EIGEN ANALYSIS FOR SOME EXAMPLES OF THE METROPOLIS ALGORITHM

BY

PERSI DIACONIS and PHIL HANLON

TECHNICAL REPORT NO. 392
MARCH 1992

PREPARED UNDER THE AUSPICES OF
NATIONAL SCIENCE FOUNDATION GRANT DMS89-05874

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
Eigen Analysis for Some Examples of the Metropolis Algorithm

By

Persi Diaconis and Phil Hanlon

Technical Report No. 392

March 1992

Prepared Under the Auspices

Of

National Science Foundation Grant DMS89-05874

Department of Statistics
Stanford University
Stanford, California
EIGEN ANALYSIS FOR SOME EXAMPLES OF THE METROPOLIS ALGORITHM

Persi Diaconis
Dept. of Mathematics
Harvard University
Cambridge, MA 02138

Phil Hanlon
Department of Mathematics
University of Michigan
Ann Arbor, Michigan 48109-1003

ABSTRACT

The Metropolis algorithm allows us to sample from a given probability distribution by running a Markov chain. We derive the eigenvalues for a class of simple chains. These appear to be the first examples where such explicit computation is possible. They thus allow us to compare exact results with currently available bounds. The eigenvalues turn out to be related to families of orthogonal and symmetric polynomials; as one varies the temperature in the Metropolis algorithm one runs through the natural parameter in the family.
1. Introduction.

Let $X$ be a finite set and let $\pi(x)$ be a probability distribution on $X$ with $\pi(x) > 0$ for all $x \in X$. The Metropolis algorithm is a strategy for sampling from $\pi$ which is effective when $\pi$ is only known up to a norming constant which is difficult to compute because $|X|$ is large. The algorithm proceeds by running a Markov chain with transition $M(x, y)$ and stationary distribution $\pi$. The chain is constructed by "thinning down" an easy to run "base chain". A more careful description is given below. The Metropolis algorithm is very widely used in statistical mechanics (Ising simulations), statistics, and as an ingredient in simulated annealing. Little is known about its non-asymptotic rate of convergence to stationarity.

This paper, along with Hanlon (1992), gives a class of examples where eigenvalues and sharp computations for rates of convergence can be derived. The examples involve natural random walks on a group "thinned down" by a natural distance function. The examples are interesting in two directions. They seem to be the first examples where explicit computation can be carried out. They can thus serve as test problems for bounds on rates of convergence. Further, they produce classical families of 1-parameter orthogonal polynomials such as Krawtchouck polynomials and symmetric functions such as Jack symmetric functions.

The Metropolis algorithm is explained more carefully in section 2. Our most interesting example appears in section 4. This analyzes a process on the symmetric group. Here, the base chain is based on randomly transposing pairs of cards. The chain is "thinned down" according to its distance to the identity using $d$ – the minimum number of transpositions metric. The resulting chain has stationary distribution proportional to $\theta^d(\pi, \pi_0)$, a measure that arises in statistical applications. The eigenvalues are the change of basis coefficients when the Jack symmetric functions are expanded in the power sum symmetric functions. Using recent results of Stanley and Macdonald, these eigenvalues are explicitly available. We determine the convergence properties and study the dependence on $\theta$.

Section 3 carries out an easier analysis for nearest neighbor walk on the cube $\mathbb{Z}_2^d$, thinning down with Hamming distance. Now the one-parameter family of Krawtchouck
polynomials appears in the eigen analysis.

In both cases, the chains are rapidly mixing: the Metropolis thinning only changes the eigenvalues in a linear way. In contrast, use of off the shelf bounds leads to exponentially perturbed eigenvalues and a poor picture of convergence rates.

The final section carries out this analysis for a twisted Markov chain on the space of matchings. Again, exact analysis and interesting special functions appear.
2. The Metropolis algorithm and eigen analysis of reversible Markov chains.

This section contains a careful description of the Metropolis algorithm along with background on the use of eigenvalues to study total variation convergence for reversible Markov chains.

Let $X$ be a finite set. Let $\pi(x)$ be a positive probability on $X$. Often, practical considerations allow easy access to the ratios $r_{yx} = \pi(y)/\pi(x)$. Let $S(x, y)$ be the transition matrix of a symmetric irreducible Markov chain on $X$. This is the base chain which is assumed to be easy to run. Let $M(x, y)$ be defined by “thinning down” $S(x, y)$ according to the following

\[
M(x, y) = \begin{cases}
S(x, y)r_{yx} & \text{if } r_{yx} < 1 \\
S(x, y) & \text{if } x \neq y \text{ and } r_{yx} \geq 1 \\
S(x, x) + \sum_{z \neq x} S(x, z)(1 - r_{zy}) & \text{if } x = y.
\end{cases}
\]

This definition has a simple implementation. If the chain is at $x$, pick $y$ with probability $S(x, y)$. If $y \neq x$ and $r_{yx} \geq 1$, the chain moves to $y$. If $y \neq x$ and $r_{yx} < 1$, flip a coin with success probability $r_{yx}$. If the coin toss succeeds, the chain moves to $y$. In all other cases the chain stays at $x$.

It is straightforward to show that $M(x, y)$ defines an irreducible aperiodic transition matrix with stationary distribution $\pi$. For further details, see e.g. Hammersley and Handscomb (1964).

The Metropolis chain $M$ is reversible: $\pi(x)M(x, y) = \pi(y)M(y, x)$. To conclude this section we derive bounds on rates of convergence for a general reversible chain in terms of eigenvalues. Thus for the remainder of this section, let $M, \pi$ be a reversible Markov chain on a finite set $X$. Let $D$ be a diagonal matrix having $\sqrt{\pi(x)}$ down the diagonal. Reversibility yields that $T = DMD^{-1}$ is symmetric. Thus $T$ can be orthogonally diagonalized $T = \Gamma\beta\Gamma^t$ with $\Gamma$ orthogonal and $\beta$ a diagonal matrix having the eigenvalues of $T$, and so $M$, on the diagonal. Thus $M = V\beta V^{-1}$ with $V = D^{-1}\Gamma, V^{-1} = \Gamma^tD$. This implies that the right eigenvectors of $M$ are the columns of $V$: $V_{xy} = \Gamma_{yx}/\sqrt{\pi(x)}$. These are orthonormal in $L^2(\pi)$. The left eigenvectors are the rows of $V^{-1}$: $V_{xy}^{-1} = \Gamma_{yx}\sqrt{\pi(y)}$. These are orthonormal in $L^2(1/\pi)$. 
Define total variation distance as

\[ \|M^k(x, \cdot) - \pi(\cdot)\| = \frac{1}{2} \sum_y |M^k(x, y) - \pi(y)|. \]

The following lemma gives bounds on total variation in terms of eigenvalues.

