Kipnis-Landim Appendix 1

Pedro Cardoso

Instituto Superior Técnico

5 de março de 2020

Summary

Hereafter, we always use the convention

$$0log(0) = \lim_{x \to 0^+} xlog(x) = \lim_{x \to 0^+} \frac{log(x)}{\frac{1}{x}} = \lim_{x \to 0^+} \frac{\frac{1}{x}}{-\frac{1}{x^2}} == \lim_{x \to 0^+} -x = 0.$$

Let π be a reference probability measure on E. For a probability measure μ denote by $H(\mu|\pi)$ the relative entropy f μ with respect to π defined by the variational formula:

$$H(\mu|\pi) = \sup_f \{ < \mu, f > - \log(< \pi, e^f >) \}.$$

In this formula the supremum is carried over all bounded functions f and $< \mu, f >$ stands for the integral of f with respect to μ. From now on, to keep notation and terminology simple, we denote $H(\mu|\pi)$ by $H(\mu)$ and refer to it as the entropy of μ.
Notice that the addition of a constant to the function f does not change the value of $< \mu, f > - \log(< \pi, e^f >)$. We may therefore restrict the supremum to bounded positive functions.

Proposition

The entropy is non-negative, convex and lower semicontinuous.

Let $c \in \mathbb{R}, f_1 : E \to \mathbb{R}$ be such that $f_1(x) = c, \forall x \in E$. Then f_1 is bounded and

$$< \mu, f_1 > - \log(< \pi, e^{f_1} >) = \sum_{x \in E} \mu(x) f_1(x) - \log\left(\sum_{x \in E} \pi(x) e^{f_1(x)}\right)$$

$$= \sum_{x \in E} \mu(x) \cdot c - \log\left(\sum_{x \in E} \pi(x) e^{c}\right)$$

$$= c \sum_{x \in E} \mu(x) - log\left(e^{c} \sum_{x \in E} \pi(x)\right) = c - log(e^{c}) = c - c = 0.$$

Therefore,

$$H(\mu) = \sup_{f}\{< \mu, f > - \log(< \pi, e^{f} >)\} \geq < \mu, f_1 > - \log(< \pi, e^{f_1} >) = 0$$

and we have that the entropy is non-negative.

Let $\alpha \in [0,1]$ and let μ_1, μ_2 be probability measures on E. Then,

$$H(\alpha\mu_1 + (1-\alpha)\mu_2) = \sup_f \{ < \alpha\mu_1 + (1-\alpha)\mu_2, f > - \log(< \pi, e^f >)\}$$

$$= \sup_f \{\alpha\big[< \mu_1, f > - \log(< \pi, e^f >)\big] + (1-\alpha)\big[< \mu_1, f > - \log(< \pi, e^f >)\big]\}$$

$$\leq \sup_f \{\alpha\big[< \mu_1, f > - \log(< \pi, e^f >)\big]\}$$

$$+ \sup_f \{(1-\alpha)\big[< \mu_1, f > - \log(< \pi, e^f >)\big]\}$$

$$= \alpha \sup_f \{\big[< \mu_1, f > - \log(< \pi, e^f >)\big]\}$$

$$+ (1-\alpha) \sup_f \{\big[< \mu_1, f > - \log(< \pi, e^f >)\big]\}$$

$$= \alpha H(\mu_1) + (1-\alpha) H(\mu_2).$$

Since for every μ_1, μ_2 probability measures on E, we have

$$H(\alpha\mu_1 + (1-\alpha)\mu_2) \leq \alpha H(\mu_1) + (1-\alpha) H(\mu_2),$$

we have that the entropy is convex.

Let $(\mu_n)_{n\in\mathbb{N}}$ a sequence of probability measures on E which converges weakly to μ, which is a probability measure on E. Assume that $H(\mu) > \liminf_{n\to\infty} H(\mu_n)$. In this case, choose

$$\varepsilon = \frac{H(\mu) - \liminf_{n\to\infty} H(\mu_n)}{3} > 0.$$

Since $H(\mu) = \sup_f \{< \mu, f > - \log(< \pi, e^f >)\}$ over all bounded functions, there exists f_0 such that f_0 is a bounded function and

$$H(\mu) < \; < \mu, f_0 > - \log(< \pi, e^{f_0} >) + \varepsilon.$$

Since μ_n converges weakly to μ, there is $n_0 \in \mathbb{N}$ such that

$$< \mu, f_0 > < \ < \mu_n, f_0 > + \varepsilon, \forall n > n_0,$$

which is the same as

$$< \mu, f_0 > - \log(< \pi, e^{f_0} >) + \varepsilon < \ < \mu_n, f_0 > - \log(< \pi, e^{f_0} >) + 2\varepsilon, \forall n > n_0,$$

and leads to

$$H(\mu) < \ < \mu_n, f_0 > - \log(< \pi, e^{f_0} >) + 2\varepsilon, \forall n > n_0.$$

Taking the supremum over all bounded positive functions bounded below by a strictly positive constant, we get

$$\begin{aligned} H(\mu) \leq & \sup_f \{ < \mu_n, f > - \log(< \pi, e^f >) + 2\varepsilon \big] \} \\ = & 2\varepsilon + \sup_f \{ < \mu_n, f > - \log(< \pi, e^f >) \big] \} = 2\varepsilon + H(\mu_n), \forall n > n_0. \end{aligned}$$

Taking the $\liminf$ above, we get

$$H(\mu) \le 2\varepsilon + \liminf_{n\to\infty} H(\mu_n) = 2\varepsilon + H(\mu) - 3\varepsilon = H(\mu) - \varepsilon < H(\mu).$$

Therefore, the assumption that $H(\mu) > \liminf_{n\to\infty} H(\mu_n)$ is false and we have

$$H(\mu) \le \liminf_{n\to\infty} H(\mu_n).$$

Since $H(\mu) \le \liminf_{n\to\infty} H(\mu_n)$ for every sequence $(\mu_n)_{n\in\mathbb{N}}$ of probability measures on E such that $(f_n)_{n\in\mathbb{N}}$ converges weakly toμ, we have that the entropy is lower semicontinuous.

We shall repeatedly use the entropy to estimate the expectation of a function with respect to a probability measure μ in terms of integrals with respect to the reference measure π. Indeed, for every positive constant α and for every bounded function $f : E \to \mathbb{R}$, the entropy inequality gives that

$$H(\mu) \geq < \mu, \alpha f > - \log(< \pi, e^{\alpha f} >),$$

which is the same as

$$\alpha < \mu, f > = < \mu, \alpha f > \leq H(\mu) + \log(< \pi, e^{\alpha f} >),$$

which leads to

$$< \mu, f > \leq \alpha^{-1}\{\log(< \pi, e^{\alpha f} >) + H(\mu)\}$$

For indicator functions this inequality takes a simple form.

Proposition

Let A be a subset of E such that $\pi[A] > 0$. *Then*

$$\mu[A] \leq \frac{\log 2 + H(\mu)}{\log(1 + \frac{1}{\pi[A]})}.$$

Choose $f = \mathbb{1}_A$. Then we have

$$< \mu, f >=< \mu, \mathbb{1}_A >= \mu[A]$$

and for every $\alpha > 0$

$$\begin{aligned} < \pi, e^{\alpha f} >= < \pi, e^{\alpha \mathbb{1}_A} >= \pi[A]e^{\alpha \cdot 1} + \pi[A^C]e^{\alpha \cdot 0} \\ =\pi[A]e^{\alpha} + (1 - \pi[A]) \cdot 1 = \pi[A](e^{\alpha} - 1) + 1. \end{aligned}$$

Then

$$< \mu, f > \leq \alpha^{-1}\{\log(< \pi, e^{\alpha f} >) + H(\mu)\}$$

leads to

$$\mu[A] \leq \frac{\log(\pi[A](e^{\alpha} - 1) + 1) + H(\mu)}{\alpha}.$$

Since $\pi[A] > 0$, we can choose α such as

$$\alpha = \log(1 + \frac{1}{\pi[A]}) > \log(1 + 0) = 0$$

which leads to

$$\log(\pi[A](e^{\alpha} - 1) + 1) = \log(\pi[A](e^{\log(1+\frac{1}{\pi[A]})} - 1) + 1)$$
$$= \log(\pi[A](1 + \frac{1}{\pi[A]}) - 1) + 1) = \log(\pi[A] + 1 - \pi[A] + 1) = \log 2.$$

Then, from

$$\mu[A] \leq \frac{\log(\pi[A](e^{\alpha} - 1) + 1) + H(\mu)}{\alpha}$$

We get

$$\mu[A] \leq \frac{\log(\pi[A](e^{\alpha} - 1) + 1) + H(\mu)}{\alpha} = \frac{\log 2 + H(\mu)}{\log(1 + \frac{1}{\pi[A]})}.$$

The following result will be useful in the proof of an explicit formula for the entropy.

Proposition

Let S be a set. Let μ, π be probability measures on S. Define the functional $\Phi : \mathbb{R}^{|S|} \to \mathbb{R}$ by

$$\Phi(f) = < \mu, f > - log(< \pi, e^f >), \forall f : S \to \mathbb{R}.$$

Then Φ is concave in $\mathbb{R}^{|S|}$.

Let $f, g : S \to \mathbb{R}$.. Let $\alpha \in [0,1]$. There are three possibilities: $\alpha = 0$ (case 1), $\alpha = 1$ (case 2) or $\alpha \in (0,1)$ (case 3).
Case 1: $\alpha = 0$. In this case, we have

$$\Phi\big(\alpha f + (1-\alpha)g\big) = \Phi\big(0 \cdot f + (1-0)g\big) = \Phi(g)$$
$$\geq 0 \cdot \Phi(f) + 1 \cdot \Phi(g) = 0 \cdot \Phi(f) + (1-0)\Phi(g) = \alpha\Phi(f) + (1-\alpha)\Phi(g).$$

Case 2: $\alpha = 1$. In this case, we have

$$\Phi\big(\alpha f + (1-\alpha)g\big) = \Phi\big(1 \cdot f + (1-1)g\big) = \Phi(f)$$
$$\geq 1 \cdot \Phi(f) + 0 \cdot \Phi(g) = 1 \cdot \Phi(f) + (1-1)\Phi(g) = \alpha\Phi(f) + (1-\alpha)\Phi(g).$$

Case 3: $\alpha \in (0,1)$. From Holder's inequality, we have

$$
\begin{aligned}
log(<\pi, e^{\alpha f+(1-\alpha)g>}) &= log\Big(\int_S e^{\alpha f} e^{(1-\alpha)g} d\pi\Big) \\
&\leq log\Bigg(\Big(\int_S (e^{\alpha f})^{\frac{1}{\alpha}} d\pi\Big)^{\alpha}\Big(\int_S (e^{(1-\alpha)g})^{\frac{1}{1-\alpha}} d\pi\Big)^{1-\alpha}\Bigg) \\
&= log\Bigg(\Big(\int_E e^{f} d\pi\Big)^{\alpha}\Big(\int_S e^{g} d\pi\Big)^{1-\alpha}\Bigg) \\
&= \alpha log\Big(\int_S e^{f} d\pi\Big) + (1-\alpha) log\Big(\int_S e^{g} d\pi\Big) \\
&= \alpha log(<\pi, e^{f}>) + (1-\alpha) log(<\pi, e^{g}>),
\end{aligned}
$$

which leads to

$$
-log(<\pi, e^{\alpha f+(1-\alpha)g>}) \geq -\alpha log(<\pi, e^{f}>) - (1-\alpha) log(<\pi, e^{g}>).
$$

Then, we get

$$\Phi\big(\alpha f+(1-\alpha)g\big) =< \mu, \alpha f+(1-\alpha)g > -log(< \pi, e^{\alpha f+(1-\alpha)g>})$$
$$\geq < \mu, \alpha f > + < \mu, (1-\alpha)g > -\alpha log(< \pi, e^{f} >) - (1-\alpha)log(< \pi, e^{g} >)$$
$$=\alpha < \mu, f > +(1-\alpha) < \mu, g > -\alpha log(< \pi, e^{f} >) - (1-\alpha)log(< \pi, e^{g} >)$$
$$=\alpha\big(< \mu, f > -log(< \pi, e^{f} >)\big) + (1-\alpha)\big(< \mu, g > -log(< \pi, e^{g} >)\big)$$
$$=\alpha\Phi(f) + (1-\alpha)\Phi(g).$$

Since f, g are arbitrary, we have

$$\Phi\big(\alpha f+(1-\alpha)g\big) \geq \alpha\Phi(f) + (1-\alpha)\Phi(g), \forall \alpha \in [0, 1], \forall f : S \to \mathbb{R}.$$

Therefore, the functional Φ is concave in $\mathbb{R}^{|S|}$.

The next results presents an explicit formula for the entropy.

Theorem

The entropy $H(\mu)$ is given by the formula

$$H(\mu) = \sum_{x \in E} \pi(x) \frac{\mu(x)}{\pi(x)} \log\left(\frac{\mu(x)}{\pi(x)}\right) = \sum_{x \in E} \mu(x) \log\left(\frac{\mu(x)}{\pi(x)}\right)$$

if μ is absolutely continuous with respect to π and is equal to ∞ otherwise.

There are two possibilities: μ is not absolutely continuous with respect to π (case 1) or μ is absolutely continuous with respect to π (case 2).

Case 1: μ is not absolutely continuous with respect to π. In this case, since E is countable, there is $x_0 \in E$ such that $\mu(x_0) > 0$ and $\pi(x_0) = 0$. For each $n \in \mathbb{N}$, consider $f_n : E \to \mathbb{R}$ given by

$$f_n(x) = \begin{cases} n, & \text{if} \;\; x = x_0, \\ 0, & \text{if} \;\; x \neq x_0. \end{cases}$$

Then, we have

$$
\begin{aligned}
< \mu, f_n > - log(< \pi, e^{f_n} >) &= \sum_{x \in E} \mu(x) f_n(x) - log\Big(\sum_{x \in E} \pi(x) e^{f_n(x)}\Big) \\
&= \mu(x_0) f_n(x_0) + \sum_{x \neq x_0} \mu(x) f_n(x) - log\Big(\pi(x_0) e^{f_n(x_0)} + \sum_{x \neq x_0} \pi(x) e^{f_n(x)}\Big) \\
&= \mu(x_0) n + \sum_{x \neq x_0} \mu(x) \cdot 0 - log\Big(0 \cdot e^n + \sum_{x \neq x_0} \pi(x) e^1\Big) \\
&= n\mu(x_0) - log\Big(e^1 \sum_{x \neq x_0} \pi(x)\Big) = n\mu(x_0) - log(e(1 - \pi(x_0))) \\
&= n\mu(x_0) - log(e(1-0)) = n\mu(x_0) - log(e) = n\mu(x_0) - 1, \forall n \in \mathbb{N}.
\end{aligned}
$$

Since f_n is bounded, $\forall n \in \mathbb{N}$, we get

$$
H(\mu) \geq \limsup_{n \to \infty} \big[< \mu, f_n > - \log(< \pi, e^{f_n} >) \big] = \limsup_{n \to \infty} [n\mu(x_0) - 1] = \infty.
$$

Case 2: μ is absolutely continuous with respect to π. Then, for every $x \in E$ with $\pi(x) = 0$, we have $\mu(x) = 0$.
There are two possibilities: E is finite (case 2.1) or E is not finite (case 2.2).

Case 2.1: E is finite.
In this case, we can write $E = \{x_1, \ldots, x_N\}$, with $N = |E|$. For every $f : E \to \mathbb{R}$, denote $y_j = f(x_j), \forall j = 1, \ldots, N$. We denote $\mu_j := \mu(x_j)$ and $\pi_j = \pi(x_j)$. We also denote the functional $\Phi : \mathbb{R}^{|E|} \to \mathbb{R}$ by

$$\begin{aligned}\Phi(f) &:= <\mu, f> - log(<\pi, e^f>)\\ &= \sum_{i=1}^{N} \mu_i y_i - log\Big(\sum_{i=1}^{N} \pi_i e^{y_i}\Big) = \Phi(y_1, \ldots, y_N).\end{aligned}$$

This leads to

$$\frac{\partial \Phi}{\partial y_j}(y_1, \ldots, y_N) = \mu_j - \frac{\pi_j e^{y_j}}{\sum_{i=1}^{N} \pi_i e^{y_i}}, \forall j = 1, \ldots, N.$$

From Proposition 3, we have that Φ is concave in $\mathbb{R}^{|E|}$, then Φ assumes its maximum where its gradient vanishes. In particular, consider $f_0 : E \to \mathbb{R}$ given by

$$f_0(x_j) := y_{0,j} = \begin{cases} log\left(\frac{\mu_j}{\pi_j}\right) \text{ if } \pi_j \neq 0; \\ 0 \text{ if } \pi_j = 0. \end{cases}$$

Then we have

$$\sum_{i=1}^{N} \pi_i e^{y_{0,i}} = \sum_{i=1}^{N} \pi_i \frac{\mu_i}{\pi_i} = \sum_{i=1}^{N} \mu_i = 1,$$

which leads to

$$\begin{aligned} \frac{\partial \Phi}{\partial y_j}(y_{0,1}, \ldots, y_{0,N}) =& \mu_j - \frac{\pi_j e^{y_{0,j}}}{\sum_{i=1}^{N} \pi_i e^{y_{0,i}}} = \mu_j - \frac{\pi_j \frac{\mu_j}{\pi_j}}{1} \\ =& \mu_j - \pi_j \frac{\mu_j}{\pi_j} = \mu_j - \mu_j = 0, \forall j = 1, \ldots, N. \end{aligned}$$

Then Φ attains its maximum at f_0, which leads to

$$\begin{aligned} H(\mu) =& \Phi(f_0) = \sum_{i=1}^{N} \mu_i y_{0,i} - log\Big(\sum_{i=1}^{N} \pi_i e^{y_{0,i}}\Big) \\ &= \sum_{i=1}^{N} \mu_i log\Big(\frac{\mu_j}{\pi_j}\Big) - log(1) \\ &= \sum_{x \in E} \mu(x) \log\left(\frac{\mu(x)}{\pi(x)}\right) = \sum_{x \in E} \pi(x) \frac{\mu(x)}{\pi(x)} \log\left(\frac{\mu(x)}{\pi(x)}\right). \end{aligned}$$

Case 2.2: E is not finite.
Since E is countable, we can write $E = \{x_1, x_2, \ldots\}$. For every $k \in \mathbb{N}$, denote $E_k := \{x_1, \ldots, x_k\}$ and $\mathcal{D}(E_k)$ for the set of functions $f : E \to \mathbb{R}$ that are constant on the complement of E_k. Then $(E_k)_{k\geq 1}$ is an increasing sequence of finite subsets of E whose union is equal to E and $\mathcal{D}(E_k) \subset \mathcal{D}(E_{k+1}), \forall k \in \mathbb{N}$. For every f bounded, denote $\Phi(f)$ as

$$\Phi(f) :=< \mu, f > - \log(< \pi, e^f >).$$

Since f is bounded for all $f \in \mathcal{D}(E_k), \forall k \in \mathbb{N}$, we have

$$\sup_{f \in \mathcal{D}(E_k)} \Phi(f) \leq H(\mu), \forall k \in \mathbb{N},$$

which leads to

$$\lim_{k\to\infty} \sup_{f \in \mathcal{D}(E_k)} \Phi(f) \leq H(\mu).$$

Suppose that $H(\mu) > \lim_{k\to\infty} \sup_{f\in\mathcal{D}(E_k)} \Phi(f)$. Then there exists f bounded such that

$$\lim_{k\to\infty} \sup_{f\in\mathcal{D}(E_k)} \Phi(f) < \ <\mu, f> - \log(<\pi, e^f>) < H(\mu).$$

Since f is bounded, there exists $M \geq 1$ such that $|f(x)| \leq M, \forall x \in E$. Since μ is a probability measure, we have

$$\begin{aligned}\lim_{k\to\infty} \Big| \sum_{x\in E_k^C} \mu(x) f(x) \Big| &\leq \lim_{k\to\infty} \sum_{x\in E_k^C} \mu(x)|f(x)| \leq \lim_{k\to\infty} \sum_{x\in E_k^C} \mu(x) M \\ &= M \lim_{k\to\infty} \sum_{x\in E_k^C} \mu(x) = M \lim_{k\to\infty} \mu(E_k^C) = 0.\end{aligned}$$

Since π is a probability measure, we have

$$\lim_{k\to\infty}\sum_{x\in E_k^C}\pi(x)e^{f(x)} \leq \lim_{k\to\infty}\sum_{x\in E_k^C}\pi(x)e^{|f(x)|} \leq \lim_{k\to\infty}\sum_{x\in E_k^C}\pi(x)e^{M}$$
$$= e^M \lim_{k\to\infty}\sum_{x\in E_k^C}\pi(x) = e^M \lim_{k\to\infty}\pi(E_k^C) = 0.$$

For every $k\in\mathbb{N}$, define $f_k : E\to\mathbb{R}$ by

$$f_k(x) = \begin{cases} f(x), & \text{if } x\in E_k; \\ 0, & \text{if } x\notin E_k. \end{cases}$$

Since $|f_k(x)| \leq |f(x)| \leq M, \forall x\in E, \forall k\in\mathbb{N}$, f_k is bounded, $\forall k\in\mathbb{N}$.
Moreover, we have $f_k\in\mathcal{D}(E_k), \forall k\in\mathbb{N}$.

