JMRF, Vol. I, No. 1, 24-28, 2023
web: https://publication.morancollege.com/journals/JMRF/2023/article4.pdf

A Study on Ergodicity and its Consequences

Himanshu Lekharu [1]*

Abstract
Ergodic theory, a profound mathematical field, delves into the long-term behavior of dynamic systems. This concise overview begins by introducing fundamental measure theory concepts, such as measure spaces and measurable functions. It then explores measure-preserving transformations with examples like Poincaré recurrence and introduces the crucial concept of ergodicity, exemplified by irrational rotations. The article culminates in an examination of ergodic theorems, prominently featuring Birkhoff's ergodic theorem, which elucidates the long-term statistical properties of evolving systems. This exploration provides valuable insights into the captivating world of ergodic theory, with broad applications across diverse domains.

Keywords
Poincaré Recurrence, Ergodicity, Birkhoff's Ergodic Theorem

[1] *Research scholar, Jawaharlal Nehru University, New Delhi*
***Corresponding author**: himanshu.lekharu@gmail.com

Contents

1. Introduction and Preliminaries

1.1 Introduction

The word *ergodic* comes from two Greek words *ergon* (meaning work) and *odos* (meaning path). Defining ergodic theory in a proper manner is not so easy as it uses techniques and examples from many fields such as statistical mechanics, probability theory, measure theory, number theory, functional analysis, group actions and many more. In a broad sense, ergodic theory is a branch of mathematics where we study the behavior of measure-preserving transformations. The word ergodic is introduced by Boltzmann in statistical mechanics.

A modern description of what ergodic theory deals with is as follows: It is the study of long-term average behavior of systems evolving in time. The collection of all states of the system form a space X and the evolution is represented by a transformation $T : X \to X$, where Tx represents the state of the system at time $t = 1$, when the system initially stated at x at time $t = 0$.

The transformation T depends on the space X and we want T to preserve the basic structure of X. For example:

- If X is a measure space, then T must be measurable.
- If X is a topological space, then T must be continuous.

1.2 Some Basic Definitions from Measure Theory

A measure space is a tuple (X, β, μ) where:

- X is any nonempty set,
- β is a σ-algebra, and
- μ is a measure on the space (X, β).

Definition 1.1 A σ-algebra on a set X is a collection of subsets β such that:

- $X, \phi \in \beta$,
- β is closed under complement, and
- β is closed under countable union.

Definition 1.2 Given a nonempty set X with a σ-algebra β on X, we define $\mu : \beta \to \mathbb{R}$ to be a measure if:

- $\forall S \in \beta,\ \mu(S) \geq 0$,
- $\mu(\phi) = 0$, and
- For all countable collections $\{S_i\}_{i=1}^{\infty}$ of pairwise disjoint sets in β:

$$\mu\left(\bigcup_{i=1}^{\infty} S_i\right) = \sum_{i=1}^{\infty} \mu(S_i)$$

Definition 1.3 Let (X,β) and (Y,β') be two measure spaces. A function $f : X \to Y$ is called measurable if for all $A \in \beta'$, $f^{-1}(A) \in \beta$.

Lebesgue Outer Measure: If $S \subseteq \mathbb{R}$, the Lebesgue outer measure of S is

$$\mu^*(S) = \inf\left\{ \sum_{k=1}^{\infty} |I_k| : S \subseteq \bigcup_{i=1}^{\infty} I_k, \text{ where } (I_k)_{k=1}^{\infty} \text{ is a collection of open intervals} \right\}$$

where $|I_k|$ is the length of the interval I_k.

Definition 1.4 If X is any topological space, then the smallest σ-algebra generated by all the open sets in X is called the Borel σ-algebra. Also, it is the smallest σ-algebra containing all the closed sets.

Definition 1.5 A property P of points of a set $S \subseteq X$ is said to hold almost everywhere (a.e.) if the set of points of S where the property fails has measure zero.

Definition 1.6 A measure μ on a measure space X is said to be complete if for $B \subseteq A \subseteq X$ and $\mu(A) = 0$, then $\mu(B) = 0$, i.e., the measure of subsets of a measure-zero set is zero.

