

Introductory Ergodic Theory and the Birkhoff Ergodic Theorem

James Pinkerton

January 14, 2014

In this exposition we'll cover an introduction to ergodic theory. Specifically, the Birkhoff Mean Theorem. Ergodic theory is generally described as the study of dynamical systems that have an invariant measure. This reaches has relations to flows (such as the Poincaré's Recurrence Theorem, etc.).

Let's begin by discussing what it means for a dynamical system to be ergodic.

We'll start by defining a probability space.

Definition 0.1. *A probability space is a triple (X, \mathcal{F}, μ) , where:*

- X is a set of possible outcomes
- \mathcal{F} is a set of events, where each event is a set containing outcomes
- μ is an assignment of probabilities to each event in \mathcal{F} .

Definition 0.2. *Ergodicity*

Let $T : X \rightarrow X$ be a measure-preserving transformation.

T is ergodic with respect to μ if $\forall F \in \mathcal{F}$ s.t. $T^{-1}(F) = F$ either $\mu(F) = 0$ or $\mu(F) = 1$.

Here are some examples:

Example 0.3. *If X is finite and has uniform measure, $T : X \rightarrow X$ is ergodic IFF it's a cycle.*

Example 0.4. *If T is ergodic, then so is T^{-1} . T^n is not necessarily ergodic for all n , as we can see from the previous example (this property total ergodicity).*

Intuitively, we can think of T as a transformation w.r.t. time. Then, an ergodic function describes a random process for which the time average of one sequence equals the total average.

Now there is a natural analogy to measurable flows; I discussed flows in my presentation earlier in the course.

Let T^t be a measurable flow on a probability space (X, \mathcal{F}, μ) . An event $F \in \mathcal{F}$ is invariant mod 0 under T^t if $\forall t \in \mathbb{R}, X(T^t(F) \Delta F) = 0$ (Δ means symmetric difference).

Now that we understand the definition of a probability space and of Ergodicity, we can begin proving the Birkhoff Theorem. We will first prove a lemma.

The best summary of the probability of an event is the frequency of its occurrence over some large span of time. This is captured in ergodic theory by the measure preserving map, T , from the probability space to itself. T is the change from one outcome of a random series of events to the next.

So suppose we have some measurable function f . Then we define the average of measures over time, the Cesáro average, which is defined as follows:

Definition 0.5. *Cesáro Average*

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

Definition 0.6. *σ -finite measure space*

A measure μ defined on a σ -algebra Σ of subsets of a set X is called finite if $\mu(X) \in \mathbb{R}$. It is σ -finite if X can be written as the countable union of measurable sets with finite measure. A set in a measure space is of σ -finite measure if it can be written as the countable union of sets with finite measure.

Lemma 0.7. *Maximal Ergodic Theorem*

Let T be a measure-preserving transformation on the σ -finite measure space (X, \mathcal{F}, μ) . Suppose $f \in L^1(\mu)$.

Set $E_n = \{x | A_j(f, x) > 0 \text{ for some } j \leq n\}$.

Then

$$\int_{E_n} f d\mu \geq 0.$$

Proof. The trick here is to find a non-decreasing series of functions and then redefine our integral in terms of differences in the series (so that we are integrating over an entirely positive space).

$$\text{Let } F_n(x) = \max\left(0, \sum_{i=0}^{j-1} f(T^i(x)) : j \leq n\right).$$

This will be our non-decreasing series.

This gives us $F_{n+1} = \max(0, f + F_n \circ T)$. But note that the second term is always at least 0 over E_{n+1} , meaning we have

$$F_{n+1} = f + F_n \circ T \implies f = F_{n+1} - F_n \circ T$$

over E_{n+1} .

Now take the integral:

$$\int_{E_{n+1}} f d\mu = \int_{E_{n+1}} (F_{n+1} - F_n \circ T) d\mu.$$

But this is convenient since $F_{n+1} = 0$ everywhere except E_{n+1} and $-F_n \circ T \leq 0$ everywhere, giving $F_{n+1} - F_n \circ T \leq 0$ everywhere but E_{n+1} .

So we can change the boundaries on the integral to prove our theorem!

$$\begin{aligned} \int_{E_{n+1}} f d\mu &= \int_{E_{n+1}} (F_{n+1} - F_n \circ T) d\mu \geq \int_X (F_{n+1} - F_n \circ T) d\mu \\ &= \int_X (F_{n+1} - F_n) d\mu \geq 0. \end{aligned}$$

□

So we've just shown that if the Cesáro average of an L^1 function is positive for a small enough time frame over a subset, then so is its integral.

Corollary 0.8. Set $E_\infty = \cup_{n=1}^\infty E_n$ and

$$\int_{E_\infty} f d\mu \geq 0.$$

Theorem 0.9. *Birkhoff Ergodic Theorem*

Let G be a measure-preserving transformation on a σ -finite measure space (X, \mathcal{F}, μ) , and $f \in L^1(\mu)$.

