250B: Commutative Algebra

Nir Elber

Spring 2022

Contents

Theme 1: Introduction

Hold tight to your geometric motivation as you learn the formal structures which have proved to be so effective in studying fundamental questions.

—Ravi Vakil

1.1 January 18

So it begins.

1.1.1 Logistics

Here are some logistic things.

- We are using Eisenbud's *Commutative Algebra: With a View Toward Algebraic Geometry*. We will follow it pretty closely.
- All exams will be open-book and at-home. The only restrictions are time constrains (1.5 hours, 1.5 hours, and 3 hours).
- The first homework will be posted on Monday, and it will be uploaded to bCourses.
- Supposedly there will be a reader for the course, but nothing is known about the reader.

1.1.2 Rings

Commutative algebra is about commutative rings.

Convention 1.1. All of our rings will have a 1_R element and be commutative, as God intended. We do permit the zero ring.

We are interested in particular kinds of rings. Here are some nice rings.

Integral domain

Definition 1.2 (Integral domain)**.** An *integral domain* is a (nonzero) ring R such that, for $a, b \in R$, $ab = 0$ implies $a = 0$ or $b = 0$.

Units

Definition 1.3 (Units)**.** Given a ring R, we define the group of *units* $R^\times$ to be the set of elements of R which have multiplicative inverses.

Field

Definition 1.4 (Field)**.** A *field* is a nonzero ring R for which $R = \{0\} \cup R^\times$.

Reduced

Definition 1.5 (Reduced)**.** A ring R is *reduced* if and only if it has no nonzero nilpotent elements.

Local

Definition 1.6 (Local)**.** A ring R is *local* if and only if it has a unique (proper) maximal ideal.

It might seem strange to have lots a unique maximal ideal; here are some examples.

Example 1.7. Any field is a local ring with maximal ideal $\{0\}$.

Example 1.8. The ring of p-adic integers $\mathbb{Z}_p$ is a maximal ring with maximal ideal (p).

Example 1.9. The ring $\mathbb{Z}/p^2\mathbb{Z}$ is a local ring with maximal ideal $p\mathbb{Z}/p^2\mathbb{Z}$.

1.1.3 Ideals

The following is our definition.

Ideal

Definition 1.10 (Ideal)**.** Given a ring R, a subset $I \subseteq R$ is an *ideal* if it contains 0 and is closed under R-linear combination.

Given a ring R, we will write

$$(S) \subseteq R$$

to be the ideal generated by the set $S \subseteq R$.

Finitely generated

Definition 1.11 (Finitely generated)**.** An ideal $I \subseteq R$ is said to be *finitely generated* if and only if there are finitely many elements $r_1, \dots, r_n \in R$ such that $I = (r_1, \dots, r_n)$.

Principal

Definition 1.12 (Principal)**.** An ideal $I \subseteq R$ is *principal* if and only if there exists $r \in R$ such that $I = (r)$.

We mentioned maximal ideals above; here is that definition.

Maximal

Definition 1.13 (Maximal)**.** An ideal $I \subseteq R$ is *maximal* if and only if $I \neq R$ and, for any ideal $J \subseteq R$, $I \subseteq J$ implies $I = J$ or $I = R$.

Alternatively, an ideal $I \subseteq R$ is maximal if and only if the quotient ring R/I is a field. We will not show this here.

Prime

Definition 1.14 (Prime)**.** An ideal $I \subseteq R$ is *prime* if and only if $I \neq R$ and, for $a, b \in R$, $ab \in I$ implies $a \in I$ or $b \in I$.

Again, we can view prime ideals by quotient: I is prime if and only if R/I is a (nonzero) integral domain.

With the above definitions in mind, we can define the following very nice class of rings.

Principal ideal

Definition 1.15 (Principal ideal)**.** An integral domain R is a *principal ideal domain* if and only if all ideals of R are principal.

Example 1.16. The ring $\mathbb{Z}$ is a principal ideal domain. The way this is showed is by showing $\mathbb{Z}$ is Euclidean. Explicitly, fix $I \subseteq \mathbb{Z}$ an ideal. Then if $I \neq (0)$, find an element of $m \in I$ of smallest absolute value and use the division algorithm to write, for any $a \in I$,

$$a = mq + r$$

for $0 \leq r < m$. But then $r \in I$, so minimality of m forces $r = 0$, so $a \in (m)$, finishing.

Example 1.17. For a field k, the ring $k[x]$ is a principal ideal domain. Again, this is because $k[x]$ is a Euclidean domain, where we measure size by degree.

1.1.4 Unique Factorization

We have the following definition.

Irreducible, prime

Definition 1.18 (Irreducible, prime)**.** Fix R a ring and $r \in R$ an element.

- We say that $r \in R$ is *irreducible* if and only if r is not a unit, not zero, and $r = ab$ for $a, b \in R$ implies that one of a or b is a unit.
- We say that $r \in R$ is *prime* if and only if r is not a unit, not zero, and (r) is a prie ideal: $ab \in (r)$ implies $a \in (r)$ or $b \in (r)$.

This gives rise to the following important definition.

Unique factorization domain

Definition 1.19 (Unique factorization domain)**.** Fix R an integral domain. Then R is a *unique factorization domain* if and only if all nonzero elements of R have a factorization into irreducible elements, unique up to permutation and multiplication by units.

Remark 1.20. Units have the "empty" factorization, consisting of no irreducibles.

Example 1.21. The ring $\mathbb{Z}$ is a unique factorization domain. We will prove this later.

Note there are two things to check: that the factorization exists and that it is unique. Importantly, existence does not imply uniqueness.

Exercise 1.22. There exists an integral domain R such that every element has a factorization into irreducibles but that this factorization is unique.

Proof. Consider the subring $R := k\left[x^2, xy, y^2\right] \subseteq k[x, y]$. Here x^2, xy, y^2 are all irreducibles because the only way to factor a quadratic nontrivially would be into linear polynomials, but R has no linear polynomials.

However, these elements are not prime:

$$x^2 \mid xy \cdot xy$$

while x^2 does not divide xy. More concretely, $(xy)(xy) = x^2 \cdot y^2$ provides non-unique factorization into irreducibles. ■

The following condition will provide an easier check for the existence of factorizations.

Ascending chain condition

Definition 1.23 (Ascending chain condition)**.** Given a collection of sets $\mathcal{S}$, we say that $\mathcal{S}$ *has the ascendinc chain condition* (ACC) if and only every chain of sets in $\mathcal{S}$ must eventually stablize.

Example 1.24 (ACC for principal ideals)**.** A ring R has the ascending chain condition for principal ideals if and only if every ascending chain of principal ideals

$$(a_1) \subseteq (a_2) \subseteq (a_3) \subseteq \cdots$$

has some N such that $(a_N) = (a_n)$ for $n \geq N$.

Now, the fact that $\mathbb{Z}$ is a unique factorization domain roughly comes from the fact that $\mathbb{Z}$ is a principal ideal domain.

Theorem 1.25. Fix R a ring. Then R is a principal ideal domain implies that R is a unique factorization domain.

Proof. We start by showing that R has the ascending chain for principal ideals. Indeed, suppose that we have some ascending chain of principal ideals

$$(a_1) \subseteq (a_2) \subseteq (a_3) \subseteq \cdots.$$

Then the key idea is to look at the union of all these ideals, which will be an ideal by following the chain condition. However, R is a principal ideal domain, so there exists $b \in R$ such that

$$\bigcup_{k=1}^{\infty} (a_k) = (b).$$

However, it follows $b \in (a_N)$ for some N, in which case $(a_n) = (a_N)$ for each $n \geq N$.

We can now show that every nonzero element in R has a factorization into irreducibles.

Lemma 1.26. Suppose that a ring R has the ascending chain condition for principal ideals. Then every nonzero element of R has a factorization into irreducibles.

Proof. Fix some $r \in R$. If $(r) = R$, then r is a unit and hence has the empty factorization.

Otherwise, note that every ideal can be placed inside of a maximal and hence prime ideal, so say that $(r) \subseteq \mathfrak{m}$ where $\mathfrak{m}_1$ is prime; because R is a principal ring, we can say that $\mathfrak{m} = (\pi_1)$ for some $\pi_1 \in R$, so $\pi_1 \mid r$. This π_1 should go into our factorization, and we have left to factor r/π_1.

The above argument can then be repeated for r/π_1, and if r/π_1 is not a unit, then we get an irreducible π_2 and consider $r/(\pi_1\pi_2)$. This process must terminate because it is giving us an ascending chain of principal ideals

$$(r) \subseteq \left(\frac{r}{\pi_1}\right) \subseteq \left(\frac{r}{\pi_1\pi_2}\right) \subseteq \cdots,$$

which must stabilize eventually and hence must be finite. Thus, there exists N so that

$$\left(\frac{r}{\pi_1\pi_2\cdots\pi_N}\right) = R,$$

so $r = u\pi_1\pi_2\cdots\pi_N$ for some unit $u \in R^\times$. ∎

It remains to show uniqueness of the factorizations. The main idea is to show that all prime elements of R are the same as irreducible ones. One direction of the implication does not need the fact that R is a principal ring.

Lemma 1.27. Fix R an integral domain. Then any prime $r \in R$ is also irreducible.

Proof. Note that r is not a unit because it is prime. Now, suppose that $r = ab$ for $a, b \in R$; this implies that $r \mid ab$, so because r is prime, without loss of generality we force $r \mid a$. Then, dividing by r (which is legal because R is an integral domain), we see that

$$1 = (a/r)b,$$

so b is a unit. This finishes showing that r is irreducible. ∎

Warning 1.28. The reverse implication of the above lemma is not true for arbitrary integral domains: in the ring $\mathbb{Z}[\sqrt{-5}]$, there is the factorization

$$\left(1+\sqrt{-5}\right)\left(1-\sqrt{-5}\right) = 2 \cdot 3.$$

One can show that all elements above are irreducible, but none of them are prime.

The other side of this is harder. Pick up some $\pi \in R$ which is irreducible, and we show that π is prime. In fact, we will show stronger: we will show that (π) is a maximal ideal. Note $(\pi) \neq R$ because π is not a unit.

Indeed, suppose that $(\pi) \subseteq (r)$ for some ideal $(r) \subseteq R$. Then

$$\pi = rs$$

for some $s \in R$. Now, one of r or s must be a unit (π is irreducible). If s is a unit, then $(\pi) = (r)$; if r is a unit then $(r) = R$. This finishes showing that (π) is maximal.

