Geometric and Functional Analysis
Vol. 4, No. 1 (1994)
1016-443X/94/01001-36$1.50+0.20/0

MODERATE GROWTH AND RANDOM WALK ON FINITE GROUPS

P. DIACONIS AND L. SALOFF-COSTE

Abstract

We study the rate of convergence of symmetric random walks on finite groups to the uniform distribution. A notion of moderate growth is introduced that combines with eigenvalue techniques to give sharp results. Roughly, for finite groups of moderate growth, a random walk supported on a set of generators such that the diameter of the group is γ requires order γ^2 steps to get close to the uniform distribution. This result holds for nilpotent groups with constants depending only on the number of generators and the class. Using Gromov's theorem we show that groups with polynomial growth have moderate growth.

1. Introduction

We begin with an example of the problem under study. Let m be a positive integer. Let $U_3(m)$ be the Heisenberg group mod m. This is the set of 3×3 matrices of form

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \quad x, y, z \in \mathbb{Z}_m \ .$$

Thus $|U_3(m)| = m^3$. A random walk can be performed on $U_3(m)$ by repeatedly choosing one of the following 5 matrices with probability $1/5$:

$$I_3, \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.1}$$

The n^{th} stage of the walk is the product of the first n chosen matrices. Let $q^{(n)}(s)$ be the chance that the random walk is at s after n steps. For n suitably large, $q^{(n)}(s)$ is approximately equal to the uniform measure $u(s) = 1/m^3$. The following result says that "suitably large" is $n \gg m^2$.

THEOREM 1.1. *Let q assign mass $1/5$ to the matrices in (1.1). Then, there are universal positive constants a, b, a', b' such that*

$$a'e^{-b'n/m^2} \leq \|q^{(n)} - u\|_{\text{T.V.}} \leq ae^{-bn/m^2} .$$

The total variation distance is defined as

$$\|q^{(n)} - u\|_{\text{T.V.}} = \tfrac{1}{2} \sum_s |q^{(n)}(s) - u(s)| . \tag{1.2}$$

Thus, the distance to the uniform distribution is exponentially small provided $n \gg m^2$ and bounded away from 0 if n is small compared to m^2.

The object of this paper is to prove theorems like Theorem 1.1 for general symmetric measures on certain finite groups which we call groups of *moderate growth*.

Let G be a finite group. Let E be a set of generators for G. Throughout, we assume that E is symmetric and $id \in E$. The Cayley graph (G, E) is the graph with vertex set G and edge set $\{(x, xe) : x \in G , e \in E\}$. The *volume growth* function $V(n)$ is defined by

$$V(n) = |E^n| . \tag{1.3}$$

The *diameter* γ of G with respect to E is defined by

$$\gamma = \min\{n : V(n) = |G|\} . \tag{1.4}$$

The group G has (A, d)-*moderate growth* with respect to E if there are positive constants A and d such that

$$\frac{V(n)}{V(\gamma)} \geq \frac{1}{A} \left(\frac{n}{\gamma}\right)^d \quad 1 \leq n \leq \gamma . \tag{1.5}$$

To state a general result, define a probability q on G by

$$q(s) = \begin{cases} 1/|E| & \text{if } s \in E \\ 0 & \text{elsewhere} \end{cases} \tag{1.6}$$

and let $q^{(n)}$ denote the n^{th} convolution power of q.

THEOREM 1.2. *Let G, E have (A, d)-moderate growth. Let q be defined by (1.6). For $n = (1 + c)|E|\gamma^2$, $c > 0$,*

$$\|q^{(n)} - u\|_{\text{T.V.}} \leq Be^{-c}$$

with $B = A^{1/2}2^{d(d+3)/4}$. *Further, for* $n = c\gamma^2/(2^{4d+2}A^2)$,

$$\|q^{(n)} - u\|_{\text{T.V.}} \geq \tfrac{1}{2}e^{-c}\,.$$

Theorem 1.2 is of interest for a family of groups of moderate growth with $|E|, A, d$ fixed as $|G|$ (and so γ) gets large. The results then show that, if the number of steps n is a large multiple of γ^2, the walk is close to uniform whereas, for n a small multiple of γ^2, the walk is far from uniform. The transition from 1 to 0 as c varies is typically smooth so that the "cutoff phenomena" observed by Aldous and Diaconis [AD] can be proved not to occur for the groups under study.

Remarks: 1. The group $\mathbb{Z}_m$ of the integers mod m (take m odd for definiteness) with $E = \{0, \pm 1\}$ has diameter $\gamma = (m-1)/2$. This gives moderate growth with $A = 1$ and $d = 1$.

2. The rate of growth can depend on the generators. The group $\mathbb{Z}_m$ with $E = \{0, \pm 1, \pm\lfloor\sqrt{m}\rfloor\}$ has γ of order $\sqrt{m}$ and moderate growth with $A = 1$ and $d = 2$.

3. The Heisenberg group $U_3(m)$ with E given by (1.1) has $\gamma = m + 1$ and moderate growth with $d = 3$ and fixed A for all m. See section 3.B.

4. Of course, any finite group has moderate growth for suitable A. The point is that there are many natural families of groups which have moderate growth for fixed A and d as $|G|$ gets large. These include nilpotent groups (and so p groups) and affine groups.

5. It is often convenient to check (1.5) by determining γ, then showing $|G| \leq \beta\gamma^d$ and $V(n) \geq \alpha n^d$, $1 \leq n \leq \gamma$. This gives (1.5) with $A = \beta/\alpha$ since $V(\gamma) = |G|$.

6. It is instructive to see the difference between the two conditions of Remark 5 and the single condition (1.5). Consider the product $\mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2}$ with $m_1 \leq m_2$ and both m_1, m_2 odd. Take E={(0,0),(1,0),(-1,0),(0,1),(0,-1)} as the generating set. The diameter is $\gamma = \frac{m_1-1}{2} + \frac{m_2-1}{2} \leq m_2$. The growth function satisfies

$$V(n) \geq n^2 \quad \text{for} \quad 1 \leq n \leq \frac{m_1 - 1}{2}$$

$$V(n) \geq m_1 n \quad \text{for} \quad \frac{m_1 - 1}{2} \leq n \leq \gamma\,.$$

Since $|G| = m_1 m_2 = V(\gamma)$, the group G, E has moderate growth with $A = \frac{1}{2}$ and $d = 2$. Theorem 1.2 shows that order m_2^2 steps are necessary and suffice to achieve randomness. However, the best that can be said using the approach of Remark 5 is $V(n) \geq n$, $1 \leq n \leq \gamma$; $|G| \leq 2m_1\gamma$. Dividing leads to

$$\frac{V(n)}{V(\gamma)} \geq \frac{1}{2m_1}\left(\frac{n}{\gamma}\right)\,.$$

Using this, Theorem 1.2 yields only that order $m_2^2 \log m_1$ steps suffice. The lower bound shows that after $(m_2/m_1)^2$ steps the variation distance is bounded below by a constant over $m_1^{1/2}$. This is virtually useless.

The structure of this paper is as follows. Section 2 develops the basic analytical tools of eigenvalues and volume growth. Some results are given for groups of exponential growth. The basic new ingredient leans heavily on an idea of Hebisch [He]. Theorems 1.1 and 1.2 are proved in section 3.

Section 4 gives examples which include the Heisenberg group and the upper triangular group.

Section 5 introduces the doubling property: $V(2n) \leq AV(n)$ for all n. This implies moderate growth and is shown to hold for all nilpotent groups with A depending only on the number of generators and the class of nilpotency.

Section 6 discusses polynomial growth: $V(n) \leq An^d$ for all n. Gromov's theorem is used to show that polynomial growth is equivalent to the doubling property and so implies moderate growth.

Section 7 treats normal extensions such as the affine group mod$p$. Some of these extensions have moderate growth but fail to have polynomial growth.

In a companion paper we treat random walk on homogeneous spaces of nilpotent groups by a different set of techniques: the walk is lifted to the free nilpotent group. Then, Harnack inequalities of Hebisch and Saloff-Coste [HeS-C] can be applied. See Diaconis and Saloff-Coste [DS-C3].

2. Volume Growth and Decay of Convolution Powers

A. Basics. Let G be a finite group with identity id. Given real valued functions φ, ψ on G, their *convolution* is the function $\varphi * \psi$ defined by

$$\varphi * \psi(x) = \sum_y \varphi(xy^{-1})\psi(y) = \sum_y \varphi(y)\psi(y^{-1}x) .$$

We denote by Ψ the operator $\Psi\varphi = \varphi * \psi$ and by $\varphi^{(n)}$ the convolution powers of φ. Let U be the operator associated with the uniform distribution $u(x) = 1/|G|$. For $1 \leq s \leq \infty$, define the usual l^s norms of a function φ as

$$\|\varphi\|_s = \left(\sum_{x \in G} |\varphi(x)|^s \right)^{\frac{1}{s}}, \quad \|\varphi\|_\infty = \max_{x \in G} \{|\varphi(x)|\} .$$

The variation distance $\|q - \tilde{q}\|_{\text{T.V.}} = \max_{I \subset G} \{|q(I) - \tilde{q}(I)|\}$ between two probabilities $q, \tilde{q}$ is half the l^1 norm of $q - \tilde{q}$.

Let q be a symmetric (i.e. $q(x) = q(x^{-1})$) probability on G. Under mild restrictions, the convolution powers $q^{(n)}$ converge to the uniform distribution in variation distance. Analytical approaches to this result proceed by bounding the l^1 norm by the l^2 norm using the Cauchy-Schwarz inequality

$$\|q^{(n)} - u\|_{\text{T.V.}} \leq \tfrac{1}{2}|G|^{1/2}\|q^{(n)} - u\|_2 \,. \tag{2.1}$$

The l^2 norm is bounded by eigenvalues. Because of symmetry, the associated Markov chain has real eigenvalues $1 = \beta_0 > \beta_1 \geq \beta_2 \geq \cdots \geq \beta_{|G|-1} > -1$. A detailed discussion of this approach is given by Diaconis and Saloff-Coste [DS-C2] which should be regarded as a companion to the present paper. The following bound introduces a fresh ingredient.

LEMMA 2.1. *Let $q(s)$ be a symmetric probability on a finite group G. Let*

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

be the second largest eigenvalue, in absolute value, of the associated Markov chain. Then for non-negative integers n, m,

$$\|q^{(n+m)} - u\|_2^2 \leq q^{(2m)}(id)\beta_*^{2n} \,.$$

Proof: Using operator notation, the following equalities are known and easy to verify.

$$q^{(2m)}(id) = \|q^{(m)}\|_2^2 = \|Q^m\|_{2\to\infty}^2 \,; \quad \|Q^n - U\|_{2\to 2} = \beta_*^n \,.$$

From these, for any function f,

$$\begin{aligned} \|f * (q^{(n+m)} - u)\|_\infty &= \|f * (q^{(n)} - u) * q^{(m)}\|_\infty \\ &\leq \|Q^m\|_{2\to\infty}\|f * (q^{(n)} - u)\|_2 \\ &\leq \|Q^m\|_{2\to\infty}\|Q^n - U\|_{2\to 2}\|f\|_2 \,. \end{aligned}$$

Then

$$\|q^{(n+m)} - u\|_2^2 = \|Q^{n+m} - U\|_{2\to\infty}^2 \leq q^{(2m)}(id)\beta_*^{2n} \,. \qquad \square$$

Remark: A variety of bounds are available for β_*. These range from exact computation using Fourier analysis, to geometric inequalities of Poincaré and Cheeger type. Diaconis [D], Diaconis and Stroock [DStr] and Diaconis and Saloff-Coste [DS-C2] give a picture of what is available. The only bound needed in the present paper is the following diameter bound derived in [DS-C2]; many versions of this bound have been published earlier by a number of authors.

LEMMA 2.2. *Let G be a finite group, E a symmetric set of generators for G, γ the diameter of G with respect to E. Let q be a probability on G such that $q(s) = q(s^{-1})$ and $\eta = \inf\{q(s) : s \in E \setminus \{id\}\} > 0$. Then the second largest eigenvalue is bounded by*

$$\beta_1 \leq 1 - \frac{\eta}{\gamma^2} .$$

Further, the smallest eigenvalue satisfies

$$\beta_{|G|-1} \geq -1 + 2q(id) .$$

Remarks: 1. Using Lemma 2.1 with $m = 0$ and the bound (2.1) gives

$$\|q^{(n)} - u\|_{\text{T.V.}} \leq \tfrac{1}{2}|G|^{\frac{1}{2}}\beta_*^n .$$

For groups with moderate growth the $|G|^{\frac{1}{2}}$ factor is extraneous. Lemma 2.1 allows the information on the decrease of $q^{(2m)}(id)$ to be used to kill this term. The next section shows how moderate growth can be used for this purpose.

