CORNELL UNIVERSITY MATHEMATICS DEPARTMENT SENIOR THESIS

The Logarithmic Sobolev Constant of Some Finite Markov Chains

A THESIS PRESENTED IN PARTIAL FULFILLMENT
OF CRITERIA FOR HONORS IN MATHEMATICS

Wai Wai Liu

May 2006

BACHELOR OF ARTS, CORNELL UNIVERSITY

THESIS ADVISOR(S)

Laurent Saloff-Coste
Department of Mathematics

The Logarithmic Sobolev Constant of some finite Markov Chains

1. Introduction

The Perron-Frobenius Theorem asserts that an ergodic Markov chain converges to its stationary distribution at an exponential rate given asymptotically by the second largest eigenvalue, in modulus. However, most applications of finite Markov chains, such as the Metropolis algorithm, requires the knowledge of how many steps are needed for the chain to be closed to equilibrium. The asymptotic result does not give answer to this question.

The logarithmic Sobolev constant was introduced in 1975 by Leonard Gross in the study of diffusion processes and Dirichlet forms. A logarithmic Sobolev inequality is an inequality of the form

$$\begin{aligned} &\int_\Omega |f(x)|^2 \ln |f(x)|^2 \nu(dx) \\ \leq\quad & \rho \int_\Omega |\mathrm{grad}\, f(x)|^2 \nu(dx) + \int_\Omega |f(x)|^2 \nu(dx) \ln \int_\Omega |f(x)|^2 \nu(dx), \end{aligned}$$

where ν is a probability measure on the space Ω and the logarithmic Sobolev constant α is the inverse of the smallest ρ such that the inequality holds for all f in the Dirichlet space. In the context of finite Markov chains, it turns out that α is closely related to the time to equilibrium (Theorem 1).

The natural question to ask is whether one can compute or estimate α. Unlike the problem of finding eigenvalues, finding α seems to be a very difficult problem. In this paper, we will present some techniques for estimating α and some examples which we compute the exact value of α.

Acknowledgements: I would like to thank both my faculty thesis advisor Professor Laurent Saloff-Coste and my peer thesis advisor Guan-Yu Chen. Without their encouragement and help, this thesis would not exist.

2. Preliminaries

A discrete time Markov chain on a finite state space $\mathcal{X}$ is a stochastic process $X \in \mathcal{X}^{\mathbb{N}}$, X_0 with some starting distribution μ on $\mathcal{X}$ that satisfies the following

$$\mathbb{P}\{X_{n+1} = y | X_n = x, X_{n-1} = x_{n-1}, ..., X_1 = x_1, X_0 = x_0\} = P_{xy}$$

for all states $x_0, x_1, ..., x_{n-1}, x, y$ in $\mathcal{X}$ and $n \geq 0$. In words, this says that the probability of making a jump to state y in the next step given the present state is x is completely independent of the past and time.

We can identify a Markov chain on a finite space $\mathcal{X}$ by a kernel K satisfying

$$K(x,y) \geq 0, \sum_{y \in \mathcal{X}} K(x,y) = 1.$$

The associated Markov operator is defined by

$$Kf(x) = \sum_{y \in \mathcal{X}} f(y) K(x,y).$$

The iterated kernel K^n is defined by

$$K^n(x,y) = \sum_{z \in \mathcal{X}} K^{n-1}(x,z) K(z,y).$$

If we set $K_x^n(y) = K^n(x,y)$, then $K_x^n(\cdot)$ defines a probability measure on $\mathcal{X}$ which

represents the distribution of the discrete Markov chain starting at x after exactly n steps.
We say a probability measure π on $\mathcal{X}$ is invariant with respect to K if
$\sum_{x\in\mathcal{X}} \pi(x)K(x,y) = \pi(y).$
That is, starting with distribution π and moving according to the kernel K leaves the distribution of the chain unchanged. If we assume K is irreducible, i.e. for each $x, y \in \mathcal{X}$ there is an n such that $K^n(x,y) > 0$, then the invariant measure π is unique and $\pi(x) > 0$ for all $x \in \mathcal{X}$ We will further assume that K is reversible, that is $\frac{K(x,y)}{\pi(y)} = \frac{K(y,x)}{\pi(x)}$.

The continuous semigroup associated with K is defined by

$$H_t f(x) = e^{-t(I-K)} = e^{-t}\sum_{i=0}^{\infty}\frac{t^i K^i f}{i!}.$$

It follows from the definition that

$$\begin{aligned} H_{t+s} &= H_t H_s \\ \lim_{t\to 0} H_t &= I \end{aligned}$$

If we set $H_t^x(y) = H_t(x,y)$. Then $H_t(\cdot)$ is a probability measure on $\mathcal{X}$ which represents the distribution at time t of the continuous time Markov chain $(X_t)_{t>0}$ associated with K and started at x. We can interpret the process as follows. The Markov chain makes a jump according to the rule described by the discrete time Markov chain with kernel K started at x after independent exponential(1) waiting times. Therefore the probability that the chain has made exactly i jumps at time t is $e^{-t}t^i/i!$ and the probability to be at y after exactly i jumps at time t is $e^{-t}t^iK^i(x,y)/i!$.
The densities of the probability measures K_x^n, H_t^x with respect to the stationary measure π will be denoted by

$$k_x^n = k^n(x,y) = \frac{K^n(x,y)}{\pi(y)}$$

and

$$h_t^x(y) = h_t(x,y) = \frac{H_t(x,y)}{\pi(y)}$$

3. The spectral gap and logarithmic Sobolev constant

Definition 1. The form $\mathcal{E}(f,g) = E_\pi[g(I-K)f]$ is called the Dirichlet form associated with the finite Markov chain(K,π).

Notice in the case reversible (K,π), the above definition implies

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

To see this, observe $\mathcal{E}(f,f) = ||f||_2^2 - E_\pi[fKf]$ and

$$\begin{aligned} \frac{1}{2}\sum_{x,y}|f(x)-f(y)|^2K(x,y)\pi(x) &= \frac{1}{2}\sum_{x,y}(f(x)^2+f(y)^2-2f(x)f(y))K(x,y)\pi(x) \\ &= ||f||_2^2 - E_\pi(fKf) \end{aligned}$$

Definition 2. Let K be a Markov kernel with Dirichlet form $\mathcal{E}$.The spectral gap $\lambda = \lambda(K)$ is defined by

$$\lambda = \min\left\{\frac{\mathcal{E}(f,f)}{\mathrm{Var}_\pi(f)}; \mathrm{Var}_\pi(f) \neq 0\right\}$$

Observe that λ is not, in general, an eigenvalue of $(I-K)$. If (K,π) is reversible then λ is the smallest non-zero eigenvalue of $I-K$.

Definition 3. Let K be an irreducible Markov chain with stationary measure π.The logarithmic Sobolev constant $\alpha = \alpha(K)$ is defined by

$$\alpha = \inf\left\{\frac{\mathcal{E}(f,f)}{\mathcal{L}(f)}; \mathcal{L}(f) \neq 0\right\}$$

where $\mathcal{L}(f) = \sum_{x\in X} |f(x)|^2 \log\left(\frac{|f(x)|^2}{||f||_2^2}\right)\pi(x)$.

Notice $\mathcal{L}(f)$ is nonnegative by Jensen's inequality applied to the convex function $\phi(t) = t^2 \log t^2$. Furthermore $\mathcal{L}(f) = 0$ if and only if f is constant.
It follows from the definition that α is the largest constant c such that the logarithmic Sobolev inequality

$$c\mathcal{L}(f) \leq \mathcal{E}(f,f)$$

holds for all functions f. Since $\mathcal{L}(f) = \mathcal{L}(|f|)$ and $\mathcal{E}(|f|,|f|) \leq \mathcal{E}(f,f)$, one can restrict f to be real nonnegative in the definition of α.