**Lemma 1.** Let \( M, \pi \) be a reversible Markov chain on a finite set \( X \). Let \( \beta_y \) denote the eigenvalues, \( \beta^* \) the second largest eigenvalue in absolute value. Let \( f_y(\cdot) \) be an orthonormal basis of right eigenfunctions in \( L^2(\pi) \). Let \( g_y(\cdot) \) be an orthonormal basis of left eigenfunctions in \( L^2(\frac{1}{\pi}) \). Then, for any starting state \( x \), the total variation \( 4\|M(x, \cdot) - \pi(\cdot)\|^2 \) is bounded above by any of the following 3 quantities

\begin{align*}
(2.2) & \quad \sum_y \beta^2_y f^2_y(x) - 1 \\
(2.3) & \quad \frac{1}{\pi^2(x)} \sum_y \beta^2_y g_y(x) - 1 \\
(2.4) & \quad \frac{1}{\pi(x)} (\beta^*)^{2k}
\end{align*}

**Proof:** For any fixed \( x \)

\[
\left( \sum_y |M^k(x, y) - \pi(y)| \right)^2 = \left( \sum_y \frac{|M^k(x, y) - \pi(y)|}{\sqrt{\pi(y)}} \sqrt{\pi(y)} \right)^2 \leq \sum_y \frac{|M^k(x, y) - \pi(y)|^2}{\pi(y)}
\]

\[
= \sum_y \left( \frac{M^k(x, y)}{\pi(y)} \right)^2 - 1 = \frac{M^{2k}(x, x)}{\pi(x)} - 1
\]

\[
= \frac{1}{\pi(x)} \sum_y \beta^2_y \Gamma^2_{xy} - 1.
\]

The inequality above is Cauchy-Schwarz. The next to last equality used \( \pi(x)M(x, y) = \pi(y)M(y, x) \). The bounds now follow by relating \( \Gamma_{xy} \) to left or right eigenvectors (multiplying or dividing by \( \sqrt{\pi(x)} \)). The final bound follows by bounding \( \beta_y \) by \( \beta^* \).

The Metropolis chains in this paper are all built by thinning down a simple chain with known spectral properties. It is natural to try to compare the eigenvalues. This can be achieved by comparing the Dirichlet forms. For a reversible chain \( M, \pi \), define

\[ \mathcal{E}(f, f) = \frac{1}{2} \sum_{x,y} (f(x) - f(y))^2 \pi(x)M(x, y). \]
This is just the quadratic form defined by I-M. The classical minimax characterization of eigenvalues yields

\[ 1 - \beta_i = \min_{W_i} \min_{f \in W_i} \frac{\mathcal{E}(f, f)}{\|f\|_W^2} \quad \text{dim } W_i = i + 1. \]

Here \( \beta_i \) is the \( i \)th largest eigenvalue \( 1 = \beta_0 > \beta_1 \geq \beta_2 \cdots \geq \beta_{|X|-1} \geq -1 \). This renders the following comparison lemma evident.

**Lemma 2.** Let \((M, \pi), (\tilde{M}, \tilde{\pi})\) be reversible Markov chains on a finite set \( X \). Suppose the associated Dirichlet forms satisfy

\[ \tilde{\mathcal{E}} \leq A\mathcal{E} \quad \text{and} \quad \tilde{\pi} \geq a\pi \quad \text{for positive } A, a. \]

Then

\[ \beta_i \leq 1 - \frac{a}{A}(1 - \tilde{\beta}_i). \]

As will emerge, these bounds can be very far off in the Metropolis setting. For lower bounds, and examples where comparison is useful, see Diaconis and Saloff-Coste (1992).
3. A simple example: nearest neighbor walk on $\mathbb{Z}_2^d$.

Let $\mathbb{Z}_2^d$ be the group of binary $d$-tuples under coordinatewise addition, thought of as the vertices of a cube in $d$-dimensions. Nearest neighbor walk has transition matrix

$$ S(x, y) = \begin{cases} 
\frac{1}{d} & \text{if } H(x, y) = 1 \\
0 & \text{otherwise}
\end{cases} $$

with $H(x, y)$ the Hamming distance: the number of coordinates where $x$ and $y$ disagree. This walk has been extensively analyzed because it gives a representation of the Ehrenfests urn. See Kac (1947), Letac and Tackas (1979), or Diaconis (1988) and references cited there. In particular, the matrix $S$ is diagonalizable with eigenvalues $\left(1 - \frac{2j}{d}\right)$ occurring with multiplicity $\binom{d}{j}$, $0 \leq j \leq d$, and eigenvectors given by the characters of $\mathbb{Z}_2^d$. Note this walk has periodicity problems since $-1$ is an eigenvalue. These will disappear for the Metropolis walk.

Consider next the family of measures on $\mathbb{Z}_2^d$ defined by

$$ P_\theta(x) = \frac{\theta^{H(x)}}{(1 + \theta)^d} $$

with $0 < \theta \leq 1$, $H(x) = H(x, 0)$. For this example, there is no problem simulating from $P_\theta$ by a variety of schemes. For example, coordinates of $x$ can be chosen as independent 1 or 0 with probability $\frac{\theta}{1 + \theta}$ and $\frac{1}{1 + \theta}$ respectively. This takes order $d$ operations.

The Metropolis algorithm can be used to generate from $P_\theta$. Starting from $S(x, y)$ at (3.1), the recipe (2.1) results in a chain with

$$ M(x, y) = \begin{cases} 
\frac{1}{d} & \text{if } H(x, y) = 1, \ H(y) < H(x) \\
\frac{\theta}{d} & \text{if } H(x, y) = 1, \ H(y) > H(x) \\
\left(1 - \frac{H(x)}{d}\right)(1 - \theta) & \text{if } H(x, y) = 0 \\
0 & \text{otherwise.}
\end{cases} $$

For example, when $d = 2$, $M$ appears as

$$
\begin{pmatrix}
00 & 01 & 10 & 11 \\
00 & \begin{pmatrix} 1 - \theta & \theta \\ \theta & \frac{1}{2} \end{pmatrix} & \begin{pmatrix} \theta \\ \frac{1}{2} \end{pmatrix} & \begin{pmatrix} \theta \\ \frac{1}{2} \end{pmatrix} \\
01 & \begin{pmatrix} 1 - \theta & \theta \\ \frac{1}{2} & \frac{1 - \theta}{2} \end{pmatrix} & \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix} & \begin{pmatrix} \theta \\ \frac{1}{2} \end{pmatrix} \\
10 & \begin{pmatrix} \theta \\ \frac{1}{2} \end{pmatrix} & \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix} & \begin{pmatrix} 1 - \theta & \theta \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \\
11 & \begin{pmatrix} \theta \\ \frac{1}{2} \end{pmatrix} & \begin{pmatrix} \theta \\ \frac{1}{2} \end{pmatrix} & \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix}
\end{pmatrix}
$$

7
The permutation group \( S_d \) operates on \( \mathbb{Z}_d^d \). It is clear that \( M(\pi x, \pi y) = M(x, y) \) for \( \pi \in S_d \). This implies that the orbit chain, which just records \( H(x) \) when the chain is in \( x \), is again a Markov chain. This chain takes values in \( (0, 1, \cdots, d) \) with transition matrix

\[
(3.4) \quad m(i, j) = \begin{cases} \frac{i}{d} & \text{if } j = i - 1 \\ (1 - \frac{i}{d})\theta & \text{if } j = i + 1 \\ (1 - \frac{i}{d})(1 - \theta) & \text{if } j = i. \end{cases}
\]

For example, when \( d = 2 \), \( m \) appears as

\[
\begin{pmatrix}
0 & 1 & 2 \\
0 & 1 - \theta & \theta & 0 \\
1 \frac{1}{2} & \frac{1-\theta}{2} & \frac{\theta}{2} & 0 \\
2 & 0 & 1 & 0
\end{pmatrix}.
\]

In general, \( m \) has \((1 - \frac{i}{d})(1 - \theta)\) down the diagonal, \(0 \leq i \leq d; \frac{i}{d}\) below the diagonal \(1 \leq i \leq d; (1 - \frac{i}{d})\theta\) above the diagonal \(0 \leq i \leq d - 1\); and zeros elsewhere. This \( m \) has stationary distribution \( \pi(i) = \frac{\theta^i}{(1+\theta)^d} \binom{d}{i} \).

