

# RAPIDLY MIXING RANDOM WALKS AND BOUNDS ON CHARACTERS OF THE SYMMETRIC GROUP

NATHAN LULOV

DIMACS Center  
Rutgers University  
Piscataway, New Jersey, 08855  
lulov@dimacs.rutgers.edu

IGOR PAK

Department of Mathematics  
MIT  
Cambridge, MA 02139  
pak@math.mit.edu

February 8, 2002

ABSTRACT. We investigate mixing of random walks on  $S_n$  and  $A_n$  generated by permutations of a given cycle structure. The approach follows methods developed by Diaconis, which requires certain estimates on characters of the symmetric group and uses combinatorics of Young tableaux. We conclude with conjectures and open problems.

## Introduction

The subject of random walks on groups has been of intense interest in the past two decades [D1,D2]. Various probabilistic, analytic and combinatorial techniques were introduced in order to obtain sharp bounds in special cases. In this paper we continue this study, analyzing mixing of the various random walks on  $S_n$  (and  $A_n$ ) by means of the noncommutative Fourier transform and combinatorics of Young tableaux. This approach originated in the seminal paper [DS] and has been used with great success in many publications (see review article [D2]). Our results improve known upper bounds for a series of conjugacy-invariant generating sets of symmetric groups.

Let  $G$  be a finite group, and let  $R$  be a generating set for  $G$ . Consider a random walk  $\{X_t\}$  on  $G$ , defined as  $X_0 = \text{id}$ ,  $X_{t+1} = X_t \cdot s$ , where  $s \in R$  is chosen uniformly

---

*Key words and phrases.* Random walks on groups, symmetric group, Young tableaux.

Typeset by  $\mathcal{A}\mathcal{M}\mathcal{S}\text{-}\mathcal{T}\mathcal{E}\mathcal{X}$

and independently. Under mild conditions  $X_t$  converges to a uniform distribution on  $G$ , and the *mixing time*  $\text{mix}$  can be defined (see section 1). Intuitively, the mixing time is the first time when the probability distribution of the random walk is close to uniform (in  $\ell_1$ -distance.) A general bound

$$(*) \quad \text{mix} \leq c \Delta(G, R)^2 |R| \log |G|$$

(see e.g. [AF, §15]) is often used to bound the mixing time, where  $\Delta(G, R)$  is the diameter of the corresponding Cayley graph, and  $c$  is a universal constant. This bound can be sharpened when  $R$  is symmetric, i.e. transitive under the action of  $\text{Aut}(G)$ :

$$(**) \quad \text{mix} \leq c \Delta(G, R)^2 \log |G|.$$

For example, when  $R$  is a conjugacy class, the above formula applies (see e.g. [AF]). As was pointed out in [DS], the diameter  $\Delta$  in this case is easy to bound from a ratio of character values.

The problem with bound (\*\*) is the factor  $\log |G|$ , which is not tight in many special cases. Some examples show that it cannot be removed completely, and the best one could hope for is  $\Omega(\log \log |G|)$  for general conjugacy classes. For example, for  $G = S_n$  the mixing time is indeed  $\theta(\log n)$  for a random walk generated by cycles of length  $n/2$ . Breaking the ‘ $\log |G|$ ’ barrier requires extreme delicacy and has been done in very few special cases. The main purpose of this work is to extend the family of examples for which  $\text{mix} = o(\log \log |G|)$ .

We consider random walks on the symmetric group  $S_n$  generated by single-cycle conjugacy classes  $[r1^{n-r}]$  of  $r$ -cycles in  $S_n$ . Understanding these walks is crucial for studying the behavior of general random walks on  $S_n$ . Many properties of  $[r1^{n-r}]$  walks may be extended to the broader type of  $[\rho 1^{n-r}]$  walks,  $\rho \vdash r$ , with a given number of fixed points  $n-r$ . A single-cycle walk  $[r1^{n-r}]$  turns out to be the easiest to analyze among all walks of this type.

Roughly, we show that the mixing time is, up to a constant, the expected time to touch all the elements in the permutation (the latter is an obvious lower bound for the mixing time [D1].) This shows that the symmetric group is very ‘robust’ with respect to mixing of random walks, a property of interest in both theoretical and practical context.

A few words about the history of the problem. Several estimates are known about mixing times of single-cycle walks. There is the pioneering result of Diaconis and Shahshahani [DS] on the mixing time for random transpositions  $[21^{n-2}]$  (the case  $r = 2$ ). The case of small  $r$  is treated in [R], and some results for large  $r$  (the case  $n-r = O(1)$ ) follow from [L]. In [R] (cf. [P]), a lower bound for the mixing time of walks with all conjugacy classes and a given number of fixed points is proved. We use it here to match our upper bound.

In this paper we compute asymptotic mixing times in case of all cycles of lengths  $r = (1 - o(1))n$ . Our results give tight upper bounds in this case. No upper bound better than  $O(\log n)$  was known before in this case (cf. [R]). Our primary tool of

evaluating the total variation distance is a special case of the Upper Bound Lemma of Diaconis and Shahshahani [DS]. We conclude with a fringe result on the mixing time of the random walk generated by *all* cycles (i.e. a union of all single-cycle conjugacy classes). Remarkably, this random walk mixes in exactly 2 steps!

A few words are due about random walks generated by conjugacy classes with no fixed points. The first results in this direction were obtained for  $n$ -cycles (see [D1,S]). A breakthrough was made in [L], where the author was able to solve the case of conjugacy classes  $[r^{n/r}]$ ,  $r \geq 2$ . We present here a few general conjectures for further study.

