Solutions.

5.3.1 Let $p_j r = P(M \geq r | X_0 = j)$. Then for $j < r$ we have

$$p_j r = p p_{(j-1)r} + q p_{(j+1)r}.$$

Solution of this equation takes form

$$p_{jr} = A_r + B_r (q/p)^r.$$

Since $p_{jr} \to 0$ as $j \to -\infty$ we have $A_r = 0$. Now $p_{rr} = 1$ implies $B_r = (p/q)^r$. Plugging $j = 0$ we obtain the required formula.

5.4.1 $Z_n = \sum_{j=1}^{Z_{n-1}} X_j$ where X_j are iid having the same distribution as Z_1. It follows that

$$E(Z_n | Z_{n-1} = k) = kE(Z_1) = k\mu \text{ and so } E(Z_n) = E(Z_{n-1})\mu$$

Thus $E(Z_n) = \mu^n$. A similar computation gives

$$E(Z_n^2 | Z_{n-1} = k) = k^2\mu^2 + k\sigma^2 \text{ and so } E(Z_n^2) = E(Z_{n-1}^2)\mu^2 + \mu^{n-1}\sigma^2.$$

Iterating this formula we get

$$E(Z_n^2) = E(Z_{n-2}^2)\mu^4 + \mu_{n-1}\sigma^2 + \mu^n\sigma^2 = E(Z_{n-3}^2)\mu^6 + \mu^{n-1}\sigma^2 + \mu\sigma^2 + \mu^{n+1}\sigma^2 = \dots$$

$$= E(Z_0^2)\mu^{2n} + \left(\mu^{n-1} + \mu^n + \dots + \mu^{2n-2}\right)\sigma^2 = \mu^{2n} + \frac{\mu^{2n-1} - \mu^{n-1}}{\mu - 1}\sigma^2.$$

Hence

$$V(Z_n) = \frac{\mu^{2n-1} - \mu^{n-1}}{\mu - 1}\sigma^2.$$

Continuing in the same way we find

$$E(Z_n Z_m) = E(Z_m^2)\mu^{n-m} = \mu^{m+n} + \frac{\mu^{m+n-1} - \mu^{n-1}}{\mu - 1}\sigma^2.$$

Hence

$$Cov(Z_n Z_m) = \frac{\mu^{m+n-1} - \mu^{n-1}}{\mu - 1}\sigma^2 \text{ and so}$$

$$\rho(m,n) = \frac{\mu^{m+n-1} - \mu^{n-1}}{\mu - 1}\sigma^2 : \left(\frac{\sqrt{(\mu^{2m-1} - \mu^{m-1})(\mu^{2n-1} - \mu^{n-1})}}{\mu - 1}\sigma^2 \right) = \sqrt{\frac{\mu^{n-1}(\mu^m - 1)}{\mu^{m-1}(\mu^n - 1)}}.$$

5.4.2 **The statement is incorrect.** Indeed

$$Z_n = \sum_{j=1}^{Z_{n-1}} X_j$$

where $X_j \sim Z_1$. Therefore

$$P(L = n-1 | Z_{n-1} = k, X_1 = x_1 \dots X_k = x_k) = \frac{\sum_{j=1}^k x_j^2}{(\sum_{j=1}^k x_j)^2}$$

so by the Law of Large numbers

$$P(L = n-1 | Z_{n-1} = k) = \frac{E(Z_1^2)}{E(Z_1)^2} \frac{1}{k} (1 + o(1)).$$

Taking the expectation we obtain

$$P(L = n-1) = \frac{E(Z_1^2)}{E(Z_1)^2} E\left(\frac{1}{Z_{n-1}}\right)(1 + o(1)).$$

which is greater than the textbook answer

$$P(L = n-1) = E\left(\frac{1}{Z_{n-1}}\right).$$

The mistake is that while given that $Z_{n-1} = k$ it is true that for randomly chosen individual has probability $1/k$ to have individual A as an ancestor, the probability that the second individual would have A as an ancestor *given that A is the ancestor of the first individual* is greater $1/k$ since A is likely to have larger than average number of children.

5.12.7 For j different from 0 we have the following recursive relation

$$p_j = pp_{j+2} + qp_{j-1} \text{ that is } p_{j+2} = \frac{p_j - qp_{j-1}}{p}.$$

This recursion can be rewritten in the matrix form

$$\begin{pmatrix} p_{j+2} \\ p_{j+1} \\ p_j \end{pmatrix} \begin{pmatrix} \frac{1}{p} & 0 & \frac{-q}{p} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} p_{j+1} \\ p_j \\ p_{j-1} \end{pmatrix}.$$

Iterating this system we obtain

$$\begin{pmatrix} p_{j+2} \\ p_{j+1} \\ p_j \end{pmatrix} \begin{pmatrix} \frac{1}{p} & 0 & \frac{-q}{p} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}^j \begin{pmatrix} p_2 \\ p_1 \\ p_0 \end{pmatrix}.$$

It follow that p_j is a linear combination of the powers of eigenvalues of the above matrix. The characteristic equation take form

$$\lambda = p\lambda^3 + q$$

with solutions $\lambda = 1$, $\lambda = \lambda_1$ or $\lambda = \lambda_2$ where

$$\lambda_1 = \frac{\sqrt{p(1+3q)} - p}{2p}, \quad \lambda_2 = -\frac{\sqrt{p(1+3q)} + p}{2p}.$$

Now consider three cases:

(I) $p < 1/3$. In this case $X_n \to -\infty$ and since X_n can not skip points going left we have $p_j = 1$ for $j \geq 0$. For $j < 0$ we have

$$p_j = A + B\lambda_1^j + C\lambda_2^j,$$

where $A = 0$ since $p_j \to 0$ as $j \to -\infty$. Conditions $p_1 = p_0 = 1$ give

$$B + C = 1, \quad B\lambda_1 + C\lambda_2 = 1.$$

Solving this system we get

$$B = \frac{\lambda_2 - 1}{\lambda_2 - \lambda_1}, \quad C = \frac{\lambda_1 - 1}{\lambda_1 - \lambda_2}.$$

(II) $p = 1/3$. In this case $EX_n \equiv 0$ and so by Theorem 5.10.17 X_n is recurrent. Hence $p_j = 1$.

