

# Stochastic Processes

Lectured by [Weijun Xu](#)

L<sup>A</sup>T<sub>E</sub>Xed by [Chengxin Gong](#)

March 23, 2023

## 目录

<a href="#">1</a>	<a href="#">Review of Martingales</a>	<a href="#">2</a>
<a href="#">2</a>	<a href="#">Markov Chains</a>	<a href="#">3</a>
<a href="#">3</a>	<a href="#">Ergodic Theorem</a>	<a href="#">4</a>

# 1 Review of Martingales

- $(X_n)_{n \geq 0}$  is  $L^2$ -bounded martingale  $\Rightarrow X_n$  converges in  $L^2$ .
- $(X_n)_{n \geq 0}$  is  $L^1$ -bounded martingale  $\Rightarrow X_n$  converges a.s.
- (1) + (2): If  $(X_n)_{n \geq 0}$  is  $L^p$ -bounded martingale for  $p > 1$ , then  $X_n$  converges in  $L^{p'}$  for  $p' \in [1, p)$ .
- Statement is false when  $p = 1$ . Example:  $\Omega = [0, 1)$ ,  $\mathcal{F}_n = \sigma\{\frac{i}{2^n}, \frac{i+1}{2^n}\}_{i=0}^{2^n-1}$ ,  $X_n(\omega) := \begin{cases} 2^n & \omega \in [0, \frac{1}{2^n}) \\ 0 & \text{otherwise} \end{cases}$ .
- Let  $p > 1$  and  $(X_n)_{n \geq 0}$  be  $L^p$  bounded martingale w.r.t.  $\mathcal{F}_n$ . Then  $\exists X \in L^p(\Omega, \mathcal{F}_\infty, P)$  s.t.  $X_n \rightarrow X$  in  $L^p$  and a.s. and  $X_n = \mathbb{E}(X|\mathcal{F}_n)$ .
- Let  $(Z_n)_{n \geq 0}$  be a nonnegative sub-martingale and  $Z_n^* = \sup_{0 \leq k \leq n} Z_k$ , then  $\mathbb{P}(Z_n^* > \lambda) \leq \frac{1}{\lambda} \mathbb{E}(Z_n 1_{\{Z_n^* > \lambda\}}) \leq \frac{1}{\lambda} \mathbb{E}Z_n$ . Corollary:  $\mathbb{P}(Z_n^* > \lambda) \leq \frac{1}{\lambda^p} \mathbb{E}(Z_n^p 1_{\{Z_n^* > \lambda\}}) \leq \frac{1}{\lambda^p} \mathbb{E}(Z_n^p)$ .
- Doob's maximal inequality: Let  $p > 1$ ,  $\exists C = C_p$  s.t.  $\forall$  martingale  $(X_n)_{n \geq 0}$ , we have  $\mathbb{E}|X_n^*|^p \leq C_p \mathbb{E}|X_n|^p$  where  $|X_n^*| = \sup_{0 \leq k \leq n} |X_k|$ .
- If  $(X_n)_{n \geq 0}$  is a martingale with  $\sup_n \mathbb{E}(|X_n| \log(1 + |X_n|)) < +\infty$ , then  $X_n$  converges in  $L^1$ .

*Proof*  $\mathbb{E}|X_n^*| = \int_0^{+\infty} \mathbb{P}(|X_n^*| > \lambda) d\lambda \leq 1 + \int_1^{+\infty} \frac{1}{\lambda} (\int_{X_n^* > \lambda} |X_n| d\mathbb{P}) d\lambda = 1 + \int |X_n| 1_{X_n^* > 1} (\int_1^{X_n^*} \frac{1}{\lambda} d\lambda) d\mathbb{P} \Rightarrow \mathbb{E}|X_n^*| \leq 1 + \mathbb{E}(|X_n| \log(X_n^* \vee 1)) \Rightarrow \mathbb{E}(X_n^* \vee 1) \leq 2 + \mathbb{E}(|X_n| \log(X_n^* \vee 1))$ . Since  $x \log y \leq 10^{10}(2+x) \log(2+x) + \frac{y}{2}$  when  $x, y$  are large enough (insight: if  $y \gg x^2$  then  $x \log y \leq \frac{y}{2}$ ; else  $x \log y \leq 10^{10}(2+x) \log(2+x)$ ),  $\mathbb{E}X_n^* \leq 10^{100}[1 + \mathbb{E}(|X_n| + 2) \log(|X_n| + 2)]$ . Then use dominated convergence theorem.  $\square$

- Two probability measures  $\mathbb{P}$  and  $\mathbb{Q}$  on  $(\Omega, \mathcal{F})$ ,  $\mathbb{Q} \ll \mathbb{P}$  on  $\mathcal{F}_n$  for every  $n$  and  $M_n = \frac{d\mathbb{Q}|_{\mathcal{F}_n}}{d\mathbb{P}|_{\mathcal{F}_n}}$ .  $(M_n)_{n \geq 0}$  is a  $\mathbb{P}$ -martingale w.r.t.  $(\mathcal{F}_n)_{n \geq 0}$ .  $\mathbb{Q} \ll \mathbb{P}$  on  $\mathcal{F}_\infty$  if and only if  $M_n \rightarrow M$  in  $L^1$ .  $\mathbb{Q}(A) = \int_A M d\mathbb{P} + \mathbb{Q}(A \cap \{M = +\infty\})$ .

*Proof* Sufficiency.  $\mathbb{Q} \ll \mathbb{P}$  on  $\mathcal{F} = \mathcal{F}_\infty$ , thus let  $Z = \frac{d\mathbb{Q}}{d\mathbb{P}}$ , we need to show  $M_n$  converges to  $Z$  in  $L^1$ .  $\forall A \in \mathcal{F}_n$ ,  $\int_A M_n d\mathbb{P} = \mathbb{Q}(A) = \int_A Z d\mathbb{P} \Rightarrow M_n = \mathbb{E}(Z|\mathcal{F}_n)$ . Thus  $M_n$  is uniformly integrable, thus converges in  $L^1$ .