The rest of the paper is constructed as follows. In section 1, we recall some standard definitions and notation. We present our main results in section 2. In section 3, we present an upper bound due to Diaconis and Shahshahani [D1,DS]. In a short section 4, we clarify the case of conjugacy classes with no fixed points. In the remaining sections we give proofs of the results.

### Acknowledgements

We would like to express extreme gratitude to our advisor Persi Diaconis for the introduction to the world of random walks, for his interest, help and support. The results of this paper were obtained about four years ago, when both authors were Ph.D. students at Harvard University. Without Diaconis' persistent encouragement, this paper would never be written.

We would like to thank Sergey Fomin, Alex Lubotzky, Richard Stanley, Yuval Roichman and Uzi Vishne for helpful remarks. We also thank an anonymous referee for pointing a number of grammatical errors and typos in the previous version of the paper.

## 1. BASIC DEFINITIONS

Let  $G$  be a finite group, let  $R$  be a generating set, and let  $\mathcal{W} = \{X_t\}$  be a random walk starting at  $\text{id}$ , and defined as in the introduction. One can think of  $\mathcal{W} = \mathcal{W}(G, R)$  as a nearest neighbor random walk on the Cayley graph  $\Gamma = \text{Cayley}(G, R)$ . Denote by  $Q^t$  the probability distribution of the random walk  $\mathcal{W}$ :  $Q^t(g) = \mathbf{P}(X_t = g)$ , for all  $g \in G$ . Unless  $\Gamma$  is bipartite, the probability distribution  $Q^t$  converges to the uniform distribution  $U(g) = 1/|G|$ :  $Q^t \rightarrow U$  as  $t \rightarrow \infty$  (see e.g. [AF]). Denote by  $\mathbf{d}(t)$  the *total variation distance* after  $t$  steps:

$$\mathbf{d}(t) = \|Q^t - U\| = \max_{A \subset G} \left| Q^t(A) - \frac{|A|}{|G|} \right| = \frac{1}{2} \sum_{g \in G} \left| Q^t(g) - \frac{1}{|G|} \right|.$$

It is easy to see that  $\mathbf{d}(t)$  is nonincreasing and has an exponential decay (see [AD,AF]). Define the *mixing time*  $\text{mix}_\epsilon = \text{mix}_\epsilon(\mathcal{W})$  as follows:

$$\text{mix}_\epsilon = \min \{t : \mathbf{d}(t) < \epsilon\}.$$

It is known [AD] (see also [AF,D1,P]) that for  $t > 2 \text{mix}_\epsilon$ ,  $\epsilon < 1/4$ , we have  $Q^t(g) > (1 - 4\sqrt{\epsilon})/|G|$ , so the walk mixes in separation distance as well. A special value  $\text{mix}_{1/e}$  or  $\text{mix}_{1/4}$  is often defined as a *mixing time* (see [AF,LW,P]). For our purposes here we use a different definition, corresponding to a somewhat stronger notion of the mixing time.

Let  $\{(G_n, R_n), n = 1, 2, \dots\}$  be a sequence of finite groups and their generating sets, and let  $\mathcal{W}_n = \mathcal{W}(G_n, R_n)$  be a sequence of random walks. We say that random walks  $\{\mathcal{W}_n\}$  have a *cutoff* at  $\kappa = (\kappa_n)$ , if for every  $\delta > 0$ , we have :

$$\begin{aligned} \|Q_n^{(1-\delta)\kappa_n} - U_n\| &\rightarrow 1 \quad \text{as } n \rightarrow \infty, \\ \|Q_n^{(1+\delta)\kappa_n} - U_n\| &\rightarrow 0 \quad \text{as } n \rightarrow \infty, \end{aligned}$$

where  $U_n$  is the uniform distribution on  $G_n$ .

Let us state the cutoff of  $\{\mathcal{W}_n\}$  in the language of the mixing times  $\text{mix}_\epsilon$ : for every  $\epsilon > 0$ ,  $\epsilon < 1/2$ , we have

$$\frac{\text{mix}_\epsilon(\mathcal{W}_n) - \text{mix}_{(1-\epsilon)}(\mathcal{W}_n)}{\kappa_n} \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

In general, whenever a sequence of random walks  $\{\mathcal{W}_n\}$  has a cutoff, one can always take  $\kappa_n = \text{mix}_{1/4}(\mathcal{W}_n)$ . Example of standard random walks without cutoff are the standard random walks on  $\mathbb{Z}_n$  (see [D1]). The cutoff is a very strong notion of mixing, actively studied in the recent years (see [D2] and references therein). It roughly corresponds to phase transition and sharp threshold notions in Statistical Physics and Probabilistic Combinatorics (cf. [PV]). Establishing cutoff is a delicate matter, which usually involves obtaining sharp lower and upper bounds (see [D2]).

For the rest of the paper, unless specified otherwise, we will assume that  $G_n$  is either an alternating group  $A_n$  or a symmetric group  $S_n$ , and that the generating set  $C_n$  is a conjugacy class, corresponding to some partition  $\lambda \vdash n$ . We will refer to “all transpositions” or “all cycles of length  $n/2$ ” to indicate a sequence of conjugacy classes.

There are two problems with random walks generated by conjugacy classes. Recall that each odd conjugacy class in the symmetric group  $S_n$  (which belongs to  $S_n \setminus A_n$ ) generates  $S_n$ , and each even class (except the identity) generates the alternating group  $A_n \subset S_n$  (see e.g. [Dv]). In this paper we consider random walks generated by the conjugacy classes in  $S_n$ , without indication whether the group they generate  $A_n$  and  $S_n$ .