Theorem 1. *Let (K,π) be an irreducible and reversible Markov kernel. Set*

$$T_p = T_p(K, 1/e) = \min\left\{t > 0 : \max_t ||h_t^x - 1||_p \leq 1/e\right\},$$

and define $\pi_* = \min\{x \in \mathcal{X} : \pi(x)\}$, $\log_+(x) = \max\{0, \log(x)\}$, then for $1 \leq p \leq 2$,

$$\begin{aligned} \frac{1}{\lambda} \leq \quad T_p \quad &\leq \frac{1}{2\lambda}\left(2 + \log\frac{1}{\pi_*}\right), \\ T_p \quad &\leq \frac{1}{4\alpha}\left(4 + \log_+ \log\frac{1}{\pi_*}\right) \end{aligned}$$

and for $2 < p \leq \infty$,

$$\begin{aligned} \frac{1}{\lambda} \quad &\leq T_p \leq \quad \frac{1}{\lambda}\left(1 + \log\frac{1}{\pi_*}\right), \\ \frac{1}{2\alpha} \quad &\leq T_p \leq \quad \frac{1}{3\alpha}\left(4 + \log_+ \log\frac{1}{\pi_*}\right). \end{aligned}$$

From the above theorem, we see that α gives a much sharper bound to the time for the Markov chain to be close to its equilibrium than the bound obtained using the spectral gap. However, unlike the spectral gap, which can be found by finding the eigenvalues of K, we have limited tools for finding the logarithmic Sobolev constant.

4. Some techniques for bounding α

Theorem 2. *For any chain K the log-Sobolev constant α and the spectral gap λ satisfy $2\alpha \le \lambda$*

Proof. Let g be real and set $f = 1 + \epsilon g$ and write, for ϵ sufficiently small

$$\begin{aligned} |f|^2 \log |f|^2 &= 2(1 + 2\epsilon g + \epsilon^2 |g|^2)(\epsilon g - \frac{\epsilon^2 |g|^2}{2} + O(\epsilon^3)) \\ &= 2\epsilon g + 3\epsilon^2 |g|^2 + O(\epsilon^3). \end{aligned}$$

and

$$\begin{aligned} |f|^2 \log ||f||_2^2 &= (1 + 2\epsilon g + \epsilon^2 |g|^2)(2\epsilon \pi(g) + \epsilon^2 ||g||_2^2 - 2\epsilon^2 (\pi(g))^2 + O(\epsilon^3)) \\ &= 2\epsilon\pi(g) + 4\epsilon^2 g\pi(g) + \epsilon^2 ||g||_2^2 - 2\epsilon^2(\pi(g))^2 + O(\epsilon^3) \end{aligned}$$

Thus,

$$|f|^2 \log \frac{|f|^2}{||f||_2^2} = 2\epsilon(g - \pi(g)) + \epsilon^2(3|g|^2 - ||g||_2^2 - 4g\pi(g) + 2(\pi(g))^2) + O(\epsilon^3)$$

and

$$\begin{aligned} \mathcal{L}(f) &= 2\epsilon^2(||g||_2^2 - (\pi(g))^2) + O(\epsilon^3) \\ &= 2\epsilon^2 \mathrm{Var}_\pi(g) + O(\epsilon^3) \end{aligned}$$

Now observe that $\mathcal{E}(f,f) = \epsilon^2 \mathcal{E}(g,g)$, use the variational characterizations of α and λ, and let ϵ tend to zero, we get the desired result. □

Corollary 1. *Let g be an eigenfunction correspond to the spectral gap λ. If $\pi(g^3) \neq 0$, then $\alpha < \lambda/2$.*

Proof. Follow the proof of Theorem 1, expand log up to ϵ^3 terms, we have

$$|f|^2 \log |f|^2 = 2\epsilon g + 3\epsilon |g|^2 + 2\epsilon^3 g^3/3 + O(\epsilon^4)$$

and

$$|f|^2 \log ||f||_2^2 = \epsilon^2 ||g||_2^2 + 2\epsilon^3 g ||g||_2^2 + O(\epsilon^4).$$

Because g is an eigenfunction corresponding to λ, it is perpendicular to the constant function and so $\pi(g) = 0$ and a lot of terms vanish. Thus,

$$|f|^2 \log \frac{|f|^2}{||f||_2^2} = 2\epsilon g + \epsilon^2(3|g|^2 - ||g||_2^2) + 2\epsilon^3 g^3/3 + O(\epsilon^4)$$

and

$$\mathcal{L}(f) = 2\epsilon^2 \mathrm{Var}_\pi(g) + 2\epsilon^3 \pi(g^3)/3 + O(\epsilon^4).$$

By choosing appropriate ϵ, we find

$$\frac{\mathcal{E}(f,f)}{\mathcal{L}(f)} = \frac{\mathcal{E}(g,g)}{2\mathrm{Var}_\pi(g) + 2\epsilon\pi(g^3)/3 + O(\epsilon^2)} < \lambda/2.$$

□

From the definition of α, it is not obvious why it cannot be zero. The next theorem ensures us that $\alpha(K) > 0$ for any finite irreducible Markov chain.

Theorem 3. *Let K be an irreducible Markov chain with stationary measure π.Let α be its log-Sobolev constant and λ its spectral gap. Then either $\alpha = \lambda/2$ or there exists a positive non-constant function u which is a solution of*

$$2u\log u - 2u\log ||u||_2 - \frac{1}{\alpha}(I-K)u = 0 \tag{3.1}$$

and such that $\alpha = \mathcal{E}(u,u)/\mathcal{L}(u)$. In particular, $\alpha > 0$

Proof. We can restrict ourselves to non-negative functions satisfying $\pi(f) = 1$ when looking for minimizer of $\mathcal{E}(f,f)/\mathcal{L}(f)$. Now, either there exists a non-constant non-negative minimizer u, or the infimum is attained at the constant function 1 where $\mathcal{E}(1,1) = \mathcal{L}(1) = 0$. In this second case, the proof of the previous theorem shows that we must have $\alpha = \lambda/2$ since for any function $g \not\equiv 0$ satisfying $\pi(g) = 0$,

$$\lim_{\epsilon\to 0}\frac{\mathcal{E}(1+\epsilon g, 1+\epsilon g)}{\mathcal{L}(1+\epsilon g)} = \lim_{\epsilon\to 0}\frac{\epsilon^2\mathcal{E}(g,g)}{2\epsilon^2\mathrm{Var}_\pi(g)} \geq \frac{\lambda}{2}$$

Hence, either $\alpha = \lambda/2$ or there must exist a non-constant non-negative function u which minimizes $\mathcal{E}(f,f)/\mathcal{L}(f)$. Notice that $\mathcal{E}(u,u)$ and $\mathcal{L}(u)$ define differentiable functions in each coordinate u_i. A minimizer must have all directional derivatives equal to zero and would satisfy (3.1). Finally, if $u \geq 0$ is not constant and satisfies (3.1) then $u > 0$. Suppose u vanishes at $x \in \mathcal{X}$ then $Ku(x) = 0$ from (3.1) and u must vanish at all points y such that $K(x,y) > 0$. By irreducibility, this would imply $u \equiv 0$, a contradiction. □

Lemma 1. *Let (K_i, π_i), i=1,...,d be reversible Markov chains on finite sets $\mathcal{X}_\rangle$ with spectral gaps λ_i and log-Sobolev constants α_i. Fix $\mu = (\mu_i)_1^d$ such that $\mu_i > 0$ and $\Sigma\mu_i = 1$. Then the product chain (K,π) on $\mathcal{X} = \prod_1^d \mathcal{X}_\rangle$ with kernel*

$$K(x,y) = \sum_{i=1}^{d}\mu_i\delta(x_1,y_1)...\delta(x_{i-1},y_{i-1})K_i(x_i,y_i)\delta(x_{i+1},y_{i+1})...\delta(x_d,y_d)$$

where $\delta(x,x) = 1$ and $\delta(x,y) = 0$ for $x \neq y$ and stationary measure $\pi = \bigotimes_1^d \pi_i$ satisfies

$$\lambda = \min_i\{\mu_i\lambda_i\}, \alpha = \min_i\{\mu_i\alpha_i\}$$

Proof. Let $\mathcal{E}_i$ denote the Dirichlet form associated to K_i, then the product chain K has Dirichlet form

$$\mathcal{E}(f,f) = \sum_1^d \mu_i\left(\sum_{x_j:j\neq i}\mathcal{E}_i(f,f)(x^i)\pi^i(x^i)\right)$$

