The Annals of Probability
1996, Vol. 24, No. 2, 987–1000

ENUMERATION AND RANDOM RANDOM WALKS ON FINITE GROUPS

BY CARL DOU AND MARTIN HILDEBRAND

J. P. Morgan & Co. and University of Texas at Austin

This paper examines random walks on a finite group G and finds upper bounds on how long it takes typical random walks supported on $(\log|G|)^a$ elements to get close to uniformly distributed on G. For certain groups, a cutoff phenomenon is shown to exist for these typical random walks. A variation of the upper bound lemma of Diaconis and Shahshahani and some counting arguments related to a group equation are used to get the upper bound. A further example which uses this variation is discussed.

1. Introduction. Random walks on finite groups have received considerable study recently. For an overview of such walks, see Diaconis [5]. One question which arises is how long does it take for such walks to become close to uniformly distributed on the finite group. One technique used for studying such walks involves studying a family of such walks; such a family can be formed by looking at all walks where the number of elements obtainable in one step from the identity is a given function of the order of the group. Sometimes bounds on the average distance of how far the random walk is at a given time can be found. Such techniques have been used by Greenhalgh [9], Hildebrand [12] and Wilson [18] to obtain results on specific groups.

In this paper, we shall use these techniques to obtain results which are valid on arbitrary groups. The only information which we use and which varies between groups is the order of the group.

Let G be an arbitrary finite group of order g with identity element labeled 1. Define a probability measure Q on G. Let $Z_1, \ldots, Z_m$ be i.i.d. random variables on G with distribution Q and let $X_0 = 1$, $X_n = Z_n X_{n-1}$, if $n \geq 1$. Let Q^{*m} be the distribution of X_m. (Note that Q^{*m} has the same meaning as in Diaconis [5].)

Let P be a probability distribution on G and let U be the uniform distribution on G. We shall define the variation distance between P and U by

$$\begin{aligned}\|P - U\| &:= \tfrac{1}{2} \sum_{s \in G} |P(s) - (1/g)| \\ &= \max_{A \subseteq G} |P(A) - U(A)|.\end{aligned}$$

We shall show the following theorem.

Received June 1994; revised February 1995.
AMS 1991 *subject classifications*. Primary 60B15; secondary 60J15, 05A18.
Key words and phrases. Random walk, finite groups, enumeration, upper bound lemma.

THEOREM 1. *Let* $k = \lfloor (\log g)^a \rfloor$, *where* $a > 1$ *is constant. Let* $\varepsilon > 0$ *be given. Suppose* S *is a random* k*-subset of* G (*chosen uniformly from all subsets of* G *with* k *elements*) *and let*

$$Q(s) := \begin{cases} 1/k, & \text{if } s \in S, \\ 0, & \text{otherwise.} \end{cases}$$

Suppose

$$m > \frac{a}{a-1} \frac{\log g}{\log k} (1 + \varepsilon).$$

Then $E[\|Q^{*m} - U\|] \to 0$ *as* $g \to \infty$.

In other words, for a typical random walk which is supported on k points, after m steps the walk's position will be close to uniformly distributed on G.

Theorem 1 is a modification of the following informal conjecture of Aldous and Diaconis [2].

CONJECTURE. *Let* G *be an arbitrary finite group of order* g *and let* Q *be a probability measure on* G. *Suppose* Q *is a random* k*-subset. If both* k *and* $\log g / \log k$ *are large, then if* $m > (\log g / \log k)(1 + \varepsilon)$, *with high probability* $\|Q^{*m} - U\| \approx 0$.

Note that if $m < (\log g / \log k)(1 - \varepsilon)$, then on the mth step of the random walk, there are no more than $g^{1-\varepsilon}$ elements reached. Thus $\|Q^{*m} - U\| \geq 1 - (g^{1-\varepsilon}/g) \to 1$ as $g \to \infty$.

This conjecture needs to be modified for two reasons. First, k must grow rather substantially, namely, $k \geq \log g / \log 2$. Otherwise if $G = \mathbf{Z}_2^d$, then $k < d$ and a random walk supported on k elements will be confined to at most half of G. Furthermore, even in the case when $k = \lfloor (\log g)^a \rfloor$, where $a > 1$ is constant, more steps are needed on all Abelian and certain non-Abelian groups. Hildebrand [12] showed this fact on Abelian groups; this fact will be proved here for certain non-Abelian groups.

The proof of Theorem 1 uses a modification of the upper bound lemma of Diaconis and Shahshahani, uses counting arguments to get bounds on the number of solutions of a group equation and then uses some bounds on Stirling numbers of the second kind. Similar techniques can be used in proving results when k is larger. We shall describe what happens when $G = \mathbf{Z}_n$, $k = n^{1/2+\varepsilon}$ and $m = 2$.

Questions related to Theorem 1 involve random walks of Cayley graphs of such groups and the question whether such graphs are good expanders. See, for example, [3], [4], [8], [15] and [16]. Random walks on Cayley graphs are essentially random walks on finite groups, where the set S on which the random walk is based contains the inverses of the elements of S. By a modification of the argument in [4], one can show a result analogous to Theorem 1. (See the article following the current article for details.) In Theorem 1, we do not assume that the set S contains the inverses of the elements of S.

