ANALYSIS OF A BOSE-EINSTEIN MARKOV CHAIN

Persi Diaconis

Department of Mathematics and Statistics

Stanford University

Abstract

This paper gives sharp rates of convergence to stationarity for a Markov chain generating Bose-Einstein configurations of n balls in k boxes. The analysis leads to curious identities for the arc sine distribution.

0. On a Personal Note In 1971, as a beginning graduate student at Harvard's Department of Statistics, I badly wanted to learn "real" probability. Someone told me that the deepest, best book was Paul-Andre Meyers' "Probability and Potential Theory". For the next year and a half, three of us ran a reading group on this book. We moved slowly, like ants on a page, without any global understanding but happy to be in the presence of a master. I can't say I internalized any abstract potential theory but I learned a lot of measure theory and the last chapter (on Choquet Theory) made a big impact on my ability to abstract deFinettis theorem. As the magisterial sequence of books by Dellacherie-Meyer evolved, my familiarity with the original made them welcome and accessible.

I only met Paul-Andre Meyer once (at Luminy in 1995). He kindly stayed around after my talk and we spoke for about an hour. I was studying rates of convergence of finite state space Markov chains. He made it clear that, for him, finite state space Markov chains is a trivial subject. Hurt but undaunted, I explained some of our results and methods. He thought about it and said, "I see, yes, those are very hard problems".

The *analytic* parts of Dirichlet space theory have played an enormous role in my recent work. I am sure that there is much to learn from the abstract theory as well. In the present paper I treat rates of convergence for a simple Markov chain. I am sorry not to have another hour with Paul-Andre Meyer. Perhaps he would say "This piece of our story might help you". Perhaps one of his students or colleagues can help fill the void.

1. Background The use of Markov chains in Monte Carlo simulations has become a mainstay of scientific computing. There are a bewildering variety of methods for constructing reversible Markov chains with a given stationary distribution. Liu [2001] is a good overview of the present state of the art. Some order has appeared by the realization that many different algorithms are special cases of one general algorithm. Known variously as auxiliary variables (Besag and Green [1993]), data augmentation (Tanner and Wong [1987]), slice sampling (Neale [2003]) or hit and run (Belisle et al. [1991]), these algorithms include the celebrated Swedsen-Wang algorithm of statistical mechanics. They seem to allow "big moves" which suggests rapidly converging chains.

There has been very little rigorous work on rates of convergence for any of these algorithms. One spectacular exception are the negative result of Gore and Jerrum (1997) and Borgs et al. [1999] showing that the Swedsen-Wang algorithm does not mix rapidly at the critical temperature. Huber [2000] proves a rapid mixing for Swedsen-Wang suitably far from the critical temperature. The discussion in Neale [2003] contains pointers to a few examples

where proofs have been possible.

The present paper studies a class of problems called the "Burnside Process", introduced by computer scientists Mark Jerrum and Leslie Goldberg. These are a special case of the algorithms above; even here, general convergence results are far off in the future, but a successful analysis is possible for a subclass of problems with Bose-Einstein stationary distributions. The Burnside process is closely connected to Polya's method of enumeration. A short overview of this is contained in Appendix 1.

Let $\mathcal{X}$ be a finite set. Let G be a finite group acting on $\mathcal{X}$. This splits $\mathcal{X}$ into disjoint orbits $\mathcal{X} = O_1 \cup O_2 \cup \ldots \cup O_k$. The problem is to choose an *orbit* uniformly at random. The problem of picking unlabeled objects at random is familiar to probabilists from Bose-Einstein statistics. Another example arises in enumerating trees. It is well known that there are n^{n-2} labeled trees on n vertices and it is easy to pick a random tree using e.g. Prüfer codes. There is no simple enumeration of unlabeled trees and generating random subtrees of a graph is an active research area.

