Theorem 1 (Thm 5.1.32)**.** *If $(X_n, \mathcal{F}_n)$ is a sub-MG (or a sup-MG or a MG) and $\theta \le \tau$ are stopping times for $\{\mathcal{F}_n\}$, then $(X_{n\wedge\tau} - X_{n\wedge\theta}, \mathcal{F}_n)$ is also a sub-MG (or sup-MG or MG, respectively). In particular, taking $\theta = 0$ we have that $(X_{n\wedge\tau}, \mathcal{F}_n)$ is then a sub-MG (or sup-MG or MG, respectively).*

Theorem 2 (Coro 5.1.33)**.** *If $(X_n, \mathcal{F}_n)$ is a sub-MG and $\tau \ge \theta$ are $\mathcal{F}_n$-stopping times, then $\mathbb{E}X_{n\wedge\tau} \ge \mathbb{E}X_{n\wedge\theta}$ for all n. The reverse inequality holds in case $(X_n, \mathcal{F}_n)$ is a sup-MG, with $\mathbb{E}X_{n\wedge\theta} = \mathbb{E}X_{n\wedge\tau}$ for all n in case $(X_n, \mathcal{F}_n)$ is a MG.*

Theorem 3 (Thm 5.2.6 Doob's Inequality)**.** *For any sub-martingale $\{X_n\}$ and $x > 0$, let $\tau_x = \inf\{k \ge 0 : X_k \ge x\}$. Then, for any finite $n \ge 0$,*

$$\mathbb{P}\left(\max_{k=0}^{m} X_k \ge x\right) \le x^{-1}\mathbb{E}[X_n \mathbb{I}_{\{\tau_x \le n\}}] \le x^{-1}\mathbb{E}[(X_n)^+].$$

Theorem 4 (Thm 5.2.18 Doob's up-crossing)**.** *If $\{X_n\}$ is a sup-MG then*

$$(b-a)\mathbb{E}[U_n[a,b]] \le \mathbb{E}[(X_n - a)_-] - \mathbb{E}[(X_0 - a)_-] \quad \forall a < b.$$

where $U_n[a,b](\omega)$ is the of up-crossings of the interval $[a,b]$ by $\{X_k(\omega), k = 0, 1, \ldots, n\}$: the largest $l \in \mathbb{Z}^+$ such that $X_{s_i}(\omega) < a$ and $X_{t_i}(\omega) > b$ for $1 \le i \le l$ and some $0 \le s_1 < t_1 < \ldots < s_l < t_l \le n$.

Theorem 5 (Thm 5.3.2 Doobs' convergence)**.** *Suppose sup-MG $(X_n, \mathcal{F}_n)$ is such that $\sup_n \mathbb{E}[(X_n)_-] < \infty$. Then, $X_n \xrightarrow{a.s.} X_\infty$ and $\mathbb{E}|X_\infty| \le \liminf_n \mathbb{E}|X_n|$ is finite.*

And we have the following equivalent conditions (Exercise 5.3.3)

- $\lim_n \mathbb{E}|X_n|$ *exists and is finite.*
- $\sup_n \mathbb{E}|X_n| < \infty$.
- $\liminf_n \mathbb{E}|X_n| < \infty$.
- $\lim_n \mathbb{E}(X_n)_+$ *exists and is finite.*
- $\sup_n \mathbb{E}(X_n)_+ < \infty$.

