Journal of Pure and Applied Mathematics: Advances and Applications
Volume 12, Number 1, 2014, Pages 1-24

ERGODIC THEORY AND MIXING PROPERTIES

HENA RANI BISWAS

Department of Mathematics
University of Barisal
Barisal
Bangladesh
e-mail: henabiswas@aiub.edu

Abstract

In this paper, we discuss ergodic measure and the various types of mixing for measure-preserving transformations. We discuss the partially understood phenomenon of mixing and indicate some of the contrast between the situations for single measure-preserving transformations. After discussion, we see that strong mixing implies weak mixing. Furthermore, weak mixing (and thus also strong mixing) implies ergodicity. We study Birkhoff ergodic theorem and mixing properties. In this article, we solve some problems of ergodic measure and mixing. We observe that Bernoulli shift is strong mixing. We also show that Markov chain is strong mixing if it is irreducible and aperiodic. Actually, this paper is intended to provide motivation for studying ergodic theory and to describe the major ideas of the subject to a general mathematical audience.

1. Introduction

The study of dynamics in a measure space is traditionally called ergodic theory (even when no ergodicity is involved), since the earliest work in this area countered around the problem of understanding the concept of ergodicity. Ergodic theory is the abstract study of the

2010 Mathematics Subject Classification: 37A25.

Keywords and phrases: measurable space, probability space, ergodicity, mixing, spectral isomorphism, discrete spectrum.

Submitted by Kh. Abdul Maleque

Received May 9, 2014

transformations, dealing in particular with their long term average behaviour. Central aspect of ergodic theory is the behaviour of a dynamical system when it is allowed to run for a long period of time. This is expressed through ergodic theorems [2] which assert that, under certain conditions, the time average of a function along the trajectories exists almost everywhere and is related to the space average.

The term ergodic was introduced by Boltzmann [5] in his work on statistical mechanics, where he was studying Hamiltonian systems with large numbers of particles.

Ergodic theory has connections to many areas of mathematics, but primarily to the area of dynamical systems. The fundamental ideas of ergodic theory have been developed in only the last century and there are still many problems of fundamental importance. I interested in ergodic theory on symbolic dynamics which utilizes results from ergodic theory in the study of shifts of finite type.

Ergodic theory have been studied by many authors, notable amongst them are Pollicott and Yuri [2], Walters [4], Parry [11]. In general, the ergodic theorems of Birkhoff and Von Neumam are used in all aspects of dynamical systems and many problems in mathematical physics.

Jason Preszler [9] apply ergodic theory is the study of the qualitative actions of a group on a space. Jakobson [14] discussed ergodic theory of one-dimensional mappings.

In 1935, Hedlund proved the ergodicity of the geodesic flow on the unit tangent bundle of a surface of constant negative curvature; in 1940, Hopf extended the result to establish ergodicity of the geodesic flow on arbitrary manifolds with negative sectional curvature. In this setting, the invariant measure is the Liouville measure.

A related property of a measure-preserving transformation is mixing. In this article, we will outline a rather full range of mixing properties with ergodicity at the weakest end and the Bernoulli property at the strongest end.

In this paper, we will give some examples of mixing. We shall also see that there are connections between the range of mixing properties that we discuss and measure-theoretic entropy. In measure-preserving transformations that arise in practice, there is a correlation between strong mixing properties and positive entropy, although many of these properties are logically independent.

2. Mathematical Preliminaries

There are two main definitions for ergodic theory that are simple to state and provide an adequate general idea of the subject. The first is that ergodic theory is the long-term, qualitative study of the dynamical systems. Alternatively, ergodic theory is the study of the qualitative actions of a group on a space. Clearly, examples are needed to fully understand these loose definitions. In this paper, we will focus on measure-theoretic ergodic theory, thus neglecting other kinds of spaces.

Definition 2.1 (σ-Algebra)**.** A family β of subsets of X is called an σ-**algebra** [3] if and only if

(i) if $B_n \in \beta$ for $n = 1, 2, 3, \ldots,$ then $\bigcap_{n=1}^{\infty} B_n \in \beta,$

(ii) for any $B \in \beta,$ then $X / B \in \beta,$

(iii) the empty set ϕ belongs to $\beta.$

The elements of β are usually referred to as measurable sets.

Definition 2.2 (Measure)**.** A function $\mu : \beta \to \Re^{+}$ is called **measure** [3] on β if and only if

(i) $\mu(B) \geq 0 \;\; \forall B \in \beta,$

(ii) $\mu(\phi) = 0,$

(iii) for any sequence $\{B_n\}$ of disjoint measurable sets $B_n \in \beta$, $n = 1, 2, \ldots$

$$\mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mu(B_n).$$

Definition 2.3 (Measurable space)**.** A measurable space is a set X with collection of subsets β of X such that

(i) $X \in \beta$,

(ii) if $B \in \beta$, then $X - B \in \beta$,

(iii) $B_n \in \beta \Rightarrow \bigcup_{n=1}^{\infty} B_n \in \beta$.