**Theorem 1.** The matrix \( m \) defined at (3.4) has eigenvalues

\[
\beta_i = 1 - \frac{i}{d}(1 + \theta), \quad 0 \leq i \leq d.
\]

The corresponding right eigenvector is the Krawtchouck polynomial

\[
P_i(j) = \left( \theta^i \binom{d}{i} \right)^{-1/2} \sum_{k=0}^{i} (-1)^k \binom{i}{k} \binom{d-j}{i-k} \theta^{i-k}.
\]

These have been normalized to be orthonormal in \( L^2(\pi) \).

**Proof:** This is essentially contained in Krawtchouck (1929). The statement is easy to verify from known properties of Krawtchouck polynomials as in McWilliams and Sloane (1977, p. 150) \( \square \)

For example, when \( d = 2 \), the eigenvalues are \( 1, 1 - \frac{(1+\theta)}{2}, -\frac{\theta}{2} \), with eigenvectors

\[
\begin{pmatrix}
1 \\
1 \\
1
\end{pmatrix}, \quad \frac{1}{\sqrt{2\theta}} \begin{pmatrix} 2\theta \\ \theta - 1 \\ -2 \end{pmatrix}, \quad \frac{1}{\theta} \begin{pmatrix} \theta^2 \\ -\theta \\ 1 \end{pmatrix}.
\]

The main point is that thinning down by distance leads to an interesting deformation of the eigenvalues and eigenvectors. The next result converts this into a rate of convergence result for the original chain \( M \).
THEOREM 2. Fix $\theta$ in $(0, 1)$. For the Metropolis chain $M$ at (3.3) started at 0, let $k = \frac{d}{2(1+\theta)}(\log d\theta + c)$. Then

$$\|M^k(0, \cdot) - P_\theta(\cdot)\| \leq f(\theta, c)$$

with $f(\theta, c)$ independent of $d$, tending to zero as $c \to \infty$, for each fixed $\theta$.

PROOF: Because of invariance, the distance $\|M^k(0, \cdot) - P_\theta(\cdot)\| = \|m^k(0, \cdot) - \pi_\theta(\cdot)\|$ where the orbit chain is defined at (3.4). Using Theorem 1 and the upper bound (2.2),

$$4\|m^k(0, \cdot) - \pi_\theta(\cdot)\|^2 \leq \sum_{j=1}^{d} \left(1 - \frac{j(1+\theta)}{d}\right)^{2k} \theta^j \binom{d}{j}.$$  (3.5)

Break the sum in (3.5) into $S_1 + S_2$ with $S_1$ summed over $1 \leq j \leq d/(1+\theta)$ and $S_2$ summed over $d/(1+\theta) < j \leq d$. For $S_1$, use $1 - x \leq e^{-x}$, $\binom{d}{j} \leq d^j / j!$, and the definition of $k$ to conclude

$$S_1 \leq \sum_{j=1}^{\infty} \frac{e^{-c}}{j!} = e^{e^{-c}} - 1.$$  

For $S_2$, replace $j$ by $d - \ell$ and use the inequalities above to conclude

$$S_2 \leq \theta^{d+2k} \sum_{\ell=1}^{d(1+\theta)} \frac{1}{\ell!} \exp\left\{-\frac{\ell}{\theta} (\log(d\theta) + c) + \ell \log(d\theta)\right\}.$$  

Now $\theta < 1$ yields $\frac{\ell}{\theta} \log d + \ell \log d < 0$. Making this replacement,

$$S_2 \leq \sum_{\ell=1}^{\infty} \frac{1}{\ell!} \exp\left\{-\ell (\log \theta^{\frac{d+1}{\theta}} + \frac{c}{\theta})\right\}.$$  

The bounds for $S_1$ and $S_2$ sum to give $f(\theta, c)$ with the stated properties. \qed

REMARKS: 1. The bound is sharp in the sense that if $k = \frac{d}{2(1+\theta)}(\log d\theta - c)$ the variation distance does not tend to zero. This can be shown by using the first eigenfunction as in Diaconis (1988, Chapter 3).

2. Theorem 2 shows that the Metropolis algorithm is only "off by a log". That is, its running time (order $d \log d$) is comparable to the optimal algorithm (order $d$) which uses the structure of the normalizing constant. Of course, the function $H(x)$ is relatively simple.
3. It is instructive to compare the second eigenvalue $1 - \frac{(1+\theta)}{d}$ with what results from comparison of Dirichlet forms (Lemma 2 of Section 2). Passing to the chain $m$ and comparing with the unthinned base chain ($\theta = 1$) requires $a = \left(\frac{1+\theta}{\theta}\right)^d$ and $A = \frac{1}{\theta} \left(\frac{1+\theta}{\theta}\right)^d$. This yields

$$\beta_i \leq 1 - \theta^{d+1}(1 - \beta_i).$$

This is off by an exponential factor and virtually useless in practice.
4. Simulating from a distribution on the symmetric group.

We begin with some motivation for the example to be studied. Statisticians sometimes have to work with ranked data as when a group of people are each asked to rank order 5 wines. To facilitate data analysis, a variety of metrics between permutations are employed. For example, the Cayley distance is defined as

\[ d(\pi, \sigma) = \text{minimum number of transpositions required to bring } \pi \text{ to } \sigma, \text{ for } \pi, \sigma \in S_n. \]

This is named after Cayley who discovered \( d(\pi, \sigma) = n - C(\pi \sigma^{-1}) \), with \( C(\tau) \) the number of cycles in \( \tau \). One use of such metrics is to build probability distributions on \( S_n \)

\[ P_\theta(\pi) = c(\theta) \theta^{d(\pi, \pi_0)} \]

where \( 0 < \theta \leq 1 \), \( c^{-1}(\theta) = \sum_{\pi} \theta^{d(\pi, \pi_0)} = \prod_{i=1}^{n} (1 + \theta(i - 1)) \) is a normalizing constant, and \( \pi_0 \) is a "location parameter". This model describes a population peaked about \( \pi_0 \) which falls off geometrically at rate \( \theta \). For \( \theta = 1 \), \( P_1 \) becomes the uniform distribution. In statistical applications, \( \pi_0 \) and \( \theta \) would be unknown and estimated from a sample of rankings. Further background can be found in Critchlow (1985) or Diaconis (1988, Chapter 6).

In any modern statistical work, the ability to sample efficiently from a probability distribution is crucial. In this section we analyze the Metropolis algorithm for this problem and compare it with the best available alternative. Without essential loss, take \( \pi_0 = id \) throughout.

