MATH5175 Ergodic Theory

Wednesday, 30 July 2014 12:03 PM

<u>Dynamical System</u>: A map $T : X \circlearrowleft$ on some space X

The dynamics is by composition $T^n = T \circ T^{n-1} = T^{n-1} \circ T$

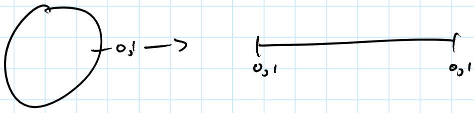
A natural qn is on asymptotic behaviour

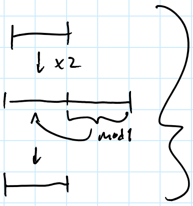
<u>Exercise 1</u>: $X = S^1$ circle w. unit circ

$T : X \circlearrowleft \qquad T(x) = x + \frac{2}{5} \pmod 1$

T is a rigid rotation by angle $\frac{2}{5}$

<u>Example 2</u>: $X = S^1 \quad T : X \circlearrowleft \quad Tx = 2x \pmod 1$





Circ. doubling map.

<u>Asymptotics</u>:

$x = 0,\ Tx = 0$ fixed.

$x = 1/2,\ Tx = 0$ fixed

$x = 1/5,\ Tx = 2/5$
$T^2x = 4/5$
$T^3x = 3/5$
$T^4x = 1/5 = x$

x has period 4

$x = \frac{1}{\pi}$, $T^n x \neq x \ \forall n \geq 1$ aperiodic (why?)

<u>Example 3</u> $X = \{0,1\}^{\mathbb{Z}^+} = \{0,1\} \times \{0,1\} \times \dots$ (binary sequences)

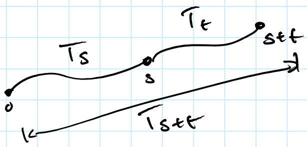
$T : X \circlearrowleft$ s.t $(T\underset{\sim}{x})_i = (\underset{\sim}{x})_{i+1}$ left shift map.

i.e. $T((x_0, x_1, x_2, \dots)) = (x_1, x_2, x_3, \dots)$

(This is a model of coin tossing where 0 is heads 1 is tails and one $\underset{\sim}{x}$ is a realisation of an infinite sequence of coin tosses!)

<u>Continuous Time Dynamics</u>: Have a family of maps T_t satisfying

$T_t \circ T_s = T_{t+s}$



Semigroup Property

The collection (X, T_t) is a flow (or semiflow if $t > 0$)

Often T_t is the solution to an ODE

$$\dot{x} = F(x) \qquad x \in X$$

Let $x(t)$ be a solution. We can write this as

$$x(t) = T_t(x(0)) \quad \text{where } x(0) \text{ is the IV.}$$

Example: $X = S^1$, $\mathcal{F}(x) = a > 0$ (constant speed)

so $\dot{x} = a \Rightarrow T_t(x) = x + at \pmod 1$

One can make a <u>discrete time system</u> from a <u>cts system</u> by stepping in units of τ: Define

$T(x) = T_\tau(x) \leftarrow$ "time τ" map of T

(X, T_τ) is the discrete dynamical system.

Example (N body prob): $X = \mathbb{R}^6 = \underset{\text{posn}}{\mathbb{R}^3} \times \underset{\text{momentum.}}{\mathbb{R}^3}$

$$\underset{\sim}{x} = (\underset{\sim}{q}, \underset{\sim}{p}) = (\underbrace{q_1, q_2, q_3}_{\text{posn}}, \underbrace{p_1, p_2, p_3}_{\text{momenta}})$$

For the i^{th} body:

$$m_i \frac{d^2 \underset{\sim}{q}_i}{dt^2} = \underbrace{\sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{G m_i m_j (\underset{\sim}{q}_j - \underset{\sim}{q}_i)}{\| \underset{\sim}{q}_j - \underset{\sim}{q}_i \|^3}}_{\text{Gravitational Interactions}} \quad ; \quad \underset{\sim}{p}_i = m_i \frac{d \underset{\sim}{q}_i}{dt}$$

$$\left(\frac{d \underset{\sim}{p}_i}{dt} \right)$$

Huuuge config. space!!!

Example 3: Turbulence in Fluids $\rightarrow$ weather/climate

V. high dimensional phase sp. X w. dynamics operating on a vast range of spatio-temporal scales.

Abstract Setup of Ergodic Theory

Defn A measure space is a triplet: $(X, \mathcal{B}, \mu)$

X: a set (Sometimes a space)
$\mathcal{B}$: a σ-algebra (Collection of subsets of X that are measurable)

$\mathcal{B}$ contains $\emptyset$, is closed under: complementation
countable unions

The elts of $\mathcal{B}$ are called "measurable sets".

$\mu: \mathcal{B} \to \mathbb{R}^{\geq 0}$ is the <u>measure</u>. It's a σ-additive fn i.e.

$\{E_i\} \in \mathcal{B}$, $E_i \cap E_j = \emptyset$ for $i \neq j$. Then

$$\mu(\cup_i E_i) = \sum_i \mu(E_i)$$

If $\mu(X) = 1$ then we call μ a <u>probability measure</u> &

$(X, \mathcal{B}, \mu)$ is a <u>probability space</u>. //

Defn: A measure preserving transformation (mpt) is a quartet $(X, \mathcal{B}, \mu, T)$ where (1) T is called measurable:

$$E \in \mathcal{B} \Rightarrow T^{-1}E \in \mathcal{B} \quad (\Leftrightarrow T^{-1}E \notin \mathcal{B} \Rightarrow E \notin \mathcal{B})$$

(T does not take non-measurable sets to measurable ones)

(2) μ is T-invariant i.e.

$$\mu(T^{-1}E) = \mu(E) \quad \forall E \in \mathcal{B} \quad \Leftrightarrow \boxed{E \in \mathcal{B} \Rightarrow \mu(T^{-1}E) = \mu(E)}$$

$$\Leftrightarrow \boxed{\mu(T^{-1}E) \neq \mu(E) \Rightarrow E \notin \mathcal{B}}$$

Some technical defns to help w. later examples:

Defn: A seq. of sets $\{A_n\}$ is increasing (resp. decreasing) if $\forall n \quad A_n \subseteq A_{n+1}$ (resp. $A_n \supseteq A_{n+1}$).

Notn/ we write: $A_n \uparrow A$ if $\{A_n\}$ increasing and $A = \bigcup_i A_i$

$A_n \downarrow A$ if $\{A_n\}$ decreasing & $A = \bigcap_i A_i$

Defn: (A monotone class) Let X be a set. A monotone class of subsets X is a collection $\mathcal{M}$ of subsets that contains all subsets of X, $\emptyset$, s.t. if $A_n \in \mathcal{M}$ (exclusive or) $\left[\begin{matrix} A_n \uparrow A \\ A_n \downarrow A \end{matrix}\right]$ then $A \in \mathcal{M}$.

Thm (Monotone Class Thm)

A monotone class $\mathcal{M}$ that contains an alg. [same def as a σ alg but closed under finite unions which is a more general notion than a σ alg], also contains the σ-alg generated by this alg.

Example: $T_\alpha : S^1 \circlearrowleft$ (unit circle)

$T_\alpha(x) = x + \alpha \pmod 1$ (α-rotn)

preserves Lebesgue measure on S^1 for each $\alpha \in \mathbb{R}$

To see this intuitively: $\ell([a,b]) = b - a \qquad 0 < a < b < 1$

$$\ell(T_\alpha^{-1}([a,b])) = \ldots \text{ (checking crossing of } 0/1)$$
$$= b - a$$
$$= \ell([a,b])$$

For any interval, T_α is measure preserving. To extend this to all of S^1 we make use of the monotone class thm:

$$\underset{\text{(monotone class)}}{\mathcal{M}} = \{E \in \mathcal{B} : \ell(T_\alpha^{-1}E) = \ell(E)\}$$

$\hookrightarrow$ contains the algebra $\mathcal{A}$ of finite disjoint unions of intervals.

algebra of a set

$\Rightarrow \mathcal{A} \subset \mathcal{M} \subset \mathcal{B}$

Exercise: Check $\mathcal{M}$ is a monotone class

Then by the thm $\mathcal{M} \supset \mathcal{B} \Rightarrow \mathcal{M} = \mathcal{B}$.

∴ T_α preserves lebesgue measure

Plan:

* Prob. view of dynamics
* Ergodicity

$(X, \mathcal{B}, \mu, T)$ prob. pres. trans (ppt).

analogues w. prob:

$X \longleftrightarrow$ Sample Sp.
$\mathcal{B} \longleftrightarrow$ Observable events
$\mu \longleftrightarrow \mathbb{P}(\cdot)$

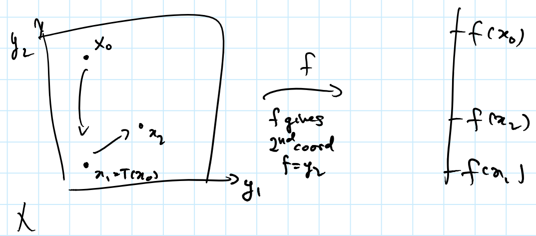
$f : X \to \mathbb{R}$ measurable, R.V.

dynamics:

$X_n = f \circ T^n$ sequence of measurements
e.g. X is the ocean
f is a measurement (temperature)

$\rightsquigarrow$ Stochastic process.

upshot: Probabilistic intuition are useful for ergodic theory.



Ergodicity:

$(X, \mathcal{B}, \mu, T)$ is a m.p.t.

Def: E is invar if $T^{-1}E = E$

Exercise: if E is invar $\Rightarrow E^c$ is invar

Defn: A mpt $(X, \mathcal{B}, \mu, T)$ is ERGODIC if every invar set E satisfies $\mu(E) = 0$ or $\mu(E^c) = 0$

in this case μ is an ergodic measure. (No nontrivial invariant sets).

Example: $X = S^1$, $R_\alpha : X \circlearrowleft$ $R_\alpha x = x + \alpha \pmod 1$

$\mu =$ lebesgue measure ℓ (length)

prop R_α ergodic $\iff \alpha \notin \mathbb{Q}$

pf/ Suppose $\alpha \in \mathbb{Q}$ then $\alpha = p/q$ for $p, q \in \mathbb{Z}$ $q \neq 0$ etc etc.

Then $(R_\alpha)^q = Id$

orbit $x,\ x+\frac{p}{q},\ x+\frac{2p}{q}, \dots, x+\frac{qp}{q} = x+p \simeq x$

So nontrivial invariant set: some ε nbhd around any x

s.t. $\underbrace{B_\varepsilon(x),\ R_\alpha(B_\varepsilon(x)), \dots, R_\alpha^{q-1}(B_\varepsilon(x))}_{\text{disjoint}}.$

Then $R_\alpha^{-1}\left(\bigcup_{j=0}^{q-1} R_\alpha^j(B_\varepsilon(x))\right) = \bigcup_{j=0}^{q-1} R_\alpha^j(B_\varepsilon(x)).$

Obviously $\mu\left(\bigcup_{j=0}^{q-1} R_\alpha^j(B_\varepsilon(x))\right)$

$= \sum_{j=0}^{q-1} \mu(R_\alpha^j(B_\varepsilon(x)))$ $\quad R_\alpha$ is meas pres.

$= \sum_{j=0}^{q-1} \mu(B_\varepsilon(x))$

$= 2q\varepsilon \neq 1, 0,$ for $\varepsilon < \frac{1}{2q}$

$\Rightarrow R_\alpha$ is not ergodic

prop/ if $\alpha \notin \mathbb{Q}$ then R_α is ergodic.

Pf/ Take $E \subseteq S^1$

which is R_α invariant. Let $f = \mathbb{1}_E$.

Write $f(x) = \sum_{n\in\mathbb{Z}} a_n e^{2\pi i n x}$ (fourier expansion)

$$\alpha_n = \int f(x) e^{-2\pi i n x}\, dx$$

$\because R_\alpha^{-1}E = E \Rightarrow f\circ R_\alpha(x) = f(x)$ (hence $a_n = b_n$)

So $f_0 \circ R_\alpha(x) = \sum_{n\in\mathbb{Z}} b_n e^{2\pi i n x}$

$b_n = \int f\circ R_\alpha(x) e^{-2\pi i n x}\, dx$

$= \int f(x+\alpha) e^{-2\pi i n x}\, dx$

$= \int f(x) e^{-2\pi i n(x-\alpha)}\, dx$

$= e^{2\pi i n \alpha} a_n$

So $b_n = a_n = e^{2\pi i n \alpha} a_n$

$a_n(1 - e^{2\pi i n\alpha}) = 0$

$\downarrow$

$a_n = 0\ \forall n$ or $e^{2\pi i n \alpha} = 1 \Rightarrow \alpha \in \mathbb{Z}$

So if $\alpha \notin \mathbb{Q}$ then it must be that $a_n = 0\ \forall n \geq 1$

$1_E(x) = f(x) = a_0$ constant

$$\hookrightarrow = \begin{cases} 1 & \Rightarrow E = S^1 \\ 0 & \Rightarrow E = \emptyset . \end{cases}$$

So $\mu(E) = 1$ or 0 $\Rightarrow R_\alpha$ is ergodic for $\alpha \notin \mathbb{Q}$

31st July 2024

Recap: A m.p.t $(X, \mathcal{B}, \mu, T)$ is <u>ergodic</u> if for every invariant $E \in \mathcal{B}$ ($T^{-1}E = E$) $\mu(E) \in \{0,1\}$. No nontrivial proper subsets that are invariant.

