

Week 7

Representations of the Symmetric Group

7.1 Specht Modules

11/6:

- Announcements.
 - Midterm description is on the Canvas page.
 - Review what he says to review, and then look at the PSets. The operator averaging stuff and S_4 , S_5 examples are most important.
 - New HW will be due next Friday (not this Friday).
- New topic: Representations of S_n .
 - We will talk about these almost until the end of the course.
 - Very hard.
 - Any specialist in rep theory will still say that they know some approaches, but nobody understands this stuff completely.
 - We'll explore some phenomena, but if we feel after this course that we still don't understand everything about S_n , that's typical; if we think we understand everything, we're probably wrong.
- Representation theory of $GL_n(\mathbb{F}_{p^k})$ is related but even worse.
 - Same with $O_n(\mathbb{F}_{p^k})$.
 - Recently, all this stuff was understood with something called linguistic (??) theory, but that's far beyond us.
- $|S_n| = n!$, and the conjugacy classes are in bijection with cyclic structures of a permutation.
 - Our good understanding of the conjugacy classes of S_n is the only thing that makes this problem the slightest bit tractable.
 - Cyclic structures are also in bijection with the **partitions** of a number; recall that we briefly talked about these in MATH 25700!
- **Partition** (of $n \in \mathbb{N}$): An ordered tuple satisfying the following constraints. *Denoted by $\lambda, (\lambda_1, \dots, \lambda_k)$.*
Constraints
 1. $\lambda_i \in \mathbb{N}$ for $i = 1, \dots, k$;
 2. $\lambda_1 \geq \dots \geq \lambda_k$;
 3. $\lambda_1 + \dots + \lambda_k = n$.

- Example: The partitions of the number “4” are (4), (3, 1), (2, 2), (2, 1, 1), and (1, 1, 1, 1).
 - This is the same way we’ve been denoting representations!
- $p(n)$: The number of possible partitions of n .
 - Hardy and Ramanujan helped understand the number $p(n)$ of partitions of n , but they’re still very hard to understand.
- One way to understand $p(n)$ is through its encoding in the **generating function**

$$\sum_{n \geq 1} p(n)x^n = 1 + x + 2x^2 + 3x^3 + 5x^4 + \dots$$

- We can think of the above generating function as an actual function of x if it converges for small x ; if it doesn’t converge, then we just think of it as a “meaningless” **formal power series**.
- To choose a partition, we need to choose a certain number of 1’s, a certain number of 2’s, a certain number of 3’s, etc. all the way up to n .
- So let’s look at

$$(1 + x + x^2 + \dots)(1 + x^2 + x^4 + \dots)(1 + x^3 + x^6 + \dots)(1 + x^4 + x^8 + \dots) \dots$$

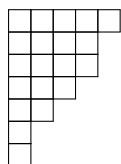
- Formally, this is

$$\prod_{i=1}^{\infty} \left(\sum_{j=0}^{\infty} x^{ij} \right)$$

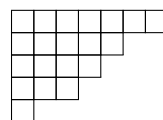
- This equals the generating function! It tells us that to compute $p(100)x^{100}$, we need only look at certain terms.
- Recall that we can write $1 + x + x^2 + \dots = 1/(1 - x)$. Doing similarly for other terms transforms the above product into

$$\frac{1}{1-x} \frac{1}{1-x^2} \frac{1}{1-x^3} \dots$$

- **Generating function**: An encoding of an infinite sequence of numbers as the coefficients of a formal power series.
- **Formal power series**: An infinite sum of terms of the form ax^n that is considered independently from any notion of convergence.
- The above discussion of $p(n)$ as a generating function is only for our fun; Rudenko is not going to use this in any way.
 - It’s just a pretty function.
 - Takeaway: We can write the generating function as something nice and then use it to prove something.
- We can visualize these partitions using something called a **Young diagram**.



(a) $\lambda = (5, 4, 4, 3, 2, 1, 1)$.



(b) $\lambda' = (7, 5, 4, 3, 1)$.

Figure 7.1: Young diagrams for a partition of 20.

