TENSOR POWERS OF THE DEFINING REPRESENTATION OF $S_n$

SHANSHAN DING

Abstract. We give a decomposition formula for tensor powers of the defining representation of $S_n$ and apply it to bound the mixing time of a Markov chain on $S_n$.

1. Introduction

The defining, or permutation, representation of $S_n$ is the $n$-dimensional representation $\rho$ where

$$\rho(\sigma)_{i,j} = \begin{cases} 1 & \sigma(j) = i \\ 0 & \text{otherwise}. \end{cases}$$

Since the fixed points of $\sigma$ can be read off of the matrix diagonal, the character of $\rho$ at $\sigma$, $\chi_\rho(\sigma)$, is precisely the number of fixed points of $\sigma$. The irreducible representations, or irreps for short, of $S_n$ are parametrized by the partitions of $n$, and $\rho$ decomposes as $S^{n-1,1} \oplus S^n$. Note that $\chi_{S^{n-1,1}}(\sigma)$ is one less than the number of fixed points of $\sigma$. In the terminology of [7], we call the $(n-1)$-dimensional irrep $S^{(n-1,1)}$ the standard representation of $S_n$.

A classic question in the representation theory of symmetric groups is how tensor products of representations decompose as direct sums of irreps. In Section 2 we will present a neat formula for the decomposition of tensor powers of $\rho$ and, as corollary, that of tensor powers of $S^{(n-1,1)}$.

Our study of tensor powers of $\rho$ arose from an investigation in the mixing time of the Markov chain on $S_n$ formed by applying a single uniformly chosen $n$-cycle to a deck of $n$ cards and following up with repeated random transpositions. This chain is a natural counterpart to the random transposition walk on $S_n$, famously shown by Diaconis and Shahshahani in [3] to mix in $O(n \ln n)$ steps, in the sense that random transpositions induce Markov chains on not just $S_n$, but the set of partitions of $n$: the time-homogeneous random transposition walk is one such chain that starts at the partition $(1^n)$, whereas the process we proposed is one that starts at the other extreme, $(n)$. Along with following the classic approach of [3], we will use the tensor decomposition formula to show in Section 3 that the mixing time for the $n$-cycle-to-transpositions chain is $O(n)$.

2. Decomposition Formula for Tensor Powers of $\rho$

Let $\lambda$ be a partition of $n$, and recall that the irreps of $S_n$, the $S^\lambda$'s, are indexed by the partitions of $n$. As promised, we give a compact formula for the decomposition of
tensor powers of $\rho$ into irreps, i.e. the coefficients $a_{\lambda,r}$ in the expression

\[ \rho^\otimes r = \bigoplus_{\lambda \vdash n} a_{\lambda,r} S^\lambda := \bigoplus_{\lambda \vdash n} (S^\lambda)^{\otimes a_{\lambda,r}}. \]

**Proposition 2.1.** Let $\lambda \vdash n$ and $1 \leq r \leq n - \lambda^2$. The multiplicity of $S^\lambda$ in the irreducible representation decomposition of $\rho^\otimes r$ is given by

\[ a_{\lambda,r} = \sum_{i=|\lambda|}^{r} \binom{i}{|\lambda|} \binom{r}{i}, \]

where $\tilde{\lambda} = (\lambda_2, \lambda_3, \ldots)$ with weight $|\tilde{\lambda}|$, $f_{\tilde{\lambda}}$ is the number of standard Young tableaux of shape $\tilde{\lambda}$, and $\left\{ \begin{array}{c} r \end{array} \right\}$ is a Stirling number of the second kind.

