May 1, 2003

ERGODIC THEORY

JASON PRESZLER

ABSTRACT. During the fall semester of 2002 I completed an independent study course under the direction of Prof. David Scott on ergodic theory. 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

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 open problems of fundamental importance. Due to the recent development of the subject and the requisite background, ergodic theory is rarely seen in any form at the undergraduate level. I first became interested in ergodic theory during the 2002 AMS short course on symbolic dynamics which utilizes results from ergodic theory in the study of shifts of finite type.

After this short course I was left with the question of "what is ergodic theory?" This paper answers this question and is intended for undergraduates and non-specialists looking to broaden their horizons.

2. Background

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 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.

Consider the unit circle $S_1 = \{z \in \mathbb{C} : |z| = 1\}$ in the complex plane and fix a point z_0 in S_1. A simple group action rotates S_1 by a fixed angle, i.e. $\phi_\alpha : S_1 \to S_1$ such that for $z = e^{i\theta}$ then $\phi_\alpha(z) = e^{i(\theta+\alpha)}$. Clearly this automorphism of S_1 preserves Lebesgue measure and is therefore referred to as a *measure preserving transformation*. Measure preserving transformations, or mpt's, are the primary objects in ergodic theory. If α is a root of unity than ϕ_α generates a finite cyclic group and the orbit of z_0 under ϕ_α is periodic and trivial to analyze. However if α is not a root of unity then the orbit of z_0 is aperiodic and ϕ_α provides our simplest example of an *ergodic* transformation. The idea of ergodicity will be formalized shortly.

The above example shows a simple kind of dynamical system that is also a group action on a finite measure space. The rotation preserves the measure, when working with another type of space (such as a topological space) more relevant kinds of transformations appear (such as homeomorphisms). For the remainder of this paper we will focus on measure-theoretic ergodic theory, thus neglecting other kinds of spaces.

2.1. **Measure Theory.** Let A be a set and $\mathfrak{B}$ a collection of subsets of A. The pair $(A, \mathfrak{B})$ is referred to as a measurable space. If $\{B_n\}$ is a sequence of pairwise disjoint members of $\mathfrak{B}$, then a finite measure is a function $m : \mathfrak{B} \to \mathbb{R}^+$ such that $m(\emptyset) = 0$ and $m(\cup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} m(B_n)$. The triple $(A, \mathfrak{B}, m)$ is a finite measure space. Often it is useful to normalize the measure so that $m(X) = 1$, in which case we will call $(A, \mathfrak{B}, m)$ a probability space.

One should note that most measures are not finite. A Common example is the Lebesgue measure on open subsets of the real line; $A = \mathbb{R}$, $\mathfrak{B}$ is the collection of open subsets of A of the form (a, b), and $m((a, b)) = b - a$. However, restricting A to the interval $[0, 1]$ in the above example produces a useful finite measure, and in fact a probability measure. Another common example is Lebesgue/Haar measure of the complex unit circle S_1. Here the collection of subsets of S_1 is the collection of all intervals $(e^{i\theta_0}, e^{i\theta_1})$ and the measure of an interval is the angle between the endpoints. As noted earlier this measure is preserved by rotations of S_1.

Since the idea behind integration is to measure something, measures are intimately related to integrals. Usually the Riemann Integral measures the area under a curve or the volume under a surface, similarly the Lebesgue integral measures the size of an interval. For the purposes of this paper we will not formalize these ideas or the relationship between integration and measures and will instead introduce useful measures and ideas from measure theory as needed.

One other useful measure is Haar measure. Haar measure is not a specific measure but rather a probability measure that is invariant under group actions. 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 1. *Let G be a topological group. There exists a probability measure m defined on the σ-algebra $\mathfrak{B}(G)$ of Borel subsets of G such that $m(xE) = m(E)\ \forall\ x \in G\ \forall\ E \in \mathfrak{B}$ and m is regular. There is only one regular rotation invariant probability measure on $(G, \mathfrak{B}(G))$.*

If our space is a metric space then any probability measure is regular, since this is the usual case we will forego the definition of a regular measure. A σ-algebra is a collection, $\mathfrak{B}$, of subsets of X, satisfying three properties. First, $X \in \mathfrak{B}$. Second, if $B \in \mathfrak{B}$ then $X \setminus B \in \mathfrak{B}$. Third, if $B_n \in \mathfrak{B}$ for $n \geq 1$ then $\cup_{n=1}^{\infty} B_n \in \mathfrak{B}$.

