MA5P2 The Symmetric Groups

Daan Krammer

November 21, 2008

Contents

In this order we do:

- Sections 7.1–7.17 except 7.8 and 7.16 in Stanley's book.
- Background on representations of finite groups.
- Section 7.18.
- Specht modules.
- More to come!

1 Background on representations of finite groups

We give the background on representations of finite groups without proofs. It is perfectly possible to use these results, as we do, without knowing how to prove them.

Throughout we fix an algebraically closed field K of characteristic 0. All our vector spaces are over K unless stated otherwise. We fix a finite group G throughout.

1.1 G-modules

If V is a vector space over K, let $\mathrm{GL}(V)$ denote the set of bijective K-linear maps $V \to V$.

A (left) *G-module* is a finite-dimensional vector space V together with an action $G \times V \to V$ such that, for all $g \in V$, the map $x \mapsto gx$ is K-linear.

A *submodule* of a G-module V is a linear subspace of V invariant under G.

A nonzero G-module is said to be *irreducible* if it doesn't have any submodules different from 0 and itself.

A *homomorphism* or simply *map* of G-modules $f\colon U \to V$ is a linear map such that $f(gx) = g(fx)$ for all $x \in U$, $g \in G$. The set of maps of G-modules $U \to V$ is written $\operatorname{Hom}(U, V) = \operatorname{Hom}_G(U, V)$.

Lemma 1 (Schur's lemma)**.** *Let U, V be irreducible G-modules. Then*

$$\dim \operatorname{Hom}(U, V) \cong \begin{cases} 1 & \text{if } U \cong V, \\ 0 & \text{otherwise.} \end{cases}$$

□

A G-module is called *isotypical* if it is isomorphic to a direct sum of a number of copies of the same irreducible module. An *isotypical component* of a G-module is a maximal isotypical submodule.

Proposition 2. *Every G-module is the direct sum of its isotypical components.*

1.2 The group algebra

Let R be an associative ring (always with identity). A *(left) R-module* is a $\mathbb{Z}$-module ${}_RV = V$ together with a *scalar multipliciation* $R \times V \to V$ such that

$$a(bx) = (ab)x, \quad a(x+y) = ax + ay, \quad (a+b)x = ax + bx$$

for all $a, b \in$, $x, y \in V$. Similar for right R-modules. In particular, if R is a field, then an R-module is nothing but a vector space over R.

The *group algebra* of a finite group G is a vector space KG with basis G. Thus, the elements of KG are formal linear combinations

$$\sum_{g \in G} a_g\, g$$

where $a_g \in K$ for all $g \in G$. The group algebra is equipped with an associative K-linear map $KG \times KG \to KG$ (that's why it is called an algebra) by extending the multiplication in G by linearity. In other words,

$$\Big(\sum_{g \in G} a_g\, g\Big)\Big(\sum_{g \in G} b_g\, g\Big) := \sum_{g \in G} c_g\, g \qquad \text{where} \qquad c_g = \sum_{h \in G} a_h\, b_{h^{-1}g}.$$

Let V be a left G-module and let $a = \sum_{g \in G} a_g\, g \in KG$ where $a_g \in K$ for all $g \in G$. Let $x \in V$. We define

$$ax := \sum_{g \in G} a_g\, (gx) \in V.$$

One can show that this makes V into a left KG-module. Conversely, every KG is of this form, for a unique G-module (not just up to isomorphism!). Thus:

Proposition 3. *There is a natural bijection between the set* [(1)] *of G-modules and the set of KG-modules.* □

[(1)] More correct would be to say *class* of G-modules

1.3 Tensor products

Let $V_R = V$ be a right R-module and ${}_RW = W$ a left R-module. A *tensor product* $V \otimes_R W$ is defined to be a pair (T, f) of a $\mathbb{Z}$-module T and a $\mathbb{Z}$-bilinear map $f\colon V \times W \to T$ such that $f(xa, y) = f(x, ay)$ for all $x \in V$, $y \in W$, $a \in R$ satisfying the *universal property* that for all pairs (T', f') with the above properties, there exists a unique $\mathbb{Z}$-linear map $g\colon T \to T'$ making the following diagram commute:

$$\begin{array}{ccc} & \overset{f}{\nearrow} & T \\ V \times W & & \downarrow g \\ & \underset{f'}{\searrow} & T'. \end{array} \tag{4}$$