**Proof.** Goupil and Chauve derived in [8] the generating function

\[ \sum_{r \geq |\lambda|} a_{\lambda,r} \frac{x^r}{r!} = \frac{f_{\tilde{\lambda}}}{|\lambda|!} e^{e^x - 1(e^x - 1)^{|\lambda|}}. \]

By (24b) and (24f) in Chapter 1 of [11],

\[ \sum_{s \geq j} \left\{ \begin{array}{c} s \\ j \end{array} \right\} \frac{x^s}{s!} = \frac{(e^x - 1)^j}{j!} \]

and

\[ \sum_{t \geq 0} B_t \frac{x^t}{t!} = e^{e^x - 1}, \]

where $B_0 := 1$ and $B_t = \sum_{q=1}^{t} \left\{ \begin{array}{c} t \\ q \end{array} \right\}$ is the $t$-th Bell number, so we obtain from (2.3) that

\[ \frac{a_{\lambda,r}}{r!} = \sum_{s+t=r} B_t \frac{B_s}{s!t!} \left\{ \begin{array}{c} s \\ |\tilde{\lambda}| \end{array} \right\}, \]

and thus

\[ \frac{a_{\lambda,r}}{f_{\tilde{\lambda}}} = \sum_{t=0}^{r-|\lambda|} B_t \binom{r}{t} \frac{r-t}{|\lambda|!} \]

\[ = \left\{ \begin{array}{c} r \\ |\lambda| \end{array} \right\} + \sum_{t=1}^{r-|\lambda|} \sum_{q=1}^{t} \binom{t}{q} \binom{r-t}{|\lambda|} \]

\[ = \left\{ \begin{array}{c} r \\ |\lambda| \end{array} \right\} + \sum_{q=1}^{r-|\lambda|} \sum_{t=q}^{r-|\lambda|} \binom{t}{q} \binom{r-t}{|\lambda|}. \]
By (24.1.3, II.A) of [1],
\[
\sum_{t=q}^{r-|\bar{\lambda}|} \binom{t}{q} \binom{r}{t} \binom{r-t}{|\bar{\lambda}|} = \left( \frac{q+|\bar{\lambda}|}{|\bar{\lambda}|} \right) \binom{r}{q+|\bar{\lambda}|},
\]
so that
\[
a_{\lambda, r} = \frac{f^{\lambda}}{|\bar{\lambda}|!} + \sum_{q=1}^{r-|\bar{\lambda}|} \left( \frac{q+|\bar{\lambda}|}{|\bar{\lambda}|} \right) \binom{r}{q+|\bar{\lambda}|}
\]
\[
= \frac{r}{|\bar{\lambda}|!} + \sum_{i=|\bar{\lambda}|+1}^{r} \left( \frac{i}{|\bar{\lambda}|} \right) \binom{r}{i} = \sum_{i=|\bar{\lambda}|}^{r} \left( \frac{i}{|\bar{\lambda}|} \right) \binom{r}{i},
\]
as was to be shown. \(\square\)

Now, let \(b_{\lambda, r}\) be the multiplicities such that
\[
(S^{(n-1,1)})^\otimes r = \bigoplus_{\lambda \vdash n} b_{\lambda, r} S^\lambda.
\]
Goupil and Chauve also derived the generating function
\[
\sum_{r \geq |\bar{\lambda}|} b_{\lambda, r} \frac{x^r}{r!} = \frac{f^{\lambda}}{|\bar{\lambda}|!} e^{x - x - 1} (e^x - 1)^{|\bar{\lambda}|},
\]
so from Proposition 2.1 we can obtain a formula for the decomposition of \((S^{(n-1,1)})^\otimes r\) as well.

**Corollary 2.2.** Let \(\lambda \vdash n\) and \(1 \leq r \leq n - \lambda_2\). The multiplicity of \(S^\lambda\) in the irreducible representation decomposition of \((S^{(n-1,1)})^\otimes r\) is given by
\[
b_{\lambda, r} = f^{\lambda} \sum_{s=|\lambda|}^{r} (-1)^{r-s} \binom{r}{s} \left( \sum_{i=|\lambda|}^{s} \binom{i}{s} \right) \left( \sum_{s \geq |\lambda|} \frac{x^s}{s!} \right) \left( \sum_{l \geq 0} \frac{(-x)^l}{l!} \right).
\]

