(In)finite Ergodic Theory

(Math 595 IEG)

Plan of the course

1) Examples of Dynamical Systems
2) Basic topological properties
3) Ergodic Theory. Basic notions, ergodic theorems, "law of large numbers"
4) Examples and applications to Number Theory
5) Infinite Ergodic Theory
6) "Central limit theorem(s)" for Dynamical Systems
7) Entropy theory
8) More about mixing when mixing doesn't make sense.
Ergodic theory studies the statistical properties of deterministic dynamical systems. "Dynamical" refers to time evolution, e.g. iteration of a function, or following the solution of an ODE/PDE.

Discrete dynamical systems: time evolves in discrete units, parametrized by $\mathbb{N}$ or $\mathbb{Z}$.
Continuous dynamical systems: the time variable changes continuously, parametrized by $\mathbb{R^+}$ or $\mathbb{R}$.

Discrete dynamical systems: obtained by iterating a map, say $T$, over a space $X$.
Examples of $X$: the unit interval $[0,1]$, the unit square $[0,1] \times [0,1]$, the circle $\mathbb{R}/\mathbb{Z}$, a Cantor set, ...

Let $T: X \to X$ be a map. For $x \in X$, consider the iterates $x, T(x), T(T(x)), T(T(T(x)))$.
... We use the notation $T^0 = \text{id}$ and $T^n = T \circ \cdots \circ T$. 
We denote by $\Theta^+_T(x)$ the forward orbit of $x$ under $T$, i.e.

$$\Theta^+_T(x) = \{ x, T(x), T^2(x), \ldots \} = \{ T^n(x), n \geq 0 \}$$

**Example 1.1**

$X = [0,1]$, $T(x) = 4x(1-x)$.

$$\Theta^+_T(\frac{1}{3}) = \left\{ \frac{1}{3}, \frac{4}{9}(1-\frac{1}{3}) = \frac{8}{9}, 4 \cdot \frac{8}{9}(1-\frac{8}{9}) = \frac{32}{81}, \ldots \right\}$$

**Example 1.2**

$x = \text{circle of radius 1}$, $T : X \rightarrow X$ clockwise rotation by angle $2\pi \alpha$.

---

If $T$ is invertible, then $T^{-1} : X \rightarrow X$ and we can consider the backward iterates

$$T^{-1}(x), T^{-2}(x), \ldots, T^{-n}(x) = (T^{-1} \circ \cdots \circ T^{-1})(x)$$

If $T$ is not invertible, then $T^{-1}(A)$, where $A \subseteq X$ denotes the set of preimages of $A$, i.e. $\{ x \in X : T(x) \in A \}$.

If $T$ is invertible, then the **full orbit** of $x$ under $T$ is
\[ \Theta_T(x) = \{ T^n(x), -T^{-n}(x), \ldots, T^{-2}(x), T^{-1}(x), T^0(x), T(x), T^1(x), \ldots \} \]

Although the (backward or forward) time evolution is deterministic, the system can be "chaotic." All notions of "chaos" require sensitive dependence on initial conditions, i.e., even if \( x \) and \( y \) are very close, there exists some integer \( n \) such that \( T^n(x) \) and \( T^n(y) \) are far apart.

Topological Dynamics and Ergodic Theory give ways to predict how chaotic a system is, and predict asymptotic (long time) behavior.

Behavior of a single orbit can be very complex, and often one can only predict the average behavior.

First dichotomy: is the orbit finite or infinite?

Def \( x \in X \) is periodic if \( \exists n > 0 \) such that \( T^n(x) = x \). If \( n = 1 \), then \( x \) is a fixed point. \( n \) is called a period for \( x \).
and $T^{n+j}(x) = T^{j}(x)$ for all $j$.

**Example**

In example 1.1. $x = \frac{3}{4}$ is a fixed point.

$$T(\frac{3}{4}) = 4 \cdot \frac{3}{4} (1 - \frac{3}{4}) = \frac{3}{4}$$

**Example**

In example 1.2., with $a = \frac{1}{4}$, we have a rotation by $\frac{\pi}{2}$ and all points are periodic with period 4.

For a periodic point $x$, the minimum period is the smallest $n \geq 1$ such that $T^n(x) = x$.

