

MARKOV CHAINS

6. Invariant/equilibrium measures and distributions. Positive and null recurrence.

Invariant distributions, statement of existence and uniqueness up to constant multiples. Mean return time, positive recurrence; equivalence of positive recurrence and the existence of an invariant distribution.

I will start with a list of definitions and basic facts about invariant/equilibrium measures and distributions: most of these facts will be proved in subsequent lectures.

Given a state $i \in I$, we can assign to it a non-negative ‘mass’ $\mu_i \geq 0$, then

$$\mu(A) = \sum_{i \in A} \mu_i \tag{6.1}$$

gives the mass of a set $A \subseteq I$. This defines a non-negative *measure* on I . If the total mass is one: $\mu(I) = \sum_{i \in I} \mu_i = 1$, we obtain a *probability distribution* on I . Otherwise, we can normalise:

$$\pi_i = \frac{\mu_i}{\sum_{j \in I} \mu_j} \tag{6.2}$$

and produce a probability distribution $\pi = \{\pi_i\}$, provided that the original measure μ had a finite total mass: $\mu(I) = \sum_{i \in I} \mu_i < 1$.

Definition 6.1 We say that a μ measure is in-
probability distribution π

variant (or equilibrium), for a transition matrix P , if

$$\begin{aligned} \mu P &= \mu, & \sum_{i \in I} \mu_i P_{ij} &= \mu_j \\ \pi P &= \pi, & \sum_{i \in I} \pi_i P_{ij} &= \pi_j \end{aligned} \quad \text{that is,} \quad \forall j \in I. \quad (6.3)$$

We see that an invariant measure (IM) $\mu = (\mu_i)$ and an equilibrium distribution (ED) $\pi = (\pi_i)$ form row-eigenvectors of matrix P , with the eigenvalue 1¹ (π normalised and μ in general not). A nonnegative multiple $a\mu$, $a \geq 0$, of an IM μ is again an IM; more generally, any nonnegative linear combination $\sum_l a_l \mu^{(l)}$ of IMs $\mu^{(l)}$ is again an IM.

If a chain is finite (i.e., $\#I < +\infty$), an ED π always exists. If a chain is countable ($\#I = +\infty$) and irreducible, we can only guarantee existence of an IM μ : if $\mu(I) = \sum_i \mu_i < +\infty$ then an ED π also exists, by (6.2).

An ED π , when it exists, **may** be **non-unique**. In fact, each closed and recurrent communicating class $C^{(l)}$ **may** ‘support’ its own ED $\pi^{(l)}$, with entries

$$\begin{cases} \pi_j^{(l)} = 0 & j \notin C^{(l)}, \\ \pi_i^{(l)} > 0 & i \in C^{(l)}, \end{cases} \quad (6.4.1)$$

and

$$\pi^{(l)}(C^{(l)}) = \sum_{i \in C^{(l)}} \pi_i^{(l)} = 1 \quad (6.4.2)$$

That is, ED $\pi^{(l)}$ is concentrated entirely on the whole of $C^{(l)}$. If, for a given class $C^{(l)}$, such an ED $\pi^{(l)}$ exists, it is **unique**.

On the other hand, open communicating classes and closed transient communicating classes do **not** support equilibrium distributions.

The distinction between recurrent CCs that support their own EDs and those that do not is delicate: it leads to the concepts of *positive recurrent* and *null recurrent* communicating classes. See below.

¹One should not confuse two equations $\mu P = \mu$ (invariance) and $Ph = h$ (the hitting time equation).

In any case, each **finite** closed CC *has* its own unique ED $\pi^{(l)}$, and for a finite MC, every ED π is a linear combination of such EDs $\pi^{(l)}$:

$$\pi = \sum_{C^{(l)} : \text{closed CC}} a_l \pi^{(l)} \quad \text{with } a_l \geq 0 \quad \text{and} \quad \sum_l a_l = 1. \quad (6.5.1)$$

For a countable MC, this becomes the decomposition over positive recurrent closed CCs:

$$\pi = \sum_{C^{(l)} : \text{positive recurrent CC}} a_l \pi^{(l)} \quad \text{with } a_l \geq 0 \quad \text{and} \quad \sum_l a_l = 1. \quad (6.5.2)$$

Example 6.1 Consider the 2×2 transition matrix

$$P = \begin{pmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{pmatrix}.$$

Then (a) if $\alpha + \beta > 0$, it has a unique equilibrium distribution

$$\pi = \left(\frac{\beta}{\alpha + \beta}, \frac{\alpha}{\alpha + \beta} \right),$$

(b) if $\alpha = \beta = 0$ then $P = \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ (a 2×2 unit matrix), and every row-vector $(x; y)$ is invariant. Thus, in case (b) every stochastic vector $\pi = (q; 1 - q)$, $0 \leq q \leq 1$, gives an ED. But the chain has two absorbing states each forming a closed communicating class, and we have the representation

$$\pi = q\delta_0 + (1 - q)\delta_1 \quad \text{where} \quad \begin{array}{l} \delta_0 = (1; 0), \text{ an ED supported by CC } \{0\}, \\ \delta_1 = (0; 1), \text{ an ED supported by CC } \{1\}. \end{array}$$

This agrees with representation (6.5.1): $\pi = a_0\pi^{(0)} + a_1\pi^{(1)}$ where $a_0 = q$, $a_1 = 1 - q$ and $\pi^{(0)} = \delta_0$, $\pi^{(1)} = \delta_1$.

An important fact is that if we choose the initial distribution $\lambda = \pi$, an ED, then it is preserved in time:

$$\mathbb{P}(X_0 = i) = \pi_i \quad \forall i \in I \quad \Rightarrow \quad \mathbb{P}(X_n = j) = \pi_j \quad \forall j \in I \quad \text{and } n \geq 0. \quad (6.6)$$

In fact, for $n = 1$:

$$\mathbb{P}(X_1 = j) = \sum_{i \in I} \mathbb{P}(X_0 = i, X_1 = j) = \sum_i \pi_i p_{ij} = \pi_j \quad (6.7)$$

And so on.

For this reason, a (π, P) -Markov chain is called *stationary*, or an MC *in equilibrium*.

From now on, until further notice, I will assume that our Markov chain is *irreducible*, i.e., has a single communicating class. Then it is recurrent or transient. If it is transient, it has no ED. If the chain is recurrent, then there will be a dichotomy: either it supports an ED π or it does not. We will study this dichotomy.

