $\mu(X \setminus \tilde{X}) = 0$, $\tilde{Y} \subset Y$, $\nu(Y \setminus \tilde{Y}) = 0$ and a 1-to-1 measurable transformation $h : \tilde{X} \to \tilde{Y}$ such that on $\tilde{X}$

$$\tau = h^{-1} \circ T \circ h,$$

and $\mu = h_* \nu$.

Definition 3.6.6. Topological Conjugation
Let $(X, \mathfrak{B}_X, \mu, \tau)$ and $(Y, \mathfrak{B}_Y, \nu, T)$ be continuous dynamical systems on compact metric spaces X and Y, respectively. We say that they are topologically conjugated if there exists a homeomorphism $h : X \to Y$ such that

$$\tau = h^{-1} \circ T \circ h,$$

and $\mu = h_* \nu$.

A measure theoretic isomorphism preserves measure theoretic properties of τ, while topological conjugation preserves both measure theoretic and topological properties of τ (e.g., periodic points).

Example 3.6.6. Let $\tau : [0,1] \to [0,1]$ be defined by $\tau(x) = 2x \pmod 1$ and let $T : \Sigma_+ \to \Sigma_+$ be the left shift on the space of $\{0,1\}$-sequences. τ preserves Lebesgue measure λ, while T preserves the product measure $\mu = \nu^{\{\mathbb{N} \cup 0\}}$, where $\nu(0) = \nu(1) = \frac{1}{2}$. For any $x \in [0,1]$, we define $h(x) = (\varepsilon_0, \varepsilon_1, \varepsilon_2, \dots)$, where $x = \frac{\varepsilon_0}{2} + \frac{\varepsilon_1}{2^2} + \cdots + \frac{\varepsilon_n}{2^{n+1}} + \dots$. To guarantee uniqueness in the definition of $h(x)$, we assume that the expansion never ends with an infinite sequence of 1's. The image $h([0,1]) = \tilde{Y} \subset \Sigma_+$ and $\mu(\tilde{Y}) = 1$. For any $x \in [0,1]$, we have $T \circ h(x) = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots)$, $h^{-1} \circ T \circ h(x) = \frac{\varepsilon_1}{2} + \frac{\varepsilon_2}{2^2} + \cdots + \frac{\varepsilon_n}{2^n} + \dots$, which is obviously equal to $\tau(x) = 2x \pmod 1$. Thus, τ and T are measure theoretic isomorphic. τ and T cannot be topologically conjugate to each other because of the different topological dimensions of $[0,1]$ and Σ_+.

3.7 Infinite and Finite Invariant Measures

Theorem 3.7.1. *(Kakutani–Hajian* [Hajian and Kakutani, 1964]*) Let ν be an infinite ergodic τ-invariant measure. Then any set of finite positive ν-measure contains a weakly wandering set of positive measure. (Weak wandering $\iff$ in the sequence $\{\tau^{-k}A\}_{k=0}^{\infty}$ there are infinitely many disjoint sets.)*

Proof. Take E with $0 < \nu(E) < +\infty$. Then $f = \chi_E$ is integrable. The limit

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ \tau^k$$

exists in $\mathfrak{L}^1$ (von Neumann's Ergodic Theorem) and is τ-invariant. By ergodicity it must be a constant and since $\nu(x) = +\infty$, it is equal to 0. Thus, for any $F \in \mathfrak{B}$, $\nu(F) < +\infty$,

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} \nu(\tau^{-k}(E) \cap F) = 0.$$

By the Koopman–von Neumann Lemma (Problem 3.5.3), there exists a subset $\mathbb{N}_0 \subset \mathbb{N}$ of density 0 such that

$$\lim_{\substack{k\to+\infty \\ k\in\mathbb{N}\setminus\mathbb{N}_0}} \nu(\tau^{-k}(E) \cap F) = 0.$$

