1. Some affine geometry

Apollonius was perhaps the greatest of the Greek geometers. He lived around 200 BC. One of his most famous results is:

The locus of points, whose distances from two fixed points are in a constant ratio, is a circle.

This is one of the harder results one can prove using only classical geometry. However using just a little algebra makes the result almost completely trivial.

Change coordinates so that the first point A is at the origin and the second point B lies on the x-axis,

$$A = (0, 0) \qquad\qquad B = (a, 0).$$

Let $P = (x, y)$ be a general point of the locus described above. Let m be the fixed ratio. By assumption

$$|PA| = m|PB|.$$

So we have

$$|PA|^2 = x^2 + y^2 \qquad\qquad |PB|^2 = (x - a)^2 + y^2.$$

Thus

$$x^2 + y^2 = m^2(x - a)^2 + m^2y^2.$$

After a little manipulation we get

$$(x - \alpha)^2 + y^2 = \alpha^2 - \beta$$

for appropriate α and β, that is, we get a circle.

There are two general principles to be gleaned from this example:

- To solve problems in analytic geometry, a little bit of algebra goes a long way.
- Since we introduce coordinates to solve this problem, we get to choose where to put the origin.

Definition 1.1. *Let K be a field.* ***Affine n-space over K****, is a copy of a K-vector space V of dimension n. An* ***affine linear subspace*** *Λ is is the translate of a linear subspace of V.*

$\mathbb{A}^n_K$ is the copy of affine space associated to the standard vector space $V = K^n$ of dimension n.

In other words, affine space is nothing more than a vector space without a preferred point and a line in the affine plane is what a calculus student would call a line but not an undergraduate math major. We will invariably drop the word affine.

Definition 1.2. *Let K be an algebraically closed field.*

*An **affine subvariety** V of $\mathbb{A}^n_K$ is any subset of $\mathbb{A}^n_K$ defined by the zeroes of a collection of polynomials.*

Example 1.3. *Any linear space is an affine variety; any conic is an affine variety (a parabola, circle, ellipse, hyperbola, etc).*

$$C = \{\, (x, y) \in \mathbb{A}^2_K \mid y^2 = x^2 + x^3 \,\}$$

*is an affine variety called a **nodal cubic**; it is the zero set of $y^2-x^2-x^3$. Similarly*

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

*is an affine variety called a **cuspidal cubic**; it is the zero set of y^2-x^3.*

If S is a set of polynomials in the polynomial ring $K[x_1, x_2, \dots, x_n]$ it is convenient to let $V(S)$ be the common zero set.

Definition-Lemma 1.4. *The **Zariski topology** on $\mathbb{A}^n_K$ is the topology whose closed subsets are the affine subvarieties.*

Proof. The empty set is defined by the polynomial 1 and $\mathbb{A}^n_K$ is defined by the polynomial 0.

If V_α are closed subsets then we may find subsets $S_\alpha \subset K[x_1, x_2, \dots, x_n]$ such that $V_\alpha = V(S_\alpha)$. If

$$S = \bigcup_\alpha S_\alpha \qquad \text{then} \qquad V = \bigcap_\alpha V(S),$$

so that the intersection of closed sets is closed.

If V_1 and V_2 are two closed sets defined by polynomial subsets S_1 and S_2 so that $V_i = V(S_i)$ then

$$V = V_1 \cup V_2 = V(S),$$

where $S = S_1 S_2$ is the set of all products. Thus the union of two closed subsets is closed and we have a topology. □

Example 1.5. *Note that any finite subset of $\mathbb{A}^n_K$ is a Zariski closed subset.*

Example 1.6. *Consider closed subsets of $\mathbb{A}^1_K$. If $f(x) \in K[x]$ is a polynomial in one variable then $f(x)$ factors,*

$$f(x) = \lambda(x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_n).$$

Thus either $f(x)$ is identically zero, or constant but never zero, or $f(x)$ vanishes at only finitely many points $\alpha_1, \alpha_2, \dots, \alpha_n$. In particular the only proper closed subsets of $\mathbb{A}^1_K$ are finite subsets. Thus the Zariski topology is not Hausdorff.

Given an affine subvariety V it is natural to wonder about the degrees of polynomials definining V. Let f be a polynomial with no constant term. Note that f and $x_1 f$, $x_2 f$, $\dots$, $x_n f$ have the same zero set, so clearly we should focus on minimising the degrees of a defining set S.

If V consists of d points in $\mathbb{A}^n_K$ then it is not hard to see that V is defined by polynomials of degree d. Moreover, in general, we cannot do better than this. If $n = 1$ then the smallest degree of a polynomial vanishing at d points is d; more generally if we take d points on a line then and if $f(x_1, x_2, \dots, x_n)$ is a polynomial vanishing on those points but not on the line then f has degree at least d.

Definition 1.7. *The* ***span of a subset*** $V \subset \mathbb{A}^n_K$ *is the smallest linear space containing* V*.*

We say that the points of V *are in* ***linear general position*** *if any subset of* $k \leq n$ *points spans a linear space of dimension* $k - 1$*.*

Remark 1.8. *If* V *has at least* $n+1$ *points then* V *is in linear general position if and only if any subset of cardinality* $n+1$ *spans the whole of* $\mathbb{A}^n_K$*.*

Note that a linear space is the same as a Zariski closed subset defined by linear polynomials. A **hyperplane** is a linear space defined by a single equation.

Theorem 1.9. *If* $V \subset \mathbb{A}^n_K$ *is any set of* $d \leq 2n$ *points in linear general position then* V *is defined by quadratic polynomials.*

Proof. We do the case $d = 2n$; the general case follows in a similar fashion.

We have

$$V = \{ p_1, p_2, \dots, p_{2n} \}.$$

Let q be a point which belongs to the zero set of every quadratic polynomial which vanishes on the whole of V. We have to show that $q \in V$.

Suppose that we decompose V into two subsets of cardinality n, $V = V_1 \cup V_2$. Both subsets V_1 and V_2 define hyperplanes H_1 and H_2, both of which are defined by linear polynomials. The union $H_1 \cup H_2$ is defined by the product, a quadratic polynomial. By what we just said, $q \in H_1 \cup H_2$, so that q must belong to either H_1 or H_2.

Let $k \leq n$ be the smallest integer such that q belongs to the span of k points of $p_1, p_2, \dots, p_{2n}$. Renumbering we may assume that q belongs to the span of $p_1, p_2, \dots, p_k$. Pick a subset Σ of cardinality $n - k + 1$ of $V - \{ p_1, p_2, \dots, p_k \}$. Let Λ be the hyperplane spanned by p_2, p_3, $\dots$, p_k and Γ. Note that q does not belong Λ, since p_1 belongs to the span of p_2, p_3, $\dots$, p_k and q and Λ doesn't contain p_1.

Thus q belongs to the hyperplane spanned by the other n points of V. Varying Γ, q must belong to the intersection of these hyperplanes, which is just p_1. Thus $q = p_1$. $\square$

Definition-Lemma 1.10. *Let $V \subset \mathbb{A}^n_K$ be any subset.*

*The **ideal** $I = I(V)$ of V is the set of all polynomials vanishing on V,*

$$I = \{\, f \in K[x_1, x_2, \dots, x_n] \,|\, f(v) = 0, v \in V \,\} \trianglelefteq K[x_1, x_2, \dots, x_n].$$

Proof. We check I is an ideal. $0 \in I$ and so I is non-empty. If f and $g \in I$ then and p and $q \in K[x_1, x_2, \dots, x_n]$ then

$$(pf + qg)(v) = p(v)f(v) + q(v)g(v) = 0 \qquad \text{for all} \qquad v \in V,$$

so that $pf + qg \in I$ and I is an ideal. $\square$

Lemma 1.11.

(1) If $S_1 \subset S_2 \subset K[x_1, x_2, \dots, x_n]$ then $V(S_2) \subset V(S_1)$.
(2) If $V_1 \subset V_2 \subset \mathbb{A}^n_K$ then $I(V_2) \subset I(V_1)$.
(3) If V_1 and V_2 are two subsets of $\mathbb{A}^n_K$ and $I_i = I(V_i)$ then

$$I(V_1 \cup V_2) = I(V_1) \cap I(V_2).$$

(4) $V \subset \mathbb{A}^n_K$ is any subset then $V(I(V)$ is the Zariski closure of V.

Proof. Straightforward. $\square$

If $S \subset K[x_1, x_2, \dots, x_n]$ is a collection of polynomials it is natural to consider the ideal of all polynomials vanishing on the zero set of S, that is, to consider $I(V(S))$. The first guess is that this is simply the ideal generated by S.

Example 1.12. *Let $I = \langle x^2 \rangle \subset K[x]$. The zero set is the origin in $\mathbb{A}^1_K$ and the ideal of the zero set is $\langle x \rangle$.*

All rings R are commutative with an identity.

Definition-Lemma 1.13. *Let $I \trianglelefteq R$ be an ideal in a ring R. The **radical of** I, denoted $\sqrt{I}$, is the ideal of all elements of r such that $r^m \in I$ for some natural number m.*

*We say that I is a **radical ideal** if it is equal to its radical.*

Proof. $0 \in \sqrt{I}$ so that the radical is non-empty. If i and $j \in \sqrt{I}$ then we may find m and n such that $i^m, j^n \in I$. Let r and $s \in R$.

$$(ri + sj)^{m+n} = \sum_l \binom{n+m}{l} r^l i^l s^{n+m-l} j^{n+m-l} \in I,$$

so that $ri + sj \in \sqrt{I}$ and $\sqrt{I}$ is an ideal. $\square$

Theorem 1.14 (Hilbert's Nullstellensatz)**.** *Let K be an algebraically closed field.*

If $I \subset K[x_1, x_2, \dots, x_n]$ then

$$I(V(I)) = \sqrt{I}.$$

Corollary 1.15. *There is an inclusion reversing correspondence between Zariski closed subsets of $\mathbb{A}^n_K$ and radical ideals of $K[x_1, x_2, \dots, x_n]$.*

Definition 1.16. *Let X be a topological space. We say that X is* ***irreducible*** *if for every pair of closed subsets X_1 and X_2, such that $X_1 \cup X_2 = X$, we have either $X = X_1$ or $X = X_2$.*

Compare this definition with the definition of connected. Clearly the definition of irreducible is stronger than connected; in practice most connected topological spaces are rarely irreducible. For example if X is irreducible (and has at least two points) then it is not Hausdorff.

Lemma 1.17. *Let X be an irreducible topological space.*

Then every non-empty open subset is dense.

Proof. Let U be a non-empty open subset. If U is not dense then there is another non-empty open subset U such that $U \cap V$ is empty. Let F and G be their complements. Then F and G are two proper closed subsets, whose union is X, a contradiction. □

Lemma 1.18. *A Zariski closed subset $X \subset \mathbb{A}^n_K$ is irreducible if and only if $I(X)$ is a prime ideal.*

Proof. Suppose that X is irreducible. Let f_1 and f_2 be two elements of $K[x_1, x_2, \dots, x_n]$ such that $f_1 f_2 \in I$. Let $X_i = V(f_i)$. Then X_i are closed subsets of X and $X = X_1 \cup X_2$ since $f_1 f_2$ vanishes on X. As X is irreducible $X = X_i$ some i and so $f_i \in I$. But then I is prime.

Now suppose that I is prime. Let $X = X_1 \cup X_2$ and let $I_i = I(X_i)$. Then

$$I = I(X) = I(X_1 \cup X_2) = I_1 \cap I_2.$$

As I is prime, $I = I_i$ some i and $X = X_i$. □

Example 1.19. *A finite set is irreducible if and only if its cardinality is less than two.*

The nodal and cuspidal cubics are irreducible.

Definition 1.20. *Let X be a topological space. We say that X is* ***Noetherian*** *if the set of closed subsets satisfies DCC (the descending chain condition). That is, any sequence of descending closed subsets*

$$\dots \subset X_n \subset X_{n-1} \subset \dots \subset X_1 \subset X_0.$$

eventually stablises, so that we may find n_0 such that $X_n = X_{n+1}$ for all $n \geq n_0$.

Theorem 1.21. *Any affine variety is Noetherian.*

Proof. Let $X \subset \mathbb{A}^n$ be a closed subset. We may as well suppose that $X = \mathbb{A}^n$. By (1.11) it suffices to check that the set of radical ideals satisfies the ACC. But Hilbert's basis theorem says that the set of all ideals satisfies the ACC. □

Principle 1.22 (Noetherian Induction)**.** *Let P be a property of topological spaces, satisfying the following inductive hypothesis: if X is a topological space and every proper closed subset $Y \subset X$ satisfies property P, then X satisfies property P.*

Then every Noetherian topological space satisfies property P.

Proof. Suppose not. Let X be a Noetherian topological space, minimal with the property that it does not satisfy property P.

Let $Y \subset X$ be a proper closed subset. By minimality of X, Y satisfies property P. By the inductive hypothesis, X then satisfies property P, a contradiction. □

Lemma 1.23. *Let X be a Noetherian topological space.*

Then X has a decomposition into closed irreducible factors

$$X = X_1 \cup X_2 \cup \cdots \cup X_n,$$

where X_i is not contained in X_j, unique up to re-ordering of the factors.

Proof. If X is irreducible there is nothing to prove. Otherwise we may assume that $X = A \cup B$, where A and B are proper closed subsets. By the principle of Noetherian Induction, we may assume that A and B are the finite union of closed irreducible factors. Taking the union, and discarding any redundant factors (that is, any subset contained in another subset), we get the existence of such a decomposition.

Now suppose that

$$X_1 \cup X_2 \cup \cdots \cup X_m = Y_1 \cup Y_2 \cup \cdots \cup Y_n.$$

Consider

$$X_m = (X_m \cap Y_1) \cup (X_m \cap Y_2) \cup (X_m \cap Y_3) \cup \cdots \cup (X_m \cap Y_n).$$

By irreducibility of X_m, there is an index j such that $X_m \subset Y_j$. Thus $m \leq n$ and for every i there is a j such that $X_i \subset Y_j$. By symmetry, for every j there is a k such that $Y_j \subset X_k$. In this case $X_i \subset X_k$ and so $i = k$, by assumption. Thus $X_i = Y_j$. □

2. Nullstellensatz

We will need the notion of localisation, which is a straightforward generalisation of the notion of the field of fractions.

Definition 2.1. *Let R be a ring. We say that a subset S of R is* ***multiplicatively closed*** *if for every s_1 and s_2 in S, $s_1 s_2 \in S$, that is,*

$$S \cdot S \subset S.$$

Definition-Lemma 2.2. *Let R be a ring and let S be a multiplicatively closed subset, which contains 1 but not zero. The* ***localisation of*** *R* ***at*** *S, denoted R_S, is a ring R_S together with a ring homomorphism*

$$\phi \colon R \longrightarrow R_S,$$

with the property that for every $s \in S$, $\phi(s)$ is a unit in R_S, which is universal amongst all such rings. That is, given any morphism

$$\psi \colon R \longrightarrow T,$$

with the property that $\psi(s)$ is a unit, for every $s \in S$, there is a unique ring homomorphism

$$\begin{tikzcd} R \arrow[r, "\psi"] \arrow[d, "\phi"'] & T. \\ R_S \arrow[ur, dashed] & \end{tikzcd}$$

Proof. This is almost identical to the construction of the field of fractions, and so we will skip most of the details. Formally we define R_S to be the set of all pairs (r, s), where $r \in R$ and $s \in S$, modulo the equivalence relation,

$$(r_1, s_1) \sim (r_2, s_2) \qquad \text{iff} \qquad s(r_1 s_2 - r_2 s_1) \text{ for some } s \in S.$$

We denote an equivalence class by $[r, s]$ (or more informally by r/s). Addition and multiplication are defined in the obvious way. $\square$

Note that if R is an integral domain, then $S = R - \{0\}$ is multiplicatively closed and the localisation is precisely the field of fractions. Note also that as we are not assuming that R is an integral domain, we need to throw in the extra factor of s, in the definition of the equivalence relation and the natural map $R \longrightarrow R_S$ is not necessarily injective.

Example 2.3. *Suppose that $\mathfrak{p}$ is a prime ideal in a ring R. Then $S = R - \mathfrak{p}$ is a multiplicatively closed subset of R. The localisation is denoted $R_{\mathfrak{p}}$. It elements consist of all fractions r/f, where $f \notin \mathfrak{p}$. On* *the other hand, suppose that $f \in R$ is not nilpotent. Then the set of powers of f,*

$$S = \{ f^n \mid n \in \mathbb{N} \},$$

is a multiplicatively closed subset. The localisation consists of all elements of the form r/f^n.

For example, take $R = \mathbb{Z}$ and $f = 2$. Then $R_f = \mathbb{Z}[1/2] \subset \mathbb{Q}$ consists of all fractions whose denominator is a power of two.

Lemma 2.4. *Let F be a field and let $f \in F[x]$ be a polynomial.*

Then $F[x]_f$ is not a field.

Proof. Suppose not.

Clearly $\deg(f) > 0$ so that $1+f \neq 0$. Therefore we may find $g \in F[x]$ such that

$$(1+f)^{-1} = \frac{g}{f^n},$$

for some n. Multiplying out, we get that $(1+f)$ divides f^n.

So f^n is congruent to 0 modulo $(1+f)$. On the other hand, f is congruent to -1 modulo $(1+f)$. The only possibility is that $1+f$ is a unit, which is clearly impossible. □

Definition 2.5. *Let $R \subset F$ be a subring of the field F.*

*We say that $c \in F$ is **integral** over R if and only if there is a monic polynomial*

$$m(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in R[x],$$

such that $m(c) = 0$.

*If $R \subset S \subset F$ is an intermediary ring, we say that S is **integral over** R if every element of S is integral over R.*

*The **integral closure** of R in F is the set of all elements integral over R.*

Lemma 2.6. *Let $R \subset F$ be a subring of the field F.*

The following are equivalent:

(1) c is integral over R,
(2) $R[c]$ is a finitely generated R-module,
(3) there is an intermediary ring $R[c] \subset C \subset F$ which is a finitely generated R-module.

Proof. Suppose that c is integral over R. Pick $m(x) \in R[x]$ monic such that $m(c) = 0$. If $m(x)$ has degree d it is easy to see that 1, c, c^2, $\dots$, c^{d-1} generate $R[c]$ as an R-module. Thus (1) implies (2).

(2) implies (3) is clear.

Now suppose that C is a finitely generated R-module. Multiplication by c defines an R-linear map

$$\phi \colon C \longrightarrow C.$$

Pick generators $c_1, c_2, \dots, c_k$ for the R-module C. Then we may find $A = (a_{ij}) \in M_k(R)$ such that

$$\phi(c_i) = \sum a_{ij} c_j.$$

Then $m(x) = \det(A - \lambda I) \in R[x]$ is a monic polynomial and $m(\phi) = 0$, by Cayley-Hamilton. But then $m(c) = m(\phi(1)) = 0$. Hence (3) implies (1). □

Lemma 2.7. *Let $R \subset F$ be a subring of the field F.*

If $S = R[r_1, r_2, \dots, r_k]$ where each $r_1, r_2, \dots, r_k$ is integral over S then S is integral over R.

Proof. By (2.6) it suffices to prove that S is a finitely generated R-module. By induction on k we may assume that $S' = R[r_1, r_2, \dots, r_{k-1}]$ is a finitely generated R-module. As S is a finitely generated S'-module (r_k is integral over S' as it is integral over R) it follows that S is a finitely generated R-module. □

We will need the following result later:

Lemma 2.8. *Let $R \subset F$ be a subring of the field F.*

The integral closure S of R in F is a ring.

Proof. Let a and b be in S. It suffices to prove that $a \pm b$ and ab are in S. But $a \pm b$ and ab belong to $R[a, b]$ and this is finitely generated over R by (2.7). □

Lemma 2.9. *Let E be a field and let R be a subring.*

If E is integral over R then R is a field.

Proof. Pick $a \in R$ and let $b \in E$ be the inverse. As E is integral over R, we may find $r_1, r_2, \dots, r_n \in R$ such that

$$b^n + r_1 b^{n-1} + \cdots + r_n = 0.$$

Multiply both sides by a^{n-1} and solve for b to get

$$b = -r_1 - r_2 a - \cdots - r_n a^{n-1} \in A.$$ □

Lemma 2.10. *Let E/F be a field extension.*

If E is finitely generated as an F-algebra then E/F is algebraic.

Proof. By assumption $E = F[f_1, f_2, \dots, f_m]$. We proceed by induction on m.

Let $f = f_m$. By induction $E = F(f)[f_1, f_2, \dots, f_{m-1}]$ is algebraic over $F(f)$. Let $m_i(x) \in F(f)[x]$ be the minimal polynomial of f_i. Clearing denominators, we may assume that $m_i(x) \in F[f][x]$. Let a_i be the leading coefficient of $m_i(x)$ and let a be the product of the a_i. Then $(1/a_i)m_i(x) \in F[f]_a[x]$ is a monic polynomial, so that f_i is integral over $F[f]_a$.

By (2.9) $F[f]_a$ is a field. But then f is algebraic over F by (2.4). □

Theorem 2.11 (Weak Nullstellensatz)**.** *Let K be an algebraically closed field.*

Then an ideal $\mathfrak{m} \lhd R = K[x_1, x_2, \dots, x_n]$ is maximal if and only if it has the form

$$\mathfrak{m}_p = \langle x_1 - a_1, x_2 - a_2, \dots, x_n - a_n \rangle,$$

for some point $p = (a_1, a_2, \dots, a_n) \in K^n$.

Proof. Let $\mathfrak{m} \lhd R$ be an ideal and let $L = R/\mathfrak{m}$. Then $\mathfrak{m}$ is maximal if and only if $L = R/\mathfrak{m}$ is a field and $L = K$ if and only if $\mathfrak{m} = \mathfrak{m}_p$ for some point p.

So we may assume that L is a field and we want to prove that $L = K$. But L is a finitely generated algebra over K (generated by the images of $x_1, x_2, \dots, x_n$) so that by (2.10) L/K is algebraic. As K is algebraically closed, $L = K$. □

Corollary 2.12 (Weak Nullstellensatz)**.** *Let K be an algebraically closed field.*

If $f_1, f_2, \dots, f_m \in R = K[x_1, x_2, \dots, x_n]$ is a sequence of polynomials then either

(1) $f_1, f_2, \dots, f_m$ have a common zero, or
(2) there are polynomials $g_1, g_2, \dots, g_m \in K[x_1, x_2, \dots, x_n]$ such that

$$f_1 g_1 + f_2 g_2 + \cdots + f_m g_m = 1.$$

Proof. Let $I = \langle f_1, f_2, \dots, f_m \rangle \trianglelefteq R$ be the ideal generated by the polynomials $f_1, f_2, \dots, f_m$. Note that (1) holds if and only if I is contained in one of the ideals $\mathfrak{m}_p$ for some $p = (a_1, a_2, \dots, a_n) \in K^n$. Indeed, in this case $f_1, f_2, \dots, f_n$ all vanish at p. On the other hand, note that (2) holds if and only if $I = R$.

So suppose that $I \neq R$. Pick a maximal ideal $\mathfrak{m}$ containing I. By (2.11) we may find $p \in K^n$ such that $\mathfrak{m} = \mathfrak{m}_p$. □

Theorem 2.13 (Strong Nullstellensatz)**.** *Let K be an algebraically closed field.*

If $f_1, f_2, \dots, f_m, g \in R = K[x_1, x_2, \dots, x_n]$ is a sequence of polynomials then either

(1) $f_1, f_2, \dots, f_m$ have a common zero, at a point where the polynomial g is not equal to zero, or

(2) there are polynomials $g_1, g_2, \dots, g_m \in K[x_1, x_2, \dots, x_n]$ such that

$$f_1g_1 + f_2g_2 + \cdots + f_mg_m = g^r,$$

for some natural number r.