<u>Prop 1</u>: Let $(X, \mathcal{B}, \mu, T)$ be a m.p.t. on a complete measure space [every null set is measurable] TFAE:

(i) μ is ergodic

(ii) $f: X \to \mathbb{R}$ measurable & $f \circ T = f$ almost everywhere (a.e.) then $\exists c \in \mathbb{R}$ s.t. $f(x) = c$ for μ a.e. $x \in X$.

<u>def</u> Two mpt $(X_i, \mathcal{B}_i, \mu_i, T_i)$ $i = 1,2$ are <u>measure theoretically isomorphic</u> if $\exists$ measurable maps $\pi: X_1 \to X_2$, $\rho: X_2 \to X_1$ such that

① $\exists X_i' \in \mathcal{B}_i$ s.t $\mu(X_i - X_i') = 0$ and $\rho \circ \pi|_{X_1'} = id|_{X_1'}$ and $\pi \circ \rho|_{X_2'} = id|_{X_2'}$ (almost everywhere inverses)

② $\mu_1(\pi^{-1}E) = \mu_2(E)$ $\forall E \in \mathcal{B}_2$
$\mu_2(\rho^{-1}F) = \mu_1(F)$ $\forall F \in \mathcal{B}_1$

③ $T_2 \circ \pi = \pi \circ T_1$ on X_1'

[equiv: $T_1 \circ \rho = \rho \circ T_2$ on X_2']

$$\begin{array}{ccc} (X_1, \mathcal{B}_1, \mu_1) & \xrightarrow{T_1} & (X_1, \mathcal{B}_1, \mu_1) \\ \pi \downarrow & \circlearrowleft & \downarrow \pi \\ (X_2, \mathcal{B}_2, \mu_2) & \xrightarrow{T_2} & (X_2, \mathcal{B}_2, \mu_2) \end{array}$$

<u>Prop 2</u>: Assume $(X_i, \mathcal{B}_i, \mu_i, T_i)$ $i = 1,2$ are measure theoretically isomorphic, and one of them is ergodic, then so is the other one.

Pf/ Suppose $(X_1, \mathcal{B}_1, \mu_1, T_1)$ is ergodic. Let $f: X_2 \to \mathbb{R}$ be measurable and s.t. $f \circ T_2 = f$ μ_2-a.e. Then $f \circ T_2 \circ \pi = f \circ \pi$ μ_1-a.e. By prop 1, since μ_1 is ergodic $f \circ \pi$ is constant μ_1 a.e. Then $f \circ \pi \circ \rho = f$ μ_2-a.e. (by (i)) and $f \circ \pi \circ \rho$ is also constant a.e. Then f is constant μ_2 a.e. thus by prop $(X_2, \mathcal{B}_2, \mu_2, T_2)$ is ergodic.

$\square$

<u>def</u>: A p.p.t. $(X, \mathcal{B}, \mu, T)$ is called mixing if $\forall E, F \in \mathcal{B}$ $\mu(E \cap T^{-n}F) \to \mu(E)\mu(F)$ as $n \to \infty$.

<u>prop 3</u> Mixing implies ergodicity

Pf/ Let $(X, \mathcal{B}, \mu, T)$ be a mixing p.p.t. Suppose $E \in \mathcal{B}_1$ is T-invariant.

$\Rightarrow E = T^{-k}E$ $\forall k \geq 0$

So $\mu(E \cap T^{-n}E) = \mu(E)$, and as $n \to \infty$ this $\to \mu(E) \cdot \mu(E)$, so $\mu(E) = \mu(E)^2 \to \mu(E) = \begin{cases} 0 \\ 1 \end{cases}$

$$\downarrow^{n\to\infty} \quad \Big\} \ \mu(E)\cdot\mu(E) \Rightarrow \mu(E)\in\{0,1\}$$
$\mu(E)\cdot\mu(E)$

MAGS ARE

So μ is ergodic. $\square$

Exercise: Mixing is invariant under measure theoretic isomorphism. yes!

Prop: A p.p.t $(X_2, \mathcal{B}_2, \mu, T)$ is mixing

$\iff$

$\forall f, g \in L^2(\mu)$ $[f \in L^2(\mu) \iff \int f^2 d\mu < \infty]$

$$\int (f\cdot g)\circ T^n \, d\mu \longrightarrow \int f \, d\mu \cdot \int g \, d\mu$$

Pf/ RHS $\Rightarrow$ mixing: Let $E, F \in \mathcal{B}$ consider $f = \mathbb{1}_E$, $g = \mathbb{1}_F$

then $f\cdot g\circ T^n = \mathbb{1}_E \cdot \mathbb{1}_F(T^n x) = \begin{cases} 1 \\ 0 \end{cases}$

$= \mathbb{1}_{E\cap T^{-n}(F)}(x)$

So $\mu(E\cap T^{-n}(F)) \to \underbrace{\int \mathbb{1}_E \, d\mu}_{\mu(E)} \cdot \underbrace{\int \mathbb{1}_F \, d\mu}_{\mu(F)}$

Pf / skipped

Example Angle doubling map

$T: S^1 \circlearrowleft$ $\underset{0}{|}\!\!\longrightarrow\!\!\underset{1}{|}$ 0,1 identified.

$T(x) = 2x \pmod 1$ w. lebesgue measure μ.

We will show that T is mixing and μ is preserved.

Pf/ Define cylinder sets $\underset{\sim}{a} = [a_1, \dots, a_n] = \{x \in [0,1) \mid x = 0.a_1 a_2 \dots a_n \dots\}$ in binary, $a_i \in \{0,1\}$.

$T(0.a_1a_2\dots) = 0.a_2\dots$

$T^{-1}([a_1\dots a_n]) = [0\, a_1 \dots a_n] \cup [1\, a_1 \dots a_n]$

$l([\underset{\sim}{a}]) = 2^{-n}$, $l(T^{-1}[a_1 \dots a_n]) = 2^{-n}$. $\square$

Poincare Recurrence Thm (Gary's bit)

Suppose $(X, \mathcal{B}, \mu, T)$ is an mpt. & μ is a prob meas.
If $E \in \mathcal{B}$, then for μ a.e $x \in E$ $\exists$ a seq $n_k \to \infty$ s.t $T^{n_k} x \in E$ [i.e. $\exists F \subset E$, $\mu(F) = \mu(E)$ s.t $\forall x \in F$, $\exists$ seq $n_1 < n_2 < \dots$ w. $T^{n_i} x \in F$ $\forall i \geq 1$.]

NB not true if $\mu(X) = \infty$

$X = \mathbb{Z}$, $x \xrightarrow{T} x+1$

Pf/ For $N \geq 0$, let $E_N = \bigcup_{n=N}^{\infty} T^{-n} E$

$\hookrightarrow$ all pts that land in E after N or more steps

Put $E^\infty = \bigcap_{N=0}^{\infty} E_N$. E^∞ is the set of all pts entering E infinitely often.

$F = E \cap E^\infty$ is the set of pts re-entering E infinitely often.

Let $x \in F$, claim T^{n_i} is not only in E, but also in F.

i.e. $T^{n_i} x$ still returns infinitely often to E.

This is true because choose $n_j > n_i$ then $\underbrace{T^{n_j} x}_{\in E}$

$= T^{n_j - n_i} (T^{n_i} x)$

So $T^{n_i} x$ returns to E in $n_j - n_i$ steps. $E_0 \supseteq E_1 \supseteq \dots$

We show that $\mu(F) = \mu(E)$. Note that.

$$T^{-1}(E_N) = T^{-1}\left(\bigcup_{n=N}^{\infty} T^{-n} E \right) = \bigcup_{n=N}^{\infty} T^{-(n+1)} E$$

$$= \bigcup_{n=N+1}^{\infty} T^{-n} E = E_{N+1}$$

$$\mu(E_{N+1}) = \mu(T^{-1} E_N) \overset{\text{inv}}{=} \mu(E_n)$$

$\therefore$ $\mu(E_{N+1}) = \mu(E_0)$ by induction, $\forall N \geq 0$.

Exercise

$$\mu(E^\infty) = \mu(E_0)$$

Thus

$$\mu(F) = \mu(E \cap E^\infty) = \mu(E \cap E_0) = \mu(E)$$

since $E \subset E_0$

Showing that F is the same as E up to measure zero.

Bernoulli Schemes

Let S be a finite set called the <u>alphabet</u> & $X = S^{\mathbb{N}}$ (sequences of S indexed by $\mathbb{N}$). Impose a metric on X:

$$d(\underset{\sim}{x}, \underset{\sim}{y}) = 2^{-\min\{k : x_k \neq y_k\}} \quad \text{for} \quad \begin{matrix} \underset{\sim}{x} = (x_i)_{i \in \mathbb{N}} \\ y = (y_i)_{i \in \mathbb{N}} \end{matrix}$$

The resulting topology is generated by <u>cylinders</u>

$$[a_0 a_1 \dots a_{n-1}] = \{ \underset{\sim}{x} \in X \mid x_i = a_i \text{ for } i = 0, \dots, n-1 \}$$

The left shift is the transformation $\sigma(\underset{\sim}{x}) = (x_{i+1})_{i \in \mathbb{N}}$

We shall construct an invariant measure on this space

Fix a prob vector $p = (p_i)_{i \in S}$ (vector of non neg numbers that sum to 1)
i ?? wtf

<u>Defn</u> The <u>Bernoulli measure</u> (μ) corresp. to p is the unique measure on the Borel σ alg of X s.t

$$\mu([a_0 \dots a_{n-1}]) = p_{a_0} \cdots p_{a_{n-1}}$$

[This is a generic scheme for indep events like a multi-sided unfair coin toss sequence]

Prop/ Let $S = \{0, 1\}^{\mathbb{N}}$, μ be the $(\frac{1}{2}, \frac{1}{2})$ Bernoulli measure and σ the left shift transformation. Then $(X, \mathcal{F}, \mu, \sigma)$ is thre measure theoretic isom to the circle doubling map $(S^1, \mathcal{B}, l, T)$.

Pf/ We seek $\pi : X \to S^1$

Define $\pi(\underset{\sim}{x}) = \sum_{n \geq 0} x_n / 2^{n+1}$ this will be our ergodic isom.

We confirm this:

π is an a.e bijection:

$$\begin{tikzcd} (X, \mathcal{F}, \mu) \arrow[r, "\sigma"] \arrow[d, "\pi"'] & (X, \mathcal{F}, \mu) \arrow[d, "\pi"] \\ (S^1, \mathcal{B}, l) \arrow[r, "T"'] & (S^1, \mathcal{B}, l) \end{tikzcd}$$

$$l(E) = \mu(\pi^{-1} E) \quad \forall E \in \mathcal{B}$$

$= \pi^* \mu(E)$ Notn? WTF

$X' = \{ x \in X : \nexists\, n \text{ s.t. } x_m = 1 \ \forall m \geq n \}$ (Do this to maintain 1-1ness since $d((0.111\dots), (1.000)) = 0$)

(No trailing 1s)

$\pi([0])$ is $[0, 1/2)$ $\quad$ $[0] \to (0, \dots)$ $\quad$ (think of this as binary reps π : binary $\to [0, 1)$)

$\pi([1])$ is $[1/2, 1)$

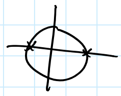
$\pi([01])$ is $[1/4, 1/2)$

[So we can partition $X = [00] \cup [01] \cup [10] \cup [11]$ etc.]

Binary representations are 1-1 w. real numbers

[a.e. bit is due to trailing 1s which is a countable set]

Now for the dynamics



$$\sigma^{-1}([0]) = \sigma^{-1}(0 \dots)$$

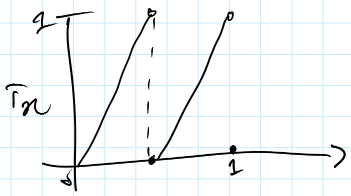
$$= [00] \cup [10]$$

$$T^{-1}([0,\tfrac{1}{2})) = [0,\tfrac{1}{4}) \cup [\tfrac{1}{2},\tfrac{3}{4}) =$$

$$[0,1/2) \cap T^{-1}[0,1/2) \qquad [\tfrac{1}{2},1) \cap T^{-1}[\tfrac{1}{2},1)$$

In general,

$\pi([a_0 \dots a_{n-1}])$ Break Here



we show diagram commutes

$$\begin{tikzcd}
(X, \mathcal{F}, \mu) \arrow[r, "\sigma"] \arrow[d, "\pi"'] & (X, \mathcal{F}, \mu) \arrow[d, "\pi"] \\
(S^1, \mathcal{B}, \ell) \arrow[r, "T"] & (S^1, \mathcal{B}, \ell)
\end{tikzcd}$$

Pf: $T \circ \pi = \pi \circ \sigma \quad \forall x \in X$

$$T(\pi(x_0 x_1 \dots)) = T\left(\sum_{n \geq 0} \frac{x_n}{2^{n+1}}\right) = \sum_{n \geq 0} \frac{x_n}{2^n} \mod 1$$

$$= \sum_{n \geq 1} \frac{x_n}{2^n}$$

$$\pi(\sigma(x_0 x_1 \dots)) = \pi(x_1 x_2 \dots) = \sum_{i \geq 1} x_i / 2^i$$

$\square$

Now

$$\pi([0]) = [0, 1/2)$$

$$\pi([1]) = [\tfrac{1}{2}, 1)$$

$$\pi([01]) = [\tfrac{1}{4}, \tfrac{1}{2}) = \underbrace{[0, \tfrac{1}{2})}_{A_0} \cap T^{-1} \underbrace{[1/2, 1)}_{A_1}$$

$$\pi([a_0 \dots a_{n-1}]) = \bigcap_{k=0}^{n-1} T^{-k} A_{a_k}$$

$$\left.\begin{array}{l} x \in A_{a_0} \\ \vdots \\ T^{n-1} x \in A_{a} \end{array}\right\} x \in$$

$\vdots$

$T^{n-1}x \in A_{a_{n-1}}$ $\Big\}$ $x \in$ ⌟

$\ell(E) = \mu(\pi^{-1}E) \quad \forall E \in \mathcal{B}$

$\mu(\pi^{-1}(\cap_{k=0}^{n-1} T^{-k} A_{a_k})) = \mu[a_0 \dots a_{n-1}] = \frac{1}{2^n}$

$\ell(\cap_{k=0}^{n-1} T^{-k} A_{a_k}) = \frac{1}{2^n}$ (by some halfway argument)

Stop here at the level of cylinders, use monotone class theorem to extend to all of the borel set!