where x^i is the sequence $(x_1,...,x_d)$ with x_i omitted, $\pi^i = \bigotimes_{l:l\neq i}\pi_l$ and

$$\mathcal{E}_i(f,f)(x^i) = \mathcal{E}_i(f(x_1,...,x_d), f(x_1,...,x_d))$$

has the meaning that $\mathcal{E}_i$ acts on the i^{th} coordinate whereas the other coordinates are fixed. It is enough to prove the result when $d = 2$. Let $f : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathbb{R}$ be a nonnegative function and set

$$F(x_2) = (\sum_{x_1} f(x_1,x_2)^2\pi_1(x_1))^{1/2}.$$

Write

$$\begin{aligned}\mathcal{L}(f) &= \sum_{x_1,x_2} |f(x_1,x_2)|^2 \log \frac{f(x_1,x_2)^2}{||f||_{2,\pi}^2}\pi(x_1,x_2) \\ &= \sum_{x_2} |F(x_2)|^2 \log \frac{F(x_2)^2}{||F||_{2,\pi_2}^2}\pi_2(x_2) \\ &\quad + \sum_{x_1,x_2} |f(x_1,x_2)|^2 \log \frac{f(x_1,x_2)^2}{F(x_2)^2}\pi(x_1,x_2) \\ &\leq [\mu_2\alpha_2]^{-1}\mu_2\mathcal{E}_2(F,F) + [\mu_1\alpha_1]^{-1}\sum_{x_2}\mu_1\mathcal{E}_1(f(\cdot,x_2),f(\cdot,x_2))\pi_2(x_2).\end{aligned}$$

Now, the triangle inequality

$$\begin{aligned}|F(x_2)-F(y_2)| &= |||f(\cdot,x_2)||_{2,\pi_1} - ||f(\cdot,y_2)||_{2,\pi_1}| \\ &\leq ||f(\cdot,x_2)-f(\cdot,y_2)||_{2,\pi_1}\end{aligned}$$

implies that

$$\mathcal{E}_2(F,F) \leq \sum_{x_1}\mathcal{E}_2(f(x_1,\cdot),f(x_1,\cdot))\pi_1(x_1).$$

Hence

$$\begin{aligned}\mathcal{L}(f) \leq\ & [\mu_2\alpha_2]^{-1}\sum_{x_1}\mu_2\mathcal{E}_2(f(x_1,\cdot),f(x_1,\cdot))\pi_1(x_1) \\ & +[\mu_1\alpha_1]^{-1}\sum_{x_2}\mu_1\mathcal{E}_1(f(\cdot,x_2),f(\cdot,x_2))\pi_2(x_2)\end{aligned}$$

which gives $\mathcal{L}(f) \leq \max_i\{1/[\mu_i\alpha_i]\}\mathcal{E}(f,f)$. Therefore $\alpha \geq \min_i[\mu_i\alpha_i]$. Testing on functions that depend only on one of the two variables shows that $\alpha = \min_i[\mu_i\alpha_i]$. The proof for λ is similar. □

Lemma 2. *Let $(\mathcal{X},\pi,(\mathcal{E},\mathcal{D}))$ and $(\tilde{\mathcal{X}},\tilde{\pi},(\tilde{\mathcal{E}},\tilde{\mathcal{D}}))$ be two Dirichlet spaces. Assume that there is a map $p:\tilde{\Omega}\to\Omega$ such that for any $f\in\mathcal{D}$ and $\tilde{f}=f\circ p\in\tilde{\mathcal{D}}$,$\tilde{\mathcal{E}}(\tilde{f},\tilde{f}) = \mathcal{E}(f,f)$. Assume further that μ is the pushforward of $\tilde{\pi}$ under p, i.e.,$\tilde{\pi}(\tilde{f})=\pi(f)$ for any measureable non-negative f on $\mathcal{X}$. Let $\tilde{\lambda}$ and $\tilde{\alpha}$ be the spectral gap and logarithmic Sobolev constant on $\tilde{\mathcal{X}}$. Then*

$$\alpha \geq \tilde{\alpha}, \lambda \geq \tilde{\lambda}.$$

In particular, if $\tilde{\lambda}=\tilde{\alpha}/2$ and $\tilde{\lambda}=\lambda$, then $\alpha=\lambda/2$.

Lemma 3. *Let (K,π) and (K',π') be two Markov chains on the same finite set $\mathcal{X}$. Assume that there exist A, $a>0$ such that*

$$\mathcal{E}' \leq A\mathcal{E}, \qquad a\pi \leq \pi'.$$

Then

$$\lambda' \leq \frac{A}{a}\lambda, \qquad \alpha' \leq \frac{A}{a}\alpha.$$

Proof. The inequality between the spectral gaps follows from the variational definition of λ and the formula

$$\mathrm{Var}(f) = \min_{c\in\mathbb{R}}\sum_x |f(x)-c|^2\pi(x).$$

To get the inequality relating the log-Sobolev constants, notice

$$\forall x, y \geq 0, \qquad x\log x - x\log y - x + y \geq 0$$

and

$$\begin{aligned}\mathcal{L}_\pi(f) &= \sum_x (|f(x)|^2 \log|f(x)|^2 - |f(x)|^2\log||f||_2^2 - |f(x)|^2 + ||f||_2^2)\pi(x) \\ &= \min_{c>0}\sum_x (|f(x)|^2\log|f(x)|^2 - |f(x)|^2\log c - |f(x)|^2 + c)\pi(x).\end{aligned}$$

Using the variational definition of α, we get the desired inequality. □

5. Examples for which α is known

Theorem 4. *Let $\mathcal{X} = \{0, 1\}$ be the two point space. Fix $0 < p \leq 1/2$. Consider the Markov kernel $K = K_p$ given by $K(0,0) = K(1,0) = p$, $K(0,1) = K(1,1) = 1-p$. The chain is reversible with respect to the stationary measure π_p where $\pi_p(0) = p$ and $\pi_p(1) = 1-p$. The logarithmic Sobolev constant of (K_p, π_p) on $\mathcal{X}$ is given by*

$$\alpha_p = \frac{1-2p}{\log[(1-p)/p]}$$

with $\alpha_{1/2} = 1/2$.

Proof. For the case $p = 1/2$, we write $f(0) = 1+s$, $f(1) = 1-s$, $0 \leq s < 1$. The fact that $\alpha_{1/2} = 1/2$ is equivalent to

$$(*) \quad \frac{1}{2}((1+s)^2\log(1+s)^2 + (1-s)^2\log(1-s)^2 - 2(1+s^2)\log(1+s^2)) \leq 2s^2.$$

Let $f(s) = (1+s)^2\log(1+s)^2 + (1-s)^2\log(1-s)^2 - 2(1+s^2)\log(1+s^2) - 4s$. Then

$$f'(s) = -8s + 4(1-s)\log(1+s) - 4\log(1-s) - 4s\log(1+s^2),$$

and

$$f''(s) = \frac{-8s}{1+s^2} - 4\log(1+s^2).$$

Now it is easy to see that on $[0,1)$, f has a maximum at 0, proving (*). For the case $p < 1/2$, equation (3.1) is equivalent to

$$pu^2 + (1-p)v^2 = 1, \tag{4-1}$$

$$2u\log u = \frac{1}{\alpha_p}(1-p)(u-v), \tag{4-2}$$

$$2v\log v = \frac{1}{\alpha_p}p(v-u). \tag{4-3}$$

Adding up (4-2) and (4-3) with weights p and $1-p$, we have

$$pu\log u + (1-p)v\log v = 0, \tag{4-4}$$

which is independent of α_p. Rewrite (4-1) and (4-4) as

$$\begin{aligned}(p, 1-p)\cdot(u^2-1, v^2-1) &= 0,\\ (p, 1-p)\cdot(u\log u, v\log v) &= 0,\end{aligned}$$

we see that (u^2-1, v^2-1) and $(u\log u, v\log v)$ are both perpendicular to $(p, 1-p)$, and so one is a multiple of the other, i.e.

$$\frac{u\log u}{u^2-1} = \frac{v\log v}{v^2-1}$$

Define

$$f(x) = \frac{x\log x}{x^2-1},$$

then

$$f'(x) = \frac{-1+x^2-\log x - x^2\log x}{(x^2-1)^2}$$

Let $g(x) = -1+x^2-\log x - x^2\log x$, then

$$\begin{aligned} g'(x) &= x - 1/x - 2x\log x \\ g''(x) &= 1 + 1/x^2 - 2 - 2\log x > 0 \quad \text{if } x < 1 \end{aligned}$$

Since $g'(1) = 0$, we conclude that $g' < 0$ on $(0,1)$, which in term shows $f' > 0$ on $(0,1)$. Therefore, f is strictly increasing on $(0,1)$.

