Ergodic theory (ET)

Ergodic theory is mathematical study of the long-term average behavior of systems.

- X - the collection of all states of a system
- Evolution of the system is represented by transformation $T: X \to X$

(Tx is the state at time 1 of the system which at time 0 is at state x)

- X - C^∞ mnfld, T - diffeo $\to$ diff-dynamics
- X - top. space, T - homeo $\to$ top-dynamics

Topic for today's lecture $\to$
- X - measure space, T measure preserving transformation (MPT) $\to$ ET

Set up: • $(X, \mathcal{B}, \mu)$ (complete) probab. space: $\mathcal{B}$ - σ-algebra of measurable subsets

$\mathcal{B}$ - Family of observable events

μ - specifies time-independent probabilities of occurences of events in $\mathcal{B}$.

μ - countably additive non-negative function on $\mathcal{B}$ subsets of X s.t. $\mu(X) = 1$, and

$\mathcal{B}$ - contains all subsets of sets of measure 0.

- $T: X \to X$ is 1-1 onto map such that T, T^{-1} are both measurable: $T^{-1}\mathcal{B} = T\mathcal{B} = \mathcal{B}$.

Rem: Since sets of measure 0 don't matter, we only care that T is well-defined and 1-1 after ~~a set~~ on $X \setminus$ set of measure 0

- T is μ-preserving (MPT): $\mu(T^{-1}E) = \mu(E)$

$$(T_* \mu = \mu)$$

Rem: Also non-invertible case ...

- For $x \in X$: $\text{orbit}(x) = \{T^n x : n \in \mathbb{Z}\}$ - complete history of the system, future + past.

1'

Poincaré recurrence : $T : X \to X$ preserves probab. measure μ then every point of each positive measure set E returns to E ∞ often:

T - MPT of (X, μ), $E \subset X$ measurable, $\mu(E) > 0$

$\Rightarrow \exists$ measurable set $F \subseteq E$ with $\mu(F) = \mu(E)$

s.t. $\forall x \in F$, $\exists\, 0 < n_1 < n_2 < \dots$ in $\mathbb{Z}$ s.t.

$T^{n_i} x \in E$

pf Let $B = \{x \in E \mid T^n x \notin E \text{ for any } n \geq 1\}$

Then $B = E \cap T^{-1}(X \setminus E) \cap \dots \cap T^{-n}(X \setminus E) \cap \dots$

is measurable

$\forall n \geq 1$: $T^{-n} B = T^{-n} E \cap T^{-n-1}(X \setminus E) \cap \dots$

is disjoint from any previous one.

all sets $B, T^{-1}B, \dots$ are disjoint and all have measure $\mu(B)$ since T preserves measure

$\Rightarrow \mu(B) = 0$ (since μ is probab. measure)

$\Rightarrow \mu(B^c) = \mu(E)$ $\square$

Q: Exactly how long a time do the images of recurrent pts spend in E? This is exactly the question about $\frac{1}{n} \sum f(T^j x)$ - Cesàro convergence of $\{f(T^n x)\}$

- Now let $f: X \to \mathbb{R}$ be a measurable function.
(It may represent a measurment made on the system.)

- Sequence $f(x), f(Tx), \ldots$ may be thought of as the values of some physically interesting variable at successive instants of time, beginning with world at initial state x.

- For a large # of successive measurements (observations) it makes sense to look at the average

$$\frac{1}{N} \sum_{k=0}^{N-1} f(T^k x)$$

- Basic question of ET: Convergence of these averages? if the limit does exist, it can be thought of as some sort of equilibrium for f.

- Boltzmann ergodic hypothesis → (CONTEXT: gas dynamics)

(*) $\lim \frac{1}{N} \sum_{k=1}^{N-1} f(T^k x) = \int_X f \, d\mu \quad \mu$ a.e. X.

TIME Mean → Space mean.