Definition 1.7 A measure space (X,β,μ) is said to be σ-finite if X can be written as a countable union of measurable sets of finite measure, i.e., $X = \bigcup_{i=1}^{\infty} A_i$ with $\mu(A_i) < \infty$ for all i.

Definition 1.8 Let (X,β,μ) be a complete measure space and $f : X \to \mathbb{R}$ be a measurable function, then for each integer $p \geq 1$, we say that $f \in L^p(\mu)$ if

$$\int_X |f|^p \, d\mu < \infty.$$

For any such $f \in L^p(\mu)$, we may define the L^p-norm as

$$\|f\|_p = \left(\int_X |f|^p \, d\mu\right)^{1/p}.$$

Identifying the functions whose values agree almost everywhere allows for defining a metric on the $L^p(\mu)$-norm. We treat $L^p(\mu)$ as the set of equivalence classes of functions which coincide almost everywhere.

Definition 1.9 Let (X,β,μ) be a measure space. If $\mu(X) = 1$, then μ is called a probability measure and (X,β,μ) is called a probability space.

Dominated Convergence Theorem Suppose $(f_n)_{n=1}^{\infty}$ is a sequence of measurable functions and $\lim_{n\to\infty} f_n(x) = f(x)$ for all $x \in X$, and $|f_n(x)| \leq g(x)$ for all $n \in \mathbb{N}, x \in \mathbb{R}$, where g is an integrable function. Then

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu.$$

Monotone Convergence Theorem Suppose $(f_n)_{n=1}^{\infty}$ is a non-decreasing sequence of non-negative measurable functions. Let $f(x) = \lim_{n\to\infty} f_n(x)$. Then,

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu.$$

2. Measure Preserving Transformations

2.1 Definition

Let (X,β,μ) be a probability space, and $T : X \to X$ measurable. The map T is called measure preserving, or we call μ is T-invariant if $\mu(T^{-1}(S)) = \mu(S)$ for all $S \in \beta$.

2.2 Examples of Measure Preserving Transformations

1. **Translations** - Let $X = [0,1)$ with the Lebesgue σ-algebra β and Lebesgue measure. Let $0 < \theta < 1$, define $T : X \to X$ by

$$Tx = x + \theta \mod 1 = x + \theta - \lfloor x + \theta \rfloor.$$

   Then clearly T is measure preserving.

2. **Multiplication by 2 modulo 1** - Let (X,β,λ) be as above and let $T : X \to X$ be defined as follows

$$Tx = \begin{cases} 2x & \text{if } 0 \leq x < \frac{1}{2} \\ 2x - 1 & \text{if } \frac{1}{2} \leq x < 1. \end{cases}$$

   Then for any interval $[a,b)$,

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

   and

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

   So, the transformation T is measure preserving.

3. **Baker's Transformation** - Consider the space $[0,1)^2$ with product Lebesgue σ-algebra $\beta \times \beta$ and product Lebesgue measure $\lambda \times \lambda$. Now define $T : [0,1)^2 \to [0,1)^2$ by

$$T(x,y) = \begin{cases} (2x, y/2) & \text{if } 0 \leq x < \frac{1}{2} \\ (2x-1, (y+1)/2) & \text{if } \frac{1}{2} \leq x < 1. \end{cases}$$

   Then T is measurable and measure preserving.

4. **β-transformation** - Let $X = [0,1)$ with the Lebesgue σ-algebra. Let $\alpha = \frac{1+\sqrt{5}}{2}$, the golden ratio, and $\alpha^2 = \alpha + 1$. Now define a transformation $T : X \to X$ by

$$Tx = \begin{cases} \alpha x & \text{if } 0 \le x < \frac{1}{\alpha} \\ \alpha x - 1 & \text{if } \frac{1}{\alpha} \le x < 1. \end{cases}$$

