Representation Theory of the Symmetric Group: The Verhsik-Okounkov Approach

Adam B Block and Roger Van Peski

July 2018

The representation theory of S_n has been well-studied classically. It has been known essentially since the dawn of the study of groups what the conjugacy classes of the symmetric groups are, and so the number of these representations was discovered early. As [OV] note, there are two main classical approaches to the study of the simple representations of S_n. The first, and most classical, is the manipulation of Young tableaux and Specht modules. The second is to study the representations of $\mathrm{GL}_{\mathbb{C}}(n)$ using Schur functors and then apply Schur-Weyl duality to get the representations of S_n. While both methods have their merits, the method of [OV] has a few primary advantages, which they list. We focus on the fact that this approach provides a natural way to use the inductive structure of this family. As such, instead of an *ad hoc* attempt to find for some simple representation of S_n the restrictions to S_k or induction to S_m for some $k \leq n \leq m$, these functorial constructions fall naturally out of the theory developed. Moreover, the key step in the proof is to show that the spectrum of a special algebra with respect to certain distinguished elements is given by integer vectors satisfying certain technical conditions coming directly from Coxeter relations. This further justifies the approach exposited below, as the family S_n is a family of coxeter groups. In the sequel, we occasionally refer to [FS], an exhaustive account of the details of [OV], especially when the latter decides to elide over proofs of nontrivial results.

We proceed in three parts. In the first, we introduce the topic and prove that the family S_n has simple branching. In the second part, we introduce the YJM elements and describe the representation theory of the degenerate affine Hecke algebra $H(2)$, using this analysis to show that the spectrum consists of integers. In the third lecture, we prove the main theorem of classification and discuss some applications.

The only prerequisites are some knowledge of the representation theory of finite groups over $\mathbb{C}$. A quick review is provided here.

1 Gelfand-Tsetlin Algebras and Simple Branching

Some of the results below apply to more general families of groups or algebras than just the family of symmetric groups. As such, we define

Definition 1. An inductive family of groups is $\{G_i\}_{i\in\mathbb{N}}$ such that $G_i < G_{i+1}$ for groups G_i. For any i, let $\hat{G}_i$ be the set of simple representations of G_i (up to isomorphism). Let the branching graph of the family $\{G_i\}$ be defined as follows. Let the vertices be

$$\bigcup_i \hat{G}_i$$

and call the $V^\lambda \in \hat{G}_i$ the i^{th} level. For $V^\mu \in G_{i-1}$ and $V^\lambda \in G_i$, the number of edges connecting V^μ and V^λ is the multiplicity of V^μ in $\mathrm{Res}_{S_n-1} V^\lambda$, or $\dim \mathrm{Hom}_{S_{n-1}}(V^\mu, V^\lambda)$. We say that the family has simple branching when there is at most one edge connecting any two vertices.

We need some notation. We say $V^\mu \nearrow V^\lambda$ if there is at least one edge connecting V^μ and V^λ. For $V^\mu \in \hat{G}_j < \hat{G}_i$ we say that $V^\mu \subset V^\lambda$ if there is an ascending path in the branching graph from V^μ to V^λ. We may write this path as $\mu \nearrow \lambda_1 \nearrow \cdots \nearrow \lambda$. Consider the following trivial example:

Example 2. Let $G_i = Z_{p^i}$ the cyclic group of order p^i. Let $G_i < G_{i+1}$ be embedded by $1 \mapsto p$. Because G_i is abelian, all simple representations are characters, given by $\chi_a : G_i \to \mathbb{C}$ by $\chi_a(1) = \zeta_{p^i}^a$. Note that the restriction of χ_a to G_{i-1} is given by $\chi_a(p) = \zeta_{p^i}^{ap} = (\zeta_{p^i}^p)^a = \zeta_{p^{i-1}}^a$. Thus χ_a restricts to $a \pmod{p^{i-1}}$. Thus the branching graph at the i^{th} level connects $\chi_a \in \hat{G}_i$ with one edge to each $\chi_{a+p^i r} \in \hat{G}_{i+1}$. Note that this family has simple branching.

The purpose of the current talk is to show that S_n has simple branching, which is the first main component in finding the simple representations. The reason for this is as follows. If a family has simple branching, there is a canonical decomposition

$$\mathrm{Res}_{G_{i-1}}^{G_i} V^\lambda = \bigoplus_{\mu \nearrow \lambda} V^\mu$$

For each of the $V^\mu \in \hat{G}_{i-1}$, we may repeat the process. By induction, we obtain

$$\mathrm{Res}_1^{G_n} V^\lambda = \bigoplus_P V_P$$

Where P is a path $\lambda_0 \nearrow \lambda_1 \nearrow \cdots \nearrow \lambda_n = \lambda$ and we are summing over all of these paths that terminate in λ. Note that V_P is a simple $G_0 = \{1\}$ representation, so is one dimensional. We may choose a G-invariant inner product and a unit vector $v_P \in V_P$ for all paths P, giving us a basis for V^λ.

Definition 3. The Gelfand-Tsetlin (GZ) basis for a simple representation V^λ of G_n for $\{G_n\}$ an inductive family of groups with simple branches is $\{v_P\}$ where P varies over all paths $\lambda_0 \nearrow \cdots \nearrow \lambda$.

Now, note that $v_P \in V^{\lambda_i}$ for all $0 \leq i \leq n$ in the path P and so, by irreducibility, we have $\mathbb{C}[G_i] \cdot v_P = V^{\lambda_i}$ for all i. If $v' \in V^\lambda$ were another vector with this property, then $v' \in V^{\lambda_i}$ for all i and so, by simple branching, v' and v_P differ only by a scalar.

With the GZ-basis defined, we now define the GZ-algebra.

Definition 4. Let $Z(n)$ be the center of $\mathbb{C}[G_n]$, the set of all elements $\alpha \in \mathbb{C}[G_n]$ that commute with all other elements. Define the Gelfand-Tsetlin (GZ) algebra to be

$$GZ(n) = \langle Z(1), Z(2), \ldots, Z(n) \rangle \subset \mathbb{C}[G_n]$$

the subalgebra generated by all of these centers.

