Electronic Journal of Probability

Vol. 4 (1999) Paper no. 1, pages 1–11.

Journal URL
http://www.math.washington.edu/~ejpecp/
Paper URL
http://www.math.washington.edu/~ejpecp/EjpVol4/paper1.abs.html

RANDOM WALKS ON FINITE GROUPS WITH FEW RANDOM GENERATORS

IGOR PAK

Department of Mathematics
Yale University
New Haven, CT 06520
paki@math.yale.edu

ABSTRACT. Let G be a finite group. Choose a set S of size k uniformly from G and consider a lazy random walk on the corresponding Cayley graph. We show that for almost all choices of S given $k = 2\,a \log_2 |G|$, $a > 1$, this walk mixes in under $m = 2\,a \log \frac{a}{a-1} \log |G|$ steps. A similar result was obtained earlier by Alon and Roichman (see [AR]), Dou and Hildebrand (see [DH]) using a different techniques. We also prove that when sets are of size $k = \log_2 |G| + O(\log \log |G|)$, $m = O(\log^3 |G|)$ steps suffice for mixing of the corresponding symmetric lazy random walk. Finally, when G is abelian we obtain better bounds in both cases.

A.M.S. Classification. 60C05, 60J15.

Key words and phrases. Random random walks on groups, random subproducts, probabilistic method, separation distance.

Submitted to EJP on June 5, 1998.
Final version accepted on November 11, 1998.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Introduction

In the past few years there has been a significant progress in analysis of random walks on groups with random support. Still for general groups G and small sets of generators, such as of size $O(\log |G|)$, more progress is yet to be made. Our results partially fill this gap.

Here is a general setup of a problem. Let G be a finite group, $n = |G|$. For a given k choose uniformly k random elements $g_1, \dots, g_k \in G$. Denote by S the set of these elements. A *lazy random walk* $\mathcal{W} = \mathcal{W}(G,S)$ is defined as a finite Markov chain X_t with state space G, and such that $X_0 = e$,

$$X_{t+1} = X_t \cdot g_i^{\epsilon_i}$$

where $g_i = g_i(t)$ are independent and uniform in $[k] = \{1, \dots, k\}$; ϵ_i are independent and uniform in $\{0,1\}$. By Q^m denote the probability distribution of X_m. If S is a set of generators, then $Q^m(g) \to 1/|G|$, i.e. the walk $\mathcal{W}$ has a uniform stationary distribution U, $U(g) = 1/n$ for all $g \in G$.

Define the *total variation distance* $\mathbf{d}(m)$ of the walk after m steps as follows:

$$\mathbf{d}(m) = \max_{A \subset G} |Q^m(A) - U(A)| = \frac{1}{2} \sum_{g \in G} \left| Q^m(g) - \frac{1}{|G|} \right|$$

Also define the *separation distance*

$$\mathbf{s}(m) = |G| \max_{g \in G} \left(\frac{1}{|G|} - Q^m(g) \right)$$

It is easy to see that $0 \le \mathbf{d}(m) \le \mathbf{s}(m) \le 1$. It is also known (see e.g. [AD2]) that $\mathbf{d}(m+1) \le \mathbf{d}(m)$, $\mathbf{s}(m+1) \le \mathbf{s}(m)$ for all $m > 0$, and $\mathbf{s}(2m) < C\,\mathbf{d}(m)$, a universal constant C and large enough m, such that $\mathbf{d}(m) < 1/16$.

The general problem is to find the smallest m such that $\mathbf{d}(m), \mathbf{s}(m) \le \varepsilon$ for almost all choices of S. Clearly, if m is small enough, then almost surely S is not a set of generators and $\mathbf{s}(m) = 1$, $\mathbf{d}(m) \ge 1/2$. The example of $G = \mathbb{Z}_2^r$ shows that if $k < r = \log_2 n$ this is the case. Thus it is reasonable to consider only the case $k \ge \log_2 n$.

Theorem 1. *Let G be a finite group, $n = |G|$. Let $\varepsilon > 0$, $a > 1$ be given. Then*

$$E[\mathbf{s}(m)] \to 0 \quad \text{as } n \to \infty\,,$$

where the expectation is taken over all choices of $S = \{g_1, \dots, g_k\}$ of size

$$k > 2\,a \log_2 n\ ,$$

and where $\mathbf{s}(m)$ is the separation distance of the lazy random walk $\mathcal{W}(G,S)$ after m steps, for

$$m > 2\,(1+\varepsilon)\,a \ln \frac{a}{a-1} \log_2 n$$

For example, when $a = 2$, $\varepsilon \to 0$, we have $m \approx 2.77 \log_2 n$ steps of the lazy walk is enough to drive the expected separation distance to 0, where the set of generators has size $k > 4 \log_2 n$ and is chosen uniformly in G.

