THE PROBABILITY OF LONG CYCLES IN INTERCHANGE PROCESSES

GIL ALON AND GADY KOZMA

1. Introduction

A well known phenomenon in the theory of mixing times[1] is that occasionally certain aspects of a system mix much faster than the system as a whole. Pemantle [12] constructed an example of a random walk on the symmetric group S_n which mixes in time $n^{1+o(1)}$ while every k elements mix in $\le C(k)\sqrt{n}$ time. Schramm [13] showed that for the interchange process on the complete graph — this is another random walk on S_n, see below for details — the structure of the large cycles mixes in time $\approx n$, and it was known before [4] that the mixing time of this graph is $\approx n\log n$. Schramm's result is related to — in physics' parlance, it is the *mean-field* case of — a conjecture of Bálint Tóth [14] that the cycle structure of the interchange process on the graph $\mathbb{Z}^d$, $d \ge 3$, exhibits a *phase-transition*. In this paper we investigate the probability of long cycles, and obtain precise formulae for any graph, using the representation theory of S_n. As an application, we analyse certain variations on Tóth's conjecture.

Let us define the interchange process. Let G be a finite graph with vertex set $\{1,\dots,n\}$, and equip each edge $\{i,j\}$ with an alarm clock that rings with exponential rate $a_{i,j}$. Put a marble on every vertex of G, all different, and whenever the clock of $\{i,j\}$ rings, exchange the two marbles. Each marble therefore does a standard continuous-time random walk on the graph but the different walks are dependent. The positions of the marbles at time t is a permutation of their original positions, and viewed this way the process is a random walk on the symmetric group. Note that we have changed the timing from the previous paragraph. For example, if our graph is the complete graph and $a_{i,j} = 1$ for all i and j, then the process mixes in time $\approx \log n$ and large cycle structure mixes in time ≈ 1. However, the added convenience of having each marble do the natural continuous time random walk outweighs the difference in notations from some of the literature.

The stronger results of this paper require representation theory to state, but let us start with two corollaries that can be stated elementarily. Let $s_k(t)$ be the number of cycles of length k in our permutation at time t. Let $0 = \lambda_0 \le \lambda_1 \le \cdots \le \lambda_{n-1}$ be the eigenvalues of the continuous time Laplacian of the random walk on the graph G. Then

Theorem 1. *We have*

$$\mathbb{P}(s_n(t) = 1) = \frac{1}{n}\prod_{i=1}^{n-1}(1 - e^{-\lambda_i t})$$

[1]We do not need the notion of mixing time in this paper, it is only used for comparison. The reader unfamiliar with it may peruse the survey [9] or the book [7]. Other notions not explicitly defined (say the symmetric group or a class function) may be found in Wikipedia.

Let us demonstrate the utility of this formula on the graph $G = \{0,1\}^d$ with weights equal to 1. There is nothing particular about this graph, but existing literature allows for easy comparison. For example, Wilson [15, §9] showed that the mixing time of the interchange process on G is $\geq cd$ (see also [10, 11]). The eigenvalues of G may be calculated explicitly: the eigenvectors are the Walsh functions, indexed by $\{0,1\}^d$ and given by $f_y(x) = (-1)^{\sum_{i=1}^d x_i y_i}$. We get that $2k$ is an eigenvalue with multiplicity $\binom{d}{k}$ for $k = 0, \ldots, d$. Inserting into the formula at times $\frac{1\pm\epsilon}{2}\log d$ gives

$$\mathbb{P}\Big(s_n\Big(\frac{1-\epsilon}{2}\log d\Big) = 1\Big) = 2^{-d}\prod_{k=1}^{d}\Big(1 - e^{-(1-\epsilon)k\log d}\Big)^{\binom{d}{k}} \leq$$

and looking only at $k = K := \lfloor d^\epsilon/2 \rfloor$,

$$\leq \exp\left(-d^{(\epsilon-1)K}\binom{d}{K}\right) \stackrel{(*)}{\leq} \exp\left(-\Big(\frac{d^\epsilon}{K}\Big)^K\right) \leq \exp\left(-\exp\left(cd^\epsilon\right)\right)$$

where $(*)$ comes from

$$\binom{d}{K} = \Big(\frac{d}{K}\Big)^K \cdot \Big(\frac{1-1/d}{1-1/K}\cdot\frac{1-2/d}{1-2/K}\cdot\cdots\Big) \geq \Big(\frac{d}{K}\Big)^K.$$

