The Mixing Time for a Random Walk on the Symmetric Group Generated by Random Involutions

Megan Bernstein[†]

Georgia Institute of Technology, USA

Abstract. The involution walk is a random walk on the symmetric group generated by involutions with a number of 2-cycles sampled from the binomial distribution with parameter p. This is a parallelization of the lazy transposition walk on the symmetric group. The involution walk is shown in this paper to mix for $\frac{1}{2} \le p \le 1$ fixed, n sufficiently large in between $\log_{1/p}(n)$ steps and $\log_{2/(1+p)}(n)$ steps. The paper introduces a new technique for finding eigenvalues of random walks on the symmetric group generated by many conjugacy classes using the character polynomial for the characters of the representations of the symmetric group. This is especially efficient at calculating the large eigenvalues. The smaller eigenvalues are handled by developing monotonicity relations that also give after sufficient time the likelihood order, the order from most likely to least likely state. The walk was introduced to study a conjecture about a random walk on the unitary group from the information theory of back holes.

Résumé. La marche d'involution est une marche aléatoire sur le groupe symétrique générée par les involutions avec un nombre de 2-cycles tiré aléatoirement en suivant une distribution binomiale de paramètre p. Il s'agit d'une parallèlisation de la marche de transposition paresseuse sur le groupe symétrique. Nous montrons dans cet article que la marche d'involution mélange pour $\frac{1}{2} \le p \le 1$ fixé, et n suffisamment grand entre $\log_{1/p}(n)$ et $\log_{2/(1+p)}(n)$ étapes. Cet article introduit une nouvelle technique pour trouver les valeurs propres de marches aléatoires sur le groupe symétrique générées par de nombreuses classes de conjugaison en utilisant les caractères des représentations du groupe symétrique. Cette méthode est particulièrement efficace pour calculer les grandes valeurs propres. Les petites valeurs propres sont traitées en développant des relations de monotonicité qui donnent aussi après suffisamment de temps l'ordre de probabilité, c'est-à-dire l'ordre de l'état le plus probable à l'état le moins probable. La marche est introduite pour étudier une conjecture concernant une marche aléatoire sur le groupe unitaire provenant de la théorie de l'information sur les trous noirs.

Keywords. random walk, Markov chain, mixing time, representation theory of the symmetric group, character polynomial

1 Introduction

This paper examines how a natural notion of "parallelization" affects the rate of convergence to stationary for a random walk on the symmetric group. The base walk that this paper parallelizes is the p lazy random

[†]Email: `bernstein@math.gatech.edu`

transposition walk. It has generators chosen as the identity with probability p and a uniformly random transposition with probability $1-p$. This is equivalent to putting n cards on the table and with probability $1-p$ swapping a random pair. This walk takes order $n\log(n)$ steps to converge to its uniform stationary distribution [Diaconis and Shahshahani(1981)]. Suppose the walk is parallelized by simultaneously transposing s disjoint pairs at the same time. This is like taking s steps of the non-lazy transposition walk, except it guarantees $2s$ distinct cards are moved. This problem can be explored in several ways. For n even, the maximum number of disjoint transpositions is $n/2$. If these are chosen via a random matching, a randomly chosen fixed point free involution results. This walk was analyzed by Lulov [Lulov(1996)], who showed that 3 steps suffice for this walk to mix. In this paper, each transposition in a fixed point free involution is discarded with some probability. This is a parallelized p-lazy transposition walk. When this probability is fixed and at least $\frac{1}{2}$ the results here show that this walk has mixing time order $\log(n)$.

More specifically, this paper studies the random walk on S_n, for n even, generated by first choosing uniformly at random a fixed point free involution, also known as a perfect matching, then discarding or keeping each 2-cycle it contains independently with probability $p, 1-p$ respectively. This means the probability an involution with s 2-cycles is selected is $\binom{n/2}{s}p^{n/2-s}(1-p)^s$. By considering general p, this gives a family of walks with pn as the expected number of fixed points of a generator. Taking $p = 1 - 2/n$ gives an "expected transposition walk" where on average, a transposition will be selected, or as $p \to 0$, an expected fixed-point free involution walk. The author conjectures that mixing occurs with cutoff at $\log_{1/p}(n)$ for any p bounded away from 0. This paper, for $p \geq 1/2$ fixed, for n sufficiently large, establishes for mixing a lower bound of $\log_{1/p}(n)$ and an upper bound of $\log_{2/(1+p)}(n)$. These are separated by just over a factor of 2. A walk is said to mix when the total variation distance between the distribution of the walk after t steps, P^{*t} and its uniform stationary distribution, U, is less an a chosen constant.

