Z. Wahrscheinlichkeitstheorie verw. Gebiete
57, 159–179 (1981)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Generating a Random Permutation with Random Transpositions

Persi Diaconis and Mehrdad Shahshahani

Stanford University, Department of Statistics, Stanford, CA 94305, USA

1. Introduction

Imagine n cards, labled 1 through n, face up on a table. Suppose card 1 is at the left, card 2 is next, and so on, with card n at the right of the row. By a random transposition, we mean the following operation: Two integers L and R are chosen independently and uniformly between 1 and n (so $L=R$ with probability $1/n$). The cards labled L and R are transposed (if $L=R$ no transposition is made). If many transpositions are made; the row of cards will tend to appear in random arrangement. The problem is: How many transpositions are needed until the permutation is close to random?

More formally, a random transposition is modeled by a probability measure T on the symmetric group S_n:

$$\begin{aligned} &T(e)=1/n \quad \text{if } e \text{ is the identity} \\ &T(\tau)=2/n^2 \quad \text{if } \tau \text{ is a transposition} \\ &T(\cdot)=0 \quad \text{otherwise.} \end{aligned} \tag{1.1}$$

The result of k random transpositions is modeled by the convolution of T with itself k times. This will be denoted T^{*k}. The uniform distribution on S_n is denoted by U, so

$$U(\pi)=1/n! \quad \text{for all } \pi \in S_n. \tag{1.2}$$

As a measure of the distance between T^{*k} and U, we use variation distance:

$$\|T^{*k}-U\| = \sum_{\pi \in S_n} \left| T^{*k}(\pi) - \frac{1}{n!} \right| = 2 \max_{A \subset S_n} |T^{*k}(A) - U(A)|. \tag{1.3}$$

The main result implies that if k is larger than $\frac{1}{2} n \log n$, T^{*k} is close to uniform:

Theorem 1. *Assume* (1.1)–(1.3). *Let*

$$c = c(k, n) = \frac{k - \frac{1}{2} n \log n}{n}. \tag{1.4}$$

0044-3719/81/0057/0159/$04.20

Suppose that $n \geqq 10$ *and* $c > 0$. *Then there is a positive constant* b *such that*

$$\| T^{*k} - U \| \leqq b e^{-2c}.$$

Theorem 1 is proved in Sect. 3 using the tools of group representations. A better upper bound which permits explicit computation of the constant b is described at the end of Sect. 3 and in Remark 1 of Sect. 5. Some needed background and technical lemmas are in Sect. 2. In Sect. 4 we show that the group theoretic approach yields an explicit formula for the eigenvalues of the transition matrix of the associated markov chain.

A lower bound for the variation distance is given by

Theorem 2. *Assume* (1.1)-(1.4). *Then, for all* k, *as* n *tends to infinity*

$$\| T^{*k} - U \| \geqq 2 \left(\frac{1}{e} - e^{-e^{-2c}} \right) + o(1).$$

Theorem 2 is proved at the end of this section. The lower bound is useful if $c < 0$.

Remarks. The problem studied here arose from two independent sources. The first source involved computer generation of a random permutation. The usual algorithm for generating a random permutation goes as follows: Choose a random integer U_1 uniformly between 1 and n, then transpose U_1 and 1. At stage j, choose a random integer U_j uniformly between j and n, then transpose U_j and j. It is not hard to show that the distribution of the permutation at stage $n-1$ is exactly uniform. A proof and discussion is on pages 124–126 of Knuth (1969). The algorithm uses $n-1$ transpositions. One of us had a programmer who used the measure T^{*k} to generate random permutations. It is not hard to see that for $n > 2$, T^{*k} is never exactly uniform (c.f. Remark 3, Sect. 5). A discussion arose about how large k should be to make T^{*k} approximately uniform. Theorems 1 and 2 imply that k must be larger than $1/2 n \log n$.

A second source for this problem is a paper by David Aldous. Aldous (1980) gives bounds on the length of time that a Markov chain takes to approach its stationary distribution. For the chain arising in Theorems 1 and 2, he gave $c_1 n < k < c_2 n^2$ and conjectured that $k = O(n \log n)$ was the right number of transpositions. Further remarks and references are in Sect. 4.

We finish this section by proving the lower bound given in Theorem 2. The argument, due essentially to Charles Stein, is useful in producing a set $A \subset S_n$ where the maximum difference between T^{*k} and U is large. The set A consists of all permutations with one or more fixed points.

Proof of Theorem 2. We get a lower bound on the variation distance between T^{*k} and the uniform distribution U by considering the set A of all permutations with one or more fixed points. Under U, the chance of one or more fixed points is well known under the name of the matching problem. The results of Sect. IV.4 of Feller (1968) imply

$$U(A) = 1 - 1/e + O(1/n!). \tag{1.5}$$

To bound $T^{*k}(A)$, consider the process for generating T^{*k} described in the introduction. This was based on making random transpositions $(L_1, R_1)\ldots(L_k, R_k)$. Let B be the event that the set of lables $\{L_i, R_i\}_{i=1}^k$ is strictly smaller than $\{1, 2, 3, \ldots, n\}$. Clearly $A \supset B$. The probability of B is the same as the probability that if $2k$ balls are dropped at random into n boxes, one or more of the boxes will be empty. The argument for Theorem 3 of Sect. IV.2 of Feller (1968) implies that the probability of B equals

$$1 - e^{-ne^{-2k/n}} + o(1) \quad \text{uniformly in } k, \text{ as } n \to \infty.$$

Using the definition of c at (1.4) we have

$$T^{*k}(A) \geqq 1 - e^{-e^{-2c}} + o(1). \tag{1.6}$$

Using (1.5), (1.6), and the definition of the variation distance at (1.3),

$$\begin{aligned} \|T^{*k} - U\| &\geqq 2|T^{*k}(A) - U(A)| \geqq 2(T^{*k}(A) - U(A)) \\ &\geqq 2\left(\frac{1}{e} - e^{-e^{-2c}}\right) + o(1). \quad \square \end{aligned}$$

Acknowledgement. We thank David Aldous, Joseph Deken, Richard Durrett, Leo Flatto, Ed Gilbert, Larry Shepp, Charles Stein, and Sandy Zabell for helpful discussions.

2. Preliminaries

Theorem 1 is proved by considering the problem as a random walk on the permutation group and using the analogue of Fourier analysis - representation theory. A slick introduction to linear representations of finite groups is Serre (1977). We use this reference whenever possible. A readable comprehensive treatment of representation theory is Curtis and Reiner (1962). A recent monograph on representations of the symmetric group is James (1978).