Recall: for a given a state $k \in I$, the random variable T_k stands for its return time. Set:

$$\begin{aligned} \gamma_i^k &= \mathbb{E}_k \sum_{n=0}^{T_k-1} \mathbf{1}(X_n = i) \\ &= \begin{cases} \mathbb{E}_k \left(\# \text{ of visits to } i \text{ before returning to } k \right), & \text{if } i \neq k \text{ (with } 1 \leq n < T_k), \\ 1, & \text{if } i = k \text{ (from } n = 0). \end{cases} \end{aligned} \quad (6.8)$$

Then $0 \leq \gamma_i^k \leq \infty$. Observe that

$$\begin{aligned} \sum_{i \in I} \gamma_i^k &= 1 + \sum_{i \in I: i \neq k} \mathbb{E}_k \left(\# \text{ of visits to } i \text{ before returning to } k \right) \\ &= 1 + \mathbb{E}_k(T_k - 1) = \mathbb{E}_k T_k. \end{aligned} \quad (6.9)$$

Consider row-vectors $\underline{\gamma}^k = (\gamma_i^k, i \in I)$, parametrised by $k \in I$.

Theorem 6.1 (a) *Suppose that the chain is irreducible and recurrent. Then, $\forall k \in I$, vector $\underline{\gamma}^k$ is invariant:*

$$\underline{\gamma}^k P = \underline{\gamma}^k. \quad (6.10.1)$$

That is, \forall state k :

$$(\underline{\gamma}^k P)_j := \sum_{i \in I} \gamma_i^k p_{ij} = \gamma_j^k, \quad j \neq k, \quad (6.10.2)$$

and

$$(\underline{\gamma}^k P)_k := \sum_{i \in I} \gamma_i^k p_{ik} = 1 = \gamma_k^k. \quad (6.10.3)$$

Furthermore,

$$0 < \gamma_i^k < \infty \quad \forall \text{ states } i, k \in I. \quad (6.11)$$

Hence, for an irreducible and recurrent matrix P , the vector $\underline{\gamma}^k$ is a ‘genuine’ invariant vector with strictly positive and finite entries, $\forall k \in I$.

(b) Suppose that the chain is irreducible and transient. Then $\forall k \in I$, equation (6.10.2) still holds, but

$$(\underline{\gamma}^k P)_k := \sum_{i \in I} \gamma_i^k p_{ik} < 1 = \gamma_k^k \quad (\text{sub-inbvariance}). \quad (6.12)$$

Hence, vector $\underline{\gamma}^k$ is invariant iff the chain is recurrent.

Proof (a) By the Markov property $\forall m \geq 2$ and states $i \neq k$ and $j \neq k$,

$$\mathbb{P}_k(T_k > m - 1, X_{m-1} = i) p_{ij} = \mathbb{P}_k(T_k > m - 1, X_{m-1} = i, X_m = j), \quad (6.13)$$

and

$$\mathbb{P}_k(T_k > m - 1, X_{m-1} = i) p_{ik} = \mathbb{P}_k(T_k = m, X_{m-1} = i). \quad (6.14)$$

Then, for $j \neq k$

$$\begin{aligned}
\gamma_j^k &= \mathbb{E}_k \sum_{0 \leq n \leq T_k - 1} \mathbf{1}(X_n = j) = \sum_{n \geq 1} \mathbb{E}_k \mathbf{1}(X_n = j, T_k > n) \\
&= \sum_{n \geq 1} \mathbb{P}_k(X_n = j, T_k > n) \\
&= p_{kj} + \sum_{n \geq 2} \left[\sum_{i: i \neq k} \mathbb{P}_k(T_k > n - 1, X_{n-1} = i, X_n = j) \right] \\
&= p_{kj} + \sum_{n \geq 2} \left[\sum_{i: i \neq k} \mathbb{P}_k(T_k > n - 1, X_{n-1} = i) p_{ij} \right] \text{ by (6.13)} \\
&= \gamma_k^k p_{kj} + \sum_{i: i \neq k} \sum_{n \geq 1} \mathbb{E}_k \mathbf{1}(T_k > n, X_n = i) p_{ij} = (\underline{\gamma}^k P)_j.
\end{aligned}$$

So far, we have not used the recurrence property. However, for $j = k$,

$$\begin{aligned}
(\underline{\gamma}^k P)_k &= \sum_{i \in I} \gamma_i^k p_{ik} = \gamma_k^k p_{kk} + \sum_{i: i \neq k} \gamma_i^k p_{ik} \\
&= p_{kk} + \sum_{i: i \neq k} \mathbb{E}_k \left[\sum_{1 \leq n < T_k} \mathbf{1}(X_n = i) \right] p_{ik} \\
&= p_{kk} + \sum_{i: i \neq k} \left[\sum_{n \geq 1} \mathbb{E}_k \mathbf{1}(T_k > n, X_n = i) p_{ik} \right] \\
&= p_{kk} + \sum_{i: i \neq k} \left[\sum_{n \geq 1} \mathbb{P}_k(T_k = n + 1, X_n = i) \right] \text{ by (6.14)} \\
&= p_{kk} + \sum_{n \geq 2} \mathbb{P}_k(T_k = n) \\
&= \sum_{n \geq 1} \mathbb{P}_k(T_k = n) = \mathbb{P}_k(T_k < \infty) := f_k.
\end{aligned}$$

Now, $f_k \equiv 1$ when the chain is irreducible and recurrent. Thus, for an irreducible and recurrent chain, all vectors $\underline{\gamma}^k$ are invariant.

We now want to check (6.11). Since P is irreducible, any pair of states $i, k \in I$ communicate. That is, $\forall i, k \in I \exists m, n \geq 0$ such that $p_{ik}^{(n)} > 0$ and

$p_{ki}^{(m)} > 0$. When P is recurrent, we know that vector $\underline{\gamma}^k$ is invariant; hence $\underline{\gamma}^k P^m = \underline{\gamma}^k P^n = \underline{\gamma}^k$. So,

$$\gamma_i^k = \sum_l \gamma_l^k p_{li}^{(m)} \geq \gamma_k^k p_{ki}^{(m)} = p_{ki}^{(m)}, \text{ hence } \gamma_i^k > 0.$$

On the other hand

$$1 = \gamma_k^k = \sum_l \gamma_l^k p_{lk}^{(n)} \geq \gamma_i^k p_{ik}^{(n)}, \text{ i.e. } \gamma_i^k \leq \frac{1}{p_{ik}^{(n)}}, \text{ hence } \gamma_i^k < +\infty.$$

(b) If the chain is irreducible and transient then the above calculation involving f_k still holds:

$$(\underline{\gamma}^k P)_k = f_k < 1 = \gamma_k^k \quad (\text{sub-invariance}).$$

