Massachusetts Institute of Technology

Notes for 18.721

INTRODUCTION TO ALGEBRAIC GEOMETRY

(This is a preliminary draft. Please don't reproduce.)

.

TABLE OF CONTENTS

Chapter 1 PLANE CURVES

Plane curves were the first algebraic varieties to be studied, so we begin with them. They provide helpful examples, and we will see in Chapter 5 how they control varieties of arbitrary dimension. Chapters 2 - 7 are about varieties of arbitrary dimension. We come back to curves in Chapter 8.

1.1 The Affine Plane

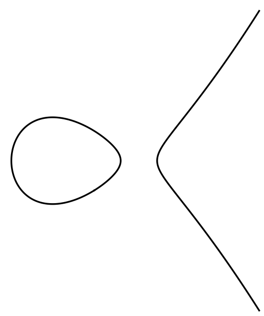
The n-dimensional *affine space* $\mathbb{A}^n$ is the space of n-tuples of complex numbers. The *affine plane* $\mathbb{A}^2$ is the two-dimensional affine space.

Let $f(x_1, x_2)$ be an irreducible polynomial in two variables with complex coefficients. The set of points of the affine plane at which f vanishes, the *locus of zeros* of f, is called a *plane affine curve*. Let's denote this locus by X. Using vector notation $x = (x_1, x_2)$,

(1.1.1) $$X = \{x \mid f(x) = 0\}$$

The *degree* of the curve X is the degree of its irreducible defining polynomial f.

1.1.2.



The Cubic Curve $y^2 = x^3 - x$ **(real locus)**

1.1.3. Note. In contrast with polynomials in one variable, most complex polynomials in two or more variables are *irreducible* – they cannot be factored. This can be shown by a method called "counting constants". For instance, quadratic polynomials in x_1, x_2 depend on the six coefficients of the monomials of degree at most two. Linear polynomials ax_1+bx_2+c depend on three coefficients, but the product of two linear polynomials depends on only five parameters, because a scalar factor can be moved from one of the linear polynomials to the other. So the quadratic polynomials cannot all be written as products of linear polynomials. This reasoning is fairly convincing. It can be justified formally in terms of *dimension*, which will be discussed in Chapter 4.□

We will get an understanding of the geometry of a plane curve as we go along, and we mention just one important point here. A plane curve is called a *curve* because it is defined by one equation in two variables. Its *algebraic* dimension is one. But because our scalars are complex numbers, it will be a surface, geometrically. This is analogous to the fact that the *affine line* $\mathbb{A}^1$ is the plane of complex numbers.

One can see that a plane curve X is a surface by inspecting its projection to the affine x_1-line $\mathbb{A}^1$. One writes the defining polynomial as a polynomial in x_2, whose coefficients $c_i = c_i(x_1)$ are polynomials in x_1:

$$f(x_1, x_2) = c_0 x_2^d + c_1 x_2^{d-1} + \cdots + c_d$$

Let's suppose that d is positive, i.e., that f isn't a polynomial in x_1 alone (in which case, since it is irreducible, it would be linear).

The *fibre* of a map $X \to Z$ over a point p of Z is the inverse image of p, the set of points of X that map to p. The fibre of the projection $X \to \mathbb{A}^1$ over the point $x_1 = a$ is the set of points (a, b) such that b is a root of the one-variable polynomial

$$f(a, x_2) = \overline{c}_0 x_2^d + \overline{c}_1 x_2^{d-1} + \cdots + \overline{c}_d$$

with $\overline{c}_i = c_i(a)$. There will be finitely many points in this fibre, and the fibre won't be empty unless $f(a, x_2)$ is a constant. So the curve X covers most of the x_1-line, a complex plane, finitely often.

(1.1.4) changing coordinates

We allow linear changes of variable and translations in the affine plane $\mathbb{A}^2$. When a point x is written as the column vector $(x_1, x_2)^t$, the coordinates $x' = (x_1', x_2')^t$ after such a change of variable will be related to x by the formula

(1.1.5)
$$x = Qx' + a$$

where Q is an invertible 2×2 matrix with complex coefficients and $a = (a_1, a_2)^t$ is a complex translation vector. This changes a polynomial equation $f(x) = 0$, to $f(Qx' + a) = 0$. One may also multiply a polynomial f by a nonzero complex scalar without changing the locus $\{f = 0\}$. Using these operations, all *lines*, plane curves of degree 1, become equivalent.

An *affine conic* is a plane affine curve of degree two. Every affine conic is equivalent to one of the loci

(1.1.6)
$$x_1^2 - x_2^2 = 1 \quad \text{or} \quad x_2 = x_1^2$$

The proof of this is similar to the one used to classify real conics. The two loci might be called a complex 'hyperbola' and 'parabola', respectively. The complex 'ellipse' $x_1^2 + x_2^2 = 1$ becomes the 'hyperbola' when one multiplies x_2 by i.

On the other hand, there are infinitely many inequivalent cubic curves. Cubic polynomials in two variables depend on the coefficients of the ten monomials $1, x_1, x_2, x_1^2, x_1x_2, x_2^2, x_1^3, x_1^2x_2, x_1x_2^2, x_2^3$ of degree at most 3 in x. Linear changes of variable, translations, and scalar multiplication give us only seven scalars to work with, leaving three essential parameters.

1.2 The Projective Plane

The n-dimensional *projective space* $\mathbb{P}^n$ is the set of equivalence classes of *nonzero* vectors $x = (x_0, x_1, ..., x_n)$, the equivalence relation being

(1.2.1) $$(x'_0,...,x'_n) \sim (x_0,...,x_n) \quad \text{if} \quad (x'_0,...,x'_n) = (\lambda x_0,...,\lambda x_n)$$

for some nonzero complex number λ. The equivalence classes are the *points* of $\mathbb{P}^n$, and one often refers to a point by a particular vector in its class.

Points of $\mathbb{P}^n$ correspond bijectively to one-dimensional subspaces of $\mathbb{C}^{n+1}$. When x is a nonzero vector, the vectors λx, together with the zero vector, form the one-dimensional subspace of the complex vector space $\mathbb{C}^{n+1}$ spanned by x.

The *projective plane* $\mathbb{P}^2$ is the two-dimensional projective space. Its points are equivalence classes of nonzero vectors (x_0,x_1,x_2).

(1.2.2) the projective line

Points of the *projective line* $\mathbb{P}^1$ are equivalence classes of nonzero vectors (x_0,x_1). If x_0 isn't zero, we may multiply by $\lambda = x_0^{-1}$ to normalize the first entry of (x_0,x_1) to 1, and write the point it represents in a unique way as $(1,u)$, with $u = x_1/x_0$. There is one remaining point, the point represented by the vector $(0,1)$. The projective line $\mathbb{P}^1$ can be obtained by adding this point, called the *point at infinity*, to the affine u-line, which is a complex plane. Topologically, $\mathbb{P}^1$ is a two-dimensional sphere.

(1.2.3) lines in projective space

A *line* in projective space $\mathbb{P}^n$ is determined by a pair of distinct points p and q. When p and q are represented by specific vectors, the set of points $\{rp+sq\}$, with r,s in $\mathbb{C}$ not both zero is a line L. Points of L correspond bijectively to points of the projective line $\mathbb{P}^1$, by

(1.2.4) $$rp+sq \quad \longleftrightarrow \quad (r,s)$$

A line in the projective plane $\mathbb{P}^2$ can also be described as the locus of solutions of a homogeneous linear equation

(1.2.5) $$s_0x_0+s_1x_1+s_2x_2=0$$

1.2.6. Lemma. *In the projective plane, two distinct lines have exactly one point in common and, in a projective space of any dimension, a pair of distinct points is contained in exactly one line.* □

(1.2.7) the standard covering of $\mathbb{P}^2$

If the first entry x_0 of a point $p = (x_0,x_1,x_2)$ of the projective plane $\mathbb{P}^2$ isn't zero, we may normalize it to 1 without changing the point: $(x_0,x_1,x_2) \sim (1,u_1,u_2)$, where $u_i = x_i/x_0$. We did the analogous thing for $\mathbb{P}^1$ above. The representative vector $(1,u_1,u_2)$ is uniquely determined by p, so points with $x_0 \neq 0$ correspond bijectively to points of an affine plane $\mathbb{A}^2$ with coordinates (u_1,u_2):

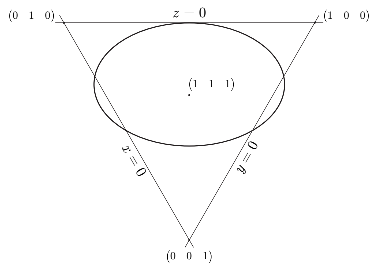
$$(x_0,x_1,x_2) \sim (1,u_1,u_2) \quad \longleftrightarrow \quad (u_1,u_2)$$

We regard the affine plane as a subset of $\mathbb{P}^2$ by this correspondence, and we denote that subset by $\mathbb{U}^0$. The points of $\mathbb{U}^0$, those with $x_0 \neq 0$, are the *points at finite distance*. The *points at infinity* of $\mathbb{P}^2$, those of the form $(0,x_1,x_2)$, are on the *line at infinity* L^0, the locus $\{x_0 = 0\}$. The projective plane is the union of the two sets $\mathbb{U}^0$ and L^0. When a point is given by a coordinate vector, we can assume that the first coordinate is either 1 or 0.

When looking at a point of $\mathbb{U}^0$, we may simply set $x_0 = 1$, and write the point as $(1,x_1,x_2)$. To write $u_i = x_i/x_0$ makes sense only when a particular vector (x_0,x_1,x_2) has already been given.

There is an analogous correspondence between points $(x_0,1,x_2)$ and points of an affine plane $\mathbb{A}^2$, and between points $(x_0,x_1,1)$ and points of $\mathbb{A}^2$. We denote the subsets $\{x_1 \neq 0\}$ and $\{x_2 \neq 0\}$ by $\mathbb{U}^1$ and $\mathbb{U}^2$, respectively. The three sets $\mathbb{U}^0, \mathbb{U}^1, \mathbb{U}^2$ form the *standard covering* of $\mathbb{P}^2$ by three *standard affine open sets*. Since the vector $(0,0,0)$ has been ruled out, every point of $\mathbb{P}^2$ lies in at least one of the standard affine open sets. Points whose three coordinates are nonzero lie in all of them.

1.2.8.



A Schematic Representation of the Projective Plane, with a Conic

This figure shows the plane W: $x+y+z=1$ in the real projective space $\mathbb{RP}^3$. If $p=(x,y,z)$ is a nonzero vector, the one-dimensional subspace of $\mathbb{R}^3$ spanned by p will meet W in a single point unless p is on the line $x+y+z=0$. So the plane is a faithful representation of most of $\mathbb{RP}^2$.

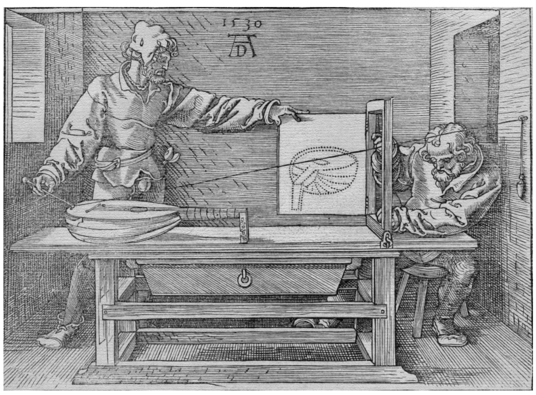
1.2.9. Note. Which points of $\mathbb{P}^2$ are at infinity depends on which of the standard affine open sets is taken to be the one at finite distance. When the coordinates are (x_0, x_1, x_2), I like to normalize x_0 to 1, as above. Then the points at infinity are those of the form $(0, x_1, x_2)$. But when coordinates are (x,y,z), I may normalize z to 1. Then the points at infinity are the points $(x,y,0)$. I hope this won't cause too much confusion. $\square$

(1.2.10) digression: the real projective plane

The points of the *real projective plane* $\mathbb{RP}^2$ are equivalence classes of nonzero real vectors $x=(x_0,x_1,x_2)$, the equivalence relation being $x' \sim x$ if $x' = \lambda x$ for some nonzero real number λ. The real projective plane can also be thought of as the space of one-dimensional subspaces of the real vector space $V=\mathbb{R}^3$.

The plane $U : \{x_0 = 1\}$ in the space V is analogous to the standard affine open subset $\mathbb{U}^0$ in the complex projective plane $\mathbb{P}^2$. We can project V from the origin $p_0 = (0,0,0)$ to U, sending a point $x=(x_0,x_1,x_2)$ of V distinct from p_0 to the point $(1,u_1,u_2)$, with $u_i = x_i/x_0$. The fibres of this projection are the lines through p_0 and x, with p_o omitted. Looking from the origin, U becomes a "picture plane".

1.2.11.



This illustration is from Dürer's book on perspective

The projection to U is undefined at the points $(0,x_1,x_2)$, which are orthogonal to the x_0-axis. The line connecting such a point to p_0 doesn't meet U. Those points correspond to the points at infinity of $\mathbb{RP}^2$.

The projection from 3-space to a picture plane goes back to the the 16th century, the time of Desargues and Dürer. Projective coordinates were introduced by Möbius, 200 years later.

(1.2.12) changing coordinates in the projective plane

An invertible 3×3 matrix P determines a linear change of coordinates in $\mathbb{P}^2$. With $x = (x_0, x_1, x_2)^t$ and $x' = (x'_0, x'_1, x'_2)^t$ represented as column vectors, the coordinate change is given by

$$Px' = x \tag{1.2.13}$$

As the next proposition shows, four special points, the three points $e_0 = (1, 0, 0)^t, e_1 = (0, 1, 0)^t, e_2 = (0, 0, 1)^t$, together with the point $\epsilon = (1, 1, 1)^t$, determine the coordinates.

1.2.14. Proposition. *Let p_0, p_1, p_2, q be four points of $\mathbb{P}^2$, no three of which lie on a line. There is, up to scalar factor, a unique linear coordinate change $Px' = x$ such that $Pp_i = e_i$ and $Pq = \epsilon$.*

proof. The hypothesis that the points p_0, p_1, p_2 don't lie on a line means that the vectors that represent those points are independent. They span $\mathbb{C}^3$. So q will be a combination $c_0p_0 + c_1p_1 + c_2p_2$, and because no three points lie on a line, the coefficients c_i will be nonzero. We can *scale* the vectors p_i (multiply them by nonzero scalars) to make $q = p_0+p_1+p_2$ without changing the points. Next, the columns of P can be an arbitrary set of independent vectors. We let them be p_0, p_1, p_2. Then $Pe_i = p_i$, and $P\epsilon = q$. The matrix P is unique up to scalar factor, as can be verified by looking the reasoning over. □

(1.2.15) conics

A polynomial $f(x_0, x_1, x_2)$ is *homogeneous* , and *of degree* d, if all monomials that appear with nonzero coefficient have (total) degree d. For example, $x_0^3 + x_1^3 - x_0x_1x_2$ is a homogeneous cubic polynomial.

A homogeneous quadratic polynomal is a combination of the six monomials

$$x_0^2,\ x_1^2,\ x_2^2,\ x_0x_1,\ x_1x_2,\ x_0x_2$$

A *conic* is the locus of zeros of an irreducible homogeneous quadratic polynomial.

1.2.16. Proposition. *For any conic C, there is a choice of coordinates so that C becomes the locus*

$$x_0x_1 + x_0x_2 + x_1x_2 = 0$$

proof. Since the conic C isn't a line, it will contain three points that aren't colinear. Let's leave the verification of this fact as an exercise. We choose three non-colinear points on C, and adjust coordinates so that they become the points e_0, e_1, e_2. Let f be the quadratic polynomial in those coordinates whose zero locus is C. Because e_0 is a point of C, $f(1, 0, 0) = 0$, and therefore the coefficient of x_0^2 in f is zero. Similarly, the coefficients of x_1^2 and x_2^2 are zero. So f has the form

$$f = ax_0x_1 + bx_0x_2 + cx_1x_2$$

Since f is irreducible, a, b, c aren't zero. By scaling the variables appropriately, we can make $a = b = c = 1$. We will be left with the polynomial $x_0x_1 + x_0x_2 + x_1x_2$. □

1.3 Plane Projective Curves

The loci in projective space that are studied in algebraic geometry are those that can be defined by systems of *homogeneous* polynomial equations. The reason for homogeneity is that the vectors $(a_0, ..., a_n)$ and $(\lambda a_0, ..., \lambda a_n)$ represent the same point of $\mathbb{P}^n$. One wants to know that if $f(x) = 0$ is a polynomial equation, and if $f(a) = 0$, then $f(\lambda a) = 0$ for every $\lambda \neq 0$. As we verify now, this will be true if and only if f is homogeneous.

A polynomial f can be written as a sum of its *homogeneous parts*:

$$f = f_0 + f_1 + \cdots + f_d \tag{1.3.1}$$

where f_0 is the constant term, f_1 is the linear part, etc., and d is the degree of f.

1.3.2. Lemma. *Let f be a polynomial of degree d, and let $x = (x_0, ..., x_n)$ be a particular nonzero vector. Then $f(\lambda x) = 0$ for every nonzero complex number λ if and only if $f_i(x)$ is zero for every $i = 0, ..., d$.*

proof. $f(\lambda x_0, ..., \lambda x_n) = f_0 + \lambda f_1(x) + \lambda^2 f_2(x) + \cdot + \lambda^d f_d(x)$. When we evaluate at a given vector x, the right side of this equation becomes a polynomial of degree at most d in λ. Since a nonzero polynomial of degree at most d has at most d roots, $f(\lambda x)$ won't be zero for *every* λ unless that polynomial is zero – unless $f_i(x) = 0$ for every i. □

1.3.3. Lemma. *If a homogeneous polynomial f is a product gh of polynomials, then g and h are homogeneous, and the zero locus $\{f = 0\}$ in projective space is the union of the two loci $\{g = 0\}$ and $\{h = 0\}$.* □

It is also true that relatively prime homogeneous polynomials f and g have only finitely many common zeros. This isn't obvious. It will be proved below, in Proposition 1.3.11.

(1.3.4) loci in the projective line

Before going to plane curves, we describe the zero locus in the projective line $\mathbb{P}^1$ of a homogeneous polynomial in two variables.

1.3.5. Lemma. *Every nonzero homogeneous polynomial $f(x,y) = a_0x^d + a_1x^{d-1}y + \cdots + a_dy^d$ with complex coefficients is a product of homogeneous linear polynomials that are unique up to scalar factor.*

To prove this, one uses the fact that the field of complex numbers is algebraically closed. A one-variable complex polynomial factors into linear factors in the polynomial ring $\mathbb{C}[y]$. To factor $f(x,y)$, one may factor the one-variable polynomial $f(1,y)$ into linear factors, substitute y/x for y, and multiply the result by x^d. When one adjusts scalar factors, one will obtain the expected factorization of $f(x,y)$. For instance, to factor $f(x,y) = x^2 - 3xy + 2y^2$, substitute $x = 1$: $2y^2 - 3y + 1 = 2(y-1)(y-\frac{1}{2})$. Substituting $y = y/x$ and multiplying by x^2, $f(x,y) = 2(y-x)(y-\frac{1}{2}x)$. The scalar 2 can be distributed arbitrarily among the linear factors. □

Adjusting scalar factors, we may write a homogeneous polynomial as a product of the form

$$f(x,y) = (v_1x - u_1y)^{r_1} \cdots (v_kx - u_ky)^{r_k} \tag{1.3.6}$$

where no factor $v_ix - u_iy$ is a constant multiple of another, and where $r_1 + \cdots + r_k$ is the degree of f. The exponent r_i is the *multiplicity* of the linear factor $v_ix - u_iy$.

A linear polynomial $vx - uy$ determines a point (u,v) in the projective line $\mathbb{P}^1$, the unique *zero* of that polynomial, and changing the polynomial by a scalar factor doesn't change its zero. Thus the linear factors of the homogeneous polynomial (1.3.6) determine points of $\mathbb{P}^1$, the *zeros* of f. The points (u_i, v_i) are zeros of *multiplicity* r_i. The total number of those points, counted with multiplicity, will be the degree of f.

The zero (u_i, v_i) of f corresponds to a root $x = u_i/v_i$ of multiplicity r_i of the one-variable polynomial $f(x,1)$, except when the zero is the point $(1,0)$. This happens when the coefficient a_0 of f is zero, and y is a factor of f. One could say that $f(x,y)$ has a zero at infinity in that case.

This sums up the information contained in an algebraic locus in the projective line. It will be a finite set of points with multiplicities.

(1.3.7) intersections with a line

Let Z be the zero locus of a homogeneous polynomial $f(x_0, ..., x_n)$ of degree d in projective space $\mathbb{P}^n$, and let L be a line in $\mathbb{P}^n$ (1.2.4). Say that L is the set of points $rp + sq$, where $p = (a_0, ..., a_n)$ and $q = (b_0, ..., b_n)$ are represented by specific vectors, so that L corresponds to the projective line $\mathbb{P}^1$ by $rp + sq \leftrightarrow (r,s)$. Let's also assume that L isn't entirely contained in the zero locus Z. The intersection $Z \cap L$ corresponds to the zero locus in $\mathbb{P}^1$ of the polynomial in r,s that is obtained by substituting $rp + sq$ into f. This substitution yields a homogeneous polynomial $\overline{f}(r,s)$ of degree d in r,s. For example, if $f = x_0x_1 + x_0x_2 + x_1x_2$, then with p and q as above, $\overline{f}$ is the following quadratic polynomial in r,s:

$$\overline{f}(r,s) = f(rp+sq) = (ra_0+sb_0)(ra_1+sb_1) + (ra_0+sb_0)(ra_2+sb_2) + (ra_1+sb_1)(ra_2+sb_2)$$
$$= (a_0a_1{+}a_0a_2{+}a_1a_2)r^2 + \big(\textstyle\sum_{i\neq j} a_ib_j\big)rs + (b_0b_1{+}b_0b_2{+}b_1b_2)s^2$$

The zeros of $\overline{f}$ in $\mathbb{P}^1$ correspond to the points of $Z \cap L$. There will be d zeros, when counted with multiplicity. □

1.3.8. Definition. With notation as above, the *intersection multiplicity* of Z and L at a point p is the multiplicity of zero of the polynomial $\overline{f}$. □

1.3.9. Corollary. *Let Z be the zero locus in $\mathbb{P}^n$ of a homogeneous polynomial f, and let L be a line in $\mathbb{P}^n$ not contained in Z. The number of intersections of Z and L, counted with multiplicity, is equal to the degree of f.* □

(1.3.10) loci in the projective plane

1.3.11. Proposition. *Homogeneous polynomials $f_1, ..., f_r$ in three variables x, y, z that have no common factor have finitely many common zeros.*

As this shows, the most interesting type of locus in the projective plane is the zero set of a single equation. The proof of the proposition is below.

The locus of zeros of an irreducible homogeneous polynomial f is called a *plane projective curve*. The *degree* of a plane projective curve is the degree of its irreducible defining polynomial.

1.3.12. Note. Suppose that a homogeneous polynomial is reducible, say $f = g_1 \cdots g_k$, where g_i are irreducible and distinct (i.e., g_i and g_j don't differ by a scalar factor when $i \neq j$). Then the zero locus C of f is the union of the zero loci V_i of the factors g_i. In this case, C may be called a *reducible curve*.

When there are multiple factors, say $f = g_1^{e_1} \cdots g_k^{e_k}$ and some e_i are greater than 1, it is still true that the locus $C : \{f = 0\}$ will be the union of the loci $V_i : \{g_i = 0\}$, but the connection between the geometry of C and the algebra of f is weakened. In this situation, the structure of a *scheme* becomes useful. We won't discuss schemes. The only situation in which we will need to keep track of multiple factors is when counting intersections with another curve D. For this purpose, one can define the *divisor* of f to be the integer combination $e_1V_1 + \cdots + e_kV_k$. □

We need a lemma for the proof of Proposition 1.3.11. The ring $\mathbb{C}[x, y]$ embeds into its field of fractions F, which is the field of rational functions $\mathbb{C}(x, y)$ in x, y. The polynomial ring $\mathbb{C}[x, y, z]$ is a subring of the one-variable polynomial ring $F[z]$. It can be useful to study a problem in $F[z]$ first because $F[z]$ is a principal ideal domain. Its algebra is simpler.

Recall that the *unit ideal* of a ring R is the ring R itself.

1.3.13. Lemma. *Let $f_1, ..., f_k$ be homogeneous polynomials in x, y, z with no common factor. Their greatest common divisor in $F[z]$ is 1, and therefore $f_1, ..., f_k$ generate the unit ideal of $F[z]$. There is an equation of the form $\sum g_i' f_i = 1$ with g_i' in $F[z]$.*

proof. **(i)** Let h' be an element of $F[z]$ that isn't a unit of $F[z]$, i.e., that isn't an element of F, and suppose that, for every i, h' divides f_i in $F[z]$, say $f_i = u_i'h'$. The coefficients of h' and u_i' have denominators that are polynomials in x, y. We clear denominators from their coefficients, to obtain elements of $\mathbb{C}[x, y, z]$. This will give us equations of the form $d_if_i = u_ih$, where d_i are polynomials in x, y and u_i, h are polynomials in x, y, z.

Since h isn't in F, it will have positive degree in z. Let g be an irreducible factor of h of positive degree in z. Then g divides d_if_i but doesn't divide d_i which has degree zero in z. So g divides f_i, and this is true for every i. This contradicts the hypothesis that $f_1, ..., f_k$ have no common factor. □

proof of Proposition 1.3.11.

We are to show that homogeneous polynomials $f_1, ..., f_r$ in x, y, z with no common factor have finitely many common zeros. Lemma 1.3.13 tells us that we may write $\sum g'_i f_i = 1$, with g'_i in $F[z]$. Clearing denominators from g'_i gives us an equation of the form

$$\sum g_i f_i = d$$

where d is a polynomial in x, y and g_i are polynomials in x, y, z. Taking suitable homogeneous parts of d and g_i produces an equation $\sum g_i f_i = d$ in which all terms are homogeneous.

Lemma 1.3.5 asserts that d is a product of linear polynomials, say $d = \ell_1 \cdots \ell_r$. A common zero of $f_1, ..., f_k$ is also a zero of d, and therefore it is a zero of ℓ_j for some j. It suffices to show that, for every j, $f_1, ..., f_r$ and ℓ_j have finitely many common zeros.

Since $f_1, ..., f_k$ have no common factor, there is at least one f_i that isn't divisible by ℓ_j. Corollary 1.3.9 shows that f_i and ℓ_j have finitely many common zeros. Therefore $f_1, ..., f_k$ have finitely many common zeros for every j. □

1.3.14. Corollary. *Every locus in the projective plane $\mathbb{P}^2$ that can be defined by a system of homogeneous polynomial equations is a finite union of points and curves.* □

The next corollary is a special case of the Strong Nullstellensatz, which will be proved in the next chapter.

1.3.15. Corollary. *Let f be an irreducible homogeneous polynomial in x, y, z that vanishes on an infinite set S of points of $\mathbb{P}^2$. If another homogeneous polynomial g vanishes on S, then f divides g. Therefore, if an irreducible polynomial vanishes on an infinite set S, that polynomial is unique up to scalar factor.*

proof. If the irreducible polynomial f doesn't divide g, then f and g have no common factor, and therefore they have finitely many common zeros. □

(1.3.16) the classical topology

The usual topology on the affine space $\mathbb{A}^n$ will be called the *classical topology*. A subset U of $\mathbb{A}^n$ is open in the classical topology if, whenever U contains a point p, it contains all points sufficiently near to p. We call this the classical topology to distinguish it from another topology, the *Zariski topology*, which will be duscussed in the next chapter.

The projective space $\mathbb{P}^n$ also has a classical topology. A subset U of $\mathbb{P}^n$ is open if, whenever a point p of U is represented by a vector $(x_0, ..., x_n)$, all vectors $x' = (x'_0, ..., x'_n)$ sufficiently near to x represent points of U.

(1.3.17) isolated points

A point p of a topological space X is *isolated* if both $\{p\}$ and its complement $X - \{p\}$ are closed sets, or if $\{p\}$ is both open and closed. If X is a subset of $\mathbb{A}^n$ or $\mathbb{P}^n$, a point p of X is isolated in the classical topology if X doesn't contain points p' distinct from p, but arbitrarily close to p.

The proof of the next proposition is below.

1.3.18. Proposition *Let n be an integer greater than one. The zero locus of a polynomial in $\mathbb{A}^n$ or in $\mathbb{P}^n$ contains no points that are isolated in the classical topology.*

An element $f(x_1, ..., x_n)$ of the polynomial ring $\mathbb{C}[x_1, ..., x_n]$ is a *formal polynomial*, but the next lemma shows that we needn't be careful to distinguish formal polynomials from polynomial functions.

1.3.19. Lemma. *A formal polynomial $f(x_1, ..., x_n)$ is determined by the function that it defines on $\mathbb{A}^n$.*

proof. We are to show that if two formal polynomials f and g define the same function, they are equal. We replace f by $f - g$. Then what we must show is that, if a formal polynomial f defines the zero function, then

$f = 0$. We use induction on the number n of variables. We label the variable x_n as y, so that our formal polynomial becomes

$$f(x_1, ..., x_{n-1}, y) = c_k y^k + c_{k-1} y^{k-1} + \cdots + c_0$$

where c_i are polynomials in $x_1, ..., x_{n-1}$ and c_k isn't the zero polynomial. By induction, c_k doesn't define the zero function. We choose $a_1, ..., a_{n-1}$ so that $c_k(a) \neq 0$. Then $f(a_1, ..., a_{n-1}, y)$ is a nonzero polynomial in y of degree k, that has at most k zeros. Choosing b so that $f(a_1, ..., a_{n-1}, b)$ isn't zero shows that the function defined by f isn't identically zero. □

1.3.20. Lemma. *Let f be a polynomial of degree d in the variables $x_1, ..., x_n$. There is a linear change of variable $Px' = x$, where P is an invertible $n \times n$ matrix, such that $f(Px')$ is a monic polynomial of degree d in the variable x'_n.*

proof. We write $f = f_0 + f_1 + \cdots + f_d$, where f_i is the homogeneous part of f of degree i, and we choose a point p of $\mathbb{A}^n$ at which f_d isn't zero. We change variables so that p becomes the point $(0, ..., 0, 1)$. We call the new variables $x_1, .., .x_n$ and the new polynomial f. Then $f_d(0, ..., 0, x_n)$ will be equal to cx_n^d for some nonzero constant c. When we adjust x_n by a scalar factor to make $c = 1$, f will be monic. □

proof of Proposition 1.3.18. The proposition is true for loci in affine space and also for loci in projective space. We look at the affine case. Let $f(x_1, ..., x_n)$ be a polynomial with zero locus Z, and let p be a point of Z. We adjust coordinates so that p is the origin $(0, ..., 0)$ and f is monic in x_n. We relabel x_n as y, and write f as a polynomial in y. Let's write $f(x, y) = \widetilde{f}(y)$:

$$\widetilde{f}(y) = f(x, y) = y^d + c_{d-1}(x) y^{d-1} + \cdots + c_0(x)$$

where c_i is a polynomial in $x_1, ..., x_{n-1}$. For fixed x, $c_0(x)$ is the product of the roots of $\widetilde{f}(y)$. Since p is the origin and $f(p) = 0$, $c_0(0) = 0$. So $c_0(x)$ will tend to zero with x. Then at least one root y of $\widetilde{f}(y)$ will tend to zero. This gives us points (x, y) of Z that are arbitrarily close to p. □

1.3.21. Corollary. *Let C' be the complement of a finite set of points in a plane curve C. In the classical topology, a continuous function g on C that is zero at every point of C' is identically zero.* □

1.4 Tangent Lines

(1.4.1) homogenizing and dehomogenizing

We will often want to inspect a plane curve $C : \{f(x_0, x_1, x_2) = 0\}$ in a neighborhood of a particular point p. To do this we may adjust coordinates so that p becomes the point $(1, 0, 0)$, and look in the standard affine open set $\mathbb{U}^0 : \{x_0 \neq 0\}$. The intersection C^0 of C with $\mathbb{U}^0$ will be the zero locus of the non-homogeneous polynomial $f(1, x_1, x_2)$, and p will be the origin in the affine x_1, x_2-plane. The polynomial $f(1, x_1, x_2)$ is the *dehomogenization* of f.

A simple procedure, *homogenization*, inverts dehomogenization. Suppose given a non-homogeneous polynomial $F(x_1, x_2)$ of degree d. To *homogenize* F, we replace the variables x_i, $i = 1, 2$, by $u_i = x_i/x_0$. Then since u_i have degree zero in x, so does $F(u_1, u_2)$. When we multiply by x_0^d, the result will be a homogeneous polynomial of degree d in x_0, x_1, x_2 not divisible by x_0,

1.4.2. Lemma. *A homogeneous polynomial $f(x_0, x_1, x_2)$ not divisible by x_0 is irreducible if and only if its dehomogenization $f(1, x_1, x_2)$ is irreducible.* □

We will come back to homogenization in Chapter 2.

(1.4.3) smooth points and singular points

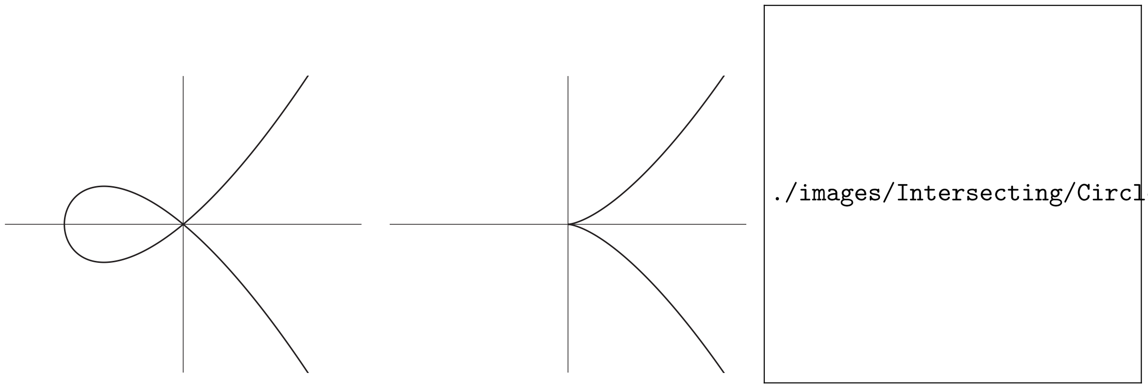
Let C be the plane curve defined by an irreducible homogeneous polynomial $f(x_0, x_1, x_2)$, and let f_i denote the partial derivative $\frac{\partial f}{\partial x_i}$, computed by the usual calculus formula, and a point of C at which the partial derivatives f_i aren't all zero is called a *smooth point* of C, and a point at which all partial derivatives are zero is a *singular point*. A curve is *smooth*, or *nonsingular*, if it contains no singular point. Otherwise it is a *singular* curve.

The *Fermat curve*

(1.4.4) $$x_0^d + x_1^d + x_2^d = 0$$

is smooth because the only common zero of the partial derivatives $dx_0^{d-1}, dx_1^{d-1}, dx_2^{d-1}$, which is $(0,0,0)$, doesn't represent a point of $\mathbb{P}^2$. The cubic curve $x_0^3 + x_1^3 - x_0x_1x_2 = 0$ is singular at the point $(0,0,1)$.

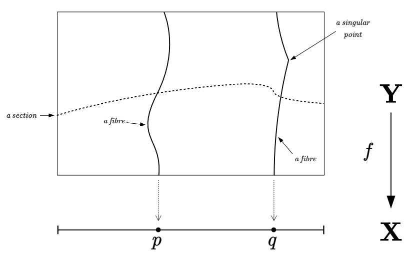
The Implicit Function Theorem explains the meaning of smoothness. Suppose that $p = (1,0,0)$ is a point of C. We set $x_0 = 1$ and inspect the locus $f(1, x_1, x_2) = 0$ in the standard affine open set $\mathbb{U}^0$. If $f_2(p)$ isn't zero, the Implicit Function Theorem tells us that we can solve the equation $f(1, x_1, x_2) = 0$ for x_2 locally as an analytic function φ of x_1. Sending x_1 to $(1, x_1, \varphi(x_1))$ inverts the projection from C to the affine x_1-line, locally. So at a smooth point, C is locally homeomorphic to the affine line.



1.4.5. intersecting the lines $y = 0, 1$.

Note. *(about figures)* In algebraic geometry, dimensions are too big to allow realistic figures. Even with a plane curve, one is dealing with a locus in the space $\mathbb{A}^2$, whose dimension as a real vector space is four. In some cases, such as in the figures above, depicting the real locus can be helpful, but in most cases, even the real locus is too big, and one must make do with a schematic figure. The figure below is an example. My students tell me that all of my figures look more or less like this:

1.4.6.



A Typical Schematic Figure

1.4.7. Euler's Formula. Let f be a homogeneous polynomial of degree d in the variables $x_0, ..., x_n$. Then

$$\sum_i x_i \frac{\partial f}{\partial x_i} = d\, f.$$

proof. It is enough to check this formula when f is a monomial. As an example, let f be the monomial x^2y^3z, then

$$xf_x + yf_y + zf_z = x(2xy^3z) + y(3x^2y^2z) + z(x^2y^3) = 6x^2y^3z = 6\,f \qquad \square$$

1.4.8. Corollary. (i) *Let p be a point of $\mathbb{P}^2$. If all partial derivatives of an irreducible homogeneous polynomial f are zero at p, then f is zero there, and therefore p is a singular point of the curve defined by f.*

(ii) *The partial derivatives of an irreducible polynomial have no common (nonconstant) factor.*

(iii) *A plane curve has finitely many singular points.* $\square$

(1.4.9) tangent lines and flex points

Let C be the plane projective curve defined by an irreducible homogeneous polynomial f. A line L is *tangent* to C at a smooth point p if the intersection multiplicity of C and L at p is at least 2 (see (1.3.8)). As we will see, there is a unique tangent line at a smooth point.

A smooth point p of C is a *flex point* if the intersection multiplicity of C and its tangent line at p is at least 3, and p is an *ordinary* flex point if the intersection multiplicity is equal to 3.

Let L be a line through a point p and let q be a point of L distinct from p. We represent p and q by specific vectors (p_0, p_1, p_2) and (q_0, q_1, q_2), to write a variable point of L as $p + tq$, and we expand the restriction of f to L in a Taylor's Series. Let $f_i = \frac{\partial f}{\partial x_i}$ and $f_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$. Then

(1.4.10) $$f(p+tq) \;=\; f(p) \;+\; \Big(\sum_i f_i(p)\,q_i\Big)t \;+\; \tfrac{1}{2}\Big(\sum_{i,j} q_i\,f_{ij}(p)\,q_j\Big)t^2 \;+\; O(3)$$

where the symbol $O(3)$ stands for a polynomial in which all terms have degree at least 3 in t. The point q is missing from this parametrization, but this isn't important.

Note. The Taylor expansion carries over to complex polynomials because it is an identity. It can be used to show that the derivative has some properties analogous to properties of the derivative of a real polynomial.

Let ∇ be the *gradient vector* (f_0, f_1, f_2), let H be the *Hessian matrix* of f, the matrix of second partial derivatives

(1.4.11) $$H \;=\; \begin{pmatrix} f_{00} & f_{01} & f_{02} \\ f_{10} & f_{11} & f_{12} \\ f_{20} & f_{21} & f_{22} \end{pmatrix}$$

and let ∇_p and H_p be the evaluations of ∇ and H, respectively, at p. Regarding p and q as column vectors, Equation 1.4.10 can be written as

(1.4.12) $$f(p+tq) = \; f(p) \;+\; \big(\nabla_p\,q\big)t \;+\; \tfrac{1}{2}(q^tH_p\,q)t^2 \;+\; O(3)$$

in which $\nabla_p q$ and q^tH_pq are matrix products.

The intersection multiplicity of C and L at p (**1.3.7**) is the lowest power of t that has nonzero coefficient in $f(p + tq)$. The intersection multiplicity is at least 1 if p lies on C, i.e., if $f(p) = 0$.

Suppose that p is a smooth point of C. Then L is tangent to C at p if the coefficient (∇_pq) of t is zero, and p is a flex point if (∇_pq) and (q^tH_pq) are both zero.

The equation of the tangent line L at a smooth point p is $\nabla_px = 0$, or

(1.4.13) $$f_0(p)x_0 + f_1(p)x_1 + f_2(p)x_2 = 0$$

It tells us that a point q lies on L if the linear term of (1.4.12) is zero.

By the way, Taylor's formula shows that the restriction of f to every line through a singular point has a multiple zero. However, we will speak of tangent lines only at smooth points of the curve.

It is convenient to introduce the notation $\langle u, v\rangle$ for the symmetric form $u^tH_p\,v$ on $\mathbb{C}^3 \times \mathbb{C}^3$. It makes sense to say that this form vanishes on a pair of points of $\mathbb{P}^2$, because the relation $\langle u, v\rangle = 0$ doesn't depend on the vectors that represent those points.

Note that p is a smooth point of C if and only if $\langle p, p\rangle = 0$ but $\langle p, x\rangle$ is not identically zero.

1.4.14. Proposition. *Equation (1.4.10) can also be written as*

$$f(p+tq) \;=\; \tfrac{1}{d(d-1)}\langle p,p\rangle \;+\; \tfrac{1}{d-1}\langle p,q\rangle t \;+\; \tfrac{1}{2}\langle q,q\rangle t^2 \;+\; O(3)$$

proof. This follows from the two formulas

$p^tH_p = (d-1)\nabla_p$ and $\nabla_p p = df(p)$,, which can be obtained by applying Euler's formula to the entries of H_p and ∇_p. □

1.4.15. Corollary. *Let p be a smooth point of $\mathbb{P}^2$, let L be the line with the equation $\nabla_p x = 0$, and let q be a point of L distinct from p. Then*

(i) *L is tangent to C at p if and only if $\langle p,p\rangle = \langle p,q\rangle = 0$, and*

(ii) *p is a flex point of C with tangent line L if and only if $\langle p,p\rangle = \langle p,q\rangle = \langle q,q\rangle = 0$.* □

1.4.16. Theorem. *A smooth point p of the curve C is a flex point if and only if the determinant $\det H_p$ of the Hessian matrix at p is zero.*

proof. Let p be a smooth point of C, so that $\langle p,p\rangle = 0$ but $\langle p,v\rangle$ isn't identically zero. If $\det H_p = 0$, the form is degenerate. There is a nonzero null vector q, and because $\langle p,v\rangle$ isn't identically zero, q is distinct from p. But $\langle p,q\rangle = \langle q,q\rangle = 0$, so p is a flex point.

Conversely, suppose that p is a flex point and let q be a point on the tangent line at p and distinct from p, so that $\langle p,p\rangle = \langle p,q\rangle = \langle q,q\rangle = 0$. The restriction of the form to the two-dimensional space W spanned by p and q will be zero. A form on a space V of dimension 3 that restricts to zero on a two-dimensional subspace W is degenerate: If (p,q,v) is a basis with p,q in W, the matrix of the form on V will look like this:

$$\begin{pmatrix} 0 & 0 & * \\ 0 & 0 & * \\ * & * & * \end{pmatrix}$$

□

1.4.17. Proposition.

(i) *Let $f(x,y,z)$ be an irreducible homogeneous polynomial of degree at least two. The Hessian determinant $\det H$ isn't divisible by f. In particular, $\det H$ isn't identically zero.*

(ii) *A curve that isn't a line has finitely many flex points.*

proof. **(i)** Let C be the curve defined by f. If f divides the Hessian determinant, every smooth point of C will be a flex point. We set $z = 1$ and look on the standard affine $\mathbb{U}^2$, choosing coordinates so that the origin p is a smooth point of C, and $\frac{\partial f}{\partial y} \neq 0$ at p. The Implicit Function Theorem tells us that we can solve the equation $f(x,y,1) = 0$ for y locally, say $y = \varphi(x)$. The graph $\Gamma : \{y = \varphi(x)\}$ will be equal to C in a neighborhood of p (see below). A point of Γ is a flex point if and only if $\frac{d^2\varphi}{dx^2}$ is zero there. If this is true for all points near to p, then $\frac{d^2\varphi}{dx^2}$ will be identically zero, which implies that φ is linear: $y = ax$. Then $y = ax$ solves $f = 0$, and therefore $y{-}ax$ divides $f(x,y,1)$. But since $f(x,y,z)$ is irreducible, and so is $f(x,y,1)$. Therefore $f(x,y,1)$ and $f(x,y,z)$ are linear, contrary to hypothesis.

(ii) This follows from **(i)** and (1.3.11). The irreducible polynomial f and the Hessian determinant have finitely many common zeros. □

1.4.18. Review. (*about the Implicit Function Theorem*)

Let $f(x,y)$ be a polynomial such that $f(0,0) = 0$ and $\frac{df}{dy}(0,0) \neq 0$. The Implicit Function Theorem asserts that there is a unique analytic function $\varphi(x)$, defined for small x, such that $\varphi(0) = 0$ and $f(x,\varphi(x))$ is identically zero.

We make some further remarks. Let $\mathcal{R}$ be the ring of functions that are defined and analytic for small x. In the ring $\mathcal{R}[y]$ of polynomials in y with coefficients in $\mathcal{R}$, the polynomial $y{-}\varphi(x)$, which is monic in y, divides $f(x,y)$. To see this, we do division with remainder of f by $y - \varphi(x)$:

(1.4.19) $$f(x,y) = (y - \varphi(x))q(x,y) + r(x)$$

The quotient $q(x,y)$ is in $\mathcal{R}[y]$, and the remainder $r(x)$ has degree zero in y, so it is in $\mathcal{R}$. Setting $y = \varphi(x)$ in the equation, one sees that $r(x) = 0$.

Let Γ be the graph of φ in a suitable neighborhood U of the origin in x, y-space. Since $f(x,y) = (y - \varphi(x))q(x,y)$, the locus $f(x,y) = 0$ in U has the form $\Gamma \cup \Delta$, where Γ is the graph of φ and Δ is the zero locus of $q(x,y)$. Differentiating, we find that $\frac{\partial f}{\partial y}(0,0) = q(0,0)$. So $q(0,0) \neq 0$. Then Δ doesn't contain the origin, while Γ does. This implies that Δ is disjoint from Γ, locally. A sufficiently small neighborhood U of the origin won't contain any of points Δ. In such a neighborhood, the locus of zeros of f will be Γ. $\square$

1.5 Nodes and Cusps

Let C be the projective curve defined by an irreducible homogeneous polynomial $f(x,y,z)$ of degree d, and let p be a point of C. We choose coordinates so that $p = (0,0,1)$, and we set $z = 1$. This gives us an affine curve C_0 in $\mathbb{A}^2_{x,y}$, the zero set of the polynomial $\widetilde{f}(x,y) = f(x,y,1)$, and p becomes the origin $(0,0)$. We write

(1.5.1) $$\widetilde{f}(x,y) = f_0 + f_1 + f_2 + \cdots + f_d,$$

where f_i is the homogeneous part of $\widetilde{f}$ of degree i, which is also the coefficient of z^{d-i} in $f(x,y,z)$.

If the origin p is a point of C_0, the constant term f_0 will be zero. Then the linear term f_1 will define the tangent direction to C_0 at p, If f_0 and f_1 are both zero, p will be a singular point of C.

It seems permissible to drop the tilde and the subscript 0 in what follows, denoting $f(x,y,1)$ by $f(x,y)$, and C_0 by C.

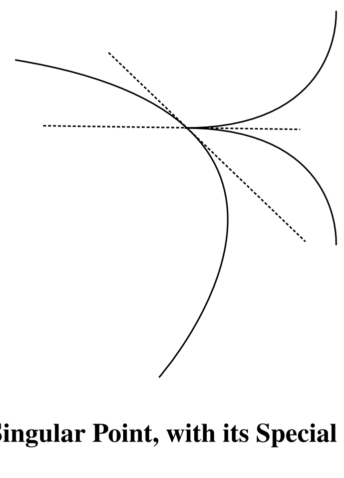
(1.5.2) the multiplicity of a singular point

Let $f(x,y)$ be an analytic function, defined for small x, y, and let C denote the locus of zeros of f in a neighborhood of $p = (0,0)$. To describe the singularity of C at p, one expands f as a series in x, y and looks at the part of f of lowest degree. The smallest integer r such that $f_r(x,y)$ isn't zero is the *multiplicity* of p. When the multiplicity of p is r, f will have the form

(1.5.3) $$f(x,y) = f_r + f_{r+1} + \cdots$$

Let L be a line $\{vx = uy\}$ through p. The intersection multiplicity of C and L at the origin p will be r unless $f_r(u,v)$ is zero. If $f_r(u,v) = 0$, the intersection multiplicity will be greater than r. Such lines are *special*. They correspond to the zeros of f_r in $\mathbb{P}^1$. Because f_r has degree r, there will be at most r of them.

1.5.4.



a Singular Point, with its Special Lines

(1.5.5) double points

Suppose that the origin p is a point of multiplicity 2, a *double point*,. Let the quadratic part of f be

$$f_2 = ax^2 + bxy + cy^2 \tag{1.5.6}$$

The point p is called a *node* if f_2 has distinct zeros in $\mathbb{P}^1$. A node is the simplest singularity that a curve can have.

When the discriminant $b^2 - 4ac$ is zero, f_2 will be a square. We may arrange coordinates so that $c \neq 0$. Then

$$4c(ax^2 + bxy + cy^2) \;=\; (bx + 2cy)^2 \tag{1.5.7}$$

The line L: $\{bx + 2cy = 0\}$ is special. The point p is a *cusp* if the multiplicity of intersection of C and L at p is 3. This will be true if and only if $bx + 2cy$ doesn't divide $f_3(x, y)$.

When the discriminant is zero, one can adjust coordinates to make $f_2 = y^2$. Then p is a cusp if the coefficient of the monomial x^3 in f_3 isn't zero. The *standard cusp* is the locus $y^2 = x^3$.

The definitions of nodes and cusps are made in terms of particular coordinates x, y, though they don't depend on the choice of coordinates.

1.5.8. Lemma. *Let $f(x, y)$ be an analytic function, let $x = x(u, v)$, $y = y(u, v)$ be an analytic change of variable, and let $g(u, v) = f(x(u, v), y(u, v))$. Also, let C and D be the local zero sets of f and g at p. Then C has a node or a cusp if and only D does.* $\square$

The simplest example of a double point that isn't a node or cusp is a *tacnode*, a point at which two smooth branches of a curve intersect with the same tangent direction (see Figure **??**).

1.6 Transcendence degree

Let $F \subset K$ be a field extension. A set $\alpha = \{\alpha_1, ..., \alpha_n\}$ of elements of K is *algebraically dependent* over F if there is a nonzero polynomial $f(x_1, ..., x_n)$ with coefficients in F, such that $f(\alpha) = 0$. If there is no such polynomial, the set α is *algebraically independent* over F.

An infinite set is called algebraically independent if every finite subset is algebraically independent – if there is no polynomial relation among any finite set of its elements.

The set consisting of a single element α_1 of K will be algebraically dependent if α_1 is algebraic over F. Otherwise, the set $\{\alpha_1\}$ of one element will be algebraically independent, and then α_1 is said to be *transcendental* over F.

An algebraically independent set $\alpha = \{\alpha_1, ..., \alpha_n\}$ that isn't contained in a larger algebraically independent set is called a *transcendence basis* for K over F. If there is a finite transcendence bases, its order is the *transcendence degree* of the field extension K of F. Lemma 1.6.2 below shows that all transcendence basis for K over F have the same order. Therefore the transcendence degree is well-defined. If there is no finite transcendence basis, the transcendence degree of K over F is infinite.

For example, let $K = F(x_1, ..., x_n)$ be the field of rational functions in n variables. The variables form a transcendence basis of K over F, and the transcendence degree of K over F is n.

A nonzero ring with no zero divisors will be called a *domain*, and a domain that contains the field F as a subring will be called an F-algebra. We use the customary notation $F[\alpha_1, ..., \alpha_n]$ or $F[\alpha]$ for the F-algebra generated by a set $\alpha = \{\alpha_1, ..., \alpha_n\}$, and we may denote the field of fractions of $F[\alpha]$ by $F(\alpha_1, ..., \alpha_n)$ or by $F(\alpha)$.

The set $\{\alpha_1, ..., \alpha_n\}$ is algebraically independent over F if and only if the surjective map from the polynomial algebra $F[x_1, ..., x_n]$ to $F[\alpha_1, ..., \alpha_n]$ that sends x_i to α_i is bijective.

1.6.1. Lemma. *Let K/F be a field extension, and let $\alpha = \{\alpha_1, ..., \alpha_n\}$ be a set of elements of K that is algebraically independent over F, and let $F(\alpha)$ be the field of fractions of $F[\alpha]$.*

(i) *Every element of the field $F(\alpha)$ that isn't in F is transcendental over F.*

(ii) *If β is another element of K, the set $\{\alpha_1, ..., \alpha_n, \beta\}$ is algebraically dependent if and only if β is algebraic over $F(\alpha)$.*

(iii) *The algebraically independent set α is a transcendence basis if and only if every element of K is algebraic over $F(\alpha)$.*

proof. **(i)** We write an element z of $F(\alpha)$ as a fraction $p/q = p(\alpha)/q(\alpha)$, where $p(x)$ and $q(x)$ are relatively prime polynomials. Suppose that z satisfies a nontrivial polynomial relation $c_0z^n + c_1z^{n-1} + \cdots + c_n = 0$ with c_i in F. We may assume that $c_0 = 1$. Substituting $z = p/q$ and multiplying by q^n gives us the equation

$$p^n = -q(c_1p^{n-1} + \cdots + c_nq^{n-1})$$

Because α is an algebraically independent set, this equation is equivalent with a polynomial equation in $F[x]$. It shows that q divides p^n, which contradicts the hypothesis that p and q are relatively prime. □

1.6.2. Lemma.

(i) *Let K/F be a field extension. If K has a finite transcendence basis, then all algebraically independent subsets of K are finite, and all transcendence bases have the same number of elements. Therefore the transcendence degree is well-defined.*

(ii) *If $L \supset K \supset F$ are fields and if the degree $[L : K]$ of L over K is finite, then K and L have the same transcendence degree over F.*

proof. **(i)** Let $\alpha = \{\alpha_1, ..., \alpha_r\}$ and $\beta = \{\beta_1, ..., \beta_s\}$. Assume that K is algebraic over $F(\alpha)$ and that the set β is algebraically independent. We show that $s \le r$. The fact that all transcendence bases have the same order will follow: If both α and β are transcendence bases, then $s \le r$, and since we can interchange α and β, $r \le s$.

The proof that $s \le r$ proceeds by reducing to the trivial case that β is a subset of α. Suppose that some element of β, say β_s, isn't in the set α. The set $\beta' = \{\beta_1, ..., \beta_{s-1}\}$ is algebraically independent, but it isn't a transcendence basis. So K isn't algebraic over $F(\beta')$. Since K is algebraic over $F(\alpha)$, there is at least one element of α, say α_r, that isn't algebraic over $F(\beta')$. Then $\gamma = \beta' \cup \{\alpha_s\}$ will be an algebraically independent set of order s, and it will contain more elements of the set α than β does. Induction shows that $s \le r$. □

1.7 The Dual Curve

(1.7.1) the dual plane

Let $\mathbb{P}$ denote the projective plane with coordinates x_0, x_1, x_2, and let L be the line in $\mathbb{P}$ whose equation is

(1.7.2) $$s_0x_0 + s_1x_1 + s_2x_2 = 0$$

The solutions of this equation determine the coefficients s_i only up to a common nonzero scalar factor, so the line L determines a point (s_0, s_1, s_2) in another projective plane $\mathbb{P}^*$ called the *dual plane*. We denote that point by L^*. Moreover, a point $p = (x_0, x_1, x_2)$ in $\mathbb{P}$ determines a line in the dual plane, the line with the equation (1.7.2), when s_i are regarded as the variables and x_i as the scalar coefficients. We denote that line by p^*. The equation exhibits a duality between $\mathbb{P}$ and $\mathbb{P}^*$. A point p of $\mathbb{P}$ lies on a line L if and only if the equation is satisfied, and this means that, in $\mathbb{P}^*$, the point L^* lies on the line p^*.

(1.7.3) the dual curve

Let C be a plane projective curve of degree at least two, and let U be the set of its smooth points. This is the complement of a finite subset of C. We define a map

$$U \xrightarrow{t} \mathbb{P}^*$$

If p is a point of U and L be the tangent line to C at p, we define $t(p) = L^*$, where L^* is the point of $\mathbb{P}^*$ that corresponds to L.

Denoting the partial derivative $\frac{\partial f}{\partial x_i}$ by f_i as before, the tangent line L at a smooth point $p = (x_0, x_1, x_2)$ of C has the equation $f_0x_0 + f_1x_1 + f_2x_2 = 0$. Therefore L^* is the point

(1.7.4) $$(s_0, s_1, s_2) = \big(f_0(x), f_1(x), f_2(x)\big)$$

Let U^* be the image of U in $\mathbb{P}^*$. Points of U^* correspond to tangent lines at smooth points of C.

$$??figure??$$

We assume that C has degree at least two because, if C were a line, U^* would be a point.

1.7.5. Lemma. *Let $\varphi(s_0, s_1, s_2)$ be a homogeneous polynomial, and let $g(x_0, x_1, x_2) = \varphi(f_0(x), f_1(x), f_2(x))$. Then $\varphi(s)$ is identically zero on U^* if and only if $g(x)$ is identically zero on U. This is true if and only if f divides g.*

proof. The first assertion comes from the fact that (s_0, s_1, s_2) and $(f_0(x), f_1(x), f_2(x))$ represent the same point of $\mathbb{P}^*$, and the last one follows from Corollary 1.3.21, because U is the complement of a finite set. □

1.7.6. Theorem. *Let C be the plane curve defined by an irreducible homogeneous polynomial f of degree at least two. With notation as above, the closure C^* of the image $U^* = t(U)$ is a curve in the dual space $\mathbb{P}^*$.*

The curve C^* referred to in the theorem is the *dual curve* .

proof. If an irreducible homogeneous polynomial $\varphi(s_0, s_1, s_2)$ vanishes on U^*, it will be unique up to scalar factor (Corollary 1.3.15). Its zero locus will be the dual curve.

We show first that there is a nonzero polynomial φ, not necessarily irreducible or homogeneous, that vanishes on U^*. The field $\mathbb{C}(x_0, x_1, x_2)$ has transcendence degree three over $\mathbb{C}$. Therefore the four polynomials f_0, f_1, f_2, and f are algebraically dependent. There is a nonzero polynomial $\psi(s_0, s_1, s_2, t)$ such that $\psi(f_0(x), f_1(x), f_2(x), f(x))$ is the zero polynomial in x. We can cancel factors of t, so we may assume that ψ isn't divisible by t. Let $\varphi(s_0, s_1, s_2) = \psi(s_0, s_1, s_2, 0)$. When t doesn't divide ψ, φ isn't the zero polynomial.

Let $\overline{x} = (\overline{x}_1, \overline{x}_2, \overline{x}_3)$ be a vector that represents a point of U, and let $f(\overline{x}) = \overline{f}$ and $f_i(\overline{x}) = \overline{f}_i$. Then $\overline{f} = 0$, and therefore

$$\varphi(\overline{f}_0, \overline{f}_1, \overline{f}_2) = \psi(\overline{f}_0, \overline{f}_1 \overline{f}_2, 0) = \psi(\overline{f}_0, \overline{f}_1, \overline{f}_2, \overline{f})$$

Since $\psi(f_0, f_1, f_2, f)$ is identically zero. $\varphi(\overline{f}_0, \overline{f}_1, \overline{f}_2) = 0$ for all $\overline{x}$ that represent points of U. The vector $\lambda\overline{x}$ represents the same point of U, and if f is homogeneous of some degree d, the derivatives f_i are homogeneous of degree $d-1$. Therefore $\varphi(f_0(\lambda\overline{x}), f_1(\lambda\overline{x}), f_2(\lambda\overline{x})) = \lambda^{d-1}\varphi(\overline{f}_0, \overline{f}_1, \overline{f}_2) = 0$.

Since λ^{d-1} can be any complex number, Lemma 1.3.2 tells us that the homogeneous parts of φ vanish at $(\overline{f}_0, \overline{f}_1, \overline{f}_2)$, and the homogeneous parts of $\varphi(s)$ vanish on U^* (1.7.5). So we may assume that φ is homogeneous, of some degree r. Then if f has degreee d, the polynomial $g(x) = \varphi(f_0(x), f_1(x), f_2(x))$ will be homogeneous, of degree $r(d-1)$. It will vanish on U, and therefore on C (1.3.21). So f will divide g. Finally, if $\varphi(s)$ factors, then $g(x)$ factors accordingly, and because f is irreducible, it will divide one of the factors of g. The corresponding factor of φ will vanish on U^* (1.7.5). So we may replace the polynomial φ, now homogeneous, by one of its irreducible factors. □

In principle, the proof of the theorem gives a method for finding a polynomial that vanishes on C^*. That is to find a polynomial relation among f_x, f_y, f_z, f, and then set $f = 0$. But it can be painful to determine the defining polynomial of the dual curve explicitly. The degrees of C and C^* will often be different, and several points of the dual curve C^* may correspond to a singular point of C, and vice versa.

However, the computation is simple for a conic.

1.7.7. Examples.

(i) (*the dual of a conic*) Let $f = x_0x_1 + x_0x_2 + x_1x_2$ and let C be the conic $f = 0$. Let $(s_0, s_1, s_2) = (f_0, f_1, f_2) = (x_1+x_2, x_0+x_2, x_0+x_1)$. Then

(1.7.8) $$s_0^2 + s_1^2 + s_2^2 - 2(x_0^2 + x_1^2 + x_2^2) = 2f \quad \text{and} \quad s_0s_1 + s_1s_2 + s_0s_2 - (x_0^2 + x_1^2 + x_2^2) = 3f$$

We eliminate $(x_0^2 + x_1^2 + x_2^2)$ from the two equations.

(1.7.9) $$(s_0^2 + s_1^2 + s_2^2) - 2(s_0s_1 + s_1s_2 + s_0s_2) = -4f$$

Setting $f = 0$ gives us the equation of the dual curve. It is another conic.

(ii) (*the dual of a cuspidal cubic*) It isn't easy to compute the dual of a smooth cubic, whose equation has degree 6. We compute the dual of a cubic with a cusp instead. The curve C defined by the irreducible polynomial $f = y^2z + x^3$ has a cusp at $(0, 0, 1)$. The Hessian matrix of f is

$$H = \begin{pmatrix} 6x & 0 & 0 \\ 0 & 2z & 2y \\ 0 & 2y & 0 \end{pmatrix}$$

and the Hessian determinant $h = \det H$ is $-24xy^2$. The common zeros of f and h are the cusp point $(0,0,1)$ and a single flex point $(0,1,0)$.

We scale the partial derivatives of f to simplify notation. Let $u = f_x/3 = x^2$, $v = f_y/2 = yz$, and $w = f_z = y^2$. Then

$$v^2w - u^3 = y^4z^2 - x^6 = (y^2z + x^3)(y^2z - x^3) = f(y^2z - x^3)$$

The zero locus of the irreducible polynomial $v^2w - u^3$ is the dual curve. It is another cuspidal cubic. □

(1.7.10) a local equation for the dual curve

We label the coordinates in $\mathbb{P}$ and $\mathbb{P}^*$ as x, y, z and u, v, w, respectively, and we work in a neighborhood of a smooth point p_0 of the curve C defined by a homogeneous polynomial $f(x, y, z)$, choosing coordinates so that $p_0 = (0, 0, 1)$, and that the tangent line at p_0 is the line $L_0 : \{y = 0\}$. The image L_0^* of p_0 in the dual curve C^* is $(u, v, w) = (0, 1, 0)$.

Let $\widetilde{f}(x, y) = f(x, y, 1)$. In the affine x, y-plane, the point p_0 becomes the origin $(0, 0)$. So $\widetilde{f}(0, 0) = 0$, and since the tangent line is the line $\{y = 0\}$, $\frac{\partial \widetilde{f}}{\partial x}(0, 0) = 0$, while $\frac{\partial \widetilde{f}}{\partial y}(0, 0) \neq 0$. The Implicit Function Theorem allows us to solve the equation $\widetilde{f} = 0$ for y as an analytic function $y(x)$ for small x, with $y(0) = 0$. Let $y'(x)$ denote the derivative $\frac{dy}{dx}$. Differentiating the equation $f(x, y(x)) = 0$ shows that $y'(0)$ is zero too.

