MEASURE THEORETIC ASPECTS OF DYNAMICAL SYSTEMS

WATEE SRININ

ABSTRACT. In this exposition, we will present the fundamental theorem in Ergodic theory and its applications.

CONTENTS

1. INTRODUCTION TO DYNAMICAL SYSTEM

Originally, dynamical systems is a discipline that studies the movement of some physical systems through time e.g. the movement of celestial bodies. From mathematical point of view, the movement of physical systems can be described in term of the self-mapping of a space. Hence, we come to the formal definition of a dynamical system.

Definition 1.1. A dynamical system, denoted by (X, f), consists of a non-empty set X called *phase space*, whose elements represent possible state of the system, and a collection of self-mapping $\{f^t | f^t : X \longrightarrow X\}$.

Note that the collection of maps cannot be arbitrary. In fact, the collection of maps must have a group or a semigroup structure. That is if f^s and f^t belongs to $\{f^t\}$ then $f^{s+t} := f^s \circ f^t$ belongs to $\{f^t\}$ where $\circ$ is standard composition. The associativity property thus immediately holds. Also we define f^0 to be an identity map. With these properties hold, the collection of maps have a semigroup structure. If every f^s has an inverse denoted by f^{-s}, then such collection becomes a group.

Although we have defined dynamical systems in a completely abstract setting, where a phase space X is simply a set, in practice a phase space X usually come with an additional structure that is preserved under the mapping. For example, (X, f) could be a measure space and a measure preserving map, a topological space and a continuous map, a metric space and isometry, or a smooth manifold and a differentiable map. In this paper, we will be interested in Lebesgue measurable space.

We can classify dynamical systems according to the group structure of the collection of the maps.

Definition 1.2. A **discrete-time dynamical system** is a dynamical systems whose $t \in \mathbb{Z}$. In a discrete dynamical systems, we do not have to specify the entire member in the collection since $\{f^t\}$ can be generated by single element, f.

Definition 1.3. A **continuous-time dynamical systems** is a dynamical systems whose $t \in \mathbb{R}$ or $t \in \mathbb{R}_0^+$.

With the following terminologies we can discuss the behavior of a point or a set of points through the process of iteration.

Definition 1.4. For $x \in$ X, the **positive semiorbit** $\mathcal{O}_f^+(x) = \bigcup_{t \geq 0} f^t(x)$. Likewise, the **negative semiorbit** $\mathcal{O}_f^-(x) = \bigcup_{t \leq 0} f^t(x)$

Definition 1.5. The orbit of the point $x \in$ X is $\mathcal{O}_f^+(x) \bigcup \mathcal{O}_f^-(x) = \bigcup_t f^t(x)$

Definition 1.6. A point $x \in$ X is a **periodic point** of *period* T > 0 if $f^T(x) = x$. The orbits of a period point is called a **periodic orbit**.

Definition 1.7. A subset $A \subset X$ is **f-invariant** if $f^t(A) \subset A$

Example 1.8. (Circular Rotation) Consider the unit circle $S^1 = \{(x.y)|x^2 + y^2 = 1\}$ or in complex plane $S^1 = \{z \in \mathbb{C}| \ \|z\| = 1\}$. For $\alpha \in \mathbb{R}$, let R_α be the rotation of S^1 by an angle α. Its action on a point in a space is given as

$$R_\alpha(x = e^{i\theta}) = e^{i(\theta + \alpha)}$$

If α is a rational number i.e. there exist $p, q \in \mathbb{Z}$ such that $\alpha = p/q$, then $R_\alpha^q = R_\alpha^0 = Id$. It follows that every point is a periodic point of period T $= p$. On the other hand, if α is a irrational number, then there is no periodic orbit i.e. the original point will move around in the space S^1 never to return to its original position.

However, we can formulate the following questions which are related to the behavior of non-periodic point.

Question 1 Suppose that p is a non-periodic point. How close does this point is to the original point?

Question 2 What is the distribution in the space of such point? Specifically, does it distribute evenly throughout the space? If not, which part in the space that such point spend time the most?