This completes the proof of Theorem 6.1. QED

Theorem 6.2 Suppose that $\mu = (\mu_i)$ is an IM: $\mu P = \mu$ and $\mu_i \geq 0 \forall i \in I$. Suppose in addition that $\mu_k = 1$ for some given state k . Then (a) $\forall i \in I$:

$$\mu_i \geq \gamma_i^k.$$

(b) For an irreducible and recurrent matrix P the equality holds

$$\mu_i = \gamma_i^k, \quad \forall i \in I.$$

Proof (a) Invariance plus the fact that $\mu_k = 1$ imply that $\forall j \in I$ and $n \geq 1$

$$\begin{aligned} \mu_j &= \sum_i \mu_i p_{ij} = 1 \cdot p_{kj} + \sum_{i: i \neq k} \mu_i p_{ij} = p_{kj} + \sum_{i \neq k} \sum_l \mu_l p_{li} p_{ij} \\ &= p_{kj} + \sum_{i \neq k} p_{ki} p_{ij} + \sum_{i \neq k} \sum_{l \neq k} \mu_l p_{li} p_{ij} = \dots \\ &= p_{kj} + \sum_{i \neq k} p_{ki} p_{ij} + \dots + \sum_{i_1, \dots, i_{n-1} \neq k} p_{ki_1} \dots p_{i_{n-1}j} \\ &\quad + \sum_l \sum_{i_1, \dots, i_{n-1} \neq k} \mu_l p_{li_1} \dots p_{i_{n-1}j}. \end{aligned}$$

Now, the non-negativity implies that the last expression is

$$\begin{aligned} &\geq \mathbb{P}_k(X_1 = j, T_k > 1) + \mathbb{P}_k(X_2 = j, T_k > 2) + \dots \\ &+ \mathbb{P}_k(X_n = j, T_k > n), \end{aligned}$$

which tends to γ_j^k as $n \rightarrow \infty$.

(b) Now let P be irreducible and recurrent. Then $\underline{\gamma}^k$ is invariant: $\underline{\gamma}^k P = \underline{\gamma}^k$. Then $\tilde{\mu} = \mu - \underline{\gamma}^k$ is also invariant: $\tilde{\mu} = \tilde{\mu}P$, and, owing to (a), non-negative: $\tilde{\mu}_i \geq 0 \forall i \in I$. But, for $i = k$, $\tilde{\mu}_k = \mu_k - \gamma_k^k = 1 - 1 = 0$.

Next, given $i \in I$, $\exists n \geq 1$ with $p_{ik}^{(n)} > 0$. Then, as

$$0 = \tilde{\mu}_k = \sum_l \tilde{\mu}_l p_{lk}^{(n)} \geq \tilde{\mu}_i p_{ik}^{(n)},$$

we obtain that $\tilde{\mu}_i = 0$. Hence, $\tilde{\mu} = 0$ and $\mu = \underline{\gamma}^k$. The proof of Theorem 6.2 is complete. QED

We see that for an irreducible recurrent chain, everything is fixed by the condition $\mu_k = 1$. More precisely, if μ is a non-zero IM, i.e. $\mu P = \mu$, $\mu_i \geq 0$ and $\mu_k > 0$ for some state k , then

$$\mu = \mu_k \underline{\gamma}^k.$$

This implies that **all** non-zero IMs are proportional: $\mu' = c\mu$. Next, every non-zero IM has all entries finite and strictly positive. In particular, all vectors $\underline{\gamma}^k$ are proportional:

$$\gamma_i^k \underline{\gamma}^i = \underline{\gamma}^k, \quad i, k \in I. \quad (6.15)$$

Now, for an irreducible recurrent chain we have two cases:

(i) all non-zero IMs μ have

$$\sum_{j \in I} \mu_j < \infty, \quad (6.16)$$

and (ii) all non-zero IMs μ have

$$\sum_{j \in I} \mu_j = \infty. \quad (6.17)$$

Definition 6.2 In case (i) we call the irreducible Markov chain (or matrix P) *positive recurrent*, and in case (ii) *null recurrent*.

If the number of states $\#I < \infty$ then case (ii) is impossible. Hence, an irreducible finite Markov chain is always positive recurrent and has a (unique) equilibrium distribution $\pi = (\pi_i)$. Furthermore, equilibrium probabilities π_i are strictly positive.

We now see that in general, when P is positive recurrent then normalising $\mu_j / \sum_i \mu_i = \pi_j$ yields a (unique) equilibrium distribution. It has all $\pi_i > 0$. Then vector $\underline{\gamma}^k$ is recovered by division:

$$\underline{\gamma}^k = \frac{1}{\pi_k} \pi, \text{ i.e. } \gamma_i^k = \frac{\pi_i}{\pi_k}. \quad (6.18)$$

In other words, we obtain the following

Theorem 6.3 *In an irreducible positive recurrent chain with equilibrium distribution π , \forall states $k \neq i$*

$$\mathbb{E}_k(\# \text{ of visits to } i \text{ before returning to } k) = \frac{\pi_i}{\pi_k}. \quad (6.19)$$

For $i = k$ we obtain

Theorem 6.4 *In an irreducible positive recurrent chain with equilibrium distribution π , \forall state k*

$$m_k := \mathbb{E}_k T_k = \text{the mean return time to state } k = \frac{1}{\pi_k} < \infty. \quad (6.20)$$

Proof In Eqn (6.9) we observed that

$$\mathbb{E}_k T_k = 1 + \mathbb{E}_k (T_k - 1) = 1 + \sum_{i: i \neq k} \gamma_i^k = \sum_i \gamma_i^k < \infty.$$

Hence,

$$m_k = \sum_i \frac{\pi_i}{\pi_k} = \frac{1}{\pi_k},$$

implying that $m_k = 1/\pi_k$. \square

Concluding this section, let me summarise our findings on recurrence and transience so far.

(I) Irreducible Markov chains with more than one state have transition probabilities $p_{ij}^{(n)} > 0 \forall$ states $i, j \in I$ where $n \geq 1$ depends on the pair i, j . Also, $p_{ii} < 1 \forall i \in I$ (no absorption).

(II) An irreducible Markov chain (X_n) can be transient or recurrent:

(i) Transient:

$$\mathbb{P}_i(\text{return time } T_i < \infty) < 1, \text{ i.e., } \mathbb{P}_i(T_i = \infty) > 0, \forall i \in I.$$

Equivalently:

$$\mathbb{P}_i(i \text{ not visited after some finite time}) = 1 \text{ and } \sum_{n \geq 0} p_{ii}^{(n)} < \infty, \forall i \in I.$$

Equivalently:

$$h_j^{\{i\}} = \mathbb{P}_j(\text{hit } i) < 1, \text{ for some states } j \text{ and } i.$$

(ii) Recurrent:

$$\mathbb{P}_i(\text{return time } T_i < \infty) = 1, \text{ i.e., } \mathbb{P}_i(T_i = \infty) = 0, \forall i \in I.$$

Equivalently:

$$\mathbb{P}_i(i \text{ visited at arbitrarily large times}) = 1 \text{ and } \sum_{n \geq 0} p_{ii}^{(n)} = \infty, \forall i \in I.$$

Equivalently,

$$h_j^{\{i\}} = \mathbb{P}_j(\text{hit } i) = 1, \forall \text{ states } j \text{ and } i.$$

In this case, $\forall i$, the vector $\gamma^i = (\gamma_j^i)$ from (6.8) has $0 < \gamma_j^i < \infty$ and gives an IM for (X_n) ; all such IMs are of the form $\alpha \underline{\gamma}^i$. In particular,

$$\text{vector } \underline{\gamma}^k = (\gamma_k^i)^{-1} \times \text{vector } \underline{\gamma}^i, \forall \text{ states } i, k.$$

(III) Next, an irreducible recurrent DTMC can be

(i) Null Recurrent (NR):

$$m_i = \mathbb{E}_i(\text{return time } T_i) = \infty, \quad \forall i \in I;$$

in this case there is no IM $\mu = (\mu_i)$ with $\sum_j \mu_j < \infty$. Hence, there is no ED.

