

Math 250B: Commutative Algebra

Alex Fu

Spring 2023

These are transcripts of Professor Richard Borchers' lectures for Math 250B: Commutative Algebra in spring 2023. As they are reformatted from shorthand (and Prof. Borchers tends to lecture rather quickly), they are very rough and unpolished. Remarks enclosed in (parentheticals) or [brackets] are very often but not always my own, inserted for the sake of personal clarity or elaboration. All mistakes are my own.

1 2023-01-17

Welcome to commutative algebra. All rings are commutative rings with unity.

Pure commutative algebra is dry and technical; this is really a service course for two or three other areas of math.

1. Algebraic number theory. For example, consider the Gaussian integers $\mathbb{Z}[i]$, the integers with a square root of -1 adjoined. We can ask various question about it, such as whether or not it has unique factorization.
2. Algebraic geometry. Algebraic geometers study algebraic varieties, e.g. $y^2 = x^3 - x$, which is some sort of elliptic curve, or $z^2 = x^2 + y^2$, some sort of cone [like an hourglass]. For a variety, we have a coordinate ring, consisting of all functions on the variety; simply take all polynomial functions on the plane (or space or such), and quotient out by the functions that vanish on the variety, e.g. $\mathbb{k}[x, y]/(y^2 - x^3 + x)$, or $\mathbb{k}[x, y, z]/(z^2 - x^2 - y^2)$. These are commutative rings.

Conversely, we can reconstruct the variety from this ring. Studying varieties is the same as studying coordinate rings; we can ask the same questions we did like unique factorization, e.g. $\mathbb{k}[x, y]/(y^2 - x^3 = x)$ lacks unique factorization.

How exactly do we reconstruct a variety from a ring? One way is to take a point on the variety, which gives a homomorphism from $R \rightarrow \mathbb{k}$ by evaluation. So points are identified with $\text{Hom}(R, \mathbb{k})$. Or, kernels have a homomorphism correspond to certain ideals [slightly unintelligible reconstruction of what Prof. Borchers said]. So, there's a big dictionary between commutative rings on one hand and algebraic varieties on the other hand.

These are two main ways commutative algebra is used. There is a third one, slightly old:

3. Invariant theory, which was a really great deal in the second half of the 19th century, up until Hilbert came along and solved nearly every problem in it. [:skull:]

Take something like an icosahedron, look at its rotations and reflections, a group of order I that acts on \mathbb{R}^3 and hence any polynomial function $\mathbb{R}[x, y, z]$. Ask: which elements are fixed by the action of I ? These are the **ring of invariants**, some sort of algebra over \mathbb{R} , and contains 1 and \mathbb{R} rather trivially. Setting the icosahedron at the origin, treating all points as rotations, it obviously contains the degree two invariants $x^2 + y^2 + z^2$ as well, which are preserved under rotations.

There are some more subtle and harder to see invariants, of degree 6 and 10. The ring of invariants turns out to be the polynomial ring generated by $x^2 + y^2 + z^2$ and these two slightly mysterious degree 6 and 10 elements.

Take G acting on V , and ask for the ring of invariants. Fundamental questions: Can you find a finite collection generating all invariants? And all relations between this finite generating set? So, the main problem of invariant theory is:

Given G acting on \mathbb{k}^n , is the algebra of invariants finitely generated?

This is quite complicated; sometimes it is, sometimes it isn't. Hilbert showed that most [if not all] of the rings people wanted to work with are finitely generated. In the first week, we will cover Hilbert's solution to this problem. In fact, he conjectured that the ring of invariants is always finitely generated, but decades later [McGarten? inaudible, likely actually] Nagata found a weird counterexample where it wasn't.

Some examples in invariant theory.

Example 1. Take the orthogonal group $O(n)$ acting on \mathbb{R}^n . The obvious invariant is $x_1^2 + \cdots + x_n^2$, trivially invariant more or less by definition, since orthogonal transformations preserve squared norm distance. And in fact the only invariants are polynomials of these, $\mathbb{R}[x_1^2 + \cdots + x_n^2]$. So this has a very simple solution to the fundamental question, and we don't have to worry about relations, since there's only one generator.

Example 2. From Math 250A: take $G = S_n$ acting on \mathbb{k}^n by commuting the coordinates in the most obvious way possible, so G acts on $\mathbb{k}[x_1, \dots, x_n]$. The invariants are the symmetric polynomials; recall the elementary simple polynomials

$$e_1 = \sum_{i=1}^n x_i, \quad e_n = \prod_{i=1}^n x_i, \quad e_k = \sum_{I \subseteq \{1, \dots, n\}, |I|=k} x^{(I)}.$$

[The last sum is over all possible k tuples.] And in fact all invariants are just $\mathbb{k}[e_1, \dots, e_n]$. Furthermore, there's no relations between any of these! So again, it's just a polynomial ring.

This happens to be a misleadingly easy case. The reason it's easy: S_n is a reflection group. Anything that fixes a codimension 1 hyperplane, e.g. exchanging x_i with x_j , is a reflection. For a reflection group, a theorem of Chevalley (–Shephard–Todd) says the algebra of invariants is then a polynomial ring. There's a sort of converse to this: if a finite group is not a reflection group, then the algebra of invariants is not only not a polynomial ring, but nearly always extraordinarily complicated. There are only a few cases where we are able to write down ring of non-reflection groups in a reasonable way.

Example 3. $G = A_n$ acting on \mathbb{k}^n has the obvious invariants of e_1, \dots, e_n , since A_n is a subgroup of S_n . There's one more: the [discriminant?] Vandermonde determinant $\Delta = \prod_{i < j} (x_i - x_j)$, which changes sign if we transpose i and j , but is invariant under the alternating group of even permutations.

We observed in 250A that all invariant polynomials are polynomials in e_1, \dots, e_n and Δ , so the algebra of invariants is $\mathbb{k}[e_1, \dots, e_n, \Delta]$, but this is not a polynomial ring: Δ^2 is invariant under the symmetric group, so it must be some polynomial in e_1, \dots, e_n . The algebra is the quotient $\mathbb{k}[e_1, \dots, e_n, \Delta]/(\Delta^2 - \text{discriminant}(e_1, \dots, e_n))$, where the discriminant is a special case of the resultant.

The algebra of invariants is finitely generated, but not finitely generated as a ring, e.g. if \mathbb{k} is not finitely generated. Don't confuse algebras and rings.

A relation between generators like $\Delta^2 - \text{discriminant}(e_1, \dots, e_n)$ is an example of a **syzygy**, a Greek word meaning *yoked together*, as if two oxen connected when pulling a plow, connecting the various invariants. "Syzygy is also notable because it's almost the longest word that doesn't use any vowels. There's also rhythms, which is even longer."

If we've only got two variables, the invariant $\Delta^2 = b^2 - 4c$, where $b = x_1 + x_2$, $c = x_1 x_2$, so $\Delta = (x_1 - x_2)$.

Example 4. A slightly more complicated example: $G = \mathbb{Z}/n\mathbb{Z}$, generated by g , acting on \mathbb{C}^2 . We take the 2D complex vector space so we don't have to worry about the characteristic dividing n . Let $g(x, y) = (\zeta x, \zeta y)$, where $\zeta^n = 1$ is the primitive n th root of unity $\zeta = e^{2\pi i/n}$, the most obvious oscillator.

Writing out a basis for all polynomials,

$$\begin{array}{ccccccc} 1 & x & x^2 & x^3 & \cdots & x^n \\ y & xy & x^2y & \cdots & x^{n-1}y & \\ y^2 & \cdots & & & & \\ \vdots & & & & & \\ y^n & & & & & \end{array}$$

So under the action, the first diagonal gets multiplied by 1, the second diagonal (x, y) multiplied by ζ , and so on. The diagonal which connects x^n and y^n , i.e. the basis elements homogeneous of degree n , are all invariants. And $x^{2n}, x^{2n-1}y, \dots$ can all be attained as polynomials in the lower-order invariants, so the algebra of invariants is $\mathbb{K}[x^n, x^{n-1}y, \dots, y^n]$, quotiented out by some relations between these.

To avoid confusion, we will call these c_0 up to c_n . There are lots of relations for $\mathbb{K}[c_0, \dots, c_n]$: $x^i y^j x^k y^\ell = x^a y^b x^c y^d$ whenever $i + k = a + c$; and $c_0 c_2 = c_1^2$, $c_1 c_4 = c_2 c_3$, and so on, are all relations in the quotient. So there are a huge number of syzygies. Now, we come to another question in invariant theory:

1. Is the algebra of invariants finitely generated?
2. Are the syzygies of invariants finitely generated?

Is there a finite collection of relations, such that all relations can be obtained from them? [In some sense “finitely presented” I suppose.] But there might be relations between the relations there if they’re not independent, and we might have higher-order syzygies, relations between lower-order ones, so we again ask

3. Are second-order syzygies finitely generated?

⋮

- ∞. Is there a limit to the order of syzygies?

Does this process eventually stop? Finding the invariants is hard, and syzygies even worse, and syzygies between syzygies between syzygies is a nightmare. Hilbert actually answered all these questions, positively: all are finitely generated, and he even showed that this process comes to an end after a while.

