Representation theory of the symmetric groups

Representation theory of the symmetric groups

William Woods
Selwyn College

I declare that this essay is work done as part of the Part III Examination. I have read and understood the Statement on Plagiarism for Part III and Graduate Courses issued by the Faculty of Mathematics, and have abided by it. This essay is the result of my own work, and except where explicitly stated otherwise, only includes material undertaken since the publication of the list of essay titles, and includes nothing which was performed in collaboration. No part of this essay has been submitted, or is concurrently being submitted, for any degree, diploma or similar qualification at any university or similar institution.

Signed ....

Contents

1 Introduction **1**

2 Preliminaries **1**
2.1 Notation 1
2.2 Representation theory 2
2.3 Character theory 2

3 Representations of S_n **3**
3.1 Young tableaux and tabloids 4
3.2 Specht modules 4
3.2.1 A basis for the Specht module over FS_n 6
3.2.2 A basis for the Specht module over F 7
3.3 Some combinatorial results 11
3.4 p-regularity 13
3.5 Irreducible representations of S_n 14

4 Characters of S_n **18**
4.1 Notation 18
4.2 Calculation of the ordinary character table 19
4.3 The Murnaghan-Nakayama rule 24
4.3.1 Hooks and skew-hooks 24
4.3.2 Statement of the rule 25

5 Conclusion **29**

1 Introduction

It was originally the study of permutations of finite sets and their interaction that gave rise to the fundamental ideas of group theory. The finite symmetric groups S_n are thus a classical object of study, and yet some of their most basic properties still remain elusive. A relatively modern and fruitful approach to the study of the symmetric groups has been to attempt to understand their representations – that is, to try to understand them by analysing how they act on vector spaces rather than by studying them directly. Yet despite the apparent complexity of these groups, their representation theory is particularly beautiful in its simplicity, requiring very little representation-theoretic machinery to achieve some significant results. The flavour of most of the arguments required is combinatorial, and does not require much background knowledge.

This essay will present an exposition of the very basics of the theory. We begin by introducing the necessary building blocks, which are often very simple constructs with a lot of symmetry, and from there go on to build up the irreducible representations of S_n. We proceed in such a way that does not heavily depend on the underlying field, and so in particular our approach is essentially characteristic-free. However, a very powerful tool in representation theory, which it would be unfair to neglect, is character theory; so, in the latter half of the essay, we switch to studying the characters of S_n, and so implicitly restrict our attention to representations over a field of characteristic 0.

In many cases, the detail of the proofs obscures the simplicity of the arguments and results, so examples of the theory in action have been given throughout where appropriate.

The bulk of the material in sections 3 and 4 has been adapted from James [1].

2 Preliminaries

2.1 Notation

Throughout this essay, functions will be written on the left.

2.2 Representation theory

A *representation* of a group G is a pair (ρ, V), where V is a vector space over a field F, and $\rho : G \to \mathrm{GL}(V)$ is a group homomorphism (i.e. a linear action of G on V). This can also be characterised slightly differently. Define the *group algebra* FG of G over F to be the (unital but not necessarily commutative) ring whose elements are formal finite sums of formal products $a \cdot g$ for $a \in F$, $g \in G$, with addition defined in the obvious way and multiplication inherited from G: then a representation of G over F is equivalently an FG-module.

We will be interested in the *irreducible* representations of G up to (FG-module) isomorphism – that is, the representations (ρ, V) such that $(\rho|_U, U)$ is not a representation for any non-trivial subspace U of V. In the language of FG-modules, a representation V is *irreducible* if it has no non-trivial sub-FG-module U.

We will quote a non-trivial result that will be important in finding the modular representations of S_n:

Lemma 2.2.1. The number of modular irreducible representations of a group G over a field of characteristic p is equal to the number of p-regular conjugacy classes of G.

Proof. See Robinson [2, theorem 12.39]. □

Finally, the following result will be useful in calculating the character table of S_n:

Lemma 2.2.2. If W is an irreducible $\mathbb{C}G$-module, then the number of composition factors of the $\mathbb{C}G$-module V that are isomorphic to W is $\dim \mathrm{Hom}_{\mathbb{C}G}(V, W)$.

Proof. James and Liebeck [3, corollary 11.6]. □

2.3 Character theory

Given an ordinary representation (ρ, V) of a group G, we can assign to each element $g \in G$ the *character* $\chi_\rho(g) \in \mathbb{C}$, defined to be the trace of $\rho(g)$ (which is well-defined, as $\rho(g) : V \to V$ is a linear map). Indeed, as trace is invariant under conjugation, characters are *class functions*, i.e. functions that are constant on conjugacy classes.

Define an inner product $\langle \, , \, \rangle$ on the space of all functions $G \to \mathbb{C}$ by

$$\langle \alpha, \beta \rangle = \frac{1}{|G|} \sum_{g \in G} \alpha(g) \overline{\beta(g)},$$

where the bar denotes complex conjugation. Given two irreducible characters χ_1, χ_2, it is easy to show that

$$\langle \chi_1, \chi_2 \rangle = \begin{cases} 1 & \text{if } \chi_1 = \chi_2 \\ 0 & \text{if } \chi_1 \neq \chi_2; \end{cases}$$

as a sort of converse, given a character χ, it can be shown that $\langle \chi, \chi \rangle = 1$ if and only if χ is irreducible [3, theorems 14.12 and 14.20].

Unsurprisingly, isomorphic representations afford the same character. However, more remarkably, the converse holds; that is, any two ordinary representations of G affording the same character must be isomorphic [3, theorem 14.21]. Combined with the lemma quoted in the previous section, these facts imply that the irreducible characters of a group G form an orthonormal basis for the $\mathbb{C}$-vector space of class functions on G.

If H is a subgroup of G, restriction of the G-character ψ from G to H always produces an H-character; this will be written $\psi{\downarrow_H}$. We can also define induction of the H-character θ from H to G: this will be written $\theta{\uparrow^G}$, and is defined as

$$\theta{\uparrow^G}(g) = \frac{1}{|H|} \sum_{\substack{x \in G, \\ x^{-1}gx \in H}} \theta(x^{-1}gx).$$

This produces a G-character. A reciprocity theorem of Frobenius [3, theorem 21.16] says that $\langle \theta{\uparrow^G}, \psi \rangle = \langle \theta, \psi{\downarrow_H} \rangle$, where the first inner product is taken over G and the second over H.

3 Representations of S_n

Definition. Let n be a positive integer. A *partition* of n is a finite sequence $\lambda = \langle \lambda_1, \ldots, \lambda_r \rangle$ of positive integers with $\lambda_1 \geq \cdots \geq \lambda_r$ whose sum is n.

Example. Some partitions of 7 are $\langle 7 \rangle, \langle 2, 2, 1, 1, 1 \rangle, \langle 5, 2 \rangle$.

Remark. Partitions of n are in natural bijective correspondence with cycle types in S_n, and hence also with conjugacy classes in S_n.

3.1 Young tableaux and tabloids

Definition. Let $\lambda = \langle \lambda_1, \ldots, \lambda_r \rangle$ be a partition of n. The *(Young) diagram* corresponding to λ is the $(r \times \lambda_1)$ matrix with (i,j)-entry "$\times$" if $j \leq \lambda_i$, and blank otherwise; that is, the jth row contains a $\times$ in each of the first λ_j entries. Brackets around the matrix are often omitted. It will cause no confusion to identify partitions with their Young diagrams, so we will do so frequently. Also, we will write $(i,j) \in \lambda$ if the (i,j)-entry of the Young diagram of λ is "$\times$", and $(i,j) \notin \lambda$ otherwise.

Example. $\langle 5,2 \rangle = \begin{matrix} \times & \times & \times & \times & \times \\ \times & \times & & & \end{matrix}$

Definition. Let λ be a partition of n. A *λ-tableau* is the result of replacing each $\times$ in the diagram of λ with an integer from $\{1, 2, \ldots, n\}$ such that each integer appears exactly once in the resulting matrix. (By convention, tableaux will be drawn so that their entries are separated by lines of a grid.)

Example. $t_1 = \begin{array}{|c|c|c|c|c|} \hline 1 & 6 & 5 & 3 & 2 \\ \hline 7 & 4 \\ \cline{1-2} \end{array}$ and $t_2 = \begin{array}{|c|c|c|c|c|} \hline 3 & 2 & 1 & 5 & 4 \\ \hline 7 & 6 \\ \cline{1-2} \end{array}$ are $\langle 5,2 \rangle$-tableaux.

Remark. S_n acts on λ-tableaux in the obvious way (by permuting their entries).

Definition. Given a λ-tableau t, define its *row stabiliser* to be the subgroup R_t of S_n that fixes each row setwise, and its *column stabiliser* to be the subgroup C_t of S_n that fixes each column setwise.

Definition. Two λ-tableaux t_1, t_2 are *equivalent* if one can be obtained from the other by permuting elements in each row but keeping each row fixed setwise, i.e. if $t_1 = \sigma(t_2)$ for some $\sigma \in R_{t_2}$. (Of course, this is an equivalence relation: $R_{t_1} = R_{t_2}$.) An equivalence class of a λ-tableau t is called a *λ-tabloid*, denoted $[t]$. (By convention, tabloids will be drawn so that their *rows* are separated by lines.)