These questions aim at describing the statistical behavior of a dynamical system. We can formulate these questions in a precise way.

Given any function $f : X \to X$, any orbit $x_0 \mapsto f(x_0) = x_1 \mapsto f(x_1) = x_2 \mapsto ...$ and any real valued function $\varphi : X \to \mathbb{R}$, we can try to form the limit

$$A(x_0) = \lim_{n \to \infty} \frac{1}{n}(\varphi(x_0) + \varphi(x_1) + ... + \varphi(x_{n-1})) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(x_k)$$

If this limit exist, then it is called **time average** of φ over the forward orbit of x_0

To formulate its counterpart, **space average**, is a bit more complicated. We have to borrow a tool from measure theory. Now suppose that $(X, \mathcal{B}, \mu)$ is a finite measure space, and that φ is an integrable function. Then the **space average** of φ is defined to be the ratio $(\int_X \varphi d\mu)/\mu(X)$.

The basic goal in this exposition is prove the fundamental result, commonly known as Birkhoff Ergodic Theorem, which, roughly speaking, specifies the conditions under which space averages are equal to time average.

2. Basics Measure Theory

Let X be a set. A σ-algebra of subsets of X is a set $\mathcal{B}$ of subsets of X i.e. $\mathcal{B} \in \mathcal{P}(X)$ satisfying the following conditions:

(i) $X \in \mathcal{B}$.

(ii) If $B \in \mathcal{B}$, then so is B^c.

(iii)If $B_n \in \mathcal{B}$ for all $n \in \mathbb{N}$, then so is $\bigcup_{n \in \mathbb{N}} B_n$

These properties imply that

(iv) $\emptyset \in \mathcal{B}$

(v) If $B_1, ..., B_n \in \mathcal{B}$, then so is $\bigcap_{j=1}^{n} B_j$. This is also true for infinite collection.

We denoted such space with its σ-algebra as $(X, \mathcal{B})$, and then called it **a measurable space**. The elements of $\mathcal{B}$ are called **measurable set**. If X is a topological space, there is a natural σ-algebra that we want to work with. We shall always consider the σ-algebra of a Borel sets, i.e. the smallest σ-algebra containing all open subsets of X.

A function, μ, called a **measure** in X, is a function that assigns to each elements in set $\mathcal{B}$ a non-negative number i.e. $\mu : \mathcal{B} \to \mathbb{R}_{\geq 0}$. Moreover, such measure have to satisfying the following condition:

(i) $\mu(\bigcup_{n \in \mathbb{N}} B_n) = \sum_{n \in \mathbb{N}} \mu(B_n)$, if $B_i \cap B_j = \emptyset$ for all i$\neq$j.

(ii)$\mu(\emptyset) = 0$

Definition 2.1. A measure μ is finite if $\mu(X) \leq \infty$. Measurable space together with finite measure is called finite measurable space.

Definition 2.2. if $\mu(X) = 1$, then a triple $(X, \mathcal{B}, \mu)$ is called probability space.

Theoretically, our measure can be as bizarre as we want as long as it satisfies the above conditions. However, in this exposition we will work with Lesbesgue measure.

Definition 2.3. The σ-algebra of subset of $\mathbb{R}^n$ generated by open set and null sets will be denoted by $\mathcal{M}$. Sets in $\mathcal{M}$ will be called Lebesgue measurable.

We can define a measure, $\mu : \mathcal{M} \to \mathbb{R}_{\geq 0}$, called Lebesgue measure, as the following.

Definition 2.4. For any subset A of $\mathbb{R}^n$, we can define its outer measure $\mu^*(A)$.

$$\mu^*(A) = \inf\{\sum_{B \in \mathcal{C}} Vol(B) : \mathcal{C} \text{ is a countable collection of boxes whose union covers } A\}$$

where B is a set, called **box**, of the form

$$B = \prod_{i=1}^{n}[a_i, b_i]$$

The volume vol(B) of this box is defined to be

$$\prod_{i=1}^{n}(b_i - a_i)$$

This measure also have the following properties.

