The Annals of Probability
1993, Vol. 21, No. 4, 2131–2156

COMPARISON TECHNIQUES FOR RANDOM WALK ON FINITE GROUPS

BY PERSI DIACONIS AND LAURENT SALOFF-COSTE

Harvard University and Université de Paris VI CNRS

We develop techniques for bounding the rate of convergence of a symmetric random walk on a finite group to the uniform distribution. The techniques gives bounds on the second largest (and other) eigenvalues in terms of the eigenvalues of a comparison chain with known eigenvalues. The techniques yield sharp rates for a host of previously intractable problems on the symmetric group.

1. Introduction. This paper develops techniques for bounding the rate of convergence of a symmetric random walk on a finite group. Let G be a finite group of order $|G| = g$. Let id denote the identity of G. Let E be a symmetric set of generators: $E^{-1} = E$. This E can be used to define a random walk with steps chosen uniformly from E.

Familiar examples include simple random walk on the integers $(\operatorname{mod} m)$ where $E = \{1, -1\}$, the Ehrenfest walk on the cube $\mathbb{Z}_2^d$, where $E = \{e_i: 1 \le i \le d\}$, $e_i =$ the ith standard basis vector, or the random walk on the symmetric group S_n which proceeds by repeated random transpositions where $E = \{(i, j): 1 \le i < j \le n\}$.

These examples have all been analyzed using Fourier analysis on the appropriate group. Diaconis [(1988), Chapter 3] gives background and details. Fourier analysis gives all the eigenvalues of the associated Markov chains in terms of the characters of the group. The main results of this paper show how these eigenvalues can be used to get good bounds for less symmetric measures.

As a running example, consider $G = S_n$ and the random walk generated by a transposition and an n-cycle:

$$E = \{id, (1,2), (n, n-1, n-2, \dots, 1), (1, 2, \dots, n)\}. \tag{1.1}$$

In Section 3 we show that order $n^3 \log n$ steps suffice to achieve randomness for this walk and that order n^3 steps are necessary. This result follows from comparison with the walk generated by random transpositions. The same techniques work for a host of other walks that have defied previous analysis: two different models for the familiar overhand shuffle and some "two-dimensional" shuffles where the deck is arranged in a $k \times l$ grid, a card is chosen at random, and switched with one of its nearest neighbors. These are developed in Section 4.

Received January 1992; revised September 1992.

AMS 1991 *subject classifications*. 20B30, 60B15, 60J05, 60F99.

Key words and phrases. Card shuffling, reversible Markov chains, random walk, groups, eigenvalues.

Section 2 lays out preliminaries on norms, eigenvalues and the two quadratic forms we use. It gives bounds for standard distances such as total variation in terms of eigenvalues.

Section 3 gives upper and lower bounds for eigenvalues by comparison. One can always compare with the uniform distribution and the bounds are shown to specialize to known results giving bounds on the second largest eigenvalue in terms of the diameter of the group in the generators E. These techniques are illustrated in example (1.1) and for a class of examples on $\mathbb{Z}_m$, the integers mod m.

Section 5 treats natural product random walks.

The techniques of this paper can be supplemented by volume growth estimates to give sharp results for random walks on nilpotent groups such as the Heisenberg group. This is carried out in Diaconis and Saloff-Coste (1992). Comparisons can also be carried out for reversible Markov chains where they offer a supplement to the geometric techniques of Diaconis and Stroock (1991). We use them to get sharp bounds on the eigenvalues of the simple exclusion process treated by Fill (1991) in Diaconis and Saloff-Coste (1993).

The techniques in this paper often seem to give the correct order [viz. $O(n^3 \log n)$ in example (1.1)]. They usually do not give sharp lead term constants and so do not lead to proofs of the cutoff phenomenon that so often occurs.

2. Norms, forms, and eigenvalues.

This section gives preliminaries on distances from uniformity, the two basic quadratic forms to be used and some comparison inequalities. The main result is Lemma 5 from uniformity which gives upper bounds on the L^2 distance, of one probability in terms of a second probability in the presence of a comparison between their Dirichlet forms. The results are elementary, but we hope it is convenient to have them collected together.

Norms. 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 $u(x) \equiv 1/|G|$. Thus $U(\varphi)$ is the mean of φ over G. There is one exception to our notation: The function equal to one at id and zero elsewhere is denoted δ_{id} (the Dirac mass at id). The associated operator is the identity I.

For $1 \le s \le \infty$, let

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

denote the usual l^s norms of a function φ. Recall that for $1 \le s' \le s \le \infty$:

$$g^{(1/s)-(1/s')} \|\varphi\|_{s'} \le \|\varphi\|_s \le \|\varphi\|_{s'}.$$

The variation distance $\|p-\tilde{p}\|_{TV}=\max_{I\subset G}\{|p(I)-\tilde{p}(I)|\}$ between two probabilities $p,\tilde{p}$ is just half the l^1 norm of $p-\tilde{p}$. The operator norm of Ψ from $l^{s'}$ to l^s is denoted $\|\Psi\|_{s'\to s}$. The following classical inequalities are useful:

$$\|\Psi\|_{s\to s}\le\|\psi\|_1 \quad \text{for all } 1\le s\le\infty,$$

$$\|\Psi\|_{s\to\infty}=\|\Psi\|_{1\to\bar{s}}=\|\psi\|_{\bar{s}} \quad \text{for all } 1\le s\le\infty,\ \frac{1}{s}+\frac{1}{\bar{s}}=1.$$

In particular, for $s=2$ we get $\|\psi\|_2=\|\Psi\|_{2\to\infty}=\|\Psi\|_{1\to 2}$.

Throughout this paper, p denotes a symmetric probability on G. We are interested in bounding the rate of convergence of the convolution powers $p^{(n)}$ to the uniform distribution u. We concentrate on bounds for total variation. However, these are achieved by bounding the l^1 norm by the l^2 norm using the Cauchy–Schwarz inequality. More precisely, if we define a normalized distance

$$d_s(n)=g^{1-1/s}\|p^{(n)}-u\|_s, \qquad 1\le s\le\infty,$$

our bounds on total variation are obtained from

$$\|p^{(n)}-u\|_{TV}=\tfrac{1}{2}d_1(n)\le\tfrac{1}{2}d_2(n)$$

and bounding $d_2(n)$ by eigenvalue estimates. In other words, all the bounds on $\|p^{(n)}-u\|_{TV}$ stated in this paper are in fact bounds on $(1/2)d_2(n)$. This is also true of the bounds obtained by Fourier analysis in Diaconis (1988). The distance d_1 is twice the total variation distance while d_∞ is the maximum relative error. It turns out that in many interesting examples good bounds on d_2 yield good bounds on total variation (however, see Example 1 below and Example 1 of Section 5). Note also that $d_2(n)\le d_\infty(n)$ and $d_\infty(2n)\le d_2^2(n)$, whereas it is not possible in general to obtain good bounds on d_2 or d_∞ from bounds on total variation.

Eigenvalues. Because p is symmetric, the matrix $\{p(y^{-1}x)\}_{x,y\in G}$ has real eigenvalues $1=\pi_0\ge\pi_1\ge\cdots\ge\pi_{g-1}\ge-1$. We set $\pi_*=\max\{\pi_1,|\pi_{g-1}|\}$. The importance of π_* comes from

$$\|P-U\|_{2\to 2}=\pi_*.$$

Note that $\pi_{g-1}>-1$ as soon as $p(id)>0$. Indeed:

LEMMA 1. $\pi_{g-1}\ge-1+2p(id)$.

PROOF. The result is true if $p(id)=0$. If $p(id)>0$, let $q=(1-p(id))^{-1}(p-p(id)\delta_{id})$. This is a symmetric probability with smallest eigenvalue bounded below by -1, that is, $-1\le(1-p(id))^{-1}(\pi_{g-1}-p(id))$. This gives the result. □

Let us also consider the continuous time semigroup $H_t=e^{-t(I-P)}$ and its convolution kernel

$$h_t=e^{-t}\sum_0^\infty\frac{t^n}{n!}p^{(n)}.$$

The eigenvalues of H_t are the numbers $e^{-t\lambda_i}$ where $\lambda_i = 1 - \pi_i$ are the eigenvalues of $I - P$ and

$$\|H_t - U\|_{2\to 2} = e^{-\lambda_1 t}.$$

The eigenvalues π_* and λ_1 give simple estimates on the distance to uniformity as follows:

LEMMA 2. *Let p be a symmetric probability on a finite group G. Let π_* be the second largest eigenvalue in absolute value. Then*

$$d_1(n) = 2\|p^{(n)} - u\|_{TV} \le g^{1/2}\|p^{(n)} - u\|_2 = d_2(n) \le g^{1/2}\pi_*^n, \tag{2.1}$$

$$g\|p^{(n)} - u\|_\infty = d_\infty(n) \le g\pi_*^n. \tag{2.2}$$

The same estimates hold if we replace $p^{(n)}$ by h_t and π_*^n by $e^{-t\lambda_1}$.