Our second result deals with the case when $k = \log_2 n + o(\log n)$. While we cannot show that $m = O(\log n)$ steps is enough (later we show that this is *not* true in general), we prove that $m = O(\log^3 n)$ suffices. For a technical reason, we need to use a *symmetric lazy random walk* $\mathcal{W}_\circ(G, S)$ defined as follows :

$$X_{t+1} = X_t \cdot g_i^{\epsilon_i}$$

where ϵ_i are independent and uniform in $\{\pm 1, 0\}$, and g_i are independent and uniform in S.

Theorem 2. *Let G be a finite group, $n = |G|$. Let $\varepsilon > 0$ be given. Then*

$$E[\mathbf{s}(m)] \to 0 \quad \textit{as } n \to \infty\,,$$

where the expectation is taken over all choices of $S = \{g_1, \ldots, g_k\}$ of size

$$k = \lceil \log_2 n + (1 + \varepsilon) \log_2 \log_2 n \rceil$$

and where $\mathbf{s}(m)$ is the separation distance of the symmetric lazy random walk $\mathcal{W}_\circ(G, S)$ after m steps, for

$$m > (1 + \varepsilon)\, 3 \ln 2\, (\log_2 n)^3$$

Both results are obtained as an application of the Erdős-Rényi results on random subproducts (see [ER]).

A brief history of the problem. In [AD1] Aldous and Diaconis formulated the following informal conjecture for the usual (not *lazy*) random walks :

If both k and $\log_k n$ are large, then the total variation distance $\mathbf{d}(m)$ is small with high probability for $m > (1 + \varepsilon) \log_k n$.

In a superlogarithmic case the conjecture was modified and proved by Dou and Hildebrand in [DH]. They showed that $E[\mathbf{d}(m)] \to 0$ as $n \to \infty$ if $k > (\log n)^a$, $a > 1$, and $m > \frac{a}{a-1} \log_k n(1 + \varepsilon)$. They also showed that the factor $\frac{a}{a-1}$ cannot be lowered for certain classes of groups. A different proof was later found by Roichman (see [R].)

The case $k = O(\log n)$ for general groups was first explored by Alon and Roichman in [AR], where authors showed that the second largest eigenvalue λ_2 of the Cayley graph $\Gamma(G, S)$ is bounded by a constant. This immediately implies $m = O(\log n)$ steps is enough for mixing. Formally, they showed that given $1 > \delta > 1/e$, $k \geq (1 + o(1)) 2e^4 \ln 2/(\delta\, e - 1)$ then $E(\lambda_2) < \delta$. Although our results do not imply these, for the mixing time this gives bounds that are slightly worse than ours. We shall note that authors work with symmetric sets of generators.

Another approach was introduced by Dou and Hildebrand in [DH, §5]. They showed that if $k = a \log n$, $m > b \log n$, where $a > e^2$, $b < a/4$, and $b \log(eb/a) < -1$, then $E[d(m)] \to 0$ as $n \to \infty$. The result of Theorem 1 is a somewhat stronger version of a similar result. Particularly, we require just $a > 2$. Again, the direct comparison of results is cumbersome since authors use different measures of mixing (separation vs. total variation distance), different types of walks (1/2 vs. 0 holding probability), and in addition to that the latter result expresses $m = m(k, n)$ inexplicitly. Let us point out, however, that as $k/\log_2 n \to \infty$ the result in [DH] gives an asymptotically better bound on the expected number of steps $(m/\log_2 n \to 0$ vs. $m/\log_2 n \to 2)$. On the other hand our *probabilistic method* approach seems slightly more straightforward and easier to generalize.

The case $k = \log_2 n + o(n)$ studied in Theorem 2 is also not new in this setting. In [AR] authors remarked that one can easily get $m = O(\log^4 n)$ bound using just the diameter bound. We have all reasons to believe that the power 3 can be (and should be) brought down to at least 2. However the examples show that $m = \Omega(\log n \log\log n)$ in some cases (see below).