Example. $\begin{array}{|cccc|} \hline 1 & 3 & 4 & 6 \\ \hline 2 & 5 \\ \cline{1-2} \end{array} = \begin{array}{|cccc|} \hline 6 & 4 & 1 & 3 \\ \hline 5 & 2 \\ \cline{1-2} \end{array} = \ldots$

3.2 Specht modules

Definition. Let $\lambda = \langle \lambda_1, \ldots, \lambda_r \rangle$ be a partition of n. Then define the *Young subgroup* $S_\lambda \leq S_n$ to be any subgroup isomorphic to $S_{\lambda_1} \times \cdots \times S_{\lambda_r}$, i.e. the row stabiliser of some given λ-tableau t. (Usually S_λ is taken to be the row stabiliser of the following tableau:

$$\begin{array}{|c|c|c|c|l}
\hline
1 & 2 & 3 & 4 & \cdots \\
\hline
\lambda_1{+}1 & \lambda_1{+}2 & \ddots \\
\cline{1-2}
\lambda_1{+}\lambda_2{+}1 & \ddots \\
\cline{1-1}
\vdots
\end{array}$$

but such a specific choice is entirely unnecessary for our purposes, as the row stabilisers of *any* two λ-tableaux are S_n-conjugate.)

Define also the formal F-vector space M^λ with all possible λ-tabloids as basis elements.

Example. Let $\lambda = \langle 4, 2, 1 \rangle$. A basis element of M^λ can be specified by choosing the four elements in the first row from $\{1, \ldots, 7\}$ and then choosing the two elements in the second row from the three remaining numbers - that is, M^λ is a vector space of dimension $\binom{7}{4} \cdot \binom{3}{2} = 105$.

Remark. M^λ is an FS_n-module by extending the action of S_n on tabloids linearly. Of course, M^λ is always an FS_n-module of dimension 1, as any F-basis element can be transformed into any other by applying an element of S_n.

Definition. Given a λ-tableau t, define the *polytabloid*

$$e_t = \sum_{\sigma \in C_t} \operatorname{sgn}(\sigma) \cdot \sigma([t]) \in M^\lambda.$$

That is, $e_t = \kappa_t([t])$, where $\kappa_t = \sum_{\sigma \in C_t} \operatorname{sgn}(\sigma) \cdot \sigma \in FS_n$ is the *signed column sum*.

Definition. The *Specht module* S^λ is the sub-FS_n-module of M^λ spanned by polytabloids.

Example. Take $\lambda = \langle 4, 2, 1 \rangle$ again. Take $t = \begin{array}{|c|c|c|c|} \hline 1 & 2 & 3 & 4 \\ \hline 5 & 6 \\ \cline{1-2} 7 \\ \cline{1-1} \end{array}$.

It is clear that $C_t = \operatorname{Sym}(\{1, 5, 7\}) \times \operatorname{Sym}(\{2, 6\})$, which is a group of order 12, so the orbit of t contains 12 tableaux $t = t_1, t_2, \ldots, t_{12}$. Writing $t_i = \sigma_i(t)$, and ordering the σ_i so that $\sigma_1, \ldots, \sigma_6$ are the even permutations and $\sigma_7, \ldots, \sigma_{12}$ are the odd permutations, it is clear that

$\kappa_t = (\sigma_1 + \cdots + \sigma_6) - (\sigma_7 + \cdots + \sigma_{12})$.

Hence the polytabloid e_t is a linear combination of the 12 tableaux t_i.

It is natural to ask about the properties of these Specht modules. We proceed by finding a standard basis for the Specht module as an F-vector space.

3.2.1 A basis for the Specht module over FS_n

We begin by examining what the Specht module looks like as an FS_n-module: this is considerably simpler than as an F-vector space.

Theorem 3.2.1. S^λ is generated by any one polytabloid, i.e. is an FS_n-module of dimension 1.

Proof. Let t be a λ-tableau. Any other λ-tableau can be written as $\rho(t)$ for some $\rho \in S_n$. Clearly $C_t = \rho C_{\rho(t)} \rho^{-1}$, and so

$$\begin{aligned}
\kappa_{\rho(t)} &= \sum_{\sigma \in C_{\rho(t)}} \operatorname{sgn}(\sigma) \cdot \sigma \\
&= \sum_{\rho^{-1}\sigma\rho \in C_t} \operatorname{sgn}(\sigma) \cdot \sigma \\
&= \sum_{\rho^{-1}\sigma\rho \in C_t} \operatorname{sgn}(\rho^{-1}\sigma\rho) \cdot \sigma \quad \text{(sgn is conjugation-invariant)} \\
&= \sum_{\rho^{-1}\sigma\rho \in C_t} \operatorname{sgn}(\rho^{-1}\sigma\rho) \cdot \rho(\rho^{-1}\sigma\rho)\rho^{-1} \\
&= \sum_{\sigma' \in C_t} \operatorname{sgn}(\sigma') \cdot \rho\sigma'\rho^{-1} \\
&= \rho \left[\sum_{\sigma' \in C_t} \operatorname{sgn}(\sigma') \cdot \sigma' \right] \rho^{-1} = \rho\kappa_t\rho^{-1}.
\end{aligned}$$

This gives that

$$\begin{aligned}
e_{\rho(t)} &= \kappa_{\rho(t)}([\rho(t)]) \\
&= \rho\kappa_t\rho^{-1} \cdot \rho([t]) \\
&= \rho e_t.
\end{aligned}$$

□

3.2.2 A basis for the Specht module over F

Definition. A tableau t is *standard* if the entries of t increase along each row and down each column. A tabloid (viewed as an equivalence class of tableaux) is *standard* if it contains a standard tableau. A polytabloid e_t is *standard* if t is standard.

Our aim is to show that the standard polytabloids form a basis for S^λ over F. We begin by defining a useful total order on λ-tabloids, which will help us to show linear independence of the standard polytabloids.

Definition. Suppose λ is a partition of n. The sequences of length n containing each of the numbers $1, \ldots, n$ can be endowed with the (total) lexicographic order. This induces a total ordering on the λ-tableaux in an obvious way: a λ-tableau can be considered as a sequence of length n by reading each row in turn from left to right, starting with the top row and working down towards the bottom. Hence, for example, the $\langle 5,4,1\rangle$-tableaux are totally ordered as follows:

$$\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & \\ 10 & & & & \end{array} < \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 10 & \\ 9 & & & & \end{array} < \cdots < \begin{array}{ccccc} 10 & 9 & 8 & 7 & 6 \\ 5 & 4 & 3 & 2 & \\ 1 & & & & \end{array}.$$

This then induces a total ordering on the λ-tabloids. Take two λ-tabloids $[t_1], [t_2]$, assuming without loss of generality that t_1, t_2 are the $<$-minimal tableaux in $[t_1], [t_2]$ respectively. (This assumption is critical!) Then $[t_1] \leq [t_2]$ if and only if $t_1 \leq t_2$.

Lemma 3.2.2. If t is a standard tableau, then $[t]$ is the $<$-minimal tabloid involved in e_t.

Proof. Any other tabloid involved in e_t corresponds to a tableau s which can be obtained from t by permuting entries in each column (and then, without loss of generality, reordering the elements in each row of s so that they are increasing along the row). Consider the first row in s (reading from left to right, top to bottom, as before) which differs from the corresponding row in t, and reorder the elements in this row of s so that they are in increasing order. As t is standard, a permutation of the columns cannot decrease any entry in this row; hence we must have that each entry in this row of t is less than or equal to the corresponding entry of s. So $[t] \leq [s]$. □

Lemma 3.2.3 (*linear independence*)**.** Let $e_{t_1}, \dots, e_{t_k}$ be standard polytabloids, with all t_i distinct. Then the e_{t_i} are linearly independent.

Proof. Given that the t_i are all distinct standard tableaux, we cannot obtain t_j from t_i by permuting the elements of each row (for $i \neq j$), so that all the $[t_i]$ are distinct standard tabloids. The proof is now trivial: suppose, without loss of generality, that e_{t_1} is the $<$-minimal of these k polytabloids. Then, given any linear combination $a_1 e_{t_1} + \cdots + a_k e_{t_k} \in S^\lambda$, the coefficient of $[t_1]$ is a_1, so we must have $a_1 = 0$. Now induction on k gives the required result. □

It now remains to prove that the standard polytabloids span S^λ.

Remark. The tabloid $[t]$ is the row equivalence class of t. We can also define $\{t\}$, the *column* equivalence class of a tableau t. Note that tableaux, and hence also column equivalence classes, have a total ordering $\preceq$ in a similar way to $\leq$ on the tabloids: read each column from top to bottom, starting with the leftmost column and working towards the right. (This is essentially the transpose of the $\leq$ ordering.) It is also worth noting that a column equivalence class $\{t\}$ determines a polytabloid e_t uniquely up to sign: if $\{s\} = \{t\}$, then $s = \sigma(t)$ for some $\sigma \in C_t$, so $e_s = e_{\sigma(t)} = \sigma(e_t) = \pm e_t$. In particular, if $\{s\}$ is $\preceq$-minimal, then $e_s = \pm e_{s_0}$, where s_0 is the $\preceq$-minimal tabloid; s_0 is standard, and so e_s is standard.

Lemma 3.2.4 (*spanning*)**.** Any polytabloid can be written as a linear combination of the standard polytabloids.

Proof. Let λ be a fixed partition of n.

Suppose t is a λ-tableau: we aim to show that e_t can be written as a linear combination of standard polytabloids. (Without loss of generality, we may reorder the elements in each column of t to be in increasing order down the column, since if $\sigma \in C_t$, we know that

$$e_{\sigma(t)} = \sigma(e_t) = \sigma \sum_{\tau \in C_t} \operatorname{sgn}(\tau)\tau([t]) = \pm e_t,$$

and replacing e_t by a multiple of e_t does not affect the result.)