Goldberg and Jerrum have developed a Markov chain called the Burnside process on $\mathcal{X}$ which has a uniform stationary distribution when lumped to orbits. From $x \in \mathcal{X}$, choose uniformly among all $g \in G$ with $x^g = x$. Given g, choose uniformly among all y with $y^g = y$. The chain moves from x to y. If $\mathcal{X}_g = \{x : x^g = x\}$, $G_x = \{g : x^g = x\}$, and 0_x is the G orbit containing x, it is easy to see that this Burnside process is a reversible Markov chain on $\mathcal{X}$ with transition matrix and stationary distribution

$$K(x, y) = \frac{|0_x|}{|G|} \sum_{g \in G_x \cap G_y} 1/|\mathcal{X}_g|, \quad \pi(x) = \frac{z^{-1}}{|0_x|},$$

where z is a normalizing constant (which in fact equals the number of orbits). It follows that the chain lumped to orbits has a uniform stationary distribution.

2. Bose-Einstein Statistics   If n balls are dropped into k boxes so that each configuration of unlabeled balls is equally likely, the resulting probability distribution on $\binom{n+k-1}{k-1}$ configurations is called the Bose-Einstein distribution. In statistics it is sometimes called the beta-binomial or Dirichlet-multinomial distribution. See Feller [1968] for background. To put things into the notation of Section 1, let $[k] = \{1, 2, \cdots, k\}$ and $\mathcal{X} = [k]^n$. The coordinates of a vector in $\mathcal{X}$ represent the various balls and x_i represents the box containing ball labeled i. The symmetric group $G = S_n$ acts on $\mathcal{X}$ by permuting coordinates and the problem of choosing a random orbit becomes the problem of choosing a random Bose-Einstein

2

configuration.

Let G_x be the subgroup of S_n permuting coordinates with equal entries in x. If x contains n_j entries labeled j, $1 \leq j \leq k$, then $G_x \cong S_{n_1} \times S_{n_2} \times \ldots \times S_{n_k}$. Let the set of points in $\mathcal{X}$ fixed by g be denoted $\mathcal{X}_g$. This is the set of vectors having constant values on the cycles of the permutation g. From here, the Burnside process is easy to describe explicitly:

(2.1) From x, identify the set of coordinates I_j with common value j, $1 \leq j \leq k$. Choose uniform random permutations $w_1, w_2, \ldots, w_k$ of these sets. Break each w_i into cycles and label the coordinates of each cycle with a uniform choice in $[k]$. Let the final configuration be y.

The main result of this paper may now be stated:

Theorem 1   For any fixed k and n, let $K(x,y)$ be the transition matrix defined in (2.1) on $[k]^n$. Let $\pi = 1 \Big/ \binom{n+k-1}{k-1}$ be its stationary distribution. Then, there is $c = c(k)$ such that for all n and ℓ,

$$\|K_0^\ell - \pi\|_{TV} \leq (1-c)^\ell$$

with K_0 the chain started with all coordinates equal.

Remarks   The theorem shows that for fixed k the mixing time is independent of n. In the proof we show that $c(2)$ may be taken as $1/\pi$ (with $\pi = 3.14159...$) and give bounds for other values of k. The proof of the theorem is given in Section 3. It is based on an explicit expression for the transition matrix of the lumped chain. This involves a curious appearance of the discrete arc sine distribution. Section 4 gives lower bounds and a coupling bound of Aldous. The Appendix gives a brief, self-contained overview of the needed Polya theory. Of course, there are other, easier ways to directly generate Bose-Einstein configurations. One may place the n balls sequentially into k boxes, each time choosing a box with probability proportional to its current content plus one. Starting from the empty configuration this results in a Bose-Einstein distribution for every stage. The study of the Burnside process for this example is a prelude to more substantial studies.