For every $k \in \mathbb{N}$, we have

$$\Phi(f) - \Phi(f_k) = <\mu, f> - \log(<\pi, e^f>) - \Big(<\mu, f_k> - \log(<\pi, e^{f_k}>)\Big)$$

$$= \sum_{x\in E} \mu(x)[f(x) - f_k(x)] - log\Big(\frac{\sum_{x\in E} \pi(x)e^{f(x)}}{\sum_{x\in E} \pi(x)e^{f_k(x)}}\Big)$$

$$= \sum_{x\in E_k} \mu(x)[f(x) - f_k(x)] + \sum_{x\in E_k^C} \mu(x)[f(x) - f_k(x)]$$

$$-log\Big(\frac{\sum_{x\in E_k} \pi(x)e^{f(x)} + \sum_{x\in E_k^C} \pi(x)e^{f(x)}}{\sum_{x\in E_K} \pi(x)e^{f_k(x)} + \sum_{x\in E_K^C} \pi(x)e^{f_k(x)}}\Big)$$

$$= \sum_{x\in E_k} \mu(x)[f(x) - f(x)] + \sum_{x\in E_k^C} \mu(x)[f(x) - 0]$$

$$-log\Big(\frac{\sum_{x\in E_k} \pi(x)e^{f(x)} + \sum_{x\in E_k^C} \pi(x)e^{f(x)}}{\sum_{x\in E_K} \pi(x)e^{f(x)} + \sum_{x\in E_K^C} \pi(x)e^{0}}\Big).$$

Then, we get

$$\Phi(f) - \Phi(f_k) = \sum_{x \in E_k^C} \mu(x) f(x) - log\Big(\frac{\sum_{x \in E_k} \pi(x) e^{f(x)} + \sum_{x \in E_k^C} \pi(x) e^{f(x)}}{\sum_{x \in E_K} \pi(x) e^{f(x)} + \pi(E_k^C)} \Big).$$

This leads to

$$\lim_{k\to\infty}[\Phi(f)-\Phi(f_k)]$$

$$=lim_{k\to\infty}\Big[\sum_{x\in E_k^C}\mu(x)f(x) - log\Big(\frac{\sum_{x\in E_k}\pi(x)e^{f(x)}+\sum_{x\in E_k^C}\pi(x)e^{f(x)}}{\sum_{x\in E_K}\pi(x)e^{f(x)}+\pi(E_k^C)}\Big)\Big]$$

$$=lim_{k\to\infty}\sum_{x\in E_k^C}\mu(x)f(x)$$

$$-log\Big(\frac{lim_{k\to\infty}\sum_{x\in E_k}\pi(x)e^{f(x)}+lim_{k\to\infty}\sum_{x\in E_k^C}\pi(x)e^{f(x)}}{lim_{k\to\infty}\sum_{x\in E_k}\pi(x)e^{f(x)}+lim_{k\to\infty}\pi(E_k^C)}\Big)$$

$$=0-log\Big(\frac{\sum_{x\in E}\pi(x)e^{f(x)}+0}{\sum_{x\in E}\pi(x)e^{f(x)}+0}\Big)=0-log(1)=0-0=0.$$

Since $\lim_{k\to\infty}[\Phi(f) - \Phi(f_k)] = 0$, we have a contradiction with

$$\lim_{k\to\infty} \sup_{f\in\mathcal{D}(E_k)} \Phi(f) < \; <\mu, f> - \log(<\pi, e^f>) < H(\mu).$$

Therefore, we have

$$\lim_{k\to\infty} \sup_{f\in\mathcal{D}(E_k)} \Phi(f) = H(\mu).$$

Let $k \in \mathbb{N}$. Let $f \in \mathcal{D}(E_k)$. Then there exists $y_0 \in \mathbb{R}$ such that $f(x) = y_0, \forall x \in E_k^C$. Denote $y_j := f(x_j), \forall 1 \leq j \leq k$. Therefore

$$\begin{aligned}\Phi(f) =: & \sum_{x \in E} \mu(x) f(x) - log\Big(\sum_{x \in E} \pi(x) e^{f(x)} \Big) \\ = & \sum_{x \in E_k} \mu(x) f(x) + \sum_{x \in E_k^C} \mu(x) f(x) - log\Big(\sum_{x \in E_k} \pi(x) e^{f(x)} + \sum_{x \in E_k^C} \pi(x) e^{f(x)} \Big) \\ = & \sum_{j=1}^{k} \mu_j y_j + \sum_{x \in E_k^C} \mu(x) y_0 - log\Big(\sum_{j=1}^{k} \pi_j e^{y_j} + \sum_{x \in E_k^C} \pi(x) e^{y_0} \Big).\end{aligned}$$

Then, we get

$$\Phi(f) = \sum_{j=1}^{k} \mu_j y_j + \mu(E_k^C) y_0 - log\Big(\sum_{j=1}^{k} \pi_j e^{y_j} + \pi(E_k^C) e^{y_0}\Big)$$
$$= \Phi_k(y_0, y_1, \dots, y_k),$$

where $\Phi_k : \mathbb{R}^{k+1} \to \mathbb{R}$ is defined by

$$\Phi_k(y_0, y_1, \dots, y_k) = \sum_{j=1}^{k} \mu_j y_j + \mu(E_k^C) y_0 - log\Big(\sum_{j=1}^{k} \pi_j e^{y_j} + \pi(E_k^C) e^{y_0}\Big).$$

This leads to

$$\frac{\partial \Phi_k}{\partial y_j}(y_0, y_1, \ldots, y_k) = \mu_j - \frac{\pi_j e^{y_j}}{\sum_{i=1}^{k} \pi_i e^{y_i} + \pi(E_k^C) e^{y_0}}, \forall j = 1, \ldots, k$$

and to

$$\frac{\partial \Phi_k}{\partial y_0}(y_0, y_1, \ldots, y_k) = \mu(E_k^C) - \frac{\pi(E_k^C) e^{y_0}}{\sum_{i=1}^{k} \pi_i e^{y_i} + \pi(E_k^C) e^{y_0}}.$$

Let $S_k = \{0, x_1, x_2, \ldots, x_k\}$ be a set with $k+1$ different elements. Define $\mu^k : S_k \to \mathbb{R}$ by

$$\mu^k(x) = \begin{cases} \mu(E_k^C) \geq 0, \text{if } \ x = 0; \\ \mu(x_k) \geq 0, \text{if } \ x = x_k. \end{cases}$$

Then, we have

$$\begin{aligned} \sum_{x \in S_k} \mu^k(x) =& \mu^k(0) + \sum_{j=1}^{k} \mu^k(x_j) = \mu(E_k^C) + \sum_{j=1}^{k} \mu(x_j) \\ =& \mu(E_k^C) + \sum_{x \in E_k}^{k} \mu(x) = \mu(E_k^C) + \mu(E_k) = 1. \end{aligned}$$

Therefore, μ^k is a probability measure on S_k.

Define $\pi^k : S_k \to \mathbb{R}$ by

$$\pi^k(x) = \begin{cases} \pi(E_k^C) \geq 0, \text{if } \; x = 0; \\ \pi(x_k) \geq 0, \text{if } \; x = x_k. \end{cases}$$

Then, we have

$$\begin{aligned} \sum_{x \in S_k} \pi^k(x) =& \pi^k(0) + \sum_{j=1}^{k} \pi^k(x_j) = \pi(E_k^C) + \sum_{j=1}^{k} \pi(x_j) \\ =& \pi(E_k^C) + \sum_{x \in E_k}^{k} \pi(x) = \pi(E_k^C) + \pi(E_k) = 1. \end{aligned}$$

Therefore, π^k is a probability measure on S_k.

Let $f_k \in \mathbb{R}^{|S_k|}$. If we write $f = (y_0, y_1, \ldots, y_k)$, we have

$$\begin{aligned}\Phi_k(f_k) =& \Phi_k(y_0, y_1, \ldots, y_k) \\ =& \sum_{j=1}^{k} \mu_j y_j + \mu(E_k^C) y_0 - log\Big(\sum_{j=1}^{k} \pi_j e^{y_j} + \pi(E_k^C) e^{y_0} \Big) \\ =& \mu^k(0) y_0 + \sum_{j=1}^{k} \mu^k(x_j) y_j - log\Big(\pi^k(0) e^{y_0} + \sum_{j=1}^{k} \pi^k(x_j) e^{y_j} \Big) \\ =& < \mu^k, f_k > - log(< \pi^k, e^{f_k} >).\end{aligned}$$

Since μ^k, π^k are probability measures on the finite set S_k, from Proposition 3, Φ_k is concave in $\mathbb{R}^{|S_k|}$. Then Φ_k assumes its maximum where its gradient vanishes.

In particular, define $f_{0,k} : E \to \mathbb{R}$ by

$$f_{0,k}(x) = \begin{cases} log(\frac{\mu(x)}{\pi(x)}), & \text{if } x \in E_k \text{ and } \pi(x) \neq 0; \\ 0, & \text{if } x \in E_k \text{ and } \pi(x) = 0; \\ c_0 := log\left(\frac{\mu(E_k^C)}{\pi(E_k^C)}\right), & \text{if } x \notin E_k. \end{cases}$$

Denote $y_{0,j} := f_{0,k}(x_j), \forall j = 1, \ldots, k$. Then

$$\sum_{i=1}^{k} \pi_i e^{y_{0,i}} + \pi(E_k^C) e^{c_0} = \sum_{x \in E_k} \pi(x) e^{f_{0,k}(x)} + \pi(E_k^C) e^{log(\frac{\mu(E_k^C)}{\pi(E_k^C)})}$$

$$= \sum_{x \in E_k} \pi(x) e^{log(\frac{\mu(x)}{\pi(x)})} + \pi(E_k^C) \frac{\mu(E_k^C)}{\pi(E_k^C)}$$

$$= \sum_{x \in E_k} \pi(x) \frac{\mu(x)}{\pi(x)} + \mu(E_k^C)$$

$$= \sum_{x \in E_k} \mu(x) + \mu(E_k^C) = \mu(E_k) + \mu(E_k^C) = \mu(E) = 1.$$

For every $j = 1, \ldots, k$, we have

$$\frac{\partial \Phi_k}{\partial y_j}(c_0, y_{0,1}, \ldots, y_{0,k}) = \mu_j - \frac{\pi_j e^{y_{0,j}}}{\sum_{i=1}^k \pi_i e^{y_{0,i}} + \pi(E_k^C) e^{c_0}}$$

$$= \mu_j - \frac{\pi_j \frac{\mu_j}{\pi_j}}{1} = \mu_j - \mu_j = 0.$$

Also, we have

$$\frac{\partial \Phi_k}{\partial y_0}(c_0, y_{0,1}, \ldots, y_{0,k}) = \mu(E_k^C) - \frac{\pi(E_k^C) e^{c_0}}{\sum_{i=1}^k \pi_i e^{y_{0,i}} + \pi(E_k^C) e^{c_0}}$$

$$= \mu(E_k^C) - \frac{\pi(E_k^C) \frac{\mu(E_k^C)}{\pi(E_k^C)}}{1} = \mu(E_k^C) - \mu(E_k^C) = 0.$$

Then Φ_k attains maximum in $(c_0, y_{0,1}, \dots, y_{0,k})$. This leads to

$$\begin{aligned}
\sup_{f \in \mathcal{D}(E_k)} \Phi(f) =& \Phi_k(c_0, y_{0,1}, \dots, y_{0,k}) \\
&= \sum_{j=1}^{k} \mu_j y_{0,j} + \mu(E_k^C) c_0 - log\Big(\sum_{j=1}^{k} \pi_j e^{y_{0,j}} + \pi(E_k^C) e^{c_0} \Big) \\
&= \sum_{x \in E_k} \mu(x) log\Big(\frac{\mu(x)}{\pi(x)}\Big) + \mu(E_k^C) log\Big(\frac{\mu(E_k^C)}{\pi(E_k^C)}\Big) - log\Big(1\Big) \\
&= \sum_{x \in E_k} \pi(x) \frac{\mu(x)}{\pi(x)} log\Big(\frac{\mu(x)}{\pi(x)}\Big) + \pi(E_k^C) \frac{\mu(E_k^C)}{\pi(E_k^C)} log\Big(\frac{\mu(E_k^C)}{\pi(E_k^C)}\Big) \\
&= \sum_{x \in E_k} \pi(x) g\Big(\frac{\mu(x)}{\pi(x)}\Big) + \pi(E_k^C) g\Big(\frac{\mu(E_k^C)}{\pi(E_k^C)}\Big).
\end{aligned}$$

Since $\mathcal{D}(E_k) \subset \mathcal{D}(E_{k+1}), \forall k \in \mathbb{N}$, we have that $(\sup_{f\in\mathcal{D}(E_k)} \Phi(f))_{k\in\mathbb{N}}$ is an increasing sequence. Finally, observe that

$$\lim_{k\to\infty} \pi(E_k^C) = \lim_{k\to\infty} \mu(E_k^C) = 0,$$

which leads to

$$\begin{aligned}
H(\mu) &= \lim_{k\to\infty} \sup_{f\in\mathcal{D}(E_k)} \Phi(f) \\
&= \lim_{k\to\infty} \Big[\sum_{x\in E_k} \pi(x) g\Big(\frac{\mu(x)}{\pi(x)}\Big) + \pi(E_k^C) g\Big(\frac{\mu(E_k^C)}{\pi(E_k^C)}\Big)\Big] \\
&= \sum_{x\in E} \pi(x) g\Big(\frac{\mu(x)}{\pi(x)}\Big) + 0 \\
&= \sum_{x\in E} \pi(x) \frac{\mu(x)}{\pi(x)} log\Big(\frac{\mu(x)}{\pi(x)}\Big) = \sum_{x\in E} \mu(x) log\Big(\frac{\mu(x)}{\pi(x)}\Big).
\end{aligned}$$

This explicit formula for the relative entropy involving the function $u \log u$ explains the relation between the entropy and the expectation of functions of type e^f in the entropy inequality. Indeed, we starting from the explicit formula in the previous result, we can derive the entropy inequality.

The following result is true:

Proposition

We have

$$uv \leq e^v + u\log(u) - u, \forall u \geq 0, \forall v \in \mathbb{R}. \tag{1}$$

If $u = 0$, we have

$$uv = 0 \leq e^v = e^v + 0\log(0) - 0 = e^v + u\log u - u$$

and the result holds. Consider $F : (0, \infty) \times \mathbb{R} \to \mathbb{R}$ given by

$$F(u, v) = ulog(u) + e^v - uv - u, \forall u > 0, \forall v \in \mathbb{R}.$$

We have

$$\frac{\partial F}{\partial u}(u,v) = log(u) + 1 - v - 1 = log(u) - v, \forall u > 0, \forall v \in \mathbb{R}$$

and

$$\frac{\partial F}{\partial v}(u,v) = E^{v} - u, \forall u > 0, \forall v \in \mathbb{R}.$$

Then the points (u_0, v_0) in which the gradient of F vanishes are the points of the curve $u_0 = e^{v_0}, v_0 \in \mathbb{R}$.

We also have

$$\frac{\partial^2 F}{\partial u^2}(u,v) = \frac{1}{u}, \forall u > 0, \forall v \in \mathbb{R},$$

$$\frac{\partial^2 F}{\partial u \partial v}(u,v) = -1, \forall u > 0, \forall v \in \mathbb{R},$$

and

$$\frac{\partial^2 F}{\partial v^2}(u,v) = e^v, \forall u > 0, \forall v \in \mathbb{R}.$$

In the points of the curve $u_0 = e^{v_0}$, the eigenvalues of the Hessian matrix of F are 0 and $u_0 + \frac{1}{u_0} > 0$.

Therefore, F attains its minimum when $u = e^v$, which leads to

$$\begin{aligned} F_{min}(u,v) =& F(u_0,v_0) = u_0 log(u_0) + e^{v_0} - u_0 v_0 - u_0 \\ =& u_0 v_0 + u_0 - u_0 v_0 - u_0 = 0. \end{aligned}$$

Therefore,

$$ulog(u) + e^v - uv - u \geq 0, \forall u > 0, \forall v \in \mathbb{R},$$

which leads to

$$uv \leq ulog(u) + e^v - u, \forall u \geq 0, \forall v \in \mathbb{R}.$$

Proposition

If μ is absolutely continuous with respect to π and

$$\bar{H}(\mu) := \sum_{x \in E} \pi(x) \frac{\mu(x)}{\pi(x)} log\left(\frac{\mu(x)}{\pi(x)}\right),$$

then

$$\bar{H}(\mu) \geq \sup_{f} \{< \mu, f > - \log(< \pi, e^{f} >)\},$$

where the supremum is carried over all bounded functions f and $< \mu, f >$ stands for the integral of f with respect to μ.

Let $f : E \to \mathbb{R}$ be a bounded function. Take u as the density of μ with respect to π and v as the function f plus a constant c. Then

$$\int_E uvd\pi = \int_E \frac{d\mu}{d\pi}(f+c)d\pi = \int_E (f+c)d\mu = c + \int_E fd\mu,$$

$$\begin{aligned}\int_E ulog(u)d\pi &= \sum_{x\in E} \pi(x)u(x)log(u(x)) \\ &= \sum_{x\in E} \pi(x)\frac{\mu(x)}{\pi(x)}log\Big(\frac{\mu(x)}{\pi(x)}\Big) = \bar{H}(\mu),\end{aligned}$$

$$\int_E e^v d\pi = \int_E e^{f+c}d\pi = \int_E e^c e^f d\pi = e^c \int_E e^f d\pi,$$

and

$$\int_E -ud\pi = -\int_E \frac{d\mu}{d\pi}d\pi = -\mu(E) = -1.$$

Therefore, integrating (1) with respect to π, we get

$$\int_E uvd\pi \leq \int_E ulog(u)d\pi + \int_E e^v d\pi \int_E -ud\pi,$$

which is the same as

$$\int_E fd\mu + c \leq \bar{H}(\mu) + e^c \int_E e^f d\pi - 1$$

and we get

$$\bar{H}(\mu) \geq c + 1 + \int_E fd\mu - e^c \int_E e^f d\pi = g_1(c),$$

where $g_1 : \mathbb{R} \to \mathbb{R}$ is given by

$$g_1(c) = c + 1 + \int_E fd\mu - e^c \int_E e^f d\pi, \forall c \in \mathbb{R}.$$

Since $\bar{H}(\mu) \geq g_1(c), \forall c \in \mathbb{R}$, choosing

$$c_0 = -log\Big(\int_E e^f d\pi\Big),$$

we get

$$\begin{aligned}\bar{H}(\mu) \geq & g_1(c) \geq g_1(c_0) = c_0 + 1 + \int_E f d\mu - e^{c_0} \int_E e^f d\pi \\ = & - log\Big(\int_E e^f d\pi\Big) + 1 + \int_E f d\mu - \frac{\int_E e^f d\pi}{\int_E e^f d\pi} \\ = & - log\Big(\int_E e^f d\pi\Big) + 1 + \int_E f d\mu - 1 = \int_E f d\mu - log\Big(\int_E e^f d\pi\Big).\end{aligned}$$

Taking the supremum over every bounded function $f : E \to \mathbb{R}$, we have

$$\bar{H}(\mu) \geq \sup_{f} \Big\{ \int_E f d\mu - log\Big(\int_E e^f d\pi \Big) \Big\}.$$

Consider a Markov chain on a countable space E with an invariant measure denoted by π. Let $(P_t)_{t\geq 0}$ be the semigroup associated to the Markov chain. The following result will be useful in the first Proposition of this section.

Lemma

If μ is absolutely continuous with respect to π, then μP_t is absolutely continuous with respect to π, $\forall t \geq 0$.

Let $t \geq 0$. Let $x \in E$ such that $\pi(x) = 0$. Since π is an invariant measure, we get

$$0 = \pi(x) = (\pi P_t)(x) = \sum_{y \in E} \pi(y) P_t(y, x),$$

which leads to $\pi(y)P_t(y, x) = 0, \forall y \in E$. Since μ is absolutely continuous with respect to π, $\mu(y) = 0$ if and only if $\pi(y) = 0$, for every $y \in E$.
Let $y_0 \in E$. There are two possibilities: $\mu(y_0) = 0$ (case 1) or $\mu(y_0) \neq 0$ (case 2).
Case 1: $\mu(y_0) = 0$. In this case, we have $\mu(y_0)P_t(y_0, x) = 0 \cdot P_t(y_0, x) = 0$.

Case 2: $\mu(y_0) \neq 0$. In this case, we have $\pi(y_0) \neq 0$. Since $\pi(y_0)P_t(y_0, x) = 0$, we get $P_t(y_0, x) = 0$, which leads to $\mu(y_0)P_t(y_0, x) = \mu(y_0) \cdot 0 = 0$.
Therefore, we have $\mu(y)P_t(y, x) = 0, \forall y \in E$, which leads to

$$(\mu P_t)(x) = \sum_{y \in E} \mu(y)P_t(y, x) = \sum_{y \in E} 0 = 0.$$

Then $(\mu P_t)(x) = 0$ when $\pi(x) = 0$. This means that μP_t is absolutely continuous with respect to π. Since $t \geq 0$ is arbitrary, we have that μP_t is absolutely continuous with respect to π, $\forall t \geq 0$.

The relative entropy with respect to the invariant measure plays an important role in the investigation of the time evolution of the process. indeed, since $\phi(u) = u \log u$ is strictly convex and vanish only at $s = 0$ and $s = 1$, the relative entropy of μP_t with respect to π does not increase in time. This is the content of the next proposition.

Proposition

For every probability measure μ, we have

$$H(\mu P_t) \leq H(\mu).$$

Moreover, $H(\mu P_t) = H(\mu) < \infty$ implies that $\mu = \pi$ if the chain is indecomposable.

Let $\phi : (0, \infty) \to \mathbb{R}$, $\phi(x) = x \log(x), \forall x > 0$. Observe that $\phi \in C^\infty((0, \infty))$, $\frac{dg}{dx}(x) = log(x) + 1, \forall x > 0$ and $\frac{d^2 g}{dx^2}(x) = \frac{1}{x} > 0, \forall x > 0$, therefore ϕ is strictly convex.
If μ is not absolutely continuous with respect to π, we have $H(\mu) = \infty$, which leads to

$$H(\mu P_t) \leq \infty = H(\mu), \forall t \geq 0.$$

If μ is absolutely continuous with respect to π, Lemma 1 gives that μP_t is absolutely continuous with respect to π, $\forall t \geq 0$.

