Convergence of Markov Chains via Analytic and Isoperimetric Inequalities

Dissertation zur Erlangung des Doktorgrades
an der
Fakultät für Mathematik
der
Universität Bielefeld

vorgelegt von Jörg Kienitz

Bielefeld, Oktober 2000

Contents

List of Figures

Chapter 1

Introduction

Markov chains have become very important tools in many fields of application. One very remarkable application originated in statistical physics and is now widely used for simulation. It is called Markov Chain Monte Carlo or MCMC for short.
The idea to use Markov chains for simulation is as follows. Unlike Monte Carlo methods one does not sample directly from a probability distribution π. One appropriately constructs a Markov chain having π as its equilibrium distribution. Appropriate in this context means it must not take too much steps to get close to π. Instead of drawing samples directly from the probability measure π one can draw samples from the distribution of the Markov chain at some time. This concept has proven useful for high dimensional state spaces where the time necessary to draw from a distribution close to π can be determined, e.g. in the Ising model. This method has also been applied to approximate the permanent of a matrix or to estimate the volume of convex polyhedra. For further results in this direction, see [S 93].
With regard to the question of the right time for drawing a sample we have to specify what we mean by 'the distribution of the Markov chain being close to π'. Then, we can turn to the main part of this thesis, namely estimating the 'rate of convergence' or the 'mixing rate' for various Markov processes including Markov chains on general state spaces.
A Markov chain is, in this dissertation, a collection of random variables $\mathbf{X} = (X_n)_{n\in\mathbb{N}}$ on some probability space $(\Omega, \mathcal{F}, P)$ taking values in a measure space $(S, \mathcal{S}, \mu)$. Without loss of generality we can take $\Omega = S, \mathcal{F} = \mathcal{S}, P = \mu$. To be a Markov process $\mathbf{X}$ has to fulfill the *Markov property*. This means for all $n \in \mathbb{N}$ and for all $A \in \mathcal{F}$ we have

$$P(X_{n+1} \in A | X_n, X_{n-1}, ..., X_0) = P(X_{n+1} \in A | X_n). \tag{1.1}$$

For finite, countable or general state spaces one can describe a homogenous Markov chain by considering its transition kernel K which is a mapping

$$K : S \times \mathcal{S} \to [0,1], \tag{1.2}$$

such that $K(x,\cdot)$ is a probability measure on $(S,\mathcal{S})$ and for all $A \in \mathcal{S}$, $K(\cdot,A) : S \to [0,1]$ is a measurable mapping. For $n \in \mathbb{N}$, the $n-$th step transition K^n is defined inductively by $K^n(x,dy) = \int_S K(z,dy)K^{n-1}(x,dz)$. We work with the continuous semigroup $H_t := \exp(-t(I-K))$, $t > 0$, associated with K. The corresponding process is a continuous time Markov jump process $(Y_t)_t$. The moves of $(Y_t)_t$ are those of the discrete time Markov chain $(X_n)_n$. The difference is that after one move $(Y_t)_t$ stays at the actual position according to independent exponential waiting times. Thus, the probability to be at a certain point $y \in S$ after exactly i jumps up to time t is $\exp(-t)t^i K^i(x,y)/i!$. The probability for the process $(Y_t)_t$ started at $x \in S$ to be in $A \in \mathcal{S}$, i.e. $P_x(Y_t \in A)$, is given by $H_t(x,A)$. We often write H_t^x instead of $H_t(x,\cdot)$.
For finite or countable state spaces we can identify the elements of S with $\{0,1,...,|S|\}$ or $\mathbb{N}_0$ respectively. We can identify K with a matrix with entries being the transition probabilities for a move from i to j.
For finite state spaces the well known Perron-Frobenius theorem from Linear Algebra for irreducible and aperiodic chains assures the existence of a unique vector π such that for all i

$$\lim_{k\to\infty} K^k(i,j) = \pi(j) = \lim_{t\to\infty} H_t(i,j). \tag{1.3}$$

Irreducibility is a kind of connectedness condition, i.e. for all $x,y \in S$ there exists $n \in \mathbb{N}$ such that $K^n(x,y) > 0$ whereas aperiodicity excludes periodic behaviour. The condition we need is that the greatest common divisor of the set $\{n : n \geq 0, K^n(i,i) > 0\}$ is 1.
The measure π is called the stationary distribution and the property (1.3) is often called *(ordinary) ergodicity*.
For general and countable state spaces the existence of a unique stationary distribution π is not so easy to establish but it exists under mild conditions. For an introduction to this subject we refer to the book [MT 96].
A standard way for quantifying the rate of convergence is to estimate the signed measures $H_t^x - \pi$ or $K^n(x,\cdot) - \pi$ respectively, using the well known norms of total variation, $\|\cdot\|_{TV}$, and the χ^2-norm, $\|\cdot\|_{2,\pi}$, i.e.

$$\|H_t^x - \pi\|_{TV} := \sup_{|f|\leq 1} \left| \int_S f(y)(H_t^x - \pi)(dy) \right| = \sup_{A\subset S} |(H_t^x - \pi)(A)|, \tag{1.4}$$

$$\|H_t^x - \pi\|_{2,\pi} := \left(\int_S (H_t^x(y) - 1)^2 \, \pi(dy) \right)^{1/2}. \tag{1.5}$$

Applying the Cauchy-Schwarz inequalitiy the norms are related by

$$2\|H_t^x - \pi\|_{TV} \leq \|H_t^x - \pi\|_{2,\pi}. \tag{1.6}$$

In Chapter 2 we study the ergodic behaviour of (Sub-) Markov processes on a Polish space $(S, \mathcal{S})$, i.e. a separable metric space, equipped with a countably generated σ-field $\mathcal{S}$ and a σ-finite non-negative measure μ having transition kernels of the form

$$K(x, dy) = k(x, y)\mu(dy) \text{ and } \int_S K(x, dy) \leq 1. \tag{1.7}$$

Furthermore, we assume that there exists a stationary distribution π having a density $h > 0$ with respect to μ.
Unlike the case of finite state spaces there are not many tools available for bounding the rate of convergence for (Sub-) Markov chains on general state spaces.
In the sequel we study L^2-*exponential ergodicity*. The semigroup $(H_t)_t$ is L^2-exponential ergodic if there exists a real constant $C > 0$ such that

$$\|H_t^x - \pi\|_{2,\pi} \leq \|\delta_x - \pi\|_{2,\pi} \exp(-Ct). \tag{1.8}$$

In order to understand the role of certain analytic inequalities in showing L^2-exponential ergodicity we briefly explain the connection of analytic and probabilistic notions.
A real valued function $f : S \to \mathbb{R}$ is in the domain $D(\mathcal{E})$ of the quadratic form

$$\mathcal{E}(f, f) := \frac{1}{2} \int_S \int_S (f(x) - f(y))^2 K(x, dy)\mu(dx). \tag{1.9}$$

if $\mathcal{E}(f, f) < \infty$. We call $(\mathcal{E}, D(\mathcal{E}))$ the Dirichlet form associated with the kernel K and denote by $Var_\mu[f]$ the *variance* if it exists.
As a constant in (1.8) we can take the *spectral gap*, i.e.

$$\lambda_1 := \inf \left\{ \frac{\mathcal{E}(f, f)}{Var_\mu[f]} : f \in D(\mathcal{E}), Var_\mu[f] \neq 0 \right\}. \tag{1.10}$$

(Sub-) Markov processes on general state spaces not necessarily have a spectral gap. Therefore, our results described in the sequel can also be seen as existence results for a spectral gap.
For this reason we are interested in lower bounds on the spectral gap. A lower bound on λ_1 exists if for a constant C and for all $f \in D(\mathcal{E})$ we have the inequality

$$Var_\mu[f] \leq C\mathcal{E}(f, f). \tag{1.11}$$

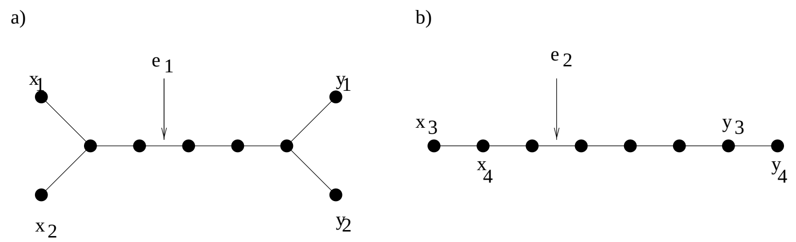

Figure 1.1: **Four paths in $\mathbb{R}^2$.**

Since such inequalities for various variances and L^2-norms of derivatives had been studied at the end of the 19th century by H. Poincaré we say in case (1.11) holds, $\mathcal{E}$ satisfies a *Poincaré inequality*. Combining (1.10) and (1.11) yields a lower bound on λ_1.
To state our results we introduce discrete paths. To this end for each tupel $(x, y) \in S \times S$ consider $N(x, y) \in \mathbb{N}$ and introduce map γ which maps a triple (x, y, n), $n \leq N(x, y)$, to some point $\gamma_{xy}(n) \in S$. We call the map $\gamma_{xy} : \{0, 1, ..., N(x, y)\} \to S$ a *discrete path*. The only restriction here is that $\gamma_{xy}(0) = x$ and $\gamma_{xy}(N(x, y)) = y$. The number $N(x, y)$ is the *length* of the path.
We say a set of paths $\Gamma = \{\gamma_{xy}\}$ is admissible if γ_{xy} is injective and adapted to the transition kernel $K(x, dy) = k(x, y)\mu(dy)$. By adapted we mean that $k(\gamma_{xy}(n-1), \gamma_{xy}(n)) > 0$. A point $e \in S \times S$ is called an *edge* of *index* n if $e = (\gamma_{xy}(n-1), \gamma_{xy}(n))$ for some path γ_{xy}.
To state our lower bounds on λ_1 we have to single out some 'nice' path sets. To this end we introduce the *path reconstruction property*, abbreviated p.r.p. Roughly speaking, a path γ is 'nice' if one is able to determine the starting point and the ending point whenever the length of the path γ_{xy} and the index of one edge e are known. Consider Figure 1.1. It shows four paths γ_{x_i,y_i}, $i = 1, 2, 3, 4$. The length $L(\gamma_{x_i,y_i})$ of each path is 6. In Figure 1.1 a) we cannot reconstruct neither the starting point nor the ending point if we know the index of e_1 which is 3. It can either be x_1 or x_2 and y_1 or y_2 respectively. In Figure 1.1 b) we are able to recover the starting points x_3, x_4 and the ending points y_3 and y_4. This example suggests that if an edge is contained in two different paths it must be of different index in each path or if the index is the same the length of the paths must not be the same.
For a better understanding of the p.r.p. and how it comes into play we briefly review the connection of a graph on the one hand and an irreducible Markov chain on a finite state space $\mathcal{X}$ with transition kernel K on the other. A set $\mathcal{A} \subset \mathcal{X} \times \mathcal{X}$ is said to be an adapted edge set to K if $\mathcal{A}$ is symmetric, connected and whenever $(x, y) \in \mathcal{A}$ one has $K(x, y) + K(y, x) > 0$. If we

define $Q(x, y) = K(x, y)\pi(x)$ then Q is a probability measure on the edge set $\mathcal{A}$. Furthermore, the boundary ∂A of a set $A \subset \mathcal{X}$ is given by

$$\partial A := \{(x, y) \in \mathcal{A} : x \in A, y \in A^c \text{ or } x \in A^c, y \in A\}. \tag{1.12}$$

Using these notions we can describe the interplay between graphs and Markov chains.
Define a weighted unidirectional graph G with weight function $Q : G \times G \to [0, 1]$. And denote by $d_x = \sum_y Q(x, y)$ the degree of a vertex. Then, the discrete Laplacian $\mathcal{L}$ of the graph G is given by

$$\mathcal{L}(x, y) := \begin{cases} 1 - \frac{Q(x,y)}{d_x} & \text{if } \ x = y, \\ -\frac{Q(x,y)}{\sqrt{d_x d_y}} & \text{if } \ x \neq y. \end{cases} \tag{1.13}$$

It reflects the 'geometry' of the graph, i.e. the adjacency relation.
This interplay can be used to establish path based upper bounds on constants appearing in analytic inequalities like Poincaré inequalities. Upper bounds obtained using this interplay have been studied for the finite setting by many authors, see for instance [DS 91], [JS 88], [JS 89], [SC 97] or [S 92].
For general state spaces we have to deal with integrals and not with finite sums. Moreover, the main difficulty is that there is no graph structure. But our path sets can be used to build up an artificial graph structure.
To build up the artificial graph structure for a given path set $\Gamma = \{\gamma_{xy}\}$ we take the image of the set

$$E_1 := \{(x, y, n) : x, y \in S, 1 \leq n \leq N(x, y)\}$$

under the map G which maps a point (x, y, n) to an *indexed edge*, i.e. $(\gamma_{xy}(n-1), \gamma_{xy}(n), n, N(x, y))$. We call $G(E_1) =: E_2$ the indexed edge set.
If $d\delta$ denotes the counting measure on $\mathbb{N}$, let ν_1 and ν_2 be the induced measures on E_1 and E_2 of $\mu \otimes \mu \otimes d\delta$ and $\mu \otimes \mu \otimes d\delta \otimes d\delta$.
Our assumption is that the map $G : E_1 \to E_2$ is bijective, which means the paths are 'nice' in the sense described earlier, and $G(\nu_1)$ has a density h_1 with respect to ν_2. Now, an admissible path set Γ having this property is said to obey the p.r.p. Under these circumstances the integration formula for image measures is

$$\int_{E_1} f \, d\nu_1 = \int_{E_2} f \circ G^{-1} h_1 \, d\nu_2. \tag{1.14}$$

Our approach using discrete paths at first glance seems like an ad hoc method but in many examples we can find path sets obeying a p.r.p., for instance

convex sets. This fact can be used to establish an extended p.r.p. for sets having a finite cover with convex sets.
We deduce a Poincaré inequality with constant κ^Γ which depends on the path set Γ. Also, a *local Poincaré inequality* can be shown. To this end let ρ denote the metric on S and for a real $r > 0$ let

$$B(x,r) := \{y : \rho(x,y) \leq r\}$$

be the ball with radius r centered at $x \in S$. Let Γ be a path set obeying the p.r.p. and denote by $f_r := \frac{1}{\mathrm{Vol}(B(x,r))} \int_S 1_{B(x,r)}(y) f(y) \mu(dy)$. we deduce inequalities of the form

$$\|f - f_r\|_2^2 \leq \eta^\Gamma \mathcal{E}(f,f). \tag{1.15}$$

Here η^Γ is a constant depending on the given path set Γ. Such inequalities are called local Poincaré inequalities, because f_r only depends on the neighbourhood $B(x,r)$ of x.
Meanwhile, similar bounds on constants appearing in Poincaré inequalities have been independently obtained by Yuen, [Y 00]. His approach is different and his proofs differ from ours. Also, his regularity conditions are stricter.
Other lower bounds for λ_1 are induced by the concept of *isoperimetry*.
Isoperimetric inequalities originated in Riemannian geometry. To explain this concept, let (M,g) be a C^∞-Riemannian manifold. Roughly speaking, these are inequalities between the volume measure μ and the surface measure σ of a relatively compact set $A \subset M$ with smooth boundary, abbreviated by r.c.s.b. set in the sequel.
The following inequalities translate to the finite setting which has been intensively studied by Diaconis and Stroock, [DS 91], Saloff-Coste, [SC 97] or Houdré, [H 00], but they also apply to our general setting of Chapter 2.
Let $L(M)$ denote the set of locally integrable real-valued functions on M and let H be a subset of $L(M) \times L(M)$. We consider functionals $\mathcal{S}$ on $C_0^\infty(M)$ of the form

$$\mathcal{S}(f) = \sup_{(V,W)\in H} \left\{ \int_M f^+ V + \int_M f^- W \right\}, \tag{1.16}$$

where $f^\pm = \pm f \vee 0$. The well known co-area formula from differential geometry is a key ingredient for the proof of Theorem 1 in [R 85] which gives the clue for applications of isoperimetric concepts for analyzing ergodic behaviour of Sub-Markov processes as it points out the connection to analytic inequalities. Namely, it reduces the study of inequalities of the form $c\mathcal{S}[f] \leq \int_M |\nabla f| \, d\mu$, for all $f \in C_0^\infty(M)$, to the study of characteristic functions, i.e. to inequalities $c \max\{\mathcal{S}[1_A], \mathcal{S}[-1_A]\} \leq \sigma(\partial A)$ for all r.c.s.b. $A \subset S$. With the gradient

∇ replaced by a discrete analogue the theorem holds true in our setting. Due to [R 85] functionals which can be considered are

$$\mathcal{L}_\mu(f) = \int f^2 \ln\left(\frac{f^2}{\|f\|_2^2}\,d\mu\right) = \sup_{V:\int\exp(V)\le 1} \int fV\,d\mu, \tag{1.17}$$

$\|\cdot\|_p$-norms and if $\mathrm{Vol}(M) < \infty$, $1 \le p \le \infty$, such that $1/p + 1/q{=}1$,

$$\mathcal{S}_1(f) = \inf_{a\in\mathbb{R}} \left(\int |f-a|^p\,d\mu\right)^{\frac{1}{p}} = \sup_{\substack{V:\int V=0\\ \|V\|_q\le 1}} \int fV\,d\mu. \tag{1.18}$$

For $A \subset S$, we have $\mathcal{E}(1_A, 1_A) = Q(\partial A)$, $\|1_A\|_p = \pi(A)$ and $Var_\pi[1_A] = \pi(A)\pi(A^c)$. To study Poincaré inequalities Rothaus' theorem suggests to study the *isoperimetric constants*

$$I_K := \inf_{\substack{A\subset S\\ \pi(A)\le 1/2}} \frac{Q(\partial A)}{\pi(A)} \text{ and } \tilde{I}_K := \inf_{A\subset S} \frac{Q(\partial A)}{(\pi(A^c))\pi(A)}. \tag{1.19}$$

The infimum in (1.19) is only taken over all measurable subsets $A \subset S$. For the case of Markov chains on general or finite state spaces several people proved *Cheeger inequalities*, see Lawler and Sokal, [LS 88], Diaconis and Stroock, [DS 91], Sinclair, [S 92] or Saloff-Coste, [SC 97]. In 1970 Cheeger proved an isoperimetric inequality involving the spectral gap of the Laplacian on Riemannian manifolds of the form (1.20) below, see [C 70].
The version for general state spaces, obtained by Lawler and Sokal, relate the constants I_K and $\tilde{I}_K$ to each other and to the spectral gap by

$$\frac{\tilde{I}_K^2}{8} \le \frac{I_K^2}{8} \le \lambda_1 \le I_K \le \tilde{I}_K. \tag{1.20}$$

We can deduce a lower bound on I_K in Theorem 2.5.2 using a constant B^Γ depending on the given path set Γ. Cheeger's inequality, (1.20) can be applied to provide lower bounds on the spectral gap in terms of B^Γ.
Here we point out that Yuen in [Y 00] needs somewhat stricter regularity conditions on the state space.
The connection of analytic and probabilistic notions can be used to compare the rate of convergence of two different (Sub-) Markov chains. In applications a given (Sub-) Markov chain can be very difficult to analyze. But sometimes it is possible to compare analytic notions with those of a known chain which already has been analyzed or is easier to analyze.
To this end consider an auxilliary Sub- Markov chain $\tilde{K}$. The general idea of comparison is to establish a relation between the Dirichlet $\mathcal{E}$ and $\tilde{\mathcal{E}}$ associated

with the (Sub-) Markov chains corresponding to K and $\tilde{K}$. We recall that admissibility depends on the transition kernel of a chain. Since we study two different kernels now, we have to change the admissibility condition to be applicable for the two different kernels at once. We call this property *c-admissibility*. In the sequel we use a constant A^{Γ} depending on the given c-admissible path set to relate the Dirichlet forms $\mathcal{E}$ and $\tilde{\mathcal{E}}$ associated with the two different chains. We will derive an inequality of the form:

$$\mathcal{E}(f,f) \leq A^{\Gamma}\tilde{\mathcal{E}}(f,f). \tag{1.21}$$

With regard to the variational characterization of the spectral gap, (1.10) and that of the *logarithmic-Sobolev constant* such inequalities can be used to compare the corresponding notions for the different chains.
Here the *logarithmic-Sobolev constant* corresponding to a chain K can be introduced as

$$\alpha := \inf\left\{\frac{\mathcal{E}(f,f)}{\mathcal{L}_{\mu}[f]} : f \in D(\mathcal{E}), 0 < \mathcal{L}_{\mu}[f] < +\infty\right\}. \tag{1.22}$$

The Dirichlet form $(\mathcal{E}, D(\mathcal{E}))$ satisfies a *logarithmic-Sobolev inequality* if there exists a constant $C > 0$ such that for all $f \in D(\mathcal{E})$ and $\mathcal{L}_{\mu}[f] < \infty$ we have

$$\mathcal{L}_{\mu}[f] \leq C\mathcal{E}(f,f). \tag{1.23}$$

These inequalities have been introduced in 1976 by Gross, see [G 76]. Some standard methods in analysis are dimension dependent and therefore not suitable to study high dimensional or even infinite dimensional problems. To overcome this difficulty Gross studied logarithmic-Sobolev inequalities.
The existence of a logarithmic-Sobolev constant is equivalent to a property of the semigroup $(H_t)_t$, called *hypercontractivity*. For numbers $p, q \in [0, \infty]$, $p \leq q$ let $\|\cdot\|_{p\to q}$ denote the usual operator norm from L^p to L^q. The common use of the notion of hypercontractivity states that for $t > 0$ we have $\|H_t\|_{2\to q} \leq 1$. Here q may depend on t.
For further reference see e.g. [G 76], [B 92],[G 93] or [St 93].
To illustrate all the bounds described above we provide several examples in Section 2.7. The first example is a $d-$dim cube. We study the effect of increasing the dimension on our lower bounds on the spectral gap. Also, we give examples for comparing two chains. For the next example, a $d-$dim cross, we can use the extended p.r.p. to analyze the transition kernel. Furthermore, we study $d-$dim balls and the Gibbs sampler, which is a standard method for setting up a Markov chain for simulation. In this case we obtain bounds improving the results in the literature by a factor 10^7. Various submanifolds of $\mathbb{R}^d$ like a circle or a torus can be used to check the performance

of our methods since we can compute the spectrum using Fourier analysis. It turns out that our results can be reasonably good.
Finally, lower bounds on the logarithmic-Sobolev constant α are derived. We consider the connection between logarithmic-Sobolev inequalities and isoperimetry to some extent. We will prove that the logarithmic-Sobolev constant α can be related to the *isoperimetric constant* I_K^{ln} given by

$$I_K^{ln} := \min_{\substack{A \subset S \\ \pi(A) \leq 1/2}} -\frac{Q(\partial A)}{\pi(A)\ln(\pi(A))}. \tag{1.24}$$

Again, the infimum is only taken over measurable subsets.
In fact in Theorem 2.8.2 we prove an analogue of Cheeger's inequality for the spectral gap. We call the inequality

$$\frac{\ln(2)}{4\sqrt{2}}(I_K^{ln})^2 \leq \alpha \leq I_K^{ln}, \tag{1.25}$$

logarithmic-Cheeger inequality. As a by-product of the proof we find lower bounds on α in terms of λ_1 and I_K^{ln}. With some more effort these bounds hold also true for non-reversible (Sub-) Markov chains. Here, λ_1 denotes the minimum of the real part of all eigenvalues. These bounds are monotone in λ_1. Hence, each lower bound on λ_1 implies a lower bound on α. In the reversible case we can therefore apply our path based methods to estimate α.

From Chapter 3 onwards we restrict ourselves to irreducible Markov chains on a finite state space $\mathcal{X}$. In Section 3.1 we summarize and review the fundamental notions.
Section 3.2 reviews coupling of Markov chains and the discrete martingale problem for $L := I - K$. This can be used to obtain new lower bounds on the logarithmic-Sobolev constant α of the chain. New ideas like coupling have been developed by Lindvall in [L 92]. These ideas and stopping time methods, see [AD 87], provide elegant and efficient results to previously untractable problems. Especially, Chen and Meise proved sharp lower bounds on the spectral gap using probabilistic ideas. Chen published numerous papers on this subject for the manifold setting as well as for Markov chains, we only mention [ChW 94], [Ch 94] and [Ch 98]. The methods developed by Meise who considered finite Markov chains can be found in [Me 96].
If there exists a *logarithmic-Sobolev function* f, i.e. a non-constant minimizer for the quotient $\frac{\mathcal{E}(f,f)}{\mathcal{L}_\pi[f]}$, it fulfills the difference equation

$$[Lf](x) = \alpha f(x)\ln(f(x)^2). \tag{1.26}$$

Regard the process $(Z_n)_n$ given by

$$Z_n := f(X_n) + 2\alpha \sum_{k=0}^{n-1} f(X_k)\ln(f(X_k)). \tag{1.27}$$

Since L is the generator of the process in (1.27), $(Z_n)_n$ is a martingale. We use the optional stopping theorem and Lebesgue's convergence theorem in Theorems 3.4.1 and 3.4.2 to show if $f^* = \max_x f(x)$, and T, T_y denote the coupling time and hitting time,

$$\alpha \geq \frac{1}{2}\frac{1}{\sup_{x,y}\mathbb{E}_{x,y}[T]}\frac{f^*-1}{f^*\ln(f^*)} \text{ and } \alpha \geq \frac{1}{2}\frac{1}{\sup_{x,y}\mathbb{E}_x[T_y]}\frac{1}{1+\ln(f^*)}. \tag{1.28}$$

Furthermore, if there exists a function $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ such that some coupling $\tilde{L}$ for the generator L satisfies $[\tilde{L}d](\cdot,\cdot) \geq c_d d(\cdot,\cdot)$ in Theorem 3.4.3 we derive:

$$\alpha \geq \frac{c_d}{1+\ln(f^*)}. \tag{1.29}$$

It is interesting that probabilistic methods can be used to study non linear problems like (1.26). All our bounds contain the maximum of a logarithmic-Sobolev function. Upper bounds on f^* in the continuous setting can be obtained using *logarithmic-Harnack inequalities*, which are not available for general weighted graphs.
To study the effectiveness of our lower bounds on α we compute them for some standard examples. Furthermore, we compare them to the exact values if they are known or to the estimates obtained in [DS-C 96a].
The last section of this chapter covers some examples to illustrate the lower and upper bounds using the logarithmic-Cheeger inequality, (1.25), and some related bounds.

The objective of Chapter 4 is to estimate higher eigenvalues for irreducible, aperiodic, reversible Markov chains on a finite state space $\mathcal{X}$. Estimating higher eigenvalues has a long history in spectral geometry. It can be traced back to the work of Weyl, see [W 11], at the beginning of the 20th century. To get an idea of the problems under consideration, consider a two dimensional drum which spans a domain D in $\mathbb{R}^2$. Let the vibration of the drum skin be governed by the equation

$$\frac{\partial^2 v}{\partial t^2} = \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} =: \Delta v. \tag{1.30}$$

Here $v = v(x, y, t)$ denotes the transverse displacement of a point (x, y) of the drum skin at time t. It is customary to separate the variables in (1.30) by setting $v = F(t)u(x,y)$. Where, for some real μ, u is a solution of

$$\Delta u + \mu u = 0 \text{ in } D. \tag{1.31}$$

Since the boundary of a drum is firmly attached, the solutions are equal to 0 on the boundary of D. In general other boundary conditions can be considered. It is a classical result that there are countably many functions $u_0, u_1, ..., u_k, ..,$ called eigenfunctions and numbers $0 = \mu_0 < \mu_1 \leq \mu_2 \leq ... \leq \mu_k \leq ...,$ called eigenvalues which satisfy

$$\Delta u_k + \mu_k u_k = 0 \text{ in } D. \tag{1.32}$$

For all k, the eigenfunctions can be taken to be orthonormal. It is well known that μ_k tends to ∞ if k does. In 1911, H. Weyl proved the following asymptotic result. For arbitrary domains $D \subset \mathbb{R}^n$, $n \geq 2$,

$$\mu_k \approx 4\pi^2 \left(\frac{k}{\omega_n \mathrm{Vol}(D)} \right)^{2/n}, \quad k \to +\infty, \tag{1.33}$$

where $\omega_n = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}$ denotes the volume of the unit ball in $\mathbb{R}^n$.
In 1961 Pólya, in [P61], showed for a class of domains that the k-th eigenvalue μ_k is bigger than the expression on the right hand side of (1.33) and conjectured that this holds true for arbitrary domains $D \subset \mathbb{R}^n, n \geq 2$.
To state our lower bounds for the higher eigenvalues denote by $0 = \lambda_0 < \lambda_1 \leq ... \leq \lambda_{|\mathcal{X}|-1}$ the eigenvalues of L. We suppose the associated discrete Dirichlet form $\mathcal{E}$ satisfies some *Nash inequality*. The discrete Dirichlet form $(\mathcal{E}, D(\mathcal{E}))$ satisfies a *Nash inequality of type I* if there exist positive constants C and d such that for all real valued functions f

$$Var_\mu[f]^{1+2/d} \leq C\mathcal{E}(f,f)\|f\|_1^{4/d}. \tag{1.34}$$

It satisfies a *Nash inequality of type II*, if there exist positive constants C, d and T such that for all real valued functions f

$$\|f\|_2^{2(1+2/d)} \leq C \left(\mathcal{E}(f,f) + \frac{1}{T}\|f\|_2^2 \right) \|f\|_1^{4/d}. \tag{1.35}$$

The Nash inequalities were introduced by Nash in [N 58]. Intuitively, they capture smoothing effects of the process. As the time t increases the distribution H_t more and more smoothly spread out according to the stationary distribution. (1.34) is stronger than the Poincaré inequality, because it implies a lower bound on the spectral gap. In case (1.35) we can only consider

the time interval $(0, T]$.
A Nash inequality of type I implies for all $t > 0$ the estimate

$$\|H_t - \pi\|_{1\to\infty} \leq \left(\frac{dC}{4t}\right)^{d/4}. \tag{1.36}$$

In contrast to the Nash inequality of type I, (1.35) reflects the short time behaviour of the (Sub-) Markov process. Nash inequalities of type II do not imply a lower bound on the spectral gap. In fact (1.35) implies for all $t \leq T$

$$\|H_t\|_{1\to\infty} \leq e\left(\frac{dC}{4t}\right)^{d/2}. \tag{1.37}$$

The decay of the semigroup $(H_t)_t$ described by the inequalities (1.36) and (1.37) has become known as *ultracontractivity*. Our main result, Theorem 4.1.1, states that in case a Nash inequality of type I holds true, we have

$$\lambda_k \geq \frac{k^{d/2}}{eC}, \quad k \in \{0, ..., |\mathcal{X}| - 1\}. \tag{1.38}$$

In the case a Nash inequality of type II holds, we have

$$\text{for } k \geq \left(\frac{dC}{2T}\right)^{d/2} e - 1: \quad \lambda_k \geq \frac{2}{dC}\frac{(k+1)^{2/d}}{e^{4/d}} \tag{1.39}$$

and

$$\text{for } k \geq \left(\frac{dCe}{2T}\right)^{d/2} e - 1: \quad \lambda_k \geq \frac{(k+1)^{2/d}}{e^{1+2/d}C}. \tag{1.40}$$

The proof uses inequalities derived in Proposition 4.1.3. They are of independent interest. If TrH_t denotes the trace of the heat kernel, we have

$$\lambda_k \geq \sup_{t>0} \frac{1}{t} \ln\left(\frac{k+1}{TrH_t}\right) \text{ and } \lambda_k \geq \sup_{t>0} \frac{1}{t} \ln\left(\frac{k}{TrH_t - 1}\right). \tag{1.41}$$

They appear in the context of finite Markov chains for the first time.
(1.38), (1.39) and (1.40) are analogues to the famous Pólya conjecture for the eigenvalues $(\mu_k)_k$ of the Laplacian for arbitrary domains $M \subset \mathbb{R}^n$. There are a few results in this direction in the literature. Let us focus on the results of Saloff-Coste, [SC 97], and Chung and Yau, [ChY95]. In [SC 97], Saloff-Coste establishes a relation between TrH_t and the eigenvalue counting function

$$N(s) := \#\{i : \{0, ..., |\mathcal{X}| - 1\}, \lambda_i \leq s\}. \tag{1.42}$$

Whenever a Nash inequality of type I or II holds $N(s)$ can be bounded by a continuous function $M(s)$. Then, he derives lower bounds on $(\lambda_k)_k$ in terms of the function $M(s)$.
In [ChY95], Chung's and Yau's main tool for estimating higher eigenvalues from below are the d-dimensional isoperimetric constant of a graph, i.e.

$$I_d := \min_{\substack{A\subset\mathcal{X}\\ \pi(A)\leq\frac{1}{2}}} \frac{\sum_{x\in A, y\in A^c} K(x,y)\pi(x)}{\pi(A)^{\frac{d-1}{d}}} = \min_{\substack{A\subset\mathcal{X}\\ \pi(A)\leq\frac{1}{2}}} \frac{Q(\partial A)}{\pi(A)^{\frac{d-1}{d}}}, \tag{1.43}$$

and discrete analogues to *Sobolev inequalities*. The discrete Dirichlet form $(\mathcal{E}, D(\mathcal{E}))$ satisfies a *Sobolev inequality of type I* if there are constants $C, d > 0$ such that for all real valued functions f

$$\|f\|^2_{2d/(d-1)} = \left(\int_S f^{2d/(d-1)} d\mu\right)^{(d-1)/d} \leq C\mathcal{E}(f,f). \tag{1.44}$$

and a *Sobolev inequality of type II* if there are constants $C, d, T > 0$ such that for all real valued functions f

$$\|f\|^2_{2d/(d-2)} \leq C\left(\mathcal{E}(f,f) + \frac{1}{T}\|f\|_2^2\right). \tag{1.45}$$

These inequalities are realted to Nash inequalities. The implication Sobolev $\rightarrow$ Nash can directly be seen using Hölder's inequality. To see the equivalence in the case $d > 2$ there have been indirect approaches using the heat kernel, see [CKS87] or [V 84], direct approaches using capacities, see [M 85], and finally direct approaches using only the inequalities itself, see [BCLSC 95]. In general one only can state this equivalence for a suitable domain of test functions. In the finite setting, considered here, a real function is in any L^p-space, because it only takes finitely many real values. Given a Nash inequality it is necessary to adjust the constant C in the corresponding Sobolev inequality to $4^{6+2d/(d-2)}C$.
Our bound on the higher eigenvalues involving I_d, Theorem 4.2.4, is

$$\lambda_k \geq \frac{I_d^2}{8e} k^{2/d}, \qquad \text{for } k \in \{0, ..., |\mathcal{X}| - 1\}. \tag{1.46}$$

Since the bounds of Chung and Yau involve I_d^4, $I_d \leq 1$, we can decrease the order by 2, which improves the bounds.
Before we consider several examples to illustrate the performance of our lower bounds on the higher eigenvalues in Section 4.4, we modify a result by Diaconis and Saloff-Coste. In [DS-C 94] and [DS-C 96b] they deduce methods

to prove Nash inequalities of type II. We will show that these methods can also be used to prove Nash inequalities of type I.
Finally, we provide several examples at the end of this chapter. There are some standard examples but we have included new ones like discrete tori and Markov chains used to study time dependent simulated annealing algorithms. Our bounds are of the right order of magnitude for the higher eigenvalues.