Let P_1 and P_2 be functions on a finite group G. We write $P_2 * P_1$ for their convolution. Thus for $\gamma \in G$,

$$P_2 * P_1(\gamma) = \sum_{\eta \in G} P_2(\gamma \eta^{-1}) P_1(\eta).$$

A *representation* ρ of G is a homomorphism of G into the group of invertible linear maps of a complex finite dimensional vector space V. We write d_ρ for the dimension of V and think of $\rho(\gamma)$ as a $d_\rho \times d_\rho$ matrix. The representation ρ is *irreducible* if V admits no $\rho(G)$ invariant subspaces other than $\{0\}$ or V. Two representations ρ and π are *equivalent* if there is a matrix M such that $M\rho(\gamma)M^{-1} = \pi(\gamma)$ for all $\gamma \in G$. If P is a function on G and ρ is a representation, define

$$\rho(P) = \sum_{\eta \in G} P(\eta)\rho(\eta).$$

The transform $\rho(P)$ is the analog of the Fourier transform. It converts convolution into multiplication:

Lemma 1. *If P_1 and P_2 are two functions on G and ρ is a representation, then*

$$\rho(P_2 * P_1) = \rho(P_2)\,\rho(P_1).$$

Proof.

$$\begin{aligned}\rho(P_2 * P_1) &= \sum_\gamma \rho(\gamma)\, P_2 * P_1(\gamma) = \sum_\gamma \rho(\gamma) \sum_\eta P_2(\gamma\eta^{-1})\, P_1(\eta) \\ &= \sum_\eta \{\sum_\gamma \rho(\gamma\eta^{-1})\, P_2(\gamma\eta^{-1})\}\, \rho(\eta)\, P_1(\eta) = \{\sum_\gamma \rho(\gamma)\, P_2(\gamma)\}\{\sum_\eta \rho(\eta)\, P_1(\eta)\} \\ &= \rho(P_2)\,\rho(P_1). \quad \square\end{aligned}$$

Let $\hat{G}$ be the set of irreducible representations of G. We sometimes regard $\rho(P)$ as a matrix valued function from $\hat{G}$. The next two results show how knowing $\rho(P)$ for $\rho * \hat{G}$ gives information about P. We write $|G|$ for the number of elements in G, $\mathrm{Tr}[\cdot]$ for trace, and $*$ for complex conjugate transpose.

Lemma 2. *Plancherel formula.*

$$\sum_{\eta \in G} |P(\eta)|^2 = \frac{1}{|G|} \sum_{\rho \in \hat{G}} d_\rho\, \mathrm{Tr}[\rho(P)\,\rho(P)^*]. \tag{2.1}$$

Inversion formula.

$$P(\eta) = \frac{1}{|G|} \sum_{\rho \in \hat{G}} d_\rho\, \mathrm{Tr}[\rho(\eta)^*\, \rho(P)]. \tag{2.2}$$

Proofs of (2.1) and (2.2) are in Section (6.2) of Serre (1977) or Theorem (28.43) of Hewitt and Ross (1970). For a finite group G, $\hat{G}$ is also finite. More precisely, Corollary 2 of Sect. 2.4 in Serre (1977) gives

Lemma 3.

$$\sum_{\rho \in \hat{G}} d_\rho^2 = |G|.$$

In particular, if $G = S_n$, the symmetric group on n letters,

$$\sum_{\rho \in \hat{S}_n} d_\rho^2 = n! \tag{2.3}$$

Therefore, for every irreducible representation ρ of S_n

$$d_\rho < \sqrt{n!}. \tag{2.4}$$

The following useful characterization of an irreducible representation is proved in Sect. 2.2 of Serre (1977).

Lemma 4. *Schur's lemma. A representation ρ of G is irreducible if and only if every $d_\rho \times d_\rho$ matrix M that satisfies $M\rho(\gamma) = \rho(\gamma)\, M$ for all $\gamma \in G$ is a constant multiple of the identity.*

Two elements γ, η in G are *conjugate* if there is an $\alpha \in G$ such that $\alpha\gamma\alpha^{-1} = \eta$. This is an equivalence relation. Equivalence classes are called *conjugacy*

classes. The character χ_ρ of the representation ρ is the function from G into the complex numbers:

$$\chi_\rho(\gamma) = \mathrm{Tr}[\rho(\gamma)].$$

It follows from the properties of the trace map that equivalent representations have identical characters and that characters are constant on conjugacy classes. Section 2.3 of Serre (1977) shows that a representation ρ is determined by its character. The following consequence of Schur's lemma is crucial in the proof of Theorem 1.

Lemma 5. *Let G be a finite group. Let ρ be an irreducible representation of G. Let P be a function from G into the complex numbers which is constant on each conjugacy class. On the i-th conjugacy class, let P_i be the value of P, n_i the cardinality of the i-th conjugacy class, and χ_i the value of $\chi_\rho(\cdot)$. Then*

$$\rho(P) = CI \quad \text{where } C = \frac{1}{d_\rho} \sum_i P_i n_i \chi_i. \tag{2.5}$$

The sum in (2.5) *is over distinct conjugacy classes.*

Proof. Let M_i denote the sum of $\rho(\eta)$ as η ranges over the i-th conjugacy class. By hypothesis,

$$\rho(P) = \sum_\gamma P(\gamma)\,\rho(\gamma) = \sum_i P_i M_i.$$

The matrix M_i satisfies $\rho(\pi) M_i \rho(\pi^{-1}) = M_i$ for all $\pi \in G_n$. Thus Schur's lemma gives $M_i = C_i I$ for some real number C_i. Taking traces,

$$\mathrm{Tr}(M_i) = n_i \chi_i = C_i d_\rho.$$

This proves (2.5). □

Corollary 1. *Let T be the probability measure on S_n defined by* (1.1). *Let ρ be a representation of S_n. Write $\chi_\rho(\tau)$ for the character of ρ at any transposition τ. Then*

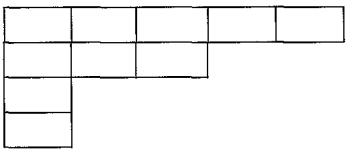
$$\rho(T) = \left(\frac{1}{n} + \frac{n-1}{n} \frac{\chi_\rho(\tau)}{d_\rho}\right) I. \tag{2.6}$$

Proof. T is constant on conjugacy classes, putting mass $1/n$ at the identity, and $2/n^2$ on each transposition. There are $\binom{n}{2}$ transpositions and the character of ρ at the identity is d_ρ. The result follows from Lemma 5. □

Remark. Corollary 1 and Lemma 1 reduce the problem of approximating $\rho(T^{*k})$ to a problem about approximating a product of real numbers.

Representation theory of the symmetric group S_n has been a subject of intensive study by Frobenius, Young, and others. We next summarize the facts we need. By a partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)$ of n we mean a sequence, $\lambda_1 \geqq \lambda_2 \geqq \ldots \geqq \lambda_m$ of positive integers with $n = \lambda_1 + \ldots + \lambda_m$. There is a one-to-one correspondence between irreducible representations of S_n and partitions of n. While the exact construction is not needed, the notion of Young diagram will

be useful. To every partition λ there corresponds a *Young diagram.* The first row of the diagram contains λ_1 squares, the second row contains λ_2 squares, and so on. For example, the diagram corresponding to the partition (5, 3, 1, 1) of 10 is



The following is well known; see Corollary 8.5 of James (1978).