A point $x$ is preperiodic if $\exists k, n \in \mathbb{N}$ such that $T^{n+k}(x) = T^k(x)$.

In this case $T^{n+j}(T^k(x)) = T^j(T^k(x))$.

**Exercise**: If $T$ is invertible, then every preperiodic point is periodic.

**Typical topological questions**: Are there fixed points? Are there periodic points? Are periodic points dense?
Do there exist orbits which are dense? Are all orbits dense?

If an orbit is dense, then it visits every part of the space. \( \Rightarrow \) Natural question: how much time does it spend in each part of the space?

Let \( A \subset X \), consider \( 1_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} \).

Number of visits up to time \( n \):

\[
\# \{0 \leq k \leq n : T^k(x) \in A \} = \sum_{k=0}^{n-1} 1_A(T^k(x)).
\]

Frequency of visits:

\[
\frac{1}{n} \sum_{k=0}^{n-1} 1_A(T^k(x))
\]

(1)

Equidistribution of the orbit \( \Theta(x) \) roughly means that, as \( n \to \infty \), this frequency approaches the "size" of \( A \) (volume, area, ...).

This means that the orbit asymptotically spends in each part of the space a time proportional to the size.
A main question in Ergodic Theory is: Are orbits equidistributed?

Remark: the limit as $n \to \infty$ of (1) need not exist!

A main result in E.T. is the Birkhoff ergodic theorem, which guarantees the existence of this limit for almost every point $x \in X$, and (if the system is "chaotic" enough) shows that the frequency converges to the expected value.

In Probability Theory, for a sequence of i.i.d. random variables
\[ \{\xi_0, \xi_1, \xi_2, \ldots \} \]
the convergence of \[ \frac{1}{n} \sum_{k=0}^{n-1} 1_A (\xi_k) \]
almost surely to $P(A)$ is called the (strong) law of large numbers.

The analogs in Ergodic Theory for \[ \{A(\xi), A(T(\xi)), A(T^2(\xi)), \ldots \} \text{ or in general} \]
\[ \{f(x), f(T(x)), f(T^2(x)), \ldots \} \text{ for some observable } f \]
is known as an ergodic theorem.  
One can view the sequence 
\[ f(x), f(Tx), f(T^2x), \ldots \] as a sequence of STRONGLY DEPENDENT random variables, 
where the only randomness comes with the choice of \( x \in X \).

In Probability \( P(X) = 1 \), and all the classical probabilistic results (law of large numbers, 
central limit theorem, \ldots) can be rephrased in the context of dynamical systems 
PROVIDED \( X \) has FINITE "Size" by normalizing 
\[
\frac{\text{Size}(A)}{\text{Size}(X)} = \tilde{\text{Size}}(A), \quad \text{where} \quad \tilde{\text{Size}}(A) \quad \text{can}
\]
be thought of as the probability of a set (event) \( A \subset X \).

However, here are very natural examples where 
\( \text{Size}(X) \) is truly infinite. This 
apparently small detail is the main problem of Infinite Ergodic Theory.
In Infinite E.T. most elementary results are simply false; however, there are weaker results (e.g., distributional ergodic theorems) that reduce to the elementary ones when $\text{Size}(X) < \infty$.

---

Examples of applications to Number Theory

The theory of continued fractions.

Every real number $x \in \mathbb{R}$ can be written as $x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots}}} = [a_0; a_1, a_2, a_3, \ldots]$ with $a_j \in \mathbb{N}$, $a_0 > 0$, $a_1, a_2, \ldots$

and approximations of $x$ can be obtained by truncating the above expression

$$x \approx a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}} = \cfrac{p_n(x)}{q_n(x)} \in \mathbb{Q}$$

A question in Diophantine approximation is "how good" is the approximation of $x$ by $\cfrac{p_n}{q_n}$
One can show, using Ergodic Theory, that for almost every \( x \in (0, 1) \)
\[
\lim_{n \to \infty} \frac{1}{n} \log \left| 1 - \frac{p_n(x)}{q_n(x)} \right| = -\frac{\pi^2}{6 \log 2},
\]
that is typically and asymptotically
\[
\left| x - \frac{p_n(x)}{q_n(x)} \right| \sim e^{-n \cdot \frac{\pi^2}{6 \log 2}}.
\]
This is related to the magnitude of the denominators \( q_n \). One can show (again using E.T.) that for almost every \( x \in (0, 1) \),
\[
\lim_{n \to \infty} \frac{1}{n} \log q_n(x) = \frac{\pi^2}{12 \log 2},
\]
i.e. \( q_n(x) \sim e^{n \cdot \frac{\pi^2}{12 \log 2}} \) for typical \( x \) and large \( n \).
A similar continued fraction representation of an arbitrary real number is