Let $\widetilde{p}_1 = (x_1, y_1)$ be a point of C_0 near to $\widetilde{p}_0$, so that $y_1 = y(x_1)$, and let $y_1' = y'(x_1)$. The tangent line L_1 at $\widetilde{p}_1$ has the equation

$$y - y_1 = y_1'(x - x_1) \tag{1.7.11}$$

Putting z back, the homogeneous equation of the tangent line L_1 at the point $(x_1, y_1, 1)$ is

$$(-y_1')x + y + (y_1'x_1 - y_1)z = 0$$

Equation (1.7.11) tells us that the point L_1^* of the dual plane that corresponds to L_1 is $(-y_1', 1, y_1'x_1 - y_1)$. Let's drop the subscript 1. As x varies, and writing $y = y(x)$ and $y' = y'(x)$,

$$(u, v, w) = (-y', 1, y'x - y), \tag{1.7.12}$$

There may be accidents: L_0 might be tangent to C at distinct smooth points q_0 and p_0, and it might pass through a singular point of C. If either of these accidents occurs, we can't analyze the neighborhood of L_0^* in C^* by this method. But, provided that they don't occur, the path (1.7.12) will trace out the dual curve C^* near to $L_0^* = (0, 1, 0)$. (See (1.4.18.) □

(1.7.13) the bidual

The *bidual* C^{**} of a curve C is the dual of the curve C^*, which is a curve in the space $\mathbb{P}^{**} = \mathbb{P}$.

1.7.14. Theorem. *A plane curve of degree greater than one is equal to its bidual:* $C^{**} = C$.

1.7.15. Lemma.
(i) *The set V of points p_0 of a curve C such that C is smooth at p_0 and its dual C^* is smooth at $t(p_0)$ is the complement of a finite subset of C.*
(ii) *Let p_1 be a point near to a smooth point p_0 of a curve C, let L_1 and L_0 be the tangent line to C at p_1 and p_0, respectively, and let q be intersection point $L_1 \cap L_0$. Then* $\lim_{p_1 \to p_0} q = p_0$.
(iii) *If p_0 is a point of V with tangent line L_0, then the tangent line to C^* at L_0^* is p_o^*.*

proof. **(i)** Let S and S^* denote the finite sets of singular points of C, and C^*, respectively. So the set U of smooth points of C is the complement of S in C, and V is obtained from U by deleting points whose images are in S^*. The fibre of t over a point L^* of C^* is the set of smooth points p of C such that the tangent line at p is L. Since L meets C in finitely many points, the fibre is finite. So the inverse image of $S^* \cap U$ will be a finite subset of U whose complement is V.

(ii) We work analytically in a neighborhood of p_0, choosing coordinates so that $p_0 = (0,0,1)$ and that L_0 is the line $\{y = 0\}$. Let $(x_q, y_q, 1)$ be the coordinates of $q = L_0 \cap L_1$. Since q is a point of L_0, $y_q = 0$. The coordinate x_q can be obtained by substituting $x = x_q$ and $y = 0$ into the equation (1.7.11) of L_1:

$$x_q \ = \ x_1 - y_1/y_1'.$$

Now, when a function has an nth order zero at the point $x = 0$, i.e, when it has the form $y = x^n h(x)$, where $n > 0$ and $h(0) \neq 0$, the order of zero of its derivative at that point is $n-1$. This is verified by differentiating $x^n h(x)$. Since the function $y(x)$ has a zero of positive order at p_0, $\lim_{p_1 \to p_0} y_1/y_1' = 0$. We also have $\lim_{p_1 \to p_0} x_1 = 0$. So $\lim_{p_1 \to p_0} x_q = 0$ and $\lim_{p_1 \to p_0} q = \lim_{p_1 \to p_0} (x_q, y_q, 1) = (0,0,1) = p_0$.

$figure$

(iii) Let p_1 be a point of C near to p_0, and let L_1 be the tangent line to C at p_1. The image of p_1 is $L_1^* = (f_0(p_1), f_1(p_1), f_2(p_1))$. Because the partial derivatives f_i are continuous,

$$\lim_{p_1 \to p_0} L_1^* = (f_0(p_0), f_1(p_0), f_2(p_0)) = L_0^*$$

With $q = L_0 \cap L_1$, q^* is the line through the points L_0^* and L_1^*. As p_1 approaches p_0, L_1^* approaches L_0^*, and therefore q^* approaches the tangent line to C^* at L_0^*. On the other hand, the lemma tells us that q^* approaches p_0^*. Therefore the tangent line at L_0^* is p_0^*. □

proof of theorem 1.7.14. Let V be the set of smooth points of C whose images in C^* are smooth, as in Lemma 1.7.15. Let U^* denote the set of smooth points of C^*, and let $U^* \xrightarrow{t^*} \mathbb{P}^{**} = \mathbb{P}$ be the map analogous to the map t. Recall that t is defined by $t(p) = L^*$. Since the tangent line to C^* at L^* is p^*, the map $U^* \xrightarrow{t^*} C$ analogous to t is $t^*(L^*) = (p^*)^* = p$. So for all points p of V, $t^*t(p) = t^*(L^*) = p$. It follows that the restriction of t to V is injective, and that it defines a bijective map from V to its image V^*, whose inverse function is t^*. So V is contained in the bidual C^{**}. Since V is dense in C and C^{**} is a closed set, $C \subset C^{**}$. Since C and C^{**} are curves, $C = C^{**}$. □

1.7.16. Corollary. *Let C be a smooth curve. The map $C \xrightarrow{t} C^*$, which is defined at all points of C, is surjective.*

proof. Let W denote the image of C in C^*. The map $C^* \xrightarrow{t^*} C^{**} = C$ is defined at the smooth points of C^*, and it inverts t at those points. Therefore W contains the smooth points of C^*. The complement S of W in C^* is a finite set. Since C is compact, its image W is compact, and therefore closed in C^*. Then its complement S is open, and since it is a finite set, S is also closed. So S consists of isolated points of C^*. Since a plane curve has no isolated point (1.3.18), S is empty. □

1.8 Resultants

Let F and G be monic polynomials in x with variable coefficients:

(1.8.1) $$F(x) = x^m + a_1 x^{m-1} + \cdots + a_m \quad \text{and} \quad G(x) = x^n + b_1 x^{n-1} + \cdots + b_n$$

The resultant $\text{Res}(F,G)$ of F and G is a certain polynomial in the coefficients. Its important property is that, when the coefficients of are in a field, the resultant is zero if and only if F and G have a common factor.

The formula for the resultant is nicest when one allows leading coefficients different from 1. We work with homogeneous polynomials in two variables to prevent the degrees from dropping when a leading coefficient happens to be zero.

Let f and g be homogeneous polynomials in x, y with complex coefficients:

(1.8.2) $$f(x,y) = a_0x^m + a_1x^{m-1}y + \cdots + a_my^m, \quad g(x,y) = b_0x^n + b_1x^{n-1}y + \cdots + b_ny^n$$

Suppose that they have a common zero $(x, y) = (u, v)$ in $\mathbb{P}^1_{xy}$. Then $vx - uy$ divides both f and g. The polynomial $h = fg/(vx - uy)$ of degree $m+n-1$ will be divisible by f and by g, say $h = pf = qg$, where p and q are homogeneous polynomials of degrees $n-1$ and $m-1$, respectively. Then h will be a linear combination pf of the polynomials x^iy^jf, with $i+j = n-1$, and it will also be a linear combination qg of the polynomials $x^ky^\ell g$, with $k+\ell = m-1$. The equation $pf = qg$ tells us that the $m+n$ polynomials of degree $m+n-1$,

(1.8.3) $$x^{n-1}f,\ x^{n-2}yf, ..., \ y^{n-1}f \ ; \ x^{m-1}g,\ x^{m-2}yg, ..., \ y^{m-1}g$$

will be dependent. For example, suppose that f has degree 3 and g has degree 2. If f and g have a common zero, the polynomials

$$\begin{aligned}
xf &= a_0x^4 + a_1x^3y + a_2x^2y^2 + a_3xy^3 \\
yf &= a_0x^3y + a_1x^2y^2 + a_2xy^3 + a_3y^4 \\
x^2g &= b_0x^4 + b_1x^3y + b_2x^2y^2 \\
xyg &= b_0x^3y + b_1x^2y^2 + b_2xy^3 \\
y^2g &= bx^2y^2 + b_1xy^3 + b_2y^4
\end{aligned}$$

will be dependent. Conversely, if the polynomials (1.8.3) are dependent, there will be an equation of the form $pf = qg$, with p of degree $n-1$ and q of degree $m-1$. Then at least one zero of g must also be a zero of f.

Let $r = m+n-1$. We form a square $(r+1)\times(r+1)$ matrix $\mathcal{R}$, the *resultant matrix*, whose columns are indexed by the monomials $x^r, x^{r-1}y, ..., y^r$ of degree r, and whose rows list the coefficients of the polynomials (1.8.3). The matrix is illustrated below for the cases $m, n = 3, 2$ and $m, n = 1, 2$, with dots representing entries that are zero:

(1.8.4) $$\mathcal{R} = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 & \cdot \\ \cdot & a_0 & a_1 & a_2 & a_3 \\ b_0 & b_1 & b_2 & \cdot & \cdot \\ \cdot & b_0 & b_1 & b_2 & \cdot \\ \cdot & \cdot & b_0 & b_1 & b_2 \end{pmatrix} \quad \text{or} \quad \mathcal{R} = \begin{pmatrix} a_0 & a_1 & \cdot \\ \cdot & a_0 & a_1 \\ b_0 & b_1 & b_2 \end{pmatrix}$$

The *resultant* of f and g is defined to be the determinant of $\mathcal{R}$.

(1.8.5) $$\mathrm{Res}(f, g) = \det \mathcal{R}$$

The coefficients of the polynomials f and g can be in any ring.

The resultant $\mathrm{Res}(F, G)$ of the monic, one-variable polynomials $F(x) = x^m + a_1x^{m-1} + \cdots + a_m$ and $G(x) = x^n + b_1x^{n-1} + \cdots + b_n$ is the determinant of the matrix $\mathcal{R}$, with $a_0 = b_0 = 1$.

1.8.6. Corollary. *Let f and g be homogeneous polynomials in two variables, or monic polynomials in one variable, of degrees m and n, respectively, and with coefficients in a field. The resultant $\mathrm{Res}(f, g)$ is zero if and only if f and g have a common factor. If so, there will be polynomials p and q of degrees $n-1$ and $m-1$ respectively, such that $pf = qg$. If the coefficients are in $\mathbb{C}$, the resultant is zero if and only if f and g have a common root.* □

When the leading coefficients a_0 and b_0 of f and g are both zero, the point $(1, 0)$ of $\mathbb{P}^1_{xy}$ will be a zero of f and of g. In this case, f and g have a common zero at infinity.

(1.8.7) weighted degree

When defining the degree of a polynomial, one may assign an integer called a *weight* to each variable. If one assigns weight w_i to the variable x_i, the monomial $x_1^{e_1} \cdots x_n^{e_n}$ gets the *weighted degree*

$$e_1 w_1 + \cdots + e_n w_n$$

For example, it is natural to assign weight k to the coefficient a_k of the polynomial $f(x) = x^n - a_1 x^{n-1} + a_2 x^{n-2} - \cdots \pm a_n$. The reason is that, if f factors into linear factors, $f(x) = (x - \alpha_1) \cdots (x - \alpha_n)$, then a_k will be the kth elementary symmetric function in $\alpha_1, ..., \alpha_n$. When written as a polynomial in α, the degree of a_k will be k.

We leave the proof of the next lemma as an exercise.

1.8.8. Lemma. *Let $f(x, y)$ and $g(x, y)$ be homogeneous polynomials of degrees m and n respectively, with variable coefficients a_i and b_i, as in (1.8.2). When one assigns weight i to a_i and to b_i, the resultant* $\text{Res}(f, g)$ *becomes a weighted homogeneous polynomial of degree mn in the variables $\{a_i, b_j\}$.* □

1.8.9. Proposition. *Let F and G be products of monic linear polynomials, say $F = \prod_i (x - \alpha_i)$ and $G = \prod_j (x - \beta_j)$. Then*

$$\text{Res}(F, G) \quad = \quad \prod_{i,j} (\alpha_i - \beta_j) \quad = \quad \prod_i G(\alpha_i)$$

proof. The equality of the second and third terms is obtained by substituting α_i for x into the formula $G = \prod (x - \beta_j)$. We prove that the first and second terms are equal.

Let the elements α_i and β_j be variables, let R denote the resultant $\text{Res}(F, G)$ and let Π denote the product $\prod_{i.j} (\alpha_i - \beta_j)$. When we write the coefficients of F and G as symmetric functions in the roots α_i and β_j, R will be homogeneous. Its (unweighted) degree in α_i, β_j will be mn, the same as the degree of Π (Lemma 1.8.8). To show that $R = \Pi$, we choose i, j and divide R by the polynomial $\alpha_i - \beta_j$, considered as a monic polynomial in α_i:

$$R = (\alpha_i - \beta_j) q + r,$$

where r has degree zero in α_i. The resultant R vanishes when we substitute $\alpha_i = \beta_j$. Looking at this equation, we see that the remainder r also vanishes when $\alpha_i = \beta_j$. On the other hand, the remainder is independent of α_i. It doesn't change when we set $\alpha_i = \beta_j$. Therefore the remainder is zero, and $\alpha_i - \beta_j$ divides R. This is true for all i and all j, so Π divides R, and since these two polynomials have the same degree, $R = c\Pi$ for some scalar c. To show that $c = 1$, one computes R and Π for some particular polynomials. We suggest using $F = x^m$ and $G = x^n - 1$. □

1.8.10. Corollary. *Let F, G, H be monic polynomials and let c be a scalar. Then*
(i) $\text{Res}(F, GH) = \text{Res}(F, G) \, \text{Res}(F, H)$, *and*
(ii) $\text{Res}(F(x-c), G(x-c)) = \text{Res}(F(x), G(x))$. □

(1.8.11) the discriminant

The *discriminant* $\text{Discr}(F)$ of a polynomial $F = a_0 x^m + a_1 x^{n-1} + \cdots a_m$ is the resultant of F and its derivative F':

$$\text{Discr}(F) = \text{Res}(F, F') \tag{1.8.12}$$

The computation of the discriminant is made using the formula for the resultant of a polynomial of degree m. The definition makes sense when the leading coefficient a_0 is zero, though the discriminant will be zero in that case.

Note. The formula for the discriminant is often normalized by a factor $\pm a_0^k$. We won't make this normalization, so our formula of the discriminant is slightly different from the usual one.

When the coefficients of F are complex numbers, the discriminant is zero if and only if either F has a double root, which happens when F and F' have a common factor, or else F has degree less than m.

For example, the discriminant of the quadratic polynomial $F(x) = ax^2 + bx + c$ is

$$(1.8.13) \qquad \det \begin{pmatrix} a & b & c \\ 2a & b & \cdot \\ \cdot & 2a & b \end{pmatrix} = -a(b^2 - 4ac).$$

The discriminant of the monic cubic $x^3 + px + q$ whose quadratic coefficient is zero is

$$(1.8.14) \qquad \det \begin{pmatrix} 1 & \cdot & p & q & \cdot \\ \cdot & 1 & \cdot & p & q \\ 3 & \cdot & p & \cdot & \cdot \\ \cdot & 3 & \cdot & p & \cdot \\ \cdot & \cdot & 3 & \cdot & p \end{pmatrix} = 4p^3 + 27q^2$$

These are the negatives of the usual formulas. The signs are artifacts of our definition.

1.8.15. Proposition. *Let K be a field of characteristic zero. The discriminant of an irreducible polynomial F with coefficients in K isn't zero. Therefore an irreducible polynomial F with coefficients in K has no multiple root.*

proof. When F is irreducible, it cannot have a factor in common with the derivative F', which has lower degree. □

This proposition is false when the characteristic of K isn't zero. In characteristic p, the derivative F' might be the zero polynomial.

1.8.16. Proposition. *Let $F = \prod(x - \alpha_i)$ be a polynomial that is a product of monic linear factors. Then*

$$\mathrm{Discr}(F) = \prod_i F'(\alpha_i) = \prod_{i \neq j}(\alpha_i - \alpha_j) = \pm \prod_{i<j}(\alpha_i - \alpha_j)^2$$

proof. The fact that $\mathrm{Discr}(F) = \prod F'(\alpha_i)$ follows from Proposition 1.8.9. We show that

$$F'(\alpha_i) = \prod_{j, j\neq i} (\alpha_i - \alpha_j) = (\alpha_i - \alpha_1) \cdots (\widehat{\alpha_i - \alpha_i}) \cdots (\alpha_i - \alpha_n)$$

where the hat $\widehat{\ }$ indicates that that term is deleted. By the product rule for differentiation,

$$F'(x) = \sum_k (x - \alpha_1) \cdots (\widehat{x - \alpha_k}) \cdots (x - \alpha_n)$$

Substituting $x = \alpha_i$, all terms in the sum, except the one with $i = k$, become zero. □

1.8.17. Corollary. $\mathrm{Discr}(F(x)) = \mathrm{Discr}(F(x - c))$. □

1.8.18. Proposition. *Let $F(x)$ and $G(x)$ be monic polynomials. Then*

$$\mathrm{Discr}(FG) = \pm \mathrm{Discr}(F)\, \mathrm{Discr}(G) \mathrm{Res}(F, G)^2$$

proof. This proposition follows from Propositions 1.8.9 and 1.8.16 for polynomials with complex coefficients. It is true for polynomials with coefficients in any ring because it is an identity. □

1.9 Hensel's Lemma

The resultant matrix (1.8.4) arises in a second context that we explain here.

Suppose given a product $P = FG$ of two polynomials, say

$$(1.9.1) \quad \big(c_0x^{m+n} + c_1x^{m+n-1} + \cdots + c_{m+n}\big) = \big(a_0x^m + a_1x^{m-1} + \cdots + a_m\big)\big(b_0x^n + b_1x^{n-1} + \cdots + b_n\big)$$

We call the relations among the coefficients that are implied by this polynomial equation the *product equations*. The product equations are

$$c_i = a_i b_0 + a_{i-1} b_1 + \cdots + a_0 b_i$$

for $i = 0, ..., m+n$. For instance, when $m = 3$ and $n = 2$, they are

1.9.2.

$$\begin{aligned}
c_0 &= a_0 b_0 \\
c_1 &= a_1 b_0 + a_0 b_1 \\
c_2 &= a_2 b_0 + a_1 b_1 + a_0 b_2 \\
c_3 &= a_3 b_0 + a_2 b_1 + a_1 b_2 \\
c_4 &= \qquad\quad a_3 b_1 + a_2 b_2 \\
c_5 &= \qquad\qquad\qquad a_3 b_2
\end{aligned}$$

Let J denote the Jacobian matrix of partial derivatives of $c_1, ..., c_{m+n}$ with respect to the variables $b_1, ..., b_n$ and $a_1, ..., a_m$, treating a_0, b_0 and c_0 as constants. When $m, n = 3, 2$,

(1.9.3)
$$J = \frac{\partial(c_i)}{\partial(b_j, a_k)} = \begin{pmatrix} a_0 & . & b_0 & . & . \\ a_1 & a_0 & b_1 & b_0 & . \\ a_2 & a_1 & b_2 & b_1 & b_0 \\ a_3 & a_2 & . & b_2 & b_1 \\ . & a_3 & . & . & b_2 \end{pmatrix}$$

1.9.4. Lemma. *The Jacobian matrix J is the transpose of the resultant matrix $\mathcal{R}$ (1.8.4).* □

1.9.5. Corollary. *Let F and G be polynomials with complex coefficients. The Jacobian matrix is singular if and only if, either F and G have a common root, or $a_0 = b_0 = 0$.* □

This corollary has an application to polynomials with analytic coefficients. Let

(1.9.6)
$$P(t, x) = c_0(t)x^d + c_1(t)x^{d-1} + \cdots + c_d(t)$$

be a polynomial in x whose coefficients c_i are analytic functions, defined for small values of t, and let $\overline{P} = P(0, x) = \overline{c}_0 x^d + \overline{c}_1 x^{d-1} + \cdots + \overline{c}_d$ be the evaluation of P at $t = 0$, so that $\overline{c}_i = c_i(0)$. Suppose given a factorization $\overline{P} = \overline{F}\,\overline{G}$, where $\overline{G} = \overline{b}_0 x^n + \overline{b}_1 x^{n-1} + \cdots + \overline{b}_n$ is a polynomial and $\overline{F} = x^m + \overline{a}_1 x^{m-1} + \cdots + \overline{a}_m$ is a monic polynomial, both with complex coefficients. Are there polynomials $F(t, x) = x^m + a_1 x^{m-1} + \cdots + a_m$ and $G(t, x) = b_0 x^n + b_1 x^{n-1} + \cdots + b_n$, with F monic, whose coefficients a_i and b_i are analytic functions defined for small t, such that $P = FG$, $F(0, x) = \overline{F}$, and $G(0, x) = \overline{G}$?

1.9.7. Hensel's Lemma. *With notation as above, suppose that $\overline{F}$ and $\overline{G}$ have no common root. Then P factors, as above.*

proof. Since F is supposed to be monic, we set $a_0(t) = 1$. The first product equation tells us that $b_0(t) = c_0(t)$. Corollary 1.9.5 tells us that the Jacobian matrix for the remaining product equations is nonsingular at $t = 0$, so according to the Implicit Function Theorem, the product equations have a unique solution in analytic functions $a_i(t), b_j(t)$ for small t. □

Note that P isn't assumed to be monic. If $\overline{c}_0 = 0$, the degree of $\overline{P}$ will be less than the degree of P. In that case, $\overline{G}$ will have lower degree than G.

figure

1.9.8. Example. Let $P = c_0(t)x^2 + c_1(t)x + c_2(t)$. The product equations for factoring P as a product FG of two linear polynomials with F monic are

$$c_0 = b_0\,, \quad c_1 = a_1 b_0 + b_1\,, \quad c_2 = a_1 b_1$$

and the Jacobian matrix is $\begin{pmatrix} 1 & b_0 \\ a_1 & b_1 \end{pmatrix}$.

Suppose that $\overline{P} = P(0,x)$ factors: $\overline{c}_0 x^2 + \overline{c}_1 x + \overline{c}_2 = (x+\overline{a}_1)(\overline{b}_0 x + \overline{b}_1) = \overline{F}\,\overline{G}$. The determinant of the Jacobian matrix at $t=0$ is $\overline{b}_1 - \overline{a}_1\overline{b}_0$. It is nonzero if and only if the two factors are relatively prime, in which case P factors too.

On the other hand, the one-variable Jacobian criterion allows us to solve the equation $P(t,x) = 0$ for x as function of t with $x(0) = -\overline{a}_1$, provided that $\frac{\partial P}{\partial x} = 2c_0 x + c_1$ isn't zero at the point $(t,x) = (0, -\overline{a}_1)$. In that case, $\overline{P}$ factors. Substituting $\overline{c}_0 = \overline{b}_0$ and $\overline{c}_1 = \overline{a}_1\overline{b}_0 + \overline{b}_1$, shows that $-2\overline{c}_0\overline{a}_1 + \overline{c}_1 = \overline{b}_1 - \overline{a}_1\overline{b}_0$. Not surprisingly, the two conditions for factoring are the same. $\square$

1.10 Plane Curves as Coverings of the Projective Line

When f and g are polynomials in several variables including a variable z, $\operatorname{Res}_z(f,g)$ and $\operatorname{Discr}_z(f)$ will denote the resultant and discriminant, computed regarding f, g as polynomials in z. The resultant and discriminant will be polynomials in the other variables.

1.10.1. Lemma. (i) *Let $F = \mathbb{C}(x,y)$ be the field of rational functions in x, y. An irreducible polynomial f in $\mathbb{C}[x,y,z]$ has positive degree in z, but isn't divisible by z is also an irreducible element of $F[z]$.*
(ii) *Let $f(x,y,z)$ be an irreducible polynomial in $\mathbb{C}[x,y,z]$ that isn't divisible by z. The discriminant $\operatorname{Discr}_z(f)$ of f with respect to the variable z is a nonzero polynomial in x, y.* $\square$

proof. **(i)** Say that $f(x,y,z)$ factors in $F[z]$, $f = g'h'$, where g' and h' are polynomials of positive degree in z. When we clear denominators from g' and h', we obtain an equation of the form $df = gh$, where g and h are polynomials in x,y,z of positive degree in z and d is a polynomial in x,y. Since g and h have positive degree in z, neither of them divides d. Then f must be reducible.

(ii) This follows from Proposition 1.8.15. $\square$

Let π denote the *projection* $\mathbb{P}^2 \longrightarrow \mathbb{P}^1$ that drops the last coordinate, sending a point (x,y,z) to (x,y). This projection is defined at all points of $\mathbb{P}^2$ except at the point $q = (0,0,1)$, which is called the *center of projection.*

The fibre of π over a point $\widetilde{p} = (x_0, y_0)$ of $\mathbb{P}^1$ is the line L_{pq} through $p = (x_0,y_0,0)$ and $q = (0,0,1)$, with the point q omitted – the set of points (x_0,y_0,z_0).

$$figure$$

Let C be a plane curve that doesn't contain the center of projection q. The projection $\mathbb{P}^2 \xrightarrow{\pi} \mathbb{P}^1$ will be defined everywhere on C. Say that C is defined by an irreducible homogeneous polynomial $f(x,y,z)$ of degree d. We write f as a polynomial in z,

$$f = c_0 z^d + c_1 z^{d-1} + \cdots + c_d \tag{1.10.2}$$

with c_i homogeneous, of degree i in x,y. Then $f(0,0,1) = c_0$, and since C doesn't contain q, c_0 will be a nonzero constant that we normalize to 1, so that f becomes a monic polynomial of degree d in z.

The fibre of C over a point $\widetilde{p} = (x_0,y_0)$ of $\mathbb{P}^1$ is the intersection of C with the line L_{pq} described above. It consists of the points (x_0,y_0,α) such that α is a root of the one-variable polynomial

$$\widetilde{f}(z) = f(x_0,y_0,z) \tag{1.10.3}$$

We call C a *branched covering* of $\mathbb{P}^1$ of degree d. Its most important property is that all but finitely many fibres of C over $\mathbb{P}^1$ consist of d points (Lemma 1.10.1). The fibres with fewer than d points are those above the zeros of the discriminant. Points of $\mathbb{P}^1$ whose fibres contain fewer points are *branch points* of the covering.

Let's suppose that coordinates are chosen so that $q = (0,0,1)$ is in general position. In algebraic geometry, the phrases *general position* and *generic* indicate an object (the point q here) that has no special 'bad' properties. Typically, the object will be parametrized somehow, and the word *generic* indicates that the parameter representing that particular object avoids a proper closed subset of the parameter space that may be described explicitly or not. For Proposition 1.10.6 below, we require that q shall not lie on any of the following lines:

(1.10.4)

flex tangent lines and bitangent lines,
lines that contain more than one singular point,
special lines through singular points (see (**1.5.2**)),
tangent lines that contain a singular point of C.

1.10.5. Lemma. *This is a list of finitely many lines that q must avoid.*

beginning of the proof. Proposition 1.4.17 shows that there are finitely many flex tangents. Since there are finitely many singular points, there are finitely many lines through pairs of singular points and finitely many special lines. To show that there are finitely many tangent lines that pass through singular points, we project C from a singular point p and apply Lemma 1.10.1. The discriminant isn't identically zero, so it vanishes finitely often.

The proof that there are finitely many bitangents will be given later, in Corollary 1.12.15.

1.10.6. Proposition. *Let $f(x, y, z)$ be a homogeneous polynomial with no multiple factors, and let C be the (possibly reducible) plane curve $\{f = 0\}$. Suppose that $q = (0, 0, 1)$ is in general position with respect to C.*
(i) *If p is a smooth point of C with tangent line L_{pq}, the discriminant $\text{Discr}_z(f)$ has a simple zero at $\widetilde{p}$.*
(ii) *If p is a node of C, $\text{Discr}_z(f)$ has a double zero at $\widetilde{p}$.*
(iii) *If p is a cusp, $\text{Discr}_z(f)$ has a triple zero at $\widetilde{p}$.*
(iv) *If p is a an ordinary flex point of C (**1.4.9**) with tangent line L_{pq}, $\text{Discr}_z(f)$ has a double zero at $\widetilde{p}$.*

proof. There are several ways to prove this, none especially simple. We'll use Hensel's Lemma. We set $x = 1$, to work in the standard affine open set $\mathbb{U}$ with coordinates y, z. In affine coordinates, the projection π is the map $(y, z) \to y$. We may suppose that p is the origin in $\mathbb{U}$. Its image $\widetilde{p}$ will be the point $y = 0$ of the affine y-line, and the intersection of the line L_{pq} with $\mathbb{U}$ will be the line $\widetilde{L} : \{y = 0\}$. We'll denote the defining polynomial of the curve C, restricted to $\mathbb{U}$, by $f(y, z)$ instead of $f(1, y, z)$. Let $\widetilde{f}(z) = f(0, z)$.

(i)–(iii) In these three cases, the polynomial $\widetilde{f}(z) = f(0, z)$ will have a double zero at $z = 0$, so we will have $\widetilde{f}(z) = z^2\widetilde{h}(z)$, with $\widetilde{h}(0) \neq 0$. Then z^2 and $\widetilde{h}(z)$ have no common root, so we may apply Hensel's Lemma to write $f(y, z) = g(y, z)h(y, z)$, where g and h are polynomials in z whose coefficients are analytic functions of y, defined for small y, g is monic, $g(0, z) = z^2$, and $h(0, z) = \widetilde{h}$. Then (1.8.18)

$$\text{Discr}_z(f) = \pm\,\text{Discr}_z(g)\,\text{Discr}_z(h)\,\text{Res}_z(g, h)^2 \qquad (1.10.7)$$

Since q is in general position, $\widetilde{h}$ will have simple zeros. Then $\text{Discr}_z(h)$ doesn't vanish at $y = 0$. Neither does $\text{Res}_z(g, h)$. So the orders of vanishing of $\text{Discr}_z(f)$ and $\text{Discr}_z(g)$ are equal. We replace f by g.

Since g is a monic quadratic polynomial, it will have the form

$$g(y, z) = z^2 + b(y)z + c(y)$$

The coefficients b and c are analytic functions of y, and $g(0, z) = z^2$. The discriminant $\text{Discr}_z(g) = b^2 - 4c$ is unchanged when we complete the square by the substitution of $z - \frac{1}{2}b$ for z, and if $\widetilde{p}$ is a node or a cusp, that property isn't affected by this change of coordinates (Lemma 1.5.8). So we may assume that g has the form $z^2 + c(y)$. The discriminant is $D = 4c(y)$.

We write $c(y)$ as a series in y:

$$c(y) = c_0 + c_1y + c_2y^2 + c_3y^3 + \cdots$$

The constant coefficient c_0 is zero because $\widetilde{p}$ is a point of C. If $c_1 \neq 0$, $\widetilde{p}$ is a smooth point with tangent line $\widetilde{L} : \{y = 0\}$, and D has a simple zero. If $\widetilde{p}$ is a node, $c_0 = c_1 = 0$ and $c_2 \neq 0$. Then D has a double zero. If $\widetilde{p}$ is a cusp, $c_0 = c_1 = c_2 = 0$, and $c_3 \neq 0$. Then D has a triple zero at $\widetilde{p}$.

(iv) In this case, the polynomial $\widetilde{f}(z) = f(0, z)$ will have a triple zero at $z = 0$. Proceding as above, we may factor: $f = gh$ where g and h are polynomials in z with analyic coefficients in y, and $g(y, z) = z^3 + a(y)z^2 + b(y)z + c(y)$. We eliminate the quadratic coefficient a by substituting $z - \frac{1}{3a}$ for z. With $g = z^3 + az + b$, the discriminant $\text{Discr}_z(g)$ is $4b^3 + 27c^2$ (1.8.14). We write $c(y) = c_0 + c_1y + \cdots$ and $b(y) = b_0 + b_1y + \cdots$. Since p is a point of C with tangent line $\{y = 0\}$, $c_0 = 0$ and $c_1 \neq 0$. Since the intersection multiplicity of C with the line $\{y = 0\}$ at $\widetilde{p}$ is three, $b_0 = 0$. The discriminant has a zero of order two. □

1.10.8. Corollary. *Let $C : \{g = 0\}$ and $D : \{h = 0\}$ be plane curves that intersect transversally at a point $p = (x_0, y_0, z_0)$. With coordinates in general position,* $\text{Res}_z(g, h)$ *has a simple zero at* (x_0, y_0).

Two curves are said to intersect *transversally* at a point p if they are smooth at p and their tangent lines there are distinct.

proof. Proposition 1.10.6 **(ii)** applies to the product fg, whose zero locus is the union $C \cup D$. It shows that the discriminant $\text{Discr}_z(fg)$ has a double zero at $\widetilde{p}$. We also have the formula (1.10.7) with $f = gh$. Since coordinates are in general position, $\text{Discr}_z(g)$ and $\text{Discr}_z(h)$ will not be zero at $\widetilde{p}$. Then $\text{Res}_z(g, h)$ has a simple zero there. □

1.11 Genus

We describe the topological structure of smooth plane curves in the classical topology here, deferring the proof of one statement.

1.11.1. Theorem. *The smooth projective plane curves of a given degree d in $\mathbb{P}^2$ are homeomorphic manifolds of dimension two. They are compact, orientable and connected.*

The fact that a smooth curve is a two-dimensional manifold follows from the Implicit Function Theorem. (See the discussion at (**1.4.3**)).

orientability: A two-dimensional manifold is orientable if one can choose one of its two sides in a continuous, consistent way. A smooth curve C is orientable because its tangent space at a point is a one-dimensional complex vector space – the affine line with the equation (1.4.12). Multiplication by i orients the tangent space by defining the counterclockwise rotation. Then the right-hand rule tells us which side of C is "up".

compactness: A plane projective curve is compact because it is a closed subset of the compact space $\mathbb{P}^2$.

The connectedness of a plane curve is a subtle fact whose proof mixes topology and algebra. Unfortunately, I don't know a proof that fits into our discussion here. It will be proved later (see Theorem 8.4.9).

If one wants to have a proof now, one can begin by showing that the Fermat curve $x^d + y^d + z^d = 0$ is connected, by studying the projection to $\mathbb{P}^1$ from the point $(0, 0, 1)$. I propose this as an exercise. Then one can show that every plane curve is connected by proving a plausible fact: If a family C_t of smooth plane projective curves of degree d is parametrized by t in an interval of the real line, the curves in the family are homeomorphic. This can be proved using a *gradient flow*, If you are interested in following this up, read about gradient flows. However, the approach has two drawbacks: It leads us far afield, and it applies only to plane curves.

The topological *Euler characteristic* e of a compact, orientable two-dimensional manifold M is the alternating sum $b^0 - b^1 + b^2$ of its Betti numbers. (The *Betti number* b^i is the rank of the ith homology group of M in the classical topology.) The Euler characteristic depends only on the topological structure of M, and it can be computed in terms of a *topological triangulation*, a subdivision of M into topological triangles, called *faces*, by the formula

(1.11.2) $$e = |\text{vertices}| - |\text{edges}| + |\text{faces}|$$

For example, a sphere is homeomorphic to a tetrahedron, which has four vertices, six edges, and four faces. Its Euler characteristic is $4 - 6 + 4 = 2$. Any other topological triangulation of a sphere, such as the one given by the icosahedron, yields the same Euler characteristic.

Every compact, connected, orientable two-dimensional manifold is homeomorphic to a sphere with a finite number of holes, or "handles". Its *genus* is the number of holes. A torus has one hole. Its genus is one. The projective line $\mathbb{P}^1$, which is a two-dimensional sphere, has genus zero.

$Figure$

The Euler characteristic and the genus are related by the formula

(1.11.3) $$e = 2 - 2g$$

The Euler characteristic of a torus is zero, and the Euler characteristic of $\mathbb{P}^1$ is two.

To compute the the Euler characteristic of a smooth curve C of degree d, we analyze a generic projection to represent C as a branched covering of the projective line: $C \xrightarrow{\pi} \mathbb{P}^1$.

figure

We choose generic coordinates x, y, z in $\mathbb{P}^2$ and project form the point $q = (0, 0, 1)$. When the defining equation of C is written as a monic polynomial in z,

$$f = z^d + c_1 z^{d-1} + \cdots + c_d$$

where c_i is a homogeneous polynomial of degree i in the variables x, y, the discriminant $\mathrm{Discr}_z(f)$ with respect to z will be a homogeneous polynomial of degree $d(d-1) = d^2 - d$ in x, y.

If $\widetilde{p}$ is the image in $\mathbb{P}^1$ of a point p of C, the covering $C \xrightarrow{\pi} \mathbb{P}^1$ will be branched at $\widetilde{p}$ when the tangent line at p is the line L_{pq} through p and the center of projection q. If so, C and L_{pq} will have $d-1$ intersections (1.10). Proposition 1.10.6 tells us that the discriminant $\mathrm{Discr}_z(f)$ has a simple zero at the image of a tangent line. So there will be $d^2 - d$ points $\widetilde{p}$ in $\mathbb{P}^1$ over which the discriminant vanishes. They are the branch points of the covering. All other fibres consist of d points.

We triangulate the sphere $\mathbb{P}^1$ in such a way that the branch points are among the vertices, and we use the inverse images of the vertices, edges, and faces to triangulate C. Then C will have d faces and d edges lying over each face and each edge of $\mathbb{P}^1$, respectively. There will also be d vertices of C lying over a vertex $\widetilde{p}$ of $\mathbb{P}^1$, except when $\widetilde{p}$ is one of the $d^2 - d$ branch points. In that case the the fibre will contain only $d-1$ vertices. The Euler characteristic of C is obtained by multiplying the Euler characteristic of $\mathbb{P}^1$ by d and subtracting the number of branch points.

(1.11.4) $$e(C) = d\, e(\mathbb{P}^1) - (d^2 - d) = 2d - (d^2 - d) = 3d - d^2$$

This is the Euler characteristic of any smooth curve of degree d, so we denote it by e_d:

(1.11.5) $$e_d = 3d - d^2$$

Formula (1.11.3) shows that the genus g_d of a smooth curve of degree d is

(1.11.6) $$g_d \;=\; \tfrac{1}{2}(d-1)(d-2) = \binom{d-1}{2}$$

Thus smooth curves of degrees $1, 2, 3, 4, 5, 6, ...$ have genus $0, 0, 1, 3, 6, 10, ...$, respectively. A smooth plane curve cannot have genus two.

1.12 Bézout's Theorem

Bézout's Theorem counts intersections of plane curves. We state it here in a form that is ambiguous because it contains a term "multiplicity" that hasn't yet been defined.

1.12.1. Bézout's Theorem. Let C and D be distinct curves of degrees m and n, respectively. When intersections are counted with the appropriate multiplicity, the number of intersections is equal to mn. Moreover, the multiplicity at a point is 1 at a transversal intersection.

As before, C and D intersect transversally at p if they are smooth at p and their tangent lines there are distinct.

1.12.2. Corollary. *Bézout's Theorem is true when one of the curves is a line.*

See Corollary 1.3.9. The multiplicity of intersection of a curve and a line is the one that was defined there. □

The proof in the general case requires some algebra that we would rather defer. It will be given later (Theorem 7.8.1). It is possible to determine the intersections by counting the zeros of the resultant with respect to one of the variables. To do this, one chooses generic coordinates x, y, z, Then neither C nor D contains the point $(0, 0, 1)$. One writes their defining polynomials f and g as polynomials in z with coefficients in $\mathbb{C}[x, y]$.

The resultant R with respect to z will be a homogeneous polynomial in x, y, of degree mn. It will have mn zeros in $\mathbb{P}^1_{x,y}$, counted with multiplicity. If $\widetilde{p} = (x_0, y_0)$ is a zero of R, $f(x_0, y_0, z)$ and $g(x_0, y_0, z)$, which are polynomials in z, have a common root $z = z_0$, and then $p = (x_0, y_0, z_0)$ will be a point of $C \cap D$. It is a fact that the multiplicity of the zero of the resultant R at the image $\widetilde{p}$ is the (as yet undefined) intersection multiplicity of C and D at p. Unfortunately, this won't be obvious, even when multiplicity is defined. However, one can prove the next proposition using this approach.

1.12.3. Proposition. *Let C and D be distinct plane curves of degrees m and n, respectively.*

(i) *The curves C and D have at least one point of intersection, and the number of intersections is at most mn.*

(ii) *If all intersections are transversal, the number of intersections is precisely mn.*

It isn't obvious that two curves in the projective plane intersect. If two curves in the affine plane have no intersection, if they are parallel lines, for instance, their closures in the projective plane meet on the line at infinity.

1.12.4. Lemma. *Let f and g be homogeneous polynomials in x, y, z of degrees m and n, respectively, and suppose that the point $(0, 0, 1)$ isn't a zero of f or g. If the resultant $\mathrm{Res}_z(f, g)$ with respect to z is identically zero, then f and g have a common factor.*

proof. Let the degrees of f and g be m and n, respectively, and let F denote the field of rational functions $\mathbb{C}(x, y)$. If the resultant is zero, f and g have a common factor in $F[z]$ (Corollary 1.8.6). There will be polynomials p and q in $F[z]$, of degrees at most $n-1$ and $m-1$ in z, respectively, such that $pf = qg$ (1.8.2). We may clear denominators, so we may assume that the coefficients of p and q are in $\mathbb{C}[x, y]$. Then $pf = qg$ is an equation in $\mathbb{C}[x, y, z]$. Since p has degree at most $n-1$ in z, it isn't divisible by g, which has degree n in z. Since $\mathbb{C}[x, y, z]$ is a unique factorization domain, f and g have a common factor. □

proof of Proposition 1.12.3. **(i)** Let f and g be irreducible polynomials whose zero sets C and D, are distinct. Proposition 1.3.11 shows that there are finitely many intersections. We project to $\mathbb{P}^1$ from a point q that doesn't lie on any of the finitely many lines through pairs of intersection points. Then a line through q passes through at most one intersection, and the zeros of the resultant $\mathrm{Res}_z(f, g)$ that correspond to the intersection points will be distinct. Since the resultant has degree mn (1.8.8), it has at least one zero, and at most mn of them. Therefore C and D have at least one and at most mn intersections.

(ii) Every zero of the resultant will be the image of an intersection of C and D. To show that there are mn intersections if all intersections are transversal, it suffices to show that the resultant has simple zeros. This is Corollary 1.10.8. □

1.12.5. Corollary. *If the curve X defined by a homogeneous polynomial $f(x, y, z)$ is smooth, then f is irreducible, and therefore X is a smooth curve.*

proof. Suppose that $f = gh$, and let p be a point of intersection of the loci $\{g = 0\}$ and $\{h = 0\}$. The previous proposition shows that such a point exists. All partial derivatives of f vanish at p, so p is a singular point of X. □

1.12.6. Corollary. **(i)** *Let d be an integer ≥ 3. A smooth plane curve of degree d has at least one flex point, and the number of flex points is at most $3d(d-2)$.*

(ii) *If all flex points are ordinary, the number of flex points is equal to $3d(d-2)$.*

Thus smooth curves of degrees $2, 3, 4, 5, \ldots$ have at most $0, 9, 24, 45, \ldots$ flex points, respectively.

proof. **(i)** The flex points are intersections of a smooth curve C with its *Hessian divisor* $D : \{\det H = 0\}$. (The definition of divisor is given in (1.3.12.) Let $C : \{f(x_0, x_1, x_2) = 0\}$ be a smooth curve of degree d. The entries of the 3×3 Hessian matrix H are the second partial derivatives $\frac{\partial^2 f}{\partial x_i \partial x_j}$. They are homogeneous polynomials of degree $d-2$, so the Hessian determinant is homogeneous, of degree $3(d-2)$. Propositions 1.4.17 and 1.12.3 tell us that there are at most $3d(d-2)$ intersections.

(ii) Recall that a flex point is ordinary if the multiplicity of intersection of the curve and its tangent line is 3. Bézout's Theorem asserts that the number of flex points is equal to $3d(d-2)$ if the intersections of C with its Hessian divisor D are transversal, and therefore have multiplicity 1. So the next lemma completes the proof.

1.12.7. Lemma. *A curve $C : \{f = 0\}$ intersects its Hessian divisor D transversally at a point p if and only p is an ordinary flex point of C.*

proof. We prove this by computation. There may be a conceptual proof, but I don't know one.

Let L be the tangent line to C at the flex point p, and let h denote the restriction of the Hessian determinant to L. The Hessian divisor D will be transversal to C at p if and only if it is transversal to L, and this will be true if and only if the order of vanishing of h at p is 1.

We adjust coordinates x, y, z so that $p = (0, 0, 1)$ and L is the line $\{y = 0\}$, and we write the polynomial f of degree d as

(1.12.8)
$$f(x, y, z) = \sum_{i+j+k=d} a_{ij} x^i y^j z^k,$$

We set $y = 0$ and $z = 1$, to restrict f to L. The restricted polynomial is

$$f(x, 0, 1) = \sum_{i \le d} a_{i0} x^i$$

Since p is a flex point with tangent line L, the coefficients a_{00}, a_{10}, and a_{20} are zero, and p is an ordinary flex point if and only if the coefficient a_{30} is nonzero.

Let h be the restiction of $\det H$ to L: $h = \det H(x, 0, 1)$. We must show that p is an ordinary flex point if and only if h has a simple zero at $x = 0$.

To evaluate the restriction $f_{xx}(x, 0, 1)$ of the partial derivative to L, the relevant terms in the sum (1.12.8) have $j = 0$. Since $a_{00} = a_{10} = 0$,

$$f_{xx}(x, 0, 1) = 6a_{30} + 12a_{40}x^2 + \cdots = 6a_{30}x + O(2)$$

Similarly,

$$f_{xz}(x, 0, 1) = 0 + O(2)$$
$$f_{zz}(x, 0, 1) = 0 + O(2)$$

For the restriction of f_{yz}, the relevant terms are those with $j = 1$:

$$f_{yz}(x, 0, 1) = (d-1)a_{01} + (d-2)a_{11}x + O(2)$$

We don't need f_{xy} or f_{yy}.

Let $v = 6a_{30}x$ and $w = (d-1)a_{01} + (d-2)a_{11}x$. The restricted Hessian matrix has the form

(1.12.9)
$$H(x, 0, 1) = \begin{pmatrix} v & * & 0 \\ * & * & w \\ 0 & w & 0 \end{pmatrix} + O(2)$$

where $*$ are entries that don't affect terms of degree at most one in the determinant. The determinant is

$$h = -vw^2 + O(2) = -6(d-1)^2 a_{30} a_{01}^2 x + O(2)$$

It has a zero of order 1 at $x = 0$ if and only if a_{30} and a_{01} aren't zero. Since C is smooth at p and $a_{10} = 0$, the coefficient a_{01} isn't zero. Thus the curve C and its Hessian divisor D intersect transversally, and C and L intersect with multiplicity 3, if and only if a_{30} is nonzero, which is true if and only if p is an ordinary flex. □

1.12.10. Corollary. *A smooth cubic curve contains exactly 9 flex points.*

proof. Let f be the irreducible cubic polynomial whose zero locus is a smooth cubic C. The degree of the Hessian divisor D is also 3, so Bézout predicts at most 9 intersections of D with C. To derive the corollary, we show that C intersects D transversally. According to Proposition 1.12.7, a nontransversal intersection would correspond to a point at which the curve and its tangent line intersect with multiplicity greater than 3. This is impossible when the curve is a cubic. □

(1.12.11) singularities of the dual curve

Let C be a plane curve. As before, an *ordinary flex point* is a smooth point p such that the intersection multiplicity of the curve and its tangent line L at p is precisely 3. A *bitangent* to C is a line L that is tangent to C at distinct smooth points p and q, and an *ordinary bitangent* is one such that neither p nor q is a flex point. A tangent line L at a smooth point p of C is an *ordinary tangent* if it isn't a flex point or a bitangent.

The line L will have other intersections with C. Most often, these other intersections will be transversal. However, it may happen that L is tangent to C at such a point, or that it is a singular point of C. Let's call such occurences *accidents*.

1.12.12. Proposition. *Let p be a smooth point of a curve C, and let L be the tangent line at p. Suppose that there are no accidents.*
(i) *If L is an ordinary tangent at p, then L^* is a smooth point of C^*.*
(ii) *If L is an ordinary bitangent, then L^* is a node of C^*.*
(iii) *If p is an ordinary flex point, then L^* is a cusp of C^*.*

proof. We refer to the map $U \xrightarrow{t} C^*$ (**1.7.3**) from the set of smooth points of C to the dual curve. We set $z = 1$ and choose affine coordinates so that p is the origin, and the tangent line L at p is the line $\{y = 0\}$. Let $\widetilde{f}(x, y) = f(x, y, 1)$. We solve $\widetilde{f} = 0$ for $y = y(x)$ as analytic function of x near zero, as before. The tangent line L_1 to C at a nearby point $p_1 = (x, y)$ has the equation (1.7.11), and L_1^* is the point $(u, v, w) = (-y', 1, y'x - y)$ of $\mathbb{P}^*$ (1.7.12). Since there are no accidents, this path traces out all points of C^* near to L^*.

If L is an ordinary tangent line, $y(x)$ will have a zero of order 2 at $x = 0$. Then $u = -y'$ will have a simple zero. So the path $(-y', 1, y'x - y)$ is smooth at $x = 0$, and therefore C^*, is smooth at the origin.

If L is an ordinary bitangent, tangent to C at two points p and p', the reasoning given for an ordinary tangent shows that the images in C^* of small neighborhoods of p and p' in C will be smooth at L^*. Their tangent lines p^* and p'^* will be distinct, so p is a node.

The case that p is an ordinary flex point of C is trickier. Most probably, we won't know the defining equation $f = 0$ of C. We write the analytic function $y(x)$ that solves $f(x, y) = 0$ as a power series. Since p is a flex point, the coefficients of x^i are zero when $i < 3$: $y(x) = cx^3 + \cdots$, and since the flex is ordinary, we may assume that $c = 1$. In the local equation $(u, v, w) = (-y', 1, y'x - y)$ for the dual curve, $u = -3x^2 + \cdots$ and $w = 2x^3 + \cdots$. In affine u, w-space, the locus

(1.12.13) $$(u, w) = (-y', y'x - y) = (-3x^2 + \cdots, 2x^3 + \cdots)$$

contains the points of C^* near to L^*.

Let X and U denote the x-line and the u-line, respectively. We substitute (1.12.13) for u and v: $u = -3x^2 + \cdots$ and $w = 2x^3 + \cdots$. This gives us a diagram of maps

$$\begin{array}{ccc} X & \xrightarrow{b} & U \\ {\scriptstyle a}\downarrow & & \| \\ C^* & \xrightarrow{c} & U \end{array}$$

that are defined in small neighborhoods of the origins in the three spaces. The map a is locally bijective, and since the leading term of $u(x)$ is $3x^2$, b has degree 2. Therefore c also has degree 2. This implies that the origin in C^* is a point of multiplicity 2, a double point.

Let $g(u, w) = \sum_{ij} g_{ij} u^i w^j$ be the irreducible polynomial equation for C^*. Substituting for u and w, the series in x that we obtain evaluates to zero for all small x, and this implies that it is the zero series. The orders of vanishing of the monomials $u^i w^j$ as functions of x are as follows:

(1.12.14)

1	u	w	u^2	uw	w^2	u^3	u^2w	uw^2	w^3	$\cdots$
0	2	3	4	5	6	6	7	8	9	$\cdots$

Looking at these orders of vanishing, one sees that the coefficients $g_{00}, g_{10}, g_{01}, g_{20}$ and g_{11} in the series $g = \sum g_{ij} u^i w^j$ must be zero, and that $g_{02} + g_{30} = 0$. Since the origin is a double point of C^*, $g_{02} \neq 0$, and therefore $g_{30} \neq 0$. The origin is a cusp of C^*. $\square$

figure

1.12.15. Corollary. *A plane curve has finitely many bitangents.*

This corollary is true whether or not the bitangents are ordinary. It follows from the fact that the dual curve C^* has finitely many singular points (1.4.8). If L is a bitangent, ordinary or not, L^* will be a singular point of C^*. $\square$

1.13 The Plücker Formulas

A plane curve C is *ordinary* if it is smooth, if all of its bitangents and flex points are ordinary (see (**1.12.11**), and if there are no accidents. The *Plücker formulas* compute the number of flexes and bitangents of an ordinary plane curve.

For the next proposition, we refer back to the notation of Section 1.10. With coordinates in general position, let $\pi : C \to X$ be the projection of a plane curve C to the projective line X from $q = (0, 0, 1)$. If $\widetilde{p} = (x_0, y_0)$ is a point of X, we denote by $L_{\widetilde{p}}$ the line in $\mathbb{P}^2$ such that the fibre of π over $\widetilde{p}$ is the complement of q in $L_{\widetilde{p}}$.

The covering π will be branched at the points $\widetilde{p} = (x_0, y_0)$ of X such that $L_{\widetilde{p}}$ tangent line to C at some point. It will also be branched the images of singular points of C.

1.13.1. Proposition. *Let C be a plane curve, projected to $\mathbb{P}^1$ from a generic point q of the plane. With notation as above:*

(i) *The number β of points $\widetilde{p}$ such that line $L_{\widetilde{p}}$ is tangent to C at a smooth point is equal to the degree d^* of the dual curve C^*.*

(ii) *If C is a smooth curve of degree d, the degree d^* of C^* is $d^2 - d$.*

proof. **(i)** We have three numbers that we will show are equal: the number β referred to in the proposition, the degree d^* of the dual curve C^*, and the number N of intersections of the line q^* with C^*.

Let L be a line in $\mathbb{P}$ that contains q and is tangent to C at a smooth point p. Then L^* is one of the N points of $q^* \cap C^*$. Since q is generic, q isn't a point of C, L isn't a bitangent or a flex tangent, and L doesn't pass through a singular point of C (1.10.5). So L is an ordinary tangent line to C at p, and $q \neq p$. The number of such lines is β. So $\beta = N$.

Proposition 1.12.12 and Lemma 1.7.15 tell us that C^* is smooth at L^*, and that the tangent line to C^* at L^* is p^*. Since p and q are points of L and $q \neq p$, L^* is the intersection $p^* \cap q^*$. So C^* intersects q^* transversally at L^*. The intersection multiplicities are equal to 1. Therefore $N = \deg\, C^* = d^*$.

(ii) When we consider a smooth curve C as a branched covering of $\mathbb{P}^2$ by projection from q, the branch points are the images of tangent lines through q, and those tangent lines are ordinary. The discriminant of the defining polynomial f with respect to the chosen variable z will have degree d^2–d. There will be $d^2 - d$ ordinary tangent lines through q, so $d^* = d^2 - d$. $\square$

1.13.2. Theorem: Plücker Formulas. *Let C be an ordinary curve of degree d at least two, and let C^* be its dual curve. Let f and b denote the numbers of flex points and bitangents of C, and let δ^* and κ^* denote the numbers of nodes and cusps of C^*, respectively. Then:*

(i) *The dual curve C^* has no flexes or bitangents. Its singularities are nodes and cusps.*

(ii) $f = \kappa^* = 3d(d-2)$, *and* $b = \delta^* = \frac{1}{2}d(d-2)(d^2-9)$.

proof. **(i)** A bitangent or a flex on C^* would produce a singularity on the bidual C^{**}, which is the smooth curve C.

(ii) Bézout's Theorem counts the flex points (see (1.12.6)). The facts that $\kappa^* = f$ and $\delta^* = b$ are dealt with in Proposition 1.12.12. Thus $\kappa^* = f = 3d(d-2)$.

We project C^* to $\mathbb{P}^1$ from a generic point s of $\mathbb{P}^*$. Let β^* be the number of branch points that correspond to tangent lines through s at smooth points of C^*. Since $C^{**} = C$, Proposition 1.13.1, applied to C^*, tells us that $\beta^* = d$, and that $d^* = d^2 - d$.

Next, let F be the defining polynomial for C^*. The discriminant $\text{Discr}_z(F)$ has degree $d^{*2} - d^*$. Proposition 1.10.6 describes the order of vanishing of the discriminant at the images of the β tangent lines, the δ nodes, and the κ cusps of C^*. It tells us that

$$d^{*2} - d^* = \beta^* + 2\delta^* + 3\kappa^*$$

Substituting the known values $d^* = d^2 - d$, $\beta^* = d$, and $\kappa^* = 3d(d-2)$ into this formula gives us

$$(d^2 - d)^2 - (d^2 - d) = d + 2\delta^* + 9d(d-2) \quad \text{or} \quad 2\delta^* = (d^2 - 2d)(d^2 - 9) \qquad \square$$

Note. It isn't easy to count the number of bitangents directly.

1.13.3. Examples.

(i) All curves of degree 2 and all smooth curves of degree 3 are ordinary.

(ii) A curve of degree 2 has no flexes and no bitangents. Its dual curve has degree 2.

(iii) A smooth curve of degree 3 has 9 flexes and no bitangents. Its dual curve has degree 6.

(iv) An ordinary curve C of degree 4 has 24 flexes and 28 bitangents. Its dual curve has degree 12. $\square$

We will make use of the fact that a quartic curve has 28 bitangents in Chapter 4 (see (4.8.15)). The Plücker Formulas are rarely used for curves of degree greater than four.

Chapter 2 AFFINE ALGEBRAIC GEOMETRY

In the next chapters, we study varieties of dimension greater than one. We use some basic terminology that is introduced in Chapter 1, including the concepts of discriminant and transcendence degree, but most of the results of Chapter 1 won't be used again until Chapter 8.

We begin by reviewing some basic facts about rings and modules, omitting proofs. Please look up information on the concepts that aren't familiar, as needed.

Need to put localization of a module back##$

2.1 Rings and Modules

By the word 'ring', we mean 'commutative ring', $ab = ba$, unless when the contrary is stated explicitly. A *domain* is a ring that has no zero divisors and isn't the zero ring,

An *algebra* is a ring that contains the field $\mathbb{C}$ of complex numbers as subring. A set of elements $\alpha = \{\alpha_1, ..., \alpha_n\}$ *generates* gnerates an algebra A if every element of A can be expressed (usually not uniquely) as a polynomial in $\alpha_1, ..., \alpha_n$, with complex coefficients. Another way to state this is that α generates A if the homomorphism $\mathbb{C}[x_1, ..., x_n] \xrightarrow{\tau} A$ that evaluates a polynomial at α is surjective. If α generates A, then A will be isomorphic to the quotient $\mathbb{C}[x]/I$ of the polynomial algebra $\mathbb{C}[x]$, where I is the kernel of τ. A *finite-type algebra* is one that can be generated by a finite set of elements.

If I and J are ideals of a ring R, the *product ideal*, which is denoted by IJ, is the ideal whose elements are finite sums of products $\sum a_i b_i$, with $a_i \in I$ and $b_i \in J$. (This is not the product set, whose elements are the products ab, with $a \in I$ and $b \in J$.) The power I^k of I is the product of k copies of I, the ideal spanned by products of k elements of I. The intersection $I \cap J$ is also an ideal, and

(2.1.1) $$(I \cap J)^2 \subset IJ \subset I \cap J$$

An ideal M of a ring R is *maximal* if there is no ideal I such that $M < I < R$ and if it isn't the unit ideal R. This is true if and only if the quotient ring R/M is a field. An ideal P is a *prime ideal* if the quotient R/P is a domain. A maximal ideal is a prime ideal.

2.1.2. Lemma. *Let P be an ideal of a ring R, not the unit ideal. The following conditions are equivalent.*
(i) *P is a prime ideal.*
(ii) *If a and b are elements of R and if $ab \in P$, then $a \in P$ or $b \in P$.*
(iii) *If A and B are ideals of R, and if the product ideal AB is contained in P, then $A \subset P$ or $B \subset P$.* □

It is sometimes convenient to state **(iii)** this way:

(iii') *If A and B are ideals that contain P, and if the product ideal AB is contained in P, then $A = P$ or $B = P$.*

2.1.3. Mapping Property of Quotient Rings. *Let R and S be rings, let K be an ideal of R, and let $R \xrightarrow{\tau} \overline{R}$ denote the canonical map from R to the quotient ring $\overline{R} = R/K$. Homomorphisms $\overline{R} \xrightarrow{\overline{\varphi}} S$ correspond bijectively to homomorphisms $R \xrightarrow{\varphi} S$ whose kernels contain K, the correspondence being $\varphi = \overline{\varphi} \circ \tau$:*

$$\begin{array}{ccc} R & \xrightarrow{\varphi} & S \\ {\scriptstyle\tau}\downarrow & & \| \\ \overline{R} & \xrightarrow{\overline{\varphi}} & S \end{array}$$

If $\ker \varphi = I$, *then* $\ker \overline{\varphi} = I/K$. □

(2.1.4) commutative diagrams

In the diagram displayed above, the maps $\overline{\varphi}\tau$ and φ from R to S are equal. This is referred to by saying that the diagram is commutative. A *commutative diagram* is one in which every map that can be obtained by composing its arrows depends only on the domain and range of the map. In these notes, all diagrams of maps are commutative. We won't mention commutativity most of the time. □

2.1.5. Correspondence Theorem.
(i) *Let $R \xrightarrow{\varphi} S$ be a* ***surjective*** *ring homomorphism with kernel K. For instance, φ might be the canonical map from R to the quotient algebra R/K. There is a bijective correspondence*

$$\{\textit{ideals of } R \textit{ that contain } K\} \quad \longleftrightarrow \quad \{\textit{ideals of } S\}$$

This correspondence associates an ideal I of R that contains K with its image $\varphi(I)$ in S and it associates an ideal J of S with its inverse image $\varphi^{-1}(J)$ in R.

If an ideal I of R that contains K corresponds to an ideal J of S, then φ induces an isomorphism of quotient rings $R/I \to S/J$. So if one of the ideals, I or J, is prime or maximal, they both are.
(ii) *Let R be a ring, and let $M \xrightarrow{\varphi} N$ be a surjective homomorphism of R-modules with kernel L. There is a bijective correspondence*

$$\{\textit{submodules of } M \textit{ that contain } L\} \quad \longleftrightarrow \quad \{\textit{submodules of } N\}$$

This correspondence associates a submodule S of M that contains L with its image $\varphi(S)$ in N and it associates a submodule T of N with its inverse image $\varphi^{-1}(T)$ in M. □

Ideals $I_1, ..., I_k$ of a ring R are said to be *comaximal* if the sum of any two of them is the unit ideal.

2.1.6. Chinese Remainder Theorem. *Let $I_1, ..., I_k$ be comaximal ideals of a ring R.*
(i) *The product ideal $I_1 \cdots I_k$ is equal to the intersection $I_1 \cap \cdots \cap I_k$.*
(ii) *The map $R \longrightarrow R/I_1 \times \cdots \times R/I_k$ that sends an element a of R to its vector of residues is a surjective homomorphism whose kernel is $I_1 \cap \cdots \cap I_k \quad (= I_1 \cdots I_k)$.*
(iii) *Let M be an R-module. The canonical homomorphism $M \to M/I_1M \times \cdots \times M/I_kM$ is surjective.* □

2.1.7. Proposition. *Let R be a product of rings, $R = R_1 \times \cdots \times R_k$, let I be an ideal of R, and let $\overline{R} = R/I$ be the quotient ring. There are ideals I_j of R_j such that $I = I_1 \times \cdots \times I_k$ and $\overline{R} = R_1/I_1 \times \cdots \times R_k/I_k$.* □

(2.1.8) Noetherian rings

A *finite module* M over a ring R is a module that is spanned, or generated, by a finite set $\{m_1, ..., m_k\}$ of elements. To say that the set generates means that every element of M can be obtained as a combination

$r_1m_1 + \cdots + r_km_k$ with coefficients r_i in R, or that the homomorphism from the free R-module R^k to M that sends a vector $(r_1, ..., r_k)$ to the combination $r_1m_1 + \cdots + r_km_k$ is surjective.

An ideal of a ring R is *finitely generated* if, when regarded as an R-module, it is a finite module. A ring R is *noetherian* if all of its ideals are finitely generated.

The ring $\mathbb{Z}$ of integers is noetherian. Fields are notherian. If I is an ideal of a noetherian ring R, the quotient ring R/I is noetherian.

2.1.9. Hilbert Basis Theorem. *If R is a noetherian ring, the ring $R[x_1, ..., x_n]$ of polynomials with coefficients in R is noetherian.* □

Thus $\mathbb{Z}[x_1, ..., x_n]$ and $F[x_a, ..., x_n]$, F a field, are noetherian rings.

2.1.10. Corollary. *Every finite-type algebra is noetherian.* □

Note. It is important not to confuse the concept of a finite-type algebra with that of a finite module. A finite R-module M is a module in which every element can be written as a *(linear) combination* $r_1m_1 + \cdots + r_km_k$ of some finite set $\{m_1, ..., m_k\}$ of elements of M, with coefficients in R. A finite-type algebra A is an algebra in which every element can be written as a *polynomial* $f(\alpha_1, ..., \alpha_k)$ in some finite set of elements $\{\alpha_1, ..., \alpha_k\}$ of A, with complex coefficients.

(2.1.11) the ascending chain condition

The condition that a ring R be noetherian can be rewritten in several ways that we review here.

Our convention is that if X' and www are sets, the notation $X' \subset X$ means that X' is a subset of X, while $X' < X$ means that X' is a subset that is different from X. A *proper subset* X' of a set X is a nonempty subset different from X – a set such that $\emptyset < X' < X$.

