1. Introduction

Practical information about the course.

Problem classes - 1 every 2 weeks, maybe a bit more
Coursework – two pieces, each worth 5%
Deadlines: 16 February, 14 March
Problem sheets and coursework will be available on my web page:
`http://wwwf.imperial.ac.uk/~morr/2016-7/teaching/alg-geom.html`
My email address: `m.orr@imperial.ac.uk`
Office hours: Mon 2-3, Thur 4-5 in Huxley 681

Bézout's theorem. Here is an example of a theorem in algebraic geometry and an outline of a geometric method for proving it which illustrates some of the main themes in algebraic geometry.

Theorem 1.1 (Bézout)**.** Let C be a plane algebraic curve

$$\{(x, y) : f(x, y) = 0\}$$

where f is a polynomial of degree m.

Let D be a plane algebraic curve

$$\{(x, y) : g(x, y) = 0\}$$

where g is a polynomial of degree n.

Suppose that C and D have no component in common (if they had a component in common, then their intersection would obviously be infinite).

Then $C \cap D$ consists of mn points, provided that

(i) we work over the complex numbers $\mathbb{C}$;
(ii) we work in the projective plane, which consists of the ordinary plane together with some points at infinity (this will be formally defined later in the course);
(iii) we count intersections with the correct multiplicities (e.g. if the curves are tangent at a point, it counts as more than one intersection). We will not define intersection multiplicities in this course, but the idea is that multiple intersections resemble multiple roots of a polynomial in one variable.

In the lecture, we motivated the various stipulations in this theorem by considering the cases where C is a line (degree 1) and D has either degree 1 or 2.

Outline of proof of Bézout's theorem. We prove a special case, where C is the union of m lines, then use this to prove the general case of the theorem.

First for the special case: Suppose we have m lines in the plane, with equations

$$a_1x + b_1y + c_1 = 0, \ldots, a_mx + b_my + c_m = 0.$$

We can multiply these equations together to get

$$(a_1x + b_1y + c_1)(a_2x + b_2y + c_2) \cdots (a_mx + b_my + c_m) = 0.$$

This is an equation of degree m and its solution set is the union of the lines.

Each line intersects D in n points (counted with multiplicities), because we can rearrange the equation of the line into the form $x = \cdots$ or $y = \cdots$ then substitute into the equation for D. This usually gives a polynomial of degree n in one variable, and this has n roots if we count them correctly. There are also special cases to worry about where the line intersects D at infinity.

Combining all the m lines, we deduce that their union intersects D in mn points.

Now we deduce the general case from the special case.

We let the curve C vary in a family of curves of degree m. What exactly we mean by "varying in a family" will be defined later in the course. As an example, consider the family of curves

$$\mathcal{F} : \{(x, y) : x^2 - y^2 = t\}$$

where t is a parameter – for different values of t we get different curves.

When the curve C varies in a family like this, the number of intersection points in $C \cap D$ does not change (counting with multiplicity). This is the core of the proof; it requires a lot of work to justify which we will not do here.

For any degree m curve C, it is possible to find a family of curves which contains both C itself and a union of m lines X. For example, if C is the hyperbola defined by the equation $x^2 - y^2 = 1$, then it is found in the family $\mathcal{F}$ (with $t = 1$). If we let $t = 0$ in this family, then the equation factors as

$$(x - y)(x + y) = 0$$

and this defines the union of two lines in the plane.

We have already proved that $X \cap D$ has mn points, and we stated that $X \cap D$ has the same number of points as $C \cap D$ because C and X are in the same family. We conclude that $C \cap D$ has mn points. $\square$

The idea that something stays the same everywhere, or almost everywhere, in a family of varying algebraic sets is a key theme in algebraic geometry.

Note that this proof uses not just curves but also higher-dimensional algebraic sets: instead of thinking of thinking about a family of curves such as $\mathcal{F}$, with coordinates (x, y) and a parameter t, we can regard x, y, t all as coordinates in three-dimensional space and consider the surface

$$\{(x, y, t) : x^2 - y^2 = t\}.$$

Then we use facts about this surface as part of the proof.

We will not prove Bézout's theorem in this course – in particular, we will not define intersection multiplicities. But we will set up many of the tools needed to fill in the gaps in this outline proof.

2. Affine algebraic sets

Course outline.

(1) Affine varieties – definition, examples, maps between varieties, translating between geometry and commutative algebra (the Nullstellensatz)
(2) Projective varieties – definition, examples, maps between varieties, rigidity and images of maps
(3) Dimension – several different definitions (all equivalent, but useful for different purposes), calculating dimensions of examples
(4) Smoothness and singularities – definition, examples, key theorems
(5) Examples of varieties and interesting theorems (depending on how much time is left)

What is not in the course?

(1) Schemes
(2) Sheaves and cohomology
(3) Divisors and curves

The base field. Let k be an algebraically closed field.

We are going to be thinking about solutions to polynomials, so everything is much simpler over algebraically closed fields (we already saw this in Bézout's theorem). Number theorists might be interested in other fields, but you generally have to start by understanding the algebraically closed case first. In this course we will stop with the algebraically closed case too.

Apart from being algebraically closed, it usually does not matter much which field we use to do algebraic geometry – except sometimes it matters what is the characteristic of the field. In this course we will mostly stick to characteristic zero, on those occasions where it matters. You will not lose much if you just assume that $k = \mathbb{C}$ throughout the course (except occasionally when it will be explicitly something else).

Indeed it is often useful to think about $k = \mathbb{C}$ because then you can use your usual geometric intuition. When I draw pictures on the whiteboard, I am usually only drawing the real solutions because it is hard to draw shapes in $\mathbb{C}^2$. This is cheating but it is often very useful – the real solutions are not the full picture but in many cases we can still see the important features there.

Definition of affine algebraic sets.

Definition. Algebraic geometers write $\mathbb{A}^n$ to mean k^n, and call it **affine n-space**.

You may think of this as just a funny choice of notation, but there are at least two reasons for it:

(i) When we write k^n, it makes us think of a vector space, equipped with operations of addition and scalar multiplication. But $\mathbb{A}^n$ means just a set of points, described by coordinates $(x_1, \ldots, x_n)$ with $x_i \in k$, without the vector space structure.

(ii) Because it usually doesn't matter much what our base field k is (as long as it is algebraically closed), it is convenient to have notation which does not prominently mention k.

On occasions when it *is* important to specify which field k we are using, we write $\mathbb{A}^n_k$ for affine n-space.

Definition. An **affine algebraic set** is a subset $V \subseteq \mathbb{A}^n$ which consists of the common zeros of some finite set of polynomials $f_1, \ldots, f_m$ with coefficients in k.

More formally, an **affine algebraic set** is a set of the form

$$V = \{(x_1, \ldots, x_n) \in \mathbb{A}^n : f_1(x_1, \ldots, x_n) = \cdots = f_m(x_1, \ldots, x_n) = 0\}$$

for some polynomials $f_1, \ldots, f_m \in k[X_1, \ldots, X_n]$.

Examples of affine algebraic sets.

Exercise 2.1. Think of some examples and non-examples of affine algebraic sets.

These are the examples and non-examples you came up with in lectures. I apologise if I forgot any – tell me and I will add them.

Examples.

(1) The empty set, defined by the polynomial $f_1 = 3$ (for example).
(2) The whole space $\mathbb{A}^n$, defined by the polynomial $f_1 = 0$.
(3) Any finite set $\{a_1, \ldots, a_n\}$ in $\mathbb{A}^1$, defined by the polynomial equation

$$(X - a_1)(X - a_2) \cdots (X - a_n) = 0.$$

More generally, any finite set in $\mathbb{A}^n$: see below.

(4) The twisted cubic curve

$$\{(t, t^2, t^2) \in \mathbb{A}^3 : t \in k\} = \{(x_1, x_2, x_3) \in \mathbb{A}^3 : y - x^2 = z - x^3 = 0\}.$$

(5) Embeddings of $\mathbb{A}^m$ in $\mathbb{A}^n$ where $m < n$:

$$\{(x_1, \ldots, x_m, 0, \ldots, 0) \in \mathbb{A}^n\} = \{(x_1, \ldots, x_n) \in \mathbb{A}^n : x_{m+1} = \cdots = x_n = 0\}.$$

More generally, the image of a linear map $\mathbb{A}^m \to \mathbb{A}^n$:

$$\{(x_1, \ldots, x_n) \in \mathbb{A}^n : \text{some linear conditions}\}.$$

Non-examples.

(1) A line segment in $\mathbb{A}^1$ or $\mathbb{A}^n$.
(2) A line with a double point (we don't have a definition for this at the moment, but whatever it means, it is not an algebraic set!)
(3) An infinite discrete set.

(4) A sine wave. You can prove this from (3): If $\{(x, y) : y = \sin x\}$ were an affine algebraic set, then $\{(x, y) : y = \sin x, y = 0\}$ would also be an affine algebraic set because it is defined by imposing an extra polynomial condition, but the latter is an infinite discrete set.

To prove that a single point $(a_1, \ldots, a_n) \in \mathbb{A}^n$ is an affine algebraic set, write it as

$$\{(x_1, \ldots, x_n) \in \mathbb{A}^n : x_1 - a_1 = \cdots = x_n - a_n = 0\}.$$

Note that this is different from the example of a finite set in $\mathbb{A}^1$, because that example had a single polynomial in one variable of degree n, while here we have n distinct polynomials in n variables of degree 1.

Questions.

(1) Prove that any finite set in $\mathbb{A}^n$ is an affine algebraic set.
(2) Prove that line segments and infinite discrete sets are not affine algebraic sets in $\mathbb{A}^1$ (or even in $\mathbb{A}^n$ if you want – we don't yet have the tools to prove that an infinite discrete set in $\mathbb{A}^n$ is not an affine algebraic set).

3. Unions, intersections and ideals

Infinite subsets of $\mathbb{A}^1$. At the end of the last lecture I asked how to prove that infinite discrete sets or line segments are not affine algebraic sets in $\mathbb{A}^1$. We can do this using this lemma.

Lemma 3.1. Every affine algebraic set in $\mathbb{A}^1$, other than $\mathbb{A}^1$ itself, is finite.

Proof. Suppose our affine algebraic set is

$$\{x \in \mathbb{A}^1 : f_1(x) = f_2(x) = \cdots = f_m(x) = 0\}.$$

The one-variable polynomial f_1 has only finitely many roots. Imposing additional polynomial conditions can only make the set smaller. □

This also tells us that

$$\{x \in \mathbb{A}^1 : x \neq 0\}$$

is not an affine algebraic set. However there is an affine algebraic set which is "isomorphic" to $\mathbb{A}^1 \setminus \{0\}$, namely

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

By looking at just the x coordinate, this set bijects to $\mathbb{A}^1 \setminus \{0\}$. (I am using the word "isomorphism" informally here as we have not yet defined it.)

Philosophical remark. The words "affine variety" mean more or less the same thing as "affine algebraic set" but there is an ontological difference. "Affine algebraic set" means a subset which lives inside $\mathbb{A}^n$ and knows how it lives inside $\mathbb{A}^n$, while "affine variety" means an object in its own right which is considered outside of $\mathbb{A}^n$. Thus we can definitely say that $\mathbb{A}^1 \setminus \{0\}$ is not an affine algebraic set, because it knows it lives inside $\mathbb{A}^1$ and we can use Lemma 3.1; but we might say that $\mathbb{A}^1 \setminus \{0\}$ is an affine variety because there is a way to re-interpret it inside $\mathbb{A}^2$ and get an affine algebraic set. I will try to use these words consistently, but the difference is quite subtle and books may not always use it consistently.

Note that some books (e.g. Reid, Hartshorne) have another difference between affine varieties and affine algebraic sets – they require varieties to be irreducible (which we will define next time). Other books (e.g. Shafarevich) do not require varieties to be irreducible. In this course we will *not* require varieties to be irreducible.

Finite sets. Consider two points $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ in $\mathbb{A}^n$. As we saw last time, each point is an affine algebraic set: $(a_1, \ldots, a_n)$ is defined by the equations

$$X_1 - a_1 = 0, \ldots, X_n - a_n = 0$$

and $(b_1, \ldots, b_n)$ is defined by the equations

$$X_1 - b_1 = 0, \ldots, X_n - b_n = 0.$$

The two-point set $\{(a_1, \ldots, a_n), (b_1, \ldots, b_n)\}$ can be defined by taking the product for each possible pair of equations, one from each list:

$$(X_i - a_i)(X_j - b_j) = 0 \text{ for all } i, j \in \{1, \ldots, n\}.$$

Note that it is necessary to consider *all* the pairs between the lists, not just the ones with $i = j$, because otherwise we would be allowing points like $(a_1, \ldots, a_{n-1}, b_n)$.

New algebraic sets from old.

Unions. The above example of two points generalises. If $V, W \subseteq \mathbb{A}^n$ are affine algebraic sets, then their union $V \cup W \subseteq \mathbb{A}^n$ is also an affine algebraic set. To prove this, we have to take the product for each possible pair of defining equations: if

$$\begin{aligned} V &= \{(x_1, \ldots, x_n) \in \mathbb{A}^n : f_1(\underline{x}) = \cdots = f_r(\underline{x}) = 0\}, \\ W &= \{(x_1, \ldots, x_n) \in \mathbb{A}^n : g_1(\underline{x}) = \cdots = g_s(\underline{x}) = 0\}. \end{aligned}$$

then $V \cup W$ is defined by the equations $f_i(\underline{x})g_j(\underline{x}) = 0$ for $1 \leq i \leq r$, $1 \leq j \leq s$.

Let's check that these equations really do define $V \cup W$.

First: suppose that $\underline{x} \in V \cup W$. Then either

(1) $\underline{x} \in V$, so $f_i(\underline{x}) = 0$ for every i, so we can multiply by $g_j(\underline{x})$ to get $f_i(\underline{x})g_j(\underline{x}) = 0$ for every i and j;
(2) or $\underline{x} \in W$, in which case the same argument works with g_j in place of f_i.

The other direction is a little trickier. Suppose that we have $\underline{x} \in \mathbb{A}^n$ satisfying $f_i(\underline{x})g_j(\underline{x}) = 0$ for all i and j. Looking just at f_1, we get:

$$\begin{aligned} f_1 g_1(\underline{x}) &= 0, \text{ so } f_1(x) = 0 \text{ or } g_1(x) = 0. \\ f_1 g_2(\underline{x}) &= 0, \text{ so } f_1(x) = 0 \text{ or } g_2(x) = 0. \\ &\vdots \\ f_1 g_s(\underline{x}) &= 0, \text{ so } f_1(x) = 0 \text{ or } g_s(x) = 0. \end{aligned}$$

Putting these all together, we get

$$f_1(\underline{x}) = 0 \text{ or } g_j(\underline{x}) = 0 \text{ for every } j.$$

We can do the same thing for f_2 to get

$$f_2(\underline{x}) = 0 \text{ or } g_j(\underline{x}) = 0 \text{ for every } j$$

and so on for each f_i. Putting all these together, we get

$$f_i(\underline{x}) = 0 \text{ for every } i \text{ or } g_j(\underline{x}) = 0 \text{ for every } j.$$

This says precisely that $\underline{x} \in V \cup W$.

Intersections. If $V, W \subseteq \mathbb{A}^n$ are affine algebraic sets, then their intersection $V \cap W \subseteq \mathbb{A}^n$ is also an affine algebraic set. To prove this, just combine the lists of defining equations. That is, say

$$V = \{(x_1, \dots, x_n) \in \mathbb{A}^n : f_1(\underline{x}) = \cdots = f_r(\underline{x}) = 0\},$$
$$W = \{(x_1, \dots, x_n) \in \mathbb{A}^n : g_1(\underline{x}) = \cdots = g_s(\underline{x}) = 0\}.$$

Then

$$V \cap W = \{(x_1, \dots, x_n) \in \mathbb{A}^n : f_1 = \cdots = f_r = g_1 = \cdots = g_s = 0\}.$$

Products. If $V \subseteq \mathbb{A}^m$ and $W \subseteq \mathbb{A}^n$ are affine algebraic sets, then their Cartesian product $V \times W \subseteq \mathbb{A}^{m+n}$ is an affine algebraic set. Write

$$V = \{(x_1, \dots, x_m) \in \mathbb{A}^m : f_1(\underline{x}) = \cdots = f_r(\underline{x}) = 0\},$$
$$W = \{(y_1, \dots, y_n) \in \mathbb{A}^n : g_1(\underline{y}) = \cdots = g_s(\underline{y}) = 0\}.$$

Then

$$V \times W = \{(x_1, \dots, x_m, y_1, \dots, y_n) \in \mathbb{A}^{m+n} : f_1(\underline{x}) = \cdots = f_r(\underline{x}) = g_1(\underline{y}) = \cdots = g_s(\underline{y}) = 0\}.$$

Note that this time the equations for V involve different variables than the equation for W.

Exercise 3.1. Is the union of infinitely many affine algebraic sets an affine algebraic set? (No! For example, an infinite discrete set in $\mathbb{A}^1$.)

Is the intersection of infinitely many affine algebraic sets an affine algebraic set?

Ideals. In order to show that the intersection of infinitely many affine algebraic set is an affine algebraic set, we cannot simply combine the lists of defining equations: in our definition we required that an affine algebraic set can be defined by a *finite* list of equations. We introduce ideals to remove this restriction.

Definition. (Recall from Commutative Algebra.) If R is a ring, an **ideal** is a subset $I \subseteq R$ with the properties that:

(1) if $f, g \in I$, then $f + g \in I$;
(2) if $f \in I$ and $q \in R$, then $qf \in I$.

Given any subset $S \subseteq R$, we define the **ideal generated by** S to be the smallest ideal which contains S, and denote it by (S). In particular, if S is the finite set $\{f_1, \dots, f_m\}$ then it generates the ideal

$$(f_1, \dots, f_m) = \{q_1 f_1 + \cdots + q_m f_m : q_1, \dots, q_m \in R\}.$$

Introduce some notation. For *any* set $S \subseteq k[X_1, \dots, X_n]$, let

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

Lemma 3.2. If $S \subseteq k[X_1, \dots, X_n]$ generates the ideal I, then $\mathbb{V}(S) = \mathbb{V}(I)$.

Proof. We have $S \subseteq I$ and so it is easy to see that $\mathbb{V}(I) \subseteq \mathbb{V}(S)$.

Suppose that $\underline{x} \in \mathbb{V}(S)$, and $f \in \mathbb{V}(I)$. Then there are $f_1, \ldots, f_m \in S$ and $q_1, \ldots, q_m \in k[X_1, \ldots, X_n]$ such that

$$f = q_1 f_1 + \cdots + q_m f_m.$$

Since $f_1(\underline{x}) = \cdots = f_m(\underline{x}) = 0$, it follows that $f(\underline{x}) = 0$.

Since this holds for every $f \in I$, $\underline{x} \in \mathbb{V}(I)$. □

Hilbert Basis Theorem.

Theorem 3.3 (Hilbert Basis Theorem)**.** (From Commutative Algebra)

If k is any field, then the polynomial ring $k[X_1, \ldots, X_n]$ is **noetherian**. That means that the following two equivalent conditions hold:

(1) Let I be an ideal in $k[X_1, \ldots, X_n]$. Then there exists a finite set
$$\{f_1, \ldots, f_m\} \subseteq k[X_1, \ldots, X_n]$$
which generates I.

(2) Let $I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots$ be an ascending chain of ideals in $k[X_1, \ldots, X_n]$. Then there is some N such that $I_n = I_N$ for every $n > N$.

Corollary 3.4. $\mathbb{V}(S)$ is an affine algebraic set for *any* set of polynomials $S \subseteq k[X_1, \ldots, X_n]$.

Proof. Let I be the ideal generated by S. By the Hilbert Basis Theorem (statement 1), we can choose a finite set $\{f_1, \ldots, f_m\}$ which generates I. Then Lemma 3.2 tells us that

$$\mathbb{V}(S) = \mathbb{V}(I) = \mathbb{V}(f_1, \ldots, f_m).$$ □

Now we are able to verify that an intersection of infinitely many affine algebraic sets is an affine algebraic set, because it is defined by the union of the lists of defining equations for the individual algebraic sets.

Question. What is the translation of Hilbert Basis Theorem (statement 2) into affine algebraic sets?

4. Zariski topology and irreducible sets

Ideals and algebraic sets: back and forth. The following statement is the translation into affine algebraic sets of statement 2 of the Hilbert basis theorem. Note that the direction of inclusion is reversed when we go from ideals to algebraic sets: if $I_1 \subseteq I_2$, then $\mathbb{V}(I_2) \subseteq \mathbb{V}(I_1)$.

Proposition 4.1. Let $V_1 \supseteq V_2 \supseteq V_3 \supseteq \cdots$ be a descending chain of affine algebraic sets in $\mathbb{A}^n$.

Then there exists N such that $V_n = V_N$ for all $n > N$.

But how to prove this?

Exercise 4.1. Say $V_n = \mathbb{V}(I_n)$. Does $V_1 \supseteq V_2$ imply that $I_1 \subseteq I_2$?

Answer: No. For example, $I_1 = (X)$ and $I_2 = (X^2)$ in $k[X]$. We have $\mathbb{V}(I_1) = \{0\} = \mathbb{V}(I_2)$.

The problem is that there is more than one ideal defining the same algebraic set: for example, (X) and (X^2). However, there is a natural choice we can make for one ideal canonically associated with an affine algebraic set: the set of *all* polynomials which vanish on that set.

Formally, if A is a subset of $\mathbb{A}^n$ (usually A will be an affine algebraic set), we define

$$\mathbb{I}(A) = \{f \in k[X_1, \dots, X_n] : f(\underline{x}) = 0 \text{ for all } \underline{x} \in A\}.$$

This is an ideal in $k[X_1, \dots, X_n]$.

We have now defined two functions

$$\mathbb{V} : \{\text{ideals in } k[X_1, \dots, X_n]\} \to \{\text{affine algebraic sets in } \mathbb{A}^n\},$$
$$\mathbb{I} : \{\text{affine algebraic sets in } \mathbb{A}^n\} \to \{\text{ideals in } k[X_1, \dots, X_n]\}.$$

These are not inverses: the example of (X) and (X^2) shows that we do not always have $\mathbb{I}(\mathbb{V}(I)) = I$. But composing them in the other order gives the identity.

Lemma 4.2. If V is an affine algebraic set, then $\mathbb{V}(\mathbb{I}(V)) = V$.

Proof. It is clear that $V \subseteq \mathbb{V}(\mathbb{I}(V))$ (and this works when V is any subset of $\mathbb{A}^n$, not necessarily algebraic).

For the reverse inclusion, we have to use the hypothesis that V is an affine algebraic set. By the definition of affine algebraic set, $V = \mathbb{V}(J)$ for some ideal $J \subseteq k[X_1, \dots, X_n]$.

Suppose that $\underline{y} \notin V$. We shall show that $\underline{y} \notin \mathbb{V}(\mathbb{I}(V))$.

Because $\underline{y} \notin V = \mathbb{V}(J)$, there exists $f \in J$ such that $f(\underline{y}) \neq 0$. Now $J \subseteq \mathbb{I}(V)$ and so $f \in \mathbb{I}(V)$. Hence $f(\underline{y}) \neq 0$ tells us that $\underline{y} \notin \mathbb{V}(\mathbb{I}(V))$. □

Proof of Proposition 4.1. The fact that

$$V_1 \supseteq V_2 \supseteq V_3 \supseteq \cdots$$

implies that

$$\mathbb{I}(V_1) \subseteq \mathbb{I}(V_2) \subseteq \mathbb{I}(V_3) \subseteq \cdots .$$

By the Hilbert Basis Theorem (statement 2), there exists N such that $\mathbb{I}(V_n) = \mathbb{I}(V_N)$ for all $n > N$.

Since $\mathbb{V}(\mathbb{I}(V_n)) = V_n$ for every n, this proves the proposition. □

Statement of the Nullstellensatz. When does $\mathbb{I}(\mathbb{V}(I)) = I$? It turns out that the only reason that this can fail is where elements of the ideal I have n-th roots which are not in I, just as with the example of $I = (X^2)$ where $X^2 \in I$ has a square root X which is not in I.

To state this precisely, we need the following definitions.

Definition. Let I be an ideal in a ring R. The **radical** of I is

$$\operatorname{rad} I = \sqrt{I} = \{f \in R : \exists n > 0 \text{ s.t. } f^n \in I\}.$$

We say that I is a **radical ideal** if $\operatorname{rad} I = I$.

Note that, if I is any ideal, then $\operatorname{rad} I$ is always a radical ideal.

Note that, to calculate $\operatorname{rad} I$, we need to add in n-th roots of all elements of I, not just the generators. For example, if $I = (X, Y^2 - X) \subseteq k[X, Y]$, then we can rewrite this as $I = (X, Y^2)$ and so $\operatorname{rad} I = (X, Y) \neq I$, even though neither of the original generators of I had any non-trivial n-th roots.

Theorem 4.3 (Hilbert's Nullstellensatz)**.** Let I be any ideal in the polynomial ring $k[X_1, \dots, X_n]$ over an algebraically closed field k. We have

$$\mathbb{I}(\mathbb{V}(I)) = \operatorname{rad} I.$$

We'll prove this in a few lectures' time (not because we need to develop more theory, just because I would like to introduce some more concepts first which will allow us to do more with examples).

Zariski topology. In the previous lecture, we proved that affine algebraic sets in $\mathbb{A}^n$ satisfy the following conditions:

(i) A finite unions of affine algebraic sets is an affine algebraic set.
(ii) Any intersection affine algebraic sets is an affine algebraic set.
(iii) $\mathbb{A}^n$ and $\emptyset$ are affine algebraic sets. (this was in lecture 2)

These are precisely the conditions satisfied by the *closed* sets in a topological space. Therefore, we can define a topological space in which the underlying set is $\mathbb{A}^n$ and the closed sets are the affine algebraic sets. This is called the **Zariski topology**.

This is a very different topology from the ones you are used to in analysis! In particular, it is a very long way from being Hausdorff.

Exercise 4.2. Describe the Zariski closed and Zariski open sets in $\mathbb{A}^1$. Prove that the Zariski topology on $\mathbb{A}^1$ is not Hausdorff.

At the moment, the Zariski topology is likely to seem very strange. It might also seem like: what is the point of such a strange topology? We will not use it in a very deep way, it is just a convenient language to be able to talk about open and closed sets. (It does get used more seriously in the theory of schemes.)

Irreducible sets. Recall the definition of a connected topological space.

Definition. A topological space S is **connected** if it is not possible to write it as the union of two disjoint non-empty open sets.

This is equivalent to: it is not possible to write S as the union of two disjoint non-empty closed sets.

It is possible to talk about connectedness in the Zariski topology. For example, a finite set of points of size greater than 1 is not connected in the Zariski topology (every subset is closed!) The union of two disjoint lines $\{(x, y) \in \mathbb{A}^2 : x(x-1) = 0\}$ is not connected (each line is a non-empty closed subset).

But there is a more refined notion for the Zariski topology.

Exercise 4.3. The set $\{(x, y) \in \mathbb{A}^2 : xy = 0\}$ has "two pieces": it is a union of two lines (intersecting at the origin).

Describe the Zariski closed subsets of $\{(x, y) \in \mathbb{A}^2 : xy = 0\}$. Is this a connected set in the Zariski topology?

The following notion gives us a way of formally understanding the example described in Exercise 4.3.

Definition. A topological space S is **irreducible** if it is not possible to write it as the union $S_1 \cup S_2$ of two closed sets, unless at least one of S_1 and S_2 is equal to S itself. (Compared to the second definition of *connected*, we no longer require S_1 and S_2 to be disjoint.)

