APPLICATIONS OF GEOMETRIC BOUNDS TO THE CONVERGENCE RATE OF MARKOV CHAINS ON $\mathbf{R}^n$

by

Wai Kong Yuen*

(December 15, 1998.)

Acknowledgments. This work is part of my doctoral research done under the direction of Jeffrey S. Rosenthal. I thank Peter Rosenthal for helpful discussions about the operator theory issues.

Abstract. Quantitative geometric rates of convergence for reversible Markov chains are closely related to the spectral gap of the corresponding operator, which is hard to calculate for general state spaces. This article describes a geometric argument to give different types of bounds for spectral gaps of Markov chains on bounded subsets of $\mathbf{R}^n$ and to compare the rates of convergence of different Markov chains.

1. Introduction. As Markov chain Monte Carlo algorithms (Gelfand and Smith (1990), Smith and Roberts (1993)) are more widely used, quantitative geometric rates of convergence for Markov chains becomes an important topic. Diaconis (1988), Sinclair and Jerrum (1989), Jerrum and Sinclair (1988), Diaconis and Stroock (1991), Sinclair (1992) and Diaconis and Saloff-Coste (1993) proved general results on *finite* state spaces. Hanlon (1992), Frieze et al. (1994), Frigessi et al. (1993), Ingrassia (1994) and Belsley (1993) proved results specifically for Markov chain Monte Carlo. On *general* state spaces, not many results have been found yet. For partial results, see Amit and Grenander (1991), Amit (1991), Hwang et al. (1993), Lawler and Sokal (1988), Meyn and Tweedie (1994), Rosenthal (1995a, 1995b, 1996a, 1996b), Baxter and Rosenthal (1995) and Roberts and Rosenthal (1997a, 1997b).

* Department of Mathematics, University of Toronto, Toronto, Ontario, Canada M5S 3G3.Internet: yuen@math.toronto.edu

In particular, Diaconis and Stroock (1991) and Sinclair (1992) use geometric arguments with paths to bound to bound the second largest eigenvalue of a self-adjoint discrete time Markov chain. On the other hand, Lawler and Sokal (1988) took an idea from the literature of differential geometry and proved the Cheeger's inequality for positive-recurrent discrete-time Markov chains and continuous-time Markov jump process. For more on the estimation of eigenvalues that applies to nice continuous situations such as the Laplacian on various domains, see e.g. Bañuelos and Carroll (1994). Jerrum and Sinclair (1988) used a geometric argument with paths to bound the Cheeger's constant for a discrete-time finite space Markov chain. In this paper, we shall use an analogous geometric argument to bound the spectral radius and the Cheeger's constant for Markov chains on bounded subsets of $\mathbf{R}^n$.

In many cases, it is very difficult to give bounds to the convergence rate of a Markov chains. However, it is still possible to compare the rates for two different Markov chains. Diaconis and Saloff-Coste (1993) use a geometric argument to compare the eigenvalues of two discrete time reversible Markov chains. Mira (1998) introduces the covariance ordering for general Markov chains. In this paper, we shall use a geometric argument to compare two reversible Markov chains on $\mathbf{R}^n$ quantitatively.

The remainder of this section introduces the definitions of $\beta_1(L)$, the Cheeger's constant and their relation to the rate of convergence for Markov chains. In section 2, we define paths and give lower bounds for the constants when the Markov chain is defined in a bounded subset of $\mathbf{R}^n$. Then discuss some special cases which the paths are easy to find. In Section 4, we use similar geometric arguments to compare two reversible Markov chains defined on $\mathbf{R}^n$ quantitatively. Section 5 gives examples for the bounds and the comparison theorem.

1.1. *Basic notations.*

Consider a discrete-time Markov chain or a continuous time Markov jump process with measurable state space $(S,\mathcal{F})$, transition probability kernel $P(x,dy)$ and invariant probability measure π. Then P induces a positivity-preserving linear contraction on $L^2(\pi)$ by

$$(Pf)(x) = \int f(y)P(x,dy).$$

P also acts to the left on measures, so that

$$\mu P(A) = \int P(x,A)\mu(dx).$$

Recall that a Markov chain is *reversible* if

$$\pi(dx)P(x,dy) = \pi(dy)P(y,dx).$$

Hence, a Markov chain is reversible iff the operator P on $L^2(\pi)$ is self-adjoint. In this case, the spectrum is real and we can define

$$\lambda_0(P) = \inf \operatorname{spec}(P\,|_{\mathbf{1}^\perp}),$$

$$\lambda_1(P) = \sup \operatorname{spec}(P\,|_{\mathbf{1}^\perp}).$$

We consider the Laplacian $L = I - P$. Since L is also real and self-adjoint, $\lambda_1(P) = 1 - \beta_1(L)$, where

$$\beta_1(L) = \inf \operatorname{spec}(L\,|_{\mathbf{1}^\perp}) = \inf\{\frac{(f, Lf)_{L^2(\pi)}}{\|f\|^2_{L^2(\pi)}} : f \in \mathbf{1}^\perp, f \neq 0\}.$$

In section 2, we shall give a class of lower bounds for $\beta_1(L)$. On the other hand, we define the Cheeger's constant as follows:

$$k \equiv \inf_{\substack{A \in \mathcal{F} \\ 0 < \pi(A) < 1}} k(A)$$

with

$$k(A) \equiv \frac{\int \chi_A(x) P(x, A^c) \pi(dx)}{\pi(A)\pi(A^c)},$$

which is the rate of probability flow, in the stationary Markov chain, from a set A to its complement A^c, normalized by the invariant probabilities of A and A^c. Intuitively, if k is very small, or equivalently, there is a set A s.t. the flow from A to A^c is very small compared to the invariant probabilities of A and A^c, then the Markov chain must converge very slowly to the invariant distribution. In fact, Lawler and Sokal(1988) proved the *Cheeger's inequality* for reversible Markov chains:

$$k^2/8 \leq 1 - \lambda_1(P) \leq k.$$

As we shall see in Section 1.2, $\lambda_1(P)$ is closely related to the convergence rate to the invariant distribution. Roughly speaking, the bigger the $\lambda_1(P)$, the lower the rate. So, the Cheeger's inequality says that when k is small, the rate is low, which matches with our intuition.

1.2. *L^2 convergence to π.*

We extend the definition of L^2 norm for signed measure. Given a signed measure μ on S,define $\|\mu\|_2$ by

$$\|\mu\|_2^2 = \begin{cases} \int_S |\frac{d\mu}{d\pi}|^2 d\pi, & \mu << \pi; \\ \infty, & \text{otherwise.} \end{cases}$$

Hence, we can also represent $L^2(\pi)$ by $\{\mu; \|\mu\|_2 < \infty\}$, so that μ and f represent the same element whenever $f = \frac{d\mu}{d\pi}$ a.e.. Finally, set

$$\|P\mid_{\mathbf{1}^\perp}\|_2 = \sup\{\|Pf\|_2; f \in \mathbf{1}^\perp, \|f\|_2 = 1\}.$$

We shall use the fact that P is a bounded self-adjoint operator and a result from spectral theory (see e.g. Rudin (1991), Kreyszig (1978), Conway (1985)) to prove the following well-known proposition.

Proposition 1.1. For a discrete-time Markov chain, if P is reversible w.r.t. π and $\mu \in L^2(\pi)$, then

$$\|\mu P^n - \pi\|_2 \le \|\mu - \pi\|_2 \rho^n,$$

where

$$\rho = \max\{|\lambda_0(P)|, |\lambda_1(P)|\}.$$

Proof. For $\mu \in L^2(\pi)$, let $f = \frac{d\mu}{d\pi}$. Since P is reversible w.r.t.π, it is easy to check $\frac{d(\mu P)}{d\pi} = Pf$. So, P acting on $\mu \in L^2(\pi)$ is the same as P acting on $f \in L^2(\pi)$, in the sense that $d\mu = f d\pi$ implies $d(\mu P) = (Pf)d\pi$. Define $\|P\mid_{\mathbf{1}^\perp}\|_2$ by

$$\|P\mid_{\mathbf{1}^\perp}\|_2 = \sup\{\|Pf\|_2; f \in \mathbf{1}^\perp \, \|f\|_2 = 1\}.$$

It is also well known(see e.g. Baxter and Rosenthal (1995)) that $\|P\mid_{\mathbf{1}^\perp}\|_2 \le 1$. From a result of spectral theory for bounded self-adjoint operator P,

$$\|P\mid_{\mathbf{1}^\perp}\|_2 = \rho.$$

Furthermore, we observe that $f - \mathbf{1} \in \mathbf{1}^\perp$. Hence,

$$\begin{aligned}\|\mu P^n - \pi\|_2 &= \|\mu P^n - \pi P^n\|_2\\ &= \|(\mu - \pi)P^n\|_2\\ &= \|P^n(f - \mathbf{1})\|_2\\ &\le \|f - \mathbf{1}\|_2 \|P^n\mid_{\mathbf{1}^\perp}\|_2\\ &= \|\mu - \pi\|_2 \|P\mid_{\mathbf{1}^\perp}\|_2^n\\ &= \|\mu - \pi\|_2 \rho^n.\end{aligned}$$

Remarks.1. ρ is known as the *spectral radius* of the operator P. From the Cheeger's inequality, we have an upper bound for $\lambda_1(P)$. However, it is hard to bound $\lambda_0(P)$ in general (for *finite* state spaces, Diaconis and Stroock (1991) gave a geometric lower bound of it). If we assume $P(x,x) \ge a > 0 \forall x \in S$, then it directly implies that $\lambda_0(P) > -1 + 2a$.

So, we have $\rho \le \max\{1-2a, \lambda_1(P)\}$. In particular, if we consider the chain $\frac{1}{2}(I+P)$, then $a \ge \frac{1}{2}$ and so $\rho = \lambda_1(P)$. In this case, we have an upper bound for the convergence rate in terms of k. In practice, however, it is hard to calculate k numerically for general state spaces. We can also bound $\beta_1(L) = \lambda_1(P)$ directly. In section 2, we shall use a geometric argument to give a lower bound for k and a class of lower bounds for $\beta_1(L)$. Hence, two types of upper bounds for the convergence rate will be found.

