B.5 Modes of convergence and renewal theory (Sheet 4 for classes)

1. (a) Let $\xi$ satisfy the detailed balance equations. Fix $j$ and sum the associated equations over $i \in S$ to see that

$$\langle \xi \rangle_j = \sum_{i \in S} \xi_i q_{ij} = \sum_{i \in S} \xi_j q_{ji} = \xi_j \sum_{i \in S} q_{ji} = 0$$

since the rows of a $Q$-matrix sum to zero, by the definition of the diagonal elements as minus the sums of the other row-entries. Therefore $\xi Q = 0$ and so $\xi$ is invariant.

(b) For the given $Q$-matrix, the detailed balance equations only have the trivial solution $\xi = (0, 0, 0)$ which is not a probability distribution. However, $\xi = (1/3, 1/3, 1/3)$ is clearly a stationary distribution as it satisfies $\xi Q = 0$. Therefore, the converse of (a) is false: there are $Q$-matrices for which there are stationary distributions, that do not satisfy the detailed balance equations. This is not surprising, since the detailed balance equations are a system of $\binom{|S|}{2}$ linear equations for only $|S|$ unknowns.

2. (a) For all $\varepsilon > 0$ we have $\mathbb{P}(\|X_n\| > \varepsilon) \leq n^{-2} \to 0$, and this implies $X_n \to 0$ in probability. $\mathbb{E}[X_n] = n^3 n^{-2} = n \to \infty$ implies that $\mathbb{E}(X_n)$ does not converge.

(b) $A_n = \{X_n > n^{-1}\}$ satisfies

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \sum_{n=1}^{\infty} n^{-2} < \infty \quad \Rightarrow \quad \mathbb{P}(A_n \text{ infinitely often}) = 0$$

by the first Borel-Cantelli lemma. This means that there is a random variable $N$ with $\mathbb{P}(N < \infty) = 1$ such that for all $n \geq N$, $0 \leq X_n < n^{-1}$. Therefore $X_n \to 0$ almost surely. So, almost sure convergence does not imply convergence of means. In particular, convergence in probability does not imply convergence of means.

(c) Of the $2^{2n}$ paths of length $2n$, precisely $\binom{2n}{n}$ end at 0, since we need $n$ upward and $n$ downward moves, and there are this many ways to choose $n$ places for the upward moves among $2n$ steps. Therefore, for all $\varepsilon > 0$,

$$\mathbb{P}(\|B_{2n}\| > \varepsilon) \leq \mathbb{P}(B_{2n} = 1) = \mathbb{P}(S_{2n} = 0) = \frac{\binom{2n}{n}}{n} 2^{-2n} \sim \frac{1}{\sqrt{n \pi}} \to 0,$$

by Stirling’s formula. As $B_{2n+1} = 0$ almost surely, $B_n \to 0$, in probability, as $n \to \infty$.

(d) Recurrence of $(S_n)_{n \geq 0}$ means that $\{n \geq 0: B_n = 1\} = \{n \geq 0: S_n = 0\}$ is unbounded with probability 1. Hence, with probability 1, there is a subsequence $(N_m)_{m \geq 0}$ such that $B_{N_m} = 1$ for all $m \geq 0$ and $N_m \to \infty$. Hence, 1 is the only possible almost sure limit. Assuming that $B_n \to 1$ almost surely, this implies $B_n \to 1$ in probability, contradicting (a). Hence $(B_n)_{n \geq 0}$ does not converge almost surely.
3. First express $\mathbb{P}\left( \frac{X_t - t/\mu}{\sqrt{t\sigma^2/\mu^3}} \leq u \right) = \mathbb{P}\left( \frac{X_t}{t/\mu + u\sqrt{t\sigma^2/\mu^3}} \leq 1 \right)$. Then apply the hint to $n = n(t)$ where $n(t) = \frac{t/\mu + u\sqrt{t\sigma^2/\mu^3}}{1} + 1$ and $\lfloor . \rfloor$ denotes the integer part. Then
$$
\mathbb{P}(X_t \leq n(t) - 1) = \mathbb{P}(T_{n(t)} > t) = \mathbb{P}\left( \frac{T_{n(t)} - \mu(n(t))}{\sigma(n(t))} > \frac{t - \mu(n(t))}{\sigma(n(t))} \right)
$$
and as $t \to \infty$, clearly $n(t) \to \infty$, and the limit is $\mathbb{P}(Z \geq -u) = \mathbb{P}(Z \leq u)$ for a $Z \sim \text{Normal}(0, 1)$ since
$$
\frac{t - \mu(n(t))}{\sigma(n(t))} \sim \frac{t - \mu(t/\mu + u\sqrt{t\sigma^2/\mu^3})}{\sqrt{t/\mu + u\sqrt{t\sigma^2/\mu^3}}} = -u \frac{1}{\sqrt{1 + u\sqrt{\sigma^2/(t\mu)}}} \to -u.
$$