On the other hand,

$$\mathbb{P}\Big(s_n\Big(\frac{1+\epsilon}{2}\log d\Big) = 1\Big) = 2^{-d}\exp\left(\sum_{k=1}^{d} O(d^{-(1+\epsilon)k})\binom{d}{k}\right) = 2^{-d}(1 + O(d^{-\epsilon})).$$

We see that the probability equilibrates at $\frac{1}{2}\log d$, before the mixing time of the whole chain. Further, the equilibration happens sharply — this is reminiscent of the cutoff phenomenon for mixing times. See [4], [8] or [7, §18] for the cutoff phenomenon.

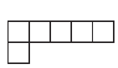
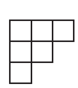
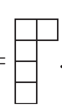
Another general, elementarily stated result is:

Theorem 2. *We have, for any graph G and any $1 \leq k \leq n$,*

$$\left|\mathbb{E}(s_k(t)) - \frac{1}{k}\right| \leq \frac{3^n}{k}e^{-t\lambda_1}$$

The point about this result is its generality — it holds for any graph. In particular examples that we tried the estimate was worse than the known or conjectured mixing time. But for general graphs it seems to be the best known.

To proceed, let us recall a few basic facts about the representations of S_n. For a full treatment see the book [6]. A representation of S_n is a group homomorphism $\tau : S_n \to \mathrm{GL}_k(\mathbb{C})$ for some k, typically denoted by $\dim\tau$. Its character, denoted by χ_τ, is an element of $L^2(S_n)$ defined by $\chi_\tau(g) = \mathrm{tr}(\tau(g))$. Now, the irreducible representations of S_n are indexed by partitions of n, namely, by sequences $\lambda = [\lambda_1, \lambda_2, \ldots, \lambda_k]$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0$ and $\sum_{i=1}^k \lambda_i = n$ (we denote this by $n \vdash \lambda$). A nice graphical representation of partitions is using *Young diagrams*, i.e., drawing each λ_i as a line of boxes from top to bottom, e.g.

$[5,1] =$  $\quad [3,2,1] =$  $\quad [2,1^3] =$ .

To each partition $n \vdash \lambda$ (and hence, for each young diagram with n boxes) corresponds an irreducible representation, which we shall denote by U_λ. For brevity, we denote the character of U_λ by χ_λ. Fix now some $1 \le k \le n$ and define

$$\alpha_k(g) = \#\{\text{cycles of length } k \text{ in } g\}. \tag{1}$$

Now, $\alpha_k(g)$ depends only on the cycle structure of g, i.e. is a class function, and hence it is a linear combinations of characters of irreducible representations. Our main result is the precise decomposition.

Theorem 3. *For any n and k,*

$$\alpha_k = \frac{1}{k}\sum_{n \vdash \rho} a_\rho \chi_\rho,$$

where

$$a_\rho = \begin{cases} 1 & \rho = [n] \\ (-1)^{i+1} & \rho = [k-i-1, n-k+1, 1^i] \text{ for some } i \in \{0, \dots, 2k-n-2\} \\ (-1)^i & \rho = [n-k, k-i, 1^i] \text{ for some } i \in \{\max\{2k-n, 0\}, \dots, k-1\} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Let us describe this verbally (ignoring the diagram $[n]$ which has a somewhat special role). If $k > (n+1)/2$, start with $[k-1, n-k+1]$, with a minus sign. Now drop boxes from the first row into the leftmost column until the first and second row are equal. Then drop in a single step two boxes, one from each of the first two rows to the leftmost column. Then start dropping boxes from the second row until you reached a hook-shaped diagram. The sign keeps changing in each step. If $k \le n/2$ start with the diagram $[n-k, k]$ with a plus sign, and drop boxes from the second row to the leftmost column until reaching a hook-shaped diagram, again switching sign at each step. The case $k = (n+1)/2$ is similar except you start from $[n-k, k-1, 1]$ with a minus sign.

