Lecture 18: Representations of symmetric groups, III.

1) Proof of character formula.

Ref: [E], Secs 5.14 & 5.15.

1) Proof of character formula.

Recall the notation. $\mathbb{F}$ is an algebraically closed field of char $=0$. To a partition λ, we assign the irreducible representation V_λ of S_n. Let $N \geqslant n$. Consider symmetric polynomials $p_m = \sum_{i=1}^N x_i^m$ $(m>0)$ and, for $\sigma \in S_n$, $p_\sigma = p_{m_1} \dots p_{m_q}$, where $(m_1, \dots m_q)$ is the cycle type of σ. Finally, consider the Vandermonde determinant $\Delta = \det(x_i^{N-j})_{i,j=1}^N = \prod_{i<j}(x_i - x_j)$

Theorem (Frobenius) $\chi_{V_\lambda}(\sigma)$ coincides w. the coefficient of $\prod_{i=1}^N x_i^{\lambda_i + N - i}$ in Δp_σ.

1.1) Formula for $\chi_{I_\lambda^+}$

The first step for proving the theorem is to get a similar in spirit formula for $\chi_{I_\lambda^+}$, where, recall, $I_\lambda^+ = \operatorname{Ind}_{S_\lambda}^{S_n} \operatorname{triv}$.

Proposition: $\chi_{I^+_\lambda}(\sigma)$ is the coefficient of $\prod_{i=1}^{N} x_i^{\lambda_i}$ in p_σ ($\sigma\in S_n$).

Proof:

We'll give a combinatorial interpretation of $\chi_{I^+_\lambda}(\sigma)$.

Recall that $I^+_\lambda=\mathrm{Fun}(S_n/S_\lambda,\mathbb{F})$, so, by Sec 2.1 of Lec 8, $\chi_{I^+_\lambda}(\sigma)=|(S_n/S_\lambda)^\sigma|$, the # of σ-fixed points in S_n/S_λ. A point of S_n/S_λ can be thought of an ordered collection of subsets $X_i\subset\{1,2,..n\}$, $i=1,\dots,k:=\lambda_1^t$, w. $|X_i|=\lambda_i$ & $\{1,2,...n\}=\bigsqcup_i X_i$: the group S_n acts by permuting the elements of $\{1,2..n\}$ (the action is transitive & S_λ is the stabilizer of the collection $X_i=\{\lambda_1+..+\lambda_{i-1}+m\,|\,m=1,2,...\lambda_i\}$ that appeared in Sec 1.1 of the previous lecture, the proof is left as an exercise).

$(X_1,\dots X_k)$ is fixed by $\sigma \Leftrightarrow \sigma(X_i)=X_i\ \forall i$. Let $\langle\sigma\rangle\subset S_n$ be the subgroup generated by σ & $Z_1,\dots Z_q$ be the $\langle\sigma\rangle$-orbits in $\{1,2,...n\}$, $Z_\ell:=\{$numbers in the ℓth cycle of $\sigma\}$, so $|Z_\ell|=m$. Of course, $\sigma(X_i)=X_i \Leftrightarrow X_i$ is the union of orbits. Therefore, the # of fixed points = # of splittings of $(m_1,\dots m_q)$ into N groups w. sums $\lambda_1,\dots\lambda_N$. This coincides w. the coefficient of $x_1^{\lambda_1}\dots x_N^{\lambda_N}$ in $\prod_{\ell=1}^{q}\sum_{i=1}^{N} x_i^{m_\ell}$, finishing the proof. $\square$

1.2) Reduction to combinatorial statement.

We now proceed to proving the theorem. In this section we reduce the proof to "Main Claim", which will be proved in the next section, entirely based on arguments that do not involve representations (manipulations w. formal power series, mostly).

First, some notation. For $\alpha = (a_1, \ldots, a_N) \in \mathbb{Z}^N$, set $x^\alpha = \prod_{j=1}^{N} x_i^{a_i}$.