From here we show the uniqueness of our factorizations. We proceed inductively, noting that two empty factorizations are of course the same up to permutation and units. Now suppose we have two factorizations of irreducibles

$$\prod_{k=1}^{m} p_k = \prod_{\ell=1}^{n} q_\ell,$$

where $k + \ell \geq 1$. Note that we cannot have exactly one side with no primes because this would make a product of irreducibles into 1, and irreducibles are not units.

Now, consider p_m. It is irreducible and hence prime and hence divides one of the right-hand factors; without loss of generality $p_m \mid q_n$. But (p_m) and (q_n) are both maximal ideals, so $(p_m) \subseteq (q_n)$ forces equality, so p_m/q_n is a unit. So we may cross off p_m and q_n and continue downwards by induction. ∎

1.1.5 Digression on Gaussian Integers

As an aside, the study of unique factorization came from Gauss's study of the Gaussian integers.

Gaussian integers

Definition 1.29 (Gaussian integers)**.** The *Gaussian integers* are the ring

$$\mathbb{Z}[i] := \{a + bi : a, b \in \mathbb{Z}\}.$$

One can in fact check that $\mathbb{Z}[i]$ is a principal ideal domain, which implies that $\mathbb{Z}[i]$ is a unique factorization domain. The correct way to check that $\mathbb{Z}[i]$ is a principal ideal domain is to show that it is Euclidean.

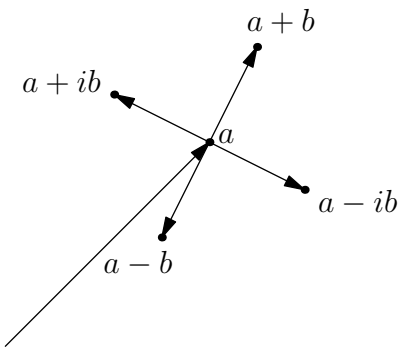
Lemma 1.30. The ring $\mathbb{Z}[i]$ is Euclidean, where our norm is $\mathrm{N}(a+bi) := a^2 + b^2$. In other words, given $\alpha, \beta \in \mathbb{Z}[i]$, we need to show that there exists $q \in \mathbb{Z}[i]$ such that

$$a = bq + r$$

where $r = 0$ or $\mathrm{N}(r) < \mathrm{N}(\beta)$.

Proof. The main idea is to view $\mathbb{Z}[i] \subseteq \mathbb{C}$ geometrically as in $\mathbb{R}^2$. We may assume that $|\beta| \leq |\alpha|$, and then it suffices to show that in this case we may find q so that $a - bq$ has smaller norm than a and induct.

Well, for this it suffices to look at $a + b, a - b, a + ib, a - ib$; the proof that one of these works essentially boils down to the following image.



Note that at least one of the endpoints here has norm smaller than a. ■

What about the primes? Well, there is the following theorem which will classify.

Theorem 1.31 (Primes in $\mathbb{Z}[i]$)**.** An element $\pi := a+bi \in \mathbb{Z}[i]$ is *prime* if and only if $\mathrm{N}(\pi)$ is a $1 \pmod 4$ prime, $(pi) = (1+i)$, or $(\pi) = (p)$ for some prime $p \in \mathbb{Z}$ such that $p \equiv 3 \pmod 4$.

We will not fully prove this; it turns out to be quite hard, but we can say small things: for example, $3 \pmod 4$ primes p remain prime in $\mathbb{Z}[i]$ because it is then impossible to solve

$$p = a^2 + b^2$$

by checking $\pmod 4$.

Remark 1.32. This sort of analysis of "sums of squares" can be related to the much harder analysis of Fermat's last theorem, which asserts that the Diophantine equation

$$x^n + y^n = z^n$$

for $xyz \neq 0$ integers such that $n > 2$.

1.1.6 Noetherian Rings

We have the following definition.

Noetherian ring

Definition 1.33 (Noetherian ring)**.** A ring R is said to be *Noetherian* if its ideals have the ascending chain condition.

There are some equivalent conditions to this.

Proposition 1.34. Fix R a ring. The following conditions are equivalent.

- R is Noetherian.
- Every ideal of R is finitely generated.

Proof. We show the directions one at a time.

- Suppos that R has an ideal which is not finitely generated, say $J \subseteq R$. Then we may pick up any $a_1 \in J$ and observe that $J \neq (a_1)$.

  Then we can pick up $a_2 \in J \setminus (a_1)$ and observe that $J \neq (a_1, a_2)$. So then we pick up $a_3 \in J \setminus (a_1, a_2)$ and continue. This gives us a strictly ascending chain

$$(a_1) \subsetneq (a_1, a_2) \subsetneq (a_1, a_2, a_3) \subsetneq \cdots,$$

contadicting the ascending chain condition.

- Suppose that every ideal is finitely generated. Then, given any ascending chain of ideals

$$I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots,$$

we need this chain to stabilize. Well, the union

$$I := \bigcup_{k=1}^{\infty} I_k$$

is also an ideal, and it must be finitely generated, so suppose $I = (a_1, a_2, \ldots, a_m)$. However, each a_k must appear in some $I_\bullet$ (and then each $I_\bullet$ after that one as well); choose N large enough to that $a_k \in I_N$ for each k. This implies that, for any $n \geq N$,

$$I_n \subseteq I = (a_1, a_2, \ldots, a_m) \subseteq I_N \subseteq I_n$$

verifying that the chain has stabilized. ■

A large class of rings turn out to be Noetherian, and in fact oftentimes Noetherian rings can build more Noetherian rings.

Proposition 1.35. Fix R a Noetherian ring and $I \subseteq R$ an ideal. Then R/I is also Noetherian.

Proof. Any chain of ideals in R/I can be lifted to a chain in R by taking pre-images along $\varphi : R \twoheadrightarrow R/I$. Then the chain must stabilize in R, so they will stabilize back down in R/I as well. ■

The above works because quotienting is an algeebraic operation. In contrast, merely being a subring is less algebraic, so it is not so surpsising that $R_1 \subseteq R_2$ with R_2 Noetherian does not imply that R_1 is Noetherian.

Example 1.36. The ring $k[x_1, x_2, \ldots]$ is not Noetherian because we have th infinite ascending chain

$$(x_1) \subseteq (x_1, x_2) \subseteq (x_1, x_2, x_3) \subseteq \cdots.$$

Howver, $k[x_1, x_2, \ldots] \subseteq k(x_1, x_2, \ldots)$, and the latter ring is Noetherian because it is a field. (Fields are Noetherian because they have finitely many ideals and therefore satisfy the ascending chain condition automatically.)

Here is another way to generate Noetherian rings.

Theorem 1.37 (Hilbert basis)**.** If R is a Noetherian ring, then $R[x]$ is also a Noetherian ring.

Corollary 1.38. By induction, if R is Noetherian, then $R[x_1, x_2, \ldots, x_n]$ is Notherian for any finite n.

Warning 1.39. Again, it is not true that $R[x_1, x_2, \ldots]$ is Noetherian, even though "inducting" with the Hilbert basis theorem might suggest that it is.

Proof of Theorem 1.37. The idea is to use the degree of polynomials to measure size. Fix $I \subseteq R[x]$ an ideal, and we apply the following inductive process.

- Pick up $f_1 \in I$ of minimal degree in I.
- If $I = (f_1)$ then stop. Otherwise find $f_2 \in I \setminus (f_1)$ of minimal degree.

- In general, if $I \neq (f_1, \ldots, f_n)$, then pick up $f_{n+1} \in I \setminus (f_1, \ldots, f_n)$ of minimal degree.

Importantly, we do not know that there are only finitely many $f_\bullet$ yet.

Now, look at the leading coefficients of the $f_\bullet$, which we name $a_\bullet$. However, the ideal

$$(a_1, a_2, \ldots) \subseteq R$$

must be finitely generated, so there is some finite N such that

$$(a_1, a_2, \ldots) = (a_1, a_2, \ldots, a_N).$$

To finish, we claim that

$$I \stackrel{?}{=} (f_1, f_2, \ldots, f_N).$$

Well, suppose for the sake of contradiction that we had some $f_{N+1} \in I \setminus (f_1, f_2, \ldots, f_N)$ of least degree. We must have $\deg f_{N+1} \geq \deg f_\bullet$ for each $f_\bullet$, or else we contradict the construction of $f_\bullet$ as being least degree.

To finish, we note $a_{N+1} \in (a_1, a_2, \ldots, a_N)$, so we are promised constants $c_1, c_2, \ldots, c_N$ such that

$$a_{N+1} = \sum_{k=1}^{N} c_k a_k.$$

In particular, the polynomial

$$g(x) := f_{N+1}(x) - \sum_{k=1}^{N} c_k a_k x^{(\deg g) - (\deg f_k)} f_k(x),$$

will be gauarnteed to kill the leading term of $f_{N+1}(x)$. But $g \equiv f_{N+1} \pmod{I}$, so g is suddenly a polynomial also not in I while of smaller degree than f_{N+1}, which is our needed contradiction. ■

1.1.7 Modules

To review, we pick up the following definition.

Module

Definition 1.40 (Module)**.** Fix R a ring. Then M is an abelian group with an R-action. Explicitly, we have the following properties; fix any $a, b \in R$ and $m, n \in M$.

- $1_R m = m$.
- $a(bm) = (ab)m$.
- $(a + b)m = am + bm$.
- $a(m + n) = am + an$.

Example 1.41. Any ideal $I \subseteq R$ is an R-module. In fact, ideals exactly correspond to the R-submodules of R.

Example 1.42. Given any two R-module M with a submodule $N \subseteq M$, we can form the quotient M/N.

Modules also have a notion of being Noetherian.

Noetherian module

Definition 1.43 (Noetherian module)**.** We say that an R-module M is *Noetherian* if and only if all R-submodules of M are finitely generated.

Remark 1.44. Equivalently, M is Noetherian if and only if the submodules of M have the ascending chain condition. The proof of the equivalence is essentially the same as Proposition 1.34.

Because modules are slightly better algebraic objects than rings, we have more ways to stitch modules together and hence more ways to make Noetherian modules. Here is one important way.

Proposition 1.45. Fix a short exact sequence

$$0 \to A \to B \to C \to 0$$

of R-modules. Then B is Noetherian if and only if A and C are both Noetherian.

Proof. We will not show this here; it is on the homework. Nevertheless, let's sketch the forwards direction, which is easier. Take B Noetherian.

- To show that A is Noetherian, it suffices to note that any submodules $M \subseteq A$ will also be a submodule of B and hence be finitely generated because B is Noetherian.
- To show that C is Noetherian, we note that C is essentially a quotient of B, so we can proceed as we did in Proposition 1.35.[1] ■

Because we like Noetherian rings, the following will be a useful way to make Noetherian modules from them.

Proposition 1.46. Every finitely generated R-module over a Noetherian ring R is Noetherian.