(i) Suppose that A$\in \mathcal{M}$, and x $\in \mathbb{R}^n$, then $\mu(A + x) = \mu(A)$

(ii) $\mu(\bigcup_{n \in \mathbb{N}} B_n) \leq \sum_{n \in \mathbb{N}} \mu(B_n)$

(iii)If A , B $\in \mathcal{M}$ and A $\subset$ B, then $\mu(A) \leq \mu(B)$

(iv) A is called a null set if and only if $\mu(A)=0$

(v) If A $\in \mathcal{M}$, then for any $\epsilon > 0$, then there exist an open set U containing A such that $\mu(U\backslash A) < \epsilon$

Not every function are integrable. In fact, there exists a family of function called **measurable function** in which we can define the Lebesgue integral.

Definition 2.5. Let $(X, \mathcal{B}, \mu)$ be a measure space. If $f : X \to \mathbb{R}_{\geq 0}$, we say that f is $\mathcal{B}$-measurable provided that

$$f^{-1}(-\infty, a) = \{x \in X | f(x) < a\} \in \mathcal{B}$$

Now, we can integrate function in this family with respect to the measure we used. The integral of a measurable function $f : X \to \mathbb{R} \cup \{\infty\}$ on a measure space $(X, \mathcal{B}, \mu)$ is usually written

$$\int_X f \ d\mu$$

It is defined as the following

(i) If $f = \chi_A$ is the **characteristic function of a set A** $\in \mathcal{B}$, then set

$$\int_X \chi_A \ d\mu = \mu(A).$$

where

$$\chi_A(x) = \begin{cases} 1 & \text{when } x \in A \\ 0 & \text{when } x \in X \backslash A \end{cases}$$

(ii) If f is a simple function i.e. f can be written as $f = \sum_{k=1}^n c_k \chi_{A_k}$ where $c_k \in \mathbb{R}$ for some finite collection $A_k \in \mathcal{B}$, then define

$$\int_X f d\mu = \sum_{k=1}^n c_k \int_X \chi_{A_k} d\mu = \sum_{k=1}^n c_k \mu(A_k).$$

(iii) If f is a nonnegative measurable function (possibly attaining the value ∞ at some points), then we define

$$\int_X f d\mu = \sup \left\{ \int_X h d\mu : h \text{ is simple and } h(x) \leq f(x) \text{ for all } \ x \in X \right\}.$$

(iv) For any measurable function f (possibly attaining the values ∞ or ?∞ at some points), write $f = f^+ - f^-$ where

$$f^+ = \max(f, 0) \quad \text{and} \quad f^- = \max(-f, 0)$$

so that $|f| = f^+ + f^-$, and define the integral of f as

$$\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu$$

provided that $\int_X f^+ \ d\mu$ and $\int_X f^+ \ d\mu$ are not both ∞

The following theorem addresses the conditions under which the limit and integral can be interchanged. We will use this theorem in proving the Birkhoff Ergodic Theorem.

Theorem 2.6. *(Lebesgue's dominated convergence theorem) Let $f_n : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a sequence of integrable functions which converges on $\mathbb{R}^n$ point wise almost everywhere to a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$. Also, assume that there is an integrable function $G : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ with $|f_n| \leq G$ for all $n \in \mathbb{B}$. Then, f is integrable and $\int_{\mathbb{R}^n} f(x)dx = \lim_{n \to \infty} \int_{\mathbb{R}^n} f_n(x)dx$*

The proof of this theorem and other topics in measure theory can be found in [4] and [5].

3. Ergodic Theory

Ergodic theory is the statistical study of groups of motions of a space with measurable structure on it. The word ergodic was introduced by Ludwig Boltzman in the context of the statistical mechanics of gas particle, and it comes from two Greek words "ergon"(work) and "odos"(path).

However,the mathematical setting in which ergodic theory is studied is $(X, \mathcal{B}, \mu, \{T^n\})$ or $(X, \mathcal{B}, \mu, \{f^t\})$, where $\{T^n\}$ and $\{f^t\}$ are measure preserving transformation.