(ii) Positive Recurrent (PR): $m_i < \infty, \forall i \in I$; in this case any invariant measure $\mu = (\mu_i)$ has $\sum_j \mu_j < \infty$, and \exists a unique equilibrium

distribution $\pi = (\pi_i)$, where $\pi_i = \mu_i / \sum_j \mu_j > 0$. In this case vector $\underline{\gamma}^k = m_k \pi$. Furthermore,

$$\mathbb{E}_i T_i = \frac{1}{\pi_i}, \quad \text{and} \quad \mathbb{E}_i(\text{time at } k \text{ before } T_i) = \frac{\pi_k}{\pi_i}, \quad \forall \text{ states } i, k.$$

Finite irreducible Markov chains are always PR.

In a general case where the chain is reducible, all above considerations are applicable to each communicating class. More precisely, we use

Definition 6.3 Set $f_i = \mathbb{P}_i(T_i < \infty)$ and $m_i = \mathbb{E}_i T_i$. A state i is called

$$\begin{aligned} \text{recurrent (R), if } \quad & f_i = 1; \Leftrightarrow \sum_n p_{ii}^{(n)} = \infty, \text{ or} \\ & \mathbb{P}_i(X_n = i \text{ for infinitely many } n) = 1, \end{aligned}$$

$$\text{positive recurrent (PR), if } \quad m_i = \mathbb{E}_i T_i < \infty, \tag{6.21}$$

$$\text{null recurrent (NR), if } \quad m_i = \mathbb{E}_i T_i = \infty, \text{ but } f_i = 1,$$

$$\begin{aligned} \text{transient (T), if } \quad & f_i < 1; \Leftrightarrow \sum_n p_{ii}^{(n)} < \infty, \text{ or} \\ & \mathbb{P}_i(X_n = i \text{ for infinitely many } n) = 0. \end{aligned}$$

As these are class properties: the states from a given CC are either all PR

or all NR or all T.

Definition 6.4 Recall, the difference

$$S_i^{(l)} = T_i^{(l)} - T_i^{(l-1)}$$

between the subsequent return (hitting) times $T_i^{(l)}$ and $T_i^{(l-1)}$ gives the *duration* of the l th *excursion* to state i , $l = 1, 2, \dots$. For $l = 1$: $S_i^{(1)} = T_i$, the first return time.

The above analysis of positive and null recurrence combined with the strong Markov property leads to the following

Theorem 6.5. Assume that chain (X_n) is R and let i be any state. Under the distribution \mathbb{P}_i , variables $S_i^{(1)}, S_i^{(2)}, \dots$ are independent and identically distributed random variables, with integer values $s \geq 1$, finite with probability 1. That is, $\forall k \geq 1, \sum_{s \geq 1} \mathbb{P}_i(S_i = s) = 1$, and \forall integers $s_1, \dots, s_k \geq 1$,

$$\mathbb{P}_i(S_i^{(l)} = s_l, 1 \leq l \leq k) = \prod_{1 \leq l \leq k} \mathbb{P}_i(T_i = s_l). \quad (6.22)$$

Furthermore, the expectation

$$m_i := \mathbb{E}_i S_i^{(l)} = \begin{cases} 1/\pi_i, & \text{if chain } (X_n) \text{ is PR,} \\ +\infty, & \text{if chain } (X_n) \text{ is NR or T.} \end{cases} \quad (6.23)$$

Here $\pi = (\pi_i)$ is the (unique) ED distribution of the PR (X_n) .

The example of independent identically distributed (IID) random variables (RVs) $S_i^{(1)}, S_i^{(2)}, \dots$ is quite intriguing, as their (common) distribution is determined by transition matrix P and varies in a rather intricate way when we change matrix P . Therefore, to analyse sequence $(S_i^{(l)})$, one needs to

develop a general theory of IID RVs (in particular, it was one of the strong motives behind a general theory of summation of IID RVs).

An example of a general statement about IID RVs which we will use in the next section is the following ‘strong’ Law of Large Numbers (LLN) for sequence $(T_i^{(n)})$:

Theorem 6.6 (Non-examinable) *Under assumptions of Theorem 6.5, \forall state i , with probability 1 the average*

$$\frac{1}{n} \left(S_i^{(1)} + S_i^{(2)} + \cdots + T_i^{(n)} \right)$$

converges, as $n \rightarrow \infty$, to the expected value m_i specified in (6.23); symbolically $\left(T_i^{(1)} + T_i^{(2)} + \cdots + T_i^{(n)} \right) / n \xrightarrow{\mathbb{P}_i\text{-a.s.}} m_i$. That is,

$$\mathbb{P}_i \left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=1}^n T_i^{(l)} = m_i \right) = 1. \quad (6.24)$$

Remark 6.1 Note the sentence “with probability 1” and the notation $\xrightarrow{\mathbb{P}_i\text{-a.s.}}$ originates from the term ‘almost sure convergence with respect to probability distribution \mathbb{P}_i ’. When the probability distribution in question is specified from the context, we write $\xrightarrow{\text{a.s.}}$. The property of convergence with probability 1 will be studied in more detail in further probabilistic courses.

Remark 6.2 We want to stress that the statement of Theorem 6.6 holds in ‘full generality’, regardless of whether the value m_i is finite or infinite, let alone existence of a finite *second moment* $\mathbb{E}(T_i)^2$ or finite *higher moments* $\mathbb{E}(T_i)^n$, $n \geq 3$. In fact, the assertion of Theorem 6.6 holds in a much wider context of *ergodic processes*.

Example 6.2 Random walks on \mathbb{Z}^d

a) *Symmetric nearest-neighbour random walk*. We know that the symmetric nearest-neighbour RW on \mathbb{Z}^d (also called the *simple* RW) is recurrent for

$d = 1$ and $d = 2$ and transient for $d = 3$. First, consider $d = 1$. The invariance equation $\mu P = \mu$ reads

$$\mu_i = \frac{1}{2} \mu_{i-1} + \frac{1}{2} \mu_{i+1}, \quad i \in \mathbb{Z};$$

a general solution is

$$\mu_i = A + Bi, \quad i = 0, \pm 1, \pm 2, \dots$$

We have an obvious non-negative solution $\mu_i \equiv A \geq 0$ which is unique, up to a positive factor. As $\sum_{i \in \mathbb{Z}} \mu_i$ diverges (unless $A = 0$ when $\mu_i \equiv 0$), the walk is NR. If $\mu_k = 1$ for some k then $\mu_i \equiv 1 \forall i$.

