

# MA262: Introduction to Stochastic Processes

Naman Mishra

January 2024

# Contents

0	The Course	2	
1	Discrete time Markov Chains	3	
1.1	Transience & Recurrence . . . . .	21	<b>Lecture</b> <b>01:</b> Thu 04 Jan '24

# Chapter 0

## The Course

### Texts:

- *Markov Chains*, J. R. Norris
- *Introduction to Stochastic Processes*, Hoel, Port, Stone
- Karlin and Taylor

### Grading:

(20%) 2 quizzes

(30%) 1 midterm

(50%) Final

# Chapter 1

## Discrete time Markov Chains

**Definition 1.1** (Stochastic matrix). Let  $S$  be a state set (at most countable). A matrix  $P = (p_{xy})_{x,y \in S}$  is called a *stochastic matrix* if  $p_{xy} \geq 0$  for all  $x, y \in S$  and  $\sum_{y \in S} p_{xy} = 1$  for all  $x \in S$ .

**Definition 1.2** (Markov chain). Let  $S$  be a state set,  $P = (p_{xy})$  a stochastic matrix, and  $\mu_0$  a probability distribution on  $S$ , i.e.,  $\mu_0(x) \geq 0$  for all  $x \in S$  and  $\sum_{x \in S} \mu_0(x) = 1$ .

Suppose  $X_0, X_1, \dots$  are random variables defined on the same probability space taking values in  $S$ . Then  $(X_n)_{n \in \mathbb{N}}$  is called a *Markov chain* with initial distribution  $\mu_0$  and transition matrix  $P$ , denoted  $MC(\mu_0, P)$ , if  $X_0$  has distribution  $\mu_0$  and for all  $n \geq 0$ ,

$$\Pr(X_{n+1} = x_{n+1} \mid X_n = x_n, \dots, X_0 = x_0) = p_{x_n x_{n+1}}$$

whenever  $\Pr(X_n = x_n, \dots, X_0 = x_0) > 0$ .

*Notation.* Whenever writing  $\Pr(X_n \in A \mid (X_0, \dots, X_{n-1}) \in B)$ , it will be understood that only  $\Pr((X_0, \dots, X_{n-1}) \in B) > 0$  is considered.

**Theorem 1.3.**  $(X_n)_{n=0}^N$  is  $MC(\mu_0, P)$  iff

$$\Pr(X_0 = x_0, \dots, X_N = x_N) = \mu_0(x_0) p_{x_0 x_1} \cdots p_{x_{N-1} x_N}$$

for all  $x_0, \dots, x_N \in S$ .

*Proof.* Both directions are proven by induction.

Suppose  $(X_n)_{n=0}^N$  is  $MC(\mu_0, P)$ . Then  $\Pr(X_0 = x_0) = \mu_0(x_0)$ .

If  $\Pr(X_0 = x_0) > 0$ , then  $\Pr(X_0 = x_0, X_1 = x_1) = \mu_0(x_0)p_{x_0x_1}$ .  
 If  $\Pr(X_0 = x_0) = 0$ , then  $\Pr(X_0 = x_0, X_1 = x_1) \leq \Pr(X_0 = x_0) = 0$ , and so  
 $\Pr(X_0 = x_0, X_1 = x_1) = 0 = \mu_0(x_0)p_{x_0x_1}$ .

**Induction:** Suppose

$$P_j := \Pr(X_0 = x_0, \dots, X_j = x_j) = \mu_0(x_0)p_{x_0x_1} \dots p_{x_{j-1}x_j}.$$

If this is zero, so is  $P_{j+1}$ , and so it is equal to  $\mu_0(x_0)p_{x_0x_1} \dots p_{x_{j-1}x_j}p_{x_jx_{j+1}}$ .  
 If not, then

$$\begin{aligned} P_{j+1} &= P_j \Pr(X_{j+1} = x_{j+1} \mid X_0 = x_0, \dots, X_j = x_j) \\ &= P_j p_{x_jx_{j+1}} \\ &= \mu_0(x_0)p_{x_0x_1} \dots p_{x_{j-1}x_j}p_{x_jx_{j+1}}, \end{aligned}$$

closing the induction. In particular,

$$\Pr(X_0 = x_0, \dots, X_N = x_N) = \mu_0(x_0)p_{x_0x_1} \dots p_{x_{N-1}x_N}.$$

Now for the converse, suppose

$$\Pr(X_0 = x_0, \dots, X_N = x_N) = \mu_0(x_0)p_{x_0x_1} \dots p_{x_{N-1}x_N}$$

for all  $x_0, \dots, x_N \in S$ . Then for any  $x_0, \dots, x_{N-1} \in S$ ,

$$\begin{aligned} \Pr(X_0 = x_0, \dots, X_{N-1} = x_{N-1}) &= \sum_{x_N \in S} \Pr(X_0 = x_0, \dots, X_N = x_N) \\ &= \sum_{x_N \in S} \mu_0(x_0)p_{x_0x_1} \dots p_{x_{N-2}x_{N-1}}p_{x_{N-1}x_N} \\ &= \mu_0(x_0)p_{x_0x_1} \dots p_{x_{N-2}x_{N-1}}. \end{aligned}$$

We have by backwards induction that for all  $1 \leq i \leq N$ ,

$$\Pr(X_0 = x_0, \dots, X_i = x_i) = \mu_0(x_0)p_{x_0x_1} \dots p_{x_{i-1}x_i}$$

and  $\Pr(X_0 = x_0) = \mu_0(x_0)$ . This allows us to deduce that

$$\Pr(X_{i+1} = x_{i+1} \mid X_0 = x_0, \dots, X_i = x_i) = p_{x_ix_{i+1}}$$

by definition of conditional probability. □

**Fact 1.4** (Strong Law of Large Numbers). Suppose  $Z_1, Z_2, \dots$  are iid  $\mathbb{R}$ -valued random variables and  $E[Z_1]$  exists. Then

$$\frac{Z_1 + \dots + Z_n}{n} \xrightarrow{\text{a.s.}} E[Z_1]$$

as  $n \rightarrow \infty$ , that is,

$$\Pr \left\{ \omega \in \Omega : \lim_{n \rightarrow \infty} \frac{Z_1(\omega) + \dots + Z_n(\omega)}{n} = E[Z_1] \right\} = 1.$$

**Fact 1.5** (Weak Law of Large Numbers). Suppose  $Z_1, Z_2, \dots$  are iid  $\mathbb{R}$ -valued random variables and  $E[Z_1]$  exists. Then

$$\frac{Z_1 + \dots + Z_n}{n} \xrightarrow{\text{P}} E[Z_1]$$

as  $n \rightarrow \infty$ , that is, for any  $\delta > 0$ ,

$$\lim_{n \rightarrow \infty} \Pr \left\{ \left| \frac{Z_1 + \dots + Z_n}{n} - E[Z_1] \right| \leq \delta \right\} = 1.$$

**Fact 1.6** (Central Limit Theorem). Suppose  $Z_1, Z_2, \dots$  are iid  $\mathbb{R}$ -valued random variables and  $E[Z_1^2]$  exists. Then

$$Y_n := \frac{\sqrt{n}}{\sqrt{\text{Var}(Z_1)}} \left( \frac{Z_1 + \dots + Z_n}{n} - E[Z_1] \right) \xrightarrow{d} N(0, 1)$$

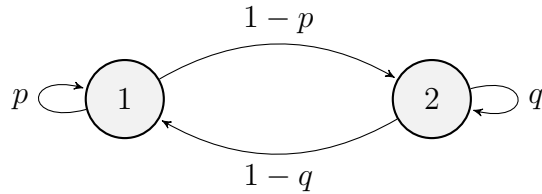
as  $n \rightarrow \infty$ , that is,

$$\lim_{n \rightarrow \infty} \Pr \{ Y_n \leq y \} = \Phi(y)$$

for all  $y \in \mathbb{R}$ , where  $\Phi$  is the pdf of the standard normal distribution.