Another obstacle is due to *parity*. For an odd conjugacy class  $C_n$  as the generating set, the random walk  $\mathcal{W}(S_n, C_n)$  is supported on the subgroup  $A_n \subset S_n$  on even steps and on the complementary subset  $S_n \setminus A_n$  on odd steps. Speaking of uniform distribution on permutations, we always mean an appropriate specification, i.e. uniform on  $A_n$  for even generating classes, and uniform on either  $A_n$  or  $S_n \setminus A_n$  after even or odd number of steps for odd generating classes.

In some cases the transition distribution may be slightly modified to include a non-zero probability of the identity in order to have a truly uniform distribution on

$S_n$  as in [DS], but we do not consider such modifications, whose rate of convergence depends heavily on the probability assigned to the identity (see [D1,P]).

We use the notation  $\lambda = (\lambda_1, \lambda_2, \dots)$ ,  $\lambda \vdash n$  to denote partitions  $\lambda_1 + \lambda_2 + \dots = n$ , where  $\lambda_1 \geq \lambda_2 \geq \dots$ . Let  $m_j = m_j(\lambda) = \#\{i : \lambda_i = j\}$ . We use the notation  $[1^{m_1} 2^{m_2} \dots]$  to denote the conjugacy class in  $S_n$  corresponding to the partition  $\lambda$ . Denote by  $T_\lambda$  the corresponding irreducible representation of  $S_n$ , and let  $d_\lambda = \dim(T_\lambda)$ .

Denote by  $f_\lambda$  the number of Standard Young tableaux of shape  $\lambda$ . It is well known (see e.g. [M,JK]) that  $f_\lambda = d_\lambda$ . By  $\chi_\lambda([\mu])$  denote the character value of  $T_\lambda$  on the conjugacy class  $[\mu]$ . The latter can be determined by the Murnagama-Nakayama rule [M,JK].

Denote by  $\lambda'$  the conjugate partition  $\lambda'_1 \geq \lambda'_2 \geq \dots$ , where  $\lambda'_j = \#\{i : \lambda_i \geq j\}$ . Clearly,  $(\lambda')' = \lambda$ . Let  $p(n)$  be the number of partitions  $\lambda \vdash n$ . It is well known that  $p(n) = \exp(O(\sqrt{n}))$  (see e.g. [An]).

## 2. MAIN RESULTS

The main result of this paper is establishing cutoff for random walks on  $S_n$  generated by the single-cycle conjugacy classes.

**Theorem 2.1** *Consider a sequence of single-cycle conjugacy classes  $C_n = [r_n 1^{m_n}]$ , where  $m_n = n - r_n < n/2$ . Then a sequence of random walks  $\mathcal{W}(S_n, C_n)$  on  $S_n$  ( $A_n$ ) generated by  $C_n$  has a cutoff at*

$$\kappa_n = \frac{\log n}{\log n - \log m_n}.$$

The case when  $m_n = n - r_n = \theta(n)$  was already investigated by Roichman in [R], where he obtained the same asymptotic bounds for all conjugacy classes with  $m_n$  fixed points. However, when  $m_n = n - r_n = o(n)$ , Roichman's technique does not work while Theorem 2.1 implies a sublogarithmic mixing time. Particularly, we have the following result:

**Corollary 2.2** *Consider mixing times  $\varkappa_n = \text{mix}_{1/4}$  of the random walks on  $S_n$  generated by (sequence of) conjugacy classes  $[r_n 1^{m_n}]$ , where  $m_n = n - r_n$  is a function of  $n$ .*

- 1) *If  $m_n = O(n/\log n)$ , then  $\varkappa_n = O\left(\frac{\log(n)}{\log \log(n)}\right)$ , as  $n \rightarrow \infty$ .*
- 2) *If  $m_n = O(n/\exp(\log^\alpha n))$ ,  $1 > \alpha > 0$ , then  $\varkappa_n = O(\log^{1-\alpha} n)$ , as  $n \rightarrow \infty$ .*
- 3) *If  $m_n = O(n^{1-\epsilon})$ , and  $0 < \epsilon < 1$ , then  $\varkappa_n = O(1)$ , as  $n \rightarrow \infty$ .*

We would like to finish this section with the side result which is of independent interest.

**Theorem 2.3** *Let  $0 < \epsilon \leq 1/4$  be a fixed constant, and let  $\varkappa = \text{mix}_\epsilon$  be the mixing time of the a random walk on  $S_n$  generated by a union of all cycles  $R = \bigcup_{r=2}^n [r1^{n-r}]$ . Then  $\varkappa = 2$ .*

In other words, the theorem says that

$$\mathbf{d}(2) = \|Q_n^2 - U_n\| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Roughly, this means that in spite of slow mixing of random walks generated by “small” cycles, the number of these cycles is small compared to the number of “long” cycles, which dominate the generating set.

### 3. UPPER AND LOWER BOUNDS

Our primary tool of evaluating the total variation distance is the following special case of the Upper Bound Lemma of Diaconis and Shahshahani [DS].

**Upper Bound Lemma 3.1** [DS] *For a random walk on a group  $G$  generated by a conjugacy class  $C \ni \sigma$ , the variation distance is bounded by the sum over all nontrivial irreducible characters  $\rho \in \widehat{G}$  of the group  $G$*

$$\|Q^\kappa - U\|^2 \leq \sum_{\rho \in \widehat{G}, \rho \neq \text{id}} d_\rho^2 \phi_\rho^{2\kappa}(\sigma),$$

where  $d_\rho$  is the dimension and  $\phi_\rho(\sigma) = \chi_\rho(\sigma)/d_\rho$  is the normalized character.

This inequality translates a random walk problem into a problem of estimating values of  $S_n$ -characters. The complete proof of Theorem 2.1 uses this combinatorial technique and will be given in sections 5 and 6.

