

Likelihood orders for the p -cycle walks on the symmetric group

Megan Bernstein*

School of Mathematics
Georgia Institute of Technology
Georgia, U.S.A.

bernstein@math.gatech.edu

Submitted: Oct 3, 2017; Accepted: Jan 8, 2018; Published: Feb 16, 2018

Mathematics Subject Classifications: 05E10, 20C30, 60J10

Abstract

Consider for a random walk on a group, the order from most to least likely element of the walk at each step, called the likelihood order. Up to periodicity issues, this order stabilizes after a sufficient number of steps. Here discrete Fourier analysis and the representations of the symmetric group, particularly formulas for the characters, are used to find the order after sufficient time for the random walks on the symmetric group generated by p -cycles for any p fixed, n sufficiently large. For the transposition walk, generated by all the 2-cycles, at various levels of laziness, it is shown that order n^2 steps suffice for the order to stabilize. Likelihood orders can aid in finding the total variation or separation distance mixing times.

Keywords: random walks; symmetric groups; character polynomial; likelihood order

1 Introduction

A random walk on the symmetric group S_n is determined by a probability distribution P for its generators. Starting at time zero at the identity, at each step a permutation is chosen according to P and appended to the left of the current state. The distribution of the t th step of the walk is given by the t th convolution power of P :

$$P^{*t}(g) = \sum_{h \in S_n} P(gh^{-1})P^{t-1}(h)$$

This is an example of Markov chain with transition matrix $K(h, g) = P(gh^{-1})$.

*Partially supported by NSF Grant DMS-1344199.

A short argument from unpublished work of Diaconis and Isaacs [6] using Cauchy-Schwartz shows that at even times the identity element of a group is the most likely element of any symmetric random walk. Assume for all $g \in G$, $P(g) = P(g^{-1})$ then:

$$P^{2t}(g) = \sum_{h \in G} P^{*t}(gh^{-1})P^{*t}(h) \leq \sum_{h \in G} (P^{*t}(h))^2 = \sum_{h \in G} P^{*t}(h^{-1})P^{*t}(h) = P^{*2t}(e)$$

Using induction, Diaconis and Isaacs go on to explore several walks, including on the cyclic group and hypercube, with in which they can describe a total order that describes the most likely to least likely element of the walk at every step. We call such an order at a given step of the walk a likelihood order, and will explore what can be said about them for random walks on the symmetric group.

Two motivating notions of distance between a random walk on a group after t steps and its uniform stationary distribution, $\pi(\cdot) = \frac{1}{|G|}$, are separation distance and total variation distance defined as (see e.g. [12]):

$$\text{sep}(t) = \max_{g \in G} 1 - |G|P^{*t}(g)$$

$$\|P^{*t} - \pi\|_{TV} = \sum_{g \in G: P^{*t}(g) > \pi(g)} P^{*t}(g) - \pi(g).$$

The separation distance is attained at the least likely element. The related l^∞ distance is attained at either the most or least likely element. For total variation, useful bounds, especially lower bounds, frequently originate from understanding the likelihood of the elements relative to the uniform distribution. It is then natural and of interest to know the most and least likely elements of a random walk as well as the likelihood order.

In his thesis, Lulov [14] found the likelihood order after sufficient time for the transposition walk on the symmetric group at even steps using a monotonicity relation on the characters of the symmetric group evaluated at transpositions utilized by Diaconis and Shahshahani [6]. Here we extend his work on the transposition walk by providing a sufficient time for the likelihood order to stabilize, and exploring how the laziness of the walk effects the likelihood order. Other walks in which likelihood orders are known include: a random walk on the symmetric group generated by involutions with the same likelihood order after sufficient time as the transposition walk [2], and random walks on Coxeter groups with P uniform on the Coxeter generators where weak Bruhat order is always a likelihood order [21].

Diaconis and Shahshahani [7] showed how the representation theory of the symmetric group can be used to diagonalize such walks. For a special class of distributions P called class functions that are constant on each conjugacy class, the theory is especially nice. For class functions, the discrete Fourier transform [5] gives a formula for the distribution of the t th step of the walk, P^{*t} in terms of the characters and dimensions of the irreducible representations of the symmetric group.

The key to the analysis in this paper is the determination of of the largest eigenvalue with non-zero coefficient in the Fourier inversion formula expansion of P^{*t} , as presented

in Proposition 7. This is accomplished through classification of representations as i -cycle detectors defined in section 3. This will yield for these walks the set of representations that determine the likelihood order. Fine control on the terms in the Fourier inversion formula for the transposition walk lead to an upper bound on when the likelihood order stabilizes.

The likelihood orders in this paper are all variants of a cycle lexicographic order. The following definitions are in terms of the cycle structures of two conjugacy classes α and β of the symmetric group, where α has a_i i -cycles.

Definition 1. [Cycle lexicographic order] Let $\alpha = (1^{a_1}, 2^{a_2}, \dots, n^{a_n}), \beta = (1^{b_1}, 2^{b_2}, \dots, n^{b_n})$ define $\alpha >_{CL} \beta$ when for $\min_k(a_k \neq b_k) = i, a_i > b_i$. $\alpha =_{CL} \beta$ exactly when $\alpha = \beta$.

Definition 2. [Alternating cycle lexicographic order] Let $\alpha = (1^{a_1}, 2^{a_2}, \dots, n^{a_n}), \beta = (1^{b_1}, 2^{b_2}, \dots, n^{b_n})$ define $\alpha >_{(-1)^{i+1}CL} \beta$ when for $\min_k(a_k \neq b_k) = i, (-1)^{i+1}a_i > (-1)^{i+1}b_i$. $\alpha =_{(-1)^{i+1}CL} \beta$ exactly when $\alpha = \beta$.

For instance, the identity is larger than a transposition, respectively $\alpha = (1^n)$ and $\beta = (1^{n-2}, 2)$ in both orders, since they both first differ at 1-cycles which is odd. While for $\alpha = (1^{n-4}, 2^2)$ two 2-cycles, $\beta = (1^{n-4}, 4^1)$, $\alpha >_{CL} \beta$ but $\beta \geq_{(-1)^{i+1}CL} \alpha$ since they first differ at a 2-cycle and $-2 < 0$.

As the level of laziness, $P(e)$, varies for the transposition walk, the importance of parity and cycle structure vary. For instance, when $P(e) = 0$ and so only transpositions are generators, at even (and respectively, odd) steps of the walk, only even (and respectively, odd) permutations can be made. When $P(e) \geq \frac{1}{2}$, parity plays no role. In between these, there is a cycle size under which differing at that size cycles trumps parity, and over which parity determines likelihood order. This will be explored in section 5.4.

Definition 3. [1-zippered cycle lexicographic order] Let $\alpha = (1^{a_1}, 2^{a_2}, \dots, n^{a_n}), \beta = (1^{b_1}, 2^{b_2}, \dots, n^{b_n})$ define $\alpha >_{1CL} \beta$ when for $\min_k(a_k \neq b_k) = i$, if $i = 1$ or $\text{sg}(\alpha) = \text{sg}(\beta)$ the same parity, $a_i > b_i$. Otherwise, $\alpha >_{1CL} \beta$ if $(-1)^t \text{sg}(\alpha) > (-1)^t \text{sg}(\beta)$. Finally, $\alpha =_{1CL} \beta$ exactly when $\alpha = \beta$.

Theorem 4. *There exists a constant C such that for $t \geq Cn^2$, the transposition walk*

1. *with $P(e) = 0$ has the likelihood order order the cycle lexicographic order restricted to even (odd) elements at even (odd) times*
2. *with $P(e) = \frac{1}{n}$ has likelihood order the 1-zippered cycle lexicographic order*
3. *with $P(e) \geq \frac{1}{2}$, bounded away from 1, has likelihood order the cycle lexicographic order*

It is suspected that for the transposition walk at even steps the n -cycles are always the least likely elements. It is trivially true up to $n - 2$ steps, and Theorem 4 shows it holds after order n^2 steps. However, the likelihood order, as a whole, breaks frequently for $t < n - 2$. Whether the n -cycles are the least likely and the likelihood order holds in this gap are open.

For p -cycle walks for fixed p and sufficiently large n , the asymptotic likelihood order is found via refining an estimate from Wasserman's thesis [20], reprinted in [8].

Theorem 5. *For t sufficiently large and $p > 1$ odd, the likelihood order for the p -cycle walk is the $(-1)^{i+1}$ cycle lexicographic order.*

For t sufficiently large and $p > 2$ even, the likelihood order for the p -cycle walk breaks into two parts. There exists a C , such that for $i \leq Cn^{\frac{p-3}{p-2}}$ for the likelihood order is the $(-1)^{i+1}$ cycle lexicographic order, for $i \geq Cn^{\frac{p-3}{p-2}}$ the order is cycle lexicographic.

As a result, the least likely element of these walks after sufficient time is a conjugacy class made up of 2-cycles and possibly another 3, 4 or 5 cycle according to parity of n and t .

Corollary 6. *For $p > 2$ fixed, n and t sufficiently large, the least likely element of the p -cycle walk at time t is:*

$$\begin{cases} (2^{n/2}) & n \text{ even}, (p-1)t \text{ even} \\ (2^{n/2-2}, 4) & n \text{ even}, (p-1)t \text{ odd} \\ (2^{(n-3)/2}, 3) & n \text{ odd}, \frac{n+1}{2} \text{ same parity as } (p-1)t \\ (2^{(n-5)/2}, 5) & n \text{ odd}, \frac{n+3}{2} \text{ same parity as } (p-1)t \end{cases}$$

Section 2 outlines the techniques, discrete Fourier analysis and character theory for the symmetric group. Section 3 establishes the i -cycle detectors as the representations that determine the likelihood orders. Following that, in section 4 discusses the cycle lexicographical orders that will appear as likelihood orders and compares them with other partial orders on partitions. Then in section 5 we find sufficient time for the likelihood order for the transposition walk to stabilize. Following this, two brief digressions, section 5.4 discusses how the likelihood order varies with degrees of laziness, while section 5.5 gives method for finding the states more likely than the stationary distribution. Finally, in section 6 careful bounds on the eigenvalues of the p -cycle walk lead to their likelihood orders.

2 Background

2.1 Notation

The letters λ, ρ, γ will always refer to partitions. The letters α, β, κ to conjugacy classes of the symmetric group on n letters denoted S_n , with with cycle structure written in one of two partition notations $\alpha = [\alpha_1, \dots, \alpha_r] = (1^{\alpha_1}, \dots, n^{\alpha_n})$ where α is a permutation with cycles lengths $\alpha_1 \geq \dots \geq \alpha_r$ and a_i i -cycles. The following partitions occur repeatedly and will be denoted by $[n-i, i] = \lambda^i, [n-i, i-k, 1, \dots, 1] = \lambda^{i,k}$, and $[n-i, 1, \dots, 1] = \rho^i = \lambda^{i, i-1}$.

Since these random walks are generated by conjugacy classes the probability function is a class function. This means that probabilities are constant on each conjugacy class. Formulas will be written in terms of conjugacy classes referring to probability of an individual element of the conjugacy class.

2.2 Discrete Fourier inversion formula

The Fourier inversion formula gives an expression for the distribution of a random walk on the symmetric group in terms of characters of irreducible representations of the symmetric group. These irreducibles are indexed by partitions of n . See chapter 2 of [5] for a more thorough treatment.

Proposition 7 (Fourier inversion formula). *For a random walk on the symmetric group starting at the identity with distribution at time one, $P(\cdot)$, a class function, the probability distribution for t^{th} step is given by*

$$P^{*t}(\alpha) = \frac{1}{n!} \sum_{\lambda} \chi_{\lambda}(\alpha) d_{\lambda} (r_{\lambda})^t$$

Where r is as follows. The sum below is over conjugacy classes κ of size $|\kappa|$,

$$r_{\lambda} = \sum_{\kappa} |\kappa| P(\kappa) \frac{\chi_{\lambda}(\kappa)}{d_{\lambda}}$$

Proof. Starting from the Fourier inversion theorem on page 13 of [5] and using $G = S_n$, for any function f on S_n ,

$$f(g) = \frac{1}{n!} \sum_{\lambda} d_{\lambda} \text{Tr}(\lambda(g^{-1}) \hat{f}(\rho_i))$$

where the Fourier transform, $\hat{f}(\lambda) = \sum_g f(g) \lambda(g)$. We will simplify this for the case $f(g) = P^{*t}(g)$ with P is a class function. For class functions, Shur's lemma implies that $\hat{P}(\lambda) = r_{\lambda} I_{d_{\lambda}}$ where $I_{d_{\lambda}}$ is the d_{λ} dimensional identity matrix and r_{λ} is as in the statement of the theorem. Moreover, the Fourier transform turns convolution into multiplication, so $\hat{P}^{*t}(\lambda) = \left(\hat{P}(\lambda)\right)^t$. Finally, by definition $\text{Tr}(\lambda(g^{-1})) = \chi_{\lambda}(g^{-1})$ using that for the symmetric group g and g^{-1} are in the same conjugacy class and the character is constant on conjugacy class, this is equal to $\chi_{\lambda}(g)$. \square

In the event that P is constant and supported on a single conjugacy class, κ , the eigenvalue $r_{\lambda} = \frac{\chi_{\lambda}(\kappa)}{d_{\lambda}}$ is the character ratio denoted $r_{\lambda}(\kappa)$.