It is now clear what is special in the case $k = n$. In this case only hook-shaped diagrams appear in the sum. For the hook-shaped diagrams there is an explicit formula for the relevant eigenvalues discovered by Bacher [2] (see also the appendix of [1]). We will explain this (i.e. the conclusion of theorem 1 from theorem 3) in section 4 below.

Let us note an important feature of the formula (2) related to Tóth's conjecture. A natural variation on Tóth's conjecture would be to assume that for the graph $\{1, \dots, m\}^d$, $d \ge 3$ one has that $\mathbb{E}(s_k(t))$ equilibrates at constant time, for $k \approx n$. The eigenvalues of the graph are all known so we can use theorem 1 and get with a simple calculation that in the case $k = n$ this probability only equilibrates in time $\approx m^2 = n^{2/d}$. The culprit for this "slow" equilibration lies in the representation $[n-1, 1]$ appearing in the sum when $k = n$. It is therefore reassuring to notice that this representation appears *only* when $k = n$. Hence we have a good basis to conjecture that in fact $\mathbb{E}(s_k(t))$ equilibrates in constant time for all $\epsilon n < k \le (1-\epsilon)n$ and that this may be demonstrated using theorem 3.

2. Notations and preliminaries

Let $A = \{a_{i,j}\}_{1 \le i < j \le n}$ be a collection of non-negative numbers which we consider as a weighted graph. The random walk on S_n associated with the weighted graph A is a process in continuous time starting from the identity permutation $\mathbf{1}$ on S_n and going from g to $(ij)g$ with rate $a_{i,j}$. Formally, consider $L^2(S_n)$, both as a hilbert space with the standard inner product, and as a ring, via the *group ring* structure. Define the Laplacian as the element of $L^2(S_n)$ given by

$$\Delta = \Delta_A = \sum_{i<j} a_{i,j}(\mathbf{1} - (ij))$$

where $\mathbf{1}$ is the element of $L^2(S_n)$ equal to 1 in the identity permutation, and 0 everywhere else; and (ij) is similarly a singleton at the transposition (ij). The distribution of the location of our process at time t is

$$e^{-t\Delta} = \sum_{k=0}^{\infty} \frac{(-t\Delta)^k}{k!}$$

In particular for α_k defined by (1),

$$\mathbb{E}(s_k(t)) = \sum_{g\in S_n} \left(e^{-t\Delta}\right)(g)\alpha_k(g) = n!\langle e^{-t\Delta}, \alpha_k\rangle$$

where here and below $\langle\cdot,\cdot\rangle$ stands for the standard inner product in $L^2(S_n)$ i.e. $\langle a, b\rangle = {}^1\!/_{n!}\sum_{g\in S_n} a(g)\overline{b(g)}$.

For the proof of theorem 3 we will need a second set of representations of S_n, this time *reducible* representations. For $n \vdash \rho$, let $T_\rho < S_n$ be the subgroup of all permutations fixing the sets $\{1, \dots, \rho_1\}$, $\{\rho_1 + 1, \dots, \rho_1 + \rho_2\}$, etc. As a group $T_\rho \cong S_{\rho_1} \times \cdots \times S_{\rho_r}$. Now, S_n acts on the left cosets of T_ρ, and using these cosets as a basis we obtain a representation of S_n, which we will denote by V_ρ. Readers familiar with exclusion processes might find it convenient to think about V_ρ as the space of distributions of the exclusion process with ρ_1 particles of color 1, ρ_2 particles of color 2 etc. — considering Δ as an operator on V_ρ it is easy to verify that one gets an identical process. We denote