Lemma 6. *The dimension d_λ of the representation corresponding to the partition λ is the number of ways of placing the numbers $1, 2, \ldots, n$ into the Young diagram of λ such that the entries in each row and column are increasing.*

Combining Lemma 6 and (2.4) we get

Corollary 2. *The dimension d_λ of the representation corresponding to the partition λ satisfies*

$$d_\lambda \leqq \binom{n}{\lambda_1} \sqrt{(n-\lambda_1)!}. \tag{2.7}$$

Proof. The first row of any of the arrangements of Lemma 6 can be chosen in at most $\binom{n}{\lambda_1}$ ways. For each choice of first row, the number of ways of choosing the remaining rows is smaller than the largest dimension of an irreducible representation of $S_{n-\lambda_1}$. By (2.4) this is at most $\sqrt{(n-\lambda_1)!}$. □

The conjugacy classes in S_n are also in one-to-one correspondence with partitions of n. The partition λ corresponds to the conjugacy class of all permutations with cyclic decomposition

$$(n_1, n_2, \ldots, n_{\lambda_1})(n_{\lambda_1+1} \ldots n_{\lambda_1+\lambda_2}) \ldots (n_{\lambda_1+\ldots+\lambda_{m-1}+1} \ldots n_n)$$

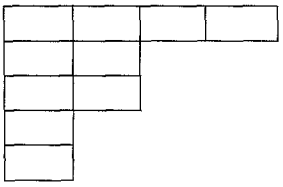
where $(n_1, \ldots, n_n)$ is a permutation of $(1, \ldots, n)$. In particular, transpositions form a single conjugacy class with $\binom{n}{2}$ elements.

The value of the characters of S_n are integers. Formulas for the characters at irreducible representations were obtained by Frobenius. These formulas are given in modern notation by MacDonald (1979). Ingham (1960) contains an accessible proof of the following special case.

Lemma 7. *The character of the irreducible representation of S_n corresponding to the partition λ, evaluated at a transposition τ satisfies:*

$$\frac{\chi_\lambda(\tau)}{d_\lambda} = \frac{1}{n(n-1)} \sum_{j=1}^{m} [(\lambda_j - j)(\lambda_j - j + 1) - j(j-1)]. \tag{2.8}$$

We need the notion of the conjugate of a representation. If λ is a partition, the Young diagram of the conjugate partition λ' is the transpose of the Young diagram of λ. Thus, if $\lambda=(5,3,1,1)$, $\lambda'=(4,2,2,1,1)$ has diagram



Two representations are *conjugate* if the corresponding partitions are conjugate.

Lemma 8. *Let ρ and ρ' be conjugate irreducible representations. Then*

$$d_\rho = d_{\rho'} \tag{2.9}$$

and, for any transposition τ

$$\chi_\rho(\tau) = -\chi_{\rho'}(\tau). \tag{2.10}$$

Proof. Lemma 6 implies (2.9) while (2.10) is a special case of results in Sect. (6.6) of James (1978). □

Because it appears in (2.6), the ratio $\chi_\rho(\tau)/d_\rho$ will be in constant use in what follows. Define

$$r(\rho) = \chi_\rho(\tau)/d_\rho. \tag{2.11}$$

We will use Frobenius' formula (2.8) to prove a monotonicity property of $r(\rho)$. Toward this end we introduce a partial order on partitions of n. Let $\lambda = (\lambda_1, \ldots, \lambda_m)$ and $\lambda' = (\lambda'_1, \ldots, \lambda'_{m'})$ be partitions of n. Define $\lambda \trianglerighteq \lambda'$ if $m \leqq m'$ and $\lambda_1 \geqq \lambda'_1$, $\lambda_1 + \lambda_2 \geqq \lambda'_1 + \lambda'_2, \ldots, \lambda_1 + \ldots + \lambda_m \geqq \lambda'_1 + \ldots + \lambda'_m$. This partial ordering is used in James (1978). An extensive discussion of the ordering is in Marshall and Olkin (1979). We need the following characterization of the ordering. For two partitions λ, λ' of n, say that λ is obtained from λ' by a single switch if for some indices $a<b$,

$$\lambda_j = \lambda'_j \quad \text{for } j \neq a \text{ or } b$$

$$\lambda_a = \lambda'_a + 1 \quad \text{and} \quad \lambda_b = \lambda'_b - 1.$$

Lemma 9. *Let λ and λ' be partitions of n. Then $\lambda \trianglerighteq \lambda'$ if and only if there is a finite sequence of partitions of n:*

$$\lambda = \lambda^0 \trianglerighteq \lambda^1 \trianglerighteq \lambda^2 \ldots \trianglerighteq \lambda^j = \lambda',$$

such that λ^i is obtained from λ^{i+1} by a single switch for all i.

Proof. This is a restatement of a basic result due to Muirhead on majorization in integers. It is proved in Sect. 5D of Marshall and Olkin (1979). □

The basic monotonically result can now be stated.

Lemma 10. *Let ρ and ρ' be irreducible representations of S_n corresponding to partitions λ and λ'. If $\lambda \trianglerighteq \lambda'$, then $r(\rho) \geqq r(\rho')$.*

Proof. Because of Lemma 9, we need only consider situations where λ is obtained from λ' by a single switch. The proof is then a straightforward computation from Frobenius' formula (2.8). Suppose that the switch involves indices $a<b$. There are two cases:

$$\text{Case 1} \quad \lambda'_b=1$$
$$\text{Case 2} \quad \lambda'_b>1.$$

In Case 1, formula (2.8) shows that $r(\rho)-r(\rho')$ equals $1/n(n-1)$ times

$$\begin{aligned}&\{(\lambda_a+1-a)(\lambda_a+1-(a-1))-a(a-1)\}\\&\quad-\{(\lambda_a-a)(\lambda_a-(a-1))-a(a-1)+(1-b)(1-(b-1))-b(b-1)\}\\&\quad=\{(\lambda_a+1)^2-(\lambda_a+1)(2a-1)\}-\{\lambda_a^2-\lambda_a(2a-1)+2(1-b)\}\\&\quad=(2\lambda_a+1)-(2a-1)+2(b-1)\\&\quad=2(\lambda_a+b-a)\geqq 4>0.\end{aligned}$$

In case 2, a similar computation shows that $r(\rho)-r(\rho')$ equals $1/n(n-1)$ times

$$2\{(\lambda_a-\lambda_b)+(b-a)\}\geqq 2>0. \quad \square$$

We make use of Lemmas 7–10 to get bounds on $r(\rho)$.