(1) π bijection almost everywhere

(2) π, π^{-1} were measurable (todo this, show ~~continuity~~ wrt metric --- then?

cylinder open set

σ maps cylinder

pull back open set to open set

open sets generate next σ alg

$\Rightarrow$ nearly open sets to open sets

σ algs match up!)

(3) $T \circ \pi = \pi \circ \sigma$

(4) $\ell = \mu \circ \pi^{-1}$ mostly.

Pf/ Doubling map is mixing

(The $(\frac{1}{2}, \frac{1}{2})$ Bernoulli shift is mixing by isomorphism)

Represent $x \in S^1$ as binary expansions

$T(0.d_0 d_1 d_2 \dots) = 0.d_1 d_2 \dots$

$[d_0 \dots d_n] = \{ x \in [0,1), x = 0.d_0 \dots d_n \varepsilon_{n+1} \varepsilon_{n+2} \dots \}$

where $\varepsilon_i \in \{0, 1\}$

Exercise: Show Lebesgue measure is preserved by T for cylinders

Mixing $f: S^1 \to \mathbb{R}$, $g: S^1 \to \mathbb{R}$, T is mixing if

$$\int f \cdot g \circ T^k \, d\ell \to \int f \, d\ell \cdot \int g \, d\ell \qquad k \to \infty$$

Interp Correl?

$f = \mathbb{1}_{[a_0 \cdots a_{n-1}]}$, $g = \mathbb{1}_{[b_0 \cdots b_{m-1}]}$

$$\int f \cdot (g \circ T^k)\, d\ell = \int \mathbb{1}_{[a_0 \cdots a_{n-1}]} \cdot \mathbb{1}_{[b_0 \cdots b_{m-1}]} T^k\, d\ell \quad \left(\begin{array}{l} \underline{\text{Aside}}: \\ \mathbb{1}_E \circ T \\ = \mathbb{1}_{T^{-1}E} \end{array} \right)$$

$$= \int \mathbb{1}_{[a_0 \cdots a_{n-1}]} \cdot \mathbb{1}_{T^{-k}[b_0 \cdots b_{m-1}]}\, d\ell$$

$$= \int \mathbb{1}_{[a_0 \cdots a_{n-1}]} \cdot \mathbb{1}_{[\underbrace{\underbrace{? \cdots ?}_{k} b_0 \cdots b_{m-1}]}}\, d\ell \quad \left(\begin{array}{l} \underline{\text{Aside}} \\ \mathbb{1}_E \cdot \mathbb{1}_F = \mathbb{1}_{E \cap F} \end{array} \right)$$

For $k \geq n$

$$= \int \mathbb{1}_{[a_0 \cdots a_{n-1}] \cap [\quad\quad]}\, d\ell$$

$$= \int \mathbb{1}_{[a_0 \cdots a_{n-1} \underbrace{? ? \cdots ?}_{k-n} b_0 \cdots b_{m-1}]}\, d\ell$$

$$= \ell([a_0 \cdots a_{n-1} \underbrace{? \cdots ?}_{k-n} b_0 \cdots b_{m-1}]) \quad \left(\begin{array}{l} \text{first} \\ n+m \\ \text{positions} \\ \text{specified} \end{array} \right)$$

$$= 2^{-(n+m)} = 2^{-n} \cdot 2^{-m}$$

$$= \ell([a_0 \cdots a_{n-1}])\, \ell([b_0 \cdots b_{m-1}])$$

$$= \int \mathbb{1}_{[a_0 \cdots a_{n-1}]}\, d\ell \int \mathbb{1}_{[b_0 \cdots b_{m-1}]}\, d\ell$$

mixing for <u>cylinders</u> (infact, exactly equal after finite steps)

Thus

$\int f \cdot (g \circ T^k)\, d\ell$ can be approximated w. arb. pres. by integral of simple fns [Every L^2 fn can be approx. by lin comb. of char fns].

(finite)
<u>Markov Chains:</u> The Bernoulli shifts are models for the simplest of stochastic processes (i.i.d). Next step up in complexity are MCs.

<u>Subshifts of Finite type</u>

Let S be a finite set and $A = (t_{ij})_{i,j \in S}$ a matrix of zeros and ones w. no col/row entirely 0.

Defn The subshift of finite type w. alphabet S & transition matrix A is the set

$$\Sigma_A^+ = \{ \underset{\sim}{x} = (x_i)_{i\in\mathbb{N}} \in S^{\mathbb{N}} \mid t_{x_n x_{n+1}} = 1 \ \forall n \}$$

eg $S = \{e,f\}$ $A = \begin{array}{c} \\ e \\ f \end{array}\!\!\begin{array}{c} \begin{array}{cc} e & f \end{array} \\ \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \end{array}$ $\left(\begin{array}{c} f \\ \text{from} \end{array} \overline{\quad\text{to}\quad} \right)$

Sequences have structure ($\to$ denotes "follows") i.e.
$e \to e$ ✓
$e \to f$ ✓
$f \to e$ ✓
$f \not\to f$

e.g. eeefe...
but not eeff...
efe

Impose metric:

$$d(\underset{\sim}{x}, \underset{\sim}{y}) = 2^{-\min\{k \,:\, x_k \neq y_k\}}$$

Exercise: Σ_A^+ is a compact metric sp. (the left shift is cts).

One can think of Σ_A^+ is the space of all infinite paths on a directed graph w. Vertices given by S and A is the adjacency matrix.

Defn: A stochastic matrix is a matrix $P = (P_{ab})_{a,b\in S}$ w.

i) Non neg entries

(ii) row sums of 1 i.e. $\sum_{j\in S} P_{ij} = 1 \quad \forall i \in S$

i.e. $P\,1_{vect} = 1_{vect}$

The matrix P is compatable with A if $A_{ij} = 0 \Rightarrow P_{ij} = 0$.

(2) A prob. vector $p = (p_a)_{a\in S}$ is a nonneg vect. s.t. $\sum_{a\in S} p_a = 1$

(3) A stationary p vect is a p-vect p s.t. $pP = p$

i.e. $\sum_{i\in S} p_i P_{ij} = p_j \quad \forall j \in S$. (Analogy M, T)

Given a prob vect p and a stoch matrix P compat. w. A

one can define a Markov Measure μ on Σ_A^+ by

$$\mu([a_0 \dots a_{n-1}]) = p_{a_0} P_{a_0 a_1} P_{a_1 a_2} \cdots P_{a_{n-2} a_{n-1}}$$

Again, as for Bernoulli measures, cylinders generate the Borel σ-alg on Σ_A^+ and one can construct a measure μ extending the above.

Propn 1. μ is shift invariant i.e. $(\Sigma_A^+, \mathcal{F}, \mu, \sigma)$ is an mpt,

$\iff$

p is stat. wrt P.

2. Any stoch matrix has a stat prob. vect.

(So we always have atleast 1 invar μ)

Pf/ μ is σ-inv if $\mu(\sigma^{-1}E) = \mu(E)$, $E \in \mathcal{F}$. Let $E = [b_0 \dots b_{n-1}]$ then

$$\mu(\sigma^{-1}E) = \mu([* \, b_0 \dots b_{n-1}]) \quad \text{and}$$

$$\mu(E) = p_{b_0} P_{b_0 b_1} \cdots P_{b_{n-2} b_{n-1}} \qquad (1)$$

Now

$$\mu([* \, b_0 \dots b_{n-1}]) = \sum_{i \in S} p_i P_{i b_0} P_{b_0 b_1} \cdots P_{b_{n-2} b_{n-1}} \qquad (2)$$

we want (1) = (2)

$$(2) = P_{b_0 b_1} \cdots P_{b_{n-2} b_{n-1}} \sum_{i \in S} p_i P_{i b_0} = p_{b_0} P_{b_0 b_1} \cdots P_{b_{n-2} b_{n-1}} = (1)$$

happy cancellation $\Rightarrow$ $p_{b_0} = \sum p_i P_{i b_0}$

$\iff$ p_{b_0} is stationary. ??

(Pf 2) $\Delta = \{ (x_1 \dots x_n) \in \mathbb{R}^n \mid x_i \geq 0 \ \sum_{i=1}^N x_i = 1 \}$

$T : \Delta \to \mathbb{R}^n$ $\quad Tx := xP$ $\quad$ Since P is stoch, claim

$T(\Delta) \subset \Delta$: for $x \in \Delta$,

$$\sum_{i=1}^N (Tx)_i = \sum_{i=1}^N \sum_{j=1}^N x_j P_{ij} = \sum_{j=1}^N x_j \underbrace{\sum_{i=1}^N P_{ij}}_{1 \text{ stoch mat}}$$

$$= \sum_{j=1}^N x_j = 1$$

By Brouwer's f.p. thm (Ex T is a cts mapping on a closed convex set, namely Δ)

T has a fp which is the stat. prob vect. $\square$

Thus, any stoch mat determines atleast one shift inv prob measure μ: is μ mixing/ergodic?

Ergodicity & Mixing

Suppose A is a transition matrix. We say a connects b in n steps and write $a \xrightarrow{n} b$ if there is a path of length n on the directed graph representing Σ_A^+ starting at a and ending at b. In terms of A, this means $\exists$ states $b_1 \dots b_{n-1}$ s.t. $A_{ab_1} A_{b_1 b_2} \cdots A_{b_{n-1} b} > 0$.

Defn/ A transit. mat. is irreducible if for any $a, b \in S$, $\exists\, n$ s.t. $a \xrightarrow{n} b$.

Defn The period of an irred trans. mat. is the number $p := \gcd\{n : a \xrightarrow{n} a\}$ $\left(\begin{smallmatrix}\text{indep}\\ a\end{smallmatrix}\right)$

An irred. trans mat. is aperiodic if $p = 1$.

Example $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ irred w. period 2

$$1 \rightleftarrows 2$$

Thm: (Ergodic Thm of Markov Chains)
Suppose P is a stoch mat w. stationary prob vect p

(1) if P is irred:

$$\frac{1}{n}\sum_{k=0}^{n-1} (P^k)_{ij} \longrightarrow p_j$$

(2) if P is irred & aperiodic, then

$$(P^k)_{ij} \longrightarrow p_j \quad \forall j \in S$$

Cor: A shift inv markov measure for a subshift of Finite type (SFT) Σ_A^+ is:

(i) Ergodic $\iff$ irred. A

(ii) Mixing $\iff$ A is irred & aperiodic.

Pf of (ii):

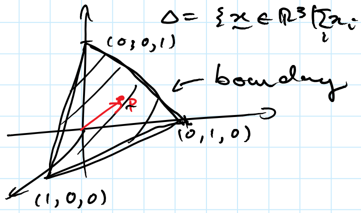
P irred & aperiodic $\Rightarrow \exists\, m$ s.t. $P^m > 0$

(See footnote in Sarig)

Let $N = |S|$, Δ as before We saw previously that

$T(x) = xP$ maps the compact, convex Δ ctsly into itself and let $p = pP$

[Eg



$\Delta = \{ \underset{\sim}{x} \in \mathbb{R}^3 \mid \sum_i x_i = 1 \}$ = 2D plane $x_1 + x_2 + x_3 = 1$ truncated as in boundary]

Consider $C = \Delta - p$

---

13 Aug

$\Delta = \{ \underset{\sim}{x} \in \mathbb{R}^n \mid x_i \geq 0, \sum_{i=1}^N x_i = 1 \}$

$Tx := xP \qquad T : \Delta \circlearrowleft$

$C := \Delta - p$ (Closed, compact, convex)

Pick an L.I set $v_1, \dots, v_{N-1} \in C$ and generate $L = sp\{v_i\}_{i=1,\dots,N-1}$

WTS $P|_L$ all has eigenvalues of modulus < 1

---

(1) All eigvals of $P|_L$ are not > 1 in modulus

$\|xP\| \leq \|x\|$, so $r_\sigma(P|_L) \leq 1$ (Spectral radius r_σ)

(2) Roots of unity are impossible: If $e^{2\pi i/k}$ were an eigval of P, then $e^{2\pi i m/k}$ would be an eigval of P^m, and 1 would be an eigval of P^{km}. Assoc. to this eigval 1 is an eig vect $y \in \mathbb{R}^N$ and we can normalise y to be in ∂C. But $P^{km} > 0$ so P^{km} cant be fixed points on ∂C

$[T^{km}(C) \subset \operatorname{int}(C)]$

(3) eigvals $e^{i\theta}$, $\theta \in 2\pi\mathbb{Q}$ are impossible

If $e^{i\theta}$ is an eigval of P, then $\exists$ 2 real vectors $u, v \in \partial C$ s.t. the action of P on $sp\{\underset{\sim}{u}, \underset{\sim}{v}\}$ is conjugate to

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$ is an irrat rotn.