Theorem 1. *For $t \geq \frac{\log(n)+c}{\log 2/(1+p)}$, $p \geq \frac{1}{2}$ fixed, and n sufficiently large,*

$$||P^{*t} - U||_{TV} \leq e^{-c/2}$$

Theorem 2. *For $p \geq \frac{1}{2}$, $t < \frac{\log(n)-c}{\log(1/p)}$, there exists $1 - \frac{\log(n)}{n} \leq A \leq 1$ so that,*

$$||P^{*t} - U||_{TV} \geq 1 - \frac{1}{\frac{1}{2} + A^2 e^{2c}}$$

This upper bound is found through a combination of two methods. Both use the expression of the eigenvalues of the walk in terms of the characters of the symmetric group. The character polynomial gives the characters of S_n as a polynomial in the cycle decomposition of a permutation. The eigenvalues of this walk, as seen in (1), are a linear combination of characters evaluated at the $n/2 + 1$ conjugacy classes of involutions. Since all these involutions only have 1- and 2-cycles, understanding the character polynomial in these cycles will give a strong bound on the large eigenvalues of the walk. To control the small eigenvalues, a recursive formula for the eigenvalues, Proposition 4 is used. This recursion constructed via the Murnaghan-Nakayama rule leads to a series of monotonicity conditions on the eigenvalues. These monotonicity conditions require that $p \geq \frac{1}{2}$.

As corollary to these monotonicity conditions, and so also restricted to $p \geq \frac{1}{2}$, a total order for the most likely to least likely permutation after sufficient time is identified in Corollary 10. At each step of a Markov chain, there is a partial ordering on the states given by the order from most likely to the least likely

state. A linear extension to a total order is called a likelihood order. With mild conditions, after sufficient time, this converges to a fixed likelihood order. For the involution walk, the limiting likelihood order is the cycle lexicographic order as defined in Definition 9. This means after sufficient time the identity will be the most likely element and an n-cycle the least likely element of the walk. This is the same likelihood order as $p \geq \frac{1}{2}$ lazy transposition walk [Bernstein(2015)].

As $p \to 0$, due to a parity problem, the walk can no longer mix in $O(\log(n))$ steps. The fixed point free involution walk at even steps is confined to A_n inside of S_n. While for any $p > 0$, the involution walk will mix to all of S_n. As $p \to 0$, the probability of selecting a fixed point free involution will grow to 1. Since it becomes more and more unlikely as $p \to 0$ anything other than a fixed point free involution is chosen, at even steps, the involution walk will be more and more prone to be stuck on even elements inside of S_n and take longer and longer to approach uniformity over all elements. This is shown in Proposition 22.

A random permutation can be made from n transpositions chosen systematically, called systematic scan, transposing in order each position with itself or a later position uniformly at random. If generating transpositions are chosen uniformly at random, this "random scan" is called the transposition walk. A random permutation takes $\frac{1}{2}n\log(n) + cn$ transpositions in the transposition walk to build [Diaconis and Shahshahani(1981)]. The largest impediment is a coupon collector problem of a constant fraction of the numbers not appearing in any transpositions from the first $\frac{1}{2}n\log(n) - cn$ steps. These never moved numbers are fixed points of the permutation, resulting in insufficiently random permutations. On the other hand, generating the random permutation from a walk generated by fixed point free involutions, takes 3 steps of the walk or $\frac{3}{2}n$ transpositions [Lulov(1996)]. By letting p vary, one can find this transition from when order n to order $n\log(n)$ transpositions are needed. The analysis here holds for fixed $p \geq \frac{1}{2}$ and n sufficiently large, for which it takes $O(n\log(n))$ transpositions to build a random permutation.

While studying the information theory of black holes, physicists became interested in a random walk on the unitary group. Information in the black hole is expressed in qubits, each an element of $\mathbb{C}^2$. Take as basis vectors $e_1 = (1,0)$, $e_0 = (0,1)$. The walk is on n qubits, and so it on a 2^n-dimensional space with a basis indexed by n-bit binary strings. At each step the walk takes a random $U(4)$ operator and applies it to two random qubits, acting on the binary strings indexing the basis. All 2^{n-2} basis vectors with the same 2-bit combination for those two qubits are effected the same way under the walk. This means each $U(4)$ operator acts 2^{n-2} times, giving rapid mixing for such a high dimensional space. This walk is known to "scramble" in $n\log(n)$ steps [Sekino and Susskind(2008)]. Recent work of Hayden and Preskill [Hayden and Preskill(2007)] and Sekino and Susskind [Sekino and Susskind(2008)] has developed interest in a version of this walk where $n/2$ commuting steps of the random walk are taken at once. A different $U(4)$ operator is chosen to each act on the different 2-cycles of a perfect matching of the qubits. This faster walk is conjectured to mix in $O(\log(n))$ steps. The involution walk was designed as a toy model to study the effects of independent random shuffling on the components of a perfect matching.

Section 2 describes the upper bound lemma from discrete Fourier analysis and the eigenvalues of the walk. Section 3 finds monotonic decay of the eigenvalues needed for the bounds. Section 4 finds an upper bound for the mixing time of the binomially distributed involution random walk. Section 5 finds the the lower bound of $\log_{\frac{1}{p}}(n)$.

2 Background

The method of establishing an upper bound for mixing time used here is spectral analysis using the character polynomial for the characters of the representations of the symmetric group. The upper-bound lemma

gives a bound on the total-variation distance between the walk and its uniform stationary distribution [Diaconis(1988)]. The mixing time is the first time that the total variation distance of the chain from its stationary distribution is at most $\frac{1}{4}$. The version of the upper-bound lemma below is specialized to conjugacy class walks on S_n, in which every element of a conjugacy class is equally likely. Partitions of n index the non-trivial irreducible representations of the symmetric group; the representation for the partition λ has dimension d_λ.