Proof. We use the *trick of Rabinowitsch*. Let

$$S = R[y] = K[x_1, x_2, \dots, x_n, y],$$

where y is an indeterminate and consider the polynomials

$$f_1, f_2, \dots, f_m, yg - 1.$$

If (1) does not hold then these equations don't have any solutions at all. By the weak Nullstellensatz (2.12) we may find polynomials $g_1, g_2, \dots, g_m, h \in S$ such that

$$f_1g_1 + f_2g_2 + \cdots + f_mg_m + h(yg - 1) = 1.$$

Let $z = 1/y$. Clearing denominators by multiplying through some large power z^r of z, and relabelling, we get

$$f_1g_1 + f_2g_2 + \cdots + f_mg_m + h(g - z) = z^r.$$

Now set $z = g$. □

Corollary 2.14 (Hilbert's Nullstellensatz)**.** *Let K be an algebraically closed field and let I be an ideal.*

Then $I(V(I)) = \sqrt{I}$.

Proof. One inclusion is clear, $I(V(I)) \supset \sqrt{I}$.

Now suppose that $g \in I(V(I))$. Pick a basis $f_1, f_2, \dots, f_k$ for I. Suppose that the point x is a common zero for $f_1, f_2, \dots, f_k$, so that $f_i(x) = 0$, for $1 \leq i \leq k$. Then $f(x) = 0$ for all $f \in I$ and so $x \in V(I)$. But then $g(x) = 0$. So we may apply the strong Nullstellensatz to $f_1, f_2, \dots, f_n, g$ to conclude that $g^r \in I$, some $r > 0$, that is, $g \in \sqrt{I}$. □

3. Categories and Functors

We recall the definition of a category:

Definition 3.1. *A category $\mathcal{C}$ is the data of two collections. The first collection is called the* ***objects*** *of $\mathcal{C}$ and is denoted $\mathrm{Obj}(\mathcal{C})$. Given two objects X and Y of $\mathcal{C}$, we associate another collection $\mathrm{Hom}(X, Y)$, called the* ***morphisms*** *between X and Y. Further we are given a law of* ***composition*** *for morphisms: given three objects X, Y and Z, there is an assignment*

$$\mathrm{Hom}(X, Y) \times \mathrm{Hom}(Y, Z) \longrightarrow \mathrm{Hom}(X, Z).$$

Given two morphisms, $f \in \mathrm{Hom}(X, Y)$ and $g \in \mathrm{Hom}(Y, Z)$, $g \circ f \in \mathrm{Hom}(X, Z)$ denotes the composition. Further this data satisfies the following axioms:

(1) Composition is associative,

$$h \circ (g \circ f) = (h \circ g) \circ f,$$

for all objects X, Y, Z, W and all morphisms $f\colon X \longrightarrow Y$, $g\colon Y \longrightarrow Z$ and $h\colon Z \longrightarrow W$.

(2) For every object X, there is a special morphism $i = i_X \in \mathrm{Hom}(X, X)$ which acts as an identity under composition. That is for all $f \in \mathrm{Hom}(X, Y)$,

$$f \circ i_X = f = i_Y \circ f.$$

We say that a category $\mathcal{C}$ is called ***locally small*** *if the collection of morphisms is a set. If in addition the collection of objects is a set, we say that $\mathcal{C}$ is* ***small****.*

There are an abundance of categories:

Example 3.2. *The category (Sets) of sets and functions; the category of (Groups) groups and group homomorphisms; the category (Vec) of vector spaces and linear maps; the category (Top) of topological spaces and continuous maps; the category (Rings) of rings and ring homomorphisms. All of these are locally small categories.*

Let X be a topological space. We can define a small category $\mathfrak{Top}X$ associated to X as follows. The objects of $\mathfrak{Top}X$ are simply the open subsets of X. Given two open subsets U and V,

$$\mathrm{Hom}(U, V) = \begin{cases} i_{UV} & \text{if } U \subset V \\ \emptyset & \text{otherwise.} \end{cases}$$

Here i_{UV} is a formal symbol. Composition of morphisms is defined in the obvious way (in fact the definition is forced, there are no choices to be made).

Definition 3.3. *We say that a category $\mathcal{D}$ is a subcategory of $\mathcal{C}$ if every object of $\mathcal{D}$ is an object of $\mathcal{C}$ and for every pair of objects X and Y of $\mathcal{D}$, $\mathrm{Hom}_{\mathcal{D}}(X, Y)$ is a subset of $\mathrm{Hom}_{\mathcal{C}}(X, Y)$ (that is, every morphism in $\mathcal{D}$ is a morphism in $\mathcal{C}$). The identity and composition of morphisms should come out the same.*

*We say that $\mathcal{D}$ is a **full subcategory** of $\mathcal{C}$, if for every pair of objects X and Y of $\mathcal{D}$, $\mathrm{Hom}_{\mathcal{D}}(X, Y)$ is equal to $\mathrm{Hom}_{\mathcal{C}}(X, Y)$.*

The category of finite sets is a full subcategory of the category $(\underline{\text{Sets}})$ of sets. Similarly the category of finite dimensional linear spaces is a full subcategory of the category $(\underline{\text{Vec}})$ of vector spaces. By comparison the category $(\underline{\text{Groups}})$ of groups is a subcategory of the category $(\underline{\text{Sets}})$ of sets (this example is a bit of a cheat) but it is not a full subcategory. In other words not every function is a group homomorphism.

It is easy construct new categories from old ones:

Definition 3.4. *Given a category $\mathcal{C}$, the **opposite category**, denoted $\mathcal{C}^{\mathrm{op}}$, is the category, whose objects are the same as $\mathcal{C}$, but whose morphisms go the other way, so that*

$$\mathrm{Hom}_{\mathcal{C}^{\mathrm{op}}}(X, Y) = \mathrm{Hom}_{\mathcal{C}}(Y, X).$$

Definition 3.5. *The **inverse** of a morphism $f\colon X \longrightarrow Y$ is a morphism $g\colon Y \longrightarrow X$, such that $f \circ g$ and $g \circ f$ are both the identity map. If the inverse of f exists, then we say that f is an **isomorphism** and that X and Y are **isomorphic**.*

Definition 3.6. *Let $\mathcal{C}$ and $\mathcal{D}$ be two categories. A **covariant functor** from F from $\mathcal{C}$ to $\mathcal{D}$ assigns to every object X of $\mathcal{C}$ an object $F(X)$ of $\mathcal{D}$ and to every morphism $f\colon X \longrightarrow Y$ in $\mathcal{C}$ a morphism $F(f)\colon F(X) \longrightarrow F(Y)$ in $\mathcal{C}$, compatible with composition and the identity.*

That is

$$F(g \circ f) = F(g) \circ F(f) \qquad \text{and} \qquad F(i_X) = i_{F(X)}.$$

*A **contravariant functor** F is the same as covariant functor, except that arrows are reversed,*

$$F(f)\colon F(Y) \longrightarrow F(X),$$

and

$$F(g \circ f) = F(f) \circ F(g).$$

In other words a contravariant functor $F\colon \mathcal{C} \longrightarrow \mathcal{D}$ is the same as a covariant functor $F\colon \mathcal{C}^{\mathrm{op}} \longrightarrow \mathcal{D}$

It is easy to give examples of functors. Let

$$F\colon (\underline{\text{Rings}}) \longrightarrow (\underline{\text{Groups}}),$$

be the functor which assigns to every ring R, the underlying additive group, and to every ring homomorphism f, the corresponding group homomorphism (the same map of course).

It is easy to check that F is indeed a functor; for obvious reasons it is called a forgetful functor and there are many such functors.

Note that we may compose functors in the obvious way and that there is an identity functor. Slightly more interestingly there is an obvious contravariant functor from a category to its opposite.

There are three non-trivial well-known functors. First there is a functor, denoted H_*, from the category $(\underline{\text{Top}})$ of topological spaces to the category of (graded) groups, which assigns to every topological space its singular homology. Similarly there is a contravariant functor from category $(\underline{\text{Top}})$ of topological spaces to the category of (graded) rings, which assigns to every topological space its singular cohomology.

The second and third are much more general.

Definition 3.7. *Let $F\colon \mathcal{C} \longrightarrow \mathcal{D}$ be a functor. We say that F is **faithful** if for every f and g, morphisms in $\mathcal{C}$, $F(f) = F(g)$ iff $f = g$. We say that F is **full** if for every morphism $h\colon F(X) \longrightarrow F(Y)$ in $\mathcal{D}$, there is a morphism f in $\mathcal{C}$ such that $F(f) = h$. We say that F is **essentially surjective** if for every object A in $\mathcal{D}$ there is an object X in $\mathcal{C}$ such that A is isomorphic to $F(X)$.*

*We say that F is an **equivalence of categories** if F is fully faithful and essentially surjective.*

For example, let $\mathcal{D}$ be the category of finite dimensional vector spaces over a field K. Let $\mathcal{C}$ be the category whose objects are the natural numbers, and such that the set of morphisms between two natural numbers m and n, is equal to the set of $m \times n$ matrices, with the obvious rule of composition. Then $\mathcal{C}$ is naturally a full subcategory of $\mathcal{D}$ (assign to n the "standard" vector space K^n) and the inclusion map, considered as a functor, is an equivalence of categories. Note however that there is no functor the other way.

More generally, given a category $\mathcal{D}$, one may form a quotient category $\mathcal{C}$. Informally the objects and morphisms of $\mathcal{C}$ are equivalence classes of objects of $\mathcal{D}$, under isomorphism.

We now turn to the third important functor. We first note that given two categories $\mathcal{C}$ and $\mathcal{D}$, where $\mathcal{C}$ is locally small, the collection of all functors from $\mathcal{C}$ to $\mathcal{D}$ is a category, denoted $\text{Fun}(\mathcal{C}, \mathcal{D})$. The objects of this category are simply functors from $\mathcal{C}$ to $\mathcal{D}$. Given two functors F and G, a morphism between them is a natural transformation:

Definition 3.8. *Let F and G be two functors from a category $\mathcal{C}$ to a category $\mathcal{D}$. A **natural transformation** u from F to G assigns to* *every object X of $\mathcal{C}$ a morphism $u_X\colon F(X) \longrightarrow G(X)$ such that for every morphism $f\colon X \longrightarrow Y$ in $\mathcal{C}$ the following diagram commutes*

$$\begin{array}{ccc} F(X) & \xrightarrow{F(f)} & F(Y) \\ {\scriptstyle u_X}\downarrow & & \downarrow{\scriptstyle u_Y} \\ G(X) & \xrightarrow{G(f)} & G(Y). \end{array}$$

It is easy to check that we may compose natural transformations, that this composition is associative and that the natural transformation which assigns to every object X, the identity map from $F(X)$ to $F(X)$ acts as an identity, so that $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ is indeed a category.

Suppose that $\mathcal{C}$ is a locally small category. Let Y be an object of $\mathcal{C}$. I claim that we get a functor $h_Y\colon \mathcal{C} \longrightarrow (\underline{\mathrm{Sets}})$. Given an object X of $\mathcal{C}$, we associate the set $h_Y(X) = \mathrm{Hom}(X, Y)$. Given a morphism $f\colon X \longrightarrow X'$, note that we get a map

$$h_Y(f) : \mathrm{Hom}(X', Y) \longrightarrow \mathrm{Hom}(X, Y),$$

which takes a morphism g and assigns the morphism $h_Y(f)(g) = g \circ f$. It is easy to check that h_Y is a contravariant functor. On the other hand, varying Y, I claim we get a functor

$$h\colon \mathcal{C} \longrightarrow \mathrm{Fun}(\mathcal{C}^{\mathrm{op}}, (\underline{\mathrm{Sets}})).$$

At the level of objects, the definition of this functor is obvious. Given $Y \in \mathcal{C}$ we assign the object $h_Y \in \mathrm{Fun}(\mathcal{C}^{\mathrm{op}}, (\underline{\mathrm{Sets}}))$. On the other hand, given a morphism $f\colon Y \longrightarrow Y'$, I claim that we get a natural transformation $h(f)$ between the two functors h_Y and $h_{Y'}$ going from $\mathcal{C}^{\mathrm{op}}$ to $(\underline{\mathrm{Sets}})$. Thus given an object X in $\mathcal{C}$, we are supposed to give a morphism

$$h(f)_X\colon h_Y(X) = \mathrm{Hom}(X, Y) \longrightarrow h_{Y'}(X) = \mathrm{Hom}(X, Y').$$

The definition of $h(f)_X$ is clear. Given $g \in \mathrm{Hom}(X, Y)$, send this to $h(f)_X(g) = f \circ g$. It is easy to check that $h(f)$ is indeed a natural transformation and that h is a functor. More significantly:

Theorem 3.9 (Yoneda's Lemma)**.** *h is fully faithful.*

The proof is left as an exercise for the reader. Yoneda's Lemma thus says that if we want to understand the category $\mathcal{C}$, we can think of it as a subcategory of the category of contravariant functors from $\mathcal{C}$ to the category $(\underline{\mathrm{Sets}})$ of sets.

In these terms obviously the must fundamental question is to ask which of these functors is in the image:

Definition 3.10. *We say that the functor $F\colon \mathcal{C}^{\mathrm{op}} \longrightarrow (\underline{Sets})$ is* ***representable*** *(by Y) if it is isomorphic to h_Y, for some object Y of $\mathcal{C}$.*

By Yoneda's Lemma, if F is representable by Y then Y is determined up to unique isomorphism.

4. Morphisms

We adopt the following working definition of a morphism between affine varieties.

Definition 4.1. *A* ***morphism***

$$f\colon V \longrightarrow W,$$

between two affine varieties, where $V \subset \mathbb{A}^m$ and $W \subset \mathbb{A}^n$ is given by picking a collection of n polynomials $f_1, f_2, \dots, f_n \in K[x_1, x_2, \dots, x_m]$ such that

$$f(x) = (f_1(x), f_2(x), \dots, f_n(x)) \in W,$$

for every $x \in V$.

Note that this gives us a category. The category of affine varieties, with maps given by morphisms.

Example 4.2. *The map*

$$f\colon \mathbb{A}^1 \longrightarrow \mathbb{A}^2,$$

given by

$$t \longrightarrow (t, t^2)$$

is a morphism.

It is interesting and instructive to look at the image. If we pick coordinates (x, y) on $\mathbb{A}^2$ then it is not too hard to check that the image is $C = V(y - x^2)$.

Now consider the morphism

$$C \longrightarrow \mathbb{A}^1 \qquad \text{given by} \qquad (x, y) \longrightarrow x.$$

This is a morphism and in fact it is pretty easy to see that the composition either way is the identity, so that it is the inverse of the first morphism and C is isomorphic to $\mathbb{A}^1$.

This example has many interesting generalisations. For example we can look at the morphism:

Example 4.3.

$$\mathbb{A}^1 \longrightarrow \mathbb{A}^3,$$

given as

$$t \longrightarrow (t, t^2, t^3)$$

The image is $C = V(y - x^2, z - x^3)$ and projection down to the x-axis is the inverse map.

It turns out that both the nodal and cuspidal cubic can be parametrised.

Example 4.4. *Consider the map*

$$f\colon \mathbb{A}^1 \longrightarrow \mathbb{A}^2,$$

given by

$$t \longrightarrow (t^2-1, t(t^2-1))$$

The image lies in the nodal cubic:

$$C = V(y^2 - x^2 - x^3) \subset \mathbb{A}^2,$$

since

$$x^2 + x^3 = x^2(x+1) = (t^2-1)^2 t^2 = t(t^2-1)^2 = y^2.$$

But in fact, if $y^2 = x^2 + x^3$ then either $x = 0$ in which case $y = 0$ and this is the image of $t = \pm 1$ or we put $t = y/x$, in which case it is easy to check that $f(t) = (x, y)$.

Example 4.5. *Consider the map*

$$g\colon \mathbb{A}^1 \longrightarrow \mathbb{A}^2,$$

given by

$$t \longrightarrow (t^2, t^3)$$

The image lies in the cuspidal cubic:

$$C = V(y^2 - x^3) \subset \mathbb{A}^2.$$

But in fact, if $y^2 = x^3$ then either $x = 0$ in which case $y = 0$ and this is the image of $t = 0$ or we put $t = y/x$, in which case it is easy to check that $f(t) = (x, y)$.

In all of these examples, note that the image is a closed subset. However

Example 4.6. *Let $C = V(xy - 1) \subset \mathbb{A}^2$ and consider the morphism*

$$C \longrightarrow \mathbb{A}^1 \qquad \text{given by} \qquad (x, y) \longrightarrow x.$$

The image is the set of all non-zero elements of $\mathbb{A}^1$. This is an open subset not a closed subset.

Definition 4.7. *If $U \subset V \subset \mathbb{A}^n$ is an open subset of the affine subvariety V then we say that U is **quasi-affine**.*

Remark 4.8. *Note that an open subset of a closed subset is the same as a locally closed subset (that is, U is the intersection of an open subset and a closed subset).*

It seems natural to wonder if quasi-affine varieties are closed under taking the image.

Example 4.9. *Consider the morphism*

$$f\colon \mathbb{A}^2 \longrightarrow \mathbb{A}^2 \qquad \text{given by} \qquad (x,y) \longrightarrow (xy,y).$$

Let $(a,b) \in \mathbb{A}^2$ be a point in the target $\mathbb{A}^2$. If $b \neq 0$ then $(x,y) = (a/b,b)$ maps to (a,b). If $f(x,y) = (a,0)$ then $y = 0$ so that $a = xy = 0$. Thus the image consists of the open subset where $y \neq 0$ union the origin. It is not hard to see that this is not quasi-affine, that is this set is not the intersection of an open and a closed set.

How should one define morphisms for quasi-affine varieties? Well, the natural thing to do is to restrict morphisms from the closure, which is an affine variety. But consider the hyperbola, $C = V(xy-1)$. The map $(x,y) \longrightarrow x$ is equivalent to the map $(x,y) \longrightarrow 1/y$, since $x = 1/y$ on C. So it is more natural to define a morphism as something which is locally the restriction of a collection of rational functions.

Definition 4.10. *Let*

$$f\colon X \longrightarrow Y,$$

be a map between two quasi-affine varieties $X \subset \mathbb{A}^m$ and $Y \subset \mathbb{A}^n$. We say that f is a morphism, if there is an open affine cover U_i for X such that for every i the restriction $f_i = f|_{U_i}$

$$f_i\colon U_i \longrightarrow Y,$$

is given by a collection of rational functions, g_j/h_j,

$$(x_1, x_2, \dots, x_m) \longrightarrow (g_1/h_1, g_2/h_2, \dots g_n/h_n)$$

where none of the h_j are zero on U_i.

Definition-Lemma 4.11. *Let $V \subset \mathbb{A}^n$ be an affine variety and let f be a polynomial. Let*

$$U_f = \{\, x \in V \mid f(x) \neq 0 \,\}.$$

Then the quasi-affine variety U_f is isomorphic to an affine variety.

Proof. Let $W \subset \mathbb{A}^{n+1}$ be the closed subset defined by the ideal of V together with the polynomial $yf - 1$. Define a morphism

$$\phi\colon W \longrightarrow V,$$

given by

$$(x_1, x_2, \dots, x_n, y) \longrightarrow (x_1, x_2, \dots, x_n).$$

The image is contained in U_f, since if we can find y such that $yf(x) = 1$ then $f(x) \neq 0$.

Define a morphism

$$U_f \longrightarrow W,$$

given by

$$\psi\colon (x_1, x_2, \dots, x_n) \longrightarrow (x_1, x_2, \dots, x_n, 1/f).$$

Then it is not hard to see that ϕ and ψ are inverses of each other. $\square$

Definition-Lemma 4.12. *Let X be a quasi-affine variety. A* ***regular function*** *is a morphism $X \longrightarrow \mathbb{A}^1 = K$.*

The set of all regular functions $\mathcal{O}_X(X)$ is a K-algebra.

Proof. Clear. $\square$

It seems worthwhile to point out that at this point one can state the very famous:

Conjecture 4.13 (Jacobian conjecture)**.** *Let*

$$\phi\colon \mathbb{A}^2 \longrightarrow \mathbb{A}^2,$$

be a morphism, given by

$$(x, y) \longrightarrow (f, g).$$

ϕ is an isomorphism if and only if the Jacobian

$$J(\phi) = \begin{vmatrix} f_x & f_y \\ g_x & g_y \end{vmatrix}$$

is a non-zero constant.

Remark 4.14. *It is easy to show that if ϕ is an isomorphism then the Jacobian $J(\phi)$ is a non-zero constant.*

5. Coordinate rings

Recall the following version of the Nullstellensatz:

Theorem 5.1 (Weak Nullstellensatz)**.** *Let K be an algebraically closed field.*

Then an ideal $\mathfrak{m} \lhd R = K[x_1, x_2, \dots, x_n]$ is maximal if and only if it has the form

$$\mathfrak{m}_p = \langle x_1 - a_1, x_2 - a_2, \dots, x_n - a_n \rangle,$$

for some point $p = (a_1, a_2, \dots, a_n) \in K^n$.

Note that with this formulation it is clear why we need K to be algebraically closed. Indeed $I = \langle x^2 + 1 \rangle$ over $\mathbb{R}$ is in fact maximal and the vanishing locus is empty.

Another way to restate the Nullstellensatz is to observe that it establishes an inclusion reversing correspondence between ideals and closed subsets of $\mathbb{A}^n$. However this is just the tip of the iceberg.

Definition 5.2. *Let $X \subset \mathbb{A}^n$ be a closed subset.*

The ***coordinate ring of*** *X, denoted $A(X)$, is the quotient*

$$K[X]/I(X).$$

Corollary 5.3. *Let $X \subset \mathbb{A}^n$ be an affine subvariety.*

There is a correspondence between the points of X and the maximal ideals of the coordinate ring $A(X)$.

Proof. Recall that there is a correspondence between ideals in $R = K[x_1, x_2, \dots, x_n]$ containing I and ideals in the quotient R/I. So there is a correspondence between maximal ideals of R/I and maximal ideals of R containing I.

But an ideal

$$\mathfrak{m}_p = \langle x_1 - a_1, x_2 - a_2, \dots, x_n - a_n \rangle,$$

contains I if and only if $p \in X$ and so we are done by (5.1). $\square$

In fact this correspondence is natural. To prove this, we have to reinterpret the coordinate ring.

Proposition 5.4. *If $X \subset \mathbb{A}^n$ is an affine variety then the ring of regular functions $\mathcal{O}_X(X)$ is isomorphic to the coordinate ring.*

Proof. Let $\pi \colon K[X] \longrightarrow \mathcal{O}_X(X)$ be the map which sends a polynomial f to the obvious regular function ϕ, $\phi(x) = f(x)$. It is clear that π is a ring homomorphism, with kernel $I(X)$. It suffices, then, to prove that π is surjective.

Let ϕ be a regular function on X. By definition there is an open cover U_i of X and rational functions f_i/g_i such that ϕ is locally given by f_i/g_i. As X is Noetherian, we may assume that each U_i is irreducible. We may assume that $U_i = U_{h_i}$ for some regular function h_i, as such subsets form a base for the topology. Replacing f_i by $f_i h_i$ and g_i by $g_i h_i$ we may assume that f_i and g_i vanish outside of U_i. There are two cases; $U_i \cap U_j$ is non-empty or empty.

Suppose that $U_i \cap U_j$ is non-empty. As U_i is irreducible it follows that $U_i \cap U_j$ is a dense subset of U_i. Now $f_i/g_i = f_j/g_j$ as functions on $U_i \cap U_j$ and so $f_i g_j = f_j g_i$ as functions on $U_i \cap U_j$. As these functions are continuous, $f_i g_j = f_j g_i$ on U_i. Suppose that $U_i \cap U_j$ is empty. Then the identity $f_i g_j = f_j g_i$ holds on U_i as both sides are zero.

By assumption, the common zero locus of $\{g_i\}$ is empty. Thus, by the Nullstellensatz, there are polynomials $h_1, h_2, \dots, h_n$ such that

$$1 = \sum_i g_i h_i.$$

Set $f = \sum_i f_i h_i$. I claim that the function

$$x \longrightarrow f(x),$$

is the regular function ϕ. It is enough to check this on U_j, for every j. We have

$$\begin{aligned} f g_j &= \left(\sum_i f_i g_j\right) h_i \\ &= \sum_i (f_i g_j) h_i \\ &= \sum_i (f_j g_i) h_i \\ &= f_j \sum_i g_i h_i = f_j. \end{aligned}$$

□

Note that this result implies that the working definition of a morphism between affine varieties is correct. Indeed, simply projecting onto the jth factor, it is clear that if the map is given as

$$(x_1, x_2, \dots, x_m) \longrightarrow (f_1(x), f_2(x), \dots, f_n(x)),$$

then each $f_j(x)$ is a regular function. By (5.4), it follows that $f_j(x)$ is given by a polynomial.