*Examples.*

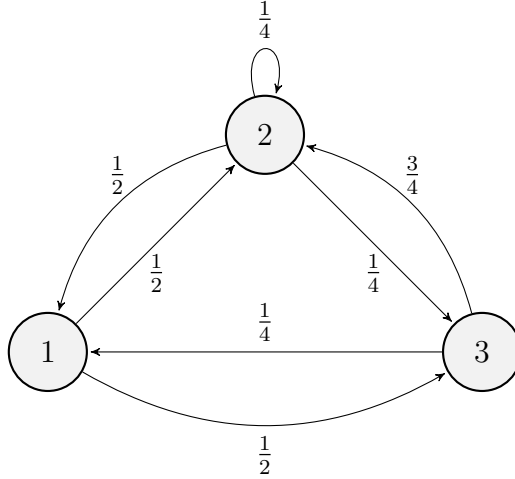
- A two-state Markov chain.



This corresponds to the matrix

$$P = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix}.$$

- A three-state Markov chain.



This has transition matrix

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 1/4 & 1/4 \\ 1/4 & 3/4 & 0 \end{pmatrix}.$$

- Simple random walk on  $\mathbb{Z}$ . Starting from some randomly chosen point, at each step, move right with probability  $p$  and left with probability  $q := 1 - p$ . That is,

$$\Pr(X_{n+1} = y \mid X_n = x) = \begin{cases} p & \text{if } y = x + 1, \\ q & \text{if } y = x - 1, \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

Such a simple random walk is called symmetric if  $p = q = \frac{1}{2}$ . A special case is where  $\mu_0 = \delta_x$  for some  $x \in \mathbb{Z}$  (where  $\delta_x$  is the Kronecker delta).

Suppose that  $Y_1, Y_2, \dots$  are iid with distribution  $\begin{pmatrix} 1 & -1 \\ p & 1-p \end{pmatrix}$ . Each  $Y_i$  has expectation  $2p - 1$ , and variance

$$\mathbb{E}[Y_1^2] - (\mathbb{E}[Y_1])^2 = 1 - (2p - 1)^2 = 4pq.$$

We have that  $(X_n)_{n \in \mathbb{N}} \stackrel{d}{=} (\sum_{j=1}^n Y_j)_{n \in \mathbb{N}}$ .

*Definition 1.7.* Suppose  $Z_1, \dots, Z_k$  are random variables taking values in a state set  $S$  defined on a probability space  $(\Omega, \mathcal{F}, \Pr)$ . and  $\tilde{Z}_1, \dots, \tilde{Z}_k$  are rvs taking values in a state set  $S$  defined on a probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ . Then  $(Z_1, \dots, Z_k)$  and  $(\tilde{Z}_1, \dots, \tilde{Z}_k)$  are said to be identically distributed if

$$\Pr(Z_1 = x_1, \dots, Z_k = x_k) = \Pr(\tilde{Z}_1 = x_1, \dots, \tilde{Z}_k = x_k).$$

This is denoted as

$$(Z_1, \dots, Z_k) \stackrel{d}{=} (\tilde{Z}_1, \dots, \tilde{Z}_k).$$

Then from the weak law of large numbers,

$$\frac{X_n}{n} \rightarrow \mathbb{E}[Y_1] = 2p - 1.$$

From the central limit theorem,

$$\frac{X_n - n(p - q)}{\sqrt{n}\sqrt{4pq}} \xrightarrow{d} N(0, 1).$$

On a graph, a simple symmetric random walk is a random walk on a graph where each

$$p_{xy} = \begin{cases} \frac{1}{\deg(x)} & \text{if } y \sim x, \\ 0 & \text{otherwise.} \end{cases}$$

On  $\mathbb{Z}^2$ , a simple random walk is given by  $p_N, p_E, p_S, p_W$ , where  $p_N + p_E + p_S + p_W = 1$ . At each step, move up with probability  $p_N$ , right with probability  $p_E$ , down with probability  $p_S$ , and left with probability  $p_W$ .

- Consider a shooting game with 4 modes:  $N$  (normal),  $D$  (distance),  $W$  (windy) and  $DW$  (distance and windy). The game changes mode randomly to a mode different from the current mode with directed graph  $K_4$  with some edge weights.

**Theorem 1.8.** If  $(X_n)_{n \in \mathbb{N}}$  is a DTMC with transition matrix  $P$ , then

$$\Pr_{\mu_0}(X_n = y) = (\mu_0 P^n)_y.$$

In particular,  $\Pr_x(X_n = y) = (P^n)_{x,y} =: p_{xy}^{(n)}$ .



Here,  $\mu_0$  is viewed as a row vector, and  $\Pr_{\mu_0}$  is the distribution under the assumption that  $X_0 \sim \mu_0$ . Also,  $\Pr_x$  is under the assumption that  $\mu_0 = \delta_x$ .

*Proof.*

$$\begin{aligned}\Pr_{\mu_0}(X_n = y) &= \sum_{x_0, \dots, x_{n-1} \in S} \Pr(X_0 = x_0, \dots, X_{n-1} = x_{n-1}, X_n = y) \\ &= \sum_{x_0, \dots, x_{n-1} \in S} (\mu_0)_{x_0} p_{x_0 x_1} \cdots p_{x_{n-1} y} \\ &= (\mu_0 P^n)_y\end{aligned}\quad \square$$

**Theorem 1.9** (Markov property). Let  $(X_n)_{n \in \mathbb{N}}$  be  $MC(\mu_0, P)$ . Then for any  $n \geq 0$ ,  $l \geq 1$ ,  $x_n, \dots, x_{n+l} \in S$  and  $A \subseteq S^n$ ,

$$\begin{aligned}\Pr_{\mu_0}(X_i = x_i, n < i \leq n+l \mid X_n = x_n, (X_0, \dots, X_{n-1}) \in A) \\ = \Pr_{x_n}(X_1 = x_{n+1}, \dots, X_l = x_{n+l})\end{aligned}$$

In other words, conditioning on  $X_n = x_n$  and  $(X_0, \dots, X_{n-1}) \in A$ , the process  $(X_n, X_{n+1}, \dots)$  is  $MC(\delta_{x_n}, P)$ .