Notice $f(x) = f(1/x)$ and so f is strictly decreasing on $(1,\infty)$. Therefore, $v = 1/u$ is the only way to have

$$\frac{u\log u}{u^2-1} = \frac{v\log v}{v^2-1}.$$

Using this relation and (4-1), we discover that the only solutions to the minimizer equation are $(u,v) = \left(\sqrt{\frac{1-p}{p}}, \sqrt{\frac{p}{1-p}}\right)$ and $(u,v) = (1,1)$. When $p \neq 1/2$, the non-constant solution gives us

$$\frac{\mathcal{E}(f,f)}{\mathcal{L}(f)} = \frac{1-2p}{\log[(1-p)/p]} < \frac{1}{2}.$$

Thus, by the virtue of Theorem 3, we have

$$\alpha_p = \frac{1-2p}{\log[(1-p)/p]}.$$

□

Theorem 5. *Let $\mathcal{X} = \mathbb{Z} \mod 2n, n \in \mathbb{N}$. Consider the simple random walk on $\mathcal{X}$, i.e. for $n > 1$, the Markov kernel K_n given by $K_n(x,y) = 1/2$ if $|x-y| = 1 \mod 2n$ and $K_n(x,y) = 0$ otherwise. $K_1(0,1) = K_1(1,0) = 1$ and $K_1(0,0) = K_1(1,1) = 0$. The chain is reversible with respect to the uniform measure and α_n is given by*

$$2\alpha_n = \lambda_n = 1 - \cos(\pi/n)$$

For proof, see [2]. With the above theorems and lemmas, we can get the following interesting results concerning logarithmic Sobolev constants.

Example 1. Fix $0 < p < 1$. Let $\mathcal{X} = \{0,1\}^d$. Define the Markov chain as follows. If the current state is x, we pick a coordinate i uniformly at random. If $x_i = 0$ we change it to 1 with probability $1-p$ and do nothing with probability $1-p$. If $x_i = 1$ we change it to 0 with probability p and do nothing with probability $1-p$.

This is the product chain of the asymmetric 2-point chain in Theorem 4 and by Lemma 1, this chain has spectral gap $\lambda = 1/d$ and log-Sobolev constant

$$\alpha = \frac{1-2p}{d\log[(1-p)/p]} \quad \text{for} \quad p \neq \frac{1}{2}$$

and

$$\alpha_{1/2} = \frac{1}{2d}$$

Example 2. Let $\mathcal{X}_0 = \{0, ..., d\}$ and define the kernel

$$K_0(i,j) = \begin{cases} 0 & \text{if } |i-j| > 1; \\ (1-p)(1-i/d) & \text{if } j = i+1; \\ pi/d & \text{if } j = i-1; \\ (1-p)i/d + p(1-i/d) & \text{if } i = j. \end{cases}$$

This Markov chain counts the number of 1's in the chain in the previous example with and has stationary measure

$$\pi_0(j) = p^j(1-p)^{d-j}\binom{d}{j}$$

If we define $p : \{0,1\}^d \to \{0, ..., d\}$ by $p(x) = \sum_{i=1}^d x_i$, then $\tilde{\psi}(x) = |x| - (1-p)d$ is an eigenfunction with eigenvalue $\tilde{\lambda} = 1/d$ that can pass to $\mathcal{X}_0$ as $\psi(x) = x-(1-p)d$. Lemma 2 gives us

$$\alpha_0 \geq \frac{1-2p}{d\log[(1-p)/p]}$$

with

$$\alpha_0 = \frac{\lambda}{2} = \frac{1}{2d} \ \text{ when } p = \frac{1}{2}.$$

Example 3. Let $\mathcal{X} = \{0, ..., n-1\}$ and

$$K(i,j) = \begin{cases} 0 & \text{if } |i-j| > 1; \\ 1/2 & \text{if } j = i+1,\ i < n-1; \\ 1/2 & \text{if } j = i-1,\ i > 0; \\ 1/2 & \text{if } i = j \text{ and } i = 0 \text{ or } i = n-1. \end{cases}$$

We compare it with the 2n-cycle. Define $p : \{0, ..., 2n-1\} \to \{0, ..., n-1\}$ by $p(x) = 2n - x - 1 \mod n$. The 2n-cycle has $\tilde{\lambda} = 1 - \cos(\pi/n)$ with eigenfunctions $e^{\pm \pi i x/n}$. The function $f(x) = \cos(\frac{\pi}{n}(x+\frac{1}{2}))$ lies in the 2 dimensional eigenspace and has the property that $f(x) = f(2n-x-1)$ and therefore passes to $\{0, ..., n-1\}$. By Lemma 2 and Theorem 4, $\alpha = \lambda/2 = (1-\cos(\pi/n))/2$

6. The 5-cycle

Chen and Shen proved that $\alpha = \lambda/2$ for the simple random walk on the discrete circle $\mathbb{Z} \mod n$ when n is even(Theorem 5). Computation of the precise value of α for odd $n \geq 5$ remains open. In this section, we prove that $\alpha = \lambda/2$ for $n = 5$.

Lemma 4. *Let f be a function defined on non-negative real numbers by $f(t) = t\log(t)$ if $t > 0$ and $f(0) = 0$ Then*

$$\forall t \geq 0, \quad f(t) \geq t - 1 \tag{4.1}$$

$$\forall t \geq s \geq 0 \ \textit{ and } \ t + s \leq 2, \quad f(t) - f(s) \leq t - s \tag{4.2}$$

$$(4.3) \qquad \forall t \geq s \geq 1, \quad f(t) - f(s) \geq t - s$$

Proof. Consider $g(t) = f(t) - (t-1)$, then

$$g'(t) = \log(t) \quad \textit{and} \quad g''(t) = 1/t > 0$$

This implies g is convex with $g'(1) = 0$ and $g(1) = 0$ and finishes the proof of (4.1)
For (4.2), we fix $s \geq 0$ and reset $g(t)$ for $t \geq s$ by

$$\begin{aligned} g(t) &= f(t) - f(s) - (t-s)f'\left(\frac{t+s}{2}\right) \\ &= t\log(t) - s\log(s) - (t-s)\left(1 + \log\left(\frac{t+s}{2}\right)\right) \end{aligned}$$

We take the first and second derivatives of g and get

$$g'(t) = \log\left(\frac{2t}{t+s}\right) - \frac{t-s}{t+s}$$

and

$$g''(t) = \frac{s(s-t)}{t(t+s)^2}$$

So g is concave and $g'(s) = 0$ implies $g(t) \leq g(s) = 0$, which means

$$f(t) - f(s) \leq (t-s)\left(1 + \log\left(\frac{t+s}{2}\right)\right)$$

By assumption $t + s \leq 2$ and so (4.2) is proved.
Apply Mean Value Theorem on f, we get for any $t \geq s$, there exists a number $c(s,t) \in (s,t)$ such that

$$f(t) - f(s) = (t-s)f'(c(s,t))$$

If $s \geq 1$, $f'(c(s,t)) \geq 1$ and (4.3) is proved. □

Theorem 6. *The logarithmic Sobolev constant of the simple random walk on 5-cycle satisfies* $\alpha = \lambda/2$*.*

Before proving the theorem, we first prove the following lemmas. They give some properties that a non-constant minimizer of the 5-cycle must satisfy if $\alpha < \lambda/2$.