2. For continuous-time Markov jump processes, the result is much simpler. Then the corresponding operator with mean 1 exponential holding times can be written as $\bar{P}^t = \sum_{n=0}^{\infty} e^{-t}\frac{t^n}{n!}P^n$. So, $\sigma(\bar{P}^t) = \{\sum_{n=0}^{\infty} e^{-t}\frac{t^n}{n!}\lambda^n : \lambda \in \sigma(P)\}$ and $\lambda_0(\bar{P}^t) \ge 0$. Hence, we have the following relation:

$$\|\mu\bar{P}^t - \pi\|_2 \le \|\mu - \pi\|_2 e^{-t\lambda_1(P)}.$$

3.For probability measures μ, we have $\|\mu - \pi\|_2^2 = \|\mu\|_2^2 - 1$.

4. We used the L^2 norm to measure the rate of convergence. For the relation between L^2 norm and other norms, see e.g. Roberts and Rosenthal (1997b).

2. Upper Bounds for $\lambda_1(P)$.

In this section, we prove a class of lower bounds for $\beta_1(L)$ and an upper bound of the Cheeger's constant for a reversible Markov chain (discrete or continuous time) on $\mathbf{R}^n$ described in 1.1. We further assume that the transition kernel is of the form $P(x, dy) = \alpha(x)\delta_x(dy) + p_x(y)dy$ where δ_x is the unit point mass on x for any $x \in \mathbf{R}^n$. Suppose the invariant distribution π has density $q(y)$ w.r.t. Lebesgue measure.

The next requirement is the existence of a set of paths satisfying some regularity conditions. Let $S = \{x \in \mathbf{R}^n : q(x) > 0\}$. $\Gamma = \{\gamma_{xy}\}$ is a *set of paths* for a Markov chain P on $\mathbf{R}^n$ if for any $x, y \in S$, there exists $b_{xy} \in \mathbf{N}$ and a path $\gamma_{xy} : \{0, \ldots, b_{xy}\} \to S$ s.t. $\gamma_{xy}(0) = x$ and $\gamma_{xy}(b_{xy}) = y$ and that $p_u(v) > 0$ for any edge (u, v) of any path where (u, v) is an *ith edge* of a path γ_{xy} iff $u = \gamma_{xy}(i-1)$ and $v = \gamma_{xy}(i)$ for some i. In this case, we write $(u, v) \in \gamma_{xy}$. Let E_i be the collection of all ith edges and $E = \cup E_i$. For any $u, v \in \mathbf{R}^n$, define $Q(u, v) = p_u(v)q(u)$ and for $\gamma_{xy} \in \Gamma$ and $\epsilon \in \mathbf{R}$, define $\|\gamma_{xy}\|_\epsilon = \sum_{(u,v)\in\gamma_{xy}} Q(u, v)^{-2\epsilon}$. Note that $Q(u, v) > 0$ for any edge (u, v).

To prove a class of lower bounds for $\beta_1(L)$, a set of paths Γ has to satisfy the following technical conditions: Let $V = \{(x, y, i) : x, y \in S, 1 \le i \le b_{xy}\}$. Define $T : V \to S^2 \times \mathbf{N}^2$ by $T(x, y, i) = (\gamma_{xy}(i-1), \gamma_{xy}(i), b_{xy}, i)$. Assume T is a 1-1 map onto $T(V)$. Fix $b, i \in \mathbf{N}$ s.t. $(u, v, b, i) \in T(V)$ for some $(u, v) \in S \times S$ and let $W_{bi} = \{(u, v) : (u, v, b, i) \in T(V)\} \subset E$. Define a 1-1 map $G_{bi} : W_{bi} \to S \times S$ by $G_{bi}(u, v) = (x, y)$ where $T(x, y, i) = (u, v, b, i)$. Suppose that G_{bi} can be extended to a bijection of open sets and has continuous partial derivatives a.e. with respect to the Lebesgue measure on $\mathbf{R}^n \times \mathbf{R}^n$ for each b, i. Let $J_{bi}(u, v)$ be the Jacobian of the change of variable $(x, y) = G_{bi}(u, v)$, given by

$$J_{bi}(u,v) = \begin{vmatrix} \frac{\partial x_1}{\partial u_1} & \cdots & \frac{\partial x_1}{\partial u_n} & \frac{\partial x_1}{\partial v_1} & \cdots & \frac{\partial x_1}{\partial v_n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial x_n}{\partial u_1} & \cdots & \frac{\partial x_n}{\partial u_n} & \frac{\partial x_n}{\partial v_1} & \cdots & \frac{\partial x_n}{\partial v_n} \\ \frac{\partial y_1}{\partial u_1} & \cdots & \frac{\partial y_1}{\partial u_n} & \frac{\partial y_1}{\partial v_1} & \cdots & \frac{\partial y_1}{\partial v_n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial y_n}{\partial u_1} & \cdots & \frac{\partial y_n}{\partial u_n} & \frac{\partial y_n}{\partial v_1} & \cdots & \frac{\partial y_n}{\partial v_n} \end{vmatrix}$$

where $u = (u_1, \ldots, u_n)$ and so on. We shall represent $J_{bi}(u, v)$ by $J_{xy}(u, v)$ as given $(u, v) \in E$, there is a 1-1 correspondence between (x, y) and (b, i). In this case, we say that Γ satisfies the *first regularity condition.* Intuitively, this condition means that the paths are 'differentiable' in the sense that slight change in the edge can only change a path containing it continuously and the corresponding Jacobian is well defined. With this condition, we can apply a change of variables to prove Theorem 2.1. We shall give some examples in Section A and B following Corollary 2.3 that if we construct some paths 'smooth' enough, the condition is easily verified.

The geometric constants that we shall use in the bounds are

$$k_\epsilon = k_\epsilon(\Gamma) = \text{esssup}_{(u,v)\in E}\{Q(u,v)^{-(1-2\epsilon)} \sum_{\gamma_{xy} \ni (u,v)} \|\gamma_{xy}\|_\epsilon q(x)q(y)|J_{xy}(u,v)|\},$$

where esssup is the essential supremum with respect to Lebesgue measure λ on $\mathbf{R}^n \times \mathbf{R}^n$, i.e., $\text{esssup}\{f(x) : x \in X\} = \inf\{a : \lambda\{x : f(x) > a\} = 0\}$, and the sum is taken over all (x, y) s.t. $\gamma_{xy} \ni (u, v)$.

Theorem 2.1. Let $\Gamma = \{\gamma_{xy}\}$ be a set of paths satisfying the first regularity condition for an irreducible Markov chain P on $\mathbf{R}^n$. Then for any $\epsilon \in \mathbf{R}$,

$$\beta_1(L) \geq \frac{1}{k_\epsilon}.$$

Hence, $\lambda_1(P) \leq \inf_\epsilon\{1 - \frac{1}{k_\epsilon}\}$.

Proof. Since L is real and self-adjoint, we need only to consider real trial functions. For real $f \in \mathbf{1}^\perp$, we have

$$(f, Lf)_{L^2(\pi)} = \frac{1}{2}\iint_{S\times S} (f(v) - f(u))^2 Q(u,v)dudv.$$

Now,

$$
\begin{aligned}
&\|f\|^2_{L^2(\pi)} \\
&= \int f^2(x)q(x)dx \\
&= \frac{1}{2}\iint_{S\times S}[f(x)-f(y)]^2 q(x)q(y)dxdy \\
&= \frac{1}{2}\iint_{S\times S}\left[\sum_{(u,v)\in\gamma_{xy}}\left(\frac{Q(u,v)}{Q(u,v)}\right)^{\epsilon}(f(v)-f(u))\right]^2 q(x)q(y)dxdy \\
&\le \frac{1}{2}\iint_{S\times S}\|\gamma_{xy}\|_{\epsilon} q(x)q(y)\sum_{(u,v)\in\gamma_{xy}} Q(u,v)^{2\epsilon}(f(v)-f(u))^2 dxdy \\
&= \frac{1}{2}\iint_{S\times S}\sum_{i=1}^{b_{xy}}\|\gamma_{xy}\|_{\epsilon} q(x)q(y)Q(\gamma_{xy}(i-1),\gamma_{xy}(i))^{2\epsilon}(f(\gamma_{xy}(i))-f(\gamma_{xy}(i-1)))^2 dxdy,
\end{aligned}
$$

where the inequality is Cauchy-Schwartz. Note that

$$
\iint_{S\times S}\sum_{i=1}^{b_{xy}} \cdot dxdy = \iiint_V \cdot\, \delta(di)dxdy,
$$

where δ is the counting measure on $\mathbf{N}$. By the first regularity condition, we can consider the change of variables $(x,y,i)=T^{-1}(u,v,b,i)$ and we have

$$
\begin{aligned}
\iiint_V \cdot\, \delta(di)dxdy &= \iiiint_{T(V)} \cdot\, |J_{bi}(u,v)|\delta(db)\delta(di)dudv \\
&\le \iint_E \sum_{\gamma_{xy}\ni(u,v)} \cdots |J_{xy}(u,v)|dudv
\end{aligned}
$$

as measure since $W_{bi}\subset E$ for each b and each i. Hence, we have

$$
\begin{aligned}
\|f\|^2_{L^2(\pi)} &\le \frac{1}{2}\iint_E \sum_{\gamma_{xy}\ni(u,v)} \|\gamma_{xy}\|_{\epsilon} q(x)q(y)Q(u,v)(f(v)-f(u))^{2\epsilon}|J_{xy}(u,v)|dudv \\
&\le \operatorname{esssup}_{(u,v)\in E}\{Q(u,v)^{-(1-2\epsilon)}\sum_{\gamma_{xy}\ni(u,v)}\|\gamma_{xy}\|_{\epsilon} q(x)q(y)|J_{xy}(u,v)|\} \\
&\quad \cdot \frac{1}{2}\iint_E Q(u,v)(f(v)-f(u))^2 dudv \\
&= k_{\epsilon}\cdot\frac{1}{2}\iint_E (f(v)-f(u))^2 Q(u,v)dudv \\
&\le k_{\epsilon}\cdot\frac{1}{2}\iint_{S\times S}(f(v)-f(u))^2 Q(u,v)dudv \\
&= k_{\epsilon}(f,Lf)_{L^2(\pi)}.
\end{aligned}
$$

Thus, $\beta_1(L) = \inf\{\frac{(f,Lf)_{L^2(\pi)}}{\|f\|^2_{L^2(\pi)}} : f \in \mathbf{1}^\perp, f \neq 0\} \geq \frac{1}{k_\epsilon}$ for any $\epsilon \in \mathbf{R}$. Since $\lambda_1(P) = 1 - \beta_1(L)$, the last statement follows trivially.