The Metropolis algorithm will be based on repeated random transpositions. Thus, consider the Markov chain on \( S_n \) with transition matrix

\[ S(\sigma, \tau) = \begin{cases} 1/(\binom{n}{2}) & \text{if } \tau = \sigma(i, j) \text{ for some } i < j \\ 0 & \text{otherwise.} \end{cases} \]

The chain \( S \) was analyzed by Diaconis and Shahshahani (1981). As given, it has periodicity problems which will disappear when it is thinned down. Define a Metropolis chain as in (2.1) with

\[ r(\sigma, \tau) = \theta^{C(\sigma) - C(\tau)}. \]
When \( n = 3 \), the transition matrix becomes

\[
M_3^\theta = \begin{pmatrix}
\text{id} & (12) & (13) & (23) & (123) & (132) \\
\text{id} & \left(1 - \theta, \frac{\theta}{3}, \frac{\theta}{3}, \frac{\theta}{3}, 0, 0\right) \\
(12) & \left(\frac{1}{3}, \frac{2}{3}(1 - \theta), 0, 0, \frac{\theta}{3}, \frac{\theta}{3}\right) \\
(13) & \left(\frac{1}{3}, 0, \frac{2}{3}(1 - \theta), 0, \frac{\theta}{3}, \frac{\theta}{3}\right) \\
(23) & \left(\frac{1}{3}, 0, 0, \frac{2}{3}(1 - \theta), \frac{\theta}{3}, \frac{\theta}{3}\right) \\
(123) & \left(0, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0\right) \\
(132) & \left(0, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0\right)
\end{pmatrix}
\]

The stationary distribution is the left eigenvector proportional to

\[(1, \theta, \theta, \theta, \theta^2, \theta^2).
\]

By construction, the matrix \( M_n^\theta \) commutes with the action of \( S_n \) on itself by conjugation. This implies that if \( \pi \) and \( \sigma \) are conjugate in \( S_n \), \( M^k(\text{id}, \pi) = M^k(\text{id}, \sigma) \) for all \( k = 1, 2, \ldots \). Thus the chain lumped to conjugacy classes is Markov. When \( n = 3 \), this lumped chain has transition matrix \( m_3^\theta \)

\[
m_3^\theta = \begin{pmatrix}
1^3 & 1, 2 & 3 \\
1 & \left(1 - \theta, \theta, 0\right) \\
1, 2 & \left(\frac{1}{3}, \frac{2}{3}(1 - \theta), \frac{\theta}{3}\right) \\
3 & \left(0, 1, 0\right)
\end{pmatrix}
\]

This has stationary distribution proportional to \((1, 3\theta, 2\theta^2)\). In general, the conjugacy classes are indexed by partitions of \( n \). The transition matrix for the lumped chain has

\[
m_n^\theta(\lambda, \mu) = \sum M_n^\theta(\pi, \sigma)
\]

summed over all \( \sigma \) in the conjugacy class \( \mu \) with \( \pi \) any permutation in conjugacy class \( \lambda \).

Any of the standard measures of speed of convergence to stationarity are the same for \( M_n^\theta \) and \( m_n^\theta \) (see, e.g., Diaconis and Zabell (1982)). The stationary distribution for \( m \) is calculated by summing over the conjugacy class. This gives

\[
\pi^\theta(\lambda) = \theta^{n-r} \frac{n!}{z_\lambda} \prod_{i=1}^{n} (\theta(i - 1) + 1)^{-1}.
\]
Here, if the partition \( \lambda = (\lambda_1 \cdots \lambda_r) \) with \( \lambda_1 \geq \lambda_2 \cdots \geq \lambda_r > 0 \) has \( n \) parts equal to \( i \),

\[
(4.5) \quad z_\lambda = \prod_{i=1}^{n} i^{a_i} a_i!
\]

Hanlon (1992) has derived the eigenvalues and eigenvectors of the Markov chain \( m \). This description involves an interesting class of symmetric functions called the Jack symmetric function. We will use the notation in Stanley (1984) and Macdonald (1974) for these functions. For each partition \( \lambda \) of \( n \) and each real \( \alpha \neq 0 \) there is a homogeneous symmetric function \( J_\lambda(\underline{x}; \alpha) \). Here \( \underline{x} = x_1 \cdots x_k \) with \( k \geq n \). For fixed \( \alpha \), \( \{J_\lambda(\underline{x}; \alpha)\}_\lambda \) is a basis of the homogeneous symmetric polynomials of degree \( n \) in \( x_1, x_2, \cdots, x_k \) as \( \lambda \) varies over partitions of \( n \). A more familiar basis is the power sum symmetric functions defined by \( P_i(\underline{x}) = x_1^i + \cdots + x_k^i \); \( P_\lambda = \prod_{i=1}^{\lambda_i} P_{\lambda_i} \). Denote the change of basis coefficients by \( C(\lambda, \mu) \):

\[
(4.6) \quad J_\lambda(\underline{x}; \alpha) = \sum_{\mu \vdash n} C^\alpha(\lambda, \mu) P_\mu(\underline{x}).
\]

The \( C^\alpha(\lambda, \mu) \) are rational functions in \( \alpha \). For example, when \( n = 3 \),

\[
J_{1,3} = P_1^3 - 3P_{1,2} + 2P_3 \\
J_{2,1} = P_1^3 + (\alpha - 1)P_{1,2} - \alpha P_3 \\
J_3 = P_1^3 - 3\alpha P_{1,2} + 2\alpha^2 P_3.
\]

When \( \alpha = 1 \), the Jack polynomials become the Schur functions. When \( \alpha = 2 \), the Jack polynomials become the zonal polynomials (spherical functions of \( GL_n/O_n \)). Until now, no interpretation of other values was known.

To state the main result, one further piece of notation is needed. Define

\[
(4.8) \quad j_\lambda(\alpha) = \prod_{s \in \lambda} h^*_s(s) h^*_\lambda(s)
\]

where the product is over square \( s \) in the diagram of \( \lambda \). If \( s = (i, j) \) has \( a \) squares strictly below it and \( b \) squares strictly to its right.

\[
h^*_s(s) = (a + 1)\alpha + b \\
h^*_s(s) = a\alpha + (b + 1)
\]

13
When $\alpha = 1$, $h^* = h_*$ is the usual hook length. For example, when $n = 3$, the upper and lower hook lengths are as shown:

\[
\begin{array}{ccc}
2 + \alpha & 1 + \alpha & \alpha \\
3 & 2 & 1
\end{array}
\quad
\begin{array}{ccc}
1 + 2\alpha & \alpha \\
2 + \alpha & 1 \\
\alpha & 1 \\
1
\end{array}
\quad
\begin{array}{c}
3\alpha \\
1 + 2\alpha \\
2\alpha \\
1 + \alpha \\
\alpha \\
1
\end{array}
\]

\[
j_\lambda = 6\alpha(1 + \alpha)(2 + \alpha) \quad \alpha^2(1 + \alpha)(2 + \alpha) \quad 6\alpha^3(1 + 2\alpha)(1 + \alpha)
\]

With this notation, our main result can be stated.

**Theorem 1.** For $0 < \theta \leq 1$, the Markov chain $m^\theta_n$ defined at (4.3) with stationary distribution $\pi^\theta_n$ at (4.4) has an eigenvalue $\beta_\lambda$ for each partition $\lambda = (\lambda_1, \lambda_2 \cdots \lambda_r)$ of $n$. These are

\[
\beta_\lambda = (1 - \theta) + \frac{\theta n(\lambda^t) + n(\lambda)}{\binom{n}{2}} \quad \text{with} \quad n(\lambda) = \sum_{i=1}^r (i - 1)\lambda_i = \sum_j \binom{\lambda^t_j}{2}.
\]

The corresponding left eigenvector, normed to be orthonormal in $L^2(1/\pi^\theta_n)$ is

\[
\frac{C^\theta(\lambda, \cdot)}{\left\{j_\lambda\Pi/(\theta^n n!))\right\}^{1/2}}
\]

with $C^\theta$ as in (4.6), $j_\lambda(\theta)$ as in (4.8), and $\Pi = \prod_{i=1}^n (\theta(i - 1) + 1)$.

**Example:** The matrix $m^\theta_3$ at (4.2) has eigenvalues and eigenvectors

\[
\begin{array}{cccc}
\lambda & 1^3 & 2,1 & 3 \\
\beta_\lambda & -\theta & \frac{2}{3}(1 - \theta) & 1 \\
C(\lambda, \cdot) & (1, -3, 2) & (1, \theta - 1, -\theta) & (1, 3\theta, 2\theta^2)
\end{array}
\]

The normalizing constants for these eigenvectors appear in (4.9).