This immediately yields a formula for the difference in probability of two permutation α, β after t steps of the walk.

Proposition 8.

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

Where r_{λ} is defined in Proposition 7.

Note that as $1 \geq r_{\lambda} \geq -1$, the terms are decay exponentially in t , and for t sufficiently large the sign of the equation is determined by the largest r_{λ} for which $\chi_{\lambda}(\alpha) \neq \chi_{\lambda}(\beta)$. For times sufficiently large so that term is larger than the sum of all the terms, the likelihood order stabilizes.

2.3 The Murnaghan-Nakayama formula

The irreducible characters for the symmetric group are computable through the recursive Murnaghan-Nakayama formula arising from the correspondence between representation theory of S_n and symmetric functions where it gives the decomposition of the Shur polynomials into power sum polynomials (see e.g. I.7 Ex 5 of [15]).

Proposition 9 (Murnaghan-Nakayama).

$$\chi_\lambda(\alpha) = \sum_S (-1)^{ht(S)}$$

summed over all sequences of partition $S = (\lambda^{(0)}, \lambda^{(1)}, \dots, \lambda^{(r)})$ such that $r = l(\alpha)$, $0 = \lambda^{(0)} \subset \lambda^{(1)} \subset \dots \subset \lambda^{(r)}$ such that each $\lambda^{(i)} - \lambda^{(i-1)}$ is a border strip of length α_i and $ht(S) = \sum_i ht(\lambda^{(i)} - \lambda^{(i-1)})$.

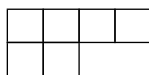
One way of viewing this process is successively removing border strips of length $\alpha_r, \dots, \alpha_1$ in all possible ways from the bottom, right of λ . Alternatively, one can envision this process as in Littlewood, [13], as successive insertions of $\alpha_1, \dots, \alpha_r$ into the top, left of λ . This reversal of the usual visualization was key in defining an i -cycle detector as it emphasizes the importance of the large pieces. The following are borrowed from Littlewood with some change in terminology to match modern usage.

Definition 10. The insertion of i nodes to a partition is called a valid insertion of i nodes if the nodes are added to any row until they are exhausted or until the number of nodes in this row exceeds the number in the preceding row by one, the nodes being then added to the preceding row according to the same rule, and so on until the i nodes are exhausted provided the final product is a valid partition. If the number of rows involved is even it is called a negative application, if odd, a positive application.

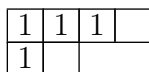
Proposition 11 (Littlewood). *If λ is a partition of n and α denotes a conjugacy class of the symmetric group with cycles of orders $\alpha_1, \dots, \alpha_r$ the $\chi_\lambda(\alpha)$ is obtained from the number of methods of building the partition λ by consecutive valid insertions of $\alpha_1, \dots, \alpha_r$ nodes by subtracting the number of ways which contain an odd number of negative applications from the number of ways which contain an even number of negative applications.*

An insertion of α_i nodes from a cycle length α_i will be shortened to an insertion of an α_i cycle.

Example 12. To see this, consider calculating $\chi_{[4,2]}(1^2, 4)$. First to choose which row to start inserting the largest cycle, the four cycle, into the shape $[4, 2]$:



Either the the first row or the second row are possible giving (the nodes are enumerated by the i of the α_i that fills it):



1	1	1	1

The first is a negative insertion since it covers an even number of rows, while the second confined to the first row is a positive insertion. It remains to place the two 1-cycles. In the first, the first 1-cycle can go into either the first or second row to be valid, and the second 1-cycle must go in the remaining spot.

1	1	1	2
1	3		

1	1	1	3
1	2		

While our second way of placing the 4-cycle leaves only the second row for each 1-cycle insertion.

1	1	1	1
2	3		

So this sums to two ways with an odd number of negative insertions and one way with an even number of negative insertions. This gives $\chi_{[4,2]}(1^2, 4) = -2 + 1 = -1$

2.4 Character Polynomials

Another useful tool for insight into the characters is the character polynomial. For more information, see [10], which forms the basis of this subsection. For each character χ_λ the character polynomial is the unique polynomial that evaluated on the cycle structure of a permutation, evaluates to the character. For example, the most well known character polynomial is the one for the irreducible representation indexed by $[n - 1, 1]$, that its character is $\chi_{[n-1,1]}(\alpha) = a_1 - 1$, so that the number of fixed points completely determines this character.

Definition 13. The character polynomial of μ a partition of n is

$$q_\mu(x_1, \dots, x_n) = \downarrow \left(\sum_{\alpha \vdash n} \frac{\chi_\mu(\alpha)}{z_\alpha} \prod_{i=1}^n (ix_i - 1)^{a_i} \right)$$

where $z_\alpha = \prod_i a_i! i^{a_i}$ and $\downarrow (x_1^{a_1} \cdots x_n^{a_n}) = (x_1)_{a_1} \cdots (x_n)_{a_n}$.

Proposition 14.

$$\chi_\lambda(\alpha) = q_{[\lambda_2, \dots, \lambda_r]}(a_1, \dots, a_{n-\lambda_1})$$

Example 15. To continue using $\chi_{[4,2]}(1^2, 4)$ as an example, find the character polynomial corresponding to $\chi_{[4,2]}(\alpha)$. This corresponds to removing the first row of the partition, so we compute $q_{[2]}$.

$$q_{[2]}(x_1, x_2) = \downarrow \left(\sum_{\alpha \vdash 2} \frac{\chi_{[2]}(\alpha)}{z_\alpha} \prod_{i=1}^2 (ix_i - 1)^{a_i} \right)$$

Since $\chi_{[2]}(\cdot) = 1$ and the only partitions of 2 are [2] and [1] with $z_{[2]} = (1!)(2), z_{[1,1]} = (2!)(1^2) = 2$, this gives,

$$q_{[2]}(x_1, x_2) = \downarrow \left(\frac{1}{2}(2x_2 - 1) + \frac{1}{2}(x_1 - 1)^2 \right) = x_2 - \frac{1}{2} + \frac{1}{2}(\downarrow x_1^2) - x_1 + \frac{1}{2} = x_2 + \binom{x_1}{2} - x_1$$

Applying this to $(1^2, 4)$ gives $\chi_{[4,2]}(1^2, 4) = 0 + \binom{2}{2} - 2 = -1$.

3 Detecting Cycle Structure

As motivated by the formula for the difference in probabilities, the goal here is to describe, for fixed conjugacy classes α, β , partitions, λ , for which we know $\chi_\lambda(\alpha) - \chi_\lambda(\beta) = 0$. Each the character for each partition has a granularity to detect cycle structure up to a size beyond which it is indiscriminate. For example, above it was noted that $\chi_{[n-1,1]}$ is determined by fixed points and it was computed that $\chi_{[n-2,2]}$ is given in terms of fixed points and 2-cycles. This leads to three equivalent conditions motivated by both Murnaghan-Nakayama and character polynomials characterizing a partition with the potential to be effected by i -cycles, to be called an i -cycle detector. In turn, a partition that is not an i -cycle detector will not be able to detect if α, β only differ in cycle decompositions for cycles $\geq i$. The definitions below reflect the property that $\chi_\lambda(\alpha) = \text{sgn}(\alpha)\chi_{\lambda'}(\alpha)$, so being an i -cycle detector is a dual statement about a partition and its conjugate.

Definition 16. An insertion of cycles lengths $\alpha_1, \dots, \alpha_k$ into λ is trivial if they insert with nodes in the same order as from inserting a cycle length $\alpha_1 + \dots + \alpha_k$.

Example 17. Recall the examples of insertions above.

$$\begin{array}{|c|c|c|c|} \hline 1 & 1 & 1 & 2 \\ \hline 1 & 3 & & \\ \hline \end{array} \quad \begin{array}{|c|c|c|c|} \hline 1 & 1 & 1 & 3 \\ \hline 1 & 2 & & \\ \hline \end{array} \quad \begin{array}{|c|c|c|c|} \hline 1 & 1 & 1 & 1 \\ \hline 2 & 3 & & \\ \hline \end{array}$$

As always, the first cycle, α_1 inserts trivially in all three examples. Now, looking at the first two cycles insertions, only the first example is trivial since in this case the second cycle was inserted following the first cycle in the same row. The second example fails to be trivial is there is no way to insert one cycle length $\alpha_1 + \alpha_2 = 5$ into the shape α_1, α_2 fill, [3, 2]. And the third fails since the first cycle does not fill the entire first column. The three cycles insert trivially in none of these examples.

Definition 18. Call λ a i -cycle detector if there is a non-trivial insertion of cycles lengths $\geq i$ into λ and λ'

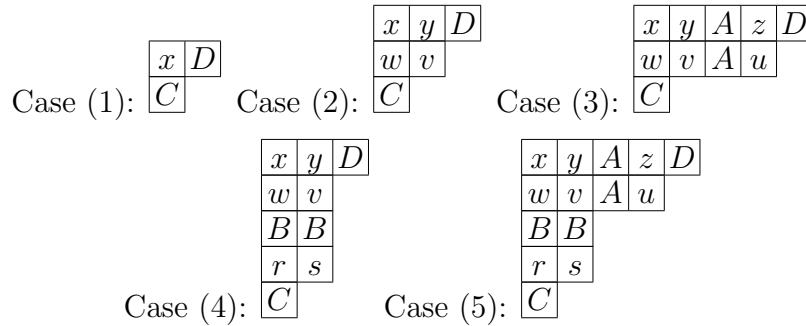
Implicit in this definition is that i -cycle detectors only exist for $i \leq \frac{n}{2}$, since it is impossible to insert two cycles size $> \frac{n}{2}$ into a partition of n .

Lemma 19. *The following are equivalent:*

1. λ is an i -cycle detector

2. $h_{2,1}, h_{1,2} \geq i$ (where $h_{x,y}$ is the length of the hook starting at (x, y) in λ)
3. some x_j for $j \geq i$ occurs in a monomial with non-zero coefficient in both $q_{[\lambda_2, \dots, \lambda_r]}(x)$, $q_{[\lambda'_2, \dots, \lambda'_r]}(x)$

Proof. The equivalence of the first two statements will mostly be a proof by diagram. The first two rows and columns of λ take on one of five shapes where the capital letters stand for any number of boxes, and the lower case a single box. These capital letters and $h_{1,2}, h_{2,1}$ will be used abusively to stand for both the boxes they represent and the number of boxes they represent.



First to show if $h_{2,1}, h_{1,2} \geq i$, then two cycles can be inserted into the first two rows and columns of λ with the second inserting non-trivially. Then it will also clearly work for λ' as it has the same property.

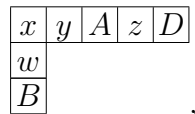
In case (1), $D = h_{1,2} \geq i$, $C = h_{2,1} \geq i$ so inserting the first i -cycle into the first row and the second into C is possible.

For case (2), $D + 2 = h_{1,2} \geq i$, $C + 2 = h_{2,1} \geq i$, so inserting the first i -cycle into the first row and the second into $h_{2,1}$ is possible.

For case (3), $h_{1,2} = 3 + A + D \geq i$, $h_{2,1} = 3 + A + C \geq i$, so the first i -cycle fits in the first row and the second in the $h_{2,1}$.

In case (4), $h_{1,2} = 3 + B + D \geq i$, $h_{2,1} = 3 + B + C \geq i$. The first column then is the same length as $h_{2,1}$, so insert the first i -cycle vertically. And the second can insert non-trivially into $h_{1,2}$.

Finally, for case (5), insert trivially into the first row and column along



which is at least i long since this is the same length as $h_{1,2}$. Insert non-trivially into the remaining squares, which are the same length as $h_{2,1}$.

If two cycles of lengths at least i can be inserted with one non-trivial into both λ, λ' , it needs to be shown that $h_{1,2}, h_{2,1} \geq i$. These same five shapes also describe the shape these insertions can fill within λ . For one shape only, it can only be shown that inserting into λ gives $h_{2,1} \geq i$. The insertion into λ' will then satisfy $h_{1,2} \geq i$. Let the insertion into λ be one of the shapes:

For case (1), since the second insertion is assumed to be non-trivial, it must be contained in C . Therefore, $h_{2,1} \geq C \geq i$. To establish that $h_{1,2} \geq i$ use that λ' satisfies the non-trivial requirement as well.

Then, for case (2), the non-trivial intersection cannot be in both C and D since this would leave only one square for the trivial insertion. This means the non-trivial insertion is contained in either $h_{2,1}$ or $h_{1,2}$, leaving the size of the other for the trivial insertion. So both $h_{2,1}$ and $h_{1,2}$ are at least i .