Proof. If M is finitely generated, then there exists some $n \in \mathbb{N}$ and surjective morphism

$$\varphi : R^n \twoheadrightarrow M.$$

Now, because R is Noetherian, R^n will be Noetherian by an induction: there is nothing to say when $n = 1$. Then the inductive step looks at the short exact sequence

$$0 \to R \to R^n \to R^{n-1} \to 0.$$

Here, the fact that R and R^{n-1} are Noetherian implies that R^n is Noetherian by Proposition 1.45. Anyways, the point is that M is the quotient of a Noetherian ring and hence Noetherian by Proposition 1.45 (again). ■

Here is the analogous result for algebras.

Algebra

Definition 1.47 (Algebra)**.** An *R-algebra* S is a ring equipped with a homomorphism $\iota : R \to S$. Equivalently, we may think of an R-algebra as a ring with an R action.

Proposition 1.48. Fix R a Noetherian ring. Then any finitely generated R-algebra is Noetherian.

Proof. Saying that S is a finitely generated R-algebra (with associated map $\iota : R \to S$) is the same as saying that there is a surjective morphism

$$\varphi : R[x_1, \dots, x_n] \twoheadrightarrow S$$

for some $n \in \mathbb{N}$. (Explicitly, $\varphi|_R = \iota$, and each x_k maps to one of the finitely many generating elements of S.) But then S is the quotient of a $R[x_1, \dots, x_n]$, which is Noetherian by Corollary 1.38, so S is Noetherian as well. ■

[1] In fact, Proposition 1.35 is exactly this in the case where $B = R$.

1.1.8 Invariant Theory

In our discussion, fix k a field of characteristic 0, and let G be a finite group or $\mathrm{GL}_n(\mathbb{C})$ (say). Now, suppose that we have a map

$$G \to \mathrm{GL}_n(k).$$

Then this gives $k[x_1, \dots, x_n]$ a G-action by writing $gf(\vec{x}) := f(g^{-1}\vec{x})$. The central question of invariant theory is then as follows.

Question 1.49 (Invariant theory)**.** Fix everything as above. Then can we describe $k[x_1, \dots, x_n]^G$?

By checking the group action, it is not difficult to verify that $k[x_1, \dots, x_n]^G$ is a subring of $k[x_1, \dots, x_n]$. For brevity, we will write $R := k[x_1, \dots, x_n]$.

Here is a result of Hilbert which sheds some light on our question.

Theorem 1.50 (Hilbert's finiteness)**.** Fix everything as above with G finite. Then $R^G = k[x_1, \dots, x_n]^G$ is a finitely generated k-algebra and hence Noetherian.

Proof. We follow Eisenbud's proof of this result. We pick up the following quick aside.

Lemma 1.51. Fix everything as above. If we write some $f \in R^G$ as

$$f = \sum_{d=0}^{\deg f} f_d$$

where f_d is homogeneous of degree d (i.e., f_d contains all terms of f of degree d), then $f_d \in R^G$ as well.

Proof. Indeed, multiplication by $\sigma \in G$ will not change the degree of any monomial (note G is acting as $\mathrm{GL}_n(k)$ on the variables themselves), so when we write

$$\sum_{d=0}^{\deg f} \sigma f_d = \sigma f = f = \sum_{d=0}^{\deg f} f_d,$$

we are forced to have $\sigma f_d = f_d$ by degree comparison arguments. ∎

Remark 1.52. In other words, the above lemma asserts that R^G may be graded by degree.

The point of the above lemma is that decomposition of an element $f \in R^G$ into its homogeneous components still keeps the homogeneous components in R^G, which is a fact we will use repeatedly.

We now proceed with the proof. The main ingredients are the Hilbert basis theorem and the Reynolds operator. Here is the Reynolds operator.

Reynolds operator

Definition 1.53 (Reynolds operator)**.** Fix everything as above. Then, given $f \in R$, we define the *Reynolds operator* $\varphi : R \to R$ as

$$\varphi(f) := \frac{1}{\#G} \sum_{\sigma \in G} \sigma f.$$

Note that division by $\#G$ is legal because k has characteristic zero.

It is not too hard to check that $\varphi : R \to R^G$ and $\varphi|_{R^G} = \mathrm{id}_{R^G}$. Additionally, we see $\deg \varphi(f) \leq \deg f$.

Let $\mathfrak{m} \subseteq R^G$ be generated by the homogeneous elements of R^G of positive degree. The input by the Hilbert basis theorem is to say that $\mathfrak{m}R \subseteq R$ is an R-ideal, and R is Noetherian (by the Hilbert basis theorem!), so $\mathfrak{m}R$ is finitely generated. So set

$$\mathfrak{m}R = (f_1, \dots, f_n) = f_1 R + \cdots + f_n R.$$

By decomposing the $f_\bullet$ into their (finitely many) homogeneous components, we may assume that the $f_\bullet$ are homogeneous.

Now we claim that the $f_\bullet$ generate R^G (as a k-algebra). Note that there is actually nontrivial difficulty turning the above finite generation of $\mathfrak{m}R$ as an R-module into finite of R^G as a k-algebra and that these notions are nontrivially different. I.e., we are claiming

$$R^G \stackrel{?}{=} k[f_1, \dots, f_n].$$

Certainly we have $\supseteq$ here. For $\subseteq$, we show that any $f \in R^G$ lives in $k[f_1, \dots, f_n]$ by induction. By decomposing f into homogeneous parts, we may assume that f is homogeneous.

We now induct on $\deg f$. If $\deg f = 0$, then $f \in k \subseteq k[f_1, \dots, f_n]$. Otherwise, f is homogeneous of positive degree and hence lives in $\mathfrak{m}$. In fact, $f \in \mathfrak{m}R$, so we may write

$$f = \sum_{i=1}^{n} g_i f_i.$$

Note that, because f and f_i are all homogeneous, we may assume that the g_i is also homogeneous because all terms in g_i of degree not equal to

$$\deg f - \deg f_i$$

will have to cancel out in the summation and hence may as well be removed entirely. In particular, each g_i has $g_i = 0$ or is homogeneous of degree $\deg f - \deg f_i$, so $\deg g_i < \deg f$ always.

We would like to finish the proof by induction, noting that $g_i \in R^G$ and $\deg g_i < \deg f$ forces $g_i \in k[f_1, \dots, f_n]$, and hence $f \in k[f_1, \dots, f_n]$ by summation. However, we cannot do that because we don't actually know if $g_i \in R^G$! To fix this problem, we apply the Reynolds operator, noting

$$f = \varphi(f) = \sum_{i=1}^{n} \varphi(g_i) f_i.$$

So now we may say that $\varphi(g_i) \in R^G$ and $\deg \varphi(g_i) < \deg f$, so $\varphi(g_i) \in k[f_1, \dots, f_n]$, and hence $f \in k[f_1, \dots, f_n]$ by summation. This finishes. ∎

The main example here is as follows.

> **Exercise 1.54.** Let S_n act on $R := k[x_1, \dots, x_n]$ as follows: $\sigma \in S_n$ acts by $\sigma x_m := x_{\sigma m}$. Then we want to describe R^G, the *homogeneous polynonmials in n letters*.

Proof. We won't work this out in detail here, but the main point is that the fundamental theorem of symmetric polynomials tells us that

$$R^G = k[e_1, e_2, \dots, e_n],$$

where the $e_\bullet$ are *elementary symmetric functions*. Namely,

$$e_m := \sum_{\substack{S \subseteq \{1,\dots,n\} \\ \#S = m}} \prod_{s \in S} x_s.$$

It is quite remarkable that R^G turned out to be a freely generated k-algebra, just like R. ∎

Here is more esoteric example.

Exercise 1.55. Let $G =: \{1, g\} \cong \mathbb{Z}/2\mathbb{Z}$ act on $R := k[x, y]$ by $g \cdot x = -x$ and $g \cdot y = -y$. Then we want to describe R^G.

Proof. Here, R^G consists of all polynomials $f(x, y)$ such that $f(x, y) = f(-x, -y)$. By checking coefficients of the various $x^m y^n$ terms, we see that $f(x, y) = f(-x, -y)$ is equivalent to forcing all terms of odd degree to have coefficient zero.

In other words, the terms of even degree are the only ones which can have nonzero coefficient. Each such term $x^a y^b$ (taking $a \geq b$ without loss of generality) can be written as

$$x^a y^b = x^{a-b}(xy)^b = \left(x^2\right)^{(a-b)/2} (xy)^b,$$

where $a - b \equiv a + b \equiv 0 \pmod 2$ justifies the last equality. so in fact we can realize R^G as

$$R^G = k\left[x^2, xy, y^2\right].$$

To see that this ring is Noetherian, we note that there is a surjection

$$\varphi : k[u, v, w] \to k\left[x^2, xy, y^2\right]$$

taking $u \mapsto x^2$ and $v \mapsto xy$ and $w \mapsto y^2$. Thus, R is the quotient of a Noetherian ring and hence Noetherian itself. In fact, we can check that[2] $\ker \varphi = \left(uw - v^2\right)$ so that

$$R^G \cong \frac{k[u, v, w]}{(uw - v^2)}.$$

Even though R^G is Noetherian, it is not a freely generated k-algebra (i.e., a polynomial ring over k) because it is not a unique factorization domain! ■

Next class we will start talking about the Nullstellensatz, which has connections to algebraic geometry.

1.2 January 20

We continue following the Eisenbud machine.

1.2.1 Affine Space

To begin our discussion, we start with some geometry.

Affine space

Definition 1.56 (Affine space)**.** Given a field k and positive integer n, we define n-dimensional *affine space* over k to be $\mathbb{A}^d(k) := k^n$.

Now, given affine space $\mathbb{A}^n(k)$, we are interested in studying subsets which are solutions to some set of polynomial equations

$$f_1, \ldots, f_n \in k[x_1, \ldots, x_d].$$

This gives rise to the following definition.

Algebraic

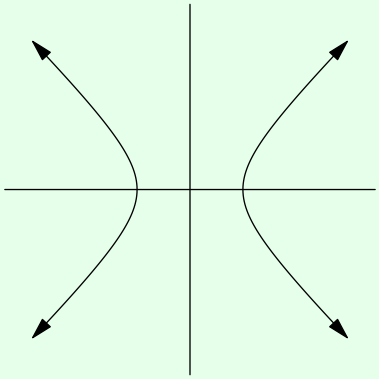
Definition 1.57 (Algebraic)**.** A subset $X \subseteq \mathbb{A}^n(k)$ is (affine) *algebraic* if and only if it is the set of solutions to some system of polynomials equations $f_1, \ldots, f_n \in k[x_1, \ldots, x_d]$.

