Moduli Theory and Invariants

Lectures by Tom Garrity
PCMI 2015 Undergraduate Summer School

Notes by David Mehrle
dfm223@cornell.edu

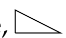
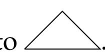
Contents

Lecture 1 — June 29, 2015

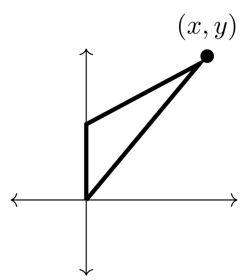
As kids, we learn that "functions describe the world". This is true in other forms of math too. For instance in algebraic geometry we study polynomials, and in complex analysis we study holomorphic functions.

Another big theme of math is **equivalence problems**, i.e. when are things the same? In topology, things are "the same" up to bending and twisting; in algebraic geometry, things are the same up to change of coordinates.

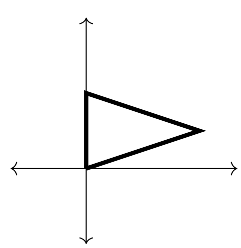
This is the idea behind **moduli spaces**. We classify some sorts of objects up to equivalence, and the equivalence classes are the moduli space. Ideally, we want to define our moduli spaces such that they have nice geometric or algebraic structure. This is **HARD**!

Example 1. We'll focus on the moduli space of triangles in an extended example. The equivalence relation here is similarity, as in the sense of basic high-school geometry. For instance,  is similar to  but not to .

For the moduli space of triangles, we may assume that one vertex is at the origin, with the shortest side the segment from $(0,0)$ to $(0,1)$. We place the triangle entirely in the first quadrant with the longest side having one of its endpoints at the origin. Because we only care about triangles up to similarity, all of these moves are okay.



Due to our clever maneuvering of these triangles, any triangle is described uniquely by the point (x,y) with a few caveats. If the triangle is isosceles, then by moving the point (x,y) either vertically up or vertically down we get the same triangle. So our moduli space is a bit redundant in that regard. In fact, we want to remove the redundant points in our moduli space to make it nicer.



This set also has some geometric structure. Namely, it's locally a manifold! If we wiggle the point (x, y) in the first picture in any direction, we get a different, equally valid scalene triangle. So there is some open neighborhood of (x, y) where this looks like $\mathbb{R}^2$.

Problem 2. What does the moduli space of triangles up to similarity properly look like (i.e. without redundancies)? What about quadrilaterals up to similarity? What about n-gons? (This last one is an open problem!)

Now for some algebraic geometry. Classical algebraic geometry sought to classify conics up to a change of coordinates (where by change of coordinates we mean a linear transformation with some translation). There are seven types of conics, of which four are degenerate.

- Ellipses $ax^2 + by^2 = 1$;
- Hyperbolae $ax^2 - by^2 = 1$;
- Parabolae $y = ax^2 + bx + c$;
- Double lines $(x - y)^2 = 0$;
- Two lines $(x + y)(x - y) = 0$;
- Points $x^2 + y^2 = 0$;
- Empty set $x^2 + y^2 = -1$.

Note that all ellipses are equivalent up to change of coordinates. For example, changing a circle $x^2+y^2 = 1$ to a ellipse $4u^2+9v^2 = 1$ is simply done by the change of coordinates $x = 2u$, $y = 3v$. Likewise, all hyperbolae are equivalent up to real change of coordinates, and all parabolae are equivalent, all double lines, etc. But a hyperbola cannot become an ellipse under any real change of coordinates.

It's different in the complex numbers. Here, by the change of coordinates $x = u$ and $y = iv$, we can turn a circle into a hyperbola. Moreover, certain degeneracies vanish. For example, $x^2 + y^2 = -1$ is no longer empty, and $x^2 + y^2 = 0$ is more than just a point.

However, we still cannot make a circle into a parabola just by working over $\mathbb{C}$. Ideally, the three nondegenerate classes would all join together and the degeneracies would go away. So we work in projective space.

Definition 3. n-dimensional **Projective space** over the complex numbers is defined as

$$\mathbb{P}^n\mathbb{C} = \mathbb{C}^{n+1}\backslash\{(0, 0, 0)\}/\sim$$

where $\sim$ is the equivalence relation $\lambda(x, y, z) \sim (x, y, z)$ for complex $\lambda \neq 0$.

We can think of $\mathbb{P}^2\mathbb{C}$ as lines through the origin in $\mathbb{C}^3$, because each equivalence class of a point in $\mathbb{P}^n\mathbb{C}$ determines a line; two points are on the same line if and only if they are scalar multiples of each other.

And in $\mathbb{P}^2\mathbb{C}$, $(x, y, z) \sim (x/y, x/z, 1)$ for $z \neq 0$, so we think of the plane $z = 1$ in $\mathbb{P}^2\mathbb{C}$ as a copy of $\mathbb{C}^2$. Each line through the origin passes through this plane in exactly one point, and hence gives a point $(x/y, x/z)$ in $\mathbb{C}^2$. Therefore, we think of $\mathbb{P}^2\mathbb{C}$ as $\mathbb{C}^2$ with some extra "points at infinity" where $z = 0$.

We can take any conic in $\mathbb{C}^2$ and **homogenize** it to get a conic in $\mathbb{P}^2\mathbb{C}$. To homogenize, replace x by x/z and y by y/z and then clear denominators.

$$x^2 + y^2 = 1 \rightsquigarrow x^2 + y^2 = z^2$$

Our change of coordinates is slightly different in projective space too.

$$\begin{aligned} x &= au + bv + cw \\ y &= du + ev + fw \\ z &= gu + hv + jw \end{aligned}$$

$$\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & j \end{bmatrix} \neq 0$$

Problem 4.

In $\mathbb{P}^2\mathbb{C}$, there are only three conics up to change of coordinates: circles, two distinct lines, and double lines. Each is described by a homogeneous equation

$$ax^2 + 2bxy + 2cxz + dy^2 + 2eyz + fz^2 = 0,$$

or equivalently a vector in $\mathbb{C}^6 \backslash \{\vec{0}\}$

$$\begin{bmatrix} a \\ 2b \\ 2c \\ d \\ 2e \\ f \end{bmatrix}.$$

As a final note, finding all higher degree curves is very very hard. Degree three are elliptic curves, and that topic is a large portion of the whole field of number theory!

Lecture 2 June 30, 2015

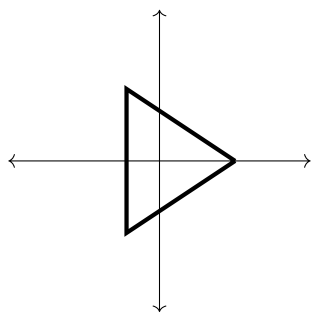
Today we'll talk about groups acting on objects we care about. Yesterday, we saw that the projective line was the right space to think about in algebraic geometry. The notion of "sameness" in algebraic geometry is equivalence under change of coordinates. We saw that a projective change of coordinates is given by

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

This matrix also defines a transformation on conics in $\mathbb{P}^2\mathbb{C}$. This is an example of a group representation!

Definition 5. A **group representation** of a group G is a group homomorphism $\rho\colon G \to \mathrm{GL}(V)$ for some vector space V.

Example 6. The cyclic group of order 3, C_3 may be realized as either $\mathbb{Z}/3\mathbb{Z}$ or $\langle t \mid t^3 = 1\rangle$, among many other ways. These are examples of different ways to represent a group, but not representations. This group can be represented as matrices that show symmetries of a triangle in a map $T\colon \langle t \mid t^3 = 1\rangle \to \mathrm{GL}_2\mathbb{C}$.



$$\begin{aligned} T(1) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ T(t) &= \begin{bmatrix} \cos(2\pi/3) & -\sin(2\pi/3) \\ \sin(2\pi/3) & \cos(2\pi/3) \end{bmatrix} = \begin{bmatrix} \frac{-1}{2} & \frac{-\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{-1}{2} \end{bmatrix} \\ T(t^2) &= \begin{bmatrix} \cos(4\pi/3) & -\sin(4\pi/3) \\ \sin(4\pi/3) & \cos(4\pi/3) \end{bmatrix} \end{aligned}$$

Problem 7. Why is

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

rotation by θ? Why does rotation by angles (rotation matrices) form a group?

Change of coordinates relevant for conics in $\mathbb{P}^2$ corresponds to a representation $\rho\colon \mathrm{GL}(3) \to GL(6)$. A conic in $\mathbb{P}^2$ is

$$ax^2 + bxy + cxz + dy^2 + eyz + fz^2 = 0.$$

This can be thought of as just the coefficients.

$$\begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} \in \mathbb{C}^6 \backslash \{(0,0,0,0,0,0)\}.$$

If our change of coordinates is given by the matrix equation

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix},$$

where $A = (a_{ij})$ is the matrix with $\det A \neq 0$. Changing coordinates is a representation of $\mathrm{GL}_3\mathbb{C}$, and in particular a map $\rho\colon \mathrm{GL}_3\mathbb{C} \to \mathrm{GL}_6\mathbb{C}$.

Example 8. Let's think about cubics in $\mathbb{P}^1\mathbb{C}$, with an equation

$$ax^3 + bx^2y + cxy^2 + dy^3 = 0 \text{ or } \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathbb{C}^4 \backslash \{\vec{0}\}.$$

The change of coordinates $x = u + v$ and $y = v$ is represented by the following matrix equation:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}.$$

After change of coordinates, we get

$$ax^3 + bx^2y + cxy^2 + dy^3 \mapsto au^3 + (3a+b)u^2v + (3a+3b+c)uv^2 + (a+b+c+d)v^3.$$

We can rewrite it as a transformation of the coefficients in a matrix equation

$$\begin{bmatrix} a \\ 3a+b \\ 3a+3b+c \\ a+b+c+d \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

Here, $\rho\colon \mathrm{GL}_2(\mathbb{C}) \to \mathrm{GL}_4\mathbb{C}$, so we get that

$$\rho\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Problem 9. It's not immediately clear that this defines a group homomorphism. Show it!

Problem 10. Do this for quartics and quintics. Do it in general for an arbitrary change of coordinates $A = (a_{ij})_{1 \leqslant i,j \leqslant 2}$ for cubics in $\mathbb{P}^1$ ($\rho\colon \mathrm{GL}_2\mathbb{C} \to \mathrm{GL}_4\mathbb{C}$). Do it in general for an arbitrary change of coordinates $A = (a_{ij})_{1 \leqslant i,j \leqslant 3}$ for conics in $\mathbb{P}^2$ ($\rho\colon \mathrm{GL}_3\mathbb{C} \to \mathrm{GL}_6\mathbb{C}$). Use Mathematica!

Definition 11. Let G be a group. Let $\rho_1\colon G \to \mathrm{GL}(V)$ and $\rho_2\colon G \to \mathrm{GL}_n\mathbb{C}$ be two representations. We say ρ_1 **is equivalent to** ρ_2 if they are the same under change of basis, i.e. there is some change of basis matrix D such that for all $g \in G$, $D\rho_1(g)D^{-1} = \rho_2(g)$. ρ_1 and ρ_2 are really the same representation, but they looked different because we chose bases for V.