To prove a bound for the Cheeger's constant k, a set of paths Γ has to satisfy a modified version of the first regularity condition: For any $A \subset S$ and $0 < \pi(A) < 1$, let $V_A = A \times A^c$. Define $T_A : V_A \to S^2 \times \mathbf{N}^2$ by

$$T_A(x,y) = (\gamma_{xy}(l-1), \gamma_{xy}(l), b_{xy}, l),$$

where

$$l = l^A_{xy} = \min\{i | \gamma_{xy}(i-1) \in A, \gamma_{xy}(i) \in A^c\}.$$

Assume T_A is a 1-1 map onto $T_A(V_A)$. Fix $b, i \in \mathbf{N}$ s.t. $(u,v,b,i) \in T_A(V_A)$ for some (u,v) and A. Let $W_{bi} = \{(u,v) : (u,v,b,i) \in T_A(V_A)$ for some $A\} \subset E$. Define a map $G_{bi} : W_{bi} \to S \times S$ by $G_{bi}(u,v) = (x,y)$ if $T_A(x,y) = (u,v,b,i)$ for some A. Suppose G_{bi} is 1-1 and satisfies the same condition as in the first regularity condition. Then we can consider the same change of variable $(x,y) = G_{bi}(u,v)$ and Jacobian $J_{xy}(u,v)$. In this case, we say that Γ satisfies the *second regularity condition.* It is somewhat similar to the first regularity condition except that we require a slight adjustment to use a different change of variables to prove Theorem 2.2.

The geometric constant that we shall use in the bound is

$$K = K(\Gamma) = \operatorname{esssup}_{(u,v)\in E}\{Q(u,v)^{-1} \sum_{\gamma_{xy} \ni (u,v)} q(x)q(y)|J_{xy}(u,v)|\}.$$

Theorem 2.2. Let $\Gamma = \{\gamma_{xy}\}$ be a set of paths satisfying the second regularity condition for an irreducible Markov chain P on $\mathbf{R}^n$. Then the Cheeger's constant

$$k \leq \frac{1}{K}.$$

Hence, $\lambda_1(P) \leq 1 - \frac{1}{8K^2}$.

Proof. For any $A \subset S$ and $0 < \pi(A) < 1$,

$$\begin{aligned}\pi(A)\pi(A^c) &= \int_A q(x)dx \int_{A^c} q(y)dy \\ &= \iint_{A\times A^c} 2q(x)q(y)dxdy \\ &= \iint_{V_A} q(x)q(y)dxdy\end{aligned}$$

.

By the second regularity condition, we can consider the change of variables $(x,y) = T_A^{-1}(u,v,b,i)$ and we have

$$\begin{aligned}\iint_{V_A} dxdy &= \iiiint_{T_A(V_A)} |J_{bi}(u,v)|\delta(db)\delta(di)dudv \\ &\leq \iint_{V_A} \sum_{\gamma_{xy}\ni(u,v)} |J_{xy}(u,v)|dudv\end{aligned}$$

as measure. Hence, we have

$$\begin{aligned}\pi(A)\pi(A^c) &\leq \iint_{V_A} \sum_{\gamma_{xy}\ni(u,v)} q(x)q(y)|J_{xy}(u,v)|dudv \\ &= \iint_{V_A} Q(u,v)^{-1} \sum_{\gamma_{xy}\ni(u,v)} q(x)q(y)|J_{xy}(u,v)|Q(u,v)dudv \\ &\leq \text{esssup}_{(u,v)\in E}\{Q(u,v) \sum_{\gamma_{xy}\ni(u,v)} q(x)q(y)|J_{xy}(u,v)|\} \iint_{V_A} Q(u,v)dudv \\ &= K\iint_{A\times A^c} Q(u,v)dudv \\ &= K\int \chi_A(x)P(x,A^c)\pi(dx).\end{aligned}$$

Thus, the Cheeger's constant

$$k = \inf_{\substack{A\in\mathcal{F} \\ 0<\pi(A)<1}} \frac{\int \chi_A(x)P(x,A^c)\pi(dx)}{\pi(A)\pi(A^c)} \geq \frac{1}{K}.$$

By Cheeger's inequality, $\lambda_1(P) \leq 1 - \frac{k^2}{8} \leq 1 - \frac{1}{8K^2}$.

Remarks. 1. In the proofs of the theorems, $\alpha(x)$ doesn't play a role except making $P(x,dy)$ a probability measure. In the following examples, unless otherwise specified, we only define $p_x(y)$, and then $\alpha(x) = 1 - \int_S p_x(y)dy$ is assumed.

2. For transition kernel of the form we considered in the theorems, the reversibility condition is equivalent to $q(x)p_x(y) = q(y)p_y(x)$ for almost all $x,y \in S$.

3. There is no guarantee that such paths should exist in Theorem 2.1 and 2.2. All we say is that *if* such paths do exist, *then* the theorem holds. Moreover, we want to choose paths such that some of the geometric constants are finite, so that we have a bound strictly less than 1 for $\lambda_1(P)$. We shall discuss a special type of paths and simplify the constants in Corollary 2.3.

4. In Theorem 2.1, k_ϵ is a bound for *any* ϵ. In particular, when $\epsilon = 0$ or $\frac{1}{2}$, we have the following nice constants:

$$k_0 = \text{esssup}_{(u,v)\in E}\{Q(u,v)^{-1} \sum_{\gamma_{xy}\ni(u,v)} b_{xy} q(x)q(y)|J_{xy}(u,v)|\},$$

and

$$k_{\frac{1}{2}} = \text{esssup}_{(u,v)\in E}\{ \sum_{\gamma_{xy}\ni(u,v)} \|\gamma_{xy}\|_{\frac{1}{2}} q(x)q(y)|J_{xy}(u,v)|\}.$$

This is a continuous version of the results in finite state spaces (see e.g. Diaconis and Stroock (1991) and Sinclair (1992)).

5. It is not necessary that $(G^{bi})^{-1}$ has continuous partial derivatives for each i on the whole W_{bi} in the regularity conditions. Measure-zero sets can be neglected. The same is true in the following corollaries.

Now we consider a particular case when $b_{xy} = b$, a constant for any $x, y \in S$. Denote the paths by γ^b_{xy}. We can then simplify the regularity conditions as follows. For $i = 1, \dots, b$, define $T_i : S \times S \to S \times S$ by $T_i(x,y) = (\gamma^b_{xy}(i-1), \gamma^b_{xy}(i))$. Assume T_i can extended to an open set V containing $S\times S$ s.t. T_i is a 1-1 map onto $T_i(V)$ and that $(T_i)^{-1}$ has continuous partial derivatives a.e. for each i. Let $J_i(u,v)$ be the Jacobian of the change of variable $(x,y) = (T_i)^{-1}(u,v)$. From here we can apply Theorem 2.1 and 2.2 directly or we can enjoy a different version by the following corollary.

Corollary 2.3. Under the above assumptions, the geometric constants in Theorem 1.1 can be simplified as

$$k_\epsilon(\Gamma) = \sum_{i=1}^{b} \text{esssup}_{(u,v)\in E_i}\{Q(u,v)^{-(1-2\epsilon)}\|\gamma^b_{xy}\|_\epsilon q(x)q(y)|J_i(u,v)| : (x,y) = (T_i)^{-1}(u,v)\},$$

and the geometric constant in Theorem 1.2 can be simplified as

$$K(\Gamma) = \sum_{i=1}^{b} \text{esssup}_{(u,v)\in E_i}\Big\{Q(u,v)^{-1} q(x)q(y)|J_i(u,v)| : (x,y) = (T_i)^{-1}(u,v)\Big\}.$$

Proof. For k_ϵ, we follow the proof of Theorem 1.1 until the change of variables. Since $b_{xy} \equiv b$, we have

$$\iint_{S\times S} \sum_{i=1}^{b_{xy}} dxdy = \sum_{i=1}^{b} \iint_{S\times S} dxdy.$$

For each i, we can consider the change of variables $(x,y) = (T_i)^{-1}(u,v)$ by assumption, which gives

$$\sum_{i=1}^{b} \iint_{S\times S} dxdy = \sum_{i=1}^{b} \iint_{E_i} |J_i(u,v)| dudv.$$

From then on, it is easy to see that k_ϵ can be simplified as in the statement. For K, we follow the proof of Theorem 1.2 until the change of variables. Redefine $T_A : V_A \to S^2$ by

$$T_A(x, y) = (\gamma_{xy}^b(l-1), \gamma_{xy}^b(l)),$$

where

$$l = l_{xy}^A = \min\{i | \gamma_{xy}^b(i-1) \in A, \gamma_{xy}^b(i) \in A^c\}.$$

Note that

$$\begin{aligned}\iint_{V_A} dxdy &= \sum_{i=1}^{b} \iint_{(T_i)^{-1}(E_i \cap T_A(V_A))} dxdy \\ &= \sum_{i=1}^{b} \iint_{(E_i \cap T_A(V_A))} |J_i(u,v)| dudv\end{aligned}$$

.

From then on, it is easy to see K can be simplified as in the statement.

