Eigen Analysis for Some Examples
of the Metropolis Algorithm

PERSI DIACONIS AND PHIL HANLON

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 π which is effective when π 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 π. 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

This paper is in final form and no version of it will be submitted for publication elsewhere .

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 one-parameter orthogonal polynomials such as Krawtchouk 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 (1989) and Macdonald (1989), these eigenvalues are explicitly available. We determine the convergence properties and study the dependence on θ.

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 Krawtchouk 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

$$(2.1) \quad M(x,y) = \begin{cases} S(x,y)r_{yx} & \text{if } r_{yx} < 1 \\ S(x,y) & \text{if } x \neq y \text{ and } r_{yx} \geq 1 \\ S(x,x) + \sum_{\substack{z \neq x \\ r_{zy}}} S(x,z)(1 - r_{zy}) & \text{if } x = y. \end{cases}$$

This definition has a simple implementation. If the chain is at x, pick y with probability $S(x,y)$. If $x \neq y$ 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 π. 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, π) 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 Γ orthogonal and β a diagonal matrix having the eigenvalues of T, and so M, on the diagonal. Thus $M = V\beta V^{-1}$ with $V = D^{-1}\Gamma, V^{-1} = \Gamma^t D$. This implies that the right eigenvectors of M are the columns of V: $V_{xy} = \Gamma_{xy}/\sqrt{\pi(x)}$. 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, π) be a reversible Markov chain on a finite set X. Let β_y denote the eigenvalues, β^* 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 three quantities:*

(2.2) $$\sum_y \beta_y^{2k} f_y^2(x) - 1$$

(2.3) $$\frac{1}{\pi^2(x)}\sum_y \beta_y^{2k} g_y(x) - 1$$

(2.4) $$\frac{1}{\pi(x)}(\beta^*)^{2k}.$$

PROOF. For any fixed x

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

$$\leq \sum_y \frac{|M^k(x, y) - \pi(y)|^2}{\pi(y)}$$

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

$$= \frac{M^{2k}(x, x)}{\pi(x)} - 1$$

$$= \frac{1}{\pi(x)}\sum_y \beta_y^{2k}\Gamma_{xy}^2 - 1.$$

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

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, π), 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

$$1 - \beta_i = \min_{W_i} \min_{f \in W_i} \frac{\mathcal{E}(f,f)}{\|f\|_\pi^2}, \qquad \dim W_i = i+1.$$

Here β_i is the ith 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* $(M, \pi), (\widetilde{M}, \widetilde{\pi})$ *be reversible Markov chains on a finite set* X. *Suppose the associated Dirichlet forms satisfy*

$$\widetilde{\mathcal{E}} \leq A\mathcal{E} \quad \text{and} \quad \widetilde{\pi} \geq a\pi \quad \text{for positive } A, a.$$

Then

$$\beta_i \leq 1 - \frac{a}{A}(1 - \widetilde{\beta}_i).$$

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

$$(3.1) \qquad S(x,y) = \begin{cases} \frac{1}{d} & \text{if } H(x,y) = 1 \\ 0 & \text{otherwise} \end{cases}$$

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 Ehrenfest urn. See Kac (1947), Letac and Takacs (1979), or Diaconis (1988) and references cited there. In particular, the matrix S is diagonalizable with eigenvalues $(1 - \frac{2j}{d})$ 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

$$(3.2) \qquad P_\theta(x) = \frac{\theta^{H(x)}}{(1+\theta)^d}$$

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

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

$$(3.3) \qquad M(x,y) = \begin{cases} \frac{1}{d} & \text{if } H(x,y) = 1, \ H(y) < H(x) \\ \frac{\theta}{d} & \text{if } H(x,y) = 1, \ H(y) > H(x) \\ (1 - \frac{H(x)}{d})(1-\theta) & \text{if } H(x,y) = 0 \\ 0 & \text{otherwise.} \end{cases}$$