Now let's talk about irreducible representations. For our representation $\rho\colon G \to \mathrm{GL}(V)$, we can make a new one. This new representation is called $\hat{\rho}\colon G \to \mathrm{GL}(V) \oplus \mathrm{GL}(V)$ and given by $\hat{\rho}(g) = \rho(g) \oplus \rho(g)$. This is an example of a **reducible** representation.

Definition 12. A representation $\rho\colon G \to \mathrm{GL}(V)$ is reducible if ρ is equivalent to a representation $\hat{\rho}$ that is a block diagonal matrix for every g. That is, for each $g \in G$, $\hat{\rho}(g)$ looks like

$$\hat{\rho}(g) = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

for matrices A, B.

Since we can build every representation from irreducible representations, then we want to find all of the irreducible representations for a given group G.

Example 13. Let's find some representations of the first interesting symmetric group.

$$S_3 = \{\mathrm{id}, (1\ 2), (1\ 3), (1\ 2\ 3), (1\ 3\ 2)\}$$

There's the boring representation $\rho_1(\sigma) = 1$ for all $\sigma \in S_3$. (called the **trivial representation**).

Another representation is the sign representation, $\rho_2(\sigma) = \mathrm{sgn}(\sigma)$ for all $\sigma \in S_3$.

The last one is the representation given by $S_3 \circlearrowright \mathbb{C}^2$ by rotations and reflections.

$$\begin{aligned}
\rho_3(\mathrm{id}) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
\rho_3(1\ 2\ 3) &= \begin{bmatrix} \cos(2\pi/3) & -\sin(2\pi/3) \\ \sin(2\pi/3) & \cos(2\pi/3) \end{bmatrix} \\
\rho_3(1\ 3\ 2) &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}
\end{aligned}$$

Why do we care? Representations of the Poincaré group, which describes symmetries of the universe, are fundamental to physics. From the irreducible representations we can rediscover Maxwell's laws and the Dirac equation, which is amazing.

Lecture 3 July 2, 2015

We've been talking about equivalence problems, in which we want to classify objects up to equivalence. The notion of equivalence we've been using is some group acting on these objects, and if they lie in the same orbit, they're the same.

Last time, we were thinking about the group $\mathrm{GL}_n\mathbb{C}$ acting by change of variables on $\mathbb{P}^n$. We got representations of $\mathrm{GL}_n\mathbb{C}$ by

$$\mathrm{GL}_n\mathbb{C} \to \mathrm{GL}(V)$$

for a much larger space V. We ended up with the definition of a representation

Definition 14. A **representation** of a group G is a group homomorphism $G \to \mathrm{GL}(V)$ for some vector space V.

Essentially, we represent elements of the group as matrices. We want to understand the irreducible representations of a group, which are building blocks for all representations.

Now we'll look at the corresponding problem for finite groups. The outline for the next few days is as follows:

1. there are only a finite number of irreducible representations for any finite group G;
2. the number of irreducible representations of G is the same as the number of conjugacy classes of the group G;
3. there is a function $\chi_T : G \to \mathbb{C}$ associated to each representation $T : G \to \mathrm{GL}(V)$, called its **character**, that contains lots of information;
4. under suitable inner products, these characters for the irreducible representations form an orthonormal basis for all such functions.

The first three are kind of reasonable. The last one is in the realm of miracles. We encode the characters for a group in something called a character table.

First, conjugacy classes.

Definition 15. Let G be a group. We say that x is **conjugate to** y if there is some $g \in G$ such that $gxg^{-1} = y$.

If G is a matrix group, this is a change of basis. So we're trying to capture the notion of change of basis in an abstract group. An important fact is that conjugacy is an equivalence relation between elements of the group. The equivalence classes under this relation are called **conjugacy classes**.

Example 16. In S_3, there are three conjugacy classes.

$$\{\text{id}\} \qquad \{(1\,2), (1\,3), (2\,3)\} \qquad \{(1\,2\,3), (1\,3\,2)\}.$$

Example 17. In any abelian group, x is conjugate to y if and only if for some $g \in G$, $gxg^{-1} = y$ if and only if $gg^{-1}x = y$ if and only if $x = y$. So each element is in its own conjugacy class.

As a bit of a preview, here's the definition of the character of a representation.

Definition 18. Let $T\colon G \to \mathrm{GL}(V)$ be a representation. Then the **character** of T is the function $\chi_T\colon G \to \mathbb{C}$ given by

$$\chi_T(g) \;:=\; \mathrm{tr}(T(g)).$$

Problem 19. Why is the character defined by the trace? Why not determinant? It's totally reasonable to define the character this way, but it seems reasonable to define it using the determinant as well. (Note: the answer "because it works" is not acceptable).

Example 20. The characters corresponding to the three representations of S_3 are as follows

S_3	1	(1 2)	(1 3)	(2 3)	(1 2 3)	(1 3 2)
χ_{T_1}	1	1	1	1	1	1
χ_{T_2}	1	-1	-1	-1	1	1
χ_{T_3}	2	0	0	0	-1	-1

this is a primitive form of what we call a **character table**.

Notice that the characters are constant on character classes. This is in fact always true, not just rigged for S_3! Real character tables will lump all conjugacy classes together in the columns.

Theorem 21. If x is conjugate to y, then $\chi_T(x) = \chi_T(y)$ for any representation T.

Proof. There is some $g \in G$ such that $gxg^{-1} = y$.

$$\begin{aligned}
\chi_T(y) &= \mathrm{tr}(T(y)) \\
&= \mathrm{tr}(T(gxg^{-1})) \\
&= \mathrm{tr}(T(g)T(x)T(g^{-1})) \\
&= \mathrm{tr}(T(g)T(x)T(g)^{-1}) \\
&= \mathrm{tr}(T(g)T(g)^{-1}T(x)) \\
&= \mathrm{tr}(T(x)) = \chi_T(x)
\end{aligned}$$

□

This means we can rewrite the character table for S_3 lumping together the conjugacy classes in the columns, as follows.

	$\{\text{id}\}$	$\{(1\ 2), (1\ 3), (2\ 3)\}$	$\{(1\ 2\ 3), (1\ 3\ 2)\}$
χ_{T_1}	1	1	1
χ_{T_2}	1	-1	1
χ_{T_3}	2	0	-1

This is usually how we write character tables. There's a huge amount of structure here, and again it's not by coincidence. The sum of the squares of entries in the first column is always the size of the group. The rows are orthonormal when weighted by the size of each of the conjugacy classes. The columns are orthonormal.

For the symmetric group S_n, the conjugacy class depends on cycle type, which is the same as the number of partitions of n, as defined below.

Definition 22. A **partitions of** n is the number of ways to summing

$$n = k_1 + k_2 + \ldots + k_\ell$$

with $k_1 \geqslant k_2 \geqslant \ldots \geqslant k_\ell$. We denote by $P(n)$ the number of partitions of n.

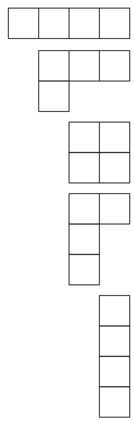
Example 23.

n	$P(n)$	how to sum to n
1	1	$1 = 1$
2	2	$2 = 2 = 1 + 1$
3	3	$3 = 3 = 2 + 1 = 1 + 1 + 1$
4	5	$4 = 4 = 3 + 1 = 2 + 2 = 2 + 1 + 1 = 1 + 1 + 1 + 1$

We should make that fact we said earlier a theorem.

Theorem 24. The number of partitions of n is equal to the number of conjugacy classes of S_n.

A diagrammatic way of representing the partitions of n is **Young Diagrams**. These are little box diagrams corresponding to the partitions. The number of diagrams with n boxes is equal to $P(n)$.



$4 = 4$

$4 = 3 + 1$

$4 + 2 + 2$

$4 = 2 + 1 + 1$

$4 = 1 + 1 + 1 + 1$

We will use these Young diagrams later to understand the irreducible representations of S_n.

Lecture 4 July 3, 2015

As always, we're thinking about equivalence problems where equivalent is moving things around with groups. We talked about how to represent groups, and then about the irreducible representations of a group. Each irreducible representation comes with a character, and the character table encodes all the information about representations of a group.

Let G be a finite group. Say we have k-many conjugacy classes, $C_1, \ldots, C_k$. This means, as last lecture, that there are necessarily k-many irreducible representations. Let $c_i = \#C_i$ be the number of elements in a conjugacy class C_i. Let g_i be the Then the inner product on characters (or really, any **class function**, i.e. a function constant on each C_i) is as follows.

Definition 25. The **inner product** on two class function $\nu, \mu \colon G \to \mathbb{C}$ is

$$\langle \nu, \mu \rangle := \frac{1}{|G|} \sum_{i=1}^{k} c_i \nu(g_i) \overline{\mu(g_i)},$$

where $g_i \in C_i \subset G$ is a representative of the conjugacy class.

This inner product is defined by a matrix

$$\begin{bmatrix} \frac{c_1}{|G|} & 0 & \ldots & 0 \\ 0 & \frac{c_2}{|G|} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \frac{c_k}{|G|} \end{bmatrix}$$

The goal of this lecture is to see why, for two representations, ρ and ψ, the characters are orthogonal.

$$\langle \chi_\rho, \chi_\psi \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

The vital piece of this theorem is **Schur's Lemma**, which says the following.

Lemma 26 (Schur's Lemma)**.** Let G be a finite group and let $T_1 \colon G \to \mathrm{GL}(V_1)$ and let $T_2 \colon G \to \mathrm{GL}(V_2)$ be irreducible representations. Let $S \colon V_1 \to V_2$ be a linear map. Assume for all $g \in G$, $ST_1(g) = T_2(g)S$.

(a) If T_1 and T_2 are inequivalent, then $S = 0$;

(b) if $V_1 = V_2$, and $T_1 = T_2$, then $S = \lambda\mathrm{id}$.

The condition that we assume is that the following diagram commutes. We call S an **intertwining operator**.

$$\begin{array}{ccc} V_1 & \xrightarrow{T_1(g)} & V_1 \\ {\scriptstyle S}\downarrow & & \downarrow{\scriptstyle S} \\ V_2 & \xrightarrow{T_2(g)} & V_2 \end{array}$$

Recall that

Definition 27. $T \colon G \to \mathrm{GL}(V)$ is **reducible** if T is equivalent to a representation $A \oplus B$ for representations $A \colon G \to \mathrm{GL}(V_1)$, $B \colon G \to \mathrm{GL}(V_2)$ such that $V_1 \oplus V_2 = V$. More concretely, $T(g)$ is equivalent to a matrix

$$\left(\begin{array}{c|c} A(g) & 0 \\ \hline 0 & B(g) \end{array} \right)$$

under some change of basis.

Notice that the subspace V_1 of V is invariant under the action of $T(g)$ for each $g \in G$, because of the block matrix form of $T(g)$. Therefore, we can make an alternative (and much neater) definition of reducible.

Definition 28. $T \colon G \to \mathrm{GL}(V)$ is **reducible** if there is a nontrivial subspace $W \subseteq V$ such that for all $g \in G$ and $w \in W$, $T(g)(w) \in W$.

Now we can prove Schur's Lemma.

