Exercise 1. A particle performs a random walk on the vertex set of a finite connected undirected graph $G$, which for simplicity we assume to have neither self-loops nor multiple edges. At each stage it moves to a neighbor of its current position, each such neighbor being chosen with equal probability. If $G$ has $\eta$ edges, show that the stationary distribution is given by $\pi_v = d_v/(2\eta)$, where $d_v$ is the degree of each vertex $v$.

**Solution:** One way to do this problem is to simply check that the proposed solution satisfies the defining equations: $\pi P = \pi$, and $\sum_v \pi_v = 1$ (we can see immediately that we have nonnegativity). We have:

$$\sum_v \pi_v = \sum_v \frac{d_v}{2\eta} = \frac{1}{2\eta} \sum_v d_v = 1,$$

since the sum of the degrees is twice the number of edges (each edge increases the sum of the degrees by exactly 2). Similarly, we can show that $\pi P = \pi$. Let us define $\delta_{vu}$ to be 1 if vertices $u$ and $v$ are adjacent, and 0 otherwise. Then, we have:

$$\sum_v \pi_v P_{vu} = \frac{1}{2\eta} \sum_v d_v \left(\frac{1}{d_v} \delta_{vu}\right) = \frac{1}{2\eta} \sum_v \delta_{vu}.$$ 

But $\sum_v \delta_{vu}$ is the number of edges incident to node $u$, that is, $\sum_v \delta_{vu} = d_u$. Therefore we have:

$$\sum_v \pi_v P_{vu} = \frac{1}{2\eta} d_u = \frac{d_u}{2\eta} = \pi_u.$$
Exercise 2. A particle performs a random walk on a bow tie $ABCDE$ drawn on Figure 1, where $C$ is the knot. From any vertex, its next step is equally likely to be to any neighbouring vertex. Initially it is at $A$. Find the expected value of:

(a) The time of first return to $A$.

(b) The number of visits to $D$ before returning to $A$.

(c) The number of visits to $C$ before returning to $A$.

(d) The time of first return to $A$, given that there were no visits to $E$ before the return to $A$.

(e) The number of visits to $D$ before returning to $A$, given that there were no visits to $E$ before the return to $A$.

Figure 1: A bow tie graph.

Solution: First, we can compute that the steady state distribution is $\pi_A = \pi_B = \pi_D = \pi_E = 1/6$, and $\pi_C = 1/3$. We can do this either by solving a system of linear equations (as usual) or just use our result from the first problem above.

(a) By the result from class, and on the handout, we have: $t_A = 1/\pi_A = 6$. Alternatively, we can solve the following system of equations (observe than $t_A$ appears in only one equation):

\[
\begin{align*}
    t_A &= \frac{1}{2}(t_B + 1) + \frac{1}{2}(t_C + 1) \\
    t_B &= \frac{1}{2} + \frac{1}{2}(t_C + 1) \\
    t_C &= \frac{1}{4} + \frac{1}{4}(t_B + 1) + \frac{1}{4}(t_D + 1) + \frac{1}{4}(t_E + 1) \\
    t_D &= \frac{1}{2}(t_C + 1) + \frac{1}{2}(t_E + 1) \\
    t_E &= \frac{1}{2}(t_C + 1) + \frac{1}{2}(t_D + 1).
\end{align*}
\]
(b) By the result from the handout on Markov Chains, we know that
\[ \pi_D = \frac{\mathbb{E}[\# \text{ transitions to } D \text{ in a cycle that starts and ends at } A]}{\mathbb{E}[\# \text{ transitions in a cycle that starts and ends at } A]}, \]
from which we find that the quantity we wish to compute is \( 6\pi_D = 1 \).

(c) Using the same method as in part (b), we find the answer to be \( 6\pi_C = 2 \).

(d) We let \( P_i(\cdot) = \mathbb{P}(\cdot|X_0 = i) \), and let \( T_j \) be the time of the first passage to state \( j \), and let \( \nu_i = \mathbb{P}_i(T_A < T_E) \). Then, as we obtained the equations above, that is, by conditioning on the first step, we have
\[
\begin{align*}
\nu_A &= \frac{1}{2}\nu_B + \frac{1}{2}\nu_C \\
\nu_B &= \frac{1}{2} + \frac{1}{2}\nu_C \\
\nu_C &= \frac{1}{4} + \frac{1}{4}\nu_B + \frac{1}{4}\nu_D \\
\nu_D &= \frac{1}{2}\nu_C.
\end{align*}
\]
Solving these, we find: \( \nu_A = 5/8, \nu_B = 3/4, \nu_C = 1/2, \nu_D = 1/4 \). Now we can compute the conditional transition probabilities, which we call \( \tau_{ij} \).