2. Upper bound lemma. The upper bound lemma of Diaconis and Shahshahani uses irreducible representations of finite groups. For a description of the representation theory of finite groups, see Serre [17] or Chapter 2 of Diaconis [5]. This lemma is the following.

LEMMA 1. *Let Q be a probability on a finite group G and let U be the uniform distribution. Then*

$$\|Q - U\|^2 \le \tfrac{1}{4} \sum_{\rho}^{*} d_\rho \operatorname{Tr}\big(\hat{Q}(\rho)\hat{Q}(\rho)^*\big),$$

where $*$ of a matrix denotes its conjugate transpose and $\sum_\rho^*$ denotes the sum over all (nonequivalent) nontrivial irreducible representations ρ of G.

This lemma is proved in Diaconis [5].

This lemma is very useful in cases where the probability is constant on conjugacy classes of a non-Abelian group; see, for example, Hildebrand [10] or Chapter 3D of Diaconis [5]. This lemma is also useful in certain random processes with a recurrence relation; see Hildebrand [11], for example. While we have neither property here, we still can adapt this upper bound lemma to a useful form:

LEMMA 2. *Let Q be a probabiltiy on G. Then for any positive integer m,*

$$4\|Q^{*m} - U\|^2 \le \sum_{\Omega} gQ(x_1)\cdots Q(x_{2m}) - \sum_{G^{2m}} Q(x_1)\cdots Q(x_{2m}),$$

where G^{2m} is the set of all $2m$-tuples $(x_1, \ldots, x_{2m})$ with $x_i \in G$ and Ω is a subset of G^{2m} consisting of all $2m$-tuples such that $x_1 x_2 \cdots x_m = x_{m+1} x_{m+2} \cdots x_{2m}$.

PROOF. Let $\rho_1, \ldots, \rho_h$ be all the nonequivalent irreducible representations of G with characters $\chi_1, \ldots, \chi_h$ and degrees $d_1, \ldots, d_h$ correspondingly. We may assume that the representations are all unitary. We also may assume that ρ_h is the trivial representation. Hence $d_h = 1$ and $\chi_h(s) = 1$ for all $s \in G$. Note that $\hat{Q}(\rho_i) = \sum_{x \in G} Q(x)\rho_i(x)$. Since ρ_i is a unitary representation, we have $\rho_i(x)^* = (\rho_i(x))^{-1} = \rho_i(x^{-1})$ for all $x \in G$. Hence

$$\hat{Q}(\rho_i)^m = \sum_{x_1, \ldots, x_m} Q(x_1)\cdots Q(x_m)\rho_i(x_1 \cdots x_m),$$

$$\big(\hat{Q}(\rho_i)^m\big)^* = \sum_{x_{m+1}, \ldots, x_{2m}} Q(x_{m+1})\cdots Q(x_{2m})\rho_i\big((x_{m+1}\cdots x_{2m})^{-1}\big).$$

Thus

$$d_i \operatorname{Tr}\big(\hat{Q}(\rho_i)^m\big(\hat{Q}(\rho_i)^m\big)^*\big) = \sum_{G^{2m}} Q(x_1)\cdots Q(x_{2m}) d_i \chi_i(s),$$

where $s = x_1 \cdots x_m (x_{m+1} \cdots x_{2m})^{-1}$. Thus four times the right side of Lemma 1 is

$$\sum_{i=1}^{h-1} d_i \operatorname{Tr}\left(\hat{Q}(\rho_i)^m \left(\hat{Q}(\rho_i)^m\right)^*\right) = \sum_{G^{2m}} Q(x_1) \cdots Q(x_{2m}) \sum_{i=1}^{h-1} d_i \chi_i(s).$$

Note that $d_h \chi_h(s) = 1$, for all $s \in G$, whereas

$$\sum_{i=1}^{h} d_i \chi_i(s) = \begin{cases} g, & \text{if } s = 1, \\ 0, & \text{otherwise.} \end{cases}$$

Thus we get

$$\sum_{i=1}^{h-1} d_i \operatorname{Tr}\left(\hat{Q}(\rho_i)^m \left(\hat{Q}(\rho_i)^m\right)^*\right) = \sum_{\Omega} Q(x_1) \cdots Q(x_{2m}) g - \sum_{G^{2m}} Q(x_1) \cdots Q(x_{2m})$$

and our proof is complete. □

An alternate proof of Lemma 2 has been found by Diaconis [6].

In addition to the proofs presented in this paper, Lemma 2 is useful in proving some upper bounds involving random walks supported on a random subset of certain Abelian groups, where the size of the support does not depend on the size of these groups. This use appears in Dou [7].

3. Counting related to the group equation. In this section, we investigate solutions to the group equation

$$x_1 \cdots x_m = x_{m+1} \cdots x_{2m}, \tag{3.1}$$

which was used in defining Ω. Obviously, $|\Omega| = g^{2m-1}$. We shall make use of the size of different subsets of Ω. These subsets will consist of the number of solutions to (3.1) such that $\{x_1, \ldots, x_{2m}\}$ consists of i distinct elements for $i = 1, \ldots, 2m$.