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

$$\begin{array}{c@{\quad}cccc} & 00 & 01 & 10 & 11 \\ \begin{array}{c} 00 \\ 01 \\ 10 \\ 11 \end{array} & \left(\begin{array}{cccc} 1-\theta & \frac{\theta}{2} & \frac{\theta}{2} & 0 \\ \frac{1}{2} & \frac{1-\theta}{2} & 0 & \frac{\theta}{2} \\ \frac{1}{2} & 0 & \frac{1-\theta}{2} & \frac{\theta}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{array}\right). \end{array}$$

The permutation group S_d operates on $\mathbb{Z}_2^d$. 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 $(0, 1, \ldots, d)$ with transition matrix

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

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

$$\begin{array}{c@{\quad}ccc} & 0 & 1 & 2 \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \left(\begin{array}{ccc} 1-\theta & \theta & 0 \\ \frac{1}{2} & \frac{1-\theta}{2} & \frac{\theta}{2} \\ 0 & 1 & 0 \end{array}\right). \end{array}$$

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

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

$$\beta_i = 1 - \frac{i}{d}(1 + \theta), \qquad 0 \le i \le d.$$

The corresponding right eigenvector is the Krawtchouk polynomial

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

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

PROOF. This is essentially contained in Krawtchouk (1929). The statement is easy to verify from known properties of Krawtchouk polynomials as given by MacWilliams 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

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad \frac{1}{\sqrt{2\theta}}\begin{pmatrix} 2\theta \\ \theta - 1 \\ -2 \end{pmatrix}, \qquad \frac{1}{\theta}\begin{pmatrix} \theta^2 \\ -\theta \\ 1 \end{pmatrix}.$$

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 θ 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*

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

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

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),

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

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

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

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

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

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

$$S_2 \leq \sum_{\ell=1}^{\infty} \frac{1}{\ell!} \exp\{-\ell[\log \theta^{\frac{\ell+1}{\theta}} + \frac{c}{\theta}]\}.$$

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

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 $a = \left(\frac{1+\theta}{2}\right)^d$ and $A = \frac{1}{\theta}\left(\frac{1+\theta}{2\theta}\right)^d$. This yields

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

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

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

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

(4.1) $$P_\theta(\pi) = c(\theta)\theta^{d(\pi, \pi_0)}$$

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 π_0 is a "location parameter." This model describes a population peaked about π_0 which falls off geometrically at rate θ. For $\theta = 1$,

P_1 becomes the uniform distribution. In statistical applications, π_0 and θ 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

$$
M_3^\theta =
\begin{array}{c}
 \\
\text{id} \\
(12) \\
(13) \\
(23) \\
(123) \\
(132)
\end{array}
\begin{array}{c}
\begin{array}{cccccc}
\text{id} & (12) & (13) & (23) & (123) & (132)
\end{array} \\
\left(
\begin{array}{cccccc}
1-\theta & \frac{\theta}{3} & \frac{\theta}{3} & \frac{\theta}{3} & 0 & 0 \\
\frac{1}{3} & \frac{2}{3}(1-\theta) & 0 & 0 & \frac{\theta}{3} & \frac{\theta}{3} \\
\frac{1}{3} & 0 & \frac{2}{3}(1-\theta) & 0 & \frac{\theta}{3} & \frac{\theta}{3} \\
\frac{1}{3} & 0 & 0 & \frac{2}{3}(1-\theta) & \frac{\theta}{3} & \frac{\theta}{3} \\
0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\
0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0
\end{array}
\right)
\end{array}.
$$

The stationary distribution is the left eigenvector proportional to

$$(1, \theta, \theta, \theta, \theta^2, \theta^2).$$

By construction, the matrix M_n^θ commutes with the action of S_n on itself by conjugation. This implies that if π and σ 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^θ

$$
(4.2) \qquad
\begin{array}{c}
1 \\
1,2 \\
3
\end{array}
\begin{array}{c}
\begin{array}{ccc}
1^3 & 1,2 & 3
\end{array} \\
\left(
\begin{array}{ccc}
1-\theta & \theta & 0 \\
\frac{1}{3} & \frac{2}{3}(1-\theta) & \frac{2}{3}\theta \\
0 & 1 & 0
\end{array}
\right)
\end{array}.
$$

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

$$(4.3) \qquad m_n^\theta(\lambda, \mu) = \sum M_n^\theta(\pi, \sigma)$$

summed over all σ in the conjugacy class μ with π any permutation in conjugacy class λ.