Necessity. Suppose  $M_n \rightarrow M$  a.s. and in  $L^1$ . We need to show  $M_n = \mathbb{E}(M|\mathcal{F}_n)$  and  $M = \frac{d\mathbb{Q}}{d\mathbb{P}}$ . It suffices to show  $\mathbb{Q}(A) = \int_A M d\mathbb{P}$  for all  $A \in \cup_n \mathcal{F}_n$ . Suppose  $A \in \mathcal{F}_N$ . Then  $\mathbb{Q}(A) = \int_A M_N d\mathbb{P} = \int_A M_{N+k} d\mathbb{P} \rightarrow \int_A M d\mathbb{P}$ . By  $\pi - \lambda$  theorem we can get the desired result.

Special situation: Suppose  $\mathbb{P} \perp \mathbb{Q}$  on  $\mathcal{F} (\exists E \text{ s.t. } \mathbb{P}(E) = 1, \mathbb{Q}(E^c) = 1)$  and  $\mathbb{P} \ll \mathbb{Q}$  on  $\mathcal{F}_n$ . Then  $\frac{1}{M_n}$  converges  $\mathbb{Q}$ -a.s. Let  $\mathbb{R} = \frac{1}{2}(\mathbb{P} + \mathbb{Q})$ ,  $\mathbb{P}, \mathbb{Q} \ll \mathbb{R}$  on  $\mathcal{F}$ ,  $\frac{d\mathbb{P}|_{\mathcal{F}_n}}{d\mathbb{R}|_{\mathcal{F}_n}} = \frac{2}{1+M_n} \rightarrow \frac{2M}{1+M}$  in  $L^1(\mathbb{R})$ ,  $\frac{d\mathbb{Q}}{d\mathbb{R}} = \frac{2M_n}{1+M_n} \rightarrow \frac{2}{1+M}$  in  $L^1(\mathbb{R})$ . Then  $\mathbb{Q}(A) = \mathbb{Q}(A \cap E^c) = \int_{A \cap E^c} \frac{2M}{1+M} d\mathbb{R} = \int_A \frac{2M}{1+M} 1_{E^c} d\mathbb{R} \stackrel{\mathbb{P}(E^c)=0}{=} 2\mathbb{R}(A \cap E^c) = 2 \int_A 1_{E^c} d\mathbb{R} \Rightarrow \int_A \frac{2M}{1+M} 1_{E^c} d\mathbb{R} = 2 \cdot 1_{E^c} \Rightarrow M = +\infty$  on  $E^c \Rightarrow \mathbb{Q}(M = +\infty) = 1$ . Similarly  $\mathbb{P}(M = 0) = \mathbb{Q}(M = +\infty) = 1$ .

General situation:  $\mathbb{Q} = \mathbb{Q}_1 + \mathbb{Q}_2$ ,  $\mathbb{Q}_1 \ll \mathbb{P}$ ,  $\mathbb{Q}_2 \perp \mathbb{P}$  on  $\mathcal{F}$ . Therefore we can decompose  $M_n$  as  $M_n = Y_n + Z_n$  where  $Y_n \rightarrow Y$  in  $L^1(\mathbb{P})$  and  $Z_n \rightarrow 0$   $\mathbb{P}$ -a.s.  $\mathbb{Q}_1(A) = \int_A Y d\mathbb{P} = \int_A M d\mathbb{P}$ .  $\mathbb{Q}_2(A) = \mathbb{Q}_2(A \cap \{Z = +\infty\})$ . Since  $Z = 0$   $\mathbb{P}$ -a.s.,  $M < +\infty$   $\mathbb{P}$ -a.s. and  $\mathbb{Q}_2(M = +\infty) = 1$ , we have  $\mathbb{Q}_2(A) = \mathbb{Q}(A \cap \{Z = +\infty\}) = \mathbb{Q}_2(A \cap \{M = +\infty\}) = \mathbb{Q}(A \cap \{M = +\infty\})$ . To sum up,  $\mathbb{Q}(A) = \int_A M d\mathbb{P} + \mathbb{Q}(A \cap \{M = +\infty\})$ .  $\square$

- Statement is false if  $M_n \not\rightarrow M$  in  $L^1$ . Example:  $\Omega = \{\omega = (\omega_1, \dots, \omega_n, \dots) \in \{\pm 1\}^{\mathbb{N}}\}$ ,  $X_n(\omega) = \omega_n$ .  $X_n$ 's are i.i.d. under  $\mathbb{P}$  and  $\mathbb{Q}$ , but  $\mathbb{P}(X_n = 1) = \frac{1}{2}$ ,  $\mathbb{P}(X_n = -1) = \frac{1}{2}$ ,  $\mathbb{Q}(X_n = 1) = \frac{1}{3}$ ,  $\mathbb{Q}(X_n = -1) = \frac{2}{3}$ .  $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$ .  $\mathbb{P}(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n X_k = 0) = 1$ ,  $\mathbb{Q}(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n X_k = -\frac{1}{3}) = 1$ .
- Monotone class theorem for functions: Suppose  $\mathcal{A}$  as a  $\pi$ -system and  $\mathcal{H}$  be a class of functions from  $\Omega$  to  $\mathbb{R}$  s.t. (1)  $1_A \in \mathcal{H}$  for every  $A \in \mathcal{A}$ , (2) if  $f, g \in \mathcal{H}$  then  $af + bg \in \mathcal{H}$ , (3) if  $f_n \in \mathcal{H}$  and  $f_n \uparrow f$  then  $f \in \mathcal{H}$ . Then all nonnegative  $\sigma(\mathcal{A})$ -measurable functions are in  $\mathcal{H}$ .
- Let  $(Y_n)_{n \geq 0}$  be i.i.d., nonnegative r.v.'s with  $\mathbb{E}Y_k = 1$ . Then  $M_n = \prod_{k=1}^n Y_k$  converges in  $L^1$  iff  $Y_n \equiv 1$ . Otherwise  $M_n \rightarrow 0$  a.s.

*Proof* Note that  $\frac{1}{n} \log M_n = \frac{1}{n} \sum_{k=1}^n \log Y_k \rightarrow \mathbb{E} \log Y$  a.s. If  $\mathbb{E} \log Y = 0$  then by Jensen's inequality we have  $Y_n \equiv 1$  which means  $M_n$  converges in  $L^1$ . If  $\mathbb{E} \log Y < 0$  then  $M_n \rightarrow 0$  a.s.  $\square$

- Kakutani's theorem:  $M_n = \prod_{k=1}^n Y_k$ ,  $Y_k \geq 0$  are independent,  $\mathbb{E}Y_k = 1$ ,  $\lambda_k = \mathbb{E}\sqrt{Y_k}$ . (1) If  $\prod_k \lambda_k > 0$ , then  $M_n \rightarrow M$  in  $L^1$ ; (2) If  $\prod_k \lambda_k = 0$ , then  $M_n \rightarrow 0$  a.s.