2.2. **Function Spaces.** Ergodic theory frequently uses Banach spaces of functions defined on a measure space. Let $(A, \mathfrak{A}, m)$ be a finite measure space and for some $p \in \mathbb{R}$, consider the set of functions $f : A \to \mathbb{C}$ 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 $L^p(A, \mathfrak{A}, m)$ 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 $||f||_p = (\int |f|^p dm)^{\frac{1}{p}}$, which makes L^p a Banach space.

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

All the functions in L^p will be measurable.

2.3. **Topology.** As mentioned earlier, ergodic theory is the study of group actions on spaces. Measure spaces are one type of space that provides nice structure to work in, another is a topological space. Ergodic theory over topological spaces has many important applications to topological dynamics and topological entropy, but topological methods are also useful in understanding measure-theoretic ergodic theory (and vice-versa).

3. Measure Preserving Transformations

As the name implies, measure preserving transformations (or mpt's) are functions on a measure space that preserve the given measure. Thus, if $(A, \mathfrak{A}, m)$ is a measure space and $f : (A, \mathfrak{A}, m) \to (A, \mathfrak{A}, m)$ such that $\forall\ B \in \mathfrak{A}\ m(B) = m(f(B))$, then f is an mpt. We will frequently write $f : A \to A$ when the measure and the σ-algebra are understood.

Also of importance are mpt's between different measure spaces since mpt's are the morphisms (structure preserving maps) of measure spaces.

Definition 2. Suppose $(X_1, \mathfrak{B}_1, m_1)$ and $(X_2, \mathfrak{B}_2, m_2)$ are two probability spaces.

(1) A transformation $T : X_1 \to X_2$ is measurable if $T^{-1}(\mathfrak{B}_2) \subseteq \mathfrak{B}_1$ (i.e. T is surjective).
(2) A transformation $T : X_1 \to X_2$ is measure-preserving if T is measurable and $m_1(T^{-1}(B_2)) = m_2(B_2)\ \forall\ B_2 \in \mathfrak{B}_2$.
(3) A transformation $T : X_1 \to X_2$ is an invertible measure-preserving transformation if T is measure-preserving, bijective, and T^{-1} is also measure-preserving.

Exercise 1 Verify that if $T_1 : X_1 \to X_2$ and $T_2 : X_2 \to X_3$ are mpt's then $T_2 \circ T_1 : X_1 \to X_3$ is also a mpt.

Note that T is a mpt between two measure spaces, then T is also a mpt between the completions of those spaces.

In ergodic theory, we are interested in long term behavior, so we will focus on mpt's from a measure space onto itself, $T : X_1 \to X_1$. Common examples of such mpt's are the identity 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, and Bernoulli and Markov Shifts.

[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 $f \sim g$ iff $m(B) = 0$.

Let us provide more detail about Bernoulli shifts as these are more complicated, but some of the more interesting mpt's. Let $k \geq 2$ and define a probability vector $(p_0, \dots, p_{k-1})$ 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 $X = \{0, \dots, k-1\}$ and $\mu(i) = p_i$. To make this space more complicated, let $(B_k^2, \mathfrak{B}, m) = \Pi_{-\infty}^{\infty}(X, 2^X, \mu)$. An 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 $(B_k^2, \mathfrak{B}, m)$ consists 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 $(\frac{1}{2}, \frac{1}{2})$ is the probability vector, or the two-sided $(p_0, \dots, p_{k-1})$-shift in general.

The above shift can also be simplified into a one-sided shift by considering the measure space $(B_k^1, \mathfrak{B}, m) = \Pi_0^{\infty}(X, 2^X, \mu)$. 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, the interested reader is referred to [4].