2. All the bounds obtained in this paper are stated for variation distance. However, the upper bounds are obtained using Lemma 2.1. Thus, our estimates hold as well for the quantity $|G|^{\frac{1}{2}}\|q^{(n)} - 1\|_2$. Moreover, the Cauchy-Schwarz inequality yields

$$|G|\|q^{(2n)} - u\|_\infty \leq |G|\|q^{(n)} - u\|_2^2$$

so that, in fact, our bounds are also valid for the relative maximum error distance

$$\sup_x \left||G|q^{(n)}(x) - 1\right| .$$

B. Decay of convolution powers. In this section a quantitative relation between volume growth and decay of $q^{(2n)}(id)$ is obtained. These results are developments of an idea of W. Hebisch [He]; see also Hebisch and Saloff-Coste [HeS-C].

THEOREM 2.3. *Let E be a set of generators of the finite group G. Let q be a symmetric probability on G and set $\eta = \inf\{q(s) : s \in E \setminus \{id\}\}$. If the volume growth $V(n)$ satisfies*

$$V(n) \geq \alpha n^d , \quad 1 \leq n \leq N ,$$

for some positive α, d, and N, then

$$q^{(2n+1)}(id) \leq q^{(2n)}(id) \leq \frac{C}{n^{d/2}} \quad \text{for all } n \leq N^2/\eta$$

with $C = 2^{2+\frac{3}{2}d+\frac{d^2}{2}}/(\alpha\eta^{d/2})$.

Theorem 2.3 follows from a sequence of lemmas which will be used further on. The first lemma says that the operator Q associated to q may be assumed positive after adjusting constants.

LEMMA 2.4. *With notation as in Theorem 2.1, let $q_+ = (\delta_{id} + q)/2$. Then Q_+ is positive on L^2, $\eta(q_+) = \eta(q)/2$, and*

$$\|q^{(2n+1)}\|_\infty \leq \|q^{(2n)}\|_\infty = \|q^{(n)}\|_2^2 \leq 2\|q_+^{(n)}\|_2^2 = 2\|q_+^{(2n)}\|_\infty .$$

Proof: For a symmetric probability $\|q^{(2n)}\|_\infty = q^{(2n)}(id)$. Thus

$$q_+^{2n}(id) = \frac{1}{2^{2n}} \sum_{j=0}^{2n} \binom{2n}{j} q^{(j)}(id) \geq \frac{1}{2^{2n}} \sum_{j=0}^{n} \binom{2n}{2j} q^{(2j)}(id)$$

$$\geq q^{(2n)}(id) \frac{1}{2^{2n}} \sum \binom{2n}{2j} = \frac{q^{(2n)}(id)}{2} .$$

The other claims are obvious. □

The following slightly mysterious Lemma is from Coulhon and Saloff-Coste [CoS-C]. It makes a crucial appearance in the proof of Lemma 2.6 below.

LEMMA 2.5. *Let q be a symmetric probability measure on a finite group G. Suppose Q is positive on L^2. Then*

$$\|(I - Q)^{\frac{1}{2}} Q^m\|_{2\to 2} \leq \frac{1}{2\sqrt{m}} .$$

Proof: Let $f_0, f_1, \ldots, f_{|G|-1}$ be an orthonormal basis of eigenvectors for Q on L^2. Let β_i be the corresponding eigenvalue. Given $f \in L^2$, write $f = \sum a_i f_i$ so that

$$(I-Q)^{\frac{1}{2}} Q^m f = \sum (1-\beta_i)^{\frac{1}{2}} \beta_i^m a_i f_i \, .$$

Thus

$$\|(I-Q)^{\frac{1}{2}} Q^m f\|_2^2 = \sum (1-\beta_i)\beta_i^{2m} a_i^2 \, .$$

Now, for $m \geq 1$, $\sup_{0 \leq x \leq 1}(1-x)x^{2m} \leq 1/(4m)$ by calculus. □

The next lemma gives a relation between the rate of decrease of convolutions and volume growth. It does not require moderate growth and is used at the end of this section to give a result for exponential growth.

LEMMA 2.6. *Let E be a set of generators of the finite group G. Let q be a symmetric probability on G such that Q is positive on L^2. Let $\eta = \inf\{q(s) : s \in E \setminus \{id\}\}$. Then, for all n, m*

$$q^{(2n+m)}(id) \leq 2/V(r(n,m))$$

where

$$r(n,m) = \left(\frac{\eta}{2}\right)^{\frac{1}{2}} m^{\frac{1}{2}} q^{(2n+m)}(id)/q^{(2n)}(id) \, .$$

Proof: For any $x \in G$ and $z \in E$

$$|q^{(2n+m)}(x) - q^{(2n+m)}(xz)| \leq \sum_y |q^{(n+m)}(y^{-1}x) - q^{(n+m)}(y^{-1}xz)| q^{(n)}(y)$$

$$\leq \left\{ \sum_y |q^{(n+m)}(y^{-1}x) - q^{(n+m)}(y^{-1}xz)|^2 \right\}^{\frac{1}{2}} \left\{ \sum_y q^{(n)}(y)^2 \right\}^{\frac{1}{2}}$$

$$\leq \frac{\|q^{(n)}\|_2}{\eta^{\frac{1}{2}}} \left\{ \sum_{\substack{y \in G \\ w \in E}} |q^{(n+m)}(y) - q^{(n+m)}(yw)|^2 q(w) \right\}^{\frac{1}{2}}$$

$$\leq (2/\eta)^{\frac{1}{2}} \|q^{(n)}\|_2 \langle (I-Q) q^{(m+n)} | q^{(m+n)} \rangle^{1/2}$$

$$= (2/\eta)^{\frac{1}{2}} \|q^{(n)}\|_2 \|(I-Q)^{\frac{1}{2}} Q^m q^{(n)}\|_2 \leq \frac{\|q^{(n)}\|_2^2}{\sqrt{2\eta m}}$$

where Lemma 2.5 was used to justify the last inequality.

Let $|x|$ denote the word length of $x \in G$ with respect to E. Writing x as a sequence of generators and using $\|q^{(n)}\|_2^2 = q^{(2n)}(id)$, we get

$$|q^{(2n+m)}(x) - q^{(2n+m)}(id)| \leq |x| q^{(2n)}(id)/\sqrt{2\eta m} \, .$$

Hence, if $|x| \leq r(n,m)$,

$$|q^{(2n+m)}(x) - q^{(2n+m)}(id)| \leq q^{(2n+m)}(id)/2 \,,$$

and thus $q^{(2m+m)}(x) \geq q^{(2n+m)}(id)/2$. Summing the last inequality over the set $|x| \leq r(n,m)$ gives the result. □

Lemma 2.6 sets up a kind of recurrence between the decay of convolution powers that appear on both sides of the inequality. This is exploited to give a proof of Theorem 2.3.

Proof of Theorem 2.3: Assume first that Q is positive on L^2. Set $A(n) = q^{(n)}(id)$. Using $V(n) \geq \alpha n^d$ for $1 \leq n \leq N$, Lemma 2.6 yields

$$\begin{aligned} A(2n+m) &\leq 2V(r(n,m))^{-1} \leq 2/(\alpha r(n,m)^d) \\ &= \frac{2}{\alpha}\left(\frac{\eta}{2}\right)^{-d/2} m^{-d/2}\left(\frac{A(2n+m)}{A(2n)}\right)^{-d} , \end{aligned} \tag{2.2}$$

this being valid for any n, m with $r(n,m) \leq N$. From the definitions in Lemma 2.6, $r(n,m) \leq \left(\frac{\eta m}{2}\right)^{1/2}$. Thus (2.2) is valid for $m \leq 2N^2/\eta$ and all n.

Rewriting (2.2), we get

$$A(2n+m) \leq \left\{\left(\frac{2}{\alpha}\right)^{\frac{1}{d}}\left(\frac{2}{\eta m}\right)^{\frac{1}{2}} A(2n)\right\}^{d/(1+d)} \tag{2.3}$$

for $m \leq 2N^2/\eta$ and all n. Fix $n_0 \leq 4N^2/\eta$. Let n be such that $2^n \leq n_0 < 2^{n+1}$. Then $\|q^{(n_0)}\|_\infty \leq A(2^n)$. Using (2.3) repeatedly, halving the argument each time,

$$\begin{aligned} A(2^n) = A(2^{n-1}+2^{n-1}) &\leq \left\{\left(\frac{2}{\alpha}\right)^{\frac{1}{d}}\left(\frac{2}{\eta 2^{n-1}}\right)^{\frac{1}{2}} A(2^{n-1})\right\}^{\frac{d}{(1+d)}} \\ &\leq \left\{\left(\frac{2}{\alpha}\right)^{\frac{1}{d}}\left(\frac{2}{\eta 2^{n-1}}\right)^{\frac{1}{2}}\right\}^{\frac{d}{(1+d)}} \left\{\left(\frac{2}{\alpha}\right)^{\frac{1}{d}}\left(\frac{2}{\eta 2^{n-2}}\right)^{\frac{1}{2}} A(2^{n-2})\right\}^{\frac{d}{(1+d)^2}} \\ &\leq \left\{\left(\frac{2}{\alpha}\right)^{\frac{1}{d}}\left(\frac{2}{\eta}\right)^{\frac{1}{2}}\right\}^{\theta+\theta^2+\ldots+\theta^{n-2}} \{2^{-[\frac{(n-1)}{2}\theta+\frac{(n-2)}{2}\theta^2+\ldots+\frac{2}{2}\theta^{n-2}]}\} A(2)^{\theta^{n-2}} \end{aligned}$$

with $\theta = d/(1+d)$. Now

$$\sum_{i=1}^{n-2}\theta^i = d(1-\theta^{n-2}) \,, \quad \sum_{i=1}^{n-2} i\theta^i = (d+d^2)(1-(n-1)\theta^{n-2}+(n-2)\theta^{n-1}) \,.$$

Using these, elementary manipulations show

$$\|q^{(n_0)}\|_\infty \leq \frac{2^{1+(3d+d^2)/2}}{\alpha(\eta n_0)^{d/2}} \quad \text{for} \quad 1 \leq n_0 \leq 4N^2/\eta \,.$$

This result was proved assuming Q is positive. For general Q, Lemma 2.4 along with the present calculations shows

$$\|q^{(2n)}\|_\infty \leq \frac{2^{2+(3d+d^2)/2}}{\alpha(\eta n)^{d/2}} \quad \text{for} \quad n \leq N^2/\eta \,.$$

This completes the proof. □

Theorem 2.3 will be used throughout this paper as a basic tool. It is natural to enquire about groups with exponential growth. The next theorem gives a result on the decrease of polynomial powers. Unfortunately, we have been unable to use it to sharpen rates of convergence.

THEOREM 2.7. *Let q be a symmetric probability measure on a finite group G. Let E be a set of generators such that $\eta = \inf\{q(s), s \in E \setminus \{id\}\} > 0$. Assume that $V(n) \geq e^{c_0 n}$ for $n \leq N$. Then, we have*

$$q^{(n)}(id) \leq 2\, e^{-\frac{3}{8}(\eta c_0^2 n)^{\frac{1}{3}}} \text{ for } n \leq 16 c_0 N^3/\eta \,.$$

Proof: No matter what the rate of volume growth, for all $1 \leq m \leq n \leq \infty$

$$q^{(4n)}(id) \leq \max\{2^{1-\frac{n}{m}} q^{(2n)}(id)\ ,\ 2/V((\eta m/8)^{\frac{1}{2}})\} \,. \tag{2.4}$$

To see this, fix $m \leq n$. Set $A(n) = q^{(n)}(id)$, note that if

$$A(2n + 2im) > A(2n + 2(i-1)m)/2$$

for an integer $i \in [1, n/m]$, then Lemma 2.6 implies

$$A(4n) \leq A(2n + 2im) \leq 2/V((\eta m/8)^{\frac{1}{2}}) \,.$$

But if for all integers $i \in [1, n/m]$ we have $A(2n+2im) \leq A(2n+2(i-1)m)/2$, then $A(4n) \leq 2^{1-n/m} A(2n)$. This proves (2.4). Choosing m of order $cn^{2/3}$, with $c = 2(\log 2)^{\frac{2}{3}}/(c_0^2\eta)^{\frac{1}{3}}$ proves Theorem 2.7. □

3. Convergence under Moderate Growth

In this section we prove a slight extension of Theorem 1.2.