It can be proved that such a pair (T, f) exists and is unique in the sense that if (T, f') is another such pair then there exists a commutative diagram (4) with g an isomorphism. We call T the *tensor product* and it is written $V \otimes_R W$. We call $f\colon V \times W \to V \otimes W$ the *natural map* and usually write $x \otimes y$ instead of $f(x, y)$.

Suppose that $\{v_i \mid i \in I\}$ is a basis of a vector space V and $\{w_j \mid j \in J\}$ is of a vector space W. Then $\{v_i \otimes w_j \mid i \in I,\ j \in J\}$ is a basis of the tensor product $V \otimes_K W$ (which we write as $V \otimes W$ henceforth).

If V, W are two G-modules then we make $V \otimes W$ into a G-module by defining the action by $g(x \otimes y) = (gx) \otimes (gy)$. The direct sum $V \oplus W = V \times W$ is made into G-module the obvious way.

1.4 Class functions

A *class function* on G is a function $G \to K$ which is constant on conjugacy classes. The vector space of class functions on G will be denoted by $\mathrm{CF}(G)$.

We define an inner product (that is, symmetric bilinear form)

$$\langle \cdot, \cdot \rangle\colon \mathrm{CF}(G) \times \mathrm{CF}(G) \to K, \qquad \langle f, g \rangle := \tfrac{1}{\#G} \sum_{w \in G} f(w)\, g(w).$$

Note that we don't use the complex conjugate of $g(w)$ as some do.

With a G-module V we associate the *character* $\chi_V\colon G \to K$ defined by $\chi_V(w) = \mathrm{tr}_V(w)$ (trace). This is a class function. We have

$$\chi_{V \oplus W} = \chi_V + \chi_W, \qquad \chi_{V \otimes W} = \chi_V\, \chi_W.$$

Proposition 5.

(a) *A G-module is determined (up to isomorphism) by its character.*
(b) *The irreducible characters form an orthonormal basis of* $\mathrm{CF}(G)$.

1.5 Restriction and induction

In this subsection, let $H \subset G$ be finite groups.

Every G-module V gives rise to an H-module by restricting the action to H; this H-module is variously denoted by

$$\mathrm{Res}^G_H(V) = V \downarrow H = V[G \downarrow H]$$

and called the *restriction* of V to H.

Conversely, if V is an H-module we define the *induced* G-module to be

$$\mathrm{Ind}_H^G(V) = V \uparrow G = V[H \uparrow G] := KG \otimes_{KH} V.$$

In order to define this tensor product one considers KG as a right KH-module; to make it into a G-module one uses the left G-action on KG.

An alternative and more elementary definition of induced module can be found in exercise 7.

If $W = V \downarrow H$ and ϕ, χ are the characters of (respectively) V, W then we also write $\chi = \phi \downarrow H$. By proposition 5 we could even extend this notation to arbitrary class functions. Likewise for induced modules.

Proposition 6 (Frobenius reciprocity). *For all $\chi \in \mathrm{CF}(G)$, $\phi \in \mathrm{CF}(H)$ we have $\langle \chi \downarrow H, \phi \rangle = \langle \chi, \phi \uparrow G \rangle$.*

Exercise 7. Let $H \subset G$ be finite groups. Let V be a (left) H-module and define

$$W := \{ p\colon G \to V \mid p(xy) = x \cdot p(y) \text{ for all } x \in H,\ y \in G \}.$$

(a) Prove that $\dim W = [G : H]\ \dim V$.
(b) For $p \in W$, $g \in G$, define $gp = q\colon G \to V$ by $q(x) = p(xg)$. Prove that $qp \in W$.
(c) Prove that the map $G \times W \to W$, $(g, p) \mapsto gp$ makes W into a left G-module.
(d) Define $G \times V \to W$, $(g, v) \mapsto g \boxtimes v$ by $(g \boxtimes v)(x) = xg(v)$ for all $x \in G$. We extend this by linearity to $KG \times V \to W$. Prove that the pair $(W, \boxtimes)$ is a tensor product $KG \otimes_{KH} V$. So W is the induced module $V \uparrow G$.
(e) Prove Frobenius reciprocity using the above.