Lemma 11. *Let $(\lambda_1, \ldots, \lambda_m)$ be a partition of n satisfying*

$$\lambda_1\leqq n/3.$$

Let ρ be the corresponding irreducible representation. Then

$$r(\rho)\leqq\frac{1}{n(n-1)}\left\{\frac{n^2}{3}+\frac{20}{3}-3n\right\}.$$

Proof. Let b denote the smallest integer greater than or equal to $n/3$.

$$b=\frac{n}{3}+\varepsilon \quad \text{where } 0\leqq\varepsilon\leqq\tfrac{2}{3}.$$

Let $\lambda'=(b, b, n-2b)$ with ρ' the corresponding representation. Clearly $\lambda'\trianglerighteq\lambda$, and so $r(\rho')\geqq r(\rho)$. Using formula (2.8), $r(\rho')$ equals $1/n(n-1)$ times

$$b(b-1)+(b-1)(b-2)-2+(n-2b-2)(n-2b-3)-6.$$

Straightforward algebra shows that the last displayed expression equals

$$\frac{n^2}{3}+6(\varepsilon^2+\varepsilon)-3n.$$

Putting $\varepsilon=2/3$ gives the asserted upper bound. $\square$

Lemma 12. *Let $(\lambda_1, \ldots, \lambda_m)$ be a partition of n satisfying*

$$\lambda_1\leqq\frac{n}{2}.$$

Let ρ be the corresponding irreducible representation. Then

$$r(\rho) \leqq \frac{1}{n(n-1)} \left\{ \frac{n^2}{2} - n \right\}.$$

Proof. Let b denote the smallest integer greater than or equal to $n/2$.

$$b = \frac{n}{2} + \varepsilon \quad \text{where } 0 \leqq \varepsilon \leqq \tfrac{1}{2}.$$

Let $\lambda' = (b, n-b)$ with ρ' the corresponding irreducible representation. Clearly $\lambda' \trianglerighteq \lambda$ and so $r(\rho') \geqq r(\rho)$. Using formula (2.8), $r(\rho')$ equals $1/n(n-1)$ times

$$b(b-1) + (n-b-1)(n-b-2) - 2.$$

Straightforward algebra shows that the last displayed expression equals

$$\frac{n^2}{2} - 2n + 2\varepsilon + 2\varepsilon^2 < \frac{n^2}{2} - n. \quad \square$$

Lemma 13. *Let $(\lambda_1, \ldots, \lambda_m)$ be a partition of n satisfying*

$$\lambda_1 > \frac{n}{2}.$$

Let ρ be the corresponding representation. Then

$$0 < r(\rho) \leqq \frac{1}{n(n-1)} \{\lambda_1(\lambda_1 - 1) + (n - \lambda_1 - 1)(n - \lambda_1 - 2) - 2\}.$$

Proof. Let ρ' be the irreducible representation corresponding to $(\lambda_1, n - \lambda_1)$. Clearly $\lambda' \trianglerighteq \lambda$ and so $r(\rho') \geqq r(l)$. Now use formuly (2.8) to get the asserted upper bound. For the lower bound, consider ρ'' the irreducible representation corresponding to $\lambda'' = (\lambda_1, 1, 1, \ldots, 1)$. Clearly $\lambda \trianglerighteq \lambda''$, so $r(\rho'') \leqq r(\rho)$. Now formula (2.8) shows that $r(\rho'')$ equals $1/n(n-1)$ times

$$\begin{aligned} &\lambda_1(\lambda_1 - 1) + \sum_{j=2}^{n-\lambda_1} \{(1-j)(2-j) - j(j-1)\} \\ &\quad = \lambda_1(\lambda_1 - 1) + 2(n - \lambda_1) - (n - \lambda_1)(n - \lambda_1 + 1) + 2. \end{aligned}$$

The last displayed expression is positive because

$$\lambda_1 \geqq (n - \lambda_1 + 1) \quad \text{and} \quad \lambda_1 - 1 \geqq n - \lambda_1. \quad \square$$

The proof of Theorem 1 requires a bound on the number of partitions of n. Chapter 3 of Ayoub (1963) contains a detailed discussion of the asymptotic behavior of the partition function. We only require the following bound, given as Eq. (2.1) in Chap. 3, Sect. 1 of Ayoub (1963).

Let (n) be the number of partitions of n. Then

$$p(n) \leqq \exp\left[\pi \left(\frac{2n}{3}\right)^{\frac{1}{2}}\right]. \tag{2.12}$$

3. Proof of Theorem 1

We first derive a convenient bound to the variation distance by using the Plancherel formula.

Lemma 14. *Assume* (1.1)–(1.3), *then*

$$\|T^{*k}-U\| \leqq \left\{ \sum_{\rho}' d_\rho^2 \left(\frac{1}{n} + \frac{n-1}{n} r(\rho) \right)^{2k} \right\}^{\frac{1}{2}}. \tag{3.1}$$

The sum in (3.1) *is over non-trivial irreducible representations* ρ.

Proof. The Cauchy-Schwarz inequality yields

$$\|T^{*k}-U\| = \sum_{\pi} \left| T^{*k}(\pi) - \frac{1}{n!} \right| \leqq \left\{ n! \sum_{\pi} \left(T^{*k}(\pi) - \frac{1}{n!} \right)^2 \right\}^{\frac{1}{2}}.$$

The Plancherel formula (2.1) gives

$$n! \sum_{\pi} \left(T^{*k}(\pi) - \frac{1}{n!} \right)^2 = \sum_{\rho} d_\rho \operatorname{Tr}\{\rho^2 (T^{*k}-U)\},$$

the right-hand sum is over all irreducible representations ρ. At the trivial representation, $\rho(\pi) \equiv 1$; $\rho(T^{*k}) = \rho(U) = 1$ so $\rho(T^{*k}-U) = 0$. For any non-trivial irreducible representation, Schur's lemma implies $\rho(U) = 0$. Lemma 1 and Corollary 1 yield

$$\rho(T^{*k}) = \left(\frac{1}{n} + \frac{n-1}{n} r(\rho) \right)^k I.$$

Hence, for any non-trivial irreducible representation ρ,

$$\operatorname{Tr}(\rho^2 (T^{*k}-U)) = d_\rho \left(\frac{1}{n} + \frac{n-1}{n} r(\rho) \right)^{2k}.$$

This completes the proof of (3.1). □

The proof of Theorem 1 proceeds by bounding the sum on the right side of (3.1). The argument involves several zones defined using the two parameters λ_1 and m of the partitions $(\lambda_1, \ldots, \lambda_m)$.

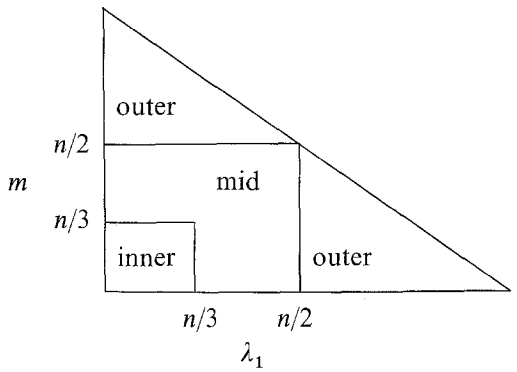
Inner zone. $\lambda_1 \leqq \frac{n}{3}$ and $m \leqq \frac{n}{3}$

Mid zone. $\left[\frac{n}{3} < \lambda_1 \leqq \frac{n}{2} \text{ and } m \leqq \frac{n}{2} \right]$

or $\left[\frac{n}{3} < m \leqq \frac{n}{2} \text{ and } \lambda_1 \leqq \frac{n}{2} \right]$

Outer zone. $\lambda_1 > \frac{n}{2}$ or $m > \frac{n}{2}$.