The lower estimate of the Theorem 2.1 is proved in [R] for any conjugacy class  $[\rho 1^{n-r_n}]$ , where  $\rho \vdash r_n$ , with a given number (a sequence of numbers) of fixed points  $m_n = n - r_n$ .

In a different direction, in [P] the lower bound for the rate of convergence was given for every  $n > 2$ . Namely, it was shown that

$$\|Q_n^\kappa - U_n\| > \frac{1}{32}$$

where  $\kappa = \kappa_n = \left\lfloor \frac{h_n}{h_n - h_m} \right\rfloor$  and  $h_i = 1 + \frac{1}{2} + \dots + \frac{1}{i} \sim \log i$ . We will not consider the lower bound any further in this paper.

## 4. OTHER CONJUGACY CLASSES

In this section we investigate the special case of conjugacy classes in  $S_n$  with no fixed points, as well as some other cases. Although we have no new results to present here, we would like to recall the main result from [L] and state some general conjectures for further study.

Let us start with an open problem, which was mentioned in [Lb], and erroneously (or, perhaps, too optimistically?) referred to as a known result, and was attributed to the first author. We present it here as the main conjecture.

**Conjecture 4.1** *Let  $C_n$  be a sequence of conjugacy classes in  $S_n$  with no fixed points. Then the mixing time  $\varkappa$  is either 2 or 3.*

While this conjecture is supported by all the available evidence, we are unable to prove even that  $\varkappa = O(1)$  in most cases. We challenge the reader to make a progress in this direction. Here is the main result from the thesis [L] of the first author:

**Theorem 4.2** [L] *For the mixing time  $\varkappa$  of the random walk on  $S_n$  generated by  $[r^{n/r}]$  ( $r \geq 2$  is fixed), we have  $\varkappa = 3$  for  $r = 2$ ,  $\varkappa = 2$  for  $r \geq 3$ .*

The proof of Theorem 4.2 is based on sharp bounds of the characters. In fact, one can reduce Conjecture 4.1 to the following innocent looking problem:

**Conjecture 4.3** *Let  $\mu \vdash n$  such that  $\mu'_1 = \mu'_2$ , i.e.  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_l \geq 2$ . Then for every  $\lambda \vdash n$  we have:*

$$|\chi_\lambda([\mu])| \leq cn^{1/4} \sqrt{d_\lambda},$$

where  $c$  is a universal constant.

The conjecture is known for  $\mu = [r^{n/r}]$  and  $r \geq 2$  [FL,L], but not in the general case. We believe that it is true also for skew Young diagrams  $\lambda$  (with  $\chi_\lambda([\mu])$  defined appropriately [M,JK]). Trying to generalize Theorem 2.1 for other conjugacy classes, it is natural to assume that the upper bound in the theorem holds for all  $\mu = [1^{m_1} 2^{m_2} \dots]$  with  $m_1 = o(n)$ . We believe that this is true indeed. This would imply an  $O(1)$  bound for the mixing time when  $m_1 = 0$  (i.e.  $[\mu]$  has no fixed points), and would have a number of other interesting consequences. Unfortunately this problem is beyond the reach of our technique at this point.

Let us add that the problem of finding mixing times of random walks generated by “large” conjugacy classes, not covered by [R], has been of interest even before [R]. Let us mention here an early paper [S] which investigates the case of long cycles. Also, recently the analogues of the main theorem in [R] have appeared for simple groups of Lie type [LS]. Let us conclude with another somewhat speculative conjecture:

**Conjecture 4.4** *Let  $\Gamma$  be an edge transitive graph with diameter  $\Delta$ . Then  $\text{mix}(\Gamma) = O(\Delta^2 \log \log |\Gamma|)$ .*

The bound  $\text{mix}(\Gamma) = O(\Delta^2 \log |\Gamma|)$  is known and follows from (\*\*) in the introduction (see e.g. [AF]). All known results seem to support this conjecture. We believe that one can attempt to prove it for all distance transitive graphs  $\Gamma$ .

## 5. PROOF OF THEOREM 2.1.

In order to estimate the number of steps  $\kappa$  sufficient to drive the variation distance  $\|Q^\kappa - U\|$  to zero we apply the Upper Bound Lemma and a simple argument allowing to do calculations in  $S_n$  instead of  $A_n$  (see [L]), where

$$(\diamond) \quad \|Q_{[r1^{n-r}]}^\kappa - U\|^2 \leq \sum_{\rho \in \tilde{A}_n, \rho \neq [n]} d_\rho^2 \psi_\rho^{2\kappa}([r1^{n-r}]) \leq 2 \sum_{\substack{\lambda \vdash n \\ \lambda \neq [n], [1^n]}} f_\lambda^2 \phi_\lambda^{2\kappa}([r1^{n-r}]),$$

where  $\psi_\rho(\cdot)$ ,  $\phi_\rho(\cdot)$  denote the reduced characters in  $A_n$  and  $S_n$ , respectively.

Estimating characters will rely on the following combinatorial lemma (proved in section 6).

**Lemma 5.1** *For any Young diagram  $\lambda$ , with  $\lambda \vdash n$  there exists at most one way to remove a rim hook of length  $r > \frac{n}{2}$ .*

After removing this unique  $r$ -rim hook from  $\lambda$  we obtain a diagram denoted  $\tilde{\lambda}$ , such that  $|\chi_\lambda([r1^{n-r}])| = f_{\tilde{\lambda}}$ . If there is no  $r$ -rim hook to remove from  $\lambda$ , then  $|\chi_\lambda([r1^{n-r}])| = 0$ . To be able to use the above character-via-dimension expression in this case also, we simply assume  $f_{\tilde{\lambda}} = 0$ , where the diagram  $\tilde{\lambda}$  is not defined.

