Chapter 5 : Drift condi° and moments bounds

In all this chapter, we fix a Markov kernel P on $(\mathcal{X}, \mathcal{X})$ and for $C \in \mathcal{X}$, we recall that we denote σ_C and τ_C, the first return and hitting times resp.

Goal today : Last time, we have seen condi° upon which $\exists$ invariant probab. measure.

① $\mathbb{P}_x(N_C = \infty) = 1 \quad \forall x \in \mathcal{X}$ for a given small set C

② $\tilde{\mathbb{E}}_{x \otimes x_0}[\tau] < \infty$

We have seen that ① is implied by $\mathbb{P}_x(\tau_C < \infty) = 1$ $\forall x$ or $\mathbb{P}_x(\sigma_C < \infty) = 1$ $\forall x$ using the Markov prop

One of these condi° is also implied by

$\forall x \quad \mathbb{E}_x[\tau_C] < \infty$ or $\mathbb{E}_x[\sigma_C] < \infty$.

Practical condi° to show such moment bounds or even other type are Lyapunov condi°.

Thm : Consider (U_n), (Y_n) and (Z_n) three ⊕ seq of r.v on $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_n)_n)$, $(\mathcal{F}_n)$-adapted. Suppose $\forall n$

$$\mathbb{E}[U_{n+1} \mid \mathcal{F}_n] \leq U_n + Y_n - Z_n.$$

Then for any $(\mathcal{F}_n)$-stopping time τ, it holds that

$$\mathbb{E}[\mathbb{1}_{\{\tau < \infty\}} U_\tau] + \mathbb{E}\Big[\sum_{k=0}^{\tau-1} Z_k\Big] \leq \mathbb{E}[U_0] + \mathbb{E}\Big[\sum_{k=0}^{\tau-1} Y_k\Big].$$

proof: We start by the case $\tau = m$

In this case, it is a simple induc°.

Consider now the generic case τ: for any $k \in \mathbb{N}$.

$$\mathbb{E}[U_{k+1\wedge\tau} \mid \mathcal{F}_k]$$

$$= \mathbb{E}[\mathbb{1}_{\{\tau > k\}} U_{k+1} \mid \mathcal{F}_k] + \mathbb{E}[\mathbb{1}_{\{\tau \le k\}} U_\tau \mid \mathcal{F}_k]$$

$$\le \mathbb{1}_{\{\tau > k\}}(U_k + Y_k - Z_k) \qquad = \mathbb{E}[\mathbb{1}_{\{\tau = \tau\wedge k\}} U_{\tau\wedge k} \mid \mathcal{F}_k].$$

$$= \mathbb{1}_{\{\tau \le k\}} U_\tau.$$

$$\mathbb{E}[U_{k+1\wedge\tau} \mid \mathcal{F}_k] \le \mathbb{1}_{\{\tau > k\}}(U_k + Y_k - Z_k) + \mathbb{1}_{\{\tau \le k\}} U_\tau.$$

$$\mathbb{E}[U_{k+1\wedge\tau} \mid \mathcal{F}_k] + \mathbb{1}_{\{\tau > k\}} Z_k \le U_{k\wedge\tau} + \mathbb{1}_{\{\tau > k\}} Y_k.$$

$$\Rightarrow \quad \mathbb{E}[U_{k+1\wedge\tau}] + \mathbb{E}[\mathbb{1}_{\{\tau > k\}} Z_k] \le \mathbb{E}[U_{k\wedge\tau}] + \mathbb{E}[\mathbb{1}_{\{\tau > k\}} Y_k]$$

Sum $k = 0$ to m.

$$\Rightarrow \mathbb{E}[U_{m+1\wedge\tau}] + \mathbb{E}\Big[\sum_{k=0}^{m\wedge(\tau-1)} Z_k\Big] \le \mathbb{E}[U_0] + \mathbb{E}\Big[\sum_{k=0}^{m\wedge(\tau-1)} Y_k\Big].$$

$$\Rightarrow$$

$$\mathbb{E}[\mathbb{1}_{\{\tau < \infty\}} U_{m+1\wedge\tau}] + \mathbb{E}\Big[\sum_{k=0}^{m\wedge(\tau-1)} Z_k\Big] \le \mathbb{E}[U_0] + \mathbb{E}\Big[\sum_{k=0}^{m\wedge(\tau-1)} Y_k\Big].$$