(III) $p > 1/3$. Then $X_n \to +\infty$. Thus for $j > 0$ we have $p_j = p_1^j$ since to go to j to 0 the particle should go from j to $j-1$, from $j-1$ to $j-2$ ... from 1 to 0. The only possibility satisfying $p_j \geq 0$, $p_j \to 0$ as $j \to \infty$ is $p_j = \lambda_1^j$. For $j < 0$ we have

$$p_j = A + B\lambda_1^j + C\lambda_2^j.$$

Condition $p_j \leq 1$ implies $B = 0$ while conditions $p_0 = 1, p_1 = \lambda_1$ give

$$A + C = 1, \quad A + C\lambda = \lambda_1.$$

Solving this system we obtain

$$A = \frac{\lambda_2 - \lambda_1}{\lambda_2 - 1}, \quad C = \frac{\lambda_2 - \lambda_1}{\lambda_2 - 1}.$$

6.1.2 (a) Yes.

$$p_{ij} = \begin{cases} 5/6 & \text{if } j = i \\ \frac{1}{6} & \text{if } j = i+1 \\ 0 & \text{otherwise} \end{cases}.$$

(b) Yes.

$$p_{ij} = \begin{cases} 5/6 & \text{if } j > i \\ \frac{i}{6} & \text{if } j = i \\ 0 & \text{otherwise} \end{cases}.$$

(c) Yes.

$$p_{ij} = \begin{cases} 5/6 & \text{if } j = i+1 \\ \frac{i}{6} & \text{if } j = 0 \\ 0 & \text{otherwise} \end{cases}.$$

(d) Yes.

$$p_{0j} = \left(\frac{5}{6}\right)^{j-1} \frac{1}{6}$$

otherwise $p_{j,j-1} = 1$.

6.1.7

$$P(Y_n = y_n | Y_{n-1} = y_{n-1}, Y_{n-2} = y_{n-2} \dots Y_0 = y_0)$$

$$= P(h(X_n) = y_n | h(X_{n-1}) = y_{n-1}, h(X_{n-2}) = y_{n-2} \dots h(Y_0) = y_0)$$

$$= P(X_n = h^{-1} y_n | X_{n-1} = h^{-1} y_{n-1} X_{n-2} = h^{-1} y_{n-2} \dots X_0 = h^{-1} y_0)$$

$$= P(X_n = h^{-1} y_n | X_{n-1} = h^{-1} y_{n-1}) = P(Y_n = y_n | Y_{n-1} = y_{n-1}).$$

Condition that h is invertible is important. Let $X_{n+1} = 2X_n$ if $X_n \geq 0$ and $X_{n+1} = X_n$ if $X_n \leq 0$. Let $Y_n = |X_n|$. Suppose that $Y_{n-1} = 6$. Then $Y_n = 12$ if $Y_{n-2} = 3$ and $Y_n = 6$ if $Y_{n-2} = 6$ however without knowing Y_{n-2} we do not have enough information to determine Y_n.

6.1.8 Let $X_{n+1} = 2X_n$, $Y_{n+1} = Y_n$ $Z_n = X_n + Y_n$. Then

$$Z_n - Z_{n-1} = X_{n-1} = 2X_{n-2} = Z_{n-1} - Z_{n-2}.$$

Thus $Z_n = 2Z_{n-1} - Z_{n-2}$ however just knowing Z_{n-1} does not give enough information.

6.2.1

$$P(X_n = j | X_0 = i) = \sum_{m=0}^{n-1} P(X_m = i | X_0 = i) P(X_n = j \text{ and } X_k \neq i \text{ for } m < k < n | X_m = i)$$

$$= \sum_{m=0}^{n-1} p_{ii}(m) l_{ij}(n-m).$$

Thus

$$P_{ij}(s) = \sum_n \sum_m p_{ii}(m) s^m l_{ij}(n-m) s^{n-m} = P_{ii}(s) L_{ij}(s)$$

proving the first claim. Comparing this with Theorem 6.2.3 (b) proves the second claim. An example of the chain satisfying this condition is a simple random walk.

6.2.5 Let $p_1 = P(T_j < T_i | X_0 = i)$, $p_2 = P(T_i < T_j | X_0 = j)$. Let V be the number of visits to j before return to i. Then for $n \neq 0$

$$P(V = n) = p_1(1-p_2)^{n-1} p_2 = p^2(1-p)^{n-1}$$

if $p_1 = p_2 = p$. The generating function takes form

$$\phi(s) = \frac{p^2 s}{1-(1-p)s} + (1-p) \text{ so } \phi'(s) = \frac{p^2(1-p)s}{(1-(1-p)s)^2} + \frac{p^2}{1-(1-p)s}.$$

Thus $\phi'(1) = (1-p) + p = 1$.

6.3.1 If there are infinitely many j such that $a_j > 0$ then the chain is irreducible since to get from k to m one can go $k \to 0$, $0 \to j$ for some $j > m$ and $j \to m$. Now from any nonzero state the particle can only go down so it eventually reaches zero and hence zero is recurrent. Since the chain is irreducible all states are recurrent. If $a_j = 0$ for $j > j_0$, $a_{j_0} = 0$ then states $0, 1, \ldots j_0$ are recurrent and all other states are transient. The equation for stationary distribution reads

$$\pi_0 = a_0\pi_0 + (1-p)\pi_1,$$

$$\pi_{j-1} = p\pi_{j-1} + (1-p)\pi_j + a_{j-1}\pi_0, \quad j > 2.$$

From the first equation

$$\pi_1 = \frac{(1-a_0)\pi_0}{1-p}.$$

From the second equation

$$\pi_j = \pi_{j-1} - \frac{a_{j-1}\pi_0}{1-p} = \pi_1 - \sum_{k=1}^{j-1} \frac{a_k pi_0}{1-p} = \frac{\pi_0(1-\sum_{k=0}^{j-1} a_k)}{1-p} = \frac{\pi_0 P(Z \geq j)}{1-p}$$

where Z takes value j with probability a_j. Using that

$$1 = \sum_j \pi_j = \frac{\pi_0}{1-p}\sum_k P(Z \geq k) = \frac{\pi_0 EZ}{1-p}$$

we get

$$\mu_0 = \frac{1}{\pi_0} = \frac{EZ}{1-p}.$$

Thus

$$\mu_j = \frac{1}{\pi_j} = \frac{1-p}{P(Z \geq j)\pi_0} = \frac{EZ}{P(Z \geq j)}.$$