In case (3), if the non-trivial insertion intersects either w, C or z, D it cannot intersect the other since this would make that insertion longer than the preceding trivial insertion. This leaves either at most $D + A + 3$ boxes or $C + A + 3$ boxes for the trivial insertion, each contained in $h_{1,2}$ and $h_{2,1}$. The remaining space then for a single insertion is at most $C + A + 3$ boxes and $D + A + 3$ boxes. So both hooks are at least i . If instead the non-trivial insertion did not intersect w, C, z, D , it must fit into v, A, u which is smaller than either hook, and both hooks are also at least i in length.

Then, in case (4), as the conjugate of the previous shape the proof follows exactly the same argument with relabeling.

Lastly, in case (5), the non-trivial intersection must contain the hook s, B, v, A, u . It may additionally contain either z, D or r, C but not both, as that would make it longer than the trivial insertion. First if it contains r, C . This makes the non-trivial in $C + 4 + B + A$ boxes, which is less than the length of $h_{2,1}$. This leaves $D + 4 + B + A$ boxes the trivial fills, less than the length of $h_{2,1}$. If the non-trivial fills z, D , then similarly it cannot also touch r, C , and must be in $D + 4 + B + A$ boxes, less than the length of $h_{1,2}$. Again, leaving $C + 4 + A + B \leq h_{2,1}$ for the trivial insertion. In the final case where the trivial is just s, B, v, A, u , $i \leq A + B + 3$ this is less than $h_{2,1} \geq A + B + D + 4$ and $h_{1,2} \geq A + B + C + 4$.

The equivalence of the latter two statements in the theorem follows from expanding the character polynomial.

$$q_{\lambda_2, \dots, \lambda_r}(x) = \downarrow \left(\sum_{\alpha \vdash n - \lambda_1} \frac{\chi_\mu(\alpha)}{z_\alpha} \prod_{i=1}^{n - \lambda_1} (ix_i - 1)^{a_i} \right)$$

The sum is over $\alpha \vdash n - \lambda_1$ but when $\chi_\mu(\alpha) = 0$ the α term is 0. By the Murnaghan-Nakayama rule, the largest cycle inserts first, and the largest cycle that can insert into $[\lambda_2, \dots, \lambda_r]$ is $h_{2,1}$. The sum can then be restricted to α with parts of size at most $h_{2,1}$. An x_i occurs in the α term only when α has a part of size i . So no x_i term can then occur in $q_{[\lambda_2, \dots, \lambda_r]}$ for $i > h_{2,1}$. Next, to show $x_{h_{2,1}}$ occurs with non-zero coefficient. Taking the sum over α with $\alpha_1 = h_{2,1}$,

$$\downarrow \left(\sum_{\alpha \vdash [n - \lambda_1]: \alpha_1 = h_{2,1}, i \neq 1, \alpha_i < h_{2,1}} \frac{\chi_{[\lambda_2, \dots, \lambda_r]}(\alpha)}{z_\alpha} \prod_{i=1}^{n - \lambda_1} (a_i x_i - 1)^{a_i} \right) \quad (1)$$

$$= \downarrow \left(\sum_{\beta \vdash n - \lambda_1 - h_{2,1}: \beta_i < h_{2,1}} \frac{(-1)^{r+1} \chi_{[\lambda_3 - 1, \dots, \lambda_r - 1]}(\beta)}{h_{2,1} z_\beta} (h_{2,1} x_{h_{2,1}} - 1) \prod_{i=1}^{h_{2,1} - 1} (ix_i - 1)^{b_i} \right) \quad (2)$$

$$= (-1)^{r+1} \left(x_{h_{2,1}} - \frac{1}{h_{2,1}} \right) \downarrow \left(\sum_{\beta \vdash n - \lambda_1 - h_{2,1}} \frac{\chi_{[\lambda_3-1, \dots, \lambda_r-1]}(\beta)}{z_\beta} \prod_{i=1}^{n-\lambda_1-h_{2,1}} (ix_i - 1)^{b_i} \right) \quad (3)$$

$$= (-1)^{r+1} \left(x_{h_{2,1}} - \frac{1}{h_{2,1}} \right) q_{[\lambda_3-1, \dots, \lambda_r-1]} \quad (4)$$

No character polynomial can be zero since no character of an irreducible is zero, so the $x_{h_{2,1}}$ will have non-zero coefficient in $q_{[\lambda_2, \dots, \lambda_r]}$. Similarly the $x_{h_{1,2}}$ term will have non-zero coefficient in $q_{[\lambda'_2, \dots, \lambda'_r]}$ □

Given a partition, the hook starting at $(2, 1)$ will be called its subhook, and its length, $h_{2,1}$ the subhook length of the partition. When $h_{2,1} \leq h_{1,2}$, the partition is a i -cycle detector if its subhook length is at least i . The proof above also shows:

Corollary 20. *For a partition λ ,*

- $h_{2,1}$ is the largest i for which x_i occurs in $q_{[\lambda_2, \dots, \lambda_r]}(x)$
- $h_{1,2}$ is the largest i for which x_i occurs in $q_{[\lambda'_2, \dots, \lambda'_r]}(x)$

Theorem 21. *If α, β with the same sign have $a_j = b_j$ for all $j < i$ and λ is not a i -cycle detector, then $\chi_\lambda(\alpha) - \chi_\lambda(\beta) = 0$.*

Proof. If λ is not an i -cycle detector, then one of $h_{2,1}, h_{1,2} < i$ by lemma 19. By corollary 20, this means one of $q_{[\lambda_2, \dots, \lambda_r]}(x), q_{[\lambda'_2, \dots, \lambda'_r]}(x)$ has no x_j terms for $j \geq i$. If it is the character polynomial for λ , then

$$\chi_\lambda(\alpha) = q_{[\lambda_2, \dots, \lambda_r]}(a_1, \dots, a_{i-1}) = q_{[\lambda_2, \dots, \lambda_r]}(b_1, \dots, b_{i-1}) = \chi_\lambda(\beta)$$

If it is the character polynomial for λ' ,

$$\begin{aligned} \chi_\lambda(\alpha) &= \text{sgn}(\alpha) \chi_{\lambda'}(\alpha) = \text{sgn}(\alpha) q_{[\lambda'_2, \dots, \lambda'_r]}(a_1, \dots, a_{i-1}) \\ &= \text{sgn}(\beta) q_{[\lambda'_2, \dots, \lambda'_r]}(b_1, \dots, b_{i-1}) = \text{sgn}(\beta) \chi_{\lambda'}(\beta) = \chi_\lambda(\beta) \end{aligned} \quad \square$$

The stronger statement that any i -cycle detector can differentiate between two conjugacy classes that first differ in their number of i -cycles is not true. For instance, the character polynomial associated with $[n - 3, 2, 1]$ has no degree two terms despite being a 2-cycle detector. This means, for example, $\chi_{[n-3,2,1]}(1^{n-4}, 2^2) = \chi_{[n-3,2,1]}(1^{n-4}, 4)$.

4 The Cycle Lexicographic Orders and Other Orders on Partitions

Two total orders on partitions will form the underpinnings for the likelihood orders for the walks generated by p -cycles. The first is cycle lexicographic order denoted \leq_{CL} :

Definition 22. [Cycle lexicographic order] Let

$$\alpha = (1^{a_1}, 2^{a_2}, \dots, n^{a_n}), \quad \beta = (1^{b_1}, 2^{b_2}, \dots, n^{b_n})$$

define $\alpha >_{CL} \beta$ when for $\min_k(a_k \neq b_k) = i$, $a_i > b_i$. $\alpha =_{CL} \beta$ exactly when $\alpha = \beta$.

Throughout the paper, i will be used to mean this first differing cycle size for any pair α, β . CL is distinct from the traditional orders on partitions: majorization/ domination/ natural, reverse lexicographical, and (an usual order) Lulov's lexicographical. Where lexicographical without the cycle prefix here refers to the α_i 's rather than the a_i in the notation $\alpha = [\alpha_1, \dots, \alpha_r] = (1^{a_1}, \dots, n^{a_n})$.

Definition 23. Majorization order is defined as $\alpha \succeq \beta$ if for all i , $\sum_{j \leq i} \alpha_j \geq \sum_{j \leq i} \beta_j$. Equivalently, $\alpha \succeq \beta$ if boxes in the Ferrers diagram of β can be moved up and to the right to get the Ferrers diagram of α .

Definition 24. In reverse lexicographical order $\alpha \succeq_{RL} \beta$ if for the minimum i that $\alpha_i \neq \beta_i$, $\alpha_i > \beta_i$. [19]

This is a refinement of majorization into a total order.

In Lulov's thesis [14] this next definition is called reverse lexicographic, which clashes with the canonical definition used above, [19], [15]. It flips the order of the λ_i . Lulov mistakenly equated this order to CL, which he established as stabilized likelihood order for the transposition walk at even times.

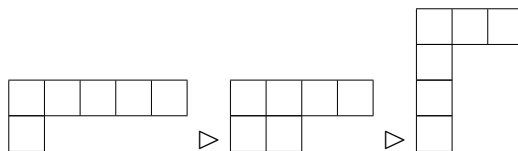
Definition 25. In Lulov's lexicographical order $\alpha \succeq_L \beta$ if the the maximum i such that $\alpha_i \neq \beta_i$, $\alpha_i < \beta_i$.

Proposition 26. *CL order is not a linear extension of majorization order (and automatically also incompatible with reverse lexicographical). Cycle lexicographical is also distinct from Lulov's lexicographical.*

Proof. The order under majorization and cycle lexicographical of the following partitions of 6 are incompatible: $[5, 1]$, $[4, 2]$, $[3, 1, 1, 1]$. In majorization order,

$$[5, 1] \succeq [4, 2] \succeq [3, 1, 1, 1]$$

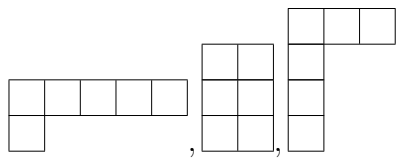
as one can see how to move boxes down and to the left to get the next shape



while under cycle lexicographic order the order is neither the same or reversed. In each case, $i = 1$ and the partitions are ordered by number of fixed points:

$$[3, 1, 1, 1] \succeq_{CL} [5, 1] \succeq_{CL} [4, 2]$$

The order under Lulov's lexicographical and cycle lexicographical of the following partitions of 6 are incompatible: $[5, 1]$, $[2, 2, 2]$, $[3, 1, 1, 1]$



In Lulov's lexicographic order these are ordered by number of parts,

$$[5, 1] \geq_L [2, 2, 2] \geq_L [3, 1, 1, 1]$$

while in cycle lexicographic the order of the three is again determined by number of fixed points:

$$[3, 1, 1, 1] \geq_{CL} [5, 1] \geq_{CL} [2, 2, 2] \quad \square$$

It is also the case that when CL order is taken on the conjugate of α and majorization or Lulov's lexicographical is taken on α , there is still incompatibility. For the first take $[5, 1]$, $[4, 2]$, $[4, 1, 1]$ and the second again $[5, 1]$, $[2, 2, 2]$, $[3, 1, 1, 1]$.

The variant of cycle lexicographical that arises in this paper as likelihood orders is as follows.

Definition 27. Define $\alpha >_{(-1)^{i+1}CL} \beta$ when for $\min_k(a_k \neq b_k) = i$, for i even $a_i < b_i$ or for i odd $a_i > b_i$. $\alpha =_{(-1)^{i+1}CL} \beta$ exactly when $\alpha = \beta$.

The largest and smallest elements under these orders will be the most and least likely elements of walks. For reference later, based on the divisibility of n and restricted to permutations with odd sign, $|_-$ or even sign, $|_+$, the largest, $\hat{1}$, and smallest, $\hat{0}$, elements of each of these orders are:

Proposition 28. When n is odd:

	CL	$(-1)^{i+1}CL$
$\hat{1}$	(1^n)	(1^n)
$\hat{1}_+$	1^n	(1^n)
$\hat{1}_-$	$(1^{n-2}, 2)$	$(1^{n-2}, 2)$
$\hat{0}$	(n)	$(2^{\frac{n-3}{2}}, 3)$
$\hat{0}_+$	(n)	if $4 (n+1)$, $(2^{\frac{n-3}{2}}, 3)$ else, $(2^{\frac{n-5}{2}}, 5)$
$\hat{0}_-$	$(\frac{n-1}{2}, \frac{n-1}{2})$	if $4 (n+1)$, $(2^{\frac{n-5}{2}}, 5)$ else, $(2^{\frac{n-3}{2}}, 3)$

When n is even:

	CL	$(-1)^{i+1}CL$
$\hat{1}$	(1^n)	(1^n)
$\hat{1}_+$	(1^n)	(1^n)
$\hat{1}_-$	$(1^{n-2}, 2^1)$	$(1^{n-2}, 2)$
$\hat{0}$	(n)	$(2^{\frac{n}{2}})$
$\hat{0}_+$	$(\frac{n}{2})$	if $4 (n)$, $(2^{\frac{n}{2}})$ else, $(2^{\frac{n-4}{2}}, 4)$
$\hat{0}_-$	(n)	if $4 (n)$, $(2^{\frac{n-4}{2}}, 4)$ else, $(2^{\frac{n}{2}})$

Proof. In cycle lexicographic order, the elements are first ordered by number of fixed points, then number of 2-cycles, etc. The largest element in the order is the identity followed by the transposition, while the smallest elements, those with only the largest possible cycles, are the n -cycles and then next largest, depending on parity of n , the permutations with two $n/2$ cycles or $(n+1)/2$ and $(n-1)/2$ cycle.