Theorem 6 (Prop 5.3.5)**.** *Suppose $\{X_n\}$ is a martingale of uniformly bounded differences. That is, almost surely $\sup_n |X_n - X_{n-1}| \le c$ for some finite non-random constant c. Then, $\mathbb{P}(A \cup B) = 1$ for the events*

$$A = \left\{\omega : \lim_n X_n(\omega) \text{ exists and is finite}\right\}, \quad B = \left\{\omega : \limsup_n X_n(\omega) = \infty \& \liminf_n X_n(\omega) = -\infty\right\}.$$

Theorem 7 (Prop 5.3.8)**.** *Suppose $(X_n, \mathcal{F}_n)$ is a non-negative sup-MG and $\tau \ge \theta$ are stopping times for the filtration $\{\mathcal{F}_n\}$. Then, $\mathbb{E}X_\theta \ge \mathbb{E}X_\tau$ are finite valued.*

Definition 1 (Defi 1.3.47 U.I.)**.** *A possibly uncountable collection of random variables $\{X_\alpha, \alpha \in I\}$ is called uniformly integrable (U.I.) if*

$$\lim_{M\to\infty} \sup_\alpha \mathbb{E}[|X_\alpha| \mathbb{I}_{|X_\alpha|>M}] = 0.$$

Theorem 8 (Thm 5.3.12)**.** *If $(X_n, \mathcal{F}_n)$ is a sub-MG, then $\{X_n\}$ is U.I, if and only if $X_n \xrightarrow{L^1} X_\infty$, in which case also $X_n \xrightarrow{a.s.} X_\infty$ and $X_n \le \mathbb{E}[X_\infty|\mathcal{F}_n]$ for all n.*

Definition 2 (Def 5.3.13 Doob's martingale)**.** *The sequence $X_n = \mathbb{E}[X|\mathcal{F}_n]$ with X an integrable R.V. and $\{\mathcal{F}_n\}$ a filtration, is called Doob's martingale of X with respect to $\{\mathcal{F}_n\}$.*

Theorem 9 (Prop 5.3.14)**.** *A martingale $(X_n, \mathcal{F}_n)$ is U.I. if and only if $X_n = \mathbb{E}[X_\infty|\mathcal{F}_n]$ is a Doob's martingale with respect to $\{\mathcal{F}_n\}$, or equivalently if and only if $X_n \xrightarrow{L^1} X_\infty$.*

Theorem 10 (Thm 5.3.15 Lévy's Upward Theorem)**.** *Suppose $\sup_m |X_m|$ is integrable, $X_n \xrightarrow{a.s.} X_\infty$ and $\mathcal{F}_n \uparrow \mathcal{F}_\infty$. Then $\mathbb{E}[X_n|\mathcal{F}_n] \xrightarrow{a.s.} \mathbb{E}[X_\infty|\mathcal{F}_\infty]$ both a.s. and in L^1.*

Theorem 11 (Coro 5.3.16 Lévy's 0-1)**.** *If $\mathcal{F}_n \uparrow \mathcal{F}_\infty$, $A \in \mathcal{F}_\infty$, then $\mathbb{E}[\mathbb{I}_A|\mathcal{F}_n] \xrightarrow{a.s.} \mathbb{I}_A$.*

Theorem 12 (Prop 5.3.22 Doob's L_p M.G. convergence)**.** *If the MG $\{X_n\}$ is such that $\sup_n \mathbb{E}|X_n|^p < \infty$ for some $p > 1$, then there exists a R.V. X_∞ such that $X_n \xrightarrow[L_p]{\text{a.s.}} X_\infty$ (so $\|X_n\|_p \to \|X_\infty\|_p$).*

Theorem 13 (Thm 5.4.1 Doob's Optional Stopping)**.** *Suppose $\theta \le \tau$ are $\mathcal{F}_n$-stopping times and $X_n = Y_n + V_n$ for sub-MGs $(V_n, \mathcal{F}_n)$, $(Y_n, \mathcal{F}_n)$ such that V_n is non-positive and $\{Y_{n\wedge\tau}\}$ is uniformly integrable.*

Then, the R.V. X_θ and X_τ are integrable and $\mathbb{E}X_\tau \ge \mathbb{E}X_\theta \ge \mathbb{E}X_0$ (where $X_\tau(\omega)$ and $X_\theta(\omega)$ are set as $\limsup_n X_n(\omega)$ in case the corresponding stopping time is infinite).