\[ x = a_0 + \frac{b_0}{a_1 + \frac{b_1}{a_2 + \frac{b_2}{a_3 + \ldots}}}, \]

where all the \( a_j \) are \text{EVEN} \( a_3 \ldots \) and \( b_j = \pm 1 \).

One can ask the same question as before: Consider the approximation of \( x \) given by

\[ a_0 + \frac{b_0}{a_1 + \frac{b_1}{a_2 + \frac{b_2}{a_3 + \ldots}} = \frac{p_n(x)}{q_n(x)} \in \mathbb{Q}. \]

How does the sequence \( q_n \) behave for large \( n \)? It follows from a deep result in Infinite Ergodic Theory that

\[ q_n \sim e^{\frac{n}{2n}} \quad \text{asymptotically but not almost surely; more precisely} \]

\[ q_n \sim e^{\frac{n}{2n}} \quad \text{asymptotically but not almost surely; more precisely} \]
\[ P(\{ x \in (0,1) : \frac{\log q_n(\alpha)}{\sqrt{n}} - \frac{\pi^2}{4} \geq \varepsilon \}) \to 0 \quad n \to \infty \]

for every \( \varepsilon > 0 \) and every continuous probability density \( p_{un} (0,1) \).

In this setting, no "almost everywhere" result will hold, except trivial ones.

In some cases, like the one shown, convergence will be in probability (weaker than a.e.)
Examples of Dynamical Systems: Circle Rotations

Consider the unit circle $S^1$, identified with the 1-torus $\mathbb{R}/\mathbb{Z} = X = [0,1)$.

Consider the rotation by angle $\alpha$

$$R_\alpha: \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}, \quad R_\alpha(x) = x + \alpha \mod 1.$$ 

Notice that $R_\alpha(x) = \begin{cases} 
  x + \alpha & \text{if } x + \alpha < 1 \\
  x + \alpha - 1 & \text{if } x + \alpha \geq 1 
\end{cases}$

-warning: Notice that it is a piecewise isometry.

Theorem (dichotomy for rotations)

1) If $\alpha = \frac{p}{q} \in \mathbb{Q}$, then all orbits are periodic with period $q$.

2) If $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, then the orbit of every point is dense.

Proof: 1) Let $\alpha = \frac{p}{q} \in \mathbb{Q}$. For every $x \in \mathbb{R}/\mathbb{Z}$

$$R^q_\alpha(x) = x + q \frac{p}{q} \mod 1 = x + p \mod 1 = x \mod 1 = x$$

Therefore, every point $x$ is periodic with period $q$. 


2) Let's use multiplicative notation for $\mathbb{R}/\mathbb{Z} = \{ e^{2\pi i x} | x \in [0, 1) \}$.

Let $\alpha$ be irrational.

Then for each $z_1 = e^{2\pi i x_1} \in S^1$,

\[ R^n_\alpha(z_1) = R^m_\alpha(z_1) \quad \forall m \neq n. \]

In fact, if they were equal $e^{2\pi i (x_1 + md)} = e^{2\pi i (x_1 + nd)}$ then $2\pi (x_1 + md) = 2\pi (x_1 + nd) + 2\pi k$ for some $k \in \mathbb{Z}$ and therefore $md = nd + k$. But since $m \neq n$ then $\alpha = \frac{k}{(m-n)} \in \mathbb{Q}$ (contradiction!)

We need to show that $\forall z_2 \in S^1$, $\forall \epsilon > 0$ there exists a point of $B(z_2, \epsilon)$ such that $R^+\alpha(z_1) \in B(z_2, \epsilon)$.

Let $N$ be big enough so that $\frac{1}{N} < \epsilon$.

Consider $z_1$, $R_\alpha(z_1)$, $R^2_\alpha(z_1)$, $\ldots$, $R^n_\alpha(z_1)$.