3. Proof of Theorem 1   The argument is given in detail for general n and $k = 2$ with indication of what is needed for generalization at the end. By construction, for any $x, y \in \mathcal{X}$ and any $g \in S_n$, $K(x,y) = K(x^g, y^g)$ and so $K^\ell(x,y) = K^\ell(x^g, y^g)$ for all ℓ. It follows that Dynkińs criterion (Kemeny and Snell [1960, Chap. 3]) is satisfied and so the chain lumped to orbits is a Markov chain. When $k = 2$, the orbits are $0_0, 0_1, \ldots, 0_n$ with 0_i the set of

3

all binary vectors of length n with i ones and $n - i$ zeros. Let $\bar{K}(i, j)$ be the transition probabilities for the lumped chain $0 \le i,\ j \le n$. Let $\bar{\pi}(i) \equiv 1/(n+1)$ be the uniform distribution. By general theory,

$$\|K_0^\ell - \pi\| = \|\bar{K}_0^\ell - \bar{\pi}\|$$

so it is enough to study the lumped chain.

Let $\alpha_k^n = \dbinom{2k}{k} \dbinom{2n - 2k}{n - k} \Big/ 2^{2n}$ be the discrete arc sine distribution $0 \le k \le n$ (Feller [1968, Chap. 3]). We show that

(3.1) $$\bar{K}(0, k) = \alpha_k^n = \bar{K}(n, k)$$

(3.2) $$\bar{K}(j, k) = \sum_\ell \alpha_\ell^j \alpha_{k-\ell}^{n-j}, \quad (j + k - n)_+ \le \ell \le j \wedge k.$$

(3.3) $$\bar{K}(j, k) = \bar{K}(k, j) = \bar{K}(n - j, k) = \bar{K}(j, n - k) \text{ for all } k, j.$$

The proof of all parts of (3.1)-(3.3) follows from the lumping argument and simple symmetry, save only the assertion for $\bar{K}(0, k)$.

To prove $\bar{K}(0, k) = \alpha_k^n$ we recall Polya's cycle index. If a permutation g has $a_i(g)$ cycles of length i, define the polynomial

(3.4) $$p_n(x_1, \ldots, x_n) = \frac{1}{n!} \sum_{g \in S_n} \prod_{i=1}^n x_i^{a_i(g)}. \quad (p_0 = 1)$$

Polya proved that the sum of these polynomials factors

$$\sum_{n=0}^\infty t^n p_n = \prod e^{t^i x_i / i}.$$

Consider first $\bar{K}(0, 0) = \bar{K}_n(0, 0)$. A zero $\to$ zero transition happens if and only if each cycle is labeled zero. We are thus claiming

$$\bar{K}_n(0, 0) = \frac{1}{n!} \sum_{g \in S_n} \left(\frac{1}{2}\right)^{c(g)} = \alpha_0^n = \frac{\dbinom{2n}{n}}{2^{2n}} \text{ with } c(g) = a_1 + \ldots + a_n$$

the number of cycles in g. This follows by setting all $x_i = \frac{1}{2}$ in the cycle index:

$$\sum_{n=0}^{\infty} t^n \bar{K}_n(0,0) = \prod_{i=1}^{\infty} e^{t^i/(2i)} = \frac{1}{\sqrt{1-t}} = \sum_{i=0}^{\infty} \binom{-\frac{1}{2}}{i}(-t)^i = \sum_{n=0}^{\infty} \frac{\binom{2n}{n}}{2^{2n}} t^n.$$

This proves 3.1 for $k = 0$.

Consider next $\bar{K}_n(0,1)$. This counts events where each cycle except for one fixed point is labeled zero and the fixed point is labeled one. There are a_1 fixed points that can be chosen. Hence

$$\bar{K}_n(0,1) = \frac{1}{n!} \sum_{g \in S_n} a_1 \left(\frac{1}{2}\right)^{a_1+\dots+a_n}.$$