Proposition 3 (Diaconis-Shahshahani [Diaconis and Shahshahani(1981)])**.** *When $K(t)$ is a class function of an aperiodic, irreducible walk on S_n,*

$$\left\| K^{*t}(\sigma) - \frac{1}{n!} \right\|_{TV} \leq \frac{1}{4} \sum_{\lambda \neq 1} d_\lambda^2 \psi_\lambda^{2t}$$

The sum below is over conjugacy classes κ of size $|\kappa|$ with $K(\kappa)$ the probability of one element of the conjugacy class,

$$\psi_\lambda = \sum_\kappa |\kappa| K(\kappa) \frac{\chi_\lambda(\kappa)}{d_\lambda}$$

For this walk, the formula for the eigenvalue ψ_λ is the sum over conjugacy classes of the probability of it being a generator times its character ratio:

$$\psi_\lambda = \sum_{s=0}^{n/2} p^{n/2-s}(1-p)^s \binom{n/2}{s} \frac{\chi_\lambda(1^{n-2s}, 2^s)}{d_\lambda} \tag{1}$$

3   Monotonicity of Eigenvalues

The eigenvalues of the involution walk show intriguing connections to the eigenvalues of the transposition walk. For example, in the transposition walk, the eigenvalues decrease according to majorization order, where λ is smaller than ρ if the blocks of λ can be moved up and to the right to get ρ. Below, this is shown for any λ and for $\rho = [n-i, i]$. This is used in the upper bound section to get a bound for λ with $\lambda_1 < \frac{n}{2}$. The result also gives that the likelihood order after sufficient time is the cycle lexicographical order, the same order as for the transposition walk.

This is based upon the following recursive construction of the eigenvalues. This is derived from a probabilistic Murnaghan-Nakayama rule, in which two 1-cycles are removed with probability p or a two-cycle with probability $1 - p$. Examining all the configurations of border-strips that can be removed and their heights amounts to:

Proposition 4. *For ψ_λ defined in (1),*

$$\psi_\lambda = \sum_{\rho:\lambda/\rho=[2]} \psi_\rho \frac{d_\rho}{d_\lambda} + (2p-1) \sum_{\rho:\lambda/\rho=[1,1]} \psi_\rho \frac{d_\rho}{d_\lambda} + 2p \sum_{\rho:\lambda/\rho=[1]\cup[1]} \psi_\rho \frac{d_\rho}{d_\lambda}$$

The following observation will be useful in the arguments that follow.

Proposition 5. *The sum of the coefficients in the expansion of ψ in Proposition 4 decreases according to majorization order.*

Proof. Examining the formula in Proposition 4 without the ψ_ρ terms gives the eigenvalue indexed by λ from the p-lazy transposition walk. Moreover, a classical result of Frobenius [Diaconis and Shahshahani(1981)], shows $\frac{\chi_\lambda(\tau)}{d_\lambda}$ decreases along majorization order. For a formula for $\frac{\chi_\lambda(\tau)}{d_\lambda}$ see [Diaconis and Shahshahani(1981)]. $\square$

Lemma 6. *For $p \geq \frac{1}{2}$ and $i \leq n/2$, $\psi_{[n-i,i]}$ decreases as i increases.*

Proof. Using the formula

$$\psi_\lambda = \sum_{\rho:\lambda/\rho=[2]} \psi_\rho \frac{d_\rho}{d_\lambda} + (2p-1) \sum_{\rho:\lambda/\rho=[1,1]} \psi_\rho \frac{d_\rho}{d_\lambda} + 2p \sum_{\rho:\lambda/\rho=[1]\cup[1]} \psi_\rho \frac{d_\rho}{d_\lambda}$$

by induction on n and i, it will follow that $\psi_{[n-i,i]} \geq \psi_{[n-i-1,i+1]}$. For the complete proof see [Bernstein(2015)] $\square$ $\square$

A similar argument gives a second stepping stone.

Lemma 7. *For all λ with $\lambda_1 = n - i$, $p \geq \frac{1}{2}$,*

$$\psi_\lambda \leq \psi_{[n-i,i]}$$

Together, these two monotonicity relations on the large eigenvalues allows the construction of a bound on the smaller eigenvalues.

Theorem 8. *For $p \geq \frac{1}{2}$ and λ such that $\lambda_1' \leq \lambda_1 < \frac{n}{2}$, $\psi_\lambda \leq \psi_{[n/2,n/2]}$*

Proof. By Proposition 5 $[n/2, n/2]$ has the maximal value of the sum of coefficients in the formula from Proposition 4 over all λ with $\lambda_1, \lambda_1' \leq n/2$ with the value $\frac{3}{4}(1 - \frac{1}{n}) + \frac{1}{4}(2p-1)(1 + \frac{3}{n-1})$. In the case that $p = \frac{1}{2}$ this value is $\frac{3}{4}(1 - \frac{1}{n-1})$. This gives a heuristic for why the eigenvalue of $[n/2, n/2]$ will be computed when $p = \frac{1}{2}$ as roughly $(3/4)^{n/2}$.