Remark 3 suggests that we have no definite way to find the paths in general. However, in the following special cases of S, we shall have at least one way to construct paths satisfying the regularity conditions. The geometric constants can be simplified as follows.

A. *S is convex.*

In this case, fix any $b \in \mathbf{N}$ and $x, y \in S$, we can choose the *uniform* path $\eta_{xy}^b : \{0, \ldots, b\} \to S$ defined by

$$\eta_{xy}^b(i) = \frac{(b-i)x + iy}{b},$$

the function onto the points on the line joining x, y so that adjacent points have equal spacings. In this case, $T_i(x, y) = (\eta_{xy}^b(i-1), \eta_{xy}^b(i))$ can be easily extended to a bijection on $\mathbf{R}^n \times \mathbf{R}^n$ and that $(T_i)^{-1}$ has continuous partial derivatives for each i. So, both the first and second regularity conditions are satisfied. It can be shown that $|J_i| = b^n$ (since for each j, x_j, y_j depend only on u_j, v_j).

On the other hand, we can also try the *linear* paths in which $b_{xy} = \lceil \frac{1}{c} \|y - x\| \rceil$ for some constant $c > 0$. Define $\gamma_{xy} : \{0, \ldots, b_{xy}\} \to S$ defined by

$$\gamma_{xy}(i) = \frac{(b_{xy} - i)x + iy}{b_{xy}}.$$

It is similar to the uniform paths except that the length of any edge $(u, v) \in E$ is less than or equal to c. Intuitively, a majority of edges have length close to c. It is easy to check that both the first and second regularity conditions are satisfied and that $|J_{xy}(u,v)| = b_{xy}^n$. We sum up the results in the following corollary:

Corollary 2.4. (a) Under the uniform paths $\{\eta_{xy}^b\}$ above, the constants can be simplified as

$$k_\epsilon = b^n \sum_{i=1}^{b} \text{esssup}_{(u,v)\in E_i}\{Q(u,v)^{-(1-2\epsilon)}\|\eta_{xy}^b\|_\epsilon q(x)q(y) : (x,y) = (T_i)^{-1}(u,v)\},$$

and

$$K = b^n \sum_{i=1}^{b} \text{esssup}_{(u,v)\in E_i}\{Q(u,v)^{-1}q(x)q(y) : (x,y) = (T_i)^{-1}(u,v)\}.$$

(b) Under the linear paths $\{\gamma_{xy}\}$ above, the constants can be simplified as

$$k_\epsilon = \text{esssup}_{(u,v)\in E}\{Q(u,v)^{-(1-2\epsilon)} \sum_{\gamma_{xy}\ni(u,v)} \|\gamma_{xy}\|_\epsilon q(x)q(y)b_{xy}^n\},$$

and

$$K = \text{esssup}_{(u,v)\in E}\{Q(u,v)^{-1} \sum_{\gamma_{xy}\ni(u,v)} q(x)q(y)b_{xy}^n\}.$$

B. *S is non-convex.*

As in the remark, it can be hard to find suitable paths. Now, we assume that S satisfies the following conditions. Let U be open in $\mathbf{R}^n$ and $S \subset U$. Suppose there exists $\phi : U \to \phi(U)$ s.t. $\phi(U)$ is open and $\phi(S)$ is convex. Assume that ϕ^{-1} exists and both ϕ, ϕ^{-1} are continuously differentiable a.e.. Since we already have uniform paths η_{xy}^b on $\phi(S)$, we can define paths γ_{xy} on S by

$$\gamma_{xy}(i) = \phi^{-1}(\frac{(b-i)\phi(\mathbf{x}) + i\phi(\mathbf{y})}{b}).$$

Define $\Phi : U \times U \to \phi(U) \times \phi(U)$ by $\Phi(x,y) = (\phi(x), \phi(y))$. So, the Jacobian of Φ exists and we can denote it by $J_\Phi(x,y)$.

To compute $J_i(u,v)$ in Corollary 1.2, we notice that $T_i = \Phi^{-1} \circ T_i^\eta \circ \Phi$ where T_i^η represents the T_i for normal paths. By the chain rule and the inverse function theorem of differentiation,

$$\begin{aligned}
|J_i(u,v)| &= |\det D(T_i)^{-1}(u,v)| \\
&= |\det D(\Phi \circ (T_i^\eta)^{-1} \circ \Phi^{-1})(u,v)| \\
&= |\det D\Phi((T_i^\eta)^{-1} \circ \Phi^{-1}(u,v))| \cdot |\det D(T_i^\eta)^{-1}(\Phi^{-1}(u,v))| \cdot |\det D\Phi^{-1}(u,v)| \\
&= |J_\Phi((T_i^\eta)^{-1} \circ \Phi^{-1}(u,v))| \cdot b^n \cdot |\frac{1}{\det D\Phi(\Phi^{-1}(u,v))}| \\
&= \frac{|J_\Phi((T_i^\eta)^{-1} \circ \Phi^{-1}(u,v))| \cdot b^n}{|J_\Phi(\Phi^{-1}(u,v))|}.
\end{aligned}$$

Moreover, if $0 < m \leq |J_\Phi| \leq M$,

$$|J_i(u,v)| \leq \frac{M \cdot b^n}{m}.$$

If $|J_\Phi|$ is constant, $|J_i(u,v)| = b^n$.

We can also use the linear paths instead of the uniform paths. In that case,

$$|J_{xy}(u,v)| = \frac{|J_\Phi((T^\eta_{xy})^{-1} \circ \Phi^{-1}(u,v))| \cdot b^n}{|J_\Phi(\Phi^{-1}(u,v))|},$$

where $T^\eta_{xy}(u,v)$ represents the corresponding bijection with respect to x, y for linear paths. Moreover, if $0 < m \leq |J_\Phi| \leq M$,

$$|J_{xy}(u,v)| \leq \frac{M \cdot b^n_{xy}}{m}.$$

If $|J_\Phi|$ is constant, $|J_{xy}(u,v)| = b^n_{xy}$.

3. Comparison of spectral radius of different Markov chains. In this section, we first discuss the comparison of λ_1 for two reversible Markov chains on general state spaces. Mira (1998) gives an ordering for chains with the same invariant probability measure. We use the min-max principal to compare λ_1 of two chains with not necessarily the same invariant probability measure. Then we use a geometric arguments to compare λ_1 for two reversible Markov chains on $\mathbf{R}^n$ quantitatively. In discrete state spaces, Diaconis and Saloff-Coste (1993) and Quastel (1992) have used some path arguments to compare reversible chains.

We consider two reversible discrete-time Markov chains (or continuous time Markov jump processes) with measurable state space $(S,\mathcal{F})$, transition probability kernels $P(x,dy)$, $\tilde{P}(x,dy)$ and invariant probability measure $\pi, \tilde{\pi}$ respectively. Recall that the spectral radius ρ of an operator H is defined as $\rho = \max\{|\lambda_0(H)|, |\lambda_1(H)|\}$. As we mentioned, it is hard to measure $\lambda_0(H)$. We shall compare λ_1 of P and $\tilde{P}$ by the operator form of min-max principle (see e.g. Reed and Simon (1978), Theorem XIII.1). Since 0 is the smallest eigenvalue for $L = I - P$ and L is bounded from below ($L \geq cI$ by taking $c = 0$), the min-max principle implies

$$1 - \lambda_1(P) = \beta_1(L) = \sup_{\phi \in L^2(\pi)} \inf_{\psi \in \phi^{\perp_\pi}} \frac{(\psi, L\psi)_\pi}{\|\psi\|^2_\pi},$$

where $\phi^{\perp_\pi}$ is shorthand for $\{\psi | (\psi,\phi)_\pi = 0\}$. The same is true by replacing L, π by $\tilde{L}, \tilde{\pi}$ respectively. Then we can extend the result of Diaconis and Saloff-Coste (1993) to a more general setting:

Theorem 3.1. Under the above assumptions, if there exists constants $a, A > 0$ s.t. (i) $(f, \tilde{L}f)_{\tilde{\pi}} \le A(f, Lf)_\pi$ for all $f \in L^2(\pi)$ and (ii) $\tilde{\pi} \ge a\pi$, then $\beta_1(L) \ge \frac{a}{A}\beta_1(\tilde{L})$, and so

$$\lambda_1(P) \le 1 - \frac{a}{A}(1 - \lambda_1(\tilde{P})).$$

Proof. By condition (ii), we have $a \cdot \int |f|^2 d\pi \le \int |f|^2 d\tilde{\pi}$, and so $L^2(\tilde{\pi}) \subset L^2(\pi)$. Define $\mathcal{D}(\pi) = \{W \subset L^2(\pi) : \mathrm{codim} W = 1\}$. Then $\mathcal{D}(\tilde{\pi}) \subset \mathcal{D}(\pi)$. It is also easy to see that $\mathcal{D}(\pi) = \{\phi^{\perp_\pi} : \phi \in L^2(\pi)\}$. Hence,

$$\begin{aligned}
1 - \lambda_1(P) = \beta_1(L) &= \sup_{\phi \in L^2(\pi)} \inf_{\psi \in \phi^{\perp_\pi}} \frac{(\psi, L\psi)_\pi}{\|\psi\|_\pi^2} \\
&= \sup_{W \in \mathcal{D}(\pi)} \inf_{\psi \in W} \frac{(\psi, L\psi)_\pi}{\|\psi\|_\pi^2} \\
&\ge \frac{1}{A} \sup_{W \in \mathcal{D}(\pi)} \inf_{\psi \in W} \frac{(\psi, \tilde{L}\psi)_{\tilde{\pi}}}{\|\psi\|_\pi^2} \\
&\ge \frac{a}{A} \sup_{W \in \mathcal{D}(\pi)} \inf_{\psi \in W} \frac{(\psi, \tilde{L}\psi)_{\tilde{\pi}}}{\|\psi\|_{\tilde{\pi}}^2} \\
&\ge \frac{a}{A} \sup_{W \in \mathcal{D}(\tilde{\pi})} \inf_{\psi \in W} \frac{(\psi, \tilde{L}\psi)_{\tilde{\pi}}}{\|\psi\|_{\tilde{\pi}}^2} \\
&= \sup_{\phi \in L^2(\tilde{\pi})} \inf_{\psi \in \phi^{\perp_{\tilde{\pi}}}} \frac{(\psi, \tilde{L}\psi)_{\tilde{\pi}}}{\|\psi\|_{\tilde{\pi}}^2} \\
&= \frac{a}{A}\beta_1(\tilde{L}) = \frac{a}{A}(1 - \lambda_1(\tilde{P})),
\end{aligned}$$

which gives the required results. The first inequality follows from condition (i), the second from $a \cdot \int |\psi|^2 d\pi \le \int |\psi|^2 d\tilde{\pi}$, and the third from $\mathcal{D}(\tilde{\pi}) \subset \mathcal{D}(\pi)$.