We shall use the following definitions.

DEFINITION. A $2m$-tuple $\nu = (x_1, \ldots, x_{2m}) \in G^{2m}$ is said to be of size i if the cardinality of $X = \{x_1, \ldots, x_{2m}\}$ is i.

DEFINITION. An i-partition of the set $\{1, 2, \ldots, 2m\}$ is a set of i disjoint subsets $\tau = \{\Delta_1, \ldots, \Delta_i\}$ such that their union is the whole set.

DEFINITION. An i-partition of the number $2m$ is an i-tuple of integers $\pi = (p_1, \ldots, p_i)$ such that

$$p_1 \geq p_2 \geq \cdots \geq p_i \geq 1 \quad \text{and} \quad \sum_{j=1}^{i} p_j = 2m.$$

Notice, first of all, that an i-partition of the set corresponds to an i-partition of the number $2m$, namely, the i-tuple of the decreasingly ordered cardinalities of the subsets in the partition of the set. Second, each $2m$-tuple in G^{2m} of size i gives rise to an i-partition of $2m$ in a natural way: For $1 \le j \le i$, let $\Delta_j \subset \{1, 2, \dots, 2m\}$ be a maximal subset of indices for which the corresponding coordinates are the same. Then the set of those Δ_j's is an i-partition of $\{1, 2, \dots, 2m\}$, and this i-partition is called the type of the $2m$-tuple. Suppose $|\Delta_1| \ge \cdots \ge |\Delta_i|$. Then the corresponding $\pi = (|\Delta_1|, \dots, |\Delta_i|)$ is an i-partition of $2m$.

EXAMPLE. Let $\nu = (0, 1, 5, 2, 2, 7, 5, 5) \in \mathbf{Z}_{10}^8$, where $\mathbf{Z}_{10}$ is all integers modulo 10. Its type is $\tau = \{\{3, 7, 8\}, \{4, 5\}, \{1\}, \{2\}, \{6\}\}$ and the corresponding 5-partition of the number 8 is $\pi = (3, 2, 1, 1, 1)$.

Now Ω can be classified and therefore counted according to the types of the $2m$-tuples. Let $\tau = \{\Delta_1, \dots, \Delta_i\}$ be a type of an i-partition π of $2m$. Write $N_\pi(\tau)$ as the number of $2m$-tuples in Ω of type τ. (The notation may seem redundant since π is uniquely determined by τ. However, this notation will be helpful in a triple sum to appear later.) A moment's thought gives the following observation:

LEMMA 3. *$N_\pi(\tau)$ is the number of i-tuples $(y_1, \dots, y_i)$ of distinct coordinates that are solutions to the induced equation obtained from* (3.1) *by substituting y_j for x_l if $l \in \Delta_j$.*

The following example should clarify Lemma 3.

EXAMPLE. Take $\tau = \{\{1, 2, 7\}, \{3, 4\}, \{5, 8\}, \{6\}, \{9\}, \{10\}\}$, $\pi = (3, 2, 2, 1, 1, 1)$ and $m = 5$. Then $N_\pi(\tau)$ is the number of 6-tuples $(y_1, y_2, y_3, y_4, y_5, y_6)$ with distinct coordinates satisfying the equation

$$y_1^2 y_2^2 y_3 = y_4 y_1 y_3 y_5 y_6.$$

The following theorem provides motivation for the above notation.

THEOREM 2. *Let $G = \mathbf{Z}_n$. Let S be a random k-subset, where k is an integer which may depend on n. Let $Q(s) = 1/k$ if $s \in G$. Then*

$$E\left[\|Q^{*2} - U\|\right] \le \frac{\sqrt{3}}{2}\left(\frac{n}{k^2}\right)^{1/2}.$$

Note that if $k = n^{1/2+\varepsilon}$ with $\varepsilon > 0$, Theorem 2 implies typical random walks on $\mathbf{Z}_n$ supported on k points take two steps to get close to uniformly distributed.

PROOF. By taking expectations of both sides of Lemma 2, we have

$$4E\left[\|Q^{*2} - U\|^2\right] \le \sum_{x_1 + x_2 = x_3 + x_4} nEQ(x) - \sum_{x \in \mathbf{Z}_n^4} EQ(x),$$

where $x = (x_1, x_2, x_3, x_4)$ and $Q(x) = Q(x_1)Q(x_2)Q(x_3)Q(x_4)$.

Let $X = \{x_1, x_2, x_3, x_4\}$ and $i = |X|$. It can be shown that $EQ(x)$ depends only on i and that

$$EQ(x) = \sum_{\substack{X \subset T \\ |T| = k}} \frac{1}{k^4} \binom{n}{k}^{-1} = \frac{1}{k^4} \binom{n-i}{k-i} \binom{n}{k}^{-1}.$$

Thus

$$4E\left[\|Q^{*2} - U\|^2\right] \leq \frac{1}{k^4} \binom{n}{k}^{-1} \sum_{i=1}^{4} \binom{n-i}{k-i} \left(nN_4^i - M_4^i\right),$$

where M_4^i is the number of solutions of $x_1 + x_2 = x_3 + x_4$ with $|X| = i$ and M_4^i is the number of 4-tuples with $|X| = i$.