Set $\rho := (N-1, N-2, \ldots, 0)$ so that

$$\Delta = \sum_{\tau \in S_n} \operatorname{sgn}(\tau)\, x^{\tau\rho}$$ permutation

For a partition λ of n & $\sigma \in S_n$, we write $\theta_\lambda(\sigma)$ for the coefficient of $x^{\lambda+\rho}$ in Δp_σ. We need to show that

$$(1) \qquad \theta_\lambda(\sigma) = \chi_{V_\lambda}(\sigma)$$

Note that by the very definition, $p_\sigma = p_{\sigma'}$ if σ, σ' have the same cycle type $\Leftrightarrow$ conjugate. So $\theta_\lambda : S_n \to \mathbb{Z}$ is a class function.

Recall that on $Cl(S_n)$ we have the symmetric bilinear form

$$(f_1, f_2) = \frac{1}{|S_n|} \sum_{\sigma \in S_n} f_1(\sigma) f_2(\sigma^{-1}) = [\sigma \,\&\, \sigma^{-1} \text{ have the same cycle type, hence conjugate}] = \frac{1}{|S_n|} \sum_{\sigma \in S_n} f_1(\sigma) f_2(\sigma).$$

The orthogonality of characters (Lec 9) tells us that

$$(\chi_{V_\lambda}, \chi_{V_\lambda}) = 1.$$

Main Claim: For all partitions λ of n, we have $(\theta_\lambda, \theta_\lambda) = 1$.

We'll prove Main Claim in the next section (w/o any representation theory).

Proof of Theorem modulo Main Claim:

The proof goes as follows:

Step 1: Check $\theta_\lambda = \chi_{I^+_\lambda} + \sum_{\mu^t < \lambda^t} a_{\mu\lambda} \chi_{I^+_\mu}$ w. $a_{\mu\lambda} \in \mathbb{Z}$.

Step 2: Check $\chi_{I^+_\lambda} = \chi_{V_\lambda} + \sum_{\mu^t < \lambda^t} K_{\mu\lambda} \chi_{V_\mu}$ w. $K_{\mu\lambda} \in \mathbb{Z}_{\geq 0}$, these are called "Kostka numbers."

Step 3: Combine Steps 1,2 w. Main Claim & orthonormality of characters of irreducibles to finish the proof.

Now the details:

Step 1: For $\alpha=(a_1,\dots,a_N)\in\mathbb{Z}^N$, let $c_\alpha(\sigma)$ be the coefficient of x^α in p_σ: $p_\sigma=\sum_\alpha c_\alpha(\sigma)x^\alpha$. Note that:

- $c_\alpha(\sigma)=0$ unless $\alpha\in\mathbb{Z}^N_{\geq 0}$ & $\sum a_j=n$
- since p_σ is symmetric, $c_\alpha(\sigma)=c_{\tau\alpha}(\sigma)$ $\forall\, \tau\in S_N$.
- if $a_1\geq a_2\geq\dots\geq a_N$ (so that α is a partition of n), we have $c_\alpha(\sigma)=\chi_{I^+_\lambda}(\sigma)$, this is Proposition in Sec 1.1.

Then $\Delta p_\sigma=\left(\sum_{\tau\in S_N}\operatorname{sgn}(\tau)x^{\tau\rho}\right)\left(\sum_{\alpha\in\mathbb{Z}^N}c_\alpha(\sigma)x^\alpha\right)=$
$=\sum_{\tau,\alpha}\operatorname{sgn}(\tau)c_\alpha(\sigma)x^{\alpha+\tau\rho}$. The coefficient of $x^{\lambda+\rho}$ is

(2) $$\theta_\lambda=\sum_{\tau\in S_N}\operatorname{sgn}(\tau)\,c_{\lambda+\rho-\tau\rho}(\sigma).$$

Now we need to deduce

(3) $$\theta_\lambda=c_\lambda+\sum_{\mu^t<\lambda^t}a_{\mu\lambda}c_\mu$$

Note that here μ is a partition: $\mu_1\geq\mu_2\geq\dots\geq\mu_N$. For $\alpha\in\mathbb{Z}^N$, we write α_+ for the unique decreasing permutation of α. We need to show that for $\mu=(\lambda+\rho-\tau\rho)_+$ w. $\tau\neq e$ we have $\mu^t<\lambda^t$ (assuming $\mu\in\mathbb{Z}^N_{\geq 0}$).