We get the generating function for these numbers by differentiating (3.4) once in x_1, multiplying by x_1, and setting all $x_i = \frac{1}{2}$. Thus

$$\sum_{n=0}^{\infty} t^n \bar{K}_n(0,1) = \frac{t}{2} \frac{1}{\sqrt{1-t}} = \sum_{n=0}^{\infty} \frac{\bar{K}_n(0,0)}{2} t^{n+1},$$

$$\bar{K}_n(0,1) = \frac{\bar{K}_n(0,0)}{2} = \frac{2\binom{2n-2}{n-1}}{2^{2n}}.$$

For the general case, $\bar{K}_n(0,j)$ we sum over partitions of j:

$$\bar{K}_n(0,j) = \frac{1}{n!} \sum_{\substack{\lambda \vdash j \\ g \in S_n}} \prod_{i=1}^{j} \binom{a_i(g)}{b_i(\lambda)} \left(\frac{1}{2}\right)^{a_1+\dots+a_n}$$

where g has $a_i(g)$ i-cycles and λ has $b_i(\lambda)$ parts equal to i. Differentiating (3.4) in x_i b_i times and multiplying by $x_i^{b_i}$ gives $a_i(a_i - 1)\dots(a_i - b_i - 1)$ in the generating function. This also brings down a factor of $\left(\frac{t^i}{i}\right)^{b_i}$. Finally, all x_i are set to $1/2$. The upshot is

$$\sum_{n=0}^{\infty} t^n \bar{K}_n(0,j) = \frac{t^j}{\sqrt{1-t}} \sum_{\lambda \vdash j} \prod_{i=1}^{j} \frac{1}{(b_i(\lambda))!(2i)^{b_i}}.$$

Multiply the sum over λ by $j!/j!$ to get

$$\frac{t^j}{j!\sqrt{1-t}} \sum_{\lambda \vdash j} \left(\frac{1}{2}\right)^{b_1+\dots+b_j} \frac{j!}{\prod_{i=1}^{j} b_i! i^{b_r}} = \frac{t^j}{\sqrt{1-t}} K_j(0,0) = \frac{t^j}{\sqrt{1-t}} \frac{\binom{2j}{j}}{2^{2j}}.$$

5

This proves (3.1) on comparing coefficients.

The proof of Theorem 1 is completed by showing that $\bar{K}$ satisfies a Doeblin Condition. As is well known, the arc sine distribution is smallest for $j = \lfloor n/2 \rfloor$ when it has the following asymptotics

$$\bar{K}_n(0, \lfloor n/2 \rfloor) \sim \frac{1}{\pi n}.$$

By a straightforward induction $\bar{K}_n(i,j) \geq \bar{K}_n(0, \lfloor n/2 \rfloor)$ for all i, j. Thus

$$\bar{K}_n(i,j) \geq c\bar{\pi}(j) \quad \text{all} \quad i, j$$

where $c \sim 1/\pi$ for large n. This Doeblin Condition shows the total variation distance of $\bar{K}^\ell$ to $\bar{\pi}$ is at most $(1 - c)^\ell$, as desired. $\qquad\qquad\qquad\square$

Remarks

1. The appearance of the discrete arc sine distribution from labeling cycles is apparently new. It can be said without mentioning permutations: Divide $[n] = \{1, 2, \ldots, n\}$ into pieces as follows: choose $j_1 \in [n]$ uniformly. The first piece is $\{1, 2, \ldots, j_1\}$. The second piece is chosen by choosing j_2 uniformly in $\{j_1+1, \ldots, n\}$. This continues until n is chosen. Call this 'discrete stick breaking.' Label the pieces by flipping a fair coin for each piece. The sum of the lengths of the 'heads' pieces has the discrete arc sine distribution. This result was guessed from the infinite version: break the unit interval into countably many pieces by uniform stick breaking. Flip a fair coin for each piece. The sum of the lengths of the 'heads' pieces has the arc sine density $\frac{1}{\pi\sqrt{x(1-x)}}$ on $[0, 1]$. This known result (Chen and Zame [1981], Kerov [1993]) suggested the discrete result. The fact that the uniform distribution on $\{0, 1, \ldots, n\}$ is stationary for $\bar{K}$ gives (after passage to the limit as $n \to \infty$) the following result for the continuous arc sine law. Let U be uniform on $[0, 1]$. Let X_1 and X_2 be independent arc sine variables. Then

