Ergodic Theory:
A Brief Introduction

Shereen Elaidi

SUMM 2020

12 January 2020

Where did ergodic theory come from?

Ergodic theory was developed in response to a statistical mechanics problem in 1880.

1. The **phase space** X is the region that we are interested in studying. X will be a compact metric space.
2. Given a region $A \subseteq X$ and a gas particle $x(t)$ beginning at $x_0 \in X$, how often is $x(t) \in A$?
3. Boltzmann's famous ergodic hypothesis:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \chi_A(T^n(x)) = \mathbb{P}(A) \tag{1}$$

Physics detour 1: what are ensembles?

Consider a phase space X with N particles, all of which have identical dynamics. Say we are interested in the **thermodynamic limit**. We can study this by fixing a time t and considering infinitely many copies of our dynamical system.

Each particle will be represented by a vector in $\mathbb{R}^{6n}$:

$$\mathbf{x} = (x, y, z, p_x, p_y, p_z)$$

The collection of the infinite copies of a dynamical system is called an **ensemble**. There are various types of ensembles that we can study (will discuss this later).

What is ergodic theory?

Ergodic theory is the study of dynamical systems X equipped with an *invariant measure* μ.

1. Intuitively, this means...

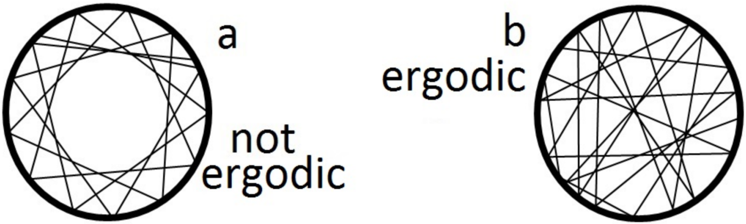

Figure: Ergodic vs. non ergodic system. Credits: Wikipedia

Talk Outline

Since ergodic theory arose out of a physics problem, we will try whenever possible to provide a physical interpretation of the mathematics.

1. Provide some basic definitions and results from measure theory.
2. Rigorously introduce the notion of "invariant measures."
3. Define ergodicity and state and prove **Birkhoff's Ergodic Theorem**.

Measures are a way for us to ascribe a size to sets. They do so by essentially weighting each point in a space X.

1. Lebesgue measure on $\mathbb{R}$ will be denoted by λ.
2. The size of an interval $[a, b]$ is obtained by computing $b - a$. The Lebesgue measure λ generalises this technique to arbitrary sets.

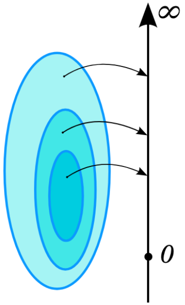

Figure: Visual depiction of what measures try to do. Credits: Wikipedia.

Measure Theory: Definitions

Definition (σ-algebra)

A collection of sets Σ is called a σ-**algebra** on X if:

1. (Contains the whole space) $X \in \Sigma$.
2. (Stable under complements) $A \in \Sigma \Rightarrow X \setminus A \in \Sigma$.
3. Stable under countable unions

Definition (Measure Space)

Let X be a set and let Σ be a σ-algebra over X. Then a set-function $\mu : \Sigma \to [0, \infty]$ is called a **measure** if:

1. $\forall E \in \Sigma$, $\mu(E) \geq 0$.
2. $\mu(\emptyset) = 0$.
3. μ is countably additive.

Measure Theory: Definitions

Definition (Probability Measure)

Let (X, Σ, μ) be a measure space. If $\mu(X) = 1$, then the triplet is called a **probability space** and μ is called a **probability measure**.

Definition (μ-almost every)

Let $P(x)$ be a property depending on a point $x \in X$. Define $N := \{x \in X \mid P(x) \text{ is false }\}$. Then, $P(x)$ holds μ-**almost everywhere** if $\mu(N) = 0$.