*Proof.*

$$\begin{aligned}\Pr_{\mu_0}(X_{n+l} = x_{n+l}, \dots, X_n = x_n, (X_0, \dots, X_{n-1}) \in A) \\ = \sum_{(x_0, \dots, x_{n-1}) \in A} p_{x_n x_{n+1}} \cdots p_{x_{n+l-1} x_{n+l}} \mu_0(x_0) p_{x_0 x_1} \cdots p_{x_{n-1} x_n} \\ = \left( \sum_{(x_0, \dots, x_{n-1}) \in A} \mu_0(x_0) p_{x_0 x_1} \cdots p_{x_{n-1} x_n} \right) p_{x_n x_{n+1}} \cdots p_{x_{n+l-1} x_{n+l}} \\ = \Pr_{\mu_0}(X_n = x_n, (X_0, \dots, X_{n-1}) \in A) \cdot p_{x_n x_{n+1}} \cdots p_{x_{n+l-1} x_{n+l}}.\end{aligned}$$

By the definition of conditional probability,

$$\begin{aligned}\Pr_{\mu_0}(X_{n+l} = x_{n+l}, \dots, X_n = x_n \mid X_n = x_n, (X_0, \dots, X_{n-1}) \in A) \\ = p_{x_n x_{n+1}} \cdots p_{x_{n+l-1} x_{n+l}} \\ = \delta_{x_n}(x_n) p_{x_n x_{n+1}} \cdots p_{x_{n+l-1} x_{n+l}} \\ = \Pr_{x_n}(X_1 = x_{n+1}, \dots, X_l = x_{n+l}).\end{aligned}\quad \square$$

**Lecture**  
**02:** Tue  
9 Jan '24

**Definition 1.10** (Sigma algebra). A  $\sigma$ -algebra over a set  $\Omega$  is a collection  $\mathcal{F}$  of subsets of  $\Omega$  such that

- (i)  $\emptyset \in \mathcal{F}$ ,
- (ii) if  $A \in \mathcal{F}$ , then  $A^c \in \mathcal{F}$ ,
- (iii) if  $\mathcal{A} \subseteq \mathcal{F}$  is countable, then  $\bigcup \mathcal{A} \in \mathcal{F}$ .

**Definition 1.11** (Probability Space). A *probability space* is a triple  $(\Omega, \mathcal{F}, \Pr)$ , where  $\Omega$  is a set,  $\mathcal{F}$  is a  $\sigma$ -algebra over  $\Omega$ , and  $\Pr$  is a probability measure over  $(\Omega, \mathcal{F})$ .

**Definition 1.12.** Let  $X_1, X_2, \dots$  and  $X$  be random variables over a probability space  $(\Omega, \mathcal{F}, \Pr)$ . We define

**Almost sure convergence.**  $X_n \xrightarrow{\text{a.s.}} X$  if

$$\Pr\{\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\} = 1.$$

**Convergence in probability.**  $X_n \xrightarrow{\text{P}} X$  if for every  $\varepsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \Pr\{|X_n - X| \leq \varepsilon\} = 1.$$

**Convergence in distribution.**  $X_n \xrightarrow{d} X$  if for every  $x$ ,

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x),$$

where  $F$ 's are cumulative distribution functions and  $F_X$  is continuous.

**Exercise 1.13.** Derive the Chapman-Kolmogorov equation

$$p_{xy}^{(m+n)} = \sum_{z \in S} p_{xz}^{(m)} p_{zy}^{(n)}$$

for all  $x, y \in S$  and  $m, n \in \mathbb{N}$ .

*Solution.* We have

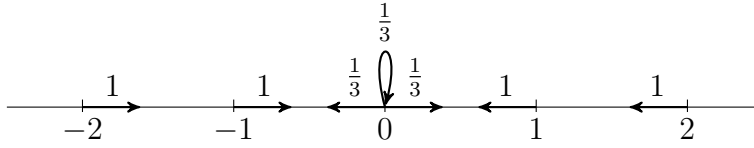
$$\begin{aligned}
p_{xy}^{(m+n)} &= P_{xy}^{m+n} \\
&= (P^m P^n)_{xy} \\
&= \sum_{z \in S} P_{xz}^m P_{zy}^n \\
&= \sum_{z \in S} p_{xz}^{(m)} p_{zy}^{(n)}.
\end{aligned}$$

**Definition 1.14** (Period). Let  $(X_n)_{n \in \mathbb{N}}$  be  $MC(S, \mu_0, P)$ . For each  $x \in S$ , let  $F_x = \{n \in \mathbb{N}^* : p_{xx}^{(n)} > 0\}$ . The *period* of  $x$  is defined as  $d_x = \gcd F_x$ , where  $\gcd \emptyset$  is considered to be 0.

A state  $x$  is *aperiodic* if  $d_x = 1$ . A Markov chain is *aperiodic* if all its states are aperiodic.

*Examples.*

- The simple random walk on  $\mathbb{Z}$  is periodic with period 2.
- Consider the walk on  $\mathbb{Z}$  given by



0 is aperiodic. 0's aperiodicity induces aperiodicity on all other states. Thus the chain is aperiodic.

**Theorem 1.15.** If  $x \leftrightarrow y$ , then  $d_x = d_y$ .

Here,  $x \leftrightarrow y$  denotes the existence of a path from  $x$  to  $y$  and from  $y$  to  $x$ . That is,  $x \leftrightarrow y$  if there exist  $n, m \in \mathbb{N}$  such that  $p_{xy}^{(n)} > 0$  and  $p_{yx}^{(m)} > 0$ .

*Proof.* Trivial when  $x = y$ . Suppose  $x \neq y$  and let  $n, m \in \mathbb{N}$  be lengths of paths from  $x$  to  $y$  and from  $y$  to  $x$ , respectively. Note that  $d_x, d_y \neq 0$ . By the Chapman-Kolmogorov equation,  $p_{xx}^{(n+m)} \geq p_{xy}^{(n)} p_{yx}^{(m)} > 0$ , so  $d_x \mid n + m$ .

Now let  $p$  be a path length from  $y$  to itself. Then  $p_{xx}^{(n+m+p)} \geq p_{xy}^{(n)} p_{yy}^{(p)} p_{yx}^{(m)} > 0$ , so  $d_x \mid n + m + p$ . This implies  $d_x \mid p$ . Since  $p$  was arbitrary,  $d_x \mid d_y$ .

By symmetry,  $d_y \mid d_x$ , so  $d_x = d_y$ .  $\square$

**Theorem 1.16.** If  $d_x \geq 1$ , then there exists an  $n_x \in \mathbb{N}^*$  such that for all  $n \geq n_x$ ,  $p_{xx}^{(nd_x)} > 0$ .  
As a special case, if  $d_x = 1$ , then  $p_{xx}^{(n)} > 0$  for all large enough  $n$ .