[2] Certainly $uw - v^2 \in \ker \varphi$. In the other direction, any term $u^a v^b w^c$ can be written $\pmod{uw - v^2}$ as a term not having both u and w. However, each $x^d y^e$ has a unique representation in exactly one of the ways $u^a v^b \mapsto x^{2a+b} y^b$ ($a > 0$) or $v^b w^c \mapsto x^b y^{b+2c}$ ($c > 0$) or $v^b \mapsto x^b y^b$, so after applying the $\pmod{uw - v^w}$ movement, we see that the kernel is trivial.

Example 1.58. The hyperbola

$$\{(x, y) \in \mathbb{R}^2 : x^2 - y^2 - 1 = 0\}$$

is an algebraic set. Geometrically, it looks like the following.



Example 1.59. The set $\varnothing \subseteq \mathbb{A}^1(\mathbb{R})$ is algebraic becasue it is the set of solutions to the equation $x^2+1 = 0$ in $\mathbb{R}$.

The above example is a little disheartening because it feels like x^2+1 really ought to have a solution, namely $i \in \mathbb{C}$. More explicitly, there are no obvious algebraic obstructions that make $x^2 + 1$ not have a solution. So with this in mind, we make the following convention.

Convention 1.60. In the following discussion on the Nullstellensatz, k will always be an algebraically closed field.

1.2.2 Nullstellensatz

The Nullstellensatz is very important.

Remark 1.61. Because the Nullstellensatz is important, its name is in German (which was the language of Hilbert).

Now, the story so far is that we can take a set of polynomials and make algebraic sets as their solution set. We can in fact go in the opposite direction.

$I(X)$

Definition 1.62 ($I(X)$)**.** If $X \subseteq \mathbb{A}^n(k)$ is an (affine) algebraic set, we define

$$I(X) := \{f \in k[x_1, \ldots, x_n] : f(X) = 0\}.$$

It is not hard to check that $I(X) \subseteq k[x_1, \ldots, x_n]$ is in fact an ideal. Namely, if $f, g \in I(X)$ and $r, s \in k[x_1, \ldots, x_n]$, then we need to know $rf + sg \in I(X)$ as well. Well, for any $x \in X$, we see

$$(rf + sg)(x) = rf(x) + sg(x) = 0,$$

so $rf + sg \in I(X)$ indeed.

One might hope that all ideals would be able to take the form $I(X)$, but this is not the case. For example, if $f^m(X) = 0$, then $f(X) = 0$ because k is a field. Thus, I will satisfy the property that $f^m \in I$ implies $f \in I$. To keep track of this obstruction, we have the following definition.

Radical

Definition 1.63 (Radical)**.** Fix R a ring. Given an R-ideal I, we define the *radical of* I to be

$$\operatorname{rad} I := \{x \in R : x^n \in I \text{ for some } n \geq 1\} \supseteq I.$$

If $I = \operatorname{rad} I$, then we call I a *radical ideal*.

To make sense, this definition requires a few sanity checks.

- We check $\operatorname{rad} I$ is in fact an ideal. Well, given $f, g \in \operatorname{rad} I$, there exists positive integers m and n such that $f^m, g^n \in I$. Then, for any $r, s \in R$, we see

$$(rf + sg)^{m+n} = \sum_{k=0}^{m+n} \left[\binom{m+n}{k} r^k s^{m+n-k} \cdot f^k g^{m+n-k} \right].$$

  However, for any k, we see that either $k \geq m$ or $m + n - k \geq n$, so all terms of this sum contain an f^m or g^n factor, so the sum is in I. So indeed, $rf + sg \in \operatorname{rad} I$.

- We check that $\operatorname{rad} I$ is a radical ideal. Well, if $f^n \in \operatorname{rad} I$ for some positive integer n, then $f^{mn} = (f^n)^m \in I$ for some positive integer m, from which $f \in \operatorname{rad} I$ follows.

It is not too hard to generate examples where the radical is strictly larger than the original ideal.

Example 1.64. Fix $R := \mathbb{Z}[\sqrt{2}]$ and $I = (2) = 2\mathbb{Z}[\sqrt{2}] = \{2a + 2b\sqrt{2} : a, b \in \mathbb{Z}\}$. Then $\left(\sqrt{2}\right)^2 = 2 \in I$ while $\sqrt{2} \notin I$, so $I \subsetneq \operatorname{rad} I$.

Here is an alternative characterization of being radical.

Lemma 1.65. Fix R a ring. Then an ideal $I \subseteq R$ is radical if and only if R/I is reduced.

Proof. This proof is akin to the one showing $I \subseteq R$ is prime if and only if R/I is an integral domain.

Anyways, I is radical if and only if $x^n \in I$ for $x \in R$ and $n \geq 1$ implies $x \in I$. Translating this condition into R/I, we are saying that $[x]_I^n \in [0]_I$ for $[x]_I \in R/I$ and $n \geq 1$ implies that $[x]_I = [0]_I$. This is exactly the condition for R/I to be radical. ■

With all of the machinery we have in place, we can now state the idea of Hilbert's Nullstellensatz.

Theorem 1.66 (Nullstellensatz, I)**.** Fix k an algebraically closed field. Then there is a bijection between radical ideals of $k[x_1, \ldots, x_n]$ and (affine) algebraic sets $\mathbb{A}^n(k)$.

So far we have defined a map from algebraic sets to radical ideals by $X \mapsto I(X)$. The reverse map is as follows.

$Z(I)$

Definition 1.67 ($Z(I)$)**.** Given a subset $S \subseteq k[x_1, \ldots, x_n]$, we define the **zero set** of S by

$$Z(S) := \{x \in \mathbb{A}^n(k) : f(x) = 0 \text{ for all } f \in I\}.$$

Note that replacing S with the ideal it generates (S) makes no difference to $Z(S)$ (i.e., linear combinations of the constraints do not make the problem harder), so we will may focus on the case where S is an ideal.

With these maps in hand, we can restate the Nullstellensatz.

Theorem 1.68 (Nullstellensatz, II)**.** Fix k an algebraically closed field. Then for ideals $I \subseteq k[x_1, \ldots, x_n]$, we have

$$I(Z(I)) = \operatorname{rad} I.$$

In particular, if I is radical, then $I(Z(I)) = I$.

Remark 1.69. Yes, it is important that k is algebraically closed here. Essentially this comes from Example 1.59: the ideal (x^2+1) is not of the form $Z(X)$ for any subset $X \subseteq \mathbb{A}^1(\mathbb{R})$ because x^2+1 has no roots and would need $X = \varnothing$, but $Z(\varnothing) = \mathbb{R}[x]$.

Example 1.70. We have that $I(Z(R)) = R$ because $Z(R) = \varnothing$ (no points satisfy $1 = 0$) and $I(\varnothing) = R$ (all functions vanish on $\varnothing$).

Remark 1.71 (Nir)**.** One might object that $I(Z(I)) = \operatorname{rad} I$ only contains one direction of the bijection, but in fact it is not too hard to show directly that $Z(I(X)) = X$ for algebraic sets X. We argue as follows.

- Each $x \in X$ will cause all polynomials in $I(X)$ to vanish by construction of $I(X)$, so $X \subseteq Z(I(X))$.
- Now set $X = Z(S)$. Each $f \in S$ has $f(x) = 0$ for each $x \in S$, so $f \in I(X)$ as well. So $S \subseteq I(X)$, so $Z(I(X)) \subseteq Z(S) = X$.

1.2.3 More on Affine Space

Let's talk about $\mathbb{A}^n(k)$ a bit more. We mentioned that this should be a geometric object, so let's give it a topology.

Zariski topology, I

Definition 1.72 (Zariski topology, I)**.** Given affine space $\mathbb{A}^n(k)$, we define the *Zariski topology* as having closed sets which are the algebraic sets.

Remark 1.73 (Nir)**.** Here is one reason why we might do this: without immediate access to better functions (the field k might have no easy geometry, like $k = \mathbb{F}_p(t)$) it makes sense to at least require polynomial functions to be continuous and k to be Hausdorff. In particular, given a polynomial f, we see that

$$Z(f) = f^{-1}(\{0\})$$

should be closed. In fact, for any subset $S \subseteq k[x_1, \dots, x_n]$ of polynomials

$$Z(S) = \bigcap_{f \in S} Z(f)$$

will also have to be closed. In particular, all algebraic sets are closed. One can then check that polynomials do remain continuous in this topology also, as promised.

We have the following checks to make sure that the algebraic sets do actually form a topology (of closed sets).

- The empty set is closed: $\varnothing$ is the set of solutions to the equation $1 = 0$.
- The full space is closed: $\mathbb{A}^n(k)$ is the set of solutions to the equation $0 = 0$.
- Arbitrary intersection of closed sets is closed: given algebraic sets $X(S)$ for given subsets $S \subseteq \mathcal{S}$ of $k[x_1, \dots, x_n]$, we note

  $$\bigcap_{S \in \mathcal{S}} X(S) = X\left(\bigcup_{S \in \mathcal{S}} S\right),$$

  so the union is in fact an algebraic set.

- Finite unions of closed sets are closed: given algebraic sets $X(S_1), \dots, X(S_n)$, we note

$$\bigcup_{i=1}^{n} X(S_i) = X\left(\prod_{i=1}^{n}(S_i)\right),$$

where (S_i) is the ideal generated by S_i. In particular, $\prod_i (S_i)$ is generated by elements $s_1 \cdot \ldots \cdot s_n$ such that $s_i \in S_i$ for each i, so any point in any of the $X(S_i)$ will show up in the given algebraic set.

Now that we've checked we actually have a topology, we remark that it is a pretty strange topology.

Proposition 1.74. Let k be an algebraically closed field. Given affine space $\mathbb{A}^n(k)$ the Zariski topology.

- The space $\mathbb{A}^n(k)$ is not Hausdorff.
- The space $\mathbb{A}^n(k)$ is compact.

Proof. We take the claims individually.

- Because $\mathbb{A}^n(k)$ has more than one point, it suffices to show that there are no disjoint nonempty Zariski open subsets of $\mathbb{A}^n(k)$. In other words, given two Zariski open sets $\mathbb{A}^n(k) \setminus Z(I)$ and $\mathbb{A}^n(k) \setminus Z(J)$, we claim that

$$(\mathbb{A}^n(k) \setminus Z(I)) \cap (\mathbb{A}^n(k) \setminus Z(J)) = \varnothing$$

implies $\mathbb{A}^n(k) \setminus Z(I) = \varnothing$ or $\mathbb{A}^n(k) \setminus Z(J) = \varnothing$. Taking complements, we know that

$$Z(IJ) = Z(I) \cup Z(J) = \mathbb{A}^n(k) = Z((0)).$$

But now, by the Nullstellensatz (!), we see that $\operatorname{rad}(IJ) = \operatorname{rad}((0))$. But $k[x_1, \dots, x_n]$ is an integral domain, so $\operatorname{rad}((0)) = (0)$.