THEOREM 3.1. *Let G be a finite group with generating set E containing the identity. Suppose G has (A,d)-moderate growth with respect to E as in (1.5). Let q be a symmetric probability on G with $\eta = \inf\{q(s), s \in E\} > 0$. Then,*

$$\|q^{(n)} - u\|_{T.V.} \le Be^{-c} \quad \text{for} \quad n = (1+c)\gamma^2/\eta \quad \text{with} \quad c > 0 \ ,$$

where $B = A^{\frac{1}{2}} 2^{d(d+3)/4}$. For a lower bound, assume in addition that q is supported on E and that $\gamma \ge A2^{2d+2}$ (so the diameter is large with respect to the constants involved). Then,

$$\|q^{(n)} - u\|_{TV} \ge \tfrac{1}{2} e^{-c} \quad \text{for} \quad n = c\gamma^2/(2^{4d+2}A^2) \ .$$

This result is useful when there is a sequence of groups with moderate growth for fixed A, d and a fixed (or slowly growing) set of generators. Theorem 3.1 then says that, if q does not vary too much on its support, the walk is close to uniformly distributed after $c\gamma^2$ steps when c is large and far from uniformly distributed if c is small.

Proof of the Upper Bound: Lemma 2.2 shows that the second eigenvalue of the walk satisfies $\beta_1 \le 1 - \eta/\gamma^2$ and that the smallest eigenvalue satisfies

$$\beta_{|G|-1} \ge -1 + 2q(id) \ge -1 + 2\eta \ge -1 + \eta/\gamma^2 \ .$$

Thus, $\beta_* \le 1 - \eta/\gamma^2$ (note that for this, only $q(id) \ge \eta/\gamma^2$ is needed). Then, Lemma 2.1 yields

$$\|q^{(n+m)} - u\|^2_{T.V.} \le \frac{|G|}{4} q^{(2m)}(id)\beta_*^{2n} \quad \text{for any} \quad n, m.$$

By assumption, $V(j) \ge |G| j^d/(A\gamma^d)$, $1 \le j \le \gamma$. Using this in Theorem 2.3 with $m = \gamma^2/\eta$ gives

$$q^{(2m)}(id) \le \frac{C}{(\gamma^2/\eta)^{d/2}} \quad \text{with} \quad C = 2^{2+\frac{3}{2}d+\frac{d^2}{2}} \left(\frac{|G|}{A\gamma^d}\eta^{d/2}\right)^{-1} \ .$$

Thus, with $n = c\gamma^2/\eta$

$$\|q^{(m+n)} - u\|^2_{T.V.} \le A2^{\frac{3}{2}d+\frac{d^2}{2}} e^{-2c} \ .$$

Taking square roots yields the desired result. □

Proof of the Lower Bound: We start with the well known equivalence

$$2\|q^{(k)} - u\|_{T.V.} = \max_{\|f\|_\infty \le 1} |Q^k(f) - U(f)| .$$

Thus, for any specific f with $\|f\|_\infty \le 1$,

$$\|q^{(k)} - u\|_{T.V.} \ge \tfrac{1}{2}|Q^{(k)}(f) - U(f)| .$$

We will choose f to be an eigenfunction for the second eigenvalue β_1 suitably normalized. Namely, let f be such that $Qf = \beta_1 f$ with $\|f\|_\infty = f(id) = 1$ say. Since $U(f) = 0$, we get for any j

$$\|q^{(j)} - u\|_{T.V.} \ge \tfrac{1}{2}|Q^{(j)}(f)(id)| = \tfrac{1}{2}\beta_1^j . \tag{3.1}$$

To conclude the argument, a lower bound for β_1 is needed. We show

$$\beta_1 \ge 1 - \frac{B}{\gamma^2} \quad \text{with } B = 4^{2d+1}A^2 . \tag{3.2}$$

To see this, consider $g(x) = |x|$, the distance with respect to the generating set E. The minimax characterization of eigenvalues gives

$$1 - \beta_1 \le \mathcal{E}(g|g)/\mathrm{Var}(g)$$

where

$$\mathcal{E}(g|g) = \tfrac{1}{2}\sum_{x,y}(g(x) - g(xy))^2 q(y)u(x) ,$$

$$\mathrm{Var}(g) = \tfrac{1}{2}\sum_{x,y}(g(x) - g(y))^2 u(x)u(y) .$$

Since q is supported on $E, (g(x) - g(xy))^2 \le 1$ for $q(y) > 0$, so $\mathcal{E}(g|g) \le \frac{1}{2}$. For the variance, write

$$\mathrm{Var}(g) = \tfrac{1}{2}\sum_{0\le i,j\le\gamma} (i-j)^2\pi(i)\pi(j) \quad \text{with } \pi(i) = u\{x : |x| = i\} .$$

Set $S = \{i : i \le \frac{\gamma}{4}\}$, $T = \{i : i \ge 3\gamma/4\}$. Clearly,

$$\mathrm{Var}(g) \ge \frac{1}{2}\,\frac{\gamma^2}{4}\,\pi(S)\pi(T) .$$

Now

$$\pi(S) = V(\gamma/4)/|G| \ge A\frac{(\gamma/4)^d}{\gamma^d} = A4^{-d} .$$

For T, choose $x_* \in G$ so $|x_*| = \gamma$. Translating a ball of radius $\gamma/4$ to x^* shows that

$$\pi(T) \geq A4^{-d} .$$

Combining bounds gives (3.2). Using this in (3.1) shows that for any j

$$\|q^{(j)} - u\|_{T.V.} \geq \frac{1}{2}\left(1 - \frac{4^{2d+1}A^2}{\gamma^2}\right)^j .$$

The hypothesis $\gamma \geq A2^{2d+2}$ insures that the eigenvalue bound is larger than $\frac{1}{2}$. Now use $1 - x \geq e^{-2x}$ for $0 < x \leq 1/2$ to complete the proof. □

Remarks: 1. In Theorem 3.1 we have assumed that the probability q puts some mass at the identity to avoid parity problems. This is not necessary. Diaconis and Saloff-Coste [DS-C2, section 2] describe several other ways to work with negative eigenvalues.

2. We have used Lemma 2.2 to bound β_* and it may well happen, in specific cases, that a better bound is known. Thus it is worth noting that the above proof yields, after minor modifications,

THEOREM 3.2. *Let G be a finite group with generating set E. Suppose G has (A, d)-moderate growth with respect to E as in (1.5). Let q be a symmetric probability on G with $\eta = \inf\{q(s), s \in E \setminus \{id\}\} > 0$. Then,*

$$\tfrac{1}{2}\beta_*^n \leq \|q^{(n)} - u\|_{TV} \leq B\beta_*^{n-\gamma^2}$$

where $B = 2^{d(d+3)/4}A^{1/2}\eta^{-d/4}$.

3. David Aldous (personal communication) has pointed out that one can use an elegant bound of Varopoulos [V] and Carne [C] to show that $\gamma^{2-\epsilon}$ steps are not enough for moderate growth problems. We observe here that even less than moderate growth is needed.

PROPOSITION 3.3. *Let G be a finite group with E a symmetric set of generators containing the identity. Suppose that G has diameter γ with respect to E and that $|G| = \beta\gamma^d$ for some positive β and d. Let q be uniform on E as in (1.6). Then,*

$$\|q^{(n)} - u\|_{T.V.} \leq \tfrac{1}{2}\beta^{\frac{1}{2}}e^{-c} \quad \text{for} \quad n = \gamma^2|E|(\tfrac{1}{2}d\log\gamma + c) \quad \text{with} \quad c > 0 ,$$

whereas

$$\|q^{(n)} - u\|_{T.V} \geq \tfrac{1}{2}(1 - 2\beta e^{-c}) \quad \text{for} \quad n = \gamma^2/8(d\log\gamma + c) .$$

Proof: The upper bound follows from the diameter bound on eigenvalues, Lemma 2.2, and $\|q^{(n)} - u\|_{T.V.} \leq \frac{1}{2}|G|^{\frac{1}{2}}\beta_*^n$.

For the lower bound, take $S = \{x : |x| \geq \gamma/2\}$. Let $B = \{x : |x| < \gamma/2\}$. If $|B| \geq \frac{1}{2}|G|$, then by translation $|S| \geq \frac{1}{2}|G|$. If $|B| < \frac{1}{2}|G|$ then also $|S| \geq \frac{1}{2}|G|$. Thus, in all cases, $u(S) = |S|/|G| \geq \frac{1}{2}$. Further, $\|q^{(n)} - u\|_{T.V} \geq |q^{(n)}(S) - u(S)|$.

Now, Carne [C] shows that if $\pi(x), P(x,y)$ is a reversible Markov chain on a countable state space X, then for all x, y, n,

$$P\{X_n = y / X_0 = x\} \leq 2\left(\frac{\pi(y)}{\pi(x)}\right)^{\frac{1}{2}} e^{-d(x,y)^2/2n},$$

with $d(x,y)$ the distance in the graph which has vertex set x and an edge from x to y if $P(x,y) > 0$. Specializing to the present situation, this bound implies that for any n

$$q^{(n)}(S) \leq 2\sum_{x \in S} e^{-|x|^2/2n} \leq 2e^{-\gamma^2/8n}|S|\,.$$

For $n = \gamma^2/8(d\log\gamma + c)$, use of this bound and $u(S) = |S|/|G| \geq \frac{1}{2}$ gives

$$\|q^{(n)} - u\|_{T.V.} \geq \tfrac{1}{2}(1 - 2\beta e^{-c})\,. \qquad \square$$

4. First Examples

This section discusses the Heisenberg group, and some other specific examples where the growth function can be estimated well enough to show that they have moderate growth.

We will use the following notation several times. Fix a positive integer N and consider the set of $N \times N$ matrices with entries in a ring with unity. For $1 \leq i, j \leq N$ we define $E_{i,j}$ to be the $N \times N$ matrix with a 1 in position (i,j) and 0 elsewhere.

EXAMPLE 1. THE HEISENBERG GROUP mod m: This is the group $U_3(m)$ of 3×3 upper triangular matrices with ones on the diagonal and entries mod m. Thus $|U_3| = m^3$. Let $x(t) = id + tE_{1,2}$, $y(t) = id + tE_{2,3}$, $z(t) = id + tE_{1,3}$ for $t \in \mathbb{Z}_m$. Elementary manipulations show

$$\begin{cases} x(s)x(t) = x(s+t),\ y(s)y(t) = y(s+t),\ z(s)z(t) = z(s+t) \\ x(s)y(t) = y(t)x(s)z(st)\,. \end{cases} \tag{4.1}$$

The generating set E in (1.1) is $E = \{Id, x(1), x(-1), y(1), y(-1)\}$. The basic geometric properties of these generators are summarized in the following lemma which, together with Theorem 3.1, gives a proof of Theorem 1.1 stated in the introduction.

LEMMA 4.1. *For $U_3(m)$ with generators given by* (1.1), *the diameter, volume growth, and group order satisfy*

$$m-1 \le \gamma \le m+2\,; \quad V(n) \ge \frac{n^3}{6}\,, \quad 1 \le n \le m\,; \quad |G| \le 8\gamma^3\,.$$

Thus $U_3(m)$ has $(48,3)$*-moderate growth.*

Proof: It is easy to see that $z(t)$ is in the center of the group. Let $x = x(1)$, $y = y(1)$. Let w be a word in x and y. Let $d(w)$ be the minimum number of pairwise adjacent switches required to bring all the x's to the left of all the y's. Thus $d(xyyxxy) = 4$. If a word w has j appearances of y, the commutation relations (4.1) show

$$w = \begin{pmatrix} 1 & j & jk+d \\ 0 & 1 & k \\ 0 & 0 & 1 \end{pmatrix}\,. \tag{4.2}$$

This easily yields that the diameter of $U_3(m)$ in the generators x and y is smaller than $4m$; any word w with $m+j$ appearances of x and $m+k$ appearances of y results in j in position (1,2) and k in position (2,3). Transposing the values of x and y allows an arbitrary d, $0 \le d < m-1$ to be achieved. A slightly more careful version of this argument shows that the diameter of $U_3(m)$ in the generators (1.1) satisfies $m-1 \le \gamma \le m+2$.