6.3.4 Denote $\tau_u = E(T_v|X_0 = u)$. Let S be a symmetry of the cube such that $Sv = v$. Then

$$P(X_1 = u_1, X_2 = u_2 \ldots X_n = u_n | X_0 = u_0)$$
$$= P(X_1 = S(u_1), X_2 = S(u_2) \ldots X_n = S(u_n) | X_0 = S(u_0)).$$

In particular

$$P(X_1 \neq v, \ldots X_{n-1} \neq v, X_n = v | X_0 = u) = P(X_1 \neq v, \ldots X_{n-1} \neq v, X_n = v | X_0 = Su),$$

so $\tau_u = \tau_{Su}$. Let x be a vertex near v and y be a vertex near w. Conditioning on X_1 we obtain

$$\tau_x = \frac{1}{4}\tau_x + \frac{2}{4}\tau_y + 1,$$
$$\tau_y = \frac{1}{4}\tau_y + \frac{1}{4}\tau_w + \frac{2}{4}\tau_x + 1,$$
$$\tau_w = \frac{1}{4}\tau_w + \frac{3}{4}\tau_y + 1.$$

Solving this system we get

$$\tau_x = \frac{28}{3}, \quad \tau_y = 12, \quad \tau_w = \frac{40}{3}.$$

Next $\tau_v = \frac{3}{4}\tau_x + 1 = 8$. Thus the answer to (a) is 8 and the answer to (b) is $\frac{40}{3}$. (Another way to solve (a) is to observe that the stationary distribution gives equal weight $1/8$ to each vertex of the cube). The answer to (c) is 1 due to problem 6.2.5.

Another way to solve (c) is the following. Denote σ_u the expected number of visits to w before T_v given that the chain starts at u. Conditioning on X_1 gives

$$\sigma_x = \frac{2}{4}\sigma_y + \frac{1}{4}\sigma_x,$$
$$\sigma_y = \frac{2}{4}\sigma_x + \frac{1}{4}\sigma_w + \frac{1}{4}\sigma_y,$$
$$\sigma_w = \frac{3}{4}\sigma_y + \frac{1}{4}\sigma_w + 1.$$

Solving this system we get

$$\sigma_x = \frac{4}{3}, \sigma_y = 2, \sigma_w = \frac{10}{3}.$$

Thus $E(\text{ visits to } w \text{ befroe return to } v) = \frac{3}{4}\tau_x = 1$.

6.4.6 Write $u \sim v$ if there is an edge from u to v. Then

$$\sum_{u\sim v} \pi_u p_{uv} = \sum_{u\sim v} \frac{\pi_u}{d_u} = \sum_{u\sim v} \frac{1}{d_u}\frac{d_u}{2\mu} = \frac{1}{2\eta}\sum_{u\sim v} 1 = \frac{d_v}{2\eta} = \pi_v.$$

6.4.7 Let Y_n be the distance from X_n to the origin. Then Y_n is a random walk on nonnegative integers with $p = \frac{2}{3}$, $q = \frac{1}{3}$ so it is nonrecurrent.

6.6.8 The charactersitic equation is

$$\lambda^3 - \frac{4\lambda^2 + 7\lambda + 1}{12} = 0.$$

Factorizing

$$12\lambda^3 - (4\lambda^2 + 7\lambda + 1) = (\lambda - 1)(12\lambda^2 + 7\lambda + 1)$$

we find the roots

$$\lambda_1 = 1, \quad \lambda_2 = -\frac{1}{2}, \quad \lambda_3 = -\frac{2}{3}.$$

Solving the equation for the stationary distribution we get

$$\pi_1 = \pi_2 = \pi_3 = \frac{1}{3}.$$

Thus we have

$$p_{ij}(n) = \frac{1}{3} + a_{ij}\left(-\frac{1}{2}\right)^n + b_{ij}\left(-\frac{2}{3}\right)^n.$$

Initial conditions

$$p_{ij}(0) = \delta_{ij}, \quad p_{ij}(1) = p_{ij}$$

imply

$$\frac{1}{3} + a_{ij} + b_{ij} = \delta_{ij}, \quad \frac{1}{3} - \frac{a_{ij}}{2} - \frac{2b_{ij}}{3} = p_{ij}.$$

This gives

$$a_{ij} = 4\delta_{ij} + 6p_{ij} - \frac{4}{3}, \quad b_{ij} = 1 - 3\delta_{ij} - 6p_{ij}.$$

6.8.4 The forward equation takes the following form

$$p_n' = -\lambda n p_n + \lambda(n-1)p_{n-1}.$$

Since this equation has unique solution it is enough to check that

$$\begin{pmatrix} n-1 \\ I-1 \end{pmatrix} e^{-n\lambda t}\left(1 - e^{-\lambda t}\right)^{n-I}.$$

Differentiating the product we get $-\lambda n p_n$ from the first factor and

$$\lambda \begin{pmatrix} n-1 \\ I-1 \end{pmatrix} e^{-n\lambda t}(n-I)\left(1 - e^{-\lambda t}\right)^{n-I-1}$$

from the second factor. Since

$$\begin{pmatrix} n-1 \\ I-1 \end{pmatrix}(n-I) = \frac{(n-1)!}{(n-I)!(I-1)!}(n-I) = \frac{(n-1)!}{(n-I-1)!(I-1)!}$$

$$= (n-1)\frac{(n-2)!}{(n-I-1)!(I-1)!}$$

the last expression contributes $\lambda(n-1)p_{n-1}$ as claimed.