In measure-theoretic ergodic theory, two standard types of problems now rear their heads. First, we would like to determine when two mpt's are isomorphic and other associated problems. The second type of problem is more external, how can we use results about mpt's to solve problems in other areas of mathematics or even outside of mathematics? The remainder of this paper will focus on the first type of problems, or the so called isomorphism problem.

4. Recurrence, Ergodicity and Mixing

Ergodicity and mixing are properties enjoyed by certain mpt's, but recurrence is a common property of a mpt as shown by Poincaré. Additionally, we will only be concerned with probability or finite measure spaces for the duration of the paper. Most results will not be valid for infinite measure spaces. Interested readers should determine why the results will not hold in an infinite measure space.

Theorem 2. *Let $T : X \to X$ be a measure-preserving transformation of a probability space $(X, \mathfrak{B}, m)$. Let $E \in \mathfrak{B}$ with $m(E) > 0$. 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 mpt as the trajectory, or orbit, of that element. Thus, for any $x \in E$ $T^n(x)$ 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 $m(F) = m(E)$.

Definition 3. Let $(X, \mathfrak{B}, m)$ be a measure space. A measure-preserving transformation T of the measure space is ergodic if the only members $B \in \mathfrak{B}$ such that $T^1 B = B$ iff $m(B) = 0$ or $m(X \setminus B) = 0$.

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 on 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 3. *If $(X, \mathfrak{B}, m)$ is a probability space and $T : X \to X$ is measure-preserving then the following are equivalent:*

1. *T is ergodic.*
2. *Whenever f is measurable and $(f \circ T)(x) = f(x)\ \forall\ x \in X$ then f is constant almost everywhere.*
3. *Whenever f is measurable and $(f \circ T)(x) = f(x)$ almost everywhere, then f is constant almost everywhere.*
4. *Whenever $f \in L^2(m)$ and $(f \circ T)(x) = f(x)\ \forall\ x \in X$ then f is constant almost everywhere.*
5. *Whenever $f \in L^2(m)$ and $(f \circ T)(x) = f(x)$ almost everywhere, then f is constant almost everywhere.*
6. *$\forall\ A, B \in \mathfrak{B}$ with $m(A) > 0$ and $m(B) > 0$ there exists $n > 0$ such that $m(T^{-n}A \cap B) > 0$.*

It is possible to relax the third statement above to be f measurable and $f(Tx) \geq f(x)$ almost everywhere then f is constant almost everywhere. The intuitive meaning of ergodicity is that if $T : X \to X$ is a continuous ergodic mpt then almost every point of X has a dense trajectory under T. Indeed, nothing is sacrificed in this intuitive understanding as it could easily be phrased as a theorem and is also the standard method of proving that a transformation is ergodic.

Now one must naturally wonder which of the above examples of mpt's are in fact ergodic. The identity transformation is ergodic iff all members of $\mathfrak{b}$ have measure 0 or 1. We have already stated that a rotation of the unit circle is ergodic iff the rotation is not by a root of unity. This generalizes to the following theorem.

Theorem 4. *Let G be a compact group and $T(x) = ax$ be a rotation of G. Then T is ergodic iff $\{a^n\}_{-\infty}^{\infty}$ is dense in G. Furthermore, if T is ergodic then G is abelian.*

Additionally the one- and two-sided $(p_0, \ldots, p_{k-1})$ shifts are ergodic. For our other examples there are known necessary and sufficient conditions for a mpt to be ergodic, but some become very detailed and will not be needed during this paper.