**Proof of Theorem 1:** Stanley (1989, Theorem 5.4) shows, in the present notation

\[
(4.10) \quad J_\lambda(1^k; \alpha) = \prod_{(i,j) \in \lambda} (k - (i - 1) + \alpha(j - 1)).
\]

On the left, the function $J_\lambda$ is evaluated when all of its arguments are equal to 1. Since this is true for all $k$, one can equate coefficients in

\[
J_\lambda(1^k; \alpha) = \sum_{\mu^\alpha} C^\alpha(\lambda, \mu)P_\mu(1^k) = \sum_{\mu^\alpha} C^\alpha(\lambda, \mu)k^{\ell(\mu)}
\]
with $\ell(\mu) = r$ if $\mu = (\mu_1, \mu_2, \ldots, \mu_r)$. Of course, this only determines $C^\alpha(\lambda, \mu)$ for certain special $\mu$, e.g., $\mu = 1^2, \mu = 1^{n-2}, 2, \mu = n, \mu = 1, n - 1$. In particular

(4.11) \[ C^\alpha(\lambda, 1^n) = 1 \quad \text{for all } \lambda \]

(4.12) \[ C^\alpha(\lambda, 1^{n-2}, 2) = \]

The argument in Hanlon (1992) shows that the $C^\theta(\lambda, \mu)$ are left eigenvectors of the matrix $m^\theta_n$. To determine the corresponding eigenvalue, observe that the first column of $m^\theta_n$ is $(1 - \theta, \frac{1}{(\theta)}, 0, 0 \cdots 0)^t$. The eigenvalue equation becomes

\[ C^\theta(\lambda, 1^n)(1 - \theta) + \frac{C^\theta(\lambda, 1^{n-2}, 2)}{\binom{n}{2}} = \beta_\lambda C^\theta(\lambda, 1^n). \]

Now (4.11), (4.12) gives the eigenvalue.

The orthonormality follows from Stanley (1989) who gives

\[ \sum_{\nu \vdash n} C^\alpha(\lambda, \nu) C^\alpha(\mu, \nu) \alpha^\ell(\nu) z_\nu = j_\lambda \delta_{\lambda \mu}. \]

\[ \square \]

The ingredients above can be combined together to yield the following reasonably
sharp bounds for rates of convergence for the original chain $M$.

**Theorem 2.** For $0 < \theta < 1$, let $P^k_\theta$ be the probability on $S_n$ associated to the Markov
chain defined from $M^\theta_n$ starting as $id$, with stationary distribution $P_\theta$ defined at (4.1). Let

\[ k = an \log n + cn, \quad \text{with } a = \frac{1}{2\theta} + \frac{1}{4\theta} \frac{1}{\theta} (\frac{1}{\theta} - \theta) \quad \text{and} \quad c > 0. \]

Then, there is a function $f(\theta, c)$ independent of $n$, with $f(\theta, c) \to 0$ for $c \to \infty$ such that

\[ \|P^k_\theta - P_\theta\| \leq f(\theta, c). \]

**Proof:** From (2.5),

\[ 4\|P^k_\theta - P_\theta\|^2 \leq \theta^n n! \prod_{i=1}^n (1 + \theta(i - 1)) \sum_{\lambda \vdash n} \frac{1}{j_\lambda} \beta_\lambda^{2k}. \]

15
Here \( \beta_{\lambda} \) and \( j_{\lambda} \) are as in Theorem 1 and \( C(\lambda, n) = 1 \) from (4.11) was used. Now, one proceeds along the lines of Diaconis and Shahshahani (1981) or Diaconis (1988, pp. 36-43) who essentially did the case \( \theta = 1 \). The eigenvalue \( \beta_{\lambda} \) is monotone if the majorizaton ordering is used on \( \lambda : \lambda \geq \lambda' \) implies \( \beta_{\lambda} \geq \beta_{\lambda'} \). This holds for all \( \theta \in (0, 1) \). It allows the terms in the upper bound to be grouped just as in the \( \theta = 1 \) case. As there, the sum is dominated by the term for \( \lambda = (n-1, 1) \). For this term, \( \beta_{n-1,1} = 1 - \frac{2\theta}{n} - \frac{2}{n(n-1)} \) and

\[
\frac{\theta^n n! \prod_{i=1}^{n}(1 + \theta(i - 1))/j_{n-1,1}}{\theta^2(n-1)!(1 + \theta)(2 + \theta) \cdots (n - 3 + \theta)(n - 1 + \theta)(n - 2 + 2\theta)} = f(\theta) n^{2+(\frac{1}{2} - \theta)}(1 + O(\frac{1}{n}))
\]

for an explicit continuous function \( f(\theta) \). It follows from this, with \( k \) as given, that

\[
\frac{\theta^n n! \prod_{i=1}^{n}(1 + \theta(i - 1))}{j_{\lambda}} \beta_{\lambda}^{2k} \leq f(\theta) n^{2+(\frac{1}{2} - \theta)} e^{-\frac{4\theta}{n}(a_n \log n - 2Cn)}(1 + O(\frac{1}{n})) = f(\theta) e^{-8\theta C}(1 + O(\frac{1}{n})).
\]

We omit the rest of the argument, except for consideration of the term corresponding to \( 1^n \). For this term \( \beta_{1^n} = -\theta \) and

\[
\frac{\theta^n n! \prod_{i=1}^{n}(1 + \theta(i - 1))/j_{1^n}}{n! \theta(1 + \theta)(2 + \theta) \cdots (n - 1 + \theta)} = \frac{\theta^n n!(1 + \theta)(1 + 2\theta) \cdots (1 + (n - 1)\theta)}{n! \theta(1 + \theta)(2 + \theta) \cdots (n - 1 + \theta)}.
\]

This last quantity is bound above by a positive continuous function \( g(\theta) \), for all \( n \). It follows that the term for \( 1^n \) tends to zero.

**Remarks:**
1) The function \( a(\theta) = \frac{1}{2\theta} + \frac{1}{4\theta}(\frac{1}{2} - \theta) \) increases as \( \theta \) decreases from 1 to 0 so it takes longer to converge for small \( \theta \). This seems curious since the walk starts at the identity which is also the most likely state. 2) We have not attempted to prove a matching lower bound but are morally certain that the variation distance does not tend to zero for \( h = a_n \log n - cn \) when \( n \) is large and \( c \) is positive. The arguments in Diaconis and Shahshahani (1981) show this for \( k = \frac{1}{2} n \log n - cn \) for all \( \theta \). This goes part of the way to explaining the curiosity in remark 1: for the walk to reach stationarity, the distribution of the number of fixed points has to be correct. In particular there has to be a reasonable chance of hitting every card. This already requires \( \frac{1}{2} n \log n + cn \) moves be made.
5. One-factors and Signed Markov Chains.

In this section, we will consider two other cases in which the Metropolis Algorithm produces twisted Markov Chains with interesting steady states. The second of these will be obtained from the first by appropriately introducing signs in the transition matrix. In view of this, it will not technically be a Markov Chain. Nevertheless, the Metropolis Algorithm still applies and gives us an interesting twisting of our “signed” Markov chain.