- Suppose we have the following partition of 20: $(5, 4, 4, 3, 2, 1, 1)$.
- Then we draw 5 cages for 5 little birds, followed by 4 cages for 4 little birds, etc.
- Thus, the i^{th} row of boxes has length λ_i .
- The same way you can denote by λ the whole *partition*, you can denote by λ the whole *diagram*.
- This is just a way to visualize partitions.
- Recall the three partitions of S_3 , corresponding to its representations: (3) , $(2, 1)$, $(1, 1, 1)$.
- Moreover, these diagrams are actually meaningful!

- **Inverse** (of λ): The partition $(\lambda'_1, \dots, \lambda'_k)$ defined as follows. *Denoted by λ' . Given by*

$$\lambda'_i = |\{\lambda_j \mid \lambda_j \geq i\}|$$

for all $i = 1, \dots, k$.

- We can see that $\lambda'_1 \geq \dots \geq \lambda'_k$.
- We can also see that the sum will still be n .
- Moreover, if we do this twice, we'll get back to λ , i.e., $(\lambda')' = \lambda$.
- We can prove $(\lambda')' = \lambda$ combinatorially, too, (that is, without Young diagrams) but that gets pretty complicated.
- Example: If $\lambda = (5, 4, 4, 3, 2, 1, 1)$ as above, then $\lambda' = (7, 5, 4, 3, 1)$.
 - See Figure 7.1b.
 - Moreover, the Young diagrams are related by a flip akin to matrix transposition.
 - Notice how the definition of inversion *exactly* specifies this flip in the picture: The number of λ_j 's that have length at least 1 is all the first column of Figure 7.1a, the number of length at least 2 is all the second column, etc.
- Onto the next question, which is the main miracle.
 - Main miracle: There exists a natural (i.e., canonical) bijection between the conjugacy classes and irreducible representations of S_n .
 - We've explored a duality for general finite groups G , before, but never a bijection.
 - In S_n , there *is* this natural bijection.
 - If you understand why intuitively, you will have started to understand the representation theory of S_n .
- If we define $\lambda \mapsto n$ (??), then there is some irrep V_λ corresponding to λ . We will look at the **Specht module** construction of V_λ .
 - Some of the proofs Rudenko will present, he stole from Etingof et al. (2011), and some of the proofs he invented himself.
 - This is *by far* the best construction, even though it's exceedingly rare in the literature.
- The usual construction.
 - Take $\mathbb{C}[S_n]$ with coefficients a_λ, b_λ , etc. similar over conjugacy classes and do something with it??
 - “Just say NO!” to this construction.
- Here is the better idea.
 - Consider an algebra of polynomials with rational coefficients: $\mathbb{Q}[x_1, \dots, x_n]$.
 - We could also do real or complex, but rational is nice.

- For symmetric groups, all representations will be integers, etc.??
- One thing to emphasize about this algebra: It is a **graded** algebra.
 - If represented by A , then it equals $A_0 \oplus A_1 \oplus A_2 \oplus \cdots$ where

$$A_m = \left\{ \sum_{k_1 + \cdots + k_n = m} a_{k_1 \dots k_n} x_1^{k_1} \cdots x_n^{k_n} \right\}$$

- I.e., A_m is the sum of all polynomials with degree equal to m .
 - Example: If we take $1 + x_1^2 x_2^3 + x_1 x_2 + x_1^{100} + x_1 x_2^{99}$, we can then break this polynomial up into polynomials of degree 1, 5, 2, and 100.
- We also have $A_{m_1} \cdot A_{m_2} \subset A_{m_1+m_2}$.
 - Example: $x_1 x_2^2 \cdot (x_1 + x_2) = x_1^2 x_2^2 + x_1 x_2^3$.
- With this algebra in hand, we may let $S_n \curvearrowright \mathbb{Q}[x_1, \dots, x_n]$ via

$$\sigma P(x_1, \dots, x_n) = P(x_{\sigma^{-1}(1)}, \dots, x_{\sigma^{-1}(n)})$$

- In other words, σ is transposing polynomials.
 - Example: $(12)(x_1^2 + x_2^3 + x_3) = x_2^2 + x_1^3 + x_3$.
- Thus, we call as A_1 the representation V_{perm}^* .
 - This is because $A_1 = \text{span}(x_1 + \cdots + x_n)$, and permuting these is much like permuting the basis of a vector space, as the typical permutation representation does.
 - It could technically be the isomorphic representation V_{perm} , but the dual fits better here for reasons??
- Then $A_2 = S^2 V^*$.
 - So if A_1 had basis e^1, \dots, e^m , A_2 has basis $e^i e^j$.
 - Why are we choosing these sets??
- Continuing, $A_3 = S^3 V^*$.
- It follows that the representation of the overall thing is

