Pacific Journal of Mathematics

CARD SHUFFLING AND THE DECOMPOSITION OF TENSOR PRODUCTS

Jason Fulman

Volume 217 No. 2 December 2004

PACIFIC JOURNAL OF MATHEMATICS
Vol. 217, No. 2, 2004

CARD SHUFFLING AND THE DECOMPOSITION OF TENSOR PRODUCTS

Jason Fulman

Let H be a subgroup of a finite group G. We use Markov chains to quantify how large r should be so that the decomposition of the r tensor power of the representation of G on cosets on H behaves (after renormalization) like the regular representation of G. For the case where G is a symmetric group and H a parabolic subgroup, we find that this question is precisely equivalent to the question of how large r should be so that r iterations of a shuffling method randomize the Robinson–Schensted–Knuth shape of a permutation. This equivalence is remarkable, if only because the representation theory problem is related to a reversible Markov chain on the set of representations of the symmetric group, whereas the card shuffling problem is related to a nonreversible Markov chain on the symmetric group. The equivalence is also useful, and results on card shuffling can be applied to yield sharp results about the decomposition of tensor powers.

1. Introduction

Let χ be a faithful character of a finite group G. A well-known theorem of Burnside and Brauer [**I**] states that if $\chi(g)$ takes on exactly m distinct values for $g \in G$, then every irreducible character of G is a constituent of one of the characters χ^j for $0 \leq j < m$. It is very natural to investigate the decomposition of χ^j, and the results in this paper are a step in that direction.

Let $\mathrm{Irr}\,(G)$ denote the set of irreducible representations of a finite group G. The Plancherel measure on $\mathrm{Irr}\,(G)$ is a probability measure that assigns mass $\dim(\rho)^2/|G|$ to ρ. The symbol χ^ρ denotes the character associated to the representation ρ. The notation Ind, Res stands for induction and restriction of class functions. We remind the reader that the character of the r-fold tensor product of a representation of G is given by raising the character to the r-th power. The inner product $\langle f_1, f_2\rangle$ denotes the usual inner product on class functions of G defined by

$$\frac{1}{|G|}\sum_{g\in G}\overline{f_1(g)}f_2(g).$$

Thus if f_1 is an irreducible character and f_2 any character, their inner product gives the multiplicity of f_1 in f_2. We let g^G denote the conjugacy class of g in G.

In Section 2 of this paper, we prove the following result:

Theorem 1.1. *Let H be a subgroup of a finite group G and let* id *denote the identity element. Let π denote the Plancherel measure of G. Suppose that $|G| > 1$. Let*

$$\beta = \max_{g \neq \mathrm{id}} \frac{|g^G \cap H|}{|g^G|} = \frac{|H|}{|G|} \max_{g \neq \mathrm{id}} \mathrm{Ind}\,_H^G(1)[g].$$

Then

$$\sum_{\rho \in \mathrm{Irr}\,(G)} \left| \left(\frac{|H|}{|G|}\right)^r \dim(\rho) \langle \chi^\rho, (\mathrm{Ind}\,_H^G(1))^r \rangle - \pi(\rho) \right| \leq |G|^{1/2} \beta^r.$$

Note that if $\beta < 1$, the right-hand side approaches 0 as $r \to \infty$. The quantity β has been carefully studied in the (most interesting) case that G is simple and H a maximal subgroup; references and an example where H is not maximal are given in Section 2.

The idea behind the proof of Theorem 1.1 is to investigate a natural Markov chain J on the set of irreducible representations of G. This chain is essentially a probabilistic reformulation of Frobenius reciprocity. This chain can be explicitly diagonalized and then Theorem 1.1 follows from spectral theory of reversible Markov chains, with $1 - \beta$ having the interpretation of a spectral gap. In fact Theorem 1.1 is a generalization of a result in our earlier paper [**F1**], where this Markov chain arose for the symmetric group case $H = S_{n-1}$ and $G = S_n$ and was combined with Stein's method to sharpen a result of Kerov on the asymptotic normality of random character ratios of the symmetric group on transpositions.

The main insight of the current paper is that when G is the symmetric group S_n and H is a parabolic subgroup, the bound of Theorem 1.1 can be improved by card shuffling. Let us describe this in detail for the case $H = S_{n-1}$. In Theorem 1.1, $\beta = 1 - \frac{2}{n}$, and one can see using Stirling's approximation for $n!$ that for

$$r > \frac{n \log(n) + 2c}{2 \log(\frac{1}{\beta})},$$

the bound in Theorem 1.1 is at most $(2\pi)^{1/4} e^{-c}$ (and hence small). Note that all logs in this paper are base e. Thus r slightly more than $\frac{1}{4} n^2 \log n$ suffices to make the bound small. The bound of Theorem 1.1 is proved by analyzing a certain Markov chain J on $\mathrm{Irr}\,(S_n)$, started at the trivial representation. The irreducible representations of S_n correspond to partitions of n (the one row partition is the trivial representation), so J is a Markov chain on partitions. Although we do not need this observation, we remark that viewing partitions as Young diagrams, this Markov chain amounts to removing a single box with certain probabilities and reattaching it somewhere. We show that the distribution on partitions given by taking r steps according to J has a completely different description. Namely starting from the identity permutation (viewed as n cards in order), perform the following procedure r times: remove the top card and insert it into a uniformly chosen random position. This gives a nonuniform random permutation, and there is a natural map called the Robinson–Schensted–Knuth or RSK correspondence (see [**Sa**] for background), which associates a partition to a permutation. We will show that applying this correspondence to the permutation obtained after r iterations of the top to random shuffle gives *exactly* the same distribution on partitions as that given by r iterations of the chain J started at the trivial representation. This will allow us to use facts about card shuffling to sharpen $\frac{1}{4}n^2 \log n$ to roughly $n \log n$, and even to see that the $n \log(n)$ is sharp to within a factor of two. Precise statements and results for more general parabolic subgroups are given in Section 3.