For the volume growth, just using x and y as generators and products with a appearances of x and b appearances of y give an interval of $a \cdot b$ distinct values of the (1,3) coordinate provided $a, b > 1$ and $ab \le m$. This implies

$$V(n) \ge \tfrac{1}{2}n^4 \text{ for } 1 \le n \le \sqrt{m}$$
$$V(n) \ge mn^2 \text{ for } \sqrt{m} \le n \le m\,.$$

These clearly imply $V(n) \ge n^3/6$ for $0 < n \le m$. □

Remark: Maria Zack [Z] suggested random walk on the Heisenberg group as a model for cascaded random number generators. Most widely used random number generators are based on a recurrence of the form $X_n = aX_{n-1} + b \pmod p$ for fixed a, b. Chung, Diaconis and Graham [ChDG] studied problems where a and b are allowed to vary randomly. Zack [Z] suggested the following scenario: Let $(\alpha_n, \beta_n, \gamma_n)$ be independent random variables with value in $\mathbb{Z}_m^3$. Define $X_0 = 0$, $X_{n+1} = X_n + \alpha_{n+1}$ and $Y_0 = 0$, $Y_{n+1} = Y_n + \beta_{n+1} \pmod p$. These are usual random walks. Define $Z_0 = 0$, $Z_{n+1} = Z_n + \beta_{n+1}X_n + \gamma_{n+1} \pmod p$. Clearly, the Z_{n+1} process proceeds like the (1,3) coordinate of a random walk on the Heisenberg group.

Theorem 1.1 shows that the walk on $U_3(p)$ takes order p^2 steps to get random for (α_n, β_n) taking values $(\pm 1, 0)$, $(0, \pm 1)$, $(0,0)$ at random, $\gamma_n = 0$. The argument shows that the same conclusion holds for $\alpha_n, \beta_n, \gamma_n$ "small" random variables, e.g. uniform on $[-k, k]$ with fixed k. On the other hand, if α_n and β_n are chosen so that the random walks they generate in the (1,2) and (2,3) positions tend to uniform at a faster rate, the walk on $U_3(p)$ gets random at this faster rate.

The walk of interest in Zack's scenario is *not* the walk on $U_3(p)$ but rather the process generated by the (1,3) coordinate. This walk gets random somewhat faster: for the generators (1.1), an argument based on the Martingale central limit theorem shows that order p steps are necessary and suffice to achieve uniformity in the (1,3) coordinate.

EXAMPLE 2. ANOTHER NON ABELIAN GROUP OF ORDER m^3: For p a prime, it is a classical theorem that there are only two non-isomorphic non Abelian groups of order p^3. When $p = 2$, these are the dihedral group D_4 and the quaternion group. For odd p one of these groups of order p^3 is the Heisenberg group $U_3(p)$. The other will be denoted here by $M_3(p)$ ([Su2, p. 54]) uses the notation $M(p^3)$). This group may be described as follows. Let $\mathbb{Z}_p$ act on $\mathbb{Z}_{p^2}$ by $j \cdot k = (1+jp)k \pmod{p^2}$. Here $\mathbb{Z}_p$ and $\mathbb{Z}_{p^2}$ are written as additive groups and the multiplication takes place in the ring $\mathbb{Z}_{p^2}$. This gives a semi-direct product description of $M_3(p)$ as

$$\{(a,b); a \in \mathbb{Z}_p,\ b \in \mathbb{Z}_{p^2}\} \quad \text{with law} \quad (a,b)(c,d) = (a+c, c \cdot b + d)\ .$$

In particular, $(a,b)^{-1} = (-a, -(1+ap)b)$. In fact, this semi-direct product construction makes sense even if $p = m$ is not a prime and can be used as a definition of $M_3(m)$ for any positive integer m. With this notation, a natural set of generators of $M_3(m)$ is given by

$$E = \{(0,0), (1,0), (-1,0), (0,1), (0,-1)\}\ . \tag{4.3}$$

It results in a walk that goes from $(x,y) \in M_3(m)$ to

$$(x,y),\ \ (x+1,y),\ \ (x-1,y),\ \ (x,(1+mx)+y),\ \ (x,-(1+mx)+y)$$

each with probability 1/5. The following lemma, coupled with Theorem 2.1 shows that $M_3(m)$ has cubic growth and that the generators in (4.3) give a random walk that gets random after order m^2 steps.

LEMMA 4.2. *For the group $M_3(m)$ with generating set* (4.3),

$$m - 2 \le \gamma \le 4m\ ;\ \ V(n) \ge \frac{n^3}{6}\ ,\ \ 1 \le n \le m\ ;\ \ |G| \le 8\gamma^3\ .$$

Proof: Observe that

$$(1,0)(x,y) = (x+1,y) \ , \qquad (x,y)(1,0) = (x+1,(1+m)y)$$
$$n(0,1)(x,y) = (x,(1+mx)+y) \ , \ (x,y)(0,1) = (x,y+1) \ .$$

Thus $(1,0)(0,1) = (1,1)$ and $(1,1)^j = (j,jm+j)$. To write (a,b), write $b = b_1 m + b_2$ with $0 \le b_1 \le m-1$. Then $(a,b) = (1,0)^{a-b}(1,1)^{b_1}(0,1)^{b_2-b_1}$. This shows $\gamma \le 4m$. This formula for (a,b) also shows that $V(n) \ge n^3/6$ for $1 \le n \le m$. The final result follows from $|G| = m^3$. □

EXAMPLE 3. UPPER TRIANGULAR MATRICES: Let $U_N(m)$ be the group of $N \times N$ upper triangular matrices with ones on the diagonal and entries mod m. Thus $|U_N(m)| = m^{\binom{N}{2}}$. For generators, take the matrices $id + E_{i,i+1}$, $1 \le i \le N-1$. This generalizes the Heisenberg group $U_3(m)$. An argument similar to the proof of Lemma 4.1 can be used to show

LEMMA 4.3. *Let $U_N(m)$ be the unipotent upper-triangular matrices. Let*

$$E = \{id \ , \ id \pm E_{i,i+1} \ \ 1 \le i \le N-1\} \ . \tag{4.4}$$

Then the diameter, volume growth, and order satisfy

$$c_1(N)m \le \gamma \le C_1(N)m; \ V(n) \ge c_2(N) n^{\binom{N}{2}}, \ 1 \le n \le \gamma; \ |G| \le C_2(N)\gamma^{\binom{N}{2}} \ .$$

This result shows that with N fixed and m large, the group $U_N(m)$ has moderate growth with $d = \binom{N}{2}$. It follows that the walk with generating set (4.4) requires order m^2 steps to get random.

Remarks: 1. This walk is studied in recent work of Stong [St]. He determined that the second eigenvalue satisfies

$$1 - \frac{A}{m^2 N} \le \beta_1 \le 1 - \frac{a}{m^2 N}$$

with a, A independent of m and N. For fixed N and m large, this eigenvalue bound and Lemma 2.1 show that order $m^2 \log m$ steps suffice. For fixed m and N large, Stong's bounds show that order N^3 steps are enough to ensure convergence. It is easy to show that at least N^2 steps are needed here by considering the last column of the random walk. The volume growth estimates are virtually useless for this fixed m large N case.

2. For prime p, Ellenberg [E] gives sharp bounds for the diameter of $U_N(p)$ with the generators (4.4). Let $f(N,p) = \frac{1}{2}Np + 6N^2 \log p$. He shows there are constants N_0, p_0, c, C, such that

$$cf(N,p) \le \gamma \le Cf(N,p) \quad \text{for} \quad N \ge N_0 \ , p \ge p_0 \ .$$

His proofs are constructive and he shows $N_0 = 10$, $p_0 = 10^4$, $c = 1/60$, $C = 32$. For fixed N, these bounds are of the same order as Theorem 4.4.

3. The walk on $M_3(m)$ takes order m^2 steps to get random, so the 2^{nd} coordinate process (which takes values in a set of size m^2) gets random at a rate faster than a usual random walk. It is not hard to see that it takes m^2 steps to get random. Essentially the same phenomena occurs in $U_3(m)$. In general, it appears that randomness comes in "waves". For example, on $U_N(m)$, for N fixed and m large, the walk generated by (4.4) has the following features. Elements just above the diagonal get random after m^2 steps. Elements two above the diagonal get random after m steps. Elements k above the diagonal get random after $m^{2/k}$ steps. This refers to the coordinates in a particular representation. Philip Hall [Ha] introduced a kind of coordinate system for nilpotent groups with his commutator process. This writes a group element as a product of generators, then first commutators, then second commutators, and so on. It would be marvellous if these coordinates had the behavior of "probability waves" as they seem to for $U_N(m)$.

5. The Doubling Property and Nilpotent Groups

A. The Doubling Property. Let G be a finite group and E a symmetric set of generators for G which contains the identity. We say that G, E satisfies the *doubling property* if for some $A \geq 1$

$$V(1) \leq A \quad \text{and} \quad V(2n) \leq AV(n)\,, \quad n = 0, 1, 2, \cdots. \tag{5.1}$$

Iterating this inequality yields

LEMMA 5.1. *Assume that G, E satisfies the doubling property* (5.1). *Then,*

$$\frac{V(n)}{V(m)} \leq A\left(\frac{n}{m}\right)^d \quad 0 \leq m \leq n < \infty\,, \quad \text{with} \quad d = \frac{\log A}{\log 2}\,. \tag{5.2}$$

In particular, G, E has (A, d)-moderate growth.

Proof: It suffices to write

$$V(n) \leq AV\left(\frac{n}{2}\right) \leq \cdots \leq A^k V\left(\frac{n}{2^k}\right) \leq A^k V(m)$$

provided $\frac{n}{2^{k+1}} < m \leq \frac{n}{2^k}$. Taking $n = \gamma$ in (5.2) gives

$$\frac{V(m)}{V(\gamma)} \geq \frac{1}{A}\left(\frac{m}{\gamma}\right)^d \quad 1 \leq m \leq \gamma, \quad \text{for} \quad d = \frac{\log A}{\log 2}\,.$$

So the doubling property implies moderate growth. □

Remarks: 1. The affine groups and other extensions discussed in section 7 give examples where the doubling property fails to hold uniformly but moderate growth holds with useful constants.

2. Taking $m = 1$ in (5.2) gives

$$V(n) \leq A^2 n^d \quad 1 \leq n < \infty\,, \quad \text{with} \quad d = \frac{\log A}{\log 2}\,.$$

This is a form of polynomial growth which has been extensively studied in the theory of discrete groups. We discuss it below in section 6. There, Gromov's theorem is used to show that the doubling property is, in a sense, a property of the group which depends only on the size of the generating set.

3. For present purposes, the doubling property (through Lemma 5.1) shows that we are in the domain of application of Theorem 3.1. In fact, if G, E satisfies the doubling property (5.1), Theorem 2.3 can be used to prove

$$q^{(n)}(id) \leq CV(\sqrt{n})^{-1} \quad \text{for all } n$$

with a constant C depending only on A and η. This can be refined. The arguments of Hebisch and Saloff-Coste [HeS-C] show that for q as in Theorem 2.3 supported on E, there are positive constants C_1, C_2, c_1, c_2, such that

$$q^{(n)}(x) \leq C_1 V(\sqrt{n})^{-1} \exp\left(-c_1 \frac{|x|^2}{n}\right) \quad \text{for all } n, x \tag{5.3}$$

$$q^{(n)}(x) \geq c_2 V(\sqrt{n})^{-1} \exp\left(-C_2 \frac{|x|^2}{n}\right) \quad \text{for } |x| \leq n\,. \tag{5.4}$$

Here, $|x|$ is the word length and the c_i, C_i's depend only on η and the doubling constant A.

Note that the volume growth function $V(n)$ can be quite erratic. The analogue of (5.3)–(5.4) with $V(\sqrt{n})$ replaced by $n^{d/2}$ for some d simply fails. Also, (5.3)–(5.4) do not hold, in general, for group of moderate growth.

B. Nilpotent Groups. The main result of this section shows that the doubling property holds for the class of nilpotent groups with A only depending on the number of generators and the degree of nilpotency.

Let G be a finite group. Define subgroups $Z_i(G)$, $i = 0, 1, 2, \cdots$ as follows: $Z_0 = \{id\}$, $Z_1 =$ center of G, and Z_i is the subgroup of G corresponding to the center of G/Z_{i-1} in the correspondence theorem. The group G is *nilpotent* if $G = Z_\ell$ for some ℓ. The smallest such ℓ is called the class of G.

Abelian groups are nilpotent of class 1. The Heisenberg group $U_3(m)$ is nilpotent of class 2. Any p-group is nilpotent and any nilpotent group is the direct product of its Sylow p-groups. The affine groups A_p of section 7 are not nilpotent because they have trivial centers for $p > 2$. An introduction to nilpotent groups is given by Rotman [Ro]. Suzuki [Su2] has a thorough treatment and the survey article by Ph. Hall [Ha] is definitive.