Now, this means $f^n \in IJ$ for some $n \in \mathbb{N}$ requires $f = 0$, which means that $IJ = (0)$, so because $k[x_1, \dots, x_n]$ is an integral domain, $I = (0)$ or $J = (0)$. (Explicitly, I and J cannot both have nonzero terms.) Without loss of generality, take $I = (0)$.

SO to finish, we see $Z(I) = Z((0)) = \mathbb{A}^n(k)$, so $\mathbb{A}^n(k) \setminus Z(I) = \varnothing$.

- Suppose we are given an open cover $\{\mathbb{A}^n(k) \setminus Z(I)\}_{I \in \mathcal{S}}$ indexed by some collection $\mathcal{S}$ of ideals of $k[x_1, \dots, x_n]$. The fact that these sets form an open cover is equivalent to saying

$$Z\left(\sum_{I \in \mathcal{S}} I\right) = \bigcap_{I \in \mathcal{S}} Z(I) = \varnothing.$$

Now, by the Nullstellensatz (we will use this trick again later on!), it follows

$$1 \in R = I(\varnothing) = I\left(Z\left(\sum_{I \in \mathcal{S}} I\right)\right) = \operatorname{rad} \sum_{I \in \mathcal{S}} I,$$

so it follows $1 \in \sum I$.

However, we can reduce this to a finite condition: $1 \in \sum_I$ merely means there are elements $\{f_i\}_{i=1}^N$ such that $f_i \in I_i$ for some $I_i \in \mathcal{S}$ such that $\sum_i f_i = 1$. This means that in fact $1 \in I_1 + \cdots + I_N$, so

$$\varnothing = Z(I_1 + \cdots + I_N) = \bigcap_{i=1}^{N} Z(I_i).$$

Thus, the finite number of sets $\mathbb{A}^n(k) \setminus Z(I_i)$ for each $1 \le i \le N$ provides us with a finite subcover of $\mathbb{A}^n(k)$. ∎

In another direction, we note can also understand algebraic sets $X \subseteq \mathbb{A}^n(k)$ by their ring of functions. Again, the only functions we have easy access to are polynomials, so we take the following definition.

Coordinate ring

Definition 1.75 (Coordinate ring)**.** Given an algebraic set $X \subseteq \mathbb{A}^n(k)$, we define the *coordinate ring* on X as

$$A(X) := k[x_1, \dots, x_n]/I(X).$$

In other words, we are looking at polynomials on $\mathbb{A}^n(k)$ and identifying them whenever they are equal on X.

Note that, because $I(X)$ is a radical ideal, the ring $A(X)$ will be reduced.

1.2.4 Corollaries of the Nullstellensatz

Let's return to talking about talking about the Nullstellensatz. To convince us that the Nullstellensatz is important, here are some nice corollaries.

Criteria for Polynomial System Solutions

The following is the feature of this subsubsection.

Corollary 1.76. A system of polynomial equations

$$\begin{cases} f_1(x_1, \dots, x_n) = 0, \\ \qquad \vdots \\ f_r(x_1, \dots, x_n) = 0, \end{cases}$$

has no solutions if and only if there exists $p_1, \dots, p_r \in k[x_1, \dots, x_n]$ such that

$$\sum_{i=1}^{r} p_i f_i = 1.$$

Proof. In the reverse direction, we proceed by contraposition: if there is a solution $x \in \mathbb{A}^n(k)$ such that $f_i(x) = 0$ for each f_i, then any set of polynomials $p_1, \dots, p_n \in k[x_1, \dots, x_n]$ will give

$$\sum_{i=1}^{r} p_i(x) f_i(x) = 0 \neq 1,$$

so it follows $\sum_{o=1}^{n} p_i f_i \neq 1$. Observe that we did not use the Nullstellensatz here.

The forwards direction is harder. The main point is that we are given $Z((f_1, \dots, f_r)) = \varnothing$, so

$$\operatorname{rad}(f_1, \dots, f_r) = I(Z((f_1, \dots, f_n))) = I(\varnothing) = R,$$

so the Nullstellensatz gives $1 \in \operatorname{rad}(f_1, \dots, f_r)$. Then it follows $1 = 1^n \in (f_1, \dots, f_r)$ for some positive integer n, so there exists $p_1, \dots, p_r \in k[x_1, \dots, x_n]$ such that

$$\sum_{i=1}^{r} p_i f_i = 1.$$

This is what we wanted. ∎

Maximal Ideals Are Points

To set up the next corollary, we claim that any point $a = (a_1, \dots, a_n) \in \mathbb{A}^n(k)$ makes a closed set corresponding to the ideal

$$I(\{a\}) \stackrel{?}{=} (x_1 - a_1, \dots, x_n - a_n) \subseteq k[x_1, \dots, x_n] = A(\mathbb{A}^n(k)).$$

Indeed, $I(\{a\})$ certainly contains $x_i - a_i$ for each i; conversely, if $f \in I(\{a\})$, then

$$f(x_1, \dots, x_n) \equiv f(a_1, \dots, a_n) = 0 \pmod{x_1 - a_1, \dots, x_n - a_n},$$

so $f \in (x_1 - a_1, \dots, x_n - a_n)$.

Example 1.77. In fact, in the case of $\mathbb{C}[x]$, it is not too hard to see that such ideals are maximal: given $z \in \mathbb{C}$, suppose that $I \subseteq \mathbb{C}[x]$ had $(x - z) \subseteq I$. If each $f \in I$ has $f(z) = 0$, then we are done; otherwise if there is $f \in I$ with $f(z) \neq 0$, then $f(x)$ and $(x - z)$ are coprime in a principal ideal domain, so

$$1 \in (f) + (x - z) \subseteq I,$$

meaning $I = \mathbb{C}[x]$.

The above example gives us the hope that maximal ideals might turn out to all be of the above form. Indeed, this is true, with the help of the Nullstellensatz.

Corollary 1.78. Fix $X \subseteq \mathbb{A}^n(k)$ an (affine) algebraic set. Then points $a = (a_1, \dots, a_n) \in X$ are in bijection with maximal ideals $\mathfrak{m}_a \subseteq A(X)$ by

$$a \mapsto \mathfrak{m}_a := I(\{a\})/I(X) = (x_1 - a_1, \dots, x_n - a_n)/I(X).$$

Proof. The input from the Nullstellensatz will come from the following lemma.

Lemma 1.79. Suppose that $I \subseteq A(\mathbb{A}^n(k))$ has $Z(I) = \varnothing$. Then $I = A(\mathbb{A}^n(k))$.

Proof. By the Nullstellensatz,

$$1 \in A(\mathbb{A}^n(k)) = I(\varnothing) = I(Z(I)) = \operatorname{rad} I,$$

so $1 \in I$ follows. ∎

Now, we have alreadys shown that $I(\{a\}) = (x_1 - a_1, \dots, x_n - a_n)$. Additionally, for $x \in X$, we have $I(X) \subseteq I(\{a\})$, so $I(\{a\})/I(X)$ is an ideal which makes sense. Thus, we may write $I(\{a\})/I(X) = (x_1 - a_1, \dots, x_n - a_n)/I(X)$.

Before continuing, we also check that $Z(I(\{a\})) = \{a\}$ as well. (This shows that $\{a\}$ is an algebraic set.) Well, set $a = (a_1, \dots, a_n)$, and we note that $x_i - a_i \in I(\{a\})$ for each i, so any $b = (b_1, \dots, b_n) \in Z(I(\{a\}))$ must vanish on each $x_i - a_i$, so

$$b_i - a_i = 0$$

for each i. Thus, $b = a$.

We now check that $a \mapsto \mathfrak{m}_a$ is a bijection.

- Well-defined: we show that $\mathfrak{m}_a$ is a maximal ideal. It is proper because $1 \notin \mathfrak{m}_a$. Now suppose we have $I \subseteq A(X)$ such that $\mathfrak{m}_a \subseteq I$. So note that $I + I(X) \subseteq A(\mathbb{A}^n(k))$ is an ideal (namely, the pre-image) containing $I(\{a\})$.

  Now, observe that $I(\{a\}) \subseteq I + I(X)$, so

  $$Z(I + I(X)) \subseteq Z(I(\{a\})) = \{a\}.$$

  We now have two cases.

  - If $Z(I + I(X)) = \varnothing$, then Lemma 1.79 gives $I + I(X) = A(\mathbb{A}^n(k))$, so $I/I(X) = A(X)$.
  - Otherwise if $Z(I + I(X)) = \{a\}$, then $I + I(X) \subseteq I(\{a\})$. Thus $I \subseteq \mathfrak{m}_a$, finishing.

- Injective: suppose $a, b \in X$ have $\mathfrak{m}_a = \mathfrak{m}_b$. But then
$$I(\{a\}) = \mathfrak{m}_a + I(X) = \mathfrak{m}_b + I(X) = I(\{b\}),$$
so $\{a\} = Z(I(\{a\})) = Z(I(\{b\})) = \{b\}$, so $a = b$ follows.
- Surjective: suppose that $\mathfrak{m} \subseteq A(X)$ is a maximal ideal. Then we look at the pre-image ideal $I := \mathfrak{m} + I(X) \subseteq A(\mathbb{A}^n(k))$. We claim that $Z(I)$ is a singleton.
  - We show that $Z(I) \neq \varnothing$. Indeed, $Z(I) = \varnothing$ implies by Lemma 1.79 that $1 \in I$, so $[1]_{I(X)} \in \mathfrak{m}$, which violates the fact that $\mathfrak{m} \subseteq A(X)$ is proper.
  - We show all elements of $Z(I)$ are equal. Suppose $a, b \in Z(I)$; because $I(X) \subseteq I$, we see $a, b \in X$ is forced by Remark 1.71. Then $\{a\}, \{b\} \subseteq Z(I)$, so
  $$I \subseteq I(\{a\}) \cap I(\{b\}),$$
  so $\mathfrak{m} = I/I(X)$ is contained in $\mathfrak{m}_a = I(\{a\})/I(X)$ and $\mathfrak{m}_b = I(\{b\})/I(X)$. But $\mathfrak{m}_a$ and $\mathfrak{m}_b$ are distinct maximal ideals, so we see $\mathfrak{m} \subseteq \mathfrak{m}_a \cap \mathfrak{m}_b \subsetneq \mathfrak{m}_a \subsetneq A(X)$, violating the fact that $\mathfrak{m}$ is maximal.