$$\psi_\rho = \chi_{V_\rho}. \tag{3}$$

It is well known that the representations V_ρ are generally reducible and their irreducible components, consist of all U_σ for $\sigma \trianglerighteq \rho$, where $\trianglerighteq$ is the *domination* order [5, Corollary 4.39] — we say that $\sigma \trianglerighteq \rho$ when you can reach ρ from σ by a series of "toppling" of a box of the Young diagram to a lower row which keep the structure of a Young diagram. Alternatively, $\sigma \trianglerighteq \rho$ is equivalent to

$$\sum_{i=1}^{j} \sigma_i \geq \sum_{i=1}^{j} \rho_i \qquad \forall j.$$

3. Character decomposition

In this section we prove theorem 3. Recall the definition of α_k, (1). Since α_k is a class function, it can be expressed as a linear combination of the characters of S_n (see, e.g., [5, Proposition 2.30]). By the character orthogonality relations (ibid.), we have

$$\alpha_k = \sum_{n\vdash\rho} \langle\alpha_k, \chi_\rho\rangle\chi_\rho \tag{4}$$

In order to prove theorem 3, we need to show that

$$\forall n \vdash \rho, \langle \alpha_k, \chi_\rho \rangle = \frac{1}{k} a_\rho$$

where a_ρ were defined in (2).

Our first step towards this goal will be to calculate the inner products $\langle \alpha_k, \psi_\rho \rangle$, where $\psi_\rho = \chi_{V_\rho}$ and V_λ are the "exclusion-like" reducible representations defined just before (3).

Let us define a function β_k on the set of partitions of n by

$$\beta_k([\lambda_1, \dots, \lambda_r]) = \#\{i : \lambda_i \geq k\}$$

Lemma 1. *We have for all* $n \vdash \rho$, $\langle \alpha_k, \psi_\rho \rangle = \frac{1}{k} \beta_k(\rho)$.

Proof. We have

$$\langle \alpha_k, \psi_\rho \rangle = \frac{1}{n!} \sum_{g \in S_n} \psi_\rho(g) \alpha_k(g)$$

Recall from §2 that V_ρ is obtained from the action of S_n on the cosets of a $T_\rho < S_n$. By the definition of trace, $\psi_\rho(g)$ equals the number of cosets of T_ρ fixed by g. A coset hT_ρ is fixed by g iff $h^{-1}gh \in T_\rho$. Hence,

$$\langle \alpha_k, \psi_\rho \rangle = \frac{1}{n!} \sum_{hT_\rho \in S_n/T_\rho} \sum_{g : h^{-1}gh \in T_\rho} \alpha_k(g).$$

Let us make a change of variables, $q = h^{-1}gh$. Since α_k is a class function, we have

$$\langle \alpha_k, \psi_\rho \rangle = \frac{1}{n! \#(T_\rho)} \sum_{h \in G} \sum_{q \in T_\rho} \alpha_k(hqh^{-1}) = \frac{1}{n! \#(T_\rho)} \sum_{h \in G} \sum_{q \in T_\rho} \alpha_k(q) = \frac{1}{\#(T_\rho)} \sum_{q \in T_\rho} \alpha_k(q).$$

The sum $\sum_{q \in T_\rho} \alpha_k(q)$ can be evaluated by summing over all possible k-cycles $c \in T_\rho$, the number of elements of T_ρ such that c is one of their cycles. For any i such that $\rho_i \geq k$, there are $\binom{\rho_i}{k} \cdot (k-1)!$ choices for a cycle c in the S_{ρ_i}-factor of T_ρ, and $\rho_1! \rho_2! \cdot (\rho_i - k)! \cdot \rho_r!$ choices for an element $g \in T_\rho$ with c as a cycle. Hence each such i contributes to the sum

$$\binom{\rho_i}{k} \cdot (k-1)! \cdot \rho_1! \rho_2! \cdot (\rho_i - k)! \cdot \rho_r! = \frac{\#(T_\rho)}{k}$$