For specific groups, such as abelian groups, the situation is well understood. Several authors have obtained sharp bounds for these random random walks (see [G, H, PV, W]). A good guidance for the case $k = O(\log n)$ is again $\mathbb{Z}_2^r$. It is known that $m = O(r \log r)$ is necessary and almost always sufficient in case $k = r + Const$ (cf. Theorem 2, see [D1, PV, W]), while in case $k = Const \cdot r$ we have $m = O(r)$ is enough (see [W]). This shows that the bound in Theorem 1 is of the right order. We should also mention that in the case $G \simeq \mathbb{Z}_2^r$ the results of Wilson in [W] give extremely sharp estimates on convergence. The paper also suggests a possible generalization to all abelian groups, but the results have yet to be published.

An interesting observation is due to Hildebrand (see also [AR, §3]). In [H] he showed that if G is abelian, $k = (\log n)^a$, where $a < 1$, then for any given $\varepsilon > 0$, $b > 0$ and $m = (\log n)^b$ we have $d(m) > 1 - \varepsilon$ for sufficiently large n. Thus there is a phase transition around $a = 1$. Therefore our results can be interpreted as a look inside this phase transition.

Although the example above cover only abelian groups, the reader should be warned that the abelian groups might give an incomplete picture. For example, besides $\mathbb{Z}_2^r$ there are *nonabelian* groups with the property that they cannot be generated by less than $\log_2 n$ generators (cf. [AP]). Random walks on them are yet to be better understood. The following result, obtained a bonus from the proof of Theorem 1, is another illustration of the *"abelian groups are easier"* principle.

Theorem 3. *Let G be a finite abelian group, $n = |G|$. Let $\varepsilon > 0$, $a > 1$ be given. Then*

$$E[\mathbf{s}(m)] \to 0 \quad \text{as } n \to \infty\,,$$

where the expectation is taken over all choices of $S = \{g_1, \dots, g_k\}$ of size

$$k > a \log_2 n\ ,$$

and where $\mathbf{s}(m)$ is the separation distance of the lazy random walk $\mathcal{W}(G, S)$ after m steps, for

$$m > (1 + \varepsilon)\, a \ln \frac{a}{a-1} \log_2 n$$

In other words, we can save a factor of 2 in Theorem 1. This makes the result tight when $G = \mathbb{Z}_2^n$ (cf. [W]). Also, we get an analog of Theorem 2 which gives much tighter bound in this case.

Theorem 4. *Let G be a finite abelian group, $n = |G|$. Let w be any functions of n such that $w \to \infty$ as $n \to \infty$. Then*

$$E[\mathbf{s}(m)] \to 0 \quad as\ n \to \infty\,,$$

where the expectation is taken over all choices of $S = \{g_1, \ldots, g_k\}$ of size

$$k > \log_2 n \cdot \left(1 + C\frac{\log\log\log n}{\log\log n}\right),$$

where C is a universal constant (independent of n), and $\mathbf{s}(m)$ is the separation distance of the lazy random walk $\mathcal{W}(G,S)$ after m steps, for

$$m > \log_2 n\,(\log\log_2 n + w)$$

For example, $w = \log\log\log n$ will work. Heuristicly, Theorem 2 corresponds to a nonexistent case $a = 1$ in Theorem 3. Roughly, let $a = 1 + 1/\log_2 n$. Then $k = \log_2 n + 1$, and $m > (1+\varepsilon)\log_2 n \log\log_2 n$, which is basically what Theorem 4 says.

Finally, let us mention a somewhat relevant conjecture of Babai. He conjectured in [B2] that there exist a universal constant c such that if G is simple, then the diameter Δ of any Cayley graph on G is at most $(\log n)^c$. Together with the standard bound $mix < C|S|\Delta^2 \log n$ on a mixing time (see [AF, DSC]), and given $k = |S| = O\big((\log n)^{c_1}\big)$ this gives us $m = C'(\log n)^{c_2}$ steps is *always* enough for convergence (assuming S is a set of generators.) On the other hand, it is known that $P = \mathbf{Pr}(\langle S\rangle = G) \to 1$ as $n \to \infty$, where P is the probability of a random S, $|S| = k \geq 2$ generating G. This is a result of Liebeck and Shalev (see [LS]), conjectured earlier by Kantor and Lubotzky (see [KL]). Therefore we conclude that Babai conjecture implies that for simple G and random S of constant size $k \geq 2$ the mixing time is almost surely polylogarithmic in n.

