The Annals of Probability
1988, Vol. 16, No. 3, 1215–1228

COVERING PROBLEMS FOR MARKOV CHAINS[1]

BY PETER MATTHEWS

University of Maryland

Upper and lower bounds are given on the moment generating function of the time taken by a Markov chain to visit at least n of N selected subsets of its state space. An example considered is the class of random walks on the symmetric group that are constant on conjugacy classes. Application of the bounds yields, for example, the asymptotic distribution of the time taken to see all $N!$ arrangements of N cards as $N \to \infty$ for certain shuffling schemes.

1. Introduction. In the coupon collector's problem as considered in Feller (1968) balls are randomly thrown into N urns. The number of throws T until all urns are nonempty satisfies

$$Ee^{(s/N)(T - N\log N)} \to \Gamma(1-s), \quad \text{as } N \to \infty \text{ for } s < 1. \tag{1.1}$$

$\Gamma(1-s)$ is the moment generating function of the extreme value (Gumbel) distribution so T, properly normalized, has asymptotically this distribution. Let U_c be the number of empty urns after $N(\log N + c)$ throws. Then

$$U_c \overset{\text{dist}}{\to} \text{Poisson}(e^{-c}), \quad \text{as } N \to \infty. \tag{1.2}$$

Kolchin, Sevast'yanov and Chistyakov (1978) give proofs of (1.1) and (1.2) and generalizations.

Next consider a similar question for an N-state Markov chain. How long until the Markov chain has covered its state space, i.e., how long until all states are visited? How many states have not been visited after $N(\log N + c)$ steps? Aldous (1983) first considered this type of problem. For a rapidly mixing random walk on a finite group, Aldous showed that the expected time taken to visit all group elements was essentially $RN \log N$, where R is the expected number of returns to the starting position in the short term.

Here another approach to these problems using auxiliary randomization is used. Matthews (1988) gives bounds applicable to mean covering times for finite Markov chains. Here an extension to bounds on moment generating functions is given. This will allow more precise statements like (1.1) and (1.2) to be made. The bounds are applicable in quite general situations, though examples where they give precise results are thus far limited to nice situations like random walks on finite groups.

Let X_k, $k = 0, \dots,$ be a finite irreducible Markov chain with state space A. Let $A_1, \dots, A_N$ be a collection of subsets of A, not necessarily disjoint. For $n \le N$ let Q_n be the first time k such that at least n of the intersections

Received October 1986; revised February 1987.
[1]This research was supported in part by NSF Grant MCS-80-24649.
AMS 1980 *subject classifications*. Primary 60G17, 60J10; secondary 60B15.
Key words and phrases. Symmetric group, random shuffle, random allocation.

$\{X_0, \dots, X_k\} \cap A_i$, $i = 1, \dots, N$, are nonempty, the first time the process has visited n of the selected subsets. For an initial position a_0 of interest define

$$\hat{A}_i = \{a_0\} \cup \bigcup_{j \neq i} A_j,$$

for $i = 1, \dots, N$. Further define $T(A_i) = \min\{k\colon X_k \in A_i\}$ and $T(A_1, \dots, A_i) = \max_{j=1,\dots,i} T(A_j)$.

Let P_a and E_a denote probability and expectation for the chain started at a. Define

$$f^-(s) = \min_{1 \le i \le N} \min_{a \in \hat{A}_i} E_a e^{sT(A_i)}$$

and

$$f^+(s) = \max_{1 \le i \le N} \max_{a \in \hat{A}_i} E_a e^{sT(A_i)}.$$

$f^+(s)$ and $f^-(s)$ are the pointwise maximum and minimum of the moment generating functions of the time taken to hit a set $A_i \in \{A_1, \dots, A_N\}$ from the initial position or another $A_j \in \{A_1, \dots, A_N\}$, $j \neq i$.

In Section 2 the following inequality will be derived.

THEOREM 1.3. *For all s such that $f^+(s) < \infty$,*

$$\begin{aligned}
&\prod_{i=1}^{N-n} \frac{i - 1 + 1/f^+(s)}{i - 1 + 1/f^-(s)} \prod_{i=N-n+1}^{N} \frac{i}{i - 1 + 1/f^-(s)}\\
&\quad \le E_{a_0} e^{sQ_n}\\
&\quad \le \prod_{i=1}^{N-n} \frac{i - 1 + 1/f^-(s)}{i - 1 + 1/f^+(s)} \prod_{i=N-n+1}^{N} \frac{i}{i - 1 + 1/f^+(s)}.
\end{aligned}$$

This can be rewritten in terms of gamma functions. In the special case $n = N$, Q_N is the time taken to visit all of $\{A_1, \dots, A_N\}$, and the bounds become

$$\frac{\Gamma(N+1)\Gamma(1/f^-(s))}{\Gamma(N + 1/f^-(s))} \le E_{a_0} e^{sQ_N} \le \frac{\Gamma(N+1)\Gamma(1/f^+(s))}{\Gamma(N + 1/f^+(s))}. \tag{1.4}$$

If $A_1, \dots, A_N$ are singletons consisting of all the possible values of X except a_0, then (1.4) gives bounds on the generating function of the time taken by X to cover its state space. In many situations the bounds will be far enough apart to be almost useless. However, in many situations the bounds can be quite tight. A special case of the example considered in Section 4 is the following. Consider a deck of M cards laid out in a row on a table. Two distinct cards are chosen randomly and switched. This is repeated independently many times. This generates a Markov chain, actually a random walk, on the set of all $M!$ arrangements of the cards. The asymptotic distribution of the time T_M to see all arrangements can be found using (1.4). The result, a special case of (4.4), is

$$Ee^{(s/M!)(T_M - M! \log M!)} \to \Gamma(1 - s), \quad \text{as } M \to \infty \text{ for } s < 1. \tag{1.5}$$

If arrangements were chosen uniformly at random, the coupon collector's problem with $N = M!$ urns would result. By (1.1) and (1.5) the time to see all arrangements is essentially the same whether arrangements are chosen randomly or by random transpositions.