To close the introduction we make some remarks. First, recall that a Markov chain M on a finite set X is called reversible with respect to the probability measure μ on X if $\mu(x)M(x,y) = \mu(y)M(y,x)$ for all x, y (this implies μ is stationary for M, i.e., $\mu(y) = \sum_x \mu(x)M(x,y)$ for all y). The top to random shuffle and its cousins that arise in connection with parabolic subgroups are nonreversible chains. Thus it is rather miraculous that the top to random shuffle has real eigenvalues; this observation is the starting point of a general theory [**BHR**]. And it is doubly surprising that the top to random shuffle should be connected with the reversible chains J. Since Proposition 3.3 shows these chains to have the same set of eigenvalues, this gives an application of the eigenvalue formulas in [**BHR**]. See [**F4**] for some other connections between the top to random shuffle and reversible Markov chains. Second, the problem of studying the convergence rate of the RSK shape after iterated shuffles to the RSK shape of a random permutation is of significant interest independent of its application in this paper. It is closely connected with random matrix theory and in some cases with Toeplitz determinants. See [**St**], [**F2**], [**F3**] and the references therein for details. Third, since Solomon's descent algebra generalizes to finite Coxeter groups, it is likely that the results in this paper can be pushed through to that setting. (However that would require an analog of the RSK correspondence for finite Coxeter groups). Fourth, we note that some of the results in this paper have now been from extended to arbitrary real valued characters of finite groups and to spherical functions of Gelfand pairs [**F5**], [**CF**]; as an application one obtains a probabilistic proof of a result of Burnside and Brauer on the decomposition of tensor products [**F6**].

2. General groups

This section proves Theorem 1.1 and gives an example. Throughout this section $X = \mathrm{Irr}\,(G)$ is the set of irreducible representations of a finite group G, endowed with Plancherel measure π_G. We also suppose that we are given a subgroup H of G.

To begin, we use H to construct a Markov chain on $\mathrm{Irr}\,(G)$ that is reversible with respect to π_G. For ρ an irreducible representation of G and τ an irreducible representation of H, we let $\kappa(\tau,\rho)$ denote the multiplicity of τ in $\mathrm{Res}\,^G_H(\rho)$. By Frobenius reciprocity, this is the multiplicity of ρ in $\mathrm{Ind}\,^G_H(\tau)$.

Proposition 2.1. *The Markov chain J on irreducible representations of G that moves from ρ to σ with probability*

$$\frac{|H|}{|G|}\frac{\dim(\sigma)}{\dim(\rho)}\sum_{\tau\in\mathrm{Irr}\,(H)}\kappa(\tau,\rho)\kappa(\tau,\sigma)$$

is in fact a Markov chain (*the transition probabilities sum to* 1), *and is reversible with respect to the Plancherel measure π_G.*

Proof. First let us check that the transition probabilities sum to 1. Indeed,

$$\begin{aligned}
&\sum_{\sigma\in\mathrm{Irr}\,(G)}\frac{|H|}{|G|}\frac{\dim(\sigma)}{\dim(\rho)}\sum_{\tau\in\mathrm{Irr}\,(H)}\kappa(\tau,\rho)\kappa(\tau,\sigma)\\
&=\frac{|H|}{|G|}\frac{1}{\dim(\rho)}\sum_{\tau\in\mathrm{Irr}\,(H)}\kappa(\tau,\rho)\sum_{\sigma\in\mathrm{Irr}\,(G)}\dim(\sigma)\kappa(\tau,\sigma)\\
&=\frac{1}{\dim(\rho)}\sum_{\tau\in\mathrm{Irr}\,(H)}\kappa(\tau,\rho)\dim(\tau)=1.
\end{aligned}$$

The second equality follows since the dimension of a representation induced from a subgroup is its original dimension multiplied by the index of the subgroup.

The reversibility with respect to Plancherel measure is immediate from the definitions. □

Next we quickly review some facts from Markov chain theory. We consider the space of real valued functions $\ell^2(\pi)$ with the norm

$$\|f\|_2=\left(\sum_x|f(x)|^2\pi(x)\right)^{1/2}.$$

If $J(x,y)$ is the transition rule for a Markov chain on a finite set X, the associated operator (also denoted by J) on $\ell^2(\pi)$ is given by $Jf(x) = \sum_y J(x,y)f(y)$. Let $J^r(x,y) = J^r_x(y)$ denote the chance that the Markov chain started at x is at y after r steps.

If the Markov chain with transition rule $J(x,y)$ is reversible with respect to π (i.e., $\pi(x)J(x,y) = \pi(y)J(y,x)$ for all x, y), then the operator J is self adjoint with real eigenvalues

$$-1 \leq \beta_{\min} = \beta_{|X|-1} \leq \cdots \leq \beta_1 \leq \beta_0 = 1.$$

Let ψ_i ($i = 0, \ldots, |X| - 1$) be an orthonormal basis of eigenfunctions such that $J\psi_i = \beta_i\psi_i$ and $\psi_0 \equiv 1$. Define $\beta = \max\{|\beta_{\min}|, |\beta_1|\}$.

The total variation distance between two probability measures Q_1, Q_2 on a set X is defined as $\|Q_1 - Q_2\|_{TV} = \frac{1}{2}\sum_{x\in X} |Q_1(x) - Q_2(x)|$. It is elementary that $\|Q_1 - Q_2\|_{TV} = \max_{A\subseteq X} |Q_1(A) - Q_2(A)|$. Thus when the total variation distance is small, the Q_1 and Q_2 probabilities of any event A are close.

The following lemma is well-known; for a proof see [**DSa**].

Lemma 2.2.

1) $2\|J_x^r - \pi\|_{TV} \leq \|(J_x^r/\pi) - 1\|_2$.

2) $J^r(x,y) = \displaystyle\sum_{i=0}^{|X|-1} \beta_i^r \psi_i(x)\psi_i(y)\pi(y)$.

3) $\|(J_x^r/\pi) - 1\|_2^2 = \displaystyle\sum_{i=1}^{|X|-1} \beta_i^{2r}|\psi_i(x)|^2 \leq \frac{1-\pi(x)}{\pi(x)}\beta^{2r}$.