PROOF. For (2.1), write

$$\begin{aligned} g^{1/2}\|p^{(n)} - u\|_2 &= g^{1/2}\|(P^n - U)\delta_{id}\|_2 = g^{1/2}\big\|(P - U)^n\delta_{id}\big\|_2 \\ &\le g^{1/2}\|P - U\|_{2\to 2}^n\|\delta_{id}\|_2 = g^{1/2}\pi_*^n. \end{aligned}$$

For (2.2), use (2.1) and

$$\|p^{(n+k)} - u\|_\infty \le \|p^{(n)} - u\|_2\|p^{(k)} - u\|_2.$$

The argument for h_t is similar. □

EXAMPLE 1. Usually, bounds that use only π_* are crude. To study this, for $0 \le \theta \le 1$, let $u_\theta = (1 - \theta)\delta_{id} + \theta u$ be a probability on G. Then

$$u_\theta^{(n)} = (1 - \theta)^n\delta_{id} + \big(1 - (1 - \theta)^n\big)u$$

and $\pi_* = 1 - \theta$. It is easy to check that

$$\begin{aligned} \|u_\theta^{(n)} - u\|_2 &= \left(1 - \frac{1}{g}\right)^{1/2}(1 - \theta)^n, \\ \|u_\theta^{(n)} - u\|_\infty &= \left(1 - \frac{1}{g}\right)(1 - \theta)^n, \\ \|u_\theta^{(n)} - u\|_1 &= 2\left(1 - \frac{1}{g}\right)(1 - \theta)^n. \end{aligned}$$

Thus the bounds (2.1) and (2.2) for the d_2 and d_∞ norms are essentially equalities but the bound for the d_1 norm can be far wrong if g is large.

The l^2 norms have the clearest connection to eigenvalues. We have

$$\begin{aligned} \|p^{(n)} - u\|_2^2 &= p^{(2n)}(id) - \frac{1}{g} = \frac{1}{g}\sum_1^{g-1}\pi_i^{2n}, \\ \|h_t - u\|_2^2 &= h_{2t}(id) - \frac{1}{g} = \frac{1}{g}\sum_1^{g-1}e^{-2t\lambda_j}. \end{aligned} \tag{2.4}$$

The continuous time process is often used to avoid parity problems. The bounds (2.4) show that negative eigenvalues π_i are important for bounding $\|p^{(n)} - u\|_2$ while they only appear in a minor way for the continuous time processes. Most examples worked out in the sequel are for discrete time and apply throughout to continuous time versions. In general, there is no easy transfer of information between h_t and $p^{(n)}$ except for the following simple result.

LEMMA 3. *Let p be a symmetric probability on a finite group G. Then*

$$\|p^{(n)} - u\|_2^2 \le \pi_{g-1}^{2n} + \|h_n - u\|_2^2, \tag{2.5}$$

$$\|h_{2n} - u\|_2^2 \le e^{-2n} + \|p^{(n)} - u\|_2^2. \tag{2.6}$$

PROOF. The first statement follows from (2.4) and the inequality $1 - x \le e^{-x}$. In more detail,

$$\|p^{(n)} - u\|_2^2 = \frac{1}{g}\sum_{1}^{g-1} \pi_i^{2n} \le \pi_{g-1}^{2n} + \frac{1}{g}\sum_{\pi_i > 0} \pi_i^{2n} \le \pi_{g-1}^{2n} + \frac{1}{g}\sum_{\lambda_i < 1} e^{-2n\lambda_i}$$

$$\le \pi_{g-1}^{2n} + \frac{1}{g}\sum_{1}^{g-1} e^{-2n\lambda_i} = \pi_{g-1}^{2n} + \|h_n - u\|_2^2.$$

The second statement follows from (2.4) and the inequality $1 - x \ge e^{-2x}$ for $0 \le x \le 1/2$. □

In this paper, eigenvalues are used to study convergence to stationarity. Let $\nu_i = 1/(1 - \pi_i)$, $1 \le i \le g - 1$. Let a random walk start at a uniformly chosen point. Let τ be the first hitting time to a previously specified point. Aldous (1989) shows $E(\tau) = \nu_1 + \nu_2 + \cdots + \nu_{g-1}$. Thus bounds on hitting times follow from bounds on eigenvalues.

Forms. The eigenvalues of symmetric probabilities can be characterized using quadratic forms. Let

$$\begin{aligned} \mathscr{E}_p(\varphi, \varphi) &= \mathscr{E}(\varphi, \varphi) = \langle (I - P)\varphi, \varphi\rangle = \tfrac{1}{2}\sum_{x,y} (\varphi(x) - \varphi(xy))^2 p(y), \\ \mathscr{F}_p(\varphi, \varphi) &= \mathscr{F}(\varphi, \varphi) = \langle (I + P)\varphi, \varphi\rangle = \tfrac{1}{2}\sum_{x,y} (\varphi(x) + \varphi(xy))^2 p(y). \end{aligned} \tag{2.7}$$

The form $\mathscr{E}_p$ is called the *Dirichlet form*. It can be used to get lower bounds on the eigenvalues $\lambda_i = 1 - \pi_i$ of $I - P$. The form $\mathscr{F}_p$ is useful for getting lower bounds on negative eigenvalues π_i.

The eigenvalues can be characterized by the minimax principle, which we briefly recall [see, e.g., Horn and Johnson (1985), page 176]. Let V be a real finite-dimensional Hilbert space and Q a symmetric linear operator on V with eigenvalues $q_0 \le q_1 \le \cdots$. Given a subspace W of V, set

$$m(W) = \min\{\langle Qf, f\rangle / \langle f, f\rangle : f \in W\},$$
$$M(W) = \max\{\langle Qf, f\rangle / \langle f, f\rangle : f \in W\}.$$

Then

$$q_i = \max\{m(W) : \dim(W^{\perp}) = i\} = \min\{M(W) : \dim(W) = i + 1\}.$$

The next lemma follows immediately from the minimax principle.

LEMMA 4. *Let p and $\tilde{p}$ be two symmetric probabilities on a finite group G with eigenvalues π_i and $\tilde{\pi}_i$. If $\tilde{\mathscr{E}} \le A\mathscr{E}$ then $\pi_i \le 1 - (1 - \tilde{\pi}_i)/A$. If $\tilde{\mathscr{F}} \le A\mathscr{F}$, then $\pi_i \ge -1 + (1 + \tilde{\pi}_i)/A$.*

Lemma 4 coupled with the previous discussion gives the following bounds which are the principal results of this section.

LEMMA 5. *If $\tilde{\mathscr{E}} \le A\mathscr{E}$, then*

$$\|p^{(n)} - u\|_2^2 \le \pi_{g-1}^{2n} + \|\tilde{h}_{n/A} - u\|_2^2 \le \pi_{g-1}^{2n} + e^{-n/A} + \|\tilde{p}^{(\lfloor n/2A\rfloor)} - u\|_2^2, \tag{2.8}$$

$$\|h_t - u\|_2^2 \le \|\tilde{h}_{t/A} - u\|_2^2. \tag{2.9}$$

If $\tilde{\mathscr{E}} \le A\mathscr{E}$, and $\tilde{\mathscr{F}} \le A\mathscr{F}$, then

$$\|p^{(n)} - u\|_2^2 \le e^{-n/A} + \|\tilde{p}^{(\lfloor n/2A\rfloor)} - u\|_2^2. \tag{2.10}$$

PROOF. All results follow from (2.4), Lemma 4 and the inequalities $1 - x \le e^{-x}$, $1 - x \ge e^{-2x}$, $0 \le x \le 1/2$. For example, from (2.4),

$$\|p^{(n)} - u\|_2^2 = \frac{1}{g}\sum_{i=1}^{g-1} \pi_i^{2n} \le \pi_{g-1}^{2n} + \frac{1}{g}\sum_{\pi_i > 0} \pi_i^{2n}.$$

By hypothesis, $\pi_i = 1 - \lambda_i \le 1 - \tilde{\lambda}_i/A \le e^{-\tilde{\lambda}_i/A}$. This yields

$$\|p^{(n)} - u\|_2^2 \le \pi_{g-1}^{2n} + \frac{1}{g}\sum_{\pi_i > 0} e^{-2n\tilde{\lambda}_i/A} \le \pi_{g-1}^{2n} + \|\tilde{h}_{n/A} - u\|_2^2,$$

which is the first claimed inequality. The second follows from (2.6). The proof of (2.9) is similar. For (2.10), set $I = \{i;\ |\tilde{\pi}_i| < 1/2\}$ and note that our hypothesis and Lemma 4 imply that $1 - |\tilde{\pi}_i| \le A(1 - |\pi_i|)$. Then write

$$\begin{aligned}\|p^{(n)} - u\|_2^2 &\le \frac{1}{g}\left(\sum_I \pi_i^{2n} + \sum_{I^c} \pi_i^{2n}\right) \le e^{-n/A} + \frac{1}{g}\sum_{I^c} e^{-2n(1-|\pi_i|)}\\ &\le e^{-n/A} + \frac{1}{g}\sum_{I^c} e^{-(2n/A)(1-|\tilde{\pi}_i|)} \le e^{-n/A} + \frac{1}{g}\sum_{I^c} |\tilde{\pi}_i|^{n/A}\\ &\le e^{-n/A} + \|\tilde{p}^{\lfloor n/2A\rfloor} - u\|_2^2.\end{aligned}$$

□

REMARKS. (a) If p_1, p_2 are symmetric probabilities, set $p = (p_1 + p_2)/2$. The estimate (2.10) yields

$$\|p^{(n)} - u\|_2^2 \le e^{-n/2} + \min_{i=1,2} \|p_i^{(\lfloor n/4 \rfloor)} - u\|_2^2.$$

We do not see how to prove such a comparison directly.

