

An Application of Harnack Inequalities to Random Walk on Nilpotent Quotients

Persi Diaconis and Laurent Saloff-Coste

RESUMÉ. On considère des marches aléatoires sur les quotients finis de groupes nilpotents finiment engendrés. On montre que ces marches atteignent leur état d'équilibre après un nombre d'itérations de l'ordre de γ^2 où γ est le diamètre du graphe de Cayley associé. La preuve repose sur une inégalité de Harnack obtenue par Hebisch et Saloff-Coste. En revanche, pour les quotients finis d'un groupe ayant la propriété T de Kazhdan, le temps d'atteinte de l'équilibre est de l'ordre de γ .

ABSTRACT. This paper shows that random walks on finite homogeneous spaces of nilpotent groups "get random" in order γ^2 steps where γ is the diameter of the associated Cayley graph. The argument uses a Harnack inequality of Hebisch and Saloff-Coste. In contrast, random walks on finite homogeneous spaces of groups satisfying Kazhdan's property T get random in order γ steps.

1. Introduction

We begin with an example of the problem under study. Let m be a positive integer, \mathbb{Z}_m the integers (mod m), and \mathbb{Z}_m^d the product space. A Markov chain on \mathbb{Z}_m^d is defined by choosing a coordinate at random and adding or subtracting the following coordinate. If the d th coordinate is chosen, ± 1 is added. More formally, if the process is currently at $x \in \mathbb{Z}_m^d$ choose $\epsilon = \pm 1$ with probability $\frac{1}{2}$ and independently an $i \in \{1, 2, \dots, d\}$ uniformly. If $1 \leq i < d$, the process goes to $(x_1, x_2, \dots, x_i + \epsilon x_{i+1}, x_{i+1}, \dots, x_d)$. If $i = d$, the process goes to $(x_1, x_2, \dots, x_d + \epsilon)$. This defines a Markov chain with transition kernel $k(x, y)$ on \mathbb{Z}_m^d . The chain is symmetric and ergodic

with the uniform distribution $u(x) = 1/m^d$ as its stationary distribution. The techniques introduced here allow us to study the convergence of the iterates k_x^n . We will prove the following.

1.1. Theorem

Fix d . There are constants $a_1, a_2, a_3, a_4 > 0$ such that

$$a_1 e^{-a_2 n/m^2} \leq \sup_{x \in X} \|k_x^n - u\|_{TV} \leq a_3 e^{-a_4 n/m^2} .$$

The constants a_i above depend on d but not on m . The result shows that order m^2 steps are necessary and sufficient to drive the variation distance to zero when d is fixed and m is large. The techniques to be introduced allow similar conclusions for some variations of the process. Instead of adding adjacent coordinates, any of a fixed collection of linear combinations can be added. The theory allows us to determine the right number of steps to achieve randomness for fixed d and large m .

The example treated in Theorem 1.1 analyzes a random walk on a homogeneous space. To introduce this in general, let G be a finite or finitely generated group and H be a subgroup. Let $X = G/H$ be the space of left cosets. Let S be a symmetric set of generators for G . Throughout, we assume that $id \in S$. Let

$$q(s) = \begin{cases} 1/|S| & \text{if } s \in S, \\ 0. & \end{cases} \tag{1.0}$$

The random walk on G generated by q induces a random walk on X by left multiplication. Formally, this has kernel

$$k(x, y) = q(yHx^{-1}) = \sum_{h \in H} q(yhx^{-1}). \tag{1.1}$$

Note that $k(x, y)$ is well defined. Iterates are defined by

$$k^n(x, y) = \sum_z k^{n-1}(x, z)k(z, y) .$$

Write k_x^n for $k^n(x, \cdot)$. In this definition, cosets $xH \in X$ are identified with coset representatives $x \in G$. Note further that $k(x, y) = k(y, x)$ and that $k(x, x) > q(id) > 0$. It follows that if X is finite, the associated chain is ergodic with $u(x) = 1/|X|$ as the unique stationary distribution.

Example 1.1. Here we put the initial example (Theorem 1.1) in a group-theoretic framework. Let $G = U_{d+1}(m)$ be the group of upper triangular $(d + 1) \times (d + 1)$ matrices with ones on the diagonal and entries (mod m) this is a group with $|U_{d+1}(m)| = m^{\binom{d+1}{2}}$. As generators, take

$$S = \{id, E_{i,i+1}(1), E_{i,i+1}(-1), 1 \leq i \leq d\} \tag{1.2}$$

with $E_{i,i+1}(a)$ having ones on the diagonal, an a in position $(i, i + 1)$, and zeros elsewhere. The effect of left multiplication by $E_{i,i+1}(a)$ is to add a times row $(i + 1)$ to row i . As a subgroup H take the matrices in $U_{d+1}(m)$ with zeros in the last column. The quotient $X = U_{d+1}/H$ can be identified with \mathbb{Z}_m^d and the walk induced on X is the chain of Theorem 1.1. We return to this example in Remark 2. \square

Random walks on homogeneous spaces can have very different rates of convergence from walks on G . The main results of this paper show that for nilpotent groups there is a relatively clean theory. To describe our result, introduce a graph with vertex set X and an edge from x to y if $y = sx$ for some $s \in S$. Thus, the edge set of our graph is $E = \{(x, sx) : x \in X, s \in S\}$. This graph may have self-loops and multiple edges. The induced walk is simply the natural random walk on this graph. Let γ_X denote the diameter. Our main result implies that order γ_X^2 steps are necessary and suffice for the random walk to reach its stationary distribution. Consider the subgroups $G_1 = G$, $G_i = [G_{i-1}, G]$, $i = 2, 3, \dots$. Here, $[x, y] = x^{-1}y^{-1}xy$ denotes the commutator of $x, y \in G$. Recall that G is nilpotent if and only if $G_{\ell+1} = \{id\}$ for some finite ℓ and that the smallest such ℓ is called the class of the nilpotent group G .

