3.7 Time reversal

Intuition suggests that the Markov chain $X_{n,n\geq 0}$ describing the diffusion of particles in Ehrenfest's model (cf Ex. 3.6.8) remains a Markov chain with time running backwards

→ **Theorem 3.7.1**: Let $X_{n,n\geq 0}$ be a irreducible and Markov(π, P) with invariant distr. $\pi = (\pi_i)_{i\in I}$. Then $Y_{n,n=0,\dots,N}$ defined as

$$Y_n \overset{\text{def}}{=} X_{N-n,n=0,\dots,N}$$

is Markov$(\pi, \hat{P})$, where $\hat{P} = (\hat{P}_{ij})$ is given by

$$\pi_j \hat{P}_{ji} = \pi_i P_{ij} \text{ for all } i,j$$

$$\left(\overset{\text{Th. 3.5.7}}{\Longleftrightarrow} \hat{P}_{ji} = \frac{\pi_i}{\pi_j} P_{ij}\right)$$

Proof: Direct consequence of Th. 3.1.3

Def. 3.7.2 (reversible Markov chains)
Let $X_{n,n\geq 0}$ be irreducible and Markov(λ, P). Then $X_{n,n\geq 0}$ is called reversible if

$$Y_n \overset{\text{def.}}{=} X_{N-n,n=0,\dots,N}$$

is also Markov(λ, P) for all $N \geq 1$

The property of reversibility is closely linked to the notion of "detailed balance".

Def. 3.7.3: Let $X_{n,n\geq 0}$ be Markov(λ, P). Then P and λ are said to be in detailed balance if for all i,j

$$\lambda_i P_{ij} = \lambda_j P_{ji} \qquad (3.7.1)$$

Lemma 3.7.4: Let P and λ be in detailed balance. Then λ is an invariant distr. i.e.

$$\lambda \cdot P = \lambda$$

Proof: $\lambda \cdot P = ((\lambda P)_i)_{i\in I}$. We have that

$$(\lambda \cdot P)_i \overset{\text{matrix mult.}}{=} \sum_{j\in I} \lambda_j P_{ji} \overset{(3.7.1)}{=} \sum_{j\in I} \lambda_i P_{ij} = \lambda_i \underbrace{\left(\sum_{j\in I} P_{ij}\right)}_{=1} = \lambda_i \Rightarrow \text{proof}$$

Def 3.7.3, L 3.7.4 → Th. 3.7.1

Theorem 3.7.5 (Characterization of reversible $X_{n,n\geq 0}$)
Let $X_{n,n\geq 0}$ be irreducible (i.e. $i \Longleftrightarrow j$ for all i,j) and Markov(λ, P). Then the following statements are equivalent:

(65) (i) $X_{n,n\geq 0}$ reversible

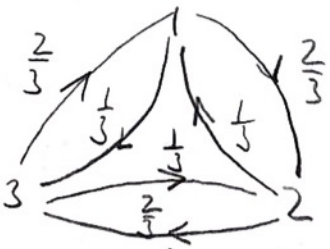
(ii) P and λ are in detailed balance

(65) (i) $X_{n,n\geq 0}$ reversible

(ii) P and λ are in detailed balance

Ex. 3.7.6 (counterexample)

$$P = \begin{pmatrix} 0 & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{pmatrix}$$



irreducible

$\longrightarrow \lambda \overset{\text{def}}{=} (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ inv. distr., i.e. $\lambda \cdot P = \lambda$ and $\lambda_1 + \lambda_2 + \lambda_3 = 1$

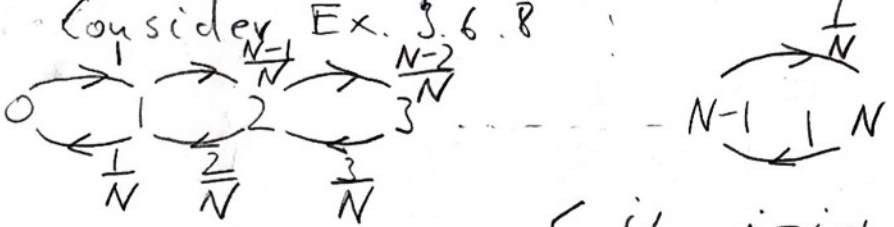
Let $X_{n,n\geq 0}$ be Markov(λ, P)