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

$$(4.4) \qquad \pi^\theta(\lambda) = \theta^{n-r} \frac{n!}{z_\lambda} \prod_{i=1}^n (\theta(i-1)+1)^{-1}.$$

Here, if the partition $\lambda = (\lambda_1, \ldots, \lambda_r)$ with $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_r > 0$ has a_i parts equal to i,

$$(4.5) \qquad z_\lambda = \prod_{i=1}^n i^{a_i} a_i!$$

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 (1979) for these functions. For each partition λ of n and each real $\alpha \neq 0$ there is a homogeneous symmetric function $J_\lambda(\mathbf{x}; \alpha)$. Here $\mathbf{x} = (x_1, \ldots, x_k)$ with $k \geq n$. For fixed α, $\{J_\lambda(\mathbf{x}; \alpha)\}_\lambda$ is a basis of the homogeneous symmetric polynomials of degree n in $x_1, x_2, \ldots, x_k$ as λ varies over partitions of n. A more familiar basis is the power sum symmetric functions defined by $P_i(\mathbf{x}) = x_1^i + \cdots + x_k^i; P_\lambda = \prod_{i=1}^r P_{\lambda_i}$. Denote the change of basis coefficients by $C(\lambda, \mu)$:

$$(4.6) \qquad J_\lambda(\mathbf{x}; \alpha) = \sum_{\mu \vdash n} C^\alpha(\lambda, \mu) P_\mu(\mathbf{x}).$$

The $C^\alpha(\lambda, \mu)$ are rational functions in α. 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

$$(4.8) \qquad j_\lambda(\alpha) = \prod_{s \in \lambda} h_*^\lambda(s) h_\lambda^*(s)$$

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

$$\begin{aligned} h^*(s) &= (a+1)\alpha + b, \\ h_*(s) &= a\alpha + (b+1). \end{aligned}$$

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

(4.9)

$$
\begin{array}{|c|c|c|}
\hline
2+\alpha & 1+\alpha & \alpha \\
\hline
3 & 2 & 1 \\
\hline
\end{array}
\qquad
\begin{array}{|c|c|}
\hline
1+2\alpha & \alpha \\
\hline
2+\alpha & 1 \\
\hline
\alpha \\
\cline{1-1}
1 \\
\cline{1-1}
\end{array}
\qquad
\begin{array}{|c|}
\hline
3\alpha \\
\hline
1+2\alpha \\
\hline
2\alpha \\
\hline
1+\alpha \\
\hline
\alpha \\
\hline
1 \\
\hline
\end{array}
$$

$j_\lambda \qquad 6\alpha(1+\alpha)(2+\alpha) \qquad \alpha^2(1+\alpha)(2+\alpha) \qquad 6\alpha^3(1+2\alpha)(1+\alpha)$

With this notation, our main result can be stated.

THEOREM 1. *For $0 < \theta \le 1$, the Markov chain m_n^θ defined at (4.3) with stationary distribution π_n^θ at (4.4) has an eigenvalue β_λ for each partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_r)$ of n. These are*

$$
\beta_\lambda = (1-\theta) + \frac{\theta n(\lambda^t) + n(\lambda)}{\binom{n}{2}}
$$

with

$$
n(\lambda) = \sum_{i=1}^r (i-1)\lambda_i = \sum_j \binom{\lambda_j^t}{2}.
$$

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

$$
\frac{C^\theta(\lambda, \cdot)}{\{j_\lambda \Pi/(\theta^n n!)\}^{1/2}}
$$

with C^θ 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_3^θ at (4.2) has eigenvalues and eigenvectors

λ	1^3	$2,1$	3
β_λ	$-\theta$	$\frac{2}{3}(1-\theta)$	1
$C(\lambda, \cdot)$	$(1, -3, 2)$	$(1, \theta - 1, -\theta)$	$(1, 3\theta, 2\theta^2)$.