The opposite: A topological space S is **reducible** if there exist closed sets $S_1, S_2 \subseteq S$ such that $S = S_1 \cup S_2$, and neither S_1 nor S_2 is equal to S.

This is not a very useful notion for the topological spaces we consider in analysis. For example, considering the real line with its usual topology, we can write it as a union of proper closed subsets:

$$\mathbb{R} = \{x \in \mathbb{R} : x \leq 1\} \cup \{x \in \mathbb{R} : x \geq 1\}$$

These subsets are not disjoint because they intersect at 1. Of course, there are many other ways to write $\mathbb{R}$ as a union of proper closed subsets in the usual topology; the same is true for any other Hausdorff space.

5. Irreducible components

Examples of irreducible sets. At the end of the last lecture I was asked:

Is the hyperbola $H = \{(x, y) : xy - 1 = 0\}$ irreducible?

The picture is misleading: it appears to have two pieces, but this is because we only draw the real solutions. For algebraic geometry, we need to look at complex solutions, and then there is only one piece. (This can be seen informally by recalling that, if we project down to the x coordinate, H bijects with $\mathbb{A}^1 \setminus \{0\}$. And $\mathbb{C} \setminus \{0\}$ is connected in the usual (analytic) topology on $\mathbb{C}$.)

For a formal proof, recall the definition of irreducible set: S is **irreducible** if we cannot write $S = S_1 \cup S_2$ where S_1 and S_2 are proper closed subsets of S (in the Zariski topology).

In order to test this, we need to work out the Zariski closed subsets of H.

Well, say that $V \subseteq H$ is a proper Zariski closed subset. Since $V \neq H$ there is some polynomial $f \in k[X, Y]$ which vanishes on V but does not vanish on all of H.

Because $V \subseteq H$ and $y = 1/x$ on H, we have

$$f(x, y) = f(x, 1/x) \text{ when } (x, y) \in V.$$

Now $f(X, 1/X)$ is almost a polynomial in the single variable X, except that it may contain negative powers of X:

$$f(X, 1/X) = \sum_{n \in \mathbb{Z}} a_n X^n.$$

We can multiply up by X^m where $-m$ is the lowest exponent of X which appears in this expression. Then $X^m f(X, 1/X)$ is a polynomial in X, which vanishes on V.

Furthermore $f(X, 1/X)$ is not identically zero because f does not vanish identically on H. Hence $X^m f(X, 1/X)$ is a non-zero single-variable polynomial, therefore it has only finitely many roots.

The roots of $X^m f(X, 1/X) = 0$ are the possible x-coordinates for points in V. For each value of x, there is only one possible y such that $(x, y) \in V$ because $y = 1/x$ on V. Therefore V is finite.

Thus we have shown that all proper Zariski closed subsets of H are finite. In particular, if V_1, V_2 are two proper Zariski closed subsets of H, they are both finite and so their union is finite. Hence $V_1 \cup V_2 \neq H$ so H is irreducible.

The same argument as in the previous paragraph also shows that $\mathbb{A}^1$ is irreducible, because we showed in lecture 4 that all proper Zariski closed subsets of $\mathbb{A}^1$ are finite.

Zariski topology on an affine algebraic set. Last time we defined the Zariski topology on $\mathbb{A}^n$. I forgot to say that we can also define the Zariski topology on any affine algebraic set $V \subseteq \mathbb{A}^n$. It is defined as the subspace topology on V, that is:

> A subset of V is Zariski closed in V if and only if it is the intersection between V and a Zariski closed subset of $\mathbb{A}^n$.

Because V itself is Zariski closed in $\mathbb{A}^n$, this is equivalent to:

> A subset of V is Zariski closed in V if and only if it is Zariski closed in $\mathbb{A}^n$.

Thus for closed sets it does not matter whether we say "Zariski closed in V" or "Zariski closed in $\mathbb{A}^n$".

On the other hand, for open sets "Zariski open in V" does not mean the same thing as "Zariski open in $\mathbb{A}^n$". For example, $H \setminus \{(1,1)\}$ is Zariski open in H but it is not Zariski open in $\mathbb{A}^2$ (there is no non-zero polynomial which vanishes on $\mathbb{A}^2 \setminus (H \setminus \{(1,1)\})$).

We can rewrite the definition of irreducible sets in terms of open subsets instead of closed subsets (in this lemma, we must use sets which are open for the topology on S itself, not for the ambient topology on $\mathbb{A}^n$):

Lemma 5.1. The following conditions on a topological space S are equivalent to irreducibility:

(i) Every pair of non-empty open subsets $U_1, U_2 \subseteq S$ have non-empty intersection.
(ii) Every non-empty open subset of S is dense in S.

Corollary 5.2. Let S be a irreducible topological space and $U \subseteq S$ a non-empty open subset. Then U is irreducible (in the subspace topology).

The corollary shows that $\mathbb{A}^1 \setminus \{0\}$ is irreducible, because it is an open subset of $\mathbb{A}^1$. Compare this to the fact that the hyperbola H is irreducible. Note that knowing that $\mathbb{A}^1 \setminus \{0\}$ does not mean we can skip the work we did in the proof that H is irreducible, because you need to do exactly the same work to prove that the Zariski topology on $H \subseteq \mathbb{A}^2$ and on $\mathbb{A}^1 \setminus \{0\} \subseteq \mathbb{A}^1$ are the same.

Irreducible components. If an affine algebraic set is reducible, then we can write it as a union of proper closed subsets. If these subsets are reducible, then we can write them in turn as unions of proper closed subsets. The following proposition says that this process eventually stops: after finitely many steps, we reach irreducible sets.

Proposition 5.3. Every affine algebraic set can be written as a union of finitely many irreducible closed subsets.

Proof. Suppose that V is an affine algebraic set which cannot be written as a union of finitely many irreducible closed subsets.

V itself must be reducible (otherwise we could write it as a union of one irreducible closed subset!) So $V = V_1 \cup W_1$, with V_1 and W_1 proper closed subsets of V.

V_1 and W_1 cannot both satisfy the proposition, otherwise we could write each of them as a union of finitely many irreducible closed subsets. Then taking the union of those decompositions would give us V as a union of finitely many irreducible closed subsets.

Thus at least one of V_1 and W_1 does not satisfy the proposition. Without loss of generality, we may suppose that V_1 does not satisfy the proposition.

Then V_1 must be reducible, so we can write $V_1 = V_2 \cup W_2$. We can repeat the argument: at least one of V_2 and W_2 does not satisfy the proposition, without loss of generality V_2, etc.

Thus we build up a chain of closed subsets $V \supset V_1 \supset V_2 \supset V_3 \supset \cdots$ where all these sets do not satisfy the proposition, and all the inclusions are strict.

This contradicts Proposition 4.1 (derived from the Hilbert Basis Theorem). $\square$

Thus for any affine algebraic set V, we can write

$$V = V_1 \cup V_2 \cup \cdots \cup V_n$$

where V_i are irreducible closed subsets. We may also assume that V_i is not contained in V_j for any $i \neq j$ (otherwise we could just drop V_i from the list without changing the union).

Subject to this non-redundancy condition, there is only one way to write V as a finite union of irreducible closed subsets (proof: Shafarevich section 3.1, Theorem 1.5). We call the V_i which appear in this decomposition the **irreducible components** of V.

For example: the irreducible components of the set $\{(x, y) : xy = 0\}$ are the lines $x = 0$ and $y = 0$.

Question. If V is an affine algebraic set, what condition on the ideal $\mathbb{I}(V)$ is equivalent to V being irreducible?

6. Regular functions

Prime ideals and irreducible sets.

Definition. (from Commutative Algebra) An ideal I in a ring R is a **prime ideal** if $I \neq R$ and for every $f, g \in R$, if $fg \in I$, then $f \in I$ or $g \in I$ (or both).

Lemma 6.1. An affine algebraic set $V \subseteq \mathbb{A}^n$ is irreducible if and only if $\mathbb{I}(V)$ is a prime ideal in $k[X_1, \dots, X_n]$.

Proof. First suppose that V is irreducible. Suppose we have $f, g \in k[X_1, \dots, X_n]$ such that $fg \in \mathbb{I}(V)$. Let

$$V_1 = \{\underline{x} \in V : f(\underline{x}) = 0\}, \ V_2 = \{\underline{x} \in V : g(\underline{x}) = 0\}.$$

For every $\underline{x} \in V$, $f(\underline{x})g(\underline{x}) = 0$ and hence either $f(\underline{x}) = 0$ or $g(\underline{x}) = 0$. Thus for every $\underline{x} \in V$, either $\underline{x} \in V_1$ or $\underline{x} \in V_2$. In other words, $V = V_1 \cup V_2$. Furthermore V_1 and V_2 are closed subsets of V. Hence as V is irreducible, either $V_1 = V$ or $V_2 = V$. If $V_1 = V$ then $f \in \mathbb{I}(V)$ and if $V_2 = V$ then $g \in \mathbb{I}(V)$.

Now suppose that V is reducible. Then we can write it as a union $V_1 \cup V_2$ of proper closed subsets. Since V_1 is a proper closed subset of V, there exists some $f \in k[X_1, \dots, X_n]$ vanishing on V_1 but not on all of V. Similarly there exists g vanishing on V_2 but not on all of V. Thus neither f nor g is in $\mathbb{I}(V)$, but the product fg vanishes on $V_1 \cup V_2$ and hence we have $fg \in \mathbb{I}(V)$. Thus $\mathbb{I}(V)$ is not prime. □

We define a **hypersurface** to be an affine algebraic set in $\mathbb{A}^n$ defined by *one* polynomial equation, that is,

$$\{\underline{x} \in \mathbb{A}^n : f(\underline{x}) = 0\}$$

for some $f \in k[X_1, \dots, X_n]$. It follows from Lemma 6.1 together with Hilbert's Nullstellensatz that a hypersurface defined by a polynomial f is irreducible if and only if f is a power of an irreducible polynomial.

For example, we can use this to prove that the circle

$$\{(x, y) : x^2 + y^2 = 1\}$$

is irreducible, by proving that the polynomial $X^2 + Y^2 - 1$ is irreducible. This is because, if $f = X^2 + Y^2 - 1 = f_1 f_2$ then we can scale f_1 and f_2 by constants to get

$$f_1 = X + g_1(Y), \quad f_2 = X + g_2(Y)$$

(since f has degree 2 in X and its X^2 term has coefficient 1). Since f has no X term, we must have $g_1 + g_2 = 0$. But then

$$f_1 f_2 = (X + g_1(Y))(X - g_1(Y)) = X^2 - g_1(Y)^2$$

so $g_1(Y)^2 = -Y^2 + 1$, and $-Y^2 + 1$ is not a square.

On the other hand, the hypersurface

$$\{(x, y) : x^2 + y^2 = 0\}$$

is reducible, because X^2+Y^2 factors as $(X-iY)(X+iY)$.

Regular functions. So far we have only considered algebraic sets as sets, sitting individually. Now we look at functions between them. Just as one uses continuous functions for topological spaces, holomorphic functions for complex manifolds, homomorphisms for groups, etc., so algebraic geometry has its own type of functions – regular functions. Of course, these are given by polynomials.

Definition. Let $V \subseteq \mathbb{A}^n$ be an affine algebraic set. A **regular function** on V is a function $f\colon V \to k$ such that there exists a polynomial $F \in k[X_1,\dots,X_n]$ with $f(x) = F(x)$ for all $x \in V$.

Note that the polynomial F is not uniquely determined by the function f: $F, G \in k[X_1,\dots,X_n]$ determine the same regular function on V if and only if $F - G$ vanishes on V, that is iff $F - G \in \mathbb{I}(V)$.

Definition. The regular functions on V form a k-algebra: they can be added, multiplied by each other and multiplied by scalars in k. This is called the **coordinate ring** of V and denoted $k[V]$.

There is a ring homomorphism $k[X_1,\dots,X_n] \to k[V]$ which sends a polynomial F to the function $F_{|V}$ which it defines on V. This homomorphism is surjective and its kernel is $\mathbb{I}(V)$, so

$$k[V] \cong k[X_1,\dots,X_n]/\mathbb{I}(V).$$

Exercise 6.1. What are the coordinate rings of the following affine algebraic sets?

(i) $\mathbb{A}^n$.
(ii) A point.
(iii) $\{x \in \mathbb{A}^1 : x(x-1) = 0\}$ (two points).
(iv) $\{(x,y) \in \mathbb{A}^2 : xy = 0\}$ (two intersecting lines).

Answers.

(i) $k[X_1,\dots,X_n]$.
(ii) k. A regular function on a point is just a single value.
(iii) $k \times k$. A regular function on two points is determined by two scalars, namely its value on each of the two points. For any pair of values $(a,b) \in k \times k$, one can easily write down a polynomial $f \in k[X]$ such that $f(1) = a$ and $f(0) = b$. Alternatively, one can check algebraically that the map
$$(a,b) \mapsto (a-1)X + b \bmod (X(X-1))$$
is a k-algebra isomorphism $k \times k \to k[X_1,\dots,X_n]/(X(X-1))$.
(iv) $\{(f,g) \in k[X] \times k[Y] : f(0) = g(0)\}$.
One can also interpret this as

$$k[X,Y]/(XY) \cong \Big\{a_0 + \sum_{r=1}^{m} b_r X^r + \sum_{s=1}^{n} c_s Y^s : a_0, b_1, \dots, b_m, c_1, \dots, c_n \in k, m, n \in \mathbb{N}\Big\}.$$

To compare these two interpretations, observe that

$$k[X] = \Big\{a_0 + \sum_{r=1}^{m} b_r X^r\Big\}, \quad k[Y] = \Big\{a_0 + \sum_{s=1}^{n} c_s Y^s\Big\},$$

and the condition that $f(0) = g(0)$ is equivalent to insisting that these two polynomials have the same constant coefficient a_0.

Example (iii) generalises: if V is a disconnected affine algebraic set, we can write V as a union $V_1 \cup V_2$ of disjoint Zariski closed subsets, and then

$$k[V] = k[V_1] \times k[V_2].$$

On the other hand, if V is merely reducible, so that the sets V_1 and V_2 are not disjoint, then $k[V]$ is a subset of $k[V_1] \times k[V_2]$ (see example (iv)).

Question. What is the coordinate ring of the hyperbola $\{(x, y) \in \mathbb{A}^2 : xy = 1\}$?

7. Regular maps

Coordinate ring of the hyperbola. Last time I ended by asking: what is the coordinate ring of the hyperbola $V = \{(x, y) \in \mathbb{A}^2 : xy = 1\}$? Of course, it is the quotient ring $k[X, Y]/(XY - 1)$.

To describe this more explicitly, observe that any term $a_{i,j}X^rY^s$ of a two-variable polynomial is congruent (mod $XY - 1$) to either $a_{r,s}X^{r-s}$ (if $r \geq s$) or $a_{r,s}Y^{s-r}$ (if $s > r$). Thus every coset in $k[X, Y]/(XY - 1)$ has a representative of the form

$$\sum_{i=0}^{m} a_i X^i + \sum_{j=1}^{n} a_j Y^j.$$

The polynomials of this form determine different functions on V, so we have written down exactly one representative of each coset.

Furthermore, since $XY = 1$ in $k[V]$, we may relabel Y as X^{-1}; then the multiplication rule will be what the notation leads us to expect. So we can write

$$k[V] = k[X, X^{-1}] = \Big\{ \sum_{j=-n}^{m} a_j X^m : a_{-n}, \dots, a_m \in k, m, n \in \mathbb{N} \Big\}.$$

Note that an affine algebraic set V is irreducible if and only if $k[V]$ is an integral domain, because we know that V is irreducible if and only if $\mathbb{I}(V)$ is a prime ideal in $k[X_1, \dots, X_n]$.

Regular maps. A regular function goes from an algebraic set V to the field k. We can also define regular maps, which go from one algebraic set V to another algebraic set W.

Definition. Let $V \subseteq \mathbb{A}^n$ and $W \subseteq \mathbb{A}^m$ be affine algebraic sets. A **regular function** $f \colon V \to W$ is a function $V \to W$ such that there exist polynomials $F_1, \dots, F_m \in k[X_1, \dots, X_n]$ such that

$$f(\underline{x}) = (F_1(\underline{x}), \dots, F_m(\underline{x}))$$

for all $\underline{x} \in V$.

Regular maps are often called **morphisms**.

Note that in order to check that a given list of polynomials $F_1, \dots, F_m$ defines a regular map $V \to W$, it is necessary to check that $(F_1(\underline{x}), \dots, F_m(\underline{x})) \in W$ for every $\underline{x} \in V$. Equivalently, we can check that the regular functions $F_{1|V}, \dots, F_{m|V} \in k[V]$ satisfy the equations

$$g(X_1, \dots, X_m) = 0$$

in the coordinate ring $k[V]$, for each polynomial $g \in \mathbb{I}(W)$.

Examples.

(1) Let $V \subseteq \mathbb{A}^n$ be an affine algebraic set. For any $m < n$, the projection $\pi\colon V \to \mathbb{A}^m$ defined by

$$\pi(x_1, \dots, x_n) = (x_1, \dots, x_m)$$

is a regular map.

(2) A regular function on V is the same thing as a regular map $V \to \mathbb{A}^1$.

(3) Let $C = \{(x, y) : y^2 = x^3\}$. Then $t \mapsto (t^2, t^3)$ is a regular map $\mathbb{A}^1 \to C$.

(4) Consider SL_n, the set of $n \times n$ matrices with determinant 1. This is an affine algebraic set in $\mathbb{A}^{n^2}$ because the determinant is a polynomial in the entries of a matrix. The map $a \mapsto a^{-1}$ is a regular map $\mathrm{SL}_n \to \mathrm{SL}_n$: Cramer's rule tells us how to write each entry of a^{-1} as a polynomial in the entries of a divided by $\det a$, and because we are only considering $a \in \mathrm{SL}_n$ we can drop the division by $\det a$.

Isomorphisms. We say that a regular map $f\colon V \to W$ is an **isomorphism** if there exists a regular map $g\colon W \to V$ such that $g \circ f = \mathrm{id}_V$ and $f \circ g = \mathrm{id}_W$.

Example. If V is the parabola $\{(x, y) : y - x^2 = 0\}$, then the regular map $f\colon V \to \mathbb{A}^1$ given by

$$f(x, y) = x$$

is an isomorphism because it has an inverse $g\colon \mathbb{A}^1 \to V$ given by

$$g(x) = (x, x^2).$$

Example. On the other hand, if H is the hyperbola $\{(x, y) : xy = 1\}$, then the projection $(x, y) \mapsto x$ is not an isomorphism $H \to \mathbb{A}^1$ because it is not surjective so it cannot possibly have an inverse. This is not enough to prove that H is not isomorphic to $\mathbb{A}^1$, because maybe there is some other regular map $H \to \mathbb{A}^1$ which is an isomorphism. (We will prove later that H is not isomorphic to $\mathbb{A}^1$.)

Example. Consider the affine algebraic set $W = \{(x, y) : y^2 - x^3 = 0\}$. The regular map $f\colon \mathbb{A}^1 \to W$ given by

$$f(t) = (t^2, t^3)$$

is a bijection but it is not an isomorphism. Note that we should expect W not to be isomorphic to $\mathbb{A}^1$ because it has a singularity at the origin.

To prove that $f\colon \mathbb{A}^1 \to W$ is not an isomorphism: Consider a regular map $g\colon W \to \mathbb{A}^1$. It must be given by a polynomial $g(X, Y) \in k[X, Y]$ and so

$$g \circ f(t) = g(t^2, t^3)$$

is a polynomial in t which can have a constant term and terms of degree 2 or greater, but no term of degree 1. Hence we cannot find g such that $g \circ f(t) = t$.

Regular maps and the coordinate ring. Suppose we have a regular map $\varphi\colon V \to W$ between affine algebraic sets. For each regular function g on W, we get a regular function $\varphi^* g$ on V defined by

$$(\varphi^* g)(x) = g(\varphi(x)).$$

We call $\varphi^* g \in k[V]$ the **pull-back** of $g \in k[W]$.

Thus φ induces a k-algebra homomorphism

$$\varphi^*\colon k[W] \to k[V].$$

Note that φ^* goes in the opposite direction to φ.

We can use this notion of pull-back to prove that regular maps are continuous in the Zariski topology. If Z is a Zariski closed subset of W, defined by the vanishing of some regular functions $g_1, \dots, g_n \in W$, then

$$\varphi^{-1}(Z) = \{x \in V : (\varphi^* g_1)(x) = 0, \dots, (\varphi * g_n)(x) = 0\}$$

and thus $\varphi^{-1}(Z)$ is a Zariski closed subset of V.

If we have two regular maps $\varphi\colon V \to W$ and $\psi\colon W \to Z$, then we can compose them to get $\psi \circ \varphi\colon V \to Z$. One can easily check that the associated pullback maps on coordinate rings satisfy

$$(\psi \circ \varphi)^* = \varphi^* \circ \psi^*\colon k[Z] \to k[V]. \tag{*}$$

For those who know category theory, we say that $V \mapsto k[V]$ is a contravariant functor

$$\{\text{affine algebraic sets}\} \to \{k\text{-algebras}\}.$$

In particular, (*) tells us that if $\varphi\colon V \to W$ is an isomorphism with inverse $\psi\colon W \to V$, then $\psi^* \circ \varphi^* = \mathrm{id}$ and $\varphi^* \circ \psi^* = \mathrm{id}$. Thus if V and W are affine algebraic sets which are isomorphic, then their coordinate rings $k[V]$ and $k[W]$ are isomorphic.

Example. Now we can prove that the hyperbola H is not isomorphic to $\mathbb{A}^1$, because $k[H] = k[X, X^{-1}]$ is not isomorphic to $k[\mathbb{A}^1] = k[X]$. To verify that these k-algebras are not isomorphic, observe that in $k[X]$ the only invertible elements are the scalars, while $k[X, X^{-1}]$ contains a non-scalar invertible element, namely X.

Example. We can similarly prove that $\mathbb{A}^1$ is not isomorphic to the singular cubic $W = \{(x, y) : y^2 = x^3\}$. We saw earlier that $k[W]$ is the ring of polynomials in one variable with no term of degree 1, that is,

$$k[W] = \{a_0 + \sum_{r=2}^{m} a_r X^r : a_0, a_2, \dots, a_r \in k\}.$$

To prove that $k[W]$ is not isomorphic to $k[\mathbb{A}^1] = k[X]$, observe that $k[X]$ is a unique factorisation domain but $k[W]$ is not because $(X^2)^3 = (X^3)^2$, and X^2 and X^3 are both irreducible in $k[W]$.

8. Rational maps

Informally, rational maps are "maps" between varieties defined by polynomial fractions, for example the "function" $x \mapsto 1/x$ on $\mathbb{A}^1$. Observe that this is not really a function $\mathbb{A}^1 \to \mathbb{A}^1$ because it is not defined at $x = 0$, but it is a genuine function on the Zariski open subset $\mathbb{A}^1 \setminus \{0\}$.

Just as with regular functions and regular maps, we first define rational functions, which take values in k, then rational maps, which go into any algebraic set.

Rational functions. Let V be an irreducible affine algebraic set. We define rational functions only for irreducible algebraic sets because, as we saw in the example of $1/x$, a rational function defines a genuine function on a Zariski open subset of V, and irreducibility guarantees that all open subsets of V are dense in V (so that a function defined on an open subset really is defined "almost everywhere" on V).

Definition. The **function field** of V is the field of fractions of the coordinate ring $k[V]$. We denote this by $k(V)$.

Note that $k[V]$ is an integral domain because V is irreducible, and therefore $k[V]$ has a field of fractions.

For example, the function field of $\mathbb{A}^1$ is $k(X)$, the fraction field of the polynomial ring $k[X]$.

Definition. A **rational function** on V is an element of the function field $k(V)$. Thus a rational function can be written in the form f/g, where f and g are regular functions. There may be many different choices for f and g which define the same rational function f/g.

We say that a rational function $\varphi \in k(V)$ is **regular** at a point $x \in V$ if there exist regular functions $f, g \in k[V]$ such that $\varphi = f/g$ and $g(x) \neq 0$. If $\varphi = f/g$ is regular at $x \in V$, then it has a value $\varphi(x) = f(x)/g(x)$.

Note that we are allowed to choose different fractions f/g representing φ at different points $x \in V$, in order to show that those points are regular. The value $\varphi(x)$ is independent of which fraction representing φ we choose, as long as it has $g(x) \neq 0$.

For example, consider the algebraic set defined by the equation $XY = ZT$ in $\mathbb{A}^4$. Let

$$\varphi = X/Z \in k(V).$$

The defining equation implies that we also have

$$\varphi = T/Y.$$

Looking at the fraction X/Z shows us that φ is regular wherever $Z \neq 0$, and looking at the fraction T/Y shows us that φ is regular wherever $Y \neq 0$. On the other hand, φ is not regular on the closed subset $Y = Z = 0$. (One can verify that there is no other fraction representing φ which is non-zero on this closed subset.)

The set of points where φ is regular is called the **domain of definition** of φ, and denoted $\operatorname{dom}\varphi$.

Lemma 8.1. The domain of definition of a rational function $\varphi \in k(V)$ is a non-empty Zariski open subset of V.

Proof. Consider the set of all possible fractions f/g with $f, g \in k[V]$ representing $\varphi \in k(V)$. The set of points at which φ is *not* regular is the intersection of the Zariski closed sets $\{x \in V : g(x) = 0\}$ across all these fractions. Hence the set of points at which φ is not regular is a Zariski closed subset of V. The domain of definition is the complement of this set, and therefore is Zariski open.

To show that the domain of definition is non-empty, pick a single fraction f/g representing $\varphi \in k(V)$. The regular function g is not equal to zero as an element of $k[V]$ (by the definition of the field of fractions), so $\{x \in V : g(x) = 0\}$ is a proper closed subset of V. The domain of definition contains the complement of this set, namely $\{x \in V : g(x) \neq 0\}$, and hence is non-empty. □

Note that every regular function $f \in k[V]$ is also a rational function $f/1 \in k(V)$, and its domain of definition is all of V. The converse also holds.

Lemma 8.2. Let $\varphi \in k(V)$ be a rational function whose domain of definition is equal to V. Then φ is a regular function on V.

Proof. Since $\operatorname{dom}\varphi = V$, for each point $x \in V$, we can choose regular functions $f_x, g_x \in k[V]$ such that $\varphi = f_x/g_x$ and $g_x(x) \neq 0$. Let $I \subseteq k[V]$ denote the ideal generated by the functions g_x. Because $k[V]$ is noetherian, we can pick finitely many of these functions $g_{x_1}, \dots, g_{x_m}$ which still generate I.

Because, for each $x \in V$, there is some $g_x \in I$ which is non-zero at x, the Zariski closed subset of V defined by I is empty. Then it follows from the Nullstellensatz that $I = k[V]$.

In particular, $1 \in I = (g_{x_1}, \dots, g_{x_m})$. Thus there exist $u_1, \dots, u_m \in k[V]$ such that

$$1 = u_1 g_{x_1} + \dots + u_m g_{x_m} \text{ in } k[V].$$

We can now calculate

$$\varphi = 1.\varphi = (u_1 g_{x_1} + \dots + u_m g_{x_m})\varphi = u_1 f_{x_1} + \dots + u_m f_{x_m}.$$

Since $u_i, f_{x_i} \in k[V]$, so is φ.