$$UX_1 + (1 - U)X_2 \text{ has a uniform distribution.}$$

2. Jim Pitman has shown us a neat generalization of the discrete arc sine results. If the original permutation is chosen from the Ewens distribution on permutations

$$P_\theta(w) = Z^{-1}\theta^{c(w)} \quad 0 < \theta \leq 1, \quad Z(\theta) = (1 + \theta)(1 + 2\theta)\ldots(1 + (n-1)\theta),$$

and the cycles are colored zero or one with probability p, $1 - p$, the sum of the lengths of the cycles labeled one has a discrete $\beta_{\theta p, \theta(1-p)}$ distribution:

$$\text{(3.5)} \qquad\qquad P(\text{length} = k) = \binom{n}{k} E(X^k(1 - X)^{n-k})$$

with X having a beta $(p\theta, (1-p)\theta)$ distribution. The integrals are easy to do and agree with the special case above. The form (3.5) is provable from the developments around the Blackwell-McQueen version of Polya's Urn and the Pitman-Dubins 'Chinese Restaurant' processes.

What is *not* obvious is why the special case treated above $(\theta = 1, p = 1/2)$ agrees with the discrete arc sine distribution from elementary probability. There must be some bijective proof that relates 2-colorings of cycles to coin-tossing paths.

3. The proof given above is for $k = 2$. For general k the lumping argument works to show it is enough to consider the orbit chain which takes values on compositions $(y_1, \ldots, y_k)$ with y_i equal to the number of times color i occurs. Thus $0 \le y_i \le n$ with $y_1 + \ldots + y_k = n$. The stationary distribution of the lumped chain is the Bose-Einstein distribution $\bar{\pi}(y_1 \ldots y_k) = \dfrac{1}{\binom{n+k-1}{k-1}}$. The transition matrix $\bar{K}$ is now indexed by compositions of n into k parts. It is determined as above by knowing the chance of going from $(n, 0, \ldots, 0)$ to $(y_1, \ldots, y_k)$. Using Polya theory, this equals

(3.6)
$$\bar{K}_n(n0\ldots0; y_1 \ldots y_k) = \binom{n}{y_1 \ldots y_k} \frac{1}{n!} \left[\Gamma\left(\frac{1}{k}\right)\right]^k \prod_{i=1}^{k} \Gamma\left(y_i + \frac{1}{k}\right) = \binom{n}{y_1 \ldots y_k} \frac{1}{k^n n!} \prod_{i=1}^{k} \prod_{j=1}^{y_i - 1} (1 + kj).$$

This can be shown to be minimal when all the y_i are within one of n/k (if k divides n all $y_i = n/k$). We need Gauss's approximation for the Gamma function in the form
$$\Gamma(z) = \lim_{n \to \infty} \frac{n! n^{z-1}}{z(z+1)\ldots(z+n-1)}, \quad \prod_{j=1}^{m} (1 + kj) \sim \frac{k^{m+1} m! m^{1/k}}{\Gamma(1/k)}.$$
From here, when all $y_i = n/k$, calculation shows that $K_n(n0\ldots0, \frac{n}{k} \ldots \frac{n}{k}) \sim \frac{k^{k-1}}{n^{k-1}\Gamma(\frac{1}{k})^k}$. When k is large $\Gamma(\frac{1}{k})^k \sim k^k$ so $K_n(n\ldots0; \frac{n}{k} \ldots \frac{n}{k}) \sim \frac{1}{kn^{k-1}}$. We thus have $\frac{1}{kn^{k-1}} \ge c(k) \frac{(k-1)!}{(n+1)\ldots n+k-1}$ provided that $\frac{(1+\frac{1}{n})\ldots(1+\frac{k-1}{n})}{k!} \ge c(k)$. This entails $c(k) \sim \frac{1}{k!}$.