Here, T is measurable but not measure preserving with respect to Lebesgue measure, as we can see $T^{-1}\left(\frac{1}{\alpha}, 1\right) = \left(\frac{1}{\alpha^2}, \frac{1}{\alpha}\right)$. But, $\lambda\left(T^{-1}\left(\frac{1}{\alpha}, 1\right)\right) \neq \lambda\left(\left(\frac{1}{\alpha^2}, \frac{1}{\alpha}\right)\right)$. But, this transformation T is measure preserving with respect to the measure μ given by

$$\mu(B) = \int_B g(x)\, dx,$$

where

$$g(x) = \begin{cases} \frac{5+3\sqrt{5}}{10} & \text{if } 0 \le x < \frac{1}{\alpha} \\ \frac{5+\sqrt{5}}{10} & \text{if } \frac{1}{\alpha} \le x < 1. \end{cases}$$

5. **Continued Fraction** - Consider (X, β) where β is the Lebesgue σ-algebra. Define $T : [0,1) \to [0,1)$ by $T0 = 0$ and for $x \neq 0$,

$$Tx = 1/x - \left\lfloor \frac{1}{x} \right\rfloor.$$

Then for any interval $[a,b]$, $T^{-1}([a,b]) = \bigcup_{k=1}^{\infty} \left[\frac{1}{b+k}, \frac{1}{a+k}\right]$. So, T is not Lebesgue measure preserving, but it preserves the Gauss probability measure μ given by

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

For any interval $[0,b]$,

$$\begin{aligned} \mu(T^{-1}([0,b])) &= \frac{1}{\log 2} \sum_{n=1}^{\infty} \int_{1/(b+n)}^{1/n} \frac{1}{1+x}\, dx \\ &= \frac{1}{\log 2} \sum_{n=1}^{\infty} (\log(1+1/n) - \log(1+1/(b+n))) \\ &= \frac{1}{\log 2} \sum_{n=1}^{\infty} \log\left(\frac{(n+1)(b+n)}{n(b+n+1)}\right) \\ &= \frac{1}{\log 2} \int_0^b \frac{1}{1+x}\, dx \\ &= \mu([0,b]). \end{aligned}$$

Taking intersections and unions of such intervals, we can show that T preserves the Gauss probability measure.

3. Recurrence

3.1 Definition

Let T be a measure-preserving transformation on a probability space (X, β, μ), and let $B \in \beta$. A point $x \in B$ is said to be B-recurrent if there exists $k \ge 1$ such that $T^k x \in B$.

3.2 Poincaré Recurrence Theorem

Theorem: Let $T : X \to X$ be a measure-preserving transformation on the probability space (X, β, μ). Let $E \in \beta$ with $\mu(E) < \infty$. Then, almost every point $x \in E$ there exists $n \ge 1$ such that $T^n(x) \in E$. Moreover, there are infinitely many values of n such that $T^n(x) \in E$.

Proof: Let E_0 be the set of points $x \in E$ that never return to E, i.e., $T^n(x) \notin E$ for all $n \ge 0$. We will show that the measure of E_0 is zero.

First, we will prove that

$$T^{-n}(E_0) \cap T^{-m}(E_0) = \emptyset \text{ for all } m \neq n \ge 1.$$

Suppose, $m > n \ge 1$ and

$$x \in T^{-n}(E_0) \cap T^{-m}(E_0).$$

Then, $T^n(x) \in E_0$ and $T^m(x) \in E_0$. Now, let $y = T^n(x)$, then

$$y \in E_0 \text{ and } T^{m-n}(y) = T^m(x) \in E_0$$

which means that y returns to E_0, contradicting the definition of E_0. This proves our first claim.

Now, since T is measure-invariant, so $\mu(T^{-n}(E_0)) = \mu(E_0)$ for all n, and hence we have

$$\mu\left(\bigcup_{i=1}^{\infty} T^{-i}(E_0)\right) = \sum_{i=1}^{\infty} \mu(T^{-i}(E_0)) = \sum_{i=1}^{\infty} \mu(E_0).$$

Since X is of finite measure, so the left side of the expression is of finite measure, hence $\mu(E_0) = 0$.