2 Specht modules

A *numbering* of the diagram of a partition is a way of putting the numbers 1, ..., n in its boxes, using each number precisely once. We shall denote such numberings by T, $\widehat{T}$.

If $\lambda \vdash n$ then S_n acts on the set of numberings of shape λ: a permutation σ takes a numbering T to σT which has $\sigma(i)$ in the box where T has i.

A numbering T gives rise to a subgroup $R(T)$ of S_n called the *row group* of T, which consists of those permutations $\sigma \in S_n$ such that $\{i, \sigma(i)\}$ is in a single row, for all $i \in [n] := \{1, \ldots, n\}$.

If T is a numbering of shape $\lambda = (\lambda_1, \lambda_2, \ldots)$ then $R(T)$ is isomorphic to $S_{\lambda_1} \times S_{\lambda_2} \times \cdots$.

Likewise we have the *column group* $C(T)$ of the permutations that preserve the columns.

Notice that for all $w \in S_n$

$$R(wT) = w\, R(T)\, w^{-1}, \qquad C(wT) = w\, C(T)\, w^{-1}.$$

Lemma 8. *Let T, $\widehat{T}$ be numberings of shapes, respectively, λ, $\widehat{\lambda}$ where $\lambda, \widehat{\lambda} \vdash n$. Assume that λ does not strictly dominate $\widehat{\lambda}$. Then exactly one of the following occurs.*

(1) *There are distinct integers which are in a single row of $\widehat{T}$ and in a single column of T.*
(2) *We have $\lambda = \widehat{\lambda}$ and there are $\widehat{p} \in R(\widehat{T})$, $q \in C(T)$ such that $\widehat{p}\,\widehat{T} = qT$.*

Proof. Suppose that (1) and (2) both hold. By (1), there are two distinct numbers in a single row of $\widehat{T}$ and in a single column of T. These numbers are then also in a single row of $\widehat{p}\,\widehat{T}$ and a single column of qT. But $\widehat{p}\,\widehat{T} = qT$, a contradiction.

Assuming now that (1) is false, it remains to prove that (2) holds.

The entries of the first row of $\widehat{T}$ must occur in different columns of T, so there is a $q_1 \in C(T)$ such that these entries occur in the first row of $q_1 T$.

The entries of the second row of $\widehat{T}$ occur in different columns of T so also of $q_1 T$. Therefore there exists $q_2 \in C(q_1 T) = C(T)$ not moving the entries equal to those in the first row of $\widehat{T}$, so that these entries all occur in the first two rows of $q_2 q_1 T$.

Continuing this way, we get $q_1, \ldots, q_k$ in $C(T)$ such that the entries in the first k rows of $\widehat{T}$ occur in the first k rows of $q_k \cdots q_1 T$. In particular since T and $q_k \cdots q_1 T$ have the same shape, it follows that

$$\widehat{\lambda}_1 + \cdots + \widehat{\lambda}_k \leq \lambda_1 + \cdots + \lambda_k.$$

Since this is true for all k we have $\widehat{\lambda} \leq \lambda$ (domination).

We assumed that λ does not strictly dominate $\widehat{\lambda}$ so $\lambda = \widehat{\lambda}$. Let ℓ be the number of rows of λ and $q = q_\ell \cdots q_1$. Then qT and $\widehat{T}$ have the same entries in each row. This means that there exists $\widehat{p} \in R(\widehat{T})$ such that $\widehat{p}\,\widehat{T} = qT$ and (2) follows. □

The *column word* of a numbering T is the sequence of integers which begins by listing the entries of the first column from bottom to top, then those in the second column, and so on. For example, the column word of

3	1	4
5	2	

is 53214.