4. (a) We have to show that the $Z_m, m \geq 0$, are independent and identically distributed. First, by (positive) recurrence of $i$, $Z_m$ is well-defined for all $m \geq 0$. The strong Markov property of the continuous-time Markov chain $X$ at $H_i^{(m)}$ yields that the post-$H_i^{(m)}$ process is independent from the past and has the same distribution as $X$ given $X_0 = i$. Therefore, $Z_m$ which is the first passage time at $i$ of the post-$H_i^{(m)}$ process, is independent of $Z_0, \ldots, Z_{m-1}$ and has the same distribution as $Z_0$. By induction, $Z_m, m \geq 0$ are all independent:
$$
\mathbb{P}(Z_0 \in A_0, \ldots, Z_m \in A_m) = \mathbb{P}(Z_0 \in A_0, \ldots, Z_{m-1} \in A_{m-1})\mathbb{P}(Z_m \in A_m) = \cdots = \mathbb{P}(Z_0 \in A_0)\cdots\mathbb{P}(Z_m \in A_m).
$$

(b) For $j = i$, we apply the Strong Law of Large Numbers to get $S_m/m \to E(Z_1) = m_i$ almost surely, as $m \to \infty$. For $j \neq i$, given $X_0 = j$, $Z_0$ does not have the same distribution as $Z_k, k \geq 1$, but irreducibility and recurrence imply $Z_0 < \infty$, so
$$
\frac{S_m}{m} = \frac{Z_0}{m} + \frac{Z_1 + \cdots + Z_{m-1}}{m-1} \to m_i \quad \text{almost surely, as } m \to \infty.
$$

(c) Denote the holding times of $X$ in $i$ by $(R_n)_{n \geq 0}$, $R_n \sim \text{Exp}(\lambda_i)$, then, by the Strong Law of Large Numbers
$$
\frac{1}{n} U_n := \frac{1}{n} \left( R_0 1_{\{X_0 = i\}} + \sum_{k=1}^{n-1} R_k \right) \to E(R_1) = \frac{1}{\lambda_i}.
$$
Therefore, still with notation $S_n = H_i^{(n)}$,
$$
\frac{1}{H_i^{(n)}} \int_{0}^{1} 1_{\{X_s = i\}} ds = \frac{1}{S_n/n} \to \frac{1}{\lambda_i m_i} = \frac{1}{m_i \lambda_i} \quad \text{almost surely, as } n \to \infty.
$$
Denote $B_t = \int_{0}^{t} 1_{\{X_s = i\}} ds$ and $N_t = \#\{n \geq 1 : S_n \leq t\}$. Note that by the Strong Law of Renewal Theory $N_t/t \to 1/m_i$ almost surely. Now, by the same argument as in Question 3(d), $U_{N_t} \leq B_t < U_{N_t+1}$. Remember that $S_{N_t} \leq t < S_{N_t+1}$, and hence $\frac{U_{N_t}}{S_{N_t+1}} \leq \frac{B_t}{S_{N_t+1}}$, from which we see $\frac{1}{\lambda_i m_i} \leq \lim_{t \to \infty} \frac{B_t}{t} \leq \frac{1}{\lambda_i m_i}$, using $N_t \to \infty$ and $S_n/S_{n+1} \to 1$. The latter can be obtained from the law of large numbers again.

5. Denote the Poisson process by $(X_t)_{t \geq 0}$. Then $E_t$ is the first time $\tilde{X} = (X_{t-r} - X_t)_{s \geq 0}$ steps up. $\tilde{X}$ is also a Poisson process, by the Markov property. Hence $E_t \sim \text{Exp}(\lambda)$.
We have $\mathbb{P}(A_t = t) = \mathbb{P}(X_t = 0) = e^{-\lambda t}$ and $\mathbb{P}(A_t > r) = \mathbb{P}(X_t - X_{t-r} = 0) = e^{-\lambda r}$ for $r < t$. Hence, $A_t$ has a partly discrete, partly continuous distribution with $\mathbb{P}(A_t = t) = e^{-\lambda t}$ and density $f_{A_t}(r) = \lambda e^{-\lambda r}, 0 < r < t$.
We have $D_t = A_t + E_t$ with $\mathbb{E}(A_t) > 0$ for all $t > 0$, so $\mathbb{E}(D_t) > \mathbb{E}(E_t) = 1/\lambda$. 
6. Let \( X \sim \text{Poi}(\lambda) \). Then \( \mathbb{E}(X) = \lambda \). By definition, \( p^{sb}_0 = 0 \) and for \( n \geq 1 \)
\[
p^{sb}_n = \frac{n p_n}{\lambda} = \frac{\lambda^{n-1} e^{-\lambda}}{(n-1)!}.
\]