Note that it might appear that we have only proved the above equation

$$\varphi = u_1 f_{x_1} + \dots + u_m f_{x_m}$$

on a Zariski open subset of V, namely the intersections of the domains of definition of $f_{x_1}/g_{x_1}, \dots, f_{x_m}/g_{x_m}$. Because V is irreducible, this open subset must be dense; but the subset where an equation of polynomials holds is closed, so it is equal to all of V. (I got this argument wrong originally, claiming that you could use the continuity of regular maps but this doesn't work; compare with Lemma 14.1.) $\square$

Rational maps. Let $V \subseteq \mathbb{A}^n$, $W \subseteq \mathbb{A}^m$ be irreducible affine algebraic sets.

Definition. A **rational map** $\varphi \colon V \dashrightarrow W$ is an m-tuple of rational functions $\varphi_1, \dots, \varphi_m \in k(V)$ such that, for all points $x \in V$ where all of $\varphi_1, \dots, \varphi_m$ are regular, the point $(\varphi_1(x), \dots, \varphi_m(x))$ is in W.

We use the broken arrow symbol $\dashrightarrow$ instead of the usual arrow because a rational map is not a function on V in the usual set-theoretic sense. It only defines a function $U \to W$ where U is the domain of definition of φ.

The **domain of definition** of a rational map $\varphi \colon V \dashrightarrow W$ is defined to be the intersection of the domains of definition of the component rational functions $(\varphi_1, \dots, \varphi_m)$. Because each the domain of definition of each φ_i is a non-empty Zariski open subset of V, and because V is irreducible, the intersection of these domains of definition is also a non-empty Zariski open subset of V.

Example. An important example of a rational map is projection from a point onto a hyperplane.

Let H be a hyperplane in $\mathbb{A}^n$ (that is, a set defined by a single *linear* equation). Let p be a point in $\mathbb{A}^n \setminus H$.

For simplicity, we shall assume that p is the origin and that

$$H = \{(x_1, \dots, x_n) \in \mathbb{A}^n : x_n = 1\}.$$

(We could always reduce to this case by a suitable change of coordinates.)

Let us also write H_0 for the hyperplane through p parallel to H, that is,

$$H_0 = \{(x_1, \dots, x_n) \in \mathbb{A}^n : x_n = 0\}.$$

For each point $x \in \mathbb{A}^n \setminus H_0$, let L_x denote the line which passes through p and x. Since $x \notin H_0$, L_x intersects H in exactly one point. We call this point $\varphi(x)$.

We can write this algebraically as

$$\varphi(x_1, \dots, x_n) = (x_1/x_n, \dots, x_{n-1}/x_n, 1)$$

and so φ is a rational map $\mathbb{A}^n \dashrightarrow H$. This map is called projection from p onto H.

9. More rational and regular maps

Composing rational maps.

Example. Let V be the circle $\{(x, y) : x^2 + y^2 = 1\}$. Consider the projection from the point $p = (1, 0)$ on to the line $x = 0$ (in the lecture, I said on to the line $x = -1$ but then the formulas would be slightly different). This is a rational map $\pi\colon V \dashrightarrow \mathbb{A}^1$ with the formula

$$\pi(x, y) = y/(1 - x).$$

We can see geometrically that this projection induces a bijection between the circle (excluding p) and the line. If we compute the formula for the inverse map, we get

$$t \mapsto \Big(\frac{t^2 - 1}{t^2 + 1}, \frac{2t}{t^2 + 1}\Big),$$

a well-known parameterisation of the circle. Thus we see that the inverse is a rational map $\varphi\colon \mathbb{A}^1 \dashrightarrow V$. Note that φ is not regular at $t = \pm i$ – we don't see this on the usual picture, which only shows the real points.

We would like to define formally what we mean by saying that the rational maps φ and π are inverse to each other, taking into account that they are not true functions between the sets V and $\mathbb{A}^1$ because they are not regular everywhere.

In order to do this, we first define what it means to compose rational maps. But it does not always make sense to compose rational maps. For example, consider the rational map $\mathbb{A}^2 \dashrightarrow \mathbb{A}^1$ defined by

$$(x, y) \mapsto \frac{1}{1 - x^2 - y^2}.$$

This map is not defined anywhere on the circle V, and hence cannot be composed with φ (considered as a rational map $\mathbb{A}^1 \dashrightarrow \mathbb{A}^2$). To get around this problem, we first define dominant rational maps.

Definition. The **image** of a rational map $\varphi\colon V \dashrightarrow W$ is the set of points

$$\{\varphi(x) \in W : x \in \operatorname{dom} \varphi\}.$$

We say that a rational map is **dominant** if its image is Zariski dense in W.

For example, φ is dominant if we consider it as a rational map $\mathbb{A}^1 \dashrightarrow V$ but it is not dominant if we consider it as a rational map $\mathbb{A}^1 \dashrightarrow \mathbb{A}^2$. (This is like surjectivity: whether a function is surjective or not depends on what codomain you use.)

Let V, W, T be irreducible affine algebraic sets. If $\varphi\colon V \dashrightarrow W$ is a *dominant* rational map and $\psi\colon W \dashrightarrow T$ is a rational map (ψ is not required to be dominant), then it makes sense to compose them because we know that $\operatorname{dom} \psi$ is a Zariski open subset of W, while $\operatorname{im} \varphi$ is a Zariski dense subset of W and so

$$\operatorname{dom} \psi \cap \operatorname{im} \varphi \neq \emptyset.$$

Thus there are at least some points where $\psi \circ \varphi$ is defined. One can check (by writing out ψ in terms of fractions of polynomials, then substituting in fractions of polynomials representing φ) that $\psi \circ \varphi$ is a rational map $V \dashrightarrow T$.

Definition. Rational maps $\varphi\colon V \dashrightarrow W$ and $\psi\colon W \dashrightarrow V$ are **rational inverses** if both are dominant and $\varphi \circ \psi = \mathrm{id}_W$ and $\psi \circ \varphi = \mathrm{id}_V$ (where these composite rational maps are defined).

A rational map $\varphi\colon V \dashrightarrow W$ is a **birational equivalence** if it is dominant and has a rational inverse.

We say that irreducible algebraic sets V and W are **birational** if there exists a birational equivalence $V \dashrightarrow W$.

Our earlier example showed that the circle is birational to $\mathbb{A}^1$.

Another example is the cuspidal cubic

$$W = \{(x, y) : y^2 = x^3\}.$$

This is also birational to $\mathbb{A}^1$, as shown by the rational maps

$$W \dashrightarrow \mathbb{A}^1 : (x, y) \mapsto y/x,$$
$$\mathbb{A}^1 \dashrightarrow W : t \mapsto (t^2, t^3).$$

Birationally equivalent affine algebraic sets look the same "almost everywhere." For example, the cuspidal cubic is the same as the affine line except for its singularity at the origin.

On the other hand, $\mathbb{A}^1$ is not birationally equivalent to $\mathbb{A}^2$ or to an elliptic curve

$$\{(x, y) : y^2 = f(x)\} \text{ where } f \text{ is a cubic polynomial with no repeated roots.}$$

We will prove this later in the course once we have more tools.

If $\varphi\colon V \dashrightarrow W$ is a dominant rational map, then we can use it to pull back rational functions from W to V (just like we used regular maps to pull back regular functions). We get a k-homomorphism of fields

$$\varphi^*\colon k(W) \to k(V)$$

defined by $\varphi^*(g) = g \circ \varphi$. (A k-homomorphism means that φ^* restricts to the identity on the copies of k which are contained in $k(W)$ and $k(V)$, namely the constant functions.)

If φ is a birational equivalence, then φ^* is a k-isomorphism of fields.

From algebra homomorphisms to regular maps. We have seen that each regular map $f\colon V \to W$ induces a k-algebra homomorphism $f^*\colon k[W] \to k[V]$, and that each dominant rational map $\varphi\colon V \dashrightarrow W$ induces a k-field homomorphism $\varphi^*\colon k(W) \to k(V)$. We can also carry out these constructions in the reverse direction: starting with a k-algebra homomorphism and getting a regular map, or similarly for rational maps.

Observe that if $f\colon V \to W$ is a regular map, we can recover f from $f^*\colon k[W] \to k[V]$ by taking the coordinate functions $X_1, \dots, X_m \in k[W]$ on W and pulling them back to get

$$f_1 = f^*(X_1),\ \dots,\ f_m = f^*(X_m) \in k[V].$$

These are precisely the regular functions on V such that $f = (f_1, \dots, f_m)$.

We generalise this to: starting from an arbitrary k-algebra homomorphism $\alpha\colon k[W] \to k[V]$, we define a regular map $s\colon V \to W$ by

$$s = (\alpha(X_1), \dots, \alpha(X_m)).$$

Then $\alpha = s^*\colon k[W] \to k[V]$.

Thus every k-algebra homomorphism $k[W] \to k[V]$ is the pull back by some regular map $V \to W$. We conclude:

Proposition 9.1. $\varphi \mapsto \varphi^*$ is a bijection

$$\{\text{regular maps } V \to W\} \longrightarrow \{k\text{-algebra homomorphisms } k[W] \to k[V]\}.$$

Corollary 9.2. Affine algebraic sets V and W are isomorphic if and only if their coordinate rings $k[V]$ and $k[W]$ are isomorphic as k-algebras.

The moral is that if we only care about affine algebraic sets up to isomorphism, then the coordinate rings contain exactly the same information as the algebraic sets themselves (in the language of category theory, the functor $V \mapsto k[V]$ is fully faithful).

One can do the same thing for rational maps:

Proposition 9.3. $\varphi \mapsto \varphi^*$ is a bijection

$$\{\text{rational maps } V \dashrightarrow W\} \longrightarrow \{k\text{-field homomorphisms } k(W) \to k(V)\}.$$

Corollary 9.4. Affine algebraic sets V and W are birationally equivalent if and only if their function fields $k(V)$ and $k(W)$ are k-isomorphic.

Question. We have seen that $V \mapsto k[V]$ leads to bijections on maps between affine algebraic sets. To fully understand the relationship between affine algebraic sets and k-algebras, there is one more question to ask:

Which k-algebras can occur as $k[V]$ where V is an affine algebraic set?

10. Equivalence between algebra and geometry

Example. In lecture 6, we calculated the ring of regular functions on a union of two intersecting lines:

$$k[X,Y]/(XY) = \{(f,g) \in k[X] \times k[Y] : f(0) = g(0)\}.$$

Someone asked me at the office hour whether the same thing works for any union of intersecting affine algebraic sets. The answer turns out to be no.

I think the simplest example is $V = \mathbb{V}(Y(Y - X^2) \subseteq \mathbb{A}^2$. This is the union of a line and a parabola, tangent to each other at the origin. We know that the parabola is isomorphic to $\mathbb{A}^1$, so each irreducible component of V has coordinate ring $k[X]$.

But

$$k[V] \ncong \{(f,g) \in k[X] \times k[X] : f(0) = g(0)\}.$$

Intuitively, because the two components are tangent to each other at $(0,0)$, any polynomial in $k[X,Y]$ must have the same derivative along the line and along the parabola at $(0,0)$. Hence we should expect that

$$k[V] \cong \{(f,g) \in k[X] \times k[X] : f(0) = g(0) \text{ and } f'(0) = g'(0)\}.$$

Exercise. Prove this description for $k[V]$. Prove that this ring is not isomorphic to $k[XY]/(XY)$.

Reduced finitely generated k-algebras. We write down some algebraic properties which obviously hold for $A = k[V]$, the coordinate ring of an affine algebraic set V:

1. A is finitely generated, because if $V \subseteq \mathbb{A}^n$ then A is generated by the coordinate functions $X_1, \dots, X_n$.
2. A is reduced (meaning that if $f \in A$ and $f^k = 0$ for some $k > 0$, then $f = 0$). This is because A is a ring of functions in the usual set-theoretic sense: if $f^k = 0$ then $f(x)^k = 0$ for all $x \in V$, so $f(x) = 0$ for all $x \in V$.

These properties are enough to completely characterise coordinate rings of affine algebraic sets (this uses the Nullstellensatz).

Proposition 10.1. Let A be a finitely generated reduced k-algebra. Then there exists an affine algebraic set V such that $k[V] \cong A$.

Proof. Pick a finite set $f_1, \dots, f_n \in A$ which generates A as a k-algebra. We can define a homomorphism $\varphi \colon k[X_1, \dots, X_n] \to A$ by $X_1 \mapsto f_1, \dots, X_n \mapsto f_n$.

Let $I = \ker \varphi$ and let $V = \mathbb{V}(I) \subseteq \mathbb{A}^n$.

The homomorphism φ is surjective because $f_1, \dots, f_n$ generate A, and so

$$A \cong k[X_1, \dots, X_n]/I.$$

Thus $k[X_1, \dots, X_n]/I$ is a reduced k-algebra. It follows that I is a radical ideal.

Hence the Nullstellensatz tells us that $I = \mathbb{I}(V)$. Thus

$$k[V] \cong k[X_1, \dots, X_n]/\mathbb{I}(V) \cong k[X_1, \dots, X_n]/I \cong A. \qquad \square$$

The notion of affine variety. Often in mathematics, it is convenient to consider objects only "up to isomorphism." For example, one might talk about "the group with 7 elements," ignoring the fact that there are many different groups with 7 elements because they are all isomorphic to each other (and therefore they all behave in the same ways).

Similarly, in algebraic geometry we often want to consider affine algebraic sets up to isomorphism. But affine algebraic sets are always defined in a concrete way: they are a subset of some specific affine space $\mathbb{A}^n$. (It is as if we had defined all finite groups to be subgroups of a symmetric group S_n.) And we have seen that affine algebraic sets can be isomorphic even when they appear to be quite different as subsets of affine space, for example the line $\mathbb{A}^1$ is isomorphic to the parabola $\mathbb{V}(Y - X^2) \subseteq \mathbb{A}^2$. Thus it is useful to use different terminology: we talk about "affine algebraic sets" when we mean subsets of $\mathbb{A}^n$, and we talk about "affine varieties" when we mean an affine algebraic set up to isomorphism, forgetting its embedding into $\mathbb{A}^n$.

Proposition 10.1 is more naturally stated in terms of affine varieties rather than affine algebraic sets: in the proof we had to choose a generating set for A, for which there is no distinguished choice. Different choices of generating set would lead to isomorphic affine algebraic sets, but embedded differently into affine space. So it is better to say that A is the coordinate ring of some affine variety V, with no distinguished choice of embedding into $\mathbb{A}^n$.

(I mentioned this philosophy about affine varieties before in lecture 3, and I will mention it again after we have defined quasi-projective varieties.)

For those who know some fancy categorical language, we can sum up all the results on the equivalence between affine geometric objects and their coordinate rings by saying that $V \mapsto k[V]$ is an *equivalence of categories*

$$\{\text{affine varieties over } k\} \longrightarrow \{\text{reduced finitely generated } k\text{-algebras}\}^{\mathrm{op}}$$

where the superscript "op" indicates that the directions of morphisms are reversed.

Dictionary between algebraic subsets and ideals. Let A be a reduced finitely generated k-algebra and V an affine variety such that $A \cong k[V]$. How can we work out the geometry of V from the algebra of A?

One example is working out the Zariski closed subsets of V. If we choose an embedding of V into $\mathbb{A}^n$, then we get an isomorphism

$$k[X_1, \dots, X_n]/\mathbb{I}(V) \to A.$$

We have proved that

$$\{\text{Zariski closed subsets of } V\} \longleftrightarrow \{\text{radical ideals in } k[X_1, \dots, X_n] \text{ containing } \mathbb{I}(V)\}.$$

But

$$\{\text{ideals in } k[X_1,\dots,X_n] \text{ containing } \mathbb{I}(V)\} \longleftrightarrow \{\text{ideals in } k[X_1,\dots,X_n]/\mathbb{I}(V)\}.$$

We conclude that

$$\{\text{Zariski closed subsets of } V\} \longleftrightarrow \{\text{radical ideals in } A\}.$$

We also get:

$$\{\text{irreducible Zariski closed subsets of } V\} \longleftrightarrow \{\text{prime ideals in } A\}$$

and

$$\{\text{points of } V\} \longleftrightarrow \{\text{maximal ideals in } A\}.$$

The Weak and Strong Nullstellensatz. There are many ways of proving the Nullstellensatz, all of which require some difficult algebra. We will roughly follow the method in Shafarevich (Appendix A), which incorporates the hard algebra into one statement which we can quote, and then do the rest as geometrically as possible.

Recall the statement of Hilbert's Nullstellensatz, also called the Strong Nullstellensatz.

Theorem 10.2 (Strong Nullstellensatz)**.** Let I be any ideal in the polynomial ring $k[X_1,\dots,X_n]$ over an algebraically closed field k. We have

$$\mathbb{I}(\mathbb{V}(I)) = \operatorname{rad} I.$$

In order to prove this, we will first prove a weaker version, which is called the Weak Nullstellensatz, then use that to deduce the Strong Nullstellensatz.

Theorem 10.3 (Weak Nullstellensatz)**.** Let I be an ideal in the polynomial ring $k[X_1,\dots,X_n]$ over an algebraically closed field k. If I is not equal to $k[X_1,\dots,X_n]$ itself, then $\mathbb{V}(I) \neq \emptyset$.

This is a statement about the existence of solutions to polynomial equations, so it is necessary to require k to be algebraically closed. As an example to show that it fails when k is not algebraically closed, consider the ideal (X^2+Y^2+1) in $\mathbb{R}[X,Y]$. This ideal is not the full polynomial ring, but there are no real solutions to the equation $x^2+y^2+1=0$.

Note that the Strong Nullstellensatz easily implies the Weak Nullstellensatz: if $\mathbb{V}(I)=\emptyset$ then the Strong Nullstellensatz tells us that

$$\operatorname{rad} I = \mathbb{I}(\emptyset) = k[X_1,\dots,X_n].$$

In particular, $1 \in \operatorname{rad} I$ but then $1 \in I$ so $I = k[X_1,\dots,X_n]$.

Proof that Weak Nullstellensatz implies Strong Nullstellensatz. We use a method called the Rabinowitsch trick, introducing an extra variable.

Let I be an ideal in $k[X_1,\dots,X_n]$ and let $V = \mathbb{V}(I) \subseteq \mathbb{A}^n$.

It is easy to see that $\operatorname{rad} I \subseteq \mathbb{I}(V)$. Thus we have to prove that $\mathbb{I}(V) \subseteq \operatorname{rad} I$.

Let $f \in \mathbb{I}(V)$. Define a new polynomial g with an extra variable Y by:

$$g(X_1, \ldots, X_n, Y) = f(X_1, \ldots, X_n)Y - 1.$$

Let J be the ideal in $k[X_1, \ldots, X_n, Y]$ generated by I and g, and consider the affine algebraic set $W = \mathbb{V}(J) \subseteq \mathbb{A}^{n+1}$.

Look at the projection map $\pi\colon \mathbb{A}^{n+1} \to \mathbb{A}^n$ (forgetting the extra Y coordinate). This maps W into V because $I \subseteq J$. But also any point $(x_1, \ldots, x_n)$ in the image $\pi(W)$ must satisfy

$$f(x_1, \ldots, x_n) \neq 0$$

in order for there to exist some y such that

$$f(x_1, \ldots, x_n)y - 1 = 0.$$

(This is generalising the fact that the hyperbola projects down to $\mathbb{A}^1 \setminus \{0\}$.) Thus

$$\pi(W) \subseteq \{\underline{x} \in V : f(\underline{x}) \neq 0\}.$$

Since $f \in \mathbb{I}(V)$, the set on the right is empty. Thus $\pi(W) = \emptyset$. This implies that W itself is empty.

Now the Weak Nullstellensatz tells us that

$$J = k[X_1, \ldots, X_n, Y].$$

In particular, $1 \in J$ and thus

$$1 = a + bg \text{ for some } a \in I.k[X_1, \ldots, X_n, Y],\ b \in k[X_1, \ldots, X_n, Y].$$

Expand out a and b as sums over powers of Y:

$$a = \sum_i a_i Y^i \text{ where } a_i \in I,$$
$$b = \sum_i b_i Y^i \text{ where } b_i \in k[X_1, \ldots, X_n].$$

The equation $1 = a + bg$ can be expanded and rearranged to give

$$1 = a_0 - b_0 + \sum_i (a_i + b_{i-1}f - b_i)Y^i.$$

Looking at the terms of degree 0 in Y gives

$$b_0 = a_0 - 1 \in I - 1$$

then terms of degree 1 in Y gives

$$b_1 = a_1 + b_0 f \in I - f$$

(using $a_1 \in I$ and $b_0 \in I - 1$). Continuing by induction, we get

$$b_j = a_j + b_j f \in I - f^j$$

for all j.

But b is a polynomial, so $b_j = 0$ once j gets large enough. Thus for large j, we get $0 \in I - f^j$, that is, $f^j \in I$. This proves that $f \in \operatorname{rad} I$.

Note: $I - s$ means the coset $\{t - s : t \in I\}$. $\square$

11. Proof of the Weak Nullstellensatz

Today we will prove the Weak Nullstellensatz.

Theorem 11.1 (Weak Nullstellensatz)**.** Let I be an ideal in the polynomial ring $k[X_1, \dots, X_n]$ over an algebraically closed field k. If I is not equal to $k[X_1, \dots, X_n]$ itself, then $\mathbb{V}(I) \neq \emptyset$.

We can restate this in elementary terms as: if $f_1, \dots, f_m \in k[X_1, \dots, X_n]$ are a finite set of polynomials (generating the ideal I), then there exists a common solution $(x_1, \dots, x_n) \in k^n$ to the equations

$$f_1(x_1, \dots, x_n) = 0, \ \dots, \ f_m(x_1, \dots, x_n) = 0.$$

We prove this in two steps:

(1) there exists some larger field K containing k such that these equations have a common solution in K^n.
(2) if the equations have a common solution in K^n, then they also have a common solution in k^n.

Finding a solution in a bigger field. The proof of step (1) is fairly short, and relies on Zorn's lemma.

Lemma 11.2. Let $f_1, \dots, f_m$ be polynomials in $k[X_1, \dots, X_n]$, such that the ideal $I = (f_1, \dots, f_m)$ is not equal to $k[X_1, \dots, X_n]$.

There exists a field K which is a finitely generated extension of k such that the equations

$$f_1(x_1, \dots, x_n) = 0, \ \dots, \ f_m(x_1, \dots, x_n) = 0$$

have a common solution $(x_1, \dots, x_n) \in K^n$.

Proof. Because $I \neq k[X_1, \dots, X_n]$, we can use Zorn's lemma to show that I is contained in some maximal ideal $M \subseteq k[X_1, \dots, X_n]$. (This is a natural way to start: we are trying to show that $\mathbb{V}(I)$ has a point, and last time we saw that points in $\mathbb{V}(I)$ correspond to maximal ideals containing I. We can't just quote the correspondence from the previous lecture because we used the Nullstellensatz in proving that correspondence, but this justifies why obtaining a maximal ideal is a good first step.)

Let $K = k[X_1, \dots, X_n]/M$. Let $x_1, \dots, x_n$ denote the images of $X_1, \dots, X_n$ in K.

K is a field because M is a maximal ideal, and it is finitely generated as an extension of k because it is generated by $x_1, \dots, x_n$.

Since $f_j(X_1, \dots, X_n) \in I \subseteq M$, we get that $f_j(x_1, \dots, x_n) = 0$ in K for each j. Thus $(x_1, \dots, x_n)$ is the required common solution to $f_1, \dots, f_m$ in K^n. $\square$

Shrinking the field required. Before proving step (2), we begin by quoting an algebraic result.

Proposition 11.3. Let k be an algebraically closed field and let K be a finitely generated extension field of k. Then there exist $t_1, \dots, t_d, u \in K$ such that

(i) $K = k(t_1, \dots, t_d, u)$;
(ii) $t_1, \dots, t_d$ are algebraically independent over k (that is, there is no non-zero polynomial in d variables with coefficients in k whose value at $(t_1, \dots, t_d)$ is zero); and
(iii) u is algebraic over $k(t_1, \dots, t_d)$ (that is, there exists a non-zero polynomial in one variable with coefficients in the field $k(t_1, \dots, t_d)$ which is zero at u).

Proof. This follows from the primitive element theorem in field theory. For a full proof, see Proposition A.7 in the Appendix to Shafarevich. □

This proposition has a nice geometric interpretation, but we need to use the Nullstellensatz in order to prove the geometric interpretation so that is postponed until later.

Theorem 11.4. Let k be an algebraically closed field and let K be a finitely generated extension field of k. Let $f_1, \dots, f_m \in k[X_1, \dots, X_n]$.

Suppose there exists a common solution $(x_1, \dots, x_n) \in K^n$ to the equations

$$f_1(x_1, \dots, x_n) = \dots = f_m(x_1, \dots, x_n) = 0.$$

Then there exists a common solution $(y_1, \dots, y_n) \in k^n$ to the equations

$$f_1(y_1, \dots, y_n) = \dots = f_m(y_1, \dots, y_n) = 0.$$

Proof. Write $K = k(t_1, \dots, t_d, u)$ as in Proposition 11.3.

Let $K' = k(t_1, \dots, t_d)$. Because $t_1, \dots, t_d$ are algebraically independent, we can identify K' with $k(T_1, \dots, T_d)$, the field of fractions of the polynomial ring $k[T_1, \dots, T_d]$. This will allow us to substitute a vector $\underline{z} \in k^d$ into an element $\alpha \in K'$ and get out an element $\alpha(\underline{z}) \in k$, as long as the denominator of α does not vanish at $\underline{z}$.

We use two facts about the finite algebraic extension K/K':

(i) There exists a minimal polynomial $p(U) \in K'[U]$ for u; that is, $p(u) = 0$, p has leading coefficient 1, and p divides every other polynomial $q(U) \in K'[U]$ such that $q(u) = 0$.
(ii) Every element of K can be written in the form $\alpha(u)$ for some polynomial $\alpha(U) \in K'[U]$.

In particular, we apply fact (ii) to $x_1, \dots, x_n \in K$ (our common solution to $f_1 = \dots = f_m = 0$): we can write $x_i = a_i(u)$ where $a_i(U) \in K'[U]$.

Because $(x_1, \dots, x_n)$ is a common solution to the polynomials $f_1, \dots, f_m$, we get

$$f_j(a_1(u), \dots, a_n(u)) = 0 \text{ in } K \text{ for each } j.$$

In other words, the single-variable polynomial $f_j(a_1(U), \dots, a_n(U)) \in K'[U]$ has u as a root.

Therefore, fact (i) tells us that this single-variable polynomial is divisible by p. That is, there exist polynomials $q_1, \dots, q_m \in K'[U]$ such that

$$f_j(a_1(U), \dots, a_n(U)) = q_j(U)\, p(U) \text{ in } K'[U]. \tag{*}$$

Each coefficient of the polynomials $p(U), \alpha_i(U), q_j(U)$ is an element of the field K' which we are identifying with $k(T_1, \dots, T_d)$; in other words each coefficient is a fraction where the numerator and denominator are polynomials in $k[T_1, \dots, T_d]$. Let σ denote the product of the denominators of all these fractions.

Because the denominator of a fraction is never zero, $\sigma \neq 0$ in $k[T_1, \dots, T_d]$. Therefore, there exists $(s_1, \dots, s_d) \in k^d$ such that

$$\sigma(s_1, \dots, s_d) \neq 0.$$

Because the denominator in each of coefficient of p divides σ, it is non-zero at $(s_1, \dots, s_d)$. Therefore we can substitute $s_1, \dots, s_d$ into these fractions of polynomials in $k[T_1, \dots, T_d]$ and get values in k. Thus evaluating the coefficients at $(s_1, \dots, s_d)$ transforms $p(U) \in K'[U]$ into a new polynomial $\tilde{p}(U) \in k[U]$. Similarly, we can evaluate the coefficients of $a_i(U)$ and $q_j(U)$ at $(s_1, \dots, s_d)$ to get polynomials $\tilde{a}_i(U), \tilde{q}_j(U) \in k[U]$.

