Ergodicity, cont

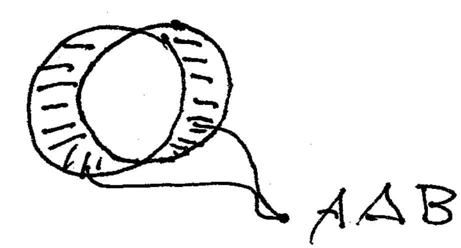
DEF: $f:(X,\mathcal{B},\mu)\circlearrowleft$ is a mpt. It is called ergodic if $A\in\mathcal{B}$, $f^{-1}(A)=A$ implies $\mu(A)=0$ or $\mu(A)=1$

Remarks (1) If $f^{-1}(A)=A$ then $f^{-1}(A^c)=A^c$ and $0<\mu(A)<1$ then $0<\mu(A^c)<1$ and f splits into a pair of subsystems each with positive measure $f|_A$ and $f|_{A^c}$

(2) Since sets of measure zero are ignorable some books define ergodicity in terms of sets which are invariant up to measure zero.

This is formulated like this:

Recall that the symmetric difference between two sets is $A \Delta B = (A-B) \cup (B-A)$



<u>Alternative Def of ergodicity</u>: f is ergodic if $A \in \mathcal{B}$, $\mu(A \Delta f^{-1}(A)) = 0$ implies $\mu(A)=0$ or $\mu(A)=1$.

(3) It is often just said briefly: "μ is an ergodic measure for f". Implicit in this is the fact that f preserves μ and there is a σ-algebra on which μ is defined.

We will have three theorems which give alternative conditions equivalent to ergodicity including the equivalence of the original and alternative definitions of ergodicity. They will allow us to prove the various examples are ergodic

<u>Theorem</u> $f: (X, \mathcal{B}, \mu)$? is a mpt, the following are equivalent

(i) f is ergodic by original definition

(ii) f is ergodic by alternative definition

(iii) If $A \in \mathcal{B}$ and $\mu(A) > 0 \Rightarrow \mu\left(\bigcup_{n=1}^{\infty} f^{-n}(A)\right) = 1$

(iv) If $A, B \in \mathcal{B}$, $\mu(A) > 0$, $\mu(B) > 0 \Rightarrow \exists\, n > 0$ with $\mu(f^{-n}A \cap B) > 0$

<u>Remark</u>: (iii) and (iv) indicate that f must "mix up" X when it is ergodic (but there is a precise condition called "mixing" which is stronger than ergodicity.

Before embarking on the proof we give a list (without proof) of some basis properties of Δ

$A \Delta B = B \Delta A$

$(A \Delta B) \Delta C = A \Delta (B \Delta C)$

$(A \Delta B) \Delta (B \Delta C) = A \Delta C$

(b) $\to$ $A \Delta C \subseteq (A \Delta B) \cup (B \Delta C)$ (a triangle inequality)

(c) $\to$ $(\cup A_\alpha) \Delta (\cup B_\alpha) \subseteq \cup (A_\alpha \Delta B_\alpha)$

$A \cap (B \Delta C) = (A \cap B) \Delta (A \cap C)$

If f is onto, $f^{-1}(A \Delta B) = f^{-1}(A) \Delta f^{-1}(B)$

Now assume μ is a probability and $A, B \in \mathcal{B}$

(e) $\to$ $|\mu(A) - \mu(B)| \leq \mu(A \Delta B)$

(g) $\to$ $\mu(A \Delta B) = 0 \Rightarrow \mu(A) = \mu(B)$

(d) $\to$ $A \subset B$, $\mu(A) = \mu(B) \Rightarrow \mu(A \Delta B) = 0$

Proof: (i) $\Rightarrow$ (ii) Assume $\mu(f^{-1}(B)\,\Delta\,B) = 0$, we construct a $B_\infty \in \mathcal{B}$ with (1) $f^{-1}(B_\infty) = B_{\infty}$, and so by (i), $\mu(B_\infty) = 0$ or 1. and (2).

$\mu(B) = \mu(B_\infty) = 0$ or 1, completeing the proof

The first step is to claim that for all $n \geq 0$, $\mu(f^{-n}(B)\,\Delta\,B) = 0$. Using (b),

$$f^{-n}(B)\,\Delta\,B \subseteq \bigcup_{i=0}^{n-1} f^{-(i+1)}(B)\,\Delta\,f^{-i}(B)$$

$$= \bigcup_{i=0}^{n-1} f^{-i}(f^{-1}(B)\,\Delta\,B)$$

so $\mu(f^{-n}(B)\,\Delta\,B) \leq n\,\mu(f^{-1}(B)\,\Delta\,B) = 0$

Let $B_\infty = \bigcap_{n=0}^{\infty} \bigcup_{i=n}^{\infty} f^{-i}(B)$

By (c), $B \,\Delta \bigcup_{i=n}^{\infty} f^{-i}(B) \subseteq \bigcup_{i=n}^{\infty} B \,\Delta\, f^{-i}(B)$

and so $\mu\left(B \,\Delta \bigcup_{i=n}^{\infty} f^{-i}(B)\right) \leq \sum_{i=n}^{\infty} \mu\left(B \,\Delta\, f^{-i}B\right) = 0$ by claim.

and so by (b), $\mu(B) = \mu\left(\bigcup_{i=n}^{\infty} f^{-i}(B)\right)$.