4. Aldous' Theorem and Lower Bounds

In Aldous and Fill [2003, Chapter 4-3] David Aldous proved a remarkable bound for the Bose-Einstein walk. His bound works for general values of n and k.

Theorem 2. (Aldous). For the Burnside walk applied to n balls dropped into k boxes defined at (2.1)
$$\|K_x^\ell - \pi\|_{tv} \le n\left(1 - \frac{1}{k}\right)^\ell.$$
The upper bound is uniform in the starting state $x \in [k]^n$.

Remarks   For k large, this is markedly better than Theorem 1. However, for fixed k (or k growing very slowly with n) Theorem 1 gives better bounds. Aldous uses an inspired coupling and a careful study of his argument as well as effort to apply it to more general problems in this class seems fully warranted.

Turn next to lower bounds. For fixed k and large n, Theorem 1 shows the Burnside walk converges in a finite number of steps. For k growing with n, Aldous' theorem shows order $k \log n$ steps suffice. The following result shows that for $k = n$ at least order $\log n$ steps are needed. We conjecture that this is the correct answer when k is of order n.

Proposition   For the Burnside walk applied to n balls dropped into n boxes, defined at (2.1). There is a fixed constant $c > 0$ such that for $\ell \leq \log n$.

$$\|K_0^\ell - \pi\|_{TV} \geq c$$

Proof   If n balls are dropped into n boxes using Bose-Einstein statistics, the configuration has the distribution of n Geometric $(\frac{1}{2})$ variables conditional on their sum being n. By standard conditioned limit arguments (see e.g. Holst (1979)), the maximum box count has the same limit distribution as the maximum of n independent identically distributed geometric $(\frac{1}{2})$ variables. This is of order $\log n$. Thus, under the stationary distribution, the maximum box count is of order $\log n$.

On the other hand consider the Burnside process started at a configuration with all balls in a given cell. The first step chooses a random permutation uniformly in S_n and labels the cycles independently with n colors. The largest cycle is of order n multiplied by a random variable with mean .61...(the length of the largest piece in uniform stick-breaking). See Billingsley [1972], Goncharov [1942] or Shepp-Lloyd [1966] for details. At step two, this largest cycle is broken into pieces, the largest of which is of order n multiplied by a product of two independent copies of L. Continuing, we see that the walk must be run order $\log n$ steps to have the sequence of largest subpieces drop to size $\log n$. Further details are omitted.

$\square$

APPENDIX: Pointers to Polya Theory

Polya theory concerns itself with question such as: How many ways can we paint ten dice with colors red, white, blue? Here the ordering of the dice and their position doesn't matter. As mathematics, the questions become: Let $\mathcal{X}$ be a finite set. Let a finite group $\mathcal{G}$ act on $\mathcal{X}$ This splits $\mathcal{X}$ into orbits $\mathcal{X} = 0_1 \cup 0_2 \ldots \cup 0_k$. What is k? In the dice example,

identify the faces of the ten dice with points in $[6]^{10}$. Let $\mathcal{X}$ be all functions from $[6]^{10}$ into $\{R, W, B\}$. The group $\mathcal{G}_1$ of rotations of the usual cube acts on $[6]$; thus $\mathcal{G}_1^{10}$ acts on $[6]^{10}$. Similarly, the symmetric group S_{10} acts on $[6]^{10}$. Putting these actions together gives an action of $\mathcal{G} = S_{10} \ltimes \mathcal{G}_1^{10}$ on $\mathcal{X}$. The number of orbits equals the number of distinct colorings.