Let $\{a_k\}_{k=1}^{\infty}$ be a strictly positive sequence with $\sum_{k=1}^{\infty} a_k < \nu(E)$. We can find integers $n_k :\ 0 < n_1 < n_2, \ldots$ such that

$$\sum_{j<k} \nu(\tau^{-(n_k-n_j)}(E) \cap E) = \sum_{j<k} \nu(\tau^{-n_k}(E) \cap \tau^{-n_j}(E)) < a_k, \quad k = 1, 2, \ldots.$$

($F = \tau^{-n_j}(E)$ to obtain n_k). Let

$$E_0 = E \setminus \bigcup_k \bigcup_{j<k} \tau^{-(n_k-n_j)}(E).$$

By the choice of $\{a_k\}_{k=1}^{+\infty}$, $\nu(E_0) > 0$. If $j < k$ we have

$$E_0 \subseteq E \setminus \tau^{-(n_k-n_j)}(E)$$

and

$$\tau^{-n_j}(E_0) \subset \tau^{-n_j}(E) \setminus \tau^{-n_k}(E) \subseteq \tau^{-n_j}(E) \setminus \tau^{-n_k}(E_0).$$

Hence $\tau^{-n_j}(E_0) \cap \tau^{-n_k}(E_0) = \emptyset$ and E_0 is weakly wandering.

□

Corollary 3.7.1. *Let τ and ν be as in Theorem 3.7.1. An ergodic infinite measure ν admits no equivalent finite invariant measure since weak wandering implies no finite invariant measure can exist.*

Problems for Chapter 3

Problem 3.1.1. Prove that

$$\begin{aligned} &\text{(i)} \quad \tau^{-1}A \cap \tau^{-1}B = \tau^{-1}(A \cap B), \\ &\text{(ii)} \quad \tau^{-1}A \cup \tau^{-1}B = \tau^{-1}(A \cup B). \end{aligned}$$

Problem 3.1.2. Prove

$$\tau(\tau^{-1}(A)) \subseteq A \tag{1}$$

and

$$\tau^{-1}(\tau(A)) \supseteq A. \tag{2}$$

Show examples with strict inclusions. Prove if τ is injective, then there is equality in (2). If τ is surjective, then there is equality in (1).

Problem 3.1.3. Give an example of a measurable transformation τ such that $\tau^{-1}\tau \neq \tau^0 = \mathrm{I}d$.

Problem 3.1.4. Let τ be a measure preserving transformation on $(X, \mathfrak{B}, \mu)$. Let $A \in \mathfrak{B}$. Prove that $\mu(A \setminus \tau^{-1}A) = \mu(\tau^{-1}A \setminus A)$.

Problem 3.1.5. Let $(X, \mathfrak{B}, \mu)$ be a measure space. Let $\tau : X \to X$ be a measurable transformation and $A \in \mathfrak{B}$. Prove

(a) if either of the sets $A \setminus \tau^{-1}(A)$ or $\tau^{-1}(A) \setminus A$ has μ-measure 0, then A is almost τ-invariant, i.e., $\mu(A \Delta \tau^{-1}(A)) = 0$;

(b) A is an almost τ-invariant set if and only if $\chi_A = \chi_{\tau^{-1}(A)}$ a.e., i.e., χ_A is an almost τ-invariant function;

(c) if τ is nonsingular and if either $\mu(A) = 0$ or $\mu(A^c) = 0$, then A is almost τ-invariant.

Problem 3.1.6. Let $\tau : \mathbb{R} \to \mathbb{R}$ be defined by $\tau(x) = 2x$. Show that τ is not measure preserving with respect to Lebesgue measure.

Problem 3.1.7. Let $X = [0, 1)$ and let us define the Gauss transformation

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

where $\{y\}$ denotes the fractional part of y. Let a measure μ be defined by

$$\mu(A) = \frac{1}{\ln 2} \int_A \frac{1}{1+x}\, dx.$$

Prove that τ is measure preserving with respect to μ.

Problem 3.1.8. Let $\tau : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $\tau(x, y) = (2x, \frac{1}{2}y)$. Prove that τ is measure preserving with respect to the Lebesgue measure on $\mathbb{R}^2$.