We first make the trivial remark that, if t is standard, we are done, so assume henceforth that t is not standard. As t is not standard, there must be two adjacent columns, say c_k and c_{k+1}, whose entries are $a_1 < a_2 < \cdots < a_r$ and $b_1 < b_2 < \cdots < b_s$ respectively, such that $a_q > b_q$ for some q and c_k is to the left of c_{k+1}, so that the qth row is non-increasing. Hence $b_1 < b_2 < \cdots < b_q < a_q < \cdots < a_r$.

Define $A = \{a_q, a_{q+1}, \dots, a_r\}$, $B = \{b_1, b_2, \dots, b_q\}$. We have two S_n-subgroups $S_A \times S_B \leq S_{A\cup B}$: take a (left) transversal $\mathcal{T}$ of $S_A \times S_B$ in $S_{A\cup B}$ containing id and define the *Garnir element* of $\mathcal{T}$ to be $G^{\mathcal{T}} = \sum_{\sigma\in\mathcal{T}} \operatorname{sgn}(\sigma)\sigma$. Define also

$$\varphi = \sum_{\tau\in S_A\times S_B} \operatorname{sgn}(\tau)\tau, \quad \psi = \sum_{\tau\in S_{A\cup B}} \operatorname{sgn}(\tau)\tau,$$

and note the following:

(i) $G^{\mathcal{T}}\varphi = \left(\sum_{\sigma\in\mathcal{T}} \operatorname{sgn}(\sigma)\sigma\right)\left(\sum_{\tau\in S_A\times S_B} \operatorname{sgn}(\tau)\tau\right)$
$= \sum_{\sigma,\tau} \operatorname{sgn}(\sigma\tau)\cdot(\sigma\tau) = \psi$ (by definition of $\mathcal{T}$)

(ii) $\varphi\kappa_t = \left(\sum_{\tau\in S_A\times S_B} \operatorname{sgn}(\tau)\tau\right)\left(\sum_{\rho\in C_t} \operatorname{sgn}(\rho)\rho\right)$
$= \sum_{\tau,\rho} \operatorname{sgn}(\tau\rho)\cdot(\tau\rho) = |S_A\times S_B|\kappa_t$
(as $S_A \times S_B \subseteq C_t$, so fixing τ and summing over ρ gives one copy of κ_t).

Hence

(iii) $\psi(e_t) \overset{\text{(i)}}{=} G^{\mathcal{T}}\varphi(e_t) = G^{\mathcal{T}}\varphi\kappa_t([t]) \overset{\text{(ii)}}{=} |S_A\times S_B| G^{\mathcal{T}}\kappa_t([t])$
$= |S_A\times S_B| G^{\mathcal{T}}(e_t)$.

But, given $\tau \in C_t$, there are two numbers, say α, β, in $A\cup B$ that are in the same row of $\tau(t)$, so that $(\alpha \;\; \beta)\,[\tau(t)] = [\tau(t)]$. Then every element of $S_{A\cup B}$ is either even or can be written uniquely as the product of some even permutation with $(\alpha \;\; \beta)$. So

$$\psi = \sum_{\substack{\sigma\in S_{A\cup B},\\ \sigma \text{ even}}} \sigma - \sum_{\substack{\sigma\in S_{A\cup B},\\ \sigma \text{ even}}} \sigma\,(\alpha \;\; \beta) = \widetilde{\sigma}(\mathrm{id} - (\alpha \;\; \beta)),$$

where $\widetilde{\sigma}$ is the sum of all the even permutations in $S_{A\cup B}$, and so $\psi([\tau(t)]) = 0$ for each $\tau \in C_t$. This shows that $\psi(e_t) = 0$, and so by (iii), we have that $G^{\mathcal{T}}(e_t) = 0$. We can write this out in full, separating the term corresponding to $\sigma = \mathrm{id}$, and noting (by theorem 3.2.1) that $\sigma(e_t) = e_{\sigma(t)}$:

$$e_t + \sum_{\mathrm{id} \neq \sigma \in \mathcal{T}} \operatorname{sgn}(\sigma) e_{\sigma(t)} = 0.$$

That is, we have proven that e_t is a linear combination of the $e_{\sigma(t)}$ for $\mathrm{id} \neq \sigma \in \mathcal{T}$. However, it still remains to show that we can write e_t as a linear combination of *standard* polytabloids: for this, we use the remark made directly before this lemma.

Take any $\mathrm{id} \neq \sigma \in \mathcal{T}$: then σ is an element of $S_{A \cup B}$, and there exist $a_i \in A, b_j \in B$ such that $\sigma(a_i) = b_j$ (otherwise σ is an element of $S_A \times S_B$. But $b_1 < b_2 < \cdots < b_q < a_q < \cdots < a_r$; that is, σ replaces some elements in the column c_k with smaller elements from c_{k+1}. Hence, with notation as in the remark, we have that $\{\sigma(t)\} \prec \{t\}$. This shows that we can write e_t as a linear combination of some $e_{\sigma(t)}$ with $\{\sigma(t)\} \prec \{t\}$.

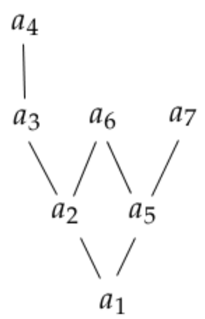
Now we simply need to recall that, when $\{s\}$ is $\preceq$-minimal, e_s is a standard polytabloid; this proves the theorem by induction on the $\preceq$ ordering, as iterating the above procedure will eventually write e_t in terms of standard polytabloids. □

To summarise, we have proven:

Theorem 3.2.5. The standard polytabloids form a basis for S^λ. □

Example. Once again, take $\lambda = \langle 4, 2, 1 \rangle$. Then the dimension of S^λ as an F-vector space is the number of standard polytabloids, which is the number of standard tableaux.

Let $t = \begin{array}{|c|c|c|c|} \hline a_1 & a_2 & a_3 & a_4 \\ \hline a_5 & a_6 \\ \cline{1-2} a_7 \\ \cline{1-1} \end{array}$ be a standard λ-tableau. That is, the a_i are partially ordered as follows:



Choosing a standard λ-tableau is equivalent to choosing an ordering for the a_i. Note that we must always have $a_1 = 1$. We can list the possible orderings of a_2, a_3, a_4, a_5, a_6 systematically according to the position of a_5:

- $a_2 < a_3 < a_4 < \underline{a_5} < a_6$
- $a_2 < a_3 < \underline{a_5} < a_4 < a_6$
- $a_2 < a_3 < \underline{a_5} < a_6 < a_4$
- $a_2 < \underline{a_5} < a_3 < a_4 < a_6$
- $a_2 < \underline{a_5} < a_3 < a_6 < a_4$
- $a_2 < \underline{a_5} < a_6 < a_3 < a_4$
- $\underline{a_5} < a_2 < a_3 < a_4 < a_6$
- $\underline{a_5} < a_2 < a_3 < a_6 < a_4$
- $\underline{a_5} < a_2 < a_6 < a_3 < a_4$

Now we need to choose a position for a_7 for each of these possibilities. But the only constraint on a_7 is that it comes later than a_5. So, for each of these possibilities, a_7 can assume any one of 2, 3, 3, 4, 4, 4, 5, 5 and 5 positions respectively.

This gives a total of 35 standard tableaux, and hence the Specht module $S^{\langle 4,2,1 \rangle}$ has dimension 35 over F.

3.3 Some combinatorial results

The results in this section are not interesting in their own right for our purposes, but are critical for several proofs later, so we derive them all together at this point for convenience.

Lemma 3.3.1. Let λ, μ be partitions of n, and suppose that t_1 is a λ-tableau. If we can find a μ-tableau t_2 such that, for every i, the numbers in the ith row of t_2 belong to different columns of t_1, then $\lambda \geq \mu$.

Proof. Pick one number from each column of t_1, and try to put them in the first row of t_2; that is, we have λ_1 numbers to fit into μ_1 spaces. If they fit, we must have $\lambda_1 \geq \mu_1$. Cross out all the used numbers in t_1. Now pick one unused number from each remaining column of t_1, and try to put them into the second row of t_2; if they fit, we have $\lambda_1 + \lambda_2 \geq \mu_1 + \mu_2$, and so $\lambda_2 \geq \mu_2$. Continuing in this way gives us $\lambda \geq \mu$. $\square$

Lemma 3.3.2. Let λ, μ be partitions of n, and suppose that t_1 is a λ-tableau and t_2 is a μ-tableau such that $\kappa_{t_1}([t_2]) \neq 0$. Then $\lambda \geq \mu$.

Proof. If a, b are two numbers in the same row of t_2, then clearly $[t_2] = [\begin{pmatrix} a & b \end{pmatrix} t_2] = \begin{pmatrix} a & b \end{pmatrix} [t_2]$.