$\lambda_1 \cdot p_{12} \neq \lambda_2 \cdot p_{21}$ ($\lambda_1 = \frac{1}{3}$, $p_{12} = \frac{2}{3}$, $\lambda_2 = \frac{1}{3}$, $p_{21} = \frac{1}{3}$) $\overset{\text{Def 3.7.3}}{\Longrightarrow}$ P and λ not in detailed balance

$\Longrightarrow$ $X_{n,n\geq 0}$ not reversible

Ex. 3.7.7 (model of Ehrenfest)

Consider Ex. 3.6.8



$\longrightarrow$ trans. prob. $p_{ij} = \begin{cases} i/N, & j = i-1 \\ (N-i)/N, & j = i+1 \\ 0 & \text{else} \end{cases}$ (+)

$\longrightarrow$ $\lambda_i = \binom{N}{i} \frac{1}{2^N}$, $i \geq 0$ (unique) invariant distr. (++)

Check $\lambda_i \cdot p_{ij} = \lambda_j \cdot p_{ji}$ for all i, j (+++)

Fix $i \geq 0$

1. case: $j \neq i-1, i+1$

$\Longrightarrow$ both sides of (+++) zero

2. case: $j = i+1$

$$\Longrightarrow \lambda_i \cdot p_{ij} \overset{(+),(++)}{=} \binom{N}{i} \frac{1}{2^N} \frac{N-i}{N} = \frac{1}{2^N} \frac{N!}{(N-i)!\, i!} \frac{N-i}{N}$$

$$= \frac{1}{2^N} \frac{N! \,(N-i)\cdot(i+1)}{(N-i-1)!\,(N-i)\, i!\,(i+1)\, N} = \frac{1}{2^N} \frac{N!}{(N-i-1)!\,(i+1)!} \frac{i+1}{N}$$

$$= \underbrace{\frac{1}{2^N} \binom{N}{i+1}}_{\lambda_{i+1}} \cdot \underbrace{\frac{i+1}{N}}_{p_{i+1,i}} \overset{j=i+1}{=} \lambda_j \cdot p_{ji}$$

3. case: $j = i-1$ similar

$\Longrightarrow$ (+++) satisfied

$\Longrightarrow$ P and λ in detailed balance $\overset{\text{Th. 3.7.5}}{\Longrightarrow}$

(66) Ehrenfest's model $X_n, n\geq 0$ with initial distr. $\lambda_i = \binom{N}{i}\frac{1}{2^N}$ is reversible, i.e. $(X_{N-n})_{0\leq n\leq N}$ is Markov(λ, P)

Ex. 3.7.8 (Random walk on a graph)

graph G $\longrightarrow$ countable collection of vertices, some of which are joined by edges

valency v_i of a vertex i $\longrightarrow$ number of edges at i

A Markov chain $X_n, n\geq 0$ with state space $I = G$ can be e.g. introduced by defining

$$p_{ij} = \begin{cases} \frac{1}{v_i}, & \text{if } (i,j) \text{ edge} \\ 0, & \text{else} \end{cases} \qquad (+)$$

provided $v_i < \infty$ for all vertices i

Assumption: $\delta \overset{\text{def}}{=} \sum_{i\in G} v_i < \infty$ and G connected