We can integrate with respect to different measures. The Riemann Integral is not the only way to "compute the area under the curve."

Probability Theory: Definitions

$\mathcal{B}(\mathbb{R})$ denotes the **Borel Sigma Algebra**.

Definition (Random Variable)

Let $(X, \mathcal{F}, \mu)$ be a probability space. Then, a measurable function $f : (X, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is called a **random variable**.

Definition (Expected Value)

Let $f : (X, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be a random variable. Then, the **expected value** is:

$$\mathbb{E}[f] := \int_X f d\mu < \infty \tag{2}$$

Motivation: invariant measures play a key role in the development of ergodic theory, as they provide the connection between space- and time-averages. Here, we consider a map $T : (X, \mathcal{F}, \mu) \to (T, \mathcal{G}, \gamma)$ be between two measure spaces.

1. T is said to be **measure-preserving** if $\mu(T^{-1}(A)) = \gamma(A)$ $\forall A \in \mathcal{G}$.
2. Example of a measure-preserving transformation. Let $X = \mathbb{Z}$. Define μ on $(\mathbb{Z}, \mathcal{F})$ as:

$$\mu(A) := \begin{cases} \text{card}(A) & \text{if } A \text{ is finite} \\ +\infty & \text{if } A \text{ is infinite} \end{cases}$$

   μ is called the **counting measure**.
   $T : \mathbb{Z} \to \mathbb{Z}$, $x \mapsto x + 1$ is measure-preserving with respect to μ.

Poincare's Recurrence Theorem

After a finite time, a dynamical system governed by a measure-preserving law will return to a state identical to the initial condition. This theorem was proven by Constantin Caratheodory in 1919.

Theorem (Poincare's Recurrence Theorem)

Let $(X, \mathcal{F}, \mu)$ be a probability space, and let $T : X \to X$ be a measure-preserving map. Let $A \in \mathcal{F}$ be such that $\mu(A) > 0$. Then, for μ-almost every point $X \in A$, $\exists\ n \in \mathbb{N}$ such that $T^n(x) \in A$. Moreover, there exist infinitely many $k \in \mathbb{N}$ for which $T^k(x) \in A$. In other words, almost every trajectory with an initial condition in A will return to A infinitely many times.

Proof.

On the blackboard. □

Question: Since Poincare's Recurrence Theorem tells us almost all trajectories must return to sets of strictly positive measure, **can we determine how often trajectories return to those sets, and if so, how?**

Answer: Birkhoff's Ergodic Theorem.

Ergodic Transformations

Definition (Ergodic Transformation)

We say that a map T is **ergodic** with respect to *mu* if $\forall\ A \in \mathcal{F}$ with $T^{-1}(A) = A$, there are exactly two possibilities:

1. $\mu(A)$ has full measure.
2. $\mu(A) = 0$

Equivalently:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n-1} \mu(T^i(A) \cap B) = \mu(A)\mu(B) \tag{3}$$

Definition (T-invariant)

Let $A \in \mathcal{F}$. Then, we say that A is **T-invariant** if $T^{-1}(A) = A$. For measurable functions f, f is **T-invariant** if $f \circ T = f$ a.e.

Theorem (Useful Properties of Ergodic Transformations)

Let $(X, \mathcal{F}, \mu)$ be a measure space and let $T : X \to X$ be a measure-preserving transformation. Then:

1. *If T is ergodic, then $T^{-1}(A) = A \Rightarrow T(A) = A$.*
2. *Let T be a measurable transformation and let f be an invariant measurable function. Then, T is ergodic $\iff$ f is constant a.e.*

Ergodic transformations: physics detour 2

Let H denote the total energy of the system. Define an **energy surface** S:

$$S_E := \{x, y \in H \mid H(x) - H(y) = 0\}$$

The surface generated by evaluating H at all points $x \in X$ is called the **Hamiltonian** of the system. It encodes the total energy of the system.