We have the following obvious result:

Lemma 5. *For all n, $GZ(n)$ is commutative.*

Proof. We induct. The statement is clear for $n = 1$. Supposing it holds for $GZ(n-1)$, we note that $GZ(n) = \langle GZ(n-1), Z(n) \rangle$ and both $GZ(n-1)$ and $Z(n)$ are commutative. But by the definition of center, because $GZ(n-1) \subset \mathbb{C}[G_{n-1}] \subset \mathbb{C}[G_n]$, we have that $GZ(n-1)$ and $Z(n)$ commute. ■

Currently, it might appear that the GZ-basis and -algebra are related by no more than name. The following proposition will dispell all such notions.

Proposition 6. *The GZ-algebra consists precisely of all operators diagonal with respect to the GZ-basis.*

Before we can prove this, we need a classical result:

Lemma 7 (Wedderburn)**.** *As an algebra, we have*

$$\mathbb{C}G_n = \bigoplus_\lambda \mathrm{End}(V^\lambda)$$

where the sum is over all simple representations V^λ.

Proof. Let $\varphi : \mathbb{C}G_n \to \bigoplus \mathrm{End}(V^\lambda)$ be given by $\varphi(g) = (\rho_\lambda(g))_\lambda$ and extended by linearity. This is clearly a morphism of algebras. Note that $\dim_{\mathbb{C}} \mathbb{C}G_n = |G_n|$ and the dimension over the right side is the sum of the squares of the dimensions of the irreducible representations of G_n, which we know to be $|G_n|$. Thus it suffices to know that φ is injective. To see this, note that the regular representation is the sum of all irreducible representations with multiplicities given by their dimensions. Thus if $\varphi(\alpha) = \varphi(\alpha')$ then α and α' must act identically on the regular representation, but the regular representation is faithful. ■

Proof of Proposition 6. By Lemma 7, we may choose $\pi_{\lambda_i} \in \mathbb{C}G_i$ that is projection to V_{λ_i}, i.e., is the identity in $\mathrm{End}(V^{\lambda_i})$ and zero elsewhere. Then it is clear that $\pi_{\lambda_i} \in Z(i) \subset GZ(n)$. Thus we may let $\pi_P = \pi_{\lambda_0}\pi_{\lambda_1}\cdots\pi_\lambda \in GZ(n)$. Then π_P is projection to V_P. Let $GZ'(n)$ be the space of operators diagonal with respect to the GZ-basis. Then $GZ'(n) = \langle \pi_P \rangle_P$ so $GZ'(n) \subset GZ(n)$. But the set of diagonal matrices is a maximal torus, so $GZ'(n) = GZ(n)$ by maximality. ■

We now reach the key criterion for simple branching. Recall that if $N \subset M$ is a subalgebra, then the centralizer $Z(M, N)$ is the set of all elements in M that commute with all elements in N. We have:

Proposition 8. *We have that $Z(M, N)$ is commutative if and only if for any $V^\mu \in \hat{N}$, $V^\lambda \in \hat{M}$, the multiplicity of V^μ in $\mathrm{Res}^M_N V^\lambda$ is at most one.*

Proof. Because M is semi-simple and the statements above factor through direct summation, we may assume that M is simple. Let V be the nontrivial simple M-module such that $M = \mathrm{End}(V)$. Then we may consider ${}_NV = \mathrm{Res}^M_N V$ and break down ${}_NV$ as a sum of isotypical components:

$$ {}_NV = \bigoplus_i V_i^{m_i} $$

Then the N-linear endomorphisms are exactly given by $Z(M, N)$ by the Double Centralizer Theorem (see [Pie]). Thus we have

$$ Z(M, N) = \mathrm{End}({}_NV) = \bigoplus_i \mathrm{End}(V_i^{m_i}) $$

Now note that the right hand side is a sum of matrix rings and so commutes if and only if $m_i \leq 1$ for all i, as desired. ■

We will use Proposition 6 to show that S_n has simple branching. We need a few results first. We recall that a $\mathbb{C}^*$-algebra is an algebra over $\mathbb{C}$ with an anti-symmetric conjugate linear involution $*$. We have

Lemma 9. *A $\mathbb{C}^*$-algebra is commutative if and only if all elements are normal. If all real elements in A a $\mathbb{C}^*$ algebra are self-conjugate, then A is abelian.*

Proof. To show the first, one direction is trivial, so suppose all elements are normal and let A be a $\mathbb{C}^*$ algebra. For any element $a \in A$, we have

$$ a = \frac{a + a^*}{2} + i\frac{a - a^*}{2i} $$

Let A^{sa} be the subspace of self-adjoint elements. Then we have $A = A^{sa} \oplus iA^{sa}$ by the above decomposition. Let $a, b \in A^{sa}$. Then we have $(a + ib)^* = a^* - ib^* = a - ib$. Thus by all elements being normal we have $(a + ib)(a - ib) = (a - ib)(a + ib)$. Thus we have

$$ \begin{aligned} a^2 - iab + iba + b^2 = a^2 + iab - iba + b^2 \qquad &\text{or} \\ ab = ba& \end{aligned} $$

Thus all self adjoint elements commute and so $xy = (x_1 + ix_2)(y_1 + iy_2) = (y_1 + iy_2)(x_1 + ix_2) = yx$.

To prove the second statement, we have $A = A_{\mathbb{R}} \oplus iA_{\mathbb{R}}$. If $a, b \in A_{\mathbb{R}}$ we have $ab \in A_{\mathbb{R}}$ so

$$ ab = (ab)^* = b^*a^* = ba $$

and so $A_{\mathbb{R}}$ is abelian. But then so is A. ■

Now we introduce a lemma specific to S_n and we have

Lemma 10. *For all permutations $\sigma \in S_n$, there is an element $\tau \in S_{n-1}$ such that $\sigma^{-1} = \tau\sigma\tau^{-1}$*

Proof. Note that σ, σ^{-1} are of the same cycle type so they are clearly conjugate in S_n. Let $\sigma' \in S_{n-1}$ be the element obtained from σ by removing n from whichever cycle in which it appears and leaving everything else the same. Then we have $\tau \in S_{n-1}$ such that $\tau\sigma'\tau^{-1} = \sigma'^{-1}$. Now consider $\tau \in S_n$ as fixing n. Then $\tau\sigma\tau^{-1} = \sigma^{-1}$. ∎

Example 11. As an example of the above method, consider $\sigma = (1\,2\,3)(4\,5) \in S_5$. Then $\sigma^{-1} = (1\,3\,2)(4\,5)$. We delete 5 from σ and get $\sigma' = (1\,2\,3)$ which is conjugate to σ'^{-1} by $\tau = (2\,3)$. Then it is elementary to check that $\sigma^{-1} = \tau\sigma\tau^{-1}$.