Introduce a parameter  $t$  to denote the vertical coordinate of the leftmost node of the unique  $r$ -rim hook removed from  $\lambda$ ,  $1 \leq t \leq \tilde{\lambda}'_1 + r$ . The pair  $(\tilde{\lambda}, r)$  completely defines  $\lambda = \tilde{\lambda} \cup R_t$ , where  $R_t$  is the  $r$ -rim hook with the tail at the level  $t$  (counting down from the first row of  $\tilde{\lambda}$ ).

The right-hand side of  $(\diamond)$  may be rewritten as follows (the asterisk indicates that the sum omits the terms for trivial characters)

$$\begin{aligned} \sum_{\lambda \vdash n}^* f_\lambda^2 \phi_\lambda^{2\kappa}([r1^{n-r}]) &= \sum_{\lambda \vdash n}^* f_\lambda^2 \left( \frac{\chi_\lambda([r1^{n-r}])}{f_\lambda} \right)^{2\kappa} = \sum_{\lambda \vdash n}^* f_\lambda^2 \left( \frac{f_{\tilde{\lambda}}}{f_\lambda} \right)^{2\kappa} = \\ \sum_{(\tilde{\lambda}, t)}^* f_\lambda^2 \left( \frac{f_{\tilde{\lambda}}}{f_\lambda} \right)^{2\kappa-2} &= \sum_{\tilde{\lambda} \vdash n-r} \sum_{t=1}^{\tilde{\lambda}'_1+r} f_\lambda^2 \left( \frac{f_{\tilde{\lambda}}}{f_\lambda} \right)^{2\kappa-2} \leq \sum_{t=1}^n \sum_{\tilde{\lambda} \vdash n-r} f_\lambda^2 \left( \frac{f_{\tilde{\lambda}}}{f_\lambda} \right)^{2\kappa-2}. \end{aligned}$$

In the last sum the terms with  $t > \tilde{\lambda}'_1 + r$  appear, which do not correspond to any rim hooks and should be simply disregarded.

To estimate the resulting sum we bound from above the values of  $f_{\tilde{\lambda}}$  and the ratio  $\frac{f_{\tilde{\lambda}}}{f_\lambda}$ . We use the following two lemmas.

**Lemma 5.2** *For any  $\mu \vdash m$ , we have  $f_\mu \leq \frac{m!}{\mu_1! \sqrt{(m-\mu_1)!}}$ .*



**Lemma 5.3** [L] *Fix a natural number  $l \leq m$ . Consider all diagrams  $\tilde{\lambda} \vdash m$  with  $\tilde{\lambda}_1 = l$  and all possible  $r$ -rim hooks  $R_t$  added to  $\tilde{\lambda}$ ,  $\tilde{\lambda} \cup R_t = \lambda \vdash n = m + r$ . Then*

1. for  $l < \frac{m}{2}$ ,  $\frac{f_{\tilde{\lambda}}}{f_{\lambda}} \leq \frac{\binom{m}{l}}{\binom{n}{l-1}}$ ,
2. for  $l < \frac{m}{2}$ ,  $\frac{f_{\tilde{\lambda}}}{f_{\lambda}} \leq \frac{\binom{m}{l}}{\binom{n}{m-l}}$ .

Lemma 5.2 is proved in section 4. Lemma 5.3 (in a slightly different form) was proved by the first author in his thesis [L]. The proof is a subtle application of the hook length formula. We refer to [L] for details.

From Lemma 5.3, the maximal values of the ratio  $\frac{f_{\tilde{\lambda}}}{f_{\lambda}}$  correspond to hook diagrams  $\tilde{\lambda} = [l \ 1^{m-l}]$  with  $r$ -rim hook  $R_t$  extending the longer of the leg or the arm of  $\tilde{\lambda}$ :  $t = 1$  or  $t = m - l + r$ .

Using the inequalities in Lemmas, we can estimate the right-hand side of the Upper Bound Lemma:

$$\begin{aligned} \sum_{t=1}^n \sum_{\tilde{\lambda} \vdash n-r} f_{\tilde{\lambda}}^2 \left( \frac{f_{\tilde{\lambda}}}{f_{\lambda}} \right)^{2\kappa-2} &\leq \sum_{t=1}^n \sum_{l=1}^m \sum_{\substack{\tilde{\lambda} \vdash m \\ \tilde{\lambda}_1=l}} \frac{m!}{l! \sqrt{(m-l)!}} \cdot \max \left( \frac{\binom{m}{l}}{\binom{n}{l-1}}, \frac{\binom{m}{l}}{\binom{n}{m-l}} \right) = \\ &\sum_{l=1}^{m/2} n \mathfrak{p}(m-l) \left( \frac{m!}{l! \sqrt{(m-l)!}} \right)^2 \left( \frac{m! (n-l)!}{n! (m-l)!} \right)^{2\kappa-2} + \\ &+ \sum_{l=m/2}^n n \mathfrak{p}(m-l) \left( \frac{m!}{l! \sqrt{(m-l)!}} \right)^2 \left( \frac{m! (l+n-m)!}{n! l!} \right)^{2\kappa-2} \leq \\ &\leq \frac{m!^{2\kappa}}{n!^{2\kappa-2}} n \mathfrak{p}(m) \sum_{l=1}^{m/2} \frac{(n-l)!^{2\kappa-2}}{(m-l)!^{2\kappa-1} l!^2} + \frac{m!^{2\kappa}}{n!^{2\kappa-2}} n \mathfrak{p}(m) \sum_{l=m/2}^n \frac{(n+l-m)!^{2\kappa-2}}{l!^{2\kappa-1} (m-l)!}. \end{aligned}$$

Denote by  $A = \sum_l A_l$  the terms in the last summation. Consider the term for  $l = \frac{m}{2}$ :

$$A_{\frac{m}{2}} = n \mathfrak{p}(m) \frac{m!^{2\kappa}}{n!^{2\kappa-2}} \frac{(n-m/2)!^{2\kappa-2}}{(m/2)^{2\kappa+1}}.$$

Compute the value of  $\kappa$  necessary for  $A_{m/2} = o(1/n)$  as  $n \rightarrow \infty$ . The calculation of other terms  $A_l$  can be treated in the same way.