We can use Fatou for the left hand side and the monotone conv. thm to establish

$$\mathbb{E}[\mathbb{1}_{\{\tau < \infty\}} U_\tau] + \mathbb{E}\Big[\sum_{k=0}^{\tau-1} Z_k\Big]$$

$$\le \liminf_{m\to\infty} \mathbb{E}[\mathbb{1}_{\{\tau < \infty\}} U_{m+1\wedge\tau}] + \mathbb{E}\Big[\sum_{k=0}^{m\wedge(\tau-1)} Z_k\Big]$$

$$\le \liminf_{m\to\infty} \mathbb{E}[U_0] + \mathbb{E}\Big[\sum_{k=0}^{m\wedge(\tau-1)} Y_k\Big] = \mathbb{E}[U_0] + \mathbb{E}\Big[\sum_{k=0}^{\tau-1} Y_k\Big].$$

Suppose that P satisfies the Lyap drift cond°

$PV \le V - g$ on $C^c \subset \mathcal{X}$ for some $V: \mathcal{X} \to \mathbb{R}_+$ and $g: \mathcal{X} \to \mathbb{R}_+$.

$C \in \mathcal{X}$ is generic not necessarily small.

V is called a Lyapunov func°.

$\Rightarrow \forall x$, $\mathbb{E}_x[V(X_{n+1}) \mid \mathcal{F}_n] = \mathbb{E}_{X_n}[V(X_1)] = PV(X_n)$.

$\leq (V-g)(X_n)\mathbb{1}\{X_n \in C\} + PV(X_n)\mathbb{1}\{X_n \in C\}$

$\leq V(X_n) - g(X_n) + \mathbb{1}\{X_n \in C\}\{PV(X_n) + g(X_n)\}$.

$\Rightarrow$ We may apply the previous result for τ_C.

$U_n = V(X_n)$ $Z_n = g(X_n)$ $Y_n = \mathbb{1}\{X_n \in C\}\{PV(X_n) + g(X_n)\}$

We deduce the following proposition

Prop: Suppose that P satisfies the Lyap drift condition

$PV \leq V - g$ on $C^c \in \mathcal{X}$ for some $V: \mathcal{X} \to \mathbb{R}_+$ and

$g: \mathcal{X} \to \mathbb{R}_+$.

Then $\forall x \in \mathcal{X}$,

$$\mathbb{E}_x\left[\mathbb{1}\{\tau_C < \infty\} V(X_{\tau_C})\right] + \mathbb{E}_x\left[\sum_{k=0}^{\tau_C - 1} g(X_k)\right]$$

$$\leq \underbrace{\mathbb{E}_x[V(X_0)]}_{V(x)} + \mathbb{E}_x\left[\sum_{k=0}^{\tau_C - 1} \mathbb{1}\{X_k \in C\}\{PV(X_k) + g(X_k)\}\right]$$

$$\leq V(x).$$

$$\mathbb{E}_x\left[\mathbb{1}\{\tau_C < \infty\} V(X_{\tau_C})\right] + \mathbb{E}_x\left[\sum_{k=0}^{\tau_C - 1} g(X_k)\right] \leq V(x)$$

Rk - Corollary: Under the same setting, if $g(x) \geq c > 0$ $\forall x$

$\Rightarrow \mathbb{E}_x\left[\sum_{k=0}^{\tau_C - 1} g(X_k)\right] \geq c\,\mathbb{E}_x[\tau_C]$ $\forall x$

$\Rightarrow \mathbb{E}_x[\tau_C] \leq V(x)/c < \infty$ $\forall x$.

Prop: Suppose again that $PV \leq V - g$ on C^c, for some $C \in \mathcal{X}$.

Then $\forall x$

$$\mathbb{E}_x\left[\mathbb{1}\{\sigma_C < \infty\} V(X_{\sigma_C})\right] + \mathbb{E}_x\left[\sum_{k=0}^{\sigma_C - 1} g(X_k)\right]$$

$$\leq \begin{cases} V(x) & \text{if } x \notin C \\ PV(x) + g(x) & \text{otherwise.} \end{cases}$$

proof: If $x \notin C$ $\mathbb{P}_x$-a.s. $\sigma_C = \tau_C$, the result is a corollary of the previous one.

Recall $\sigma_c = 1 + \tau_c \circ \theta$.