THEOREM 5.2. *Let G be a nilpotent group of class ℓ. Let E be a symmetric set of generators for G. Then (G, E) has the doubling property* (5.1) *with $A = A(|E|, \ell)$ depending only on the number of generators and the class ℓ of G.*

Proof: Let $\hat{G}$ be the free nilpotent group on $|E|$ generators of class ℓ. This is an infinite discrete group formed as the quotient of a free group F on $|E|$ generators by the normal subgroup Γ_ℓ. Here Γ_1 is the commutator subgroup $[F, F]$ and $\Gamma_i = [F, \Gamma_{i-1}]$. The group $\hat{G}$ has the property that any class ℓ nilpotent group on $|E|$ generators is a homomorphic image of $\hat{G}$. M. Hall [H] or Magnus et al [MKSo] has further details.

Bass [B] showed that any finitely generated nilpotent group has its volume growth function bounded above and below by polynomials of the same degree. It follows from his work that $\hat{G}$ has volume growth satisfying

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

with $c > 0$ and $D = \sum_{i=1}^{\ell} i f_i(|E|)$, with $f_i(x) = \frac{1}{i} \sum_{d|i} \mu(d) x^{i/d}$ for $\mu(d)$ the Mobius function of elementary number theory.

This certainly shows that $\hat{G}$, with its canonical set of $|E|$ generators satisfies the doubling property (5.1) with $\hat{A} = c^2 2^d$. To complete the proof, we show that the doubling property passes to quotients of $\hat{G}$. Guivarc'h [Gu, Lemma 1.1], specialized to the present situation, implies that for any subgroup $\hat{H}$ of $\hat{G}$, and A, B finite sets in $\hat{G}$, Y a finite subset of $\hat{G}/\hat{H}$,

$$|A||BY| \leq |BA||A^{-1}Y| \,. \tag{5.5}$$

Take $Y = id, A = B(n), B = B(2n)$ balls in $\hat{G}$ of the indicated diameter. Then (5.5) specializes to

$$V_{\hat{G}}(n) V_{\hat{G}/\hat{H}}(2n) \leq V_{\hat{G}}(3n) V_{\hat{G}/\hat{H}}(n) \,.$$

Since $V_{\hat{G}}(3n) \leq V_{\hat{G}}(4n) \leq \hat{A}^2 V_{\hat{G}}(n)$ and $\hat{G}/\hat{H} = G$, the result follows. □

COROLLARY 5.3. *For any positive integers ℓ and e, there exist two positive constants $B = B(\ell, e)$ and $C = C(\ell, e)$ such that for any finite group G nilpotent of class ℓ and any symmetric set of generators $E \subset G$ with $id \in E$ and $|E| = e$, the random walk generated by $q(s) = \frac{1}{|E|}\delta_E(s)$ satisfies*

$$\|q^{(n)} - u\|_{T.V.} \leq Be^{-c} \quad \text{if} \quad n = (1+c)\gamma^2|E| \quad \text{with} \quad c > 0\,,$$

and

$$\|q^{(n)} - u\|_{T.V.} \geq \tfrac{1}{2}e^{-c} \quad \text{for} \quad n = c\gamma^2/C\,.$$

C. *p*-groups and Frattini walks. Let p be a prime. A p-group is a group of order a power of p. These are nilpotent groups and any group of order p^a has class at most $a-1$ ([H, p. 422]) and is generated by at most a generators. Thus, Theorem 5.2 and Corollary 5.3 applies to such groups uniformly as p varies. Any group of order p^2 is Abelian. When p is odd, the two non Abelian groups of order p^3 are $U_3(p)$ and $M_3(p)$ as discussed in section 4. The groups of order p^4, p^5 have been classified. If $f(a,p)$ denotes the number of isomorphism classes of groups of order p^a, then $\log f(a,p) \sim \frac{2}{27}a^3$ for large a and there are bounds on $f(a,p)$ uniform in p. See [Su2, p. 85–95].

For p-groups, the minimal sets of generators have some structure. All minimal generating sets have the same number of elements. These are described by the Frattini subgroup. Recall that, for a group G of order p^a, the Frattini subgroup $\Phi = \Phi(G)$ is defined as the intersection of all subgroups of order p^{a-1}. As shown below, it is often easy to identify. The Burnside basis theorem says that G/Φ is an elementary Abelian p-group which may be regarded as a vector space over $\mathbb{Z}_p$. The dimension d of this vector space is the minimal number of generators of G and any set $x_1, x_2, \ldots x_d$ of coset representatives such that $x_1\Phi, x_2\Phi, \ldots, x_d\Phi$ form a basis of G/Φ give a generating set of G. Conversely, if $x_1, x_2, \ldots, x_d$ generate G then $x_1\Phi, \ldots x_d\Phi$, are a basis of G/Φ. Background, details, and examples may be found in [Su1, chapter 2, section 2; Su2, chapter 4, section 4]. We call a walk supported on a minimal set of generators a Frattini walk.

The Frattini subgroup of the Heisenberg group $U_3(p)$ is its center $\{id + tE_{1,3} : t \in \mathbb{Z}_p\}$ and the walk considered in Theorem 1.1 is a Frattini walk.

For $M_3(p) = \mathbb{Z}_p \ltimes \mathbb{Z}_{p^2}$ (see Example 2, section 4) the Frattini subgroup is $\Phi = \{(0, jp) : 0 \leq j \leq p-1\}$. Again, the walk on $M_3(p)$ considered in section 4 is a Frattini walk.

Further examples are given by $U_N(p)$ and its subgroups. Namely, let $\Gamma = \{(i,j) : 1 \leq i < j \leq N\}$. For $(i,j) \in \Gamma$, let $x_{ij}(s) = id + sE_{i,j}$.

Elementary manipulations show that

$$\begin{cases} x_{ij}(s)x_{ij}(t) = x_{ij}(s+t) \\ x_{ij}(s)x_{kl}(t) = x_{kl}(t)x_{ij}(s) \\ x_{ij}(s)x_{jk}(t) = x_{jk}(t)x_{ij}(s)x_{ik}(st)\ . \end{cases} \tag{5.6}$$

Call a subset $A \subset \Gamma$ *closed* if $(i,j) \in A$, $(j,k) \in A$ then $(i,k) \in A$. The relations (5.6) show that the matrices in $U_N(p)$ with non-zero entries only in a closed set of positions (and zero's in the remaining positions) form a subgroup $U_A(p)$. This subgroup is generated by $\{x_{ij}(\pm 1) : (i,j) \in A\}$. The following lemma identifies the Frattini subgroup of $U_A(p)$.

LEMMA 5.4. *Let $A \subset \Gamma$ be a closed set of indices. Let $A^+ = \{(i,k) :$ for some $j, (i,j)$ and $(j,k) \in A\}$. Then $\Phi(U_A(p)) = U_{A_+}(p)$.*

Proof: The commutation relations (5.6) imply that U_{A^+} is normal in U_A. Further U_A/U_{A^+} is clearly generated by $\{x_{ij}U_{A^+}\}$. These generators commute, and each has order p, so U_A/U_{A^+} is an elementary Abelian p-group. On the other hand, the commutation relations imply $U_{A^+} \subseteq [U_A, U_A]$ (and in fact $U_{A^+} = [U_A, U_A]$). This implies that U_{A^+} is the smallest normal subgroup with elementary Abelian quotient. This characterizes Φ for p-groups. □

For a fixed N and A, let

$$E = \{id,\ x_{ij}(1),\ x_{ij}(-1);\ (i,j) \in A \setminus A_+\}\ . \tag{5.7}$$

The argument for Lemma 4.1 gives

LEMMA 5.5. *Let A be a closed set of indices. The set E of (5.7) is a generating set for $U_A(p)$ which satisfies*

$$c_1(N)p \le \gamma \le C_1(N)p\ , \quad V(n) \ge c_2(N)n^{|A|}\ , \quad |G| \le C_2(N)\gamma^{|A|}\ .$$

We thus see that any of the groups U_A has $(C_2(N), |A|)$-moderate growth. It follows that the random walk takes order p^2 steps to achieve randomness if N is bounded and p is large.

For instance, take

$$\text{Let } A = \{(1,2),(1,3)\cdots(1,N),(2,N),(3,N)\cdots(N-1,N)\}\ .$$

Then U_A has non-zero entries in the first row and last column only. This example is sometimes called the N-dimensional Heisenberg group. Here $A^+ = \{(1,N)\}$. This example is not hard to analyze for large N; it is

essentially random walk on $\mathbb{Z}_p^{2(N-1)}$. For p fixed and N large, order $N \log N$ steps are necessary and suffice to achieve randomness.

As a second example, let

$$A = \{(i,j) : i \leq a, j \geq b\} \quad \text{for } 1 \leq a < b \leq N \text{ fixed.}$$

Then U_A is elementary Abelian, A^+ is empty. Diaconis and Saloff-Coste [DS-C2] showed that order $p^2 d \log d$ steps are necessary and suffice with $d = a(N - b + 1)$.

As a third example, take $A = \{(i,j) : 1 \leq i < j \leq N \text{ with } j \geq i + 2\}$. Then U_A is the Frattini subgroup of U_N. Its Frattini subgroup is U_{A^+} with $A^+ = \{(i,j) : 1 \leq i < j \leq N \text{ with } j \geq i + 4\}$. The quotient U_A/U_{A^+} has order p^{2N-5}. A minimal set of generators consists of the matrices with a ± 1 in one of the two stripes just above the diagonal.

Remarks: 1. The above considerations carry over to Chevalley groups. The analogue of $U_N(p)$ is the subgroup generated by the positive roots. The analogues of U_A are groups generated by closed sets of positive roots. All of these groups are p-groups with explicit Frattini subgroups and associated walks which can be successfully analyzed when the rank N is bounded and p is large.

2. These examples allow us to present some open problems: to what extent does the choice of generators effect the rate of convergence? To focus we recall that for the symmetric group, even restricting attention to generating sets of $(n-1)$ transpositions, the rate of convergence varies from order $n \log n$ (for generators $(12), (13), \ldots, (1,n)$) to order $n^3 \log n$ (for generators $(1,2), (2,3), \ldots, (n-1,n)$). See [DS-C2] for these results.

For the Heisenberg group $U_3(p)$ it can be shown that any automorphism of U_3/Φ lifts to an automorphism of U_3. Thus, all minimal sets of generators are equivalent. For $M_3(p)$ this is no longer true. We do not know the extent to which the choice of generators can effect things for $M_3(p)$. However, Corollary 5.3 show that things cannot vary too widely.

Experience with the circle and symmetric group suggests that if the number of generators is fixed, most sets of generators converge at the same rate. It seems like a difficult but tantalizing problem to make this precise.

3. One final rather specific question. The Burnside problem asks whether or not a finitely generated group whose elements are of order r is finite. This is known to be false for sufficiently large r. However, it is true for $r = 2, 3, 4, 6$. For such r and fixed $m \geq 1$, there is a largest finite group $B(r,m)$ with the property that $B(r,m)$ is generated by m generators and has $s^r = id$ for all s. When r is a prime power, we are in the setting of the present section and enough may be known to make progress on the problem of a random walk supported on a generating set.

D. Dihedral and generalized quaternion groups. For $m = 2^a$ the dihedral group D_m is a 2-group of class a. The constants in Theorem 5.2 and Corollary 5.3 grow exponentially in the class so that the bounds given are useless here. However, for any m, D_m with any set of 2 generators has linear growth, so order m^2 steps are enough. If D_m is represented by $\mathbf{Z}_2 \ltimes \mathbf{Z}_m$ with $\mathbf{Z}_2$ acting on $\mathbf{Z}_m$ by $x \mapsto -x$, the walk generated by $E = \{(0,0),(1,0),(0,1),(0,-1)\}$ can be analyzed by Fourier Analysis to get sharp rates of convergence. A second generating set is $E' = \{(0,0),(1,0),(0,\frac{m-1}{2}),(0,-\frac{(m-1)}{2})\}$. One can show that while both walks require order m^2 steps, the second walk is faster by a factor of 2.

Entirely similar considerations hold for Q_m, the group of generalized quaternions. Here, if w is a primitive $2m^{\text{th}}$ root of 1, let

$$x = \begin{bmatrix} w & 0 \\ 0 & \bar{w} \end{bmatrix}, \quad y = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Then x and y satisfy $x^m = y^2, y^{-1}xy = x^{-1}$, and x, y generate a group of order $4m$. When $m = 2$ this is the quaternion group. When $m = 2^a$, it is a 2-group of class a. Again, Fourier analysis or the moderate growth ideas of section 2 can be used to prove that for $E = \{id, x, x^{-1}, y, y^{-1}\}$, order m^2 steps are necessary and suffice for any m.