Alternating cycle lexicographic order can be thought of as rewarding odd cycles and punishing even cycles. Again, fixed points are the first deciding point, ordering by decreasing number of fixed points. Then 2-cycles are the next deciding factor, ordering by increasing number of 2-cycles. And so on, alternating by cycle parity. The largest element in the order is the identity, since it will have the most fixed points. The next largest is the only conjugacy class with $n-2$ fixed points, the transpositions.

When n is even, $\alpha = (2^{\frac{n}{2}})$ is last in this order since for $\beta \neq \alpha$, $\min_k(a_k \neq b_k)$ is either 1 if β has a fixed point or 2 if not. If $i = 1$, $a_1 = 0 < 1 \leq b_1$, and 1 is odd so β is larger. If $i = 2$, $a_2 = \frac{n}{2} > b_2$, but 2 is even so having more 2-cycles makes a conjugacy class less likely under this order. The next smallest element when n is even is $(2^{\frac{n-4}{2}}, 4)$ since it is also fixed point free and has the next most 2 cycles. One of these two is odd while the other is even, and so they are the respective smallest odd and even elements.

When n is odd, the smallest element is the fixed point free conjugacy class with the most 2 cycles, $(2^{\frac{n-3}{2}}, 3)$, so again in comparison with any other conjugacy class i is 1 or 2. $(2^{\frac{n-5}{2}}, 5)$ has the next most 2-cycles amongst fixed point free conjugacy classes. Exactly one of these will be odd, the other even. \square

5 Transposition Walk

Consider building a permutation by at each step appending a randomly selected transposition on the left. The goal is to find which permutations are more or less likely than others after many steps of the walk. The answer for the transposition walk, is that that the likelihood order after sufficient time is given by cycle lexicographic order, confined to even (odd) permutations at even (odd) times. The key is finding, given a pair of permutations, the partition(s) indexing the largest character ratio(s) with non-zero character difference in the decomposition given by Proposition 8. For the transposition walk, when two permutations first differ at an i -cycle the partitions are $[n-i, i] = \lambda^i$ and $(\lambda^i)'$. While this is enough for the likelihood order to asymptotically be cycle lexicographic, we will show $O(n^2)$ steps suffices. In the author's thesis[4], an explicit non-optimal constant 2.82 was found in Theorem 44, but the result has been simplified for readability. The methodology is extended in section 5.4 to the lazy version of the walk, with a discussion of how the likelihood order changes as the laziness increases. While it is clear that finding the least likely element is useful in establishing separation distance, the likelihood order is also of potential use in finding the total variation distance. In section 5.5 we find where in the order the uniform distribution would fall, with those permutations below it being the optimal bad set.

Showing cycle lexicographic order is the likelihood order after sufficient time is the best that can be hoped for as for all $n \geq 8$, the does not hold as the likelihood order

for some $t < n - 1$. For example when n is divisible by 4, the conjugacy class of an $(n - 1)$ -cycle is more likely after sufficient time than the conjugacy class of $n/2$ 2-cycles. Yet, as it requires $n - 2$ transpositions to achieve the former and only $n/2$ to achieve the latter, for $n/2 \leq t < n - 2$ the order breaks. All known cases, from simulation, of the order breaking are of this form where the order rights itself when the eventually more likely element is first possible, thus there is not a known case of the order breaking for $t \geq n - 1$.

5.1 Character Ratio Maximizing Cycle Detector

From Proposition 8, the formula for the difference in probability of two conjugacy classes α, β is:

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

After sufficient time, the sign of this expression will be determined by the partitions λ with the largest magnitude of character ratio, $\left| \frac{\chi_{\lambda}(\tau)}{d_{\lambda}} \right|$, and non-zero character difference $\chi_{\lambda}(\alpha) - \chi_{\lambda}(\beta)$. It happens that when $\min_k(a_k \neq b_k) = i$ this lead position is taken by $\lambda^i = [n - i, i], (\lambda^i)'$. The relative sizes of character ratios at a transposition are well understood see e.g. Lemma 10 of [7]:

Proposition 29. *When $\lambda \triangleright \rho$, $\frac{\chi_{\lambda}(\tau)}{d_{\lambda}} \geq \frac{\chi_{\rho}(\tau)}{d_{\rho}}$. Recall, $\lambda \triangleright \rho$ when boxes in the Ferrers diagram of ρ can be moved up and to the right to get the Ferrers diagram of λ . The difference in character ratios is the distance the boxes travel divided by $\binom{n}{2}$.*

Recall that an i cycle detector must have $h_{2,1}, h_{1,2} \geq i$. The goal is to find the i -cycle detector with largest positive character ratio. For this walk, its conjugate will have the largest in magnitude character ratio among the i -cycle detectors with negative character ratios at a transposition.

Proposition 30. *$\lambda^i = [n - i, i]$ is larger in majorization order than any other i -cycle detector, and so it and its conjugate are the i -cycle detectors with largest character ratios.*

Proof. In order for $h_{2,1}$ to be $\geq i$, at least i blocks must exist in the subhook. The positioning of these most up and to the right is as in $[n - i, i]$. Moreover, for any partition with at least i blocks in its subhook, these i blocks can be moved up and to the right to be in the second row and all other blocks below the first row moved to the first row yielding $[n - i, i]$. If any additional block is moved up and to the right from $[n - i, i]$, it must be moved out of the subhook making the partition no longer an i -cycle detector. To be incompatible to $[n - i, i]$ in majorization order, a partition must have both more blocks in the first row and some blocks in the third row. If it had more than $n - i$ blocks in the first row this would preclude it from being an i -cycle detector. Thus, all i -cycle detectors are comparable to and below $[n - i, i]$ in majorization order. \square

5.2 Cycle Lexicographical Order

The sign of 5.1 after sufficient time is determined by the signs of the terms for λ^i and $(\lambda^i)'$. The λ^i term will end up controlling the sign, while its conjugate contributes to the condition that only even partitions can be reached at even times and odd transpositions at odd times.

Proposition 31. *Let α, β be two conjugacy classes and $\alpha \geq_{CL} \beta$ with $\min_k(a_k \neq b_k) = i$, then*

$$\chi_{\lambda^i}(\alpha) - \chi_{\lambda^i}(\beta) = a_i - b_i$$

Proof. From equation (4), the characteristic polynomial has a single monomial containing x_i , namely x_i , and no terms x_j for any $j > i$. \square

And quickly some computations of the characters and dimensions of these partitions,

Proposition 32.

$$d_{\lambda^i} = \binom{n}{i} \frac{n - 2i + 1}{n - i + 1}$$

$$\frac{\chi_{\lambda^i}(\tau)}{d_{\lambda^i}} = 1 - \frac{i(n - i + 1)}{\binom{n}{2}} > 0$$

Proof. By the hook length formula [19],

$$d_{\lambda^i} = \frac{n!}{i!(n - i + 1) \cdots (n - 2i + 2)(n - 2i) \cdots 1} = \frac{n!}{i!(n - i)!} \frac{n - 2i + 1}{n - i + 1}$$

Using the formula for the difference in character ratios in Proposition 29, the difference between the character ratios of λ^0 (with character ratio 1) and λ^i is the distance in the Ferrers diagram the squares travel up and to the left from λ^i to become λ^0 , all divided by $\binom{n}{2}$. In this case, the i blocks need to travel from the second row to the first row, for an average distance traveled of $n - i + 1$. Hence,

$$\frac{\chi_{\lambda^i}(\tau)}{d_{\lambda^i}} = \frac{\chi_{\lambda^0}(\tau)}{d_{\lambda^0}} - \frac{i(n - i + 1)}{\binom{n}{2}} = 1 - \frac{i(n - i + 1)}{\binom{n}{2}} \quad \square$$

Proposition 33. *The sign of the sum of the λ^i and $(\lambda^i)'$ terms in 5.1 is positive when t is the same sign as α, β and $\alpha \geq_{CL} \beta$ with $\min_k(a_k \neq b_k) = i$. Thus, after sufficient time, CL is the likelihood order.*

Proof.

$$\begin{aligned} & (\chi_{\lambda^i}(\alpha) - \chi_{\lambda^i}(\beta)) d_{\lambda^i} \left(\frac{\chi_{\lambda^i}(\tau)}{d_{\lambda^i}} \right)^t + (\chi_{(\lambda^i)' }(\alpha) - \chi_{(\lambda^i)' }(\beta)) d_{(\lambda^i)' } \left(\frac{\chi_{(\lambda^i)' }(\tau)}{d_{(\lambda^i)' }} \right)^t \\ &= (a_i - b_i) d_{\lambda^i} \left(\frac{\chi_{\lambda^i}(\tau)}{d_{\lambda^i}} \right)^t + \text{sgn}(\alpha)(a_i - b_i) d_{\lambda^i} \left(-\frac{\chi_{\lambda^i}(\tau)}{d_{\lambda^i}} \right)^t \\ &= (a_i - b_i)(1 + \text{sgn}(\alpha)(-1)^t) d_{\lambda^i} \left(\frac{\chi_{\lambda^i}(\tau)}{d_{\lambda^i}} \right)^t \end{aligned}$$

Which is positive when $a_i > b_i$ and α, β are the same parity as t . \square

Proposition 34. *After sufficient time the following hold. The identity is the most likely element at even times, a transposition the most likely at odd times. When n is odd, an n -cycle is the least likely element at even times and an element of $(\frac{n-1}{2}, \frac{n+1}{2})$ at odd times. When n is even, at even times an element of $(\frac{n}{2})$ is least likely, while at odd times an n -cycle is least likely.*

Proof. These are the largest and smallest elements of CL order under the required parity constraints, as in Proposition 28. \square

5.3 Proof of Theorem 4 (1)

The general methodology will be to show in the Fourier inversion theorem from Proposition 7, for two conjugacy classes α, β first differing at an i -cycle, that the λ^i term in is greater than the sum of the absolute value of all other terms for partitions λ with $i \leq h_{2,1} \leq h_{1,2}$. The symmetry between conjugates handles the remaining partitions. The λ^i term will be shown to be at least a scaling factor, c_j , times greater than all the λ terms with $\lambda_1 = n - j$, where the sum over all the scaling factors will be at most 1. This breaks down into showing:

$$c_j |\chi_{\lambda^i}(\alpha) - \chi_{\lambda^i}(\beta)| d_{\lambda^i} r_{\lambda^i}(\tau)^t \geq \left(\max_{\lambda: \lambda_1 = n-j} |\chi_{\lambda}(\alpha) - \chi_{\lambda}(\beta)| \right) \left(\sum_{\lambda: \lambda_1 = n-j} d_{\lambda} \right) \left(\max_{\lambda: \lambda_1 = n-j} r_{\lambda}(\tau) \right)$$

This will be accomplished through finding good bounds on the difference in character ratios $\delta_{i,j} = r_{\lambda^i}(\tau) - (\max_{\lambda: \lambda_1 = n-j} r_{\lambda}(\tau))$ and writing (using that $1+x \geq e^{x/2}$ for $0 \leq x \leq 1$):

$$\left(\frac{r_{\lambda^i}(\tau)}{r_{\lambda}(\tau)} \right)^t \leq \left(1 + \frac{\delta_{i,j}}{r_{\lambda}(\tau)} \right)^t \leq (1 + \delta_{i,j})^t \leq e^{\frac{\delta_{i,j}}{2} t}$$

Using t of order n^2 , it suffices to show that $\log(c_j)$, $\log(\max_{\lambda: \lambda_1 = n-j} |\chi_{\lambda}(\alpha) - \chi_{\lambda}(\beta)|)$ and $\log\left(\frac{\sum_{\lambda: \lambda_1 = n-j} d_{\lambda}}{d_{\lambda^i}}\right)$ are at most order $n^2 \delta_{i,j}$.

Two sets of partitions have to be done as special cases, as their character ratio differences with λ^i can be as small as $\frac{4}{n(n-1)}$, while all other character ratio differences will be at least $\frac{1}{n}$. These special cases are the partitions $[n-i, i-k, 1^k]$ and for $i \geq n/3$, $[n-j, j]$. For the remaining partitions, a single bound on the character differences and sum of dimensions will suffice, though several cases of character ratio differences are needed. We start with the first special case of the partitions $[n-i, i-k, 1^k]$, then move onto the general case where the second special case will be handled as a separate clause in Propositions 36,37,38, respectively giving bounds for the character differences, dimension ration, and character ratio differences. The scaling factor will be $c_j = 2^{i-j-2}$ for j other than these special cases, and $\frac{1}{2}$ for the $[n-i, i-k, 1^k]$ and $\frac{1}{4}$ for the $[n-j, j]$ when $i \geq n/3$.