Lemma 5.5. *There is a contravariant functor A from the category of affine varieties over K to the category of commutative rings. Given an affine variety X we associate the ring $\mathcal{O}_X(X)$. Given a morphism*

$f\colon X \longrightarrow Y$ of affine varieties, $A(f)\colon \mathcal{O}_Y(Y) \longrightarrow \mathcal{O}_X(X)$, which sends a regular function ϕ to the regular function $A(f)(\phi) = \phi \circ f$.

It is interesting to describe the image of this functor. Clearly the ring $A(X)$ is an algebra over K (which is to say that it contains K, so that we can multiply by elements of K). Further the ring $A(X)$ is a quotient of the polynomial ring, so that it is a finitely generated algebra over K. Also since the ideal $I(X)$ is radical, the ring $A(X)$ does not have any nilpotents.

Definition 5.6. *Let R be a ring. A non-zero element r of R is said to be **nilpotent** if there is a positive integer n such that $r^n = 0$.*

Clearly if a ring has a nilpotent element, then it is not an integral domain.

Theorem 5.7. *The functor A is an equivalence of categories between the category of affine varieties over K and the category of finitely generated algebras over K, without nilpotents.*

Proof. First we show that A is essentially surjective. Suppose we are given a finitely generated algebra A over K. Pick generators $\xi_1, \xi_2, \dots, \xi_n$ of A. Define a ring homomorphism

$$\pi\colon K[x_1, x_2, \dots, x_n] \longrightarrow A,$$

simply by sending x_i to ξ_i. It is easy to check that π is an algebra homomorphism. Let I be the kernel of π. Then I is radical, as A has no nilpotents. Let $X = V(I)$. Then the coordinate ring of X is isomorphic to A, by construction. Thus A is essentially surjective.

To prove the rest, it suffices to prove that if X and Y are two affine varieties then A defines a bijection between

$$\mathrm{Hom}(X, Y) \qquad \text{and} \qquad \mathrm{Hom}(\mathcal{O}_Y(Y), \mathcal{O}_X(X)).$$

To prove this, we may as well fix embeddings $X \subset \mathbb{A}^m$ and $Y \subset \mathbb{A}^n$. In this case A naturally defines a map between

$$\mathrm{Hom}(X, Y) \qquad \text{and} \qquad \mathrm{Hom}(A(Y), A(X)),$$

which we continue to refer to as A. It suffices to prove that there is a map

$$B\colon \mathrm{Hom}(A(Y), A(X)) \longrightarrow \mathrm{Hom}(X, Y),$$

which is inverse to the map

$$A\colon \mathrm{Hom}(X, Y) \longrightarrow \mathrm{Hom}(A(Y), A(X)).$$

Suppose we are given a ring homomorphism $\alpha\colon A(Y) \longrightarrow A(X)$. Define a map

$$B(\alpha)\colon X \longrightarrow Y,$$

as follows. Let $y_1, y_2, \dots, y_n$ be coordinates on $Y \subset \mathbb{A}^n$. Let $f_1, f_2, \dots, f_n$ be the polynomials on $\mathbb{A}^n$, defined by $\alpha(y_i) = f_i$. Then define $B(\alpha)$ by the rule

$$(x_1, x_2, \dots, x_m) \longrightarrow (f_1, f_2, \dots, f_n).$$

Clearly this is a morphism. We check that the image lies in Y. Suppose that $p \in X$. We check that $q = (f_1(p), f_2(p), \dots, f_n(p)) \in Y$. Pick $g \in I(Y)$. Then

$$\begin{aligned} g(q) &= g(f_1(p), f_2(p), \dots f_n(p)) \\ &= g(\alpha(y_1)(p), \alpha(y_2)(p), \dots, \alpha(y_n)(p)) \\ &= \alpha(g)(p) \\ &= 0. \end{aligned}$$

Thus $q \in Y$ and we have defined the map B.

We now check that B is the inverse of A. Suppose that we are given a morphism $f\colon X \longrightarrow Y$. Let $\alpha = A(f)$. Suppose that f is given by $(f_1, f_2, \dots, f_n)$. Then $\alpha(y_i) = y_i \circ f = f_i$. It follows easily that $B(\alpha) = f$. Now suppose that $\alpha\colon A(Y) \longrightarrow A(X)$ is an algebra homomorphism. Then $B(\alpha)$ is given by $(f_1, f_2, \dots, f_n)$ where $f_i = \alpha(y_i)$. In this case $A(f)(y_i) = f_i$. As $y_1, y_2, \dots, y_n$ are generators of $A(Y)$, we have $\alpha = A(B(\alpha))$. $\square$

(5.7) raises an interesting question. Can we enlarge the category of affine varieties so that we get every finitely generated algebra over K and not just those without nilpotents. In fact, why stop there? Can we find a class of geometric objects, such that the space of functions on these objects, gives us any ring whatsoever (not nec. finitely generated, not nec. over K). Amazingly the answer is yes, but to do this we need the theory of schemes.

6. Sheaves

Definition 6.1. *Let X be a topological space. A **presheaf of groups** $\mathcal{F}$ on X is a function which assigns to every open set $U \subset X$ a group $\mathcal{F}(U)$ and to every inclusion $V \subset U$ a restriction map,*

$$\rho_{UV} \colon \mathcal{F}(U) \longrightarrow \mathcal{F}(V),$$

which is a group homomorphism, such that if $W \subset V \subset U$, then

$$\rho_{VW} \circ \rho_{UV} = \rho_{UW}.$$

Succintly put, a pre-sheaf is a contravariant functor from $\mathfrak{Top}(X)$ to the category $\underline{\text{(Groups)}}$ of groups. Put this way, it is clear what we mean by a presheaf of rings, etc. The elements of $\mathcal{F}(U)$ are called *sections*. We almost always denote $\rho_{UV}(s) = s|_V$. U_{ij} denotes $U_i \cap U_j$.

Example 6.2. *Let X be a topological space and let G be a group. Define a presheaf $\mathcal{G}$ as follows. Let U be any open subset of X. $\mathcal{G}(U)$ is defined to be the set of constant functions from X to G. The restriction maps are the obvious ones.*

Definition 6.3. *A **sheaf** $\mathcal{F}$ on a topological space is a presheaf which satisfies the following two axioms:*

1. *Given an open cover U_i of U an open subset of X, and a collection of sections s_i on U_i, such that $s_i|_{U_{ij}} = s_j|_{U_{ij}}$ then there is a section s on U such that $s|_{U_i} = s_i$.*
2. *Given an open cover U_i of U an open subset of X, if s is a section on U such that $s|_{U_i} = 0$, then s is zero.*

Note that we could easily combine (1) and (2) and require that there is a unique s, which is patched together from the s_i. It is very easy to give lots of examples of sheaves and presheaves. Basically, any collection of functions is a sheaf.

Example 6.4. *Let M be a complex manifold. Then there are a collection of sheaves on M. The sheaf of holomorphic functions, the sheaf of C^∞-functions and the sheaf of continuous functions. In all cases, the restrictions maps are the obvious ones, and there are obvious inclusions of sheaves.*

If X is a quasi-affine variety, then let $\mathcal{O}_X$ be the sheaf of regular functions. This is a sheaf of rings.

Note however that in general the presheaf defined in (6.2) *is not a sheaf. For example, take $X = \{a, b\}$ to be the topological space with the discrete topology and take $G = \mathbb{Z}_2$. Let $U_1 = \{a\}$ and $U_2 = \{b\}$ and suppose $s_1 = 0$ and $s_2 = 1$. Then there is no global constant function which restricts to both* 0 *and* 1.

However this is easily fixed. Take $\mathcal{F}$ to be the sheaf of locally constant functions.

Definition 6.5. *Let X be a topological space and let $\mathcal{F}$ be a presheaf on X. Given $p \in X$, the collection of open sets which contain p forms a directed set I. The **stalk of $\mathcal{F}$ at** p, denoted $\mathcal{F}_p$, is the direct limit over I, $\lim_{p\in U} \mathcal{F}(U)$.*

It is useful to untwist this definition. An element of the stalk is a pair (s, U), such that $s \in \mathcal{F}(U)$, modulo the equivalence relation,

$$(s, U) \sim (t, V)$$

if there is an open subset $W \subset U \cap V$ such that

$$s|_W = t|_W.$$

In other words, we only care about what s looks like in an arbitrarily small neighbourhood of p.

Definition 6.6. *A ring R is called a **local ring** if there is a unique maximal ideal.*

Note that when we have a sheaf of rings, the stalk is often a local ring.

Example 6.7. *Let M be a complex manifold of dimension n and let p be a point of M. Then*

$$\mathcal{O}^h_{M,p} \simeq \mathbb{C}\{z_1, z_2, \dots, z_n\},$$

the ring of convergent power series, since locally about p, M looks like $\mathbb{C}^n$ about zero, and any holomorphic function is determined by its Taylor series. On the other hand if M is a real manifold of dimension n there is a ring surjective homomorphism

$$\mathcal{C}^\infty_{M,p} \longrightarrow \mathbb{R}[[x_1, x_2, \dots, x_n]],$$

the ring of formal power series, but the kernel is simply huge. In other words, there are lots of infinitely differentiable functions with a trivial Taylor series.

Definition-Lemma 6.8. *Let X be an affine variety and let $p \in X$.*

Then the stalk of the structure sheaf of X at p, $\mathcal{O}_{X,p}$ is equal to the localisation of $A(X)$ at the maximal ideal m_p of $p \in X$.

Proof. There is an obvious ring homomorphism

$$A(X) \longrightarrow \mathcal{O}_{X,p},$$

which just sends a polynomial f to the equivalence class (f, X). Suppose that $f \notin m$. Then $p \in U_f \subset X$ and $(1/f, U_f)$ represents the inverse of (f, X) in the ring $\mathcal{O}_{X,p}$. By the universal property of the localisation there is a ring homomorphism

$$A(X)_m \longrightarrow \mathcal{O}_{X,p},$$

which is clearly injective. Now suppose that we have an element (σ, U) of $\mathcal{O}_{X,p}$. Since sets of the form U_f form a basis for the topology, we may assume that $U = U_g$. But then $\sigma = f/g^n \in A(X)_g \subset A(X)_m$, for some f and n. $\square$

Definition 6.9. *A map between presheaves is a natural transformation of the corresponding functors.*

Untwisting the definition, a map between presheaves

$$f\colon \mathcal{F} \longrightarrow \mathcal{G}$$

assigns to every open set U a group homomorphism

$$f(U)\colon \mathcal{F}(U) \longrightarrow \mathcal{G}(U),$$

such that the following diagram always commutes

$$\begin{tikzcd}
\mathcal{F}(U) \arrow[r, "f(U)"] \arrow[d, "\rho_{UV}"'] & \mathcal{G}(U) \arrow[d, "\sigma_{UV}"'] \\
\mathcal{F}(V) \arrow[r, "f(V)"] & \mathcal{G}(V).
\end{tikzcd}$$

Note that this gives us a category of presheaves, together with a full subcategory of sheaves.

Definition-Lemma 6.10. *Let $\mathcal{F}$ be a presheaf.*

*Then the **sheaf associated to the presheaf**, is a sheaf $\mathcal{F}^+$, together with a morphism of sheaves $u\colon \mathcal{F} \longrightarrow \mathcal{F}^+$ which is universal amongst all such morphisms of sheaves: that is, given any morphism of presheaves*

$$f\colon \mathcal{F} \longrightarrow \mathcal{G},$$

where $\mathcal{G}$ is a sheaf, there is a unique induced morphism of sheaves which makes the following diagram commute

$$\begin{tikzcd}
\mathcal{F} \arrow[r, "f"] \arrow[d, "u"'] & \mathcal{G}. \\
\mathcal{F}^+ \arrow[ur, dashed] &
\end{tikzcd}$$

Proof. We just give the construction of $\mathcal{F}^+$ and leave the details to the reader. Let H be the disjoint union of all the stalks of $\mathcal{F}$. Let $U \subset X$ be an open set. A section s of $\mathcal{F}^+$ is by definition a function $U \longrightarrow H$ which sends a point p to an element of H, that is, a germ

$s(p) = s_p \in \mathcal{F}_p$, which is locally given by sections of $\mathcal{F}$. That is, for every $q \in U$, we require that there is an open subset $V \subset U$ containing q, and a section $t \in \mathcal{F}(V)$ such that (t, V) represents s_p in $\mathcal{F}_p$ for all $p \in V$. $\square$

For example the sheaf associated to the presheaf of constant functions to G, is the sheaf of locally constant functions to G.

Proposition 6.11. *Let $\phi\colon \mathcal{F} \longrightarrow \mathcal{G}$ be a morphism of sheaves.*

Then ϕ is an isomorphism if and only if the induced map on stalks is always an isomorphism.

Proof. One direction is clear. So suppose that the map on stalks is an isomorphism. It suffices to prove that $\phi(U)\colon \mathcal{F}(U) \longrightarrow \mathcal{G}(U)$ is an isomorphism, for every open subset $U \subset X$, since then the inverse morphism ϕ is given by setting $\psi(U) = \phi(U)^{-1}$.

We first show that $\phi(U)$ is injective. Let $s \in \mathcal{F}(U)$ and suppose that $\phi(U)(s) = 0$. Then surely $\phi_p(s_p) = 0$, where $s_p = (s, U) \in \mathcal{F}_p$ and $p \in U$ is arbitrary. Since ϕ_p is injective by assumption, it follows that there is an open set $V_p \subset U$ containing p such that $s|_{V_p} = 0$. $\{V_p\}_{p \in U}$ is an open cover of U and as $\mathcal{F}$ is a sheaf, it follows that $s = 0$. Hence $\phi(U)$ is injective, for every U.

Now we show that $\phi(U)$ is surjective. Suppose that $t \in \mathcal{F}(U)$. Since ϕ_p is surjective, for every p, we may find an open set $p \in U_p \subset U$ and a section $s_p \in \mathcal{F}(U_p)$ such that $\phi(U_p)(s_p) = t|_{U_p}$. Pick p and $q \in U$ and set $V = U_p \cap U_q$. Then $\phi(V)(s_p|_V) = \phi(V)(s_q|_V)$. Since $\phi(V)$ is injective, it follows that $s_p|_V = s_q|_V$. As $\mathcal{F}$ is a sheaf, it follows that there is a section $s \in \mathcal{F}(U)$ such that $s|_{U_P} = s$. But then $\phi(U)(s) = t$ and so $\phi(U)$ is surjective. $\square$

Example 6.12. *Let $X = \mathbb{C} - \{0\}$, let $\mathcal{F} = \mathcal{O}_X$, the sheaf of holomorphic functions and let $\mathcal{G} = \mathcal{O}_X^*$, the sheaf of non-zero holomorphic functions.*

There is a natural map

$$\phi\colon \mathcal{F} \longrightarrow \mathcal{G},$$

which just sends a function f to its exponential. Then ϕ is surjective on stalks; this just says that given a non-zero holmorphic function g, then $\log(g)$ makes sense in a small neighbourhood of any point.

On the other hand $\phi(X)$ is not surjective. Indeed $z \in \mathcal{G}(X)$ is a function which is not in the image, since $\log(z)$ is not globally single valued.

Definition 6.13. *Let $f\colon X \longrightarrow Y$ be a continuous map of topological spaces. Let $\mathcal{F}$ be a sheaf on X. The* ***pushforward of $\mathcal{F}$****, denoted*

$f_*\mathcal{F}$, is defined as follows

$$f_*\mathcal{F}(U) = \mathcal{F}(f^{-1}(U)),$$

where $U \subset Y$ is an open set.

*Let $\mathcal{G}$ be a sheaf on Y. The **inverse image of** $\mathcal{G}$, denoted $f^{-1}\mathcal{G}$, is the sheaf assigned to the presheaf*

$$U \longrightarrow \lim_{f(U)\subset V} \mathcal{G}(V),$$

where U is an open subset of X and V ranges over all open subets of Y which contain $f(U)$.

Definition 6.14. *A pair $(X, \mathcal{O}_X)$ is called a **ringed space**, if X is a topological space, and $\mathcal{O}_X$ is a sheaf of rings. A morphism $\phi\colon X \longrightarrow Y$ of ringed spaces is a pair $(f, f^{\#})$, consisting of a continuous function $f\colon X \longrightarrow Y$ and a sheaf morphism $f^{\#}\colon \mathcal{O}_Y \longrightarrow f_*\mathcal{O}_X$.*

*A **locally ringed space**, is a ringed space $(X, \mathcal{O}_X)$ such that in addition every stalk $\mathcal{O}_{X,x}$ of the structure sheaf is a local ring. A morphism of locally ringed spaces is a morphism of ringed spaces, such that for every point $x \in X$, the induced map*

$$f^{\#}_x\colon \mathcal{O}_{Y,y} \longrightarrow \mathcal{O}_{X,x},$$

where $y = f(x)$ is a morphism of local rings (that is, the inverse image of the maximal ideal of $\mathcal{O}_{Y,y}$ is the maximal ideal of $\mathcal{O}_{X,x}$).

Note that we get a category of ringed spaces, whose objects are ringed spaces and whose morphisms are morphisms of ringed spaces. Further the category of locally ringed spaces is a subcategory, not necessarily a full subcategory.

Definition 6.15. *Let $(X, \mathcal{O}_X)$ be a ringed space. An $\mathcal{O}_X$**-module** is a sheaf $\mathcal{F}$ such that for every open set $U \subset X$, $\mathcal{F}(U)$ has the structure of an $\mathcal{O}_X(U)$-module, compatible with the restriction map, in an obvious way.*

Using (6.10) we may define various natural operations on sheaves. For example, let $\mathcal{F}$ and $\mathcal{G}$ be two $\mathcal{O}_X$-modules. The tensor product of $\mathcal{F}$ and $\mathcal{G}$, denoted $\mathcal{F} \underset{\mathcal{O}_X}{\otimes} \mathcal{G}$, is the sheaf associated to the presheaf

$$U \longrightarrow \mathcal{F}(U) \underset{\mathcal{O}_X(U)}{\otimes} \mathcal{G}(U),$$

and curly hom, denoted $\mathbf{Hom}_{\mathcal{O}_X}(\mathcal{F}, \mathcal{G})$, is the sheaf associated to the presheaf

$$U \longrightarrow \mathrm{Hom}_{\mathcal{O}_U = \mathcal{O}_X|_U}(\mathcal{F}|_U, \mathcal{G}|_U).$$

Let $f\colon \mathcal{F} \longrightarrow \mathcal{G}$ be a morphism of sheaves. The kernel of f is the sheaf which assigns to every open set U the kernel of the homomorphism $f(U)\colon \mathcal{F}(U) \longrightarrow \mathcal{G}(U)$. Similarly the image is the sheaf associated to the presheaf which assigns to every open set U the image of the homomorphism $f(U)\colon \mathcal{F}(U) \longrightarrow \mathcal{G}(U)$. We say that ϕ is injective if and only if $\operatorname{Ker}(\phi) = 0$ and we say that ϕ is surjective if and only if $\operatorname{Im}(\phi) = \mathcal{G}$.

Given a morphism of ringed spaces $\phi\colon X \longrightarrow Y$, and a sheaf $\mathcal{G}$ of $\mathcal{O}_Y$-modules, the pullback of $\mathcal{G}$, denoted $\phi^*\mathcal{G}$, is the sheaf of $\mathcal{O}_X$-modules,

$$\phi^{-1}\mathcal{G} \underset{f^{-1}\mathcal{O}_Y}{\otimes} \mathcal{O}_X.$$

7. Affine schemes I

Schemes were introduced by Grothendieck more than fifty years ago into the world of algebraic geometry. It might help to quickly review the reasons why schemes were introduced in the first place. Then in the course of these lectures we will see how the theory of schemes deals with the limitations of working with varieties.

Geometrically there are two compelling reasons to work with more general objects than varieties. Firstly, it is desirable to have a definition of an affine variety which is independent of any embedding into affine space. Compare this with the definition of a group. Originally groups were thought of as subsets of the set of permutations of a set, which are closed under composition and inverses. It is clearly much better to have the abstract definition of a group and then consider all the possible ways of embedding the group into permutation groups. This way one can think about groups being isomorphic, without worrying about a particular embedding. Similarly one of the big conceptual advances of the twentieth century was a definition of an abstract manifold.

Secondly if one looks at even the simplest families of varieties, some fibres (or members of the family) are not really varieties. For example, consider $S = V(xy - t) \subset \mathbb{A}^3$, a surface in $\mathbb{A}^3$. Projection down to the t-axis,

$$\pi \colon S \longrightarrow \mathbb{A}^1,$$

realises this surface as a family of curves in $\mathbb{A}^2$. If $a \neq 0$ then $C = V(xy - a) \subset \mathbb{A}^2$ is a hyperbola but if $a = 0$ then $C = V(xy) \subset \mathbb{A}^2$ is a pair of lines. This example is okay (but only because we allow reducible varieties). Now consider the $S = V(x^2 - ty) \subset \mathbb{A}^3$, a surface in $\mathbb{A}^3$. Projection down to the t-axis,

$$\pi \colon S \longrightarrow \mathbb{A}^1,$$

realises this surface as a family of curves in $\mathbb{A}^2$. If $a \neq 0$ then $C = V(x^2 - ay) \subset \mathbb{A}^2$ is a parabola but if $a = 0$ then $C = V(x^2) \subset \mathbb{A}^2$ is a line. But something is wrong here; a line is not really a conic, it is defined by a linear polynomial not a quadratic polynomial. We want something geometric corresponding to a doubled line.

Moreover there are other equally compelling reasons to enlarge the category of varieties, coming from other areas of mathematics. Suppose that we want to understand the polynomial equation

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

In terms of arithmetic, we are interested in those 3-tuples (x, y, z), where x, y and $z \in \mathbb{Z}$. It is well known that determining the integral solutions is very hard, and it is natural to attack such problems by considering what happens over $\mathbb{C}$ and also what happens when we reduce modulo p, which are both considerably easier and shed light on what happens over the integers. In these terms, it seems that we have a single object X (determined by the equation) and we seek to understand X, by computing what happens when we look at the set

$$X(R) = \{ (x, y, z) \in R^3 \mid x^n + y^n = z^n \},$$

where R is a commutative ring. Note also in this context, that even over a field K, it is not enough to work with zero sets over the field. For example consider the field $\mathbb{R}$. Then the family of curves

$$x^2 + y^2 = t,$$

inside $\mathbb{R}^2$, where $t \in \mathbb{R}$, is not well behaved. For $t > 0$, we get a circle, for $t = 0$ we get a single point and for $t < 0$, we get the empty set. In other words, if we have an algebraic variety, it is not enough to consider the ordinary points over $\mathbb{R}$. This becomes even clearer if we work over a finite field. It is clear that different geometric objects, which have very different dimensions, will have the same zero set.

Finally, it is often useful to attack problems in commutative algebra, by considering the corresponding affine variety. In these terms, restricting to finitedly generated algebras without nilpotents is unnecessarily restrictive.

The definition of an affine scheme is motivated by the correspondence between affine varieties and finitely generated algebras over a field, without nilpotents. The idea is that we should be able to associate to any ring R, a topological space X, and a set of continuous functions on X, which is equal to R. In practice this is too much to expect and we need to work with a slightly more general object than a continuous function.

Now if X is an affine variety, the points of X are in correspondence with the maximal ideals of the coordinate ring $A = A(X)$. Unfortunately if we have two arbitrary rings R and S, then the inverse image of a maximal ideal won't be maximal. However it is easy to see that the inverse image of a prime ideal is a prime ideal.

Definition 7.1. *Let R be a ring. $X = \operatorname{Spec} R$ denotes the set of prime ideals of R. X is called the* ***spectrum*** *of R.*

Note that given an element of R, we may think of it as a function on X, by considering it value in the quotient. It is interesting to see what these functions look like in specific cases.