For this, it's convenient to introduce a <u>partial</u> order on the set of partitions of n, often called the dominance order. For partitions λ,μ of n, we set $\lambda\leq\mu$ if

$$\sum_{i=1}^{k}\lambda_i \leq \sum_{i=1}^{k}\mu_i \quad \forall\, k=1,\dots N.$$

Two remarks are in order:

- $\lambda \prec \mu \Rightarrow \lambda < \mu$.
- $\lambda \prec \mu \Leftrightarrow$ as a diagram μ is obtained from λ by moving some boxes down (and to the right) $\Leftrightarrow \mu^t \prec \lambda^t$

Return to $\mu = (\lambda+\rho-\tau\rho)_+$. Note that $\mu \succeq \lambda$:

$$\sum_{i=1}^{k}\mu_i \geq \sum_{i=1}^{k}\left(\lambda_i + N-i-N+\tau^{-1}(i)\right) \geq \sum_{i=1}^{k}\lambda_i$$

w 2nd $\geq$ being $>$ for $\tau\neq e$ b/c $\sum_{i=1}^{k}\tau^{-1}(i) \geq \sum_{i=1}^{k} i \ \forall k$ w. $= \Leftrightarrow \tau=e$.

So $\mu^t \prec \lambda^t$ if $\tau\neq e \Rightarrow \mu^t < \lambda^t$. This finishes Step 1.

Step 2: We have $I^+_\lambda = \bigoplus_\mu V_\mu^{\oplus K_{\mu\lambda}}$ for some $K_{\mu\lambda}\in\mathbb{Z}_{\geq 0}$. Recall that V_μ occurs in I^-_μ $\forall \mu$ &

(4) $$\dim \operatorname{Hom}_{S_n}(I^+_\lambda, I^-_\mu) = \begin{cases} 1 & \text{if } \lambda=\mu \\ 0 & \text{if } \mu^t > \lambda^t \end{cases}$$

(4) implies $K_{\mu\lambda}\neq 0 \Rightarrow \mu^t \leq \lambda^t$. Moreover, $K_{\lambda\lambda}>0$ by the construction of V_λ & ≤ 1 by (4). So $I^+_\lambda = V_\lambda \oplus \bigoplus_{\mu^t<\lambda^t} V_\mu^{\oplus K_{\mu\lambda}} \Rightarrow$

(5) $$\iota_\lambda = \chi_{V_\lambda} + \sum_{\mu^t<\lambda^t} K_{\mu\lambda}\chi_{V_\mu}.$$

Step 3: From (3) & (5) we deduce

(6) $\theta_\lambda = \chi_{V_\lambda} + \sum_{\mu^t < \lambda^t} b_{\mu\lambda} \chi_{V_\mu} \quad (b_{\lambda\mu} \in \mathbb{Z})$

According to Main Claim, $(\theta_\lambda, \theta_\lambda) = 1 \Rightarrow [(6) + (\chi_{V_{\lambda'}}, \chi_{V_{\lambda''}})$ $= \delta_{\lambda', \lambda''}] = 1 + \sum_{\mu^t < \lambda^t} b_{\mu\lambda}^2 \Rightarrow \theta_\lambda = \chi_{V_\lambda}$ $\square$

Remark: Here is a combinatorial interpretation of $K_{\mu\lambda}$. By a (semistandard) Young tableau of shape μ and weight λ we mean a filling of the Young diagram μ w. λ_1 1's, λ_2 2's,... so that the numbers weakly increase left to right and strictly increase bottom to top. E.g. for $\mu = (3,1)$ & $\lambda = (2,1,1)$, have $K_{\mu\lambda} = 2$:

$$\begin{array}{|c|cc} \hline 3 & \multicolumn{2}{c}{} \\ \hline 1 & \multicolumn{1}{|c|}{1} & \multicolumn{1}{c|}{2} \\ \hline \end{array} \qquad \begin{array}{|c|cc} \hline 2 & \multicolumn{2}{c}{} \\ \hline 1 & \multicolumn{1}{|c|}{1} & \multicolumn{1}{c|}{3} \\ \hline \end{array}$$

1.3) Proof of Main Claim.