Proposition 2.3. *Let G be a finite group and H a subgroup of G. Then the eigenvalues and eigenfunctions of the operator J are indexed by conjugacy classes C of G.*

1) *The eigenvalue parameterized by C is $|C \cap H|/|C|$.*

2) *An orthonormal basis of eigenfunctions ψ_C is defined by*

$$\psi_C(\rho) = \frac{|C|^{\frac{1}{2}}\chi^\rho(C)}{\dim(\rho)}.$$

Proof. First, note that the transition probability in the definition of J can be rewritten as

$$\begin{aligned}
&\frac{|H|\dim(\sigma)}{|G|\dim(\rho)}\left\langle \chi^\sigma, \mathrm{Ind}_H^G \mathrm{Res}_H^G(\chi^\rho)\right\rangle \\
&\qquad = \frac{|H|\dim(\sigma)}{|G|\dim(\rho)}\frac{1}{|G|}\sum_{g\in G}\overline{\chi^\sigma(g)}\frac{1}{|H|}\sum_{\substack{t\in G\\ t^{-1}gt\in H}}\chi^\rho(t^{-1}gt) \\
&\qquad = \frac{\dim(\sigma)}{\dim(\rho)}\frac{1}{|G|}\sum_{g\in G}\overline{\chi^\sigma(g)}\chi^\rho(g)\frac{|g^G\cap H|}{|g^G|}.
\end{aligned}$$

The first equality used the well-known formula for induced characters [**I**].

Now to see that ψ_C is an eigenfunction with the asserted eigenvalue, one calculates that

$$\begin{aligned}\sum_{\sigma\in\mathrm{Irr}\,(G)} \frac{\dim(\sigma)}{\dim(\rho)}\frac{1}{|G|}\sum_{g\in G}\overline{\chi^\sigma(g)}\chi^\rho(g)\frac{|g^G\cap H|}{|g^G|}|C|^{\frac{1}{2}}\frac{\chi^\sigma(C)}{\dim(\sigma)} \\ = \frac{|C|^{\frac{1}{2}}}{\dim(\rho)}\sum_{g\in G}\frac{|g^G\cap H|}{|g^G|}\frac{\chi^\rho(g)}{|G|}\sum_{\sigma\in\mathrm{Irr}\,(G)}\overline{\chi^\sigma(g)}\chi^\sigma(C) \\ = \frac{|C|^{\frac{1}{2}}\chi^\rho(C)}{\dim(\rho)}\sum_{g\in C}\frac{|g^G\cap H|}{|g^G|}\frac{1}{|g^G|} \\ = \frac{|C|^{\frac{1}{2}}\chi^\rho(C)}{\dim(\rho)}\frac{|C\cap H|}{|C|}.\end{aligned}$$

The second inequality used the orthogonality relations of the characters of G.

Finally, the fact that ψ_C are orthonormal follows from the orthogonality relations for irreducible characters. They are a basis since the number of irreducible representations of a finite group is equal to its number of conjugacy classes. $\square$

Next we prove Theorem 1.1 from the introduction.

Proof. First note that the equivalence of the definitions of β follows from the general formula for induced characters. Now let 1 denote the trivial representation of G. From Proposition 2.3 and part 2 of Lemma 2.2,

$$\begin{aligned}J_1^r(\rho) &= \dim(\rho)\sum_C\left(\frac{|C\cap H|}{|C|}\right)^r\frac{|C|\chi^\rho(C)}{|G|} \\ &= \dim(\rho)\frac{1}{|G|}\sum_{g\in G}\left(\frac{|g^G\cap H|}{|g^G|}\right)^r\chi^\rho(g) \\ &= \dim(\rho)\left(\frac{|H|}{|G|}\right)^r\langle\chi^\rho, (\mathrm{Ind}\,_H^G(1))^r\rangle,\end{aligned}$$

where in the third equality we have used the well-known formula for induced characters used in the proof of Proposition 2.3. The theorem now follows from part 1 of Proposition 2.3 and parts 1 and 3 of Lemma 2.2. $\square$

Remarks.

1) The quantity β has been well studied in the case that G is simple and H is a maximal subgroup of G. See for instance [**GK**], [**LSh**] and the references therein. We defer discussion of the case that $G = S_n$ and H is a parabolic subgroup to Section 3. The remarkable paper [**GM**] classifies all pairs (G,H) where G is a finite group, H is maximal in G, and β is at least 1/2.

2) Observe that if $\beta = 1$ the upper bound of Theorem 1.1 is useless. And it can happen that $\beta = 1$. For instance if H is a nontrivial normal subgroup of G, there are conjugacy classes of G contained in H. On the representation theory side, suppose for simplicity that H is normal of index 2. Then except in trivial cases, the state space of the Markov chain J isn't connected, so the the quantity bounded in Theorem 1.1 won't go to 0 as $r \to \infty$. Indeed, either $\mathrm{Ind}_H^G \mathrm{Res}_H^G(\rho)$ is two copies of ρ or else the sum of ρ and ρ', where the character of ρ' is equal to the character of ρ on H but takes opposite values on $G - H$ [**FH**, p. 64].
3) If $\beta = 0$, then $|H| = 1$, which implies that the decomposition of $\mathrm{Ind}_H^G(1)$ is given exactly by Plancherel measure. Then the bound in Theorem 1.1 is an equality.
4) Note that Propositions 2.1 and 2.3 involve the idea of first restricting a representation of G to H and then inducing. There is a similar (but less natural) result for inducing and then restricting. Namely the Markov chain in Proposition 2.1 becomes a reversible Markov chain on irreducible representations of H (with respect to the Plancherel measure of H), where one moves from ρ to σ with probability
$$\frac{|H|}{|G|} \frac{\dim(\sigma)}{\dim(\rho)} \sum_{\tau \in \mathrm{Irr}\,(G)} \kappa(\rho, \tau)\kappa(\sigma, \tau).$$
If G conjugacy classes of H coincide with conjugacy classes of H, then the eigenvalues are parameterized by conjugacy classes C of H: the eigenvalue is $|C|/|C^G|$ (the denominator is the size of the conjugacy class of C in G), and the eigenvector is $|C|^{\frac{1}{2}} \chi^\rho(C)/\dim(\rho)$. For the pair (S_n, S_{n+1}) this was applied in [**F1**] and the proof method is similar to that of Proposition 2.3. However we believe that it is more natural to restrict and then induce as this involves only the internal structure of the group. Hence we do not develop this remark further.