(b) Although the form $\mathscr{F}$ is useful in bounding negative eigenvalues, it is sometimes harder to use than $\mathscr{E}$. Here is a trick that shows that the negative eigenvalues other than π_{g-1} do not play much of a role. Indeed, observe that

$$0 \le g p^{(2n+1)}(id) = \sum_0^{g-1} \pi_i^{2n+1}.$$

It follows that

$$\sum_{\pi_i < 0} \pi_i^{2n+2} \le \sum_{\pi_i > 0} \pi_i^{2n}.$$

Hence, we can bound $\|p^{(N)} - u\|_2^2$ by writing $N = n + n' + 1$ and

$$\|p^{(N)} - u\|_2^2 \le g^{-1}\Bigl(\sum_{\pi_i > 0} \pi_i^{2n'}\Bigr)\pi_{g-1}^{2n} + g^{-1} \sum_{0 < \pi_i < 1} \pi_i^{2N}.$$

From this, we deduce as before the following:

LEMMA 6. *If $\tilde{\mathscr{E}} \le A\mathscr{E}$, we have for $N = n + n' + 1$,*

$$\begin{aligned}\|p^{(N)} - u\|_2^2 \le{}& g^{-1}\pi_{g-1}^{2n} + e^{-n'/A}\pi_{g-1}^{2n} + \|\tilde{p}^{(\lfloor n'/2A \rfloor)} - u\|_2^2 \pi_{g-1}^{2n}\\ &+ e^{-N/A} + \|\tilde{p}^{(\lfloor N/2A \rfloor)} - u\|_2^2.\end{aligned}$$

Examples that use this lemma are given in subsections 4A, 4B and 4C.

REMARK. In Lemmas 3, 5 and 6 we used π_{g-1} to bound the negative eigenvalues of p. When p has no negative eigenvalues, π_{g-1} can be replaced by 0 in these lemmas.

3. Comparison of forms and first examples. This section develops bounds of the type $\tilde{\mathscr{E}} \le A\mathscr{E}$ for Dirichlet forms associated with symmetric probabilities $\tilde{p}$ and p on a finite group G. The constant A is an average length which can often be usefully bounded in examples of interest. The techniques are illustrated for the chain generated by a transposition and an n-cycle, for a class of walks on $\mathbb{Z}_m$ and for the Ehrenfest walk.

Let E be a symmetric set of generators of the finite group G. For $y \in G$, write $y = z_1 z_2 \cdots z_k$ with $z_i \in E$. The smallest such k is called the length of y and denoted $|y| = |y|_E$. By definition the identity id has length 0. Let

$$N(z, y) = \text{number of times } z \in E \text{ occurs in the chosen representation of } y. \tag{3.1}$$

Clearly

$$N(z,y) \le |y|. \tag{3.2}$$

THEOREM 1. *Let $\tilde{p}$ and p be symmetric probabilities on a finite group G. Let E be a symmetric set of generators. Suppose that the support of p contains E. Then the Dirichlet forms defined in* (2.7) *satisfy*

$$\tilde{\mathscr{E}} \le A\mathscr{E}$$

with

$$A = \max_{x \in E} \frac{1}{p(z)} \sum_{y \in G} |y| N(z,y) \tilde{p}(y). \tag{3.3}$$

PROOF. Given $x, y \in G$, suppose $y = z_1 z_2 \cdots z_k$ with $z_i \in E$. Then

$$\varphi(x) - \varphi(xy) = \{(\varphi(x) - \varphi(xz_1)) + (\varphi(xz_1) - \varphi(xz_1 z_2)) + \cdots + (\varphi(xz_1 \cdots z_{k-1}) - \varphi(xy))\}.$$

Squaring both sides and using the Cauchy–Schwarz inequality,

$$(\varphi(x) - \varphi(xy))^2 \le |y|\{(\varphi(x) - \varphi(xz_1))^2 + \cdots + (\varphi(xz_1 \cdots z_{k-1}) - \varphi(xy))^2\}.$$

Summing in x gives

$$\sum_{x \in G} (\varphi(x) - \varphi(xy))^2 \le |y| \sum_{\substack{x \in G \\ z \in E}} (\varphi(x) - \varphi(xz))^2 N(z,y).$$

The result follows after multiplying both sides by $\tilde{p}(y)$, summing over $y \in G$, and dividing by 2. Then the left-hand side is $\tilde{\mathscr{E}}(\varphi, \varphi)$ while the right-hand side is

$$\frac{1}{2} \sum_{\substack{x \in G \\ z \in E}} (\varphi(x) - \varphi(xz))^2 p(z) \frac{1}{p(z)} \sum_{y \in G} |y| N(z,y) \tilde{p}(y) \le A\mathscr{E}(\varphi, \varphi). \qquad \square$$

Before developing further bounds we give some examples. The first result follows by choosing $\tilde{p} = u$, the uniform distribution:

COROLLARY 1. *Let G be a finite group and E a symmetric set of generators. Let p be any symmetric distribution with support of p containing E. Let $\eta = \min_{z \in E} p(z)$, $r = \max\{|y|: y \in G\}$. Then*

$$\pi_1(p) \le 1 - \frac{1}{A} \le 1 - \frac{\eta}{r^2}. \tag{3.4}$$

Versions of this inequality have been given by Aldous (1987), Babai (1990), Gangolli (1991) and Mohar (1989).

Example 1. For our running example, $G = S_n$, $E = \{id, (1,2), (1,2,\dots,n), (n, n-1,\dots,1)\}$. Then $\eta = 1/4$. It is straightforward to show that the generators $(1,2), (1,2,\dots,n)$ require at worst $3\binom{n}{2}$ steps to represent any permutation. The idea is to work from the bottom up. If the bottom i cards are in the correct order, with the $i+1$st card somewhere above them, move cards from top to bottom until this $i+1$st card is at the top. Then, transpose and shift repeatedly to bring this card just next to the original bottom block of i. Then cut these $i+1$ cards to the bottom. Thus $r \le 3n^2/2$ and $\pi_1 \le 1 - 1/(9n^4)$. This will now be improved to $1 - c/n^3$ by comparison with a measure supported on transpositions.

Let $\tilde{p}(id) = 1/n$, $\tilde{p}(s) = 2/n^2$ for s any transposition and $\tilde{p}(\pi) = 0$ otherwise. Diaconis and Shahshahani (1981) analyzed this chain, determining all the eigenvalues using Fourier analysis on the symmetric group. The bound (3.3) gives

$$A \le 4\sum_y |y|^2 \tilde{p}(y) \le 36n^2.$$

The last inequality follows because any transposition can be written with at most $3n$ generators. Diaconis and Shahshahani (1981) showed $\tilde{\pi}_1 = 1 - 2/n$. Using this in Lemma 4 gives

$$\pi_1 \le 1 - 1/(18n^3).$$

To see that this bound is of the right order, consider $\varphi(\pi)$ as the circular distance between $\pi^{-1}(1)$ and $\pi^{-1}(2)$. If permutations are associated to arrangements of n cards in such a way that $\pi(i)$ denotes the label of the card at position i, then $\pi^{-1}(i)$ is the position of the card labeled i and $\varphi(\pi)$ is the circular distance between cards labeled 1 and 2. Now the minimax characterization of eigenvalues gives

$$\pi_1 \ge 1 - \mathscr{E}(\varphi,\varphi)/\|\varphi - U\varphi\|_2^2.$$

It is straightforward to show that $\|\varphi - U\varphi\|_2^2 \sim n^2 n!/48$. On the other hand,

$$\mathscr{E}(\varphi,\varphi) = \tfrac{1}{8}\sum_{\substack{x\in G\\ y\in E}} (\varphi(x) - \varphi(xy))^2 \sim \tfrac{1}{8}4(n-1)!.$$

These bounds give

$$\pi_1 \ge 1 - \frac{24 + o(1)}{n^3}.$$

To finish this example, observe that use of just the second eigenvalue together with $\pi_{g-1} \ge -1/2$ (from Lemma 1 of Section 2) shows that order $n^4 \log n$ steps suffice to drive the variation distance close to zero. This can be improved by making full use of the comparison as in (2.8).

THEOREM 2. *Let p be the uniform distribution on $E = \{id, (1,2), (1,2,\dots,n), (n, n-1, n-2,\dots,1)\}$ in the symmetric group S_n. If $k = 36n^3(\log n + c)$, then, for $c > 0$,*

$$\|p^{(k)} - u\|_{TV} \le \alpha e^{-c}$$

for a universal positive constant α.

If $k = cn^3$, then

$$\liminf_{n\to\infty} \|p^{(k)} - u\|_{TV} \ge 1 - f(c)$$

with $f(c)$ tending to zero as c tends to zero.

PROOF. Compare with $\tilde{p}$, the random transposition measure. The upper bound follows from $\tilde{\mathscr{E}} \le 36n^2\mathscr{E}$, $\pi_{g-1} \ge -1/2$ (from Lemma 1 of Section 2), and (2.8). Diaconis and Shahshahani (1981) showed that $g\|\tilde{p}^{(m)} - u\|_2^2 \le \beta e^{-2c}$ for an explicit universal $\beta > 0$ when $m = (1/2)n(\log n + c)$.

For the lower bound, consider $\varphi(\pi)$, the circular distance between $\pi^{-1}(1)$ and $\pi^{-1}(2)$ as above. This takes values in $\{1, 2, \dots, n/2\}$. It changes by at most 1, doing this only when at least one of the cards labeled 1 or 2 is on top or in the second position. Elementary considerations, comparing with random walk on an interval with geometric wait size, show that this distance requires order n^3 steps to have an appreciable chance of being of order n, its size under the uniform distribution. Further details are omitted. □

We turn next to bounds involving the form $\mathscr{F}$ defined in (2.7). The results here use paths in the following way: If x and y are elements of G and $y = z_1 z_2 \cdots z_k$ with k odd, then

$$\begin{aligned}\varphi(x) + \varphi(xy) = {} & (\varphi(x) + \varphi(xz_1)) - (\varphi(xz_1) + \varphi(xz_1z_2)) \\ & + \cdots + (\varphi(xz_1 \cdots z_{k-1}) + \varphi(xy)).\end{aligned}$$

Thus, let $|y|_*$ be the length of the shortest representation of y as a product of an *odd* number of generators (if the identity is in E, then $|y|_* \le |y| + 1$). We set $|y|_* = \infty$ if y cannot be so expressed. Now, $|id|_* > 0$. The function $N_*(z, y)$ is defined as in (3.1). With this notation, the proof of Theorem 1 goes through word for word to give the following:

THEOREM 3. *Let $\tilde{p}$ and p be symmetric probabilities on a finite group G with support of p containing E, a symmetric set of generators. Then the forms $\tilde{\mathscr{F}}$, $\mathscr{F}$ defined in (2.7) satisfy*

$$\tilde{\mathscr{F}} \le A_* \mathscr{F}$$

with

$$A_* = \max_{z\in E} \frac{1}{p(z)} \sum_{y\in G} |y|_* N_*(z,y)\tilde{p}(y).$$

Choosing $\tilde{p}$ as the uniform distribution gives a lower bound for the smallest eigenvalue:

Corollary 2. *Let G be a finite group and E a symmetric set of generators. Let p be a symmetric probability with support of p containing E. Let $\eta = \min_{z \in E} p(z)$ and $r_* = \max\{|y|_*: y \in G\}$. Then*

$$\pi_{g-1}(p) \ge -1 + \frac{1}{A_*} \ge -1 + \frac{\eta}{r_*^2}.$$

Remark. It follows from Corollaries 1 and 2 that $\pi_*(p) \le 1 - \eta/r_*^2$.