Proof of Schur's Lemma 26. We first prove part (a). Assume that T_1 is not equivalent to T_2, which means there is no matrix P such that $PT_1(g)P^{-1} = T_2(g)$. If $S = 0$, then we're done!

So now suppose that $S \neq 0$. We will show that S is both injective and surjective, which means that it's bijective and therefore $ST_1(g)S^{-1} = T_2(g)$. This is a contradiction, because we assumed that T_1 and T_2 were *not* equivalent.

Let $W = \ker S \subseteq V$. S is injective if and only if $\ker S = 0$. We will show that W is an invariant subspace, i.e. for all $g \in G$ and $w \in W$, $T_1(g)(w) \in W$. If true, then $\dim W = 0$ or $\dim W = \dim V_1$. The latter implies $S = 0$, which contradicts our assumption that $S \neq 0$.

So now we show that for all $g \in G$, $w \in W$, $T_1(g)(w) \in W$. Because $W = \ker S$, we now show that $S(T_1(g)(w)) = 0$.

$$S(T_1(g)(w)) = T_2(g)(S(w)) = T_2(g)(0) = 0.$$

So $T_1(g)(w) \in \ker S = W$. Hence W is an invariant subspace. Therefore, either $\dim W = 0$ or $\dim W = \dim V_1$. In the latter case, $\ker S = W = V_1$, and so $S = 0$, which we assumed was not the case. In the former case, $\dim W = 0$, so $\ker S = 0$ and S must be injective.

Now we show that S is surjective. Let $W = \operatorname{im} S$. The strategy here is the same. We will show that for all $W \in W$, $T_2(g)(w) \in W$, which implies by a similar argument to the above that either $W = 0$ or $W = V_2$.

To that end, let $v_2 \in W \subseteq V_2$. Since $W = \operatorname{im} S$, there is some $v_1 \in V_1$ with $S(v_1) = v_2$. Then

$$T_2(g)(v_2) = T_2(g)(S(v_1)) = S(T(g)(v_1)) \in \operatorname{im} S$$

Hence, $T_2(g)(v_2) \in \operatorname{im} S = W$. So W is an invariant subspace under the action of $T_2(g)$ for each $g \in G$. Because T_2 is an irreducible representation, this means that either $W = V_2$ or $W = 0$. If $W = 0$, then $S = 0$, which is a contradiction because we assumed $S \neq 0$. This leaves the former possibility, namely that $W = V_2$, then S is surjective, as desired.

So S is both surjective and injective, so it is an isomorphism. Hence, $ST_1(g) = T_2(g)S$ implies that $ST_1(g)S^{-1} = T_2(g)$, so T_1 and T_2 are equivalent. This contradicts our assumption that they were inequivalent!

Now we prove part (b). Here, $V_1 = V_2 = V$, and $T_1 = T_2 = T$. We want to show that $S = \lambda \mathrm{id}_V$ for some $\lambda \in \mathbb{C}$.

Let λ be an eigenvalue of $S\colon V \to V$. Hence, λ is a root of $\det(x\mathrm{id}_V - S) = 0$. Let w be an eigenvector for λ, so $Sw = \lambda w$. Define $\hat{S} = S - \lambda \mathrm{id}_V$. We will show that for all $v \in V$, $v \in \ker(\hat{S})$, and therefore $Sv = \lambda v$ for all $v \in V$.

This will follow if we can show for all $g \in G$, $\hat{S}T(g) = T(g)\hat{S}$ and then apply part (a).

$$\begin{aligned}
\hat{S}T(g)(v) &= ST(g)v - \lambda \mathrm{id}_V T(g)v \\
&= T(g)Sv - T(g)(\lambda \mathrm{id}_V v) \\
&= T(g)\Big(Sv - \lambda \mathrm{id}_V v\Big) \\
&= T(g)\hat{S}v
\end{aligned}$$

Now we can apply part (a) to claim that $\hat{S} = 0$. Expanding the definition of $\hat{S}$, we see that $S - \lambda \mathrm{id}_V = 0$, so therefore $S = \lambda \mathrm{id}_V$. $\square$

This proof doesn't use the fact that G is a finite group. In fact, it holds in much more generality.

Lecture 5 July 6, 2015

Last week, we talked about sets of objects up to equivalence, which are called Moduli spaces. The equivalence relations we care about are the orbits of a group action.

Let G be a finite group. Today, we will show that characters of irreducible representations form an orthonormal basis for the space of class functions subject to the inner product

$$\langle \chi_T, \chi_S \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_T(g) \overline{\chi_S(g)}.$$

Here, T, S are irreducible representations. Recall that

$$\chi_T(s) = \mathrm{tr}(T(g)) = \text{ sum of diagonal of matrix } = \text{ sum of eigenvalues } .$$

Let $g \in G$, and let $T \colon G \to \mathrm{GL}_n\mathbb{C}$. Let λ be an eigenvalue of $T(g)$. We claim that λ lies on the unit circle. Let $\vec{v}$ be an eigenvector for $T(g)$ corresponding to λ. Then,

$$T(g^2)\vec{v} = T(g)T(g)\vec{v} = T(g)\lambda\vec{v} = \lambda^2\vec{v}$$

Thus, we can conclude

$$T(g^k)\vec{v} = \lambda^k\vec{v}.$$

Since G is a finite group, g has a finite order, say $g^n = 1$. So,

$$\vec{v} = T(1)\vec{v} = T(g^n)\vec{v} = \lambda^n\vec{v},$$

and therefore $\lambda^n = 1$. Doing some more manipulations, we find that

$$\frac{1}{\lambda} = \lambda^{n-1}.$$

Now if $\lambda = a + bi$, then

$$\frac{1}{\lambda} = a - bi = \overline{\lambda}.$$

What does this tell us about characters? Well, the trace is the sum of the eigenvalues, so

$$\chi_T(g) = \mathrm{tr}(T(g)) = \lambda_1 + \ldots + \lambda_n$$

Taking conjugates,

$$\overline{\chi_T(g)} \;=\; \overline{\lambda_1} + \ldots + \overline{\lambda_n}.$$

Using that $\frac{1}{\lambda} = \overline{\lambda}$, then

$$\overline{\chi_T(g)} \;=\; \frac{1}{\lambda_1} + \ldots + \frac{1}{\lambda_n} = \chi_T(g^{-1}),$$

because the eigenvalues of the inverse of a matrix are the reciprocals of the eigenvalues.

That was just the first step. Let T and S be inequivalent irreducible representations.

$$T\colon G \to \mathrm{GL}_n\mathbb{C}$$
$$S\colon G \to \mathrm{GL}_m\mathbb{C}$$

Our goal is to show that

$$0 \;=\; \frac{1}{|G|}\sum_{g\in G}\chi_T(g)\overline{\chi_S(g)}. \tag{1}$$

Notice first that

$$\frac{1}{|G|}\sum_{g\in G}\chi_T(g)\overline{\chi_S(g)} \;=\; \frac{1}{|G|}\sum_{g\in G}\chi_T(g)\chi_S(g^{-1}).$$

We're going to rewrite this in terms of not eigenvalues, but instead as the sums of diagonals of matrices. Let

$$T(g) = (t_{ij}(g))_{1\leqslant i,j\leqslant n} \qquad\qquad S(g) = (s_{ij}(g))_{1\leqslant i,j\leqslant m},$$

so for example

$$\chi_T(g) \;=\; \sum_{i=1}^{m} t_{ii}(g).$$
$$\chi_S(g^{-1}) \;=\; \sum_{i=1}^{m} s_{ii}(g^{-1}).$$

Let's plug these into the equation 1 from earlier. We now want to show that

$$0 \;=\; \frac{1}{|G|}\sum_{g\in G}\left(\sum_{i=1}^{m} t_{ii}(g)\right)\left(\sum_{i=1}^{m} s_{ii}(g^{-1})\right)$$

The strategy here is to plug in things to a polynomial equation and get zero, and from there conclude that all coefficients (these nasty sums) have to be zero. This is like saying that the polynomial $ax^2+bx+c=0$ for all x implies that $a=b=c=0$.

Let $X\colon \mathbb{C}^n \to \mathbb{C}^m$ be a matrix $X = (x_{ij})$ for $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$. X is our matrix of variables. Define $X^0\colon \mathbb{C}^n \to \mathbb{C}^m$ by

$$X^0 \;=\; \sum_{g\in G} S(g)^{-1} X T(g).$$

This new function is the composition

$$X^0\colon \mathbb{C}^n \xrightarrow{T(g)} \mathbb{C}^n \xrightarrow{X} \mathbb{C}^m \xrightarrow{S(g)} \mathbb{C}^m.$$

The key point here is that X^0 is an intertwining operator! This point is highlighted in the following fact.

Fact 29. For all $g \in G$, $X^0 T(g) = S(g) X^0$.

If we know this, then because S and T are inequivalent, X^0 must be zero.

Proof. Compute both sides.

$$\begin{aligned} X^0 T(h) \;&=\; \sum_{g\in G} S(g^{-1}) X T(g) T(h) \\ &=\; \sum_{g\in G} S(g^{-1}) X T(gh) \\ &=\; \sum_{k=gh} S(hk^{-1}) X T(k) \\ &=\; \sum_{k\in G} S(h) S(k^{-1}) X T(g) \\ &=\; S(h) \sum_{k\in G} S(k^{-1}) X T(g) \;=\; S(h) X^0 \end{aligned}$$ $\square$

Now we apply Schur's Lemma[1] and conclude $X^0 = 0$. Therefore, examining each individual entry of the matrix X^0,

$$0 \;=\; x^0_{ij} \;=\; \sum_{g\in G} \sum_{a=1}^{m} \sum_{b=1}^{n} s_{ia}(g^{-1}) x_{ab} t_{bj}(g)$$

Now let $a = i$, $b = j$. The coefficient of x_{ij} in the above sum is

$$0 \;=\; \sum_{g\in G} s_{ii}(g^{-1}) t_{jj}(g) \tag{2}$$

We're not quite there yet, but this looks a lot like the trace formula! It's enough to use it.

[1] without which we are nothing

$$
\begin{aligned}
\langle \chi_T, \chi_S \rangle &= \frac{1}{|G|} \sum_{g \in G} \mathrm{tr} T(g) \mathrm{tr} S(g^{-1}) \\
&= \frac{1}{|G|} \sum_{g \in G} \left(t_{11}(g) + t_{22}(g) + \ldots + t_{nn}(g)\right) \left(s_{11}(g^{-1}) + s_{22}(g^{-1}) + \ldots + s_{mm}(g^{-1})\right) \\
&= \frac{1}{|G|} \sum_{g \in G} \sum_{i=1}^{n} \sum_{j=1}^{m} t_{ii}(g) s_{jj}(g^{-1}) \\
&= \frac{1}{|G|} \sum_{i=1}^{n} \sum_{j=1}^{m} \left(\sum_{g \in G} t_{ii}(g) s_{jj}(g^{-1}) \right) \\
&= \frac{1}{|G|} \sum_{i=1}^{n} \sum_{j=1}^{m} \begin{pmatrix} \text{coefficient} \\ \text{of } x_{ij} \text{ as in (2)} \end{pmatrix} = 0
\end{aligned}
$$

Lecture 6 July 7, 2015

Last time we talked about the Schur orthogonality relations among characters of a finite group. This says that the rows of the character table form an orthonormal basis for the space of class functions on a group G, that is,

$$\{f \colon G \to \mathbb{C} \mid f(xgx^{-1}) = f(g) \text{ for all } g \in G\}.$$

The inner product on this space is given by

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

If χ_S, χ_T are characters of irreducible representations, then

$$\langle \chi_T, \chi_S \rangle = \begin{cases} 1 & T \cong S \\ 0 & \text{otherwise} \end{cases}.$$

There is a similar orthogonality relation among the columns of the character table, but there is a slightly different inner product.