1.2. Theorem

Let G be a finitely generated nilpotent group, H a subgroup of finite index in G , and $S \subset G$ a symmetric set of generators containing the identity. For the random walk defined by (1.1) there exist constants $a_1, a_2, a_3, a_4 > 0$ such that

$$a_1 e^{-a_2 n / \gamma_X^2} \leq \sup_{x \in X} \|k_x^n - u\|_{T.V.} \leq a_3 e^{-a_4 n / \gamma_X^2}.$$

The constants a_i depend on $|S|$ and the class of G , but not otherwise on G, H , or S .

Remarks. 1. Here is a simple example illustrating Theorem 1.2. Take $G = \mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2}$ with $m_1 \leq m_2$. Take $S = \{(0, 0), (1, 0), (-1, 0), (0, 1), (0, -1)\}$. A random walk on G takes order m_2^2 steps to get close to the uniform distribution. Next, take $H = \{(0, s), s \in \mathbb{Z}_{m_2}\}$. The induced walk on G/H is simple random walk on \mathbb{Z}_{m_1} (with holding probability $3/5$). This takes order m_1^2 steps to get close to the uniform distribution. Thus, using the diameter of the quotient graph can make a big difference.

2. For Example 1.1, the group $U_{d+1}(m)$ is nilpotent (see, e.g., Suzuki [Su2, p. 22]). For H as given and fixed d , the diameter γ_X is of order m , so Theorem 1.1 follows from Theorem 1.2. This example is generalized in §6, which contains further examples and remarks.

3. The upper bound in Theorem 1.2 is proved in §2. The key tool is a Harnack inequality of Hebisch and Saloff-Coste that works for general discrete nilpotent groups. We use this to derive Harnack inequalities for quotients and, from these, bounds for eigenvalues and decay of convolution powers. The results proved are a bit more general, including quotients of groups of polynomial growth that need not be nilpotent. The lower bound in Theorem 1.2 is proved in §3. Section 4 relaxes the condition of nilpotency to polynomial growth. Section 5 develops results for groups having Kazhdan’s property T.

4. In [DS-C1, 2] we discuss broadly applicable techniques (comparison inequalities for eigenvalues and volume growth techniques) for bounding rates of convergence of random walks on groups. This paper gives a different proof of the main result of Diaconis and Saloff-Coste [DS-C2] for finite

nilpotent groups. The techniques of that paper work for a wider class of groups (groups with moderate growth such as finite affine groups). The present techniques work for homogeneous spaces. \square

2. The Upper Bound

This section proves the upper bound in Theorem 1.2. The main result is Theorem 2.4, which is stated at the end of the section.

2.1. Infinite Nilpotent Groups

While the focus of this paper is on finite groups, the proofs proceed by lifting to an infinite group. Let Γ be a finitely generated discrete group. Let S be a symmetric set of generators of Γ . Throughout we assume $id \in S$. Define the volume growth function as

$$V(n) = |S^n|.$$

Although $V(n)$ depends on the choice of generators, changing old generators for new gives comparable volume growth: for some $c > 0$

$$c^{-1}V_{\text{new}}(c^{-1}n) \leq V_{\text{old}}(n) \leq cV_{\text{new}}(cn).$$

A finitely generated group Γ has *polynomial volume growth* if there exist $A, B > 0$ such that

$$V(n) \leq An^B \quad \text{for all } n > 0.$$

As examples, \mathbb{Z}^d has polynomial volume growth with $B = d$. The Heisenberg group $U_3(\mathbb{Z})$ has polynomial volume growth with $B = 4$. Any finitely generated nilpotent group Γ has polynomial growth according to a Theorem of H. Bass [B]. He showed that for such groups there exist C and D such that

$$C^{-1}n^D \leq V(n) \leq Cn^D \quad \text{for all } n > 0$$

where

$$D = \sum_{i=1}^{\infty} i \operatorname{rank}(\Gamma_i / \Gamma_{i+1}).$$

Here $\Gamma_1 = \Gamma$ and $\Gamma_i = [\Gamma, \Gamma_{i-1}]$ are successive commutator subgroups and $\operatorname{rank}(\Gamma_i / \Gamma_{i+1})$ denotes the rank of the torsionfree part of the Abelian group Γ_i / Γ_{i+1} .

A celebrated Theorem of M. Gromov [Gr] says that any finitely generated group Γ having polynomial growth (2.0) contains a nilpotent group N of finite index and satisfies

$$C^{-1}n^D \leq V_\Gamma(n) \leq Cn^D \quad \text{for all } n > 0 \quad (2.1)$$

where

$$D = \sum_{i=1}^{\infty} i \operatorname{rank}(N_i/N_{i+1}).$$

Simple random walk on Γ proceeds by picking a generator each time at random. Hebisch and Saloff-Coste [HS-C] have carried out a detailed study of such walks. They proved the following Harnack type inequality [HS-C, Theorem 6.3], which is crucial in what follows.

2.1. Theorem

Let Γ be a finitely generated group. Fix a symmetric set of generators $S \subset \Gamma$ with $id \in S$. Assume that (Γ, S) has polynomial volume growth (2.1). Let q be the uniform distribution on S defined at (1.0) and Q the corresponding Markov operator. Then there exists $c > 0$ and a positive integer a such that for all $n > 0$, $g \in \Gamma$, and any sequence U_i , $1 < i < \infty$, of nonnegative functions satisfying $Q(U_i) = U_{i+1}$,

$$\sup_{B(g, \sqrt{n})} U_n \leq c \inf_{B(g, \sqrt{n})} U_{an}.$$

Here a and c depend only on the constants C, D in (2.1).

Remarks. 1. In typical applications, the functions U_i are taken to be $k^i(x, \cdot)$, the i th step transition kernels on a homogeneous space lifted to the group. These satisfy $QU_i = U_{i+1}$ by construction. Theorem 2.1 shows they are flattening out to constants in a well-specified sense.