The polynomial $p(U) \in K'[U]$ has leading coefficient 1, so $\tilde{p}(U) \in k[U]$ still has leading coefficient 1. Thus $\tilde{p}(U)$ is not a constant polynomial, so we can choose $z \in k$ such that $\tilde{p}(z) = 0$.

Let

$$y_i = \tilde{a}_i(z) \in k.$$

Then (*) tells us that

$$f_j(y_1, \dots, y_n) = \tilde{q}_j(z)\, \tilde{p}(z) \text{ for each } j.$$

But we chose z such that $\tilde{p}(z) = 0$, and so we conclude that $(y_1, \dots, y_n) \in k^n$ is a common solution to

$$f_1(y_1, \dots, y_n) = \cdots = f_m(y_1, \dots, y_n) = 0. \qquad \square$$

Combining Proposition 11.3 and Theorem 11.4 proves the Weak Nullstellensatz.

Hypersurfaces and birational equivalence. Proposition 11.3 also has a geometrical version, which we will now prove.

Proposition 11.5. Let $V \subseteq \mathbb{A}^n$ be an irreducible affine algebraic set. Then there exists an irreducible hypersurface $H \subseteq \mathbb{A}^m$ for some m such that V is birationally equivalent to H.

This tells us that, even if V is a complicated algebraic set defined by many equations, provided we only care about properties of V which are preserved by birational equivalence, we can replace V by a simpler set defined by just one equation, that is, a hypersurface. Note that it is *not* true that every affine algebraic set is *isomorphic* to a hypersurface.

Proof. Let K denote the function field $k(V)$. Write $K = k(t_1, \ldots, t_d, u)$ as in Proposition 11.3, and let $K' = k(t_1, \ldots, t_d)$.

Because u is algebraic over K', it has a minimal polynomial $f(U) \in K'[U]$. Each coefficient is a fraction whose numerator and denominator are polynomials in $t_1, \ldots, t_d$. We can multiply up by a suitable element of $k[t_1, \ldots, t_d]$ to clear the denominators, and also replace $t_1, \ldots, t_d$ by indeterminates $T_1, \ldots, T_d$ to get a polynomial $g \in k[T_1, \ldots, T_d, U]$ such that

$$g(t_1, \ldots, t_d, u) = 0 \text{ in the field } K.$$

Assuming we multiplied up by a lowest common denominator for the coefficients of f, g is irreducible.

Let H be the hypersurface in $\mathbb{A}^{d+1}$ defined by the polynomial g. Because g is irreducible, it generates a radical ideal and so the (Strong) Nullstellensatz implies that

$$\mathbb{I}(H) = (g).$$

Thus the coordinate ring is given by

$$k[H] = k[T_1, \ldots, T_d, U]/(g).$$

There is a k-algebra homomorphism $\alpha\colon k[T_1, \ldots, T_d, U] \to K$ defined by

$$T_1 \mapsto t_1,\ \ldots,\ T_d \mapsto t_d,\ U \mapsto u.$$

A little algebra (using Gauss's lemma) shows that the kernel of α is generated by g, so α induces an injection $k[H] \hookrightarrow K$. Furthermore, the image of α generates K as a field, so α induces an isomorphism from the fraction field of $k[H]$ to K.

The fraction field of $k[H]$ is the function field $k(H)$. Thus we have shown that $k(H) \cong k(V)$. By Corollary 9.4, this implies that V is birationally equivalent to H. $\square$

12. Projective algebraic sets

Projective space. Projective space consists of affine space together with "points at infinity," one for each direction. The purpose for adding extra points is that it avoids special cases where a point "disappears to infinity." For example, a pair of parallel lines do not intersect in affine space but they do intersect at a point at infinity in projective space.

Definition. Projective n-space, $\mathbb{P}^n$, is the set of lines through the origin in $\mathbb{A}^{n+1}$.

A convenient way to label points in $\mathbb{P}^n$ is via **homogeneous coordinates**. These are just coordinates in $k^{n+1} \setminus \{(0, \dots, 0)\}$: any sequence of coordinates $\underline{x} \in k^{n+1} \setminus \{(0, \dots, 0)\}$ represents the unique line through the origin and $\underline{x}$ in $\mathbb{A}^{n+1}$. Two sequences of homogeneous coordinates $(x_0, \dots, x_n)$ and $(y_0, \dots, y_n)$ represent the same point in $\mathbb{P}^n$ if and only if there exists $\lambda \in k \setminus \{0\}$ such that

$$(x_0, \dots, x_n) = (\lambda y_0, \dots, \lambda y_n).$$

To demonstrate that we are working with homogeneous coordinates instead of ordinary coordinates, we usually write them as

$$[x_0 : x_1 : \cdots : x_n].$$

Observe that we can embed $\mathbb{A}^n$ into $\mathbb{P}^n$ by the map

$$(x_1, \dots, x_n) \mapsto [1 : x_1 : \cdots : x_n].$$

Any other homogeneous coordinates where the first coordinate is non-zero can be re-scaled to have first coordinate 1. So we are left with the points with first coordinate equal to 0: these are the "points at infinity." A point $[0 : x_1 : \cdots : x_n]$ can be seen as a point in $\mathbb{P}^{n-1}$, by just dropping the initial zero. Thus

$$\mathbb{P}^n = \mathbb{A}^n \cup \mathbb{P}^{n-1}.$$

Similarly

$$\mathbb{P}^1 = \mathbb{A}^1 \cup \{\text{a point}\}.$$

We can embed $\mathbb{A}^1$ by the map $x \mapsto [1 : x]$, and then the point at infinity is $[0 : 1]$. Over the complex numbers, $\mathbb{P}^1_{\mathbb{C}}$ is also called the Riemann sphere.

Thinking about projective space as affine space plus points at infinity can be useful if we want to make use of our geometric intuition about affine space or the algebraic tools we have developed for working with affine algebraic sets. On the other hand, thinking about projective space in terms of homogeneous coordinates emphasises that all points of projective space look the same: we can only distinguish points at infinity from points in affine space after choosing a convention for how we embed $\mathbb{A}^n$ into $\mathbb{P}^n$ (for example, we could have used $[x_1 : \cdots : x_n : 1]$ instead); throughout this lecture we will use the convention above.

Projective algebraic sets. A projective algebraic set is a subset of projective space defined by the vanishing of a finite list of polynomials. What does it mean for a polynomial to vanish at a point in projective space? Because a single point in $\mathbb{P}^n$ can be represented by many different homogeneous coordinates, it does not make sense to evaluate a polynomial in $k[X_0, \dots, X_n]$ at a point of $\mathbb{P}^n$. We have to restrict attention to homogeneous polynomials.

Definition. A polynomial $f \in k[X_0, \dots, X_n]$ is **homogeneous** if every term of f has the same degree.

For example, $X_0^3 + X_0^2 X_1 + 3X_2^3 - X_0 X_1 X_2$ is homogeneous of degree 3 while $X_0 X_1 - X_2$ is not homogeneous because it has a term of degree 2 and a term of degree 1.

If $[x_0 : \dots : x_n]$ and $[y_0 : \dots : y_n]$ represent the same point $p \in \mathbb{P}^n$, then

$$(x_0, \dots, x_n) = \lambda(y_0, \dots, y_n) \text{ with } \lambda \in k \setminus \{0\}.$$

Hence if $f \in k[X_0, \dots, X_n]$ is a homogeneous polynomial of degree d, then

$$f(x_0, \dots, x_n) = \lambda^d f(y_0, \dots, y_n)$$

Thus the actual value of f at p does not make sense, but it does make sense to ask whether f is zero at p.

Definition. A **projective algebraic set** is a set of the form

$$\{[x_0 : \dots : x_n] \in \mathbb{P}^n : f_1(x_0, \dots, x_n) = 0,\ \dots,\ f_m(x_0, \dots, x_n) = 0\}$$

for some finite list of homogeneous polynomials $f_1, \dots, f_m \in k[X_0, \dots, X_n]$.

Example. Consider the affine parabola $V = \{(x, y) \in \mathbb{A}^2 : y - x^2 = 0\}$. Under our embedding $\mathbb{A}^2 \to \mathbb{P}^2$, this becomes

$$V = \{[1 : x : y] \in \mathbb{P}^2 : y - x^2 = 0\}.$$

The polynomial $Y - X^2$ is not homogeneous. Consider instead the homogeneous polynomial $WY - X^2$; when $w = 1$, this restricts to $Y - X^2$. Therefore if we let

$$V' = \{[w : x : y] \in \mathbb{P}^2 : wy - x^2 = 0\},$$

then $V' \cap \mathbb{A}^2 = V$. That takes care of the points of V' where $w \neq 0$ (we can scale the homogeneous coordinates of such a point to get $w = 1$).

But V' contains extra points where $w = 0$. Substituting $w = 0$ into $wy - x^2 = 0$ gives also $x = 0$. There is only one point of $\mathbb{P}^2$ with $w = x = 0$: the point $[0 : 0 : 1]$ (any other value for y could be scaled to 1). So

$$V' = V \cup \{[0 : 0 : 1]\}.$$

Geometrically, V' consists of V together with a point at infinity along the y-axis (informally, the two arms of the parabola close up at infinity).

Example. Consider the affine hyperbola $H = \{(x, y) \in \mathbb{A}^2 : xy = 1\}$. This time, in order to find a homogeneous polynomial in $k[W, X, Y]$ which restricts to $XY - 1$ when $W = 1$, we have to replace the constant 1 by W^2. Thus we consider

$$H' = \{[w : x : y] \in \mathbb{P}^2 : xy = w^2\}.$$

Again, when $w \neq 0$, we can scale to get $w = 1$ so just get back H. When $w = 0$, the equation becomes $xy = 0$, so we now get two points at infinity: either $x = 0$, giving the point $[0 : 0 : 1] \in \mathbb{P}^2$, or $y = 0$, giving the point $[0 : 1 : 0] \in \mathbb{P}^2$. Thus

$$H' = H \cup \{[0 : 0 : 1], [0 : 1 : 0]\}.$$

Geometrically, H' consists of H together with points at infinity along the x- and y-axes. These axes are the asymptotes of H.

Compare the two above examples: V' had equation $wy - x^2$, H' had equation $xy - w^2$. These equations differ only by relabelling the coordinates. Thus V' and H' are isomorphic (we have not yet defined isomorphism of projective algebraic sets, but just relabelling the coordinates should certainly be an isomorphism). From the point of view of projective geometry, the only difference between the hyperbola and the parabola is that the parabola has one point at infinity while the hyperbola has two points at infinity.

It turns out that V' and H' are also isomorphic to the projective line $\mathbb{P}^1$ (we will need to define isomorphism of projective algebraic sets before we can prove this).

Homogenisation. The process we went through above to obtain V' from V and H' from H can be generalised.

Definition. For any polynomial $f \in k[X_1, \dots, X_n]$, we define the **homogenisation** of f to be the polynomial in $\bar{f} \in k[X_0, \dots, X_n]$ obtained by the following procedure: let d be the maximum degree of terms of f. Then multiply each term of f by X_0^{d-e}, where e is the degree of this term in f.

For example: if $f(X_1, X_2, X_3) = X_1^3 + 4X_1X_2X_3 - X_1^2 - X_2^2 + 5X_3 + 8$, then the homogenisation is

$$\bar{f}(X_0, X_1, X_2, X_3) = X_1^3 + 4X_1X_2X_3 - X_1^2X_0 - X_2^2X_0 + 5X_3X_0^2 + 8X_0^3.$$

Let $V \subseteq \mathbb{A}^n$ be an affine algebraic set. Let $W \subseteq \mathbb{P}^n$ be the set defined by the homogenisations of all polynomials in $\mathbb{I}(V)$. Then W is a projective algebraic set (this is not entirely obvious, because we have defined it using infinitely many homogeneous polynomials; we will prove this in the next lecture). When we substitute $x_0 = 1$ into the polynomials defining W, we just get back $\mathbb{I}(V)$, so

$$W \cap \{[1 : x_1 : \cdots : x_n]\} = V.$$

This proves that every affine algebraic set V is of the form $W \cap \mathbb{A}^n$ for some projective algebraic set W.

We call W the **projective closure** of V.

Example. When defining the projective closure, it is not enough to just take the homogenisations of some finite list of polynomials which define V; you must take all of $\mathbb{I}(V)$. The standard example of this is the twisted cubic

$$C = \{(t, t^2, t^3)\} = \mathbb{V}(Y - X^2,\, Z - XY) \subseteq \mathbb{A}^3.$$

Parametrically, we can write this as

$$C = \{[1 : t : t^2 : t^3] \in \mathbb{P}^3\}.$$

Homogenising the parametric description, we might expect the projective closure to be

$$C' = \{[s^3 : s^2t : st^2 : t^3] \in \mathbb{P}^3\} = C \cup \{[0 : 0 : 0 : 1]\}.$$

One can check that C' is a projective algebraic set.

But if we homogenise the two defining polynomials $Y - X^2$ and $Z - XY$, we get the projective algebraic set

$$C'' = \{[w : x : y : z] \in \mathbb{P}^3 : wy = x^2,\, wz = xy\}.$$

One can check that

$$C'' = C \cup \{[w : x : y : z] \in \mathbb{P}^3 : w = x = 0\}.$$

Thus $C'' \neq C'$: it contains an extra line at infinity.

If we homogenised all polynomials in $\mathbb{I}(C)$ and not just the two generators, then we would see that the projective closure of C is in fact C'. For example, the three polynomials

$$Y - X^2,\, Z - XY,\, XZ - Y^2$$

are a generating set for $\mathbb{I}(C)$ and their homogenisations define C'.

Zariski topology on $\mathbb{P}^n$. We can define the Zariski topology on $\mathbb{P}^n$ by saying that the closed subsets are the projective algebraic sets. Observe that $\mathbb{A}^n$ is embedded as a Zariski open subset in $\mathbb{P}^n$, because the complement $\mathbb{P}^n \setminus \mathbb{A}^n$ is described by the homogeneous polynomial equation $X_0 = 0$.

The existence of projective closures shows that that the Zariski topology on $\mathbb{A}^n$ is the same as the subspace topology on $\mathbb{A}^n \subseteq \mathbb{P}^n$.

The terminology "projective closure" is justified by noting that the projective closure of $V \subseteq \mathbb{A}^n$ is simply the closure of V in the Zariski topology on $\mathbb{P}^n$.

13. Regular maps between projective algebraic sets

Homogeneous ideals. Someone pointed out to me that last time, I did not actually prove that the projective closure $\overline{V}$ of $V \subseteq \mathbb{A}^n$ is a projective algebraic set – because we constructed $\overline{V}$ as the zero set of infinitely many homogeneous polynomials, but said that a projective algebraic set must be defined using finitely many homogeneous polynomials. We can prove that these are equivalent using the Hilbert Basis Theorem, but it is a little more subtle than in the affine case.

Definition. A **homogeneous ideal** in $k[X_0, \dots, X_n]$ is an ideal which can be generated by homogeneous polynomials.

Note that a homogeneous ideal does not contain only homogeneous polynomials: one can take a homogeneous polynomial f in the ideal and multiply it by $X_0 + 1$ to get a non-homogeneous in I.

If f is any polynomial in $k[X_0, \dots, X_n]$, we can write f (uniquely) as

$$f = \sum_{i=0}^{d} f_i$$

where f_i is homogeneous of degree i. The f_i are called the **homogeneous components** of f.

Lemma 13.1. An ideal $I \subseteq k[X_0, \dots, X_n]$ is a homogeneous ideal if and only if, for each $f \in I$, every homogeneous component of f is in I.

Proof. Just some algebraic manipulation. □

Proposition 13.2. Let $I \subseteq k[X_0, \dots, X_n]$ be a homogeneous ideal. Then there exists a finite set $f_1, \dots, f_m$ of homogeneous polynomials which generate I.

Proof. By the Hilbert Basis Theorem, there exists a finite set $g_1, \dots, g_r$ of polynomials (not necessarily homogeneous) which generate I. In total, the g_i have finitely many homogeneous components. By Lemma 13.1, all homogeneous components are in I; clearly they generate I. □

Thus any set of homogeneous polynomials, even an infinite set, defines a projective algebraic set.

We can use this proposition to prove that every projective algebraic set is a finite union of irreducible components, by the smae proof as for affine algebraic sets.

Projective Nullstellensatz. Which homogeneous ideals can occur as the ideal of functions vanishing on a projective algebraic set? Clearly they have to be radical ideals. Is there a projective version of the Nullstellensatz?

Yes, but it turns out that there is an exceptional case to deal with. Consider the homogeneous ideal $I_1 = (X_0, \dots, X_n) \subseteq k[X_0, \dots, X_n]$. The only solution in k^{n+1} to the equations $x_0 = 0, \dots, x_n = 0$ is $(0, \dots, 0)$. But this is not the homogeneous coordinates of any point in $\mathbb{P}^n$. So the projective algebraic set defined by I_1 is the empty set. Thus the ideals I_1 and $k[X_0, \dots, X_n]$ both define the empty set in $\mathbb{P}^n$, even though they are both radical homogeneous ideals. So we have to modify the statement of the Nullstellensatz slightly from the affine case.

This turns out to be the only special case.

Theorem 13.3 (Projective Weak Nullstellensatz)**.** Let $I \subseteq k[X_0, \dots, X_n]$ be a homogeneous ideal such that $\operatorname{rad} I$ is not equal to either $k[X_0, \dots, X_n]$ or I_1. Then the projective algebraic set defined by I is non-empty.

Proof. Let $V \subseteq \mathbb{P}^n$ denote the projective algebraic set defined by I.

We can also consider denote the affine algebraic set in $\mathbb{A}^{n+1}$ defined by I, which we label C.

Since $\operatorname{rad} I \neq k[X_0, \dots, X_n]$ or I_1, the affine Strong Nullstellensatz tells us that C is not equal to their associated affine algebraic sets, namely $\emptyset$ or $\{(0, \dots, 0)\}$.

Therefore C contains some point $(x_0, \dots, x_n) \in \mathbb{A}^{n+1}$ other than the origin.

But then the point of $\mathbb{P}^n$ with homogeneous coordinates $[x_0 : \cdots : x_n]$ is in V. □

The set C which appears in the above proof is called the **affine cone** of V – it is the union of the lines through the origin in $\mathbb{A}^{n+1}$ which correspond to points of $V \subseteq \mathbb{P}^n$.

One can apply the affine Nullstellensatz to the affine cones of projective algebraic sets to deduce the following bijection between ideals and algebraic sets.

Theorem 13.4. The map sending a homogeneous ideal to the corresponding projective algebraic set is a bijection between the following sets:

$$\left\{\begin{matrix} \text{radical homogeneous ideals} \\ \text{in } k[X_0, \dots, X_n] \\ \text{other than } (X_0, \dots, X_n) \end{matrix}\right\} \longrightarrow \{\text{projective algebraic sets in } \mathbb{P}^n\}.$$

A remark on compactness. Over the complex numbers, every projective algebraic set is compact in the analytic topology. This is because they are closed subsets of $\mathbb{P}^n_{\mathbb{C}}$, which is compact. (In the Zariski topology, the notion of compactness is not very interesting: every algebraic set is compact in the Zariski topology, even affine algebraic sets which definitely do not behave in ways matching our intuition about compactness. This is because most of the usual theory of compact sets is only valid when the sets are Hausdorff.)

There is a converse to this, which tells us that there is a very close relationship between analytic and algebraic geometry in $\mathbb{P}^n_{\mathbb{C}}$:

Theorem 13.5 (Chow's theorem)**.** Let V be an analytic subset of $\mathbb{P}^n_{\mathbb{C}}$ which is closed in the analytic topology. Then V is a projective algebraic set.

I won't define analytic subsets here, but roughly it means a set defined by zeroes of holomorphic functions. This theorem is much harder to prove than to state, and is beyond this course.

One can prove analytically that every holomorphic function on a connected compact complex manifold is constant (for example, this holds on the Riemann sphere, which is equal to $\mathbb{P}^1_{\mathbb{C}}$). Polynomials are holomorphic, so every regular function on a connected projective algebraic set over $\mathbb{C}$ is constant. Once we define regular functions on projective algebraic sets, it will turn out that the same is true over any field.

Regular maps. In the affine case, we defined regular functions first and then used them to define regular maps. However, as remarked above the only regular functions on an irreducible projective algebraic set are constants so they are not useful for defining regular maps. Therefore we will jump directly to defining regular maps.

Let $V \subseteq \mathbb{P}^n$ and $W \subseteq \mathbb{P}^m$ be projective algebraic sets. We expect a regular map $\varphi\colon V \to W$ to be a function which can be expressed as polynomials:

$$\varphi([x_0 : \cdots : x_n]) = [f_0(x_0, \dots, x_n) : \cdots : f_m(x_0, \dots, x_n)].$$

Because we are working with homogeneous coordinates, in order for this to be a well-defined function, all the f_i must be homogeneous polynomials of the same degree. Then, if we re-scale the input coordinates $[x_0 : \cdots : x_n]$ by λ, we get

$$\begin{aligned}&[f_0(\lambda x_0, \dots, \lambda x_n) : \cdots : f_m(\lambda x_0, \dots, \lambda x_n)]\\ &= [\lambda^d f_0(x_0, \dots, x_n) : \cdots : \lambda^d f_m(x_0, \dots, x_n)].\end{aligned}$$

Thus all the output coordinates are re-scaled by the same value λ^d, so they define the same point in $\mathbb{P}^m$.

There is another condition which must be imposed to get a well-defined function $V \to \mathbb{P}^m$: we must never have

$$f_0(x_0, \dots, x_n) = \cdots = f_m(x_0, \dots, x_n) = 0$$

because $[0 : \cdots : 0]$ is not the homogeneous coordinates of a point in $\mathbb{P}^m$.

However it turns out that often, there is not a single sequence of polynomials which will define a regular map at every point of V – whatever polynomials we try, there might be some points where they all vanish. Just like with rational maps, we have to allow different sequences of polynomials to represent our regular map at different points of V. (It is the scaling of homogeneous coordinates that allows different sequences of polynomials to represent the same regular map at places where both are defined.)

Therefore, a regular map is defined to be a map which can be represented by some homogeneous polynomials at every point of V. It is not enough just to say that for each point $x \in V$, there exist some polynomials which give the correct value at x, because then we could get every set-theoretic map by choosing different polynomials at different points. To relate the values of the map at different points, we require that there is some list of polynomials which defines the map on an open neighbourood of x.

Definition. A **regular map** $\varphi\colon V \to W$ is a function $V \to W$ such that for every point $x \in V$, there exist a Zariski open set $U \subseteq V$ containing x and a sequence of polynomials $f_0, \dots, f_m \in k[X_0, \dots, X_n]$ such that:

(i) $f_0, \dots, f_m$ are homogeneous of the same degree;
(ii) for every $y \in U$, $f_0, \dots, f_m$ are not all zero at y;
(iii) for every $y = [y_0 : \dots : y_n] \in U$, $\varphi(y) = [f_0(y_0, \dots, y_n) : \dots : f_m(y_0, \dots, y_n)]$.

Example. Let $V = \mathbb{P}^1$ and let W be the projective closure of the parabola

$$W = \{[w : x : y] \in \mathbb{P}^2 : wy = x^2\}.$$

There is a regular map $\varphi\colon V \to W$ defined by

$$[s : t] \mapsto [s^2 : st : t^2].$$

In this case, this single expression is sufficient to define the regular map everywhere on $V = \mathbb{P}^1$, because s^2, st, t^2 are never simultaneously zero for $[s : t] \in \mathbb{P}^1$ (recall that s and t cannot both be zero at the same point).

We could have constructed φ by starting from the regular map of affine algebraic sets $\mathbb{A}^1 \to$ (the affine parabola) given by $t \mapsto (t, t^2)$. Writing this in homogeneous coordinates as

$$[1 : t] \mapsto [1 : t : t^2]$$

and then homogenising gives φ (this is similar to what we did with the parameterisation of the twisted cubic in the previous lecture).

This homogenisation procedure often allows us to extend a regular map between affine algebraic sets into a regular map between their projective closures, but it does not always work – sometimes we might find that the homogeneous polynomials involved become simultaneously zero at some point at infinity.

Example. Taking V and W as in the previous example, we can define a regular map ψ in the opposite direction $W \to V$ by

$$[w : x : y] \mapsto [w : x] \text{ if } w \neq 0,$$
$$[w : x : y] \mapsto [x : y] \text{ if } w = 0.$$

At this point I ran out of time in the lecture. I will finish this example next time.

14. More on maps between projective algebraic sets

Examples of regular maps.

Example. At the end of the last lecture, I stated this example of a regular map between projective algebraic sets.

Let $V = \mathbb{P}^1$ and let $W = \{[w : x : y] \in \mathbb{P}^2 : wy = x^2\}$ (the projective parabola). Define a map $\psi\colon W \to V$ by

$$[w : x : y] \mapsto [w : x] \text{ if } w \neq 0,$$
$$[w : x : y] \mapsto [x : y] \text{ if } y \neq 0.$$

Each of these expressions is well-defined on a Zariski open subset of W – they are made up of homogeneous polynomials of the same degree, and they never give a point with homogeneous coordinates $[0 : 0]$ within the specified open sets.

To check that this does indeed define a map $W \to V$, we have to check that:

(1) We have defined the map at every point of $\mathbb{P}^1$. This is true because every point of W must have satisfy at least one of $w \neq 0$ or $y \neq 0$ (if $w = y = 0$, then the equation $wy = x^2$ implies that $x = 0$ but $[0 : 0 : 0]$ is not a point of $\mathbb{P}^2$).

(2) On the overlap between the two open sets, both expressions define the same map. This is true because, if w and y are both non-zero and $[w : x : y] \in W$, then x is also non-zero. We can then use the re-scaling property of homogeneous coordinates and the equation $wy = x^2$ to see that

$$[x : y] = [wx : wy] = [wx : x^2] = [w : x].$$

In this example, there is no single sequence of homogeneous polynomials which defines ψ everywhere on V.

Definition. A regular map $\varphi\colon V \to W$ between projective algebraic sets is an **isomorphism** if there exists a regular map $\psi\colon W \to V$ such that $\varphi \circ \psi = \mathrm{id}_W$ and $\psi \circ \varphi = \mathrm{id}_V$.

Observe that the map $\psi\colon W \to V$ which we just defined is inverse to the map $\varphi\colon V \to W$ defined by

$$\varphi([s : t]) = [s^2 : st : t^2].$$

Thus $\mathbb{P}^1$ is isomorphic to the projective parabola.

We already remarked that the projective parabola is isomorphic to the projective hyperbola

$$H = \{[w : x : y] \in \mathbb{P}^2 : xy = w^2\}$$

(by relabelling coordinates), so we deduce that the projective hyperbola is also isomorphic to $\mathbb{P}^1$. In fact, we can show that all irreducible conics in $\mathbb{P}^2$ (a **conic** is a subset of $\mathbb{P}^2$ defined by a homogeneous polynomial of degree 2) are isomorphic to $\mathbb{P}^1$ – by using a projection as in problem 5 on problem sheet 2, and checking that in the projective setting (in this case) the projection gives regular maps and not just rational ones.

Regular maps equal on a dense subset. The following lemma is useful as it tells us that we only need to test equality between regular maps on a dense subset (for example, if V is irreducible, then it is sufficient to look at the open set where a single expression for the regular map is defined).

I made use of this lemma previously for regular maps between affine algebraic sets, and claimed that it was true because regular maps are continuous. But someone pointed out to me that this claim does not hold for continuous maps between non-Hausdorff topological spaces, so we need a proof which uses the fact that our maps are regular.