$$\bigoplus_{m \geq 0} (S^m V_{\text{perm}}^*)$$

- This is called the **symmetric algebra**.
- **Graded** (algebra): An algebra for which the underlying additive group is a direct sum of abelian groups A_i such that $A_i A_j \subset A_{i+j}$.
- So how do we construct representations?
 - For S_2 , $x_1 - x_2$ changes sign when we apply S_2 .
 - For $S_3 \dots$
 - The trivial's polynomial is 1 and *tableaux*.
 - The standard is $(2, 1)$. When we apply S_3 to $(x_1 - x_2)$, we get

$$\langle (x_1 - x_2), (x_2 - x_1), (x_1 - x_3), (x_3 - x_1), (x_2 - x_3), (x_3 - x_2) \rangle$$
 - If we let $a = x_1 - x_2$, $b = x_2 - x_3$, then some elements equal $a + b$. This is another way to think about the action.
 - What about the alternating representation? We have $(x_1 - x_2)(x_2 - x_3)(x_1 - x_3) = \Delta_{123}$, which changes sign when we apply any element with sign -1 !
 - For $S_4 \dots$

- (4) is 1.
- (3, 1) is $S_4(x_1 - x_2) = \Delta_{12}$.
- (1, 1, 1, 1) is $(x_1 - x_2)(x_1 - x_3)(x_1 - x_4)(x_2 - x_3)(x_2 - x_4)(x_3 - x_4) = \Delta_{1234}$.
- (2, 1, 1) is $(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)$.
 - We got this polynomial by guessing; the same way $(x_1 - x_2)$ worked in multiple cases, maybe this one does too! And it does.
 - Something to check is that $\Delta_{123} - \Delta_{124} - \Delta_{134} - \Delta_{234} = 0$.
- (2, 2) is $(x_1 - x_2)(x_3 - x_4)$.
 - Something related we can prove is that

$$(x_1 - x_2)(x_3 - x_4) - (x_1 - x_3)(x_2 - x_4) - (x_1 - x_4)(x_2 - x_3) = 0$$
 - This formula appears in **cross ratios**, which we can discuss in Rudenko's algebraic geometry course next quarter.
- For $\lambda = (4, 3, 1)$, we have $\Delta_{123}\Delta_{45}\Delta_{67}$, and we act by S_8 upon this! Explicitly, we have $S_8(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)(x - 4 - x_5)(x_6 - x_7)$.
- Takeaway: It all depends on column length!
- These polynomials are called **Vandermonde determinants**; those are the little Δ things with subscripts. We'll talk about these next times.
- We need to prove reducibility and not pairwise isomorphic to make sure that this construction is valid, but that's easy!

7.2 Vandermonde Determinants

11/8:

- Announcements.
 - OH tonight at 6:00 PM.
- Consider S_n .
 - Recall the symmetric algebra $R = \mathbb{Q}[x_1, \dots, x_n]$, which is a graded ring $\bigoplus_{d \geq 0} R_d$ where $R_d = S^d V_{\text{perm}}^*$.
 - The action of $S_n \curvearrowright \mathbb{Q}[x_1, \dots, x_n]$ is $\sigma P(x_1, \dots, x_n) = P(x_{\sigma^{-1}(1)}, \dots, x_{\sigma^{-1}(n)})$.
- With the definitions of last class behind us, we can now look at the **space of invariants** R^{S_n} , isotypical components of which σ acts on trivially.