2. Theorem 2.1 is an analog of Harnack's principle for parabolic equations. For example, let Δ be the Laplace operator $\Delta f = -\sum \frac{\partial^2 f}{\partial x_i^2}$ on \mathbb{R}^d . Let $U(t, x)$ be a nonnegative solution of the heat equation $(\partial_t + \Delta)U = 0$ in $\mathbb{R}_+ \times \mathbb{R}^d$. Then the Harnack inequality says for some $c > 0$, for all $t > 0$, $x \in \mathbb{R}^d$,

$$\sup_{B(x, \sqrt{t})} U(t, \cdot) \leq c \inf_{B(x, \sqrt{t})} U(2t, \cdot).$$

Moser [Mos] gives proofs and further discussion of Harnack inequalities for parabolic equations. The idea that a Harnack inequality can easily be projected from a covering space to its quotients was introduced in Varopoulos [V] for the study Laplace operators on nilpotent Lie groups. See also Varopoulos et al. [VS-CC]. For a simple random walk, the analog of Δ is $(I - Q)$. The sequence of functions U_i , $1 \leq i < \infty$, considered in Theorem 2.1 satisfies

$$U_{i+1} - U_i = -(I - Q)U_i.$$

This is the analog of $\partial_t U = -\Delta U$. \square

As a first application we bound the decay rate of Markov chains induced on Γ/H . Here H is a subgroup of Γ and the chain with transitions $k^n(x, y)$ is defined from a walk on Γ just as in (1.1). For example, if $\Gamma = \mathbb{Z}$, then $H = m\mathbb{Z}$, $\Gamma/H \cong \mathbb{Z}_m$, and the induced random walk is random walk on \mathbb{Z}_m . The next result, which is due to Hebisch and Saloff-Coste [HS-C, Theorem 10.2], does not require that H have finite index in Γ .

2.2. Theorem

Let (Γ, S) be as in Theorem 2.1. For any subgroup $H \subset \Gamma$, the chain k^n defined at (1.1) on $X = \Gamma/H$ satisfies

$$k^n(x, y) \leq \frac{c}{V(x, \sqrt{n})} \quad \text{for all } x, y \in X, \quad n \geq 0. \quad (2.2)$$

In (2.2), $V(x, n)$ is the volume of the ball in the graph (X, E) and c is the constant given by Theorem 2.1.

Proof. Fix $x \in X$ and consider $U_i(y) = k^i(x, y)$. Lift U_i to functions on Γ by $\tilde{U}_i = U_i \circ \pi$ where π is the canonical projection $\pi : \Gamma \rightarrow \Gamma/H$. Observe that \tilde{U}_i satisfies $\tilde{U}_{i+1} = Q\tilde{U}_i$ with Q the convolution operator for the uniform distribution on S . Indeed, for $z \in X$, let $\bar{z} \in \Gamma$ be a coset representative. Then

$$\begin{aligned} \tilde{U}_{i+1}(\bar{z}) &= U_{i+1}(z) = \sum_{t \in X} k(t, z) U_i(t) \\ &= \sum_{t \in X} \sum_{h \in H} q(\bar{z}(\bar{t}h)^{-1}) \tilde{U}_i(\bar{t}h) = \sum_{g \in \Gamma} q(\bar{z}g^{-1}) \tilde{U}_i(g). \end{aligned}$$

Thus Theorem 2.1 yields $c > 0$ and a positive integer a such that for any $\bar{y} \in \Gamma$ and all $n > 0$

$$\sup_{B(\bar{y}, \sqrt{n})} \tilde{U}_n \leq c \inf_{B(\bar{y}, \sqrt{n})} \tilde{U}_{an}. \quad (2.3)$$

Now, fix $y \in X$ and $\bar{y} \in \Gamma$ with $\pi(y) = \bar{y}$. By definition, any element in $B(y, \sqrt{n})$ can be lifted to an element in $B(\bar{y}, \sqrt{n})$ and so $\pi(B(\bar{y}, \sqrt{n})) = B(y, \sqrt{n})$. This together with (2.3) give

$$\sup_{B(y, \sqrt{n})} U_n \leq c \inf_{B(y, \sqrt{n})} U_{an}.$$

As a consequence, for any n and $y \in X$

$$U_n(y) \leq \frac{c}{V(y, \sqrt{n})} \sum_{z \in B(y, \sqrt{n})} U_{an}(z) \leq \frac{c}{V(y, \sqrt{n})}.$$

This and the symmetry of $k^n(x, y)$ prove (2.2). \square

Remark. The constant c in Theorem 2.2 is the same as the constant in Theorem 2.1. This in turn is given effectively in terms of C, D of (2.1) by Hebisch and Saloff-Coste [HS-C]. They give c as exponential in D^2 . This renders the present results useless when D grows with the problem description (as in Example (1.1) with growing D).

2.2. Finite Quotients

In this section, Γ is a finitely generated group (countable or finite) having polynomial growth. H is a subgroup having finite index in Γ , so $X = \Gamma/H$ is a finite set. Let S be a symmetric set of generators of Γ containing the identity. Let $q(s) = \delta_S(s)/|S|$ be the uniform distribution on S , and let $k(x, y)$ be the Markov kernel induced on X as in (1.1). This k is a symmetric kernel and so has eigenvalues $1 > \beta_1 \geq \beta_2 \geq \dots \geq \beta_{|X|-1} > -1$. Let

$$\beta_* = \max\{|\beta_{|X|-1}|, |\beta_1|\}.$$

The first step in bounding the rate of convergence is to derive bounds on β_* .

2.3. Theorem

With notation as above

$$\beta_* \leq 1 - \frac{B}{\gamma_X^2}$$

where γ_X refers to the diameter of the graph (X, E) and $B = 1/(2\lceil a/2 \rceil c)$, with a and c as in Theorem 2.1. Thus B depends on the constants C, D in (2.1) but not otherwise on Γ, H , or S .

Proof. For any given $x \in X$ and positive integer n , there exists $x_0 \in X$ with $k^n(x, x_0) \geq \frac{1}{|X|}$. From this and Theorem 2.1 there is a positive integer a and $\epsilon > 0$ such that for all $x, y \in X$

$$k_N(x, y) \geq \frac{\epsilon}{|X|} \quad \text{for } N \geq a\gamma_X^2, \quad \text{with } \epsilon = 1/c. \quad \square \quad (2.4)$$

Theorem 2.3 follows from this and the following lemma (with $\tilde{k} = u$).