Let $t \geq 0$. From Theorem 1, we get

$$\begin{aligned} H(\mu P_t) &= \sum_{x \in E} \pi(x) \phi\Big(\frac{1}{\pi(x)} (\mu P_t)(x) \Big) \\ &= \sum_{x \in E} \pi(x) \phi\Big(\frac{1}{\pi(x)} \sum_{y \in E} \mu(y) P_t(y, x) \Big) \\ &= \sum_{x \in E} \pi(x) \phi\Big(\sum_{y \in E} \frac{\mu(y)}{\pi(y)} \frac{\pi(y) P_t(y, x)}{\pi(x)} \Big). \end{aligned}$$

Since π is an invariant measure, $\pi = \pi P_t \forall t \geq 0$. Then, for every $t \geq 0$, $x \in E$, we have

$$\frac{\pi(y) P_t(y, x)}{\pi(x)} \geq 0, \forall y \in E$$

and

$$\sum_{y \in E} \frac{\pi(y) P_t(y, x)}{\pi(x)} = \frac{1}{\pi(x)} \sum_{y \in E} \pi(y) P_t(y, x) = \frac{(\pi P_t)(x)}{\pi(x)} = \frac{\pi(x)}{\pi(x)} = 1.$$

Therefore, $\alpha_{t,x} : E \to \mathbb{R}$ given by $\alpha_{t,x}(y) = \frac{\pi(y)P_t(y,x)}{\pi(x)}$ is a probability measure. Then, Jensen's inequality leads to

$$\begin{aligned} H(\mu P_t) &= \sum_{x \in E} \pi(x)\phi\Big(\frac{1}{\pi(x)}(\mu P_t)(x)\Big) \\ &= \sum_{x \in E} \pi(x)\phi\Big(\sum_{y \in E} \frac{\mu(y)}{\pi(y)}\frac{\pi(y)P_t(y,x)}{\pi(x)}\Big) \\ &= \sum_{x \in E} \pi(x)\phi\Big(\sum_{y \in E} \frac{\mu(y)}{\pi(y)}\alpha_{t,x}(y)\Big) \\ &= \sum_{x \in E} \pi(x)\phi\Big(E_{\alpha_{t,x}}\Big[\frac{\mu}{\pi}\Big]\Big) \\ &\leq \sum_{x \in E} \pi(x)E_{\alpha_{t,x}}\Big[\phi\Big(\frac{\mu}{\pi}\Big)\Big]. \end{aligned}$$

Then, we get

$$\begin{aligned}
H(\mu P_t) &\leq \sum_{x\in E} \pi(x) E_{\alpha_{t,x}}\left[\phi\left(\frac{\mu}{\pi}\right)\right].\\
&= \sum_{x\in E} \pi(x) \sum_{y\in E} \phi\left(\frac{\mu(y)}{\pi(y)}\right)\alpha_{t,x}(y)\\
&= \sum_{x\in E} \pi(x) \sum_{y\in E} \phi\left(\frac{\mu(y)}{\pi(y)}\right)\frac{\pi(y)P_t(y,x)}{\pi(x)}\\
&= \sum_{y\in E} \pi(y)\phi\left(\frac{\mu(y)}{\pi(y)}\right) \sum_{x\in E} \pi(x)\frac{P_t(y,x)}{\pi(x)}\\
&= \sum_{y\in E} \pi(y)\phi\left(\frac{\mu(y)}{\pi(y)}\right) = H(\mu).
\end{aligned}$$

For the remainder of this section, $P_t^\star$ stands for the adjoint of P_t in $L^2(\pi)$, f stands for the density of μ with respect to π and f_t stands for the density of μP_t with respect to π. Then, we have

Proposition

$$f_t(x) = (P_t^\star f)(x).$$

In particular, the density f_t is solution of

$$\begin{cases} f_0 = f \\ \partial_t f_t = L^\star f_t. \end{cases} \tag{2}$$

We know that the adjoint of L in $L^2(\mu)$, denoted by $L^\star$, is a generator with $P_t^\star = e^{tL^\star}$ is also the adjoint of P_t in $L^2(\mu)$. Then, for $g \in L^2(\pi)$, we have

$$\begin{aligned} < g, P_t^\star f >_\pi = < P_t g, f >_\pi = \int_E (P_t g) f d\pi = \int_E (P_t g) \frac{d\mu}{d\pi} d\pi \\ = \int_E (P_t g) d\mu = \sum_{x \in E} \mu(x)(P_t g)(x) = \sum_{x \in E} \mu(x) \sum_{y \in E} P_t(x,y) g(y) \\ = \sum_{x \in E} \sum_{y \in E} \mu(x) P_t(x,y) g(y). \end{aligned}$$

We also have

$$
\begin{aligned}
< g, f_t >_\pi &= \int_E g f_t d\pi = \int_E g \frac{d(\mu P_t)}{d\pi} d\pi \\
&= \int_E g d(\mu P_t) = \sum_{y \in E} (\mu P_t)(y) g(y) = \sum_{y \in E} \sum_{x \in E} \mu(x) P_t(x,y) g(y) \\
&= \sum_{x \in E} \sum_{y \in E} \mu(x) P_t(x,y) g(y) = < g, P_t^\star f >_\pi \, .
\end{aligned}
$$

Since $< g, f_t >_\pi = < g, P_t^\star f >_\pi, \forall g \in L^2(\pi)$, we have

$$f_t(x) = (P_t^\star f)(x).$$

From the definition of f_t, we get

$$f_0 = \frac{d\mu P_0}{d\pi} = \frac{d\mu}{d\pi} = f.$$

From $P_t^\star = e^{tL^\star}$, we get

$$\partial_t P_t^\star = \partial_t(e^{tL^\star}) = L^\star e^{tL^\star} = L^\star P_t^\star,$$

which leads to

$$\partial_t f_t = \partial_t(P_t^\star f) = (\partial_t P_t^\star) f = (L^\star P_t^\star) f = L^\star (P_t^\star f) = L^\star f_t.$$

Therefore, the density f_t is solution of

$$\begin{cases} f_0 = f \\ \partial_t f_t = L^\star f_t. \end{cases}$$

The following result will be useful.

Lemma

We have

$$x[\log(y) - \log(x)] \leq 2\sqrt{x}[\sqrt{y} - \sqrt{x}], \forall x, y \geq 0.$$

For $x = 0$, $y \geq 0$, we have

$$x[\log(y) - \log(x)] = x\log(y) - x\log(x) = 0 - 0$$
$$=0 = 2\sqrt{0}[\sqrt{y} - \sqrt{0}] = 2\sqrt{x}[\sqrt{y} - \sqrt{x}].$$

For $x > 0$, $y = 0$, we have

$$x[\log(y) - \log(x)] = x\log(0) - x\log(x) = -\infty - x\log(x)$$
$$= -\infty < -2x = 2\sqrt{x}[\sqrt{0} - \sqrt{x}] = 2\sqrt{x}[\sqrt{y} - \sqrt{x}].$$

Define $F : (0, \infty)^2 \to \mathbb{R}$ by

$$F(x, y) = 2\sqrt{x}\sqrt{y} - 2x + x\log(x) - x\log(y), \forall x, y > 0.$$

Then we have

$$\frac{\partial F}{\partial x}(x, y) = \frac{\sqrt{y}}{\sqrt{x}} - 2 + \log(x) + 1 - log(y) = \sqrt{\frac{y}{x}} + \log\left(\frac{x}{y}\right) - 1, \forall x, y > 0,$$

and

$$\frac{\partial F}{\partial y}(x, y) = \frac{\sqrt{x}}{\sqrt{y}} - \frac{x}{y} = \sqrt{\frac{x}{y}}\left(1 - \sqrt{\frac{x}{y}}\right), \forall x, y > 0,$$

We also have

$$\frac{\partial^2 F}{\partial x^2}(x,y) = \frac{1}{x} - \frac{\sqrt{y}}{2x\sqrt{x}}, \forall x, y > 0,$$

$$\frac{\partial^2 F}{\partial y \partial x}(x,y) = \frac{\partial^2 F}{\partial x \partial y}(x,y) = \frac{\sqrt{1}}{2\sqrt{xy}} - \frac{1}{y}, \forall x, y > 0,$$

and

$$\frac{\partial^2 F}{\partial y^2}(x,y) = \frac{x}{y^2} - \frac{\sqrt{x}}{2y\sqrt{y}}, \forall x, y > 0.$$

Then, the points (x_0, y_0) such that

$$\frac{\partial F}{\partial x}(x_0, y_0) = \frac{\partial F}{\partial y}(x_0, y_0) = 0$$

are such that $x_0 = y_0 > 0$. In such points, we have

$$\begin{aligned} F(x_0, y_0) =& 2\sqrt{x_0}\sqrt{y_0} - 2x_0 + x_0 \log(x_0) - x_0 \log(y_0) \\ =& 2\sqrt{x_0}\sqrt{x_0} - 2x_0 + x_0 \log(x_0) - x_0 \log(x_0) = 2x_0 - 2x_0 = 0. \end{aligned}$$

Moreover, it holds

$$\frac{\partial^2 F}{\partial x^2}(x_0, y_0) = \frac{1}{x_0} - \frac{\sqrt{y_0}}{2x_0\sqrt{x_0}}$$
$$= \frac{1}{x_0} - \frac{\sqrt{x_0}}{2x_0\sqrt{x_0}} = \frac{1}{x_0} - \frac{1}{2x_0} = \frac{1}{2x_0},$$

$$\frac{\partial^2 F}{\partial y \partial x}(x_0, y_0) = \frac{\partial^2 F}{\partial x \partial y}(x_0, y_0) = \frac{\sqrt{1}}{2\sqrt{x_0 y_0}} - \frac{1}{y_0}$$
$$= \frac{\sqrt{1}}{2\sqrt{x_0 x_0}} - \frac{1}{x_0} = \frac{1}{2x_0} - \frac{1}{x_0} = -\frac{1}{2x_0},$$

and

$$\frac{\partial^2 F}{\partial y^2}(x_0, y_0) = \frac{x_0}{y_0^2} - \frac{\sqrt{x_0}}{2y_0\sqrt{y_0}}$$
$$= \frac{x_0}{x_0^2} - \frac{\sqrt{x_0}}{2x_0\sqrt{x_0}} = \frac{1}{x_0} - \frac{1}{2x_0} = \frac{1}{2x_0}.$$

In the points of the curve $y_0 = x_0 > 0$, the eigenvalues of the Hessian matrix of F are 0 and $\frac{1}{x_0} > 0$. Then, F attains its minimum in the points (x_0, y_0) such that $x_0 = y_0$. Then we have

$$2\sqrt{x}\sqrt{y} - 2x + x\log(x) - x\log(y) = F(x, y) \geq F(x_0, y_0) = 0, \forall x, y > 0,$$

which is the same as

$$x[\log(y) - \log(x)] \leq 2\sqrt{x}[\sqrt{y} - \sqrt{x}], \forall x, y > 0.$$

In this way, we get

$$x[\log(y) - \log(x)] \leq 2\sqrt{x}[\sqrt{y} - \sqrt{x}], \forall x, y \geq 0.$$

Finally, we will deduce a estimate for the time derivative of the entropy of μP_t.

Theorem

Let μ be a probability measure with finite entropy: $H(\mu) < \infty$. For every $t, h \geq 0$, we have that

$$H(\mu P_{t+h}) - H(\mu P_t) = \int_t^{t+h} < f_s, L \log f_s >_\pi ds$$

$$\leq \int_t^{t+h} 2 < \sqrt{f_s}, L\sqrt{f_s} >_\pi ds.$$

Moreover,

$$2 < \sqrt{f_s}, L\sqrt{f_s} >_\pi = - \sum_{x,y \in E} \pi(x) L(x,y) [\sqrt{f_s(y)} - \sqrt{f_s(x)}]^2.$$

We have

$$
\begin{aligned}
\frac{\partial}{\partial}(f_s \log\left(f_s(x)\right) =& \partial_s\big(f_s(x)\big)\log(f_s(x)) + f_s(x)\partial_s \log(f_s(x)) \\
=& \partial_s\big(f_s(x)\big)\log\left(f_s(x)\right) + f_s(x)\frac{\partial_s\big(f_s(x)\big)}{f_s(x)} \\
=& \partial_s\big(f_s(x)\big)[1 + \log\left(f_s(x)\right)] = L^\star f_s(x)[1 + \log\left(f_s(x)\right)].
\end{aligned}
$$

By the explicit formula for the entropy, the difference $H(\mu P_{t+h}) - H(\mu P_t)$ is equal to

$$\begin{aligned} H(\mu P_{t+h}) - H(\mu P_t) &= \sum_{x\in E} \pi(x) f_{t+h}(x) \log\left(f_{t+h}(x)\right) - \sum_{x\in E} \pi(x) f_t(x) \log\left(f_t(x)\right) \\ &= \sum_{x\in E} \pi(x) [f_{t+h}(x) \log\left(f_{t+h}(x)\right) - f_t(x) \log\left(f_t(x)\right)] \\ &= \sum_{x\in E} \pi(x) \int_t^{t+h} \frac{\partial}{\partial s}(f_s \log(f_s(x))) ds \\ &= \sum_{x\in E} \pi(x) \int_t^{t+h} L^\star f_s(x)[1 + \log\left(f_s(x)\right)] ds. \end{aligned}$$

Since π is an invariant probability measure, we observe that for every $x \in E$,

$$\sum_{y \in E} p^\star(x,y) = \sum_{y \in E} \frac{\lambda(y)\pi(y)p(y,x)}{\lambda(x)\pi(x)}$$

$$= \frac{1}{\lambda(x)\pi(x)} \sum_{y \in E} \lambda(y)\pi(y)p(y,x)$$

$$= \frac{1}{\lambda(x)\pi(x)} \lambda(x)\pi(x) = 1.$$

We denote the upper bound of the jump rate $\lambda(\cdot)$ by $\bar{\lambda}$.

We make the following claim.

Claim

The positive function $g_1 : E \to \mathbb{R}$ given by

$$g_1(x) = \sum_{y \neq x} L^{\star}(x, y) f_s(y) + f_s(x)\lambda(x), \forall x \in E,$$

is such that $\int_E g_1 d\pi \leq 2\bar{\lambda} < \infty$.

Indeed, we have

$$\int_E g_1 d\pi = \sum_{x\in E} g_1(x)\pi(x) = \sum_{x\in E}\Big[\sum_{y\neq x} L^\star(x,y)f_s(y) + f_s(x)\lambda(x)\Big]\pi(x)$$

$$= \sum_{x\in E}\sum_{y\neq x} L^\star(x,y)f_s(y)\pi(x) + \sum_{x\in E} f_s(x)\lambda(x)\pi(x)$$

$$= \sum_{x\in E}\Big[\sum_{y\in E} L^\star(x,y)f_s(y)\pi(x) - L^\star(x,x)f_s(x)\pi(x)\Big] + \sum_{x\in E} f_s(x)\lambda(x)\pi(x)$$

$$= \sum_{x\in E}\sum_{y\in E} L^\star(x,y)f_s(y)\pi(x) - \sum_{x\in E}\big(-\lambda(x)\big)f_s(x)\pi(x) + \sum_{x\in E} f_s(x)\lambda(x)\pi(x)$$

$$= \sum_{x\in E}\Big[\sum_{y\in E} L^\star(x,y)\big[[f_s(y) - f_s(x)] + f_s(x)\big]\Big]\pi(x) + 2\sum_{x\in E} f_s(x)\lambda(x)\pi(x).$$

This leads to

$$\int_E g_1 \, d\pi$$

$$\leq \sum_{x\in E}\Big[\sum_{y\in E} L^\star(x,y)[f_s(y)-f_s(x)] + \sum_{y\in E} L^\star(x,y) f_s(x)\Big]\pi(x) + 2\sum_{x\in E} f_s(x)\bar{\lambda}\pi(x)$$

$$= \sum_{x\in E}\Big[\sum_{y\in E} L^\star(x,y)[f_s(y)-f_s(x)] + f_s(x)\sum_{y\in E} L^\star(x,y)\Big]\pi(x) + 2\bar{\lambda}\sum_{x\in E} f_s(x)\pi(x)$$

$$= \sum_{x\in E}\Big[\sum_{y\in E} L^\star(x,y)[f_s(y)-f_s(x)] + f_s(x)\cdot 0\Big]\pi(x) + 2\bar{\lambda}\sum_{x\in E}(\mu P_s)(x)$$

$$= \sum_{x\in E} 1\Big[\sum_{y\in E} L^\star(x,y)[f_s(y)-f_s(x)]\Big]\pi(x) + 2\bar{\lambda}\cdot 1$$

$$= \sum_{x\in E} 1\cdot(L^\star f_s)(x) + 2\bar{\lambda}$$

$$= <\mathbf{1}, L^\star f_s>_\pi + 2\bar{\lambda} = < L\mathbf{1}, f_s>_\pi + 2\bar{\lambda} = <\mathbf{0}, f_s>_\pi + 2\bar{\lambda} = 2\bar{\lambda} < \infty.$$

For every $x \in E$, we have

$$\begin{aligned}
|L^\star f_s(x)| &= \Big| \sum_{y \in E} L^\star(x,y)[f_s(y) - f_s(x)] \Big| \\
&= \Big| \sum_{y \neq x} L^\star(x,y)[f_s(y) - f_s(x)] \Big| \\
&= \Big| \sum_{y \neq x} L^\star(x,y) f_s(y) - \sum_{y \neq x} L^\star(x,y) f_s(x) \Big| \\
&= \Big| \sum_{y \neq x} L^\star(x,y) f_s(y) - f_s(x) \sum_{y \neq x} L^\star(x,y) \Big| \\
&= \Big| \sum_{y \neq x} L^\star(x,y) f_s(y) - f_s(x)\lambda(x) \Big| \\
&\leq \Big| \sum_{y \neq x} L^\star(x,y) f_s(y) \Big| + |f_s(x)\lambda(x)| \\
&\leq \sum_{y \neq x} L^\star(x,y) f_s(y) + f_s(x)\lambda(x) = g_1(x).
\end{aligned}$$

This leads to

$$\int_t^{t+h}\Big[\int_E |L^\star f| d\pi\Big] ds \leq \int_t^{t+h}\Big[\int_E g_1 d\pi\Big] ds$$
$$= \int_t^{t+h}\sum_{x\in E} g_1(x)\pi(x)dx \leq \int_t^{t+h} 2\bar{\lambda} ds = 2\bar{\lambda}h.$$

Then, by Fubini Theorem, we get

$$\sum_{x\in E}\pi(x)\int_t^{t+h} L^\star f_s(x) ds = \int_E\Big[\int_t^{t+h} L^\star f_s ds\Big] d\pi$$
$$= \int_t^{t+h}\Big[\int_E L^\star f_s d\pi] ds = \int_t^{t+h} < L^\star f_s, \mathbf{1} >_\pi ds$$
$$= \int_t^{t+h} < f_s, L\mathbf{1} >_\pi ds = \int_t^{t+h} < f_s, \mathbf{0} >_\pi ds = \int_t^{t+h} 0 ds = 0.$$

By Fubini Theorem, we also get

$$\sum_{x\in E}\pi(x)\int_t^{t+h}\log\left(f_s(x)\right)L^\star f_s(x)ds=\int_E\left[\int_t^{t+h}\log(f_s)L^\star f_s ds\right]d\pi$$

$$=\int_t^{t+h}\left[\int_E\log(f_s)L^\star f_s d\pi\right]ds=\int_t^{t+h}<L^\star f_s,\log(f_s)>_\pi ds$$

$$=\int_t^{t+h}<f_s,L\log(f_s)>_\pi ds.$$

This leads to

$$H(\mu P_{t+h})-H(\mu P_t)=\sum_{x\in E}\pi(x)\int_t^{t+h}L^\star f_s(x)[1+\log\left(f_s(x)\right)]ds$$

$$=\sum_{x\in E}\pi(x)\int_t^{t+h}L^\star f_s(x)ds+\sum_{x\in E}\pi(x)\int_t^{t+h}\log\left(f_s(x)\right)L^\star f_s(x)ds$$

$$=0+\int_t^{t+h}<f_s,L\log(f_s)>_\pi ds=\int_t^{t+h}<f_s,L\log(f_s)>_\pi ds.$$

From Lemma 2, we get

$$\int_t^{t+h} < f_s, L\log(f_s) >_{d\pi} ds = \int_t^{t+h} \sum_{x\in E} f_s(x)\big(L\log(f_s)\big)(x)\pi(x)ds$$

$$= \int_t^{t+h} \sum_{x\in E} f_s(x)\Big[\sum_{y\in E} \lambda(x)p(x,y)[\log\left((f_s)(y)\right) - \log\left((f_s)(x)\right)]\Big]\pi(x)ds$$

$$= \int_t^{t+h} \sum_{x\in E} \Big[\sum_{y\in E} \lambda(x)p(x,y)f_s(x)[\log\left((f_s)(y)\right) - \log\left((f_s)(x)\right)]\Big]\pi(x)ds$$

$$\leq \int_t^{t+h} \sum_{x\in E} \Big[\sum_{y\in E} \lambda(x)p(x,y)2\sqrt{f_s(x)}[\sqrt{f_s(y)} - \sqrt{f_s(x)}]\Big]\pi(x)ds$$

$$=2\int_t^{t+h} \sum_{x\in E} \sqrt{f_s(x)}\Big[\sum_{y\in E} \lambda(x)p(x,y)[\sqrt{f_s(y)} - \sqrt{f_s(x)}]\Big]\pi(x)ds$$

$$=2\int_t^{t+h} \sum_{x\in E} (\sqrt{f_s})(x)(L\sqrt{f_s})(x)\pi(x)ds = \int_t^{t+h} 2 < \sqrt{f_s}, L\sqrt{f_s} >_\pi ds.$$

Finally, we will prove the final claim of the Theorem. We have

$$
\begin{aligned}
2 < \sqrt{f_s}, L\sqrt{f_s} >_\pi &= < \sqrt{f_s}, L\sqrt{f_s} >_\pi + < \sqrt{f_s}, L\sqrt{f_s} >_\pi \\
&= < \sqrt{f_s}, L\sqrt{f_s} >_\pi + < L^\star\sqrt{f_s}, \sqrt{f_s} >_\pi \\
&= \sum_{x\in E} \sqrt{f_s(x)}(L\sqrt{f_s})(x)\pi(x) + \sum_{x\in E} \sqrt{f_s(x)}(L^\star\sqrt{f_s})(x)\pi(x) \\
&= \sum_{x\in E} \sqrt{f_s(x)} \sum_{y\in E} L(x,y)[\sqrt{f_s(y)} - \sqrt{f_s(x)}]\pi(x) \\
&+ \sum_{x\in E} \sqrt{f_s(x)} \sum_{y\in E} L^\star(x,y)[\sqrt{f_s(y)} - \sqrt{f_s(x)}]\pi(x).
\end{aligned}
$$