Lemma 5. *Let* $u = (a,b,c,d,e)$ *be a non-constant positive vector satisfying* $a \geq b \geq c \geq d \geq e \geq 0$ *and no three consecutive numbers are the same. Let* $f(x)$ *be defined as in Lemma 4.Suppose they satisfy the following system of equations*

$$(1) \qquad 2a - (b+c) = 4\alpha f(a),$$

$$(2) \qquad 2b - (a+d) = 4\alpha f(b),$$

$$(3) \qquad 2c - (a+e) = 4\alpha f(c),$$

$$(4) \qquad 2d - (c+e) = 4\alpha f(d),$$

$$(5) \qquad 2e - (c+d) = 4\alpha f(e),$$

$$(6) \qquad a^2 + b^2 + c^2 + d^2 + e^2 = 5$$

where α is a constant. If $\alpha < \frac{1}{2}(1-\cos(2\pi/5))$, then

$$b > 1; \tag{5.1}$$

$$2c > a+e > 2 > b+d, if \;\; c > 1; \tag{5.2}$$

$$d \leq 1; \tag{5.3}$$

Proof. First, observe that $\frac{1}{2}(1-\cos(2\pi/5))$ is the smaller root of the equation $g(t)=0$, where $g(t)=16t^2-20t+5=(2-4t)(3-4t)-1$. It is easy to verify that $g(\alpha)>0$.

To show (5.1), we suppose $b \leq 1$. By (4.1), the RHS of (2) is bounded below by $4\alpha(b-1)$ and the LHS can be rewritten as $2(b-1)-(a+d-2)$. This implies $(2-4\alpha)(b-1) \geq (a+d-2)$. Since $\alpha < 1/2$ and $b \leq 1$, we have $a+d \leq 2$, which implies $c+e \leq 2$. Subtract (4) from (1) and apply (4.2), we have

$$2(a-d)-(c-e) \leq 4\alpha(a-d)$$

or equivalently

$$(2-4\alpha)(a-d)-(c-e) \leq 0. \tag{7}$$

Similarly subtracting (5) from (3) and apply (4.2) will produce

$$(3-4\alpha)(c-e)-(a-d) \leq 0. \tag{8}$$

The contradiction comes from adding (7) $\times$ $(3-4\alpha)$ + (8), which gives $g(\alpha)(a-d) \leq 0$ or $g(\alpha) < 0$.

For (5.2), suppose $c > 1$. By (3), $a+e < 2c$. Subtract (2) from (1) and apply (4.3), we get

$$(3-4\alpha)(a-b)-(c-d) \geq 0. \tag{9}$$

Similarly, subtract (5) from (4) and apply (4.2), we get

$$(3-4\alpha)(e-d)-(c-b) \geq 0. \tag{10}$$

Now add up (3),(9),(10) produces

$$(2-4\alpha)(a+e-b-d) \geq 4\alpha c\log(c) \geq 0.$$

Hence, $a+e \geq b+d$. Suppose $a+e \leq 2$, then apply (4.2) to (1) – (5) and (2) – (4) respectively will produce

$$(2-4\alpha)(a-e)-(b-d) \leq 0, \tag{11}$$

$$(3-4\alpha)(b-d)-(a-e) \leq 0. \tag{12}$$

$(3-4\alpha) \times (11)-(12)$, we get $g(\alpha)(a-e) \leq 0$, which gives $g(\alpha) \leq 0$. The last inequality comes from the fact that $a+b+c+d+e \leq 5$

For (5.3), if $c > 1$, by (5.2) we have $b+d < 2$. But $b > 1$ by (5.1), so $d < 1$. If $c \leq 1$, the result is trivial. □

Remark 1. If $\alpha < \lambda/2$, a positive minimizer $u=(a,b,c,d,e)$ of the 5-cycle with $||u||_2 = 1$ must satisfy $a \geq b \geq c \geq d \geq e$. The arrangement of (a,b,c,d,e) is illustrated in the picture below.

Recall

$$\alpha = \inf\left\{\frac{\mathcal{E}(f,f)}{\mathcal{L}(f)}; \mathcal{L}(f) \neq 0\right\}.$$

If u is a minimizer of the 5-cycle with $||u||_2 = 1$, we have

$$\alpha = \frac{(a-b)^2 + (a-c)^2 + (b-d)^2 + (c-e)^2 + (d-e)^2}{a^2 \log a^2 + b^2 \log b^2 + c^2 \log c^2 + d^2 \log d^2 + e^2 \log e^2}.$$

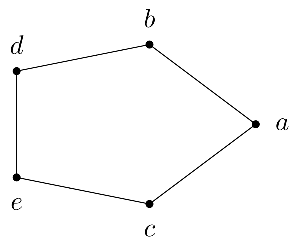
Notice the relative order of a, b, c, d, e does not change the value of the denominator, and without loss of generality, we can assume a is the largest. By symmetry, suppose b is the smallest. It is not difficult to see that the order $a \geq c \geq e \geq d \geq b$ will minimize the numerator. However, by swapping b and e, the difference in the numerator is

$$(a-b)^2 + (c-e)^2 - (a-e)^2 - (c-b)^2.$$

Using the fact that

$$x + y = x' + y' \text{ and } |x-y| > |x'-y'| \Rightarrow x^2 + y^2 > x'^2 + y'^2$$

We conclude that $(a-b)^2 + (c-e)^2 - (a-e)^2 + (c-b)^2 > 0$, and so b can not be the smallest. Therefore, by symmetry, we may assume e is the smallest and conclude $a \geq c \geq e$ and $a \geq b \geq d \geq e$. Suppose $d > c$, by swapping d and c, the difference we get is $(a-c)^2 + (b-d)^2 - (a-d)^2 - (b-c)^2$. If $a \geq b \geq d \geq c$, this will be positive, meaning the α we get is not optimal. If $b < c$, by swapping b and c, the difference we get is $(b-d)^2 + (c-e)^2 - (c-d)^2 - (b-e)^2$, which is clearly positive. Therefore, we must have $a \geq b \geq c \geq d \geq e$.



Lemma 6. *Let* (a, b, c, d, e) *be a vector satisfying* $a^2 + b^2 + c^2 + d^2 + e^2 = 5$ *and* $0 \leq e \leq d \leq c \leq b \leq a$. *Suppose they satisfy the following system of equations*

$$g(a) = b + c, \tag{13}$$

$$g(b) = a + d, \tag{14}$$

$$g(c) = a + e, \tag{15}$$

$$g(d) = b + e, \tag{16}$$

$$g(e) = c + d, \tag{17}$$

where $g(x) = 2x - 4\alpha x \log x$, α *is a constant. If* $\alpha < \frac{1}{2}(1 - \cos(2\pi/5))$, *then* $d = e$. *In other words, if* $\alpha < \lambda/2$, *the minimizer of the 5-cycle is symmetric.*

Proof. By equations (15) and (17), we get $a = g(g(e) - d) - e$. Similarly, by equations (14) and (16), we get $a = g(g(d) - e) - d$. If d, e solves the above system, we must have $g(g(e) - d) - g(g(d) - e) = e - d$. By (5.3) in the previous lemma, it suffices to consider $d \leq 1$.

For any fixed $d \leq 1$, define

$$g_1 : [0, d] \bigcap \{x | g(x) \geq d\} \to \mathbb{R}$$

by

$$g_1(e) = g(g(e) - d) - g(g(d) - e).$$

The idea of the proof is to show that for any fixed $d \leq 1$, $g_1(e)$ and $g_2(e) = e - d$ intersect only at the point $e = d$.
To achieve this goal, we utilize the convexity of g_1 and g_2. If we can show that for any fixed $d \leq 1$, g_1 is concave with slope greater than 1, it follows that g_1 and g_2 intersect only at the point $e = d$. Taking derivatives of g_1, we have

$$\begin{aligned} g_1'(e) &= g'(g(e) - d) \times g'(e) + g'(g(d) - e) \\ g_1''(e) &= g'(g(e) - d) \times g''(e) + g''(g(e) - d) \times g'(e)^2 - g''(g(d) - e) \\ g_1'''(e) &= g'(g(e) - d) \times g'''(e) + g''(g(e) - d) \times g'(e) \times g''(e) \\ &\quad + g''(g(e) - d) \times 2g'(e) \times g''(e) + g'''(g(e) - d) \times g'(e)^3 \\ &\quad + g'''(g(d) - e). \end{aligned}$$