Now suppose that a, b are in the same column of t_1; then $\begin{pmatrix} a & b \end{pmatrix} \in C_{t_1}$. Let H be the subgroup of C_{t_1} generated by $\begin{pmatrix} a & b \end{pmatrix}$; then H is a group of order 2, and each left coset contains precisely one even permutation and one odd permutation. Let $\sigma_1, \ldots, \sigma_k$ be a complete list of the even permutations without repetitions; then $\sigma_1 \begin{pmatrix} a & b \end{pmatrix}, \ldots, \sigma_k \begin{pmatrix} a & b \end{pmatrix}$ is a complete list of the odd permutations without repetitions. Hence

$$\begin{aligned}
\kappa_{t_1} &= \sum_{\sigma \in C_t} \operatorname{sgn}(\sigma) \cdot \sigma \\
&= \underbrace{\sum_{i=1}^{k} \operatorname{sgn}(\sigma_i) \cdot \sigma_i}_{\text{even terms}} + \underbrace{\sum_{i=1}^{k} \operatorname{sgn}(\sigma_i \begin{pmatrix} a & b \end{pmatrix}) \cdot \sigma_i \begin{pmatrix} a & b \end{pmatrix}}_{\text{odd terms}} \\
&= \sum_{i=1}^{k} \sigma_i (\mathrm{id} - \begin{pmatrix} a & b \end{pmatrix}),
\end{aligned}$$

so that

$$\begin{aligned}
\kappa_{t_1}([t_2]) &= \sum_{i=1}^{k} \sigma_i (\mathrm{id} - \begin{pmatrix} a & b \end{pmatrix})([t_2]) \\
&= \sum_{i=1}^{k} \sigma_i ([t_2] - \begin{pmatrix} a & b \end{pmatrix} [t_2]) \\
&= 0,
\end{aligned}$$

which contradicts our initial assumption. Hence all elements in any given row of t_2 lie in different columns of t_1; now invoke lemma 3.3.1. □

Lemma 3.3.3. Suppose that t is a λ-tableau and $u \in M^\lambda$. Then $\kappa_t(u)$ is a multiple of e_t.

Proof. First we will prove the statement in the case where u is a single tabloid. Let $t_1 = t$, and let t_2 be another λ-tableau. If $\kappa_{t_1}([t_2]) = 0$, then clearly it is a multiple of e_t; otherwise, the same argument as the start of the proof of lemma 3.3.2 shows that the numbers in each row of t_2 belong to different columns of t_1. Hence there is some $\sigma \in C_{t_1}$ such that the elements in each row of $\sigma(t_1)$ are also in the corresponding row of t_2 – that is, $\sigma([t_1]) = [t_2]$. So $\kappa_{t_1}([t_2]) = \kappa_{t_1}\sigma([t_1])$. But $\kappa_{t_1}\sigma = \pm\kappa_{t_1}$, as $\sigma \in C_{t_1}$. So $\kappa_{t_1}([t_2])$ is a multiple of $\kappa_{t_1}([t_1]) = e_{t_1}$.

Now, given a general $u \in M^\lambda$, we can simply write u as a formal linear combination of λ-tabloids and apply the above to each tabloid, and the result follows immediately. □

3.4 p-regularity

Throughout this section, F is a field of characteristic p.

Remark. This is the first point in section 3 at which we take into account the characteristic of the field F.

Definition. A partition λ of n is *p-singular* if it contains p rows all of the same length; otherwise it is *p-regular*.

Example. Consider $\lambda = \langle 5, 3, \underbrace{2, 2, 2, 2, 2, 2}_{\text{6 copies of the same number}}, 1 \rangle$:

$$
\begin{array}{lllll}
\times & \times & \times & \times & \times \\
\times & \times & \times & & \\
\times & \times & & & \\
\times & \times & & & \\
\times & \times & & & \\
\times & \times & & & \\
\times & \times & & & \\
\times & \times & & & \\
\times & & & &
\end{array}
$$

As λ has 6 rows of the same length, it is p-singular for all $p \leq 6$ and p-regular for all $p > 6$.

Definition. A conjugacy class in S_n is *p-singular* if any element has order divisible by p; otherwise it is *p-regular*.

Theorem 3.4.1. The p-regular partitions of n are in bijection with the p-regular conjugacy classes of S_n (justifying the overloading of the term "p-regular").

Proof. Consider the formal power series $q(X)$ whose coefficient of X^n is the number of p-regular conjugacy classes of S_n. Equivalently, this is the number of cycle types not including any cycles of length divisible by p, i.e. the number of partitions of n not containing any rows of length divisible by p. It is easily checked that

$$q(X) = \prod_{p \nmid i} (1 + X^i + (X^i)^2 + (X^i)^3 + \dots),$$

where the partition $\langle \underbrace{\alpha_1, \dots, \alpha_1}_{m_1 \text{ times}}, \alpha_2, \dots, \alpha_{r-1}, \underbrace{\alpha_r, \dots, \alpha_r}_{m_r \text{ times}} \rangle$ of n (where $\alpha_1 > \alpha_2 > \dots > \alpha_r$) corresponds to the contribution of $(X^{\alpha_1})^{m_1} \dots (X^{\alpha_r})^{m_r}$ (and all other factors 1) towards the coefficient of X^n, since $p \nmid \alpha_1, \dots, \alpha_r$.

By multiplying both sides by the 'missing' factors corresponding to $p|i$, we get

$$q(X) \cdot \prod_{i=1}^{\infty} (1 + X^{pi} + (X^{pi})^2 + \dots) = \prod_{i=1}^{\infty} (1 + X^i + (X^i)^2 + \dots).$$

But $1 + X^j + (X^j)^2 + \dots = (1 - X^j)^{-1}$, and so

$$q(X) \cdot \prod_{i=1}^{\infty} (1 - X^{pi})^{-1} = \prod_{i=1}^{\infty} (1 - X^i)^{-1}.$$

That is,

$$\begin{aligned} q(X) &= \prod_{i=1}^{\infty} \left(\frac{1 - X^{ip}}{1 - X^i} \right) \\ &= \prod_{i=1}^{\infty} (1 + X^i + X^{2i} + \dots + X^{(p-1)i}). \end{aligned}$$

As before, the partition $\langle \underbrace{\alpha_1, \dots, \alpha_1}_{m_1 \text{ times}}, \alpha_2, \dots, \alpha_{r-1}, \underbrace{\alpha_r, \dots, \alpha_r}_{m_r \text{ times}} \rangle$ of n can be seen to correspond to the contribution of $(X^{m_1 \alpha_1}) \dots (X^{m_r \alpha_r})$ (and all other factors 1) towards the coefficient of X^n. But a partition is p-regular if and only if all $m_i < p$. So, from this new product formula for $q(X)$, we can deduce that the coefficient of X^n is the number of p-regular partitions of n. □

3.5 Irreducible representations of S_n

Definition. Given a partition λ of n, there is a unique form $\langle\ ,\ \rangle$ on M^λ which is defined on tabloids by

$$\langle [t_1],[t_2]\rangle = \begin{cases} 1 & \text{if } [t_1]=[t_2] \\ 0 & \text{if } [t_1]\neq[t_2] \end{cases}$$

and extended linearly. (This form is clearly symmetric, bilinear, S_n-invariant and non-degenerate.) For a submodule V of M^λ, write $V^\perp$ to mean the submodule of M^λ consisting of all elements u with $\langle u,v\rangle = 0$ for all $v \in V$. (This is a slight abuse of notation: $\langle\ ,\ \rangle$ is not an inner product unless we are working over a field of characteristic 0, so $V \cap V^\perp$ may be a non-trivial subspace of M^λ!)

Lemma 3.5.1. For $u, v \in M^\lambda$ and any λ-tableau t, we have $\langle \kappa_t(u), v\rangle = \langle u, \kappa_t(v)\rangle$.

Proof.

$$\begin{aligned} \langle \kappa_t(u), v\rangle &= \langle \sum_{\sigma\in C_t} \operatorname{sgn}(\sigma)\sigma(u), v\rangle \\ &= \sum_{\sigma\in C_t} \operatorname{sgn}(\sigma)\langle\sigma(u), v\rangle \\ &= \sum_{\sigma\in C_t} \operatorname{sgn}(\sigma)\langle u, \sigma^{-1}(v)\rangle \\ &= \sum_{\sigma^{-1}\in C_t} \operatorname{sgn}(\sigma^{-1})\langle u, \sigma^{-1}(v)\rangle \\ &= \langle u, \sum_{\sigma^{-1}\in C_t} \operatorname{sgn}(\sigma^{-1})\sigma^{-1}(v)\rangle \\ &= \langle u, \kappa_t(v)\rangle. \end{aligned}$$

□

We will need one final combinatorial result:

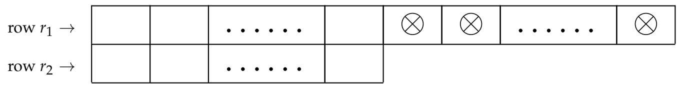
Lemma 3.5.2. Let t be a λ-tableau; let t' be the λ-tableau obtained by reversing the order of the entries in each row of t (the *reverse* of t). Then $\kappa_t(e_{t'}) = he_t$ for some constant h, and $p|h$ if and only if λ is p-singular.

Proof. Lemma 3.3.3 immediately gives that $\kappa_t(e_{t'}) = he_t$ for some h. It is also easy to see that $h = \langle he_t, [t]\rangle = \langle \kappa_t(e_{t'}), [t]\rangle$,

which (by lemma 3.5.1) is equal to $\langle e_{t'}, \kappa_t([t])\rangle$, which is just $\langle e_{t'}, e_t\rangle$. We now attempt to find h by evaluating this inner product, which simply counts the number of tabloids involved in both $e_{t'}$ and e_t.