Then, $\forall i \neq k$

$$\gamma_i^k = \mathbb{E}_k(\text{number of visits to } i \text{ before returning to } k) = 1.$$

[You may find it surprising as it could be suspected that

$$1 < \gamma_{k+1}^k < \gamma_{k+2}^k < \dots .]$$

More precisely,

$$\begin{aligned} & \mathbb{P}_k(\text{number of visits to } i \text{ before returning to } k \text{ is } n) \\ &= \left(\frac{1}{2|k-i|} \right)^2 \left(1 - \frac{1}{2|k-i|} \right)^{n-1}, \end{aligned}$$

see Example 6.3 below. Also

$$m_k = \mathbb{E}_k(\text{return time to } k) = \infty, \quad k \in \mathbb{Z}.$$

However, the probability $f_k = \mathbb{P}_k(T_k < +\infty) = 1$.

For $d = 2$, the invariance equations are similar

$$\mu_{\underline{i}} = \frac{1}{4} \sum (\mu_{(i_1 \pm 1, i_2)} + \mu_{(i_1, i_2 \pm 1)}), \quad \underline{i} = (i_1, i_2) \in \mathbb{Z}^2,$$

and again have $\mu_{\underline{i}} \equiv A \geq 0$ as a non-negative solution. Hence, the walk is NR, and as before, $f_{\underline{k}} \equiv 1$, $m_{\underline{k}} \equiv +\infty$ and

$$\gamma_{\underline{i}}^{\underline{k}} \equiv 1.$$

For $d = 3$, $\mu_{\underline{i}} \equiv 1$ is still an IM (this remains true for all d). However, as the walk is transient, vectors $\underline{\gamma}^k$ are sub-invariant, not invariant. Hence, it is no longer true that $\gamma_{\underline{i}}^k \equiv 1$, although $m_{\underline{k}}$ is still $\equiv +\infty$. Also, return probability $f_{\underline{k}} \equiv f_{\underline{0}} < 1$.

b) *Asymmetric homogeneous nearest-neighbour random walk on \mathbb{Z}* . Here the invariance equations are

$$\mu_i = p\mu_{i-1} + (1-p)\mu_{i+1}, \quad i \in \mathbb{Z},$$

and $p \neq 1/2$. The RW is transient. A general non-negative solution

$$\mu_i = A + B \left(\frac{p}{1-p} \right)^i, \quad i = 0, \pm 1, \pm 2, \dots,$$

contains two parameters, $A, B \geq 0$, and violates $\sum_i \pi_i < \infty$. We see that not all IMs λ are proportional. Again, the γ^k are sub-invariant, not invariant. Also, it is not true that γ_i^k is of the form λ_i/λ_k for some IM λ . But still $f_k \equiv f_0 < 1$ and $m_k \equiv \infty$, as

$$1 - f_k = \mathbb{P}_k(\text{no return to } k \text{ in a finite time}) > 0.$$

Example 6.3 (Homogeneous birth-and-death process) This is a RW on state space $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$, with

$$p_{ii+1} = p, \quad p_{ii-1} = 1-p, \quad i \geq 1, \quad p_{01} = r, \quad p_{00} = 1-r,$$

where $0 \leq p, r \leq 1$. Consider the case $0 < r \leq 1$ and $0 < p < 1$, when the chain is irreducible. Then the answer is

$$\begin{aligned} 0 < p < 1/2 &: \text{ PR,} \\ p = 1/2 &: \text{ NR,} \\ 1/2 < p < 1 &: \text{ T,} \end{aligned}$$

regardless of $r \in (0, 1]$.

In fact, the invariance equation $\mu = \mu P$ gives:

$$\begin{aligned} \mu_i &= p\mu_{i-1} + (1-p)\mu_{i+1}, \quad i > 1, \\ \mu_1 &= r\mu_0 + (1-p)\mu_2, \\ \mu_0 &= (1-r)\mu_0 + (1-p)\mu_1, \end{aligned}$$

and still admits the solution $\mu_i = A + B(p/(1-p))^i$, $i \geq 1$.

For $p < 1/2$, a further reduction seems reasonable: $A = 0$. At $i = 0, 1$ we obtain the same equation

$$r\mu_0 = pB.$$

To normalise $\sum_i \mu_i = 1$, write

$$\begin{aligned} 1 &= B \left(\frac{p}{r} + \frac{p}{1-p} + \frac{p^2}{(1-p)^2} + \dots \right) = B \left(\frac{p}{r} + \frac{p/(1-p)}{1-p/(1-p)} \right) \\ &= B \frac{p(1-2p+r)}{r(1-2p)}, \text{ whence } B = \frac{r(1-2p)}{p(1+r-2p)}. \end{aligned}$$

Therefore, the ED probabilities

$$\pi_0 = \frac{1-2p}{1+r-2p}, \quad \pi_i = \frac{r}{p} \frac{1-2p}{1+r-2p} \left(\frac{p}{1-p} \right)^i, \quad i \geq 1,$$

and the chain is positive recurrent, as claimed.

Further, for $p < 1/2$

$$\gamma_i^k = \mathbb{E}_k(\text{number of visits to } i \text{ before returning to } k) = \frac{\pi_i}{\pi_k}$$

$$= \begin{cases} \left(\frac{p}{1-p} \right)^{i-k}, & 0 < i, k < \infty, i \neq k, \\ \frac{r}{p} \left(\frac{p}{1-p} \right)^i, & 0 = k < i < \infty, \\ \frac{p}{r} \left(\frac{1-p}{p} \right)^k, & 0 = i < k < \infty, \end{cases}$$

and

$$m_k = \mathbb{E}_k(\text{return time to } k) = \frac{1}{\pi_k}, \quad k \in \mathbb{Z}_+;$$

more precisely:

$$m_k = \begin{cases} (1+r-2p)/(1-2p), & k = 0, \\ [(1-p)^k p(1+r-2p)]/[r(1-2p)p^k], & k \geq 1. \end{cases}$$

For $p > 1/2$ we know that the chain is transient: $f_k < 1$ and $m_k \equiv +\infty$. The IMs $\mu_i = A + B(p/(1-p))^i$, $i \geq 1$, contain two parameters A, B and are not all proportional.

It remains to check the case $p = 1/2$: we know that in this case $f_i = 1$ and the chain is recurrent. The invariance equations

$$\mu_i = \frac{1}{2} \mu_{i-1} + \frac{1}{2} \mu_{i+1}, \quad i > 1,$$

again have the general solution $\mu_i = A + Bi$, $i \geq 1$. At $i = 1, 0$ they have the form

$$\mu_1 = r\mu_0 + \frac{1}{2} \mu_2, \quad \mu_0 = (1-r)\mu_0 + \frac{1}{2} \mu_1,$$

which yields $B = 0$ and

$$\mu_i \equiv A, \quad i \geq 1, \quad \mu_0 = \frac{1}{2r} A,$$

and the non-negative IMs correspond to $A \geq 0$. We see that inequality $\sum_i \mu_i < \infty$ is impossible unless $A = 0$. Thus, the chain does not have an equilibrium distribution, and hence is NR.