And we have the following equivalent conditions for $\{Y_{n\wedge\tau}\}$ being U.I. (Prop 5.4.4)

- *$\mathbb{E}\tau < \infty$ and a.s. $\mathbb{E}[|Y_n - Y_{n-1}||\mathcal{F}_{n-1}] \le c$ for some finite, non-random c.*
- *$\{Y_n \mathbb{I}_{\tau>n}\}$ is uniformly integrable and $Y_\tau \mathbb{I}_{\tau<\infty}$ is integrable.*
- *$(Y_n, \mathcal{F}_n)$ is a uniformly integrable sub-MG (or sup-MG).*

Exercise 1 (Exer 5.4.6-5.4.7)**.** *(5.4.6) Show that if $\{X_n\}$ is a sub-martingale such that $\mathbb{E}X_0 \ge 0$ and $\inf_n X_n < 0$ a.s. then necessarily $\mathbb{E}[\sup_n X_n] = \infty$.*

(5.4.7) Fixing $b > 0$, let $\tau_b = \inf\{n \ge 0 : S_n \ge b\}$ for the random walk $\{S_n\}$ of Definition 5.1.6 and suppose $\xi_n = S_n - S_{n-1}$ are uniformly bounded, of zero mean and positive variance.

- *Show that τ_b is almost surely finite.*
- *Show that $\mathbb{E}[\min\{S_n : n \le \tau_b\}] = -\infty$.*

Proof. We first prove (5.4.6): Since max is a convex function, we have that $Y_n := \max\{X_n, -1\}$ is still a sub-M.G. Now assume that $\mathbb{E}[\sup |Y_n|] < \infty$ i.e. Y_n is integrable. We consider the stopping time $\tau := \inf\{n : Y_n < 0\}$. Since $\inf_n X_n < 0$ a.s., we have $\tau < \infty$ a.s., for which $Y_\tau < 0$, and thus we further have by Doob's optional stopping theorem that

$$0 > \mathbb{E}[Y_\tau] \ge \mathbb{E}[Y_0] \ge 0$$

which is a contradiction. Thus we have that $\mathbb{E}[\sup |Y_n|] = \infty$. Now since $\sup |Y_n| = \max\{\sup Y_+, \sup Y_-\}$ while $\sup Y_- \le 1$, we have

$$\begin{aligned}\infty = \mathbb{E}[\sup |Y_n|] &\le \mathbb{E}[\max\{\sup \max\{X_n, -1\}_+, 1\}] \\ &\le \mathbb{E}[\max\{1, \sup X_n\}]\end{aligned}$$

for this to hold, we must have $\mathbb{E}[\sup X_n] = \infty$. Thus we have proved (5.4.6):

- For random walk, we know that we have $S_n/\sqrt{n} \xrightarrow{\text{d}} N(0,1)$ which is a non-degenerate distribution. Thus we have that $\mathbb{P}(\lim_n S_n \text{ exists}) = 0$, because for the event $\{\lim_n S_n \text{ exists}\}$, we must have that

$S_n(\omega)/\sqrt{n} \to 0$. On the other hand, for such M.G. with bounded difference, by (prop 5.3.5) we have that

$$\lim_n S_n \text{ exists }, \textit{or} \liminf_n S_n = -\infty, \limsup_n S_n = \infty, \quad \text{a.s.}$$

and from the above argument we have that the first case is w.p. 0, thus we have that $\limsup_n S_n = \infty \& \liminf_n S_n = -\infty$ a.s., which means that $\tau_b < \infty$ a.s.

- Note that we have $S_{\tau_b} \geq b > 0$, i.e. $\sup S_{n\wedge\tau_b} > 0$. And since τ_b is a stopping time, we also have that $S_{n\wedge b}$ is a M.G. (with $S_{0\wedge\tau_b} = 0$), thus by the lemma we have that $\mathbb{E}\left[\min\{S_n : n \leq \tau_b\}\right] = \mathbb{E}\left[\inf_n S_{n\wedge\tau_b}\right] = \infty$.