A sequence $X_1, X_2, ...$, finite or infinite, of subsets of a set Z forms an *increasing chain* if $X_n \subset X_{n+1}$ for all n, equality $X_n = X_{n+1}$ being permitted. If $X_n < X_{n+1}$ for all n, the chain is *strictly increasing*.

Let $\mathcal{S}$ be a set whose elements are subsets of a set Z. A member M of $\mathcal{S}$ is a *maximal member* if there is no member M' of $\mathcal{S}$ such that $M < M'$. For example, the set of proper subsets of a set of five elements contains five maximal members, the subsets of order four. The set of finite subsets of the set of integers contains no maximal member.

A maximal ideal of a ring R is a maximal member of the set of ideals of R different from the unit ideal.

2.1.12. Proposition. *The following conditions on a ring R are equivalent:*
(i) *R is **noetherian**: Every ideal of R is finitely generated.*
(ii) *The **ascending chain condition**: Every strictly increasing chain $I_1 < I_2 < \cdots$ of ideals of R is finite.*
(iii) *Every nonempty set of ideals of R contains a maximal member.* □

The next corollary follows from the ascending chain condition, but the conclusions are true whether or not R is noetherian.

2.1.13. Corollary. *Let R be a noetherian ring.*
(i) *If R isn't the zero ring, every ideal of R except the unit ideal is contained in a maximal ideal.*
(ii) *A nonzero ring R contains at least one maximal ideal.*
(iii) *An element of a ring R that isn't in any maximal ideal is a unit – an invertible element of R.* □

2.1.14. Corollary. *Let $s_1, ..., s_k$ be elements that generate the unit ideal of a noetherian ring R. For any positive integer n, the powers $s_1^n, ..., s_k^n$ generate the unit ideal.* □

2.1.15. Proposition. *Let R be a noetherian ring, and let M be a finite R-module.*
(i) *Every submodule of M is a finite module.*
(ii) *The set of submodules of M satisfies the ascending chain condition.*
(iii) *Every nonempty set of submodules of M contains a maximal member.* □

This concludes our review of rings and modules.

2.2 The Zariski Topology

As before, the affine space $\mathbb{A}^n$ is the space of n-tuples $(a_1, ..., a_n)$ of complex numbers. Algebraic geometry studies polynomial equations in terms of their solutions in affine space. If $f_1, ..., f_k$ are polynomials in $x_1, ..., x_n$, the set of points of $\mathbb{A}^n$ that solve the system of equations

(2.2.1) $$f_1 = 0\,,\ \ldots,\ f_k = 0$$

is a *Zariski closed* subset of $\mathbb{A}^n$. A *Zariski open* subset U is a subset whose complement in $\mathbb{A}^n$, the set of points not in U, is Zariski closed.

When it seems unlikely to cause confusion, we may abbreviate the notation for an indexed set, using a single letter. The polynomial algebra $\mathbb{C}[x_1, ..., x_n]$ may be denoted by $\mathbb{C}[x]$, and the system of equations (2.2.1) by $f = 0$. The locus of solutions of the equations $f = 0$ may be denoted by $V(f_1, ..., f_k)$ or by $V(f)$. Its points are called the *zeros* of the polynomials f.

We use analogous notation for infinite sets. If $\mathcal{F}$ is any set of polynomials, $V(\mathcal{F})$ denotes the set of points of affine space at which all elements of $\mathcal{F}$ are zero. In particular, if I is an ideal of the polynomial ring, $V(I)$ denotes the set of points at which all elements of I vanish.

The ideal I *generated* by the polynomials $f_1, ..., f_k$ is the set of combinations $r_1f_1 + \cdots + r_kf_k$ with *polynomial* coefficients r_i. Some notations for this ideal are $(f_1, ..., f_k)$ and (f). All elements of this ideal vanish on the zero set $V(f)$, so $V(f) = V(I)$. The Zariski closed subsets of $\mathbb{A}^n$ are the sets $V(I)$, where I is an ideal.

We note a few simple relations among ideals and their zero sets here. To begin with, we note that an ideal I isn't determined by its zero locus $V(I)$. For any $k > 0$, the power f^k has the same zeros as f.

The *radical* of an ideal I of a ring R, which will be denoted by $\operatorname{rad} I$, is the set of elements α of R such that some power α^r is in I.

(**2.2.2**) $$\operatorname{rad} I = \{\alpha \in R \,|\, \alpha^r \in I \text{ for some } r > 0\}$$

The radical of I is an ideal that contains I. An ideal that is equal to its radical is a *radical ideal*. A prime ideal is a radical ideal.

The radical describes the ideals that define the same closed set.

2.2.3. Lemma. *If I is an ideal of the polynomial ring $\mathbb{C}[x]$, then $V(I) = V(\operatorname{rad} I)$.* □

Consequently, if I and J are ideals and if $\operatorname{rad} I = \operatorname{rad} J$, then $V(I) = V(J)$. The converse of this statement is also true: If $V(I) = V(J)$, then $\operatorname{rad} I = \operatorname{rad} J$. This is a consequence of the *Strong Nullstellensatz* that will be proved later in this chapter. (See (2.4.7).)

Because $(I \cap J)^2 \subset IJ \subset I \cap J$,

(2.2.4) $$\operatorname{rad}(IJ) = \operatorname{rad}(I \cap J)$$

and $\operatorname{rad}(I \cap J) = (\operatorname{rad} I) \cap (\operatorname{rad} J)$.

2.2.5. Lemma. *Let I and J be ideals of the polynomial ring $\mathbb{C}[x]$.*
(i) *If $I \subset J$, then $V(I) \supset V(J)$.*
(ii) $V(I^k) = V(I)$.
(iii) $V(I \cap J) = V(IJ) = V(I) \cup V(J)$.
(iv) *If I_ν are ideals, then $V(\sum I_\nu) = \bigcap V(I_\nu)$.*

proof. **(iii)** $V(I \cap J) = V(IJ)$ because the two ideals have the same radical, and because I and J contain IJ, $V(IJ) \supset V(I) \cup V(J)$. To prove that $V(IJ) \subset V(I) \cup V(J)$, we note that $V(IJ)$ is the locus of common zeros of the products fg with f in I and g in J. Suppose that a point p is a common zero: $f(p)g(p) = 0$ for all f in I and all g in J. If $f(p) \neq 0$ for some f in I, we must have $g(p) = 0$ for every g in J, and then p is a point of $V(J)$. If $f(p) = 0$ for all f in I, then p is a point of $V(I)$. In either case, p is a point of $V(I) \cup V(J)$. □

Zariski closed sets are the closed sets in the *Zariski topology* on $\mathbb{A}^n$. This topology is very useful in algebraic geometry, though it is very different from the classical topology.

To verify that the Zariski closed sets are the closed sets of a topology, one must show that

- the empty set and the whole space are Zariski closed,
- the intersection $\bigcap C_\nu$ of an arbitrary family of Zariski closed sets is Zariski closed, and
- the union $C \cup D$ of two Zariski closed sets is Zariski closed.

The empty set and the whole space are the zero sets of the elements 1 and 0, respectively. The other conditions follow from Lemma 2.2.5. □

2.2.6. Example. The proper Zariski closed subsets of the affine line, or of a plane affine curve, are finite sets. The proper Zariski closed subsets of the affine plane $\mathbb{A}^2$ are finite unions of points and curves. We omit the proofs of these facts. The corresponding facts for loci in the projective line and the projective plane have been noted before (see (**1.3.4**) and (1.3.14)). □

figure

(Caption: A Zariski closed subset of the affine plane (real locus).)

A subset S of a topological space X becomes a topological space with the *induced topology*. The closed (or open) subsets of S in the induced topology are intersections $S \cap Y$, where Y is closed (or open) in X.

The induced topology on a subset S of $\mathbb{A}^n$ will be called its *Zariski topology* too. A subset of S is closed in the Zariski topology if it has the form $S \cap Y$ for some Zariski closed subset Y of $\mathbb{A}^n$. If S itself is a Zariski closed subset of $\mathbb{A}^n$, a closed subset of S will be a closed subset of $\mathbb{A}^n$ that is contained in S.

Affine space also has a *classical topology*. A subset U of $\mathbb{A}^n$ is open in the classical topology if, whenever a point p is in U, all points sufficently near to p are in U. Since polynomial functions are continuous, their zero sets are closed in the classical topology. Therefore Zariski closed sets are closed in the classical topology too.

When two topologies T and T' on a set X are given, T' is said to be *coarser* than T if it contains fewer closed sets or fewer open sets, and *finer* than T if it contains more closed sets or more open sets. The Zariski topology is coarser than the classical topology. The next proposition shows that it is much coarser.

2.2.7. Proposition. *Every nonempty Zariski open subset of $\mathbb{A}^n$ is dense and path connected in the classical topology.*

proof. The (complex) *line* L through distinct points p and q of $\mathbb{A}^n$ is a Zariski closed set whose points can be written as $p + t(q - p)$, with t in $\mathbb{C}$. It corresponds bijectively to the one-dimensional affine t-space $\mathbb{A}^1$, and the Zariski closed subsets of L correspond to Zariski closed subsets of $\mathbb{A}^1$. They are the finite subsets of L, and L itself.

Let U be a nonempty Zariski open set, and let C be its Zariski closed complement. To show that U is dense in the classical topology, we choose distinct points p and q of $\mathbb{A}^n$, with p in U. If L is the line through p and q, $C \cap L$ will be a Zariski closed subset of L that doesn't contain p, a finite set. In the classical topology, the closure of the complement of this finite set, which is $U \cap L$, will be the whole line L. Therefore the closure of U contains q, and since q was arbitrary, the closure of U is $\mathbb{A}^n$.

Next, let L be the line through two points p and q of U. As before, $C \cap L$ will be a finite set. In the classical topology, L is a complex plane. The points p and q can be joined by a path in L that avoids a finite set. □

Though we will refer to the classical topology from time to time, the Zariski topology will appear more often. For this reason, we will refer to a Zariski closed subset simply as a *closed set*. Similarly, by an *open set* we mean a Zariski open set. We will mention the adjective "Zariski" only for emphasis.

(**2.2.8**) **irreducible closed sets**

The fact that the polynomial algebra is a noetherian ring has important consequences for the Zariski topology that we discuss here.

A topological space X satisfies the *descending chain condition* on closed subsets if there is no infinite, strictly descending chain $C_1 > C_2 > \cdots$ of closed subsets of X. The descending chain condition on closed subsets is equivalent with the *ascending chain condition* on open sets.

A topological space that satisfies the descending chain condition on closed sets is called a *noetherian space*. In a noetherian space, every nonempty family $\mathcal{S}$ of closed subsets has a minimal member, one that doesn't contain any other member of $\mathcal{S}$, and every nonempty family of open sets has a maximal member. (See (**2.1.11**).)

2.2.9. Proposition. *With its Zariski topology, $\mathbb{A}^n$ is a noetherian space.*

This follows from the ascending chain condition for ideals in $\mathbb{C}[x_1, ..., x_n]$. □

2.2.10. Definition. A topological space X is *irreducible* if it isn't the union of two proper closed subsets. Another way to say that X is irreducible is this:

If C and D are closed subsets of X, and if $X = C \cup D$, *then* $X = C$ *or* $X = D$.

The concept of irreducibility is useful primarily for noetherian spaces. The only irreducible subsets of a Hausdorff space are its points. In particular, with the classical topology, the only irreducible subsets of affine space are points.

The *closure* of a subset S of a topological space X is the smallest closed subset that contains S. The closure is the intersection of all closed subsets that contain S.

2.2.11. Lemma. **(i)** *The following conditions on topological space X are equivalent.*
- *X is irreducible.*
- *The intersection $U \cap V$ of two nonempty open subsets U and V of X is nonempty.*
- *Every nonempty open subset U of X is dense – its closure is X.*

(ii) *A noetherian topological space is quasicompact: Every open covering has a finite subcovering.* □

2.2.12. Lemma. **(i)** *Let Z be a subspace of a topological space X, let S be a subset of Z, and let $\overline{S}$ denote the closure of S in X. The closure of S in Z is the intersection $\overline{S} \cap Z$.*
(ii) *The closure $\overline{Z}$ of a subspace Z of a topological space X is irreducible if and only if Z is irreducible.*
(iii) *A nonempty open subspace W of an irreducible space X is irreducible.*

proof. **(ii)** Let Z be an irreducible subset of X, and suppose that its closure $\overline{Z}$ is the union $\overline{C} \cup \overline{D}$ of two closed sets $\overline{C}$ and $\overline{D}$. Then Z is the union of the sets $C = \overline{C} \cap Z$ and $D = \overline{D} \cap Z$, and they are closed in Z. Therefore Z is one of those two sets; say $Z = C$. Then $Z \subset \overline{C}$, and since $\overline{C}$ is closed, $\overline{Z} \subset \overline{C}$. Because $\overline{C} \subset \overline{Z}$ as well, $\overline{C} = \overline{Z}$. Conversely, suppose that the closure $\overline{Z}$ of a subset Z of X is irreducible, and that Z is a union $C \cup D$ of closed subsets. Then $\overline{Z} = \overline{C} \cup \overline{D}$, and therefore $\overline{Z} = \overline{C}$ or $\overline{Z} = \overline{D}$, say $\overline{Z} = \overline{C}$ So $Z = \overline{C} \cap Z = C$, and C is not a proper subset.

(iii) The closure of W is the irreducible space X. □

Irreducibility is somewhat analogous to connectedness. A topological space is *connected* if it isn't the union $C \cup D$ of two proper *disjoint* closed subsets. However, the condition that a space be irreducible is much more restrictive because, in Definition 2.2.10, the closed sets C and D aren't required to be disjoint. In the Zariski topology on the affine plane, the union of two intersecting lines is connected, but not irreducible.

2.2.13. Theorem. *In a noetherian topological space, every closed subset is the union of finitely many irreducible closed sets.*

proof. Let C_0 be a closed subset of a topological space X that isn't a union of finitely many irreducible closed sets. Then C_0 isn't irreducible, so it is a union $C_1 \cup D_1$, where C_1 and D_1 are proper closed subsets of C_0, and therefore closed subsets of X. Since C_0 isn't a finite union of irreducible closed sets, C_1 and D_1 cannot both be finite unions of irreducible closed sets. Say that C_1 isn't such a union. We have the beginning $C_0 > C_1$ of a chain of closed subsets. We repeat the argument, replacing C_0 by C_1, and we continue in this way, to construct an infinite, strictly descending chain $C_0 > C_1 > C_2 > \cdots$. So X isn't a noetherian space. □

2.2.14. Definition. An *affine variety* is an irreducible closed subset of affine space $\mathbb{A}^n$.

Theorem 2.2.13 tells us that every closed subset of $\mathbb{A}^n$ is a finite union of affine varieties. Since an affine variety is irreducible, it will be a connected set in the Zariski topology. It will also be connected in the classical topology, but this isn't very easy to prove. We may not get to it.

(2.2.15) the coordinate algebra of a variety

2.2.16. Proposition. *Let P be a radical ideal (**2.2.2**) of the polynomial algebra $\mathbb{C}[x_1, ..., x_n]$ and let V be the locus of zeros $V(P)$ in affine space $\mathbb{A}^n$. Then V is irreducible if and only if P is a prime ideal.*

Thus the affine varieties in $\mathbb{A}^n$ are the sets $V(P)$, where P is a prime ideal of the polynomial algebra $\mathbb{C}[x]$. We will use this proposition in the next section, but we defer the proof to Section 2.5.

As before, an *algebra* is a ring that contains the complex numbers.

2.2.17. Definition. Let P be a prime ideal of the polynomial ring $\mathbb{C}[x_1, ..., x_n]$, and let V be the affine variety $V(P)$ in $\mathbb{A}^n$. The *coordinate algebra* of V is the quotient algebra $A = \mathbb{C}[x]/P$.

Geometric properties of the variety are reflected in algebraic properties of its coordinate algebra and vice versa. In a primitive sense, one can regard the geometry of an affine variety V as given by closed subsets and incidence relations – the inclusion of one closed set into another, as when a point lies on a line. A finer study of the geometry takes into account things such as tangency, but it is reasonable to begin by studying incidences $C' \subset C$ among closed subvarieties. Such incidences translate into inclusions $P' \supset P$ in the opposite direction among prime ideals. This is one reason that prime ideals are important.

2.3 Some affine varieties

This section contains a few simple examples of varieties.

2.3.1. A *point* $p = (a_1, \ldots, a_n)$ of affine space $\mathbb{A}^n$ is the set of solutions of the n equations $x_i - a_i = 0$, $i = 1, \ldots, n$. A point is a variety because the polynomials $x_i - a_i$ generate a maximal ideal in the polynomial algebra $\mathbb{C}[x]$, and a maximal ideal is a prime ideal. We denote that maximal ideal by $\mathfrak{m}_p$. It is the kernel of the substitution homomorphism $\pi_p : \mathbb{C}[x] \to \mathbb{C}$ that evaluates a polynomial $g(x_1, ..., x_n)$ at p: $\pi_p(g(x)) = g(a_1, ..., a_n) = g(p)$. As here, we usually denote that homomorphism by π_p.

The coordinate algebra of a point p is the quotient algebra $\mathbb{C}[x]/\mathfrak{m}_p$. It is also called the *residue field* at p, and it will be denoted by $k(p)$. The residue field at p is isomorphic to the image of π_p, the field $\mathbb{C}$ of complex numbers, but $k(p)$ is a particular quotient of the polynomial ring.

2.3.2. The varieties in the affine line $\mathbb{A}^1$ are its points and the whole line $\mathbb{A}^1$. The varieties in the affine plane $\mathbb{A}^2$ are points, plane affine curves, and the whole plane.

This is true because the varieties correspond to the prime ideals of the polynomial ring. The prime ideals of $\mathbb{C}[x_1, x_2]$ are the maximal ideals, the principal ideals generated by irreducible polynomials, and the zero ideal. The proof of this is a good exercise.

2.3.3. The set X of solutions of a single irreducible polynomial equation $f_1(x_1, ..., x_n) = 0$ is a variety, called an *affine hypersurface*.

The *special linear group* SL_2, the group of complex 2×2 matrices with determinant 1, is a hypersurface in $\mathbb{A}^4$. It is the locus of zeros in $\mathbb{A}^4$ of the irreducible polynomial $x_{11}x_{22} - x_{12}x_{21} - 1$.

The reason that an affine hypersurface is a variety is that an irreducible element of a unique factorization domain is a prime element, and a prime element generates a prime ideal. The polynomial ring $\mathbb{C}[x_1, ..., x_n]$ is a unique factorization domain.

2.3.4. A hypersurface in the affine plane $\mathbb{A}^2$ is a plane affine curve.

A *line* in the plane, the locus of a linear equation $ax + by - c = 0$, is a plane affine curve. Its coordinate algebra is isomorphic to a polynomial ring in one variable. Every line is isomorphic to the affine line $\mathbb{A}^1$.

2.3.5. Let $p=(a_1,\dots,a_n)$ and $q=(b_1,\dots,b_n)$ be distinct points of $\mathbb{A}^n$. The *point pair* (p,q) is the closed set defined by the system of n^2 equations $(x_i-a_i)(x_j-b_j)=0$ with $1\le i,j\le n$. A point pair isn't a variety because the ideal I generated by the polynomials $(x_i-a_i)(x_j-b_j)$ isn't a prime ideal. The next proposition, which follows from the Chinese Remainder Theorem 2.1.6, describes the ideal I.

2.3.6. Proposition. *The ideal of polynomials that vanish on a point pair is the product of the maximal ideals $\mathfrak{m}_p\mathfrak{m}_q$, and the quotient algebra $\mathbb{C}[x]/I$ is isomorphic to the product algebra $\mathbb{C}\times\mathbb{C}$.* □

2.4 Hilbert's Nullstellensatz

2.4.1. Nullstellensatz (version 1). *Let $\mathbb{C}[x]$ be the polynomial algebra in the variables $x_1,\dots,x_n$. There are bijective correspondences between the following sets:*

- *points p of the affine space $\mathbb{A}^n$,*
- *algebra homomorphisms $\pi_p:\mathbb{C}[x]\to\mathbb{C}$,*
- *maximal ideals $\mathfrak{m}_p$ of $\mathbb{C}[x]$.*

If $p=(a_1,...,a_n)$ is a point of $\mathbb{A}^n$, the corresponding homomorphism π_p evaluates a polynomial at p: $\pi_p(g)=g(a_1,....,a_n)\ \ \big(=g(p)\big)$, and the maximal ideal $\mathfrak{m}_p$ is the kernel of π_p. It is generated by the linear polynomials $x_1-a_1,\dots,x_n-a_n$. □

It is obvious that every algebra homomorphism $\mathbb{C}[x]\to\mathbb{C}$ is surjective and that its kernel is a maximal ideal. It isn't obvious that every maximal ideal of $\mathbb{C}[x]$ is the kernel of such a homomorphism. The proof can be found manywhere.[1]

The Nullstellensatz gives us a way to describe the closed set $V(I)$ of zeros of an ideal I in affine space in terms of maximal ideals. The points of $V(I)$ are those at which all elements of I vanish. Thus

(2.4.2) $$V(I)=\{p\in\mathbb{A}^n\mid I\subset\mathfrak{m}_p\}$$

2.4.3. Proposition. *Let I be an ideal of the polynomial ring $R=\mathbb{C}[x]$. If the zero locus $V(I)$ is empty, then I is the unit ideal of R.*

proof. Every ideal I that is not the unit ideal is contained in a maximal ideal (Corollary 2.1.13). □

2.4.4. Strong Nullstellensatz. *Let I be an ideal of the polynomial algebra $\mathbb{C}[x_1,\dots,x_n]$, and let V be the locus of zeros of I in $\mathbb{A}^n$: $V=V(I)$. If a polynomial g vanishes at every point of V, then I contains a power of g.*

proof. This beautiful proof is due to Rainich. Let $g(x)$ be a polynomial that is identically zero on V. We are to show that I contains a power of g. If g is the zero polynomial, it is in I. So we may assume that g isn't zero.

The Hilbert Basis Theorem tells us that I is a finitely generated ideal; let $f=f_1,\dots,f_k$ be a set of generators. In the $n+1$–dimensional affine space with coordinates $(x_1,\dots,x_n,y)$, let W be the locus of solutions of the $k+1$ equations

(2.4.5) $$f_1(x)=\cdots=f_k(x)=0\quad\text{and}\quad g(x)y-1=0$$

Suppose that we have a solution x of the equations $f(x)=0$, say $(x_1,...,x_n)=(a_1,...,a_n)$. Then a is a point of V, and our hypothesis tells us that $g(a)=0$ too. So there can be no b such that $g(a)b=1$. There is no point $(a_1,...,a_n,b)$ that solves the equations (2.4.5): The locus W is empty. Proposition 2.4.3 tells us that the polynomials $f_1,...,f_k,gy-1$ generate the unit ideal of $\mathbb{C}[x_1,...,x_n,y]$. There are polynomials $p_1(x,y),\dots,p_k(x,y)$ and $q(x,y)$ such that

(2.4.6) $$p_1f_1+\cdots+p_kf_k+q(gy-1)=1$$

The ring $R=\mathbb{C}[x,y]/(gy-1)$ can be described as the one obtained by adjoining an inverse of g to the polynomial ring $\mathbb{C}[x]$. The residue of y is the inverse of g. Since g isn't zero, $\mathbb{C}[x]$ is a subring of R. In R,

[1] While writing a paper, the mathematician Nagata decided that the English language needed this word, and then he managed to find it in a dictionary.

$gy - 1 = 0$, so the equation (2.4.6) becomes $p_1 f_1 + \cdots + p_k f_k = 1$. When we multiply both sides of this equation by a large power g^N of g, we can use the equation $gy = 1$, which is true in R, to cancel all occurences of y in the polynomials $p_i(x, y)$. Let $h_i(x)$ denote the polynomial in x that is obtained by cancelling y in $g^N p_i$. Then

$$h_1(x)f_1(x) + \cdots + h_k(x)f_k(x) = g^N(x)$$

is a polynomial equation that is true in R and in its subring $\mathbb{C}[x]$. Since $f_1, ..., f_k$ are in I, this equation shows that g^N is in I. □

2.4.7. Corollary. *Let I and J be ideals of the polynomial ring $\mathbb{C}[x_1, ..., x_n]$.*

(i) *Let P be a prime ideal of $\mathbb{C}[x]$, and let $V = V(P)$ be the variety of zeros of P. If a polynomial g vanishes at every point of V, then g is an element of P.*

(ii) *Let f be an irreducible polynomial in $\mathbb{C}[x]$. If a polynomial g vanishes at every point of $V(f)$, then f divides g.*

(iii) $V(I) \supset V(J)$ *if and only if* $\operatorname{rad} I \subset \operatorname{rad} J$, *and* $V(I) > V(J)$ *if and only if* $\operatorname{rad} I > \operatorname{rad} J$ *(see* **(2.2.2)***).* □

As before, a finite-type algebra is an algebra that can be generated by a finite set of elements.

2.4.8. Nullstellensatz (version 2). Let A be a finite-type algebra. There are bijective correspondences between the following sets:

- algebra homomorphisms $\overline{\pi} : A \to \mathbb{C}$,
- maximal ideals $\overline{\mathfrak{m}}$ of A.

The maximal ideal $\overline{\mathfrak{m}}$ that corresponds to a homomorphism $\overline{\pi}$ is the kernel of $\overline{\pi}$.

If A is presented as a quotient of a polynomial ring, say $A \approx \mathbb{C}[x_1, ..., x_n]/I$, then these sets also correspond bijectively to points of the set $V(I)$ of zeros of I in $\mathbb{A}^n$.

(We use the symbol $\approx$ to indicate an isomorphism.)

proof. We choose a presentation of A as a quotient of a polynomial ring to identify A with a quotient $\mathbb{C}[x]/I$. The Correspondence Theorem tells us that maximal ideals of A correspond to maximal ideals of $\mathbb{C}[x]$ that contain I. Those maximal ideals correspond to points of $V(I)$ (see (2.4.2)).

Let τ denote the canonical homomorphism $\mathbb{C}[x] \to A$. The Mapping Property 2.1.3, applied to τ, tells us that homomorphisms $A \xrightarrow{\overline{\pi}} \mathbb{C}$ correspond to homomorphisms $\mathbb{C}[x] \xrightarrow{\pi} \mathbb{C}$ whose kernels contain I. Those homomorphisms also correspond to points of $V(I)$.

(2.4.9)
$$\begin{array}{ccc} \mathbb{C}[x] & \xrightarrow{\pi} & \mathbb{C} \\ {\scriptstyle\tau}\big\downarrow & & \big\| \\ A & \xrightarrow{\overline{\pi}} & \mathbb{C} \end{array}$$

□

2.5 The Spectrum

The Nullstellensatz allows us to associate a set of points to a finite-type domain A without reference to a presentation. We can do this because the maximal ideals of A and the homomorphisms $A \to \mathbb{C}$ don't depend on the presentation. If A is presented as a quotient $\mathbb{C}[x]/P$ of a polynomial ring, P a prime ideal, it becomes the coordinate algebra of the variety $V(P)$ in affine space. Then the points of $V(P)$ correspond to maximal ideals of A and also to homomorphisms $A \to \mathbb{C}$.

When a finite-type domain A is given without a presentation, we replace the variety $V(P)$ by an abstract set of points, the *spectrum* of A, that we denote by $\operatorname{Spec} A$ and call an *affine variety*. We put one point into the spectrum for every maximal ideal of A, and then we turn around and denote the maximal ideal that corresponds to a point p by $\overline{\mathfrak{m}}_p$. The Nullstellensatz tells us that p also corresponds to a homomorphism $A \to \mathbb{C}$ whose kernel is $\overline{\mathfrak{m}}_p$. We denote that homomorphism by $\overline{\pi}_p$. The domain A is the *coordinate algebra* of the affine variety $\operatorname{Spec} A$ (see (2.2.17)). To work with $\operatorname{Spec} A$, we may interpret its points as maximal ideals or as homomorphisms to $\mathbb{C}$, whichever is convenient.

When defined in this way, the variety $\operatorname{Spec} A$ isn't embedded into affine space, but if we present A as a quotient $\mathbb{C}[x]/P$, points of $\operatorname{Spec} A$ correspond to points of the subset $V(P)$ in $\mathbb{A}^n$. Even when the coordinate ring A of an affine variety is presented as $\mathbb{C}[x]/P$, we may denote the variety by $\operatorname{Spec} A$ rather than by $V(P)$.

Let $X = \operatorname{Spec} A$. The elements of A define (complex-valued) functions on X: A point p of X corresponds to a homomorphism $A \xrightarrow{\overline{\pi}_p} \mathbb{C}$. If α is an element of A, the *value* of the function α at p is defined to be $\overline{\pi}_p(\alpha)$:

(2.5.1)
$$\alpha(p) \stackrel{def}{=} \overline{\pi}_p(\alpha)$$

Then the kernel $\overline{\mathfrak{m}}_p$ of $\overline{\pi}_p$ is the set of elements α of the coordinate algebra A such that $\alpha(p) = 0$:

$$\overline{\mathfrak{m}}_p = \{\alpha \in A \,|\, \alpha(p) = 0\}$$

The functions defined by the elements of A are the *regular functions* on X. (See Proposition 2.6.2 below.)

For example, the spectrum $\operatorname{Spec} \mathbb{C}[x_1, ..., x_n]$ of the polynomial algebra is the affine space $\mathbb{A}^n$. The homomorphism $\pi_p : \mathbb{C}[x] \to \mathbb{C}$ that corresponds to a point $p = (a_1, ..., a_n)$ of $\mathbb{A}^n$ is evaluation at p. So $\pi_p(g) = g(a_1, ..., a_n) = g(p)$. The function defined by a complex polynomial $g(x)$ is the polynomial function.

2.5.2. Lemma. *Let A be a quotient $\mathbb{C}[x]/P$ of the polynomial ring $\mathbb{C}[x_1, ..., x_n]$ modulo a prime ideal P, so that $\operatorname{Spec} A$ becomes the closed subset $V(P)$ of $\mathbb{A}^n$. Then a point p of $\operatorname{Spec} A$ corresponds to a point $(a_1, ..., a_n)$ of $\mathbb{A}^n$. When an element α of A is represented by a polynomial $g(x)$, the value of α at p is $\alpha(p) = g(a_1, ..., a_n) = g(p)$.*

proof. The point p of $\operatorname{Spec} A$ gives us a diagram (2.4.9), with $\overline{\pi} = \overline{\pi}_p$ and $\pi = \pi_p$, and where τ is the canonical map $\mathbb{C}[x] \to A$. Then $\alpha = \tau(p)$, and

(2.5.3)
$$g(p) \stackrel{defn}{=} \pi_p(g) = \overline{\pi}_p \tau(g) = \overline{\pi}_p(\alpha) \stackrel{defn}{=} \alpha(p). \qquad \square$$

Thus the value $\alpha(p)$ at a point p of $\operatorname{Spec} A$ can be obtained by evaluating a polynomial g at p. However, the polynomial g that represents the regular function α won't be unique unless P is the zero ideal.

(2.5.4) the Zariski topology on an affine variety

Let $X = \operatorname{Spec} A$ be an affine variety with coordinate algebra A. An ideal $\overline{J}$ of A defines a locus in X, a *closed subset*, that we denote by $V_X(\overline{J})$:

(2.5.5)
$$V_X(\overline{J}) = \{p \in \operatorname{Spec} A \,|\, \overline{J} \subset \overline{\mathfrak{m}}_p\}$$

When a presentation $\mathbb{C}[x]/P \approx A$, is given, the ideal $\overline{J}$ of A corresponds to an ideal J of $\mathbb{C}[x]$ that contains P. Then if $V_{\mathbb{A}^n}(J)$ denotes the zero locus of J in $\mathbb{A}^n$, $V_X(\overline{J}) = V_{\mathbb{A}^n}(J)$.

The properties of closed sets in affine space that are given in Lemmas 2.2.3 and 2.2.5 are true for closed subsets of an affine variety. In particular, $V_X(\overline{J}) = V_X(\operatorname{rad} \overline{J})$, and $V_X(\overline{IJ}) = V_X(\overline{I} \cap \overline{J}) = V_X(\overline{I}) \cup V_X(\overline{J})$.

2.5.6. Proposition. *Let $\overline{J}$ be an ideal of a finite-type domain A, and let $X = \operatorname{Spec} A$. The zero set $V_X(\overline{J})$ is empty if and only if $\overline{J}$ is the unit ideal of A. If X is empty, then A is the zero ring.*

proof. The zero ring is the only ring with no maximal ideals. $\square$

2.5.7. Note. We have put bars on the symbols $\overline{\mathfrak{m}}$ and $\overline{\pi}$ here in order to distinguish maximal ideals of A from maximal ideals of $\mathbb{C}[x]$ and homomorphisms $A \to \mathbb{C}$ from homomorphisms $\mathbb{C}[x_1, \ldots, x_n] \to \mathbb{C}$. In the future, we will put bars over the letters only when there is a danger of confusion. $\square$

(2.5.8) ideals whose zero sets are equal

2.5.9. Lemma. *An ideal I of a noetherian ring R contains a power of its radical.*

proof. Since R is noetherian, the ideal $\operatorname{rad} I$ is generated by a finite set of elements $\alpha = \{\alpha_1, ..., \alpha_k\}$, and for large r, α_i^r is in I. We can use the same large integer r for every i. A monomial $\beta = \alpha_1^{e_1} \cdots \alpha_k^{e_k}$ of sufficiently large degree n in α will be divisible α_i^r for at least one i, and therefore it will be in I. The monomials of degree n generate $(\operatorname{rad} I)^n$, so $(\operatorname{rad} I)^n \subset I$ $\big($and $I \subset \operatorname{rad} I\big)$. □

2.5.10. Corollary. *Let I and J be ideals of a finite-type domain A, and let $X = \operatorname{Spec} A$. Then $V_X(I) \supset V_X(J)$ if and only if $\operatorname{rad} I \subset \operatorname{rad} J$.*

This follows from Corollary 2.4.7. □

For example, there is a bijective correspondence between radical ideals in the polynomial ring $\mathbb{C}[x_1, ..., x_n]$ and closed subsets of $\mathbb{A}^n$.

The next proposition includes Proposition 2.2.16 as a special case.

2.5.11. Proposition. *Let $X = \operatorname{Spec} A$, where A is a finite-type domain. The closed subset $V_X(P)$ defined by a radical ideal P is irreducible if and only if P is a prime ideal.*

proof. Let P be a radical ideal of A, and let $Y = V_X(P)$. Let C and D be closed subsets of X such that $Y = C \cup D$. Say $C = V_X(I)$, $D = V_X(J)$. We may suppose that I and J are radical ideals. Then the inclusion $C \subset Y$ implies that $I \supset P$. Similarly, $J \supset P$. Because $Y = C \cup D$, we also have $Y = V_X(I \cap J) = V_X(IJ)$. So $IJ \subset P$ (Corollary 2.5.10). If P is a prime ideal, then $I = P$ or $J = P$, and therefore $C = Y$ or $D = Y$. So Y is irreducible. Conversely, suppose that P is not a prime ideal. Then there are ideals A, B strictly larger than P, such that $AB \subset P$ (2.1.2). Then Y will be the union of the two proper closed subsets $V_X(A)$ and $V_X(B)$, and is not irreducible (2.5.10). □

2.5.12. Examples.
(i) Let I be the ideal generated by y^5 and $y^2 - x^3$ in the polynomial algebra $\mathbb{C}[x, y]$ in two variables. The origin $y = x = 0$ is the only common zero of these polynomials in the affine plane, and the polynomial x also vanishes at the origin. The Strong Nullstellensatz predicts that I contains a power of x. This is verified by the following equation:

$$yy^5 - (y^4 + y^2x^3 + x^6)(y^2 - x^3) \ = \ x^9$$

(ii) We may regard pairs A, B of $n \times n$ matrices as points of an affine space $\mathbb{A}^{2n^2}$ with coordinates a_{ij}, b_{ij}, $1 \le i, j \le n$. The pairs of *commuting matrices* ($AB = BA$) form a closed subset of $\mathbb{A}^{2n^2}$, the locus of common zeros of the n^2 polynomials c_{ij} that compute the entries of the matrix $AB - BA$:

(2.5.13)
$$c_{ij}(a, b) = \sum_\nu a_{i\nu} b_{\nu j} - b_{i\nu} a_{\nu j}$$

Let I denote the ideal of the polynomial algebra $\mathbb{C}[a, b]$ generated by the polynomials c_{ij}. Then $V(I)$ is the set of pairs of commuting complex matrices. The Strong Nullstellensatz asserts that if a polynomial $g(a, b)$ vanishes on every pair of commuting matrices, some power of g is in I. Is g itself in I? It is a famous conjecture that I is a prime ideal. If so, g would be in I. Proving the conjecture would establish your reputation as a mathematician, but I don't recommend spending very much time on it right now. □

(2.5.14) the nilradical

The *nilradical* of a ring is the set of its nilpotent elements. It is the radical of the zero ideal. The nilradical of a domain is the zero ideal. If a ring R is noetherian, its nilradical will be *nilpotent*: some power of will be the zero ideal (Lemma 2.5.9).

2.5.15. Proposition. *The nilradical of a noetherian ring R is the intersection of the prime ideals of R.*

proof. Let x be an element of the nilradical N. So some power of x is zero. Since the zero element is in every prime ideal, x is in every prime ideal. Therefore N is contained in every prime ideal. Conversely, let x be an element not in N, i.e., not nilpotent. We show that there is a prime ideal that doesn't contain any power of x. Let $\mathcal{S}$ be the set of ideals that don't contain a power of x. The zero ideal is one such ideal, so $\mathcal{S}$ isn't empty. Since R is noetherian, $\mathcal{S}$ contains a maximal member P (**2.1.11**). We show that P is a prime ideal by showing that, if two ideals A and B are strictly larger than P, their product AB isn't contained in P. Since P is a maximal member of $\mathcal{S}$, A and B aren't in $\mathcal{S}$. They contain powers of x, say $x^k \in A$ and $x^\ell \in B$. Then $x^{k+\ell}$ is in AB but not in P. Therefore $AB \not\subset P$. □

The conclusion of this proposition is true whether or not the ring R is noetherian.

2.5.16. Corollary.
(i) *Let A be a finite-type algebra. An element that is in every maximal ideal of A is nilpotent.*
(ii) *Let A be a finite-type domain. The intersection of the maximal ideals of A is the zero ideal.*

proof. **(i)** Say that A is presented as $\mathbb{C}[x]/I$. Let α be an element of A that is in every maximal ideal, and let $g(x)$ be a polynomial whose residue in A is α. Then α is in every maximal ideal of A if and only if $g = 0$ at all points of $V_{\mathbb{A}}(I)$. If so, the Strong Nullstellensatz asserts that some power g^n is in I. Then $\alpha^n = 0$. □

2.5.17. Corollary. *An element α of a finite-type domain A is determined by the function that it defines on $X = \operatorname{Spec} A$.*

proof. It is enough to show that an element α that defines the zero function is the zero element. Such an element is in every maximal ideal (2.5.6), so α is nilpotent, and since A is a domain, $\alpha = 0$. □

(**2.5.18**) localization

Let s be a nonzero element of a domain A. The ring $A[s^{-1}]$ obtained by adjoining an inverse of s to A is called a *localization* of A. The localization is isomorphic to the quotient $A[z]/(sz-1)$ of the polynomial ring $A[z]$ in the variable z by the principal ideal generated by $sz-1$, and it will be denoted by A_s. If A is a finite-type domain, the variety $\operatorname{Spec} A_s$ will also be called a *localization* of X, and it may be denoted by X_s.

2.5.19. Proposition. (i) *With terminology as above, points of the variety $X_s = \operatorname{Spec} A_s$ correspond bijectively to the open subset of X of points at which the value of s is nonzero.*
(ii) *When we identify a localization X_s with a subset of X, the Zariski topology on X_s is the induced topology from X. So X_s is an open subspace of X.*

proof. **(i)** Let p be a point of X, let $A \xrightarrow{\pi_p} \mathbb{C}$ be the corresponding homomorphism, and let $c = s(p)$ $\big(= \pi_p(s)\big)$. If $c \neq 0$, π_p extends uniquely to a homomorphism $A_s \to \mathbb{C}$ that sends s^{-1} to c^{-1}. This gives us a unique point of X_s whose image in X is p. On the other hand, if $c = 0$, then π_p doesn't extend to A_s. So X_s identifies with the complement of the zero locus of s.

(ii) If C is closed in $X = \operatorname{Spec} A$, say $C = V_X(I)$, then $C \cap X_s$ is the zero set of I in X_s. So it is closed in X_s. We must show that if D is a closed subset of X_s, there is a closed subset C of X such that $D = C \cap X_s$. Say that D is the zero set of an ideal J of A_s, and let $\alpha_1, ..., \alpha_k$ be generators for J that are contained in A. Let I be the ideal $(\alpha_1, ..., \alpha_k)A$, and let $C = V_X(I)$. If $\mathfrak{m}_p$ is the maximal ideal of A at a point p of X_s, the maximal ideal of A_s at p is the extended ideal $(\mathfrak{m}_p)_s$, the ideal of A_s generated by $\mathfrak{m}_p$. Its elements are fractions as^{-k}, with a in $\mathfrak{m}_p$. The point p is in D if and only if $J \subset (\mathfrak{m}_p)_s$, and this is true if and only if $I \subset \mathfrak{m}_p$, i.e., if and only if p is a point of C. □

We usually identify X_s as an open subset of X. Then the effect of adjoining the inverse is to throw out the points of X at which s vanishes. For example, the spectrum of the Laurent polynomial ring $\mathbb{C}[t, t^{-1}]$ becomes the complement of the origin in the affine line $\mathbb{A}^1 = \operatorname{Spec} \mathbb{C}[t]$.

This illustrates the benefit of working with an affine variety without a fixing an embedding into affine space. If X is embedded into $\mathbb{A}^n$, the localization X_s wants to be in $\mathbb{A}^{n+1}$.

As is true for many open sets, the complement X' of the origin in the affine plane $\operatorname{Spec} \mathbb{C}[x_1, x_2]$ isn't a localization. Every polynomial that vanishes at the origin vanishes on an affine curve, which has points different from the origin. Its inverse doesn't define a function on X'.

It may be hard to tell whether or not a given open set is a localization. The general facts at our disposal are that the intersection of localizations is a localization, because $X_s \cap X_t = X_{st}$, and the next proposition.

2.5.20. Proposition. *The localizations X_s of an affine variety X form a basis for the Zariski topology on X.*

A *basis* for the topology on a topological space X is a family $\mathcal{B}$ of open sets such that every open set is a union of open sets that are members of $\mathcal{B}$.

proof of Proposition 2.5.20. We must show that every open subset U of X can be covered by localizations of X, i.e., that for every point p of U, there is a localization X_s that is contained in U and that contains p.

Let C be the complement of U in X, and let A be the coordinate algebra of X. Because C is closed, it is the zero locus of some elements of A. Since p isn't in C, at least one of those elements, say s, will be nonzero at p though it is identically zero on C. Then X_s contains no point of C, so $X_s \subset U$, and because $s(p) \neq 0$, $p \in X_s$. □

2.5.21. Lemma. *Let U and V be open subsets of an affine variety X.*
(i) *If V is a localization of U and U is a localization of X, then V is a localization of X.*
(ii) *If U and V are affine, $V \subset U$, and if V is a localization of X, then V is a localization of U.*
(iii) *Let p be a point of $U \cap V$. If U and V are affine, there is an open set Z containing p that is a localization of U and also a localization of V.*

proof. **(i)** Say that $X = \operatorname{Spec} A$, $U = X_s = \operatorname{Spec} A_s$ and $V = U_t = \operatorname{Spec}(A_s)_t$. Then t is an element of A_s, say $t = s^{-k}r$ with r in A. The localizations $(A_s)_t$, $(A_s)_r$ are equal, and $(A_s)_r = A_{sr}$. So $V = X_{sr}$.

(ii) Say that $X = \operatorname{Spec} A$, $U = \operatorname{Spec} B$, and $V = \operatorname{Spec} A_t$, where t is a nonzero element of A. The elements of B are the fractions $t^{-k}a$ with $a \in A$, and t is an element of B. So $B_t = A_t$.

(iii) Since localizations form a basis for the topology, $U \cap V$ contains a localization X_s of X that contains p. By **(ii)**, X_s is a localization of U and of V. □

(2.5.22) extension and contraction of ideals

Let $A \subset B$ be the inclusion of a ring A as a subring of a ring B. The *extension* of an ideal I of A is the ideal IB of B generated by I. Its elements are finite sums $\sum_i z_i b_i$ with z_i in I and b_i in B. The *contraction* of an ideal J of B is the intersection $J \cap A$. It is an ideal of A.

If A_s is a localization of A and I is an ideal of A, the elements of the extended ideal IA_s are fractions of the form zs^{-k}, with z in I. We denote this extended ideal by I_s.

2.5.23. Lemma.
(i) *Let $A \subset B$ be rings, let I be an ideal of A and let J be an ideal of B. Then $I \subset (IB) \cap A$ and $(J \cap A)B \subset J$.*
(ii) *Let A_s be a localization of A, let I' be an ideal of A_s and let $I = I' \cap A$. Then $I' = IA_s$. Every ideal of A_s is the extension of an ideal of A.*
(iii) *Let P be a prime ideal of A. If s is an element of A that isn't in P, the extended ideal P_s is a prime ideal of A_s. If s is in P, the extended ideal is the unit ideal.* □

2.6 Morphisms of Affine Varieties

Morphisms are the allowed maps between varieties. Morphisms between affine varieties are defined below. They correspond to algebra homomorphisms in the opposite direction between their coordinate algebras. Morphisms of projective varieties require more thought. They will be defined in the next chapter.

(2.6.1) regular functions

The *function field* K of an affine variety $X = \operatorname{Spec} A$ is the field of fractions of A. A *rational function* on X is a nonzero element of the function field K.

As we have seen, (2.5.1) elements of the coordinate algebra A define functions on X, the rule being $\alpha(p) = \pi_p(\alpha)$, where π_p is the homomorphism $A \to \mathbb{C}$ that corresponds to p. A rational function $f = a/s$ with a and s in A is an element of A_s, and it defines a function on the open subset X_s. A rational function f is *regular* at a point p of X if it can be written as a fraction a/s such that $s(p) \neq 0$. A rational function is a *regular function* on X if it is regular at every point of X.

2.6.2. Proposition. *The regular functions on an affine variety* $X = \operatorname{Spec} A$ *are the elements of the coordinate algebra* A*.*

proof. Let f be a rational function that is regular on X. So for every point p of X, there is a localization $X_s = \operatorname{Spec} A_s$ that contains p, such that f is an element of A_s. Because X is quasicompact, a finite set of these localizations, say $X_{s_1}, \ldots, X_{s_k}$, will cover X. Then $s_1, ..., s_k$ have no common zeros on X, so they generate the unit ideal of A (2.5.6). Since f is in A_{s_i}, we can write $f = s_i^{-n} b_i$, or $s_i^n f = b_i$, with b_i in A, and we can use the same exponent n for each i. Since the elements s_i generate the unit ideal of A, so do the powers s_i^n. Say that $\sum s_i^n c_i = 1$, with c_i in A. Then $f = \sum s_i^n c_i f = \sum c_i b_i$ is an element of A. $\square$

(2.6.3) morphisms

Let $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} B$ be affine varieties, and let $A \xrightarrow{\varphi} B$ be an algebra homomorphism. A point q of Y corresponds to an algebra homomorphism $B \xrightarrow{\pi_q} \mathbb{C}$. When we compose $\pi_q\, \varphi$, we obtain a homomorphism $A \xrightarrow{\pi_q \varphi} \mathbb{C}$. The Nullstellensatz tells us that there is a unique point p of X such that $\pi_q \varphi$ is the homomorphism π_p:

(2.6.4)
$$\begin{array}{ccc} A & \xrightarrow{\varphi} & B \\ {\scriptstyle \pi_p}\big\downarrow & & \big\downarrow{\scriptstyle \pi_q} \\ \mathbb{C} & = & \mathbb{C} \end{array}$$

2.6.5. Definition. Let $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} B$. A *morphism* $Y \xrightarrow{u} X$ is a map defined, as above, by an algebra homomorphism $A \xrightarrow{\varphi} B$: If q is a point of Y, then uq is the point p of X such that $\pi_p = \pi_q \varphi$.

Then if α is an element of A and $\beta = \varphi(\alpha)$,

(2.6.6)
$$\beta(q) = \pi_q(\beta) = \pi_q(\varphi\alpha) = \pi_p(\alpha) = \alpha(p)$$

The morphism $Y \xrightarrow{u} X$ is an *isomorphism* if and only if there is an inverse morphism. This will be true if and only if $A \xrightarrow{\varphi} B$ is an isomorphism of algebras. $\square$

The relationship between a homomorphism $A \xrightarrow{\varphi} B$ and the associated morphism $Y \xrightarrow{u} X$ can be summed up by the next formula. If q is a point of Y and α is an element of A, then

(2.6.7)
$$\alpha[u(q)] = [\varphi\alpha](q)$$

Thus the homomorphism φ is determined by the map u. But most maps $Y \to X$ aren't morphisms.

The description of a morphism can be confusing because the direction of the arrow is reversed. It will become clearer as we expand the discussion.

Morphisms to the affine line.

A morphism $Y \xrightarrow{u} \mathbb{A}^1$ from a variety $Y = \operatorname{Spec} B$ to the affine line $\operatorname{Spec} \mathbb{C}[t]$ is defined by an algebra homomorphism $\mathbb{C}[x] \xrightarrow{\varphi} B$, which substitutes an element β of B for x. The morphism u that corresponds to φ sends a point q of Y to the point of the x-line at which $x = \beta(q)$.

For example, let Y be the space of 2×2 matrices, so that $B = \mathbb{C}[y_{ij}]$, $1 \leq i, j \leq 2$. The determinant defines a morphism $Y \to \mathbb{A}^1$ that sends a matrix $\beta = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$ to its determinant $b_{11}b_{22} - b_{12}b_{21}$. The corresponding algebra homomorphism $\mathbb{C}[x] \xrightarrow{\varphi} \mathbb{C}[y_{ij}]$ substitutes $y_{11}y_{22}-y_{12}y_{21}$ for x. It sends a polynomial $f(x)$ to $f(y_{11}y_{22} - y_{12}y_{21})$.

In the other direction, a morphism from $\mathbb{A}^1$ to a variety Y is a (complex) polynomial path in Y. For example, if Y is the space of matrices, a morphism $\mathbb{A}^1 \to Y$ corresponds to a homomorphism $\mathbb{C}[y_{ij}] \to \mathbb{C}[x]$, which substitutes polynomials in x for the variables y_{ij}.

Morphisms to affine space.

A morphism from an affine variety $Y = \operatorname{Spec} B$ to affine space $\mathbb{A}^n$ will be defined by a homomorphism $\mathbb{C}[x_1, ..., x_n] \xrightarrow{\Phi} B$ which substitutes elements β_i of B for x_i: $\Phi(f(x)) = f(\beta)$. The corresponding morphism $Y \xrightarrow{u} \mathbb{A}^n$ sends a point q of Y to the point $(\beta_1(q), ..., \beta_n(q))$ of $\mathbb{A}^n$.

Morphisms to affine varieties.

Let $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} B$ be affine varieties. Say that we have chosen a presentation $A = \mathbb{C}[x_1, ..., x_m]/(f_1, ..., f_k)$ of A, so that X becomes the closed subvariety $V(f)$ of affine space $\mathbb{A}^m$. There is no need to choose a presentation of B. A natural way to define a morphism from a variety Y to X is as a morphism $Y \xrightarrow{u} \mathbb{A}^m$ to affine space, whose image is contained in X. We check that this agrees with Definition 2.6.5:

As above, a morphism $Y \xrightarrow{u} \mathbb{A}^m$ corresponds to a homomorphism $\mathbb{C}[x_1, ..., x_m] \xrightarrow{\Phi} B$, and defined by a set $(\beta_1, ..., \beta_m)$ of elements of B. Since X is the locus of zeros of the polynomials f, the image of Y will be contained in X if and only if $f_i(\beta_1(q), ..., \beta_m(q)) = 0$ for every point q of Y and every i, i.e., $f_i(\beta)$ is in every maximal ideal of B, in which case $f_i(\beta) = 0$ for every i (2.5.16)**(i)**. Another way to say this is:

The image of Y is contained in X if and only if $\beta = (\beta_1, ..., \beta_m)$ solves the equations $f(x) = 0$.

And, if β is a solution, the map Φ defines a map $A \xrightarrow{\varphi} B$.

$$\begin{array}{ccc} \mathbb{C}[x] & \xrightarrow{\Phi} & B \\ \downarrow & & \| \\ A & \xrightarrow{\varphi} & B \end{array}$$

This is an elementary, but important, principle:

• *Homomorphisms from an algebra $A = \mathbb{C}[x]/(f)$ to an algebra B correspond to solutions of the equations $f = 0$ in B.*

2.6.8. Corollary. *Let $X = \operatorname{Spec} A$ and let $Y = \operatorname{Spec} B$ be affine varieties. Suppose that A is presented as the quotient $\mathbb{C}[x_1, ..., x_m]/(f_1, ..., f_k)$ of a polynomial ring. There are bijective correspondences between the following sets:*

- *algebra homomorphisms $A \to B$, or morphisms $Y \to X$,*
- *morphisms $Y \to \mathbb{A}^n$ whose images are contained in X,*
- *solutions of the equations $f_i(x) = 0$ in B,* □

The second and third sets refer to an embedding of the variety X into affine space, but the first one does not. It shows that a morphism depends only on the varieties X and Y, not on the embedding of X into affine space.

The geometry of a morphism will be described more completely in Chapters 4 and 5. We note a few more facts about them here.

2.6.9. Proposition. *Let $X \xleftarrow{u} Y$ be the morphism of affine varieties that corresponds to a homomorphism of coordinate algebras $A \xrightarrow{\varphi} B$.*

(i) *Let $Y \xleftarrow{v} Z$ be another morphism, that corresponds to another homomorphism $B \xrightarrow{\psi} R$ of finite-type domains. The the composition $Z \xrightarrow{uv} X$. is the morphism that corresponds to the composed homomorphism $A \xrightarrow{\psi\varphi} R$.*

(ii) *Suppose that $B = A/P$, where P is a prime ideal of A, and that φ is the canonical homomorphism $A \to A/P$. Then u is the inclusion of the closed subvariety $Y = V_X(P)$ into X.*

(iii) *φ is surjective if and only if u maps Y isomorphically to a closed subvariety of X.* □

It is useful to rephrase the definition of the morphism $Y \xrightarrow{u} X$ that corresponds to a homomorphism $A \xrightarrow{\varphi} B$ in terms of maximal ideals. Let $\mathfrak{m}_q$ be the maximal ideal of B at a point q of Y. The inverse image of $\mathfrak{m}_q$ in A is the kernel of the composed homomorphism $A \xrightarrow{\varphi} B \xrightarrow{\pi_q} \mathbb{C}$, so it is a maximal ideal of A: $\varphi^{-1}\mathfrak{m}_q = \mathfrak{m}_p$ for some point p of X. That point is the image of q: $p = uq$.

In the other direction, let $\mathfrak{m}_p$ be the maximal ideal at a point p of X, and let J be the ideal generated by the image of $\mathfrak{m}_p$ in B. This ideal is called the *extension* of $\mathfrak{m}_p$ to B. Its elements are finite sums $\sum \varphi(z_i)b_i$ with z_i in $\mathfrak{m}_p$ and b_i in B. If q is is a point of Y, then $uq = p$ if and only if $\mathfrak{m}_p = \varphi^{-1}\mathfrak{m}_q$, and this will be true if and only if J is contained in $\mathfrak{m}_q$.

Recall that, if $Y \xrightarrow{u} X$ is a map of sets, the *fibre* of Y over a point p of X is the set of points q of Y that map to p.

2.6.10. Corollary. *Let $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} B$, and let $Y \xrightarrow{u} X$ be the morphism corresponding to a homomorphism $A \xrightarrow{\varphi} B$, let $\mathfrak{m}_p$ be the maximal ideal at a point p of X, and let $J = \mathfrak{m}_p B$ be the extended ideal.*

(i) *The fibre of Y over p is the set $V_Y(J)$ of points q such that $J \subset \mathfrak{m}_q$.*

(ii) *The fibre of Y over p is empty if and only if J is the unit ideal of B.* □

2.6.11. Example. *(blowing up the plane)*

Let Z and Y be the affine planes with coordinates x, z and x, y, respectively. The map $Z \xrightarrow{\pi} Y$ defined by $y = xz$, the morphism that corresponds to the algebra homomorphism $\mathbb{C}[x, y] \xrightarrow{\varphi} \mathbb{C}[x, z]$ defined by $\varphi(x) = x$, $\varphi(y) = xz$.

The morphism π is bijective at points (x, y) with $x \neq 0$. At such a point, $y = x^{-1}z$. The fibre of Z over a point of Y of the form $(0, y)$ is empty unless $y = 0$, and the fibre over the origin $(0, 0)$ in Y is the z-axis $\{(0, z)\}$ in the plane Z Because the origin in Y is replaced by a line in Z, this morphism is called a *blowup* of the affine plane Y. □

figure

2.6.12. Proposition. *A morphism $Y \xrightarrow{u} X$ of affine varieties is a continuous map in the Zariski topology and also in the classical topology.*

proof. First, the Zariski topology: Let $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} B$, so that u corresponds to an algebra homomorphism $A \xrightarrow{\varphi} B$. A closed subset C of X will be the zero locus of a set $\alpha = \{\alpha_1, ..., \alpha_k\}$ of elements of A. Let $\beta_i = \varphi\alpha_i$. The inverse image $u^{-1}C$ is the set of points q such that $p = uq$ is in C, i.e., such that $\alpha_i(uq) = 0$, and $\alpha_i(uq) = \beta_i(q)$ (2.6.5). So $u^{-1}C$ is the zero locus in Y of the elements $\beta_i = \varphi(\alpha_i)$ of B. It is a closed set.

Next, for the classical topology, we use the fact that polynomials are continuous functions. A morphism if affine spaces $\mathbb{A}^n_y \xrightarrow{\widetilde{u}} \mathbb{A}^m_x$ is defined by an algebra homomorphism $\mathbb{C}[x_1, ..., x_m] \xrightarrow{\Phi} \mathbb{C}[y_1, ..., y_n]$, and this homomorphism is determined by the polynomials $h_1(y), ..., h_m(y)$ that are the images of the variables x. The morphism $\widetilde{u}$ sends the point $(y_1, ..., y_n)$ of $\mathbb{A}^n$ to the point $(h_1(y), ..., h_m(y))$ of $\mathbb{A}^m$. It is continuous.

The morphism $Y \xrightarrow{u} X$ is defined by a homomorphism $A \xrightarrow{\varphi} B$. We choose presentations $A = \mathbb{C}[x]/I$ and $B = \mathbb{C}[y]/J$, and we form a diagram of homomorphisms and the associated diagram of morphisms:

$$\begin{array}{ccc} \mathbb{C}[x] & \xrightarrow{\Phi} & \mathbb{C}[y] \\ {\scriptstyle\alpha}\downarrow & & \downarrow{\scriptstyle\beta} \\ A & \xrightarrow{\varphi} & B \end{array} \qquad \begin{array}{ccc} \mathbb{A}^n_x & \xleftarrow{\widetilde{u}} & \mathbb{A}^m_y \\ \uparrow & & \uparrow \\ X & \xleftarrow{u} & Y \end{array}$$

Here α and β are the canonical maps of a ring to a quotient ring. The map α sends $x_1, ..., x_n$ to $\alpha_1, ..., \alpha_n$. Then Φ is obtained by choosing elements h_i whose images in B are the same as the images off α_i. In the diagram of morphisms, $\widetilde{u}$ is continuous, and the vertical arrows are the embeddings of X and Y into their affine spaces. Since the topologies on X and Y are induced from their embeddings, u is continuous. □

As we see here, every morphism of affine varieties can be obtained by restricting a morphism of affine spaces. However, in the diagram above, the morphism $\widetilde{u}$ isn't unique. It depends on the choice of the polynomials h_i.

2.7 Finite Group Actions

Let G be a finite group of automorphisms of a finite-type domain B. An *invariant* element of B is an element that is sent to itself by every element σ of G. For example, the product and the sum

(2.7.1)
$$\prod_{\sigma\in G}\sigma b \quad , \quad \sum_{\sigma\in G}\sigma b$$

are invariant elements. The invariant elements form a subalgebra of B that is often denoted by B^G. Theorem 2.7.5 below asserts that B^G is a finite-type algebra, and that points of $\operatorname{Spec} B^G$ correspond bijectively to G-orbits in $\operatorname{Spec} B$.

2.7.2. Examples.
(i) The symmetric group $G = S_n$ operates on the polynomial ring $R = \mathbb{C}[x_1, ..., x_n]$ by permuting the variables, and the Symmetric Functions Theorem asserts that the elementary symmetric functions

$$s_1 = \sum_i x_i\,, \quad s_2 = \sum_{i<j} x_ix_j\,,\ \ldots,\ s_n = x_1x_2\cdots x_n$$

generate the ring R^G of invariant polynomials. Moreover, $s_1, ..., s_n$ are algebraically independent, so R^G is the polynomial algebra $\mathbb{C}[s_1, ..., s_n]$. The inclusion of R^G into R gives us a morphism from affine x-space $\mathbb{A}^n_x$ to affine s-space $\mathbb{A}^n_s = \operatorname{Spec} R^G$. If $a = (a_1, ..., a_n)$ is a point of $\mathbb{A}^n_s$, the points $b = (b_1, ..., b_n)$ of $\mathbb{A}^n_x$ that map to a are those such that $s_i(b) = a_i$. They are the roots of the polynomial $x^n - a_1x^{n-1} + \cdots \pm a_n$. Since the roots form a G-orbit, the set of G-orbits of points of $\mathbb{A}^n_x$ maps bijectively to $\mathbb{A}^n_s$.

(ii) Let $\zeta = e^{2\pi i/n}$, let σ be the automorphism of the polynomial ring $B = \mathbb{C}[y_1, y_2]$ defined by $\sigma y_1 = \zeta y_1$ and $\sigma y_2 = \zeta^{-1}y_2$. Let G be the cyclic group of order n generated by σ, and let A denote the algebra B^G of invariant elements. A monomial $m = y_1^iy_2^j$ is invariant if and only if n divides $i - j$, and an invariant polynomial is a linear combination of invariant monomials. You will be able to show that the three monomials

(2.7.3)
$$u_1 = y_1^n\,,\ u_2 = y_2^n\,,\ \text{and } w = y_1y_2$$

generate A. We'll use the same symbols u_1, u_2, w to denote variables in the polynomial ring $\mathbb{C}[u_1, u_2, w]$. Let J be the kernel of the canonical homomorphism $\mathbb{C}[u_1, u_2, w] \xrightarrow{\tau} A$ that sends u_1, u_2, w to y_1^n, y_2^n, y_1y_2.

2.7.4. Lemma. *With notation as above, the kernel J of τ is the principal ideal of $\mathbb{C}[u_1, u_2, w]$ generated by the polynomial $f = w^n - u_1u_2$.*

proof. First, f is an element of J. Let $g(u_1, u_2, w)$ be an element of J. So $g(y_1^n, y_2^n, y_1y_2) = 0$. We divide g by f, considered as a monic polynomial in w, say $g = fq + r$, where the remainder r has degree $< n$ in w. The remainder will be in J too: $r(y_1^n, y_2^n, y_1y_2) = 0$. We write r as a polynomial in w: $r = r_0(u_1, u_2) + r_1(u_1, u_2)w + \cdots + r_{n-1}(u_1, u_2)w^{n-1}$. When we substitute y_1^n, y_2^n, y_1y_2, the term $r_i(u_1, u_2)w^i$ becomes $r_i(y_1^n, y_2^n)(y_1y_2)^i$. The degree in y_1 of every monomial that appears here will be congruent to i modulo n, and the same is true for y_2. Since $r(y_1^n, y_2^n, y_1y_2) = 0$, and because the indices i are distinct, $r_i(y_1^n, y_2^n)$ will be zero for every i. And if $r_i(y_1^n, y_2^n)$ is zero, then $r_i(u_1, u_2) = 0$. So $r = 0$, which means that f divides g. □

We go back to the operation of the cyclic group on B. Let Y denote the affine plane $\operatorname{Spec} B$, and let $X = \operatorname{Spec} A$. The group G operates on Y, and except for the origin, which is a fixed point, the orbit of a point (y_1, y_2) consists of the n points $(\zeta^iy_1, \zeta^{-i}y_2)$, $i = 0, \ldots, n-1$. To show that G-orbits in Y correspond bijectively to points of X, we fix complex numbers u_1, u_2, w with $w^n = u_1u_2$, and we look for solutions of the equations (2.7.3). When $u_1 \neq 0$, the equation $u_1 = y_1^n$ has n solutions for y_1, and then y_2 is determined by the equation $w = y_1y_2$. So the fibre has order n. Similarly, there are n points in the fibre if $u_2 \neq 0$. If $u_1 = u_2 = 0$, then $y_1 = y_2 = w = 0$. In all cases, the fibres are the G-orbits. □

2.7.5. Theorem. *Let G be a finite group of automorphisms of a finite-type domain B, and let A denote the algebra B^G of invariant elements. Let $Y = \operatorname{Spec} B$ and $X = \operatorname{Spec} A$.*

(i) *A is a finite-type domain and B is a finite A-module.*
(ii) *G operates by automorphisms on Y.*
(iii) *The morphism $Y \to X$ defined by the inclusion $A \subset B$ is surjective, and its fibres are the G-orbits of points of Y.*

When a group G operates on a set Y, one often denotes the set of G-orbits of Y by Y/G. With that notation, the theorem asserts that there is a bijective map $Y/G \to X$.

proof of 2.7.5 **(i)**: *The invariant algebra $A = B^G$ is a finite-type algebra, and B is a finite A-module.*

This is an interesting indirect proof. To show that A is a finite-type algebra, one constructs a finite-type subalgebra R of A such that B is a finite R-module.