Now, let F be the set of points $x \in E$ that return to E only a finite number of times. By direct consequence of the definition, every point $x \in F$ has some iterate $T^k(x) \in E_0$. That is,

$$F \subseteq \bigcup_{i=1}^{\infty} T^{-i}(E_0).$$

Then,

$$\mu(F) = \mu\left(\bigcup_{i=1}^{\infty} T^{-i}(E_0)\right) = \sum_{i=1}^{\infty} \mu(T^{-i}(E_0)) = \sum_{i=1}^{\infty} \mu(E_0) = 0.$$

Note: In the above theorem, it is necessary to assume that the space is of finite measure. For example, consider $T : \mathbb{R} \to \mathbb{R}$ defined by $T(x) = x + 1$, then T is Lebesgue measure preserving but there is no recurring point under T.

3.3 Topological Flavours of Recurrence

3.3.1 Definition

Let X be a topological space endowed with the Borel σ-algebra. We say that a point $x \in X$ is recurrent for the transformation $T : X \to X$ if there exists a sequence (n_k) of natural numbers such that $T^{n_k}(x) \to x$.

3.3.2 Theorem

Let $T : X \to X$ be a continuous transformation in a compact metric space X. Then, there exists some point $x \in X$ recurrent for T.

Proof: Consider the family $\mathscr{I}$ of all non-empty closed sets $M \subset X$ that are invariant under T, i.e., $T(M) \subset M$. As $X \in \mathscr{I}$, so $\mathscr{I}$ is non-empty.

We say that an element $M \in \mathscr{I}$ is minimal for the inclusion relation if and only if the orbit of the point $x \in M$ is dense in M.

Certainly, since M is closed and invariant, then M contains the closure of the orbits. Hence, M is minimal if it coincides with any of the orbits' closures. Likewise, if M coincides with the closure of the orbit of any of its points, then it coincides with any closed invariant subset, i.e., M is minimal. This proves our claim. In particular, any point x in a minimal set is recurrent. Hence, to prove the theorem, it suffices to show that there exists a minimal set.

Now we claim that an ordered set $\{M_\alpha\} \subset \mathscr{I}$ admits a lower bound. Indeed, consider $M = \bigcap_\alpha M_\alpha$. Notice that M is non-empty since $\{M_\alpha\}$ are compact, and the family is ordered. Clearly, M is closed and invariant under T, and it is also a lower bound for the set $\{M_\alpha\}$. This proves our claim. Now by Zorn's lemma, $\mathscr{I}$ contains a minimal element. Hence the theorem.

4. Ergodicity

4.1 Definition

Let T be a measure-preserving transformation on a probability space (X, β, μ). The map T is said to be ergodic if for every measurable set A satisfying $T^{-1}A = A$, we have $\mu(A) = 0$ or $\mu(A) = 1$.

The definition of ergodicity for (X, β, μ, T) means that it is impossible to split X into two subsets of positive measure, each of which is invariant under T.

4.2 Examples of Ergodic Transformations

1. Irrational Rotation - Consider $([0,1), \beta, \lambda)$, where β is the Lebesgue σ-algebra and λ is the Lebesgue measure. For $\theta \in (0,1)$, consider the transformation $T_\theta : [0,1) \to [0,1)$ defined by $T_\theta x = x + \theta (\text{mod } 1)$. Then T_θ is measure-preserving with respect to λ. We can see that if θ is rational, then T_θ is not ergodic, as if we take $\theta = p/q$, $\gcd(p,q) = 1$, then T_θ^q is the identity map. Now if we pick ε to be sufficiently small such that ε-neighborhoods of $x + k\theta$, $k = 0, 1, 2, \ldots, q-1$, are disjoint. Then the union of these neighborhoods is an invariant set of positive measure. So T_θ is not ergodic when θ is rational.

2. Multiplication by 2 modulo 1 - Let (X, β, μ) as above and define $T : X \to X$ given by

$$Tx = \begin{cases} 2x & \text{if } 0 \leq x < \frac{1}{2} \\ 2x - 1 & \text{if } \frac{1}{2} \leq x < 1 \end{cases}$$

Then, T is measure-preserving and ergodic.