**Proof.** Comparing (2.11) with (2.3) gives
\[
\sum_{r \geq |\lambda|} b_{\lambda, r} \frac{x^r}{r!} = \left( \sum_{s \geq |\lambda|} \frac{x^s}{s!} \right) e^{-x} = \left( \sum_{s \geq |\lambda|} \frac{x^s}{s!} \right) \left( \sum_{l \geq 0} \frac{(-x)^l}{l!} \right),
\]
so that
\[
b_{\lambda, r} = \sum_{s+l=r} \frac{(-1)^s a_{\lambda, s}}{s! l!} = \sum_{s=|\lambda|}^{r} \frac{(-1)^{r-s}}{s!(r-s)!} \left( f^{\lambda} \sum_{i=|\lambda|}^{s} \binom{i}{s} \right),
\]
and the result follows. \(\square\)
Remark. Corollary 2.2 is very similar to Proposition 2 of [8], but our result is cleaner, as it does not involve associated Stirling numbers of the second kind. For another approach to the decomposition of tensor powers of $\rho$, see [6].

3. Connection to Markov Chain Mixing Time

Consider the Markov chain on $S_n$ formed by first applying a random $n$-cycle to a deck of $n$ cards and then following with repeated random transpositions. Formally, form a Markov chain $\{X_k\}$ on the symmetric group $S_n$ as follows: let $X_0$ be the identity, set $X_1 = \pi X_0$, where $\pi$ is a uniformly selected $n$-cycle, and for $k \geq 2$ set $X_k = \tau_k X_{k-1}$, where $\tau_k$ is a uniformly selected transposition. Observe that $X_k \in A_n$ when $n$ and $k$ are of the same parity. Otherwise, $X_k \in S_n \setminus A_n$. Let $\mu_k$ be the law of $X_k$, and let $U_k$ be the uniform measure on $A_n$ if $X_k \in A_n$ and the uniform measure on $S_n \setminus A_n$ if $X_k \in S_n \setminus A_n$. What is the total variation distance between $\mu_k$ and $U_k$?

The goal of this section is to prove the following:

Theorem 3.1. For any $c > 0$, after one $n$-cycle and $cn$ transpositions,

$$\frac{e^{-2cn}}{e} - o(1) \leq \|\mu_{cn+1} - U_{cn+1}\|_{TV} \leq \frac{e^{-2cn}}{2\sqrt{1 - e^{-4cn}}} + o(1)$$

as $n$ goes to infinity.

The upper bound follows from the approach of [3]. For the (lazy) random transposition shuffle on $n$ cards, the time-homogeneous chain on $S_n$ with increment measure $\nu$ that assigns mass $\frac{1}{n}$ to the identity and $\frac{2}{n^2}$ to each of the $\frac{n(n-1)}{2}$ transpositions $\tau$, Diaconis and Shahshahani derived the bound

$$4\|\mu_k - U\|_{TV}^2 \leq \sum_{\rho \in S_n^* \setminus \{\text{triv} \}} d^2(\rho) \left( \frac{1}{n} + \frac{n-1}{nd^2(\rho)} \right)^2 d(\rho) \left( \chi(\tau) \right)^{2k},$$

where $U$ is the uniform measure on $S_n$, $\hat{S}_n$ is the set of irreps of $S_n$, and $d(\rho)$ and $\chi(\tau)$ denote the dimension and the character at $\tau$ of the representation $\rho$, respectively. Careful computations of the terms on the RHS of (3.2) gave a mixing time of $O(n \ln n)$, and explicit constants were later calculated by Saloff-Coste and Zúñiga in [10].

Inequality (3.2) comes from the theory of non-commutative Fourier analysis on $S_n$. It carries the following routine extension (carefully spelled out in Chapter 2 of [4]) to the $n$-cycle-to-transpositions chain:

$$4\|\mu_{k+1} - U_{k+1}\|_{TV}^2 \leq \frac{1}{2} \sum_{\rho \in S_n^* \setminus \{\text{triv, sign} \}} d^2(\rho) \left( \frac{\chi(\tau)}{d(\rho)} \right)^{2k} \left( \frac{\chi(\pi)}{d(\rho)} \right)^{2k},$$

Proposition 3.2. For any $c > 0$, after one $n$-cycle and $cn$ transpositions,

$$4\|\mu_{cn+1} - U_{cn+1}\|_{TV}^2 \leq \frac{e^{-4cn}}{1 - e^{-4cn}} + o(1)$$
as $n$ goes to infinity.