Note here that the above conjecture of Babai as well as its application to convergence is open even for $G = A_n$. The best known result is due to Babai and Hetyei (see [BH]) who found $\Delta \leq (\log n)^{\log n (1/2 + o(1))}$ bound for *almost all* pairs of even permutations.

1. Proof of Theorem 1

Let G be a finite group, $n = |G|$. Throughout the paper we will ignore a small difference between *random subsets* S and *random sequences* J of group elements. The reason is that the two concepts are virtually identical since probability of repetition of elements (having $g_i = g_j$, $1 \leq i < j \leq k$) when $k = O(\log n)$ is exponentially small. Thus in the future we will substitute uniform sets S of size k by the uniform sequences $J \in G^k$, which, of course, can have repeated elements.

Fix a sequence $J = (g_1, \ldots, g_k) \in G^k$. *Random subproducts* are defined as

$$g_1^{\epsilon_1} \cdot \ldots \cdot g_k^{\epsilon_k}$$

where $\epsilon_i \in \{0,1\}$ are given by independent unbiased coin flips. Denote by P_J the probability distribution of the random subproducts on G. Erdős and Rényi showed in [ER] that if $g_1, \dots, g_k$ are chosen uniformly and independently, then :

$$(*)\ \mathbf{Pr}\left(\max_{g\in G}\left|P_J(g) - \frac{1}{n}\right| \le \frac{\varepsilon}{n}\right) > 1-\delta \text{ for } k \ge 2\log_2 n + 2\log_2 1/\varepsilon + \log_2 1/\delta$$

Proofs of Theorems 1, 3 are based on $(*)$.

Let $m > 2\log_2|G|$, and let J be as above. Denote by Q_J the probability distribution Q_J^m of the lazy random walk $\mathcal{W}(G,S)$ after m steps, where $S = S(J)$ is a set of elements in J. Suppose we can show that with probability $> 1-\alpha/2$ we have $\mathbf{s}_J(m) = n\max_{g\in G}(1/n - Q_J^m(g)) \le \alpha/2$, where $\alpha \to 0$ as $n\to\infty$. This would imply the theorem. Indeed, we have

$$\begin{aligned} E[\mathbf{s}_J(m)] &\le \mathbf{Pr}\big(\mathbf{s}_J \le \alpha/2\big)\cdot \alpha/2 + \mathbf{Pr}\big(\mathbf{s}_J > \alpha/2\big)\cdot 1 \\ &< (1-\alpha/2)\alpha/2 + \alpha/2 < \alpha \to 0 \end{aligned}$$

By definition, Q_J is distributed as random subproducts

$$g_{i_1}^{\epsilon_1}\cdot \ldots \cdot g_{i_m}^{\epsilon_m}$$

where $i_1, \dots, i_m$ are uniform and independent in $[k] = \{1,\dots,k\}$.

Let $J = (g_1,\dots,g_k)$ be fixed. For a given $I = (i_1,\dots,i_m) \in [k]^m$, consider $J(I) = (g_{i_1},\dots,g_{i_m})$ and $R_I = P_{J(I)}$. By definition of a lazy random walk we have

$$Q_J = \frac{1}{k^m}\sum_{I\in[k]^m} R_I$$

We will show that for almost all choices of J and I, the probability distribution R_I is almost uniform.

Let $I = (i_1,\dots,i_m) \in [k]^m$ be a sequence. Define an $\mathbf{L}$-*subsequence* $I' = (i_{r_1},\dots,i_{r_l})$ to satisfy $1 \le r_1 < \dots < r_l \le m$, and for all j, $1\le j\le m$, there exist a unique t, $1 \le t \le m$, such that $r_t \le j$ and $i_{r_t} = i_j$. In other words, we read numbers in I, and whenever we find a new number, we add it to I'. For example, if $I = (2,7,5,1,2,3,2,5,6)$, then $I' = (2,7,5,1,3,6)$ is an $\mathbf{L}$-subsequence of length 6. Note that by definition $\mathbf{L}$-subsequence is always unique.