This is a bit of a complicated mess. We should think about how to reformulate this in terms of abstract algebra in a slightly clearer way. So we have a polynomial ring mapping onto a ring of invariants A , generated by some finite set, the kernel of which is going to be some sort of ideal (of syzygies).

$$0 \rightarrow I \rightarrow \mathbb{K}[i_1, \dots, i_n] \rightarrow A \rightarrow 0.$$

“How do you spell syzygy?”

We should be a bit careful about this term *finitely generated*, which has two different meanings. One is as an algebra over \mathbb{K} , i.e. every element is a polynomial in some basis set of invariants. On the other hand, we’re not asking whether the syzygies are finitely generated as an algebra, rather as an ideal, where we want every element to be of the form $r_1 g_1 + \dots + r_n g_n$, g_i is a fixed set of syzygies, and the r_i in the polynomial ring $R = \mathbb{K}[i_1, \dots, i_n]$. The ring of invariants finitely generated as an algebra, and the syzygies finitely generated as an ideal.

Now higher-order syzygies. Call $I = I_0$, where an ideal is a special case of a module over a ring. To be finitely generated as an ideal is to ask the question: is there a map

$$R^n \rightarrow I \rightarrow 0?$$

The generators of R^n are syzygies!; the relations between syzygies is the kernel of this map,

$$R^{n_1} \rightarrow R^{n_0} \rightarrow I \rightarrow 0.$$

Can we find a map from a finitely generated free module onto the kernel of this map? So R^{n_1} has a basis of second order syzygies. And asking again and again, repeating,

$$\dots \rightarrow R^{n_2} \rightarrow R^{n_1} \rightarrow R^{n_0} \rightarrow I \rightarrow 0.$$

This is an exact sequence called the *free resolution* of I . We can also ask if it terminates at some point:

$$0 \rightarrow R^{n_k} \rightarrow \dots \rightarrow R^{n_0} \rightarrow I \rightarrow 0.$$

Sometimes it does, but there are counterexamples too. If yes, then it forms a *finite* free resolution of I . And in commutative algebra, for any module, not just an ideal, we can ask if it has a finite free resolution; it's great as a complete description of what the module is. In fact, finite free resolutions originally came from higher-order syzygies in invariant theory.

Invariant theorists didn't usually work with finite groups; they were more interested in e.g. special linear groups, the main example being:

Example 5. Binary quantics, meaning *two variables* and *of some unknown degree*. If the degree is known, we would call it quadric, cubic, quartic, quintic, sextic, heptic, etc. If you're not quite sure what the degree is, just call it a quantic. A typical binary quantic is

$$a_n x^n + a_{n-1} x^{n-1} y + \cdots + a_1 x y^{n-1} + a_0 y^n.$$

What is the invariant theory of binary quantics? We need a group and a vector space, $G = \mathrm{SL}_2(\mathbb{k}) = \{(a, b; c, d)\}$ and $V = \mathbb{k}^{n+1}$ with basis a_0, \dots, a_n . So $(a, b; c, d) \in G$ acts on (x, y) by

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}.$$

And so it maps on the quantics, taking a binary quantic to

$$a_n (ax + by)^n (cx + dy)^0 + \cdots,$$

which is going to be some other quantic,

$$x^n (a_n a^n, a_{n-1} a^{n-1} c + \cdots) + x^{n-1} y (\cdots).$$

Thus, the action of G takes the coefficient a_n to some complicated linear combination of the a_i 's, so it forms an action on V .

Problem: find the invariants of $\mathrm{SL}_2(\mathbb{k})$ acting on $\mathbb{k}[a_0, \dots, a_n]$. And you think it's a horrible mess! The action is already hard to write explicitly, and here we want finite generating set of invariant polynomials in a_0, \dots, a_n .

This was the main problem tackled in the 19th century, finally solved by Gordon, by incredibly complicated arguments, that this is finitely generated. I was going to show you how complicated this argument is; Sylvester's calculations took two papers, the first for order up to 10, the second doing order 12. Pages of horrendous numerical coefficients; these aren't actually invariants, but polynomials that tells the dimension of the number of invariants of each degree. Incredibly hairy.

For example, ternary cubics, three variables, of degree 3. Sylvester worked in the 19th century, long before computers, so all these calculations were done by hand. "People must've had a lot of spare time in those days."

$$a_{3,0,0} x^3 + a_{2,1,0} x^2 y + a_{2,0,1} x^2 z + \cdots + a_{0,3,0} y^3 + \cdots + a_{0,0,3} z^3$$

which has 10 coefficients [10 being the number of nonnegative partitions of 3, or some stars-and-bars type argument, the number of $i + j + k = 10$, $i, j, k \geq 0$].

$\mathrm{SL}_3(\mathbb{k})$ acts on x, y, z , so it acts on the ring $\mathbb{k}[a_{0,0,3}, \dots, a_{3,0,0}]$. What does the ring of invariants look like? It's too complicated to write down explicitly, but fortunately Sturmfels has done it for me. The ring is unusually easy in this case, a polynomial ring. There are two invariants of degrees 4 and 6; the latter takes up most of this page and half of the next. "Pass this around, see why I am not going to write down invariants explicitly." And this is still a simple case, 3 and 3. The general case of higher-degree variables, higher quantics, are even worse than this.

Note: the two books passed around were *The Collected Mathematical Papers of James Joseph Sylvester, Volume 2* and *Algorithms in Invariant Theory* by Sturmfels. Here was the usual 5-minute break.

—

Finding the ring of invariants is hard, finding the syzygies harder, and higher-order syzygies worse and worse. [There are many dependencies on each other.] And we want to show that these are finitely generated. The way Hilbert solved this problem: showing that syzygies are finitely generated isn't actually all that bad if the ring of invariants is finitely generated. So he proved this implication.

The algebra of invariants is $\mathbb{k}[g_1, \dots, g_n]/I$, where g_i are the basic invariants or generators, and I is the ideal of syzygies.

The following is the fundamental theorem of 250B; its statement forgets entirely about invariant theory.

Theorem 1 (Hilbert's theorem).

Any ideal of a polynomial ring $\mathbb{k}[x_1, \dots, x_n]$ is finitely generated.

Ideals of $\mathbb{Z}[x_1, \dots, x_n]$ are also finitely generated. First, time for some review; this theorem is "so important, but most of you probably forgot what I said about it."

Proposition 1 (Noether).

If all ideals of a ring R are finitely generated, so are all ideals of $R[x]$. (By induction, so it is for $R[x_1, \dots, x_n] = R[x_1, \dots, x_{n-1}][x_n]$.)

The hypothesis says R is **Noetherian**. This idea was not first invented by Noether [I might have been pronouncing her name wrong before; pronounced like "note her" in lecture], but by Hilbert; however, Noether was the one who pointed out this property.

An explanation of the difference between being finitely generated as an algebra (without 1) and as an ideal:

Example 6. Take $R = \mathbb{k}[x, y]$ and the ideal $(y) = Ry$. We draw a picture of R by writing down a basis of R :

...				
y^3	...			
y^2	xy^2	...		
y	xy	x^2y	...	
1	x	x^2	x^3	...

And the ideal I is the region containing everything except the bottom row. The IDEAL I is generated by y , while the algebra I [over \mathbb{k}] is — when I said that all rings and algebras have a 1, I was lying; this one doesn't — not finitely generated, e.g. (y, xy, x^2y) doesn't get x^3y . It's not true that subalgebras of polynomial rings are finitely generated as algebras, but this is true for ideals.

I'm quickly going to sketchily recall parts of the proof of Hilbert's theorem from 250A. The following are equivalent:

- a R is Noetherian, i.e. all ideals are finitely generated.
- b Any nonempty set of ideals has a maximal element.
- c Any strictly increasing chain of ideals $I_1 \subset I_2 \subset \dots$ is finite. (ACC)

(The equivalence of b and c is true for any poset; it has nothing to do with rings and ideals.)

Proof. Suppose R is Noetherian; we will show that $I \subseteq R[x]$ is finitely generated. Look at $I_0 \subset I_1 \subset I_2 \subset \cdots$, where each I_n is the leading term of polynomials in I of degree $\leq n$. These are ideals of R , and $I_{n-1} \subset I_n$ because we can always multiply by x . Thus $I_n = I_{n+1} = \cdots$ eventually stabilizes, as otherwise ACC is violated. For a finite set of generators for I , take

- 0. A finite set of polynomials of $\deg \leq 0$ whose leading coefficients generate I_0 ,
- 1. $\deg \leq 1$ generate I_1 , and so on,
- \vdots
- m. There are an infinite number of possible degrees, but fortunately we can now stop at step m .

This is a finite set and generates I . (The key for finiteness in this proof: isolating coefficients and using the ascending chain condition.) To check that this does generate I , we induct on the degree of f in I . We can find some linear combination of elements in the generating set with the same leading coefficient as f . If its degree is $\geq m$, multiply by a suitable power of x to get the same leading coefficient. Subtracting the leading term, we make the degree of $(f - \text{linear combination})$ one smaller. \square

So, the first theorem of Hilbert: if the ring of invariants is finitely generated, then the ideal of syzygies is automatically finitely generated. Next, the harder part: the ring of invariants is finitely generated.