We define a total ordering $<$ on the set of numberings of diagrams of all shapes by $T < \widehat{T}$ if and only if $\mathrm{sh}(T) <_R \mathrm{sh}(\widehat{T})$ or $\mathrm{sh}(T) = \mathrm{sh}(\widehat{T})$ and the largest entry in different boxes occurs earlier in the column word for $\widehat{T}$ than for T.

Clearly, if T is a standard tableau, then for all $p \in R(T)$ and $q \in C(T)$,

$$T \leq pT, \qquad qT \leq T. \tag{9}$$

Corollary 10. *If T, $\widehat{T}$ are standard tableaux and $\widehat{T} > T$ then there are two distinct integers in a single row of $\widehat{T}$ and in a single column of T.*

Proof. We have not $\mathrm{sh}(\widehat{T}) < \mathrm{sh}(T)$ because otherwise $\mathrm{sh}(\widehat{T}) <_R \mathrm{sh}(T)$ and therefore $\widehat{T} < T$, contradicting the assumption $\widehat{T} > T$. By lemma 8, if the pair doesn't exist, we are in case (2). Using (9), it follows that $\widehat{T} \leq \widehat{p}\,\widehat{T} = qT \leq T$, a contradiction. $\square$

A *tabloid* is an equivalence class of numberings of a diagram, two numberings T, $\widehat{T}$ being equivalent if corresponding rows contain the same entries. The tabloid containing T is written $\{T\}$. Thus, $\{T\} = \{\widehat{T}\}$ if and only if there exists $p \in R(T)$ with $pT = \widehat{T}$.

Tabloids are sometimes displayed without the vertical lines between the boxes:

$$\begin{array}{|lll|}\hline 5 & 1 & 6 \\ \hline 4 & 7 & \\ \cline{1-2} 2 & 3 & \\ \cline{1-2}\end{array} \quad = \quad \begin{array}{|lll|}\hline 5 & 6 & 1 \\ \hline 4 & 7 & \\ \cline{1-2} 2 & 3 & \\ \cline{1-2}\end{array}\,.$$

The symmetric group S_n acts on the set of tabloids by $w\{T\} = \{wT\}$. As S_n-set, the orbit of $\{T\}$ is isomorphic to $S_n/R(T)$ on which S_n acts by left multiplication.

For a numbering T we define

$$a_T := \sum_{p \in R(T)} p \in \mathbb{C}S_n, \qquad b_T := \sum_{q \in C(T)} \varepsilon_q\, q \in \mathbb{C}S_n.$$

The element $c_T := b_T\, a_T$ is called the *Young symmetrizer.*

Exercise 11.

(a) For all $p \in R(T)$ and $q \in C(T)$ we have

$$p\, a_T = a_T\, p = a_T, \qquad q\, b_T = b_T\, q = \varepsilon_q\, b_T.$$

(b) We have $a_T a_T = [\#R(T)]\, a_T$ and $b_T b_T = [\#C(T)]\, b_T$.

Definition 12. We define M^λ to be the complex vector space whose basis is the set of tabloids of shape λ. In other words, elements of M^λ are formal complex linear combinations of tabloids.

The symmetric group S_n acts on the set of tabloids and thereby on M^λ. So M^λ is an S_n-module. As S_n-module, M^λ is isomorphic to $1[S_\lambda \uparrow S_n]$ which we used before.

All we need to know about modules is that if $a = \sum_{w \in S_n} a_w\, w \in \mathbb{C}S_n$ is an element of the group algebra and $x \in V$ is an element in some S_n-module V, then we define, as a very useful shorthand, $ax := \sum_{w \in S_n} a_w\,(wx) \in \mathbb{C}V$.

For every numbering T we define

$$v_T := b_T\{T\} = \sum_{q \in C(T)} \varepsilon_q\, \{qT\} \in M^\lambda.$$

Exercise 13. We have $w\, v_T = v_{wT}$ for all $w \in S_n$ and all numberings T of shape $\mathrm{sh}(T) \vdash n$.