Notice $g(e) - d \leq g(d) - d \leq 4\alpha \exp(1/4\alpha - 1)$ and $g'(x) = 2 - 4\alpha - 4\alpha \log x$, so

$$\begin{aligned} g'(g(e) - d) &\geq 2 - 4\alpha - 4\alpha(\log(4\alpha) + 1/4\alpha - 1) \\ &= 1 - 4\alpha \log(4\alpha) \\ &\geq 1 - 2(1 - \cos(2\pi/5)) \log(2(1 - \cos(2\pi/5))) > 0. \end{aligned}$$

Combing the facts that $g''(x) = -4\alpha/x < 0$ and $g'''(x) = 4\alpha/x^2 > 0$, we have $g_1''' > 0$. Therefore, if we can show $g_1''(d) < 0$, g_1 will be concave.
We define $h(d) = g_1''(d) = g'(g(d) - d) \times g''(d) + g''(g(d) - d) \times g'(d)^2 - g''(g(d) - d)$

$$\begin{aligned} h'(d) &= g''(g(d) - d) \times [g'(d) - 1] \times g''(d) + g'(g(d) - d) \times g'''(d) \\ &\quad + g'''(g(d) - d) \times [g'(d) - 1' \times g'(d)^2 + g''(g(d) - d) \times 2g'(d) \times g''(d) \\ &\quad - g'''(g(d) - d) \times [g'(d) - 1] \\ &= [g'(d) - 1] \times g'''(g(d) - d) \times (g'(d)^2 - 1) \\ &\quad + g''(g(d) - d) \times g''(d) \times [3g'(d) - 1] + g'''(d) g'(g(d) - d) \\ &= [g'(d) - 1]^2 \times g'''(g(d) - d) \times [g'(d) + 1] \\ &\quad + g''(g(d) - d) \times g''(d) \times [3g'(d) - 1] + g'''(d) g'(g(d) - d). \end{aligned}$$

Since $g'(d) \geq g'(1) = 2 - 4\alpha > 2 - 2(1 - \cos(2\pi/5))$, $3g'(d) - 1 > 0$, and so $h'(d) > 0$ for all $d \leq 1$.

$$\begin{aligned} h(1) &= (2 - 4\alpha)(-2\alpha) + (-2\alpha)(2 - 4\alpha)^2 + 2\alpha \\ &= -2\alpha[(2 - 4\alpha)(3 - 4\alpha) - 1] < 0. \end{aligned}$$

So we conclude that $h(d) = g_1''(d) < 0$ for all $d \leq 1$. Therefore, for each fixed $d \leq 1$, the function g_1 defined is concave.
Our proof of the lemma will be complete once we show that for any fixed $d \leq 1$, $g_1'(d) > 1$. Define $h_1(d) = g_1'(d) = g'(g(d) - d) \times [1 + g'(d)]$. Then
$h_1'(d) = g''(g(d) - d) \times [g'(d)^2 - 1] + g''(d) \times g'(g(d) - d)$.
If $g'(d) \geq 1$, $h_1'(d) < 0$.
If $g'(d) < 1$, since $g(d) - d \geq d$ if $d \leq 1$, $h_1'(d) \leq g''(d)[g'(g(d) - d) + g'(d)^2 - 1]$.
Define $h_2(d) = g'(g(d) - d) + g'(d)^2 - 1$,

$$\begin{aligned} h_2'(d) &= g''(g(d) - d) \times [g'(d) - 1] + 2 \times g'(d) \times g''(d) \\ &\leq g''(d) \times [3g'(d) - 1] < 0. \end{aligned}$$

Notice $h_2(1) > 0$, so $h_2(d) > 0$ if $d \leq 1$, which implies $h_1'(d) < 0$ if $d \leq 1$. Therefore, $h_1(d) > h_1(1) > 1$ if $d \leq 1$. This shows $g_1'(d) > 1$ for any fixed $d \leq 1$, proving the fact that g_1 and g_2 intersect only at $e = d$. Using (16) and (17), we conclude that a minimizer must be symmetric. □

Lemma 7. *Let* $K : \{1,2,3\} \times \{1,2,3\}$ *be the Markov kernel defined by*

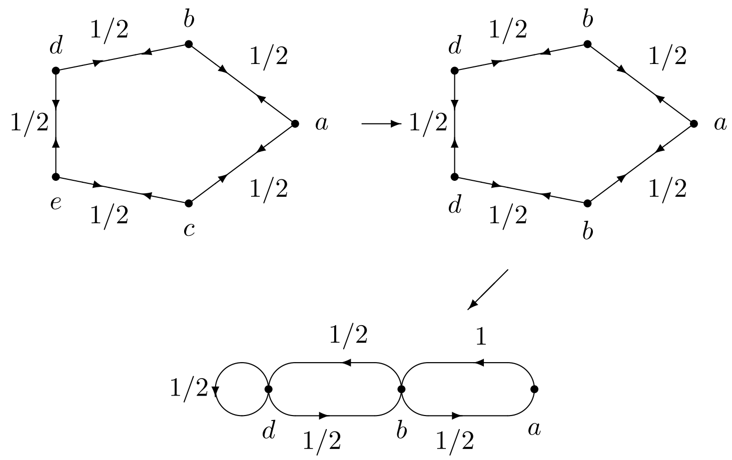
$$K = \begin{pmatrix} 0 & 1 & 0 \\ .5 & 0 & .5 \\ 0 & .5 & .5 \end{pmatrix}$$

with stationary measure $\pi = (1/5, 2/5, 2/5)$. *The log Sobolev constant* α *satisfies*

$$\alpha = \lambda/2 = \frac{1}{2}(1 - \cos(2\pi/5))$$

Remark 2. Notice Lemma 7 is a corollary of Theorem 6 by Lemma 2. With the result of Lemma 6, Lemma 7 is equivalent to Theorem 6. The collapse of the 5-cycle to the 3-point stick is illustrated in the figure below.

FIGURE 1. Collapse of 5-cycle to the 3-point stick



Proof. Suppose the log Sobolev constant $\alpha < \frac{1}{2}(1 - \cos(\frac{2\pi}{5}))$. By Theorem 3, there exists a positive non-constant vector $u = (u_1, u_2, u_3)$ that satisfies (3.1). We may assume without loss of generality that $u_1^2 + 2u_2^2 + 2u_3^2 = 5$ since $\mathcal{E}(kf, kf) = k^2\mathcal{E}(f, f)$ and $\mathcal{L}(kf) = k^2\mathcal{L}(f)$ so the ratio is not changed. Equation (3.1) is then equivalent to the following

$$2\alpha u_1 \log(u_1) = u_1 - u_2$$
$$4\alpha u_2 \log(u_2) = 2u_2 - u_1 - u_3$$
$$4\alpha u_3 \log(u_3) = u_3 - u_2$$

Define

$$f(x, \alpha) = x - 2\alpha x \log(x).$$

It is easy to check that if (u_1, u_2, u_3) is a non-constant vector satisfying the above system, then $u_2 = f(u_1, \alpha)$ and $u_3 = 2f(f(u_1, \alpha), \alpha) - u_1$, so there exists a $u_1 \neq 1$ such that

$$F(u_1, \alpha) = u_1^2 + 2[f(u_1, \alpha)]^2 + 2[2f(f(u_1, \alpha), \alpha) - u_1]^2 - 5 = 0.$$

By Lemma 6, since the minimizer is symmetric, if $u = (u_1, u_2, u_3)$ is a minimizer for the chain above, then $\tilde{u} = (u_1, u_2, u_2, u_3, u_3)$ will be a minimizer for the simple random walk on 5-cycle. If $u_1 < 1$, then $u_1 < u_2 = f(u_1, \alpha) < 1$ and so u_3 must be greater than 1 by the normalization constraint. But Lemma 6 tells us that the two smallest entries of the minimizer of the 5-cycle are equal, i.e. $u_1 = u_2$ which implies $u_1 = u_2 = u_3 = 1$ by the minimizer equation. Therefore, it suffice to consider $u_1 > 1$.