2.1. Lemma

Let k, \tilde{k} be two symmetric Markov kernels on a finite state space X . Suppose there exist $\epsilon > 0$ and positive integer j such that for all x, y

$$\epsilon \tilde{k}(x, y) \leq k_{2j}(x, y).$$

Then

$$\beta_* \leq 1 - \frac{\epsilon(1 - \tilde{\beta}_1)}{2j}.$$

Proof. Consider the Dirichlet form associated to \tilde{k} :

$$\tilde{\mathcal{E}}(f|f) = \langle (I - \tilde{K})f|f \rangle = \frac{1}{2} \sum_{x,y} (f(x) - f(y))^2 \tilde{k}(x, y).$$

Set $\mathcal{E}_i(f|f) = \langle (I - K^i)f|f \rangle$ for $i = 1, 2, \dots$. The hypothesis implies

$$\tilde{\mathcal{E}}(f|f) \leq \frac{1}{\epsilon} \mathcal{E}_{2j}(f|f).$$

Now

$$\begin{aligned} \mathcal{E}_{2j}(f|f) &= \|f\|_2^2 - \|K^j f\|_2^2 = \sum_{i=0}^{j-1} \{\|K^i f\|_2^2 - \|K^{i+1} f\|_2^2\} \\ &= \sum_{i=0}^{j-1} \|K^i (I - K^2)^{1/2} f\|_2^2 \leq j \| (I - K^2)^{1/2} f \|_2^2 = j \mathcal{E}_2(f|f). \end{aligned}$$

Thus

$$\tilde{\mathcal{E}}(f|f) \leq \frac{j}{\epsilon} \mathcal{E}_2(f|f).$$

The second largest eigenvalue of K^2 is β_*^2 . Now, the minimax characterization of eigenvalues gives $1 - \tilde{\beta}_1 \leq \frac{j}{\epsilon} (1 - \beta_*^2)$ or

$$\beta_* \leq \sqrt{1 - \frac{\epsilon}{j} (1 - \tilde{\beta}_1)} \leq 1 - \frac{\epsilon(1 - \tilde{\beta}_1)}{2j}. \quad \square$$

Remark. Lemma 2.1 can be generalized to the case where \tilde{K} and K are reversible with respect to probabilities $\tilde{\pi}$ and π . \square

Combining bounds gives the following theorem, which proves the upper bound of Theorem 1.2 as a special case.

2.5. Theorem

Let Γ be a finitely generated group. Let S be a symmetric set of generators that includes the identity. Assume that (Γ, S) has polynomial growth (2.1). There exist $c_1, c_2 > 0$ such that for any subgroup H of Γ with $X = \Gamma/H$ finite, the Markov chain $k(x, y)$ of (1.1) satisfies

$$2\|k^n(x, \cdot) - u\|_{T.V.} \leq |X|^{1/2}\|k^n(x, \cdot) - u\|_2 \leq c_1 e^{-c_2 \theta}$$

for $n \geq \gamma_X^2(\frac{1}{2} + \theta)$, $\theta > 0$, and any $x \in X$. Here $c_1 = \sqrt{c}$ and $c_2 = 1/(2\lceil a/2 \rceil c)$ with a and c as in Theorem 2.1. Thus c_1, c_2 depend on the volume growth constants C, D at (2.1) but not otherwise on Γ, S , or H .

Proof. From the Cauchy-Schwartz inequality

$$4\|k^n(x, \cdot) - u(\cdot)\|_{T.V.}^2 = \left\{ \sum_{y \in X} |k^n(x, y) - u(y)| \right\}^2 \leq |X| \|k_x^n - u\|_2^2.$$

For any decomposition $n_1 + n_2 = n$,

$$\|k_x^n - u\|_2^2 = \|(K^{n_1} - U)k_x^{n_2}\|_2^2 \leq \|K^{n_1} - U\|_{2 \rightarrow 2}^2 \|k_x^{n_2}\|_2^2$$

where Uf is the mean of f over X . Then

$$\|K^{n_1} - U\|_{2 \rightarrow 2}^2 = \beta_*^{2n_1} \leq \left(1 - \frac{B}{\gamma_X^2}\right)^{2n_1}$$

from Theorem 2.3. Finally, (2.2) with $2n_2 \geq \gamma_X^2$ gives

$$k^{2n_2}(x, y) \leq \frac{c}{|X|}.$$

From this $\|k_x^{n_2}\|_2^2 = k^{2n_2}(x, x) \leq c/|X|$.

Combining bounds, for $n = n_1 + n_2$ with $n_2 \geq \gamma_X^2$,

$$|X| \|k_x^n - u\|_2^2 \leq |X| \frac{c}{|X|} \left(1 - \frac{B}{\gamma_X^2}\right)^{2n_1} \leq c e^{-2Bn_1/\gamma_X^2}.$$

Taking $n_1 = \theta \gamma_X^2$ and $n_2 = \frac{1}{2} \gamma_X^2$ gives the stated result. \square

Remark. The argument for Theorem 2.4 gives a bound on the L^2 norm that is useful for comparison with other chains on the same state space (see [DS-C1, 2]). There is a second argument that avoids eigenvalues: From (2.4), there is a positive integer a such that

$$K_N(x, y) \geq \frac{\epsilon}{|X|} \quad \text{for } N \geq a\gamma_X^2.$$

Here $\epsilon = 1/c$, with a and c from Theorem 2.1. The result now follows from the subadditivity of variation distance

$$\|k^n(x, \cdot) - u(\cdot)\|_{T.V.} \leq (1 - \epsilon)^{n/(a\gamma_X^2)}, \quad n \geq a\gamma_X^2.$$

This proves the theorem with $c_1 = 1$, $c_2 = 1/(ac)$ for $n \geq a\gamma_X^2$. A similar argument works to bound separation distance

$$s(n) = \max_y \left\{ 1 - \frac{K_n(x, y)}{u(y)} \right\},$$

for this is subadditive as well (cf. [AD]).