*Proof* Let  $Z_n = \prod_{k=1}^n \frac{\sqrt{Y_k}}{\lambda_k}$ . Then  $Z_n$  is a martingale and has an a.s. limit  $Z$ , and  $M_n = (\prod_{k=1}^n \lambda_k)^2 Z_n^2$ . If  $\prod_k \lambda_k > 0$ , then  $Z_n$  is  $L^2$  bounded and then convergence in  $L^2$ , which implies  $M_n \rightarrow M$  in  $L^1$ . If  $\prod_k \lambda_k = 0$ , it is obvious that  $M_n \rightarrow 0$  a.s.  $\square$

- Martingale LLN: Let  $(M_n)_{n \geq 0}$  be a martingale s.t.  $\sum_{k=1}^{+\infty} \frac{\mathbb{E}(M_k - M_{k-1})^2}{k^2} < +\infty$ . Then  $\frac{M_n}{n} \rightarrow 0$  a.s.

*Proof* Let  $Y_n = \sum_{k=1}^n \frac{X_k}{k}$ . Then  $(Y_n)_{n \geq 0}$  is an  $L^2$  bounded martingale, thus  $Y_n \rightarrow Y$  a.s. Then use Kronecker's lemma.  $\square$

- Martingale CLT: Let  $(M_n)_{n \geq 0}$  be a martingale with  $M_0 = 0$  and  $\sigma_n^2 = \sum_{k=1}^n \mathbb{E}X_k^2 = \mathbb{E}\langle M \rangle_n$ . Assume that  $\frac{1}{\sigma_n^2} \max_{1 \leq k \leq n} (\mathbb{E}X_k^2) \rightarrow 0$ ,  $\frac{1}{\sigma_n^2} \sum_{k=1}^n \mathbb{E}(X_k^2 1_{\{|X_k| > \epsilon \sigma_n\}} | \mathcal{F}_{k-1}) \xrightarrow{P} 0$  for all  $\epsilon > 0$ ,  $\frac{1}{\sigma_n^2} \langle M \rangle_n \xrightarrow{P} 1$ . Then  $\frac{M_n}{\sigma_n} \Rightarrow \mathcal{N}(0, 1)$ .

## 2 Markov Chains

- Let  $(X_n)_{n \geq 0}$  be a homogeneous Markov chain on a discrete space  $S$ .  $\mathbb{P}^x$ : law of  $(X_n)_{n \geq 0}$  conditioned on  $X_0 = x$ .  $\mathbb{P}(X_{n+1} \in A | \mathcal{F}_n) = \mathbb{P}^{X_n}(X_1 \in A) = \mathbb{P}(X_1 \in A | X_0 = X_n)$ .  $\mathbb{E}^x$ : expectation under  $\mathbb{P}^x$ .  $\mathbb{P}^x(X_1 = y) = p(x, y)$ .
- For every  $f : S \rightarrow \mathbb{R}$  bounded, define  $(\mathcal{P}f)(x) = \sum_{y \in S} p(x, y)f(y) = \mathbb{E}^x(f(X_1))$ ,  $(\mathcal{L}f)(x) = \sum_{y \in S} p(x, y)f(y) - f(x)$ .  $\mathcal{L} = \mathcal{P} - \text{id}$ , the generator.
- Let  $(X_n)_{n \geq 0}$  be a homogeneous Markov chain with generator  $\mathcal{L}$ . Then for every bounded  $f : S \rightarrow \mathbb{R}$ ,  $M_n = f(X_n) - f(X_0) - \sum_{k=0}^{n-1} (\mathcal{L}f)(X_k)$  is a martingale. Conversely, let  $(X_n)_{n \geq 0}$  be a process and  $\mathcal{L}$  be an operator on  $\mathcal{B}(S)$  s.t.  $M_n^f$  is a martingale for every  $f$ , then  $(X_n)_{n \geq 0}$  is a Markov chain with generator  $\mathcal{L}$ .
- Given operator  $\mathcal{L}$  on  $\mathcal{B}(S)$ , we say  $f : S \rightarrow \mathbb{R}$  is (1) harmonic for  $\mathcal{L}$  if  $\mathcal{L}f = 0$ ; (2) sub-harmonic for  $\mathcal{L}$  if  $\mathcal{L}f \geq 0$ ; (3) super-harmonic for  $\mathcal{L}$  if  $\mathcal{L}f \leq 0$ .
- Let  $f$  be the generator of a Markov chain  $(X_n)_{n \geq 0}$ . Then  $f$  is (sub-/super-)harmonic  $\Leftrightarrow f(X_n)_{n \geq 0}$  is a (sub-/super-) martingale.
- $f$  is (sub-/super-)harmonic on  $D \subset S$  if  $\mathcal{L}f \geq / \leq / = 0$  on  $D$ . Let  $\tau = \inf\{k \geq 0 : X_k \in D^c\}$ , then  $(f(X_{n \wedge \tau}))_{n \geq 0}$  is a (sub-/super-)martingale.
- Maximum principle: Let  $(X_n)_{n \geq 0}$  be a Markov chain and  $D \subset S$  s.t. the stopping time  $\tau = \inf\{k \geq 0, X_k \in D^c\}$  is a.s. finite. If  $f$  is bounded and sub-harmonic on  $D$ , then  $\sup_{x \in D} f(x) \leq \sup_{x \in D^c} f(x)$ .