Proposition 30. Let $g, h \in G$, and let χ_i for $i = 1, \ldots, k$ index the irreducible representations of G. Then if g is in conjugacy class C,

$$\frac{1}{|G|} \sum_{i=1}^{k} \chi_i(g) \overline{\chi_i(h)} = \begin{cases} |C| & \text{if } g \text{ and } h \text{ are conjugate} \\ 0 & \text{otherwise} \end{cases}.$$

Example 31. Let's go back to S_3. Before we have anything in the character table, we can determine what the dimensions of the representations are. Note that

$$\chi_V(\mathrm{id}) = \mathrm{tr}(\rho_V(\mathrm{id})) = \mathrm{tr}\,\mathrm{id}_V = \dim V.$$

Therefore, the first column is the dimensions. And by the column orthogonality, we know that the sum of squares of dimensions is $|S_3| = 6$. The only way to write 6 as the sum of three squares $6 = a^2 + b^2 + c^2$ is $6 = 1^2 + 1^2 + 2^2$, so the dimensions must be $1, 1$ and 2.

S_3	$\{\mathrm{id}\}$	$\{(1\ 2), (1\ 3), (2\ 3)\}$	$\{(1\ 2\ 3), (1\ 3\ 2)\}$
trivial	$a = 1$		
sgn	$b = 1$		
χ	$c = 2$		

Of course, we know that the first row is $(1, 1, 1)$ and the last row is $(2, 0, -1)$. We can write the inner product on characters as a bilinear form:

$$\langle \text{trivial}, \chi \rangle = \left\langle \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} \right\rangle = \frac{1}{6}[\overline{2}, \overline{0}, \overline{-1}] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{3}$$

Notice that the matrix has sizes of conjugacy classes on the diagonal!

Now let $T\colon S_3 \to \mathrm{GL}(n)$ be any representation, with character $\chi_T(\mathrm{id}) = 11$, $\chi_T(1\ 2) = -3$, and $\chi_T(1\ 2\ 3) = -1$. We want to decompose T into irreps. We know that $T(g)$ is similar to a matrix

$$\begin{bmatrix} T_1(g) & & & \\ & T_2(g) & & \\ & & \ddots & \\ & & & T_\ell(g) \end{bmatrix}$$

where each T_i is an irrep. So form this, we know that

$$\chi_T = \alpha \,\text{trivial} + \beta \,\text{sgn} + \gamma\chi,$$

and we want to find α, β, γ. So we just take some inner products to find that $\alpha = \langle \chi_T, \text{trivial} \rangle$, $\beta = \langle \chi_T, \text{sgn} \rangle$, and $\gamma = \langle \chi_T, \chi \rangle$. Let's compute it using the bilinear form from (3).

$$\langle \chi_T, \text{trivial} \rangle = \frac{1}{6}[1, 1, 1] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 11 \\ -3 \\ -1 \end{bmatrix} = \frac{1}{6}[1, 1, 1] \begin{bmatrix} 11 \\ -9 \\ -2 \end{bmatrix} = 0$$

$$\langle \chi_T, \text{sgn} \rangle = \frac{1}{6}[1, -1, 1] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 11 \\ -3 \\ -1 \end{bmatrix} = \frac{1}{6}[1, -1, 1] \begin{bmatrix} 11 \\ -9 \\ -2 \end{bmatrix} = 3$$

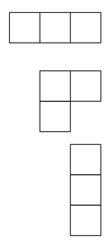
$$\langle \chi_T, \chi \rangle = \frac{1}{6}[2, 0, -1] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 11 \\ -3 \\ -1 \end{bmatrix} = \frac{1}{6}[2, 0, -1] \begin{bmatrix} 11 \\ -9 \\ -2 \end{bmatrix} = 4$$

From this, we know that the representation T is composed of 3 copies of the sign representation and 4 copies of the 2-dimensional representation.

Remark 32. Fourier analysis is the representation theory of an abelian infinite group, namely the unit circle S^1. This is an abelian group, so even though the group is infinite, each of the irreducible representations is one-dimensional. These representations are just $e^{i\theta}$. This is related to a construction on groups called the **Pontryagin Dual**.

For now, we'll be done with character tables and finite group representations. Instead, we'll focus on explicitly finding irreducible representations of the symmetric groups S_n. This is important to our eventual goal of going back to algebraic geometry, because the irreducible representations of S_n are very closely related to the irreducible representations of $\mathrm{GL}(n)$. And that's what we really want to know in algebraic geometry.

The first thing we need to know are the conjugacy classes of S_n. These correspond to the cycle types of elements, and are represented by Young diagrams. For instance, for $n = 3$, we get



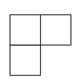
identity

two cycles

three-cycles

Definition 33. A **Young tabloid** is a Young diagram filled in with numbers from 1 to n, where n is the number of boxes in the diagram. They are considered equivalent up to permutation of the rows.

Example 34. For instance, with the Young diagram



there are 3 possibilities,

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

These correspond to the permutations (1 2)(3), (1 3)(2), and (2 3)(1), respectively.

Definition 35. A **partition** of n is a sequence $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ with $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_k$ such that $\lambda_1 + \lambda_2 + \ldots + \lambda_k = n$. A partition of n is often written $\lambda \dashv n$.

Young diagrams correspond to partitions, as described in a previous lecture. Briefly, the partition $(2,1)$ of 3 corresponds to $\ydiagram{2,1}$. We say a Young diagram **is of shape** λ where λ is the corresponding partition of n.

The irreps of S_n correspond to conjugacy classes, which correspond to Young diagrams. So it's natural to search for irreps in the Young diagrams.

To that end, define a representation space for S_n for λ a partition of n by

$$F(M_\lambda) = \bigoplus_{\substack{\text{Young tabloids } Y \\ \text{of shape } \lambda}} \mathbb{C}Y.$$

It's the $\mathbb{C}$-vector space with basis all of the Young tabloids of shape λ.

For example, when $\lambda = 2 + 1$ a partition of 3, we have an element of $F(M_{2+1})$ given by

$$(3,2,6) = 3\,\begin{ytableau} 1 & 2 \\ 3 \end{ytableau} + 2\,\begin{ytableau} 1 & 3 \\ 2 \end{ytableau} + 6\,\begin{ytableau} 2 & 3 \\ 1 \end{ytableau}.$$

What is the action of S_3 on this space? It just permutes the basis vectors by permuting the corresponding boxes in the Young tabloid. For instance, if $\sigma = (1\ 2) \in S_3$,

$$\sigma \cdot \begin{ytableau} 1 & 2 \\ 3 \end{ytableau} = \begin{ytableau} 2 & 1 \\ 3 \end{ytableau}.$$

Lecture 7 July 9, 2015

Last time, we talked about representations of the symmetric group. Our goal this time is to introduce an algorithm to explicitly find all irreducible representations for any S_n. We're going to start with an example and do it for S_3 and on your problem set you will do it for S_4.

The first thing we should know for our algorithm is the termination condition. How many irreducible representations should we have? We know that the number of irreducible representations is equal to the number of conjugacy classes. For S_n, the conjugacy classes are given by cycle type, which is captured by the Young diagrams. This number is equal to $P(n)$, the partition number of n. This is equal to the number of Young diagrams.

Example 36. For S_3, we need to find the partitions of 3. There are three of them: $3 = 3 = 2 + 1 = 1 + 1 + 1$. These correspond to Young diagrams.

$$\ydiagram{3} \quad (3)$$

$$\ydiagram{2,1} \quad (2,1)$$

$$\ydiagram{1,1,1} \quad (1,1,1)$$

Our goal here is to match each Young diagram with the corresponding irreducible representation of S_n. First, we introduce a partial ordering on Young diagrams.

Definition 37. The **partial ordering on partitions / Young diagrams** is given for two partitions of n, called $\mu = (\mu_1, \ldots, \mu_n)$ and $\lambda = (\lambda_1, \ldots, \lambda_n)$ as follows. We say $\lambda \geqslant \mu$ if and only if

$$\begin{aligned} \lambda_1 &\geqslant \mu_1 \\ \lambda_1 + \lambda_2 &\geqslant \mu_1 + \mu_2 \\ \lambda_1 + \lambda_2 + \lambda_3 &\geqslant \mu_1 + \mu_2 + \mu_3 \\ \lambda_1 + \ldots + \lambda_n &\geqslant \mu_1 + \ldots + \mu_n \end{aligned}$$

Example 38. For example, among partitions of 3, the partition $(3) \geqslant (2+1)$ and $(2+1) \geqslant (1+1+1)$.

This is only a partial ordering. There may be some incomparable partitions. For example, the partitions $(3,3)$ and $(4,1,1)$ are not compatible. Although $4 \geqslant 3$, we also have $4+1 \leqslant 3+3$.

Problem 39. Describe this ordering in terms of the Young diagrams.

Last time, we introduced Young diagrams and Young tabloids. The Young diagrams are the boxes, and the Young tabloids are the boxes with numbers 1 through n inside of them, but among the rows it doesn't matter in what order the numbers appear.

Example 40. For the Young diagram corresponding to the partition $(2,1)$, there are three associated Young tabloids

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

For each Young diagram, we want an irreducible representation – a vector space with an action of S_n. We have some number of Young tabloids for each partition λ of n. For notation, let's list the different Young tabloids corresponding to λ as $\delta_1, \ldots, \delta_k$. These are going to be the basis elements for a vector space on which S_n will acts.

$$M_\lambda = \bigoplus_{i=1}^{k} \mathbb{C}\delta_i.$$

This will not be the final vector space on which S_n will act to give an irreducible representation, but it will contain the irreducible representation.

Example 41. Back to S_3. For the partition $3 = 2+1$, we get the vector space

$$M_{(2,1)} = \left\{ \alpha \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} + \beta \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} + \gamma \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \;\middle|\; \alpha, \beta, \gamma \in \mathbb{C} \right\} \cong \mathbb{C}^3$$

For each such vector space, we get an action of S_n. Let's illustrate with an example.

Example 42. There is a natural action of S_3 on $M_{(2,1)}$. Let $\sigma = (1\ 2)$. There is a natural action on $M_{(2,1)}$ given by

$$\sigma \cdot \begin{array}{|c|c|}\hline i & j \\ \hline k & \\ \cline{1-1}\end{array} = \begin{array}{|c|c|}\hline i' & j' \\ \hline k' & \\ \cline{1-1}\end{array},$$

where $i' = \sigma(i)$, $j' = \sigma(j)$ and $k' = \sigma(k)$.