Obviously, if $\rho_i < k$ then there are no k-cycles in the S_{ρ_i}-factor, and the contribution is 0. Hence,

$$\langle \alpha_k, \psi_\rho \rangle = \frac{1}{\#(T_\rho)} \sum_{i : \rho_i \geq k} \frac{\#(T_\rho)}{k} = \frac{1}{k} \beta_k(\rho). \qquad \square$$

Now, recall that by Young's rule [5, Corollary 4.39], the characters ψ_ρ and the characters χ_ρ are related by the linear equations

$$\psi_\rho = \sum_{n \vdash \sigma} K_{\sigma\rho} \chi_\sigma$$

Where the numbers $K_{\sigma\rho}$, called the Kostka numbers, are defined as follows: Let $\rho = [\rho_1, \dots, \rho_r]$, then $K_{\sigma\rho}$ is the number of ways the Young diagram of σ can be filled with ρ_1 1's, ρ_2 2's, etc., such that each row is nondecreasing, and each column is strictly increasing. The numbers $K_{\sigma\rho}$ satisfy $K_{\sigma\rho}=0$ whenever $\sigma<\rho$ (with respect to the lexicographic order), and $K_{\sigma\sigma}=1$. (See [5], appendix A). Hence,

$$\frac{1}{k}\beta_k(\rho)=\langle\alpha_k,\psi_\rho\rangle=\sum_{n\vdash\sigma}K_{\sigma\rho}\langle\alpha_k,\chi_\sigma\rangle$$

In other words the numbers $\langle\alpha_k,\chi_\sigma\rangle$ satisfy a system of linear equations. Further, the matrix $(K_{\sigma\rho})$ is triangular with 1's on the diagonal, hence invertible. It follows that it is enough to show that the numbers $\frac{1}{k}a_\rho$ satisfy the same equations, i.e.

$$\sum_{n\vdash\sigma}K_{\sigma\rho}a_\sigma=\beta_k(\rho)\qquad\forall n\vdash\rho \tag{5}$$

(with a_σ defined by (2)) to demonstrate theorem 3. Let us start with the case of $k=n$.

Lemma 2. *For all* $n\vdash\rho$, $\sum_{i=0}^{n-1}(-1)^iK_{[n-i,1^i]\rho}=\beta_n(\rho)$

Proof. Let ρ have r rows. We have $K_{[n-i,1^i]\rho}=\binom{r-1}{i}$, and $K_{[n-i,1^i]\rho}=0$ for $i\geq r$, since the top left box of $[n-i,1^i]$ has to be numbered 1, and the whole configuration is determined by the choice of distinct i numbers out of $2,\dots,r$ to be placed in the leftmost column in ascending order. The result follows by the binomial identity. □

We immediately conclude:

Corollary 1. We have

$$\alpha_n=\frac{1}{n}\sum_{i=0}^{n-1}(-1)^i\chi_{[n-i,1^i]}$$

Which is theorem 3 for $k=n$.

Let us now treat the general case. It is comfortable to treat variables indexed by partitions in the realm of *symmetric polynomials*. Fix an integer $m\geq n$ (whose value is not important), and consider the ring of symmetric polynomials in m variables $x_1,\dots,x_m$ over $\mathbb{C}$.

Consider the following homogeneous symmetric polynomials of degree n (see [5], ibid. for more details):

- For $n\vdash\lambda=[\lambda_1,\dots,\lambda_r]$, $M_\lambda=\sum_\alpha x^\alpha$, where $\alpha=(\alpha_1,\dots,\alpha_n)$ goes over all the possible permutations of $(\lambda_1,\dots\lambda_r,0,\dots,0)$.
- The Schur polynomials $S_\mu=\sum_\lambda K_{\mu\lambda}M_\lambda$
- The full homogeneous polynomial H_n, defined as the sum of all monomials of degree n. It is easy to see that for all $n\vdash\lambda$, $K_{[n]\lambda}=1$. Hence, $H_n=\sum_{n\vdash\lambda}M_\lambda=\sum_{n\vdash\lambda}K_{[n]\lambda}M_\lambda=S_{[n]}$.