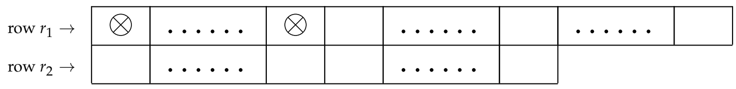
Suppose $\lambda = \langle \underbrace{\alpha_1,\ldots,\alpha_1}_{m_1 \text{ times}}, \alpha_2,\ldots,\alpha_{r-1}, \underbrace{\alpha_r,\ldots,\alpha_r}_{m_r \text{ times}}\rangle$ as above. Take some element $\sigma \in C_t$: clearly $[\sigma(t)]$ is involved in e_t, and we wish to determine whether it is involved in $e_{t'}$, i.e. whether it is $[\tau(t')]$ for some $\tau \in C_{t'}$.

Suppose σ moves an element in one row to a row of a different length; without loss of generality, i is an element of a row r_1 of length α_1 and $\sigma(i)$ lies in a row r_2 of length α_2. Then i cannot be any of the $(\alpha_1 - \alpha_2)$ elements at the end of r_1 in t, i.e. the cells marked with a $\otimes$ here:



.

Now let $\sigma(i) = j$, and consider t'. The disallowed values for i found above are still marked:



.

Now $\tau^{-1}(j)$ must be back in row r_1, which shows that j cannot be any entry in r_2 *below* a cell marked with a $\otimes$ in r_1 (as they would give rise to disallowed values for i). Hence we have reduced the possible allowed values for both i and j by at least 1, and r_1 still contains more allowed values than r_2. Repeating this algorithm eventually shows that no value of i is allowed!

This contradicts our initial assumption that σ could move elements between rows of different lengths. Hence any σ such that $[\sigma(t)]$ is involved in $e_{t'}$ must only permute elements in rows of the same length. It is also easy to see that, if σ only permutes elements in rows of the same length, then $\sigma \in C_t \cap C_{t'}$, so $[\sigma(t)] = [\sigma(t')]$ is certainly involved in $e_{t'}$.

Hence the number of tabloids common to both e_t and $e_{t'}$ is the number of such $\sigma \in C_t$, which is $h = \prod_{i=1}^{r} (m_i!)^{\alpha_i}$, since within each block of rows of length α_i, C_t can permute each column freely, i.e. can act like a copy of S_{m_i} on each column. This explicit formula for h shows that $p|h$ if and only if some $m_i \geq p$, i.e. if λ is p-singular. □

Theorem 3.5.3. If U is a submodule of M^λ, then either $S^\lambda \subseteq U$ or $U \subseteq (S^\lambda)^\perp$.

Proof. If every $u \in U$ and every λ-tableau t has $\kappa_t(u) = 0$, then

$$\langle u, e_t \rangle = \langle u, \kappa_t([t]) \rangle = \langle \kappa_t(u), [t] \rangle = 0$$

by lemma 3.5.1. As S^λ is generated by e_t, this shows that $U \subseteq (S^\lambda)^\perp$. But if we can find some u and t such that $\kappa_t(u) \neq 0$, then by lemma 3.3.3, we know that $\kappa_t(u)$ is a non-zero multiple of e_t, and hence is in S^λ. □

Definition. For each partition λ of n, define $D^\lambda = \dfrac{S^\lambda}{S^\lambda \cap (S^\lambda)^\perp}$.

Theorem 3.5.4. If $D^\lambda \neq 0$, then D^λ is absolutely irreducible.

Proof. Let $e_1, \ldots, e_k$ be a basis of polytabloids for S^λ, and let $\varepsilon_1, \ldots, \varepsilon_k$ be the dual basis for $(S^\lambda)^*$. Define $\theta : S^\lambda \to (S^\lambda)^*$ by $u \mapsto \psi_u$, where $\psi_u(v) = \langle u, v \rangle$. Then $\psi_{e_i}(e_j) = \langle e_i, e_j \rangle$, so $\psi_{e_i} = \sum_j \langle e_i, e_j \rangle \varepsilon_j$. So the matrix of θ with respect to these bases is $G = (\langle e_i, e_j \rangle)_{i,j}$, and the rank of G is the dimension of the image of θ.

However, the kernel of θ is just the set of elements $u \in S^\lambda$ such that $\psi_u(v) = 0$ for all v – that is, $S^\lambda \cap (S^\lambda)^\perp$, and $S^\lambda / \ker \theta = D^\lambda$. From this we see that $\dim D^\lambda = \mathrm{rank}(G)$. But the entries of G are inner products of polytabloids, which are sums of $+1$ and -1, so the entries of G always lie in the prime subfield of F. Hence the rank of G will not change if F is extended to a bigger field, so the dimension of D^λ will also remain constant. □

From now on, we will assume that the ground field F has characteristic p, where p is either a prime or ∞. (The case $p = \infty$ is the case where F contains $\mathbb{Q}$ as a subfield, usually called "characteristic 0", but this non-standard notation is more convenient for our purposes.)

Lemma 3.5.5. $D^\lambda = 0$ if and only if λ is p-singular.

Proof. $D^\lambda = 0$, i.e. $S^\lambda \subseteq (S^\lambda)^\perp$, if and only if $\langle e_{t_1}, e_{t_2} \rangle = 0 \in F$ for every pair of polytabloids in S^λ. Equivalently, $D^\lambda = 0$ if and only if $p | g^\lambda$, where g^λ is defined as the highest common factor of all of the $\langle e_{t_1}, e_{t_2} \rangle \in \mathbb{Z}$ for every pair of polytabloids in S^λ, when the inner product is calculated over a field of characteristic 0.

Clearly $g^\lambda \,\Big|\, \prod_{i=1}^{r} (m_i!)^{\alpha_i}$ as in the proof of lemma 3.5.2, because $g^\lambda | \langle e_t, e_{t'} \rangle$ for some tableau t and its reverse t'; so, if λ is p-regular, then $D^\lambda \neq 0$.

Conversely, we can define an equivalence relation $\sim$ on the set of λ-tabloids: $[t_1] \sim [t_2]$ if we can obtain t_2 from t_1 by swapping rows of equal size; clearly, if $[t_1]$ is involved in e_t and $[t_1] \sim [t_2]$ then $[t_2]$ is also involved in e_t. Using the notation of lemma 3.4.1, the equivalence classes have size $\prod_{i=1}^{r} m_i!$, so any two polytabloids e_{s_1}, e_{s_2} have a multiple of this many tabloids in common up to sign. But the coefficients of $[t_1]$ and $[t_2]$ will either always be the same or always be opposite, depending only on $[t_1]$ and $[t_2]$, not on s_1 and s_2. If they are the same, the equivalence class will contribute $\prod m_i!$ to the inner product of e_{s_1} and e_{s_2}; otherwise the class will contribute $-\prod m_i!$.

This shows that the inner product will evaluate to a multiple of $\prod m_i!$, and so $\prod m_i! | g^\lambda$. Hence if λ is p-singular, then $D^\lambda = 0$. □

Lemma 3.5.6. Let λ and μ be partitions of n, with λ p-regular. Let $U \leq M^\mu$ be a submodule. If there exists a non-zero FS_n-homomorphism $\theta : S^\lambda \to M^\mu / U$, then $\lambda \geq \mu$.

Proof. Let t be a λ-tableau, and let t' be its reverse.

Then $\kappa_t(e_{t'}) = he_t$ for some constant $h \neq 0 \in F$ (lemma 3.5.2), and so $\kappa_t(\theta(e_{t'})) = \theta(\kappa_t(e_{t'})) = h\theta(e_t)$. But S^λ is generated by e_t (lemma 3.2.1), and θ is non-zero, so $\kappa_t(\theta(e_{t'})) \neq 0 \in M^\mu/U$. Hence there must be some μ-tableau s such that $\kappa_t([s]) \neq 0 \in M^\mu$, and so lemma 3.3.2 applies, showing that $\lambda \geq \mu$. □

Theorem 3.5.7. The irreducible representations of S_n are precisely the D^λ, where λ ranges through the p-regular partitions of n.

Proof. We have shown that the D^λ are irreducible; it remains to show that they are distinct. Suppose that $D^\lambda \simeq D^\mu$ as FS_n-modules. Then we have a map

$$S^\lambda \overset{\pi_\lambda}{\twoheadrightarrow} \frac{S^\lambda}{S^\lambda \cap (S^\lambda)^\perp} = D^\lambda \overset{\sim}{\to} D^\mu = \frac{S^\mu}{S^\mu \cap (S^\mu)^\perp} \hookrightarrow \frac{M^\mu}{S^\mu \cap (S^\mu)^\perp},$$

where π_λ is the canonical projection map. D^λ is non-zero, so π_λ has non-zero image, and all of the other maps are injective, so this composite is non-zero; it is an FS_n-homomorphism as each of its factors is. So the previous lemma applies, showing that $\lambda \geq \mu$. Similarly $\mu \geq \lambda$, so in fact $\lambda = \mu$. □

4 Characters of S_n

4.1 Notation

In this section, we will use the following notation.

- The ordinary irreducible character of S_n corresponding to the partition λ will be denoted by $\chi^\lambda : S_n \to \mathbb{C}$.

- The trivial character of a group G will be denoted by $\mathbb{1}_G$.
- The value of the character χ^λ on an element of cycle type μ will be denoted $\chi^\lambda(\mu)$.
- If λ is a partition of n, the S_n-conjugacy class of elements of cycle type λ will be written $\mathrm{ccl}(\lambda)$.