The bound on ψ_λ is by induction. Since $p \geq \frac{1}{2}$ the only possibly negative terms in the expansion

$$\psi_\lambda = \sum_{\rho:\lambda/\rho=[2]} \psi_\rho \frac{d_\rho}{d_\lambda} + (2p-1) \sum_{\rho:\lambda/\rho=[1,1]} \psi_\rho \frac{d_\rho}{d_\lambda} + p \sum_{\rho:\lambda/\rho=[1]\cup[1]} \psi_\rho \frac{d_\rho}{d_\lambda}$$

are the ψ_ρ. An upper bound on $|\psi_\rho|$ suffices to pull it in front of the expression.

The first n with partition λ with $\lambda_1' \leq \lambda_1 < \frac{n}{2}$ existing is $n = 8$, with a single partition $\lambda = [3, 3, 2]$. $\psi_{[3,3,2]} \leq \psi_{[4,4]}$ as for each s, $0 \leq \frac{\chi_{[3,3,2]}(2^s)}{d_{[3,3,2]}} \leq \frac{\chi_{[4,4]}(2^s)}{d_{[4,4]}}$. This can be seen in the character table of S_8 [Littlewood(2006)]. Then assume by induction the bound holds for ρ a partition of $n - 2$ with $\rho_1' \leq \rho_1 < \frac{n-4}{2}$, and it was proven before for $\rho_1 \leq \frac{n-2}{2}$ as well. The ρ's that appear in the expression have $\rho_1 \leq \lambda_1 \leq \frac{n-2}{2}$, so $\psi_\rho \leq \psi_{[(n-1)/2,(n-1)/2]}$. Finally, $\left(\frac{3}{4}(1 - \frac{1}{n-1}) + (2p-1)\frac{1}{4}(1 + \frac{3}{n-1}) \right) \psi_{[(n-1)/2,(n-1)/2} \leq \psi_{[n/2,n/2]}$ since

$$\psi_{[n/2,n/2]} = \psi_{[n/2,n/2-2]} \frac{d_{[n/2,n/2-2]}}{d_{[n/2,n/2]}} + \psi_{[n/2-1,n/2-1]}(2p-1)\frac{d_{[n/2-1,n/2-1]}}{d_{[n/2,n/2]}}$$

And, since $\frac{d_{[n/2,n/2-2]}}{d_{[n/2,n/2]}} + (2p-1)\frac{d_{[n/2-1,n/2-1]}}{d_{[n/2,n/2]}} = \frac{3}{4}(1 - \frac{1}{n-1}) + (2p-1)\frac{1}{4}(1 + \frac{3}{n-1})$ and $\psi_{[n/2,n/2-2]} \geq \psi_{[n/2-1,n/2-1]}$ by Lemma 6. Therefore, $\psi_\lambda \leq \psi_{[n/2,n/2]}$. $\square$

These monotonicity relations on the eigenvalues, which will be used to form a upper bound on the mixing time of the walk, also show a quite granular structure inside of the mixing. After many steps, the walk will approach being uniform on all permutations, but some permutations will always be more likely than others. The likelihood order for the walk is a total order that describes the relative likelihood of the permutations. For the $p \geq \frac{1}{2}$ the transposition walk and involution walk, the likelihood order after sufficient time is the cycle lexicographic order on permutations defined as:

Definition 9. Let $\alpha = (1^{a_1}, 2^{a_2}, ..., n^{a_n}), \beta = (1^{b_1}, 2^{b_2}, ..., n^{b_n})$ be two conjugacy classes of S_n. Define $\alpha >_{CL} \beta$ when for $\min_k(a_k \neq b_k) = i$, $a_i > b_i$.

Corollary 10. *The likelihood order for the involution walk for $p \geq \frac{1}{2}$ and for t sufficiently large is the cycle lexicographical order.*

Proof. The difference in likelihood of two permutations α and β can be studied through the discrete Fourier transform. For the involution walk at two permutations α and β,

$$P^{*t}(\alpha) - P^{*t}(\beta) = \frac{1}{n!} \sum_\lambda \left(\chi_\lambda(\alpha) - \chi_\lambda(\beta) \right) d_\lambda (\psi_\lambda)^t$$

The trivial representation has eigenvalue and coefficient one in the discrete Fourier decomposition for both α and β and so vanishes. Other partitions for which $\chi_\lambda(\alpha) = \chi_\lambda(\beta)$ will also not contribute to this quantity. After sufficient time, the terms for the partitions with largest eigenvalue in magnitude with $\chi_\lambda(\alpha) \neq \chi_\lambda(\beta)$ will be exponentially larger than any other terms and hence will will determine the sign of $P^{*t}(\alpha) - P^{*t}(\beta)$. In lazy walks the largest eigenvalue in magnitude almost always occurs for a single partition.