The action of σ permutes the basis vectors as

$$\begin{aligned} \sigma \cdot \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \\ \cline{1-1}\end{array} &= \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \\ \cline{1-1}\end{array} \\ \sigma \cdot \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \\ \cline{1-1}\end{array} &= \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 & \\ \cline{1-1}\end{array} \\ \sigma \cdot \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 & \\ \cline{1-1}\end{array} &= \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \\ \cline{1-1}\end{array} \end{aligned}$$

Therefore, σ acts on any vector in $M_{(2,1)}$ by

$$\sigma \cdot \left(\alpha \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \\ \cline{1-1}\end{array} + \beta \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \\ \cline{1-1}\end{array} + \gamma \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 & \\ \cline{1-1}\end{array} \right) = \beta \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \\ \cline{1-1}\end{array} + \alpha \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \\ \cline{1-1}\end{array} + \gamma \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 & \\ \cline{1-1}\end{array}.$$

So the matrix corresponding to σ is

$$\sigma = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

In general, we get a representation $S_n \to \mathrm{GL}(M_\lambda)$ in this way. However, this representation is not irreducible, but it has an irreducible representation corresponding to λ and all other irreducible representations making up this representation correspond to Young diagrams $\mu \geqslant \lambda$. This means that by starting with the largest partition under the partial order, we can find one irrep. And at each next step, the representation decomposes as only irreps we have already found and exactly one irrep we have not yet seen.

Example 43. So the steps of our algorithm are as follows, illustrated for S_3. Here we have

$$(3) \geqslant (2,1) \geqslant (1,1,1).$$

So we begin with the representation for the largest partition, which is (3). We first start by finding $S_3 \circlearrowright M_{(3)}$, and then we find $S_3 \circlearrowright M_{(2,1)}$ and strip out parts corresponding to $S_3 \circlearrowright M_{(3)}$. And then we move on to $S_3 \circlearrowright M_{(1,1,1)}$ but remove parts equivalent to $S_3 \circlearrowright M_{(3)}$ and $S_3 \circlearrowright M_{(2,1)}$.

Now, we need to talk about the Young Tableaux. To distinguish the Young tabloids from the Young tableaux, set

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} = \delta_3 \qquad \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array} = \delta_2 \qquad \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1}\end{array} = \delta_1$$

Definition 44. The **Young Tableaux** are the different ways to put the numbers 1 through n in the different boxes of a Young diagram of size n.

Example 45. For the partition $(2, 1)$ of 3, there are six Young Tableaux.

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

These project by a map π onto the Young tabloids, as follows:

$$\begin{array}{ccc}
\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} & & \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 \\ \cline{1-1}\end{array} \\
\downarrow \pi & \swarrow \pi & \\
\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} = \delta_3 & &
\end{array}
\qquad
\begin{array}{ccc}
\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array} & & \begin{array}{|c|c|}\hline 3 & 1 \\ \hline 2 \\ \cline{1-1}\end{array} \\
\downarrow \pi & \swarrow \pi & \\
\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array} = \delta_2 & &
\end{array}
\qquad
\begin{array}{ccc}
\begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1}\end{array} & & \begin{array}{|c|c|}\hline 3 & 2 \\ \hline 1 \\ \cline{1-1}\end{array} \\
\downarrow \pi & \swarrow \pi & \\
\begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1}\end{array} = \delta_1 & &
\end{array}$$

Let t be a Young Tableaux for some λ a partition of n. Let

$$C_t \;=\; \{\sigma \in S_n \mid \sigma \text{ permutes elements in the same column }\}.$$

Example 46.

$$C_{\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}} = \{\mathrm{id}, (1\ 3)\} \subset S_3$$

We also define elements $e_t \in M_\lambda$ via

$$e_t \;:=\; \sum_{\sigma \in C_t} \mathrm{sgn}(\sigma)\pi(\sigma(t)).$$

Let's see what this is in the case of S_3.

Example 47.

$$e_{\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}} \;=\; \pi\left(\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}\right) - \pi\left(\begin{array}{|c|c|}\hline 3 & 2 \\ \hline 1 \\ \cline{1-1}\end{array}\right) \;=\; \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} - \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1}\end{array} = \delta_3 - \delta_1$$

Our irreducible representation corresponding to a Young diagram λ will be

$$\mathrm{GL}\Big(\mathrm{Span}(e_t \mid t \text{ Young tableaux of } \lambda)\Big)$$

Example 48. Let's do it again for S_3. Set

$$t = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array}.$$

Then $C_t = \{\mathrm{id}, (1\,2)\}$. We find

$$\begin{aligned} \mathrm{id} \cdot t &= \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array} \\ (1\,2) \cdot t &= \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1}\end{array}. \end{aligned}$$

Therefore, $e_t = \delta_2 - \delta_1$. Now set

$$t = \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 \\ \cline{1-1}\end{array}.$$

Then $C_t = \{\mathrm{id}, (2\,3)\}$. So

$$\begin{aligned} \mathrm{id} \cdot t &= \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 2 \\ \cline{1-1}\end{array} \\ (2\,3) \cdot t &= \begin{array}{|c|c|}\hline 3 & 1 \\ \hline 2 \\ \cline{1-1}\end{array}. \end{aligned}$$

Therefore, $e_t = \delta_3 - \delta_2$.

So the three vectors we get are $\delta_3 - \delta_1$, $\delta_2 - \delta_1$, and $\delta_3 - \delta_2$. These are not linearly independent, so we get a representation of S_3 of dimension 2.

Lecture 8 — July 10, 2015

Today we'll explicitly find all of the irreducible representations of the symmetric group for the case of S_3, and along the way illustrate the general method.

Young diagrams index the conjugacy classes of S_n, and therefore the irreducible representations. Let λ be a Young diagram, and let $\delta_1, \ldots, \delta_k$ denote Young Tabloids for λ. Set

$$M_\lambda = \mathbb{C}\delta_1 \oplus \mathbb{C}\delta_2 \oplus \ldots \oplus \mathbb{C}\delta_k.$$

The representation $S_n \to \mathrm{GL}(M_\lambda)$ contains one new irreducible representation.

Let t be a Young Tableaux for λ. Set

$$C_t = \{\sigma \in S_n \mid \sigma \text{permutes columns of } t\}$$

for example,

$$C_{\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}} = \{\mathrm{id}, (1\,3)\}.$$

The irreducible representation is GL(Span$\{e_t\}$), where

$$e_t = \sum_{\sigma \in C_t} \mathrm{sgn}(\sigma)\pi(\sigma(t)).$$

The first half of this lecture is basically contained in the following example.

Example 49. For S_3 and $\lambda = \ydiagram{2,1}$, the Young Tableaux are

$$\begin{ytableau} 1 & 2 \\ 3 \end{ytableau} \qquad \begin{ytableau} 2 & 1 \\ 3 \end{ytableau} \qquad \begin{ytableau} 1 & 3 \\ 2 \end{ytableau} \qquad \begin{ytableau} 3 & 1 \\ 2 \end{ytableau} \qquad \begin{ytableau} 2 & 3 \\ 1 \end{ytableau} \qquad \begin{ytableau} 3 & 2 \\ 1 \end{ytableau}$$

These project by a map π onto the Young tabloids, as follows:

$$\begin{tikzcd} \begin{ytableau} 1 & 2 \\ 3 \end{ytableau} \arrow[d, "\pi"] & \begin{ytableau} 2 & 1 \\ 3 \end{ytableau} \arrow[dl, "\pi"] \\ \begin{ytableau} 1 & 2 \\ 3 \end{ytableau} = \delta_3 & \end{tikzcd} \qquad \begin{tikzcd} \begin{ytableau} 1 & 3 \\ 2 \end{ytableau} \arrow[d, "\pi"] & \begin{ytableau} 3 & 1 \\ 2 \end{ytableau} \arrow[dl, "\pi"] \\ \begin{ytableau} 1 & 3 \\ 2 \end{ytableau} = \delta_2 & \end{tikzcd} \qquad \begin{tikzcd} \begin{ytableau} 2 & 3 \\ 1 \end{ytableau} \arrow[d, "\pi"] & \begin{ytableau} 3 & 2 \\ 1 \end{ytableau} \arrow[dl, "\pi"] \\ \begin{ytableau} 2 & 3 \\ 1 \end{ytableau} = \delta_1 & \end{tikzcd}$$

We found last time that

$$\text{Span}(\{e_t \mid t \text{ Young Tableaux for } \lambda\}) = \text{Span}(\delta_1 - \delta_3, \delta_1 - \delta_2).$$

There is an irreducible representation $T\colon S_3 \to \text{GL}(2, \mathbb{C})$ with character given by

$$\chi_T(\text{id}) = 2, \qquad \chi_T(1\ 2) = 0, \qquad \chi_T(1\ 2\ 3) = -1.$$

Let's write

$$\delta_1 - \delta_3 = v_1 \sim \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \delta_1 - \delta_2 = v_2 \sim \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

How do the elements of S_3 act on the vectors v_1 and v_2?

element of S_3	$v_1 = \delta_1 - \delta_3$	$v_2 = \delta_1 - \delta_2$
id	v_1	v_2
$(1\ 2)$	$v_1 - v_2$	$-v_2$
$(1\ 3)$		
$(2\ 3)$		
$(1\ 2\ 3)$	$-v_2$	$v_1 - v_2$
$(1\ 3\ 2)$		

This shows that the matrix of the element $(1\ 2)$ is given by

$$T(1\ 2) \ = \ \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix},$$

so we can confirm that $\chi_T(1\ 2) = \text{tr}(T(1\ 2)) = 0$. Similarly, the matrix for $(1\ 2\ 3)$ is

$$T(1\ 2\ 3) \ = \ \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix},$$

and as expected, $\chi_T(1\ 2\ 3) = \text{tr}(T(1\ 2\ 3)) = -1$.

Remember that $\lambda = \ydiagram{2,1}$. The representation M_λ was three dimensional! So we found a two-dimensional irreducible representation, so can we decompose M_λ into irreducibles?

Well, there's a subspace $\mathbb{C}(\delta_1 + \delta_2 + \delta_3)$, which we didn't look at before. It is invariant because for any $\alpha \in \mathbb{C}$,

$$\sigma(\alpha(\delta_1 + \delta_2 + \delta_3)) = \alpha(\delta_{\sigma(1)} + \delta_{\sigma(2)} + \delta_{\sigma(3)})$$

Switching gears – invariant theory.

We are interested about equivalence questions. Given a space X, we often consider elements of X equivalent up to action of some group G. Elements $x, y \in X$ are equivalent if there is some $g \in G$ such that $g \cdot x = y$. For example, in the discussion of conics, we found that there were three equivalence classes corresponding to double lines, two distinct lines, and circles.

Definition 50. Let $G \circlearrowright X$. The **Orbit** of $x \in X$ is the set

$$\{g \cdot x \mid g \in G\}.$$

Definition 51. An **invariant** of a space X is a function $f : X \to \mathbb{C}$ such that if $x \sim y$, then $f(x) = f(y)$.

Example 52. The **genus** of a surface is a topological invariant. We can classify compact two-dimensional surfaces, for example, by their genus.

We can instead think about invariants in terms of category theory too. Instead of a map to $\mathbb{C}$, invariants can take values in a different category. Given any category **C**, we can have as an invariant some functor $F : \mathbf{C} \to \mathbf{D}$ for some other category **D**.