$\square$

Theorem 14 (Coro 5.4.8 Gambler's Ruin). *Fixing positive integers a and b the probability that a SRW $\{S_n\}$, starting at $S_0 = 0$, hits $-a$ before first hitting $+b$ is $r = (e^{\lambda b} - 1)/(e^{\lambda b} - e^{-\lambda a})$ for $\lambda = \log[(1-p)/p] \neq 0$. For the symmetric SRW, i.e. when $p = 1/2$, this probability is $r = b/(a+b)$.*

Definition 3 (Def 6.1.1 Markov Chain). *Given a filtration $\{\mathcal{F}_n\}$, an $\mathcal{F}_n$-adapted stochastic process $\{X_n\}$ taking values in a measurable space $(\mathbb{S}, \mathcal{S})$ is called an $\mathcal{F}_n$-Markov chain with state space $(\mathbb{S}, \mathcal{S})$ if for any $A \in \mathcal{S}$,*

$$\mathbb{P}[X_{n+1} \in A | \mathcal{F}_n] = \mathbb{P}[X_{n+1} \in A | X_n] \quad \forall n, \quad a.s.$$

Theorem 15 (Prop 6.1.16 Strong Markov Property). *Fix a homogeneous $\mathcal{F}_n$-Markov chain $\{X_n\}$ with transition probabilities $p(\cdot,\cdot)$. Identifying via $X_n(\omega) \mapsto \omega_n$ the restriction of $\mathbb{P}$ to $\mathcal{F}_X = \sigma(X_k, k \geq 1)$ with the probability space $(S_\infty, \mathcal{S}_c, \mathbb{P}_\nu)$, set the shift operator $\theta : S_\infty \to S_\infty$ such that $(\theta\omega)_k = \omega_{k+1}$ for all $k \geq 0$ (with the corresponding iterates $(\theta^n\omega)_k = \omega_{k+n}$ for $k, n \geq 0$). Then, for any $\{h_n\} \subseteq \mathcal{F}_X$ with $\sup_{n,\omega} |h_n(\omega)|$ finite, and any $\mathcal{F}_n$-stopping time τ,*

$$\mathbb{E}[h_\tau(\theta_\tau\omega) | \mathcal{F}_\tau] \mathbb{I}_{\{\tau < \infty\}} = \mathbb{E}[h_\tau] \mathbb{I}_{\{\tau < \infty\}}.$$

And in the case of $\tau = n$ and $h_k = h$ we have Markov property

$$\mathbb{E}[h(\theta^n\omega) | \mathcal{F}_n] = \mathbb{E}_{X_n}[h].$$

Exercise 2 (Exer 6.1.18). *Consider a homogeneous Markov chain $\{X_n\}$ with B-isomorphic state space $(\mathbb{S}, \mathcal{S})$. Fixing $\{B_l\} \subseteq \mathcal{S}$, let $\Gamma_n = \bigcup_{l>n}\{X_l \in B_l\}$ and $\Gamma = \{X_l \in B_l \ i.o.\}$.*

- *Using the Markov property and Lévy's upward theorem (Theorem 5.3.15), show that $\mathbb{P}(\Gamma_n | X_n) \xrightarrow{a.s.} \mathbb{I}_\Gamma$.*
- *Show that $\mathbb{P}(\{X_n \in A_n \ i.o.\} \cap \Gamma) = 0$ for any $\{A_n\} \subseteq \mathcal{S}$ such that for some $\eta > 0$ and all n, with probability one, $\mathbb{P}(\Gamma_n | X_n) \geq \eta \mathbb{I}_{\{X_n \in A_n\}}$.*
- *Suppose $A, B \in \mathcal{S}$ are such that $\mathbb{P}_x(X_l \in B$ for some $l \geq 1) \geq \eta$ for some $\eta > 0$ and all $x \in A$. Deduce that*

$$\mathbb{P}(\{X_n \in A \textit{ finitely often}\} \cup \{X_n \in B \ i.o.\}) = 1.$$

Proof.