The pair (X, β) is then called **a measurable space.**

Definition 2.4. The triple (X, β, μ) is then called a **finite measure space** [12]. We will usually normalize a finite measure by assuming that $\mu(X) = 1$. With this normalization, μ is called **a probability measure** on (X, β) and (X, β, μ) is called a **probability space.**

For a **probability measure**, note that $0 \leq \mu(B) \leq 1 \; \forall B \in \beta$.

Definition 2.5 (Invariant measures)**.** Let (X, β, μ) be a measure space. Assume that μ is a probability measure, that is, $\mu(X) = 1$.

A measurable map $T : X \to X$ $\left(\text{that is, } T^{-1}\beta \subset \beta\right)$ is said to **preserve the measure** μ if for any $B \in \beta$, we have $\mu(B) = \mu\left(T^{-1}B\right)$. Alternatively, we say that μ is T-invariant.

We note that most measures are not finite. A common example is the Lebesgue measure on open subsets of the real line; $A = \mathfrak{R}$, β is the collection of open subsets of A of the form (a, b) and $\mu(a, b) = b - a$. However, restricting A to the interval $[0, 1]$ in the above example produces a useful finite measure, and in fact a probability measure.

Haar measure is one of the useful measures. Haar measure is not a specific measure but rather a probability measure that is invariant under group actions. Consider the unit circle $S_1 = \{z \in C : |z| = 1\}$ in the complex plane. Thus normalized Lebesgue measure on S_1 corresponds to the Haar measure on S_1. Furthermore, Haar measure is unique, as the following theorem shows.

Theorem 2.1. *Let G be a topological group. There exists a probability measure* μ *defined on the* σ*-algebra* $\beta(G)$ *of Borel subsets of G such that* $\mu(x\,E) = \mu(E)\ \forall\, x \in G\ \forall E \in \beta$ *and* μ *is regular. There is only one regular rotation invariant probability measure on* $(G, \beta(G))$.

Definition 2.6 (Function spaces)**.** Ergodic theory frequently uses Banach spaces of functions defined on a measure space. Let (X, β, μ) be a finite measure space and for some $p \in \mathfrak{R}$, consider the set of functions $f : X \to C$ such $|f|^p$ is integrable. This set forms a vector space and there is an equivalence relation $f \sim g$ iff $f = g$ almost everywhere[1]. We will denote the space of equivalence classes as $L^P(X, \beta, \mu)$ or simply L^P if the measure space is obvious. The function space L^P also has an associated complete norm given by the formula $\|f\|_P = \left(\int |f|^P\, d\mu\right)^{\frac{1}{P}}$, which makes L^P a Banach space. All the functions in L^P will be measurable.

Definition 2.7. A function $f : (X, \beta, \mu) \to \mathfrak{R}$ is measurable if $f^{-1}(D) \in \beta$ whenever D is an open subset of $\mathfrak{R}$.

[1]Almost everywhere, abbr. a.e., means that sets of measure zero can be ignored. Thus if $B = \{x \in A : f(x) \neq g(x)\}$, then $f \sim g$ if and only if $m(B) = 0$.

3. Measure Preserving Transformations, Isomorphism and Conjugacy

The measure preserving transformations (or mpt's) are functions on a measure space that preserve the given measure. Consider a measurable transformation T from (X, β) to itself. Also, T is a **measure preserving** if $T_*(\mu) = \mu,$ or in the other words, if $\mu(B) = \mu\left(T^{-1}(B)\right)$ for every $B \in \beta.$

We say that T is an invertible **measure preserving transformation** if T is bijective and both T and T^{-1} are measure preserving.

We use the notation $T : (X, \beta, \mu) \to (X, \beta, \mu)$ to denote a measure preserving transformation of a probability space to itself. For instance, if X is a topological structure, then β is always the Borel σ-algebra (that is, the σ-algebra generated by open sets).

Definition 3.1. Suppose (X_1, β_1, μ_1) and (X_2, β_2, μ_2) are two probability spaces.

(i) A transformation $T : X_1 \to X_2$ is measurable if $T^{-1}(\beta_2) \subseteq \beta_1$ (that is T is surjective).

(ii) A transformation $T : X_1 \to X_2$ is measure-preserving if T is measurable and $\mu_1\left(T^{-1}(B_2)\right) = \mu_2(B_2)\ \forall B_2 \in \beta_2.$

(iii) A transformation $T : X_1 \to X_2$ is invertible measure-preserving transformation if T is measure-preserving, bijective and T^{-1} is also measure-preserving.

In ergodic theory, we are interested in long term behaviour, so we will focus on measure-preserving transformations from a measure space onto itself, $T : X_1 \to X_1.$ Common examples of such measure-preserving transformations are the identify transformation (which preserve any measure), and rotations of a compact $T(x) = ax$ which preserve Haar measure. Also, a continuous endomorphism of a compact group preserves Haar measure, affine transformations of a compact group, Bernoulli and Markov shifts.