In the same random transposition Markov chain, let U_c denote the number of unvisited permutations after $M!(\log M! + c)$ transpositions. Since, excluding the initial position, there are $N = M! - 1$ permutations to visit,

$$P(U_c = j) = P(Q_{M!-1-j} \le k) - P(Q_{M!-j} \le k). \tag{1.6}$$

This can be combined with Theorem 1.3 to yield a result analogous to (1.2), namely,

$$U_c \overset{\text{dist}}{\to} \text{Poisson}(e^{-c}), \quad \text{as } M \to \infty. \tag{1.7}$$

Whenever $1/f^{+}(s/N)$ and $1/f^{-}(s/N)$ differ by $o(\log N)$ the bounds in Theorem 1.3 will be tight asymptotically, and results like (1.5) and (1.7) will hold. In some other situations the bounds fail to be tight because hitting times are stochastically too small for a small number of (initial position, set to be hit) pairs. In these cases ad hoc improvements to Theorem 1.3 can be given. Matthews (1985) contains an example of this for a random walk on the cube Z_2^N.

Finally, again for random transpositions, consider the spatial distribution of the unvisited permutations after k transpositions. A first question is, are the unvisited permutations more or less clumped together than if permutations were chosen as in the coupon collector's problem? In a weak sense, neither distribution is clumpier.

Formally, for two permutations π and σ let $d(\pi, \sigma)$ be Cayley's distance, the minimum number of transpositions needed to transform π into σ. For a set of M cards let

$$C_M = \inf\{k\colon d(\pi, \sigma) > 1, \forall\, \pi \neq \sigma, \text{such that } T(\tau) > k \text{ and } T(\alpha) > k\}. \tag{1.8}$$

In words C_M is the first time all unvisited permutations differ by more than one transposition and is, using Cayley's distance, the time taken until all unvisited permutations are isolated from each other. Theorem 1.3 can be used to show

$$P\big(C_M > \tfrac{1}{2}M!\big(M! + \log^2 M\big)\big) \to 0 \tag{1.9}$$

and

$$P\big(C_M < \tfrac{1}{2}M!\big(M! - \log^2 M\big)\big) \to 0, \quad \text{for } \varepsilon > 0 \text{ as } M \to \infty.$$

Thus the distribution of C_M has a sharp cutoff at $\frac{1}{2}$M!log $M!$ for large M. The asymptotics in (1.9) hold for independent allocations as well, so at this level of investigation the unvisited permutations are equally clumpy in these cases.

The outline for the remainder of this article follows. Section 2 gives a proof of Theorem 1.3. In Section 3 some background material on hitting times for random walks on finite groups is given. Random walks on the symmetric group that are constant on conjugacy classes are the examples considered in Section 4. Specialization to random transpositions yields the examples stated above.

2. General results. In this section Theorem 1.3 is proven. The proof can be generalized to arbitrary time homogeneous strong Markov processes. For simplicity, only finite Markov chains will be considered here.

Consider a Markov chain X with state space A and initial position $a_0 \in A$. Consider $(A_1, \dots, A_N)$, a collection of subsets of A of interest. Assume for later use that these subsets have labels $1, 2, \dots, N$. Define $T(A_i)$, Q_n, P_a, E_a, $f^+(s)$ and $f^-(s)$ as in the Introduction. P and E without subscripts will stand for P_{a_0} and E_{a_0}. Let σ be a uniformly distributed random permutation that is independent of X under P.

Let F_i be the σ-field generated by σ and $\{X_j,\ j = 0, \dots, i\}$ for $i = 0, 1, \dots$. For any permutation π of order N let A_i^π denote the π_ith member of $(A_1, \dots, A_N)$.

PROPOSITION 2.1. *$T(A_1^\sigma, \dots, A_i^\sigma)$ is a stopping time with respect to $\{F_j,\ j = 0, 1, \dots\}$ for $i = 1, \dots, N$.*

PROOF. $T(A_j^\sigma \le k) = \bigcup_{j=1}^N \{\sigma_j = i\} \cap \{T(A_i) \le k\}$. Each event on the right is in F_k. □

Let F^0 be the σ-field generated by σ, and for $i = 1, 2, \dots, N$ let F^i be the σ-field generated by σ and $\{X_k,\ k = 0, \dots, T(A_1^\sigma, \dots, A_i^\sigma)\}$. Further define $R_1 = T(A_1^\sigma)$ and for $i = 2, \dots, N$, $R_i = T(A_1^\sigma, \dots, A_i^\sigma) - T(A_1^\sigma, \dots, A_{i-1}^\sigma)$. Thus $T(A_1, \dots, A_N) = \sum_{i=1}^N R_i$. For $i = 1, \dots, N$ define a random variable r_i as follows. Let $(A_1^\omega, A_2^\omega, \dots, A_N^\omega)$ be a listing of $(A_1, \dots, A_N)$ in the order in which they are visited, with ties broken by the convention that if $T(A_i) = T(A_j)$, then A_i appears before A_j in the list if the label i is less than the label j. Let $r_i = 1$ if A_i^σ is further to the right in $(A_1^\omega, \dots, A_N^\omega)$ than all of $(A_1^\sigma, \dots, A_{i-1}^\sigma)$. Otherwise $r_i = 0$. If ties are impossible, then $r_i = I\{R_i \ne 0\}$.