- By property of sets we have $\Gamma_n \to \Gamma$ thus $\mathbf{1}_{\Gamma_n} \xrightarrow{\text{a.s.}} \mathbf{1}_\Gamma$, then by Lévy's upward theorem we have:

$$\mathbb{P}\left(\Gamma_n | X_n\right) = \mathbb{E}\left[\mathbf{1}_{\Gamma_n} | X_n\right] = \mathbb{E}\left[\mathbf{1}_{\Gamma_n} | \mathcal{F}_n\right] \xrightarrow[L_1]{\text{a.s.}} \mathbb{E}\left[\mathbf{1}_\Gamma | \mathcal{F}_\infty\right] = \mathbf{1}_\Gamma$$

- Denote $K := \{\omega : X_n(\omega) \in A_n \, i.o.\}$. Then we have that $\forall N > 0$, $\exists n > N$ s.t. $\mathbb{P}(\Gamma_n \cap K | X_n) \geq \eta > 0$. On the other hand we have

$$\eta < \mathbb{P}(\Gamma_n \cap K | X_n) \xrightarrow{\text{a.s.}} \mathbb{P}(\Gamma \cap K | X_\infty) = \mathbf{1}_{\Gamma \cap K} = 1 = \mathbb{P}(\Gamma \cap K)$$

which gives $\mathbb{P}(K \backslash \Gamma) = 0$. In the above we applied Lévy's upward theorem to $\Gamma_n \cap K | X_n$.

- Use $A_n \equiv A$ and $B_n \equiv B$ and we have using the precedence:

$$\begin{aligned} 1 \leq & \mathbb{P}\left(\{X_n \in A \text{ finitely often}\} \cup (\{X_n \in A \, i.o.\} \backslash \Gamma) \cup \Gamma\right) \\ \leq & \mathbb{P}\left(\{X_n \in A \text{ finitely often}\} \cup \Gamma\right) + \mathbb{P}\left(\{X_n \in A \, i.o.\} \backslash \Gamma\right) \\ = & \mathbb{P}\left(\{X_n \in A \text{ finitely often}\} \cup \Gamma\right) + 0 \end{aligned}$$

where $\Gamma = \{X_n \in B \, i.o.\}$ so thus we have proved the claim.

□

Theorem 16 (Prop 6.2.1 Chapman-Kolmogorov)**.** *For any $x, y \in \mathbb{S}$ and non-negative integers $k \leq n$,*

$$\mathbb{P}_x(X_n = y) = \sum_{z \in \mathbb{S}} \mathbb{P}_x(X_k = z) \mathbb{P}_z(X_{n-k} = y).$$

Proof. Using the canonical construction of the chain whereby $X_n(\omega) = \omega_n$, we combine the tower property with the Markov property for $h(\omega) = \mathbb{I}_{\{\omega_{n-k} = y\}}$ followed by a decomposition according to the value z of X_k to get that

$$\mathbb{P}_x(X_n = y) = \mathbb{E}_x\left[h(\theta^k \omega)\right] = \mathbb{E}_x\left[\mathbb{E}_x\left[h(\theta^k \omega) | \mathbf{F}_k\right]\right] = \mathbb{E}_x\left[\mathbb{E}_{X_k}[h]\right] = \sum_{z \in \mathbb{S}} \mathbb{P}_x(X_k = z) \mathbb{P}_z(X_{n-k} = y).$$

This concludes the proof as $\mathbb{E}_z(h) = \mathbb{P}_z(X_{n-k} = y)$.