It can easily be shown that $M_4^4 = n(n-1)(n-2)(n-3)$, $M_4^3 = 6n(n-1)(n-2)$, $M_4^2 = 7n(n-1)$ and $M_4^1 = n$.

To find N_4^i, we need to examine the individual types.

If $i = 1$, there is one i-partition of 4 and one type $\tau = \{\{1,2,3,4\}\}$. The induced equation is $y_1 + y_1 = y_1 + y_1$. This holds regardless of the value for y_1. So here $N_\pi(\tau) = n$.

If $i = 2$, there are two partitions of the number 4. If $\pi = (3,1)$, there are four types. For instance, τ may be $\{\{1,2,3\},\{4\}\}$. The induced equation is $y_1 + y_1 = y_1 + y_2$, and hence $y_1 = y_2$. We assume $y_1 \neq y_2$, and so there are no solutions to (3.1) here. Hence $N_\pi(\tau) = 0$. The other types for this partition of 4 are similar. The other partition of 4 is $\pi = (2,2)$. This partition has three types. If $\tau = \{\{1,2\},\{3,4\}\}$, then the induced equation is $y_1 + y_1 = y_2 + y_2$ with $y_1 \neq y_2$. If n is odd, there are no solutions, but if n is even, there are n solutions. For each value y_1, let $y_2 = y_1 + n/2 \pmod{n}$. Let $\beta = N_\pi(\tau)$ for this type τ. If $\tau = \{\{1,3\},\{2,4\}\}$, the induced equation is $y_1 + y_2 = y_1 + y_2$ and there are $n(n-1)$ solutions here. There are also $n(n-1)$ solutions to the equation induced by the other type.

Via similar reasoning, one can show that $N_4^3 = 2(n(n-1) - \beta)$ and $N_4^4 = n(n-1)(n-4) + (n-1)n + \beta$.

The theorem follows by elementary algebra, which we omit, and the Schwarz inequality. □

Although getting precise expressions for the $N_\pi(\tau)$'s can be very difficult, in general, we can find some useful information about their asymptotic behavior. This information is in the following lemma.

LEMMA 4. *Let π be an i-partition of $2m$ and τ a type of π. Let $N_\pi(\tau)$ be as before. Then the following inequalities hold*:

$$|gN_\pi(\tau) - [g]_i| \leq \begin{cases} g[g]_i, & \text{if } 1 \leq i \leq m, \\ \dfrac{(i-1)!}{(m-1)!} g[g]_m, & \text{if } m \leq i \leq 2m, \end{cases}$$

where $[g]_i := g(g-1)\cdots(g-i+1)$.

PROOF. The first case follows trivially from Lemma 3 and the fact that the number of i-tuples $(y_1, \ldots, y_i)$ with distinct coordinates is $[g]_i$.

We use induction to prove the second case. If $i = m$, the result is true by the first case. Now consider $i \geq m + 1$. Let $\tau = \{\Delta_1, \ldots, \Delta_i\}$. For at least one value $i_0 \leq i$, $|\Delta_{i_0}| = 1$ since $i > m$, $|\Delta_j| \geq 1$, for $j = 1, \ldots, i$, and $\sum_{j=1}^{i} |\Delta_j| = 2m$. Without loss of generality, assume $|\Delta_i| = 1$. By Lemma 3.1, consider the equation in $(y_1, \ldots, y_i)$ induced by τ. For each of the $[g]_{i-1}$ choices of the $(i-1)$-tuples $(y_1, \ldots, y_{i-1})$ with distinct coordinates, there exists a unique solution for y_i which satisfies the induced equation because y_i appears only once in the equation and all values y_i are invertible. (In the example where $y_1^2 y_2^2 y_3 = y_4 y_1 y_3 y_5 y_6$, we would get $y_6 = y_5^{-1} y_3^{-1} y_1^{-1} y_4^{-1} y_1^2 y_2^2 y_3$.) Of these $[g]_{i-1}$ possible candidates for solutions with distinct coordinates, some may have y_i being one of the values $y_1, \ldots, y_{i-1}$. So we need to count the number of these bad candidates and subtract this number to get $N_\pi(\tau)$. Let A_l be the set of solutions with $y_l = y_i$ (and with $y_1, \ldots, y_{i-1}$ distinct) for $l = 1, \ldots, i-1$. It is not hard to see that $|A_l| = N_{\pi_l}(\tau_l)$, where

$$\tau_l := \{\Delta_1, \ldots, \Delta_{l-1}, \Delta_l \cup \Delta_i, \Delta_{l+1}, \ldots, \Delta_{i-1}\}$$

is an $(i-1)$-partition of the set and π_l is the corresponding $(i-1)$-partition of $2m$. Since the sets A_l are pairwise disjoint, we may conclude

$$N_\pi(\tau) = [g]_{i-1} - \sum_{l=1}^{i-1} N_{\pi_l}(\tau_l).$$

Furthermore, the function $[g]_i$ satisfies the following recurrence:

$$[g]_i = [g]_{i-1}(g - i + 1) = g[g]_{i-1} - (i-1)[g]_{i-1}.$$

Combining the above recurrences, we get

$$\begin{aligned}
|gN_\pi(\tau) - [g]_i| &= \left| \sum_{l=1}^{i-1} [g]_{i-1} - gN_{\pi_l}(\tau_l) \right| \\
&\leq \sum_{l=1}^{i-1} |gN_{\pi_l}(\tau_l) - [g]_{i-1}| \\
&\leq \sum_{l=1}^{i-1} \frac{(i-2)!}{(m-1)!} g[g]_m \quad \text{(by the induction hypothesis)} \\
&= \frac{(i-1)!}{(m-1)!} g[g]_m.
\end{aligned}$$