PROPOSITION 2.2. *For $i = 1, \dots, N$, $r_i \in F^{i-1}$.*

PROOF. Write $\{r_i = 1\} = \{T(A_i^\sigma) > T(A_1^\sigma, \dots, A_{i-1}^\sigma)\} \cup \{\{T(A_i^\sigma) = T(A_1^\sigma, \dots, A_{i-1}^\sigma)\} \cap \bigcap_{j=1}^{i-1} \{\{T(A_j^\sigma) < T(A_1^\sigma, \dots, A_{i-1}^\sigma)\} \cup \{\sigma_j < \sigma_i\}\}\}$. Each of these events is in F^{i-1}. □

PROPOSITION 2.3. *For $i = 1, \dots, N$, $E(r_i) = P(r_i = 1) = 1/i$.*

PROOF. The random order $(A_1^\omega, \dots, A_N^\omega)$ depends only on X, and hence is independent of σ. Conditional on a particular value of $(A_1^\omega, \dots, A_N^\omega)$, $\{r_i = 1\}$ is just the event that σ_i appears further to the right in $\omega_1, \dots, \omega_N$ than $\sigma_1, \dots, \sigma_{i-1}$. By independence and uniformity of σ, this has conditional, hence unconditional, probability $1/i$. □

PROPOSITION 2.4. *For $i = 1, \dots, N$, r_i and $T(A_1^\sigma, \dots, A_i^\sigma)$ are independent.*

PROOF. Let $\{\sigma_1,\dots,\sigma_i\}$ be the unordered set containing $\sigma_1,\dots,\sigma_i$ and let $\{\pi_1,\dots,\pi_i\}$ be an arbitrary set of i members of $\{1,2,\dots,N\}$. Write

$$\begin{aligned} &P(T(A_1^\sigma,\dots,A_i^\sigma) \le k \cap r_i = 1) \\ &= \binom{N}{i}^{-1} \sum_{(\pi_1,\dots,\pi_i)} P(T(A_1^\sigma,\dots,A_i^\sigma) \\ &\qquad\qquad \le k \cap r_i = 1 | \{\sigma_1,\dots,\sigma_i\} = \{\pi_1,\dots,\pi_i\}) \\ &= \binom{N}{i}^{-1} \sum_{(\pi_1,\dots,\pi_i)} P(T(A_1^\pi,\dots,A_i^\pi) \\ &\qquad\qquad \le k \cap r_i = 1 | \{\sigma_1,\dots,\sigma_i\} = \{\pi_1,\dots,\pi_i\}). \end{aligned} \tag{2.5}$$

Condition as well on $X_0,\dots,X_k$. Then $\{T(A_1^\pi,\dots,A_i^\pi) \le k\}$ will have conditional probability 0 or 1, hence $\{T(A_1^\pi,\dots,A_0^\pi) \le k\}$ and $\{r_j = 1\}$ will be conditionally independent. As in the proof of Proposition 2.3, $P(r_i = 1 | X_0,\dots,X_k, \{\sigma_1,\dots,\sigma_i\} = \{\pi_1,\dots,\pi_i\}) = 1/i$. Thus (2.5) can be rewritten

$$\begin{aligned} &\binom{N}{i}^{-1} \sum_{(\pi_1,\dots,\pi_i)} P(T(A_1^\pi,\dots,A_i^\pi \le k) | \{\sigma_1,\dots,\sigma_i\} = \{\pi_1,\dots,\pi_i\}) 1/i \\ &\quad = P(T(A_1^\pi,\dots,A_i^\pi) \le k) P(r_i = 1). \end{aligned}$$

□

To prove Theorem 1.3 it will be convenient to prove (1.4) first. To prove (1.4) first note

$$\begin{aligned} \frac{1}{1 - 1 + 1/f^-(s)} &= f^-(s) \le Ee^{sT(A_1^\sigma)} \\ &\le f^+(s) = \frac{1}{1 - 1 + 1/f^+(s)}, \end{aligned} \tag{2.6}$$

by definition of $f^-(s)$ and $f^+(s)$. Infinite values of $f^+(s)$ and $f^-(s)$ present no difficulty if $1/0$ is interpreted as ∞. For brevity let T_i denote $T(A_1^\sigma,\dots,A_i^\sigma)$. For $i \ge 2$ write

$$Ee^{sT_i} = E\left(e^{sT_{i-1}} E\left(e^{sR_i} | F^{i-1}\right)\right), \tag{2.7}$$

by definition of F^{i-1}. Consider the inner conditional expectation. By Proposition 2.2 and the fact that $R_i = 0$ whenever $r_i = 0$ it can be written

$$1 - r_i + r_i E\left(e^{sR_i} | F^{i-1}\right).$$

On the set $\{r_i = 1\}$, R_i is the time taken to hit A_i^σ from $X_{T_{i-1}}$. By the Markov property the conditional expectation is

$$1 - r_i + r_i E_{X_{T_{i-1}}} e^{sT(A_i^\sigma)}.$$

If $r_i = 1$, then $X_{T_{i-1}}$ must be in $\hat{A}_i^\sigma$, so by the definitions of $f^-(s)$ and $f^+(s)$,

$$1 - r_i + r_i f^-(s) \le E\left(e^{sR_i} | F^{i-1}\right) \le 1 - r_i + r_i f^+(s). \tag{2.8}$$