Proof of Theorem 1.2. To conclude this section, we explain how the upper bound in Theorem 1.1 follows from Theorem 2.4. Consider the free nilpotent group $N_{s,\ell}$ of class ℓ with s generators. This is obtained as a quotient $N_{s,\ell} = \Gamma(s)/\Gamma(s, \ell + 1)$ where $\Gamma(s)$ is the free group on s generators and $\Gamma(s, j)$ is the commutator, defined inductively as $\Gamma(s, 1) = \Gamma(s)$, $\Gamma(s, j) = [\Gamma(s, j-1), \Gamma(s)]$. Any nilpotent group G of class ℓ generated by $s = |S|$ generators is a homomorphic image of $N_{s,\ell}$. Nilpotent groups have polynomial growth by Bass's theorem. Thus the upper bound given in Theorem 1.1 follows from Theorem 2.4. \square

3. Lower Bounds

This section proves the lower bound of Theorem 1.2 in a slightly more general setting. Throughout, Γ is a finitely generated discrete group, $S \subset \Gamma$ a fixed symmetric set of generators with $id \in S$, and $H \subset \Gamma$ a subgroup of finite index. We let $X = \Gamma/H$ be the corresponding homogeneous space and consider the associated graph (X, E) . The main result is Theorem 3.1, stated at the end of this section. This gives a lower bound on the variation distance by using the second eigenvalue and a test function. The argument yields a lower bound on the second eigenvalue.

The first lemma, a variation on a theorem of Guivarc'h [G], shows that quotients of groups with polynomial growth satisfy the doubling property.

3.1. Lemma

Assume that (Γ, S) has polynomial growth (2.1). For C, D as in (2.1), the volume growth function on the graph (X, E) satisfies

$$V(x, 2r) \leq C^2 3^D V(x, r)$$

for all $x \in X$ and $r > 0$.

Proof. Guivarc'h [G, Lemma 1.1] proved that if A, B are finite subsets of Γ and Y is a finite subset of X , then $|A||BY| \leq |BA||A^{-1}Y|$. Take $Y = \{x\}$, $A = B(r)$, $B = B(2r)$ balls in Γ of radius r and $2r$. The bound becomes

$$V_\Gamma(r)V(x, 2r) \leq V_\Gamma(3r)V(x, r) \leq C^2 3^D V(x, r).$$

The last inequality follows from (2.1). \square

Remark. The constant $C^2 3^D$ does not depend on the subgroup H . The argument works for subgroups of infinite index. \square

The next lemma gives a lower bound for the second eigenvalue that matches the upper bound of Theorem 2.3 "up to constants."

3.2. Lemma

Assume that (Γ, S) has polynomial growth (2.1). The second largest eigenvalue β_1 of the chain on X defined at (1.1) satisfies

$$\beta_1 \geq 1 - 32C^4 3^{2D} / \gamma_X^2$$

where γ_X is the diameter of the graph (X, E) .

Proof. From the minimax characterization of eigenvalues

$$\beta_1 \geq 1 - \mathcal{E}(f|f) / \|f\|_2^2$$

for any f with $\sum f(x) = 0$ and \mathcal{E} the Dirichlet form (cf., Lemma 2.1). We construct a good test function f by choosing x_1, x_2 at a distance γ_X apart. Define

$$f(x) = (\lfloor \gamma_X / 2 \rfloor - \rho(x_1, x))_+ - A(\lfloor \gamma_X / 2 \rfloor - \rho(x_2, x))_+$$

with ρ the distance function on (X, E) and A chosen so that $\sum f(x) = 0$. Now,

$$\mathcal{E}(f|f) = \frac{1}{2} \sum_{x,y} |f(x) - f(y)|^2 k(x, y) \leq (1 + A)^2 |X|.$$

Using Lemma 3.1, we get $V(x_i, \gamma_X/4) \geq |X|/C_1$ with $C_1 = C^4 3^{2D}$. This gives

$$\|f\|_2^2 \geq \left(\frac{\gamma_X}{4}\right)^2 \frac{|X|}{C_1} + A^2 \left(\frac{\gamma_X}{4}\right)^2 \frac{|X|}{C_1} = (1 + A^2) \frac{\gamma_X^2 |X|}{16 C_1}.$$

Now $(1 + A)^2/(1 + A^2) \leq 2$, so the result follows. \square

The final result proves the lower bound in Theorem 1.1.

3.3. Theorem

Assume that (Γ, S) has polynomial growth (2.1). Let H be a subgroup of finite index. Suppose the diameter γ_X of $X = \Gamma/H$ satisfies $\gamma_X \geq \beta = 64C^4 3^{2D}$ with C, D from (2.1). Then, for the Markov chain defined at (1.1), there is $x_0 \in X$ such that

$$2\|k_{x_0}^n - u\|_{TV} \geq e^{-n\beta/\gamma_X^2} \quad \text{for all } n \geq 1.$$

Proof. Let β_1 be the second largest eigenvalue of the matrix $k(x, y)$. Choose f to be an eigenfunction for β_1 . Then $|f|$ takes its maximum value at some $x_0 \in X$ and we may assume f is normalized so that $1 = f(x_0) = \sup_X |f|$. Now it follows from Lemma 3.2 that

$$\begin{aligned} 2\|k_{x_0}^n - u\|_{TV} &= \sup_{\|g\|_\infty \leq 1} |K_{x_0}^n(g) - u(g)| \\ &\geq K^n f(x_0) = \beta_1^n \geq \left(1 - \frac{\beta}{\gamma_X^2}\right)^n \geq e^{-\beta n/\gamma_X^2}. \end{aligned}$$

This is the desired result. \square

4. Quotients Under the Condition $V(n) \leq An^B$

The results described in Theorems 2.4 and 3.1 apply uniformly to any finite quotient X of a finitely generated group Γ equipped with a symmetric generating set S satisfying (2.1) (i.e., $C^{-1}n^D \leq V(n) \leq Cn^D \forall n \geq 1$). These theorems show that the random walk on X defined at (1.1) achieves randomness after order γ_X^2 steps with constants depending only on C, D at (2.1).