We are now ready to prove our final result, giving simple branching of S_n.

Proposition 12. *The centralizer $Z(\mathbb{C}S_n, \mathbb{C}S_{n-1})$ is commutative.*

Proof. By Lemma 9, it suffices to show that if $\alpha \in Z(n, n-1)$ is real then $\alpha^* = \alpha$, where $(ag)^* = \overline{a}g^{-1}$. But let $\alpha = \sum a_g g$ with $a_g \in \mathbb{R}$. We know that α commutes with $\mathbb{C}S_{n-1}$ and by Lemma 10 we may choose $\tau_g \in S_{n-1} \subset \mathbb{C}S_{n-1}$ such that $\tau_g g \tau_g^{-1} = g^{-1} = (g_i)^*$. But note that α commutes with τ_g so $\tau_g \alpha \tau_g^{-1} = \alpha$ and so, because a_g is real, we have $(a_g g)^* = a_g g^{-1}$. But if $\alpha = \tau_g \alpha \tau_g^{-1}$ then $a_g = a_{g^{-1}}$ and so $\alpha^* = \alpha$ as desired. ∎

We now, finally, have our result.

Theorem 13. *The branching graph of the symmetric groups is simple.*

Proof. Combine Propositions 8 and 12. ∎

2 Young-Jucys-Murphy elements and representations of $H(2)$

Recall: last lecture, we proved that $\mathrm{Res}^{S_n}_{S_{n-1}} V^\lambda$ splits into distinct irreducibles. In this lecture we will introduce the Young-Jucys-Murphy (YJM) elements of $\mathbb{C}[S_n]$, prove they generate $GZ(n)$, and study their action on the irreps of S_n. This gives an alternate proof that $\mathrm{Res}^{S_n}_{S_{n-1}} V^\lambda$ splits into distinct irreducibles, but also allows further analysis of the explicit action of S_n on its irreps.

Definition 14. The YJM element $X_i \in \mathbb{C}[S_i]$ is given by

$$X_i := (1\,i) + (2\,i) + \ldots + (i-1\,i) = \sum_{\sigma \in S_i \text{ a 2-cycle}} \sigma - \sum_{\sigma \in S_{i-1} \text{ a 2-cycle}} \sigma \tag{1}$$

for $i > 1$, and when $i = 1$ by $X_1 = 0$.

The second part of the definition serves to motivate the YJM element a little bit. We wish to show that the YJM elements generate $GZ(n) = \langle Z(1), \ldots, Z(n) \rangle$, so certainly the YJM elements should be related to central elements in some way. We will show that $\mathbb{C}[S_n]$ is algebraically generated by elements of the form $\sum_{\sigma \in S_n \text{ a } k\text{-cycle}} \sigma$, of which the simplest nontrivial case is $\sum_{\sigma \in S_n \text{ a 2-cycle}} \sigma$. It might be natural to look for elements X_i which lie in $Z(i)$, but since we wish to in some way induct from $GZ(n) = \langle Z(1), \ldots, Z(n) \rangle$ to $GZ(n+1) = \langle Z(1), \ldots, Z(n+1) \rangle$, so hopefully it is at least believable at this stage that taking the difference of the simplest nontrivial generator of $Z(n+1)$ with the simplest nontrivial generator of $Z(n)$ might be a good choice. We will use the fact that X_i is a difference of central elements in $\mathbb{C}[S_i]$ and $\mathbb{C}[S_{i+1}]$ quite frequently in what follows.

Theorem 15. *For any $n \geq 2$, $GZ(n) = \langle X_1, \ldots, X_n \rangle$.*

Proof. $\supset$ follows immediately since X_i is a difference of an element of $Z(i)$ and $Z(i-1)$ as discussed above, and $Z(i), Z(i-1) \subset GZ(n)$.

For $\subset$ we induct. For the base case, $\mathbb{C}[S_2]$ is commutative and $X_2 = (12)$ generates it, so assume $GZ(n-1) = \langle X_1, \ldots, X_{n-1} \rangle$. Hence we must prove $Z(n) \subset \langle GZ(n-1), X_n \rangle$, and clearly it suffices to show $Z(n) \subset \langle Z(n-1), X_n \rangle$. It is easy to check that for any finite group G, $Z(\mathbb{C}[G])$ is generated as a $\mathbb{C}$-vector space by elements of the form $K_C := \sum_{g \in C} g$, as C ranges over all conjugacy classes. Hence we must show that for any partition λ, the associated element K_C lies in $\langle GZ(n-1), X_n \rangle$.

Define the *support* of a permutation $\sigma \in S_n$ to be the number of elements in $\{1, \ldots, n\}$ which it does not fix, and define the support of a conjugacy class to be the size of the support of any permutation it contains (note that this only refers to a single element of the conjugacy class rather than the class itself, since each element is moved by some element of any non-identity conjugacy class). We show that each K_C lies in $\langle GZ(n-1), X_n \rangle$ by inducting on the size of the support. For the base case where C is the conjugacy class of the identity, $K_C = 1$ which clearly lies in $\langle GZ(n-1), X_n \rangle$, so suppose for the inductive hypothesis that $K_C \in \langle GZ(n-1), X_n \rangle$ for all C of support $\leq k$. Let C_λ be the conjugacy class corresponding to $\lambda = (\lambda_1, \ldots, \lambda_t, 1, \ldots, 1)$, where λ is any partition such that C_λ has support $k+1$, and has more than one nontrivial cycle (so $t \geq 2$). Letting $C_{\ell,n}$ be the conjugacy class of ℓ-cycles in S_n, the element

$$K_{C_{\lambda_1,n}} \cdot K_{C_{\lambda_2,n}} \cdots K_{C_{\lambda_t,n}} = K_{C_\lambda} + (\text{terms of support} \leq k),$$

where the second term denotes a linear combination of elements K_C for C of support $\leq k$. By inductive hypothesis, the latter terms lie in $\langle Z(n-1), X_n \rangle$, and each term $K_{C_{\lambda_i,n}}$ does as well, using the fact that $\sum_i \lambda_i = k+1$ and each λ_i is at least 1. Hence $K_{C_\lambda} = K_{C_{\lambda_1,n}} \cdots K_{C_{\lambda_t,n}} - (\text{terms of support} \leq k) \in \langle Z(n-1), X_n \rangle$.