Then, for $p = 1/2$, (i) all non-negative IMs $\mu = (\mu_i)$ are proportional to each other, and each such measure different from the identical 0 has $\mu_i = A > 0$ for $i \geq 1$ and $\mu_0 = A/(2r) > 0$. Furthermore, (ii) all vectors $\underline{\gamma}^k$, $k \geq 0$, must be invariant and hence proportional to each other. With the normalisation $\gamma_k^k = 1$, the only possibility is that (i) $\gamma_i^k \equiv 1$ and $\gamma_0^k = 1/(2r) \forall k, i \geq 1$ and (ii) $\gamma_i^0 = 2r \forall i \geq 1$. (Again it seems unexpected as one might expect that for $k \geq 1$

$$\gamma_0^k < \dots < \gamma_{k-2}^k < \gamma_{k-1}^k < 1 < \gamma_{k+1}^k < \gamma_{k+2}^k < \dots,$$

and there is no reason to believe that $\gamma_{k-1}^k = \gamma_{k+1}^k$ because of the asymmetry of the model.)

Finally, it is not difficult to check that $\forall i > k \geq 1$

$$\begin{aligned} & \mathbb{P}_k(\text{number of visits to } i \text{ before returning to } k \text{ is } n) \\ &= \left(\frac{1}{2(i-k)} \right)^2 \left(1 - \frac{1}{2(i-k)} \right)^{n-1}, \end{aligned}$$

as in the case of the symmetric RW on \mathbb{Z} .

Example 6.4. (Math Tripos, Markov Chains, Part IIA, 1997, A301J)

(i) Let $X = (X_n : n \geq 0)$ be a random walk on the integers, which moves one step rightwards and one step leftwards with probability $\frac{1}{2}$, at each time point. Show that

$$\mathbb{P}(X_{2n} = 0 | X_0 = 0) = \binom{2n}{n} \left(\frac{1}{2}\right)^{2n},$$

and deduce that X is recurrent.

(ii) Let X be given as above, and assume that $X_0 = 0$. Let m be a strictly positive integer, and let N be the number of visits to the point m before returning to 0.

Find $\mathbb{P}(N \geq 1)$, and deduce that

$$\mathbb{P}(N = n) = \left(\frac{1}{2m}\right)^2 \left(1 - \frac{1}{2m}\right)^{n-1}, \quad n \geq 1.$$

Solution (i) $\mathbb{P}(X_{2n} = 0 | X_0 = 0) = p_{00}^{(2n)}$ is the probability of all paths of length $2n$ starting at and returning to 0. Each such path must have n steps right and n steps left. The total number of such paths is $\binom{2n}{n}$ and each of them has the same probability $(1/2)^{2n}$. Hence the formula for $\mathbb{P}(X_{2n} = 0 | X_0 = 0)$.

The sum $\sum_{n=1}^{\infty} \mathbb{P}(X_n = 0 | X_0 = 0)$ coincides with $\sum_{n=1}^{\infty} \mathbb{P}(X_{2n} = 0 | X_0 = 0)$ (return at odd times is not possible), and is analysed via Stirling's formula: $n! \approx \sqrt{2\pi n} n^{n+1/2} e^{-n}$. This leads to the series $\sum_n 1/(\sqrt{\pi n})$ which diverges. So, by Theorem 5.2, state 0 is recurrent. The same argument works for every state i . Hence, the chain is recurrent. (The same conclusion holds because recurrence is a class property; see Theorem 5.3.)

(ii) Symbol \mathbb{P} here means \mathbb{P}_0 , the distribution of the (δ_0, P) chain. Then $\mathbb{P}(N \geq 1) = \mathbb{P}_0(\text{hit } m \text{ before returning to } 0)$. By conditioning on the first

step, write:

$$\mathbb{P}(N \geq 1) = \frac{1}{2} \mathbb{P}_1(\text{hit } m \text{ before visiting } 0),$$

where \mathbb{P}_i stands for the distribution of the (δ_i, P) chain. Set

$$h_i = \mathbb{P}_i(\text{hit } m \text{ before visiting } 0),$$

then

$$h_i = \frac{1}{2} h_{i-1} + \frac{1}{2} h_{i+1}, \quad 1 \leq i < m.$$

The general solution $h_i = A + Bi$ is specified by $h_0 = 0$, $h_m = 1$: $A = 0$, $B = 1/m$. Hence, $h_1 = 1/m$, and $\mathbb{P}(N \geq 1) = 1/(2m)$.

Clearly, $1 - 1/(2m) = \mathbb{P}(N = 0) = \mathbb{P}_0(\text{hit } 0 \text{ again before visiting } m)$. By symmetry,

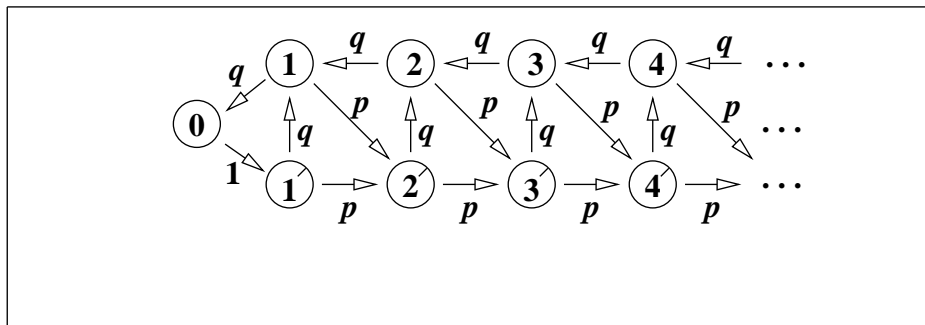
$$\mathbb{P}_m(\text{hit } m \text{ again before visiting } 0) = 1 - \frac{1}{2m}.$$

To be in event $\{N = n\}$, a path from 0 must hit m before returning to 0, return to m $n - 1$ times without visiting 0 and then proceed to 0 without return to m . By the strong Markov property,

$$\mathbb{P}(N = n) = \frac{1}{2m} \left(1 - \frac{1}{2m}\right)^{n-1} \frac{1}{2m};$$

the last factor being $\mathbb{P}_m(\text{hit } 0 \text{ before returning to } m)$, again by symmetry. Hence the result.

Example 6.5. (Math Tripos, Markov Chains, Part IB, 2004, 222H)
 Consider a Markov chain on the state space $I = \{0, 1, 2, \dots\} \cup \{1', 2', 3', \dots\}$ with transition probabilities as illustrated in Fig. 8.2 where $0 < q < 1$ and $p = 1 - q$.