To conclude this section we compute β in the case that $G = \mathrm{GL}(n, q)$ and $H = \mathrm{GL}(n-1, q)$ (which is not a maximal subgroup). There are clearly more examples in this direction that can be worked out using Wall's formulas for conjugacy class sizes [**W**] — though as in Proposition 2.4 below some (minor) effort is required to determine when $|g^G \cap H|/|g^G|$ is largest for nontrivial g. However as we have no need for them we stop here.

Proposition 2.4. *Suppose that* $G = \mathrm{GL}(n, q)$ *and* $H = \mathrm{GL}(n-1, q)$*, and that* $n \geq 2$*. Then*
$$\beta = \frac{(1 - 1/q^{n-1})}{q^2(1 - 1/q^n)} \quad \textit{for } q > 2 \quad \textit{and} \quad \beta = \frac{(1 - 1/q^{n-2})}{q^2(1 - 1/q^n)} \quad \textit{for } q = 2.$$

Proof. The conjugacy classes C of $\mathrm{GL}(n, q)$ are parameterized by all ways of associating a partition λ_ϕ to each monic irreducible polynomial $\phi(z)$ with coefficients in F_q such that $|\lambda_z| = 0$ and $\sum_\phi \deg(\phi)|\lambda_\phi| = n$. Here $|\lambda|$ denotes the size of a partition λ and $\deg(\phi)$ denotes the degree of the polynomial ϕ. Moreover the size of the conjugacy class with this data is ([**M**], p. 181)

$$\frac{|\mathrm{GL}(n,q)|}{\prod_\phi \prod_{j\geq 1} q^{\deg(\phi)(\lambda'_{\phi,j})^2}(1-1/q^{\deg(\phi)})\cdots(1-1/q^{\deg(\phi)m_j(\lambda_\phi)})}.$$

Here $m_j(\lambda_\phi)$ is the number of parts of λ_ϕ of size j, and $\lambda'_{\phi,j}$ is the number of parts of λ_ϕ of size at least j. In order that $\frac{|g^G\cap H|}{|g^G|}$ is nonzero, it is necessary that g has its conjugacy data satisfying $m_1(\lambda_{z-1}(g)) \geq 1$. Then $g^G \cap H$ is a single conjugacy class of H, with conjugacy data the same as for g except that a part of size 1 is removed from the partition corresponding to the polynomial $z-1$. Thus one sees that

$$\frac{|g^G \cap H|}{|g^G|} = \frac{|\mathrm{GL}(n-1,q)|}{|\mathrm{GL}(n,q)|}(1-1/q^{m_1(\lambda_{z-1}(g))})q^{2\lambda'_{z-1,1}(g)-1}.$$

Thus to find β, it is necessary study the maximum of the function

$$(1-1/q^{m_1(\lambda)})q^{2\lambda'_1}$$

among partitions λ of size at most n having at least 1 part equal to 1, but excluding the partition of size n that consists of all 1's. Here $m_1(\lambda)$ denotes the number of parts of λ of size 1, and λ'_1 denotes the number of parts of λ. It is straightforward to see that if $|\lambda| < n$, this function is maximized when $|\lambda| = n-1$ and λ consists of $n-1$ 1's. For $|\lambda| = n$ it is straightforward that the function is maximized for the partition consisting of 1 part of size 2 and $n-2$ parts of size 1. Comparing these two cases one sees that the maximum occurs for the first case. The first case occurs for $q > 2$ but can not occur for $q = 2$ (since $z-1$ is the only polynomial of degree 1 with nonzero constant term), and for $q = 2$ it is straightforward to see that the second case is the maximum. □

3. Symmetric groups

This section considers the Markov chain J in the case of the symmetric group and develops connections with card shuffling. We assume throughout that the reader is familiar with the Robinson–Schensted–Knuth (RSK) correspondence. See [**Sa**] for background on this topic.

Consider the symmetric group S_n. Let $\Pi = \{\epsilon_1-\epsilon_2, \cdots, \epsilon_{n-1}-\epsilon_n\}$ be a set of simple roots for the root system consisting of the $n(n-1)$ vectors $\epsilon_i - \epsilon_j$, where $1 \leq i \neq j \leq n$. The positive roots are $\epsilon_i - \epsilon_j$ where $i < j$ and the negative roots are those with $i > j$. The descent set of a permutation g consists of the elements in Π that g maps to negative roots. For $L \subseteq \Pi$, let X_L denote the set of permutations whose descent set is disjoint from L. It is well-known [**H**] that $|X_L| = n!/|S_L|$, where $|S_L|$ is the parabolic subgroup generated by adjacent transpositions corresponding to the roots in L. Consequently, if the $p_L \geq 0$ satisfy the equality $\sum_{L \subseteq \Pi} p_L = 1$, the element $\sum_{L \subseteq \Pi} \frac{p_L |S_L|}{n!} X_L$ defines a probability measure on the symmetric group.

Given an element $\sum_{g \in S_n} c_g g$ of the group algebra of the symmetric group, by the inverse element we mean $\sum_{g \in S_n} c_g g^{-1}$. It is known that the RSK correspondence associates the same partition to g and to g^{-1}, so when discussing the RSK correspondence one need not be concerned with whether we are considering an element in the group algebra or its inverse. The inverse of the element $\sum_{L \subseteq \Pi} \frac{p_L |S_L|}{n!} X_L$ can be thought of as a shuffle. For instance if $p_{\Pi - \{\epsilon_1 - \epsilon_2\}} = 1$, this shuffle is simply the top to random shuffle. One reason these shuffles are important is a result of Solomon [**So**] that states that $x_L x_K = \sum_{N \subseteq \Pi} a_{LKN} x_N$ for certain constants a_{LKN}. Thus one can at least in principle compute powers

$$\left(\sum_{L \subseteq \Pi} \frac{p_L |S_L|}{n!} X_L \right)^r,$$

which corresponds to understanding iterates of shuffles.