2 2023-01-19

Recall we showed that R Noetherian implies $R[x]$ is Noetherian, and so $R[x_1, \dots, x_n]$ is Noetherian as well. More generally, if R is a Noetherian ring, $R \subseteq S$, and S is finitely generated as an R -algebra, then S is Noetherian.

(As usual, we warn that being “finitely generated” is a bit ambiguous: here, it means that we can find $s_1, \dots, s_n \in S$ such that all elements in S are polynomials in $R[s_1, \dots, s_n]$.) As an exercise, we check that this follows quite easily from the fact that $R[x_1, \dots, x_n]$ maps onto S .

Proposition 2 (The surjective image of a Noetherian ring is Noetherian).

Let $\varphi: R \rightarrow S$ be a surjective ring homomorphism. If R is Noetherian, then so is S .

Proof. Let I be an ideal in S . We know that $\varphi^{-1}(I)$ is an ideal in R , and by hypothesis $\varphi^{-1}(I) = (a_1, \dots, a_n)$ as an ideal (i.e. as an R -module, such that every $j \in \varphi^{-1}(I)$ equals some $r_1 a_1 + \dots + r_n a_n$). Then $I = (\varphi(a_1), \dots, \varphi(a_n))$, because φ is a ring homomorphism: every $i \in I$ has j such that

$$i = \varphi(j) = \varphi(r_1)\varphi(a_1) + \dots + \varphi(r_n)\varphi(a_n).$$

In other words, every ideal in S is finitely generated as well, so S is Noetherian. \square

If the algebra of invariants is finitely generated as an algebra, then the syzygies are finitely generated (as an ideal). These are relations between g_i , the kernel of the map from the polynomial ring $\mathbb{k}[x_1, \dots, x_n]$ to the ring of invariants. So, as a special case of the fact that every ideal of a polynomial ring is finitely generated, this shows that first-order syzygies are finitely generated. What about higher-order syzygies?

We know that there are free modules [a free resolution]

$$\dots \rightarrow R^{n_2} \rightarrow R^{n_1} \rightarrow R^{n_0} \rightarrow I,$$

where I is the ideal of higher-order syzygies, but are all the n_i finite? It's not much harder to show that yes, all of these are finite. Recall from last lecture that if we try to write them down, they're a ghastly mess; even invariants are complicated. But Hilbert's theorem shows that all of these also finite. Recall:

Definition 1 (Noetherian modules).

M is a Noetherian R -module if any or all of the following equivalent conditions hold.

1. Every nonempty set of submodules has a maximal element.
2. Any strictly increasing chain of submodules $M_0 \subseteq M_1 \subseteq \dots$ is finite.
3. Any submodule is finitely generated (as an R -module).

These conditions are equivalent in the same way that they were for rings. [For example, conditions 1 \equiv 2 for posets in general.] In fact, R is Noetherian as ring iff R is Noetherian as an R -module, and I is an ideal iff it is a submodule of the R -module R . [This is a basic part of the dictionary between the languages of rings and modules.]

Now, suppose we have the following exact sequence of R -modules.

$$0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$$

Something that happens quite often common for modules: something in the middle $[B]$ has some property iff both the left side and right side $[A$ and $C]$ do. Here, B is Noetherian iff A and C are Noetherian.

Proof. The forward direction is trivial (A is a sub- R -module of B , e.g. A an ideal of $B = R$, and C is a surjective image of B ; we have shown both cases before.)

Conversely, suppose $M \subseteq B$. [Our proof is the same as a proof of the rank-nullity theorem.] The image of M is a submodule N of C , and similarly for $M \cap A$. (Strictly speaking, the isomorphic image of A .) Pick a finite set $\subseteq M$ whose images generate N in C , and a finite set of generators for $M \cap A$, using the fact that A, C Noetherian implies the same for $M \cap A, N$; combine the two sets by union, and this union generates M . \square

Corollary 1.

Easy corollary: any finitely generated module over a Noetherian ring is Noetherian.

1. R^n is Noetherian by induction, taking the exact sequence $0 \rightarrow R^{n-1} \rightarrow R^n \rightarrow R \rightarrow 0$.
2. Any finitely generated module M is a quotient of some R^n .

Now, even higher-order syz are finitely generated. More generally, suppose M is a finitely generated module over a Noetherian ring. Then we can find an exact sequence

$$R^{n_2} \rightarrow \dots \rightarrow R^{n_0} \rightarrow M \rightarrow 0,$$

all n_i finite. It's kind of obvious how to do this. If

$$R^{n_k} \rightarrow R^{n_{k-1}} \rightarrow \dots \rightarrow M \rightarrow 0,$$

all finite, R^{n_k} is Noetherian, so $\ker(R^{n_k} \rightarrow R^{n_{k-1}})$ is finitely generated, by say n_{k+1} elements. So we can extend to

$$R^{n_{k+1}} \rightarrow R^{n_k} \rightarrow \dots.$$

Thus, the free resolution of any Noetherian module by finite-dimensional free modules is finite. This solves the question of whether high-order syzygies are finitely generated using a much more general theory, which quite frankly had nothing to do with invariant theory.

Now, some examples and non-examples of Noetherian rings. Any ring finitely generated over something Noetherian is Noetherian, so for something that isn't Noetherian, we need something not finitely generated. This makes it obvious how to find non-Noetherian rings: for instance, polynomials $\mathbb{k}[x_1, x_2, \dots]$ in infinitely many variables. We can see quite explicitly that all three conditions fail:

1. (x_1, x_2, \dots) , all polynomials with zero constant term, is not finitely generated.
2. $(x_1) \subset (x_1, x_2) \subset (x_1, x_2, x_3) \subset \dots$ is an infinite strictly ascending chain of ideals.
3. Borrowing, the same example above is a set of ideals without a maximal element.

So far, we have considered increasing chains and maximal elements; what about decreasing chains $I_1 \supset I_2 \supset \dots$? A result about posets is that every decreasing chain finite iff every nonempty set of ideals has a minimal element. If being Noetherian is equivalent to the ACC (ascending chain condition), then

Definition 2 (Decreasing chain condition; Artinian rings).

A ring is **Artinian** if every strictly decreasing chain of ideals is finite (the descending chain condition, or DCC).

Examples of Artinian rings are much rarer; e.g. \mathbb{Z} is not Artinian by $(2) \supset (4) \supset (8) \supset \dots$, or many other such chains, but is Noetherian. In fact, many think being Noetherian and Artinian are dual to each other; but Artinian implies Noetherian [!].


Let us look at the following chain of rings.

1. Polynomials $\mathbb{R}[x]$.
2. Entire functions (holomorphic on \mathbb{C}), $\exp(x)$.
3. Germs of analytic functions on \mathbb{R} . (That is, functions analytic near 0; we don't care what they are away from 0, undefined or not.)
4. Smooth functions on \mathbb{R} .
5. Germs of smooth functions.
6. Formal power series $\mathbb{R}[[x]]$.

These all seem quite similar, having something to do with functions on reals.

- 1 is a subset of 2;
- 2 is almost a subset of 3,
- 3 is sort of a subset of 4, or rather there's a natural map from 3 to 4;
- Similarly, there are natural maps from $4 \rightarrow 5$ and $5 \rightarrow 6$.

Which are Noetherian? Some are, some aren't. 1, 3, and 6 are Noetherian; 2, 4, and 5 are not.

1. Notice that $R[x]$ is a PID; not only are all ideals finitely generated, but in fact principal, generated by one element.
2. Entire functions admit an infinite increasing chain of ideals: let I_1 be the functions vanishing on $\{1, 2, 3, \dots\}$, I_2 those vanishing on $\{2, 3, \dots\}$, and so on. Real polynomials are a dense subring of entire functions, but one is Noetherian, and by a seemingly small change, the other stops being Noetherian.
3. For germs of analytic functions, the only ideals are (0) , (x) , (x^2) , (x^3) , \dots
4. [I don't recall any comments on this particular item]
5. For germs of smooth functions, let I be the ideal of all functions vanishing to infinite order at 0. You can in fact have function vanish to ∞ order at 0 that aren't actually 0, e.g. $e^{-1/x^2} \cdot \mathbb{1}_{x \neq 0}$, a traditional example in analysis which looks like  [I'm a \LaTeX artist what can I say] and has a zero of order ∞ at the origin.

As an exercise, show that this ideal is not finitely generated. "The reason I'm making this an exercise is because I can't actually remember how to prove it."

6. For formal power series with coefficients in a Noetherian R :

Proposition 3.

If R is Noetherian, so is $R[[x]]$, and by extension $R[[x_1, \dots, x_n]]$.

So, $R[x]$ and $R[[x_1, \dots, x_n]]$ are both Noetherian, but lots of intermediate rings are not, e.g. the set of all power series converging on the complex plane is not Noetherian.

Proof. We can try to copy Hilbert's proof for $R[x]$, since polynomials are a bit like power series. A bit of a problem: we want to define I_n generated by leading terms, but power series have no highest-order term. So the key idea: look at the lowest-order nonzero coefficient.