4.2 Calculation of the ordinary character table

Suppose we wish to evaluate the character table $X = (\chi^\lambda(\mu))$ of S_n, where λ and μ run through a complete set of partitions of n. To begin, it will be helpful to give the set of partitions of n a total ordering: we will take the lexicographic ordering, so that, for example,

$$\langle 1,1,1,1,\ldots,1\rangle < \langle 2,1,1,\ldots,1\rangle < \langle 2,2,\ldots,1\rangle < \cdots < \langle n-1,1\rangle < \langle n\rangle.$$

The method by Fox and Mullineux requires two auxiliary matrices A and B, defined as follows:

- $A = (a_{\lambda,\mu})$, where $a_{\lambda,\mu} = |S_\lambda \cap \mathrm{ccl}(\mu)|$, i.e. the number of elements of cycle type μ in S_λ, and
- $B = (b_{\lambda,\mu})$, where $b_{\lambda,\mu} = |S_\mu|\langle \chi^\lambda, \mathbb{1}_{S_\mu}\uparrow^{S_n}\rangle$, i.e. the number of times χ^λ appears as a summand of $\mathbb{1}_{S_\mu}\uparrow^{S_n}$ when written as a sum of irreducibles,

where S_λ and S_μ are Young subgroups, as defined earlier.

The matrix A is easy to calculate.

Example. The partitions of 4 are $\langle 1,1,1,1\rangle, \langle 2,1,1\rangle, \langle 2,2\rangle, \langle 3,1\rangle, \langle 4\rangle$. Call these $\lambda_1,\ldots,\lambda_5$ respectively. For example, to evaluate a_{λ_4,λ_2}, we need to count the number of elements of cycle type λ_2 in S_{λ_4}, i.e. the number of transpositions in S_3, which is 3. The matrix A for S_4 is given below:

$a_{\lambda,\mu}$	λ_1	λ_2	λ_3	λ_4	λ_5
λ_1	1	0	0	0	0
λ_2	1	1	0	0	0
λ_3	1	2	1	0	0
λ_4	1	3	0	2	0
λ_5	1	6	3	8	6

The matrix B is slightly more difficult to calculate. Introduce a new matrix:

- $M = (m_{\lambda,\mu})$, where $m_{\lambda,\mu} = \delta_{\lambda\mu}\left(\dfrac{n!}{|\mathrm{ccl}(\lambda)|}\right)$ and δ is the Kronecker delta.

We will calculate B with the help of a few lemmas:

Lemma 4.2.1. $XA^T = B$.

Proof.

$$
\begin{aligned}
b_{\lambda,\nu} &= |S_\nu|\langle \chi^\lambda, \mathbb{1}_{S_\nu}\uparrow^{S_n}\rangle \\
&= |S_\nu|\langle \chi^\lambda\downarrow_{S_\nu}, \mathbb{1}_{S_\nu}\rangle && \text{(Frobenius reciprocity)} \\
&= \sum_{g\in S_\nu} (\chi^\lambda(g)\cdot \underbrace{\mathbb{1}_{S_\nu}(g)}_{=1}) && \text{(by definition of inner product)} \\
&= \sum_\mu \left(\chi^\lambda(\mu)\cdot |S_\nu \cap \mathrm{ccl}(\mu)|\right) && \text{(as } \chi^\lambda \text{ is constant on } S_\nu\cap\mathrm{ccl}(\mu)) \\
&= \sum_\mu \left(\chi^\lambda(\mu)\cdot a_{\nu,\mu}\right) \\
&= (XA^T)_{\lambda,\nu}.
\end{aligned}
$$

□

Lemma 4.2.2. $B^TB = AMA^T$.

Proof. The characters $\{\chi^\mu | \mu \text{ a partition of } n\}$ form a basis for the class functions on S_n, and so we may write $\mathbb{1}_{S_\lambda}\uparrow^{S_n} = \sum_\mu c_\mu \chi^\mu$, $\mathbb{1}_{S_\nu}\uparrow^{S_n} = \sum_\mu d_\mu\chi^\mu$, for some constants c_μ, d_μ. By noting that this basis is orthonormal with respect to the inner product on characters, it is clear that

$$\sum_\mu \langle \chi^\mu, \mathbb{1}_{S_\lambda}\uparrow^{S_n}\rangle\langle \chi^\mu, \mathbb{1}_{S_\nu}\uparrow^{S_n}\rangle = \sum_\mu c_\mu d_\mu = \langle \mathbb{1}_{S_\lambda}\uparrow^{S_n}, \mathbb{1}_{S_\nu}\uparrow^{S_n}\rangle.$$

Hence

$$
\begin{aligned}
(B^T B)_{\lambda,\nu} &= \sum_{\mu} b_{\mu,\lambda} b_{\mu,\nu} \\
&= \sum_{\mu} |S_\lambda| \langle \chi^\mu, \mathbb{1}_{S_\lambda}\uparrow^{S_n} \rangle \cdot |S_\nu| \langle \chi^\mu, \mathbb{1}_{S_\nu}\uparrow^{S_n} \rangle \\
&= |S_\lambda||S_\nu| \langle \mathbb{1}_{S_\lambda}\uparrow^{S_n}, \mathbb{1}_{S_\nu}\uparrow^{S_n} \rangle && \text{(by the above)} \\
&= |S_\lambda||S_\nu| \langle (\mathbb{1}_{S_\lambda}\uparrow^{S_n})\downarrow_{S_\nu}, \mathbb{1}_{S_\nu} \rangle && \text{(Frobenius reciprocity)} \\
&= |S_\lambda| \sum_{g \in S_\nu} \mathbb{1}_{S_\lambda}\uparrow^{S_n}(g) \mathbb{1}_{S_\nu}(g) && \text{(by definition of inner product)} \\
&= |S_\lambda| \sum_{g \in S_\nu} \mathbb{1}_{S_\lambda}\uparrow^{S_n}(g) \\
&= |S_\lambda| \sum_{g \in S_\nu} \left[\frac{1}{|S_\lambda|} \sum_{\substack{x \in S_n, \\ x^{-1}gx \in S_\lambda}} \mathbb{1}_{S_\lambda}(x^{-1}gx) \right] && \text{(by definition of induction)} \\
&= \sum_{g \in S_\nu} \sum_{\substack{x \in S_n, \\ x^{-1}gx \in S_\lambda}} \mathbb{1}_{S_\lambda}(x^{-1}gx) \\
&= \sum_{g \in S_\nu} \sum_{\substack{x \in S_n, \\ x^{-1}gx \in S_\lambda}} \mathbb{1}_{S_\lambda}(g) && \text{(as } \mathbb{1}_{S_\lambda} \text{ is constant on ccl}(g)) \\
&= \sum_{g \in S_\nu} \#\{x \in S_n | x^{-1}gx \in S_\lambda\} && .
\end{aligned}
$$

Now note that the set $\{x \in S_n | x^{-1}gx \in S_\lambda\}$ forms a group which acts on S_λ by conjugation, and so by the orbit-stabiliser theorem, its order is

$\#\{x^{-1}gx \in S_\lambda\} \cdot \#\{x \in S_n | x^{-1}gx = g\}$

$= |S_\lambda \cap \mathrm{ccl}(g)| \cdot |\mathrm{stab}_{S_n}(g)|$.

But we also have that S_n acts on itself by conjugation, and

$|S_n| = |\mathrm{orb}_{S_n}(g)| \cdot |\mathrm{stab}_{S_n}(g)|$, i.e.

$|\mathrm{stab}_{S_n}(g)| = \dfrac{n!}{|\mathrm{ccl}(g)|}$.

Hence

$$
\begin{aligned}
(B^T B)_{\lambda,\nu} &= \sum_{g \in S_\nu} \#\{x \in S_n | x^{-1} g x \in S_\lambda\} \\
&= \sum_{g \in S_\nu} |S_\lambda \cap \mathrm{ccl}(g)| \cdot \frac{n!}{|\mathrm{ccl}(g)|} \\
&= \sum_{\mu} \frac{n!}{|\mathrm{ccl}(\mu)|} \cdot |S_\lambda \cap \mathrm{ccl}(\mu)| \cdot |S_\nu \cap \mathrm{ccl}(\mu)| \\
&= \sum_{\mu} \frac{n!}{|\mathrm{ccl}(\mu)|} \cdot a_{\lambda,\mu} \cdot a_{\nu,\mu} \\
&= \sum_{\mu} \frac{n!}{|\mathrm{ccl}(\mu)|} \cdot a_{\lambda,\mu} \cdot \left(\sum_{\xi} \delta_{\mu,\xi} a_{\nu,\xi} \right) \\
&= \sum_{\mu,\xi} a_{\lambda,\mu} \cdot m_{\mu,\xi} \cdot a_{\nu,\xi} = (AMA^T)_{\lambda,\nu}.
\end{aligned}
$$

□

Lemma 4.2.3. B is lower triangular with non-negative entries.

Proof. $\langle \chi^\lambda, \mathbb{1}_{S_\mu} \uparrow^{S_n} \rangle$, the number of times χ^λ appears in the character of M^μ, is always non-negative. Additionally, lemmas 3.5.6 and 2.2.2 show that the composition factors of M^μ can only be some of the S^λ for $\lambda \geq \mu$, so $\langle \chi^\lambda, \mathbb{1}_{S_\mu} \uparrow^{S_n} \rangle = 0$ whenever $\lambda < \mu$. □

If we know A, we can try to evaluate B. It follows from these two lemmas that B is easy to evaluate - starting from the bottom-right element and working from right-to-left, we can evaluate the entire last row of B element-by-element, with each equation having only one unknown (as B is lower triangular) and a unique solution (as the elements of B are non-negative). We can then continue up the rows. This is most easily illustrated with an example!