Then, $2 < \sqrt{f_s}, L\sqrt{f_s} >_\pi$ is equal to

$$\sum_{x\in E}\sqrt{f_s(x)}\sum_{y\in E}L(x,y)[\sqrt{f_s(y)}-\sqrt{f_s(x)}]\pi(x)$$
$$+\sum_{x\in E}\sqrt{f_s(x)}\sum_{y\in E}L^\star(x,y)[\sqrt{f_s(y)}-\sqrt{f_s(x)}]\pi(x)$$
$$=\sum_{x\in E}\sum_{y\in E}\pi(x)L(x,y)\sqrt{f_s(x)}[\sqrt{f_s(y)}-\sqrt{f_s(x)}]$$
$$+\sum_{x\in E}\sum_{y\in E}\pi(x)L^\star(x,y)\sqrt{f_s(x)}[\sqrt{f_s(y)}-\sqrt{f_s(x)}]$$
$$=\sum_{x\in E}\sum_{y\in E}\pi(x)L(x,y)\sqrt{f_s(x)}[\sqrt{f_s(y)}-\sqrt{f_s(x)}]$$
$$+\sum_{y\in E}\sum_{x\in E}\pi(y)L^\star(y,x)\sqrt{f_s(y)}[\sqrt{f_s(x)}-\sqrt{f_s(y)}].$$

Since $L^\star$ is the adjoint of L in $L^2(\pi)$, we have

$$\pi(x)L(x,y) = \pi(y)L^\star(x,y), \forall x, y \in E,$$

which leads to

$$\begin{aligned}
&2 < \sqrt{f_s}, L\sqrt{f_s} >_\pi \\
&= \sum_{x\in E}\sum_{y\in E} \pi(x)L(x,y)\sqrt{f_s(x)}[\sqrt{f_s(y)} - \sqrt{f_s(x)}] \\
&+ \sum_{y\in E}\sum_{x\in E} \pi(y)L^\star(y,x)\sqrt{f_s(y)}[\sqrt{f_s(x)} - \sqrt{f_s(y)}] \\
&= \sum_{x\in E}\sum_{y\in E} \pi(x)L(x,y)\sqrt{f_s(x)}[\sqrt{f_s(y)} - \sqrt{f_s(x)}] \\
&+ \sum_{y\in E}\sum_{x\in E} \pi(x)L(x,y)\sqrt{f_s(y)}[\sqrt{f_s(x)} - \sqrt{f_s(y)}].
\end{aligned}$$

Finally, we get that $2 < \sqrt{f_s}, L\sqrt{f_s} >_\pi$ is equal to

$$\sum_{x\in E}\sum_{y\in E}\pi(x)L(x,y)\sqrt{f_s(x)}[\sqrt{f_s(y)}-\sqrt{f_s(x)}]$$

$$+\sum_{y\in E}\sum_{x\in E}\pi(x)L(x,y)\sqrt{f_s(y)}[\sqrt{f_s(x)}-\sqrt{f_s(y)}]$$

$$=\sum_{x,y\in E}\pi(x)L(x,y)\sqrt{f_s(x)}[\sqrt{f_s(y)}-\sqrt{f_s(x)}]$$

$$+\sum_{x,y\in E}\pi(x)L(x,y)(-\sqrt{f_s(y)})[\sqrt{f_s(y)}-\sqrt{f_s(x)}]$$

$$=\sum_{x,y\in E}\pi(x)L(x,y)[\sqrt{f_s(x)}-\sqrt{f_s(y)}][\sqrt{f_s(y)}-\sqrt{f_s(x)}]$$

$$=-\sum_{x,y\in E}\pi(x)L(x,y)[\sqrt{f_s(y)}-\sqrt{f_s(x)}][\sqrt{f_s(y)}-\sqrt{f_s(x)}]$$

$$=-\sum_{x,y\in E}\pi(x)L(x,y)[\sqrt{f_s(y)}-\sqrt{f_s(x)}]^2.$$

We introduce, for every function $f \in L^2(\pi)$, the Dirichlet form $\mathcal{D}(f)$ of f defined by

$$\mathcal{D}(f) := - < f, Lf >_\pi = - \sum_{x \in E} f(x) Lf(x) \pi(x).$$

The sum is well defined because the generator L is a bounded operator in $L^2(\pi)$.

Proposition

The Dirichlet form of a function $f \in L^2(\pi)$ is positive and equal to

$$\mathcal{D}(f) = \frac{1}{2} \sum_{x,y \in E} \pi(x) L(x,y) [f(y) - f(x)]^2.$$

Denote the adjoint of L in $L^2(\pi)$ by $L^\star$. Then

$$
\begin{aligned}
& 2 < f_s, Lf_s >_\pi = < f_s, Lf_s >_\pi + < f_s, Lf_s >_\pi \\
&= < f_s, Lf_s >_\pi + < L^\star f_s, f_s >_\pi \\
&= \sum_{x \in E} f_s(x)(Lf_s)(x)\pi(x) + \sum_{x \in E} f_s(x)(L^\star f_s)(x)\pi(x) \\
&= \sum_{x \in E} f_s(x) \sum_{y \in E} L(x,y)[f_s(y) - f_s(x)]\pi(x) \\
&+ \sum_{x \in E} f_s(x) \sum_{y \in E} L^\star(x,y)[f_s(y) - f_s(x)]\pi(x).
\end{aligned}
$$

Then, $2 < f_s, Lf_s >_\pi$ is equal to

$$\sum_{x\in E} f_s(x) \sum_{y\in E} L(x,y)[f_s(y) - f_s(x)]\pi(x)$$
$$+ \sum_{x\in E} f_s(x) \sum_{y\in E} L^\star(x,y)[f_s(y) - f_s(x)]\pi(x)$$
$$= \sum_{x\in E}\sum_{y\in E} \pi(x)L(x,y)f_s(x)[f_s(y) - f_s(x)]$$
$$+ \sum_{x\in E}\sum_{y\in E} \pi(x)L^\star(x,y)f_s(x)[f_s(y) - f_s(x)]$$
$$= \sum_{x\in E}\sum_{y\in E} \pi(x)L(x,y)f_s(x)[f_s(y) - f_s(x)]$$
$$+ \sum_{y\in E}\sum_{x\in E} \pi(y)L^\star(y,x)f_s(y)[f_s(x) - f_s(y)].$$

Since $L^\star$ is the adjoint of L in $L^2(\pi)$, we have

$$\pi(x)L(x,y) = \pi(y)L^\star(x,y), \forall x, y \in E,$$

which leads to

$$\begin{aligned}
& 2 < f_s, Lf_s >_\pi \\
= & \sum_{x\in E}\sum_{y\in E} \pi(x)L(x,y)f_s(x)[f_s(y) - f_s(x)] \\
+ & \sum_{y\in E}\sum_{x\in E} \pi(y)L^\star(y,x)f_s(y)[f_s(x) - f_s(y)] \\
= & \sum_{x\in E}\sum_{y\in E} \pi(x)L(x,y)f_s(x)[f_s(y) - f_s(x)] \\
+ & \sum_{y\in E}\sum_{x\in E} \pi(x)L(x,y)f_s(y)[f_s(x) - f_s(y)].
\end{aligned}$$

Then, we get that $2 < f_s, Lf_s >_\pi$ is equal to

$$\sum_{x\in E}\sum_{y\in E}\pi(x)L(x,y)f_s(x)[f_s(y)-f_s(x)]$$
$$+\sum_{y\in E}\sum_{x\in E}\pi(x)L(x,y)f_s(y)[f_s(x)-f_s(y)]$$
$$=\sum_{x\in E}\sum_{y\in E}\pi(x)L(x,y)f_s(x)[f_s(y)-f_s(x)]$$
$$+\sum_{y\in E}\sum_{x\in E}\pi(x)L(x,y)\big(-f_s(y)\big)[f_s(y)-f_s(x)]$$
$$=\sum_{x,y\in E}\pi(x)L(x,y)[f_s(x)-f_s(y)][f_s(y)-f_s(x)]$$
$$=-\sum_{x,y\in E}\pi(x)L(x,y)[f_s(y)-f_s(x)][f_s(y)-f_s(x)]$$
$$=-\sum_{x,y\in E}\pi(x)L(x,y)[f_s(y)-f_s(x)]^2.$$

Finally, we have

$$\begin{aligned}\mathcal{D}(f) :=& - < f, Lf >_\pi \\ =& -\frac{1}{2}[2 < f_s, Lf_s >_\pi] \\ =& -\frac{1}{2}\Big[- \sum_{x,y\in E} \pi(x)L(x,y)[f_s(y) - f_s(x)]^2 \Big] \\ =& \frac{1}{2}\sum_{x,y\in E} \pi(x)L(x,y)[f_s(y) - f_s(x)]^2.\end{aligned}$$

Notice that if $\mathcal{D}(f) = 0$ and the process is indecomposable, then f is constant.

Proposition

If a function $F : \mathbb{R} \to \mathbb{R}$ is a contraction, (i.e., $|F(a) - F(b)| \leq |a - b|$), then

$$\mathcal{D}(F \circ f) \leq \mathcal{D}(f). \tag{3}$$

Indeed, we have

$$\begin{aligned}
\mathcal{D}(F\circ f) &= \frac{1}{2}\sum_{x,y\in E}\pi(x)L(x,y)[(F\circ f)(y)-(F\circ f)(x)]^2\\
&= \frac{1}{2}\sum_{x,y\in E}\pi(x)L(x,y)|F\big(f(y)\big)-F\big(f(x)\big)|^2\\
&\le \frac{1}{2}\sum_{x,y\in E}\pi(x)L(x,y)|f(y)-f(x)|^2\\
&= \frac{1}{2}\sum_{x,y\in E}\pi(x)L(x,y)[f(y)-f(x)]^2=\mathcal{D}(f).
\end{aligned}$$

Proposition

Let M be a fixed real number. Then the function $F(x) = \min\{x, M\}$ is a contraction .

If M is a fixed real number and $F(x) = \min\{x, M\}$, we have

$$|F(a) - F(b)| = \begin{cases} |M - M| = |0| = 0 \leq |a - b|, \text{ if } a \geq M \text{ and } b \geq M; \\ |M - b| = M - b \leq a - b = |a - b|, \text{ if } a \geq M \text{ and } b < M; \\ |a - M| = M - a \leq b - a = |a - b|, \text{ if } a < M \text{ and } b \geq M; \\ |a - b| \leq |a - b|, \text{ if } a < M \text{ and } b < M; \end{cases}$$

Proposition

The function $F(x) = |x|$ is a contraction.

Indeed, we have

$$|F(a) - F(b)| = \big||a| - |b|\big| \leq |a - b|.$$

Another interesting result is the convexity of the Dirichlet form.

Proposition

Let p be a probability measure on $\mathbb{N}$ and $(f_j)_{j\in\mathbb{N}} \subset L^2(\pi)$. Then

$$\mathcal{D}\Big(\sum_{j\in\mathbb{N}} p_j f_j\Big) \leq \sum_{j\in\mathbb{N}} p_j \mathcal{D}(f_j).$$

For every $(x, y) \in E^2$, let $\alpha_{x,y} : \mathbb{N} \to \mathbb{R}$ be the random variable which is $f_y(y) - f_j(x)$ with probability p_j. Then

$$\begin{aligned}\Big(E[\alpha_{x,y}]\Big)^2 =& \Big(\sum_{j\in\mathbb{N}} p_j[f_y(y) - f_j(x)]\Big)^2 \\ \leq& \sum_{j\in\mathbb{N}} p_j[f_j(y) - f_j(x)]^2 \\ =& E[\alpha^2_{x,y}], \forall x, y \in E.\end{aligned}$$

This leads to

$$\begin{aligned}
\mathcal{D}\Big(\sum_{j\in\mathbb{N}} p_j f_j\Big) &= \frac{1}{2}\sum_{x,y\in E}\pi(x)L(x,y)\Big[\Big(\sum_{j\in\mathbb{N}} p_j f_j\Big)(y) - \Big(\sum_{j\in\mathbb{N}} p_j f_j\Big)(x)\Big]^2 \\
&= \frac{1}{2}\sum_{x,y\in E}\pi(x)L(x,y)\Big[\sum_{j\in\mathbb{N}} p_j f_j(y) - \sum_{j\in\mathbb{N}} p_j f_j(x)\Big]^2 \\
&= \frac{1}{2}\sum_{x,y\in E}\pi(x)L(x,y)\Big[\sum_{j\in\mathbb{N}} p_j [f_j(y) - f_j(x)]\Big]^2 \\
&\leq \frac{1}{2}\sum_{x,y\in E}\pi(x)L(x,y)\sum_{j\in\mathbb{N}} p_j [f_j(y) - f_j(x)]^2 \\
&= \sum_{j\in\mathbb{N}} p_j\Big(\frac{1}{2}\sum_{x,y\in E}\pi(x)L(x,y)[f_j(y) - f_j(x)]^2\Big),
\end{aligned}$$

which leads to

$$\mathcal{D}\Big(\sum_{j\in\mathbb{N}} p_j f_j\Big) \leq \sum_{j\in\mathbb{N}} p_j \mathcal{D}(f_j).$$

If the probability measure π is reversible, there exists a variational formula for the Dirichlet form $\mathcal{D}(f)$.

Theorem

*(**Variational formula for the Dirichlet form**) Assume π is reversible. For every non-negative function $f \in L^2(\pi)$,*

$$\mathcal{D}(f) := -\inf_g \sum_{x \in E} \pi(x) \frac{f^2(x)}{g(x)} Lg(x).$$

In this formula, the infimum is taken over all bounded positive functions g which are bounded below by a strictly positive constant.

Fix a function g that is bounded and is bounded below by a strictly positive constant. Set $\alpha = \frac{g}{f}\mathbb{1}\{f > 0\}$, so that $\alpha(x) = 0$ if and only if $f(x) = 0$. With this definition,

$$< \frac{f^2}{g}, P_t g >_\pi = \sum_{x \in E} \frac{f^2(x)}{g(x)}(P_t g)(x)\pi(x)$$

$$= \sum_{\substack{x \in E \\ f(x)>0}} \frac{f^2(x)}{g(x)}\Big(\sum_{y \in E} P_t(x,y)g(y)\Big)\pi(x)$$

$$= \sum_{\substack{x \in E \\ f(x)>0}} \frac{f(x)}{\frac{g(x)}{f(x)}}\Big(\sum_{\substack{y \in E \\ f(y)>0}} P_t(x,y)f(y)\frac{g(y)}{f(y)}\Big)\pi(x)$$

$$= \sum_{x \in E} \frac{f(x)}{\frac{g(x)}{f(x)}\mathbb{1}\{f(x) > 0\}}\Big(\sum_{y \in E} P_t(x,y)f(y)\frac{g(y)}{f(y)}\mathbb{1}\{f(y) > 0\}\Big)\pi(x).$$

Then, we have

$$
\begin{aligned}
&< \frac{f^2}{g}, P_t g >_\pi \\
&= \sum_{x\in E} \frac{f(x)}{\frac{g(x)}{f(x)} \mathbb{1}\{f(x) > 0\}} \Big(\sum_{y\in E} P_t(x,y) f(y) \frac{g(y)}{f(y)} \mathbb{1}\{f(y) > 0\} \Big) \pi(x) \\
&= \sum_{x\in E} \frac{f(x)}{\alpha(x)} \Big(\sum_{y\in E} P_t(x,y) f(y) \alpha(y) \Big) \pi(x) \\
&= \sum_{x\in E} \Big(\frac{f}{\alpha} \Big)(x) \Big(\sum_{y\in E} P_t(x,y) (f\alpha)(y) \Big) \pi(x) \\
&= \sum_{x\in E} \Big(\frac{f}{\alpha} \Big)(x) \big(P_t(f\alpha) \big)(x) \pi(x) = < \frac{f}{\alpha}, P_t(f\alpha) >_\pi, \forall t > 0.
\end{aligned}
$$

Since the probability measure π is reversible, P_t is self-adjoint, which leads to

$$
\begin{aligned}
&< \frac{f^2}{g}, P_t g >_\pi = < \frac{f}{\alpha}, P_t(f\alpha) >_\pi \\
=& \frac{1}{2}\Big[< \frac{f}{\alpha}, P_t(f\alpha) >_\pi + < \frac{f}{\alpha}, P_t(f\alpha) >_\pi \Big] \\
=& \frac{1}{2}\Big[< \frac{f}{\alpha}, P_t(f\alpha) >_\pi + < f\alpha, P_t\Big(\frac{f}{\alpha}\Big) >_\pi \Big] \\
=& \frac{1}{2}\Big[\sum_{x\in E} \Big(\frac{f}{\alpha}\Big)(x)\big(P_t(f\alpha)\big)(x)\pi(x) + \sum_{x\in E}(f\alpha)(x)\Big(P_t\Big(\frac{f}{\alpha}\Big)\Big)(x)\pi(x)\Big] \\
=& \frac{1}{2} \sum_{\substack{x\in E \\ f(x)>0}} \Big[\Big(\frac{f}{\alpha}\Big)(x)\big(P_t(f\alpha)\big)(x) + (f\alpha)(x)\Big(P_t\Big(\frac{f}{\alpha}\Big)\Big)(x)\Big]\pi(x).
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
&< \frac{f^2}{g}, P_t g >_\pi \\
=&\frac{1}{2} \sum_{\substack{x \in E \\ f(x)>0}} \Big[\Big(\frac{f}{\alpha}\Big)(x) \big(P_t(f\alpha)\big)(x) + (f\alpha)(x) \Big(P_t\Big(\frac{f}{\alpha}\Big)\Big)(x) \Big] \pi(x) \\
=&\frac{1}{2} \sum_{\substack{x \in E \\ f(x)>0}} \Big[\frac{f(x)}{\alpha(x)} \sum_{\substack{y \in E \\ f(y)>0}} P_t(x,y) f(y) \alpha(y) + f(x)\alpha(x) \sum_{\substack{y \in E \\ f(y)>0}} P_t(x,y) \frac{f(y)}{\alpha(y)} \Big] \pi(x) \\
=&\frac{1}{2} \sum_{\substack{x,y \in E \\ f(x)f(y)>0}} \pi(x) f(x) f(y) P_t(x,y) \Big(\frac{\alpha(y)}{\alpha(x)} + \frac{\alpha(x)}{\alpha(y)} \Big).
\end{aligned}
$$

Since $x + x^{-1} \geq 2, \forall x > 0$, we get

$$< \frac{f^2}{g}, P_t g >_\pi$$

$$= \frac{1}{2} \sum_{\substack{x,y \in E \\ f(x)f(y)>0}} \pi(x) f(x) f(y) P_t(x,y) \Big(\frac{\alpha(y)}{\alpha(x)} + \frac{\alpha(x)}{\alpha(y)} \Big)$$

$$\geq \frac{1}{2} \sum_{\substack{x,y \in E \\ f(x)f(y)>0}} \pi(x) f(x) f(y) P_t(x,y) 2$$

$$= \sum_{\substack{x,y \in E \\ f(x)f(y)>0}} \pi(x) f(x) f(y) P_t(x,y) = \sum_{x,y \in E} \pi(x) f(x) f(y) P_t(x,y)$$

$$= \sum_{x \in E} f(x) \Big(\sum_{y \in E} P_t(x,y) f(y) \Big) \pi(x) = \sum_{x \in E} f(x) (P_t f)(x) \pi(x) =< f, P_t f >_\pi .$$

Then, we have

$$< \frac{f^2}{g}, P_t g >_\pi \;\geq\; < f, P_t f >_\pi, \forall t > 0,$$

and we get

$$-\frac{1}{t} < \frac{f^2}{g}, P_t g >_\pi \;\leq\; -\frac{1}{t} < f, P_t f >_\pi, \forall t > 0.$$

This leads to

$$\begin{aligned}\frac{1}{t} < \frac{f^2}{g}, (g - P_t g) >_\pi &= \frac{1}{t} < \frac{f^2}{g}, g >_\pi - \frac{1}{t} < \frac{f^2}{g}, P_t g >_\pi \\ &= \frac{1}{t} < f, f >_\pi - \frac{1}{t} < \frac{f^2}{g}, P_t g >_\pi \\ &\leq \frac{1}{t} < f, f >_\pi - \frac{1}{t} < f, P_t f >_\pi \\ &= \frac{1}{t} < f, (f - P_t f) >_\pi, \forall t > 0.\end{aligned}$$

Since g is bounded, we have that the sequence $\{t^{-1}(P_t g - g), t > 0\}$ converges uniformly to Lg as $t \downarrow 0$. Since g is bounded below by a strictly positive constant, there is $C > 0$ such that $\frac{1}{g(x)} < C, \forall x \in E$. We claim that

$$\lim_{t \to 0^+} \frac{1}{t} < \frac{f^2}{g}, (g - P_t g) >_\pi = < \frac{f^2}{g}, -Lg >_\pi .$$

Indeed, let $\varepsilon > 0$. Since the sequence $\{t^{-1}(P_t g - g), t > 0\}$ converges uniformly to Lg as $t \downarrow 0$, there exists $t_0 > 0$ such that

$$|t^{-1}(P_t g - g)(x) - (Lg)(x)| < \frac{\varepsilon}{C(< f, f >_\pi + 1)}, \forall x \in E, \forall\, 0 < t < t_0.$$

Then, for all $0 < t < t_0$, we have

$$\Big|\frac{1}{t} < \frac{f^2}{g}, (g - P_t g) >_\pi - < \frac{f^2}{g}, -Lg >_\pi \Big|$$

$$= \Big| < \frac{f^2}{g}, t^{-1}(P_t g - g) - Lg >_\pi \Big|$$

$$= \Big| \sum_{x \in E} f^2(x) \frac{1}{g(x)} \big(t^{-1}(P_t g - g)(x) - (Lg)(x)\big) \pi(x) \Big|$$

$$\leq \sum_{x \in E} f^2(x) \frac{1}{g(x)} |t^{-1}(P_t g - g)(x) - (Lg)(x)| \pi(x)$$

$$< \sum_{x \in E} f^2(x) C \frac{\varepsilon}{C(< f, f >_\pi + 1)} \pi(x) = \frac{\varepsilon < f, f >_\pi}{(< f, f >_\pi + 1)} < \varepsilon,$$

and we have

$$\lim_{t \to 0^+} \frac{1}{t} < \frac{f^2}{g}, (g - P_t g) >_\pi = < \frac{f^2}{g}, -Lg >_\pi .$$

Since $f \in L^2(\pi)$, we have that

$$\lim_{t \to 0^+} < t^{-1}(f - P_t f) - Lf, t^{-1}(f - P_t f) - Lf >_\pi = 0.$$

From Holder inequality, we have

$$\begin{aligned} & \left| \frac{1}{t} < f, (f - P_t f) >_\pi - < f, -Lf >_\pi \right|^2 \\ = & < f, t^{-1}(f - P_t f) - Lf >_\pi^2 \\ \leq & < f, f >_\pi < t^{-1}(f - P_t f) - Lf, t^{-1}(f - P_t f) - Lf >_\pi , \end{aligned}$$

which leads to

$$\lim_{t \to 0^+} \left| \frac{1}{t} < f, (f - P_t f) >_\pi - < f, -Lf >_\pi \right|^2 = 0,$$

which is the same as

$$\lim_{t \to 0^+} \frac{1}{t} < f, (f - P_t f) >_\pi = < f, -Lf >_\pi = - < f, Lf >_\pi = \mathcal{D}(f).$$

Since

$$\frac{1}{t} < \frac{f^2}{g}, (g - P_t g) >_\pi \leq \frac{1}{t} < f, (f - P_t f) >_\pi, \forall t > 0,$$

Making $t \to 0^+$, we get

$$- < \frac{f^2}{g}, Lg >_\pi = < \frac{f^2}{g}, -Lg >_\pi = \lim_{t \to 0^+} \frac{1}{t} < \frac{f^2}{g}, (g - P_t g) >_\pi$$
$$\leq \lim_{t \to 0^+} \frac{1}{t} < f, (f - P_t f) >_\pi = \mathcal{D}(f).$$

This is the same as

$$\mathcal{D}(f) \geq - \sum_{x \in E} \pi(x) \frac{f^2(x)}{g(x)} Lg(x).$$

Since g is a arbitrary function that is bounded and is bounded below by a strictly positive constant, we get

$$\mathcal{D}(f) \geq \sup_{g} - \sum_{x \in E} \pi(x) \frac{f^2(x)}{g(x)} Lg(x) = - \inf_{g} \sum_{x \in E} \pi(x) \frac{f^2(x)}{g(x)} Lg(x),$$

where we take the supremum and the infimum over all bounded positive functions g which are bounded below by a strictly positive constant.