(equilibrium of the measurment coincides with the weighted average over all possible states of the system)

- This is PRECISELY true for ERGODIC SYSTEMS. Equivalent statements of ergodicity:
  - Every point visits each set of positive measure
  - $\mu(E) > 0, \mu(F) > 0 \Rightarrow \mu(T^n E \cap F) > 0$ for some $n > 0$

  - If E is: $\mu(E) > 0$, $T^{-1}E = E \Rightarrow \mu(E) = 0$ or 1
  - Cannot split the space into components of pos measure, which are dynamically independent
  - Every L^2-inv. function is constant a.e.

$\boxed{\text{Prop}}$ T is ergodic $\Leftrightarrow$ $I_T := \{f \in L^2(X) : f(x) = f(Tx) \text{ a.e.}\}$
is $= \mathbb{C}\chi_X$

$\underline{\text{Pf}}$ If $E \in \mathcal{B}$ is T-invariant $\Rightarrow \chi_E$ is T-invariant

$\Rightarrow \chi_E \in I_T$ $\Rightarrow \chi_E$ is either 1 or 0 a.e.

i.e. $\mu(E) = 0$ or 1

Suppose T is ergodic. If $f \in I_T$: $f = U_T f$.

Then $\exists a$ s.t. the sets $E_+ = \{x \in X : f(x) < a\}$
$E_- = \{x \in X : f(x) \geq a\}$
are T-invariant $\underbrace{}_{\text{positive measure}}$, E_+, E_- are measurable $\Rightarrow$

$\mu(E_\pm) = 1$ or 0 $\lightning$

1870 → For large systems of interacting particles in equilibrium, time averages are close to the ensemble or equilibrium average (2)

[Boltz]mann ergodic hypothesis: Mechanical system (ex: gas dynamics)

[st]arting from any point, under time evolution, passes eventually through every state on energy surface.

εργον = energy $\qquad \frac{1}{N}\sum_{n=1}^{N-1} f(T^n x) = \int_X f\,d\mu(x)$ a.e. μ.

οδος = path $\qquad$ take $f = \chi_E$

Maxwell called this: "continuity of path"

Gibbs observed this in measurements.

- At the time: measure theory of Borel, Lebesgue and modern topology were not sufficiently developed for thorough set-up of statistical mechanics.

1930's von Neumann, Birkhoff, Koopman → 1929: Ergodic theorems
Stone, A. Weil, ...

Oxtoby-Ulam: On a compact polyhedron M with Lebesgue-Stiltjes measure, the set of all ergodic measure preserving homeos is dense G_δ or residual among all measure preserving homeos.

But: 1974 (Markus Meyer) (earlier: KAM)

Thm In the space of smooth Hamiltonians
(1) The non-ergodic ones form a dense open subset
(2) the non-integrable ones form a dense open subset.

(Hamiltonian is ergodic if for almost every value of the energy, the system is ergodic on the corresponding submanifold in the phase space.

original problem (Boltzmann) about gas dynamics being ergodic is still open!!!

(Hiperbolicity : Hedlund, Hopf, ...
'38, '39

Boltzmann - Sinai Ergodic hypothesis : 1963 (Sinai)

The system of N hard balls given on $\mathbb{T}^2$ or $\mathbb{T}^3$ is ergodic $\forall N \geq 2$.

(Ergodicity is expected to hold on (connected components of the submanifold of the phase space specified by the invariants of motion.

Sinai proved this for 2d 2 balls on $\mathbb{T}^2$

Sinai : billiards

D. Szasz, Simanyi, ... classic hard ball systems.

Arriving at similar questions from another direction:

Q1: Given $N \in \mathbb{N}$, does there exist $n \in \mathbb{N}$ such that
$2^n = N\ldots$ (2^n starts with digits of N)

Q2 Which digit among $\{1, \ldots, 9\}$ is most common as a leading digit of 2^n

Q1: ex: $N = 12345$

$12345 \cdot 10^p \leq 2^n < (12345+1)\, 10^p$ / $\log_{10}$?