Lemma 1. *Let I, J be as above, $n = |G|$. Let I' be a $\mathbf{L}$-subsequence of I. Then for all $\alpha, \beta > 0$ we have $\max_{g\in G}|R_{I'}(g) - 1/n| \le \alpha/n$ with probability $1-\beta$ implies $\max_{g\in G}|R_I(g) - 1/n| \le \alpha/n$ with probability $1-\beta$.*

Lemma 2. *Let $\beta > 0$, $a > 1$, $k = a\,l$, and $m = (1+\beta)\,k\,\ln\frac{a}{(a-1)^2}$. Consider the probability $P(l)$ a sequence $I \in [k]^m$ contains an $\mathbf{L}$-subsequence I' of length l. Then $P(l) \to 1$ as $l\to\infty$.*

First we deduce Theorem 1 from the lemmas and then prove the lemmas.

Proof of Theorem 1. Let I' be a $\mathbf{L}$-subsequence of I of length $l > 2\log_2 n + 3\log_2 1/\delta$. Since numbers in I' are all different, for at least $(1-\delta)$ fraction of all $J = \{g_1, \ldots, g_k\}$, we have

$$\max_{g \in G} \left| R_{I'}(g) - \frac{1}{n} \right| \leq \frac{\delta}{n}$$

Indeed, this is a restatement of $(*)$ with $\varepsilon = \delta$.

Note here that we do not require the actual group elements g_{i_j}, $i_j \in I'$ be different. By coincidence they can be the same. But we do require that *numbers* in I' are all different, so that the corresponding group elements are independent.

Let $l = \lceil 2\log_2 n + 3\log_2 1/\delta \rceil$, $k > a\,l$, and $m > (1+\varepsilon)\,a\,l\,\ln\frac{a}{a-1}$. Denote by $P(l)$ the probability that a uniformly chosen $I \in [k]^m$ contains an $\mathbf{L}$-subsequence of length l. By Lemma 1, with probability $> P(I)(1-\delta)$ we have

$$\max_{g \in G} \left(\frac{1}{n} - R_I(g) \right) \leq \frac{\delta}{n}$$

where the the probability is taken over all $I \in [k]^m$ and all $J \in G^k$. Setting $\delta = \delta(\alpha, \varepsilon, n)$ small enough we immediately obtain $\mathbf{s}_J(m) \leq \alpha/2$ with probability $> (1-\alpha/2)$. where the the probability is taken over all $J \in G^k$. By observations above, this is exactly what we need to prove the theorem.

Now take $\delta = \alpha/4$, $\beta = \varepsilon/2$. By Lemma 2, and and since $l > \log_2 n$ we have $P(I) > 1-\alpha/4$ for n large enough. We conclude $P(I)(1-\delta) > (1-\alpha/4)^2 > 1-\alpha/2$. This finishes proof of Theorem 1. $\square$

2. Proof of Lemmas

Proof of Lemma 1. For any $x, y \in G$ denote by y^x the element $xyx^{-1} \in G$. Clearly, if y is uniform in G and independent of x, then y^x is also uniform in G.

Let Q be a distribution on a group G which depends on $J \in G^m$ and takes values in G. We call Q (α,β)-*good* if with probability $> (1-\beta)$ it satisfies inequality $\max_{g \in G} |Q(g) - 1/n| \leq \alpha/n$.

Consider the following random subproducts:

$$h = g_1^{\epsilon_1} \cdot \ldots \cdot g_r^{\epsilon_r} \cdot x \cdot g_{r+1}^{\epsilon_r+1} \cdot \ldots \cdot g_l^{\epsilon_l}$$

where x is fixed, while $g_1, \ldots, g_l$ are uniform and independent in G, and $\epsilon_1, \ldots, \epsilon_l$ are uniform and independent in $\{0,1\}$. We have

$$h = g_1^{\epsilon_1} \cdot \ldots \cdot g_r^{\epsilon_r} \cdot (g_{r+1}^x)^{\epsilon_{r+1}} \cdot \ldots \cdot (g_l^x)^{\epsilon_l} \cdot x$$

Thus $h \cdot x^{-1}$ is distributed as R_I, $I = (1, 2, \ldots, l)$. Therefore if R_I is (α,β)-good, then distribution of h is also (α,β)-good.