Let $\{z_1, \ldots, z_k\}$ be the G-orbit of an element z_1 of B. The orbit is the set of roots of the polynomial

$$f(t) \;=\; (t - z_1) \cdots (t - z_k) \;=\; t^k - s_1 t^{k-1} + \cdots \pm s_k$$

whose coefficients are the elementary symmetric functions in $\{z_1, ..., z_k\}$. Let R_1 denote the algebra generated by those symmetric functions. Because the symmetric functions are invariant, $R_1 \subset A$. Using the equation $f(z_1) = 0$, we can write any power of z_1 as a polynomial in z_1 of degree less than k, with coefficients in R_1.

We choose a finite set of generators $\{y_1, \ldots, y_r\}$ for the algebra B. If the order of the orbit of y_j is k_j, then y_j will be the root of a monic polynomial f_j of degree k_j with coefficients in A. Let R denote the finite-type algebra generated by all of the coefficients of all of the polynomials $f_1, ..., f_r$. We can write any power of y_j as a polynomial in y_j with coefficients in R, and of degree less than k_j. Using such expressions, we can write every monomial in $y_1, ..., y_r$ as a polynomial $y_1, ..., y_r$ with coefficients in R, whose degree in each variable y_j is less than k_j. Since $y_1, ..., y_r$ generate B, we can write every element of B as such a polynomial. Then the finite set of monomials $y_1^{e_1} \cdots y_r^{e_r}$ with $e_j < k_j$ spans B as an R-module. Therefore B is a finite R-module.

Since R is a finite-type algebra, it is noetherian. The algebra A of invariants is a subalgebra of B that contains R. So when regarded as an R-module, A is a submodule of the finite R-module B. Since R is noetherian, A is also a finite R-module. When we put a finite set of algebra generators for R together with a finite set of R-module generators for A, we obtain a finite set of algebra generators for A. So A is a finite-type algebra. And, since B is a finite R-module, it is also a finite module over the larger ring A.

*proof of 2.7.5***(ii)**: *The group G operates on Y.*

A group element σ is a homomorphism $B \xrightarrow{\sigma} B$, which defines a morphism $Y \xleftarrow{u_\sigma} Y$, as in Definition 2.6.5. Since σ is an invertible homomorphism, i.e., an automorphism, u_σ is also an automorphism. Thus G operates on Y. However, there is a point that should be mentioned.

Let's write the operation of G on B on the left as usual, so that a group element σ maps an element β of B to σb. Then if σ and τ are two group elements, the product $\sigma\tau$ acts as *first do τ, then σ*: $(\sigma\tau)\beta = \sigma(\tau\beta)$.

(2.7.6) $$B \xrightarrow{\tau} B \xrightarrow{\sigma} B$$

We substitute $u = u_\sigma$ into Definition 2.6.5: If q is a point of Y, the morphism $Y \xleftarrow{u_\sigma} Y$ sends q to the point p such that $\pi_p = \pi_q \sigma$. It seems permissible to drop the symbol u, and to write the morphism simply as $Y \xleftarrow{\sigma} Y$. But since arrows are reversed when going from homomorphisms of algebras to morphisms of their spectra, the maps displayed in (2.7.6), give us morphisms

(2.7.7) $$Y \xleftarrow{\tau} Y \xleftarrow{\sigma} Y$$

On $Y = \operatorname{Spec} B$, the product $\sigma\tau$ acts as *first do σ, then τ.*

To get around this problem, we can put the symbol σ on the right when it operates on Y, so that σ sends a point q to $q\sigma$. Then if q is a point of Y, we will have $q(\sigma\tau) = (q\sigma)\tau$, as required of an operation.

- *If G operates on the left on B, it operates on the right on* $\operatorname{Spec} B$.

This is important only when one wants to compose morphisms. In Definition 2.6.5, we followed custom and wrote the morphism u that corresponds to an algebra homomorphism φ on the left. We will continue to write morphisms on the left when possible, but not here.

Let β be an element of B and let q be a point of Y. The value $[\sigma\beta](q)$ of the function $\sigma\beta$ at q is the same as the value of β at $q\sigma$: $[\sigma\beta](q) \overset{defn}{=} \pi_q(\sigma\beta) = \pi_{q\sigma}(\beta) \overset{defn}{=} \beta(q\sigma)$ (2.6.6):

(2.7.8)
$$[\sigma\beta](q) = \beta(q\sigma)$$
□

proof of 2.7.5 **(iii)**: *The fibres of the morphism $Y \to X$ are the G-orbits in Y.*

We go back to the subalgebra $A = B^G$. For σ in G, we have a diagram of algebra homomorphisms and the corresponding diagram of morphisms

(2.7.9)
$$\begin{array}{ccc} B & \xrightarrow{\sigma} & B \\ \uparrow & & \uparrow \\ A & = & A \end{array} \qquad \begin{array}{ccc} Y & \xleftarrow{\sigma} & Y \\ \downarrow & & \downarrow \\ X & = & X \end{array}$$

The diagram of morphisms shows that the elements of Y forming a G-orbit have the same image in X, and therefore that the set of G-orbits in Y, which we denote by Y/G, maps to X. We show that the map $Y/G \to X$ is bijective.

2.7.10. Lemma. (i) *Let $p_1, \ldots, p_k$ be distinct points of affine space $\mathbb{A}^n$, and let $c_1, \ldots, c_k$ be complex numbers. There is a polynomial $f(x_1, \ldots, x_n)$ such that $f(p_i) = c_i$ for $i = 1, \ldots, n$.*
(ii) *Let B be a finite-type algebra, let $q_1, \ldots, q_k$ be points of* Spec B, *and let $c_1, \ldots, c_k$ be complex numbers. There is an element β in B such that $\beta(q_i) = c_i$ for $i = 1, \ldots, k$.* □

injectivity of the map $Y/G \to X$:

Let O_1 and O_2 be distinct G-orbits. Lemma 2.7.10 tells us that there is an element β in B whose value is 0 at every point of O_1, is 1 at every point of O_2. Since G permutes the orbits, $\sigma\beta$ will also be 0 at points of O_1 and 1 at points of O_2. Then the product $\gamma = \prod_\sigma \sigma\beta$ will be 0 at points of O_1 and 1 at points of O_2, and γ is invariant. If p_i denotes the image in X of the orbit O_i, the maximal ideal $\mathfrak{m}_{p_i}$ of A is the intersection $A \cap \mathfrak{m}_q$, where q is any point in O_i. Therefore γ is in the maximal ideal $\mathfrak{m}_{p_1}$, but not in $\mathfrak{m}_{p_2}$. The images of the two orbits are distinct.

surjectivity of the map $Y/G \to X$:

It suffices to show that the map $Y \to X$ is surjective.

2.7.11. Lemma. *If I is an ideal of the invariant algebra A, and if the extended ideal IB is the unit ideal of B, then I is the unit ideal of A.*

As before, the extended ideal IB is the ideal of B generated by I.

Let's assume the lemma for the moment, and use it to prove surjectivity of the map $Y \to X$. Let p be a point of X. The lemma tells us that the extended ideal $\mathfrak{m}_pB$ isn't the unit ideal. So it is contained in a maximal ideal $\mathfrak{m}_q$ of B, where q is a point of Y. Then $\mathfrak{m}_p \subset (\mathfrak{m}_pB) \cap A \subset \mathfrak{m}_q \cap A$.

The contraction $\mathfrak{m}_q \cap A$ is an ideal of A, and it isn't the unit ideal because 1 isn't in $\mathfrak{m}_q$. Since $\mathfrak{m}_p$ is a maximal ideal, $\mathfrak{m}_p = \mathfrak{m}_q \cap A$. This means that q maps to p in X. □

proof of the lemma. If $IB = B$, there will be an equation $\sum_i z_i b_i = 1$, with z_i in I and b_i in B. The sums $\alpha_i = \sum_\sigma \sigma b_i$ are invariant, so they are elements of A, and the elements z_i are invariant. Therefore $\sum_\sigma \sigma(z_i b_i) = z_i \sum_\sigma \sigma b_i = z_i \alpha_i$ is in I. Then

$$\sum_\sigma 1 = \sum_\sigma \sigma(1) = \sum_{\sigma,i} \sigma(z_i b_i) = \sum_i z_i \alpha_i$$

The right side is in I, and the left side is the order of the group which, because A contains the complex numbers, is an invertible element of A. So I is the unit ideal. □

Chapter 3 PROJECTIVE ALGEBRAIC GEOMETRY

3.1 Projective Varieties

The projective space $\mathbb{P}^n$ of dimension n was defined in Chapter 1. Its points are equivalence classes of nonzero vectors $(x_0, ..., x_n)$, the equivalence relation being that, for any nonzero complex number λ,

(3.1.1) $$(x_0, ..., x_n) \sim (\lambda x_0, ..., \lambda x_n).$$

A subset of $\mathbb{P}^n$ is *Zariski closed* if it is the set of common zeros of a family of *homogeneous* polynomials $f_1, ..., f_k$ in the coordinate variables $x_0, ..., x_n$, or if it is the set of zeros of the homogeneous ideal $\mathcal{I}$ generated by such a family (see Section 3.2). Homogeneity is required because the vectors (x) and (λx) represent the same point of $\mathbb{P}^n$. As explained in (1.3.1), $f(\lambda x) = 0$ for all λ if and only if f is homogeneous. We usually omit the word 'Zariski', and refer to a Zariski closed set simply as a closed set.

Because the polynomial ring $\mathbb{C}[x_0, ..., x_n]$ is noetherian, $\mathbb{P}^n$ is a noetherian space: Every strictly increasing family of ideals of $\mathbb{C}[x]$ is finite, and every strictly decreasing family of closed subsets of $\mathbb{P}^n$ is finite. Therefore every closed subset of $\mathbb{P}^n$ is a finite union of irreducible closed sets (2.2.13). The irreducible closed sets are the *closed subvarieties* of $\mathbb{P}^n$, the *projective varieties.*

Thus a projective variety X is an irreducible closed subset of some projective space. We will also want to know when two projective varieties are isomorphic. This will be explained in Section 3.4, where morphisms are defined.

The closed subsets of $\mathbb{P}^n$ are the closed sets in the *Zariski topology* on $\mathbb{P}^n$, and the Zariski topology on a projective variety X is induced from the topology on the projective space that contains it. Since a projective variety X is closed in $\mathbb{P}^n$, a subset of X is closed in X if it is closed in $\mathbb{P}^n$.

3.1.2. Lemma. *The one-point subsets of projective space are closed.*

proof. This simple proof illustrates a general method. Let p be the point $(a_0, ..., a_n)$. The first guess might be that the one-point set $\{p\}$ is defined by the equations $x_i = a_i$, but the polynomials $x_i - a_i$ aren't homogeneous in x. This is reflected in the fact that, for any $\lambda \neq 0$, the vector $(\lambda a_0, ..., \lambda a_n)$ represents the same point, though it won't satisfy those equations. The equations that define the set $\{p\}$ are

(3.1.3) $$a_i x_j = a_j x_i,$$

for $i, j = 0, ..., n$, which show that the ratios a_i/a_j and x_i/x_j are equal. □

3.1.4. Lemma. *The proper closed subsets of the projective line are the nonempty finite subsets, and the proper closed subsets of the projective plane are finite unions of points and curves.* □

Though affine varieties are important, most of algebraic geometry concerns projective varieties. It isn't very clear why this is so, but one property of projective space gives a hint of its importance: With its classical topology, projective space is *compact*.

A topological space is compact if it has these properties:

Hausdorff property: Distinct points p, q of X have disjoint open neighborhoods, and
quasicompactness: If X is covered by a family $\{U^i\}$ of open sets, then a finite subfamily covers X.

By the way, when we say that the sets $\{U^i\}$ *cover* a topological space X, we mean that X is the union $\bigcup U^i$. We don't allow U^i to contain elements that aren't in X, though that would be a customary English usage.

In the classical topology, affine space $\mathbb{A}^n$ is a Hausdorff space, but it isn't quasicompact, and therefore it isn't compact. The *Heine-Borel Theorem* asserts that a subset of $\mathbb{A}^n$ is compact if and only if it is closed and bounded.

We'll show that $\mathbb{P}^n$ is compact, assuming that the Hausdorff property has been verified. The $2n+1$-dimensional sphere $\mathbb{S}$ of unit length vectors in $\mathbb{A}^{n+1}$ is a bounded set, and because it is the zero locus of the equation $\overline{x}_0x_0 + \cdots + \overline{x}_nx_n = 1$, it is closed. The Heine-Borel Theorem tells us that $\mathbb{S}$ is compact. The map $\mathbb{S} \to \mathbb{P}^n$ that sends a vector $(x_0, ..., x_n)$ to the point of projective space with that coordinate vector is continuous and surjective. The next lemma of topology shows that $\mathbb{P}^n$ is compact.

3.1.5. Lemma. *Let $Y \xrightarrow{f} X$ be a continuous map. Suppose that Y is compact and that X is a Hausdorff space. Then the image $Z = f(Y)$ is a closed and compact subset of X.*

proof. The image Z gets the induced topology. Since it is a subspace of X, Z is a Hausdorff space. If $\{V^i\}$ is an open covering of Z, the inverse images $U^i = f^{-1}V^i$ form an open covering of Y that has a finite subcovering $U^{i_1}, ..., U^{i_k}$. The open sets V^{i_ν} are the images of U^{i_ν}, and they cover Z. So Z is quasicompact. $\square$

The rest of this section contains a few examples of projective varieties.

(3.1.6) linear subspaces

Let V denote the affine space $\mathbb{A}^{n+1}$. The complement of the origin in V is mapped to the projective space $\mathbb{P}^n$ by sending a vector $(x_0, ..., x_n)$ to the point of $\mathbb{P}^n$ it defines. This map can be useful when one studies projective space.

If W is a subspace of dimension $r+1$ of the vector space V, the points of $\mathbb{P}^n$ that are represented by the nonzero vectors in W form a *linear subspace* L of $\mathbb{P}^n$, of dimension r. If $(w_0, ..., w_r)$ is a basis of W, the linear subspace L corresponds bijectively to a projective space of dimension r, by

$$c_0w_0 + \cdots + c_rw_r \quad \longleftrightarrow \quad (c_0, ..., c_r)$$

For example, the set of points $(x_0, ..., x_r, 0, ..., 0)$ is a linear subspace of dimension r. $\square$

(3.1.7) a quadric surface

A *quadric* in $\mathbb{P}63$ is the locus of zeros of an irreducible homogeneous quadratic equation in four variables. We describe a bijective map from the product $\mathbb{P}^1 \times \mathbb{P}^1$ of projective lines to a quadric.

Let coordinates in the two copies of $\mathbb{P}^1$ be (x_0, x_1) and (y_0, y_1), respectively, and let the four coordinates in $\mathbb{P}^3$ be w_{ij}, with $0 \leq i, j \leq 1$. The map is defined by $w_{ij} = x_iy_j$. Its image is the quadric Q whose equation is

(3.1.8) $$w_{00}w_{11} = w_{01}w_{10}$$

Let's check that the map $\mathbb{P}^1 \times \mathbb{P}^1 \to Q$ is bijective. If w is a point of Q, one of the coordinates, say w_{00}, will be nonzero. Then if (x, y) is a point of $\mathbb{P}^1 \times \mathbb{P}^1$ whose image is w, so that $w_{ij} = x_iy_j$, the coordinates x_0 and y_0 must be nonzero. When we normalize w_{00}, x_0 and y_0 to 1, there is a unique solution for x and y such that $w_{ij} = x_iy_j$, namely $x_1 = w_{10}$ and $y_1 = w_{01}$.

The quadric with the equation (3.1.8) contains two families of *lines* (one dimensional linear subspaces), the images of the subsets $x\times\mathbb{P}^1$ and $\mathbb{P}^1\times y$ of $\mathbb{P}\times\mathbb{P}$.

Note. Equation (3.1.8) can be diagonalized by the substitution $w_{00} = s+t$, $w_{11} = s-t$, $w_{01} = u+v$, $w_{10} = u-v$. This substitution changes the equation (3.1.8) to $s^2 - t^2 = u^2 - v^2$. When we look at the affine open set $\{u = 1\}$, the equation becomes $s^2+v^2-t^2 = 1$. The real locus of this equation is a one-sheeted hyperboloid in $\mathbb{R}^3$, and the two families of complex lines in the quadric correspond to the familiar rulings of this hyperboloid by real lines.

$$figure:\ hyperboloid\, with\, rulings.\ perhaps\, actual\, picture\, of\, the\, locus\, above \qquad \square$$

(3.1.9) hypersurfaces

A *hypersurface* is the locus of zeros in a projective space $\mathbb{P}^n$ of an irreducible homogeneous polynomial $f(x_0, ..., x_n)$. Plane projective curves and quadric surfaces are hypersurfaces.

(3.1.10) the Segre embedding of a product

The product $\mathbb{P}^m_x\times\mathbb{P}^n_y$ of projective spaces can be embedded by its *Segre embedding* into a projective space $\mathbb{P}^N_w$ that has coordinates w_{ij}, with $i = 0, ..., m$ and $j = 0, ..., n$. So $N = (m+1)(n+1)-1$. The Segre embedding is defined by

(3.1.11)
$$w_{ij} = x_i y_j.$$

We call the coordinates w_{ij} the *Segre variables*.

The map from $\mathbb{P}^1\times\mathbb{P}^1$ to $\mathbb{P}^3$ that was described in (**3.1.7**) is the simplest case of a Segre embedding.

3.1.12. Proposition. *The Segre embedding maps the product $\mathbb{P}^m\times\mathbb{P}^n$ bijectively to the locus Z of the* **Segre equations**

(3.1.13)
$$w_{ij}w_{k\ell} - w_{i\ell}w_{kj} = 0.$$

proof. When one substitutes (3.1.11) into the Segre equations, one obtains equations in $\{x_i, y_j\}$ that are true. So the image of the Segre embedding is contained in Z.

Let $\mathbb{U}^i$, $\mathbb{V}^j$ and $\mathbb{W}^{ij}$ denote the standard affine open subsets $\{x_i \neq 0\}$, $\{y_j \neq 0\}$ and $\{w_{ij} \neq 0\}$ of $\mathbb{P}^m$, $\mathbb{P}^n$ and $\mathbb{P}^N$, respectively. Say that we have a point p of Z that lies in W^{00}, and that p is the image of (x, y). Then w_{00} is nonzero, so x_0 and y_0 are also nonzero. We normalize w_{00}, x_0, and y_0 to 1. Then $w_{ij} = w_{i0}w_{0j}$ for all i, j, $x_i = w_{i0}$ and $y_j = w_{0j}$. $\square$

The Segre embedding is important because it makes the product of projective spaces into a projective variety, the closed subvariety of $\mathbb{P}^N$ defined by the Segre equations. However, there is a point that should be discussed. To show that the product is a variety, we need to show that the locus of the Segre equations is irreducible. We defer discussion of this point to Section 3.3 (see Proposition 3.3.1).

(3.1.14) the Veronese embedding of projective space

Let the coordinates in $\mathbb{P}^n$ be x_i, and let those in $\mathbb{P}^N$ be v_{ij}, with $0\le i\le j\le n$. Then $N = \binom{n+2}{2} - 1$. The Veronese embedding is the map $\mathbb{P}^n \xrightarrow{f} \mathbb{P}^N$ defined by $v_{ij} = x_i x_j$. The Veronese embedding resembles the Segre embedding, but in the Segre embedding, there are distinct sets of coordinates x and y, and there is no requirement that $i\le j$.

The proof of the next proposition is similar to the proof of (3.1.12).

3.1.15. Proposition. *The Veronese embedding f maps $\mathbb{P}^n$ bijectively to the locus X in $\mathbb{P}^N$ of the equations*

$$v_{ij}v_{k\ell} = v_{i\ell}v_{kj} \quad \textit{for} \quad 0\le i\le k\le j\le \ell\le n \qquad \square$$

For example, the Veronese embedding maps $\mathbb{P}^1$ bijectively to the conic $v_{00}v_{11} = v_{01}^2$ in $\mathbb{P}^2$.

(3.1.16) the twisted cubic

There are higher order Veronese embeddings, defined in an analogous way by the monomials of some degree $d > 2$. The first example is the embedding of $\mathbb{P}^1$ by the cubic monomials in two variables, which maps $\mathbb{P}^1_x$ to $\mathbb{P}^3_v$. Let the coordinates in $\mathbb{P}^3$ be $v_0, ..., v_3$. The cubic Veronese embedding is defined by

$$v_0 = x_0^3, \quad v_1 = x_0^2 x_1, \quad v_2 = x_0 x_1^2, \quad v_3 = x_1^3$$

Its image is a *twisted cubic* in $\mathbb{P}^3$, the locus $(v_0, v_1, v_2, v_3) = (x_0^3, x_0^2 x_1, x_0 x_1^2, x_1^3)$, which is the set of common zeros of the three polynomials

(3.1.17) $$v_0 v_2 - v_1^2 \ , \ \ v_1 v_2 - v_0 v_3 \ , \ \ v_1 v_3 - v_2^2$$

These polynomials are the 2×2 minors of the 2×3 matrix

(3.1.18) $$\begin{pmatrix} v_0 & v_1 & v_2 \\ v_1 & v_2 & v_3 \end{pmatrix}$$

A 2×3 matrix has rank ≤ 1 if and only if its 2×2 minors are zero. So a point (v_0, v_1, v_2, v_3) lies on the twisted cubic if (3.1.18) has rank one. This means that the vectors (v_0, v_1, v_2) and (v_1, v_2, v_3), if both are nonzero, represent the same point of $\mathbb{P}^2$. Setting $x_0 = v_0 = 1$ and $x_1 = v_1 = t$, the twisted cubic becomes the locus of points $(1, t, t^2, t^3)$. There is also one point at which $v_0 = 0$, the point $(0, 0, 0, 1)$. □

3.2 Homogeneous Ideals

We denote the polynomial algebra $\mathbb{C}[x_0, ..., x_n]$ by R here.

3.2.1. Lemma. *Let $\mathcal{I}$ be an ideal of R. The following conditions are equivalent.*
(i) *$\mathcal{I}$ can be generated by homogeneous polynomials.*
(ii) *A polynomial is in $\mathcal{I}$ if and only if its homogeneous parts are in $\mathcal{I}$.* □

An ideal $\mathcal{I}$ of R that satisfies these conditions is a *homogeneous ideal*.

3.2.2. Lemma. *The radical of a homogeneous ideal is homogeneous.*

proof. Let $\mathcal{I}$ be a homogeneous ideal, and let f be an element of its radical $\operatorname{rad}\mathcal{I}$. So f^r is in $\mathcal{I}$ for some r. When f is written as a sum $f_0 + \cdots + f_d$ of its homogeneous parts, the highest degree part of f^r is $(f_d)^r$. Since $\mathcal{I}$ is homogeneous, $(f_d)^r$ is in $\mathcal{I}$ and f_d is in $\operatorname{rad}\mathcal{I}$. Then $f_0 + \cdots + f_{d-1}$ is also in $\operatorname{rad}\mathcal{I}$. By induction on d, all of the homogeneous parts $f_0, ..., f_d$ are in $\operatorname{rad}\mathcal{I}$. □

If f is a set of homogeneous polynomials, the set of its zeros in $\mathbb{P}^n$ may be denoted by $V(f)$ or $V_{\mathbb{P}^n}(f)$, and the set of zeros of a homogeneous ideal $\mathcal{I}$ by $V(\mathcal{I})$ or $V_{\mathbb{P}^n}(\mathcal{I})$. (This is the same notation as is used for closed subsets of affine space.)

A homogeneous ideal $\mathcal{I}$ has a zero locus in projective space $\mathbb{P}^n$ and a zero locus in affine space $\mathbb{A}^{n+1}$. We can't use the $V(\mathcal{I})$ notation for both of them, so let's denote these two loci by V and W, respectively. Unless $\mathcal{I}$ is the unit ideal, the origin $x = 0$ will be a point of W, and the complement of the origin will map surjectively to V. If a point x other than the origin is in W, then because a homogeneous polynomial f vanishes at x if and only if it vanishes at λx, every point of the line through 0 and x is in W. An affine variety that is the union of lines through the origin is called an *affine cone*. If the locus W contains a point x other than the origin, it is an affine cone.

The familiar locus $x_0^2 + x_1^2 - x_2^2 = 0$ is a cone in $\mathbb{A}^3$. The zero locus of the polynomial $x_0^3 + x_1^3 - x_2^3$ is also called a cone.

Note. The real locus $x_0^2 + x_1^2 - x_2^2 = 0$ in $\mathbb{R}^3$ decomposes into two parts when the origin is removed. Because of this, it is sometimes called a "double cone". However, the complex locus doesn't decompose.

(3.2.3) the irrelevant ideal

In the polynomial algebra $R = \mathbb{C}[x_0, ..., x_n]$, the maximal ideal $\mathcal{M} = (x_0, ..., x_n)$ generated by the variables is called the *irrelevant ideal* because its locus of zeros in projective space is empty.

3.2.4. Proposition. *The zero locus in $\mathbb{P}^n$ of a homogeneous ideal $\mathcal{I}$ of R is empty if and only if $\mathcal{I}$ contains a power of the irrelevant ideal.*

Another way to say this is that $V(\mathcal{I})$ is empty if and only if either $\mathcal{I}$ is the unit ideal R, or its radical is the irrelevant ideal.

proof of Proposition 3.2.4. Let Z be the zero locus of $\mathcal{I}$ in $\mathbb{P}^n$. If $\mathcal{I}$ contains a power of $\mathcal{M}$, it contains a power of each variable. Powers of the variables have no common zeros in projective space, so Z is empty.

Suppose that Z is empty, and let W be the locus of zeros of $\mathcal{I}$ in the affine space $\mathbb{A}^{n+1}$. Since the complement of the origin in W maps to the empty locus Z, it is empty. The origin is the only point that might be in W. If W is the one point space consisting of the origin, then $\operatorname{rad}\mathcal{I}$ is the irrelevant ideal $\mathcal{M}$. If W is empty, $\mathcal{I}$ is the unit ideal. □

3.2.5. Lemma. *Let $\mathcal{P}$ be a homogeneous ideal in the polynomial algebra R, not the unit ideal. The following conditions are equivalent:*

(i) *$\mathcal{P}$ is a prime ideal.*

(ii) *If f and g are homogeneous polynomials, and if $fg \in \mathcal{P}$, then $f \in \mathcal{P}$ or $g \in \mathcal{P}$.*

(iii) *If $\mathcal{A}$ and $\mathcal{B}$ are homogeneous ideals, and if $\mathcal{AB} \subset \mathcal{P}$, then $\mathcal{A} \subset \mathcal{P}$ or $\mathcal{B} \subset \mathcal{P}$.*

Thus a homogeneous ideal is a prime ideal if the usual conditions for a prime ideal are satisfied when the polynomials or ideals are homogeneous.

proof of the lemma. The facts that **(i)** implies **(ii)** and **(iii)** follow from the analogous statements for nonhomogeneous ideals, and the implication **(iii)** ⇒ **(ii)** is proved by considering the principal ideals generated by f and g.

(ii) ⇒ (i) Suppose that a homogeneous ideal $\mathcal{P}$ satisfies the condition stated in **(ii)**, and that the product fg of two polynomials, not necessarily homogeneous, is in $\mathcal{P}$. If f has degree d and g has degree e, the highest degree part of fg is the product $f_d g_e$ of the homogeneous parts of f and g of maximal degree. Since $\mathcal{P}$ is a homogeneous ideal, it contains $f_d g_e$. Therefore one of the factors, say f_d, is in $\mathcal{P}$. Let $h = f - f_d$. Then hg is in $\mathcal{P}$, and it has lower degree than fg. By induction on the degree of fg, h or g is in $\mathcal{P}$, and if h is in $\mathcal{P}$, so is f. □

3.2.6. Proposition. *Let Y be the zero locus in $\mathbb{P}^n$ of a homogeneous radical ideal $\mathcal{I}$, not the irrelevant ideal. Then Y is a projective variety (an irreducible closed subset of $\mathbb{P}^n$) if and only if $\mathcal{I}$ is a prime ideal. Thus a subset Y of $\mathbb{P}^n$ is a projective variety if and only if it is the zero locus of a homogeneous prime ideal that isn't the irrelevant ideal.*

proof. Let W be the locus of zeros of $\mathcal{I}$ in the affine space $\mathbb{A}^{n+1}$. Then W is irreducible if and only if Y is irreducible. This is easy to see. Proposition 2.2.16 tells us that W is irreducible if and only if the radical ideal $\mathcal{I}$ is a prime ideal. □

As before, $V(\mathcal{I})$ stands for the zero locus of a homogeneous ideal $\mathcal{I}$ in projective space.

3.2.7. Strong Nullstellensatz, projective version.

(i) *Let g be a nonconstant homogeneous polynomial in $x_0, ..., x_n$, and let $\mathcal{I}$ be a homogeneous ideal of $\mathbb{C}[x]$. If g vanishes at every point of $V(\mathcal{I})$, then $\mathcal{I}$ contains a power of g.*

(ii) *Let f and g be homogeneous polynomials. If f is irreducible and if $V(f) \subset V(g)$, then f divides g.*

(iii) *Let $\mathcal{I}$ and $\mathcal{J}$ be homogeneous ideals, and suppose that $\operatorname{rad}\mathcal{I}$ isn't the irrelevant ideal or the unit ideal. Then $V(\mathcal{I}) = V(\mathcal{J})$ if and only if $\operatorname{rad}\mathcal{I} = \operatorname{rad}\mathcal{J}$.*

proof. **(i)** Let W be the locus of zeros of $\mathcal{I}$ in the affine space $\mathbb{A}^{n+1}$ with coordinates x. The polynomial g vanishes at every point of W different from the origin, and since g isn't a constant, it vanishes at the origin too. So the affine Strong Nullstellensatz applies to W.

(ii) If f is irreducible, the principal ideal (f) will be a prime ideal. So f divides a power of g if and only if f divides g.

(iii) If $\operatorname{rad}\mathcal{I} = \operatorname{rad}\mathcal{J}$, then because $V(\mathcal{I}) = V(\operatorname{rad}\mathcal{I}) = V(\operatorname{rad}\mathcal{J}) = V(\mathcal{J})$. Suppose that $\operatorname{rad}\mathcal{I} > \operatorname{rad}\mathcal{I}$. Then $V(\mathcal{I}) \subset V(\mathcal{I})$. Let g be a homogeneous polynomial in $\operatorname{rad}\mathcal{I}$ that isn't in $\operatorname{rad}\mathcal{I}$. Then g is zero on $V(\mathcal{I})$, but by **(i)**, g isn't zero on $V(\mathcal{I})$. So $V(\mathcal{I}) < V(\mathcal{I})$. □

(3.2.8) quasiprojective varieties

A nonempty (Zariski) open subset X of a projective variety is called a *quasiprojective variety*. For instance, a projective variety is quasiprojective. The complement of a point in a projective variety is a quasiprojective variety. An affine variety $X = \operatorname{Spec} A$ may be regarded as a quasiprojective variety by embedding it as a closed subvariety of the standard affine space $\mathbb{U}^0$. It becomes an open subvariety of its closure in $\mathbb{P}^n$, which is a projective variety (Lemma 2.2.12 **(ii)**).

The topology on a quasiprojective variety is induced from the topology on projective space.

3.2.9. Lemma. *The topology on the affine open subset $\mathbb{U}^0 : x_0 \neq 0$ of $\mathbb{P}^n$ that is induced from the Zariski topology on $\mathbb{P}^n$ is the same as the Zariski topology obtained by viewing $\mathbb{U}^0$ as the affine space $\operatorname{Spec}\mathbb{C}[u_1, ..., u_n]$, $u_i = x_i/x_0$.* □

Here is the description of a quasiprojective variety X in terms of equations:

3.2.10. Let $\overline{X}$ be the closure of X in projective space $\mathbb{P}^n$, and let C be the (closed) complement of X in $\overline{X}$. The closed set $\overline{X}$ will be the zero set of a family $f_1, ..., f_k$ of homogeneous polynomials, and C will be the zero set of another family $g_1, ..., g_\ell$. Then a point p of X will be a point of projective space that solves the equations $f = 0$ but doesn't solve $g = 0$: All of the polynomials f_i vanish at p, and there is at least one polynomial g_j that doesn't vanish there.

For example, if an affine variety X is embedded as a closed subvariety of $\mathbb{U}^0$ and the locus of zeros of $f = 0$ is the closure $\overline{X}$, then a point of X is a zero of f, but x_0 isn't zero at p. □

These days, it is customary to define varieties without reference to an embedding into projective space, as we did for affine varieties in Chapter 2 (**??**). However, to do this requires work. Most operations that one wants to make preserve the quasiprojective property, and though there are varieties that cannot be embedded into any projective space, they aren't very important. In fact, it is hard enough to find convincing examples of such varieties that we won't try to give one here. All varieties that we consider will be quasiprojective. In order to simplify terminology, and because the word "quasiprojective" is ugly, we will henceforth use the word "variety" to mean "quasiprojective variety".

3.3 Product Varieties

The properties of products of varieties seem intuitive, but some of the proofs aren't obvious.

3.3.1. Proposition. *Let X and Y be irreducible topological spaces, and suppose that a topology is given on the product $P = X \times Y$, such that*

• *the topology on $X \times Y$ is at least as fine as the product topology, i.e., the projections $P \xrightarrow{\pi_1} X$ and $P \xrightarrow{\pi_2} Y$ are continuous, and*

• *for all x in X and all y in Y, the fibres ${}_xP = x \times Y$ and $P_y = X \times y$, with topologies induced from P, are homeomorphic to Y and X, respectively.*

Then P is an irreducible topological space.

3.3.2. Lemma. *Let X, Y, P be as in the proposition. If W is an open subset of P, its image U via the projection $P \to Y$ is an open subset of Y.*

proof. Since W is open, the intersection ${}_xW = W \cap {}_xP$ is an open subset of the fibre ${}_xP$, whose image ${}_xU$ in the homeomorphic space Y is also open. Since W is the union of the sets ${}_xW$, U is the union of the open sets ${}_xU$. So U is open. □

proof of Proposition 3.3.1. Let C and C' be closed subsets of the product P. Suppose that $C < P$ and $C' < P$, and let $W = P - C$ and $W' = P - C'$ be the open complements of C and C' in P. To show that P is irreducible, we must show that $C \cup C' < P$. We do this by showing that $W \cap W'$ is nonempty.

Since $C < P$, W is nonempty, and similarly, W' is nonempty. The lemma tells us that the images U and U' of W and W' via projection to Y are nonempty open subsets of Y. Since Y is irreducible, $U \cap U'$ is nonempty. Let y be a point of $U \cap U'$. Then the open subsets $W_y = W \cap P_y$ and $W'_y = W' \cap P_y$ of P_y are nonempty. Since P_y is homeomorphic to the irreducible space X, $W_y \cap W'_y$ is nonempty. Therefore $W \cap W'$ is nonempty, as was to be shown. □

(3.3.3) products of affine varieties

We inspect the product $X \times Y$ of the affine varieties $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} B$. Say that X is embedded as a closed subvariety of $\mathbb{A}^m$, so that $A = \mathbb{C}[x]/P$ for some prime ideal P, and that Y is embedded similarly into $\mathbb{A}^n$, and $B = \mathbb{C}[y]/Q$. Then in affine x,y-space $\mathbb{A}^{m+n}$, $X \times Y$ is the locus of the equations $f(x) = 0$ and $g(y) = 0$ with $f \in P$ and $g \in Q$. Proposition 3.3.1 shows that $X \times Y$ is irreducible. Therefore it is a variety.

Let $I = (P, Q)$ denote the ideal of $\mathbb{C}[x, y]$ generated by the elements $f(x)$ of P and $g(y)$ of Q. The zero locus of I in $\mathbb{A}^{m+n}$ is the variety $X \times Y$.

3.3.4. Proposition. *Let P and Q be prime ideals of $\mathbb{C}[x]$ and $\mathbb{C}[y]$, respectively, and let $X = V(P)$ and $Y = V(Q)$ be their zero sets in $\mathbb{A}^m$ and $\mathbb{A}^n$, respectively. Then $I = (P, Q)$ is the ideal of all elements of $\mathbb{C}[x, y]$ that vanish on the variety $X \times Y$. Therefore I is a prime ideal.*

The fact that $V(I)$ is the variety $X \times y$ tells us only that the radical of I is a prime ideal.

proof of Proposition 3.3.4 The ideal generated by P in $\mathbb{C}[x, y]$ consists of finite sums of products of elements of P with polynomials in x, y. Let's denote that ideal by P too. Similarly, let Q denote the ideal generated in $\mathbb{C}[x, y]$ as well as the ideal of $\mathbb{C}[y]$. Let $A = \mathbb{C}[x]/P$, $B = \mathbb{C}[y]/Q$, and $R = \mathbb{C}[x, y]/I$. (The ring R is the tensor product algebra $A \otimes B$ (see(**??**).)

When we evaluate polynomials $p(x, y)$ at a point y^0 of Y, we obtain an algebra R^0 that is isomorphic to A. Therefore the homomorphism $A \to R$ defined by the inclusion $\mathbb{C}[x] \subset \mathbb{C}[x, y]$ is injective. Similarly, we have an injective map $B \to R$. Let's identify A and B with their images in R. Together, these images generate R. Any element p of R can be written as a finite sum

$$p = \sum_{i=1}^{k} a_i b_i \tag{3.3.5}$$

with a_i in A and b_i in B. We show that if p vanishes identically on $X \times Y$, then $p = 0$. To do this, we show that the same element p can be written as a sum of $k - 1$ products.

If a_k is zero, then $p = \sum_{i=1}^{k-1} a_i b_i$. Suppose that $a_k \neq 0$. Then a_k isn't identically zero on X. We choose a point x^0 of X such that $a_k(x^0) \neq 0$. Then, writing $a_i(x^0) = a_i^0$ and $p^0(y) = p(x^0, y)$, we have the equation $p^0(y) = \sum_{i=1}^{k} a_i^0 b_i$. Since p vanishes on $X \times Y$, p^0 vanishes on Y. Therefore $p^0 = 0$. Then since $a_k^0 \neq 0$, we can solve the equation $\sum_{i=1}^{k} a_i^0 b_i = 0$ for b_k: $b_k = \sum_{i=1}^{k-1} c_i b_i$, where $c_i = -a_i^0/a_k^0$. Substituting into p gives us an expression for p as a sum of $k - 1$ terms. Finally, when $k = 1$, $a_1^0 b_1 = 0$. Therefore $b_1 = 0$, and $p = 0$. □

(3.3.6) the Zariski topology on $\mathbb{P}^m \times \mathbb{P}^n$

As mentioned above (**3.1.10**), the product of projective spaces $\mathbb{P}^m \times \mathbb{P}^n$ is made into a projective variety by identifying it with its Segre image, the locus of the Segre equations $w_{ij} w_{k\ell} = w_{i\ell} w_{kj}$. However, its Zariski topology merits discussion.

The first examples of closed subsets of $\mathbb{P}^m \times \mathbb{P}^n$ are products of the form $X \times Y$, where X is a closed subset of $\mathbb{P}^m$ and Y is a closed subset of $\mathbb{P}^n$. These are the closed sets in the *product topology* on $\mathbb{P}^m \times \mathbb{P}^n$. The product topology is much coarser than the Zariski topology. For example, the proper closed subsets of $\mathbb{P}^1$ are the nonempty finite subsets. In the product topology, the proper closed subsets of $\mathbb{P}^1 \times \mathbb{P}^1$ are finite unions of points and sets of the form $\mathbb{P}^1 \times q$ and $p \times \mathbb{P}^1$ ('horizontal' and 'vertical' lines). Most Zariski closed subsets of $\mathbb{P}^1 \times \mathbb{P}^1$ aren't of this form.

Since $\mathbb{P}^m \times \mathbb{P}^n$, with its Segre embedding, is a projective variety, we don't really need a separate definition of its Zariski topology. Its closed subsets are the zero sets of families of homogeneous polynomials in the Segre variables w_{ij} that include the Segre equations.

One can also describe the closed subsets of $\mathbb{P}^m \times \mathbb{P}^n$ directly, in terms of bihomogeneous polynomials. A polynomial $f(x,y)$ is *bihomogeneous* if it is homogeneous in the variables x and also in the variables y. For example, the polynomial $x_0^2 y_0 + x_0 x_1 y_1$ is bihomogeneous, of degree 2 in x and degree 1 in y.
Because (x,y) and $(\lambda x, \mu y)$ represent the same point of $\mathbb{P}^m \times \mathbb{P}^n$ for all nonzero λ and μ, we want to know that $f(x,y) = 0$ if and only if $f(\lambda x, \mu y) = 0$, and this is true for all nonzero λ and μ if and only if f is bihomogeneous.

3.3.7. Proposition. *Let x and y be coordinates in $\mathbb{P}^m$ and $\mathbb{P}^n$, respectively.*

(i) *A subset of $\mathbb{P}^m \times \mathbb{P}^n$ is closed if and only if it is the locus of zeros of a family of bihomogeneous polynomials.*

(ii) *If X and Y are closed subsets of $\mathbb{P}^m$ and $\mathbb{P}^n$, respectively, then $X \times Y$ is a closed subset of $\mathbb{P}^m \times \mathbb{P}^n$.*

(iii) *The projections $\mathbb{P}^m \times \mathbb{P}^n_y \to \mathbb{P}^m$ and $\mathbb{P}^m \times \mathbb{P}^n_y \to \mathbb{P}^n$ are continuous maps.*

(iv) *For all x in $\mathbb{P}^m$ and all y in $\mathbb{P}^n$, the fibres $x \times \mathbb{P}^n$ and $\mathbb{P}^m \times y$, with topologies induced from $\mathbb{P}^m_x \times \mathbb{P}^n_y$, are homeomorphic to $\mathbb{P}^n$ and $\mathbb{P}^m$, respectively.*

proof. **(i)** For the proof, we denote the Segre image of $\mathbb{P}^m \times \mathbb{P}^n$ by Π. Let $f(w)$ be a homogeneous polynomial in the Segre variables w_{ij}. When we substitute $w_{ij} = x_i y_j$ into into f, we obtain a polynomial $f(x_i y_j)$ that is bihomogeneous and that has the same degree as f in x and in y. Let's denote that bihomogeneous polynomial by $\widetilde{f}(x,y)$. The inverse image of the zero set of f in Π is the zero set of $\widetilde{f}$ in $\mathbb{P}^m \times \mathbb{P}^n$. Therefore the inverse image of a closed subset of Π is the zero set of a family of bihomogeneous polynomials in $\mathbb{P}^m \times \mathbb{P}^n$.

Conversely, let $g(x,y)$ be a bihomogeneous polynomial, say of degrees r and s in x and y, respectively. If $r = s$, we may collect variables that appear in g in pairs $x_i y_j$ and replace each pair $x_i y_j$ by w_{ij}. We will obtain a homogeneous polynomial G in w such that $G(w) = g(x,y)$ when $w_{ij} = x_i y_j$. The zero set of G in Π is the image of the zero set of g in $\mathbb{P}^m \times \mathbb{P}^n$.

Suppose that $r \geq s$, and let $k = r - s$. Because the variables y cannot all be zero at any point of $\mathbb{P}^n$, the equation $g = 0$ on $\mathbb{P}^m \times \mathbb{P}^n$ is equivalent with the system of equations $g y_0^k = g y_1^k = \cdots = g y_n^k = 0$. The polynomials $g y_i^k$ are bihomogeneous, of same degree in x as in y.

(ii) A polynomial $f(x)$ can be viewed as a bihomogeneous polynomial of degree zero in y, and a polynomial $g(y)$ is bihomogeneous of degreee zxero in x. So $X \times Y$, which is the locus $f = g = 0$ in $\mathbb{A}^{m+n}$, is a closed subset of $\mathbb{P}^m \times \mathbb{P}^n$.

(iii) We look at the projection $\mathbb{P}^m \times \mathbb{P}^n \to \mathbb{P}^m$. If X is the closed subset of $\mathbb{P}^m$ defined by a system of homogeneous polynomials $f_i(x)$, its inverse image in $\mathbb{P}^m \times \mathbb{P}^n$ is the zero set of the system of bihomogeneous equations $f_i(x) y_0, ..., f_i(x) y_n$. So the inverse image is closed, which shows that the projection is continuous.

(iv) It suffices to show that the inclusion map $X \to \mathbb{P}^m \times \mathbb{P}^n$ that sends X to $\mathbb{P}^m \times y$ is continuous. If $f(x,y)$ is a bihomogeneous polynomial and y^0 is a point of Y, the zero set of f in $\mathbb{P}^m \times y^0$ is the zero set of $f(x, y^0)$. This polynomial defines a closed subset of $\mathbb{P}^m$. □

3.3.8. Corollary. *Let X and Y be projective varieties, and let P denote the product $X \times Y$, a closed subset of $\mathbb{P}^m \times \mathbb{P}^n$.*

(i) *The projections $P \to X$ and $P \to Y$ are continuous.*

(ii) *For all x in X and all y in Y, the fibres ${}_xP = x \times Y$ and $P_y = X \times y$, with topologies induced from P, are homeomorphic to Y and X, respectively.* □

The next corollary follows from Proposition 3.3.1 and Corollary 3.3.8.

3.3.9. Corollary. *If X and Y are projective varieties, so is $X \times Y$.* □

We will discuss the mapping property of a product in Chapter 5.

3.4 Morphisms and Isomorphisms

When defining morphisms varieties, one must keep in mind that points of projective space are equivalence classes of vectors, not the vectors themselves. This is a complication that turns out to be very useful.

Some morphisms are sufficiently obvious that they don't require discussion. They include the projection from a product variety $X\times Y$ to X, the inclusion of X into the product $X\times Y$ as the set $X\times y$ for some point y of Y, the morphism of products $X\times Y\to X'\times Y$ when a morphism $X\to X'$ is given, and of course, the analogous maps when Y replaces X.

If X and Y are subvarieties of projective spaces $\mathbb{P}^m$ and $\mathbb{P}^n$, respectively, a morphism $Y\to X$ will be determined by a morphism from Y to $\mathbb{P}^m$ whose image is contained in X. However, it is an important fact that a morphism $Y\xrightarrow{f}X$ needn't be the restriction of a morphism from $\mathbb{P}^n$ to $\mathbb{P}^m$. There will often be no way to extend the morphism from Y to $\mathbb{P}^n$. Or, put another way, it may not be possible to define f using polynomials in the coordinate variables of $\mathbb{P}^n$.

For example, The Veronese map from the projective line $\mathbb{P}^1$ to $\mathbb{P}^2$, defined by $(x_0,x_1)\rightsquigarrow(x_0^2,x_0x_1,x_1^2)$, is an obvious morphism. Its image is the conic $C:\ v_{00}v_{11}-v_{01}^2=0$ in the projective plane $\mathbb{P}^2$ with coordinates v_{00},v_{01},v_{11}. The Veronese defines a bijective morphism $\mathbb{P}^1\xrightarrow{f}C$. Its inverse function sends a point (v_{00},v_{01},v_{11}) of C with $v_{00}\neq 0$ to the point $(x_0,x_1)=(v_{01},v_{11})$. There is no way to extend this inverse function to $\mathbb{P}^2$, though it is a morphism. In fact, there is no nonconstant morphism from $\mathbb{P}^2$ to $\mathbb{P}^1$.

In order to have a definition that includes all cases, we will define morphisms using points with values in a field.

(3.4.1) the function field

Let X be a projective variety, and let X^i be its intersection with the standard affine open subset $\mathbb{U}^i$ of projective space. Then if nonempty, X^i will be an irreducible closed subset of $\mathbb{U}^i$, an affine variety. Let's omit the indices for which X^i is empty. Then the intersection $X^{ij}=X^i\cap X^j$ will be a localization of X^i and also a localization of X^j. If $X^i=\operatorname{Spec}A_i$ and if $u_{ij}=x_j/x_i$, then $X^{ij}=\operatorname{Spec}A_{ij}$, and $A_{ij}=A_i[u_{ij}^{-1}]=A_j[u_{ji}^{-1}]$. The fields of fractions of the coordinate algebras A_i are equal for all i such that X^i isn't empty.

3.4.2. Definition. The *function field* K_X of a projective variety X is the field of fractions of the coordinate algebra A_i of any one of its nonempty affine open subsets $X^i=X\cap\mathbb{U}^i$. A *rational function* on a variety X is a nonzero element of its function field.

A point p of a projective variety X will lie in one of the nonempty affine open sets $X^i=X\cap\mathbb{U}^i$. A rational function α on X is *regular* at p if it is a regular function at p on one of those affine open sets. □

Suppose that X is affine: $X=\operatorname{Spec}A$. As has been noted, we may regard X as a quasiprojective variety by embedding it as a closed subset of $\mathbb{U}^0$. The function field of X will be the field of fractions of its coordinate algebra A, and a rational function α on X will be *regular* at a point p of X if it can be written as a fraction $\alpha=a/s$, where a and s are in A and s isn't zero at p. Thus α is regular at p if it is an element of the coordinate algebra of some localization X_s that contains p.

We note in passing that if α is regular at p and $\alpha(p)\neq 0$, then α^{-1} will be regular at p too.

3.4.3. Lemma. *The regularity of a rational function at p doesn't depend on the choice of the open set X^i that contains p.* □

Finally, let X be a (quasiprojective) variety. The *function field* K_X of X is the function field of its closure $\overline{X}$ in projective space.

(3.4.4) the function field of a product

To define the function field of a product $X\times Y$ of projective varieties, we use the Segre embedding $\mathbb{P}^m_x\times\mathbb{P}^n_y\to\mathbb{P}^N$. We use notation as in (**3.1.10**), and let's denote the product $X\times Y$ by Π. So x_i, y_j, and w_{ij} are coordinates in the three projective spaces, and the Segre map is defined by $w_{ij}=x_iy_j$. Let $\mathbb{U}^i$, $\mathbb{V}^j$, and $\mathbb{W}^{ij}$

be the standard affine open sets $x_i \neq 0$, $y_j \neq 0$ and $w_{ij} \neq 0$. The function field will be the field of fractions of the nonempty intersections $\Pi \cap \mathbb{W}^{ij} = \Pi^{ij}$, and $\Pi^{ij} \approx X^i \times Y^j$, where $X^i = X \cap \mathbb{U}^i$ and $Y^j = Y \cap \mathbb{V}^j$. Since Π^{ij}, X^i, and Y^j are affine varieties, the function field of the product $\Pi = X \times Y$ is the field of fractions of any one of the nonempty affine open sets Π^{iJ}.

Since $\Pi^{ij} = X^i \cap Y^j$, all that remains to do is to describe the field of fractions of a product of affine varieties $\Pi = X \times Y$, when $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} B$. If $A = \mathbb{C}[x]/P$ and $B = \mathbb{C}[y]/Q$, the coordinate algebra of Π is the algebra $\mathbb{C}[x, y]/(P, Q)$. As mentioned before, this is the tensor product algebra $A \otimes B$. We don't need to know much about the tensor product algebra here, but let's use the tensor product notation.

The function field K_X of X is the field of fractions of the coordinate algebra A. Similarly, K_Y is the field of fractions of B and $K_{X \times Y}$ is the field of fractions of $A \otimes B$. The one important fact to note is that $K_{X \times Y}$ isn't generated by K_X and K_Y. For example, if $A = \mathbb{C}[x]$ and $B = \mathbb{C}[y]$ (one x and one y), then $K_{X \times Y}$ is the field of rational functions in two variables $\mathbb{C}(x, y)$. The algebra generated by the fraction fields $\mathbb{C}(x)$ and $\mathbb{C}(y)$ consists of the rational functions $p(x, y)/q(x, y)$ in which q is a product fg, where f is a polynomial in x and g is a polynomial in y. Most rational functions, $1/(x + y)$ forexample, aren't of this type.

But, $K_{X \times Y}$ is the fraction field of $A \otimes B$.

(3.4.5) interlude: rational functions on projective space

Let R denote the polynomial ring $\mathbb{C}[x_0, ..., x_n]$. If f is a homogeneous polynomial of positive degree d, it makes sense to say that f vanishes at a point of $\mathbb{P}^n$, because $f(\lambda x) = \lambda^d f(x)$. But f doesn't define a function on $\mathbb{P}^n$. On the other hand, a fraction g/h of homogeneous polynomials of the same degree d does define a function wherever h isn't zero, because

$$g(\lambda x)/h(\lambda x) = \lambda^d g(x)/\lambda^d h(x) = g(x)/h(x)$$

A *homogeneous fraction* f is a fraction of homogeneous polynomials. The *degree* of a homogeneous fraction $f = g/h$ is the difference of degrees: $\deg f = \deg g - \deg h$.

3.4.6. Definition. A homogeneous fraction f is *regular* at a point p of $\mathbb{P}^n$ if, when it is written as a fraction g/h of relatively prime homogeneous polynomials, the denominator h isn't zero at p, and f is *regular* on a subset U if it is regular at every point of U. □

This definition agrees with the one given above, in Definition 3.4.2.

3.4.7. Lemma. (i) *Let h be a homogeneous polynomial of positive degree d, and let V be the open subset of $\mathbb{P}^n$, of points at which h isn't zero. The nonzero rational functions that are regular on V are those of the form g/h^k, where $k \geq 0$ and g is a homogeneous polynomial of degree dk.*
(ii) *The only rational functions that are regular at every point of $\mathbb{P}^n$ are the constant functions.*

For example, the homogeneous polynomials that are nonzero at every point of the standard affine open set $\mathbb{U}^0$ are the scalar multiples of powers of x_0. So the rational functions that are regular on $\mathbb{U}^0$ are those of the form g/x_0^k, g homogeneous of degree k. This agrees with the fact that the coordinate algebra of $\mathbb{U}^0$ is the polynomial ring $\mathbb{C}[u_1, ..., u_n]$, with $u_i = x_i/x_0$: $g(x_0, ..., x_m)/x_0^k = g(u_0, ..., u_n)$ (with $u_0 = 1$).

proof of Lemma 3.4.7 **(i)** Let α be a regular function on the open set U, say g_1/h_1, where g_1 and h_1 are relatively prime homogeneous polynomials. Then h_1 doesn't vanish on U, so its zero locus in $\mathbb{P}^n$ is contained in the zero locus of h. According to the Strong Nullstellensatz 3.2.7, h_1 divides a power of h, say $h^k = fh_1$. Then $g_1/h_1 = fg_1/fh_1 = fg_1/h^k$.

(ii) If a rational function f is regular at every point of $\mathbb{P}^n$, then it is regular on $\mathbb{U}^0$. It will have the form g/x_0^k, g has degree k and isn't divisible by x_0. And since f is also regular on $\mathbb{U}^1$, it will have the form h/x_1^ℓ, where x_1 doesn't divide h. Then $gx_1^\ell = hx_0^k$. Since x_0 doesn't divide g, $k = 0$, g is a constant, and $f = g$. □

It is also true that the only rational functions on a projective variety X that are regular at every point of X are the constant functions. The proof of this will be given later (see Corollary 8.4.8). When studying projective varieties, the constant functions are useless. One has to look at at regular functions on open subsets.

One way that affine varieties appear in projective algebraic geometry is as open subsets that have enough regular functions.

(3.4.8) points with values in a field

Let K be a field that contains the complex numbers. A *point* of projective space $\mathbb{P}^n$ *with values in* K is an equivalence class of nonzero vectors $\alpha = (\alpha_0, ..., \alpha_n)$ with α_i in K, the equivalence relation being analogous to the one for ordinary points: $\alpha \sim \alpha'$ if $\alpha' = \lambda\alpha$ for some λ in K.

If a closed subvariety X of $\mathbb{P}^n$ is defined by a set of homogeneous polynomial equations $f(x) = 0$, a point α of X with values in a field K is a point of $\mathbb{P}^n$ with values in K, such that $f(\alpha) = 0$.

If X is quasiprojective and C is the complement of X in its closure $\overline{X}$ in projective space, then as for ordinary points (3.2.10), a *point of* X *with values in* K_X is a point of $\overline{X}$ with values in K_X that isn't a point of C.

3.4.9. Lemma. *Let X be a projective variety, a subvariety of $\mathbb{P}^n$, and let K_X be the function field of X.*
(i) *The projective embedding of X defines a point $(\alpha_0, ..., \alpha_n)$ of $\mathbb{P}^n$ with values in the function field K_X.*
(ii) *A homogeneous polynomial $f(x_0, ..., x_n)$ vanishes at every point of X if and only if $f(\alpha) = 0$.*

proof. **(i)** Let $X^i = X \cap \mathbb{U}^i$ be the intersection with a standard affine open set, and assume that X^i is nonempty. Let A_i be its coordinate algebra. The embedding $X^i \subset \mathbb{U}^i$ is defined by a homomorphism $\mathbb{C}[u] \to A_i$, $u = u_0, ..., u_n$ and $u_j = x_j/x_i$. The function field K_X is the field of fractions of A_i, so we have a composed homomorphism $\mathbb{C}[u] \to A_i \to K_X$. The point α is the image of $(u_0, ..., u_n)$ via this homomorphism.

(ii) If $f(x) = 0$, then $f(u) = 0$ and therefore $f(\alpha) = 0$. □

In what follows, it will be helpful to have a separate notation for the point with values in K determined by a nonzero vector α. We'll denote that point by $\underline{\alpha}$. Thus $\underline{\alpha} = \underline{\alpha}'$ if $\alpha' = \lambda\alpha$ for some nonzero λ in K. We'll drop this notation later.

(3.4.10) morphisms to projective space

Let K be the function field of a variety Y, and let $\alpha = (\alpha_0, ..., \alpha_n)$ be a nonzero vector with entries in K. We try to define a morphism from Y to projective space $\mathbb{P}^n$ using the point $\underline{\alpha}$. To define the image $\underline{\alpha}(q)$ of a point q of Y (an ordinary point), we look for a vector $\alpha' = (\alpha'_0, ..., \alpha'_n)$, with $\underline{\alpha}' = \underline{\alpha}$, i.e., $\alpha' = \lambda\alpha$, such that the rational functions α'_i are all regular and not all zero at q. Such a vector may exist or not. If it exists, we define

$$\underline{\alpha}(q) = (\alpha'_0(q), ..., \alpha'_n(q)) \quad \big(= \underline{\alpha}'(q)\big) \tag{3.4.11}$$

If such a vector α' exists for every point q of Y, we call $\underline{\alpha}$ a *good point*.

3.4.12. Lemma. *A point $\underline{\alpha}$ of $\mathbb{P}^n$ with values in the function field K_Y of Y is a good point if either one of the two following conditions holds for every point q of Y:*

- *There is an element λ in K_Y such that the rational functions $\alpha'_i = \lambda\alpha_i$, $i = 0, ..., n$, are regular and not all zero at q.*
- *There is an index j, $0 \le j \le n$, such that the rational functions α_i/α_j, $j = 0, ..., n$, are regular at q.*

proof. The first condition simply restates the definition. We show that it is equivalent with the second one.

Suppose that α_i/α_j are regular at q for all i. Let $\lambda = \alpha_j^{-1}$, and let $\alpha'_i = \lambda\alpha_i = \alpha_i/\alpha_j$. The rational functions α'_i are regular at q, and they aren't all zero there because $\alpha'_j = 1$.

Conversely, suppose that $\alpha'_i = \lambda\alpha_i$ are all regular at q and that α'_j isn't zero there. Then ${\alpha'_j}^{-1}$ is a regular function at q, so the rational functions α'_i/α'_j, which are equal to α_i/α_j, are regular at q for all i. □

3.4.13. Lemma. *With notation as in (3.4.11), the point $\underline{\alpha}(q)$ is independent of the choice of the vector α'.*

proof. Let $\alpha' = \lambda\alpha$ and $\alpha'' = \mu\alpha$. Suppose that α'_i are all regular and not all zero at q, and that α''_i are also regular and not all zero there. We need to show that $\underline{\alpha}' = \underline{\alpha}''$, and we may assume that $\alpha'' = \alpha$. Then $\alpha' = \lambda\alpha$. The rational functions α_i and α'_i are all regular at q, and there are indices j, k such that $\alpha_j(q)$ and $\alpha'_k(q)$ are nonzero. Then $\lambda = \alpha'_j/\alpha_j$ and $\lambda^{-1} = \alpha_k/\alpha'_k$ are both regular at q. So $\lambda(q) \neq 0$, $\alpha'(q) = \lambda(q)\alpha(q)$, and $\underline{\alpha}'(q) = \underline{\alpha}(q)$. □

3.4.14. Definition. Let Y be a variety with function field K_Y. A *morphism* from Y to projective space $\mathbb{P}^n$ is a map that is defined by a good point $\underline{\alpha}$ with values in K_Y, as in (3.4.11).

3.4.15. Example. *The identity map* $\mathbb{P}^1 \to \mathbb{P}^1$.
Let $X = \mathbb{P}^1$, and let (x_0, x_1) be coordinates in X. The function field of X is the field $K = \mathbb{C}(t)$ of rational functions in the variable $t = x_1/x_0$. The identity map $X \to X$ is the map $\underline{\alpha}$ defined by the point $\alpha = (1, t)$ with values in K. For every point p of X except the point $(1, 0)$, $\alpha(p)$ is defined an not zero, so $\underline{\alpha}(p) = \alpha(p)$. At the point $(0, 1)$, $\alpha' = (t^{-1}, 1) = t^{-1}\alpha$ defines $\underline{\alpha}$. □

(3.4.16) morphisms to quasiprojective varieties

3.4.17. Definition. Let Y be a variety, and let X be a subvariety of a projective space $\mathbb{P}^n$. A morphism of varieties $Y \xrightarrow{\alpha} X$ is the restriction of a morphism $Y \xrightarrow{\alpha} \mathbb{P}^n$ whose image is contained in X.

Thus if a projective variety X is the locus of zeros of a family f of homogeneous polynomials, a morphism $Y \xrightarrow{\alpha} \mathbb{P}^n$ defines a morphism $Y \to X$ if $f(\alpha) = 0$.

A word of caution: A morphism $Y \xrightarrow{\alpha} X$ won't define a map on function fields $K_X \to K_Y$ unless the image of Y is dense in X.

3.4.18. Proposition. *A morphism of varieties $Y \xrightarrow{\alpha} X$ is a continuous map in the Zariski topology, and also in the classical topology.*

proof. Let $\mathbb{U}^i$ be the standard affine open subset of $\mathbb{P}^m$, and let Y^i be an affine open subset of the inverse image of $\mathbb{U}^i$. If $X = \mathbb{P}^m$, the restriction $Y^i \to \mathbb{U}^i$ of $\underline{\alpha}$ is continuous in both topologies because it is a morphism of affine varieties. Since Y can be covered by affine open sets such as Y^i, $\underline{\alpha}$ is continuous. Continuity for a morphism to a subvariety X of $\mathbb{P}^m$ follows because the topology on X is the induced topology. □

3.4.19. Proposition. *Let X, Y, and Z be varieties and let $Z \xrightarrow{\beta} Y$ and $Y \xrightarrow{\alpha} X$ be morphisms. The composed map $Z \xrightarrow{\alpha\beta} X$ is a morphism.*

proof. Say that X is a subvariety of $\mathbb{P}^m$. The morphism α is the restriction of a morphism $Y \to \mathbb{P}^m$ whose image is in X, and that is defined by a good point $\underline{\alpha}$, $\alpha = (\alpha_0, ..., \alpha_m)$ of $\mathbb{P}^m$ with values in the function field K_Y of Y. Similarly, if Y is a subvariety of $\mathbb{P}^n$, the morphism $\underline{\beta}$ is the restriction of a morphism $Z \to \mathbb{P}^n$ whose image is contained in Y, and that is defined by a good point $\underline{\beta}$, $\beta = (\beta_0, ..., \beta_n)$ of $\mathbb{P}^n$ with values in the function field K_Z of Z.