  Thus, set $Z(I) = \{a\}$; note $a \in X$ because $I(X) \subseteq I$ (by Remark 1.71 again). Now, $I \subseteq I(\{a\})$, so we see $\mathfrak{m} = I/I(X) \subseteq I(\{a\})/I(X) = \mathfrak{m}_a$, so the maximality of $\mathfrak{m}$ forces $\mathfrak{m} = \mathfrak{m}_a$. ∎

The reason the above is nice is because, instead of having to look at the geometry of X, it is now legal to study the algebra of $A(X)$.

1.2.5 The Spectrum of a Ring

We continue trying to move the geometry of affine sets $X \subseteq \mathbb{A}^n(k)$ into the coordinate ring $A(X)$.

Later in life we will want to consider maps $\varphi : X \to Y$ between affine sets. In affine space, we again remark that really only the functions we have access to are polynomials, so our only morphisms will be functions which are polynomials in each coordinate.

Now let's move φ to geometry. Note that $A(X)$ and $A(Y)$ arre intended to describe functions $X \to k$ and $Y \to k$ respectively, so a morphism $\varphi : X \to Y$ induces a ring homomorphism
$$\varphi : A(Y) \to A(X)$$
by $f \mapsto f \circ \varphi$. (This is a ring homomorphism because φ is made of polynomials.) So under the paradigm that points should become maximal ideals, we would like to recover φ as some kind of map of maximal ideals $A(X) \to A(Y)$. The natural way is to simply pull back along φ, writing
$$\mathfrak{m} \subseteq A(X) \mapsto \varphi^{-1}\mathfrak{m} \subseteq A(Y).$$
However, this is a problem: $\varphi^{-1}\mathfrak{m}$ need not be maximal!

Example 1.80. If $\mathfrak{p} \subseteq R$ is a prime but not maximal ideal (e.g., $(x) \subseteq k[x, y]$), we can define the composite
$$R \twoheadrightarrow R/\mathfrak{p} \hookrightarrow \operatorname{Frac}(R/\mathfrak{p}).$$
Now, (0) is maximal in $\operatorname{Frac}(R/\mathfrak{p})$, but its pre-image in R is $\mathfrak{p}$, which is not maximal by construction.

However, if we weaken requiring our points to be prime ideals $\mathfrak{p}$ instead of maximal ideals, we do have that $\varphi^{-1}\mathfrak{p}$ is a prime ideal: $ab \in \varphi^{-1}\mathfrak{p}$ implies $\varphi(a)\varphi(b) = \varphi(ab) \in \mathfrak{p}$ implies $a \in \varphi^{-1}\mathfrak{p}$ or $b \in \varphi^{-1}\mathfrak{p}$.

So instead of making our geometry on $A(X)$ defined by maximal ideals, we use prime ideals. This gives the following definition.

Spectrum of a ring

Definition 1.81 (Spectrum of a ring)**.** Given a ring R, we define *spectrum of* R by
$$\operatorname{Spec} R := \{\mathfrak{p} \subseteq R : \mathfrak{p} \text{ is a prime ideal}\}.$$

In fact, $\operatorname{Spec} R$ also has a Zariski topology as follows.

Zariski topology, II

Definition 1.82 (Zariski topology, II)**.** Given a ring R, we define the *Zariski topology* to have closed sets

$$X(I) := \{\mathfrak{p} \in \operatorname{Spec} R : I \subseteq \mathfrak{p}\}$$

for R-ideals I.

Remark 1.83 (Nir)**.** As for motivation for why we might define our topology like this, recall the case of affine varieties: we have $a \in X(I)$ if and only if $I \subseteq I(\{a\})$. So when we translate $X(I)$ into the algebraic side, we call the maximal ideal $\mathfrak{m}_a = I(\{a\})$ our "point" and see that

$$X(I) = \{\mathfrak{m}_a : I \subseteq \mathfrak{m}_a\}.$$

It is a different story why we use prime ideals instead of maximal ones, which we discussed above.

The checks that the $X(I)$ do actually define closed sets for a topology are essentially the same as for the first version of the Zariski topology. The main points are that

$$\bigcap_{I\in\mathcal{S}} X(I) = X\left(\sum_{I\in\mathcal{S}} I\right) \qquad \text{and} \qquad \bigcup_{k=1}^{N} X(I_k) = X\left(\prod_{k=1}^{N} I_k\right)$$

give that arbitrary intersection of closed sets is closed and finite union of closed sets is closed.[3]

Again, the Zariski topology is very weird, like with affine space.

Proposition 1.84. Fix R a ring. Given $\operatorname{Spec} R$ the Zariski topology.

- If R is an integral domain which is not a field, then $\operatorname{Spec} R$ is not Hausdorff.
- The space $\operatorname{Spec} R$ is compact.

Proof. We take the claims one at a time.

- The fact that R is a field means that $\operatorname{Spec} R$ has more than one point. So again, it suffices to show that there are no disjoint open subsets of $\operatorname{Spec} R$. Indeed, suppose
$$(\operatorname{Spec} R \setminus X(I)) \cap (\operatorname{Spec} R \setminus X(J)) = \varnothing,$$
and we claim $\operatorname{Spec} R \setminus X(I) = \varnothing$ or $\operatorname{Spec} R \setminus X(J) = \varnothing$.
Again, we know that $X(IJ) = X(I) \cup X(J) = \operatorname{Spec} R$, so by definition, we see $IJ \subseteq \mathfrak{p}$ for each prime $\mathfrak{p}$, or
$$IJ \subseteq \bigcap_{\mathfrak{p}} \mathfrak{p}.$$
Now, because R is an integral domain, we see that (0) is a prime ideal, so $IJ = (0)$ follows. Thus, because R is an integral domain again, $I = (0)$ or $J = (0)$, so without loss of generality, we take $I = (0)$. But then
$$\operatorname{Spec} R \setminus X(I) = \operatorname{Spec} R \setminus \operatorname{Spec} R = \varnothing,$$
as desired.
- Suppose that the Zariski open sets $\{\operatorname{Spec} R \setminus X(I)\}_{I\in\mathcal{S}}$ cover $\operatorname{Spec} R$, for some collection $\mathcal{S}$ of ideals. Now, the sets $\{\operatorname{Spec} R \setminus X(I)\}_{I\in\mathcal{S}}$ covering $\mathbb{A}^n(k)$ is equivalent to
$$X\left(\sum_{I\in\mathcal{S}} I\right) = \bigcap_{I\in\mathcal{S}} X(I) = \varnothing.$$

[3] The second equality requires some care. The main point is to show, for $\mathfrak{p}$ prime, $IJ \subseteq \mathfrak{p}$ is equivalent to $I \subseteq \mathfrak{p}$ or $J \subseteq \mathfrak{p}$. The reverse is easy. For the forwards, suppose $IJ \subseteq \mathfrak{p}$ and $J \not\subseteq \mathfrak{p}$ so that we have $j \in J \setminus \mathfrak{p}$. Then $jI \subseteq IJ \subseteq \mathfrak{p}$ forces $I \subseteq \mathfrak{p}$.

However, $X\left(\sum I\right) = \varnothing$ implies that there is no prime ideal $\mathfrak{p}$ such that $\sum I \subseteq \mathfrak{p}$, but any proper ideal is contained in some maximal and hence prime ideal. Thus, we must have that

$$\sum_{I \in \mathcal{S}} I = R.$$

In particular, 1 is in this ideal, so we can express 1 as the sum of some elements $x_i \in I_i$ for $\{I_i\}_{i=1}^N \subseteq \mathcal{S}$; i.e.,

$$1 = \sum_{i=1}^N x_i \in \sum_{i=1}^N I_i.$$

Thus, $\sum_{i=1}^N I_i = R$, meaning $X\left(\sum_{i=1}^N I_i\right) = \varnothing$, so reversing the argument we see that $\{\operatorname{Spec} R \setminus X(I_i)\}_{i=1}^N$ will be a finite subcover. This finishes. ■

1.2.6 Projective Space

To define projective varities, we need to define projective space first.

Projective space

Definition 1.85 (Projective space)**.** Fix k a field and n a positive integer. Then we define n-dimensional *projective space* $\mathbb{P}^n(k)$ to be the one-dimensional subspaces of k^{n+1}.

Concretely, we will think about lines in homogeneous coordinates, in the form

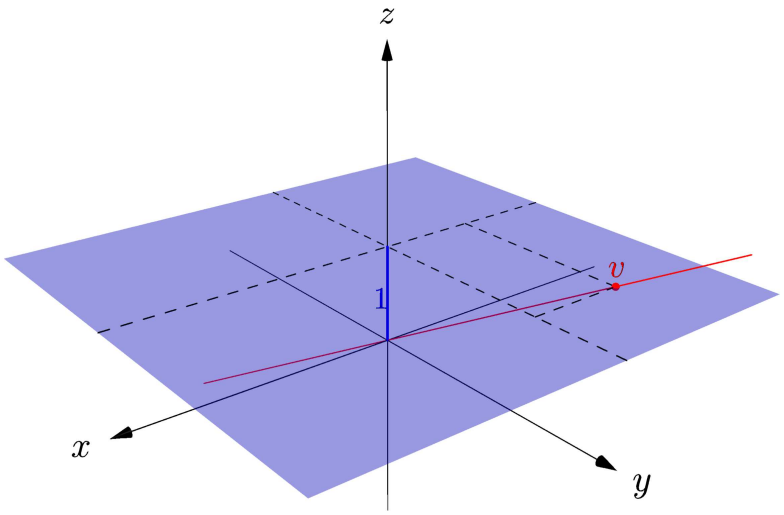
$$(a_0 : a_1 : \ldots : a_n) \in \mathbb{P}^n(k)$$

to represent the subspace $k(a_0, a_1, \ldots, a_n) \subseteq \mathbb{A}^{n+1}(k)$. As such multiplying the point $(a_0 : a_1 : \ldots : a_n)$ by some constant $c \in k^\times$ will give the same line and should be the same point in $\mathbb{P}^n(k)$. Additionally, we will ban the point $(0 : 0 : \ldots : 0)$ from projective space because it is not the basis for any line.

We would like to have a better geometry understanding of $\mathbb{P}^n(k)$. Note that we have a sort of embedding $\mathbb{A}^n(k) \hookrightarrow \mathbb{P}^n(k)$ by

$$(x_1, x_2, \ldots, x_n) \mapsto (x_1 : x_2 : \ldots : x_n : 1).$$

For geometric concreteness, we can imagine $\mathbb{A}^2(\mathbb{R}) \hookrightarrow \mathbb{P}^2(\mathbb{R})$ as the plane $z = 1$ in $\mathbb{R}^3$, where each point on the plane gives rise to a unique line in $\mathbb{R}^3$. Here is the image, with a chosen red line going through a point v on the $z = 1$ plane.