By the (weak) Markov prop, we get that

$$A = \mathbb{E}_x[\mathbb{1}_{\{\sigma_c < \infty\}} V(X_{\sigma_c})] = \mathbb{E}_x[\mathbb{E}_x[\mathbb{1}_{\{1+\tau_c\circ\theta < \infty\}} V(X_{\tau_c})\circ\theta \mid \mathcal{F}_1]]$$

$$= \mathbb{E}_x[\mathbb{E}_{X_1}[\mathbb{1}_{\{\tau_c < \infty\}} V(X_{\tau_c})]].$$

$$B = \mathbb{E}_x\left[\sum_{k=0}^{\tau_c - 1} g(X_k)\right] = \mathbb{E}_x\left[\sum_{k=0}^{\infty} \mathbb{1}_{\{\sigma_c > k\}} g(X_k)\right].$$

$$= \mathbb{E}_x\left[\sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_c \circ \theta > k-1\}} g(X_{k-1}\circ\theta) + g(x)\right].$$

by Markov

$$= \mathbb{E}_x\left[\mathbb{E}_{X_1}\left[\sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_c > k-1\}} g(X_{k-1})\right] + g(x)\right].$$

(1) $A + B = \mathbb{E}_x[H(X_1) + g(x)] = PH(x) + g(x)$

$$H(x) = \mathbb{E}_x[\mathbb{1}_{\{\tau_c < \infty\}} V(\tau_c)] + \mathbb{E}_x\left[\sum_{k=0}^{\tau_c - 1} g(X_k)\right].$$

The previous remark implies $H(x) \le V(x) \quad \forall x$.

(2) $\Rightarrow PH(x) \le PV(x)$.

(1) + (2) $\Rightarrow A + B \le PV(x) + g(x)$.

Prop: Suppose again that $PV \le V - g$ on C^c, for some $C \in \mathcal{X}$.

Suppose now $\sup_{x \in C} PV(x) + g(x) < \infty$

Then

1) $\forall x \quad \mathbb{E}_x[\tau_C] < \infty$ if $g(x) \ge c > 0 \quad \forall x$

2) $\pi(g) = \int g \, d\pi < \infty$ $\forall$ inv prob π of P.

proof: 1) is immediate

2) Combining the two cond°

$PV \le V - g + b \mathbb{1}_C$ where $b = \sup_{x \in C} PV(x) + g(x)$.

Intuition: If we iterate P in this ineq

$$\mathbb{E}_x[V(X_{n+1})] = \mathbb{E}_x[PV(X_n)] \leq \mathbb{E}_x[V(X_n) - g(X_n) + b\mathbb{1}_C(X_n)]$$

Sum over n and divide N, we get

$$(*). \quad \frac{1}{N}\mathbb{E}_x\left[\sum_{n=0}^{N-1} g(X_n)\right] + \mathbb{E}[V(X_{N+1})/N] \leq \frac{V(x)}{N} + b$$

$$\frac{1}{N}\sum_{n=0}^{N-1} P^n g(x).$$

If π inv for P we get

$$\pi(g) \leq \pi(V)/N + b \qquad \forall N.$$

Pb: $\pi(V)$ may be ∞.

Instead $(*)$, we deduce in fact from it that $\forall M$.

$$\frac{1}{N}\sum_{n=0}^{N-1} P^n g_M(x) \leq \left(\frac{1}{N}\sum_{n=0}^{N-1} P^n g(x)\right) \wedge M \qquad \forall x.$$

where $g_M(x) = g(x) \wedge M$

$$\leq \left(\frac{V(x)}{N} + b\right) \wedge M \qquad \forall x$$

Taking int wrt π: $\forall M, N$

$$\pi(g_M) \leq \pi(V_M/N) + b \quad \text{where } V_M = V \wedge M$$

By the dominated conv $\forall M \lim_N \pi(V_M/N) = 0$

we deduce $\forall M \quad \pi(g_M) \leq b$

we conclude by the monotone conv thm.

—

RR: Basically if we have a "simple" Lyap cond

if $g(x) \geq c > 0$. Then, $\forall x \quad \mathbb{E}_x[\tau_C] < \infty$

$\Rightarrow \mathbb{P}_x(N_C = \infty) = 1$

Prop: Suppose $PV \leq V - c$ on C^c and $\sup_{x \in C} PV(x) + g(x) < \infty$

and $g(x) \geq c > 0 \quad \forall x$, $\Rightarrow V(x) \geq c > 0$ on C^c.