Lemma 14.1. Let $\varphi, \psi \colon V \to W$ be regular maps. If there exists a Zariski dense subset $A \subseteq V$ such that $\varphi_{|A} = \psi_{|A}$, then $\varphi = \psi$.

Proof. Let $Z = \{x \in V \colon \varphi(x) = \psi(x)\}$. By hypothesis, Z contains a dense subset of V. Hence in order to show that $Z = V$, it suffices to show that Z is closed in V.

Fact. Let S be any topological space (not necessarily Hausdorff). If $\{U_\alpha\}$ is a collection of open subsets of S, whose union is all of S, and Z is any subset of S such that $Z \cap U_\alpha$ is closed in the subspace topology on U_α for every α, then Z is closed in S.

From the definition of regular maps, we know that we can cover V by Zariski open sets U_α such that on each U_α, both φ and ψ are defined by sequences of homogeneous polynomials:

$$\varphi_{|U_\alpha} = [f_{\alpha,0} : \dots : f_{\alpha,m}], \quad \psi_{|U_\alpha} = [g_{\alpha,0} : \dots : g_{\alpha,m}].$$

By the general topological property, it suffices to show that $Z \cap U_\alpha$ is relatively closed in U_α for every α.

Now

$$Z \cap U_\alpha = \{x \in U_\alpha : [f_{\alpha,0}(x) : \dots : f_{\alpha,m}(x)] = [g_{\alpha,0}(x) : \dots : g_{\alpha,m}(x)]\}.$$

This is the same as the set of $x \in U_\alpha$ where the vectors $(f_{\alpha,0}(x), \dots, f_{\alpha,m}(x))$ and $(g_{\alpha,0}(x), \dots, g_{\alpha,m}(x))$ are proportional (for any choice of homogeneous coordinates for x), or in other words where the matrix

$$\begin{pmatrix} f_{\alpha,0}(x) & \cdots & f_{\alpha,m}(x) \\ g_{\alpha,0}(x) & \cdots & g_{\alpha,m}(x) \end{pmatrix}$$

has rank 1. A little linear algebra shows that this condition is equivalent to all the 2×2 minors of this matrix vanishing, that is,

$$f_{\alpha,i}(x) g_{\alpha,j}(x) - f_{\alpha,j}(x) g_{\alpha,i}(x) = 0 \text{ for all } i, j \in \{0, \dots, m\}.$$

This last condition is defined by homogeneous polynomials, and therefore defines a closed subset in the subspace topology on U_α. $\square$

Quasi-projective algebraic sets. So far, we have defined affine algebraic sets and projective algebraic sets, as separate types of object. It is very convenient to have a single notion that unifies both affine and projective algebraic sets.

Definition. A **quasi-projective algebraic set** is the intersection between an open subset and a closed subset of $\mathbb{P}^n$ (in the Zariski topology).

A projective algebraic set is quasi-projective (just take the open subset to be $\mathbb{P}^n$ itself). An affine algebraic set V is also quasi-projective, because it is the intersection between $\mathbb{A}^n$ (which is open in $\mathbb{P}^n$) and the projective closure $\overline{V}$.

We define a regular map between quasi-projective algebraic sets by the same definition as a regular map between projective algebraic sets (it is a map which has a well-defined expression by homogeneous polynomials on a neighbourhood of every point).

If V and W are affine algebraic sets, we now have two ways to define regular maps $V \to W$:

(a) the original definition of regular maps between affine algebraic sets;
(b) view V and W as quasi-projective algebraic sets, and use the new definition of regular maps between quasi-projective algebraic sets.

Fortunately, these two definitions turn out to be equivalent. One has to do a bit of work to check this (the problem is that a regular map of affine algebraic sets must be defined by the same list of polynomials at every point, but a regular map of quasi-projective algebraic sets may be defined by the same polynomials at every point; proving that actually one list of polynomials is enough if the set happens to be affine is similar to the proof of Lemma 8.2).

This gives us a way of defining regular maps from a projective algebraic set to an affine algebraic set or vice versa: just view them both as quasi-projective algebraic sets. For example, we can now define a **regular function** on a projective algebraic set V to be a regular map $V \to \mathbb{A}^1$ (thus it is a function from the algebraic set V taking values in the base field k). As remarked last lecture, we will later prove that the only regular functions on a projective algebraic set are the constants.

We can now make rigorous the claim that "$\mathbb{A}^1 \setminus \{0\}$ looks the same as the affine hyperbola $H = \{(x, y) \in \mathbb{A}^2 : xy = 1\}$." The set

$$\mathbb{A}^1 \setminus \{0\} = \mathbb{P}^1 \setminus \{[1 : 0], [0 : 1]\}$$

is a Zariski open subset of $\mathbb{P}^1$, because its complement is finite. Hence $\mathbb{A}^1 \setminus \{0\}$ is a quasi-projective algebraic set. The map $\varphi\colon \mathbb{A}^1 \setminus \{0\} \to H$ given by $\varphi(t) = (t, 1/t)$ can be written in homogeneous coordinates as

$$\varphi([1 : t]) = [1 : t : 1/t] = [t : t^2 : 1]$$

so homogenising, we get

$$\varphi([s : t]) = [st : t^2 : s^2].$$

So long as $[s : t] \in \mathbb{A}^1 \setminus \{0\}$, this does give a point in

$$H = \{[w : x : y] \in \mathbb{P}^2 : xy = w^2\} \cap \mathbb{A}^2$$

so φ is a regular map $\mathbb{A}^1 \setminus \{0\} \to H$. The projection $(x, y) \mapsto x$ is a regular inverse to φ. Hence $\mathbb{A}^1 \setminus \{0\}$ and H are isomorphic as quasi-projective algebraic sets.

Varieties. As mentioned previously, we use the word "variety" to mean an algebraic set considered up to isomorphism, not caring about how it is embedded into affine or projective space. For example, $\mathbb{A}^1 \setminus \{0\}$ is isomorphic (as a quasi-projective algebraic set) to the affine algebraic set H, so we may say that $\mathbb{A}^1 \setminus \{0\}$ is an affine variety, even though $\mathbb{A}^1 \setminus \{0\}$ is definitely not an affine algebraic set.

There exist quasi-projective algebraic sets which are not isomorphic to anything either projective or affine, for example $\mathbb{A}^2 \setminus \{(0, 0)\}$ (see problem sheet 2).

Rational maps. Let $V \subseteq \mathbb{P}^n$ be an irreducible quasi-projective algebraic set. Just like for affine algebraic sets, a rational map $V \dashrightarrow \mathbb{P}^m$ is something which is almost a regular map, except that it is allowed to have some points where it is not defined. Unlike for affine sets, we don't need to use fractions of polynomials in the definition of a rational map: because our coordinates are homogeneous, we can always multiply up by a common denominator and get an expression involving only polynomials. Rational maps of affine algebraic sets were undefined at points where the denominator was zero; for projective algebraic sets, the points where they are undefined are when all the coordinates of the map become zero.

Once again, we have to pay attention that a rational map can be expressed in terms of polynomials in more than one way, and it might be necessary to use more than one expression to see the full domain of definition of the rational map. So our definition begins by saying which sequences of polynomials determine rational maps, and then specifies when two sequences of polynomials determine the same rational map.

Definition. A **rational map** $\varphi \colon V \dashrightarrow \mathbb{P}^m$ is defined by a sequence of homogeneous polynomials $f_0, \dots, f_m \in k[X_0, \dots, X_n]$ of the same degree such that $f_0, \dots, f_m$ are not all identically zero on V.

We write this as $\varphi = [f_0 : \dots : f_m]$.

Two sequences of polynomials $[f_0 : \dots : f_m]$ and $[g_0 : \dots : g_m]$ represent the same rational map if the homogeneous coordinates

$$[f_0(x) : \dots : f_m(x)], [g_0(x) : \dots : g_m(x)]$$

represent the same point in $\mathbb{P}^m$ wherever both expressions make sense. (Using the fact that V is irreducible, we can check that this is an equivalence relation on sequences of homogeneous polynomials.)

This is exactly the same as the definition of a regular map $V \to \mathbb{P}^m$, except that we are allowing there to be points where no expression for the map is defined.

15. Rational maps between quasi-projective algebraic sets

Let $V \subseteq \mathbb{P}^n$ be an irreducible quasi-projective algebraic set.

At the end of the last lecture, we defined rational maps $V \dashrightarrow \mathbb{P}^m$. Now we define rational maps $V \dashrightarrow W$, where $W \subseteq \mathbb{P}^m$ is any quasi-projective algebraic set. The definition is mostly what you would expect: a rational map is determined by a sequence of homogeneous polynomials $[f_0 : \cdots : f_m]$. There are allowed to be points where these polynomials are all zero, but they cannot be all zero everywhere on V so that the rational map is defined somewhere.

When we say that the rational map goes into W instead of into $\mathbb{P}^m$, we require there to be a Zariski dense set $A \subseteq V$ on which $[f_0(x) : \cdots : f_m(x)]$ lies in W, but we do not require $[f_0(x) : \cdots : f_m(x)] \in W$ at every point of V – this is the difference between regular and rational maps.

Definition. Let $V \subseteq \mathbb{P}^n$ and $W \subseteq \mathbb{P}^m$ be irreducible quasi-projective algebraic sets. A **rational map** $\varphi\colon V \dashrightarrow W$ is determined by a sequence of homogeneous polynomials $f_0, \dots, f_m \in k[X_0, \dots, X_n]$ of the same degree such that:

(1) $f_0, \dots, f_m$ are not all identically zero on V;
(2) there is a Zariski dense set $A \subseteq V$ such that, for all $x \in A$, the homogeneous coordinates $[f_0(x) : \cdots : f_m(x)]$ make sense and define a point in W.

Two sequences of polynomials $[f_0 : \cdots : f_m]$ and $[g_0 : \cdots : g_m]$ represent the same rational map if the homogeneous coordinates

$$[f_0(x) : \cdots : f_m(x)], [g_0(x) : \cdots : g_m(x)]$$

represent the same point in $\mathbb{P}^m$ wherever both expressions make sense.

See also Appendix B for a mroe formal statement of this definition, using equivalence classes.

Definition. A rational map $\varphi\colon V \dashrightarrow W$ is **regular** at a point $x \in V$ if there exists some representation $[f_0 : \cdots : f_m]$ for φ such that $f_0(x), \dots, f_m(x)$ are not all zero and

$$[f_0(x) : \cdots : f_m(x)] \in W.$$

If φ is regular at $x \in V$, then it makes sense to talk about the point $\varphi(x) \in W$, with homogeneous coordinates given by $[f_0(x) : \cdots : f_m(x)]$. This point is independent of the choice of polynomials representing φ, and of the choice of homogeneous coordinates representing x.

The **domain of definition** of φ is the set of points where φ is regular.

Note that, just as in the affine case, when checking whether φ is rational at a point x, it is not enough to check whether the representation $[f_0 : \cdots : f_m]$ which we first used to define the map is regular at x. We have to check whether there exists *any* representation $[g_0 : \cdots : g_m]$ for φ which is defined at x.

Furthermore, the domain of definition can change if we change the target set W. For example, consider the map $\mathbb{P}^1 \to \mathbb{P}^2$ defined by

$$[s : t] \mapsto [s^2 : st : t^2].$$

This is regular at every point. We could interpret the same formula as defining a rational map $\mathbb{P}^1 \dashrightarrow W$ where $W \subseteq \mathbb{P}^2$ is the open set

$$W = \{[w : x : y] : w \neq 0\}.$$

As a rational map $\mathbb{P}^1 \dashrightarrow W$, this is not regular at the point $[0 : 1]$ because this point maps to $[0 : 0 : 1] \notin W$.

Lemma 15.1. Let $\varphi\colon V \dashrightarrow W$ be a rational map. The domain of definition of φ is a non-empty Zariski open subset of V.

Proof. Similar to the affine case (Lemma 8.1). $\square$

It follows immediately from the definition of regular maps between projective algebraic sets that if a rational map is regular at every point, then it is a regular map. (In the affine case (Lemma 8.2), we had to work to prove that if a rational map is regular at every point, then there is a *single* polynomial expression which defines the map everywhere. In the projective case, we don't need to do this because our definition of regular map allows different expressions at different points.)

Example. Let C denote the affine algebraic set

$$C = \{(x, y) \in \mathbb{A}^2 : y = x^3\}.$$

This has projective closure

$$\overline{C} = \{[w : x : y] \in \mathbb{P}^2 : w^2y = x^3\} = C \cup \{[0 : 0 : 1]\}.$$

Consider the regular map of affine algebraic sets $\varphi\colon C \to \mathbb{A}^1$ given by

$$\varphi(x, y) = x.$$

If we try to extend this to a map of projective algebraic sets $\overline{\varphi}\colon \overline{C} \to \mathbb{P}^1$, we would say that for points $[1 : x : y] \in C \subseteq \overline{C}$,

$$\varphi([1 : x : y]) = [1 : x]$$

and this homogenises to

$$\overline{\varphi}([w : x : y]) = [w : x].$$

Thus $\overline{\varphi}$ is a rational map $\overline{C} \dashrightarrow \mathbb{P}^1$.

The above expression for $\overline{\varphi}$ is not defined at the point $[0 : 0 : 1] \in \overline{C}$. We can prove that there is no other expression for $\overline{\varphi}$ which is defined at that point, and so $\overline{\varphi}$ is not regular at $[0 : 0 : 1]$.

Thus a regular map of affine algebraic sets extends to a rational map between their projective closures, but the extended map is not necessarily regular at the points at infinity.

16. Products of projective algebraic sets

Birational maps. This lecture will mainly be about products of projective algebraic sets, but first we say a few words about birational maps of quasi-projective algebraic sets.

Just as in the affine case, if we have irreducible quasi-projective sets V, W, T and rational maps $\varphi\colon V \dashrightarrow W$ and $\psi\colon W \dashrightarrow T$, if the image of φ is dense in W, then the composite $\psi \circ \varphi$ is a rational map $V \dashrightarrow T$.

Definition. A rational map $\varphi\colon V \dashrightarrow W$ is **dominant** if its image is dense in W.

A rational map $\varphi\colon V \dashrightarrow W$ is a **birational equivalence** if it is dominant and there exists a dominant rational map $\psi\colon W \dashrightarrow V$ such that $\psi \circ \varphi = \mathrm{id}_V$ and $\varphi \circ \psi = \mathrm{id}_W$ (where these composite rational maps are defined).

Irreducible algebraic sets V and W are **birational** if there exists a birational equivalence $V \dashrightarrow W$.

Note that $\mathbb{A}^n$ is birational to $\mathbb{P}^n$: consider the regular map

$$\varphi\colon \mathbb{A}^n \to \mathbb{P}^n : (x_1, \dots, x_n) \mapsto [1 : x_1 : \dots : x_n]$$

and the rational map

$$\psi\colon \mathbb{P}^n \dashrightarrow \mathbb{A}^n : [x_0 : \dots : x_n] \mapsto (x_1/x_0, \dots, x_n/x_0).$$

Each of these is dominant and composing them in either direction gives the identity, so these are birational equivalences.

Observe that φ is an isomorphism from $\mathbb{A}^n$ to an open subset of $\mathbb{P}^n$. We can generalise this to show that if V is any irreducible quasi-projective variety and U is a Zariski open subset of V (thus U is also an irreducible quasi-projective variety), then U is birational to V. Indeed, this is a corollary of the following stronger result, which makes precise the intuition that varieties are birational if and only if they are the same "almost everywhere."

Note that we need the concept of quasi-projective varieties to state this lemma, even if V and W are both affine or both projective, because it is necessary to interpret the statement that A and B are isomorphic.

Lemma 16.1. Let V and W be quasi-projective varieties. V is birational to W if and only if there exist non-empty Zariski open subsets $A \subseteq V$ and $B \subseteq W$ such that A is isomorphic to B (as quasi-projective varieties).

Proof. Let $\varphi\colon V \dashrightarrow W$ and $\psi\colon W \dashrightarrow V$ be an inverse pair of rational maps. Let $A_1 = \operatorname{dom} \varphi$ and $B_1 = \operatorname{dom} \psi$. B_1 is a non-empty open subset of W.

Since φ induces a continuous map $A_1 \to W$, $A = \varphi^{-1}(B_1)$ is an open subset of V. Furthermore, since φ is dominant, its image intersects the open set $B_1 \subseteq W$. Therefore A is non-empty.

Similarly $B = \psi^{-1}(A_1)$ is a non-empty open subset of W.

One can now check that $\varphi_{|A}$ and $\psi_{|B}$ form an inverse pair of isomorphisms between A and B. $\square$

If V is a quasi-projective algebraic set, we define a **rational function** on V to be a rational map $\varphi\colon V \dashrightarrow \mathbb{A}^1$. By definition, this is the same as a rational map $\varphi'\colon V \dashrightarrow \mathbb{P}^1$ except that we declare φ to be non-regular at points where $\varphi'(x) = \infty = [0 : 1] \in \mathbb{P}^1$. We can therefore say

$$\varphi(x) = [f(x)(x) : g(x)] = [1 : g(x)/f(x)] = \frac{g(x)}{f(x)} \in \mathbb{A}^1$$

whenever $f(x) \neq 0$, for suitable polynomials f, g. Of course, as always with rational maps, we might need to use different polynomials to evaluate it at different points.

The rational functions on V form a field $k(V)$. Just as in the affine case, V is birational to W if and only if $k(V)$ is isomorphic to $k(W)$. This allows us to calculate

$$k(\mathbb{P}^n) = k(\mathbb{A}^n) = k(X_1, \dots, X_n).$$

Products of projective algebraic sets. It is often useful to work with products $V \times W$ of algebraic sets. For affine algebraic sets, the product is easy to define – $V \times W$ is an affine algebraic subset of $\mathbb{A}^m \times \mathbb{A}^n \cong \mathbb{A}^{m+n}$.

For projective algebraic sets, it is harder to define products because $\mathbb{P}^m \times \mathbb{P}^n \not\cong \mathbb{P}^{m+n}$. To see informally why $\mathbb{P}^1 \times \mathbb{P}^1 \not\cong \mathbb{P}^2$, recall that $\mathbb{P}^1 = \mathbb{A}^1 \cup \{\text{pt}\}$ so

$$\mathbb{P}^1 \times \mathbb{P}^1 = (\mathbb{A}^1 \times \mathbb{A}^1) \cup (\mathbb{A}^1 \times \{\text{pt}\}) \cup (\{\text{pt}\} \times \mathbb{A}^1) \cup (\{\text{pt}\} \times \{\text{pt}\}) = \mathbb{A}^2 \cup \mathbb{A}^1 \cup \mathbb{A}^1 \cup \{\text{pt}\}.$$

Meanwhile

$$\mathbb{P}^2 = \mathbb{A}^2 \cup \mathbb{P}^1 = \mathbb{A}^2 \cup \mathbb{A}^1 \cup \{\text{pt}\}.$$

Thus $\mathbb{P}^1 \times \mathbb{P}^1$ contains an extra copy of $\mathbb{A}^1$ compared to $\mathbb{P}^2$.

We could try defining $\mathbb{P}^m \times \mathbb{P}^n$ by hand. It is fairly clear what algebraic subsets of $\mathbb{P}^m \times \mathbb{P}^n$ should mean (sets defined by polynomials in the two sets of homogeneous coordinates $[x_0 : \cdots : x_m], [y_0 : \cdots : y_n]$; in order for the zero set of such a polynomial to be well-defined, it must be homogeneous in the xs and homogeneous in the ys, but the x and y degrees can be different – such polynomials are called bihomogeneous). Similarly, we could give a definition of regular maps between subvarieties of $\mathbb{P}^m \times \mathbb{P}^n$ involving polynomials in both sets of homogeneous coordinates. But it would be annoying to have just defined quasi-projective varieties, unifying affine and projective varieties, and then immediately have to introduce *ad hoc* definitions for another different kind of variety. So we aim to construct the product in a way which makes it a quasi-projective set, and then we can just reuse the definitions from before.

Furthermore, projective varieties have special properties of their own (in particular, next week we will prove that the image of a projective variety under a regular map is always closed). By showing that the product of projective varieties is itself a projective variety, we will be able to apply these properties to products too.

To construct the product $\mathbb{P}^m \times \mathbb{P}^n$ as a projective algebraic set, we will embed it inside some larger $\mathbb{P}^N$. To do this, let

$$N = (m + 1)(n + 1) - 1 = mn + m + n.$$

Thus the number of homogeneous coordinates needed to specify a point in $\mathbb{P}^N$ is $(m+1)(n+1)$. We will arrange the homogeneous coordinates of points in $\mathbb{P}^N$ in an $(m+1)\times(n+1)$ matrix; thus we label them as $[(z_{ij} : 0 \leq i \leq m, 0 \leq j \leq n)]$ rather than $[z_0 : \cdots : z_{N-1}]$.

Define a map $\sigma_{m,n} \colon \mathbb{P}^m \times \mathbb{P}^n \to \mathbb{P}^N$ by sending $([x_0 : \cdots : x_m], [y_0 : \cdots : y_n])$ to the point in $\mathbb{P}^N$ whose homogeneous coordinates $[(z_{ij})]$ are given by

$$z_{ij} = x_i y_j$$

for each pair of indices i, j. Another way to describe this is to say that the homogeneous coordinates of $\sigma_{m,n}([x_0 : \cdots : x_m], [y_0 : \cdots : y_n])$ are given by the product matrix

$$\begin{pmatrix} x_0 \\ \vdots \\ x_m \end{pmatrix} \begin{pmatrix} y_0 & \cdots & y_n \end{pmatrix}.$$

Observe that this matrix has rank 1.

Let

$$\Sigma_{m,n} = \{[z_{00} : \cdots : z_{mn}] \in \mathbb{P}^N : \text{the matrix } (z_{ij}) \text{ has rank } 1\}.$$

Some linear algebra shows that we can also describe $\Sigma_{m,n}$ as the subset of $\mathbb{P}^N$ where all 2×2 submatrices of the matrix (z_{ij}) have zero determinant. Thus $\Sigma_{m,n}$ is a projective algebraic set, defined by the equations

$$z_{ij} z_{k\ell} = z_{kj} z_{i\ell} \text{ for } 0 \leq i, k \leq m,\ 0 \leq j, \ell \leq n.$$

Lemma 16.2. $\sigma_{m,n}$ is a bijection from $\mathbb{P}^m \times \mathbb{P}^n$ to $\Sigma_{m,n}$.

Proof. We can define an inverse to $\sigma_{m,n}$ as follows:

Let $a \in \Sigma_{m,n}$, and let A be a matrix giving homogeneous coordinates for a. A is not the zero matrix (because it is a set of homogeneous coordinates), so we can pick j such that the j-th column of A contains a non-zero entry. Define $\pi_1(a) \in \mathbb{P}^m$ to be the point with homogeneous coordinates given by the j-th column of A, that is,

$$\pi_1(a) = [A_{1j} : \cdots : A_{mj}].$$

This is independent of the choice of j because the matrix has rank 1 (every non-zero column is a multiple of every other non-zero column).

Similarly we can pick i such that the i-th row of A contains a non-zero entry, and define $\pi_2(a) \in \mathbb{P}^n$ to be the point with homogeneous coordinates given by the i-th row of A. Again this is independent of the choice of i.

Now $(\pi_1, \pi_2) \colon \Sigma_{m,n} \to \mathbb{P}^m \times \mathbb{P}^n$ is an inverse to $\sigma_{m,n}$. □

This construction shows that the projections $\pi_1 \colon \Sigma_{m,n} \to \mathbb{P}^m$ and $\pi_2 \colon \Sigma_{m,n} \to \mathbb{P}^n$ are regular maps (each column of the matrix is non-zero on a Zariski open subset of $\Sigma_{m,n}$).

The map $\sigma_{m,n} \colon \mathbb{P}^m \times \mathbb{P}^n \to \mathbb{P}^N$ is called the **Segre embedding** and its image $\Sigma_{m,n} \subseteq \mathbb{P}^N$ is called the **Segre variety**.

Example. When $m = n = 1$, $N = 3$. The Segre variety $\Sigma_{m,n} \subseteq \mathbb{P}^3$ is defined by the single equation

$$\det \begin{pmatrix} z_{00} & z_{01} \\ z_{10} & z_{11} \end{pmatrix} = z_{00}z_{11} - z_{10}z_{01} = 0.$$

The Segre embedding is given by

$$\sigma_{m,n}([x_1 : x_2], [y_1 : y_2]) = [x_1y_1 : x_1y_2 : x_2y_1 : x_2y_2].$$

We see that $\Sigma_{m,n}$ is an irreducible quadric hypersurface in $\mathbb{P}^3$. Therefore by problem sheet 2, problem 5, it is birational to $\mathbb{A}^2$ (see also sheet 3, problem 4). This is not surprising, because of course $\mathbb{P}^1 \times \mathbb{P}^1$ should have an open subset isomorphic to $\mathbb{A}^1 \times \mathbb{A}^1 \cong \mathbb{A}^2$.

We gave an informal argument earlier for why $\mathbb{P}^1 \times \mathbb{P}^1$ should not be isomorphic to $\mathbb{P}^2$. I don't think we have quite enough tools to prove this rigorously yet.

Because $\Sigma_{m,n}$ is a projective algebraic set, it has a Zariski topology and so we get a Zariski topology on $\mathbb{P}^m \times \mathbb{P}^n$. One can check that this topology is the same as what we expect, namely:

Lemma 16.3. Let $V \subseteq \mathbb{P}^m \times \mathbb{P}^n$ is closed if and only if

$$V = \{([x_0 : \cdots : x_m], [y_0 : \cdots : y_n]) : f_i(x_0, \ldots, x_m, y_0, \ldots, y_n = 0) \text{ for } 1 \leq i \leq s\}$$

where $f_1, \ldots, f_s \in k[X_0, \ldots, X_m, Y_0, \ldots, Y_n]$ are bihomogeneous polynomials.

We say that a polynomial $f \in k[X_0, \ldots, X_m, Y_0, \ldots, Y_n]$ is **bihomogeneous** of degree (d, e) if every term of f has degree d with respect to the X variables and degree e with respect to the Y variables.

17. Images of regular maps

Graphs of regular functions. If $V \subseteq \mathbb{P}^m$ and $W \subseteq \mathbb{P}^n$ are projective algebraic sets, then their product $V \times W$ is a closed subset of $\mathbb{P}^m \times \mathbb{P}^n$ and therefore $V \times W$ is itself a projective variety. Similarly, if $V \subseteq \mathbb{P}^m$ and $W \subseteq \mathbb{P}^n$ are quasi-projective, then $V \times W$ is the intersection between an open and a closed subset of $\mathbb{P}^m \times \mathbb{P}^n$ and therefore is itself a quasi-projective variety.

Example. One useful example of a subvariety of a product is the graph of a regular function.

Let $V \subseteq \mathbb{P}^n$ and $W \subseteq \mathbb{P}^m$ be quasi-projective algebraic sets, and let $\varphi \colon V \to W$ be a regular map. The **graph** of φ is

$$\Gamma = \{(x, y) \in V \times W : y = \varphi(x)\}.$$

To check that this is closed in $V \times W$, observe that Γ is the preimage of the diagonal $\Delta \subseteq \mathbb{P}^m \times \mathbb{P}^m$ under the regular map

$$(\iota \circ \varphi, \iota) \colon V \times W \to \mathbb{P}^m \times \mathbb{P}^m$$

where ι denotes the inclusion map $W \to \mathbb{P}^m$.