$\theta_\lambda(\sigma)$ is defined as the coefficient of a monomial in some polynomial. We want to give a similar interpretation of $(\theta_\lambda, \theta_\lambda)$. For this we need two collections of variables: $x_1, \dots x_N$ & $y_1, \dots y_N$.

Lemma 1: $(\theta_\lambda, \theta_\lambda)$ coincides w. the coefficient of $x^{\lambda+\rho}y^{\lambda+\rho}$ in the formal power series expansion of

$$\Delta(x)\Delta(y)\prod_{i,j=1}^{N}(1-x_iy_j)^{-1}$$

where $\Delta(x), \Delta(y)$ are the Vandermondes in $x_1, \dots x_N$ & $y_1, \dots y_N$.

Proof:

Since $\theta_\lambda(\sigma)$ is the coefficient of $x^{\lambda+\rho}$ in $\Delta(x)p_\sigma(x)$, then $\theta_\lambda(\sigma)^2$ is the coefficient of $x^{\lambda+\rho}y^{\lambda+\rho}$ in $\Delta(x)\Delta(y)p_\sigma(x)p_\sigma(y)$. To get to $(\theta_\lambda, \theta_\lambda) = \frac{1}{n!}\sum_{\sigma\in S_n}\theta_\lambda(\sigma)^2$ we need to rewrite the r.h.s. appropriately. We encode a conjugacy class in S_n as a sequence $\underline{i} = (i_m)_{m\geq 1}$ w. $\sum_{m=1}^{\infty} m i_m = n$ (this equality means, in particular, that only finitely many of i_m's are nonzero): to this collection we assign the class w. cycle type consisting of i_m cycles of length m, $\forall m$. We write $\theta_\lambda(\underline{i})$ for $\theta_\lambda(\sigma)$ w. σ in the corresponding conjugacy class and $z(\underline{i})$ for the order of $Z_{S_n}(\sigma)$ so that the number of elements in the conjugacy class is $\frac{n!}{z(\underline{i})}$. So

$$(\theta_\lambda, \theta_\lambda) = \frac{1}{n!}\sum_{\sigma\in S_n}\theta_\lambda(\sigma)^2 = \sum_{\underline{i}\,|\,\sum m i_m = n}\frac{\theta_\lambda(\underline{i})^2}{z(\underline{i})}$$

Exercise: $z(\underline{i}) = \prod_{m \geq 1} i_m! \, m^{i_m}$ (all factors but finitely many are 1).

What we've got so far is that $(\theta_\lambda, \theta_\lambda)$ is the coefficient of $x^{\lambda+\rho} y^{\lambda+\rho}$ in $\Delta(x)\Delta(y) \sum_{\underline{i}} \dfrac{p_{\underline{i}}(x) p_{\underline{i}}(y)}{\prod_{m \geq 1} i_m! \, m^{i_m}}$. Here the sum is taken over all $\underline{i}$ s.t. $\sum m i_m = n$ (and $p_{\underline{i}}(x) = \prod_{\ell=0}^{\infty} (\sum x_j^\ell)^{i_\ell}$, it's equal to $p_\sigma(x)$ for σ in the conjugacy class corresponding to $\underline{i}$). The key observation is that $x^{\lambda+\rho} y^{\lambda+\rho}$ can only appear in $p_{\underline{i}}(x) p_{\underline{i}}(y)$ if $\sum m i_m = n$ — for degree reasons. So we can sum over all $\underline{i}$.