*Proof*  $f$  is sub-harmonic implies  $(f(X_{n \wedge \tau}))$  is a sub-martingale, hence for  $x \in D$  we have  $f(x) \leq \mathbb{E}^x(f(X_{n \wedge \tau})) \rightarrow \mathbb{E}^x(f(X_\tau)) \leq \sup_{x \in D^c} f(x)$ .  $\square$

- $A \subset S, \tau_A = \sup\{k \geq 0 : X_k \in A\}$ . (1)  $u(x) = \mathbb{P}^x(\tau_A < +\infty) \Rightarrow \begin{cases} \mathcal{L}u = 0 & \text{on } A^c \\ u = 1 & \text{on } A \end{cases}$ . (2)  $u(x) = \mathbb{P}(\tau_A < \tau_B) \Rightarrow$

$$\begin{cases} \mathcal{L}u = 0 & \text{on } (A \cup B)^c \\ u = 1 & \text{on } A \\ u = 0 & \text{on } B. \end{cases} \quad (3) \quad u(x) = \mathbb{E}^x[\tau_A] \Rightarrow \begin{cases} \mathcal{L}u = -1 & \text{on } A^c \\ u = 0 & \text{on } A \end{cases}.$$

- Any nonnegative solution  $v$  to  $\begin{cases} \mathcal{L}v = 0 & \text{on } A^c \\ v = 1 & \text{on } A \end{cases}$  satisfies  $v \geq u$ . Furthermore, if  $u \equiv 1$ , then  $\exists$  1 bounded solution

$$\text{to } \begin{cases} \mathcal{L}v = 0 & \text{on } A^c \\ v = f & \text{on } A \end{cases} \quad \text{with } v(x) = \mathbb{E}^x(f(X_{\tau_A})).$$

*Proof* Let  $v(x)$  be a non-negative solution, then  $v(X_{n \wedge \tau_A})_{n \geq 0}$  is a martingale.  $v(x) = \mathbb{E}^x v(X_{n \wedge \tau_A}) = \mathbb{E}^x v(X_{n \wedge \tau_A}) 1_{\tau_A < \infty} + \mathbb{E}^x v(X_{n \wedge \tau_A}) 1_{\tau_A = \infty} \geq \mathbb{E} v(X_{n \wedge \tau_A}) 1_{\tau_A < \infty}$ . Let  $n \rightarrow \infty$  and by Fatou's lemma, we have  $v(x) \geq \mathbb{E}^x v(X_{\tau_A}) 1_{\tau_A < \infty} = \mathbb{P}^x(\tau_A < \infty) = u(x)$ . If  $u(x) \equiv 1$  and  $v(x)$  is bounded, then by bounded convergence theorem,  $v(x) = \mathbb{E}^x v(X_{n \wedge \tau_A}) \rightarrow \mathbb{E}^x v(X_{\tau_A}) = \mathbb{E}^x f(X_{\tau_A})$ .  $\square$

- Doob's  $h$ -transform: Let  $h$  be nonnegative, harmonic with  $h(x_0) = 1$  for some  $x_0 \in S$ . Then  $(h(X_n))_{n \geq 0}$  is a martingale with  $\mathbb{E}^{\mathbb{P}^{x_0}}(h(X_n)) = 1$ . Then  $\exists$  1 measure  $\mathbb{Q}^h$  on  $\mathcal{F}_\infty$  s.t.  $\frac{d\mathbb{Q}^h}{d\mathbb{P}^{x_0}|_{\mathcal{F}_n}} = h(X_n), \forall n \geq 0$ .  $\mathbb{Q}^h(X_0 = x_0) = 1$ ,  $(X_n)_{n \geq 0}$  never visits the set  $D = \{x : h(x) = 0\}$ . Under  $\mathbb{Q}^h$ ,  $(X_n)_{n \geq 0}$  is again a Markov chain on  $S \setminus D$  with transition probability  $q(x, y) = \frac{p(x, y)h(y)}{h(x)}$  (or equivalently,  $(\mathcal{L}^h f)(x) = \frac{1}{h(x)}(\mathcal{L}(hf))(x)$ ).

*Proof* The first two props are trivial.  $\mathbb{Q}(X_{n+1} = y | \mathcal{F}_n) = \frac{\mathbb{Q}(X_{n+1}=y, X_n=x_n, \dots, X_0=x_0)}{\mathbb{Q}(X_n=x_n, \dots, X_0=x_0)} = \frac{\int_{\{X_{n+1}=y, X_n=x_n, \dots, X_0=x_0\}} h(X_{n+1}) d\mathbb{P}^{x_0}}{\int_{\{X_n=x_n, \dots, X_0=x_0\}} h(X_n) d\mathbb{P}^{x_0}} = \frac{h(y)\mathbb{P}^{x_0}(X_{n+1}=y, X_n=x_n, \dots, X_0=x_0)}{h(x_n)\mathbb{P}^{x_0}(X_n=x_n, \dots, X_0=x_0)} = \frac{h(y)p(x_n, y)}{h(x_n)}$ . Next we show  $M_n^f := f(X_n) - f(X_0) - \sum_{k=0}^{n-1} (\mathcal{L}^h f)(X_k)$  is a  $\mathbb{Q}$ -martingale for any bounded  $f$ . Let  $Z_n = \mathbb{E}^{\mathbb{Q}} f(X_{n+1}) | \mathcal{F}_n$ .  $\forall A \in \mathcal{F}_n$ ,  $\int_A Z_n h(X_n) d\mathbb{P}^{x_0} = \int_A Z_n d\mathbb{Q} = \int_A f(X_{n+1}) d\mathbb{Q} = \int_A f(X_{n+1}) h(X_{n+1}) d\mathbb{P}^{x_0} = \mathbb{E}^{\mathbb{P}^{x_0}} [\mathbb{E}^{\mathbb{P}^{x_0}}(f(X_{n+1})h(X_{n+1})1_A | \mathcal{F}_n)] = \mathbb{E}^{\mathbb{P}^{x_0}} [1_A \mathbb{E}^{\mathbb{P}^{x_0}}(f(X_{n+1})h(X_{n+1}) | \mathcal{F}_n)] = \int_A \mathcal{P}(hf)(X_n) d\mathbb{P}^{x_0}$ . Thus  $Z_n = \frac{\mathcal{P}(hf)(X_n)}{h(X_n)}$  only depends on  $X_n$ , i.e.  $(X_n)_{n \geq 0}$  is a MC on  $\mathbb{Q}$  with generator  $\mathcal{L}^h$ .  $\square$