This section twists a chain studied by Diaconis (1986) who gave the following interpretation: consider $f$ pairs of mathematicians who come to a party. They arrive in pairs as $\{(1,2),(3,4)\cdots (2f-1,2f)\}$. Being mathematicians, they stand there talking to the person they arrived with. A host decides to mix things up by picking a pair, say $(i_1,i_2)$, at random, then a second pair, say $(j_1,j_2)$, at random, and switching, say, to $(i_1,j_1),(i_2,j_2)$ or $(i_1,j_2),(i_2,j_1)$. This defines a process on $\Omega_f$ the partitions of $2f$ into $f$ 2-element blocks where order within or between doesn’t matter.

The symmetric group $S_{2f}$ acts transitively on $\Omega_f$. The isotropy subgroup can be identified with $B_f = Z_2^f \alpha S_f$, the hyperoctahedral group. The pair $S_{2f},B_f$ is a Gelfand pair with spherical functions given by the coefficients $C^2(\lambda,\mu)$ of section 4. This allowed a complete analysis when $\theta = 1$. The present section refines this analysis to allow for twisting.

5.1 Switching Random Pairs in One-factors.

A one-factor on $2f$ points is a graph with $2f$ points in which every point has degree 1. Let $\Omega_f$ denote the set of 1-factors on $(2f)$ points. It is easy to see that $|\Omega_f| = (2f)!/2^f f!$. We usually draw a one-factor $\delta \in \Omega_f$ by putting the points $1,2,\cdots,f$ in a top row and the points $(f+1),\cdots,2f$ in a bottom row. For example, the one-factor $\Delta_0$ which has an edge from $i$ to $i+f$ for $i = 1,2,\ldots,f$ is drawn as:

\[
\begin{array}{cccc}
  i & 2 & \cdots & f \\
  \cdot & \cdot & \cdots & \cdot \\
\end{array}
\]

\[\Delta_0 = \begin{array}{cccc}
  & \cdot & \cdots & \cdot \\
  f+1 & f+2 & \cdots & 2f \\
\end{array}\]
Given \( \delta_1, \delta_2 \in \Omega_f \). Let \( \delta_1 \cup \delta_2 \) be the graph obtained by taking the union of their edges. It is easy to check that \( \delta_1 \cup \delta_2 \) is a disjoint union of cycles of even lengths. Let \( \Lambda(\delta_1, \delta_2) \) be the partition of \( f \) whose parts are half the cycle length of \( \delta_1 \cup \delta_2 \).

Define a Markov chain \( M \) with states \( \Omega_f \) in the following way. Given a 1-factor \( \delta \), the probability \( R_f(1)_{\delta, \tau} \) of moving to another 1-factor \( \tau \) is given by

\[
R_f(1)_{\delta, \tau} = \begin{cases} \frac{1}{2f-1} & \text{if } \delta = \tau \\ \frac{2}{f(2f-1)} & \text{if } \tau \text{ can be obtained from } \delta \text{ by switching a pair of non-adjacent points} \\ 0 & \text{otherwise.} \end{cases}
\]

For example, with \( f = 2 \) we have

\[
R_2(1) = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}
\]

Let \( \delta \) be in \( \Omega_f \). The type of \( \delta \) is the partition \( \Lambda(\delta, \Delta_0) \). Type for one-factors is analogous to cycle type for permutations. Let \( c(\delta) \) denote the number of parts of \( \Lambda(\delta, \Delta_0) \).

Define a probability distribution \( \Gamma \) on \( \Omega_f \) by

\[
\Gamma(\delta) = \alpha^{f-c(\delta)} / G
\]

where \( \alpha \) is a real number greater than 1 and where \( G \) is the appropriate normalizing constant \( G = \prod_{i=0}^{f-1} (1 + i(2\alpha)) \). We now have an \( M = M(1) \) and a probability distribution \( \Gamma \) so we are in a position to apply the Metropolis algorithm to obtain a twisted Markov chain \( M(\alpha) \) whose stable distribution is \( \Gamma \). Let \( R_f(\alpha) \) denote the transition matrix of \( M(\alpha) \). For example,

\[
R_2(\alpha) = \begin{pmatrix} \alpha & \alpha & \alpha \\ 1 & 3\alpha - 2 & 1 \\ 1 & 1 & 3\alpha - 2 \end{pmatrix} \frac{1}{3\alpha}.
\]

The next proposition we state without proof because we will prove something more general later in this section. The proposition says that the Markov chain \( R_f(\alpha) \) can be reduced (or lumped) to a Markov chain on the types.
PROPOSITION. For each partition $\lambda$ of $f$ let $v_\lambda$ be defined by

$$v_\lambda = 2^{-t(\lambda)} \left\{ \sum_{\delta \in \Delta_\lambda} \delta \right\},$$

and let $V$ denote the span of the $v_\lambda$. Then

$$R_f(\alpha)v_\lambda \subseteq V.$$

Let $R_f^{(f, \phi)}(\alpha)$ denote the restriction of $R_f(\alpha)$ to $V$ written with respect to the basis \{v_\lambda\}. Recall from Section 2 that $J_\lambda(\alpha)$ denotes the Jack symmetric function indexed by $\lambda$ and that $c_{\lambda, \mu}^\phi$ is the coefficient of $p_\mu(x)$ when $J_\lambda(x; \alpha)$ expanded in terms of power sums. The main result of this section can now be stated.

THEOREM 1. For each $\lambda$, the vector $\sum_{\mu + f} c_{\lambda, \mu}^{2f} v_\mu$ is a left eigenvector of $R_f^{(f, \phi)}(\alpha)$ with corresponding eigenvalue $(f\alpha + 2\{2n(\lambda') - n(\lambda)\})/\alpha(2f)_2$.

PROOF: To begin, we will take a closer look at the entries of the matrix $R_f^{(f, \phi)}(\alpha)$. Fix a partition $\lambda = (\lambda_1, \ldots, \lambda_\ell)$, let $\pi_\lambda$ be the permutation in $S_f$ given by

$$\pi = (1, 2, \ldots, \lambda_1)(\lambda_1 + 1, \ldots, \lambda_1 + \lambda_2) \cdots (\lambda_1 + \cdots + \lambda_{\ell-1} - 1, \cdots, f)$$

(so $\pi_\lambda$ has cycle type $\lambda$) and let $\delta_\lambda$ be the 1-factor with an edge from $i$ to $f + \pi(i)$ for $i = 1, 2, \cdots, f$. So $\delta_\lambda$ looks like

$$\delta_\lambda =$$

It is easy to check that $\delta_\lambda$ has type $\lambda$ (so $\delta_\lambda$ is amongst the one-factors appearing in the sum $v_\lambda$).

We can make use of the one-factors $\delta_\lambda$ to compute the entries of $R_f^{(f, \phi)}(\alpha)$. The following lemma is derived easily from the definition of $R_f^{(f, \phi)}(\alpha)$.

LEMMA 1. Let $\lambda$ and $\mu$ be partitions of $f$. The $\lambda, \mu$ entry of $R_f^{(f, \phi)}(\alpha) := \alpha(2f)_2 R_f^{(f, \phi)}(\alpha)$ is given by:
(a) If $\ell(\mu) \geq \ell(\lambda)$ then $(\hat{R}_f^{(f,\phi)}(\alpha))_{\lambda\mu}$ is $a^{2\ell(\mu) - \ell(\lambda)}$ times the number of $\delta$ of type $\lambda$ which can be obtained from $\delta_\mu$ by switching a pair of points.