Let us provide more detail about Bernoulli shifts as these are more complicated, but some of the more interesting measure-preserving transformations. Let $k \geq 2$ and define a probability vector $(p_0, \cdots, p_{k-1})$ to be a vector such that $p_i > 0$ and $\sum_{i=0}^{k-1} p_i = 1$. A simple measure space is $(X, 2^X, \mu)$, where $X = \{0, \cdots, k-1\}$ and $\mu(i) = p_i$. To make this space is more complicated, let $(B_k^2, \beta, m) = \prod_{-\infty}^{\infty}(X, 2^X, \mu)$. As mpt on this new space is given by $T : B_k^2 \to B_k^2$ such that $T(\{x_n\}) = \{x_{n+1}\}$. As a concrete example, if $k = 2$, then (B_k^2, β, m) consist of all bi-infinite binary sequences and T shifts a given sequence one space to the left. The transformation T is called the two-sided Bernoulli shift if $k = 2$ and $\left(\frac{1}{2}, \frac{1}{2}\right)$ is the probability vector, or the two-sided $(p_0, \cdots, p_{k-1})$-shift in general.

The above shift can also be simplified into a one-sided shift by considering the measure space $(B_k^2, \beta, m) = \prod_{0}^{\infty}(X, 2^X, \mu)$. In this case, T simply erases the left-most element of an infinite sequence. Markov shifts also occur in one-and two-sided varieties and generalize the above shifts using stochastic matrices [4].

It will be helpful to present a selection of simple examples, relative to which we will be able to explore ergodicity and the various notions of mixing. These examples necessary to show that they are measure-preserving transformations as claimed may be found in the books of Walters [4], Rudolph [6], and Petersen [13].

Measure-theoretic isomorphism is the basic notion of ‘sameness’ in ergodic theory. As an example of measure-theoretic isomorphism, it may be seen that the doubling map is isomorphic to the one-sided Bernoulli shift on $\{0, 1\}$ with $p_0 = p_1 = \frac{1}{2}$ and the inverse map h takes a sequence of 0’s and 1’s to the point in $[0, 1)$ with that binary expansion.

Since in ergodic theory, measure-theoretic isomorphism is the basic notion of sameness, all properties that are used to describe measure-preserving systems are required to be invariant under measure-theoretic isomorphism.

When two measure-preserving transformations are isomorphic, there are certain invariant properties, ergodic theory is very interested in theses invariants and if a weaker relation than isomorphism could also preserve certain properties such as ergodicity or mixing. Let us begin this investigation with the definition of when two measure spaces are isomorphic.

Definition 3.2. Let (X_1, β_1, μ_1) and (X_2, β_2, μ_2) be two probability spaces. These spaces are isomorphic if $\exists\ M \in \beta_1$ and $N \in \beta_2$ such that $\mu_1(M) = 1 = \mu_2(N)$ and there is an invertible measure-preserving transformation $\phi : M \to N$.

The criteria for isomorphism is actually much stronger than is needed to consider two measure spaces "equivalent". All that is needed to consider two measure spaces as equivalent is what is called conjugacy, defined below. Both isomorphism and conjugacy of measure-spaces can be slightly modified for isomorphism and conjugacy of measure-preserving transformations.

Definition 3.3. Let (X_1, β_1, μ_1) and (X_2, β_2, μ_2) be two probability spaces with measure algebras $(\tilde{\beta}_1, \tilde{\mu}_1)$ and $(\tilde{\beta}_2, \tilde{\mu}_2)$. These measure algebras are isomorphic if $\exists$ a bijective measure-preserving transformation $\phi : \tilde{\beta}_2 \to \tilde{\beta}_1$ that preserves complements countable unions and intersections. Two probability spaces are conjugate if their measure algebras are isomorphic.

Definition 3.4. Let (X_1, β_1, μ_1) and (X_2, β_2, μ_2) be two probability spaces with measure-preserving transformation $T_1 : X_1 \to X_1$ and $T_2 : X_2 \to X_2$. Then T_1 and T_2 are isomorphic if $\exists\, M \in \beta_1$ and $N \in \beta_2$ with $\mu_1(M) = 1 = \mu_2(N)$ such that

(i) $T_1 M \subseteq M$, $T_2 N \subseteq N$, and

(ii) $\exists$ an invertible measure-preserving transformation $\phi : M \to N$ such that

$$\phi(T_1(x)) = T_2(\phi(x)) \quad \forall x \in M.$$

Isomorphism is as usual an equivalence relation and if T_1 is isomorphic to T_2, denoted $T_1 \approx T_2$, then all higher iterations of T_1 and T_2 are isomorphic.

Definition 3.5. Let (X_1, β_1, μ_1) and (X_2, β_2, μ_2) be two probability spaces with measure-preserving transformation $T_1 : X_1 \to X_1$ and $T_2 : X_2 \to X_2$. Then T_1 and T_2 are conjugate if there is a measure algebra isomorphism $\phi : \widetilde{\beta}_2 \to \widetilde{\beta}_1$ such that $\phi \widetilde{T}_2^{-1} = \widetilde{T}_1^{-1} \phi$.

Conjugacy is also an equivalence relation and all isomorphic measure-preserving transformations are conjugate. In some cases, conjugacy can also imply isomorphism.