Since $(\iota \circ \varphi, \iota)$ is a regular map, it is continuous. Therefore it suffices to check that the diagonal is a Zariski closed subset of $\mathbb{P}^m \times \mathbb{P}^m$. This is true because we can describe the diagonal by bihomogeneous equations as follows:

$$\Delta = \{([x_0 : \cdots : x_m], [y_0 : \cdots : y_m]) : x_i y_j = x_j y_i \text{ for all } i, j\}.$$

Images of projective varieties. The following is a key property of projective algebraic varieties.

Theorem 17.1. Let V be a projective variety. Let $\varphi \colon V \to W$ be a regular map into any quasi-projective variety. Then the image of φ is Zariski closed.

Clearly the theorem is false if V is not projective: for example consider the projection of the hyperbola $\{(x, y) : xy = 1\}$ onto one of the axes.

This theorem shows that projective varieties are similar to compact spaces in topology: if S is a compact topological space and T is a Hausdorff topological space, then the image of any continuous map $S \to T$ is closed.

By applying Theorem 17.1 to $\iota \circ \varphi$, where ι is an embedding $W \to \mathbb{P}^m$, we see that the image of a projective variety under a regular map is again a projective variety.

To prove Theorem 17.1, we will use the graph $\Gamma \subseteq V \times W$ of φ. The image of φ is the same as the projection of Γ onto W. Hence Theorem 17.1 can be deduced from the following theorem.

Theorem 17.2. Let V be a projective variety. For any quasi-projective variety W, the second projection map $\pi_2 \colon V \times W \to W$ maps closed sets to closed sets.

Again, we can see that Theorem 17.2 does not apply when V is not projective by taking $V = W = \mathbb{A}^1$ and taking the hyperbola as a closed subset of $V \times W$.

At first sight, Theorem 17.2 looks stronger than Theorem 17.1 because it applies to all closed subsets of $\subseteq V \times W$, not just the graphs of regular maps. In fact it is easy to deduce Theorem 17.2 from Theorem 17.1, by applying it to $\pi_2 \circ \iota \colon Z \to W$ where ι is the inclusion map $Z \to V \times W$ for any closed subset $Z \subseteq V \times W$.

We say that a variety is **complete** if it satisfies the conclusion of Theorem 17.2. For quasi-projective varieties, complete is equivalent to projective, but if we go beyond the world of quasi-projective varieties it is possible to find algebraic varieties which are complete but not projective. Completeness is the natural analogue in algebraic geometry for compactness in topology; this is justified by the following result from topology.

Lemma 17.3. Let S be a topological space. S is compact if and only if, for every topological space T, the second projection map $S \times T \to T$ maps closed sets to closed sets.

Note that in Lemma 17.3, we use closed sets for the product topology on $S \times T$, while in Theorem 17.2 we use closed sets for the Zariski topology on $V \times W$ (coming from the Segre embedding). These are not the same thing: we have seen (in the case $\mathbb{A}^1 \times \mathbb{A}^1$) that the Zariski topology on a product has more closed sets than the product topology.

We remark that, over the complex numbers, an algebraic variety is complete if and only if it is compact for the analytic topology (this is hard to prove).

Consequences of completeness. Before proving Theorem 17.2, we shall state some important corollaries.

Lemma 17.4. Every regular function on an irreducible projective variety is constant.

Proof. Let V be an irreducible projective variety and $\varphi \colon V \to \mathbb{A}^1$ a regular function. Let $\iota \colon \mathbb{A}^1 \to \mathbb{P}^1$ be the natural inclusion.

Then $\iota \circ \varphi \colon V \to \mathbb{P}^1$ is a regular map, so by Theorem 17.2, its image is a closed subset of $\mathbb{P}^1$. But the image of $\iota \circ \varphi$ is contained in $\mathbb{A}^1$, so it cannot be all of $\mathbb{P}^1$. Therefore the image of φ is finite.

Since V is irreducible, its image is also irreducible and therefore consists of a single point. $\square$

Corollary 17.5. The image of a regular map from an irreducible projective variety to an affine variety is a point.

Proof. Suppose we have a regular map $\varphi \colon V \to W$, where V is projective and irreducible and W is affine. We can suppose that $W \subseteq \mathbb{A}^m$, and let $X_1, \ldots, X_m$ denote the coordinate functions on W. Then $X_1 \circ \varphi, \ldots, X_m \circ \varphi$ are all constant by Lemma 17.4, and so φ is constant. $\square$

Lemma 17.6. Let $V \subseteq \mathbb{P}^n$ be an infinite projective algebraic set and let $H \subseteq \mathbb{P}^n$ be a hypersurface. The intersection $V \cap H$ is non-empty.

Proof. We use the following fact. (This fact is proved for $\mathbb{P}^2$ on problem sheet 3 using the Veronese embedding, and the proof generalises to arbitrary $\mathbb{P}^n$.)

Fact. If $H \subseteq \mathbb{P}^n$ is a hypersurface, then the complement $\mathbb{P}^n \setminus H$ is isomorphic to an affine algebraic set.

Suppose that the lemma were false: then $V \subseteq \mathbb{P}^n \setminus H$. By the fact, $\mathbb{P}^n \setminus H$ is isomorphic to an affine algebraic set. Hence there is an injective regular map $\varphi \colon \mathbb{P}^n \setminus H \to \mathbb{A}^m$ for some m.

Pick an infinite irreducible component $V_1 \subseteq V$. Then V_1 is a projective algebraic set so, by Corollary 17.5, φ maps V_1 to a point.

Since V_1 is infinite, this contradicts the fact that φ is injective. $\square$

Images of quasi-projective varieties. Completeness tells us that images of regular maps of projective algebraic sets are closed. We know that this is false for affine algebraic sets: consider our favourite example of the hyperbola and its projection to $\mathbb{A}^1$.

So what can we say about the images of regular maps of affine, or more generally quasi-projective algebraic sets? We might speculate that they would always be quasi-projective i.e. the intersection of an open and a closed set. But this is not true either: consider the regular map $\varphi \colon \mathbb{A}^2 \to \mathbb{A}^2$ given by

$$\varphi(x, y) = (x, xy).$$

The image of φ is

$$\{(x, y) : x \neq 0\} \cup \{(0, 0)\}.$$

This is the union of an open set with a closed set, not their intersection.

It turns out that this is more or less as bad as things can get.

Definition. Let S be any topological space.

A **locally closed** subset of S is the intersection between an open and a closed set. (For example, quasi-projective algebraic sets are locally closed subsets of $\mathbb{P}^n$.)

A **constructible subset** of S is a finite union of locally closed sets.

Equivalently, a constructible set is any set which can be obtained by starting with a finite list of open and closed sets, and combining them in any way using unions and intersections.

The image of the map $(x, y) \mapsto (x, xy)$ considered above is the typical example to keep in mind for a constructible set which is not locally closed.

Chevalley's theorem tells us that the image of a regular map between quasi-projective varieties is constructible. Indeed it tells us slightly more – the image of a constructible set is constructible.

Theorem 17.7 (Chevalley's theorem)**.** Let $\varphi\colon V \to W$ be a regular map of quasi-projective algebraic sets. The image of any constructible set in V is a constructible set in W.

We will prove completeness of projective varieties and Chevalley's theorem in the next lecture. Their proofs are linked but neither theorem is an easy consequence of the other.

Aside: Chevalley's theorem and mathematical logic. This is not a formal part of the course, so I will not define things carefully. For anyone with an interest in mathematical logic, we remark on a logical interpretation of Chevalley's theorem.

In logic, we consider "formulas" made out of some algebraic operations (in the context of algebraic geometry, these will be polynomial equations over an algebraically closed field) and combine these using logical operations – AND, OR, NOT, $\exists$, $\forall$. Thus a logical formula might look like

$$(\text{not } xy = z) \text{ and } \exists(u, v) \text{ s.t. } (x = u^2 \text{ and } y = uv \text{ and } z = v^2).$$

If we allow just AND, OR, NOT then formulas like this define unions and intersections of Zariski open and closed sets in $\mathbb{A}^n$ – that is, constructible sets. If we also allow quantifiers, then we can also get images of regular maps – for example, the part of the formula above starting with $\exists(u, v)$ defines the image of the regular map

$$(u, v) \mapsto (u^2, uv, v^2).$$

But Chevalley's theorem tells us that images of regular maps are actually also constructible sets, so we deduce that:

Fact. Every formula (made out of polynomials over an algebraically closed field) is equivalent to a formula without quantifiers.

This is called "elimination of quantifiers for algebraically closed fields."

18. Proof of completeness

Proof of completeness. We will now prove the completeness of projective varieties. We recall the version of the theorem which we shall prove.

Theorem 18.1. Let V be a projective variety. For any quasi-projective variety W, the second projection map $\pi_2\colon V \times W \to W$ maps closed sets to closed sets.

Let Z be a closed subset of $V \times W$.

By problem 4 of sheet 4, we may cover W by open sets U_α such that each U_α is an affine variety. We previously observed that, if $\pi_2(Z) \cap U_\alpha$ is closed in U_α for every α, then $\pi_2(Z)$ is closed in W (see proof of Lemma 14.1). Since

$$\pi_2(Z) \cap U_\alpha = \pi_2(Z \cap (V \times U_\alpha)),$$

we conclude that it suffices to prove Theorem 18.1 for the case where W is affine.

Then we can replace $V \subseteq \mathbb{P}^n$ by $\mathbb{P}^n$ and $W \subseteq \mathbb{A}^m$ by $\mathbb{A}^m$ (because V is closed in $\mathbb{P}^n$ and W is closed in $\mathbb{P}^m$, Z will still be closed in $\mathbb{P}^n \times \mathbb{A}^m$).

Thus it will suffice to prove the following.

Theorem 18.2. The second projection map $\pi_2\colon \mathbb{P}^n \times \mathbb{A}^m \to \mathbb{A}^m$ maps closed sets to closed sets.

Proof. We can concretely describe a Zariski closed subsets $Z \subseteq \mathbb{P}^n \times \mathbb{A}^m$ as the zero set of some polynomials $f_1, \dots, f_r \in k[X_0, \dots, X_n, Y_1, \dots, Y_m]$ which are homogeneous with respect to $X_0, \dots, X_n$. (The coordinates $Y_1, \dots, Y_m$ are affine coordinates, so there is no homogeneity condition with respect to them.)

For each point $(y_1, \dots, y_m) \in \mathbb{A}^m$, we can substitute the values $(y_1, \dots, y_n)$ into these polynomials and get a projective algebraic set

$$Z_{(y_1,\dots,y_m)} = \{[x_0 : \cdots : x_n] \in \mathbb{P}^n : f_i(x_0, \dots, x_n, y_1, \dots, y_m) = 0 \text{ for all } i\}.$$

Observe that $(y_1, \dots, y_m) \in \pi_2(Z)$ if and only if $Z_{(y_1,\dots,y_m)}$ is non-empty.

Let $I_{(y_1,\dots,y_m)}$ denote the ideal in $k[X_0, \dots, X_n]$ generated by the polynomials

$$f_0(X_0, \dots, X_n, y_1, \dots, y_m), \dots, f_r(X_0, \dots, X_n, y_1, \dots, y_m).$$

By the Projective Nullstellensatz, $Z_{(y_1,\dots,y_m)}$ is non-empty if and only if $\operatorname{rad} I_{(y_1,\dots,y_m)}$ is not equal to either the full ring $k[X_0, \dots, X_n]$ or to the ideal $(X_0, \dots, X_n)$. It is easy to see that this is equivalent to: $I_{(y_1,\dots,y_n)}$ does not contain S_d for any positive integer d, where S_d denotes the set of all homogeneous polynomials of degree d in $k[X_0, \dots, X_n]$.

Write

$$W_d = \{(y_1, \dots, y_m) : I_{(y_1,\dots,y_m)} \not\supseteq S_d\}.$$

We have shown that $\pi_2(Z) = \bigcap_{d \in \mathbb{N}} W_d$. Since the W_d are a descending chain of sets, it suffices to prove that W_d is a closed subset of $\mathbb{A}^m$ for all d greater than some positive integer e.

Let the polynomials $f_0, \dots, f_r$ have degrees $d_0, \dots, d_r$ with respect to the X variables. We shall show that W_d is closed for $d \geq \max(d_0, \dots, d_r)$.

Now we just need some linear algebra to finish the proof.

If $g \in S_d$, then $g \in I_{(y_1,\dots,y_m)}$ if and only if we can write

$$g(X_0,\dots,X_n) = \sum_{i=1}^{r} f_i(X_0,\dots,X_n,y_1,\dots,y_m)\, h_i(X_0,\dots,X_n)$$

for some homogeneous polynomials $h_1,\dots,h_r$, where $\deg h_i = d - d_i$. Hence $S_d \cap I_{(y_1,\dots,y_m)}$ is the image of the linear map

$$\alpha_{d,y_1,\dots,y_m} \colon \bigoplus_{i=1}^{r} S_{d-d_i} \to S_d$$

given by

$$\alpha_{d,y_1,\dots,y_m}(h_1,\dots,h_r) = \sum_{i=1}^{r} f_i(X_0,\dots,X_n,y_1,\dots,y_m)\, h_i(X_0,\dots,X_n).$$

Therefore

$$\begin{aligned} W_d &= \{(y_1,\dots,y_m) : \alpha_{d,y_1,\dots,y_m} \text{ is not surjective}\} \\ &= \{(y_1,\dots,y_m) : \operatorname{rk}\alpha_{d,y_1,\dots,y_m} < \dim S_d\}. \end{aligned}$$

If we choose bases for S_d and $\bigoplus_i S_{d-d_i}$ and write $\alpha_{d,y_1,\dots,y_m}$ as a matrix with respect to these bases, then W_d consists of the points $(y_1,\dots,y_m)$ where all $\dim S_d \times \dim S_d$ submatrices of this matrix have zero determinant. The determinants of these submatrices are polynomials in $y_1,\dots,y_m$, proving that W_d is Zariski closed in $\mathbb{A}^m$. $\square$

The resultant. Theorem 18.2 has the following application to roots of polynomials: We want to describe the set of pairs of polynomials $f, g \in k[S,T]$, homogeneous of degrees d and e respectively, for which the set of common zeroes

$$\{[s:t] \in \mathbb{P}^1 : f(s,t) = 0, g(s,t) = 0\}$$

is non-empty.

We can identify the space of homogeneous polynomials of degree d in two variables with $\mathbb{A}^{d+1}$, by associating $\underline{a} = (a_0,\dots,a_d) \in \mathbb{A}^{d+1}$ with the polynomial

$$f_{\underline{a}}(S,T) = \sum_{i=0}^{d} a_i S^i T^{d-i}.$$

We can define a Zariski closed subset of $\mathbb{P}^1 \times \mathbb{A}^{(d+1)+(e+1)}$ by:

$$Z_{d,e} = \{([s:t],\underline{a},\underline{b}) \in \mathbb{P}^1 \times \mathbb{A}^{(d+1)+(e+1)} : \sum_{i=0}^{d} a_i s^i t^{d-i} = 0 \text{ and } \sum_{i=0}^{e} b_i s^i t^{e-i} = 0\}.$$

For any point $(\underline{a},\underline{b}) \in \mathbb{A}^{(d+1)+(e+1)}$ the fibre $\pi_2^{-1}(\underline{a},\underline{b}) \cap Z_{d,e}$ is simply the set of common zeroes of $f_{\underline{a}}$ and $f_{\underline{b}}$ in $\mathbb{P}^1$. Hence

$$\pi_2(Z_{d,e}) = \{(\underline{a},\underline{b}) \in \mathbb{A}^{(d+1)+(e+1)} : f_{\underline{a}} \text{ and } f_{\underline{b}} \text{ have a common zero in } \mathbb{P}^1\}.$$

By Theorem 18.2, $\pi_2(Z_{d,e})$ is a Zariski closed subset of $\mathbb{A}^{(d+1)+(e+1)}$.

In other words, there is some list of polynomials $p_1, \dots, p_r$ such that the condition "homogeneous polynomials f, g in two variables of given degrees have a common zero in $\mathbb{P}^1$" is equivalent to $p_1, \dots, p_r$ all vanishing at the coefficients of f and g. It turns out that this condition is equivalent not just to the vanishing of a list of polynomials in the coefficients, but to a single polynomial called the **resultant** $\mathrm{Res}_{d,e}$:

Theorem 18.3. Fix positive integers d, e. There exists a polynomial $\mathrm{Res}_{d,e} \in k[A_0, \dots, A_d, B_0, \dots, B_e]$ such that the polynomials

$$\sum_{i=0}^{d} a_i S^i T^{d-i} \text{ and } \sum_{i=0}^{e} b_i S^t T^{e-i}$$

have a common root in $\mathbb{P}^1$ if and only if

$$\mathrm{Res}_{d,e}(a_0, \dots, a_d, b_0, \dots, b_e) = 0.$$

This can be proved by going through the linear algebra from the end of the proof of Theorem 18.2 – indeed it is possible to work out the polynomial $\mathrm{Res}_{d,e}$ explicitly in this way.

We shall just quote this as a result of algebra. The algebra actually gives us something more: this works not just for polynomials over an algebraically closed field, but for polynomials over any integral domain, provided we replace "have a common zero" by "have a common factor of positive degree."

Theorem 18.4. Fix positive integers d, e. There exists a universal polynomial $\mathrm{Res}_{d,e} \in \mathbb{Z}[A_0, \dots, A_d, B_0, \dots, B_e]$ such that, for any integral domain R and any values $a_0, \dots, a_d, b_0, \dots, b_e \in R$, the homogeneous polynomials

$$f = \sum_{i=0}^{d} a_i S^i T^{d-i}, \quad g = \sum_{i=0}^{e} b_i S^i T^{e-i} \in R[S, T]$$

have a common factor of positive degree in $R[S, T]$ if and only if

$$\mathrm{Res}_{d,e}(a_0, \dots, a_d, b_0, \dots, b_e) = 0.$$

Aside: The resultant in elementary algebra. From the perspective of elementary algebra, when stating Theorem 18.3, it is simpler to look at inhomogeneous polynomials in one variable and roots in $\mathbb{A}^1$ instead of homogeneous polynomials in two variables and roots in $\mathbb{P}^1$. Of course, we can convert back and forth by homogenising and dehomogenising the polynomials (replacing $f(S, T)$ by $f(1, T)$ and vice versa) but we have to be a little bit careful. We have to worry about the possibility that $f(S, T)$ and $g(S, T)$ might have a common root at $\infty = [0 : 1] \in \mathbb{P}^1$ but not anywhere in $\mathbb{A}^1 = \mathbb{P}^1 \setminus \{\infty\}$. It turns out that $f(S, T)$ has a root at ∞ if and only if the dehomogenised polynomial $f(1, T)$ has degree strictly less than $\deg f(S, T)$.

Thus, Theorem 18.3 implies that $\mathrm{Res}_{d,e}$ vanishes on the coefficients of two single variable polynomials

$$f(T) = \sum_{i=0}^{d} a_i T^i, \quad g(T) = \sum_{i=0}^{e} b_i T^i \tag{*}$$

if and only if f and g have a common root in k, *as long as* f and g have degrees exactly d and e respectively. If $\deg f < d$, but we still write out f as in (*) with $a_d = 0$, then looking at $\mathrm{Res}_{d,e}$ might give the wrong answer for whether f and g have a common root (and similarly if $\deg g < e$).

19. Proof of Chevalley's theorem

We shall now prove Chevalley's theorem. We recall the statement.

Theorem 19.1. Let $\varphi\colon V \to W$ be a regular map of quasi-projective algebraic sets. The image φ maps constructible sets to constructible sets.

We will reduce the proof to projections by using graphs, in a similar way to what we did for completeness. We can also reduce to the case where V and W are affine spaces, and perform a few more simplifications.

The real work of the proof lies in the first two lemmas; after proving these lemmas, we will show the sequence of simpler steps to get from there to Theorem 19.1.

Lemma 19.2. Let $\pi\colon \mathbb{P}^1 \times \mathbb{A}^m \to \mathbb{A}^m$ denote the second projection map.

Let V be an irreducible closed subset of $\mathbb{P}^1 \times \mathbb{A}^m$ and let T be a proper closed subset of V. Then either:

(i) $\pi(T)$ is strictly contained in $\pi(V)$; or
(ii) $V = \mathbb{P}^1 \times \pi(V)$.

Proof. Let $W = \pi(V)$. By the completeness of $\mathbb{P}^1$, W is closed in $\mathbb{A}^m$ (Theorem 18.1). Since V is irreducible, W is also irreducible. Hence the ring of regular functions $k[W]$ is an integral domain.

We write $k[W][X, Y]$ for the ring of polynomials in two variables X, Y with coefficients in $k[W]$. Let $I \subseteq k[W][X, Y]$ denote the homogeneous ideal of polynomials which vanish on $V \subseteq \mathbb{P}^1 \times W$.

Assume that $V \neq \mathbb{P}^1 \times W$ (otherwise conclusion (ii) holds so we have nothing to prove). Thus $I \neq \{0\}$.

Let f be a homogeneous polynomial in I which has minimum degree with respect to X, Y. Since V is irreducible, f has no factorisation into factors of positive degree (if $f = f_1 f_2$, then $V \cup \{f_1 = 0\}$ and $V \cup \{f_2 = 0\}$ would be two closed subsets which cover V).

Since T is closed and properly contained in V, there exists $g \in k[W][X, Y]$ which vanishes on T but not on V. Since g does not vanish on V, f does not divide g.

Since f has no factors of positive degree, we conclude that f and g have no common factors of positive degree. Hence we can use the resultant over the integral domain $k[W]$ (Theorem 18.4) to say that

$$\mathrm{Res}_{d,e}(f, g) \neq 0 \text{ in } k[W]$$

where $d = \deg f$, $e = \deg g$.

Now $\mathrm{Res}_{d,e}(f, g)$ is an element of $k[W]$, i.e. it is a regular function on W. Since it is not identically zero, we can choose $w \in W$ such that

$$\mathrm{Res}_{d,e}(f, g)(w) \neq 0.$$

This is the same as saying that $\mathrm{Res}_{d,e}(f_w, g_w) \neq 0$, where $f_w, g_w \in k[X, Y]$ are the polynomials obtained from f and g by evaluating their coefficients at w.

By the defining property of the resultant over the algebraically closed field k, we conclude that f_w and g_w have no common root in $\mathbb{P}^1$. But $\pi^{-1}(w) \cap T$ is contained in the set of common roots of f_w and g_w. Thus $\pi^{-1}(w) \cap T = \emptyset$, that is, $w \notin \pi(T)$. $\square$

Lemma 19.3. Let $\pi\colon \mathbb{A}^{1+m} \to \mathbb{A}^m$ denote projection onto the last m coordinates.

Let $Z \subseteq \mathbb{A}^{1+m}$ be an irreducible locally closed subset. Let W be the Zariski closure of $\pi(Z)$ in $\mathbb{A}^m$.

Then $\pi(Z)$ contains a non-empty open subset of W.

Proof. We embed $\mathbb{A}^{1+m} = \mathbb{A}^1 \times \mathbb{A}^m$ into $\mathbb{P}^1 \times \mathbb{A}^m$, in order to be able to use Lemma 19.2. (The previous lemma needed $\mathbb{P}^1 \times \mathbb{A}^m$ in order to use completeness, here we need $\mathbb{A}^{1+m} \to \mathbb{A}^m$ so that we can set up an induction $\mathbb{A}^{2+m} \to \mathbb{A}^{1+m} \to \mathbb{A}^m$ etc.)

Let V be the closure of Z in $\mathbb{P}^1 \times \mathbb{A}^m$. By completeness (Theorem 18.1), $\pi(V)$ is closed in $\mathbb{A}^m$, so $\pi(V) = W$.

Let $T = V \setminus Z$. Since Z is locally closed, T is a closed subset of $\mathbb{P}^1 \times \mathbb{A}^m$.

We can now apply Lemma 19.2. We get two cases:

Case (i). $\pi(T)$ is strictly contained in $\pi(V)$.

Completeness tells us that $\pi(T)$ is closed in $\mathbb{A}^m$. Hence $W \setminus \pi(T)$ is an open subset of W. Since $Z = V \setminus T$, it is clear that $W \setminus \pi(T)$ is contained in $\pi(Z)$. And $W \setminus \pi(T)$ is non-empty because $\pi(T) \neq W$.

Case (ii). $V = \mathbb{P}^1 \times W$.

In this case, an element $w \in W$ is in the image of Z if and only if $\mathbb{P}^1 \times \{w\}$ is not contained in T. That is

$$\pi(Z) = \{w \in W : \mathbb{P}^1 \times \{w\} \not\subseteq T\}.$$

It is easy to see that the complement $\{w \in W : \mathbb{P}^1 \times \{w\} \subseteq T\}$ is a closed subset of W, so $\pi(Z)$ is open in W. And $\pi(Z)$ is certainly non-empty. $\square$

Corollary 19.4. Let $\pi\colon \mathbb{A}^{n+m} \to \mathbb{A}^m$ denote projection onto the last m coordinates.

Let $Z \subseteq \mathbb{A}^{n+m}$ be a constructible subset. Let W be the Zariski closure of $\pi(Z)$ in $\mathbb{A}^m$.

Then $\pi(Z)$ contains a dense open subset of W.

Proof. It suffices to assume that Z is an irreducible locally closed set; if it was not, we could simply break it up first into finitely many locally closed sets (by the definition of constructible sets), and then break up each of these into finitely many irreducible components.

The proof is by induction on n. (The base case is Lemma 19.3.) We factor $\pi\colon \mathbb{A}^{n+m} \to \mathbb{A}^m$ as

$$\mathbb{A}^{n+m} \xrightarrow{p} \mathbb{A}^{1+m} \xrightarrow{q} \mathbb{A}^m \,.$$

Let W' be the Zariski closure of $p(Z)$. By induction, $p(Z)$ contains a dense open subset $U \subseteq W'$. Since U is dense in W', $q(U)$ is dense in $q(W') = \pi(Z)$ which in turn is dense in W. Therefore W is equal to the closure of $q(U)$ in $\mathbb{A}^m$.

U is an open subset of the closed subset W' in $\mathbb{A}^{1+m}$, so U is locally closed in $\mathbb{A}^{1+m}$. Hence we can apply Lemma 19.3 to conclude that $q(U)$ contains a non-empty open subset of W. Because Z, and hence W, is irreducible, this non-empty open subset is dense in W. □

Lemma 19.5. Let $\pi\colon \mathbb{A}^{n+m} \to \mathbb{A}^m$ denote projection onto the last m coordinates. If $Z \subseteq \mathbb{A}^{n+m}$ is a constructible subset, then $\pi(Z)$ is constructible.

Proof. Let W_1 be the Zariski closure of $\pi(Z)$. By Corollary 19.4, $\pi(Z)$ contains a dense open subset $U_1 \subseteq W_1$. Then $Z_1 = \pi^{-1}(W_1 \setminus U_1)$ is a proper closed subset of Z. In particular Z_1 is itself constructible.

Now repeat this argument: let W_2 be the Zariski closure of $\pi(Z_1)$. By Corollary 19.4, $\pi(Z_1)$ contains a dense open subset $U_2 \subseteq W_2$. Then $Z_2 = \pi^{-1}(W_2 \setminus U_2)$ is a proper closed subset of Z_1.

We repeat this, getting W_3, U_3, Z_3, etc.

Now $\pi(Z_1)$ is contained in the closed set $W_1 \setminus U_1$, so $W_2 \subseteq W_1 \setminus U_1$. Since $U_1 \neq \emptyset$, $W_2 \neq W_1$. Similarly, $W_3 \neq W_2$ etc.

Hence we get a strictly descending chain of closed subsets of $\mathbb{A}^m$:

$$W_1 \supsetneq W_2 \supsetneq W_3 \supsetneq \cdots$$

By the noetherian property of the Zariski topology (Proposition 4.1), this chain must terminate. So we eventually get to $W_r = \emptyset$.