It remains to show that $K_C \in \langle Z(n-1), X_n \rangle$ where C is the conjugacy class of all $(k+1)$-cycles. For this, we have by inductive hypothesis that $K_{C_{k,n}} \in \langle Z(n-1), X_n \rangle$, and clearly $K_{C_{k,n-1}} \in \langle Z(n-1), X_n \rangle$, hence their difference $\sum_{1 \leq i_1, \ldots, i_{k-1} \leq n \text{ distinct}} (i_1, \ldots, i_{k-1}, n)$ lies in $\langle Z(n-1), X_n \rangle$. Now,

$$X_n \cdot \sum_{1 \leq i_1, \ldots, i_{k-1} \leq n \text{ distinct}} (i_1, \ldots, i_{k-1}, n) \tag{2}$$

is a product of terms $(j\, n)(i_1\, \ldots, i_{k-1}\, n)$, for $i_1, \ldots, i_{k-1}$ distinct and $j \neq n$ possibly equal to one of the i_ℓ. Such terms fall into two cases:

- If $j \notin \{i_1, \ldots, i_{k-1}\}$, then $(j, n)(i_1\, \ldots\, i_{k-1}\, n) = (i_1\, \ldots\, i_{k-1}\, j\, n)$ is a single $(k+1)$-cycle.
- If $j = i_\ell$ for some ℓ, then $(j\, n)(i_1\, \ldots\, i_{k-1}\, n) = (n\, i_1\, \ldots\, i_{\ell-1})(i_\ell\, \ldots\, i_{k-1})$ is a product of two cycles, though one cycle may be trivial (or both may be trivial when $k = 2$ as we saw before). The size of the support is k or $k-1$.

Hence (2) is equal to

$$C_{k+1,n} + (\text{terms of support} \leq k), \tag{3}$$

so $C_{k+1,n} \in \langle Z(n-1), X_n \rangle$, completing the proof. ∎

We now relate the YJM elements to the GZ basis, starting with the claim that the GZ basis is a simultaneous eigenbasis for $GZ(n)$. This follows because each element of $GZ(n)$ is a linear combination of elements $\prod_i x_i$ where $x_i \in Z(n_i)$ for some sequence of integers $1 \leq n_i \leq n$, and each x_i acts by a scalar on any irrep of S_{n_i} and hence acts as a scalar on the GZ basis.

Definition 16. For a vector v in the GZ basis of some irreducible representation of S_n, we say its *weight* is $\alpha(v) = (a_1, \ldots, a_n) \in \mathbb{C}^n$ where a_i is the eigenvalue of X_i on v. We denote the set of all weights by $\mathrm{Spec}(n)$.

The following definition and proposition begin to get to the heart of why the YJM elements capture the inductive structure of the chain $S_1 \subset S_2 \subset \ldots$. To motivate the definition, note that if $s_i = (i\, i+1)$ is a Coxeter generator, s_i commutes with X_j for $j \neq i, i+1$; for $i > j+1$ this is immediate, and for $i < j$ we have that $(i\, i+1)((i\, n) + (i+1\, n)) = (i+1\, n) + (i\, n)$. Furthermore, $s_i X_i + 1 = X_{i+1} s_i$.

Definition 17. The *degenerate affine Hecke algebra* $H(2)$ is the $\mathbb{C}$-algebra with generators Y_1, Y_2, s satisfying relations $s^2 = 1, Y_1Y_2 = Y_2Y_1$, and $sY_1 + 1 = Y_2s$.

Proposition 18. *$\mathbb{C}[S_n]$ is generated by $\mathbb{C}[S_n]$ and a copy of $H(2)$ with generators $Y_1 = X_{n-1}, Y_2 = X_n, s = s_n$. (note that there are also relations between elements of this copy of $H(2)$ and of $\mathbb{C}[S_n]$, e.g. the Coxeter relations $s_{n-1}s_ns_{n-1} = s_ns_{n-1}s_n$).*

This motivates the study of the representations of $H(2)$ ($H(2)$-modules), which in view of the previous proposition can be used to understand the action of $GZ(n) = \langle X_1, \ldots, X_n\rangle$ on irreps of S_n, and hence understand the possible weight vectors. We note that each nonzero simple $H(2)$-module contains a simultaneous eigenvector v of Y_1 and Y_2, and by the defining relations of $H(2)$ we see that the module must be spanned by v and sv. In particular, all simple $H(2)$-modules have dimension at most 2.

Proposition 19. *Let $V = \mathrm{Span}(v, s\cdot v)$ be a simple representation of $H(2)$ such that $Y_1 \cdot v = a_1v$ and $Y_2v = a_2v$ and one (and thus both) of Y_1, Y_2 are diagonalizable on V. Then*

1. *$a_1 \neq a_2$.*
2. *V is one dimensional if and only if $a_2 = a_1 \pm 1$.*
3. *If $a_2 \neq a_1 \pm 1$ then, with respect to the basis $\{v, sv\}$*

$$Y_1 = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}, \qquad Y_2 = \begin{pmatrix} a_2 & 0 \\ 0 & a_1 \end{pmatrix}, \qquad s = \begin{pmatrix} \frac{1}{a_2-a_1} & 1 - \frac{1}{(a_2-a_1)^2} \\ 1 & \frac{1}{a_1-a_2} \end{pmatrix}.$$

Proof. Since Y_1 and Y_2 commute, they have a simultaneous eigenbasis in any simple $H(2)$-module. Letting v be a simultaneous eigenvector, we have that $Y_1v = a_1v, Y_2v = a_2v$ for some a_1, a_2. Then by the relation $sY_1 + 1 = Y_2s$,

$$Y_1(sv) = (sY_2 - 1)v = a_2(sv) - v \tag{4}$$
$$Y_2(sv) = (sY_1 + 1)v = a_1(sv) + v. \tag{5}$$

Hence V is spanned by v, sv, so its dimension is ≤ 2. If V is one-dimensional, then there is a constant $c \in \mathbb{C}$ for which $sv = cv$. Because $s^2 = 1$, $c^2 = 1$ so $c = \pm 1$. We now have that $Y_1(sv) = a_1cv$ and $Y_2(sv) = a_2cv$, so (4) becomes $a_1cv = a_2cv - v$, so $c(a_2 - a_1) = 1$. Hence $a_2 - a_1 = \pm 1$. Conversely, suppose $a_2 = a_1 \pm 1$, and so we have from $1 = Y_2s - sY_1$ that, by applying this equality to v,

$$v = 1 \cdot v = (Y_2s - sY_1)v = a_2(sv) - a_1(sv) = \pm sv$$

This proves the backward direction of (2).