4. Ergodicity and Mixing

Ergodicity and mixing are properties enjoyed by certain measure-preserving transformations. A related (stronger) property of a measure-preserving transformation is mixing. Here one is investigating the correlation between the states of the system at different times. The system is mixing if the states are asymptotically independent: As the times between the measurements increase to infinity, the observed values of the measurements at those times become independent.

We discuss in this article, this definition of ergodicity applies also to infinite measure-preserving transformations and even to certain non-measure-preserving transformations.

Theorem 4.1. *Let* $T : X \to X$ *be a measure-preserving transformation of a probability space* (X, β, μ). *Let* $E \in \beta$ *with* $\mu(E) > 0$. *Then almost all points of E return infinitely often to E under positive iteration of T.*

We will sometimes refer to the path of an element under positive iteration of an measure-preserving transformation as the trajectory, or orbit, of that element. Thus, for any $x \in ET^n(x)$ gives the n-th position in the trajectory of x for a specified n. In general, $T^n(x)$ will be used to denote the entire trajectory, therefore by the above theorem, $T^n(x) \in E$ for infinitely many n if $x \in F$, where $F \subseteq E$ and $\mu(F) = \mu(E)$. It is not hard to see that T is ergodic, if and only if T invariant measurable functions.

Definition 4.1. Let (X, β, μ) be a measure space. A measure-preserving transformation T of the measure space is **ergodic** [12] if the only members $B \in \beta$ such that $T^{-1}B = B$ iff $\mu(B) = 0$ or $\mu(X \setminus B) = 0$.

As for example, Tent map $T(x) = 2x \pmod 1$, $x \in [0, 1)$ is ergodic.

Clearly the above definition holds for infinite measure spaces. Ergodicity is equivalent to a number of different criteria, some in terms of the measure of sets and others in terms of functions. The criteria of functions will be the most useful and easily understood, and are therefore stated below along with the more useful of the measure theoretic versions.

Theorem 4.2. *If* (X, β, μ) *is a probability space and* $T : X \to X$ *is measure-preserving* [12], *then of the following equivalent*:

(i) *T is ergodic.*

(ii) *Whenever f is a measurable and* $(f \circ T)(x) = f(x)\ \forall x \in X$, *then f is constant almost everywhere.*

(iii) *Whenever f is measurable and* $(f \circ T)(x) = f(x)$ *almost everywhere, then f is constant almost everywhere.*

(iv) *Whenever* $f \in L^2(\mu)$ *and* $(f \circ T)(x) = f(x)\ \forall x \in X$, *then f is constant almost everywhere.*

(v) *Whenever* $f \in L^2(\mu)$ *and* $(f \circ T)(x) = f(x)$ *almost everywhere, then f is constant almost everywhere.*

(vi) $\forall A, B \in \beta$ *with* $\mu(A) > 0$ *and* $\mu(B) > 0$ *there exists* $n > 0$ *such that* $\mu\left(T^{-n}A \cap B\right) > 0$.

The intuitive meaning of ergodicity is that if $T : X \to X$ is a continuous ergodic measure-preserving transformation then almost every point of X has a dense trajectory under T. Indeed, nothing is sacrificed in this intuitive understanding as it could be phrased as a theorem and is also the standard method of providing that a transformation is ergodic.

The identity transformation is ergodic if and only if all members have measure 0 or 1. We have already stated that a rotation of the unit circle is ergodic if and only if the rotation is not by a root of unity. This generalizes to the following theorem:

Theorem 4.3. *Let G be a compact group and* $T(x) = ax$ *be a rotation of G. Then T is ergodic if and only if* $\left\{a^n\right\}_{-\infty}^{\infty}$ *is dense in G. Furthermore, if T is ergodic then G is abelian.*

Additionally, the one-and two-sided $(p_0, \cdots, p_{k-1})$ shifts are ergodic. For our other examples, there are known necessary and sufficient conditions for a mpt to be ergodic.

Example 4.1 (Deterministic ergodicity: The logistic map)**.** We have seen that the logistic map, $Tx = 4x(1 - x)$ has an invariant density (with respect to Lebesgue measure). It has an infinite collection of invariant sets, but the only invariant interval is the whole state $[0, 1]$, any smaller interval is not invariant. From this, it is easy to show that all the invariant sets either have measure 0 or measure 1 and they differ from ϕ or from $[0, 1]$ by only a countable collection of points. Hence, the invariant measure is ergodic. Notice that the Lebesgue measure on $[0, 1]$ is ergodic, but not invariant.

Example 4.2 (Invertible ergodicity: Rotations)**.** Let $Tx = x + \phi \bmod 1$, and let μ be the Lebesgue measure on $[0, 1)$. Clearly, T preserves μ. If ϕ is rational, then for any x, the sequence of iterates will visit only finitely many points, and the process is not ergodic, because we can construct invariant sets whose measure is neither 0 nor 1. If, on the other hand, ϕ is irrational, then $T^n x$ never repeats, and it is easy to show that the process is ergodic, because it is metrically transitive.

This example is very important in physics, because many mechanical systems can be represented in terms of "action-angle" variables, the speed of rotation of the angular variables being set by the actions, which are conserved, energy-like quantities.

We have now arrived at the first central results of ergodic theory, which were first proven independently by Birkhoff and Von Neumann.