But then $\pi(Z_{r-1}) \subseteq W_r$, so Z_{r-1} is empty (we can't go back any further than that: it is entirely possible that $U_{r-1} = W_{r-1}$).

Then

$$\pi(Z) = U_1 \cup \pi(Z_1) = U_1 \cup U_2 \cup \pi(Z_2) = \cdots = U_1 \cup \cdots \cup U_{r-1}.$$

Each U_i is an open subset of a closed subset of $\mathbb{A}^n$, so we conclude that $\pi(Z)$ is constructible. Since $q(U) \subseteq \pi(Z)$, this completes the proof.

(This argument, building a descending chain of closed subsets and concluding that it terminates, is called *noetherian induction.*)

I got this proof wrong in the earlier version of the notes (and probably also in the lecture) – I tried to use a descending chain $Z \supsetneq Z_1 \supsetneq Z_2 \supsetneq \cdots$ but this is only a chain of constructible sets, not closed sets, so we need to use $W_1 \supsetneq W_2 \supsetneq \cdots$ instead. □

Now we can finish the proof of Theorem 19.1. By considering the graph of a regular map $\varphi\colon V \to W$, just as for completeness, we can reduce it to:

Theorem 19.6. Let V and W be quasi-projective algebraic sets. The second projection $\pi_2\colon V \times W \to W$ maps constructible sets to construtible sets.

Let Z be a constructible subset of $V \times W$. Suppose that $V \subseteq \mathbb{P}^n$ and $W \subseteq \mathbb{P}^m$. Then Z is also constructible as a subset of $\mathbb{P}^n \times \mathbb{P}^m$, so we can replace V and W by $\mathbb{P}^n$ and $\mathbb{P}^m$.

Let
$$U_i = \{[x_0 : \cdots : x_m] \in \mathbb{P}^m : x_i \neq 0\}.$$
The U_i form an open cover for $\mathbb{P}^m$, and each of them is isomorphic to $\mathbb{A}^m$. If each set $Z \cap (V \times U_i)$ has a constructible image, then $\pi(Z)$ is a finite union of constructible sets, so is itself constructible. So it suffices to prove Theorem 19.6 for $W = \mathbb{A}^m$.

A similar argument allows us to replace $V = \mathbb{P}^n$ by $V = \mathbb{A}^n$. Thus the proof of Theorem 19.1 is reduced to Lemma 19.5.

The real hard work in this argument was in the proof of Lemmas 19.2 and 19.3, while the rest consisted of a series of relatively simple steps allowing us to gradually simplify the varieties and regular maps which we had to work with. A sequence of reductions like this, before you do the hard work on a simpler case of the original problem, is a very common technique in algebraic geometry – common enough to be given its own name, *dévissage*.

20. Dimension

Dimension and transcendence degree. We want to define the dimension of algebraic varieties. There are several different definitions, all equivalent but each being useful in different situations. Note of these definitions is particularly obvious, so we begin by listing some properties that the "dimension" of an irreducible quasi-projective variety V ought to have. (We only consider irreducible varieties here, because a reducible variety might have components of different dimensions so it is less clear what properties it should have.)

(1) $\dim V$ is a nonnegative integer.
(2) $\dim V = 0$ if and only if V is a point (remember that we are assuming that V is irreducible).
(3) $\dim \mathbb{A}^n = \dim \mathbb{P}^n = n$.
(4) A hypersurface in $\mathbb{A}^n$ or $\mathbb{P}^n$ has dimension $n - 1$.
(5) If U is an open subset of V, then $\dim U = \dim V$ (note that this holds for manifolds in differential geometry).
(6) If V and W are birational, then $\dim V = \dim W$ (this follows from (5)).

By Proposition 11.5, every irreducible quasi-projective variety is birational to a hypersurface in some affine space. (Proposition 11.5 was only for affine varieties, but given an arbitrary irreducible quasi-projective V we can use sheet 4 of problem 4 to get an affine open set $U \subseteq V$, and then apply Proposition 11.5 to U.) Hence properties (4) and (6) are enough to tell us the dimension of every irreducible quasi-projective variety:

Quasi-definition. The dimension of an irreducible quasi-projective variety V is d if V is birational to a hypersurface in $\mathbb{A}^{d+1}$.

I have described this as a quasi-definition instead of a definition, because there is one problem with it: V might be birational to lots of different hypersurfaces. How do we know that they will all live in affine space of the same dimension?

We solve this by using the notion of transcendence degree from algebra.

Theorem 20.1. Let k and K be fields, with $k \subseteq K$. All maximal k-algebraically independent sets in K have the same cardinality.

The cardinality of maximal k-algebraically independent sets in K is called the **transcendence degree** of the extension K/k.

In the proof of Proposition 11.5, we took a maximal k-algebraically independent set $z_1, \dots, z_d$ in $k(V)$, then proved that V is birational to a hypersurface in $\mathbb{A}^{d+1}$. Theorem 20.1 shows that the value of d here is the same for all maximal k-algebraically independent sets in $k(V)$. This establishes that the following definition is equivalent to the above quasi-definition:

Definition. The **dimension** of an irreducible quasi-projective variety V is the transcendence degree (over k) of the field of rational functions $k(V)$.

It is easy to see that this definition satisfies all the above desired properties. (In $k(\mathbb{A}^n) = k(X_1, \ldots, X_n)$, $X_1, \ldots, X_n$ form a maximal algebraically independent set, so $\dim \mathbb{A}^n = n$.) In particular, if V and W are birational, then they have the same dimension because $k(V) \cong k(W)$.

Generically finite maps. Actually, we can prove something stronger this: we don't need a birational map $V \dashrightarrow W$, just a generically finite dominant rational map.

Definition. Let V and W be irreducible quasi-projective varieties. A dominant rational map $\varphi\colon V \to W$ is **generically finite** if there is a non-empty open set $U \subseteq W$ such that $\varphi^{-1}(x)$ is finite for every $x \in U$. (Note: there are two reasonable ways to define "generically finite" for non-dominant rational maps, and both are used by different authors. We shall allow ourselves to use the words "generically finite" only when the map is dominant.)

Lemma 20.2. If $\varphi\colon V \dashrightarrow W$ is a generically finite dominant rational map, then $\dim V = \dim W$.

Proof. We can replace V by the open subset $\operatorname{dom} \varphi$, so that φ becomes a regular map. We can then replace V and W by affine open subsets, and then replace V by the graph of φ in $V \times W$. Hence it suffices to assume that $\varphi = \pi_{|V}$, where π is the projection $\mathbb{A}^{n+m} \to \mathbb{A}^m$ and V is a closed subset of $\mathbb{A}^{n+m}$, W is the Zariski closure of $\pi(V)$.

By induction, we may reduce to the case $n = 1$.

Since φ is a dominant rational map, it induces an injection of fields $\varphi^*\colon k(W) \to k(V)$. We have to prove that the resulting field extension $k(V)/\varphi^*(k(W))$ is algebraic (since then $\operatorname{trdeg}(k(V)/k) = \operatorname{trdeg}(k(W)/k)$).

Look at the coordinate function X_1 on V (this is the coordinate which is discarded by π). Since φ is generically finite, V is strictly contained in $\mathbb{A}^1 \times W$. Hence there is a non-zero polynomial $f \in k[W][X_1]$ which vanishes on V. This gives an $k(W)$-algebraic relation satisfied by X_1 in $k(V)$. Therefore $k(V)$ is algebraic over $k(W)$ as required. $\square$

This allows us to restate the quasi-definition by saying: $\dim V = d$ if and only if there exists a generically finite dominant rational map $V \dashrightarrow \mathbb{P}^d$. Such a map $V \dashrightarrow \mathbb{P}^d$ exists, because V is birational to a hypersurface $H \subseteq \mathbb{P}^{d+1}$ and then then projection from any point $p \in \mathbb{P}^{d+1} \setminus H$ onto a hyperplane gives a generically finite dominant rational map $H \dashrightarrow \mathbb{P}^d$.

Dimension of a reducible variety. So far we have defined the dimension of an irreducible quasi-projective variety. The **dimension** of a reducible variety is defined to be the maximum of the dimensions of the irreducible components.

This makes sense because, if $V = V_1 \cup \cdots \cup V_r$ are the irreducible components of V, then $V_i \subseteq V$ implies that $\dim V_i \leq \dim V$ for each i. Meanwhile, we could find W_i for each i such that $V_i \subseteq W_i$ and $\dim W_i = \max(\dim V_1, \ldots, \dim V_r)$. Then every irreducible component of $W_1 \cup \cdots \cup W_r$ has the same dimension, so it makes sense to declare that $\dim(W_1 \cup \cdots \cup W_r)$ is equal to $\dim W_i$ for all i. But then $V \subseteq W_1 \cup \cdots \cup W_r$ so $\dim V \leq \dim W_i = \max(\dim V_1, \ldots, \dim V_r)$.

Facts about dimension. We begin with some simple facts.

Facts.

(1) If $\varphi\colon V \to W$ is a dominant rational map, then $\dim W \leq \dim V$. This follows from the fact that φ^* is an injection $k(W) \to k(V)$.

(2) $\dim(V \times W) = \dim V + \dim W$. This holds because if $\varphi\colon V \dashrightarrow \mathbb{A}^d$ and $\psi\colon W \dashrightarrow \mathbb{A}^e$ are generically finite dominant rational maps, then $(\varphi, \psi)\colon V \times W \dashrightarrow \mathbb{A}^{d+e}$ is a generically finite dominant rational map.

21. Dimension of closed subsets

Our aim today is to prove that if V is irreducible and W is a proper closed subset of V, then $\dim W < \dim V$. This is surprisingly hard!

Note that the condition that V is irreducible is necessary: otherwise we could take W to be an irreducible component of V. The key step in the proof relies on Lemma 19.2, which we proved using the resultant during the proof of Chevalley's theorem.

We will reduce to affine sets in $\mathbb{A}^n$ and use projections to do an induction on n. In order to carry this out, we need to know that it is possible to choose coordinates so that the projection becomes generically finite (and hence preserves dimension).

Lemma 21.1. Let $V \subseteq \mathbb{A}^n$ be an irreducible affine algebraic set, with $V \neq \mathbb{A}^n$. Then there exists a basis for $\mathbb{A}^n$ such that, if $\pi\colon \mathbb{A}^n \to \mathbb{A}^{n-1}$ means projection onto the first $n-1$ coordinates (with respect to this basis), then

$$\pi_{|V}\colon V \to W \text{ is generically finite,}$$

where W is the Zariski closure of $\pi(V)$.

Proof. V does not contain all lines through the origin, because $V \neq \mathbb{A}^n$. So we can pick a line L through the origin which is not contained in V but which does intersect V. After a linear change of coordinates, we may assume that L is the X_n-axis.

Letting $\pi\colon \mathbb{A}^n \to \mathbb{A}^{n-1}$ be projection onto the first $n-1$ coordinates with respect to this basis, we have $L = \pi^{-1}(0,\dots,0)$. Let W be the closure of $\pi(V)$.

Since L intersects V, $(0,\dots,0) \in W$. Because $L \not\subseteq V$, $(0,\dots,0)$ is in

$$U = \{x \in W : \pi^{-1}(x) \not\subseteq V\}.$$

Hence U is non-empty.

Furthermore U is open in W. To prove this, write equations for V as $f_i = \sum_j a_{ij} X_n^j$ where $a_{ij} \in k[X_1,\dots,X_{n-1}$. Then $W \setminus U$ is the set of points where the line $\pi^{-1}(x)$ is contained in V, that is,

$$W \setminus U = \{x \in W : a_{ij}(x) = 0 \text{ for all } i,j\}.$$

Hence $W \setminus U$ is closed.

For each $x \in U$, $\pi_{|V}^{-1}(x)$ is a proper closed subset of $\mathbb{A}^1$, so it is finite.

Thus we have found a non-empty subset of W above which all fibres of $\pi_{|V}$ are finite. $\square$

Lemma 21.2. Let V be an irreducible quasi-projective variety and let W be a closed subset of V.

(a) $\dim W \leq \dim V$.
(b) If $W \neq V$, then $\dim W < \dim V$.

Proof. It suffices to prove the lemma for each irreducible component of W, so we may assume that W is irreducible.

Pick a point $x \in W$. Then x is also in V, so we can apply problem 4 of sheet 4, to get an open set $U \subseteq V$ containing x which is affine. Since $x \in W$, $W \cap U$ is a non-empty open subset of W.

Replacing V by U and W by $W \cap U$ does not change their dimensions. So we may assume that V is affine. Suppose that $V \subseteq \mathbb{A}^n$.

We shall prove both parts of the lemma by induction on n.

In the base case, $W = \mathbb{A}^n$. Then also V must be equal to $\mathbb{A}^n$, so we have $\dim W = \dim V = n$ and (a) holds. Part (b) does not arise when $W = \mathbb{A}^n$.

Otherwise, we can apply Lemma 21.1 to W (not to V!). We can choose coordinates so that the projection $\pi \colon \mathbb{A}^n \to \mathbb{A}^{n-1}$ restricts to a generically finite map $\pi_{|W} \colon W \to W_1$, where W_1 is the Zariski closure of $\pi(W)$ in $\mathbb{A}^{n-1}$

Let V_1 be the Zariski closure of $\pi(V)$. By Lemma 20.2 and by fact (1) from the end of the previous lecture, we get

$$\dim W = \dim W_1, \quad \dim V_1 \leq \dim V. (*)$$

Part (a). By induction we get $\dim W_1 \leq \dim V_1$. Combined with (*) this proves part (a).

Part (b). Embed $\mathbb{A}^n = \mathbb{A}^{n-1} \times \mathbb{A}^1$ in $\mathbb{A}^{n-1} \times \mathbb{P}^1$, and write $\overline{V}$, $\overline{W}$ for the closures of V and W respectively in $\mathbb{A}^{n-1} \times \mathbb{P}^1$. By completeness of $\mathbb{P}^1$, $\pi(\overline{V}) = V_1$ and $\pi(\overline{W}) = W_1$.

Since W is strictly contained in V, we can apply Lemma 19.2 and get two cases:

(1) W_1 is strictly contained in V_1.
By induction, $\dim W_1 < \dim V_1$. Combined with (*), this proves (b).

(2) $\overline{V} = \mathbb{P}^1 \times V_1$.
By fact (2) from the end of the previous lecture, $\dim \overline{V} = 1 + \dim V_1$. By part (a), $\dim W_1 \leq \dim V_1$. Since V is open in $\overline{V}$, $\dim V = \dim \overline{V}$. Again combined with (*), this proves (b). $\square$

22. Dimension and intersections

Intersection with a hyperplane. We begin by studying intersections between a projective variety and hyperplanes. This is much simpler for projective varieties than for quasi-projective varieties, because then we know that there can be no intersections "hiding at infinity."

In order for the induction in the proof of the next proposition to work, it is essential to prove it for reducible varieties as well as irreducible ones.

Proposition 22.1. Let $V \subseteq \mathbb{P}^n$ be a projective variety.

There exists a sequence of hyperplanes $H_1, H_2, \dots, H_r \subseteq \mathbb{P}^n$ such that

$$V \cap H_1 \cap \dots \cap H_r = \emptyset$$

with $r \leq \dim V + 1$.

Proof. We prove this by induction on $\dim V$.

Choose a hyperplane H_1 which does not contain any irreducible component of V. Note that this is always possible: it suffices to choose one point in each component of V, giving a finite set of points; then we can choose a hyperplane H_1 which avoids that finite set.

We claim that $\dim(V \cap H_1) < \dim V$. To prove this, let W be an irreducible component of $V \cap H_1$. Then W must be contained in some irreducible component W' of V. We chose H_1 such that $W' \not\subseteq H_1$, and so W is a proper closed subset of W. Therefore by Lemma 21.2, $\dim W' < \dim W$. Since this holds for every irreducible component of $V \cap H_1$, we conclude that

$$\dim(V \cap H_1) < \dim V.$$

By induction, we can find hyperplanes $H_2, \dots, H_r \subseteq \mathbb{P}^n$ such that

$$(V \cap H_1) \cap H_2 \cdots \cap H_r = \emptyset.$$

We started these hyperplanes from 2, so there are $r-1$ of them, so we get

$$r - 1 \leq \dim(V \cap H_1) + 1 \leq \dim V. \qquad \square$$

Lemma 22.2. Let $V \subseteq \mathbb{P}^n$ be a projective variety.

If there exist hyperplanes $H_1, \dots, H_r \subseteq \mathbb{P}^n$ such that $V \cap H_1 \cdots \cap H_r$ is empty, then $r \geq \dim V + 1$.

Proof. Let $L = H_1 \cap \dots \cap H_r$. We may assume without loss of generality that the hyperplanes $H_1, \dots, H_r$ are independent (that is, L cannot be obtained as the intersection of any subset of these hyperplanes). Then L is a linear space of dimension $n - r$.

Let $f_1, \dots, f_r \in k[X_0, \dots, X_n]$ be linear homogeneous polynomials defining the hyperplanes $H_1, \dots, H_r$. (Independence of the hyperplanes is equivalent to linear independence of $f_1, \dots, f_r$.) Then

$$\varphi = [f_1 : \dots : f_r]$$

defines a rational map $\mathbb{P}^N \dashrightarrow \mathbb{P}^{r-1}$.

Since $V \cap H_1 \cdots \cap H_r = \emptyset$, the polynomials $f_1, \dots, f_r$ are never simultaneously zero on V. Hence φ restricts to a regular map on V. By completeness, the image $W = \varphi(V)$ is closed in $\mathbb{P}^{r-1}$.

We claim that $\varphi_{|V} \colon V \to W$ is generically finite. To prove this, let $w = [w_1 : \cdots : w_r]$ be a point in W. Assume without loss of generality that $w_1 \neq 0$. Then the preimage $\varphi_{|V}^{-1}(w)$ is given by

$$\{v \in V : w_j f_1(v) = w_1 f_j(v) \text{ for } 2 \leq j \leq r\}.$$

Each of the equations

$$w_j f_1(v) - w_1 f_j(v)$$

is a linear homogeneous polynomial, so it defines a hyperplane H'_j (depending on w). Let

$$L' = H'_2 \cap \cdots \cap H'_r.$$

A little linear algebra shows that $L = L' \cap H_1$. So

$$\varphi_{|V}^{-1}(w) \cap H_1 = V \cap L' \cap H_1 = V \cap L$$

and this is empty by assumption. By Lemma 17.6, every infinite closed subset of $\mathbb{P}^n$ has non-empty intersection with H_1, so $\varphi_{|V}^{-1}(w)$ must be finite.

Thus every fibre of $\varphi \colon V \to W$ is finite, so the map is certainly generically finite. By Lemma 20.2, we conclude that $\dim V = \dim W$. But $W \subseteq \mathbb{P}^{r-1}$ so $\dim W \leq r - 1$, completing the proof. $\square$

Let V be a projective variety of dimension $d > 0$. If we take a sequence of hyperplanes as in Proposition 22.1, we know that $V \cap H_1 \cap \cdots \cap H_r$ must become empty by the time r gets to $d + 1$. But by Lemma 22.2, it cannot happen earlier than that. It follows that in the sequence of inequalities

$$d = \dim V > \dim(V \cap H_1) > \dim(V \cap H_1 \cap H_2) > \cdots > \dim(V \cap H_1 \cap \cdots \cap H_d) \geq 0,$$

the dimension must simply go down by 1 each time, that is,

$$\dim(V \cap H_1 \cap \cdots \cap H_i) = d - i. \tag{*}$$

In Proposition 22.1, the only condition imposed on H_1 is that it does not contain an irreducible component of V. So we get the following:

Theorem 22.3. Let $V \subseteq \mathbb{P}^n$ be a projective variety of positive dimension.

Let $H \subseteq \mathbb{P}^n$ be a hyperplane which does not contain any irreducible component of V.

Then $\dim(V \cap H) = \dim V - 1$.

It is possible to prove Theorem 22.3 for the intersection between V and a hypersurface by using a Veronese embedding to reduce this to the case of an intersection with a hyperplane.

On the other hand, it is much harder to tell what the dimension will be for the intersection between V and 2 or more hyperplanes: the problem is that in order for (*) to apply, H_2 must satisfy the condition that it does not contain any irreducible component of $V \cap H_1$, and it may be hard to tell whether this happens or not.

Dimension and equations. Apply this to $\mathbb{P}^n$ itself. The zero set of a single homogeneous polynomial is a hypersurface of dimension $n-1$. The zero set of two homogeneous polynomials f_1, f_2 is an intersection $H_1 \cap H_2$ of two hypersurfaces. $H_1 \cap H_2$ has dimension $n-2$, as long as H_2 does not contain any irreducible component of H_1; this is equivalent to saying that f_1 and f_2 have no common factor. But once we look at three homogeneous polynomials f_1, f_2, f_3, there is no easy condition to tell whether the dimension of their zero set is equal to $n-3$ or not. So in general, all we get is the following inequality.

Corollary 22.4. If $f_1, \dots, f_r$ are homogeneous polynomials and V is the zero set of these polynomials, then
$$\dim V \geq n - r.$$
In particular, if $r \leq n$, then $V \neq \emptyset$.

Proof. Let H_i be the hypersurface defined by the equation $f_i = 0$. By the "hypersurface" version of Theorem 22.3, if H_i does not contain any irreducible component of $H_1 \cap \cdots \cap H_{i-1}$, then
$$\dim(H_1 \cap \cdots \cap H_i) = \dim(H_1 \cap \cdots \cap H_{i-1}) - 1.$$
On the other hand, if H_i does contain an irreducible component of $H_1 \cap \cdots \cap H_{i-1}$, then the dimension might not go down at all. In any case,
$$\dim(H_1 \cap \cdots \cap H_i) \geq \dim(H_1 \cap \cdots \cap H_{i-1}) - 1.$$
Iterating this proves the corollary. $\square$

23. Topological definition of dimension

Complete intersections. We saw last time that if you take r homogeneous polynomials, their zero set in $\mathbb{P}^n$ has dimension $\geq n - r$. We can't insist that the dimension be equal to $n - r$, because the zero set of one of the polynomials might contain the zero set of the others.

In reverse, we can ask: if $V \subseteq \mathbb{P}^n$ is a projective algebraic set of dimension $n-r$, do there exist r homogeneous polynomials which define V? Answer: not always.

There are two relevant definitions. The first one is more in the style of this course, but the second one turns out to be more natural.

Definition. Let $V \subseteq \mathbb{P}^n$ be an algebraic set of dimension $n - r$.

V is a **set-theoretic complete intersection** if there exist r homogeneous polynomials such that V is the zero set of these polynomials.

V is a **complete intersection** if there exist r homogeneous polynomials which generate the ideal of V.

Being a complete intersection is a stronger property than being a set-theoretic complete intersection.

Examples.

(1) The set of three non-collinear points in $\mathbb{P}^2$ is a set-theoretic complete intersection but not a complete intersection: there exist 2 polynomials defining this set, but you need 3 polynomials to generate its ideal.

(2) An irreducible example is the twisted cubic which we saw earlier. It is the 1-dimensional algebraic set $C \subseteq \mathbb{P}^3$ defined by the three equations

$$WY - X^2 = 0,\ WZ - XY = 0,\ XZ - Y^2 = 0.$$

Any two of these equations define a 1-dimensional algebraic set which has C as an irreducible component, but also has another irreducible component.

It is possible to find two polynomials which define the set C, for example

$$WY - X^2 = 0,\ WZ^2 - 2XYZ + Y^3 = 0.$$

But two polynomials cannot generate the ideal of C (the earlier three polynomials do generate the ideal of C).

(3) An example of a set which is not a set-theoretic complete intersection: take the two planes in $\mathbb{P}^4$:

$$P_1 = \{x \in \mathbb{P}^4 : x_1 = x_2 = 0\},\ P_2 = \{x \in \mathbb{P}^4 : x_3 = x_4 = 0\}.$$

These intersect in only one point, namely $[1 : 0 : 0 : 0 : 0]$.

The union $P_1 \cup P_2$ has dimension 2 but it needs 4 equations to define it.

(4) One can find examples of irreducible 2-dimensional algebraic sets of $\mathbb{P}^4$ which are not set-theoretic complete intersections, with a singularity which looks like two planes intersecting in one point.

We had to go to $\mathbb{P}^4$ to give explicit examples of non-set-theoretic complete intersections. It is unknown whether every irreducible curve in $\mathbb{P}^3$ is a set-theoretic complete intersection.

Topological definition of dimension. Now we can describe the dimension of a projective variety in terms of its topology.

Theorem 23.1. Let V be a projective variety.

The dimension of V is the maximum integer d such that there exists a chain of irreducible closed subsets

$$V \supseteq V_d \supsetneq V_{d-1} \supsetneq \cdots \supsetneq V_0 \supsetneq \emptyset.$$

Some care is required in the statement of this theorem to get the numbering right! The point is that $\dim V_i = i$, so V_0 is still non-empty. Note that $V = V_d$ if and only if V is irreducible; all the other inclusions must be strict. In Theorem 23.1, it is essential to require all the V_i to be irreducible. Otherwise we could make the chain arbitrarily long by inserting reducible sets with more and more components, all of dimension i, in between V_i and V_{i+1}.

Proof. First we prove that such a sequence with $d = \dim V$ exists.

Choose V_d to be an irreducible component of V whose dimension is equal to $\dim V$. Choose H as in Theorem 22.3 applied to V_d. Let V_{d-1} be an irreducible component in $V_d \cap H$ such that

$$\dim V_{d-1} = \dim(V_d \cap H) = \dim V - 1.$$

We can repeat this procedure, getting $V_i \subsetneq V_{i+1}$ with $\dim V_i = i$ until we get to V_0 with $\dim V_0 = 0$.

In the other direction, to show that there is no such sequence with $d > \dim V$, this follows immediately from the fact that $\dim V_i < \dim V_{i+1}$ (Lemma 21.2). □

Extending this to quasi-projective varieties. Theorem 23.1 holds for quasi-projective varieties as well as projective varieties. We will omit the proof, but it is not much harder. The idea is to apply the same argument to the projective closure $\overline{V}_d$. Almost all hyperplanes give $\dim(\overline{V}_d \cap H) = \dim V - 1$, but only a few hyperplanes case trouble by having a component of $\overline{V}_d \cap H$ which does not intersect V_d. So it is possible to find some hyperplane which gives $\dim(V_d \cap H) = \dim V - 1$ and then repeat.

Theorem 22.3 also applies to irreducible quasi-projective algebraic set $V \subseteq \mathbb{P}^n$, except that it could happen for a quasi-projective variety that $V \cap H = \emptyset$ (unlike the projective case – see Lemma 17.6). The precise statement is as follows:

Theorem 23.2. Let $V \subseteq \mathbb{P}^n$ be an irreducible quasi-projective algebraic set.

Let $H \subseteq \mathbb{P}^n$ be a hyperplane which does not contain V.

If $V \cap H \neq \emptyset$, then

$$\dim(V \cap H) = \dim V - 1.$$

This is much harder to prove than for projective varieties, requiring a result from algebra called Krull's Hauptidealsatz. The problem is that if we write V as $\overline{V} \cap U$, where $\overline{V}$ is the closure of V in $\mathbb{P}^n$ and U is an open set, then some components of $\overline{V} \cap H$ might be contained in the complement of U. *A priori*, it could therefore happen that the components of $V \cap H$ all have smaller dimension than $\overline{V} \cap H$ and the hard work is required to rule this out.