When m is not a power of 2, neither D_m nor Q_m is nilpotent.

6. Polynomial Growth and Gromov's Theorem

Let G be a finitely generated group (possibly infinite). Let E be a symmetric set of generators containing the identity. We say that the Cayley graph G, E has (A, d)-*polynomial growth* if

$$V(n) \leq An^d \quad n = 1, 2, \cdots. \tag{6.1}$$

For future reference we note that if (G, E) has (A, d)-polynomial growth then $|E| \leq A$.

The main result of this section shows that for finite groups, polynomial growth implies the doubling property which (cf. section 5) implies moderate growth. Thus random walk on such groups reach stationarity in order γ^2 steps. In the process, we also show that small extensions of groups of polynomial growth have polynomial growth. The main tool is a celebrated theorem of M. Gromov [Gr]. This asserts that a group G of polynomial growth is a finite extension of a nilpotent group (i.e. G contains a nilpotent

group of finite index). Using results of Bass described in Section 5, one can conclude that there exist constants C, D such that

$$C^{-1}n^D \leq V(n) \leq Cn^D , \quad n = 1, 2, \cdots .$$

For infinite groups of polynomial growth this gives a precise and precious description of the volume growth. The book of Varopoulos et al. [VS-CCo] and the article of Hebisch and Saloff-Coste [HeS-C] give applications to random walks.

For finite groups, these bounds contain no information since $D = 0$ works for an appropriate C. However, Gromov has given a version of his theorem which does yield useful information for finite groups having (A, d)-polynomial growth.

THEOREM 6.1 [Gromov]. *Given $A, d > 0$, there is a $C = C(A, d) > 0$ such that any finitely generated group G of (A, d)-polynomial growth contains a nilpotent subgroup N with*

$$[G : N] \leq C \quad \textit{and} \quad \textit{class}(N) \leq C .$$

This result will be applied in the following sections. At present the computation of C as a function of A and d is not effective. This means that the results based on Gromov's theorem have a rather theoretical flavour.

A. Subgroups of small index. Let H be a subgroup of a finite group G. In this section we start with a given set of generators for G and construct generators for H such that the diameter and volume growth are comparable. This construction is applied in section B below in cases where H is nilpotent. Then, H has the doubling property by Lemma 5.1 and so G has the doubling property by comparability.

We begin with a preliminary lemma. The main result is Proposition 6.3 below.

LEMMA 6.3. *Let $R = R^{-1}$ be a generating set of G and let H be a subgroup of G. Assume that $\Omega \subset G$ is such that $\Omega^{-1} = \Omega$, $id \in \Omega$ and $G = H\Omega$. Then, $T = \Omega R \Omega \cap H$ generates H and the diameters γ_G of (G, R) and γ_H of (H, T) satisfy $\gamma_H \leq \gamma_G$.*

Proof (adapted from [Su1, p. 180]): Let $h \in H$ and write

$$h = r_1 \dots r_n$$

with $r_i \in R$. Set $v_0 = \text{id}$, $v_i = r_1 \dots r_i$. By hypothesis there exist $h_i \in H$, $\omega_i \in \Omega$ such that $v_i = h_i \omega_i$. Moreover, we can choose $\omega_0 = \omega_n = \text{id}$. Thus, we have

$$h = \omega_0 r_1 \omega_1^{-1} \omega_1 r_2 \omega_2^{-1} \dots \omega_{n-1} r_n \omega_n^{-1} = w_1 \dots w_n$$

with

$$w_i = \omega_{i-1} r_i \omega_i^{-1} = h_{i-1}^{-1} r_1 \dots r_{i-1} r_i r_i^{-1} \dots r_1^{-1} h_i$$
$$= h_{i-1}^{-1} h_i \in H \cap \Omega R \Omega = T$$

Thus, T generates H and $\gamma_H \leq \gamma_G$. □

It seems difficult to compare the growth of (G, R) and (H, T) without further hypotheses.

PROPOSITION 6.3. *Let G be a group. Assume that $E = E^{-1}$ is a finite symmetric set of generators containing the identity. Let H be a subgroup of G such that $[G : H] = k$. Then, there is an integer $1 \leq \nu \leq k$ such that $G = HE^\nu$ and $\Sigma = E^{3\nu} \cap H$ generates H. Moreover, the diameters and growth functions of (G, E) and (H, Σ) satisfy*

$$\gamma_H \leq \gamma_G \; ; \quad V_H(n) \leq V_G(3\nu n) \; , \quad n = 1, 2, \dots \tag{6.2}$$

$$\gamma_G \leq 2(k+1)(3\nu\gamma_H + 1) \; ; \quad V_G(\nu n) \leq |S|^\nu V_H(n) \; , \quad n = 1, 2, \dots . \tag{6.3}$$

Proof: First, we claim that $G = HE^k$. To see this, consider the quotient space $X = \{Hg : g \in G\}$. The graph (G, E) induces a graph structure on X with edge set $\{(Hg, Hge) : g \in G, e \in E\}$. This graph is connected and has diameter smaller or equal to $k = |X|$. This shows that E^k contains a set of coset representatives. Thus there is an integer $1 \leq \nu \leq k$ such that $G = HE^\nu$. We can now apply Lemma 2 with $R = \Omega = E^\nu$. This shows that $\Sigma = E^{3\nu} \cap H$ generates H with $\gamma_H \leq \lceil \gamma_G / \nu \rceil$. Moreover, $\Sigma^n \subset E^{3\nu n}$ shows that $V_H(n) \leq V_G(3\nu n)$. This proves (6.2).

In order to prove the diameter bound in (6.3), write $\gamma_G = 2\alpha(3\nu\gamma_H + 1)$ for some $\alpha > 0$. Then, we can find at least $\lfloor \alpha \rfloor$ disjoint balls of radius $3\nu\gamma_H$ in (G, E). But, each of these balls has volume at least $|H|$ since $H = \Sigma^{\gamma_H} \subset E^{3\nu\gamma_H}$. Thus, $\lfloor a \rfloor |H| \leq |G|$. This shows that $\alpha \leq k + 1$ and thus $\gamma_G \leq 2(k+1)(3\nu\gamma_H + 1)$.

Finally, to prove the volume bound in (6.3), we claim that $E^{\nu(n+1)} \subset \Sigma^n E^\nu$. This easily follows by induction from $E^{2\nu} \subset \Sigma E^\nu$. But, for $r, r' \in R = E^\nu$, there exist $t \in R$, $h \in H$ such that $rr' = ht$ because $G = HR$. Hence, $rr't^{-1} \in E^{3\nu} \cap H = \Sigma$ and $rr' = rr't^{-1}t \in \Sigma S^\nu$. This ends the proof of Proposition 6.3.

□

B. Examples. Simple random walk on the "circle" $\mathbf{Z}_m$ is well understood. Order m^2 steps are necessary and suffice for stationarity. We begin by studying all extensions of $\mathbf{Z}_m$ of degree 2. Even here there are some surprises: simple random walk so extended can get random in order m steps. We start by delineating all extensions of degree 2. Let D_m be the dihedral group. This can be constructed by letting $\mathbf{Z}_2$ act on $\mathbf{Z}_m$ by $x \mapsto -x$.

PROPOSITION 6.4. *Let $m = p_1^{a_1} \dots p_l^{a_l}$ be a product of distinct odd prime powers. There are 2^l non-isomorphic extensions of degree 2 of $\mathbb{Z}_m$. Each one of them is isomorphic to a direct product $D_{m_1} \times \mathbb{Z}_{m_2}$ where m_1 is a product of a subset of $\{p_1^{a_1}, \dots, p_l^{a_l}\}$ and $m_2 = m/m_1$.*

Proof: Let G be an extension of degree 2 of $\mathbb{Z}_m$. Then $G = \mathbb{Z}_2 \ltimes \mathbb{Z}_m$ by the Schur-Zassenhaus theorem. Thus, G is specified once we specify an action of $\mathbb{Z}_2$ on $\mathbb{Z}_m$; that is, an automorphism of $\mathbb{Z}_m$ of order 2. An automorphism is determined by the image of 1 in $\mathbb{Z}_m$ and must be of the form $j \to bj$ for $b \in \mathbb{Z}_m$ with $b^2 = 1$. Write $b = (b_1, \dots, b_l)$ with $b_i \in \mathbb{Z}_{p_i^{a_i}}$ using the Chinese remainder theorem. Thus $b_i = \pm 1$ and any choice is possible. Different choices lead to non-isomorphic extensions. Finally, fix a b and so an action. If m_1 is the product of prime powers where $b_i = -1$, then $G \cong D_{m_1} \times \mathbb{Z}_{m_2}$. □

For instance, if $m = 15$, there are 4 choices for b of order 2: $\{1, 4, 11, 14\}$. The 4 extensions are $\mathbb{Z}_2 \times \mathbb{Z}_{15}$, D_{15}, $D_3 \times \mathbb{Z}_5$, $D_5 \times \mathbb{Z}_3$. Even m can be handled by the same techniques. The extensions of degree 2 of cyclic groups of size 2^k are classified in [Su2, Theorem 4.1, p. 54].

THEOREM 6.5. *Let m be an odd integer. Let $G = \mathbb{Z}_2 \ltimes \mathbb{Z}_m$ be an extension of degree 2 of $\mathbb{Z}_m$, with generating set*

$$E = \{(0,0), (1,0), (0,1)(0,-1)\}.$$

Let m_1, m_2 be the odd integers given by Proposition 6.2 such that $G = D_{m_1} \times \mathbb{Z}_{m_2}$ and set $m_* = \max\{m_1, m_2\}$. Then, (G, E) has diameter of order m_* and satisfies the doubling property uniformly in m. Thus, order m_*^2 steps are necessary and suffice for the associated random walk to be close to equilibrium.

Proof: The isomorphism $\mathbb{Z}_2 \ltimes \mathbb{Z}_m \cong D_{m_1} \times \mathbb{Z}_{m_2}$ can be taken of the form

$$(x, y) \to ((x, y_1), y_2)$$

with $y_i \equiv y \pmod{m_i}$, $i = 1, 2$. Under this map the generating set E becomes

$$E = \{((0,0),0), ((1,0),0), ((0,1),1), (0,-1),-1)\}.$$

It is an easy matter to see that E^3 contains the elements $((0, \pm 1), 0)$ and $((0,0), \pm 1)$ and to deduce that (G, E) has diameter of order m_*.

We now show that all the graphs (G, E) in Theorem 6.5 have the doubling property (uniformly). By Proposition 6.3, this boils down to the fact that the cyclic groups $\mathbb{Z}_m$ have the doubling property uniformly for any possible choice of at most 64 generators (here, 64 is a crude estimate for the size of $E^3 \cap \mathbb{Z}_m$). This was proved in section 5. □

Remark: As the action of $\mathbb{Z}_2$ on $\mathbb{Z}_m$ varies, m_* varies in $[\sqrt{m}, m]$ and essentially any values in this interval can be obtained.

It seems like a natural project to try to understand better the extensions of $\mathbb{Z}_m$ by $\mathbb{Z}_r$, for fixed r. Namely, let $G = G_\theta = \mathbb{Z}_r \ltimes_\theta \mathbb{Z}_m$ for some action θ, and fix

$$E = \{(0,0),(0,1),(0,-1),(1,0),(-1,0)\} \tag{6.4}$$

as generating set. Proposition 6.3 and further elementary considerations show that G_θ, E has (A,d)-moderate growth for some constant $A = A(r)$ and $d = d(r) \leq r$, uniformly in m and θ. It follows that the corresponding random walk is approximately uniformly distributed after order γ^2 steps. This, however, does not tell us what the diameter γ is in terms of r, m, θ. Moreover, a precise study of the volume growth would be of interest.

We will not pursue this here in complete generality but we now describe two examples with $r = 3$ to illustrate further what is going on.