Remarks. 1. In particular, if $\pi = \tilde{\pi}$, we can take $a = 1$ and the condition is reduced to comparing (f, Lf) and $(f, \tilde{L}f)$. In fact, Mira (1998) defines the covariance ordering which preserves the ordering of limiting variance of a function in $L_0^2(\pi)$.

2. If the state space is discrete and finite, we can compare the chains eigenvalue by eigenvalue. See Diaconis and Saloff-Coste (1993). Even in this general setting, we can have more refined comparison. However, for simplicity, we focus on β_1 which is directly related to the spectral radius.

3. It is usually not hard to manipulate condition (ii). However, this is not the case for condition (i). In what follows, we shall again apply some path arguments to get condition (i) for certain type of chains on $\mathbf{R}^n$.

Now we develop geometric constants that satisfy condition (i) in Theorem 3.1 for reversible Markov chains (discrete or continuous time) on $\mathbf{R}^n$ described in 1.1. We further assume that the transition kernels are of the form $P(x, dy) = \alpha(x)\delta_x(dy) + p_x(y)dy$ and $\tilde{P}(x, dy) = \tilde{\alpha}(x)\delta_x(dy) + \tilde{p}_x(y)dy$. Suppose the invariant distributions $\pi, \tilde{\pi}$ have density $q(y), \tilde{q}(y)$ w.r.t. Lebesgue measure respectively.

We also require the existence of a set of paths $\Gamma = \{\gamma_{xy}\}$ satisfying some regularity conditions related to both P, π and $\tilde{P}, \tilde{\pi}$. For each $x \neq y$ with $\tilde{q}(x)\tilde{p}_x(y) > 0$, there exists $b_{xy} \in \mathbf{N}$ and a path $\gamma_{xy} : \{0, \ldots, b_{xy}\} \to \mathbf{R}^n$ s.t. $\gamma_{xy}(0) = x$ and $\gamma_{xy}(b_{xy}) = y$ and that $p_u(v) > 0$ for any edge (u, v) of any path. It is slightly different from the *set of paths* defined in Section 2. We shall call it a *set of* $(P, \tilde{P})$ *paths.*

We also slightly modify the first regularity condition as follows: We replace $V = \{(x, y, i) : x, y \in S, 1 \leq i \leq b_{xy}\}$ by $V = \{(x, y, i) : \tilde{q}(x)\tilde{p}_x(y) > 0, 1 \leq i \leq b_{xy}\}$ and keep the rest unchanged. We call the new condition $(P, \tilde{P})$ *regularity condition.* Again, for paths 'nice' enough, this condition is easily satisfied.

The geometric constants we shall use are

$$A_\epsilon = A_\epsilon(\Gamma) = \mathrm{esssup}_{(u,v)\in E}\{Q(u, v)^{-(1-2\epsilon)} \sum_{\gamma_{xy} \ni (u,v)} \|\gamma_{xy}\|_\epsilon \tilde{q}(x)\tilde{p}_x(y)|J_{xy}(u, v)|\},$$

where the sum is taken over all (x, y) s.t. $\gamma_{xy} \ni (u, v)$.

Theorem 3.2. Let $\tilde{P}, \tilde{\pi}$ and P, π be reversible Markov chains on $\mathbf{R}^n$ and $\Gamma = \{\gamma_{xy}\}$ be a set of $(P, \tilde{P})$ paths satisfying the $(P, \tilde{P})$ regularity condition. Then for any $\epsilon \in \mathbf{R}$ and $f \in L^2(\pi)$,

$$(f, \tilde{L}f)_{\tilde{\pi}} \leq A_\epsilon(f, Lf)_\pi.$$

Proof. Note that

$$(f, \tilde{L}f)_{\tilde{\pi}} = \frac{1}{2}\iint |f(x) - f(y)|^2\tilde{q}(x)\tilde{p}_x(y)dxdy.$$

With the modifications of the regularity conditions, we can follow the exact same proof of Theorem 2.1 and get the required result.

Remarks. 1. To apply Theorem 3.1 and 3.2, we require a chain of reference: a chain $\tilde{P}, \tilde{\pi}$ with known $\lambda_1(\tilde{P})$. In the discrete case, we know how to solve the eigenvalues of some chains. In the general case, however, most chains are still unsolved.

2. As in Theorem 2.1, there is no definite procedure to find the *paths.* However, for special cases mentioned in Section 2, A_ϵ can be simplified considerably in a similar fashion.

4.Examples. Example 4.1 to 4.5 below can be considered as reversible discrete-time Markov chains or continuous-time Markov jump processes on S and transition probability kernel defined as in Section 2. They are all random walk chains, a particular case from the Metropolis-Hasting algorithm (see e.g. Tierney (1994)). The Markov chains we shall compare in Example 4.6 and 4.7 can be considered as chains described before Theorem 3.2.

Example 4.1. One dimensional case: $S = [-a, a]$.

(a) Let $p_x(y) = \frac{1}{2}$ (the uniform p.d.f.) on $[x-1, x+1] \cap S$ and 0 otherwise. As $p_x(y) = p_y(x)$, the chain is reversible w.r.t. the invariant distribution π with density $q(x) = \frac{1}{2a}\mathbf{1}_S(x)$, uniform over S. Since S is convex, we can choose the uniform paths or the linear paths and apply Corollary 2.4. For the uniform paths, we need to choose b big enough s.t. the expression inside the esssup is finite for all $(u,v) \in E_i$. So, we require whenever $(x,y) = (T_i)^{-1}(u,v)$, $Q(u,v) > 0$ for K to be finite and in addition $\|\eta^b_{xy}\|_\epsilon$ is finite for k_ϵ to be finite. For the best bounds, we take $b = \lceil 2a \rceil$. In this case, the distance between two adjacent points in a path is less than 1. So, $Q(u,v) = \frac{1}{4a} > 0$ and $\|\eta^b_{xy}\|_\epsilon = \sum_{(u,v)\in\eta^b_{xy}} Q(u,v)^{-2\epsilon} = b(4a)^{2\epsilon}$ which is finite. By Corollary 2.4(a),

$$k_\epsilon = b^1 \sum_{i=1}^{b} (\frac{1}{4a})^{-(1-2\epsilon)} b(4a)^{2\epsilon} (\frac{1}{2a})^2 = \frac{b^3}{a},$$

and

$$K = b^1 \sum_{i=1}^{b} (\frac{1}{4a})^{-1} (\frac{1}{2a})^2 = \frac{b^2}{a}.$$

In particular, if $2a$ is an integer, $b = 2a$ and so $k_\epsilon = 8a^2 = O(a^2)$ and $K = 4a = O(a)$. So, $1 - \frac{1}{k_\epsilon} = 1 - \frac{1}{8a^2}$ and $1 - \frac{1}{8K^2} = 1 - \frac{1}{2^7 a^2}$ are upper bounds of $\lambda_1(P)$ by Theorem 2.1 and 2.2 respectively. In this case, $1 - \frac{1}{k_\epsilon}$ turns out to be a better bound. However, $\frac{1}{k_\epsilon}$ and $\frac{1}{8K^2}$ differ only by a constant.

For the linear paths, we choose $c = 1$ for the same reason as for the uniform paths and use Corollary 2.4(b). We shall give an upper bound for K. The sum in the essential supremum $\sum_{\gamma_{xy}\ni(u,v)}$ is in fact a double sum: For $(u,v) \in E$ and $I \geq 1$, there are at most I pairs of x, y s.t. $\gamma_{xy} \ni (u,v)$ with $b_{xy} = I$. Moreover, $I \leq (1 - |v-u|)^{-1} \leq \lceil 2a \rceil$ for (u,v) to be an edge. So,

$$Q(u,v)^{-1} \sum_{\gamma_{xy}\ni(u,v)} q(x)q(y)b^n_{xy} \leq (\frac{1}{4a})^{-1} \sum_{I=1}^{\lceil 2a \rceil} \sum_{i=1}^{I} (\frac{1}{2a})^2 I = \frac{1}{6a} \lceil 2a \rceil (\lceil 2a \rceil + 1)(2\lceil 2a \rceil + 1).$$

In particular, if $2a$ is an integer, $K \leq \frac{1}{3}(2a+1)(4a+1) = O(a^2)$. Similarly, $k_\epsilon \leq O(a^3)$. Both bounds are much worse (different orders for K and k_ϵ) than the corresponding ones we get from the uniform paths. This illustrates different paths can give very different bounds.

If we consider $\tilde{P} = \frac{1}{2}(I + P)$ as the transitional kernel for a discrete time Markov chain, then $\rho \leq 1 - \frac{1}{16a^2}$ (with k_ϵ for the normal paths). Then by Proposition 1.1, for any $\mu \in L^2(\pi)$,

$$\|\mu\tilde{P}^n - \pi\|_2 \leq \|\mu - \pi\|_2(1 - \frac{1}{16a^2})^n.$$

We can also consider the continuous-time Markov jump process with mean 1 holding times with operator $\bar{P}^t = \sum_{n=0}^{\infty} e^{-t}\frac{t^n}{n!}P^n$. By Remark 2 after Proposition 1.1, we have

$$\|\mu\bar{P}^t - \pi\|_2 \leq \|\mu - \pi\|_2 e^{-t(1-\frac{1}{16a^2})}.$$

The following examples can all be treated in the same way.