□

Definition 4 (Def 5.1.25 Harmonic)**.** *A lower semi-continuous function $f : \mathbb{R}^d \to \mathbb{R}$ is superharmonic if for any x and $r > 0$,*

$$f(x) \geq \frac{1}{|B(0,r)|} \int_{B(x,r)} f(y) dy.$$

Definition 5 (Def 6.2.4 Harmonic)**.** *Extending Definition 5.1.25 we say that $f : \mathbb{S} \to \mathbb{R}$ which is either bounded below or bounded above is super-harmonic for the transition probability $p(x,y)$ at $x \in \mathbb{S}$ when $f(x) \geq \sum_{y \in \mathbb{S}} p(x,y) f(y)$. Likewise, $f(\cdot)$ is sub-harmonic at x when this inequality is reversed and harmonic at x in case an equality holds. Such a function is called super-harmonic (or sub-harmonic, harmonic, respectively) for $p(\cdot,\cdot)$ (or for the corresponding chain $\{X_n\}$), if it is super-harmonic (or, subharmonic, harmonic, respectively), at all $x \in \mathbb{S}$. Equivalently, $f(\cdot)$ which is either bounded below or bounded above is harmonic provided $\{f(X_n)\}$ is a martingale whenever the initial distribution of the chain is such that $f(X_0)$ is integrable. Similarly, $f(\cdot)$ bounded below is super-harmonic if $\{f(X_n)\}$ is a super-martingale whenever $f(X_0)$ is integrable.*

Exercise 3 (Exer 6.2.5)**.** *Suppose $\mathbb{S} \backslash C$ is finite, $\inf_{x \notin C} \mathbb{P}_x(\tau_C < \infty) > 0$ and $A \subset C$, $B = C \backslash A$ are both non-empty.*

(a) *Show that there exist $N < \infty$ and $\epsilon > 0$ such that $\mathbb{P}_y(\tau_C > kN) \leq (1-\epsilon)^k$ for all $k \geq 1$ and $y \in \mathbb{S}$.*
(b) *Show that $g(x) = \mathbb{P}_x(\tau_A < \tau_B)$ is harmonic at every $x \notin C$.*
(c) *Show that if a bounded function $g(\cdot)$ is harmonic at every $x \notin C$ then $g(X_{n\wedge\tau_C})$ is a martingale.*
(d) *Deduce that $g(x) = \mathbb{P}_x(\tau_A < \tau_B)$ is the only bounded function harmonic at every $x \notin C$ for which $g(x) = 1$ when $x \in A$ and $g(x) = 0$ when $x \in B$.*
(e) *Show that if $f : \mathbb{S} \to \mathbb{R}^+$ satisfies $f(x) = 1 + \sum_{y\in\mathbb{S}} p(x,y)f(y)$ at every $x \notin C$ then $M_n := n\wedge\tau_C + f(X_{n\wedge\tau_C})$ is a martingale, provided $\mathbb{P}(X_0 \in C) = 0$. Deduce that if in addition $f(x) = 0$ for $x \in C$ then $f(x) = \mathbb{E}_x[\tau_C]$ for all $x \in \mathbb{S}$.*

Proof. (c) By the harmonic property we have that

$$\begin{aligned}
\mathbb{E}\left[g(X_{n\wedge\tau_C+1})|\mathcal{F}_n\right] =&\mathbb{E}\left[g(X_{n\wedge\tau_C+1})(\mathbf{1}_{\tau_C\leq n} + \mathbf{1}_{\tau_C>n})|\mathcal{F}_n\right] \\
=&g(X_{n\wedge\tau_C})\mathbf{1}_{\tau_C\leq n} + \mathbb{E}\left[g(X_{n\wedge\tau_C+1})\mathbf{1}_{\tau_C>n}|\mathcal{F}_n\right] \\
=&g(X_{n\wedge\tau_C})\mathbf{1}_{\tau_C\leq n} + \mathbb{E}\left[g(\theta^1 X_{n\wedge\tau_C})\mathbf{1}_{\tau_C>n}|\mathcal{F}_n\right] \\
\overset{\text{SMP}}{=}&g(X_{n\wedge\tau_C})\mathbf{1}_{\tau_C\leq n} + \mathbb{E}_{X_{n\wedge\tau_C}=x\notin C}\left[g(X_{n\wedge\tau_C+1})\right]\mathbf{1}_{\tau_C>n} \\
=&g(X_{n\wedge\tau_C})\mathbf{1}_{\tau_C\leq n} + g(X_{n\wedge\tau_C})\mathbf{1}_{\tau_C>n} \\
=&g(X_{n\wedge\tau_C})
\end{aligned}$$