Plugging back into (2.7) yields

$$\begin{aligned}(2.9)\qquad & Ee^{sT_{i-1}}\left[(1-r_i)(1-f^-(s))+f^-(s)\right]\\ &\le Ee^{sT_i}\\ &\le Ee^{sT_{i-1}}\left[(1-r_i)(1-f^+(s))+f^+(s)\right].\end{aligned}$$

Restrict attention to the upper bound of (2.9); the lower bound can be treated in exactly the same manner. On the set $\{r_i = 0\}$, $R_i = 0$ and hence $T_{i-1} = T_i$. Thus (2.9) is

$$Ee^{sT_i} \le (1-f^+(s))Ee^{sT_i}(1-r_i) + f^+(s)Ee^{sT_{i-1}}.$$

By Proposition 2.4 e^{sT_i} and $1-r_i$ are independent, and by Proposition 2.3 $E(1-r_i) = 1 - 1/i$. The bound can be rearranged to

$$(2.10)\qquad Ee^{sT_i} \le \frac{i}{i-1+1/f^+(s)}Ee^{sT_{i-1}}.$$

Combining (2.10) for $i = 2,\dots,N$ with (2.6) yields

$$Ee^{sT_N} = Ee^{sQ_N} \le \prod_{i=1}^{N} \frac{i}{i-1+1/f^+(s)},$$

which is the upper bound of (1.4). The lower bound follows in the same manner.

Theorem 1.3 can now be proven. Consider Q_n, the time taken to visit at least n members of $\{A_1,\dots,A_N\}$. Q_n is a stopping time with respect to $F_0, F_1,\dots$. Let G^n be the σ-field generated by σ and $X_0,\dots,X_{Q_n}$. Suppose $n+m$ members of $\{A_1,\dots,A_N\}$ have been visited at time Q_n. If ties among $T(A_1),\dots,T(A_N)$ are impossible, then $m = 0$. Let $\{A_1^\theta,\dots,A_{N-n}^\theta\}$ be the set of $N-n-m$ members of $\{A_1,\dots,A_N\}$ that are unvisited at time R_n and the m members of $\{A_1,\dots,A_N\}$ among the $m+1$ members first visited at time R_n that have the largest labels. $\{A_1^\theta,\dots,A_{N-n}^\theta\}$ has $N-n$ members, and membership in $\{A_1^\theta,\dots,A_{N-n}^\theta\}$ is determined by G^n.

Write $Q_N = Q_n + (Q_N - Q_n)$. $(Q_N - Q_n)$ is the time taken, starting at X_{Q_n}, to visit all of $\{A_1^\theta,\dots,A_{N-n}^\theta\}$. Write

$$(2.11)\qquad Ee^{sQ_N} = E\left(e^{sQ_n}E\left(e^{s(Q_N-Q_n)}|G^n\right)\right).$$

As in the proof of (1.4),

$$\begin{aligned}(2.12)\qquad & \prod_{i=1}^{N-n}\frac{i}{i-1+1/f^-(s)}\\ &\le E\left(e^{s(Q_N-Q_n)}|G^n\right) \le \prod_{i=1}^{N-n}\frac{i}{i-1+1/f^+(s)}.\end{aligned}$$

As long as $f^+(s)$, and hence $f^-(s)$, is finite, (2.12) and (1.4) can be inserted into (2.11) and terms can be rearranged, yielding Theorem 1.3.

3. Finite group hitting times. Here some basic facts about random walks on finite groups are reviewed. Diaconis (1988) gives a detailed treatment of random walks on finite groups. The transition matrix of a random walk on a finite group has nice properties that make the calculation of the generating functions $f^+(s)$ and $f^-(s)$ needed in Theorem 1.3 tractable. The main purpose of this section is to give conditions sufficient for $f^+(s)$ to exist for some $s > 0$, which allows asymptotic distributions of covering times to be deduced from bounds on moment generating functions.

Let $G = \{g_1, g_2, \dots, g_{|G|}\}$ be a finite group with group operation $\cdot$. A random walk on G is essentially described by a probability measure μ on G. If $Y_1, Y_2, \dots$ are independent with distribution μ and X_0 is the initial position of the random walk, then $X_1 = Y_1 \cdot X_0$, $X_2 = Y_2 \cdot X_1, \dots$ are the successive group elements visited by the random walk. The random walk is a finite Markov chain with transition matrix P having entries $P_{ij} = \mu(g_j g_i^{-1})$ since to go to g_j at step k from g_i, $Y_k = g_j g_i^{-1}$ is needed.

Certain properties of the Markov chain follows from properties of μ. If the support of μ generates G, then the Markov chain is irreducible. The measure μ is constant on conjugacy classes if $\mu(ghg^{-1}) = \mu(h)$ for all $g, h \in G$. The measure μ is symmetric if $\mu(g) = \mu(g^{-1})$ for all $g \in G$. This implies the transition matrix P is symmetric. The examples of random walks on symmetric groups considered in Section 4 will all be constant on conjugacy classes. This simplifies the analysis considerably and, for the symmetric group, implies symmetry.