ACKNOWLEDGEMENTS

First of all I would like to mention the Stochastics group at Bielefeld. My supervisor Prof. Götze is acknowledged for the opportunity to take part in the DFG Schwerpunktprogramm *Interagierende stochastische Systeme hoher Komplexität* and the possibility to do part of my research at Bristol, UK.
Dr. Peter Eichelsbacher and Dr. Franz Merkl must be mentioned for fruitfull discussions and the first one for reading an early manuscript of the paper [K 00a] and a version of this dissertation.
As mentioned, part of the work was done from January to July 2000 at the University of Bristol. I would like to thank Prof. M. van den Berg for his kind invitation and the School of Mathematics of the University of Bristol for its hospitality.
Furthermore, I would like to thank the Creans. They not only provided accomodation but really made me a home at Bristol. Special thanks to Dick Crean who carefully read the introduction of this thesis.
Support by the DFG and a DAAD grant is grateful acknowledged.
Many thanks to my parents Sigrid and Paul for their support during school and my years at university.
Also, I thank my grandfather Gustav who taught me that education is so important. I never had the possibility to tell him that.
I would like to mention all my friends and, of course, my wife Beatrice. She encourages me all day and makes life worth living.

Chapter 2

Convergence rates for Sub-Markov chains on general state spaces

The purpose of this chapter is to develop methods to estimate the rate of convergence for reversible, homogenuous discrete-time Sub-Markov chains and time-continuous Sub-Markov jump processes $(X_t)_t$, $t \in \mathbb{N}$ or $t \in \mathbb{R}^+$ respectively, on a Polish state space $(S, \mathcal{S}, \mu)$ with countably generated σ-field.
Let K be a kernel, i.e. a map $K : S \times \mathcal{S} \to [0, 1]$ such that $K(x, \cdot)$ is a sub-probability measure on $(S, \mathcal{S})$ and for all $A \subset S$, $K(\cdot, A)$ is measurable. We assume K is of the form

$$K(x, dy) = k(x, y)\mu(dy), \quad \text{such that} \int K(x, dy) \leq 1. \tag{2.1}$$

Furthermore, for $x \in S$, $A \in \mathcal{S}$ and $n \in \mathbb{N}$ define K^n inductively by

$$K^n(x, dy) = \int_S K^{n-1}(x, dz)K(z, dy).$$

We say a collection of random variables $(X_n)_{n\in\mathbb{N}}$ is a Sub-Markov chain with transition kernel K if $P(X_n \in A : X_k, k < n; X_0 = x) = K^n(x, A)$. A probability measure π on $(S, \mathcal{S})$ is called invariant or stationary measure with respect to K if it has the property

$$\pi(A) = \int_S K^n(x, A)\pi(dy).$$

Conditions for uniqueness and existence of a stationary measure can be found in [MT 96].

Remark 2.0.1 *We can introduce a Markovian transition kernel $\overline{K}$ corresponding to K by*

$$\overline{K}(x,dy) = a(x)\delta_x(dy) + k(x,y)\mu(dy). \tag{2.2}$$

Here δ_x denotes the Dirac measure with point-mass 1 in $x \in S$ and $a : S \to \mathbb{R}$ is a numerical function which ensures that $\overline{K}(x,dy)$ is a probability transition kernel, i.e.

$$a(x) = 1 - \int k(x,y)\,\mu(dy).$$

We finally assume that a *stationary measure* π of the Sub-Markov chain exists and that it has a density $h > 0$ with respect to the measure μ, i.e. for $A \in \mathcal{S}, \pi(A) = \int 1_A(x)h(x)\,d\mu(x)$.

Example

For a positive real number $r \geq 1$ take $I := [-r,r]$ and $d\mu(y) = dy$, the Lebesgue measure on I. Consider

$$k(x,y) := \begin{cases} \frac{1}{2}, & x,y \in I \cap I_x := \{y \| x-y| \leq 1\} \\ 0, & \text{otherwise} \end{cases}.$$

The kernel $\overline{K}$ is given by

$$\overline{K}(x,dy) = \left(1 - \int_{-r}^{r} 1_{I_x}(y)\,dy\right)\delta_x(dy) + k(x,y)dy.$$

According to estimate the rate of convergence, i.e. the time necessary for K^n being close to π, we establish techniques to obtain bounds on constants appearing in analytic inequalities considered in the introduction.
In analogy to the finite case, see [DS 91], [S 92], [JS 88], [JS 89] or [SC 97], we introduce path sets adapted to the Sub-Markov chains under consideration. Let a function N be given by

$$\begin{aligned} N : S \times S &\to \mathbb{N}, \\ (x,y) &\mapsto N(x,y) \end{aligned} \tag{2.3}$$

and consider the map γ,

$$\begin{aligned} \gamma : \{(x,y,n) : x,y \in S, n \leq N(x,y)\} &\to S, \\ (x,y,n) &\mapsto \gamma_{xy}(n). \end{aligned} \tag{2.4}$$

For fixed $x, y \in S$ a *discrete path* which connects x and y is the map

$$\begin{aligned} \gamma_{xy} : \{0, 1, ..., N(x,y)\} &\rightarrow S, \\ n &\mapsto \gamma_{xy}(n), \end{aligned} \tag{2.5}$$

such that $\gamma_{xy}(0) = x$ and $\gamma_{xy}(N(x,y)) = y$.

Definition 2.0.2 *A path γ_{xy} connecting x and y is called admissible with respect to $(X_t)_t$ if the map γ_{xy} is injective and adapted, i.e.*

$$k(\gamma_{xy}(n-1), \gamma_{xy}(n)) > 0 \text{ for } n = 1, ..., N(x,y). \tag{2.6}$$

A path set $\Gamma = \{\gamma_{xy}\}$ is called admissible if each path γ_{xy} is admissible.

Furthermore, we define the *adapted edge set*.

Definition 2.0.3 *The adapted edge set $\mathcal{A}^\Gamma$ corresponding to $(X_t)_t$ is*

$$\begin{aligned} &\mathcal{A}^\Gamma := \\ &\{(v,w) : v, w \in S, \text{ it exists } \gamma_{xy} \in \Gamma, \text{ s.t. } v = \gamma_{xy}(n-1), \, w = \gamma_{xy}(n)\}. \end{aligned} \tag{2.7}$$

Given an edge $e = (v,w) \in \mathcal{A}^\Gamma$, such that $v = \gamma_{xy}(n-1)$ and $w = \gamma_{xy}(n)$ for some $\gamma_{xy} \in \Gamma$ we say e is of index n.
The boundary ∂A of a set $A \subset S$ is the set

$$\partial A := \{e = (v,w) \in S \times S : v \in A, \quad w \in A^c \text{ or vice versa}\}. \tag{2.8}$$

To simplify the notation, for an edge $e = (v,w)$ let us introduce

$$Q(v,w) = k(v,w)h(v). \tag{2.9}$$

Definition 2.0.4 *Let $\Gamma = \{\gamma_{xy}\}$ be an admissible path set and $p \in \mathbb{R}^+$. For $\gamma_{xy} \in \Gamma$ we set*

$$|||\gamma_{xy}|||_p := \sum_{n=1}^{N(x,y)} Q(\gamma_{xy}(n-1), \gamma_{xy}(n))^{-p}. \tag{2.10}$$

Remark 2.0.5 *1) $|||\cdot|||_p$ is well defined, since every edge $e = (v,w)$ of an admissible path satisfies $Q(e) := Q(v,w) > 0$.*
2) If we take $p = 0$ we have $|||\gamma_{xy}|||_0 = \sum_{(v,w)\in\gamma_{xy}} 1 = N(x,y)$. Thus, $|||\cdot|||_0$ is nothing but the length of a path.

The assumption on the state space S is that one is able to find a set of paths $\Gamma = \{\gamma_{xy}\}$ obeying a, as we call it, *path reconstruction property*. Once having introduced this property we are able to state our lower bounds on the spectral gap λ_1, (1.10), and the logarithmic-Sobolev constant α, (1.22).
The first section provides a precise definition of the path reconstruction property, abbreviated p.r.p. in the sequel. One can think of the p.r.p. in the following geometric terms:
Let $e = (v, w) \in \mathcal{A}$ be an arbitrary edge. Whenever the index of e and the length of a path containing e are known one is able to determine the starting-point and the ending-point of the path.
It should be pointed out that such path sets not necessarily exist. We will put forward a general framework in Section 2.2 to handle not only convex sets, for which it turns out that such path sets exist, but also sets having a finite convex covering.

2.1 Dirichlet forms

Firstly, to derive our lower bounds on the spectral gap and logarithmic-Sobolev constant we have to invoke a general framework. To this end we introduce Dirichlet forms associated with our processes. We start by introducing some linear operators.

The Gradient

For $f \in \mathcal{B}(S)$ define the linear operator $[\nabla\cdot]$ by

$$[\nabla\cdot] : \mathcal{B}(S) \to \mathcal{B}(S) \times \mathcal{B}(S) \tag{2.11}$$

$$f \mapsto [\nabla f], \tag{2.12}$$

with $[\nabla f](x, y) = f(x) - f(y)$. We call ∇ the discrete gradient.
Take the set $\tilde{\mathcal{B}} := \{\tilde{f} \in \mathcal{B}(S) \times \mathcal{B}(S) : \tilde{f}(x, y) = -\tilde{f}(y, x)\}$ and define

$$[\nabla^*\cdot] : \tilde{\mathcal{B}} \to \mathcal{B} \tag{2.13}$$

$$f \mapsto \int_S \tilde{f}(\cdot, y) K(\cdot, dy). \tag{2.14}$$

We call ∇^* the divergence.
For the generator L of the Sub-Markov chain, respectively jump process, we have in the sense of equality of operators

$$L = \nabla^*\nabla. \tag{2.15}$$

Now, we are in a position to define the corresponding Dirichlet form. For $f, g \in \mathcal{B}(S)$ and $\langle \nabla f, \nabla g \rangle_K := \int_S [\nabla f][\nabla g] K(x, dy)$ take the closure of

$$\begin{aligned}
(2.16) \quad \mathcal{E}(f,g) &:= \int_S \langle \nabla f, \nabla g \rangle_K \, d\pi(x) \\
&= \int_S \int_S (f(x) - f(y))(g(x) - g(y)) K(x, dy) d\pi(x).
\end{aligned}$$

on $L^2(S, \mu)$ with domain $D(\mathcal{E})$, i.e. the closure of $\mathcal{B}(S)$ in $L^2(S, \mu)$ with respect to the norm $\mathcal{E}(\cdot, \cdot) = \| \cdot \|_2$, and denote it also by $\mathcal{E}$.

Remark 2.1.1 *We have*

$$\begin{aligned}
\mathcal{E}(f,g) &= \int_S [Lf](x) g(x) \, \pi(dx) \\
&= \int_S [\nabla^* \nabla f](x) g(x) \, \pi(dx).
\end{aligned}$$

The first equality is a straightforward computation and to see the second we apply (2.15).

Definition 2.1.2 *Define the heat kernel corresponding to the Dirichlet form $\mathcal{E}$ of (2.16) by*

$$H_t(x, dy) := \exp(-t) \sum_{j=0}^{\infty} \frac{t^j K^j(x, dy)}{j!}.$$

The heat kernel is the transition kernel corresponding to the time-continuous process with exponential holding with respect to K.

We are now interested in measuring $\|K^n - \pi\|_{TV}$ or $\|K^n - \pi\|_{2,\pi}$ and $\|H_t - \pi\|_{TV}$ or $\|H_t - \pi\|_{2,\pi}$ respectively. If it is clear with respect to which measure we take the χ^2-norm we simply write $\| \cdot \|_2$.
The following proposition holds true.

Proposition 2.1.3 *If $\mathcal{E}_K$ and $\mathcal{E}_{\overline{K}}$ denote the Dirichlet forms corresponding to the kernels K and $\overline{K}$ in (2.1) and (2.2), we have for $f \in D(\mathcal{E}_K) \cap D(\mathcal{E}_{\overline{K}})$*

$$\mathcal{E}_K(f, f) = \mathcal{E}_{\overline{K}}(f, f) \tag{2.17}$$

and $\overline{K}(x, dy)$ is reversible iff $K(x, dy)$ is.

Proof:
A straightforward computation yields

$$\begin{aligned}
\mathcal{E}_K(f,f) &= \int_S\int_S (f(y)-f(x))^2 K(x,dy)\pi(dx) \\
&= \underbrace{\int_S\int_S (f(y)-f(x))^2 \alpha(x)\delta_x(dy)\pi(dx)}_{=0} \\
&\quad + \int_S\int_S (f(y)-f(x))^2 \overline{K}(x,y)\mu(dy)\pi(dx) \\
&= \int_S\int_S (f(y)-f(x))^2 \overline{K}(x,y)\mu(dy)\pi(dx) = \mathcal{E}_{\overline{K}}(f,f).
\end{aligned}$$

This proves Proposition 2.1.3. $\square$

Remark 2.1.4 *We will prove several inequalities involving the Dirichlet form $\mathcal{E}_K$ associated with K in the next few sections. By Proposition 2.1.3 and $D(\mathcal{E}_{\overline{K}}) \subset D(\mathcal{E}_K)$ all inequalities hold true for $\mathcal{E}_{\overline{K}}$.*

2.2 Path Reconstruction Properties

Suppose we are given an admissible collection of paths $\Gamma = \{\gamma_{xy}\}$ in the sense of Definition 2.0.2. Let $d(\delta n)$ denote the counting measure on $\mathbb{N}$. For the set $S \times S \times \mathbb{N}$ equipped with the measure $\tilde{\nu}_1$ given by $\tilde{\nu}_1 = \mu \otimes \mu \otimes d(\delta n)$ let $E_1 \subset (S \times S \times \mathbb{N})$ denote the set

$$E_1 := \{(x,y,n) : x,y \in S, 1 \leq n \leq N(x,y)\}. \tag{2.18}$$

Furthermore, denote by $\nu_1 := \tilde{\nu}_1|_{E_1}$ the restriction of $\tilde{\nu}_1$ to E_1. Let us now introduce a map G which maps each point of E_1 to an indexed edge. To this end, take a path $\gamma_{xy} \in \Gamma$ connecting x and y and define G by

$$G : E_1 \rightarrow S \times S \times \mathbb{N} \times \mathbb{N} \tag{2.19}$$

$$(x,y,n) \mapsto \left(\underbrace{\gamma_{xy}(n-1), \gamma_{xy}(n)}_{\in \mathcal{A}^\Gamma}, \underbrace{n}_{index}, \underbrace{N(x,y)}_{length} \right). \tag{2.20}$$

We call $G(E_1) =: E_2$ the indexed edge set and introduce a measure ν_2 on E_2 as the restriction of $\tilde{\nu}_2 = \mu \otimes \mu \otimes d(\delta n) \otimes d(\delta N)$ to E_2.
The following property is crucial for our further considerations

(D) Suppose the map G is a measurable injection. Thus, $G: E_1 \to E_2$ is a bijection. Suppose E_2 is a measurable subset of $S \times S \times \mathbb{N} \times \mathbb{N}$. Furthermore, suppose there exists a density h_1 of the image measure $G(\nu_1)$ with respect to ν_2 and that G^{-1} is measurable.

Definition 2.2.1 *(Path-Reconstruction-Property, p.r.p.)*
An admissible collection of paths $\Gamma = \{\gamma_{xy}\}$ obeys the path reconstruction property if it has property (D).

Let us explain how to apply the p.r.p. in the sequel. For an integrable function $f: E_1 \to \mathbb{R}$ consider the integral

$$\int_S \int_S \sum_{n=1}^{N(x,y)} f(x,y,n)\, d\mu(x) d\mu(y).$$

Using the measure space E_1 we can write this integral as

$$\int_{E_1} f\, d\mu d\mu d\delta = \int_{E_1} f\, d\nu_1.$$

Now, we can use the p.r.p. and obtain

$$\int_{E_1} f\, d\nu_1 = \int_{E_2} f \circ G^{-1}\, dG(\nu_1) = \int_{E_2} f \circ G^{-1} h_1\, d\nu_2.$$

Using Fubini's theorem the latter integral can be written as

$$\int_{E_2} f \circ G^{-1} h_1\, d\mu d\mu d\delta d\delta = \int_S \int_S \sum_{\mathbb{N}\times\mathbb{N}} f \circ G^{-1} 1_{E_2} h_1\, d\mu d\mu.$$

Because the density h_1 is only definied on E_2 we have to multiply the integrand with 1_{E_2}.

Examples

Let S be a convex subset of $\mathbb{R}^d$, $1 \leq d \in \mathbb{N}$. Define admissible path sets:
<u>1:</u> Fix $N \in \mathbb{N}$ and take $N(x,y) = N$. Define

$$\Gamma^1 = \left\{\gamma^1_{xy}\right\}, \gamma^1_{x,y}(n) = \frac{(N-n)x + ny}{N}, n \in \{0,..,N\}. \tag{2.21}$$

<u>2:</u> For $\eta \in \mathbb{R}^+$ take $N(x,y) = \lceil \eta d(x,y) \rceil$ and define

$$\Gamma^2 = \left\{\gamma^2_{xy}\right\}, \gamma^2_{x,y}(n) = \frac{(N(x,y)-n)x + ny}{N(x,y)}, n \in \{0,...,N(x,y)\}. \tag{2.22}$$

Remark 2.2.2 *1) The parameter η appearing above is necessary in applications to ensure that consecutive points in a discrete path v and w in γ^2_{xy} remain in the domain of the kernel K. Hence, η ensures that the paths are admissible.*
2) We will sometimes abbreviate $N(x, y)$ by N^{xy}.

In the first example the paths are chosen to be the discrete straight line segments connecting two points $x, y \in S$. A path contains N edges. Let e be an edge of index k. To recover the starting-point we have to choose the left vertex of the edge appearing before $k-1$ steps. To recover the ending-point we take the right vertex of the edge appearing after $N-k$ steps.
Formally, given an edge $e = (v, w)$ of index k we can solve the following equations (2.23) and (2.24) to recover x and y from v and w:

$$v = \frac{(N-k+1)x + (k-1)y}{N} \tag{2.23}$$

$$w = \frac{(N-k)x + ky}{N}. \tag{2.24}$$

We can go through similar computations and the same reasoning holds true for the second example. Here, we have not chosen the equidistant partition but one which depends on the distance between x and y.
In both examples for fixed $l, L \in \mathbb{N}$ the map $G(\cdot, \cdot, l, L)$ is a continuous linear map. The densities h_1 are the determinants of the corresponding Jacobian of the linear transformation. For fixed l, L for the first example L can only be the fixed length N and therefore $h_1 = N^d$.
To compute the density h_1 of the second example we have to introduce projections.
For $i = 1, 2$ let

$$\begin{aligned} pr_i : E_1 &\to S, \\ (x_1, x_2, n) &\mapsto x_i \end{aligned}$$

denote the projection on the i-th coordinate.
The density is $h_1 = N(pr_1(G^{-1}(v, w, l, L)), pr_2(G^{-1}(v, w, l, L))^d$. In contrast to the first example h_1 explicitely depends on l and L.
But nevertheless it can be bounded by the length of the longest path, i.e. $\sup_{x,y\in S}\lceil \eta \|x-y\|_{eukl.}\rceil$. We will use this fact in Section 2.7 when we present several examples.

Extended Path Reconstruction Property

There are simple sets which do not obey the p.r.p. for simple path systems. For instance star- or barbell-shaped sets are not convex and therefore not any two points can be joined by a straight line. In the sequel we want to extentd the p.r.p. to cover a large class of sets including the latter examples. To this end we consider a set S fulfilling

1. S has a finite covering $S = \bigcup_{i=1}^{N} S_i$ such that each set S_i, $i \in \{1, 2, ..., N\}$ obeys the p.r.p. in the sense of Definition 2.2.1.

2. For each $i \in \{1, 2, ..., N\}$ there exists $j \in \{1, 2, ..., N\}$, $i \neq j$, such that $S_i \cap S_j \neq \emptyset$.

Let us consider adapted edge sets to the process $(X_t)_t$ on each S_i. For an admissible path set Γ^i we take its adapted edge set $\mathcal{A}_i$. Because of condition 2. from above, we obtain an adapted edge set $\mathcal{A}^\Gamma$ if we set $\mathcal{A}^\Gamma = \bigcup_{i=1}^{N} \mathcal{A}_i$. The path set obeys an extended p.r.p. in the following sense:

Definition 2.2.3 *(Extended path reconstruction property)*
Let S satisfy conditions 1. and 2. from above. An admissible collection of paths $\Gamma = \{\gamma_{xy}\}$ obeys the extended path reconstruction property if for all γ_{xy} there exists a decomposition $\gamma_{xy} = \gamma_1 + ... + \gamma_N$ such that γ_i is a path of a path set Γ^i obeying the p.r.p.

We denote by

$$E_1^i := \{(s_{i-1}, s_i, n) : s_{i-1}, s_i \in S_i, 1 \leq n \leq N(s_{i-1}, s_i)\} \tag{2.25}$$

and

$$\begin{aligned} G_i : E_1^i &\rightarrow S_i \times S_i \times \mathbb{N} \times \mathbb{N} && (2.26)\\ (v, w, n) &\mapsto (\gamma_{xy}(n-1), \gamma_{xy}(n), n, N(s_{i-1}, s_i)) && (2.27) \end{aligned}$$

the corresponding sets and maps to those in Subsection 2.2. The p.r.p. on S_i guarantees the existence of a density h_1^i and with $E_2^i = G(E_1^i)$ the change of measure formula becomes

$$\int_{E_1^i} f \, d\nu_1^i = \int_{E_2^i} f \circ G_i^{-1} h_1^i \, d\nu_2^i. \tag{2.28}$$

Let us illustrate the definition of the extended p.r.p. To this end we consider a set $S \subset \mathbb{R}^d$, $d \geq 1$, such that $S = \bigcup_i^N S_i$, S_i convex and fulfilling condition 2. from above.
For all $i \in \{1, 2, ..., N\}$ the set S_i is convex and therefore it obeys the p.r.p. in the sense of Definition 2.2.1. For an admissible collection of paths and the corresponding adapted edge set we have $\bigcup_{i=1}^{N} \mathcal{A}_i^{\Gamma^i} \subset \mathcal{A}^\Gamma$. We take for

$s_0 = x, s_1 \in S_1, ..., s_{d-1} \in S_d, s_d = y$, denoting $A(l) := \sum_{k=1}^{l} N_k$ and consider the path set

$$\Gamma^3 := \{\gamma^3_{xy}\} \text{ with} \tag{2.29}$$

$$\gamma^3_{xy} = \begin{cases} \frac{(A(1)-n)x+ns_1}{N_1} & , \quad 0 \le n \le A(1), \\ \frac{(A(2)-n)s_1+(n-A(1))s_2}{N_2} & , \quad A(1) \le n \le A(2), \\ \dots & \\ \frac{(A(d)-n)s_{d-1}+(n-(A(d-1))y}{N_d} & , \quad A(d-1) \le n \le A(d). \end{cases}$$

Due to Example 1 each path component, as it is a discrete straight line segment, obeys the p.r.p. in the sense of Definition 2.2.1. That means we are able to compute the density h_1 for each path component constituting the discrete polygonal path.

For another example take $s_0 = x, s_1 \in S_1, ..., s_{d-1} \in S_d, s_d = y$, denote by $A(l) := \sum_{k=1}^{l} N_k^{s_{k-1},s_k}$ and consider the path set

$$\Gamma^4 := \{\gamma^4_{xy}\} \text{ with} \tag{2.30}$$

$$\gamma^4_{xy} = \begin{cases} \frac{(A(1)-n)s_0+ns_1}{N_1^{s_0,s_1}} & , \quad 0 \le n \le A(1), \\ \frac{(A(2)-n)s_1+(n-A(1))s_2}{N_2^{s_1,s_2}} & , \quad A(1) \le n \le A(2), \\ \dots & \\ \frac{(A(d)-n)s_{d-1}+(A(d-1)-n)s_d}{} & , \quad A(d-1) \le n \le A(d). \end{cases}$$

As in the first example of this subsection, the path set Γ^4 fulfills the requirements of Definition 2.2.3, because we already observed that each path component obeys the p.r.p., see the examples of the previous section.

We will use the extended p.r.p. for $N = 2$ in the examples given in Section 2.7 for $d-$dim crosses.

2.3 Poincaré Inequalities

The first classical inequality under consideration is the Poincaré inequality, see (1.11) of the introduction. For $C > 0$ and all $f \in D(\mathcal{E})$

$$Var_\mu[f] \le C\mathcal{E}(f, f).$$

Using the variational characterization of the spectral gap λ_1, (1.10), the constant C is a lower bound on λ_1, i.e. $\lambda_1 \ge 1/C$.

To state our results we have to introduce another projection map on E_1. Let

$$\begin{aligned} pr_{(1,2)} : E_1 &\rightarrow S \times S, \\ (x, y, n) &\mapsto (x, y) \end{aligned}$$

denote the projection on the first two coordinates.
In the sequel we will abbreviate

$$\begin{array}{lll} pr_{(1,2)}(G^{-1}(v,w,m,n)) & \text{by} & pr_{1,2}G^{-1}, \\ pr_1(G^{-1}(v,w,m,n)) & \text{by} & pr_1 G^{-1}, \\ pr_2(G^{-1}(v,w,m,n)) & \text{by} & pr_2 G^{-1}. \end{array}$$

With these notational simplifications we can state our first theorem.

Theorem 2.3.1 *Let $(X_t)_t$ be a reversible Sub-Markov chain in discrete time or a time continuous Sub-Markov jump process on a Polish space $(S, \mathcal{S}, \mu)$. Assume the transition kernel of the chain is given by (2.1) and there exists an admissible path set Γ obeying the p.r.p.*
Then, $(\mathcal{E}, D(\mathcal{E}))$ satisfies a Poincaré inequality with constant

(2.31) $\kappa^{\Gamma}(p) :=$

$$\operatorname*{esssup}_{(v,w)\,\in\,\mathcal{A}^{\Gamma}} \left\{ \sum_{m,n\in\mathbb{N}} |||\gamma_{pr_{1,2}G^{-1}}|||_p h(pr_1 G^{-1}) h(pr_2 G^{-1}) \frac{h_1(v,w,m,n) 1_{E_2}}{Q(v,w)^{1-p}} \right\}.$$

__**Proof:**__
For arbitrary points $x, y \in S$ take the path γ_{xy} connecting x and y. For a bounded measurable function f using the triangle inequality, we obtain

$$|f(x) - f(y)| \leq \sum_{n=1}^{N(x,y)} |f(\gamma_{xy}(n-1)) - f(\gamma_{xy}(n))|.$$

Since,

$$2Var_{\pi}[f] = \int_S \int_S |f(x) - f(y)|^2 \, d\pi(x) d\pi(y)$$

we obtain the inequality

$$2Var_{\pi}[f] \leq \int_S \int_S \left(\sum_{n=1}^{N(x,y)} |f(\gamma_{xy}(n-1)) - f(\gamma_{xy}(n))| \right)^2 d\pi(x) d\pi(y).$$

Applying the Cauchy-Schwarz inequality for the counting measure to

$$\left(\sum_{n=1}^{N(x,y)} |f(\gamma_{xy}(n-1)) - f(\gamma_{xy}(n))| \frac{Q(\gamma_{xy}(n-1), \gamma_{xy}(n))^{p/2}}{Q(\gamma_{xy}(n-1), \gamma_{xy}(n))^{p/2}} \right)^2,$$

we obtain that $2Var_\pi[f]$ is bounded by

$$\int_S \int_S \left(\sum_{n=1}^{N(x,y)} |f(\gamma_{xy}(n-1)) - f(\gamma_{xy}(n))|^2 Q(\gamma_{xy}(n-1), \gamma_{xy}(n))^p \right) \times \underbrace{\left(\sum_{[n=1}^{N(x,y)} Q(\gamma_{xy}(n-1), \gamma_{xy}(n))^{-p} \right)}_{=|||\gamma_{xy}|||_p} d\pi(x) d\pi(y).$$

With denoting by F the measurable function

$$F(x,y,n) := |f(\gamma_{xy}(n-1)) - f(\gamma_{xy}(n))|^2 Q(\gamma_{xy}(n-1), \gamma_{xy}(n))^p |||\gamma_{xy}|||_p h(y) h(x),$$

we can apply the p.r.p. and obtain

$$\begin{aligned} \int_S \int_S \sum_{n=1}^{N(x,y)} F(x,y,n)\, d\mu(y) d\mu(x) &= \int_{E_2} F \circ G^{-1} h_1 \, d\mu d\mu d\delta d\delta \\ &= \int_S \int_S \sum_{m,n \in \mathbb{N}} F \circ G^{-1} 1_{E_2} h_1 \, d\mu d\mu =: (*). \end{aligned}$$

For the last equalitiy we used the Fubini theorem. The factor 1_{E_2} has to be added because the density h_1 is only defined on E_2.
The last term (*) is equal to:

$$\begin{aligned} (*) &= \int_S \int_S \sum_{m,n \in \mathbb{N}} |||\gamma_{pr_{1,2}G^{-1}}|||_p h(pr_1 G^{-1}) h(pr_2 G^{-1}) \\ &\quad \times Q(v,w)^{p-1} h_1(v,w,m,n) 1_{E_2} \times Q(v,w) |f(v) - f(w)|^2 d\mu(v) d\mu(w). \end{aligned}$$

And this can now be estimated by

$$\underset{(v,\,w)\,\in\,\mathcal{A}^\Gamma}{\text{esssup}} \left\{ \sum_{m,n \in \mathbb{N}} |||\gamma_{pr_{1,2}G^{-1}}|||_p h(pr_1 G^{-1}) h(pr_2 G^{-1}) h_1(v,w,m,n) Q(v,w)^{p-1} \right\} \times \int_S \int_S Q(v,w) |f(v) - f(w)|^2 \, d\mu(v) d\mu(w).$$

Therefore, we have shown

$$Var_\pi[f] \leq \kappa^\Gamma \mathcal{E}(f,f).$$

This proves the Theorem. □
We apply the bound from Theorem 2.3.1 to various examples in Section 2.7.