Put I_0 the ideal of elements a_0 for $a_0 + a_1x + \dots \in I$, where I is an ideal of $R[[x]]$, and define I_1, \dots similarly. Note that $I_0 \subseteq I_1$: we can multiply by x just as before. However, we can't quite fully copy the proof of $R[x]$; as an exercise: why does the proof of $R[[x]]$ not work for $R[x]$? If you figure out why that is, you will understand this [lowest-order term] proof. \square

Warning: while the result holds for submodules, the *subring* of a Noetherian need not be Noetherian. For example, $k[x_1, x_2, \dots] \subseteq$ the field of quotients $\mathbb{k}(x_1, x_2, \dots)$, which is certainly Noetherian as a field. Any ring with no zero divisors is automatically contained in a field, which is Noetherian, so being subring doesn't tell you much. [Being an ideal or submodule is more meaningful.]

Warning: since every ideal of $\mathbb{k}[x_1]$ is generated by 1 element, we might guess: every ideal of $\mathbb{k}[x_1, x_2]$ is generated by 2 elements. But this is wrong! In fact, there's no bound on the number of elements needed. [In some sense, a 2D plane allows too much freedom and is too large, as we will see.] E.g., in

$$\begin{array}{ccccccc} & & & & & & \dots \\ & & & & & & y^3 & \dots \\ & & & & & & y^2 & xy^2 & \dots \\ & & & & & & y & xy & x^2y & \dots \\ & & & & & & 1 & x & x^2 & x^3 & \dots \end{array}$$

take the ideal generated by the four elements on the diagonal (x^3, x^2y, xy^2, y^3) [which includes everything "above or to the right" of this diagonal]. This cannot be generated by less than four elements: it's obvious that no proper subset of $\{x^3, x^2y, xy^2, y^3\}$ will generate it, but maybe there's some clever way to come up with some complicated polynomials to get all of those.

Nope. Take I/I^2 , which is a vector space over \mathbb{k} of dimension 4. Any set of generators of ideal I has to generate $I/(y^4, y^3x, y^2x^2, x^3, x^4)$ as a vector space over \mathbb{k} .

Finally, getting back to Hilbert's theorem:

Theorem 2 (Hilbert's theorem).

The algebra of invariants of a finite group on \mathbb{C}^n are finitely generated.

Proof. First, the ring of invariants $A \subseteq \mathbb{k}[x_1, \dots, x_n] := R$, where R is a graded ring

$$R = \bigoplus_n R_n, \quad R_n R_m \subseteq R_{m+n}.$$

In this case, R_n are the homogeneous polynomials of (combined or total) degree n . [In fact, I believe polynomial rings are the motivating example for the definition of a graded ring.]

Thus, the ring of invariants is also graded; let $A = A_0 \oplus A_1 \oplus \dots$. We know that ideals are finitely generated, so let I be the ideal generated by $A_1 \oplus A_2 \oplus \dots$. I is finitely generated as an ideal by say i_1, \dots, i_n . We can assume that these are homogeneous and invariant.

Idea: show that i_1, \dots, i_n generate the algebra A .

This turns out to be rather difficult to prove. “The main reason it’s difficult to prove is that it’s not actually true.” For example, take the polynomial ring

$$\begin{array}{ccccccc} & & & & & & \dots \\ & & & & & y^3 & \dots \\ & & y^2 & & xy^2 & & \dots \\ y & & xy & & x^2y & & \dots \\ 1 & & x & & x^2 & & x^3 & \dots \end{array}$$

We might take the algebra A to have basis of [everything above which contains a y , as well as 1], and take the ideal $I = (y)$ [everything above the bottom row]. Then $i_1 = y$ [the “bottom left corner of the ideal”] is a generator for I , but this does not generate the graded subalgebra A .

So, the fact that i_1, \dots, i_n generates A fails for some A , but works for the algebra of invariants. What special property do they have? Rings of invariants have **Reynolds operators** ρ , which we will talk about after the break. \square

Remark. The first question is *who is Reynolds?* He didn’t work with invariants; in fact, he wasn’t even a pure mathematician. He worked in fluid dynamics, and he had no interest whatsoever in pure mathematics. You may know him as the guy of the famous Reynolds number, for any flow. And he introduced something called a Reynolds operator, originally in the context of fluid dynamics, for example if you have a fluid flow varying with time.

These are mind-numbingly complicated; they satisfy the Navier–Stokes equation, a monstrosity / mightily complicated nonlinear differential equation. And his idea was to replace the flow at a point with the *average* [over time] of the flow at that point. So applied mathematicians just assume that all integrals converge, just

$$\lim_{t \rightarrow \infty} \frac{1}{2t} \int_{-t}^t \text{flow } dt.$$

And if you’re an applied mathematician, all limits exist, no problem. If you’re a pure mathematician, it’s hard to tell if the limit exists. But, [in any case,] this is the average of the fluid flow under time-translation, which is the group action of \mathbb{R} . And the same could be defined for any group action.

Definition 3 (Reynolds operator).

More generally, if G acts on V , then $\rho: v \in V \mapsto$ the average of v under G , i.e. $\frac{1}{\text{vol}(G)} \int_{g \in G} g(v)$. This might not exist in general, but for a finite group,

$$\rho(v) := \frac{1}{|G|} \sum_{g \in G} g(v)$$

is well-defined. Well, almost. We got to divide by the order of G , so this is well-defined if $|G| \neq 0$ in field we are working with. In \mathbb{C} , this is a positive integer, no problem. But in a field of characteristic p , if p divides $|G|$ is problematic, so we really want $p \nmid |G|$.

Proposition 4 (Properties of ρ).

1. Linearity. $\rho(u + v) = \rho(u) + \rho(v)$, and $\rho(\lambda v) = \lambda \rho(v)$ for all scalars λ .
2. “Invariance.” $\rho(u) = u$ if u is fixed by G . [This is clear; we can permute the terms of any finite sum, and addition is commutative.]

3. Preserves multiplication. Since $\rho(v)$ preserves addition, taking $V = \mathbb{k}[x_1, \dots, x_n]$ and working in algebra, we might guess that $\rho(uv) = \rho(u)\rho(v)$. This is false; there's no particular reason why this should be true. A minor variation that is correct: $\rho(uv) = \rho(u)\rho(v)$ if v is fixed by G , or $\rho(v) = v$. (And this is clear.)

Let's draw a picture of what ρ is doing:

$$\rho: A \leftarrow \mathbb{k}[x_1, \dots, x_n],$$

where the ring of invariants A is a subset of $\mathbb{k}[x_1, \dots, x_n]$. ρ is not a homomorphism of rings in general, but it *is* a homomorphism of A -modules from condition 3 (it preserves the multiplication of elements in A .)

So, given ρ , we get a splitting of an exact sequence of A -modules:

$$0 \rightarrow A \rightarrow \mathbb{k}[x_1, \dots, x_n] \rightarrow \mathbb{k}[x_1, \dots, x_n]/A \rightarrow 0,$$

a sequence which does not split in general, but does with Reynolds operators. That is, $\mathbb{k}[x_1, \dots, x_n] = A \oplus \ker(\rho)$. And **splitting** will turn out to be the key fact to proving Hilbert's theorem.

Proof continued. Suppose that i_1, \dots, i_n generate the ideal I (the ideal generated by $A_1 \oplus A_2 \oplus \dots$). Let us show that i_1, \dots, i_n generate the algebra A . Pick $a \in A$ (homogeneous). Put

$$a = r_1 i_1 + \dots + r_n i_n, \quad r_i \in R,$$

which is because $a \in I = (i_1, \dots, i_n)$ (generated as an ideal). But problem. The r_i are in R , but we really want them in A . How do we get them in A ? Just apply the magical ρ operator. [Recall that $\rho: \mathbb{k}[x_1, \dots, x_n] \rightarrow A$]

Applying ρ , by linearity,

$$\begin{aligned} a &= \rho(a) = \rho(r_1)\rho(i_1) + \dots + \rho(r_n)\rho(i_n) \\ &= \rho(r_1)i_1 + \dots + \rho(r_n)i_n. \end{aligned} \quad (*)$$

Note that the generators $i_k = \rho(i_k)$ are invariant and homogeneous, and $a = \rho(a)$ is invariant as well. Well, $\rho(r_k)$ are elements of A , and furthermore has degree $\leq \deg a$, as $\deg i_1, \dots, i_n \leq 1$. So by induction on the degree, $\rho(r_1), \dots, \rho(r_n)$ are in the algebra generated by i_1, \dots, i_n , i.e. $A = \mathbb{k}[i_1, \dots, i_n]$. \square

This is *magical*: the key point of Hilbert's proof is this one line [the starred line above] which wipes out thousands of pages of complicated calculations in invariant theory.

So, this proves finiteness for finite groups in characteristic = 0. Before leaving this topic, I wanted to look at some other cases.