Lemma 14. *Let T, $\widehat{T}$ be numberings of shapes λ, $\widehat{\lambda}$. Assume that λ does not strictly dominate $\widehat{\lambda}$. If there is a pair of integers in the same row of $\widehat{T}$ and in the same column of T, then $b_T\{\widehat{T}\} = 0$. Otherwise, $b_T\{\widehat{T}\} \in \{v_T, -v_T\}$.*

Proof. If such a pair exists, let $s \in S_n$ be the transposition that interchanges them. Then $b_T s = -b_T$ because $s \in C(T)$ (and using exercise 11). Also $s\{\widehat{T}\} = \{\widehat{T}\}$ because $s \in R(\widehat{T})$. It follows that

$$b_T\{\widehat{T}\} = b_T\{s\,\widehat{T}\} = (b_T\, s)\{\widehat{T}\} = -b_T\{\widehat{T}\}$$

whence $b_T\{\widehat{T}\} = 0$. If such a pair doesn't exist, let $\widehat{p}$ and q be as in (2) of lemma 8. We find

$$b_T\{\widehat{T}\} = b_T\{\widehat{p}\,\widehat{T}\} = b_T\{qT\} = (b_T\, q)\{T\} = \varepsilon_q\, b_T\{T\} = \varepsilon_q\, v_T. \qquad \square$$

Corollary 10 immediately gives:

Corollary 15. *If T, $\widehat{T}$ are standard tableaux and $\widehat{T} > T$ then $b_T\{\widehat{T}\} = 0$.* □

Definition 16. The *Specht module* S^λ is the subspace of M^λ spanned by the v_T where T varies over the numberings of λ.

By exercise 13, S^λ is invariant under the S_n-action. Note that S^λ is spanned (as vector space) $S_n\, v_T = \{wT \mid w \in S_n\}$ for any numbering T of λ. We say that S^λ is generated by v_T as S_n-module.

We have $v_T \neq 0$ for all T.

Exercise 17. For $\lambda = (n) \vdash n$ the module $S^{(n)}$ is the 1-dimensional trivial S_n-module. For $\lambda = (1^n) \vdash N$ the module $S^{(1^n)}$ is isomorphic to $\mathbb{A}_n$, the alternating S_n-module.

Proposition 18. *The S^λ are pairwise non-isomorphic irreducible S_n-modules. Every irreducible S_n-module is isomorphic to one of them.*

Proof. For all $w \in S_n$ and all numberings T such that $\operatorname{sh}(T) \vdash n$ we have $b_T\,\{wT\} = (b_T\, w)\,\{T\} = \varepsilon_w\, b_T\,\{T\} = \varepsilon_w\, v_T$ which shows that

$$b_T\{wT\} = \varepsilon_w\, v_T. \tag{19}$$

Morever, $b_T\, v_T = b_T b_T\{T\} = [\#C(T)]\, b_T\{T\} = [\#C(T)]\, v_T$ which proves that

$$b_T\, v_T = [\#C(T)]\, v_T. \tag{20}$$

Next, we prove that if $\operatorname{sh}(T) = \lambda$, $\operatorname{sh}(\widehat{T}) = \widehat{\lambda}$ then:

- $b_T\, M^\lambda = b_T\, S^\lambda = \mathbb{C}\, v_T.$ (21)
- $b_T\, M^{\widehat{\lambda}} = b_T\, S^{\widehat{\lambda}} = 0$ if $\widehat{\lambda} >_R \lambda.$ (22)

To prove (21), observe that $b_T\, S^\lambda \subset b_T\, M^\lambda \overset{(19)}{\subset} \mathbb{C}\, v_T \overset{(20)}{\subset} b_T\, S^\lambda$. In order to prove (22), let $\widehat{\lambda} >_R \lambda$. Then λ does not strictly dominate $\widehat{\lambda}$. By lemma 8, we have precisely one of two possibilities (1), (2). But (2) implies $\lambda = \widehat{\lambda}$, contradicting our assumption that $\widehat{\lambda} >_R \lambda$. Thus we have (1), that is, there is a pair of distinct integers in a single row of $\widehat{T}$ and a single column of T. By lemma 14 it follows that $b_T\{\widehat{T}\} = 0$. This finishes the proof of (22).