Since $2 \leqq \lambda_1 + m \leqq n+1$, these zones can be pictured as parts of a triangle:



Bounds for the Inner Zone. Note that the inner zone is empty unless $n \geqq 9$. Lemma 11 implies that for ρ in the inner zone,

$$\frac{1}{n} + \frac{n-1}{n} r(\rho) \leqq \frac{1}{3} + \frac{1}{n^2}\frac{20}{3} - \frac{2}{n} < \frac{1}{3}.$$

Using Lemma 8 and Lemma 11, for any ρ in the inner zone,

$$\frac{1}{n} + \frac{n-1}{n} r(\rho) \geqq -\frac{1}{3} - \frac{1}{n^2}\frac{20}{3} + \frac{4}{n} > -\frac{1}{3}.$$

Now use (2.3) to argue that the sum over the inner zone of

$$d_\rho^2 \left| \frac{1}{n} + \frac{n-1}{n} r(\rho) \right|^{2k}$$

is at most

$$(\tfrac{1}{3})^{2k} n!. \tag{3.2}$$

This error term tends to zero exponentially fast if $k \geqq \frac{1}{2} n \log n$. We discuss this carefully at the end of this section.

Bounds for the Mid Zone. Using Lemmas 8, 10, and 12 shows that for ρ in the mid zone

$$\left| \frac{1}{n} + \frac{n-1}{n} r(\rho) \right| \leqq \tfrac{1}{2}.$$

Next use (2.7); for ρ in the mid zone, suppose $n/3 < \lambda_1 \leqq n/2$.

$$d_\rho \leqq \binom{n}{\lambda_1} \sqrt{(n-\lambda_1)!} < 2^n \sqrt{(n-\lambda_1)!} \leqq 2^n n^{n/3}.$$

Because of (2.9), this same bound holds if ρ satisfies $n/3 < m \leqq n/2$. Combining bounds, we see that for ρ in the mid zone

$$d_\rho^2 \left| \frac{1}{n} + \frac{n-1}{n} r(\rho) \right|^{2k} \leqq 4^n (\tfrac{1}{2})^{2k} n^{2n/3}.$$

Using (2.12) to bound the number of partitions, the error from the mid zone is at most

$$e^{\pi(2n/3)^{\frac{1}{2}}} 4^n (\tfrac{1}{2})^{2k} n^{2n/3}. \tag{3.3}$$

Again, this error tends to zero exponentially fast if $k \geqq \frac{1}{2} n \log n$.

Bounds for the Outer Zone. The argument for the outer zone is complicated. We proceed by breaking the outer zone into three disjoint zones.

$$\text{Zone I.} \quad \frac{n}{2} < \lambda_1 \leqq .7n \quad \text{or} \quad \frac{n}{2} < m \leqq .7n$$

$$\text{Zone II.} \quad .7n < m \leqq n$$

$$\text{Zone III.} \quad .7n < \lambda_1 \leqq n-1.$$

The representation corresponding to $\lambda_1 = n$ is the trivial representation which does not appear in the sum (3.1).

Bounds for Zone I. We first argue that for all ρ in Zone I,

$$\left| \frac{1}{n} + \frac{n-1}{n} r(\rho) \right| \leqq .58 + \frac{.2}{n} + \frac{4}{n^2}. \tag{3.4}$$

To show (3.4), first suppose that ρ satisfies $n/2 < \lambda_1 \leqq .7n$. Let b denote the smallest integer larger than $.7n$, so

$$b = .7n + \varepsilon \quad \text{where} \quad 0 \leqq \varepsilon \leqq \tfrac{9}{10}.$$

Let $\lambda' = (b, n-b)$ with ρ' the associated irreducible representation. Since $\lambda' \trianglerighteq \lambda$, we have $r(\rho') \geqq r(\rho)$. Formula (2.8) gives $r(\rho')$ as $1/n(n-1)$ times

$$b(b-1) + (n-b-1)(n-b-2) - 2.$$

The last displayed expression equals

$$(.7n)^2 + (.3n)^2 + 2(\varepsilon^2 + \varepsilon) + n(.8\varepsilon - 1.6).$$

This last expression can be bounded above by setting $\varepsilon = 1$, giving

$$n^2 \left[.58 - \frac{.8}{n} + \frac{4}{n^2} \right].$$

Since Lemma 13 implies $r(\rho)$ is positive, the inequality (3.4) follows.

For ρ in Zone I with $n/2 < m \leqq .7n$, Lemmas 8 and 13 imply that (3.4) holds.

Next, Corollary 2 shows that for ρ in Zone I satisfying $n/2 < \lambda_1 \leqq .7n$,

$$d_\rho \leqq \binom{n}{\lambda_1} \sqrt{(n-\lambda_1)!} \leqq 2^n \sqrt{[n/2]!}$$

with $[.]$ denoting greatest integer. The same bound holds if $n/2 < m \leqq .7n$ because of Lemma 8. Combining bounds and using (2.12) to bound the number

of partitions, the error from Zone I is at most

$$e^{\pi(\frac{2}{3}n)^{\frac{1}{2}}} 4^n \left[\frac{n}{2}\right]! \left(.58+\frac{.2}{n}+\frac{4}{n^2}\right)^{2k}. \tag{3.5}$$

Again, this error term tends to zero exponentially fast when $k \geqq \frac{1}{2} n \log n$.

Bounds for Zone II. For $.7n < m \leqq n$, Lemmas 8 and 13 imply that $r(\rho)$ is negative and bounded below by $-1/n(n-1)$ times

$$\begin{aligned} &m(m-1)+(n-m-1)(n-m-2)-2 \\ &\quad = n^2+2m^2-2mn+2m-3n. \end{aligned}$$

Suppose first that $r(\rho) < -1/n$. Then, for $n \geqq 4$,

$$\begin{aligned} \left|\frac{1}{n}+\frac{n-1}{n} r(\rho)\right| &\leqq 1-2\frac{m}{n}+2\left(\frac{m}{n}\right)^2+2\frac{m}{n^2}-\frac{4}{n} \\ &\leqq 1-2\frac{m}{n}+2\left(\frac{m}{n}\right)^2-\frac{2}{n}. \end{aligned}$$

The restriction $n \geqq 4$ insures that the right-hand inequality is positive. Using the elementary inequality $1-x \leqq e^{-x}$ and raising the last displayed inequality to the $2k$-th power gives

$$\left|\frac{1}{n}+\frac{n-1}{n} r(\rho)\right|^{2k} \leqq \exp\left\{-2k\left[\frac{2m}{n}-2\left(\frac{m}{n}\right)^2+\frac{2}{n}\right]\right\}. \tag{3.6}$$

For the dimension, Lemma 8 and formula (2.7) give

$$d_\rho \leqq \binom{n}{m} \sqrt{(n-m)!} \leqq \frac{n^{n-m}}{\sqrt{(n-m)!}}.$$

Let $n-m=j$ so $0 \leqq j \leqq .3n$. Rewrite inequality (3.6) in terms of j and multiply by the bound for d_ρ^2 to get that the general term in Zone II is bounded above by $e^{-4k/n}$ times

$$\frac{1}{j!} \exp\left\{-2k\left[2\frac{j}{n}-2\left(\frac{j}{n}\right)^2\right]+2j \log n\right\}.$$

Since $k \geqq (n/2) \log n$, the exponent in the last displayed expression is larger than

$$\exp\left\{2j^2 \frac{\log n}{n}\right\}.$$

There are at most $p(j)$ irreducible representations for each value of j. It follows that the error term summed over Zone II is bounded above by $e^{-4k/n}$ times

$$\sum_{j=0}^{.3n} \frac{p(j)}{j!} \exp\left\{2j^2 \frac{\log n}{n}\right\}. \tag{3.7}$$

The bound (3.7) was derived under the assumption that $r(\rho) < -1/n$. If $-1/n < r(\rho) < 0$, then

$$\left|\frac{1}{n} + \frac{n-1}{n} r(\rho)\right|^{2k} \leqq \left(\frac{1}{n}\right)^{2k}$$

and the general term is bounded above by

$$\frac{p(j)}{j!} n^{2j-2k}.$$

It is easy to see that for any j the j-th term in (3.7) is larger than $p(j)$ times the term last displayed. Thus (3.7) bounds the error for Zone II. We will show that the sum (3.7) is bounded for all n. Hence the error from Zone II tends to zero like $e^{-4k/n}$.