Example 2. Take $G = \mathbb{Z}_m$, the integers modulo m with m odd. Take $E = \{1, -1\}$ with $p(1) = p(-1) = 1/2$. This is the classical gambler's walk. Compare with the uniform distribution. Clearly $A_* \le (2/m)\Sigma_y |y|_*^2 \le 2m^2$, so $\pi_{m-1} \ge -1 + 1/2m^2$. For this walk the eigenvalues are $\cos(2\pi j/m)$, $0 \le j \le m$. Thus the smallest eigenvalue is $\cos(\pi - \pi/m) = -1 + \pi^2/2m^2 + O(1/m^4)$. Thus the bound is of the correct order for large m. Upper bounds on π_1 for this example are discussed in Example 4 below.

Example 3. Consider $G = S_n$ and $E = \{(1, 2), (1, 2, \dots, n), (n, n-1, \dots, 1)\}$ with p uniform on E. This is example (1.1) with the identity deleted. To avoid parity problems, suppose n is odd. Compare with the random transposition measure p described before Theorem 2. The quantity A_* is bounded above by

$$A_* \le 3\left\{\frac{n^2}{n} + \frac{2}{n^2}\binom{n}{2}9n^2\right\} \le 3(n + 9n^2).$$

The first term in curly brackets comes from the identity, the second term comes from the $\binom{n}{2}$ transpositions. Diaconis and Shahshahani (1981) found $\tilde{\pi}_{g-1} = -1 + 2/n$. Now Theorem 3 yields $\pi_{g-1} \ge -1 + 1/15n^3$. An upper bound for A (see 3.3) follows as in Theorem 2. Using these results and Lemma 5 of Section 2 shows that order $n^3 \log n$ steps suffice for this set of generators.

The next two examples show how the function $N(z, y)$ enters the bounds.

Example 4. Let $G = \mathbb{Z}_m$. Let $a \in G$ with $a \le m^{1/2}$ and choose $E = \{-a, -a + 1, \dots, a\}$. Take p uniform on E, so $p(x) = 1/(2a + 1)$ for $x \in E$ and zero otherwise. We derive bounds on the second eigenvalue π_1 using Theorem 1. We show that, for a universal $c > 0$,

$$\pi_1 \le 1 - \frac{ca^2}{m^2}, \tag{3.5}$$

this result being uniform in $|a| \le m^{1/2}$. Here, Fourier analysis can be used to show that (3.5) is an equality up to a constant.

For definiteness, suppose m and a are even. Break $\mathbb{Z}_m$ into right and left halves working symmetrically with the two parts. Identify the right half with $0, 1, \dots, m/2$. For y in the right half, write $y = \alpha a + \beta$, $0 \le \beta < a$. Represent y as

$$y = \overbrace{(0 + a) + (1 + a - 1) + \cdots + (j + a - j)}^{\alpha \text{ terms}} + \beta.$$

Here, if $\alpha > ak$, the terms repeat cyclically. For example, if $m = 27$ and $a = 3$, represent $13 = (0 + 3) + (1 + 2) + (0 + 3) + (1 + 2) + 1$. This is not the minimum length description but $|y|$ in the definition of A in (3.3) (with $\tilde{p} = u$) can be defined as $2\alpha + 1$ and the bound goes through as stated. This representation of y may use pairs $(0, a), (1, a - 1) \cdots$ cyclically, and of course a given pair may need to be used many times. With these conventions, consider

$$A = \max_{z} \frac{(2a + 1)}{m} \sum_{y=0}^{m-1} |y| N(z, y).$$

For any z, $N(z, y) \le N(a, y) + 1$. For $y = \alpha a + \beta$, $\alpha = \lfloor y/a \rfloor$ and $N(a, y) = \lceil \alpha/a \rceil \le y/a^2 + 1$. The sum over y in the right half is thus bounded above by

$$\sum_{y=0}^{m/2} |y| N(a, y) \le \sum_{y=0}^{m/2} \left(2\frac{y}{a} + 1\right)\left(\frac{y}{a^2} + 2\right).$$

This implies that $A \le Cm^2/a^2$ (recall that $a \le m^{1/2}$), which leads to the bound (3.5). Using more naive paths:

$$\overbrace{a + a + \cdots + a}^{\alpha \text{ terms}} + \beta$$

leads to a bound of the wrong order of magnitude for a large.

EXAMPLE 5. Let $G = \mathbb{Z}_2^d$ be the "cube" and choose $E = \{e_i, 1 \le i \le d\}$ with e_i the usual ith basis vector. Take p uniform on E so $p(x) = 1/d$ for $x \in E$ and zero elsewhere. There is a unique minimum length path up to order and for any choice of $z \in E$,

$$\frac{d}{2^d} \sum_{y} |y| N(z, y) = \frac{d}{2^d}(d + 1)2^{d-2}.$$

Corollary 1 gives

$$\pi_1 \le 1 - \frac{4}{d(d + 1)}.$$

As is well known [see, e.g. Diaconis (1988), Chapter 3], $\pi_1 = 1 - 2/d$ so that the bound is "off" by a factor of order d. Diaconis and Stroock (1991) show how to use paths on a collapsed chain to get the correct answer.

REMARKS. (a) There has been considerable work in the computer science community deriving diameter bounds for groups. For example, Driscoll and

Furst (1987) show that the diameter of a permutation group of degree n generated by cycles of bounded degree is $0(n^2)$. Babai, Hetyii, Kantor, Lubotzky and Seress (1990) contains a survey. Babai, Kantor and Lubotzky (1989) give generating sets of size less than 7 for the classical families of finite simple groups. Their paper contains many examples where one generating set is written in terms of a second.

(b) Comparison bounds can be developed for reversible Markov chains. Such bounds are used in Diaconis and Saloff-Coste (1992b) to get sharp rates for a variety of exclusion processes. The bounds specialize to those given here when the Markov chain is symmetric random walk on a group and give a geometric interpretation to A as a measure of "bottlenecks" along the lines of Diaconis and Stroock (1991).

4. Examples in the symmetric group.

This section presents analysis of shuffling schemes on the symmetric group. They are arranged as: shuffles involving transpositions, shuffles involving cycles, overhand shuffles and other shuffles. A few of the shuffles have been analyzed before so we can evaluate the new techniques on problems with known answers. Many results below represent the first analysis of a natural shuffling scheme that has previously defied analysis. In most cases the new techniques give the right answer up to small numerical constants (e.g., $3n \log n$ where the right answer is $n \log n$). We have not attempted to get the sharpest possible constants.

A. Shuffles involving transpositions. Let $\mathscr{G}$ be an undirected graph on $\{1, 2, \dots, n\}$ with edge set E. Each edge (i, j) can be thought of as a transposition in the symmetric group S_n. It is well known that a set of transpositions generates S_n if and only if $\mathscr{G}$ is connected. For example, any spanning tree gives rise to a set of generators. A random walk on S_n generated by $\mathscr{G}$ can be described as follows. To start, place cards labelled $1, 2, \dots, n$ at the vertices of $\mathscr{G}$. At each stage, an edge is chosen at random and the two cards at the ends of the edge are switched.

The main result of this section gives bounds on the rate of convergence to the uniform distribution in terms of the geometry of the underlying graph. The results follow by comparison with known results for the complete graph.

To describe things, for each pair x, y let γ_{xy} be a path from x to y in $\mathscr{G}$. Sometimes the choice of such paths is forced, as when $\mathscr{G}$ is a tree, but in general, it is still a matter of art to choose good paths. Let

γ be the length (number of edges) of the longest path,

$$b = \max_{e \in E} \left| \{ (x, y) \colon e \in \gamma_{xy} \} \right|.$$

The comparison bound may be formalized as follows.

THEOREM 1. *Let $\mathscr{G}$ be a connected graph on $\{1, 2, \dots, n\}$ with edge set E. Define a probability p on the symmetric group S_n by $p(id) = 1/n$, $p(i, j) =$*

$(n-1)/|E|n$ for $(i,j) \in E$ and $p(\pi)=0$ otherwise. Let

$$k = \{8|E|\gamma b/(n-1)+n\}(\log n + c), \qquad c>0.$$

Then there is a universal constant $\alpha > 0$ such that

$$\|p^{(k)} - u\|_{TV} \le \alpha e^{-c}.$$

PROOF. Let $\tilde{p}$ be the random transpositions measure corresponding to the complete graph. The comparison Theorem 1 of Section 3 gives $\tilde{\mathscr{E}} \le A\mathscr{E}$ where

$$A = \max_{e\in E} \frac{1}{p(e)} \sum_{y\in S_n} |y| N(e,y)\tilde{p}(y).$$

Here $p(e)=(n-1)/(n|E|)$, $\tilde{p}(y)=2/n^2$ for y a transposition. Any transposition (i,j) can be realized by transposing successive pairs corresponding to edges in γ_{ij}, starting at i, and then reversing all but the final transposition. This gives $|(i,j)| \le 2|\gamma_{ij}| \le 2\gamma$. A fixed edge e appears at most twice in such a series of moves, so $N(e,y) \le 2$. Using these bounds gives

$$A \le \frac{8|E|\gamma b}{n(n-1)}.$$

Further, all paths described above have odd length. Considering further *id* leads to

$$A_* \le \frac{8|E|\gamma b}{n(n-1)} + 1.$$

Using (2.10) and results of Diaconis and Shahshahani (1981) for random transpositions completes the proof. □

REMARK. The quantity γb of Theorem 1 can be replaced by $\Delta = \max_{e\in E} \sum_{\gamma_{xy}\ni e} |\gamma_{xy}| \le \gamma b$.