So $\exists$ seq $k_n \to \infty$ s.t. $\underset{\sim}{u} P^{mk_n} \to \underset{\sim}{u} \in \partial C$.

Again, this is impossible since P^m maps C into $\operatorname{int}(C)$. and by compactness can't intersect C.

$\therefore$ $P|_L$ has spec. rad. < 1

Thm P is a stoch mat.

(Part 2) P is irred & aperiodic then

$$(P^k)_{ij} \to p_j \quad \text{as } k \to \infty \quad \forall i,j \in S.$$

We can write

$$\mathbb{R}^N = \mathrm{sp}\{p\} \oplus L$$

For any $y \in \mathbb{R}^N$, decompose as $y = tp + q$, $q \in L$.

Then $y P^n = (tp+q) P^n = tp + qP^N \to tp$.

So then $(P^k)_{ij} \to p_j$ as $k \to \infty$.

Finally, we show that p is the unique stationary vect.

Suppose $\exists\, q \neq p$ w. $qP = q$. We know

$(P^n)_{ij} \to p_j$, $\forall i,j \in S$ so define $Q_{ij} = p_j$

$\Rightarrow (P^n)_{ij} \to Q_{ij} \quad \forall i,j \in S.$

So stationarity implies

$$q = qP^n \to qQ \quad \text{so} \quad q = qQ$$

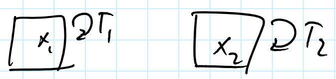
But $(qQ)_j = \sum_{i=1}^{N} q_i Q_{ij} = \sum_{i=1}^{N} q_i p_j = p_j.$

$\Rightarrow q = p.$

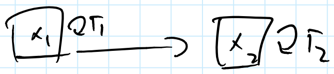
New Section

Products:

Suppose we have 2 dynamical sys that we want to combine

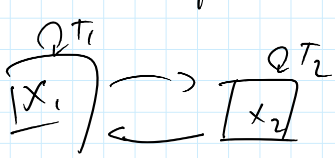


* We can combine by cartesian product (Sys don't interact)
  (called a skew product)
* We can couple one to the other



(Skew product)

* We can couple in both directions



Products:

-

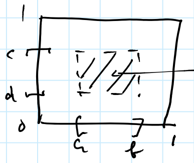
Recall that the product of 2 meas. sp. $(X_i, \mathcal{B}_i, \mu_i)$ $i=1,2$ is the measure sp

$$(X_1 \times X_2, \mathcal{B}_1 \otimes \mathcal{B}_2, \mu_1 \times \mu_2)$$

where $\mathcal{B}_1 \otimes \mathcal{B}_2$ is the smallest σ alg containing all sets of the form $B_1 \times B_2$ (i.e. its the σ alg generated by $\mathcal{B}_1 \times \mathcal{B}_2$). $\mu_1 \times \mu_2$ is the unique (not proved) measure s.t.

$$(\mu_1 \times \mu_2)(B_1 \times B_2) = \mu_1(B_1)\,\mu_2(B_2).$$

eg: $X_1 = [0,1] = X_2$



$(\mu_1 \times \mu_2)([a,b] \times [c,d])$
$= \mu([a,b])\,\mu([c,d])$
$= (b-a)(d-c)$
$=$ area!

Rem: Probabalistic interp: The products capture the notion of independence of stoch processes!
if $(X_i, \mathcal{B}_i, \mu_i)$ $i=1,2$ are prob. models of 2 random experiments that are independent, then $(X_1 \times X_2, \mathcal{B}_1 \otimes \mathcal{B}_2, \mu_1 \times \mu_2)$ is the model for the joint experiment.
So if $E_1 \in \mathcal{B}_1$, $F_1 = E_1 \times X_2$ is the event that "E_1 happened, outcome of 2nd experiment is ab" etc.

Defn/ The prod of 2 mpts $(X_i, \mathcal{B}_i, \mu_i, T_i)$ $i=1,2$ is the mpt $(X_1 \times X_2, \mathcal{B}_1 \otimes \mathcal{B}_2, \mu_1 \times \mu_2, T_1 \times T_2)$
where

$$(T_1 \times T_2)(x_1, x_2) = T(x_1) \times T(x_2)$$

(NB: $\times$ is cartesian prod).

Ex: Check that $T_1 \times T_2$ preserves $\mu_1 \times \mu_2$ for sets of the form $E_1 \times E_2$, $E_i \in \mathcal{B}_i$.

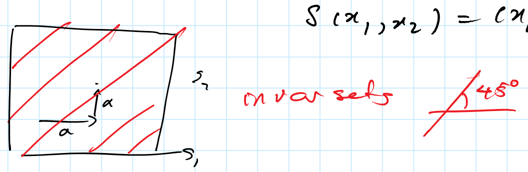
Propn: (1) The product of 2 ergodic mpts is not necessarily ergodic.
(2) The product of 2 mixing mpts is always mixing
(See sarig's notes for pf)

eg(1): $S: T_\alpha \times T_\alpha : S^1 \times S^1$, $\alpha \notin \mathbb{Q}$

$S(x, y) = (x+\alpha, y+\alpha)$ [illegible]

Eg (1) : $S : T_\alpha \times T_\alpha : S^1 \times S^1$, $\alpha \notin \mathbb{Q}$

$S(x_1, x_2) = (x_1 + \alpha, x_2 + \alpha) \pmod 1$.



An invariant, non constant function f s.t.

$f : S^1 \times S^1 \to \mathbb{R}$ s.t. $f \circ S$ is

$f(x, y) = x - y \pmod 1$

$f \circ S(x_1, x_2) = f(x_1 + \alpha, x_2 + \alpha) = ((x_1 + \alpha) - (x_2 + \alpha)) \bmod 1$

$= x_1 - x_2 \bmod 1$.

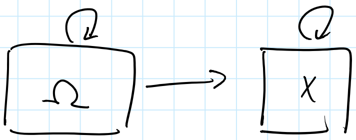
another f is

$f(x, y) = 1_A$ A is any $45°$ band.

Exercise: T_α, T_β are s.t. α, β not rationally related.

Skew products

Models of forced/driven/time dependent/random/non autonomous dynamics.



Let $(\Omega, \mathcal{F}, p)$ be a prob sp. and σ is an mpt preserving p.

σ is the driving dynamics. Let $(X, \mathcal{B}, \mu)$ be another prob sp. and $\{T_\omega\}_{\omega \in \Omega}$ a family of maps on X.

Easy case is where for each $\omega \in \Omega$,

$(X, \mathcal{B}, \mu, T_\omega)$ is an mpt.

Define the skew product

$$(\Omega \times X, \mathcal{F} \otimes \mathcal{B}, \nu, \tau)$$

where

$$\tau(\omega, x) = (\sigma\omega, T_\omega x)$$

$$\tau^2(\omega, x) = \tau(\sigma\omega, T_\omega x)$$

$$= (\sigma^2\omega, T_{\sigma\omega} \circ T_\omega x)$$

$$T^n(\omega, x) = (\sigma^n\omega, T_{\sigma^{n-1}\omega} \circ \cdots \circ T_\omega x)$$

The skew product is sometimes called a random dynamical system.

Defn The base $(\Omega, \mathcal{B}, p, \sigma)$ is the driving system, the "noise".

Examples

(1) $\Omega = \Sigma_2^+ = \{0,1\}^{\mathbb{N}}$, p is the Bernoulli $(\frac{1}{2}, \frac{1}{2})$ measure

$X = \mathbb{Z}$, $T_\omega = x + (-1)^{\omega_0}$.

(Model of a random walk)

$\omega_i := (\sigma^i\omega)_0$

$$T^n(\omega, x) = (\sigma^n\omega, x + \sum_{i=0}^{n-1} (-1)^{\omega_i})$$

Can write this as a stoch proc.

$x + \sum_{i=0}^{n-1} X_i \qquad X_i = (-1)^{\omega_i}$ and say that the $\{X_i\}$ are an indep random walk on $\mathbb{Z}$.

Ex 2 Driven fluid flow.

Ω is some configuration space or noise model X is some compact subset $\mathbb{R}^3$, the atmos, the ocean ...

σ updates the driving system. T_ω describes the fluid evolution over some time unit when the current driving config is ω.

Ex 3 Noisy rotn

$\Omega = S^1 \qquad \sigma : S^1 \circlearrowleft \qquad \sigma(\omega) = \omega + \alpha \pmod 1) \quad \alpha \notin \mathbb{Q}$

$X = S^1 \qquad T_\omega : S^1 \circlearrowleft \qquad T_\omega(x) = x + \beta + 2\varepsilon(\omega - \frac{1}{2})$ (mod 1)

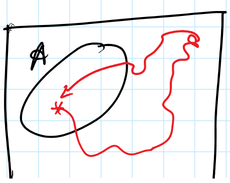
$\beta \notin \mathbb{Q}$

$\varepsilon > 0$ small.

Gary's bit

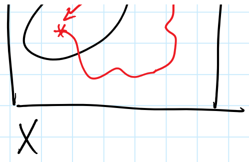
(1) An application of the Ergodic thm of Markov Chains to the Internet Search.

(2) Induced transformations.



A: set of places in X where we can measure things

We want to define a dynamical system on A.



We want to define a dynamical system on A.

Let $(X, \mathcal{B}, \mu, T)$ be a ppt let $A \in \mathcal{B}$ be such that $\mu(A) > 0$. By Poincaré recurrence thm, for a.e $x \in A$, $\exists n \geq 1$ s.t $T^n x \in A$. Let

$$\varphi_A(x) = \min\{n \geq 1 : T^n x \in A\} \quad [\varphi_A(x) = \infty \text{ if no such } n \text{ exists}]$$

<u>defn</u> The induced transformation on A is $(A_0, \mathcal{B}_A, \mu_A, T_A)$

where $A_0 = \{x \in A : \varphi_A(X) < \infty\}$

$$\mathcal{B}_A = \{E \cap A : E \in \mathcal{B}\}$$

$$T_A(x) = T^{\varphi_A(x)}(x)$$

$$\mu_A(E) = \frac{\mu(E \cap A)}{\mu(A)} = \mu(E \mid A)$$

<u>Thm</u> suppose $(X, \mathcal{B}, \mu, T)$ is a ppt, let $A \in \mathcal{B}$, $\mu(A) > 0$ then

(I) μ_A is invar under T_A $(\mu_A \circ T_A^{-1} = \mu_A)$

(II) If T is ergodic then so is T_A

<u>Remark</u>: T_A is not always mixing if T is mixing

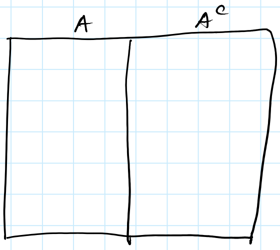
(III) <u>Kac's formula</u>: if μ is ergodic then

$$\int f \, d\mu = \int_A \sum_{k=0}^{\varphi_A - 1} f \circ T^k \, d\mu$$

In particular, the expected return time to A is $\frac{1}{\mu(A)}$

[follows from choosing $f = 1$].

<u>Pf of (I)</u>: <u>Let $E \subset A$</u>, $E \in \mathcal{B}$



$$T^{-1}E = \underbrace{[(T^{-1}E) \cap A]}_{[\varphi_A(E) = 1]} \,\dot{\cup}\, [(T^{-1}E) A^c]$$

$$T^{-1}(T^{-1}E \cap A^c) = T^{-2}E \cap T^{-1}A^c$$

$$= (T^{-1}E \cap T^{-1}A^c \cap A) \cup (T^{-1}E \cap T^{-1}A^c \cap A^c)$$

$$\mu(E) = \mu(T^{-1}E) = \mu(T^{-1}E \cap [\varphi_A = 1]) + \mu(T^{-1}E \cap A^c)$$

$$= \mu(T^{-1}E \cap [\varphi_A = 1]) + \mu(T^{-2}E \cap [\varphi_A = 2]) + \mu(T^{-2}E \cap T^{-1}A^c \cap A^c)$$

$$= \cdots$$

$$= \sum_{j=1}^{n} \mu(T^{-j} E \cap [\varphi_A = j]) + \mu(T^{-n} E \cap \bigcap_{j=0}^{n-1} T^{-j} A^c)$$

in particular,

$$\mu(E) \geq \sum_{j=1}^{\infty} \mu(T^{-j} E \cap [\varphi_A = j])$$

$T_A^{-1}(E) = \bigcup_{j=1}^{\infty} T^{-j}(E \cap [\varphi_A = j])$ disjoint union.

$\mu(T_A^{-1} E) = \sum_{j\geq 1} \mu(T^{-j}(E \cap [\varphi_A = j]))$

$\mu(E) \geq \mu(T_A^{-1} E)$

So $\mu_A(E) \geq \mu_A(T_A^{-1} E)$

Arguing similarly, for $A \setminus E$

$\mu_A(A \setminus E) \geq \mu_A(T_A^{-1}(A \setminus E))$

$\Rightarrow \mu_A(A) - \mu_A(E) \geq \mu_A(T_A^{-1} A \setminus T_A^{-1} E)$

$= \mu_A(A) - \mu_A(T_A^{-1} E)$

$\Rightarrow \mu_A(T_A^{-1} E) \geq \mu_A(E)$

Combining inequalities, we get equality so μ_A is T_A invar.

(ii) Pf of Ergodicity: Suppose T is ergodic. WTS T_A is ergodic.