Proof. Let $\chi_\lambda^\gamma$ denote the character of $S^\lambda$ on the cycle type $\gamma$. The first and most critical step of the proof is the observation that, discounting $(n)$ and $(1^n)$, $\chi_\lambda^{(n)} = 0$ for all $\lambda$ except the hook-shaped ones, for which $\lambda_2 = 1$. This is an almost trivial consequence of the Murnaghan-Nakayama rule, as it is impossible to remove a rim hook of size $n$ from a Young diagram of size $n$ unless the Young diagram itself is the rim hook. Moreover, for a hook-shaped $\lambda$, it is clear that $\chi_\lambda^{(n)}$ is equal to 1 if $\lambda$ has an odd number of rows and $-1$ if $\lambda$ has an even number of rows. Thus we arrive at a significant simplification of (3.3), namely that

$$4 \| \mu_{k+1} - U_{k+1} \|_{TV}^2 \leq \frac{1}{2} \sum_{\lambda \in \Lambda_n} \left( \frac{\chi_\lambda^{(2,1^{n-2})}}{\dim S^\lambda} \right)^{2k},$$

where

$$\Lambda_n = \{ \lambda \vdash n : \lambda_1 > 1 \text{ and } \lambda_2 = 1 \}.$$ 

The normalized characters $\frac{\chi_\lambda^{(2,1^{n-2})}}{\dim S^\lambda}$ have a simple description when $\lambda \in \Lambda_n$: let $j$ be one less than the number of rows of $\lambda$, then for $1 \leq j \leq \left\lfloor \frac{n-1}{2} \right\rfloor$,

$$\frac{\chi_\lambda^{(n-j,1^j)}}{\dim S^{(n-j,1^j)}} = \frac{n-1-2j}{n-1}.$$ 

This is a special case of the identity

$$\frac{\chi_\lambda^{(2,1^{n-2})}}{\dim S^\lambda} = \frac{\sum_i (\lambda_i^2 - (2i-1)\lambda_i)}{n(n-1)},$$

known as early as to Frobenius in [3].

Fix any $c > 0$. By calculus, for $n-1-2j > 0$,

$$\lim_{n \to \infty} \left( \frac{n-1-2j}{n-1} \right)^{2cn} = e^{-4cj}.$$ 

Thus (3.7) and the fact that $\chi_\gamma^\lambda = \pm \chi_{\gamma'}^\lambda$, where $\lambda'$ is the conjugate partition of $\lambda$ (see p. 25 of [9]), imply that

$$\sum_{\lambda \in \Lambda_n} \left( \frac{\chi_\lambda^{(2,1^{n-2})}}{\dim S^\lambda} \right)^{2cn} \sim \begin{cases} 2 \sum_{j=1}^{(n-2)/2} e^{-4cj} & n \text{ is even} \\ 2 \sum_{j=1}^{(n-3)/2} e^{-4cj} & n \text{ is odd} \end{cases}.$$ 

Summing the geometric series gives

$$4 \| \mu_{cn+1} - U_{cn+1} \|_{TV}^2 \leq \frac{1}{2} \sum_{\lambda \in \Lambda_n} \left( \frac{\chi_\lambda^{(2,1^{n-2})}}{\dim S^\lambda} \right)^{2cn} \sim \frac{e^{-4c}}{1 - e^{-4c}}.$$
as was to be shown.