From [Bernstein(2015)], a partition is called an i-cycle detector if $\lambda_2 + \lambda_1' - 2 \geq i$ and $\lambda_1 + \lambda_2' - 2 \geq i$. If λ is not an i-cycle detector and the smallest cycle differing in the cycle decomposition of α and β is an i-cycle, then $\chi_\lambda(\alpha) - \chi_\lambda(\beta) = 0$ [Bernstein(2015)]. Therefore, one must only examine the i-cycle detecting partitions for each value of i from 1 to $n/2$ in order to find the eventual likelihood order. By Lemma 6, Lemma 7, and Theorem 8, the partition $[n-i, i]$ has the largest magnitude of eigenvalue of all i-cycle detecting partitions. More over, when α and β first differ at an i-cycle, $\chi_{[n-i,i]}(\alpha) - \chi_{[n-i,i]}(\beta) = a_i - b_i \neq 0$ [Bernstein(2015)]. In this case, the term for $[n-i, i]$ in the discrete Fourier transform, $(\chi_{[n-i,i]}(\alpha) - \chi_{[n-i,i]}(\beta)) d_\lambda \psi_{[n-i,i]}^t$, determines the sign of $P^{*t}(\alpha) - P^{*t}(\beta)$ for sufficiently large t. Since $\chi_{[n-i,i]}(\alpha) - \chi_{[n-i,i]}(\beta) = a_i - b_i$ in this case and $\psi_{[n-i,i]} > 0$, the permutation with more i-cycles is more likely after sufficient time. This is the cycle lexicographic order from Definition 9. $\quad\square\quad\square$

4 Upper Bound on Mixing

This section will be working towards bounds on ψ_λ to use in the upper bound formula,

$$\left\| K^{*t}(\sigma) - \frac{1}{n!} \right\|_{TV}^2 \leq \frac{1}{4} \sum_{\lambda \neq 1} d_\lambda^2 \psi_\lambda^{2t}$$

Recall that,

$$\psi_\lambda = \sum_{s=0}^{n/2} p^{n/2-s}(1-p)^s \binom{n/2}{s} \frac{\chi_\lambda(1^{n-2s}, 2^s)}{d_\lambda}$$

Instead of bounding $\chi_\lambda(1^{n-2s}, 2^s)$ for each s individually, the character polynomial will give an expression for the character as a polynomial in $n - 2s$ and s. The character polynomial, $q_\rho(x_1, ..., x_k)$ for the partition ρ of k is a polynomial in variables $x_1, ..., x_k$ so that

$$\chi_{n-k,\rho}(1^{x_1}, ..., k^{x_k}, ..., n^{x_n}) = q_\rho(x_1, ..., x_k)$$

for any conjugacy class $(1^{x_1}, ..., n^{x_n})$ of S_n. Garsia and Goupil [Garsia and Goupil(2009)] give a formula for the character polynomial akin to the Murnaghan-Nakayama rule run backwards from its traditional order, peeling off border strips of the largest cycles first.

$$q_\rho(x_1, ..., x_i, 0, ..., 0) = \sum_j \binom{x_i}{j} \sum_{P=(\rho^0, ..., \rho^j)} (-1)^{ht(P)} q_{\rho^j}(x_1, ..., x_{i-1}, 0, ..., 0),$$

where P ranges over all possible ways of removing border strips of size i from ρ so that a Ferrers diagram remains at each step, as in Murnaghan-Nakayama. The formula says, choose j i-cycles of the x_i i-cycles and attempt to peel them off from below the first row of λ, and take the remaining $x_i - j$ i-cycles from the first row of λ. Recurse on the remaining shape with the next largest cycle size. In Murnaghan-Nakayama, the first row does not receive this special treatment. Letting $i = 2$ gives the character polynomial for an involution as:

$$q_\rho(n - 2s, s, 0, ...0) = \sum_j \binom{s}{j} \sum_{P=(\rho^0, ..., \rho^j)} (-1)^{ht(P)} q_{\rho^j}(n - 2s, 0, ..., 0),$$

where the last term can be expanded as

$$q_{\rho^j}(n - 2s) = d_{n-2s-|\rho^j|, \rho^j} = \binom{n - 2s}{|\rho^j|} d_{\rho^j} \prod_{k=1}^{\rho_1^j} \frac{n - 2s - |\rho^j| - k + 1}{n - 2s - |\rho^j| - k + (\rho^j)_k' + 1}$$

Then an upper bound on q_{ρ^j} that is more computationally tractable comes from ignoring the sign associated with the insertions, rounding $\frac{n-2s-|\rho^j|-k-i+1}{n-2s-|\rho^j|-k-i+(\rho^j)_i'+1}$ to 1, and upper bounding the ways of inserting one and two cycles by the dimension of ρ giving:

$$q_\rho(n - 2s, s) \leq \sum_j \binom{s}{j} \sum_{\rho^0, ..., \rho^j} \binom{n - 2s}{|\rho| - 2j} d_{\rho^j} \leq \sum_j \binom{s}{j} \binom{n - 2s}{|\rho| - 2j} d_\rho$$

Then using this in ψ_λ and splitting s into j_1 and j_2 gives:

$$\psi_\lambda \leq \sum_{s=0}^{n/2} p^{n/2-s} (1-p)^s \binom{n/2}{s} \sum_j \binom{s}{j} \binom{n - 2s}{n - \lambda_1 - 2j} \frac{d_{\lambda/\lambda_1}}{d_\lambda} \tag{2}$$

$$= \frac{d_{\lambda/\lambda_1}}{d_\lambda} p^{n/2} \sum_{j_1, j_2} \left(\frac{1-p}{p} \right)^{j_1+j_2} \binom{n/2}{j_1, j_2} \binom{n - 2j_1 - 2j_2}{n - \lambda_1 - 2j_2}, \tag{3}$$

which says to approximate ψ take the expectation over the binomial distribution over all ways to choose j_1 and j_2 of the $n/2$ 2-cycles to insert into the first row and the remaining partition and to split the remaining unused numbers into either the first row or the remaining partition. The d_{λ/λ_1} factor takes into account that there may be may ways to arrange things in the lower part of the partition. When $\lambda_1 \geq n/2$, the maximum value of $\frac{d_{\lambda/\lambda_1}}{d_\lambda}$ occurs at the partition $[n-i, i]$ where this a very good approximation.