Theorem 4.4 (Birkhoff ergodic theorem)**.** *Suppose* $T : X \to X$ *is measure-preserving and* $f \in L^1(\mu)$. *Then* $\frac{1}{N}\sum_{n=0}^{N-1} f(T^n(x))$ *converges almost everywhere to a function* $f^* \in L^1(\mu)$. *Also* $f^* \circ T = f^*$ *almost everywhere and if* $\mu(X) < \infty$, *then* $\int f^* d\mu = \int f \, d\mu$.

Clearly if T is ergodic, then f^* is constant almost everywhere and if $\mu(X) < \infty$, then $f^* = \left(\frac{1}{\mu(X)}\right)\int f d\mu$. Furthermore, if (X, β, μ) is a probability space and T is ergodic, then

$$\forall f \in L^1(\mu) \quad \lim_{N\to\infty}\left(\frac{1}{N}\right)\sum_{n=0}^{N-1} f(T^n x) = \int f\, d\mu \text{ a.e..} \tag{1}$$

The Birkhoff ergodic theorem has many uses, particularly in statistical mechanics, but also to number theory and dynamical systems. Theses applications can be found in the literature.

The following corollary is due to Von Neumann.

Corollary 4.1 {(L^p Ergodic theorem of Von Neumann) [8]}**.** *Let $1 \le p < \infty$ and let T be a measure-preserving transformation of the probability space (X, β, μ). If $f \in L^P(\mu)$ $\exists$ $f^* \in L^P(\mu)$ with*

$$f^* \circ T = f^* \text{ a.e. } \textit{and } \left\| \left(\frac{1}{N}\right)\sum_{n=0}^{N-1} f\left(T^n(x)\right) - f^*(x) \right\|_p \to 0.$$

The theorem of Von Neumann was published a year before Birkhoff's result.

The next corollary provides yet another criterion for ergodicity.

If the measure-preserving transformation T is ergodic, then by virtue of Theorem 4.2 (iii), the limit functions appearing in the ergodic theorems are constant.

Corollary 4.2. *Let (X, β, μ) be a probability space and let $T : X \to X$ be a measure-preserving transformation. Then T is ergodic if and only if $\forall A, B \in \beta$*

$$\frac{1}{N}\sum_{n=0}^{N-1} \mu\left(T^{-n}A \cap B\right) \to \mu(A)\,\mu(B). \tag{2}$$

The convergence in the above corollary can be changed to yield different notions: weak and strong mixing.

Definition 4.2. Let T be a measure-preserving transformation of a probability space (X, β, μ).

(i) T is weak-mixing if $\forall A,\ B \in \beta$

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

(ii) T is strong mixing if $\forall A,\ B \in \beta$

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

We will see that every strong-mixing transformation is weak-mixing, and every weak-mixing transformation is ergodic. This can be proven by simply considering the limit of a sequence of real numbers. However, the converse is not true since a rotation of the unit circle by a non-root of unity is ergodic but not weak-mixing. The details of this are worked out in [4] and from it we should gain the intuitive understanding that a weak-mixing transformation does some stretching of the input set. Examples of weak-mixing transformations that are not strong-mixing exist, but are very complicated. Despite this, we can topologize the space of invertible mpt's of a measure space with weak topology and then show that the class of strong-mixing transformations is of first category; thus the most transformations are weak-mixing.

Now we will solve some problems with solutions related to mixing.

Problem 4.1. Show that mixing implies weak mixing and that weak mixing implies ergodicity.

Solution. Suppose that a probability measure μ is mixing [15] with respect to a mapping $f : X \to X$. Let A, B be two measurable sets. Mixing then means that $\left|\mu\left(f^{-n}(A) \cap B\right) - \mu(A)\mu(B)\right| \to 0$ as $n \to \infty$.

Hence, the averages of this sequence also converge to 0, which means that μ is weakly mixing. To prove that weak mixing implies ergodicity, apply the definition of weak mixing for $A = B$: given that $f^{-1}(A) = A$, we conclude $\mu(A) = \mu(A)^2$, which implies ergodicity.

Problem 4.2. Show that the Lebesgue measure is not weakly mixing for any circle rotation.

Solution. It suffices to prove this for irrational circle rotations f_α, since the Lebesgue measure is not ergodic for rational circle rotations. Set $A := \left\{e^{2\pi x} : x \in [0, \frac{1}{4}]\right\}$ and $B := \left\{e^{2\pi x} : x \in [\frac{1}{2}, \frac{3}{4}]\right\}$. Since the Lebesgue measure is ergodic with respect to $f_{-\alpha}$, Birkhoff's ergodic theorem implies that there exists a $y \in A$ such that $\frac{1}{n}\sum_{i=0}^{n-1} \lambda_B\left(f_{-\alpha}^{i}(y)\right) \to \frac{1}{4}$ as $n \to \infty$. The set $\{i \in N_0 : f_\alpha^{-i}(A) \cap A = \phi\}$ thus has density of at least $\frac{1}{4}$. This means that $\left|\lambda\left(f_\alpha^{-i}(A) \cap A\right) - \lambda(A)^2\right| = \frac{1}{16}$ for all i from a set of density of at least $\frac{1}{4}$. Hence, the Lebesgue measure is not weakly mixing with respect to f_α.