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

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

$$
J_\lambda(1^k; \alpha) = \sum_{\mu \vdash n} C^\alpha(\lambda, \mu) P_\mu(1^k) = \sum_{\mu \vdash n} C^\alpha(\lambda, \mu) k^{\ell(\mu)}
$$

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 μ, e.g., $\mu = 1^2, \mu = 1^{n-2}, 2, \mu = n, \mu = 1, n - 1$. In particular

$$(4.11) \qquad\qquad C^\alpha(\lambda, 1^n) = 1$$

and

$$(4.12) \qquad C^\alpha(\lambda, 1^{n-2}, 2) = \alpha \sum_j \binom{\lambda'_j}{2} - \sum_i \binom{\lambda_i}{2}$$

for all λ. The argument in Hanlon (1992) shows that the $C^\theta(\lambda, \mu)$ are left eigenvectors of the matrix m_n^θ. To determine the corresponding eigenvalue, observe that the first column of m_n^θ is $(1 - \theta, \frac{1}{\binom{n}{2}}, 0, 0, \ldots, 0)^t$. 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_θ^k *be the probability on* S_n *associated to the Markov chain defined from* M_n^θ *starting as* id, *with stationary distribution* P_θ *defined at (4.1). Let*

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

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

$$\|P_\theta^k - P_\theta\| \leq f(\theta, c).$$

PROOF. From (2.5),

$$4\|P_\theta^k - P_\theta\|^2 \leq \theta^n n! \prod_{i=1}^{\cdot n}(1 + \theta(i - 1)) \sum_{\substack{\lambda \vdash n \\ \lambda \neq (n)}} \frac{1}{j_\lambda} \beta_\lambda^{2k}.$$

Here β_λ and j_λ 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 β_λ is monotone if the majorization 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

$$\theta^n n! \prod_{i=1}^n (1 + \theta(i - 1))/j_{n-1,1}$$

$$= \frac{\theta^n n!(1+\theta)(1+2\theta)\cdots(1+(n-1)\theta)}{\theta^2(n-1)!(1+\theta)(2+\theta)\cdots(n-3+\theta)(n-1+\theta)(n-2+2\theta)}$$

$$= f(\theta)n^{2+(\frac{1}{\theta}-\theta)}(1 + O(\frac{1}{n}))$$

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

$$\frac{\theta^n n! \prod_{i=1}^n (1 + \theta(i-1))}{j_\lambda}\beta_\lambda^{2k} \le f(\theta)n^{2+(\frac{1}{\theta}-\theta)}e^{-\frac{4\theta}{n}(an\log n - 2C^n)}(1 + O(\frac{1}{n}))$$

$$= f(\theta)e^{-8\theta C}(1 + O(\frac{1}{n})).$$

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

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

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. $\square$

REMARKS. 1) The function $a(\theta) = \frac{1}{2\theta} + \frac{1}{4\theta}(\frac{1}{\theta} - \theta)$ increases as θ decreases from 1 to 0 so it takes longer to converge for small θ. 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 $h = an\log n - cn$ 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 θ. 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
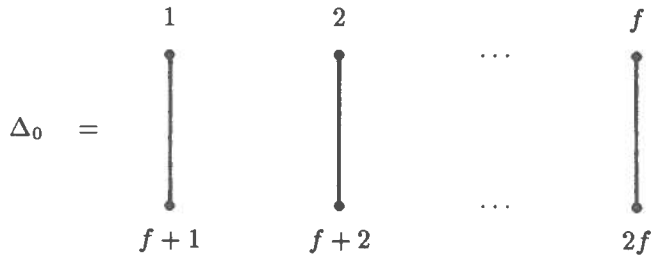
of these will be obtained from the first by appropriately introducing signs in the transition matrix. In view of this, it will not technically be a Markov chain. Nevertheless, the Metropolis Algorithm still applies and gives us an interesting twisting of our "signed" Markov chain.