Let $S = \{x \in X : T^n x \in A$ for infinitely many $n \geq 0\}$.

Then $\mu(S) \geq \mu(A) > 0$ by the Poincaré recur. thm.

S is T invar. (obviously). By ergodicity

$\mu(S) = 1$ (can't be 0 since $\mu(S) > 0$)

Thus,

$$r_A(x) := \min\{n \geq 0 \mid T^n x \in A\} < \infty \quad \mu\text{. a.e.}$$

Suppose $f: A \to \mathbb{R}$ is T_A-invar. Let $F(x) = f(T^{r_A(x)} x)$

(well defined μ a.e.)

$$r_A(Tx) = \begin{cases} r_A(x) - 1 & r_A(x) \neq 0 \\ \varphi_A(x) - 1 & r_A(x) = 0,\ r_A(Tx) \neq 0 \\ \varphi_A(x) + \varphi_A(Tx) - 1 & r_A(x) = 0,\ r_A(Tx) = 0 \end{cases}$$

$$F(Tx) = f(T^{r_A(Tx)} Tx)$$

$$= \begin{cases} f(T^{r_A(x)} x) = F(x) \\ f(T^{\varphi_A(x)} x) = f(T_A x) = f(x) = F(x) & \text{since } f \text{ is } T_A \text{ invar. } (x \in A) \\ f(T^{\varphi_A(x) + \varphi_A(Tx)} x) = f(T_A^2 x) = f(x) = F(x) \end{cases}$$

So shown that F is T invar.

So F is constant a.e.

So if $x \in A$, $F(x) = f(T^0 x) = f(x)$

so f is a.e const

Then T_A is ergodic $\blacksquare$.

20th Aug 2014 von Neumann's L^2 (mean) Ergodic Theorem (1931)

Suppose $(X, \mathcal{B}, \mu, T)$ is a p.p.t. if $f \in L^2(X, \mu)$ then

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} f\circ T^k \to \bar{f} \text{ in } L^2 \text{ where } \bar{f} \text{ is } T\text{-invariant} \quad (*)$$

If T is ergodic, then $\bar{f} = \int f \, d\mu$ μ.a.e

Pf/ Observe that T p.p.t $\Rightarrow$ $\|f\circ T\|_2 = \|f\|_2$ $\forall f \in L^2$:

[$f = \mathbb{1}_A$ then $\|\mathbb{1}_A \circ T\|_2 = \left(\int (\mathbb{1}_A \circ T)^2 \, d\mu\right)^{1/2} = \left(\int \mathbb{1}_{T^{-1}A} \, d\mu\right)^{1/2}$

$= \sqrt{\mu(T^{-1}A)} = \sqrt{\mu(A)} = \dots = \|f\|_2$.

Now approximate arb $f \in L^2$ w. simple functions from below w. simple functions etc...]

We split L^2 into 2 pieces: $\mathcal{C} \subseteq L^2$ where $\mathcal{C} = \{f \in L^2 \mid f = g - g\circ T, \text{ some } g \in L^2\}$

chet $f \in \mathcal{C}$, then note that

$$\left\|\frac{1}{N}\sum_{k=0}^{N-1} f\circ T^k - 0\right\|_2 = \left\|\frac{1}{N}(f + f\circ T + f\circ T^2 + \dots)\right\|_2$$

$$= \left\|\frac{1}{N}\left((g - \cancel{g\circ T}) + (\cancel{g\circ T} - \cancel{g\circ T^2}) + \dots + (\cancel{g\circ T^{N-1}} - g\circ T^N)\right)\right\|_2$$

$$= \left\|\frac{1}{N}(g - g\circ T^N)\right\|_2$$

$$= \frac{1}{N}\|g - g\circ T^N\|_2 \leq \frac{1}{N}(\|g\| + \|g\circ T^N\|)$$

$$= \frac{1}{N}\cdot 2\|g\| \to 0 \text{ as } N\to\infty$$

So $\bar{f} = 0$ when $f \in \mathcal{C}$.

Recall that $L^2(X,\mu)$ is a Hilbert space w. $\mathcal{C}$ as a linear subspace.

We can write $L^2(X,\mu) = \bar{\mathcal{C}} \oplus \mathcal{C}^\perp$.

We prove the L^2 Ergodic theorem holds in each subsp individually:

$\forall f \in \bar{\mathcal{C}}$, given $\varepsilon > 0$, $\exists F \in \mathcal{C}$ s.t $\|f - F\|_2 < \varepsilon$

Choose N_0 s.t $\forall N > N_0$, $\left\|\frac{1}{N}\sum_{k=0}^{N-1} F\circ T^k\right\| < \varepsilon$.

Then $\forall N > N_0$, $\|\frac{1}{N}\sum_{k=0}^{N-1} f\circ T^k\| \le \|\frac{1}{N}\sum_{k=0}^{N-1}(f-F)\circ T^k\| + \|\frac{1}{N}\sum_{k=0}^{N-1} F\circ T^k\|_2$

$$< \frac{1}{N}\sum_{k=0}^{N-1}\|(f-F)\circ T^k\| + \varepsilon$$
$$= \frac{1}{N}\sum_{k=0}^{N-1}\|f-F\| + \varepsilon < \varepsilon + \varepsilon$$

Thus $\lim_{N\to\infty}\|\frac{1}{N}\sum_{k=0}^{N-1} f\circ T^k - 0\| = 0$ for $f \in \overline{\mathcal{C}}$.

Now for $\mathcal{C}^\perp$: We show that f is invariant: Let $f \in \mathcal{C}^\perp$

$\|f - f\circ T\|_2^2 = \langle f - f\circ T, f - f\circ T\rangle = \|f\|_2^2 - 2\langle f, f\circ T\rangle + \|f\circ T\|_2^2$

$= 2\|f\|_2^2 - 2\langle f, f\circ T\rangle$

$= 2\|f\|_2^2 - 2\langle f, f - f + f\circ T\rangle$

$= 2\|f\|_2^2 - 2\|f\|_2^2 + 2\langle f, f - f\circ T\rangle$

$= 2\langle f, f - f\circ T\rangle$ Now $f \in \mathcal{C}^\perp$ so $f - f\circ T \in \mathcal{C}$ by def while $f \in \mathcal{C}^\perp$

$= 0$

$\Rightarrow f = f\circ T \quad \forall f \in \mathcal{C}$

Clearly, for $f \in \mathcal{C}^\perp$, $\lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{N-1} f\circ T^k = \lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{N-1} f = f$

Thus $\circledast$ holds $\forall f \in L^2$ since $f = a_1 f_1 + a_2 f_2$ $\quad f_1 \in \overline{\mathcal{C}}$, $f_2 \in \mathcal{C}^\perp$

Infact
$$\lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{N-1} f\circ T^k = \begin{cases} 0 & \text{if } f \in \overline{\mathcal{C}} \\ f & \text{if } f \in \mathcal{C}^\perp \end{cases}$$
is a projection onto $\mathcal{C}^\perp$ (the invariant fcts).

Finally, if T is ergodic, we know invariant fcts are constant a.e.

By the Hölder, $\|(f_N - \bar{f})\cdot 1\|_1 \le \|(f_N - \bar{f})\|_2 \cdot \|1\|_2$

So L^2 convergence $\Rightarrow$ L^1 convergence so

$$\int f_N\, d\mu = \int \frac{1}{N}\sum_{k=1}^{N-1} f\circ T^k\, d\mu = \frac{1}{N}\sum_{k=1}^{N-1}\int f\circ T^k d\mu = \underbrace{\int f\, d\mu}_{\text{the constant}}.$$

(shorthand for N^{th} Birkoff sum)

$\square$

Cor A ppt $(X, \mathcal{B}, \mu, T)$ is ergodic $\iff$

$$\frac{1}{N}\sum_{k=0}^{N-1}\mu(A\cap T^{-k}B) \longrightarrow \mu(A)\mu(B) \quad \forall A, B \in \mathcal{B}$$

Pf/ Put $f_N = \frac{1}{N}\sum_{k=0}^{N-1} \underbrace{1_B}_{:= f}\circ T^k$, $g = 1_A$

L^2 Ergodic Thm $\Rightarrow$ $f_N \to \bar{f}$, $\bar{f} \in L^2$, $\bar{f}$ invariant

so $\langle f_n, g\rangle \longrightarrow \langle \bar{f}, g\rangle^{(**)}$ $\left[\because |\langle f_n - \bar{f}, g\rangle| \overset{HI}{\le} \underbrace{\|f_n - \bar{f}\|_2}_{\to 0} \cdot \underbrace{\|g\|_2}_{\text{finite}}\right]$

$(\Rightarrow)$ T is ergodic. By L^2 ergodic thm then

$$\bar{f} = \int f \, d\mu = \mu(B)$$

So $\langle \bar{f}, g \rangle = \int \bar{f} \cdot g \, d\mu = \int \mu(B) \mathbb{1}_A \, d\mu = \mu(B)\mu(A)$

$$\langle f_N, g \rangle = \int \frac{1}{N} \sum_{k=0}^{N-1} \mathbb{1}_{T^{-k}B} \mathbb{1}_A \, d\mu = \frac{1}{N} \sum_{k=0}^{N-1} \int \mathbb{1}_{T^{-k}B \cap A} \, d\mu$$

$$= \frac{1}{N} \sum_{k=0}^{N-1} \mu(T^{-k} B \cap A)$$

So by ergodicity, $\frac{1}{N} \sum_{k=0}^{N-1} \mu(T^{-k} B \cap A) \longrightarrow \mu(A)\mu(B)$

($\Leftarrow$) If T $\neg$ergodic $\Rightarrow \neg(*)$

$\exists$ A' s.t $\mu(A') \in (0,1)$, A' invar.

~~Set~~ $B = A'$ and $A = (X \setminus A')$

Then $\frac{1}{N} \sum_{k=0}^{N-1} \mu((X \setminus A') \cap T^{-k} A') = 0$ by invar of A'.

But $\mu(X \setminus A')\mu(A') > 0$. So this is the counterexample. $\square$

Birkhoff's Ergodic Theorem

Let $(X, \mathcal{B}, \mu, T)$ be a p.p.t. If $f \in L^1(X, \mu)$ then the following limit exists μ a.e. $x \in X$:

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

The limit is equal μ-a.e. to a T invar fn.

If T is ergodic, then the limit is $\int f d\mu$ μ-a.e.

Applications

(1) Trajectory distributions (T ergodic)

Put $f = 1_A$, $A \in \mathcal{B}$. Then $\lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} 1_A(T^k x)$

is $\lim_{N \to \infty} \dfrac{\#\text{ times } T^k x \in A, \ k = 0, \ldots, N-1}{N}$

which is the asymptotic freq. that the orbit of x enters A.

The BET says this is $\int f \, d\mu = \mu(A)$ for μ-a.e. init. pts. This is an interpret of $\mu(A)$.

eg For both circ. rotns / doubling/triply maps, for any meas. set. A, the freq w. which the orbits of Lebesgue a.a initial points visit A is the Leb. measure of A.

This is despite rotns / doublings having different dynamics.

(2) Space & Time averages

One can interp.

$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(T^k x)$ as the time average of x along its orbit. eg T is evolution of particles in the ocean/atmosphere and f is temperature/density etc. And x is an initial particle posn. Then the average tmp of the particle along its orbit is equal to the average temperature of the ocean/atmos.

So if we know μ we can compute ⊛ "simply" by integrat. wrt μ.

Eg $T: [0,1]\circlearrowleft$ is diff'ble.

$f(x) = \log|T'(x)|$

$$LE(x) = \lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} \log|T'(T^k x)|$$

$$\begin{cases} > 0 & \text{"expansion on average along orbit of } x\text{"} \\ < 0 & \text{"contraction on average"} \end{cases}$$

Then BET $\Rightarrow LE(x) = \int \log|T'(x)|\, d\mu$.

(3) Borel's Normal Number Thm

Lebesgue a.a real numbers in $[0,1)$ are normal i.e. the freq. of any digit $k \in \{0,1,\dots,9\}$ in the decimal exp. is $\frac{1}{10}$.

Though π, $\log 2$, $\sqrt{2}$, e not known if normal or not.

Pf/ $T:[0,1]\circlearrowleft$ $Tx = 10x \pmod 1$ preserves Lebesgue & is ergodic. (Delete all trailing 9s: let Y be the set of pts w. unique decimal expansion). Y has a countable complement which is all expansions w. trailing 9s. $\Rightarrow \mu(Y)=1$)

Suppose $x = \sum_{k\ge 1} a_k 10^{-k}$

Then $Tx = \sum_{k \geq 1} a_{k+1} 10^{-k}$

$f = \mathbb{1}_{[\frac{k}{10}, \frac{k+1}{10})}$ $\quad k = 0, \ldots, 9$

Then $f(T^i x) = f(\sum_{j \geq 1} a_{j+i} 10^{-j})$

$$= \begin{cases} 1 & a_{i+1} = k \\ 0 & \text{else} \end{cases}$$

Thus $\sum_{i=0}^{N-1} f(T^i x) = \#$ ks in the first N digits of x.

So $\lim_{N \to \infty} \frac{1}{N} \sum_{i=0}^{N-1} f(T^i x)$ is the overall average propn of ks in the expansion. So by BET this is equal to $\int f \, d\mu = \mu([\frac{k}{10}, \frac{k+1}{10})) = \frac{1}{10}$.