(b) If $\ell(\mu) < \ell(\lambda)$ then $(\hat{R}_f^{(f,\phi)}(\alpha))_{\lambda\mu}$ is $2^{\ell(\mu) - \ell(\lambda)}$ times the number of $\delta$ of type $\lambda$ which can be obtained from $\delta_\mu$ by switching a pair of points (you count $\delta$ more than once if it can be obtained more than once).

Return now to the proof of Theorem 1. We will fix a partition $\mu = (\mu_1, \ldots, \mu_\ell)$ and compute the $\lambda, \mu$ entries of $(\hat{R}_f^{(f,\phi)}(\alpha))_{\lambda\mu}$ for all $\lambda$. According to the lemma we need to count how many $\delta'$ of each type $\lambda$ arise by switching a pair of points $(x, y)$ in $\delta_\mu$.

**CASE 1**: $x$ and $y$ come from different cycles of $\delta_\mu \cup \Delta_0$, say $x$ comes from the cycle of length $\mu_i$ and $y$ comes from the cycle of length $\mu_j$ (where $i \neq j$).

In $((x, y) \cdot \delta_\mu) \cup \Delta_0$ the points that were in these cycles of lengths $\mu_i$ and $\mu_j$ now form a cycle of length $2(\mu_i + \mu_j)$. So the partition "corresponding" to partition $\lambda$ is

$$\lambda = \mu[\mu_i + \mu_j \leftarrow \mu_i, \mu_j]$$

(this notation means that $\mu_i$ and $\mu_j$ are removed and replaced by $\mu_i + \mu_j$).

For fixed $\mu_i$ and $\mu_j$ we had $(2\mu_i)(2\mu_j)$ choices of the pair $(x, y)$. In this case, $\ell(u) = \ell(\lambda) + 1$ so the $\lambda, \mu$ entry of $(\hat{R}_f^{(f,\phi)}(\alpha))_{\lambda\mu}$ contains a factor of $2\alpha$.

**CASE 2**: $x$ and $y$ come from the same cycle $C$ of $\delta_\mu \cup \Delta_0$. Let $\mu_i$ be the part of $\mu$ corresponding to $C$ so $C$ consists of $2\mu_i$ points, $\mu_i$ from each row. It is straightforward to check the following fact.

(5.3) A) Suppose $x$ and $y$ are in the same row and are separated (cyclically) by gaps of $b$ and $\mu_i$. Then $((x, y)\delta_\mu) \cup \Delta_0$ corresponds to the partition

$$\lambda = \mu[b, \mu_i - b \leftarrow \mu_i].$$

B) Suppose $x$ and $y$ are in different rows. Then $((x, y)\delta_\mu) \cup \Delta_0$ corresponds to the partition $\mu$.

Note that in case (A) $(\hat{R}_f^{(f,\phi)}(\alpha))_{\lambda\mu}$ is equal to $\frac{1}{2}$ whereas in case (B) $(\hat{R}_f^{(f,\phi)}(\alpha))_{\lambda\mu}$ is equal to $\alpha$.
Applying the above facts we see that the $\mu$th column of $(\hat{R}_f^{(f, \phi)}(\alpha))_{\lambda \mu}$ is given by the above transformation:

\[
\mu\text{th column of } \hat{R}_f^{(f, \phi)}(\alpha)
\]

\[
v_\mu \longrightarrow \sum_{i < j} (2\mu_i \mu_j) v_{\mu[\mu_i + \mu_j - \mu_i, \mu_j]} (f\alpha + 2(2\alpha - 1)(\sum_i \binom{\mu_i}{2})) v_\mu \\
+ (2\alpha) \left( \sum_i \mu_i \sum_{b = 1}^{\mu_i - 1} v_{\mu[b, \mu_i - b - \mu_i]} \right).
\] (5.4)

It is worth commenting on how we computed the coefficient of $v_\mu$ on the right-hand side. For starters there are $\sum_i \mu_i^2$ pairs $(x, y)$ that yield the same $\mu$. Of these, $\sum_i \mu_i(\mu_i - 1)$ satisfy $(x, y)\delta_\mu \neq \delta_\mu$ and $f$ satisfy $(x, y)\delta_\mu = \delta_\mu$. The former pairs each contribute 1 to the $v_\mu, v_{\mu}$ entry. The latter contribute a more complicated factor, namely, the $\delta_\mu, \delta_\mu$ entry of $R_f(\alpha)$. We need to compute the entry. First observe that the $\delta_\mu, \delta_\mu$ entry in $R_f(1)$ is $f$ (pairs $(x, y)$ which fix $\delta_\mu$ must be endpoints of an edge of $\delta_\mu$). The off-diagonal entries of $\hat{R}_f(\alpha)$ are either 0, 1, or $\alpha$. It follows that the $\delta_\mu, \delta_\mu$ entry of $\hat{R}_f(\alpha)$ is $f\alpha + H(\alpha - 1)$ where $H$ is the number of off-diagonal entries equal to 1 in the $\delta_\mu$ row. This is exactly the number of pairs $(x, y)$ with $c((x, y)\delta_\mu) > c(\delta_\mu)$. As seen above, this number is $2 \sum_i \binom{\mu_i}{2}$. So, the $\delta_\mu, \delta_\mu$ entry in $\hat{R}_f(\alpha)$ is

\[
2\alpha \sum_i \binom{\mu_i}{2} + f\alpha + 2(\alpha - 1)(\sum_i \binom{\mu_i}{2}) \\
= 2\{\frac{\alpha}{2} + (2\alpha - 1)(\sum_i \binom{\mu_i}{2})\}.
\]

Let $\Lambda_n$ be the ring of symmetric functions in $x_1, \ldots, x_n$ and let $\Lambda_n^f$ be the subspace spanned by polynomials that are homogeneous of degree $f$. Define $\psi : V \rightarrow \Lambda_n^f$ to be the linear map satisfying

\[
\psi(v_\lambda) = p_\lambda(\bar{x}).
\]

Now let $\rho$ be the linear transformation on $\Lambda_n^f$ given by

\[
\rho = \psi \circ \hat{R}_f(\alpha) \circ \psi^{-1}.
\]

The computation we've just done tells us how to compute the $\mu$th column of $\hat{R}_f(\alpha)$, i.e., how to compute the linear map given by $\hat{R}_f(\alpha)^t$. The matrix for $\rho$ is the dual $\hat{R}_f(\alpha)^t$ with
respect to the inner product \( \langle p_{\lambda}(x), p_{\mu}(x) \rangle = x^{-1} \delta_{\lambda \mu} \). A straightforward computation shows that

\[
\rho(p_{\lambda}(x)) = \{(f \alpha) + 2(2\alpha - 1)n(\lambda')\} p_{\lambda}(x) \\
+ 2(2\alpha) \sum_{u \neq v} \lambda_u \lambda_v p_{\lambda[\lambda_u, \lambda_v - \lambda_u + \lambda_v]}(x) \\
+ 2\left(\frac{1}{2} \sum_k \lambda_k^{k-1} \sum_{j=1} p_{\lambda[j, \lambda_k - j]}(x)\right).
\]

(5.5)

The result now follows by comparing (5.5) with either formula (3.6) in Hanlon (1988) or the first formula in the proof of Theorem 3.1 in Stanley (1989).

We refrain from carrying out further computations for this example. Preliminary analysis shows order \( a n \log n + cn \) steps are needed with \( a \) depending on \( \alpha \).

5.2 Signed One-factors.

We are going to modify the situation in the previous section so that each one-factor comes with an "orientation". Changing the orientation alters the one-factor by a sign. This change in orientation will sometimes occur when we switch points and the result will be that the transition matrix \( S_f(1) \) for the untwisted Markov chain will have some negative entries. So \( S_f(1) \) will not really represent a Markov chain but we can still apply the Metropolis algorithm to it as given in (2.1).