In our setting, i.e. general state spaces, it is not obvious that a spectral gap must exist. As we already mentioned in the introduction all our results can also be seen as existence results for a spectral gap.

Simplification of the constant
The constant $\kappa^\Gamma(p)$ in (2.31) seems to be very complicated. It can be simplified by using another notation. Furthermore, for the path sets introduced in (2.21) and (2.22) the density h_1 is a constant and (2.31) becomes more simple.
For a given edge $e = (v, w)$ there is a $1 : 1$ correspondance of the index of e and the length of the path in which e appears. Hence, to have a more simple notation we suppress the dependence of v, w, m and n in the arguments of h and $|||\cdot|||_p$ in (2.31) and in the sequel by writing x, y instead of $pr_{1,2}G^{-1}(v, w, m, n)$, x instead of $pr_1 G^{-1}(v, w, m, n)$ and y instead of $pr_2 G^{-1}(v, w, m, n)$. With these notational simplification the constant becomes

$$\kappa^\Gamma(p) = \operatorname*{esssup}_{(v,w)\,\in\,\mathcal{A}^\Gamma} \left\{ \sum_{m,n} |||\gamma_{xy}|||_p h(x)h(y)h_1(v, w, m, n)Q(v, w)^{p-1}1_{E_2} \right\}.$$

In the sequel we derive a local Poincaré inequality. We denote by $B_x(r)$ the ball with radius r centered at $x \in S$ with volume $\mathrm{Vol}(B_x(r)) := \int 1_{B_x(r)}(y)d\pi(y)$. We further define

$$f_r(x) := \frac{1}{\mathrm{Vol}(B_x(r))} \int 1_{B_x(r)}(y)f(y)\,d\pi(y).$$

Definition 2.3.2 *Let $(\mathcal{E}, D(\mathcal{E}))$ be a Dirichlet form on $L^2(S, \mu)$ and denote by $diam(S)$ the diameter of S. $(\mathcal{E}, D(\mathcal{E}))$ satisfies a local Poincaré inequality if for $r \in [0, diam(S)]$, there exists a constant C_r which only depends on r such that for all $f \in D(\mathcal{E})$*

$$\|f - f_r\|_2^2 \le C_r \mathcal{E}(f, f). \tag{2.32}$$

Remark 2.3.3 *For bounded S we can deduce a Poincaré inequality if we only consider $r = diam(S)$.*

We further introduce the measure space $E_{1,r} \subset E_1$ by

$$E_{1,r} := \{(x, y, n) : x \in S, y \in B_x(r), 1 \le n \le N(x, y)\}$$

and define $E_{2,r} := G(E_{1,r})$. Having chosen a path set obeying the p.r.p. one can derive local Poincaré inequalities by the following theorem.

Theorem 2.3.4 *Let $(X_t)_t$ be a reversible Sub-Markov chain in discrete time or a time continuous Sub-Markov jump process on a bounded Polish space $(S, \mathcal{S}, \mu)$. Assume the transition kernel of the chain is given by (2.1) and there exists an admissible path set Γ obeying the p.r.p.*
Then, $(\mathcal{E}, D(\mathcal{E}))$ satisfies a local Poincaré inequality with constant

$$\eta^{\Gamma}(p,r) := \tag{2.33}$$

$$\operatorname*{esssup}_{(v,\,w)\,\in\,\mathcal{A}^\Gamma} \left\{ \frac{1}{Q(v,w)^{1-p}} \sum_{\substack{m,n\in\mathbb{N}\\ d(x,y)\le r}} |||\gamma_{xy}|||_p \frac{h(x)h(y)}{Vol(B_x(r))} h_1(v,w,m,n) 1_{E_{2,r}} \right\}.$$

<u>Proof:</u>
We abbreviate $\mathrm{Vol}(B_x(r))$ by $V_x(r)$. Jensen's inequality implies

$$|f(x) - f_r(x)|^2 \le \frac{1}{V_x(r)} \int_{B_x(r)} |f(x) - f(y)|^2 \, d\pi(y).$$

Consider the integrand on the right hand side. For each $y \in B_x(r)$ take the path γ_{xy} connecting x and y. As in the latter proof we obtain using the triangle inequality and the Cauchy-Schwarz inequality

$$\begin{aligned} (**) &= |f(x) - f(y)|^2 \\ &\le \sum_{n=1}^{N(x,y)} |f(\gamma_{xy}(n-1)) - f(\gamma_{xy}(n))|^2 Q(\gamma_{xy}(n-1), \gamma_{xy}(n))^p |||\gamma_{xy}|||_p. \end{aligned}$$

With

$$F_r(x,y,n) := |f(\gamma_{xy}(n-1)) - f(\gamma_{xy}(n))|^2 Q(\gamma_{xy}(n-1), \gamma_{xy}(n))^p |||\gamma_{xy}|||_p \frac{h(x)h(y)}{V_x(r)}$$

and integrating (**) over S with respect to π yields

$$2\|f - f_r\|_2^2 \le \int_S \int_{B_x(r)} \sum_{n=1}^{N(x,y)} F_r(x,y,n) \, d\mu(y) d\mu(x).$$

Applying the p.r.p. and the simplification of the constant described just before stating the Theorem we find

$$\int_S \int_{B_x(r)} \sum_{n=1}^{N(x,y)} F_r(x,y,n) \, d\mu(y) d\mu(x) = \int_{E_{2,r}} F_r \circ G^{-1} h_1 \, d\mu d\mu$$

$$
\begin{aligned}
&\leq \quad \operatorname*{esssup}_{(v,w)\,\in\,\mathcal{A}^\Gamma} \left\{ Q(v,w)^{p-1} \sum_{\substack{m,n\in\mathbb{N}\\ d(x,y)\leq r}} |||\gamma_{xy}|||_p \frac{h(x)h(y)}{\mathrm{Vol}(B_x(r))} h_1(v,w,m,n) 1_{E_{2,r}} \right\} \\
&\quad \times \int_{S\times S} Q(v,w)(f(v)-f(w))^2 \, d\mu(v) d\mu(w) \\
&= \quad 2\eta^\Gamma(p,r)\mathcal{E}(f,f)
\end{aligned}
$$

This proves the Theorem. □

In the finite setting it has been proven useful to consider general measures on path sets. We refer to [S 92] for further explanations and examples.

With no further difficulties we are able to prove similar results in the setting for general state spaces. In Definition 2.3.5 we introduce the notions of a weight and use it to prove a more sophisticated version of a Poincaré inequality in Theorem 2.3.6. All the results in the next two sections can be generalized to weights but we do not apply these bounds to any concrete examples.

Definition 2.3.5 *For a path set $\Gamma = \{\gamma_{xy}\}$ a weight function ω is a positive function*

$$
\begin{aligned}
&(2.34) \qquad \omega : \mathcal{A}^\Gamma \;\to\; (0,\infty) \\
&(2.35) \qquad\qquad e \;\mapsto\; \omega(e).
\end{aligned}
$$

The ω-length of a path $\gamma_{xy} \in \Gamma$ is given by

$$
(2.36) \qquad |\gamma_{xy}|_\omega := \sum_{n=1}^{N(x,y)} \omega((\gamma_{xy}(n-1), \gamma_{xy}(n)))^{-1}.
$$

We remind the reader of the simplification of the constant earlier in this chapter. Using this simplification we can state the following theorem.

Theorem 2.3.6 *Let $(X_t)_t$ be a reversible Sub-Markov chain in discrete time or a time continuous jump Sub-Markov process on a bounded Polish space $(S, \mathcal{S}, \mu)$. Assume the transition kernel of the chain is given by (2.1) and there exists an admissible path set Γ obeying the p.r.p.*
Then, a weighted Poincaré inequality with constant

$$
(2.37) \quad \kappa^{\Gamma,\omega} = \operatorname*{\mathit{esssup}}_{(v,w)\,\in\,\mathcal{A}^\Gamma} \left\{ \frac{\omega((v,w))}{Q(v,w)} \sum_{m,n\in\mathbb{N}} |\gamma_{xy}|_\omega h(x)h(y)h_1(v,w,m,n)1_{E_2} \right\}
$$

holds true.

Proof:
For $x, y \in S$ take the path γ_{xy} connecting x and y and observe that by triangle and the Cauchy-Schwarz inequality

$$\begin{aligned}
&|f(y) - f(x)|^2 \\
&\leq \left(\sum_{n=1}^{N(x,y)} \frac{1}{\omega(\gamma_{xy}(n-1), \gamma_{xy}(n))} \right) \\
&\quad \times \left(\sum_{n=1}^{N(x,y)} |f(\gamma_{xy}(n-1)) - f(\gamma_{xy}(n))|^2 \omega(\gamma_{xy}(n-1), \gamma_{xy}(n-1)) \right) \\
&= |\gamma_{xy}|_\omega \sum_{n=1}^{N(x,y)} |f(\gamma_{xy}(n-1)) - f(\gamma_{xy}(n))|^2 \omega(\gamma_{xy}(n-1), \gamma_{xy}(n))
\end{aligned}$$

Integration over $S \times S$ with respect to $\pi \otimes \pi$ gives

$$\begin{aligned}
&2\|f - \mathbb{E}_\pi[f]\|_2^2 \leq \\
&\int_S \int_S |\gamma_{xy}|_\omega \sum_{n=1}^{N(x,y)} |f(\gamma_{xy}(n-1)) - f(\gamma_{xy}(n))|^2 \omega(\gamma_{xy}(n-1), \gamma_{xy}(n)) \, d\pi(x) d\pi(y).
\end{aligned}$$

Then continuing as in the proof of Theorem 2.3.1. proves the theorem. □

2.4 Moderate Growth and Nash Inequalities

We further explore the range of inequalities obtainable using the constant $\eta^\Gamma(p, r)$. In the sequel we consider Nash-inequalities of type I and II, i.e.

$$Var_\pi[f]^{1+2/d} \leq C_1 \mathcal{E}(f, f) \|f\|_1^{4/r} \quad \text{(type I)} \tag{2.38}$$

and

$$\|f\|_2^{2+4/d} \leq C_2 \left[\mathcal{E}(f, f) + \frac{1}{aR^2} \|f\|_2^2 \right] \|f\|_1^{4/d} \quad \text{(type II).} \tag{2.39}$$

Definition 2.4.1 *Given a Sub-Marvkov chain $(X_t)_t$ in discrete time or a time continuous Sub-Markov jump process and a probability measure π having a density with respect to μ on M, we say for constants $A, d \geq 1$ the chain has (A, d)-moderate growth if for all $x \in S, r \in \{0, 1, ..., diam(M)\}$*

$$Vol(B_x(r)) \geq \frac{1}{A} \left(\frac{r+1}{diam(M)} \right)^d. \tag{2.40}$$

This leads to the following

Theorem 2.4.2 *Let $(X_t)_t$ be a reversible Sub-Markov chain in discrete time or a time continuous jump Sub-Markov process on a bounded Polish space $(S, \mathcal{S}, \mu)$. Assume the transition kernel of the chain is given by (2.1) and there exists an admissible path set obeying the p.r.p. If there exist real numbers $M, d \geq 1, a > 0$ such that for some $R \in \mathbb{N}$*

$$B_x(r) \geq \frac{(r+1)^d}{M} \text{ and } \|f - f_r\|_2^2 \leq \eta^\Gamma(p,r)\mathcal{E}(f,f). \tag{2.41}$$

Suppose $\eta^\Gamma(p,r) \leq ar^2$ for $r \in [0,R)$. Then, for all $x \in S$, $r \in [0,R]$ and all functions $f \in D(\mathcal{E}) \cap L^1(S,\mu)$ (2.38) and (2.39) hold true for $C_1 = C_2 = (1+1/d)2(1+d)^{2/d}M^{2/d}a$.

Since the proof is essentially the same as in the finite setting we postpone the proof of this theorem to Chapter 4.

2.5 Isoperimetric Inequalities

In the previous sections we derived Poincaré inequalities of various types. As mentioned in the introduction, in [LS 88], Lawler and Sokal proved a version of Cheeger's inequality. For the transition kernel in (2.1) it reads

Lemma 2.5.1 *([LS 88]) The spectral gap of a Sub-Markov chain with transition kernel given in (2.1) and the isoperimetric constant*

$$I_K := \min_{A \subset S} \frac{Q(\partial(A))}{\pi(A)\pi(A^c)},$$

where the infimum is taken over all measurable subsets $A \subset S$, are related by

$$\frac{I_K^2}{8} \leq \lambda_1(K) \leq I_K. \tag{2.42}$$

In the sequel we use paths to state a lower bound on I_K and therefore on λ_1. To prove the following theorem we need a well known variational characterization of the isoperimetric constant I_K, i.e.

$$I_K = \inf_{\substack{f \in L^2(S,\pi) \\ f \text{ non constant}}} \left\{ \frac{\int\int |f(x) - f(y)|\, K(x,dy) d\pi(x)}{\int \left| f(x) - \int f(x)\, d\pi(x) \right|\, d\pi(x)} \right\}. \tag{2.43}$$

To see how isoperimetric and analytic notions are related we present a proof for the variational characterization.

Proof:
Take $f \in L^2(S,\pi)$ and consider $f = f_+ + f_-$ and $F_t^{\pm} := \{x : f_+(x) \geq t\}$. Where $f_+ = f \vee 0$ and $f = -[f \wedge 0]$.
We obtain the following estimate

$$\int |f_\pm(x) - f_\pm(y)| k(x,y)\, d\pi(x) \geq 2 I_K \int_0^\infty \pi(F_t^\pm)(1-\pi(F_t^\pm))\, dt.$$

Owing to an identity proved by Rothaus, in [R 85], we have

$$\begin{aligned} 2\pi(F_t^\pm)(1-\pi(F_t^\pm)) &= \int |1_{F_t^\pm(x)} - \pi(F_t^\pm)|\, d\pi(x) \\ &= \sup_{\substack{g:\int_g (x)\, d\pi(x)=0 \\ \inf_\alpha |g-\alpha|\leq 1}} \int 1_{F_t^\pm}(x) g(x)\, d\pi(x). \end{aligned}$$

Therefore, we have for $g \in L^2(S,\pi), \int g(x)\, d\pi(x) = 0, \inf |g-\alpha| \leq 1$

$$\begin{aligned} \int |f_\pm(x) - f_\pm(y)| k(x,y)\, d\pi(x) &\geq I_K \int (\int_0^\infty 1_{F_t^\pm}(x)\, dt) g(x)\, d\pi(x) \\ &\geq I_K \int f_\pm(x) g(x)\, d\pi(x). \end{aligned}$$

By adding the two inequalities for f_+ and f_-, we obtain

$$\int |f(x) - f(y)| k(x,y)\, d\pi(x) \geq I_K \int \left| f(x) - \int f(x)\, d\pi(x) \right| d\pi(x).$$

To prove $I_K \geq \frac{\int\int |f(x)-f(y)|K(xdy)d\pi(x)}{\int |f(x) - \int f(x) d\pi(x)| d\pi(x)}$, take $f := 1_A$. □
Again, we remind the reader of the simplification of the constant obtained using the p.r.p. The following theorem provides a lower bound on the isoperimetric constant I_K which, via Cheeger's inequality, leads to a lower bound on the spectral gap.

Theorem 2.5.2 *Let $(X_t)_t$ be a reversible Sub-Markov chain in discrete time or a time continuous jump Sub-Markov process on a measure space $(S, \mathcal{S}, \mu)$. Assume the transition kernel of the chain is given by (2.1) and there exists an admissible path set Γ obeying the p.r.p.*
Then, for

$$(2.44)\quad B^\Gamma := \operatorname*{esssup}_{(v,w) \in \mathcal{A}^\Gamma} \left\{ \sum_{m,n\in\mathbb{N}} Q(v,w)^{-1} h(x) h(y) h_1(v,w,m,n) 1_{E_2} \right\},$$

we have

$$(2.45)\quad I_K \geq \frac{1}{B^\Gamma}.$$

Proof:
By the usual argument invoving the triangle inequality we have

$$\int_S \int_S |f(x) - f(y)|\, d\pi(x) d\pi(y) \leq$$
$$\int_S \int_S \sum_{n=1}^{N(x,y)} |f(\gamma_{xy}(n-1)) - f(\gamma_{xy}(n))|\, d\pi(x) d\pi(y).$$

Denote by LS and RS the left hand side and respectively right hand side of the previous inequality.
The LS is an upper bound on $\int_S |f(x) - \mathbb{E}_\pi[f]|\, d\pi(x)$, whereas using the p.r.p. and the simplification of the constant described before Theorem 2.3.4, i.e. writing x instead of $pr_1 G^{-1}(v,w,m,n)$ and y instead of $pr_2 G^{-1}(v,w,m,n)$, we have

$$\begin{aligned} RS &= \int_S \int_S \sum_{m,n\in\mathbb{N}} \frac{Q(v,w)}{Q(v,w)} |f(v) - f(w)| h(x) h(y) \\ &\qquad \times h_1(v,w,m,n) 1_{E_2}\, d\mu(v) d\mu(w) \\ &\leq \operatorname*{esssup}_{(v,\,w)\,\in\,\mathcal{A}^\Gamma} \left\{ \sum_{m,n\in\mathbb{N}} \frac{h(x)h(y)}{Q(v,w)} h_1(v,w,m,n) 1_{E_2} \right\} \\ &\qquad \times \int_S \int_S |f(v) - f(w)| Q(v,w)\, d\mu(v) d\mu(w). \end{aligned}$$

This proves Theorem 2.5.2 □

Remark 2.5.3 *Not even for Markov chains on finite state spaces it is clear if lower bounds on the spectral gap obtained using Theorem 2.3.1 perform always better than those applying Theorem 2.5.2. Recently, Fulman and Wilmer in [FW 99] answered this question for several examples in the finite case in the affirmative.*
In Section 2.7 we also observe this fact.

2.6 Comparison

Next, we consider a measure space $(S, \mathcal{S}, \mu)$ and two reversible Sub-Markov chains $(X_t)_t, (Y_t)_t$ in discrete time or time continuous Sub-Markov jump processes having transition kernels $K(x,dy)$ and $\tilde{K}(x,dy)$. The corresponding stationary probability measures are denoted by π and $\tilde{\pi}$, having densities h

and $\tilde{h}$ with respect to μ.

PATH SETS:
We have to change the definition of admissibility to be able to compare two Sub-Markov chains. This is necessary because the domain of the corresponding transition kernels can be different.

Definition 2.6.1 *We say an admissible collection of paths $\Gamma = \{\gamma_{xy}\}$ is an c-admissible collection of paths (comparison admissible) if the conditions of Definition 2.0.2 are fulfilled for the transition kernels $\tilde{K}$ and K.*

The following theorem can be used to compare the spectral gap and the logarithmic-Sobolev constant for different Sub-Markov chains.

Theorem 2.6.2 *If there exist constants $a, A \in \mathbb{R}^+$ such that*

(i) $\tilde{\mathcal{E}}(f,f) \leq A\mathcal{E}(f,f)$, *for all* $f \in L^2(X,\pi)$ *and*

(ii) $a\pi \leq \tilde{\pi}$,

then, we have

$$\tilde{\lambda}_1 \geq \frac{A}{a}\lambda_1 \text{ and } \tilde{\alpha} \geq \frac{A}{a}\alpha \tag{2.46}$$

<u>**Proof:**</u>
We use the fact that the variance $Var_\pi[f]$ is given by

$$Var_\pi[f] = \inf_c \left\{ \int_S |f(x) - c|^2 \, d\pi(x) \right\}.$$

Therefore,

$$Var_\pi[f] \leq \frac{1}{a} Var_{\tilde{\pi}}[f].$$

Finally, this leads to

$$\frac{\mathcal{E}(f,f)}{Var_\pi[f]} \geq \frac{a}{A} \frac{\tilde{\mathcal{E}}(f,f)}{Var_{\tilde{\pi}}[f]}.$$

By the variational characterization of the spectral gap λ_1 the first assertion of the theorem follows.
To prove the second assertion we use the variational characterization of the logarithmic-Sobolev functional proposed by Holley and Stroock in [HS87], i.e.

$$\mathcal{L}_\pi[f] = \inf_c \left\{ \int_S \left(f(x)^2 \log(f(x)^2) - f(x)^2 \log(c) - f(x)^2 + c \right) \, d\pi(x) \right\}.$$

Consequently,

$$\mathcal{L}_\pi[f] \leq \frac{1}{a}\mathcal{L}_{\tilde{\pi}}[f].$$

Similar to the case for the spectral gap we find

$$\frac{\mathcal{E}(f,f)}{\mathcal{L}_\pi[f]} \geq \frac{a}{A}\frac{\tilde{\mathcal{E}}(f,f)}{\mathcal{L}_{\tilde{\pi}}[f]}.$$

Therefore, the theorem is proved. □

Remark 2.6.3 *1) The state space may be infinite dimensional.*
2) If $\tilde{\pi} = \pi$ we can take $a = 1$ and we can restrict ourselves to compare the Dirichlet forms.
3) If the state space is finite we recover the results obtained by Diaconis and Saloff-Coste. They also obtain lower bounds, but in this case the number of elements of the state space comes in.

In applications the necessity for comparing Sub-Markov chains on different state spaces occurs. This is possible with respect to the following remark.

Remark 2.6.4 *Let K and $\tilde{K}$ be transition kernels on $(S, \mathcal{S}, \mu)$ and $(\tilde{S}, \tilde{\mathcal{S}}, \tilde{\mu})$ respectively. Assume there exists a linear transformation T, such that*

$$\begin{array}{rcl} T : L^2(S,\pi) & \to & L^2(\tilde{\mathcal{S}}, \tilde{\pi}) \\ f & \mapsto & \tilde{f}. \end{array}$$

Then, if there are $A, a_1, a_2, B_1, B_2 \in \mathbb{R}^+$, such that for all $f \in L^2(S,\pi)$

$$\begin{array}{rcl} \tilde{\mathcal{E}}(\tilde{f},\tilde{f}) & \leq & A\mathcal{E}(f,f), \\ a_1 Var_\pi[f] & \leq & Var_{\tilde{\pi}}[\tilde{f}] + B_1\mathcal{E}(f,f), \\ a_2\mathcal{L}_\pi[f] & \leq & \mathcal{L}_{\tilde{\pi}}[\tilde{f}] + B_2\mathcal{E}(f,f), \end{array}$$

the following inequalities hold true

$$\frac{a_1\lambda_1(\tilde{K})}{A + B_1\lambda_1(\tilde{K})} \leq \lambda_1(K) \text{ and } \frac{a_2\alpha(\tilde{K})}{A + B_2\alpha(\tilde{K})} \leq \alpha(K). \tag{2.47}$$

The statements of Remark 2.6.4 follow immediately using the linear tranformation T and the variational characterizations of λ_1, α, $\tilde{\lambda_1}$ and $\tilde{\alpha}$. As in the finite setting these considerations, the simplification of the constant and Theorem 2.6.2 lead to

Theorem 2.6.5 *Let $(S, \mathcal{S}, \mu)$ be a measure space and $(X_t)_t, (Y_t)_t$ be two reversible discrete-time Sub-Markov chains or time-continuous Sub-Markov jump processes having transition kernels $K(x, dy)$ and $\tilde{K}(x, dy)$. The corresponding stationary probability measures are denoted by π and $\tilde{\pi}$, having densities h and $\tilde{h}$ w.r.t μ. Assume there exists c-admissible path set in the sense of Definition 2.6.1 on S obeying the p.r.p. In the above setting we have for $f \in D(\mathcal{E})$*

$$(f, \tilde{H}f)_{\tilde{\pi}} \leq A^{\Gamma}(p)(f, Hf)_{\pi} \tag{2.48}$$

with constant

$$A^{\Gamma}(p) := \underset{(v,w) \in \mathcal{A}^{\Gamma}}{\text{esssup}} \left\{ \sum_{m,n \in \mathbb{N}} |||\gamma_{xy}|||_p \tilde{h}(x)\tilde{k}(x,y) h_1(v,w,m,n) Q(v,w)^{p-1} 1_{E_2} \right\}.$$

<u>**Proof:**</u>
The corresponding Dirichlet forms are $\mathcal{E}(f, f) = 2(f, Hf)_{\pi}$ and $\tilde{\mathcal{E}}(f, f) = 2(f, \tilde{H}f)_{\tilde{\pi}}$. Applying the Cauchy-Schwarz inequality and the p.r.p., we obtain

$$\begin{aligned}
\tilde{\mathcal{E}}(f,f) &= 2(f, \tilde{H}f)_{\tilde{\pi}} \\
&\leq \int_S \int_S \left(\sum_{n=1}^{N(x,y)} \left(\frac{Q(\gamma_{xy}(n-1), \gamma_{xy}(n))}{Q(\gamma_{xy}(n-1), \gamma_{xy}(n))} \right)^{\frac{p}{2}} |f(\gamma_{xy}(n-1)) - f(\gamma_{xy}(n))| \right)^2 \\
&\qquad \times \tilde{h}(x)\tilde{k}(x,y)\, d\mu(x) d\mu(y) \\
&\leq \int_S \int_S |||\gamma_{xy}|||_p \tilde{h}(x)\tilde{k}(x,y) \\
&\qquad \times \sum_{n=1}^{N(x,y)} Q(\gamma_{xy}(n-1), \gamma_{xy}(n))^p) |f(\gamma_{xy}(n-1)) - f(\gamma_{xy}(n))|^2 \, d\mu(x) d\mu(y) \\
&\overset{p.r.p.}{=} \int_S \int_S \sum_{m,n \in \mathbb{N}} |||\gamma_{xy}|||_p \tilde{h}(x) Q(v,w)^p |f(v) - f(w)|^2 \\
&\qquad \times h_1(v,w,m,n) 1_{E_2} \, d\mu(v) d\mu(w) \\
&\leq A^{\Gamma}(p) \int_S \int_S Q(v,w)(f(v) - f(w))^2 \, d\mu(v) d\mu(w) \\
&= 2A^{\Gamma}(p)(f, Hf)_{\pi}
\end{aligned}$$

Where we used the simplification of the constant described before Theorem 2.3.4 in the step indicated with 'p.r.p'. This means we suppressed the dependence of v, w, m and n in the arguments of $\tilde{h}$ and $||| \cdot |||_p$. Then we write x, y instead of $pr_{1,2}G^{-1}(v, w, m, n)$, x instead of $pr_1 G^{-1}(v, w, m, n)$ and y instead of $pr_2 G^{-1}(v, w, m, n)$. Hence, the result follows. □

In applications given transition kernels can be very complicated. Computations involving such kernels can be very tricky or even impossible. In these situations one may take advantage of a comparison argument. Sometimes it is possible to find lower and/or upper bounds of the associated Dirichlet form in terms of a Dirichlet form corresponding to a simpler kernel. Then, one can apply Theorem 2.6.2. Whereas Theorem 2.6.5 can be useful to find a lower and/or upper bound. We apply both Theorems in Section 2.7 to some examples.

2.7 Examples

In the sequel we apply our results to some examples.
We should not expect that our lower bounds are always of the right order of magnitude. Already in the finite case one can find counterexamples, the 'n-dog' for instance, see Example 5 in Chapter 4, where neither the bound using a Poincaré inequality nor an isoperimetric inequality yields the right order of magnitude. But nevertheless, we observe that there are examples where the bounds can be sharp.
Firstly, we consider $d-$dim cubes and our main concern is to study the influence of increasing the dimension. Also, we give an example for the comparison argument.
The second example, a $d-$dim cross, illustratetes the use of the extended p.r.p. of Section 2.2.
Example 3 deals with the Gibbs sampler. There are not many tools available for estimating the rate of convergence for the Gibbs sampler on general state spaces as pointed out in the introduction. We will compare our results to an estimate obtained in [Ro 96].
The next few examples deal with submanifolds of $\mathbb{R}^d$. We are mainly interested in the cases where the set S is a circle or a torus. Here the spectrum can be computed explicitely using Fourier analysis. Therefore, we can check the performance of our methods.
For $n, k \in \mathbb{N}$ we need to calculate the sum $S_k(n) := \sum_{l=1}^{k} l^k$. This can be done using the following remark.

Remark 2.7.1 *Let $n \in \mathbb{N}$ and $k \leq n$. the sums $S_k(n)$ satisfy*

$$\binom{k+1}{1} S_k(n) + \binom{k+1}{2} S_{k-1}(n) + ... + \binom{k+1}{k+1} S_0(n) = (n+1)^{k+1} - 1.$$

We state the explicit formulas for $k = 0, 1, ..., 5$.

$$S_0(n) = n, S_1(n) = \frac{n(n+1)}{2}, S_2(n) = \frac{n(n+1)(2n+1)}{6}, S_3(n) = \frac{n^2(n+1)^2}{4},$$

$$S_4(n) = \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30}, S_5(n) = \frac{(2n+1)^3(n+1)^3}{48}.$$

Before we start with the first example we want to remind the reader that the density h_1 for subsets of $\mathbb{R}^d$ and path sets Γ^i, $i = 1, 2$, are N^d and $N(\cdot,\cdot)^d$ respectively. This was shown in Section 2 after Remark 2.2.2. We use this fact in the sequel with no further explaination.

The general idea is to compute the notions which are necessary to evalutate the constants appearing in the theorem. Then, we choose the parameters N and η for admissible path sets such that the bounds become optimal. Recall that admissiblility puts restrictions on the values of N and η.

Example 1: Cubes

We consider the measure space $(S, \mathcal{S}, \mu) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), dx)$. For $r > 0$ let the cube , Q_0^r, with side length $2r$ and center 0 (see Figure 2.1), called $(r, 0)$-cube, Q_0^r, be given by

$$Q_0^r := \left\{x \in \mathbb{R}^d : |x_i| \leq r\right\}.$$

For $r \geq \epsilon > 0$ and $x \in Q_0^r$ consider the (x, ϵ)-cube, Q_x^ϵ given by

$$Q_x^\epsilon := \left\{y \in \mathbb{R}^d : |x_i - y_i| \leq \epsilon\right\}.$$

We consider the transition kernel

$$K(x, dy) = \frac{1}{(2\epsilon)^d} 1_{Q_x^\epsilon \cap Q_0^r}(y) dy.$$

For K there is a stationary measure π which has density

$$h(x) := \frac{1}{(2r)^d} 1_{Q_0^r}(x)$$

with respect to Lebesgue measure dx.

Firstly, we observe that the path sets Γ^1 and Γ^2 introduced in (2.21) and (2.22) obey the p.r.p., because the $(r, 0)$-cube Q_0^r is convex. The map G is even differentiable and therefore measurable.