Example 7.2. *Suppose that we take $X = \operatorname{Spec} k[x, y]$. Now any element $f = f(x, y) \in k[x, y]$ defines a function on X. Suppose that we consider a maximal ideal of the form $\mathfrak{p} = \langle x - a, x - b\rangle$. Then the value of f at $\mathfrak{p}$ is equal to the class of f inside the quotient*

$$R/\mathfrak{p} = \frac{k[x, y]}{\langle x - a, x - b\rangle}.$$

If we identify the quotient with k, under the obvious identification, then this is the same as evaluating f at (a, b).

Example 7.3. *Now consider $\mathbb{Z}$. Suppose that we choose an element $n \in \mathbb{Z}$. Then the value of n at the prime ideal $\mathfrak{p} = \langle p\rangle$ is equal to the value of n modulo p. For example, consider $n = 60$. Then the value of this function at the point 7 is equal to $60 \mod 7 = 4 \mod 7$. Moroever 60 has zeroes at 2, 3 and 5, where both 3 and 5 are ordinary zeroes, but 2 is a double zero.*

Example 7.4. *Suppose that we take the ring $R = k[x]/\langle x^2\rangle$. Then the spectrum contains only one element, the prime ideal $\langle x\rangle$. Consider the element $x \in R$. Then x is zero on the unique element of the spectrum, but it is not the zero element of the ring.*

Now we wish to define a topology on the spectrum of a ring. We want to make the functions above continuous. So given an element $f \in R$, we want the set

$$\{\, \mathfrak{p} \in \operatorname{Spec} R \mid f(\mathfrak{p}) = 0 \,\} = \{\, \mathfrak{p} \in \operatorname{Spec} R \mid \langle f\rangle \subset \mathfrak{p} \,\},$$

to be closed. Given that any ideal $\mathfrak{a}$ is the union of all the principal ideals contained in it, so that the set of prime ideals which contain $\mathfrak{a}$ is equal to the intersection of prime ideals which contain every principal ideal contained in $\mathfrak{a}$ and given that the intersection of closed sets is closed, we have an obvious candidate for the closed sets:

Definition 7.5. *The **Zariski topology** on X is given by taking the closed sets to be*

$$V(\mathfrak{a}) = \{\mathfrak{p} \in \operatorname{Spec} R \mid \mathfrak{a} \subset \mathfrak{p}\,\},$$

where $\mathfrak{a}$ is any ideal of R.

Lemma 7.6. *Let R be a ring.*

Then $X = \operatorname{Spec} R$ is a topological space. Moreover the open sets

$$U_f = \{\, \mathfrak{p} \in R \mid f \notin \mathfrak{p} \,\},$$

form a base for the topology.

Proof. Easy check. □

By what we said above, the Zariski topology is the weakest topology so that the zero sets of $f \in R$ are closed.

Example 7.7. *Let k be a field. Then* $\operatorname{Spec} k$ *consists of a single point.*

Example 7.8. *Now consider* $\operatorname{Spec} k[x]$*. If k is an algebraically closed field, then by the Nullstellensatz, the maximal ideals are in correspondence with the points of k. However, since $k[x]$ is an integral domain, the zero ideal is a prime ideal. Since $k[x]$ is a PID, the proper closed sets of X consist of finite unions of maximal ideals. The closure of the point $\xi = \langle 0 \rangle$ is then the whole of X. In particular, not only is the Zariski topology, for schemes, not Hausdorff or T_2, it is not even T_1.*

Example 7.9. *Now consider $k[x, y]$, where k is an algebraically closed field. Prime ideals come in three types. The maximal ideals correspond to points of k^2. The zero ideal, whose closure consists of the whole of X. But there are also the prime ideals which correspond to prime elements $f \in k[x, y]$. The zero locus of f is then an irreducible curve C, and in fact the closure of the point $\xi = \langle f \rangle$ is the curve C. The proper closed sets thus consist of a finite union of maximal ideals, union infinite sets of the maximal ideals which consist of all points belonging to an affine curve C, together with the ideal of each such curve.*

Example 7.10. *Now suppose that k is not algebraically closed. For example, consider* $\operatorname{Spec} \mathbb{R}[x]$*. As before the closure of the zero ideal consists of the whole of X. The maximal ideals come in two flavours. First there are the ideals $\langle x - a \rangle$, where $a \in \mathbb{R}$. But in addition there are also the ideals*

$$\langle x^2 + ax + b = (x - \alpha - i\beta)(x - \alpha + i\beta) \rangle,$$

where a, b, α and $\beta > 0$ are real numbers, so that $b^2 - 4a < 0$.

Example 7.11. *There is a very similar (but more complicated) picture inside* $\operatorname{Spec} \mathbb{R}[x, y]$*. The set $V(x^2+y^2 = -1)$ does not contain any ideals of the first kind, but it contains many ideals of the second kind. In the classical picture, the conic does $x^2 + y^2 = -1$ does not contain any points but it does contain many points if you include all prime ideals.*

Example 7.12. *Now suppose that we take $\mathbb{Z}$. In this case the maximal ideals correspond to the prime numbers, and in addition there is one point whose closure is the whole spectrum. In this respect* $\operatorname{Spec} \mathbb{Z}$ *is very similar to* $\operatorname{Spec} k[t]$*.*

We will need one very useful fact from commutative algebra:

Lemma 7.13. *If $\mathfrak{a} \trianglelefteq R$ is an ideal in a ring R then the radical of $\mathfrak{a}$ is the intersection of all prime ideals containing $\mathfrak{a}$.*

Proof. One inclusion is clear; every prime ideal $\mathfrak{p}$ is radical (that is, equal to its own radical) and so the intersection of all prime ideals containing $\mathfrak{a}$ is radical.

Now suppose that r does not belong to the radical of $\mathfrak{a}$. Let $\mathfrak{b}$ be the ideal generated by the image of $\mathfrak{a}$ inside the ring R_r. Then the image of r inside the quotient ring $R_r/\mathfrak{b}$ is non-zero. Pick an ideal in this ring, maximal with respect to the property that it does not contain the image of r. Then the inverse image $\mathfrak{p}$ of this ideal is a prime ideal which does not contain r. $\square$

Lemma 7.14. *Let X be the spectrum of the ring R and let $f \in R$.*

If $U_f = \bigcup U_{g_i}$ then $f^n = \sum b_i g_i$, where $b_1, b_2, \dots, b_k \in R$. In particular, U_f is a finite union of the U_{g_i} and U_f is compact.

Proof. Taking complements, we see that

$$V(\langle f \rangle) = \bigcap_i V(\langle g_i \rangle) = V(\langle \sum_i g_i \rangle).$$

Now $V(\mathfrak{a})$ consists of all prime ideals that contain $\mathfrak{a}$, and the radical of $\mathfrak{a}$ is the intersection of all the prime ideals that contain $\mathfrak{a}$. Thus

$$\sqrt{\langle f \rangle} = \sqrt{\langle \sum_i g_i \rangle}.$$

But then, in particular, f^n is a finite linear combination of the g_i and the corresponding open sets cover U_f. $\square$

8. Affine Schemes II

As pointed out in §7, we need a slightly more general notion of a function than the one given above:

Definition 8.1. *Let R be a ring. We define a sheaf of rings $\mathcal{O}_X$ on the spectrum X of R as follows. Let U be any open set of X. A section $s \in \mathcal{O}_X(U)$ is by definition any function*

$$s \colon U \longrightarrow \coprod_{\mathfrak{p} \in U} R_{\mathfrak{p}},$$

where $s(\mathfrak{p}) \in R_{\mathfrak{p}}$, which is locally represented by a quotient. More precisely, given a point $\mathfrak{q} \in U$, there is an element $f \in R$ such that $\mathfrak{q} \in U_f \subset U$ and such that the section $s|_{U_f}$ is represented by a/f^n, for some $a \in R$ and $n \in \mathbb{N}$.

*An **affine scheme** is then any locally ringed space isomorphic to the spectrum of a ring with its associated sheaf. A **scheme** is a locally ringed space, which is locally isomorphic, as locally ringed space, to an affine scheme.*

It is not hard to see that $\mathcal{O}_X(U)$ is a ring (sums and products are defined in the obvious way) and that we do in fact have a sheaf rather than just a presheaf.

The key result is the following:

Lemma 8.2. *Let X be an affine scheme, isomorphic to the the spectrum of R and let $f \in R$.*

1. *For any $\mathfrak{p} \in X$, the stalk $\mathcal{O}_{X,\mathfrak{p}}$ is isomorphic to the local ring $R_{\mathfrak{p}}$.*
2. *The ring $\mathcal{O}_X(U_f)$ is isomorphic to R_f.*

In particular $\mathcal{O}_X(X) \simeq R$.

Proof. We first prove (1). There is an obvious ring homomorphism

$$\mathcal{O}_{X,\mathfrak{p}} \longrightarrow R_{\mathfrak{p}},$$

which just sends a germ (g, U) to its value $g(\mathfrak{p})$ at $\mathfrak{p}$.

On the other hand, there is an obvious ring homomorphism,

$$R \longrightarrow \mathcal{O}_{X,\mathfrak{p}},$$

which sends an element $r \in R$ to the pair (r, X). Suppose that $f \notin \mathfrak{p}$. Then $(1/f, U_f)$ defines an element of $\mathcal{O}_{X,\mathfrak{p}}$, and this element is an inverse of (f, X). It follows, by the universal property of the localisation, that there is a ring homomorphism,

$$R_{\mathfrak{p}} \longrightarrow \mathcal{O}_{X,\mathfrak{p}},$$

which is the inverse map. Hence (1).

Now we turn to the proof of (2). As before there is an obvious ring homomorphism,

$$R \longrightarrow \mathcal{O}_X(U_f),$$

which induces a ring homomorphism

$$R_f \longrightarrow \mathcal{O}_X(U_f).$$

We have to show that this map is an isomorphism. We first consider injectivity. Suppose that $a/f^n \in R_f$ is sent to zero. Then for every $\mathfrak{p} \in \operatorname{Spec} R$, $f \notin \mathfrak{p}$, the image of a/f^n is equal to zero in $R_{\mathfrak{p}}$. For each such prime $\mathfrak{p}$ there is an element $h \notin \mathfrak{p}$ such that $ha = 0$ in R. Let $\mathfrak{a}$ be the annihilator of a in R. Then $h \in \mathfrak{a}$ and $h \notin \mathfrak{p}$, so that $\mathfrak{a}$ is not a subset of $\mathfrak{p}$. Since this holds for every $\mathfrak{p} \in U_f$, it follows that $V(\mathfrak{a}) \cap U_f = \emptyset$. But then $f \in \sqrt{\mathfrak{a}}$ so that $f^l \in \mathfrak{a}$, for some l. It follows that $f^l a = 0$, so that a/f^n is zero in R_f. Thus the map is injective.

Now consider surjectivity. Pick $s \in \mathcal{O}_X(U_f)$. By assumption, we may cover U_f by open sets V_i such that s is represented by $a_i/g_i^{n_i}$ on V_i. Replacing g_i by $g_i^{n_i}$ we may assume that $n_i = 1$. By definition $g_i \notin \mathfrak{p}$, for every $\mathfrak{p} \in V_i$, so that $V_i \subset U_{g_i}$. Now since sets of the form U_h form a base for the topology, we may assume that $V_i = U_{h_i}$. As $U_{h_i} \subset U_{g_i}$ it follows that $V(g_i) \subset V(h_i)$ so that

$$\sqrt{\langle h_i \rangle} \subset \sqrt{\langle g_i \rangle}.$$

But then $h_i^{n_i} \in \langle g_i \rangle$, so that $h_i^{n_i} = c_i g_i$. In particular

$$\frac{a_i}{g_i} = \frac{c_i a_i}{h_i^{n_i}}.$$

Replacing h_i by $h_i^{n_i}$ and a_i by $c_i a_i$, we may assume that U_f is covered by U_{h_i}, and that s is represented by a_i/h_i on U_{h_i}.

We have already shown that $f^n = \sum b_i h_i$, where $b_1, b_2, \dots, b_k \in R$ and U_f can be covered by finitely many of the sets U_{h_i}. Thus we may assume that we have only finitely many h_i. Now on $U_{h_i h_j} = U_{h_i} \cap U_{h_j}$, there are two ways to represent s, one way by a_i/h_i and the other by a_j/h_j. By injectivity, we have $a_i/h_i = a_j/h_j$ in $R_{h_i h_j}$ so that for some n,

$$(h_i h_j)^n (h_j a_i - h_i a_j) = 0.$$

Since there are only finitely many i and j, we may assume that n is independent of i and j. We may rewrite this equation as

$$h_j^{n+1}(h_i^n a_i) - h_i^{n+1}(h_j^n a_j) = 0.$$

If we replace h_i by h_i^{n+1} and a_i by $h_i^n a_i$, then s is still represented by a_i/h_i and moreover

$$h_j a_i = h_i a_j.$$

Let $a = \sum_i b_i a_i$, where $f^n = \sum_i b_i h_i$. Then for each j,

$$\begin{aligned} h_j a &= \sum_i b_i a_i h_j \\ &= \sum_i b_i h_i a_j \\ &= f^n a_j. \end{aligned}$$

But then $a/f^n = a_j/h_j$ on U_{h_j}. But then a/f^n represents s on the whole of U_f. □

Note that by (2) of (8.2), we have achieved our aim of constructing a topological space from an arbitrary ring R, which realises R as a natural subset of the continuous functions.

Definition 8.3. *A morphism of schemes is simply a morphism between two locally ringed spaces which are schemes.*

The gives us a category, the category of schemes. Note that the category of schemes contains the category of affine schemes as a full subcategory and that the category of schemes is a full subcategory of the category of locally ringed spaces.

Theorem 8.4. *There is an equivalence of categories between the category of affine schemes and the category of commutative rings with unity.*

Proof. Let F be the functor that associates to an affine scheme, the global sections of the structure sheaf. Given a morphism

$$(f, f^{\#})\colon (X, \mathcal{O}_X) = \operatorname{Spec} B \longrightarrow (Y, \mathcal{O}_Y) = \operatorname{Spec} A,$$

of locally ringed spaces then let

$$\phi\colon A \longrightarrow B,$$

be the induced map on global sections. It is clear that F is then a contravariant functor and F is essentially surjective by (8.2).

Now suppose that $\phi\colon A \longrightarrow B$ is a ring homomorphism. We are going to construct a morphism

$$(f, f^{\#})\colon (X, \mathcal{O}_X) \longrightarrow (Y, \mathcal{O}_Y),$$

of locally ringed spaces. Suppose that we are given $\mathfrak{p} \in X$. Then $\mathfrak{p}$ is a prime ideal of B. But then $\mathfrak{q} = \phi^{-1}(\mathfrak{p})$ is a prime ideal of A. Thus we get a function $f : X \longrightarrow Y$. Now if $\mathfrak{a}$ is an ideal of A, then

$f^{-1}(V(\mathfrak{a})) = V(\langle\phi(\mathfrak{a})\rangle)$, so that f is certainly continuous. For each prime ideal $\mathfrak{p}$ of B, there is an induced morphism

$$\phi_{\mathfrak{p}}\colon A_{\phi^{-1}(\mathfrak{p})} \longrightarrow B_{\mathfrak{p}},$$

of local rings. Now suppose that $V \subset Y$ is an open set. We want to define a ring homomorphism

$$f^{\#}(V)\colon \mathcal{O}_Y(V) \longrightarrow \mathcal{O}_X(f^{-1}(V)).$$

Suppose first that $V = U_g$, where $g \in A$. Then $\mathcal{O}_Y(V) = A_g$ and $f^{-1}(V) \subset U_{\phi(g)}$. But then there is a restriction map

$$\mathcal{O}_X(U_{\phi(g)}) \simeq B_{\phi(g)} \longrightarrow \mathcal{O}_X(f^{-1}(V)).$$

On the other hand, composing there is a ring homomorphism

$$A \longrightarrow B_{\phi(g)}.$$

Since the image of g is invertible, by the universal property of the localisation, there is an induced ring homomorphism

$$A_g \longrightarrow B_{\phi(g)}.$$

Putting all of this together, we have defined $f^{\#}(V)$ when $V = U_g$. Since the sets U_g form a base for the topology, and these maps are compatible in the obvious sense, this defines a morphism

$$f^{\#}\colon \mathcal{O}_Y \longrightarrow f_*\mathcal{O}_X,$$

of sheaves. Clearly the induced map on local rings is given by $\phi_{\mathfrak{p}}$, and so $(f, f^{\#})$ is a morphism of local rings.

Finally it suffices to prove that these two assignments are inverse. The composition one way is clear. If we start with ϕ and construct $(f, f^{\#})$ then we get back ϕ on global sections. Conversely suppose that we start with $(f, f^{\#})$, and let ϕ be the map on global sections. Given $\mathfrak{p} \in X$, we get a morphism of local rings on stalks, which is compatible with ϕ and localisation, so that we get a commutative diagram

$$\begin{tikzcd} A \arrow[r, "\phi"] \arrow[d] & B \arrow[d] \\ A_{f(\mathfrak{p})} \arrow[r, "f^{\#}_{\mathfrak{p}}"] & B_{\mathfrak{p}}. \end{tikzcd}$$

Let's compare $f(\mathfrak{p})$ and $\phi^{-1}(\mathfrak{p})$. If $r \notin f(\mathfrak{p})$ then the image of r in $A_{f(\mathfrak{p})}$ is a unit, so that $f^{\#}_{\mathfrak{p}}(r)$ is a unit. Hence $\phi(r) \notin \mathfrak{p}$, that is, $r \notin \phi^{-1}(\mathfrak{p})$. On the other hand, as $f^{\#}_{\mathfrak{p}}$ is a local ring homomorphism, it follows that the inverse image of a unit in $B_{\mathfrak{p}}$ is a unit in $A_{f(\mathfrak{p})}$. Pick $r \notin \phi^{-1}(\mathfrak{p})$. Then $\phi(r) \notin \mathfrak{p}$, and this is sent to a unit in $B_{\mathfrak{p}}$. Thus the image of r in $A_{f(\mathfrak{p})}$ is a unit and so $r \notin f(\mathfrak{p})$. Thus $f(\mathfrak{p}) = \phi^{-1}(\mathfrak{p})$.

Now let's compare $f^{\#}$ and the map $g^{\#}$ associated to ϕ. Their difference is a morphisms of sheaves,

$$f^{\#} - g^{\#} \colon \mathcal{O}_Y \longrightarrow f_*\mathcal{O}_X,$$

This morphism is zero on stalks, as we have seen, so that it is the zero morphism. Thus $f^{\#} = g^{\#}$. $\square$

9. Enumerative geometry

Here is a typical question in enumerative geometry:

Question 9.1. *How many circles, in the usual real plane, pass through three points which are not collinear?*

The answer is one. Probably the easiest way to see this is to use synthetic geometry. Suppose the points are p, q and r. Let L and M be the bisectors of the two lines $\langle p, q\rangle$ and $\langle q, r\rangle$ (if S is a set of points then let $\langle S\rangle$ denote the span of S). Then it is easy to see that the point of intersection $L \cap M$ is the centre of the circle we are looking for and that this is the only circle through p, q and r.

However there are two entirely different ways to proceed, both of which will prove more fruitful, as they are more general.

Here is the first. Imagine moving the points around. Clearly the answer won't change (or better, if it did the original question does not really make sense). Now suppose that the points become collinear. In this case the only circle through these points is the straight line (a circle of infinite radius) containing them. Supposing that the answer does not change the answer must then be one in general. It is convenient to state more clearly the underlying assumption.

Principle 9.2. *(Principle of continuity) If we are given a problem in enumerative geometry, then the number of solutions is invariant under a continuous change of parameters.*

This is a very useful principle; unfortunately as stated it is clearly false, as there are some obvious counterexamples. The point is to change the definitions, so that this principle does indeed hold.

Question 9.3. *In how many points do two lines intersect?*

At first sight the answer would seem to be one; unfortunately some lines are parallel. In fact it is clear that the principle of continuity fails as well.

Example 9.4. *Let L be the line $y = 0$ and let M_t be the line $y = tx+1$, where $t \in K$. Then as t approaches zero, M approaches a line parallel to L, so that the number of points*

$$L \cap M_t,$$

is not constant, it is one for $t \neq 0$ but zero when $t = 0$.

Consider how the principle of continuity fails in this case. We have a sequence of points, $L \cap M_t$, without a limit. If we had a topological space (for example take $K = \mathbb{R}$), then this can only happen if the space is not compact. So we could fix the problem if we can compactify $\mathbb{A}^2$, by adding some points at infinity.

Definition 9.5. *Let K be a field and let V be a vector space V of dimension $n+1$. $\mathbb{P}(V)$ denotes the space of lines in V.* ***Projective space of dimension*** *n, denoted by $\mathbb{P}^n_K$, is the case $V = K^{n+1}$.*

Note that, as V is a vector space and not affine space, a line in V contains the origin.

Let us examine this definition more closely. Let V be a vector space of dimension $n+1$. Pick $v \in V - \{0\}$. Then v determines a line $\langle v\rangle$, in the usual way. On the other hand, if w is another non-zero vector, proportional to v, that is, $w = \lambda v$, for some $\lambda \neq 0 \in K$, then $\langle v\rangle = \langle w\rangle$. Thus we have proved:

Definition-Lemma 9.6. *Let V be a vector space. $\mathbb{P}(V)$ is equal to the set of points V, modulo the equivalence relation $\sim$, defined as $v \sim w$ iff $v = \lambda w$, $\lambda \in K^*$.*

The equivalence class of the vector v is denoted $[v]$.

Let us see what happens for small values of n. If $n+1 = 0$, then V does not contain any non-zero vectors, and so $\mathbb{P}^{-1}$ is empty. If $n+1 = 1$, then V contains a unique line and so $\mathbb{P}^0$ is a point.

The first interesting case is $\mathbb{P}^1$. Let $V = K^2$. Then $\mathbb{P}^1$ is the set of lines in the plane K^2. Suppose that $v = (X, Y) \in K^2 - \{0\}$. We denote the corresponding point of $\mathbb{P}^1$, by $[v] = [X : Y]$. Then the line spanned by v has a slope, provided $X \neq 0$, and this uniquely determines the line.

The slope $m = Y/X$ takes any value in K. Thus $\mathbb{A}^1 \subset \mathbb{P}^1$. On the other hand, we are only missing one point, corresponding to the line with slope infinity. Thus $\mathbb{P}^1 = \mathbb{A}^1 \cup \{p\}$, and we have compactified $\mathbb{A}^1$, by adding a single point. In fact, we sometimes refer to p as the point at infinity and even denote it by ∞ (the value of Y/X as it were). As an equivalence class, $p = [0 : 1]$.

Note that this situation is completely symmetric. Instead of looking at $y = Y/X$, we could consider $x = X/Y$. In this case we compactify $\mathbb{A}^1$, with coordinate x, by adding the point $q = [1 : 0]$.

It is useful to introduce some more notation to handle this. We denote by U_0 the locus of points of $\mathbb{P}^1$ where $X \neq 0$. As we have already seen, U_0 is a copy of $\mathbb{A}^1$. In this case $\mathbb{P}^1 = U_0 \cup \{[0 : 1]\}$.

Similarly we denote by U_1 the locus of points where $Y \neq 0$. Thus $\mathbb{P}^1 = U_1 \cup \{[1 : 0]\}$. The two sets U_0 and U_1 obviously intersect, along the locus $XY \neq 0$.

Let us see what happens for $\mathbb{P}^2$. Introduce coordinates (X, Y, Z) on $V \simeq K^3$. There are three obvious loci to consider, $X \neq 0$, $Y \neq 0$ and $Z \neq 0$. These induce three subsets of $\mathbb{P}^2$, U_0 and U_1 and U_2. I claim that that U_i is a copy of $\mathbb{A}^2$.

It is easy to see this algebraically. If $(X, Y, Z) \in K^3$ and $X \neq 0$, then $[X : Y : Z] = [1 : Y/X : Z/X]$. Thus the ratios $y = Y/X$ and $z = Z/X$ define coordinates on U_0 and identify U_0 with $\mathbb{A}^2$.