Pf of Ergodic thm:

We prove this only for $L^\infty(\mu)$ functions. Note it is enough to treat non negative f w. $\|f\|_\infty \leq 1$ (Since WOLA we can normalise fns $\to \frac{f}{\|f\|_\infty}$)

Similarly, we can write $f = f^+ - f^-$ where

$f^+ = f \cdot \mathbb{1}_{\{f \geq 0\}}$ and $f^- = f \cdot \mathbb{1}_{\{f < 0\}}$

And $(f \circ T)^+ = f^+ \circ T$ and $(f \circ T)^- = f^- \circ T$.

(Exercise!)

Fix some $f \in L^1$, $f \geq 0$, define

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

$\bar{A} := \limsup_{N \to \infty} A_N(x)$

$\underline{A} := \liminf_{N \to \infty} A_N(x)$

Easy to see that $\bar{A}$ and $\underline{A}$ are invariant by telescopy (compare $\bar{A} - \bar{A} \circ T$) (Exercise!)

Also $0 \leq \underline{A} \leq \bar{A} \leq \|f\|_\infty \leq 1$ $(***)$ (otherwise).

We show that $\int \bar{A} - \underline{A} \, d\mu = 0$ and so by $(***)$,

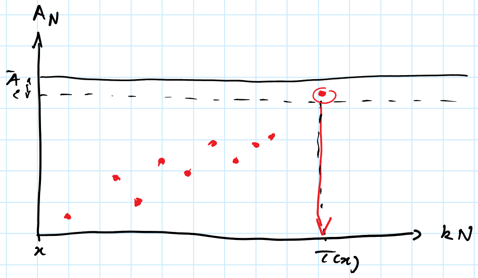
$\bar{A}(x) = \underline{A}(x)$ μ.a.e. [L^1 convergence & boundedness $\Rightarrow$ a.e. convergence]

$\therefore \lim_{N \to \infty} A_N(x)$ exists μ.a.e.

This limit is invar (telescoping & $\|f\| < 1$ exercise!).

So WTS $\int \bar{A} - \underline{A} \, d\mu = 0$: Fix $\varepsilon > 0$, M large, $N \gg M$ and

define $\tau(x) = \min\{N \geq 0 \; A_N(x) > \bar{A}_n - \varepsilon\}$



$$R = \{0 \leq k \leq N-M \mid \tau(T^k x) > M\}$$

$$\sum_{k \in R} \left[f(T^k x) + \mathbb{1}_{\{\tau > M\}} \right] \geq |R| \geq |R| (\bar{A} - \varepsilon) \quad \because \bar{A} < 1$$

(1)

Set $\quad$ first elt in seg $\hat{B}$

$$B = \{0 \leq k \leq N-M \mid \tau(T^k x) \leq M, [\tau(T^{k-1}x) > M \text{ or } k-1 \in B']\}$$

$$\hat{B} = \bigcup_{b \in B} \{b, b+1, \dots, b + \tau(b) - 1\}$$

$$B' = \{b + \tau(b) - 1 : b \in B\}$$

Now $R \cup \hat{B} = \{0, 1, 2, \dots, N-M\}$

$$\sum_{k \in \hat{B}} f(T^k x) + \mathbb{1}_{\{\tau > M\}}(T^k x)$$

$$= \sum_{k \in \hat{B}} f(T^k x) = \sum_{b \in B} \sum_{k=b}^{\tau(b)-1} f(T^k x)$$

$$= \sum_{b \in B} \tau(b) A_{\tau(T^b x)}(T^b x) \qquad \left[\text{Recall } A_n(x) = \frac{1}{N} \sum_{k=0}^{N-1} f(T^k x)\right]$$

$$> \sum_{b \in B} \tau(b) (\bar{A}(T^b x) - \varepsilon) \text{ by defn of } \tau.$$

$$= \sum_{b \in \hat{B}} (\bar{A}(x) - \varepsilon) \quad \text{using invariance of } \bar{A}$$

$$= |\hat{B}| (\bar{A}(x) - \varepsilon)$$

(2)

(1) and (2) $\Longrightarrow$

$$\sum_{k=0}^{N-1} f(T^k x) + \mathbb{1}_{\{\tau > M\}}(T^k x)$$

$$\geq \sum_{k=0}^{N-M} \dots$$

$$\geq (N-M)(\bar{A} - \varepsilon)$$

Integrating, we get

$$N\left(\int f \, d\mu + \mu(\{\tau > M\})\right) \geq (N-M)\left(\int \bar{A} \, d\mu - \varepsilon\right)$$

So $\int f \, d\mu \geq (1 - \frac{M}{N})(\int \bar{A} \, d\mu - \varepsilon) - \mu(\{\tau > M\})$

Send $N \to \infty$

$$\int f \, d\mu \geq \left(\int \bar{A} \, d\mu - \varepsilon\right) - \mu(\{\tau > M\}).$$

Now send $M \to \infty$

$$\int f \, d\mu \geq \int \bar{A} \, d\mu - \varepsilon$$

Finally, send $\varepsilon \to 0$,

$$\int f \, d\mu \geq \int \bar{A} \, d\mu$$

Repeat prev arguments for $\underline{A}$:

$$\tau'(x) = \min\{N \geq 0 \mid A_N \leq \underline{A} + \varepsilon\}$$

to get $\int f \, d\mu \leq \int \underline{A} \, d\mu$

Since $\bar{A}(x) \geq \underline{A}(x)$ $\forall x \in X$ and we have shown

$$\int \bar{A} \, d\mu \leq \int f \, d\mu \leq \int \underline{A} \, d\mu, \text{ we have}$$

$\int \bar{A} - \underline{A} \, d\mu = 0$ as required $\square$

Rem BET ...

Furthermore, $\bar{f}$ is integrable and $\int \bar{f} \, d\mu = \int f \, d\mu$ even if T is not ergodic.

Cedric's lect

Ergodic decomposition.

Let $(X, \mathcal{B}, \mu, T)$ be ppt. Where X is a compact metric space and $\mathcal{B}$ is the borel σ-alg. Then for each $x \in X$, $\exists$ a T invar measure μ_x such that:

(i) $\forall f \in L^1(X, \mu) = \int f \, d\mu = \int \left(\int f \, d\mu_x\right) d\mu(x)$

(ii) " " and μ a.e. x, we have

$$A_n f(x) \to \int f \, d\mu_x \quad \text{as } n \to \infty.$$

(iii) μ_x is invariant and ergodic prob. measure.

Sketch pf/: [We will assume that T is cts] (Based on Quas' & Dwarig's notes).

Step 1: Construction of μ_x

By BET, if $f \in L^1(\mu)$, then $\{A_N f(x)\}_{N \geq 1}$ converges for μ-a.e. x to a limit $\bar{f}(x)$. Fix a seq. of c.t.s. fns $\{f_j\}_{j \geq 0}$ which is dense in the space of continuous functions.

[i.e. for every $\varepsilon > 0$, $f \in C(X)$, $\exists\ j$ s.t

$\sup_{x \in X} |f(x) - f_j(x)| < \varepsilon$ i.e. $\|f - f_j\|_\infty < \varepsilon$

this follows from separability of $C(X)$]

So for each n, $\exists B_n \in X$ w. $\mu(B_n) = 1$ s.t

$A_N f_n(x)$ is convergent for any $x \in B_n$. (by the BET)

Let
$$B = \bigcap_{n \geq 1} B_n$$

$\mu(B) = 1$

and $\forall x \in B$: $\forall n$ $\{A_N f_n(x)$ converge$\}$ is convergent.

Consider the map $L_x : C(X) \to \mathbb{R}$ given by

$$L_x(f) = \begin{cases} \bar{f}(x) & x \in B \\ 0 & \text{else.} \end{cases}$$

L_x is linear and $L_x(f) \geq 0$ if $f \geq 0$ by defn.

$L_x(1) = 1$

So by Reizs representation thm, $\exists$ a probability measure μ_x s.t $L_x(f) = \int f \, d\mu_x$

[Define $\mu_x = \mu$ if $x \notin B$].

If μ is ergodic then $L_x(f) = \int f \, d\mu$

Examples:

① $(X, \mathcal{B}, \mu, T)$ is p.p.t. Then $\mu_x = \mu$.

② $(X, \mathcal{B}, l, T)$ where $X = S^1$
$Tx = x + 1/2 \pmod 1$

$$A_N f(x) = \frac{1}{N}\big(\underbrace{f(x) + f(x + \tfrac{1}{2}) + f(x) + \cdots}_{n \text{ terms.}}\big)$$

$$\bar{f}(x) = \lim_{N\to\infty} A_N f(x) = \frac{1}{2}\left(f(x) + f\left(x+\frac{1}{2}\right)\right) = \mu_x(f)$$

$$\left[\text{i.e. } \mu_x = \frac{1}{2}\left(\delta_x + \delta_{x+1/2}\right)\right].$$

Sketch pf of ergodic decomp:

Last time, we had found $\mathcal{B} \subseteq X$ s.t $\mu(\mathcal{B}) = 1$ and for any $x \in \mathcal{B}$, we identified a probability measure μ_x s.t. for μ-a.e. y & every cts function f,

$$\int f \, d\mu_x = \bar{f}(x) = \lim_{N\to\infty} A_N f(x)$$

(This is (ii) for cts functions f)

Let's check (i) for cts functions f: If f is cts then $\{A_N(f)\}_{N\geq 1}$ is a bounded sequence (bounded by maximum of f)

and $A_N f(x) \longrightarrow \int f \, d\mu$ for μ a.e x. So since this sequence is bounded by the boundd convergence thm.

$$\begin{array}{ccc} \int A_N f(x) \, d\mu & \longrightarrow & \int \left(\int f \, d\mu_x\right) d\mu. \\ \Big\Downarrow {\scriptstyle \because \mu \text{ is T inver}} & & \\ \int f \, d\mu & & \end{array}$$

So this gives (i) for cts functions

When $f \in L^1(\mu)$, (i) and (ii) also hold. (See Sarig's / Quas' notes?

(iii) μ_x is invariant:

$$\int f \, d\mu_x = \bar{f}(x) \text{ for } \mu \text{ a.e } x \text{ and cts } f.$$

$$\int f\circ T \, d\mu_x = \bar{f}(Tx) = \bar{f}(x) \quad (\text{by Birkhoff Ergodic Thm})$$

μ_x is Ergodic: (Cor of Birkhoff Ergodic Thm)

(I) $(X, \mathcal{B}, \mu, T)$ is p.p.t is ergodic

$$\Longleftrightarrow$$

$$\forall f \in L^1(\mu), \ \boxed{\lim_{N\to\infty} A_N f(x) \longrightarrow \int f \, d\mu \quad \text{for } \mu \text{ a.e}} \ \circledast$$

(II) If X is a compact Met. Sp. & $\mathcal{B}$ is Borel σ alg, then $(X, \mathcal{B}, \mu, T)$ is ergodic $\Longleftrightarrow$ $\circledast$ holds for every cts fn f.

Sketch Pf of μ_x is ergodic:

Claim (See Quas' Notes): Let $\{f_k\}$ be a countable dense set of cts fns, then

$$\mathcal{D} = \left\{ x \in X \mid \forall k \ \ f_k(y) = \int f_k \, d\mu_x \text{ for } \mu_x\text{-a.e } y \right\}$$

has full measure. ($\mu(D) = 1$)

Claim 2: For $x \in \mathcal{D}$, μ_x is ergodic

Pf/ Let $x \in \mathcal{D}$, $f \in C(X)$. Let k s.t.

$$\sup_{y \in X} | f(y) - f_k(y) | < \varepsilon . \text{ Then}$$

$$\left| \int f \, d\mu_x - \int f_k \, d\mu_x \right| \le \int | f - f_k | \, d\mu_x \le \varepsilon .$$

Also $| A_N f(y) - A_N f_k(x) | \le \varepsilon$ and

$\bar{f}_k(y) = \int f_k \, d\mu_x$ for μ_x a.e. y.

Then
$$\left| \int f \, d\mu_x - \bar{f}(y) \right| \le \left| \int f d\mu_x - \int f_k \, d\mu_x \right| + \left| \int f_k \, d\mu_x - \bar{f}_k(y) \right| + \left| \bar{f}_k(y) - \bar{f}(y) \right| \le \varepsilon + \varepsilon$$

Since ε is arb, $\int f \, d\mu_x = \bar{f}(y)$ for μ_x a.e y

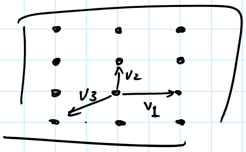
and by ~~cor~~ above, μ_x is ergodic.

(Non examinable)

Subadditive Ergodic Theorem

Example: A random walk on lattice $\mathbb{Z}^2$.

$(X, \mathcal{B}, \mu)$ Bernoulli scheme w. 3 symbols $\{3, 1, 2\}$ and prob vect (p_1, p_2, p_3) w. $\sum_{i=1}^{3} p_i = 1$



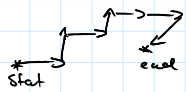
Def of V_i $i = 1, \dots, 3$

Consider

$f : X \to \mathbb{Z}^2$ s.t

$f(x_0, x_1, x_2) = V_{x_0}$

So 1, 2, 1, 2, 1, 1, 3 gives



Then $f_n(x) := \sum_{k=0}^{n-1} f(\sigma^k x)$ is our position on $\mathbb{Z}^2$ after n steps

Birkhoff ergodic theorem says that

$$\frac{1}{n} f_n(x) \to \int f \, d\mu = \sum_{i=1}^{3} p_i V_i$$

So the RW has "speed" $\| \sum_{i=1}^{3} p_i V_i \|$

and direction $\dfrac{\sum_{i=1}^{3} p_i V_i}{\| \sum_{i=1}^{3} p_i V_i \|}$

Eg : derivative cocycle

Suppose $T : V \circlearrowleft$ is a diffeomorphism from an open set $V \subset \mathbb{R}^d$. By the chain rule,

$$DT^n(x) = DT(T^{n-1}x) \circ \dots \circ DT(Tx) \circ DTx \quad (\text{chain rule})$$

$\Rightarrow$ multiplications along the orbit of x.