We proceed by computing the weights $|||\cdot|||_p$ for paths of the two different path sets.

$$\begin{aligned}
|||\gamma_{xy}^1|||_p &= \sum_{n=1}^{N} Q(\gamma_{xy}^1(n-1), \gamma_{xy}^1(n))^{-p} = N(4\epsilon r)^{dp} \\
|||\gamma_{xy}^2|||_p &= \sum_{n=1}^{N(x,y)} Q(\gamma_{xy}^2(n-1), \gamma_{xy}^2(n))^{-p} = N(x, y)(4\epsilon r)^{dp}.
\end{aligned}$$

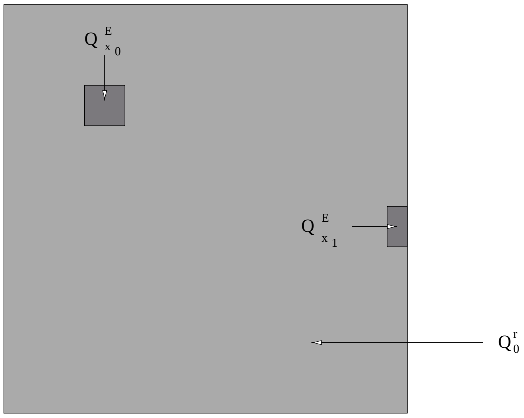

Figure 2.1: **The state space Q_0^r and the domain of k for two points** $x_0, x_1 \in Q_0^r$

Now, we have computed the relevant notions needed to calculate the constants of Theorem 2.3.1 and Theorem 2.5.2.
Since for the path set Γ^1 each path is of length N and there are edges of index 1 up to index N in each path the sum in

$$\kappa^{\Gamma^1}(p) = \operatorname*{esssup}_{(v,w)\in\mathcal{A}^{\Gamma^1}} \left\{ \sum_{m,n\in\mathbb{N}} |||\gamma_{xy}|||_p h(x)h(y)h_1(v,w,m,n)Q(v,w)^{p-1}1_{E_2} \right\}$$

is a sum over the first N natural numbers in m. Because the summands are all constant and the density is N^d we find

$$\kappa^{\Gamma^1} \leq N^{d+2}\left(\frac{\epsilon}{r}\right)^d.$$

Therefore, N has to be chosen such that $\frac{2\sqrt{d}r}{N} \leq \epsilon$ holds. This is necessary, because the path can be the longest one, which is the diagonal. And the condition assures that consecutive points on the path are in the domain of the kernel, i.e. it ensures admissibility.
The constant B^{Γ^1} providing a lower bound on the isoperimetric constant I_K can be estimated as follows. As in the case for computing the constant κ^{Γ^1} the sum appearing in the constant is over the first N natural numbers in n. Therefore, we have

$$B^{\Gamma^1} \leq N^d \sum_{n=1}^{N} \frac{h(x)h(y)}{Q(v,w)} = N^{d+1}\left(\frac{\epsilon}{r}\right)^d.$$

Owing to admissibility we have $\frac{2\sqrt{d}r}{N} \leq \epsilon$.
Considering the bounds on the spectral gap λ_1 obtained using Theorem 2.3.1 and Theorem 2.5.2, we find

$$\lambda_1 \geq \frac{r^d}{\epsilon^d N^{d+2}} \approx \frac{\epsilon^d}{r^d 4^d d d^{d/2}}$$

and

$$\lambda_1 \geq \frac{r^{2d}}{N^{2d+2}\epsilon^{2d}}/8 \approx \frac{\epsilon^d}{32 d^{d+1} 4^d r^d}.$$

As expected, the bound obtained applying Theorem 2.3.1 performs better. For both lower bounds we observe that an increase in the dimension causes an decrease in the lower bound. Since we applied path based methods and at some stage the length of the longest path comes in, we expected this phenomenom.
Now, consider the admissible path set Γ^2 defined in (2.22). The sum appearing in the expression

$$\kappa^{\Gamma^2}(p) = \operatorname*{esssup}_{(v,\,w)\,\in\,\mathcal{A}^{\Gamma^2}} \left\{ \frac{1}{Q(v,w)^{(1-p)}} \sum_{m,n\in\mathbb{N}} |||\gamma_{xy}^2|||_p h(x)h(y)N^d(x,y)1_{E_2} \right\}$$

is finite because the longest path, the diagonal, is of finite length. Hence, by calculating the length of the longest path and observing that a path of length L can at most have L edges. We find for the constants

$$\kappa^{\Gamma^2}(p) \leq \lceil \frac{2\sqrt{d}r}{\epsilon} \rceil \left(\frac{\epsilon}{r}\right)^d \sum_{n=1}^{\lceil 2\eta\sqrt{d}r \rceil} \sum_{m=1}^{n} n^{d+1} = \lceil \frac{2\sqrt{d}r}{\epsilon} \rceil \left(\frac{\epsilon}{r}\right)^d \sum_{n=1}^{\lceil \eta 2\sqrt{d}r \rceil} n^{d+2}.$$

The sums on the right hand sight can be calculated recursively by Remark 2.7.1.
Taking into account that the paths should be admissible, we need $\frac{1}{\lceil 2\eta\sqrt{d}r \rceil} \leq \epsilon$, i.e.

$$\kappa^{\Gamma^2}(p) \leq \left\lceil \frac{2\sqrt{d}r}{\epsilon} \right\rceil \left(\frac{\epsilon}{r}\right) \sum_{n=1}^{\lceil \frac{2\sqrt{d}r}{\epsilon} \rceil} n^{d+2}.$$

The same lines of thought lead to an upper bound on B^{Γ^2}, namely

$$B^{\Gamma^2} \leq \left(\frac{\epsilon}{r}\right)^d \sum_{n=1}^{\lceil \frac{2\sqrt{d}r}{\epsilon} \rceil} n^{d+2}.$$

We only state the lower bound for $d = 2$ and 3. With respect to admissibility we have for $d = 2$:

$$\lambda_1 \geq \frac{30}{\frac{2\sqrt{2}r}{\epsilon}(\frac{(2\sqrt{2}r}{\epsilon}+1)(2\frac{2\sqrt{2}r}{\epsilon}+1)(3\left(\frac{2\sqrt{2}r}{\epsilon}\right)^2+3\frac{2\sqrt{2}r}{\epsilon}-1)}\left(\frac{r}{\epsilon}\right)^2\frac{\epsilon}{2\sqrt{2}r} \approx \frac{\epsilon^4}{108r^4}$$

and

$$\lambda_1 \geq \left(\frac{30}{\frac{2\sqrt{2}r}{\epsilon}(\frac{2\sqrt{2}r}{\epsilon}+1)(2\frac{2\sqrt{2}r}{\epsilon}+1)(3\left(\frac{2\sqrt{2}r}{\epsilon}\right)^2+3\frac{2\sqrt{2}r}{\epsilon}-1)}\left(\frac{r}{\epsilon}\right)^2\right)^2 \approx \frac{\epsilon^6}{800r^6}.$$

For $d = 3$ we obtain

$$\lambda_1 \geq \left(\frac{r}{\epsilon}\right)^3\frac{\epsilon}{2\sqrt{3}r}\frac{48}{(2\lceil\frac{2\sqrt{3}r}{\epsilon}\rceil+1)(\lceil\frac{2\sqrt{3}r}{\epsilon}\rceil+1)} \approx \frac{\epsilon^4}{288r^4}$$

and

$$\lambda_1 \geq \left(\frac{48}{(2\lceil\frac{2\sqrt{3}r}{\epsilon}\rceil+1)(\lceil\frac{2\sqrt{3}r}{\epsilon}\rceil+1)}\frac{\epsilon^3}{r^3}\right)^2 \approx \frac{\epsilon^6}{2304r^6}.$$

Therefore, the bound using the Poincaré inequalities performs better than the bound using the lower bound on the isoperimetric constant I_K and Cheeger's inequality. With respect to the considerations in Section 2.5 we expected the bound using Theorem 2.3.1 to be better.
In the sequel we want to apply the comparison techniques, Theorems 2.6.2 and 2.6.5, to different kernels on the same cube.
For $r \geq \epsilon_1 \geq \epsilon_2 \geq 1$ let

$$\mathcal{E}_i(f,f) = \int_S\int_S (f(x)-f(y))^2 k_i(x,y)dy\pi_i(dx), \quad i = 1,2,$$

where $\pi_1 = \pi_2 = 1_{Q_0^r}(x)dx$, $k_1(x,y) = \frac{1}{2\epsilon_1}1_{Q_x^{\epsilon_1}\cap Q_0^r}(y)$ and $k_2(x,y) = 1_{Q_x^{\epsilon_2}\cap Q_0^r}(y)$. Taking $\beta = \frac{\epsilon_1}{\epsilon_2}$ we obtain

$$\begin{aligned}
\mathcal{E}_2(f,f) &= \int_S\int_S (f(x)-f(y))^2\frac{1}{2\epsilon_1}1_{Q_x^{\epsilon_1}\cap Q_0^r}(y)dy\pi(dx) \\
&\leq \beta\int_S\int_S (f(x)-f(y))^2\frac{1}{2\epsilon_2}1_{Q_x^{\epsilon_2}\cap Q_0^r}(y)dy\pi(dx) \\
&= \beta\mathcal{E}_1(f,f).
\end{aligned}$$

Therefore, by Theorem 2.6.2

$$\lambda(K_1) \geq \frac{\epsilon_2}{\epsilon_1}\lambda(K_2) \text{ and } \alpha(K_1) \geq \frac{\epsilon_2}{\epsilon_1}\alpha(K_2).$$

Take another transition kernel, $\tilde{K}$. For $x \in Q_0(r)$ we define

$$\tilde{k}(x,y) := \begin{cases} \frac{\exp\left(-\frac{\|x-y\|^2}{2}\right)}{(2\pi)^{\frac{d}{2}}} & , \quad y \in Q_0(r) \\ 0 & , \quad \text{otherwise.} \end{cases}$$

The stationary measure for $\tilde{k}$ is the equidistribution, i.e. $\tilde{\pi} = \pi$. Since we need c-admissibility, an edge can at most be of length $\sqrt{d}\epsilon$. We compute the constant $A^{\Gamma^1}(p)$ of Theorem 2.6.5. For the following computation we have to keep in mind that we already computed the weights $|||\cdot|||_p$ with respect to the transition kernel K.

$$\begin{aligned} A^{\Gamma^1}(p) &= \operatorname*{esssup}_{(v,\,w)\,\in\,\mathcal{A}^{\Gamma^1}} \left\{ \sum_{m,n\in\mathbb{N}} |||\gamma^1_{xy}|||_p \tilde{h}(x)\tilde{k}(x,y)h_1(v,w,m,n)1_{E_2}Q(v,w)^{p-1} \right\} \\ &\leq \frac{N^{d+2}2^{d/2}\epsilon^d}{\pi^{d/2}} \operatorname*{esssup}_{(v,\,w)\,\in\,\mathcal{A}^{\Gamma^1}} \left\{ \exp\left(-\frac{\|v-w\|^2}{2}\right)\right\}. \end{aligned}$$

Since admissibility is assumed we must have $\frac{2\sqrt{d}r}{N} \leq \epsilon$. Choosing the smallest N satisfying the condition, Theorem 2.6.5 yields

$$\tilde{\lambda}_1 \geq \frac{\pi^{d/2}\epsilon^2\exp(\epsilon^2/2)}{r^{d+2}2^{d+2+d/2}d^{d/2+1}}\lambda_1.$$

The previous example points out the usefulness of a comparison argument. The spectral gap corresponding to the more intricate kernel $\tilde{K}$ can be estimated from below using a lower bound for the spectral gap of the more simple kernel K.

Example 2: Crosses

We consider the measure space $(S, \mathcal{S}, \mu) = (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), dx)$ where dx as usual denotes the Lebesgue measure. For $l_1, l_2, b_1, b_2 \in \mathbb{R}^+$ consider the sets

$$A := \{(x,y) \in \mathbb{R}^2 : -b_2 - l_2 \leq y \leq b_2 + l_2, -b_1 \leq x \leq b_1\}$$

and

$$B := \{(x,y) \in \mathbb{R}^2 : -b_1 - l_1 \leq x \leq b_1 + l_1, -b_2 \leq y \leq b_2\}.$$

The cross centered at 0 is the set $C_{l_1,l_2,b_1,b_2} := A \cup B$. In the sequel we abbreviate C_{l_1,l_2,b_1,b_2} by C. Figure 2.2. illustrates the situation.
Let $\epsilon > 0$, $x \in C$ and take the (x,ϵ)-cube Q^ϵ_x, i.e.

$$Q^\epsilon_x := \left\{ y \in \mathbb{R}^d : |x_i - y_i| \leq \epsilon, i = 1, 2 \right\}.$$

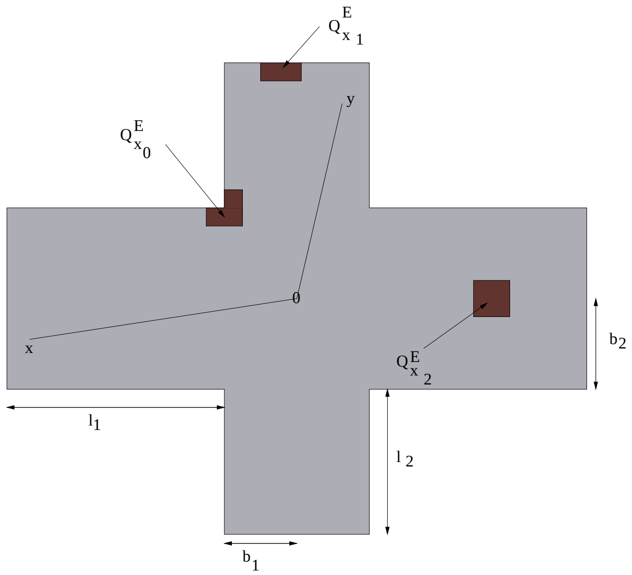

Figure 2.2: **The state space C_{l_1,l_2,b_1,b_2} and the domain of k for three points $x_0, x_1, x_2 \in C_{l_1,l_2,b_1,b_2}$**

We consider the transition kernel

$$K(x, dy) = \frac{1}{4\epsilon^2} 1_{Q_x^\epsilon \cap C}(y) dy.$$

With respect to Lebesgue measure, dx, the stationary measure π has density

$$h(x) = \frac{1}{2(l_1 b_2 + l_2 b_1) + b_1 b_2} 1_C(x) = \frac{1}{Vol(C)} 1_C(x).$$

We observe that the length L of the longest path component connecting a point $x \in C$ and 0 can at most be $d_1 \vee d_2$. Choose $d_1 = \sqrt{(l_1 + b_1)^2 + b_2^2}$ and $d_2 = \sqrt{(l_2 + b_2)^2 + b_1^2}$. Without loss of generality we may assume $d_1 \geq d_2$ and hence $L \leq \sqrt{2} d_1$.
Now, we consider the path set Γ^3 introduced in (2.29) and (2.30). Since we have for the convex sets A and B from above $C = A \cup B$ and $A \cap B \neq \emptyset$, the path set Γ^3 obeys therefore the extended p.r.p.
The constants can be computed by considering the components of a path $\gamma^3_{xy} = \gamma^1_{x0} + \gamma^1_{0y}$ where γ^1_{x0} and $\gamma^1_{0,y}$ are the discrete straight line segments connecting x, 0, and 0, y respectively. We have

$$|||\gamma^1_{x0}|||_p = \sum_{n=1}^{N_1} Q(\gamma^1_{x0}(n-1), \gamma^1_{x0}(n))^{-p} = N_1 \left(\frac{1}{4\epsilon^2 \mathrm{Vol}(C)} \right)^{-p},$$

$$|||\gamma^1_{0y}|||_p = \sum_{n=1}^{N_2} Q(\gamma^1_{0y}(n-1), \gamma^1_{0y}(n))^{-p} = N_2 \left(\frac{1}{4\epsilon^2 \mathrm{Vol}(C)} \right)^{-p}.$$

Since we already computed the constants $\kappa^{\Gamma^1}(p)$ and the path set Γ^3 consists of segments of paths of the path set Γ^1, we obtain

$$\kappa^{\Gamma^3}(p) \leq \frac{4\epsilon^2}{\mathrm{Vol}(C)}(N_1^4 + N_2^4).$$

For the lower bound on the isoperimetric constant I_K we have to estimate the constant B^{Γ^3} and with the same considerations we find

$$B^{\Gamma^3} \leq (N_1^3 + N_2^3)\frac{4\epsilon^2}{\mathrm{Vol}(C)}.$$

Since admissibility is assumed, we need $N_1, N_2 \geq d_1/\epsilon$. Thus,

$$\lambda_1 \geq \frac{Vol(C)}{4\epsilon^2(N_1^4 + N_2^4)} \approx \frac{\mathrm{Vol}(C)\epsilon^2}{8d_1^4}$$

and respectively

$$\lambda_1 \geq \left(\frac{\mathrm{Vol}(C)^2}{4\epsilon^2(N_1^3 + N_2^3)} \right)^2 /8 \approx \frac{\mathrm{Vol}(C)\epsilon^4}{256 d_1^6}.$$

Now, we take the admissible path set Γ^4. Firstly, we compute the path measures and find

$$\begin{aligned}
|||\gamma^2_{x0}|||_p &= \sum_{n=1}^{N_1(x,0)} Q(\gamma^2_{x0}(n-1), \gamma^2_{x0}(n))^{-p} \\
&= N_1(x,0)(4\epsilon^2\mathrm{Vol}(C))^p, \\
|||\gamma^2_{0y}|||_p &= N_2(0,y)(4\epsilon^2\mathrm{Vol}(C))^p.
\end{aligned}$$

Using this, we obtain

$$\begin{aligned}
\kappa^{\Gamma^4}(p) &\leq 2\frac{4\epsilon^2}{\mathrm{Vol}(C)} \sum_{n=1}^{\lceil \eta d_1 \rceil} \sum_{m=1}^{n} n^3 \\
&= \frac{8\epsilon^2}{15\mathrm{Vol}(C)} \lceil \eta d_1 \rceil \left(2\lceil \eta d_1 \rceil + 1\right)\left(\lceil \eta d_1 \rceil + 1\right)\left(3\lceil \eta d_1 \rceil^2 + 3\lceil \eta d_1 \rceil - 1\right).
\end{aligned}$$

This leads to a upper bound on the isoperimetric constant I_K:

$$B^{\Gamma^4} \leq 2\frac{4\epsilon^2}{\mathrm{Vol}(C)} \sum_{n=1}^{\lceil \eta d_1 \rceil} \sum_{m=1}^{n} n^2 = \frac{4\epsilon^2}{\mathrm{Vol}(C)} \lceil \eta d_1 \rceil^2 \left(\lceil \eta d_1 \rceil + 1\right)^2.$$

Since admissiblility is assumed we need $\eta \geq 1/\epsilon$. Therefore, we obtain

$$\begin{aligned}\lambda_1 &\geq \frac{15\mathrm{Vol}(C)}{8\epsilon^2\lceil\eta d_1\rceil\left(2\lceil\eta d_1\rceil+1\right)\left(\lceil\eta d_1\rceil+1\right)\left(3\lceil\eta d_1\rceil^2+3\lceil\eta d_1\rceil-1\right)}\\ &\approx \frac{3}{4}\frac{\mathrm{Vol}(C)\epsilon^2}{d_1^4(d_1+\epsilon)}\end{aligned}$$

and

$$\lambda_1 \geq \left(\frac{\mathrm{Vol}(C)}{4\epsilon^2\lceil\eta d_1\rceil^2\left(\lceil\eta d_1\rceil+1\right)^2}\right)/8 \approx \frac{\mathrm{Vol}(C)^2\epsilon^4}{104 d_1^6}$$

respectively.
In this case the spectrum cannot be computed explicitely because C has a boundary and one has to observe that the kernel K is not homogenous in the space variable, i.e. it depends explicitely on the starting point. We have not found any results concerning the last example in the literature. But we expected a close relation to Example 1, because we used two path sets similar to those in Example 1.
To study the effect of an increase in the dimension on the lower bounds on the spectral gap, we consider $d-$dimensional crosses.
To this end consider the $d-$dim cross $C^d_{r_1,\ldots,r_d,b_1,\ldots,b_d}$, centered at 0. We find by similar computations as above

$$\begin{aligned}\kappa^{\Gamma^3} &= (N_1^{d+2}+N_2^{d+2})\frac{2^d\epsilon^d}{\mathrm{Vol}(C)} \approx 2^{d+1}\frac{d_1^{d+2}}{\epsilon^2\mathrm{Vol}(C^d)},\\ B^{\Gamma^3} &= (N_1^{d+1}+N_2^{d+1})\frac{2^d\epsilon^d}{\mathrm{Vol}(C^d)} \approx 2^{d+1}\frac{d_1^{d+1}}{\epsilon\mathrm{Vol}(C^d)}.\end{aligned}$$

We observe that increasing the dimension by 1 yields a factor $2d_1$ in our upper bounds on the corresponding constants.
The length D of the longest line connecting a point $x \in C^d_{l_1,\ldots,l_d,b_1,\ldots,b_d}$ and 0 used to state upper bounds on the constants κ^{Γ^4} and B^{Γ^4} can be obtained by computing the length of the diagonals connecting a corner point and 0. Let the diagonals be denoted by $d_1, d_2, \ldots, d_d$ and if we take $D = \max_{i=1,\ldots,d} d_i$ this yields

$$\kappa^{\Gamma^4} \leq \frac{2^{d+1}\epsilon^d}{\mathrm{Vol}(C^d)}\sum_{n=1}^{\lceil\eta D\rceil} n^{d+2} \text{ and } B^{\Gamma^4} \leq \frac{2^{d+1}\epsilon^d}{\mathrm{Vol}(C^d)}\sum_{n=1}^{\lceil\eta D\rceil} n^{d+1}.$$

With regards to the latter inequalities, the effect of increasing the dimension can directly be seen. It increases the upper bounds on the corresponding constants and therefore decreases the corresponding lower bounds on the spectral gap.

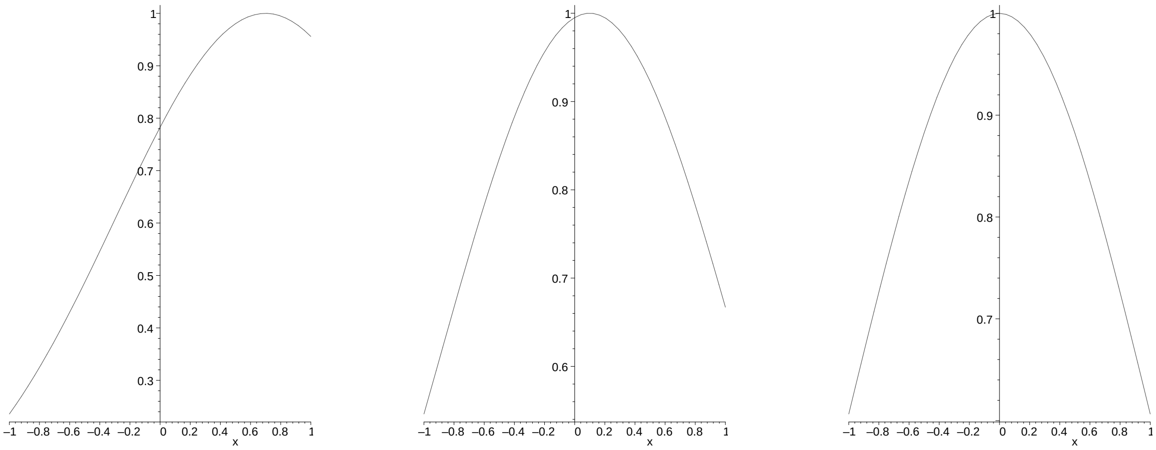

Figure 2.3: **Density plot for the kernel of Example 3**

Example 3: Balls

We consider the case where the state space S is the d-dimensional ball $B_d(r)$ with radius r and center 0, i.e.

$$B_d(r) = \{x \in \mathbb{R}^d : \|x\|_{eukl.} \leq r\}.$$

Take a reversible Sub-Markov transition kernel K, given by

$$K(x, dy) = k(x, y)dy, \text{ where}$$

$$k(x,y) := \begin{cases} \frac{\exp\left(-\frac{\|x-y\|^2}{2}\right)}{(2\pi)^{\frac{d}{2}}} & , \quad y \in B_d(r), \\ 0 & , \quad \text{otherwise.} \end{cases}$$

Figure 3.3 illustrates the situation. The chain has stationary measure

$$\pi(dx) = \frac{1}{\mathrm{Vol}(B_d(r))} 1_{B_d(r)}(x)dx, \text{ where}$$

$$\mathrm{Vol}(B_d(r)) = \frac{\pi^{\frac{d}{2}} r^d}{\Gamma\left(\frac{d}{2}+1\right)}.$$

We consider the path sets Γ^1 and Γ^2. We obtain for the corresponding weights on the paths:

$$|||\gamma^1_{xy}|||_p = N \left(\frac{\exp\left(-\frac{2r^2}{N^2}\right)}{\mathrm{Vol}(B_d(r))(2\pi)^{\frac{d}{2}}} \right)^{-p},$$

$$|||\gamma_{xy}^2|||_p \quad = \quad N(x,y)\left(\frac{-\left(\frac{\|x-y\|}{N(x,y)}\right)^2}{\mathrm{Vol}(B_d(r))(2\pi)^{\frac{d}{2}}}\right)^{-p}$$

Firstly, we estimate the constant appearing in the Poincaré inequality, Theorem 2.3.1 from above. For a path of length N we take into account that two consecutive vertices v and w satisfy $\|v-w\| \le r/N$. Therefore,

$$\kappa^{\Gamma^1}(p) \le \frac{N^{d+2}(2\pi)^{\frac{d}{2}}}{\mathrm{Vol}(B_d(r))}\exp\left(\frac{2r^2}{N^2}\right).$$

This implies a lower bound on the spectral gap which is

$$\lambda_1(K) \ge \frac{\mathrm{Vol}(B_d(r))}{N^{d+2}(2\pi)^{\frac{d}{2}}}\exp\left(-\frac{2r^2}{N^2}\right).$$

If we regard the right hand sight as a function of N, we can solve for N_0 with $f(N_0) = \max_N f(N)$.
This gives $N_0 = \sqrt{\frac{2r^2}{d+2}}$ and we can take $N_{max} := \left\lceil \sqrt{\frac{2}{d+2}}r \right\rceil$ to obtain

$$\lambda_1 \ge \frac{\mathrm{Vol}(B_d(r))}{N_{max}^{d+2}(2\pi)^{\frac{d}{2}}}\exp\left(-\frac{2r^2}{N_{max}^2}\right) \approx \frac{\mathrm{Vol}(B_d(r))\sqrt{d+2}^{d+2}}{\sqrt{2}^{d+2}r^{d+2}(2\pi)^{d+2}}\exp(-d+2).$$

The same considerations for the path set Γ^2 and the fact that the length of the longest path can be at most $\lceil 2r \rceil$ lead to

$$\begin{aligned}\kappa^{\Gamma^2}(p) \quad &\le \quad \sum_{n=1}^{\lceil 2r \rceil}\sum_{m=1}^{n}\frac{(2\pi)^{d/2}}{\mathrm{Vol}(B_d(r))}n^{d+1}\exp\left(\frac{2r^2}{n^2}\right) \\ &\le \quad \frac{(2\pi)^{d/2}}{\mathrm{Vol}(B_d(r))}\lceil 2r \rceil^{d+2}\exp\left(\frac{2r^2}{\lceil 2r \rceil^2}\right).\end{aligned}$$

In the sequel we will restrict ourselves to the case $d = 1$ and consider the symmetric interval $I = [-r, r]$. For

$$k(x,y) := \begin{cases} \frac{\exp\left(-\frac{|x-y|^2}{2}\right)}{(2r)} & , \quad y \in I, \\ \qquad 0 & , \quad \text{otherwise}, \end{cases}$$

we consider the Gibbs sampler used in Markov chain Monte Carlo simulation. In the language of 'Metropolis algorithms' the acceptance probabilities are given by

$$\alpha_{xy} = \begin{cases} \min\{1, \exp(-y^2/2), \exp(-x^2/2)\} & \text{for} \quad |y| \le r \\ 0 & \text{for} \quad |y| > r \end{cases}$$

The Markov transition kernel for the Gibbs sampler is

$$K_G(x,dy) = \alpha_x \delta_x(dy) + \alpha_{xy} k(x,y)dy = \alpha_x \delta_x(dy) + f(x,y)dy,$$

$$\text{where } f(x,y) = \begin{cases} \frac{1}{\sqrt{2\pi}} \exp\left(-\max \frac{\{((y-x)^2, 2y^2-2yx\}}{2}\right) & , \quad |y| \leq r, \\ 0 & , \quad \text{otherwise} \end{cases}$$

and α_x is defined as in Remark 2.0.1. The chain has stationary measure

$$\pi(dx) = \frac{1}{2r} 1_I(x) dx.$$

Choosing the path set Γ^1 and $p = 1$ we obtain:

$$\begin{aligned} |||\gamma^1_{xy}|||_1 &= \sum_{i=1}^{N} \left(\frac{1}{2r\sqrt{2\pi}} \exp\left(-\max \frac{\{(y-x)^2, 2y^2 - 2yx\}}{2}\right)\right)^{-1} \\ &= N2r\sqrt{2\pi} \exp\left(\max \frac{\{(y-x)^2, 2y^2 - 2yx\}}{2}\right). \end{aligned}$$

Now we apply $|v - w| \leq r$. By definition of $\kappa^\Gamma(1)$ this leads to

$$\kappa^\Gamma(1) \leq \frac{N^3\sqrt{2\pi}}{2r} \exp\left(\frac{r^2}{2N^2}\right).$$

Minimizing with respect to N yields

$$\lambda_{K_G} \geq \frac{12\sqrt{6}}{r^2\sqrt{2\pi}} \exp(-3).$$

In [Ro 96], Rosenthal obtains

$$\lambda_{K_G} \geq \frac{(1 - e^{-1/2})^8}{8(6e^2 r)^2}.$$

Our bound is of the same order as Rosenthal's but the constant factor is improved by a factor 10^7.

Example 4: Submanifolds

In certain cases it is possible to use the path sets Γ^i, $i = 1, 2, 3, 4$, to obtain estimates in a more general setting, for instance submanifolds.
Suppose $M \subset \mathbb{R}^d$, $d \in \mathbb{N}$, is a submanifold and $dim(M) = k \leq d$. Furthermore, we suppose that there exist local coordinates , i.e. a map $\varphi : V \subset$

$\mathbb{R}^k \to M$, such that $\varphi^{-1}(M) = V$ and V is convex.
We can define a path set, Γ^1_φ, obeying the p.r.p. if $\varphi^{-1}(M)$ is convex. Then, let $N \in \mathbb{N}$ and take

$$\Gamma^1_\varphi = \left\{\gamma^{1,\varphi}_{xy}\right\}, \tag{2.49}$$

$$\gamma^{1,\varphi}_{xy}(n) = \varphi\left(\frac{(N-n)\varphi^{-1}(x) + n\varphi^{-1}(y)}{N}\right), \quad n \in \{0, .., N\}. \tag{2.50}$$

Here we use local coordinates to lift the path set Γ^1 to the submanifold.
Let us now consider the path set obtained using local coordinates. We consider the map

$$G^M = (\varphi^{-1}, \varphi^{-1}, id_{\mathbb{N}}, id_{\mathbb{N}}) \circ G \circ (\varphi, \varphi, id_{\mathbb{N}}), \text{ i. e.}$$

$$\begin{aligned} G^M : E^M_1 &\to M \times M \times \mathbb{N} \times \mathbb{N} \\ (x, y, n) &\mapsto (\gamma^\varphi_{xy}(n-1), \gamma^\varphi_{xy}(n), n, N(x, y)). \end{aligned}$$

To simplify notation we use $\tilde{\varphi} := \begin{pmatrix} \varphi \\ \varphi \end{pmatrix}$ and $\tilde{\varphi}^{-1} := \begin{pmatrix} \varphi^{-1} \\ \varphi^{-1} \end{pmatrix}$.
We suppose for fixed index m and length n that the map $G^{-1}(\cdot, \cdot, m, n)$ is differentiable. Fixing m and n, this leads to

$$\begin{aligned} D(G^M)^{-1} &= D\left(\tilde{\varphi} \circ G \circ \tilde{\varphi}^{-1}\right)^{-1} = D\tilde{\varphi}^{-1} \circ G^{-1} \circ \tilde{\varphi} \\ &= D\tilde{\varphi}^{-1} \circ G^{-1}(\tilde{\varphi}) D\tilde{\varphi} = \tilde{\varphi}^{-1}(G^{-1}(\varphi)) DG^{-1}(\tilde{\varphi}) D\tilde{\varphi}. \end{aligned}$$

Here we used the chain rule for differentiable functions. Therefore, we have

$$\det DG^{M^{-1}} = \det G^{-1}. \tag{2.51}$$

Homothetien and Translations

Using the latter computations we are able to study the corresponding density function h_1 for homothetien and translations. Since we deal with Lebesgue measure we expect invariance of h_1 with respect to homothetien and translations.
Let $r \in \mathbb{R}$ and consider the map Φ_r and its inverse Φ_r^{-1} which clearly exists, given by

$$\begin{aligned} \Phi_r : \mathbb{R}^d &\to \mathbb{R}^d & \qquad \Phi_r^{-1} : \mathbb{R}^d &\to \mathbb{R}^d \\ x &\mapsto rx & x &\mapsto x/r. \end{aligned}$$

Furthermore, we for $a \in \mathbb{R}^d$ we consider the map T_a and its inverse given by

$$\begin{aligned} T_a : \mathbb{R}^d &\to \mathbb{R}^d & \qquad T_a^{-1} : \mathbb{R}^d &\to \mathbb{R}^d \\ x &\mapsto x + a & x &\mapsto x - a. \end{aligned}$$

Using (2.51) the density h_1 for subsets of $\mathbb{R}^d$ and path sets Γ^1, Γ^2 are invariant with respect to homothetien Φ_r and translations T_a.