Now the main theorem of this section can be stated. Recall that the irreducible representations of the symmetric group S_n are parameterized by partitions λ of n.

Theorem 3.1. *Suppose that $p_L \geq 0$ satisfy $\sum_{L \subseteq \Pi} p_L = 1$. For $L \subseteq \Pi$, let $J[L]$ denote the Markov chain associated to the pair $G = S_n$ and $H = S_L$, and let $J[\vec{p}] = \sum_L p_L J[L]$ denote the mixture of the Markov chains $J[L]$. Then $J[\vec{p}]_1^r(\lambda)$ (the chance that the mixed chain started at the trivial representation is at the representation parameterized by λ after r steps) is equal to the chance that an element of the symmetric group distributed as $\left(\sum_{L \subseteq \Pi} \frac{p_L |S_L|}{n!} X_L\right)^r$ has* RSK *shape λ.*

Proof. From Proposition 2.3, the functions $\psi_C(\lambda)$ are a common orthonormal basis of eigenfunctions for the chains $J[L]$. Hence they are an orthonormal basis of eigenfunctions for the mixed chain $J[\vec{p}]$. This allows one to compute $J[\vec{p}]_1^r(\lambda)$ by the same method used in the proof of Theorem 1.1, and one concludes that it is equal to

$$\dim(\lambda) \left\langle \chi^\lambda, \left(\sum_L \frac{p_L |S_L|}{n!} \operatorname{Ind}_{S_L}^{S_n}(1) \right)^r \right\rangle .$$

As explained in the preliminary remarks of Section 4 of [**BBHT**], the coefficients a_{LKN} are related to tensor products of representations:

$$\operatorname{Ind}_{S_L}^{S_n}(1) \times \operatorname{Ind}_{S_K}^{S_n}(1) = \sum_{N \subseteq \Pi} a_{LKN} \operatorname{Ind}_{S_N}^{S_n}(1).$$

Letting $c_{N,r,\vec{p}}$ denote the coefficient of X_N in

$$\left(\sum_{L\subseteq\Pi} \frac{p_L|S_L|}{n!}X_L\right)^r,$$

it follows that $J[\vec{p}]_1^r(\lambda)$ is equal to

$$\dim(\lambda)\sum_{N\subseteq\Pi} c_{N,r,\vec{p}}\langle\chi^\lambda, \operatorname{Ind}_{S_N}^{S_n}(1)\rangle.$$

Letting μ denote the type of N (that is S_N is the direct product of symmetric groups whose sizes are the parts of the partition μ), the multiplicity of λ in $\operatorname{Ind}_{S_N}^{S_n}(1)$ is by definition the Kostka–Foulkes number $K_{\lambda\mu}$ discussed in [**Sa**]. Thus $J[\vec{p}]_1^r(\lambda)$ is equal to

$$\dim(\lambda)\sum_{\mu} K_{\lambda,\mu}\sum_{N:\mathrm{type}(N)=\mu} c_{N,r,\vec{p}}$$

where the sum is over all partitions μ of n.

Next it is necessary to show this is equal to the chance that an element of the symmetric group distributed as $\left(\sum_{L\subseteq\Pi}\frac{p_L|S_L|}{n!}X_L\right)^r$ has RSK shape λ. By the definition of $c_{N,r,\vec{p}}$, we know that

$$\left(\sum_{L\subseteq\Pi} \frac{p_L|S_L|}{n!}X_L\right)^r = \sum_{N\subseteq\Pi} c_{N,r,\vec{p}}X_N.$$

So it suffices to show that the number of summands of the element X_N (or equivalently the inverse of X_N) that the RSK correspondence maps to λ is $\dim(\lambda)K_{\lambda,\mathrm{type}(N)}$. But writing $S_N = S_{a_1}\times S_{a_2}\cdots\times S_{a_r}$ the summands of the inverse of x_N correspond (in an RSK shape preserving way) to words on the letters $\{1,\dots,r\}$ in which the letter l appears a_l times. But such words with RSK shape λ correspond to pairs (P,Q) of Young tableau with Q standard of shape λ and P semistandard of shape λ and content type(N). Since the number of these is $\dim(\lambda)K_{\lambda,\mathrm{type}(N)}$, the theorem is proved. □

Corollary 3.2 is an important consequence of Theorem 3.1.

Corollary 3.2. *Let $tv(r,\vec{p})$ denote the total variation distance between the probability measure $\left(\sum_{L\subseteq\Pi}\frac{p_L|S_L|}{n!}X_L\right)^r$ on the symmetric group and the uniform distribution on the symmetric group. Let π be the Plancherel measure of S_n. Then*

$$\frac{1}{2}\sum_{\lambda\in\mathrm{Irr}\,(S_n)}\left|\dim(\lambda)\left\langle\chi^\lambda, \left(\sum_{L\subseteq\Pi}\frac{p_L|S_L|}{n!}\operatorname{Ind}_{S_L}^{S_n}(1)\right)^r\right\rangle - \pi(\lambda)\right| \le tv(r,\vec{p}).$$

Proof. From the proof of Theorem 3.1, we know that

$$\frac{1}{2}\sum_{\lambda\in \mathrm{Irr}\,(S_n)}\left|\dim(\lambda)\Big\langle\chi^\lambda,\Big(\sum_{L\subseteq\Pi}\frac{p_L|S_L|}{n!}\mathrm{Ind}\,_{S_L}^{S_n}(1)\Big)^r\Big\rangle-\pi(\lambda)\right|$$

is equal to the total variation distance between the measure $J[\vec{p}]_1^r$ and the Plancherel measure of the symmetric group. Theorem 3.1 gives that this is equal to the total variation distance between the RSK pushforward of the measure $\big(\sum_{L\subseteq\Pi}\frac{p_L|S_L|}{n!}X_L\big)^r$ and the Plancherel measure. Since the Plancherel measure is the RSK pushforward of the uniform distribution on the symmetric group, the corollary follows. □

The significance of Corollary 3.2 is that it allows one to apply work on convergence rates of shuffles to the study of tensor products. We now give some examples showing that the bound of Corollary 3.2 can be much sharper than that of Theorem 1.1. Note that here we only treat examples with $p_L=1$ as these are the most natural from the viewpoint of decomposition of tensor products. The convergence rate of the RSK shape for other shuffles is considered in [**F2**], [**F3**].