We have:
\[
\begin{align*}
\tau_{AB} &= \mathbb{P}_A(X_1 = B|T_A < T_E) \\
&= \frac{\mathbb{P}_A(X_1 = B)\mathbb{P}_B(T_A < T_E)}{\mathbb{P}_A(T_A < T_E)} \\
&= \frac{\nu_B}{2\nu_A} = \frac{3}{5}.
\end{align*}
\]
Similarly, we find: \( \tau_{AC} = 2/5, \tau_{BA} = 2/3, \tau_{BC} = 1/3, \tau_{CA} = 1/2, \tau_{CB} = 3/8, \tau_{CD} = 1/8, \tau_{DC} = 1 \). Now we have essentially reduced to a problem like part (a). We can compute the conditional expectation by solving a system of linear equations using the new transition probabilities:
\[
\begin{align*}
\tilde{i}_A &= 1 + \frac{3}{5}\tilde{i}_B + \frac{2}{5}\tilde{i}_C \\
\tilde{i}_B &= 1 + \frac{2}{3}(1) + \frac{1}{3}\tilde{i}_C \\
\tilde{i}_C &= 1 + \frac{1}{2}(1) + \frac{3}{8}\tilde{i}_B + \frac{1}{8}\tilde{i}_D \\
\tilde{i}_D &= 1 + \tilde{i}_C.
\end{align*}
\]
Solving these equations, yields \( \tilde{i}_A = 14/5 \).
(e) We can use the conditional transition probabilities above, to reduce to a problem essentially like that in part (b). Let \( N \) be the number of visits to \( D \). Then, denoting by \( \eta_i \) the expected value of \( N \) given that we start at \( i \), and that \( T_A < T_E \), we have the equations:

\[
\begin{align*}
\eta_A &= \frac{3}{5} \eta_B + \frac{2}{5} \eta_B \\
\eta_B &= 0 + \frac{1}{3} \eta_C \\
\eta_C &= 0 + \frac{3}{8} \eta_B + \frac{1}{8} (1 + \eta_D) \\
\eta_D &= \eta_C.
\end{align*}
\]

Solving, we obtain: \( \eta_A = 1/10 \).

**Exercise 3.** Let \((\Omega, F) = (\mathbb{R}^\infty, \mathcal{B}^\infty)\), \(X_k(\omega) = \omega_k, k \in \mathbb{N}\), be the canonical coordinate functions and \(\{F_k\}\) a filtration of \(F\). Recall that a filtration is a sequence of increasing \(\sigma\)-algebras \(F_1 \subset F_2 \subset \cdots\) contained in \(F, F_k \subset F\). We say that \(\tau\) is a stopping time of the filtration \(\{F_k\}\) if

(a) \(\tau\) is a positive integer

(b) for every \(k \geq 1\) we have \(\{\tau \leq k\} \in F_k\)

Let \(\tau : \Omega \to \mathbb{N}\) be \((F, \mathcal{B})\) measurable. Show that \(\tau\) is a stopping of \(\{F_k\}\) if and only if for every \(\omega, \omega_0 \in \Omega\) and for every \(n \geq 1\)

\[
\tau(\omega) = n, \ X_k(\omega) = X_k(\omega') \ \forall \ 1 \leq k \leq n \implies \tau(\omega') = n. \quad (1)
\]

**Solution:** A positive integer valued random variable \(\tau\) is a stopping time if and only if \(\{\tau = n\} \in F_n\) for all \(n\). The forward direction follows from \(\{\tau \leq n\} = \bigcup_{k=1}^n \tau = k\) and the reverse direction follows from \(\{\tau = n\} = \{\tau \leq n\} \setminus \{\tau \leq n - 1\}\). The relation \(\omega \overset{n}{\sim} \omega'\) if

\[
X_k(\omega) = X_k(\omega') \quad 1 \leq k \leq n
\]

is an equivalence relation, i.e. reflexive, symmetric, and transitive. For all \(E \subset \Omega\) define

\[
[E]_n = \{\omega \in \Omega | \exists \omega' \in E \text{ s.t. } \omega' \overset{n}{\sim} \omega\}.
\]