Example 53. The fundamental group $\pi_1 : \mathbf{Top} \to \mathbf{Groups}$ is an invariant on topological spaces taking values in the category of groups.

Given some space X and an action $G \circlearrowright X$, we want to find all the functions which are constant on orbits, that is, all the possible invariants of the space. To that end, let R be the ring of all functions on X.

Definition 54. Let $G \circlearrowright X$, and let R be the ring of all functions on X. Then the **ring of invariant functions** on X is

$$R^G := \Big\{ f \in R \mid f(g \cdot x) = f(x) \text{ for all } x \in X, g \in G \Big\}.$$

In general, there may be infinitely many invariant functions in R^G – if f is invariant, so is $f \circ f$, and $f \circ f \circ f$, and so on. So instead of asking if R^G is finite, we ask if it's finitely generated. And even if R^G is finitely generated, we want to know if there are finitely many relations between generators? Or can we get all of the relations from finitely many? So we want to know the answer to the following questions

(1) is R^G finitely generated?

(2) is the ring of relations between generators finitely generated?

In invariant theory, there are two important theorems which answer these questions, respectively. They are given the boring names

(1) The First Fundamental Theorem of Invariant Theory;

(2) The Second Fundamental Theorem of Invariant Theory.

They are really more schema for questions to ask rather than honest-to-god theorems. Often called syzygies[2].

Example 55. Let $R = \mathbb{C}[x_1, \ldots, x_n]$ and let $S_n \circlearrowright R$ by permuting the coordinates.

$$\sigma \cdot f(x_1, \ldots, x_n) \;=\; f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

For example, with $\sigma = (1\ 2)$ and $R = \mathbb{C}[x_1, x_2, x_3]$,

$$f = x_1^2 + x_2^2 + x_2 \mapsto \sigma \cdot f = x_2^2 + x_1^2 + x_3.$$

What is the ring of invariants? It is generated by the **elementary symmetric polynomials**

$$\begin{aligned} \sigma_1 &= x_1 + x_2 + \ldots + x_n \\ \sigma_2 &= x_1x_2 + x_2x_3 + \ldots + x_{n-1}x_n \\ \vdots \quad & \quad \vdots \\ \sigma_n &= x_1x_2 \cdots x_n \end{aligned}$$

This answers the First Fundamental Theorem of Invariant Theory for $S_n \circlearrowright R$. The Second Fundamental Theorem is much easier to answer – no relations.

Why do we care? Let's switch gears for a while and talk about science (is math not science? Are we not doing science now?). To a mathematician, the following completely describes science.

Joke 56.

$$\textbf{Science} = \textbf{Numbers in boxes}$$

For any physical phenomena in an experiment, you should get the same numbers (up to units). This is an equivalence relation. The invariants are the symmetries of a system. These symmetries are very important to science, because they can help us determine the structure of physical phenomena. We can describe the relations among these ideas.

[2]worth 24 points in Scrabble!

$$
\begin{tikzcd}
 & \text{Lagrangian of your system} \arrow[dl, leftrightarrow] \arrow[dr, leftrightarrow] & \\
\text{Allowable Symmetries} \arrow[rr, leftrightarrow] & & \text{Invariants (Conservation Laws)}
\end{tikzcd}
$$

These are connected by Noether's Theorem, which is possibly the ultimate invariant theorem in Physics.

Lecture 9

July 13, 2015

Problem 57. Make a pop-up book about Schur's lemma.

Last time we talked about invariants and invariant theory. Given a space X and a finite group G acting on X, we want to find objects up to equivalence. An invariant for the action of G on X is a function $f: X \to \mathbb{C}$ that is constant on each orbit of $G \circlearrowright X$.

Given a space X, we consider the ring R of functions on X. The subring R^G is the ring of invariant functions, which is constant on each orbit (alternatively, functions on the orbit space $X /\!\!/ G$).

There are two fundamental theorems of invariant theory. The first describes if R^G is finitely generated, and the second describes the number of relations among the generators. These are not really theorems, but questions that you ask of a particular action of a group on some space.

Today we're going to talk about the alternating group acting on polynomials and find the first two fundamental theorems.

$$A_n = \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = 1\}$$

We let the alternating group act on

$$R = \mathbb{C}[x_1, \dots, x_n].$$

Of course $A_n \subseteq S_n$, so if some polynomial is invariant under the action of S_n, then it is also invariant under action of A_n.

Hence, the ring of invariants R^{A_n} still contains all of the elementary symmetric polynomials,

$$
\begin{aligned}
\sigma_1 &= x_1 + x_2 + \ldots + x_n \\
\sigma_2 &= x_1x_2 + x_2x_3 + \ldots + x_{n-1}x_n \\
&\vdots \qquad \vdots \\
\sigma_n &= x_1x_2\cdots x_n.
\end{aligned}
$$

But it also contains a new invariant. Claim that the new invariant is

$$\Delta(x_1, \ldots, x_n) \;=\; \prod_{i<j}(x_i - x_j).$$

Example 58. When $n = 3$, we have

$$\Delta = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3)$$

Notice that Δ is not invariant under S_3, because

$$(1\;2) \cdot \Delta = (x_2 - x_1)(x_2 - x_3)(x_1 - x_3) = -\Delta,$$

but it is invariant under A_3. For instance,

$$(1\;2\;3)\cdot\Delta = (1\;2\;3)\cdot(x_1-x_2)(x_1-x_3)(x_2-x_3) = (x_2-x_3)(x_2-x_1)(x_3-x_1) = \Delta$$

Notice in the above example that the action of $\sigma \in S_3$ on Δ picks up a factor of $\operatorname{sgn}(\sigma)$. This is true in general, and gives us a clue as that there is an algebraic relation between Δ and the symmetric polynomials. In particular, Δ^2 is a symmetric polynomial in R^{S_n}, because acting on Δ might pick up a negative sign, but squaring it eliminates the sign difficulties.

Example 59. For $n = 2$,

$$\begin{aligned}\Delta^2 &= (x_2 - x_1)^2 \\ &= x_2^2 - 2x_1x_2 + x_1^2 \\ &= (x_1 + x_2)^2 - 4x_1x_2 = \sigma_1^2 - 4\sigma_2\end{aligned}$$

Something similar happens for $n = 3, 4, \ldots$

Are there any other polynomial generators needed for R^{A_n}? The answer is no! This is the first fundamental theorem of invariant for $A_n \circlearrowright \mathbb{C}[x_1, \ldots, x_n]$.

Claim 59.1.

$$R^{A_n} \;=\; \mathbb{C}[\sigma_1, \ldots, \sigma_n, \Delta]$$

Proof. Let $f \in R^{A_n}$, which means that for all $\tau \in A_n$, $\tau \cdot f = f$. Our goal is to write f as

$$f = \text{ symmetric polynomial } + \Delta\,(\text{ symmetric polynomial })\,,$$

and we know that we don't need any other powers of Δ because Δ^2 is symmetric.

Let $\hat{f} = (1\;2) \cdot f$. Then we claim that $\hat{f} = (i\;j)f$ for any transposition $(i\;j)$. To show this, notice that

$$(i\;j) = (1\;2)(1\;2)(i\;j),$$

and moreover $(1\ 2)(i\ j)$ is an even permutation. So $(1\ 2)(i\ j) \cdot f = f$ because f is invariant under the action of the alternating group, and $(1\ 2)(i\ j) \in A_n$.

$$\begin{aligned}(i\ j)f &= (1\ 2)(1\ 2)(i\ j)f \\ &= (1\ 2)f \\ &= \hat{f}.\end{aligned}$$

Now we can think about $f + \hat{f}$ and $f - \hat{f}$ and rewrite

$$f = \frac{1}{2}(f + \hat{f}) + \frac{1}{2}(f - \hat{f}).$$

This is closer to what we want if one of those two terms is symmetric. Let's claim that $f + \hat{f}$ is symmetric. For $\sigma \in S_n$ an even permutation,

$$\begin{aligned}\sigma(f + \hat{f}) &= \sigma f + \sigma(1\ 2)f \\ &= f + (1\ 2)\hat{\sigma} f && \hat{\sigma} \text{ even} \\ &= f + (1\ 2)f \\ &= f + \hat{f}.\end{aligned}$$

Now that we've shown that $f + \hat{f}$ is invariant under A_n, we just need to show that it's invariant under transpositions to show that $f + \hat{f}$ is invariant under S_n. To that end, let $(i, j) \in S_n$.

$$\begin{aligned}(i\ j)(f + \hat{f}) &= (i\ j)(f + (i\ j)f) \\ &= (i\ j)f + (i\ j)(i\ j)f \\ &= \hat{f} + f.\end{aligned}$$

Therefore, $f + \hat{f}$ is a symmetric polynomial, that is, an element of R^{S_n}.

Now let's think about $f - \hat{f}$. What happens when we act on it by $(1\ 2)$?

$$\begin{aligned}(1\ 2)(f - \hat{f}) &= (1\ 2)(f - (1\ 2)f) \\ &= (1\ 2)f - (1\ 2)(1\ 2)f \\ &= (1\ 2)f - f \\ &= \hat{f} - f = -(f - \hat{f})\end{aligned}$$

Of course, there was nothing special about $(1\ 2)$, so this works in fact for any transposition in S_n. Acting on $f - \hat{f}$ by a transposition $(i\ j)$ picks up a sign. We can think of $f - \hat{f}$ as a one-variable polynomial in x_i, and when we plug in $x_i = x_j$, the polynomial is equal to its own opposite and therefore identically zero. Since x_j is a root of $(f - \hat{f})$ as a polynomial in x_i, there must be a factor of $(x_i - x_j)$ in $f - \hat{f}$.

$$(x_i - x_j) \mid (f - \hat{f})$$

Since this is true for all transpositions, and the $(x_i - x_j)$ are relatively prime (by degree considerations), we conclude that

$$\Delta \mid (f - \hat{f})$$

Hence, we write

$$f - \hat{f} = \Delta g$$

for some polynomial g.

Finally, claim that g is symmetric. For $\sigma \in S_n$, write $\sigma = (1\ 2)\tau$ where τ is even. We know that

$$\sigma \cdot (f - \hat{f}) = -(f - \hat{f}),$$

but also,

$$\begin{aligned}\sigma \cdot (f - \hat{f}) &= \sigma \cdot (\Delta g)\\ &= \sigma \cdot \Delta \sigma \cdot g\\ &= -\Delta(\sigma \cdot g),\end{aligned}$$

so we conclude that

$$-(\Delta g) = -(f - \hat{f}) = -\Delta(\sigma \cdot g),$$

so $\sigma \cdot g = g$ and g is symmetric.

Hence, we write

$$f = \frac{1}{2}(f + \hat{f}) + \frac{1}{2}(f - \hat{f}) = \frac{1}{2}(f + \hat{f}) + \frac{1}{2}\Delta g,$$

and $(f + \hat{f}), g \in R^{S_n}$ are symmetric. $\square$

Definition 60. Let $\mathrm{GL}(n)$ act on X with ring of functions R. Let $g \in \mathrm{GL}(n)$, and let $f \in R^{\mathrm{GL}(n)}$. Then f is **covariant** with weight k if

$$g(f) = (\det g)^k f.$$

Example 61. The action of $\mathrm{GL}(n)$ on $n \times n$ matrices given by

$$\phi(B) = A^T BA$$

has a covariant of weight 2 given by $\det\colon \mathrm{M}_n \to \mathbb{C}$, because

$$\det(\sigma(B)) = \det(A^T BA) = \det(A^T)\det(B)\det(A) = \det(A)^2 \det B$$

The invariants here are coefficients of $\det(\lambda I - B)$.