A heuristic for when the likelihood stabilizes is:

$$t \geq \max_{\lambda} \left(\frac{\chi_{\lambda^i}(\tau)}{d_{\lambda^i}} - \frac{\chi_{\lambda}(\tau)}{d_{\lambda}} \right)^{-1} \log(d_{\lambda}/d_{\lambda^i})$$

This is analogous to the heuristic estimate $\hat{\tau}_n$ of the mixing time of a random walk with sufficient symmetry on a group based on its spectral gap, $1 - \lambda_2$ (where $1 = \lambda_1, \lambda_2, \dots$ are the eigenvalues), and multiplicity of the second eigenvalue, δ_n , [1].

$$\hat{\tau}_n = \tau_e(n) \log(\delta_n) \text{ where } \tau_e(n) = -1/\log(\lambda_2)$$

In these cases, the closest character will have a larger dimension than λ^i by a factor between a $O(i)$ and n . While the minimal character ratio difference, playing the part of the spectral gap, varies inversely at the same time between $O(\frac{O(i)}{n^2})$ and $O(\frac{1}{n})$. This analog of the heuristic describes all the bounds achieved for by the author in her thesis for the transposition walk, three cycle walk, and n -cycle walk [4]. Note that there exists cases where the spectral gap heuristic is off by a constant, such as the random-to-random shuffle [3].

Proposition 35. *The $\lambda^{i,k}$ term $(\chi_{\lambda^{i,k}}(\alpha) - \chi_{\lambda^{i,k}}(\beta)) d_{\lambda^{i,k}r_{\lambda^{i,k}}}(\tau)^t$ is:*

$$(-1)^k (a_i - b_i) \binom{n}{i} \binom{i-1}{k} \frac{n-2i+k+1}{n-i+k+1} \left(1 - \frac{i(n-i+k+1)}{\binom{n}{2}} \right)^t$$

The sum of all the $\lambda^{i,k}$ terms is

$$(a_i - b_i) \binom{n}{i} \sum_{k=1}^{i-1} (-1)^k \binom{i-1}{k} \frac{n-2i+k+1}{n-i+k+1} \left(1 - \frac{i(n-i+k+1)}{\binom{n}{2}} \right)^t.$$

This is less than half of the λ^i term for $t \geq \frac{n^2(\log(i-1)+1)}{i}$.

Proof. From the Ferrers diagram of λ^i , to get to the Ferrers diagram of $\lambda^{i,k}$ k boxes must be moved each a distance of i . This gives the character ratio. The hook length formula here gives

$$\begin{aligned} d_{\lambda^{i,k}} &= \frac{n!}{(n-i+k+1)(n-i) \cdots (n-2i+k+2)(n-2i+k) \cdots (1)(i)(i-k-1)!k!} \\ &= \binom{n}{i} \binom{i-1}{k} \frac{n-2i+k+1}{n-i+k+1} \end{aligned}$$

And finally for the character difference, in equation (4) we see the coefficient of x_i in the character polynomial for $[n-i, i-k, 1^k]$ is $(-1)^r = (-1)^{k+2}$. All small cycle terms will cancel in the character difference leaving, $(-1)^k (a_i - b_i)$.

For $t \geq \frac{n^2(\log(i-1)+1)}{i}$ consecutive terms will be shown to fall by half in magnitude. Taken with the alternating signs this gives the desired result. The $i = 1$ case is trivial. Consider $2 \leq i \leq n/3$, the character differences are the same, and the ratio of dimension is at least:

$$d_{\lambda^{i,k}}/d_{\lambda^{i,k+1}} = \binom{i-1}{k} / \binom{i-1}{k+1} \frac{n-2i+k+1}{n-i+k+1} \frac{n-i+k+2}{n-2i+k+2} \geq \frac{1}{2(i-1)}$$

Meanwhile, the ratio of eigenvalues gives:

$$\frac{\left(1 - \frac{i(n-i+k+1)}{\binom{n}{2}}\right)^t}{\left(1 - \frac{i(n-i+k+2)}{\binom{n}{2}}\right)^t} \geq \left(1 + \frac{i}{\binom{n}{2}}\right)^t \geq e^{\frac{i}{n^2}t}$$

Taking $t \geq \frac{n^2(\log(i-1)+2)}{i}$ suffices for the product of these terms to be at least 2. Finally, $\sum_{k=0}^{i-1} \left(\frac{-1}{2}\right)^k \geq \frac{1}{2}$, for $i \geq 2$. \square

This is a central obstruction to cycle lexicographic order holding at the mixing time $\frac{1}{2}n \log(n) \pm cn$, let alone by $n - 2$ steps. When $i = 1$ this issue does not exist since it is the sole term with $\lambda_1 = n - 1$. However, for all other i simulation suggests that although the sum of the λ^i and $\lambda^{i,k}$ terms may be positive by $O(n \log(n))$ time, it will not be a constant fraction of the λ^i term until $O\left(\frac{n^2 \log(i)}{i}\right)$. For $i \geq \frac{n}{3}$ there is another impediment to showing the likelihood order holds after $O(n \log(n))$ steps.

We now turn to showing general bounds for the character differences, dimension ratios, and character ratio differences in the next three propositions. These will handle all the remaining cases of partitions.

Proposition 36. For $\lambda = [n - j, \dots]$ with $j > i$ and a i -cycle detector and α, β conjugacy classes of the same parity with $\min_k(a_k \neq b_k) = i$,

$$\chi_\lambda(\alpha) - \chi_\lambda(\beta) \leq (n - 2i)_{j-i} (j - i + 1)^2 \frac{n - i}{i}$$

Where there exists a constant C such that,

$$\log(n - 2i)_{j-i} (j - i + 1)^2 \frac{n - i}{i} \leq C((j - i)(n - 2i) + n)$$

Proof. We bound the valid insertions from the Murnaghan-Nayama rule containing non-trivial i -cycle insertions independent of sign. Each insertion is determined by how the cycles that insert below the first row of λ insert. The first insertion must be trivial and at least i in length, as both cycles must contain an i or larger cycle. There are at most $j - i + 1$ places in the first row to start this insertion leaving room for an i -cycle. Any insertions with no i or larger cycles inserted non-trivially, the first of which is below the first row, will cancel in the difference, so the sum in Murnaghan-Nakayama is restricted to insertions where at least one i or larger cycle is inserted non-trivially. This means that one of α, β has at least two cycles of size $\geq i$. Then since $\sum_{j < i} j(a_j) = \sum_{j < i} j(b_j)$ both are at most $n - 2i$. Following this first non-trivial i or larger cycle at most $j - i$ cycles can then additionally insert below the first row and there are at most $n - 2i$ of these. There are at most $(j - i)!$ ways of arranging these. And there are at most $\frac{n}{i} - 1$ choices for the first non-trivial i or larger cycle and $j - i + 1$ rows in which to insert it.

Finally, $\log(n - 2i)_{j-i} \leq (j - i) \log(n - 2i) \leq (j - i)(n - 2i)$, and $\log((j - i + 1)^2 \frac{n - i}{i}) \leq 3n$. \square

Proposition 37. Taking the sum over $\lambda = [n - j, \dots]$ with $j > i$ and $h_{2,1} \geq i$,

$$\sum_{\lambda} d_{\lambda} \leq \binom{n}{j} \binom{j}{a,b} e_2(j-i)$$

where if $j \geq \frac{3}{2}i$, $a = b = j/3$ otherwise $a = b = i/2$.

There exists a constant C such that summing over the same partitions,

$$\log \left(\left(\sum_{\lambda} d_{\lambda} \right) / d_{\lambda^i} \right) \leq C((j-i)(n-2i) + n(1 + |n-2j|_+))$$

Proof. The number of standard Young tableau of shape λ is d_{λ} . To upper bound this count, λ will be separated into three pieces. The first piece is $\lambda_1 = n - j$. The second piece is $h_{2,1}$, of which it is restricted to $i \leq h_{2,1} \leq j$ and is made of λ_2 and $\lambda'_1 - 2$. This leaves a partition consisting of all blocks not in $\lambda_1, \lambda_2, \lambda'_1$, which contains at most $j - i$ blocks. Consider all possible ways of placing $1, 2, \dots, n$ into these parts so that each row/column inside part is increasing. Not all of these will give valid tableau but all possible tableau will be present. First choose $n - j$ of n to place in λ_1 . Then λ_2 and λ'_1 of the remaining j . Let a, b be the values of $\lambda_2, \lambda'_1 - 2$ that maximize this over all λ . Finally this leaves a partition of size at most $j - i$. The sum of the dimensions of all partitions of $j - i$ is a combinatorial object called a telephone number and denoted $e_2(j - i)$ (the name comes from a bijection with involutions of S_n). From [11], $e_2(j - i) \leq \left(\frac{j-i}{e}\right)^{(j-i)/2} e^{\sqrt{j-i}}$. This gives

$$\sum_{\lambda} d_{\lambda} \leq \binom{n}{j} \binom{j}{a,b} e_2(j-i)$$

This leaves finding values of a, b . $\binom{j}{a,b}$ is maximized when the parts are as equal as possible in size, so clearly $a = b$. Also, $i \leq a + b \leq j$. So if $j \geq \frac{3}{2}i$, all parts can be equal letting $a = b = j/3$. Otherwise, it is maximized when $a = b = i/2$, letting the remaining part be as large as possible.

Recall from Proposition 32 that $d_{\lambda^i} = \binom{n}{i} \frac{n-2i+1}{n-i+1} \geq \binom{n}{i} \frac{1}{n}$. First take the case that $j \leq n - i$. The largest terms from the dimension of each mostly cancel:

$$\log \left(\binom{n}{j} / \binom{n}{i} \right) \leq (j-i) \log \left(\frac{n-i}{i} \right) \leq (j-i) \log(1+n-2i) \leq n + (j-i)(n-2i).$$

It remains to show a similar bound on $n, \binom{j}{a,b}$ and $e_2(j-i)$. The first is trivial, and then $\log \left(\binom{j}{a,b} \right) \leq \log(3^j) \leq 2n$. While, using the bound from Knuth for the telephone numbers, $\log(e_2(j-i)) \leq \frac{(j-i)}{2} \log(j-i) + \sqrt{j-i}$. Since $j \leq n - i$, then $\log(j-i) \leq n - 2i$.

Second, we consider the case $j > n - i$. Now, $\binom{n}{j} \leq \binom{n}{i}$, and these terms can be bounded by 1. The logarithm of n and $\binom{j}{a,b}$ are both bounded by a constant times n . Lastly, $\log(e_2(j-i)) \leq (j-i) \log(j-i)$, which can be broken into:

$$\left(j - \frac{n}{2}\right) \log(j-i) + \left(\frac{n}{2} - i\right) \log(j-i) \leq \left(j - \frac{n}{2}\right) n + (j-i) \left(\frac{n}{2} - i\right). \quad \square$$

We finish by showing a difference in the character ratios that matches the bounds in the previous two propositions.

Proposition 38. *For $1 \leq i \leq n/2$, $\lambda_1 = n - j$ with $j \geq i$ and either $i \leq n/3$ or $\lambda \neq [n - j, j]$, there exists a constant C such that,*

$$Cn^2(r_{\lambda^i} - r_\lambda) \geq (j - i)(n - 2i + 1) + n(1 + |n - 2j|_+)$$

When $i < j \leq n/2$ and $\lambda = [n - j, j]$, there exists a constant C so that $Cn^2(r_{\lambda^i} - r_{\lambda^j}) \geq (j - i)(n - 2i + 1)$.

Proof. For fixed $i, j \leq n/2$, the smallest character ratio difference occurs for $\lambda = \lambda^j$ with:

$$r_{\lambda^i} - r_{\lambda^j} = \frac{(j - i)(n - i - j + 1)}{\binom{n}{2}} \geq \frac{2(j - i)(n/2 - i + 1)}{n(n - 1)} \geq \frac{(j - i)(n - 2i + 1)}{n^2}$$

This finishes the second statement of the proposition. When $i \leq n/3$, $n - 2i + 1 \geq n/3$, so the extra can be used to show the first case. If $n/3 < i < j \leq n/2$, but $\lambda \neq [n - j, j]$, there is a box in the partition λ in the third row. Moving this box down from the second row of $[n - j, j]$, it travels distance at least j . Then since j is order n , the statement holds.