Next if $G(t,s)$ is the generating function for $p_n(t)$ then the forward equations imply

$$\frac{\partial G}{\partial t} = \lambda(s^2 - s)\frac{\partial G}{\partial s}.$$

Differentiating with respect to s we get

$$\frac{\partial}{\partial t}\frac{\partial G}{\partial s} = \lambda\left(2s-1\right)\frac{\partial G}{\partial s} + \lambda\left(s^2 - s\right)\frac{\partial^2 G}{\partial s^2}.$$

Plugging $s = 1$ we see that $\mu(t) = E(X(t))$ satisfies $\mu' = \lambda\mu$. Hence $\mu(t) = Ie^{\lambda t}$. Taking another derivative we get

$$\frac{\partial}{\partial t}\frac{\partial^2 G}{\partial s^2} = 2\lambda\frac{\partial G}{\partial s} + 2\lambda\left(2s-1\right)\frac{\partial^2 G}{\partial s^2} + \lambda\left(s^2 - s\right)\frac{\partial^3 G}{\partial s^3}.$$

Hence $W(t) = \frac{\partial^2 G}{\partial s^2}(t,1)$ satisfies

$$W'(t) = 2\lambda\mu(t) + 2\lambda W(t).$$

Solving this equation with the boundary condition $W(0) = I^2 - I$ we obtain

$$W(t) = (I^2 + I)e^{2\lambda t} - 2Ie^{\lambda t}.$$

Hence $V(t) = \mathrm{Var}(X(t))$ satisifies

$$V(t) = W(t) + \mu(t) - \mu^2(t) = I(e^{2\lambda t} - e^{\lambda t}).$$

6.9.1 (a) The forward equation gives

$$p_1' = -\lambda p_1 + \mu p_2.$$

Substituting $p_2 = 1 - p_1$ we get

$$p_1 = -(\lambda+\mu)p_1 + \mu.$$

Hence

$$p_1(t) = \frac{\mu}{\lambda+\mu} + \left(p_1(0) - \frac{\mu}{\lambda+\mu}\right)e^{-(\lambda+\mu)t}.$$

Likewise

$$p_2(t) = \frac{\lambda}{\lambda+\mu} + \left(p_2(0) - \frac{\lambda}{\lambda+\mu}\right)e^{-(\lambda+\mu)t}.$$

Hence

$$P(t) = \begin{pmatrix} \frac{\mu}{\lambda+\mu} + \left(\frac{\lambda}{\lambda+\mu}\right)e^{-(\lambda+\mu)t} & \frac{\lambda}{\lambda+\mu}\left(1 - e^{-(\lambda+\mu)t}\right) \\ \frac{\mu}{\lambda+\mu}\left(1 - e^{-(\lambda+\mu)t}\right) & \frac{\lambda}{\lambda+\mu} + \left(\frac{\mu}{\lambda+\mu}\right)e^{-(\lambda+\mu)t} \end{pmatrix}.$$

(b) A direct computation shows that $G^2 = -(\lambda+\mu)G$. Hence $G^n = (-(\lambda+\mu))^{n-1}G$ and

$$I+\sum_{k=1}\frac{t^kG^k}{k!} = I-\frac{G}{\lambda+\mu}\left[\left(\sum_{k=0}^{\infty}\frac{(-(\lambda+\mu)t)^k}{k!}\right)-1\right] = I+\frac{G}{\lambda+\mu}\left[1-e^{-(\lambda+\mu)t}\right]$$

which coincides with the answer obtained in part (a).

(c) Equation for stationary distribution takes form $-\pi_1\mu+\pi_2\lambda=0$, that is $\pi_1 = \frac{\lambda}{\mu}\pi_2$. Thus

$$\pi_2 + \frac{\lambda}{\mu}\pi_2 = 1 \text{ so } \pi_2 = \frac{\mu}{\mu+\lambda}.$$

Likewise $\pi_1 = \frac{\lambda}{\mu+\lambda}$ which coincides with the limiting values obtained in (a) and (b).

Another way to get the same result is the following. Let U_n be the lengths of consecutive visits to 1 and V_n be the lengths of consecutive visits to 2. Then U_n are iid Exp(μ) random variables and V_n are iid Exp(μ) random variables. By law of large numbers

$$\sum_{n=1}^{N} U_n \approx \frac{N}{\mu} \quad \sum_{n=1}^{N} V_n \approx \frac{N}{\lambda}.$$

Thus the ratio of the time spent at 1 to the time spent at 2 equals λ/μ explaining the result of (c).

6.9.9 Let T be the total time spent at site i and let Z_t be the first time when the time spent in i equals t. Then for any $u>t$

$$P(T>u) = P(T>t)P(X \text{ will spend time } (u-t) \text{ at } i \text{ after } Z_t|X_{Z_t}=i)$$

$$= P(T>t)P(T>(u-t)).$$

Thus T is memoryless and so it is exponentially distributed.

6.11.3 The backward equation reads

$$p_{10}' = -(\lambda+\mu)p_{10} + \lambda p_{20} + \mu p_{00} = -(\lambda+\mu)p_{10} + \lambda p_{20} + mu.$$

Since $\lambda_n = n\lambda_1$, $\mu_n = n\mu_1$, each individual either does or divides into 2 with the same intensities independently of others. Thus if divide individuals into k groups each group would evolve independently. If we start with k individuals then the population die if offsprings of each individual die which are independent evens. So that $p_{k0} = \eta(t)^k$. In particular the forward equation takes form

$$\eta' = \mu - (\lambda+\mu)\eta + \lambda\eta^2.$$

$$\frac{d\eta}{\mu - (\lambda+\mu)\eta + \lambda\eta^2} = dt.$$

Decomposing the right hand side as

$$-\frac{1}{\mu-\eta}\left[\frac{1}{1-\eta} - \frac{1}{\frac{\mu}{\lambda}-\eta}\right]$$

and integrating we obtain

$$C - (\mu-\lambda)t = \ln\left(\frac{1-\eta}{\frac{\mu}{\lambda}-\eta}\right).$$