(b) Let $p_x(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{(y-x)^2}{2}}$ (the normal p.d.f. of $n(x,1)$) for $y \in S$ and 0 otherwise. As $p_x(y) = p_y(x)$, the chain is reversible w.r.t. the invariant distribution with density $q(x) = \frac{1}{2a}\mathbf{1}_S(x)$, uniform over S. We consider the normal paths η_{xy}^b in Corollary 2.4 for $b \in \mathbf{N}$. For any $(u,v) \in E_i$ and $(x,y) = (T_i)^{-1}(u,v)$,

$$\|\eta_{xy}^b\|_\epsilon = \sum_{i=1}^{b}(\frac{1}{2a}\frac{1}{\sqrt{2\pi}}e^{-\frac{(v-u)^2}{2}})^{-2\epsilon} = b(\frac{1}{2a\sqrt{2\pi}}e^{-\frac{(v-u)^2}{2}})^{-2\epsilon}.$$

Note that $p_x(y)$ is a decreasing function of $|y-x|$. So,

$$k_\epsilon \leq b^1\sum_{i=1}^{b} b(\frac{1}{2a\sqrt{2\pi}}e^{-\frac{(\frac{2a}{b})^2}{2}})^{-1}(\frac{1}{2a})^2 = \frac{\sqrt{2\pi}b^3 e^{\frac{2a^2}{b^2}}}{2a}.$$

To get the best upper bound, we minimize the expression over b and we get $b = \lceil\frac{2a}{\sqrt{3}}\rceil$ or $\lfloor\frac{2a}{\sqrt{3}}\rfloor$. In particular, if $\frac{2a}{\sqrt{3}}$ is an integer,

$$k_\epsilon \leq \frac{4\sqrt{2\pi}a^2}{3\sqrt{3}}.$$

We can go through similar calculations for K to get

$$K \leq \frac{\sqrt{2\pi}b^2 e^{\frac{2a^2}{b^2}}}{2a}.$$

Minimizing over b and assuming that $\sqrt{2}a$ is an integer,

$$K \leq \frac{1}{e\sqrt{2\pi}a}.$$

In this example, the bound we obtained from k_ϵ is better than that from K.

Remarks. From the above calculations, it is easy to show that for random walk chains, i.e., $p_x(y) = p(y-x)$, k_ϵ is independent of ϵ with uniform or linear paths. For simplicity, we shall only compute k_0 in the following examples if this is the case.

Example 4.2. Two dimensional convex case: $S = B(0,a) \subset \mathbf{R}^2$.

(a) Let $p_{(u_1,u_2)}(v_1,v_2) = \frac{1}{\pi}$ (the uniform p.d.f.) on $B(x,1)\cap S$ and 0 otherwise. Then the chain is reversible w.r.t. the invariant distribution with density $q(x) = \frac{1}{a^2\pi}\mathbf{1}_S(x)$ As in Example 4.1(a), we use the uniform paths η^b_{xy} and take $b = \lceil 2a \rceil$. Since $\frac{2a}{b} \leq 1$, for any $(u,v) \in E_i$, $p_u(v) > 0$. With similar calculations,

$$k_\epsilon = k_0 = b^2 \sum_{i=1}^{b} (\frac{1}{\pi}\frac{1}{a^2\pi})^{-1} b(\frac{1}{a^2\pi})^2 = \frac{b^4}{a^2}.$$

Similarly, $K = \frac{b^3}{a^2}$. In particular, if $2a$ is an integer, $b = 2a$ and by Theorem 2.1 and 2.2, $\lambda_1(P) \leq \min\{1-\frac{1}{k_0}, 1-\frac{1}{8K^2}\} = 1-\frac{1}{16a^2}$. Again, k_ϵ gives a better bound. So, unless otherwise specified, we shall only calculate the most useful bound in the following examples.

(b) Let $p_{(u_1,u_2)}(v_1,v_2) = \frac{1}{2\pi}e^{-\frac{(v_1-u_1)^2+(v_2-u_2)^2}{2}}$ (the bivariate normal p.d.f. for two independent normal distributions $n(u_1,1)$,$n(u_2,1)$) for $(v_1,v_2) \in S$ and 0 otherwise. It is also clear that $p_{(u_1,u_2)}(v_1,v_2) = p_{(v_1,v_2)}(u_1,u_2)$. Then the invariant p.d.f. is $q(x) = \frac{1}{a^2\pi}\mathbf{1}_S(x)$, uniform over S. Similar to Example 4.1(b), we consider the uniform paths and let $b \in \mathbf{N}$. Note that $p_{(u_1,u_2)}(v_1,v_2)$ is constant on any circle on S with center (u_1,u_2) and the value is decreasing as the radius increases. So, the esssup that defines k_ϵ in Corollary 2.4(a) is reached when $\|(v_1,v_2)-(u_1,u_2)\| = \frac{D}{b} = \frac{2a}{b}$ for some $(u,v) \in E_i$. So,

$$k_\epsilon \leq \frac{2b^4}{a^2}e^{\frac{2a^2}{b^2}}.$$

Again, minimizing the R.H.S. gives $b = \lceil a \rceil$ or $\lfloor a \rfloor$. In particular, if a is an integer,

$$k_\epsilon \leq 2a^2e^2.$$

Example 4.3. Family of right-angled triangle

Let $S_{r,\theta}$ be the right-angled triangle with hypotenuse r and an angle θ with transition kernel as in 3.2(a). Then the invariant p.d.f. is $q_{r,\theta}(x) = \frac{4}{r^2\sin 2\theta}\mathbf{1}_{S_{r,\theta}}(x)$, uniform over $S_{r,\theta}$. Take the uniform paths with $b = b_{r,\theta} = \lceil r \rceil$. By Corollary 2.4 (a),

$$k_\epsilon(r,\theta) = k_0(r,\theta) = \frac{4\pi\lceil r \rceil^4}{r^2\sin 2\theta}.$$

So, $1-\frac{1}{k_\epsilon(r,\theta)}$, the upper bounds for $\lambda_1(P)$ in Theorem 2.1, go to 1 as $\theta\to 0$. Intuitively, as the angle becomes sharper, the convergence rate (for the continuous time Markov chain or the discrete time Markov chain with transition kernel $\frac{1}{2}(I+P)$) becomes slower.

Example 4.4. Two dimensional non-convex case.

(a) Suppose $S=\{(x,y)\in(-a,a)\times\mathbf{R}: a-|x|<y<2(a-|x|)\}$, an open inverted 'V' shape with vertices $(a,0),(0,2a),(-a,0),(0,a)$, which is a non-convex set. Let $p_{(u_1,u_2)}(v_1,v_2)=\frac{1}{\pi}$ (the uniform p.d.f.) on $B(x,1)\cap S$ and 0 otherwise. Then the invariant p.d.f. $q(x)=\frac{1}{a^2}\mathbf{1}_S(x)$. We shall follow the discussion in Section 2B and then apply Corollary 2.3. So, we need to construct a bijection ϕ from S (which is open already) to a convex set, in which we already have the uniform paths. In this example, we define $\phi: S\to S'$ s.t. $\phi(x,y)=(x,2(y-a+|x|))$ where $S'=\{(x,y)\in[-a,a]\times\mathbf{R}: 0<y<2(a-|x|)\}$, an open triangle with vertices $(a,0),(0,2a),(-a,0)$. So, $\phi^{-1}(x,y)=(x,a-|x|+\frac{y}{2})$. It is obvious that both ϕ,ϕ^{-1} are continuously differentiable a.e.. Recall that $\Phi: S\times S\to S'\times S'$ is defined by

$$\Phi((x_1,x_2),(y_1,y_2))=(\phi(x_1,x_2),\phi(y_1,y_2))=(x_1,2(x_2-a+|x_1|),y_1,2(y_2-a+|y_1|)).$$

Note that $(T_i)^{-1}$ defined in the discussion before Corollary 2.3 has continuous partial derivatives except on a set M of λ(the Lebesgue measure on $\mathbf{R}^2\times\mathbf{R}^2$) measure zero. To see this, let $Y=\{(0,y): 1\le y\le 2\}$, which is Lebesgue measure zero on $\mathbf{R}^2$). Then M is the union of $Y\times S, S\times Y, T_i(Y\times S), T_i(S\times Y)$, which is clearly of measure zero. It is easy to show that the Jacobian of Φ, $|J_\Phi((x_1,x_2),(y_1,y_2))|\equiv 4$. Since $|J_\Phi|$ is a constant, $|J_i(u,v)|=b^2$ as in Section 2B. To apply Corollary 2.3, we need to find b s.t. $Q(u,v)>0$ for any edge (u,v). Since $p_{(u_1,u_2)}(v_1,v_2)=0$ when $\|(v_1,v_2)-(u_1,u_2)\|>1$, we have to choose b big enough that the distance between any two such points is less than or equal to 1. According to the definition of ϕ, it is easy to see that if $b\ge 2\sqrt{2}a$, the horizontal distance between two adjacent points is less than $\frac{1}{\sqrt{2},}$ and so the actual distance is less than 1. Hence, by Corollary 2.3, taking $b=\lceil 2\sqrt{2}a\rceil$,

$$k_\epsilon\le\frac{\pi\lceil 2\sqrt{2}a\rceil^4}{a^2}=O(a^2).$$

(b) Suppose $S=\{(x,y)\in(-a,a)\times\mathbf{R}: \frac{1}{2}\sqrt{a^2-x^2}<y<\sqrt{a^2-x^2}\}$, an open crescent shape, which is non-convex. Let $p_{(u_1,u_2)}(v_1,v_2)=\frac{1}{\pi}$ (the uniform p.d.f.) on $B(x,1)\cap S$ and 0 otherwise. Then the invariant p.d.f. $q(x)=\frac{4}{a^2\pi}\mathbf{1}_S(x)$. Define bijection $\phi: S\to S'$ by $\phi(x,y)=(x,2y-\sqrt{a^2-x^2})$, where $S'=\{(x,y)\in(-a,a)\times\mathbf{R}: 0<y<\sqrt{a^2-x^2}\}$, an open semi-disc. So, $\phi(S)=S'$ is convex. It is also obvious that both ϕ,ϕ^{-1} are continuously differentiable. To apply Corollary 2.3, recall that $\Phi : S \times S \to S' \times S'$ is defined by

$$\Phi((x_1,x_2),(y_1,y_2)) = (\phi(x_1,x_2),\phi(y_1,y_2)) = (x_1, 2x_2 - \sqrt{a^2-x_1^2}, y_1, 2y_2 - \sqrt{a^2-y_1^2}).$$