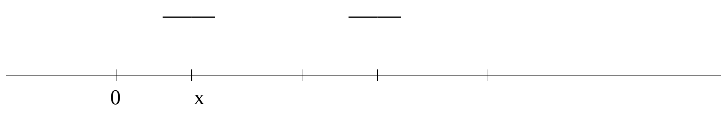

Figure 2.4: **The state space** $C(0,r)$

The Circle

Firstly, we take the circle $C(0,1)$ with radius 1 centered at 0.
We consider the transition kernel

$$K(x,dy) = \frac{1}{2\epsilon} 1_{I_x^\epsilon}(y), \quad 0<\epsilon<\pi, \quad I_x^\epsilon := \{z \in C(0,1) : d(x,z) < \epsilon\}.$$

Here dy denotes the Lebesgue measure on the circle, i.e. $dy(C(0,1)) = 2\pi$.
Then, we have as the stationary measure $\pi(dx) = \frac{1}{2\pi} dx$.
We take as a path set Γ^C the shortest geodesics connecting two points $x, y \in C(0,1)$. These are nothing but the straight line segments connecting two points using local coordinates.
Take a path $\gamma_{xy} \in \Gamma^C$. We have

$$|||\gamma_{xy}^1|||_p = \sum_{n=1}^{N} Q(\gamma_{xy}^1(n-1), \gamma_{xy}^1(n))^{-p} = \left(4\epsilon\pi \sum_{j=1}^{N} 1_{\{\frac{\pi}{N}<\epsilon\}}\right)^p = \left(4\epsilon\pi N 1_{\{\frac{\pi}{N}\le\epsilon\}}\right)^p.$$

Hence,

$$\kappa^{\Gamma^C} = \left\{ \underset{(v,w)\,\in\,\mathcal{A}^\Gamma}{\mathrm{esssup}} \sum_{m,n\in\mathbb{N}} |||\gamma_{xy}^1|||_p h(x)h(y)h_1(v,w,m,n) 1_{E_2} Q(v,w)^{p-1} \right\} \le N^3 \frac{\epsilon}{\pi}.$$

Since admissibility is assumed, i.e. $\frac{\pi}{N} \le \epsilon$ we find

$$\lambda_1 \ge \frac{\pi}{\epsilon}\frac{1}{N} \approx \frac{\epsilon^2}{\pi^2}.$$

We can explicitely calculate the eigenvalues corresponding to the operator under consideration. We take as a basis of orthonormal functions for the circle the system $\left(1, \left(\frac{\cos(kx)}{\sqrt{2\pi}}\right)_{k\in\mathbb{N}}, \left(\frac{\sin(kx)}{\sqrt{2\pi}}\right)_{k\in\mathbb{N}^+}\right)$. It follows that

$$\begin{aligned}
\int_C k(x,y)\frac{\cos(ky)}{\sqrt{2\pi}} dy &= \frac{1}{\sqrt{2\pi}}\frac{1}{2\epsilon}\int_0^{2\pi} 1_{U_\epsilon(x)}(y)\cos(ky)dy \\
&= \frac{1}{\sqrt{2\pi}}\frac{1}{2\epsilon}\int_{x-\epsilon}^{x+\epsilon} \cos(ky)dy
\end{aligned}$$

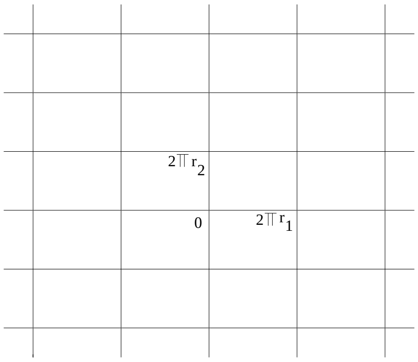

Figure 2.5: **The state space T**

$$
\begin{aligned}
&= \frac{1}{\sqrt{2\pi}} \frac{1}{2\epsilon} \frac{1}{k} \sin(ky)|_{x-\epsilon}^{x+\epsilon} \\
&= \frac{1}{\sqrt{2\pi}} \frac{1}{2k\epsilon} \left(\sin(k(x+\epsilon)) - \sin(k(x-\epsilon))\right) \\
&= \frac{\sin(k\epsilon)}{k\epsilon} \frac{\cos(kx)}{\sqrt{2\pi}}.
\end{aligned}
$$

The same computation using the functions 1 and $\frac{\sin(kx)}{\sqrt{2\pi}}$ shows that the eigenvalues of L are given by

$$0, \left(1 - \frac{\sin(k\epsilon)}{k\epsilon}\right)_{k \in \mathbb{N}^+}.$$

This compares well with our estimate because $\lambda_1 \approx \frac{\epsilon^2}{6}$.
Now, take the circle $C(0, r)$ with radius $r \in \mathbb{R}^+$ centerd a 0. We consider the transition kernel

$$K(x, dy) = \frac{1}{2\epsilon} 1_{I_x^\epsilon}(y), \quad 0 < \epsilon < \pi, \quad I_x^\epsilon := \{z \in C(0, r) : d(x, z) < \epsilon\}.$$

Similar computations as in the case $r = 1$ yield

$$\lambda_1(C(0, r)) \geq \left(\frac{\epsilon}{\pi r}\right)^2.$$

They are off by a factor $\approx 1/3$.

The Torus

Let $\epsilon, r_1, r_2 \in \mathbb{R}^+$ and assume $r_1 \geq r_2$ and $\pi r_2/2 \geq \epsilon$. The torus T is the two dimensional submanifold obtained by rotating a circle of radius r_2 around a

line lying in the same plane as the circle and having distance r_1 from the circle's midpoint.
We have $\mathrm{Vol}(T) = 4\pi^2 r_1 r_2$. If we take $k(x,y) = \frac{1}{4\epsilon^2} 1_{U_\epsilon(x)}(y)$, then the stationary measure is $d\pi(x) = \frac{1}{\mathrm{Vol}(T)} dx$. We choose as a path set Γ^T the shortest geodesics γ_{xy} connecting two points $x, y \in T, x \neq y$. A path γ_{xy} corresponds to the image of the shortest straight line connecting $\varphi(x)$ and $\varphi(y) + \begin{pmatrix} kr_1 \\ lr_2 \end{pmatrix}$, $k, l \in \mathbb{N}$ in local coordinates. See Figure 2.5.
If $L(\gamma)$ denotes the length of a path γ, we set $L = \sup_{\gamma \in \Gamma^T} L(\gamma)$. For an upper bound on L consider the midpoint x_0 of a rectangle $[2\pi k r_1, 2\pi(k+1)r_1) \times [2\pi l r_2, 2\pi(l+1)r_2)$ and one of the four corner points. Owing to the assumption $r_1 \geq r_2$ we obtain

$$L = \sqrt{(\pi r_1)^2 + (\pi r_2)^2} \leq \sqrt{2}\pi r_1.$$

Now, we are able to state our estimates on the spectral gap λ_1.
Firstly, we calculate $|||\cdot|||_p$ for a path γ^1_{xy} connecting $x, y \in T$ and find

$$\begin{aligned} |||\gamma^1_{xy}|||_p &= \sum_{n=1}^{N} Q(\gamma^1_{xy}(n-1), \gamma^1_{xy}(n))^{-p} = N(16\epsilon^2\pi^2 r_1 r_2)^p \\ |||\gamma^2_{xy}|||_p &= N(x,y)(16\epsilon^2\pi^2 r_1 r_2)^p. \end{aligned}$$

For the constant $\kappa^{\Gamma^{T,1}}$ we obtain

$$\kappa^{\Gamma^{T,1}} = N^2 \sum_{m=1}^{N} \frac{N}{16\pi^2 r_1^2 r_2^2} 16\epsilon^2\pi^2 r_1 r_2 = N^4 \frac{\epsilon^2}{\pi^2 r_1 r_2}.$$

Here we used the fact that all paths have length N and in each path there are edges of index 1 up to N.
For the constant $B^{\Gamma^{T,1}}$, appearing in the estimate of I_K,we get

$$B^{\Gamma^{T,1}} = N^3 \frac{\epsilon^2}{\pi^2 r_1 r_2}.$$

Choose N such that $\frac{\pi\sqrt{r_1^2 + r_2^2}}{\epsilon} \leq N$. Now, taking $N_0 \geq \left\lceil \frac{\sqrt{2}\pi r_1}{\epsilon} \right\rceil$ would suffice to ensure admissibility. This leads to

$$\lambda_1 \geq \frac{\pi^2 r_1 r_2}{\epsilon^2 N_0^4} \approx \frac{\epsilon^2 r_2}{4\pi^2 r_1^3} \text{ and } \lambda_1 \geq \frac{\pi^2 r_1 r_2}{\epsilon^2 N_0^3} \approx \frac{\epsilon^2 r_2^2}{64\pi^2 r_1^4}, \text{ respectively.}$$

To compute the constant $\kappa^{\Gamma^{T,2}}$ for the path set Γ^2 we can use that the length of the longest path in T is $\lceil \eta L \rceil$. This leads to

$$\kappa^{\Gamma^{T,2}} = \frac{\epsilon^2}{\pi r_1 r_2} \sum_{n=1}^{\lceil \eta L \rceil} n^4 = \frac{\epsilon^2}{30\pi r_1 r_2} \lceil \eta L \rceil (2\lceil \eta L \rceil + 1)(\lceil \eta L \rceil + 1)(3\lceil \eta L \rceil^2 + 3\lceil \eta L \rceil - 1).$$

and

$$B^{\Gamma^{T,2}} = \frac{\epsilon^2}{\pi r_1 r_2} \sum_{n=1}^{\lceil \eta L \rceil} n^3 = \frac{\epsilon^2}{\pi r_1 r_2} \frac{\lceil \eta L \rceil^2 (\lceil \eta L \rceil + 1)^2}{4}$$

Taking the admissibility into account, i.e. we require $\eta \geq 1/\epsilon$, we obtain

$$\begin{aligned} \lambda_1 &\geq \frac{r_2}{30\epsilon\sqrt{2}(2\sqrt{2}\pi r_1/\epsilon + 1)(\sqrt{2}\pi r_1/\epsilon + 1)(3\sqrt{2}\pi r_1/\epsilon + 3\sqrt{2}\pi r_1/\epsilon - 1)} \\ &\approx \frac{\epsilon^3 r_2}{\pi^4 r_1^4} \end{aligned}$$

and

$$\lambda_1 \geq \left(\frac{2r_2}{\pi(\sqrt{2}\pi r_1/\epsilon + 1)^2} \right)^2 /8 \approx \frac{\epsilon^4 r_2}{4\pi^4 r_1^3}.$$

To compute the spectrum of L explicitely we take an orthonormal system of functions on T: $\frac{1}{2\pi}$, $\left(\frac{\cos(xk/r_1)\cos(yl/r_2)}{2\pi}\right)_{k,l\in\mathbb{N}}$, $\left(\frac{\cos(xk/r_1)\sin(yl/r_2)}{2\pi}\right)_{k\in\mathbb{N},l\in\mathbb{N}^+}$, $\left(\frac{\sin(xk/r_1)\cos(yl/r_2)}{2\pi}\right)_{k\in\mathbb{N}^+,l\in\mathbb{N}}$, $\left(\frac{\sin(xk/r_1)\sin(yl/r_1)}{2\pi}\right)_{k,l\in\mathbb{N}^+}$. Proceeding as in the previous example, we obtain for the spectrum of the generator: 0, $\left(1-\frac{r_1\sin(\epsilon k/r_1)}{k\epsilon}\right)_{k\in\mathbb{N}^+}$, $\left(\frac{r_2\sin(\epsilon k/r_2)}{k\epsilon}\right)_{k\in\mathbb{N}^+}$, $\left(1-r_1\frac{\sin(\epsilon k/r_1)r_2\sin(\epsilon l/r_2)}{kl\epsilon^2}\right)_{k,l\in\mathbb{N}^+}$. Hence, $\lambda_1 \approx \frac{\epsilon^2}{6r_1^2}$. Applying our methods the lower bounds are off by a factor $\approx 1/12 \cdot r_2/r_1$. The factor r_2/r_1 is the consequence of the estimate $\sqrt{r_1^2+r_2^2} \leq \sqrt{2}r_1$. For r_2 being close to r_1 the estimate is of the right order. One can think of tori with a 'small hole'. Whereas if r_2 is much smaller than r_1 the bound becomes worse. This is the case for tori having a very 'big hole'.

Thetasums

We consider the circle $C(0,1)$ again but now we take the transition kernel

$$K(x,dy) = \frac{1}{\sqrt{2\pi}} \sum_{k\in\mathbb{Z}} \exp\left(-\frac{|2\pi k + |x-y||^2}{2}\right) dy.$$

Let us proceed by using the path set Γ^1 and computing the weights $|||\cdot|||_p$ for a path $\gamma^1_{xy} \in \Gamma^1$.

$$\begin{aligned} |||\gamma^1_{xy}|||_p &= \sum_{n=1}^{N} Q(\gamma^1_{xy}(n-1), \gamma^1_{xy}(n))^{-p} = \sum_{n=1}^{N} \frac{(2\pi)^{3p/2}}{\left(\sum_{k\in\mathbb{Z}} \exp\left(-\frac{|2\pi k+|x-y||^2}{2}\right)\right)^p} \\ &\leq N \frac{(2\pi)^{3p/2}}{\left(\sum_{k\in\mathbb{Z}} \exp\left(-\frac{(2\pi(k+1/N))^2}{2}\right)\right)^p}. \end{aligned}$$

We take $p = 1$. The constant $\kappa^1(1)$ can be estimated as follows:

$$\kappa^1(1) \leq \frac{N^3}{\sum_{k\in\mathbb{Z}} \exp\left(-\frac{(k+1/N)^2 4\pi^2}{2}\right)\sqrt{2\pi}}.$$

The estimate for $\kappa^1(p)$ leads to

$$\lambda_1 \geq \frac{\sum_{k\in\mathbb{Z}} \exp\left(-\frac{(2\pi(k+1/N))^2}{2}\right)\sqrt{2\pi}}{N^3}.$$

If we take the summand $\exp(-(2\pi(k+1/N))^2/2)\sqrt{2\pi}/N^3$ it is maximized for $N = 2\pi/3(\pi k + \sqrt{3})$. Since the summands are monotone decreasing we maximize with respect to the summand corresponding to $k = 0$. This yields $N = 4$ and we find

$$\lambda_1 \geq \frac{\sum_{k\in\mathbb{Z}} \exp\left(-\frac{(2\pi(k+1/4)^2}{2}\right)\sqrt{2\pi}}{4^3} \approx 0.023.$$

As in the two preceeding examples we can calculate the spectrum explicitely. We consider the orthonormal system $(\frac{\exp(i2\pi lx)}{\sqrt{2\pi}})_{l\in\mathbb{Z}}$.

$$\begin{aligned}
&\frac{1}{\sqrt{2\pi}}\int_0^{2\pi}\sum_{k\in\mathbb{Z}}\exp\left(-\frac{|2\pi k + |x-y||^2}{2}\right)\exp(i2\pi ly)dy\\
&= \frac{1}{\sqrt{2\pi}}\sum_{k\in\mathbb{Z}}\int_0^{2\pi}\exp\left(-\frac{|2\pi k + |x-y||^2}{2}\right)\exp(i2\pi ly)dy\\
&= \frac{\exp(i2\pi lx)}{\sqrt{2\pi}}\sum_{k\in\mathbb{Z}}\int_{2\pi k - x}^{2\pi(k+1)-x}\exp\left(-\frac{z^2}{2}\right)\exp(i2\pi(-l)z)dz\\
&= \frac{\exp(i2\pi lx)}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\exp\left(-\frac{z^2}{2}\right)\exp(i2\pi(-l)z)dz\\
&= \exp\left(-\frac{l^2}{2}\right)\frac{\exp(i2\pi lx)}{\sqrt{2\pi}}.
\end{aligned}$$

Therefore, the spectrum are the numbers $\left(1 - \exp\left(-\frac{l^2}{2}\right)\right)_{l\in\mathbb{Z}}$. All eigenvalues, except the case $l = 0$, have multiplicity 2. Especially, we find

$$\lambda_1 = 1 - \exp(-1/2) \approx 0.394.$$

Thus, our bound is far from being of the right order of magnitude!
The latter can be generalized to circles of radius r and tori. For instance we find for the spectrum in the case of a circle with radius r:

$$\left(1 - \exp\left(-\frac{l^2}{2r^2}\right)\right)_{l\in\mathbb{Z}}.$$

2.8 Log-Sobolev - Isoperimetry

As mentioned in the introduction there is a close relation between the analytic inequalities and isoperimetric notions. The well known Cheeger inequality for the spectral gap determines an interval I, namely $I = [I_K^2/4, I_K]$, such that $\lambda_1 \in I$. in the sequel we want to deduce an analogues inequality which we call *Logarithmic-Cheeger inequality.*
As a consequence of the proof we find lower bounds on the logarithmic-Sobolev constant in terms of $I_K^{\ln}$, introduced in (1.24), and λ_1 which we call mixed isoperimetric bounds. They are summarized in Remark 2.8.3.
To proceed we observe that Theorem 2.3.3 of [R 85] also holds true in our setting and we have for the discrete gradient $[\nabla_K f](x,y) = f(x) - f(y)$,

Theorem 2.8.1 *Let $c \geq 0$. Then the following are equivalent*
(i) $c\mathcal{L}_\pi[f] \leq \int_M |\nabla_K f|\, d\pi(x)$, for all $f \in \mathcal{B}(S)$
(ii) $c\max\{\mathcal{L}_\pi[1_A], \mathcal{L}_\pi[-1_A]\} \leq Q(\partial A)$ for $A \subset S$, A measurable, $\pi(A) \leq \frac{1}{2}$.

The proof is essentially the same as in [R 85], Theorem 2.3.3 and therefore omitted.
Let us shortly motivate how one is led to study I_K^{ln}.
Setting $0\ln(0) := 0$, for finding the best possible constant $c > 0$ in Theorem 2.8.1, (i), for a measurable subset $A \subset S$ we only need to consider the indicator functions 1_A. We have

$$-c\pi(A)\ln(\pi(A)) = c\mathcal{L}_\pi[1_A] \leq \mathbb{E}_\pi[\int_S |1_A(y) - 1_A(x)|K(x,dy)] = Q(\partial A).$$

Thus, we are led to consider the isoperimetric constant

$$I_K^{ln} = \inf_{A\subset S} -\frac{Q(\partial A)}{\pi(A)\ln(\pi(A))}. \tag{2.52}$$

Here the infimum is taken over all measurable sets $A \subset S$.

Theorem 2.8.2 *Let $(X_t)_t$ be a reversible Sub-Markov chain in discrete time or a time continuous jump Sub-Markov process on a measure space $(S, \mathcal{S}, \mu)$. Assume the transition kernel of the chain is given by (2.1).*
For the logarithmic-Sobolev constant we have

$$\frac{\ln(2)}{4\sqrt{2}}(I_K^{\ln})^2 \leq \alpha \leq I_K^{\ln}. \tag{2.53}$$

Proof:
Take a function $f \in D(\mathcal{E})$, s.t. $\mathcal{L}_\pi[f] < \infty$, applying the triangle inequality we have

$$
\begin{aligned}
&I_K^{\ln}\mathcal{L}_\pi[f]\\
&\leq \mathbb{E}_\pi[\int_S |f^2(y)-f^2(x)|K(x,dy)]\\
&\leq \mathbb{E}_\pi[\int_S |f(y)-f(x)|^2 K(x,dy)] + 2\mathbb{E}_\pi[|f(x)|\int_S |f(y)-f(x)|K(x,dy)].
\end{aligned}
$$

Applying the Cauchy-Schwarz inequality as well as Jensen's inequality yields

$$
\begin{aligned}
&I_K^{\ln}\mathcal{L}_\pi[(f-\mathbb{E}_\pi[f])]\\
&\leq \mathbb{E}_\pi[\int_S |f(y)-\mathbb{E}_\pi[f]-f(x)+\mathbb{E}_\pi[f]|^2 K(x,dy)]\\
&\quad +2\mathbb{E}_\pi[|f-\mathbb{E}_\pi[f]|\int_S |f(y)-\mathbb{E}_\pi[f]-f(x)+\mathbb{E}_\pi[f]|K(x,dy)]\\
&\leq \mathcal{E}(f,f)+2\sqrt{Var[f]}\sqrt{\mathbb{E}_\pi[\int_S |f(y)-f(x)|K(x,dy)]^2}\\
&\leq \mathcal{E}(f,f)+\frac{2}{\sqrt{\lambda_1}}\mathbb{E}_\pi[\int_S |f(y)-f(x)|^2 K(x,dy)]\\
&= \left(1+\frac{2}{\sqrt{\lambda_1}}\right)\mathcal{E}(f,f)
\end{aligned}
$$

Using an inequality due to Rothaus, [R 85], Lemma 9, i.e.

$$\mathcal{L}_\pi[f] \leq \mathcal{L}_\pi[f-\mathbb{E}_\pi[f]] + 2Var[f],$$

and because the chain is reversible we have

$$\mathbb{E}_\pi[\int_S |f^2(y)-f^2(x)|K(x,dy)] \leq 2\mathbb{E}_\pi[f(x)\int_S |f(y)-f(x)|K(x,dy)].$$

Hence, we obtain

$$2\alpha \geq \frac{\lambda_1 I_K^{\ln}}{\sqrt{\lambda_1}+I_K^{\ln}}. \tag{2.54}$$

Using $x\ln(2) \leq -x\ln(x)$, $x \in (0,1/2)$, it follows that $\lambda_1 \geq \frac{\ln(2)}{8}(I_K^{\ln})^2$. Together with the previous considerations this leads to

$$\alpha \geq \frac{\ln(2)}{4\sqrt{2}}(I_K^{\ln})^2.$$

The upper bound follows immediately by the definiton of $I_K^{\ln}$, the definition of the logarithmic-Sobolev constant and the variational characterization of

$\mathcal{L}_\pi$, see (1.17). This proves the Theorem. □

For general Markov chains we define the spectral gap by

$$\lambda_1 := \min{}^+ \{|\lambda| : \lambda \text{ eigenvalue of } L\}, \tag{2.55}$$

where $\min^+ = \min(A\backslash\{0\})$ denotes the minimum only over the positive elements for the set A. This conicides with our definition for λ_1 for reversible chains.

Remark 2.8.3 *From the latter proof we find the mixed isoperimetric bound*

$$\alpha \geq \begin{cases} \frac{\lambda_1 I_K^{\ln}}{2\left(\sqrt{\lambda_1}+I_K^{\ln}\right)+\lambda_1} & : \; K \textit{ is non-reversible,} \\ \frac{\lambda_1 I_K^{\ln}}{2(\sqrt{\lambda_1}+I_K^{\ln})} & : \; K \textit{ is reversible.} \end{cases} \tag{2.56}$$

Proof:

It remains to show the bound for the non reversible case. As in the proof of the preceding theorem we use Lemma 9 of [R 85] and obtain

$$\alpha \geq \frac{\lambda_1 I_K^{\ln}}{2\left(\sqrt{\lambda_1}+I_K^{\ln}\right)+\lambda_1}. \tag{2.57}$$

This complements the proof of the statements in Remark 2.8.3. □

Remark 2.8.4 *Since the right hand side is a nondecreasing function in λ_1, every lower bound on λ_1 leads to a lower bound on α. Therefore, we can use our path-based lower bounds in the reversible case to estimate the logarithmic-Sobolev constant α.*

Chapter 3

Estimating the Logarithmic-Sobolev constant for Markov chains on finite sets

From now on we consider irreducible and aperiodic Markov chains having a finite state space $\mathcal{X}$. After identifying the elements of $\mathcal{X}$ with $\{0, ..., |\mathcal{X}|-1\}$, the transition kernel is given by an $|\mathcal{X}| \times |\mathcal{X}|$-matrix $K = (p_{ij})_{i,j \in \{0,...,|\mathcal{X}|-1\}}$. The entries p_{ij} are the one-step probabilities.

In Section 3.1 we define operators ∇ and ∇^*. They are discrete analogues to the gradient and the divergence operator. We use these operators to define discrete Dirichlet forms corresponding to the Markov chains under consideration.

In 3.2 we recall the definition of coupling of Markov chains. In his 1938 paper, [Do 38], Doeblin proved the asymptotic stationarity of irreducible and aperiodic Markov chains on finite state spaces. More precisely, he considered the classical coupling. Given the original Markov chain $(X_n)_n$ let $(\tilde{X}_n)_n$ be a differently started version of $(X_n)_n$ whose transitions are independent of the transitions of $(X_n)_n$ unless both chains coincide.

The first time they coincide is referred to as the coupling time T of $(X_n)_n$ and $(\tilde{X}_n)_n$. After that time both chains keep together. Of fundamental importance for proving limit results is the coupling inequality

$$\|P(X_n \in \cdot) - P(\tilde{X}_n \in \cdot)\|_{TV} \leq 2P(T > n). \tag{3.1}$$

Aldous, Diaconis and Meise used coupling techniques to estimate the rate of convergence, see [AD 87], [Me 96], [Me 96], [Me97] and [Me 99].

In Section 3.3 the notion of a logarithmic-Sobolev function is reviewed. These functions fulfill a certain difference equation.

We will apply coupling methods in Section 3.4 to deduce lower bounds on

the logarithmic-Sobolev constant α of a Markov chain. However, it turns out that these lower bounds do not lead to very sharp estimates.
Section 3.5 contains several examples and there we compare our lower bounds with the best known bounds or the exact values if they are known.
In the final section of this chapter we apply Theorem 2.8.2, the logarithmic-Cheeger inequality, as well as Remark 2.8.3, the mixed isoperimetric bound, to estimate α. Our examples show that these bounds can be of the right order of magnitude. For the finite setting a similar bound as in Remark 2.8.3 has been obtained in [H 00]. But there no examples have been given.

3.1 Calculus for finite Markov chains

Let $\mathcal{X}$ denote a finite set and suppose that K is an irreducible, aperiodic transition kernel. It is well known that in this case there is a uniquely determined stationary measure π corresponding to K that charges every point.
We call the transition kernel K reversible if it satisfies the detailed balance equation, i.e. $K(x,y)\pi(x) = K(y,x)\pi(y)$. For notational convenience we consider a measure on $\mathcal{X} \times \mathcal{X}$ given by

$$Q(x,y) := \begin{cases} K(x,y)\pi(x) & : \ K \text{ is reversible,} \\ \frac{1}{2}(K(x,y)\pi(x) + K(y,x)\pi(y)) & : \ \text{otherwise.} \end{cases}$$

Furthermore, let us introduce the L^2-spaces $L^2(\mathcal{X},\pi)$ and $L^2(\mathcal{X}\times\mathcal{X},Q)$, i.e. the equivalence classes of

$$l^2(\mathcal{X},\pi) := \{f : \mathcal{X} \to \mathbb{R} | \|f\|_{2,\pi} < +\infty\}$$

and

$$l^2(\mathcal{X}^2,Q) := \left\{f : \mathcal{X}^2 \to \mathbb{R} | f(x,y) = -f(y,x), \|f\|_{2,Q} < +\infty\right\}.$$

Definition 3.1.1 *We introduce the following operators*

$$\begin{aligned} \nabla : L^2(\mathcal{X},\pi) &\to L^2(\mathcal{X}\times\mathcal{X},Q) && (3.2) \\ [\nabla f](x,y) &= f(x) - f(y) && (3.3) \end{aligned}$$

and

$$\begin{aligned} \nabla^* : L^2(\mathcal{X}\times\mathcal{X},Q) &\to L^2(\mathcal{X},\pi) && (3.4) \\ [\nabla^*]f(x) &= \sum_{y\in\mathcal{X}} f(x,y)K(x,y). && (3.5) \end{aligned}$$

Finally, consider

$$
\begin{aligned}
L : L^2(\mathcal{X},\pi) \quad &\to \quad L^2(\mathcal{X},\pi) && (3.6)\\
[Lf](x) := [\nabla^*\nabla f](x) \quad &= \quad \sum_{y\in\mathcal{X}} (f(x)-f(y))\,K(x,y). && (3.7)
\end{aligned}
$$

L is called the discrete generator of the Markov chain.

Remark 3.1.2 *1) The L^2-space $L^2(\mathcal{X}\times\mathcal{X},Q)$ can be seen as an discrete analogue to the space of 1-forms and the operators correspond to the gradient, the divergence and the Laplacian.*
2) Alternatively, by considering the associated linear operator to the transition kernel $[K\cdot]$ on the Hilbert space $l^2(\mathcal{X},\pi)$ given by

$$
[Kf](x) := \sum_{y\in\mathcal{X}} K(x,y)f(y), \quad f\in l^2(\mathcal{X},\pi), \tag{3.8}
$$

we could have introduced L as the operator $Id-K$.

There is a one-to-one correspondance between the eigenvalues of K and the eigenvalues of L, namely β is an eigenvalue of K if and only if $1-\beta$ is an eigenvalue of L. In the case K is time reversible, we call the first positive eigenvalue of L the spectral gap.

Definition 3.1.3 *The quadratic form,*

$$
\mathcal{E}(f,g) := \langle [Lf], g\rangle_2, \quad f,g\in L^2(\mathcal{X},\pi), \tag{3.9}
$$

is called the (discrete) Dirichlet form associated to K. We have the explicit representation on the diagonal:

$$
\mathcal{E}(f,f) = \frac{1}{2}\sum_{x,y\in\mathcal{X}} |f(y)-f(x)|^2 K(x,y)\pi(x). \tag{3.10}
$$

Definition 3.1.4 *Given an irreducible transition kernel K, we define the continuous time semigroup $(H_t)_t$ associated with K by*

$$
H_t f(x) := \exp(-t[Lf](x)) := \exp(-t)\sum_{n=0}^{\infty} \frac{t^n[K^n f](x)}{n!}. \tag{3.11}
$$

Obviously,

$$
H_t(x,y) = \exp(-t)\sum_{n=0}^{\infty}\frac{t^n K^n(x,y)}{n!}.
$$

Remark 3.1.5 *1) Let $(X_t)_{t\in\mathbb{R}^+}$ be the time continuous Markov process associate to K. We have $P_x(X_t = y) = H_t(x,y)$, i.e. $H_t(x,y)$ is the probability for the time continuous process started at $x \in \mathcal{X}$ to be at $y \in \mathcal{X}$ at time $t > 0$. It is called the heat semigroup. The following is well known:*

$$\frac{\partial}{\partial t}\|H_t f\|_2^2 = -2\mathcal{E}(H_t f, H_t f), \tag{3.12}$$

$$\mathcal{E}(f,f) = -\partial_t \|H_t f\|_2^2\Big|_{t=0} = \lim_{t\to 0} -\frac{1}{2t}\langle [(I-H_t)f], f\rangle_2. \tag{3.13}$$

2) If we work with the heat semigroup the assumption of aperiodicity is not needed anymore.

3.2 Coupling

Recall the definition of a coupling:

Definition 3.2.1 *(Coupling of probability measures)*
Let $(\Omega_i, \mathcal{A}_i, P_i)$, $i \in \{1,2\}$, be probability spaces. A probability measure $\tilde{P}$ on

$$(\Omega, \mathcal{A}) := (\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$$

is called a coupling of (P_1, P_2), if the following marginal conditions are satisfied: i) $\tilde{P}(A_1 \times \Omega_2) = P_1(A_1)$ for all $A_1 \in \mathcal{A}_1$
ii) $\tilde{P}(\Omega_2 \times A_2) = P_2(A_2)$ for all $A_2 \in \mathcal{A}_2$.

For a measurable space $(S, \mathcal{S})$ let $\Omega_1 = \Omega_2 = S^{\mathbb{N}}$ and let $K : S \times \mathcal{S} \to [0,1]$ be a transition kernel on S. We consider the projections Y_n^i, $n \geq 0$, given by $Y_n(\omega) = (Y_n^1(\omega), Y_n^2(\omega)) = (\omega_n^1, \omega_n^2)$, $\omega \in \Omega_1 \times \Omega_2$. Given initial distributions ν and μ, a measure $\tilde{P}$ is a coupling measure on the path space $(\Omega, \mathcal{A}) := ((S^2)^{(\mathbb{N}_0)}, \otimes_{k=0}^{\infty}\mathcal{S}^2)$ with respect to the induced measures P_ν and P_ν if for $A \in \sigma(X_n, n \geq 0)$ the marginal conditions

$$\begin{aligned}\tilde{P}(A \times \Omega_2) &= P_\mu(A)\\ \tilde{P}(\Omega_2 \times A) &= P_\nu(A)\end{aligned}$$

are fulfilled. Let T be the hitting time for the diagonal in $S \times S$, i.e.

$$T := \inf\{n \geq 0 : Y_n^1 = Y_n^2\}. \tag{3.14}$$

In the sequel we will be interested in measures $\tilde{P}$ on the sequence space $(\Omega, \mathcal{A})$, such that the sequence $(Z_n)_n$ given by

$$Z_n := f(X_n) + 2\alpha \sum_{k=0}^{n-1} f(X_n)\ln(f(X_n))$$

is a martingale.
In order to specify those we introduce coupling of transition kernels.