Example 1 (The defining representation)**.** The first example is when $p_L=1$ for $L=\Pi-\{\epsilon_1-\epsilon_2\}$. Then $G=S_n$ and $H=S_{n-1}$. The representation theory problem in this case is the study of decompositions of the r-th tensor power of the defining (n-dimensional) representation, and the card shuffling problem is the r fold iteration of the top to random shuffle.

Consider the bound of Theorem 1.1. Letting $n_1(g)$ denote the number of fixed-points of g, it is clear that $|g^G\cap H|/|g^G|=n_1(g)/n$. Thus $\beta=1-\frac{2}{n}$. It follows that

$$\sum_{\lambda\in \mathrm{Irr}\,(S_n)}\left|\frac{\dim(\lambda)}{n^r}\langle\chi^\lambda,(\mathrm{Ind}\,_H^G(1))^r\rangle-\pi(\lambda)\right|\le\sqrt{n!}\Big(1-\frac{2}{n}\Big)^r.$$

Using Stirling's approximation [**Fe**]

$$n!\le\sqrt{2\pi}e^{-n+\frac{1}{12n}+(n+\frac{1}{2})\log(n)},$$

one sees that for $r>\frac{n\log(n)+2c}{2\log(\frac{1}{\beta})}$, this is at most

$$(2\pi)^{1/4}e^{r\log(\beta)+\frac{n\log(n)}{2}}\le(2\pi)^{1/4}e^{-c}.$$

For c fixed and large n, $\frac{n\log(n)+2c}{2\log(\frac{1}{\beta})}$ is roughly $\frac{1}{4}n^2\log n$.

The bound from Corollary 3.2 is much sharper. Indeed, it is known [**AD**] that for $r=n\log(n)+cn$, the total variation distance between r iterations of the top to random shuffle and the uniform distribution is at most e^{-c}, for $c\ge0,n\ge2$.

Next consider lower bounds for

$$\frac{1}{2}\sum_{\lambda\in \mathrm{Irr}\,(S_n)}\left|\frac{\dim(\lambda)}{n^r}\left\langle\chi^\lambda,(\mathrm{Ind}\,{}_H^G(1))^r\right\rangle-\pi(\lambda)\right|.$$

By Theorem 3.1, this is equal to the total variation distance between the RSK pushforward of r iterations of the top to random shuffle and the Plancherel measure. A result of Chapter 5 of [**U**] is that for large n at least $\frac{1}{2}n\log(n)$ iterations of the top to random shuffle are needed to randomize the length of the longest increasing subsequence (actually he states the result for the random to top shuffle, but this is the inverse of top to random). Since the longest increasing subsequence is a function of the RSK shape, it follows that

$$\frac{1}{2}\sum_{\lambda\in \mathrm{Irr}\,(S_n)}\left|\frac{\dim(\lambda)}{n^r}\left\langle\chi^\lambda,(\mathrm{Ind}\,{}_H^G(1))^r\right\rangle-\pi(\lambda)\right|$$

requires r at least $\frac{1}{2}n\log(n)$ to be small. Thus the upper bound on r in the previous paragraph is sharp to within a factor of two.

The next two examples generalize Example 1, but in different directions.

Example 2 ($S_{n-k}\subset S_n$)**.** This example is the case that

$$L=\Pi-\{\epsilon_1-\epsilon_2,\ldots,\epsilon_k-\epsilon_{k+1}\}$$

where $k\le n-1$. Then $G=S_n$ and $H=S_{n-k}$. The representation theory problem is to study the decomposition of the rth tensor power of $\mathrm{Ind}\,{}_{S_{n-k}}^{S_n}(1)$, and the relevant card shuffling is the top k to random shuffle, which proceeds by removing the top k cards from the deck and sequentially inserting them into random positions (this is equivalent to thoroughly mixing the top k cards and then riffling them with the rest of the deck — i.e., choosing a random interleaving).

First consider the bound of Theorem 1.1. Using the fact that two elements in a symmetric group are conjugate if and only if they have the same structure, and that a conjugacy class with n_i cycles of length i for all i has size $n!/\prod_i i^{n_i}n_i!$, one finds that $\beta=\frac{(n-k)(n-k-1)}{n(n-1)}$. By the same argument as Example 1, it follows that

$$\sum_{\lambda\in \mathrm{Irr}\,(S_n)}\left|\frac{\dim(\lambda)}{(n\cdots(n-k+1))^r}\left\langle\chi^\lambda,(\mathrm{Ind}\,{}_H^G(1))^r\right\rangle-\pi(\lambda)\right|\le(2\pi)^{1/4}e^{-c}$$

when $r>\frac{n\log(n)+2c}{2\log(\frac{1}{\beta})}$. For fixed c,k and large n, $\frac{n\log(n)+2c}{2\log(\frac{1}{\beta})}$ is roughly $\frac{n^2\log(n)}{4k}$.

The convergence rate of the card shuffling problem was studied in [**DFiP**], where it was shown that for k fixed and large n, the total variation distance is small for $r=\frac{n}{k}(\log(n)+c)$. Thus the bound from Corollary 3.2 is much sharper. The argument for the lower bound also generalizes, showing that r must be at least $\frac{1}{2k} n \log(n)$ for

$$\frac{1}{2} \sum_{\lambda \in \mathrm{Irr}\,(S_n)} \left| \frac{\dim(\lambda)}{(n(n-1)\cdots(n-k+1))^r} \langle \chi^\lambda, (\mathrm{Ind}\,{}_H^G(1))^r \rangle - \pi(\lambda) \right|$$

to be small.