It is easy to show that the Jacobian of Φ, $|J_\Phi((x_1,x_2),(y_1,y_2))| \equiv 4$, a constant. Similar to (a), we have to choose b s.t. $Q(u,v) > 0$ for any edge (u,v). According to the definition of ϕ, it is not hard to see that for any fixed b, the distance between any two adjacent points must be less than the distance $d(b)$ between $(a,0)$ and $(a-\frac{2a}{b}, \sqrt{a^2-(a-\frac{2a}{b})^2})$. So, if $b \geq 4a^2$, or $d(b) = \frac{4a^2}{b} \leq 1$, the distance between any adjacent points is less than 1. Hence, by Corollary 2.3, taking $b = \lceil 4a^2 \rceil$,

$$k_\epsilon \leq \frac{4\lceil 4a^2\rceil^4}{a^2} = O(a^6).$$

This value is much bigger than that in part (a). Intuitively, since the angles at the vertices are 0, it is relatively hard to 'escape' from the vertices. This accounts for the slower convergence rate. Similarly,

$$K = O(a^4).$$

It is interesting to note that $\frac{1}{k_\epsilon} = O(a^{-6})$ while $\frac{1}{8K^2} = O(a^{-8})$. This is example of using uniform paths and that k_ϵ, K give very different bounds.

(c) Suppose $A > a > 0$ and $S = \{(x,y) \in (-A,A) \times \mathbf{R} : \sqrt{a^2-x^2} < y < \sqrt{A^2-x^2}, y > 0\}$ is the open 'C' shape and the transition kernel is as in (b). Then the invariant p.d.f. $q(x) = \frac{2}{(A^2-a^2)\pi}\mathbf{1}_S(x)$. Instead of using the type of bijections in (a) and (b), we consider the natural polar bijection ϕ. Define bijection $\phi : S \to (a,A) \times (0,\pi)$ by $\phi(x,y) = (\sqrt{x^2+y^2}, \theta)$, where $\theta \in (0,\pi)$ is given $\tan\theta = \frac{y}{x}$. Hence, $\phi^{-1}(r,\theta) = (r\cos\theta, r\sin\theta)$. Similar to (a) and (b), we can apply Corollary 2.3. By direct calculations, $|\det D\phi^{-1}(r,\theta)| = r$, and so $|\det D\phi(x,y)| = \frac{1}{\sqrt{x^2+y^2}}$. Hence,

$$|J_\Phi((x_1,x_2),(y_1,y_2))| = \frac{1}{\sqrt{x_1^2+x_2^2}\sqrt{y_1^2+y_2^2}},$$

and so $\frac{1}{A^2} < |J_\Phi| < \frac{1}{a^2}$. From the discussion in Section 2B,

$$|J_i(u,v)| \leq \frac{\frac{1}{a^2}b^2}{\frac{1}{A^2}} = \frac{A^2b^2}{a^2}.$$

Now, we have to choose b s.t. $g(b) > 0$. For any fixed b, it is easy to see that the distance $d(b)$ between $(A,0)$ and $\frac{(b-1)A+a}{b}(\cos\frac{\pi}{b}, \sin\frac{\pi}{b})$ is the longest among the distances between adjacent points. Therefore, $Q(u,v) = \frac{2}{(A^2-a^2)\pi}\frac{1}{\pi}$ if $d(b) \leq 1$ for any edge (u,v). For fixed A and a, we can use numerical methods to find the smallest b s.t. $d(b) \leq 1$. For such b, say b_0, by Corollary 2.3,

$$k_\epsilon \leq \frac{2A^2b_0^4}{a^2(A^2-a^2)}.$$

For very large A and a, b_0 will also be large. So, the triangle with vertices $(A,0)$, $\frac{(b_0-1)A+a}{b_0}(\cos\frac{\pi}{b_0}, \sin\frac{\pi}{b_0})$ and $A(\cos\frac{\pi}{b_0}, \sin\frac{\pi}{b_0})$ is approximately right-angled at the last vertex and so

$$d(b_0) \approx \sqrt{(\frac{A-a}{b_0})^2 + 2A^2(1-\cos\frac{\pi}{b_0})} \approx \sqrt{(\frac{A-a}{b_0})^2 + (A\frac{\pi}{b_0})^2}.$$

Since b_0 is the smallest b s.t. $d(b) \leq 1$, we have

$$b_0 \approx \sqrt{(A-a)^2 + (A\pi)^2}.$$

In particular, if $A = 2a$,

$$k_\epsilon \leq \frac{2A^2b_0^4}{a^2(A^2-a^2)} \approx \frac{8}{3}a^2(1+4\pi^2)^2.$$

This value differs that in (a) by only a constant. Intuitively, S has no 'sharp' ends. In fact, all the angles at the vertices are $\frac{\pi}{2}$.

Example 4.5. n dimensional case: $S = [-a,a]^n$.

Let $p_{\mathbf{x}}(\mathbf{y}) = 2^{-n}$ (the uniform p.d.f.) on $C(\mathbf{x},1) \cap S$ and 0 otherwise, where $C(\mathbf{x},1) = \{\mathbf{y} \in \mathbf{R}^n : |y_i - x_i| \leq 1, i = 1,\ldots,n\}$ is the n-cube with width 2 and center $\mathbf{x}$. Then the chain is reversible w.r.t. the invariant distribution with density $q(\mathbf{x}) = (2a)^{-n}\mathbf{1}_S(\mathbf{x})$. By choosing $b = \lceil 2a \rceil$, it is easy to see that for any two adjacent points $\mathbf{u}$,$\mathbf{v}$ in a uniform path, we have $\mathbf{v} \in C(\mathbf{u},1)$. So, by Corollary 2.4(a),

$$k_\epsilon(n) = k_0(n) = \frac{b^{n+2}}{a^n} \approx 2^{n+2}a^2.$$

Example 4.6. Comparison of different one dimension chains with bounded state spaces.

(a) We consider two Markov chains on $S = [-a,a]$ of the type described in Section 3. $\tilde{p}_x(y) = \frac{1}{2}$ (the uniform p.d.f.) on $[x-1, x+1] \cap S$ and 0 otherwise. $p_x(y) = \frac{b}{2}$ (the uniform p.d.f.) on $[x-\frac{1}{b}, x+\frac{1}{b}] \cap S$ and 0 otherwise for some $b \geq 1$. For simplicity, let b be an integer. As $p_x(y) = p_y(x)$ and $\tilde{p}_x(y) = \tilde{p}_y(x)$, both chains are reversible w.r.t. the invariant distribution $\pi = \tilde{\pi}$ with density $q(x) = \tilde{q}(x) = \frac{1}{2a}\mathbf{1}_S(x)$, uniform over S. We consider the *uniform* paths η^b_{xy} as usual. Then for each $x \neq y$ with $\tilde{q}(x)\tilde{p}_x(y) > 0$ and for any edge (u,v) of η^b_{xy}, $|v-u| \leq \frac{b}{b} = 1$ and so $p_u(v) > 0$. Therefore, it is a set of $P, \tilde{P}$ paths. Similar to the uniform paths in Section 2A, it is easy to show that the $P, \tilde{P}$ regularity condition is also satisfied with $J_{xy}(u,v) = b$. We are now ready to apply Theorem 3.2. We have

$$A_0 = \text{esssup}_{(u,v)\in E}\{Q(u,v)^{-1} \sum_{\gamma_{xy}\ni(u,v)} b_{xy}\tilde{q}(x)\tilde{p}_x(y)|J_{xy}(u,v)|\} \leq (\frac{1}{2a}\frac{b}{2})^{-1}b(b\frac{1}{2a}\frac{1}{2}b) = b^2.$$

So, by Theorem 3.2,

$$(f, \tilde{L}f)_{\tilde{\pi}} \leq b^2(f, Lf)\pi$$

for any $f \in L^2(\pi)$. Since $\tilde{\pi} = \pi$, Theorem 3.1 implies that

$$\lambda_1(P) \leq 1 - \frac{1}{b^2}(1 - \lambda_1(\tilde{P})).$$

Note that from Example 4.1(a), $\lambda_1(\tilde{P}) \leq 1 - \frac{1}{8a^2}$. So,

$$\lambda_1(P) \leq 1 - \frac{1}{b^2}(\frac{1}{8a^2}) = 1 - \frac{1}{8(ab)^2}.$$

Observe that the chain P is equivalent to the chain in Example 4.1(a), replacing a by ab. So, the bound we get from there is also $1 - \frac{1}{8(ab)^2}$. More importantly, the comparison we got is a comparison of the *actual* values of $\lambda_1(P)$ and $\lambda_1(\tilde{P})$.