We first prove a general number-theoretic result.

**Theorem 1.17** (Schur's Lemma). Suppose  $S \subseteq \mathbb{N}^*$  and denote  $\gcd(S)$  by  $g_S$ . Then there exists an  $m_s \in \mathbb{N}^*$  such that for all  $m \geq m_s$ , there exist  $k \in \mathbb{N}^*$ ,  $e_1, \dots, e_k \in \mathbb{N}^*$  and  $s_1, \dots, s_k \in S$  such that  $mg_S = \sum_{i=1}^k e_i s_i$ .

We prove the following lemma to restrict  $S$  to a finite set.

**Lemma 1.18.** Let  $S \subseteq \mathbb{N}^*$ . Then there exists a finite set  $S' \subseteq S$  such that  $\gcd(S) = \gcd(S')$ .

*Proof.* Let  $g_S = \gcd(S)$ . For any finite set  $S' \subseteq S$ , we either have  $\gcd(S') = g_S$  in which case we are done, or  $\exists s \in S \setminus S'$  such that  $\gcd(S') \nmid s$ . In the latter case, we can add  $s$  to  $S'$  and continue, producing a sequence of finite sets with *strictly decreasing* gcds. Since the gcd can decrease only a finite number of times, this process must terminate with a finite set whose gcd is  $g_S$ .  $\square$

We will also use the following characterization of the gcd.

**Lemma 1.19.** Let  $X \subseteq \mathbb{N}^*$  and let  $Y = X \cup \{n\}$ . Then  $\gcd(Y) = \gcd\{\gcd(X), n\}$ .

*Proof.* Let  $g = \gcd(Y)$  and  $\tilde{g} = \gcd\{\gcd(X), n\}$ .

- Since  $\tilde{g} \mid \gcd(X)$  and  $\tilde{g} \mid n$ , we have  $\tilde{g} \mid y$  for all  $y \in Y$ . Thus  $g \mid \tilde{g}$ .
- Since  $g \mid y$  for all  $y \in Y$ , we have  $g \mid \gcd(X)$  and  $g \mid n$ . Thus  $\gcd\{\gcd(X), n\} = \tilde{g} \mid g$ .  $\square$

We are now ready to prove Schur's Lemma.

*Proof of Schur's Lemma.* Let  $S = \{s_1, s_2, \dots, s_k\}$ . Define  $\tilde{g}_S$  to be the minimum positive linear combination of  $S$  over  $\mathbb{Z}$ . That is,

$$\tilde{g}_S = \min \left( [1, \infty) \cap \left\{ \sum_{i=1}^j a_i x_i \mid 1 \leq j \leq k, a_i \in \mathbb{Z}, x_i \in S \right\} \right).$$

We claim that  $\tilde{g}_S = g_S$ .

- $g_S \mid \tilde{g}_S$  by definition.
- Let  $s \in S$  be decomposed as  $s = q\tilde{g}_S + r$  with  $0 \leq r < \tilde{g}_S$ . Then  $r = s - q\tilde{g}_S$ . However, this is a linear combination of  $S$  over  $\mathbb{Z}$ , so  $r = 0$ . Thus  $\tilde{g}_S \mid g_S$ .

Thus we can write  $g_S = \sum_{s \in S} a_s s$  where  $a_s \in \mathbb{Z}$ . First consider the case  $|S| = 2$ . Let  $S = \{s_1, s_2\}$ . We know that  $g_S = as_1 + bs_2$  for some  $a, b \in \mathbb{Z}$ . Now for any  $m \in \mathbb{N}^*$ ,

$$\begin{aligned} mg_S &= mas_1 + mbs_2 + ks_1s_2 - ks_1s_2 \\ &= (ma - ks_2)s_1 + (mb + ks_1)s_2 \end{aligned}$$

Choose  $k \in \mathbb{N}$  such that  $0 \leq ma - ks_2 < s_2$ . We can write  $mg_S = a_ms_1 + b_ms_2$  where  $0 \leq a_m < s_2$ .

Let  $m_0$  be such that  $m_0g_S > s_1s_2$ . Then for all  $m \geq m_0$ ,  $mg_S - a_ms_1 > (s_2 - a)s_1 > 0$ , so that  $b_m > 0$ . Thus  $m_S = m_0$  works.

Suppose the lemma holds for all sets of size  $l - 1$ .

Let  $S = \{s_1, s_2, \dots, s_l\}$  and  $F = S \setminus \{s_l\}$ . Then by the previous lemma,  $g_S = \gcd(g_F, s_l)$ .

Let  $m_0$  be such that  $mg_S - m_Fg_F \geq m_{\{g_F, s_l\}}$ . Then

$$\begin{aligned} mg_S - m_Fg_F &= ag_F + bs_l \text{ for some } a \in \mathbb{N}, b \in \mathbb{N}^* \\ mg_S &= (a + m_F)g_F + bs_l \end{aligned}$$

but  $a + m_F \geq m_F$ , so we can write

$$mg_S = \sum_{i=1}^{l-1} a_i s_i + bs_l$$

where all  $a_i$ s are non-negative integers. This closes the induction.  $\square$

We can now prove theorem 1.16.

*Proof.* Applying Schur's lemma to  $F_x$ , we have the existence of an  $n_x$  such that  $nd_x$  can be written as a non-negative integer combination of elements of

**Lecture**  
**03: Thu**  
**11 Jan**  
**'24**

$F_x$ , for all  $n \geq n_x$ . Let  $nd_x = \sum_{i=1}^k a_i f_i$ . Then by the Chapman-Kolmogorov equation,

$$\begin{aligned} p_{xx}^{(nd_x)} &\geq \underbrace{p_{xx}^{(f_1)} \cdots p_{xx}^{(f_1)}}_{a_1 \text{ times}} \cdots \underbrace{p_{xx}^{(f_k)} \cdots p_{xx}^{(f_k)}}_{a_k \text{ times}} \\ &\geq \prod_{i=1}^k (p_{xx}^{(f_i)})^{a_i} \\ &> 0 \end{aligned}$$

□

**Definition 1.20** (The Extended Reals). The *extended real line* is the set of real numbers along with 2 formal symbols  $+\infty$  and  $-\infty$ , denoted by

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}.$$

$\overline{\mathbb{R}}$  will be endowed with the order

$$-\infty < x < \infty \text{ for all } x \in \mathbb{R},$$

along with the usual order on  $\mathbb{R}$ . We extend the algebraic operations on  $\mathbb{R}$  to  $\overline{\mathbb{R}}$ .

- $x + \infty = +\infty$ ,  $x - \infty = -\infty$  for all  $x \in \mathbb{R}$ .
- $x \cdot (+\infty) = +\infty$ ,  $x \cdot (-\infty) = -\infty$  for all  $x \in \mathbb{R}$ ,  $x > 0$ .
- $x \cdot (+\infty) = -\infty$ ,  $x \cdot (-\infty) = +\infty$  for all  $x \in \mathbb{R}$ ,  $x < 0$ .
- $\frac{x}{+\infty} = \frac{x}{-\infty} = 0$ , for all  $x \in \mathbb{R}$ .