Since they are all distinct, by Pigeonhole principle there exist $n, m$ such that $0 \leq n < m \leq N$ with

\[ d(R^n_\alpha(z_1), R^m_\alpha(z_1)) \leq \frac{1}{N} < \epsilon \]

[Here $d(e^{2\pi i x}, e^{2\pi i y}) = \min\{|x-y|, 1-|x-y|\}$]

$\Rightarrow \exists \theta$ such that $|\theta| \leq \frac{1}{N}$ and

\[ R^m_\alpha(z_1) = e^{2\pi i \theta} R^n_\alpha(z_1) \quad \Rightarrow \forall \epsilon \quad z_1 = e^{2\pi i x_1} \epsilon \quad z_1 < \epsilon$
\[ e^{2\pi i m} = e^{2\pi i n} \]

Claim: \( R_d^{m-n} \) is a rotation by angle multiple of \( \epsilon \).

In fact, \( R_d^{m-n} (z_1) = e^{2\pi i (m-n)} z_1 = e^{2\pi i \theta} z_1 = R_d^\theta (z_1) \)
and \( |\theta| < \frac{\epsilon}{2} \).

Now consider \( z_1, R_d^{m-n} (z_1), R_d^{2(m-n)} (z_1), R_d^{3(m-n)} (z_1), \ldots \)
\( i.e \quad e^{2\pi i x_1}, e^{2\pi i (x_1+\theta)}, e^{2\pi i (x_1+2\theta)} \)
and these are points on \( S^1 \) with spacing less than \( \epsilon \).

\( \Rightarrow \forall z_2 \in S^1 \exists j > 0 \text{ s.t. } R_d^{j(m-n)} (z_1) \in B(z_2, \epsilon) \)

\( \Rightarrow \) the orbit \( \Omega^+_{R_d} (z_1) \) is dense \( \square \)

An application to Number Theory

Theorem (Dirichlet Theorem)

If \( a \) is irrational, for each \( \delta > 0 \) there exists an integer \( g \), \( 0 < q < \delta \) and \( p \in \mathbb{Z} \) such that

\[ \left| \frac{p}{q} - \frac{a}{g} \right| < \frac{\delta}{q} \]
Proof: Orbit of 0: \( R^n(0) = na \mod 1, n \geq 0. \)

By Theorem 1.1 for irrational \( \alpha \), the orbit \( R^*_n(0) \) consists of distinct points.

Choose \( N \) s.t. \( \frac{1}{N} \leq \frac{1}{\delta} \) and by Pigeonhole Principle

\[ \exists 0 \leq m < N \text{ s.t. } \]

\[ d(R^m(0), R^n(0)) < \frac{1}{N} \]

Define \( q = m - n \).

Since \( R_\alpha \) is an isometry

\[ d(0, R^q_\alpha(0)) = d(0, R^{m-n}_\alpha(0)) = d(R^m_\alpha(0), R^n_\alpha(0)) < \frac{1}{N} \]

\[ \implies \exists p \in \mathbb{Z} \text{ s.t. } 1q\alpha - p \leq \frac{1}{N} \leq 1q\alpha - p \]

\[ \implies |\alpha - \frac{p}{q}| \leq \frac{\delta}{N} \]

\[ \text{i.e. } |\alpha - p/q| \leq \frac{\delta}{N} \]

Corollary: If \( \alpha \) is irrational, then there exist infinitely many rational numbers \( p/q \) with \( p \in \mathbb{Z}, q \in \mathbb{N} \) that solve the inequality \( |\alpha - p/q| \leq \frac{1}{q^2} \).
Examples of Dynamical Systems: the quadratic map

$X = [0, 1]$, \( T_\mu(x) = \mu x (1-x) \), \( 0 < \mu \leq 4 \)

As the parameter \( \mu \) varies, we have a family of maps called the quadratic family or the logistic family.