If f is bounded and bounded below by a strictly positive constant, we can make $g = f$, which leads to

$$-\inf_g \sum_{x\in E} \pi(x)\frac{f^2(x)}{g(x)} Lg(x) = \sup_g - \sum_{x\in E} \pi(x)\frac{f^2(x)}{g(x)} Lg(x)$$

$$\geq - \sum_{x\in E} \pi(x)\frac{f^2(x)}{f(x)} Lf(x) = - \sum_{x\in E} \pi(x) f(x) Lf(x) = - < f, Lf >_\pi = \mathcal{D}(f),$$

which leads to

$$\mathcal{D}(f) \leq -\inf_g \sum_{x\in E} \pi(x)\frac{f^2(x)}{g(x)} Lg(x),$$

where we take the infimum over all bounded positive functions g which are bounded below by a strictly positive constant.

However, in the general case f is neither bounded or bounded below by a strictly positive constant. In this case, we need to approximate f by bounded positive functions bounded below by strictly positive constants.

For each positive integer M, let $f_M : E \to \mathbb{R}$ be the function defined by

$$f_M(x) = M^{-1} + \min\{f(x), M\}, \forall x \in E.$$

Since f is positive, $0 \leq \min\{f(x), M\} \leq M, \forall x \in E$, which leads to

$$M^{-1} = M^{-1} + 0 \leq M^{-1} + \min\{f(x), M\} = f_M(x) \leq M^{-1} + M, \forall x \in E.$$

Then, f_M is bounded and bounded below by a strictly positive constant, for every $M \in \mathbb{N}$.

We claim that $\lim_{M\to\infty} f_M(x) = f(x), \forall x \in E$. Indeed, let $x \in E$. Also, let $\varepsilon > 0$. Choosing M_0 such that $M_0 > \max\{\varepsilon^{-1}, f(x)\}$, we get

$$f_M(x) = M^{-1} + \min\{f(x), M\} = M^{-1} + f(x) > f(x), \forall M > M_0,$$

which leads to

$$|f_M(x) - f(x)| = f_M(x) - f(x) = M^{-1} < \varepsilon, \forall M > M_0.$$

Since for every $\varepsilon > 0$, there exists $M_0 \in \mathbb{N}$ such that $|f_M(x) - f(x)| < \varepsilon, \forall M > M_0$, we have $\lim_{M\to\infty} f_M(x) = f(x)$. Since $x \in E$ is arbitrary, we have

$$\lim_{M\to\infty} f_M(x) = f(x), \forall x \in X.$$

Define the measure $\mu : E^2 \to \mathbb{R}$ on E^2 by

$$\mu(x, y) = \begin{cases} \frac{\pi(x)L(x,y)}{2} = \frac{\pi(x)\lambda(x)p(x,y)}{2}, & \text{if } x \neq y; \\ 0, & \text{if } x = y; \end{cases}$$

Define $F : E^2 \to \mathbb{R}$ by

$$F(x, y) = \big(F(y) - F(x)\big)^2, \forall x, y \in E.$$

Then $F(x,y) \geq 0$ and we have

$$\mathcal{D}(f) = \frac{1}{2} \sum_{x,y \in E} \pi(x) L(x,y)[f(y) - f(x)]^2 = \sum_{\substack{x,y \in E \\ x \neq y}} \frac{\pi(x)L(x,y)}{2} F(x,y)$$

$$= \sum_{\substack{x,y \in E \\ x \neq y}} \mu(x,y) F(x,y) = \sum_{(x,y) \in E^2} \mu(x,y) F(x,y) = \int_{E^2} F d\mu.$$

For each positive integer M, let $F_M : E^2 \to \mathbb{R}$ be the function defined by

$$F_M(x,y) = \Big[\frac{\big(f(y)\big)^2}{f_M(y)} - \frac{\big(f(x)\big)^2}{f_M(x)} \Big] [f_M(y) - f_M(x)], \forall (x,y) \in E^2.$$

Since f is positive, for every $(x, y) \in E^2$, we get

$$\begin{aligned}
\lim_{M\to\infty} F_M(x,y) &= \lim_{M\to\infty}\Big[\frac{\big(f(y)\big)^2}{f_M(y)} - \frac{\big(f(x)\big)^2}{f_M(x)}\Big][f_M(y) - f_M(x)] \\
&=\Big[\frac{\big(f(y)\big)^2}{\lim_{M\to\infty} f_M(y)} - \frac{\big(f(x)\big)^2}{\lim_{M\to\infty} f_M(x)}\Big][\lim_{M\to\infty} f_M(y) - \lim_{M\to\infty} f_M(x)] \\
&=\Big[\frac{\big(f(y)\big)^2}{f(y)} - \frac{\big(f(x)\big)^2}{f(x)}\Big][f(y) - f(x)] \\
&=[f(y) - f(x)][f(y) - f(x)] = [f(y) - f(x)]^2 = F(x,y).
\end{aligned}$$

We claim that $F_M(x, y) \geq 0, \forall x, y \in E$. For every $(x, y) \in E^2$, we have four possibilities: $f(x) \geq M$ and $f(y) \geq M$ (case 1), $f(x) \geq M$ and $f(y) < M$ (case 2), $f(x) < M$ and $f(y) \geq M$ (case 3), $f(x) < M$ and $f(y) < M$ (case 4).

Case 1: $f(x) \geq M$ and $f(y) \geq M$. In this case, we have

$$f_M(x) = M^{-1} + \min\{f(x), M\} = M^{-1} + M$$

and

$$f_M(y) = M^{-1} + \min\{f(y), M\} = M^{-1} + M.$$

Then we get

$$\begin{aligned} F_M(x,y) =& \Big[\frac{(f(y))^2}{f_M(y)} - \frac{(f(x))^2}{f_M(x)} \Big] [f_M(y) - f_M(x)] \\ =& \Big[\frac{(f(y))^2}{M^{-1}+M} - \frac{(f(x))^2}{M^{-1}+M} \Big] [M^{-1} + M - M^{-1} - M] \\ =& \Big[\frac{(f(y))^2 - (f(x))^2}{M^{-1}+M} \Big] \cdot 0 = 0. \end{aligned}$$

Case 2: $f(x) \geq M$ and $f(y) < M$. In this case, we have

$$f_M(x) = M^{-1} + \min\{f(x), M\} = M^{-1} + M$$

and

$$f_M(y) = M^{-1} + \min\{f(y), M\} = M^{-1} + f(y).$$

Then we get

$$F_M(x,y) = \Big[\frac{\big(f(y)\big)^2}{f_M(y)} - \frac{\big(f(x)\big)^2}{f_M(x)}\Big][f_M(y) - f_M(x)]$$

$$= \Big[\frac{\big(f(y)\big)^2}{M^{-1} + f(y)} - \frac{\big(f(x)\big)^2}{M^{-1} + M}\Big][M^{-1} + f(y) - M^{-1} - M]$$

$$= \frac{f(y) - M}{\big(M^{-1} + f(y)\big)\big(M^{-1} + M\big)}\Big[\big(M^{-1} + M\big)\big(f(y)\big)^2 - \big(M^{-1} + f(y)\big)\big(f(x)\big)^2\Big].$$

Since $f(y) - M < 0$ and $f(y) > 0$, we have $\frac{f(y)-M}{\left(M^{-1}+f(y)\right)\left(M^{-1}+M\right)} < 0$.

Also,

$$\begin{aligned}
&\left(M^{-1}+M\right)\left(f(y)\right)^2 - \left(M^{-1}+f(y)\right)\left(f(x)\right)^2 \\
=&M^{-1}\left(f(y)\right)^2 + M\left(f(y)\right)^2 - M^{-1}\left(f(x)\right)^2 - f(y)\left(f(x)\right)^2 \\
=&M^{-1}\left[\left(f(y)\right)^2 - \left(f(x)\right)^2\right] + f(y)\left[Mf(y) - \left(f(x)\right)^2\right] < 0,
\end{aligned}$$

since $f(y) < M \leq f(x)$ leads to $\left(f(y)\right)^2 - \left(f(x)\right)^2 < 0$ and to

$$Mf(y) - \left(f(x)\right)^2 < M.M - \left(f(x)\right)^2 \leq 0.$$

Therefore,

$$\begin{aligned}
&\frac{f(y)-M}{\left(M^{-1}+f(y)\right)\left(M^{-1}+M\right)}\left[\left(M^{-1}+M\right)\left(f(y)\right)^2 - \left(M^{-1}+f(y)\right)\left(f(x)\right)^2\right] \\
=&F_M(x,y) > 0.
\end{aligned}$$

Case 3: $f(x) < M$ and $f(y) \geq M$. In this case, we have

$$f_M(x) = M^{-1} + \min\{f(x), M\} = M^{-1} + f(x)$$

and

$$f_M(y) = M^{-1} + \min\{f(y), M\} = M^{-1} + M.$$

Then we get

$$F_M(x,y) = \Big[\frac{\big(f(y)\big)^2}{f_M(y)} - \frac{\big(f(x)\big)^2}{f_M(x)}\Big][f_M(y) - f_M(x)]$$

$$= \Big[\frac{\big(f(y)\big)^2}{M^{-1}+M} - \frac{\big(f(x)\big)^2}{M^{-1}+f(x)}\Big][M^{-1} + M - M^{-1} - f(x)]$$

$$= \frac{M - f(x)}{\big(M^{-1}+M\big)\big(M^{-1}+f(x)\big)}\Big[\big(M^{-1}+f(x)\big)\big(f(y)\big)^2 - \big(M^{-1}+M\big)\big(f(x)\big)^2\Big].$$

Since $M - f(x) > 0$ and $f(x) > 0$, we have $\frac{M-f(x)}{\left(M^{-1}+M\right)\left(M^{-1}+f(x)\right)} > 0$.

Also,

$$\begin{aligned}
&\left(M^{-1} + f(x)\right)\left(f(y)\right)^2 - \left(M^{-1} + M\right)\left(f(x)\right)^2 \\
=&M^{-1}\left(f(y)\right)^2 + f(x)\left(f(y)\right)^2 - M^{-1}\left(f(x)\right)^2 - M\left(f(x)\right)^2 \\
=&M^{-1}\left[\left(f(y)\right)^2 - \left(f(x)\right)^2\right] + f(x)\left[\left(f(y)\right)^2 - Mf(x)\right] > 0,
\end{aligned}$$

since $f(x) < M \leq f(y)$ leads to $\left(f(y)\right)^2 - \left(f(x)\right)^2 > 0$ and to

$$\left(f(y)\right)^2 - Mf(x) > \left(f(y)\right)^2 - MM \geq 0.$$

Therefore,

$$\begin{aligned}
&\frac{M - f(x)}{\left(M^{-1} + M\right)\left(M^{-1} + f(x)\right)}\left[\left(M^{-1} + f(x)\right)\left(f(y)\right)^2 - \left(M^{-1} + M\right)\left(f(x)\right)^2\right] \\
=&F_M(x, y) > 0.
\end{aligned}$$

Case 4: $f(x) < M$ and $f(y) < M$. In this case, we have

$$f_M(x) = M^{-1} + \min\{f(x), M\} = M^{-1} + f(x)$$

and

$$f_M(y) = M^{-1} + \min\{f(y), M\} = M^{-1} + f(y).$$

Then, we have

$$F_M(x,y) = \Big[\frac{\big(f(y)\big)^2}{f_M(y)} - \frac{\big(f(x)\big)^2}{f_M(x)}\Big][f_M(y) - f_M(x)]$$

$$= \Big[\frac{\big(f(y)\big)^2}{M^{-1} + f(y)} - \frac{\big(f(x)\big)^2}{M^{-1} + f(x)}\Big][M^{-1} + f(y) - M^{-1} - f(x)]$$

$$= \frac{f(y) - f(x)}{\big(M^{-1} + f(y)\big)\big(M^{-1} + f(x)\big)}\Big[\big(M^{-1} + f(x)\big)\big(f(y)\big)^2 - \big(M^{-1} + f(y)\big)\big(f(x)\big)^2\Big].$$

Then, $F_M(x,y)$ is equal to

$$\frac{f(y)-f(x)}{\left(M^{-1}+f(y)\right)\left(M^{-1}+f(x)\right)}\left[\left(M^{-1}+f(x)\right)\left(f(y)\right)^2-\left(M^{-1}+f(y)\right)\left(f(x)\right)^2\right]$$

$$=\frac{[f(y)-f(x)]\left[M^{-1}\left(f(y)\right)^2+f(x)\left(f(y)\right)^2-M^{-1}\left(f(x)\right)^2-f(y)\left(f(x)\right)^2\right]}{\left(M^{-1}+f(y)\right)\left(M^{-1}+f(x)\right)}$$

$$=\frac{[f(y)-f(x)]\left[M^{-1}\left[\left(f(y)\right)^2-\left(f(x)\right)^2\right]+f(x)f(y)[f(y)-f(x)]\right]}{\left(M^{-1}+f(y)\right)\left(M^{-1}+f(x)\right)}$$

$$=\frac{[f(y)-f(x)]\left[M^{-1}[f(y)+f(x)][f(y)-f(x)]+f(x)f(y)[f(y)-f(x)]\right]}{\left(M^{-1}+f(y)\right)\left(M^{-1}+f(x)\right)}$$

$$=\frac{[f(y)-f(x)]^2\left[M^{-1}[f(y)+f(x)]+f(x)f(y)\right]}{\left(M^{-1}+f(y)\right)\left(M^{-1}+f(x)\right)}\geq 0.$$

Therefore, we have $F_M(x,y)\geq 0$, $\forall (x,y)\in E^2$, $\forall M\in\mathbb{N}$.

We know that

$$- < \frac{f^2}{f_M}, Lf_M >_\pi = -\sum_{x\in E} \pi(x)\Big(\frac{f^2}{f_M}\Big)(x)(Lf_M)(x)$$

$$= -\sum_{x\in E} \pi(x)\frac{\big(f(x)\big)^2}{f_M(x)} \sum_{y\in E} L(x,y)[f_M(y) - f_M(x)]$$

$$= -\sum_{x\in E}\sum_{y\in E} \pi(x)L(x,y)\frac{\big(f(x)\big)^2}{f_M(x)}[f_M(y) - f_M(x)]$$

$$= -\frac{1}{2}\sum_{x\in E}\sum_{y\in E} \pi(x)L(x,y)\frac{\big(f(x)\big)^2}{f_M(x)}[f_M(y) - f_M(x)]$$

$$-\frac{1}{2}\sum_{x\in E}\sum_{y\in E} \pi(x)L(x,y)\frac{\big(f(x)\big)^2}{f_M(x)}[f_M(y) - f_M(x)].$$

Interchanging the variables in the second double summation, we get

$$-< \frac{f^2}{f_M}, Lf_M >_\pi = -\frac{1}{2}\sum_{x\in E}\sum_{y\in E}\pi(x)L(x,y)\frac{\left(f(x)\right)^2}{f_M(x)}[f_M(y)-f_M(x)]$$

$$-\frac{1}{2}\sum_{x\in E}\sum_{y\in E}\pi(x)L(x,y)\frac{\left(f(x)\right)^2}{f_M(x)}[f_M(y)-f_M(x)]$$

$$= -\frac{1}{2}\sum_{x\in E}\sum_{y\in E}\pi(x)L(x,y)\frac{\left(f(x)\right)^2}{f_M(x)}[f_M(y)-f_M(x)]$$

$$-\frac{1}{2}\sum_{y\in E}\sum_{x\in E}\pi(y)L(y,x)\frac{\left(f(y)\right)^2}{f_M(y)}[f_M(x)-f_M(y)].$$

Since π is reversible, $\pi(x)L(x,y) = \pi(y)L(y,x) \forall x, y \in E$, and

$$
\begin{aligned}
-<\frac{f^2}{f_M}, Lf_M>_\pi = &-\frac{1}{2}\sum_{x\in E}\sum_{y\in E}\pi(x)L(x,y)\frac{(f(x))^2}{f_M(x)}[f_M(y)-f_M(x)] \\
&-\frac{1}{2}\sum_{y\in E}\sum_{x\in E}\pi(y)L(y,x)\frac{(f(y))^2}{f_M(y)}[f_M(x)-f_M(y)] \\
=&\frac{1}{2}\sum_{x\in E}\sum_{y\in E}\pi(x)L(x,y)\Big(-\frac{(f(x))^2}{f_M(x)}\Big)[f_M(y)-f_M(x)] \\
&-\frac{1}{2}\sum_{y\in E}\sum_{x\in E}\pi(x)L(x,y)\frac{(f(y))^2}{f_M(y)}[f_M(x)-f_M(y)].
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
-<\frac{f^2}{f_M}, Lf_M>_\pi =& \frac{1}{2}\sum_{x\in E}\sum_{y\in E}\pi(x)L(x,y)\Big(-\frac{\big(f(x)\big)^2}{f_M(x)}\Big)[f_M(y)-f_M(x)] \\
&-\frac{1}{2}\sum_{y\in E}\sum_{x\in E}\pi(x)L(x,y)\frac{\big(f(y)\big)^2}{f_M(y)}[f_M(x)-f_M(y)] \\
=& \frac{1}{2}\sum_{x\in E}\sum_{y\in E}\pi(x)L(x,y)\Big(-\frac{\big(f(x)\big)^2}{f_M(x)}\Big)[f_M(y)-f_M(x)] \\
&+\frac{1}{2}\sum_{x\in E}\sum_{y\in E}\pi(x)L(x,y)\frac{\big(f(y)\big)^2}{f_M(y)}[f_M(y)-f_M(x)] \\
=& \frac{1}{2}\sum_{x,y\in E}\pi(x)L(x,y)\Big[\frac{\big(f(y)\big)^2}{f_M(y)}-\frac{\big(f(x)\big)^2}{f_M(x)}\Big][f_M(y)-f_M(x)].
\end{aligned}
$$

This leads to

$$
\begin{aligned}
-<\frac{f^2}{f_M}, Lf_M>_\pi &= \frac{1}{2}\sum_{x,y\in E}\pi(x)L(x,y)\Big[\frac{\big(f(y)\big)^2}{f_M(y)}-\frac{\big(f(x)\big)^2}{f_M(x)}\Big][f_M(y)-f_M(x)] \\
&= \sum_{\substack{x,y\in E\\ x\neq y}}\frac{\pi(x)L(x,y)}{2}\Big[\frac{\big(f(y)\big)^2}{f_M(y)}-\frac{\big(f(x)\big)^2}{f_M(x)}\Big][f_M(y)-f_M(x)] \\
&= \sum_{\substack{x,y\in E\\ x\neq y}}\mu(x,y)\Big[\frac{\big(f(y)\big)^2}{f_M(y)}-\frac{\big(f(x)\big)^2}{f_M(x)}\Big][f_M(y)-f_M(x)] \\
&= \sum_{\substack{x,y\in E\\ x\neq y}}\mu(x,y)F_M(x,y) \\
&= \sum_{(x,y)\in E^2}F_M(x,y)\mu(x,y) = \int_{E^2}F_M\,d\mu.
\end{aligned}
$$

Since $F_M(x,y) \geq 0, \forall (x,y) \in E^2, \forall M \in \mathbb{N}$, and we have

$$\liminf_{M\to\infty} F_M(x,y) = \lim_{M\to\infty} F_M(x,y) = F(x,y), \forall (x,y) \in E^2,$$

Fatou's Lemma gives

$$\begin{aligned}\mathcal{D}(f) &= \int_{E^2} F d\mu = \int_{E^2} \liminf_{M\to\infty} F_M d\mu \\ &\leq \liminf_{M\to\infty} \int_{E^2} F_M d\mu = \liminf_{M\to\infty} - < \frac{f^2}{f_M}, Lf_M >_\pi .\end{aligned}$$

This leads us to

$$\begin{aligned}\mathcal{D}(f) &\leq \liminf_{M\to\infty} - < \frac{f^2}{f_M}, Lf_M >_\pi = \sup_{k\in\mathbb{N}} \inf_{M\geq k} - < \frac{f^2}{f_M}, Lf_M >_\pi \\ &\leq \sup_{k\in\mathbb{N}} - < \frac{f^2}{f_M}, Lf_M >_\pi \leq \sup_{g} - < \frac{f^2}{g}, Lg >_\pi \\ &= \sup_{g} - \sum_{x\in E} \pi(x) \frac{f^2(x)}{g(x)} Lg(x),\end{aligned}$$

leading to

$$\mathcal{D}(f) \leq \sup_{g} - \sum_{x\in E} \pi(x) \frac{f^2(x)}{g(x)} Lg(x) = - \inf_{g} \sum_{x\in E} \pi(x) \frac{f^2(x)}{g(x)} Lg(x),$$

where we take the supremum and the infimum over all bounded positive functions g which are bounded below by a strictly positive constant.