Define

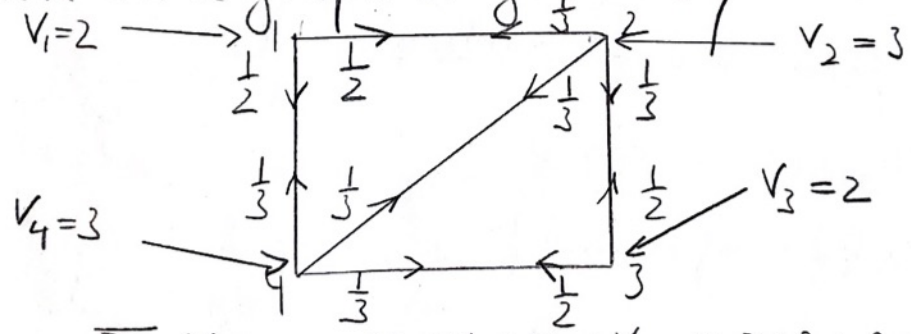
$$\lambda_i \overset{\text{def}}{=} \frac{v_i}{\delta}, \quad i \in G \qquad (++)$$

$$\Rightarrow \sum_i \lambda_i p_{ij} \overset{(+),(++)}{=} \sum_{(i,j)\text{edge}} \frac{v_i}{\delta}\cdot\frac{1}{v_i} = \frac{1}{\delta}\left(\sum_{(i,j)\text{edge}} 1\right) = \frac{1}{\delta} v_j \overset{(++)}{=} \lambda_j$$

($\sum_{(i,j)\text{edge}} 1$ = number of edges at j)

$$\Leftrightarrow \lambda \cdot P = \lambda \Rightarrow \lambda \text{ inv. distr.}$$

e.g. X_n position of a particle at time n which performs a rand. walk on a graph G given by



$$\Rightarrow \delta = \sum_{i\in G} v_i = v_1+v_2+v_3+v_4 = 2+3+2+3 = 10$$

$$\overset{(++)}{\Rightarrow} \lambda = (\lambda_1,\lambda_2,\lambda_3,\lambda_4) = \left(\frac{2}{10}, \frac{3}{10}, \frac{2}{10}, \frac{3}{10}\right) \text{ inv. distr.}$$

$$\Rightarrow m_3 \overset{\text{def}}{=} E[T_3 \mid X_0 = 3] \overset{\text{Th. 3.5.7}}{=} \frac{1}{\lambda_3} = \frac{10}{2} = 5$$

$\longrightarrow$ expected time that the particle starting in vertex 3 returns to this vertex is 5

One also checks that

Th. 3.7.5 $\Rightarrow$ $\lambda_i \cdot p_{ij} = \lambda_j \cdot p_{ji}$ for all i, j

$\Rightarrow$ $X_n, n\geq 0$ reversible

3.8 Ergodic theorem

→ We want to study the long-time average behaviour of Markov chains $X_n, n \geq 0$ as e.g.

$$\frac{X_0 + X_1 + \dots + X_n}{n}$$

for large n

→ ergodic theorems are a tool to analyze this problem

→ special version of an ergodic theorem

Theorem 3.8.1 (Strong law of large numbers) (indep.; identically distr.)

Let $Y_n \geq 0$, $n \geq 0$ be a seq. of i.i.d. rand. variables with

$$\mu \stackrel{\text{def}}{=} E[Y_1] \quad (\in [0, \infty])$$

Then

$$\frac{Y_1 + Y_2 + \dots + Y_n}{n} \xrightarrow[n \to \infty]{} \mu$$

with prob. 1

Proof: See basic course of statistics.

We need the following def.:

$$V_i(n) \stackrel{\text{def}}{=} \sum_{k=0}^{n-1} \mathbb{1}_{\{X_k = i\}} \stackrel{\text{def}}{=} \begin{cases} 1, & \text{if } X_k = i \\ 0 & \text{else} \end{cases} \tag{3.8.1}$$

→ number of visits to i before n

→ $\frac{V_i(n)}{n}$ proportion of time spent in state i

→ For large n $\frac{V_i(n)}{n}$ is the long-run proportion of time spent by $X_n, n \geq 0$ in i

→ **Theorem 3.8.2** (Ergodic theorem)

Let $X_n, n \geq 0$ be irreducible (i.e. $i \leftrightarrow j$ for all i, j) and Markov (λ, P). Then