3. Consider the space $([0,1)^2, \beta \times \beta, \lambda \times \lambda)$, where β is the Lebesgue σ-algebra and λ is the normalized Lebesgue measure. Suppose $\theta \in (0,1)$ is irrational and define $T_\theta \times T_\theta : [0,1) \times [0,1) \to [0,1) \times [0,1)$ by

$$T_\theta \times T_\theta(x,y) = (x + \theta (\text{mod } 1), y + \theta (\text{mod } 1))$$

Here, $T_\theta \times T_\theta$ is measure-preserving but not ergodic.

4.3 Some Characterizations of Ergodic Maps

4.3.1 Theorem:

Let (X, β, μ) be a probability space and $T : X \to X$ be measure-preserving. Then the following are equivalent: 1. T is ergodic. 2. If $B \in \beta$ with $\mu(T^{-1}B \triangle B) = 0$, then $\mu(B) = 0$ or $\mu(B) = 1$. 3. If $A \in \beta$ with $\mu(A) > 0$, then $\mu\left(\bigcup_{n=1}^{\infty} T^{-n}A\right) = 1$ (i.e., if A is a set of positive measure, almost every $x \in X$ will visit A infinitely often). 4. If $A, B \in \beta$ with $\mu(A) > 0$ and $\mu(B) > 0$, then there exists $n > 0$ such that $\mu(T^{-n}A \cap B) > 0$ (i.e., elements of B will eventually enter A).

Proof: (1) $\Rightarrow$ (2): Let $B \in \beta$ such that $\mu(B \triangle T^{-1}B) = 0$. We will define a measurable set C with $C = T^{-1}C$ and $\mu(C \triangle B) = 0$. Let

$$C = \{x \in X : T^n x \in B \text{ i.o.}\} = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} T^{-k}B$$

Then, $T^{-1}C = C$, hence by (1) $\mu(C) = 0$ or $\mu(C) = 1$. Furthermore,

$$\begin{aligned} \mu(C \triangle B) &= \mu\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} T^{-k}B \triangle B\right) \\ &\leq \mu\left(\bigcup_{k=1}^{\infty} T^{-k}B \triangle B\right) \\ &\leq \sum_{k=1}^{\infty} \mu(T^{-k}B \triangle B). \end{aligned}$$

Using induction and the fact that $\mu(E \triangle F) \leq \mu(E \triangle G) + \mu(G \triangle F)$, we can show that for each $k \geq 1$, we have $\mu(T^{-k}B \triangle B) = 0$. Hence, $\mu(C \triangle B) = 0$, which implies $\mu(C) = \mu(B)$. Therefore, $\mu(B) = 0$ or $\mu(B) = 1$.

(2) $\Rightarrow$ (3): Let $\mu(A) > 0$ and $B = \bigcup_{n=1}^{\infty} T^{-n}A$. Then, $T^{-1}B \subset B$. Since T is measure-preserving, then $\mu(B) > 0$ and

$$\mu(T^{-1}B \triangle B) = \mu(B \setminus T^{-1}B) = \mu(B) - \mu(T^{-1}B) = 0$$

Thus, by (2), $\mu(B) = 1$.

(3) $\Rightarrow$ (4): Suppose $\mu(A)\mu(B) > 0$. By (3),

$$\mu(B) = \mu\left(\bigcup_{n=1}^{\infty} T^{-n}A\right) = 1$$

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

(4) $\Rightarrow$ (1): Suppose $T^{-1}A = A$ with $\mu(A) > 0$. If $\mu(A^c) > 0$, then by (4), there exists $k \geq 1$ such that $\mu(A^c \cap T^{-k}A) > 0$. Since $T^{-k}A = A$, it follows that $\mu(A^c \cap A) > 0$, a contradiction. Hence, $\mu(A) = 1$, and T is ergodic.

4.3.2 Theorem:

Let (X,β,μ) be a probability space and $T:X\to X$ be measure-preserving. Then T is ergodic if and only if for every real-valued measurable function f that is invariant under T, f is constant almost everywhere.