Diaconis and Shashahani (1981) showed that the decomposition of P into eigenvalues and eigenvectors reduces to the same decomposition of the Fourier transforms of μ at the irreducible representations of G. For $z \in C$, $|z| < 1$, let

$$P(z) = \sum_{n=0}^{\infty} z^n P^n.$$

Assume the chain is irreducible. Then from classical Markov chain theory the generating function of the time taken to hit g_j from g_i is

$$F_{ij}(z) = E_{g_i}\left(z^{T(g_j)}\right) = \frac{P_{ij}(z)}{P_{jj}(z)}.$$

Use of the decomposition of P or an argument of Flatto, Odlyzko and Wales (1985) yields

$$F_{ij}(z) = \frac{\sum_{r=1}^{R} d_r \operatorname{Tr}\left(\rho_r\left(g_i g_j^{-1}\right)\left(I - z\rho_r(\mu)\right)^{-1}\right)}{\sum_{r=1}^{R} d_r \operatorname{Tr}\left(I - z\rho_r(\mu)\right)^{-1}}, \tag{3.1}$$

for $z < 1$, where ρ_r, $r = 1, \dots, R$, are the inequivalent irreducible representations of G and d_r, $r = 1, \dots, R$, their dimensions. Tr denotes trace, and $\rho_r(\mu)$ is the Fourier transform of μ at ρ_r,

$$\rho_r(\mu) = \sum_{g \in G} \rho_r(g)\mu(g).$$

The structure of a random walk on a finite group permits calculation of $f^+(s)$ and $f^-(s)$ as the maximal and minimal generating functions of a restricted class of hitting times. To study the time taken by the random walk to visit every group element, in principle $F_{ij}(z)$ must be calculated for all $g_i, g_j \in G$. However, the time to hit g_j from g_i has the same distribution as the time to hit the identity from $g_i g_j^{-1}$. Let

$$F_i(z) = E_{g_i} z^{T(\text{id})},$$

where id is the group identity. In calculating maximal and minimal generating functions of hitting times of singletons, attention can and will be restricted to $F_i(e^s)$ for $g_i \neq \text{id}$.

For the bounds in Theorem 1.3 to give the asymptotic distribution of Q_N as $N \to \infty$, $f^+(s)$, and hence $f^-(s)$, must exist in an interval containing 0, which will follow from $F_i(z)$ existing in a ball of radius greater than 1 for all i. Let $F_{ij}^*(z)$ be the function defined by (3.1) for all z. $F_{ij}^*(z)$ is a rational function with no poles in the ball $|z| < 1$. An analytic continuation argument shows that the generating function $F_{ij}(z)$ exists and is equal to $F_{ij}^*(z)$ for $|z| < |z_0|$, where z_0 is the pole of $F_{ij}^*(z)$ of smallest magnitude. The following proposition gives sufficient conditions for $F_{ij}(z)$ to exist for all i, j if μ is symmetric.

Proposition 3.2. *Suppose μ is a symmetric measure on G and the support of μ generates G. If $-1 \le \lambda_1 < \lambda_2 < \cdots < \lambda_{p-1} < \lambda_p = 1$ are the ordered eigenvalues of the transition matrix P with multiplicities $m_1, \dots, m_{p-1}, m_p = 1$, then for $F_{ij}(z)$ to exit for all $g_i, g_j \in G$ it is sufficient that the following hold*:

(i) $$1/|z| > \lambda_{p-1},$$

(ii) $$-1/|z| < \lambda_2,$$

(iii) $$1 + (1 - |z|) \sum_{i=1}^{p-1} \frac{m_i}{1 - |z|\lambda_i} > 0,$$

(iv) $$-1/|z| < \lambda_1 \quad \text{or} \quad m_1 + (1 + \lambda_1|z|) \sum_{i=2}^{p} \frac{m_i}{1 + |z|\lambda_i} > 0.$$

Proof. Consider an irreducible representation ρ_r of G of dimension d_r. Since $\mu(g^{-1}) = \mu(g)$ and $\rho_r(g^{-1}) = \rho_r^*(g)$ for all $g \in G$, $\rho_r(\mu)$ is a real symmetric matrix. Thus $\rho_r(\mu)$ has real eigenvalues and there is a unitary matrix U such that $U\rho_r(\mu)U^*$ is diagonal. Since $U\rho_r U^*$ and ρ_r are equivalent irreducible representations of G, without loss of generality assume that each ρ_r has been chosen so that $\rho_r(\mu)$ is diagonal. Let $\Lambda(\mu)$ be the $|G|$ by $|G|$ diagonal matrix with diagonal blocks $\rho_r(\mu)$ each appearing d_r times. Also let $\Lambda(g)$ be the block diagonal matrix with diagonal blocks $\rho_r(g)$ appearing with multiplicity d_r, in the same order as $\Lambda(\mu)$.

For $z \in C$ (3.1) can be written

$$F_{ij}^*(z) = \frac{\operatorname{Tr}\left(\Lambda\left(g_i g_j^{-1}\right)(I - z\Lambda(\mu))^{-1}\right)}{\operatorname{Tr}(I - z\Lambda(\mu))^{-1}}.$$

That is,

$$\frac{\prod_{i=1}^{p}(1 - z\lambda_i)\operatorname{Tr}\left(\Lambda\left(g_i g_j^{-1}\right)(I - z\Lambda(\mu))^{-1}\right)}{\prod_{i=1}^{p}(1 - z\lambda_i)\sum_{i=1}^{p}\dfrac{m_i}{(1 - z\lambda_i)}}.$$

The denominator is a polynomial in z. The numerator is also, since it is $\prod_{i=1}^{p}(1 - z\lambda_i)$ times a linear combination of terms of the form $(1 - z\lambda_i)^{-1}$. Thus for all i, j, the smallest pole of $F_{ij}^*(z)$ is no smaller than the smallest zero of

$$\prod_{i=1}^{p}\left(\frac{1}{z} - \lambda_i\right)\sum_{i=1}^{p}\frac{m_i}{\left(\dfrac{1}{z} - \lambda_i\right)}. \tag{3.3}$$