For each value of q , determine whether the chain is transient, null recurrent or positive recurrent.

When the chain is positive recurrent, calculate the invariant distribution.

Solution For $i \geq 1$ set

$$a = \mathbb{P}_i(\text{hit } i - 1), \quad b = \mathbb{P}_{i'}(\text{hit } i).$$

(these probabilities do not depend on the value of i because of the homogeneous property of the chain). Conditioning on the first jump and using the strong Markov property

$$a = q + pba^2, \quad b = q + pba,$$

whence

$$b = \frac{q}{1 - pa}, \quad \text{and} \quad a = q + \frac{pqa^2}{1 - pa}.$$

Thus,

$$p(1 + q)a^2 - (pq + 1)a + q = 0,$$

and the solutions are

$$a = 1 \quad \text{and} \quad a = \frac{q}{1 - q^2}.$$

We are interested in the minimal solution

$$\frac{q}{1 - q^2} < 1 \quad \text{if and only if} \quad q < \frac{\sqrt{5} - 1}{2}.$$

Therefore, the chain is recurrent if and only if $q \geq (\sqrt{5} - 1) / 2$ and transient if and only if $q < (\sqrt{5} - 1) / 2$.

To decide whether it is NR or PR, consider the invariance equation $\pi = \pi P$

$$\begin{aligned}\pi_0 &= \pi_1 q, & \pi_i &= \pi_{i+1} q + \pi_{i'} q, & i &\geq 1, \\ \pi_{1'} &= \pi_0, & \pi_{i'} &= \pi_{(i-1)'} p + \pi_{i-1} p, & i' &\geq 2.\end{aligned}$$

It admits a recursive solution:

$$\pi_1 = \frac{1}{q} \pi_0, \quad \pi_{1'} = \pi_0,$$

$$\pi_2 = \left(\frac{1}{q^2} - 1 \right) \pi_0 = \frac{1}{q} \frac{1 - q^2}{q} \pi_0, \quad \pi_{2'} = (1 - q) \left(1 + \frac{1}{q} \right) \pi_0 = \frac{1 - q^2}{q} \pi_0,$$

and similarly,

$$\pi_3 = \frac{1}{q} \left(\frac{1 - q^2}{q} \right)^2 \pi_0, \quad \pi_{3'} = \left(\frac{1 - q^2}{q} \right)^2 \pi_0.$$

By induction, one gets the general formulas

$$\pi_i = \frac{1}{q} \left(\frac{1 - q^2}{q} \right)^{i-1} \pi_0, \quad \pi_{i'} = \left(\frac{1 - q^2}{q} \right)^{i-1} \pi_0,$$

and the equilibrium distribution will exist if and only if both series converge, that is, $(1 - q^2)/q < 1$, i.e. $q > (\sqrt{5} - 1) / 2$. Hence, the chain is NR when $q = (\sqrt{5} - 1) / 2$ and PR when $q > (\sqrt{5} - 1) / 2$. In the latter case

$$\pi_0 = \left[1 + \sum_{i \geq 1} \left(\frac{1}{q} + 1 \right) \left(\frac{1 - q^2}{q} \right)^{i-1} \right]^{-1} = \frac{q^2 + q - 1}{q^2 + 2q}.$$

Example 6.6. (Math Tripos, Markov Chains, Part IIA, 2004, A101I and Part IIB, 2004, B101I) **(i)** Give the definitions of a *recurrent* and a *null recurrent* irreducible Markov chain.

Let (X_n) be a recurrent Markov chain with state space I and irreducible transition matrix $P = (p_{ij})$. Prove that the vectors $\underline{\gamma}^k = (\gamma_j^k, j \in I)$, $k \in I$, with entries $\gamma_k^k = 1$ and

$$\gamma_i^k = \mathbb{E}_k(\text{number of visits to } i \text{ before returning to } k), \quad i \neq k,$$

are P -invariant

$$\gamma_j^k = \sum_i \gamma_i^k p_{ij}.$$

(ii) Let (W_n) be the birth-and-death process on $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ with the following transition probabilities

$$\begin{aligned} p_{i,i+1} &= p_{i,i-1} = \frac{1}{2}, \quad i \geq 1 \\ p_{01} &= 1. \end{aligned}$$

By relating (W_n) to the symmetric simple random walk (Y_n) on \mathbb{Z} , or otherwise, prove that (W_n) is a recurrent Markov chain. By considering invariant measures, or otherwise, prove that (W_n) is null recurrent.

Calculate the vectors $\underline{\gamma}^k = (\gamma_i^k, i \in \mathbb{Z}_+)$ for the chain (W_n) , $k \in \mathbb{Z}_+$.

Finally, let $W_0 = 0$ and let N be the number of visits to 1 before returning to 0. Show that $\mathbb{P}_0(N = n) = (1/2)^n$, $n \geq 1$.

[You may use properties of the random walk (Y_n) or general facts about Markov chains without proof but should clearly state them.]

Solution (i) The definition of an irreducible R Markov chain: all states communicate and

$$f_i := \mathbb{P}_i(\text{return to } i) = 1 \quad \forall \text{ state } i. \quad (*)$$

The definition of an irreducible NR Markov chain: all states communicate, Eqn (*) holds, and

$$m_i := \mathbb{E}_i(\text{return time to } i) = \infty \quad \forall \text{ state } i,$$

which is equivalent to the fact that the chain has no invariant distribution.

Invariance of vectors $\underline{\gamma}^k = (\gamma_i^k)$ is checked as follows. For $j \neq k$,

$$\begin{aligned} \gamma_j^k &= \mathbb{E}_k \left(\sum_{n \geq 1} \mathbf{1}(X_n = j \text{ and } X_l \neq k \text{ for } l = 1, \dots, n-1) \right) \\ &= \sum_{n \geq 1} \mathbb{P}_k(X_n = j \text{ and } X_l \neq k \text{ for } l = 1, \dots, n-1) \end{aligned}$$

$$\begin{aligned}
&= p_{kj} + \sum_{n \geq 2} \sum_{i \neq k} \mathbb{P}_k(X_n = j, X_{n-1} = i \text{ and } X_l \neq k \text{ for } l = 1, \dots, n-2) \\
&= \gamma_k^k p_{kj} + \sum_{i \neq k} \sum_{n \geq 1} \mathbb{P}_k(X_l \neq k \text{ for } l = 1, \dots, n-1 \text{ and } X_n = i) p_{ij} \\
&= \gamma_k^k p_{kj} + \sum_{i \neq k} \gamma_i^k p_{ij} = (\gamma^k P)_j.
\end{aligned}$$

For $j = k$, a similar calculation yields

$$(\gamma^k P)_k = \sum_{n \geq 1} \mathbb{P}_k(X_l = k \text{ for } l = n \text{ but not for } l = 1, \dots, n-1) = f_j$$

which equals 1 when j is an R state.

This gives the invariance equation $\gamma^k P = \gamma^k$ for an R chain.