If \( 0 < \mu \leq 4 \), then \( T_\mu([0, 1]) \subseteq [0, 1] \)

**Example:** \( \mu = \frac{5}{2} \)

\[
T_\mu(0) = 0 = T_\mu(1) \\
T_\mu'(x) = \frac{5}{2} - 5x \\
T_\mu'(x) = 0 \quad \Rightarrow \quad x = \frac{1}{2} \\
\text{and} \quad T_\mu\left(\frac{1}{2}\right) = \frac{5}{8}
\]

Fixed points are solutions of \( T_\mu(x) = x \), i.e. \( \frac{5}{2}x - \frac{5}{2}x^2 = x \)

\( \Rightarrow \quad x \left( \frac{3}{2} - 5x \right) = 0 \quad \Rightarrow \quad x = 0 \quad \text{or} \quad x = \frac{3}{5} \)

Behavior of orbit near the fixed points

- \( x = 0 \) (repelling)
- \( x = \frac{3}{5} \) (attracting)
A fixed point \( x \) is **attracting** if \( \exists \) a ball \( U = B(x, \varepsilon) \) around \( x \) s.t. \( T(U) \subseteq U \)

and \( \bigcap_{n \geq 0} T^n(U) = \{ x \} \).

A fixed point is **repelling** if \( \exists \) a ball \( U = B(x, \varepsilon) \) around \( x \) s.t.

\[
\overline{U} \subseteq T(U) \quad \text{and} \quad \bigcap_{n \geq 0} T^{-n}(U) = \{ x \}
\]

**Remark:** If \( T \) is invertible, \( x \) is attracting for \( T \) \( \iff \) \( x \) is repelling for \( T^{-1} \) and vice versa.

**Theorem 1.2** (criterion to determine a fixed pt.) is repelling or attracting.

Let \( X \subset \mathbb{R} \) be an interval, and let \( T : X \to X \) differentiable with continuous derivative.

Let \( x = T(x) \) be a fixed pt.

1) \( \text{If } |T'(x)| < 1 \text{, then } x \text{ is attracting.} \)

2) \( \text{If } |T'(x)| > 1 \text{, then } x \text{ is repelling.} \)

**Remark:** If \( |T'(x)| = 1 \text{, it's not possible to determine } \)
whether \( x \) is attracting or repelling just from the information about \( f'(x) \).

**Proof** Since \( T \) is continuous and \( |T'(x)| < 1 \), \( \exists \varepsilon > 0 \) \( \forall x \in (x-\varepsilon, x+\varepsilon) = B(x, \varepsilon) \)

we have \( |T'(y)| < \rho < 1 \). Then \( \forall y \in B(x, \varepsilon) \), since \( x \neq T(x) \), by mean value theorem

\[ \exists \xi \in B(x, \varepsilon) \text{ s.t.} \]

\[ |T(y) - x| = |T(y) - T(x)| = |T'(\xi)| \cdot |y - x| \leq \rho \cdot |y - x| \leq \rho \varepsilon \]

This means \( f|_B(x, \varepsilon) \) \( \forall y \in B(x, \varepsilon) \), i.e. \( T(B(x, \varepsilon)) \subset B(x, \varepsilon) \). We claim now that

\[ |T^n(y) - x| \leq \rho^n \varepsilon. \]

We already knew this for \( n=1 \). By induction \( \forall y \in B(x, \varepsilon) \exists \xi \in B(x, \varepsilon) \text{ s.t.} \)

\[ |T^{n+1}(y) - T^{n+1}(x)| = |T(T^n(y)) - T(T^n(x))| =
\]

\[ = |T'(\xi)| \cdot |T^n(y) - T^n(x)| \leq \rho |T^n(y) - x|. \] and

by inductive hypothesis.
\[ |T^{n+1}(y) - x| = |T^{n+1}(y) - T^{n+1}(x)| \leq p \cdot |T^n(y) - x| \leq p^{n+1} \cdot \epsilon \]

This shows the claim.

Now, since \( p < 1 \), \( \lim_{n \to \infty} p^n = 0 \) and

\[ \lim_{n \to \infty} T^n(y) = x \quad \text{and} \quad \bigcap_{n \geq 0} T^n(B(x, \delta)) = \{x\} \]

2) Exercise

---

Fact: For small value of the parameter, say \( 0 < p < 3 \), then all points are either attracted or repelled.

For \( \mu \in (3, 4) \) the dynamics is very complicated, exhibiting a phenomenon called period doubling.

Route to chaos.
At most the map $T_4: [0, 1] \rightarrow [0, 1]$ is "chaotic". We will see that it is (topologically) semi-conjugated to the doubling map $T: [0, 1] \rightarrow [0, 1]$.

$$T(x) = 2x \mod 1.$$