The characteristic polynomial of P is $Q(x) = \prod_{i=1}^{p}(x - \lambda_i)^{m_i}$. The derivative of $Q(x)$, $\prod_{i=1}^{p}(x - \lambda_i)^{m_i}\sum_{i=1}^{p}(m_i/(x - \lambda_i))$ has a zero of multiplicity $m_i - 1$ at λ_i for $i = 1,\dots,p$ and a zero in each of the $p - 1$ intervals (λ_1, λ_2), $(\lambda_2, \lambda_3),\dots,(\lambda_{p-1}, 1)$. Thus (3.3), as a function of $x = 1/z$, has a zero in each of $p - 1$ intervals $(\lambda, \lambda_2),\dots,(\lambda_{p-1}, 1)$.

The proof of the proposition is now reduced to the following problem. For $z \in C$, $|z| \geq 1$, for $F_{ij}(z)$ to exist for all i and j, it is sufficient that the function

$$H(x) = \prod_{i=1}^{p}(x - \lambda_i)\sum_{i=1}^{p}\frac{m_i}{x - \lambda_i}$$

have no zeros in the intervals $[-1, -1/|z|]$ and $[1/|z|, 1]$.

Under condition (i) of Proposition 3.2 the only zero $H(x)$ could have in the interval $[1/|z|, 1]$ is the zero between λ_{p-1} and 1. Call this zero x_0. $H(x)$, and hence

$$J(x) = (x - 1)\sum_{i=1}^{p}\frac{m_i}{x - \lambda_i} = 1 + (x - 1)\sum_{i=1}^{p-1}\frac{m_i}{x - \lambda_i},$$

has only one sign change in $(\lambda_{p-1}, 1)$. $J(x)$ clearly does not change sign to the right of 1 and is positive to the right of 1. Thus if $J(x) > 0$ and $x > \lambda_{p-1}$, then $x > x_0$. Condition (iii) asserts the positivity of $J(1/|z|)$, implying $1/|z| > x_0$, implying $H(x)$ has no zeros in $[1/|z|, 1)$. Conditions (ii) and (iv) similarly imply that $H(x)$ has no zeros in $(-1, -1/|z|]$. Thus together (i)–(iv) imply $F_{ij}(z)$ exists for all i, j. □

4. Examples on the symmetric group. The symmetric group Σ_N is the group of all $N!$ permutations of N elements. Naimark and Stern (1982) and Flatto, Odlyzko and Wales (1985) are good references on analysis on Σ_N. We will only consider covering problems for random walks on Σ_N that are constant on conjugacy classes. The first sets to be visited will be all $N!$ singletons. The time to visit all these singletons and the number unvisited after a large number of steps will be discussed. Then more complex covering problems will be used to study the spatial distribution of unvisited permutations. For any fixed N, the bounds in Theorem 1.3 are quite messy. As $N \to \infty$ precise asymptotic results are available, and only these will be given.

The conjugacy classes of Σ_N are determined by cycle structure. Let $n_i(\pi)$ denote the number of i-cycles of a permutation π for $i = 1, \dots, N$. Then π_1 and π_2 are conjugate if $n_i(\pi) = n_i(\pi_2)$ for all $i = 1, 2, \dots, N$. As usual, denote the conjugacy class with n_1 1-cycles, $\dots$, n_N N-cycles by $1^{n_1}2^{n_2} \cdots N^{n_N}$. For example, the identity permutation is the only member of the class 1^N and any transposition is in the class $1^{N-2}2^1$.

Any permutation can be written as a product of transpositions. Unlike cycle structure, this decomposition is nonunique. However, any permutation must be either even (a product of an even number of transpositions) or odd (a product of an odd number of transpositions). In fact, there is a one-dimensional irreducible representation of Σ_N, denoted Alt, with

$$\text{Alt}(\pi) = 1, \quad \text{if } \pi \text{ is even}$$

and

$$\text{Alt}(\pi) = -1, \quad \text{if } \pi \text{ is odd.}$$

$\text{Alt}(\pi) = \text{Alt}(\sigma^{-1}\pi\sigma)$ so all members of a conjugacy class are simultaneously even or odd. Thus conjugacy classes can be called even or odd.

Consider probability measures on Σ_M of the following form. Let $L_1, \dots, L_J$ be conjugacy classes of Σ_M. Choose $\mu_1, \dots, \mu_J$ with $\mu_j > 0$ for $j = 1, \dots, J$ and $\sum_{j=1}^{J} \mu_j = 1$. Let μ put mass $\mu_j/|L_j|$ on each member of L_j for $j = 1, \dots, J$. To have the support of μ generate Σ_M it is necessary and sufficient that μ puts positive mass on at least one odd conjugacy class. Suppose μ puts total mass $q > 0$ on odd conjugacy classes. Also suppose μ puts mass $p \geq 0$ on the identity.

To give clean asymptotics, a sequence of random walks on Σ_N, as $N \to \infty$ must be defined. The measure μ defined on Σ_M can be naturally extended to a measure on Σ_N for $N > M$ by taking each conjugacy class L_i of Σ_M and padding it with $N - M$ one cycles, making it a conjugacy class of Σ_N. For example, if μ originally defined on Σ_4 puts total mass 1 on the $\binom{4}{2}$ transpositions of Σ_4, then the version of μ defined on Σ_6 would also put total mass 1 on the $\binom{6}{2}$ transpositions of Σ_6.