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 $\{(1,2),(3,4),\ldots,(2f-1,2f)\}$. 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 Ω_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 Ω_f. The isotropy subgroup can be identified with $B_f = \mathbb{Z}_2^f \alpha S_f$, 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 Ω_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 $1, 2, \ldots, f$ in a top row and the points $(f+1), \ldots, 2f$ in a bottom row. For example, the one-factor Δ_0 which has an edge from i to $i+f$ for $i = 1, 2, \ldots, f$ is drawn as:

$$
\Delta_0 \quad = \qquad
\begin{array}{cccc}
1 & 2 & & f \\
\bullet & \bullet & \cdots & \bullet \\
\mid & \mid & & \mid \\
\bullet & \bullet & \cdots & \bullet \\
f+1 & f+2 & & 2f
\end{array}
$$

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 Ω_f in the following way. Given a 1-factor δ, the probability $R_f(1)_{\delta,\tau}$ of moving to another 1-factor τ is given by

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

For example, with $f = 2$ we have

$$R_2(1) \quad = \quad \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix} .$$

Let δ be in Ω_f. The *type of* δ 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 Γ on Ω_f by

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

where α 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 Γ so we are in a position to apply the Metropolis algorithm to obtain a twisted Markov chain $M(\alpha)$ whose stable distribution is Γ. Let $R_f(\alpha)$ denote the transition matrix of $M(\alpha)$. For example,

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

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 λ of f let v_λ be defined by*

$$v_\lambda = 2^{-\ell(\lambda)} \left\{ \sum_{\substack{\delta \\ \Lambda(\delta, \Delta_n) = \lambda}} \delta \right\},$$

and let V denote the span of the v_λ. 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 $J_\lambda(\mathbf{x}; \alpha)$ denotes the Jack symmetric function indexed by λ and that $c_{\lambda,\mu}^\alpha$ is the coefficient of $p_\mu(x)$ when $J_\lambda(\mathbf{x}; \alpha)$ expanded in terms of power sums. The main result of this section can now be stated.

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

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 π_λ be the permutation in S_f given by

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

(so π_λ has cycle type λ) and let δ_λ be the 1-factor with an edge from i to $f+\pi(i)$ for $i = 1, 2, \ldots, f$. So δ_λ looks like
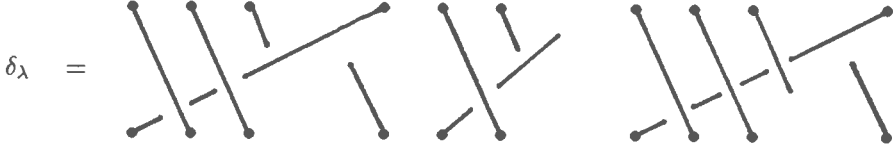
It is easy to check that δ_λ has type λ (so δ_λ is amongst the one-factors appearing in the sum v_λ).

We can make use of the one-factors δ_λ 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 λ and μ be partitions of f. The λ, μ entry of $\widehat{R}_f^{(f,\phi)}(\alpha) :=$ $\alpha\binom{2f}{2} R_f^{(f,\phi)}(\alpha)$ is given by:*

(a) *If $\ell(\mu) \geq \ell(\lambda)$ then $(\widehat{R}_f^{(f,\phi)}(\alpha))_{\lambda\mu}$ is $\alpha 2^{\ell(\mu)-\ell(\lambda)}$ times the number of δ of type λ which can be obtained from δ_μ by switching a pair of points.*

(b) *If $\ell(\mu) < \ell(\lambda)$ then $(\widehat{R}_f^{(f,\phi)}(\alpha))_{\lambda\mu}$ is $2^{\ell(\mu)-\ell(\lambda)}$ times the number of δ of type λ which can be obtained from δ_μ by switching a pair of points (you count δ 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 λ, μ entries of $(\widehat{R}_f^{(f,\phi)}(\alpha))_{\lambda\mu}$ for all λ. According to the lemma we need to count how many δ' of each type λ arise by switching a pair of points (x, y) in δ_μ.

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

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

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

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

For fixed μ_i and μ_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 λ, μ entry of $(\widehat{R}_f^{(f,\phi)}(\alpha))_{\lambda\mu}$ contains a factor of 2α.