Problem 3.1.9. Let $(X, \mathfrak{B}, \lambda)$ be a probability space, where $X = [0, 1]^2 \subset \mathbb{R}^2$, $\mathfrak{B}$ is the Borel σ-algebra, and λ is Lebesgue measure on X. Let $\tau : X \to X$ be the baker transformation defined by

$$\tau(x, y) = \begin{cases} (2x, \frac{1}{2}y), & x \in [0, \frac{1}{2}),\ y \in [0, 1], \\ (2x - 1, \frac{1}{2}y + \frac{1}{2}), & x \in [\frac{1}{2}, 1],\ y \in [0, 1]. \end{cases}$$

Show that τ preserves λ.

Problem 3.1.10. Let $(X, \mathfrak{B}, \mu)$ be a measure space and let $\tau : X \to X$ be a measurable transformation. Show that the set function η defined by $\eta(A) = \mu(\tau^{-1}(A))$ defines a measure.

Problem 3.1.11. Let $(X, \mathfrak{B}, \mu)$ be a measure space and let $\tau : X \to X$ be a measurable transformation. Let $A \subseteq X$ be a subset of X. Define $B \equiv \lim_{n\to\infty} \tau^{-n} A \equiv \cap_{k=1}^{\infty} (\cup_{n=k}^{\infty} \tau^{-n} A)$. Prove that
(a) $B = \{x \in X \mid x \in \tau^{-n} A \text{ for infinitely many } n\}$;
(b) B is τ-invariant;
(c) if τ is measure preserving and A is measurable and invariant, then $\mu(A) = \mu(B)$.

Problem 3.1.12. Let $(X, \mathfrak{B}, \mu)$ be a measure space and let $\tau : X \to X$ be a measurable transformation. Prove that if μ is τ^n-invariant, then the measure η defined by $\eta(A) = \frac{1}{n} \sum_{i=0}^{n-1} \mu(\tau^{-i} A)$ is τ-invariant.

Problem 3.1.13. (Two-sided $(p_0, \ldots, p_{k-1})$-shift or Bernoulli scheme.) Let $Y = \{0, ..., k-1\}$ and μ be a measure on Y such that $\mu(\{i\}) = p_i$, $i = 0, ..., k-1$ and $\sum_{i=0}^{k-1} p_i = 1$. Consider $X = \Pi_{-\infty}^{\infty} Y$ with the product measure μ. Define $\tau : X \to X$ by

$$\tau(\{x_i\}) = \{y_i\},$$

where $y_i = x_{i+1}$. Show that μ is τ-invariant.

Problem 3.1.14. (One-sided $(p_0, \ldots, p_{k-1})$-shift.) Let Y be as in 3.1.13, $X = \Pi_0^\infty Y$ with the product measure μ. Let $\tau : X \to X$ be defined by $(x_0, x_1, ...) \to (x_1, x_2, ...)$. Show that μ is τ-invariant.

Problem 3.1.15. Let $X = \{\frac{1}{n} : n = 1, 2, \ldots\} \cup 0$ and let $\tau(\frac{1}{n}) = \frac{1}{n+1}$ for $n = 1, 2, \ldots$ and $\tau(0) = 1$. Prove that τ does not preserve any finite measure, i.e., that Theorem (3.1.3) does not hold without continuity of τ.

Problem 3.2.1. Let $X = [0, 1]$. Let $\mathfrak{B}$ be the Borel σ-algebra on $[0, 1]$ and let λ be Lebesgue measure on $[0, 1]$. Define $\tau_\alpha : X \to X$ by $\tau_\alpha(x) = x + \alpha \pmod 1$, where $\alpha > 0$. Prove that τ is ergodic if and only if α is irrational.

Problem 3.2.2. Let (Y, μ, τ) be the Bernoulli scheme of Problem 3.1.13. Prove that τ is ergodic.