- An irreducible Markov chain  $(X_n)_{n \geq 0}$  (1) is transient if  $\exists x$  and  $A \subset S$  s.t.  $\mathbb{P}(\tau_A < \infty | X_0 = x) < 1$ ; (2) is recurrent if  $\exists$  a finite set  $A \subset S$  s.t.  $\mathbb{P}(\tau_A < \infty) = 1$  for all  $x \in S$ . (3) is positive recurrent if  $\exists$  a finite set  $A \subset S$  s.t.  $\mathbb{E}(\tau_A) < \infty$  for all  $x \in S$ .
- Foster-Lyapunov criterion: An irreducible MC on a countable state space  $S$  (1) is transient iff  $\exists v : S \rightarrow \mathbb{R}^+$  and  $A \subset S$  non-empty s.t.  $\mathcal{L}v \leq 0$  on  $A^c$  and  $v(x) < \inf_{y \in A} v(y)$  for some  $x \in A^c$ ; (2) is recurrent iff  $\exists v : S \rightarrow \mathbb{R}^+$  s.t.  $\mathcal{L}v \leq 0$  on  $A^c$  where  $A$  is a finite set and  $\{x : v(x) \leq N\}$  is finite for every  $N$ ; (3) is positive recurrent iff  $\exists v : S \rightarrow \mathbb{R}^+$ ,  $A \subset S$  finite,  $\exists \epsilon > 0$  s.t.  $\mathcal{L}v \leq -\epsilon$  on  $A^c$  and  $\sum_{y \in S} p(x, y)V(y) < +\infty$  for all  $x \in A$ .

*Proof* (1)  $v(X_{n \wedge \tau_A})_{n \geq 0}$  is a super-martingale, hence  $v(x) \geq \mathbb{E}v(X_{n \wedge \tau_A}) \geq \mathbb{E}v(X_{n \wedge \tau_A})1_{\tau_A < \infty}$ . Let  $n \rightarrow \infty$  we know  $v(x) \geq \mathbb{E}v(X_{\tau_A}1_{\tau_A < \infty}) \geq (\inf_{y \in A} v(y))\mathbb{P}(\tau_A < \infty) \Rightarrow \mathbb{P}(\tau_A < \infty) < \frac{v(x)}{\inf_{y \in A} v(y)} < 1$ . (2) On  $\{\tau_A = \infty\}$ ,  $\limsup_{n \rightarrow \infty} v(X_{n \wedge \tau_A}) = +\infty$  a.s. Since  $(v(X_{n \wedge \tau_A}))_{n \geq 0}$  is a nonnegative super-martingale, hence converges a.s., therefore  $\lim_{n \rightarrow \infty} v(X_{n \wedge \tau_A}) = +\infty$  a.s. Note that  $v(x) \geq \mathbb{E}v(X_{n \wedge \tau_A})1_{\tau_A = \infty}$ . Since LHS is a finite number, we have  $\mathbb{P}(\tau_A = \infty) = 0$ . (3)  $\mathbb{E}v(X_{n \wedge \tau_A}) | \mathcal{F}_{n-1} \leq v(X_{(n-1) \wedge \tau_A}) - \epsilon 1_{\tau_A \geq n}$ . Taking expectation on the both sides,  $\mathbb{E}v(X_{n \wedge \tau_A}) \leq \mathbb{E}v(X_{(n-1) \wedge \tau_A}) - \epsilon \mathbb{P}(\tau_A \geq n) \leq \dots \leq v(x) - \epsilon \sum_{k=1}^n \mathbb{P}(\tau_A \geq k) \Rightarrow \mathbb{E}^x \tau_A = \sum_{k=1}^{\infty} \mathbb{P}(\tau_A \geq k) \leq \frac{v(x)}{\epsilon} < \infty$ .

Conversely, (1) Let  $v(x) = \mathbb{P}^x(\tau_A < \infty)$ . (2) Let  $u(x) = \mathbb{P}^x(\tau_B < \tau_A)$ . We have shown that if  $x \in (A \cup B)^c$  then  $\mathcal{L}u \leq 0$ . When  $x \in B$ ,  $(\mathcal{L}u)(x) = \sum_{y \in S} p(x, y)u(y) - 1 \leq 0$ . Take  $B_N \downarrow \emptyset$  s.t.  $B_N^c$  is finite for every  $N$ . Via a diagonal argument  $\Rightarrow \exists$  subsequence  $\{N_k\}$  s.t.  $v(x) := \sum_{k \geq 1} \mathbb{P}^x(\tau_{B_{N_k}} < \tau_A) < +\infty$  for every  $x \in S$ . (3) Let  $v(x) = \mathbb{E}^x \tau_A$ .  $\square$

- e.g.  $h(x) = \frac{\mathbb{P}^x(\tau_A < \tau_B)}{\mathbb{P}^{x_0}(\tau_A < \tau_B)}$  is harmonic on  $(A \cup B)^c$  with  $h(x_0) = 1$  ( $x_0 \in (A \cup B)^c$ ). Then  $\forall x, y \in (A \cup B)^c, q(x, y) = \frac{h(y)p(x, y)}{h(x)} = \frac{\mathbb{P}^y(\tau_A < \tau_B)p(x, y)}{\mathbb{P}^x(\tau_A < \tau_B)} = \frac{\mathbb{P}^x(X_1=y, \tau_A < \tau_B)}{\mathbb{P}^x(\tau_A < \tau_B)} = \mathbb{P}^x(X_1 = y | \tau_A < \tau_B)$ .
- e.g.  $\mathbb{P}$  is simple symmetric random walk on  $\mathbb{Z}$  starting from  $X_0 = 0$ . Question: what is the law of  $(X_n)_{n \geq 0}$  conditioned on  $X_n \geq 0$  for all  $n$ ? Let  $\tau_k = \inf\{n \geq 0, X_n = k\}$ . On  $\{\tau_N < \tau_{-1}\}$ ,  $\frac{h(y)}{h(x)} = \frac{\mathbb{P}^y(\tau_N < \tau_{-1})}{\mathbb{P}^x(\tau_N < \tau_{-1})} = \frac{y+1}{x+1}$ . Thus  $q_N(x, y) = \frac{1}{2} \frac{y+1}{x+1}, |x - y| = 1, x \in \{0, \dots, N-1\} \Rightarrow q(x, y) = \frac{1}{2} \frac{y+1}{x+1}, x \geq 0, |x - y| = 1$ .