Condition 1 is equivalent to \([\{\tau = n\}]_n \subset \{\tau = n\}\). Therefore, it suffices to show that, for all \(n\), \(\{\tau = n\} \in F_n\) if and only if \([\{\tau = n\}]_n \subset \{\tau = n\}\).
Suppose \( \tau \) is a stopping time. Let

\[
D = \{ E \subset \Omega \mid [E]_n \subset E \}.
\]

By definition, \( D \) contains the empty set and sets of the form \( X_j^{-1}(B) \) for \( B \subset \mathbb{R} \) and \( 1 \leq j \leq n \). Moreover, let \( \{E_j\} \in D \), then

\[
\left[ \bigcup_{j=1}^{\infty} E_j \right]_n = \left[ \bigcup_{j=1}^{\infty} [E_j]_n \right] \subset \left[ \bigcap_{j=1}^{\infty} [E_j]_n \right],
\]

and therefore, \( D \) is a monotone class. Let

\[
C = \{ X_j^{-1}(B) \mid B \in \mathcal{B}, 1 \leq j \leq n \}.
\]

Then, the minimal algebra containing \( C \) is \( \alpha(C) \) is the set of finite unions of finite intersections of sets of the form \( X_j^{-1}(B) \) or \( X_j^{-1}(B)^c \). As the inverse image respects complements and \( D \) is closed under intersections and unions, \( D \) contains \( \alpha(C) \) and by the monotone class theorem \( D \supset \sigma(C) = \mathcal{F}_n \). Hence \( \{ \tau = n \} \in D \).

Conversely, suppose that condition 1 is satisfied. By definition, \( \{ \tau = n \} \supset \{ \tau = n \} \) and thusly \( \{ \tau = n \} \subset \{ \tau = n \} \). Therefore, \( \Omega \) decomposes as a union of equivalence classes \( \Omega = \bigcup_{\alpha \in I} U_\alpha \), for some indexing set \( I \) where \( [U_\alpha]_n = U_\alpha \) for all \( \alpha \) and \( U_\alpha \cap U_\beta = \emptyset \) for \( \alpha \neq \beta \). For each \( \alpha \in I \) choose a representative \( \omega_\alpha \in U_\alpha \). Let \( f : \Omega \to \Omega \) with \( f|_{U_\alpha} \equiv \omega_\alpha \). To show that \( f \) is measurable it suffices to check on a generating collection. Let \( S \subset \mathcal{N} \) be a finite set and \( B = \prod_{s \in S} B_s \) with \( B_s \in \mathcal{B}(\mathbb{R}) \), then \( f^{-1}(B) = \bigcap_{k=1}^{\infty} X_k^{-1}(X_k(B)) \in \mathcal{F}_n \) since \( X_k(B) \) is either \( B_k \) or \( \emptyset \) and \( X_k \) is measurable. Therefore, \( f \) is \( (\mathcal{F}_n, \mathcal{F}) \) measurable and, as \( \{ \tau = n \} \subset \{ \tau = n \} \) and \( \tau \) is \( (\mathcal{F}, \mathcal{B}) \) measurable, \( \{ \tau = n \} = f^{-1}(\{ \tau = n \}) \in \mathcal{F}_n \). Hence \( \tau \) is a stopping time.

**Exercise 4.** Let \( \tau \) be a stopping time of a filtration \( \mathcal{F}_n \). Recall that the \( \sigma \)-algebra \( \mathcal{F}_\tau \) of “past until \( \tau \)” is defined as

\[
\mathcal{F}_\tau = \{ E : E \cap \{ \tau \leq n \} \in \mathcal{F}_n \ \forall n \},
\]

Show that for every random variable \( V \) measurable with respect to \( \mathcal{F}_\tau \) there exists a stochastic process \( \{ G_n, n = 1, \ldots \} \), with \( G_n \) measurable with respect to \( \mathcal{F}_n \), such that

\[
V = G_\tau.
\]

(Hint: First consider simple \( V \)).
Solution: Let $V$ be a random variable measurable with respect to $\mathcal{F}_\tau$. Then $V$ decomposes as

$$V = V \mathbb{1}\{V > 0\} + V \mathbb{1}\{V = 0\} - (-V) \mathbb{1}\{V < 0\} = V_+ - V_-.$$  