Finally, if $n/3 < i \leq n/2 < j$, we will estimate the largest the character ratio for $\lambda = [n - j, \dots]$ and then find the difference. the largest eigenvalue corresponds to the partition λ with $\lfloor \frac{n}{n-j} \rfloor = m$ rows of $n - j$ boxes and a final row of $n - (n - j)m$ blocks. These all have non-negative character ratios since they are above their conjugates in majorization order. Moving boxes down and to the left from $[n]$ to λ , we will count the distance they travel. In the first row, the first $n - j$ stay in place, while the next $n - j$ each travel a distance of $n - j + 1$ from the second row. The third group of $n - j$ travel $2(n - j + 1)$ from the third row, and so on, ending with the spare $n - (n - j)m$ blocks each traveling $m(n - j + 1)$. This gives a total distance traveled of

$$\binom{n}{2} \left(1 - \frac{\chi_\lambda(\tau)}{d_\lambda} \right) \geq \left(\sum_{k=0}^{m-1} k(n - j)(n - j + 1) \right) + m(n - j + 1)(n - (n - j)m) \quad (5)$$

$$= \frac{1}{2}(m - 1)m(n - j + 1)(n - j) + m(n - j + 1)(n - (n - j)m) \quad (6)$$

$$= m(n - j + 1) \left(\frac{1}{2}(m - 1)(n - j) + (n - (n - j)m) \right) \quad (7)$$

$$= (m)(n - j + 1) \left(n - \frac{1}{2}(m + 1)(n - j) \right) \quad (8)$$

$$= \frac{1}{2}(m)(n - j + 1)(n + j - m(n - j)) \quad (9)$$

$$\geq \frac{1}{2}j(n + 2) \quad (10)$$

The inequality (10) follows from viewing it as a downwards parabola in m , and saying the minimum value must occur at either $m = \frac{n}{n-j}$ or $\frac{n}{n-j} - 1$. The former reduces to $\frac{1}{2}j(n + n/(n - j)) \geq j(n + 2)$ and the latter to $\frac{1}{2}(j + \frac{n}{n-j} - 1)(n + 2j)$.

Then,

$$\binom{n}{2} (r_{\lambda^i} - r_\lambda) = \frac{1}{2}j(n+2) - i(n-i+1) = (j-i)(n/2 - i + 1) + (j - n/2)i,$$

which since $i > n/3$, is at least as big as the claimed difference in the statement for some constant. \square

5.4 The Lazy Version

When the walk is modified to be lazy, the likelihood order integrates even and odd permutations into one total order. The representations λ and λ' originally worked jointly to impose the parity restrictions, but now λ will be the sole lead i -cycle detector. However, the partition $[1^n]$ will now act as a sign detector. It will function to make even permutations more likely than odd or vice versa depending on laziness and parity of time. Modifications to the previous results give a bound for sufficient time for the likelihood order to hold. If during a step the chain stays with probability p , and moves via a transposition with probability $1 - p$, the probability function becomes as from Proposition 8:

$$P^{*t}(\alpha) = \sum_{\lambda} \chi_{\lambda}(\alpha) d_{\lambda} \left(p + (1-p) \frac{\chi_{\lambda}(\tau)}{d_{\lambda}} \right)^t$$

The i -cycle detector with largest eigenvalue is now solely λ^i with eigenvalue $1 - (1 - p) \frac{2i(n-i+1)}{n(n-1)}$ instead of jointly λ^i and $(\lambda^i)'$, as for $\lambda \supseteq \lambda'$:

$$p + (1-p) \frac{\chi_{\lambda}(\tau)}{d_{\lambda}} \geq \left| p - (1-p) \frac{\chi_{\lambda}(\tau)}{d_{\lambda}} \right|$$

The sign detector $[1^n]$ has eigenvalue $2p - 1$, and term in the Fourier Inversion formula $(\text{sg}(\alpha) - \text{sg}(\beta))(2p - 1)^t$.

When $p \geq \frac{1}{2}$, all eigenvalues are positive, and it is always the case that $1 - (1 - p) \frac{2i(n-i+1)}{n(n-1)} > |2p - 1|$, and so cycle lexicographic order will hold on all permutations after a sufficient time.

When $p < \frac{1}{2}$, for any i with $1 - (1 - p) \frac{2i(n-i+1)}{n(n-1)} \geq 1 - 2p$, normal CL order holds (the greater dimension of λ^i determines it at equality). The inequality simplifies to $\frac{i(n-i+1)}{n(n-1)} < \frac{p}{1-p}$. When the inequality is reversed, if $\alpha \beta$ first differ at i cycles and are of opposite sign, the even will be more likely at even times, the odd more likely at odd times. CL order will still hold restricted to the evens and odds respectively after sufficient time. In particular, for $p = \frac{1}{n}$, the traditional value for the lazy transposition walk, only $i = 1$ escapes this split by have the same magnitude of eigenvalue as $[1^n]$. This split can be thought of as a recalcitrant unzipped zipper. The force you pull on the zipper with is proportional to p . If p is large, you zip all of the CL_{even} , CL_{odd} together into CL . When p is small, the zipper catches above the i were that inequality triggers, leaving some portion of the two orders zipped together. To read the order, first read the zipped part, then if t is odd the CL_{odd} part before the CL_{even} , and vice versa if t is even.

The $[1^n]$ term will dominate the λ^i when $\frac{i(n-i+1)}{n(n-1)} > \frac{p}{1-p}$ and for two conjugacy classes α, β of opposite parity first differing at i . The time bound depends particularly strongly on the choice of p , as one can make the eigenvalue of $[1^n]$ arbitrarily close to that of λ^i .

We will now extend the proof of order n^2 time being sufficient for the likelihood order to have stabilized to two special cases of laziness, first $P(e) = \frac{1}{n}$ the traditional laziness for the transposition walk as used by Diaconis and Shahshahani [7], and second a large but constant laziness of $P(e) \geq \frac{1}{2}$. Both proofs extend the proof of Theorem 4 (1). Both have to handle the new case of α, β different signs, and transposes of partitions, which were handled by symmetry before.

Proof of Theorem 4 (2). When α and β are the same sign and first differ at an i -cycle, we need to show the $[n-i, i]$ term in Proposition 8 dominates all other terms. The eigenvalues for this walk are $r_\lambda = \frac{1}{n} + \frac{n-1}{n}r_\lambda(\tau)$. For $\lambda \supseteq \lambda'$, since $r_\lambda(\tau) = -r_{\lambda'}(\tau)$ and $r_{\lambda'}(\tau) \geq r_\lambda(\tau)$, $r_\lambda(\tau) \geq 0$ and $r_\lambda \geq 0$. Therefore the difference in magnitude between the eigenvalues of $[n-i, i]$ and such a λ , is $\frac{n-1}{n}$ of the old difference in magnitude and the arguments of Theorem 4 (1) carry through with the same scaling coefficients.

We still need to show the $[n-i, i]$ term dominates the term for those partitions λ with $\lambda \triangleleft \lambda'$. In particular, we have to handle $[n-i, i]' = [2^i, 1^{n-i}]$. As noted in the proof of Proposition 33, $\chi_\lambda(\alpha) = \text{sg}(\alpha)\chi_{\lambda'}(\alpha)$, so for the case of α, β the same parity, we can reuse the bound on the character difference from Proposition 38. The dimensions bound from Propositions 37 will also be recycled. It suffices to show a large difference in magnitude of eigenvalues. Since $r_\lambda(\tau) < 0$, if r_λ is positive its bounded above by $\frac{1}{n}$, and there is at least a constant order gap between it and $r_{[n-i, i]}$. Otherwise, if $\lambda'_1 = n-j$ with $j \geq i$, there exists a constant C so,

$$n^2 (|r_{[n-i, i]}| - |r_\lambda|) = 2n + n(n-1)(r_{[n-i, i]}(\tau) - r_{\lambda'}(\tau)) \leq C(2n + (j-i)(n-2i) + n|j-n/2|_+) \quad (11)$$

Note this now holds even if $j = i$, where we now check that the dimension and character difference bounds suffice. Note that with a tiny scaling coefficient for these terms, $c'_j = 2^{-n-j}$, $-\log(c'_j)$ is still bounded by a constant times n^2 times the eigenvalue gap. The $\lambda' = [n-i, i-k, 1^k]$ terms have dimension $\binom{n}{i} \binom{i-1}{k} \frac{n-2i+k+1}{n-i+k+1}$ versus $\lambda = [n-i, i]$ with dimension $\binom{n}{i} \frac{n-2i+1}{n-i+1}$. The dimension ratio is thus bounded by $n2^{i-1}$. The character differences are equal up to sign. Using the eigenvalue difference times n^2 is order n , this easily tops the logarithm of the dimension ratio.

When α and β are of difference sign we need to show (a) if they first differ at a 1-cycle that the $[n-1, 1]$ term dominates or (b) otherwise that the $[1^n]$ term dominates.

In this case if $a_1 > b_1$, then the difference of the $[n-1, 1]$ and $[1^n]$ terms is

$$((a_1 - b_1)(n-1) - 2) \left(1 - \frac{2}{n}\right)^t > 0.$$

The same argument as from Theorem 4 (1) handles the remaining partitions with $\lambda \supseteq \lambda'$. We reuse the argument above for the remaining partitions, except we need to adjust the character difference term. Since $\chi_\lambda(\alpha) = -\chi_{\lambda'}(\alpha)$ and α, β are of different signs,

$$|\chi_\lambda(\alpha) - \chi_\lambda(\beta)| = |\chi_\lambda(\alpha) + \chi_\lambda(\beta)| \leq 2d_\lambda$$

Fortunately, since we only have to handle $i = 1$, the eigenvalue difference bound from (11) suffices. Using if $\lambda'_1 = n - j$, $d_\lambda \leq \binom{n}{j} e_2(j) \leq n^j$, then there exists a constant C such that $\log(n^j) = j \log(n) \leq C(j - 1)(n - 2)$.

Finally, we show that the $[1^n]$ term dominates the remaining terms where $\max(\lambda_1, \lambda'_1) \leq n - 2$. It has the same eigenvalue gap as $[n - 1, 1]$ just a smaller coefficient of 2 rather than $(a_1 - b_1)(n - 1) \leq n^2$. The argument for the $i = 1, j \geq 2$ terms from Theorem 4 (1) can absorb the difference. Similarly, so can the argument in the prior paragraph. \square

When laziness is very large, all eigenvalues are positive, and it is much simpler to show order n^2 time is sufficient for the likelihood order to stabilize to cycle lexicographic.

Proof of Theorem 4 (2). The eigenvalues for this walk are $r_\lambda = p + (1 - p)r_\lambda(\tau) \geq 0$. As all the eigenvalues are positive, the difference in magnitude between the eigenvalue for $[n - i, i]$ and the eigenvalue for λ is just the difference between their eigenvalues, and is:

$$(p + (1 - p)r_{[n-i,i]}(\tau)) - (p + (1 - p)r_\lambda(\tau)) = (1 - p)(r_{[n-i,i]}(\tau) - r_\lambda(\tau))$$

When $\lambda \succeq \lambda'$, this gap is within a constant of what it is in the proof of Theorem 4 (0), and $O(n^2)$ time still suffices for $[n - i, i]$ to dominate these λ .

Otherwise, if $\lambda \triangleleft \lambda'$, we use the following bounds. First $|\chi_\lambda(\alpha) - \chi_\lambda(\beta)| \leq 2d_\lambda$. Incorporating this into the sum of dimensions, the sum of dimensions squared of all partitions of n is $n!$ by the RSK algorithm (see e.g. [18]). Second, $r_\lambda(\tau) < 0$ when $\lambda \triangleleft \lambda'$ since $r_\lambda(\tau) = -r_{\lambda'}(\tau)$ and $r_{\lambda'}(\tau) \geq r_\lambda(\tau)$, and so

$$r_{[n-i,i]} - r_\lambda \geq r_{[[n/2],[n/2]]} - p \geq \frac{1}{2}(1 - p) \left(1 - \frac{3}{n - 1}\right)$$

Since this is bounded from below by a constant, to show t of order n^2 is sufficient, this means we only need an order n^2 bound on the logarithm of the scaling factor, character differences, and sum of dimensions. Letting the scaling factor for these be $2^{-(n+2)}$, we see $\log(2^{(n+2)}) \leq n + 2$ and $\log(n!) \leq n \log(n) \leq n^2$. \square

5.5 More or Less Likely than Stationary

This style of analysis can also be insightful into total variation distance as well as separation distance. One of the equivalent definitions of total variation distance between the walk at time t and stationary is

$$\|P^{*t} - \pi\|_{TV} = \sum_{\sigma \in S_n, P^{*t}(\sigma) \geq \pi(\sigma)} [P^{*t}(\sigma) - \pi(\sigma)]$$

So knowing the permutations that are more likely than stationary leads to being able to calculate the total variation distance for each of the variants on the transposition walk discussed: non-lazy at even times, non-lazy at odd times, and lazy. The stationary distributions are respectively uniform over even permutations, odd permutations, and all

permutations. The long time behavior is seen in the Fourier inversion formula in the $[n]$ term giving the stationary distribution, and if $P(e) = 0$ in the $[1^n]$ term as well controlling periodicity. After sufficient time, the sign of the next non-zero term determines whether the Fourier inversion formula evaluated at a permutation will be more or less likely than uniform.

The next non-zero term will be the first of $[n-1, 1]$, $[n-2, 2]$, $[n-2, 1, 1]$ with non-zero character at α .