### 3 Ergodic Theorem

- Basic setup: a measurable map  $T : (\Omega, \mathcal{F}) \rightarrow (\Omega, \mathcal{F})$ . Examples: (1) circle rotations:  $\Omega = \mathbb{R}/\mathbb{Z}, T : x \mapsto x + \alpha$ ; (2) doubling map:  $\Omega = \mathbb{R}/\mathbb{Z}, x \mapsto 2x$ ; (3) shift map:  $\Omega = S^{\mathbb{N}}, (T\omega)_n = \omega_{n+1}$ .
- Let  $T : (\Omega, \mathcal{F}) \rightarrow (\Omega, \mathcal{F})$  measurable and  $\mathbb{P}$  be a probability measure on  $(\Omega, \mathcal{F})$ . We say  $T$  is measure-preserving if  $\mathbb{P}(T^{-1}(A)) = \mathbb{P}(A)$  for every  $A \in \mathcal{F}$  (or  $\mathbb{P} \circ T^{-1} = \mathbb{P}$ ).
- Question: what if we define by  $\mathbb{P}(T(A)) = \mathbb{P}(A)$  for every  $A \in \mathcal{F}$  instead?  $\mathbb{P} \circ T = \mathbb{P} \Rightarrow \mathbb{P} \circ T^{-1} = \mathbb{P}$  while the converse proposition is false.
- $(X_n)_{n \geq 0}$  be i.i.d.  $\sim \mu$ . We can build  $(\Omega, \mathcal{F}, \mathbb{P})$  and  $X_n : \Omega \rightarrow \mathbb{R}$  measurable s.t.  $(X_n)_{n \geq 0}$  i.i.d.  $\sim \mu$  under  $\mathbb{P}$ : (1)  $\Omega = \mathbb{R}^{\mathbb{N}} = \{\omega : \omega = (\omega_0, \omega_1, \dots)\}$ ; (2)  $X_n(\omega) = \omega_n$ ; (3)  $\mathcal{F} = \sigma(X_0, X_1, \dots, X_n, \dots)$ ; (4)  $\mathbb{P} = \mu^{\otimes \mathbb{N}}$ . It is easy to show that the shift map is measure-preserving:  $\mathcal{F}$  is generated by sets of the form  $A = \{\omega_{k_1} \in I_1, \dots, \omega_{k_N} \in I_N\}$ ,  $T^{-1}(A) = \{\omega : (T\omega)_{k_1} \in I_1, \dots, (T\omega)_{k_N} \in I_N\} = \{\omega : \omega_{k_1+1} \in I_1, \dots, \omega_{k_N+1} \in I_N\}$ . Key: the only thing used is that  $(X_{k_1}, \dots, X_{k_N}) \stackrel{\text{law}}{=} (X_{k_1+1}, \dots, X_{k_N+1})$  for every  $N$  and every  $k_1, \dots, k_N$ .

- A sequence of random variables is stationary if  $(X_n)_{n \in J} \stackrel{\text{law}}{=} (X_{n+k})_{n \in J}$  for all  $k$  and finite set  $J$ .
- Let  $T : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\Omega, \mathcal{F}, \mathbb{P})$  be measure-preserving and  $X : \Omega \rightarrow \mathbb{R}$  be measurable. Then  $X_n(\omega) := X(T^n \omega)$  defines a stationary sequence.
- *Proof* It suffices to show that for every  $N$ , every  $I_1, \dots, I_N \subset \mathbb{R}$  and every  $k_1 < k_2 < \dots < k_N$ , we have  $\mathbb{P}(X_{k_1} \in I_1, \dots, X_{k_N} \in I_N) = \mathbb{P}(X_{k_1+1} \in I_1, \dots, X_{k_N+1} \in I_N)$ .  $\mathbb{P}(\{\omega : X_{k_1}(\omega) \in I_1, \dots, X_{k_N}(\omega) \in I_N\}) = \mathbb{P}(T^{-1}\{\omega : X_{k_1}(\omega) \in I_1, \dots, X_{k_N}(\omega) \in I_N\}) = \mathbb{P}(\{\omega : X_{k_1}(T\omega) \in I_1, \dots, X_{k_N}(T\omega) \in I_N\}) = \mathbb{P}(\{\omega : X_{k_1+1}(\omega) \in I_1, \dots, X_{k_N+1}(\omega) \in I_N\})$ .  $\square$
- Let  $(\Omega, \mathcal{F}, \mathbb{P}, T)$  be a measure-preserving system. (1) A set  $A \in \mathcal{F}$  is invariant if  $\mathbb{P}(A \Delta T^{-1}(A)) = 0$ . (2) A random variable  $X : \Omega \rightarrow \mathbb{R}$  is invariant if  $X = X \circ T$   $\mathbb{P}$ -a.e.
- The collection of invariant sets  $\mathcal{I} = \{A \in \mathcal{F} : A \text{ is invariant}\}$  is a  $\sigma$ -algebra and  $X : \Omega \rightarrow \mathbb{R}$  is invariant iff it is  $\mathcal{I}$ -measurable.
- We say  $T : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\Omega, \mathcal{F}, \mathbb{P})$  measurable-preserving is ergodic if  $\mathbb{P}(A) = 0$  or  $1$  for all  $A \in \mathcal{I}$ .
- Let  $T : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\Omega, \mathcal{F}, \mathbb{P})$  be measure preserving and  $f \in L^p(p \geq 1)$ . Then  $\frac{1}{N} \sum_{k=0}^{N-1} f \circ T^k \rightarrow \mathbb{E}(f|\mathcal{I})$  a.s. and in  $L^p$ . In particular,  $\mathbb{E}(f|\mathcal{I}) = \mathbb{E}f$  if  $T$  is ergodic.

*Proof* We first show **convergence in  $L^p$** .

*Lemma 1* If  $(\Omega, \mathcal{F}, \mathbb{P}, T)$  is a measure-preserving system and  $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ . Then  $\int_{\Omega} X d\mathbb{P} = \int_{\Omega} X \circ T d\mathbb{P}$ . In fact,  $\|X\|_{L^p} = \|X \circ T\|_{L^p}$ ,  $p \in [1, +\infty]$ .

*Proof* Take  $X = 1_A$ . LHS =  $\mathbb{P}(A) = \mathbb{P}(T^{-1}(A)) = \int_{\Omega} 1_A(T\omega) d\mathbb{P}$ .  $\square$

Let  $\mathcal{U}_T : L^p(\Omega, \mathcal{F}, \mathbb{P}) \rightarrow L^p(\Omega, \mathcal{F}, \mathbb{P})$  be defined by  $(\mathcal{U}_T f)(\omega) := f(T\omega)$  (or  $\mathcal{U}_T f = f \circ T$ ).