For measures \( \mu \) and \( \nu \) on a set \( G \), a classic approach to finding a lower bound for \( \| \mu - \nu \|_{TV} \) is to identify a subset \( A \) of \( G \) where \( |\mu(A) - \nu(A)| \) is close to maximal. In many mixing problems involving the symmetric group, it is convenient to make \( A \) either the set of fixed-point-free permutations or its complement, since it is well-known that the distribution of the number of fixed points with respect to the uniform measure on \( S_n \) is asymptotically \( \mathcal{P}(1) \), the Poisson distribution of mean one. The same is true for the distribution of fixed points with respect to the uniform measure on either \( A_n \) or \( S_n \setminus A_n \). See Theorem 4.3.3 of [4] for a proof.

For Diaconis and Shahshahani’s random transposition shuffle, \( A \) is the set of permutations with one or more fixed points, and finding \( \mu_k(A) \) boils down to a coupon collector’s problem. Let \( B \) be the event that, after \( k \) transpositions, at least one card is untouched. It is not difficult to see that \( \mu_k(A) \geq \mathbb{P}(B) \), where \( \mathbb{P}(B) \) is equal to the probability that at least one of \( n \) coupons is still missing after \( 2k \) trials. The coupon collector’s problem is well-studied, so this immediately gives a lower bound for \( \mu_k(A) \), which in turn produces a lower bound for \( \| \mu_k(A) - U(A) \|_{TV} \).

The above argument is so short and simple that it was tagged onto the end of the introduction of [3], as if an afterthought. Unfortunately, it is inapplicable to our problem, since the initial \( n \)-cycle obliterates the core of the argument. Instead, we will fully characterize the distribution of \( \chi_\emptyset \) with respect to \( \mu_{k+1} \) by deriving all moments of \( \chi_\emptyset \) with respect to \( \mu_k \). Let \( E_\mu \) denote expectation with respect to \( \mu \), then as observed in Chapter 3D of [2],

\[
(3.12) \quad E_\mu(\chi_\rho) = \sum_{\sigma \in S_n} \mu(\sigma) \text{tr}(\rho(\sigma)) = \text{tr} \left( \sum_{\sigma \in S_n} \mu(\sigma) \rho(\sigma) \right) = \text{tr}(\hat{\mu}(\rho)),
\]

so that

\[
(3.13) \quad E_\mu((\chi_\emptyset)^r) = \sum_{\lambda \vdash n} a_{\lambda,r} \text{tr}(\hat{\mu}(S^\lambda)),
\]

where \( \hat{\mu} \) is the Fourier transform of \( \mu \) and

\[
(3.14) \quad \text{tr}(\hat{\mu}_{k+1}(S^\lambda)) = \chi_\emptyset^\lambda(n) \left( \frac{\chi_\emptyset^{\lambda(2,1^{n-2})}}{\dim S^\lambda} \right)^k.
\]

**Proposition 3.3.** Fix any \( c > 0 \). As \( n \) approaches infinity, the distribution of the number of fixed points after one \( n \)-cycle and \( cn \) transpositions converges to \( \mathcal{P}(1 - e^{-2c}) \).

**Proof.** One can deduce from the moment-generating function that the \( r \)-th moment of \( \mathcal{P}(\nu) \) is \( \sum_{i=1}^r \{r\} \nu^i \). It is a standard result that \( \hat{\mu}_{cn+1}(S^{(n)}) = 1 \), and we will ignore the alternating representation because it suffices to consider the first \( n - 2 \) moments, in which the alternating representation does not appear. For the non-trivial and non-alternating representations, we take advantage of previous computations and synthesize
(3.7), (3.9) with \( n \) instead of \( 2n \), and (3.14) to obtain

\[
(3.15) \quad \hat{\mu}_{cn+1}(S^\lambda) \sim \begin{cases} 
(-1)^{|\lambda|} e^{-2c|\lambda|} & \lambda \in \Lambda_n \\
0 & \text{otherwise}
\end{cases}
\]