Let $G_n = V \mathbb{1}\{\tau \leq n\}$. Then $G_\tau = V$ and

$$G_n = V_+ \mathbb{1}\{\tau \leq n\} - V_- \mathbb{1}\{\tau \leq n\}.$$

As random variables are closed under addition and scalar multiplication, it suffices to show that $G_n$ is measurable with respect to $\mathcal{F}_n$ for positive $V$. If $V > 0$ then $G_n \geq 0$. Let $x \geq 0$. Then

$$\{G_n > x\} = \{V \mathbb{1}\{\tau \leq n\} > x\} = \{V > x\} \cap \{\tau \leq n\} \in \mathcal{F}_n$$

since $V$ is measurable with respect to $\mathcal{F}_\tau$. As $\{(x, \infty)\}$ is a generating $p$-system for the Borel sigma algebra on the real numbers, $G_n$ is measurable with respect to $\mathcal{F}_n$.

Exercise 5. (Cover time of $C_n$) For a MC with state space $\mathcal{X}$ we define $\tau_{cov}$ to be the first time that every element of $\mathcal{X}$ was visited. The covering time $t_{cov} = \max_{x \in \mathcal{X}} E_x^{\tau_{cov}}$. Consider a MC that is a simple random walk on an $n$-cycle: it moves with probability $1/2$ to one of the neighbors each time. Show that $t_{cov}(n) = \frac{n(n-1)}{2}$ (Lovász’93). (Hint: Let $\tau_n$ be the first time a simple random walk on $\mathbb{Z}$ started at 0 visits $n$ distinct states. Relate to $t_{cov}$ and gambler’s ruin.)

Solution: Clearly, by symmetry, it does not matter what vertex we start from. Let us define $\sigma_k$ to be the first time that at least $k$ distinct vertices have been visited; obviously $\sigma_1 = 0$. We now note that $t_{cov} = E[\sigma_n]$; we can also telescope these like so:

$$\sigma_n = (\sigma_n - \sigma_{n-1}) + (\sigma_{n-1} - \sigma_{n-2}) + \cdots + (\sigma_2 - \sigma_1)$$

(note that we omit the “$\cdots + \sigma_1$” because it’s just 0). This of course means that $t_{cov} = \sum_{k=1}^{n-1} E[\sigma_{k+1} - \sigma_k]$ (by linearity).

Now let us examine what the situation is like at time $\sigma_k$ for $k < n$. We have $k$ visited vertices, which obviously are contiguous (and so form a path); furthermore, $X_{\sigma_k}$ must be at an endpoint of the path since by definition of $\sigma_k$, it must be the first visit we made to this vertex.

Now we ask: how long from then until $\sigma_{k+1}$? Well, we have a Gambler’s Ruin problem: exiting either end of the path of visited vertices gives us a new
one. To be precise, it’s a Gambler’s Ruin starting with 1 dollar and ending either with 0 dollars or \( k + 1 \) dollars; we know that the expected number of steps for this is \( j(k + 1 - j) \) where \( j = 1 \), which gives \( k \) steps. Therefore,

\[
\mathbb{E}[\sigma_{k+1} - \sigma_k] = k
\]

Plugging this in to the above, we get

\[
t_{\text{cov}} = \sum_{k=1}^{n-1} k = \frac{n(n - 1)}{2}
\]

**Exercise 6.** *(Last visited vertex of \( C_n \)) Consider a simple random walk \( X_t \) on an \( n \)-cycle \( C_n \) and let \( \tau_{\text{cov}} \) be the first time that every vertex was visited. Show that given that \( X_0 = v \) the distribution of \( X_{\tau_{\text{cov}}} \) is uniform on \( \{v\}^c \). *(Hint: Notice that to have \( X_{\tau_{\text{cov}}} = k \) the random walk should visit the states \( k - 1 \) and \( k + 1 \) before \( k \).)"

Fun fact: cycles and cliques are the only graphs with this property (Lovász-Winkler’93).