If V is 2-dimensional then (4) yields

$$Y_1 = \begin{pmatrix} a_1 & -1 \\ 0 & a_2 \end{pmatrix}, Y_2 = \begin{pmatrix} a_2 & 1 \\ 0 & a_1 \end{pmatrix}, s = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{6}$$

in the basis v, sv. If $a_1 = a_2$ then Y_1, Y_2 cannot be diagonalized, but since the X_i are diagonalizable on irreps V^λ of S_n, such a representation cannot occur in the decomposition of V^λ into irreps of $\langle X_i, X_{i+1}, s_i\rangle$. This proves (1).

Now, if $a_2 \neq a_1 \pm 1$ then $\{v, s\cdot v\}$ span V still and thus we may write the operators with respect to this basis as

$$s = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad Y_1 = \begin{pmatrix} a_1 & -1 \\ 0 & a_2 \end{pmatrix}, \qquad Y_2 = \begin{pmatrix} a_2 & 1 \\ 0 & a_1 \end{pmatrix}$$

by the identities above. But we have the matrices for Y_1, Y_2 are diagonalizable if and only if $a_1 \neq a_2$, in which case we have an eigenvector $(1,\ 0)$ corresponding to eigenvalues of a_1 and a_2 for Y_1 and Y_2 respectively, and $(1,\ a_1 - a_2)$ corresponding to eigenvalues of a_2 and a_1 respectively. Thus we immediately get that $v' = v + (a_1 - a_2)(sv)$ is an eigenvector of both Y_2, Y_1 with eigenvalues a_2, a_1 respectively.

If $|a_1 - a_2| = 1$ then $v \pm sv$ is an eigenvecor of Y_1 and Y_2; but it is clearly also an eigenvector of s, hence s, Y_1, Y_2 commute, so V is one-dimensional. This proves the forward direction of (2). ∎

Corollary 20. *Let $\alpha = (a_1, \ldots, a_n) \in \mathrm{Spec}(n)$ be a weight with corresponding eigenvector v_α in the GZ basis. Then*

1. *$a_i \neq a_{i+1}$ for all i.*

2. *For all i, $a_{i+1} = a_i \pm 1$ if and only if $s_i v_\alpha = \pm v_\alpha$.*

3. *If $a_{i+1} \neq a_i \pm 1$ then $\alpha' = s_i\alpha = (a_1, a_2, \ldots, a_{i-1}, a_{i+1}, a_i, a_{i+2}, \ldots, a_n) \in \mathrm{Spec}(n)$. Moreover,*

$$v_{\alpha'} = s_i v_\alpha - \frac{1}{a_{i+1} - a_i} v_\alpha$$

and $\mathrm{Span}(v_\alpha, v_{\alpha'})$ is invariant under the action of X_i, X_{i+1}, s_i, with these actions represented by

$$X_i = \begin{pmatrix} a_i & 0 \\ 0 & a_{i+1} \end{pmatrix}, \qquad X_{i+1} = \begin{pmatrix} a_{i+1} & 0 \\ 0 & a_i \end{pmatrix}, \qquad s_i = \begin{pmatrix} \frac{1}{a_{i+1}-a_i} & 1 - \frac{1}{(a_{i+1}-a_i)^2} \\ 1 & \frac{1}{a_i - a_{i+1}} \end{pmatrix}.$$

Proof. Note that we have X_i, X_{i+1}, s_i satisfy the relations to define $H(2)$. Also note that by definition, X_i, X_{i+1} are simultaneously diagonalizable, with respect to the GZ-basis. Thus all results follow immediately from Proposition 19 except for the last statement in (3). To see that $\alpha' = s \cdot \alpha \in \mathrm{Spec}(n)$, we note that for $j \notin \{i, i+1\}$ we have $X_j s_i = s_i X_j$ and so $X_j v_{\alpha'} = X_j s \cdot v_\alpha = s X_j \cdot v_\alpha = a_j v_{\alpha'}$. Thus $\alpha' \in \mathrm{Spec}(n)$ and $v' = v_{\alpha'}$ as desired. ■

Remark 21. The reader who has seen the representation theory of Lie algebras or Lie groups may note that the method of finding subalgebras isomorphic to $H(2)$ and studying how they act on representations of S_n is very similar to the way representations of Lie algebras (resp. Lie groups) are built from copies of representations of $\mathfrak{sl}_2$ (resp. With point (3) of Corollary 20 in mind, we have the following definition:

Definition 22. A permutation $\sigma \in S_n$ is admissible for $\alpha \in \mathbb{C}^n$ if it is in the subgroup of S_n generated by all of the s_i such that $a_{i+1} \neq a_i \pm 1$.

The reason for this definition should be clear: the set of admissible permutations acts on $\mathrm{Spec}(n)$ by (3) in Corollary 20 (and we will see later that it acts on $\mathrm{Cont}(n)$ too).

3 Main Theorems

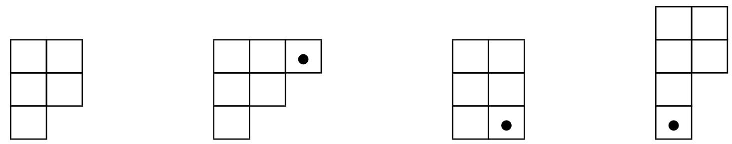
We now change tacks slightly to introduce some of the requisite combinatorics. After this short interlude, we will be prepared to prove the main theorem. Recall that given a partition λ of some integer n, we may form the Young diagram associated to λ by creating an array of squares with λ_i squares in the i^{th} row. See Figure 1 for an example. We denote by $\mathbb{Y}$ the Young graph, which has vertices the Young tableau for any partition

$$\begin{array}{|c|c|}\hline 0 & 1 \\ \hline -1 & 0 \\ \hline -2 \\ \cline{1-1}\end{array} \qquad \begin{array}{|c|c|c|}\hline 0 & 1 & 2 \\ \hline -1 & 0 \\ \cline{1-2} -2 \\ \cline{1-1}\end{array} \qquad \begin{array}{|c|}\hline 0 \\ \hline -1 \\ \hline -2 \\ \hline\end{array} \qquad \begin{array}{|c|c|c|c|}\hline 0 & 1 & 2 & 3 \\ \hline\end{array}$$

(a) (b) (c) (d)