Problem 4.3. Show that any Bernoulli shift is strong mixing.

Solution. We have to show that any Bernoulli shift is strong mixing. To see this, let A and B be arbitrary measurable sets. By standard measure-theoretic arguments, A and B may each be approximated arbitrarily closely by a finite union of cylinder sets. Since if A' and B' are finite unions of cylinder sets, we have that $\mu\left(A' \cap T^{-n}B'\right)$ is equal to $\mu(A')\mu(B')$ for large n, it is easy to deduce that $\mu\left(A \cap T^{-n}B\right) \to \mu(A)\mu(B)$ as required. Since the doubling map is measure-theoretically isomorphic to a one-sided Bernoulli shift, it follows that the doubling map is also strong-mixing.

Problem 4.4. Show that a Markov chain is strong mixing when it is irreducible and aperiodic.

Solution. If a Markov chain is irreducible (that is for any states i and j, there exists an $n \geq 0$ such that $P_{ij}^n > 0$) and aperiodic (there is a state i such that $\gcd\{n : P_{ii}^n > 0\} = 1$), then given any pair of cylinder sets A' and B', we have by standard theorems of Markov chains $\mu(A' \cap T^{-n}B') \to \mu(A')\mu(B')$. The same argument as above then shows that an aperiodic irreducible Markov chain is strong mixing [16].

On the other hand, if a Markov chain is periodic $(d = \gcd\{n : P_{ii}^n > 0\} > 0)$, then letting $A = B = \{x : x_0 = i\}$, we have that $\mu(A \cap T^{-n}B) = 0$. It follows that T^d is not ergodic, so that T is not weak mixing.

We will give intuitive ideas of ergodicity, weak-mixing, and strong-mixing after the next theorem, which gives a useful characterization of weak-mixing.

Theorem 4.5. *If T is a measure-preserving transformation of a probability space (X, β, μ), then the following are equivalent*:

(i) *T is weak-mixing.*

(ii) $\exists J \subset Z^+$ *of density zero such that*

$$\forall A, B \in \beta \lim_{N \notin J \to \infty} \mu(T^N A \cap B) = \mu(A)\mu(B).$$

(iii) $\forall A, B \in \beta$

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \left| \mu(T^{-n}A \cap B) - \mu(A)\mu(B) \right|^2 = 0.$$

Intuitively, given any set A, for T to be strong-mixing means that $T^{-n}A$ becomes asymptotically independent of any other set B. If T is ergodic, then this independence is on the average rather than asymptotically. If T is weak-mixing, then this independence again occurs, provided that a few points in the trajectory are ignored.

The following theorem gives an interesting result:

Theorem 4.6. *If* T *is a measure-preserving transformation and* (X, β, μ) *is a probability space, then the following are equivalent*:

(i) *T is weak-mixing.*

(ii) $T \times T$ *is weak-mixing.*

(iii) $T \times T$ *is ergodic.*

Furthermore, T *is strong-mixing if and only if* $T \times T$ *is strong-mixing.*

The following theorem compares ergodicity and mixing in terms of functions. We will use (f, g) to denote the inner product of two functions f and g with respect to the norm on L^P.

Theorem 4.7. *Suppose* (X, β, μ) *is a probability space and* $T : X \to X$ *is measure-preserving. Then*

(a) *The following are equivalent*:

(i) *T is ergodic.*

(ii) $\forall f, g \in L^2(\mu)$

$$\lim_{N\to\infty}\left(\frac{1}{N}\right)\sum_{n=0}^{N-1}\left(f \circ T^n, g\right) = (f, 1)(1, g).$$

(iii) $\forall f \in L^2(\mu)$

$$\lim_{N\to\infty}\left(\frac{1}{N}\right)\sum_{n=0}^{N-1}\left(f \circ T^n, f\right) = (f, 1)(1, f).$$

(b) *The following are equivalent*:

(i) *T is weak-mixing.*

(ii) $\forall f,\ g \in L^2(\mu)$

$$\lim_{N\to\infty}\left(\frac{1}{N}\right)\sum_{n=0}^{N-1}\left|\left(f \circ T^n,\ g\right) - (f,\ 1)(1,\ g)\right| = 0.$$

(iii) $\forall f \in L^2(\mu)$

$$\lim_{N\to\infty}\left(\frac{1}{N}\right)\sum_{n=0}^{N-1}\left|\left(f \circ T^n,\ f\right) - (f,\ 1)(1,\ f)\right| = 0.$$

(iv) $\forall f \in L^2(\mu)$

$$\lim_{N\to\infty}\left(\frac{1}{N}\right)\sum_{n=0}^{N-1}\left|\left(f \circ T^n,\ f\right) - (f,\ 1)(1,\ f)\right|^2 = 0.$$

(c) *The following are equivalent*:

(i) *T is strong-mixing*.