First, the proof works whenever we have the Reynolds operator. E.g. $G = \mathbb{R}$, acting on \mathbb{R}^2 by $x \mapsto \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}$ has no Reynolds operator. The space of things fixed by \mathbb{R} is just one-dimensional. For the exact sequence

$$0 \rightarrow \mathbb{R} \rightarrow \mathbb{R}^2 \rightarrow \mathbb{R} \rightarrow 0,$$

G acts trivially on the left- and right-hand \mathbb{R} , but acts nontrivially on G , so it does not split. We cannot split the space \mathbb{R}^2 into a direct sum of a space of invariant vectors and another space fixed by G .

We could even have G finite, for example $\mathbb{Z}/p\mathbb{Z}$ acting on \mathbb{k}^2 , where $\mathbb{k} = \mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. Then G acts on \mathbb{k} by the same map as before, and we still have no ρ . So for a finite group of finite nonzero characteristic, we still don't get a Reynolds operator.

But, later in the course, we will see that it is Noetherian / Noether's insight [too ambiguous to tell what was said from my notes, sorry; but likely the latter]: if G is a finite group, even in characteristic > 0 , then the ring of invariants is finitely generated. However, this requires some extra ideas about [principal ideal] domains.

Hilbert wasn't working with finite groups; he was working with groups like $SL_2(\mathbb{R})$ or such. How do we deal with this case?

Example 1 (Compact groups).

The same works if G is a compact group, e.g. the orthogonal group $O_n(\mathbb{R})$. In this case,

$$\rho(v) = \frac{1}{\text{vol}(G)} \int_G g(v).$$

The integral of a continuous function over a compact group is well-defined.

Anything for finite groups can quite often extend to compact groups. But the groups that Hilbert worked with were not actually compact, so next, we will deal with non-compact groups.

3 2023-01-24

There were three ingredients in the proof of Hilbert's finiteness (of invariants) theorem:

1. $\mathbb{K}[x_1, \dots, x_n]$ is Noetherian.
2. G has a Reynolds operator ρ .
3. Everything is *graded*.

For G finite (in characteristic 0), the Reynolds operator was defined. But Hilbert's original theorem wasn't for finite groups; we can extend to compact groups by the remarks above, where $\text{vol}(G) = \int_{g \in G} 1$. For compact things, integration behaves very much like summation.

What about the non-compact $\text{SL}_2(\mathbb{R})$? As we recall, for the invariants of binary quantics $a_0 x^n y^0 + \dots + a_n x^0 y^n$, $G = \text{SL}_2(\mathbb{R})$ takes $x \rightarrow ax + by$, $y \rightarrow cx + dy$, and thus acts on a_0, \dots, a_n [via matrix multiplication]. So G acts on $\mathbb{C}[a_0, \dots, a_n]$.

Gordon was the guy who proved finiteness of invariants for this particular case. Not clear whether G has a ρ [abbreviation for has a Reynolds operators]; in fact it does. For this, we need a fourth idea:

4. Weyl's unitarian trick.

First, $\text{SU}_2(\mathbb{R})$, the group of all unitary transformations, IS compact, so it has ρ . Second, $\text{SU}_2(\mathbb{R})$ and $\text{SL}_2(\mathbb{R})$ have the "same complexification," i.e. the "same" finite-dimensional representations, which means that $\text{SL}_2(\mathbb{R})$ has ρ — the two groups are almost identical.

What does it mean for two groups to have the same complexification? Strictly speaking, there is no such thing as the complexification of a linear group in general. For the next few minutes, we have an extended comment on Lie algebras and Lie groups.

$$\text{group } \text{SL}_2(\mathbb{R}) \leftrightarrow \text{Lie of } \text{SL}_2(\mathbb{R}) \rightarrow \text{Lie of } \text{SL}_2(\mathbb{C}) \leftarrow \text{Lie of } \text{SU}_2(\mathbb{R}) \leftrightarrow \text{group } \text{SU}_2(\mathbb{R})$$

["Lie of" means "Lie algebra of."] Label the objects above 1, 2, 3, 4, 5 for simplicity.

The Lie algebra of $\text{SL}_2(\mathbb{R})$, which we could write $\mathfrak{sl}_2(\mathbb{R})$, is all traceless matrices, i.e. all entries $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ of trace 0. If $\det(a, b; c, d) = 1$, then $\det(I + \varepsilon A) = \det(1 + \varepsilon a, \varepsilon b; \varepsilon c, 1 + \varepsilon d) = 1 + \text{tr } A\varepsilon + \varepsilon^2 \dots$. So the elements of trace 0 are precisely the elements "close to the identity" (under \det).

And in the same way, we have $\text{SU}_2(\mathbb{C})$ and its Lie algebra. Now, both 2 and 4 \rightarrow 3 by complexification, i.e. by taking the tensor product $- \otimes \mathbb{C}$, by which they become a complex Lie algebra, the same complex Lie algebra of $\text{SL}_2(\mathbb{C})$.

Now, 1 \leftrightarrow 2, and likewise 4 \leftrightarrow 5, have the same finite-dimensional representations; the representations of Lie algebras are closely related to representations of groups. 2 \rightarrow 3, and 4 \rightarrow 3, have the same complex representations, where representations of $\text{SL}_2(\mathbb{C})$ preserve complex structure. (We add this condition that preserves complex multiplication because 3 is a vector space over \mathbb{C} , even though 2 isn't.)

Putting it all together, $\text{SL}_2(\mathbb{R})$ and $\text{SU}_2(\mathbb{R})$ have the same finite-dimensional representations, and ρ carries over. So Hilbert's finiteness theorem also works for $\text{SL}_2(\mathbb{R})$. There's nothing special about it; works for any semisimple Lie group.

Warning: the finiteness of dimension here is absolutely critical; this totally fails for infinite-dimensional representations. The representation theory of these two groups looks completely different. $\text{SU}_2(\mathbb{R})$ still has a Reynolds operator of

infinite dimensions, while $\mathrm{SL}_2(\mathbb{R})$ does not. What goes wrong is that $\exp(A)$ is well-defined if A is finite-dimensional, but it usually doesn't make sense in infinite dimensions. And when going from a Lie algebra to Lie group, we need to take \exp ; the exponential of a matrix doesn't converge in general is where everything breaks down.

—

So Hilbert proved the finiteness of invariants for almost any semisimple Lie group. The fact that ideals of polynomial rings are very general, independent of the group, and he seemed to think that we could push past semisimple Lie groups.

Hilbert's 14th problem.

Given any group G acting on a finite-dimensional space, is the algebra of invariants finitely generated?

Hilbert had a list of 23 problems in 1900. The 14th problem was actually solved unexpectedly by Nagata in 1958, the answer being no. [Note: Borchers says naGAta, but the wrong emphasis is on the wrong syllable.] Nagata gave an example; the group itself wasn't actually that complicated, \mathbb{C}^{13} .

So Hilbert essentially showed this is true for many groups in characteristic 0; works for reductive groups. Roughly, reductive means that there are no normal subgroups of the form \mathbb{C}^n . To define reductive, I would have to spend 10 minutes explaining algebraic groups, but just this is just a passing comment.

Think of the easiest possible groups: $(\mathbb{C}^n, +)$, it's hard to think of easier. But these are the groups that cause the most problems, cause all of the trouble. Reductive groups include all semisimple Lie groups. In the theory of Lie groups, abelian groups are easy, while simple Lie groups are very complicated, but in invariant theory, semisimple groups are easy to deal with, while abelian ones give lots of trouble.

Haboush and Nagata showed that finiteness also works for characteristic > 0 if the group is reductive. This is a lot harder: there is usually no ρ . The standard counterexample in this case is $(1, *; 0, 1)$. Haboush found a sort of nonlinear substitute of ρ , which is a linear map. You can't find a $\mathrm{char} p$ linear map, but there are more complicated nonlinear maps. Haboush's theorem is much too difficult to prove, and doesn't involve not much commutative algebra. An easy special case of this, which was done or mentioned earlier, is the following.

Theorem 3 (Noether).

If G is a finite group in $p > 0$, then the algebra of invariants is finitely generated.

The ring of invariants is not usually finitely generated.

For the proof, recall the following:

1. Any finitely generated module over a Noetherian ring is Noetherian. [We used this fact to prove that all higher-order syzygies were finitely generated; we reuse it here.]
2. Various sorts of finite generation, i.e. its three different meanings: as module, as algebra, and as field.

A quick example to distinguish them: $\mathbb{k}[x]$ is f.g. over \mathbb{k} as a module and algebra; $\mathbb{k}(x)$, the rational functions, is f.g. over \mathbb{k} as a field but not a module or algebra.

3. Not quite recalling since it hasn't come up yet, but *integral extensions*.

Let $R \subseteq S$ be rings. $s \in S$ is *integral* over R if s is a root of a monic (with leading coefficient 1) polynomial $x^n + a_{n-1}x^{n-1} \cdots + a_0 = 0$ in $R[x]$ with s as a root. [In other words, s is “algebraic” over R , but this definition is for rings in general, not just fields.]

As an example, in $\mathbb{Z} \subseteq \mathbb{Q}$, the only elements of \mathbb{Q} integral over \mathbb{Z} are integers. This is where the name “integral” comes from; integers in rationals are those that satisfy monic polynomials in \mathbb{Q} , left as an exercise.