**Solution:** Fix a vertex \( x \); let \( \sigma_x \) be the first time that a neighbor of \( x \) is visited. For \( x \neq v \), obviously a neighbor of \( x \) must be visited before \( x \) is (keeping in mind that \( v \) itself could be this neighbor). Let \( u = X_{\sigma_u} \) (the first neighbor visited) and \( w \) be the other neighbor, which by definition has not been visited by time \( \sigma_x \).

Now note that if \( x \) is visited before \( w \), then \( x \) cannot be the last vertex, i.e. \( X_{\tau_{\text{cov}}} \neq x \); but if \( w \) is visited before \( x \), then every other vertex must have also been visited before \( x \) since there is no way to get from \( u \) to \( w \) without either passing through \( x \) or passing through literally every other vertex.

Finally, note that this is simply a Gambler’s Ruin problem - where the gambler starts with 1 dollar (since \( u \) is next to \( x \)) and wins if he gets to \( n - 1 \) dollars (since \( w \) is the target). The probability of winning is just \( \frac{1}{n-1} \). Since this holds regardless of what \( x \) is (provided \( x \neq v \) of course) we get that every non-\( v \) vertex has an equal probability of being the final vertex.

*(Sanity check: The probabilities should sum up to 1, which they do because there are \( n - 1 \) non-starting vertices, each with \( \frac{1}{n-1} \) probability of being the last visited.)*

**Exercise 7.** Let \( B_k \) be iid with law \( \mathbb{P}[B_k = +1] = p = 1 - \mathbb{P}[B_k = -1] \).

Answer the following:

- Let \( X_n = B_n B_{n+1}, n \geq 0 \). Is it Markov? If yes, find its transition kernel.
• Let \( Y_n = \frac{1}{2} (B_n - B_{n-1}) \), \( n \geq 1 \). Is it Markov? If yes, find its transition kernel.

• Let \( Z_n = | \sum_{k=1}^{n} B_k | \), \( n \geq 1 \). Is it Markov? If yes, find its transition kernel.

• If \( \{ V_i, i \geq 0 \} \) is a Markov process with state space \( \mathcal{X} \), and \( E_j \) are some subsets of \( \mathcal{X} \), is it true that

\[
P[V_n \in E_n | V_{n-1} \in E_{n-1}, V_{n-2} \in E_{n-2}, \ldots, V_0 \in E_0] = P[V_n \in E_n | V_{n-1} \in E_{n-1}],
\]

provided that \( P[V_{n-1} \in E_{n-1}, \ldots, V_0 \in E_0] > 0 \)?

• Suppose that \( P(x, y) \) is a kernel of an irreducible Markov chain. If \( P(\cdot, x_1) = P(\cdot, x_2) \) show that \( \pi(x_1) = \pi(x_2) \), where \( \pi \) is a stationary distribution.

What if the chain is not irreducible?

**Solution:**
1) It is not Markov (a couple exceptions, listed at the end). Let \( p = 0.99 \), and consider \( P[X_3 = 1 | X_2 = -1] \). Note that \( X_2 = -1 \) means either \( B_2 = -1 \) and \( B_3 = 1 \) or vice versa; and (given no other information) these two cases are equally probable. So no matter what \( B_4 \) happens to be, \( P[X_3 = 1 | X_2 = -1] = 1/2 \). But now suppose that we add the information that \( X_1 = -1 \) as well. If \( X_1 = X_2 = -1 \), then we have one of the following two cases:

1. \((B_1, B_2, B_3) = (-1, 1, -1)\);
2. \((B_1, B_2, B_3) = (1, -1, 1)\).

Note that the second case is vastly more probable than the first; therefore,

\[
P[X_3 = 1 | X_2 = -1, X_1 = -1] > 1/2
\]

(we could calculate it precisely using Baye’s Theorem, but we don’t really need to go to the trouble). Therefore \( \{X_n\} \) does not satisfy the Markov property.

(Remark: The exceptions are when \( p = 1/2 \) or, if we’ll allow such a thing, \( p = 0 \) or 1.)

2) Same as for 1 - a counterexample can be easily constructed, so it is not Markovian.

3) Yes it is Markov, although this is far from obvious. We’ll be using the reflection principle to see this. First, note that if \( Z_n = 0 \), then \( Z_{n+1} = 1 \) for
sure, so that \( P(0, 1) = 1 \); also note that \( Z_n \) can never move except by 1, so \( P(i, j) = 0 \) for all \( |i - j| \neq 1 \).