Thus

$$\frac{1-\eta}{\frac{\mu}{\lambda}-\eta} = Ae^{-(\mu-\lambda)t} = \frac{\lambda}{\mu}e^{(\lambda-\mu)t}$$

since $\eta(0) = 0$. Solving this equation we get

$$\eta(t) = \frac{e^{(\lambda-\mu)t}-1}{\frac{\lambda}{\mu}e^{(\lambda-\mu)t}-1}.$$

$$P(X(t) = 0|X(u) = 0) = \frac{\eta(t)}{\eta(u)}.$$

8.2.1 (*b*) $P(X_r = 1) = P(X_0 = 0)p_{01}(n) + P(X_0 = 1)p_{11}(n).$

Since the chain is irreducible $p_{01}(n) \to \pi$ and $p_{11}(n) \to \pi$, where π_1 is a stationary probability. Thus

$$\lim_{r\to\infty} P(X_r = 1) = P(X_0 = 0)\pi_1 + \pi_1 p_{11}(n) = \pi_1(P(X_0 = 0) + P(X_0 = 1)) = \pi_1.$$

Thus also the average

$$\frac{1}{n}\lim_{r=1}^{n} P(X_r = 1) \to \pi_1.$$

To find π_1 explicitly combine the stationarity equation

$$\pi_1 = \alpha\pi_0 + (1-\beta)\pi_1 \text{ with } \pi_0 + \pi_0 = 1$$

to find $\pi_1 = \frac{\alpha}{\alpha+\beta}$. If the process is strictly stationary we need $P(X_r = s)$ not to depend on r that is

$$P(X_r = 0) = \frac{\beta}{\alpha+\beta}, \quad P(X_r = 1) = \frac{\alpha}{\alpha+\beta}.$$

On the other hand if this condition holds then the process is strictly stationary since

$$P(X_r = s_r, X_{r+1} = s_{r+1} \dots X_{r+m} = s_{r+m}) = \pi_{s_r} \times p_{s_r s_{r+1}} \times \cdots \times p_{s_{r+m-1}s_{r+m}}.$$

$$(a) \quad \rho(m, m{+}n) = \frac{Cov(X_m X_{m+n})}{\sqrt{V(X_m X_{m+n})}} = \frac{P(X_{m=1} X_{m+n} = 1) - P(X_{m=1}) P(X_{m+n} = 1)}{\sqrt{P(X_m = 1) P(X_m = 0) P(X_{m+n} = 1) P(X_{m+n} = 0)}}$$
$$= \frac{P(X_{m=1} p_{11}(n) - P(X_{m=1}) P(X_{m+n} = 1)}{\sqrt{P(X_m = 1) P(X_m = 0) P(X_{m+n} = 1) P(X_{m+n} = 0)}}$$

where

$$P(X_r = s) = P(X_0 = 0) p_{0s}(r) + P(X_0 = 1) p_{1s}(r).$$

To find the transition probabilities explicitly observe that the trace of of the transition matrix equals $2 - (\alpha + \beta)$ and so its eigenvalues are 1 and $(1 - (\alpha + \beta)$. Hence $p_{ij}(r) = a_{ij} + b_{ij}(1 - (\alpha + \beta)^r$. Using that $p_{ij}(0) = \delta_{ij}$ and $\lim_{r\to\infty} = \pi_j$ we find $a_{ij} = \pi_j$, $b_{ij} = \delta_{ij} - \pi_j$. Finally arguing as in part (b) we find

$$\lim_{m\to\infty} \rho(m, m+n) = \frac{\pi_1 p_{11}(n) - \pi_1^2}{\pi_1 \pi_0} = \frac{p_{11} - \pi_1}{\pi_0} = (1 - (\alpha + \beta))^n.$$

8.4.3 Let M be the number of vehicle passing before the possibility to crossing occurs. Then $P(M = m) = p(1-p)^m$ where $p = e^{-\lambda a}$. Thus

$$T = \sum_{j=1}^{M} Y_j$$

where Y_j are exponential random variables conditioned to be less than a. Thus Y has density

$$p(y) = \frac{1}{1-p} \lambda e^{-\lambda x} 1_{0 \le y \le a}.$$

It follows that

$$EY = \int_0^a p(y) dy = \frac{1}{1-p} \left(\frac{1}{\lambda} - \frac{e^{-\lambda a}}{\lambda} - a e^{-\lambda a} \right).$$

Therefore

$$ET = EY EM = \left(\frac{1}{p} - 1 \right) \left(\frac{1}{1-p} \right) \left(\frac{1}{\lambda} - \frac{e^{-\lambda a}}{\lambda} - a e^{-\lambda a} \right)$$
$$= \frac{1}{\lambda} \left(e^{\lambda a} - 1 - a\lambda \right) = \sum_{k=2}^{\infty} \frac{a^k \lambda^{k-1}}{k!}.$$

(**Observe that** the answer given in the textbook (the summation starts from $k = 1$) is not correct. For example if there are no cars $\lambda = 0$ then $ET = 0$ not 1 as given in the text.)