One can also see this geometrically. Any line through the origin of K^3 is determined by its intersection with the locus $X = 1$ (assuming it does intersect, that is, assuming the line lies in U_0). But the locus $X = 1$ is surely a copy of $\mathbb{A}^2$.

What is missing? In other words, what is $\mathbb{P}^2 - U_0$? This is the set of points with zero first coordinate, in other words all points of the form $[0 : Y : Z]$. But this is surely a copy of $\mathbb{P}^1$.

In other words we can compactify $\mathbb{A}^2$ by adding a copy of $\mathbb{P}^1$, to get $\mathbb{P}^2$. This copy of $\mathbb{P}^1$ is sometimes called the *line at infinity.*

As before, the situation is completely symmetric. Moreover, all of this generalises in an obvious fashion.

Definition 9.7. *Pick coordinates $X_0, X_1, \dots, X_n$ on K^{n+1}. We will refer to $[X_0 : X_1 : \cdots : X_n]$ as* ***homogeneous coordinates*** *on $\mathbb{P}^n$. The subsets $U_0, U_1, \dots, U_n$, given as $X_i \neq 0$, which are copies of $\mathbb{A}^n$, are called the* ***standard open affine subsets****. Indeed the ratios $x_i = \frac{X_j}{X_i}$ define coordinates $x_1, x_2, \dots, \hat{x}_i, \dots, x_n$ on U_i.*

The locus $X_i = 0$ is called the ***hyperplane at infinity****. $\mathbb{P}^n = U_i \cup \{X_i = 0\}$.*

Note that what is at infinity, depends on our point of view. Note also that the term homogeneous coordinates is a bit of a misnomer. In fact $X_0, X_1, \dots, X_n$ are not functions at all, since they are not invariant under rescaling. The only thing that does make sense, is to ask where they are zero (which is invariant under rescaling).

Definition 9.8. *A subset Λ of a projective space $\mathbb{P}(V)$ is called* ***linear*** *if it is given as $\mathbb{P}(W)$, where $W \subset V$ is a linear subspace. The* ***dimension*** *of Λ is the dimension of W minus one.*

In other words a line l in $\mathbb{P}^2$ is the same as a plane W in the corresponding three dimensional vector space K^3.

Lemma 9.9. *Let Λ_1 and Λ_2 be two linear subspaces of $\mathbb{P}^n$ of dimension r and s.*

Then the dimension of the intersection is at least $r + s - n$.

Proof. Let W_1 and W_2 be the corresponding linear subspaces of V, where $\mathbb{P}^n = \mathbb{P}(V)$. Then W_1 has dimension $r+1$, W_2 has dimension $s+1$ and V has dimension $n+1$.

Clearly $\Lambda_1 \cap \Lambda_2 = \mathbb{P}(W_1 \cap W_2)$. On the other hand

$$\begin{aligned}\dim(W_1 \cap W_2) &\geq (r+1) + (s+1) - (n+1)\\ &= r+s-n+1.\end{aligned}$$

□

The following example shows that we have fixed out problem concerning parallel lines.

Example 9.10. *Let l_1 and l_2 be two lines in $\mathbb{P}^2$. Then $l_1 \cap l_2$ intersect. Indeed the dimension of the intersection is at least zero ($= 1+1-2$) and the empty set has dimension -1.*

We will see later what happens when we take two parallel lines in $\mathbb{A}^2$ and compactify to $\mathbb{P}^2$. In practice it is often more efficient to work with the codimension and not the dimension.

Definition 9.11. *Let $\Lambda \subset \mathbb{P}^n$ be linear subspace. The* ***codimension*** *of Λ is equal to the difference $n-d$, where d is the dimension of Λ.*

The following is a simple restatement of (9.9); its virtue lies in the fact that is easier to remember and apply:

Lemma 9.12. *Let Λ_1 and Λ_2 be two linear subspaces of $\mathbb{P}^n$ of codimension r and s.*

Then the codimension of the intersection is at most $r+s$. That is the codimension of the intersection is at most the sum of the codimensions.

Let us go back to the principle of continuity. Unfortunately there is another problem.

Question 9.13. *In how many points do a line and a circle meet?*

Example 9.14. *Let L be the line $x = \sqrt{2}$ and C the circle $x^2 + y^2 = 1$ in $\mathbb{A}^2_{\mathbb{R}}$. Then L and C don't intersect.*

Now consider the family of lines L_t, $x = t$. Then $L_t \cap C$ depends on $t \in \mathbb{R}$. If $|t| < 1$ we get two points, if $t = \pm 1$, we get one and if $|t| > 1$ none at all. Thus the principle of continuity does not hold up.

The important thing to realise is that the problem here has nothing to do with the points of intersection moving off to infinity. The problem is that $\mathbb{R}$ is not algebraically closed, so that the equation $y^2 = -1$ has no solutions.

The solution is simple, we should replace $\mathbb{R}$ with $\mathbb{C}$. Now we always get two points (ignoring the possibility that $t = \pm 1$, which we will come back to), $x = t$, $y = \pm\sqrt{1-t^2}$.

In practice, when we are considering problems in enumerative geometry, we will work almost exclusively over an algebraically closed field of characteristic zero, which for all intents and purposes means we work over $\mathbb{C}$. In fact working with other fields normally poses extra technical problems, so that working over $\mathbb{C}$ is the most convenient.

10. Conics in $\mathbb{P}^2$

We want to talk about curves in $\mathbb{P}^2$. For that we need to look at polynomials. The problem is that polynomials in X, Y and Z don't define functions on $\mathbb{P}^2$, since polynomials are not invariant under rescaling. However, we don't really care what the value of the polynomial is, all we care about is whether or not the polynomial is zero.

Definition 10.1. *Let $F(X) \in K[X]$ be a polynomial in the variables $X_0, X_1, \dots, X_n$. We say that F is* ***homogeneous*** *if every non-zero term of F has the same degree d.*

Lemma 10.2. *Let $F(X) \in K[X]$.*

(1) If F is homogeneous of degree d, then $F(\lambda X) = \lambda^d F(X)$, for all $\lambda \in K$ (we adopt the convention here that $0^0 = 1$).
(2) Conversely, if $F(\lambda X) = \lambda^d F(X)$, for all $\lambda \in K$ and K is infinite then F is homogeneous of degree d.

Proof. (1) is clear. Suppose now that F is any polynomial. Then $F = \sum_i F_i$ has a unique decomposition, where F_i is homogeneous of degree i. If $F(\lambda X) = \lambda^d F(X)$, then this forces $\lambda^i F_i(X) = \lambda^d F_i(X)$, for every i. If K is infinite then for every $i \neq d$, we can pick λ, so that $\lambda^i \neq \lambda^d$. Thus $F_i(X) = 0$, for all $i \neq d$. □

Definition-Lemma 10.3. *Let S be a set of homogeneous polynomials. The zero set $X = V(S)$ is called a* ***projective subvariety*** *of $\mathbb{P}^n$.*

The ***Zariski topology*** *on $\mathbb{P}^n$ is the topology whose closed subsets are the projective subvarieties.*

Lemma 10.4. *$\Lambda \subset \mathbb{P}^n$ is a linear subspace if and only if it is defined by a collection of homogeneous linear equations.*

In particular every linear subspace of $\mathbb{P}^n$ is a projective subvariety.

Proof. Clear, since a subset $W \subset V$ is a linear subspace if and only if it is defined by homogeneous linear equations. □

One of the key points, is that we can go backwards and forwards between affine and projective varieties.

First let us suppose that we are given a subset $V \subset \mathbb{P}^n$. Clearly we can form $V \subset U_0$ simply by intersecting V with U_0. Suppose that V is a closed subvariety, say defined by $F_\alpha(X)$ homogenous. Define $f_\alpha(x)$ by replacing X_i by X_i/X_0. It is pretty easy to see that V_0 is defined by the f_α.

Conversely suppose we are given f_α, which defines $V \subset \mathbb{A}^n$. Then we can form $F_\alpha(X)$ homogeneous, simply by topping up each term of

f_α, by the appropriate power of X_0. This defines $\bar{V}$ in X. Again, it is not hard to see that $\bar{V} \cap U_0 = V$.

In both cases, the best way to see what is going on, is to look at some examples.

Example 10.5. *Suppose we consider $x^2 + y^2 = 1$ inside $\mathbb{A}^2$. We may think of this as $U_2 \subset \mathbb{P}^3$, with coordinates X, Y and Z. Replace $x^2+y^2 = 1$ by x^2+y^2-1. This has degree two. The first two terms have degree two, and there is nothing to do (apart from replacing lower caps by upper). The last term has degree zero. To make this homogeneous then, we need to multiply by Z^2. We get $X^2 + Y^2 - Z^2$. Now suppose we want to work on U_0. Then we divide through by X^2 and replace Y/X by y and Z/X by z, to get $1 + y^2 - z^2$. Note the quick way to do this is simply to replace X by 1 and replace upper caps by lower.*

Example 10.6. *Consider $y = x^3$. We get $y - x^3$. Consider this inside $\mathbb{P}^2$, with coordinates X, Y and Z. We get $YZ^2 - X^3$. Now work inside U_1. We get $z^2 - x^3$.*

It is interesting to see what happens to parallel lines in $\mathbb{A}^2$.

Example 10.7. *Let L be the line $y = 0$ and let M_t be the line $y = tx + 1$, where $t \in K$. Then L becomes the line $Y = 0$ and M_t the line $Y = tX + Z$. When $t = 0$, we get $Y = Z$. Thus $Z = 0$, and we get the point $[1 : 0 : 0]$. Thus our two parallel lines intersect along the line at infinity, at the point $[1 : 0 : 0]$, corresponding to the fact that both lines are horizontal.*

In fact it is interesting to consider the family in the coordinate patch $Y \neq 0$. We get $x = 0$ and $x = t + z$, which is equivalent to $x = 0$ and $z = -t$.

Note that these processes are not quite inverse.

Example 10.8. *Suppose we start with $X = 0$ inside $\mathbb{P}^2$. If we go to the coordinate patch U_0 then we get the empty set. Going back to $\mathbb{P}^2$, we get the empty set. The whole point is that the whole of $X = 0$ completely avoids the set U_0.*

One of the beautiful results of classical projective geometry is the following:

Lemma 10.9. *Let $f \in \mathbb{R}[x, y]$ be a polynomial of degree two. Suppose that $f = 0$ contains more than one real point. Let F be the homogenisation of f.*

Then $f = 0$ is a circle if and only if $F = 0$ contains the points $[1 : \pm i : 0]$.

Proof. Suppose that $f = 0$ defines a circle. Then $f(x, y)$ has the form

$$(x - a)^2 + (y - b)^2 = r^2.$$

Thus F is equal to

$$(X - aZ)^2 + (Y - bZ)^2 = r^2 Z^2.$$

Set $Z = 0$. Then $X^2 + Y^2 = 0$, which has the solution $[1 : \pm i : 0]$.

Conversely suppose that $F = 0$ contains the points $[1 : \pm i : 0]$. Then

$$F(X, Y, 0) = aX^2 + bXY + cY^2,$$

vanishes at $[1 : \pm i : 0]$. Thus

$$ax^2 + bx + c = 0,$$

has roots $\pm i$, which is only possible if $b = 0$ and $a = c$. Hence $F(X, Y, 0)$ is a non-zero multiple of $X^2 + Y^2$. Possibly rescaling, we may assume that

$$F(X, Y, Z) = X^2 + Y^2 + ZG(X, Y, Z)$$

where $G(X, Y, Z)$ is a linear polynomial. Thus

$$f(x, y) = x^2 + y^2 + g(x, y),$$

for some linear polynomial g. Completing the square, we can put this in the form

$$(x - a)^2 + (y - b)^2 = k.$$

The condition that $f = 0$ contains more than one point is equivalent to requiring that $k > 0$, so that $k = r^2$, some $r > 0$ and we have the equation of a circle. $\square$

Since we want to work over $\mathbb{C}$, it turns out that we want to reinvent the wheel:

Definition 10.10. *The curve $C \subset \mathbb{P}^2_{\mathbb{C}}$, given as $F = 0$, is a **circle** if F has degree two and C contains the points* $[1 : \pm i : 0]$.

Let us consider the general polynomial of degree two in X, Y and Z,

$$F(X, Y, Z) = aX^2 + bY^2 + cZ^2 + dYZ + eXZ + fXY,$$

where a, b, c, d, e and f are in K. Thus polynomials of degree two are naturally in correspondence with K^6. On the other, if $F = \lambda G$, $\lambda \neq 0$, then F and G define the same zero locus. Over an algebraically closed field, the converse is true. Thus the set of conics in $\mathbb{P}^2$ is naturally in bijection with K^6 modulo scalars, that is, $\mathbb{P}^5$.

Given that we want to count how many circles pass through two points and that a circle is nothing more than a conic that passes through two fixed points, the natural problem is to identify the following locus:

$$H_p = \{ [a : b : c : d : e : f] \in \mathbb{P}^5 \mid F = 0 \text{ passes through } p \},$$

where $p \in \mathbb{P}^2$ is a point.

Lemma 10.11. *$H_p \subset \mathbb{P}^5$ is a hyperplane (that is, a linear space defined by a single equation).*

Proof. Indeed, if $p = [u : v : w]$, then $[a : b : c : d : e : f] \in H_p$ if and only if $[a : b : c : d : e : f]$ satisfies the linear equation

$$u^2 A + v^2 B + w^2 C + (vw)D + (uw)E + (uv)F = 0. \qquad \square$$

For example, the conic passes through $p = [0 : 0 : 1]$ if and only if the coefficient of Z^2 is zero if and only if $c = 0$.

Lemma 10.12. *Suppose we are given five points p_1, p_2, p_3, p_4 and p_5, and we are working over an infinite field.*

Then, either there is a unique conic through these points, or infinitely many.

Proof. Let $H_i \subset \mathbb{P}^5$ be the hyperplane corresponding to p_i. Then the set of conics passing through the given points corresponds to the intersection of the five hyperplanes. As the intersection of linear spaces is a linear space, the result follows. $\square$

Theorem 10.13. *There is a unique conic passing through five points in linear general position.*

Proof. Suppose not. Then the intersection of the five hyperplanes H_1, H_2, H_3, H_4 and H_5 would contain a line, call it $l \subset \mathbb{P}^5$. Pick two points of this line, corresponding to two quadratic polynomials F and G. As any two points on l, span l, the general point of l is given as $[sF + tG]$, for $[s : t] \in \mathbb{P}^1$. Thus the curve $sF + tG = 0$ contains the five given points p_1, p_2, p_3, p_4 and p_5.

Pick any point $p \in \mathbb{P}^2$. Then we may find $[s : t] \in \mathbb{P}^1$ such that $(sF + tG)(p) = 0$. Indeed, if $G(p) = 0$, take $[s : t] = [0 : 1]$, else set $s = 1$ and

$$t = -\frac{F(p)}{G(p)}.$$

Now pick p collinear with p_1 and p_2. Then the conic C corresponding to $sF + tG = 0$ contains the three points p_1, p_2 and p of the line $m = \langle p_1, p_2 \rangle$. Pick coordinates so that m is given as $Z = 0$. Then the quadratic polynomial

$$F(X, Y, 0),$$

has three zeroes. It follows that $F(X,Y,0) = 0$, so that $F(X,Y,Z) = ZG(X,Y,Z)$. In other words the curve C is the union of the two lines $Z = 0$ and $G = 0$. But then one of the two lines contains three of our five points, which contradicts our assumption that the points are in linear general position. $\square$

Corollary 10.14. *There is a unique circle passing through three non-collinear points in* $\mathbb{R}^2$.

Proof. Note that the line spanned by the points $[1 : \pm i : 0]$ is the line at infinity of $\mathbb{P}^2_{\mathbb{C}}$. Thus given three points p, q and r in $\mathbb{R}^2$, which are not collinear, then the five points p, q, r and $[1 : \pm i : 0]$ are in linear general position in $\mathbb{P}^2_{\mathbb{C}}$.

By (10.13) there is a unique conic through the five given points. Now of the five hyperplanes that define this conic, three are defined by linear equations with real coefficients and even though the other two have complex roots, the equations of the hyperplanes are complex conjugates. Since the set of solutions to a set of equations which is invariant under complex conjugation, is invariant under complex conjugation, it follows that this unique solution has coefficients which are invariant under complex conjugation, which is to say that it is a point with real coordinates. In particular the defining equation of the unique conic passing through the five given points is real. On the other hand the corresponding curve contains three real points. Therefore by (10.9) there is a unique circle through the three real points. $\square$

Note the fancy footwork needed to deal with the problem of working over non algebraically closed fields.

It turns out there are other ways to prove (10.13).

11. Morphisms between varieties I

We adopt the following working definition of a morphism between projective varieties:

Definition 11.1. *A* ***morphism***

$$f\colon V \longrightarrow W,$$

between two projective varieties, where $V \subset \mathbb{P}^m$ and $W \subset \mathbb{P}^n$, is given by picking a collection of $n+1$ homogenous polynomials $F_0, F_1, \dots, F_n \in K[X_0, X_1, \dots, X_m]$ of the same degree such that

$$f(x) = [F_0(x) : F_1(x) : \dots : F_n(x)] \in W,$$

for every $x \in V$, ***where for every $x \in V$ there is an i such that*** *$F_i(x) \neq 0$.*

Note that this gives us the category of projective varieties, with maps given by projective morphisms.

Observe that if $w = \lambda v$, and each F_i has degree d, then

$$F_i(w) = \lambda^d F_i(v),$$

so that

$$\begin{aligned}[F_0(w) : F_1(w) : \dots : F_n(w)] &= [\lambda^d F_0(v) : \lambda^d F_1(v) : \dots : \lambda^d F_n(v)] \\ &= [F_0(v) : F_1(v) : \dots : F_n(v)].\end{aligned}$$

Example 11.2. *The map*

$$f\colon \mathbb{P}^1 \longrightarrow \mathbb{P}^2,$$

given by

$$[S : T] \longrightarrow [S^2 : ST : T^2],$$

is a morphism. Indeed we only need to check that S^2, ST and T^2 cannot be simultaneously zero, which is clear.

Consider the image. Suppose that we pick coordinates $[X : Y : Z]$ on $\mathbb{P}^2$. On the image we have

$$\begin{aligned}Y^2 &= (ST)^2 \\ &= S^2T^2 \\ &= XZ.\end{aligned}$$

Thus the image lies in the locus $Y^2 - XZ = 0$. On the other hand suppose we have a point $[X : Y : Z] \in \mathbb{P}^2$. If $X \neq 0$, set $S = X$ and

$T = Y$. Then

$$\begin{aligned}[S^2 : ST : T^2] &= [X^2 : XY : Y^2] \\ &= [X^2 : XY : XZ] \\ &= [X : Y : Z],\end{aligned}$$

as $X \neq 0$. We need to worry about the case $X = 0$. One way to proceed is to observe that then $Y = 0$ so that we have the point $[0 : 0 : 1]$, the image of $[0 : 1]$. Or observe that X and Z cannot simultaneously be zero, and if $Z \neq 0$, we set $S = Y$ and $T = Z$ and argue as before (using the obvious symmetry). Either way we have established that the image of $\mathbb{P}^1$ is a conic in $\mathbb{P}^2$.

This example has many interesting generalisations.

Example 11.3. *Consider the morphism*

$$\mathbb{P}^1 \longrightarrow \mathbb{P}^3,$$

given as

$$[S : T] \longrightarrow [S^3 : S^2T : ST^2 : T^3].$$

*The image C is known as the **twisted cubic**.*

Consider the image. We have $[X : Y : Z : W] = [S^3 : S^2T : ST^2 : T^3]$. Thus certainly $Y^2 = XZ$, $XW = YZ$ and $Z^2 = YW$. It is an interesting exercise to check that these equations define the image.

Example 11.4. *More generally still, we can look at the morphism*

$$\mathbb{P}^1 \longrightarrow \mathbb{P}^d,$$

given as

$$[S : T] \longrightarrow [S^d : S^{d-1}T : \cdots : ST^{d-1} : T^d].$$

*The image is called a **rational normal curve of degree** d.*

Example 11.5. *More generally still, there is a morphism*

$$\mathbb{P}^n \longrightarrow \mathbb{P}^N,$$

given as

$$[X_0 : X_1 : \cdots : X_n] \longrightarrow [X^I],$$

*where given an n-tuple $I = (i_0, i_1, \dots, i_n)$, X^I denotes the monomial $X_0^{i_0}X_1^{i_1}\dots X_n^{i_n}$. Here we choose coordinates Z_I, where I ranges over all n-tuples of positive integers, whose sum is d. N is equal to the number of such n-tuples, minus one. Note that not every X^I can be zero. Indeed if $X_0^d = X_1^d = \dots X_n^d = 0$, then $X_0 = X_1 = \cdots = X_n = 0$. This morphism is called the d**-uple embedding**.*

Note that for every I, J, I' and J' such that $I + J = I' + J'$, the image lies in the hypersurface

$$Z_I Z_J = Z_{I'} Z_{J'},$$

since

$$X^I X^J = X^{I+J} = X^{I'+J'} = X^{I'} X^{J'}.$$

Once again, in fact the image is cut out by these equations.

Perhaps the most interesting example is to take $n = d = 2$.

Example 11.6. *In this case we get a morphism*

$$\mathbb{P}^2 \longrightarrow \mathbb{P}^5,$$

given as

$$\begin{aligned}[X : Y : Z] \longrightarrow [X^2 : Y^2 : Z^2 : YZ : XZ : XY] = \\ [Z_{(2,0,0)} : Z_{(0,2,0)} : Z_{(0,0,2)} : Z_{(0,1,1)} : Z_{(1,0,1)} : Z_{(1,1,0)}].\end{aligned}$$

This morphism is called the **Veronese morphism** and the image is called the **Veronese surface**. It turns out that the Veronese surface is an exception to practically every (otherwise) general statement about projective varieties.

Finally it seems worthwhile to point out that there are other ways to construct rational normal curves.

Definition 11.7. *Let k be a positive integer. A subset X of projective space is a **determinental variety** if X is the locus where a matrix $M = (F_{ij})$ of homogeneous polynomials F_{ij} has rank at most k.*

For example, consider the matrix

$$\begin{pmatrix} X_0 & X_1 & X_2 & \dots & X_n - 1 \\ X_1 & X_2 & \dots & X_{n-1} & X_n \end{pmatrix}$$

The locus where this matrix has rank one is precisely a rational normal curve. Indeed if

$$[X_0 : X_1 : \dots : X_n] = [S^n : S^{n-1}T : \dots : T^n],$$

then clearly the second row is nothing more than the first row times T/S. Conversely if the given matrix has rank 1, then the second row is a scalar multiple of the first, and it is easy to get the result.

On the other hand, the locus where a matrix has rank at most one, is exactly the locus where the 2×2 minors are all zero. For example, for $n = 3$, we recover the three quadrics containing a twisted cubic.

We can do a similar thing for the Veronese. In this case, we look at the locus where the matrix

$$\begin{pmatrix} Z_{2,0,0} & Z_{1,1,0} & Z_{1,0,1} \\ Z_{1,1,0} & Z_{0,2,0} & Z_{0,1,1} \\ Z_{1,0,1} & Z_{0,1,1} & Z_{0,0,2} \end{pmatrix},$$

has rank one.

12. Change of coordinates

Definition 12.1. $\mathrm{PGL}_n(K)$ *denotes the space of invertible* $n\times n$ *matrices with entries in* K*, modulo the normal subgroup of scalar matrices, that is*

$$\mathrm{PGL}_n(K) = \frac{\mathrm{GL}_n(K)}{K^*}.$$

Note that the canonical action of $\mathrm{GL}_{n+1}(K)$ on K^{n+1} descends to an action of $\mathrm{GL}_{n+1}(K)$ on $\mathbb{P}^n$, in an obvious way. Clearly the set of scalar matrices acts trivially and in fact it is not hard to see that the scalar matrices are the kernel of the induced homomorphism. On the other hand, it is also easy to see that if we fix a matrix A, then the induced bijection

$$\mathbb{P}^n \longrightarrow \mathbb{P}^n$$

is in fact a morphism. Thus the group $\mathrm{PGL}_n(K)$ is a subgroup of the group of all automorphisms of $\mathbb{P}^n$.