Let z be a point (an ordinary point) of Z. Since $\underline{\beta}$ is a good point, we may adjust β by a factor in K_Z so that the rational functions β_i are regular and not all zero at z. Then $\underline{\beta}(z)$ is the point $(\beta_0(z), ..., \beta_n(z))$ of Y. Let's denote that point by $q = (q_0, ..., q_n)$. So $q_i = \beta_i(z)$. The elements α_j are rational functions on Y. We may adjust α by a factor in K_Y, so that they are regular and not all zero at q. Then $[\underline{\alpha\beta}](z) = \underline{\alpha}(q) = (\alpha_0(q), ..., \alpha_m(q))$, and $\alpha_j(q) = \alpha_j(\beta_0(z), ..., \beta_n(z)) = \alpha_j(\beta(z))$ are not all zero. When these adjustments have been made, the point of $\mathbb{P}^m$ with values in K_Z that defines $\underline{\alpha\beta}$ is $(\alpha_0(\beta(z)), ..., \alpha_m(\beta(z)))$. □

This next is a lemma of topology.

3.4.20. Lemma. *Let $\{X^i\}$ be a covering of a topological space X by open sets. A subset Y of X is open (or closed) if and only if $Y \cap X^i$ is open (or closed) in X^i for every i. In particular, if $\{\mathbb{U}^i\}$ is the standard affine cover of $\mathbb{P}^n$, a subset Y of $\mathbb{P}^n$ is open (or closed) if and only if $Y \cap \mathbb{U}^i$ is open (closed) in $\mathbb{U}^i$ for every i.* □

3.4.21. Lemma.

(i) *The inclusion of an open or a closed subvariety Y into a variety X is a morphism.*

(ii) *Let $Y \xrightarrow{f} X$ be a map whose image lies in an open or a closed subvariety Z of X. Then f is a morphism if and only if its restriction $Y \to Z$ is a morphism.*

(iii) *Let $\{Y^i\}$ be open an open covering of a variety Y, and let $Y^i \xrightarrow{f^i} X$ be morphisms. If the restrictions of f^i and f^j to the intersections $Y^i \cap Y^j$ are equal for all i, j, there is a unique morphism f whose restriction to Y^i is f^i.* □

(3.4.22) isomorphisms

A bijective morphism $Y \xrightarrow{u} X$ of quasiprojective varieties whose inverse function is also a morphism is an *isomorphism*. Isomorphisms are important because they allow us to identify different incarnations of the "same variety", i.e., to describe an isomorphism class of varieties. For example, the projective line $\mathbb{P}^1$, a conic in $\mathbb{P}^2$, and the twisted cubic in $\mathbb{P}^3$ are isomorphic.

3.4.23. Example.
Let y_0, y_1 be coordinates in $Y = \mathbb{P}^1$. As before, the function field of Y is the field $K = \mathbb{C}(t)$ of rational functions in $t = y_1/y_0$. The degree 3 Veronese map $Y \longrightarrow \mathbb{P}^3$ (**3.1.16**) defines an isomorphism of Y to its image, a twisted cubic X. The Veronese map is defined by the point $\alpha = (1, t, t^2, t^3)$ of $\mathbb{P}^3$ with values in K. On the open set $\{y_0 \neq 0\}$ of Y, the rational functions $1, t, t^2, t^3$ are regular and not all zero. Let $\lambda = t^{-3}$ and $\alpha' = \lambda\alpha = (t^{-3}, t^{-2}, t^{-1}, 1)$. The functions t^{-k} are regular on the open set $\{y_1 \neq 0\}$. So α is a good point. It defines a morphism $Y \xrightarrow{\alpha} X$.

The twisted cubic X is the locus of zeros of the equations (3.1.17).

$$v_0v_2 = v_1^2 \ , \ v_2v_1 = v_0v_3 \ , \ v_1v_3 = v_2^2$$

To identify the function field K_1 of X, we put $v_0 = 1$, obtaining relations $v_2 = v_1^2, v_3 = v_1^3$. Then K_1 is the field $\mathbb{C}(v_1)$, and the point of $Y = \mathbb{P}^1$ with values in K_1 that defines the inverse of the morphism $\underline{\alpha}$ is $\beta = (1, v_1)$. □

3.4.24. Lemma. *Let $Y \xrightarrow{f} X$ be a morphism of varieties, let $\{X^i\}$ be an open covering of X, and let $Y^i = f^{-1}X^i$. If the restrictions $Y^i \xrightarrow{f^i} X^i$ of f are isomorphisms, then f is an isomorphism.*

proof. Let g^i denote the inverse of the morphism f^i. Then $g^i = g^j$ on $X^i \cap X^j$ because $f^i = f^j$ on $Y^i \cap Y^j$. By (3.4.21) **(iii)**, there is a unique morphism $X \xrightarrow{g} Y$ whose restriction to Y^i is g^i. That morphism is the inverse of f. □

(3.4.25) the diagonal

Let X be a variety. The *diagonal* X_Δ, the set of points (p, p) in $X \times X$ is an example of a subset of $X \times X$ that is closed in the Zariski topology, but not in the product topology.

3.4.26. Proposition. *Let X be a variety. The diagonal X_Δ is a closed subvariety of the product variety $X \times X$.*

proof. Let $\mathbb{P}$ denote the projective space $\mathbb{P}^n$ that contains X, and let $x_0, ..., x_n$ and $y_0, ..., y_n$ be coordinates in the two factors of $\mathbb{P} \times \mathbb{P}$. The diagonal $\mathbb{P}_\Delta$ in $\mathbb{P} \times \mathbb{P}$ is the closed subvariety defined by the bilinear equations $x_iy_j = x_jy_i$, or in the Segre variables, by the equations $w_{ij} = w_{ji}$, which show that the ratios x_i/x_j and y_i/y_j are equal.

Next, suppose that X is the closed subvariety of $\mathbb{P}$ defined by a system of homogeneous equations $f(x) = 0$. The diagonal X_Δ can be identified as the intersection of the product $X \times X$ with the diagonal $\mathbb{P}_\Delta$ in $\mathbb{P} \times \mathbb{P}$, so it is a closed subvariety of $X \times X$. As a closed subvariety of $\mathbb{P} \times \mathbb{P}$, the diagonal X_Δ is defined by the equations

$$x_iy_j = x_jy_i \quad \text{and} \quad f(x) = 0 \tag{3.4.27}$$

The equations $f(y) = 0$ are redundant. Finally, X_Δ is irreducible because it is homeomorphic to X. $\square$

It is interesting to compare Proposition 3.4.26 with the Hausdorff condition for a topological space. The proof of the next lemma is often given as an exercise in topology.

3.4.28. Lemma. *A topological space X is a Hausdorff space if and only if, when $X\times X$ is given the product topology, the diagonal X_Δ is a closed subset of $X\times X$.* $\square$

Though a variety X with its Zariski topology isn't a Hausdorff space unless it is a point, Lemma 3.4.28 doesn't contradict Proposition 3.4.26 because the Zariski topology on $X\times X$ is finer than the product topology.

(3.4.29) the graph of a morphism

Let $Y \xrightarrow{f} X$ be a morphism of varieties. The *graph* Γ of f is the subset of $Y\times X$ of pairs (q, p) such that $p = f(q)$.

3.4.30. Proposition. *The graph Γ_f of a morphism $Y \xrightarrow{f} X$ is a closed subvariety of $Y\times X$, and it is isomorphic to Y.*

proof. We form a diagram of morphisms

(3.4.31)
$$\begin{array}{ccc} \Gamma_f & \longrightarrow & Y\times X \\ {\scriptstyle v}\downarrow & & \downarrow{\scriptstyle f\times id} \\ X_\Delta & \longrightarrow & X\times X \end{array}$$

where v sends a point (q, p) of Γ_f with $f(q) = p$ to (p, p). The graph Γ_f is the inverse image in $Y\times X$ of the diagonal X_Δ. Since the diagonal is closed in $X\times X$, Γ_f is closed in $Y\times X$.

Let π_1 denote the projection from $X\times Y$ to Y. The composition of the morphisms $Y \xrightarrow{(id,f)} Y\times X \xrightarrow{\pi_1} Y$ is the identity map on Y, and the image of the map (id, f) is the graph Γ_f. Therefore Y maps bijectively to Γ_f. The two maps $Y \to \Gamma_f$ and $\Gamma_f \to Y$ are inverses, so Γ_f is isomorphic to Y. $\square$

(3.4.32) projection

The map

(3.4.33)
$$\mathbb{P}^n \xrightarrow{\pi} \mathbb{P}^{n-1}$$

that drops the last coordinate of a point: $\pi(x_0, ..., x_n) = (x_0, ..., x_{n-1})$ is called a *projection*. It is defined at all points of $\mathbb{P}^n$ except at the point $q = (0, ..., 0, 1)$, which is called the *center of projection*. So π is a morphism from the complement $U = \mathbb{P}^n - \{q\}$ to $\mathbb{P}^{n-1}$.

Let the coordinates in $\mathbb{P}^n$ and $\mathbb{P}^{n-1}$ be $x = x_0, ..., x_n$ and $y = y_0, ..., y_{n-1}$, respectively. The fibre $\pi^{-1}(y)$ over a point $(y_0, ..., y_{n-1})$ is the set of points $(x_0, ..., x_n)$ such that $(x_0, ..., x_{n-1}) = (\lambda y_0, ..., \lambda y_{n-1})$, while x_n is arbitrary. It is the line in $\mathbb{P}^n$ through the points $(y_1, ..., y_{n-1}, 0)$ and $q = (0, ..., 0, 1)$, with the center of projection q omitted.

The graph Γ of π in $U\times\mathbb{P}^{n-1}_y$ is the locus of solutions of the equations $w_{ij} = w_{ji}$ for $0\le i, j\le n-1$, which imply that the vectors $(x_0, ..., x_{n-1})$ and $(y_0, ..., y_{n-1})$ are proportional.

3.4.34. Proposition. *In $\mathbb{P}^n_x \times \mathbb{P}^{n-1}_y$, the locus $\overline{\Gamma}$ of the equations $x_iy_j = x_jy_i$, or $w_{ij} = w_{ji}$, with $0 \le i, j \le n-1$, is the closure of the graph Γ of π.*

proof. The equations are true at points (x, y) of Γ at which $x \ne q$, and also at all points (q, y). So the locus $\overline{\Gamma}$, a closed set, is the union of the graph Γ and the set $q\times\mathbb{P}^{n-1}$. We must show that a homogeneous polynomial $g(w)$ that vanishes on Γ vanishes at all points of $q\times\mathbb{P}^{n-1}$. Given y, let $x = (ty_0, ..., ty_{n-1}, 1)$. For all $t \ne 0$, the point (x, y) is in Γ and therefore $g(x, y) = 0$. Since g is a continuous function, $g(x, y)$ approaches $g(q, y)$ as $t \to 0$. So $g(q, y) = 0$. $\square$

The projection $\overline{\Gamma} \to \mathbb{P}^n_x$ that sends a point (x, y) to x is bijective except when $x = q$. The fibre over q, which is $q\times\mathbb{P}^{n-1}$, is a projective space of dimension n–1. Because the point q of $\mathbb{P}^n$ is replaced by a projective space in $\overline{\Gamma}$, the map $\overline{\Gamma} \to \mathbb{P}^n_x$ is called a *blowup* of the point q.

figure: projection with closure of graph??

3.4.35. Proposition. *Let* $Y \xrightarrow{\alpha} X$ *and* $Z \xrightarrow{\beta} W$ *be morphisms of varieties. The product map* $Y\times Z \xrightarrow{\alpha\times\beta} X\times W$ *that sends* (y, z) *to* $(\underline{\alpha}(y), \underline{\beta}(z))$ *is a morphism*

proof. Let P and q be points of X and Y, respectively. We may assume that α_i are regular and not all zero at p and that β_j are regular and not all zero at q. Then, in the Segre coordinates w_{ij}, $[\alpha\times\beta](p, q)$ is the point $w_{ij} = \alpha_i(p)\beta_j(q)$. We must show that $\alpha_i\beta_j$ are all regular at (p, q) and are not all zero there. This follows from the analogous properties of α_i and β_j. □

3.5 Affine Varieties

We have used the term 'affine variety' in several contexts:

A closed subset of affine space $\mathbb{A}^n_x$ is an affine variety, the set of zeros of a prime ideal P of $\mathbb{C}[x]$. Its coordinate algebra is $A = \mathbb{C}[x]/P$.

The spectrum $\operatorname{Spec} A$ of a finite type domain A is an affine variety that becomes a closed subvariety of affine space when one chooses a presentation $A = \mathbb{C}[x]/P$.

An affine variety becomes a quasiprojective variety by identifying the ambient affine space $\mathbb{A}^n$ with the open subset $\mathbb{U}^0$ of projective space.

We combine these definitions now: An *affine variety* X is a variety that is isomorphic to a variety of the form $\operatorname{Spec} A$.

If $X = \operatorname{Spec} A$ is an affine variety with function field K, its coordinate algebra A will be the subalgebra of K of regular functions on X. So A and therefore $\operatorname{Spec} A$, are determined uniquely by X, and the isomorphism $\operatorname{Spec} A \to X$ is determined uniquely too. When X is affine, it seems permissible to identify X with $\operatorname{Spec} A$.

(3.5.1) regular functions on affine varieties

Let $X = \operatorname{Spec} A$ be an affine variety. Its function field K is the field of fractions of A. A rational function α is regular at a point p of X if it can be written as a fraction a/s where a, s are in A and $s(p) \neq 0$, and α is regular on X if it is regular at every point of X. On the other hand, in Chapter 2 (**2.6.1**), α is defined to be a regular function on X if and only if it is an element of the coordinate algebra A. The next lemma shows that the two conditions are equivalent.

3.5.2. Lemma. *The regular functions on an affine variety* $X = \operatorname{Spec} A$*, as defined in (3.4.2), are the elements of its coordinate algebra* A*.*

proof. Let α be a regular function on X, as defined above. So for every point p of X, there is a localization $X_s = \operatorname{Spec} A_s$ that contains p, such that α is an element of A_s. Because X is quasicompact, a finite set of these localizations, say $X_{s_1}, \dots, X_{s_k}$, will cover X. Then $s_1, ..., s_k$ have no common zeros on X, so they generate the unit ideal of A. Since α is in A_{s_i}, we can write $\alpha = s_i^{-n} b_i$ with b_i in A, and we can use the same exponent n for all i. Since the elements s_i generate the unit ideal of A, so do the powers s_i^n. Say that $\sum s_i^n a_i = 1$, with a_i in A. Then $\alpha = \sum s_i^n a_i \alpha = \sum a_i b_i$ is in A. □

3.5.3. Proposition.
(i) *Let* R *be the algebra of regular functions on a variety* Y*, and let* A *be a finite-type domain. A homomorphism* $A \to R$ *defines a morphism* $Y \xrightarrow{f} \operatorname{Spec} A$*.*
(ii) *When* X *and* Y *are affine varieties, say* $X = \operatorname{Spec} A$ *and* $Y = \operatorname{Spec} B$*, morphisms* $Y \to X$*, as defined in Definition 3.4.17 correspond bijectively to algebra homomorphisms* $A \to B$*, as in Definition 2.6.5.*

Note. Since Y isn't affine, all that we know about the algebra R is that its elements are rational functions that are regular on Y.

proof of Proposition 3.5.3. **(i)** Let $\{Y^i\}$ be an affine open covering of Y, and let R_i be the coordinate algebra of Y^i. The inclusions $A \subset R \subset R_i$ define morphisms $Y^i = \operatorname{Spec} R_i \xrightarrow{f^i} \operatorname{Spec} A$. It is true that $f^i = f^j$ on $Y^i \cap Y^j$, so Lemma 3.4.21 **(iii)** applies. □

3.5.4. Lemma. *Let X and Y be affine varieties, say $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} B$. Morphisms $Y \xrightarrow{\alpha} X$, as defined in (3.4.14) and (3.4.17), correspond bijectively to algebra homomorphisms $A \xrightarrow{\varphi} B$.*

proof. We choose a presentation of A, to embed X as a closed subvariety of affine space, and we identify that affine space with the standard affine open set $\mathbb{U}^0$ of $\mathbb{P}^n$. Let K be the function field of Y – the field of fractions of B. A morphism $Y \xrightarrow{u} X$ is determined by a good point α with values in K, and since $\alpha_0 \neq 0$, we may suppose that this point has the form $\alpha = (\alpha_0, ..., \alpha_n)$. Then the rational functions $\alpha_i/\alpha_0 = \alpha_i$ will be regular at every point of Y. So they are elements of B. The coordinate algebra A of X is generated by the residues of the coordinate variables $x_1, ..., x_n$, with $x_0 = 1$. Sending $x_i \to \alpha_i$ defines a homomorphism $A \xrightarrow{\varphi} B$. Conversely, if φ is such a homomorphism, the good point that defines the morphism $Y \xrightarrow{u} X$ is $(1, \varphi(x_1), ..., \varphi(x_n))$. □

(3.5.5) affine open subsets

An *affine open subset* of a variety X is an open subset that is an affine variety. If V is a nonempty open subset of X and R is the algebra of rational functions that are regular on V, then V is an affine open subset if and only if

- R is a finite-type domain and
- V is isomorphic to $\operatorname{Spec} R$.

In Theorem 3.5.9 below, we prove an important fact, that the intersection of two affine open sets is again an affine open set.

3.5.6. Proposition. *The complement of a hypersurface is an affine open subvariety of $\mathbb{P}^n$.*

proof. Let V be the complement of the hypersurface $\{f = 0\}$, where f is an irreducible homogeneous polynomial of degree d, let K be its fraction field, and let R be the algebra of regular functions on V.

The regular functions on V are the homogeneous fractions of degree zero of the form g/f^k (**3.4.5**), and the fractions m/f, where m is a monomial of degree d, generate R. Since there are finitely many monomials of degree d, R is a finite-type domain. Let w be an arbitrary monomial of degree $d-1$, and let $s_i = x_i w/f$. The point $(x_0, ..., x_n)$ of V can also be written as $(s_0, ..., s_n)$. The fractions s_i are among the generators for R. So if $W = \operatorname{Spec} R$, $(s_0, ..., s_n)$ is a point with values in K that defines a morphism $W \xrightarrow{z} V$. We show that z is an isomorphism.

3.5.7. Lemma. *Let $\mathbb{U}^i$ be the standard affine open subset of $\mathbb{P}^n$. With s_i as above, the intersection $V^i = V \cap \mathbb{U}^i$ is isomorphic to the localization W_{s_i} of W.*

proof. Say that $i = 0$, and let denote $S = s_0 = x_0^d/f$ and $t = s^{-1} = f/x_0^d$. Let A be the coordinate algebra of $\mathbb{U}^0$. Then $V^0 = V \cap \mathbb{U}^0$ is the set of points of $\mathbb{U}^0$ at which t isn't zero. Its coordinate algebra is the localization A_t, and V^0 is the affine variety $\operatorname{Spec} A_t$.

It suffices to show that A_t is the localization R_s of R. With coordinates $u_j = x_j/x_0$ for $\mathbb{U}^0$, a fraction m/f, where $m = x_{j_1} \cdots x_{j_d}$, can be written as $u_{j_1} \cdots u_{j_d}/t$. These fractions generate R, so $R \subset A_t$, and since $s^{-1} = t$ is in A_t, $R_s \subset A_t$. For the other inclusion, we write $u_j = (x_j x_0^{d-1}/f)\, s^{-1}$. Because $x_j x_0^{d-1}/f$ is in R, u_j is in R_s. Therefore $A \subset R_s$ and $A_t \subset R_s$. So $A_t = R_s$, as claimed. □

We go back to the proof of Proposition 3.5.6. The sets $V^i = V \cap \mathbb{U}^i$ for $i = 0, ..., n$ cover V, and the morphism z restricts to an isomorphism $V^i \to \operatorname{Spec} R_{s_i}$. So the morphism z defined above is an isomorphism (3.4.24). □

3.5.8. Lemma. *The affine open subsets of a variety X form a basis for the topology on X.*

proof. See Proposition 2.5.20. □

Let $[A, B]$ denote the algebra generated by two subalgebras A and B of the function field K of X. The elements of this algebra are finite sums of products of elements of A and B. If $A = \mathbb{C}[a]$, $a = a_1, ..., a_r$, and $B = \mathbb{C}[b]$, $b = b_1, ..., b_s$, then $[A, B]$ is the finite-type subalgebra of K generated by the set $\{a, b\}$.

3.5.9. Theorem

Let U and V be affine open subvarieties of a variety X, say $U \approx \operatorname{Spec} A$ and $V \approx \operatorname{Spec} B$. The intersection $U \cap V$ is an affine open subvariety whose coordinate algebra is generated by the two algebras A and B.

proof. With A and B as in the statement of the theorem, let $R = [A, B]$, and let $W = \operatorname{Spec} R$. We are to show that W is isomorphic to $U \cap V$. The inclusions of coordinate algebras $A \to R$ and $B \to R$ give us morphisms $W \to U$ and $W \to V$. We also have inclusions $U \subset X$ and $V \subset X$, and X is a subvariety of a projective space $\mathbb{P}^n$.

Let α be the point of $\mathbb{P}^n$ with values in K that defines the projective embedding $X \xrightarrow{\varphi} \mathbb{P}^n$. The maps from U and V to $\mathbb{P}^n$ defined by α are restrictions of φ. The variety W also has function field K, and α defines a morphism $W \xrightarrow{\psi} \mathbb{P}^n$ whose image is in $U \cap V$. This gives us a morphism $W \xrightarrow{\epsilon} U \cap V$. We show that ϵ is an isomorphism.

Let p be a point of $U \cap V$. We choose an affine open subset Z of $U \cap V$ that is a localization of U and of V, and that contains p (2.5.21)**(ii)**. Let S be the coordinate ring of Z. So $S = A_s$ for some nonzero s in A and also $S = B_t$ for some nonzero t in B. Then

$$R_s = [A, B]_s = [A_s, B] = [B_t, B] = B_t = S$$

So ϵ maps the localization $W_s = \operatorname{Spec} R_s$ of W isomorphically to the open subset Z of $U \cap V$. Since we can cover $U \cap V$ by open sets such as Z, Lemma 3.4.21 **(ii)** shows that ϵ is an isomorphism. □

3.6 Lines in Projective Three-Space

The *Grassmanian* $G(m, n)$ is a variety whose points correspond to subspaces of dimension m of the vector space $\mathbb{C}^n$, and to linear subspaces of dimension $m-1$ of $\mathbb{P}^{n-1}$. One says that $G(m, n)$ *parametrizes* those subspaces. For example, the Grassmanian $G(1, n+1)$ is the projective space $\mathbb{P}^n$. Points of $\mathbb{P}^n$ parametrize one-dimensional subspaces of $\mathbb{C}^{n+1}$.

The Grassmanian $G(2, 4)$ parametrizes two-dimensional subspaces of $\mathbb{C}^4$, and lines in $\mathbb{P}^3$. In this section we describe this Grassmanian, which we denote by $\mathbb{G}$. The point of $\mathbb{G}$ that corresponds to a line ℓ in $\mathbb{P}^3$ will be denoted by $[\ell]$.

One can get some insight into the structure of $\mathbb{G}$ using row reduction. Let $V = \mathbb{C}^4$, let u_1, u_2 be a basis of a two-dimensional subspace U of V and let M be the 2×4 matrix whose rows are u_1, u_2. The rows of the matrix M' obtained from M by row reduction span the same space U, and the row-reduced matrix M' is uniquely determined by U. Provided that the left hand 2×2 submatrix of M is invertible, M' will have the form

(3.6.1)
$$M' = \begin{pmatrix} 1 & 0 & * & * \\ 0 & 1 & * & * \end{pmatrix}$$

So the Grassmanian $\mathbb{G}$ contains, as an open subset, a four-dimensional affine space whose coordinates are the variable entries of M'.

In any 2×4 matrix M with independent rows, some pair of columns will be independent. Those columns can be used in place of the first two in a row reduction. So $\mathbb{G}$ is covered by six four-dimensional affine spaces that we denote by $\mathbb{W}^{ij}$, $1 \le i < j \le 4$, $\mathbb{W}^{ij}$ being the space of 2×4 matrices such that $column_i = (1, 0)^t$ and $column_j = (0, 1)^t$. Since $\mathbb{P}^4$ and the Grassmanian are both covered by affine spaces of dimension four, they may seem similar, but they aren't the same.

(3.6.2) the exterior algebra

Let V be a complex vector space. The *exterior algebra* $\bigwedge V$ (read 'wedge V') is a noncommutative ring that contains the complex numbers and is generated by the elements of V, with the relations

(3.6.3)
$$vw = -wv \quad \text{for all } v, w \text{ in } V.$$

3.6.4. Lemma. *The condition (3.6.3) is equivalent with: $vv = 0$ for all v in V.*

proof. To get $vv = 0$ from (3.6.3), one sets $w = v$. Suppose that $vv = 0$ for all v in V. Then $(v+w)(v+w) = vv = ww = 0$, and since $(v+w)(v+w) = vv + vw + wv + ww$, $\;vw + wv = 0$. $\square$

To familiarize yourself with computation in $\bigwedge V$, verify that $v_2v_3v_1v_4 = v_1v_2v_3v_4$ and that $v_2v_3v_4v_1 = -v_1v_2v_3v_4$.

Let $\bigwedge^r V$ denote the subspace of $\bigwedge V$ spanned by products of length r of elements of V. The exterior algebra $\bigwedge V$ is the direct sum of the subspaces $\bigwedge^r V$. An algebra A that is a direct sum of subspaces A^i, and such that multiplication maps $A^i{\times}A^j$ to A^{i+j} is called a *graded algebra*. Since its multiplication law isn't commutative, the exterior algebra is a noncommutative graded algebra.

3.6.5. Proposition. *If $(v_1, ..., v_n)$ is a basis for V, the products $v_{i_1} \cdots v_{i_r}$ of length r with increasing indices $i_1 < i_2 < \cdots < i_r$ form a basis for $\bigwedge^r V$.*

The proof is at the end of the section.

3.6.6. Corollary. *Let $v_1, ..., v_r$ be elements of V. The product $v_1 \cdots v_r$ in $\bigwedge^r V$ is zero if and only if the set $(v_1, ..., v_r)$ is dependent.* $\square$

For the rest of the section, we let V be a vector space of dimension four with basis $(v_1, ..., v_4)$. Proposition 3.6.5 tells us that

(3.6.7)

$\bigwedge^0 V = \mathbb{C}$ is a space of dimension 1, with basis $\{1\}$
$\bigwedge^1 V = V$ is a space of dimension 4, with basis $\{v_1, v_2, v_3, v_4\}$
$\bigwedge^2 V$ is a space of dimension 6, with basis $\{v_iv_j \mid i < j\} = \{v_1v_2, v_1v_3, v_1v_4, v_2v_3, v_2v_4, v_3v_4\}$
$\bigwedge^3 V$ is a space of dimension 4, with basis $\{v_iv_jv_k \mid i < j < k\} = \{v_1v_2v_3, v_1v_2v_4, v_1v_3v_4, v_2v_3v_4\}$
$\bigwedge^4 V$ is a space of dimension 1, with basis $\{v_1v_2v_3v_4\}$
$\bigwedge^q V = 0$ when $q > 4$

The elements of $\bigwedge^2 V$ are combinations

(3.6.8)
$$w = \sum_{i<j} a_{ij}v_iv_j$$

We regard $\bigwedge^2 V$ as an affine space of dimension 6, identifying the combination w with the vector whose coordinates are the six coefficients a_{ij} $(i < j)$. We use the same symbol w to denote the point of the projective space $\mathbb{P}^5$ with those coordinates: $w = (a_{12}, a_{13}, a_{14}, a_{23}, a_{24}, a_{34})$.

3.6.9. Definition. An element w of $\bigwedge^2 V$ is *decomposable* if it is a product of two elements of V.

3.6.10. Proposition. *The decomposable elements $w = \sum_{i<j} a_{ij}v_iv_j$ of $\bigwedge^2 V$ are those such that $ww = 0$, and the relation $ww = 0$ is given by the following equation in the coefficients a_{ij}:*

(3.6.11)
$$a_{12}a_{34} - a_{13}a_{24} + a_{14}a_{23} = 0$$

proof. If w is decomposable, say $w = u_1u_2$, then $w^2 = u_1u_2u_1u_2 = -u_1^2u_2^2$ is zero because $u_1^2 = 0$. For the converse, we compute w^2 when $w = \sum_{i<j} a_{ij}v_iv_j$. The answer is

$$ww = 2\big(a_{12}a_{34} - a_{13}a_{24} + a_{14}a_{23}\big)v_1v_2v_3v_4$$

To show that w is decomposable if $w^2 = 0$, it seems simplest to factor w explictly. Since the assertion is trivial when $w = 0$, we may suppose that some coefficient of w, say a_{12}, is nonzero. Then if $w^2 = 0$, w is the product

(3.6.12)
$$w = \frac{1}{a_{12}}\big(a_{12}v_2 + a_{13}v_3 + a_{14}v_4\big)\big(-a_{12}v_1 + a_{23}v_3 + a_{24}v_4\big)$$
$\square$

3.6.13. Corollary. (i) *Let w be a nonzero decomposable element of $\bigwedge^2 V$, say $w = u_1u_2$, with u_i in V. Then (u_1, u_2) is a basis for a two-dimensional subspace of V.*

(ii) *If (u_1, u_2) and (u_1', u_2') are bases for the same subspace U of V, then $w = u_1u_2$ and $w' = u_1'u_2'$ differ by a scalar factor. Their coefficients represent the same point of $\mathbb{P}^5$.*

(iii) *Let u_1, u_2 be a basis for a two-dimensional subspace U of V, and let $w = u_1u_2$. The rule $\epsilon(U) = w$ defines a bijection ϵ from $\mathbb{G}$ to the quadric Q in $\mathbb{P}^5$ whose equation is (3.6.11).*

Thus $\mathbb{G}$ can be represented as the quadric (3.6.11).

proof. **(i)** If an element w of $\bigwedge^2 V$ is decomposable, say $w = u_1u_2$, and if w is nonzero, then u_1 and u_2 must be independent (3.6.6). They span a two-dimensional subspace.

(ii) When we write the second basis in terms of the first one, say $(u_1', u_2') = (au_1{+}bu_2, cu_2{+}du_2)$, the product $u_1'u_2'$ becomes $(ad{-}bc)u_1u_2$, and $ad{-}bc \neq 0$.

(iii) In view of **(i)** and **(ii)**, all that remains to show is that, if (u_1, u_2) and (u_1', u_2') are bases for distinct two-dimensional subspaces U and U', then $u_1u_2 \neq u_1'u_2'$ in $\bigwedge^2 V$.

Since $U \neq U'$, the intersection $W = U \cap U'$ has dimension at most 1, o at least three of the vecotrs u_1, u_2, u_1', u_2' will be independent. Therefore $u_1u_2 \neq u_1'u_2'$. □

For the rest of this section, we use the algebraic dimension of a variety, a concept that will be studied in the next chapter. We refer to the algebraic dimension simply as the *dimension*. The *dimension* of a variety X can be defined as the length d of the longest chain $C_0 > C_1 > \cdots > C_d$ of closed subvarieties of X.

As was mentioned in Chapter 1, the topological dimension of X its dimension in the classical topology, is always twice the algebraic dimension. Because the Grassmanian $\mathbb{G}$ is covered by affine spaces of dimension 4, its algebraic dimension is 4 and its topological dimension is 8.

3.6.14. Proposition. *Let $\mathbb{P}^3$ be the projective space associated to a four dimensional vector space V. In the product $\mathbb{P}^3 \times \mathbb{G}$, the locus Γ of pairs $p,[\ell]$ such that the point p of $\mathbb{P}^3$ lies on the line ℓ is a closed subset of dimension* 5.

proof. Let ℓ be the line in $\mathbb{P}^3$ that corresponds to the subspace U with basis (u_1, u_2), and say that p represented by the vector x in V. Let $w = u_1u_2$. Then $p \in \ell$ means $x \in U$, which is true if and only if (x, u_1, u_2) is a dependent set, and this happens if and only if $xw = 0$ (3.6.5). So Γ is the closed subset of points (x, w) of $\mathbb{P}^3 \times \mathbb{P}^5$ defined by the bihomogeneous equations $w^2 = 0$ and $xw = 0$.

When we project Γ to $\mathbb{G}$, The fibre over a point $[\ell]$ of $\mathbb{G}$ is the set of points $p, [\ell]$ such that p is a point of the line ℓ. The fibre over the point $[\ell]$ of $\mathbb{G}$ is the line ℓ. Thus Γ can be viewed as a family of lines, parametrized by the four-dimensional variety $\mathbb{G}$. Its dimension is $\dim \ell + \dim \mathbb{G} = 1 + 4 = 5$. □

(3.6.15) lines on a surface

One may ask whether or not a given surface in $\mathbb{P}^3$ contains a line. One surface that contains lines is the quadric Q in $\mathbb{P}^3$ with equation $w_{01}w_{10} = w_{00}w_{11}$, the image of the Segre embedding $\mathbb{P}^1 \times \mathbb{P}^1 \to \mathbb{P}^3_w$ (**3.1.7**). It contains two families of lines, corresponding to the two "rulings" $p \times \mathbb{P}^1$ and $\mathbb{P}^1 \times q$ of $\mathbb{P}^1 \times \mathbb{P}^1$. There are surfaces of arbitrary degree that contain lines, but, that a generic surface of degree four or more doesn't contain any line.

We use coordinates x_i with $i = 1, 2, 3, 4$ for $\mathbb{P}^3$ here. There are $N = \binom{d+3}{3}$ monomials of degree d in four variables, so homogeneous polynomials of degree d are parametrized by an affine space of dimension N, and surfaces of degree d in $\mathbb{P}^3$ by a projective space of dimension $N-1$. Let $\mathbb{S}$ denote that projective space, and let $[S]$ denote the point of $\mathbb{S}$ that corresponds to a surface S. The coordinates of $[S]$ are the coefficients of the monomials in the defining polynomial f of S. Speaking infomally, we say that a point of $\mathbb{S}$ "is" a surface of degree d in $\mathbb{P}^3$. (When f is reducible, its zero locus isn't a variety. Let's not worry about this.)

Consider the line ℓ_0 defined by $x_3 = x_4 = 0$. Its points are those of the form $(x_1, x_2, 0, 0)$, so a surface $S : \{f = 0\}$ will contain ℓ_0 if and only if $f(x_1, x_2, 0, 0) = 0$ for all x_1, x_2. Substituting $x_3 = x_4 = 0$ into f leaves us with a polynomial in two variables:

(3.6.16) $$f(x_1, x_2, 0, 0) = c_0x_1^d + c_1x_1^{d-1}x_2 + \cdots + c_dx_2^d,$$

where the coefficients c_i are among the coefficients of the polynomial f. If $f(x_1, x_2, 0, 0)$ is identically zero, all of its coefficients must be zero. So the surfaces that contain ℓ_0 correspond to the points of the linear subspace $\mathbb{L}_0$ of $\mathbb{S}$ defined by the equations $c_0 = \cdots = c_d = 0$. Its dimension is $(N-1)-(d+1) = N-d-2$. This is a satisfactory answer to the question of which surfaces contain ℓ_0, and we can use it to make a guess about lines in a generic surface of degree d.

3.6.17. Lemma. *In the product variety $\mathbb{G}\times\mathbb{S}$, the set Γ of pairs $[\ell],[S]$ such that $\ell \subset S$ is a closed subset.*

proof. Let $\mathbb{W}^{ij}$, $1 \le i < j \le 4$ denote the six affine spaces that cover the Grassmanian, as at the beginning of this section. It suffices to show that the intersection $\Gamma^{ij} = \Gamma \cap (\mathbb{W}^{ij} \times \mathbb{S})$ is closed in $\mathbb{W}^{ij} \times \mathbb{S}$ (3.4.20). We inspect the case $i, j = 1, 2$.

A line ℓ such that $[\ell]$ is in $\mathbb{W}^{12}$ corresponds to a subspace of $\mathbb{C}^2$ with basis of the form $u_1 = (1, 0, a_2, a_3)$, $u_2 = (0, 1, b_2, b_3)$ and ℓ is the line $\{ru_1 + su_2\}$. Let $f(x_1, x_2, x_3, x_4)$ be the polynomial that defines a surface S. The line ℓ is contained in S if and only if $f(r, s, ra_2 + sb_2, ra_3 + sb_3)$ is zero for all r and s. This is a homogeneous polynomial of degree d in r, s. Let's call it $\widetilde{f}(r, s)$. If we write $\widetilde{f}(r, s) = z_0r^d + z_1r^{d-1}s + \cdots + z_ds^d$, the coefficients z_ν will be polynomials in a_i, b_i and in the coefficients of f. The locus $z_0 = \cdots = z_d = 0$ is the closed set Γ^{12} of $\mathbb{W}^{12} \times \mathbb{S}$. □

The set of surfaces that contain our special line ℓ_0 corresponds to the linear space $\mathbb{L}_0$ of $\mathbb{S}$ of dimension $N-d-2$, and ℓ_0 can be carried to any other line ℓ by a linear map $\mathbb{P}^3 \to \mathbb{P}^3$. So the sufaces that contain another line ℓ also form a linear subspace of $\mathbb{S}$ of dimension $N-d-2$. They are the fibres of Γ over $\mathbb{G}$. The dimension of the Grassmanian $\mathbb{G}$ is 4. Therefore the dimension of Γ is $\dim \Gamma = \dim \mathbb{L}_0 + \dim \mathbb{G} = (N-d-2) + 4$. Since $\mathbb{S}$ has dimension $N-1$,

$$\dim \Gamma = \dim \mathbb{S} - d + 3. \tag{3.6.18}$$

We project the product $\mathbb{G}\times\mathbb{S}$ and its subvariety Γ to $\mathbb{S}$. The fibre of Γ over a point $[S]$ is the set of pairs $[\ell],[S]$ such that ℓ is contained in S – the set of lines in S.

When the degree d of the surfaces we are studying is 1, $\dim \Gamma = \dim \mathbb{S}+2$. Every fibre of Γ over $\mathbb{S}$ will have dimension at least 2. In fact, every fibre has dimension equal to 2. Surfaces of degree 1 are planes, and the lines in a plane form a two-dimensional family.

When $d = 2$, $\dim \Gamma = \dim \mathbb{S}+1$. We can expect that most fibres of Γ over $\mathbb{S}$ will have dimension 1. This is true: A smooth quadric contains two one-dimensional families of lines. (All smooth quadrics are equivalent with the quadric (3.1.8).) But if a quadratic polynomial $f(x_1, x_2, x_3, x_4)$ is the product of linear polynomials, its locus of zeros will be a union of planes. It will contain two-dimensional families of lines. Some fibres have dimension 2.

When $d \ge 4$, $\dim \Gamma < \dim \mathbb{S}$. The projection $\Gamma \to \mathbb{S}$ cannot be surjective. Most surfaces of degree 4 or more contain no lines.

The most interesting case is that $d = 3$. In this case, $\dim \Gamma = \dim \mathbb{S}$. Most fibres will have dimension zero. They will be finite sets. In fact, a generic cubic surface contains 27 lines. We have to wait to see why the number is precisely 27 (see Theorem 4.8.17).

Our conclusions are intuitively plausible, but to be sure about them, we need to study dimension carefully. We do this in the next chapters.

proof of Proposition 3.6.5. Let $v = (v_1, ..., v_n)$ be a basis of the vector space V. The proposition asserts that the products $v_{i_1} \cdots v_{i_r}$ of length r with increasing indices $i_1 < i_2 < \cdots < i_r$ form a basis for $\bigwedge^r V$.

To prove this, we need to be more precise about the definition of the exterior algebra $\bigwedge V$. We start with the algebra $T(V)$ of noncommutative polynomials in the basis v, which is also called the *tensor algebra* on V. The part $T^r(V)$ of $T(V)$ of degree r has as basis the n^r noncommutative monomials of degree r, products $v_{i_1} \cdots v_{i_r}$ of length r of elements of the basis v. Its dimension is n^r. When $n = r = 2$, $T^2(V)$ the basis is $(x_1^2, x_1x_2, x_2x_1, x_2^2)$.

The exterior algebra $\bigwedge V$ is the quotient of $T(V)$ obtained by forcing the relations $vw+wv = 0$ (3.6.3). Using the distributive law, one sees that the relations $v_iv_j+v_jv_i = 0$, $1 \le i, j \le n$, are sufficient to define this quotient. The relations $v_iv_i = 0$ are included when $i = j$.

To obtain $\bigwedge^r V$, we multiply the relations $v_iv_j+v_jv_i$ on left and right by arbitrary noncommutative monomials p and q in $v_1,...,v_n$ whose degrees add to $r-2$. The noncommutative polynomials

$$p(v_iv_j+v_jv_i)q \tag{3.6.19}$$

span the kernel of the linear map $T^r(V) \to \bigwedge^r V$. So in $\bigwedge^r V$, $p(v_iv_j)q = -p(v_jv_i)q$. Using these relations, any product $v_{i_1}\cdots v_{i_r}$ in $\bigwedge^r V$ is, up to sign, equal to a product in which the elements v_{i_ν} are listed in increasing order. Thus the products with indices in increasing order span $\bigwedge^r V$, and because $v_iv_i = 0$, such a product will be zero unless the indices are strictly increasing.

We go to the proof now. Let $v = (v_1,...,v_n)$ be a basis for V. We show first that the product $w = v_1\cdots v_n$ in increasing order of the basis elements of V is a basis of $\bigwedge^n V$. We have shown that this product spans $\bigwedge^n V$, and it remains to show that $w \neq 0$, or that $\bigwedge^n V \neq 0$.

Let's use multi-index notation: $(i) = (i_1,...,i_r)$, and $v_{(i)} = v_{i_1}\cdots v_{i_r}$. We define a surjective linear map $T^n(V) \xrightarrow{\varphi} \mathbb{C}$ on the basis of $T^n(V)$ of products $v_{(i)} = (v_{i_1}\cdots v_{i_n})$ of length n. If there is no repetition among the indices $i_1,...,i_n$, then (i) will be a permutation of the indices $1,...,n$. In that case, we set $\varphi(v_{(i)}) = \varphi(v_{i_1}\cdots v_{i_n}) = sign(i)$. If there is a repetition, we set $\varphi(v_{(i)}) = 0$.

Let p and q be noncommutative monomials whose degrees add to $n-2$. If the product $p(v_iv_j)q$ has no repeated index, the indices in $p(v_iv_j)q$ and $p(v_jv_i)q$ will be permutations of $1,...,n$, and those permutations will have opposite signs. Then $p(v_iv_j+v_jv_i)q$ will be in the kernel of φ. Since these elements span the space of relations, φ defines a surjective linear map $\bigwedge^n V \to \mathbb{C}$. Therefore $\bigwedge^n V \neq 0$.

To prove (3.6.5), we must show that for $r \leq n$, the products $v_{i_1}\cdots v_{i_r}$ with $i_1 < i_2 < \cdots < i_r$ form a basis for $\bigwedge^r V$, and we know that those products span $\bigwedge^r V$. We must show that they are independent. Suppose that a combination $z = \sum c_{(i)}v_{(i)}$ is zero, the sum being over sets of strictly increasing indices. We choose a set $(j_1,...,j_r)$ of strictly increasing indices, and we let $(k) = (k_1,...,k_{n-r})$ be the set of indices not occuring in (j), listed in arbitrary order. Then all terms in the sum $zv_{(k)} = \sum c_{(i)}v_{(i)}v_{(k)}$ will be zero except the term with $(i) = (j)$. On the other hand, since $z = 0$, $zv_{(k)} = 0$. Therefore $c_{(j)}v_{(j)}v_{(k)} = 0$, and since $v_{(j)}v_{(k)}$ differs by sign from $v_1\cdots v_n$, it isn't zero. It follows that $c_{(j)} = 0$. This is true for all (j), so $z = 0$. $\square$

Chapter 4 STRUCTURE OF VARIETIES I: DIMENSION

4.1 Dimension

Let X be a variety and let K be its function field. The *dimension* of X, which will be denoted by $\dim X$, is the transcendence degree of K over $\mathbb{C}$. The dimension of a finite-type domain A is the transcendence degree of its field of fractions. Thus, if X' is an open subvariety of X, then $\dim X' = \dim X$.

A proper closed subvariety of an affine variety X will have lower dimension than X, but it isn't obvious how much lower its dimension will be. Krull's Theorem is a tool that helps to determine the drop in dimension.

Krull's Principal Ideal Theorem. Let X be an affine variety of dimension n, and let α be a nonzero element of its coordinate algebra. Every irreducible component of the zero locus of α in X has dimension $n-1$.

So a single equation drops dimension by precisely 1. The proof of Krull's Theorem will be given in Section 4.2. We use that Theorem here to derive properties of dimension.

(4.1.1) chains of subvarieties

A *chain* of subvarieties of X *of length* k is a strictly decreasing sequence of closed subvarieties (irreducible closed sets)

(4.1.2) $$C_0 > C_1 > C_2 > \cdots > C_k$$

The chain is *maximal* if it cannot be lengthened by inserting another closed subvariety. This will be true if $C_0 = X$, if there is no closed subvariety $\widetilde{C}$ with $C_i > \widetilde{C} > C_{i+1}$ when $i < k$, and if C_k is a point.

The chain

(4.1.3) $$\mathbb{P}^n > \mathbb{P}^{n-1} > \cdots > \mathbb{P}^0$$

in which $\mathbb{P}^k$ is the subspace of $\mathbb{P}^n$ of points $(x_0, ..., x_k, 0, ..., 0)$, is a maximal chain in $\mathbb{P}^n$. The maximal chains in $\mathbb{P}^2$ have the form $\mathbb{P}^2 > C > p$, where C is a plane curve and p is a point.

4.1.4. Lemma. *Let X' be an open subset of a variety X. There is a bijective correspondence between chains $C_0 > \cdots > C_k$ of closed subvarieties of X such that $C_k \cap X'$ is nonempty and chains $C'_0 > \cdots > C'_k$ of closed subvarieties of X'. Moreover, the chain C_i is maximal if and only if C'_i is maximal.*

Given a chain $C_0 > \cdots > C_k$ in X such that $C_k \cap X' \neq \emptyset$, the corresponding chain in X' is $C'_i = C_i \cap X'$, and given a chain C'_i in X', the corresponding chain in X consists of the closures C_i of the varieties C'_i in X.

proof. Suppose given a chain $C_0 > C_1 > \cdots > C_k$ in X, and that $C_k \cap X'$ isn't empty. Then the intersection $C'_i = C_i \cap X'$ will be nonempty for every i. It will be a dense open subset of the irreducible closed set C_i that is closed in X', and its closure in X will be C_i. Since C_i is irreducible and $C_i > C_{i+1}$, it is also true that C'_i is irreducible and that $C'_i > C'_{i+1}$. Therefore $C'_0 > \cdots > C'_k$ is a chain of closed subsets of X'. Conversely, let $C'_0 > \cdots > C'_k$ be a chain in X', and let C_i be the closure of C'_i in X. Then $C'_i = C_i \cap X$, so $C_i > C_{i+1}$. The closures in X form a chain of closed subsets of X. □

The *codimension* of a closed subvariety Y of a variety X is defined to be the difference $\dim X - \dim Y$. Krull's Theorem tells us that if $X = \operatorname{Spec} A$ and Y is a component of the zero locus of a nonzero element α of A, the codimension of Y will be 1.

4.1.5. Corollary. **(i)** *Every proper closed subvariety of a variety X is contained in a closed subvariety of codimension* 1.

(ii) *A closed subvariety Y of a variety X has codimension* 1 *if and only if $X > Y$, and there is no closed subvariety Y such that $X > C > Y$.* □

4.1.6. Theorem. *Let X be a variety of dimension n. All chains of closed subvarieties of X have length at most n, and all maximal chains have length n.*

proof. Induction allows us to assume that the theorem is true for a variety of dimension less than n. The base case $n = 0$ is that X is a point. That case is trivial.

Let X be a variety of dimension $n > 0$ and let $C_0 > C_1 > \cdots > C_k$ be a chain in X. We are to show that $k \le n$ and that $k = n$ if the chain is maximal. We choose an affine open subset X' of X whose intersection with C_k is nonempty. Lemma 4.1.4 shows that the intersections $C'_i = C_i \cap X'$ form a chain in X'. We may replace X by X', so we may assume that X is affine, say $X = \operatorname{Spec} A$.

We can insert closed subvarieties into our chain when possible. This will increase k. So we may assume that $C_0 = X$. If $k = 0$, there is nothing to show. Otherwise $C_0 > C_1$. Some nonzero element α of A will vanish on C_1. Then C_1 will be contained in a component C' of the zero locus of α. Krull's Theorem tells us that the dimension of C' is equal to $n-1$.

If $C' > C_1$ we insert C' into the chain. So we may assume that $C' = C_1$, i.e., that C_1 is a component of the zero locus of the element α, and that $\dim C_1 = n - 1$. Induction applies to the chain $C_1 > \cdots > C_k$ of closed subvarieties of C_1. The length of that chain is at most $n-1$, and it is equal to $n-1$ if the chain is maximal. Moreover, there is no closed subvariety D such that $X > D > C_1$ (Corollary 4.1.5). Therefore the chain $C_1 > \cdots > C_k$ is maximal if and only if the given chain $C_0 > C_1 > \cdots > C_k$ is maximal. So $k \le n$ and $k = n$ if the given chain is maximal. □

On an affine variety, Theorem 4.1.6 can be stated in terms of chains of prime ideals. A chain $C_1 > \cdots > C_k$ of closed subvarieties of $X = \operatorname{Spec} A$ will correspond to an increasing chain

(4.1.7)
$$P_0 < P_1 < P_2 < \cdots < P_k,$$

of prime ideals of A, a *prime chain*. For instance, in the polynomial algebra $\mathbb{C}[x_1, \ldots, x_n]$, the prime chain

(4.1.8)
$$0 < (x_n) < (x_n, x_{n-1}) < \cdots < (x_n, \ldots, x_1)$$

corresponds to a chain $\mathbb{A}^n > \mathbb{A}^{n-1} > \cdots > \mathbb{A}^0$ in affine space $\mathbb{A}^n$.

A prime ideal P of a noetherian domain A has *codimension* 1 if it is not the zero ideal, and there is no prime ideal $\widetilde{P}$ such that $(0) < \widetilde{P} < P$. Thus $C = V(P)$ is a closed subvariety of codimension 1 in $\operatorname{Spec} A$ if and only if P is a prime ideal of codimension 1 of A.

The prime ideals of codimension 1 in the polynomial algebra $\mathbb{C}[x_1, \ldots, x_n]$ are the principal ideals generated by irreducible polynomials.

4.2 Proof of Krull's Theorem

###ERROR HERE###

We restate the theorem for reference.

4.2.1. Krull's Theorem. *Let X be an affine variety of dimension n, and let α be a nonconstant element of its coordinate algebra A. Every irreducible component of the zero locus of α in X has dimension $n-1$.*

proof. Let W be a component of the zero locus $V_X(\alpha)$, and let Z be the union of the components of $V_X(\alpha)$ that are distinct from W. We eliminate Z by localizing. We choose an element s in A that is identically zero on Z, but not identically zero on W. Then the localization X_s contains points of W, but no point of Z. Since X and X_s have the same function field, their dimensions are equal. Similarly, W and W_s have the same fraction field F, and their dimensions are equal. We replace X by X_s and W by W_s.

We may now suppose that $W = V_X(\alpha)$ is an irreducible closed subset – a closed subvariety. Let B be its coordinate algebra. The inclusion $W \subset X$ gives us a surjective homomorphism $A \to B$, whose kernel P is a prime ideal that contains α. Moreover, $V_X(\alpha) = V_X(P)$. Therefore P is the radical of the principal ideal (α), and (α) contains a power P^k of P (2.5.9).

We extend the element α to a transcendence basis $\alpha_1, ..., \alpha_n$ of K, with α_i in A and $\alpha_n = \alpha$. Let $\overline{\alpha}_i$ denote the residue of α_i in B. Then $\overline{\alpha}_n = 0$. We show that $\overline{\alpha}_1, ..., \overline{\alpha}_{n-1}$ is a transcendence basis of the function field F of W, and therefore that W has dimension $n-1$.

Suppose given a polynomial $f(x_1, ..., x_{n-1})$ such that $f(\overline{\alpha}_1, ..., \overline{\alpha}_{n-1}) = 0$. Then $f(\alpha_1, ..., \alpha_{n-1})$ is in P, so $f(\alpha_1, ..., \alpha_{n-1})^k$ is in (α_n). There is a polynomial $g(x_1, ..., x_n)$ such that $f(\alpha_1, ..., \alpha_{n-1})^k = \alpha_n g(\alpha_1, ..., \alpha_n)$. Because $\alpha_1, ..., \alpha_n$ are algebraically independent, $f(x_1, ..., x_{n-1})^k - x_n g(x_1, ..., x_n)$ is the zero polynomial. Then, since $f(x)^k$ isn't divisible by x_n, f and g must be zero too. This shows that $\overline{\alpha}_1, ..., \overline{\alpha}_{n-1}$ are algebraically independent. The dimension of W is at least $n-1$.

To show that the dimension of W is $n-1$, we show that if $\overline{\beta}$ is any element of B, then $\overline{\alpha}_1, ..., \overline{\alpha}_{n-1}, \overline{\beta}$ is algebraically dependent. We represent $\overline{\beta}$ by an element β of A. Since $\alpha_1, ..., \alpha_n$ is a transcendence basis of K, there is a nontrivial polynomial $g(x_1, ..., x_n, y)$ such that $g(\alpha_1, ..., \alpha_n, \beta) = 0$. We may cancel a power of x_n from $g(x, y)$. So we may assume that $g(x_1, ..., x_{n--1}, 0, y)$ isn't the zero polynomial. But $g(\overline{\alpha}_1, ..., \overline{\alpha}_{n-1}, 0, \overline{\beta}) = 0$. So $\{\overline{\alpha}_1, ..., \overline{\alpha}_{n-1}, \overline{\beta}\}$ isn't algebraically independent. Therefore $\overline{\alpha}_1, ..., \overline{\alpha}_{n-1}$ is a transcendence basis for $\overline{K}$. □

4.3 The Nakayama Lemma

This lemma is a cornerstone of the theory of modules.

4.3.1. Nakayama Lemma. Let M be a finite module over a ring A, and let J be an ideal of A such that $M = JM$. There is an element z in J such that $m = zm$ for all m in M, or such that $(1-z)M = 0$.

It is always true that $M \supset JM$, so the hypothesis $M = JM$ can be replaced by $M \subset JM$.

(4.3.2) eigenvectors

It won't surprise you that eigenvectors are important, but the way that they are used to study modules may be unfamiliar.

Let P be an $n \times n$ matrix with entries in a ring A. The concept of an eigenvector for P makes sense when the entries of a vector are in a module. A column vector $v = (v_1, ..., v_n)^t$ with entries in an A-module M is an *eigenvector* of P with *eigenvalue* λ if $Pv = \lambda v$.

When the entries of a vector are in a module, it becomes hard to adapt the usual requirement that an eigenvector must be nonzero, so we drop it, though the zero eigenvector tells us nothing.

4.3.3. Lemma. *Let $p(t)$ be the characteristic polynomial* $\det(tI - P)$ *of a square matrix P. If v is an eigenvector of P with eigenvalue λ, then $p(\lambda)v = 0$.*

The usual proof, in which one multiplies the equation $(\lambda I - P)v = 0$ by the cofactor matrix of $(\lambda I - P)$, carries over. □

proof of the Nakayama Lemma. By definition, JM denotes the set of (finite) sums $\sum a_i m_i$ with a_i in J and m_i in M. Let $v_1, ..., v_n$ be generators for the finite A-module M, and let v be the vector $(v_1, ..., v_n)^t$. The equation $M = JM$ tells us that there are elements p_{ij} in J such that $v_i = \sum p_{ij} v_j$. In matrix notation, $v = Pv$. So v is an eigenvector of P with eigenvalue 1, and if $p(t)$ is the characteristic polynomial of P, then $p(1)v = 0$. Since the entries of P are in J, inspection of the determinant of $I-P$ shows that $p(1)$ has the form $1-z$, with z in J. Then $(1-z)v_i = 0$ for all i. Since $v_1, ..., v_n$ generate M, $(1-z)M = 0$. □

4.3.4. Corollary. *With notation as in the Nakayama Lemma, let $s = 1-z$, so that $sM = 0$. The localized module M_s is the zero module.*

4.3.5. Corollary. **(i)** *Let I and J be ideals of a noetherian domain A. If $I = JI$, then either I is the zero ideal or J is the unit ideal.*

(ii) *Let $A \subset B$ be rings, and suppose that B is a finite A-module. If J is an ideal of A, and if the extended ideal JB is the unit ideal of B, then J is the unit ideal of A.*

(iii) *Let x be an element of a noetherian domain A, and let J be the ideal xA. The intersection $\bigcap J^n$ is the zero ideal. Therefore, if y is a nonzero element of A, the integers k such that x^k divides y in A are bounded.*

proof. **(i)** Since A is noetherian, I is a finite A-module. If $I = JI$, the Nakayama Lemma tells us that there is an element z of J such that $zx = x$ for all x in I. Suppose that I isn't the zero ideal. We choose a nonzero element x of I. Because A is a domain, we can cancel x from the equation $zx = x$, obtaining $z = 1$. Then 1 is in J, and J is the unit ideal.

(ii) Suppose that $B = JB$. The Nakayama Lemma tells us that there is an element z in J such that $zb = b$ for all b in B. Setting $b = 1$ shows that $z = 1$. So J is the unit ideal.

(iii) Let $I = \bigcap J^n$. The elements of I are those that are divisible by x^n for every n. Let y be an element of I. So for every n, there is an element a_n in A such that $y = a_n x^n$. Then $y/x = a_n x^{n-1}$, which is an element of J^{n-1}. This is true for every n, so y/x is in I, and y is in JI. Since y can be any element of I, $I = JI$. But J isn't the unit ideal, so **(i)** tells us that $I = 0$. □

The proof of the next corollary is left as an exercise

4.3.6. Corollary. *Let $A \subset B$ be noetherian domains and suppose that B is a finite A-module. Then A is a field if and only if B is a field.* □

Since there are many subrings of fields that aren't fields, we see that the hypothesis that one is dealing with a finite module cannot be dropped from the Nakayama Lemma.

4.4 Integral Extensions

Let A be a domain. An *extension* B of A is a ring that contains A as a subring. An element β of an extension B is *integral over* A if it is a root of a monic polynomial

(4.4.1) $$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0,$$

with coefficients a_i in A, and an extension B is an *integral extension* if all of its elements are integral over A.

4.4.2. Lemma. *Let $A \subset B$ be an extension of domains.*

(i) *An element b of B is integral over A if and only if the subring $A[b]$ of B generated by b is a finite A-module.*

(ii) *The set of elements of B that are integral over A is a subring of B.*

(iii) *If B is generated as A-algebra by finitely many integral elements, it is a finite A-module.*

(iv) *Let $R \subset A \subset B$ be rings, and suppose that A is an integral extension of R. An element of B is integral over A if and only if it is integral over R.* □

4.4.3. Corollary. *An extension $A \subset B$ of finite-type domains is an integral extension if and only if B is a finite A-module.* □

4.4.4. Definition. Let $Y \xrightarrow{u} X$ be a morphism of affine varieties $Y = \operatorname{Spec} B$ and $X = \operatorname{Spec} A$, and let $A \xrightarrow{\varphi} B$ be the corresponding homomorphism of finite-type domains. If φ makes B into a finite A-module, we call u a *finite morphism* of affine varieties. If $A \subset B$ and B is an integral extension of A, we call u an *integral morphism* of affine varieties.

Thus an integral morphism is a finite morphism whose associated algebra homomorphism $A \xrightarrow{\varphi} B$ is injective.

4.4.5. Example. If G is a finite group of automorphisms of a finite-type domain B and $A = B^G$ is the algebra of invariants, then B is an integral extension of A. (See Theorem 2.7.5.)

If $A \subset B$ is an integral extension of finite-type domains with fraction fields $K \subset L$, then L will be a finite field extension of K, and the transcendence degrees of K and L will be equal. Therefore $\dim Y = \dim X$.

The next example is helpful for an intuitive understanding of the geometric meaning of integrality.

4.4.6. Example. Let $f(x, y)$ be an irreducible polynomial in $\mathbb{C}[x, y]$, let $B = \mathbb{C}[x, y]/(f)$ (one x and one y), and let $A = \mathbb{C}[x]$. So $X = \operatorname{Spec} A$ is the affine line $\mathbb{A}^1_x$, and $Y = \operatorname{Spec} B$ is an affine plane curve. The canonical map $A \to B$ defines a morphism $Y \xrightarrow{u} X$, which can be described as a restriction of the projection $\mathbb{A}^2_{x,y} \to \mathbb{A}^1_x$.

We write f as a polynomial in y, whose coefficients are polynomials in x:

$$f(x, y) = a_0(x)y^n + a_1(x)y^{n-1} + \cdots + a_n(x) \tag{4.4.7}$$

The fibre of Y over a point $x = x_0$ of X is the set of points (x_0, y_0) such that y_0 is a root of the one-variable polynomial $f(x_0, y) = \widetilde{f}(y)$. Because f is irreducible, its discriminant with respect to the variable y isn't identically zero (1.10.1). For all but finitely many values x_0, $\widetilde{f}(y)$ its discriminant will be nonzero and it will have degree n. Then $\widetilde{f}(y)$ will have n distinct roots.

When $f(x, y)$ is a monic polynomial in y, u will be an integral morphism. If so, the leading term y^n of f will be the dominant term, when y is large. For x near to any point x_0, there will be a positive real number B such that

$$|y^n| > |a_1 y^{n-1} + \cdots + a_n|$$

when $|y| > B$. Then $\widetilde{f}(y) \neq 0$ when $|y| > B$. Therefore the roots y of $f(x, y)$ are bounded for all x near to any point x_0.

On the other hand, when the leading coefficient $a_0(x)$ isn't a constant, B won't be integral over A, and when x_0 is a root of a_0, $f(x_0, y)$ will have degree less than n. In this case, as a point x_1 of X approaches x_0, at least one root of $f(x_1, y)$ tends to infinity. In calculus, one says that the locus $f(x, y) = 0$ has a vertical asymptote at x_0.

To see this, we divide f by its leading coefficient. Let $g(x, y) = f(x, y)/a_0 = y^n + c_1 y^{n-1} + \cdots + c_n$ with $c_i(x) = a_i(x)/a_0(x)$. For any x at which $a_0(x)$ isn't zero, the roots of g are the same as those of f. However, let x_0 be a root of a_0. Because f is irreducible, there is at least one coefficient $a_j(x)$ that isn't divisible by $x - x_0$. Then $c_j(x)$ is unbounded near x_0, and because the coefficient c_j is an elementary symmetric function in the roots, the roots can't all be bounded.

This is the general picture: The roots of a polynomial remain bounded where the leading coefficient isn't zero. If the leading coefficient vanishes at a point, some roots are unbounded near that point. □

$figure$: nonmonic polynomial, but compare with figure for Hensel's Lemma

The next theorem is a useful tool.

4.4.8. Noether Normalization Theorem. Let A be a finite-type algebra over an infinite field k. There exist elements $y_1, \ldots, y_n$ in A that are algebraically independent over k, such that A is a finite module over its polynomial subalgebra $k[y_1, \ldots, y_n]$.

The Noether Normalization Theorem is also true when k is a finite field, though the proof given below needs to be modified. When $K = \mathbb{C}$, the theorem can be stated by saying that every affine variety X admits an integral morphism to an affine space.

4.4.9. Lemma. *Let k be an infinite field, and let $f(x)$ be a nonzero polynomial of degree d in $x_1, \ldots, x_n$, with coefficients in k. After a suitable linear change of variable, the coefficient of x_n^d in f will be nonzero.*

proof. Let f_d be the homogeneous part of f of maximal degree d. We regard f_d as a polynomial function. Since k is infinite, this function isn't identically zero. We choose coordinates $x_1, ..., x_n$ so that the point $q = (0, ..., 0, 1)$ isn't a zero of f_d. Then $f_d(0, ..., 0, x_n) = cx_n^d$, and the coefficient c, which is $f_d(0, ..., 0, 1)$, will be nonzero. By scaling x_n, we can make $c = 1$. □

proof of the Noether Normalization Theorem. Say that the finite-type algebra A is generated by elements $x_1, \ldots, x_n$. If those elements are algebraically independent over k, A will be isomorphic to the polynomial algebra $\mathbb{C}[x]$, and we will be done. If not, they will satisfy a polynomial relation $f(x) = 0$ of some degree d, with coefficients in k. The lemma tells us that, after a suitable change of variable, the coefficient of x_n^d in f will be 1. Then f will be a monic polynomial in x_n with coefficients in the subalgebra R generated by $x_1, \ldots, x_{n-1}$. So x_n will be integral over R, and A will be a finite R-module. By induction on n, we may assume that R is a finite module over a polynomial subalgebra P. Then A will be a finite module over P too. □

The next proposition is an example of a general principle: Any construction involving finitely many operations can be done in a simple localization (see **5.1.17**).