Similarly, let $x, y, \ldots$ be fixed group elements. Then random subproducts

$$h = g_1^{\epsilon_1} \cdot \ldots \cdot x \cdot g_r^{\epsilon_r} \cdot \ldots \cdot y \cdot g_l^{\epsilon_l} \cdot \ldots$$

are distributed as $R_I \cdot f(x, y, \ldots)$, $I = (1, \ldots, r, \ldots, l, \ldots)$. Indeed, pull the right-most fixed element all the way to the right, then pull the previous one, etc. We conclude that if R_I is (α,β)-good, then distribution of h is also (α,β)-good. Note that in the observation above we can relax a condition that the elements $x, y, \dots$ are fixed. Since we do not have to change their relative order, it is enough to require that they are independent of the elements g_i to the right of them.

Now let $I = (i_1, \dots, i_m) \in [k]^m$, and let I' be an **L**-subsequence of I. Define $Q(h)$ to be a distribution of random subproducts

$$h = g_{i_1}^{\epsilon_1} \cdot \ldots \cdot g_{i_m}^{\epsilon_m}$$

where all the powers ϵ_j are fixed except for those of $j \in I'$. We claim that if $R_{I'}$ is (α,β)-good, then $Q(h)$ is also (α,β)-good. Indeed, pull all the elements that are not in I' to the right. By definition of the **L**-subsequence, the elements in I' to the right of those that are not in I' must be different and thus independent of each other. Thus by the observation above $Q(h)$ is also (α,β)-good.

Now, the distribution R_I is defined as an average of the distributions $Q(h)$ over all of the 2^{m-l} choices of values ϵ_s of elements not in $I' = (i_{r_1}, \dots, i_{r_l})$. Observe that for fixed $g_1, \dots, g_k$ and different choices the ϵ_s, $s \neq r_j$ the distributions of subproducts h can be obtained by a shift from each other (i.e. by multiplication on a fixed group element). Therefore each of these distributions has the same separation distance. In other words, each of the J is either "good" altogether or "bad" altogether for all 2^{m-l} choices. Therefore after averaging we obtain an (α,β)-good distribution R_I. This finishes proof of the lemma. $\square$

Proof of Lemma 2. The problem is equivalent to the following question. What is the probability that in the usual coupon collector's problem with k coupons, after m trials we have at least l different coupons? Indeed, observe that if all m chosen coupons correspond to elements in a sequence $I \in [k]^m$, then *distinct* coupons correspond to **L**-subsequence I' of length l. Note that in our case $k = a\,l$ and $m = (1+\beta)\,k\,\log\frac{a}{a-1}$.

Let τ be the first time we collect l out of k possible coupons. Let us compute the expected time $E(\tau)$. By the usual argument (see [F]) we have

$$E(\tau) = \frac{k}{k} + \frac{k}{k-1} + \frac{k}{k-2} + \cdots + \frac{k}{k-l+1} = k\left(\ln k - \ln(k-l) + o(1)\right)$$

The $o(1)$ in the last equality comes from cancellation of the Euler-Mascheroni constant γ when $k, (k-r) \to \infty$, and the formula (see e.g. [WW], §12.1) :

$$\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n} = \ln n + \gamma + o(1)$$

When $k = a\,l$. We obtain

$$E(\tau) = a\,l\left(\log\frac{a}{a-1} + o(1)\right)$$

Let us compute $Var(\tau)$. We have

$$Var(\tau) = \sum_{i=0}^{l-1} \frac{k}{k-i}\left(\frac{k}{k-i} - 1\right) \le l\,\frac{k}{k-l}\left(\frac{k}{k-l} - 1\right) = l\,\frac{a}{(a-1)^2}$$

Now let $m = (1+\beta)\, E(\tau)$. The probability $P(l)$ that after m trials we collect l coupons is equal to $\mathbf{Pr}(\tau \le m)$. Use Chebyshev inequality:

$$\mathbf{Pr}(\tau > (1+\beta)\, E(\tau)) \le \mathbf{Pr}(|\tau - E(\tau)| > \beta\, E(\tau)) < \frac{Var(\tau)}{\beta^2\, (E(\tau))^2} \to 0 \ \text{ as } l \to \infty$$

This finishes the proof of the lemma. □

3. Proof of Theorem 2

The proof of Theorem 2 is based on a different result of Erdős and Rényi. In [ER] along with $(*)$ they proved that if a sequence $J = (g_1, \dots, g_k)$ is chosen uniformly and independently, then :

$$(**)\ \mathbf{Pr}\left(\min_{g \in G} P_J(g) > 0\right) > 1 - \delta \text{ for } k \ge \log_2 n + \log_2 \log_2 n + 2 \log_2 1/\delta + 5$$

Here P_J is a distribution of random subproducts as in $(*)$.