Definition 3.2.2 *(Coupling of transition kernels and generators)*
Let $K_i : \Omega_i \times \mathcal{A}_i \to [0,1]$, $i \in \{1,2\}$ be transition kernels and let $L_i : L^2(S,\pi) \to L^2(S,\pi)$, $i \in \{1,2\}$, be the corresponding generators. A transition kernel, respectively generator

$$\tilde{K} : (\Omega_1 \times \Omega_2) \times (\mathcal{A}_1 \otimes \mathcal{A}_2) \to [0,1],$$

$$\tilde{L} : L^2(S \times S, \pi \otimes \pi) \to L^2(S \times S, \pi \otimes \pi),$$

is called a coupling of (K_1, K_2) or (L_1, L_2) respectively if
i) $\tilde{K}((\omega_1,\omega_2), A_1 \times \Omega_2) = K_1(\omega_1, A_1)$ for all $(\omega_1,\omega_2) \in \Omega_1 \times \Omega_2$, $A_1 \in \mathcal{A}_1$,
ii) $\tilde{K}((\omega_1,\omega_2), \Omega_1 \times A_2) = K_2(\omega_2, A_2)$ for all $(\omega_1,\omega_2) \in \Omega_1 \times \Omega_2$, $A_2 \in \mathcal{A}_2$.

or respectively,

i) $\tilde{L}(\tilde{f}) = [L_1 f](x)$ for all $\tilde{f} = f(x)$,
ii) $\tilde{L}(\tilde{f}) = [L_2 f](y)$ for all $\tilde{f} = f(y)$.

Let $(S, \mathcal{S})$ be a measurable space and for $n \geq 0$ let $K^n : S^{n+1} \times \mathcal{S} \to [0,1]$ be transition kernels, where $S^n := \underbrace{S \times ... \times S}_{\otimes n}$. Introduce the discrete generator $L^n := I^n - K^n$. Here I^n denotes the kernel defined by

$$I^n(x_0, ..., x_n, A) = I(x^n, A) = \begin{cases} 1 & : \quad x^n \in A, \\ 0 & : \quad x^n \notin A. \end{cases}$$

Furthermore, we denote by Y_n the coordinate process. Recall that for any initial distribution μ on $(S, \mathcal{S})$ there exists exactly one measure P_μ on $(\Omega, \otimes_{n=0}^{\infty} \mathcal{S})$, such that for any bounded $\mathcal{S}^{n+1}$-measurable function $f : S^{n+1} \to \mathbb{R}$

$$\int_\Omega f \circ (Y_0, ..., Y_n)\, dP_\mu = \int_{S^{n+1}} f\, dP_\mu^{(n)}$$

holds. Here $P_\mu^{(n)}$ is defined recursively:
For $n = 0$ set $P_\mu^{(0)} = \mu$. Given $P_\mu^{(n-1)}$ on $(S^{(n+1)}, \mathcal{S}^{(n+1)})$ by

$$P_\mu^{(n+1)}(A) = P_\mu^{(n)} \otimes K^n(A) = \int_{S^n} \int_S 1_A(x, y) P_\mu^{(n)}(dx).$$

We have the following lemma.

Lemma 3.2.3 *For any bounded $\mathcal{S}$-measurable function f and any initial distribution μ on $(S, \mathcal{S})$ the sequence $(Z_n)_{n\geq 0}$ defined by*

$$Z_n := f(Y_n) - \sum_{k=0}^{n-1} [L^k f](Y_0, ..., Y_k), \tag{3.15}$$

is a martingale under P_μ with respect to the natural filtration $(\mathcal{F}_n)_{n\geq 0}$, i.e. $\mathcal{F}_n := \sigma(Y_k : 0 \leq k \leq n)$.

Proof:
The function $(Y_0, ..., Y_k) : \Omega \to S^{k+1}$ is $\mathcal{F}_k$-measurable and so is the map $[L_k f] : S^{k+1} \to \mathbb{R}$. As a composition of $\mathcal{F}_k$-measurable functions the map $[L_k f](Y_0, ..., Y_k) : \Omega \to \mathbb{R}$ is $\mathcal{F}_k \subset \mathcal{F}_n$-measurable. Therefore $(Z_n)_n$ is $(\mathcal{F}_n)_n$-adapted.
To obtain the martingale property we show that

$$\mathbb{E}[f(Y_n + 1)|\mathcal{F}_n] = [K_n f](Y_0, ..., Y_n)$$

holds true.
If we take $A = A_n \times S \times S \times ... \in \mathcal{F}_n$, where $A_n \in \bigotimes_{k=0}^n S$, we can conclude

$$\begin{aligned} \int_A f(Y_0, ..., Y_{n+1})\, dP_\mu &= \int_{A_n} \int_S f(y) K_n(x, dy)\, P_\mu^{(n)}(dx) \\ &= \int_{A_n} [K_n f](x)\, P_\mu^{(n)}(dx) \\ &= \int_A [K_n f]\, dP_\mu. \end{aligned}$$

This implies

$$\mathbb{E}[Z_{n+1} - Z_n|\mathcal{F}_n] = \mathbb{E}[f(Y_n + 1) - f(Y_n) + [L_n f](Y_0, ..., Y_n)|\mathcal{F}_n] = 0.$$

This is the claimed result. □

Examples of Coupling Kernels

1) Classical Coupling
Let $\mathcal{X}$ be a discrete set and K a transition kernel on $\mathcal{X}$. The classical coupling $\tilde{K}$ is now given by

$$\begin{aligned} \tilde{K}((x_1, x_2), (y_1, y_2)) &:= K(x_1, y_1) K(x_2, y_2), \quad y_1 \neq y_2, \\ \tilde{K}((x, x), (y, y)) &:= K(x, y). \end{aligned}$$

The classical coupling describes a random walk of two particles each of them marginally behaving like K. After they have met for the first time they behave like one particle and move according to K.
2) Wasserstein Coupling
In this case the coupling kernel $\tilde{K}$ is given by

$$\begin{aligned}
\tilde{K}((x_1,x_2),(y,y)) &:= K(x_1,y)\wedge K(x_2,y)),\\
\tilde{K}((x_1,x_2),(y_1,y_2)) &:= \frac{(K(x_1,y_1)-K(x_2,y_1))^+(K(x_2,y_2)-K(x_1,y_2))^+}{1-\sum_{x\in\mathcal{X}}K(x_1,z)\wedge K(x_2,z)}\\
&\quad\ \text{for } x_1\neq x_2,\ y_1\neq y_2.
\end{aligned}$$

There are situations where 1) and 2) coincide, e.g. it has been shown for nearest-neighbour random walk on the star defined in Example 3 of this chapter.

3.3 Logarithmic-Sobolev Functions

We apply coupling to derive lower bounds on the logarithmic-Sobolev constant α. Since it is a nonlinear problem the lower bound contains the maximum of a logarithmic-Sobolev function to be introduced in Proposition 3.3.1 which characterizes logarithmic-Sobolev functions using a difference equation.
Here we have to remind the reader that the spectral gap is defined for general, not necessarily reversible Markov chains, see (2.55).

Proposition 3.3.1 *Let (K,π) be an irreducible, aperiodic Markov Chain on $\mathcal{X}$. Let α be its logarithmic-Sobolev constant and λ_1 be its spectral gap. Then either $\frac{\lambda_1}{2}=\alpha$ or there exists a positive non-constant function f_α which is the solution of*

$$\alpha f_\alpha \ln(f_\alpha^2) = L f_\alpha, \qquad \|f_\alpha\|_2 = 1 \tag{3.16}$$

and therefore $\alpha=\frac{\mathcal{E}(f_\alpha,f_\alpha)}{\mathcal{L}_\pi[f_\alpha]}$. In particular $\alpha>0$.

Proof:
According to the variational characterization of the logarithmic-Sobolev constant we have to minimize $\frac{\mathcal{E}(f,f)}{\mathcal{L}_\pi(f)}$ and can restrict ourselves to the case $\|f\|_2=1$. Now, there either exists a non-negative minimizer m or the minimum is attained at the constant function 1, i.e. $\mathcal{E}(f,f)=0=\mathcal{L}_\pi(f)$. In the latter case using $\frac{\lambda_1}{2}\geq\alpha$ shows that $\alpha=\frac{\lambda_1}{2}$. Because for a function $g\neq 0$, $\mathbb{E}_\pi[g]=0$

$$\lim_{\epsilon\to 0}\frac{\mathcal{E}(1+\epsilon g,1+\epsilon g)}{\mathcal{L}_\pi(1+\epsilon g)} = \lim_{\epsilon\to 0}\frac{\epsilon^2\mathcal{E}(g,g)}{2\epsilon^2 Var_\pi[g]} \geq \frac{\lambda_1}{2}.$$

If $\alpha \neq \frac{\lambda_1}{2}$, there has to be a minimizer m, which clearly has to fulfill (3.16). If the minimizer m vanishes at one point, $x_0 \in \mathcal{X}$ say, we have $[Km](x_0) = 0$ and m must vanish at all points x, with $K(x_0, x) > 0$. By irreducibility, this would imply $m(x) = 0$ for all $x \in \mathcal{X}$. But we actually assumed that $m \neq 0$. This proves the proposition. □

Remark 3.3.2 *The case $\frac{\lambda_1}{2} = \alpha$ occurs, whenever the difference equation in Proposition 3.3.1 is only solvable for constant functions.*

3.4 Universal Lower Bounds

We are in a position to state our theorems.

Theorem 3.4.1 *Let (K, π) be an irreducible Markov chain on $\mathcal{X}$, T the coupling time introduced in (3.14) and f_α a positive Log-Sobolev function for K. If f^* denotes the maximum of f_α we have*

$$\alpha \geq \frac{1}{2} \frac{1}{\sup_{x,y \in \mathcal{X}} \mathbb{E}_{x,y}[T]} \frac{f^* - 1}{f^* \ln(f^*)}. \tag{3.17}$$

Proof:
Take $x_0, y_0 \in \mathcal{X}$ such that f_α attains its maximum at x_0 and its minimum at y_0. We have

$$f(x_0) - f(y_0) = f(X_n) - f(Y_n) + \sum_{k=1}^{n-1} \left([Lf](X_k) - [Lf](Y_k)\right),$$

Here we used the martingale property for $Z_n = f(W_n) + \sum_{k+1}^{n-1} [Lf](W_k)$, $W_0 = w$ for $(W_n)_n$ having the law of K. Applying the optional sampling theorem for martingales, we obtain

$$f(x_0) - f(y_0) = f(X_{T \wedge n}) - f(Y_{T \wedge n}) + \sum_{k=1}^{T \wedge n-1} \left([Lf](X_k) - [Lf](Y_k)\right).$$

Now, using the difference charactarization of f_α, (3.16) and Lebesgue's convergence theorem we obtain

$$f(x_0) - f(y_0) = \underbrace{f(X_T) - f(Y_T)}_{=0} + 2\alpha \sum_{k=1}^{T-1} \left(f(X_k)\ln(f(X_k)) - f(Y_k)\ln(f(Y_k))\right).$$

This gives

$$|f(x_0) - f(y_0)| \leq 2\alpha \sum_{k=1}^{T-1} |f(X_k)\ln(f(X_k)) - f(Y_k)\ln(f(Y_k))| .$$

Thus,

$$|f(x_0) - f(y_0)| \leq 2\alpha\delta(f\ln(f))\mathbb{E}^{x_0,y_0}[T],$$

where $\delta(f\ln(f)) := \max_{x,y\in\mathcal{X}}(f(x)\ln(f(x)) - f(y)\ln(f(y)))$.
If $\max_{x\in\mathcal{X}} f(x) < 1$, we have $\delta(f\ln(f)) \leq 1$ and otherwise

$$\begin{aligned}
\delta(f\ln(f)) &\leq \max_{x\in\mathcal{X}} f(x)\ln(f(x)) - \min_{x\in\mathcal{X}} f(x)\ln(f(x)) \\
&= f(x_0)\ln(f(x_0)) - \left(f(y_0)\vee e^{-1}\right)\ln\left(f(y_0)\vee e^{-1}\right) \\
&= \int_{(f(y_0)\vee e^{-1})}^{f(x_0)} (1+\ln(s))\,ds \\
&\leq \left(f(x_0) - f(y_0)\vee e^{-1}\right)\frac{1}{f(x_0)-1}\int_1^{f(x_0)} (1+\ln(f(s)))\,ds \\
&\leq f(x_0) - f(y_0)\frac{f(x_0)\ln(f(x_0))}{f(x_0)-1}
\end{aligned}$$

By using $f(x_0) = f^*$ and dividing by $f(x_0) - f(y_0)$ leads to

$$1 \leq 2\alpha\mathbb{E}^{x_0,y_0}[T]\frac{f^*\ln(f^*)}{f^*-1}.$$

Thus, the theorem is proved. □

Theorem 3.4.2 *Let (K,π) be an irreducible and aperiodic Markov chain on $\mathcal{X}$. Let $T_y := \inf\{n \geq 0 : X_n = y\}$ and suppose that for all $x, y \in \mathcal{X}$, $\mathbb{E}_x[T_y] < \infty$, then, we have*

$$\alpha \geq \frac{1}{2}\frac{1}{\sup_{x,y}\mathbb{E}_x[T_y]}\frac{1}{1+\ln(f^*)}. \tag{3.18}$$

Proof:
Using that $Z_n := f(Y_n) + \sum_{k=1}^{n-1}[Lf](Y_k)$ is a martingale, the optional stopping theorem and Lebesgue's convergence theorem we have

$$f(x) = f(y) + \mathbb{E}_x\left[\sum_{k=1}^{T_y}[Lf](Y_k)\right].$$

If f is a logarithmic-Sobolev function then there are $x, y \in \mathcal{X}$, such that $f(x) > 1$ and $f(y) < 1$, because $\|f\|_2 = 1$ and f is not constant. Take $\tilde{f(x, y)} := f(x)1_A(x) - f(y)1_B(y)$, where

$$A := \{x : f(x) > 1\}, \quad B := \{x : f(x) < 1\}.$$

Then, $\tilde{f}$ fulfills (3.16) and we have

$$1 \leq |\tilde{f}(x, y)| \leq 2\alpha\mathbb{E}_x[T_y](1 + \ln(f^*))).$$

This gives the result. □

Theorem 3.4.3 *Let (K, π) be an irreducible and aperiodic Markov Chain on $\mathcal{X}$. Let $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$ be a function, such that for some coupling $\tilde{L}$ of (L, L) and $0 \leq c_d \in \mathbb{R}$ the inequality*

$$\left[\tilde{L}d\right](x, y) \geq c_d d(x, y), \text{ for all } x \neq y \tag{3.19}$$

holds true. Then, we have for the logarithmic-Sobolev constant α of the chain K,

$$\alpha \geq \frac{c_d}{2\,(1 + \ln(f^*)))}. \tag{3.20}$$

Proof:
We denote by

$$\mathrm{lip}_f := \sup\left\{\left|\frac{f(v) - f(w)}{d(v, w)}\right|, v \neq w\right\}$$

the Lipschitz-constant of f with respect to d. Take $x, y \in \mathcal{X}, x \neq y$, such that $\mathrm{lip}_f = \frac{f(x)-f(y)}{d(x,y)}$. Using $\beta = 1 + \sup_{v\in\mathcal{X}}\{\ln(f(v))\}$ and $\tilde{f}(x, y) = f(x) - f(y)$, we find

$$\begin{aligned}
2\alpha\frac{|f(x) - f(y)|}{d(x, y)}d(x, y)\beta &= 2\alpha|f(x) - f(y)|\beta \\
&\geq 2\alpha|f(x)\ln(f(x)) - f(y)\ln(f(y))| \\
&= \left|\left[\tilde{L}\tilde{f}\right](x, y)\right| \\
&= \left|\left[\tilde{L}\tilde{f}/d \cdot d\right](x, y)\right| \\
&= |\mathrm{lip}_f[\,\tilde{L}d\,](x, y)| \\
&\geq |\mathrm{lip}_f c_d d(x, y)|.
\end{aligned}$$

Dividing by $d(x, y) \neq 0$ and 2β gives

$$\alpha \geq \frac{c_d}{2\,(1 + \ln(f^*))}.$$

Therefore, we obtain the desired inequality. □

Proposition 3.4.4 *(i) If* $\pi^* := \max_{x\in\mathcal{X}} \pi(x)^{-1}$*, then the maximum of* f_α *is bounded by* $\sqrt{\pi^*}$*.*
(ii) Let (K,π) *be reversible. If* I_d *denotes the* d*-dimensional isoperimetric constant of the chain* (K,π) *we have* $\max_{x\in\mathcal{X}} f_\alpha(x) \le \exp\left(\frac{2d}{I_d^2}\right)$.

Proof:
(i) is obvious because $\|f_\alpha\|_2 = 1$.
(ii) To prove this statement, we firstly observe that if the Markov chain is reversible we have the following lower bound on the logarithmic-Sobolev constant α by [SC 97], Theorem 2.3.18,

$$\alpha \ge \frac{I_d^2}{4d}.$$

By the difference characterization of f_α, (3.16), we have

$$2\alpha f_\alpha(x)\ln(f_\alpha(x)) = [Lf_\alpha](x) = f_\alpha(x) - \sum_{y\in\mathcal{X}} f_\alpha(y)K(x,y)$$

and thus,

$$2\alpha f_\alpha(x)\ln(f_\alpha(x)) \le f_\alpha(x).$$

Because $f_\alpha(x) > 0$ for all $x \in \mathcal{X}$ we can divide by $f_\alpha(x)$. Since we can find a lower bound on α in terms of the d-dimensional isoperimetric I_d, we have

$$\ln(f_\alpha(x)) \le \frac{1}{2\alpha} \le \frac{2d}{I_d^2}.$$

But this shows that

$$\max_{x\in\mathcal{X}} f_\alpha(x) \le \exp\left(\frac{2d}{I_d^2}\right).$$

Therefore, the proposition has been proven. □

Remark 3.4.5 *Sometimes the upper bound in terms of* I_d *gives slightly better estimates then the obvious bound in (i). This is the case for the discrete cube, Example 4 in the next section.*

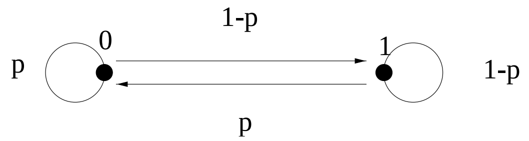

Figure 3.1: **The two point space** $\{0,1\}$

3.5 Examples

Example 1: The two point space $\{0,1\}$

For $0 < p < 1$, the non-symmetric transition kernel K on $\{0,1\}$ is given by

$$K = \begin{pmatrix} p & 1-p \\ p & 1-p \end{pmatrix}.$$

Then, the stationary measure corresponding to K is $\pi(0) = p, \pi(1) = 1-p$. To state our lower bounds we use the classical coupling introduced in Section 3.2. For the expected time to hit the diagonal we have

$$\begin{aligned} \mathbb{E}_{(0,1)}[T] &= \mathbb{E}_{(1,0)}[T] \\ &= \sum_{n=0}^{\infty} nP[T=n] \\ &= \sum_{n=0}^{\infty}\sum_{k=0}^{n-1} \binom{n-1}{k} (p(1-p))^k (p(1-p))^{n-1-k}(p^2+(1-p)^2) \\ &= \sum_{n=0}^{\infty} n(2p(1-p))^{n-1}(p^2+(1-p)^2) \\ &= \frac{(p^2+(1-p)^2)}{(1-2p(1-p))^2}. \end{aligned}$$

Thus, applying Theorem 3.4.1

$$\alpha \geq \frac{(1-2p(1-p))^2}{(p^2+(1-p)^2)} \frac{1-\sqrt{p}}{|\ln(\sqrt{p})|}.$$

To apply the distance method, Theorem 3.4.3, we use the distance $\delta_{xy} = \begin{cases} 1, x = y \\ 0, x \neq y \end{cases}$. If $\tilde{K}$ denotes the kernel corresponding to the classical coupling we used

$$\sum_{(x_1,y_1)\in\{0,1\}\times\{0,1\}} \tilde{K}((x,y),(x_1,y_1))\delta_{x_1y_1} \leq (1-(1+p(1-p)))\delta_{xy}.$$

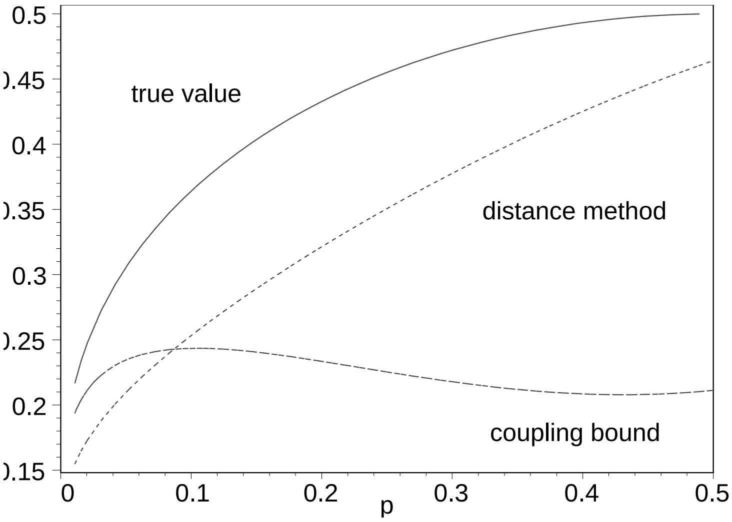

Figure 3.2: **Performance of the lower bounds for Example 1**

Thus, the distance method, Theorem 3.4.3, yields

$$\alpha \geq \frac{1+p(1-p)}{2(1+|\ln(\sqrt{p})|)}.$$

The exact value of α depends on p and is

$$\alpha = \begin{cases} \frac{1-2p}{\ln((1-p)/p)} & : \ 0 < p < 1/2, \\ 1/2 & : \ p = 1/2. \end{cases}$$

Figure 3.2 illustrates the performance of our lower bounds with respect to the exact value of α.

Example 2: The discrete circle $\mathbb{Z}_p$

Firstly, consider the random walk $(\tilde{X}_n)_n$ on the integers $\mathbb{Z}$ with transition kernel

$$K_{\mathbb{Z}}(i,j) = \begin{cases} 1/2 & : \ |i-j| = 1, \\ 0 & : \ \text{otherwise}. \end{cases}$$

Take $(X_n)_n := mod(\tilde{X}_n)_n$ for $mod(i) = i \bmod p$. Thus, results for the discrete circle $\mathbb{Z}_p$ can be deduced from results for $\mathbb{Z}$. For instance, with T_k, the hitting time for k, defined in Theorem 3.4.2, we find

$$\mathbb{E}_0[T_k] = k(p-k).$$

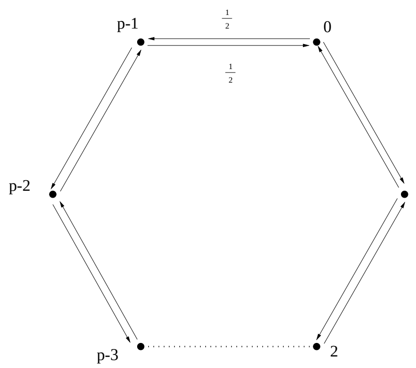

Figure 3.3: **The discrete circle $\mathbb{Z}_p$**

Because this is the mean exit time from the interval $(k-n, k)$ for the random walk on $\mathbb{Z}$. The latter expression is maximized for $k = \lfloor p/2 \rfloor$, $p \geq 3$. By rotational symmetry we have this result for all possible starting points $x \in \mathbb{Z}_p$. Using the latter, the point-to-point hitting bound, Theorem 3.4.2, implies

$$\alpha \geq \frac{2}{\lfloor p^2/4 \rfloor} \frac{1}{1 + \ln(\sqrt{p}))}.$$

The best known lower bound has been obtained in [DS-C 96b]. They show

$$\alpha \geq \frac{8\pi^2}{25p^2}.$$

Thus, for small p our bound is better but for large p it is off by a factor $1 + \ln(\sqrt{p})$ and therefore not of the right order.

Example 3: The star S_p

For $p \in \mathbb{N}$ let S_p be the set $\{0, 1, 2, ..., p\}$ and for $\theta \in (0, 1)$ let the transition kernel K on S_p be given by

$$K = \begin{pmatrix} \theta & \frac{\theta}{p+1} & \dots & \frac{\theta}{p+1} \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & 0 \end{pmatrix}.$$

The set S_p together with the graph structure induced by K is called the star with p elements and center 0.

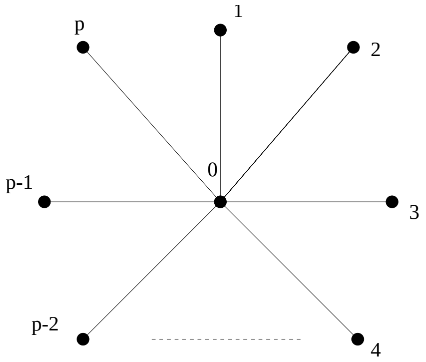

Figure 3.4: **The star S_p**

The stationary distribution is $\pi(0) = 1/2$, and $\pi(x) = 1/2n$ for $x \neq 0$. To apply the distance method, Theorem 3.4.3, we use the classical coupling kernel $\tilde{K}$ of (K, K) introduced in 3.2 and the distance function δ_{xy}. For any $0 \neq x, y \in S$ we have

$$[\tilde{K}\delta](x, y) = 0.$$

For $x = 0, y \neq 0$ we compute

$$[\tilde{K}\delta](x, y) = (1 - \theta)\delta_{xy}.$$

Therefore, we can use θ as the appropriate constant in Theorem 3.4.3 and obtain

$$\alpha \geq \frac{\theta}{2 + \ln(p)}.$$

We have not found a lower bound for this case in the literature but we suppose that the bound is not of the right order of magnitude because of the factor $1/(2 + \ln(p))$.

Example 4: The discrete cube $\mathbb{Z}_2^d$

We consider the nearest neighbour random walk K on the discrete cube $\mathbb{Z}_2^d$ without holding and define the following coupling $\tilde{K}$ of (K, K).
For $x, y \in \mathbb{Z}_2^d$ denote

$$V(x, y) = \{1 \leq j \leq d : x(j) \neq y(j)\}.$$

The set of all indices at which x and y differ. Then, we define for $\overline{x} = (x_1, x_2) \in \mathbb{Z}_2^d \times \mathbb{Z}_2^d$

$$\tilde{K}(\overline{x}, \overline{x} + (e_i, e_j)) = \frac{1}{d}, \quad i \in \{1, ..., d\} \backslash V(x_1, x_2),$$

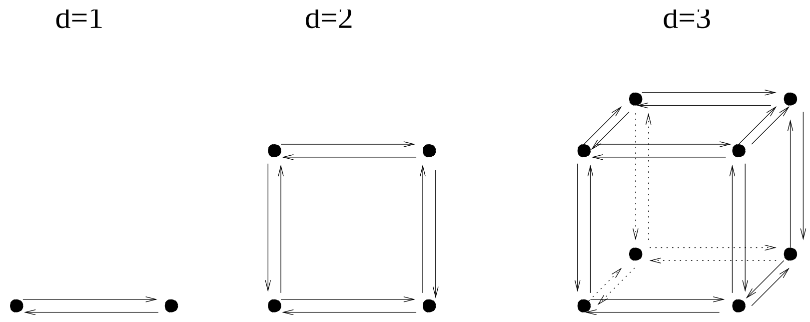

Figure 3.5: **The discrete cube $\mathbb{Z}_2^d$ for $d = 1, 2, 3$**

$$\tilde{K}(\overline{x}, \overline{x} + (e_i, e_j)) = \frac{1}{d(|V(x_1, x_2)| - 1)}$$

for $i, j \in V(x_1, x_2), i \neq j, |V(x_1, x_2)| \geq 2$ and

$$\tilde{K}(\overline{x}, \overline{x} + (e_i, e_i)) = \frac{1}{d}$$

for $1 \leq i \leq d$, $|V(x_1, x_2)| \leq 1$. For the Hamming distance $d_H(x_1, x_2) = \sum_{j=1}^{d} 1_{\{x_i \neq y_i\}}$, Meise in [Me 96] considers the function $h : \mathbb{Z}_2^d \times \mathbb{Z}_2^d \to \mathbb{N}$ given by

$$h(x, y) = \begin{cases} d_H(x, y) & : \ f(x) - f(y) > 0, \\ 0 & : \ \text{otherwise.} \end{cases}$$

and he proved

$$[\tilde{L}h](x, y) \geq \frac{1}{d} d(x, y).$$

Furthermore, by [SC 97], Theorem 2.3.15, $I_d \geq \frac{1}{2d2^{1-1/d}}$. Therefore, the distance method yields

$$\alpha \geq \frac{1}{2d} \frac{1}{1 + d\ln(2)}.$$

For the random walk on the discrete cube with holding probability $\frac{1}{d+1}$ we use the Hamming distance d_H instead of the function h and obtain

$$\alpha \geq \frac{1}{2(d+1)} \frac{1}{1 + d\ln(2)}.$$

Because Diaconis and Saloff-Coste in [DS-C 96a] showed that $\alpha = 1/2d$ in the case without holding and $1/2(d+1)$ in the case with holding, our bounds are off by the factor $\frac{1}{1+d\ln(2)}$.

3.6 Log Sobolev - Isoperimetry

In the sequel we give examples to illustrate Theorem 2.8.2, the logarithmic-Cheeger inequality, and Remark 2.8.3, the mixed isoperimetric bound. As we will see the lower bounds are of the right order.

Example 1: The discrete Circle $\mathbb{Z}_p$

We consider the discrete circle $\mathbb{Z}_p$ as in Example 3 of Section 3.5.
Choosing two subsets $A_1, A_2 \subset \mathcal{X}$ such that $\pi(A_1) = \pi(A_2)$ and A_1 is connected we find $Q(\partial A_1) \leq Q(\partial A_2)$, because $|\partial A_1| \leq |\partial A_2|$. Therefore, we can restrict ourselves to connected subsets. Let $A \subset \mathcal{X}$ be a connected subset then $Q(\partial A) = \frac{1}{p+1}$. In order to state a lower bound on I_K^{ln} we have to minimize $\frac{1}{|A| \ln\left(\frac{p}{|A|}\right)}$ with respect to the number of elements $|A|$. To this end we consider the function

$$f(t) := t \ln\left(\frac{p}{t}\right),\ t \in [0, p]. \tag{3.21}$$

Because $f(0) = f(p) = 0$ and $0 \leq f|_{(0,p)}$ is differentiable there has to be t_{max} such that $f(t_{max}) = \max_{t \in [0,p]} f(t)$. The point $t_{max} = \frac{p}{e}$. Thus, we can take A to be a connected set such that $|A| = [p/e]$. Let

$$h(p) := \begin{cases} \frac{2}{p \ln(2)} & : \ p \text{ even}, \\ \frac{2}{(p+1)(\ln(2p) - \ln(p+1))} & : \ p \text{ odd}. \end{cases}$$

Then, we have

$$h(p) \geq I_K^{ln} \geq \frac{e}{p}.$$

Using the logarithmic-Cheeger inequality, we obtain

$$h(p) \geq \alpha \geq \frac{\ln(2) e^2}{4\sqrt{2} p^2}.$$

By using $\lambda_1 = 1 - \cos\left(\frac{\pi}{p}\right) \approx \frac{\pi^2}{2p^2}$, the mixed isoperimetric bound, Remark 2.8.3, yields

$$\frac{\pi^2}{2p^2} \geq \alpha \geq \frac{\lambda_1 \frac{e}{p}}{2(\sqrt{\lambda_1} + \frac{e}{p})} \approx \frac{\pi^2 e}{2\sqrt{2}(\pi + \sqrt{2} e) p^2}.$$

We recover that the spectral gap and the logarithmic-Sobolev constant are of the same order, i.e. $1/p^2$. But we improve the lower bound on α of Diaconis and Saloff-Coste, see Section 3.5, Example 2, by a factor 3. The exact value for α in this case is unknown.