Figure 1: Young tableaux with the contents of their boxes corresponding to (a) $\lambda = (2, 2, 1)$, (b) $\lambda = (3, 2, 1)$, (c) $\lambda = (1, 1, 1)$, and (d) $\lambda = (4)$

λ and has an edge between μ, λ if and only if $\mu \subset \lambda$ and λ/μ is exactly one box; in this case we write $\mu \nearrow \lambda$. See Figure 2 for an example. Given a box in a Young diagram, we define the content of the box to be the difference of the x- and y-coordinates. See Figure 1 for an example. We may define $\mathrm{Tab}(\lambda)$ to be the set of paths in $\mathbb{Y}$ ending in λ and $\mathrm{Tab}(n)$ to be the union of all $\mathrm{Tab}(\lambda)$ such that $|\lambda| = n$. A convenient way to bookkeep this information is to put the number i into $\lambda_i/\lambda_0 i - 1$ if we have a path $\emptyset \nearrow \lambda_1 \nearrow \cdots \nearrow \lambda$. We have the following definition:



(a) (b) (c) (d)

Figure 2: (a) is the partition $\mu = (2,2,1)$ and (b), (c), (d) are the partitions that μ has an edge to, with the extra box marked by $\bullet$.

Definition 23. For $T \in \mathrm{Tab}(n)$ we define the content vector

$$c(T) = (c(\lambda_1/\lambda_0), c(\lambda_2/\lambda_1), \ldots, c(\lambda_i/\lambda_0 i - 1), \ldots, c(\lambda_n/\lambda_{n-1})) \in \mathbb{Z}^n$$

This defines a map $c : \mathrm{Tab}(n) \to \mathbb{Z}^n$ and we define the image $\mathrm{Im}\, c = \mathrm{Cont}(n) \subset \mathbb{Z}^n$. We say that two paths are equivalent, $T \approx T'$ if and only if they end on the same vertex in $\mathbb{Y}$, i.e., are the same shape.

It is immediate that the content map $c : \mathrm{Tab}(n) \to \mathbb{Z}^n$ is injective; indeed, counting the number of appearances for each value in the content vector yields the number of boxes on each diagonal, which uniquely determines the tableau. Our end goal is to show that $\mathrm{Cont}(n) = \mathrm{Spec}(n)$, but we have a number of results to establish before this. We first give a criterion for when a vector is a content vector.

Proposition 24. *For $\alpha \in \mathbb{C}^n$, we have $\alpha \in \mathrm{Cont}(n)$ if and only if the following conditions hold:*

1. $a_1 = 0$.
2. *For all $q > 1$, we have $\{a_q + 1, a_q - 1\} \cap \{a_1, \ldots, a_{q-1}\} \neq \emptyset$.*
3. *For all $q > 1$, if there is some $p < q$ such that $a_p = a_q = a$, then $\{a-1, a+1\} \subset \{a_{p+1}, \ldots, a_{q-1}\}$*

Proof. To show necessity, note that the first box is always in position $(1,1)$ so has content 0. For the second condition, note that we can add boxes only at the ends of rows and at the bottoms of columns. The third condition follows from the same reasoning: if $a_p = a_q$ then a_p, a_q lie on the same diagonal and there must be both a box above and a box to the left of λ_q/λ_{q-1} which must have been placed after the p^{th} box. To show sufficiency, merely note that we can inductively construct a path $T \in \mathrm{Tab}(n)$ by placing boxes corresponding to these contents in the unique place allowed. ■

An example path is shown in Figure 3 We need one more result, regarding the equivalence relation defined

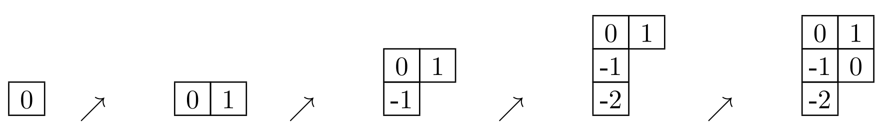


Figure 3: A sample path corresponding to the content vector $(0, 1, -1, -2, 0)$ ending in a partition of shape $(2, 2, 1)$

above, before we can finally unify the theory.

Proposition 25. *Let $\alpha, \beta \in \mathrm{Cont}(n)$. Then $\alpha \approx \beta$ if and only if there is an admissible permutation $\sigma \in S^n$ such that the $\sigma \cdot \alpha = (a_{\sigma \cdot i})_{1 \leq i \leq n} = \beta$.*

Proof. We show that if α is of shape λ, then by under some admissible permutation σ, $\sigma \cdot \alpha$ is equivalent to $T_\lambda \in \mathrm{Tab}(\lambda)$, the path that gives rise to the standard Young tableau of shape λ, where box i in row j is given the number $\lambda_1 + \cdots + \lambda_{j-1} + i$, which has content vector

$$c(T_\lambda) = (0, 1, 2, \ldots, \lambda_1 - 1, -1, 0, \ldots, \lambda_2 - 2, -2, \ldots, \lambda_n - n)$$

We proceed by induction. Clearly the result holds for $n = 1$. Suppose it holds for all $m < n$. Let i be the last box of the last row of T_α, the path such that $c(T_\alpha) = \alpha$. Then we cannot have $a_{i+1} = a_i \pm 1$ because the boxes above and to the left of box i must already have been filled in and, because we are in the last box of the last row, we cannot have box $i+1$ be to the right or below box i. Thus s_i is admissible and $\alpha \approx s_i \cdot \alpha$. But now $i+1$ is admissible with respect to $s_i \cdot \alpha$ for the same reason and $\alpha \approx (s_{i+1}s_i) \cdot \alpha$. Continuing, we have $\alpha \approx (s_{n-1} \dots s_i)\alpha = \alpha'$. We then have that α' has box n as the last box of the last row. Let box i' be the farthest box to the right of the farthest down row such that α' and $c(T_\lambda)$ do not agree. Then we may apply our inductive hypothesis to the content vector $(a_1, \dots a_{i'})$ and we are done. ∎

Thus we have that admissible permutations preserve $\approx$ on $\mathrm{Cont}(n)$. Just as we have an equivalence relation on $\mathrm{Cont}(n)$, so, too, do we have one on $\mathrm{Spec}(n)$:

Definition 26. Let $\alpha, \alpha' \in \mathrm{Spec}(n)$. Then we have $\alpha \sim \alpha'$ if and only if there is some $V^\lambda \in \hat{S}_n$ such that $\alpha, \alpha' \in V^\alpha$, i.e., they belong to the same simple representation of S_n.