There exists an \bar{f} with

$$A_n(f, x) \rightarrow \bar{f}(x)$$

almost everywhere.

Proof. This is the deep theorem which we've been aiming for.

First let's consider some rational u, v . And let's set $E_{u,v} = \{x | \overline{\lim} A_n(f, x) > v > u > \underline{\lim} A_n(f, x) > 0\}$. Now we can say: if $x \notin E_{u,v}$ for any u, v then $\lim_{n \rightarrow \infty} A_n(f, x)$ exists.

So assuming $v > 0$, otherwise replace f by $-f$ and $-u > 0$. Assume $\mu(E_{u,v}) > 0$. From its definition, $T(x) \in E_{u,v}$, then x is also. In other words, $T^{-1}(E_{u,v}) \subseteq E_{u,v}$, so we may, w.l.o.g. assume $E_{u,v}$ is the entire measure space.

So $\forall x \in X = E_{u,v}, \exists n$ s.t.

$$\frac{1}{n} \sum_{i=0}^{n-1} (f(T^i(x)) - v) > 0$$

So, by the maximal ergodic lemma(!), we have

$$\int_X (f - v \chi_A) d\mu \geq 0.$$

Meaning,

$$\int_X f d\mu \geq v \mu(A)$$

because $X = E_\infty$ from the first Corollary.

□

Now there are several interesting corollaries that result from Birkhoff's Theorem.

Corollary 0.10. *Defining the map $L(f) = \bar{f}$, for $f \in L^1(\mu), \|L(f)\|_1 \leq \|f\|_1$ and so $L(f)$ is a continuous projection from $L^1(\mu)$ onto the subspace of T -invariant L^1 -functions.*

Proof. As $\|A_n(g)\|_1 = \|g\|_1$ for $g \geq 0$, and since $A_n(f)$ converges pointwise to f ,

$$\int |L(f)| d\mu \leq \int L(|f|) d\mu \leq \underline{\lim} \int A_n(|f|) d\mu = \int |f| d\mu.$$

and

$$\|L(F)\|_1 \leq \|f\|_1.$$

So we have:

$$\begin{aligned}
L(f)(T(x)) &= \lim A_n(f, T(x)) \\
&= \lim \frac{1}{n} \sum_{i=1}^n f(T^i(x)) \\
&= \lim \left(\frac{n+1}{n} A_{n+1}(f, x) - \frac{f(x)}{n} \right) \\
&= L(f)(x).
\end{aligned}$$

So $L(f)$ is T -invariant. As $\|L(f - g)\|_1 = \|L(f) - L(g)\|_1 \leq \|f - g\|_1$, L is a continuous projection onto the T -invariant L^1 functions. \square

Corollary 0.11. *If (X, \mathcal{F}, μ) has no T -invariant subsets of finite measure, then $L(f) \equiv 0, \forall f \in L^1(\mu)$.*

Proof. If X has no T -invariant sets of finite measure, the only T -invariant L^1 function is identically equal to 0. \square

Corollary 0.12. *If $\mu(X) < \infty$, then $\|A_n(f) - L(f)\|_1 \rightarrow 0$.*

Proof. Define $A = \{f \in L^1(\mu) : \|A_n(f) - L(f)\|_1 \rightarrow 0\}$.

As the operator L is a contraction in L^1 , \mathcal{A} is L^1 -closed (if f_i is Cauchy so is $L(f_i)$).

$f < \infty \implies$ all $A_n(f)$ have the same bound.

By the dominated convergence theorem, $A_n(f) \rightarrow L(f)$ in L^1 . $L^1(\mu)$ is the only closed subspace of $L^1(\mu)$ that contains all bounded functions. \square

Last, we should understand why the Birkhoff Theorem is important in Ergodic Theory.

This lies in the relationship to ergodic maps. In fact we can use the theorem to directly characterize ergodic maps!

Corollary 0.13. *Application of Birkhoff's Theorem*

If $\{A_i\}$ is a countable collection of sets L^1 dense in the collection of all sets and

$$\frac{1}{n} \sum_{i=1}^{n-1} \chi_{A_j}(T^i(x)) \rightarrow \mu(A_j)$$

for all i and almost every x then T is ergodic.

Proof. Pick A measure-invariant and A_j arbitrarily from our collection.

$$\begin{aligned}\mu(A)\mu(A_j^c) &= \int_A \mu(A_i^c) d\mu \\ &= \lim_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{i=0}^{n-1} \int_A \chi_{A_j^c}(T^i(x)) d\mu \right) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^n \int \chi_{T^{-i}(A \cap A_j^c)}(x) d\mu \\ &= \mu(A \cap A_j^c)\end{aligned}$$

Selecting A_j so $\mu(A \Delta A_j) \rightarrow 0$, $\mu(A)\mu(A^c) = 0$. So $\mu(A) \in \{0, 1\}$.

□

References:

An Introduction to Ergodic Theory by Walters

Fundamentals of Measurable Dynamics by Rudolph