In addition, suppose that C is a small set.

$$P(x, \cdot) \geq \mathbb{1}_C(x)\, \varepsilon\, \nu(\cdot) \quad \text{for } \nu \text{ probability measure on } (\mathcal{X}, \mathcal{X})$$

Then using the results of the previous chapter

$$\tilde{E}_{(x,0)}[\tau] < \infty \qquad \tilde{E}_{\nu \otimes \delta}[\tau] < \infty$$

proof:

Recall we have defined $Q(x, A) = P(x, A) - \mathbb{1}_C(x)\, \varepsilon\, \nu(A)$.

and shown $E_{(x,0)}[\tau] = \sum_{i \geq 0} Q^i \mathbb{1}(x)$

It remains to show that the sum converges

Note now $Q V(x) = PV(x) - \mathbb{1}_C(x)\, \varepsilon\, \nu(V)$

$$= PV(x) - \varepsilon\, \nu(V) \quad \forall x \in C.$$

If $\underline{C^c = \mathcal{X}} + V > 0$, we can conclude

$$Q V(x) \leq V(x) - g(x) - \varepsilon\, \nu(V) \leq V(x) - \varepsilon\, \nu(V)$$

$$\Rightarrow Q V(x) + \varepsilon\, \nu(V) \leq V(x)$$

Applying Q^n and summing

$$\varepsilon\, \nu(V) \sum_{k=0}^{n-1} Q^k \mathbb{1} \leq V(x) - Q^n V(x)$$

$$\sum_{k \geq 0} Q^k \mathbb{1} \leq V(x) / \varepsilon\, \nu(V) \quad \forall x$$

$$< \infty.$$

If $C \neq \mathcal{X}$ and $\forall V$, we still have on C^c

$$Q V(x) \leq V(x) - g(x) \leq V(x) - c$$

on C

$$Q V(x) \leq PV(x) - \varepsilon\, \nu(V).$$

$\Rightarrow$ For any $A \geq 0$.

$$Q(V + \mathbb{1} A)(x) \leq (V(x) - g(x) + \mathbb{1}(x) A)\, \mathbb{1}\{x \in C^c\}$$

$$+ (PV(x) - \varepsilon\, \nu(V) + (1 - \varepsilon)\, \mathbb{1}(x) A)\, \mathbb{1}\{x \in C\}$$

using $Q \mathbb{1}(x) = 1 - \varepsilon$ for $x \in C$.

$\int$ now A is large enough

(*) $\quad QV(x) + \mathbb{1}_A(x) \le V(x) + \mathbb{1}_A(x) - \tilde{c} \quad$ for $\tilde{c} > 0$.

Indeed we have

$$(V(x) - g(x) + \mathbb{1}_{(x)} A)\, \mathbb{1}_{\{x \in C^c\}} \le (V(x) - c + \mathbb{1}_{(x)} A)\, \mathbb{1}_{\{x \in C^c\}}.$$

$$(PV(x) - \varepsilon \nu(V) + (1-\varepsilon)\, \mathbb{1}_{(x)} A)\, \mathbb{1}_{\{x \in C\}}.$$

$$\le (\sup_y PV(y) + (1-\varepsilon)\, \mathbb{1}_{(x)} A)\, \mathbb{1}_{\{x \in C\}}.$$

$$\le A\, \mathbb{1}_{\{x \in C\}} - \frac{\varepsilon}{2} \quad \text{if} \quad A = 2 \sup_y PV(y) / \varepsilon$$

$$\le (V(x) + A)\, \mathbb{1}_{\{x \in C\}} - \varepsilon/2$$

Then we have shown (*) holds for the chosen A and $\tilde{c} = c \wedge \varepsilon/2$.

We then have

$$QV_A(x) \le V_A(x) - \tilde{c} \quad \forall x.$$

and therefore applying Q^m and summing, we get

$$\sum_{j \ge 0} Q^j \mathbb{1}(x) \le V_A(x) / \tilde{c} \quad \forall x.$$

For the last point, leave as an exercise.

Chap: Law of large numbers for MC.

We show LLN for MC based on Birkhoff ergodic theorem, which is a result from dynamical systems.

① Dynamical system and ergodicity.

Def: Let $(E, \mathcal{A}, \lambda)$ be a probability space.
A meas map $T: E \to E$ is said to be meas preserving (for λ) if

$$\lambda \circ T^{-1} = \lambda. \qquad (*)$$

$(E, \mathcal{A}, \lambda, T)$ is call a dynamical system.

Rk. 1) Condi° $(*) \iff \lambda(T^{-1}(A)) = \lambda(A) \quad \forall A \in \mathcal{A}$.

$\iff \forall g$ meas and bounded $\int g \circ T \, d\lambda = \int g \, d\lambda$.

2) If T is meas preserv, $T^m = \underbrace{T \circ \dots \circ T}_{m \text{ times}}$ is also meas pres.