Example. Continuing with the example of S_4, we now know that B takes the form

$$
B = \begin{pmatrix}
b_{1,1} & 0 & 0 & 0 & 0 \\
b_{2,1} & b_{2,2} & 0 & 0 & 0 \\
b_{3,1} & b_{3,2} & b_{3,3} & 0 & 0 \\
b_{4,1} & b_{4,2} & b_{4,3} & b_{4,4} & 0 \\
b_{5,1} & b_{5,2} & b_{5,3} & b_{5,4} & b_{5,5}
\end{pmatrix},
$$

where $b_{i,j}$ means b_{λ_i,λ_j}. It is trivial to evaluate M and hence AMA^T, which in this case turns out to be

$$AMA^T = \begin{pmatrix} 24 & 24 & 24 & 24 & 24 \\ 24 & 28 & 32 & 36 & 48 \\ 24 & 32 & 48 & 48 & 96 \\ 24 & 36 & 48 & 72 & 144 \\ 24 & 48 & 96 & 144 & 576 \end{pmatrix}.$$

We need to solve the matrix equation $B^T B = AMA^T$ for B. In full, we need to solve:

$$\begin{pmatrix} b_{1,1} & b_{2,1} & b_{3,1} & b_{4,1} & b_{5,1} \\ 0 & b_{2,2} & b_{3,2} & b_{4,2} & b_{5,2} \\ 0 & 0 & b_{3,3} & b_{4,3} & b_{5,3} \\ 0 & 0 & 0 & b_{4,4} & b_{5,4} \\ 0 & 0 & 0 & 0 & b_{5,5} \end{pmatrix} \begin{pmatrix} b_{1,1} & 0 & 0 & 0 & 0 \\ b_{2,1} & b_{2,2} & 0 & 0 & 0 \\ b_{3,1} & b_{3,2} & b_{3,3} & 0 & 0 \\ b_{4,1} & b_{4,2} & b_{4,3} & b_{4,4} & 0 \\ b_{5,1} & b_{5,2} & b_{5,3} & b_{5,4} & b_{5,5} \end{pmatrix} = \begin{pmatrix} 24 & 24 & 24 & 24 & 24 \\ 24 & 28 & 32 & 36 & 48 \\ 24 & 32 & 48 & 48 & 96 \\ 24 & 36 & 48 & 72 & 144 \\ 24 & 48 & 96 & 144 & 576 \end{pmatrix}$$

for each $b_{i,j}$. (In principle, this is 15 equations – one of which is quintic – in 15 unknowns, but choosing the equations in a sensible order reduces them to 15 linear equations each in one unknown!) Proceed as follows:

- Start in the bottom-right corner and equate the $(5,5)$-elements of $B^T B$ and AMA^T: this gives us the equation $b_{5,5}^2 = 576$, from which we can deduce that $b_{5,5} = 24$.
- Now we can consider the $(5,i)$-elements for each $i = 1,2,3,4$. These give us the values of $b_{5,i} b_{5,5}$, i.e. $24 b_{5,i}$, so we can immediately calculate the $b_{5,i}$. Now we have found the whole of the bottom row!
- Continue with the $(4,4)$-element and then the fourth row...

Eventually, we get:

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 3 & 2 & 0 & 0 & 0 \\ 2 & 2 & 4 & 0 & 0 \\ 3 & 4 & 4 & 6 & 0 \\ 1 & 2 & 4 & 6 & 24 \end{pmatrix}.$$

Lemma 4.2.4. *A is invertible.*

Proof. A is lower triangular: if $\lambda < \mu$ in the lexicographic ordering, then S_λ contains no elements of cycle type μ; however, for each λ, S_λ clearly contains at least one element of cycle type λ, so the diagonal elements of A are all non-zero. Hence the determinant of A is non-zero. □

Combining this result with lemma 4.2.1, we have proven:

Theorem 4.2.5. The ordinary character table is given by $X = B(A^T)^{-1}$. □

Example. Finally, the character table for S_4 is

$$X = \begin{pmatrix} 1 & -1 & 1 & 1 & -1 \\ 3 & -1 & -1 & 0 & 1 \\ 2 & 0 & 2 & -1 & 0 \\ 3 & 1 & -1 & 0 & -1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Of course, the character table is much more commonly seen with a slightly different ordering of the rows:

		λ_1	λ_2	λ_3	λ_4	λ_5
$\mathbb{1}_{S_5} = \chi_1$	λ_5	1	1	1	1	1
sgn $= \chi_2$	λ_1	1	-1	1	1	-1
χ_3	λ_4	3	1	-1	0	-1
$\chi_2 \cdot \chi_3 = \chi_4$	λ_2	3	-1	-1	0	1
χ_5	λ_3	2	0	2	-1	0

This method is a fairly efficient method of finding the entire character table, as it mostly consists of basic matrix manipulation that a computer can handle easily. It is, of course, horrendously inefficient if we only need to work out a *single* element in the table. However, due to Murnaghan and Nakayama, there is a much more efficient method of doing this, often even efficient enough be performed fairly quickly by hand for reasonably small symmetric groups. The following sections will introduce and demonstrate this method.

4.3 The Murnaghan-Nakayama rule

4.3.1 Hooks and skew-hooks

Definition. Let λ be a partition of n. If $(i, j) \in \lambda$, the (i, j)-*hook* of λ is the union of the (i, j)-node of the Young diagram of λ along with all those nodes $(i, j + a)$, $(i + b, j) \in \lambda$ for $a, b > 0$, i.e. the nodes to the right of and below the (i, j)-node. The *height* of the (i, j)-hook h is the number of nodes strictly below (i,j) that are in λ, denoted $\mathrm{ht}(h)$; the *size* of the (i,j)-hook is the number of nodes in the hook.

Example. The $(1,2)$-hook of $\langle 6,3,3,2\rangle$ has height 3 and size 8:

$$\begin{array}{ccccccl} \times & \otimes & \otimes & \otimes & \otimes & \otimes & \leftarrow \text{row 1} \\ \times & \otimes & \times & & & & \\ \times & \otimes & \times & & & & \\ \times & \otimes & & & & & \\ & \uparrow & & & & & \\ & \text{col 2} & & & & & \end{array}$$

Definition. Similarly, the (i,j)*-skew hook* of λ is the unique connected part of λ whose endpoints are the same as those of the (i,j)-hook but which must crawl strictly along the lower-right edge of λ. The *height* and *size* of the (i,j)-skew hook are the same as those of the (i,j)-hook.

Example. The $(1,2)$-skew hook of $\langle 6,3,3,2\rangle$:

$$\begin{array}{ccccccl} \times & \times & \otimes & \otimes & \otimes & \otimes & \leftarrow \text{row 1} \\ \times & \times & \otimes & & & & \\ \times & \otimes & \otimes & & & & \\ \times & \otimes & & & & & \\ & \uparrow & & & & & \\ & \text{col 2} & & & & & \end{array}$$

It is easy to see that, if we remove every node of a skew hook from the Young diagram of a partition, we are left with a valid diagram that corresponds to another partition. For example, removing the $(1,2)$-skew hook from $\langle 6,3,3,2\rangle$ above would leave us with the diagram for $\langle 2,2,1,1\rangle$. We will denote this removal of the skew-hook h from the partition λ by $\lambda - h$; so, if λ is a partition of n, and h is a skew-hook of λ of size k, then $\lambda - h$ is a partition of $n-k$.

4.3.2 Statement of the rule

The Murnaghan-Nakayama rule gives a simple relation between entries in the character table of S_n and entries in the character tables of smaller symmetric groups - in particular, it relates the character associated with a partition λ to the characters associated with the partitions obtained by removing skew hooks from λ. It can be used recursively to calculate a single entry in the character table of S_n.

Theorem 4.3.1. (The Murnaghan-Nakayama rule.) Suppose λ and $\mu = \langle \mu_1, \ldots, \mu_s\rangle$ are partitions of n, and $\sigma \in S_n$ is an element of cycle type μ. Further, let $\mu' = \langle \mu_2, \ldots, \mu_s\rangle$ be the partition of $n - \mu_1$ obtained by simply removing the first row from μ, and let $\sigma' \in S_{n-\mu_1}$ be an element of cycle type μ'. The value of the character χ^λ on (the conjugacy class of) σ is given by

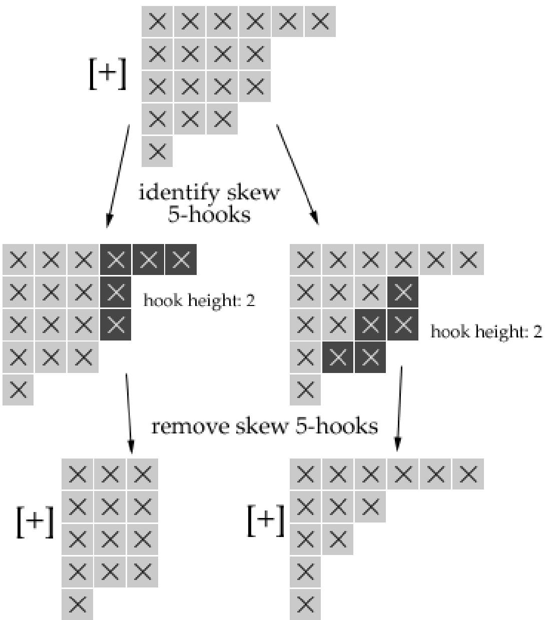
$$\chi^\lambda(\sigma) = \sum_h (-1)^{\mathrm{ht}(h)} \chi^{(\lambda - h)}(\sigma')$$

where the sum ranges over all skew hooks h in λ of size μ_1. (We adopt the convention that the only partition of 0 is the empty partition $\mathbf{0}$, and that S_0 is a group of order 1, with $\chi^{\mathbf{0}}$ taking the value 1 on this unique element.)