$\{\log_{10} 12345\} \leq \{n \log_{10} 2\} \leq \{\log_{10} (12346)\}$

Q2 $k \cdot 10^p \leq 2^n < (k+1)\, 10^p$ / $\log_{10}$

$k = 1, 2, \ldots, 9$

$\log_{10} k \leq \{n \log_{10} 2\} < \log_{10}(k+1)$

Let $\log_{10} 2 = \alpha \in [0,1]$

Define $T : [0,1]/_\sim \to [0,1]/_\sim \simeq S^1$

$T : x \mapsto x + \alpha \mod 1$

Q2: $T^n(0) \in [\underbrace{\log_{10} k}_{a}, \underbrace{\log_{10}(k+1)}_{b})$

$$\frac{\#\{i = 0, \ldots, N-1 \mid n\alpha \mod 1 \in (a,b)\}}{N} \to ?$$

$$\frac{\sum_{n=0}^{N-1} \chi_{(a,b)}(T^n(0))}{N} \to ?$$

Equidistribution theorem for $\alpha \notin \mathbb{Q}$:

$\forall x$ $\longrightarrow \int \chi_{(a,b)}\, d\lambda = b - a$ $\Rightarrow$ Answer to Q2: digit 1 occurs most often.

4

another way to view time averages: $\lim \frac{1}{N}\sum \chi_A(T^n x) = \lim \frac{1}{N}\sum \delta_{T^n x}(A)$

$\int f \, d\delta_x = f(x)$

if $f \in C(X)$ then

$$\lim \frac{1}{N}\sum f(T^n x) = \lim \int f \, d\left(\frac{1}{N}\sum \delta_{T^n x}\right)$$

$$\downarrow ???$$

$$\int f \, d\mu$$

This is weak* convergence in the space of probab. measures $\mathcal{M}(X)$:

so the question is whether

$$\frac{1}{N}\sum \delta_{T^n x} \xrightarrow{w^*} \mu .$$

Fact: if $\nearrow$ has a weak* limit, then the limit is a probability measure which is T-invariant

Moreover: if $T^{-1}A = A$ then $\delta_{T^n x}(A) = \delta_x(A)$, so $\mu(A) = 1$ or 0.

So: these weak* limits are all ergodic measures

$M^T(X)$ denote T-invariant probab. measures

$\boxed{\text{Thm}}$ $T: X \to X$ cts, X-cpct metric space.

ν_n – any sequence in $\mathcal{M}(X)$, let

$$\mu_n = \frac{1}{N}\sum_{k=0}^{N-1} T_*^k \nu_n$$

Then every weak* limit point of $\{\mu_n\}$ is a member of $M^T(X)$.

Since $\mathcal{M}(X)$ is weak* compact $\Rightarrow$ μ_n must have a limit point

$\boxed{\text{Corollary}}$ [Kryloff – Bogoliuboff thm] $T: X \to X$ cts, X-cpct metric space

$$\mathcal{M}^T(X) \neq \emptyset$$

Fact (1) $\mathcal{M}^T(X)$ is convex (exercise)

(2) Ergodic measures in $\mathcal{M}^T(X)$ are exactly extremal points in $\mathcal{M}^T(X)$:

$$\mu\text{-ergodic for } T \Longleftrightarrow \mu \text{ is extremal in } \mathcal{M}_T$$

Pf Suppose: $\mu = \lambda\mu_1 + (1-\lambda)\mu_2$ and μ-ergodic

$\mu_1 \ll \mu \xRightarrow{\text{Radon-Nykodim}} \mu_1(E) = \int_E g\,d\mu$ for some g.

Since $\mu_1 \in \mathcal{M}_T \Rightarrow \forall E$ $\int_E g\,d\mu = \int_{T^{-1}E} g\,d\mu = \int_E g\circ T\,d\mu \Rightarrow g = g\circ T$ μ a.e.

$\Rightarrow g = \text{Const}$ μ a.e

$\Rightarrow \mu_1(X) = \text{Const}\cdot\mu(X) \Rightarrow \text{Const} = 1$ and $\mu_1 = \mu$.