3) If T is a bij with meas inverse, T^{-1} is also meas preserving.

Example: Consider $E = \mathcal{X}^{\mathbb{N}}$ $\mathcal{A} = \mathcal{X}^{\otimes \mathbb{N}}$, then for some choice of distribu° on E, the shift is meas preserv.

Prop: Consider either

① $\lambda = \mu^{\otimes \mathbb{N}}$ for μ prob meas on $\mathcal{X}$.

② $\lambda = \mathbb{P}_\pi^P$ for a Markov kernel P and π an invariant distribu° for P.

Then $(\mathcal{X}^{\mathbb{N}}, \mathcal{X}^{\otimes \mathbb{N}}, \lambda, \Theta)$ is a dyn system.

More generally if the canonical process $(X_k)_k$ is stationnary under λ, then $(\mathbb{X}^{\mathbb{N}}, \mathcal{X}^{\otimes \mathbb{N}}, \lambda, \theta)$ is a dyn system.

proof: $(X_k)_k$ statio under $\lambda \iff \forall m$ and k.

$$\text{Law}(X_0, \dots X_m) = \lambda \circ (X_0 \dots X_m)^{-1}$$

$$= \text{Law}(X_k, \dots, X_{m+k}) = \lambda \circ \theta_k^{-1} \circ (X_0, \dots, X_m)^{-1}.$$

Therefore this implies that λ and $\lambda \circ \theta^{-1}$ admits the same finite dim distribution $\Rightarrow \lambda = \lambda \circ \theta^{-1}$.

RR: The proof also implies if $(\mathbb{X}^{\mathbb{N}}, \mathcal{X}^{\otimes \mathbb{N}}, \lambda, \theta)$ is a dynamical system $\Rightarrow (X_k)_k$ is statio under λ.

Def: Consider $(E, \mathcal{A}, \lambda, T)$ be a dyn syst
we say that $A \in \mathcal{A}$ is T-invariant or an invariant set / event if $T^{-1}(A) = A$.
we denote

$$I_T = \{ A \in \mathcal{A} : A \text{ is } T\text{-invariant} \}.$$

$$I_T^\lambda = \{ A \in \mathcal{A} : \lambda(A \Delta T^{-1}(A)) = 0 \}$$

where $A \Delta B = (A \cap B^c) \cup (A^c \cap B)$

Prop: I_T and I_T^λ are sub-σ-field of $\mathcal{A}$ and $I_T \subset I_T^\lambda$.

proof: $\emptyset \in I_T$

$A \in I_T$ using $T^{-1}(A^c) = T^{-1}(A)^c \Rightarrow A^c \in I_T$

Similarly we show if $(A_n)_n \in I_T^{\mathbb{N}} \Rightarrow \bigcup_n A_n \in I_T$.

Def: A meas map / r.v $Y : \Omega \to \mathbb{R}$ is said to be T-inv if $Y \circ T = Y$

Prop: (1) Y inv $\iff$ Y is $\mathcal{I}_T / \mathcal{B}(\mathbb{R})$ - meas

(2) if $Y = Y \circ T$ λ-a.e then $\exists\, \tilde{Y}$ T-inv such that $Y = \tilde{Y}$ λ-a.e

As a consequ $\forall A \in \mathcal{I}_T^{\lambda}$, $A = \tilde{A} \cup N$ for $\tilde{A} \in \mathcal{I}_T$ and N such that $\lambda(N) = 0$.

proof:

(1) $\Rightarrow$ Let $A \in \mathcal{B}(\mathbb{R})$ and let's show $Y^{-1}(A) \in \mathcal{I}_T$.

$$T^{-1}(Y^{-1}(A)) = (Y \circ T)^{-1}(A) \underset{Y \text{ inv}}{=} Y^{-1}(A)$$

$$\Rightarrow Y^{-1}(A) \in \mathcal{I}_T.$$

$\Leftarrow$ If Y is $\mathcal{I}_T / \mathcal{B}(\mathbb{R})$ - meas $\exists$ a seq of simple $\mathcal{I}_T$ - meas such that $e_n \to Y$ pointwise.

Yet if e is a simple $\mathcal{I}_T$ - meas using $\mathbb{1}_A \circ T = \mathbb{1}_A$ for $A \in \mathcal{I}_T$.

we have $e \circ T = e$

Taking the limit, we get $Y \circ T = Y$

(2)

RR: we may construct "easily" inv r.v

If Z is a r.v $\limsup Z \circ T^n$ and $\liminf Z \circ T^n$ are two inv r.v.