4.4.10. Proposition. *Let $A \subset B$ be finite-type domains. There is a nonzero element s in A such that B_s is a finite module over a polynomial subring $A_s[y_1, ..., y_r]$.*

proof. Let S be the set of nonzero elements of A, so that $K = AS^{-1}$ is the fraction field of A, and let $B_K = BS^{-1}$ be the ring obtained from B by inverting all elements of S. Also, let $\beta = (\beta_1, ..., \beta_k)$ be a set of elements of the finite-type algebra B that generates B as algebra. Then B_K is a finite-type K-algebra, generated as K-algebra by β. (A K-algebra is a ring that contains K as subring. The Noether Normalization Theorem tells us that B_K is a finite module over a polynomial subring $P = K[y_1, ..., y_r]$. So B_K is an integral extension of P. Any element of B will be in B_K, and therefore it will be the root of a monic polynomial, say

$$f(x) = x^n + c_{n-1}(y)x^{n-1} + \cdots + c_0(y) = 0$$

where the coefficients $c_j(y)$ are elements of P. Each coefficient is a combination of finitely many monomials in y, with coefficients in K. If $d \in A$ is a common denominator for all of those coefficients, then $c_j(x)$ will have coefficients in $A_d[y]$. Since the generators β of B are integral over P, we may choose a denominator s so that all of the generators $\beta_1, ..., \beta_k$ are integral over $A_s[y]$. The algebra B_s is generated over A_s by β, so it will be an integral extension of $A_s[y]$. □

4.5 Normalization

Let A be a domain with fraction field K, and let L be a finite field extension of K. The *integral closure of A in L* is the set of elements of L that are integral over A. It follows from Lemma 4.4.2 **(ii)** that the integral closure is a domain that contains A.

The *normalization* $A^\#$ of A is the integral closure of A in its fraction field K – the set of elements of K that are integral over A. A domain A is *normal* if it is equal to its normalization.

A variety X is a *normal variety* if it has an affine covering $\{X^i = \operatorname{Spec} A_i\}$ in which the algebras A_i are normal domains. To justify this definition, we need to show that if an affine variety $X = \operatorname{Spec} A$ has an affine covering $X^i = \operatorname{Spec} A_i$, in which A_i are normal domains, then A is normal. This follows from Lemma 4.5.3 **(iii)** below.

Our goal here is the next theorem, whose proof is at the end of the section.

4.5.1. Theorem. *Let A be a finite-type domain with fraction field K of characteristic zero, and let L be a finite field extension of K. The integral closure of A in L is a finite A-module, and therefore a finite-type domain. In particular, the normalization of A is a finite A-module and a finite-type domain.*

Thus, if B is the integral closure of A in L, there will be an integral morphism $\operatorname{Spec} B \to \operatorname{Spec} A$.

The proof given here makes use of the characteristic zero hypothesis, though the theorem is true for a finite-type k-algebra when k is a field of characteristic p.

4.5.2. Example. *(normalization of a nodal cubic curve)* The algebra $A = \mathbb{C}[u, v]/(v^2 - u^3 - u^2)$ can be embedded into the one-variable polynomial algebra $B = \mathbb{C}[x]$, by $u = x^2 - 1$ and $v = x^3 - x$. The fraction fields of A and B are equal because $x = v/u$, and the equation $x^2 - (u+1) = 0$ shows that x is integral over A. Since B is normal, it is the normalization of A (see Lemma 4.5.3 **(i)**).

In this example, $\operatorname{Spec} B$ is the affine line $\mathbb{A}^1_x$, and the plane curve $C = \operatorname{Spec} A$ has a node at the origin $p = (0,0)$. The inclusion $A \subset B$ defines an integral morphism $\mathbb{A}^1_x \to C$ whose fibre over p is the point pair $x = \pm 1$. The morphism is bijective at all other points. I think of C as the variety obtained by gluing the points $x = \pm 1$ of the affine line together.

figure: curve, not quite glued

In this example, the effect of normalization can be visualized geometrically. This isn't always so. Normalization is an algebraic process whose effect on geometry may be subtle. □

4.5.3. Lemma. **(i)** *A unique factorization domain is normal. In particular, a polynomial algebra over a field is normal.*

(ii) *If s is a nonzero element of a normal domain A. The localization A_s is normal.*

(iii) *Let $s_1, ..., s_k$ be nonzero elements of a domain A that generate the unit ideal. If the localizations A_{s_i} are normal for all i, then A is normal.*

proof. **(i)** Let A be a unique factorization domain, and let β be an element of its fraction field that is integral over A. Say that

$$\beta^n + a_1\beta^{n-1} + \cdots + a_{n-1}\beta + a_n = 0 \tag{4.5.4}$$

with a_i in A. We write $\beta = r/s$, where r and s are relatively prime elements of A. Multiplying by s^n gives us the equation

$$r^n = -s\,(a_1 r^{n-1} + \cdots + a_n s^{n-1})$$

This equation shows that if a prime element of A divides s, it also divides r. Since r and s are relatively prime, there is no such element. So s is a unit, and β is in A.

(ii) Let β be an element of the fraction field of A that is integral over A_s. There will be a polynomial relation of the form (4.5.4), except that the coefficients a_i will be elements of A_s. The element $\gamma = s^k\beta$ satisfies the polynomial equation

$$\gamma^n + (s^k a_1)\gamma^{n-1} + \cdots + (s^{(n-1)k} a_{n-1})\gamma + (s^{nk} a_n) = 0$$

Since a_i are in A_s, all coefficients of this polynomial will be in A when k is sufficiently large, and then γ will be integral over A. Since A is normal, γ will be in A, and $\beta = s^{-k}\gamma$ will be in A_s.

(iii) This proof follows a familiar pattern. Suppose that A_{s_i} is normal for every i. If an element β of K is integral over A, it will be in A_{s_i} for all i, and $s_i^n\beta$ will be an element of A if n is large. We can use the same exponent n for all i. Since $s_1, ..., s_k$ generate the unit ideal, so do their powers $s_i^n, ..., s_k^n$. Say that $\sum r_i s_i^n = 1$, with r_i in A. Then $\beta = \sum r_i s_i^n \beta$ is in A. □

4.5.5. Lemma. *Let A be a normal noetherian domain with fraction field K of characteristic zero, and let L be a field extension of K. An element β of L that is algebraic over K is integral over A if and only if the coefficients of the monic irreducible polynomial f for β over K are in A.*

proof. If the monic polynomial f has coefficients in A, then β is integral over A. Suppose that β is integral over A. Since we may replace L by any field extension that contains β, we may assume that L is a finite extension of K. A finite extension embeds into a Galois extension, so we may assume that L/K is a Galois extension. Let G be its Galois group, and let $\{\beta_1, \dots, \beta_r\}$ be the G-orbit of β, with $\beta = \beta_1$. The irreducible polynomial for β over K is

$$f(x) = (x - \beta_1)\cdots(x - \beta_r) \tag{4.5.6}$$

Its coefficients are symmetric functions of the roots. If β is integral over A, then all elements of the orbit are integral over A, and therefore the symmetric functions are integral over A. Since A is normal, they are in A. So f has coefficients in A. □

4.5.7. Example. A polynomial in $A = \mathbb{C}[x, y]$ is *square-free* if it has no nonconstant square factors and isn't a constant. Let $f(x, y)$ be a square-free polynomial, and let B denote the integral extension $\mathbb{C}[x, y, w]/(w^2 - f)$ of A. Let K and L be the fraction fields of A and B, respectively. Then $L = K[w]/(w^2 - f)$ is a Galois extension of K. Its Galois group is generated by the automorphism σ of order 2 defined by $\sigma(w) = -w$. The elements of L have the form $\beta = a + bw$ with $a, b \in K$, and $\sigma(\beta) = \beta' = a - bw$.

We show that B is the integral closure of A in L. Suppose that $\beta = a+bw$ is integral over A. If $b = 0$, then $\beta = a$. This is an element of A and therefore it is in B. If $b \neq 0$, the irreducible polynomial for $\beta = a + bw$ will be

$$(x - \beta)(x - \beta') = x^2 - 2ax + (a^2 - b^2 f)$$

Because β is integral over A, $2a$ and $a^2 - b^2 f$ are in A. Because the characteristic isn't 2, this is true if and only if a and $b^2 f$ are in A. We write $b = u/v$, with u, v relatively prime elements of A, so $b^2 f = u^2 f / v^2$. If v weren't a unit, then since f is square-free, it couldn't cancel v^2. So from $b^2 f$ in A we can conclude that b is in A. Summing up, β is integral if and only if a and b are in A, which means that β is in B. □

(4.5.8) trace

Let L be a finite field extension of a field K and let β be an element of K. When L is viewed as a K-vector space, multiplication by β becomes a linear operator $L \xrightarrow{\beta} L$. The *trace* of this operator will be denoted by $\mathrm{tr}(\beta)$. The trace is a K-linear map $L \to K$.

4.5.9. Lemma. *Let L be a finite field extension of K, let $f(x) = x^r + a_1 x^{r-1} + \cdots + a_r$ be the irreducible polynomial for an element β of L over K, and let $K(\beta)$ be the extension of K generated by β. Say that $[L\,{:}\,K(\beta)] = d$ and $[L\,{:}\,K] = n$. Since $[K(\beta) : K] = r$, $n = rd$, and $\mathrm{tr}(\beta) = -da_1$. If β is an element of K, then $\mathrm{tr}(\beta) = n\beta$.*

proof. The set $(1, \beta, \ldots, \beta^{r-1})$ is a K-basis for $K(\beta)$, and on this basis, the matrix M of multiplication by β has the form illustrated below for the case $r = 3$.

$$M = \begin{pmatrix} 0 & 0 & -a_3 \\ 1 & 0 & -a_2 \\ 0 & 1 & -a_1 \end{pmatrix}.$$

For all r, the trace of the matrix M is $-a_1$. Next, let $(u_1, \ldots, u_d)$ be a basis for L over $K(\beta)$. Then $\{\beta^i u_j\}$, with $i = 0, \ldots, r-1$ and $j = 1, \ldots, d$, will be a basis for L over K. When this basis is listed in the order

$$(u_1, u_1\beta, ..., u_1\beta^{n-1}; u_2, u_2\beta, \ldots u_2\beta^{n-1}; \ \ldots \ ; u_d, u_d\beta, ..., u_d\beta^{n-1}),$$

the matrix of multiplication by β will be made up of d blocks of the matrix M. □

4.5.10. Corollary. *Let A be a normal domain with fraction field K and let L be a finite field extension of K. If an element β is integral over A, its trace is an element of A.*

This follows from Lemmas 4.5.5 and 4.5.9. □

4.5.11. Lemma. *Let A be a normal noetherian domain with fraction field K* of characteristic zero*, and let L be a finite field extension of K. The form $L \times L \to K$ defined by $\langle \alpha, \beta \rangle = \mathrm{tr}(\alpha\beta)$ is K-bilinear, symmetric, and nondegenerate. If α and β are integral over A, then $\langle \alpha, \beta \rangle$ is an element of A.*

proof. The form is obviously symmetric, and it is K-bilinear because multiplication is K-bilinear and trace is K-linear. A form is nondegenerate if its nullspace is zero, which means that when α is a nonzero element, there is an element β such that $\langle \alpha, \beta \rangle \neq 0$. We let $\beta = \alpha^{-1}$. Then $\langle \alpha, \beta \rangle = \mathrm{tr}(1)$, which, according to (4.5.9), is the degree $[L\,{:}\,K]$ of the field extension. It is here that the hypothesis on the characteristic of K enters: The degree is a nonzero element of K.

If α and β are integral over A, so is their product $\alpha\beta$ (Lemma 4.4.2 **(ii)**). Corollary 4.5.10 shows that $\langle \alpha, \beta \rangle$ is an element of A. □

4.5.12. Lemma. *Let A be a domain with fraction field K, let L be a field extension of K, and let β be an element of L that is algebraic over K. If β is a root of a polynomial $f = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ with a_i in A, then $\gamma = a_n\beta$ is integral over A.*

proof. One finds a monic polynomial with root γ by substituting $x = y/a_n$ into f and multiplying by a_n^{n-1}. □

proof of Theorem 4.5.1. Let A be a finite-type domain with fraction field K of characteristic zero, and let L be a finite field extension of K. We are to show that the integral closure of A in L is a finite A-module.

Step 1. We may assume that A is normal.

We use the Noether Normalization Theorem to write A as a finite module over a polynomial subalgebra $R = \mathbb{C}[y_1, \ldots, y_d]$. Let F be the fraction field of R. Then K and L are finite extensions of F. An element of L will be integral over A if and only if it is integral over R ((4.4.2) **(iv)**). So the integral closure of A in L is the same as the integral closure of R in L. We replace A by the normal algebra R and K by F.

Step 2. Bounding the integral extension.

We assume that A is normal. Let $(v_1, \ldots, v_n)$ be a K-basis for L whose elements are integral over A. Such a basis exists because we can multiply any element of L by a nonzero element of K to make it integral (Lemma 4.5.12). Let

(4.5.13) $$T : L \to K^n$$

be the map $T(\beta) = \big(\langle v_1, \beta\rangle, \ldots, \langle v_n, \beta\rangle\big)$, where $\langle\ ,\ \rangle$ is the form defined in Lemma 4.5.11. This map is K-linear. If $\langle v_i, \beta\rangle = 0$ for all i, then because $(v_1, \ldots, v_n)$ is a basis for L, $\langle \gamma, \beta\rangle = 0$ for all γ in L, and since the form is nondegenerate, $\beta = 0$. Therefore T is injective.

Let B be the integral closure of A in L. The basis elements v_i are in B, and if β is in B, $v_i\beta$ will be in B too. Then $\langle v_i, \beta\rangle$ will be in A, and $T(\beta)$ will be in A^n (4.5.11). When we restrict T to B, we obtain an injective map $B \to A^n$ that we denote by T_0. Since T is K-linear, T_0 is a A-linear. It is an injective homomorphism of A-modules. It maps B isomorphically to its image, a submodule of A^n. Since A is noetherian, every submodule of the finite A-module A^n is finitely generated. Therefore the image of T_0 is a finite A-module, and so is the isomorphic A-module B. □

4.6 Geometry of Integral Morphisms

The main geometric properties of an integral morphism of affine varieties are summarized below, in Theorem 4.6.6 below, which shows that the geometry is as nice as could be expected.

We use the following notation:

(4.6.1) $$Y \xrightarrow{u} X$$

will be an integral morphism of the affine varieties $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} B$. So B is an integral extension of A. Also, C will be a closed subvariety of X, D will be a component of its inverse image $u^{-1}C$, and the prime ideals of A and B corresponding to C and D, respectively, will be P and Q, and $P \subset Q$. Let $\overline{A} = A/P$, $\overline{B} = B/Q$, $C = \operatorname{Spec} \overline{A}$, and $D = \operatorname{Spec} \overline{B}$.

The homomorphisms and the corresponding morphisms form the diagrams

$$\begin{array}{ccccccc} B & \longrightarrow & \overline{B} & \qquad & Y & \longleftarrow & D \\ \uparrow & & \uparrow & & \downarrow & & \downarrow \\ A & \longrightarrow & \overline{A} & & X & \longleftarrow & C \end{array}$$

The images in $\overline{B}$ of a set of A-module generators for B will generate $\overline{B}$ as $\overline{A}$-module. So $\overline{B}$ is a finite $\overline{A}$-module, and the induced morphism $D \to C$ is a finite morphism of affine varieties. Therefore $\dim D \le \dim C$. In many cases, though not always, $\dim D = \dim C$, and if so, the map $D \to C$ will be an integral morphism.

We say that a closed subvariety D of Y *lies over* the closed subvariety C of X if the image of D is C, and we say that a prime ideal Q of B *lies over* the prime ideal P of A if the contraction $Q \cap A$ of Q is P.

For instance, a point q of Y lies over its image p, and the maximal ideal $\mathfrak{m}_q$ of B at q lies over the maximal ideal $\mathfrak{m}_p$ of A at p.

4.6.2. Lemma. *An integral morphism $Y \xrightarrow{u} X$ of affine varieties is surjective.*

proof. Let $\mathfrak{m}_p$ be the maximal ideal at point p of X. Corollary 4.3.5 **(ii)** shows that the extended ideal $\mathfrak{m}_p B$ isn't the unit ideal of B. So it is contained in a maximal ideal of B, say in $\mathfrak{m}_q$, where q is a point of Y. Then $\mathfrak{m}_q \cap A \supset \mathfrak{m}_p B \cap A \supset \mathfrak{m}_p$. Since $\mathfrak{m}_q \cap A$ doesn't contain 1, it isn't the unit idea, and since $\mathfrak{m}_p$ is a maximal ideal, the inclusions are equalities. So q lies over p. □

The next lemma will allow us to replace X and Y by C and D in some situations.

4.6.3. Lemma. *With notation as at the beginning of the section,*
(i) *If Q lies over P then $\overline{B}$ is an integral extension of $\overline{A}$.*
(ii) *Q lies over P if and only if D lies over C.* □

proof. **(i)** The kernel of the composed map $A \to B \to \overline{B}$ is $Q \cap A$. If Q lies over P, i.e., if $Q \cap A = P$, the mapping property of quotient rings gives us an injective map $\overline{A} \to \overline{B}$. Since $\overline{B}$ is a finite $\overline{A}$-module, it is an integral extension of $\overline{A}$.

(ii) Suppose that Q lies over P. Then **(i)** shows that $\overline{B}$ is an integral extension of $\overline{A}$, and Lemma 4.6.2 shows that the integral morphism $D \to C$ is surjective. Thus D lies over C. Conversely, if D lies over C, an element α of A will be zero at every point of C if and only if it is zero at every point of D. This means that an element α of A is in P if and only if it is in Q. So $P = Q \cap A$, i.e., Q lies over P. □

4.6.4. Lemma. *Let $A \subset B$ be an integral extension of finite-type domains, and let J be a nonzero ideal of B. The contraction $J \cap A$ is a nonzero ideal of A.*

proof. A nonzero element β of J will be integral over A. There will be a polynomial relation of the form $\beta^n + a_1\beta^{n-1} + \cdots + a_n = 0$, with a_i in A. If $a_n = 0$, then because B is a domain, we can cancel β from the equation. So there is a polynomial relation with constant term $a_n \neq 0$. The equation shows that a_n is in J, and since it is also in A, it is a nonzero element of $J \cap A$. □

4.6.5. Lemma. *Let $Y \xrightarrow{u} X$ be an integral morphism of affine varieties, with $Y = \operatorname{Spec} B$ and $X = \operatorname{Spec} A$. The fibres of u have bounded cardinality. If B is generated as A-module by a set of n elements, there are at most n points in the fibre over a point p of X.*

proof. Let $q_1, ..., q_r$ be the points of Y that lie over p, and let k_i denote the residue field of B at q_i. The maximal ideals $\mathfrak{m}_i$ of B at q_i contain the maximal $\mathfrak{m}_p$ of A at p, and they are comaximal. By the Chinese Remainder Theorem, the map $B \to k_1 \times \cdots \times k_r$ is surjective. The images of the n generators β for B generate the product $k_1 \times \cdots \times k_r$, which is a vector space of dimension r over the residue field k of A at p. Therefore $n \geq r$. □

4.6.6. Theorem. *Let $Y \xrightarrow{u} X$ be an integral morphism of affine varieties, $Y = \operatorname{Spec} B$ and $X = \operatorname{Spec} A$.*
(i) *The image of a closed subset of Y is a closed subset of X.*
(ii) *Let D' and D be closed subvarieties of Y that lie over the same closed subvariety C of X. If $D' \supset D$, then $D' = D$.*
(iii) *The set of closed subvarieties D of Y that lie over a closed subvariety C of X is finite and nonempty.*

proof. **(i)** *(The image of a closed set is closed.)*

It suffices to show that the image of a closed subvariety D of Y is closed. Let Q be the prime ideal of B that corresponds to D, and let P be the contraction $Q \cap A$. Lemma 4.6.3 **(ii)** shows that D lies over $C = V_X(P)$. So the image of D is the closed set C.

(ii) *(inclusions among subvarieties that lie over C).*

We rename D' as Y and C as X. With this notation, what must be shown is that the image of a proper closed subset of Y is a proper closed subset of X. Or, the contraction $P = Q \cap A$ of a nonzero prime ideal Q of B is a nonzero prime ideal of A. The contraction of P is a prime ideal, and Lemma 4.6.4 shows that $P \neq 0$.

(iii) *(subvarieties that lie over a closed subvariety)*

Let C be a closed subvariety of X. Its inverse image $Z = u^{-1}C$ is a closed subset of Y, the union of finitely many irreducible closed subsets, say $Z = \bigcup D_i$. As **(i)** shows, the image C_i of D_i will be an irreducible closed subset of X. Since the map u is surjective, $C = \bigcup C_i$, and since C is irreducible, it is equal to at least one C_i. The components D_i of Z such that $C_i = C$ are subvarieties that lie over C.

Next, any subvariety D' that lies over C will be contained in Z, and since it is irreducible, D' will be contained in D_i for some i. Part **(ii)** shows that $D' = D_i$. Therefore the closed subsets that lie over C are among the finitely many closed sets D_i. □

4.7 Chevalley's Finiteness Theorem

(4.7.1) finite morphisms

The concepts of a finite morphism and an integral morphism of affine varieties were defined in Section 4.4. A morphism $Y \xrightarrow{u} X$ of affine varieties $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} B$ is a finite morphism if the homomorphism $A \xrightarrow{\varphi} B$ that corresponds to u makes B into a finite A-module. As was explained, the difference between a finite morphism and an integral morphism of affine varieties is that for a finite morphism, the homomorphism φ needn't be injective. If φ is injective, B will be an integral extension of A, and u will be an integral morphism. We extend the definitions to varieties that aren't necessarily affine here.

By the *restriction* of a morphism $Y \xrightarrow{u} X$ to an open subset X' of X, we mean the induced morphism $Y' \to X'$, where Y' is the inverse image of X'.

4.7.2. Definition. A morphism of varieties $Y \xrightarrow{u} X$ is a *finite morphism* if X can be covered by affine open subsets to which the restriction of u is a finite morphism of affine varieties, as defined in (4.4.4). A morphism u is an *integral morphism* if there is a covering of X by affine open sets to which the restriction of u is an integral morphism of affine varieties.

4.7.3. Corollary. *An integral morphism is a finite morphism. The composition of finite morphisms is a finite morphism. The inclusion of a closed subvariety into a variety is a finite morphism.* □

When X is affine, Definition 4.4.4 and Definition 4.7.2 both apply. The next proposition shows that the two definitions are equivalent.

4.7.4. Proposition. *Let $Y \xrightarrow{u} X$ be a finite or an integral morphism, as in (4.7.2), and let X' be an affine open subset of X. The restriction of u to X' is a finite or an integral morphism of affine varieties, as defined in (4.4.4).*

4.7.5. Lemma. **(i)** *Let $A \xrightarrow{\varphi} B$ be a homomorphism of finite-type domains that makes B into a finite A-module, and let s be a nonzero element of A. Then B_s is a finite A_s-module.*
(ii) *The restriction of a finite (or an integral) morphism $Y \xrightarrow{u} X$ to an open subset of X is a finite (or an integral) morphism, as in Definition 4.7.2.*

proof. **(i)** In the statement, B_s denotes the localization of B as A-module. This localization can also be obtained by localizing the algebra B with respect to the image $s' = \varphi(s)$, provided that it isn't zero. If s' is zero, then s annihilates B, so $B_s = 0$. In either case, a set of elements that spans B as A-module will span B_s as A_s-module, so B_s is a finite A_s-module.

(ii) Say that X is covered by affine open sets to which the restriction of u is a finite morphism. The localizations of these open sets form a basis for the Zariski topology on X. So X' can be covered by such localizations. Part **(i)** shows that the restriction of u to X' is a finite morphism. □

proof of Proposition 4.7.4. We'll do the case of a finite morphism. The proof isn't difficult, but there are several things to check. This makes the proof a bit longer than one would like.

Step 1. Preliminaries.

We are given a morphism $Y \xrightarrow{u} X$, X is covered by affine open sets X^i, and the restrictions of u to these open sets are finite morphisms of affine varieties. We are to show that the restriction to any affine open set X' is a finite morphism of affine varieties.

The affine open set X' is covered by the affine open sets $X'^i = X' \cap X^i$, and the restrictions fo X'^i are finite morphisms ((4.7.5) **(ii)**). So we may replace X by X'. Since the localizations of an affine variety form a basis for its Zariski topology, we see that what is to be proved is this:

A morphism $Y \xrightarrow{u} X$ is given in which $X = \operatorname{Spec} A$ is affine, and there are elements $s_1, ..., s_k$ that generate the unit ideal of A, such that for every i, the inverse image Y^i of $X^i = X_{s_i}$ if nonempty, is affine, and its coordinate algebra B_i is a finite module over the localized algebra $A_i = A_{s_i}$. We must show that Y is affine, and that its coordinate algebra B is a finite A-module.

Step 2. The algebra B of regular functions on Y.

If Y is affine, the algebra B of regular functions on Y will be a finite-type domain and Y will be its spectrum. Since Y isn't assumed to be affine, we don't know very much about B other than that it is a subalgebra of the function field L of Y. On the other hand, the inverse image Y^i of X^i, if nonempty, is affine. It is the spectrum of a finite-type domain B_i. Since the localizations X^i cover X, the affine varieties the Y^i cover Y. We throw out the indices i such that Y^i is empty. Then a function is regular on Y if and only if it is regular on each Y^i, and

$$B = \bigcap B_i$$

the intersection being in the function field L.

Let's denote the image of s_i in B by the same symbol.

Step 3. For any index j, B_j is the localization $B[s_j^{-1}]$ of B.

The intersection $Y^j \cap Y^i$ is an affine variety. It is the set of points of Y^j at which s_i isn't zero, and its coordinate algebra is the localization $B_j[s_i^{-1}]$. Then

$$B[s_j^{-1}] \overset{(1)}{=} \bigcap \left(B_i[s_j^{-1}]\right) \overset{(2)}{=} \bigcap B_j[s_i^{-1}] \overset{(3)}{=} B_j[s_j^{-1}] \overset{(3)}{=} B_j$$

where the explanation of the numbered equalities is as follows:

(1) A rational function β is in $B_i[s_j^{-1}]$ if $s_j^n\beta$ is in B_i for large n, and we can use the same exponent n for all $i = 1, ..., r$. So β is in $\bigcap \left(B_i[s_j^{-1}]\right)$ if and only if $s_j^n\beta$ is in $\bigcap B_i = B$, i.e., if and only if β is in $B[s_j^{-1}]$.

(2) This is true because $Y^j \cap Y^i = Y^i \cap Y^j$.

(3) For all i, $B_j \subset B_j[s_i^{-1}]$. Moreover, s_j doesn't vanish on Y^j. It is a unit in B_j, and therefore $B_j[s_j^{-1}] = B_j$: $B_j \subset \bigcap B_j[s_i^{-1}] \subset B_j[s_j^{-1}] \subset B_j$.

Step 4. B is a finite A-module.

We choose a finite set $b = (b_1, ..., b_n)$ of elements of B that generates the A_i-module B_i for every i. We can do this because we can span the finite A_i-module B_i by finitely many elements of B, and there are finitely many algebras B_i. We show that the set b generates the A-module B.

Let x be an element of B. Since x is in B_i, it is a combination of the elements b with coefficients in A_i. Then for large k, $s_i^k x$ will be a combination of b with coefficients in A, say

$$s_i^k x = \sum_\nu a_{i,\nu} b_\nu$$

with $a_{i,\nu} \in A$. We can use the same exponent k for all i. Then with $\sum r_i s_i^k = 1$,

$$x = \sum_i r_i s_i^k x = \sum_i r_i \sum_\nu a_{i,\nu} b_\nu$$

The right side is a combination of b with coefficients in A.

Step 5. Y is affine.

The algebra B of regular functions on Y is a finite-type domain because it is a finite module over the finite-type domain A. Let $\widetilde{Y} = \operatorname{Spec} B$. The fact that B is the algebra of regular functions on Y gives us a morphism $Y \xrightarrow{\epsilon} \widetilde{Y}$ (Corollary 3.5.3). Restricting to the open subset X^j of X gives us a morphism $Y^j \xrightarrow{\epsilon^j} \widetilde{Y}^j$ in which Y^j and $\widetilde{Y}^j$ are both equal to $\operatorname{Spec} B_j$. Therefore ϵ^j is an isomorphism. Corollary 3.4.21 **(ii)** shows that ϵ is an isomorphism. So Y is affine and by *Step 4*, its coordinate algebra B is a finite A-module. □

We come to Chevalley's theorem now. Let $\mathbb{P}$ denote the projective space $\mathbb{P}^n$ with coordinates $y_0, ..., y_n$.

4.7.6. Chevalley's Finiteness Theorem. *Let X be a variety, let Y be a closed subvariety of the product $\mathbb{P}\times X$, and let π denote the projection $Y \to X$. If all fibres of π are finite sets, then π is a finite morphism.*

4.7.7. Corollary. *Let Y be a projective variety and let $Y \xrightarrow{u} X$ be a morphism whose fibres are finite sets. Then u is a finite morphism. In particular, if Y is a projective curve, any nonconstant morphism $Y \xrightarrow{u} X$ is a finite morphism.*

This corollary follows from the theorem when one replaces Y by the graph of u in $Y\times X$. If Y is embedded as a closed subvariety of $\mathbb{P}$, the graph will be a closed subvariety of $\mathbb{P}\times X$ (Proposition 3.4.30). □

In the next lemma, A denotes a finite-type domain, B denotes a quotient of the algebra $A[u]$ of polynomials in n variables $u_1, ..., u_n$ with coefficients in A, and $A \xrightarrow{\varphi} B$ denotes the canonical homomorphism. We'll use capital letters for nonhomogeneous polynomials here. If $G(u)$ is a polynomial in $A[u]$, we denote its image in B by $G(u)$. too.

4.7.8. Lemma. *Let k be a positive integer. Suppose that, for each $i = 1, ..., n$, there is a polynomial $G_i(u_1, ..., u_n)$ of degree at most $k-1$ in n variables with coefficients in A, such that $u_i^k = G_i(u)$ in B. Then B is a finite A-module.*

proof. Any monomial in $u_1, ..., u_n$ of degree at least nk will be divisible by u_i^k for at least one i. So if m is a monomial of degree $d \geq nk$, the relation $u_i^k = G_i(u)$ shows that, in B, m is equal to a polynomial in $u_1, ..., u_n$ of degree less than d, with coefficients in A. By induction, it follows that the monomials of degree at most $nk-1$ span B. □

Let $y_0, ..., y_n$ be coordinates in $\mathbb{P}^n$, and let $A[y_0, ..., y_n]$ be the algebra of polynomials in y with coefficients in A. A *homogeneous element* of $A[y]$ is an element that is a homogeneous polynomial in y with coefficients in A. A *homogeneous ideal* of $A[y]$ is an ideal that can be generated by homogeneous polynomials.

4.7.9. Lemma. *Let Y be a closed subset of $\mathbb{P}\times X$, where $X = \operatorname{Spec} A$ is affine,*

(i) *The ideal $\mathcal{I}$ of elements of $A[y]$ that vanish at every point of Y is a homogeneous ideal of $A[y]$.*

(ii) *If the zero locus of a homogeneous ideal $\mathcal{I}$ of $A[y]$ is empty, then $\mathcal{I}$ contains a power of the irrelevant ideal $\mathcal{M} = (y_0, ..., y_n)$ of $A[y]$.*

proof. **(i)** Let's write a point of $\mathbb{P}\times X$ as $q = (y_0, ..., y_n, x)$, with x representing a point of X. So $(y, x) = (\lambda y, x)$. Then the proof for the case $A = \mathbb{C}$ that is given in (1.3.2) carries over.

(ii) Let V be the complement of the origin in the affine $n+1$-space with coordinates y. The complement of the origin in $V\times X$ maps to $\mathbb{P}\times X$ (see 3.2.4). If the locus of zeros of $\mathcal{I}$ in $\mathbb{P}\times X$ is empty, its locus of zeros in $V\times X$ will be contained in $o\times X$, o being the origin in $\mathbb{P}$. Then the ideal of $o\times X$, which is the radical ideal generated by the elements $y_0, ..., y_n$, will contain $\mathcal{I}$. □

proof of Chevelley's Finiteness Theorem. This proof is adapted from a proof by Schelter.
We abbreviate the notation for a product $Z\times X$ of a variety Z with X, denoting $X\times X$ by $\widetilde{Z}$.

We are given a closed subvariety Y of $\widetilde{\mathbb{P}} = \mathbb{P}\times X$, and the fibres over X are finite sets. We are to prove that the projection $Y \to X$ is a finite morphism. We may assume that X is affine, say $X = \operatorname{Spec} A$, and by induction on n, we may assume that the theorem is true when $\mathbb{P}$ is a projective space of dimension $n-1$.

Case 1. There is a hyperplane H in $\mathbb{P}$ such that Y is disjoint from $\widetilde{H} = H\times X$ in $\widetilde{\mathbb{P}} = \mathbb{P}\times X$.

This is the main case. We adjust coordinates $y_0, ..., y_n$ in $\mathbb{P}$ so that H is the hyperplane $\{y_0 = 0\}$. Because Y is a closed subvariety of $\widetilde{\mathbb{P}}$ disjoint from $\widetilde{H}$, Y is also a closed subvariety of $\widetilde{\mathbb{U}}^0 = \mathbb{U}^0\times X$, $\mathbb{U}^0$ being the standard affine $\{y_0 \neq 0\}$. So Y is affine.

Let $\mathcal{P}$ and $\mathcal{Q}$ be the (homogeneous) prime ideals in $A[y]$ that define Y and $\widetilde{H}$, respectively, and let $\mathcal{I} = \mathcal{P} + \mathcal{Q}$. So $\widetilde{q}$ is the principal ideal of $A[y]$ generated by y_0. A homogeneous element of $\mathcal{I}$ of degree k has the form $f(y) + y_0 g(y)$, where f is a homogeneous polynomial in $A[y]$ of some degree k, and g is a homogeneous polynomial of degree $k-1$.

The closed subsets Y and $\widetilde{H}$ are disjoint. Since $Y \cap \widetilde{H}$ is empty, the sum $\mathcal{I} = \mathcal{P} + \mathcal{Q}$ contains contains a power of the irrelevant ideal $\mathcal{M} = (y_0, ..., y_n)$, say $\mathcal{M}^k \subset \mathcal{I}$. Then y_i^k is in $\mathcal{I}$ for $i = 0, ..., n$. So we may write

(4.7.10) $$y_i^k = f_i(y) + y_0 g_i(y)$$

with f_i of degree k in $\mathcal{P}$ and g_i of degree $k-1$ in $A[y]$. We omit the index $i = 0$. For that index, one can take $f_0 = 0$ and $g_0 = y_0^{k-1}$.

We dehomogenize these equations with respect to the variables y, substituting $u_i = y_i/y_0$ for y_i, $i = 1, ..., n$ with $u_0 = 1$. Writing dehomogenizations with capital letters, the dehomogenized equations that correspoind to the equations $y_i^k = f_i(y) + y_0 g(y)$ have the form

$$u_i^k = F_i(u) + G_i(u) \tag{4.7.11}$$

The important point is that the degree of G_i is at most $k-1$.

Recall that Y is also a closed subset of $\mathbb{U}^0$. Let P be its (nonhomogenous) ideal in $A[u]$, which contains the polynomials $F_1, ..., F_n$. The coordinate algebra of Y is $B = A[u]/P$. In the quotient algebra B, the terms F_i drop out, leaving us with equations $u_i^k = G_i(u)$. These equations are true in B. Since G_i has degree at most $k-1$, Lemma 4.7.8 tells us that B is a finite A-algebra, as was to be shown. This completes the proof of Case 1.

Case 2. the general case.

We have taken care of the case in which there exists a hyperplane H such that Y is disjoint from $\widetilde{H}$. The next lemma shows that we can cover the given variety X by open subsets to which this special case applies. Then Lemma 4.7.4 and Proposition 4.7.4 will complete the proof.

4.7.12. Lemma. *Let Y be a closed subvariety of $\widetilde{\mathbb{P}} = \mathbb{P}^n \times X$, and suppose that the projection $Y \xrightarrow{\pi} X$ has finite fibres. Suppose also that Chevalley's Theorem has been proved for closed subvarieties of $\mathbb{P}^{n-1} \times X$. For every point p of X, there is an open neighborhood X' of p in X, and there is a hyperplane H in $\mathbb{P}$, such that the inverse image $Y' = \pi^{-1}X'$ is disjoint from $\widetilde{H}$.*

proof. Let p be a point of X, and let $\widetilde{q} = (\widetilde{q}_1, ..., \widetilde{q}_r)$ be the finite set of points of Y making up the fibre over p. We project $\widetilde{q}$ from $\mathbb{P} \times X$ to $\mathbb{P}$, obtaining a finite set $q = (q_1, ..., q_r)$ of points of $\mathbb{P}$, and we choose a hyperplane H in $\mathbb{P}$ that avoids this finite set. Then $\widetilde{H}$ avoids the fibre $\widetilde{q}$. Let W denote the closed set $Y \cap \widetilde{H}$. Because the fibres of Y over X are finite, so are the fibres of W over X. By hypothesis, Chevalley's Theorem is true for subvarieties of $\mathbb{P}^{n-1} \times X$, and $\widetilde{H}$ is isomorphic to $\mathbb{P}^{n-1} \times X$. It follows that, for every component W' of W, the morphism $W' \to X$ is a finite morphism, and therefore its image is closed in X (Theorem 4.6.6). Thus the image Z of W is a closed subset of X, and it doesn't contain p. Then $X' = X - Z$ is the required neighborhood of p. □

figure: ??I'm not sure

4.8 Double Planes

(4.8.1) affine double planes

Let A be the polynomial algebra $\mathbb{C}[x, y]$, and let X be the affine plane $\operatorname{Spec} A$. An *affine double plane* is a locus of the form $w^2 = f(x, y)$ in affine 3-space with coordinates w, x, y, where f is a square-free polynomial in x, y, as in Example 4.5.7. Let $B = \mathbb{C}[w, x, y]/(w^2 - f)$. So the affine double plane is $Y = \operatorname{Spec} B$.

We'll denote by w, x, y both the variables and their residues in B.

4.8.2. Lemma *The algebra B is a normal domain of dimension two, and a free A-module with basis $(1, w)$. It has an automorphism σ of order 2, defined by $\sigma(a + bw) = a - bw$.* □

The fibres of Y over X are the σ-orbits in Y. If $f(x_0, y_0) \neq 0$, the fibre consists of two points, and if $f(x_0, y_0) = 0$, itconsists of one point. The reason that Y is called a *double plane* is that most points of the plane X are covered by two points of Y. The *branch locus* of the covering, which will be denoted by Δ, is the (possibly reducible) curve $\{f = 0\}$ in X. The fibres over the *branch points*, points of Δ, are single points.

We study the closed subvarieties D of Y that lie over a curve C in X. These subvarieties will have dimension one, and we call them curves too. If D lies over C, and if $D = D\sigma$, then D is the only curve lying over C. Otherwise, there will be the two curves that lie over C, namely D and $D\sigma$. In that case we say that C *splits* in Y.

A curve C in X will be the zero set of a principal prime ideal P of A, and if D lies over C, it will be the zero set of a prime ideal Q of B that lies over P (4.6.3). The prime ideal Q isn't always a principal ideal.

4.8.3. Example. Let $f(x,y)=x^2+y^2-1$. The double plane $Y=\{w^2=x^2+y^2-1\}$ is an *affine quadric* in $\mathbb{A}^3$. In the affine plane, its branch locus Δ is the curve $\{x^2+y^2=1\}$.

The line $C_1:\{y=0\}$ in X meets the branch locus Δ transversally at the points $(x,y)=(\pm1,0)$, and y generates a prime ideal of B. When we set $y=0$ in the equation for Y, we obtain the irreducible polynomial w^2-x^2+1. On the other hand, the line $C_2:\{y=1\}$ is tangent to Δ at the point $(0,1)$, and it splits. When we set $y=1$ in the equation for Y, we obtain $w^2=x^2$. The locus $\{w^2=x^2\}$ is the union of the two lines $\{w=x\}$ and $\{w=-x\}$ that lie over C_1. The prime ideals of B that correspond to these lines aren't principal ideals.

□

figure circle with two lines

This example illustrates a general principle: If a curve intersects the branch locus transversally, it doesn't split. We explain this now.

(4.8.4) local analysis

Suppose that a plane curve $C:\{g=0\}$ and the branch locus $\Delta:\{f=0\}$ of a double plane $w^2=f$ meet at a point p. We adjust coordinates so that p becomes the origin $(0,0)$, and we write

$$f(x,y)=\sum a_{ij}x^iy^j=a_{10}x+a_{01}y+a_{20}x^2+\cdots$$

Since p is a point of Δ, the constant coefficient of f is zero. If the two linear coefficients aren't both zero, p will be a smooth point of Δ, and the tangent line to Δ at p will be the line $\{a_{10}x+a_{01}y=0\}$. Similarly, writing $g(x,y)=\sum b_{ij}x^iy^j$, the tangent line to C, if defined, is the line $\{b_{10}x+b_{01}y=0\}$.

Let's suppose that the two tangent lines are defined and distinct – that Δ and C intersect transversally at p. We change coordinates once more, to make the tangent lines the coordinate axes. After adjusting by scalar factors, the polynomials f and g will have the form

$$f(x,y)=x+u(x,y)\quad\text{and}\quad g(x,y)=y+v(x,y),$$

where u and v are polynomials all of whose terms have degree at least 2.

Let $X_1=\operatorname{Spec}\mathbb{C}[x_1,y_1]$ be another affine plane. We consider the map $X_1\to X$ defined by the substitution $x_1=x+u$, $y_1=y+v$. In the classical topology, this map is invertible analytically near the origin, because the Jacobian matrix

$$\left(\frac{\partial(x_1,y_1)}{\partial(x,y)}\right)_{(0,0)} \tag{4.8.5}$$

at p is the identity matrix. When we make this substitution, Δ becomes the locus $\{x_1=0\}$ and C becomes the locus $\{y_1=0\}$. In this local analytic coordinate system, the equation $w^2=f$ that defines the double plane becomes $w^2=x_1$. When we restrict it to C by setting $y_1=0$, x_1 becomes a local coordinate function on C. The restriction of the equation remains $w^2=x_1$. So the inverse image Z of C doesn't split analytically near p. Therefore it doesn't split globally either.

4.8.6. Corollary. *A curve that meets the branch locus transversally at some point doesn't split.* □

This isn't a complete analysis. When C and Δ are tangent at every point of intersection, C may split or not, and which possibility occurs cannot be decided locally in most cases. However, one case in which a local analysis suffices to decide splitting is that C is a line. Let t be a coordinate in a line C, so that $C\approx\operatorname{Spec}\mathbb{C}[t]$. Let'sassume that C does't intersect Δ at $t=\infty$. The restriction of the polynomial f to C will give us a polynomial $\overline{f}(t)$ in t. A root of $\overline{f}$ corresponds to an intersection of C with Δ, and a multiple root corresponds to an intersection at which C and Δ are tangent, or at which Δ is singular. The line C will split if and only if $\overline{f}$ is a square in $\mathbb{C}[t]$, and this will be true if and only if the multiplicity of every root of $\overline{f}$ is even.

A *rational curve* is a curve whose function field is a rational function field $\mathbb{C}(t)$ in one variable. One can make a similar analysis for any rational plane curve, a conic for example, but one needs to inspect its points at infinity and its singular points as well as the smooth points at finite distance.

(4.8.7) projective double planes

Let X be the projective plane $\mathbb{P}^2$, with coordinates x_0, x_1, x_2. A *projective double plane* is a locus of the form

$$y^2 = f(x_0, x_1, x_2), \tag{4.8.8}$$

where f is a square-free, homogeneous polynomial *of even degree* $2d$. To regard this as a homogeneous equation, we must assign weight d to the variable y (see **1.8.7**). Then, since we have weighted variables, we must work in a *weighted projective space* $\mathbb{WP}$ with coordinates x_0, x_1, x_2, y, where x_i have weight 1 and y has weight d. A point of this weighted space $\mathbb{WP}$ is represented by a nonzero vector (x_0, x_1, x_2, y) with the relation that, for all $\lambda \neq 0$, $(x_0, x_1, x_2, y) \sim (\lambda x_0, \lambda x_1, \lambda x_2, \lambda^d y)$. The points of the projective double plane Y are the points of $\mathbb{WP}$ that solve the equation (4.8.8).

The projection $\mathbb{WP} \to X$ that sends (x, y) to x is defined at all points except at $(0, 0, 0, 1)$. If (x, y) solves (4.8.8) and if $x = 0$, then $y = 0$ too. So $(0, 0, 0, 1)$ isn't a point of Y. The projection is defined at all points of Y. The fibre of the morphism $Y \to X$ over a point x consists of points (x, y) and $(x, -y)$, which will be equal if and only if x lies on the *branch locus* of the double plane, the (possibly reducible) plane curve $\Delta : \{f = 0\}$ in X. The map $\sigma : (x, y) \rightsquigarrow (x, -y)$ is an automorphism of Y, and points of X correspond bijectively to σ-orbits in Y.

Since the double plane Y is embedded into a weighted projective space, it isn't presented to us as a projective variety in the usual sense. However, it can be embedded into a projective space in the following way: The projective plane X can be embedded by a Veronese embedding of higher order, using as coordinates the monomials $m = (m_1, m_2, \ldots)$ of degree d in the variables x. This embeds X into a projective space $\mathbb{P}^N$ where $N = \binom{d+2}{2} - 1$. When we add a coordinate y of weight d, we obtain an embedding of the weighted projective space $\mathbb{WP}$ into $\mathbb{P}^{N+1}$ that sends the point (x, y) to (m, y). The double plane can be realized as a projective variety by this embedding.

If $Y \to X$ is a projective double plane, then, as happens with affine double planes, a curve C in X may split in Y or not. If C has a transversal intersection with the branch locus Δ, it will not split. On the other hand, if C is a line, and if C intersects the branch locus Δ with multiplicity 2 at every intersection point, it will split. For example, when the branch locus Δ is a generic quartic curve, the lines that split will be the bitangent lines (see Section 1.13).

(4.8.9) homogenizing an affine double plane

To construct a projective double plane from an affine double plane, we write the affine double plane as

$$w^2 = F(u_1, u_2) \tag{4.8.10}$$

for some nonhomogeneous polynomial F. We suppose that F has even degree $2d$, and we homogenize F, setting $u_i = x_i/x_0$. We multiply both sides of this equation by x_0^{2d} and set $y = x_0^d w$. This produces an equation of the form (4.8.8), where f is the homogenization of F.

If F has odd degree $2d - 1$, one needs to multiply F by x_0 in order to make the substitution $y = x_0^d w$ permissible. When we do this, the line at infinity $\{x_0 = 0\}$ becomes a part of the branch locus.

(4.8.11) cubic surfaces and quartic double planes

We use coordinates x_0, x_1, x_2, z for the (unweighted) projective 3-space $\mathbb{P}^3$ here, and X will denote the projective x-plane $\mathbb{P}^2$. Let $\mathbb{P}^3 \xrightarrow{\pi} X$ denote the projection that sends (x, z) to x. It is defined at all points except at the center of projection $q = (0, 0, 0, 1)$, and its fibres are the lines through q, with q omitted.

Let S be a cubic surface in $\mathbb{P}^3$, the locus of zeros of an irreducible homogeneous cubic polynomial $g(x, z)$. We'll denote the restriction of π to S by the same symbol π.

Let's suppose that q **is** a point of S. Then the coefficient of z^3 in g will be zero, and g will be quadratic in z: $g(x,z) = az^2 + bz + c$, where the coefficienta a, b, c are homogeneous polynomials in x, of degrees $1, 2, 3$, respectively. The equation for S becomes

$$az^2 + bz + c = 0 \tag{4.8.12}$$

The discriminant $f = b^2 - 4ac$ of g is a homogeneous polynomial of degree 4 in x. Let Y be the projective double plane

$$y^2 = b^2 - 4ac \tag{4.8.13}$$

.

We denote by V the affine space of polynomials a, b, c of degrees $1, 2, 3$ in x, and by W the affine space of homogeneous quartic polynomials in x. Sending g to its discriminant f defines a morphism $V \xrightarrow{u} W$ (4.8.12).

4.8.14. Lemma. *The image of the morphism u contains all quartic polynomials f such that the divisor $D : f = 0$ has at least one bitangent line. Therefore the image of u is dense in W.*

proof. Given such a quartic polynomial f, let a be a linear polynomial such that the line $\ell_1 : \{a = 0\}$ is a bitangent to $D : \{f = 0\}$. Then, as noted above, ℓ_1 splits in the double plane $y^2 = f$. So f is congruent to a square, modulo a. Let b be a quadratic polynomial such that $f \equiv b^2$ modulo a. When we take this polynomial as b, we will have $f = b^2 - 4ac$ for some cubic polynomial c.

Conversely, if $g(x, y) = az^2 + bz + c$, the line $\ell_1 : \{a = 0\}$ will be a bitangent to D provides that the locus $b = 0$ meets ℓ_1 in two distinct points. □

It follows from the lemma that, if $g(x, z) = az^2 + bz + c$ is a polynomial in which a, b, c are generic homogeneous polynomials in x, of degrees $1, 2, 3$, respectively, the discriminant $b^2 - 4ac$ will be a generic homogeneous quartic polynomial in x.

We go back to the generic cubic surface $S : az^2 + bz + c = 0$ and the generic double plane $Y : y^2 = b^2 - 4ac$.

4.8.15. Theorem. *A generic cubic surface S in $\mathbb{P}^3$ contains precisely 27 lines.*

This theorem follows from next lemma, which relates the 27 lines in S to the 28 bitangents of the generic quartic curve $\Delta : \{b^2 - 4ac = 0\}$ in the plane X (1.13.3).

As noted above, the line ℓ_1 defined by the linear equation $a = 0$ is a bitangent to the quartic curve Δ.

4.8.16. Lemma. *Let S be a generic cubic surface. The 27 bitangent lines in X that are distinct from ℓ_1 are the images of the lines in S, and distinct lines in S have distinct images.*

proof. Because the cubic surface S is generic, it contains finitely many lines (3.6). When we project to X from a generic point q of S, q won't lie on any of those lines. The fibres of the projection $\mathbb{P}^3 \to X$ are lines through q, and they aren't contained in S. So a line in S projects bijectively to a line in X.

A line in X is defined by a homogeneous linear equation in the variables x. The same linear equation defines a plane H in $\mathbb{P}^3$ that contains q, and the intersection $C = S \cap H$ will be a cubic curve in H. At least one of the irreducible components of C contains q, and that component isn't a line. So if C is reducible, it will be a union $Q \cup L$, where Q is a conic that contains q and L is a line in S. Thus lines L in S correspond bijectively to lines in X such that the corresponding cubic C is reducible.

Referring to (4.8.12) and (4.8.13), the quadratic formula solves for z in terms of y whenever $a \neq 0$:

$$z = \frac{-b + y}{2a} \quad \text{or} \quad y = 2az + b \tag{4.8.17}$$

These equations define a bijection $S' \longleftrightarrow Y'$ between the open subsets S' and Y' of points of S and Y at which $a \neq 0$.

If ℓ is a line in X, not the line ℓ_1, the intersection $\ell \cap \ell_1$ will be a point p. The bijection $S' \longleftrightarrow Y'$ will be defined at all points that lie over ℓ except those whose images are p. If ℓ is the image of a line in S, the cubic curve $C = S \cap H$ is reducible. Because ℓ splits in Y, it is a bitangent to the quartic curve Δ. Conversely, if ℓ splits in Y, then C will be reducible. It will be the union of a line and a conic. So every bitangent line distinct from ℓ_1 is the image of a unique line in S.

The line $\ell_1 : \{a=0\}$ is special. Its inverse image C in S is the locus of zeros of the two polynomials a and $az^2 + bz + c$, or equivalently, the locus $a = 0$ and $bz + c = 0$.

Let's adjust coordinates so that a becomes the polynomial x_0. The locus $\{x_0 = 0\}$ in $\mathbb{P}^3$ is the projective plane P with coordinates x_1, x_2, z, and in P C is the locus $\overline{g} = 0$ in that plane, where $\overline{g} = \overline{b}z + \overline{c}$, $\overline{b}, \overline{c}$ being the polynomials obtained from b, c by substituting $x_0 = 0$. In P, the point q becomes $(0, 0, 1)$, and C becomes the cubic curve $\overline{g} = 0$. The cubic curve C is singular at q because $\overline{g}$ has no term of degree > 1 in z. As we have noted, C cannot be the union $Q \cup L$, of a conic and a line that meet at q. Therefore C is irreducible. It doesn't contain a line, so ℓ_1 doesn't split. □

Summing up: The 27 bitangents distinct from ℓ_1 are images of lines in S, but ℓ_1 is not the image of a line in S.

Chapter 5 STRUCTURE OF VARIETIES II: CONSTRUCTIBLE SETS

The goal of this chapter is to explain how algebraic curves control the geometry of higher dimensional varieties. We do this, beginning in Section 5.5. We begin with a short review about modules that we will use in the chapter. We omit most proofs.

5.1 Modules

(5.1.1) exact sequences

A sequence

$$\cdots \to V^{n-1} \xrightarrow{d^{n-1}} V^n \xrightarrow{d^n} V^{n+1} \xrightarrow{d^{n+1}} \cdots$$

of homomorphisms of R-modules is *exact* if the image of d^{k-1} is equal to the kernel of d^k. For example, a sequence $0 \to V \xrightarrow{d} V'$ is exact if and only if the map d is injective, and a sequence $V \xrightarrow{d} V' \to 0$ is exact if and only if d is surjective.

Any homomorphism $V \xrightarrow{d} V'$ can be embedded into an exact sequence

$$0 \to K \to V \xrightarrow{d} V' \to C \to 0,$$

where K and C are the kernel and cokernel of d, respectively.

A *short exact sequence* is an exact sequence of the form

$$0 \to V \xrightarrow{a} V' \xrightarrow{b} V'' \to 0.$$

The statement that this sequence is exact asserts that the map a is injective, and that V'' is isomorphic to the quotient group V'/aV.

5.1.2. Proposition. *(functorial property of the kernel and cokernel) Suppose given a (commutative) diagram of R-modules*

$$\begin{array}{ccccccc} & & V & \xrightarrow{u} & V' & \longrightarrow & V'' & \longrightarrow & 0 \\ & & {\scriptstyle f}\downarrow & & {\scriptstyle f'}\downarrow & & {\scriptstyle f''}\downarrow & & \\ 0 & \longrightarrow & W & \longrightarrow & W' & \underset{v}{\longrightarrow} & W'' & & \end{array}$$

whose rows are exact sequences. Let K, K', K'' and C, C', C'' denote the kernels and cokernels of f, f', and f'', respectively.

(i) (kernel is left exact) *The kernels form an exact sequence* $K \to K' \to K''$. *If* u *is injective, the sequence* $0 \to K \to K' \to K''$ *is exact.*
(ii) (cokernel is right exact) *The cokernels form an exact sequence* $C \to C' \to C''$. *If* v *is surjective, the sequence* $C \to C' \to C'' \to 0$ *is exact.*
(iii) (Snake Lemma) *There is a canonical homomorphism* $K'' \xrightarrow{d} C$ *that combines with the above sequences to form an exact sequence*

$$K \to K' \to K'' \xrightarrow{d} C \to C' \to C''.$$

If u *is injective and/or* v *is surjective, the sequence remains exact with zeros at the appropriate ends.* □

(5.1.3) tensor products

Let U and V be modules over a ring R. The *tensor product* $U \otimes_R V$ is an R-module that is generated by elements $u \otimes v$ called tensors, one for each u in U and v in V. Its elements are combinations of tensors with coefficients in R.

The defining relations among the tensors are the *bilinear relations*:

(5.1.4) $$(u_1+u_2) \otimes v = u_1 \otimes v + u_2 \otimes v\,, \quad u \otimes (v_1+v_2) = u \otimes v_1 + u \otimes v_2$$

and $$r(u \otimes v) = (ru) \otimes v = u \otimes (rv)$$

for all u in U, v in V, and r in R. The tensor symbol $\otimes$ is used as a reminder that the elements $u \otimes v$ are to be manipulated using these relations.

One can absorb a coefficient from R into either one of the factors of a tensor, so every element of $U \otimes_R V$ can be written as a finite sum $\sum u_i \otimes v_i$ with u_i in U and v_i in V.

5.1.5. Example. Let U be the space of m dimensional (complex) column vectors, and let V be the space of n-dimensional row vectors. Then $U \otimes_{\mathbb{C}} V$ identifies naturally with the space of $m \times n$-matrices. If U and V are free R-modules with bases $\{u_i\}$ and $\{v_j\}$, respectively, then $U \otimes_R V$ is a free R-module with basis $\{u_i \otimes v_j\}$.

There is an obvious map of sets $U \times V \xrightarrow{\beta} U \otimes_R V$ from the product *set* to the tensor product, that sends (u, v) to $u \otimes v$. This map isn't a module homomorphism. The defining relations (5.1.4) show that it is *R-bilinear*, not linear. It is a universal bilinear map.

5.1.6. Corollary. *Let* U, V, *and* W *be* R*-modules. Homomorphisms of* R*-modules* $U \otimes_R V \to W$ *correspond bijectively to* R*-bilinear maps* $U \times V \to W$. □

Any R-bilinear map $U \times V \xrightarrow{f} W$ to a module W can be obtained from a module homomorphism $U \otimes_R V \xrightarrow{\widetilde{f}} W$ by composition with the bilinear map β defined above: $U \times V \xrightarrow{\beta} U \otimes_R V \xrightarrow{\widetilde{f}} W$.

This follows from the defining relations. □

5.1.7. Proposition. *There are canonical isomorphisms*
- $U \otimes_R R \approx U$, *defined by* $u \otimes r \leftrightsquigarrow ur$
- $(U \oplus U') \otimes_R V \approx (U \otimes_R V) \oplus (U' \otimes_R V)$, *defined by* $(u_1 + u_2) \otimes v \leftrightsquigarrow u_1 \otimes v + u_2 \otimes v$
- $U \otimes_R V \approx V \otimes_R U$, *defined by* $u \otimes v \leftrightsquigarrow v \otimes u$
- $(U \otimes_R V) \otimes_R W \approx U \otimes_R (V \otimes_R W)$, *defined by* $(u \otimes v) \otimes w \leftrightsquigarrow u \otimes (v \otimes w)$ □

5.1.8. Proposition. *Tensor product is right exact: Let* $U \xrightarrow{f} U' \xrightarrow{g} U'' \to 0$ *be an exact sequence of* R*-modules, then for any* R*-module* V, *the sequence below is exact:*

$$U \otimes_R V \xrightarrow{f \otimes id} U' \otimes_R V \xrightarrow{g \otimes id} U'' \otimes_R V \to 0$$

□

Tensor product operation isn't left exact. For example, Let $R = \mathbb{C}[x]$. Then $R/xR \approx \mathbb{C}$, so there is an exact sequence $0 \to R \xrightarrow{x} R \to \mathbb{C} \to 0$. When we tensor with $\mathbb{C}$ we get the sequence $0 \to \mathbb{C} \xrightarrow{0} \mathbb{C} \to \mathbb{C} \to 0$. The zero map isn't injective.

5.1.9. Corollary. *Let U and V be modules over a domain R and let s be a nonzero element of R. Let R_s, U_s, V_s be the (simple) localizations of R, U, V, respectively.*
(i) *There is a canonical isomorphism $U_s \approx U \otimes_R (R_s)$.*
(ii) *Tensor product is compatible with localization: $U_s \otimes_{R_s} V_s \approx (U \otimes_R V)_s$* □

We note that the product *module* $U \times V$ and the tensor product module $U \otimes_R V$ are very different. For instance, when U and V are free modules of ranks r and s, $U \times V$ is free of rank $r+s$, while $U \otimes_R V$ is free of rank rs.

(5.1.10) extension of scalars in a module

Let $R \xrightarrow{\rho} S$ be a ring homomorphism. Extension of scalars constructs an S-module from an R-module.

Let's write scalar multiplication on the right. So M will be a right R-module. Then $M \otimes_R S$ be comes an S-module, multiplication by $s \in S$ being $(m \otimes a)s = m \otimes (as)$. This gives the functor

$$R{-}modules \overset{\otimes S}{\longrightarrow} S{-}modules$$

that is called *extension of scalars*.

(5.1.11) localization, again

If s is a nonzero element of a domain A, the *simple localization* A_s, which is often referred to simply as a *localization*, is the ring obtained by adjoining an inverse of s, and to work with the inverses of finitely many nonzero elements, one may simply adjoin the inverse of their product. For working with an infinite set of inverses, the concept of a multiplicative system is useful. A *multiplicative system* S in a domain A is a subset that consists of nonzero elements, is closed under multiplication, and contains 1. If S is a multiplicative system, the ring of S*-fractions* AS^{-1} is the ring obtained by adjoining inverses of all elements of S. Its elements are equivalence classes of fractions as^{-1} with a in A and s in S, the equivalence relation and the laws of composition being the usual ones for fractions. The ring AS^{-1} called a *localization* too..

5.1.12. Examples. **(i)** The set consisting of the powers of a nonzero element s of a domain A is a multiplicative system whose ring of fractions is the simple localization $A_s = A[s^{-1}]$.

(ii) The set S of all nonzero elements of a domain A is a multiplicative system whose ring of fractions is the field of fractions of A.

(iii) An ideal P of a domain A is a prime ideal if and only if its complement, the set of elements of A **not** in P, is a multiplicative system. □

5.1.13. Definition. Let $A \subset B$ be a ring extension, and let I and J be ideals of A and B, respectively. The *extension* of I is the ideal IB of B generated by I, whose elements are finite sums $\sum_i z_i b_i$ with z_i in I and b_i in B. The *contraction* of J is the intersection $J \cap A$, which is an ideal of A.

The next lemma explains what happens when one combines extension and contraction.

5.1.14. Lemma. *Let $A \subset B$ be rings, and let I and J be ideals of A and B, respectively. Then $I \subset (IB) \cap A$, and $(J \cap A)B \subset J$.* □

5.1.15. Proposition. *Let S be a multiplicative system in a domain A, and let A' be the localization AS^{-1}.*
(i) *Let I be an ideal of A. The extended ideal IA' is the set IS^{-1} whose elements are classes of fractions xs^{-1}, with x in I and s in S. The extended ideal is the unit ideal if and only if I contains an element of S.*

(ii) *Let J be an ideal of the localization A' and let I denote its contraction $J \cap A$. The extended ideal IA' is equal to J: If J is an ideal of A', then $J = (J \cap A)A'$.*
(iii) *If Q is a prime ideal of A and if $Q \cap S$ is empty, the extended ideal $Q' = QA'$ is a prime ideal of A', and the contraction $Q' \cap A$ is equal to Q. If $Q \cap S$ isn't empty, the extended ideal is the unit ideal. Thus prime ideals of AS^{-1} correspond bijectively to prime ideals of A that don't meet S.* □

5.1.16. Corollary. *Every localization AS^{-1} of a noetherian domain A is noetherian.* □

(5.1.17) a general principle

An important, though elementary, principle for working with fractions is that any finite sequence of computations in a localization AS^{-1} will involve only finitely many denominators, and can therefore be done in a simple localization A_s, where s is a common denominator for the fractions that occur. This principle has been mentioned before, in Proposition 4.4.10.

For example, let $A \subset B$ be finite-type domains, and let S be the multiplicative system of nonzero elemets of A. Then $AS^{-1} = K$ is the field of fractions of A, and $B_K = BS^{-1}$ is a finite-type K-algebra. The Noether Normalization Theorem tells us that B_K is a finite module over a polynomial subring $K[y_1, ..., y_n]$, Therefore there is a nonzero element s in A such that B_s is a finite module over the polynomial ring $A_s[y_1, ..., y_n]$.

###ugh fix this ###

(5.1.18) module homomorphisms

Beginning in Chapter 6, we will work with modules over various rings. Let $R \xrightarrow{\rho} R'$ be a ring homomorphism, let M be an R-module, and let M' be an R'-module. A *homomorphism $M \xrightarrow{\varphi} M'$ associated to the ring homomorphism ρ* is an additive group homomorphism $M \xrightarrow{\varphi} M'$ compatible with scalar multiplication. If m in M and a in R, then

(5.1.19) $$\varphi(am) = \rho(a)\varphi(m)$$

For example, if M is a module over a domain R and s is a nonzero element of R, the localization M_s is an R_s-module. The homomorphism $M \to M_s$ is compatible with the localization map $R \to R_s$.

If $A \xrightarrow{\rho} B$ is a ring homomorphism, a B-module N can be made into an A-module by *restriction of scalars*, scalar multiplication by an element a of A being defined by the formula

(5.1.20) $$an = \rho(a)n$$

If it seems necessary in order to avoid confusion, we may denote the B-module N and the A-module obtained from it by restriction of scalars by ${}_BN$ and ${}_AN$, respectively.

Let $M \xrightarrow{\varphi} M'$ be a homomorphism compatible with a ring homomorphism $R \xrightarrow{\rho} R'$. When M' is made into an R-module by restriction of scalars, φ becomes a homomorphism of R-modules.