Key point:

Proposition 5.

s is integral over R iff $R[s]$ is a finitely generated module over R . Obviously, s integral implies that $R[s]$ is spanned by $1, s, s^2, \dots, s^n$, where $s^n = -a_{n-1}s^{n-1} - \cdots - a_0$, so it has a finite basis by powers of s .

Conversely, if $R[s]$ is finitely generated as a module, it is generated by polynomials p_1, p_2, \dots of bounded degree $\leq n$ for some n . But then $s^n \in R[s]$, so it is in the module generated by p_1, p_2, \dots . We write

$$\begin{aligned} sp_1 &= a_{1,1}p_1 + a_{1,2}p_2 + \cdots \\ sp_2 &= a_{2,1}p_1 + a_{2,2}p_2 + \cdots \\ &\vdots \end{aligned}$$

and taking this system of equations A , the matrix $A - sI$ has a nontrivial solution, thus determinant 0, which gives a degree n equation for s with leading coefficient 1.

(Minor addendum from 2023-01-26 on the missed out explanation: the system of equations $Ap = 0$, where $A = (a_{i,j})$ and $p = (p_i)$, implies that $A^{\text{ad}}Ap = 0$. The adjoint A^{ad} times the original matrix is $A^{\text{ad}}A = (\det A)I$. In other words, $\det(A)p = 0$, or $\det(A)p_i = 0$ for $i = 1, \dots, n$.)

An application: if $R \subseteq S$, the integral closure of R , the set of elements in S integral over R , forms a ring. If R is Noetherian, it is enough for $R[x]$ to be contained in a finitely generated module (as any submodule is also finitely generated). So, if R is Noetherian, the set of elements integral over R is also a ring. “ R being Noetherian is not necessary, but I’m blanking out over the proof of the non-Noetherian case.”

If s_1, s_2 are integral, then s_1, s_2 are contained in finitely generated modules M_1, M_2 . s_1s_2 is contained in M_1M_2 , finitely generated by the generating set $m_{1,-}m_{2,-}$. And similarly, $s_1 + s_2$ is also contained in $M_1 + M_2$. So the sums and products of integral elements are always integral.

A special case you’ve come across in number theory: for $R = \mathbb{Z}$ and $S = \mathbb{C}$, the integral elements are the algebraic integers, and these form a ring. Not obvious from the definition; e.g., $\sqrt{2} + \sqrt[3]{2} + \sqrt[5]{2}$ satisfies a monic polynomial in $\mathbb{Z}[x]$, but if you try and write it down explicitly, you will give up. The polynomial has degree 30, and the coefficients are a real pain, so it doesn’t work very well. The abstract definition of lying in a finitely generated module is much easier to deal with.

With the background results done, let us proceed with the proof of Noether’s theorem.

Proof of Noether’s theorem. Finite group G acts on a vector space with basis x_1, \dots, x_n , so acts on $\mathbb{k}[x_1, \dots, x_n] = S$. Let S^G be the algebra of invariants. Look at the following polynomial: $p_i = \prod_{g \in G} (x - gx_i)$.

First of all, x_i is a root by taking $g = 1$. Secondly, the coefficients are in S^G , as the product is over all of G . Thirdly, the leading coefficient is 1. Put $R =$ the algebra generated by all coefficients of p_1, \dots, p_n . We have $R \subseteq S^G$ [see secondly] $\subseteq S$.

Now, some relations between rings.

1. $(R \leftrightarrow S.)$ S is integral over R , and finitely generated as an algebra by x_1, \dots, x_n , all integral elements.
2. $(R \leftrightarrow S.)$ S is finitely generated as a module over R , generated by a finite number of integral elements. For instance, the module is spanned by all elements of the form $x_1^{i_1} \cdots x_n^{i_n}$, where $i_k \leq \deg$ of (monic) polynomial of x_k .
3. R is finitely generated as an algebra over \mathbb{k} ; there are a finite number of coefficients of a finite number of polynomials.
4. S^G is finitely generated as an R -module. It is contained in the finitely generated module S , and R is Noetherian by Hilbert's theorem. With the fact that S is finitely generated as an algebra, S is Noetherian, and submodules of finitely generated modules are finitely generated.
5. (Follows from 3 and 4.) S^G is finitely generated as a \mathbb{k} -algebra. Take the generators for R as a \mathbb{k} -algebra, and generators for S^G as \mathbb{k} -module. If these generate as a \mathbb{k} -module, then they generate as a \mathbb{k} -algebra.

And that's what we wanted to prove: S^G is finitely generated as a \mathbb{k} -algebra.

□

As a summary, we have this result of Hilbert / Haboush / Nagata.

Theorem 4.

The following are equivalent for an algebraic group $G \subseteq \mathrm{GL}_n(\mathbb{k})$.

1. G is reductive (has no normal subgroups \mathbb{k}^n).
2. If G acts on a finitely generated algebra over \mathbb{k} , the *ring* of invariants is a finitely generated algebra.
3. If G acts on \mathbb{k}^n and has a fixed vector $v \neq 0$, we can find some polynomials fixed by G with different values on $0, v$.

3 essentially says there are enough polynomials fixed by \mathbb{k} in order to separate points fixed by G . A sort of nonlinear ρ , a polynomial of degree 1. In char 0, we can get away with polynomials of degree 1, but we need higher degrees otherwise, which becomes a lot harder to define. 2 implies 1 involves Nagata's counterexample, and we have shown 1 implies 2.

This is a really complete solution to Hilbert's 14th problem, specifies almost exactly which groups have this finiteness property.

Actually, there's one major problem in Hilbert's proof.

1. Is the algebra of invariants finitely generated? (Often yes.)
2. Can we calculate the invariants?
3. Efficiently?

Say I gave you the icosahedral action on \mathbb{R}^3 and asked for a finite basis over the invariants. You suddenly realize, even after listening to all my lectures, that you have no idea how to find the invariants. The problem is that Hilbert's proof is nonconstructive. Let's see why.

His proof of the ideal being finitely generated is nonconstructive; we can keep finding generators for the ideal, but we never really know when we've found enough. E.g. $\mathbb{K}[x_1, \dots, x_n]$, which contains the ring of invariants $A = A_0 \oplus A_1 \oplus A_2 \cdots$. To explicitly find generators for A : First, find invariants of A_k for fixed k . This is in principle easy. $A_k \subseteq$ polynomials of $\deg k$, the fixed subspace of a finite dimensional vector space under some operators, a routine linear algebra problem.

In practice, there's some problems. The main one is that these spaces have very large dimension. For example, for ternary cubics, polynomials in 10 variables; invariants of degree 50 or 100; polynomials of degree 100 in 10 variables; vector space of a million or even billions of dimensions. The spaces you run into become very soon so big, that even on modern computers you can't even do algebra. Linear algebra in a million dimensions means you get 10^{12} coefficients, lots of lots of matrix calculations, starts taking hours and hours.

And that's just ternary cubics, a rather small case. With 50 or 100 variables, it's out of reach even for computers. So you can find invariants of various degrees, provided you have a big budget.

Another problem: Find the invariants of A_1, A_2, A_3 , so on. When do we stop??? The proof doesn't give any help with this. We need a bound for the degrees of the sets of generators of invariants. Without a bound, Hilbert's proof is nonconstructive and doesn't give an algorithm.

This caused quite a stinker near the end of the 19th century. There were some questions about whether Hilbert's proof was even a solution or not, since there was no way, even in principle, of finding the set of invariants. A famous quote, sometimes attributed to Gordon, was that "This is not math; this is theology." Just claiming the existence of something and not constructing it, even in principle.

(There's no evidence that Gordon said this as anything except for a joke. Gordon was not against Hilbert; he actually thought very highly of Hilbert's work, and continued on his work later. Gordon was a referee of Hilbert's original paper, but blocked its publication, though for very good reason: there were some gaps in Hilbert's proof.)

This problem was fixed by Hilbert, who found bounds on these degrees. However, these bounds are so large as to be impractical to find.

Example. If G is a finite group acting on \mathbb{K}^n , the ring of invariants is *Cohen–McCauley*, which we will cover near the end of course — it is a finitely generated FREE module over a [finite] polynomial ring. You get some idea about how complicated a ring of invariants is by the dimension of the free module. Huffman–Sloane worked out some examples of lower bounds of the dimension of the free module: for \mathbb{K}^{24} , the dimension of this module has to be at least 205679393714995200, for just a particular case, Conway's group.

There's still a lot of work to do with computer calculations of invariants. People in the 19th century did a lot of easy cases; if you go too far, the cases are horrendous even for computers, but there's a lot of work to do in the middle, to make computer algebra more efficient. Several people using computers often got rather embarrassed: they took hours to find rings of invariants, and discovered that some guys in the 19th century did it by hand. Invariant theorists had a lot of smart tricks for working out rings of invariants.

We finish our discussion of invariant theory with one application of Hilbert's finiteness theorem: the construction of moduli spaces.