EXAMPLE 1. Let $\mathscr{G}$ be a "star" with $E=\{(1,j)\colon 2\le j\le n\}$. This corresponds to the random walk on S_n which transposes a random card with the first card. This walk has been treated using Fourier analysis by Flatto, Odlyzko and Wales (1985), Diaconis (1989) and Diaconis and Greene (1989). These authors show that $n\log n + cn$ is the right number of steps, the variation distance tending to zero for c large, and tending to one for c small. The geometric bounds give the right answer "up to constants": This graph is a tree, so paths are forced. Clearly $\gamma=2$, $b=2(n-1)$, $|E|=n-1$. This shows that for $k=33n(\log n + c)$, $\|p^{(k)}-u\|_{TV} \le \alpha e^{-c}$.

EXAMPLE 2. Let $\mathscr{G}$ be a "path" with $E=\{(i,i+1)\colon 1\le i\le n-1\}$. This corresponds to the random walk on S_n which begins with the cards in a row, picks a position $1\le i\le n-1$ at random and transposes the card there with the card to its right. This graph is a tree with $\gamma = n-1 = |E|$ and $b \le 2(n/2)^2$. Proposition 1 shows that if $k = n(4n^2+1)(\log n + c)$, $\|p^{(k)}-u\|_{TV} \le \alpha e^{-c}$.

A lower bound showing that the total variation distance is bounded away from 0 if $k = cn^3$, for c fixed, follows from considering a fixed card (say the card labelled 1). This performs a nearest neighbor random walk with steps occurring at rate $1/n$. As is well known, random walk takes $\gg n^2$ steps to get random on a path of length n, so this entails $\gg n^3$ steps. We conjecture that order $n^3 \log n$ steps is the correct answer for this problem.

EXAMPLE 3. Let $\mathscr{G}$ be a "double star": Take $n = 2m$ and $E = \{(1, m), (2, m), \dots, (m - 1, m), (m, m + 1), (m + 1, m + 2), \dots, (m + 1, n)\}$. This graph is a tree with $\gamma = 3$, $|E| = n - 1$, and $b = 2m^2$. Proposition 1 shows that for $k = n(12n + 1)(\log n + c)$,

$$\|p^{(k)} - u\|_{TV} \le \alpha e^{-c}.$$

A lower bound showing that order $n^2 \log n$ steps are required follows from the following rough argument. Consider vertices $\{1, 2, \dots, m\}$ as forming "urn 1" and vertices $\{m + 1, \dots, n\}$ as forming "urn 2." Transfer between the two urns occurs at rate $1/n$. The transfer process is essentially the Bernoulli Laplace process analyzed by Diaconis and Shahshahani (1987). Their results imply that $(n/4)(\log n + c)$ transfers must occur to ensure that the proportions in each urn are close to $1/2$. This shows that for $k = k(n) = (n^2/4)(\log n + c)$ with c fixed, $\liminf_{n\to\infty} \|p^{(k)} - u\|_{TV} > 0$.

EXAMPLE 4 (A two-dimensional shuffle). Consider n cards in an $l \times m$ grid. A random walk proceeds by picking a card at random and transposing it with one of its nearest neighbors. Observe that this walk has a dimensional aspect: If $l = 1$, it reduces to random transpositions on a "path" treated in Example 2. As shown below, the extra dimension speeds things up.

To write things out, identify the grid with the integer lattice points in the positive quadrant between $(0, 0)$ and $(l - 1, m - 1)$. The lattice points in the grid will be denoted $v = (x, y)$. Each edge on the graph gives a transposition. Let E be the associated set of transpositions. Thus $|E| = (l - 1)m + (m - 1)l$ transpositions are involved altogether. The measure p_1 described above is not uniform on E. Rather, for neighboring (v, v'),

$$p_1(v, v') = \begin{cases} 5/(6n), & \text{if } v \text{ is a corner and } v' \text{ is an edge cell or vice versa,} \\ 2/(3n), & \text{if } v \text{ and } v' \text{ are both edge cells,} \\ 7/(12n), & \text{if } v \text{ is an edge and } v' \text{ is an internal cell or vice versa,} \\ 1/(2n), & \text{if } v \text{ and } v' \text{ are internal.} \end{cases}$$

To avoid parity problems, let

$$p = \frac{1}{n}\delta_{id} + \left(1 - \frac{1}{n}\right)p_1. \tag{4.1}$$

THEOREM 2. *Let l and m be positive integers with $lm = n$. Define a measure p on S_n by* (4.1). *Let $k = (B + 1)n(\log n + c)$ with $B = 16(l + m)\max(l, m)$ and $c > 0$. Then, there is a universal constant $\alpha > 0$ such that*

$$\|p^{(k)} - u\|_{TV} \leq \alpha e^{-c}.$$

PROOF. The result essentially follows from Theorem 1. The measure at (4.1) is slightly different from the measure of Theorem 1, but we omit the details. □

REMARKS FOR THEOREM 2. (a) When $l = n$ and $m = 1$, this shows order $n^3 \log n$ steps are enough as in Example 2. For $l = m = \sqrt{n}$ it gives order $n^2 \log n$ steps. The two walks have a comparable number of generators (order n). This provides a sense in which there is "more freedom" in two dimensions.

(b) The technique of following a single card gives a lower bound showing that for fixed $c > 0$, $k = cn(\max(l, m))^2$ steps do not suffice. For $l = m = \sqrt{n}$, this shows order n^2 steps are not enough. We conjecture that order $n^2 \log n$ is the right answer.

(c) The argument can clearly be generalized to higher dimensions. In particular, take $n = 2^d$, and use the graph of the "cube." Theorem 1 shows that order $n(\log n)^2$ steps suffice to achieve randomness. Following a single card gives a lower bound of order $n(\log n)(\log\log n)$.

(d) Pemantle (1992) has used the techniques of the present paper to study a different two-dimensional shuffle in which random subrectangles of a grid of cards are rotated in place.

GENERAL REMARKS. (a) It is possible to show that for any tree on n vertices, $4(n - 1) \leq \gamma b \leq (n - 1)n^2/2$, the minimum occurring for a star, as in Example 1, the maximum occurring for a path, as in Example 2.

(b) Preliminary considerations indicate that a random tree has γb concentrated near $n^{5/2}$.

(c) For trees, a coupon collector's analysis of the number of fixed points shows that $k = (n/2)(\log n - c)$ steps can never be enough to drive the variation distance to zero.

(d) Trees exist for which the bound is of order $f(n)$, with $n \log n \leq f(n) \leq n^3$. Take a path of length m connected to a star of size $n - m$. These have $|E| = n - 1$, γ of order m, and b of order $m(n - m)$. By appropriate choice of m, essentially any bound occurs.

(e) There is another random walk classically associated with a graph $\mathscr{G}$. This has a single particle hopping around on the graph by choosing its nearest neighbor. Call this the classical walk on $\mathscr{G}$. There are several connections between the walks. For simplicity, suppose the graph is regular. Following a single card in the permutation walk gives a classical walk run at rate $2/n$. If K_c, K_π denote the smallest k such that variation distance is smaller than $1/e$ for the two walks, this shows $K_\pi \geq (n/2)K_c$.

Random Insertions. A different class of walks can be associated with a graph, and the same analysis applies. Let c_{ij} be the permutation in S_n resulting from taking card i and inserting it into position j. Thus for $i < j$, $c_{ij} = (j, j-1, \dots, i)$ and for $i > j$, $c_{ij} = c_{ji}^{-1}$. Given a connected graph, a walk can be performed by choosing an edge $\{i, j\}$ at random and performing c_{ij} or c_{ji} with probability $1/2$. For the complete graph this is essentially the random to random shuffle. For a star with vertex 1 at the center, this becomes random to top or top to random (with probability 1/2 each). Even a star with a different vertex at the center has resisted analysis. For a path, it becomes the nearest neighbor transposition walk analyzed in Example 2.

THEOREM 3. *Let $\mathscr{G}$ be a connected graph on $\{1, 2, \dots, n\}$ with edge set E. For c_{ij} defined above, let $q(id) = 1/n$, $q(c_{ij}) = (n-1)/2|E|n$ for $\{i, j\} \in E$ and $q = 0$ otherwise. With b and γ defined as in Theorem 1, let*

$$k = (8|E|\gamma b/(n-1) + n)(\log n + c)$$

with $c > 0$. Then, there is a universal $\alpha > 0$ such that

$$\|q^{(k)} - u\|_{TV} \le \alpha e^{-c}.$$

PROOF. The first step is to bound the rate of convergence (and Dirichlet form) for the random insertion process based on the complete graph. Here, a random card is removed and inserted in a random position. A straightfoward comparison with random transpositions shows that this process requires order $n \log n$ steps to achieve randomness. Bounds for more general graphs now follow by comparison with random insertions based on the complete graph: choosing paths, c_{ij} can be represented by a sequence of insertions along the paht from i to j. There is no need to "clean up" afterward. The stated bounds follow from these considerations. □

We have not investigated lower bounds except in a few instances where we found the results sharp "up to constants." Observe that c_{12} and c_{1n} generate S_n so the graph need not be connected. We have not investigated this direction.