(5.1.21) localizing a module

Let S be a multiplicative system in a domain A. The *localization MS^{-1}* of an A-module M is defined in a natural way, as the AS^{-1}-module whose elements are equivalence classes of fractions ms^{-1} with m in M and s in S, and there will be a homomorphism $M \to MS^{-1}$ that sends an element m to the fraction $m/1$. The only complication comes from the fact that M may have S-torsion elements – nonzero elements m such that $ms = 0$ for some s in S. If $ms = 0$ and s is in S, then m must map to zero in MS^{-1}, because in MS^{-1}, we will have $m = mss^{-1}$.

To define MS^{-1}, it suffices to modify the equivalence relation. Two fractions $m_1s_1^{-1}$ and $m_2s_2^{-1}$ are defined to be equal if there is an element $\widetilde{s} \in S$ such that $m_1s_2\widetilde{s} = m_2s_1\widetilde{s}$. Then $ms_1^{-1} = 0$ if and only if $m\widetilde{s} = 0$ for some $\widetilde{s}$ in S. This takes care of torsion, and MS^{-1} becomes an AS^{-1}-module.

This is also how one localizes a ring that isn't a domain.

When S is the set of powers of an element s, the localized A_s-module will be denoted by M_s. Its elements have the form ms^{-k}, $k \geq 0$, and $m_1s^{-k_1} = m_2s^{-k_2}$ if for sufficiently large n, $m_1s^{k_2+n} = m_2s^{k_1+n}$.

5.1.22. Proposition. *Let S be a multiplicative system in a domain A.*
(i) *Localization is an exact functor: A homomorphism $M \xrightarrow{\varphi} N$ of A-modules induces a homomorphism $MS^{-1} \xrightarrow{\varphi'} NS^{-1}$ of AS^{-1}-modules, and if $M \xrightarrow{\varphi} N \xrightarrow{\psi} P$ is an exact sequence of A-modules, the localized sequence $MS^{-1} \xrightarrow{\varphi'} NS^{-1} \xrightarrow{\psi'} PS^{-1}$ is exact.*
(ii) *Let M be an A-module. and let N be an AS^{-1}-module. Homomorphisms of AS^{-1}-modules $MS^{-1} \to N$ correspond bijectively to homomorphisms of A-modules $M \to N$.*
(iii) *If multiplication by s is an injective map $M \to M$ for every s in S, then $M \subset S^{-1}M$. If multiplication by every s is a bijective map $M \to M$, then $M \approx S^{-1}M$.* □

(5.1.23) local rings

A *local ring* is a noetherian ring that contains just one maximal ideal. If an element of a local ring R isn't in its maximal ideal M, then it isn't in any maximal ideal, so it is a unit. A local ring R will have a *residue field* R/M. The case that is most important for us is that the residue field is the field of complex numbers.

We make a few general comments about local rings here though we will be interested mainly in some special local rings, discrete valuation rings that are discussed below.

The Nakayama Lemma 4.3.1 has a useful version for local rings:

5.1.24. Local Nakayama Lemma. *Let R be a local ring with maximal ideal M, let V be a finite R-module, and let $\overline{V}$ be the quotient V/MV, which is a vector space over the residue field k of R as well as an R-module.*
(i) *If $\overline{V} = 0$, then $V = 0$.*
(ii) *Let $v = (v_1, ..., v_r)$ be a set of elements of V, and let $\overline{v} = (\overline{v}_1, ..., \overline{v}_r)$ be the residues of v in $\overline{V}$. If $\overline{v}$ spans $\overline{V}$, then v spans V.*

proof. **(i)** If $\overline{V} = 0$, then $V = MV$. The Nakayama Lemma tells us that M contains an element z such that $1-z$ annihilates V. Then $1-z$ isn't in M, so it is a unit. A unit annihilates V, and therefore $V = 0$.

(ii) Let V' be the submodule of V spanned by v, let $W = V/V'$, and let $\overline{W} = W/MW$. To show that v spans V, we show that $W = 0$, and according to **(i)**, it suffices to show that $\overline{W} = 0$.

We inspect the iagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & MV & \longrightarrow & V & \xrightarrow[a]{} & \overline{V} & \longrightarrow & 0 \\ & & \downarrow & & \downarrow{\scriptstyle b} & & \downarrow{\scriptstyle c} & & \\ 0 & \longrightarrow & MW & \longrightarrow & W & \xrightarrow[d]{} & \overline{W} & \longrightarrow & 0 \end{array}$$

Its rows are exact, the maps labelled a, b, c, d are surjective, and the kernel of b is V'. If $\overline{v}$ spans $\overline{V}$, i.e., if $aV' = \overline{V}$, then $caV' = \overline{W} = dbV'$. Since $bV' = 0$, it follows that $\overline{W} = 0$, as required. □

5.1.25. Corollary. *Let R be a local ring with maximal ideal M and residue field k, and let $m = \{m_1, ..., m_r\}$ be a set of elements of M. If the residues of m span the k-vector space M/M^2, then m spans M.* □

(5.1.26) the local ring at a point

Let $\mathfrak{m}$ be the maximal ideal at a point p of an affine variety $X = \operatorname{Spec} A$. The complement S of $\mathfrak{m}$ is a multiplicative system (5.1.12)**(iii)**), and the prime ideals P of the localization AS^{-1} (the ring obtained by inverting the elements of S) are extensions of the prime ideals Q of A that are contained in $\mathfrak{m}$: $P = QS^{-1}$ (5.1.15). Thus AS^{-1} is a local ring whose maximal ideal is $\mathfrak{m}S^{-1}$. This ring is called the *local ring of A at p* , and is often denoted by A_p.

For example, let $X = \operatorname{Spec} A$ be the affine line, $A = \mathbb{C}[t]$, and let p be the point $t = 0$. The local ring A_p is the ring whose elements are fractions $f(t)/g(t)$ with $g(0) \neq 0$.

Any finite set $\alpha_1, ..., \alpha_k$ of elements of the local ring A_p at p will be contained in a simple localization A_s, for some s in S. It will be in the coordinate algebra of the affine open neighborhood X_s of p.

5.2 Valuations

A local domain R with maximal ideal M has *dimension one* if (0) and M are distinct, and if those ideals are the only prime ideals of R. In this section, we describe the *normal* local domains of dimension one. They are the discrete valuation rings that are defined below.

Let K be a field. A *discrete valuation* v on K is a surjective homomorphism

(5.2.1) $$K^\times \xrightarrow{\text{v}} \mathbb{Z}^+$$

from the multiplicative group of nonzero elements of K to the additive group of integers such that, if a, b are elements of K and if a, b and $a{+}b$ aren't zero, then

- $\text{v}(a{+}b) \geq \min\{\text{v}(a), \text{v}(b)\}$.

The word "discrete" refers to the fact that $\mathbb{Z}^+$ has the discrete topology. Other valuations exist. They are interesting, but less important, and we won't use them. To simplify terminology, we refer to a discrete valuation simply as a *valuation*.

Let k be a positive integer. If v is a valuation and if $\text{v}(a) = k$, then k is the *order of zero of* a, and if $\text{v}(a) = -k$, then k is the *order of pole of* a, with respect to the valuation.

5.2.2. Lemma. *If* v *is a valuation on a field* K *that contains the complex numbers, every nonzero complex number has value zero.*

proof. This is true because $\mathbb{C}$ contains n th roots. If γ is an n th root of a nonzero complex number c, then because v is a homomorphism, $\text{v}(\gamma) = \text{v}(c)/n$. The only integer that is divisible by every positive integer n is zero. □

The *valuation ring* R associated to a valuation v on a field K is the subring of elements of K with non-negative values, together with zero:

(5.2.3) $$R = \{a \in K^\times \mid \text{v}(a) \geq 0\} \cup \{0\}.$$

Valuation rings are usually called "discrete valuation rings", but since we have dropped the word discrete from the valuation, we drop it from the valuation ring too.

5.2.4. Proposition. *Valuations of the field* $\mathbb{C}(t)$ *of rational functions in one variable correspond bijectively to points of the projective line* $\mathbb{P}^1_t$*. The valuation ring that corresponds to a point* $p \neq \infty$ *is the local ring of the polynomial ring* $\mathbb{C}[t]$ *at* p.

beginning of the proof. Let K denote the field $\mathbb{C}(t)$, and let a be a complex number. To define the valuation v that corresponds to the point $p:\ t = a$ of $\mathbb{P}^1$, we write a nonzero polynomial f as $(t-a)^k h$, where $t-a$ doesn't divide h, and we define, $\text{v}(f) = k$. Then we define $\text{v}(f/g) = \text{v}(f) - \text{v}(g)$. You will be able to check that with this definition, v becomes a valuation whose valuation ring is the local ring at p. The valuation that corresponds to the point at infinity of $\mathbb{P}^1$ is obtained by working with t^{-1} in place of t.

The proof that these are all of the valuations of $\mathbb{C}(t)$ will be given at the end of the section.

5.2.5. Proposition. *Let* v *be a valuation on a field* K*, let* R *be its valuation ring, and let* x *be an element of the multiplicative group* $K^\times$ *with value* $\text{v}(x) = 1$.

(i) *The ring* R *is a normal local domain of dimension one. Its maximal ideal* M *is the principal ideal* xR*. The elements of* M *are those that have positive value:*

$$M = \{a \in K^\times \mid \text{v}(a) > 0\} \cup \{0\}$$

(ii) *The units of* R *are the elements of* $K^\times$ *with value zero. Every element* z *of* $K^\times$ *has the form* $z = x^k u$*, where* u *is a unit and* $k = \text{v}(z)$ *is an integer.*

(iii) *The proper* R*-submodules of* K *are the sets* $x^k R$*, where* k *is a positive or negative integer. The set* $x^k R$ *consists of zero and the elements of* $K^\times$ *with value* $\geq k$*. The nonzero ideals of* R *are the principal ideals* $x^k R$ *with* $k \geq 0$*, the powers of the maximal ideal.*

(iv) *There is no ring properly between* R *and* K*: If* R' *is a ring and if* $R \subset R' \subset K$*, then either* $R = R'$ *or* $R' = K$.

proof. We prove **(i)** last.

(ii) Since v is a homomorphism, $\mathrm{v}(u^{-1}) = -\mathrm{v}(u)$. So u and u^{-1} are both in R, i.e., u is a unit, if and only if $\mathrm{v}(u) = 0$. If z is a nonzero element of K with $\mathrm{v}(z) = k$, then $u = x^{-k}z$ has value zero, so it is a unit, and $z = ux^k$.

(iii) The R-module x^kR consists of the elements of K of value at least k. Suppose that an R-submodule N of K contains an element z with value k. Then $z = ux^k$, where u is a unit, and therefore N contains x^k and x^kR. If k is the smallest integer such that N contains an element z with value k, then $N = x^kR$. If there is no minimum value of the elements of N, then N contains x^kR for every k, and $N = K$.

(iv) This follows from **(iii)**. The ring R' will be an R-submodule of K. If $R' \neq K$, then $R' = x^kR$ for some k, and since R' contains R, $k \le 0$. If $k < 0$ then x^kR isn't closed under multiplication. So $k = 0$ and $R' = R$.

(i) First, R is noetherian because **(iii)** tells us that it is a principal ideal domain, and it follows from **(ii)** that the only prime ideals of R are $\{0\}$ and $M = xR$. So R is a local ring of dimension 1. If the normalization of R were larger than R, then according to **(iv)**, it would be equal to K. Then x^{-1} would be integral over R. It would satisfy a polynomial relation $x^{-r} + a_1x^{-(r-1)} + \cdots + a_r = 0$ with a_i in R. When one multiplies this relation by x^r, one sees that 1 would be a multiple of x. Then x would be a unit, which it is not. □

5.2.6. Theorem.
(i) *A local domain whose maximal ideal is a nonzero principal ideal is a valuation ring.*
(ii) *Every normal local domain of dimension* 1 *is a valuation ring.*

proof. **(i)** Let R be a local domain whose maximal ideal M is a nonzero principal ideal, say $M = xR$, with $x \neq 0$, and let y be a nonzero element of R. The integers k such that x^k divides y are bounded (4.3.5). Let x^k be the largest power that divides y. Then $y = ux^k$, where $k \ge 0$ and u isn't in M. It is a unit. Then any nonzero element z of the fraction field K of R will have the form $z = ux^r$ where u is a unit and r is an integer, possibly negative. This is shown by writing the numerator and denominator of a fraction in such a form and dividing.

The valuation whose valuation ring is R is defined by $\mathrm{v}(z) = r$ when $z = ux^r$ as above. If $z_i = u_ix^{r_i}$, $i = 1, 2$, where u_i are units and $r_1 \le r_2$, then $z_1 + z_2 = \alpha x^{r_1}$, where $\alpha = u_1 + u_2x^{r_2-r_1}$ is an element of R. Therefore $\mathrm{v}(z_1 + z_2) \ge r_1 = \min\{\mathrm{v}(z_1), \mathrm{v}(z_2)\}$. We also have $\mathrm{v}(z_1z_2) = \mathrm{v}(z_1) + \mathrm{v}(z_2)$. Thus v is a surjective homomorphism. The requirements for a valuation are satisfied.

(ii) The fact that a valuation ring is a normal, one-dimensional local ring is Proposition 5.2.5 **(i)**. We show that a normal local domain R of dimension 1 is a valuation ring by showing that its maximal ideal is a principal ideal. The proof is a bit tricky.

Let z be a nonzero element of M. Because R is a local ring of dimension 1, M is the only prime ideal that contains z, so M is the radical of the principal ideal zR, and $M^r \subset zR$ if r is large. Let r be the smallest integer such that $M^r \subset zR$. Then there is an element y in M^{r-1} that isn't in zR, but such that $yM \subset zR$. We restate this by saying that $w = y/z$ isn't in R, but $wM \subset R$. Since M is an ideal, multiplication by an element of R carries wM to wM. So wM is an ideal. Since M is the maximal ideal of the local ring R, either $wM \subset M$, or $wM = R$. If $wM \subset M$, the lemma below shows that w is integral over R. This can't happen because R is normal and w isn't in R. Therefore $wM = R$ and $M = w^{-1}R$. This implies that w^{-1} is in R and that M is a principal ideal. □

5.2.7. Lemma. *Let I be a nonzero ideal of a noetherian domain A, and let B be a domain that contains A. An element w of B such that $wI \subset I$ is integral over A.*

proof. This is the Nakayama Lemma once more. Because A is noetherian, I is finitely generated. Let $v = (v_1, ..., v_n)^t$ be a vector whose entries generate I. The hypothesis $wI \subset I$ allows us to write $wv_i = \sum p_{ij}v_j$ with p_{ij} in A, or in matrix notation, $wv = Pv$. So w is an eigenvalue of P. If $p(t)$ denotes the characteristic polynomial of P, $p(w)v = 0$. Since $I \neq 0$, at least one v_i is nonzero. Since A is a domain, $p(w)v_i = 0$ implies that $p(w) = 0$. The characteristic polynomial is a monic polynomial with coefficients in A, so w is integral over A. □

5.2.8. Lemma. *A rational function α on a variety X is regular on X if it is in the local ring of X at every point p.*

This is true because a function α is in the local ring at p if and only if it is in the coordinate algebra of some affine neighborhood of p (**5.1.26**). □

5.2.9. Corollary. *Let $X = \operatorname{Spec} A$ be an affine variety.*

(i) *The coordinate algebra A is the intersection of the local rings A_p at points of X.*

$$A = \bigcap_{p \in X} A_p$$

(ii) *The coordinate algebra A is normal if and only if all of its local rings A_p are normal.*

See Lemma 4.5.3 for **(ii)**. □

5.2.10. Note. (*about the overused word local*) A property is true *locally* on a topological space X if every point p of X has an open neighborhood U such that the property is true on U.

The words *localize* and *localization* refer to the process of adjoining inverses. The localizations X_s of an affine variety $X = \operatorname{Spec} A$ form a basis for the topology on X. So if some property is true locally on X, one can cover X by localizations on which the property is true. There will be elements $s_1, ..., s_k$ of A that generate the unit ideal, such that the property is true on each of the localizations X_{s_i}.

An A-module M is *locally free* if there are elements $s_1, ..., s_k$ that generate the unit ideal of A, such that M_{s_i} is a free A_{s_i}-module for each i. If a locally free A-module U that is locally isomorphic to A^k, then U has *rank* k.

An ideal I of A is *locally principal* if there are elements s_i that generate the unit ideal, such that I_{s_i} is a principal ideal of A_{s_i}. □

5.2.11. Corollary. *Let M be a finite module over a finite-type domain A. If for some point p of $X = \operatorname{Spec} A$ the localized module M_p (5.1) is a free module, there is an element s not in $\mathfrak{m}_p$ such that M_s is free.*

proof. See the general principle (**5.1.17**). □

We finish the proof of Proposition 5.2.4 now, by showing that every valuation v of the function field $K = \mathbb{C}(t)$ of $\mathbb{P}^1$ corresponds to a point of $\mathbb{P}^1$.

Let R be the valuation ring of v. If $\mathrm{v}(t) < 0$, we replace t by t^{-1}. So we may assume that $\mathrm{v}(t) \geq 0$. Then t is an element of R, and therefore $\mathbb{C}[t] \subset R$. The maximal ideal M of R isn't zero. It contains a nonzero element of K, a fraction $\alpha = f/g$ of polynomials in t. The denominator g is in R, so M also contains the nonzero polynomial $f = g\alpha$. Since M is a prime ideal, it contains an irreducible factor of f. The irreducible polynomials in t are linear, so M contains $t - a$ for some complex number a. Then $t - c$ isn't in M when $c \neq a$, because the scalar $c - a$ cannot be in M. Since R is a local ring, $t - c$ is a unit of R for all $c \neq a$. The localization R_0 of $\mathbb{C}[t]$ at the point $t = a$ is a valuation ring that is contained in the valuation ring R (5.2.4). There is no ring properly containing R_0 except K, so $R_0 = R$. □

5.3 Smooth Curves

A *curve* is a variety of dimension 1. The proper closed subsets of a curve are its nonempty finite subsets.

5.3.1. Definition. A point p of a curve X is a *smooth point* if the local ring at p is a valuation ring. Otherwise, p is a *singular point*. A curve X is *smooth* if all of its points are smooth.

##be careful about smooth curve and smooth affine curve##

Let p be a smooth point of a curve X, and let v_p be the corresponding valuation. As with any valuation, we say that a rational function α on X has a *zero of order* $k > 0$ at p if $\mathrm{v}_p(\alpha) = k$, and that it has a *pole of order* k at p if $\mathrm{v}_p(\alpha) = -k$.

5.3.2. Lemma. **(i)** *An affine curve X is smooth if and only if its coordinate algebra is a normal domain.*

(ii) *A curve has finitely many singular points.*

(iii) *The normalization $\widetilde{X}$ of a curve X is a smooth curve, and the canonical morphism $\widetilde{X} \to X$ becomes an isomorphism when the finite set of singular points of X and their inverse images are deleted.*

proof. **(i)** This follows from Theorem 5.2.6 and Proposition 4.5.3.

(ii),(iii) Any nonempty open subset of a curve X will be the complement of a finite set, so we may replace X by an affine open subset, say $\operatorname{Spec} A$. The normalization $\widetilde{A}$ of A will be a finite A-module, and therefore a finite-type algebra with the same fraction field as A, and $\operatorname{Spec} \widetilde{A}$ will be a smooth curve. It follows from the principle **5.1.17** that $\widetilde{A}$ and A have a common localization, say A_s. The open subset $X_s = \operatorname{Spec} A_s$ of X will be smooth. □

5.3.3. Proposition. *Let X be a smooth curve with function field K. Every point of $\mathbb{P}^n$ with values in K defines a morphism $X \to \mathbb{P}^n$.*

proof. A point $(\alpha_0, ..., \alpha_n)$ of $\mathbb{P}^n$ with values in K determines a morphism $X \to \mathbb{P}^n$ if and only if, for every point p of X, there is an index j such that the functions α_i/α_j are regular at p for every i (3.4.14). This will be true when j is chosen so that the order of zero $\mathrm{v}_p(\alpha_j)$ of α_j at p is minimal. □

As the next example shows, the analog of this proposition isn't true for varieties X of dimension greater than one.

5.3.4. Example. Let Y be the complement of the origin in the affine plane $X = \operatorname{Spec} \mathbb{C}[x, y]$, and let $K = \mathbb{C}(x, y)$ be the function field of X. The vector (x, y) defines a point of $\mathbb{P}^1_{x,y}$ with values in K. This point can be written as $(1.y/x)$ and also as $(x/y, 1)$. So (x, y) defines a morphism to $\mathbb{P}^1$ wherever at least one of the functions x/y or y/x is regular, which is true at all points of Y. However, there is no way to extend the morphism to X. □

5.3.5. Proposition. *Let $X = \operatorname{Spec} A$ be a smooth affine curve with function field K. The local rings of X are the valuation rings of K that contain A. Therefore the maximal ideals of A are locally principal, and if R is a valuation ring with maximal ideal M, its residue field R/M is isomorphic to $\mathbb{C}$.*

proof. Since A is a normal domain of dimension one, its local rings are valuation rings that contain A (Theorem 5.2.6). Let R be a valuation ring of K that contains A, let v be the associated valuation, and let M be the maximal ideal of R. The intersection $M \cap A$ is a prime ideal of A. Since A has dimension 1, the zero ideal is the only prime ideal of A that isn't a maximal ideal. We can clear the denominator of an element of M, multiplying by an element of R, to obtain an element of A while staying in M. So $M \cap A$ isn't the zero ideal. It is the maximal ideal $\mathfrak{m}_p$ of A at a point p of X. The elements of A that aren't in $\mathfrak{m}_p$ aren't in M either, so they are invertible in R. Therefore the local ring A_p, at p, a valuation ring, is contained in R. This implies that $A_p = R$ (5.2.5) **(iii)**. □

5.3.6. Proposition. *Let X' and X be smooth curves with the same function field K.*

(i) *Any morphism $X' \xrightarrow{f} X$ that is the identity on the function field K maps X' isomorphically to an open subvariety of X.*

(ii) *If X is projective, every smooth curve X' with function field K is isomorphic to an open subvariety of X.*

(iii) *If X' and X are both projective, they are isomorphic.*

(iv) *If X is projective, every valuation ring of K is the local ring at a point of X.*

proof. **(i)** Let q be a point of X', let U be an affine open neighborhood of $p = fq$, and let V be an affine open neighborhood of q in X' that is contained in the inverse image of U. Say $U = \operatorname{Spec} A$ and $V = \operatorname{Spec} B$. The morphism f gives us a homomorphism $A \to B$, and since q maps to p, this homomorphism extends to an inclusion of local rings $A_p \subset B_q$. These rings are valuation rings with the same field of fractions, so $A_p = B_q$. Since B is a finite-type algebra, there is an element s in A, with $s(q) \neq 0$, such that $A_s = B_s$. The open subsets $\operatorname{Spec} A_s$ of X and $\operatorname{Spec} B_s$ of X' are the same. Since q is an arbitrary point of X', X' is covered by open subvarieties of X. So it is an open subvariety of X too.

(ii) The projective embedding $X \subset \mathbb{P}^n$ is defined by a point $(\alpha_0, ..., \alpha_n)$ with values in K, and that same point defines a morphism $X' \to \mathbb{P}^n$. If $f(x_0, ..., x_n) = 0$ is a set of defining equations of X in $\mathbb{P}^n$, then $f(\alpha) = 0$ in K, and therefore f vanishes on X' too (3.4.9). So the image of X' is contained in the zero locus of f, which is X. Then **(i)** shows that X' is an open subvariety of X.

(iii) This follows from **(ii)**.

(iv) The local rings of X are normal and of dimension one. They are valuation rings. We prove the converse. Let $\beta = (\beta_0, ..., \beta_n)$ be the point with values in K that defines the projective embedding of X. and let R be a valuation ring of K, and let v be the corresponding valuation. We order the coordinates so that $\text{v}(\beta_0)$ is minimal. Then the ratios $\gamma_j = \beta_j/\beta_0$ will be in R. The coordinate algebra A_0 of the affine variety $X^0 = X \cap \mathbb{U}^0$ is generated by the coordinate functions γ_j. So $A_0 \subset R$, and R is the local ring of X^0 at some point 5.3.5. □

5.3.7. Proposition. < *Let* $X = \operatorname{Spec} A$ *be an affine curve, and let* $\mathfrak{m}$ *and* v *be the maximal ideal and valuation, respectively, at a smooth point* p. *Let* R *be the valuation ring of* v *and let* M *be its maximal ideal.*
(i) *The power* $\mathfrak{m}^k$ *of* $\mathfrak{m}$ *consists of the elements of* A *whose values are at least* k. *If* I *is an ideal of* A *whose radical is* $\mathfrak{m}$, *then* $I = \mathfrak{m}^k$ *for some* $k > 0$.
(ii) *The algebras* $A/\mathfrak{m}^{n+1}$ *and* R/M^{n+1} *are isomorphic to the truncated polynomial ring* $\mathbb{C}[t]/(t^{n+1})$.
(iii) *If* X *is a smooth affine curve, every nonzero ideal* I *of* A *is a product* $\mathfrak{m}_1^{e_1} \cdots \mathfrak{m}_k^{e_k}$ *of powers of maximal ideals.*

proof. **(i)** The nonzero ideals of R are powers of M. Let I be an ideal of A whose radical is $\mathfrak{m}$, and let k be the minimal value $\text{v}(x)$ of the nonzero elements x of I. We will show that I is the set of all elements of A with value $\geq k$, i.e., that $I = M^k \cap A$. Since we can apply the same reasoning to $\mathfrak{m}^k$, it will follow that $I = \mathfrak{m}^k$.

Let x be an element of I with value k, and let y be an element with value at least k. Then x divides y in R, say $y/x = u$, with u in R. The element u will be a fraction a/s with s, a in A and s not in $\mathfrak{m}$, and $sy = ax$. The element s will vanish at a finite set of points $q_1, ..., q_r$, but not at p. We choose an element z of A that vanishes at p but not at any of the points $q_1, ..., q_r$. Then z is in $\mathfrak{m}$, and since the radical of I is $\mathfrak{m}$, some power of z is in I. We replace z by that power. Then z is in I. By our choice, z and s have no common zeros in X. They generate the unit ideal of A. We write $1 = cs + dz$ with c and d in A. Then $y = csy + dzy = cax + dzy$. Since x and z are in I, so is y.

(ii) Since p is a smooth point, the local ring of A at p is the valuation ring R, and A contains an element t with value $\text{v}(t) = 1$. Let P be the subring $\mathbb{C}[t]$ of A, and let $\overline{P}_k = P/(t)^k, \overline{A}_k = A/\mathfrak{m}^k$, and $\overline{R}_k = R/M^k$. Since $\mathfrak{m}$ isn't the zero ideal, $\mathfrak{m}^{k-1} < \mathfrak{m}^k$ (Corollary 4.3.5**(ii)**). It follows from **(i)** that $t\mathfrak{m}^{k-1} = \mathfrak{m}^k$. Therefore $\mathfrak{m}^{k-1}/\mathfrak{m}^k$ has $\mathbb{C}$-dimension 1. The map labelled g_{k-1} in the diagram below is bijective.

$$\begin{array}{ccccccccc}
0 & \longrightarrow & (t^{k-1})/(t^k) & \longrightarrow & \overline{P}_k & \longrightarrow & \overline{P}_{k-1} & \longrightarrow & 0 \\
 & & {\scriptstyle g_{k-1}}\downarrow & & {\scriptstyle f_k}\downarrow & & {\scriptstyle f_{k-1}}\downarrow & & \\
0 & \longrightarrow & \mathfrak{m}^{k-1}/\mathfrak{m}^k & \longrightarrow & \overline{A}_k & \longrightarrow & \overline{A}_{k-1} & \longrightarrow & 0
\end{array}$$

Induction on k shows that the map labelled f_{k-1} is bijective, so f_k is bijective. A similar argument shows that $\overline{P}_k$ and $\overline{R}_k$ are isomorphic

(iii) Let I be a nonzero ideal of A. Because X has dimension one, the locus of zeros of I is a finite set $\{p_1, ..., p_k\}$. Therefore the radical of I is the intersection $\mathfrak{m}_1 \cap \cdots \cap \mathfrak{m}_k$ of the maximal ideals $\mathfrak{m}_j$ at p_j, which, by the Chinese Remainder Theorem, is the product ideal $\mathfrak{m}_1 \cdots \mathfrak{m}_k$, Moreover, I contains a power of that product, say $I \supset \mathfrak{m}_1^N \cdots \mathfrak{m}_k^N$ Let $J = \mathfrak{m}_1^N \cdots \mathfrak{m}_k^N$. The quotient algebra A/J is the product $B_1 \times \cdots \times B_k$, with $B_j = A/\mathfrak{m}_j^N$, and A/I is a quotient of A/J. Proposition 2.1.7 tells us that A/I is a product $\overline{A}_1 \times \cdots \times \overline{A}_k$, where $\overline{A}_j$ is a quotient of B_j. By part **(ii)**, each B_j is a truncated polynomial ring. Then the quotients $\overline{A}_j$ must also be truncated polynomial rings. So the kernel I of the map $A \to \overline{A}_1 \times \cdots \times \overline{A}_k$ is a product of powers of the maximal ideals $\mathfrak{m}_j$. □

(5.3.8) isolated points

5.3.9. Proposition. *A curve, smooth or not, contains no point that is isolated in the classical topology.*

This was proved before for plane curves (Proposition 1.3.18).

5.3.10. Lemma.
(i) *Let* Y' *be an open subvariety of a variety* Y. *Then* q *is an isolated point of* Y *if and only if it is an isolated point of* Y'.

(ii) *Let $Y' \xrightarrow{u'} Y$ be a nonconstant morphism of curves and let q' be a point of Y'. If the image of q' is an isolated point of Y, then q' is an isolated point of Y'.*

proof. **(i)** A point q of Y is isolated if $\{q\}$ is an open subset of Y. If $\{q\}$ is open in Y' and Y' is open in Y, then $\{q\}$ is open in Y. If $\{q\}$ is open in Y, it is open in Y'.

(ii) Because Y' has dimension one, the fibre over q will be a finite set, say $\{q'\} \cup F$, where F is finite. Let Y'' denote the (open) complement $Y' - F$ of F in Y', and let u'' be the restriction of u' to Y''. The fibre of Y'' over q is the point q'. If $\{q\}$ is open in Y, then because u'' is continuous, $\{q'\}$ will be open in Y'', and therefore open in Y'. □

proof of Proposition 5.3.9. Let q be a point of a curve Y. Part **(i)** of Lemma 5.3.10 allows us to replace Y by an affine neighborhood of q. Let Y' be the normalization of Y. Part **(ii)** of the lemma allows us to replace Y by Y'. So we may assume that Y is a smooth affine curve, say $Y = \operatorname{Spec} B$. We can still replace Y by an open neighborhood of q, so we may assume that the maximal ideal $\mathfrak{m}_q$ is a principal ideal.

Say that $B = \mathbb{C}[x_1, ..., x_n]/(f_1, ..., f_k)$, q is the origin $(0, ..., 0)$ in $\mathbb{A}^n_x$, and that the maximal ideal $\mathfrak{m}_q$ is generated by the residue of a polynomial f_0 in B. Then $f_0, ..., f_k$ generate the maximal ideal $(x_1, ..., x_n)$ in $\mathbb{C}[x_1, ..., x_n]$. Let's write $f_i = \sum_1^n c_{ij} x_j + O(2)$, where $O(2)$ denotes an undetermined polynomial, all of whose terms have degree ≥ 2 in x. The coefficient c_{ij} is the partial derivative $\frac{\partial f_i}{\partial x_j}$, evaluated at q. If J denotes $(k+1) \times n$ Jacobian matrix $\left(\frac{\partial f_i}{\partial x_j}\right)$ at q, we have $(f_0, ..., f_k)^t = J(x_1, ..., x_n)^t + O(2)$. Since $f_0, ..., f_k$ generate the maximal ideal, there is a matrix P with polynomial entries such that $Pf^t = x^t$. Then $x^t = PJx^t + O(2)$. If P_0 is the constant term of P, $P_0 J$ will be the identity matrix. So J has rank n.

Let J^1 be the matrix obtained by deleting the column with index 0 from J. This matrix has rank at least $n-1$, and we may assume that the submatrix with indices $1 \leq i, j \leq n-1$ is invertible. The Implicit Function Theorem says that the equations $f_1, ..., f_{n-1}$ can be solved for the variables $x_1, ..., x_{n-1}$ as analytic functions of x_n, for small x_n. The locus Z of zeros of $f_1, ..., f_{n-1}$ has dimension at most 1, it is locally homeomorphic to the affine line (1.4.18), and it contains Y. Since Y has dimension 1, the component of Z that contains q must be equal to Y. So Y is locally homeomorphic to $\mathbb{A}^1$, which has no isolated point. Therefore q isn't an isolated point of Y.

5.4 Constructible Sets

In this section, X will denote a noetherian topological space. Every closed subset of X is a finite union irreducible closed sets (2.2.13).

The intersection $L = C \cap U$ of a closed set C and an open set U is a *locally closed* set. Open sets and closed sets are examples of locally closed sets.

A *constructible set* is a set that is the union of finitely many locally closed sets.

5.4.1. Lemma. *The following conditions on a subset L of A are equivalent.*
- *L is locally closed.*
- *L is a closed subset of an open subset U of X.*
- *L is an open subset of a closed subset C of X.* □

5.4.2. Examples.
(i) A subset S of a curve X is constructible if and only if it is either a finite set or the complement of a finite set. Thus S is constructible if and only if it is either closed or open.
(ii) Let C be the line $\{y = 0\}$ in the affine plane $X = \operatorname{Spec} \mathbb{C}[x, y]$, let $U = X - C$ be its open complement, and let $p = (0, 0)$. The union $U \cup \{p\}$ is constructible, but not locally closed. □

We will use the following notation: L is a locally closed set, C is a closed set, and U is an open set.

5.4.3. Theorem. *The family of constructible subsets of a noetherian topological space X, which we denote by $\mathbb{S}$, is the smallest family of subsets that contains the open sets and is closed under the three operations of finite unions, finite intersections, and complementation.*

proof. Let $\mathbb{S}_1$ denote the family of subsets obtained from the open sets by the three operations mentioned in the statement. Open sets are constructible, and with those three operations, one can produce any constructible set from the open sets. So $\mathbb{S} \subset \mathbb{S}_1$. To show that $\mathbb{S} = \mathbb{S}_1$, we show that the family of constructible sets is closed under the three operations.

It is obvious that a finite union of constructible sets is constructible. The intersection of two locally closed sets $L_1 = C_1 \cap U_1$ and $L_2 = C_2 \cap U_2$ is locally closed because $L_1 \cap L_2 = (C_1 \cap C_2) \cap (U_1 \cap U_2)$. If $S = L_1 \cup \cdots \cup L_k$ and $S' = L_1' \cup \cdots \cup L_r'$ are constructible sets, the intersection $S \cap S'$ is the union of the locally closed intersections $(L_i \cap L_j')$, so it is constructible.

Let S be the constructible set $L_1 \cup \cdots \cup L_k$. Its complement is the intersection of the complements of L_i: $S^c = L_1^c \cap \cdots \cap L_k^c$. We have shown that intersections of constructible sets are constructible. So to show that the complement S^c is constructible, it suffices to show that the complement of a locally closed set is constructible. Let L be the locally closed set $C \cap U$, and let C^c and U^c be the complements of C and U, respectively. Then C^c is open and U^c is closed. The complement L^c of L is the union $C^c \cup U^c$ of constructible sets, so it is constructible. □

5.4.4. Proposition. *Let X be a noetherian topological space. Every constructible subset S is a union $L_1 \cup \cdots \cup L_k$ of locally closed sets $L_i = C_i \cap U_i$, in which the closed sets C_i are irreducible and distinct.*

proof. Suppose that $L = C \cap U$ is a locally closed set, and let $C = C_1 \cup \cdots \cup C_r$ be the decomposition of C into irreducible components. Then $L = (C_1 \cap U) \cup \cdots \cup (C_r \cap U)$, which is constructible. So every constructible set S is a union of locally closed sets $L_i = C_i \cap U_i$ in which the C_i are irreducible. Next, suppose that two of the irreducible closed sets are equal, say $C_1 = C_2$. Then $L_1 \cup L_2 = (C_1 \cap U_1) \cup (C_1 \cap U_2) = C_1 \cap (U_1 \cup U_2)$ is locally closed. So we can find an expression in which the closed sets are distinct as well. □

5.4.5. Lemma.

(i) *Let X_1 be a closed subset of a variety X, and let X_2 be its open complement. A subset S of X is constructible if and only if $S \cap X_1$ and $S \cap X_2$ are constructible.*

(ii) *Let X' be an open or a closed subvariety of a variety X.*

a) *If S is a constructible subset of X, then $S' = S \cap X'$ is a constructible subset of X'.*

b) *If S' is a constructible subset of X', then it is a constructible subset of X.*

proof. **(i)** This follows from Theorem 5.4.3.

(iia) It suffices to prove that the intersection $L' = L \cap X'$ of a locally closed subset L of X is a locally closed subset of X'. If $L = C \cap U$, then $C' = C \cap X'$ is closed in X', and $U' = U \cap X'$ is open in X'. So $L' = C' \cap U'$ is locally closed.

(iib) It suffices to show that a locally closed subset $L' = C' \cap U'$ of X' it is locally closed in X. If X' is closed in X, then C' is closed in X, and $U' = X \cap U$ for some open subset U of X. If X' is open in X, then U' is open in X, and if C is the closure of C' in X, then $C \cap U' = C' \cap U'$. So $L' = C \cap U'$ is locally closed in X. □

The next theorem illustrates a general fact, that sets arising in algebraic geometry tend to be constructible.

5.4.6. Theorem. *Let $Y \xrightarrow{f} X$ be a morphism of varieties. The inverse image of a constructible subset of X is a constructible subset of Y. The image of a constructible subset of Y is a constructible subset of X.*

proof. The fact that a morphism is continuous implies that the inverse image of a constructible set is constructible. To prove that the image of a constructible set is constructible, one keeps reducing the problem until there is nothing left to do.

Let S be a constructible subset of Y. Noetherian induction allows us to assume that the theorem is true when S is contained in a proper closed subvariety of Y, and also when its image $f(S)$ is contained in a proper closed subvariety of X.

Suppose that Y is the union of a proper closed subvariety Y_1 and its open complement Y_2, and let $S_i = S \cap Y_i$. It suffices to show that S_i is a constructible subset of Y_i, $i = 1, 2$, and induction applies to Y_1. So we may replace Y by any nonempty open subvariety.

Let X_1 be a proper closed subvariety of X and let X_2 be its open complement. The inverse image $Y_1 = f^{-1}(X_1)$ will be closed in Y, and its open complement will be the inverse image $Y_2 = f^{-1}(X_2)$.

A constructible subset S of Y is the union of the constructible sets $S_1 = S \cap Y_1$ and $S_2 = S \cap Y_2$. So it suffices to show that $f(S_i)$ is constructible. To show this, it suffices to show that $f(S_i)$ is a constructible subset of X_i for $i = 1, 2$ (5.4.5) **(iib)**. Moreover, induction applies to X_1. So we need only show that $f(S_2)$ is a constructible subset of X_2. This means that we can replace X and Y by nonempty open subsets finitely many times.

Since S is a finite union of locally closed sets, it suffices to treat the case that S is locally closed. Moreover, we may suppose that $S = C \cap U$, where C is irreducible. Then Y is the union of the closed subset $C = Y_1$ and its complement Y_2. Since $S \cap Y_2 = \emptyset$, it suffices to treat Y_1. We may replace Y by C. So we may assume that $S = Y \cap U = U$, and we may replace Y by U. We are thus reduced to the case that $S = Y$.

At this point, we may still replace X and Y by nonempty open subsets, so we may assume that they are affine, say $Y = \operatorname{Spec} B$ and $X = \operatorname{Spec} A$. Then the morphism $Y \to X$ corresponds to an algebra homomorphism $A \xrightarrow{\varphi} B$. If the kernel of φ were nonzero, the image of Y would be contained in a proper closed subset of X to which induction would apply. So we may assume that φ is injective.

Proposition 4.4.10 tells us that, for suitable nonzero s in A, B_s will be a finite module over a polynomial subring $A_s[y_1, ..., y_k]$. Then the maps $Y_s \to \operatorname{Spec} A_s[y]$ and $\operatorname{Spec} A_s[y] \to X_s$ are both surjective, so Y_s maps surjectively to X_s. When we replace X and Y by X_s and Y_s, the map $Y \to X$ becomes surjective, and we are done. □

5.5 Closed Sets

Limits of sequences are often used to analyze subsets of a topological space. In the classical topology, a subset Y of $\mathbb{C}^n$ is closed if, whenever a sequence of points in Y has a limit in $\mathbb{C}^n$, the limit is in Y. In algebraic geometry one uses curves as substitutes.

We use the following notation:

(5.5.1) *C is a smooth affine curve, q is a point of C, and C' is the complement of q in C.*

The closure of C' will be C, and we think of q as a limit point. Theorem 5.5.3, which is below, asserts that a constructible subset of a variety is closed if it contains all such limit points.

The next theorem tells us that there are enough curves to do the job.

5.5.2. Theorem. *(enough curves) Let Y be a constructible subset of a variety X, and let p be a point of its closure $\overline{Y}$. There exists a morphism $C \xrightarrow{f} X$ from a smooth affine curve to X, and a point q of C with $f(q) = p$, such that the image of $C' = C - \{q\}$ is contained in Y.*

proof. We use Krull's Theorem to slice Y down to dimension 1. If $X = p$, then $Y = p$ too. In this case, we may take for f the constant morphism from any curve C to p. So we may assume that X has dimension at least one. Next, we may replace X by an affine open subset X' that contains p, and Y by $Y' = Y \cap X'$. The closure $\overline{Y}'$ of Y' in X' will be the intersection $\overline{Y} \cap X'$, and it will contain p. So we may assume that X is affine, say $X = \operatorname{Spec} A$.

Since Y is constructible, it is a union $L_1 \cup \cdots \cup L_k$ of locally closed sets, say $L_i = Z_i \cap U_i$ where Z_i are irreducible closed sets and U_i are open sets. (We use Z_i in place of C_i here to avoid confusion with a curve.) The closure of Y is the union $Z_1 \cup \cdots \cup Z_k$, and p is in one of the closed sets Z_i. We may replace X by Z_i and Y by L_i, so we may assume that Y is a nonempty open subset of X.

Suppose that the dimension n of X is at least two. Let $D = X - Y$ be the (closed) complement of the open set Y. The components of D have dimension at most $n-1$. We choose an element α of the coordinate algebra A of X that is zero at p and isn't identically zero on any component of D except p itself, if p happens to be a component. Krull's Theorem tells us that every component of the zero locus of α has dimension $n-1$, and at least one of those components, call it V, contains p. If V were contained in D, it would be a component of D because $\dim V = n-1$ and $\dim D \le n-1$. By our choice of α, this isn't the case. So $V \not\subset D$, and therefore $V \cap Y \ne \emptyset$. Because V is irreducible and Y is open, $V \cap Y$ is an open dense subset of V, and p is a point of its closure V. We replace X by V and Y by $V \cap Y$. The dimension of X is thereby reduced to $n-1$.

Thus it suffices to treat the case that X has dimension one. Then X will be a curve that contains p and Y will be a nonempty open subset of X. The normalization of X will be a smooth curve x_1 that comes with an integral and therefore surjective morphism to Y. Finitely many points of X_1 will map to p. We choose for C an affine open subvariety of X_1 that contains just one of those points, and we call that point q. $\square$

5.5.3. Theorem *(curve criterion for a closed set) Let Y be a constructible subset of a variety X. The following conditions are equivalent:*

(a) *Y is closed.*

(b) *For every morphism $C \xrightarrow{f} X$ from a smooth affine curve to X, the inverse image $f^{-1}Y$ is closed in C.*

(c) *Let q be a point of a smooth affine curve C, let $C' = C - \{q\}$, and let $C \xrightarrow{f} X$ be a morphism. If $f(C') \subset Y$, then $f(C) \subset Y$.*

The hypothesis that Y be constructible is necessary. For example, let X be the affine line $\mathbb{A}^1$. The set Z of points of X with integer coordinates isn't constructible, but it satisfies the curve criterion. Any morphism $C' \to X$ whose image is in Z will map C' to a single point, and therefore it will extend to C.

proof. The implications **(a)** $\Rightarrow$ **(b)** $\Rightarrow$ **(c)** are obvious. We prove the contrapositive of the implication **(c)** $\Rightarrow$ **(a)**. Suppose that Y isn't closed. We choose a point p of the closure $\overline{Y}$ that isn't in Y, and we apply Theorem 5.5.2. There exists a morphism $C \xrightarrow{f} X$ from a smooth curve to X and a point q of C such that $f(q) = p$ and $f(C') \subset Y$. Since $q \notin Y$, this morphism shows that **(c)** doesn't hold either. $\square$

5.5.4. Theorem. *A constructible subset Y of a variety X is closed in the Zariski topology if and only if it is closed in the classical topology.*

proof. A Zariski closed set is closed in the classical topology because the classical topology is finer than the Zariski topology.

Suppose that Y is closed in the classical topology. Let q be a point of the Zariski closure $\overline{Y}$ of Y, and let $C \xrightarrow{f} X$ be a morphism from a smooth affine curve to X that maps the complement C' of q to Y. Let $Y' = f^{-1}Y$. Then Y' contains C', so it is ieither C' or C. A morphism is a continuous map in the classical topology. Since Y is closed in the classical topology, Y' is closed in C. If Y' were equal to C', then $\{q\}$ would be open as well as closed. It would be an isolated point of C. Since a curve contains no isolated point, the closure is C. Therefore the curve criterion (5.5.3) is satisfied, and Y is closed in the Zariski topology. $\square$

5.6 Fibred Products

(5.6.1) the mapping property of a product

The product $X\times Y$ of two sets X and Y has a mapping property that is easy to verify: Maps from a set T to the product set $X\times Y$, correspond bijectively to pairs of maps $T \xrightarrow{f} X$ and $T \xrightarrow{g} Y$. The map $T \xrightarrow{(f,g)} X\times Y$ defined by the pair of maps f, g sends a point t to the point pair $(f(t), g(t))$.

Let $X\times Y \xrightarrow{\pi_1} X$ and $X\times Y \xrightarrow{\pi_2} Y$ denote the projection maps. If $T \xrightarrow{h} X\times Y$ is a map to the product, the corresponding maps to X and Y are the compositions with the projections: $T \xrightarrow{\pi_1\circ h} X$ and $T \xrightarrow{\pi_2\circ h} Y$:

The analogous statements are true for morphisms of varieties.

5.6.2. Proposition. *Let X and Y be varieties, and let $X\times Y$ be the product variety.*

(i) *The projections $X\times Y \xrightarrow{\pi_1} X$ and $X\times Y \xrightarrow{\pi_2} Y$ are morphisms.*

(ii) *Morphisms from a variety T to the product variety $X\times Y$ correspond bijectively to pairs of morphisms $T \to X$ and $T \to Y$, the correspondence being the same as for maps of sets.* $\square$

It was proved in Proposition 3.4.35 that if $X \xrightarrow{f} Z$ and $Y \xrightarrow{g} W$ are morphisms of varieties, the product map $X\times Y \xrightarrow{f\times g} Z\times W$ defined by $[f\times g](x,y) = (f(x), g(y))$ is a morphism.

(5.6.3) fibred products of sets

If $X \xrightarrow{f} Z$ and $Y \xrightarrow{g} Z$ are maps of sets, the *fibred product* $X \times_Z Y$ is the subset of the product $X \times Y$ consisting of pairs of points x, y such that $f(x) = g(y)$. It fits into a diagram

(5.6.4)
$$\begin{array}{ccc} X \times_Z Y & \xrightarrow{\pi_2} & Y \\ {\scriptstyle \pi_1}\big\downarrow & & {\scriptstyle g}\big\downarrow \\ X & \xrightarrow{f} & Z \end{array}$$

in which π_1 and π_2 are the projections. Many important subsets of a product can be described as fibred products. If a map $Y \to Z$ is given, and if $p \to Z$ is the inclusion of a point into Z, then $p \times_Z Y$ is the fibre of Y over p. The diagonal in $X \times X$ is the fibred product $X \times_X X$.

The reason for the term "fibred product" is that the fibre of $X \times_Z Y$ over a point x of X maps bijectively to the fibre of Y over the image $z = f(x)$, and the analogous statement is true for fibres over points of Y.

(5.6.5) fibred products of varieties

Since we are working with varieties, not schemes, we have a small problem: The fibred product of varieties will be a scheme, but it needn't be a variety.

5.6.6. Example. Let $X = \operatorname{Spec} \mathbb{C}[x]$, $Y = \operatorname{Spec} \mathbb{C}[y]$ and $Z = \operatorname{Spec} \mathbb{C}[z]$ be affine lines, let $X \xrightarrow{f} Z$ and $X \xrightarrow{g} Z$ be the maps defined by $z = x^2$ and $z = y^2$, respectively. The fibred product $X \times_Z Y$ is the closed subset of the affine x, y-plane consisting of points (x, y) such that $x^2 = y^2$. It is the union of the two lines $x = y$ and $x = -y$. □

The next proposition will be enough for our purposes.

5.6.7. Proposition. *Let $X \xrightarrow{f} Z$ and $Y \xrightarrow{g} Z$ be morphisms of varieties. The fibred product $X \times_Z Y$ is a closed subset of the product variety $X \times Y$.*

proof. The graph Γ_f of a morphism $X \xrightarrow{f} Z$ of varieties is a closed sub*variety* of $X \times Z$ isomorphic to X (Proposition 3.4.30). Next, let u and v be two morphisms from a variety X to another variety: Z. We show that the set W consisting of points x in X such that $u(x) = v(x)$ is a closed subset of X. In $X \times Z$, let W' be the intersection of the graphs of u and v: $W' = \Gamma_u \cap \Gamma_v$. A point (x, z) is in W' if $z = ux = vx$. This is an intersection of closed sets, so it is closed in Γ_u (and in Γ_v). The projection $\Gamma_u \to X$, which is an isomorphism, carries W' to W, so W is closed in X.

With reference to Diagram 5.6.4, $X \times_Z Y$ is the subset of the product $X \times Y$ of points at which the maps $f\pi_X$ and $g\pi_Y$ to Z are equal, so it is closed in $X \times Y$. □

5.7 Projective Varieties are Proper

As has been noted (3.1), an important property of projective space with the classical topology is that it is a compact space. A variety isn't compact in the Zariski topology unless it is a single point. Howver, in the Zariski topology, projective varieties have a property closely related to compactness: They are *proper*.

Before defining the concept of a proper variety, we explain an analogous property of compact spaces.

5.7.1. Proposition. *Let X be a compact space, let Z be a Hausdorff space, and let C be a closed subset of $Z \times X$. The image of C in Z is a closed subset of Z.*

proof. Let D be the image of C. We show that if a sequence of points z_i of D has a limit $\underline{z}$ in Z, then $\underline{z}$ is in D. For each i, we choose a point p_i of C that lies over z_i. So p_i is a pair (z_i, x_i), x_i being a point of X. Since X is compact, there is a subsequence of the sequence x_i that has a limit $\underline{x}$ in X. Passing to subsequences, we may suppose that x_i has limit $\underline{x}$. Then p_i will have the limit $\underline{p} = (\underline{z}, \underline{x})$. Since C is closed, $\underline{p}$ is in C, and therefore $\underline{z}$ is in D. □

5.7.2. Definition. A variety X is *proper* if has the following property: Let $Z\times X$ be the product with another variety Z, let π denote the projection $Z\times X \longrightarrow Z$, and let C be a closed subvariety of $Z\times X$. The image $D=\pi(C)$ of C is a closed subvariety of Z.

(5.7.3)
$$\begin{array}{ccc} C & \xrightarrow{\subset} & Z\times X \\ \downarrow & & \downarrow{\scriptstyle \pi} \\ D & \xrightarrow{\subset} & Z \end{array}$$

If X is proper, then because every closed set is a finite union of closed subvarieties, the image of any closed subset of $Z\times X$ will be closed in Z.

5.7.4. Theorem. *Projective varieties are proper.*

This is the most important application of the use of curves to characterize closed sets.

proof. Let X be a projective variety. With notation as in Definition 5.7.2, suppose we are given a closed subvariety C of the product $Z\times X$. We must show that its image D is a closed subvariety of Z. If the image is a closed set, it will be irreducible. So it suffices to show that D is closed, and to do this, it suffices to show that D is closed in the classical topology (Theorem 5.5.4). Theorem 5.4.6 tells us that D is a constructible set, and since X is closed in projective space, it is compact in the classical topology. Proposition 5.7.1 tells us that D is closed in the classical topology. □

The next examples show how the theorem can be used.

5.7.5. Example. (*singular curves*) We parametrize the plane curves of a given degree d. The number of distinct monomials $x_0^ix_1^jx_2^k$ of degree $d\ =i{+}j{+}k$ is the binomial coefficient $\binom{d+2}{2}$. We order those monomials arbitrarily, and label them as $m_0,...,m_r$, with $r=\binom{d+2}{2}-1$. A homogeneous polynomial of degree d will be a combination $\sum z_im_i$ of monomials with complex coefficients z_i, so the homogeneous polynomials f of degree d in x, taken up to scalar factors, are parametrized by the projective space of dimension r with coordinates z. Let's denote that projective space by Z. Points of Z correspond bijectively to divisors of degree d in the projective plane.

The product variety $Z\times\mathbb{P}^2$ represents pairs (D,p), where D is a divisor of degree d and p is a point of $\mathbb{P}^2$. A variable homogeneous polynomial of degree d in x will be a bihomogeneous polynomial $f(z,x)$ of degree 1 in z and degree d in x. So the locus Γ: $\{f(z,x)=0\}$ in $Z\times\mathbb{P}^2$ is a closed set. Its points are pairs (D,p) such that D is the divisor of f and p is a point of D.

Let Σ be the set of pairs (D,p) such that p is a singular point of D. This is also a closed set. It is defined by the system of equations $f_0(z,x)=f_1(z,x)=f_2(z,x)=0$, where f_i are the partial derivatives $\frac{\partial f}{\partial x_i}$. Euler's Formula shows that then $f(x,z)=0$. The partial derivatives f_i are bihomogeneous, of degree 1 in z and degree $d{-}1$ in x.

The next proposition isn't very easy to prove directly, but the proof becomes easy when one uses the fact that projective space is proper.

5.7.6. Proposition *The singular divisors of degree d, the divisors containing at least one singular point, form a closed subset S of the projective space Z of all divisors of degree d.*

proof. The points of S are the images of points of the set Σ via projection to Z. Theorem 5.7.4 tells us that the image of Σ is closed. □

5.7.7. Example. (*surfaces that contain a line*) We go back to the discussion of lines in a surface, as in (3.6). Let $\mathbb{S}$ denote the projective space that parametrizes surfaces of degree d in $\mathbb{P}^3$, as before.

5.7.8. Proposition *In $\mathbb{P}^3$, the surfaces of degree d that contain a line form a closed subset of the space $\mathbb{S}$.*

proof. Let $\mathbb{G}$ be the Grassmanian $G(2,4)$ of lines in $\mathbb{P}^3$, and let Ξ be the subset of $\mathbb{G}\times\mathbb{S}$ of pairs of pairs $[\ell]$, $[S]$ such that $\ell\subset S$. Lemma 3.6.17 tells us that Ξ is a closed subset of $\mathbb{G}\times\mathbb{S}$. Therefore its image in $\mathbb{S}$ is closed.□

5.8 Fibre Dimension

A function $Y \xrightarrow{\delta} \mathbb{Z}$ from a variety to the integers is *constructible* if, for every integer n, the set of points of Y such that $\delta(p) = n$ is constructible, and δ is *upper semicontinuous* if for every n, the set of points such that $\delta(p) \geq n$ is closed. For brevity, we refer to an upper semicontinuous function as *semicontinuous*, though the term is ambiguous. since a function might be lower semicontinuous.

A function δ on a curve C is semicontinuous if and only if for every integer n, there is a nonempty open subset C' of C such that $\delta(p) = n$ for all points p of C' and $\delta(p) \geq n$ for all points not in C'.

The next curve criterion for semicontinuous functions follows from the criterion for closed sets.

5.8.1. Proposition. *(curve criterion for semicontinuity) Let Y be a variety. A function $Y \xrightarrow{\delta} \mathbb{Z}$ is semicontinuous if and only if it is a constructible function, and for every morphism $C \xrightarrow{f} Y$ from a smooth curve C to Y, the composition $\delta \circ f$ is a semicontinuous function on C.* □

Let $Y \xrightarrow{f} X$ be a morphism of varieties, let q be a point of Y, and let Y_p be the fibre of f over $p = f(q)$. The *fibre dimension* $\delta(q)$ of f at q is the maximum among the dimensions of the components of the fibre that contain q.

5.8.2. Theorem. *(semicontinuity of fibre dimension) Let $Y \xrightarrow{u} X$ be a morphism of varieties, and let $\delta(q)$ denote the fibre dimension at a point q of Y.*

(i) *Suppose that X is a smooth curve, that Y has dimension n, and that the image of u is a point. Then δ is constant: Every nonempty fibre has constant dimension $n - 1$.*

(ii) *Suppose that the image of Y contains a nonempty open subset of X, and let the dimensions of X and Y be m and n, respectively. There is a nonempty open subset X' of X such that $\delta(q) = n-m$ for every point q in the inverse image of X'.*

(iii) *δ is a semicontinuous function on Y.*

We leave the proof of this theorem as an exercise. When you have done it, you will have understood the chapter.

Chapter 6 MODULES

This chapter explains how modules on a variety are defined. variety.

We will need few facts about localization. Recall that, if s is a nonzero element of a domain A, the symbol A_s stands for the localization $A[s^{-1}]$, and if $\operatorname{Spec} A = X$, then $\operatorname{Spec} A_s = X_s$.

- Let $U = \operatorname{Spec} A$ be an affine variety. The intersection of two localizations $U_s = \operatorname{Spec} A_s$ and $U_t = \operatorname{Spec} A_t$ is the localization $U_{st} = \operatorname{Spec} A_{st}$.
- Let $W \subset V \subset U$ be affine open subsets of a variety X. If V is a localization of U and W is a localization of V, then W is a localization of U (2.5.21).
- The affine open subsets of a variety X form a basis for the topology on a variety X. The localizations of an affine variety form a basis for its topology (2.5.20).
- If U and V are affine open subsets of X, the open sets W that are localizations, both of U and of V, form a basis for the topology on $U \cap V$. (2.5.21).

6.1 The Structure Sheaf.

We introduce two categories associated to a variety X. The first is the category *(opens)*. Its objects are the open subsets of X, and its morphisms are inclusions: If U and V are open sets and if $V \subset U$, there is a unique morphism $V \to U$ in *(opens)*. If $V \not\subset U$ there is no morphism $V \to U$.

We also introduce a subcategory *(affines)* of the category *(opens)*. Its objects are the affine open subsets of X, and its morphisms are localizations. A morphism $V \to U$ in *(opens)* $-$ an inclusion $V \subset U$ of open subsets $-$ is a morphism in *(affines)* if U is affine and V is a localization of U $-$ if V is an open subset of the form U_s, where s is a nonzero element of the coordinate algebra of U.

The *structure sheaf* $\mathcal{O}_X$ on a variety X is the functor

(6.1.1)
$$\textit{(affines)}^\circ \xrightarrow{\mathcal{O}_X} \textit{(algebras)}$$

from affine open sets to algebras, that sends an affine open set $U = \operatorname{Spec} A$ to its coordinate algebra A, which is then denoted by $\mathcal{O}_X(U)$.

As has been noted before, inclusions $V \to U$ of affine open subsets needn't be localizations. We focus attention on localizations because the relationship between the coordinate algebras of an affine variety and a localization is easy to understand. However, the structure sheaf can be extended without much difficulty to the category *(opens)*, (See Corollary 6.1.3 below.)

A brief review about regular functions: The *function field* F of a variety X is the field of fractions of the coordinate algebra of any one of its affine open subsets, and a *rational function* on X is a nonzero element of

F. A rational function is *regular* on an affine open set $U = \operatorname{Spec} A$ if it is an element A, and is *regular* on any nonempty open set that can be covered by affine open sets on which it is regular. The function field of a variety X contains the regular functions on every nonempty open subset, and the regular functions on X are governed by the regular functions on its affine open subsets.

An affine variety is determined by its regular functions, but regular functions don't suffice to determine a variety that isn't affine. For instance, the only rational functions that are regular everywhere on the projective line $\mathbb{P}^1$ are the constant functions, which are useless. We will be interested in regular functions on non-affine open sets, especially in functions that are regular on the whole variety, but one should always work with the affine open sets, where the definition of a regular function is clear.

6.1.2. Lemma. *Let U and V be open subsets of a variety X, with $V \subset U$. If a rational function is regular on U, it is also regular on V.* □

Thus if $U \subset V$ is an inclusion of affine open subsets, say $U = \operatorname{Spec} A$ and $V = \operatorname{Spec} B$, then $A \subset B$. However, it won't be clear how to construct B from A unless B is a localization. If $V = U_s$, then $B = A[s^{-1}]$. If B isn't a localization, the exact relationship between A and B remains obscure.

6.1.3. Corollary. *When $\mathcal{O}_X(U)$ is defined to be the algebra of regular functions on U, the structure sheaf $\mathcal{O}_X$ on a variety X extends to a functor*

$$\textit{(opens)}^\circ \xrightarrow{\mathcal{O}_X} \textit{(algebras)}$$

from all open subsets to algebras, . □

If it is clear which variety is being studied, we may write $\mathcal{O}$ for $\mathcal{O}_X$.

As for affine open sets, the algebra of regular functions on an open set U is denoted by $\mathcal{O}_X(U)$. Its elements are called *sections* of the structure sheaf $\mathcal{O}_X$ on U.

When $V \to U$ is a morphism in *(opens)*. Lemma 6.1.2 tells us that $\mathcal{O}_X(U)$ is contained in $\mathcal{O}_X(V)$. This gives us the homomorphism, an inclusion,

$$\mathcal{O}_X(U) \to \mathcal{O}_X(V)$$

that makes $\mathcal{O}_X$ into a functor.

Note that arrows are reversed by $\mathcal{O}_X$. If $V \to U$, then $\mathcal{O}_X(U) \to \mathcal{O}_X(V)$. A functor that reverses arrows is a *contravariant* functor. The superscript $\circ$ in (6.1.1) and (6.1.4) is a customary notation to indicate that a functor is contravariant.

6.1.4. Proposition *The (extended) structure sheaf has the following* **sheaf property***:*

- *If an open subset Y of X is covered by affine open subsets $U^i = \operatorname{Spec} A_i$, then*

$$\mathcal{O}_X(Y) = \bigcap \mathcal{O}_X(U^i) \quad \big(= \bigcap A_i \big)$$

This sheaf property is especially simple because regular functions are elements of the function field. It is more complicated for $\mathcal{O}$-modules, which will be defined in the next section.

The proposition is quite simple. By definition, if f is a regular function on X, there is a covering by affine open sets U^i such that f is regular on each of them, i.e., that f is in $\mathcal{O}(U^i)$ for every i.

6.1.5. Lemma. *Let Y be an open subset of a variety X. The intersection $\bigcap \mathcal{O}_X(U^i)$ is the same for every affine open covering $\{U^i\}$ of Y.*

We prove the lemma first in the case of a covering of an affine open set by localizations.

6.1.6. Sublemma. *Let $U = \operatorname{Spec} A$ be an affine variety, and let $\{U^i\}$ be a covering of U by localizations, say $U^i = \operatorname{Spec} A_{s_i}$. Then $A = \bigcap A_{s_i}$, i.e., $\mathcal{O}(U) = \bigcap \mathcal{O}(U^i)$.*

proof. A finite subset of the set $\{U^i\}$ will cover U, so we may assume that the index set is finite.

It is clear that A is a subset of $\bigcap A_{s_i}$. Let α be an element of $\bigcap A_{s_i}$. So $\alpha = s_i^{-r} a_i$, or $s_i^r \alpha = a_i$ for some a_i in A and some integer r, and we can use the same r for every i. Because $\{U^i\}$ covers U, the elements s_i

generate the unit ideal in A, and so do their powers s_i^r. There are elements b_i in A such that $\sum b_i s_i^r = 1$. Then $\alpha = \sum b_i s_i^r \alpha = \sum b_i a_i$ is in A. □

proof of Lemma 6.1.5. Say that Y is covered by affine open sets $\{U^i\}$ and also by affine open sets $\{V^j\}$. We cover the intersections $U^i \cap V^j$ by open sets $W^{ij\nu}$ that are localizations of U^i and also localizations of V^j. Fixing i and letting j and ν vary, the set $\{W^{ij\nu}\}_{j,\nu}$ will be a covering of U^i by localizations, and the sublemma shows that $\mathcal{O}(U^i) = \bigcap_{j,\nu} \mathcal{O}(W^{ij\nu})$. Then $\bigcap_i \mathcal{O}(U^i) = \bigcap_{i,j,\nu} \mathcal{O}(W^{ij\nu})$. Similarly, $\bigcap_j \mathcal{O}(V^j) = \bigcap_{i,j,\nu} \mathcal{O}(W^{ij\nu})$. □

6.1.7. Example.

Let A denote the polynomial ring $\mathbb{C}[x, y]$, and let Y be the complement of a point p in affine space $X = \operatorname{Spec} A$. We cover Y by two localizations of X, $X_x = \operatorname{Spec} A[x^{-1}]$ and $X_y = \operatorname{Spec} A[y^{-1}]$. A regular function on Y will be regular on X_x and on X_y, so it will be in the intersection of their coordinate algebras. The intersection $A[x^{-1}] \cap A[y^{-1}]$ is A. So the sections of the structure sheaf $\mathcal{O}_X$ on Y are the same as the sections on X. They are the elements of A.