λ	$\chi_\lambda(\alpha)$	$a_1 > 1$	$a_1 = 0$	$a_1 = 1, a_2 > 1$	$a_1 = 1, a_2 = 0$	$a_1 = 1, a_2 = 1$
$n-1, 1$	$a_1 - 1$	> 0	< 0	0	0	0
$n-2, 2$	$a_2 + \binom{a_1}{2} - a_1$			> 0	< 0	0
$n-2, 1, 1$	$-a_2 + \binom{a_1}{2} - a_1 + 1$					< 0

So after sufficient time, a permutation is more likely than uniform if it has at least two fixed points, or one fixed point and at least two 2-cycles. And a permutation is less likely than uniform if it has no fixed points or one fixed point and at most one 2-cycle.

An analysis similar to section 5.3 can be used to show this relationship holds by $O(n \log(n))$ steps, except for the case $a_1 = 1, a_2 > 1$, when to show the $[n-2, 2]$ eigenvalue overwhelms the $[n-2, 1^2]$ eigenvalue $O(n^2)$ steps are needed, similar to Proposition 35. The strategy would be similar to the above analysis but with only $\lambda^1, \lambda^2, \lambda^{2,1}$ dominating, and with Proposition 7 not Proposition 8, $\chi_\lambda(\alpha)$ replacing $\chi_\lambda(\alpha) - \chi_\lambda(\beta)$. The interested reader pointed towards section 3.5.5 of the author's thesis in which such analysis is conducted.

6 Likelihood orders for cycle walks

The likelihood order will be found for any p -cycle walk, for n and t sufficiently large by extending an approximation for the character ratio $r_\lambda(p)$ from Wasserman's thesis. This utilizes a generating function formula for the character ratios introduced by Frobenius[9]. One case that Frobenius explicitly computed was the character ratio for the transpositions. Formulas for other small values of p , and small mixed cycle structures can be found in [17]. Unfortunately, though the first order approximation, with error $O(\frac{1}{n})$, Wasserman computed extends to general cycle structures, the step with multiple cycles cannot be extended to error of $O(\frac{1}{n^2})$. As we saw for the transposition walk with the eigenvalues for the $[n-2, 2]$ and $[n-2, 1^2]$ partitions with which differ by $\frac{4}{n(n-1)}$, small order differences in eigenvalues can determine the i -cycle detector.

In his thesis Wasserman [20] found the following formula, reprinted with proof by Flatto, Odlyzko, and Walkes[8]. It utilizes Frobenius notation for a partition λ , in which $a_i := \lambda_i - i$, $b_i := \lambda'_i - i$.

Theorem 39. [Wasserman] *Let γ be a permutation of $1, \dots, m$ with γ_2 2-cycles, γ_3 3-cycles, etc; thus γ may be considered an element of S_n for $n \geq m$. Let $\lambda = \begin{pmatrix} a_1 & \dots & a_s \\ b_1 & \dots & b_s \end{pmatrix}$.*

Then,

$$r_\lambda(\gamma) = \prod_{p \geq 2} \left(\sum_{i=1}^s [(a_i + 1/2)^p - (-b_i + 1/2)^p] / n^p \right)^{\gamma_p} + O(1/n)$$

Extending this formula to an error margin of $O(1/n^2)$ for p -cycles will be sufficient to show when λ is $[n - i, i - k, 1^k]$ for some k this character ratio is maximized over all i -cycle detectors. Extending to $O(1/n^3)$ will show which $[n - i, i - k, 1^k]$ is dominant. It will always be either $[n - i, i]$ or $[n - i, 1^i]$. Through line 6 of the following proof is from the argument by Wasserman.

Corollary 40.

$$r_\lambda(p) = \frac{s_p - p/2 \sum_{k=1}^{p-2} s_k s_{p-k-1}}{(n)_p} + O(1/n^2)$$

$$r_\lambda(p) = \frac{1}{(n)_p} \left(s_p - p/2 \sum_{k=1}^{p-2} s_k s_{p-k-1} + \left(\frac{p^3(p-1)^2}{8} - \frac{p^2(p-1)}{24} \right) s_{p-2} + \frac{p^2}{6} \sum_{l,m \geq 1} s_l s_m s_{p-2-l-m} \right) + O(1/n^3)$$

where $s_k = \sum_{i=1}^s [(a_i + 1/2)^k - (-b_i + 1/2)^k]$ and $s_1 = n$. The constants are functions of p .

Proof. From Frobenius [9] as also described in Murnaghan [16], letting $x_i := a_i + 1/2, y_i := b_i + 1/2, F(x) := \prod_{i=1}^s \frac{x+y_i}{x-x_i}$, then $r_\lambda(p)$ is the coefficient of $1/x$ in

$$\frac{-(x+1/2) \cdots (x+p-1/2) F(x+p)}{p(n)_p F(x)}.$$

For $|x|$ sufficiently large to expand log as a Laurent expansion,

$$\log F(x) = \sum_{i=1}^s \left[\log \left(1 + \frac{y_i}{x} \right) - \log \left(1 - \frac{x_i}{x} \right) \right] = \sum_{j=1}^{\infty} \frac{s_j}{j x^j}.$$

Where s_j is as defined above. Then expanding $\frac{F(x+p)}{F(x)}$ gives:

$$\exp \left\{ \sum_{j=1}^{\infty} \frac{s_j}{j x^j} \left[\left(1 + \frac{p}{x} \right)^{-j} - 1 \right] \right\} = \sum_{k=0}^{\infty} \frac{1}{k!} \left\{ \sum_{j=1}^{\infty} \frac{s_j}{j x^j} \left[\sum_{l=1}^{\infty} \binom{-j}{l} \left(\frac{p}{x} \right)^l \right] \right\}^k$$

So $r_\lambda(p)$ is the coefficient of $1/x$ in

$$\frac{-(x+1/2) \cdots (x+p-1/2)}{p(n)_p} \sum_{k=0}^{\infty} \frac{1}{k!} \left\{ \sum_{j=1}^{\infty} \frac{s_j}{j x^j} \left[\sum_{l=1}^{\infty} \binom{-j}{l} \left(\frac{p}{x} \right)^l \right] \right\}^k.$$

Note that $|s_j| \leq \sum_{i=1}^s (x_i^j + y_i^j) \leq [\sum_{i=1}^s (x_i + y_i)]^j = n^j$. For p fixed and n sufficiently large, the largest contributions to $r_\lambda(p)$ can be ranked by J from terms $s_{j_1} \cdots s_{j_k}$ where

$j_1 + \dots + j_k = J \leq p + 1 - k$. There is a unique way to make this sum p , so as Wasserman observed,

$$r_\lambda(p) = \frac{s_p}{(n)_p} + O(1/n).$$

The coefficient of s_{p-1} is zero, as

$$\begin{aligned} & \frac{-(1/2 + \dots + (p-1/2))x^{p-1}}{p(n)_p} \frac{1}{1!} \frac{s_{p-1}}{(p-1)x^{p-1}} \binom{-(p-1)}{1} \frac{p}{x} \\ & + \frac{-x^p}{p(n)_p} \frac{1}{1!} \frac{s_{p-1}}{(p-1)x^{p-1}} \binom{-(p-1)}{2} \frac{p^2}{x^2} = 0. \end{aligned}$$

The remaining solutions to $j_1 + \dots + j_k = p - 1 \leq p + 1 - k$ give the terms:

$$\frac{-x^p}{p(n)_p} \frac{1}{2!} \left(\sum_{j=1}^{p-2} \frac{s_j}{jx^j} \binom{-j}{1} \frac{p}{x} \frac{s_{p-1-j}}{(p-1-j)x^{p-1-j}} \binom{-(p-1-j)}{1} \frac{p}{x} \right) = \frac{1-p/2}{x} \frac{\sum_{j=1}^{p-2} s_j s_{p-1-j}}{(n)_p}$$

There are three ways to arrive at a coefficient of $\frac{1}{x}$ from s_{p-2} . These correspond to the coefficients of x^p, x^{p-1}, x^{p-2} in $(x + \frac{1}{2}) \cdots (x + p - \frac{1}{2})$ and correspondingly $\binom{-(p-2)}{3} (p/x)^3, \binom{-(p-2)}{2} (p/x)^2, \binom{-(p-2)}{1} (p/x)$.

$$\begin{aligned} & \frac{-x^p}{p} \frac{s_{p-2}}{(p-2)x^{p-2}} \binom{-(p-2)}{3} (p/x)^3 + \frac{-x^{p-1} \frac{p^2}{2}}{p} \frac{s_{p-2}}{(p-2)x^{p-2}} \binom{-(p-2)}{2} (p/x)^2 \\ & + \frac{-\frac{p(p-1)(3p^2-p-1)}{24} x^{p-2}}{p} \frac{s_{p-2}}{(p-2)x^{p-2}} \binom{-(p-2)}{1} (p/x)^1 \\ & = \frac{s_{p-2}}{x} \left(\frac{p^3(p-1)^2}{8} - \frac{p^2(p-1)}{24} \right) \end{aligned}$$

In the same way that the s_{p-1} term vanishes, the $\sum s_l s_{p-1-l}$ terms also cancel.

$$\begin{aligned} & \frac{-x^p}{p} \frac{1}{2} \left(\sum_{j=1}^{p-1} \frac{s_j}{jx^j} \binom{-j}{2} \left(\frac{p}{x} \right)^2 \frac{s_{p-1-j}}{(p-1-j)x^{p-1-j}} \binom{-(p-1-j)}{1} \frac{p}{x} \right. \\ & \quad \left. + \frac{s_j}{jx^j} \binom{-j}{1} \left(\frac{p}{x} \right) \frac{s_{p-1-j}}{(p-1-j)x^{p-1-j}} \binom{-(p-1-j)}{2} \left(\frac{p}{x} \right)^2 \right) + \\ & \frac{-1}{p} \frac{p^2}{2} x^{p-1} \frac{1}{2} \left(\sum_{j=1}^{p-2} \frac{s_j}{jx^j} \binom{-j}{1} \left(\frac{p}{x} \right) \frac{s_{p-1-j}}{(p-1-j)x^{p-1-j}} \binom{-(p-1-j)}{1} \left(\frac{p}{x} \right) \right) = 0 \end{aligned}$$

The last that can be built, $s_l s_m s_{p-2-l-m}$, is through a single way:

$$\begin{aligned} & \frac{-1}{p} x^p \frac{1}{6} \sum_{m,l} \frac{s_m}{mx^m} \binom{-m}{1} \frac{p}{x} \frac{s_l}{lx^l} \binom{-l}{1} \frac{p}{x} \frac{s_{p-2-l-m}}{(p-2-l-m)x^{p-2-l-m}} \binom{-(p-2-l-m)}{1} \frac{p}{x} \\ & = \frac{p^2}{6} \frac{\sum_{l,m} s_l s_m s_{p-2-l-m}}{x} \quad \square \end{aligned}$$

Now that we have fine estimates for the character ratios at p -cycles, we need to compare them at two partitions. The following corollary takes advantage of some cancellation of terms to simplify that process.

Corollary 41.

$$\begin{aligned} r_\lambda(p) - r_\rho(p) &= \frac{1}{(n)_p} \left((s_p - s'_p) - \frac{p}{2} \sum_{l>1} (s_l - s'_l)(s_{p-1-l} + s'_{p-1-l}) \right. \\ &\quad + \left(\frac{p^3(p-1)^2}{8} - \frac{p^2(p-1)}{24} \right) (s_{p-2} - s'_{p-2}) \\ &\quad + \frac{p^2}{2} \sum_{l>1, m \geq 1} ((s_l - s'_l)((s_m + s'_m)(s_{p-2-l-m} + s_{p-2-l-m}) \\ &\quad \left. + \frac{1}{3}(s_m - s'_m)(s_{p-2-l-m} - s'_{p-2-l-m}))) \right) + O(1/n^3) \end{aligned}$$

Where s_p corresponds to λ and s'_p corresponds to ρ .

Proof. It suffices to check that $\sum_{l \geq 1} s_l s_{p-1-l} - s'_l s'_{p-1-l} = \sum_{l>1} (s_l - s'_l)(s_{p-1-l} + s'_{p-1-l})$ and $\sum_{l, m \geq 1} s_l s_m s_{p-2-l-m} - s'_l s'_m s'_{p-2-l-m} = \sum_{l>1, m \geq 1} (s_l - s'_l)((s_m + s'_m)(s_{p-2-l-m} + s'_{p-2-l-m}) + \frac{1}{3}(s_m - s'_m)(s_{p-2-l-m} - s'_{p-2-l-m}))$. \square

Now to compare $r_\lambda(p)$ for i -cycle detecting λ . The i -cycle partitions with maximal first row are of the form $[n-i, i-k, 1^k]$ which translates in Frobenius notation to, for $k \leq i-1$, $\begin{pmatrix} n-i-1 & i-k-2 \\ k+1 & 0 \end{pmatrix}$ and for $k=i$, $\begin{pmatrix} n-i-1 \\ i \end{pmatrix}$. This gives respectively,

$$\begin{aligned} s_j &= (n-i-1/2)^j + (i-k-3/2)^j - (-(k+3/2))^j - (-1/2)^j \text{ or} \\ s_j &= (n-i-1/2)^j - (-(i+1/2))^j \end{aligned}$$

We the process of finding the likelihood order by showing that the lead i -cycle detector determining the likelihood order for these walks has a subhook of length i .