This completes the proof. □

Let $M_\pi(\tau)$ be the number of $2m$-tuples of type τ in G^{2m}, where π is the corresponding i-partition of the number $2m$. It is easy to show that $M_\pi(\tau) = [g]_i$.

The following lemma shows where Lemma 4 is useful in finding expectations of variation distances.

LEMMA 5.

$$4E\big(\|Q^{*m} - U\|^2\big) \le \sum_{i=1}^{2m} \sum_{\pi \in P(i)} \frac{1}{k^{2m}} \frac{[k]_i}{[g]_i} \sum_{\tau \in T(\pi)} (gN_\pi(\tau) - [g]_i),$$

where $P(i)$ is the set of all i-partitions of $2m$ and $T(\pi)$ is the set of all types of π.

PROOF. Observe from Lemma 2 that

$$4E\big(\|Q^{*m} - U\|^2\big) \le \sum_{\Omega} gE\big(Q(x_1) \cdots Q(x_{2m})\big) - \sum_{G^{2m}} E\big(Q(x_1) \cdots Q(x_{2m})\big).$$

We shall evaluate the right side of the above equation very carefully. If π is an i-partition of $2m$, then a $2m$-tuple of π is a $2m$-tuple whose type corresponds to π. Let $D_1(\pi)$ be the set of all $2m$-tuples of π in Ω and let $D_2(\pi)$ be the set of all $2m$-tuples of π in G^{2m}. Let $T(\pi)$ be all types of π. Then

$$|D_1(\pi)| = \sum_{\tau \in T(\pi)} N_\pi(\tau), \qquad |D_2(\pi)| = \sum_{\tau \in T(\pi)} M_\pi(\tau)$$

and

$$4E\big(\|Q^{*m} - U\|^2\big) \le \sum_{i=1}^{2m} \sum_{\pi \in P(i)} \left(\sum_{x \in D_1(\pi)} gEQ(x) - \sum_{x \in D_2(\pi)} EQ(x) \right), \tag{3.2}$$

where $x := (x_1, \ldots, x_{2m})$ and $Q(x) := Q(x_1) \cdots Q(x_{2m})$.

We shall next evaluate $EQ(x)$. Its value only depends on the partition π associated with the $2m$-tuple.

The probability that a given i-tuple $(y_1, \ldots, y_i)$ with distinct elements is contained in a random k-subset of G is $[k]_i/[g]_i$. Thus if x corresponds to an i-partition of $2m$,

$$EQ(x) = \frac{1}{k^{2m}} \frac{[k]_i}{[g]_i}.$$

Thus we may conclude

$$\sum_{x \in D_1(\pi)} gEQ(x) = \frac{1}{k^{2m}} \frac{[k]_i}{[g]_i} \sum_{\tau \in T(\pi)} gN_\pi(\tau)$$

and

$$\begin{aligned} \sum_{x \in D_2(\pi)} EQ(x) &= \frac{1}{k^{2m}} \frac{[k]_i}{[g]_i} \sum_{\tau \in T(\pi)} M_\pi(\tau) \\ &= \frac{1}{k^{2m}} \frac{[k]_i}{[g]_i} \sum_{\tau \in T(\pi)} [g]_i. \end{aligned}$$

Thus the right side of (3.2) can be rewritten

$$\sum_{i=1}^{2m} \sum_{\pi \in P(i)} \frac{1}{k^{2m}} \frac{[k]_i}{[g]_i} \sum_{\tau \in T(\pi)} (gN_\pi(\tau) - [g]_i)$$

and the lemma is proven. □

The following lemma gives an upper bound which uses Stirling numbers of the second kind. Such numbers are described in combinatorics texts such as Aigner [1]. We shall denote such numbers $S_{2m,i}$, where $S_{2m,i}$ is the number of ways to place $2m$ labeled balls in i unlabeled boxes such that there are no empty boxes.