Definition 3.1. Suppose that $\mathcal{B}$ is a σ-algebra of X and μ is a finite measure defined on $\mathcal{B}$.
(i) A function $T : X \to X$ is called a measure preserving provided that for each $B \in \mathcal{B}$ the set $T^{-1}(B) \in \mathcal{B}$ and $\mu(T^{-1}(B)) = \mu(B)$.
(ii) A function $\varphi : X \to \mathbb{C}$ which satisfies $\varphi(x) = \varphi(T(x))$ for μ-almost all x is called T-invariant.
(iii) A set A is called T-invariant if $\chi_A(x)$ is a T-invariant i.e. $\chi(x) = \chi(T(x))$ for μ-almost all x.

Definition 3.2. Suppose that $T : X \to X$ is a μ-measure preserving transformation. A point $x \in A \subset$ X is said to be recurrent for T with respect to μ-measurable set A provided that the set of return times, $R(x) = \{n | T^n(x) \in A, n \in \mathbb{N}\}$ is infinite.

With the terminology given above we can discuss the solution to the first question stated in the introduction part.

Theorem 3.3. *(Poincaré recurrence) Suppose $T : X \to X$ is a μ-measure preserving transformation of a finite measure space and suppose that $A \subset X$ is μ-measurable. Then μ almost all $x \in A$ are recurrent for T with respect to A.*

Proof. Let $\acute{A} = $ A $\cap \bigcap_{N=0}^{\infty} \bigcup_{n=N}^{\infty} T^{-n}(A)$. We want to show that $\mu(\acute{A}) = \mu(A)$ since this implies that the iterates $\{T^n\}$ map almost every $x \in$ A infinitely back into A. For each $x \in \acute{A}$, there exist arbitrarily large m $\in \mathbb{N}$ with $T^m(x) \in A$, and the iterates of T thus map any such x infinitely many times to A. Since

$$T^{-1}(\bigcup_{n=N}^{\infty} T^{-n}(A)) = \bigcup_{n=N+1}^{\infty} T^{-n}(A) \subset \bigcup_{n=N}^{\infty} T^{-n}(A)$$

and T is a measure preserving transformation, it follows that

$$\mu(\bigcup_{n=0}^{\infty} T^{-n}(A)) = \mu(\bigcup_{n=N}^{\infty} T^{-n}(A)) = \bigcap_{N=0}^{\infty} \bigcup_{n=N}^{\infty} T^{-n}(A) \text{for all n} \in \mathbb{N}$$

and because A $\subset \bigcup_{n=o}^{\infty} T^{-n}(A)$, therefore $\mu(\acute{A}) = \mu(A)$. $\square$

However, we can answer the same question in a more precise manner.This lead us to the main theorem in this exposition which aim at answering the second question stated in the introduction part.

Theorem 3.4. *(Birkhoff Ergodic Theorem) Let T be a measure-preserving transformation in a finite measure space $(X, \mathcal{B}, \mu)$. For any integrable function φ i.e. $\varphi \in L^1(X, \mathcal{B}, \mu)$, then the time average*

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

exits for almost every x. Moreover, A is integrable with respect to μ and is a T-invariant function i.e. $A \circ T(x) = A(x)$ for all most every x, and satisfies

$$\int_X A(x)d\mu = \int_X \varphi(x)d\mu$$

The argument used in this proof will be based on the following lemma.

Lemma 3.5. *(Existence of the "positive" orbits) Let T be a measure-preserving transformation, and let φ be integrable with$\int_X \varphi(x)d\mu > 0$. Then there exists an orbit $x_0 \mapsto T(x_0) = x_1 \mapsto T(x_1) = x_2 \mapsto ...$ which satisfies the inequality*

$$\varphi(x_0) + \varphi(x_1) + ... + \varphi(x_{n-1}) > 0$$

for all $n \geq 1$

Proof. **(Set up)** This lemma can be proved by contradiction. Suppose to the contrary. It follows that for every $x_0 \in X$ there exists an integer n with $\varphi(x_0) + \varphi(x_1) + ... + \varphi(x_{n-1}) \leq 0$.We must show that $\int_X \varphi(x)d\mu \leq 0$.