(i)
$$\frac{V_i(n)}{n} \xrightarrow[n \to \infty]{} \frac{1}{m_i} \quad \text{with prob. 1},$$

where

$$m_i = E[T_i \mid X_0 = i] \quad (\in [0, \infty])$$

is the expected return time to i

(ii) If $X_n, n \geq 0$ is positive recurrent (i.e. $m_i < \infty$) then

$$\frac{1}{n} \sum_{k=0}^{n-1} f(X_k) \xrightarrow[n \to \infty]{} \overline{f} \quad \text{with prob. 1},$$

for all bounded functions $f : S \to \mathbb{R}$,

where

$$\overline{f} \overset{\text{def}}{=} \sum_{i \in \mathbb{N}_0} \pi_i f(i)$$

for the unique inv. distr. $\pi = (\pi_i)$ of X_n, $n \geq 0$ (see Th. 3.5.7)

Rem. 3.8.3: $\mathbb{N}_0$ finite $\Rightarrow$ Th. 3.8.2 holds for all f, since all $f: \mathbb{N}_0 \to \mathbb{R}$ are bounded

Ex. 3.8.4: Car insurance of Ex. 3.2.7

$\longrightarrow$ X_n risk category of the customer in year n

with
$$P = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.6 & 0.2 & 0.2 \\ 0 & 0.6 & 0.4 \end{pmatrix}$$

ergodic $\overset{\text{Theorem}}{\Longrightarrow}$
$$\frac{V_3(n)}{n} \xrightarrow[n \to \infty]{} \frac{1}{m_3} \quad \text{with. prob. 1}$$

$m_3 \overset{\text{Th. 3.5.7}}{=} \frac{1}{\pi_3} \overset{\text{Ex. 3.5.8}}{=} 5.8$

$\longrightarrow$ long-run proportion of time spent by the customer in the highest risk category is 17.2%

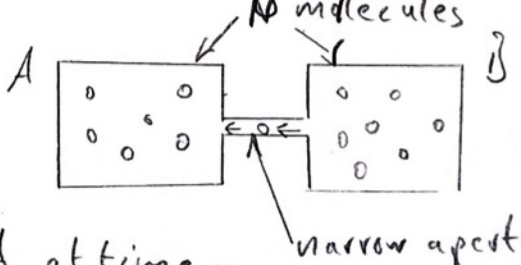
From (iii) of Th. 3.8.2 and Rem. 3.8.3 it follows for $f(x) = x$ that

$$\frac{X_0 + X_1 + \ldots + X_{n-1}}{n} \xrightarrow[n \to \infty]{} \overline{f} \quad \text{with. prob. 1,}$$

where
$$\overline{f} \overset{\text{def}}{=} \sum_{i=1}^{3} \pi_i \cdot i \overset{\text{Ex. 3.5.8}}{=} \frac{3.6}{5.8} + 2 \cdot \frac{1.2}{5.8} + \frac{3}{5.8} \approx 1.552$$

$\longrightarrow$ long-run average category of the insured is 1.552, i.e. between category 1 and 2

Ex. 3.8.5 (Ehrenfest's model)



$\longrightarrow$ X_n number of molecules in A at time n

$\longrightarrow$ $\pi_i = \binom{N}{i} \frac{1}{2^N} = \binom{N}{i} \left(\frac{1}{2}\right)^{N-i} \left(\frac{1}{2}\right)^i$ inv. distr. (i.e. a binom. distr.)