Proposition 11. *For* $1 \leq i \leq \frac{n}{2}$,

$$\max_{\lambda:\lambda_1=n-i} \frac{d_{\lambda/\lambda_1}}{d_\lambda} = \frac{d_{[i]}}{d_{[n-i,i]}} = \binom{n}{i}^{-1} \frac{n-i+1}{n-2i+1}$$

Proof. Using the hook length formula [Stanley(1999)],

$$\frac{d_{\lambda/\lambda_1}}{d_\lambda} = \frac{\prod_{k=1}^{\lambda_1} \lambda_1 - k + \lambda_k'}{(n)_{\lambda_1}} = \binom{n}{\lambda_1} \prod_{k=1}^{\lambda_1} \frac{\lambda_1 - k + \lambda_k'}{\lambda_1 - k + 1}$$

The λ_k' are decreasing, and the product is maximized if these are taken to be as even as possible. So for $1 \leq k \leq i$, $\lambda_k' = 1$, for $k > i$, $\lambda_k' = 0$. This is the partition $[n-i, i]$.   □   □

The bound used above on the character polynomial, $\sum_{P=\{\rho_0,...,\rho_j\}}(-1)^P \leq d_\rho$, was sufficiently strong for the partitions with first row at least $n/2$, but not for those with smaller first row. However, by Proposition 8, the eigenvalues for λ with $\lambda_1 < n/2$ are bounded by the eigenvalue for $[n/2, n/2]$.

The next step is to handle the sum (4). Instead of counting how the the two cycles $(1,2), ..., (n-1, n)$ and unchosen cycles used as fixed points are arranged separately, an easier approach exists. Consider instead, splitting the numbers $1, 2, ..., n$ into two parts. When $2i - 1$ and $2i$ are in the same part, this could have happened using them as a single two cycle, or separately as fixed points, for a total weight under the binomial distribution of $\frac{1-p}{p} + 1 = \frac{1}{p}$. And when $2i - 1$ and $2i$ are not in the same part, this could only have happened from 1-cycle insertion but two different ways, for a weight of 2.

Proposition 12.

$$\sum_{j_1,j_2} \left(\frac{1-p}{p}\right)^{j_1+j_2} \binom{n/2}{j_1, j_2} \binom{n-2j_1-2j_2}{n-\lambda_1-2j_1-2j_2} = \sum_j 2^j \frac{1}{p^{n/2-j}} \binom{n/2}{j, \frac{n-i-j}{2}, \frac{i-j}{2}}$$

Note that j must be such that $i - j, n - i - j$ are both even. So,

Proposition 13. *For* $\lambda_1 \geq \frac{n}{2}$ *let* $i = n - \lambda_1$, *for* $\lambda_1 \leq \frac{n}{2}$ *let* $i = \frac{n}{2}$

$$\psi_\lambda \leq \binom{n}{i}^{-1} \frac{n-i+1}{n-2i+1} \sum_{j \leq i, i-j \, even} (2p)^j \binom{n/2}{j, \frac{n-i-j}{2}, \frac{i-j}{2}}$$

Now to approximate the sum, one can use that it is less than $i/2$ times its largest term, except for i small where the largest term is the last and the other terms will be exponentially smaller. Note that nothing is assumed about p in this bound. The remaining proofs are a series of technical estimates and bounds. For the proofs see [Bernstein(2015)].

Proposition 14. *When*

$$\alpha = \sqrt{1 + \frac{1-p^2}{p^2}\frac{4(n-i)(i)}{n^2}}$$

$$\sum_j (2p)^j \binom{n/2}{j,\ \frac{n-i-j}{2},\ \frac{i-j}{2}} \leq \binom{n/2}{i/2}\left(1 - \frac{n-i}{n}\frac{2}{1+\alpha}\right)^{-\frac{i}{2}}\left(1 - \frac{i}{n}\frac{2}{1+\alpha}\right)^{-\frac{n-i}{2}}$$

$$\frac{i}{2}\left(1_{i=n/2}\frac{1}{\sqrt{2\pi i}} + 1_{i\neq n/2}\sqrt{\frac{n-2i+1}{n-i+1}}\right)$$

This gives for λ with $\lambda_1 = n - i > n/2$:

$$\psi_\lambda \leq \frac{\binom{n/2}{i/2}}{\binom{n}{i}}\left(\frac{n-i+1}{n-2i+1}\sqrt{\frac{n-i}{n-2i}}\frac{i+1}{2}\frac{e^2}{2^{3/2}}\right)\left(1 - \frac{n-i}{n}\frac{2}{1+\alpha}\right)^{-\frac{i+1}{2}}\left(1 - \frac{i}{n}\frac{2}{1+\alpha}\right)^{-\frac{n-i}{2}}$$