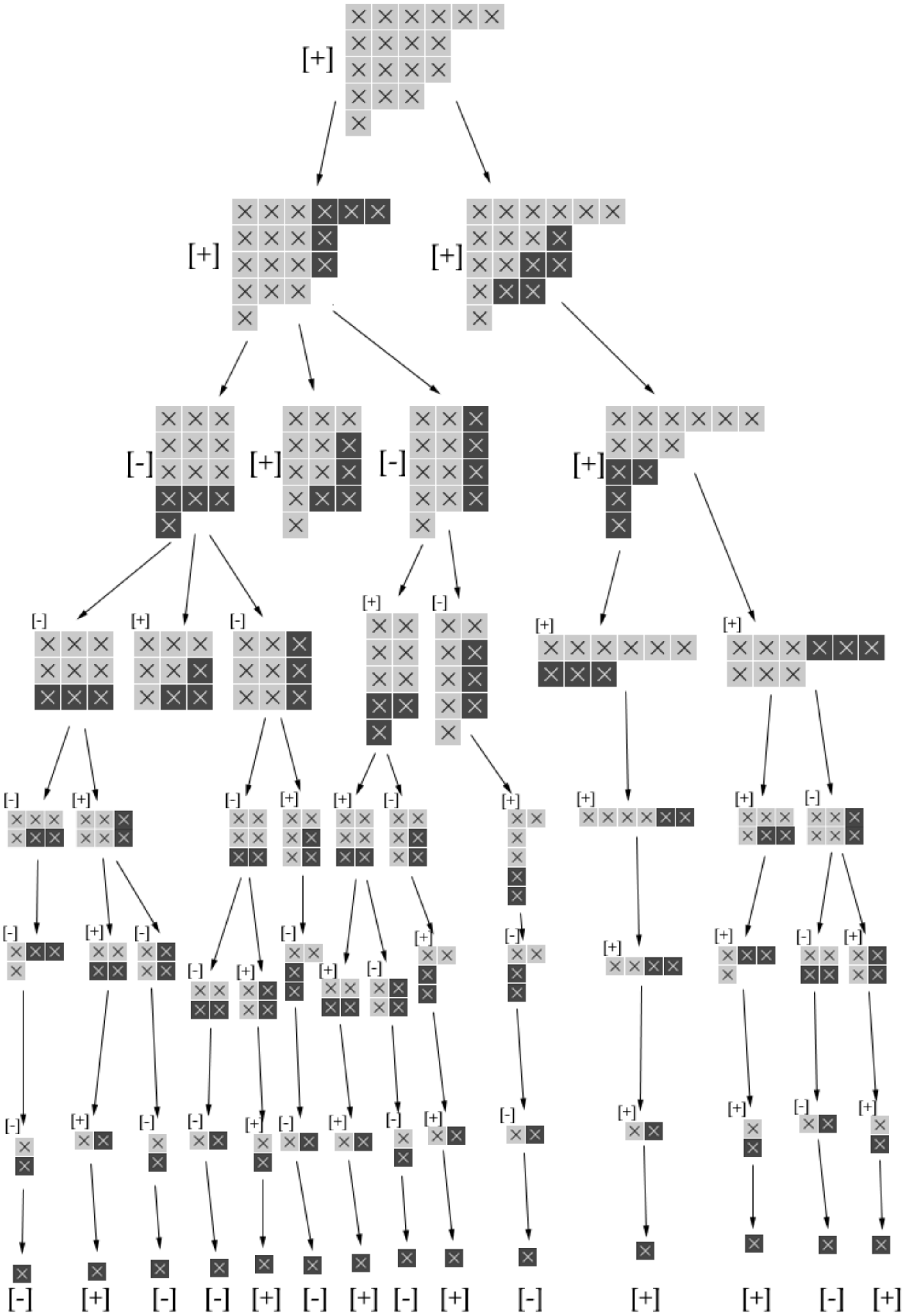
Proof. Omitted. (See James [1].) □

The proof of the Murnaghan-Nakayama rule follows nicely from a lot of beautiful theory about the composition of the modules M^λ. Unfortunately, the development of the machinery required to do this is long and beyond the scope of this essay, and any proof of the Murnaghan-Nakayama rule that does not explicitly rely on this theory is artificial and unenlightening, so we omit the proof.

Example. To calculate the character of $D^{\langle 6,4,4,3,1 \rangle}$ evaluated at an element of S_{18} of cycle type $(5, 4, 3, 2, 2, 1, 1)$, we must first draw the Young diagram of the partition, and attach the sign $+$ to it. We then remove a skew hook of size 5 in each way possible, and record the partitions of S_{13} that this leaves us with; we switch the sign (from $+$ to $-$ or vice-versa) if the hook has odd height, and we do not switch the sign if it has even height:



We must then calculate the character of each of these remaining partitions at an element of cycle type $(4, 3, 2, 2, 1, 1)$ by further removing all skew 4-hooks and appending the appropriate sign, etc. Eventually, the diagram looks like this:



Character value = -1 + 1 - 1 - 1 + 1 - 1 + 1 - 1 + 1 - 1 + 1 + 1 - 1 + 1 = 0.

Example. Earlier, we found an explicit basis for the Specht module, and determined that $S^{\langle 4,2,1 \rangle}$ had dimension 35 over $\mathbb{C}$ (actually, we showed the dimension was the same over any field F). Since we are working over $\mathbb{C}$, the bilinear form on polytabloids is an inner product, so $S^\lambda = D^\lambda$, and an alternative method would have been to calculate the character of $D^{\langle 4,2,1 \rangle}$ on the conjugacy class containing the identity, i.e. on $\langle 1,1,1,1,1,1,1 \rangle$, using the Murnaghan-Nakayama rule. We verify that this gives the same answer.

$$
\begin{aligned}
&\chi^{(4,2,1)} && \text{on } \langle 1,1,1,1,1,1,1 \rangle \\
&= \chi^{(3,2,1)} + \chi^{(4,1,1)} + \chi^{(4,2)} && \text{on } \langle 1,1,1,1,1,1 \rangle \\
&= (\chi^{(2,2,1)} + \chi^{(3,1,1)} + \chi^{(3,2)}) + (\chi^{(3,1,1)} + \chi^{(4,1)}) + (\chi^{(3,2)} + \chi^{(4,1)}) && \text{on } \langle 1,1,1,1,1 \rangle \\
&= \chi^{(2,2,1)} + 2\chi^{(3,1,1)} + 2\chi^{(3,2)} + 2\chi^{(4,1)} && \text{(tidying up)} \\
&= (\chi^{(2,1,1)} + \chi^{(2,2)}) + 2(\chi^{(2,1,1)} + \chi^{(3,1)}) + 2(\chi^{(2,2)} + \chi^{(3,1)}) && \\
&\qquad +2(\chi^{(3,1)} + \chi^{(4)}) && \text{on } \langle 1,1,1,1 \rangle \\
&= 3\chi^{(2,1,1)} + 3\chi^{(2,2)} + 6\chi^{(3,1)} + 2\chi^{(4)} && \text{(tidying up)} \\
&= 3(\chi^{(1,1,1)} + \chi^{(2,1)}) + 3(\chi^{(2,1)}) + 6(\chi^{(2,1)} + \chi^{(3)}) + 2(\chi^{(3)}) && \text{on } \langle 1,1,1 \rangle \\
&= 3\chi^{(1,1,1)} + 12\chi^{(2,1)} + 8\chi^{(3)} && \text{(tidying up)} \\
&= 3(\chi^{(1,1)}) + 12(\chi^{(1,1)} + \chi^{(2)}) + 8(\chi^{(2)}) && \text{on } \langle 1,1 \rangle \\
&= 15\chi^{(1,1)} + 20\chi^{(2)} && \text{(tidying up)} \\
&= 15(\chi^{(1)}) + 20(\chi^{(1)}) && \text{on } \langle 1 \rangle \\
&= 35\chi^{(1)} && \text{(tidying up)} \\
&= 35\chi^{\mathbf{0}}(\mathrm{id}) = 35.
\end{aligned}
$$

5 Conclusion

Much more work has been done on the representation theory of the symmetric groups than it was possible to include in a short essay, but it is worth pointing out that the field still contains many basic open problems. For example, the composition factors of M^λ are all known when the field has characteristic 0, and some results are known when the field has prime characteristic, but the problem of characterising the composition factors in general is still far from having been solved – this is perhaps one of the biggest open problems in representation theory.

Nonetheless, the theory has already seen many applications within representation theory, for example in finding the representations of $\mathrm{Sym}(\mathbb{N})$ and $\mathrm{GL}_n(F)$.

References

[1] G. D. James, *The representation theory of the symmetric groups*. Springer, 1978.

[2] G. de B. Robinson, *Representation theory of the symmetric group*. Edinburgh University Press, 1961.

[3] G. D. James and M. W. Liebeck, *Representations and characters of groups*. Cambridge University Press, 1993.