The best short introduction to Polya theory is in De Bruijn [1968] where one finds extensions to counting functions between two finite sets $\mathcal{A}, \mathcal{B}$ with groups acting on each side. Polya's original article was aimed at chemistry problems (count the number of labelings of a Benzine ring with two colors up to dihedral symmetry). A translation and a long survey of developments is in Polya and Read [1987]. Chemists are still interested – see Brocas-Gielen-Willem [1983]. An extensive mathematical development appears in Kerber [1994].

Polya theory can be seen as a chapter of symmetric function theory; indeed it is thus treated in the last section of the last chapter of Stanley (1999). To see the connection, consider $\mathcal{G}$ acting on $\mathcal{X}$. The *Cycle Index* is

$$Z_{\mathcal{G}}(P_1, P_2, \ldots, P_{|\mathcal{X}|}) = \frac{1}{|\mathcal{G}|} \sum_g \prod P_i^{a_i(g)}.$$

Here P_i are indeterminates and for permutations $g \in \mathcal{G}$, $a_i(g)$ is the number of i-cycles. As above, let $\mathcal{C} = \{c_1, c_2, \ldots\}$ be a set of colors and let $\mathcal{F} = \{f : \mathcal{X} \to \mathcal{C}\}$. Then $\mathcal{G}$ acts on $\mathcal{F}$ by $gf(x) = f(gx)$. For variables $z_1, z_2, \ldots$, define the *weight of f* as $wt(f) = z_1^{c_1(f)} z_2^{c_2(f)} \ldots$ with $c_i(f)$ the number of x with color c_i $(c_i(f) = |f^{-1}(c_i)|)$. The generating function

$$F_{\mathcal{G}} = \sum_{0 \in g/\mathcal{G}} wt(f)$$

summed over orbits 0 of $\mathcal{F}$, with any choice of $f \in 0$, has the coefficient of $z_1^{b_1} z_2^{b_2} \ldots$ the number of orbits with color i occurring b_i times. In the dice example, the number of colorings with one red, three white and six blue. The main theorem of Polya theory states

Theorem $Z_{\mathcal{G}}(P_1, P_2, \ldots) = F_{\mathcal{G}}(z_1, z_2, \ldots) = chIND_{\mathcal{G}}^{S_n}(1)$.

Here $P_i = \sum_j z_j^i$ is the ith power sum symmetric function. The Group $\mathcal{G}$ is a subgroup of S_n with $n = |\mathcal{X}|$. $IND_{\mathcal{G}}^{S_n}(1)$ is the permutation character and for any character χ, the characteristic map is $ch(\chi) = \sum_g \chi(g) \prod_i P_{a_i(g)}$.

Corollary   The number of inequivalent colorings of $\mathcal{X}$ with m colors equals $Z_{\mathcal{G}}(m, m, \ldots, m)$.

The symmetric function formulation of Polya theory allows tools such as character theory and Schur Functions. For a proof of the main theorem and remarkable applications, see Stanley (1999).

9

Computer science theorists have opened a new chapter in Polya theory by proving that the evaluation of $Z_{\mathcal{G}}(2, 2, \ldots 2)$ is #-P complete, even for $\mathcal{G}$ an Abelian 2-group. They also show that computing a single coefficient in $Z_{\mathcal{G}}(P_1, P_2, \ldots P_n)$ is intractable. The problems are reduced to counting the number of colorings of a graph. For these results see Goldberg [1993]. The probabilistic approach to approximate counting develops Markov chains to sample problem instances. If these chains can be proved to mix rapidly, then accurate, efficient approximation can be proved. This was the genesis of the Burnside process. Goldberg and Jerrum [2002] relate these problems to approximation of the partition functions of Ising and Potts Models. They use these connections to give examples where the Burnside process mixes slowly. More precisely, they show there is an infinite family of permutation groups G such that the mixing time of the Burnside process is exponential in the degree of G. The present example shows that for some group actions, the Burnside process mixes rapidly.

References

Aldous, D. and Fill, J. (2003). Reversible Markov chains and random walks and graphs. www.stat.berkeley.edu/users/aldous.