Conversely: If $\exists A$ s.t. $T^{-1}A = A$ and $\mu(A) \neq 0, 1$ $\Rightarrow$

$\chi_A \circ T = \chi_A$. Define $\mu_1(E) = \frac{\mu(E\cap A)}{\mu(A)}$, $\mu_2(E) = \frac{\mu(E\cap A^c)}{\mu(A)}$

$\Rightarrow \mu = \mu(A)\mu_1 + (1-\mu(A))\mu_2 \Rightarrow \mu$ is not extremal perator theory point of view

$$U_T : f \longmapsto f \circ T \; ; \text{ then } \int f \, d\mu = \int U_T f \, d\mu$$

oopman operator

- If $f \in L^2$: $\langle U_T f_1, U_T f_2 \rangle = \int f_1 \circ T \cdot \overline{f_2 \circ T} \, d\mu$
  $= \int f_1 \overline{f_2} \, d\mu \quad (T_* \mu = \mu)$
  $= \langle f_1, f_2 \rangle$

In particular: $\| U_T f \| = \| f \|$

- If T is invertible then U_T is invertible

and $\langle U_T f, g \rangle = \langle f, U_T^* g \rangle$

$\langle U_T f, g \rangle = \langle U_T U_T^{-1} f, U_T^{-1} g \rangle = \langle f, U_T^{-1} g \rangle$

$\Big\} \Rightarrow U_T^* = U_T^{-1}$

- Any invertible MPT on a measure space $(X, \mathcal{B}, T)$ induces linear operator $U_T : \mathcal{L}^2(\mu) \circlearrowleft$

and U_T is an isometric isomorphism, $U_T^* = U_T^{-1}$

So U_T is <u>UNITARY</u>

<u>Prop</u> MPT T is ergodic wrt μ iff 1 is a simple eigenvalue of U_T

Γ $\{ f \in L^2(X) \mid U_T f = f \} = \mathbb{C} \chi_X \iff \mu$-ergodic

$\Updownarrow$

1 is a simple eigenvalue of U_T

$\left(\dim \left(\bigcup_{k=1}^{\infty} \ker (\lambda I - U_T)^k \right) = 1 \iff \lambda \text{ is simple eigenvalue} \right)$

$\lrcorner$ T - ergodic and invertible $\Rightarrow$

<u>EVERY</u> eigenvalue of U_T is simple and

the set of all eigenvalues $\subset S^1$

$\ulcorner$ $U_T f = \lambda f$

$$\left.\begin{array}{l} \langle U_T f, f \rangle = \langle \lambda f, f \rangle = \lambda \langle f, f \rangle \\ \quad \| \\ \langle f, U_T^* f \rangle = \langle f, U^{-1} f \rangle = \langle f, \frac{1}{\lambda} f \rangle = \frac{1}{\lambda} \langle f, f \rangle \end{array}\right\} \Rightarrow |\lambda| = 1$$

<u>Ex1</u> R_α, $\alpha \notin \mathbb{Q}$ has PURELY DISCRETE SPECTRUM (pure point spectrum) i.e. $\exists$ ONB in $L^2(X, \mu)$ made up of eigenfunctions e_n

<u>Ex2</u> E_2 (- not invertible so U_{E_2} is not unitary, but it is an isometry)

$f = \sum a_n \chi_n$

$f \circ E_2 = \lambda f$

$$\sum a_n e^{2\pi i n \cdot 2x} = \lambda \sum a_n e^{2\pi i n x}$$

$$\sum a_n e^{2\pi i 2n \cdot x} = \sum \lambda a_n e^{2\pi i n x}$$

$a_{\frac{n}{2}} = \lambda a_n$, n - even ;

Von Neumann's ergodic theorem

$U : \mathcal{H} \to \mathcal{H}$ - Hilbert space is unitary (enough that U is isometry)

$P : \mathcal{H} \to W_U$ projection from $\mathcal{H}$ to U-invariant vectors.