If  $E \subseteq \mathbb{R}$  is not bounded above in  $\mathbb{R}$ , we say  $\sup E = +\infty$ . If  $E = \emptyset$ , we say  $\inf E = +\infty$ .

**Definition 1.21** (Filtration). Let  $(\Omega, \mathcal{F}, \Pr)$  be a probability space. A collection  $(\mathcal{F}_n)_{n \in \mathbb{N}}$  of  $\sigma$ -algebras over  $\Omega$  is called a *filtration* if  $\mathcal{F}_n \subseteq \mathcal{F}_{n+1} \subseteq \mathcal{F}$  for all  $n \in \mathbb{N}$ .

**Definition 1.22** (Natural filtration). Let  $(X_n)_{n \in \mathbb{N}}$  be a sequence of  $S$ -valued random variables defined on  $(\Omega, \mathcal{F}, \Pr)$ . For  $n \in \mathbb{N}$ , define

$$\begin{aligned} \mathcal{F}_n &= \{(X_0, X_1, \dots, X_n)^{-1}(A) \mid A \subseteq S^{n+1}\} \cap \mathcal{F} \\ &= \sigma(X_0, X_1, \dots, X_n) \end{aligned}$$

Here,  $(X_0, \dots, X_n)^{-1}(A) = \{\omega \in \Omega \mid (X_0(\omega), \dots, X_n(\omega)) \in A\}$ .

This sequence of  $\sigma$ -algebras is called the *natural filtration* of  $(X_n)_{n \in \mathbb{N}}$ .

*Remark.* Note that each  $\mathcal{F}_n$  is a subset of  $2^\Omega$ , not  $\Omega$ .

Why is  $\mathcal{F}_n$  a  $\sigma$ -algebra? The empty set is in  $\mathcal{F}_n$  because  $\emptyset \in S^{n+1}$ , and the set of pre-images of the null set is the null set. The complement of any set in  $\mathcal{F}_n$  is in  $\mathcal{F}_n$  because  $(X_0, \dots, X_n)^{-1}(A)^c = (X_0, \dots, X_n)^{-1}(A^c)$ .

Why is  $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$ ? For any  $A \subseteq S^{n+1}$ , we have

$$(X_0, \dots, X_n)^{-1}(A) = (X_0, \dots, X_n, X_{n+1})^{-1}(A \times S).$$

The set  $\overline{\mathcal{F}} = 2^\Omega$  works as a  $\sigma$ -algebra containing each  $\mathcal{F}_n$ . But is this the desired closure? Or did we intend  $\mathcal{F}_n$  to be intersected with  $\mathcal{F}$ ?

**Definition 1.23** (Stopping time). Suppose  $(X_n)_{n \in \mathbb{N}}$  is a sequence of  $S$ -valued random variables on  $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \in \mathbb{N}}, \Pr)$  where  $(\mathcal{F}_n)_{n \in \mathbb{N}}$  is the natural filtration of  $(X_n)_{n \in \mathbb{N}}$ .

Then  $\tau: \Omega \rightarrow \mathbb{N} \cup \{\infty\}$  is called a *stopping time* with respect to  $(\mathcal{F}_n)_{n \in \mathbb{N}}$  if for all  $n \in \mathbb{N}$ ,

$$\{\omega \in \Omega \mid \tau(\omega) \leq n\} \in \mathcal{F}_n.$$

This is equivalent to saying that for all  $n \in \mathbb{N}$ ,

$$\mathbf{1}_{\{\tau \leq n\}} = \mathbf{1}_{\{(X_0, \dots, X_n) \in A\}} \quad \text{for some } A \in S^{n+1}.$$

Intuitively, a stopping time is a time at which we can decide whether or not to stop the process based on the information available up to that time (in measurable terms).

Consider the simple random walk  $(X_n)_{n \in \mathbb{N}}$  on  $\mathbb{Z}$ . Then the event that the hitting time of 10 is at most  $n$  is

$$\{T_{10} \leq n\} = \bigcup_{i=1}^n \{X_i = 10\}.$$

*Examples.*

- Let  $(X_n)_{n \in \mathbb{N}}$  be an  $S$ -valued stochastic process and let  $A \subseteq S$ . Let  $T_A := \inf\{n \geq 1 \mid X_n \in A\}$  (where we take  $\inf \emptyset$  to be  $+\infty$ ). Then  $T_A$  is a stopping time with respect to the natural filtration associated with  $(X_n)_{n \in \mathbb{N}}$ . That is, for all  $n \in \mathbb{N}$ ,

$$\{T_A \leq n\} = \bigcup_{i=1}^n \{X_i \in A\} \in \mathcal{F}_n.$$

Intuitively, say we stop as soon as we hit a desired state. Then we

can decide whether or not to stop at time  $n$  based on the information available up to time  $n$ .

- SRW( $p$ ) started at the origin. Then  $L = \sup\{n \geq 1 \mid X_n < 7\}$  is NOT a stopping time. Intuitively,  $L$  is the *last* time we are below 7, which cannot be determined based on the information available from the past.

**Proposition 1.24.**  $\tau$  is a stopping time iff for all  $n \in \mathbb{N}$ ,

$$\{\tau = n\} \in \mathcal{F}_n.$$

*Proof.* Suppose  $\tau$  is a stopping time. Then  $\{\tau = n\} \in \mathcal{F}_n$  for all  $n \in \mathbb{N}$ . This is because

$$\{\tau = n\} = \{\tau \leq n\} \cap \{\tau \leq n-1\}^c$$

where both sets are in  $\mathcal{F}_n \supseteq \mathcal{F}_{n-1}$ , and therefore so is their intersection (de Morgan's law).

Conversely, suppose  $\{\tau = n\} \in \mathcal{F}_n$  for all  $n \in \mathbb{N}$ . Since  $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots \subseteq \mathcal{F}_n$ , we have  $\{\tau = i\} \in \mathcal{F}_n$  for all  $i \leq n$ . Hence so is

$$\{\tau \leq n\} = \bigcup_{i=0}^n \{\tau = i\}. \quad \square$$

**Proposition 1.25.** If  $\tau_1$  and  $\tau_2$  are stopping times, then so are  $\tau_1 \wedge \tau_2$ ,  $\tau_1 \vee \tau_2$  and  $\tau_1 + \tau_2$ .

*Proof.* We have

$$\begin{aligned} \{\tau_1 \wedge \tau_2 \leq n\} &= \{\tau_1 \leq n\} \cup \{\tau_2 \leq n\} \\ \{\tau_1 \vee \tau_2 \leq n\} &= \{\tau_1 \leq n\} \cap \{\tau_2 \leq n\} \\ \{\tau_1 + \tau_2 \leq n\} &= \bigcup_{i=0}^n \{\tau_1 \leq i\} \cap \{\tau_2 \leq n-i\} \end{aligned} \quad \square$$

We can interpret the above operations as follows.