To begin, we will generalize the notion of one-factors. Let \( B_f \) be the automorphism group of \( \Delta_0 \). \( B_f \) is the hyperoctahedral group of order \( f!2^f \) and can be thought of as the set of signed \( f \times f \) permutation matrices in the following way. An element \( \sigma \in B_f \) must permute the \( f \) edges of \( \Delta_0 \) which gives us the underlying \( f \times f \) permutation matrix \( \bar{\sigma} \). If \( \sigma \) maps the \( i^{th} \) edge of \( \Delta_0 \) to the \( j^{th} \) edge (so there is a 1 in the \( i,j \) entry of \( \bar{\sigma} \)). Then \( \sigma_{ij} = 1 \), if \( \sigma \) maps the point in the top row of column \( j \) and \( \sigma_{ij} = -1 \) if the point in the top row of column \( i \) goes to the point in the bottom row of column \( j \).

The group \( B_f \) has four linear characters \( \delta_0, \delta_1, \delta_2, \delta_3 \). To describe them, let \( \sigma \in B_f \)
(think of $\sigma$ as an $f \times f$ signed permutation matrix). Then

$$\delta_0(\sigma) = 1$$

$$\delta_1(\sigma) = \text{sign}(\sigma)$$

$$\delta_2(\sigma) = \text{det}(\sigma)$$

$$\delta_3(\sigma) = \delta_1(\sigma)\delta_2(\sigma).$$

There is an alternative description of the character $\delta_3$. If we think of $B_f$ as a subgroup of $S_{2f}$ then $\delta_3$ is the restriction of the sign character of $S_{2f}$ to $B_f$.

We will now follow a construction given in Stembridge (1992). For each $i$, let $e_i$ denote the idempotent given by $\delta_i$,

$$e_i = \frac{1}{|B_f|} \sum_{\sigma \in B_f} \delta_i(\sigma)\sigma.$$ 

Let $X_i$ denote the left ideal $e_iCS_{2f}$ and let $V_i$ denote the two-sided ideal $e_iCS_{2f}e_i$. Note that $X_0$ and $V_0$ are isomorphic to the spaces of left $B_f$-cosets and double $B_f$-cosets in $S_{2f}$.

There is a combinatorial method for identifying a basis of the $X_i$ and $V_i$. Let $\tau$ be a permutation in $S_{2f}$ written in 1-line form,

$$\tau = a_1a_2\cdots a_{2f}.$$ 

Assign to $\tau$ a one-factor $\delta(\tau)$ by putting an edge between $i$ and $j$ in $\delta(\tau)$ iff $|a_i - a_j| = f$.

The following result is well known for $i = 0, 3$ and can be found in Stembridge (1992) for $i = 1, 2$.

**Theorem 5.**

(A) For $i = 0, 1, 2, 3$ we have $e_i\tau = \pm e_i\pi$ iff $\delta(\tau) = \delta(\pi)$.

(B) For $i = 0, 3$ we have $e_i\tau e_i = \pm e_i\pi e_i$ iff $\Lambda(\delta(\tau), \Delta_0) = \Lambda(\delta(\pi), \Delta_0)$.

(C) For $i = 1, 2$ we have

(i) $e_i\tau c_i \neq 0$ iff $\Lambda(\delta(\tau), \Delta_0)$ has all odd parts.

(ii) Suppose $e_i\tau e_i$ and $e_i\pi e_i$ are nonzero.

Then $e_i\tau e_i = e_i\pi e_i$ iff $\Lambda(\delta(\tau), \Delta_0) = \Delta(\delta(\pi), \Delta_0)$. For the moment we will construct only cases $i = 0, 3$. Theorem 5 tells us that the spaces $\Omega_f$ and $V$ from Section 4.1 can be identified with the cosets $X_0$ and double cosets $V_0$ of $B_f$.

The stable distribution $\Gamma$ is a simple function that is constant on double cosets. It remains to understand the Markov chain $R_f(1)$ in this context.
Lemma 2. Let $\Omega_f$ be identified with $X_0$ as above. Then $R_f(1)$ is the matrix for right multiplication by $\frac{1}{|2f|} \sum_{1 \leq i \leq j \leq 2f} (i, j)$.

We will now follow the same construction as in Section 5.1 but with $X_0$ replaced by $X_3$. The stable distribution $\Gamma$ will be the same and the untwisted Markov chain will again be multiplication by $\frac{1}{(2f)} \sum (i, j)$. To make this precise, we need to specify bases for $X_3$ and $V_3$ (because the natural bases are determined only up to sign).

Definition.

(A) For each one-factor, $\delta$, let $\tau(\delta)$ be the lexicographically minimal element $\alpha$ of $S_{2f}$ with $\delta(\alpha) = \delta$.

(B) For each partition $\lambda$ of $f$, let $\pi(\lambda)$ be the lexicographically minimal element $\beta$ of $S_{2f}$ with $\Lambda(\delta(\beta), \Delta_0) = \lambda$.

Let $B$ be the basis for $X_3$ given by

$$B = \{ e_3 \tau(\delta) : \delta \in \Omega_f \}$$

and let $C$ be the vector space for $V_3$ given.

$$C = \{ 2^{f(\lambda)} e_3 \pi(\lambda) e_3 : \lambda \vdash f \}.$$ 

We henceforth let $v^{(3)}_\lambda$ denote $2^{f(\lambda)} e_3$. Let $Q_f(1)$ be the matrix for right multiplication by $(2f)^{-1} \sum (i, j)$ with respect to the basis $B$. Let $Q_f(\alpha)$ be the matrix obtained by applying the Metropolis algorithm to $Q_f(1)$ and $\Gamma$. Since $Q_f(1)$ corresponds to right multiplication by a conjugacy class in $S_{2f}$, it commutes with right multiplication by $e_3$. It follows that $Q_f(\alpha)$ restricts to the space $V_3$. Let $Q_f^{(\phi,1^f)}(\alpha)$ be this restriction with respect to the basic $C$. We can now state the analogue of Theorem 4.

Theorem 2. For each $\lambda$, the vector $\sum_{\mu+f} c^{(\alpha/2)}_{\mu} v^{(3)}_{\mu}$ is a left eigenvector of $Q_f^{(\phi,1^f)}(\alpha)$ with corresponding eigenvalue $(-f \alpha + 2\{ \frac{\alpha}{2} n(\lambda') - n(\lambda) \})/\alpha(2f)$.

To prove Theorem 2 one follows much the same procedure as in the proof of Theorem 1 (but care must be taken with signs).
One reason for the discussion of this signed case is to bring up the question of what happens for the characters $\delta_1$ and $\delta_2$. A simple computation shows that the restriction of right multiplication by $\left(\frac{2t}{3}\right)^{-1} \sum(i,j)$ to $V_1$ (or $V_2$) is the identity map. So this is the wrong (untwisted) Markov chain to begin with. Another possibility is to let the untwisted Markov chain be right multiplication by $\left(2\left(\frac{2t}{3}\right)^{-1} \sum(i,j,k)\right)$. A recent result of Stembridge tells us that the restriction of this Markov chain to $V_1$ has Schur's $Q$-functions as eigenvectors. So applying the Metropolis algorithm to this restriction will produce a Markov chain with eigenvectors that are perhaps interesting deformations of the $Q$-functions.

Acknowledgement. We thank Jim Fill for help with Section 3.
REFERENCES