Next

$$E(e^{\theta Y}) = \left(\frac{1}{1-p} \right) \left(\frac{\lambda}{\theta - \lambda} \right) \left(e^{(\theta - \lambda) a} - 1 \right).$$

Thus

$$E(e^{\theta T}|M=m) = \left(\frac{1}{1-p}\right)^m \left(\frac{\lambda}{\theta-\lambda}\right)^m \left(e^{(\theta-\lambda)a}-1\right)^m.$$

Therefore

$$E(e^{\theta T}) = \sum_{m=0}^{\infty} p(1-p)^m \left(\frac{1}{1-p}\right)^m \left(\frac{\lambda}{\theta-\lambda}\right)^m \left(e^{(\theta-\lambda)a}-1\right)^m$$

$$= \sum_{m=0}^{\infty} p \left(\frac{\lambda}{\theta-\lambda}\right)^m \left(e^{(\theta-\lambda)a}-1\right)^m = \frac{\theta-\lambda}{(\theta-2\lambda)e^{\lambda a}-\lambda e^{\theta a}}.$$

Finally, let $f(\lambda) = \sum_{k=2}^{\infty} \frac{a^k \lambda^{k-1}}{k!}$. In case (a) we need to cross the first lane and then the second lane so

$$ET_a = f(\lambda) + f(\mu).$$

In case (b) the union of cars form Poisson process with intensity $\lambda + \mu$ so

$$ET_b = f(\lambda+\mu).$$

Clearly $ET_b \geq ET_a$ since the island can not hurt (we can always ignore the island and cross both lane in one run). To see this analytically observe that f is convex since $f'' > 0$. Combining this with $f(0) = 0$ we get

$$f(\lambda) < f(\lambda+\mu)\frac{\lambda}{\lambda+\mu} + f(0)\frac{\mu}{\lambda+\mu} = f(\lambda+\mu)\frac{\lambda}{\lambda+\mu}.$$

Likewise

$$f(\mu) < f(\lambda+\mu)\frac{\mu}{\lambda+\mu}.$$

Adding these two equations we obtain

$$f(\mu) + f(\lambda) < f(\lambda+\mu)$$

as claimed.

8.5.3 Let $W(t) = aW_1(t) + bW_2(t)$. Then $W(t)$ is Gaussian since $W_1(t)$ and $W_2(t)$ are independent Gaussians and for $t_1 \leq t_2 \leq t_3 \leq t_4$ the increments $W(t_4) - W(t_3)$ and $W(t_2) - W(t_1)$ are independent since the random variables

$$W_1(t_4)-W_1(t_3), \quad W_1(t_2)-W_1(t_1), \quad W_2(t_4)-W_2(t_3), \quad W_2(t_2)-W_2(t_1)$$

are independent. Finally

$$E(W(t)) = aE(W_1(t))+bE(W_2(t)) = 0, \quad V(W(t)) = a^2V(W_1(t))+b^2V(W_2(t)) = (a^2+b^2)t.$$

Hence $W(t)$ is a standard Wiener process iff $a^2 + b^2 = 1$.

$$9.7.2 \quad Cov(X_0, X_n) = \sum_{k_1,k_2=1}^{r} a_{k_1}\overline{a_{k_2}}Cov(X_{-k_1}X_{n-k_2}) = \sum_{k_1,k_2=1}^{r} a_{k_1}a_{k_2}\sigma_Y^2\rho_Y(n+k_1-k_2)$$

$$= \int \sum_{k_1,k_2=1}^{r} a_{k_1}\overline{a_{k_2}}\sigma_Y^2 e^{i\lambda(n+k_1-k_2)} f_Y(\lambda)d\lambda$$

$$= \int \left(\sum_{k_1=1}^{r} e^{i\lambda k_1}a_{k_1}\right)\overline{\left(\sum_{k_2=1}^{r} e^{i\lambda k_2}a_{k_1}\right)} e^{i\lambda n} f_Y(\lambda)d\lambda = \int |G(\lambda)|^2\sigma_Y^2 f_Y(\lambda)e^{i\lambda n}d\lambda.$$

Since $\rho_X(n) = Cov(X_0, X_n)/\sigma_X^2$ we get

$$f_X(\lambda) = \frac{\sigma_Y^2}{\sigma_X^2} f_Y(\lambda)|G(\lambda)|^2.$$

In case $r = \infty$ $a_k = (1-\mu)\mu_k$

$$G(\lambda) = \sum_k (1-\mu)\mu^k e^{ik\lambda} = \frac{1-\mu}{1-\mu e^{i\lambda}}.$$

Hence

$$|G(\lambda)|^2 = \frac{(1-\mu^2)}{|1-\mu e^{i\lambda}|^2} = \frac{(1-\mu^2)}{(1-\mu\cos\lambda)^2 + \mu\sin^2\lambda} = \frac{(1-\mu^2)}{1-2\mu\cos\lambda+\mu^2}.$$

9.7.3 We have that $Cov(X_k, X_{n+1} - aX_n - bX_{n-1}) = 0$ for all $k \le n$. In particular taking $k = 0$ we get

$$\rho(n+1) = a\rho(n) + b\rho(n-1) \tag{1}$$

Conversely if (1) holds then $\rho(n+1-k) = a\rho(n-k) + b\rho(n-1-k)$ and hence $Cov(X_k, X_{n+1} - aX_n - bX_{n-1}) = 0$ for all $k \le n$. To solve (1) rewrite it

$$\begin{pmatrix} \rho(n+1) \\ \rho(n) \end{pmatrix} = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \rho(n) \\ \rho(n-1) \end{pmatrix} = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} \rho(1) \\ \rho(0) \end{pmatrix}.$$

Therefore $\rho(n) = c_1\lambda_1^n + c_2\lambda_2^n$ where λ_1 and λ_2 are the eigenvalues of the above matrix. It follows that

$$\rho(x) = \frac{1}{2\pi}\left(1 + \sum_{k=1}\left[c_1\lambda_1^k e^{ikx} + c_2\lambda_2^k e^{ikx}\right] + \sum_{k=1}\overline{\left[c_1\lambda_1^k e^{ikx} + c_2\lambda_2^k e^{ikx}\right]}\right)$$

$$= \frac{1}{2\pi}\left(1 + \frac{c_1 2\Re\hat{\lambda}_1(x)}{(1-\Re(\hat{\lambda}_1(x)))^2 + (\Im\hat{\lambda}_1(x))^2} + \frac{c_2 2\Re\hat{\lambda}_2(x)}{(1-\Re(\hat{\lambda}_2(x)))^2 + (\Im\hat{\lambda}_2(x))^2}\right)$$

where $\hat{\lambda}_j(x) = \lambda_j e^{ix}$. Here c_1 and c_2 can be find from the equations

$$\rho(0) = 1, \quad \rho(1) = a\rho(0) + b\overline{\rho(1)}$$

that is

$$c_1 + c_2 = 1, \quad c_1\lambda_1 + c_2\lambda_2 = a + b(\overline{c_1\lambda_1} + \overline{c_2\lambda_2}).$$