The key of the proof is to recognize that for any fixed u_1, $F(u_1, \alpha)$ is increasing decreasing in α if $1 < u_1 < 1.42$. Therefore, we can pick $F(u_1, \frac{1}{2}[1 - \cos(2\pi/5)])$ as a reference function and if $u_1 = 1$ is the only root of $F(u_1, \frac{1}{2}[1 - \cos(2\pi/5)])$ on $(1, 1.42)$, we may conclude that no non-constant minimizer exists, and so $\alpha = \lambda/2$. To see F is decreasing in α for $u_1 > 1$, take the first partial derivative of F with respect to α,

$$\begin{aligned} F_\alpha(u_1, \alpha) &= 4f(u_1, \alpha) \times f_\alpha(u_1, \alpha) \\ &\quad +4[2f(f(u_1, \alpha), \alpha) - u_1] \times 2f_\alpha(f(u_1, \alpha), \alpha) \times [f_\alpha(u_1, \alpha) + 1]. \end{aligned}$$

Observe that for any fixed α, the set of u_1 that gives $u_2, u_3 > 0$ is an interval. On this interval, we have

$$F_\alpha(u_1, \alpha) = 4u_2 \times f_\alpha(u_1, \alpha) + 4u_3 \times 2f_\alpha(u_2, \alpha) \times [f_\alpha(u_1, \alpha) + 1].$$

Since $f_\alpha(x, \alpha) = -2x \log x$, the first term in the sum is negative if $u_1 > 1$. $f_\alpha(u_2, \alpha) < 0$ as $u_1 > 1$ implies $u_2 > 1$. Since we consider only $u_1 < 1.42$, it is easy to check that $f_\alpha(u_1, \alpha) + 1 > 0$ on this interval and so the second term is also negative.

From now on, I will denote $f(x, \frac{1}{2}[1 - \cos(2\pi/5)])$ by $f(x)$.

The first and second derivatives of $F(u_1, \frac{1}{2}(1 - \cos(2\pi/5)))$ are

$$\begin{aligned} F' &= 2u_1 + 4f(u_1)f'(u_1) + 4(2f(f(u_1)) - u_1)(2f'(f(u_1))f'(u_1) - 1) \\ F'' &= 2 + 4f'(u_1)^2 + 4f(u_1)f''(u_1) + 4(2f'(f(u_1))f'(u_1) - 1)^2 \\ &\quad +4(2f(f(u_1)) - u_1)(2f''(u_1)f'(f(u_1)) + 2f''(f(u_1))f'(u_1)^2). \end{aligned}$$

When evaluated at 1,

$$\begin{aligned} F' &= 2[(3 - 4\alpha)(2 - 4\alpha) - 1] = 0 \\ F'' &= 2[(3 - 4\alpha)(2 - 4\alpha) - 1](8\alpha^2 - 8\alpha + 1) = 0. \end{aligned}$$

since $\alpha = (1/2)(1 - \cos(2\pi/5))$ is a root of $(3 - 4\alpha)(2 - 4\alpha) - 1$.

The third derivative of F is

$$F''' = 12f'(u_1)f''(u_1) + 4f(u_1)f'''(u_1) + h(u_1)$$

where

$$\begin{aligned} h(u_1) &= 4f(u_1)f'''(u_1) \\ &\quad +12[2f'(f(u_1))f'(u_1) - 1][2f''(f(u_1))f'(u_1)^2 + 2f'(f(u_1))f''(u_1)] \\ &\quad +4[2f(f(u_1)) - u_1][2f'''(u_1)f'(f(u_1)) + 2f''(u_1)f''(f(u_1))f'(u_1) \\ &\qquad +2f'''(f(u_1))f'(u_1)^3 + 4f''(f(u_1))f'(u_1)f''(u_1)]. \end{aligned}$$

Using the facts that $f(u_1) > u_1$ if $u_1 > 1$, f' is decreasing, f'' is negative and f''' is positive, one can see that h is positive on the interval $(1, 1.42)$.

To see $F''' > 0$ on (1,1.42), notice

$$\begin{aligned} & 12f'(u_1)f''(u_1) + 4f(u_1)f'''(u_1) \\ = \quad & 12(1-2\alpha-2\alpha\log u_1)\frac{-2\alpha}{u_1} + 4(u_1 - 2\alpha u_1 \log u_1)\frac{2\alpha}{u_1^2} \\ = \quad & \frac{8\alpha}{u_1}[1-2\alpha\log u_1 - (3-6\alpha-6\alpha\log u_1)] \\ = \quad & \frac{8\alpha}{u_1}(-2+6\alpha+4\alpha\log u_1) > 0 \end{aligned}$$

when $\alpha = \frac{1}{2}(1-\cos(2\pi/5))$.

Therefore, $F(u_1, \frac{1}{2}[1-\cos(2\pi/5)]) > 0$ on $(1, 1.42)$ by the convexity of F. Combining all the results, for any fixed $\alpha < \frac{1}{2}(1-\cos(2\pi/5))$, the only solution to $F(u_1, \alpha) = 0$ is $u_1 = 1$, thus the lemma is proved. □

Remark 3. In the search of a minimizer, we only need to consider $u_1 < 1.42$. Recall if $u = (u_1, u_2, u_3)$ is a minimizer for the above chain, $\tilde{u} = (u_1, u_2, u_2, u_3, u_3)$ is a minimizer for the 5-cycle. By (5.1) in Lemma 5, we have $u_2 > 1$ and by (5.3) in Lemma 5 we have $u_1 + u_3 > 2$. Therefore, if $u_1 > 1.42$, $u_1^2 + 2u_3^2 > 2.6892$ and $u_2 = f(u_1) > 1.0759$, so $u_1^2 + 2u_2^2 + 2u_3^2 > 5$

7. Some other examples of 3 point chain

Theorem 7. *For $0 \le p < 1$, let $K_p : \{1,2,3\} \times \{1,2,3\}$ be the Markov kernel defined by*

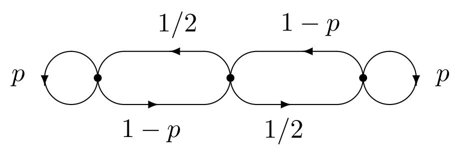
$$K_p = \begin{pmatrix} p & 1-p & 0 \\ .5 & 0 & .5 \\ 0 & 1-p & p \end{pmatrix}$$

with stationary distribution

$$\pi_p = (\frac{1}{4-2p}, \frac{2-2p}{4-2p}, \frac{1}{4-2p}).$$

Then $\alpha_p = \lambda_p/2$.

Figure 2. A family of 3 point chain with $\alpha = \lambda/2$



Proof. We proceed by contradiction. Observe that $\lambda_p = 1-p$. Suppose $\alpha_p < \lambda_p/2$. By Theorem 3, there exists a non-constant positive $u = (u_1, u_2, u_3)$ that satisfies (3.1). By symmetry, we may assume $u_1 \geq u_3$. Notice if u is a minimizer, ku is also a minimizer for all $k > 0$. Therefore, we may assume

$$u_1^2 + (2-2p)u_2^2 + u_3^2 = 4 - 2p \tag{7-1}$$

and (3.1) is equivalent to

$$\frac{2\alpha_p}{1-p} f(u_1) = u_1 - u_2 \tag{7-2}$$

$$4\alpha_p f(u_2) = 2u_2 - u_1 - u_3 \tag{7-3}$$

$$\frac{2\alpha_p}{1-p} f(u_3) = u_3 - u_2 \tag{7-4}$$

where $f(t) = t \log t$ is the function used in Lemma 4. Subtract (7-4) from (7-2) and rearrange the terms, we get

$$f(u_1) - f(u_3) = \frac{1-p}{2\alpha_p}(u_1 - u_3) > u_1 - u_3$$

since we assume $2\alpha_p < \lambda_p = 1-p$. By (4.2) in Lemma 4, we get $u_1 + u_3 > 2$. Since $u_1 + u_3 > 2$ implies $u_1^2 + u_3^2 > 2$, from (7-1) we get $u_2 < 1$.
By considering the combination (7-2)/u_1+(6-4)/u_3-(6-3)/u_2, we obtain

$$\frac{2\alpha}{1-p}\log(u_1u_3) - 4\alpha\log(u_2) = \frac{u_1}{u_2} - \frac{u_2}{u_1} + \frac{u_3}{u_2} - \frac{u_2}{u_3}.$$