Finally, we have

$$\mathcal{D}(f) = -\inf_{g} \sum_{x \in E} \pi(x) \frac{f^2(x)}{g(x)} Lg(x),$$

where we take the infimum over all bounded positive functions g which are bounded below by a strictly positive constant.

The next result is a simple consequence of this proposition.

Corollary

If π is reversible, the functional

$$D(f) = \mathcal{D}(\sqrt{f})$$

defined for all densities with respect to π is convex and lower semicontinuous.

Since π is reversible, we have by the previous result that

$$D(f) = \mathcal{D}(\sqrt{f}) = -\inf_g \sum_{x \in E} \pi(x) \frac{f(x)}{g(x)} (Lg)(x) = \sup_g - \sum_{x \in E} \pi(x) \frac{f(x)}{g(x)} (Lg)(x),$$

where we take the supremum and the infimum over all bounded positive functions g which are bounded below by a strictly positive constant.

If $\alpha \in [0,1]$ and f_1, f_2 are densities with respect to π, we have

$$D(\alpha f_1 + (1-\alpha)f_2) = \sup_g - \sum_{x\in E} \pi(x) \frac{(\alpha f_1 + (1-\alpha)f_2)(x)}{g(x)} (Lg)(x)$$

$$= \sup_g \Big[-\alpha \sum_{x\in E} \pi(x) \frac{f_1(x)}{g(x)} (Lg)(x) + \Big(-(1-\alpha) \sum_{x\in E} \pi(x) \frac{f_2(x)}{g(x)} (Lg)(x) \Big) \Big]$$

$$\leq \sup_g -\alpha \sum_{x\in E} \pi(x) \frac{f_1(x)}{g(x)} (Lg)(x) + \sup_g -(1-\alpha) \sum_{x\in E} \pi(x) \frac{f_2(x)}{g(x)} (Lg)(x)$$

$$= \alpha \sup_g - \sum_{x\in E} \pi(x) \frac{f_1(x)}{g(x)} (Lg)(x) + (1-\alpha) \sup_g - \sum_{x\in E} \pi(x) \frac{f_2(x)}{g(x)} (Lg)(x)$$

$$= \alpha D(f_1) + (1-\alpha) D(f_2).$$

Since for every f_1, f_2 densities with respect to π, we have

$$D(\alpha f_1 + (1-\alpha)f_2) \leq \alpha D(f_1) + (1-\alpha) D(f_2), \forall \alpha \in [0,1],$$

the functional $D(f)$ is convex.

Let$(f_n)_{n\in\mathbb{N}}$ be a sequence of densities with respect to π such that f_n converges weakly to f. Assume that $D(f) > \liminf_{n\to\infty} D(f_n)$. In this case, choose

$$\varepsilon = \frac{D(f) - \liminf_{n\to\infty} D(f_n)}{3} > 0.$$

Since $D(f) = \sup_g -\sum_{x\in E} \pi(x)\frac{f(x)}{g(x)}(Lg)(x)$ over all bounded positive functions g which are bounded below by a strictly positive constant, there exists g_0 such that g_0 is a bounded positive function, is bounded below by a strictly positive constant and

$$D(f) < -\sum_{x\in E} \pi(x)\frac{f(x)}{g_0(x)}(Lg_0)(x) + \varepsilon.$$

Since f_n converges weakly to f, there is $n_0 \in \mathbb{N}$ such that

$$-\sum_{x\in E} \pi(x)\frac{f(x)}{g_0(x)}(Lg_0)(x) < -\sum_{x\in E} \pi(x)\frac{f_n(x)}{g_0(x)}(Lg_0)(x) + \varepsilon, \forall n > n_0,$$

which is the same as

$$-\sum_{x\in E} \pi(x)\frac{f(x)}{g_0(x)}(Lg_0)(x) + \varepsilon < -\sum_{x\in E} \pi(x)\frac{f_n(x)}{g_0(x)}(Lg_0)(x) + 2\varepsilon, \forall n > n_0,$$

and leads to

$$D(f) < -\sum_{x\in E} \pi(x)\frac{f_n(x)}{g_0(x)}(Lg_0)(x) + 2\varepsilon, \forall n > n_0.$$

Taking the supremum over all bounded positive functions g which are bounded below by a strictly positive constant, we get

$$D(f) \leq \sup_g \Big[-\sum_{x \in E} \pi(x) \frac{f_n(x)}{g(x)} (Lg)(x) + 2\varepsilon \Big]$$

$$= 2\varepsilon + \sup_g - \sum_{x \in E} \pi(x) \frac{f_n(x)}{g(x)} (Lg)(x) = 2\varepsilon + D(f_n), \forall n > n_0.$$

Taking the $\liminf$ above, we get

$$D(f) \leq 2\varepsilon + \liminf_{n \to \infty} D(f_n) = 2\varepsilon + D(f) - 3\varepsilon = D(f) - \varepsilon < D(f).$$

Therefore, the assumption that $D(f) > \liminf_{n \to \infty} D(f_n)$ is false and we have

$$D(f) \leq \liminf_{n \to \infty} D(f_n).$$

Since $D(f) \leq \liminf_{n \to \infty} D(f_n)$ for every sequence $(f_n)_{n \in \mathbb{N}}$ of densities with respect to π such that f_n converges weakly to f, the functional $D(f)$ is lower semicontinuous.

We conclude the Appendix 1 with a maximal inequality for reversible Markov processes. We assume throughout this section that X_t is a reversible Markov process with respect to some invariant state π.

Theorem

Fix $g : E \to \mathbb{R}$*. For each* $T > 0$ *and* $A > 0$*, we have that*

$$P_\pi\Big[\sup_{0\le t\le T} |g(X_t)| \ge A\Big] \le \frac{e}{A}\sqrt{< g, g >_\pi + T\mathcal{D}(g)}. \qquad (4)$$

Fix $g : E \to \mathbb{R}$, $T > 0$ and $A > 0$. Denote the subset $G := \{x \in E : |g(x)| \geq A\}$ by G and the hitting time of G by τ, i.e., $\tau = \inf\{t \geq 0, X_t \in G\}$. Denote $\mathcal{E}(G_\infty)$ for the set of functions $f : E \to \mathbb{R}$ such that $f \in L^2(\pi)$ and $f(x) = 1, \forall x \in G$.
Let $\lambda > 0$ and define the function $\phi_\lambda : E \to \mathbb{R}$ by

$$\phi_\lambda(x) := \phi(\lambda, x) = E_x[e^{-\lambda\tau}] \leq E_x[e^{-0}] = 1, \forall x \in \mathbb{R}.$$

This leads to

$$\sum_{x\in E}\left(\phi_\lambda(x)\right)^2\pi(x) \leq \sum_{x\in E}\left(1\right)^2\pi(x) = \sum_{x\in E}\pi(x) = 1 < \infty,$$

and $\phi_\lambda \in L^2(\pi)$.
Since $\tau \geq 0$, for $x \in G$, we have

$$P_x(\tau = 0) = \mathbb{P}(\tau = 0|X_0 = x) = \mathbb{P}(\inf\{t \geq 0, X_t \in G\} = 0|X_0 = x \in G) = 1,$$

which leads to $P_x(\tau > 0) = 0$.

Then, we get

$$0 \leq E_x[e^{-\lambda\tau}\mathbb{1}_{\{\tau>0\}}] \leq E_x[e^{-\lambda 0}\mathbb{1}_{\{\tau>0\}}] = E_x[\mathbb{1}_{\{\tau>0\}}] = P_x(\tau > 0) = 0,$$

which leads to $E_x[e^{-\lambda\tau}\mathbb{1}_{\{\tau>0\}}] = 0$ and to

$$\begin{aligned}\phi_\lambda(x) =& E_x[e^{-\lambda\tau}] = E_x[e^{-\lambda\tau}\mathbb{1}_{\{\tau=0\}}] + E_x[e^{-\lambda\tau}\mathbb{1}_{\{\tau>0\}}]\\ =& E_x[e^{-\lambda 0}\mathbb{1}_{\{\tau=0\}}] + E_x[e^{-\lambda\tau}\mathbb{1}_{\{\tau>0\}}]\\ =& E_x[\mathbb{1}_{\{\tau=0\}}] + 0 = P_x(\tau = 0) + 0 = 1 + 0 = 1.\end{aligned}$$

Therefore,

$$\phi_\lambda(x) = 1, \forall x \in G. \tag{5}$$

and $\phi_\lambda \in \mathcal{E}(G_\infty)$.

Now let us consider the case x not in G. Let $t > 0$. We will decompose the chain according to the first site visited. If T_1 is the instant when the chain changes from the initial state to the first state and ξ_1 is the first state, we have

$$
\begin{aligned}
& E_x[\mathbb{1}_{\{T_1 \leq t\}} e^{-\lambda\tau} \mathbb{1}\{\xi_1 = y\}] \\
&= \int_0^t \mathbb{P}(\xi_1 = y, s \leq T_1 \leq s + ds | \xi_0 = x) E_x[e^{-\lambda\tau} | \xi_1 = y] ds \\
&= \int_0^t \mathbb{P}(\xi_1 = y, s \leq T_1 \leq s + ds | \xi_0 = x, T_0 = 0) E_x[e^{-\lambda\tau} | \xi_1 = y] ds \\
&= \int_0^t \mathbb{P}(\xi_1 = y, s \leq T_1 \leq s + ds | \xi_0 = x, T_0 = 0) E_x[e^{-\lambda s} e^{-\lambda(\tau - s)} | \xi_1 = y] ds \\
&= \int_0^t \mathbb{P}(\xi_1 = y, s \leq T_1 \leq s + ds | \xi_0 = x, T_0 = 0) e^{-\lambda s} E_x[e^{-\lambda(\tau - s)} | \xi_1 = y] ds.
\end{aligned}
$$

This leads to

$$\begin{aligned}
&E_x[\mathbb{1}_{\{T_1 \le t\}} e^{-\lambda\tau} \mathbb{1}\{\xi_1 = y\}] \\
&= \int_0^t \mathbb{P}(\xi_1 = y, s \le T_1 \le s + ds | \xi_0 = x, T_0 = 0) e^{-\lambda s} E_x[e^{-\lambda(\tau - s)} | \xi_1 = y] ds \\
&= \int_0^t \mathbb{P}(\xi_1 = y, s \le T_1 \le s + ds | \xi_0 = x, T_0 = 0) e^{-\lambda s} E_y[e^{-\lambda\tau}] ds \\
&= \int_0^t p(x, y)\lambda(x) e^{-\lambda(x)(s-0)} \mathbb{1}_{\{s>0\}} e^{-\lambda s} E_y[e^{-\lambda\tau}] ds \\
&= \int_0^t p(x, y)\lambda(x) e^{-\lambda(x)s} e^{-\lambda s} E_y[e^{-\lambda\tau}] ds
\end{aligned}$$

With Fubini's Theorem, we get

$$
\begin{aligned}
& E_x[\mathbb{1}_{\{T_1 \leq t\}} e^{-\lambda\tau}] \\
= & \sum_{y \in E} E_x[\mathbb{1}_{\{T_1 \leq t\}} e^{-\lambda\tau} \mathbb{1}\{\xi_1 = y\}] \\
& \sum_{y \in E} \int_0^t p(x,y)\lambda(x) e^{-\lambda(x)s} e^{-\lambda s} E_y[e^{-\lambda\tau}] ds \\
= & \int_0^t \Big[\sum_{y \in E} p(x,y)\lambda(x) e^{\{-\lambda(x)-\lambda\}s} E_y[e^{-\lambda\tau}] \Big] ds.
\end{aligned}
$$

Then, we have

$$E_x[\mathbb{1}_{\{T_1 \le t\}} e^{-\lambda\tau}] = \int_0^t \Big[\sum_{y \in E} p(x,y)\lambda(x) e^{\{-\lambda(x)-\lambda\}s} \phi_\lambda(y)\Big] ds. \qquad (6)$$

Since $0 \le e^{-\lambda\tau} \le 1$ and $\{T_0 \le t\} = \{0 \le t\}$ has probability one, the Markov property gives that

$$\begin{aligned} E_x[\mathbb{1}_{\{T_1 > t\}} e^{-\lambda(\tau - t)}] &= E_x[\mathbb{1}_{\{T_1 > t\}} e^{-\lambda(\tau - t)} | (X_0 = x, T_0 = 0)] \\ &= e^{-\lambda(x)}(t-0) E_x[e^{-\lambda\tau}] = e^{-\lambda(x)t} \phi_\lambda(x), \end{aligned}$$

which leads to

$$\begin{aligned} E_x[\mathbb{1}_{\{T_1 > t\}} e^{-\lambda\tau}] &= E_x[\mathbb{1}_{\{T_1 > t\}} e^{-\lambda(\tau - t)} e^{-\lambda t}] \\ &= e^{-\lambda t} E_x[\mathbb{1}_{\{T_1 > t\}} e^{-\lambda(\tau - t)}] \\ &= e^{-\lambda t} e^{-\lambda(x)t} \phi_\lambda(x), \end{aligned}$$

and we have

$$E_x[\mathbb{1}_{\{T_1 > t\}} e^{-\lambda\tau}] = e^{\{-\lambda(x)-\lambda\}t} \phi_\lambda(x). \qquad (7)$$

Equations (6) and (7) give

$$\begin{aligned}\phi_\lambda(x) =& E_x[e^{-\lambda\tau}]\\ =& E_x[\mathbb{1}_{\{T_1\le t\}}e^{-\lambda\tau}] + E_x[\mathbb{1}_{\{T_1>t\}}e^{-\lambda\tau}]\\ =& \int_0^t \Big[\sum_{y\in E} p(x,y)\lambda(x)e^{\{-\lambda(x)-\lambda\}s}\phi_\lambda(y)\Big]ds + e^{\{-\lambda(x)-\lambda\}t}\phi_\lambda(x),\end{aligned}$$

which is the same as

$$\int_0^t \Big[\sum_{y\in E} p(x,y)\lambda(x)e^{\{-\lambda(x)-\lambda\}s}\phi_\lambda(y)\Big]ds$$
$$+(e^{\{-\lambda(x)-\lambda\}t}-1)\phi_\lambda(x) = 0, \forall t>0.$$

Differentiating with respect to t, we get

$$\sum_{y \in E} p(x,y)\lambda(x)e^{\{-\lambda(x)-\lambda\}t}\phi_\lambda(y) - (\lambda(x)+\lambda)e^{\{-\lambda(x)-\lambda\}t}\phi_\lambda(x) = 0, \forall t > 0.$$

Making $t \to 0^+$, we get

$$\sum_{y \in E} p(x,y)\lambda(x)\phi_\lambda(y) - (\lambda(x)+\lambda)\phi_\lambda(x) = 0.$$

This is the same as

$$\begin{aligned}
\lambda\phi_\lambda(x) &= -\lambda(x)\phi_\lambda(x)\cdot 1 + \sum_{y\in E} p(x,y)\lambda(x)\phi_\lambda(y) \\
&= -\lambda(x)\phi_\lambda(x)\sum_{y\in E} p(x,y) + \sum_{y\in E} p(x,y)\lambda(x)\phi_\lambda(y) \\
&= \sum_{y\in E} p(x,y)\lambda(x)[\phi_\lambda(y)-\phi_\lambda(x)] \\
&= \sum_{\substack{y\in E\\ y\neq x}} p(x,y)\lambda(x)[\phi_\lambda(y)-\phi_\lambda(x)] \\
&= \sum_{\substack{y\in E\\ y\neq x}} L(x,y)[\phi_\lambda(y)-\phi_\lambda(x)] \\
&= \sum_{y\in E} L(x,y)[\phi_\lambda(y)-\phi_\lambda(x)].
\end{aligned}$$

Therefore, we have

$$(L\phi_\lambda)(x) = \lambda\phi_\lambda(x), \forall x \in G^C. \tag{8}$$

We are interested in finding out how many functions $h \in \mathcal{E}(G_\infty)$ satisfy

$$(Lh)(x) = \lambda h(x), \forall x \notin G. \tag{9}$$

Assume that $h_1, h_2 \in \mathcal{E}(G_\infty)$ satisfy (9). Let $h_3 : E \to \mathbb{R}$ be such that $h_3(x) = h_1(x) - h_2(x)$, $\forall x \in E$. Then

$$h_3(x) = h_1(x) - h_2(x) = 1 - 1 = 0, \forall x \in G.$$

In particular, we have

$$(Lh_3)(x)h_3(x)\pi(x) = \lambda h_3(x)h_3(x)\pi(x),\ \forall x \in G.$$

We also have

$$\begin{aligned}(Lh_3)(x) =& \sum_{y\in E} L(x,y)[h_3(y) - h_3(x)] \\ =& \sum_{y\in E} L(x,y)\big[\big(h_1(y) - h_2(y)\big) - \big(h_1(x) - h_2(x)\big)\big] \\ =& \sum_{y\in E} L(x,y)[h_1(y) - h_1(x)] - \sum_{y\in E} L(x,y)[h_2(y) - h_2(x)] \\ =& (Lh_1)(x) - (Lh_2)(x) = \lambda h_1(x) - \lambda h_2(x) \\ =& \lambda\big(h_1(x) - h_2(x)\big) = \lambda h_3(x), \forall x \notin G.\end{aligned}$$

This leads to

$$(Lh_3)(x)h_3(x)\pi(x) = \lambda h_3(x)h_3(x)\pi(x), \ \forall x \notin G$$

and to

$$(Lh_3)(x)h_3(x)\pi(x) = \lambda h_3(x)h_3(x)\pi(x), \ \forall x \in E.$$

Then we have

$$-\mathcal{D}(h_3) = < Lh_3, h_3 >_\pi = \sum_{x\in E}(Lh_3)(x)h_3(x)\pi(x)$$
$$= \sum_{x\in E}\lambda h_3(x)h_3(x)\pi(x) = \lambda\sum_{x\in E}h_3(x)h_3(x)\pi(x) = \lambda < h_3, h_3 >_\pi \geq 0.$$

Since $-\mathcal{D}(h_3) \leq 0$, we have that $\lambda < h_3, h_3 >_\pi = 0$, which means that $h_3(x) = 0, \forall x \in E$, which is the same as $h_1(x) = h_2(x), \forall x \in E$. Then, if h_1, h_2 satisfy (9), $h_1 = h_2$. This means that there is at most one function $h \in \mathcal{E}(G_\infty)$ that satisfies (9). By (5) and (8), we have that ϕ_λ satisfies (9). Therefore, ϕ_λ is the unique function on $\mathcal{E}(G_\infty)$ which satisfies (9).

By definition of the stopping time τ, the events $\{\sup_{0\leq t\leq T}|g(X_t)|\geq A\}$ and $\{\tau\leq T\}$ are the same, which leads to

$$P_\pi(\sup_{0\leq t\leq T}|g(X_t)|\geq A)=P_\pi(\tau\leq T).$$

We have

$$\begin{aligned}P_x(\tau\leq T)=&E_x[\mathbb{1}_{\{\tau\leq T\}}]\leq E_x[e^{\lambda(T-\tau)}\mathbb{1}_{\{\tau\leq T\}}]\\ \leq&E_x[e^{\lambda(T-\tau)}]=e^{\lambda T}E_x[e^{-\lambda\tau}]=e^{\lambda T}\phi_\lambda(x).\end{aligned}$$

Schwartz's inequality leads to

$$\begin{aligned}
P_\pi(\sup_{0\le t\le T}|g(X_t)| \ge A) =& P_\pi(\tau \le T) = \sum_{x\in E}\pi(x)P_x(\tau \le T)\\
\le& \sum_{x\in E}\pi(x)e^{\lambda T}\phi_\lambda(x) = e^{\lambda T}\sum_{x\in E}\pi(x)\phi_\lambda(x)\\
=& e^{\lambda T}E_\pi[\phi_\lambda] \le e^{\lambda T}\sqrt{E_\pi[\phi_\lambda^2]}\\
=& e^{\lambda T}\sqrt{\sum_{x\in E}\pi(x)\phi_\lambda^2(x)}.
\end{aligned}$$

Since the Dirichlet form is non-negative, it holds

$$\sum_{x\in E}\pi(x)\phi_\lambda^2(x) \le \sum_{x\in E}\pi(x)\phi_\lambda^2(x) + \frac{1}{\lambda}\mathcal{D}(\phi_\lambda). \tag{10}$$

Define the functional $J_\lambda(f)$ by

$$J_\lambda(f) = \sum_{x\in E}\pi(x)f^2(x) + \frac{1}{\lambda}\mathcal{D}(f), \tag{11}$$

among all functions $f \in \mathcal{E}(G_\infty)$.