It is interesting to see what happens for $\mathbb{P}^1$. Suppose we take a matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Then A sends $[X : Y]$ to

$$[aX + bY : cX + dY].$$

Suppose we work in the affine chart $z = X/Y$. Then A sends z to

$$\begin{aligned} \frac{aX + bY}{cX + dY} &= \frac{a(X/Y) + b}{c(X/Y) + d} \\ &= \frac{az + b}{cz + d}. \end{aligned}$$

In the case when $K = \mathbb{C}$, we recover the Möbius group, the group of Möbius transformations.

Perhaps the most interesting property of $\mathrm{PGL}_n(K)$ is the following:

Theorem 12.2. *Let* $p_1, p_2, \dots, p_{n+2}$ *and* $q_1, q_2, \dots, q_{n+2}$ *be two sets of* $n+2$ *in* $\mathbb{P}^n$ *in linear general position.*

Then there is a unique element of $\phi \in \mathrm{PGL}_n(K)$ *such that*

$$\phi(p_i) = q_i.$$

Using this, in the case $n = 1$, we can give a synthetic construction of the unique conic through five points p_1, p_2, p_3, p_4 and p_5 in linear general position, which is known as the *Steiner construction*. Fix two points $p = p_1$ and $q = p_2$. Consider the set of lines through p.

Definition 12.3. *Suppose that* $\mathbb{P}^n = \mathbb{P}(V)$. *Then* $\hat{\mathbb{P}^n} = \mathbb{P}(V^*)$ *is called the* ***dual projective space***.

The whole point of $\hat{\mathbb{P}^n}$ is that it parametrises hyperplanes in $\mathbb{P}^n$. Indeed an element of V^* is a linear functional on V. Its zero locus is a hyperplane in V and this defines a hyperplane in $\mathbb{P}^n$. Conversely a hyperplane in $\mathbb{P}^n$ corresponds to a hyperplane in V. This defines a linear functional on V, up to scalars, that is, an element of $\mathbb{P}(V^*)$.

Another way of putting this is as follows. Pick coordinates $X_0, X_1, \dots, X_n$ on V. These form a basis of V^*. A general element of V^* is then of the form

$$a_0X_0 + a_1X_1 + \cdots + a_nX_n,$$

and its zero locus is a hyperplane in $\mathbb{P}^n$.

Lemma 12.4. *Let* $\Lambda \subset \mathbb{P}^n$ *be a linear subspace of* $\mathbb{P}^n$ *of dimension* k.

Then the set of linear spaces Γ *of dimension* $k+1$ *(or of dimension* $n-1$*) containing* Λ *is a copy of projective space of dimension* $n-k-1$.

Proof. We will give three different proofs of this result and we will also show that these two cases are duals of each other.

The first is geometric. Pick Λ' a complimentary linear subspace (that is, Λ' has the property that it is disjoint from Λ and of maximal dimension with this property). Then Λ' is of dimension $n-k-1$, so that it is a copy of projective space of dimension $n-k-1$.

Claim 12.5. *The points of* Λ' *are in bijection with linear spaces of dimension* $k+1$ *containing* Λ.

Proof of (12.5). One direction is clear. Given a point of Λ' the span of this point and Λ is a linear space of dimension $k+1$ containing Λ.

On the other hand, a linear space Γ of dimension $k+1$ containing Λ must meet Λ' in a unique point. Indeed the dimension of the intersection of Γ and Λ' is at least zero. On the other hand, if it were positive dimensional, then there would be a line l in the intersection. This line is contained in Γ and Λ is a hyperplane in Γ, so that l and Λ must meet in a point. But this contradicts the fact that Λ and Λ' are disjoint. □

The second is algebraic. Pick coordinates so that Λ is given as $Z_{k+1} = \dots Z_n = 0$. Then a hyperplane containing Λ is given by an equation of the form

$$a_{k+1}Z_{k+1} + \cdots + a_nZ_n = 0.$$

Thus the set of hyperplanes containing Λ is naturally in bijection with $\mathbb{P}^{n-k-1}$, with coordinates $[a_{k+1} : a_{k+2} : \cdots : a_n]$.

The third uses a little linear algebra. Suppose that $\mathbb{P}^n = \mathbb{P}(V)$. Then V has dimension $n+1$ and $\Lambda = \mathbb{P}(W)$, where W is of dimension

$k+1$. Suppose that $\Gamma = \mathbb{P}(U)$. Then the set of U containing W is in bijection with the set of U' of dimension one in V/W. But the latter is by definition $\mathbb{P}(V/W)$ and as V/W has dimension $n-k$, the result follows. By duality, hyperplanes in $\mathbb{P}(V/W)$ correspond to lines in $\mathbb{P}(V^*/W^*)$ and so the two results are indeed dual. $\square$

Thus the set of lines H_p through p is naturally a copy of $\mathbb{P}^1$. Similarly for the set H_q of lines through q. Choose parametrisations L_t and M_t of these set of lines. Formally we pick isomorphisms $\mathbb{P}^1 \longrightarrow H_p$ and $\mathbb{P}^1 \longrightarrow H_q$. The two lines L_t and M_t intersect in a point p_t. Varying t, the locus of points p_t sweeps out a curve, call it C. First note that C contains p and q, provided that the line $\langle p, q\rangle$ does not correspond to the same parameter value (we will check later that our choice of parametrisations satisfies this condition).

Note that we have three degrees of freedom left. Indeed we may choose our parametrisation of H_p so that $t=0$ corresponds to the line $\langle p, p_3\rangle$, $t=1$ to $\langle p, p_4\rangle$ and $t=\infty$ to $\langle p, p_5\rangle$, using (12.2). Similarly for H_q. It follows then that C_t passes through p_3, p_4 and p_5.

It remains to check that C is a conic. There are two ways to see this. The first is by direct computation. If L_t is given by $aX+bY+cZ$ and M_t is given by $dX+eY+fZ$ then the point of intersection of L_t and M_t may be determined as follows. Let A be a square $n \times n$ matrix and let B be the adjugate matrix (the tranpose of the matrix of $n-1 \times n-1$ minors). Then

$$AB = BA = (\det A) I_n.$$

Now let

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 0 \end{pmatrix}.$$

Then $\det A = 0$ and computing BA we see that the last column of B gives the intersection point of L_t and M_t. So this point has coordinates which are quadratic in a-f. These are in turn linear in S and T, so we get three quadratic polynomials F, G and H. C is then the image of the morphism

$$\mathbb{P}^1 \longrightarrow \mathbb{P}^2,$$

given by

$$[S:T] \longrightarrow [F:G:H].$$

It is now easy to see that C is a conic:

Lemma 12.6. *Let $C \subset \mathbb{P}^d$ be the image of a morphism*

$$[S:T] \longrightarrow [F_0 : F_1 : \dots : F_d],$$

where $F_0, F_1, \dots, F_d$ have degree at most d.

If C is not contained in a hyperplane then C is projectively equivalent to a rational normal curve of degree d.

Proof. Note that C is not contained in a hyperplane if and only if $F_0, F_1, \dots, F_d$ are linearly independent. Since the space of homogeneous polynomials of degree d in S and T has dimension $d+1$, it follows that the polynomials $F_0, F_1, \dots, F_d$ are a basis for the homogeneous polynomials of degree d. But then we may find a linear transformation taking $S^i T^{d-i}$ to F_i, that is an element of $\mathrm{GL}(d)$, and this defines a projective equivalence with the rational normal curve of degree d. □

Note that in our case if C were to lie in a hyperplane then it would be a line, which is not the case, since C contains p_1, p_2, p_3, p_4 and p_5 and any three of these points are not collinear.

Actually there is another way to check that F, G and H have degree 2. The basic idea is that to find the degree of a curve C, just intersect with a typical line L. The number of points $|C \cap L|$ will just be the degree of the curve. In fact if the line is given by $aX + bY + cZ$ then we just need to find the solutions to the equation

$$aF + bG + cH = 0.$$

If F, G and H have degree d then this equation ought to have d solutions. Borrowing a result from later in the course, in fact we can always choose L with the property that $|L \cap C| = d$ (this is equivalent to saying that not every line is a tangent line).

So pick a typical line L (in particular a line that does not contain any of p_1, p_2, p_3, p_4 and p_5). We calculate $|C \cap L|$.

Since the set L is in (non-canonical) bijection with H_p and H_q, a moments thought will convince the reader that picking L determines an automorphism $\phi \in \mathrm{PGL}(2)$ (indeed send $L_t \cap L$ to $M_t \cap L$) and we want to calculate the number of fixed points of ϕ .

Lemma 12.7. *Let $\phi \in \mathrm{PGL}(2)$.*

Then ϕ is conjugate to

1. *The identity,*
2. *$z \longrightarrow az$, some $a \in K^*$,*
3. *$z \longrightarrow z + 1$.*

Moreover the three cases are determined by the number of fixed points; at least three; two, one.

Thus the degree d of F, G and H is at most 2. If $d = 1$ then C is a line, a contradiction. Thus $d = 2$ and C is a conic.

Now suppose that the line $\langle p, q\rangle$ corresponds to the same parameter value. In this case $d \leq 1$ since one of the fixed points of ϕ corresponds to $L \cap \langle p, q\rangle$, a contradiction. Another way to proceed, which generalises better to higher dimensions, is to consider the line $H = \langle p_3, p_4\rangle$. We get an automorphism of this line, by sending the point $L_t \cap H$ to $M_t \cap H$. This automorphism would have three fixed points, p_3, p_4 and $H \cap \langle p, q\rangle$. But then this automorphism would be the identity. This can only happen if $C = H$ and so p_5 would also lie on H, a contradiction.

This result has the following interesting generalisation:

Theorem 12.8. *Let $p_1, p_2, \dots, p_{n+3}$ be $n+3$ points in linear general position in $\mathbb{P}^n$.*

Then there is a unique rational normal curve through these points.

Proof. We will do the case of $n = 3$ (the general case is no harder, just notationally more involved). Let l be the line $\langle p_1, p_2\rangle$, m be the line $\langle p_2, p_3\rangle$ and n the line $\langle p_1, p_3\rangle$. The set of planes that contains l, H_l, is a copy of $\mathbb{P}^1$. Pick three parametrisations of the three copies of $\mathbb{P}^1$, H_l, H_m and H_n. We choose these parametrisations subject to the condition that the plane spanned by p_1, p_2 and p_3 corresponds to three different parameter values. Given $t \in \mathbb{P}^1$, the three planes corresponding to t intersect in a point, and so we get a curve C in $\mathbb{P}^3$.

Once again we have three degrees of freedom. We may choose our parametrisations, so that $t = 0$ corresponds to the three planes $\langle l, p_4\rangle$, $\langle m, p_4\rangle$ and $\langle n, p_4\rangle$. In this way, we may pick C so that it contains the six points $p_1, p_2, \dots, p_6$ (we check our non-degenerary condition at the end).

It remains to check that C is a twisted cubic. As before we could use the adjugate matrix to conclude that C is the image of

$$[S : T] \longrightarrow [F : G : H : I],$$

where F, G, H and I all have degree three and then we just apply (12.6).

Instead, we could also the geometric argument. As before, it suffices to check that C meets a general plane P in three points. We use the same argument. The planes P_t and Q_t intersect P in a single point x. Similarly the planes Q_t and R_t intersect P in a single point y. H intersects C at the point corresponding to t iff $x = y$. The assignment $x \longrightarrow y$ is an automorphism of P, a copy of $\mathbb{P}^2$, and any automorphism of $\mathbb{P}^2$, not equal to the identity, can have at most three fixed points.

Now suppose that $P_{t_0} = Q_{t_0}$ (necessarily $\langle p_1, p_2, p_3\rangle$), for some t_0. Consider the line $L = \langle p_4, p_5\rangle$. The automorphism given by sending

$P_t \cap L$ to $Q_t \cap L$ would have three fixed points, p_3, p_4 and $L \cap \langle p_1, p_2, p_3 \rangle$. But then p_6 must also lie on L, a contradiction. $\square$

13. Morphisms between varieties II

Consider the following very classical problem.

Determine all triples of integers a, b and c such that $a^2 + b^2 = c^2$.

One way to solve this problem is to use some geometry. Consider the circle with equation $x^2 + y^2 = 1$. Pick a point of this conic, say $(0, 1)$ in the ordinary plane. Consider picking a line l passing through p. This line will intersect the circle at one further point, say q. Let us find the coordinates of q.

Now the general line through $(0, 1)$ is of the form

$$y - 1 = -tx,$$

for some $t \in K$, at least if the line is not vertical. Substituting into

$$x^2 + y^2 = 1,$$

we get

$$x^2 + (-tx + 1)^2 = 1,$$

so that

$$(t^2 + 1)x^2 - 2tx = 0.$$

Thus either $x = 0$, the solution we already have, or

$$(t^2 + 1)x - 2t = 0,$$

so that

$$x = \frac{2t}{t^2 + 1}.$$

In this case

$$y = -tx + 1 = \frac{1 - t^2}{t^2 + 1}.$$

In this way, we get a morphism

$$f \colon \mathbb{A}^1 - \{\pm i\} \longrightarrow \mathbb{A}^2,$$

where

$$t \longrightarrow \left(\frac{2t}{t^2 + 1}, \frac{1 - t^2}{t^2 + 1} \right)$$

Note that we can reverse this process. That is, we can start with a point (x, y) of C and obtain a point of the x-axis, simply by projection. In fact this map is defined for any point away from the line $y = 1$. If we have a point (x, y) then we send this to a point z, where the three points $(0, 1)$, (x, y) and $(z, 0)$ are collinear. Now the reciprocal of the slope of the line connecting $(0, 1)$ to $(z, 0)$ is $-z$, so that

$$\frac{x}{y - 1} = -z.$$

Thus the map is

$$\mathbb{A}^2 - \{y = 1\} \longrightarrow \mathbb{A}^1,$$
$$(x, y) \longrightarrow \frac{x}{1-y}.$$

When we restrict to C we get a morphism outside of $(0, 1)$. What happens when we projectivise?

In this case, $x = X/Z$, $y = Y/Z$ and, at least symbolically we get

$$[X : Y : Z] \longrightarrow [1 : \frac{X/Z}{1 - Y/Z}] = [1 : \frac{X}{Z - Y}].$$

Now note that, at least outside the locus $Z = Y$,

$$[1 : \frac{X}{Z - Y}] = [Z - Y : X].$$

So that it makes sense to extend the map by the rule

$$[X : Y : Z] \longrightarrow [Z - Y : X].$$

This gives us a well-defined map, except on the locus $Z = Y$, $X = 0$, that is the point $[0 : 1 : 1]$. The key point to observe is that even though there is no way to extend this morphism, which is defined on $\mathbb{P}^2 - [0 : 1 : 1]$, to the whole of $\mathbb{P}^2$ (geometrically this is clear, since we would somehow be picking out one line from amongst all lines through $[0 : 1 : 1]$, and the natural symmetry says we cannot do this), in fact it does make sense to extend this map to the whole of $C = V(X^2 + Y^2 - Z^2)$. Again geometrically this is clear. In the limit as y tends to one, the line tends to a horizontal line, a line of zero slope. In fact this is equally clear algebraically. On the curve $X^2 + Y^2 = Z^2$, so that

$$(Z - Y)(Z + Y) = X^2.$$

Given this, on the locus where $Z + Y \neq 0$ and $X \neq 0$,

$$\begin{aligned} [Z - Y : X] &= [(Z + Y)(Z - Y) : (Z + Y)X] \\ &= [X^2 : (Y + Z)X] \\ &= [X : Y + Z]. \end{aligned}$$

Thus we could have defined the morphism equally well by using $[X : Y + Z]$. Note this gives us what we expect geometrically; the point $[0 : 1 : 1]$ would be sent to the point $[0 : 2] = [0 : 1]$, using the last prescription.

In other words, a priori, the map

$$C - \{[0 : 1 : 1]\} \longrightarrow \mathbb{P}^2,$$

given as

$$[X : Y : Z] \longrightarrow [Z - Y : X],$$

would not seem to extend to $[0 : 1 : 1]$. In fact, using our working definition, it does not. The point is, then, that our working definition is not quite right. Let us think carefully through how we want to change our definition, so as to add this map as a morphism. We are given $X \subset \mathbb{P}^n$ and a morphism defined on two "large" subsets of X, a la working definition. That is we are given $U \subset X$, and $F_0, F_1, \dots, F_m$ without common zeroes on U, similarly $V \subset X$, with $G_0, G_1, \dots, G_m$ without common zeroes on V. On the intersection, we have

$$[F_0 : F_1 : \dots : F_m] = [G_0 : G_1 : \dots : G_m]$$

so that they represent the same function on the intersection $U \cap V$. We would then like to extend the morphism to the whole of $U \cup V$, definining the map piecewise.

Definition 13.1. *A **quasi-projective variety** is a locally closed subset of $\mathbb{P}^n$.*

Example 13.2. *Every affine variety $V \subset \mathbb{A}^n$ is a quasi-projective variety. Indeed V is closed subset of $\mathbb{A}^n$ so that it is the intersection of the closure W of V with the open set $U_0 \simeq \mathbb{A}^n$. Similarly, every quasi-affine variety is a quasi-projective variety.*

Conversely, note that if $V \subset \mathbb{P}^n$ is a quasi-projective variety then $V_i = V \cap U_i \subset U_i \simeq \mathbb{A}^n$ is a quasi-affine variety.

Definition 13.3. *Let*

$$f \colon V \longrightarrow W,$$

be a map between two quasi-projective varieties $V \subset \mathbb{P}^m$ and $W \subset \mathbb{P}^n$. Let V_α and W_i be the quasi-affine covers defined above and let

$$U_{\alpha,i} = f^{-1}(W_i) \cap V_\alpha.$$

We say that f is a morphism, if the resriction

$$f|_{W_{i,\alpha}} \colon W_{\alpha,i} \longrightarrow U_i$$

is a morphism of quasi-affine varieties.

Note that the coordinates of $f|_{W_{i,\alpha}}$ are regular functions on a quasi-affine variety.

This gives us a category, the category of quasi-projective varieties and morphisms.

We now relate this definition, to our previous working definition. For example, suppose we are given a map between projective varieties

$$f \colon X \longrightarrow Y,$$

$X \subset \mathbb{P}^m$ and $Y \subset \mathbb{P}^n$, which is given by a collection of homogeneous polynomials $F_0, F_1, \dots, F_n$ of the same degree d, which don't vanish simultaneously,

$$[X_0 : X_1 : \dots : X_m] \longrightarrow [F_0 : F_1 : \dots : F_n].$$

On the open subsets $X_\alpha \neq 0$ and $Y_i \neq 0$ this reduces to the map

$$(x_0, x_1, \dots, \hat{x_\alpha}, \dots, x_n) \longrightarrow (f_0, f_1, \dots, \hat{f}_i, \dots, f_n),$$

where

$$x_j = \frac{X_j}{X_\alpha} \qquad \text{and} \qquad f_j = \frac{Y_j}{Y_i}.$$

We note one very curious

Proposition 13.4. *Let $X \subset \mathbb{P}^n$ be a projective variety.*

Then we may find an embedding of X into some $\mathbb{P}^N$ such that X is defined by linear polynomials and quadratic polynomials of rank at most 4.

Lemma 13.5. *The d-uple embedding*

$$\mathbb{P}^n \longrightarrow \mathbb{P}^N,$$

is an isomorphism and a homeomorphism with its image Y, which is a closed subset of $\mathbb{P}^N$ with equations

$$Z_I Z_J = Z_{I'} Z_{J'} \qquad \textit{for all labels} \qquad I + J = I' + J'$$

Proof. The d-uple embedding is certainly a morphism. We already showed that $Y \subset \mathbb{P}^N$ is closed. One can check that Y is cut out by the given equations.

We now try to write down the inverse map. The image Y is contained in the open affine cover $V_i = V_I$, where I ranges over the pure powers, $I = (0, 0, \dots, 0, d, 0, \dots 0)$ (that is, $X^I = X_i^d$). The inverse image is then the open affine U_i, $X_i \neq 0$, and we get morphisms

$$f_i \colon U_i \longrightarrow V_i.$$

We write down do the inverse image in the case of f_0, so that we want to define a map $g_0 \colon V_0 \longrightarrow U_0$ a morphism, which will turn out to be the inverse of f_0. In this case, we send

$$[Z_I] \longrightarrow [Z_{(d,0,0,\dots,0)} : Z_{(d-1,1,0,\dots,0)} : Z_{(d-1,0,1,\dots,0)} : \dots : Z_{(d-1,0,0,\dots,1)}].$$

This map is certainly a morphism. Moreover it is not hard to see that it is the inverse of f_0. By symmetry this gives us maps g_i, inverses of f_i. Since the inverse map $g = f^{-1}$ is unique, provided it exists, compatibility on overlaps is guaranteed. □

Lemma 13.6. *Let X be a projective variety and let F be a homogeneous polynomial of degree d. Then the set*

$$U_F = \{ x \in X \mid F(x) \neq 0 \},$$

is an open affine subset of X.

Proof. Let Y be the image in $\mathbb{P}^N$ of X under the d-uple embedding. By (13.5) Y is isomorphic to X. Suppose that

$$F(X) = \sum_I a_I X^I.$$

Then $F = 0$ corresponds to the locus

$$L = \sum a_I Z_I = 0,$$

which is a hyperplane in $\mathbb{P}^N$. Since the complement of any hyperplane is a copy of affine space (just change coordinates so that the hyperplane is given as $Z_N = 0$) the corresponding subset of Y, U_L, is an affine subset. As U_L is isomorphic to U_F, the result follows. $\square$

Proof of (13.4). Suppose that X is defined by $F_1, F_2, \dots, F_k$. Note that we may assume that $F_1, F_2, \dots, F_k$ all have the same degree. Indeed the vanishing locus of $X_0, X_1, \dots, X_n$ is empty so that the vanishing locus of F and X_0F, X_1F, $\dots$, X_nF coincide. So just top up the degrees until they are all the same.

Now consider the d-uple embedding ν_d. The image of X inside $\nu_d(\mathbb{P}^n)$ is defined by a collection of linear polynomials; if $F_i = \sum a_I X^I$ then we need the linear polynomial $\sum a_I Z_I$. But we already know that $\nu_d(\mathbb{P}^n)$ is defined by quadratic polynomials of the form $Z_I Z_J = Z_{I'} Z_{J'}$, which has rank at most four. $\square$

14. Limits

One of the more interesting notions of category theory, is the theory of limits.

Definition 14.1. *Let $\mathbb{I}$ be a category and let $F\colon \mathbb{I} \longrightarrow \mathcal{C}$ be a functor. A **prelimit** for F is an object L of $\mathcal{C}$, together with morphisms $f_I\colon L \longrightarrow F(I)$, for every object I of $\mathbb{I}$, which are compatible in the following sense: Given a morphism $f\colon I \longrightarrow J$ in $\mathbb{I}$, the following diagram commutes*

$$\begin{tikzcd} L \arrow[r, "f_I"] \arrow[dr, "f_J"'] & F(I) \arrow[d, "F(f)"] \\ & F(J). \end{tikzcd}$$

*The **limit of** F, denoted $L = \varprojlim_{\mathbb{I}} F$ is a prelimit L, which is universal amongst all prelimits in the following sense: Given any prelimit L' there is a unique morphism $g\colon L' \longrightarrow L$, such that for every object I in $\mathbb{I}$, the following diagram commutes*

$$\begin{tikzcd} L' \arrow[r, "g"] \arrow[dr, "f'_I"'] & L \arrow[d, "f_I"] \\ & F(I). \end{tikzcd}$$

Informally, then, if we think of a prelimit as being to the left of every object $F(I)$, then the limit is the furthest prelimit to the right. Note that limits, if they exist at all, are unique, up to unique isomorphism, by the standard argument. Note also that there is a dual notion, the notion of colimits. In this case, F is a contravariant functor and all the arrows go the other way (informally, then, a prelimit is to the right of every object $F(I)$ and a limit is any prelimit which is furthest to the left).

Let us look at some special cases. First suppose we take for $\mathbb{I}$ the category with one object and one morphism. In this case a functor picks out an object. It is clear that in this case the limit is the same object. Similarly for the colimit.