Therefore, for \( X^{sb} \) with probability function \( p^{sb} \), we get, for \( m \geq 0 \),
\[
\mathbb{P}(X^{sb} - 1 = m) = \mathbb{P}(X^{sb} = m + 1) = p^{sb}_{m+1} = \frac{\lambda^m e^{-\lambda}}{m!} = p_m,
\]
as required.

7. (a) First, for \( X \sim \text{Gamma}(a, \lambda) \), we have \( \mathbb{E}(X) = a/\lambda \), therefore
\[
f^{sb}(x) = \frac{\lambda x f(x)}{a} = \frac{\lambda^{a+1} x^a e^{-\lambda x}}{a \Gamma(a)} = \frac{\lambda^{a+1} x^a e^{-\lambda x}}{\Gamma(a+1)}
\]
by well-known properties of the Gamma function, or by identifying the density of the \( \text{Gamma}(a+1, \lambda) \) distribution up to a constant: since size-biasing always gives probability densities that integrate to 1, the last equality must hold.

(b) From the lectures, we know that for a stationary renewal process \( X \) the process of excess lifetimes \( E_t = T_{X_t+1} - t \) is stationary, and its stationary distribution is that of \( LU \), where \( L \sim f^{sb} \) and \( U \sim \text{Unif}((0,1)) \) are independent. Therefore
\[
m_{\text{stat}} = \mathbb{E}(E_t) = \mathbb{E}(LU) = \mathbb{E}(L)\mathbb{E}(U) = \frac{a+1}{\lambda} \frac{1}{2} = \frac{a+1}{2\lambda}.
\]

(c) Now the waiting time is a full inter-arrival time with \( \text{Gamma}(a, \lambda) \) distribution, so \( m_{\text{ren}} = a/\lambda \) and therefore
\[
m_{\text{stat}} > m_{\text{ren}} \iff \frac{a+1}{2\lambda} > \frac{a}{\lambda} \iff a+1 > 2a \iff a < 1.
\]

It is paradoxical that someone who is waiting for a full inter-arrival time is served earlier, on average, than someone who arrives in the middle of an inter-arrival time. The reason is the size-biasing effect: somebody who arrives in stationarity, picks a size-biased inter-arrival time, and this is typically a long one. Here, even half of this is still longer than a non-size-biased inter-arrival time.

8. (a) We insert \( r \) into the right hand side to get
\[
H + r \ast f = H + H \ast f + H \ast \sum_{k \geq 1} f^{* (k)} \ast f = H + H \ast \sum_{j \geq 1} f^{* (j)} = r.
\]

(b) As for the uniqueness, first note that boundedness of \( H \) on bounded intervals implies
\[
\sup_{0 \leq t \leq T} |r(t)| \leq \sup_{0 \leq t \leq T} |H(t)| + \sup_{0 \leq t \leq T} \left| \int_0^t H(t-x) m'(t) dt \right| \leq (1 + m(T)) \sup_{0 \leq t \leq T} |H(t)| < \infty,
\]
i.e. \( r \) is bounded on bounded intervals. Suppose that \( s \) is another solution that is bounded on bounded intervals, and \( \delta = r - s \), then \( \delta \) is also bounded on bounded intervals and solves \( \delta = \delta \ast f \). Inductively applying convolution properties, one easily sees that also \( \delta = \delta \ast f^{* (k)} \) for all \( k \geq 1 \) and hence
\[
|\delta(t)| \leq \int_0^t f^{* (k)}(s) ds \sup_{0 \leq u \leq t} |\delta(u)| = \mathbb{P}(T_k \leq t) \sup_{0 \leq u \leq t} |\delta(u)| \to 0 \quad \text{as} \quad k \to \infty,
\]
so that \( \delta \equiv 0 \) and hence \( \mu \equiv \nu \).
(c) By one-step analysis,
\[ \mathbb{P}(E_t > y) = \int_0^\infty \mathbb{P}(E_t > y|T_1 = x)f(x)dx \]
\[ = \mathbb{P}(T_1 > t + y) + \int_0^t \mathbb{P}(E_{t-x} > y)f(x)dx \]
is a renewal type equation with \( H(t) = \mathbb{P}(T_1 > t + y) \) and \( r(t) = \mathbb{P}(E_t > y) \). By (a) and (b), the unique solution is
\[ \mathbb{P}(E_t > y) = r(t) = H(t) + \int_0^t H(t-x)m'(x)dx \]
\[ = \mathbb{P}(T_1 > t + y) + \int_0^t \mathbb{P}(T_1 > t + y - x)m'(x)dx \]
as required.