Then, we have

$$\begin{aligned} P_\pi(\sup_{0\le t\le T} |g(X_t)| \geq A) \le& e^{\lambda T}\sqrt{\sum_{x\in E}\pi(x)\phi_\lambda^2(x)} \\ \le& e^{\lambda T}\sqrt{\sum_{x\in E}\pi(x)\phi_\lambda^2(x) + \frac{1}{\lambda}\mathcal{D}(\phi_\lambda)} \\ =& e^{\lambda T}\sqrt{J_\lambda(\phi_\lambda)}. \end{aligned}$$

Let $h : E \to \mathbb{R}$ be such that

$$h(x) = A^{-1}\min\{|g(x)|, A\} \le A^{-1}\cdot A = 1, \forall x \in E.$$

Since $|g(x)| \geq A, \forall x \in G$, we get

$$h(x) = A^{-1}\min\{|g(x)|, A\} = A^{-1}A = 1, \forall x \in G.$$

Then $h \in \mathcal{E}(G_\infty)$ and we have

$$\begin{aligned} J_\lambda(h) &= \sum_{x\in E} \pi(x)h^2(x) + \frac{1}{\lambda}\mathcal{D}(h) \\ &= \sum_{x\in E} \pi(x)\big(A^{-1}\min\{|g(x)|, A\}\big)^2 + \frac{1}{\lambda}\mathcal{D}(A^{-1}\min\{|g|, A\}) \\ &= \sum_{x\in E} A^{-2}\pi(x)\big(\min\{|g(x)|, A\}\big)^2 + A^{-2}\frac{1}{\lambda}\mathcal{D}(\min\{|g|, A\}) \\ &\leq \sum_{x\in E} A^{-2}\pi(x)g^2(x) + A^{-2}\frac{1}{\lambda}\mathcal{D}(\min\{|g|, A\}) \\ &= A^{-2}\Big[\sum_{x\in E} \pi(x)g^2(x) + \frac{1}{\lambda}\mathcal{D}(\min\{|g|, A\})\Big]. \end{aligned}$$

By Proposition 10 and Propositon 11, we get

$$\begin{aligned} J_\lambda(h) \le & A^{-2}\Big[\sum_{x\in E}\pi(x)g^2(x) + \frac{1}{\lambda}\mathcal{D}(\min\{|g|, A\})\Big] \\ \le & A^{-2}\Big[\sum_{x\in E}\pi(x)g^2(x) + \frac{1}{\lambda}\mathcal{D}(|g|)\Big] \\ \le & A^{-2}\Big[\sum_{x\in E}\pi(x)g^2(x) + \frac{1}{\lambda}\mathcal{D}(g)\Big]. \end{aligned}$$

We claim that

Claim

A function which minimizes the functional J_λ must satisfy (9).

Assume that the claim holds. Since ϕ_λ is the unique function on $\mathcal{E}(G_\infty)$ which satisfies the (9), ϕ_λ is the minimizer of J_λ. In particular,

$$J_\lambda(\phi_\lambda) \le J_\lambda(h) \le A^{-2}\Big[\sum_{x\in E}\pi(x)g^2(x) + \frac{1}{\lambda}\mathcal{D}(g)\Big],$$

which leads to

$$\begin{aligned} P_\pi(\sup_{0\le t\le T}|g(X_t)| \ge A) \le & e^{\lambda T}\sqrt{J_\lambda(\phi_\lambda)} \\ \le & e^{\lambda T}\sqrt{A^{-2}\Big[\sum_{x\in E}\pi(x)g^2(x) + \frac{1}{\lambda}\mathcal{D}(g)\Big]} \\ = & \frac{e^{\lambda T}}{A}\sqrt{< g, g >_\pi + \frac{1}{\lambda}\mathcal{D}(g)}. \end{aligned}$$

Since $\lambda > 0$ is arbitrary, we have

$$P_\pi(\sup_{0\leq t\leq T}|g(X_t)| \geq A) \leq \frac{e^{\lambda T}}{A}\sqrt{< g, g >_\pi + \frac{1}{\lambda}\mathcal{D}(g)}, \forall \lambda > 0.$$

In particular, choosing $\lambda = \frac{1}{T}$, we have

$$\begin{aligned} P_\pi(\sup_{0\leq t\leq T}|g(X_t)| \geq A) \leq & \frac{e^{\lambda T}}{A}\sqrt{< g, g >_\pi + \frac{1}{\lambda}\mathcal{D}(g)} \\ = & \frac{e^{\frac{1}{T} T}}{A}\sqrt{< g, g >_\pi + \frac{1}{\frac{1}{T}}\mathcal{D}(g)} \\ = & \frac{e}{A}\sqrt{< g, g >_\pi + T\mathcal{D}(g)}, \end{aligned}$$

and the theorem is proved.

We only need to prove the claim. We have

$$\begin{aligned} J_\lambda(f) &= < f, f >_\pi + \frac{1}{\lambda}\mathcal{D}(f) \\ &= \sum_{x\in E} \pi(x)f(x)f(x) + \frac{1}{\lambda}\frac{1}{2}\sum_{x,y}\pi(x)L(x,y)[f(y)-f(x)]^2 \\ &= \sum_{x\in E}\pi(x)f^2(x) + \frac{1}{2\lambda}\sum_{x,y\in E}\pi(x)L(x,y)[f(y)-f(x)]^2. \end{aligned}$$

Since J_λ is defined over the functions $f : E \to \mathbb{R}$ such that $f(x) = 1 \ \forall x \in G$, we get

$$\begin{aligned} \sum_{x\in E}\pi(x)f^2(x) &= \sum_{x\in G}\pi(x)f^2(x) + \sum_{x\in G^C}\pi(x)f^2(x) \\ &= \sum_{x\in G}\pi(x)1^2 + \sum_{x\in G^C}\pi(x)f^2(x) = \pi(G) + \sum_{x\in G^C}\pi(x)f^2(x). \end{aligned}$$

Since X_t is a Markov process reversible with respect to the probability measure π, $\pi(x)L(x,y) = \pi(y)L(y,x)$, $\forall x, y \in E$ and we have

$$\sum_{x,y\in E} \pi(x)L(x,y)[f(y)-f(x)]^2$$
$$= \sum_{x,y\in G} \pi(x)L(x,y)[f(y)-f(x)]^2 + \sum_{x\in G}\sum_{y\in G^C} \pi(x)L(x,y)[f(y)-f(x)]^2$$
$$+ \sum_{x\in G^C}\sum_{y\in G} \pi(x)L(x,y)[f(y)-f(x)]^2 + \sum_{x,y\in G^C} \pi(x)L(x,y)[f(y)-f(x)]^2$$
$$= \sum_{x,y\in G} \pi(x)L(x,y)[1-1]^2 + \sum_{x\in G}\sum_{y\in G^C} \pi(y)L(y,x)[f(y)-f(x)]^2$$
$$+ \sum_{x\in G^C}\sum_{y\in G} \pi(x)L(x,y)[f(y)-f(x)]^2 + \sum_{x,y\in G^C} \pi(x)L(x,y)[f(y)-f(x)]^2.$$

Interchanging x and y in the second summation,

$$
\begin{aligned}
&\sum_{x,y\in E} \pi(x)L(x,y)[f(y)-f(x)]^2 \\
=&\sum_{x,y\in G} \pi(x)L(x,y)0^2 + \sum_{y\in G}\sum_{x\in G^C} \pi(x)L(x,y)[f(x)-f(y)]^2 \\
&+\sum_{x\in G^C}\sum_{y\in G} \pi(x)L(x,y)[f(y)-f(x)]^2 + \sum_{x,y\in G^C} \pi(x)L(x,y)[f(y)-f(x)]^2 \\
=&0 + 2\sum_{x\in G^C}\sum_{y\in G} \pi(x)L(x,y)[f(x)-f(y)]^2 + \sum_{x,y\in G^C} \pi(x)L(x,y)[f(y)-f(x)]^2 \\
=&2\sum_{x\in G^C}\sum_{y\in G} \pi(x)L(x,y)[f(x)-1]^2 + \sum_{x,y\in G^C} \pi(x)L(x,y)[f(y)-f(x)]^2 \\
=&2\sum_{x\in G^C} \pi(x)[f(x)-1]^2 \sum_{y\in G} L(x,y) + \sum_{\substack{x,y\in G^C \\ y\neq x}} \pi(x)L(x,y)[f(y)-f(x)]^2.
\end{aligned}
$$

Define $q : G^C \to \mathbb{R}$ by

$$q(x) = \sum_{y \in G} L(x, y) \geq 0, \forall x \in G^C.$$

Then

$$\begin{aligned}
&\sum_{x,y \in E} \pi(x) L(x, y)[f(y) - f(x)]^2 \\
=&2 \sum_{x \in G^C} \pi(x)[f(x) - 1]^2 \sum_{y \in G} L(x, y) + \sum_{\substack{x,y \in G^C \\ y \neq x}} \pi(x) L(x, y)[f(y) - f(x)]^2 \\
=&2 \sum_{x \in G^C} \pi(x)[f(x) - 1]^2 q(x) + \sum_{\substack{x,y \in G^C \\ y \neq x}} \pi(x) L(x, y)[f(y) - f(x)]^2.
\end{aligned}$$

This leads to

$$
\begin{aligned}
J_\lambda(f) =& \sum_{x\in E}\pi(x)f^2(x) + \frac{1}{2\lambda}\sum_{x,y\in E}\pi(x)L(x,y)[f(y)-f(x)]^2 \\
=& \pi(G) + \sum_{x\in G^C}\pi(x)f^2(x) + \frac{1}{\lambda}\sum_{x\in G^C}\pi(x)[f(x)-1]^2 q(x) \\
&+ \frac{1}{2\lambda}\sum_{\substack{x,y\in G^C\\ y\neq x}}\pi(x)L(x,y)[f(y)-f(x)]^2.
\end{aligned}
$$

There are two possibilities: G^C is finite (case 1) or G^C is not finite (case 2).

Case 1: G^C is finite. In this case, we can write $G^C = \{x_1, \ldots, x_N\}$, with $N = |G^C|$. For every $f : E \to \mathbb{R}$ such that $f(x) = 1$, $\forall x \in G$, denote $y_i = f(x_i)$, $\forall i = 1, \ldots, N$. We also denote $\pi_i := \pi(x_i) > 0$, $\forall i = 1, \ldots, N$, $q_i := q(x_i)$, $\forall i = 1, \ldots, N$ and $L(i,j) := L(x_i, x_j)$, $\forall i, j = 1, \ldots, N$. Then, we have

$$\begin{aligned}
J_\lambda(f) =& \pi(G) + \sum_{x \in G^C} \pi(x) f^2(x) + \frac{1}{\lambda} \sum_{x \in G^C} \pi(x)[f(x) - 1]^2 q(x) \\
&+ \frac{1}{2\lambda} \sum_{\substack{x, y \in G^C \\ y \neq x}} \pi(x) L(x, y)[f(y) - f(x)]^2 \\
=& \pi(G) + \sum_{i=1}^{N} \pi(x_i) f^2(x_i) + \frac{1}{\lambda} \sum_{i=1}^{N} \pi(x_i)[f(x_i) - 1]^2 q(x_i) \\
&+ \frac{1}{2\lambda} \sum_{\substack{i,j=1 \\ j \neq i}}^{N} \pi(x_i) L(x_i, x_j)[f(x_j) - f(x_i)]^2.
\end{aligned}$$

Since $\pi(x_j)L(x_j,x_i) = \pi(x_i)L(x_i,x_j), \forall i,j \in \{1,\dots,N\}$,

$$\begin{aligned}
J_\lambda(f) =& \pi(G) + \sum_{i=1}^{N} \pi(x_i) f^2(x_i) + \frac{1}{\lambda}\sum_{i=1}^{N} \pi(x_i)[f(x_i)-1]^2 q(x_i) \\
&+ \frac{1}{2\lambda}\sum_{\substack{i,j=1 \\ j<i}}^{N} \left[\pi(x_i)L(x_i,x_j)[f(x_j)-f(x_i)]^2 + \pi(x_j)L(x_j,x_i)[f(x_i)-f(x_j)]^2\right] \\
=& \pi(G) + \sum_{i=1}^{N} \pi(x_i) f^2(x_i) + \frac{1}{\lambda}\sum_{i=1}^{N} \pi(x_i)[f(x_i)-1]^2 q(x_i) \\
&+ \frac{1}{2\lambda}\sum_{\substack{i,j=1 \\ j<i}}^{N} \left[\pi(x_i)L(x_i,x_j)[f(x_j)-f(x_i)]^2 + \pi(x_i)L(x_i,x_j)[f(x_i)-f(x_j)]^2\right] \\
=& \pi(G) + \sum_{i=1}^{N} \pi_i y_i^2 + \frac{1}{\lambda}\sum_{i=1}^{N} \pi_i [y_i-1]^2 q_i + \frac{1}{\lambda}\sum_{\substack{i,j=1 \\ j<i}}^{N} \pi_i L(i,j)[y_i-y_j]^2.
\end{aligned}$$

Then, we have

$$J_\lambda(f) = \Phi(y_1, \ldots, y_N) = \Phi(f(x_1), \ldots, f(x_N)),$$

where we define $\Phi : \mathbb{R}^N \to \mathbb{R}$ by

$$\begin{aligned}&\Phi(y_1, \ldots, y_N)\\ =&\pi(G) + \sum_{i=1}^{N} \pi_i y_i^2 + \frac{1}{\lambda}\sum_{i=1}^{N} \pi_i [y_i - 1]^2 q_i + \frac{1}{\lambda}\sum_{\substack{i,j=1\\ j<i}}^{N} \pi_i L(i,j)[y_i - y_j]^2.\end{aligned}$$

We observe that $x \mapsto x^2$, $x \mapsto (x-1)^2$ and $(x,y) \mapsto (x-y)^2$ are convex functions in $\mathbb{R}$, $\mathbb{R}$ and $\mathbb{R}^2$, respectively.
Since $\pi_i \geq 0, \forall i = 1, \ldots, N$, $q_i \geq 0, \forall i = 1, \ldots, N$, $L(i,j) \geq 0, \forall i \neq j \in 1, \ldots, N$,$\lambda > 0$ and a finite linear combination (with non-negative coefficients) convex functions is convex, we have that Φ is convex in $\mathbb{R}^N$. Then Φ assumes its minimum where its gradient vanishes.

For every $i = 1, \ldots, N$ we have

$$\begin{aligned}
\frac{\partial \Phi}{\partial y_i}(y_1, \ldots, y_N) =& 2\pi_i y_i + \frac{2\pi_i [y_i - 1] q_i}{\lambda} + \frac{1}{\lambda} \sum_{\substack{j=1 \\ j \neq i}}^{N} \pi_i L(i,j) 2 [y_i - y_j] \\
=& 2\pi_i \Big(y_i + \frac{1}{\lambda}[y_1 - 1] q_i + \frac{1}{\lambda} \sum_{\substack{j=1 \\ j \neq i}}^{N} L(i,j)[y_i - y_j] \Big).
\end{aligned}$$

This leads to

$$\frac{\partial \Phi}{\partial y_i}(f(x_1),\ldots,f(x_N))$$

$$=2\pi_i\Big(f(x_i)+\frac{1}{\lambda}[f(x_i)-1]q(x_i)+\frac{1}{\lambda}\sum_{\substack{j=1\\ j\neq i}}^{N}L(x_i,x_j)[f(x_i)-f(x_j)]\Big)$$

$$=2\pi_i\Big(f(x_i)+\frac{1}{\lambda}[f(x_i)-1]\sum_{y\in G}L(x_i,y)+\frac{1}{\lambda}\sum_{\substack{y\in G^C\\ y\neq x_i}}L(x_i,y)[f(x_i)-f(y)]\Big)$$

$$=2\pi_i\Big(f(x_i)+\frac{1}{\lambda}\sum_{y\in G}L(x_i,y)[f(x_i)-1]+\frac{1}{\lambda}\sum_{\substack{y\in G^C\\ y\neq x_i}}L(x_i,y)[f(x_i)-f(y)]\Big)$$

$$=2\pi_i\Big(f(x_i)+\frac{1}{\lambda}\sum_{y\in G}L(x_i,y)[f(x_i)-f(y)]+\frac{1}{\lambda}\sum_{\substack{y\in G^C\\ y\neq x_i}}L(x_i,y)[f(x_i)-f(y)]\Big).$$

Then, we have

$$\frac{\partial\Phi}{\partial y_i}(f(x_1),\ldots,f(x_N))$$
$$=2\pi_i\Big(f(x_i)+\frac{1}{\lambda}\sum_{y\in G}L(x_i,y)[f(x_i)-f(y)]+\frac{1}{\lambda}\sum_{\substack{y\in G^C\\ y\neq x_i}}L(x_i,y)[f(x_i)-f(y)]\Big)$$
$$=2\pi_i\Big(f(x_i)+\frac{1}{\lambda}\sum_{\substack{y\in E\\ y\neq x_i}}L(x_i,y)[f(x_i)-f(y)]\Big)$$
$$=2\pi_i\Big(f(x_i)+\frac{1}{\lambda}\sum_{y\in E}L(x_i,y)[f(x_i)-f(y)]\Big)$$
$$=2\pi_i\Big(f(x_i)-\frac{1}{\lambda}\sum_{y\in E}L(x_i,y)[f(y)-f(x_i)]\Big)$$
$$=2\pi_i\big(f(x_i)-\frac{1}{\lambda}(Lf)(x_i)\big),\forall i=1,\ldots,N.$$

Since $\pi_i > 0$, $\forall i = 1, \ldots, N$, we have that $\frac{\partial \Phi}{\partial y_i}(f(x_1), \ldots, f(x_N)) = 0$ if and only if $(Lf)(x_i) = \lambda f(x_i)$. Since Φ attains its mininum where its gradient vanish and $J_\lambda(f) = \Phi(f(x_1), \ldots, f(x_N))$, we have that a function f which minimizes the functional J_λ must be such that $(Lf)(x) = \lambda f(x)$, $\forall x \in G^C$ and therefore satisfies (9).

Case 2: G^C is not finite.
Since G^C is countable, we can write $G^C = \{x_1, x_2, \ldots\}$. For every $k \in \mathbb{N}$, denote $G_k := \{x_1, \ldots, x_k\}$ and $\mathcal{E}(G_k)$ for the set of functions $f \in \mathcal{E}(G_\infty)$ that are constant on $G^C - G_k$. Then $(G_k)_{k\geq 1}$ is an increasing sequence of finite subsets of G^C whose union is equal to G^C and $\mathcal{E}(G_k) \subset \mathcal{E}(G_{k+1}), \forall k \in \mathbb{N}$. Observe that

$$0 \leq \sum_{y \in G} L(x, y) = q(x) \leq \sum_{\substack{y \in E \\ y \neq x}} L(x, y) = \lambda(x) \leq \bar{\lambda}, \forall x \in G^C.$$

Since $\mathcal{E}(G_k) \subset \mathcal{E}(G_{k+1}) \subset \mathcal{E}(G_\infty), \forall k \in \mathbb{N}$, we have

$$\inf_{f\in\mathcal{E}(G_k)} J_\lambda(f) \geq \inf_{f\in\mathcal{E}(G_\infty)} J_\lambda(f), \forall k \in \mathbb{N},$$

which leads to

$$\lim_{k\to\infty} \inf_{f\in\mathcal{E}(G_k)} J_\lambda(f) \geq \inf_{f\in\mathcal{E}(G_\infty)} J_\lambda(f).$$

Suppose that $\inf_{f\in\mathcal{E}(G_\infty)} J_\lambda(f) < \lim_{k\to\infty} \inf_{f\in\mathcal{E}(G_k)} J_\lambda(f)$. Then there exists $f_\infty \in \mathcal{E}(G_\infty)$ such that

$$\inf_{f\in\mathcal{E}(G_\infty)} J_\lambda(f) < J_\lambda(f_\infty) < \lim_{k\to\infty} \inf_{f\in\mathcal{E}(G_k)} J_\lambda(f).$$

From the definition of q, we have that

$$0 \leq q(x) = \sum_{y \in G} L(x,y) \leq \sum_{\substack{y \in E \\ y \neq x}} L(x,y) = \lambda(x) \leq \bar{\lambda}, \forall x \in G^C.$$

For every $k \in \mathbb{N}$, define $f_k : E \to \mathbb{R}$ by

$$f_k(x) = \begin{cases} f_\infty(x) = 1, & \text{if } x \in G; \\ f_\infty(x), & \text{if } x \in G_k; \\ 0, & \text{if } x \in G^C - G_k. \end{cases}$$

Since $f_k(x) = 1, \forall x \in G, \forall k \in \mathbb{N}$ and $|f_k(x)| \leq |f_\infty(x)|, \forall x \in E, \forall k \in \mathbb{N}$, $f_k(x) \in \mathcal{E}(G_\infty)$, $\forall k \in \mathbb{N}$. Moreover, since $f_k(x) = 0, \forall x \in G^C - G_k$, we have $f_k \in \mathcal{E}(G_k), \forall k \in \mathbb{N}$. Also, for every $k \in \mathbb{N}$, we have

$$J_\lambda(f_\infty) - J_\lambda(f_k) = \pi(G) + \sum_{x \in G^C} \pi(x) f_\infty^2(x) + \frac{1}{\lambda} \sum_{x \in G^C} \pi(x)[f_\infty(x) - 1]^2 q(x)$$

$$+\frac{1}{2\lambda} \sum_{\substack{x,y \in G^C \\ y \neq x}} \pi(x) L(x,y)[f_\infty(y) - f_\infty(x)]^2$$

$$-\pi(G) - \sum_{x \in G^C} \pi(x) f_k^2(x) - \frac{1}{\lambda} \sum_{x \in G^C} \pi(x)[f_k(x) - 1]^2 q(x)$$

$$-\frac{1}{2\lambda} \sum_{\substack{x,y \in G^C \\ y \neq x}} \pi(x) L(x,y)[f_k(y) - f_k(x)]^2$$

Then, $J_\lambda(f_\infty) - J_\lambda(f_k)$ is equal to

$$\sum_{x \in G^C - G_k} \pi(x) f_\infty^2(x) + \frac{1}{\lambda} \sum_{x \in G^C - G_k} \pi(x) \big[[f_\infty(x) - 1]^2 - 1 \big] q(x)$$

$$+ \frac{1}{2\lambda} \sum_{x \in G_k} \sum_{y \in G^C - G_k} \pi(x) L(x,y) \big[[f_\infty(y) - f_\infty(x)]^2 - f_\infty^2(x) \big]$$

$$+ \frac{1}{2\lambda} \sum_{x \in G^C - G_k} \sum_{y \in G_k} \pi(x) L(x,y) \big[[f_\infty(y) - f_\infty(x)]^2 - f_\infty^2(y) \big]$$

$$+ \frac{1}{2\lambda} \sum_{\substack{x,y \in G^C - G_k \\ y \neq x}} \pi(x) L(x,y) [f_\infty(y) - f_\infty(x)]^2.$$