(Step 1) Let us suppose for a moment that we have a stronger condition. That is there exist a constant $k > 1$ so that for every $x_0 \in X$ there exists an integer $1 \leq n \leq k$ with $\varphi(x_0) + \varphi(x_1) + ... + \varphi(x_{n-1}) \leq 0$. This assumption guarantees that our n for each x can not exceed the value k.

(Step 1.1) Under this stronger assumption, we will prove that the inequality

$$\sum_{j=0}^{N-1} \varphi(x_j) \leq \sum_{j=N-k}^{N-1} |\varphi(x_j)|$$

is true for some fixed orbit and any positive integer N. It follows from this assumption that for each integer $p > 0$, we can always find an integer q with $p < q \leq p + k$ such that $\sum_{j=p}^{q-1} \varphi(x_j) \leq 0$. From this relation, we can construct a sequence of integers $0 = p_0 < p_1 < p_2 < ...$ with $p_{i+1} \leq p_i + k$, and with $\sum_{j=p_i}^{p_{i+1}-1} \varphi(x_j) \leq 0$. Summing this last inequality for $0 \leq i < l$, it follows that

$$\sum_{j=0}^{p_l-1} \varphi(x_j) \leq 0$$

Now given an arbitrary large N, we can choose p_j so that $N - k \leq p_j \leq N$. It follows that

$$\sum_{j=0}^{N-1} \varphi(x_j) = \sum_{j=0}^{p_l-1} \varphi(x_j) + \sum_{p_l}^{N-1} \varphi(x_j) \leq 0 + \sum_{p_l}^{N-1} (|\varphi(x_j)|) \leq \sum_{N-k}^{N-1} (|\varphi(x_j)|)$$

This proves the inequality above.

(Step 1.2) Now consider both sides of inequality (3.5) as a function of $x_0 \in X$. We can then integrate both side with measure, μ, over the whole space, X. We obtain

$$N \int_X \varphi(x)d\mu \leq k \int_X |\varphi(x)|d\mu$$

$$\int_X \varphi(x)d\mu \leq \frac{k}{N} \int_X |\varphi(x)|d\mu$$

Since k is fixed and N can be arbitrary large, this proves that $\int_X \varphi(x)d\mu \leq 0$

(Step 2) We consider the sequence of measurable real-valued functions φ_k where

$$\varphi_k(x_0) = \begin{cases} \varphi(x_0) & \text{if there exist} 1 \leq n \leq k \quad \text{with} \quad \varphi(x_0) + \varphi(x_1) + ... + \varphi(x_{n-1}) \leq 0, \\ 0 & \text{otherwise.} \end{cases}$$

From this construction it is clear that φ_k is integrable since $|\varphi_k| \leq |\varphi|$. We can see that $\varphi_1 \leq \varphi_2 \leq \varphi_3 \leq ... \leq \varphi$. Since each φ_k satisfies the assumption assume in step 1.1, it follows that $\int_X \varphi_k(x)d\mu \leq 0$. On the other hand, it follows from original assumption that φ_k converge pointwise to φ. Therefore, from Lebesgue's dominated convergence theorem, $\int_X \varphi(x)d\mu = \lim_{n\to\infty} \int_X \varphi_k(x)d\mu \leq 0$. $\qquad\square$

We now provide a proof of Birkhoff Ergodic Theorem

Proof. We can form the upper an lower time averages

$$A^+ = \limsup_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(T^k(x)) \qquad A^- = \liminf_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(T^k(x))$$

Where

$$-\infty \leq A^-(x) \leq A^+(x) \leq +\infty$$

We can see that both A^+ and A^- are measurable functions since $A_n = \frac{1}{n} \sum_{k=0}^{n-1} \varphi(T^k(x)$ is measurable, and that both are T-invariant i.e. $A^\pm(T(x)) = A^\pm(x)$.