Since $b_T\, S^\lambda \neq 0$ and $b_T\, S^{\widehat{\lambda}} = 0$ whenever $\widehat{\lambda} >_R \lambda$ the Specht modules are pairwise non-isomorphic.

Let us next prove that the Specht modules are irreducible. Suppose that $S^\lambda = V \oplus W$ for S_n-submodules V, W. Then

$$\mathbb{C}\, v_T = b_T\, S^\lambda = b_T V \oplus b_T W$$

so v_T is contained in V or W, say in V. But S^λ is generated, as S_n-module, by v_T so $V = S^\lambda$. This proves that the Specht modules are irreducible.

It is known that for every finite group G, the number of isomorphism classes of irreducible complex G-modules is the number of conjugacy classes in G. The proof is thus finished by observing that the number of Specht modules S^λ is $\#\{\lambda \mid \lambda \vdash n\}$, that is, the number of conjugacy classes in S_n. □

Before we can compute the character of a Specht module, we need two lemmas.

Lemma 23. *Let $\theta\colon M^\lambda \to M^{\widehat{\lambda}}$ be a map of S_n-modules. If S^λ is not contained in the kernel of θ, then $\widehat{\lambda} \leq \lambda$.*

Proof. Let T be a numbering of λ. Since v_T is not in the kernel of θ, we have $0 \neq \theta(v_T) = \theta(b_T\{T\}) = b_T\, \theta\{T\}$. Therefore $b_T\{\widehat{T}\} \neq 0$ for some numbering $\widehat{T}$ of $\widehat{\lambda}$. If $\lambda \neq \widehat{\lambda}$ and λ does not dominate $\widehat{\lambda}$ then we are in case (1) of lemma 8. This contradicts lemma 14. □

Lemma 24. *There are nonnegative integers $k_{\nu\lambda}$ for $\nu > \lambda$ such that*

$$M^\lambda \cong S^\lambda \oplus \Big(\bigoplus_{\nu > \lambda} k_{\nu\lambda}\, S^\nu \Big).$$

Proof. For each ν, let $k_{\nu\lambda}$ be the number of times the irreducible representation S^ν occurs in the decomposition of M^λ. To see that $k_{\nu\nu} = 1$, take any numbering T of λ and use equation (21). Since every S^ν occurs in M^ν, there is a projection from M^ν to S^ν. Suppose S^ν also occurs in the decomposition of M^λ. Then the projection from M^ν to S^ν followed by an imbedding of S^ν in M^λ is a homomorphism θ from M^ν to M^λ that does not contain S^ν in its kernel. Lemma 23 implies that $\lambda \leq \nu$, concluding the proof. □

Proposition 25. *We have* $\operatorname{ch} S^\lambda = s_\lambda$*. Equivalently, the character of S^λ is χ^λ.*

Proof. By lemma 24 we have

$$\operatorname{ch} M^\lambda = \operatorname{ch} S^\lambda + \sum_{\nu > \lambda} k_{\nu\lambda} \operatorname{ch} S^\nu.$$

By proposition 7.10.5 we have

$$h_\lambda = s_\lambda + \sum_{\nu > \lambda} K_{\nu\lambda}\, s_\nu.$$

Since $\operatorname{ch} M^\lambda = h_\lambda$, it follows that

$$s_\lambda = \operatorname{ch} S^\lambda + \sum_{\nu > \lambda} m_{\nu\lambda} \operatorname{ch} S^\nu$$

for some integers $m_{\nu\lambda}$. But the ch S^μ are an orthonormal basis because ch is an isometry, so that

$$1 = \langle s_\lambda, s_\lambda \rangle = 1 + \sum (m_{\nu\lambda})^2.$$

The coefficients $m_{\nu\lambda}$ must therefore all vanish, giving the desired equation ch $S^\lambda = s_\lambda$. □