In $\mathbb{P}^n$, we can write down a chain of closed subsets

$$\mathbb{P}^n \supsetneq \mathbb{P}^{n-1} \supsetneq \mathbb{P}^{n-2} \supsetneq \cdots \supsetneq \mathbb{P}^1 \supsetneq \{\mathrm{pt}\} \supsetneq \emptyset.$$

This chain is *maximal* – we cannot insert another irreducible closed subset anywhere in the middle of it. But this chain, together with Theorem 23.1, is not enough to prove that $\dim \mathbb{P}^n = n$ – maybe there is a completely different chain which is longer. It turns out that that can't happen: one can prove that every maximal chain of irreducible closed subsets in an irreducible projective variety V has length equal to $\dim V$. But this is another hard theorem. It requires the same hard work as proving Theorem 23.2 for quasi-projective algebraic sets.

Fibre dimension theorem. (non-examinable)

We have now seen several definitions of dimension: via transcendence degree, via rational maps to hypersurfaces or to affine space, via chains of closed subsets. None of these is easy to compute except in simple cases (knowing that the chain of closed subsets definition works for any maximal chain means that it is sometimes usable). When we want to calculate the dimension of a particular variety, we often use the following powerful theorem.

Theorem 23.3. Let V, W be irreducible quasi-projective varieties and let $\varphi \colon V \to W$ be a surjective regular map. Then:

(1) For every $w \in W$, $\dim \varphi^{-1}(w) \geq \dim V - \dim W$.
(2) There exists a non-empty open subset $U \subseteq W$ such that $\dim \varphi^{-1}(w) = \dim V - \dim W$ for all $w \in U$.

Consequently,

$$\dim V - \dim W = \min_{w \in W} \dim \varphi^{-1}(w).$$

We generally use this theorem in situations where we know the dimension of either V or W and want to work out the other. If we can work out $\dim \varphi^{-1}(w)$ for just a single $w \in W$, then we get an inequality. If we can work out $\dim \varphi^{-1}(w)$ for w in some open set then we can work out the desired dimension exactly.

24. Tangent space and singular points

Zariski tangent space. We want to define what it means for a point in a variety to be singular. We will do this by using the tangent space to the variety at a point.

Let $V \subseteq \mathbb{A}^n$ be an affine algebraic set. We consider it with an embedding $\mathbb{A}^n$ because we will use coordinates in our definition of the tangent space. There is also an intrinsic definition using local rings, which we will not cover.

Choose a point $x \in V$. For each polynomial $f \in k[X_1, \dots, X_n]$, we let df_x denote the linear map $k^n \to k$ given by

$$df_x(a_1, \dots, a_n) = \sum_{i=1}^{n} \left.\frac{\partial f}{\partial X_i}\right|_x a_i.$$

Informally: df_x sends a vector $\underline{a} \in k^n$ to the "directional derivative" of f at x along that vector. Thus $df_x(\underline{a}) = 0$ precisely for those directions in which f is stationary at x. Since the polynomials in $\mathbb{I}(V)$ are zero on V, we should expect polynomials in $\mathbb{I}(V)$ to be stationary along "tangent directions" to V. This motivates the following definition.

Definition. Let $V \subseteq \mathbb{A}^n$ be an affine algebraic set and let $x \in V$. The **tangent space** to V at x is

$$T_x V = \bigcap_{f \in \mathbb{I}(V)} \ker df_x \subseteq k^n.$$

This is sometimes called the **Zariski tangent space** when it is necessary to distinguish it from other kinds of tangent space.

In our definition of df_x, we used partial derivatives. Because we are only differentiating polynomials, this can be defined purely algebraically and therefore makes sense over any field, even in positive characteristic where there is no analysis. However derivatives of polynomials can behave surprisingly in positive characteristic: over a field of characteristic p we have

$$\frac{\mathrm{d}}{\mathrm{d}X} X^p = pX^{p-1} = 0$$

so it is possible for a non-constant polynomial to have derivative equal to zero.

Similarly the informal motivation for the definition relied on our intuition from analysis about what happens over $\mathbb{C}$. Even over $\mathbb{C}$, our analytic intuition only works correctly if the variety is non-singular at x, as we will see later when we define non-singular points.

Let $f_1, \dots, f_n$ be a finite list of polynomials which generate $\mathbb{I}(V)$. It is easy to prove that the tangent space $T_x V$ can be calculated just by looking at this finite list of functions:

$$T_x V = \bigcap_{i=1}^{n} \ker d(f_i)_x.$$

Thus it is straightforward to calculate tangent spaces in practice. There is just one thing which needs to be careful of: the functions $f_1, \dots, f_n$ must generated the ideal $\mathbb{I}(V)$. It is not enough to take a list of functions which define V as an algebraic set but do not generate the ideal.

Example. As a very simple example, consider the line L in $\mathbb{A}^2$ defined by the equation $X = 0$. At the point $(0, 0)$, we have

$$dX_0(a_1, a_2) = \left.\frac{\partial X}{\partial X}\right|_0 a_1 + \left.\frac{\partial X}{\partial Y}\right|_0 a_2 = 1.a_1 + 0.a_2 = a_1.$$

Since X generates $\mathbb{I}(L)$, we get

$$T_0V = \ker dX_0 = \{(a_1, a_2) \in k^2 : a_1 = 0\}.$$

This is what we should expect: the tangent space to a line is a line in the same direction.

However, noting that L is also the zero set of the polynomial $g = X^2$, we could try to calculate

$$dg_0(a_1, a_2) = \left.\frac{\partial X^2}{\partial X}\right|_0 a_1 + \left.\frac{\partial X^2}{\partial Y}\right|_0 a_2 = 0.a_1 + 0.a_2 = 0.$$

So $\ker dg_0 = k^2$ which is too big. Thus using functions which do not generate the whole ideal of the variety may give the wrong answer.

Example. Our intuitive idea of tangent spaces is less useful at singular points.

Let V be the nodal cubic curve defined by the polynomial

$$f(X, Y) = Y^2 - X^2(X + 1).$$

We have

$$\frac{\partial f}{\partial X} = -3X^2 - 2X, \quad \frac{\partial f}{\partial Y} = 2Y.$$

Hence

$$\left.\frac{\partial f}{\partial X}\right|_{(0,0)} = \left.\frac{\partial f}{\partial Y}\right|_{(0,0)} = 0$$

and so $df_{(0,0)}$ is the zero map. Therefore $T_{(0,0)}V = k^2$.

Perhaps this is unsurprising: the curve V has two tangent lines at the origin, $y = x$ and $y = \pm x$ so both of these need to be contained in $T_{(0,0)}V$, but by definition $T_{(0,0)}V$ is a vector space so this forces it to contain all of k^2.

Example. Consider the cuspidal cubic curve defined by the polynomial

$$g(X, Y) = Y^2 - X^3.$$

We have

$$\frac{\partial f}{\partial X} = -3X^2, \quad \frac{\partial f}{\partial Y} = 2Y.$$

Again both of these vanish at $(0, 0)$, so $T_{(0,0)}V = k^2$.

This is more surprising than the previous example, because looking at a picture suggests that this curve has only the x-axis as a tangent line at the origin. This demonstates that we cannot rely on geometric intuition to calculate the tangent space at singular points: it is necessary to use the algebraic definition.

Singular points. In the previous two examples, I have referred informally to the idea of singular points. Now we are ready to define these.

Intuition suggests that, at non-singular points, the dimension of the tangent space should be equal to the dimension of the algebraic set. The above examples indicate that this breaks down at singular points. This motivates us to define a singular point to be a point $x \in V$ where $\dim T_x V \neq \dim V$.

However this simple definition only works correctly for irreducible algebraic sets. This is because, if we take a union of two irreducible components of different dimensions, say $V = V_1 \cup V_2$ where $\dim V_1 = 1$ and $\dim V_2 = 2$, then by definition $\dim V = 2$. If we take a point $x \in V_1$ which is not in the intersection $V_1 \cap V_2$, then whether x is a singular point of V should not care about V_2 – we need to compare $\dim T_x V$ against $\dim V_1 = 1$, not against $\dim V = 2$..

In order to fix this and corerctly define singular points of reducible algebraic sets, we introduce a new definition:

Definition. Let V be a quasi-projective variety and let x be a point of V. The **local dimension** of V at x, written $\dim_x V$, is the maximum of the dimensions of those irreducible components of V which contain x.

(Taking the maximum of the dimensions of components fits with the way we defined the dimension of a reducible variety, but now we are ignoring components which do not contain x.)

Thus in our previous example $V = V_1 \cup V_2$, $\dim_x V = 2$ if $x \in V_2$ (including if $x \in V_1 \cap V_2$) while $\dim_x V = 1$ if $x \in V_1 \setminus (V_1 \cap V_2)$.

If we consider the topological definition of dimension (Theorem 23.1), we can use this to define local dimension as follows: $\dim_x V$ is the maximum integer d such that there exists a chain of irreducible closed subsets

$$V \supseteq V_d \supsetneq V_{d-1} \supsetneq \cdots \supsetneq V_1 \supsetneq V_0 = \{x\} \supsetneq \emptyset.$$

Now we can define singular points of a reducible algebraic set by using local dimension.

Definition. Let V be an affine algebraic set and let $x \in V$. Then x is a **singular point** of V if

$$\dim T_x V \neq \dim_x V.$$

(We may also express this as "V is **singular** at V.")

We will prove later that $\dim T_x V \geq \dim_x V$ always, so we could equivalently state this definition as: x is a singular point of V if $\dim T_x V > \dim_x V$.

If $V = \bigcup_{i=1}^r V_i$ is a union of irreducible components, then for any point x which lies in only one irreducible component V_i, we have

$$\dim_x V = \dim_x V_i$$

by definition. A little algebra also shows that $T_xV = T_xV_i$ and so V is singular at x if and only if V_i is singular at x.

On the other hand, if x lies in an intersection of two or more irreducible components of V, then it turns out that x is always a singular point of V. This is intuitively sensible, but requires too much algebra to prove in this course (specifically, it requires Nakayama's lemma).

Independence of embedding. We have defined the tangent space only for affine algebraic sets, depending on the embedding into $\mathbb{A}^n$. Thus we have defined the tangent space as a subspace of k^n (where n is the dimension of the affine space into which we embed V). We can canonically identify k^n with the tangent space of $\mathbb{A}^n$ at any point, so one way of looking at this is to say: we have $V \subseteq \mathbb{A}^n$ and we have identified T_xV as a subspace of $T_x\mathbb{A}^n$.

There is an intrinsic way to define the tangent space which does not depend on an embedding into affine space, using a bit more algebra. We will omit this here, and just state:

Fact. The dimension of T_xV is independent of the embedding of V into $\mathbb{A}^n$.

(In other words, if $V \subseteq \mathbb{A}^n$ and $W \subseteq \mathbb{A}^m$ are isomorphic, with $\varphi\colon V \to W$ being an isomorphism, then $\dim T_xV = \dim T_{\varphi(x)}W$.)

One can prove this by showing that an isomorphism $\varphi\colon V \to W$ induces an isomorphism of vector spaces $d\varphi_x\colon T_xV \to T_{\varphi(x)}W$ (using the chain rule for partial derivatives).

This implies that whether a point is singular or not is independent of the embedding.

The dimension of the tangent space actually contains slightly more information than just whether the point is singular or not – for example, if we have a curve in $\mathbb{A}^3$ then its tangent space could have dimension 1 (non-singular) or 2 or 3 (both singular, but if $\dim T_xV = 3$ then it is a "bigger" or "more complicated" singularity). Again this extra information is independent of the embedding.

25. Singular points on quasi-projective varieties

I began this lecture with some remarks on rational maps. In order to preserve the logical ordering of these notes, I have put those remarks in Appendix B (after the mastery material in Appendix A) and carry on with singular points here.

Tangent spaces and isomorphisms of affine algebraic sets. Last time we observed that an isomorphism of affine algebraic sets induces isomorphisms between the tangent spaces at corresponding points. This is proved as follows:

Lemma 25.1. Let $V \subseteq \mathbb{A}^n$ and $W \subseteq \mathbb{A}^m$ be affine algebraic sets. Let $\varphi\colon V \to W$ be an isomorphism.

For any $x \in V$, if $y = \varphi(x) \in W$, then φ induces an isomorphism

$$d\varphi_x\colon T_xV \to T_yW.$$

Outline proof. Choose polynomials $f_1, \dots, f_m$ such that $\varphi = (f_1, \dots, f_m)$. Define a linear map $k^n \to k^m$ by the matrix

$$\left(\frac{\partial f_i}{\partial X_j}\bigg|_x \right).$$

We define $d\varphi_x$ to be the restriction of this map to T_xV. (Recall that T_xV is a subspace of k^n and T_yW is a subspace of k^m.)

Using the chain rule for partial derivatives, one can check that:

(i) $d\varphi_x$ maps T_xV into T_yW.
(ii) $d\varphi_x$ is independent of the choice of polynomials representing φ.
(iii) Because φ is an isomorphism of algebraic sets, $d\varphi_x$ is an isomorphism of vector spaces. □

We don't actually need an isomorphism between V and W themselves, just an isomorphism between open subsets (in other words, a birational map $V \dashrightarrow W$).

Lemma 25.2. Let $V \subseteq \mathbb{A}^n$ and $W \subseteq \mathbb{A}^m$ be affine algebraic sets. Let $U_1 \subseteq V$ and $U_2 \subseteq W$ be open subsets, and let $\varphi\colon U_1 \to U_2$ be an isomorphism of quasi-projective varieties.

For any $x \in U_1$, if $y = \varphi(x) \in U_2$, then φ induces an isomorphism

$$d\varphi\colon T_xV \to T_yW.$$

Note that in this lemma, we talk about T_xV and T_yW rather than T_xU_1 and T_yU_2, even though the isomorphism is between U_1 and U_2. We have to do this because U_1 and U_2 need not be affine algebraic sets, so we have not yet defined their tangent spaces. Nevertheless, since U_1 is open in V, we intuitively expect $T_xU_1 = T_xV$ for $x \in U_1$ (the tangent space at x only depends on what happens near x, and "near x" there is no difference between U_1 and V). Hence the lemma makes sense.

The proof of Lemma 25.2 is much the same as Lemma 25.1 – we just write down fractions of polynomials which represent φ instead of polynomials. Because these fractions will have non-zero denominators at x, the calculus still works out fine.

Singular points on quasi-projective varieties. Let V be a quasi-projective variety. We know from problem sheet 4, problem 4, that for any point $x \in V$ we can find an open subset $U \subseteq V$ which contains x and such that there is an isomorphism $\varphi \colon U \to W$ where W is an affine algebraic set.

We define the "dimension of the tangent space to V at x" to be the dimension of $T_{\varphi(x)W}$. Of course there are many choices for U, W and φ, but Lemma 25.2 guarantees that $\dim T_{\varphi(x)} W$ will be independent of the choice. However, we do not get a specific vector space independent of choices which we can call $T_x V$. Of course, knowing the dimension is enough to determine a vector space up to isomorphism, but that is not as good as having a concrete vector space. There is a way of defining a concrete vector space which is "the tangent space" to V at x using local rings, but we will not do that here.

Knowing the dimension of tangent spaces is enough to define singular points. Just as in the case of an affine algebraic set, we define a **singular point** of V to be a point $x \in V$ where $\dim T_x V \neq \dim_x V$. The set of singular points of V is called the **singular locus** of V and denoted $\operatorname{Sing} V$.

Appendix A. Mastery material: The singular locus

Our aim is to prove the following theorem:

Theorem A.1. Let V be an irreducible affine algebraic set. Then $\operatorname{Sing} V$ is a proper closed subset of V.

The theorem is true for irreducible quasi-projective varieties as well as affine ones – this can easily be checked by using a cover by affine open subsets. Furthermore, it is also true for reducible varieties, but we will not prove this as it requires more algebra.

A key intermediate step in the proof, which is also interesting in its own right, is the fact that $\dim T_x V \geq \dim_x V$ for every point $x \in V$.

The singular locus of a hypersurface. We begin by proving Theorem A.1 for a hypersurface.

Let $V \subseteq \mathbb{A}^n$ be a hypersurface and let f be a polynomial which generates $\mathbb{I}(V)$. Since f generates $\mathbb{I}(V)$, the tangent space $T_x V$ is just $\ker df_x$. In other words $T_x V$ is the kernel of a linear map $k^n \to k$, and so

$$\begin{aligned} &\dim T_x V = n - 1 \text{ if } df_x \text{ is not the zero map;} \\ &\dim T_x V = n \text{ if } df_x \text{ is the zero map.} \end{aligned}$$

For any point $x \in V$, we have $\dim_x V = \dim V = n - 1$. Hence

$$\operatorname{Sing} V = \{x \in V : df_x = 0\}.$$

Going back to the definition of df_x, we can write this as

$$\operatorname{Sing} V = \left\{ x \in V : \left. \frac{\partial f}{\partial X_i} \right|_x = 0 \text{ for } i = 1, \dots, n \right\}.$$

This may look like the definition of singular points which you have seen before for curves in $\mathbb{A}^2$.

For each i, $\partial f / \partial X_i$ is a polynomial. Therefore:

Lemma A.2. For any hypersurface $V \subseteq \mathbb{A}^n$, $\operatorname{Sing} V$ is a closed subset of V.

We want to show that for a hypersurface V, $\operatorname{Sing} V \neq V$. First we prove a lemma on derivatives in positive characteristic. We saw last time that X^p has derivative zero in characteristic p, and clearly this applies also to X^{ip} for any positive integer i. We prove that these span all the polynomials with zero derivative.

Lemma A.3. Let k be a field of characteristic $p > 0$. Let $f \in k[X]$ be a polynomial. If $\frac{\mathrm{d}f}{\mathrm{d}X} = 0$, then for every term of f, the exponent of X is a multiple of p, that is,

$$f = \sum_{i=0}^{d} a_{ip} X^{ip}.$$

Proof. Consider a term a_jX^j in f. This term differentiates to ja_jX^{j-1}. No other term of f differentiates to a scalar multiple of X^{j-1}, so this term can never cancel with another term in $\mathrm{d}f/\mathrm{d}X$.

Hence if $\mathrm{d}f/\mathrm{d}X = 0$, then $ja_j = 0$ (in k) for every j. If j is not a multiple of p, then j is invertible in k so this forces $a_j = 0$. Thus only terms where j is a multiple of p can appear in f. □

Proposition A.4. If V is a non-empty hypersurface, then $\operatorname{Sing} V$ is strictly contained in V.

Proof. Assume for contradiction that $\operatorname{Sing} V = V$. Then $\partial f/\partial X_1, \dots, \partial f/\partial X_n$ are all zero on V.

Since f generates $\mathbb{I}(V)$, this implies that f divides $\partial f/\partial X_i$ for each i. But $\partial f/\partial X_i$ has strictly smaller X_i-degree than f. This forces $\partial f/\partial X_i = 0$ for each i (as a polynomial in $k[X_1, \dots, X_n]$).

Over a field of characteristic zero, this implies that f is constant. But then V would be empty, contradicting the hypothesis.

Over a field of characteristic $p > 0$, by Lemma A.3, the fact that $\partial f/\partial X_i = 0$ implies that very term of f must have its X_i-exponent being a multiple of p. Since this holds for all i, each term of f is a p-th power (the constant in the term must be a p-th power because k is algebraically closed).

But the binomial expansion implies that

$$(a + b)^p = a^p + b^p$$

over a field of characteristic p. So if every term of f is a p-th power, then f itself is a p-th power. But then the ideal generated by f is not a radical ideal. Via the Nullstellensatz, this contradicts the assumption that f generates $\mathbb{I}(V)$. □

The singular locus of an irreducible variety.

Lemma A.5. Let $V \subseteq \mathbb{A}^n$ be an affine algebraic set. For any integer d, the set

$$\Sigma_d(V) = \{x \in V : \dim T_xV > d\}$$

is a closed subset of V.

Proof. Choose polynomials $f_1, \dots, f_m$ which generate $\mathbb{I}(V)$. Recall that

$$T_xV = \bigcap_{i=1}^{m} \ker d(f_i)_x.$$

In other words, T_xV is the kernel of the matrix

$$M_x = \left(\frac{\partial f_i}{\partial X_j}\bigg|_x \right)_{ij}$$

which represents a linear map $k^n \to k^m$

By the rank–nullity theorem, $\dim T_xV$ is equal to $n - \operatorname{rk} M_x$. Hence

$$\Sigma_d(V) = \{x \in V : \operatorname{rk} M_x < n - d\}.$$

By linear algebra, $\operatorname{rk} M_x < n-d$ is equivalent to: every $(n-d)\times(n-d)$ submatrix of M_x has determinant zero.

The determinant of a submatrix of M_x is a polynomial, hence this gives us polynomial equations defining $\Sigma_d(V)$. □

Lemma A.6. Let V be an irreducible affine algebraic set. Then the non-singular points of V are dense in V.

Proof. By Proposition 11.5, V is birational to a hypersurface $H \subseteq \mathbb{A}^{d+1}$. By Lemma 16.1 we can find non-empty open sets $U \subseteq V$ and $J \subseteq H$ such that there is an isomorphism $\varphi \colon U \to J$.

By Lemma A.2 and Proposition A.4, the non-singular points of H form a non-empty open subset $H_{ns} \subseteq H$. Since H is irreducible, H_{ns} must intersect J.

Since φ is continuous, $A = \varphi^{-1}(H_{ns} \cap J)$ is an open subset of U. Since $H_{ns} \cap J \neq \emptyset$ and since φ is surjective onto J, A is non-empty. Since V is irreducible, we conclude that A is dense in V.

For any $x \in A$, let $y = \varphi(x) \in H_{ns} \cap J$. Lemma 25.2 tells us that T_xV is isomorphic to T_yH. Since $y \in H_{ns}$, we have $\dim T_yH = \dim H$. Thus

$$\dim T_xV = \dim T_yH = \dim H = \dim V$$

so V is non-singular at x. □

Lemma A.7. Let V be an irreducible affine algebraic set. For every $x \in V$, $\dim T_xV \geq \dim_x V$.

Proof. Since V is irreducible, $\dim_x V = \dim V$ for every $x \in V$ so we can work with $\dim V$ instead of $\dim_x V$. (The lemma is true for reducible V as well, but we do not have the tools to prove it when $\dim_x V$ is not constant.)

Let $d = \dim V$ and consider the set $\Sigma_{d-1}(V)$ as in Lemma A.5. By Lemma A.5, $\Sigma_{d-1}(V)$ is closed. Every non-singular point of V is in $\Sigma_{d-1}(V)$, so Lemma A.6 implies that $\Sigma_{d-1}(V)$ is dense in V.

Since $\Sigma_{d-1}(V)$ is closed and dense in V, we conclude that $\Sigma_{d-1}(V) = V$. □

Theorem A.8. Let V be an irreducible affine algebraic set. Then $\operatorname{Sing} V$ is a proper closed subset of V.

Proof. By Lemma A.7, $x \in V$ is a singular point if and only if $\dim T_xV > \dim_x V$. Again since V is irreducible, we can replace $\dim_x V$ by $\dim V$.

Thus

$$\operatorname{Sing} V = \Sigma_d(V)$$

where $d = \dim V$. So Lemma A.5 tells us that $\operatorname{Sing} V$ is closed in V. Lemma A.6 implies that $\operatorname{Sing} V$ is properly contained in V. □

Appendix B. Rational maps as equivalence classes

I made these remarks came at the beginning of lecture 25.

The way in which I defined rational maps between quasi-projective varieties was a bit confusing, so I will try to clear this up. Recall the definition I gave in lecture 15:

Definition. Let $V \subseteq \mathbb{P}^n$ and $W \subseteq \mathbb{P}^m$ be irreducible quasi-projective algebraic sets. A **rational map** $\varphi\colon V \dashrightarrow W$ is determined by a sequence of homogeneous polynomials $f_0, \dots, f_m \in k[X_0, \dots, X_n]$ of the same degree such that:

(1) $f_0, \dots, f_m$ are not all identically zero on V;
(2) there is a Zariski dense set $A \subseteq V$ such that, for all $x \in A$, the homogeneous coordinates $[f_0(x) : \dots : f_m(x)]$ make sense and define a point in W.

Two sequences of polynomials $[f_0 : \dots : f_m]$ and $[g_0 : \dots : g_m]$ represent the same rational map if the homogeneous coordinates

$$[f_0(x) : \dots : f_m(x)], [g_0(x) : \dots : g_m(x)]$$

represent the same point in $\mathbb{P}^m$ wherever both expressions make sense.

This was intended to be an informal way of defining rational maps as equivalence classes for a certain equivalence relation. The part "a rational map is determined by ..." describes the set on which we put the equivalence relation, and "two sequences represent the same rational map if ..." defines the equivalence relation itself. I will now define this equivalence relation formally, in the hope of reducing confusion.

Field of fractions. Before defining rational maps as equivalence classes, let's ask: why is that a sensible thing to do? Think back to the definition of rational functions on an *affine* variety V. They are defined as the field of fractions of $k[V]$.

The field of fractions of an integral domain R is defined as a set of equivalence classes – namely, you take the set

$$\{(a, b) \in R^2 : b \neq 0\}$$

and the equivalence relation

$$(a, b) \sim (c, d) \text{ if } ad = bc.$$

The field of fractions of R is defined to be the set of equivalence classes for this relation.

But normally we don't think of fractions as equivalence classes. We just write down one representative, with the special notation $\frac{a}{b}$, and then manipulate it by the normal rules for manipulating fractions.

If $R = \mathbb{Z}$, then often it may make sense to reduce fractions to "lowest terms" representatives (and if we impose the condition $b > 0$, then every fraction has a unique lowest terms representative). But if R is not a UFD, then we do not have special "lowest terms" representatives for fractions.

Rational maps as equivalence classes. Let $V \subseteq \mathbb{P}^n$ and $W \subseteq \mathbb{P}^m$ be irreducible quasi-projective algebraic sets.

Let S denote the set of sequences $(f_0, \dots, f_m) \in k[X_0, \dots, X_n]^{m+1}$ such that:

(1) $f_0, \dots, f_m$ are homogeneous of the same degree;
(2) $f_0, \dots, f_m$ are not all identically zero on V (note that this looks a little like the $b \neq 0$ condition in defining the field of fractions);
(3) there exists a Zariski dense set $A \subseteq V$ such that, for all $x \in A$, the homogeneous coordinates$[f_0(x) : \dots : f_m(x)]$ make sense and define a point in W.

Define an equivalence relation $\sim$ on S by: $(f_0, \dots, f_m) \sim (g_0, \dots, g_m)$ if

$$[f_0(x) : \dots : f_m(x)] = [g_0(x) : \dots : g_m(x)] \in \mathbb{P}^m$$

for all $x \in V$ where both expressions make sense. We could write this more algebraically as: $(f_0, \dots, f_m) \sim (g_0, \dots, g_m)$ if

$$f_i g_j = f_j g_i \text{ for all } i, j.$$

Observe that this resembles the equivalence relation used in defining the field of fractions.

These two definitions of S and $\sim$ are more formal ways of writing the two parts of the definition of a rational map (note that the conditions in the definitions of S and $\sim$ are the same as the conditions in the earlier definition of rational maps). Having defined S and $\sim$, we then define a **rational map** $\varphi\colon V \dashrightarrow W$ to be an equivalence class for $\sim$.

One needs to check that $\sim$ really is an equivalence relation – this is a detail which was hidden in my "informal" statement of the definition. This uses the fact that V is irreducible and that if two polynomials are equal on a Zariski dense set, then they are equal everywhere.

The **domain of definition** of a rational map is defined to be the union of the sets where each representative of the equivalence class makes sense as a map.

Just as with fractions, we usually just write down a single representative for a rational map. There is a special notation for representatives of rational maps, namely $[f_0 : \dots : f_m]$.

We have seen examples on problem sheets of UFD-like situations where one can choose a "lowest terms" representative for the rational map, but this is not always possible.