thus we have $g(X_{n\wedge\tau_C})$ is a M.G.

(e) Similar to (c) we prove the following:

$$\begin{aligned}
\mathbb{E}\left[M_{n+1}|\mathcal{F}_n\right] =&\mathbb{E}\left[M_{n+1}\mathbf{1}_{\tau_C\leq n}|\mathcal{F}_n\right] + \mathbb{E}\left[M_{n+1}\mathbf{1}_{\tau_C>n}|\mathcal{F}_n\right] \\
=&M_n\mathbf{1}_{\tau_C\leq n} + \mathbb{E}\left[M_{n+1}\mathbf{1}_{\tau_C>n}|\mathcal{F}_n\right] \\
=&M_n\mathbf{1}_{\tau_C\leq n} + (n+1+\mathbb{E}\left[f(X_{n+1})|\mathcal{F}_n\right])\mathbf{1}_{\tau_C>n} \\
=&M_n\mathbf{1}_{\tau_C\leq n} + (n+1+\mathbb{E}_{X_n=x\notin C}\left[f(X_{n+1})\right])\mathbf{1}_{\tau_C>n} \\
=&M_n\mathbf{1}_{\tau_C\leq n} + (n+1+f(X_n)-1)\mathbf{1}_{\tau_C>n} \\
=&M_n
\end{aligned}$$

thus M_n is a M.G. On the other hand we notice that

$$\mathbb{E}\left[M_n\right] =\mathbb{E}\left[n\wedge\tau_C + f(X_{n\wedge\tau_C})\right] \leq \mathbb{E}\left[\tau_C\right] + \sum_{x\in\mathbb{S}} f(x)\mathbb{P}\left(X_{n\wedge\tau_C} = x\right) \ \leq \mathbb{E}\left[\tau_C\right] + \sum_{x\in\mathbb{S}\setminus C} f(x) + \sum_{x\in C} 0 < \infty$$

then by DCT we have that

$$\mathbb{E}_x\left[\tau_C\right] = \mathbb{E}\left[\lim_{n\to\infty} n\wedge\tau_C + f(X_{n\wedge\tau_C})\right] = \lim_{n\to\infty}\mathbb{E}\left[n\wedge\tau_C + f(X_{n\wedge\tau_C})\right] = \mathbb{E}_x\left[0 + f(X_{0\wedge\tau_C})\right] = f(x)$$

□

Definition 6 (Def 6.2.7, 6.2.9)**.** $\rho_{x,y} := \mathbb{P}_x\left(T_y < \infty\right)$. *Call A state $y \in \mathbb{S}$ is called **recurrent** (or persistent) if $\rho_{yy} = 1$ and **transient** if $\rho_{yy} < 1$.*

*State y is said to be accessible from state $x \neq y$ if $\rho_{xy} > 0$ (or alternatively, we then say that x leads to y). Two states $x \neq y$, each accessible to the other, are said to intercommunicate, denoted by $x \leftrightarrow y$. A non-empty collection of states $C \subseteq \mathbb{S}$ is called **irreducible** if each two states in C intercommunicate, and **closed** if there is no $y \notin C$ and $x \in C$ such that y is accessible from x.*