THEOREM 6.6. *Let $m = k^3 - 1$ or $m = k^2 + k + 1$ with $k = 2, 3, \ldots$. For such m and $i \in \mathbb{Z}_3$, $j \in \mathbb{Z}_m$, $\theta_i(j) = k^i j \pmod m$ defines an action of $\mathbb{Z}_3$ on $\mathbb{Z}_m$. Let $G = \mathbb{Z}_3 \ltimes_\theta \mathbb{Z}_m$ with generating set E at (6.4).*

1. *If $m = k^3 - 1$, the Cayley graph G, E has diameter of order $k = m^{1/3}$ and cubic growth.*
2. *If instead $m = k^2 + k + 1$, the Cayley graph G, E has diameter of order $k = m^{1/2}$ and quadratic growth.*

Proof: For $(x, y) \in G$ we have

$$(x,y)(x',y') = (x + x', k^{x'} y + y') \,,$$

and thus

$$(0,y)(x,0)(0,-y) = (0, k^x y) \quad \text{for} \quad x = 0,1,2 \,, \quad y \in \mathbb{Z}_m \,. \tag{6.5}$$

In case 1 where $m = k^3 - 1$, we can write any $0 \leq y \leq m - 1$ as $y = a_1 + a_2 k + a_3 k^2$ with $0 \leq a_i \leq k - 1$. Using (6.5), it follows that the diameter γ satisfies $\gamma \leq 3(k+1)$. Also, elementary considerations give $\gamma \geq (k-1)/2$. A similar argument shows that G, E has cubic growth.

For case 2 where $m = k^2 + k + 1$, we can write any $0 \leq y \leq m - 1$ as $y = a_1 + a_2 k$ with $0 \leq a_i \leq k + 1$. Using (6.5) again, we get $\frac{1}{2}(k-1) \leq \gamma \leq 2(k+1)$. Similar further arguments show that the growth is quadratic in this case. □

Here is another class of examples which seems worth making explicit.

THEOREM 6.7. *Fix $A > 0$. Let G be a group of order mp^n where p is a prime and $(m,p) = 1$. Also, let E be a symmetric set of generators of G containing the identity. Then, there exists a constant $C = C(A)$ such that the Cayley graph (G,E) has the C-doubling property provided that*

$$m \leq A\,, \quad n \leq A \quad \text{and} \quad |E| \leq A\,.$$

Thus, if the Cayley graph (G,E) satisfies the above hypotheses and has diameter γ, order γ^2 steps are necessary and suffice for the simple random walk on (G,E) to be close to uniform.

Proof: Let N be a p-Sylow subgroup of G. It has order p^n. Thus, N is a nilpotent group of class $\mathrm{cl}(N) \leq n \leq A$; see section 5.C. By Proposition 6.3, the volume growth of (G,E) is comparable to the volume growth of (N,Σ) where $\Sigma = E^{3m} \cap N$. Since Σ has at most $|E|^{3m} \leq A^{3A}$ and N is nilpotent of class at most A, we know that (N,Σ) is $C(A)$-doubling for some constant $C(A)$. This clearly yields the desired result. □

Remark: The constant $C(A)$ in Theorem 6.7 can be made explicit in principle. It is probably of the type A^{A^A}.

C. Groups of polynomial growth. We now present some theoretical consequences of Proposition 6.3 and Gromov's theorem.

THEOREM 6.8. *Let (G,E) be a Cayley graph of (A,d)-polynomial growth where E is a symmetric set of generators containing the identity. Let γ be the diameter. There exist constants $C_i = C_i(A,d)$, $1 \leq i \leq 5$ such that:*

1. *The graph (G,E) has the C_1-doubling property.*
2. *Any probability q such that $\eta = \inf_{s\in E}\{q(s)\} > 0$ satisfies*

$$\|q^{(n)} - u\|_{TV} \leq C_2 e^{-c} \quad \text{for} \quad n = (1+c)\gamma^2/\eta \quad \text{with} \quad c > 0\,.$$

3. *If, moreover, q is supported in E and γ is large enough, we have*

$$\|q^{(n)} - u\|_{TV} \geq \tfrac{1}{2} e^{-c} \quad \text{for} \quad n = c\gamma^2/C_3\,.$$

4. *Let $\widetilde{E}$ be another symmetric set of generators such that $|\widetilde{E}| \leq A$. Then, $(G,\widetilde{E})$ is C_5-doubling. It follows that there exist $A' = A'(A,d)$ and $d' = d'(A,d)$ such that $(G,\widetilde{E})$ has (A',d')-polynomial growth. This also implies that the diameters $\gamma, \tilde{\gamma}$ satisfies*

$$C_5^{-1} \log\gamma \leq \log\tilde{\gamma} \leq C_5 \log\gamma\,.$$

Proof: The first assertion follows from Proposition 6.3 and Gromov's theorem as in the proof of Theorem 6.7. Assertions 2 and 3 then follows from section 5.A. The last statement about diameters follows from

$$\gamma \leq |G| \leq A\gamma^d \ , \quad \tilde{\gamma} \leq |G| \leq A'\tilde{\gamma}^{d'} \ . \qquad \square$$

Remark: It would be nice to have a proof of this theorem that yields explicit constants. This could be achieved either by getting explicit constants in Gromov's Theorem 6.1 or by avoiding the use of Gromov's result in the proof of Theorem 6.8. However, proving that a group having (A, d)-polynomial growth satisfies the C_1-doubling property without using Gromov's theorem seems to be a serious challenge!

7. Semi-direct Products and Normal Extensions

This section treats the affine group $\mathrm{mod}\, p$ and other extensions of groups of moderate growth. These often do not have polynomial growth but the moderate growth theory applies.

A. The affine group. Let p be an odd prime. Let A_p be the "$ax + b$" group $(\mathrm{mod}\, p)$. This is the group of pairs (a, b) with $b \in \mathbf{Z}_p$, $a \in \mathbf{Z}_p^*$ and

$$(a, b)(a', b') = (aa', a'b + b'); \ , \quad id = (1, 0)$$

and thus $(a, b)^{-1} = (a^{-1}, -a^{-1}b)$. Random walks on A_p arise in the study of random number generators as explained in [ChDG]. They have also been studied by Hildebrand [Hi] as discussed below.

Let α be a generator of $\mathbf{Z}_p^*$, let β be any non-zero element of $\mathbf{Z}_p$. Consider

$$E = \{(1, 0), \ (\alpha, 0), \ (\alpha^{-1}, 0), \ (1, \beta), \ (1, -\beta)\} \ . \tag{7.1}$$

The diameter of $\mathbf{Z}_p$ with β as generator is $(p-1)/2$. The diameter of $\mathbf{Z}_p^*$ with α as generator is $(p-1)/2$. It follows that the diameter γ of A_p is at most p (and at least $(p-1)/2$). It is clear that $V(n) \geq n^2$ for $1 \leq n \leq p$ and that $|A_p| \leq \gamma^2$. It follows that A_p has moderate growth with $A = 1$, $d = 2$. From this, Theorem 3.2 shows that order p^2 steps are necessary and suffice to drive the variation distance to zero.

THEOREM 7.1. *For random walk on A_p with generating set given by* (7.1), *there are universal constants $a, a', b, b' > 0$ such that*

$$a'e^{-b'k/p^2} \leq \|q^{(k)} - u\|_{\mathrm{T.V.}} \leq ae^{-bk/p^2} \ .$$

Remarks: 1. Diaconis [D, chapter 3] studied this problem using the character theory of A_p. The analysis only managed to show that order $p^2 \log p$ steps suffice. Using highly original methods, Hildebrand [Hi] showed that the second coordinate of walks generated by a set like (7.1) get random extremely rapidly (order $(\log p)^2$ steps). The first coordinate is performing simple random walk and so takes p^2 steps. A separate study of the second coordinate does not seem possible with the technique of this paper.

2. For general odd m the group A_m is defined as $\mathbb{Z}_m^* \ltimes \mathbb{Z}_m$. The minimum number of generators depends on the prime decomposition of m. If $m = p_1^{a_1} p_2^{a_2} \cdots p_\ell^{a_\ell}$ as a product of distinct odd prime powers then $\mathbb{Z}_m^* = U(p_1^{a_1}) U(p_2^{a_2}) \cdots U(p_\ell^{a_\ell})$ with $U(p^a) \cong \mathbb{Z}(p^a - p^{a-1})$ (see e.g. [IR, p. 46].) This group is generated by ℓ generators (and this is the minimum since $\mathbb{Z}_m^*$ has Z_2^ℓ as a homomorphic image). With a fixed number ℓ of generators, these groups have moderate growth with $d = \ell$ for an explicit constant A as m varies. It follows that m^2 steps are necessary and suffice to reach uniformity (for bounded ℓ). The problem is open for situations like $m = p_1 p_2 \cdots p_\ell$, the product of the first ℓ primes, as ℓ varies.

3. Let p be an odd prime. Suppose r divides $p-1$. Then $\mathbb{Z}_p^*$ is a cyclic group of order $p-1$ and so contains a cyclic group of order r. The associated walk on $\mathbb{Z}_r \ltimes \mathbb{Z}_p$ can be studied by the methods of section 6 if r is "small" (e.g. $r = 2$ giving a dihedral group) or by the method of this section if r is "large" (e.g. $r = (p-1)/2$). For other values of r (e.g. $r \sim \log p$ or $\sqrt{p}$) we do not know how to use present techniques to get the right answer. Perhaps, Hildebrand's method can be used to show that the walk on the second coordinate requires order $(\log p)^2$ steps whenever $r \geq \log p$. The walk on the first coordinate always requires order r^2 steps.

B. Normal Extensions. We now examine various ways of putting together two groups G and K having moderate growth. A succinct, readable account of normal group extensions appears in [H, chapter 15]. One must specify an action of G on K here denoted $g \cdot k$. Further, a 'factor set' must be specified. This gives, for $g, g' \in G$ an element $\langle g, g' \rangle \in K$ which satisfies

$$g \cdot (g' \cdot k) = \langle g, g' \rangle^{-1} (gg' \cdot k) \langle g, g' \rangle$$

and

$$\langle gg', g'' \rangle (g'' \cdot \langle g, g' \rangle) = \langle g, g'g'' \rangle \langle g, g'' \rangle \, .$$

From these ingredients, an extension L of K by G can be constructed:

$$L = \{(g, k), g \in G, h \in K\}$$

with product

$$(g, k)(g', k') = (gg', \langle g, g' \rangle (g' \cdot k) k') \, .$$

The set $\{(id,h) : h \in K\}$ is isomorphic to K which is normal in L with $L/K \cong G$. Further, all normal extensions arise from this construction.

If the factor set is trivial, $\langle g,g'\rangle \equiv id$, the extension is called a semi-direct product. The affine group A_p is given by this construction. The quaternions are a non trivial extension of $\mathbb{Z}_4$ by $\mathbb{Z}_2$.

Suppose E_G and E_K are symmetric sets of generators of G and K. Identify these with subsets of L as $\{(g,id) : g \in E_G\}$, $\{(id,k);\ h \in E_K\}$. The following theorem gives a sufficient condition for all extensions to have moderate growth. It is followed by examples and a counter example showing how things can change if its condition is violated.

THEOREM 7.2. *Suppose G, E_G and K, E_K have diameters γ_G, γ_K with $\gamma_K \le \theta\gamma_G$ for some $\theta \ge 1$. If G and K have moderate growth with constants A_G, d_G and A_K, d_K then any extension L with generating set $E_L = E_G \cup E_K$ has diameter γ_L satisfying $\gamma_G \le \gamma_L \le (1+\theta)\gamma_G$ and moderate growth with $A_L = 2^{d_G+d_K}\theta^{d_K}A_GA_K$, $d_L = d_G + d_K$.*

Proof: For any extension L, we have

$$(g,id)(g',id) = (gg',\langle g,g'\rangle) \quad \text{and} \quad (id,k)(id,k') = (id,kk)\ .$$

These yield $\gamma_G \le \gamma_L \le \gamma_G+\gamma_K \le (1+\theta)\gamma_G$. Next, fix n and $n_1+n_2 = n$ with $n_1 \le \gamma_G$, $n_2 \le \gamma_K$.

$$\frac{V_L(n)}{|L|} = \frac{V_L(n)}{|G||K|} \ge \frac{V_G(n_1)}{|G|}\frac{V_K(n_2)}{|K|} \ge \frac{1}{A_GA_K}\left(\frac{n_1}{\gamma_G}\right)^{d_G}\left(\frac{n_2}{\gamma_K}\right)^{d_K}$$
$$\ge \frac{1}{A_GA_K\theta^{d_K}}\left(\frac{n_1}{\gamma_L}\right)^{d_G}\left(\frac{n_2}{\gamma_L}\right)^{d_k}\ .$$

This inequality will be used in several cases.