The analysis of hitting times on Σ_N needed here has been done in Flatto, Odlyzko and Wales (1985), hereafter abbreviated FOW. The following proposition is an easy consequence of their work.

PROPOSITION 4.1. *For a sequence of random walks on symmetric groups as described above, the time to hit the identity from a group element $g \neq \text{id}$ has moment generating function F_g defined by*

$$E_g \exp\left(\frac{s(1-p)}{N!} T_{\text{id}}\right) = F_g\left(\exp\left(\frac{s(1-p)}{N!}\right)\right) = \frac{1}{1-s} + O(N^{-2}),$$

uniformly in g for $s < 1$.

Proof. The proof for $s < 0$ for a measure concentrated on a single conjugacy class is given by FOW in the proof of their Theorem 5.4. As they mention, the proof extends easily to the types of measures considered here.

The case $0 < s < 1$ requires only slightly more care. The eigenvalue of the transition matrix corresponding to the alternating representation is $1 - 2q$. The eigenvalue corresponding to the trivial representation of μ is 1. By (5.4) of FOW all other eigenvalues of μ satisfy $1 - |\lambda| \geq \theta/N^{\theta}$ for some $\theta > 0$ depending on μ. Thus other eigenvalues of the transition matrix approach one in magnitude at a rate that is at most polynomial in N. The existence of F_g at $\exp(s(1-p)N!) = 1 + O(N!^{-1})$ needs to be shown. Requirements (i) and (ii) of Proposition 3.2 are clearly satisfied. Requirements (iii) and (iv) are implicit in the calculation of (5.27) in FOW. □

From Proposition 4.1 both $f^{+}(s(1-p)/N!)$ and $f^{-}(s(1-p)N!)$ exist and are $(1-s)^{-1} + O(N^{-2})$. Write

$$1/f^{+}(s(1-p)/N!) = 1 - s + O^{+}$$

and

$$1/f^{-}(s(1-p)/N!) = 1 - s + O^{-},$$

where O^{+} and O^{-} are different $O(N^{-2})$ terms.

Consider the time Q_n taken by this kind of random walk on Σ_N to visit n of the $N! - 1$ permutations excepting the starting point. From Theorem 1.3

$$\begin{aligned}\prod_{i=1}^{N!-1-n} \frac{i - s + O^{+}}{i - s + O^{-}} \prod_{i=N!-n}^{N!-1} \frac{i}{i - s + O^{-}} \\ \leq E \exp\left(\frac{s(1-p)Q_n}{N!}\right) \\ \leq \prod_{i=1}^{N!-1-n} \frac{i - s + O^{-}}{i - s + O^{+}} \prod_{i=N!-n}^{N!-1} \frac{i}{i - s + O^{+}},\end{aligned}$$

for all s for which $f^{+}(s(1-p)/N!)$ exists. In terms of gamma functions this is

$$\begin{aligned}&\frac{\Gamma(N! - n - s + O^{+})}{\Gamma(N! - n)} \frac{\Gamma(1 - s + O^{-})}{\Gamma(1 - s + O^{+})} \frac{\Gamma(N!)}{\Gamma(N! - s + O^{+})} \\ &\leq E \exp\left(\frac{s(1-p)Q_n}{N!}\right) \qquad (4.2) \\ &\leq \frac{\Gamma(N! - n - s + O^{-})}{\Gamma(N! - n)} \frac{\Gamma(1 - s + O^{+})}{\Gamma(1 - s + O^{-})} \frac{\Gamma(N!)}{\Gamma(N! - s + O^{-})}.\end{aligned}$$

First let $n = N! - 1$, so the quantity under consideration is the time taken by the random walk to visit every member of Σ_N. Stirling's formula yields

$$\frac{\Gamma(N!)N!^{-s}}{\Gamma(N! - s + O(N^{-2}))} \to 1$$

and

$$\frac{\Gamma(1-s+O^+)}{\Gamma(1-s+O^-)} \to 1, \quad \text{for } s < 1 \text{ as } N \to \infty.$$

The bounds in (4.2) become

$$E\exp\left(\frac{s(1-p)}{N!}\left(Q_{N!-1} - \frac{N!\log N!}{1-p}\right)\right) \to \Gamma(1-s), \tag{4.3}$$

for $s < 1$ as $N \to \infty$. Recognizing $\Gamma(1-s)$ as the moment generating function of the extreme value distribution yields

$$P\left(Q_{N!-1} < \frac{N!}{1-p}(\log N! + c)\right) \to e^{-e^{-c}}, \quad \text{as } N \to \infty. \tag{4.4}$$

If $p = 0$, then the random walk moves nontrivially at every step. A comparison of (1.1) with (4.3) shows that the time taken by a random walk of this form to visit every group element is asymptotically the same as if permutations were selected uniformly and independently.