B. *Overhand shuffles.* The second most popular way of mixing cards is the overhand shuffle in which one drops small packets of cards from hand to hand reversing the order of the packets. A realistic model of this shuffle was analyzed by Pemantle (1989) who showed that order $n^2 \log n$ of his shuffles suffice while order n^2 are not enough to achieve randomness. We here analyze two different models with neater shuffles. The results are somewhat surprising; neater shuffles mix cards faster.

Neat overhand shuffle. Let t_i be the permutation that reverses the top i cards in place. Thus

$$t_i = (1, i)(2, i-1)(3, i-3) \cdots .$$

Here t_1 is taken as the identity. Define

$$p(\pi) = \begin{cases} \frac{1}{n}, & \text{for } \pi = t_i,\ 1 \le i \le n \\ 0, & \text{otherwise}. \end{cases} \tag{4.2}$$

THEOREM 4. *For p defined by* (4.2) *let $k = 48n(\log n + c)$ for $c > 0$. Then there is a universal constant $\alpha > 0$ such that*

$$\|p^{(k)} - u\|_{TV} \le \alpha e^{-c}.$$

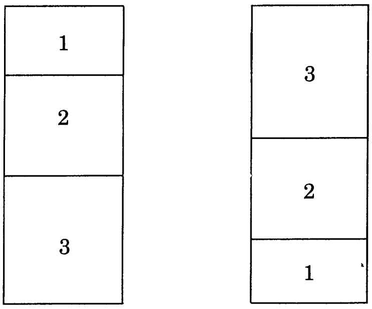
PROOF. Let us compare with $\tilde{p}$, the measure based on "transpose random with top" analyzed in Example 1. The permutation $t_{j-1}t_j$ (first perform t_j, then perform t_{j-1}) has the effect of bringing the top card to position j. It follows that $(1, i) = t_{i-1}t_{i-2}t_{i-1}t_i$ for $3 \le i \le n$ while $(1, 2) = t_2$. The quantity A becomes

$$A \le \max_i \sum_{j=1}^{n} 4N(t_i, (1, j)).$$

Since each t_i is involved in at most three transpositions $(1, j)$ and $N(t_i, (1, j)) \le 2$, $A \le 24$. The result now follows using Lemma 6 of Section 2 and known results for $\tilde{p}$. □

REMARK. A lower bound of order $n \log n$ follows by considering $T(\pi)$, the number of i such that $|\pi(i) - \pi(i+1)| = 1$. Under the uniform distribution, $T(\pi)$ is approximately Poisson(2) by standard arguments. If $k = n \log n - cn$ shuffles are performed for $c = c(n)$ tending to ∞ arbitrarily slowly, the coupon collector's problem shows $T(\pi) > 0$ with probability approaching 1.

Crude overhand shuffle. Here is a second simple model for an overhand shuffle. For $1 \le a \le b \le n$, let $t(a, b)$ move cards as in the following picture:



Thus the top a cards are cut off and placed on the table. Then a packet of size $b - a$ is cut off and placed on the original top packet. Finally, the remaining $n - b$ cards are placed on top. There are $\binom{n+1}{2}$ possible choices. These are made by choosing a uniformly in $1 \le a \le n$ and choosing b uniformly in

$a \le b \le n$. Thus

$$p(t(a,b)) = \frac{1}{n(n+1-a)}. \tag{4.3}$$

This implies $p(id) = 1/n$.

THEOREM 5. *For p defined by* (4.3), *let $k = 672n(\log n + c)$. Then there is a universal $\alpha > 0$ such that*

$$\|p^{(k)} - u\|_{TV} \le \alpha e^{-c}.$$

If $k = (n/2)(\log n + c)$, then

$$\|p^{(k)} - u\|_{TV} \ge \frac{1}{e^2} - e^{-e^{-c}} + o(1).$$

PROOF. The following elegant argument uses paths suggested by Pemantle. The argument is based on comparison with the measure associated to random transpositions $\tilde{p}(id) = 1/n$, $\tilde{p}(\pi) = 2/n^2$ if π is a transposition. Transpositions are represented by first representing "a to bottom" and "b to top" and their inverses. There are many ways to do this and the different choices must be taken in a balanced manner to get a good bound. The steps are easy but we find it helps to have a deck of cards on hand to check details.

The cycle "a to bottom" can be represented as

$$(n, n-1, \dots, a) = t(n-b+1, n-a+1)t(a-1, b) \qquad \text{for any } b,\ 2 \le a \le b \le n;$$

the cycle "b to top" can be represented as

$$(1, 2, \dots, b) = t(n-b, n-(a+1))t(a,b) \quad \text{for any } a,\ 1 \le a \le b \le n.$$

To avoid bottlenecks, transpositions (i,j) are split into 2 groups:

$$\text{Group I: } i + j \le n; \qquad \text{Group II: } i + j > n.$$

In Group I, for $1 \le i < j \le n$, the transposition (i,j) can be represented by the following:

(a) Move card i to bottom by $t(n-j+1, n-i+1)t(i-1, j)$.
(b) Move card i to position j by $t(i, n-j+1)t(j-1, n-i-1)$.
(c) Move card j to bottom by $t(i+1, n-j+2)t(j-2, n-i)$.
(d) Move card j to position i by $t(n-j-1, n-i+1)t(i-1, j)$.

Observe that when choice was possible in a representation, the second variable was used to make the choice. Clearly, in Group I, $|(i,j)| \le 8$ and no generator appears in more than seven transpositions. Further, $N(t(a,b),(i,j)) \le 3$. In Group II, for $i < j$, the transposition (i,j) can be represented by the following:

(a) Move card i to top by $t(n-i, j-1)t(n-j, i)$.
(b) Move card i to position j by $t(n-j, n-i+1)t(i,j)$.
(c) Move card j to top by $t(n-j+1, n-i-1)t(i, j-1)$.
(d) Move card j to position i by $t(n-i, j+1)t(n-j, i)$.

Again, in Group II, $|(i, j)| \leq 8$ and no generator appears in more than seven transpositions. Using these observations,

$$A \leq \max_a \frac{2n(n+1-a)}{n^2} 336 = 672.$$

Using this and the known results for random transpositions in Lemma 6 of Section 2 completes the proof of the upper bound.

For the lower bound, let $T(\pi) = |\{i\colon |\pi(i+1) - \pi(i)| = 1\}|$. Under the uniform distribution, $T(\pi)$ has a limiting Poisson(2) distribution. In particular, $u\{T(\pi) = 0\} = e^{-2} + o(1)$. On the other hand, each shuffle breaks at most two "bonds" where pairs $(1,2), (2,3) \cdots (n-1, n)$ are initially considered bonded. An easy variant of the coupon collector's problem shows that $p^{(k)}\{T(\pi) = 0\} \leq e^{-e^{-c}} + o(1)$ for $k = (1/2)n(\log n + c)$ with $c < 0$. This proves the lower bound. □

Both overhand shuffles analyzed above require order $n \log n$ repetitions. This is perhaps surprising in light of Pemantle's results: he analyzed a shuffle with many more underlying generators (order 2^n) yet found at least n^2 repetitions were needed.

C. Other shuffles. The next example solves a problem posed by Borel and Chéron [(1940), pages 8–10 and 254–256].

Borel shuffle. The basic step in this shuffle may be described as follows: Remove a random packet and place it on top. More precisely, for any a, b, $1 \leq a \leq b \leq n$, let π_{ab} be

$$\begin{matrix} 1 & 2 & \cdots & b-a+1 & b-a+2 & \cdots & b & b+1 & b+2 & \cdots & n \\ a & a+1 & \cdots & b & 1 & \cdots & a-1 & b+1 & b+2 & \cdots & n \end{matrix}$$

Let

$$(4.4) \qquad p(\pi) = \begin{cases} 1\Big/\binom{n+1}{2}, & \text{if } \pi = \pi_{ab} \text{ for some } a, b, \\ 0, & \text{otherwise.} \end{cases}$$

Observe $p(id) = 2/(n+1)$.

THEOREM 6. *Let p be defined by* (4.4). *Let $k = 16n(\log n + c)$ for $c > 0$. Then, there is a universal constant $\alpha > 0$ such that*

$$\|p^{(k)} - u\|_{TV} \leq \alpha e^{-c}.$$

If $k = (1/2)n(\log n + c)$ for $c < 0$,

$$\|p^{(k)} - u\|_{TV} \geq e^{-2} - e^{-e^{-c}} + o(1).$$

Thus, the variation distance does not tend to zero.

PROOF. The upper bound is achieved by comparison with the shuffle associated with removing a random card and inserting it at a random position. Here, for $1 \leq a, b \leq n$, let c_{ab} be the cycle $(b, b-1, \ldots, a)$ if $a \leq b$, and $(b, b+1, \dots, a)$ if $a \geq b$. Define $\tilde{p}(\pi) = 1/n^2$ if $\pi = c_{ab}$ for some a, b with $\tilde{p}(\pi) = 0$ otherwise. Theorem 3 of this section implies that if $m = 4n(\log n + c)$,

$$\|\tilde{p}^{(m)} - u\|_{TV}^2 \leq n!\|\tilde{p}^{(m)} - u\|_2^2 \leq \alpha e^{-c}. \tag{4.5}$$

For the comparison, write c_{ab} in terms of π_{cd} as

$$c_{ab} = \pi_{b-a+1, b-1}\pi_{a+1, b}, \quad \text{if } a < b,$$

$$c_{ab} = \pi_{a-b+1, a-1}\pi_{b+1, a}, \quad \text{if } a > b.$$

Here,

$$A = \max_{a, b} \frac{1}{p(\pi_{ab})} \sum_{\pi} |\pi| N(\pi_{ab}, \pi)\tilde{p}(\pi).$$

Clearly $|\pi| \leq 2$ and any fixed π_{ab} appears in the expression of at most two c_{cd}'s, whence $N(\pi_{ab}, \pi)$ is 1 for two different terms in the sum. These bounds give $A \leq 2(1 + 1/n)$. Using this, $\pi_{g-1} \geq -1 + 4/(n+1)$, and (4.5) in Lemma 6 of Section 2 proves the upper bound.