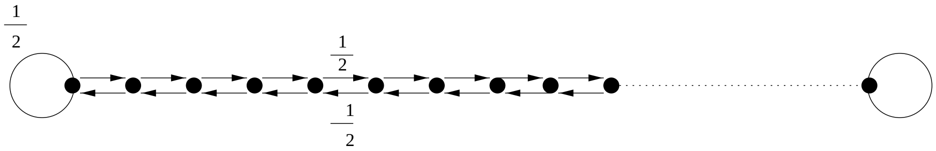

Figure 3.6: **The discrete line segment** $\{0, 1, ..., n\}, n \in \mathbb{N}$

Example 2: The discrete line segment $\{0, 1, ..., n\}, n \in \mathbb{N}$

We consider nearest neighbour random walk on the discrete line segment $\mathcal{X} = \{0, 1, ..., n\}$ with loops at the ends. The transition kernel K is

$$K(x,y) = \begin{cases} \frac{1}{2} & , \quad x, y \text{ adjacent}, \\ \frac{1}{2} & , \quad x = y, \ x \text{ endpoint}, \\ 0 & , \quad x = y, \ x \text{ interior point}, \\ 0 & , \quad x, y \text{ not adjacent}. \end{cases}$$

The stationary measure for the Markov chain is $\pi(x) = \frac{1}{n+1}$. By the same arguments as in the previous example we can restrict ourselves to connected subsets $A \subset \mathcal{X}$. For a connected subset A the measure of its boundary $Q(\partial A)$ is minimized if A contains either 0 or n. Thus, we can conclude $Q(\partial A) = \frac{1}{2(n+1)}$.
Hence, we have to find the maximum of

$$f(t) := 2t \ln\left(\frac{n+1}{t}\right), \ t \in [0, n+1].$$

As in the previous example, we find $t_{max} = \frac{n+1}{e}$. If we denote

$$h(n) := \begin{cases} \frac{2}{n \ln(2)} & : \ n \text{ even}, \\ \frac{1}{(n+1)\ln(2)} & : \ n \text{ odd}. \end{cases}$$

it follows that

$$h(n) \geq I_K^{ln} \geq \frac{e}{n+1}.$$

We obtain, applying the logarithmic-Cheeger inequality, Theorem 2.8.2,

$$h(n) \geq \alpha \geq \frac{\ln(2)e}{4\sqrt{2}(n+1)^2}.$$

Whereas the mixed isoperimetric bound yields

$$\alpha \geq \frac{\lambda_1 e}{2((1+n)\sqrt{\lambda_1} + e)}.$$

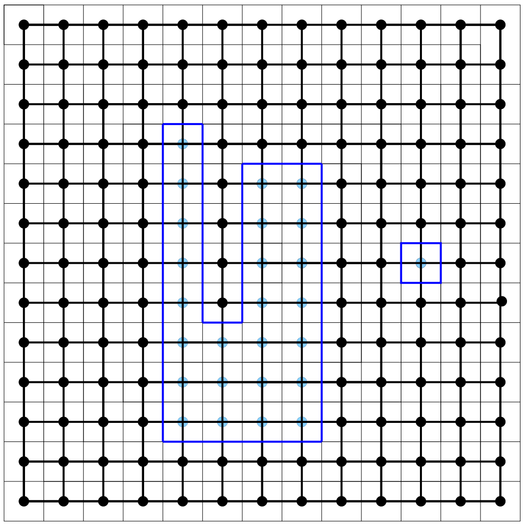

Figure 3.7: **The square grid** $\{0, 1, ..., n\}^2, n \in \mathbb{N}$

Since the eigenvalues in this case are the numbers $1-\cos\left(\frac{\pi j}{n+1}\right)$, $j = 0, 1, ..., n$, we remark that our lower bounds are of the right order of magnitude namely $1/(n+1)^2$. The exact value for α is unknown in this case.

Example 3: The square grid

We consider the square grid $\mathcal{X} = \{0, 1, ..., n\}^2$ and the transition kernel K given by

$$K(x,y) = \begin{cases} \frac{1}{4} & , \quad |x-y| = 1, \\ 0 & , \quad |x-y| > 1, \\ 0 & , \quad x = y, x \text{ interior point}, \\ \frac{1}{2} & , \quad x = y, x \text{ side point}, \\ \frac{1}{4} & , \quad x = y, x \text{ corner point}. \end{cases}$$

The adapted edge set is given by $\mathcal{A} = \{(x, y) \in \mathcal{X} : |x-y| = 1\}$ and the stationary measure is $\pi(x) = \frac{1}{(n+1)^2}$.
The geometric interpretation of the boundary ∂A of a set $A \subset \mathcal{X}$:
We view each vertex as the center of a square as in Figure 3.6. Then,
$\pi(A) \propto$ surface of those unit squares with center in A.
$Q(\partial A) \propto$ the length of the interior part of the boundary ∂A.
We minimize the measure of the boundary by the following procedure:

(i) As in the previous examples we can restrict ourselves to connected sets.

(ii) Let the number of elements in a set $A \subset \mathcal{X}$ be N. We can write

$N = k(n+1) + l$. Fill the k bottom rows and then start from the left filling the $k+1$-th row.

Following this procedure minimizes the boundary measure for a subset of N elements. We obtain

$$Q(\partial A) = \begin{cases} \frac{N+1}{4(n+1)^2} & , \quad |A| = N, N \leq n+1, \\ \frac{n+1}{4(n+1)^2} & , \quad n+1 \leq |A| = N, N = k(n+1), \\ \frac{n+2}{4(n+1)^2} & , \quad n+1 \leq |A| = N, |A| \text{ does not divide } n+1. \end{cases}$$

This means we have to consider the following functions and their minima:

$$\begin{aligned} f_1(t) &:= \frac{t+1}{4t \ln\left(\frac{(n+1)^2}{t}\right)}, \quad t < n+1, \\ f_2(t) &:= \frac{n+1}{4t \ln\left(\frac{(n+1)^2}{t}\right)}, \quad t = k(n+1), k \geq 1, \\ f_3(t) &:= \frac{n+2}{4t \ln\left(\frac{(n+1)^2}{t}\right)}, \quad t = k(n+1) + l, l = 1, 2, ..., n, \ k \geq 1. \end{aligned}$$

Tedious computation lead to: $\min_{1 \leq t \leq n+1} f_1(t) \geq \max_{i \in \{2,3\}} \min_{t \geq n+1} f_i(t)$. Both minima are equal and attained at $t = (n+1)^2/e$. Denoting

$$h(n) := \begin{cases} \frac{(n+1)^2+2}{4(n+1)^2 \ln(2)} & : \quad n+1 \text{ even}, \\ \frac{(n+1)^2+1}{4((n+1)^2-1) \ln(2)} & : \quad n+1 \text{ odd}. \end{cases}$$

This leads to

$$h(n) \geq I_K^{ln} \geq \frac{e}{4(n+1)}.$$

Thus, Theorem 2.8.2 leads to the following lower and upper bounds

$$h(n) \geq \alpha \geq \frac{\ln(2) e^2}{64\sqrt{2}(n+1)^2}.$$

If we take into account that $\lambda_1 = \frac{1}{(n+1)^2}$, Remark 2.8.3 yields

$$\frac{1}{2(n+1)^2} \geq \alpha \geq \frac{e}{8(n+1)^2 + (1 + e/4)} \geq \frac{e}{9(n+1)^2}.$$

Thus, α is of the same order as λ_1 and therefore our lower bounds are of the right order of magnitude.

Following the same lines of thought, we can show that the nearest neighbour random walk on a d-dimensional discrete parallelepiped, i.e. $\mathcal{X} = \{0, 1, ..., n_1\} \times ... \times \{0, 1..., n_d\}$, we have:

$$I_K^{ln} \geq \frac{e}{(1+d)(n^*+1)}.$$

Here $n^* := \max_{i=1,...,d} n_i$. Therefore, the bounds using Theorem 2.8.2 and Remark 2.8.3 are

$$h(n^*) \geq \alpha \geq \frac{\ln(2)e^2}{4\sqrt{2}(1+d)^2(n^*+1)^2} \tag{3.22}$$

and

$$\frac{1}{2(n^*+1)^2} \geq \alpha \geq \frac{e}{2(1+d+e)(n^*+1)^2}. \tag{3.23}$$

To obtain the latter bounds we used $\lambda_1 = \frac{1}{(n^*+1)^2}$. The bounds in (3.22) and (3.23) are off by a factor $\frac{(d+1)^2 4\sqrt{2}}{e^2 \ln(2)}$ and $\frac{2(1+d+e)}{e}$ respectively. Therefore, the spectral gap and the logarithmic Sobolev constant are of the same order namely $1/(n^*+1)^2$. We have to remark that an increase in the dimension d decreases our lower bounds. Actually α does not depend on the dimension. This has been shown in [DS-C 96a] that $\alpha = 1/2(n^*+1)^2$.

Chapter 4

Spectral estimates including higher eigenvalues

The purpose of this chapter is to develop new lower bounds for higher eigenvalues of finite Markov chains.
We compare the performance of our lower bounds with those developed in [SC 97] and [ChY95]. It turns out that our bounds in general improve upon the bounds in the cited papers by a factor bigger than 1. In examples the factor can be significantly bigger.
We study the relation of the trace TrH_t of the heat kernel H_t introduced in (3.11) with Nash inequalities. To obtain our main result, Theorem 4.1.1, we have to prove a universal lower bound on all eigenvalues in Proposition 4.1.3. Then, we use Nash inequalities to state explicit lower bounds.
Our approach is different from those considered in [SC 97] and [ChY95]. Saloff-Coste examines the relation of TrH_t to the eigenvalue counting function, $N(s) = \#\{i : \{0, ..., |\mathcal{X}| - 1\}, \lambda_i \leq s\}$, introduced in (1.42). Bounds on $N(s)$ lead to bounds on TrH_t. Chung and Yau use isoperimetry and difference calculus. Applying Theorem 4.2.4 we improve upon their lower bounds as well.
Since Nash inequalities of type I are used in our considerations we extent a useful result to prove Nash inequalities of type II to be applicable. Moderate growth, as it is called, has been contributed to the theory of finite Markov chains by Diaconis and Saloff-Coste in [DS-C 96b]. This is the contents of Theorem 4.3.2. Furthermore, we check the performance to obtain Nash inequalities which lead to reasonable lower bounds on the eigenvalues.
Several examples like discrete tori or time dependent simulated annealing algorithms, illustrate our bounds. They all improve upon known estimates and compare well with the exact values if they are known. Moreover, the bounds can easily be calculated for our examples.

4.1 Lower bounds on higher eigenvalues

The main objecitve of this section is to prove our main theorem.

Theorem 4.1.1 *Let (K, π) be an irreducible, reversible, aperiodic Markov chain on a finite state space $\mathcal{X}$. Let $0 = \lambda_0 < \lambda_1 \le \lambda_2 \le ... \le \lambda_{|\mathcal{X}|-1}$ denote the eigenvalues of L.*
Assume (K, π) satisfies a Nash inequality of type I. Then,

$$\lambda_k \ge \frac{k^{\frac{2}{d}}}{eC}, \quad k \in \{0, ..., |\mathcal{X}| - 1\}. \tag{4.1}$$

Assume (K, π) satisfies a Nash inequality of type II. Then,

$$\text{for } k \ge \left(\frac{dC}{2T}\right)^{\frac{d}{2}} e - 1: \quad \lambda_k \ge \frac{2}{dC} \frac{(k+1)^{\frac{2}{d}}}{e^{\frac{4}{d}}} \tag{4.2}$$

and

$$fork \ge \left(\frac{dCe}{2T}\right)^{\frac{d}{2}} e - 1: \quad \lambda_k \ge \frac{(k+1)^{\frac{2}{d}}}{e^{1+\frac{2}{d}} C}. \tag{4.3}$$

Remark 4.1.2 *If (K, π) satisfies a Nash inequality of type I our results improve upon the lower bound of [SC 97], Theorem 2.3.8 (1) and Corollary 2.3.9 (1), by a factor $\frac{de^{\frac{2-d}{d}}}{2} \ge 1$.*
If (K, π) satisfies a Nash inequality of type II our results improve upon the lower bound of [SC 97], Theorem 2.8.3 (2) and Corollary 2.3.9 (2), by a factor $\frac{e^{\frac{2}{d}}}{2} \ge 1$ and holds true for more eigenvalues λ_k, $k \in \{0, 1..., |\mathcal{X}| - 1\}$. The other lower bound improves upon Saloff-Coste's lower bound by a factor $\frac{de^{\frac{4-d}{d}}}{2} \ge 2$ but if $d > 2$ it holds true for less eigenvalues.

In order to prove Theorem 4.1.1 we prove the following proposition which is a key ingredient for the proof of the theorem and new in the context of finite Markov chains.

Proposition 4.1.3 *Let (K, π) be an irreducible, reversible, aperiodic Markov chain on a finite state space $\mathcal{X}$. Let $0 = \lambda_0 < \lambda_1 \le \lambda_2 \le ... \le \lambda_{|\mathcal{X}|-1}$ denote the eigenvalues of L.*
For $k \in \{0, 1, ..., |\mathcal{X}| - 1\}$ we have

$$\lambda_k \ge \sup_{t>0} \frac{1}{t} \ln\left(\frac{k+1}{TrH_t}\right) \tag{4.4}$$

and

$$\lambda_k \ge \sup_{t>0} \frac{1}{t} \ln\left(\frac{k}{TrH_t - 1}\right). \tag{4.5}$$

Proof:
For $n = 0, ..., |\mathcal{X}| - 1$ let f_n be an eigenfunction to λ_n such that $\|f_n\|_2 = 1$. The heat kernel is given by $H_t(x, y) = \sum_{n=0}^{|\mathcal{X}|-1} \exp(-t\lambda_n) f_n(x) f_n(y) \pi(y)$. Therefore, for $0 \leq k \leq |\mathcal{X}|-1$ we have $H_t(x, x) \geq \exp(-t\lambda_k) \sum_{n=0}^{k} f_n^2(x)\pi(x)$. Thus, for all $k \in \{0, ..., |\mathcal{X}| - 1\}$

$$\begin{aligned} TrH_t &= \sum_{x \in \mathcal{X}} H_t(x, x) \\ &\geq \sum_{x \in \mathcal{X}} \exp(-t\lambda_k) \sum_{n=0}^{k} f_n^2(x)\pi(x) \\ &= \exp(-t\lambda_k) \sum_{n=0}^{k} \sum_{x \in \mathcal{X}} f_n^2(x)\pi(x) \\ &= (k + 1) \exp(-t\lambda_k). \end{aligned}$$

For all $t \in \mathbb{R}^+$, the inequality $TrH_t \geq (k + 1) \exp(-t\lambda_k)$ yields

$$\lambda_k \geq \frac{1}{t} \ln \left(\frac{k + 1}{TrH_t} \right).$$

By observing that $H_t(x, x) \geq (1 + e^{-t\lambda_k} \sum_{n=1}^{k} f_n^2(x))\pi(x)$ the same computation yields $TrH_t - 1 \geq k \exp(-t\lambda_k)$ and the second claim follows. □

Corollary 4.1.4 *Suppose an irreducible, reversible, aperiodic Markov chain (K, π) satisfies a Nash inequality of type I or II. Then, we have*

$$\lambda_k \geq \begin{cases} \sup_{t>0} \frac{1}{t} \ln \left(k \left(\frac{2t}{dC} \right)^{\frac{d}{2}} \right) & , \text{ type I} \\ \sup_{0<t\leq T} \frac{1}{t} \ln \left(\frac{k+1}{exp(1)} \left(\frac{2t}{dC} \right)^{\frac{d}{2}} \right) & , \text{ type II} \end{cases} \tag{4.6}$$

Proof:
The classical argument for establishing regularity of solutions of certain partial differential equations put forward by Nash in [N 58] uses the differential characterization of the heat kernel, see (3.12). Let a function f satisfy $\|f\|_1 = 1$. Since the kernel K is Markovian this implies $\|H_t f\|_1 \leq 1$ and together with (1.34) it yields

$$\begin{aligned} Var_\pi[H_t f] &= \|H_t(f - \mathbb{E}_\pi[f])\|_2^2 \\ &\leq \|H_t(f - \mathbb{E}_\pi[f])\|_2^{2(1+2/d)} \\ &\leq -\frac{c}{2} \frac{\partial}{\partial t} \|H_t(f - \mathbb{E}_\pi[f]\|_2^2. \end{aligned}$$

Since $\frac{\partial}{\partial t}\frac{dC}{4}\|H_t(f-\mathbb{E}[f])\|_2^{-2/d} \geq 1$ and $\frac{dC}{4}\|H_0(f-\mathbb{E}[f])\|_2^{-2/d} \geq 0$ the latter yields

$$\|H_t(f-\mathbb{E}_\pi[f])\|_2^2 \leq \left(\frac{dC}{4t}\right)^{d/2}.$$

Since we work with reversible Markov chains, we use the fact that H_t is self-adjoint and the semigroup property to find

$$H_t(x,x) \leq \left(\left(\frac{dC}{4t}\right)^{d/2} + 1\right)\pi(x).$$

Assuming (1.35) a similar argument leads to $\|H_t\|_{1\to 2} \leq e^{t/T}\frac{dC}{4t}^{d/4}$ for all $t \leq T$. Again, using the adjoint $H_t^* = H_t$ and $e^{t/T} \leq e$ yields

$$H_t(x,x) \leq e\left(\frac{dC}{2t}\right)^{d/2}.$$

Eventually, this yields

$$TrH_t \leq \begin{cases} \left(\frac{dC}{2t}\right)^{\frac{d}{2}} + 1 & , \quad \text{Nash of type I,} \\ e\left(\frac{dC}{2t}\right)^{\frac{d}{2}} & , \quad \text{Nash of type II.} \end{cases}$$

Applying Proposition 4.1.3 we obtain the result. □

We are now in the position to prove Theorem 4.1.1.

<u>**Proof of Theorem 4.1.1:**</u>

The best lower bounds on λ_k using Proposition 4.1.3 can be calculated explicitely.

We have to calculate the maximum of the function $f(t) = \frac{c_1}{t}\ln(t) + \frac{c_2}{t}$ with constants $c_1 := \frac{d}{2}$ and $c_2 := \ln\left(k\left(\frac{2}{dC}\right)^{\frac{d}{2}}\right)$. By taking the derivative $f'(t) = \frac{1}{t^2}(c_1 - c_1\ln(t) - c_2)$, we observe that the maximum is attained at the point $t_0 = \exp\left(1 - \frac{c_2}{c_1}\right)$. Now, substituting the constants c_1 and c_2 gives

$$\begin{aligned} f(t_0) &= \frac{d}{2}\exp\left(\frac{2\ln\left(k\left(\frac{2}{dC}\right)^{\frac{d}{2}}\right)}{d} - 1\right) \\ &= \frac{d}{2}\exp\left(\ln\left(k^{\frac{2}{d}}\frac{2}{dC}\right)\right) \\ &= \frac{k^{\frac{2}{d}}}{eC}. \end{aligned}$$

To obtain the first lower bound in case a Nash inequality of type II holds true, take $t_0 = \frac{dC}{2}\frac{e^{\frac{4}{d}}}{k^{\frac{2}{d}}}$. Whenever $t_0 \leq T$, i.e. $k \geq \left(\frac{dC}{2T}\right)^{\frac{d}{2}} e - 1$, we have

$$\lambda_k \geq \frac{1}{t_0} \ln\left(\frac{k+1}{e}\left(\frac{2t_0}{dC}\right)^{\frac{d}{2}}\right) = \frac{2}{dC}\frac{(k+1)^{\frac{2}{d}}}{e^{\frac{4}{d}}}.$$

To obtain the second estimate we can calculate the maximum of the function $f(t)$ from above but as constants we take $c_1 = \frac{d}{2}$ and $c_2 = \ln\left(\frac{k+1}{e}\left(\frac{2}{dC}\right)^{\frac{d}{2}}\right)$. We observe that the maximum is attained at $t_0 = \exp(1 - \frac{c_2}{c_1}) = \frac{e^{1+\frac{2}{d}}dC}{2(k+1)^{\frac{2}{d}}}$. If $t_0 \leq T$, i.e. $k \geq \left(\frac{dCe}{2T}\right)^{\frac{d}{2}} e - 1$. The lower bound is

$$\lambda_k \geq f(t_0) = \frac{(k+1)^{\frac{2}{d}}}{e^{1+\frac{2}{d}}C}.$$

This proves Theorem 4.1.1 □

4.2 Bounds via Isoperimetry

In this section we give lower bounds on higher eigenvalues in terms of the d-dimensional isoperimetric constant I_d, introduced in (1.43), i.e.

$$I_d = \min_{\substack{\pi(A)\leq\frac{1}{2}\\ A\subset\mathcal{X}}} \frac{Q(\partial A)}{\pi(A)^{\frac{d-1}{d}}}.$$

Our method states lower bounds in terms of I_d^2, whereas the bounds obtained by Chung and Yau in [ChY95] are of order I_d^4. Since $I_d \leq 1$ our results improve upon those in [ChY95].
By [SC 97], Theorem 3.3.11, the following holds true.

Theorem 4.2.1 *[SC 97] Let (K, π) be an irreducible, reversible, aperiodic Markov chain on a finite state space $\mathcal{X}$. Let $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_{|\mathcal{X}|-1}$ denote the eigenvalues of L.*
Assume, for some positvie $S \in \mathbb{R}^+$, (K, π) satisfies

$$\pi(A)^{\frac{d-1}{d}} \leq SQ(\partial A),$$

for all $A \subset \mathcal{X}$ such that $\pi(A) \leq \frac{1}{2}$ and for $d \geq 1$. Then for all $f \in L^2(\mathcal{X}, \pi)$

$$Var_\pi[f]^{(1+2/d)} \leq 8S^2\mathcal{E}(f, f)\|f\|_1^{\frac{4}{d}}.$$

Remark 4.2.2 *The same reasoning holds true for generators of Markov processes in continous space and time if the corresponding Dirichlet form satisfies a Nash inequality and the spectrum of the generator is discrete.*

Corollary 4.2.3 *Let I_d denote the d-dimensional isoperimetric constant of the chain (K, π). For all $f \in L^2(\mathcal{X}, \pi)$*

$$Var_\pi[f]^{(1+2/d)} \leq \frac{8}{I_d^2}\mathcal{E}(f,f)\|f\|_1^{\frac{4}{d}}.$$

Using our lower bounds we obtain

Theorem 4.2.4 *Let (K, π) be an irreducible, reversible, aperiodic Markov chain on a finite state space $\mathcal{X}$. Let $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_{|\mathcal{X}|-1}$ denote the eigenvalues of L.*
Let I_d denote the d-dimensional isoperimetric constant of (K, π). Then

$$\lambda_k \geq \frac{I_d^2}{8e}k^{\frac{2}{d}}. \tag{4.7}$$

Proof:
Theorem 4.1.1 together with Corollary 4.2.3 yields the desired result. □

Remark 4.2.5 *The lower bounds obtained in [ChY95] are*

$$\lambda_k \geq \frac{I_d^4(d-1)^2}{2e9^{2/d}d^2}k^{2/d}, \quad k \in \{0, 1, ..., |\mathcal{X}| - 1\}.$$

Our bounds improve these by a factor $\frac{9^{\frac{2}{d}}d^2}{4I_d^2(d-1)^2} \geq 1$, because $I_d \leq I_K$ and $2^{\frac{1}{d}} \leq I_K \leq 1$, i.e. $I_d \leq 2^{-\frac{1}{d}}$.

4.3 How to prove Nash inequalities of type I

Before we put forward the examples, let us recall a useful tool to establish Nash inequalities. Diaconis and Saloff-Coste contributed it to the theory of finite Markov chains in [DS-C 96b] to prove Nash inequalities of type II.
We will prove that the same conditions imply Nash inequalities of type I.
Let $d(\cdot, \cdot)$ denote the length of a shortest path connecting two points in the state space. As in the general case, let $B(x, r) = \{y : d(x, y) \leq r\}$ be the closed ball around $x \in \mathcal{X}$ with radius r. The volume of $B(x, r)$ is

$$V(x,r) = \sum_{y \in B(x,r)} \pi(y).$$

Furthermore, set

$$f_r(x) := \frac{1}{V(x,r)} \sum_{y \in B(x,r)} f(y)\pi(y).$$

Definition 4.3.1 *For a Markov chain (K,π) fix the edge set $\mathcal{A}$ and a probability measure π on $\mathcal{X}$. For $A, d \geq 1$, $(\mathcal{X}, \mathcal{A}, \pi)$ and $\gamma = \max_{x,y \in \mathcal{X}} d(x,y)$ we say (K,π) has (A,d)-moderate growth if*

$$V(x,r) \geq \frac{1}{A}\left(\frac{r+1}{\gamma}\right)^d$$

and the Dirichlet form $\mathcal{E}$ associated with (K,π) satisfies a local Poincaré inequality if there exists a constant $a > 0$ such that for any real function f and integer r

$$\|f - f_r\|^2_{L^2(\mathcal{X},\pi)} \leq ar^2 \mathcal{E}(f,f).$$

Now, we are able to state the main result of this section.

Theorem 4.3.2 *Let (K,π) be an irreducible, reversible, aperiodic Markov chain on a finite set $\mathcal{X}$. Suppose that for some integer R there exist real constants $M, d \geq 1$, $a > 0$ such that for all $x \in \mathcal{X}$, $r \in [0,R]$ and all functions f*

$$V(x,r) \geq \frac{(r+1)^d}{M} \quad \text{and} \quad \|f - f_r\|^2_{L^2(\mathcal{X},\pi)} \leq ar^2\mathcal{E}(f,f).$$

Then, with $C = (1+1/d)^2(d+1)^{2/d}M^{2/d}a$ we have

$$Var_\pi[f]^{1+\frac{2}{d}} \leq C\mathcal{E}(f,f)\|f\|_1^{\frac{4}{d}}.$$

Proof:
Observe that

$$\|f\|_2^2 = \langle f,f\rangle_{L^2(\mathcal{X},\pi)} = \langle f - f_r, f\rangle_{L^2(\mathcal{X},\pi)} + \langle f_r, f\rangle_{L^2(\mathcal{X},\pi)}. \tag{4.8}$$

Thus, for all $r \in [0,R]$ we can estimate the first summand on the right hand side and get

$$\langle f - f_r, f\rangle_{L^2(\mathcal{X},\pi)} \leq \|f - f_r\|_2 \|f\|_2 \leq \sqrt{a} r \mathcal{E}(f,f)^{\frac{1}{2}} \|f\|_2.$$

By definition of f_r and our assumption, we have for $\langle f_r, f\rangle_{L^2(\mathcal{X},\pi)}$ in (4.8)

$$\langle f_r, f\rangle_{L^2(\mathcal{X},\pi)} \leq \|f_r\|_\infty \|f\|_1 \leq M(r+1)^{-d}\|f\|_1^2.$$

Hence, for all $r \in [1, \infty)$,

$$\|f\|_2^2 \leq \sqrt{a\mathcal{E}(f,f)}r\|f\|_2 + Mr^{-d}\|f\|_1^2.$$

The minimum of the function $g(r) = c_1 r + c_2 r^{-d}$ with constants $c_1 = \sqrt{a\mathcal{E}(f,f)}\|f\|_2$ and $c_2 = M\|f\|_1^2$ is attained at $r_0 = d^{1+1/d}\left(\frac{c_2}{c_1}\right)^{1/(d+1)}$. Substituting r_0 in the above equation yields

$$\|f\|_2^{2(1+\frac{2}{d})} \leq M^{2/d}a(1+1/d)^2(d+1)^{2/d}\mathcal{E}(f,f)\|f\|_1^{4/d}.$$

Since the latter inequality holds for all f it can be applied to $f - c$, where c is a median for f. Using $\mathcal{E}(f,f) = \mathcal{E}(f-c, f-c)$ we find

$$\begin{aligned}\|f-c\|_2^{2(1+2/d)} &\leq M^{2/d}a(1+1/d)^2(d+1)^{2/d}\mathcal{E}(f,f)\|f-c\|_1^{4/d} \\ &\leq M^{2/d}a(1+1/d)^2(d+1)^{2/d}\mathcal{E}(f,f)\|f\|_1^{4/d}.\end{aligned}$$

Since $Var_\pi[f] = \inf_c \|f(x) - c\|_2^2$ we obtain for all f:

$$Var_\pi[f]^{(1+\frac{2}{d})} \leq M^{2/d}a(1+1/d)^2(d+1)^{2/d}\mathcal{E}(f,f)\|f\|_1^{4/d}.$$

This proves the theorem. □

Remark 4.3.3 *Under the assumptions of Theorem 4.3.2 Diaconis and Saloff-Coste in [DS-C 96b] derive a Nash inequality of type II with constants d, $C = (1+1/d)^2(1+d)^{2/d}M^{2/d}a$ and $T = aR^2$.*

One might wonder if this procedure to deduce Nash inequalities yields 'good' Nash inequalities, i.e. sharp upper bounds on the heat kernel.
In the sequel we consider the discrete circle $\mathbb{Z}_p$ and derive Nash inequalities of type I by using bare hands on calculation and the procedure outlined above. The discrete circle $\mathbb{Z}_p$ has been introduced in Example 3.3.3. We can explicitely compute the Dirichlet form.

$$\begin{aligned}\mathcal{E}(f,f) &= \frac{1}{2}\sum_{j=0}^{p-1} K(j,j+1)\pi(j)(f(j)-f(j+1))^2 \\ &\quad + \frac{1}{2}\sum_{j=1}^{p} K(j,j-1)\pi(j)(f(j)-f(j-1))^2 \\ &= \frac{1}{4p}\sum_{j=0}^{p-1}(f(j)-f(j+1))^2 + \frac{1}{4p}\sum_{j=1}^{p}(f(j)-f(j-1))^2 \\ &= \frac{1}{2p}\sum_{j=0}^{p-1}(f(j)-f(j+1))^2.\end{aligned}$$

For $v, w \in \mathbb{Z}_p$ choose a shortest path γ connecting the vertices v and w. Let $i, i+1, ..., i+k$ be vertices on the chosen path, then

$$|f(v) - f(w)| \leq \sum_{j=i}^{k-1} |f(j) - f(j+1)|.$$

For an arbitrary function $\tilde{f}$ let $f := sgn(\tilde{f} - c)(\tilde{f} - c)^2$ where c is a median of $\tilde{f}$. In the sequel we need the function under consideration to be not of constant sign. In the case that $\tilde{f}$ is of constant sign, we take f as defined above.
We have

$$\|f\|_\infty \leq \sum_{j=0}^{p-1} |f(j) - f(j+1)|.$$

Now, take u, such that $\|f\|_\infty = f(u)$ and v such that $f(v)f(u) \leq 0$. Thus, $|f(u) - f(v)| \geq |f(u)|$. Choose a function g such that g has median 0 ($\pi(g > 0) \leq 1/2$ and $\pi(g < 0) \leq 1/2$) and define $f := sgn(g)g^2$. Clearly, f changes sign and

$$\begin{aligned} |f(i+1) - f(i)| &= |sgn(g(i+1))g(i+1)^2 - sgn(g(i))g(i)^2| \\ &\leq |g(i+1) - g(i)|(|g(i+1)| + |g(i)|). \end{aligned}$$

Hence, by the definition of g and applying the triangle and the Cauchy-Schwarz inequality leads to

$$\begin{aligned} \|f\|_\infty &\leq \sum_{j=0}^{p-1} |f(j) - f(j+1)| \\ &\leq \sum_{j=0}^{p-1} |g(j+1) - g(j)|(|g(j+1)| + |g(j)|) \\ &\leq \left(\sum_{j=0}^{p-1} |g(j+1) - g(j)|^2 \right)^{1/2} \left(\sum_{j=0}^{p-1} (|g(j+1)| + |g(j)|)^2 \right)^{1/2} \\ &\leq \sqrt{2}p\sqrt{\mathcal{E}(g,g)}\|g\|_2. \end{aligned}$$

Since $\|f\|_\infty = \|g\|_\infty^2$, we have

$$\|g\|_2^4 \leq \|g\|_\infty^2 \|g\|_1^2 \leq \sqrt{2}p\sqrt{\mathcal{E}(g,g)}\|g\|_2\|g\|_1^2.$$

These considerations lead to

$$\|g\|_2^6 \leq 2p^2\mathcal{E}(g,g)\|g\|_1^4.$$

Now, we use the fact that if c is a median for f, $g := f - c$ has median 0. Hence,

$$\begin{aligned} \|f-c\|_2^6 &\leq 2p^2\mathcal{E}(f-c, f-c)\|f-c\|_1^4 \\ &\leq 2p^2\mathcal{E}(f,f)\|f\|_1^4. \end{aligned}$$

Owing to the variational characterization of Var_π this leads to

$$Var_\pi[f]^6 \leq 2p^2\mathcal{E}(f,f)\|f\|_1^4,$$

which is a Nash inequality of type I.
Using the above Theorem, for $r \in \{0, ..., (p-1)/2\}$ and $x \in \mathbb{Z}_p$ we have to compute $V(x,r)$. By elementary calculations we observe

$$2\frac{r+1}{p} \geq V(x,r) = \frac{2r+1}{p} \geq \frac{r+1}{p}.$$

Thus, $\mathbb{Z}_p$ has $(p,1)$-moderate growth. The constant appearing in the local Poincaré inequality can be estimated from above:

$$\begin{aligned} \eta(r) &= \max_{e\in\mathcal{A}}\left\{\frac{2}{Q(e)}\sum_{|\gamma_{xy}|\leq r}|\gamma_{xy}|\frac{\pi(x)\pi(y)}{V(x,r)}\right\} \\ &\leq 4p\sum_{|\gamma_{xy}|\leq r}|\gamma_{xy}|\frac{1}{p2(r+1)} \\ &\leq \frac{2r}{r+1}\sum_{|\gamma_{xy}|\leq r}1 \leq r^2. \end{aligned}$$

Therefore, we have

$$\|f-f_r\|_2^2 \leq r^2\mathcal{E}(f,f).$$

We obtain the Nash inequalities

$$Var_\pi[f]^3 \leq 16p^2\mathcal{E}(f,f)\|f\|_1^4$$

and

$$\|f\|_2^6 \leq 16p^2\left(\mathcal{E}(f,f+\frac{4}{p^2}\|f\|_2^2\right)\|f\|_1^4.$$

This example shows that the Nash inequalities obtained using moderate growth and local Poincaré inequalities do not lead to very sharp constants. But nevertheless, these methods can often be applied when no bare hands on calculations are possible.