Theorem 17 (Prop 6.2.10)**.** *With $T_y^0 = 0$, let $T_y^k = \inf\{n > T_y^{k-1} : X_n = y\}$ for $k \geq 1$ denote the time of the k-th return to state $y \in \mathbb{S}$ (so $T_y^1 = T_y > 0$ regardless of X_0). Then, for any $x, y \in \mathbb{S}$ and $k \geq 1$,*

$$\mathbb{P}_x\left(T_y^k < \infty\right) = \rho_{xy}\rho_y^{k-1}.$$

Further, let $N_\infty(y)$ denote the number of visits to state y by the Markov chain at positive times. Then, $\mathbb{E}_x[N_\infty(y)] = \rho_{xy}/(1-\rho_{yy})$ is positive if and only if $\rho_{xy} > 0$, in which case it is finite when y is transient and infinite when y is recurrent.

Theorem 18 (Coro 6.2.12)**.** *The following are equivalent for a state y being recurrent:*

- $\rho_{yy} = 1$.
- $\mathbb{P}_y(T_y^k < \infty) = 1$ *for all k.*
- $\mathbb{P}_y(X_n = y, i.o.) = 1$.
- $\mathbb{P}_y(N_\infty(y) = \infty) = 1$.
- $\mathbb{E}_y[N_\infty(y)] = \infty$.

Theorem 19 (Prop 6.2.15)**.** *If F is a finite set of transient states then for any initial distribution $\mathbb{P}_\nu(X_n \in F\ i.o.) = 0$. Hence, any finite closed set C contains at least one recurrent state, and if C is also irreducible then C is recurrent.*

Theorem 20 (Prop 6.2.21)**.** *Suppose $\mathbb{S}$ is irreducible for a chain $\{X_n\}$ and there exists $h : \mathbb{S} \to [0, \infty)$ of finite level sets $G_r = \{x : h(x) < r\}$ that is super-harmonic at $\mathbb{S} \backslash G_r$ for this chain and some finite r. Then, the chain $\{X_n\}$ is recurrent.*

f-divergence quantifies the difference between a pair of distributions over a measurable space $(\mathcal{X}, \mathcal{F})$. A formal definition is as follows:

Definition 7 (f-divergence)**.** *Let P and Q be two probability distributions on $\mathcal{X}$. Then for any convex function $f : (0, \infty) \to \mathbb{R}$ such that it is strictly convex at 1 and $f(1) = 0$, the f-divergence of P from Q with $Q \ll P$ is defined as*

$$D_f(Q\|P) = \mathbb{E}_P[f(\frac{\mathrm{d}Q}{\mathrm{d}P})],$$

where $\dfrac{\mathrm{d}Q}{\mathrm{d}P}$ is the Radon-Nikodym derivative of Q with respect to P, whenever $Q \ll P$. And in the case that $\mathcal{X}$ is discrete, we use the notation $D_f(Q\|P) = \sum_{x\in\mathcal{X}} P(x) f(\dfrac{Q(x)}{P(x)})$.

Some frequently used f functions and the corresponding divergences are as follows:

- **(KL-divergence)** $f(t) = t\log t$;
$$D(Q\|P) := \mathbb{E}_P\left[\frac{Q}{P}\log\frac{Q}{P}\right] = \mathbb{E}_Q\left[\log\frac{P}{Q}\right].$$
- **(Total variation)** $f(t) = \frac{1}{2}|t-1|$;
$$d_{\mathrm{TV}}(P, Q) := \frac{1}{2}\mathbb{E}_P\left[\left|\frac{Q}{P} - 1\right|\right] = \frac{1}{2}\int |\,\mathrm{d}Q - \,\mathrm{d}P|$$
- **(χ^2-divergence)** $f(t) = (t-1)^2$
$$\chi^2(Q\|P) := \mathbb{E}_P\left[(\frac{Q}{P} - 1)^2\right] = \int \frac{P^2}{Q} - 1$$