Example. The classification of elliptic curves. Any elliptic curve is given by a cubic in \mathbb{P}^2 , $a_{3,0,0}x^3 + a_{2,0,1}x^2y + \cdots + a_{0,0,3}z^3$. The problem is different polynomials can correspond to the same elliptic curve; a change of variables by $\mathrm{SL}_3(\mathbb{C})$ just gives the same elliptic curve. Take the set of polynomials, more or less isomorphic to \mathbb{P}^7 , and quotient it out by the action of $\mathrm{SL}_3(\mathbb{C})$. And if we can take this quotient, we can find the space classifying elliptic curves.

This leads to the general problem of defining the quotient variety/group, of an algebraic variety given by a subset of equations quotiented out by a group. This is in general a rather complicated operation, but one way to do it is to ask what is the ring of functions on this? This should be the functions on the variety invariant under G . And this is finitely generated by Hilbert's theorem.

If you take an algebraic geometry course, any algebra over a field, with no zero divisors, finitely generated, gives you another algebraic variety. No zero divisors follows from being over a field [may have misheard this], so the key point is to prove finite generation. From this, we can construct moduli spaces of various geometric objects we want to classify.

Next, some ways of visualizing rings.

4 2023-01-26

There are at least three methods of visualizing rings, which we will cover in the next two lectures.

1. Draw a point for each element of the ring.
2. Draw a point for each basis element of the ring.
3. Draw a point for each prime ideal of the ring.

The third is by far the most powerful and useful, but also the most mysterious, not at all clear at first.

Obvious examples. The prototypical examples: \mathbb{Z} as the dotted line, \mathbb{R} as a line, \mathbb{C} as the complex plane. If you've done Math 185, the entire course is about how to visualize rings like that. $\mathbb{Z}/5\mathbb{Z}$ is five points around the circle in the shape of a regular pentagon.

Or, the Gaussian integers $\mathbb{Z}[i] \subseteq \mathbb{C}$, which overlaps with 250A. It's difficult to precisely separate 250A and B. These form a square lattice, $x + yi \rightarrow (x, y)$. This has unique factorization into primes; recall that Euclidean ring implies PID implies UFD. Recall that Euclidean rings have a Euclidean algorithm for division-remainder: given a and $b \neq 0$, we can write $a = bq + r$ such that $|r| < |b|$ for the Euclidean norm or absolute value $|\cdot| : R \rightarrow \mathbb{Z}^+$. [I believe should be $\mathbb{R} \rightarrow \mathbb{N}$.]

Define $|m + ni| = m^2 + n^2$, the square of the usual absolute value. (Squaring makes it an integer.) We want $a/b = q + r/b$, where $|r/b| < 1$ and $q \in \mathbb{Z}[i]$, and $a/b \in \mathbb{Q}[i] \subseteq \mathbb{C}$ as a quotient of two Gaussian integers. Well, this [a stronger statement] just says that every complex number is distance < 1 from some Gaussian integer, which is kind of geometrically obvious.

Draw an open disk of radius 1 about each Gaussian integer $m + ni$; these cover the complex plane. Unique factorization is a rather complicated and nonobvious question, but we've reduced it to a completely obvious geometric fact. So looking at \mathbb{C} geometrically is a powerful way to see they've got unique factorization.

Example. Beyond $\mathbb{Z}[i] = \mathbb{Z}[\sqrt{-1}]$, we can also look at $\mathbb{Z}[\sqrt{-2}]$, which stretches out the lattice vertically. The complex plane is also covered by open disks of radius < 1 , so also a UFD. $\mathbb{Z}[\sqrt{-3}]$ is where things go wrong; the "center point" $(-1 + \sqrt{-3})/2$ (and all of its translated copies) that is distance 1 from all lattice points is not covered.

So the obvious thing breaks down. But this doesn't prove that $\mathbb{Z}[\sqrt{-3}]$ is not a UFD. Still, it's not a UFD: $2 \times 2 = (1 + \sqrt{-3})(1 - \sqrt{-3}) = 4$. We see that this only just fails, closed unit disks would cover it but not open disks, yet this is enough to make the whole argument fail. [The reason for failure is the Pythagorean theorem $(\frac{1}{2})^2 + (\frac{\sqrt{-3}}{2})^2 \geq 1$, where the lattice is stretched out by a factor of $\sqrt{-3}$ vertically.]

$\mathbb{Z}[\sqrt{-3}]$ is not a PID. What do non-principal ideals look like? PIs are $aR = (a)$; from complex analysis, multiplication is rotation by its argument and scaling by its modulus, so these have the "same shape" as R , rotated and magnified, but still the same. R looks "rectangular," and aR preserves "rectangularity," [relative orientation / position / arrangement between points,] even if not location after rotation and scaling.

[aR is a simple transformation, left multiplication by an element, a structure-preserving homomorphism on the additive group by definition of a ring, so the shape is kept.]

Non-principal ideal: $(2, 1 + \sqrt{-3}) = m + n(1 + \sqrt{-3})$. The picture of this ideal is the checkerboard containing the origin, or the moveset range of a bishop starting at 0, i.e. the points $0, 2, 1 + i\sqrt{3}, -1 + i\sqrt{3}$, etc., a sort of triangular or (sheared) parallelogram lattice in which one can only jump around "diagonally" or by 2's. Not rectangular, and thus non-principal. We could prove that is non-principal by doing half a page of algebraic calculations, but here it is

immediately obvious.

An easy way to make it into a PID: $\mathbb{Z}[\sqrt{-3}] \subseteq \mathbb{Z}[\omega]$, where $\omega = \frac{1+\sqrt{-3}}{2}$, $\omega^2 + \omega + 1 = 0$, $\omega^3 = 1$. This is a UFD since disks of radius 1 cover \mathbb{C} . We took the points that were outside the disks, made them elements of the ring, and even added disks around those new points, so obviously will cover the plane.

Example. PID that is not Euclidean. These are quite difficult to find; there are loads of UFDs that are not Euclidean, but finding PIDs is hard. Integers and polynomial rings are all Euclidean. The simplest is $R = \mathbb{Z}[\frac{1+\sqrt{-19}}{2}]$. The plane is not covered by unit disks; well, it's not even for $\frac{1+\sqrt{-15}}{2}$, but it turns out to not be a PID. First, the easy part: a proof that R is not Euclidean.

Proof. Suppose R is. Look at the values of $|r|$: $0, a_1, a_2, \dots$, where a_1 is the smallest possibility and a_2 the next smallest. We can check that if $|r| = a_1$, then r is a unit; the remainder must be 0. If $|r| = a_2$, then every element of R/r is represented by 0 or a unit (represented by some element x , $|x| < |r| = a_2$). We can now check that for $R = \mathbb{Z}[\frac{1+\sqrt{-19}}{2}]$, if $r \neq 0$ or unit, then $R/(r)$ has at least 4 different elements.

The number (≥ 4) is $|r|^2$. We can check that it is never 2 or 3. "When I say check, I mean you can check this, and I'm not gonna check this because it's an easy exercise." Next thing to check: the only units are ± 1 . Another "easy exercise." So $\{0, \text{units}\}$ has size 3, and thus R is not Euclidean, as it has no elements other than 0, units such that all elements of $R/(r)$ are represented by 0 or units.

There are plenty of rings that are not Euclidean for the obvious absolute value, but is Euclidean for other Euclidean norms. But this proof shows that R cannot be Euclidean for any norm. \square

Next, to show that R is a PID [and thus a UFD].

Proof. We need to understand possible shapes of ideals. It's not obvious where to start. First, notice any ideal I is some lattice in $\mathbb{C} \simeq \mathbb{R}^2$, and we generate it by 2 elements ω_1, ω_2 . Any lattice or discrete surface spanning \mathbb{R}^2 is the free abelian group of linear combinations of two elements or basis vectors, so $I = \{m\omega_1 + n\omega_2\}, m, n \in \mathbb{Z}$.

To pin things down a bit more, an extra property: I is closed under multiplication by the element $x = \frac{1+\sqrt{-19}}{2}$, 'cause it's an ideal. What extra information this gives: $x\omega_1 = A\omega_1 + B\omega_2$, $x\omega_2 = C\omega_1 + D\omega_2$ (i.e. $(A, B), (C, D)$ in the (ω_1, ω_2) coordinate system), for integers A, B, C, D . So x is a linear transformation of \mathbb{R}^2 with trace $2\Re(x)$, the trace of complex multiplication, and determinant $|x|^2$, depending on what definition you're using. The trace is $1 = A + D$, the determinant $5 = AD - BC$.

And there are far too many variables, so we get rid of ω_1, ω_2 . Put $\tau = \omega_2/\omega_1$, so $x = A + B\tau$, $x\tau = C + D\tau$. Eliminate x . Well, we know its value, but we still want to eliminate the $\sqrt{-19}$, so this is $(A + B\tau)\tau = C + D\tau$, a quadratic equation for τ .

$B\tau^2 + (A - D)\tau - C = 0$. For psychological reasons, it's much easier to write this as $a\tau^2 + b\tau + c = 0$, where $a = B, b = A - D, c = -C$. If you've got a quadratic equation, the first thing you look at is its discriminant, $d = b^2 - 4ac = (A - D)^2 + 4BC = (A + D)^2 - 4(AD - BC) = -19$. This -19 is the same as the -19 up there.