We will use $(**)$ to prove that with probability $> 1 - \alpha$ taken over all choices of $J \in G^k$ we have

$$\mathbf{s}_J \le n \max_{g \in G} \left(Q_J^m(g) - \frac{1}{n} \right) < \alpha$$

where $m > 3 \ln 2\, k^2$ and k as above. By the same reasoning as in the proof of Theorem 1, this implies Theorem 2.

Consider group elements $g_1, \dots, g_k$ such that every $g \in G$ is given by a subproduct

$$g = g_1^{\epsilon_1} \cdot \ldots \cdot g_k^{\epsilon_k}\,,$$

where $\epsilon_i \in \{0, 1\}$. Denote $p(k)$ the probability of this event given $g_1, \dots, g_k$ are chosen uniformly in G.

We will use the subproducts above as paths on a Cayley graph generated by $g_i^{\pm 1}$. Note here that every generator occurs in each of the subproducts at most $N = 1$ time. Also, the *diameter* Δ of the Cayley graph above is bounded by maximum length of subproducts:

$$\Delta \le k$$

Let us use the path arguments (see [DSC, AF]) to compare the reversible Markov chain (which corresponds to our symmetric random walk) and a trivial Markov chain which at each step sends a chain to a uniform group element. For the second largest eigenvalue λ we get :

$$\lambda \le 1 - 1/A$$

where

$$A = \max_i \frac{1}{Q^1(g_i)} \sum_{g \in G} \Delta\, N\, \frac{1}{|G|} = 3\, k\, \Delta\, N \le 3\, k^2$$

Therefore if $m = (1+\beta) A \ln n$, $n = |G|$ we get

$$\mathbf{s}(m) \le n \max_{g \in G} |Q^m - 1/n| \le n \lambda^m \le (1 - 1/A)^{(1+\beta) A \ln n} n = 1/n^\beta \to 0$$

as $n \to \infty$, where the second inequality can be found e.g. in [B1].

Now take $k = \lceil \log_2 n + (1+\varepsilon) \log_2 \log_2 n \rceil$. Since $\varepsilon \log_2 \log_2 n - 5 \to \infty$, we obtain that $p(k) \to 1$ as $n \to \infty$, where $p(k)$ is the probability of choosing $J \in G^k$ such that the left hand side of $(**)$ holds. But in this case $\mathbf{s}_J(m) \to 0$ as $n \to \infty$, and where $m = (1+\beta)\, 3\, k^2 \ln n$ as above. Now take $\beta = \varepsilon/2$ and express m in terms of $\log_2 n$. This finishes proof of the Theorem. $\square$

4. Proof of Theorems 3 and 4

Proof of Theorems 3, 4 is very much similar to the proof of Theorem 1, so we will just point out the differences.

When G is an abelian group, we can use a result of Erdős and Hall in [EH]. They showed that there exist a universal constant C such that for any $\varepsilon, \delta > 0$, $n \to \infty$ we have

$$(*') \ \mathbf{Pr}\left(\max_{g \in G} \left| P_J(g) - \frac{1}{n} \right| \le \frac{\varepsilon}{n}\right) > 1 - \delta \text{ for } k \ge \log_2 n \left(1 + C \frac{\log\log\log n}{\log\log n}\right)$$

Using the arguments from the proof of Theorem 1 we immediately obtain Theorem 3. $\square$

Similarly, to obtain Theorem 4 we need a proper analog of Lemma 2. Let us estimate how big the m we need to get l distinct numbers in I. This is just coupon collector's problem. If there is a total of k different coupons, $k \ge l$, let X be the time to get l different coupons. Then the the expected time $E = E(X)$ is given by

$$E = \frac{k}{k} + \frac{k}{k-1} + \cdots + \frac{k}{k-l+1} \le \frac{l}{l} + \frac{l}{l-1} + \cdots + \frac{l}{1} = l \log l + O(l)$$

Also, since the variance of the coupon collector's problem is given by $Var(X) = O(l^2)$ (see [F, §9.9]) the Chebyshev inequality gives

$$\mathbf{Pr}\big(X > l \log l + t\, l\big) < \left(\frac{C_1}{t - C_2}\right)^2$$

where C_1, C_2 are universal constants. Thus if

$$k \ge l = \left\lceil \log_2 n + C \frac{\log_2 n \log\log\log n}{\log\log n} \right\rceil$$

we have $\delta \to 0$. Now take $\delta = \varepsilon$ is as in the proof of Theorem 1, and let $m > l(\log l + w)$. Proceeding as in the proof of Theorem 1, we finish the proof of Theorem 4. $\square$

Acknowledgements

We would like to thank Gene Cooperman, Martin Hildebrand, László Lovász and Van Vu for helpful remarks. Special thanks to the anonymous referee for pointing out an error in the original proof and suggesting a useful shortcut.