Given a homogeneous polynomial of degree two, $P(x, y) = ax^2 + bxy + cy^2$, a covariant under change of coordinates is $f(a, b, c) = b^2 - 4ac$. This is the discriminant of the quadratic formula! A big problem in the 19th century among German mathematicians was the following. This is what algebraic geometry was all about in that era.

Problem 62 (Gordon's problem)**.** Solve the first fundamental problem of invariant theory for change of coordinates of homogeneous polynomials of each degree and number of variables.

After many years of people trying to find algorithms to find the generators of the rings of invariants, David Hilbert solved the whole thing in about ten pages: all invariants are finitely generated, but he didn't produce an algorithm. Gordon said of this proof "Das ist nicht Mathematik, das ist Theologie." Two years later, Hilbert found an algorithm to compute all of the generators of the ring of invariants.

Much of this theory leads to moduli spaces and David Mumford's work in algebraic geometry, which we will talk about for the rest of this week.

Lecture 10 — July 14, 2015

Today we're going to give an intro to algebraic geometry in haunting similarity with the morning lecture. We will see that the notion of objects up to equivalence is very closely related to invariant functions on some space.

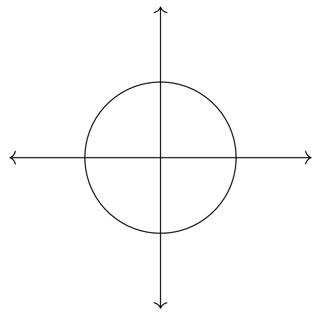
In terms of algebraic geometry, there is a close relationship between geometric objects called algebraic sets and algebraic objects: namely, commutative rings. Functions describe the world, and these commutative rings are indeed rings of functions.

We will use the ring of invariants to create a new algebraic object, which will capture the orbits of the objects of some space under action of a group.

Definition 63. Let $S \subset \mathbb{C}[x_1, \ldots, x_n]$. An **algebraic set** is the set of common zeros of polynomials in S.

$$\{(a_1, \ldots, a_n) \in \mathbb{C}^n \mid P(a_1, \ldots, a_n) = 0 \text{ for all } P \in S\}$$

Example 64. The quintessential example is the zeros of $x^2 + y^2 - 1$ in $\mathbb{C}^2$. This is the circle.



For our algebraic sets, we may assume that $S = I$ is an ideal of $\mathbb{C}[x_1, \ldots, x_n]$. The notation is

$$\mathbb{V}(I) := \{(a_1, \ldots, a_n) \mid P(a_1, \ldots, a_n) = 0,\ \forall\ P \in I\}$$

The next question we want to consider is: "What are the (polynomial) functions defined on $\mathbb{V}(I)$?", that is, functions $\mathbb{V}(I) \to \mathbb{C}$.

Example 65. What are the polynomial functions on the circle? For example, $Q(x, y) = x^3$ is an example. This is a function on all of $\mathbb{C}^2$, but also on the circle. Additionally, $P(x, y) = x^3 + x^2 + y^2 - 1$ is a polynomial function on all of $\mathbb{C}^2$. In general, $P \neq Q$ in $\mathbb{C}^2$. But on the other hand, for points $(x, y) \in \mathbb{V}(x^2 + y^2 - 1)$, $P(x, y) = Q(x, y)$. They're the same at all values on the circle! So logically, we should say that $P = Q$ on the circle. Indeed, P, Q represent the same coset in

$$\mathbb{C}[x,y] \big/ \langle x^2+y^2-1 \rangle$$

In general, the functions on $\mathbb{V}(I)$ will be described by

$$\mathbb{C}[x_1,\ldots,x_n] \big/ I.$$

Technically, we want I to be a **radical ideal**, that is, if $f^n \in I$, then $f \in I$.

Definition 66. Given an algebraic set V associated to an ideal $\mathbb{I}(V)$, the **ring of regular functions** on V is

$$\mathcal{O}_V = \mathbb{C}[x_1,\ldots,x_n] \big/ \mathbb{I}(V).$$

So we know how to go one direction: given an algebraic set V, we can produce a ring associated to it. We want to be able to do the converse as well: how can we go from knowing a commutative ring to knowing the algebraic set associated to it? This theorem is a partial answer.

Theorem 67. Let R be a commutative ring (with some additional assumptions). The points in the corresponding algebraic set are in bijection with the maximal ideals of R.

Example 68. Let $R = \mathbb{C}[x]$. What is the algebraic set associated to it? Well, the maximal ideals of R are $\langle x - a \rangle$ for $a \in \mathbb{C}$, so for each maximal ideal we get a point for each $a \in \mathbb{C}$.

For $\mathbb{C}[x, y]$, the maximal ideals are $\langle x - a, y - b \rangle$, so correspond to points in $\mathbb{C}^2$.

Definition 69. Given a ring R, the associated geometric object is the **spectrum** of R,

$$\operatorname{Spec}(R) = \{I \lhd R \mid I \text{ prime}\}.$$

The maximal ideals are called **geometric points**.

The spectrum of a ring R has an associated topology, called the **Zariski topology**, and we can do much much more with these spaces. But we'll leave it here for now.

Example 70.

$$\operatorname{Spec} \mathbb{Z} = \big\{\langle 0 \rangle, \langle 2 \rangle, \langle 3 \rangle, \langle 5 \rangle, \langle 7 \rangle, \dots\big\}$$

To summarize, for any algebraic set V, there is an associated ring $\mathcal{O}_V$. To this ring $\mathcal{O}_V$, we associate it's field of fractions

$$K_V = \left\{ \frac{f}{g} \;\middle|\; f, g \in \mathcal{O}_V,\ g \neq 0 \right\}.$$

This field is the set of rational functions on the algebraic set.

Definition 71. An algebraic set V is a **variety** if

$$\mathbb{I}(V) = \{P \mid \forall\, \vec{a} \in V,\ P(\vec{a}) = 0\}$$

is a prime ideal.

Now let's go back to equivalence problems.

Definition 72. There are two notions of equivalence between two algebraic varieties V and W.

(a) V is **isomorphic** to W if and only if $\mathcal{O}_V$ is isomorphic to $\mathcal{O}_W$;

(b) V is **birationally equivalent** to W if and only if K_V is isomorphic to K_W as a field.

While we may naïvely think that these are the same, but they are very much not the same. The first is a very rigid notion of equivalence, while birational equivalence isn't so rigid. In recent years, there have been a number of breakthroughs in the area of birational geometry called the **minimal model program**, which aims to find a canonical element in each birational equivalence class.

Example 73. $\mathbb{P}^2\mathbb{C}$ is birationally equivalent to $\mathbb{P}^1\mathbb{C} \times \mathbb{P}^1\mathbb{C}$, yet these are not isomorphic.

Projective algebraic geometry

So far, everything we've done is in affine space. But we want to work over projective space, which is where we've been thinking about equivalences. For example, there are lots of conics in affine space, but only three types in projective space.

So just replace every instance of Spec in the above with Proj, right?

Okay, we'll do it carefully. In $\mathbb{P}^n\mathbb{C}$, $(x_0, \ldots, x_n) = (\lambda x_0, \ldots, \lambda x_n)$ for any $\lambda \neq 0$. So how does it make sense to define a polynomial on $\mathbb{P}^n\mathbb{C}$?

Example 74. Let $P(x, y, z) = x^2 + y^2 - z^2$. This polynomial is poorly behaved in projective space, where $(1, 1, 1)$ and $(2, 2, 2)$ are the same point.

$$P(1, 1, 1) = 1^2 + 1^2 - 1^2 = 1$$

$$P(2, 2, 2) = 2^2 + 2^2 - 2^2 = 2$$

So instead we work with **homogeneous polynomials** (every monomial has the same degree), where at least the zero sets are well-defined.

Definition 75. Let I be an ideal of $\mathbb{C}[x_0, x_1, \ldots, x_n]$. I is **homogeneous** if it is generated by homogeneous polynomials.

Here, we think of $\mathbb{C}[x_0, x_1, \ldots, x_n]$ as a **graded ring**

$$\mathbb{C}[x_0, x_1, \ldots, x_n] \;=\; \bigoplus_{n=0}^{\infty} \mathbb{C}[x_0, x_1, \ldots, x_n]_{\deg = n},$$

where $\mathbb{C}[x_0, x_1, \ldots, x_n]_{\deg=n}$ is the polynomials of degree n.

Definition 76. Let I be a homogeneous ideal. Then the **projective algebraic set** defined by P is

$$\mathbb{V}(I) = \{(a_0, \ldots, a_n) \mid P(a_0, \ldots, a_n) = 0 \text{ for all } P \in I\}$$

Definition 77. Let R be a graded ring

$$\text{Proj}(R) \;=\; \big\{ \text{homogeneous, prime ideals of } R \big\}$$

Lecture 11 July 16, 2015

Let's recall what's been going on here. Let X be a geometric object with an action of a group G. We consider $x \sim y$ if there is some $g \in G$ such that $g \cdot x = y$. Let R be the ring of functions on X (in algebraic geometry, $\mathcal{O}_X$). Then the ring of invariants is $R^G = \{f \in R \mid f(g \cdot x) = f(x) \text{ for all } x \in X, g \in G\}$. We will relate the space of orbits $X /\!\!/ G$ to Spec(R^G).

Suppose that R^G is generated by $f_1, \dots, f_k$. There's a map $X \to \mathbb{C}^k$ defined by

$$p \mapsto (f_1(p), f_2(p), \dots, f_k(p)).$$

We hope that for each orbit of $G \circlearrowright X$, there is a corresponding single point in $\mathbb{C}^k$ under the image of this map.

Let's look at some concrete examples. Sometimes this works out really well, but sometimes it fails horribly.

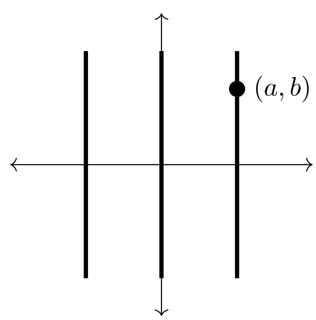
Example 78. Let $G = (\mathbb{C}, +)$. Let $X = \mathbb{C}^2$ and let $t \in \mathbb{C}$. Define an action of G on X by $t \mapsto \sigma_t \colon X \to X$,

$$\sigma_t(a, b) = (a, b + t).$$

For concreteness,

$$\sigma_3(8, 11) = (8, 14).$$

What do the orbits of this action look like?



Here, $R = \mathbb{C}[x, y]$ is the polynomial functions on X. The functions which are invariant are those that don't depend on y, so $R^G = \mathbb{C}[x]$. This ring is generated by $f(x) = x$.