CASE 2. x and y come from the same cycle $\mathcal{C}$ of $\delta_\mu \cup \Delta_0$. Let μ_i be the part of μ corresponding to $\mathcal{C}$ so $\mathcal{C}$ consists of $2\mu_i$ points, μ_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 μ_i. Then $((x, y)\delta_\mu) \cup \Delta_0$ corresponds to the partition

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

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

Note that in case (A), $(\widehat{R}_f^{(f,\phi)}(\alpha))_{\lambda\mu} = \frac{1}{2}$; whereas in case (B), $(\widehat{R}_f^{(f,\phi)}(\alpha))_{\lambda\mu} = \alpha$.

Applying the above facts we see that the μth column of $(\widehat{R}_f^{(f,\phi)}(\alpha))_{\lambda\mu}$ is given by the above transformation:

$$\boxed{\mu^{\text{th}} \text{ column of } \widehat{R}_f^{(f,\phi)}(\alpha)}$$

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

It is worth commenting on how we computed the coefficient of v_μ on the right-hand side. For starters there are $\sum_i\mu_i^2$ pairs (x,y) that yield the same μ. 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_μ, v_μ entry. The latter contribute a more complicated factor, namely, the δ_μ, δ_μ entry of $R_f(\alpha)$. We need to compute the entry. First observe that the δ_μ, δ_μ entry in $R_f(1)$ is f (pairs (x,y) which fix δ_μ must be endpoints of an edge of δ_μ). The off-diagonal entries of $\widehat{R}_f(\alpha)$ are either $0, 1$, or α. It follows that the δ_μ, δ_μ entry of $\widehat{R}_f(\alpha)$ is $f\alpha + H(\alpha-1)$ where H is the number of off-diagonal entries equal to 1 in the δ_μ^{th} 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 δ_μ, δ_μ entry in $\widehat{R}_f(\alpha)$ is

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

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

$$\psi(v_\lambda) = p_\lambda(\mathbf{x}).$$

Now let ρ be the linear transformation on Λ_n^f given by

$$\rho = \psi \circ \widehat{R}_f(\alpha) \circ \psi^{-1}.$$

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

(5.5)
$$\rho(p_\lambda(\mathbf{x})) = \{(f\alpha) + 2(2\alpha-1)n(\lambda')\}p_\lambda(\mathbf{x})$$
$$+ 2(2\alpha)\sum_{u\neq v}\lambda_u\lambda_v p_{\lambda[\lambda_u,\lambda_v\leftarrow\lambda_u+\lambda_v]}(\mathbf{x})$$
$$+ 2(\frac{1}{2}\sum_k\lambda_k\sum_{j=1}^{\lambda_k-1}p_{\lambda[\lambda_k\leftarrow j,\lambda_k-j]}(\mathbf{x})).$$

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 α.

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 Δ_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 Δ_0 which gives us the underlying $f \times f$ permutation matrix $\hat{\sigma}$. If σ maps the ith edge of Δ_0 to the jth (so there is a 1 in the i,j entry of $\hat{\sigma}$). Then $\sigma_{ij} = 1$, if σ 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 σ as an $f \times f$ signed permutation matrix). Then

$$\delta_0(\sigma) = 1$$
$$\delta_1(\sigma) = \text{sign}(\hat{\sigma})$$
$$\delta_2(\sigma) = \det(\sigma)$$
$$\delta_3(\sigma) = \delta_1(\sigma)\delta_2(\sigma).$$

There is an alternative description of the character δ_3. If we think of B_f as a subgroup of S_{2f} then δ_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 δ_i,

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

Let X_i denote the left deal $e_i \mathbb{C}S_{2f}$ and let V_i denote the two-sided ideal $e_i \mathbb{C}S_{2f} e_i$. 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 τ be a permutation in S_{2f} written in 1-line form,

$$\tau = a_1 a_2 \cdots a_{2f}.$$

Assign to τ 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 Ω_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 Γ 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 Ω_f be identified with X_0 as above. Then $R_f(1)$ is the matrix for right multiplication by $\frac{1}{|2f|} \displaystyle\sum_{1 \leq i \leq j \leq 2f} (i, j)$.*