Bounds for Zone III. Throughout this zone, $r(\rho)$ is positive so that increasing $r(\rho)$ increases $\left|\frac{1}{n} + \frac{n-1}{n} r(\rho)\right|$. Lemma 13 implies that $r(\rho)$ is smaller than $1/n(n-1)$ times

$$\begin{aligned} &\lambda_1(\lambda_1 - 1) + (n - \lambda_1 - 1)(n - \lambda_1 - 2) - 2 \\ &\quad = \lambda_1^2 + (n - \lambda_1)^2 - 3n + 2\lambda_1. \end{aligned}$$

Thus

$$\left|\frac{1}{n} + \frac{n-1}{n} r(\rho)\right| \leqq \left(\frac{\lambda_1}{n}\right)^2 + \left(1 - \frac{\lambda_1}{n}\right)^2 - \frac{2}{n^2}(n - \lambda_1).$$

Write $j = n - \lambda_1$, so $1 \leqq j \leqq .3n$. Making this substitution and using $1 - x \leqq e^{-x}$ leads to

$$\left|\frac{1}{n} + \frac{n-1}{n} r(\rho)\right| \leqq \exp\left\{-2\frac{j}{n} + 2\left(\frac{j}{n}\right)^2 - 2\frac{j}{n^2}\right\}.$$

In this notation, the bound (2.7) for the dimension yields

$$d_\rho^2 \leqq \frac{n^{2j}}{j!}.$$

Combining bounds, and multiplying by $p(j)$ to account for the number of distinct representations with a fixed value of j, gives $p(j)/j!$ times

$$\exp\left\{2j \log n - 2k\left[\frac{2j}{n} - 2\left(\frac{j}{n}\right)^2 + \frac{2j}{n^2}\right]\right\}.$$

Next write $k = \frac{n}{2}\log n + cn$. The term in the exponent is

$$2\frac{\log n}{n} j(j-1) - 4cn\left[\frac{j}{n} - \left(\frac{j}{n}\right)^2 + \frac{j}{n^2}\right].$$

Calculus arguments show that

$$\frac{j}{n}-\left(\frac{j}{n}\right)^2+\frac{j}{n^2}\geqq\frac{1}{n}.$$

Thus, we have that the error summed over Zone III is bounded above by

$$e^{-4c}\sum_{j=1}^{.3n}\frac{p(j)}{j!}\exp\left\{2j(j-1)\frac{\log n}{n}\right\}. \tag{3.13}$$

To complete the proof of Theorem 1 it only remains to collect the error bounds.

From (3.2) $(\frac{1}{3})^{2k}n!$

From (3.3) $e^{\pi(\frac{2}{3}n)^{\frac{1}{2}}}4^n(\frac{1}{2})^{2k}n^{2n/3}$

From (3.5) $e^{\pi(\frac{2}{3}n)^{\frac{1}{2}}}4^n\left[\frac{n}{2}\right]!\left(.58+\frac{.2}{n}+\frac{4}{n^2}\right)^{2k}$

From (3.7) $e^{-4k/n}\sum_{j=0}^{.3n}\frac{p(j)}{j!}\exp\left\{2j^2\frac{\log n}{n}\right\}$

From (3.13) $e^{-4c}\sum_{j=1}^{.3n}\frac{p(j)}{j!}\exp\left\{2j(j-1)\frac{\log n}{n}\right\}$.

We now show that the sum of these bounds is smaller than Be^{-2c} for some universal constant B, provided $c>0$ and $n\geqq 10$. First,

$$\left(\frac{1}{3}\right)^{2k}n!\leqq\left(\frac{1}{3}\right)^{2cn}\left(\frac{e}{3}\right)^{n\log n}<e^{-2cn}\leqq e^{-4c}.$$

Next, the term from (3.3) equals

$$\exp\{\pi(\tfrac{2}{3}n)^{\frac{1}{2}}+n\log 4+(\tfrac{2}{3}-\log 2)n\log n-2cn\log 2\}.$$

The term in the exponent which does not involve c tends to $-\infty$ as n tends to infinity because $2/3<\log 2$. Thus the last displayed expression is bounded above by B_1e^{-4c}. Next consider (3.5). The term $(.58+.2/n+4/n^2)$ is smaller than 1 for all $n\geqq 4$. When $n=10$, it is 0.66 and numerical computation shows $(.66)^{2cn}<e^{-4c}$ for $n\geqq 10$. Next observe that

$$e^{\pi(2/3n)^{\frac{1}{2}}}4^n\left[\frac{n}{4}\right]!\left(.58+\frac{.2}{n}+\frac{4}{2}\right)^{n\log n}$$

is bounded and tends to zero as n tends to infinity. This shows that the error term from (3.5) is bounded above by B_2e^{-4c}.

For (3.7) and (3.13) we must show that

$$S(n)=\sum_{j=0}^{.3n}\frac{p(j)}{j!}\exp\left\{2j^2\frac{\log n}{n}\right\}<B_3.$$

Using Stirling's formula and the inequality for $p(j)$ given in (2.12),

$$\frac{p(j)}{j!}=\exp[-j\log j+O(j)]. \tag{3.14}$$

Now consider the sum broken into three parts:

Part 1. $0\leqq j\leqq\sqrt{n/\log n}$. Then, the exponential term in $S(n)$ is bounded above by e^2, and $p(j)/j!$ sums to a finite limit using (3.14). For Parts 2 and 3 of the sum, express the general term as

$$\exp\left\{j\log n\left[\frac{2j}{n}-\frac{\log j}{\log n}+O\left(\frac{1}{\log n}\right)\right]\right\}. \tag{3.15}$$

Part 2. $\sqrt{n/\log n}<j\leqq n/\log n$. Then, the term in square brackets in the last display is bounded above by

$$\frac{2}{\log n}-\frac{1}{2}\left(1-\frac{\log\log n}{\log n}\right)+O\left(\frac{1}{\log n}\right)=-\frac{1}{2}+O\left(\frac{\log\log n}{\log n}\right).$$

This is bounded and tends to $-1/2$, so the sum over Part 2 is bounded for all n.