It is in fact more interesting to take for $\mathbb{I}$ the empty category, that is the category with no objects and no morphisms. Then every object is a prelimit and so a limit has the property that every object has a unique map to it. For obvious reasons this is called a **terminal object**. The category (Sets) of sets has as terminal object any set with one object; the category (Vec) of vector spaces any space of dimension zero. The colimit has the property that it has a unique map to every object and it is called an **initial object**. The empty set is an initial object of the category (<u>Sets</u>) of sets; the group with one element is an initial object in the category (<u>Groups</u>) of groups.

At the other extreme one can take the identity functor, so that $\mathbb{I} = \mathcal{C}$. A limit, if it exists at all, is an object to which all other objects map (in a compatible fashion). In the case that a category has an initial object, then the limit of the identity functor is the initial object. Dually, a colimit, if it exists at all, is an object which maps to all other objects. In the case that a category has a terminal object, then the colimit of the identity functor is the terminal object.

Now take as category two objects, with two morphisms (that is, the two identity maps). A functor picks out two objects, call them X and Y. First consider the case of the limit. A prelimit is the data of an object Z, together with a pair or morphisms, $f\colon Z \longrightarrow X$ and $g\colon Z \longrightarrow Y$. This prelimit is a limit if and only if it is universal amongst all such prelimits. That is, suppose we are given two morphisms $f'\colon Z' \longrightarrow X$ and $g'\colon Z' \longrightarrow Y$, then there is a unique induced morphism $h\colon Z' \longrightarrow Z$, such that the following diagram commutes

$$\begin{tikzcd} & & X \\ Z' \arrow[urr, "f'"] \arrow[r, dashed] \arrow[drr, "g'"'] & Z \arrow[ur, "f"'] \arrow[dr, "g"] & \\ & & Y. \end{tikzcd}$$

Dually, consider the case of a colimit, where all the arrows are reversed. A prelimit is the data of an object Z, together with a pair of morphisms, $f\colon X \longrightarrow Z$ and $g\colon Y \longrightarrow Z$. This prelimit is a limit if and only if it is universal amongst all such prelimits. That is, suppose we are given two morphisms $f'\colon X \longrightarrow Z'$ and $g'\colon Y \longrightarrow Z'$, then there is a unique induced morphism $h\colon Z \longrightarrow Z'$, such that the following diagram commutes

$$\begin{tikzcd} X \arrow[dr, "f"'] \arrow[drr, "f'"] & & \\ & Z \arrow[r, dashed] & Z'. \\ Y \arrow[ur, "g"] \arrow[urr, "g'"'] & & \end{tikzcd}$$

Definition 14.2. *Let X and Y be two objects of a category $\mathcal{C}$. The* ***product*** *is the limit and the* **coproduct** *is the colimit, of the functor above.*

The product of two sets is the ordinary cartesian product; the product of two topological spaces is the product of the spaces and so on. The coproduct of two sets is their disjoint union; similarly for topological spaces; the coproduct of two vector spaces is the direct sum; similarly for groups and rings. Note that for groups, rings and vector spaces, the coincidence that the product and coproduct are in fact isomorphic, even though they satisfy two quite different universal properties.

Now let us be a little more ambitious. Take a category with three objects and five morphisms. The two non-trivial morphisms should have the same target, but different domains.

Definition 14.3. *Suppose we are given a diagram*

$$\begin{tikzcd} & Y \arrow[d, "g"] \\ X \arrow[r, "f"] & B. \end{tikzcd}$$

The limit of the corresponding functor, denoted $X \underset{B}{\times} Y$, is known as the ***fibre product*** *or* ***fibre square****.*

As with the definition of the product, there is an accompanying commutative diagram

$$\begin{tikzcd} Z' \arrow[dr, dashed] \arrow[drr, bend left] \arrow[ddr, bend right] & & \\ & Z \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner"] & Y \arrow[d, "g"] \\ & X \arrow[r, "f"] & B. \end{tikzcd}$$

Note that if B is a terminal object, then the fibre product is nothing more than the product.

Lemma 14.4. *The category (Sets) of sets admits fibre products.*

Proof. It is easy to check that

$$X \underset{B}{\times} Y = \{\, (x, y) \in X \times Y \mid f(x) = g(y) \,\},$$

does the trick. $\square$

The fibre product is sometimes also known as the pullback. In other words we think of the morphism

$$X \underset{B}{\times} Y \longrightarrow X,$$

as the pullback of the map $g\colon Y \longrightarrow B$ along the map $f\colon X \longrightarrow B$. In particular the fibre of the former map over the point $x \in X$ is equal to the fibre of the map g over the point $f(x)$.

The dual notion is that of pushout. Basically take the diagram above, flip about the $Y - X$-diagonal and reverse the arrows. Thus if we start with the diagram

$$\begin{tikzcd} B \arrow[r, "g"] \arrow[d, "f"'] & Y, \\ X & \end{tikzcd}$$

the pushout Z enjoys the universal property encoded in the following commutative diagram:

$$\begin{tikzcd} B \arrow[r, "g"] \arrow[d, "f"'] & Y \arrow[d] \arrow[ddr] & \\ X \arrow[r] \arrow[drr] & Z \arrow[dr, dashed] \arrow[ul, phantom, "\ulcorner", very near start] & \\ & & Z'. \end{tikzcd}$$

For example, consider the category of rings. Suppose we are given two ring homomorphisms $A \longrightarrow B$ and $A \longrightarrow C$, and two ring homomorphisms $B \longrightarrow P$ and $C \longrightarrow P$. Then we get a bilinear map $B \times C \longrightarrow P$, using multiplication in P. It is then easy to see that the pushout is the tensor product $B \underset{A}{\otimes} C$.

15. Products and fibre products

Definition 15.1. *Let*

$$\mathbb{P}^m \times \mathbb{P}^n \longrightarrow \mathbb{P}^{mn+m+n},$$

denote the map given by

$$([X_0, X_1, \dots, X_m], [Y_0, Y_1, \dots, Y_n]) \longrightarrow [X_i Y_j].$$

This map is easily seen to be a bijection and the image is a closed subset, defined by the quadratic polynomials

$$Z_{ij} Z_{kl} = Z_{il} Z_{kj},$$

(where of course Z_{ij} corresponds to $X_i Y_j$). The image V is called the **Segre variety**, and we define the product using this map, that is, we are aiming for:

Proposition 15.2. *Let $X \subset \mathbb{P}^m$ and $Y \subset \mathbb{P}^n$. Then the image of $X \times Y$ under the Segre map is the product (in the sense of category theory) of X and Y.*

Lemma 15.3. *The Segre Variety V is the product, in the sense of category theory, of $\mathbb{P}^m$ and $\mathbb{P}^n$.*

Proof. We have to exhibit two morphisms $p\colon V \longrightarrow \mathbb{P}^m$ and $q\colon V \longrightarrow \mathbb{P}^n$ and show that they satisfy the universal property. Fix l and let $U_l \subset V$ be the open subset where at least one of Z_{il} is non-zero. Define a map

$$U_l \longrightarrow \mathbb{P}^m,$$

by sending $[Z_{ij}]$ to $[Z_{il}]$. This is clearly a morphism, and these maps agree on overlaps. Moreover, varying l, the U_l cover V so that we get a morphism on the whole of V.

By symmetry, this gives us two morphisms p and q. Moreover, under the identification of V with $\mathbb{P}^m \times \mathbb{P}^n$, it is clear that p and q are the ordinary projection maps. Since $\mathbb{P}^m \times \mathbb{P}^n$ is a product in the category of sets, given any morphisms $p'\colon Z \longrightarrow \mathbb{P}^m$ and $q'\colon Z \longrightarrow \mathbb{P}^n$, there is an induced unique function

$$f\colon Z \longrightarrow V.$$

It suffices to check that f is a morphism. We check this locally. Let $U_{ij} \subset V$ be the locus where $Z_{ij} \neq 0$. Then U_{ij} corresponds to $U_i \times U_j$. We first check that U_{ij} is isomorphic to $\mathbb{A}^{m+n}$. By symmetry, we may assume that $i = j = 0$. In this case, dehomogenising, the equations of $U = U_{00}$ become

$$z_{ij} = z_{i0} z_{0j} \qquad \text{and} \qquad z_{ij} z_{kl} = z_{il} z_{kj}.$$

Define a morphism

$$\mathbb{A}^{n+m} \longrightarrow U,$$

by the rule

$$(z_{10}, z_{20}, \dots, z_{m0}, z_{01}, z_{02}, \dots z_{0n}) \longrightarrow (z_{i0}z_{0j}).$$

This is clearly a morphism with image U and it is not hard to show that projection on the first $m+n$ factors is the inverse. Thus $U_{ij} \simeq \mathbb{A}^{n+m}$.

It is easy to check that $\mathbb{A}^{m+n}$ is the product of $\mathbb{A}^m$ and $\mathbb{A}^n$. Thus, by the universal property of the product, f_{ij}, the restriction of f to the inverse image of U_{ij}, is a morphism. □

The general case, follows by the same argument, provided we can prove that the image of $X \times Y$ is a closed subset. In other words we have to say something about which subsets of V are closed.

Definition 15.4. *Let $F(X, Y)$ be a polynomial in $X_0, X_1, \dots, X_m$ and $Y_0, Y_1, \dots, Y_n$. We say that $F(X, Y)$ is* ***bihomogeneous*** *of bi-degree (d, e) if it is homogeneous of degree d in the variables $X_0, X_1, \dots, X_m$ and of degree e in the variables $Y_0, Y_1, \dots, Y_n$.*

For example, $X_0Y_1^2 + X_1Y_0Y_1$ is bihomogeneous of bi-degree $(1, 2)$.

Note that the zero locus of a bihomogeneous polynomial is a well-defined subset of the product.

Lemma 15.5. *Let $Z \subset V$ be a subset defined by bihomogeneous polynomials.*

Then Z is a closed subset.

Proof. Topping up the degree, we may as well assume that X is defined by bihomogeneous polynomials F of bi-degree (d, d). It suffices then to prove that there is a polynomial G on $\mathbb{P}^{mn+m+n}$ which pulls back to F. By linearity, it suffices to prove this for monomials. But Z_{ij} pulls back to X_iY_j, and we can clearly build any monomial X^IY^J, as a product of such monomials, provided that X^I and Y^J have the same degree. □

Proof of (15.2). By (15.5) the image of $X \times Y$ is a closed subset of the Segre variety, under the Segre map, and the rest of the the proof goes through as before. □

It is interesting to see what happens in a specific example. Suppose we take the twisted cubic in $\mathbb{P}^3$. This lies in the quadric $XW = YZ$, that is, it lies in the Segre variety. Now it also lies in the quadric $Y^2 - XZ$. Pulling back to $\mathbb{P}^1 \times \mathbb{P}^1$ we get the bihomogeneous polynomial

$$(X_0Y_1)^2 - (X_0Y_0)(X_1Y_0) = X_0(X_0Y_1^2 - X_1Y_0^2).$$

Now the equation $X_0 = 0$ corresponds to a line in the quadric (see below), and what is left defines the twisted cubic. Thus the twisted cubic is defined by a bi-homogeneous polynomial of type $(1, 2)$.

It is also interesting to see what happens to $\{p\} \times \mathbb{P}^1$ and $\mathbb{P}^1 \times \{q\}$. Indeed $[\lambda_0 : \lambda_1] \times [Y_0 : Y_1]$ is sent to $[\lambda_0 Y_0 : \lambda_0 Y_1 : \lambda_1 Y_0 : \lambda_1 Y_1]$, which is the parametric form of a line. In fact the equations of the line are $\lambda_1 X = \lambda_0 Z$ and $\lambda_1 Y = \lambda_0 W$.

Similarly $[X_0 : X_1] \times [\mu_0 : \mu_1]$ is sent to the line $[\mu_0 X_0 : \mu_1 X_0 : \mu_0 X_1 : \mu_1 X_1]$. This line has equations $\mu_1 X = \mu_0 Y$ and $\mu_1 Z = \mu_0 W$.

It follows that the Segre variety $\mathbb{P}^1 \times \mathbb{P}^1 \subset \mathbb{P}^3$ is covered by two 1-parameter families of lines.

There is in fact another way to look at all of this. Let V and W be two vector spaces of dimension two. Consider the natural map

$$V \times W \longrightarrow V \otimes W.$$

This induces a map

$$\mathbb{P}^1 \times \mathbb{P}^1 \longrightarrow \mathbb{P}^3.$$

Let us calculate what this map is in terms of coordinates. A general vector $v \in V$ has the form $v = ae + bf$, where $\{e, f\}$ is a basis of V. Similarly a general vector $w \in W$ is of the form $cg + dh$, where $\{g, h\}$ is a basis of W. Thus the pair (v, w) is sent to

$$\begin{aligned} v \otimes w &= (ae + bf) \otimes (cg + dh) \\ &= ac(e \otimes g) + ad(e \otimes h) + bc(f \otimes g) + bd(f \otimes h). \end{aligned}$$

The induced map is then

$$\begin{aligned} ([v], [w]) &\longrightarrow [ac(e \otimes g) + ad(e \otimes h) + bc(f \otimes g) + bd(f \otimes h)] \\ ([a : b], [c : d]) &\longrightarrow [v \otimes w] = [ac : ad : bc : bd], \end{aligned}$$

which is clearly the Segre map. Thus the Segre variety consists of all tensors of rank one. The two families of lines, are given as $[v] \times \mathbb{P}^1$ and $\mathbb{P}^1 \times [w]$.

Clearly this generalises in an obvious way to the general Segre variety, which is covered by two families of linear spaces. A family of linear spaces of dimension m, parametrised by $\mathbb{P}^n$ and a family of linear spaces of dimension n, parametrised by $\mathbb{P}^m$.

This also sheds some light on the fact that the twisted cubic is not the intersection of two quadrics. First a quick digression on quadrics, the zero locus of a quadratic polynomial. A quadratic polynomial in the variables $X_0, X_1, \dots, X_n$ is the same as an element of $\mathrm{Sym}^2(V^*)$. These can be identified with symmetric bilinear forms. As such any quadratic polynomial (whence a quadric) has a rank. Since the only invariant of a symmetric bilinear form over an algebraically closed field is its rank, quadrics are classified up to projective equivalence by their rank.

It is interesting to see how this works in practice. If $Q \subset \mathbb{P}^n$ is a quadric of rank $r \leq n+1$ then we can always choose coordinates so that Q is given as

$$X_0^2 + X_1^2 + \dots + X_{r-1}^2 = 0.$$

(Actually not quite; in characteristic two one must look at $X_0X_1 + X_1X_2 + \dots$ with a square at the end depending on the parity of r). Now if any variety X is defined by homogeneous polynomials which don't involve the last variable then X contains the point $q = [0 : 0 : \dots : 1]$. In fact if $p \in X$ then the line $\langle p, q \rangle$ is contained in X. More generally, if the equations defining X don't involve the last $n - r$ variables, then X contains the linear space Λ given by $X_0 = X_1 = \dots X_{r-1} = 0$ and if $p \in X$ then so is the linear space $\langle p, \Lambda \rangle$. In this case we say that X is a **cone** over Λ and we call Λ the **vertex**. Note that to specify X, look at the variety Y you get by considering the zero locus of the polynomials in $\mathbb{P}^{r-1}$ and then joining every point of Y to every point of Λ.

Back to quadrics. In $\mathbb{P}^1$ there are two possibilities. If the rank is 1 then we must have X^2 and we get one point. If the rank is 2 we get $X^2 + Y^2$ (or XY in characteristic two) and we get two points. So there are three possibilities for a conic in $\mathbb{P}^2$. If the rank is at most two, we get the cone over one or two points, that is a line (counted twice, as it were) or a pair of lines. Otherwise we get a smooth conic. In $\mathbb{P}^3$ there are four possibilties. A double plane, a pair of planes, a cone (in the classic sense) or the Segre variety.

Back to twisted cubics. If one of the quadrics has maximal rank 4 (or better one of the quadrics in the pencil, which is in fact always true), then it is projectively equivalent to the Segre variety. In this case the other quadric cuts out a curve of bi-degree $(2, 2)$ on $\mathbb{P}^1 \times \mathbb{P}^1$. As the twisted cubic has bi-degree $(1, 2)$, it follows that we get not only the twisted cubic, but a line (something of bi-degree $(1, 0)$), so that the union has bi-degree $(2, 2)$. Now the line is a fibre of one of the rulings, and a general fibre meets the cubic in two points (since a quadratic polynomial has two roots in general).

In fact projecting a curve C of bi-degree (d, e) to either factor defines a morphism $C \longrightarrow \mathbb{P}^1$ which has degree d (respectively e), that is, the typical fibre contains d points (at least in characteristic zero).

Theorem 15.6. *Let $\phi\colon X \longrightarrow B$ and $\psi\colon Y \longrightarrow B$ be two morphisms of quasi-projective varieties.*

Then the set

$$I = \{ (x, y, b) \in X \times Y \mid b = \phi(x) = \psi(y) \} \subset X \times Y \times B,$$

is the fibre product of ϕ and ψ in the category of varieties.

Proof. We have already seen that the two projection maps $p\colon X \times Y \times B \longrightarrow X$ and $q\colon X \times Y \times B \longrightarrow Y$ are morphisms. Suppose we are given two morphisms $f\colon Z \longrightarrow X$ and $g\colon Z \longrightarrow Y$ such that $\phi \circ f = \psi \circ g$. By composition we get two morphisms to B and by the universal property of the products $X \times B$ and $Y \times B$, we get two morphisms $f'\colon Z \longrightarrow X \times B$ and $g'\colon Z \longrightarrow Y \times B$. By the universal property of the product $X \times B \times Y \times B$, there is an induced morphism $Z \longrightarrow X \times Y \times B \times B$. Note that under projection to the last two factors, the image always lies in the diagonal

$$\{ (b, b) \mid b \in B \},$$

which is easily to be a copy of B. So we get a morphism $Z \longrightarrow X \times Y \times B$. By the universal property of the fibre product, the image lands in I.

So the only thing to check is that I is a quasi-projective variety. $\square$

Definition-Lemma 15.7. *Let $\phi\colon X \longrightarrow B$ be a morphism of quasi-projective varieties. The* ***graph*** *of ϕ is the closed set*

$$\Gamma_\phi = \{ (x, b) \mid \phi(x) = b \} \subset X \times B.$$

It is isomorphic to X via the first projection map.

Proof. The only things to check are that Γ_ϕ is closed and the first projection map is an isomorphism. Since we can check this locally, we may assume that X and B are affine. We may then assume that $B = \mathbb{A}^n$ and that ϕ is given as

$$(x_1, x_2, \dots, x_m) \longrightarrow (f_1/g, f_2/g, ..., f_n/g),$$

where g does not vanish on X. In this case the graph is given by the equations

$$gy_i = f_i,$$

where $y_1, y_2, \dots, y_n$ are coordinates on $\mathbb{A}^n$. To see that the first map is an isomorphism, one can use the fact that the graph is in fact the fibre product of the identity $X \longrightarrow X$ and the morphism $\phi\colon X \longrightarrow B$ over B. The inverse map of the first projection is

$$(x_1, x_2, \dots, x_m) \longrightarrow (x_1, x_2, \dots, x_m, f_1/g, f_2/g, ..., f_n/g). \qquad \square$$

Lemma 15.8. *$I \subset X \times Y \times B$ is closed.*

Proof. This is easy. The graphs of ϕ and ψ define two closed subsets of $X \times Y \times B$,

$$\{\,(x, y, b) \mid \phi(x) = b\,\} \qquad \text{and} \qquad \{\,(x, y, b) \mid \psi(y) = b\,\},$$

and I is the intersection. $\square$

16. More examples of schemes

Definition 16.1. *Let X be a scheme and let $x \in X$ be a point of X. The **residue field of** X **at** x is the quotient of $\mathcal{O}_{X,x}$ by its maximal ideal.*

We recall some basic facts about valuations and valuation rings.

Definition 16.2. *Let K be a field and let G be a totally ordered abelian group. A **valuation** of K with values in G, is a map*

$$\nu\colon K - \{0\} \longrightarrow G,$$

such that for all x and $y \in K - \{0\}$ we have:

(1) $\nu(xy) = \nu(x) + \nu(y)$.
(2) $\nu(x + y) \geq \min(\nu(x), \nu(y))$.

Definition-Lemma 16.3. *If ν is a valuation, then the set*

$$R = \{\, x \in K \,|\, \nu(x) \geq 0 \,\} \cup \{0\},$$

*is a subring of K, which is called the **valuation ring** of ν. The set*

$$\mathfrak{m} = \{\, x \in K \,|\, \nu(x) > 0 \,\} \cup \{0\},$$

is an ideal in R and the pair $(R, \mathfrak{m})$ is a local ring.

Proof. Easy check. $\square$

Definition 16.4. *A valuation is called a **discrete valuation** if $G = \mathbb{Z}$ and ν is surjective. The corresponding valuation ring is called a **discrete valuation ring**. Any element $t \in R$ such that $\nu(t) = 1$ is called a **uniformising parameter**.*

Lemma 16.5. *Let R be an integral domain, which is not a field.*

The following are equivalent:

- *R is a DVR.*
- *R is a local ring and a PID.*

Proof. Suppose that R is a DVR. Then R is certainly a local ring. Suppose that a and $b \in R$ and $\nu(a) = \nu(b)$. Then $\nu(b/a) = \nu(b) - \nu(a) = 0$ and so $\langle a \rangle = \langle b \rangle$. It follows that the ideals of R are of the form

$$I_k = \{\, a \in R \,|\, \nu(a) \geq k \,\}.$$

As ν is surjective, there is an element $t \in R$ such that $\nu(t) = 1$. Then

$$I_k = \langle t^k \rangle = \mathfrak{m}^k.$$

Thus R is a PID.

Now suppose that R is a local ring and a PID. Let $\mathfrak{m}$ be the unique maximal ideal. As R is a PID, $\mathfrak{m} = \langle t \rangle$, for some $t \in R$. Define a map

$$\nu\colon K \longrightarrow \mathbb{Z},$$

by sending a to k, where $a \in \mathfrak{m}^k - \mathfrak{m}^{k+1}$ and extending this to any fraction a/b in the obvious way. It is easy to check that ν is a valuation and that R is the valuation ring. □

There are two key examples of a DVR. First let k be field and let $R = k[t]_{\langle t \rangle}$. Then R is a local ring and a PID so that R is a DVR. t is a uniformising parameter. Note that R is the stalk of the struture sheaf of the affine line at the origin.

Now let

$$\Delta = \{\, z \in \mathbb{C} \mid |z| < 1 \,\},$$

be the unit disc in the complex plane. Then the stalk $\mathcal{O}_{\Delta,0}$ of the sheaf of holomorphic functions is a local ring. The order of vanishing realises this ring as a DVR. z is a uniformising parameter.

In fact if C is a smooth algebraic curve, an algebraic variety of dimension one, then $\mathcal{O}_{C,p}$ is a DVR.

Example 16.6. *Let R be the local ring of a curve over an algebraically closed field (or more generally a discrete valuation ring). Then* $\operatorname{Spec} R$ *consists of two points; the maximal ideal, and the zero ideal. The first t_0 is closed and has residue field the groundfield k of C, the second t_0 has residue field the quotient ring K of R, and its closure is the whole of X. The inclusion map $R \longrightarrow K$ corresponds to a morphism which sends the unique point of* $\operatorname{Spec} K$ *to t_1.*

Example 16.7. *There is another morphism of ringed spaces which sends the unique point of* $\operatorname{Spec} K$ *to t_0 and uses the inclusion above to define the map on structure sheaves.*

Since there is only one way to map R to K, this does not come from a map on rings. In fact the second map is not a morphism of locally ringed spaces, and so it is not a morphism of schemes.