Case 1. $n = 2m$ with $m \le \min(\gamma_G,\gamma_K)$. Then, with $n_1 = n_2 = m$.

$$\frac{V_L(n)}{|L|} \ge \frac{1}{A_GA_K\theta^{d_K}}\left(\frac{n}{2\gamma_L}\right)^{d_G}\left(\frac{n}{2\gamma_L}\right)^{d_K} = \frac{1}{A_GA_K\theta^{d_K}2^{d_G+d_K}}\left(\frac{n}{\gamma_L}\right)^{d_G+d_K}\ .$$

Case 2. $n = 2m-1$ with $m \le \min(\gamma_G,\gamma_K)$. Then, with $n_1 = m, n_2 = m-1$,

$$\frac{V_L(n)}{|L|} \ge \frac{1}{A_GA_K\theta^{d_K}}\left(\frac{2m}{2\gamma_L}\right)^{d_G}\left(\frac{2(m-1)}{2\gamma_L}\right)^{d_K}$$
$$\ge \frac{1}{A_GA_K\theta^{d_K}2^{d_G+d_K}}\left(\frac{2(m-1)}{2m-1}\right)^{d_K}\left(\frac{n}{\gamma_L}\right)^{d_G+d_K}\ .$$

Case 3. $\gamma_K \leq \frac{n}{2} \leq \gamma_G$. Then,

$$\frac{V_L(n)}{|L|} \geq \frac{V_G\left(\frac{n}{2}\right)}{|G|}\frac{V_K\left(\frac{n}{2}\right)}{|K|} \geq \frac{1}{A_G}\left(\frac{n}{2\gamma_G}\right)^{d_G} \geq \frac{1}{A_G}\left(\frac{n}{2\gamma_G}\right)^{d_G+d_K}$$
$$\geq \frac{1}{A_G 2^{d_G+d_K}}\left(\frac{n}{\gamma_L}\right)^{d_G+d_K} .$$

Case 4. $\gamma_G \leq \frac{n}{2} \leq \gamma_K$ and $n \leq \gamma_L$. Then

$$\frac{V_L(n)}{|L|} \geq \frac{V_G\left(\frac{n}{2}\right)}{|G|}\frac{V_K\left(\frac{n}{2}\right)}{|K|} \geq \frac{1}{A_K}\left(\frac{n}{2\gamma_K}\right)^{d_K} \geq \frac{1}{A_K(2\theta)^{d_K}}\left(\frac{n}{\gamma_L}\right)^{d_G+d_K} .$$

This exhausts the cases since $\gamma_K \leq \gamma_G < \frac{n}{2}$ or $\gamma_G \leq \gamma_K < \frac{n}{2}$ implies $n = \frac{n}{2} + \frac{n}{2} > \gamma_K + \gamma_G \geq \gamma_L$. □

Remarks: 1. Theorem 7.2 is useful for γ_K small or moderate with respect to γ_G. The results then show that all extensions have moderate growth so that the convergence results of Theorem 1.2 are in force. The affine groups offer examples where Theorem 7.2 is effective. Other examples are described below.

2. If γ_G is small with respect to γ_K, Theorem 7.2 is not very useful because the constant θ is large. The examples of extensions of degree 2 and 3 of $\mathbb{Z}_m$ described in section 6.B show that the degrees of volume growth needn't simply add: $\mathbb{Z}_2$ has growth of degree 0, $\mathbb{Z}_m$ with the canonical generators has linear growth, but $\mathbb{Z}_2 \ltimes \mathbb{Z}_m$ may have quadratic growth. Thus, the condition $\gamma_K \leq \theta\gamma_G$ in Theorem 7.2 cannot be removed.

3. Semi-direct products give examples of groups of moderate growth which do not have polynomial growth. For instance, for the affine group A_p with generating set (7.1), one can show that $V(n) \geq e^{cn}$ for $1 \leq n \leq \log p$ where c is a universal positive constant. This proves that (6.1) cannot hold with fixed A, d when p tends to infinity.

4. Elementary considerations show that when L above is a semi-direct product, the center Z_L of L can be identified as

$$Z_L = \{(id, k) : \ k \in Z_K \ \text{ and } \ g \cdot k = k \text{ for all } \ g \in G\} .$$

Thus $Z_L = Z_K \cap C_L(G)$. For example, the affine group is centerless for $p \neq 2$ and the dihedral groups D_m are centerless unless m is a power of 2. In these cases, the groups are not nilpotent so the present arguments offer the only currently available route to studying random walk.

C. Further examples. A large class of examples can be obtained by using the action of $\mathbf{Z}_r$ on $\mathbf{Z}_p$, where (say) p is a prime and r divides $p-1$. Here, for $j \in \mathbf{Z}_r$ and $a \in \mathbf{Z}_p^*$ satisfying $a^r = 1$, define $\theta_j(x) = a^j x$ for $x \in \mathbf{Z}_p$. For instance, consider $\mathbf{Z}_r \ltimes U_3(p)$ where the action of $\mathbf{Z}_r$ on $U_3(p)$ is given by

$$\Theta_j\left(\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}\right) = \begin{pmatrix} 1 & a^j x & a^{2j} z \\ 0 & 1 & a^j y \\ 0 & 0 & 1 \end{pmatrix} \quad x, y, z \in \mathbf{Z}_p \,.$$

For r comparable to p (e.g. $r = p-1$) Theorem 7.2 applies whereas for small fixed r Theorem 6.7 can be used.

A second class of examples uses the natural action of $G = U_d(p)$ on $K = \mathbf{Z}_p^d$ (see section 4 for notation). For fixed d and large p and with their natural generators, these groups have comparable diameters of order p. Then, Theorems 7.2 and 2.3 combine to show that order p^2 steps are required for randomness. Diaconis and Graham [DG] studied walks of this type with $p = 2$ and d large as examples of repetitive computer algorithms operating in the presence of a bad bit.

The upper triangular group T_N. Let p be a prime. The group $T_N(p)$ of upper triangular, invertible, $N \times N$ matrices with entries mod p has order $|T_N(p)| = (p-1)^N p^{\binom{N}{2}}$. It contains $U_N(p)$ as a normal subgroup. The quotient $T_N/U_N \cong (\mathbf{Z}_p^*)^N$ and T_N is a semi-direct product of U_N by $(\mathbf{Z}_p^*)^N$. Let α be a generator of $\mathbf{Z}_p^*$. Let $E_i(\alpha)$ be a diagonal matrix with α in the i^{th} place and ones elsewhere. Let $G = (\mathbf{Z}_p^*)^N$, with generating set $\{E_1(\alpha), E_1(\alpha^{-1}), \cdots, E_N(\alpha^{-1})\} = E_G$ and $K = U_N(p)$ with generating set as in (4.4). Then, for N fixed, $L = T_N$ with $E_L = E_G \cup E_K$ has diameter of order p. Thus, for N fixed and p large the walk on T_N with these generators gets random after order p^2 steps.

Polynomials under composition. Consider the set G_N of polynomials with coefficients (mod p) taken (mod x^{N+1}) of form $a_1 x + a_2 x^2 + \cdots + a_N x^N$ with $a_1 \in \mathbf{Z}_p^*$ and $a_i \in \mathbf{Z}_p$, $2 \le i \le N$. This set forms a group under composition. The subset G'_N with $a_1 = 1$ is a normal subgroup of order p^{N-1}. As generators for G'_N choose $x + x^2, x + x^3, \cdots x + x^N$. G'_N in these generators has moderate growth since it is a p-group. From here, G_N can be handled by the theory of this section. For N fixed and p large the natural walk requires order p^2 steps to get random.

The group G_N can be realized as a subgroup of T_N by mapping $f \in G_N$

into a matrix as follows. Write

$$f = a_1 x + \cdots + a_N x^N$$
$$\underbrace{f \circ f \cdots \circ f}_{k} = a_1^k x + a_2^k x^2 + \cdots + a_N^k x^N \ .$$

Define $m(f)$ to have i^{th} row beginning with $i-1$ zeros and then $a_1^i, a_2^i, \cdots, a_{N-i}^i$. The map $f \mapsto m(f)$ is an injective homomorphism. For small N, G_N is familiar: $G_1 \cong \mathbb{Z}_p^*, G_2 \cong A_p$, $G_3 \cong \mathbb{Z}_p^* \ltimes \mathbb{Z}_p \times \mathbb{Z}_p$, $G_4 \cong \mathbb{Z}_p^* \ltimes U_3(p)$. Johnson (1988) has further information about these groups.

Acknowledgement. We thank David Aldous, Jordan Ellenberg, Paul Fong, Jeffrey Silver, and Maria Zack for their help.

References

[AD] D. Aldous, P. Diaconis, Shuffling cards and stopping times, Amer. Math. Monthly 93 (1986), 333–348.

[B] H. Bass, The degree of polynomial growth of finitely generated nilpotent groups, Proc. London Math Soc. 25 (1982), 603–614.

[C] T. Carne, A transmutation formula for Markov chains, Bull. Sc. Math. 2^{nd} série 109 (1985), 399–405.

[ChDG] F. Chung, P. Diaconis, R.L. Graham, A random walk problem arising in random number generation, Ann. Prob. 15 (1987), 1148–1165.

[CoS-C] T. Coulhon, L. Saloff-Coste, Puissance d'un operateur regularisant, Ann. Inst. H. Poincare Proba. Stat. 26 (1990), 419–436.

[D] P. Diaconis, Group Representations in Probability and Statistics, Institute of Mathematical Statistics, Hayward, CA (1988).

[DG] P. Diaconis, R.L. Graham, An affine walk on the hypercube, Jour. Comp. Appl. Math 41 (1992), 215–235.

[DS-C1] P. Diaconis, L. Saloff-Coste, Nash's inequality and convergence of finite Markov chains to equilibrium, Technical report (1992).

[DS-C2] P. Diaconis, L. Saloff-Coste, Comparison theorems for random walk on groups, To appear in Ann. Prob. 21 (1993).

[DS-C3] P. Diaconis, L. Saloff-Coste, An application of Harnack inequalities to random walk on nilpotent quotients, Technical report, Dept. of Mathematics, Harvard University (1993).

[DStr] P. Diaconis, D. Stroock, Geometric bounds for reversible Markov chains, Ann. Appl. Prob. 1 (1991), 36–61.

[E] J. Ellenberg, A sharp diameter bound for upper triangular matrices, Senior honors thesis, Dept. Math., Harvard University (1993).

[Gr] M. Gromov, Groups of polynomial growth and expanding maps, Publ. Math. I.H.E.S. 53 (1981), 53–73.

[Gu] Y. Guivarc'h, Croissance polynomiale et periodes des fonctions harmoniques, Bull. Soc. Math. France, 101 (1973), 333–379.

[H] M. HALL, The Theory of Groups, 2nd ed, Chelsea, New York (1976).
[Ha] P. HALL, Nilpotent Groups. In Collected Works of Philip Hall, Oxford University Press, Oxford (1957), 417–462.
[He] W. HEBISCH, On heat kernels on Lie groups. Math Zeit. 210 (1992), 593–605.
[HeS-C] W. HEBISCH, L. SALOFF-COSTE, Gaussian estimates for Markov chains, Ann. Proba. 21 (1993), 673-709.
[Hi] M. HILDEBRAND, Random processes of the form $X_{n+1}=a_n X_n+b_n \pmod p$, Ann. Proba. 21 (1993), 710-720.
[IR] K. IRELAND, M. ROSEN, Elements of Number Theory, Bogdon and Quigley, Tarrytown, New York (1972).
[J] D.L. JOHNSON, The groups of formal power series under substitutions, Jour. Austral. Math Soc. (A) 45 (1988), 296–302.
[MKSo] W. MAGNUS, A. KARRASS, D. SOLITAR, Combinatorial Group Theory, 2nd ed., Dover, New York (1976).
[Ro] J. ROTMAN, The Theory of Groups: An Introduction, 3rd ed., Allyn and Bacon, Boston (1984).
[St] R. STONG, A random walk on the Heisenberg group, Technical report, Department of Mathematics, U.C.L.A. (1992).
[Su1] M. SUZUKI, Group Theory I, Springer Verlag, New York (1982).
[Su2] M. SUZUKI, Group Theory II, Springer Verlag, New York (1986).
[T] J. TITS, Appendix to Gromov's paper, Publ. Math. I.H.E.S. 58 (1981), 74–78.
[V] N. VAROPOULOS, Long range estimates for Markov chains, Bull. Sc. Math. 109 (1985), 225–252.
[VS-CCo] N. VAROPOULOS, L. SALOFF-COSTE, TH. COULHON, Analysis and geometry on groups, Cambridge University Press, Cambridge (1993).
[Z] M. ZACK, A random walk on the Heisenberg group. Ph.D Thesis, Dept. Math., University of California, La Jolla (1989).

Persi Diaconis
Department of Mathematics
Harvard University
Cambridge, MA 02138
USA

Laurent Saloff-Coste
CNRS, Université Paul Sabatier
Laboratoire de Statistique et Probabilités
118 route de Narbonne
31062 Toulouse, Cedex
France

Submitted: July 1993
Final Version: October 1993