So the ideal has the shape of $m + n\tau$, where $a\tau^2 + b\tau + c = 0$, $b^2 - 4ac = -19$. Well, there are an awful lot of quadratic equations with discriminant 19, so to cut them down a bit, take ω_1, ω_2 to be the shortest and next shortest (that is not a multiple of the previous, i.e. not linearly dependent) vector in the lattice.

The first implies that $|\tau| \geq 1$, since $|\omega_2| \geq |\omega_1|$, and the second implies that $\Re(\tau) \leq \frac{1}{2}$. Looking at it, consider the parallelogram lattice where $\omega_1 \in \mathbb{R}$ is to the right of the origin 0, and ω_2 is above and to the right of 0. The projection of ω_2 onto the x -axis must be within the vertical boundaries of $\pm \frac{1}{2}\omega_1$. [Otherwise, the point in the fourth quadrant is closer to 0 than ω_2 .] And $\Re(|\omega_2|) \leq \omega_1$.

So might as well enforce these two conditions. τ is in the region above the unit circle and within the thin vertical strip of radius $\frac{1}{2}$ about the y -axis, i.e. the famous fundamental domain of $SL_2(\mathbb{Z})$ acting on the upper-half plane. [I may have misheard the last few words.]

Now, $|\tau| \geq 1$ implies that $|a| \leq |c|$, $a\tau^2 + b\tau + c = 0$, and $\Re(\tau) \leq \frac{1}{2}$ implies that $|b| \leq |a|$, where we can take $a > 0$ (otherwise, just flip all the signs). So $|b| \leq |a| \leq |c|$, $b^2 - 4ac = -19$, $a, b, c \in \mathbb{Z}$ the equations to be solved.

We finally reduced to a case where we can reduce the solutions. The conditions imply that $3a^2 \leq -19$, $a^2 \leq 19/3 \approx 6$, so $a = 1$ or 2 , and $|b| \leq a$ and has the same parity, which means that $b = \pm 1$ or $0, \pm 2$. And $c = 5$ or no solutions in the two cases for a .

So τ is a root of $\tau^2 \pm \tau + 5 = 0$, and the solutions are $(\pm 1 \pm \sqrt{-19})/2$. These all give the ideal $R = \mathbb{Z}[x]$, as we're just changing the sign a bit. So all ideals are principal. \square

And that gives you some idea of how hard it is showing that something is a PID that's not Euclidean. I'll leave it as a couple exercises to try:

1. Show that $\mathbb{Z}[\frac{1+\sqrt{-23}}{2}]$ has exactly three types of ideals. In fact you find that τ are roots of the following, $\tau^2 + \tau + 6 = 0$ — but you also get 2 more — $2\tau^2 + \tau + 3 = 0$, $2\tau^2 - \tau + 3 = 0$, the shape of 3 different ideals.
2. If you're feeling ambitious, a slightly more complicated example: show that $\mathbb{Z}[\frac{1+\sqrt{-163}}{2}]$ is a PID. This is the largest number for which you still get PID for this.

4.1 Basis elements

Method 1 has a problem: we can only draw rings which can be embedded as subrings of \mathbb{C} , or extended by embedding as $\subseteq \mathbb{R}^n$, which works in number theory, but applies only for a limited number of rings. Method 2 is more general, e.g. $\mathbb{k}[x, y]$, where we only drew the 2D array of basis elements $x^i y^j$.

Example. We used this picture from earlier in seeing that the ring $\mathbb{k}[y, xy, x^2y, \dots]$ is not a finitely generated subalgebra. Meditate upon the picture of the region that contains everything except x, x^2, x^3, \dots

Example. What is the dimension of the ring $\mathbb{k}[x, y]/(x^3, x^2y^2, y^5)$? Quite why I'd want to know the dimension of this ring I have no idea why. By algebraic calculation, it could work in 18 years. But really obvious if you just draw a picture. Identify the points $1 = (0, 0)$, $y^5 = (0, 5)$, $x^2y^2 = (2, 2)$, etc., on the 2D grid. The ideal (x^2y^2) consists of the infinite region with x^2y^2 as its bottom left corner. Killing off all of these things, the three regions, i.e. the multiples of x^3 , x^2y^2 , and y^5 , we now just look at all the monomials that are left. There are 12, which form a basis for the quotient; 12 is its dimension. We made the problem trivial by drawing the right picture and looking at it, without messing around with algebra.

Example. There are a huge number of rings that are finite-dimensional vector spaces over \mathbb{k} , which implies it's an Artinian ring. You can't even have infinitely descending chains in vector spaces. This is obvious: draw the same 2D grid of basis elements. Pick a random ideal, a random *graded* ideal with a zigzag border, like the boundary path of a fence for a sheep enclosure, or like an inverse staircase. The quotient is finite-dimensional. These correspond to partitions, a very large number. So Artinian rings are very common. In fact, they are even far more common than this diagram suggests; there's no reason at all why ideals should be graded at all.

This is a general theme in mathematics; Artinian things tend to be too messy to classify, there's a huge number of them and too wild. Sort of related to finite groups of prime power order, and the absurdly large number of them. [Recall the book *Groups of Order 2^n* , $n \leq 6$.] Reason why there are so many is closely related to why the number of Artinian rings is absurdly large.

Example. Related to one of the homework exercises. Consider the invariants of $G = (\mathbb{Z}/2\mathbb{Z}) = (g)$ acting on $\mathbb{Z}[x, y]$ by the simple diagonal action $gx = -x$, $gy = -y$. The picture of invariant ring: draw the same picture of $\mathbb{k}[x, y]$ again. The ring of invariants is the checkerboard pattern, $1, xy, x^2, \text{ etc.}$, and the ring of invariants is generated by $a = x^2, b = xy, c = y^2$. This is not a polynomial ring: $b^2 = ac$ is a relation between the generators. But it can give a description of the structure of the ring.

The ring is (obviously) a free module of rank 2 over a polynomial ring. Obvious by picture: $\mathbb{k}[x^2, y^2]$, and the remaining green squares ($xy, x^3y, xy^3, \text{ etc.}$, like a double-spaced checkerboard) have basis $1, xy$. Rings that are free modules of finite rank over polynomial rings have a beautiful property called being *Cohen–Macaulay*, so we proved a ring is CC just by drawing a picture of it.

Example. What else can we do? The quotient field is a field of rational functions in 2 variables. $\mathbb{k}(x^2, xy, y^2)$. $\mathbb{k}[x^2, xy, y^2]$ is not a polynomial ring; if it was, we could just take two generators, and the quotient ring would be all rational functions of those two variables. So we have to think carefully about why the quotient field is a field of rational functions, even though the ring isn't polynomial.

Look at the picture of a slightly bigger ring, $\mathbb{k}[x, y, x^{-1}, y^{-1}]$, polynomials, but aren't polynomials. The grid is now the full \mathbb{Z}^2 , where $x^n y^m$ is at (n, m) . Now, the ring of invariants of G acting on $\mathbb{k}[x, y, x^{-1}, y^{-1}]$ is the larger checkerboard extending infinitely in all directions, where the "red squares" include the origin and move diagonally, $xy, x^{-1}y, x^2, x^2y^2, \text{ etc.}$ We notice that the red points form a (sub)lattice \mathbb{Z}^2 , and the basis can consist of the expected vectors at ± 45 degrees, xy and xy^{-1} . So the quotient field is the ring of rational functions in xy, xy^{-1} , now independent variables, and isomorphic to $\mathbb{k}(z_1, z_2)$.

What's happening is a little subtle. The ring of invariants of G on $\mathbb{k}[x, y]$ is not a free monoid on two variables; we can't pick two red points such that every red point is linear combination of them. But once we extend to the lattice in the whole of \mathbb{Z}^2 , then the red points *do* become a free abelian group on 2 variables. The ring of variables is not a polynomial ring, but the quotient field is a field of rational functions in 2 variables. Easy to see if you draw a picture, hard if doing random algebraic calculations.

4.2 Prime ideals

The first two methods have the disadvantage of only working for special rings, but the third method works for all rings. The first obvious question: why on earth would it work??

Idea: try to think of a ring as some sort of space of functions on some topological space. If you can represent a ring like this, it becomes easy to think about. Suppose you have the ring of all continuous functions on \mathbb{R} ; then you can visualize its elements easily.

You can't represent *anything* like this; e.g., if there are nilpotent elements, there is no good way to represent as a ring of functions. But we can get pretty close. First, the set of prime ideals has the Zariski topology and is called the **spectrum** of a ring.

Why is it called the spectrum? This is closely related to the spectrum in physics, e.g. of a star, which has certain spectral lines. Black lines at certain points indicate whether the star contains various elements or whatever. Hilbert suggested these lines should be the eigenvalues of some operator on a Hilbert space. His guess turned out to be correct, when quantum mechanics was invented a few decades later. Pretty good guess on his part. This leads to the spectrum of an operator, the set of its eigenvalues.

Next lecture, we will see how to go from the spectrum of an operator on a vector space to the spectrum of a ring.

5 2023-01-31