The case of $k = \frac{n}{2}$ is also of interest. Then for n large, β is roughly $\frac{1}{4}$, and the upper bound of Theorem 1.1 shows that r roughly $\frac{n \log(n)}{2 \log(4)}$ is sufficient. Again the bound of Corollary 3.2 is shaper. To see this note that one wants an upper bound on the total variation distance between r iterates of the shuffle and the uniform distribution. The shuffle is a special case of the Bidigare–Hanlon–Rockmore walks on chambers of hyperplane arrangements, and a convenient upper bound for total variation distance is in [**BD**] (this bound is somewhat weaker than the bound in [**BHR**] but is easier to apply). In the case at hand the bound turns out to be $\binom{n}{2}\beta^r$, which shows that r roughly $\frac{2\log(n)}{\log(4)}$ is sufficient.

Example 3 (Action on k-sets)**.** The next example is the case that $p_L = 1$, where $L = \Pi - \{\epsilon_k - \epsilon_{k+1}\}$, and $1 \leq k \leq n/2$. Then $G = S_n$ and $H = S_k \times S_{n-k}$. The representation theory problem in this case is the study of decompositions of the r-th tensor power of the permutation representation on k-sets, and the card shuffling problem is the r-fold iteration of the shuffle that proceeds by cutting off exactly k cards, and then riffling them with the other $n-k$ cards (i.e., choosing a random interleaving).

First consider the bound of Theorem 1.1. The value of β is calculated in [**GM**] for $n \geq 5$ and shown to occur for the conjugacy class of transpositions, where it is $\frac{\binom{n-2}{k}+\binom{n-2}{k-2}}{\binom{n}{k}}$. For k fixed and large n, $\log(\frac{1}{\beta})$ is roughly $\frac{2k}{n}$, so that r slightly more than $\frac{n^2\log(n)}{4k}$ will make

$$\sum_{\lambda \in \mathrm{Irr}\,(S_n)} \left| \frac{\dim(\lambda)}{\binom{n}{k}^r} \langle \chi^\lambda, (\mathrm{Ind}\,{}_H^G(1))^r \rangle - \pi(\lambda) \right|$$

small.

Now consider the bound from Corollary 3.2. To apply it we require an upper bound on the total variation distance between the uniform distribution and r iterations of the shuffle that cuts off exactly k cards and riffles them with the rest of the deck. This shuffle too is a special case of the Bidigare–Hanlon–Rockmore walks on chambers of hyperplane arrangements, and a convenient upper bound for total variation distance is in [**BD**]. In the case at hand one can check that the total variation bound becomes

$$\binom{n}{2} \left(\frac{\binom{n-2}{k-2} + \binom{n-2}{k}}{\binom{n}{k}} \right)^r,$$

which is better than the bound $\frac{\sqrt{n!}}{2}\left(\frac{\binom{n-2}{k}+\binom{n-2}{k-2}}{\binom{n}{k}}\right)^r$ from Theorem 1.1. One concludes that r slightly more than $\frac{n\log(n)}{k}$ makes

$$\frac{1}{2}\sum_{\lambda\in \mathrm{Irr}\,(S_n)}\left|\frac{\dim(\lambda)}{\binom{n}{k}^r}\left\langle\chi^\lambda,(\mathrm{Ind}\,_H^G(1))^r\right\rangle-\pi(\lambda)\right|$$

small. Moreover, the argument for the lower bound in the other examples generalizes, showing that r must be at least $\frac{n\log(n)}{2k}$.

The case of $k=\frac{n}{2}$ is also of interest. Then for n large, β is roughly $\frac{1}{2}$, and the upper bound of Theorem 1.1 shows that r roughly $\frac{n\log(n)}{2\log(2)}$ is sufficient. Again the bound of Corollary 3.2 is shaper, showing that r roughly $\frac{2\log(n)}{\log(2)}$ is sufficient.

We remark that the fact that nonreversible Markov chains such as top to random are related to the reversible Markov chain J by means of Theorem 3.1 is quite mysterious. As a further result in this direction, we show that the Markov chains $J[\vec{p}]$ and $\sum_{L\subseteq\Pi}\frac{p_L|S_L|}{n!}X_L$ have the same set of eigenvalues (of course the multiplicities are different).

Proposition 3.3. *The Markov chain $J[\vec{p}]$ and the element $\sum_{L\subseteq\Pi}\frac{p_L|S_L|}{n!}X_L$ have the same set of eigenvalues.*

Proof. Since the chains $J[L]$ have a common basis of eigenvectors, the eigenvalues of $J[\vec{p}]$ are linear functions in the p's. Similarly [**BHR**] finds a formula for the eigenvalues of the element $\sum_{L\subseteq\Pi}\frac{p_L|S_L|}{n!}X_L$ and shows that they are linear in the p's. Hence it is enough to prove the result when $p_L=1$ for some L.

From Corollary 2.2 of [**BHR**], the eigenvalues of the element $\frac{|S_L|}{n!}X_L$ are indexed by permutations $g\in S_n$. Let μ be such that the orbits of S_L on $\{1,\dots,n\}$ are $\{1\cdots\mu_1\}$, $\{\mu_1+1\cdots\mu_1+\mu_2\}$, etc.; hence μ is a composition of n. A block ordered partition of the set $\{1,\dots,n\}$ is by definition a set partition with an ordering on the blocks of the partition. We say that a block ordered partition has type μ if the first block has size μ_1, the second block has size μ_2 and so on. The result of [**BHR**] is that the eigenvalue corresponding to g is the proportion of block ordered partitions of type μ that are fixed by g in the sense that each block is sent to itself. This is equivalent to requiring that each block is a union of cycles of g. Letting n_i denote the number of i-cycles of g, it follows that this proportion is

$$\frac{\mu_1!\mu_2!\cdots}{n!}\sum_{\substack{\sum_k a_i^{(k)}=n_i\\ \sum_i ia_i^{(k)}=\mu_k\text{ all }k}}\prod_{i\geq 1}\binom{n_i}{a_i^{(1)},a_i^{(2)},\dots}.$$

On the other hand, by Proposition 2.3, we know that the eigenvalues of J are parameterized by conjugacy classes C of S_n. Let n_i denote the number of cycles of length i for elements in the class C. Using the fact that $|C| = n!/\prod_i i^{n_i} n_i!$, it follows that

$$\frac{|C \cap S_L|}{|C|} = \frac{\prod_i i^{n_i} n_i!}{n!} \sum_{\substack{\sum_k a_i^{(k)} = n_i \\ \sum_i i a_i^{(k)} = \mu_k \text{ all } k}} \prod_k \frac{\mu_k!}{\prod_i i^{a_i^{(k)}} a_i^{(k)}!}.$$

This is equal to the expression of the previous paragraph, so the proof is complete. □

References

[AD] D. Aldous and P. Diaconis, *Shuffling cards and stopping times*, Amer. Math. Monthly, **93** (1986), 333–348, MR 0841111 (88a:60021), Zbl 0603.60006.