A physical system is said to be **ergodic** if almost all trajectories will flow "close" to almost every point on the same energy surface as the initial condition of the flow.

The ensembles of a dynamical system are the invariant measures of the dynamical system.

Ergodic transformations: physics detour 2

1. An **ensemble density** is a distribution of the dynamical states on an energy surface S_E. It ascribes a "probability" to each outcome.
2. Let $R \subseteq S_E$. An ensemble density ρ of an ensemble is
$$\int_R p(x)dx \tag{4}$$
3. For the microcanonical ensemble, ρ is chosen such that $\forall x \in S_E$, $\rho(x) = c$. It is an invariant ensemble.
4. **Ergodic Hypothesis in Statistical Mechanics**: For Hamiltonian systems, our system is ergodic on S_E if and only if the microcanonical ensemble is the *only* invariant ensemble.
    1. Similar to the previous theorem, where ergodicity is equivalent to f being constant a.e.

Birkhoff's Ergodic Theorem

Theorem (Birkhoff's Ergodic Theorem)

Let $(X, \mathcal{F}, \mu)$ be a probability space and let $T : X \to X$ be a probability-preserving map. Define the σ-algebra of T-invariant sets $\mathcal{G} := \sigma(\{A \in \mathcal{F} \mid T^{-1}(A) = A\})$. Let $f : X \to \mathbb{R}$ be an integrable random variable. Then:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N-1} f(T^n(x)) = \mathbb{E}[f|\mathcal{G}](x) \text{ a.s.} \tag{5}$$

If T is ergodic with respect to μ, we do not need the conditional expectation.

Corollary (Birkhoff's Theorem)

If, moreover, T is ergodic with respect to mu, then the time-average asymptotically equals the space-average:

$$\lim_{N\to\infty} \sum_{n=0}^{N-1} f(T^n(x)) = \mathbb{E}[f] \ a.s \tag{6}$$

1. This is nothing more than a rigorous formulation of Boltzmann's hypothesis in 1880!
2. Time-average encoded by $\lim_{N\to\infty} \sum_{n=0}^{N-1} f(T^n(x))$.
3. Space-average encoded by $\mathbb{E}[f]$.

Proof.

On the blackboard. □

Eigenfunctions and Eigenvalues: Another way to look at ergodicity

$L^2(X, \mu, \mathbb{C}) := f$ such that $(\int_X f^2 d\mu)^{\frac{1}{2}} < \infty$ is a banach space (a complete normed vector space).
The analogous concept to eigenvectors/eigenvalues are eigenvalues and **eigenfunctions**.

Definition (Eigenvalues/Eigenfunctions)

Let $(X, \mathcal{F}, \mu)$ be a probability space, and let $T : X \to X$ be a measure-preserving transformation. $\lambda \in \mathbb{C}$ is an **eigenvalue** of T if $\exists$ a non-zero $f \in L^2(X, \mu, \mathbb{C})$ such that:

$$f(T(x)) = \lambda f(x) \text{ for } \mu - \text{a.e.}$$

f is called the **eigenfunction** corresponding to λ of T.

If T is measure-preserving, then...

1. $|\lambda| = 1$ (eigenvalues measure how transformation stretch space).
2. T is ergodic $\iff$ for all eigenvalues λ of T, the sub-space $E(\lambda)$ is one-dimensional.
3. If T is ergodic and f is an eigenfunction, then $|f|$ is constant a.e.

References

1. "Ergodic Hypotheses in Classical Statistical Mechanics." (Thiago Werland Cesar R. de Oliveira (2007)).
2. "Ergodic Theory (Lecture Notes)" (Joan Andreu Lazaro Cami (2010))
3. "Ergodic Theory I (Max Plank Institute)" (YouTube videos)
4. "The Second Law of Thermodynamics: Comments from Ergodic Theory" (Michael C. Mackey).
5. "An Invitation to Ergodic Theory " (Silvia)
6. "Modern classical physics" (Thorne and Blandford)

Thank you for listening!