Next let $n = N! - 1 - J$ for a fixed J. Another calculation with (4.2) gives

$$E\exp\left(\frac{s(1-p)}{N!}\left(Q_{N!-1-J} - \frac{N!\log N!}{1-p}\right)\right) \to \frac{\Gamma(J+1-s)}{\Gamma(J+1)},$$

for $s < 1$. With U_c as in (1.2) the number of unvisited permutations after $N!/(1-p)(\log N! + c)$ steps, (1.6) gives

$$P(U_c = J) \to \frac{1}{2\pi i}\int_{d-i\infty}^{d+i\infty} \frac{\lambda^z}{z}\left(\frac{\Gamma(J+1-z)}{\Gamma(J+1)} - \frac{\Gamma(J-z)}{\Gamma(J)}\right) dz, \tag{4.5}$$

for any $d \in (0,1)$ and $J > 0$ where $\lambda = e^{-c}$. Let $y = 1 - J + z$. Then (4.5) is

$$\frac{1}{2\pi i}\frac{\lambda^{J-1}}{J!}\int_{d-i\infty}^{d+i\infty} \lambda^y \Gamma(1-y)\, dy, \tag{4.6}$$

for any $d \in (0,1)$. Except for the term $\lambda^{J-1}/J!$, (4.6) is the extreme value density at $-\log\lambda$, which is $\lambda e^{-\lambda}$. Thus

$$P(U_c = J) \to \frac{\lambda^J}{J!}e^{-\lambda}, \quad \text{for all } J > 1, \tag{4.7}$$

and for $J = 0$ as well from (4.4).

Therefore U_c has a limiting Poisson distribution for any fixed c. Specializing again to the case $p = 0$ and comparing (4.7) with (1.2). says that if the random walk moves nontrivially at each step, then the number of unvisited permutations after a large number of steps has asymptotically the same distribution as if permutations has been picked independently and uniformly.

Finally, consider (1.8). Theorem 1.3 can be used by bounding C_n above and below by other random variables. Consider the set A_N of all $\frac{1}{2}N!\binom{N}{2}$ pairs of permutations of Cayley distance one apart. Again $T(A_N)$ denotes the first time all the members of A_N have been visited. Let B_N denote the set of $\frac{1}{2}N! - 1$ pairs of permutations

$$\{(\pi, \pi \cdot (12)) : \pi \text{ is even}, \pi \neq a_0 \text{ and } \pi \cdot (12) \neq x_0\}.$$

The members of B_N are disjoint. $T(B_N)$ is the first time all members of B_N have been visited. Then

$$T(B_N) \leq C_N \leq T(A_N).$$

Theorem 1.3 with $N = 2$ and $n = 1$ shows that the time to hit a pair of permutations from any point outside the pair has moment generating function

$$f(s/N!) = \frac{1}{1 - \dfrac{s}{2(1-p)} + O(N^{-2})}.$$

Applying (1.4) yields a lower bound on $T(B_N)$ and an upper bound on $T(A_N)$, giving

$$\begin{aligned} &\frac{\Gamma\left(\frac{1}{2}N!\right)\Gamma\left(1 - \dfrac{s}{2(1-p)} + O(N^{-2})\right)}{\Gamma\left(\frac{1}{2}N! - \dfrac{s}{2(1-p)} + O(N^{-2})\right)} \\ &\quad \leq E_{a_0}\exp\left(\frac{sC_N}{N!}\right) \\ &\quad \leq \frac{\Gamma\left(\frac{1}{2}N!\binom{N}{2} + 1\right)\Gamma\left(1 - \dfrac{s}{2(1-p)} + O(N^{-2})\right)}{\Gamma\left(\frac{1}{2}N!\binom{N}{2} + 1 - \dfrac{s}{2(1-p)} + O(N^{-2})\right)}, \end{aligned} \tag{4.8}$$

$$\text{for } s < \frac{1}{1-p}.$$

The left-hand side of (4.8) is $(\frac{1}{2}N!)^{s/2(1-p)}(1 + O(1))$, and the right-hand side of (4.8) is $\left(\frac{1}{2}N!\binom{N}{2}\right)^{s/2(1-p)}(1 + O(1))$. Markov's inequality for s positive and negative applied to $\exp(sC_N/N!)$ yields

$$P\left(\left|\frac{C_N}{N!} - \frac{\log N!}{2(1-p)}\right| > \log^2 N\right) \to 0, \quad \text{as } N \to \infty.$$

In the case considered in the Introduction $p = 0$ and (1.9) follows.

It is interesting to wonder how far the similarity between the coupon collector's problem and rapidly mixing random walks holds up. Here it has been shown that in what Kolchin, Sevast'yanov and Chistyakov (1978) call the right-hand domain, the distribution of the number of empty cells is asymptotically the same, at least in this case. Whether the relationship also holds in the central domain and whether finer aspects of the right-hand domain, such as the relative locations of unvisited permutations are the same in both problems, are interesting questions.

Acknowledgment. Most of this work is a part of the author's doctoral dissertation, supervised by Professor Persi Diaconis, and he would like to thank Professor Diaconis for a wealth of advice and encouragement.

REFERENCES

ALDOUS, D. J. (1983). On the time taken by random walks on finite groups to visit every state. *Z. Wahrsch. verw. Gebiete* **62** 361–393.

DIACONIS, P. (1988). *Group Theory in Statistics*. IMS, Hayward, Calif. To appear.

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

FELLER, W. (1968). *An Introduction to Probability Theory and Its Applications* **1**, 3rd ed. Wiley, New York.

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

KOLCHIN, V. F., SEVAST'YANOV, B. A. and CHISTYAKOV, V. P. (1978). *Random Allocations*. V. H. Winston and Sons, Washington.

MATTHEWS, P. (1985). Covering problems for random walks on spheres and finite groups. Ph.D. thesis, Stanford Univ.

MATTHEWS, P. (1988). Covering problems for Brownian motion on spheres. *Ann. Probab.* **16** 189–199.

NAIMARK, M. and STERN, A. (1982). *Theory of Group Representations*. Springer, Berlin.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND BALTIMORE COUNTY
CATONSVILLE, MARYLAND 21228