$$R^{S_n} := \{P(x_1, \dots, x_n) \mid \forall \sigma \in S_n, P(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = P(x_1, \dots, x_n)\}$$
 - This is the ring of symmetric polynomials.
 - Example: If $n = 3$, then $x_1^3 + x_2^3 + x_3^3 - 3x_1x_2x_3 \in \mathbb{Z}^{S_3}$.
- We now define some stuff to help us prove a major result: The n elementary symmetric polynomials.
 - $\sigma_1 = x_1 + \dots + x_n = \sum_{1 \leq i \leq n} x_i$.
 - $\sigma_2 = x_1x_2 + \dots + x_1x_n + x_2x_3 + \dots = \sum_{1 \leq i < j \leq n} x_ix_j$.
 - $\sigma_k = \sum_{1 \leq i_1 < \dots < i_k \leq n} x_{i_1} \dots x_{i_k}$.
 - $\sigma_n = x_1 \dots x_n$.
- With these definitions, we can say that

$$x^n + a_{n-1}x^{n-1} + \dots + a_0 = (x - x_1) \dots (x - x_n)$$

where

$$a_{n-1} = -\sigma_1 \qquad a_{n-2} = \sigma_2 \qquad \dots \qquad a_0 = (-1)^n \sigma_n$$

- Fundamental theorem.
 - Basic statement: Every polynomial is a polynomial in these polynomials.
 - A more precise statement follows.

- Theorem (Fundamental theorem of symmetric polynomials): We have that

$$\mathbb{Q}[x_1, \dots, x_n]^{S_n} = \mathbb{Q}[\sigma_1, \dots, \sigma_n]$$

- Before we prove the fundamental theorem, there are a few points we need to discuss.

- Example:

- Take $x^2 + px + q = 0$.
- If it has two roots x_1, x_2 , then $\sigma_1 = x_1 + x_2 = -p$ and $\sigma_2 = x_1x_2 = q$.
- Then $x_1^2 + x_2^2 = \sigma_1^2 - 2\sigma_2 = p^2 - 2q$.
- Thus, if we take $(x_1 - x_2)^2 = (x_1 + x_2)^2 - 4x_1x_2 = \sigma_1^2 - 4\sigma_2 = p^2 - 4q$, which is the discriminant.
- In general, $x_1^n + x_2^n$ has an expression as a polynomial in σ_1, σ_2 . This will be a homework problem.
- What is going on here?? Is $x^2 + px + q$ even in $\mathbb{Q}[x]^{S_n}$? If so, why do we factor it into σ_1, σ_2 instead of just σ ? What are the other examples about?

- **Lexicographic order** (on monomials): An ordering of monomials based on the following rule. *Denoted by \succ . Given by*

$$x_1^{a_1} \cdots x_n^{a_n} \succ x_1^{b_1} \cdots x_n^{b_n} \dots$$

1. If $a_1 > b_1$ OR...
2. If $a_1 = b_1$ and $a_2 > b_2$ OR...
3. $a_1 = b_1$ and $a_2 = b_2$ and $a_3 > b_3$ OR...
4. So on and so forth.

- Notes on the lexicographic ordering.

- Don't think of this order like an ordering on integers.
- This allows us to define the key notion for a number of proofs we'll see in the coming days.
- Although it may seem counterintuitive, the lexicographic ordering is still determined for polynomials such as σ_1 . For example, we may look at

$$\sigma_1 = x_1 + \cdots + x_n$$

and think, "Wait a second — all these terms have the same order: They all have the same exponent of 1." However, we would be discounting the fact that the lexicographic ordering views σ_1 as

$$\sigma_1 = x_1^1 x_2^0 \cdots x_n^0 + \cdots + x_1^0 \cdots x_{n-1}^0 x_n^1$$

From here, we can see that $LM(\sigma_1) = x_1^1 x_2^0 \cdots x_n^0 = x_1$.

- **Largest monomial** (of $P \neq 0$): The monomial in $P(x_1, \dots, x_n) \neq 0$ that is the largest lexicographically. *Denoted by $LM(P)$.*

- $C_{LM}(P)$: The coefficient of $LM(P)$.

- Example: Consider the polynomial $P = x_1^2 + x_1^3 x_2 x_3 - 7x_1^3 x_2 x_3^{100}$.

- Then $LM(P) = x_1^3 x_2 x_3^{100}$ and $C_{LM}(P) = -7$.

- Properties.

1. $P, Q \neq 0$ implies that $LM(PQ) = LM(P)LM(Q)$.

- Using inductive reasoning, try multiplying the example above by $Q = x_1^2 + x_2^2 + x_3^2$!
- Rudenko will not give rigorous proofs of any of these properties; they will just confuse us. It's better to do everything intuitively here.