Rearranging the terms yields

$$\frac{4\alpha p}{1-p}\log(u_2) = \left(\frac{u_1}{u_2} - \frac{u_2}{u_1} - \frac{2\alpha}{1-p}\log\frac{u_1}{u_2}\right) - \left(\frac{u_2}{u_3} - \frac{u_3}{u_2} - \frac{2\alpha}{1-p}\log\frac{u_2}{u_3}\right). \tag{7-5}$$

Define $h(t) = t - t^{-1} - k\log(t)$, and note that

$$h'(t) = \frac{(t-1)^2}{t^2} + \frac{2-k}{t} > 0 \quad if \quad k < 2$$

The LHS of (7-5) is always negative since $u_2 < 1$. By assumption,$2\alpha/(1-p) < 1 < 2$, so $h(u_1/u_2) - h(u_2/u_3) < 0$, which implies $u_1/u_2 < u_2/u_3$. This means $u_1u_3 < u_2^2 < 1$. Using this relation in (7-1) yields

$$\begin{aligned} 4-2p &> u_1^2 + (2-2p)u_1u_3 + u_3^2 \\ &> (u_1+u_3)^2 - 2p(u_1u_3) \\ &> 4 - 2p(u_1u_3). \end{aligned}$$

Rearranging the terms, we get $u_1u_3 > 1$, a contradiction. □

Theorem 8. *For $0 \leq p < 1$, let $K_p : \{0,1,2\} \times \{0,1,2\}$ be the Markov kernel defined by*

$$K_p = \begin{pmatrix} q & p & 0 \\ q & 0 & p \\ 0 & q & p \end{pmatrix}$$

with stationary distribution

$$\pi_p = \left(\frac{1}{1+p/q+(p/q)^2}, \frac{p/q}{1+p/q+(p/q)^2}, \frac{(p/q)^2}{1+p/q+(p/q)^2}\right),$$

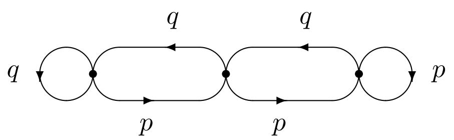
$p+q=1$. *Then*

$$\alpha_p = \frac{p-q}{2(\log p - \log q)}$$

with minimizer

$$u = (p/q, 1, q/p).$$

FIGURE 3. A family of 3 point chain with α known and not equal to $\lambda/2$



Proof. We compare this chain with another 3 point chain,

$$\tilde{K}_p = \begin{pmatrix} q & p & 0 \\ q/2 & 1/2 & p/2 \\ 0 & q & p \end{pmatrix}$$

with stationary distribution

$$\tilde{\pi}_p = \left(\frac{1}{1+2p/q+(p/q)^2}, \frac{2p/q}{1+2p/q+(p/q)^2}, \frac{(p/q)^2}{1+2p/q+(p/q)^2}\right),$$

The Dirichlet form associated with $(\tilde{K}, \tilde{\pi})$ is

$$\tilde{\mathcal{E}}_p(u,u) = \frac{(u_1-u_2)^2 p + (u_2-u_3)^2 p^2/q}{1+2p/q+(p/q)^2}.$$

The Dirichlet form associated with (K, π) is

$$\mathcal{E}_p(u,u) = \frac{(u_1-u_2)^2 p + (u_2-u_3)^2 p^2/q}{1+p/q+(p/q)^2}.$$

It is easy to see that

$$\tilde{\mathcal{E}}_p \le \frac{1+p/q+(p/q)^2}{1+2p/q+(p/q)^2}\mathcal{E}_p \;\text{ and }\; \frac{1+p/q+(p/q)^2}{1+2p/q+(p/q)^2}\pi_p \le \tilde{\pi}_p.$$

By Lemma 3, $\alpha_p \ge \tilde{\alpha}_p$.

Next, consider the product chain of K'_p, where K'_p is the asymmetric two point chain in Theorem 4 on $\{0,1\}^2$, with $\mu = (1/2, 1/2)$. Define

$$p : \{0,1\}^2 \to \{0,1,2\}$$

by

$$p(x) = |x|.$$

It follows from Lemma 2 that

$$\tilde{\alpha}_p \ge \frac{p-q}{2(\log p - \log q)}.$$

Therefore,

$$\alpha_p \geq \frac{p-q}{2(\log p - \log q)}.$$

By letting $u = (p/q, 1, q/p)$, we get

$$\frac{\mathcal{E}(u,u)}{\mathcal{L}(u)} = \frac{p-q}{2(\log p - \log q)}.$$

Thus,

$$\frac{p-q}{2(\log p - \log q)} \leq \tilde{\alpha_p} \leq \alpha_p \leq \frac{p-q}{2(\log p - \log q)},$$

proving the theorem. □

Proposition 1. *Fix $0 \leq p < 1$.Let $K : \{1,2,3\} \times \{1,2,3\}$ be the Markov kernel defined by*

$$K_p = \begin{pmatrix} 0 & 1 & 0 \\ .5 & 0 & .5 \\ 0 & 1-p & p \end{pmatrix}$$

with stationary measure

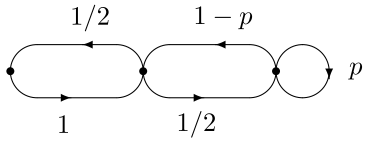
$$\pi_p = (\frac{1-p}{4-3p}, \frac{2-2p}{4-3p}, \frac{1}{4-3p}).$$

Then the log Sobolev constant α_p satisfies

$$\alpha_p = \lambda_p/2$$

only when $p = 0$ or $p = 1/2$.

FIGURE 4. A family of 3 point chain with $\alpha < \lambda/2$ except when $p = 0$ or $p = 1/2$



Proof. We utilize the result of Corollary 1. The spectral gap λ is

$$\lambda(K_p) = \frac{p - 1 + \sqrt{1+p^2}}{2}$$

with eigenfunction

$$\psi = (1, \frac{p-1+\sqrt{1+p^2}}{2}, (p-1)(p+\sqrt{1+p^2})).$$

$$E(\psi^3) = \frac{p(1-p)(p-1/2)[3-3p+6p^2-4p^3+\sqrt{1+p^2}(-2+12p-8p^2)]}{4-3p}.$$

Notice $3-3p+6p^2-4p^3$ is strictly decreasing in p and $\sqrt{1+p^2}(-2+12p-8p^2)$ is strictly increasing on $(0,.75)$ and not less than 2 on $[.75,1)$. Combining these observations, it is easy to see that

$$3-3p+6p^2-4p^3+\sqrt{1+p^2}(-2+12p-8p^2)>0 \text{ on } (0,1).$$

Therefore, $E(\psi^3)\neq 0$ unless $p=0$ or $p=1/2$. By Corollary 1, we have $\alpha_p<\lambda_p$ for $p\neq 0,1/2$.

The equality for $p=0,1/2$ is proved in Lemma 7 and Theorem 7. □

References

[1] Bobkov S. and Tetali P. *Modified log-Sobolev inequalities, mixing and hypercontractivity.* Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing, 287-296, ACM, New York 2003.

[2] Chen G-Y. and Sheu Y-C. *On the log-Sobolev constant for the simple random walk on the n-cycle: the even cases.* J. Funct. Anal 202 (2003), 473-485.

[3] Diaconis P. and Saloff-Coste L. *Comparison theorems for reversible Markov chains.* Ann. Appl. Probab. 3 (1993), 696-730.

[4] Diaconis P. and Saloff-Coste L. *Logarithmic Sobolev inequalities for finite Markov chains* Ann. Appl. Probab. 6 (1996), 695-750.

[5] Goel S. *Modified logarithmic Sobolev inequalities for some models of random walk.* Stochastic Process. Appl. 114 (2004), 51-79.

[6] Gross L. *Logarithmic Sobolev inequalities.* Amer. J. Math. 97 (1976), 1061-1083.

[7] Saloff-Coste L. *Lectures on finite Markov chains.* In Lectures on probability theory and statistics (Saint-Flour, 1996), 301-413, Lecture Notes in Math., 1665, Springer, Berlin, 1997.