For the lower bound, take $T(\pi) = |\{i\colon |\pi(i+1) - \pi(i)| = 1\}|$. Under the uniform distribution, $T(\pi)$ has an approximate Poisson distribution with mean 2. In particular, $u\{T(\pi) = 0\} = e^{-2} + o(1)$. On the other hand, for $k = (1/2)n \log n + cn$, $p^{(k)}\{T(\pi) = 0\} = e^{-e^{-c}} + o(1)$. □

A rapidly mixing shuffle with small support. This example gives a probability supported on six permutations which achieves randomness extremely rapidly. The generators were developed by Babai, Hetyii, Kantor, Lubotzky and Seress (1990) for group-theoretic algorithms. They can be described as the two types of pêrfect shuffles of an even deck together with a single transposition. More precisely, suppose n is even. Let S_n act on the n set $X = \mathbb{Z}_{n-1} \cup \{\infty\}$. Let $\pi_0\colon x \mapsto 2x$ and $\pi_1\colon x \mapsto 2x + 1$ be two permutations of X (both fix ∞). Let $\pi_2 = (0, \infty)$. Take

$$E = \{id, \pi_0, \pi_0^{-1}, \pi_1, \pi_1^{-1}, \pi_2\}.$$

Define

$$p(\pi) = 1/6 \quad \text{if } \pi \in E \text{ and zero otherwise}. \tag{4.6}$$

THEOREM 7. *Let p be defined by* (4.6). *Let $k = 24n(\log n)^2(\log n + c)$ for $c > 0$. Then there is a universal constant $\alpha > 0$ such that*

$$\|p^{(k)} - u\|_{TV} \leq \alpha e^{-c}.$$

PROOF. Babai et al. (1990) show that for any fixed j, $0 \leq j \leq n-2$, there is a product of at most $\log n$ terms involving π_0 and π_1 that gives a permutation σ_j taking j to position 0. Thus, $\sigma_j^{-1}\pi_2\sigma_j = (j, \infty)$. Thus any transposition (j, ∞) can be written using at most $2\log n + 1$ generators.

Comparing with $\tilde{p}(id) = \tilde{P}(j, \infty) = 1/n, 0 \le j \le n-2$,

$$A = \max_{\pi \in E} \frac{1}{p(\pi)} \sum |(j,\infty)| N(\pi, (j,\infty)) \tilde{p}(j,\infty) \le 12(\log n + 1) \log n.$$

Now, use of (2.8) together with known results for $\tilde{p}$ give the result. □

REMARK. Any fixed number of generators require order $n \log n$ steps. Thus, up to logarithmic factors, this is as fast as possible. We do not know if part (or all) of the extra $\log^2$ is necessary.

5. Products. Random walk on the hypercube (Example 5 of Section 3) is an example of a natural walk on a product group. In this section we analyze two natural walks on the d-fold product of an arbitrary group. We were led to study them because they are natural problems where comparison techniques do not work well (see Example 1 below). Our analysis is based on elementary use of eigenvalues combined with Lemma 6 of Section 2 in which the corresponding continuous time process is crucial. No comparison argument is used. However, the walks we analyze here can be used to study other walks on products by comparison. We had been unable to get the results of this section (in particular, Example 2) by Fourier analysis or any other technique.

Let G_0 be a finite group with $|G_0| = g_0$. Let u_0 be the uniform probability on G_0: $u_0 \equiv g_0^{-1}$. For $d \ge 1$, let $G = G_0^d$ be the product of d copies of G_0. Also let $u = u_0^{\otimes d} \equiv g_0^{-d}$ be the uniform probability on G.

Given an arbitrary symmetric probability p_0 on G_0, consider the symmetric probability on G defined by

$$p = \frac{1}{d} \sum_{i=1}^{d} \underbrace{\delta_{id} \otimes \cdots \otimes}_{i-1} p_0 \otimes \underbrace{\cdots \otimes \delta_{id}}_{d-i},$$

when δ_{id} is point mass at id in G_0. The probability p has a simple interpretation: Pick a coordinate at random and put a random choice from p_0 in that coordinate.

Define

$$q = p_0 \otimes \cdots \otimes p_0 = p_0^{\otimes d}.$$

Let $\pi_0 = 1 \ge \pi_1 \ge \cdots \ge \pi_{g-1} \ge -1$ be the eigenvalues of p_0 and let φ_i be the eigenfunction associated with π_i. We make the following observation:

The eigenvalues of p are the g_0^d numbers

$$\pi_I = \frac{1}{d} \sum_{i \in I} \pi_i, \qquad I \in \{0, \ldots, g_0 - 1\}^d,$$

whereas the eigenvalues of q are the g_0^d numbers

$$\gamma_I = \prod_{i \in I} \pi_i, \qquad I \in \{0, \ldots, g_0 - 1\}^d.$$

Both π_I and γ_I are associated with the eigenfunction $\varphi_I(x_1, \dots, x_d) = \varphi_{i_1}(x_1) \cdots \varphi_{i_d}(x_d)$ where $I = (i_1, \dots, i_d)$.

A bound for q follows easily.

THEOREM 1. *Assume that* $g_0^{1/2}\|p_0^{(k)} - u_0\|_2 \le \alpha e^{-c}$ *for some* $\alpha > 0$ *and* $k = B_0(B_1 + c)$, *for all* $c > 0$. *Then, for* $K = B_0(B_1 + (1/2)\log d + c)$, $c > 0$, *we have*

$$2\|q^{(K)} - u\|_{TV} \le g^{1/2}\|q^{(K)} - u\|_2 \le \alpha e^{(\alpha^2/2)-c}.$$

PROOF. Using eigenvalues and the inequalities

$$(1 + x)^d - 1 = d\textstyle\int_0^x (1 + y)^{d-1}\,dy \le dx(1 + x)^{d-1} \le dx\,e^{dx},$$

we see that

$$\begin{aligned} g\|q^{(K)} - u\|_2^2 &= \left(1 + g_0\|p_0^{(K)} - u_0\|_2^2\right)^d - 1 \\ &\le dg_0\|p_0^{(K)} - u_0\|_2^2 e^{dg_0\|p_0^{(K)} - u_0\|_2^2}. \end{aligned}$$ □

In order to study p, consider first the semigroup kernels

$$h_{0,t} = e^{-t}\sum_0^\infty \frac{t^n}{n!}p_0^{(n)}$$

and

$$h_t = e^{-t}\sum_0^\infty \frac{t^n}{n!}p^{(n)}.$$

Observe that $h_t = h_{0,t/d} \otimes \cdots \otimes h_{0,t/d}$. Indeed, this can be checked on eigenfunctions. Note also that the smallest nonzero eigenvalue of $\delta_{id} - p$ is λ_1/d, where $\lambda_1 = 1 - \pi_1$ is the smallest non zero eigenvalue of $\delta_{id} - p_0$.

THEOREM 2. *Assume that* $g_0^{1/2}\|h_{0,t} - u_0\|_2 \le \alpha e^{-c}$ *for some constant* $\alpha > 0$ *and all* $t = B_0(B_1 + c)$, $c > 0$. *Then, for* $T = dB_0(B_1 + (1/2)\log d + c)$, $c > 0$, *we have*

$$2\|h_T - u\|_{TV} \le g^{1/2}\|h_T - u\|_2 \le e^{\alpha^2/2 - c}.$$

PROOF. Using the above observation and eigenvalues, write, for $T = s' + s$,

$$\|h_T - u\|_2^2 \le \|h_{s'}\|_2^2 e^{-2s\lambda_1/d} = \left(h_{0,2s'/d}(id)\right)^d e^{-2s\lambda_1/d}.$$

By hypothesis, we have $\lambda_1 \ge 1/B_0$ and $h_{0,2s'/d}(id) \le g_0^{-1}(1 + \alpha^2 e^{-2C})$, for $s'/d = B_0(B_1 + C)$ and any $C > 0$. Take $s' = dB_0(B_1 + (1/2)\log d)$, $s = dB_0c$. Putting this in the above estimate yields the desired result. □

THEOREM 3. *Assume that* $g_0^{1/2}\|p_0^{(k)} - u_0\|_2 \le \alpha e^{-c}$ *for some constant* $\alpha > 0$ *and* $k = B_0(B_1 + c)$, *for all* $c > 0$. *Then, for*

$$K = 2dB_0\big(\max\{B_1, B_0^{-1}\log(g_0^{1/2})\} + \log(d^{1/2}) + c\big) + 1, \qquad c > 0,$$

we have

$$2\|p^{(K)} - u\|_{TV} \le g^{1/2}\|p^{(K)} - u\|_2 \le \big(1 + 2e^{1+\alpha^2}\big)^{1/2} e^{-c}.$$