Theorem 42. For each i , $1 \leq i \leq n/2 - 1$, for some k , $0 \leq k \leq i-1$, $r_{[n-i, i-k, 1^k]}(p) > r_\lambda(p)$ for each i -cycle detecting λ .

Proof. When p is even, boxes above the diagonal count positively towards s_p and boxes below count negatively. Then s_p is maximized by having the most boxes up and to the right in the Ferrers diagram of λ . In this case, the maximum s_p over $[n-i, i-k, 1^k]$ occurs when $k=0$ and is $(n-i-\frac{1}{2})^p + (i-\frac{3}{2})^p - (\frac{1}{2})^p - (\frac{3}{2})^p$. Over all ρ with $\rho_1 < n-i$, the largest s'_p would be for $[n-i-1, i+1]$ with $(n-i-\frac{3}{2})^p + (i-\frac{1}{2})^p - (\frac{1}{2})^p - (\frac{3}{2})^p$. The difference between these two is of order $p(n-2i)((n-i-\frac{1}{2})^{p-2} + \dots + (i-\frac{3}{2})^{p-2})$.

When p is odd, all boxes count positively. The longer a row or column extends from the diagonal the more it contributes. The maximum s_p over $[n-i, i-k, 1^k]$ is for $k=i-1$ with its column of length i with $s_p = (n-i-\frac{1}{2})^p + (i+\frac{1}{2})^p$. Over $\rho' \leq \rho$ with $\rho_1 < n-i$,

the next largest occurs at $[n - i - 1, 1^{i+1}]$ with $s'_p = (n - i - \frac{3}{2})^p + (i + \frac{3}{2})^p$. The difference is at least $p(n - 2i - 2)(n - i - \frac{3}{2})^{p-2}$.

Its necessary to check that the $s_l s_{p-1-l}$ terms do not overwhelm these. The only partitions at risk have one block moved from the first row of a $[n - i, i - k, 1^k]$. For $[n - i - 1, a, 1^b]$, compare it against $[n - i, a, 1^{b-1}]$ or $[n - i, a - 1, 1^b]$. For each p , $s_p - s'_p$ is of order $p(n - i)^{p-1} - pa^{p-1}$ or $p(n - i)^{p-1} + (-1)^p pb^{p-1}$. Since $s_l - s'_l$ appears in the product of two, this difference will also step down a power. Removing from b is avoidable in each case other than $[n - i, 1^i]$ and $[n - i - 1, 1^{i+1}]$. When i is large, p is odd, and l is even, the sign change could present a problem. However, since $p - 1 - l$ will also be even, $s_{p-1-l} + s'_{p-1-l}$ is of order $(n - 2i)(n - i)^{p-2-l}$. This means the product of two terms is of order $(n - 2i)(n - i)^{p-3}$ compared to the $s_p - s'_p$ at $(n - 2i)(n - i)^{p-2}$.

For the case that i is large, its also necessary to check the terms with weight $p - 2$. The s_{p-2} term has the same sign and parity as the s_p term and will behave the same way. This leaves the products of three terms. A product of the form $(s_l - s'_l)(s_m + s'_m)(s_{p-2-l-m} + s'_{p-2-l-m})$ will drop, as in the product of two. For the second kind of product of three, $(s_l - s'_l)(s_m - s'_m)(s_{p-2-l-m} - s'_{p-2-l-m})$, the only concern is again the p odd $[n - i, 1^i]$ versus $[n - i - 1, 1^{i+1}]$. In this event, $p - 2$ is odd, so one of $l, m, p - 2 - l - m$ is odd introducing the same step down in degree. □

Next, we determine which of the partitions with a subhook of length i is the lead i -cycle detector that determines the likelihood order.

Proposition 43. *Let p be fixed and n sufficiently large. The lead i -cycle detector is $[n - i, 1^i]$ for all odd p and even $p \geq 4$ for $i < Cn^{\frac{p-3}{p-2}}$. For these i , the likelihood order for cycles of size i is the $(-1)^{i-1}$ cycle lexicographic order. For even $p \geq 4$ and $i \geq Cn^{\frac{p-3}{p-2}}$, the lead i -cycle detector $[n - i, i]$ and the likelihood order for cycles of size i is the cycle lexicographic order.*

Proof. The difference between $r_{n-i,i}$ and $r_{n-i,1^i}$ when p is even is

$$\frac{1}{n^p} \left((2i^p) - \frac{p}{2} \sum_{l \text{ even}} 2i^l (n - i)^{p-l-1} \left(1 - \frac{p}{n - i} \right) + \sum_{l \text{ even}, m} (i^l (n - i)^{p-2-l}) \right) + O\left(\frac{1}{n^3}\right)$$

When $i < Cn^{\frac{p-3}{p-2}}$ for some C the second terms dominate, principally the $l = 2$ term of order $2i^2(n - i)^{p-3}$, and this will be negative. When i is above this threshold, the $2i^p$ term dominates and $r_{n-i,i}$ will be larger than $r_{n-i,1^i}$.

When p is odd, the difference is

$$\begin{aligned} & \frac{1}{n^p} \left((-2pi^{p-1} - \frac{p}{2} \sum_{l \text{ even}} 2i^l ((n - i)^{p-l-1} - p(n - i)^{p-l-2}) \right. \\ & \left. + \frac{p^2}{6} \sum_{l, m} 2i^l (n - i)^{p-l-2} \left(1 - \frac{p}{n - i} \right) \right) + O\left(\frac{1}{n^3}\right). \end{aligned}$$

This is always negative for n sufficiently large, meaning $\rho_i = [n - i, 1^i]$ is the lead i -cycle detector, which gives $(-1)^{i-1}$ cycle lexicographic order.

Its left to rule out the other $[n - i, i - k, 1^k]$ terms. The difference between the s_{k-1} and s_k terms is

$$\left((i - k - \frac{1}{2})^{p-1} - (i - k - \frac{3}{2})^{p-1}\right) + (-1)^p \left(\left(k + \frac{1}{2}\right)^{p-1} + \dots + \left(k + \frac{3}{2}\right)^{p-1}\right)$$

In the case p is even, this is always positive and $k = 0$ is maximal, $k = i - 1$ minimal. When p is odd this is of order $p(i - 2k - 2)((k^{p-2} + \dots + (i - k)^{p-2})$. This makes $k = 0$ or $i - 1$ maximal. When i is small, $\frac{-p}{2}(s_2 - s'_2)(s_{p-3} - s'_{p-3})$ dominates. Since $k = i - 1$ minimizes s_2 , $[n - i, 1^i]$ will beat out each other k . When i is large, $s_p - s'_p$ instead is largest, giving $[n - i, i]$ for p even, $[n - i, 1^i]$ for p odd. \square

As a result, we can identify the least likely elements of the walk as described in Corollary 6 of the introduction.

Proof of Corollary 6. Both parities of p have $(-1)^{i+1}CL$ as the likelihood order for $p \geq 4$, $i \leq 5$ for n sufficiently large. When p is odd, the walk is restricted to A_n , so the least likely element must be of odd parity. The above four elements have the maximal number of 2-cycles with no fixed points for the given parity of n and parity of Cayley distance. Therefore, when compared to any other element of S_n the minimal i -cycle they differ at will be 1 or 2. Since all four are fixed point free, the other permutation must have more one cycles or fewer two cycles and be larger in $(-1)^{i+1}CL$ order. \square

7 Other Remarks

Many other random walks generated by a conjugacy class fail in various ways to have such nice likelihood orders after sufficient time. There may not be a unique largest i -cycle detector or various values of i may share a single partition. Or indeed, the largest i -cycle detector may be blind to i -cycles (as observed following Theorem 21) $[n - 3, 2, 1]$ is to 2-cycles. When the lead i -cycle detector fails to differentiate conjugacy classes, it falls to the next largest i -cycle detector and so on. Another issue is in adding a lazy component to a walk the likelihood order may change decidedly when for some i the largest i -cycle detectors all have negative character ratios. $[n - i, i], [n - i, 1^i]$ are particularly nice in these regards, though the latter fails to be nice for lazy walks. These failures can lead to partial or partially describable likelihood orders rather than total ones.

For example, in the case of the random walk generated by $(n - 1)$ -cycles, the character ratio is zero except at $\lambda = [n], [1^n], [n - i, 2, 1^{i-2}]$. Where $[n - i, 2, 1^i]$ is a i -cycle detector. $[n - 2, 2]$ doubles as the lead 1 and 2-cycle detector where $\chi_{[n-2,2]}(\alpha) = a_2 + \binom{a_1}{2} - a_1$. And the sign of a_i in $\chi_{[n-i,2,i]}(\alpha)$ is $(-1)^i$. The order after sufficient time is first based on $(a_2 - b_2) + \frac{(a_1 - b_1)(a_1 + b_1 - 3)}{2} > 0$, but if this does not resolve the 1,2-cycle detection, it falls to the 3-cycle detector now the lead 1, 2, 3-cycle detector to resolve the dispute and so on. In the case that both conjugacy classes have no 1 or 2-cycles at even times it becomes just $(-1)^i CL$ order.

For the case of non-conjugacy class walks, though a version of the Fourier inversion theorem holds, the walk does not act as a constant on the irreducible representations, and so $\hat{P}(\lambda)^t$ is not a constant times the identity, and $\text{Tr}(\lambda(g^{-1})\hat{P}(\lambda)^t)$ does not simplify to $\chi_\lambda(g)(r_\lambda)^t$. With sufficient knowledge of the eigenvectors of such a walk, some information about the likelihood order might still be recoverable. In particular, $\chi_{[n-1,1]}$ is likely not a good test function to find a tight lower bound. Diagonalizing just the permutation representation and understanding those terms in the Fourier inversion formula, might yield a very strong lower bound.

Acknowledgements

Thank you to my advisor, Persi Diaconis, for suggesting this problem.

References

- [1] David Aldous and Persi Diaconis. Strong uniform times and finite random walks. *Adv. in Appl. Math.*, 8(1):69–97, 1987.
- [2] M. Bernstein. A random walk on the symmetric group generated by random involutions. [arXiv:1606.09588](https://arxiv.org/abs/1606.09588).
- [3] M. Bernstein and E. Nestoridi. Cutoff for random to random card shuffle. [arXiv:1703.06210](https://arxiv.org/abs/1703.06210).
- [4] Megan Bernstein. *Random Walks on the Symmetric Group, Likelihood Orders, and Involutions*. PhD thesis, Stanford University, 2015.
- [5] Persi Diaconis. *Group representations in probability and statistics*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11. Institute of Mathematical Statistics, Hayward, CA, 1988.
- [6] Persi Diaconis and I. Martin Isaacs. Least likely elements for random walk on finite groups. Preprint.
- [7] Persi Diaconis and Mehrdad Shahshahani. Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete*, 57(2):159–179, 1981.
- [8] L. Flatto, A. M. Odlyzko, and D. B. Wales. Random shuffles and group representations. *Ann. Probab.*, 13(1):154–178, 1985.
- [9] G. Frobenius. Uber die charaktere der symmetrischen gruppe. *Sitz. Pruess. Akad. Berlin*, pages 516–534, 1900.
- [10] A.M. Garsia and A. Goupil. Character polynomials, their q-analogs and the kronecker product. *The Electronic Journal of Combinatorics*, 16(2), #R19 2009.
- [11] Donald E. Knuth. *The art of computer programming. Volume 3*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1973. Sorting and searching, Addison-Wesley Series in Computer Science and Information Processing.

- [12] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.
- [13] Dudley E. Littlewood. *The theory of group characters and matrix representations of groups*. AMS Chelsea Publishing, Providence, RI, 2006. Reprint of the second (1950) edition.
- [14] Nathan Lulov. *Random Walks on the Symmetric Group Generated by Conjugacy Classes*. PhD thesis, Stanford University, 1996.
- [15] I. G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, second edition, 1995. With contributions by A. Zelevinsky, Oxford Science Publications.
- [16] Francis D. Murnaghan. *The theory of group representations*. Dover Publications, Inc., New York, 1963.
- [17] S.J R.E. Ingram. Some characters of the symmetric group. *Proceedings of the AMS*, 1(3):358–369, 1950.
- [18] Richard 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.
- [19] Richard P. Stanley. *Enumerative combinatorics. Volume 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2012.
- [20] A.J. Wasserman. *Automorphic actions of compact groups on operator algebras*. PhD thesis, University of Pennsylvania, 1981.
- [21] G. White. The weak Bruhat order for random walks on Coxeter groups. [arXiv:1611.04098](https://arxiv.org/abs/1611.04098).