Note that we have shown from point (3) in Corollary 20 that if $\sigma \in S_n$ is admissible, then we have $\alpha \sim v_{\sigma\cdot\alpha}$. We will show that this relation agrees with our relation on $\mathrm{Cont}(n)$, in fact,

Proposition 27. *For all n, $\mathrm{Spec}(n) \subset \mathrm{Cont}(n)$ and, moreover, if $\alpha, \beta \in \mathrm{Spec}(n)$ such that $\alpha \approx \beta$ then $\alpha \sim \beta$.*

In fact, we will show equality above of both of these relations. Before we do this, however, we need two lemmata.

Lemma 28. *If $\alpha = (a_1, \dots, a_n) \in \mathbb{C}^n$ such that $a_i = a_{i+2} = a_{i+1} - 1$ for some $1 \le i \le n-2$ then $\alpha \notin \mathrm{Spec}(n)$.*

Proof. We know from (2) in Corollary 20 that $s_i v_\alpha = v_\alpha$ and $s_{i+1}v_\alpha = -v_\alpha$ and so we have $(s_i s_{i+1} s_i) \cdot v_\alpha = v_\alpha$ while $(s_{i+1}s_i s_{i+1}) \cdot v_\alpha = -v_\alpha$. But it is easy to see that

$$s_i s_{i+1} s_i = (i\ \ i+1)(i+1\ \ i+2)(i\ \ i+1) = (i\ \ i+2) = (i+1\ \ i+2)(i\ \ i+1)(i+1\ \ i+2) = s_{i+1}s_i s_{i+1}$$

. Thus we have a contradiction. ∎

Lemma 29. *If $\alpha = (a_1, \dots, a_n) \in \mathrm{Spec}(n)$ then $a_1 = 0$ and $\alpha' = (a_1 \dots, a_{n-1}) \in \mathrm{Spec}(n-1)$.*

Proof. First, $X_1 = 0$ so $a_1 = 0$. Now, by definition of X_i, a_i, v_α, we have $X_i \cdot v_\alpha = a_i \cdot v_\alpha$ for all $1 \le i \le n-1$. Thus, $\alpha' \in \mathrm{Spec}(n-1)$. ∎

Example 30. Consider $\mathrm{Spec}(2)$ and $\mathrm{Cont}(2)$. It is clear that $\mathrm{Cont}(2) = \{(0,1),(0,-1)\}$ because there are only two partitions of 2. By definition, $X_2 = (1\ \ 2)$ and so if $v \in V_+$ the trivial representation then $X_2 \cdot v = v$ and otherwise $X_2 \cdot v = -v$. Thus we have $\mathrm{Spec}(2) = \{(0,1),(0,-1)\} = \mathrm{Cont}(n)$.

We are now ready to prove the key proposition.

Proof of Proposition 27. By Proposition 24, it suffices to show those three conditions. By Lemma 29, we have $a_1 = 0$ so the first condition holds. We prove the other two by induction. We have already established the result for $n = 2$ in the example above, and the case $n = 1$ is trivial. Thus we may assume that the result holds for all $m < n > 2$ and in particular $\mathrm{Spec}(n-1) \subset \mathrm{Cont}(n-1)$. Let $\alpha = (a_1, \dots, a_n) \in \mathrm{Spec}(n)$. Then, by Lemma 29, $\alpha' = (a_1, \dots, a_{n-1}) \in \mathrm{Spec}(n-1)$. Thus in conditions (2) and (3) of Proposition 24, we may assume that $q = n$. Suppose that $\{a_n - 1, a_n + 1\} \cap \{a_1, \dots, a_{n-1}\} = \emptyset$. Then s_{n-1} is admissible so $(a_1, \dots, a_n, a_{n-1}) \in \mathrm{Spec}(n)$ and so, again by Lemma 29, we have $(a_1, \dots, a_{n-2}, a_n) \in \mathrm{Spec}(n-1) \subset \mathrm{Cont}(n-1)$. But then we have $\{a_n - 1, a_n + 1\} \cap \{a_1, \dots a_{n-2}\} = \emptyset$ which contradicts Proposition 24. Thus criterion (2) holds.

Now that there exists some $p < n$ such that $a_p = a_n = a$ and $a - 1 \notin \{a_{p+1}, \dots, a_{n-1}\}$. Without loss of generality, we may choose p maximal subject to these conditions, i.e., $a \notin \{a_{p+1}, \dots, a_{n-1}\}$. Now, because p is maximal and we have $(a_1, \dots, a_{n-1}) \in \mathrm{Cont}(n-1)$, we have that $a+1$ can appear at most once in $\{a_{p+1}, \dots a_{n-1}\}$. Suppose $a+1$ does not appear in this set. Then we may continually apply admissible transpositions to α to get $\alpha \sim \widetilde{\alpha} = (a_1, \dots, a_{p-1}, \dots a_p, a_n, \dots) \in \mathrm{Spec}(n)$. But $a_p = a_n$ and this contradicts (1) of Corollary 20. Now suppose that $a+1 \in \{a_{p+1}, \dots, a_{n-1}\}$. Then, by a sequence of admissible transopositions, we have $\alpha \sim \hat{\alpha} = (a_1, \dots, a_{p-1}, \dots, a_p, a+1, a_n, \dots) \in \operatorname{Spec}(n)$, but this contradicts Lemma 28. The argument applies mutatis mutandis for $a-1$. Thus, $\operatorname{Spec}(n) \subset \operatorname{Cont}(n)$ for all n.

For the last statement, we have by Proposition 25 that $\alpha \approx \beta$ if and only if there is an admissible permutation σ such that $\sigma \cdot \alpha = \beta$. But by (3) of Corollary 20, this implies that $\alpha \sim \beta$. ■

Remark 31. Note that the above proof immediately yields the fact that $\operatorname{Spec}(n) \subset \mathbb{Z}^n$, a statement which is not at all easy. Despite the fact that [OV] cites this fact without proof, we were unable to replicate this result independent of the fact that every irreducible representation of S_n is realizable over $\mathbb{Q}$, a fact that we derive from the classification below.

We are now ready to prove the final result and give an application.

Theorem 32. *We have* $\operatorname{Spec}(n) = \operatorname{Cont}(n)$*, and on this set,* $\approx$ *and* $\sim$ *agree. The branching graph of* $\{S_n\}$ *is given by* $\mathbb{Y}$.