LEMMA 6. *If* $k < \sqrt{2g}$ *and* $m < k/4$, *then*

$$4E\left[\|Q^{*m} - U\|^2\right] \leq \frac{1}{k^{2m}} [k]_m g \left(\sum_{i=1}^{m} \frac{[k]_i}{[k]_m} S_{2m,i} + \sum_{i=m+1}^{2m} S_{2m,i} \right).$$

PROOF. Use Lemma 5. Note that by Lemma 4, if $1 \leq i \leq m$, then

$$\frac{[k]_i}{[g]_i} |gN_\pi(\tau) - [g]_i| \leq \frac{[k]_i}{[g]_i} g[g]_i = [k]_m \frac{[k]_i}{[k]_m} g.$$

If $m + 1 \leq i \leq 2m$, then

$$\begin{aligned} \frac{[k]_i}{[g]_i} |gN_\pi(\tau) - [g]_i| &\leq \frac{(i-1)!}{(m-1)!} \frac{[k]_i}{[g]_i} g[g]_m \\ &\leq g[k]_m, \end{aligned}$$

since if $k \leq \sqrt{2g}$ and $m < k/4$,

$$\frac{(i-1)!}{(m-1)!} \frac{[k]_i}{[g]_i} \leq \frac{[k]_m}{[g]_m}.$$

Observe that

$$\sum_{\pi \in P(i)} \sum_{\tau \in T(\pi)} 1 = S_{2m,i}.$$

Putting these results together completes the proof of this lemma. □

4. Proof of Theorem 1. First off, note that $(\log g)^\alpha < \sqrt{2g}$ for sufficiently large values of g and $m < k/4$. Thus Lemma 6 may be used.

Observe that $S_{2m,i} \leq i^{2m}/i!$ since there are no more than i^{2m} ways to place $2m$ labeled balls in i *labeled* boxes with no empty boxes.

Thus

$$\sum_{i=1}^{m} \frac{[k]_i}{[k]_m} S_{2m,i} \le \sum_{i=1}^{m} \frac{f(m,i)}{k^{m-i}} \frac{i^{2m}}{i!}$$
$$\le \sum_{i=1}^{m} i^m \frac{i^m}{k^m} \frac{k^i}{\sqrt{2\pi i}\, e^{-i} i^i} \frac{f(m,i)}{g(i)}$$
$$\le \sum_{i=1}^{m} i^m \left(\frac{i}{k}\right)^{m-i} e^i \frac{f(m,i)}{\sqrt{2\pi i}\, g(i)},$$

where $g(i) \to 1$ as $i \to \infty$, $f(m,i) < 2^{m-i}$ since $m \ll k$ and $\sqrt{2\pi i}\, g(i) \ge 1$, for all $i \ge 1$. Thus for large enough m,

$$\sum_{i=1}^{m} \frac{[k]_i}{[k]_m} S_{2m,i} \le \sum_{i=1}^{m} i^m \left(\frac{2i}{k}\right)^{m-i} e^i \le (em)^m 2$$

since $i^m \le m^m$, $e^i \le e^m$ and $\sum_{i=1}^{m} (2i/k)^{m-i} \le \sum_{j=0}^{\infty} (1/2)^j = 2$.

Observe that if $i > m$,

$$\frac{(i+1)^{2m}/(i+1)!}{i^{2m}/i!} = \frac{((i+1)/i)^{2m}}{i+1}$$
$$= \frac{(1+1/i)^{2m}}{i+1}$$
$$< \frac{e^{2m/i}}{i+1} < \frac{e^2}{i+1} < 0.5,$$

if $m > 2e^2$.

Thus for sufficiently large m,

$$\sum_{i=m+1}^{2m} S_{2m,i} \le \sum_{i=m+1}^{2m} \frac{i^{2m}}{i!} \le \sum_{i=1}^{m} (0.5)^i \frac{m^{2m}}{m!} \le (em)^m$$

and

$$\left(\sum_{i=1}^{m} \frac{[k]_i}{[k]_m} S_{2m,i} + \sum_{i=m+1}^{2m} S_{2m,i} \right) \le 3(em)^m.$$

Thus

$$4E\left(\|Q^{*m} - U\|^2\right) \le \frac{1}{k^{2m}} k^m g 3(em)^m$$
$$= 3g\left(\frac{em}{k}\right)^m.$$

Suppose $m = (a/(a-1))(\log g/\log k)(1+\varepsilon)$. (In this argument, we shall omit explicit reference to the floor notation for k and m since such omission will not affect the conclusion.) Then $e^m = g^{o(1)}$ and $k^m = g^{(a/(a-1))(1+\varepsilon)}$. Observe that

$$m^m = (\log g)^m((a/(a-1))(1/\log k)(1+\varepsilon))^m.$$

Since $k = (\log g)^a$,

$$((a/(a-1))(1/\log k)(1+\varepsilon))^m = g^{o(1)},$$

whereas

$$(\log g)^m = \exp\left(\log\log g \frac{a}{a-1}\frac{\log g}{a\log\log g}(1+\varepsilon)\right) = g^{(1/(a-1))(1+\varepsilon)}.$$

Thus

$$\begin{aligned} E(\|Q^{*m} - U\|^2) &\le \frac{3}{4}\frac{gg^{(1/(a-1))(1+\varepsilon)}g^{o(1)}}{g^{(a/(a-1))(1+\varepsilon)}} \\ &= \frac{3}{4}\frac{1}{g^{\varepsilon - o(1)}} \to 0 \end{aligned}$$

as $g \to \infty$. By the Schwarz inequality, we conclude $E(\|Q^{*m} - U\|) \to 0$ as $g \to \infty$. Since $\|Q^{*m} - U\|$ is nonincreasing as m increases, we may conclude Theorem 1. □

5. Another theorem. The techniques used in proving Theorem 1 are useful even if k is an appropriate multiple of $\log g$ instead of an appropriate power of $\log g$. The following theorem is the result of such techniques. (We omit the use of the floor notation since such omissions do not alter the conclusion.)