Now $\bigcup_{i=n}^{\infty} f^{-i}B$ indexed by n is a decreasing chain in n and each $\bigcup_{i=n}^{\infty} f^{-i}B$ has the same measure as B. Thus using the definition of B_∞

$\mu(B_\infty) = \mu(B)$ proving (2) above

To finish, again using the definition of B_∞

$$f^{-1}(B_\infty) = \bigcap_{n=0}^{\infty} \bigcup_{L=n}^{\infty} f^{-(L+1)}(B)$$

$$= \bigcap_{n=0}^{\infty} \bigcup_{L=n+1}^{\infty} f^{-n}(B) = B_\infty,$$ proving (i).

(ii) $\Rightarrow$ (iii) Let $A \in \mathcal{B}$ with $\mu(A) > 0$ and let $A_1 = \bigcup_{n=1}^{\infty} f^{-n}(A)$. By definition, $f^{-1}(A_1) \subseteq A_1$ and by invariance of the measure, $\mu(f^{-1}(A_1)) = \mu(A_1)$, so by (d) $\mu(f^{-1}A_1 \Delta A_1) = 0$. Thus by the assumption of (ii), $\mu(A_1) = 0$ or 1. Now $f^{-1}(A) \subseteq A_1$ and $\mu(f^{-1}(A)) = \mu(A) > 0$ so $\mu(A_1) > 0$ and so $\mu(A_1) = 1$.

(iii) $\Rightarrow$ (iv) Assume $\mu(A), \mu(B) > 0$. By (iii) we have $\mu\left(\bigcup_{n=1}^{\infty} f^{-n}(A)\right) = 1$. Now in general, $W, C \in \mathcal{B}$ and $\mu(C) = 1 \Rightarrow \mu(W \wedge C) = \mu(W)$ for any

so $0 < \mu(B) = \mu\left(B \wedge \bigcup_{n=1}^{\infty} f^{-n}(A)\right) = \mu\left(\bigcup_{n=1}^{\infty} B \wedge f^{-n}(A)\right)$

and so for some n, $\mu(B \wedge f^{-n}(A)) > 0$

(iv) $\Rightarrow$ (i) We prove the contrapositive, so assume $B \in \mathcal{B}$ with $f^{-1}(B) = B$ and so $f^{-n}(B) = B \ \forall n \geq 0$ and $0 < \mu(B) < 1$. Now this implies $0 < \mu(B^c) < 1$ and certainly $B^c \in \mathcal{B}$ and

$0 = \mu(B \wedge B^c) = \mu(f^{-n}B \wedge B^c)$, for all $n \geq 0$

contradicting (iv) ∎

Recall the Bernoulli Shift on Σ_n is based on a probability vector $\vec{p} = (p_0, \dots, p_{n-1})$ with $\sum p_i = 1$ and we assume now all $p_i > 0$. A cylinder set

$$C = {}_k[b_0 b_1 \dots b_\ell] \text{ has measure}$$

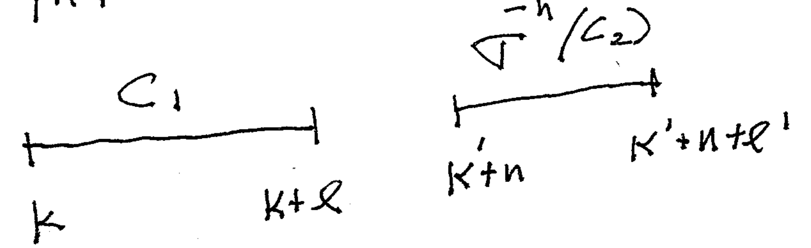
$$\mu_p(C) = p_{b_0} p_{b_1} \cdots p_{b_\ell}$$

<u>Theorem</u> Bernoulli Shifts are ergodic

First the idea of the proof. Let $C_1 = {}_k[b_0 \dots b_\ell]$ and $C_2 = {}_{k'}[b_0' \dots b_{\ell'}']$ be two cylinder sets

Pick n so that $k' + n > k + 3$

So schematically



- Now because the measure is a product measure

$\mu(C_1 \cap \sigma^{-n}(C_2)) = \mu(C_1) \cdot \mu(\sigma^{-n}(C_2)) = \mu(C_1)\,\mu(C_2) > 0$

so by part (iv) of the Theorem we would like to say we are done

- The problem is that we must verify (iv) for every pair of measurable sets.

- We first need another basic extension Theorem

<u>Theorem</u>: Let $(X, \mathcal{B}, \mu)$ be a probability space and the algebra $\mathcal{A}$ generates $\mathcal{B}$. Then $\forall \varepsilon > 0$, $\forall B \in \mathcal{B}$

$\exists A \in \mathcal{A}$ with $\mu(A \Delta B) < \varepsilon$

- NOTE: By (e) $|\mu(A) - \mu(B)| \leq \mu(A \Delta B) < \varepsilon$ also

- Via approximation we could make the above idea work but it is better to use a more direct clever method (next time)