($DT : V \longrightarrow GL(d, \mathbb{R})$ the set of invertible $d \times d$ matrices).

is there a "speed" or "direction" ?

17th September 2014 – Perron Frobenius Operators etc

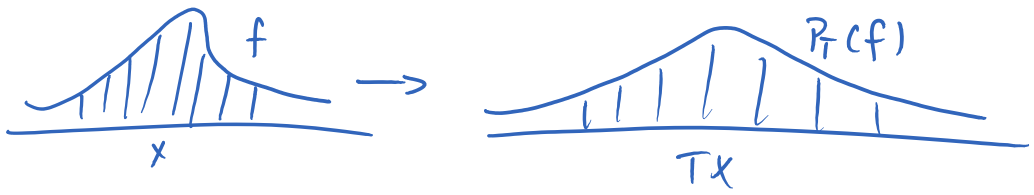
Review: $T : (X, \mathcal{B}, \mu) \circlearrowleft$ ergodic

The Perron - Frobenius (transfer operator) of T is

$$P_T : L^1(\mu) \circlearrowleft \qquad (\text{Notn}/ \quad P_T(f) = \frac{d}{d\mu}(\mu \circ T^{-1}) \quad)$$

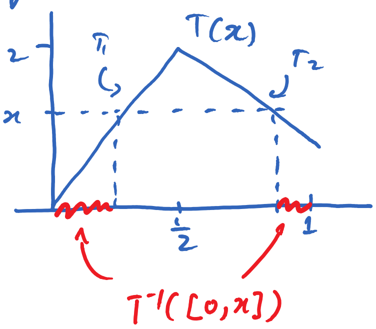
In particular: $\displaystyle\int_E P_T(f)\, d\mu = \int_{T^{-1}E} f \, d\mu$

So P_T describes the evolution of densities under T (the dynamics)



Example: $X = [0,1]$, $\mu = l$ (Lebesgue measure)

$$T(x) = \begin{cases} 2x & x \in [0, \frac{1}{2}] \\ \frac{5}{3} - \frac{4}{3}x & x \in [\frac{1}{2}, 1] \end{cases}$$



$T^{-1}([0,x]) = [0, \frac{1}{2}x]$ if $x < \frac{1}{3}$

$T^{-1}([0,x]) = [0, \frac{1}{2}x] \cup [\frac{5}{4} - \frac{3}{4}x, 1]$ $\quad x \geq \frac{1}{3}$

Combining:

$$T^{-1}[0,x] = [0, \tfrac{1}{2}x] \cup \left([\tfrac{5}{4} - \tfrac{3}{4}x, 1] \cap B \right) \text{ where } B = [\tfrac{1}{2}, 1]$$

Hence, for $f \in L^1(l)$:

$$\int_{T^{-1}[0,x]} f\, d\mu = \int_0^{\frac{1}{2}} f(x)\, dx + \int_{\frac{1}{4}(5-3x)}^{1} f(x)\, \mathbb{1}_B(x)\, dx$$

So $\displaystyle P_T(f) = \frac{d}{dx}\int_{T^{-1}[0,x]} f(x)dx = \frac{1}{2} f\left(\frac{x}{2}\right) + \frac{3}{4} f\left(\frac{5}{4} - \frac{3}{4}x\right) \cdot \mathbb{1}_{[0,\frac{1}{2})}\left(\frac{5}{4} - \frac{3}{4}x\right)$

$$= \frac{1}{2} f\left(\frac{x}{2}\right) - \frac{3}{4} f\left(\frac{5}{4} - \frac{3}{4}x\right) \mathbb{1}_{[1/3, 1]}(x)$$

by the chain rule.

There is an easier way to get the transfer operator however...

Properties of the Perron-Frobenius Operator

Let $X = [0,1] = I$. $\mathcal{B}$ the Borel σ-alg on I, μ the Lebesgue meas.

$T: I \circlearrowleft$ nonsingular and measurable. Let P_T be the Perron-Frobenius Operator associated to T.

Prop 1 P_T is linear : $P_T(af + bg) = aP_T(f) + bP_T(g)$, $f, g \in L^1(\mu)$

Pf/ Let $a, b \in \mathbb{R}$ and $f, g \in L^1(\mu)$ and consider $A \subset I$.

$$\int_A P_T(af+bg)\, d\mu = \int_{T^{-1}A} af + bg\, d\mu = a\int_{T^{-1}A} f\, d\mu + b\int_{T^{-1}A} g\, d\mu$$

$$= a\int_A P_T(f)\, d\mu + b\int_A P_T(g)\, d\mu$$

$$= \int_A aP_T(f) + bP_T(g)\, d\mu \qquad (\mu\text{-a.e})$$

Since this holds for arb $A \subset I$, $P_T(af+bg) = aP_T(f) + bP_T(g)$ a.e $\square$

Prop 2 (Positivity) Let $f \in L^1(\mu)$, $f \geq 0$ then $P_T(f) \geq 0$

Pf/ For $A \in \mathcal{B}$, $\int_A P_T(f)\, d\mu = \int_{T^{-1}A} f\, d\mu \geq 0$. Since A is arb. $P_T(f) \geq 0$ for a.e x.

$\square$

Prop 3 (Preservation of Integrals)

$$\int_I P_T f\, d\mu = \int_{T^{-1}I} f\, d\mu = \int_I f\, d\mu \qquad \square$$

Prop 4 (Contraction property) $P_T : L^1(\mu) \circlearrowleft$ is a contraction. That is,

$$\|P_T(f)\|_1 \leq \|f\|_1 \quad \forall f \in L^1(\mu) \quad \left(\text{Recall } \|f\|_1 = \int_I |f|\, d\mu\right)$$

Pf/ Let $f \in L^1(\mu)$, $f_+ := \max(f, 0)$, $f_- := -\min(0, f) \Rightarrow f_\pm \geq 0$

Then $f = f_+ - f_- \in L^1(\mu)$, $|f| = f_+ + f_-$. By linearity,

$P_T(f) = P_T(f_+ - f_-) = P_T(f_+) - P_T(f_-)$. Hence, we bound:

$|P_T(f)| \leq |P_T(f_+)| + |P_T(f_-)| = P_T(f_+) + P_T(f_-) = P_T(f_+ + f_-) = P_T(|f|)$.

So $\|P_T(f)\|_1 = \int_I |P_T(f)|\, d\mu \leq \int_I P_T(|f|)\, d\mu \overset{P.3}{=} \int_I |f|\, d\mu = \|f\|_1$ $\square$

Prop 5 (Composition) Let $S: I \circlearrowleft$ nonsingular and measurable. Then

$$P_{T\circ S}(f) = (P_T \circ P_S)(f)$$

In particular, $P_{T^n}(f) = (P_T)^n(f)$.

Aside: $\nu \ll \mu$ (abs. cts) is important to define the Peron-Frob. op via Radon Nykodym Thm

Pf/ Let $f \in L^1(\mu)$. Let $\nu(A) = \int_{(T\circ S)^{-1}A} f \, d\mu$

So $\nu(A) = \int_A P_{T\circ S}(f) \, d\mu$.

Also, $\int_A P_T \circ P_S f \, d\mu = \int_{T^{-1}A} P_S f \, d\mu = \int_{S^{-1}\circ T^{-1}A} f \, d\mu = \int_{(T\circ S)^{-1}A} f \, d\mu$

$$= \int_A P_{T\circ S}(f) \, d\mu = \nu(A)$$

Since A is arb, $P_{T\circ S}(f) = P_T \circ P_S f$ μ.a.e.

Definition: The Koopman Operator aka. composition operator, $U_T: L^\infty(\mu) \circlearrowleft$ defined as $U_T(g) = g\circ T$.

Property 6: (Adjoint property) If $f \in L^1(\mu)$, $g \in L^\infty(\mu)$. Then

$$\int_I P_T f \, d\mu = \int_I f \cdot U_T(g) \, d\mu$$

Pf/ Consider $f \in L^1(\mu)$, $A \in \mathcal{B}$, $g = \mathbb{1}_A$, then

$$\int_I f \cdot U_T(g) \, d\mu = \int_I f \cdot g\circ T \, d\mu = \int_I f \cdot \mathbb{1}_{T^{-1}A} \, d\mu$$

$$= \int_{T^{-1}A} f \, d\mu =$$

So then the adjoint property holds if g is a char. fn. Since char fns are dense in $L^\infty(\mu)$, then the adjoint property holds for any $f \in L^1(\mu)$, and $g \in L^\infty(\mu)$.

Fixed points of P_T are densities of T-invariant measures (abs. cts. wrt. μ).

Prop 7: Let $f^* \in L^1(\mu)$, $f^* \geq 0$, $\|f\|_1 = 1$. Then

$P_T f^* = f^* \iff$ measure ν defined by $\frac{d\nu}{d\mu} = f^*$ [i.e. $\nu(A) = \int_A f^* d\mu$]

is a T invar prob. measure.

Fixed points of P_T are densities of T-invariant measures (abs. cts. wrt. μ).

Prop 7: Let $f^* \in L^1(\mu)$, $f^* \geq 0$, $\|f\|_1 = 1$. Then
$P_T f^* = f^* \iff$ measure ν defined by $\frac{d\nu}{d\mu} = f^*$ [i.e. $\nu(A) = \int_A f^* d\mu$]
is a T invar prob. measure.

Pf/ Suppose $P_T f^* = f^*$. Then $\int_A P_T f^* d\mu = \int_A f \, d\mu = \nu(A)$.

Also, $\int_A P_T f^* d\mu = \int_{T^{-1}A} f^* d\mu = \nu(T^{-1}A)$ so stat. distribs. f^* yield
T-invar measures ν. Moreover ν is a prob. measure:
$\nu(I) = \int_I f^* d\mu = \|f^*\|_1 = 1 \quad (f^* \geq 0)$.

In the other direction, now assume $\nu(A) = \nu(T^{-1}A)$ $\forall A \in \mathcal{B}$,
then $\int_{T^{-1}A} f^* d\mu = \int_A f^* d\mu \Rightarrow \int_A P_T f^* d\mu = \int_A f^* d\mu$. Since
A is arb, $P_T f^* = f^*$.

$\square$

Relation between ergodic properties of T and P.F. operators.

Def Let $\mathcal{D}(I, \mathcal{B}, \mu) = \{ f \in L^1(\mu) \mid f \geq 0, \int f \, d\mu = 1 \iff \|f\|_1 = 1 \}$ be
the set of density functions.

Def A sequence of $L^1(\mu)$ functions $\{f_n\}_{n\geq 1}$ converges weakly in L^1 to f
if for every $g \in L^\infty(\mu)$, $\int_I f_n \cdot g \to \int f g \, d\mu$ as $n \to \infty$.

Prop Let $(I, \mathcal{B}, \mu, T)$ be a ppt. Then

① T ergodic $\iff$ $\forall f \in \mathcal{D}(I, \mathcal{B}, \mu)$, the seq. $\{ \frac{1}{n} \sum_{k=0}^{n-1} P_T^k(f) \}_{n \geq 1}$
converges weakly in L^1.

② T is mixing $\iff$ $\forall f \in \mathcal{D}(I, \mathcal{B}, \mu)$ $\{P_T^n(f)\}_{n\geq 1}$ converges weakly in L^1.

Pf (1)/ T ergodic $\iff$ $\forall A, B \in \mathcal{B}$, $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(A \cap T^{-k}B) = \mu(A)\mu(B)$
In fact, T ergodic $\iff$ $\forall f \in L^1(\mu)$, $g \in L^\infty(\mu)$ $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int f \cdot g \circ T^k d\mu = \int f \, d\mu \int g \, d\mu$
Using the adjoint property, T is ergodic $\iff$ $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int P_T^k f \cdot g = \int f \, d\mu \int g \, d\mu$.

This is exactly what $\{\frac{1}{n}\sum_{k=0}^{n-1} P_T^k f\}_{n\geq 1}$ converges weakly to in $L^1(\mu)$:

$$\frac{1}{n}\sum_{k=0}^{n-1} P_T^k f \to \int f\,d\mu \int g\,d\mu = \int g\,d\mu \quad (\because f \in \mathcal{D}(I,\mathcal{B},\mu))$$

$\square$

Pf(2) / Similar to above, left as exercise.

Representation of Perron - Frobenius Operators

Def Let $I = [a,b]$, $T: I \circlearrowleft$ is <u>piecewise $C^{r\geq 1}$ monotonic</u> if $\exists$ partition of I
$a = a_0 < a_1 < \cdots < a_q = b$ such that:

(i) $T|_{(a_{i-1},a_i)}$ is a C^r function, extendable to $[a_{i-1}, a_i]$ $\forall i = 1,\ldots,q$.

(ii) $|T'(x)| > 0 \quad \forall x \in (a_{i-1}, a_i)$

Prop Suppose $T: I \circlearrowleft$ is piecewise C^r monotonic. Then

$$P_T f(x) = \sum_{z\in T^{-1}x} \frac{f(z)}{|T'(z)|}$$

END OF LECTURE

24th Sept 2014

Recap from last time:

We showed that fixed points of PF operator:

$$P_T f = f \Rightarrow \exists\, T\text{-invar measures abs. cts wrt. Lebesgue}$$

Next: We show $\exists$ of non neg f.p. of P_T for T piecewise smooth, expanding interval maps. The main tool we will use is the notion of functions of bounded variation.