So $(\theta_\lambda, \theta_\lambda)$ is the coefficient of $x^{\lambda+\rho} y^{\lambda+\rho}$ in $\Delta(x)\Delta(y) \cdot (*)$, where

$$(*) = \sum_{\underline{i}} \prod_{\ell=0}^{\infty} \frac{(\sum x_j^\ell)^{i_\ell} (\sum y_j^\ell)^{i_\ell}}{\ell^{i} \, i_\ell!} = \sum_{\underline{i}} \prod_{\ell=0}^{\infty} \Big(\sum_{j,k=1}^{N} x_j^\ell y_k^\ell / \ell\Big)^{i_\ell} / i_\ell! =$$

$$= \prod_{\ell=0}^{\infty} \sum_{i_\ell=0}^{\infty} \Big(\sum_{j,k=1}^{N} x_j^\ell y_k^\ell / \ell\Big)^{i_\ell} / i_\ell! = \prod_{\ell=0}^{\infty} \exp\Big(\sum_{j,k=1}^{N} (x_j y_k)^\ell / \ell\Big) =$$

$$= \exp\Big(\sum_{j,k=1}^{N} \sum_{\ell=0}^{\infty} (x_j y_k)^\ell / \ell\Big) = \prod_{j,k=1}^{N} \exp(-\log(1 - x_j y_k)) = \prod_{j,k=1}^{N} (1 - x_j y_k)^{-1}.$$

This finishes the proof. $\square$

Lemma 2 (Cauchy's determinantal identity)

$$(7) \qquad \Delta(x)\Delta(y)\prod_{j,k=1}^{N}(1-x_j y_k)^{-1} = \det\left(\frac{1}{1-x_j y_k}\right)_{j,k=1}^{N}$$

Proof: Set $z_j = x_j^{-1}$, $\varepsilon = (-1)^{N(N-1)/2}$. (7) is equivalent to

$$\frac{\varepsilon\Delta(z)\Delta(y)}{\prod_{j,k=1}^{N}(z_j-y_k)} = \det\left(\frac{1}{z_j-y_k}\right) \Leftrightarrow \varepsilon\Delta(z)\Delta(y) = \det\left(\frac{1}{z_j-y_k}\right)\prod_{j,k=1}^{N}(z_j-y_k)$$

Both sides are polynomials in z_j, y_k of deg N^2-N. Both vanish when $z_j = z_{j'}$ for $j \neq j'$ or when $y_k = y_{k'}$ for $k \neq k'$. So, the polynomials are proportional.

We need to show the coefficient of proportionality is 1. In order to do this, set $y_k = z_k$ $\forall$ $k = 1, \dots, N$. In the l.h.s. we get $\varepsilon\Delta(y)^2$. In the r.h.s. we get $\prod_{j\neq k}(y_j - y_k)$. The two are equal. $\square$

Proof of Main Claim:

Combining Lemmas 1 & 2 we see that $(\theta_\lambda, \theta_\lambda)$ is the coefficient of $x^{\lambda+\rho} y^{\lambda+\rho}$ in the power series expansion of

$$\det\left(\frac{1}{1-x_j y_k}\right)_{j,k=1}^{N} = \sum_{\tau\in S_N} \frac{\mathrm{sgn}(\tau)}{\prod_j (1-x_j y_{\tau(j)})} = \sum_{\tau\in S_N} \mathrm{sgn}(\tau) \prod_{j=1}^{N} \sum_{\ell=0}^{\infty} x_j^{\ell} y_{\tau(j)}^{\ell}$$

For $\tau\neq e$, the coefficient of $x^{\lambda+\rho} y^{\lambda+\rho}$ in $\prod_{j=1} \sum_{\ell=0}^{\infty} x_j^{\ell} y_{\tau(j)}^{\ell}$ is zero b/c $\lambda+\rho$ is strictly decreasing & the monomials in this formal power series are of the form $x^{\alpha} y^{\tau\alpha}$ for some α. And the coefficient of $x^{\lambda+\rho} y^{\lambda+\rho}$ in $\prod_{j=1}^{N} \sum_{\ell=0}^{\infty} (x_j y_j)^{\ell}$ is 1. $\square$

Remark: This finishes our study of group representations - with exception of bonus lectures, where will discuss more things around representations of symmetric groups. One remark is in order. We've seen that in the study of representations of S_n induction plays an important role. Overall, induction is a quite powerful tool. It can be used to understand representations of semi-direct products of the form $K \ltimes A$, where K & A are finite groups & A is abelian. See Sec 5.27 in [E]. Induction can also be used to classify the irreducible representations of groups like $GL_n(\mathbb{F}_q)$. See Sec 5.25 in [E] for the $n=2$ case.