9.7.15 Consider a transformation S of $[0, 1)$ given by $S(x) = \{x+\alpha\}$ where $\{\dots\}$ denotes a fractional part. Then S preserves length since it is piecewise translation. We claim that S is ergodic. Indeed if A is an S invariant set. Consider the Fourier series

$$1_A(x) = \sum_k c_k e^{2\pi ikx}.$$

Then

$$1_A(Sx) = \sum_k c_k e^{2\pi ik(x+\alpha)} = \sum_k (c_k e^{2\pi i\alpha}) e^{2\pi ikx}.$$

By uniqueness of Fourier series we have

$$c_k = c_k e^{2\pi ik\alpha} \text{ that is } c_k = 0 \text{ or } e^{2\pi k\alpha} = 1.$$

The last alternative means $k\alpha \in \mathbb{Z}$. Since α is irrational $k = 0$. Thus $1_A = c_0$ which means $1_A = 0$ or $1_A = 1$ and so S is ergodic. Thus for almost every x

$$\frac{1}{n} g(x + j\alpha) = \frac{1}{n} g(\{x + j\alpha\}) \to \int_0^1 g(u)du$$

where the first equality follows since g is 1-periodic. This proves the result.

One can in fact obtain a stronger conclusion by representing

$$g(x) = \sum_{k=-N}^{N} c_k e^{2\pi ikx} + \tilde{g}(x)$$

where $|\tilde{g}(x)| \le \varepsilon$. Then using that $\int_0^1 g(u)du = c_0$ we obtain

$$\frac{1}{n} \sum_{j=1} g(x + j\alpha) - \int_0^1 g(u)du$$

$$= \frac{1}{n} \sum_{0<|k|\le N} \frac{c_k e^{2\pi ikx}(e^{2\pi ik\alpha} - e^{2\pi i(n+1)\alpha})}{1 - e^{2\pi ik\alpha}} + \frac{1}{n} \sum_{j=1}^{n} \tilde{g}(x + j\alpha).$$

Here the first sum tends to 0 as $n \to \infty$ and the second sum is less than ε in absolute value. Thus for all x

$$\frac{1}{n} \sum_{j=1} g(x + j\alpha) - \int_0^1 g(u)du$$

can be made as small as we wish by choosing ε small and n large. In other words the converges takes place for all x not merely almost all x.

11.8.1. Let n be the total number of people in the office (including the ones being served).

The invariant distribution satisfies $G\pi = 0$.

For one server this reduces to

$$\lambda\pi_n - (\mu+\lambda)\pi_{n+1} + \mu\pi_{n+2} = 0 \quad n > 0.$$

The general solution of this recurrence is

$$\pi_n = A\xi_1^n + B\xi_2^n$$

where ξ_j satisfy $\mu\xi_j^2 - (\lambda+\mu)\xi_j + \lambda = 0$. Thus $xi_1 = 1$, $\xi_2 = \lambda/\mu$. Denote $\xi = \lambda/\mu$. The boundary condition is

$$-\lambda\pi_0 + \mu\pi_1 = 0 \text{ that is} \pi_1 = \xi\pi_0.$$

Hence $A = .0$ The condition

$$\sum_{n=0}^{N} \pi_n = 1 \text{ gives } B = \frac{1-\xi}{1-\xi^{N+1}}.$$

Therefore

$$\pi_n = \frac{(1-\xi)\xi^n}{1-\xi^{N+1}}.$$

For two servers the equation becames

$$\lambda\pi_n - (2\mu+\lambda)\pi_{n+1} + 2\mu\pi_{n+2} = 0 \quad n > 0.$$

The general solution is

$$\pi_n = A + B\left(\frac{\xi}{2}\right)^n.$$

The boundary condition still reads

$$\pi_1 = \xi\pi_0$$

(since if there is only 1 person in the office only 1 server is busy). This gives

$$A = \frac{\xi}{2(1-\xi)}B.$$

The condition

$$\sum_{n=0}^{N} \pi_n = 1 \text{ gives } B = \frac{2(1-\xi)}{(N+1)\xi + 2\left((1-\left(\frac{\xi}{2}\right)^{N+1}\right)}.$$

Therefore

$$\pi_n = \frac{2(1-\xi)\left(\frac{\xi}{2}\right)^n + \xi}{(N+1)\xi + 2\left((1-\left(\frac{\xi}{2}\right)^{N+1}\right)}.$$

11.8.6 Queue length generating function equals

$$G(s) = (1-\rho)(s-1)\frac{M_S(\lambda(s-1))}{s - M_S(\lambda(s-1))}.$$

Denote

$$\phi(s) = \frac{s - M_S(\lambda(s-1))}{s-1}.$$

Using the Taylor series of the numerator we find that

$$\phi(1) = (s - M_S(\lambda(s-1)))'|_{s=1} = (1-\rho),$$

$$\phi(1)' = \frac{1}{2}(s - M_S(\lambda(s-1)))''|_{s=1} = \frac{\lambda^2}{2}E(S^2).$$

Now using

$$G(s) = (1-\rho)\frac{M_S(\lambda(s-1))}{\phi(s)}$$

we obtain the following expression for mean queue length

$$E(Q) = G'(1) = (1-\rho)\frac{\lambda M_S'(0)\phi(1) - M_S(0)\phi'(1)}{\phi^2(1)}$$

$$= (1-\rho)\frac{\rho(1-\rho) - \lambda^2 E(S^2)/2}{(1-\rho)^2} = \rho - \frac{\lambda^2 E(S^2)}{1-\rho}.$$

Next waiting time moment generating function takes form

$$M_W(s) = \frac{(1-\rho)s}{\lambda + s - \lambda M_S(s)}.$$

Letting

$$\psi(s) = \frac{\lambda + s - \lambda M_S(s)}{s}$$

we obtain

$$\psi(0) = (\lambda + s - \lambda M_S(s))'|_{s=0} = (1-\rho),$$

$$\phi(0)' = \frac{1}{2}(\lambda + s - \lambda M_S(s))''|_{s=0} = -\frac{\lambda}{2}E(S^2).$$

Therefore

$$E(W) = (1-\rho)\frac{d}{ds}|_{s=0}\left(\frac{1}{\psi(s)}\right) = \frac{-(1-\rho)\psi'(0)}{\psi^2(0)} = \frac{\lambda E(S^2)}{2(1-\rho)}.$$

Since $E(S^2) = (ES)^2 + \text{Var}(S)$ the waiting time is minimized if $\text{Var}(S) = 0$ that is if the service is deterministic.