And for $\lambda_1 \leq n/2$,

$$\psi_\lambda \leq \frac{\binom{n/2}{n/4}}{\binom{n}{n/2}}\left((n/2+1)^2/2\sqrt{n/2}\frac{e^2}{2^{3/2}}\right)\left(1 - \frac{1}{2}\frac{2}{1+\alpha}\right)^{-\frac{n}{2}}$$

The next proposition brings things together into one expression:

Proposition 15. *For λ with $\lambda_1 = n - i$, $i < n/2$:*

$$\psi_\lambda \leq e^{-i\log(\frac{2}{1+p})+\log\left(\frac{e^2(i+1)}{2^{5/2}}\left(\frac{n-i}{n-2i}\right)^{3/2}\right)}$$

For $\lambda_1 \leq \frac{n}{2}$,

$$\psi_\lambda \leq e^{-\frac{n}{2}\log\frac{2}{1+p}+\log(\frac{n^{3/2}(n+2)e^2}{8})}$$

Proposition 16. *For $i \leq p\sqrt{n-2i+2}$,*

$$\sum_{j\leq i,\ i-j\ even} (2p)^j \binom{n/2}{j,\ \frac{n-i-j}{2},\ \frac{i-j}{2}} \leq \frac{1}{1 - \frac{i(i-1)}{2p^2(n-2i+2)}}(2p)^i\binom{n/2}{i}$$

Proposition 17.

$$\psi_{1^n} = (2p-1)^{n/2}$$

Proof. By Murnaghan-Nakayama, $\chi_{1^n}(1^{n-2s}, 2^s) = (-1)^s$ since all the 2-cycles insert vertically and these are exactly insertions covering an even number of rows.

$$\psi_{1^n} = \sum_{s=0}^{n/2} p^{n/2-s}(1-p)^s\binom{n/2}{s}\frac{\chi_{1^n}(1^{n-2s}, 2^s)}{d_{1^n}} = \sum_{s=0}^{n/2}\binom{n/2}{s}p^{n/2-s}(1-p)^s(-1)^s = (2p-1)^{n/2}$$

$\square$ $\square$

And with this bound on ψ_λ, the upper bound lemma is at hand.

Theorem 18. *For* $t = \log_{\frac{2}{1+p}}(n) + \frac{c}{\log(\frac{2}{1+p})}$, n *such that* $\frac{10 \log(n+2)}{\sqrt{(n+2)/2-1}} \leq \log\left(\frac{2}{1+p}\right)$ *and* $n - 1 > \sqrt{n/2}(1 + \log(n))$, *then*

$$||P^{*t} - U||_{TV} \leq e^{-c/2}$$

5 Lower Bound on Mixing

The representation slowest to vanish for this walk is $[n-1, 1]$, so its character gives a random variable where $P^{*t}(\cdot)$ and $\pi(\cdot)$ differ significantly. Using a lower bound formula similar to Chebychev's inequality after calculating the first and second moments of this character will give a lower bound on mixing of $\log_{\frac{1}{p}}(n)$.

Proposition 19. *[Levin et al.(2009)Levin, Peres, and Wilmer] For* γ, ν *two probability distributions on* Ω, *and* f *a real valued function on* Ω, *if*

$$|E_\gamma(f) - E_\nu(f)| \geq r\sigma$$

where $\sigma^2 = [Var_\gamma(f) + Var_\nu(f)]/2$, *then*

$$||\gamma - \nu||_{TV} \geq 1 - \frac{4}{4 + r^2}$$

In this case, $\nu = U$ is the stationary distribution of the walk, uniform over all permutations. As seen in [Diaconis(1988)],

$$E_U\left(\chi_{[n-1,1]}\right) = 0, Var_U(\chi_{n-1,1}) = 1$$

These follow for any non-trivial characters by basic tenets of representation theory. For the first, by orthogonality of characters, $\sum_{g \in G} \chi_\lambda(g) = 0$. For the second, $\sum_{g \in G} \chi_\lambda g^2 = |G|$.

Proposition 20.

$$E_{P^{*t}}\chi_{[n]} = 1$$

$$E_{P^{*t}}\left(\chi_{[n-1,1]}\right) = (n-1)\left(p - (1-p)\frac{1}{n-1}\right)^t$$

$$E_{P^{*t}}\chi_{[n-2,2]} = \frac{n(n-3)}{2}\left(p^2 - \frac{(1-p)^2}{n-3}\right)^t$$

$$E_{P^{*t}}\chi_{[n-2,1,1]} = \frac{(n-1)(n-2)}{2}\left(p^2 - \frac{1-p^2}{n-1} - \frac{2}{(n-1)(n-2)}\right)^t$$

Proof. For an irreducible representation λ, since P is a class function, by Schur's Lemma, the Fourier transform of P is a constant ψ_λ times the identity matrix.

$$\hat{P}(\lambda) = \psi_\lambda I_{d_\lambda}$$

Moreover, $\hat{P^{*t}}(\lambda) = \left(\hat{P}(\lambda)\right)^t$. This leads to the following formula for the expected value of a character over the walk:

$$\mathbb{E}_{P^{*t}}(\chi_\lambda) = \sum P^{*t}(g)\operatorname{tr}(\lambda(g)) = \operatorname{tr}(\sum P^{*t}(g)\lambda(g)) = \operatorname{tr}\hat{P^{*t}}(\lambda) = d_\lambda \psi_\lambda^t$$