This section outlines the proof of the similar result under the condition of polynomial growth

$$V(n) \leq An^B \quad \text{for all } n \geq 1. \quad (4.0)$$

4.1. Theorem

Let Γ be a finitely generated group. Let S be a symmetric generating set with $id \in S$. Assume that (Γ, S) satisfies (4.0). Then there exist a_1, a_2, a_3, a_4 depending only on A, B at (4.0) such that, for any finite quotient X of Γ with diameter γ_X , the Markov chain defined at (1.1) satisfies

$$a_1 e^{-a_2 n / \gamma_X^2} \leq \sup_{x \in X} \|k_x^n - u\|_{TV} \leq a_3 e^{-a_4 n / \gamma_X^2}.$$

Outline of the proof. The basic ideas are the same as in §§2 and 3. The problem we face is that condition (4.0) *does not* imply (2.1) with constants C, D depending only on A, B . For instance, finite groups cannot satisfy (2.1) with $D \neq 0$. Here it is worth emphasizing that the constants appearing in Theorems 2.4 and 3.1 can be *effectively* bounded in terms of C, D whereas, at present time, the constants a_i in Theorem 4.1 cannot be effectively bounded in terms of A, B .

Nevertheless, Theorem 6.8 in [DS-C2] shows that there is a constant $C_0 = C_0(A, B)$ such that if (Γ, S) satisfies (2.0) it also satisfies the doubling condition

$$V(2n) \leq C_0(A, B)V(n) \quad \text{for all } n \geq 1. \quad (4.1)$$

This result is based on Gromov's theorem [Gr]. It is easy to see that all we used for the lower bound in §3 is the doubling property, and thus the same argument applies here.

More work is needed for the upper bound. Using the method of Hebisch and Saloff-Coste [HS-C] and §5 of Diaconis and Saloff-Coste [DS-C2], one can prove the following bounds.

4.2. Theorem

Let Γ be a finitely generated group (finite or countable), and let S be a symmetric generating set with $id \in S$. Assume that (Γ, S) satisfies (4.0). Then there exist c_1, c_2, c_3, c_4 depending only on A, B at (4.0) and such that the probability q defined in (1.0) satisfies

$$q^{(n)}(x) \leq c_1 V(\sqrt{n})^{-1} \exp(-c_2 |x|^2 / n) \quad \text{for all } x \in \Gamma, n = 1, \dots$$

and

$$q^{(n)}(x) \geq c_3 V(\sqrt{n})^{-1} \exp(-c_4 |x|^2 / n) \quad \text{for all } x, n \text{ with } |x| \leq n.$$

Here $|x|$ is the word length of $x \in \Gamma$ with respect to the generating set S .

Using this result, the doubling property 4.1, and the technique of Hebisch and Saloff-Coste [HS-C §6], it is not hard to prove the Harnack inequality of Theorem 2.1 for groups satisfying (4.0) and with constants a, c depending only on A, B at (4.0). Once this is done, the upper bound in Theorem 4.1 follows from the arguments of §2. \square

5. Quotients Under Kazhdan's Property T

Let (X, E) be an r -regular graph, and let $k(x, y)$ be the Markov kernel of the simple random walk on (X, E) . Let γ_X be the diameter of (X, E) . This section presents a number of elementary facts that show if the second largest eigenvalue in absolute value, β_* , of the matrix $k(x, y)$ satisfies $\beta_* \leq 1 - \varepsilon$ for some fixed $\varepsilon > 0$, then order γ_X steps are necessary and sufficient to reach equilibrium. This contrasts with the results obtained for quotients of groups having polynomial growth where order γ_X^2 steps are necessary and sufficient. Examples of large graphs satisfying $\beta_* \leq 1 - \varepsilon$ are given by the finite quotients of any countable group having Kazhdan's property. These are discussed at the end of this section.

The following result gives an upper bound for variation distance in terms of the size of X and a lower bound in terms of the diameter γ_X . It also says that $\log |X|$ and γ_X are comparable when $\beta_* \leq 1 - \varepsilon$.

5.1. Theorem

Let (X, E) be an r -regular graph. Suppose the second eigenvalue in absolute value satisfies $\beta_ \leq 1 - \varepsilon$. Then*

$$2\|k_x^n - u\|_{TV} \leq e^{-\theta} \quad \text{for } n \geq \frac{1}{2\varepsilon}(\log |X| + 2\theta), \quad (5.1)$$

$$\sup_{x \in X} \|k_x^n - u\|_{TV} \geq 1/2 \quad \text{for } n \leq \gamma_X/4, \quad (5.2)$$

$$\frac{\log |X|}{\log r} \leq \gamma_X \leq \frac{2}{\varepsilon} \log |X|. \quad (5.3)$$

Proof. Recall the well-known universal bound (see, e.g., Sinclair and Jerrum [SJ])

$$2\|k_x^n - u\|_{TV} \leq \sqrt{|X|} \beta_*^n. \quad (5.4)$$

Using $\beta_* \leq 1 - \varepsilon$ yields (5.1). For (5.2) and (5.3), observe that the number of vertices in any ball of radius n in an r -regular graph is at most r^n . Hence,

$$\gamma_X \geq \log |X| / \log r.$$

Denote by $B(x, t)$ the ball of radius t and center $x \in X$. Clearly, there exists $x_0 \in X$ such that $|X \setminus B(x_0, \gamma_X/4)| \geq |X|/2$. This implies (5.2). Using this, the final inequality (5.3) follows by taking $\theta = 0$ in (5.1). \square

Remarks. 1. Combining (5.2) and (5.3) gives

$$\sup_{x \in X} \|k_x^n - u\|_{TV} \geq 1/2 \quad \text{for } n \leq \frac{\log |X|}{4 \log r}.$$

Similarly, the upper and lower bounds on variation distance can be given in terms of γ_X .