11.8.8 The associated random walk S_n moves to -2 or 1 with probability 1/2. $M_n = \theta^{S_n}$ is a martingale provided that

$$E(\theta S_{n+1}|S_n) = M_n\left(\frac{\theta + \theta^{-2}}{2}\right) = M_n$$

that is $\theta^3 + 1 = 2\theta$. Since

$$\theta^3 - 2\theta + 1 = (\theta - 1)(\theta^2 + \theta - 1)$$

the possible values of θ are 1, $\frac{\sqrt{5}-1}{2}$ and $\frac{\sqrt{5}+1}{2}$. Choose $\theta = \frac{\sqrt{5}+1}{2}$. Let τ_N be the first time when either $S_n = 1$ or $S_n \le -N$ and let $p_N = P(S_{\tau_N}) = 1$. By optimal sampling theorem

$$p_N \frac{\sqrt{5}+1}{2} + o(1) = 1,$$

so

$$\lim_{N\to\infty} p_N = \frac{2}{\sqrt{5}+1} = \frac{\sqrt{5}-1}{2}.$$

In other words

$$P(S_n \text{ visits } 1) = \frac{\sqrt{5}-1}{2}.$$

By translational symmetry

$$P(S_n \text{ visits } k) = \left(\frac{\sqrt{5}-1}{2}\right)^k$$

and so

$$P(\max S_n = k) = \left(\frac{\sqrt{5}-1}{2}\right)^k - \left(\frac{\sqrt{5}-1}{2}\right)^{k+1} = \left(\frac{\sqrt{5}-1}{2}\right)^k \frac{3-\sqrt{5}}{2}.$$

Hence the limiting waiting time distribution is

$$P(W = k) = \left(\frac{\sqrt{5}-1}{2}\right)^k \frac{3-\sqrt{5}}{2}.$$

12.5.7 We have

$$E(M_{n+1}|F_n) = \sum_{r=0}^{n} S_r + E(S_{n+1}|F_n) - \frac{1}{3}E(S_{n+1}|F_n).$$

The second term here equals S_n while

$$E(S_{n+1}^3|F_n) = E((S_n^3 + 3S_n^2 X_{n+1} + 3S_n X_n^2 + X_n^3)|F_n) = S_n^3 + 3S_n.$$

Therefore M_n is a martingale.

Next, observe that there exists a constant c such that for any x, n

$$P(T \ge n + K | S_n = x) < (1 - c).$$

Iterating we obtain

$$P(T \ge n) \le (1 - c)^{[n/K]}.$$

Since $|M_T| \le TK + K^3$ M_T satisfies the conditions of the optimal sampling theorem. Therefore

$$E(M_0) = a - \frac{a^3}{3} = E(M_T) = E\left(\sum_{n=0}^{T} S_n\right) - \frac{E(S_T^3)}{3}.$$

Next $P(S_T = K) = \frac{a}{K}$ so

$$E\left(\sum_{n=0}^{T} S_n\right) = a + \frac{E(S_T^3)}{3} - \frac{a^3}{3} = a + \frac{a}{K}\frac{K^3}{3} - \frac{a^3}{3} = a + \frac{aK^2 - a^3}{3}.$$

12.5.8 The proof that $S_n - n$ and T satisfy the conditions of the optimal sampling theorem is the same as the proof for M_n in problems 12.5.7.

13.12.3 We have

$$\Delta U = -\beta h e^{-\beta t} hW + e^{-\beta t}\Delta W + O(h^2).$$

Taking expectaition we obtain

$$E(\Delta U|F_t) = -\beta h e^{-\beta t} hW(t) + O(h^2) = -\beta h U(t)$$

that is $a(u) = -\beta u$. Next using that $e^{2\beta(t+h)} - e^{2\beta t} = 2\beta h e^{2\beta t} + O(h^2)$ we get

$$E((\Delta U)^2|F_t) = e^{-2\beta t}E((\Delta W)^2) + O(h^2) = e^{-2\beta t}2\beta h e^{2\beta t} + O(h^2) = 2\beta h + O(h^2)$$

that is $b(u) = 2\beta$.

13.12.4 Problem 13.12.3 gives

$$U(t) = e^{-\beta t}W(e^{2\beta t} - 1)$$

where $\sigma_U^2 = 2\beta\sigma_W^2$. Thus $\sigma_W^2 = \frac{\sigma^2}{2\beta}$. Writing

$$U(t) = e^{-\beta t}U(0) + e^{-\beta t}\Delta W$$

where

$$e^{-\beta t}\Delta W \sim N\left(0, \frac{\sigma^2(e^{2\beta t} - 1)}{2e^{2\beta t}\beta}\right)$$

we obtain the first claim. As $t \to \infty$ the last fraction converges to $\sigma^2/(2\beta)$ establisihing the second claim. Next if $U(0) \sim N(0, \sigma^2/(2\beta))$ then $U(t)$ is normal with zero mean and variance

$$e^{-2\beta t}\frac{\sigma^2}{2\beta} + \frac{\sigma^2(1 - e^{-2\beta t})}{2\beta} = \frac{\sigma^2}{2\beta}$$

proving stationarity.