THEOREM 3. *Suppose $k = a\log g$ and $m = b\log g$, where $a > e^2$, $b < a/4$ and $b\log(eb/a) < -1$. Then $E[\|Q^{*m} - U\|] \to 0$ if Q is as in Theorem 1.*

PROOF. The proof is similar to that of Theorem 1.

Although we cannot say $m \ll k$, we still may conclude $f(m,i) < 2^{m-i}$ since $m < (1/2)k$. Since $2i/k < 2m/k < 1/2$, we may again conclude

$$\sum_{i=1}^{m} \frac{[k]_i}{[k]_m} S_{2m,i} \le 2(em)^m.$$

We may also conclude that

$$\sum_{i=m+1}^{2m} S_{2m,i} \le (em)^m$$

by the same arguments as in the proof of Theorem 1. Thus we may conclude

$$\begin{aligned} E\left[\|Q^{*m} - U\|^2\right] &\leq \frac{3}{4} g \left(\frac{em}{k}\right)^m \\ &= \frac{3}{4} g \left(\frac{eb}{a}\right)^{b \log g} \\ &= \frac{3}{4} g g^{b \log(eb/a)} \\ &\to 0 \end{aligned}$$

since $b \log(eb/a) < -1$.

The theorem follows by the Schwarz inequality. □

Observe that if $a = e^2$, then $b \log(eb/a)$ has minimum value -1, and if $a < e^2$, $b \log(eb/a)$ has minimum value larger than -1. Thus our techniques are not useful if $a \leq e^2$.

6. Lower bound for certain groups. Hildebrand [12] used straightforward arguments to show that if G is an Abelian group with n elements, $k = \lfloor (\log n)^a \rfloor$ with $a > 1$, $\varepsilon > 0$ is given and

$$m < \frac{a}{a-1} \frac{\log n}{\log k} (1 - \varepsilon),$$

then $\|Q^{*m} - U\| \to 1$ as $n \to \infty$ regardless of the choice of k points. We shall generalize this lower bound to some families of finite groups with irreducible representations of bounded degree.

Such groups have received previous study in, for example, Isaacs and Passman [13] and Kaplansky [14]. In particular, Isaacs and Passman [13] showed that if the maximum degree of an irreducible representation of a finite group G is bounded by m, then there exists a function $g(m)$ such that there is a normal Abelian subgroup N of G with $[G:N] \leq g(m)$.

For our lower bounds, we shall make the following assumption:

ASSUMPTION 1. *The degree of all irreducible representations of G is less than $d_{\max}$. Furthermore, all entries of G can be expressed by $b_i n_i$, where $n_i \in N$, the Abelian normal subgroup of G (of bounded index by Isaacs and Passman [13]) and where the order of the subgroup generated by the b_i's is bounded by a constant $h(d_{\max})$.*

Note that Assumption 1 is satisfied by the dihedral groups. In the notation of Section 5.3 of Serre [17], all elements of dihedral groups are of the form r^k or sr^k. The subgroup generated by r is N, and $s^2 = 1$, so the order of the subgroup generated by the b_i's is 2 in this example.

It is not a priori clear whether there exists a function $h(d_{\max})$ such that Assumption 1 holds for all groups G with the degree of all irreducible representations of G less than $d_{\max}$.

THEOREM 4. *Suppose G satisfies Assumption* 1. *Let $n = |G|$. Let $\varepsilon > 0$ be given. Let $k = \lfloor (\log n)^a \rfloor$, $a > 1$. Let Q be as in Theorem* 1. *Then*

$$\|Q^{*m} - U\| \to 1$$

uniformly over all choices of the set S defined in Theorem 1 *if*

$$m \leq \frac{a}{a-1} \frac{\log n}{\log k} (1 - \varepsilon).$$

PROOF. The proof is a modification of the proof of the lower bound in Theorem 3 of Hildebrand [12].

Suppose the elements of G chosen in the random walk's first m steps are $b_1 n_1, \ldots, b_m n_m$. After m steps, the walk is at $b_m n_m \cdots b_2 n_2 b_1 n_1$. Since N is normal, $n_2 b_1 = b_1 n_2'$, $n_3 b_2 b_1 = b_2 b_1 n_3'$ and so forth. There are $kh(d_{\max})$ possible values for $n_1, n_2', \ldots$. Via arguments similar to those in the proof of Theorem 3 of Hildebrand [12], it can be shown that with probability approaching 1, the proportion of the values $n_1, n_2', \ldots, n_m'$ which are duplicates is under some function which approaches 0. Since N is Abelian, we may use the arguments in the proof of Theorem 3 of Hildebrand [12] to show that, except with probability approaching 0, there are at most $n^{1-\varepsilon+o(1)}$ values for $n_m' \cdots n_2' n_1$. Since $h(d_{\max})$ is a constant and there are finitely many possible elements for $b_m \cdots b_2 b_1$, we may conclude, except with probability approaching 0, there are at most $n^{1-\varepsilon+o(1)}$ possible elements reached in the group, and so the theorem follows. □

Note that Theorems 1 and 4 imply for these groups that typical random walks will have a "cutoff phenomenon" around $(a/(a-1))(\log n/\log k)$ when $k = (\log n)^a$ and $a > 1$. Further examples of this phenomenon appear in Diaconis [5] and Hildebrand [10].