Problem 3.2.3. Let $(X, \mathfrak{B}, \tau, \mu)$ be a measure preserving dynamical system. Let $B_0 \in \mathfrak{B}$ and define $B_k = \tau^{-k} B_0$, $k = 1, 2, \ldots$. Let $\tilde{B}_k = X \backslash B_k$. Prove that

$$A = \cap_{k=0}^\infty \tilde{B}_k$$

is an invariant set.

Problem 3.2.4. Let $\tau : X \to X$ be a measure preserving transformation. Suppose each τ almost invariant set has measure 0 or $\mu(X)$. Show that τ is ergodic with respect to μ.

Problem 3.2.5. Let $\tau : [0, 1] \times [0, 1] \to [0, 1] \times [0, 1]$ be defined by $\tau(x, y) = (f(x), f(y))$, where

$$f(t) = \begin{cases} 2t, & \text{if } 0 \le t \le \frac{1}{2}, \\ 2 - 2t, & \text{if } \frac{1}{2} < t \le 1. \end{cases}$$

Show that τ is ergodic.

Problem 3.2.6. Let $(X, \mathfrak{B}, \mu, \tau)$ be a dynamical system. Assume μ is the unique τ-invariant measure. Prove that τ is ergodic.

Problem 3.2.7. Let $\tau(x) = (x - 1)^2$, $x \in [0, 1]$. Does τ have a continuous invariant measure, i.e., one for which every point has measure 0?

Problem 3.2.8. Prove Kac's Lemma using Birkhoff's Ergodic Theorem.

Problem 3.2.9. Is it possible to have $R_\tau \neq - {}_\tau$?

Problem 3.2.10. In the proof of Theorem 3.2.4 we claimed that $\tau^{-1}(C_k) = B_{k+1}$, for $k \geq 1$. Prove it.

Problem 3.3.1. Let $X \subseteq \mathbb{R}^n$ be an open set, $\lambda(X) < \infty$. Let $\tau : X \to X$ be an ergodic transformation. Show that for almost every $x \in X$ the set $\{\tau^k(x)\}_{k=1}^{\infty}$ is dense in X.

Problem 3.3.2. Let $(X, \mathfrak{B}, \tau, \mu_1)$ and $(X, \mathfrak{B}, \tau, \mu_2)$ with $\mu_1 \neq \mu_2$ be two dynamical systems such that $\mu_1(X) = \mu_2(X) = 1$, where τ is ergodic with respect to both μ_1 and μ_2. Prove that there exist sets $A_1, A_2 \in \mathfrak{B}$ such that $A_1 \cap A_2 = \emptyset$ and $\mu_1(A_1) = \mu_2(A_2) = 1$.

Problem 3.3.3. Discuss the Birkhoff Ergodic Theorem as it pertains to a finite space $X = \{a_1, a_2, \cdots, a_m\}$ with counting measure μ.

Problem 3.3.4. Let τ be a measure preserving transformation on $(X, \mathfrak{B}, \mu)$, where $\mu(X) = 1$. Given $E \in \mathfrak{B}$ and $x \in X$, define

$$\chi_E^*(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \chi_E\big(\tau^k(x)\big).$$

Prove that τ is ergodic if and only if $\chi_E^*(x) = \mu(E)$ for almost every $x \in X$.

Problem 3.3.5. Suppose $\tau : (X, \mathfrak{B}, \mu) \to (X, \mathfrak{B}, \mu)$ be an invertible (i.e., both τ and τ^{-1} are measurable) and measure preserving transformation. Prove that τ is ergodic if and only if for each $A, B \in \mathfrak{B}$,

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

Problem 3.3.6. Suppose $\tau : (X, \mathfrak{B}, \mu) \to (X, \mathfrak{B}, \mu)$ is measure preserving. Then prove that τ is ergodic if and only if for all $f, g \in \mathfrak{L}^2(X, \mathfrak{B}, \mu)$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \quad U^k f, g\rangle = \quad f, 1\rangle \cdot \quad g, 1\rangle \tag{1}$$

where $U^k f = f \circ \tau^k$.