- $\tau_1 \wedge \tau_2$  is the stopping time if we stop when either of our conditions are met.
- $\tau_1 \vee \tau_2$  is the stopping time if we stop when both of our conditions are met.



- $\tau_1 + \tau_2$  is the stopping time if we stop when we have waited for  $\tau_1$  before we started looking for  $\tau_2$ .

**Exercise 1.26.** Give an example of two stopping times  $\tau_1$  and  $\tau_2$  such that  $\Pr(\tau_1 \leq \tau_2) = 1$  but  $\tau_2 - \tau_1$  is not a stopping time.

*Solution.* Consider the SRW( $p$ ) started at the origin, with

$$\begin{aligned}\tau_1 &= \inf\{n \geq 1 \mid X_n = 1\} \\ \tau_2 &= \inf\{n \geq 1 \mid X_n = 2\}\end{aligned}$$

**Theorem 1.27** (Strong Markov property). Let  $(X_n)_{n \in \mathbb{N}}$  be in  $MC(\mu_0, p)$ , and let  $\tau$  be a stopping time. Let  $A \subseteq \mathbb{N}$ . Then

$$\begin{aligned}\Pr_{\mu_0}(X_{\tau+1} = x_1, \dots, X_{\tau+n} = x_n \mid \tau \in A, X_\tau = x) \\ = \Pr_x(X_1 = x_1, \dots, X_n = x_n)\end{aligned}$$

*Remark.* The SMP is equivalent to

$$\mathbb{E}_{\mu_0} [f((X_{\tau+j})_{j \in \mathbb{N}}) \mid \tau \in A, X_\tau = x] = \mathbb{E}_x [f((X_j)_{j \in \mathbb{N}})]$$

for any bounded function  $f: S^\infty \rightarrow \mathbb{R}$ .

*Proof.*

$$\begin{aligned}\Pr_{\mu_0}(X_{\tau+1} = x_1, \dots, X_{\tau+n} = x_n, \tau \in A, X_\tau = x) \\ = \sum_{m \in A} \Pr_{\mu_0}(X_{m+1} = x_1, \dots, X_{m+n} = x_n, \tau = m, X_m = x) \\ = \sum_{m \in A} \Pr_{\mu_0}(\tau = m, X_m = x) \\ \quad \times \Pr_{\mu_0}(X_{m+1} = x_1, \dots, X_{m+n} = x_n \mid \tau = m, X_m = x) \\ = \sum_{m \in A} \Pr_{\mu_0}(\tau = m, X_m = x) \Pr_x(X_1 = x_1, \dots, X_n = x_n)\end{aligned}$$

by the [Markov property](#), since  $\tau = m$  is equivalent to  $(X_0, \dots, X_m)$  belonging to some set in  $\mathcal{F}_m$ . Summing over  $A$  gives

$$= \Pr_{\mu_0}(\tau \in A, X_\tau = x) \Pr_x(X_1 = x_1, \dots, X_n = x_n)$$

and dividing by  $\Pr_{\mu_0}(\tau \in A, X_\tau = x)$  yields the result.  $\square$

**Definition 1.28.** Suppose  $X$  takes values in  $\mathbb{N} \cup \{\infty\}$ , and let  $p_k := \Pr(X = k)$ ,  $k \in \mathbb{N} \cup \{\infty\}$ . Then the *probability generating function* of  $X$  is defined as

$$\begin{aligned} G_X(s) &= p_0 + p_1 s + p_2 s^2 + \dots, \quad s \in (-1, 1) \\ &= \mathbb{E}[s^X] \end{aligned}$$

where we take  $s^\infty = 0$  for  $|s| < 1$ .

*Remark.* The left limit  $G_X(1^-) = \lim_{s \uparrow 1} G_X(s) = 1 - p_\infty$ .

If  $p_\infty > 0$ , then  $\mathbb{E} X = \infty$ . Otherwise,  $\mathbb{E} X = \sum_{k=0}^{\infty} k p_k = G'_X(1^-)$ .

**Fact 1.29.** Let  $X, Y$  be two random variables taking values in  $\mathbb{N} \cup \infty$  and  $G_X(s) = G_Y(s) \forall s \in (-1, 1)$ , then  $X \stackrel{d}{=} Y$ .

**Exercise 1.30.** Suppose  $(X_n)_{n \in \mathbb{N}}$  is an SRW( $p$ ) on  $\mathbb{Z}$  started at the origin. Find  $G = G_{T_{-1}}$ , where  $T_{-1}$  is the first hitting time of  $-1$ .

*Solution.*

$$\begin{aligned} G(s) &= \mathbb{E}_0[s^{T_{-1}}] \\ &= p \mathbb{E}_0[s^{T_{-1}} \mid X_1 = 1] + q \mathbb{E}_0[s^{T_{-1}} \mid X_1 = -1] \\ &= p \mathbb{E}_1[s^{1+T_{-1}}] + q s \\ &= p s \mathbb{E}_0[s^{T_{-2}}] + q s \end{aligned}$$

Since  $s^\infty = 0$  by our convention, we have

$$\begin{aligned} \mathbb{E}_0[s^{T_{-2}}] &= \mathbb{E}_0[s^{T_{-2}} \mathbf{1}_{T_{-1} < \infty}] \\ &= \sum_m \mathbb{E}_0[s^{T_{-2}} \mathbf{1}_{T_{-1}=m}] \\ &= \sum_m \Pr_0(T_{-1} = m) \mathbb{E}_{-1}[s^{m+T_{-2}}] \\ &= \sum_m \Pr_0(T_{-1} = m) s^m \mathbb{E}_{-1}[s^{T_{-2}}] \\ &= \mathbb{E}_0[s^{T_{-1}}] \sum_m \Pr_0(T_{-1} = m) s^m \\ &= G(s)^2 \end{aligned}$$

Thus

$$\begin{aligned} G(s) &= psG(s)^2 + qs \\ G(s) &= \frac{1 \pm \sqrt{1 - 4pqs^2}}{2ps} \end{aligned}$$

**Claim:**  $G(s) = \frac{1 - \sqrt{1 - 4pqs^2}}{2ps}$  for all  $s \in (-1, 1) \setminus \{0\}$ .

We get several results from this exercise. The probability of ever hitting  $-1$  is

$$\begin{aligned} 1 - \Pr_0(T_{-1} = \infty) &= \lim_{s \uparrow 1} G(s) \\ &= \frac{1 - \sqrt{1 - 4pq}}{2p} \\ &= \frac{1 - |2p - 1|}{2p} \\ &= \begin{cases} 1 & \text{if the walk is left-biased,} \\ \frac{q}{p} & \text{otherwise.} \end{cases} = \frac{q}{p} \wedge 1 \end{aligned}$$

Another way to see the left-biased case is to note that  $X_n$ 's are sums of iid  $\text{Ber}_{\pm}(p)$  random variables, and so by the strong law of large numbers,

$$\frac{X_n}{n} \xrightarrow{\text{a.s.}} p - q.$$

Thus if  $p < q$ , then  $X_n \xrightarrow{\text{a.s.}} -\infty$ .