This leads to

$$
\begin{aligned}
&|J_\lambda(f_\infty) - J_\lambda(f_k)| \\
&\leq \sum_{x \in G^C - G_k} \pi(x) f_\infty^2(x) + \frac{1}{\lambda} \sum_{x \in G^C - G_k} \pi(x) \big[[f_\infty(x) - 1]^2 + 1 \big] q(x) \\
&+ \frac{1}{2\lambda} \sum_{x \in G_k} \sum_{y \in G^C - G_k} \pi(x) L(x,y) \big[[f_\infty(y) - f_\infty(x)]^2 + f_\infty^2(x) \big] \\
&+ \frac{1}{2\lambda} \sum_{x \in G^C - G_k} \sum_{y \in G_k} \pi(x) L(x,y) \big[[f_\infty(y) - f_\infty(x)]^2 + f_\infty^2(y) \big] \\
&+ \frac{1}{2\lambda} \sum_{\substack{x,y \in G^C - G_k \\ y \neq x}} \pi(x) L(x,y) [f_\infty(y) - f_\infty(x)]^2.
\end{aligned}
$$

Since $q(x) \leq \bar{\lambda} \; \forall x \in G^C$, we get

$$
\begin{aligned}
&|J_\lambda(f_\infty) - J_\lambda(f_k)| \\
&\leq \sum_{x \in G^C - G_k} \pi(x) f_\infty^2(x) + \frac{1}{\lambda} \sum_{x \in G^C - G_k} \pi(x) \big[[f_\infty(x) - 1]^2 + 1 \big] \bar{\lambda} \\
&+ \frac{1}{2\lambda} \sum_{x \in G_k} \sum_{y \in G^C - G_k} \pi(x) L(x,y) \big[[f_\infty(y) - f_\infty(x)]^2 + f_\infty^2(x) \big] \\
&+ \frac{1}{2\lambda} \sum_{y \in G^C - G_k} \sum_{x \in G_k} \pi(y) L(y,x) \big[[f_\infty(x) - f_\infty(y)]^2 + f_\infty^2(x) \big] \\
&+ \frac{1}{\lambda} \sum_{\substack{x,y \in G^C - G_k \\ y \neq x}} \pi(x) L(x,y) [f_\infty(y) - f_\infty(x)]^2.
\end{aligned}
$$

Since $\pi(x)L(x,y) = \pi(y)L(y,x), \forall x, y \in E$, we get

$$\begin{aligned}
&|J_\lambda(f_\infty) - J_\lambda(f_k)| \\
&\le \sum_{x \in G^C - G_k} \pi(x) f_\infty^2(x) + \frac{\bar{\lambda}}{\lambda} \sum_{x \in G^C - G_k} \pi(x) \left[[f_\infty(x) - 1]^2 + 1 \right] \\
&+ \frac{1}{\lambda} \sum_{x \in G_k} \sum_{y \in G^C - G_k} \pi(x) L(x,y) \left[[f_\infty(y) - f_\infty(x)]^2 + f_\infty^2(x) \right] \\
&+ \frac{1}{\lambda} \sum_{\substack{x,y \in G^C - G_k \\ y \neq x}} \pi(x) L(x,y) [f_\infty(y) - f_\infty(x)]^2.
\end{aligned}$$

Then, $|J_\lambda(f_\infty) - J_\lambda(f_k)|$ is smaller or equal to

$$\sum_{x\in G^C - G_k} \pi(x) f_\infty^2(x) + \frac{\bar{\lambda}}{\lambda} \sum_{x\in G^C - G_k} \pi(x)\big[[f_\infty(x) - 1]^2 + 1\big]$$
$$+\frac{1}{\lambda}\Big\{ \sum_{x\in G_k} \sum_{y\in G^C - G_k} \pi(x)L(x,y)[f_\infty(y) - f_\infty(x)]^2$$
$$+ \sum_{x,y\in G^C - G_k} \pi(x)L(x,y)[f_\infty(y) - f_\infty(x)]^2 \Big\}$$
$$+\frac{1}{\lambda} \sum_{x\in G_k} \sum_{y\in G^C - G_k} \pi(x)L(x,y) f_\infty^2(x).$$

Since $f_\infty \in L^2(\pi)$, we have

$$\lim_{k\to\infty} \sum_{x\in G_k} \pi(x) f_\infty^2(x) = \sum_{x\in G^C} \pi(x) f_\infty^2(x) \leq \sum_{x\in E} \pi(x) f_\infty^2(x) = E_\pi[f_\infty^2] < \infty,$$

which leads to

$$\lim_{k\to\infty} \sum_{x\in G^C - G_k} \pi(x) f_\infty^2(x) = \sum_{x\in G^C} \pi(x) f_\infty^2(x) - \lim_{k\to\infty} \sum_{x\in G_k} \pi(x) f_\infty^2(x) = 0.$$

We also have that

$$
\begin{aligned}
&\lim_{k\to\infty} \sum_{x\in G_k} \pi(x)\big[f_\infty(x) - 1]^2 + 1\big] \\
&= \sum_{x\in G^C} \pi(x)\big[f_\infty(x) - 1]^2 + 1\big] \le \sum_{x\in E} \pi(x)\big[f_\infty(x) - 1]^2 + 1\big] \\
&\le \sum_{x\in E} \pi(x)\big[[2f_\infty^2(x) + 2] + 1\big] \\
&= 3\sum_{x\in E} \pi(x) + 2\sum_{x\in E} \pi(x) f_\infty^2(x) = 3 + 2\sum_{x\in E} \pi(x) f_\infty^2(x) < \infty,
\end{aligned}
$$

which leads to

$$
\lim_{k\to\infty} \sum_{x\in G^C - G_k} \pi(x)\big[f_\infty(x) - 1]^2 + 1\big] = 0.
$$

We know that

$$\begin{aligned}
&\lim_{k\to\infty} \sum_{x,y\in G_k} \pi(x)L(x,y)[f_\infty(y) - f_\infty(x)]^2 \\
&= \sum_{x,y\in G^C} \pi(x)L(x,y)[f_\infty(y) - f_\infty(x)]^2 \\
&\leq \sum_{x,y\in E} \pi(x)L(x,y)[f_\infty(y) - f_\infty(x)]^2 = \mathcal{D}(f_\infty) < \infty.
\end{aligned}$$

Then, we have

$$
\begin{aligned}
0 \leq \lim_{k\to\infty} \Big\{ &\sum_{x\in G_k} \sum_{y\in G^C - G_k} \pi(x)L(x,y)[f_\infty(y) - f_\infty(x)]^2 \\
&+ \sum_{x,y\in G^C - G_k} \pi(x)L(x,y)[f_\infty(y) - f_\infty(x)]^2 \Big\} \\
\leq \lim_{k\to\infty} \Big\{ &\sum_{x,y\in G^C} \pi(x)L(x,y)[f_\infty(y) - f_\infty(x)]^2 \\
&- \sum_{x,y\in G_k} \pi(x)L(x,y)[f_\infty(y) - f_\infty(x)]^2 \Big\} = 0.
\end{aligned}
$$

We know that

$$\lim_{k\to\infty} \sum_{\substack{x,y\in G_k \\ x\neq y}} \pi(x)L(x,y)f_\infty^2(x)$$

$$= \sum_{\substack{x,y\in G^C \\ x\neq y}} \pi(x)L(x,y)f_\infty^2(x)$$

$$\leq \sum_{\substack{x,y\in E \\ x\neq y}} \pi(x)L(x,y)f_\infty^2(x) = \sum_{x\in E} \pi(x)f_\infty^2(x)\sum_{\substack{y\in E \\ y\neq x}} L(x,y)$$

$$= \sum_{x\in E} \pi(x)f_\infty^2(x)\lambda(x) \leq \sum_{x\in E} \pi(x)f_\infty^2(x)\bar{\lambda} < \infty.$$

Then, we have

$$\sum_{x\in G_k}\sum_{y\in G^C-G_k}\pi(x)L(x,y)f_\infty^2(x)\le\sum_{x\in G_k}\sum_{y\in G^C-G_k}\pi(x)L(x,y)f_\infty^2(x)$$
$$+\sum_{x\in G^C-G_k}\sum_{y\in G_k}\pi(x)L(x,y)f_\infty^2(x)+\sum_{\substack{x,y\in G^C-G_k\\ x\neq y}}\pi(x)L(x,y)f_\infty^2(x)$$
$$=\sum_{\substack{x,y\in G^C\\ x\neq y}}\pi(x)L(x,y)f_\infty^2(x)-\sum_{\substack{x,y\in G_k\\ x\neq y}}\pi(x)L(x,y)f_\infty^2(x),$$

which leads to

$$\lim_{k\to\infty}\sum_{x\in G_k}\sum_{y\in G^C-G_k}\pi(x)L(x,y)f_\infty^2(x)$$
$$\le\lim_{k\to\infty}\left[\sum_{\substack{x,y\in G^C\\ x\neq y}}\pi(x)L(x,y)f_\infty^2(x)-\sum_{\substack{x,y\in G_k\\ x\neq y}}\pi(x)L(x,y)f_\infty^2(x)\right]=0.$$

Finally, we get

$$\lim_{k\to\infty}[J_\lambda(f_\infty) - J_\lambda(f_k)] = \lim_{k\to\infty}|J_\lambda(f_\infty) - J_\lambda(f_k)| = 0,$$

which is the same as

$$J_\lambda(f_\infty) = \lim_{k\to\infty} J_\lambda(f_k) \geq \lim_{k\to\infty} \inf_{f\in\mathcal{E}(G_k)} J_\lambda(f),$$

which is a contradiction with

$$\inf_{f\in\mathcal{E}(G_\infty)} J_\lambda(f) < J_\lambda(f_\infty) < \lim_{k\to\infty} \inf_{f\in\mathcal{E}(G_k)} J_\lambda(f).$$

Therefore, we have

$$\lim_{k\to\infty} \inf_{f\in\mathcal{E}(G_k)} J_\lambda(f) = \inf_{f\in\mathcal{E}(G_\infty)} J_\lambda(f).$$

Let $\bar{x} \in G^C$. Since $G^C = \{x_1, x_2, \ldots, \}$, there exists $m \in \mathbb{N}$ such that $x_m = \bar{x}$. Choose $k \geq m \in \mathbb{N}$. Let $f \in \mathcal{E}(G_k)$. Then there exists $y_0 \in \mathbb{R}$ such that $f(x) = y_0, \forall x \in G^C - G_k$. Denote $y_j := f(x_j), \forall 1 \leq j \leq k$. Therefore

$$\begin{aligned} J_\lambda(f) =& \pi(G) + \sum_{x \in G^C} \pi(x) f^2(x) + \frac{1}{\lambda} \sum_{x \in G^C} \pi(x)[f(x) - 1]^2 q(x) \\ &+ \frac{1}{2\lambda} \sum_{\substack{x,y \in G^C \\ y \neq x}} \pi(x) L(x,y)[f(y) - f(x)]^2. \end{aligned}$$

Then, we get

$$
\begin{aligned}
J_\lambda(f) =& \pi(G) + \sum_{x \in G_k} \pi(x) f^2(x) + \sum_{x \in G^C - G_k} \pi(x) f^2(x) \\
&+ \frac{1}{\lambda} \sum_{x \in G_k} \pi(x)[f(x) - 1]^2 q(x) + \frac{1}{\lambda} \sum_{x \in G^C - G_k} \pi(x)[f(x) - 1]^2 q(x) \\
&+ \frac{1}{2\lambda} \sum_{\substack{x, y \in G_k \\ y \neq x}} \pi(x) L(x,y)[f(y) - f(x)]^2 \\
&+ \frac{1}{2\lambda} \sum_{x \in G_k} \sum_{y \in G^C - G_k} \pi(x) L(x,y)[f(y) - f(x)]^2 \\
&+ \frac{1}{2\lambda} \sum_{x \in G^C - G_k} \sum_{y \in G_k} \pi(x) L(x,y)[f(y) - f(x)]^2 \\
&+ \frac{1}{2\lambda} \sum_{\substack{x, y \in G^C - G_k \\ y \neq x}} \pi(x) L(x,y)[f(y) - f(x)]^2.
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
J_\lambda(f) =& \pi(G) + \sum_{i=1}^{k} \pi(x_i) y_i^2 + y_0^2 \sum_{x \in G^C - G_k} \pi(x) + \frac{1}{\lambda} \sum_{i=1}^{k} \pi(x_i)[y_i - 1]^2 q(x_i) \\
&+ \frac{[y_0 - 1]^2}{\lambda} \sum_{x \in G^C - G_k} \pi(x) q(x) + \frac{1}{2\lambda} \sum_{\substack{i,j=1 \\ j \neq i}}^{k} \pi(x) L(x_i, x_j)[y_j - y_i]^2 \\
&+ \frac{1}{2\lambda} \sum_{x \in G_k} \sum_{y \in G^C - G_k} \pi(x) L(x, y)[f(y) - f(x)]^2 \\
&+ \frac{1}{2\lambda} \sum_{x \in G_k} \sum_{y \in G^C - G_k} \pi(x) L(x, y)[f(x) - f(y)]^2.
\end{aligned}
$$

This leads to

$$
\begin{aligned}
J_\lambda(f) =& \pi(G) + \sum_{i=1}^{k} \pi(x_i) y_i^2 + y_0^2 \pi(G^C - G_k) + \frac{1}{\lambda} \sum_{i=1}^{k} \pi(x_i)[y_i - 1]^2 q(x_i) \\
&+ \frac{[y_0 - 1]^2}{\lambda} \sum_{x \in G^C - G_k} \pi(x) q(x) \\
&+ \frac{1}{2\lambda} \sum_{\substack{i,j=1 \\ j<i}}^{k} \Big[\pi(x_i) L(x_i, x_j)[y_j - y_i]^2 + \pi(x_i) L(x_i, x_j)[y_i - y_i j]^2 \Big] \\
&+ \frac{1}{\lambda} \sum_{i=1}^{k} \sum_{y \in G^C - G_k} \pi(x_i) L(x_i, y)[y_0 - y_i]^2
\end{aligned}
$$

Then, we get

$$J_\lambda(f) = \pi(G) + \sum_{i=1}^{k} \pi(x_i) y_i^2 + y_0^2 \pi(G^C - G_k) + \frac{1}{\lambda} \sum_{i=1}^{k} \pi(x_i)[y_i - 1]^2 q(x_i)$$

$$+ \frac{[y_0 - 1]^2}{\lambda} \sum_{x \in G^C - G_k} \pi(x) q(x) + \frac{1}{\lambda} \sum_{\substack{i,j=1 \\ j<i}}^{k} \pi(x_i) L(x_i, x_j)[y_j - y_i]^2$$

$$+ \frac{1}{\lambda} \sum_{i=1}^{k} \sum_{y \in G^C - G_k} \pi(x_i) L(x_i, y)[y_0 - y_i]^2.$$

Define $\Phi_k : \mathbb{R}^{k+1} \to \mathbb{R}$ by

$$\Phi_k(y_0, y_1, \ldots, y_k) = \pi(G) + \sum_{i=1}^{k} \pi(x_i) y_i^2 + y_0^2 \pi(G^C - G_k)$$

$$+\frac{1}{\lambda} \sum_{i=1}^{k} \pi(x_i)[y_i - 1]^2 q(x_i)$$

$$+\frac{[y_0 - 1]^2}{\lambda} \sum_{x \in G^C - G_k} \pi(x) q(x) + \frac{1}{\lambda} \sum_{\substack{i,j=1 \\ j<i}}^{k} \pi(x_i) L(x_i, x_j)[y_j - y_i]^2$$

$$+\frac{1}{\lambda} \sum_{i=1}^{k} \sum_{y \in G^C - G_k} \pi(x_i) L(x_i, y)[y_0 - y_i]^2.$$

Then $J_\lambda(f) = \Phi_k(y_0, y_1, \ldots, y_k)$.

We observe that $x \mapsto x^2$, $x \mapsto (x-1)^2$ and $(x,y) \mapsto (x-y)^2$ are convex functions in $\mathbb{R}$, $\mathbb{R}$ and $\mathbb{R}^2$, respectively.
Since $\pi(x) \geq 0, \forall x \in E$, $q(x) \geq 0, \forall x \in E$, $L(x,y) \geq 0, \forall x \neq y \in E$,$\lambda > 0$ and a finite linear combination (with non-negative coefficients) convex functions is convex, we have that Φ_k is convex in $\mathbb{R}^{k+1}$. Then Φ_k assumes its minimum where its gradient vanishes.

For every $i = 1, \ldots, k$ we have

$$\frac{\partial \Phi_k}{\partial y_i}(y_0, y_1, \ldots, y_k) = 2\pi(x_i)y_i + \frac{2\pi(x_i)[y_i - 1]q(x_i)}{\lambda}$$

$$+\frac{1}{\lambda}\sum_{\substack{j=1 \\ j\neq i}}^{k}\pi(x_i)L(x_i, x_j)2[y_i - y_j] + \frac{1}{\lambda}\sum_{y\in G^C - G_k}\pi(x_i)L(x_i, y)2[y_i - y_0]$$

$$=2\pi(x_i)\Big\{y_i + \frac{1}{\lambda}[y_i - 1]q(x_i) + \frac{1}{\lambda}\sum_{\substack{j=1 \\ j\neq i}}^{k}L(x_i, x_j)[y_i - y_j]$$

$$+\frac{1}{\lambda}\sum_{y\in G^C - G_k}L(x_i, y)[y_i - y_0]\Big\}.$$

We have

$$y_i + \frac{1}{\lambda}[y_i - 1]q(x_i) + \frac{1}{\lambda}\sum_{\substack{j=1 \\ j\neq i}}^{k} L(x_i, x_j)[y_i - y_j] + \frac{1}{\lambda}\sum_{y\in G^C - G_k} L(x_i, y)[y_i - y_0]$$

$$= f(x_i) + \frac{1}{\lambda}[f(x_i) - 1]\sum_{y\in G} L(x_i, y) + \frac{1}{\lambda}\sum_{\substack{j=1 \\ j\neq i}}^{k} L(x_i, x_j)[f(x_i) - f(x_j)]$$

$$+ \frac{1}{\lambda}\sum_{y\in G^C - G_k} L(x_i, y)[f(x_i) - f(y)]$$

$$= f(x_i) + \frac{1}{\lambda}\sum_{y\in G} L(x_i, y)[f(x_i) - 1] + \frac{1}{\lambda}\sum_{\substack{y\in G_k \\ y\neq x_i}} L(x_i, y)[f(x_i) - f(y)]$$

$$+ \frac{1}{\lambda}\sum_{y\in G^C - G_k} L(x_i, y)[f(x_i) - f(y)].$$

Then, we get

$$y_i + \frac{1}{\lambda}[y_i - 1]q(x_i) + \frac{1}{\lambda}\sum_{\substack{j=1 \\ j\neq i}}^{k} L(x_i, x_j)[y_i - y_j] + \frac{1}{\lambda}\sum_{y\in G^C - G_k} L(x_i, y)[y_i - y_0]$$

$$= f(x_i) + \frac{1}{\lambda}\sum_{y\in G} L(x_i, y)[f(x_i) - f(y)] + \frac{1}{\lambda}\sum_{\substack{y\in G^C \\ y\neq x_i}} L(x_i, y)[f(x_i) - f(y)]$$

$$= f(x_i) + \frac{1}{\lambda}\sum_{\substack{y\in E \\ y\neq x_i}} L(x_i, y)[f(x_i) - f(y)]$$

$$= f(x_i) + \frac{1}{\lambda}\sum_{y\in E} L(x_i, y)[f(x_i) - f(y)]$$

$$= f(x_i) - \frac{1}{\lambda}\sum_{y\in E} L(x_i, y)[f(y) - f(x_i)] = f(x_i) - \frac{1}{\lambda}(Lf)(x_i).$$

This leads to

$$\frac{\partial \Phi_k}{\partial y_i}(y_0, f(x_1), \ldots, f(x_k)) = \frac{\partial \Phi_k}{\partial y_i}(y_0, y_1, \ldots, y_k)$$

$$= 2\pi(x_i)\Big\{ y_i + \frac{1}{\lambda}[y_i - 1]q(x_i) + \frac{1}{\lambda}\sum_{\substack{j=1 \\ j \neq i}}^{k} L(x_i, x_j)[y_i - y_j]$$

$$+ \frac{1}{\lambda} \sum_{y \in G^C - G_k} L(x_i, y)[y_i - y_0]\Big\}$$

$$= 2\pi(x_i)\Big\{ f(x_i) - \frac{1}{\lambda}(Lf)(x_i)\Big\}.$$

Since $\pi(x) > 0, \forall x \in G^C$, we have that $\frac{\partial \Phi_k}{\partial y_i}(y_0, f(x_1), \ldots, f(x_k)) = 0$ if and only if $(Lf)(x_i) = \lambda f(x_i)$. Since Φ_k attains its mininum where its gradient vanish and $J_\lambda(f) = \Phi_k(y_0, f(x_1), \ldots, f(x_k))$, we have that a function f_k which minimizes the functional J_λ on $\mathcal{E}(G_k)$ must be such that $(Lf_k)(x) = \lambda f_k(x), \forall x \in G_k$. In particular, $(Lf_k)(\bar{x}) = \lambda f_k(\bar{x})$.

Then, for every $k \geq m$, we have that a function f_k which minimizes the functional J_λ on $\mathcal{E}(G_k)$ must be such that $(Lf_k)(\bar{x}) = \lambda f_k(\bar{x})$. Since

$$\lim_{k \to \infty} \inf_{f \in \mathcal{E}(G_k)} J_\lambda(f) = \inf_{f \in \mathcal{E}(G_\infty)} J_\lambda(f),$$

we have that a function f_∞ which minimizes the functional J_λ on $\mathcal{E}(G_\infty)$ must be such that $(Lf_\infty)(\bar{x}) = \lambda f_\infty(\bar{x})$. Since $\bar{x}$ is an arbitrary element in G^C, we hava that

$$(Lf_\infty)(x) = \lambda f_\infty(x), \forall x \in G^C.$$

Therefore, f_∞ satisfies (9) and the claim is proved.