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

Theorem 5. *(Birkhoff Ergodic Theorem) Suppose $T : X \to X$ is measure-preserving and $f \in L^1(m)$. Then $\frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x))$ converges almost everywhere to a function $f^* \in L^1(m)$. Also $f^* \circ T = f^*$ almost everywhere and if $m(X) < \infty$, then $\int f^* dm = \int f dm$.*

Clearly if T is ergodic then f^* is constant almost everywhere and if $m(X) < \infty$ then $f^* = (\frac{1}{m(X)}) \int f dm$. Furthermore if $(X, \mathfrak{B}, m)$ is a probability space and T is ergodic then $\forall\ f \in L^1(m)$

$$\lim_{n\to\infty} \left(\frac{1}{n}\right) \sum_{i=0}^{n-1} f(T^i(x)) = \int f dm \quad \text{a.e.} \tag{1}$$

The Birkhoff ergodic theorem has many uses, particularly in statistical mechanics, but also to number theory and dynamical systems. These applications can be found in the literature. The following corollary is due to Von Neumann.

Corollary 1 (L^p Ergodic Theorem of Von Neumann)**.** *Let $1 \leq p < \infty$ and let T be a measure-preserving transformation of the probability space $(X, \mathfrak{B}, m)$. If $f \in L^p(m)$ $\exists\ f^* \in L^p(m)$ with $f^* \circ T = f^*$ a.e. and $||(\frac{1}{n}) \sum_{i=0}^{n-1} f(T^i(x)) - f^*(x)||_p \to 0$.*

Interestingly enough the theorem of Von Neumann was published a year before Birkhoff's result. The next corollary provides yet another criteria for ergodicity.

Corollary 2. *Let $(X, \mathfrak{B}, m)$ be a probability space and let $T : X \to X$ be a measure-preserving transformation. Then T is ergodic iff $\forall\ A, B \in \mathfrak{B}$*

$$\frac{1}{n} \sum_{i=0}^{n-1} m(T^{-i} A \cap B) \to m(A)m(B). \tag{2}$$

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

Definition 4. Let T be a measure-preserving transformation of a probability space $(X, \mathfrak{B}, m)$.

(1) T is weak-mixing if $\forall\ A, B \in \mathfrak{B}$

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

(2) T is strong mixing if $\forall\ A, B \in \mathfrak{B}$

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

One should be be able to see the 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 one 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, one can topologize the space of invertible mpt's of a measure space with the weak topology and then show that the class of weak-mixing transformations is of second category while the class of strong-mixing transformations is of first category; thus most transformations are weak-mixing, but not strong-mixing. We shall give intuitive ideas of ergodicity, weak-, and strong-mixing after the next theorem, which gives a useful characterization of weak-mixing.

Theorem 6. *If T is a measure-preserving transformation of a probability space $(X, \mathfrak{B}, m)$ then the following are equivalent:*

(1) *T is weak-mixing.*
(2) *$\exists\ J \subset \mathbb{Z}^+$ of density zero such that $\forall\ A, B \in \mathfrak{B}$*

$$\lim_{n \notin J \to \infty} m(T^n A \cap B) = m(A)m(B).$$

(3) *$\forall\ A, B \in \mathfrak{B}$*

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

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 and somewhat unusual result.

Theorem 7. *If T is a measure-preserving transformation and $(X, \mathfrak{B}, m)$ is a probability space then the following are equivalent:*

(1) *T is weak-mixing.*
(2) *$T \times T$ is weak-mixing.*
(3) *$T \times T$ is ergodic.*

Furthermore, T is strong-mixing iff $T \times T$ 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 8. *Suppose $(X, \mathfrak{B}, m)$ is a probability space and $T : X \to X$ is measure-preserving. Then:*

(1) *The following are equivalent:*
 (a) *T is ergodic.*
 (b) $\forall\ f, g \in L^2(m)\ \lim_{n\to\infty}(\frac{1}{n})\sum_{i=0}^{n-1}(f \circ T^i, g) = (f, 1)(1, g)$.
 (c) $\forall\ f \in L^2(m)\ \lim_{n\to\infty}(\frac{1}{n})\sum_{i=0}^{n-1}(f \circ T^i, f) = (f, 1)(1, f)$.
(2) *The following are equivalent:*
 (a) *T is weak-mixing.*
 (b) $\forall\ f, g \in L^2(m)\ \lim_{n\to\infty}(\frac{1}{n})\sum_{i=0}^{n-1}|(f \circ T^i, g) - (f, 1)(1, g)| = 0$.
 (c) $\forall\ f \in L^2(m)\ \lim_{n\to\infty}(\frac{1}{n})\sum_{i=0}^{n-1}|(f \circ T^i, f) - (f, 1)(1, f)| = 0$.
 (d) $\forall\ f \in L^2(m)\ \lim_{n\to\infty}(\frac{1}{n})\sum_{i=0}^{n-1}|(f \circ T^i, f) - (f, 1)(1, f)|^2 = 0$.
(3) *The following are equivalent:*