(ii) $\forall f,\ g \in L^2(\mu) \quad \lim_{N\to\infty}\left(f \circ T^N,\ g\right) = (f,\ 1)(1,\ g).$

(iii) $\forall f \in L^2(\mu) \quad \lim_{N\to\infty}\left(f \circ T^N,\ f\right) = (f,\ 1)(1,\ f).$

There is a nice relation between the weak-mixing of a transformation T and the spectral property of the operation of composition of a function in $L^2(\mu)$ with T.

Our examples have a variety of mixing properties. The identity transformation is strong mixing if and only if it is ergodic. No rotation of a compact group is weak-mixing but for endomorphism of compact groups ergodicity, weak-and strong-mixing are equivalent. Both the one-and two-sided $(p_0,\ \cdots,\ p_{k-1})$-shifts are strong-mixing. Thus, we have no simple example of a weak-mixing transformation that is not strong-mixing.

On the other hand, we shall see that some mixing-type properties are invariant under spectral isomorphism, with others are not. If a property is invariant under spectral isomorphism, we say that it is a spectral property. It turns out that given a non-invertible measure-preserving transformations, there is a natural way to uniquely associate an invertible measure-preserving transformation sharing almost all of the ergodic properties of the original transformation.

5. Spectral Isomorphism and Discrete Spectrum

In the previous section, we saw that any measure-preserving transformations that were conjugate also spectrally isomorphic. If the two spectrally isomorphic measure-preserving transformations have discrete spectrum, then they are conjugate. Thus, the property of discrete spectrum is very important and depends upon the eigenvalues of the measure-preserving transformation. Note that if two measure-preserving transformations are spectrally isomorphic, then they have the same eigenvalues.

Definition 5.1. Let T be measure-preserving transformation of a probability space (X, β, μ). A complex number λ is an eigenvalue of T if $\exists f \neq 0 \in L^2(\mu)$ such that $f \circ T(x) = \lambda f(x)$ almost everywhere. Such an f is an eigenfunction of T corresponding to λ.

Definition 5.2. Let T be measure-preserving transformation of a probability space (X, β, μ). Then T is said to have continuous spectrum if the only eigenvalue of T is 1 and the only eigenfunctions are constants.

Clearly, the only eigenfunctions of an ergodic transformation are constants. Also, we can see that all eigenvalues of a mpt are on the unit circle (i.e., $|\lambda| = 1$). This is because

$$\|f\|^2 = \|f \circ T\|^2 = (f \circ T,\ f \circ T) = (\lambda f,\ \lambda f) = |\lambda|^2 \|f\|^2.$$

Thus $\lambda = 1$ is always an eigenvalue and any nonzero constant function is a corresponding eigenfunction. The connection with weak-mixing is as follows.

Theorem 5.1. *If T is an invertible measure-preserving transformation of a probability space, then T is weak-mixing if and only if T has continuous spectrum.*

Theorem 5.2. *The measure-preserving transformation T is weak-mixing if and only if U_T has non-constant eigenfunctions.*

Of course this also shows that weak-mixing is a spectral property. Equivalently, this says that the transformation T is weak-mixing if and only if the apart from the constant eigenfunction, the operator U_T has only continuous spectrum (that is, the operator has no other eigenfunctions).

Using this theory, we can establish the following:

Theorem 5.3. (i) *T is weak-mixing if and only if $T \times T$ is ergodic*;

(ii) *If T and S are ergodic, then $T \times S$ is ergodic if and only if U_S and U_T have no common eigenvalues other than* 1.

The major internal problem of ergodic theory is to determine when to measure-preserving transformations are isomorphic or conjugate. The primary way of solving this problem to date has been to determine certain properties that are invariant under conjugation or isomorphism and show that if both measure-preserving transformations are conjugate or isomorphic, then they must both have this property. The two properties most commonly used are the eigenvalues of the measure-preserving transformations and the entropy of the measure-preserving transformations. Kolmogorov and Sinai developed the idea of entropy and Ornstein proved that two Bernoulli shifts with the same entropy are isomorphic. Entropy is a very complicated idea and beyond the scope of this paper, so we will conclude with the idea of spectral isomorphism and discrete spectrum.

Definition 5.3. Let (X_1, β_1, μ_1) and (X_2, β_2, μ_2) be two probability spaces with measure-preserving transformations $T_1 : X_1 \to X_1$ and $T_2 : X_2 \to X_2$. Then T_1 and T_2 are spectrally isomorphic if there is a linear operator $W : L^2(\mu_2) \to L^2(\mu_1)$ such that

(i) W is invertible,

(ii) $(Wf, Wg) = (f, g) \quad \forall f, g \in L^2(\mu_2)$,

(iii) $U_{T_1}W = WU_{T_2}$ where U_T is the linear operator $\circ T$ of composition with T.

Spectral isomorphism is weaker than conjugacy, thus if two measure-preserving transformations are isomorphic they are conjugate and two conjugate measure-preserving transformations are spectrally isomorphic.

Definition 5.4. A property P of a measure-preserving transformation is an isomorphism, or conjugacy or spectral, invariant if the following holds: Given T_1 has P and T_2 is isomorphic, or conjugate or spectrally isomorphic, to T_1, then T_2 has property P.