Now let’s start with the difficult part. Since

\[
Z_n = \sum_{k=1}^{n} B_n
\]

it is obvious that \( P(i, j) = 0 \) if \( j \neq i - 1, i + 1 \). Furthermore, we can easily see that \( P(0, 1) = 1 \) (and that this obviously does not depend on the history), and that \( Z_n \) can never be negative. Now we just have to examine \( P(i, i + 1) \) (noting that \( P(i, i - 1) = 1 - P(i, i + 1) \)).

We define \( W_n := \sum_{k=1}^{n} B_k \). Now note that if we know whether \( W_n \) is positive or negative, we could immediately determine \( \mathbb{P}[Z_{n+1} = Z_n + 1] \) – it would be \( p \) if \( W_n > 0 \), and \( 1 - p \) if \( W_n < 0 \) – and therefore the transition probabilities would only be determined by the current position \( Z_n \).

Now suppose that \( Z_k = z_k \) for all \( k = 0, 1, \ldots, n \), and \( z_n = \ell \) (the current state). Then we can define a possible history of \( W_k \)'s as a sequence \( w = (w_0, w_1, \ldots, w_n) \) such that

- \( w_k \in \{-z_k, z_k\} \) (so that \( |w_k| = z_k \)) for all \( k \);
- \( |w_k - w_{k-1}| = 1 \) for all \( k = 1, 2, \ldots, n \).

Define \( S \) to be the set of all such sequences (and obviously it is finite); define

\[
S_- := \{ w \in S : w_n = -\ell \} \quad \text{and} \quad S_+ := \{ w \in S : w_n = \ell \}
\]

Note the following:

- this is a partition of \( S \) – every \( w \in S \) is in exactly one of \( S_- \), \( S_+ \);
- \( |S_-| = |S_+| \) because for any \( w \in S_- \), we have \( -w \in S_+ \) (this is the “reflection” we were talking about). So let’s call

\[
m := |S_-| = |S_+|
\]

- for any \( w \in S_- \), we have \( \frac{n-\ell}{2} \) increments (corresponding to \( B_k = 1 \)) and \( \frac{n+\ell}{2} \) decrements (corresponding to \( B_k = -1 \)), and for any \( w \in S_+ \), we have \( \frac{n+\ell}{2} \) increments and \( \frac{n-\ell}{2} \) decrements. Therefore,

\[
\mathbb{P}[W_k = w_k \text{ for all } k \leq n] = \begin{cases} 
\frac{n+\ell}{2} p (1 - p) \frac{n+\ell}{2} & \text{if } w \in S_+ \\
\frac{n-\ell}{2} p (1 - p) \frac{n-\ell}{2} & \text{if } w \in S_-
\end{cases}
\]

Note that this only depends on the value of \( w_n \).
Note, therefore, that

\[
\mathbb{P}[Z_k = z_k \text{ for } k \leq n] = \sum_{w \in S} \mathbb{P}[W_k = w_k \text{ for } k \leq n] = \sum_{w \in S_+} \mathbb{P}[W_k = w_k \text{ for } k \leq n] + \sum_{w \in S_-} \mathbb{P}[W_k = w_k \text{ for } k \leq n] = \sum_{w \in S_+} p^{n+\ell} (1-p)^{n-\ell} + \sum_{w \in S_-} p^{n-\ell} (1-p)^{n+\ell} = m(p^{n+\ell} (1-p)^{n-\ell} + p^{n-\ell} (1-p)^{n+\ell}) = m(p(1-p))^{\ell} (p^{\ell} + (1-p)^{\ell})
\]

Now we can apply Bayes’ Theorem (remember that \(z_n = \ell\) here):

\[
\mathbb{P}[W_n = \ell | Z_k = z_k \text{ for } k \leq n] = \frac{\mathbb{P}\{W_k \in S_+ \mid Z_k = z_k \text{ for } k \leq n\}}{\mathbb{P}\{Z_k = z_k \text{ for } k \leq n\}} = \frac{\mathbb{P}\{W_k \in S_+\} \cdot \mathbb{P}[Z_k = z_k \mid \{W_k\} \in S_+]}{\mathbb{P}[Z_k = z_k \text{ for } k \leq n]} = \frac{\mathbb{P}[Z_k = z_k \text{ for } k \leq n]}{m(p(1-p))^{\ell} (p^{\ell} + (1-p)^{\ell})} = \frac{p^{\ell}}{p^{\ell} + (1-p)^{\ell}}
\]

(part of this was noting that \(\mathbb{P}[Z_k = z_k \mid \{W_k\} \in S_+] = 1\) by definition of \(S_+\).