The author was supported in part by the NSF Postdoctoral Research Fellowship.

References

[AD1] D. Aldous, P. Diaconis, *Shuffling cards and stopping times*, Amer. Math. Monthly **93** (1986), 333–348.

[AD2] D. Aldous, P. Diaconis, *Strong uniform times and finite random walks*, Advances in Applied Math. **8** (1987), 69–97.

[AF] D. Aldous, J. Fill, *Reversible Markov Chains and Random Walks on Graphs*, monograph in preparation, 1996.

[AR] N. Alon, Y. Roichman, *Random Cayley graphs and expanders*, Random structures and algorithms **5** (1994), 271–284.

[AP] A. Astashkevich, I. Pak, *Random walks on nilpotent and supersolvable groups*, preprint (1997).

[B1] L. Babai, *Local expansion of vertex-transitive graphs and random geneartion in finite groups*, in Proc 23^{rd} ACM STOC (1991), 164–174.

[B2] L. Babai, *Automorphism groups, isomorphism, reconstruction*, in Handbook of Combinatorics (R. L. Graham, M. Groetschel, and L. Lovasz, eds.) (1996), Elsevier.

[BH] L. Babai, G. Hetyei, *On the Diameter of Random Cayley Graphs of the Symmetric Group*, Combinatorics, Probability & Computing **1** (1992), 201–208.

[D1] P. Diaconis, *Group Representations in Probability and Statistics*, IMS, Hayward, California, 1988.

[D2] P. Diaconis, *The cutoff phenomenon in finite Markov chains*, Proc. Nat. Acad. Sci. U.S.A. **93** (1996), 1659–1664.

[DGM] P. Diaconis, R. Graham, J. Morrison, *Asymptotic analysis of a random walk on a hypercube with many dimensions*, Random structures and algorithms **1** (1990), 51–72.

[DSC] P. Diaconis, L. Saloff–Coste, *Comparison techniques for random walk on finite groups*, Ann. Prob. **21** (1993), 2131–2156.

[DH] C. Dou, M. Hildebrand, *Enumeration and random random walks on finite groups*, Ann. Prob. **24** (1996), 987–1000.

[EH] P. Erdős, R.R. Hall, *Probabilistic methods in group theory. II*, Houston J. Math. **2** (1976), 173–180.

[ER] P. Erdős, A. Rényi, *Probabilistic methods in group theory*, Jour. Analyse Mathématique **14** (1965), 127–138.

[F] W. Feller, *An introduction to Probability theory and its applications, Vol. 1* (third edition), John Wiley, New York, 1968.

[G] A. Greenhalgh, *A model for random random-walks on finite groups*, Combin. Probab. Comput. **6** (1997), 49–56.

[H] M. Hildebrand, *Random walks supported on random points of* $\mathbb{Z}/n\mathbb{Z}$, Probability Theory & Related Fields **100** (1994), 191–203.

[KL] W. M. Kantor, A. Lubotzky, *The probability of generating a finite classical group*, Geom. Dedicata **36** (1990), 67–87.

[LS] M. W. Liebeck, A. Shalev, *The probability of generating a finite simple group*, Geom. Dedicata **56** (1995), 103–113.

[P] I. Pak, *Random walks on groups: strong uniform time approach*, Ph.D. Thesis, Harvard U., 1997.

[PV] I. Pak, V. H. Vu, *On finite geometric random walks*, preprint (1998).

[R] Y. Roichman, *On random random walks*, Ann. Prob. **24** (1996), 1001–1011.

[WW] E. T. Whittaker, G. N. Watson, *A Course of Modern Analysis* (Fourth Edition), Cambridge University Press, Cambridge, UK, 1927.

[W] D. Wilson, *Random random walks on* $\mathbb{Z}_2^d$, Probability Theory & Related Fields **108** (1997), 441–457.