Since isomorphism $\Rightarrow$ conjugacy $\Rightarrow$ spectral isomorphism, a spectral invariant is a conjugacy invariant and a conjugacy invariant is an isomorphism invariant.

Theorem 5.4. *Ergodicity, weak-mixing and strong-mixing are spectral invariants of measure-preserving transformations.*

Definition 5.6. An ergodic measure-preserving transformation T and the probability space (X, β, μ) has discrete spectrum if $\exists$ an orthonormal basis for $L^2(\mu)$ consisting of eigenfunctions of T.

The following theorem is the essence of discrete spectrum and was proven by Halmos and Von Neumann in 1942.

Theorem 5.5 (Discrete spectrum theorem)**.** *Let* T_1 *and* T_2 *be ergodic measure-preserving transformations of the probability spaces* (X_1, β_1, μ_1) *and* (X_2, β_2, μ_2), *respectively. Then the following are equivalent*:

(i) T_1 *and* T_2 *are spectrally isomorphic.*

(ii) T_1 *and* T_2 *have the same eigenvalues.*

(iii) T_1 *and* T_2 *are conjugate.*

Let us now turn our attention to a collection of ergodic mpt's that have discrete spectrum. Denote S_1 as the complex unit circle. Suppose $T : S_1 \to S_1$ defined by $T(z) = az$, where a is not a root of unity. Clearly T is ergodic and is a rotation of a compact group. Consider the collection of functions $f_n : S_1 \to C$ defined by $f_n(z) = z^n$. Then f_n is an eigenfunction of T corresponding to the eigenvalue a^n and $\{f_n\}$ form a basis for $L^2(\mu)$ so T has discrete spectrum. Of more general interest is that these rotations on any compact abelian groups are the canonical class of measure-preserving transformations with discrete spectrum.

6. Conclusion

Mordern ergodic theory was started by Andrei Kolmogorov with the formal development of Boltzmann's notion of entropy, and developed in the 60's and 70's to include many differentiable actions. Applications include fluid dynamics, coding theory, number theory, complex dynamics, and cellular automata.

Entropy provides a strictly isomorphic invariant. However, the notion of entropy is also one of the most difficult concepts in ergodic theory. There are also difficult types of entropy and all are used in ergodic theory, many building off of others. We simply refer the interested reader to [4] for a more rigorous background in ergodic theory and a decent presentation of entropy and other advanced topics.

Another area of ergodic theory of great importance is the external problems. Fields such as topological dynamics and differentiable dynamics make extensive use of ergodic theory.

In this paper, it is proved that mixing implies weak mixing and that weak mixing implies ergodicity and also proved that Bernoulli shift is strong mixing and Markov chain is strong mixing when it is irreducible and aperiodic.

In this article, we have seen that any measure-preserving transformations that were conjugate also spectrally isomorphic. If the two spectrally isomorphic measure-preserving transformations have discrete spectrum, then they are conjugate. So in this paper, we have seen that if two measure-preserving transformations are spectrally isomorphic, then they have the same eigenvalues.

References

[1] P. R. Halmos, Lectures on Ergodic Theory, First Edition, Chelsea Publishing Company, 1956.

[2] Mark Pollicott and Michiko Yuri, Dynamical Systems and Ergodic Theory, Cambridge University Press, 1998, 179 pages.

[3] H. L. Royden, Real Analysis, Second Edition, MacMillan Company, 1968.

[4] Peter Walters, An Introduction to Ergodic Theory, First Edition, Springer-Verlag, 2000.

[5] L. Boltzmann, Einige allgemeine Satze uber Warmegleichgewicht, Wiener Berichte 63 (1871), 679-711.

[6] D. J. Rudolph, Fundamentals of Measurable Dynamics, Oxford, 1990.

[7] G. D. Birkhoff, Proof of the ergodic theorem, Proc. Nat. Acad. Sci. 17 (1931), 656-660.

[8] J. Von Neumann, Proof of the quasi-ergodic hypothesis, Proc. Nat. Acad. Sci. USA, 18 (1932), 70-82.

[9] Jason Preszler, Ergodic Theory, University of Puget, Sound Math, 2003.

[10] L.-S. Young, Ergodic theory of differentiable dynamical systems, In Real and Complex Dynamical Systems, (Hillerod, 1993), Kluwer (1995), 293-336.

[11] Parry William, Topics in Ergodic Theory, Cambridge University Press, Cambridge, 1981, 124 pages.

[12] H. R. Biswas and M. S. Islam, Ergodic theory of one-dimensional map, Bangladesh J. Sci. Ind. Res. 47(3) (2012), 321-326.

[13] Karl Petersen, Ergodic Theory, First Edition, Cambridge University Press, 1989.

[14] M. V. Jakobson, Ergodic theory of one-dimensional mappings, dynamical systems, ergodic theory and applications, Encyclopedia of Math., Springer, Chapter 9, 100(II) (2000), 234-263.

[15] R. V. Chacon, Weakly mixing transformations which are not strongly mixing, Proc. Amer. Math. Soc. 22 (1969), 559-562.

[16] V. Bergelson, Weakly mixing PET, Ergodic Theory and Dynamical Systems 7 (1987), 337-349.

■