Case 1 Suppose that A^+ and A^- are bounded. Then, they are integrable because we have assume that $\mu(X)$ is finite. We will show that

$$\int_X A^+(x)d\mu \leq \int_X \varphi(x)d\mu.$$

Suppose not. This implies that we could chose $\epsilon > 0$ so that

$$\int_X A^+(x)d\mu > \int_X (\varphi + \epsilon)(x)d\mu.$$

By lemma 3.4, using A^+ as φ, we could find an orbit $x_0 \mapsto T(x_0) = x_1 \mapsto T(x_1) = x_2 \mapsto ...$ such that

$$\sum_{k=0}^{n-1} A^+(T^k(x_0)) = \sum_{j=0}^{n-1} A^+(x_j) > \sum_{j=0}^{n-1} (\varphi+\epsilon)(x_j) = \sum_{k=0}^{n-1} (\varphi+\epsilon)(T^k(x_0)) \quad \text{for every } n > 0$$

Since A is T-invariant, the LHS of this inequality equals $nA^+(x_0)$. Now, dividing by n, this yields

$$A^+(x_0) > \frac{1}{n} \sum_{k=0}^{n-1} (\varphi + \epsilon)(T^k(x_0))$$

Now take lim sup

$$A^+(x_0) = \limsup_{n\to\infty} A^+(x_0) > \limsup_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (\varphi + \epsilon)(T^k(x_0)) \geq A^+(x_0) + \epsilon$$

which is impossible. Using analogous statement for the lower time average, we can show that

$$\int_X A^+(x)d\mu \leq \int_X \varphi(x)d\mu \leq \int_X A^-(x)d\mu \leq \int_X A^+(x)d\mu$$

Hence all three integral are equal. Consequently, $A^+(x) = A^-(x)$ except on a set of measure zero.

Case 2 Suppose that A^+ and A^- are unbounded. For each positive integer n, let X_n be the set of points for which

$$-n \leq A^-(x) \leq A^+(x) \leq n$$

Both A^+ and A^- defined this way are measurable and T-invariant set. Hence, we can apply the same argument to conclude that $\lim A(x)$ exists for almost all $x \in X_n$, and that $\int_{X_n} A(x)d\mu = \int_{X_n} \varphi(x)d\mu$. It follows that this is true for union of X_n, that is, for the set of all x satisfying

$$-\infty \leq A^-(x) \leq A^+(x) \leq +\infty$$

We need only check that the functions $A^{\pm}$ take finite value, except on a set of measure zero. Suppose not. Let N be the invariant set consisting of points x for which $A^+ = +\infty$. Since $\mu(N) > 0$, we can choose a finite constant c such that $\int_N c > \int_N \varphi(x)$. Applying lemma 3.4, there exist an orbit in N such that

$$nc > \sum_{j=0}^{n-1}(\varphi(x_j)) = \sum_{k=0}^{n-1}(\varphi(T^k(x_0))) \quad \text{for every } n > 1$$

Now, dividing by n and the $\limsup$, this yields

$$c \geq A^+(x_0)$$

which contradicts to the assumption that $A^+(x_0) = \infty$ where $x_0 \in \mathbb{N}$ □

4. APPLICATION

Poincaré Recurrence theorem tells us the conditions under which the elements in a measurable subset A of X return again an again to a measurable set A. However, using Birkhoff Ergodic Theorem, we can describe such phenomena more precise at least in the special case where T is an ergodic transformation.(Note that when we formulate the theorem we require T to be only measure preserving transformation.)

Definition 4.1. Suppose that $T : X \to X$ is a measure preserving transformation for a finite measure μ defined on σ-algebra $\mathcal{B}$ of subsets of X. Then, T is called **ergodic** if every T-invariant set $A \in \mathcal{B}$ is either $\mu(A) = 0$ or $\mu(A^c) = 0$.