- Lemma: If $P \in \mathbb{Q}[x_1, \dots, x_n]^{S_n}$ and $LM(P) = x_1^{a_1} \cdots x_n^{a_n}$, then $a_1 \geq \cdots \geq a_n$.

Proof. Let $i < j$. Suppose for the sake of contradiction that $a_j > a_i$. Let $\sigma = (ij) \in S_n$. Since P is symmetric, $\sigma P = P$. But then in particular, the monomial $\sigma LM(P)$ in P is lexicographically larger than $LM(P)$. Thus, $LM(P)$ is not the lexicographically largest monomial in P , a contradiction.

Here's a simple example to illustrate the idea behind this proof: Let $P = x^2y + xy^2 \in \mathbb{Q}[x, y]^{S_2}$. Suppose we pick $LM(P) = xy^2$ (obviously this is the wrong choice, but that's the contradiction we'll see). We observe that $2 = a_2 > a_1 = 1$ in this case. Let $\sigma = (12)$. Then $\sigma LM(P) = yx^2 = x^2y \succ xy^2 = LM(P)$. So $\sigma LM(P) \succ LM(P)$. Thus, $LM(P)$ is not the lexicographically largest monomial in P , and we have formally proven that our initial choice of $LM(P)$ was incorrect. \square

- We now have everything we need to prove the fundamental theorem. As such, we will restate and prove it.
- Theorem (Fundamental theorem of symmetric polynomials): We have that

$$\mathbb{Q}[x_1, \dots, x_n]^{S_n} = \mathbb{Q}[\sigma_1, \dots, \sigma_n]$$

Proof. We will prove this theorem using the well-ordering principle (every set of natural numbers has a smallest element), which is equivalent to induction. Let's begin.

Suppose for the sake of contradiction that there exists a symmetric polynomial that cannot be expressed via $\sigma_1, \dots, \sigma_n$. Given this counterexample, factor out as many terms as we want (successively reducing the degree) until it ceases to be a counterexample, thus yielding the counterexample of smallest degree. Similarly, get to the counterexample with smallest LM . Call this counterexample $P(x_1, \dots, x_n)$. Let

$$P(x_1, \dots, x_n) = C_{LM}(P) \underbrace{x_1^{a_1} \cdots x_n^{a_n}}_{LM(P)} + \text{smaller monomials}$$

Since P is symmetric and the term above is the lexicographically largest monomial, the Lemma implies that $a_1 \geq \cdots \geq a_n$. We now construct a polynomial Q out of the σ_i such that $LM(P) = LM(Q)$. To begin, note that

$$LM(\sigma_1) = x_1 \quad LM(\sigma_2) = x_1x_2 \quad \cdots \quad LM(\sigma_n) = x_1 \cdots x_n$$

Now consider $\sigma_n^{a_n}$. This clearly divides $LM(P)$ since a_n is minimal. Now multiply by $\sigma_{n-1}^{a_{n-1}-a_n}$. Continuing on, we get

$$Q = \sigma_n^{a_n} \sigma_{n-1}^{a_{n-1}-a_n} \sigma_{n-2}^{a_{n-2}-a_{n-1}} \cdots \sigma_1^{a_1-a_2}$$

Now it follows that

$$LM(P - C_{LM}(P) \cdot Q) \prec LM(P)$$

Since $C_{LM}(P) \in \mathbb{Q}$ and $Q \in \mathbb{Q}[x_1, \dots, x_n]^{S_n}$, it also follows that $P - C_{LM}(P) \cdot Q \in \mathbb{Q}[x_1, \dots, x_n]^{S_n}$. But then by the assumption that P was the counterexample with smallest LM , we know that $P - C_{LM}(P) \cdot Q \in \mathbb{Q}[\sigma_1, \dots, \sigma_n]$. It follows that $P \in \mathbb{Q}[\sigma_1, \dots, \sigma_n]$, a contradiction. \square

- Note: This is an effective proof; we can write an algorithm to do this for us, and it's actually pretty fast and efficient.
- ?? (word in blackboard picture) to show: $\sigma_1, \dots, \sigma_n$ are algebraically independent $P(\sigma_1, \dots, \sigma_n) = 0$ implies that $P = 0$.