We have the map $X = \mathbb{C}^2 \to \mathbb{C}$ defined by

$$(a, b) \mapsto f(a, b) = a,$$

and for each orbit we have one point on the X-axis. We see this from $X /\!\!/ G = \mathbb{C}^2/\mathbb{C} = \mathbb{C}$.

Alternatively, we could have thought about $\mathrm{Spec}(R^G) = \mathrm{Spec}\,\mathbb{C}[x]$ which corresponds to $\mathbb{C}$.

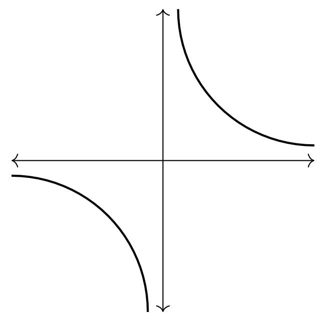
Here's another example that doesn't work out so well.

Example 79. Let $G = \mathbb{C}^\times$, and let $X = \mathbb{C}^2$. Let G act on X by $t \mapsto \sigma_t \colon X \to X$,

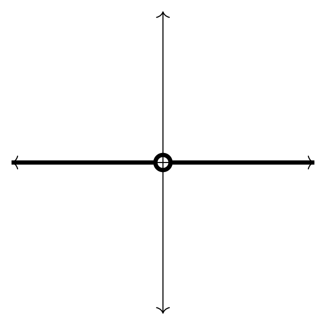
$$\sigma_t(a, b) = \left(ta, {}^{b}\!/_{t}\right).$$

There are three types of orbits.

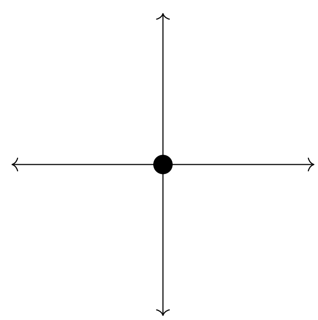
(1) If $ab \neq 0$, then we get a hyperbola centered at the origin.



(2) if exactly one of a, b is zero (that is, $0 \in \{a, b\}$), then we get one of the axes without the origin. For example, if $a \neq 0$, $b = 0$, then we get the following picture. Note that there are *two* distinct versions of this orbit: the x-axis and the y-axis. These are kind of degenerate hyperbolae.



(3) If $a = b = 0$, then we get the origin.



The ring of functions here is $R = \mathbb{C}[x, y]$ (this is also $\mathcal{O}_X = R$). An example of an invariant function on a hyperbola orbit $xy = 1$ is the function $f(x, y) = xy$. These are the only types of invariant functions.

So $R^G = \mathbb{C}[xy]$ and is generated by $f(x, y) = xy$. Our function $X \to \mathbb{C}$ is $(a, b) \mapsto ab$. Again, $\operatorname{Spec}(R^G)$ is just the complex numbers.

Unfortunately, this fails to distinguish all of the orbits. It distinguishes all of the hyperbola orbits as xy =constant, but fails to distinguish the other orbits: each gets mapped to zero.

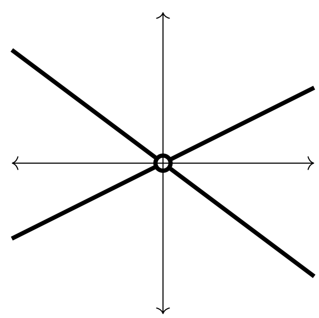
What's the failure here? It turns out to be the fact that the orbits of the second type are not closed, yet the hyperbola orbits are closed.

An invariant function on the orbit of $(0, 1)$ has to be constant, and maintain that value at the origin too because any invariant function is continuous. Thus, this invariant function has to be the same on the orbit of $(1, 0)$ – if it was not, then the invariant function would be discontinuous on the orbit of $(1, 0)$.

In general, orbits whose closures intersect will be identified.

Example 80. Let $G = \mathbb{C}^\times$, let $X = \mathbb{C}^n$ and let $t \in \mathbb{C}^\times$. Let $\sigma_t \colon \mathbb{C}^n \to \mathbb{C}^n$ be defined by $\sigma_t(a_1, \ldots, a_n) = (ta_1, \ldots, ta_n)$.

What are the orbits? Lines through the origin skipping the origin or the origin itself.



None of these orbits are closed, and yet *all* of the orbit closures intersect at the origin. So the only invariant functions are the constants!

In general, given a space X (generally algebraic), we consider the ring of regular functions on X, called $R = \mathcal{O}_X$. We have a group G acting on X. We want to find all invariants of the orbit space $X /\!\!/ G$, and identify the "bad points" on X. These are called the **nonstable points**. There are also **semistable** points, and **stable** points (which are a subset of the semistable ones).

The set of stable points X^{stable} of X is a big open set of X, and the orbit space $X^{\text{stable}} /\!\!/ G$ is well-behaved – we can distinguish all orbits within there. The set of semistable points is not too much worse, and in addition it's the closure of the stable points.

Here's a formal definition of stable.

Definition 81. A point $p \in X$ is **stable** if the orbit of p is closed and the stabilizer

$$\text{Stab}_G(p) = \{g \in G \mid g \cdot p = p\}$$

is finite.

The condition that the stabilizer is finite means that we can distinguish things nearby a point p from the point p itself.

Example 82. Going back to example 79, the stable points of $X = \mathbb{C}^2$ is the set

$$X^{\text{stable}} = \{(a, b) \mid ab \neq 0\}.$$

Lecture 12 July 17, 2015

We've been studying equivalence problems: when are things the same?

For a space X, we have an action of G on X. We consider things equivalent if they lie in the same orbit. When does the orbit space $X /\!\!/ G$ have good structure? We split this up into **stable**, **semi-stable**, and **unstable** points. Today, we're going to give a numerical criterion for the stable, semi-stable, and unstable points of $X /\!\!/ G$.

Definition 83.

$$V_{n,d} = \{ \text{ polynomials, homogeneous of degree } d \text{ in } n \text{ variables } \}$$

Example 84.

$$V_{3,2} = \{ax^2 + bxy + cxz + dy^2 + eyz + fz^2\} = \left\{ \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} \right\} \cong \mathbb{C}^6$$

We saw a few weeks ago that the change of variables on x, y, z is given by a 3×3 matrix. But changing variables also changes coefficients in $V_{3,2}$, so we get a representation

$$\rho\colon \mathrm{GL}(3) \to \mathrm{GL}(6).$$

Now we hope to understand this for any n and any d. Note that

$$\dim V_{n,d} = \binom{d+n-1}{n-1}.$$

Let's begin with $V_{2,d}$, which are single variable polynomials of the form

$$a_d x^d + a_{d-1}x^{d-1} + \ldots + a_0$$

The zero loci of these polynomials are just points on $\mathbb{P}^1\mathbb{C}$. Our space is X, and we'll restrict attention to the action of $G = \mathrm{SL}(n, \mathbb{C})$ by change of coordinates. We want to find $V_{n,d} /\!\!/ \mathrm{SL}(n, \mathbb{C})$. This is a representation $\mathrm{SL}(n) \to \mathrm{SL}\left(\binom{d+n-1}{n-1}\right)$.

Theorem 85 (Numerical Stability Criterion)**.** Mumford found criteria for points in the orbit space $V /\!/ \mathrm{SL}(n,\mathbb{C})$ to be unstable, semistable, or stable, for any action of $\mathrm{SL}(n,\mathbb{C})$ on the vector space V. Let $x \in V$.

x unstable	x semistable	x stable
$0 \in$ closure of x-orbit	$0 \notin$ closure of x-orbit	$0 \notin$ closure of x-orbit *and* $\mathrm{Stab}(x)$ finite
for all nonconstant $\mathrm{SL}(n,\mathbb{C})$-invariant f $f(x) = 0$	there is nonconstant $\mathrm{SL}(n)$-invariant f, $f(x) \neq 0$	there is nonconstant $\mathrm{SL}(n)$-invariant f, $f(x) \neq 0$ *and* $\mathrm{trdeg}(V^{\mathrm{SL}(n)}) = \dim V - \dim \mathrm{SL}(n)$

1. x **unstable**: there is a 1-parameter subgroup $\lambda\colon \mathbb{C}^\times \to \mathrm{SL}(n)$ such that the weights of x with respect to λ are all positive
2. x **semistable**: for all 1-parameter subgroups λ, not all weights are positive;
3. x **stable**: for all 1-parameter subgroups $\lambda\colon \mathbb{C}^\times \to \mathrm{SL}(n)$, there is a negative weight

Definition 86. A **1-parameter subgroup** $\lambda\colon \mathbb{C}^\times \to \mathrm{SL}(n)$ is a map such that, up to change of basis,

$$\lambda(t) = \begin{bmatrix} t^{r_1} & & & \\ & t^{r_2} & & \\ & & \ddots & \\ & & & t^{r_n} \end{bmatrix}$$

Example 87. In $\mathrm{SL}(2)$, a one parameter subgroup is

$$t \mapsto \begin{bmatrix} t & 0 \\ 0 & 1/t \end{bmatrix}.$$

Definition 88. A **weight** of λ with respect to $\vec{x} \in V$ is

$$\mu(x,\lambda) = \text{ weight } = -\max\{-r_i : x_i \neq 0\},$$

where

$$\lambda(t) = \begin{bmatrix} t^{r_1} & & & \\ & t^{r_2} & & \\ & & \ddots & \\ & & & t^{r_n} \end{bmatrix} \qquad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Example 89. Let's go back to our action of $\mathrm{SL}(2)$ by change of coordinates in $V_{2,4}$. A vector is, for example,

$$a_4x^4 + a_3x^3y + a_2x^2y^2 + a_1xy^3 + a_0y^4$$

and we have a change of coordinates defined by the matrix λ

$$\lambda \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} t & 0 \\ 0 & 1/t \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} tx \\ y/t \end{bmatrix}.$$

This polynomial transforms as

$$a_4t^4x^4 + a_xt^2x^3y + a_2x^2y^2 + \frac{a_1}{t^2}xy^3 + \frac{a_0}{t^4}y^4,$$

so the representation of the matrix is

$$\begin{bmatrix} t^4 & 0 & 0 & 0 & 0 \\ 0 & t^2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & t^{-2} & 0 \\ 0 & 0 & 0 & 0 & t^{-4} \end{bmatrix}$$

This is still in $\mathrm{SL}(5)$. What are the weights?

$$\mu((1,t,1,1,1),\lambda) = -4$$

$$\mu((1,1,1,1,0),\lambda) = -2$$

$$\mu((1,1,1,0,0),\lambda) = 0$$

Using the criteria for stability, we can figure out when things are stable, semi-stable, or unstable. It turns out to be unstable when there is a root of very high multiplicity. The cutoff for "very high" is half of the degree of the polynomial. Using the theorem, we can calculate exactly which points are stable and which are unstable, which is what made Hilbert famous.

What to take from these lectures

This is all very classical stuff. It was known to Hilbert and his contemporaries. If we had another week, we would go on to talk about the modern version of this stuff, which leads into Lie groups, Lie algebras, and much more algebraic geometry. This is the beginnings of

We began with equivalence problems, and then moved on to representation theory, invariant theory, some algebraic geometry, and now finally we have talked about moduli spaces. Hopefully you've got some idea of what's going on in the general picture.