It is interesting to see an example of an affine scheme, in a seemingly esoteric case. Consider the case of a number field k (that is, a finite extension of $\mathbb{Q}$, with its ring of integers $A \subset k$, that is, the integral closure of $\mathbb{Z}$ inside k). As a particular example, take

Example 16.8. $k = \mathbb{Q}(\sqrt{3})$. *Then $A = \mathbb{Z} \oplus \mathbb{Z}\langle\sqrt{3}\rangle$. The picture is very similar to the case of $\mathbb{Z}$. There are infinitely many maximal ideals, and only one point which is not closed, the zero ideal. Moreover, as there is a natural ring homomorphism $\mathbb{Z} \longrightarrow A$, by our equivalence* *of categories, there is an induced morphism of schemes* $\operatorname{Spec} A \longrightarrow \operatorname{Spec} \mathbb{Z}$*. We investigate this map. Consider the fibre over a point* $\langle p \rangle \in \operatorname{Spec} \mathbb{Z}$*. This is just the set of primes in* A *containing the ideal* pA*. It is well known by number theorists, that three things can happen:*

(1) *If* p *divides the discriminant of* $k/\mathbb{Q}$ *(which in this case is* 12*), that is,* $p = 2$ *or* 3*, then the ideal* $\langle p \rangle$ *is a square in* A*.*
$$\langle 2 \rangle A = \langle -1 + \sqrt{3} \rangle^2,$$
and
$$\langle 3 \rangle A = \langle \sqrt{3} \rangle^2.$$
(2) *If* 3 *is a square modulo* p*, the prime* $\langle p \rangle$ *factors into a product of distinct primes,*
$$\langle 11 \rangle A = \langle 4 + 3\sqrt{3} \rangle \langle 4 - 3\sqrt{3} \rangle,$$
or
$$\langle 13 \rangle A = \langle 4 + \sqrt{3} \rangle \langle 4 - \sqrt{3} \rangle,$$
(3) *If* $p > 3$ *and* 3 *is not a square mod* p *(e.g* $p = 5$ *and* 7*), the ideal* $\langle p \rangle$ *is prime in* A*.*

Definition 16.9. *Let* $\mathcal{C}$ *be a category and let* X *be an object of* $\mathcal{C}$*. Let* $\mathcal{D} = \mathcal{C}|_X$ *be the category whose objects consist of pairs* $f\colon Y \longrightarrow X$*, where* f *is a morphism of* $\mathcal{C}$*, and whose morphisms, consist of commutative diagrams*

$$\begin{tikzcd} Y \arrow[rr, "\phi"] \arrow[dr, "f"'] & & Z \arrow[dl, "g"] \\ & X. & \end{tikzcd}$$

$\mathcal{D}$ *is known as the category over* X*. If* X *is a scheme, then a scheme over* X *is exactly an object of the category of schemes over* X*. Let* R *be a ring.* ***Affine*** n***-space over*** R*, denoted* $\mathbb{A}^n_R$*, is the spectrum of the polynomial ring* $R[x_1, x_2, \dots, x_n]$*.*

One of the key ideas of schemes, is to work over arbitrary bases. Note that since there is an inclusion $R \longrightarrow R[x_1, x_2, \dots, x_n]$ of rings, affine space over R is a scheme over $\operatorname{Spec} R$. Thus we may define affine space over any affine scheme.

17. First properties of schemes

We start with some basic properties of schemes.

Definition 17.1. *We say that a scheme is* ***connected*** *(respectively* ***irreducible****) if its topological space is connected (respectively irreducible).*

Definition 17.2. *We say that a scheme is* ***reduced*** *if $\mathcal{O}_X(U)$ contains no nilpotent elements, for every open set U.*

Remark 17.3. *It is straightforward to prove that a scheme is reduced if and only if the stalk of the structure sheaf at every point contains no nilpotent elements.*

Definition 17.4. *We say that a scheme X is* ***integral*** *if for every open set $U \subset X$, $\mathcal{O}_X(U)$ is an integral domain.*

Proposition 17.5. *A scheme X is integral if and only if it is irreducible and reduced.*

Proof. Suppose that X is integral. Then X is surely reduced. Suppose that X is reducible. Then we can find two non-empty disjoint open sets U and V. But then

$$\mathcal{O}_X(U \cup V) \simeq \mathcal{O}_X(U) \times \mathcal{O}_X(V),$$

which is surely not an integral domain.

Now suppose that X is reduced and irreducible. Let $U \subset X$ be an open set and suppose that we have f and $g \in \mathcal{O}_X(U)$ such that $fg = 0$. Set

$$Y = \{ x \in U \mid f_x \in m_x \} \qquad \text{and} \qquad Z = \{ x \in U \mid g_x \in m_x \}.$$

Then Y and Z are both closed and by assumption $Y \cup Z = U$. As X is irreducible, one of Y and Z is the whole of U, say Y. We may assume that $U = \operatorname{Spec} A$ is affine. But then $f \in A$ belongs to the intersection of all the prime ideals of A, which is the zero ideal, as A contains no nilpotent elements. □

Definition 17.6. *We say that a scheme X is* ***locally Noetherian****, if there is an open affine cover, such that the corresponding rings are Noetherian. If in addition the topological space is compact, then we say that X is* ***Noetherian****.*

Remark 17.7. *There are examples of schemes whose topological space is Noetherian which are not Noetherian schemes.*

A key issue in this definition is whether or not we can replace an open cover, by every affine cover.

Proposition 17.8. *A scheme X is locally Noetherian if and only if for every open affine $U = \operatorname{Spec} A$, A is a Noetherian ring.*

Proof. It suffices to prove that if X is locally Noetherian, and $U = \operatorname{Spec} A$ is an open affine subset then A is a Noetherian ring.

We first show that U is locally Noetherian. Suppose that $V = \operatorname{Spec} B$ is an open affine on X where B is a Noetherian ring. Then $U \cap V$ can be covered by open sets of the form $V_f = \operatorname{Spec} B_f$, where $f \in B$. As B is a Noetherian ring then so is B_f. As open sets of the form V cover X, U is covered by open affines, which are the spectra of Noetherian rings. So U is locally Noetherian.

Replacing X by U, we are reduced to proving that if $X = \operatorname{Spec} A$ is locally Noetherian then A is Noetherian. Let $V = \operatorname{Spec} B$, be an open subset of X, where B is a Noetherian ring. Then there is an element $f \in A$ such that $U_f \subset V$. Let g be the image of f in B. As

$$X \supset U_f = U_g \subset V,$$

we have an isomorphism of rings $A_f \simeq B_g$, so that A_f is Noetherian. So we can cover X by open subsets $U_f = \operatorname{Spec} A_f$, with A_f Noetherian. As X is compact, we may assume that we have a finite cover. Now apply (17.9). □

Lemma 17.9. *Let A be a ring, and let $f_1, f_2, \dots, f_r$ be elements of A which generate the unit ideal.*

If A_{f_i} is Noetherian, for $1 \leq i \leq r$ then so is A.

Proof. Suppose that we have an ascending chain of ideals,

$$\mathfrak{a}_1 \subset \mathfrak{a}_2 \subset \mathfrak{a}_3 \subset \dots,$$

of A. Then for each i,

$$\phi_i(\mathfrak{a}_1) \cdot A_{f_i} \subset \phi_i(\mathfrak{a}_2) \cdot A_{f_i} \subset \phi_i(\mathfrak{a}_3) \cdot A_{f_i} \subset \dots,$$

is an ascending chain of ideals inside A_{f_i}, where $\phi_i \colon A \longrightarrow A_{f_i}$ is the natural map. As each A_{f_i} is Noetherian, all of these chains stabilise. But then the first chain stabilises, by (17.10). □

Lemma 17.10. *Let A be a ring, and let $f_1, f_2, \dots, f_r$ be elements of A which generate the unit ideal. Suppose that $\mathfrak{a}$ is an ideal and let $\phi_i \colon A \longrightarrow A_{f_i}$ be the natural maps. Then*

$$\mathfrak{a} = \bigcap_{i=1}^{r} \phi_i^{-1}(\phi_i(\mathfrak{a}) \cdot A_{f_i}).$$

Proof. The fact that the LHS is included in the RHS is clear. Conversely suppose that b is an element of the RHS. In this case

$$\phi_i(b) = \frac{a_i}{f^{n_i}},$$

for some $a_i \in \mathfrak{a}$ and some positive integer n_i. As there are only finitely many indices, we may assume that $n = n_i$ is fixed. But then

$$f^{m_i}(f^n b - a_i) = 0,$$

for $1 \leq i \leq r$. Once again, we may assume that $m = m_i$ is fixed. It follows that $f_i^N b \in \mathfrak{a}$, for $1 \leq i \leq r$, where $N = n + m$. Let I be the ideal generated by the Nth powers of $f_1, f_2, \dots, f_r$. As the radical of I contains 1, I contains 1. Hence we may write

$$1 = \sum_i c_i f_i^N.$$

But then

$$b = \sum_i c_i f_i^N b \in \mathfrak{a}. \qquad \square$$

Definition 17.11. *A morphism $f\colon X \longrightarrow Y$ is **locally of finite type** if there is an open affine cover $V_i = \operatorname{Spec} B_i$ of Y, such that $f^{-1}(V_i)$ is a union of affine sets $U_{ij} = \operatorname{Spec} A_{ij}$, where each A_{ij} is a finitely generated B_i-algebra. If in addition, we can take U_{ij} to be a finite cover of $f^{-1}(V_i)$, then we say that f is of **finite type**.*

Definition 17.12. *We say that a morphism $f\colon X \longrightarrow Y$ is **finite** if we may cover Y by open affines $V_i = \operatorname{Spec} B_i$, such that $f^{-1}(V_i) = \operatorname{Spec} A_i$ is an affine set, where A_i is a finitely generated B_i-module.*

In both cases, it is straightforward to prove that we can take V_i to be any affine subset of Y.

Example 17.13. *Let*

$$f\colon \mathbb{A}^1_k - \{0\} \longrightarrow \mathbb{A}^1_k,$$

by the natural map given by the natural localisation map

$$k[x] \longrightarrow k[x]_x.$$

As an algebra over $k[x]$, the ring $k[x]_x \simeq k[x, x^{-1}]$ is generated by x^{-1}, so that f is of finite type. However the $k[x]$-module $k[x, x^{-1}]$ is not finitely generated (there is no way to generate all the negative powers of x), so that f is not finite.

Definition 17.14. *Let X be a scheme and let U be an open subset of X. Then the pair $(U, \mathcal{O}_U = \mathcal{O}_X|_U)$ is a scheme, which is called an* ***open subscheme*** *of X. An* ***open immersion*** *is a morphism $f\colon X \longrightarrow Y$ which induces an isomorphism of X with an open subset of Y.*

Definition 17.15. *A* ***closed immersion*** *is a morphism of schemes $\phi = (f, f^{\#})\colon Y \longrightarrow X$ such that f induces a homeomorphism of Y with a closed subset of X and futhermore the map $f^{\#}\colon \mathcal{O}_X \longrightarrow f_*\mathcal{O}_Y$ is surjective. A* ***closed subscheme*** *of a scheme X is an equivalence class of closed immersions, where we say that two closed immersions $f\colon Y \longrightarrow X$ and $f'\colon Y' \longrightarrow X$ are equivalent if there is an isomorphism $i\colon Y' \longrightarrow Y$ such that $f' = f \circ i$.*

Despite the seemingly tricky nature of the definition of a closed immersion, in fact it is easy to give examples of closed subschemes of an affine variety.

Lemma 17.16. *Let A be a ring and let $\mathfrak{a}$ be an ideal of A. Let $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} A/\mathfrak{a}$.*

Then Y is a closed subscheme of X.

Proof. The quotient map map $A \longrightarrow A/\mathfrak{a}$ certainly induces a morphism of schemes $\phi\colon Y \longrightarrow X$. f is certainly a homeomorphism of Y with $V(\mathfrak{a})$ and $f^{\#}\colon \mathcal{O}_X \longrightarrow f_*\mathcal{O}_Y$ is surjective as the map on stalks is induced by the quotient map, which is surjective. □

In fact, it turns out that every closed subscheme of an affine scheme is of this form. It is interesting to look at some examples.

Example 17.17. *Let $X = \mathbb{A}^2_k$. First consider $\mathfrak{a} = \langle y^2 \rangle$. The support of Y is the x-axis. However the scheme Y is not reduced, even though it is irreducible.*

It is clear from this example that in general there are many closed subschemes with the same support (equivalently there are many ideals with the same radical). For example,

Example 17.18. *consider the ideals $\langle x^2, xy, y^2 \rangle$, the double of the maximal ideal of a point and the ideal $\langle x, y^2 \rangle$. They both defines subschemes of $\mathbb{A}^2_k$ supported at the origin.*

Finally consider:

Example 17.19. *$\langle x^2, xy \rangle$. The support of this ideal is the y-axis. But this time the only local ring which has nilpotents is the local ring of the origin.*

We call the origin an **embedded point.**

Definition 17.20. *Let V be an irreducible affine variety with coordinate ring A and let W be a closed irreducible subvariety, defined by the prime ideal $\mathfrak{p}$. Then we can associate two affines schemes $Y \subset X$ to $W \subset V$. Let $X = \operatorname{Spec} A$ and define Y by $\mathfrak{p}$. The n**th infinitessimal neighbourhood of** Y **in** X, denoted Y_n, is the closed subscheme of X corresponding to $\mathfrak{p}^n$.*

Note that the nth infinitessimal neighbourhood of Y in X is a closed subscheme whose support coincides with Y, but whose structure sheaf contains lots of nilpotent elements. As the name might suggest, Y_n carries more information about how Y sits inside X, than does Y itself.

Note that if a scheme X has a topological space with one point, then X must be affine, and the stalk of the structure sheaf at the unique point completely determines X, and this ring has exactly one prime ideal. Moreover a morphism of X into another scheme Y, is equivalent to picking a point y of Y and a morphism of local rings

$$\mathcal{O}_{Y,y} \longrightarrow \mathcal{O}_{X,x}.$$

But to give a morphism of local rings is the same as to give an inclusion of the quotients of the maximal ideals. Thus to give a morphism of $X = \{x\}$ into Y, sending x to y, we need to specify an inclusion of the residue field of x into the residue field of y.

18. Fibre products of schemes

The main result of this section is:

Theorem 18.1. *The category of schemes admits fibre products.*

A key part of the proof is to pass from the local case (in which case all three schemes are affine) to the global case. To do this, we need to be able to construct morphisms, by constructing them locally. We will need:

Theorem 18.2. *Let $f_i \colon U_i \longrightarrow Y$ be a collection of morphisms of schemes, with a varying domain, but a fixed target.*

Suppose that for each pair of indices i and j, we are given open subsets $U_{ij} \subset U_i$, and isomorphisms $\phi_{ij} \colon U_{ij} \longrightarrow U_{ji}$, such that $f_i|_{U_{ij}} = f_j \circ \phi_{ij} \colon U_{ij} \longrightarrow Y$ and

$$\phi_{ik} = \phi_{jk} \circ \phi_{ij},$$

on the intersection $U_{ij} \cap U_{ik}$, for all i, j and k (we adopt the convention that $U_{ii} = U_i$, so that ϕ_{ii} is the identity and moreover $\phi_{ij}^{-1} = \phi_{ji}$).

Then there is a morphism of schemes $f \colon X \longrightarrow Y$, open immersions $\psi_i \colon U_i \longrightarrow X$, whose images cover X, such that $f_i = f \circ \psi_i \colon U_i \longrightarrow Y$ and $\psi_i|_{U_{ij}} = \psi_j \circ \phi_{ij} \colon U_{ij} \longrightarrow Y$.

X is unique, up to unique isomorphism, with these properties.

We prove (18.2) in two steps (one of which can be further broken down into two substeps):

- Construct the scheme X.
- Construct the morphism f.

In fact, having constructed X, it is straightforward to construct f. Since a scheme consists of two parts, a topological space and a sheaf, we can break the first step into two smaller pieces:

- Construct the underlying topological space.
- Construct the structure sheaf.

We first show how to patch a sheaf, which is the hardest part:

Lemma 18.3. *Let X be a topological space, and let $\{X_i\}$ be an open cover of X. Suppose that we are given sheaves $\mathcal{F}_i$ on X_i and for each i and j an isomorphism*

$$\phi_{i,j} \colon \mathcal{F}_i|_{X_{ij}} \longrightarrow \mathcal{F}_j|_{X_{ij}},$$

such that

$$\phi_{ik} = \phi_{jk} \circ \phi_{ij},$$

on the triple intersection X_{ijk}, for all i, j and k.

Then there is a sheaf $\mathcal{F}$ on X, together with isomorphisms,

$$\psi_i \colon \mathcal{F}|_{X_i} \longrightarrow \mathcal{F}_i,$$

which satisfy $\psi_j = \phi_{ij} \circ \psi_i$. Further $\mathcal{F}$ is unique up to unique isomorphism, with these properties.

Proof. We just show how to define $\mathcal{F}$ and leave the rest to the interested reader. Let $U \subset X$ be any open set, and let $U_i = U \cap X_i$.

$$\mathcal{F}(U) = \{ (s_i) \in \prod_i \mathcal{F}_i(U_i) \mid \phi_{ij}(s_i|_{U_{ij}}) = s_j|_{U_{ji}} \}. \qquad \square$$

Using (18.3), one can put a natural scheme structure on any closed subset of a scheme (natural means the smallest possible scheme structure):

Definition-Lemma 18.4. *Let X be scheme and let Y be a closed subset. Then Y has a unique reduced subscheme structure, called the* ***reduced induced subscheme structure****.*

Proof. We first assume that $X = \operatorname{Spec} A$ is affine. Let $\mathfrak{a}$ be the ideal obtained by intersecting all the prime ideals in Y. Then $\mathfrak{a}$ is the largest ideal for which $V(\mathfrak{a}) = Y$. The induced scheme structure on Y is reduced, that is, the stalks of $\mathcal{O}_Y$ have no nilpotent elements, as $\mathfrak{a}$ is a radical ideal.

Now suppose that X is an arbitrary scheme. For each open affine subset $U_i \subset X$, let $Y_i \subset U_i$ be the reduced induced subscheme structure on $Y \cap U_i$. This gives us a sheaf $\mathcal{O}_{Y_i}$ on each Y_i and we want to construct a sheaf $\mathcal{O}_Y$ on the whole of Y. By (18.3) it suffices to prove that the sheaves $\mathcal{O}_{Y_i}$ agree on overlaps.

It is not hard to reduce to the case where $U = \operatorname{Spec} A$, $V = \operatorname{Spec} A_f$. We want to show that the reduced induced subscheme structure on V is the same as restricting the reduced induced subscheme structure from U to V. But this is the same as to say that if $\mathfrak{a}$ is the intersection of those prime ideals of A which are contained in Y, then $\mathfrak{a}A_f$ is the intersection of those prime ideals of A_f which are contained in Y, which is clear. $\square$

The next step is to bump this up to schemes:

Lemma 18.5. *Suppose that we are given schemes U_i, and subschemes $U_{ij} \subset U_i$, together with isomorphisms,*

$$\phi_{ij} \colon U_{ij} \longrightarrow U_{ji},$$

which satisfy

$$\phi_{ik} = \phi_{jk} \circ \phi_{ij},$$

on the intersection $U_{ij} \cap U_{ik}$, for all i, j and k.

Then there is a scheme X and open immersions $\psi_i \colon U_i \longrightarrow X$, whose images cover X, which satisfy $\psi_i|_{U_{ij}} = \psi_i \circ \phi_{ij} \colon U_{ij} \longrightarrow X$.

Proof. We first construct the topological space X. Let

$$X = \coprod_i U_i / \sim \qquad \text{where} \qquad x_i \in U_{ij} \sim x_j \in U_{ji} \text{ iff } \phi_{ij}(x_i) = x_j.$$

Here $\sim$ denotes the equivalence relation generated by the rule on the RHS, and X is just the quotient topological space (which always exists). Note that

$$X_i = U_i / \sim,$$

is an open subset of X and there are homeomorphisms $\phi_i \colon U_i \longrightarrow X_i$. Now construct a sheaf $\mathcal{O}_X$ on X, using (18.3). This gives us a locally ringed space $(X, \mathcal{O}_X)$ and the remaining properties can be easily checked. □

There are a couple of interesting examples of the construction of schemes. The first is to take U_{ij} empty (so that there are no patching conditions at all). The resulting scheme is called the **disjoint union** and is denoted

$$\coprod_i X_i.$$

Another more interesting example proceeds as follows. Take two copies U_1 and U_2 of the affine line. Let $U_{12} = U_{21}$ be the complement of the origin, and let ϕ_{12} be the identity. Then X is obtained by identifying every point, except the origin. Note that this is like the classical construction of a topological space, which is locally a manifold, but which is not Hausdorff. Of course no scheme is ever Hausdorff (apart from the most trivial examples) and it turns out that there is an appropriate condition for schemes (and in fact morphisms of schemes) which is a replacement for the Hausdorff condition for topological spaces.

Finally we turn to the problem of glueing morphisms, which is the easiest bit:

Proof of (18.2). Let X be the scheme constructed in (18.5). It is clear that to give a morphism $f \colon X \longrightarrow Y$ is the same as to give morphisms $f_i \colon X_i \longrightarrow Y$, compatible on overlaps. □

Lemma 18.6. *Let X and Y be schemes over S. Suppose that X has an open cover $\{X_i\}$ such that the fibre product $F_i = X_i \underset{S}{\times} Y$ exists.*

Then the fibre product $F = X \underset{S}{\times} Y$ exists.

Proof. Let $p_i\colon F_i \longrightarrow X$ be the natural morphism and let $F_{ij} = p_i^{-1}(X_j)$. Note that F_{ij} is isomorphic to the fibre product of $X_i \cap X_j$ and Y over S. Indeed if Z maps to X_{ij} and Y over S, it maps to X_i and Y over S. But then Z maps to F_i, by the universal property of the fibre product. It is clear that the image of Z lands in F_{ij}, so that F_{ij} is the fibre product. But then there are natural isomorphisms $\phi_{ij}\colon F_{ij} \longrightarrow F_{ji}$ such that

$$\phi_{ik} = \phi_{jk} \circ \phi_{ij},$$

on the intersection $F_{ij} \cap F_{ik}$, for all i, j and k, and $p_i|_{F_{ij}} = p_j \circ \phi_{ij}$.

(18.2) implies that we may patch F_i to a scheme F, and patch the morphisms $F_i \longrightarrow X$ to a morphism $F \longrightarrow X$. Similarly we may construct a morphism $F \longrightarrow Y$, from the individual morphisms $q_i\colon F_i \longrightarrow Y$.

Now suppose we are given $Z \longrightarrow X$ and $Z \longrightarrow Y$ morphisms over S. The open cover $\{X_i\}$ induces an open cover $\{Z_i\}$ of Z. We get morphisms $Z_i \longrightarrow F_i$, by the universal property of F_i and so we get morphisms $Z_i \longrightarrow F$ by composition. It is easy to check that these patch to a morphism $Z \longrightarrow F$. But then F is the fibre product. $\square$

Proof of (18.1). Let X and Y be two schemes over S. We want to construct the fibre product.

First suppose that $X = \operatorname{Spec} A$, $Y = \operatorname{Spec} B$ and $S = \operatorname{Spec} R$. Then there are ring homomorphisms $R \longrightarrow A$ and $R \longrightarrow B$ and so A and B are R-algebras. As $C = A \underset{R}{\otimes} B$ is the pushout in the category of rings, it follows that $Z = \operatorname{Spec} C$ is the fibre product in the category of affine schemes; in fact it is also the fibre product in the category of schemes, since a morphism to an affine scheme is the same as a ring homomorphism the other way on global sections.

We now bump this result up to the global case. First suppose that S and Y are affine. Since an arbitrary scheme X can be covered by open affines $\{X_i\}$, (18.6) implies that the fibre product of X and Y over S exists.

Now suppose that S is affine. Since Y can be covered by open affines $\{Y_j\}$ and the fibre product is obviously symmetric in X and Y, (18.6) implies that the fibre product of X and Y over S exists.

Now take an affine cover S_i of S. Let X_i and Y_i be the inverse image of S_i (meaning take the open subscheme on the open set $p_j^{-1}(S_i)$). Then the fibre product $X_i \underset{S}{\times} Y_i$ exists. But in fact this is also a fibre product for $X_i \underset{S}{\times} Y$, since anything lying over X_i automatically lies over Y_i. Since X_i forms an open cover of X we are done by one more application of (18.6). $\square$