7. Problems for further study. The bounds in Theorem 3 may not have the best possible constants. Perhaps techniques can be developed to improve these constants. For random random walks on $(\mathbf{Z}/2\mathbf{Z})^d$, Greenhalgh [9] and Wilson [18] obtain better constants; whether such constants can be extended to arbitrary finite groups is another question.

Another question worth studying is the factor $a/(a-1)$ in Theorem 1. This factor is required for certain groups, for example, Abelian groups. Can this factor be eliminated by appropriate choice of the finite group G? Such questions are worth exploring, but require more knowledge of the group be utilized than was in the proof of Theorem 1. A related question is to explore the extent to which Assumption 1 holds; for groups where this assumption holds, the factor $a/(a-1)$ cannot be eliminated.

One may wish to explore questions similar to those explored here, albeit on other Markov chains. For example, one may wish to explore random walks on random regular graphs where there are n vertices and each vertex has degree $(\log n)^a$. Dou [7] has explored random walks on random regular graphs but with larger degrees.

Acknowledgments. The authors would like to acknowledge that this paper is based partially on the first author's Ph.D. thesis. Furthermore, the authors would like to thank Persi Diaconis for his encouragement and his suggestions. The second author would like to thank David Wilson for showing a proof of the lower bound for the dihedral groups. The second author would like to acknowledge research support as a postdoctoral research associate at the Institute for Mathematics and Its Applications of the University of Minnesota. The first author would like to thank Peter Huber for his advice and suggestions, including a suggestion to try induction in the proof of Lemma 4. The first author also would like to thank Richard Stanley and Daniel Kleitman for their help with the combinatorics. The authors would also like to thank the referee for some suggestions.

REFERENCES

[1] AIGNER, M. (1979). *Combinatorial Theory*. Springer, New York.
[2] ALDOUS, D. and DIACONIS, P. (1985). Shuffling cards and stopping times. Technical Report 231, Dept. Statist., Stanford Univ.
[3] ALON, N. and MILMAN, V. (1985). λ_1 isoperimetric inequalities for graphs and superconcentrators. *J. Combin. Theory Ser. B* **38** 73–88.
[4] ALON, N. and ROICHMAN, Y. (1994). Random Cayley graphs and expanders. *Random Structures Algorithms* **5** 271–284.
[5] DIACONIS, P. (1988). *Group Representations in Probability and Statistics*. IMS, Hayward, CA.
[6] DIACONIS, P. (1991). Personal communication.
[7] DOU, C. (1992). Studies of random walks on groups and random graphs. Ph.D. dissertation, Dept. Mathematics, Massachusetts Institute of Technology.
[8] FRIEDMAN, J. (1991). On the second eigenvalue and random walks in random d-regular graphs. *Combinatorica* **11** 331–362.
[9] GREENHALGH, A. (1990). On a model for random random-walks on finite groups. Preprint.
[10] HILDEBRAND, M. (1992). Generating random elements in $\mathrm{SL}_n(\mathbf{F}_q)$ by random transvections. *J. Alg. Comb.* **1** 133–150.
[11] HILDEBRAND, M. (1993). Random processes of the form $X_{n+1} = a_n X_n + b_n \pmod p$. *Ann. Probab.* **21** 710–720.
[12] HILDEBRAND, M. (1994). Random walks supported on random points of $\mathbf{Z}/n\mathbf{Z}$. *Probab. Theory Related Fields* **100** 191–203.
[13] ISAACS, I. M. and PASSMAN, D. S. (1964). Groups with representations of bounded degree. *Canad. J. Math.* **16** 299–309.
[14] KAPLANSKY, I. (1949). Groups with representations of bounded degree. *Canad. J. Math.* **1** 105–112.
[15] LUBOTZKY, A. (1994). *Discrete Groups, Expanding Graphs, and Invariant Measures*. Birkhäuser, Boston.
[16] LUBOTZKY, A. and WEISS, B. (1993). Groups and expanders. In *Expanding Graphs: Proceedings of a DIMACS Workshop* (J. Friedman, ed.) 95–109. Amer. Math Soc., Providence, RI.
[17] SERRE, J.-P. (1979). *Linear Representations of Finite Groups*. Springer, New York.
[18] WILSON, D. (1996). Random random walks on Z_2^d. *Probab. Theory Related Fields*. To appear.

J. P. MORGAN & CO., INC.
60 WALL STREET, 4TH FLOOR
NEW YORK, NEW YORK 10260

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TEXAS AT AUSTIN
AUSTIN, TEXAS 78712
E-MAIL: mvh@math.utexas.edu