Proof. Note that the last statement follows immediately from the first two and the preceding results. Let p_n be the number of integer partitions of n. Then we have that $|\operatorname{Cont}(n)/\approx| = p_n$ because there is a bijection between shapes of Young diagrams of total weight n and partitions of n. We also know that $|\operatorname{Spec}(n)/\sim| = p_n$ because the left hand side is the number of simple representations of S_n, which is given by the number of conjugacy classes of the same: p_n. Thus we have $|\operatorname{Cont}(n)/\approx| = |\operatorname{Spec}(n)/\sim|$. From Proposition 27, we have that $|\operatorname{Spec}(n)/\sim| \le |\operatorname{Spec}(n)/\approx|$ but also since $\operatorname{Spec}(n) \subset \operatorname{Cont}(n)$, we have

$$|\operatorname{Spec}(n)/\sim| \le |\operatorname{Spec}(n)/\approx| \le |\operatorname{Cont}(n)/\approx| = |\operatorname{Spec}(n)/\sim|$$

and so each of the inequalities must be an equality, which then implies that $\operatorname{Spec}(n) = \operatorname{Cont}(n)$ and $\approx$ and $\sim$ agree, as desired. ■

We now give an example of S_3. Now, in reality, we already know what the representations are, but we shall proceed as if we were to be ignorant of this. First, note that we must have three simple representations, corresponding to each of the three shapes in Figure 4 with shapes (3), $(1,1,1)$, and $(2,1)$. In the first

$$\begin{ytableau} 0 & 1 & 2 \end{ytableau} \qquad \begin{ytableau} 0 \\ -1 \\ -2 \end{ytableau} \qquad \begin{ytableau} 0 & 1 \\ -1 \end{ytableau}$$

Figure 4: The possible simple representations of S_3 with contents filled in.

case, there is only one way of filling in boxes and so the conjugacy class under $\approx$ of a content vector in this representation is the singleton given by $(0,1,2)$. Thus by Theorem 32 the corresponding representation $V_{\ydiagram{3}}$ is one dimensional. We get more than this, however, because we now have the action of $X_2 = (1\ 2)$ and $X_3 = (1\ 3) + (2\ 3)$ on this element. We see that the actions of each of these permutations is trivial, and, because these permutations generate S_3, we see that $V_{\ydiagram{3}}$ is trivial. Similarly, we have that there is a unique way of filling in the boxes for the shape $\lambda = (1,1,1)$ with content vector $(0,-1,-2)$. And we see in a similar way that each transposition acts as negation, giving $V_{\ydiagram{1,1,1}}$ as the alternating representation. For the last representation, we have two content vectors $v = (0,1,-1)$ and $w = (0,-1,1)$. We may use the YJM elements to see exactly what the transpositions do to our vectors. Note that $(1\ 2) \cdot v = v$ while $((1\ 3) + (2\ 3)) \cdot v = -v$ and the opposite for w. It might be instructive to match this to what we already know. In particular, we must have $V_{\ydiagram{2,1}}$ is the representation given by $W = \{a_1e_1 + a_2e_2 + a_3e_3 \in \mathbb{C}^2 | a_1 + a_2 + a_3 = 0\}$ which has as a basis $v = -e_1 - e_2 + 2e_3$ and $w = e_1 - e_2$, which is easily checked. Even without this knowledge, by Corollary 20, we know exactly what the action of each Coxeter generator s_i is on the v_α, so, in principle, we know the action of every permutation $\sigma \in S_3$ on each of the basis vectors.

In the introduction, we promised a better way to think about induction and restriction. This is summarized in the following:

Corollary 33. *Let $k \leq n$, μ a partition of k and λ a partition of n. Define $m_{\mu\lambda}$ as the number of paths in $\mathbb{Y}$ from μ to λ. We have the following:*

$$\operatorname{Res}^{S_n}_{S_k} V^\lambda = \bigoplus_{\mu \subset \lambda} (V^\mu)^{\oplus m_{\mu\lambda}} \qquad \operatorname{Ind}^{S_n}_{S_k} V^\mu = \bigoplus_{\mu \subset \lambda} (V^\lambda)^{\oplus m_{\mu\lambda}}$$

Proof. The first is immediate from the definition of a branching graph and the fact that S_n has simple branching, coupled with the fact that the restriction functor is transitive. The second follows from Frobenius reciprocity. ∎

We conclude by showing that S_n satisfies a very special rationality property.

Proposition 34. *All simple representations of S_n are realizable over $\mathbb{Q}$.*

Proof. For $\sigma \in S_n$ let $\ell(\sigma)$ be the number of $i < j$ such that $\sigma \cdot i > \sigma \cdot j$ and for $T \in \mathrm{Tab}(\lambda)$, define $\ell(T) = \ell(\sigma)$, where $\sigma \cdot T_\lambda = T$, where T_λ was defined in the proof of Proposition 25. If π_T denotes orthogonal projection to V_T as in Proposition 6. Then we have $V_T = \pi_T \sigma \cdot V_{T^\lambda}$ and σ is admissible so we have

$$\sigma \cdot v_{T^\lambda} = v_T + \sum_{\substack{T' \in \mathrm{Tab}(\lambda) \\ \ell(T') < \ell(T)}} \gamma_{T'} v_{T'}$$

where each of these $\gamma_{T'} \in \mathbb{Q}$ by the fact that $\mathrm{Spec}(n) \subset \mathbb{Z}^n$ and (3) of Corollary 20 ∎

Note that we have now accomplished all of our goals that we set out to. We now have a complete description of the simple representations of S_n over $\mathbb{C}$, with a very explicit basis for each one. Because the elements of the basis are indexed by Young tableaux of shape λ, we in particular have a partition λ canonically associated to each irreducible representation, hence for S_n the bijection between conjugacy classes and irreducible representations is canonical. Moreover, we have a natural way to induce and restrict representations, that follows immediately from the construction (see Corollary 33). Now, if we wished, we could do the reverse of the normal construction and apply Schur-Weyl duality to construct Schur functors, the simple representations of $\mathrm{GL}(V)$.

References

[FS] Filippo Tolli Fabio Scarabotti, Tullio Ceccherini-Silberstein. *Representation Theory of the Symmetric Groups.*

[OV] Andrei Okounkov and Anatoly Vershik. A new appraoch to the representation theory of symmetric groups. (4):581–605.

[Pie] Richard S Pierce. *Associative Algebras.*