Taking the logarithm of the product in  $A$  and using the Stirling asymptotic formula

$$\log x! = x \log x - x + \frac{1}{2} \log x + \frac{1}{2} \log 2\pi + O\left(\frac{1}{x}\right),$$

we obtain a sum of many terms. We treat separately the leading terms coming from  $x \log x$  and the smaller terms coming from the rest of the Stirling formula. Let's first dispose of the sum of small terms starting with the linear one,

$$(2\kappa-2) \left( -\left(n - \frac{m}{2}\right) + n \right) + 2\kappa \left( -m + \frac{m}{2} \right) + \frac{m}{2} = (2\kappa-2) \frac{m}{2} + 2\kappa \left( -\frac{m}{2} \right) + \frac{m}{2} = -\frac{m}{2}.$$

The leading terms sum is

$$\begin{aligned}
& (2\kappa-2) \left( \left( n - \frac{m}{2} \right) \log \left( n - \frac{m}{2} \right) - n \log n \right) + 2\kappa \left( m \log m - \frac{m}{2} \log \frac{m}{2} \right) - \frac{m}{2} \log \frac{m}{2} = \\
& (2\kappa-2) \left( n \log \left( 1 - \frac{m}{2n} \right) - \frac{m}{2} \log \left( n - \frac{m}{2} \right) \right) + 2\kappa \left( \frac{m}{2} \log m - \frac{m}{2} \log 2 \right) - \frac{m}{2} \log \frac{m}{2} = \\
& (2\kappa-2) \left( n \left( -\frac{m}{2n} + o\left(\frac{m}{2n}\right) \right) - \frac{m}{2} \log \left( n - \frac{m}{2} \right) \right) + 2\kappa \left( \frac{m}{2} \log m - \frac{m}{2} \log 2 \right) - \frac{m}{2} \log \frac{m}{2} = \\
& (2\kappa-2) \left( -\frac{m}{2} + o(m) \right) - \frac{m}{2} \log \left( n - \frac{m}{2} \right) + 2\kappa \left( \frac{m}{2} \log m - \frac{m}{2} \log 2 \right) - \frac{m}{2} \log \frac{m}{2} = \\
& \kappa \left( -m \log \left( \frac{n - \frac{m}{2}}{m} \right) - m + o(m) + m \log 2 \right) + m - o(m) + m \log \left( \frac{n - \frac{m}{2}}{m} \right) + \frac{m}{2} \log m = \\
& \kappa \left( -m \log \left( \frac{n}{m} - \frac{1}{2} \right) + O(m) \right) + m \log \left( \frac{n}{m} - \frac{1}{2} \right) + \frac{m}{2} \log m + O(m) .
\end{aligned}$$

This quantity tends to minus infinity whenever

$$\begin{aligned}
\kappa & > \frac{m \log \left( \frac{n}{m} - \frac{1}{2} \right) + \frac{m}{2} \log m + O(m)}{m \log \left( \frac{n}{m} - \frac{1}{2} \right) + O(m)} = 1 + \frac{\log m}{2 \log \left( \frac{n}{m} - \frac{1}{2} \right)} \\
& > \frac{\log m}{2 \log(2n - m) - 2 \log 2m} = \frac{1}{2} \frac{\log n}{\log n - \log m + O(1)} =: K(m, n).
\end{aligned}$$

Now take  $\kappa > (1 + \delta)K(m, n)$ . From above and  $\log p(m) = O(\sqrt{m})$ , we obtain  $A_{m/2} = \exp O(-\sqrt{n})$ . This is stronger than the desired  $A_{m/2} = o(1/n)$  bound. For the other terms we obtain similar bounds verbatim. Namely, the constant  $1/2$  is replaced by  $(1 - \beta)$  for the term  $A_l$ ,  $l = m(1 - \beta)$ . Therefore  $\kappa > 2(1 + \delta)K(m, n)$  will always suffice for  $A_l = o(1/n)$ . This implies Theorem 2.1  $\square$

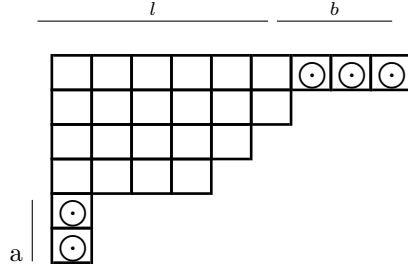
## 6. PROOFS OF LEMMAS.

We start with a definition used in [L].

**Definition 6.1** A diagram  $\lambda$  is called *thick*, if first two rows of  $\lambda$  are of equal length and first two columns of  $\lambda$  are of equal length:  $\lambda_1 = \lambda_2$ ,  $\lambda'_1 = \lambda'_2$ .