Part 3. $n/\log n<j\leqq 3n$. Then, the term in square brackets in (3.15) is bounded above by

$$.6-\left(1-\frac{\log\log n}{\log n}\right)+O\left(\frac{1}{\log n}\right)=-.4+O\left(\frac{\log\log n}{\log n}\right).$$

Thus the sum over Part 3 is bounded for all n. Using this, the error from (3.7) is bounded above by $B_3 e^{-4c}/n^2$. The error from (3.13) is bounded above by $B_3 e^{-4c}$. Theorem 1 follows by using the bounds just derived in (3.1). □

4. Markov Chains

A random walk on a finite group can be analyzed as a discrete Markov chain. In this section we show that the eigenvalues of the associated transition matrix are determined by the eigenvalues of the Fourier transforms at irreducible representations. In the walk determined by the probability T of (1.1), the transition matrix has an eigenvalue λ_ρ corresponding to each irreducible representation ρ. Using the notation of Sect. 2, $\lambda_\rho=(1/n+(n-1)/n\,r(\rho))$. The multiplicity of λ_ρ is d_ρ^2.

Let G be a finite group of order $|G|=g$. Let $x_1, x_2, \ldots, x_g$ be an enumeration of the elements of G. Let P be a probability measure on G. The *transition matrix* M of the random walk determined by P is a $g\times g$ matrix with i,j entry the probability of x_j in on step starting from x_i:

$$M_{ij}=P(x_j x_i^{-1}).$$

Theorem 3. *Let ρ be an irreducible representation of the finite group G. Let P be a probability measure on G. Let $\mathscr{E}_\rho$ denote the set of eigenvalues of the linear map $\rho(P)$. Then*

The set of eigenvalues of the transition matrix M equals (4.1)

$$\bigcup_\rho \mathscr{E}_\rho.$$

If the eigenvalue λ occurs with multiplicity $m(\lambda, \rho)$ in $\rho(P)$, the multiplicity of λ in M is (4.2)

$$\sum_\rho d_\rho m(\lambda, \rho).$$

When P is constant on conjugacy classes, the eigenvalues can be given more explicitly.

Corollary 3. *Let G be a finite group, P a probability measure on G which is constant on conjugacy classes, and M the associated transition matrix. Then, there is an eigenvalue λ_ρ of M corresponding to each irreducible representation ρ of G.*

$$\lambda_\rho = \frac{1}{d_\rho} \sum p_i n_i \chi_\rho^i$$

where the sum is over distinct conjugacy classes. On the i-th class, p_i denotes the value of P, n_i denotes the cardinality, and χ_ρ^i the value of the character χ_ρ. The eigenvalue λ_ρ occurs with multiplicity d_ρ^2.

Corollary 4. *Let M be the transition matrix of the probability T defined in* (1.1). *Then M has eigenvalues*

$$\left(\frac{1}{n} + \frac{n-1}{n} r(\rho)\right) \quad \text{with multiplicity } d_\rho^2$$

where $r(\rho)$ is defined by (2.11).

Theorem 3 and the corollaries are proved at the end of this section.

Remarks 1. From Corollary 4, the second largest eigenvalue of the transition matrix corresponding to the probability T of (1.1) is $(1-2/n)$. Just using this, the usual Perron-Frobenius argument gives $k \gg n^2 \log n$ as a rate for convergence to the uniform distribution. To get a result like Theorem 1 all the eigenvalues must be used.

2. Theorem 3 was discovered in an interesting way. Joseph Deken began computing the eigenvalues of the transition matrix corresponding to T in closed form using the MIT Macsyma system. For n smaller than 10, the second largest eigenvalue was $(1-2/n)$. This suggested a connection between the two approaches and led to Theorem 3.

3. When the group G is a cyclic group, the transition matrices are circulants. Corollary 3 suggests an interesting generalization of circulants that we hope to pursue elsewhere.

The argument for Theorem 3 proceeds by giving several equivalent ways to define the transition matrix M. We begin with a coordinate free version of M.

Let $L(G)$ denote the space of complex valued functions on G. The probability P defines a linear map from $L(G)$ into $L(G)$ by convolution: For $f \in L(G)$ and $x \in G$, let

$$P * f(x) = \sum_{y \in G} P(xy^{-1}) f(y).$$

Choose as a basis for $L(G)$ the functions $\delta_i \in L(G)$, defined by

$$\delta_i(x_j) = \delta_{ij} \quad \text{(Kroneeker delta)}.$$

An easy computation gives

$$P * \delta_i = \sum_j M_{ij} \delta_j. \tag{4.3}$$

Thus M is the matrix corresponding to convolution with P. Consider next the vector space $\mathbb{C}(G)$ with basis x_i $1 \leqq i \leqq g$. Defining multiplication in the obvious way, $\mathbb{C}(G)$ becomes an algebra, *the group algebra of* G. Let $Q \in \mathbb{C}(G)$ be defined by

$$Q = \sum_j P(x_j) x_j.$$

Left multiplication by Q defines a linear map from $\mathbb{C}(G)$ into $\mathbb{C}(G)$. We now show that M is the matrix of this map with respect to the basis x_j. Indeed,

$$Q x_i = \sum_j P(x_j x_i^{-1}) x_j$$

and

$$(P * \delta_i) x_j = \sum_x P(x_j x^{-1}) \delta_i(x) = P(x_j x_i^{-1}).$$

In view of (4.3), $(P * \delta_i) x_j = M_{ij}$. This proves the claim. Finally, *the left regular representation* π *of* G assigns a linear map of $\mathbb{C}(G)$ into itself to each $x \in G$ via

$$\pi(x) x_j = x x_j.$$

Recall that we write $\pi(P) = \sum_j P(x_j) \pi(x_j)$. With this notation we can state the basic lemma.

Lemma 15. *The following four* $g \times g$ *matrices are identical.*

(1) *The transition matrix* M.

(2) *The matrix of the linear map* $P*: L(G) \to L(G)$ *with respect to the basis* δ_j.

(3) *The matrix of the linear map given by left multiplication by* Q *with respect to the basis* x_i *of* $\mathbb{C}(G)$.

(4) *The matrix of the linear map* $\pi(P)$ *with respect to the basis* x_j *of* $\mathbb{C}(G)$.

Proof. We have already shown that (1), (2), and (3) are identical. We now show that (2) and (4) are identical. Take the π transform of (4.3) to obtain

$$\pi(P) \pi(\delta_i) = \sum_j M_{ij} \pi(\delta_j).$$

Using $\pi(\delta_i) = \pi(x_i)$ we obtain

$$\pi(P) \pi(x_i) = \sum_j M_{ij} \pi(x_j). \tag{4.4}$$

Let H be the space of linear operators on the vector space $\mathbb{C}(G)$ and define

$$V=\pi(\mathbb{C}(G))\subset H.$$

Define a representation $\mathcal{O}$ of G on V by

$$\mathcal{O}(x)\,\pi(y)=\pi(x)\,\pi(y)=\pi(xy).$$

Then for all $x\in G$ (resp. $f\in L(G)$) we have the following commutative diagram

$$\begin{array}{ccc} \mathbb{C}(G) & \xrightarrow{\pi} & V \\ {\scriptstyle \pi(f)}\downarrow & & \downarrow{\scriptstyle \mathcal{O}(f)} \\ \mathbb{C}(G) & \xrightarrow{\pi} & V \end{array}$$