[BBHT] F. Bergeron, N. Bergeron, R.B. Howlett, and D.E. Taylor, *A decomposition of the descent algebra of a finite Coxeter group*, J. Alg. Combin., **1** (1992), 23–44, MR 1162640 (93g:20079), Zbl 0798.20031.

[BHR] P. Bidigare, P. Hanlon and D. Rockmore, *A combinatorial description of the spectrum for the Tsetlin library and its generalization to hyperplane arrangements*, Duke Math J., **99** (1999), 135–174, MR 1700744 (2000m:52032), Zbl 0955.60043.

[BD] K. Brown and P. Diaconis, *Random walks and hyperplane arrangements*, Ann. Probab., **26** (1998), 1813–1854, MR 1675083 (2000k:60138), Zbl 0938.60064.

[CF] S. Chatterjee and J. Fulman, *Stein's method for chi-squared approximation and spectral measure of Gelfand pairs*, preprint.

[DFiP] P. Diaconis, J. Fill and J. Pitman, *Analysis of top to random shuffles*, Combin. Probab. Comput., **1** (1992), 135–155, MR 1179244 (93f:60011), Zbl 0798.60008.

[DSa] P. Diaconis and L. Saloff-Coste, *Comparison techniques for random walk on finite groups*, Ann. Probab., **21** (1993), 2131–2156, MR 1245303 (95a:60009), Zbl 0790.60011.

[Fe] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. I, 2nd edition, John Wiley and Sons, Inc., New York; Chapman and Hall, Ltd., London, 1957, MR 0088081 (19,466a), Zbl 0077.12201.

[F1] J. Fulman, *Stein's method and Plancherel measure of the symmetric group*, to appear in Trans. Amer. Math. Soc.

[F2] J. Fulman, *Applications of symmetric functions to cycle and increasing subsequence structure after shuffles*, J. Algebraic Combin., **16** (2002), 165–194, MR 1943587 (2004b:05200), Zbl 1012.05154.

[F3] J. Fulman, *A card shuffling analysis of deformations of the Plancherel measure of the symmetric group*, Electron. J. Combin., **11** (2004), Research Paper 21, 15 pp. (electronic), MR 2035315.

[F4] J. Fulman, *Stein's method and nonreversible Markov chains*, IMS Lecture Notes, **46** (2004), 69–77.

[F5] J. Fulman, *Stein's method, Jack measure, and the Metropolis algorithm*, to appear in J. Combin. Theory Ser. A, math.CO/0311290.

[F6] J. Fulman, *Martingales and character ratios*, to appear in Trans. Amer. Math. Soc., math.RT/0402409.

[FH] W. Fulton and J. Harris, *Representation Theory: A First Course*, Graduate Texts in Mathematics, **129**, Springer-Verlag, New York, 1991, MR 1153249 (93a:20069), Zbl 0744.22001.

[GK] R. Guralnick and W. Kantor, *Probabilistic generation of finite simple groups*, J. Algebra, **234** (2000), 743–792, MR 1800754 (2002f:20038), Zbl 0973.20012.

[GM] R. Guralnick and K. Magaard, *On the minimal degree of a primitive permutation group*, J. Algebra, **207** (1998), 127–145, MR 1643074 (99g:20014), Zbl 0911.20003.

[H] J. Humphreys, *Reflection Groups and Coxeter Groups*, Cambridge Studies in Advanced Mathematics, **29**, Cambridge University Press, Cambridge, 1990, MR 1066460 (92h:20002), Zbl 0725.20028.

[I] I.M. Isaacs, *Character Theory of Finite Groups*, Pure and Applied Mathematics, **69**, Academic Press, New York, 1976, MR 0460423 (57 #417), Zbl 0337.20005.

[LSh] M. Liebeck and A. Shalev, *Simple groups, permutation groups, and probability*, J. Amer. Math. Soc., **12** (1999), 497–520, MR 1639620 (99h:20004), Zbl 0916.20003.

[M] I. Macdonald, *Symmetric Functions and Hall Polynomials*, Oxford Mathematical Monographs, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1995, MR 1354144 (96h:05207), Zbl 0824.05059.

[Sa] B. Sagan, *The Symmetric Group. Representations, Combinatorial Algorithms, and Symmetric Functions*, The Wadsworth & Brooks/Cole Mathematics Series, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1991, MR 1093239 (93f:05102), Zbl 0823.05061.

[So] L. Solomon, *A Mackey formula in the group ring of a Coxeter group*, J. Algebra, **41** (1976), 255–264, MR 0444756 (56 #3104), Zbl 0355.20007.

[St] R. Stanley, *Generalized riffle shuffles and quasisymmetric functions*, Ann. Comb., **5** (2001), 479–491, MR 1897637 (2003c:60016), Zbl 1010.05078.

[U] J.C. Uyemura-Reyes, *Random Walk, Semidirect Products, and Card Shuffling*, Ph.D. Thesis, Stanford University, 2002.

[W] G.E. Wall, *On conjugacy classes in the unitary, symplectic, and orthogonal groups*, J. Austral. Math. Soc., **3** (1963), 1–62, MR 0150210 (27 #212), Zbl 0122.28102.

Received July 3, 2003 and revised February 25, 2004. The author was partially supported by NSA grant MDA 904-03-1-0049.

Department of Mathematics
301 Thackeray Hall
University of Pittsburgh
Pittsburgh PA 15260
E-mail address: fulman@math.pitt.edu