On the other hand, since if $p_x(y) > 0$, $\tilde{p}_x(y) = \frac{1}{b}p_x(y) > 0$ and $q = \tilde{q}$, we have

$$\begin{aligned}(f, Lf)_\pi &= \frac{1}{2}\iint |f(x) - f(y)|^2 q(x)p_x(y)dxdy \\ &\leq \frac{1}{2}\iint |f(x) - f(y)|^2 \tilde{q}(x)b\tilde{p}_x(y)dxdy \\ &= b(f, \tilde{L}f)_{\tilde{\pi}}\end{aligned}$$

By Theorem 3.1,

$$\lambda_1(\tilde{P}) \leq 1 - \frac{1}{b}(1 - \lambda_1(P)).$$

However, this is not a good comparison as we cannot recover the bound $1 - \frac{1}{8a^2}$ of $\lambda_1(\tilde{P})$ from the fact that $\lambda_1(P) \leq 1 - \frac{1}{8(ab)^2}$. Even so, we still manage to get both lower and upper bound for $\lambda_1(P)$ through $\lambda_1(\tilde{P})$:

$$1 - b(1 - \lambda_1(\tilde{P})) \leq \lambda_1(P) \leq 1 - \frac{1}{b^2}(1 - \lambda_1(\tilde{P})).$$

(b) We consider chains similar to (a). Redefine $\tilde{p}_x(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-x)^2}{2}}$ (the normal p.d.f. of $n(x,1)$) and $p_x(y) = \frac{b}{\sqrt{2\pi}} e^{-\frac{b^2(y-x)^2}{2}}$ (the normal p.d.f. of $n(x, \frac{1}{b^2})$) for $y \in S$ and 0 otherwise. Both chains are reversible w.r.t. the invariant distribution $\pi = \tilde{\pi}$ with density $q(x) = \tilde{q}(x) = \frac{1}{2a}\mathbf{1}_S(x)$, uniform over S by Example 4.1(b). Again, for simplicity, let b be an integer. As in (a), we consider the set of uniform paths η^b_{xy} which satisfies the regularity conditions. For any $(u,v) \in \eta^b_{xy}$,

$$\tilde{q}(x)\tilde{p}_x(y) = \frac{1}{2a}\frac{1}{\sqrt{2\pi}} e^{-\frac{(y-x)^2}{2}} = q(u)\frac{1}{\sqrt{2\pi}} e^{-\frac{b^2(v-u)^2}{2}} = \frac{1}{b} q(u) p_u(v) = \frac{1}{b} Q(u,v).$$

Hence,

$$A_0 \le \operatorname{esssup}_{(u,v)\in E}\{Q(u,v)^{-1} b(b\frac{1}{b}Q(u,v)b)\} = b^2.$$

By Theorem 3.2 and then 3.1,

$$\lambda_1(P) \le 1 - \frac{1}{b^2}(1 - \lambda_1(\tilde{P})).$$

In this example, we can try uniform paths with different number of steps. By choosing b steps, the constant A_0 is easy to calculate.

(c) We consider chains similar to (a), (b). Redefine $\tilde{p}_x(y) = \frac{1}{2}$ and $p_x(y) = 1 - |y - x|$ on $[x-1, x+1] \cap S$. The chains are also both reversible w.r.t. $\pi = \tilde{\pi}$ with density $q(x) = \tilde{q}(x) = \frac{1}{2a}\mathbf{1}_S(x)$. Apply the uniform paths η^b_{xy} and we get

$$A_0(b) \le \operatorname{esssup}_{(u,v)\in E}\{\frac{1}{1-|v-u|} b(\frac{1}{2}b)\} = \sup_{0 \le a \le \frac{1}{b}} \frac{b^2}{2(1-a)} = \frac{b^3}{2(b-1)}.$$

Minimizing over integer b, we have $b = 2$ and $A_0(2) = 4$. Hence,

$$\lambda_1(P) \le 1 - \frac{1}{4}(1 - \lambda_1(\tilde{P})).$$

Example 4.7. Comparison of different one dimension chains with unbounded state space.

Now consider two Markov chains on the real line, both with invariant distribution with standard normal p.d.f.: $q(x) = \tilde{q}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. The proposal distributions of the chains from a point x are uniform on $[x - \frac{1}{b}, x + \frac{1}{b}]$ and $[x-1, x+1]$ respectively. Formally, by the Metropolis algorithm (see e.g. Tierney (1994)), we can define the two Markov chains as in Section 3, where $\tilde{p}_x(y) = \frac{1}{2}\min\{\frac{q(y)}{q(x)}, 1\}$ on $[x-1, x+1]$ and 0 otherwise, and $p_x(y) = \frac{b}{2}\min\{\frac{q(y)}{q(x)}, 1\}$ on $[x - \frac{1}{b}, x + \frac{1}{b}]$ and 0 otherwise. Then both chains are reversible w.r.t. to the target density $q = \tilde{q}$. For simplicity, let b be an integer and apply the uniform paths η^b_{xy}. Simplifying the geometric constant, we have

$$A_0 = \text{esssup}_{(u,v)\in E} \frac{b \sum_{\eta^b_{xy} \ni (u,v)} \min\{e^{-\frac{x^2}{2}}, e^{-\frac{y^2}{2}}\}}{\min\{e^{-\frac{u^2}{2}}, e^{-\frac{v^2}{2}}\}}$$

Observe that for any $(u, v) \in \eta^b_{xy}$, $\min\{e^{-\frac{x^2}{2}}, e^{-\frac{y^2}{2}}\} \leq \min\{e^{-\frac{u^2}{2}}, e^{-\frac{v^2}{2}}\}$. So, $A_0 \leq b^2$. By Theorem 3.2 and then 3.1, we have

$$\lambda_1(P) \leq 1 - \frac{1}{b^2}(1 - \lambda_1(\tilde{P})).$$

This is our first example with unbounded state space. Even though we fail to bound the convergent rate of both chains with the geometric method from Section 2, we can still make a comparison with the geometric constants in Section 3.

REFERENCES

Y. Amit, On the rates of convergence of stochastic relaxation for Gaussian and non-Gaussian distributions, J. Multivar. Anal. 38 (1991) 89-99.

Y. Amit and U. Grenander, Comparing sweep strategies for stochastic relaxation, J. Multivar. Anal. 37(2), (1991) 197-222.

R. Bañuelos and T. Carroll, Brownian motion and the fundamental frequency of a drum, Duke Math. J. 75, (1994) 575-602.

J.R. Baxter and J.S. Rosenthal, Rates of convergence for everywhere-positive Markov chains, Statist. Probab. Lett. 22 (1995) 333-338.

E.D. Belsley, Rates of convergence of Markov chains related to association schemes, Ph.D. dissertation, Department of Mathematics, Harvard University (1993).

P. Billingsley, Probability and measure, 3^{nd} ed. (John Wiley & Sons, New York, 1995).

J.B. Conway, A course in functional analysis (Springer-Verlag, New York, 1985).

P. Diaconis, Group representations in probability and statistics, IMS Lecture Series 11, Institute of Mathematical Statistics, Hayward (California, 1988).

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

P. Diaconis and D. Stroock, Geometric bounds for eigenvalues of Markov chains, Ann. Appl. Probab. 1 (1991) 36-61.

A. Frieze, R. Kannan and N.G. Polson, Sampling from log-concave distributions, Ann. Appl. Probab. 4 (1994) 812-837.

A. Frigessi, C.-R. Hwang, S.J. Sheu and P. Di Stefano, Convergence rates of the Gibbs sampler, the metropolis algorithm, and other single-site updating dynamics, J. Royal Statist. Soc. Ser. B 55 (1993) 205-220.

A.E. Gelgand and A.F.M. Smith, Sampling based approaches to calculating marginal densities, J. Amer. Statist. Assoc. 85 (1990) 398-409.

C.-R. Hwang, S.-Y. Hwang-Ma and S.-J. Sheu, Accelerating Gaussian diffusions, Ann. Appl. Probab. 3 (1993) 897-913.

S. Ingrassia, On the rate of convergence of the metropolis algorithm and Gibbs sampler by geometric bounds, Ann. Appl. Probab. 2 (1994) 347-389.

M. Jerrum and A. Sinclair, Approximating the permanent, SIAM J. Comput. 18 (1988) 1149-1178.

E. Kreyszig, Introductory functional analysis with applications (John Wiley & Sons, Toronto, 1978).

G.F. Lawler and A.D. Sokal, Bounds on the L^2 spectrum for Markov chains and Markov processes: a generalization of Cheeger's inequality, Trans. Amer. Math. Soc. 309 (1988) 557-580.

S.P. Meyn and R.L. Tweedie, Computable bounds for convergence rates of Markov chains, Ann. Appl. Prob. 4 (1994) 981-1011.

A. Mira, Ordering, slicing and splitting Monte Carlo Markov chains, Ph.D. dissertation, School of Statistics, University of Minnesota (1998).

J. Quastel, Diffusion of colour in the simple exclusion process, Comm. Pure Appl. Math. 45 (1992) 623-679.

M. Reed and B. Simon, Methods of mathematical physics IV (Academic Press 1978).

G.O. Roberts and J.S. Rosenthal, Shift-coupling and convergence rates of ergodic averages, Communications in Statistics – Stochastic Models 13, No. 1 (1997a) 147-165.

G.O. Roberts and J.S. Rosenthal, Geometric ergodicity and hybrid Markov chains, Electronic Communications in Probability 2 (1997b) 13-25.

J.S. Rosenthal, Rates of convergence for Gibbs sampler for variance components models, Ann. Statist. 23 (1995a) 740-761.

J.S. Rosenthal, Minorization conditions and convergence rates for Markov chain Monte Carlo, J. Amer. Statist. Assoc. 90 (1995b) 558-566.

J.S. Rosenthal, Analysis of Gibbs samples for a model related to James-Stein estimators, Statist. Comput. 6 (1996a), 269-275.

J.S. Rosenthal, Markov chain convergence: From finite to infinite, Stochastic processes and their Applications 62 (1996b) 55-72.

W. Rudin, Functional Analysis, 2^{nd} ed. (McGraw-Hill, New York, 1991).

A. Sinclair, Improved bounds for mixing rates of Markov chains and multicomodity flow, Combinatories, Probab. Comput. 1 (1992) 351-370.

A. Sinclair and M. Jerrum, Approximate counting, uniform generation and rapidly mixing Markov chains, Inform. Comput. 82 (1989) 93-133.

A.F.M. Smith and G.O. Roberts, Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods (with discussion), J. Roy. Statist. Soc. Ser. B 55 (1993) 3-24.

L. Tierney, Markov chains for exploring posterior distributions, Ann. Statist. 22 (1994) 1701-1762.