The method of choice to compute the expectation for $\chi_{[n-1,1]}$ will be to directly compute ψ_λ. Recall,

$$\psi_\lambda = \sum_{s=0}^{n/2} p^{n/2-s}(1-p)^s \binom{n/2}{s} \frac{\chi_\lambda(1^{n-2s}, 2^s)}{d_\lambda}$$

Further, the character polynomials of the representation gives that:

$$\chi_{[n]}(1^{n-2s}, 2^s) = 1, \chi_{[n-1,1]}(1^{n-2s}, 2^s) = n - 2s - 1$$

$$\chi_{[n-2,2]}(1^{n-2s}, 2^s) = \binom{n-2s}{2} - (n-2s) + s$$

$$\chi_{[n-2,1,1]}(1^{n-2s}, 2^s) = \binom{n-2s}{2} - (n-2s) - s$$

Combining these formulas and completing the calculations gives the above formulas. □

□

This gives:

$$\sigma^2 = \frac{1}{2}\left(1 + 1 + (n-1)\left(p - (1-p)\frac{1}{n-1}\right)^t + \frac{n(n-3)}{2}\left(p^2 - \frac{1}{n-3}(1-p)^2\right)^t \right. \tag{4}$$

$$\left. + \frac{(n-1)(n-2)}{2}\left(p^2 - \frac{1-p^2}{n-1} - \frac{2}{(n-1)(n-2)}\right)^t - (n-1)^2\left(p - (1-p)\frac{1}{n-1}\right)^{2t}\right) \tag{5}$$

After a series of estimates, for details see [Bernstein(2015)], it follows that:

Theorem 21. *For $p \geq \frac{1}{2}$, $t < \frac{\log(n)-c}{\log(1/p)}$, there exists $1 - \frac{\log(n)}{n} \leq A \leq 1$ so that,*

$$\|P^{*t} - U\|_{TV} \geq 1 - \frac{1}{\frac{1}{2} + A^2 e^{2c}}$$

For $p \geq \frac{1}{2}$ this gives a lower bound for mixing of $\frac{\log(n)-c}{\log(1/p)}$ which is off by just over a factor of two from the upper bound of $\frac{\log(n)+c}{\log(2/(1+p))}$. When p is small, less than $\frac{1}{n-\sqrt{n}}$, $\psi_{[n-1,1]} = p - \frac{1-p}{n-1}$ is no longer the largest eigenvalue in magnitude as

$$\lim_{n\to\infty} |\psi_{[1^n]}| = \lim_{n\to\infty}\left(1 - \frac{2}{n}\right)^{n/2} = \frac{1}{e}$$

This cross over happens around $\frac{1}{p} = W(e^n) \approx n - \log(n) + o(1)$ where W is the product log function, also known as the Lambert W-function, as:

$$\log((1 - 2p)^{n/2} = n/2 \log(1 - 2p) \approx pn$$

$$\log\left(p - \frac{1-p}{n-1}\right) = \log(p) + \log\left(1 - \frac{1-p}{pn}\right) \approx \log(p) - \frac{1-p}{pn}$$

For $p \approx \frac{1}{n}$ that last term contributes at most a constant, leaving the equation $pn \approx \log(p)$ with solution $p = W(e^n)$.

Proposition 22. *For the involution walk, t even, $||P^{*t} - U||_{TV} \geq 1 - \frac{2}{2 + (2p-1)^{tn/2}}$. When $t \leq \frac{1}{np}$, total variation distance at least $\frac{1}{1+2e}$.*

References

[Bernstein(2015)] M. Bernstein. *Random Walks on the Symmetric Group, Likelihood Orders, and Involutions*. PhD thesis, Stanford University, 2015.

[Diaconis(1988)] P. Diaconis. *Group representations in probability and statistics*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11. Institute of Mathematical Statistics, Hayward, CA, 1988.

[Diaconis and Shahshahani(1981)] P. Diaconis and M. Shahshahani. Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete*, 57(2):159–179, 1981.

[Garsia and Goupil(2009)] A. M. Garsia and A. Goupil. Character polynomials, their q-analogs and the Kronecker product. *Electron. J. Combin.*, 16(2, Special volume in honor of Anders Bjorner): Research Paper 19, 40, 2009.

[Hayden and Preskill(2007)] P. Hayden and J. Preskill. Black holes as mirrors: quantum information in random subsystems. *JHEP*, 2007.

[Levin et al.(2009)Levin, Peres, and Wilmer] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.

[Littlewood(2006)] D. E. Littlewood. *The theory of group characters and matrix representations of groups*. AMS Chelsea Publishing, Providence, RI, 2006. Reprint of the second (1950) edition.

[Lulov(1996)] N. Lulov. *Random Walks on the Symmetric Group Generated by Conjugacy Classes*. PhD thesis, Harvard University, 1996.

[Sekino and Susskind(2008)] Y. Sekino and L. Susskind. Fast scramblers. *JHEP*, 2008.

[Stanley(1999)] R. P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.