For an arbitrary non-hook diagram  $\lambda$  define its *thick core*  $\widehat{\lambda}$  as a thick subdiagram of  $\lambda$  with first row and first column cut down to the size of the second row and column respectively:  $\widehat{\lambda}_1 = \lambda_2$ ,  $\widehat{\lambda}'_1 = \lambda'_2$ . All other rows and columns are of the same length as in  $\lambda$ . Define the thick core of a hook diagram to be empty. Denote  $a = \lambda'_1 - \lambda'_2$ ,  $b = \lambda_1 - \lambda_2$ . Any diagram  $\lambda \vdash n$  is uniquely defined by a triple:  $\lambda = (a, \widehat{\lambda}, b)$ . We sometimes refer to  $a$  and  $b$  as the “thin” parts of diagram  $\lambda$ .

**Example 6.2**  $[96541^2]$  is defined by  $(2, [6^254], 3)$  in this notation.



**Proof of Lemma 5.1.**

Consider the entire border of  $\lambda$ , denote its length  $\text{br}(\lambda) = \lambda_1 + \lambda'_1 - 1$ . If  $\text{br}(\lambda) < r$ , then there is no way to remove an  $r$ -rim hook. If  $\text{br}(\lambda) \geq r > \frac{n}{2}$ , then  $\lambda$  cannot be a thick diagram (since for any thick diagram  $|\lambda| \geq 2 \text{br}(\lambda)$ ) and may be (up to transposing) either  $\lambda = (0, \widehat{\lambda}, b), b \neq 0$  or  $\lambda = (a, \widehat{\lambda}, b), a, b \neq 0$ , where  $\widehat{\lambda}$  is the thick core of  $\lambda$ .

The entire border of  $\lambda$  consists of two parts: the thin part and the border of the thick core  $\widehat{\lambda}$ . Any rim hook of  $\lambda$  may not have its leftmost (rightmost) point at any point of the vertical (horizontal) thin part, but at the end one. For a rim hook of length  $k > \frac{n}{2}$  the ending points cannot both lie on the boundary of  $\widehat{\lambda}$ , since  $\text{br}(\widehat{\lambda}) < \frac{n}{2}$ , therefore it starts at the ending point of a thin part.

If  $\lambda$  has only one thin part, i.e.  $\lambda = (0, \widehat{\lambda}, b)$ , this uniquely determines the rim hook. In case  $\lambda = (a, \widehat{\lambda}, b)$ , suppose we have two rim hooks of length  $r$ , one starting at the end of the vertical thin part, another starting at the end of the horizontal thin part. By the above both end at some nodes on the boundary of  $\widehat{\lambda}$ , which has length  $\text{br}(\widehat{\lambda}) < \frac{n}{2}$ . Hence,  $\text{br}(\lambda) \geq 2r - \text{br}(\widehat{\lambda}) > n - \text{br}(\widehat{\lambda})$ , therefore  $\text{br}(\lambda) + \text{br}(\widehat{\lambda}) > n \Rightarrow \text{br}(\lambda) + |\lambda \setminus \text{br}(\lambda)| > n$ . Contradiction.  $\square$

**Proof of Lemma 5.2.**

Denote by  $\nu$  the subdiagram  $[\mu_2, \mu_3, \dots, \mu_k] \subset \mu$ , i.e. the diagram  $\mu$  with the first row  $[\mu_1]$  removed. Obviously, hook lengths with vertices inside  $\nu$  are the same for  $\mu$  and  $\nu$ :  $h_\nu(t) = h_\mu(t), t \in \nu$ . By the hook length formula (see e.g. [M,JK]) for the dimension of an irreducible  $S_n$ -character

$$f_\nu = \frac{(m - \mu_1)!}{\prod_{t \in \nu} h_\nu(t)}.$$

By the Burnside lemma [JK], we have

$$\sum_{\xi \vdash m - \mu_1} f_\xi^2 = (m - \mu_1)!.$$

Hence  $f_\nu^2 < (m - \mu_1)!$ , i.e.

$$\frac{(m - \mu_1)!^2}{\prod_{t \in \nu} h_\nu(t)^2} < (m - \mu_1)!.$$

Therefore

$$\frac{1}{\prod_{t \in \nu} h_\nu(t)^2} < \frac{1}{\sqrt{(m - \mu_1)!}}.$$

From the hook length formula for  $\mu$  and  $\prod_{t \in [\mu_1]} h_\mu(t) > \mu_1!$ , we obtain:

$$f_\mu = \frac{m!}{\prod_{t \in \mu} h_\mu(t)} = \frac{m!}{\prod_{t \in [\mu_1]} h_\mu(t) \prod_{t \in \nu} h_\nu(t)} < \frac{m!}{\mu_1! \sqrt{(m - \mu_1)!}}. \quad \square$$

## 7. PROOF OF THEOREM 2.3

As before, let  $\lambda$  be a partition of  $n$ , and let  $f_\lambda$  denote the dimension of the irreducible representation  $\pi_\lambda$  of the symmetric group  $S_n$ . By Burnside's identity,  $f_\lambda = d_\lambda < \sqrt{n!}$ . By  $\chi_\lambda(\mu)$  we denote value of the character of  $\pi_\lambda$  on a conjugacy class with cycle structure  $(\mu)$ .

By the Upper Bound Lemma 3.1, we obtain

$$\mathbf{d}(2) \leq \frac{1}{4} \sum_{\lambda: |\lambda|=n, \lambda \neq (n)} f_\lambda^2 \left( p + \sum_{l=2}^n p_l \frac{\chi_\lambda(l, 1^{n-l})}{f_\lambda} \right)^{2m} = (\diamond)$$

where  $p$  is the holding probability and

$$p_l = (1-p) \frac{\binom{n}{l} (l-1)!}{D_n} \leq \frac{Cn}{l(n-l)!},$$

$D_n$  is the total number of cycles, and  $C$  is a universal constant.