The we set $\tilde{Y} = \left(\limsup_n Y \circ T^n\right) \mathbb{1}_{\{\limsup Y \circ T^n < \infty\}}$

By assump° λ-a.e $Y \circ T^n = Y \Rightarrow \lambda$-a.e $Y = \tilde{Y}$.

In addi° $\tilde{Y}$ is a measurable of $\limsup_n Y \circ T^n$ so inv.

Def: A dynamical system is said ergodic if

$\forall A \in \mathcal{I}_\theta \quad \lambda(A) \in \{0, 1\}$.

Example of ergodic dyn syst.

If $(X^{\mathbb{N}}, \mathcal{X}^{\otimes \mathbb{N}}, \mu^{\otimes \mathbb{N}}, \theta)$, then the σ-field of invariant sets satisfies $A \in \mathcal{I}_\theta \Rightarrow \mu^{\otimes \mathbb{N}}(A) \in \{0, 1\}$

and so the dyn system is ergodic.

Lem: If we consider $\mathcal{T} = \bigcap_{n \in \mathbb{N}} \sigma(X_k : k \geq n)$ the tail σ-field. Then

$\mathcal{I}_\theta \subset \mathcal{T}$.

Proof: We show $\mathcal{I}_\theta \in \sigma(X_k : k \geq n) \quad \forall n$.

$$\sigma(X_k, k \geq n) = \sigma(X_k \circ \theta_n : k \geq 0) \quad \forall n$$

$$= \theta_n^{-1}(\sigma(X_k : k \geq 0)) \quad \forall n.$$

$$= \theta_n^{-1}(\mathcal{X}^{\otimes \mathbb{N}}). \quad \forall n.$$

If not $A \in \mathcal{I}_\theta \quad A = \theta_n^{-1}(A) \quad \forall n$

$$\in \theta_n^{-1}(\mathcal{X}^{\otimes \mathbb{N}}) = \sigma(X_k : k \geq n)$$

$$\Rightarrow \mathcal{I}_\theta \in \sigma(X_k : k \geq n) \quad \forall n.$$

—

Now, by the 0-1 Kolmogorov law which implies

if $A \in \mathcal{T}$ then $\mu^{\otimes \mathbb{N}}(A) \in \{0, 1\}$

$\Rightarrow \forall A \in \mathcal{I}_\theta \quad \mu^{\otimes \mathbb{N}}(A) \in \{0, 1\}$

Remind the proof of the 0-1 Kolm law. It consists in showing $\mathcal{T} \perp\!\!\!\perp \mathcal{T}$

If $A \in \mathcal{T}$, then $\forall$ cylinder $C = \bigcap_{i=0}^{m} X_i^{-1}(A_i) \quad A_i \in \mathcal{X}$

$$\mu^{\otimes \mathbb{N}}(A \cap C) = \mathbb{E}_{\mu^{\otimes \mathbb{N}}}\left[\mathbb{1}_A((X_k)_k) \prod_{i}^{m} \mathbb{1}_{A_i}(X_i) \right]$$

since $(X_i)_i$ are indep under $\mu^{\otimes \mathbb{N}}$ and $A \in \mathcal{T}$

$\Rightarrow \mu^{\otimes\mathbb{N}}(A \cap C) = \mathbb{E}_{\mu^{\otimes\mathbb{N}}}[\mathbb{1}_A]\,\mathbb{E}_{\mu^{\otimes\mathbb{N}}}[\mathbb{1}_C]$.

$\Rightarrow$ By the π-λ Thm that in fact $\forall B \in \mathcal{X}^{\otimes\mathbb{N}}$

$$\mu^{\otimes\mathbb{N}}(A \cap B) = \mu^{\otimes\mathbb{N}}(A)\,\mu^{\otimes\mathbb{N}}(B).$$

Any $A \in \mathcal{T} \perp\!\!\!\perp B \in \mathcal{X}^{\otimes\mathbb{N}}$.

$$\Rightarrow \mu^{\otimes\mathbb{N}}(A \cap A) = \mu^{\otimes\mathbb{N}}(A)^2$$

$$\mu^{\otimes\mathbb{N}}(A)$$

$\Rightarrow \mu^{\otimes\mathbb{N}}(A) \in \{0, 1\}$.

Thm: If π is the unique invariant distrib of P

Then $(\mathcal{X}^{\mathbb{N}}, \mathcal{X}^{\otimes\mathbb{N}}, \mathbb{P}_\pi^P, \Theta)$ is ergodic.