Proof: Suppose T is ergodic, and let f be a real-valued measurable function such that $f(Tx)=f(x)$ for all $x\in X$. So the sets $C_a=\{x\mid f(x)<a, a\in\mathbb{R}\}$ are invariant under T, so by ergodicity, $\mu(C_a)=0$ or $\mu(X-C_a)=0$. Now if $a\le b$, then $C_a\subseteq C_b$. Let $\alpha=$ l.u.b. $\{a\mid \mu(C_a)=0\}$. If $a<\alpha<b$, then $\mu(C_a)=0$, $\mu(X-C_b)=0$, so that

$$\mu\{x\mid a<f(x)<b\}=\mu(C_b-C_a)=\mu(C_b)=\mu(X).$$

So the set

$$\{x\mid f(x)=\alpha\}=\bigcap_{n=1}^{\infty}\{x\mid \alpha-1/n<f(x)<\alpha+1/n\}$$

has full measure. So f is constant almost everywhere.

Conversely, suppose every T-invariant real-valued measurable function on X is constant almost everywhere. Let $A\in\beta$ such that $T^{-1}A=A$. Then the function 1_A is real-valued, measurable, and T-invariant, hence must be constant almost everywhere. The constant must be zero or one because 1_A takes values 0 or 1. If $1_A(x)=0$ almost everywhere, then $\mu(A)=0$, and if $1_A(x)=1$ almost everywhere, then $\mu(X-A)=0$. Hence, the proof.

5. Ergodic Theorems and Their Consequences

5.1 Ergodic Theorem for Permutations

Before going to the measure-theoretic setup, let us check one form of the ergodic theorem on a finite set with a permutation of its elements. Suppose X is a finite set $X=\{x_1,x_2,\dots,x_N\}$ and σ is an irreducible permutation on X. Let f be a real-valued function on X. Then, we can see that the limit

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

exists for all $x\in X$ and is equal to $\frac{1}{N}(f(x_1)+\dots+f(x_N))$. To see this, we have from the division algorithm $n=Nl+r$, where $0\le r<N$. Since N is fixed and $0\le r<N$, if $n\to\infty$, then $l\to\infty$.

Now, since σ is irreducible, for any $x\in X$, $\sigma^N x=x$, and likewise $\sigma^{2N}x=x,\dots,\sigma^{lN}x=x$. Hence,

$$\begin{aligned}&\sum_{k=0}^{n-1}f(\sigma^k x)\\ &=f(x)+\dots+f(\sigma^{N-1}x)\\ &\quad+f(x)+\dots+f(\sigma^{N-1}x)\\ &\quad\vdots\\ &\quad+f(x)+\dots+f(\sigma^{r-1}x)\end{aligned}$$

Therefore,

$$\frac{1}{n}\sum_{k=0}^{n-1}f(\sigma^k x)=\frac{1}{Nl+r}l(f(x)+\dots+f(\sigma^{N-1}x))+\frac{1}{Nl+r}(f(x)+\dots+f(\sigma^{r-1}x)).$$

Now, since r is bounded and $l\to\infty$ as $n\to\infty$, we have

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}f(\sigma^k x)=\frac{1}{N}(f(x_1)+\dots+f(x_N)).$$

Now, we will move to the measure-theoretic setup. Let $(X,\mathscr{B},\mu)$ be a probability space, T be a measure-preserving transformation on X, and f be a real-valued function on X. As we have done in the finite set with the permutation as above, here also we can raise the same question: Does the

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

exist, or find the conditions under which this limit exists?

This question was first raised by Boltzmann and Gibbs in their statistical mechanics work. In 1931, Birkhoff proved that this limit exists for any T and f almost everywhere.

5.2 Birkhoff Ergodic Theorem

Theorem: Let $(X,\mathscr{B},\mu)$ be a probability space, and $T:X\to X$ be a measure-preserving transformation. Then, for any $f\in L^1(\mu)$, the limit

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

exists almost everywhere, is T-invariant, and $\int_X f\,d\mu=\int_X f^*\,d\mu$. Moreover, if T is ergodic, then f^* is constant almost everywhere, and $f^*=\int_X f\,d\mu$.