- This will be a homework problem; hint, it's pretty easy.
- Back to representation theory.
- **Antisymmetric** (polynomial): A polynomial $P(x_1, \dots, x_n)$ such that

$$\sigma P = (-1)^\sigma P$$

- Example.
 - $n = 2$: $x_1 - x_2$.
 - $n = 3$: $(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)$.
- These main examples are the **Vandermonde determinant** from last time!
- **Vandermonde determinant**:

$$\Delta(x_1, \dots, x_n) = \prod_{1 \leq i < j \leq n} (x_j - x_i)$$

- Exercise: $\Delta(x_1, \dots, x_n)$ is antisymmetric.
- One of the nicest definitions of sign comes from these determinants!

$$(-1)^\sigma = \frac{\prod_{i < j} (x_{\sigma(i)} - x_{\sigma(j)})}{\prod_{i < j} (x_i - x_j)}$$

- Theorem: If $P \in \mathbb{Q}[x_1, \dots, x_n]^{\text{alt}}$ (i.e., $P \in \mathbb{Q}[x_1, \dots, x_n]$ and P is antisymmetric), then $P = P' \Delta(x_1, \dots, x_n)$ where P' is symmetric (i.e., $P' \in \mathbb{Q}[x_1, \dots, x_n]^{S_n}$).
- Corollary: If P is antisymmetric and $\deg(P) < n(n-1)/2$, then $P = 0$.
 - We'll use this many times, this fact that “antisymmetric polynomials have a smallest possible degree.”
- We now prove the Theorem.

Proof. Let P be antisymmetric. Then $(12)P = -P$. It follows that $P(x_1, \dots, x_n)|_{x_1=x_2} = 0$. Now, rewrite P as a polynomial in one variable where all of the coefficients are polynomials in other variables. In particular, let

$$P = P_d(x_1 - x_2)^d + P_{d-1}(x_1 - x_2)^{d-1} + \dots + P_0$$

where each $P_i \in \mathbb{Q}[x_1, \dots, x_d]$. What is d ?? (Less than n , I'm assuming, but any other constraints?) Plugging in $x_1 = x_2$ once again, we get $0 = P = P_0$. But this implies that $P_0 = 0$. Thus, P is divisible by $x_1 - x_2$. Similarly, for all $i < j$, $(x_i - x_j) \mid P$. But since the $x_i - x_j$ are irreducible polynomials, we have that $\prod_{i < j} (x_j - x_i) \mid P$. This is justified because we are in a unique factorization domain (how is this relevant??). Thus, we have that $P = P' \cdot \Delta(x_1, \dots, x_n)$. Lastly, it follows that $P' \in \mathbb{Q}[x_1, \dots, x_n]^{S_n}$ because under any sign -1 permutation, $\Delta(x_1, \dots, x_n)$ will flip signs and P will still be equal, so P' had better just stay itself under this permutation (i.e., be symmetric). \square

- Remark: Where does the name Vandermonde *determinant* come from?
 - We have that

$$\Delta(x_1, \dots, x_n) = \begin{vmatrix} 1 & 1 \\ x_1 & x_n \\ \vdots & \vdots \\ x_1^{n-1} & x_n^{n-1} \end{vmatrix}$$

- Final reminder before the final.
 - Don't forget our awesome central construction!
 - If $\lambda = (\lambda_1, \dots, \lambda_k)$ is a partition with $\lambda_1 \geq \dots \geq \lambda_k$, then we can draw a Young Diagram and construct an associated representation $V_\lambda \in \mathbb{C}[x_1, \dots, x_n]$.
 - But what we do is $V_\lambda = \mathbb{C}[S_n]\Delta_\lambda$, where

$$\Delta_\lambda = \Delta(x_1, \dots, x_{\lambda'_1})\Delta(x_{\lambda'_1+1}, \dots, x_{\lambda'_1+\lambda'_2}) \cdots$$

- Example: For $\lambda = (2, 2, 1)$, we have $V_{(2,2,1)} = \mathbb{C}[S_n](x_1 - x_2)(x_1 - x_3)(x_2 - x_3)(x_4 - x_5)$.
 - Next time, we'll prove that $V_{(2,2,1)}$ is irreducible.
- This Specht construction is in a tiny footnote of Fulton and Harris (2004), but that's about it!