4.4 Examples

We state examples to illustrate our lower bounds on the higher eigenvalues obtained in the previous sections and compare them not only to lower bounds obtained by [ChY95] and [SC 97] but also to the exact values if available.

Example 1

Let $\mathcal{X} = \{-n, ..., n\}$ and $K(x, y) = 0$ unless $|x - y| = 1$ or $x = y = n$ or $-n$. In the latter cases $K(x, y) = \frac{1}{2}$. The chain has $\pi(i) = \frac{1}{2n+1}$ as its stationary measure.
The following Nash inequality holds true by [SC 97], Example 2.3.1,

$$Var_\pi[f]^3 \leq 2(2n + 1)^2 \mathcal{E}(f, f) \|f\|_1^4.$$

Thus, we can use $C = 2(2n + 1)^2$, $d = 1$ in Theorem 4.1.1 and obtain

$$\lambda_k \geq \frac{k^2}{e(2n + 1)^2}, \quad k \in \{0, 1, ..., |\mathcal{X}| - 1\}.$$

The lower bounds for λ_k, $k \in \{0, 1, ..., |\mathcal{X}| - 1\}$ obtained by Saloff-Coste in [SC 97] are improved by a factor e^3.
The exact values of λ_k, $k \in \{0, 1, ..., |\mathcal{X}| - 1\}$ are

$$\lambda_k = 1 - \cos\left(\frac{\pi k}{2n + 1}\right), \quad k \in \{0, 1, ..., |\mathcal{X}| - 1\}$$

and therefore compare well with our lower bounds.

Example 2

Take the square grid $\mathcal{X} = \{0, ..., n\}^2$ and the transition kernel

$$K(x, y) = \begin{cases} \frac{1}{4} & , \ |x - y| = 1 \\ 0 & , \ |x - y| > 1 \\ 0 & , \ x = y, \ x \text{ interior point} \\ \frac{1}{2} & , \ x = y, \ x \text{ side point} \\ \frac{1}{4} & , \ x = y, \ x \text{ corner point.} \end{cases}.$$

For the stationary distribution we have $\pi(x) = \frac{1}{(n+1)^2}$ for all $x \in \mathcal{X}$.
The following Nash inequality holds true, see [SC 97], Theorem 3.3.14,

$$Var_\pi[f]^2 \leq 64(n + 1)^2 \mathcal{E}(f, f) \|f\|_1^2.$$

Thus, we obtain applying Theorem 4.1.1

$$\lambda_k \geq \frac{k+1}{64(n+1)^2 e}, \quad k \in \{0, 1, ..., |\mathcal{X}| - 1\}$$

This improves the result in [SC 97], Example 2.3.5 by a factor 2.
The spectrum in this case are the numbers

$$1 - \frac{1}{2}\left(\cos\left(\frac{\pi k}{n+1}\right) + \cos\left(\frac{\pi l}{n+1}\right)\right), \quad k, l \in \{0, 1, ..., |\mathcal{X}| - 1\}, k + l \leq |\mathcal{X}| - 1.$$

These are distributed like $\frac{l^2+k^2}{(n+1)^2}$, $l, k \in \{0, 1, ..., |\mathcal{X}| - 1\}$ and $|\{(l, k) \in \{0, 1, ..., |\mathcal{X}| - 1\}^2 : l^2 + k^2 \leq j\}| \approx j$. Therefore, our bounds are of the right order of magnitude compared to the spectrum.

Example 3

Suppose (K, π) is reversible and satisfies a Nash inequality of type I with $C = \frac{A}{\lambda_1}$, where λ_1 denotes the spectral gap. Then we have

$$\lambda_k \geq \frac{\lambda_1}{Ae} k^{\frac{2}{d}}, \quad k \in \{0, 1, ..., |\mathcal{X}| - 1\}.$$

This improves the result in [SC 97], Example 2.3.3, by a factor $de^{(2-d)/d}$.

Example 4

Take $\mathcal{X} = \{1, ..., n\}$ and for $2 \leq i \leq n-1$, $K(i, i+1) = K(i, i-1) = 1/2$, for $j \in \{2, n\}$, $K(1, j) = 1/2$, $K(n, n-1) = 1$ and $K(i, j) = 0$ otherwise. The stationary measure π is given by $\pi(i) = \frac{2i}{n^2}$ for $i \neq n$ and $\pi(n) = \frac{1}{n}$. Easy calculations yield

$$V(x, r) \geq \frac{1}{4}\left(\frac{r+1}{n-1}\right)^2 \text{ and } \|f - f_r\| \leq 24r^2\mathcal{E}(f, f).$$

This means it has $(4, 2)$-moderate growth and satisfies a local Poincaré inequality. By [DS-C 96b] this implies a Nash inequality of type II, thus

$$\|f\|_2^4 \leq 641(n-1)^2\left[\mathcal{E}(f, f) + \frac{1}{24n^2}\|f\|_2^2\right]\|f\|_1^2.$$

Therefore, we have

$$\text{for } k \geq \frac{\sqrt{641}(n-1)}{4\sqrt{2}\sqrt{3}n}e - 1 \geq 7 : \lambda_k \geq \frac{2(k+1)^2}{641(n-1)^2 e^4}$$

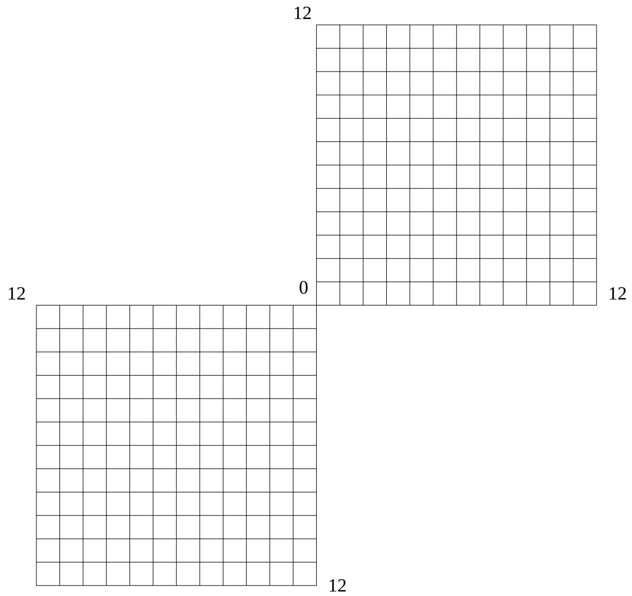

Figure 4.1: **The 'n-dog' for** $n = 12$**.**

and

$$\text{for } k \geq \frac{\sqrt{641e}(n-1)}{4\sqrt{2}\sqrt{3}n} e - 1 \geq 11 : \lambda_k \geq \frac{(k+1)^2}{641(n-1)^2 e^3}.$$

Let us compare this result with the bound using Theorem 4.3.2. Due to Theorem 4.3.2 the $(4,2)$-moderate growth and the local Poincaré inequality from above imply a Nash inequality of type I. In this case we have for all $k \in \{0, ..., |\mathcal{X}| - 1\}$

$$\lambda_k \geq \frac{k^2}{3298(n-1)^2} = \frac{k^2}{2^3 \cdot 432(n-1)^2}.$$

This not only improves upon the lower bound but also holds for all $k \in \{0, ..., |\mathcal{X}| - 1\}$.

Example 5

Let $\mathcal{X}$ be the so called 'n-dog'. $\mathcal{X}$ is the vertex set of two $n \times n$ square grids $\{0, ..., n\}^2$ and $\{-n, ..., 0\}^2$ attached by identification of the two corners $o = (0,0) \in \mathcal{X}$ such that $|\mathcal{X}| = 2(n+1)^2 - 1$. See Figure 4.1. We consider the Markovian transition kernel

$$K(x,y)=\begin{cases} 0 & \text{if } |x-y|>1 \\ 1/4 & \text{if } |x-y|=1 \\ 0 & \text{if } x=y \text{ is a point inside or } x=y=0 \\ 1/4 & \text{if } x=y \text{ is a point on the boundary but not a corner point} \\ 1/2 & \text{if } x=y \text{ is a corner point.} \end{cases}$$

The stationary measure is the equidistibution on $\mathcal{X}$, i.e. for $x\in\mathcal{X}$, $\pi(x)=\frac{1}{2(n+1)^2-1}$.
By [SC 97], Example 3.3.5, we have

$$\|f\|_2^4 \le 128(n+1)^2\left(\mathcal{E}(f,f)+\frac{1}{32(n+1)^2}\|f\|_2^2\right)\|f\|_1^2.$$

For $k\ge 11$ we find using Theorem 4.1.1 $\lambda_k \ge \frac{k+1}{128e^2(n+1)^2}$ and for $k\ge 35$: $\lambda_k \ge \frac{k+1}{e^2 128(n+1)^2}$.
This improves upon the result in [SC 97] by a factor $\frac{9(n+1)^2}{n^2}$ and there only eigenvalues λ_k with $k\ge 10000$ can be considered.

Example 6

As in Example 2 let $\mathcal{X}=\{0,...,n\}^2$ and K be the natural random walk on $\mathcal{X}$. Then we have for the 2-dimensional isoperimetric constant, see [SC 97]

$$I_2=\begin{cases} \frac{1}{2^{3/2}(n+1)} & ,\quad n+1 \text{ is even} \\ \frac{1}{2^{3/2}\sqrt{n}(n+1)} & ,\quad n+1 \text{ is odd.} \end{cases}$$

Therefore, for $k\in\{0,1,...,|\mathcal{X}|-1\}$ Theorem 4.2.4 yields

$$\lambda_k \ge \begin{cases} \frac{k}{64e(n+1)^2} & ,\quad n+1 \text{ is even} \\ \frac{k}{64en(n+1)^2} & ,\quad n+1 \text{ is odd.} \end{cases}$$

This lower bound improves upon the one obtained in [ChY95] by a factor $288(n+1)^2$ and $288(n+1)$, respectively.

Example 7

For $d\ge 1$ consider $\mathcal{X}=\{0,1,...,n_1\}\times ...\times\{0,1,...,n_d\}$ with $n_1\ge n_2\ge ...\ge n_d$ and the natural random walk.
We have for the d-dimensional isoperimetric constant, using results by [SC 97],

$$I_d \ge \frac{1}{d2^{1-1/d}(n_1+1)}.$$

Therefore, again applying Theorem 4.2.4,

$$\lambda_k \geq \frac{k^{2/d}}{d^2 2^{2-2/d}(n_1+1)^2 8e}.$$

Thus, our bounds improve upon those in [ChY95] by a factor $\frac{(n_1+1)^2 9^{2/d} d^4}{2(d-1)^2}$.

Remark 4.4.1 *Our methods to obtain lower bounds on higher eigenvalues for a Markov chain on a finite state space $\mathcal{X}$ can also be used to examine random walks on finite groups.*
In these cases moderate growth can often be established, see for instance [DS-C 94] and [DS-C 96b].

Together with the following lemma which is due to Diaconis and Saloff-Coste and proved in [DS-C 96b] we can apply Theorem 4.1.1 for random walks on finite groups.

Lemma 4.4.2 *Let G be a finite group and E a set of generators. Define a random walk on G as*

$$K(x,y) := \begin{cases} 1/|E| & : \ yx^{-1} \in E \\ 0 & : \ yx^{-1} \notin E. \end{cases}$$

The stationary measure is the equidistribution, i.e. $\pi(x) = 1/|G|$. Consider the Cayley graph with edge set $\mathcal{A} := \{(x, sx) : x \in G, s \in E \cup E^{-1}\}$. If the set of generators E is symmetric, i.e. $E = E^{-1}$ we have

$$\|f - f_r\|_2^2 \leq 2|E|r^2 \mathcal{E}(f,f) \text{ for all integers } r$$

and otherwise

$$\|f - f_r\|_2^2 \leq 4|E|r^2 \mathcal{E}(f,f) \text{ for all integers } r.$$

Example 8

For a natural number m we consider the field $\mathbb{Z}_m$ and all 3×3 matrices of the form

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}, \quad x, y, z \in \mathbb{Z}_m.$$

If we take this set together with the normal matrix multiplication we obtain a group structure. Actually, this group is called the Heisenberg group and is

denoted by $U_3(m)$.
The set of all 3×3 matrices is generated by the set

$$E = \left\{ Id, \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \right\}.$$

Diaconis and Saloff-Coste show in [DS-C 94] that

$$V(x,r) \geq \frac{(r+1)^3}{48\gamma^3} = \frac{(r+1)^3}{48(m+1)^3}.$$

Since $E = E^{-1}$ applying Lemma 4.4.2 we get

$$\|f - f_r\|_2^2 \leq 2|E|r^2\mathcal{E}(f,f) = 10r^2\mathcal{E}(f,f).$$

Thus, using Theorem 4.3.2 we have the Nash inequality

$$Var_\pi[f]^{1+2/3} \leq 1491(m+1)^2\mathcal{E}(f,f)\|f\|_1^{4/3}.$$

Now, applying Theorem 4.1.1 leads to

$$\lambda_k \geq \frac{k^{2/3}}{1491e(m+1)^2}.$$

Example 9

We consider the disrete circle $\mathbb{Z}_p$, p odd. As we saw the Nash inequality

$$Var_\pi[f]^6 \leq 2p^2\mathcal{E}(f,f)\|f\|_1^4$$

holds true. Therefore, Theorem 4.1.1 implies for all $k \in \{0,1,...,p-1\}$

$$\lambda_k \geq \frac{k^2}{2ep^2}.$$

The our lower bounds are of the right order of magnitude since the eigenvalues are

$$\left(\lambda_k = 1 - \cos\left(\frac{\pi k}{p}\right)\right)_{k\in\{0,1,..,p-1\}}.$$

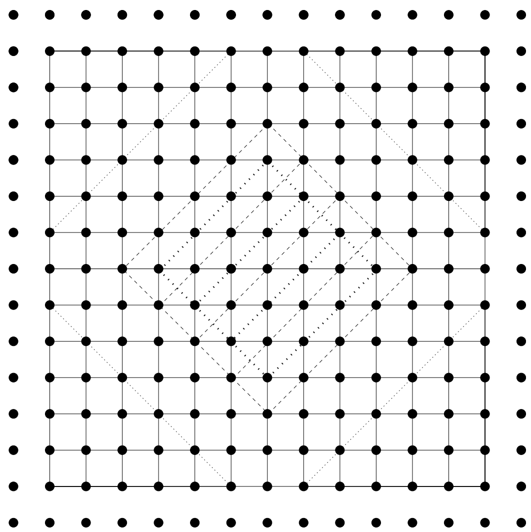

Figure 4.2: **The discrete torus $\mathbb{Z}_p^2$. The figure illustrates the calculation of the volumes** $V(x,r)$

Example 10

We consider the discrete torus $\mathbb{Z}_p^d$, p odd, and the nearest neighbour random walk on it.

$$K(x,y) = \begin{cases} 1/(2d) & , \quad \text{if adjacent,} \\ 0 & , \quad \text{otherwise.} \end{cases}$$

The stationary measure is $\pi(x) = \frac{1}{p^d}$. As shown in Figure 4.2 we find for the Volume $V(x,r)$: $r \leq (p-1)/2 : V(x,r) = (r+1)^d + r^d \cdot 1/p^d$ and $\gamma \geq r > (p-1)/2 : V(x,r) \geq (r+1)^d/p^d$. The latter holds true because the second summand in each step reduces till it reaches 0. Thus,

$$\frac{2(r+1)^d}{p^d} \geq V(x,r) \geq \frac{(r+1)^d}{p^d}.$$

For the constant $\eta(r)$ we find

$$\begin{aligned} \eta(r) &= \max_{e \in \mathcal{A}} \left\{ \frac{2}{Q(e)} \sum_{|\gamma| \leq r} \frac{\pi(x)\pi(y)}{V(x,r)} \right\} \\ &\leq 4dp^d \max_{e \in \mathcal{A}} \sum_{|\gamma| \leq r} |\gamma| \frac{1}{p^d(2r+1)^d} \\ &= 4d \frac{1}{(2r+1)^d} \max_{e \in \mathcal{A}} \sum_{|\gamma| \leq r} |\gamma| \end{aligned}$$

$$\leq \; 4d\left(\frac{2r}{2r+1}\right)^d 4r^2 \leq 16dr^2.$$

Let e be an edge in $\mathbb{Z}_p^d$. To bound the number of paths of length at most r that use this edge, suppose the edge appears after time $i-1$ and before time i, $i \in \{1,...,d\}$. There are at most r^i starting points x and at most $(2r)^{d+1-i}$ ending points y. Hence, bounding $|\gamma|$ by r yields the estimate in the last inequality. By Theorem 4.1.1 we have

$$\lambda_k \geq \frac{k^{2/d}}{256dp^d}.$$

Because we already observed that the eigenvalues of the discrete circle $\mathbb{Z}_p$ are $\left(1-\cos\left(\frac{\pi k}{p}\right)\right)_{k\in\{0,1,\ldots,p\}}$, the eigenvalues of the $d-$dimensional discrete torus are

$$\lambda_{i_1,\ldots,i_d} = \frac{1}{d}\sum_{l=1}^{d}\left(1-\cos\left(\frac{\pi i_l}{p}\right)\right),\; 0 \leq i_l \leq p-1.$$

Let us now use symmetry arguments to prove a Nash inequality. The Dirichlet form in this case can be computed and we find

$$\begin{aligned}
\mathcal{E}(f,f) &= \frac{1}{2}\left(\sum_{j=1}^{p^2-1}|f(j)-f(j-1)|^2K(j,j-1)\pi(j)\right. \\
&\quad +\sum_{j=0}^{p^2}|f(j)-f(j+1)|^2K(j,j+1)\pi(j) \\
&\quad +\sum_{j=1}^{p^2+p-1}|f(j)-f(j-p)|^2K(j,j-p)\pi(j) \\
&\quad \left.+\sum_{j=0}^{p^2-p-1}|f(j)-f(j+p)|^2K(j,j+p)\pi(j)\right) \\
&= \frac{1}{4p^2}\left(\sum_{j=0}^{p^2-1}|f(j)-f(j+1)|^2+\sum_{j=0}^{p^2-p-1}|f(j)-f(j+p)|^2\right).
\end{aligned}$$

For $v=\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, w=\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \mathbb{Z}_p^2$, choose a path γ as follows:
With no loss of generality we can assume $v_1 \leq w_1$ and $v_2 \leq w_2$ with one

inequality to be strict.

Start at v and proceed by taking the vertices $\begin{pmatrix} v_1+1 \\ v_2 \end{pmatrix}$, $\begin{pmatrix} v_1+2 \\ v_2 \end{pmatrix}$,..., $\begin{pmatrix} v_1+k \\ v_2 \end{pmatrix} = \begin{pmatrix} w_1 \\ v_2 \end{pmatrix}$. Then, proceed by $\begin{pmatrix} w_1 \\ v_2+p \end{pmatrix}$, $\begin{pmatrix} w_1 \\ v_2+2p \end{pmatrix}$,..., $\begin{pmatrix} w_1 \\ v_2+lp \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$. The path γ is a shortest path by definition. Using γ we have

$$|f(v)-f(w)| \leq \sum_{j=0}^{k-1} |f(j+i)-f(j+i+1)| + \sum_{j=0}^{l-1} |f(jp+i+k)-f(i+k+(j+1)p)|.$$

For an arbitrary function $\tilde{f}$ let $f := sgn(\tilde{f}-c)(\tilde{f}-c)^2$, where c is a median for $\tilde{f}$. In the sequel we need the function under consideration to be not of constant sign. Therefore, if $\tilde{f}$ is of constant sign, we take f as defined above, we find

$$\|f\|_\infty \leq \sum_{j=0}^{p^2-1} |f(j)-f(j+1)| + \sum_{j=0}^{p^2-p-1} |f(j)-f(j+p)|.$$

As in the case of a discrete circle we choose a function g such that g has median 0 and define $f := sgn(g)g^2$. Applying the Cauchy-Schwarz inequality yields

$$\begin{aligned}
\|f\|_\infty &\leq \sum_{j=0}^{p^2-1} |g(j+1)-g(j)|(|g(j+1)|+|g(j)|) \\
&\quad + \sum_{j=0}^{p^2-p-1} |g(j)-g(j+p)|(|g(j+p)|+|g(j)|) \\
&\leq \left(\sum_{j=0}^{p^2-1} |g(j+1)-g(j)|^2\right)^{1/2} \left(\sum_{j=0}^{p^2-1} (|g(j+1)|+|g(j)|)^2\right)^{1/2} \\
&\quad + \left(\sum_{j=0}^{p^2-p-1} |g(j+p)-g(j)|^2\right)^{1/2} \left(\sum_{j=0}^{p^2-p-1} (|g(j+p)|+|g(j)|)^2\right)^{1/2} \\
&\leq 4p^2\sqrt{\mathcal{E}(g,g)}\|g\|_2.
\end{aligned}$$

Summarizing the previous considerations we have

$$\|g\|_2^4 \leq \|g\|_\infty^2\|g\|_1^2 \leq 4p^2\sqrt{\mathcal{E}(g,g)}\|g\|_2\|g\|_1^2.$$

Consequently, following the same lines of thought as in the case of the discrete circle we find the Nash inequality

$$Var_\pi[f]^6 \leq 16p^4 \mathcal{E}(f,f)\|f\|_1^4$$

and therefore the eigenvalue bounds

$$\lambda_k \geq \frac{k^2}{16ep^4}.$$

The argument can be applied to $d-$dimensional tori $\mathbb{Z}_p^d$ as well and we find the Nash inequality

$$Var_\pi[f]^6 \leq 2^d d^2 p^{2d} \mathcal{E}(f,f)\|f\|_1^4.$$

By Theorem 4.1.1 this implies the eigenvalue bounds

$$\lambda_k \geq \frac{k^2}{2^d e p^{2d} d^2}.$$

Example 11

Suppose for all $t > 0$, $U_t : \mathcal{X} \to \mathbb{R}$ is a function. We consider the well known Simulated Annealing algorithm with time-dependent energy function U_t. The idea applying this kind of algorithm is to choose a Markov process on a finite state space $\mathcal{X}$ in such a way that the transition kernel at time t has at its invariant measure the Gibbs distribution with energy U_t and temperature β_t^{-1}, i.e.

$$\pi_t(x) := \frac{\exp(-\beta_t U_t(x))}{Z_t}\pi(x). \tag{4.9}$$

Here $Z_t := \sum_{x\in\mathcal{X}} \exp(-\beta_t U_t(x))\pi(x)$ is the partition function.
Let (K,π) be an irreducible transition kernel on $\mathcal{X}\times\mathcal{X}$. Moreover, let $t \mapsto \beta_t$ be a monotone increasing function which is strictily positive. We define a transition kernel on $\mathcal{X} \times \mathcal{X}$ by

$$K_t(x,y) = \begin{cases} \exp(-\beta_t(U_t(y) - U_t(x))^+)K(x,y) & , \quad x \neq y \\ 1 - \sum_{z\neq y} K_t(x,z) & , \quad x = y. \end{cases} \tag{4.10}$$

K_t is reversible w.r.t. the Gibbs measure π_t as defined in (4.9). We denote by $(\lambda_{k,t})_k$ the eigenvalues of the chain at time t. Let us assume that U_t fulfills the following conditions:

1. For a constant $M > 0$ we have for all $x \in \mathcal{X}$ and $t \geq 0$

$$0 \leq U_t(x) \leq M.$$

2. For constants $a, C > 0$ and for all $x \in \mathcal{X}$ we have

$$\frac{\partial}{\partial t} U_t(x) \leq \frac{C}{(1+t)^\alpha}.$$

3. For a constant $b \geq 0, m > 0$ for all $t \geq 0$ we have

$$\lambda_{1,t} \geq b \exp(-\beta_t m).$$

Under these assumption, for $d > 2$ the corresponding process satisfies the Sobolev-inequality

$$\|f - \mathbb{E}_{\pi_t}[f]\|^2_{2d/(d-2),t} \leq \frac{1}{b\pi_*^{2/d}} \exp\left(\beta_t \left(m + 2M/d\right)\right) \mathcal{E}_t(f,f). \tag{4.11}$$

Here we used the notation $\pi_* := \min_{x \in \mathcal{X}} \pi(x)$. The proof follows the same lines of thought as the proof of Theorem 2.11 in [HS88] and was carried out in [L 96]. Since Hölder's inequality implies a Nash inequality with the same constants, as in (4.11), we can apply Theorem 4.1.1 and obtain

$$\lambda_{k,t} \geq \frac{b\pi_*^{2/d} k^{2/d}}{\exp\left(1 + \beta_t \left(m + 2M/d\right)\right)}. \tag{4.12}$$

This shows that under the same circumstances as for estimating the spectral gap one can obtain lower bounds for the higher eigenvalues of the time dependent operators. The appearing constants are the same as for the spectral gap estimates.

Bibliography

[AD 87] *Aldous D., Diaconis P.*, (1987), Strong uniform times and finite fandom walks, Ad. Appl. Math. 8, 69-97.

[B 92] *Bakry D.*, (1992), L'hypercontractivité et son utilisation en théorie des semigroups. École d'Été de Saint Flour 1992. Lecture Notes Math. 1581. Springer, Berlin.

[BCLSC 95] *Bakry D., Coulhon T, Ledoux M, Saloff-Coste L.*, (1995), Sobolev inequalities in disguise, Ind. U. Math. J. 44, 4, 1043-1074.

[CKS87] *Carlen E., Kusuoka S., Stroock D.*, (1987), Upper bounds for symmetric Markov transition functions, Ann. Inst. H. Poincaré, Prob. Stat. 23, 245-287.

[C 70] *Cheeger J.*, (1970), A lower bound for the smallest eigenvalue of the Laplacian, in: Problems in Analysis, Princeton Univ. Press, 195-199.

[ChW 94] *Chen M. F., Wang F. Y.*, (1994), Application of coupling to the first eigenvalue on manifold, Sci. Sin. (A), 37, 1, 1-14.

[Ch 94] *Chen M. F.*, (1994), Optimal couplings and application to Riemannian geometry, Prob. Theory and Math. Stat., Vol. 1, VPS/TEV, 121-142.

[Ch 98] *Chen M. F.*, (1998), Triology of couplings and general formulas for lower bound of spectral gap, in 'Probability towards 2000', Lecture Notes in Statis. 128, Springer, 123-136.

[ChY95] *Chung F.R.K., Yau S.T.* , (1995), Eigenvalues of graphs and Sobolev inequalities, Comb. Prob. Comp. 4, 11-26.

[DeS 89] *Deuschel J. D., Stroock D.*, Large deviations. Academic Press, New York, 1989.

[DS 91] *Diaconis P., Stroock D.*, (1991), Geometric bounds for eigenvalues of Markov chains, Ann. Appl. Prob., 1, 36-61.

[DS-C 93] *Diaconis P., Saloff-Coste L.*, (1993) Comparison theorems for reversible Markov chains, Ann. Appl. Prob. 3, 696-730.

[DS-C 94] *Diaconis P., Saloff-Coste L.*, (1994), Moderate growth and random walk on finite groups, Geom. Func. Anal. 4, 1-34.

[DS-C 96a] *Diaconis P., Saloff-Coste L.*, (1996), Logarithmic Sobolev inequalities for finite Markov chains, Ann. Appl. Prob., 3, 695-750.

[DS-C 96b] *Diaconis P., Saloff-Coste L.*, (1996), Nash inequalities for finite Markov chains, J. Theo. Prob., 9(2), 459-510.

[Do 38] *Doeblin W.*, (1938), Exposé de la théorie des chaines simples constantes de Markov à un numbre fini d'états, Rev. Math. de l'Union Interbalkanique 2, 77-105.

[FW 99] *Fulman J., Wilmer E.*, (1999), Comparing eigenvalue bounds for Markov chains: When does Poincaré beat Cheeger?, Ann. Appl. Probab. 9, no.1, 1-13.

[GT 83] *Gilborg D., Trudinger N.*, Elliptic partial differential equations of second order, (2nd ed.), Springer, 1983.

[G 76] *Gross L.*, (1976), Logarithmic Sobolev inequalities, Amer. J. Math. 97, 1061-1083.

[G 93] *Gross L.*, Logarithmic Sobolev inequalities and contractivity properties of semigroups. Lecture Notes Math. 1563. Springer, Berlin, 1993.

[HS87] *Holley R., Stroock D.*, (1987), Logarithmic Sobolev inequalities and stochastic Ising models, J. Stat. Phys., 46, 1159-1194.

[HS88] *Holley R., Stroock D.*, (1988), Simulated annealing via Sobolev inequalities, Comm. Math. Phys. 115, 553-569.

[H 00] *Houdré C.*, (2000), Mixed and isoperimetric esitmates on the Log-Sobolev constant of graphs and Markov chains, Preprint 2000.

[JS 88] *Jerrum M., Sinclair A.*, (1988), Approximating the permanent, SIAM J. Comp., 18, 1149-1178.

[JS 89] *Jerrum M., Sinclair A.*, (1989), Approximate counting, uniform generation and rapidly mixing Markov chains, Inf. Comp., 82, 93-113.

[K 00a] *Kienitz J.*, (2000) Spectral estimates: From finite to general state spaces, submitted

[K 00b] *Kienitz J.*, (2000), A note on higher eigenvalues for finite Markov chains, submitted.

[L 92] *Lindvall T.*, Lectures on the coupling method. Wiley, New York, 1992.

[LS 88] *Lawler G. F., Sokal A. D.*, (1988), Bounds on the L^2 spectrum for Markov chains and Markov processes: a generalization of Cheeger's inequality, Trans. Amer. Math. Soc. 309, 557-580.

[L 96] *Löwe, M.*, (1996), Simulated annealing with time-dependent energy function via Sobolev inequalities. Stochastic Proc. Appl. 63, 2, 221-233.

[M 85] *Maz'ja V.*, Sobolev spaces, Springer 1985.

[Me 96] *Meise C.*, Spectral estimates and genetic algorithms. PhD dissertation University of Bielefeld, 1996.

[Me 96] *Meise, C.*, (1996), Lower Bounds for the Spectral Gap of Markov chains, SFB343 Preprint 96-019.

[Me97] *Meise C.*, (1997), Spectral gap estimates via distance arguments, submitted.

[Me 99] *Meise C.*, (1999), On spectral gap estimates of a Markov chain via hitting times and coupling, J. Appl. Prob. 36, 310-319.

[MT 96] *Meyn S.P., Tweedie R.L.*, Markov Chains and stochastic stability. Springer (2nd ed.), 1996.

[N 58] *Nash J.*, (1958), Continuity of solutions of parabolic and elliptic equations, Amer. J. Math. 80, 931-954.

[P61] *Pólya G.*, (1961) On the eigenvalues of vibrating membranes, Proc. London Math. Soc. 11, 419-433.

[Ro 96] *Rosenthal J. S.*, (1996), Markov chain convergence: From finite to infinite, Stoch. Proc. Appl., 62, 55-72.

[R 85] *Rothaus O.*, (1985), Analytic inequalities, isoperimetric inequalities and logarithmic Sobolev inequalities, J. Funct. Anal., 64, 296-313.

[SC 97] *Saloff-Coste L.*, (1997), Lectures on finite Markov-chains. École d'Été de Saint Flour 1994. Lecture Notes Math., 1665. Springer, Berlin, 301-413.

[S 92] *Sinclair A.*, (1992), Improved bounds for mixing rates of Markov chains and multicommodity flow, Comb. Prob. Comp. 1, 351-370.

[S 93] *Sinclair A.*, Algorithms for random generation and counting: A Markov chain approach, Birkhäuser 1992.

[St 93] *Stroock D. W.*, (1993), Logarithmic Sobolev inequalities for Gibbs states. Springer, Lecture Notes in Math. 1563.

[V 84] *Varopoulos N.*, (1984), Une généralisation du théoreme de Hardy-Littlewood-Sobolev pour les espaces de Dirichlet, C.R. Acad. Sci. Paris Ser. I 299, 651-654.

[W 11] *Weyl H.*, (1911), Über die asymptotische Verteilung der Eigenwerte, Gött. Nach., 110-117.

[Y 00] *Yuen W.K.*, (2000), Applications of geometric bounds to the convergence rate of Markov chains on $\mathbb{R}^n$, Stoch. Proc. Appl. 87, 1-23.