To estimate the right hand side of  $(\diamond)$  we break the inner summation inside into two parts :  $l \leq n/2$  and  $l > n/2$ . In the first case we simply bound each ratio by 1. In the second case, the value of the character turns out to be 0 except when Lemma 5.1 applies. This will simplify the problem and eventually give us the desired bound.

Indeed, the proof of Lemma 5.1 and the Murnaghan-Nakayama rule immediately imply that if  $l > n/2$ , we have  $|\chi_\lambda(k1^{n-k})| = f_{\tilde{\lambda}}$  if  $k = \lambda_1 + l(\lambda) - 1$ , and  $|\chi_\lambda(k1^{n-k})| = 0$  otherwise. We have:

$$(\diamond) \leq \sum_{k=1}^n \sum_{\lambda \neq (n), l+\lambda_1-1=k} f_\lambda^2 \left( p_k \frac{f_{\tilde{\lambda}}}{f_\lambda} + \sum_{j=2}^{\lceil n/2 \rceil} p_j \right)^{2m}$$

Now let  $m = 2$ . Using the formula for  $p_k$ , and  $\sum_{j=2}^{\lceil n/2 \rceil} p_j \leq \frac{Cn^2}{\lceil n/2 \rceil!}$ , we have

$$(\diamond) \leq \sum_{k=1}^n \sum_{\lambda \neq (n), l+\lambda_1-1=k} f_\lambda^2 \left( \frac{Cn}{k(n-k)!} \cdot \frac{\sqrt{(n-k)!}}{f_\lambda} + \frac{Cn^2}{\lceil n/2 \rceil!} \right)^4$$

$$\leq \sum_{k=1}^n \sum_{\lambda \neq (n), l+\lambda_1-1=k} \left( \frac{Cn}{k\sqrt{(n-k)!}\sqrt{f_\lambda}} + \frac{Cn^2\sqrt{f_\lambda}}{\lceil n/2 \rceil!} \right)^4,$$

where the last inequality follows from  $f_\lambda < \sqrt{(n-k)!}$ . Recall that  $(a+b)^4 \leq 8(a^4+b^4)$ . This breaks each summand into two, which breaks our double sum into two double sums. The second double sum goes to 0 as  $n \rightarrow \infty$ , which easily follows from  $\sqrt{f_\lambda} < (n!)^{1/4}$ , Stirling's formula, and since the total number of partitions  $p(n) = \exp O(\sqrt{n})$ .

For the first double sum, it was shown by the first author in [L] that

$$\sum_{\lambda \neq (n), (1^n); |\lambda|=n} \frac{1}{f_\lambda} \rightarrow 0 \text{ as } n \rightarrow \infty$$

It is not hard to see that this implies that the first double sum also goes to 0 as  $n \rightarrow \infty$ . This completes the proof.  $\square$

#### REFERENCES

- [AD] D. Aldous, P. Diaconis, *Strong uniform times and finite random walks*, Advances in Applied Math. **8** (1987), 69–97.
- [AF] D. Aldous, J. Fill, *Reversible Markov Chains and Random Walks on Graphs*, monograph in preparation, 1996.
- [An] G. Andrews, *The Theory of Partitions*, Addison-Wesley, New York, 1976.
- [D1] P. Diaconis, *Group Representations in Probability and Statistics*, IMS, Hayward, California, 1988.
- [D2] P. Diaconis, *The cutoff phenomenon in finite Markov chains*, Proc. Nat. Acad. Sci. U.S.A. **93** (1996), 1659–1664.
- [DS] P. Diaconis, M. Shahshahani, *Generating a random permutation with random transpositions*, Z. Wahr. verw. Gebiete **57** (1981), 159–179.
- [Dv] Y. Dvir, *Covering properties of permutation groups*, Lecture Notes in Math. **1112**, Springer-Verlag, 1985.
- [FL] S. Fomin, N. Lulov, *On the number of rim hook tableaux*, Zap. Nauchn. Sem. POMI **223** (1995), 219–226 (available also from <http://www.math.lsa.umich.edu/~fomin/Papers/>).
- [JK] G. James, A. Kerber, *The representation theory of the symmetric group*, *Encyclopedia of Mathematics and its Applications*, Vol. 16, G.-C. Rota, Ed., Addison-Wesley, Reading, Mass., 1981.
- [LS] M. Liebeck, A. Shalev, *Diameters of finite simple groups: sharp bounds and applications*, Ann. of Math. (2) **154** (2001), 383–406.
- [LW] L. Lovász, P. Winkler, *Mixing Times* (1998), AMS DIMACS Series, vol. 41, 189–204.
- [Lb] A. Lubotzky, *Cayley graphs: eigenvalues, expanders and random walks*, LMS Lecture Note Ser., 218, Cambridge Univ. Press, Cambridge, 1995.
- [L] N. Lulov, *Random Walks on the Symmetric Group Generated by Conjugacy Classes*, Ph.D. Thesis, Harvard University, 1996.
- [M] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, Oxford University Press, London, 1979.
- [P] I. Pak, *Random walks on permutation: strong uniform time approach*, Ph.D. Thesis, Harvard U., 1997.
- [PV] I. Pak, V. H. Vu, *On mixing of certain random walks, cutoff phenomenon and sharp threshold of random matroid processes*, Discrete Appl. Math., **110** (2001), 251–272.
- [R] Y. Roichman, *Upper bound on characters of the symmetric groups*, Invent. Math. **125** (1996), 451–485.
- [S] R. P. Stanley, *Factorization of permutations into  $n$ -cycles*, Discrete Math. **37** (1981), 255–262.