Proposition 4.2. *Suppose that T is a measure preserving in a finite measure space μ. Then T is ergodic if and only if every measurable function φ which is T-invariant is constant except on a set μ of measure 0.*

Proof. ($\Longleftarrow$) Suppose that only T-invariant functions are μ-almost every where constant. A set A is T-invariant only if the function χ_A is T-invariant. Since $\chi_A(x)$ takes on only 1 and 0, the function χ_A must equal to 0, except on a set whose measure is 0, or to 1, except on a set whose measure is 0. Hence, $\mu(A) = 0$ or $\mu(A^c) = 0$.
($\Longrightarrow$) Suppose not. If φ is a T-invariant measurable function that is not μ almost

everywhere constant, then there is a constant $c \in \mathbb{R}$ such that if $A = \varphi^{-1}([0, c))$, then $\mu(A) > 0$ and $\mu(A^c) > 0$. Hence, the set A is T-invariant. $\square$

We ask the following question. If we consider the $\{T^k(x)|0 < k < n - 1\}$, the first n points in forward orbit of $x \in A$ where A is a measurable set , and let $N_n(x)$ denote the number of those points which lie in A, then we would like to know if the limit

$$\lim_{n \to \infty} \frac{N_n(x)}{n}$$

exists. If so, how is this value of the limit compare to $\mu(A)$. Using Birkoff Ergodic Theorem, we are able to answer such question in the case that T is ergodic transformation.

Proposition 4.3. *Suppose that T is an ergodic transformation in a finite measure space $(X, \mathcal{B}, \mu)$. For any integrable function $\varphi(x)$, then*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(T^k(x)) = \frac{1}{\mu(X)} \int_X \varphi(x) d\mu$$

Proof. Using Birkhoff Ergodic Theorem, we get the following

$$\int_X A(x) d\mu = \int_X \varphi(x) d\mu$$

Since T is an ergodic transformation, from proposition 4.2, every measurable function φ is constant almost everywhere. This implies that time average is also T-invariant, and therefore constant almost everywhere.

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

$$[\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(T^k(x))] \int_X d\mu = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(T^k(x)) \mu(X) = \int_X \varphi(x) d\mu$$

Divide the RHS by $\mu(X)$,we got the desired result. $\square$

Proposition 4.4. *Suppose that T is an ergodic transformation in a finite measure space and $A \in X$ is a μ-measurable. Let $N_n(x)$ denote the number of points in the set $A \cap \{T^k(x)|0 < k < n - 1\}$. Then for μ almost all $x \in X$*

$$\lim_{n \to \infty} \frac{N_n(x)}{n} = \frac{\mu(A)}{\mu(X)}$$

Proof. Let $\varphi(x) = \chi_A(x)$. We then apply Birkoff Ergodic Theorem and corollary 4.3. We get the following

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \chi_A(T^k(x)) = \lim_{n \to \infty} \frac{N_n(x)}{n} = \frac{\int \chi_A(x) d\mu}{\mu(X)} = \frac{\mu(A)}{\mu(X)}$$

$\square$

Poincaré Recurrence theorem asserts that for μ almost all point $x \in A$ the forward orbit of x for a measure preserving transformation T returns to A infinitely often. However, for an ergodic transformation T we can do much better that is we

can measure how often the forward orbit of a point x not necessarily in A visits the set A.

Acknowledgments. I would like to thank my mentor, David Constantine, for his guidance throughout the summer.

References

[1] W. Milnor, John. Birkhoff Ergodic Theorem. www.math.sunysb.edu/ jack/DYNOTES/dn9.pdf
[2] Brin, Micheal; Stuck, Garrett. Introduction to Dynamical Systems. Cambridge University Press. 2004.
[3] Jürgen, Jost. Dynamical Systems: Examples of Complex Behaviour. Springer. 2005.
[4] Franks, John M. A (Terse) Introduction to Lebesgue Integration. American Mathematical Society Press. 2009.
[5] L. F. Richardson. Measure and Integration: A Concise Introduction to Real Analysis. John Wiley & Sons, Inc. 2009.