By Proposition 2.1 (second line below) and (3.15) (fourth line), for \( 1 \leq r \leq n-2 \),

\[
E_{\mu_{cn+1}}((\chi_\iota)^r) = a_{(n),r} + \sum_{\lambda \in \Lambda_n} a_{\lambda,r} \hat{\mu}_{cn+1}(S^\lambda)
\]

\[
= \sum_{i=1}^{r} \binom{r}{i} + \sum_{|\lambda| = 1}^{n-2} \sum_{i=|\lambda|}^{r} \binom{r}{i} \left( \frac{i}{|\lambda|} \right) \hat{\mu}_{cn+1}(S^\lambda)
\]

\[
= \sum_{i=1}^{r} \binom{r}{i} + \sum_{i=1}^{r} \sum_{|\lambda| = 1}^{i} \binom{r}{i} \left( \frac{i}{|\lambda|} \right) (-e^{-2c}|\lambda|)
\]

\[
\sim \sum_{i=1}^{r} \binom{r}{i} \left( 1 + \sum_{|\lambda| = 1}^{i} \left( \frac{i}{|\lambda|} \right) (-e^{-2c}|\lambda|) \right)
\]

\[
= \sum_{i=1}^{r} \binom{r}{i} (1 - e^{-2c})^i.
\]

This shows that the first \( n - 2 \) moments of \( \chi_\iota \) with respect to \( \mu_{cn+1} \) approach those of \( P(1 - e^{-2c}) \), and convergence follows from the method of moments. \( \square \)

Corollary 3.4. For any \( c > 0 \), after one \( n \)-cycle and \( cn \) transpositions,

\[
(3.17) \quad \| \mu_{cn+1} - U_{cn+1} \|_{TV} \geq \frac{e^{-2c}}{e} - o(1)
\]

as \( n \) goes to infinity.

\textbf{Proof.} Let \( A \) be the set of fixed-point-free permutations. Then

\[
(3.18) \quad \| \mu_{cn+1} - U_{cn+1} \|_{TV} \geq |\mu_{cn+1}(A) - U_{cn+1}(A)|
\]

\[
\sim e^{-2c} - 1 - \frac{1}{e} = \frac{1}{e} \left( e^{-2c} + \frac{(e^{-2c})^2}{2!} + \cdots \right) \geq \frac{e^{-2c}}{e},
\]

as was to be shown. \( \square \)

Together with Proposition 3.3 Corollary 3.4 completes the proof of Theorem 3.1.
SHANSHAN DING

REFERENCES

[1] M. Abramowitz and I. A. Stegun, eds., Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Dover, New York, 1965.

[2] P. Diaconis, Group Representations in Probability and Statistics, IMS Lecture Notes Monogr. Ser. 11, Inst. Math. Statist., Hayward, CA, 1988.

[3] P. Diaconis and M. Shahshahani, Generating a random permutation with random transpositions, Z. Wahrsch. verw. Geb. 57 (1981), no. 2, 159-179.

[4] S. Ding, A Random Walk in Representations, Ph.D. Thesis, University of Pennsylvania, 2014.

[5] F. G. Frobenius, Über die Charaktere der symmetrischen Gruppen, Sitz. Konig. Preuss. Akad. Wissen. (1900), 516-534.

[6] J. Fulman, Separation cutoffs for random walk on irreducible representations, Ann. Comb. 14 (2010), no. 3, 319-337.

[7] W. Fulton and J. Harris, Representation Theory: A First Course, GTM 129, Springer-Verlag, New York, 1991.

[8] A. Goupil and C. Chauve, Combinatorial operators for Kronecker powers of representations of $S_n$, Séminaire Lotharingien de Combinatoire, 54 (2006), B54j.

[9] G. D. James, The Representation Theory of the Symmetric Groups, LNM 682, Springer-Verlag, Berlin, 1978.

[10] L. Saloff-Coste and J. Zúñiga, Refined estimates for some basic random walks on the symmetric and alternating groups, ALEA Lat. Am. J. Probab. Math. Stat. 4 (2008), 359-392.

[11] R. P. Stanley, Enumerative Combinatorics, Vol. I, Wadsworth, Monterey, CA, 1986, Cambridge Stud. Adv. Math. 49, reprinted by Cambridge Univ. Press, Cambridge, 1997.

E-mail address: dish@sas.upenn.edu