However, not all lines in $\mathbb{A}^3(\mathbb{R})$ can be described like this, for there are still lots of points of the form $(x : y : 0)$, which are "points at infinity." Nevertheless, we can collect the remaining points into $\mathbb{P}^2(\mathbb{R})$, which visually just means the lines that live on the xy-plane in the above diagram.

In general, we see that we can decompose $\mathbb{P}^n(k)$ into an $\mathbb{A}^n(k)$ component as a "$z = 1$ hyperplane" and then the points at infinity living on $\mathbb{P}^{n-1}(k)$. Namely,

$$\mathbb{P}^n(k) \text{ "=" } \mathbb{A}^n(k) \sqcup \mathbb{P}^{n-1}(k).$$

Note that the above decomposition is not canonical: one has to choose which points to get to be infinity.

Anyways, as usual we interested in studying the algebraic sets but this time of projective space, but because of the constant factors allowed to wiggle, we see that we really should only be looking at homogeneous equations. More concretely, if $f \in k[x_0, \dots, x_n]$, we want

$$f(a_0 : \dots : a_n) = 0$$

to be unambiguous, so $f(a_0, \dots, a_n) = 0$ should be imply $f(ca_0, \dots, ca_n) = 0$ for any $c \in k^\times$. The easiest way to ensure this is to force all monomials of f to have some fixed degree, say d, so that

$$f(cx_0, \dots, cx_n) = c^d f(x_0, \dots, x_n).$$

These polynomials are the homogeneous ones, and they give the following definition.

Projective variety

Definition 1.86 (Projective variety)**.** A subset $X \subseteq \mathbb{P}^n(k)$ is a *projective variety* if and only if it is the solution set to some set of homogeneous (!) polynomials equations of $k[x_0, \dots, x_n]$.

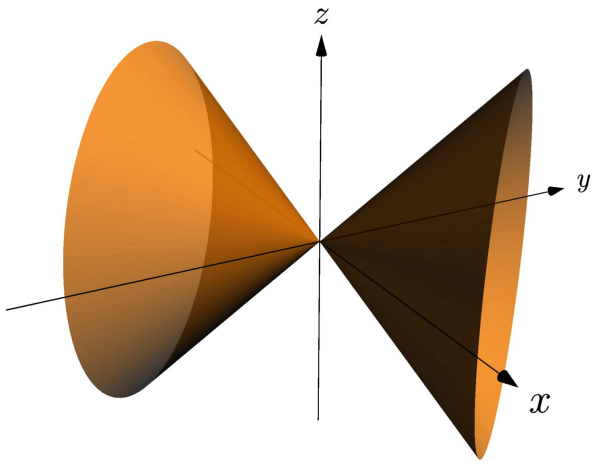
Here is an example.

Exercise 1.87. We view the solutions to $xy - 1 = 0$ in $\mathbb{A}^2(\mathbb{R}) \subseteq \mathbb{P}^2(\mathbb{R})$ in projective space.

Proof. More explicitly, we are viewing $\mathbb{A}^2(k) \subseteq \mathbb{P}^2(k)$ by sending $(x, y) \mapsto (x : y : 1)$. We can make the coordinates more familiar by setting $x, y \mapsto x/z, y/z$ so that we are looking for solutions $(x/z : y/z : 1) = (x : y : z)$ to the equation

$$xy = z^2.$$

In $\mathbb{R}^3$, this curve looks like the following.



The hyperbola for $xy = 1$ comes from slicing the $z = 1$ plane from this cone. ∎

1.2.7 Graded Rings

We have the following definition.

Graded ring

Definition 1.88 (Graded ring)**.** A ring R is *graded* by the abelian groups $R_0, R_1, \dots$ if and only if

$$R \cong \bigoplus_{d=0}^{\infty} R_d$$

as abelian groups and $R_i R_j \subseteq R_{i+j}$ for any $i, j \in \mathbb{N}$.

Remark 1.89 (Nir)**.** In fact, R_0 turns out to be a subring of R_0. We can check this directly, as follows.

- Certainly $0 \in R_0$ and $R_0 + R_0 \subseteq R_0$ because $R_0 \subseteq R$ is an additive subgroup.
- If $1_R \in R_i$, then $R_i \subseteq R_i R_i \subseteq R_{2i}$, so $i = 0$ or $R_i = R_{2i} = \{0\}$ by disjointness. So either $1 \in R_0$ or $1 \in R_0 = \{0\}$ forces $R = \{0\}$, so $1 \in R_0$ anyways.
- We see $R_0 R_0 \subseteq R_0$, so R_0 is closed under multiplication.

Alternatively, we could set $I := \{0\} \oplus R_1 \oplus R_2 \cdots$, remark that I is an ideal, and then we see $R_0 \cong R/I$.

Example 1.90. The ring $R = k[x_1, \dots, x_n]$ is "graded by degree" by setting R_d to be the space of all homogeneous n-variable polynomials of degree d (unioned with $\{0\}$).

With graded rings, it is natural to ask what other ring-theoretic constructions we can grade.

Graded ideal

Definition 1.91 (Graded ideal)**.** Fix R a graded ring. We say that an ideal I is *graded* if and only if

$$I \cong \bigoplus_{d=0}^{\infty} (R_d \cap I),$$

where the isomorphism is the natural one (i.e., $(x_0, x_1, \dots) \mapsto x_0 + x_1 + \cdots$).

Example 1.92. Given the graded ring $R = R_0 \oplus R_1 \oplus R_2 \oplus \cdots$, the ideal

$$I := R_1 \oplus R_2 \oplus R_3 \oplus \cdots$$

is called the *irrelevant ideal*; it is graded because look at it. To check I is an ideal, it is closed under addition by construction; it is closed under multiplication by R because $R_i R_j \subseteq R_{i+j}$ for $i \geq 1$ implies $i + j \geq 1$.

Remark 1.93. The above ideal is called irrelevant because, in the case where $R = k[x_0, \dots, x_n]$,

$$Z(I) = \{(a_0 : \dots : a_n) \in \mathbb{P}^n(k) : f(a_0, \dots, a_n) \text{ for each homogeneous } f \in I\} = \varnothing.$$

Indeed, any element of $Z(I)$ would have to satisfy $x_i = 0$ for each x_i, which is illegal in projective space.

The point of the definition of a graded ideal is that, when $I \subseteq R$ is a graded ideal,

$$\frac{R}{I} \cong \bigoplus_{d=0}^{\infty} \frac{R_d}{R_d \cap I}$$

will also be a graded ring, with the given grading. This isomorphism comes from combining the isomorphisms $R \cong \bigoplus_d R_d$ and $I \cong \bigoplus_d (R_d \cap I)$.

Here is another ring-theoretic construction which we can grade.

Graded module

Definition 1.94 (Graded module)**.** Fix $R = R_0 \oplus R_1 \oplus \cdots$ a graded ring. Then an R-module M is *graded* if and only if we can write

$$M \cong \bigoplus_{d \in \mathbb{Z}} M_d$$

such that $R_i M_j \subseteq M_{i+j}$ for any $i \in \mathbb{N}$ and $j \in \mathbb{Z}$.

As a quick application, here is one reason to care about graded rings: they play nice with the Noetherian condition.

Proposition 1.95. A graded ring $R = R_0 \oplus R_1 \oplus \cdots$ is Noetherian if and only if R_0 is Noetherian and R is a finitely generated R_0-algebra.

Proof. The backwards direction is Proposition 1.48. For the forwards direction, take $R = R_0 \oplus R_1 \oplus \cdots$ a Noetherian, graded ring. We note that quotienting R by the irrelevant ideal reveals that R_0 is a quotient of R, so R_0 is a Noetherian ring.

It remains to show that R is a finitely generated R_0-algebra. The idea is to imitate the Hilbert's finiteness theorem. Before doing anything, we adopt the convention that, for an arbitrary element

$$f = f_0 + f_1 + \cdots \in R,$$

we let $\deg f$ equal the largest d for which $f_d \neq 0$.

We now proceed with the proof. Because R is Noetherian, the irrelevant ideal

$$I := R_1 \oplus R_2 \oplus \cdots$$

is finitely generated over R, so fix $I := (r_1, \ldots, r_N)$. We claim that

$$R \stackrel{?}{=} R_0[r_1, \ldots, r_N].$$

For $\supseteq$, there is nothing to say. For $\subseteq$, pick up some $f \in R$, and we show that $f \in R_0[r_1, \ldots, r_N]$. By decomposing f into its grading $f = f_0 + f_1 + \cdots$, we may assume that f lives in one of the R_d.

So now we induct on d. For $d = 0$, we have $f \in R_0 \subseteq R_0[r_1, \ldots, r_N]$ and are done immediately. So take $d > 0$. Then $f \in I = (r_1, \ldots, r_N)$, so we may write

$$f = \sum_{i=1}^{N} g_i r_i$$

for some $g_1, \ldots, g_N \in R$. By decomposing the $g_\bullet$ into their gradings, we see that we may assume that only the $\deg f - \deg r_i$ component is nonzero because all other components will cancel anyways.

In particular, g_i is homogeneous with degree $\deg f - \deg r_i$, so $g_i \in R_i$ with $i < d$. So by our induction, $g_i \in R_0[r_1, \ldots, r_N]$, and $f \in R_0[r_1, \ldots, r_N]$ by the decomposition of f in I. This finishes the proof. ∎

1.2.8 The Hilbert Function

For this subsection, let $R := k[x_0, \ldots, x_n]$ (note the zero-indexing!) be a ring graded by degrree, and let $M = \cdots \oplus M_{-1} \oplus M_0 \oplus M_1 \oplus \cdots$ be a finitely generated graded R-module. It follows that

$$\dim_k M_d < \infty$$

for each $d \in \mathbb{Z}$. Indeed, R is Noetherian, so M is Noetherian (M is finitely generated over R), so we note that the R-submodule

$$M'_d := \bigoplus_{e \geq d} M_e \subseteq M$$

is a finitely generated as an R-module. (This is an R-submodule because it is closed under addition, and $R_i M_j \subseteq M_{i+j}$ for $i \in \mathbb{N}$ and $j \in \mathbb{Z}$ gives closure under R-multiplication.) But the only way $rm \in M_d$ for $r \in R$ and $m \in M'_d$ is for $r \in R_0 = k$ and $m \in M_d$, so the (finite number of) generators of M'_d in M_d will generate M_d as a k-module.

This gives us the following definition.