Thm: $\forall f \in \mathcal{H}$ $\quad \frac{1}{n} \sum_{j=0}^{n-1} U^n(f) \xrightarrow[L^2]{} P(f)$

proof

- Thm holds trivially for f which is U-invariant.
- Thm holds for coboundaries: $f = g - Ug$

$$Pf = Pg - \underbrace{P(Ug)}_{= Pg} = 0 \qquad \begin{cases} g = \underset{Pg}{f_1} + f_2 \\ Ug = Uf_1 + Uf_2 \quad Uf_1 = f_1 \text{ b/c } f_1 \text{ is } U\text{-inv.} \end{cases}$$

by the telescoping argument

- Then the thm holds for L^2 limits of coboundaries.
- Crucial in the proof is to show that

$$\overline{\text{coboundaries}}^{L^2} = W_U^{\perp}$$

So suppose $\langle f, g - Ug \rangle = 0 \quad \forall g$

$$\langle f - U^* f, g \rangle = 0 \quad \forall g$$

$$\langle f, g \rangle = \langle U^* f, g \rangle = \langle U^{-1} f, g \rangle \quad \forall g$$

$$\Rightarrow f = U^{-1} f \quad \Rightarrow \quad Uf = f \quad \Rightarrow f \text{ is } U\text{-inv.}$$

- Since thm. holds for W_U and for $W_U^{\perp}$ then it holds in all of $\mathcal{H}$.

$\triangle$

Rem: If $U = U_T$ and U_T is ergodic, then $W_U = \mathbb{C} \chi_X$

and $P(f) = \int_X f \, d\mu$

nm | Pointwise ergodic theorem

T - MPT on X, $f \in L^1$ then

$\frac{1}{n}\sum_{j=0}^{n-1} f(T^j x)$ converges a.e. to some L^1 function $f^*(x)$

and f^* is T-invariant. and $\int f^* = \int f$

Corollary: μ-ergodic $\Rightarrow$

$$\underbrace{\frac{1}{n}\sum f(T^j x)}_{\text{time mean}} \longrightarrow \underbrace{\int_X f\,d\mu}_{\text{space mean.}} \quad \text{a.e. } \mu$$

Corollaries: • $f = \chi_E$ then time average $\to \mu(F)$ a.e. x

which means: average time which orbit of a pt. spends in a positive measure set F is $\mu(F)$ for a.e. orbit

• $\frac{1}{n}\sum \underset{\chi_F}{f}(T^j x)\, \underset{\chi_G}{g(x)} \longrightarrow f^*(x)\, g(x)$

$\Rightarrow \frac{1}{n}\sum m(T^{-j}F \cap G) \longrightarrow m(F)\; m(G)$

Which means: the probability that orbit of a pt of F falls into G converges in Cesàro sense to the product of probabilities of F & G. i.e. a moving set F tends on average to become stochastically independent of each fixed set F

• In particular: for any two positive measure sets $\exists N > 0$ s.t $\mu(T^{-N}F \cap G)$

Def Unique ergodicity: time averages $\to$ space average for every x and every cts f.

($\Leftrightarrow$ ~~Here~~ $\mathcal{M}_T(X)$ contains exactly 1 ergodic measure)

Thm Weyl's polynomial equidistribution theorem: p - polynomial - then $p(n)$ is equidistributed mod 1 if at least one of the coefficients is irrational.

Mixing:

- $\forall F, G$ of positive measure $\exists N > 0$ s.t.

$$\mu(T^{-n} F \cap G) > 0 \text{ for all } n \geq N.$$

2. $$\mu(T^{-n} F \cap G) \longrightarrow \mu(F)\mu(G)$$

i.e. "events" $T^{-n}F$ and G become independent.

2. $\forall f, g \in L^2$: $\langle U_T^n f, g \rangle \longrightarrow 0$ as $n \to \infty$

Weak mixing $\Leftrightarrow$ $\forall F, G$:

$$\lim \frac{1}{N} \sum |\mu(T^{-n} F \cap G) - \mu(F)\mu(G)| = 0$$

$$\Leftrightarrow \lim \frac{1}{N} \sum \left| \int f \circ T^n \cdot g \, d\mu - \int_X f d\mu \cdot \int g \, d\mu \right| = 0$$

for every $f, g \in L^2(X, \mu)$

$\Leftrightarrow$ 1 is simple eigenvalue of U_T and there are no other eigenvalues.