Proof: The proof involves using a lemma and showing that f^* exists, is integrable, and T-invariant.

Lemma: Let M be a positive integer, and $\{a_n\}_{n\ge0}$ and $\{b_n\}_{n\ge0}$ be two sequences of non-negative real numbers such that for each $n=0,1,2,\dots$, there exists an integer $1\le m\le M$ with

$$a_n+\dots+a_{n+m-1}\ge b_n+\dots+b_{n+m-1}.$$

Then, for each positive integer $N>M$, we have

$$a_0+\dots+a_{N-1}\ge b_0+\dots+b_{N-M-1}.$$

Proof of the Lemma: By using the hypothesis recursively, we can find integers $m_0<m_1<\dots<N$ such that

$$m_0\le M,\quad m_{i+1}-m_i\le M \text{ for } i=0,1,\dots,k-1, \text{ and } N-m_k<M,$$

$$\begin{aligned}a_0+\dots+a_{m_0-1}&\ge b_0+\dots+b_{m_0-1},\\ a_{m_0}+\dots+a_{m_1-1}&\ge b_{m_0}+\dots+b_{m_1-1},\\ &\vdots\\ a_{m_{k-1}}+\dots+a_{m_k-1}&\ge b_{m_{k-1}}+\dots+b_{N-M-1}.\end{aligned}$$

Therefore,

$$\begin{aligned} a_0 + \ldots + a_{N-1} &\geq a_0 + \ldots + a_{m_k-1} \\ &\geq b_0 + \ldots + b_{m_k-1} \\ &\geq b_0 + \ldots + b_{N-M-1}. \end{aligned}$$

Proof of the Theorem: The proof involves showing that $\int_X f^* \, d\mu \leq \int_X f \, d\mu$ and $\int_X f \, d\mu \leq \int_X f^* \, d\mu$.

Corollary: Let $(X, \mathscr{B}, \mu)$ be a probability space, and $T : X \to X$ be a measure-preserving transformation. Then, T is ergodic if and only if for all $A, B \in \mathscr{B}$, we have

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

Proof: The proof of this corollary involves showing the equivalence between ergodicity and the given condition.

6. Conclusion

The notion of ergodicity and ergodic theorems has a lot of applications in different branches of mathematics. In this chapter, we have seen the ergodic behavior of the Continued Fraction Transformation. By using the Birkhoff Ergodic Theorem, we can prove that for almost every $x \in [0,1]$, the elements of the continued fraction expansion of x are unbounded. This is an application of ergodic theorems in number theory. Likewise, we can use ergodic theory to solve many problems in number theory. In probability theory, we have many applications of ergodic theory. For example, the strong law of large numbers can be viewed as a special case of the Birkhoff Ergodic Theorem. Other branches of mathematics like hyperbolic geometry, differential geometry, statistical mechanics, functional analysis, etc., also have a great demand for ergodic theorems. A limitation of the Birkhoff Ergodic Theorem is that it only gives assurance of almost everywhere convergence. To extend the statement of the Birkhoff Ergodic Theorem to have convergence everywhere on the measure space, we need some stronger condition like unique ergodicity on the transformation.

References

[1] J.F.C. Kingman and S.J. Taylor. *Introduction to Measure and Probability*. Cambridge Press, 1966.

[2] P. Walters. *An Introduction to Ergodic Theory*. Graduate Texts in Mathematics. Springer-Verlag, New York-Berlin, 1982.

[3] C. S. Aravinda and V. S. .Bhat. Basic ergodic theory. Lecture Note in 'Elements of Dynamical Systems' (Eds. R. Shah, A. Nagar, and S. Sridharan), 2015.

[4] M.G. Nadkarni. *Basic Ergodic Theory*. Hindustan Book Agency, New Delhi, India, third edition, 2012.

[5] M. Einsiedler and T. Ward. *Ergodic Theory with a View Towards Number Theory*. Springer, New York, 2010.