(a) *T is strong-mixing.*
(b) $\forall\ f, g \in L^2(m)\ \lim_{n\to\infty}(f \circ T^n, g) = (f, 1)(1, g)$.
(c) $\forall\ f \in L^2(m)\ \lim_{n\to\infty}(f \circ T^n, f) = (f, 1)(1, f)$.

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 $L^2(m)$ with T.

Definition 5. Let T be a measure-preserving transformation of the probability space $(X, \mathfrak{B}, m)$. A complex number λ is an eigenvalue of T if $\exists\ f \neq 0 \in L^2(m)$ such that $f \circ T(x) = \lambda f(x)$ a.e. Such an f is an eigenfunction of T corresponding to λ.

Definition 6. Let T be a measure-preserving transformation of the probability space $(X, \mathfrak{B}, m)$. 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 one should 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 non-zero constant function is a corresponding eigenfunction. The connection with weak-mixing is as follows.

Theorem 9. *If T is an invertible measure-preserving transformation of a probability space then T is weak-mixing iff T has continuous spectrum.*

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

5. Isomorphism, Conjugacy, and Spectral Isomorphism

Clearly two measure spaces can be isomorphic, but of more interest to us is when two mpt's are isomorphic. When two mpt's are isomorphic there are certain invariant properties, ergodic theory is very interested in these 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 7. Let $(X_1, \mathfrak{B}_1, m_1)$ and $(X_2, \mathfrak{B}_2, m_2)$ be two probability spaces. These spaces are isomorphic if $\exists\ M \in \mathfrak{B}_1$ and $N \in \mathfrak{B}_2$ such that $m_1(M) = 1 = m_2(N)$ 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 mpt's.

Definition 8. Let $(X_1, \mathfrak{B}_1, m_1)$ and $(X_2, \mathfrak{B}_2, m_2)$ be probability spaces with measure algebras $(\tilde{\mathfrak{B}}_1, \tilde{m}_1)$ and $(\tilde{\mathfrak{B}}_2, \tilde{m}_2)$. These measure algebras are isomorphic if $\exists$ a bijective measure-preserving transformation $\Phi : \tilde{\mathfrak{B}}_2 \to \tilde{\mathfrak{B}}_1$ that preserves complements, countable unions and intersections. Two probability spaces are conjugate if their measure algebras are isomorphic.

Definition 9. Let $(X_1, \mathfrak{B}_1, m_1)$ and $(X_2, \mathfrak{B}_2, m_2)$ be 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 isomorphic if $\exists\ M \in \mathfrak{B}_1$ and $N \in \mathfrak{B}_2$ with $m_1(M) = 1 = m_2(N)$ such that

(1) $T_1 M \subseteq M$, $T_2 N \subseteq N$ and
(2) $\exists$ an invertible measure-preserving transformation $\phi : M \to N$ such that $\phi(T_1(x)) = T_2(\phi(x))\ \forall\ x \in M$.

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

Definition 10. Let $(X_1, \mathfrak{B}_1, m_1)$ and $(X_2, \mathfrak{B}_2, m_2)$ be 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 conjugate if there is a measure algebra isomorphism $\Phi : \tilde{\mathfrak{B}}_2 \to \tilde{\mathfrak{B}}_1$ such that $\Phi \tilde{T}_2^{-1} = \tilde{T}_1^{-1} \Phi$.

Conjugacy is also an equivalence relation and all isomorphic mpt's are conjugate. In some cases, conjugacy can also imply isomorphism.