2. The upper bound in (5.3) gives bounds for the diameter in terms of the eigenvalue and size of a regular graph. The argument is a variant of arguments of Alon and Milman [AM], Chung [C], and Mohar [Moh]. These authors give many variants that give sharper bounds in special circumstances. We emphasize that it can be difficult to bound the diameter directly (cf., examples below). There are other circumstances (e.g., approximating the permanent) where one has reasonable bounds on γ_X but where it is hard to bound $|X|$. \square

We next give two examples where Theorem 5.1 is useful. The statements are elementary, but the proofs involve quite sophisticated ideas of representation theory and number theory that we will not develop in detail. Lubotzky [L] and de la Harpe and Valette [HV] gives accessible treatments.

Example 5.1. If Γ is a countable group having Kazhdan's property T (e.g., $SL_d(\mathbb{Z})$ with $d \geq 3$) and S is a finite set of symmetric generators containing the identity, then there is $\varepsilon > 0$ depending only on Γ and S such that any finite quotient of Γ satisfies $\beta_* \leq 1 - \varepsilon$. For example, the group $SL_3(\mathbb{Z}_p)$, p a prime, with generating set

$$S = \left\{ id, \begin{pmatrix} 1 & \pm 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \pm 1 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \right\},$$

satisfies $\beta_* \leq 1 - \varepsilon$ for an explicit ε , uniformly in p . Further details can be found in Margulis [M], Alon and Milman [AM], Lubotzky [L], and de la Harpe and Valette [HV]. \square

Example 5.2. Finite quotients of $SL_2(\mathbb{Z})$ sometimes satisfy $\beta_* \leq 1 - \varepsilon$. For example, $SL_2(\mathbb{Z}_p)$ with generating set

$$S = \left\{ id, \begin{pmatrix} 1 & \pm 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & \pm 1 \\ \mp 1 & 0 \end{pmatrix} \right\}$$

satisfies $\beta_* \leq 1 - \varepsilon$ uniformly in p . An "elementary" proof of this can be found in Brooks [Br]. Lubotzky [L] has more sophisticated proofs. \square

6. Examples

We conclude by treating two classes of examples. The first class gives homogeneous spaces with small diameter compared to the ambient groups. The second class gives examples where adding a single extra move to a walk of polynomial growth gives rapid mixing.

Example 6.1. Let U_d be the group of upper triangular $d \times d$ matrices with 1's on the diagonal and entries in \mathbb{Z} . Let $U_d(m)$ be a similar group with entries mod m . The group $U_d(m)$ is the quotient of U_d by the normal subgroup $H_d(m)$ of the matrices

$$\begin{pmatrix} 1 & mx & mz \\ 0 & 1 & my \\ 0 & 0 & 1 \end{pmatrix} \quad \text{with } x, y, z \in \mathbb{Z}.$$

Let $J(d, m)$ be the subgroup with entries k above the diagonal a multiple of m^k , $0 \leq k \leq d - 1$. Thus, $J(3, m)$ consists of matrices of the form

$$\begin{pmatrix} 1 & mx & m^2z \\ 0 & 1 & my \\ 0 & 0 & 1 \end{pmatrix} \quad \text{with } x, y, z \in \mathbb{Z}.$$

The subgroup $J(d, m)$ is not normal in U_d . Coset representatives may be thought of as $d \times d$ matrices with entries k above the diagonal taken mod m^k , $0 < k < d$. Note that the quotient $X = U_d/J(d, m)$ can also be realized as the quotient of the finite group $U_d(m^{d-1})$ by the subgroup $\underline{J}(d, m)$ similar to $J(d, m)$ but with entries mod (m^{d-1}) .

With generators as in (1.2), one sees easily that for d fixed and m large,

$$X = U_d/J(d, m) = U_d(m^{d-1})/\underline{J}(d, m)$$

has diameter of order m while $U_d(m^{d-1})$ has diameter of order m^{d-1} . Thus, Theorem 1.2 shows that a random walk on the quotient is close to stationarity after order m^2 steps while random walk on the group requires order $m^{2(d-1)}$ steps.

This example can be generalized along the following lines. Let $d = 4$, for instance, and fix integers a_1, a_2, a_3 . Let $K(a_1, a_2, a_3)$ be the subgroup of U_4 containing the matrices of the form

$$\begin{pmatrix} 1 & a_1x & a_1a_2u & a_1a_2a_3w \\ 0 & 1 & a_2y & a_2a_3v \\ 0 & 0 & 1 & a_3z \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{with } x, y, z, u, v, w \in \mathbb{Z}. \quad (6.0)$$

Consider the quotient $U_4/K(a_1, a_2, a_3)$. Choosing different rates of growth for the a_i 's leads to a rich collection of behaviors. Further variations are possible. We can pick b_1, b_2 such that b_i divides $a_i a_{i+1}$ for $i = 1, 2$ and choose c_1 , which divides $a_1 b_2, a_2 b_1$ and $b_1 b_2$. Then we can define $K(a, b, c)$ where $a = (a_i), b = (b_i), c = (c_i)$ by replacing the products $a_1 a_2, a_2 a_3$, and $a_1 a_2 a_3$ in (6.0) by b_1, b_2 , and c_1 , respectively. For example, if $d = 3$ the subgroup

$$K = K((m^2, 1), m) = \left\{ \begin{pmatrix} 1 & m^2x & mz \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} : x, y, z \in \mathbb{Z} \right\}$$

is normal in U_3 . When $m = p$ is a prime, $U_3/H_3(p) = U_3(p)$ and $U_3/K = M_3(p)$ are the only two nonabelian groups of order p^3 . For a direct analysis of random walk on $U_3(m)$ and $M_3(m)$, see Diaconis and Saloff-Coste [DS-C2]. \square

Example 6.2. Consider the opening example with $d = 2$. A Markov chain on \mathbb{Z}_m^2 was defined with the following transitions:

$$(x, y) \rightarrow \begin{cases} (x, y) \\ (x + y, y) & (x - y, y) \\ (x, y + 1) & (x, y - 1) \end{cases} \text{ each with probability } 1/5. \quad (6.1)$$

Theorem 1.1 states that for m large, this chain takes order m^2 steps to achieve randomness. Suppose now that the moves in (6.1) are supplemented by a single extra move

$$(x, y) \rightarrow (y, x).$$

Then it can be shown that order $\log m$ steps suffice to achieve randomness as follows.