Hilbert function

Definition 1.96 (Hilbert function)**.** Let M be a finitely generated module over $R := k[x_0, \dots, x_n]$, where R is graded by degree. Then we define the *Hilbert function* of M as

$$H_M(d) := \dim_k M_d.$$

Exercise 1.97. Let $M = R = k[x_0, \dots, x_n]$; i.e., view R as an R-module. Then we compute $H_M(d)$.

Proof. Here, M and R have the same grading (because $M = R$), so we are computing

$$H_M(d) = \dim_k R_d.$$

To see this, we note that we can expand any polynomial $f \in R_d$ as a unique k-linear combination of the degree-d monomials: after all, we can express generic polynomials in a unique k-linear combination of monomials, and R_d requires everything involved to have degree d.

Thus, $\dim_k R_d$ has basis consisting of the degree-d monomials in $k[x_0, \dots, x_n]$. Thus, we are counting tuples $(a_0, \dots, a_n)$ of nonnegative integers (uniquely) associated to the monomial

$$x_0^{a_0} \cdots x_n^{a_n}$$

such that $a_0 + \cdots + a_n = d$. But this is now merely a combinatorics problem! We claim that that this is $\binom{n+d}{d}$.

Indeed, for any such tuple $(a_0, \dots, a_n)$, imagine placing (in a single row) a_0 stones, then a stick, then a_1 stones, then a stick, and so on, ending by placing the last a_n stones. In total, we are placing d stones and n sticks, and the arrangement of sticks and stones uniquely describes the tuples. So now we see there are

$$\binom{n+d}{d}$$

ways to put down d sticks among $n + d$ "slots" of either sticks or stones. So indeed, we find that

$$\boxed{H_M(d) = \binom{n+d}{d}},$$

as desired. ∎

The above example found that $H_m(d)$ is a polynomial in d of degree r. This happens in general.

Theorem 1.98. Let M be a finitely generated graded module over the ring $R := k[x_0, \dots, x_n]$, where R is graded by degree. Then there exists a polynomial $P_M(d)$ of degree at most $n - 1$ which agrees with $H_M(d)$ for sufficiently large d.

Proof. The proof is by induction on n, where we will apply dimension-shifting of the grading for the inductive step. Our base case is $n = -1$, which makes M into a graded $R = R_0 = k$-vector space. But M is thus finite-dimensional, the summation

$$M = \bigoplus_{d \in \mathbb{Z}} M_d$$

of $R_0 = k$-vector spaces M_d must have only finitely many nonzero terms, so $H_M(d) = 0$ for sufficiently large d. So indeed, H_M agrees with the polynomial $P_M \equiv 0$ of degree $-\infty \leq -1$ for sufficiently large inputs.

Now, we will need to dimension-shift our grading in the proof that follows, so we have the following definition.

Twist

Definition 1.99 (Twist)**.** Given a graded R-module M, we define the dth *twist* $M(d)$ of M to be the same underlying module but with grading given by

$$M(d)_e := M_{d+e}.$$

To sanity check, we remark that $M = \bigoplus_{e \in \mathbb{Z}} M(d)_e = \bigoplus_{e \in \mathbb{Z}} M_{d+e}$ and $R_i M(d)_e = R_i M_{d+e} \subseteq M_{i+d+e} = M(d)_{i+e}$ verifies that we have in fact graded M.

Note the Hilbert function is well-behaved by shifting: $H_{M(d)}(e) = \dim_k M(d)_e = \dim_k M_{d+e} = H_M(e+d)$.

For the inductive step, the main point is to kill the x_n coordinate in creative ways. Namely, M being finitely generated over $k[x_0, \dots, x_n]$ implies that $M/x_n M$ will be finitely generated over $k[x_0, \dots, x_{n-1}]$ because any summation involving the x_n letter got killed. So we start with exact sequence

$$M \to M/x_n M \to 0.$$

We do take a moment to remark $M/x_r M$ is in fact a graded module by

$$\frac{M}{x_n M} \cong \frac{\bigoplus_{d \in \mathbb{Z}} M_d}{\bigoplus_{d \in \mathbb{Z}} x_n M_d} = \frac{\bigoplus_{d \in \mathbb{Z}} M_d}{\bigoplus_{d \in \mathbb{Z}} x_n M_{d-1}} \cong \bigoplus_{d \in \mathbb{Z}} \frac{M_d}{x_n M_{d-1}},$$

so $M \twoheadrightarrow M/x_n M$ is a map of graded modules. In particular, by disjointness, the pre-image of M_d under multiplication by x_n lives in M_{d-1}; note $x_n M_{d-1} \subseteq M_d$.

Now, to take our sequence backwards, we would like to prepend by $M \xrightarrow{x_n}$, but this is not legal because multiplication by x_n map will change the grading: we have $x_r M_{d-1} \subseteq M_d$. So instead we have to write down

$$M(-1) \xrightarrow{x_n} M \to M/x_n M \to 0.$$

This is in fact exact as graded modules because $M(-1)_d = M_{d-1}$ goes to $x_n M_{d-1}$ goes to 0 in $M_d/x_n M_{d-1}$.

To finish our short exact sequence, we let $K(-1)$ be the (twisted) kernel of $M(-1) \xrightarrow{x_n} M$ multiplication by x_n, and we get to write

$$0 \to K(-1) \to M(-1) \xrightarrow{x_n} M \to M/x_r M \to 0. \tag{$*$}$$

We take a moment to recognize $K(-1) \subseteq M(-1)$ is finitely generated over $k[x_0, \dots, x_n]$ because it is a submodule of the Noetherian module $M(-1)$. But any generator of $K(-1)$ multiplied by x_n will simply vanish, so the same generators will finitely generate $K(-1)$ over $k[x_0, \dots, x_{n-1}]$.

Now, taking the Hilbert function everywhere in $(*)$, counting dimensions gives

$$H_{K(-1)}(d) - H_{M(-1)}(d) + H_M(d) - H_{M/x_r M}(d) = 0.$$

We can rewrite this as

$$H_M(d) - H_M(d-1) = H_{M/x_n M}(d) - H_K(d-1),$$

so we see that the first finite difference of H_M agrees with $H_{M/x_n M}(d) - H_K(d-1)$, and the latter agrees with a polynomial of degree at most $n-1$ for sufficiently large d by inductive hypothesis. So theory of finite differences tells us that $H_M(d)$ will agree with a polynomial of degree at most n, finishing the induction. ∎

Theorem 1.98 justifies the following definition.

Hilbert polynomial

Definition 1.100 (Hilbert polynomial)**.** Let M be a finitely generated graded module over the ring $R := k[x_0, \dots, x_n]$, where R is graded by degree. The polynomial promised by Theorem 1.98 is called the *Hilbert polynomial* of M.

Remark 1.101. Geometrically, most of the time M will end up being the coordinate ring of a projective variety, in which case the degree of the above Hilbert polynomial is the "degree" of the projective variety. So heuristically, most of the time the degree of the Hilbert polynomial will not achieve its maximum.

Let's do some examples.

Exercise 1.102. Take $M := k[x,y,z]/\left(x^2+y^2+z^2\right)$ as a $R := k[x,y,z]$-submodule. We compute the Hilbert function for M.

Proof. For brevity, set $I := \left(x^2+y^2+z^2\right)$. Note that I is a graded ideal: if $finik[x,y,z]$ is divisible by $x^2+y^2+z^2$, then we can write $f(x,y,z) = \left(x^2+y^2+z^2\right) q(x,y,z)$. Expanding $q = q_0 + q_1 + \cdots$ into its homogeneous parts, we see that

$$f(x,y,z) = \sum_{d=2}^{\infty} \left(x^2+y^2+z^2\right) q_{d-2}$$

provides a decomposition of f into homogeneous parts, and by uniqueness this must be the decomposition of f. But each of these parts is manifestly divisible by $\left(x^2+y^2+z^2\right)$, so we have decomposed f into $(I \cap R_0) \oplus (I \cap R_1) \oplus \cdots$.

We have the following.

- We see $M_0 = R_0/(I \cap R_0)$ is simply k, so $\dim M_0 = 1$.
- We see $M_1 = R_1/(I \cap R_1)$ has basis $\{x,y,z\}$ because I hasn't killed anything yet, so it has dimension $\dim M_1 = 3$.
- We see R_2 has basis $\left\{xy, yz, zx, x^2, y^2, z^2\right\}$, but $z^2 \equiv -x^2 - y^2 \pmod{I}$ means that in $M_2 = R_2(I \cap R_2)$, we can kill z^2. However, we can do this anywhere else (more rigorous justification below), so $\dim M_2 = 5$.

For the general case, fix a degree $d \geq 2$. We note that there is a short exact sequence

$$0 \to R_{d-2} \xrightarrow{x^2+y^2+z^2} R_d \to \frac{R_d}{\left(x^2+y^2+z^2\right) R_{d-2}} \to 0.$$

Note the first map is well-defined because $\left(x^2+y^2+z^2\right) R_{d-2} \subseteq R_2 R_{d-2} \subseteq R_d$. In fact, we claim that $\left(x^2+y^2+z^2\right) R_{d-2} = I \cap R_d$, for any $f \in I \cap R_d$ has $f(x,y,z)/\left(x^2+y^2+z^2\right)$ homogeneous of degree $d-2$. So this short exact sequence is actually

$$0 \to R_{d-2} \to R_d \to M_d \to 0.$$

Thus, the short exact sequence gives $\dim M_d = \dim R_d - \dim R_{d-2}$, which by Exercise 1.97 is $\binom{n+2}{2} - \binom{n}{2} = \frac{n^2+3n+2}{2} - \frac{n^2-n}{2} = 2n+1$. ∎

Remark 1.103. Continuing with the previous remark, we see the degree of the Hilbert polynomial of M above is 1, so the associated projective variety $Z\left(x^2+y^2+z^2\right)$ ought have dimension 1. Well, $x^2+y^2+z^2=0$ defines a cone in affine 3-space (more or less), which is dimension one of projective 2-space upon recalling that lines becomes points.

Exercise 1.104 (Eisenbud 1.19)**.** Define $M := k[x,y,z]/\left(xz-y^2, yx-z^2, xw-yz\right)$ as a $R := k[x,y,z]$-module. We compute the Hilbert function for M.

Proof. We outline. For brevity, we set $I := \left(xz-y^2, yx-z^2, xw-yz\right)$. The key observation is that it happens that I is a free $k[x,w]$-module, with basis $\{1,y,z\}$.

Thus, viewing M as a $T := k[x,w]$-module, checking the basis, gives that $M = T \oplus T(-1) \oplus T(-1)$ corresponding to our basis elements $\{1,y,z\}$. (Multiplication by y or z will shift the grading, hence $T(-1)$.) It follows that the Hilbert function is $H_M(n) = 3n+1$. ■

We will start with localization next class.