The major internal problem of ergodic theory is to determine when to mpt's are isomorphic or conjugate. The primary way of solving this problem to date has been to determine certain properties that are in variant under conjugation or isomorphism and show that if both mpt's are conjugate or isomorphic, then they must both have this property. The two properties most commonly used are the eigenvalues of the mpt's and the entropy of the mpt's. 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 11. Let $(X_1, \mathfrak{B}_1, m_1)$ and $(X_2, \mathfrak{B}_2, m_2)$ be 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 $W : L^2(m_2) \to L^2(m_1)$ such that

(1) W is invertible
(2) $(Wf, Wg) = (f, g)\ \forall\ f, g \in L^2(m_2)$
(3) $U_{T_1} W = W U_{T_2}$ where U_T is the linear operator $\circ T$ of composition with T.

Spectral isomorphism is weaker then conjugacy, thus if two mpt's are isomorphic they are conjugate and two conjugate mpt's are spectrally isomorphic. As with conjugacy there are times when spectral isomorphism implies conjugacy, these will be the content of the next section.

Definition 12. 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 $\implies$ conjugacy $\implies$ spectral isomorphism, a spectral invariant is a conjugacy invariant and a conjugacy invariant is an isomorphism invariant.

Theorem 10. *Ergodicity, weak-mixing, and strong-mixing are spectral invariants of measure-preserving transformations.*

6. Discrete Spectrum

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

Definition 13. An ergodic measure-preserving transformation T and the probability space $(X, \mathfrak{B}, m)$ has discrete spectrum if $\exists$ an orthonormal basis for $L^2(m)$ 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 11. *(Discrete Spectrum Theorem) Let T_1 and T_2 be ergodic measure-preserving transformations of the probability spaces $(X_1, \mathfrak{B}_1, m_1)$ and $(X_2, \mathfrak{B}_2, m_2)$ respectively. Then the following are equivalent:*

(1) *T_1 and T_2 are spectrally isomorphic.*
(2) *T_1 and T_2 have the same eigenvalues.*
(3) *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 $f_n : S_1 \to \mathbb{C}$ 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 $L^2(m)$ so T has discrete spectrum. A reader familiar with character theory should easily be able to generalize these ideas to rotations on any compact abelian group. Of more general interest is that these rotations of compact abelian groups are the canonical class of mpt's with discrete spectrum.

Theorem 12. *(Representation Theorem) An ergodic measure-preserving transformation T with discrete spectrum on a probability space $(X, \mathfrak{B}, m)$ is conjugate to an ergodic rotation on some compact abelian group. The group will be metrisable iff $(X, \mathfrak{B}, m)$ has a countable basis.*

Theorem 13. *(Existence Theorem) Every subgroup G of S_1 is the group of eigenvalues of an ergodic measure-preserving transformation with discrete spectrum.*

In the proof of the representation theorem, the desired transformation in the existence theorem is constructed. See [4] for the proof. These last two theorems completely solve the conjugacy problem for ergodic rotation with discrete spectrum.

7. Conclusion

This is by no means a comprehensive glimpse of ergodic theory. The most important topic has only been briefly mentioned, namely the development of entropy by Kolmogorov and Sinai in the late 1950's. Entropy provides a strictly isomorphic invariant. However, the notion of entropy is also one of the most difficult concepts in ergodic theory, but can be intuitively described as a measure of the uncertainty removed by performing an experiment with known outcomes. There are also different 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 are the external problems. Fields such as topological dynamics and differentiable dynamics make extensive use of ergodic theory. In general the ergodic theorems of Birkhoff and Von Neumann are used in all aspects of dynamical systems and many problems in mathematical physics.

References

1. P.R. Halmos, *Lectures on ergodic theory*, first ed., Chelsea Publishing Company, 1956.
2. Karl Petersen, *Ergodic theory*, first ed., Cambridge University Press, 1989.
3. H.L. Royden, *Real analysis*, second ed., MacMillan Company, 1968.
4. Peter Walters, *An introduction to ergodic theory*, first ed., Springer-Verlag, 2000.

E-mail address: jpreszler@member.ams.org