(d) In (c) we have a formula of the form where the Key Renewal Theorem applies (non-negative, integrable, non-increasing \( g(x) = \mathbb{P}(T_1 > y + x) \) so that \( \mathbb{P}(E_t > y) = g(t) + \int_0^t g(t-x)m'(x)dx = 0 + \frac{1}{\mu} \int_0^\infty g(x)dx = \frac{1}{\mu} \int_y^\infty \mathbb{P}(T_1 > z)dz. \)

9. (i) First, \( E_n = T_{X_{n+1}} - n = m \geq 2 \) implies
\[ X_{n+1} = X_n \quad \text{and} \quad E_{n+1} = T_{X_{n+1}} - (n+1) = m - 1, \]
so \( \pi_{m,m-1} := \mathbb{P}(E_{n+1} = m-1|E_n = m) = 1 \) does not depend on \( n \). For \( m = 1 \), note that \( E_n = T_{X_{n+1}} - n = 1 \) implies
\[ X_{n+1} = X_n + 1 \quad \text{and} \quad E_{n+1} = T_{X_{n+1}} + X_n - (n+1) = X_n + 1 \]
so \( \pi_{1,j} := \mathbb{P}(E_{n+1} = j|E_n = 1) = \mathbb{P}(Z_{X_n+1} = j) = \mathbb{P}(Z_1 = j) =: p_j \) does not depend on \( n \); here we applied the renewal property at the first renewal time after \( n \). Remember that \( Z_{X_n} \) does not have the same distribution as \( Z_1 \), because of the size-biasing effect of the inter-renewal time containing \( n \).
To prove the claim that \( (E_n)_{n \geq 0} \) is a Markov chain with transition matrix \( \Pi = (\pi_{ij})_{i,j \in \{1,2,\ldots\}} \), note that clearly for all admissible paths, \( m \geq 2, \)
\[ \mathbb{P}(E_{n+1} = m-1|E_0 = i_0, \ldots, E_{n-1} = i_{n-1}, E_n = m) = 1 = \pi_{m,m-1}. \]
Also, for all \( (i_0, \ldots, i_{n-1}, 1) \) with \( \mathbb{P}(E_0 = i_0, \ldots, E_n = i_n) > 0 \), we have \( X_n = \#\{k = 0, \ldots, n-1 : E_k = 1\} \) and so for \( \ell = \#\{k = 0, \ldots, n-1 : i_k = 1\} \), we have
\[ \mathbb{P}(E_{n+1} = j|E_0 = i_0, \ldots, E_{n-1} = i_{n-1}, E_n = 1) = \mathbb{P}(Z_\ell = j) = \pi_{1,j}, \]
as required.

(ii) Solve \( \eta \Pi = \eta, \) i.e. for \( j \geq 1 \)
\[ \eta_1 p_j + \eta_{j+1} = \eta_j \Rightarrow \eta_{j+1} = \eta_j - \eta_1 p_j = \cdots = (1 - p_j - \cdots - p_1)\eta_1, \]
so that \( \eta_j = \mathbb{P}(Z_1 \geq j)\eta_1 \) and as, summed over \( j \geq 1 \), this must be 1, we get
\[ \frac{1}{\eta_1} = \sum_{j \geq 1} \mathbb{P}(Z_1 \geq j) = \mathbb{E}(Z_1) = \mu \Rightarrow \eta_j = \frac{1}{\mu} \mathbb{P}(Z_1 \geq j). \]
The claimed convergence holds by the discrete Convergence Theorem, as the chain is clearly irreducible and positive recurrent, aperiodic, because we assume that \( Z_1 \) is 1-arithmetic.
(iii) $S$ has probability function $\mu^{-1}kp_k$, and so

$$\mathbb{P}(U = j) = \sum_{k=j}^{\infty} \mathbb{P}(U = j|S = k) \mathbb{P}(S = k) = \sum_{k=j}^{\infty} \frac{1}{k} \mu^{-1}kp_k = \mu^{-1} \mathbb{P}(Z_1 \geq j).$$