PROOF. Lemma 3 of Section 2 and the hypothesis implies that

$$\begin{aligned} g_0\|h_{0,2k} - u_0\|_2^2 &\le g_0 e^{-2k} + g_0\|p^{(k)} - u_0\|_2^2 \\ &\le g_0 e^{-2k} + \alpha^2 e^{-2k'} \end{aligned}$$

for $k = B_0(B_1 + k')$, $k' > 0$. Reasoning as in the proof of Theorem 2, we get

$$g\|h_{2t} - u\|_2^2 \le \big(1 + g_0 e^{-2k/d} + \alpha^2 e^{2k'/d}\big)^d e^{-2s/dB_0}$$

for $t = k + s$, $s > 0$. This gives

$$g\|h_{2K_1} - u\|_2^2 \le e^{1+\alpha^2-2c}$$

for $K_1 = dB_0(\max\{B_1, B_0^{-1}\log(g_0^{1/2})\} + \log d^{1/2} + c)$. Next, the smallest eigenvalue of p is $(1/d)\Sigma_{i=1}^{d}\pi_{g_0-1} = \pi_{g_0-1}$, which satisfies $|\pi_{g_0-1}| \le e^{-1/B_0}$. Now, the argument for Lemma 6 of Section 2 shows that

$$g\|p^{(N)} - u\|_2^2 \le \Big(\sum_{\pi_I>0} \pi_I^{2n'}\Big)\pi_{g-1}^{2n} + \sum_{0<\pi_I<1} \pi_I^{2N},$$

where $N = n + n' + 1$. Hence,

$$g\|p^{(N)} - u\|_2^2 \le \pi_{g-1}^{2n} + g\|h_{n'} - u\|_2^2\pi_{g-1}^{2n} + g\|h_N - u\|_2^2.$$

Using $\pi_{g-1} = \pi_{g_0-1}$ and K as defined gives

$$g\|p^{(K)} - u\|_2^2 \le e^{-2dc} + e^{1+\alpha^2-2dc} + e^{1+\alpha^2-2c}. \qquad \square$$

EXAMPLE 1. When p_0 is uniform on G_0 (i.e., $p_0 = u_0$), one finds that the eigenvalues of p are the numbers i/d with multiplicity $\binom{d}{i}(g_0 - 1)^{d-i}$, $i = 0, \dots, d$. In this case, it follows from direct consideration of eigenvalues that there is $\alpha > 0$ such that

$$g\|p^{(k)} - u\|_2^2 \le \alpha e^{-c}$$

where $K = (1/2)d\{\log(d(g_0 - 1)) + c\}$, $c > 0$. This is of the same order of magnitude as the value given by Theorem 3 in this case. Also,

$$g\|p^{(k)} - u\|_2^2 \ge d(g_0 - 1)\left(1 - \frac{1}{d}\right)^{2k}, \qquad k > 0.$$

Thus, order $d\log(dg_0)$ steps are necessary and sufficient to drive the d_2 distance to zero.

Note that, in this case, the random walk associated with p proceeds by choosing a coordinate at random and picking a random element of G_0. The first time that each coordinate has been chosen is a strong stationary time in the sense of Aldous and Diaconis (1986). Thus, the coupon collectors' bounds give universal $\alpha > 0$ such that

$$\|p^{(k)} - u\|_{TV} \le \alpha e^{-c} \quad \text{for } k = d(\log d + c).$$

Here total variation converges more quickly than d_2 distance when g_0 grows with d. This gives an example where the usual use of Cauchy–Schwarz is "off" for bounding total variation. [See Stong (1991) for more of this.]

Example 2. Take $G_0 = \mathbb{Z}_m$, $p_0(0) = P_0(\pm 1) = 1/3$. Here, the eigenvalues are known to be $1/3 + (2/3)\cos(2\pi j/m)$, $0 \le j \le m - 1$. This yields

$$m\|p_0^{(k)} - u_0\|_2^2 \le \alpha^2 e^{-2\beta k/m^2} \quad \text{when } k > \frac{m^2}{\beta},$$

for universal $\alpha, \beta > 0$. Theorem 3 yields

$$\|p^{(k)} - u\|_{TV} \le \tfrac{3}{2}e^{-c},$$

where $k = 2\beta^{-1}m^2 d[(1/2)\log d + \max(\alpha, \beta) + c]$, $c > 0$. This is sharp, up to constants.

Example 3. Take $G_0 = S_n$, $p_0(id) = 1/n$, $p_0((i, j)) = 2/n^2$ (random transpositions). It is known that $g_0\|p_0^{(k)} - u_0\|_2^2 \le \alpha^2 e^{-2c}$ for $k = (1/2)n(\log n + c)$. Theorem 3 yields

$$\|p^{(k)} - u\|_{TV} \le \sqrt{1 + 2e^{1+\alpha^2}}\, e^{-c}$$

when $k = (3/2)nd(\log(nd^{1/2}) + c) + 1$, $c > 0$.

Remarks. (a) Our original approach to bounding $\|p^{(k)} - u\|_{TV}$ used comparison with the version of p having p_0 replaced by u_0. We later realized that all the eigenvalues of p were available and could be used to get sharper results.

(b) The techniques and results of this section carry over to products of reversible Markov chains: Bounds on the rate of convergence of components give bounds on the rate of convergence of the product. See Diaconis and Saloff-Coste (1993).

(c) One can interpolate between p and q: For $1 \le j \le d$, define p_j on G_0^d by choosing a random subset of j indices out of $\{1, 2, \dots, d\}$, placing independent, identically distributed elements in these coordinates, and the identity in the remaining coordinates. Arguing as at the beginning of this section, for each $I \in \{0, 1, \dots, g_0 - 1\}^d$ there is an eigenvalue $\pi_I = 1/\binom{d}{j}\Sigma_s \Pi_{i=1}^j \pi_{i_{s_r}}$ with the sum over subsets $s = \{s_1 \cdots s_j\}$ of size j from the set $\{1, 2, \dots, d\}$ and the product over the coordinates of $I = (i_1, \dots, i_d)_a$ determined by this subset.

Acknowledgments. We thank Robin Pemantle for the argument for Theorem 5 of Section 4. We thank Jim Fill for extensive technical help and for having the patience of a saint.

REFERENCES

ALDOUS, D. (1987). On the Markov-chain simulation method for uniform combinatorial simulation and simulated annealing. *Prob. Engng. Info. Sci.* **1** 33–46.

ALDOUS, D. (1989). Hitting times for random walks on vertex transitive graphs. *Proc. Cambridge Philos. Soc.* **106** 179–191.

BABAI, L. (1990). Local expansion of vertex transitive graphs and random generation in finite groups. Technical report, Dept. Computer Science, Univ. Chicago.

BABAI, L. HETYII, G., KANTOR, W., LUBOTZKY, A. and SERESS, A. (1990). On the diameter of finite groups. In *Proc. 31st Ann. Symp. Foundations Computer Sci.* IEEE Computer Soc., Los Alamitos CA.

BABAI, L., KANTOR, W. and LUBOTZKY, A. (1992). Small diameter Cayley graphs for finite simple groups. *European J. Combin.* **10** 507–522.

BOREL, E. and CHÉRON, A. (1940). *Théorie Mathématique du Bridge à la Porté de Tous.* Gauthier-Villars, Paris.

DESAI and RAO (1991). On the convergence of reversible Markov chains. *SIAM Journal of Matrix Anal. Appl.*. To appear.

DIACONIS, P. (1988). *Group Representations in Probability and Statistics.* IMS, Hayward, CA.

DIACONIS, P. (1989). *Applications of Non-Commutative Fourier Analysis to Probability Problems. Lecture Notes in Math.* **366** 51–100. Springer, Berlin.

DIACONIS, P. and GREENE, C. (1989). An analysis of Murphy's elements. Technical report, Dept. Statistics, Stanford Univ.

DIACONIS, P. and SALOFF-COSTE, L. (1992). Moderate growth and random walk on finite groups. Technical report, Dept. Mathematics, Harvard Univ.

DIACONIS, P. and SALOFF-COSTE, L. (1993). Comparison theorems for reversible Markov chains. *Ann. Appl. Probab.* **3** 696–730.

DIACONIS, P. and SHAHSHAHANI, M. (1981). Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete* **57** 159–179.

DIACONIS, P. and SHAHSHAHANI, M. (1987). Time to reach stationarity in the Bernoulli–Laplace diffusion model. *SIAM J. Math. Anal.* **18** 208–218.

DIACONIS, P. and STROOCK, D. (1991). Geometric bounds for eigenvalues of Markov Chains. *Ann. Appl. Probab.* **1** 36–61.

DRISCOLL, J. and FURST, M. (1987). Computing short generator sequences. *Inform. and Comput.* **72** 117–132.

FILL, J. (1991). Eigenvalue bounds on convergence to stationarity for nonreversible Markov chains with an application to the exclusion processes. *Ann. Appl. Probab.* **1** 62–87.

FLATTO, L., ODLYZKO, A. and WALES, D. (1985). Random shuffles and group representations. *Ann. Probab.* **13** 154–178.

GANGOLLI, A. (1991). Convergence bounds for Markov chains and applications to sampling. Ph.D. dissertation, Dept. Computer Science, Stanford Univ.

HORN, R. and JOHNSON, C. (1985). *Matrix Analysis.* Cambridge Univ. Press.

MOHAR, B. (1989). Eigenvalues, diameter, and mean distance in graphs. Technical report, Dept. Mathematics, Univ. E. K. Ljubjanos, Jachansha, Yugoslavia.

PEMANTLE, R. (1989). An analysis of the overhand shuffle. *J. Theoret. Probab.* **2** 37–50.

PEMANTLE, R. (1991). Some two-dimensional shuffling processes. *Random Structures and Algorithms.* To appear.

STONG, R. (1991). Choosing a random spanning subtree: A case study. *J. Theoret. Probab.* **4** 753–766.

DEPARTMENT OF MATHEMATICS
HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS 02138

UNIVERSITÉ DE PARIS VI CNRS
ANALYSE COMPLEXE ET GÉOMÉTRIE
4 PLACE JUSSIEU-TOUR 46
75252 PARIS CEDEX 05
FRANCE