6.1. Theorem

Consider the Markov chain with state space \mathbb{Z}_m^2 and transitions

$$(x, y) \rightarrow \begin{cases} (x, y) & (y, x) \\ (x + y, y) & (x - y, y) \\ (x, y + 1) & (x, y - 1) \end{cases} \text{ each with probability } 1/6. \quad (6.2)$$

Then there is a constant C such that for any starting state x in \mathbb{Z}_m^2 ,

$$\|k_x^n - u\|_{TV} \leq e^{-\theta} \text{ for } n \geq C(\theta + \log m).$$

Proof. Let $\overline{SL}_2(m)$ be the group of the 2×2 matrices with determinant $\pm 1 \pmod{m}$. Consider the group $G = \overline{SL}_2(m) \times \mathbb{Z}_m^2$ with generating set

$$S = \left\{ \left(\begin{pmatrix} 1 & \pm 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right), \left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right), \right. \\ \left. \left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \pm 1 \end{pmatrix} \right), \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right) \right\}.$$

The original walk can be realized as the evolution of the \mathbb{Z}_m^2 coordinates in this extended walk. Thus, it is the quotient walk on $G/\overline{SL}_2(m)$. The argument presented in Lubotzky [L] shows that the second largest eigenvalue of the walk on the group G is bounded by a constant $0 < 1 - \varepsilon < 1$ uniformly in

m . This is because the pair (G, \mathbb{Z}_m^2) has the restricted property T (see Lubotzky [L] and de la Harpe and Valette [HV] for details). It follows easily that the second largest eigenvalue (in absolute value) β_* of the original walk satisfies

$$\beta_* \leq 1 - \varepsilon .$$

The bound in Theorem 6.1 then follows from (5.4). \square

Remarks. 1. The graph corresponding to (6.2) is essentially the first explicit example of expander constructed by Margulis [M].

2. A similar argument shows that for any fixed $d \geq 2$, if the basic moves of Example 1.1 are supplemented by a single cyclic shift and its inverse, then order $\log m$ steps are enough to reach equilibrium. See, for example, Lubotzky [L] or de la Harpe and Valette [HV]. The arguments for these results are *not* currently elementary. \square

References

-
- [AD] Aldous, D., and Diaconis, P., (1986). Shuffling cards and stopping times. *Amer. Math. Monthly* **93**, 333–348.
- [AM] Alon, N., and Milman, V., (1985) λ_1 , isoperimetric inequalities for graphs and superconcentrators. *J. Combin. Theory, Ser. B* **38**, 78–88.
- [B] Bass, H., (1972). The degree of polynomial growth of finitely generated nilpotent groups. *Proc. London Math Soc.* **25**, 603–614.
- [Br] Brooks, R., (1989). Some relations between spectral geometry and number theory. Preprint. Department of Mathematics, University of Southern California.
- [C] Chung, F., (1989). Diameters and eigenvalues. *J. Amer. Math. Soc.* **2**, 187–196.
- [DS-C1] Diaconis, P., and Saloff-Coste, L., (1993). Comparison theorems for random walk on groups. *Ann. Probab.* **21**, 2131–2156.
- [DS-C2] ———, (1994). *Moderate growth and random walk on finite groups.* *Geom. Funct. Anal.* **4**, 1–36.
- [G] Guivarch, Y., (1973). Croissance Polynomiale et periodes des fonctions harmoniques. *Bull. Soc. Math. France* **101**, 333–379.
- [Gr] Gromov, M., (1981). Groups of polynomial growth and expanding maps. *Inst. Hautes Études Sci. Publ. Math.* **53**, 53–73.
- [H] Hall, P., (1957). Nilpotent groups. *Collected Works of Philip Hall.* Oxford University Press, Oxford, 417–462.
- [HV] de la Harpe, P., and Valette, A., (1989). *La propriété (T) de Kazhdan pour les groupes localement compacts.* *Astérisque*, vol. 175, Soc. Math. France, Paris.
- [HS-C] Hebisch, W., and Saloff-Coste, L., (1993). Gaussian estimates for Markov chains and random walks on groups. *Ann. Probab.* **21**, 673–709.
- [L] Lubotzky, A., (1994). *Discrete groups, expanding graphs and invariant measures.* Birkhauser, Basel.

- [M] Margulis G., (1975). Explicit construction of cocentrators. *Problems of Inform. Transmission* **10**, 325–332.
- [Moh] Mohar, B., (1991). Eigenvalues, diameter, and mean distance in graphs. *Graphs Combin.* **7**, 53–64.
- [Mos] Moser, J., (1964). A Harnack inequality for parabolic differential equations. *Comm. Pure Appl. Math.* **17**, 101–134.
- [SJ] Sinclair, A., and Jerrum, M., (1989). Approximate counting, uniform generation and rapidly mixing Markov chains. *Inform. and Comput.* **82**, 93–133.
- [Su1] Suzuki, M., (1982). *Group theory I*. Springer-Verlag, New York.
- [Su2] ———, (1986). *Group theory II*. Springer-Verlag, New York.
- [T] Tits, J., (1981). Appendix to Gromov's paper, *Inst. Hautes Études Sci. Publ. Math.* **53**, 74–78.
- [V] Varopoulos, N., (1986). Analysis on nilpotent groups. *J. Funct. Anal.* **66**, 406–431.
- [VS-CC] Varopoulos, N., Saloff-Coste, L., and Coulhon Th., (1993). *Analysis and geometry on groups*. Cambridge University Press, Cambridge.

Received September, 1993

Department of Mathematics, Harvard University, Cambridge, Massachusetts 02138

CNRS, Université Paul Sabatier, Lab Statistique Probabilitié, 31062 Toulouse, Cedex, France