**For  $p = 2$** ,  $\mathcal{U}_T : L^2 \rightarrow L^2$  is an isometry in the sense that  $\langle f, g \rangle = \langle \mathcal{U}_T f, \mathcal{U}_T g \rangle$ . LHS =  $\frac{1}{N} \sum_{k=0}^N \mathcal{U}_T^k f$ ,  $f = \mathbb{E}(f|\mathcal{I}) + (f - \mathbb{E}(f|\mathcal{I}))$   
 $\Rightarrow$  LHS =  $\underbrace{\mathbb{E}(f|\mathcal{I})}_{\text{Ker}(\mathcal{U}_T - \text{Id})} + \frac{1}{N} \sum_{k=0}^N \mathcal{U}_T^k (f - \mathbb{E}(f|\mathcal{I}))$ . Since  $\mathcal{H} = \text{Ker}(A) \oplus \overline{\text{Im}(A^*)}$ ,  $\exists g \in \mathcal{H}$  s.t.  $\|f - \mathbb{E}(f|\mathcal{I}) - (\mathcal{U}_T^* - \text{Id})g\| < \epsilon$ .

*Lemma 2* Let  $A : \mathcal{H} \rightarrow \mathcal{H}$  be an isometry. If  $Af = f$ , then  $A^*f = f$ .

*Proof*  $\langle A * f, g \rangle = \langle f, Ag \rangle = \langle Af, Ag \rangle = \langle f, g \rangle$   $\square$

*Proposition 1*  $\mathcal{H} = \text{Ker}(A^*) \oplus \overline{\text{Im}(A)}$ .

*Proof* We show that  $\text{Ker}(A^*) = (\text{Im}(A))^{\perp}$ . (i)  $f \in \text{Ker}(A^*) \Rightarrow A^*f = 0 \Rightarrow \langle f, Ag \rangle = \langle A^*f, g \rangle = 0$ . (ii)  $f \in (\text{Im}(A))^{\perp} \Rightarrow \langle f, Ag \rangle = 0$  for all  $g \in \mathcal{H} \Rightarrow \langle A * f, g \rangle = 0$  for all  $g \in \mathcal{H} \Rightarrow A^*f = 0$ .  $\square$

$\mathcal{H} = L^2(\omega, \mathcal{F}, \mathbb{P}) = \text{Ker}(\mathcal{U}_T^* - \text{Id}) + \overline{\text{Im}(\mathcal{U}_T - \text{Id})}$ .  $\Rightarrow \forall f \in \mathcal{H}, \forall \epsilon > 0, \exists g, h \in \mathcal{H}$  s.t.  $\|h\|_{L^2} < \epsilon$  and  $f = \mathbb{E}(f|\mathcal{I}) + (\mathcal{U}_T - \text{Id})g + h \Rightarrow$   
 $\frac{1}{N} \sum_{k=0}^{N-1} \mathcal{U}_T^k f = \mathbb{E}(f|\mathcal{I}) + \underbrace{\frac{1}{N}(\mathcal{U}_T^N g - g)}_{\| \cdot \|_{L^2} \leq \frac{1}{N} \|g\|_{L^2} \rightarrow 0} + \underbrace{\frac{1}{N} \sum_{k=0}^{N-1} \mathcal{U}_T^k h}_{\| \cdot \|_{L^2} < \epsilon} \Rightarrow \limsup_{N \rightarrow \infty} \|\frac{1}{N} \sum_{k=0}^{N-1} \mathcal{U}_T^k f - \mathbb{E}(f|\mathcal{I})\|_{L^2} < \epsilon$ .

**For  $p \neq 2$** , let  $S_N f = \sum_{k=0}^{N-1} f \circ T^k$  and  $A_N f = \frac{1}{N} S_N f$ .

(1) If  $f \in L^{\infty}$ , then  $\|A_N f\|_{L^{\infty}} \leq \|f\|_{L^{\infty}}$ ,  $\|A_N f - \mathbb{E}(f|\mathcal{I})\|_{L^2} \rightarrow 0 \Rightarrow A_N f \rightarrow \mathbb{E}(f|\mathcal{I})$  in  $L^p$  for every  $p \in [1, +\infty)$  (for  $p \geq 2$ ,  $\|A_N f - \mathbb{E}(f|\mathcal{I})\|_{L^p}^p \leq \|f\|_{L^{\infty}}^{p-2} \|A_N f - \mathbb{E}(f|\mathcal{I})\|_{L^2}^2$ ; for  $1 \leq p < 2$ ,  $\|A_N f - \mathbb{E}(f|\mathcal{I})\|_{L^p}^p \leq \|A_N f - \mathbb{E}(f|\mathcal{I})\|_{L^2}^p \|1\|_{L^2}^{2-p}$ ).

(2) If  $f \in L^p(p \geq 1)$ , then  $\forall \epsilon > 0, \exists g \in L^{\infty}$  s.t.  $\|f - g\|_{L^p} < \epsilon$ ,

$$\|A_N f - \mathbb{E}(f|\mathcal{I})\|_{L^p} \leq \underbrace{\|A_N(f - g)\|_{L^p}}_{< \epsilon} + \underbrace{\|A_N g - \mathbb{E}(g|\mathcal{I})\|_{L^p}}_{\rightarrow 0 \text{ as } N \rightarrow +\infty} + \underbrace{\|\mathbb{E}(g - f|\mathcal{I})\|_{L^p}}_{< \epsilon} \Rightarrow \forall \epsilon > 0, \limsup_{N \rightarrow \infty} \|A_N f - \mathbb{E}(f|\mathcal{I})\|_{L^p} < 2\epsilon.$$

We next show **convergence a.s.**

*Maximum ergodic theorem*  $f \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ ,  $S_n = \sum_{k=0}^{n-1} f \circ T^k$ ,  $M_n = \max\{S_1, \dots, S_n\}$ . Then  $\int_{\{M_n \geq 0\}} f(\omega) \mathbb{P}(d\omega) \geq 0$ .