Besag, J. and Green, P. (1993). Spatial statistics and Bayesian computation. *Jour. Roy. Statist. Soc.* **B** 25-37.

Billingsley, P. (1972). On the distribution of large prime divisors. *Period. Math. Hungar* **2** 283-289.

Borgs, C., Chayes, J. Frieze, A. Kim, J. Tetali, P. Vigoda, E. Vu, V. (1999). Torpid mixing of some Monte Carlo Markov chains in statistical physics, 40th Symp. Foundations of Computer Science 218-229, IEEE Compu. Soc. Los Alamaitos CA.

Brocas, J., Gielen, M. and Willem, R. (1983). *The Permutational Approach to Dynamical Sterochemistry*, McGraw-Hill, New York.

de Bruijn, N.G. (1968). Polya's theory of counting, in *Applied Combinatorial Mathematics* (E. Beckenbeck, ed.) Wiley, New York.

Chen, R. and Zame, A. (1981). Another arc sine law. *Sankhyā* **43** 371-373.

Dellacherie, C. and Meyer, P.A. (1975). *Probabilités et Potentiel*, Herman, Paris.

Dellacherie, C. and Meyer, P.A. (1980). *Probabilités et Potentiel: Theorie des Martingales*, Herman, Paris.

Dellacherie, C. and Meyer, P.A. (1983). *Probabilité et Potentiel: Theorie Discrete du Potentiel*, Herman, Paris.

Feller, W. (1968). *An Introduction to Probability and its Applications*, Vol. 1, 3rd Ed., Wiley, New York.

Holst, L. (1979). Two conditioned limit theorems with applications, *Ann. Statist.* **7** 551-557.

Huber, M. (2000). Efficient sampling from the Ising model using Swendsen-Wang.

Goldberg, L. (1993). Automating Polya theory: The computational complexity of the cycle index polynomial, *Info and Computation* **105** 268-288.

Goncharov, V. (1942). On the distribution of the cycles in permutations. Dokl. AKademii. Nauk. SSSR, 35, 229-310.

Gore, V. and Jerrum, M. (1997). The Swendsen-Wang process does not always mix rapidly. In Proc. 29th *Acm. Symp. Th. of Comput.* 674-681.

Goldberg, L. and Jerrum, M. (2002). The Burnside process mixes slowly. *Combinatorics, Probability, and Computing* **11** 21-34.

Jerrum, M. (1993). Uniform sampling Modulo a group of symmetries using Markov chain simulation, it Dimacs Series in Discrete Math **10** 37-47.

Jerrum, M. (1995). Computational Polya-theory in surveys in combinatorics, *London Math. Soc. Lecture Notes* **218**, Cambridge University Press, Cambridge. pgs. 103-108.

Kemeny, J. and Snell, L. (1960). *Finite Markov Chains*, Van Nostrand, New York.

Kerber, A. (1999). *Applied Finite Group Actions*, Springer, Berlin.

Kerov, S. (1993). Transition probabilities for continual Young diagrams and Markov moment problems, *Func. Anal. Appl.* **27** 104-117.

Liu, J. (2001). *Monte Carlo Techniques in Scientific Computing*, Springer, New York.

Meyer, P.A. (1966). *Probability and Potentials*, Blaisdell, Walthan, Mass.

Neale, R. (2003). Slice sampling (With discussion). *Ann. Statist.* **31** 705-767.

Polya, C. and Read, R. (1987). *Combinatorial enumeration of groups, graphs and chemical compounds*, Springer, New York.

Stanley, R. (1999). *Enumerative Combinatorics* Vol. 2, Cambridge Press, Cambridge.

Shepp, L. and Lloyd, S. (1966). Ordered cycle lengths in random permutations, *Trans. Amer. Math. Soc.* **121** 340-357.

Tanner, M. and Wong, W. (1987). The calculation of posterior distributions using data augmentation. *Jour. Amer. Statist. Asoc.* **82** 528-550.