We can now restate the system of linear equations (5) as a single linear equation whose coefficients are polynomials,

$$\sum_\mu a_\mu S_\mu(x)=\sum_\lambda\beta_k(\lambda)M_\lambda(x).$$

A moment's reflection shows that the right hand side is equal to:

$$\sum_\lambda \beta_k(\lambda)M_\lambda(x) = \left(\sum_i x_i^k\right) H_{n-k}.$$

According to lemma 2 (applied to the case of diagrams of size k),

$$\sum_i x_i^k = \sum_{i=0}^{k-1}(-1)^i S_{[k-i,1^i]}.$$

We conclude that (5) is equivalent to

$$\sum_\mu a_\mu S_\mu(x) = \left(\sum_{i=0}^{k-1}(-1)^i S_{[k-i,1^i]}\right) H_{n-k}.$$

We now apply Pieri's formula (see ibid.), according to which, $S_{[k-i,1^i]}H_{n-k}(x)$ is the sum of all polynomials of the form $S_{\lambda'}$, where λ' is obtained by adding $n-k$ boxes to $[k-i,1^i]$, without adding two boxes in the same column. Since we have a hook-shaped diagram, our possibilities are rather limited: we may add a box at the leftmost column or not, and the rest of the boxes go in the first two rows. Denote therefore

$$S_{[k-i,1^i]}H_{n-k} = A_i + B_i$$

where A_i is the sum when one does not add a box at the leftmost column, and B_i is when one does. Denote also $x(i,j) = S_{[n-i-j,1+j,1^{i-1}]}$ (the contribution coming from adding j boxes to the second row of $[k-i,1^i]$, and the remaining $n-k-j$ boxes to the first row). Then

$$A_0 = S_{[n]}, \qquad A_i = \sum_{j=0}^{\min(n-k,k-i-1)} x(i,j)$$

and

$$B_i = \sum_{j=0}^{\min(n-k-1,k-i-1)} x(i+1,j).$$

We now sum over i and get,

$$\left(\sum_i x_i^k\right) H_{n-k} = \sum_{i=0}^{k-1}(-1)^i(A_i+B_i)$$

Our next goal is to find the alternating sum $\sum_{i=0}^{k-1}(-1)^i(A_i+B_i)$. There are further cancellations here because B_i and A_{i+1} are quite similar — B_i corresponds to adding a box to the first column of $[k-i,1^i]$ while A_{i+1} corresponds to not adding a box to the first column of $[k-i-1,1^{i+1}]$. Hence most of the terms cancel out. We get

$$A_{i+1} = \begin{cases} \sum_{j=0}^{n-k} x(i+1,j) & 0 \le i \le 2k-n-2 \\ \sum_{j=0}^{k-i-2} x(i+1,j) & 2k-n-1 \le i \le k-2 \end{cases}$$

$$B_i = \begin{cases} \sum_{j=0}^{n-k-1} x(i+1,j) & 0 \le i \le 2k-n-1 \\ \sum_{j=0}^{k-i-1} x(i+1,j) & 2k-n \le i \le k-1 \end{cases}$$

Hence (putting $A_k = 0$),

$$B_i - A_{i+1} = \begin{cases} \sum_{j=0}^{n-k-1} x(i+1,j) - \sum_{j=0}^{n-k} x(i+1,j) = -x(i+1,n-k) & 0 \le i \le 2k-n-2 \\ \sum_{j=0}^{n-k-1} x(i+1,j) - \sum_{j=0}^{n-k-1} x(i+1,j) = 0 & i = 2k-n-1 \\ \sum_{j=0}^{k-i-1} x(i+1,j) - \sum_{j=0}^{k-i-2} x(i+1,j) = x(i+1,k-i-1) & 2k-n \le i \le k-1 \end{cases}$$