Indeed, the image of y belonging to $\mathbb{C}(G)$ is $\pi(f)\,\pi(y)$ via either route. In particular, the matrix of $\pi(p)$ with respect to the basis $\{x_j\}$ is identical to the matrix $\mathcal{O}(p)$ with respect to the basis $\{\pi(x_j)\}$. In view of (4.4), the latter is M. □

Proof of Theorem 3. As in Sect. (2.4) of Serre (1977), the vector space $\mathbb{C}(G)$ can be decomposed as a direct sum of invariant subspaces

$$\mathbb{C}(G)=\bigoplus_{\rho} V_\rho.$$

The terms are indexed by irreducible representations ρ. The subspace V_ρ is itself a direct sum of d_ρ copies of the subspace W_ρ which is isomorphic to the vector space W'_ρ corresponding to ρ. Serre shows that the regular representation π restricted to any of the W_ρ is equivalent to the representation ρ. It follows that the eigenvalues of $\pi(P)$ acting on W_ρ are the same as the eigenvalues of $\rho(P)$. □

Proof of Corollary 3. When P is constant on conjugacy classes, the linear map $\rho(P)$ was shown to be a constant times the identity in Lemma 5. □

5. Final Remarks

1. Better Bounds and Small n. The bounds used at the end of Sect. 3 have been reasonably crude. For numerical computation, direct use of n and k in one of the preliminary inequalities in each of the five zones involved in the proof of Theorem 1 gives much tighter bounds. For $n\leqq 10$, the values of $\chi_\rho(\tau)$ and d_ρ are given exactly in Littlewood (1958). These can be used directly in (3.1).

2. Different Measures. While limited, the approach used in this paper can be used to get bounds for some other measures on S_n. The measure T assigns much larger probability to the identity than to any transposition. The method of proof allows a similar analysis for the probability measure T_p where $T_p(id)=p$, $T_p(\tau)=(1-p)\Big/\binom{n}{2}$ for all transpositions τ. If $p=0$, T^{*k} does not converge

to the uniform distribution since T^{*k} (even permutations)$=1$ when k is even. The problem is quite sensitive to the choice of p. For example, let p be chosen to make the chance of any transposition equal to the chance of the identity $\left(p=1\Big/\left[1+\binom{n}{2}\right]\right)$. Then, it may be shown that k must be chosen as $n^2 C_n$ where $C_n \to \infty$, in order to have convergence to the uniform distribution.

3. Other Approaches to Proof. The problem discussed in this paper is a special case of a classical problem. How many times should a deck of cards be shuffled until it is close to random? The classical texts of Poincare, Doob, and Feller each devote several pages to this problem. They use the methods of discrete Markov chains, approximating the second largest eigenvalue of the transition matrix. We have discussed this in Sect. 4.

Shuffles very similar to the shuffle of Theorem 1 are attacked by direct combinatorial methods in Robbins and Bolker (1980). As an example of this approach in our problem, notice that the permutation resulting from k random transpositions will be even if and only if the number of times $L_i \neq R_i$ is even. The chance of this is easily seen to be $\frac{1}{2}\left(1+(-1)^k\left(1-\frac{2}{n}\right)^k\right)$. This proves our earlier remark: the measure T^{*k} is never exactly uniform except when $n=2$. As a second example, Bob Bell (personal communication) has shown that the chance that card 1 is in position 1 after k transpositions equals $1/n+[(n-1)/n](1-2/n)^k$.

There is some literature on the rate of convergence for random walk on a compact group. See Bhattacharya (1972), Heyer (1978), and Major and Shlosman (1979). These approaches give convergence in very general situations. They do not seem particularly aimed at accurate rates in special problems. For example, the main inequality of Major and Shlosman (1979) - their Lemma 1 - in connection with our results in Sect. 2, results in an upper bound for the variation distance between T^{*k} and U which suggests $k \gg n^2 \log n$ is needed.

As mentioned in the introduction, Aldous (1980) and, independently Durrett (personal communication), used coupling techniques to show that $k \gg n^2$ was sufficient for the problem treated in Theorem 1. A theorem of Griffeath (1976) implies that some coupling method exists which gives the optimal rate. It seems that this maximal coupling must be fairly complex.

We mention two other approaches. Aldous (1980) and Diaconis, Flatto, and Shepp (1980) have used the method of stopping times - find a time t such that the distribution of t transpositions is exactly uniform - in some shuffling problems. Reeds (1980) has an ingenious approach which leads to rates of convergence for shuffling methods close to real riffle shuffling.

References

Aldous, D.: Mixing time inequalities for finite Markov chains. Manuscript circulated, January 1980

Ayoub, R.: An Introduction to the Analytic Theory of Numbers. Amer. Math. Soc. Rhode Island: Providence 1963

Bhattacharya, R.N.: Speed of convergence of the n-fold convolution of a probability measure on a compact group. Z. Wahrscheinlichkeitstheorie verw. Gebiete **25**, 1-10 (1972)
Curtis, C.W., Reiner, I.: Representation Theory of Finite Groups and Associative Algebras. New York: Interscience 1962
Diaconis, P., Flatto, L., Shepp, L.: On generating a random permutation with transpositions. Unpublished manuscript 1980
Feller, W.: An Introduction to Probability Theory and its Applications, Vol. 1, 3rd ed. New York: Wiley 1968
Frobenius, F.G.: Über die Charakter der Symmetrischen Gruppen. Berliner Berichte (1901)
Griffeath, D.: A maximal coupling for Markov chains. Z. Wahrscheinlichkeitstheorie verw. Gebiete **31**, 15-106 (1975)
Hewitt, E., Ross, K.A.: Abstract Harmonic Analysis II. Berlin: Springer-Verlag 1970
Heyer, H.: Probability Measures on Locally Compact Groups. Berlin: Springer-Verlag 1977
Ingram, R.E.: Some characters of the symmetric group. Proc. Amer. Math. Soc. **1**, 358-369 (1950)
James, G.D.: The Representation Theory of the Symmetric Groups. Lecture Notes in Mathematics 682. Berlin-Heidelberg-New York: Springer-Verlag 1978
Knuth, D.: The Art of Computer Programming, Vol. II. Reading, Massachusetts: Addison Wesley 1969
Littlewood, D.E.: The Theory of Group Characters, 2nd ed. London: Oxford University Press 1958
MacDonald, I.G.: Symmetric Functions and Hall Polynomials. London: Oxford University Press 1979
Major, P., Shlosman, S.B.: A local limit theorem for the convolution of a probability measure on a compact group. Z. Wahrscheinlichkeitstheorie verw. Gebiete **50**, 137-148 (1979)
Marshall, A.W., Olkin, I.: Inequalities: Theory of Majorization and its Applications. New York: Academic Press 1974
Reeds, J.: An analysis of Riffle shuffles. (Unpublished manuscript 1980)
Robbins, D.P., Bolker, E.D.: The bias of three pseudo random shuffles. (to appear in Aequations Mathematicae)
Serre, J.P.: Linear Representations of Finite Groups. New York-Berlin-Heidelberg: Springer 1977

Received April 15, 1980