**Exercise 1.31.** Consider the SRW( $p$ ) on  $\mathbb{Z}$  started at the origin. Show that  $T_{-2} \stackrel{d}{=} \tau_1 + \tau_2$ , where  $\tau_1$  and  $\tau_2$  are iid copies of  $T_{-1}$ .

*Solution.*

$$\begin{aligned}
\mathbb{E}_0[s^{T-2}] &= \mathbb{E}_0[s^{T-2} \mathbf{1}_{T-1 < \infty}] \\
&= \sum_m \mathbb{E}_0[s^{T-2} \mathbf{1}_{T-1=m}] \\
&= \sum_m \Pr_0(T-1 = m) \mathbb{E}_{-1}[s^{m+T-2}] \\
&= \sum_m \Pr_0(T-1 = m) s^m \mathbb{E}_{-1}[s^{T-2}] \\
&= \mathbb{E}_0[s^{T-1}] \sum_m \Pr_0(T-1 = m) s^m \\
&= G(s)^2
\end{aligned}$$

On the other hand,  $\mathbb{E}_0[s^{\tau_1+\tau_2}] = \mathbb{E}_0[s^{\tau_1}]^2 = G(s)^2$  by independence.

Thus,  $T_{-2} \stackrel{d}{=} \tau_1 + \tau_2$  by fact 1.29.

**Exercise 1.32.** Consider the SRW( $p$ ) on  $\mathbb{Z}$  started at the origin. Find  $\Pr_0(T_{-1} = n)$ ,  $n \in \mathbb{N}$ .

*Solution.* We have  $G(s) = \frac{1 - \sqrt{1 - 4ps^2}}{2ps}$ . We first need a nice expression for  $\binom{1/2}{k}$  in order to use the binomial theorem.

$$\begin{aligned}
\binom{1/2}{k} &= \frac{\frac{1}{2}(\frac{1}{2} - 1) \dots (\frac{1}{2} - k + 1)}{k!} \\
&= \frac{1}{2^k} \frac{1(1-2) \dots (1-2k+2)}{k!} \\
&= \frac{(-1)^{k-1}}{2^k} \frac{(2k-3)!!}{k!}
\end{aligned}$$

We can rewrite  $(2k-3)!!$  as

$$(2k-3)!! = \frac{(2k-3)!}{(2k-4)!!} = \frac{(2k-3)!!}{2^{k-2}(k-2)!}$$

so that

$$\begin{aligned}
\binom{1/2}{k} &= \frac{(-1)^{k-1}}{2^{2k-2}} \frac{(2k-3)!}{k!(k-2)!} \\
&= \frac{(-1)^{k-1}}{2^{2k-2}k} \binom{2k-3}{k-2}
\end{aligned}$$

but multiplying by  $\frac{2k-2}{2(k-1)}$  yields an even nicer

$$= \frac{(-1)^{k-1}}{2^{2k-1}} \frac{1}{k} \binom{2k-2}{k-1}$$

The expression doesn't make sense for  $k = 0$ , for which the coefficient is 1, and the derivation doesn't make sense for  $k = 1$ , but for which the expression happens to match the coefficient.

But if we look closely, we see that this is just

$$\frac{(-1)^{k-1}}{2^{2k-1}} C_{k-1}$$

where  $C_{k-1}$  is the  $(k-1)$ th [Catalan number](#)!

From the binomial theorem,

$$\begin{aligned} (1-x)^{\frac{1}{2}} &= \sum_{k=0}^{\infty} \binom{1/2}{k} (-x)^k \\ &= 1 - \sum_{k=1}^{\infty} \frac{1}{2^{2k-1}} C_{k-1} x^k \end{aligned}$$

but more interestingly,

$$(1-4x)^{\frac{1}{2}} = 1 - 2 \sum_{k=1}^{\infty} C_{k-1} x^k$$

In fact, this gives that

$$\frac{1 - \sqrt{1-4x}}{2x} \text{ is the generating function for } (C_k)_{k \in \mathbb{N}}.$$

Getting back to the problem at hand, we have

$$\begin{aligned} G(s) &= \frac{1 - \sqrt{1-4pqs^2}}{2ps} \\ &= qs \sum_{k=0}^{\infty} C_k (pqs^2)^k \end{aligned}$$

So we have

$$\Pr_0(T_{-1} = n) = \begin{cases} 0 & \text{if } n \text{ is even} \\ p^k q^{k+1} C_k & \text{if } n = 2k+1 \end{cases}$$

In the case of  $p = q = \frac{1}{2}$ , this in fact proves that the number of SRW paths of length  $2n$  from the origin to the origin, that never go below the origin, is the  $n$ th Catalan number. Why? Because each such path can be extended

bijectively to a path of length  $2n + 1$  that hits  $-1$  for the first time at time  $2n + 1$ .

**Lecture**  
**04:** Sun  
04 Feb  
'24

## 1.1 Transience & Recurrence

**Definition 1.33.** Let  $(X_n)_{n \in \mathbb{N}}$  be  $MC_S(\mu_0, P)$ . Define  $T_y = ??? = \inf\{n \in \mathbb{N}^* \mid X_n = y\}$ , where we take  $\inf \emptyset$  to be  $+\infty$ .

For  $x, y \in S$ , define  $f_{xy} = \Pr_x(T_y < \infty)$ . A state  $x \in S$  is said to be *recurrent* if  $f_{xx} = 1$ , and *transient* otherwise.

A state  $x$  is said to be *absorbing* if  $f_{xy} > 0$  only when  $x = y$ .

We further define

$$N_y = \#\{n \in \mathbb{N}^* \mid X_n = y\}$$

$$G(x, y) = E_x[N_y]$$

**Lemma 1.34.** For all  $x, y \in S$ ,  $G(x, y) = \sum_{n \in \mathbb{N}} p_{xy}^{(n)}$ .

*Proof.* We write  $N_y$  as  $\sum_{n=1}^{\infty} \mathbf{1}_{X_n=y}$ . Then,

$$\begin{aligned} G(x, y) &= E_x \left[ \sum_{n=1}^{\infty} \mathbf{1}_{X_n=y} \right] \\ &= \sum_{n=1}^{\infty} E_x[\mathbf{1}_{X_n=y}] && \text{(MCT)} \\ &= \sum_{n=1}^{\infty} \Pr_x(X_n = y) \\ &= \sum_{n=1}^{\infty} p_{xy}^{(n)} \end{aligned}$$

The interchange of the sum and the expectation is justified by the monotone convergence theorem stated below.  $\square$

**Theorem 1.35** (Monotone convergence theorem). Let  $(\Omega, \mathcal{F}, \Pr)$  be a probability space. Let  $X_n: \Omega \rightarrow [0, \infty]$  be a sequence of random variables and  $X: \Omega \rightarrow [0, \infty]$  be another random variable. Suppose that  $X_n(\omega) \leq X_{n+1}(\omega)$  for each  $n$  and  $\omega$ , and that  $X_n(\omega) \rightarrow X(\omega)$  for each  $\omega$ . Then,  $E[X_n] \rightarrow E[X]$ .