We will now follow the same construction as in Section 5.1 but with X_0 replaced by X_3. The stable distribution Γ will be the same and the untwisted Markov chain will again be multiplication by $\binom{2f}{2}^{-1} \sum (i, j)$. 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, δ, let $\tau(\delta)$ be the lexicographically minimal element α of S_{2f} with $\delta(\alpha) = \delta$.*
(B) *For each partition λ of f, let $\pi(\lambda)$ be the lexicographically minimal element β 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 $\mathcal{C}$ be the vector space for V_3 given by

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

We henceforth let $v_\lambda^{(3)}$ denote $2^{\ell(\lambda)}e_3$. Let $Q_f(1)$ be the matrix for right multiplication by $\binom{2f}{2}^{-1} \sum (i, j)$ with respect to the basis B. Let $Q_f(\alpha)$ be the matrix obtained by applying the Metropolis algorithm to $Q_f(1)$ and Γ. 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 $\mathcal{C}$. We can now state the analogue of Theorem 4.

THEOREM 2. *For each λ, the vector $\sum_{\mu \vdash f} c_{\lambda\mu}^{(\alpha/2)} v_\mu^{(3)}$ is a left eigenvector of $Q_f^{(\phi,f)}(\alpha)$ with corresponding eigenvalue $(-f\alpha + 2\{\frac{\alpha}{2}n(\lambda') - n(\lambda)\})/\alpha(\binom{2f}{2i})$.*

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 δ_1 and δ_2. A simple computation shows that the restriction of right multiplication by $\binom{2f}{2}^{-1} \sum(i,j)$ 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 $\left(2\binom{2f}{3}^{-1} \sum(i,j,k)\right)$. 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.

REFERENCES

1. D. Critchlow, *Metric Methods for Analyzing Partially Ranked Data*, Lecture Notes in Statistics, No. 34, Springer-Verlag, Berlin, 1985.
2. P. Diaconis, *A random walk on partitions and zonal polynomials*, Unpublished manuscript, 1986.
3. _____, *Group Representations in Probability and Statistics*, Institute for Mathematical Statistics, Hayward, CA, 1988.
4. P. Diaconis and L. Saloff-Coste, *Comparison theorems for reversible Markov chain*, Technical report, Dept. of Mathematics, Harvard University, 1992.
5. P. Diaconis and M. Shahshahani, *Generating a random permutation with random transpositions*, Z. Wahrscheinlichkeitstheorie Verw. Gebiete **57** (1981), 183-195.
6. P. Diaconis and S. Zabell, *Updating subjective probability*, J. Amer. Statist. Assoc. **77** (1982), 822-830.
7. J. Hammersley and D. Handscomb, *Monte Carlo Methods*, Chapman and Hall, London, 1964.
8. P. Hanlon, *Jack symmetric functions and some combinatorial properties of Young symmetrizers*, J. Combin. Th. A **47** (1988), 37-70.
9. _____, *A Markov chain on the symmetric group and Jack's symmetric functions*, to appear, Discrete Math. (1992).
10. M. Kac, Amer. Math. Monthly **54** (1947), 369-391.
11. M. Krawtchouk, *Sur une généralisation des polynomes d'Hermite*, Comptes Rendus Acad. Sci. Paris **189** (1929), 620-622.
12. G. Letac and L. Takacs, *Random walks on the m-dimensional cube*, J. reine angew. math. **310** (1979), 187-195.
13. I. Macdonald, *Symmetric Functions and Hall Polynomials*, Clarendon Press, Oxford, 1979.
14. _____, *Orthogonal polynomials associated with root systems*, unpublished manuscript, 1989.
15. F. MacWilliams and N. Sloane, *The Theory of Error Correcting Codes*, North-Holland, Amsterdam, 1977.
16. R. Stanley, *Some combinatorial properties of the Jack polynomial*, Adv. Math. **77** (1989), 76-115.
17. J. Stembridge, *On Schur's Q-functions and the primitive idempotents of a commutative Hecke Algebra*, technical report, Department of Mathematics, University of Michigan, 1992.

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY, CAMBRIDGE, MA 02138
DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, ANN ARBOR, MICHIGAN 48109