Th. 3.8.2(ii) $\overset{\text{Rem. 3.8.2}}{\Longrightarrow}$
$$\frac{X_0 + X_1 + \ldots + X_{n-1}}{n} \xrightarrow[n \to \infty]{} \overline{f},$$

where
$$\overline{f} \overset{\text{def}}{=} \sum_{i=0}^{N} \pi_i \cdot i \overset{\text{def. of expect.}}{=} E[Y]$$

for a binomially distributed r.v. Y with parameters N and $p = \frac{1}{2}$

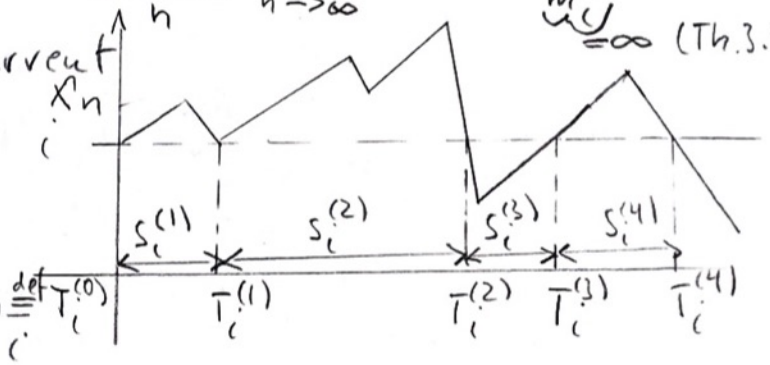
(69) $\overset{\text{basic course in statistics}}{\Longrightarrow}$ $E[Y] = N \cdot p = N \cdot \frac{1}{2}$

$N=10 \Longrightarrow$ long-run average number of molecules in container A is 5

Proof of Th. 3.8.2 : Without loss of generality $P(X_0 = i) = 1$

(i) 1. case : states i are transient $\Longrightarrow$ state i is atmost finitely often visited $\Longrightarrow$ $\frac{V_i(n)}{n} \xrightarrow[n\to\infty]{} 0 = \frac{1}{m_i} = 0$ $(m_i = \infty$ (Th. 3.5.7))

2. case : states i are recurrent



$T_i^{(r)}$ r-th passage time to i

$S_i^{(r)}$ length of the r-th excursion to i

$\overset{\text{recurrence}}{\longrightarrow}$ $T_i^{(r)} < \infty$ with prob. 1 for all i, r

$\overset{\text{L.3.4.3}}{\longrightarrow}$ $S_i^{(1)}, S_i^{(2)}, \ldots$ i.i.d with $E[S_i^{(r)}] = m_i$ for all r

From the def. $V_i(n)$ we see that

$\underbrace{S_i^{(1)} + \ldots + S_i^{(V_i(n)-1)}}_{\text{last visit to } i \text{ before } n} \leq n-1$

as well as

$\underbrace{S_i^{(1)} + \ldots + S_i^{(V_i(n))}}_{\text{the first visit to } i \text{ after } n-1} \geq n$

$\Longrightarrow$ $\frac{S_i^{(1)} + \ldots + S_i^{(V_i(n)-1)}}{V_i(n)} \leq \frac{n}{V_i(n)} \leq \frac{S_i^{(1)} + \ldots + S_i^{(V_i(n))}}{V_i(n)}$ (+)

Because of Th. 3.8.1 we know that

$\frac{S_i^{(1)} + \ldots + S_i^{(n)}}{n} \xrightarrow[n\to\infty]{} m_i$ with prob. 1 (++)

On the other hand, since $X_n, n \geq 0$ is recurrent we see that

$V_i(m) \xrightarrow[m\to\infty]{} \infty$ with prob. 1

Replacing n in (++) by $V_i(m)$ or $V_i(m)-1$ shows that

$\frac{S_i^{(1)} + \ldots + S^{(V_i(m))}}{V_i(m)} \xrightarrow[m\to\infty]{} m_i$

as well as

$\frac{S_i^{(1)} + \ldots + S_i^{(V_i(m)-1)}}{V_i(m)} \xrightarrow[m\to\infty]{} m_i$

$\overset{(+)}{\Longrightarrow}$ $\frac{m}{V_i(m)} \xrightarrow[m\to\infty]{} m_i \Longrightarrow \frac{V_i(m)}{m} \xrightarrow[m\to\infty]{} \frac{1}{m_i} \Longrightarrow$ (i)