*Remark.* The statement holds even if  $X_n \xrightarrow{\text{a.s.}} X$ .

**Theorem 1.36.** For all  $x, y \in S$ ,

$$\Pr_x(N_y = m) = \begin{cases} 1 - f_{xy} & \text{if } m = 0 \\ f_{xy}f_{yy}^{m-1}(1 - f_{yy}) & \text{if } m \in \mathbb{N}^* \\ f_{xy}[f_{yy} = 1] & \text{if } m = +\infty \end{cases}$$

*Proof.*  $N_y = 0$  if and only if  $T_y = +\infty$ . This occurs with probability  $1 - f_{xy}$ .

We define  $T_y^{(1)} = T_y$  and for  $m \geq 1$ ,

$$T_y^{(m+1)} = \inf\{n > T_y^{(m)} \mid X_n = y\}.$$

Note that  $T_y^{(m)} = +\infty$  implies  $T_y^{(m+1)} = +\infty$ . Now

$$\begin{aligned} \Pr_x(T_y^{(m+1)} < \infty) &= \Pr_x(T_y^{(m)} < \infty \text{ and } T_y^{(m+1)} < \infty) \\ &= \Pr_x(T_y^{(m)} < \infty) \Pr_y(T_y < \infty) \quad (\text{Strong Markov property}) \end{aligned}$$

and by induction,

$$\begin{aligned} \Pr_x(N_y \geq m) &= \Pr_x(T_y^{(m)} < \infty) \\ &= f_{xy}f_{yy}^{m-1}. \end{aligned}$$

The result follows by taking the difference. Or more directly,

$$\Pr_x(N_y = m) = \Pr_x(T_y^{(m)} < \infty) \Pr_y(T_y = +\infty) = f_{xy}f_{yy}^{m-1}(1 - f_{yy}).$$

Finally,

$$\begin{aligned} \Pr_x(N_y = +\infty) &= 1 - \sum_{m=0}^{\infty} \Pr_x(N_y = m) \\ &= 1 - (1 - f_{xy}) - f_{xy}(1 - f_{yy}) \sum_{m=0}^{\infty} f_{yy}^m \\ &= \begin{cases} f_{xy} & \text{if } f_{yy} = 1 \\ 0 & \text{if } f_{yy} < 1 \end{cases} \\ &= f_{xy}[f_{yy} = 1]. \end{aligned}$$

□

**Theorem 1.37.**

- (1) Suppose  $y$  is transient. Then for all  $x \in S$ ,  $\Pr_x(N_y < \infty) = 1$  and  $G(x, y) = \frac{f_{xy}}{1-f_{yy}} < \infty$ .
- (2) Suppose  $y$  is recurrent. Then  $\Pr_y(N_y = \infty) = 1$  and  $G(y, y) = +\infty$ . Further, for all  $x \in S \setminus \{y\}$ ,  $\Pr_x(N_y = \infty) = f_{xy}$  and

$$G(x, y) = \begin{cases} 0 & \text{if } f_{xy} = 0, \\ \infty & \text{if } f_{xy} > 0. \end{cases}$$

*Proof.*

- (1) Since  $y$  is transient,  $f_{yy} < 1$ . Thus by the previous theorem,  $\Pr_x(N_y = \infty) = 0$ . Then,

$$\begin{aligned} G(x, y) &= \sum_{m=1}^{\infty} m \Pr_x(N_y = m) \\ &= f_{xy}(1 - f_{yy}) \sum_{m=1}^{\infty} m f_{yy}^{m-1} \\ &= f_{xy}(1 - f_{yy}) \frac{1}{(1 - f_{yy})^2} \\ &= \frac{f_{xy}}{1 - f_{yy}}. \end{aligned}$$

- (2) Since  $y$  is recurrent,  $f_{yy} = 1$ . By the previous theorem, for any  $x \in S$ ,

$$\Pr_x(N_y = m) = \begin{cases} 1 - f_{xy} & \text{if } m = 0, \\ 0 & \text{if } m \in \mathbb{N}^*, \\ f_{xy} & \text{if } m = +\infty. \end{cases}$$

Thus  $G(x, y) = +\infty$  if  $f_{xy} > 0$  and 0 otherwise.  $\square$

**Corollary 1.38.** A state  $x$  is recurrent iff  $G(x, x) = \sum_{m=1}^{\infty} p_{xx}^{(m)} = +\infty$ .

**Definition 1.39.** A DTMC is said to be *recurrent* (resp. *transient*) if all states are recurrent (resp. transient).

**Theorem 1.40.** If  $|S| < \infty$ , then there exists a recurrent state.



*Proof.* Suppose not. Then for all  $x, y \in S$ ,  $G(x, y) = \sum_{m=0}^{\infty} p_{xy}^{(m)} < \infty$ . Then the individual terms of the series must tend to 0. Thus

$$\begin{aligned} \sum_{y \in S} \lim_{m \rightarrow \infty} p_{xy}^{(m)} &= 0 \\ \implies \lim_{m \rightarrow \infty} \sum_{y \in S} p_{xy}^{(m)} &= 0. \end{aligned}$$

The interchange of the limit and the sum is justified since the sum is finite. But  $\sum_{y \in S} p_{xy}^{(m)} = 1$  for all  $m$ . Thus we have a contradiction.  $\square$

**Theorem 1.41.** Suppose  $x \neq y \in S$ ,  $x$  is recurrent and  $x \rightsquigarrow y$ . Then  $y$  is recurrent,  $y \rightsquigarrow x$  and  $f_{xy} = f_{yx} = 1$ .

By  $x \rightsquigarrow y$ , we mean that there exists  $n \in \mathbb{N}^*$  such that  $p_{xy}^{(n)} > 0$ . In other words,  $f_{xy} > 0$ .

*Proof.* Since  $x \rightsquigarrow y$ , there exists  $n \in \mathbb{N}^*$  and  $x_1, \dots, x_{n-1}$  distinct from  $x$  such that  $p_{xx_1} p_{x_1 x_2} \dots p_{x_{n-1} y} > 0$ . Since  $x$  is recurrent,

$$0 = \Pr_x(T_x = +\infty) \geq p_{xx_1} p_{x_1 x_2} \dots p_{x_{n-1} y} \Pr_y(T_x = +\infty)$$

so  $\Pr_y(T_x = +\infty)$  must be 0. Thus  $y \rightsquigarrow x$  with  $f_{yx} = 1$ . If  $y$  is recurrent, then  $f_{xy}$  would be 1 by the same argument. Thus we need only show that  $y$  is recurrent.  $\square$

**Lecture**  
**05:** Sun  
04 Feb  
'24