*Proof*  $M_{n-1}(T\omega) = \max\{S_1(T\omega), \dots, S_{n-1}(T\omega)\} = \max\{S_2(\omega), S_n(\omega)\} - f(\omega) \Rightarrow \max\{0, M_{n-1}(T\omega)\} = M_n(\omega) - f(\omega) \Rightarrow$   
 $f(\omega) = M_n(\omega) - \max\{0, M_{n-1}(T\omega)\}$ .  $\int_{\{M_n > 0\}} f d\mathbb{P} = \int_{\{M_n > 0\}} M_n d\mathbb{P} - \int_{\{M_n > 0\}} \max\{0, M_{n-1}(T\omega)\} d\mathbb{P} = \int_{\{M_n > 0\}} M_n d\mathbb{P} -$   
 $\int_{\{M_n > 0\} \cap \{M_{n-1} \circ T > 0\}} M_{n-1} \circ T d\mathbb{P} \Rightarrow \int_{\{M_n > 0\}} f d\mathbb{P} \geq \int_{\{M_n \geq 0\}} M_n d\mathbb{P} - \int_{\{\dots\}} M_n \circ T d\mathbb{P} = \int_{\{M_n \geq 0\}} M_n d\mathbb{P} - \int_{T\{\dots\}} M_n d\mathbb{P} \geq 0$ .  $\square$

*Corollary 1*  $\mathbb{P}(\omega : \sup_{n \geq 1} (A_n f)(\omega) > \lambda) \leq \frac{\mathbb{E}|f|}{\lambda}$ .

*Proof* Let  $E_N = \{\omega : \sup_{1 \leq n \leq N} (A_n f)(\omega) > \lambda\} = \{\omega : \sup_{1 \leq n \leq N} (A_n(f - \lambda))(\omega) > 0\} = \{\omega : \sup_{1 \leq n \leq N} (S_n(f - \lambda))(\omega) > 0\}$ .  
 $E_N \uparrow E = \{\omega : \sup_{n \geq 1} (A_n f)(\omega) > \lambda\}$ .  $\int_{E_N} (f - \lambda) d\mathbb{P} \geq 0 \Rightarrow \mathbb{P}(E_N) \leq \frac{\int_{E_N} f d\mathbb{P}}{\lambda} \leq \frac{\mathbb{E}|f|}{\lambda} \Rightarrow \mathbb{P}(E) \leq \frac{\mathbb{E}|f|}{\lambda}$ .  $\square$

# ERGODIC THEOREM

Goal:  $f \in L^1$  (for finite measure  $\mathbb{P}$ ,  $L^p \subset L^1$ ), need to show  $\frac{1}{N} \sum_{k=0}^{N-1} f \circ T^k \rightarrow \mathbb{E}(f|\mathcal{I})$  a.s.

(1) If  $f \in L^2$  is  $\mathcal{I}$ -measurable, then  $A_N f = f = \mathbb{E}(f|\mathcal{I})$  a.s.

(2) If  $f = (\mathcal{U}_T - \text{Id})g$  for some  $g \in L^\infty$ , then  $(A_N f)(\omega) = \frac{1}{N}(g(T^N \omega) - g(\omega)) \leq \frac{2\|g\|_{L^\infty}}{N} \rightarrow 0$ . Check  $\mathbb{E}((\mathcal{U}_T - \text{Id})g|\mathcal{I}) = 0 : \forall A \in \mathcal{I}, \int_A (g \circ T - g)d\mathbb{P} = \int_{T^{-1}(A)} g \circ T d\mathbb{P} - \int_A g d\mathbb{P} = \int_A g d\mathbb{P} - \int_A g d\mathbb{P} = 0$ .

(3)  $\Lambda = \{f = \mathbb{E}(f_0|\mathcal{I}) + (\mathcal{U}_T - \text{Id})g : f_0 \in L^2, g \in L^\infty\}$  is dense in  $L^1$ . If  $f \in L^1$ , then  $\exists f_j \in \Lambda$  s.t.  $f_j \rightarrow f$  in  $L^1$ . We need to show  $\mathbb{P}(\limsup_{N \rightarrow \infty} |A_N f - \mathbb{E}(f|\mathcal{I})| > \epsilon) = 0$ .  $|A_N f - \mathbb{E}(f|\mathcal{I})| \leq |A_N(f - f_j)| + \underbrace{|A_N f_j - \mathbb{E}(f_j|\mathcal{I})|}_{\rightarrow 0 \text{ a.s.}} + |\mathbb{E}(f_j - f|\mathcal{I})| \Rightarrow$

$$\mathbb{P}(\limsup_{N \rightarrow \infty} |A_N f - \mathbb{E}(f|\mathcal{I})| > \epsilon) \leq \mathbb{P}(\limsup_{N \rightarrow +\infty} |A_N(f - f_j)| > \frac{\epsilon}{2}) + \mathbb{P}(|\mathbb{E}(f_j - f|\mathcal{I})| > \frac{\epsilon}{2}) \leq \frac{2\mathbb{E}|f_j - f|}{\epsilon} + \frac{2\mathbb{E}|f_j - f|}{\epsilon} \rightarrow 0. \quad \square$$

- Kingman's subadditive ergodic theorem: Let  $(\Omega, \mathcal{F}, \mathbb{P}, T)$  be a measure-preserving space and  $\{g_n\} \in L^1$  subadditive in the sense that  $g_{n+m} \leq g_n + g_m \circ T^n$  for every  $n, m$ . Then (1)  $\lim_{n \rightarrow \infty} \frac{\mathbb{E}(g_n)}{n} \rightarrow \inf_{k \geq 1} \frac{\mathbb{E}(g_k)}{k}$  (possibly  $-\infty$ ); (2)  $\frac{g_n}{n}$  convergence a.s. to  $F$  where  $F$  is  $\mathcal{I}$ -measurable and  $\mathbb{E}F = \inf_{k \geq 1} \frac{\mathbb{E}(g_k)}{k}$ ; (3) If  $\mathbb{E}F > -\infty$ , then the convergence is also in  $L^1$ .

*Proof* Recall an elementary version. If  $\{a_n\} \in \mathbb{R}$  s.t.  $a_{n+m} \leq a_n + a_m, \forall n, m$ , then  $\frac{a_n}{n} \rightarrow \inf_{k \geq 1} \frac{a_k}{k}$  as  $n \rightarrow \infty$ .  $\square$