Def: <u>Functions of b.d. variation</u> in 1D: Let $I = [a,b]$, λ denote the Lebesgue measure. For sequence $a = x_0 < x_1 < \cdots < x_n = b$, $n \geq 1$ we define the partition:

$$\mathcal{S} = \{ I_i \mid I_i = [x_{i-1}, x_i),\ 1 \leq i \leq n \}$$

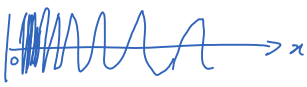
(Sometimes denoted $\mathcal{S}\{x_0, \ldots, x_n\}$). Then $f: [a,b] \to \mathbb{R}$ s.t $\exists M > 0$ so that for <u>any</u> partition $\mathcal{S} = \mathcal{S}\{x_0, \ldots, x_n\}$, we have

$$\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| \leq M$$

is a function of bounded variation.

Rem If f is monotonic, or if f satisfies a Lipschitz condition

$$|f(x) - f(y)| \leq K|x-y|$$

then f is of B.V. A fn that is <u>not</u> bounded in variation is for example: $f(x) = \sin\frac{1}{x}$ 

Def Let $f: [a,b] \to \mathbb{R}$ be a function of b.d. var. then

$$V_{[a,b]} f = \sup_{\mathcal{S}\{x_0, \ldots, x_n\}} \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|$$

is the <u>variation</u> of f on $[a,b]$.

Lem 1 If f is of B.V. on I, then it is bounded on I.

Pf/ (Tutorial Exercise)

Lem 2 Let f be of B.V on I s.t $\|f\|_1 < \infty$ [i.e. $\int_I |f|\, d\lambda < \infty$], then

$$|f(x)| \leq V_{[a,b]} f + (b-a)^{-1} \|f\|_1 \quad \forall x \in I.$$

Pf/ over pg $\to$

Lem2 Let f be of B.V on I s.t $\|f\|_1 < \infty$ [i.e. $\int_I |f| \, d\lambda < \infty$], then

$$|f(x)| \le V_{[a,b]} f + (b-a)^{-1} \|f\|_1 \quad \forall x \in I.$$

Pf/ ① $\exists$ $y \in I$ s.t. $|f(y)| \le (b-a)^{-1} \|f\|_1$. (Otherwise, $|f(x)| > (b-a)^{-1} \|f\|_1$ then $\|f\|_1 = \int_I |f| \, d\lambda > \int_I (b-a)^{-1} \|f\|_1 \, d\lambda = \|f\|_1$ contradiction!)

② By $\triangle$-ineq: $|f(x)| \le |f(y) - f(x)| + |f(x)|$, then

$$|f(x)| \le V_{[a,b]} f + \underbrace{(b-a)^{-1} \|f\|_1}_{\because ①}$$

$\square$

Thm Let f, g be fns of B.V. on I then $f+g$, $f-g$ and fg are of B.V on I. Moreover:

$$V_I(f \pm g) \le V_I f + V_I g \quad \underline{\text{and}} \quad V_I(fg) \le V_I f \, \|g\|_\infty + V_I g \, \|f\|_\infty$$

where $\|\cdot\|_\infty$ is the sup norm.

Cor If $c \in (a,b)$ then $V_{[a,b]} f = V_{[a,c]} f + V_{[c,b]} f$

Thm (Helly Selection Thm)

Let $\{f_n\}_{n \ge 1}$ be a sequence of fns of B.V. on I and $|f_n(x)| \le K$, $V_I f_n \le K$ then exists a subsequence $\{f_{n_k}\}_{k \ge 1}$ converging at every pt. of $[a,b]$ to a function f^* of B.V on I. Moreover, $f^* \in L^1$ w. $V_I f^* \le K$.

Def Let $B.V(I) = \{ f \in L^1 \mid \inf_{g = f \text{ a.e}} V_I g < \infty \}$, we define a $\underline{\text{norm}}$ on $B.V(I)$ as follows.

$$\|f\|_{BV} = \|f\|_1 + \inf_{\substack{g=f \\ \text{a.e.}}} V_I g$$

(Infact B.V(I) is a Banach Space w. B.V(I) $\underline{\text{dense}}$ in L^1.

If $V_I f_n \le K \;\; \forall n$ and $f_n \to f \in L^1 \Rightarrow V_I f \le K$)

Now, we look at existence of a.c.i.m. (abs. cts invar. measures):

Def: Let $\mathcal{T}(I)$ be the set of transformations $T: I \circlearrowleft$ s.t.

(i) T is piecewise expanding: $\exists$ $\mathcal{S}$ $\{a_1, \dots, a_q\}$ s.t. $T|_{I_i}$ is C^1 where $I_i = [a_{i-1}, a_i)$ and $|T'(x)| \ge \alpha > 1 \quad \forall \; 1 \le i \le q$, $x \in (a_{i-1}, a_i)$

(ii) Let $g(x) = |T'(x)|^{-1}$, and assume g is of $B.V(I)$. Recall if $T \in \mathcal{T}(I)$, let $\varphi_i = (T|_{I_i})^{-1}$, then

$$P_T f(x) = \sum_{y \in T^{-1}(x)} \frac{f(y)}{|T'(y)|} = \sum_{i=1}^{q} f(\varphi_i(x)) \, g(\varphi_i(x)) \, \mathbb{1}_{T(I_i)}(x)$$

Thm Let $T \in \mathcal{T}(I)$, then T admits an a.c.i.m. with density of B.V.

Lem Let $T \in \mathcal{T}(I)$, $g(x) = |T'(x)|^{-1}$ and $\delta = \min_{1 \le i \le q} \lambda(I_i)$ then for $f \in B.V(I)$:

$$V_I(P_T f) \le A V_I f + B \|f\|_1 \quad \text{where} \begin{cases} A = 3\alpha^{-1} + \max_{1 \le i \le q} V_{I_i} g \\ B = 2(\alpha\delta)^{-1} + \delta^{-1} \max_{1 \le i \le q} V_I g \end{cases}$$

(Recall $1 < \alpha < |T'(x)| \quad \forall x \in (a_{i-1}, a_i)$)

Pf/ Let $a = x_0 < x_1 < \dots < x_r = b$. Let $\varphi_i = (T|_{I_i})^{-1}$. Then,

$$S = \sum_{j=1}^r |P_T f(x_j) - P_T f(x_{j-1})|$$
$$= \sum_{j=1}^r \sum_{i=1}^q | g(\varphi_i(x_j)) f(\varphi_i(x_j)) 1\!\!1_{T(I_i)}(x_j) - g(\varphi_i(x_{j-1})) f(\varphi_i(x_{j-1})) 1\!\!1_{T(I_i)}(x_{j-1}) |$$

Split into 3 cases:

$$\left.\begin{array}{l} (i)\ 1\!\!1_{T(I_i)}(x_j) = 1 = 1\!\!1_{T(I_i)}(x_{j-1}) \\ (ii)\ 1\!\!1_{T(I_i)}(x_j) = 1,\ 1\!\!1_{T(I_i)}(x_{j-1}) = 0 \\ (iii)\ 1\!\!1_{T(I_i)}(x_j) = 0,\ 1\!\!1_{T(I_i)}(x_{j-1}) = 1 \end{array}\right\} \Rightarrow S = S_{(i)} + S_{(ii)} + S_{(iii)}$$

In case (i)

$$|f(\varphi_i(x_j)) g(\varphi_i(x_j)) - f(\varphi_i(x_{j-1})) g(\varphi_i(x_{j-1}))|$$
$$\le | f(\varphi_i(x_j)) [g(\varphi_i(x_j)) - g(\varphi_i(x_{j-1}))] | + | g(\varphi_i(x_{j-1})) [f(\varphi_i(x_j)) - f(\varphi_i(x_{j-1}))] |$$
$$\le \sup_{I_i} |f| \, | g(\varphi_i(x_j)) - g(\varphi_i(x_{j-1})) | + | f(\varphi_i(x_j)) - f(\varphi_i(x_{j-1})) | \sup_{I_i} |g|$$

Then: $S_{(i)} \le \sum_{i=1}^q \sup_{I_i} |f| V_{I_i} g + \sup_{I_i} |g| V_{I_i} f$ and by Lemma

$$S_{(i)} \le \max_{1 \le i \le q} V_{I_i} g \left(\sum_{i=1}^q V_{I_i} f + \lambda(I_i)^{-1} \int_{I_i} |f| \, d\lambda \right) + \sup_I |g| \sum_{i=1}^q V_{I_i} f$$
$$\le \max_{1 \le i \le q} V_{I_i} g \left(V_I f + \delta^{-1} \int_I |f| \, d\lambda \right) + \alpha^{-1} V_I f$$

For cases (ii), (iii), notice that (ii) only happens if $x_j \in T(I_i)$ and $x_{j-1} \notin T(I_i)$ so x_j, x_{j-1} are on different sides of an end p.t of I_i. So for each i, there is at most one pair of points x_{j-1}, x_j with this property. Similarly for (iii), this can only happen for a pair x'_{j-1}, x'_j. Then,

$$S_{(ii)} + S_{(iii)} \le \sum_{i=1}^q | g(\varphi_i(x_j)) f(\underbrace{\varphi_i(x_j)}_{s_i}) | + | g(\varphi_i(x'_{j-1})) f(\underbrace{\varphi_i(x'_{j-1})}_{r_i}) |$$
$$\le \alpha^{-1} \sum_{i=1}^q | f(s_i) | + | f(r_i) |$$
$$\le \alpha^{-1} \sum_{i=1}^q 2|f(v_i)| + | f(s_i) - f(v_i) | + | f(r_i) - f(v_i) | \quad \text{w. } v_i \text{ s.t. } |f(v_i)| \le \frac{\int_{I_i} |f| \, d\lambda}{\lambda(I_i)}$$
$$\le \alpha^{-1} \sum_{i=1}^q 2 \lambda(I_i)^{-1} \int_{I_i} |f| \, d\lambda + 2 V_{I_i} f$$
$$\le 2\alpha^{-1} V_I f + 2(\alpha\delta)^{-1} \|f\|_1 .$$

Combining, $S = S_{(i)} + S_{(ii)} + S_{(iii)}$

$$\le (3\alpha^{-1} + \max_{1\le i\le q} V_{I_i} g) V_I f + (2(\alpha\delta)^{-1} + \delta^{-1} \max_{1\le i\le q} V_{I_i} g) \|f\|_1$$

So $V_I(P_T f) \le (3\alpha^{-1} + \max_{1\le i\le q} V_{I_i} g) V_I f + (2(\alpha\delta)^{-1} + \delta^{-1}\max_{1\le i\le q} V_{I_i} g)\|f\|_1$.

Lem4 Let $T \in \mathcal{T}(I)$ and $g_n(x) = |(T^n)'(x)|^{-1} = |T'(T^{n-1}x)\dots T'(x)|^{-1}$ and $S^{(n)}$ be partition associated to T^n (see tute sheet 5). Then

$$W_n = \max_{J\in S^{(n)}} V_J g_n \le (n/\alpha^{n-1}) W_1$$

Pf/ Exercise!

Lem5 Let $T \in \mathcal{T}(I)$, then there exists $r \in (0,1)$, $C>0$, $R>0$ s.t. $\forall f \in B.V(I)$

$$\|P_T^n f\|_{BV} \le C r^n \|f\|_{BV} + R\|f\|_1$$

Pf/ Let $\alpha_n = (\sup_I g_n(x))^{-1}$, W_n as in lem 4 and $\delta_n = \min_{J\in S^{(n)}} \lambda(J)$. By the chain rule, $\alpha_n \ge \alpha^n$. By Lem 4, $W_n \le n\alpha^{-n+1} W_1$. Since $\alpha>1$ $\exists$ $k\ge 1$ s-t $r_k := 3\alpha_k^{-1} + W_k < 1$. Lets fix such k. Let $R_k = r_k \delta_k^{-1}$, $C_1 = \max\{r_1, \dots, r_k\}$ $C_2 = \max\{r_2\delta_2^{-1}, \dots, r_{k-1}\delta_{k-1}^{-1}\}$, $n = jk + i$, $0 \le i \le k-1$. Then $P_T^n = (P_{T^k})^j P_{T^i}$.

By Lem 3 multiple applications yield:

$$\begin{aligned}
V_I(P_{T^n} f) &= V_I((P_{T^k})^j (P_{T^i} f)) \\
&\le r_k V_I((P_{T^k})^{j-1} P_{T^i} f) + R_k \|f\|_1 \\
&\vdots \\
&\le r_k^j V_I(P_{T^i} f) + (1 + r_k + r_k^2 + \dots + r_k^{j-1}) R_k \|f\|_1 \\
&\le C_1 r_k^j V_I(f) + (C_2 + \tfrac{1}{1-r_k}) R_k \|f\|_1
\end{aligned}$$

So $\|P_T^n f\|_{BV} \le \|P_T^n f\|_1 + V_I(P_T^n f)$

$$\le C_1 r_k^j V_I f + (1 + C C_2 + \tfrac{1}{1-r_k}) R_k) \|f\|_1$$

Let $r = r_k^{1/k}$, $C = C_1 r^{-k+1}$, $R = R_k(C_2 + \frac{1}{1-r_k}) + 1$

Then $\|P_{T^n} f\|_{BV} \le C r^n \|f\|_{BV} + R\|f\|_1$