Problem 3.3.7. Suppose $X = \{a, b, c, d, e\}$, $\mathfrak{B}$ is the set of all subsets of X, $\mu(a) = \mu(b) = \mu(c) = 1$, $\mu(d) = \mu(e) = 2$ and τ is the permutation which takes a to b, b to c, c to d, d to e and e to d. Show that τ is measure preserving but not ergodic. Let $f(x) = \chi_{a,b,e}(x)$. Find the f^* in the Birkhoff Ergodic Theorem.

Problem 3.3.8. Assume that μ is a normalized τ-invariant measure. Let $\mathfrak{A} \subset \mathfrak{B}$ be a σ-algebra of τ-invariant (μ-a.e.) subsets of X. Prove that the operator $L : \mathfrak{L}^1(X, \mathfrak{B}, \mu) \to \mathfrak{L}^1(X, \mathfrak{B}, \mu)$ defined by $L(f) = f^*$ is actually an operator of conditional expectation:

$$L(f) = E(f|\mathfrak{A}).$$

Problem 3.3.9. Let $\tau(x) = r \cdot x \pmod 1$, $x \in [0, 1]$, where $r \geq 2$ is a positive integer. Generalize Example 3.3.2 for this transformation.

Problem 3.4.1. Prove that every Bernoulli scheme is strongly mixing.

Problem 3.4.2. If the system $(X, \mathfrak{B}, \tau, \mu)$ is weakly mixing, prove that it is ergodic.

Problem 3.4.3. Let τ be a measure preserving transformation on $(X, \mathfrak{B}, \mu)$ where $\mu(X) = 1$. Show that τ is weakly mixing if and only if for every $f, g \in \mathfrak{L}^2(X, \mathfrak{B}, \mu)$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \Big| \ U^k f, g\rangle - \ f, g\rangle \ g, 1\rangle \Big| = 0. \tag{1}$$

Problem 3.4.4. Let $(X, \mathfrak{B}, \mu, \tau)$ be a dynamical system. Let $\mathfrak{B}_0$ be an algebra which generates $\mathfrak{B}$. If $\lim_{n\to\infty} \mu(\tau^{-n} A \cap B) = \mu(A)\mu(B)$, for all $A, B \in \mathfrak{B}_0$, prove that the same is true for all $A, B \in \mathfrak{B}$.

Problem 3.4.5. Show that the dynamical system $(X, \mathfrak{B}, \mu, \tau)$, where $X = [0, 1]$ and $\tau(x) = 2x \pmod 1$ is strongly mixing.

Problem 3.5.1. Let $\eta \neq 1$, $|\eta| = 1$. Prove that

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\eta^k - a| \neq 0,$$

for any $0 < a \leq 1$.

Problem 3.5.2. Let $\{a_n\}_{n=0}^{\infty}$ be a bounded sequence of positive numbers. Prove that

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} a_k = 0 \quad \Leftrightarrow \quad \lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} a_k^2 = 0.$$

Problem 3.5.3. Let $\{a_n\}_{n=0}^{\infty}$ be a bounded sequence of positive numbers. Show that $\lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} a_k \to 0$ if and only if there exists a subset $\mathbb{N}_0 \subset \mathbb{N}$ of density 0 such that $\lim_{\substack{n\to+\infty \\ n\in\mathbb{N}\setminus\mathbb{N}_0}} a_n = 0$.

Problem 3.6.1. Let $\tau : X \to X$ preserve the measure μ. Let f be an integrable function on X. Show that integral transformation τ^f preserves the measure μ^f defined as follows:

$$\mu^f(A, i) = \frac{1}{\int_x f d\mu}\mu(A),$$

where $A \in \mathfrak{B},\ i \leq f$.

Problem 3.6.2. Under the assumptions of Problem 3.6.1, show that (τ, μ) is ergodic if and only if (τ^f, μ^f) is ergodic.