and

$$\begin{aligned} \sum_{i=0}^{k-1}(-1)^i(A_i+B_i) &= A_0 + \sum_{i=0}^{k-1}(-1)^i(B_i - A_{i+1}) = \\ &= S_{[n]} - \sum_{i=0}^{2k-n-2}(-1)^i x(i+1,n-k) + \sum_{i=2k-n}^{k-1}(-1)^i x(i+1,k-i-1) = \\ &= S_{[n]} - \sum_{i=0}^{2k-n-2}(-1)^i S_{[k-i-1,n-k+1,1^i]} + \sum_{i=2k-n}^{k-1}(-1)^i S_{[n-k,k-i,1^i]} = \sum_\rho a_\rho S_\rho. \end{aligned}$$

In summary, we get

$$\sum_\lambda \beta_k(\lambda) M_\lambda(x) = \left(\sum_i x_i^k\right) H_{n-k} = \sum_{i=0}^{k-1}(-1)^i(A_i+B_i) = \sum_\rho a_\rho S_\rho$$

and (5) is proved.

This ends the proof of theorem 3. □

4. The probability of long cycles

Let ρ be a partition of n, and let $U_\rho : S_n \to \mathrm{GL}(\mathbb{C}^{\dim U_\rho})$ be the corresponding irreducible representation. Let $D = \sum d_g g$ be any element of the group ring. Then $U_\rho(D)$ is the element of $\mathrm{GL}(\mathbb{C}^{\dim U\rho})$ given by

$$\sum_g d_g U_\rho(g).$$

In the case that $D = \Delta_A$ we will denote the eigenvalues of this matrix by $0 \le \lambda_1(A,\rho) \le \ldots \le \lambda_{\dim(\rho)}(A,\rho)$ (it is well-known that $U_\rho(\Delta_A)$ is positive semidefinite and in particular diagonalizable, see e.g. [1]).

Lemma 3. *For any n and k we have*

$$\mathbb{E}(s_k(t)) = \frac{1}{k}\sum_{n\vdash\rho} a_\rho \sum_{j=1}^{\dim U_\rho} e^{-t\lambda_j(A,\rho)}$$

where a_ρ are as in theorem 3.

Proof. As discussed in §2,

$$\mathbb{E}(s_k(t)) = n!\langle \alpha_k, e^{-t\Delta_A}\rangle = \frac{1}{k}\sum_\rho a_\rho n!\langle e^{-t\Delta_A}, \chi_\rho\rangle$$

By definition, χ_ρ attaches to each $g \in S_n$ the trace of g acting on the representation U_ρ. By the linearity of the trace,

$$n!\langle e^{-t\Delta_A}, \chi_\rho\rangle = \operatorname{tr}\left(U_\rho\left(e^{-t\Delta_A}\right)\right)$$

where $U_\rho(\cdot)$ is the action of a representation on an element of the group ring as above. Further, for every representation U and any element D of the group ring,

$$U(e^D) = e^{U(D)}$$

where the exponentiation on the left-hand side is in the group ring while on the right-hand side we have exponentiation of matrices. Since $U_\rho(-t\Delta)$ is diagonalizable,

$$\operatorname{tr}\left(U_\rho\left(e^{-t\Delta}\right)\right) = \sum_j e^{-t\lambda_j(A,\rho)}.$$

The proof now follows from theorem 3. □

Proof of theorem 1. $s_n(t)$ can take only the values 0 and 1. Hence, using lemma 3 for $k = n$, we get

$$\mathbb{P}(s_n(t) = 1) = \mathbb{E}(s_n(t)) = \frac{1}{n}\sum_{i=0}^{n-1}(-1)^i\sum_j e^{-t\lambda_j(A,[n-i,1^i])}$$