Note that this depends only on the value of \(z_n = \ell\), and not on any other \(Z_k\)’s or even on \(n\) – so therefore we can conclude that it is Markovian!

Now we have to compute the transition kernel. We have:

\[
\mathbb{P}[W_n = -\ell | Z_k = z_k \text{ for } k \leq n] = 1 - \frac{p^{\ell}}{p^{\ell} + (1-p)^{\ell}} = \frac{(1-p)^{\ell}}{p^{\ell} + (1-p)^{\ell}}
\]

Therefore (letting \(z_n = \ell\) below), we get

\[
\mathbb{P}[Z_{n+1} = \ell + 1 \mid Z_k = z_k \text{ for } k \leq n] = \mathbb{P}[Z_{n+1} = \ell + 1 \text{ and } W_n = \ell \mid Z_k = z_k \text{ for } k \leq n] + \mathbb{P}[Z_{n+1} = \ell + 1 \text{ and } W_n = -\ell \mid Z_k = z_k \text{ for } k \leq n]
\]
Dealing with each piece here on its own, we get:

\[
\mathbb{P}[Z_{n+1} = \ell + 1 \text{ and } W_n = \ell \mid Z_k = z_k \text{ for } k \leq n] = \mathbb{P}[Z_{n+1} = \ell + 1 \mid W_n = \ell \text{ and } Z_k = z_k \text{ for } k \leq n] \cdot \mathbb{P}[W_n = \ell \mid Z_k = z_k \text{ for } k \leq n] = p \cdot \mathbb{P}[W_n = \ell \mid Z_n = \ell] = \frac{p^{\ell+1}}{p^\ell + (1-p)^\ell}
\]

An analogous computation for the other piece gives

\[
\mathbb{P}[Z_{n+1} = \ell + 1 \text{ and } W_n = -\ell \mid Z_k = z_k \text{ for } k \leq n] = \frac{(1-p)^{\ell+1}}{p^\ell + (1-p)^\ell}
\]

We then finally put all of this together to get

\[
P(\ell, \ell + 1) = \mathbb{P}[Z_{n+1} = \ell + 1 \mid Z_n = \ell] = \frac{p^{\ell+1} + (1-p)^{\ell+1}}{p^\ell + (1-p)^\ell}
\]

(and of course \(P(\ell, \ell - 1) = 1 - \frac{p^{\ell+1} + (1-p)^{\ell+1}}{p^\ell + (1-p)^\ell}\)). As noted at the very top, we have also \(P(0, 1) = 1\) and \(P(i, j) = 0\) for all \(j \neq i + 1, i - 1\).

(\textbf{Remark:} A common error was to assume that because the answer differs depending on the history of the \(B_k\)’s, it cannot be Markov. But when evaluating whether the \(Z_n\)’s are Markov, you cannot look at the history of the \(B_k\)’s, only on the history of the \(Z_n\)’s.)

4) Not always. An easy example is a random walk on a 6-cycle (labeled in order \(a, b, c, d, e, f\)) with uniformly-randomly-chosen starting point \(V_0\); let \(E_n = \{a\}\) and \(E_{n-2} = \{d\}\) and \(E_{n-1} = \mathcal{X}\) (the rest of the \(E_k\) don’t matter, but if we want to feel better about ourselves we can set them to \(\mathcal{X}\) as well). Then

\[
\mathbb{P}[V_n \in E_n \mid V_{n-1} \in E_{n-1}] = \mathbb{P}[V_n = a] = 1/6
\]

because the condition \(V_{n-1} \in \mathcal{X}\) says nothing. But of course if \(V_{n-2} \in E_{n-2}\) (i.e. \(V_{n-2} = d\)), there’s no way that \(V_n = a\) since you can’t reach it in time. So

\[
\mathbb{P}[V_n \in E_n \mid V_k \in E_k \text{ for all } k < n] = 0 \neq 1/6
\]

5) This follows easily from the equation \(\pi^T = \pi^T P\). If the chain is not irreducible, that does not alter the previous statement, so it remains true.