(ii) Now, (W_n) is an irreducible Markov chain. Also, $W_n = |Y_n|$ where (Y_n) is the nearest-neighbour symmetric random walk on \mathbb{Z} . Hence, $\forall i \in \mathbb{Z}$

$$\mathbb{P}_{|i|}((W_n) \text{ returns to } i) \geq \mathbb{P}_i((Y_n) \text{ returns to } i),$$

but the RHS equals 1 as (Y_n) is R. Hence, the LHS equals 1, and (W_n) is R.

To check NR, it suffices to prove that (W_n) has no equilibrium distribution. Consider the invariance equations

$$\begin{aligned}
\mu_0 &= \frac{1}{2} \mu_1, \quad \mu_1 = \mu_0 + \frac{1}{2} \mu_2, \\
\mu_i &= \frac{1}{2} \mu_{i-1} + \frac{1}{2} \mu_{i+1}, \quad i \geq 2.
\end{aligned}$$

The second line has a general solution $\mu_i = A + Bi$, $i \geq 1$. From the first line, $B = 0$ and $\mu_0 = A/2$. Hence, any IM μ is of the form

$$\mu_i = A, \quad i \geq 1, \quad \mu_0 = \frac{1}{2} A,$$

where $A \geq 0$. It has $\sum_i \mu_i = \infty$ unless $A = 0$. Thus, no equilibrium distribution could exist, and (W_n) is NR.

Therefore, for chain (W_n)

$$\gamma_i^k = \frac{\mu_i}{\mu_k} = \begin{cases} 1, & i, k \geq 1 \text{ or } i = k = 0, \\ 1/2, & i = 0, k \geq 1, \\ 2, & i \geq 1, k = 0. \end{cases}$$

Next, by the strong Markov property

$$\begin{aligned}\mathbb{P}_0(N = n) &= \mathbb{P}_0(N \geq 1) \\ &\quad \times (\mathbb{P}_1(\text{return to 1 without visiting 0}))^{n-1} \\ &\quad \times \mathbb{P}_1(\text{hit 0 without returning to 1}) \\ &= \frac{1}{2^n}.\end{aligned}$$

In fact,

$$\mathbb{P}_0(N \geq 1) = 1 \quad (\text{as } p_{01} = 1),$$

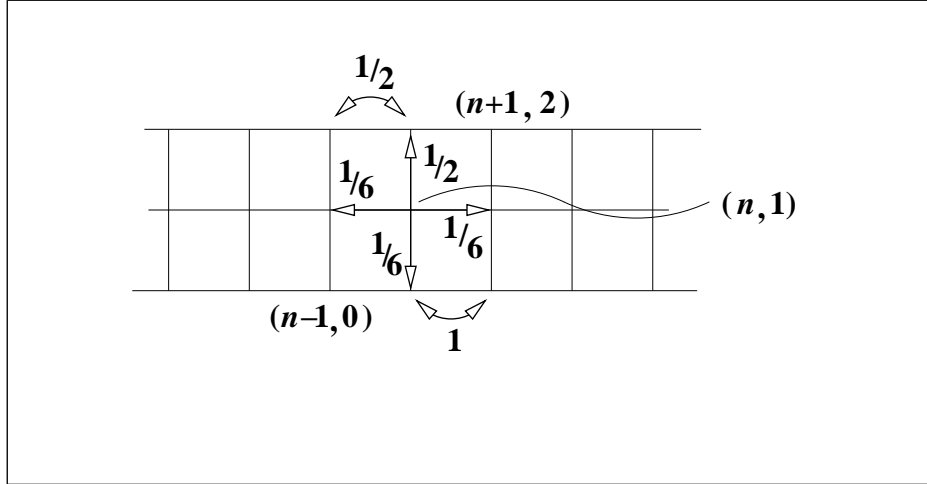
$$\mathbb{P}_1(\text{return to 1 without visiting 0}) = 1 - p_{10} = \frac{1}{2} \quad (\text{as the chain hits 0 from 1 with probability } 1/2 \text{ and is R}),$$

and

$$\mathbb{P}_1(\text{hit 0 without returning to 1}) = p_{10} = \frac{1}{2}.$$

Example 6.7. (Math Tripos, Markov Chains, Part IIA, 1992, 308B)
A snail crawls on an infinite fence, which may be taken to be a lattice with vertices at the points of $\mathbb{Z} \times \{0, 1, 2\}$. From a vertex of type $(n, 2)$, the snail crawls left or right (that is, to $(n - 1, 2)$ or $(n + 1, 2)$) with equal probability. From one of type $(n, 1)$, it crawls up to $(n, 2)$ with probability $1/2$, and in any of the other three directions with probability $1/6$. From $(n, 0)$, it necessarily moves to the left if n is even, and to the right if n is odd. Classify the states of the Markov chain corresponding to the sequence of vertices visited by the snail. If it starts at $(0, 1)$, what is the probability that it eventually reaches a positive recurrent state? What is the probability that it eventually visits $(0, 0)$?

Solution. The states $(n, 2)$, $n \in \mathbb{Z}$, form a closed communicating class and are all NR. For each n , the pair of states $(2n - 1, 0)$ and $(2n, 0)$ is a closed communicating class, hence all these states are PR. Finally, every state $(n, 1)$ is T: $p_{(n,1)(n,2)}, p_{(n,1)(n,0)} > 0$ but there is no return from level 2 or 0 to level 1.



Suppose the snail starts on level 1 and consider the probability

$$\mathbb{P}_{(i,1)}(\text{eventually reaches level 0}).$$

It equals

$$\begin{aligned} & \sum_{n=0}^{\infty} \mathbb{P}_{(i,1)}(\text{stays on level 1 for } n \text{ steps, then moves to level 0}) \\ &= \sum_{n=0}^{\infty} \left(\frac{1}{3}\right)^n \frac{1}{6} = \frac{1}{6} \left(1 - \frac{1}{3}\right)^{-1} = \frac{1}{4}. \end{aligned}$$

The snail will eventually visit $(0, 0)$ iff it eventually crawls down from either $(0, 1)$ or $(-1, 1)$. Denote

$$h_i = \mathbb{P}(\text{eventually visits } (0,0) \mid \text{currently at } (i, 1)).$$

Then $h_{-1} = h_0$ and

$$\begin{aligned} h_0 &= \frac{1}{6} + \frac{1}{6}h_0 + \frac{1}{6}h_1, \\ h_n &= \frac{1}{6}h_{n-1} + \frac{1}{6}h_{n+1}, \quad n \geq 1. \end{aligned}$$

Substituting $h_n = At^n$ in the second equation gives $t^2 - 6t + 1 = 0$, i.e., $t = 3 \pm 2\sqrt{2}$. As $h_n \leq 1$, we obtain that $h_n = A(3 - 2\sqrt{2})^n$. From the first equation: $A = h_0 = 1/2(1 + \sqrt{2}) = (\sqrt{2} - 1)/2$.