Since $[n-i, 1^i]$ is a hook-shaped diagram, the eigenvalues $\lambda_j(A, [n-i, 1^i])$ are simply all the sums of i-tuples of the eigenvalues $\lambda_1(A), \ldots, \lambda_{n-1}(A)$. (See [2] and also the appendix of [1]). Hence,

$$\mathbb{P}(s_n(t) = 1) = \frac{1}{n}\left(1 + \sum_{i=1}^{n-1}(-1)^i \sum_{1\le j_1<j_2<\ldots<j_i\le n-1} e^{-t(\lambda_{j_1}+\ldots+\lambda_{j_i})}\right) = \frac{1}{n}\prod_{i=1}^{n-1}(1 - e^{-\lambda_i t}). \quad \square$$

Proof of theorem 2. The partitions that appear in lemma 3 are of the form $[a, b, 1^c]$, where $a+b+c = n$, $a \ge b > 0$, and $c \ge 0$. For such a partition a simple calculation with the hook formula [5, §4.12] gives

$$\dim U_\lambda = \frac{b(a-b+1)}{(b+c)(a+c+1)}\frac{n!}{a!b!c!} \le \binom{n}{a, b, c}$$

Hence, the total number of summands in lemma 3 is bounded by $\sum_{a+b+c=n}\binom{n}{a,b,c} = 3^n$. Also, by the celebrated Caputo-Liggett-Richthammer theorem [3], we have for all j,

$$\lambda_j(A, [a, b, 1^c]) \ge \lambda_1(A)$$

The result now follows from lemma 3. □

References

[1] Gil Alon and Gady Kozma, *Ordering the representations of S_n using the interchange process*. Preprint (2010), available from http://arxiv.org/abs/1003.1710

[2] Roland Bacher, *Valeur propre minimale du laplacien de Coxeter pour le groupe symétrique* [French, Minimal eigenvalue of the Coxeter Laplacian for the symmetric group]. J. Algebra **167:2** (1994), 460–472.

[3] Pietro Caputo, Thomas M. Liggett and Thomas Richthammer, *Proof of Aldous' spectral gap conjecture*. J. Amer. Math. Soc. **23:3** (2010), 831–851.

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

[5] William Fulton and Joe Harris, *Representation theory. A first course*. Graduate Texts in Mathematics, 129. Readings in Mathematics. Springer-Verlag, New York, 1991.

[6] Gordon James and Adalbert Kerber, *The representation theory of the symmetric group. With a foreword by P. M. Cohn. With an introduction by Gilbert de B. Robinson*. Encyclopedia of Mathematics and its Applications, 16. Addison-Wesley Publishing Co., Reading, Mass., 1981.

[7] David A. Levin, Yuval Peres and Elizabeth L. Wilmer, *Markov chains and mixing times. With a chapter by James G. Propp and David B. Wilson*. American Mathematical Society, Providence, RI, 2009.

[8] Eyal Lubetzky and Allan Sly, *Explicit expanders with cutoff phenomena*. Preprint (2010), available from http://arxiv.org/abs/1003.3515

[9] Ravi Montenegro and Prasad Tetali, *Mathematical aspects of mixing times in Markov chains*. Found. Trends Theor. Comput. Sci. **1:3** (2006). Available from https://www.math.gatech.edu/~tetali/PUBLIS/survey.pdf

[10] Ben Morris, *The mixing time for simple exclusion*. Ann. Appl. Probab. **16:2** (2006), 615–635.

[11] Roberto Imbuzeiro Oliveira, *Mixing of the symmetric exclusion processes in terms of the corresponding single-particle random walk*. Preprint (2010), available from http://arxiv.org/abs/1007.2669

[12] Robin Pemantle, *A shuffle that mixes sets of any fixed size much faster than it mixes the whole deck*. Random Structures Algorithms **5:5** (1994), 609–626.

[13] Oded Schramm, *Compositions of random transpositions*. Israel J. Math. **147** (2005), 221–243.

[14] Bálint Tóth, *Improved lower bound on the thermodynamic pressure of the spin* $1/2$ *Heisenberg ferromagnet*. Lett. Math. Phys. **28:1** (1993), 75–84.

[15] David Bruce Wilson, *Mixing times of Lozenge tiling and card shuffling Markov chains*. Ann. Appl. Probab. **14:1** (2004), 274–325.