□

6.2 $\mathcal{O}$-Modules

A module on an affine variety $X = \operatorname{Spec} A$ is simply an A-module. On an arbitrary variety X, an *$\mathcal{O}_X$-module* associates an A-module to every affine open subset $U = \operatorname{Spec} A$.

6.2.1. Definition. An *$\mathcal{O}$-module* $\mathcal{M}$ on a variety X is a (contravariant) functor

$$\textit{(affines)}^\circ \xrightarrow{\mathcal{M}} \textit{(modules)}$$

such that $\mathcal{M}(U)$ is an $\mathcal{O}(U)$-module for every affine open set U, and such that, if s is a nonzero element of $\mathcal{O}(U)$, the module $\mathcal{M}(U_s)$ is the localization of $\mathcal{M}(U)$:

$$\mathcal{M}(U_s) = \mathcal{M}(U)_s$$

If $\mathcal{M}(U)$ is a finite $\mathcal{O}(U)$-module for every affine open set U, $\mathcal{M}$ is called a *finite* $\mathcal{O}$-module. A *section* of an $\mathcal{O}$-module $\mathcal{M}$ on an affine open set U is an element of $\mathcal{M}(U)$.

A *homomorphism* $\mathcal{M} \xrightarrow{\varphi} \mathcal{N}$ of $\mathcal{O}$-modules consists of homomorphisms of $\mathcal{O}(U)$-modules

$$\mathcal{M}(U) \xrightarrow{\varphi(U)} \mathcal{N}(U)$$

for each affine open subset U of X such that, if s is a nonzero element of $\mathcal{O}(U)$, the homomorphism $\varphi(U_s)$ is the localization of $\varphi(U)$.

A sequence of homomorphisms

$$\mathcal{M} \to \mathcal{N} \to \mathcal{P} \tag{6.2.2}$$

of $\mathcal{O}$-modules on a variety X is *exact* if the sequence of sections $\mathcal{M}(U) \to \mathcal{N}(U) \to \mathcal{P}(U)$ is exact for every affine open subset U of X. □

Note. When stating that $\mathcal{M}(U_s)$ **is** the localization of $\mathcal{M}(U)$, it would be more correct to say that $\mathcal{M}(U_s)$ and $\mathcal{M}(U)_s$ are canonically isomorphic. Let's not worry about this.

One example of an $\mathcal{O}$-module is the *free module* $\mathcal{O}^k$. The sections of the free module on an affine open set U are the elements of the free $\mathcal{O}(U)$-module $\mathcal{O}(U)^k$. In particular, $\mathcal{O}$ can be considered as an $\mathcal{O}$-module.

The *kernel, image*, and *cokernel* of a homomorphism $\mathcal{M} \xrightarrow{\varphi} \mathcal{N}$ are among the operations that can be made on $\mathcal{O}$-modules. The kernel $\mathcal{K}$ of φ is the $\mathcal{O}$-module defined by $\mathcal{K}(U) = \ker(\mathcal{M}(U) \xrightarrow{\varphi(U)} \mathcal{N}(U))$ for every affine open set U, and the image and cokernel are defined analogously. Many operations, such as these, are compatible with localization.

At first glance, the definition of $\mathcal{O}$-module seems rather complicated. However, when a module has a natural definition, the data involved in the definition take care of themselves. This will become clear as we go along.

6.3 The Sheaf Property

In this section, we extend an $\mathcal{O}$-module $\mathcal{M}$ on a variety X to a functor *(opens)*$^\circ \xrightarrow{\widetilde{\mathcal{M}}}$ *(modules)* on all open subsets of X, such that $\widetilde{\mathcal{M}}(Y)$ is an $\mathcal{O}(Y)$-module for every open subset Y, and when U is an affine open set, $\widetilde{\mathcal{M}}(U) = \mathcal{M}(U)$.

The tilde is used for clarity here. We will drop it when we have finished with the discussion, and use the same notation $\mathcal{M}$ for the functor on *(affines)* and for its extension to *(opens)*.

6.3.1. Terminology. Let *(opens)*$^\circ \xrightarrow{\widetilde{\mathcal{M}}}$ *(modules)* be a functor and let U be an open subset. An element of $\widetilde{\mathcal{M}}(U)$ is a *section* of $\widetilde{\mathcal{M}}$ on U. If $V \xrightarrow{j} U$ is an inclusion of open subsets, the associated homomorphism $\widetilde{\mathcal{M}}(U) \to \widetilde{\mathcal{M}}(V)$ is the *restriction* from U to V.

The restriction to V of a section m on U may be denoted by $j^\circ m$. However, the operation of restriction occurs very often. Because of this, we often abbreviate, using the same symbol m for a section and for its restriction. Also, if an open set V is contained in two open sets U and U', and if m, m' are sections of $\widetilde{\mathcal{M}}$ on U and U', respectively, we may say that m and m' are *equal on* V if their restrictions to V are equal. $\square$

6.3.2. Theorem. *An $\mathcal{O}$-module $\mathcal{M}$ extends uniquely to a functor*

$$\textit{(opens)}^\circ \xrightarrow{\widetilde{\mathcal{M}}} \textit{(modules)}$$

that has the sheaf property described below. Moreover, for every open set U, $\widetilde{\mathcal{M}}(U)$ is an $\mathcal{O}(U)$-module, and for every inclusion $V \to U$ of nonempty open sets, the map $\widetilde{\mathcal{M}}(U) \to \widetilde{\mathcal{M}}(V)$ is compatible with scalar multiplication in this sense:

Let m be a section of $\widetilde{\mathcal{M}}$ on U, let α be a regular function on U, and let m' and α' denote the restrictions to V. The restriction of αm is $\alpha' m'$.

In order not to break up the discussion, we have put the proof of this theorem into Section 6.7 at the end of this chapter.

(6.3.3) the sheaf property

The sheaf property is the key requirement that determines the extension of an $\mathcal{O}$-module $\mathcal{M}$ to a functor $\widetilde{\mathcal{M}}$ on *(opens)*.

Let Y be an open subset of X, and let $\{U^i\}$ be a covering of Y by affine open sets. The intersections $U^{ij} = U^i \cap U^j$ are also affine open sets, so $\mathcal{M}(U^i)$ and $\mathcal{M}(U^{ij})$ are defined. The sheaf property asserts that an element $m \in \widetilde{\mathcal{M}}(Y)$ corresponds to a set of elements m_i in $\mathcal{M}(U^i)$ such that the restrictions of m_j and m_i to U^{ij} are equal.

If the affine open subsets U^i are indexed by $i = 1, ..., n$, the sheaf property asserts that an element of $\widetilde{\mathcal{M}}(Y)$ is determined by a vector $(m_1, ..., m_n)$ with m_i in $\mathcal{M}(U^i)$, such that the restrictions of m_i and m_j to U^{ij} are equal. This means that $\widetilde{\mathcal{M}}(Y)$ is the kernel of the map

$$\prod_i \mathcal{M}(U^i) \xrightarrow{\beta} \prod_{i,j} \mathcal{M}(U^{ij}) \tag{6.3.4}$$

that sends the vector $(m_1, ..., m_n)$ to the $n \times n$ matrix (z_{ij}), where z_{ij} is the difference $m_j - m_i$ of the restrictions of m_j and m_i to U^{ij}. The analogous description is true when the index set is infinite.

In short, the sheaf property tells us that sections of $\widetilde{\mathcal{M}}$ are determined locally: A section on an open set Y is determined by its restrictions to the open subsets U^i of an affine covering of Y.

Note. The morphisms $U^{ij} \to U^i$ needn't be localizations, and if not the restriction maps $\mathcal{M}(U^i) \to \mathcal{M}(U^{ij})$ aren't a part of the structure of an $\mathcal{O}$-module. We need a definition of the restriction map for an arbitrary inclusion $V \to U$ of affine open subsets. This point will be taken care of by the proof of Theorem 6.3.2. (See Step 2 in Section 6.7.) Let's not worry about it here. $\square$

We drop the tilde now.

The next corollary follows from Theorem 6.3.2.

6.3.5. Corollary. *Let $\{U^i\}$ be an affine open covering of a variety X.*
(i) *An $\mathcal{O}$-module $\mathcal{M}$ is zero if and only if $\mathcal{M}(U^i) = 0$ for every i.*
(ii) *A homomorphism $\mathcal{M} \xrightarrow{\varphi} \mathcal{N}$ of $\mathcal{O}$-modules is injective, surjective, or bijective if and only if the maps $\mathcal{M}(U^i) \xrightarrow{\varphi(U^i)} \mathcal{N}(U^i)$ are injective, surjective, or bijective, respectively, for every i.*

proof. **(i)** Let V be any open subset of X. We can cover the intersections $V \cap U^i$ by affine open sets $V^{i\nu}$ that are localizations of U^i, and these sets, taken together, cover V. If $\mathcal{M}(U^i) = 0$, then the localizations $\mathcal{M}(V^{i\nu})$ are zero too. The sheaf property shows that the map $\mathcal{M}(V) \to \prod \mathcal{M}(V^{i\nu})$ is injective, and therefore $\mathcal{M}(V) = 0$.

(ii) This follows from **(i)** because a homomorphism φ is injective or surjective if and only if its kernel or its cokernel is zero. □

(6.3.6) families of open sets

It is convenient to have a compact notation for the sheaf property. For this, one can use symbols to represent families of open sets. Say that $\mathbb{U}$ and $\mathbb{V}$ represent families of open sets $\{U^i\}$ and $\{V^\nu\}$, respectively. A *morphism* of families $\mathbb{V} \to \mathbb{U}$ consists of a morphism from each V^ν to one of the subsets U^i. Such a morphism will be given by a map $\nu \rightsquigarrow i_\nu$ of index sets, such that $V^\nu \subset U^{i_\nu}$.

There may be more than one morphism $\mathbb{V} \to \mathbb{U}$, because a subset V^ν may be contained in more than one of the subsets U^i. To define a morphism, one must make a choice among those subsets. For example, let $\mathbb{U} = \{U^i\}$ be a family of open sets, and let V be another open set. There is a morphism $V \to \mathbb{U}$ that sends V to U^i whenever $V \subset U^i$. In the other direction, there is a unique morphism $\mathbb{U} \to V$ provided that $U^i \subset V$ for all i.

A functor *(opens)*$^\circ \xrightarrow{\mathcal{M}}$ *(modules)* can be extended to families $\mathbb{U} = \{U^i\}$ by defining

$$\mathcal{M}(\mathbb{U}) = \prod \mathcal{M}(U^i). \tag{6.3.7}$$

Then a morphism of families $\mathbb{V} \xrightarrow{f} \mathbb{U}$ defines a map $\mathcal{M}(\mathbb{V}) \xleftarrow{f^\circ} \mathcal{M}(\mathbb{U})$ in a way that is fairly obvious, though notation for it is clumsy. Say that f is given by a map $\nu \rightsquigarrow i_\nu$ of index sets, with $V^\nu \to U^{i_\nu}$. A section of $\mathcal{M}$ on $\mathbb{U}$, an element of $\mathcal{M}(\mathbb{U})$, can be thought of as a vector (u_i) with $u_i \in \mathcal{M}(U^i)$, and a section of $\mathcal{M}(\mathbb{V})$ as a vector (v_ν) with $v_\nu \in \mathcal{M}(V^\nu)$. If the map f° sends $(u_i) \to (v_\nu)$, then v_ν is the restriction of u_{i_ν} to V^ν.

We can write the sheaf property in terms of families of open sets. Let $\mathbb{U}_0 = \{U^i\}$ be an affine open covering of an open set Y, and let $\mathbb{U}_1$ denote the family $\{U^{ij}\}$ of intersections: $U^{ij} = U^i \cap U^j$, which are also affine. . Then we have a morphism $\mathbb{U}_0 \to Y$, and the two sets of inclusions

$$U^{ij} \subset U^i \quad \text{and} \quad U^{ij} \subset U^j$$

define two morphisms of families $\mathbb{U}_1 \xrightarrow{d_0, d_1} \mathbb{U}_0$ of affine open sets, $U^{ij} \xrightarrow{d_0} U^j$ and $Uij \xrightarrow{d_1} U^i$. The two composed morphisms $\mathbb{U}_1 \xrightarrow{d_i} \mathbb{U}_0 \to Y$ are equal. These morphisms form what we all a *covering diagram*

$$Y \longleftarrow \mathbb{U}_0 \leftleftarrows \mathbb{U}_1 \tag{6.3.8}$$

When we apply a functor *(opens)* $\xrightarrow{\mathcal{M}}$ *(modules)* to this diagram, we obtain a sequence

$$0 \to \mathcal{M}(Y) \xrightarrow{\alpha_{\mathbb{U}}} \mathcal{M}(\mathbb{U}_0) \xrightarrow{\beta_{\mathbb{U}}} \mathcal{M}(\mathbb{U}_1) \tag{6.3.9}$$

where $\alpha_{\mathbb{U}}$ is the restriction map and $\beta_{\mathbb{U}}$ is the difference $\mathcal{M}(d_0) - \mathcal{M}(d_1)$ of the maps induced by the two morphisms $\mathbb{U}_1 \rightrightarrows \mathbb{U}_0$. The sheaf property for the covering $\mathbb{U}_0$ of Y is the assertion that this sequence is exact, which means that $\alpha_{\mathbb{U}}$ is injective, and that its image is the kernel of $\beta_{\mathbb{U}}$.

6.3.10. Note. One can suppose that the open sets U^i that make a covering are distinct. However, the intersections won't be distinct, because $U^{ij} = U^{ji}$. Also, $U^{ii} = U^i$. These coincidences lead to redundancy in the statement (6.3.9) of the sheaf property. If the indices are $i = 1, ..., k$, we only need to look at intersections U^{ij} with $i < j$. The product $\mathcal{M}(\mathbb{U}_1) = \prod_{i,j} \mathcal{M}(U^{ij})$ that appears in the sheaf property can be replaced by the product $\prod_{i<j} \mathcal{M}(U^{ij})$ with increasing pairs of indices. For instance, suppose that an open set Y is covered by two affine open sets U and V. Then, with $\mathbb{U}_0 = \{U, V\}$, the sheaf property is the exact sequence

$$0 \to \mathcal{M}(Y) \xrightarrow{\alpha} \mathcal{M}(U)\times\mathcal{M}(V) \xrightarrow{\beta} \mathcal{M}(U\cap U)\times\mathcal{M}(U\cap V)\times\mathcal{M}(V\cap U)\times\mathcal{M}(V\cap V)$$

is equivalent with the exact sequence

(6.3.11) $$0 \to \mathcal{M}(\to \mathcal{M}(U)\times\mathcal{M}(V) \xrightarrow{+,-} \mathcal{M}(U\cap V)$$ □

6.3.12. Example.

We go back to Proposition 6.4.26, which describes the correspondence between an $\mathcal{O}$-module $\mathcal{M}$ on an affine variety $X = \operatorname{Spec} A$ and an A-module M. Namely, if $U = \operatorname{Spec} B$ is an affine open subset of X, then $\mathcal{M}(U) = B \otimes_A M$. The next example shows that, when a subset U isn't affine, defining $\mathcal{M}(U) = B \otimes_A M$ may be wrong.

Let X be the affine plane $\operatorname{Spec} A$, $\;A = \mathbb{C}[x, y]$, let U be the complement of the origin in X, and let M be the A-module A/yA. This module can be identified with $\mathbb{C}[x]$, which becomes an A-module when scalar multiplication by y is defined to be zero. Here $\mathcal{O}(U) = \mathcal{O}(X) = A$ (6.1.7). If we followed the method used for affine open sets, we would set $\mathcal{M}(U) = A \otimes_A M = \mathbb{C}[x]$.

To identify $\mathcal{M}(U)$ correctly, we cover U by the two affine open sets $U_x = \operatorname{Spec} A[x^{-1}]$ and $U_y = \operatorname{Spec} A[y^{-1}]$. Then $\mathcal{M}(U_x) = M[x^{-1}]$ while $\mathcal{M}(U_y) = 0$. The sheaf property of $\mathcal{M}$ shows that $\mathcal{M}(U) \approx \mathcal{M}(U_x) = M[x^{-1}] = \mathbb{C}[x, x^{-1}]$. □

We have tacitly assumed that our open sets aren't empty. The next lemma takes care of the empty set.

6.3.13. Lemma. *The only section of an $\mathcal{O}$-module $\mathcal{M}$ on the empty set is the zero section: $\mathcal{M}(\emptyset) = \{0\}$. In particular, $\mathcal{O}(\emptyset)$ is the zero ring.*

proof. This follows from the sheaf property. The empty set is covered by the empty covering, the covering indexed by the empty set. Therefore $\mathcal{M}(\emptyset)$ is contained in an empty product. Since we want the empty product and $\mathcal{M}(\emptyset)$ to be modules, we have no choice but to set them equal to $\{0\}$.

If you find this reasoning pedantic, you can take $\mathcal{M}(\emptyset) = \{0\}$ as an axiom. □

(6.3.14) the coherence property

In addition to the sheaf property, an $\mathcal{O}$-module on a variety X has a property called *coherence*.

6.3.15. Proposition. *(the coherence property) Let Y be an open subset of a variety X, let s be a nonzero regular function on Y, and let $\mathcal{M}$ be an $\mathcal{O}_X$-module. Then $\mathcal{M}(Y_s)$ is the localization $\mathcal{M}(Y)_s$ of $\mathcal{M}(Y)$.*

Compatibility with localization is a requirement for an $\mathcal{O}$-module when Y is affine. The coherence property is an extension to all open subsets.

proof of Proposition 6.3.15. Let $\mathbb{U}_0 = \{U^i\}$ be a family of affine open sets that covers an open set Y. The intersections U^{ij} will be affine open sets too. We inspect the covering diagram $Y \leftarrow \mathbb{U}_0 \leftleftarrows \mathbb{U}_1$. If s is a nonzero regular function on Y, the localization of this diagram forms a covering diagram $Y_s \leftarrow \mathbb{U}_{0,s} \leftleftarrows \mathbb{U}_{1,s}$, in which $\mathbb{U}_{0,s} = \{U_s^i\}$ is an affine covering of Y_s. Therefore $\mathcal{M}(\mathbb{U}_0)_s \approx \mathcal{M}(\mathbb{U}_{0,s})$. The sheaf property gives us exact sequences

$$0 \to \mathcal{M}(Y) \to \mathcal{M}(\mathbb{U}_0) \to \mathcal{M}(\mathbb{U}_1) \qquad \text{and} \qquad 0 \to \mathcal{M}(Y_s) \to \mathcal{M}(\mathbb{U}_{0,s}) \to \mathcal{M}(\mathbb{U}_{1,s})$$

and the localization of the first sequence maps to the second one:

$$\begin{array}{ccccccc} 0 & \longrightarrow & \mathcal{M}(Y)_s & \longrightarrow & \mathcal{M}(\mathbb{U}_0)_s & \longrightarrow & \mathcal{M}(\mathbb{U}_1)_s \\ & & \big\downarrow{\scriptstyle a} & & \big\downarrow{\scriptstyle b} & & \big\downarrow{\scriptstyle c} \\ 0 & \longrightarrow & \mathcal{M}(Y_s) & \longrightarrow & \mathcal{M}(\mathbb{U}_{0,s}) & \longrightarrow & \mathcal{M}(\mathbb{U}_{1,s}) \end{array}$$

The bottom row is exact, and since localization is an exact operation, the top row of the diagram is exact too. Since $\mathbb{U}_0$ and $\mathbb{U}_1$ are families of affine open sets, the vertical arrows b and c are bijections. Therefore a is a bijection. This is the coherence property. □

6.4 Some $\mathcal{O}$-Modules

6.4.1. modules on a point

Let's denote a point, the affine variety $\operatorname{Spec}\mathbb{C}$, by p. The point has only one nonempty open set: the whole space p, and $\mathcal{O}_p(p) = \mathbb{C}$. Let $\mathcal{M}$ be an $\mathcal{O}_p$-module. The space of global sections $\mathcal{M}(p)$ is an $\mathcal{O}_p(p)$-module, a complex vector space. To define $\mathcal{M}$, that vector space can be assigned arbitrarily. One may say that a module on the point *is* a complex vector space. □

6.4.2. the residue field module κ_p.

Let p be a point of a variety X. A residue field module κ_p is defined as follows: If U is an affine open subset of X that contains p, then $\mathcal{O}(U)$ has a residue field $k(p)$ at p, and $\kappa_p(U) = k(p)$. If U doesn't contain p, then $\kappa_p(U) = 0$.

6.4.3. ideals.

An ideal $\mathcal{I}$ of the structure sheaf is an $\mathcal{O}$-submodule of $\mathcal{O}$.

Let p be a point of a variety X. The *maximal ideal* at p, which we denote by $\mathfrak{m}_p$, is the ideal of $\mathcal{O}$ defined as follows: If an affine open subset U contains p, its coordinate algebra $\mathcal{O}(U)$ will have a maximal ideal consisting of the elements that vanish at p. That maximal ideal is the module of sections $\mathfrak{m}_p(U)$. If U doesn't contain p, then $\mathfrak{m}_p(U) = \mathcal{O}(U)$.

When $\mathcal{I}$ is an ideal of $\mathcal{O}$, we denote by $V_X(\mathcal{I})$ the closed set of points p such that $\mathcal{I} \subset \mathfrak{m}_p$ – such that all elements of $\mathcal{I}$ vanish at p.

6.4.4. examples of homomorphisms

(i) There is a homomorphism of $\mathcal{O}$-modules $\mathcal{O} \to \kappa_p$, whose kernel is the maximal ideal $\mathfrak{m}_p$.

(ii) Homomorphisms $\mathcal{O}^n \to \mathcal{O}^m$ of free $\mathcal{O}$-modules correspond to $m \times n$-matrices of global sections of $\mathcal{O}$.

(iii) Let $\mathcal{M}$ be an $\mathcal{O}$-module. Then $\mathcal{O}$-module homomorphisms $\mathcal{O} \xrightarrow{\varphi} \mathcal{M}$ correspond bijectively to global sections of $\mathcal{M}$.

This is analogous to the fact that, when M is a module over a ring A, A-module homomorphisms $A \to M$ correspond to elements of M. To be explicit: If m is a global section of $\mathcal{M}$, the homomorphism $\mathcal{O}(U) \xrightarrow{\varphi} \mathcal{M}(U)$ is multiplication by the restriction of m to U.

(iv) If f is a global section of $\mathcal{O}$, scalar multiplication by f defines a homomorphism $\mathcal{M} \xrightarrow{f} \mathcal{M}$.

6.4.5. kernel

As we have remarked, many operations that one makes on modules over a ring are compatible with localization, and therefore can be made on $\mathcal{O}$-modules. However, when applied to sections over non-affine open sets the operations are almost *never* compatible with localization. One important exception is the kernel of a homomorphism.

6.4.6. Proposition. *Let X be a variety, and let $0 \to \mathcal{K} \to \mathcal{M} \to \mathcal{P}$ be an exact sequence of $\mathcal{O}$-modules. For every open subset Y of X, the sequence of sections*

(6.4.7) $$0 \to \mathcal{K}(Y) \to \mathcal{M}(Y) \to \mathcal{N}(Y)$$

is exact.

proof. We choose a covering diagram $Y \longleftarrow \mathbb{U}_0 \leftleftarrows \mathbb{U}_1$, and we inspect the diagram

$$\begin{array}{ccccccc} 0 & \longrightarrow & \mathcal{K}(\mathbb{U}_0) & \longrightarrow & \mathcal{M}(\mathbb{U}_0) & \longrightarrow & \mathcal{N}(\mathbb{U}_0) \\ & & \downarrow & & \downarrow & & \downarrow \\ 0 & \longrightarrow & \mathcal{K}(\mathbb{U}_1) & \longrightarrow & \mathcal{M}(\mathbb{U}_1) & \longrightarrow & \mathcal{N}(\mathbb{U}_1) \end{array}$$

where the vertical maps are the ones described in (6.3.9). The rows are exact because $\mathbb{U}_0$ and $\mathbb{U}_1$ are families of affines, and the sheaf property asserts that the kernels of the vertical maps form the sequence (6.4.7), which is exact because taking kernels is a left exact operation. □

The section functor isn't right exact. When

$$0 \to \mathcal{K} \to \mathcal{M} \to \mathcal{N} \to 0$$

is a short exact sequence of $\mathcal{O}$-modules and Y is a non-affine open set, the sequence (6.4.7) may fail to be exact when a zero is added on the right. There is an example below. Cohomology, which will be discussed in the next chapters, is a substitute for right exactness.

6.4.8. modules on the projective line

The projective line $\mathbb{P}^1$ is covered by the standard open sets $\mathbb{U}^0$ and $\mathbb{U}^1$, and the intersection $\mathbb{U}^{01} = \mathbb{U}^0 \cap \mathbb{U}^1$ is a localization of $\mathbb{U}^0$ and of $\mathbb{U}^1$. The coordinate algebras of these affine open sets are $\mathbb{C}[u] = A_0$ and $\mathbb{C}[v] = A_1$, respectively, with $v = u^{-1}$, and $\mathcal{O}(\mathbb{U}^{01}) = \mathbb{C}[u, u^{-1}] = A_{01}$. The algebra A_{01} is the Laurent polynomial ring, whose elements are (finite) combinations of powers of u, negative powers included. The sheaf property asserts that a global section of $\mathcal{O}$ is determined by polynomials $f(u)$ in A_0 and $g(v)$ in A_1 such that $f(u) = g(u^{-1})$ in A_{01}. The only such polynomials f, g are the constants. The constants are the only rational functions that are regular everywhere on $\mathbb{P}^1$. I think we knew this.

If $\mathcal{M}$ is an $\mathcal{O}$-module, $\mathcal{M}(\mathbb{U}^0) = M_0$ and $\mathcal{M}(\mathbb{U}^1) = M_1$ will be modules over the algebras A_0 and A_1, and the A_{01}-module $\mathcal{M}(\mathbb{U}^{01}) = M_{01}$ can be obtained by localizing M_0 and also by localizing M_1: $M_0[u^{-1}] \approx M_{01} \approx M_1[v^{-1}]$. A global section of $\mathcal{M}$ is determined by a pair of elements m_1, m_2 in M_1, M_2 that become equal in the common localization M_{01}.

Suppose that M_0 and M_1 are free modules of rank r over A_0 and A_1,. Then M_{01} will be a free A_{01}-module of rank r. A basis $\mathbf{B}_0$ of the free A_0-module M_0 will also be a basis of the A_{01}-module M_{01}, and a basis $\mathbf{B_1}$ of M_1 will be a basis of M_{01}. When regarded as bases of M_{01}, $\mathbf{B_0}$ and $\mathbf{B_1}$ will be related by an $r \times r$ invertible A_{01}-matrix P, and that matrix determines $\mathcal{M}$ up to isomorphism. When $r = 1$, P will be an invertible 1×1 matrix in the Laurent polynomial ring A_{01} – a unit of that ring. The units in A_{01} are scalar multiples of powers of u. Since the scalar can be absorbed into one of the bases, an $\mathcal{O}$-module of rank 1 is determined, up to isomorphism, by a power of u. It is one of the *twisting modules* that will be described in Section 6.6.

The Birkhoff-Grothendieck Theorem, which will be proved in Chapter 8, describes the $\mathcal{O}$-modules on the projective line whose sections on $\mathbb{U}^0$ and on $\mathbb{U}^1$ are free, as direct sums of free $\mathcal{O}$-modules of rank one. This means that by changing the bases $\mathbf{B}_i$, one can diagonalize the matrix P. Such a change of basis is given by an invertible A_0-matrix Q_0 and an invertible A_1-matrix Q_1, respectively. In down-to-Earth terms, the Birkhoff-Grothendieck Theorem asserts that, for any invertible A_{01}-matrix P, there exist an invertible A_0-matrix Q_0 and an invertible A_1-matrix Q_1 , such that $Q_0^{-1} P Q_1$ is diagonal. This can be proved by matrix operations. □

6.4.9. tensor products

Tensor products are compatible with localization. If M and N are modules over a domain A and s is a nonzero element of A, the canonical map $(M \otimes_A N)_s \to M_s \otimes_{A_s} N_s$ is an isomorphism. Therefore the tensor product $\mathcal{M} \otimes_{\mathcal{O}} \mathcal{N}$ of $\mathcal{O}$-modules $\mathcal{M}$ and $\mathcal{N}$ can be defined.

Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{O}$-modules, let $\mathcal{M} \otimes_{\mathcal{O}} \mathcal{N}$ be the tensor product module, and let V be an open subset of X. For every open set V, there is a canonical map

$$\mathcal{M}(V) \otimes_{\mathcal{O}(V)} \mathcal{N}(V) \to [\mathcal{M} \otimes_{\mathcal{O}} \mathcal{N}](V) \tag{6.4.10}$$

By definition of the tensor product module, $\mathcal{M}(V) \otimes_{\mathcal{O}(V)} \mathcal{N}(V) = [\mathcal{M} \otimes_{\mathcal{O}} \mathcal{N}](V)$ when V is affine. For arbitrary V, we cover by a family $\mathbb{U}_0$ of affine open sets. The family $\mathbb{U}_1$ of intersections also consists of affine open sets. One forms a diagram

$$\begin{array}{ccccccc} & & \mathcal{M}(V)\otimes_{\mathcal{O}(V)}\mathcal{N}(V) & \longrightarrow & \mathcal{M}(\mathbb{U}_0)\otimes_{\mathcal{O}(\mathbb{U}_0)}\mathcal{N}(\mathbb{U}_0) & \longrightarrow & \mathcal{M}(\mathbb{U}_1)\otimes_{\mathcal{O}(\mathbb{U}_1)}\mathcal{N}(\mathbb{U}_1) \\ & & \downarrow{\scriptstyle a} & & \downarrow{\scriptstyle b} & & \downarrow{\scriptstyle c} \\ 0 & \longrightarrow & [\mathcal{M}\otimes_{\mathcal{O}}\mathcal{N}](V) & \longrightarrow & [\mathcal{M}\otimes_{\mathcal{O}}\mathcal{N}](\mathbb{U}_0) & \longrightarrow & [\mathcal{M}\otimes_{\mathcal{O}}\mathcal{N}](\mathbb{U}_1) \end{array}$$

The composition of the two arrows in the top row is zero, the bottom row is exact, and the vertical maps b and c are equalities. The map a is induced by the diagram. It is a bijective map when V is affine, but when V isn't affine, it may fail to be either injective or surjective.

6.4.11. Examples. (i) Let p and q be distinct points of the projective line X, and let κ_p and κ_q be the residure field modules on X. Then $\kappa_p(X)=\kappa_q(X)=\mathbb{C}$, so $\kappa_p(X)\otimes_{\mathcal{O}(X)}\kappa_q(X)\approx\mathbb{C}\otimes_{\mathbb{C}}\mathbb{C}=\mathbb{C}$. But $\kappa_p\otimes_{\mathcal{O}}\kappa_q=0$. The canonical map (6.4.10) is the zero map. It isn't injective.

(ii) Let p a point of a variety X, and let $\mathfrak{m}_p$ and κ_p be the maximal ideal and residue field modules at p. There is an exact sequence of $\mathcal{O}$-modules

(6.4.12) $$0\to\mathfrak{m}_p\to\mathcal{O}\xrightarrow{\pi_p}\kappa_p\to 0$$

In this case, the sequence of global sections is exact.

(iii) Let p_0 and p_1 be the points $(1,0)$ and $(0,1)$ of the projective line $\mathbb{P}^1$. We form a homomorphism

$$\mathfrak{m}_{p_0}\times\mathfrak{m}_{p_1}\xrightarrow{\varphi}\mathcal{O}$$

φ being the map $(a,b)\mapsto b-a$. On the open set $\mathbb{U}^0$, $\mathfrak{m}_{p_1}\to\mathcal{O}$ is bijective and therefore surjective. Similarly, $\mathfrak{m}_{p_0}\to\mathcal{O}$ is surjective on $\mathbb{U}^1$. Therefore φ is surjective. The only global section of $\mathfrak{m}_{p_0}\times\mathfrak{m}_{p_1}$ is zero, while $\mathcal{O}$ has the nonzero global section 1. So the map φ isn't surjective on global sections. □

6.4.13. the function field module

Let F be the function field of a variety X. The module of sections of the function field module $\mathcal{F}$ on any nonempty open set is the field F. This is an $\mathcal{O}$-module. It is called a *constant $\mathcal{O}$-module* because the modules of sections $\mathcal{F}(U)$ are the same for every nonempty open set U. It isn't a finite module unless X is a point.

Tensoring with the function field module: Let $\mathcal{M}$ be an $\mathcal{O}$-module on a variety X, and let $\mathcal{F}$ be the function field module. Then $\mathcal{M}\otimes_{\mathcal{O}}\mathcal{F}$ is a constant $\mathcal{O}$-module whose sections on any affine open set U form an F-vector space (that might be zero).

(6.4.14) annihilators

Let A be a ring, and let m be an element of an A-module M. The *annihilator* I of an element m of M is the set of elements α of A such that $\alpha m=0$. This is an ideal of A that we may denote by $\operatorname{ann}(m)$.

The annihilator of the A-module M is the set of elements of A such that $aM=0$. This annihilator is also an ideal, but we will have little use for it.

6.4.15. Lemma. *Let I be the annihilator of an element m of M, and let s be a nonzero element of A. The annihilator of the image of m in the localized module M_s is the localized ideal I_s* □

This allows us to extend the concept of annihilator to sections of a finite $\mathcal{O}$-module on a variety X.

(6.4.16) maximal annihilators

Let m be an element of a module M over a noetherian ring A, and let $I=\operatorname{ann}(m)$: $I=\{a\in A\,|\,am=0\}$. The same ideal I will annihilate every element of the submodule Am spanned by m.

Let $\mathcal{S}$ be the set whose members are the annihilators of nonzero elements of M. A *maximal annihilator* is a maximal member of $\mathcal{S}$. Because A is noetherian, the annihilator of any nonzero element will be contained in a maximal annihilator.

6.4.17. Proposition. *Let M be a finite, module over a noetherian ring A, not the zero module..*

(i) *Let m be a nonzero element of M whose annihilator P is a maximal annihilator, and let b be an element of A. Then P is also the annihilator of bm, if $bm \neq 0$.*

(ii) *The maximal annihilators are prime ideals.*

(iii) *Let P_i be distinct maximal annihilators, let m_i be elements of M whose annihilators are P_i, and let $N_i = Am_i$. The submodules N_i of M are independent, i.e., their sum $\sum N_i$ is the direct sum $\bigoplus N_i$.*

(iv) *The set of maximal annihilators is finite and nonempty.*

proof. Let's denote by $\operatorname{ann}(m)$ the annihilator of an element m of M.

(i) An element a that annihilates m also annihilates bm. So $\operatorname{ann}(m) \subset \operatorname{ann}(bm)$. If $\operatorname{ann}(m)$ is maximal and $bm \neq 0$, then $\operatorname{ann}(m) = \operatorname{ann}(bm)$.

(ii) Let $P = \operatorname{ann}(m)$ be a maximal annihilator, and let a, b be elementsof A such that $ab \in P$. If $b \notin P$, then $bm \neq 0$ but $abm = 0$. So $a \in \operatorname{ann}(bm) = P$.

(iii) We must show that if n_i are elements of N_i such that $n_1 + \cdots + n_k = 0$, then $n_i = 0$ or all i. We use induction on k. Let a be a nonzero element of the annihilator P_k of N_k. Then $an_1 + \cdots + an_{k-1} + 0 = 0$. By induction $an_i = 0$ for all i. Therefore a is in P_i for all i. Unless $k = 1$, this contradicts the assumption that P_i are distinct maximal annihilators.

(iv) The submodules N_i are nonzero and independent. Since $\bigoplus N_i$ is a submodule of the finite module A and since A is noetherian, $\bigoplus N_i$ is a finite module. So there can be only finitely many indices i. □

6.4.18. Corollary. *Let M be a finite module over a noetherian domain A, and let s be an element of A that isn't contained in any of the maximal annihilators of M. The multiplication map $M \xrightarrow{s} M$ is injective, and therefore the map from M to its localization M_s is injective.* □

6.4.19. Support Let M be a finite module over a finite-type domain A and $X = \operatorname{Spec} A$. The *support* of M is the locus $C = V_X(I)$ of zeros of its annihilator I in X. The support is a closed subset of X. ### , and the support of the localization M_s is the intersection $C_s = C \cap X_s$.

The *support* of a finite $\mathcal{O}_X$-module is the closed subset $V_X(\mathcal{I})$ of points such that $\mathcal{I} \subset \mathfrak{m}_p$.

If M is a finite module over a finite-type domain A and s is a nonzero element of A, the annihilator of the localized module M_s is the localization I_s of the annihilator I of M, and the support of the localization M_s is the intersection $C_s = C \cap X_s$.

For example, the support of the residue field module κ_p is the point p. The support of the maximal ideal $\mathfrak{m}_p$ at p is the whole variety X.

(6.4.20) $\mathcal{O}$-modules with support of dimension zero

6.4.21. Proposition. *Let $\mathcal{M}$ be a finite $\mathcal{O}$-module on a variety X.*

(i) *Suppose that the support of $\mathcal{M}$ is a single point p, let $M = \mathcal{M}(X)$, and let U be an affine open subset of X. If U contains p, then $\mathcal{M}(U) = M$, and if U doesn't contain p, then $\mathcal{M}(U) = 0$.*

(ii) *(Chinese Remainder Theorem) If the support of $\mathcal{M}$ is a finite set $\{p_1, ..., p_k\}$, then $\mathcal{M}$ is the direct sum $\mathcal{M}_1 \oplus \cdots \oplus \mathcal{M}_k$ of $\mathcal{O}$-modules supported at the points p_i.*

proof. **(i)** Let $\mathcal{I}$ be the annihilator of $\mathcal{M}$. The locus $V_X(\mathcal{I})$ is p. If p isn't contained in U, then when we restrict $\mathcal{M}$ to U, we obtain an $\mathcal{O}_U$-module whose support is empty. Therefore the restriction to U is the zero module.

Next, suppose that p is contained in U, and let V denote the complement of p in X. We cover X by a set $\{U^i\}$ of affine open sets with $U = U^1$, and such that $U^i \subset V$ if $i > 1$. By what has been shown, $\mathcal{M}(U^i) = 0$if $i > 0$ and $\mathcal{M}(U^{ij}) = 0$ if $j \neq i$. The sheaf axiom for this covering shows that $\mathcal{M}(X) \approx \mathcal{M}(U)$.

(ii) This follows from the ordinary Chinese Remainder Theorem. □

(6.4.22) limits of $\mathcal{O}$-modules

6.4.23. A *directed set* $M_\bullet$ is a sequence of maps of sets $M_0 \to M_1 \to M_2 \to \cdots$. Its *limit* $\varinjlim M_\bullet$ is the set of equivalence classes on the union $\bigcup M_k$, the equivalence relation being that elements m in M_i and m' in M_j are equivalent if they have the same image in M_n when n is sufficiently large. Any element of $\varinjlim M_\bullet$ will be represented by an element of M_i for some i.

6.4.24. Example. Let $R = \mathbb{C}[x]$ and let $\mathfrak{m}$ be the maximal ideal xR. Repeated multiplication by x defines a directed set

$$R \xrightarrow{x} R \xrightarrow{x} R \xrightarrow{x} R \cdots$$

whose limit is isomorphic to the Laurent Polynomial Ring $R[x^{-1}] = \mathbb{C}[x, x^{-1}]$. Proving this is a simple exercise. □

A *directed set* of $\mathcal{O}$-modules on a variety X is a sequence $\mathcal{M}_\bullet = \{\mathcal{M}_0 \to \mathcal{M}_1 \to \mathcal{M}_2 \to \cdots\}$ of homomorphisms of $\mathcal{O}$-modules. For every affine open set U, the $\mathcal{O}(U)$-modules $\mathcal{M}_n(U)$ form a directed set, as defined in (6.4.23). The *direct limit* $\varinjlim \mathcal{M}_\bullet$ is defined simply, by takin the limit for each affine open set: $[\varinjlim \mathcal{M}_\bullet](U) = \varinjlim [\mathcal{M}_\bullet(U)]$. This limit operation is compatible with localization, so $\varinjlim \mathcal{M}_\bullet$ is an $\mathcal{O}$-module.

6.4.25. Lemma. (i) *The limit operation is exact. If $\mathcal{M}_\bullet \to \mathcal{N}_\bullet \to \mathcal{P}_\bullet$ is an exact sequence of directed sets of $\mathcal{O}$-modules, the limits form an exact sequence.*

(ii) *Tensor products are compatible with limits: If $\mathcal{N}_\bullet$ is a directed set of $\mathcal{O}$-modules and $\mathcal{M}$ is another $\mathcal{O}$-module, then $\varinjlim [\mathcal{M} \otimes_{\mathcal{O}} \mathcal{N}_\bullet] \approx \mathcal{M} \otimes_{\mathcal{O}} [\varinjlim \mathcal{N}_\bullet]$.*

6.4.26. Proposition. *Let $X = \operatorname{Spec} A$ be an affine variety. Sending an $\mathcal{O}$-module $\mathcal{M}$ to the A-module $\mathcal{M}(X)$ of its global sections defines a bijective correspondence between $\mathcal{O}$-modules and A-modules.*

proof. We must invert the functor $\mathcal{O}$-*(modules)* $\to$ A-*(modules)* that sends $\mathcal{M}$ to $\mathcal{M}(X)$. Given an A-module M, the corresponding $\mathcal{O}$-module $\mathcal{M}$ is defined as follows: Let $U = \operatorname{Spec} B$ be an affine open subset of X. The inclusion $U \subset X$ corresponds to an algebra homomorphism $A \to B$. We define $\mathcal{M}(U)$ to be the B-module $B \otimes_A M$. This gives us an $\mathcal{O}$-module because, when s is a nonzero element of B, then $B_s \otimes_A M$ is the localization $(B \otimes_A M)_s$ of $B \otimes_A M$. □

6.5 Direct Image

Let $Y \xrightarrow{f} X$ be a morphism of varieties, and let $\mathcal{N}$ be an $\mathcal{O}_Y$-module. The *direct image* $f_*\mathcal{N}$ is an $\mathcal{O}_X$-module that is defined as follows: If U is an affine open subset of X and $V = f^{-1}U$, then

$$[f_*\mathcal{N}](U) = \mathcal{N}(V)$$

In particular, the direct image $f_*\mathcal{O}_Y$ of the structure sheaf $\mathcal{O}_Y$ is defined by $[f_*\mathcal{O}_Y](U) = \mathcal{O}_Y(Y)$. It is a functor

$$\mathcal{O}_Y - modules \xrightarrow{f_*} \mathcal{O}_X - modules$$

The direct image generalizes restriction of scalars in modules over rings. If $A \xrightarrow{\varphi} B$ is an algebra homomorphism and N is a B-module, one can *restrict scalars* to make N into an A-module. Scalar multiplication by an element a of A on the restricted module N is defined to be scalar multiplication by its image $\varphi(a)$ in B. For clarity, we sometimes denote the given B-module by N_B and the A-module obtained by restriction of scalars by N_A. The additive groups N_B and N_A are the same.

When one replaces the algebras A and B by their spectra $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} B$, the algebra homomorphism φ defines a morphism $Y \xrightarrow{f} X$, and an $\mathcal{O}_Y$-module $\mathcal{N}$ is determined by a B-module N_B. Then $f_*\mathcal{N}$ is the $\mathcal{O}_X$-module determined by the A-module N_A.

proof. Let $U' \to U$ be an inclusion of affine open subsets of X, and let $V = f^{-1}U$ and $V' = f^{-1}U'$. These inverse images are open subsets of X, but they aren't necessarily affine. The inclusion $V' \to V$ gives us a homomorphism $\mathcal{N}(V) \to \mathcal{N}(V')$, and therefore a homomorphism $f_*\mathcal{N}(U) \to f_*\mathcal{N}(U')$. So $f_*\mathcal{N}$ is a functor. Its $\mathcal{O}_X$-module structure is explained as follows: Composition with f defines a homomorphism $\mathcal{O}_X(U) \to \mathcal{O}_Y(V)$, and $\mathcal{N}(V)$ is an $\mathcal{O}_Y(V)$-module. Restriction of scalars makes $[f_*\mathcal{N}](U) = \mathcal{N}(V)$ into a module over $\mathcal{O}_X(U)$.

To show that $f_*\mathcal{N}$ is an $\mathcal{O}_X$-module, we must show that if s is a nonzero element of $\mathcal{O}_X(U)$, then $[f_*\mathcal{N}](U_s)$ is obtained by localizing $[f_*\mathcal{N}](U)$. Let s' be the image of s in $\mathcal{O}_V(V)$. Scalar multiplication by s on $[f_*\mathcal{N}](U)$ is given by restriction of scalars, so it is the same as scalar multiplication by s' on $\mathcal{N}(V)$. If $s' \neq 0$, the localization $V_{s'}$ is the inverse image of U_s. So $[f_*\mathcal{N}](U_s) = \mathcal{N}(V_{s'})$. The coherence property (**6.3.14**) tells us that $\mathcal{N}(V_{s'}) = \mathcal{N}(V)_{s'}$. Then $[f_*\mathcal{N}](U_s) = \mathcal{N}(V_{s'}) = \mathcal{N}(V)_{s'} = [[f_*\mathcal{N}](U)]_s$.

If $s' = 0$, then $\mathcal{N}(V)_{s'} = 0$. In this case, because scalar multiplication is defined by restricting scalars, s annihilates $[f_*\mathcal{N}](U)$, and therefore $[f_*\mathcal{N}](U)_s = 0$ too. □

6.5.1. Lemma. *Let $Y \xrightarrow{f} X$ be a morphism of varieties. The direct image $f_*\mathcal{N}$ of an $\mathcal{O}_Y$-module $\mathcal{N}$ is an $\mathcal{O}_X$-module. Moreover, for all open subsets U of X, not only for affine open subsets, $[f_*\mathcal{N}](U) = \mathcal{N}(f^{-1}U)$.* □

6.5.2. Lemma. *Direct images are compatible with limits: If $\mathcal{M}_\bullet$ is a directed set of $\mathcal{O}$-modules, then $\varinjlim (f_*\mathcal{M}_\bullet) \approx f_*(\varinjlim \mathcal{M}_\bullet)$.* □

(6.5.3) extension by zero

When $Y \xrightarrow{i} X$ is the inclusion of a **closed** subvariety into a variety X, the direct image $i_*\mathcal{N}$ of an $\mathcal{O}_Y$-module $\mathcal{N}$ is also called the *extension* of $\mathcal{N}$ *by zero*. If U is an open subset of X then, because i is an inclusion map, $i^{-1}U = U \cap Y$. Therefore

$$[i_*\mathcal{N}](U) = \mathcal{N}(U \cap Y)$$

The term "extension by zero" refers to the fact that, when an open set U of X doesn't meet Y, the intersection $U \cap Y$ will be empty, and the module of sections of $[i_*\mathcal{N}](U)$ will be zero. So $i_*\mathcal{N}$ is zero outside of the closed set Y.

6.5.4. Examples.

(i) Let $p \xrightarrow{i} X$ be the inclusion of a point into a variety. We may view the residue field $k(p)$ as an $\mathcal{O}$-module on p. Then its extension by zero $i_*k(p)$ is the residue field module κ_p.

(ii) Let $Y \xrightarrow{i} X$ be the inclusion of a closed subvariety, and let $\mathcal{I}$ be the ideal of Y. The extension by zero of the structure sheaf on Y fits into an exact sequence of $\mathcal{O}_X$-modules

$$0 \to \mathcal{I} \to \mathcal{O}_X \to i_*\mathcal{O}_Y \to 0$$

So the extension by zero $i_*\mathcal{O}_Y$ is isomorphic to the quotient module $\mathcal{O}_X/\mathcal{I}$. □

6.5.5. Proposition. *Let $Y \xrightarrow{i} X$ be the inclusion of a closed subvariety Y into a variety X, and let $\mathcal{I}$ be the ideal of Y. Let $\mathbb{M}$ denote the subcategory of the category of $\mathcal{O}_X$-modules that are annihilated by $\mathcal{I}$. Extension by zero defines an equivalence of categories*

$$(\mathcal{O}_Y - modules) \xrightarrow{i_*} \mathbb{M}$$

proof. Let f be a section of $\mathcal{O}_X$ on an affine open set U, let $\overline{f}$ be its restriction to $U \cap Y$, and let α be an element of $[i_*\mathcal{N}](U)$ $\big(= \mathcal{N}(U \cap Y)\big)$. If f is in $\mathcal{I}(U)$, then $\overline{f} = 0$ and therefore $f\alpha = \overline{f}\alpha = 0$. So the extension by zero of an $\mathcal{O}_Y$-module is annihilated by $\mathcal{I}$. The direct image $i_*\mathcal{N}$ is an object of $\mathbb{M}$.

We construct a quasi-inverse to the direct image. Starting with an $\mathcal{O}_X$-module $\mathcal{M}$ that is annihilated by $\mathcal{I}$, we construct an $\mathcal{O}_Y$-module $\mathcal{N}$ such that $i_*\mathcal{N}$ is isomorphic to $\mathcal{M}$.

Let Y' be an open subset of Y. The topology on Y is induced from the topology on X, so $Y' = X_1 \cap Y$ for some open subset X_1 of X. We try to set $\mathcal{N}(Y') = \mathcal{M}(X_1)$. To show that this is well-defined, we show that if X_2 is another open subset of X, and if $Y' = X_2 \cap Y$, then $\mathcal{M}(X_2)$ is isomorphic to $\mathcal{M}(X_1)$. Let $X_3 = X_1 \cap X_2$. Then it is also true that $Y' = X_3 \cap Y$. Since $X_3 \subset X_1$, we have a map $\mathcal{M}(X_1) \to \mathcal{M}(X_3)$, and It suffices to show that this map is an isomorphism. The same reasoning will give us an isomorphism $\mathcal{M}(X_2) \to \mathcal{M}(X_3)$.

The complement $U = X_1 - Y'$ of Y' in X_1 is an open subset of X_1 and of X, and $U \cap Y = \emptyset$. We cover U by a set $\{U^i\}$ of affine open sets. Then X_1 is covered by the open sets $\{U^i\}$ together with X_3. The restriction of $\mathcal{I}$ to each of the sets U^i is the unit ideal, and since $\mathcal{I}$ annihilates $\mathcal{M}$, $\mathcal{M}(U^i) = 0$. The sheaf property shows that $\mathcal{M}(X_1)$ is isomorphic to $\mathcal{M}(X_3)$. The rest of the proof is boring. □

(6.5.6) inclusion of an open set

Let $Y \xrightarrow{j} X$ be the inclusion of an **open** subvariety Y into a variety X.

First, let $\mathcal{M}$ be an $\mathcal{O}_X$-module. Since open subsets of Y are also open subsets of X, we can *restrict* $\mathcal{M}$ from X to Y. By definition, the sections of the restricted module on a subset U of Y are simply the elements of $\mathcal{M}(U)$. For example, the restriction of the structure sheaf $\mathcal{O}_X$ is the structure sheaf $\mathcal{O}_Y$ on Y. We extend the subscript notation to $\mathcal{O}$-modules, writing $\mathcal{M}_Y$ for the restriction of an $\mathcal{O}_X$-module $\mathcal{M}$ to Y and denoting the given module $\mathcal{M}$ by $\mathcal{M}_X$. Then if U is an open subset of Y,

$$\mathcal{M}_X(U) = \mathcal{M}_Y(U)$$

Now the direct image: Let $Y \xrightarrow{j} X$ be the inclusion of an open subvariety Y, and let $\mathcal{N}$ be an $\mathcal{O}_Y$-module. The inverse image of an open subset U of X is the intersection $Y \cap U$, so

$$[j_*\mathcal{N}](U) = \mathcal{N}(Y \cap U)$$

For example, $[j_*\mathcal{O}_Y](U)$ is the algebra of rational functions on X that are regular on $Y \cap U$. They needn't be regular on U.

6.5.7. Example. Let $X_s \xrightarrow{j} X$ be the inclusion of a localization into an affine variety $X = \operatorname{Spec} A$. Modules on X correspond to their global sections, which are A-modules. Similarly, modules on X_s correspond to A_s-modules. We restrict an $\mathcal{O}_X$-module $\mathcal{M}_X$ to the open set X_s, obtaining an $\mathcal{O}_{X_s}$-module $\mathcal{M}_{X_s}$. Then if M denotes the A-module of global sections $\mathcal{M}_X(X)$, the module of global sections of the direct image $j_*\mathcal{M}_{X_s}$ is the localization M_s:

$$[j_*\mathcal{M}_{X_s}](X) = \mathcal{M}_{X_s}(X_s) = \mathcal{M}_X(X_s) = M_s$$

The localization M_s is made into an A-module by restriction of scalars. □

6.5.8. Proposition. *Let $Y \xrightarrow{j} X$ be the inclusion of an open subvariety Y into a variety X.*

(i) *The restriction $\mathcal{O}_X$-modules $\to \mathcal{O}_Y$-modules is an exact operation.*

(ii) *If Y is an* ***affine*** *open subvariety of X, the direct image functor j_* is exact.*

(iii) *Let $\mathcal{M}_X$ be an $\mathcal{O}_X$-module. There is a canonical homomorphism $\mathcal{M}_X \to j_*[\mathcal{M}_Y]$.*

proof. **(ii)** Let U be an affine open subset of X, and let $\mathcal{M} \to \mathcal{N} \to \mathcal{P}$ be an exact sequence of $\mathcal{O}_Y$-modules. The sequence $j_*\mathcal{M}(U) \to j_*\mathcal{N}(U) \to j_*\mathcal{P}(U)$ is the same as the sequence $\mathcal{M}(U \cap Y) \to \mathcal{N}(U \cap Y) \to \mathcal{P}(U \cap Y)$, except that the scalars have changed. Since U and Y are affine, $U \cap Y$ is affine. By definition of exactness, this last sequence is exact.

(iii) Let U be open in X. Then $j_*\mathcal{M}_Y(U) = \mathcal{M}_X(U \cap Y)$. Since $U \cap Y \subset U$, $\mathcal{M}_X(U)$ maps to $\mathcal{M}_X(U \cap Y)$. □

6.5.9. Example. Let $X = \mathbb{P}^n$ and let j denote the inclusion $\mathbb{U}^0 \subset X$ of the standard affine open subset into X. The direct image $j_*\mathcal{O}_{\mathbb{U}^0}$ is the algebra of rational functions that are allowed to have poles on the hyperplane at infinity.

The inverse image of an open subset W of X is its intersection with $\mathbb{U}^0$: $j^{-1}W = W \cap \mathbb{U}^0$. The sections of the direct image $j_*\mathcal{O}_{\mathbb{U}^0}$ on an open subset W of X are the regular functions on $W \cap \mathbb{U}^0$:

$$[j_*\mathcal{O}_{\mathbb{U}^0}](W) \;=\; \mathcal{O}_{\mathbb{U}^0}(W \cap \mathbb{U}^0) \;=\; \mathcal{O}_X(W \cap \mathbb{U}^0)$$

Say that we write a rational function α as a fraction g/h of relatively prime polynomials. Then α is an element of $\mathcal{O}_X(W)$ if h doesn't vanish at any point of W, and α is a section of $[j_*\mathcal{O}_{\mathbb{U}^0}](W) = \mathcal{O}_X(W \cap \mathbb{U}^0)$ if h doesn't vanish on $W \cap \mathbb{U}^0$. Arbitrary powers of x_0 can appear in the denominator h. □

6.6 Twisting

The twisting modules that we define here are among the most important modules on projective space.

Let X denote the projective space $\mathbb{P}^n$ with coordinates $x_0, ..., x_n$. As before, a homogeneous fraction of degree d is a fraction g/h of homogeneous polynomials with $\deg\, g - \deg\, h = d$. When g and h are relatively prime, the fraction g/h is *regular* on an open set V if h isn't zero at any point of V.

The definition of the twisting module is this: The sections of $\mathcal{O}(d)$ on an open subset V of $\mathbb{P}^n$ are the homogeneous fractions of degree d that are regular on V.

6.6.1. Proposition.
(i) *Let V be an affine open subset of $\mathbb{P}^n$ that is contained in the standard affine open set $\mathbb{U}^0$. The sections of $\mathcal{O}(d)$ on V form a free module of rank one with basis x_0^d, over the coordinate algebra $\mathcal{O}(V)$.*
(ii) *The twisting module $\mathcal{O}(d)$ is an $\mathcal{O}$-module.*

proof. **(i)** Let α be a section of $\mathcal{O}(d)$ on an affine open set V that is contained in $\mathbb{U}^0$. Then $f = \alpha x_0^{-d}$ has degree zero. It is a rational function. Since $V \subset \mathbb{U}^0$, x_0 doesn't vanish at any point of V. Since α is regular on V, f is a regular function on V, and $\alpha = fx_0^d$.

(ii) It is clear that $\mathcal{O}(d)$ is a contravariant functor. We verify compatibility with localization. Let $V = \operatorname{Spec} A$ be an affine open subset of X and let s be a nonzero element of A. We must show that $[\mathcal{O}(d)](V_s)$ is the localization of $[\mathcal{O}(d)](V)$, and it is true that $[\mathcal{O}(d)](V)$ is a subset of $[\mathcal{O}(d)](V_s)$. What has to be shown is that if β is a section of $\mathcal{O}(d)$ on V_s, then $s^k\beta$ is a section on V, if k is sufficiently large.

We cover V by the affine open sets $V^i = V \cap \mathbb{U}^i$. To show that $s^k\beta$ is a section on V, it suffices to show that it is a section on $V \cap \mathbb{U}^i$ for every i. This is the sheaf property. We apply **(i)** to the open subset V_s^0 of V^0. Since V_s^0 is contained in $\mathbb{U}^0$, β can be written (uniquely) in the form fx_0^d, where f is a rational function that is regular on V_s^0. We know already that $\mathcal{O}$ has the localization property. Therefore $s^k f$ is a regular function on V^0 if k is large, and then $s^k\alpha = s^k f x_0^d$ is a section of $\mathcal{O}(d)$ on V^0. The analogous statement is true for every index i. □

Part **(i)** of the proposition shows that $\mathcal{O}(d)$ is quite similar to the structure sheaf. However, $\mathcal{O}(d)$ is only *locally* free. Its sections on the standard open set $\mathbb{U}^1$ form a free $\mathcal{O}(\mathbb{U}^1)$-module with basis x_1^d. That basis is related to the basis x_0^d on $\mathbb{U}^0$ by the factor $(x_0/x_1)^d$, a rational function that isn't invertible on on $\mathbb{U}^0$ or on $\mathbb{U}^1$.

6.6.2. Proposition. *When $d \geq 0$, the global sections of the twisting module $\mathcal{O}(d)$ on $\mathbb{P}^n$ $(n > 0)$ are the homogeneous polynomials of degree d. When $d < 0$, the only global section of $\mathcal{O}(d)$ is zero.*

proof. A nonzero global section u of $\mathcal{O}(d)$ will restrict to a section on the standard affine open set $\mathbb{U}^0$. Since $[\mathcal{O}(d)](\mathbb{U}^0)$ is a free module over $\mathcal{O}(\mathbb{U}^0)$ with basis x_0^d, and $u = g/x_0^m$ for some some homogeneous polynomial g not divisible by x_0 and some m. Similarly, restriction to $\mathbb{U}^1$ shows that $u = h/x_1^n$. It follows that $m = n = 0$ and that $u = g$. Since u has degree d, g will be a polynomial of degree d. □

6.6.3. Examples.
The product uv of homogeneous fractions of degrees r and s is a homogeneous fraction of degree $r+s$, and if u and v are regular on an open set V, so is their product uv. Therefore multiplication defines a homomorphism of $\mathcal{O}$-modules

(6.6.4) $$\mathcal{O}(r) \times \mathcal{O}(s) \to \mathcal{O}(r+s)$$

Multiplication by a homogeneous polynomial f of degree d defines an injective homomorphism

(6.6.5) $$\mathcal{O}(k) \xrightarrow{f} \mathcal{O}(k+d).$$

When $k = -d$, this becomes a homomorphism $\mathcal{O}(-d) \xrightarrow{f} \mathcal{O}$. □

The twisting modules $\mathcal{O}(n)$ have a second interpretation. They are isomorphic to the modules that we denote by $\mathcal{O}(nH)$, of rational functions on projective space with poles of order at most n on the hyperplane $H : \{x_0 = 0\}$ at infinity.

By definition, the nonzero sections of $\mathcal{O}(nH)$ on an open set V are the rational functions f such that $x_0^n f$ is a section of $\mathcal{O}(n)$ on V. Thus multiplication by x_0^n defines an isomorphism

(6.6.6) $$\mathcal{O}(nH) \xrightarrow{x_0^n} \mathcal{O}(n)$$

If f is a section of $\mathcal{O}(nH)$ on an open set V, and if we write f as a homogeneous fraction g/h of degree zero, with g, h relatively prime, the denominator h may have x_0^k, with $k \leq n$, as factor. The other factors of h cannot vanish anywhere on V. If $f = g/h$ is a global section of $\mathcal{O}(nH)$, then $h = cx_0^k$, with $k \leq n$, so a global section can be represented as a fraction g/x_0^k.

Since x_0 doesn't vanish at any point of the standard affine open set $\mathbb{U}^0$, the sections of $\mathcal{O}(nH)$ on an open subset V of $\mathbb{U}^0$ are simply the regular functions on V. The restrictions of $\mathcal{O}(nH)$ and $\mathcal{O}$ to $\mathbb{U}^0$ are equal. Using the subsctript notation (**6.5.6**) for restriction to an open set,

(6.6.7) $$\mathcal{O}(nH)_{\mathbb{U}^0} = \mathcal{O}_{\mathbb{U}^0}$$

Let V be an open subset of one of the other standard affine open sets, say of $\mathbb{U}^1$. The ideal of $H \cap \mathbb{U}^1$ in $\mathbb{U}^1$ is principal, generated by $v_0 = x_0/x_1$, and v_0 generates the ideal of $H \cap V$ in V too. If f is a rational function, then because x_1 doesn't vanish on $\mathbb{U}^1$, the function fv_0^n will be regular on V if and only if the homogeneous fraction fx_0^n is regular there. So f will be a section of $\mathcal{O}(nH)$ on V if and only if fv_0^n is a regular function. Because v_0 generates the ideal of H in V, we say that such a function f has a pole of order at most n on H.

The isomorphic $\mathcal{O}$-modules $\mathcal{O}(n)$ and $\mathcal{O}(nH)$ are interchangeable. The twisting module $\mathcal{O}(n)$ is often better because its definition is independent of coordinates. On the other hand, $\mathcal{O}(nH)$ can be convenient because it restricts to the structure sheaf $\mathcal{O}$ on $\mathbb{U}^0$.

6.6.8. Proposition. *Let Y be the zero locus of an irreducible homogeneous polynomial f of degree d, a hypersurface of degree d in $\mathbb{P}^n$, let $\mathcal{I}$ be the ideal of Y, and let $\mathcal{O}(-d)$ be the twisting module on X. Multiplication by f defines an isomorphism $\mathcal{O}(-d) \xrightarrow{f} \mathcal{I}$.*

proof. If α is a section of $\mathcal{O}(-d)$ on an open set V, then $f\alpha$ will be a rational function that is regular on V and vanishes on Y. Therefore the image of the multiplication map $\mathcal{O}(-d) \xrightarrow{f} \mathcal{O}$ is contained in $\mathcal{I}$. This map is injective because $\mathbb{C}[x_0, ..., x_n]$ is a domain. To show that it is an isomorphism, it suffices to show that its restrictions to the standard affine open sets $\mathbb{U}^i$ are isomorphisms (6.3.5). As usual, we work with $\mathbb{U}^0$.

We choose coordinates in X so that the coordinate variables x_i don't divide f. Then $Y \cap \mathbb{U}^0$ will be a dense open subset of Y. The sections of $\mathcal{O}$ on $\mathbb{U}^0$ are the homogeneous fractions g/x_0^k of degree zero. Such a fraction is a section of $\mathcal{I}$ on $\mathbb{U}^0$ if and only if g vanishes on $Y \cap \mathbb{U}^0$. If so, then since $Y \cap \mathbb{U}^0$ is dense in Y, it will vanish on Y, and therefore it will be divisible by f: $g = fq$. The sections of $\mathcal{I}$ on $\mathbb{U}^0$ have the form fq/x_0^k. They are in the image of $\mathcal{O}(-d)$. $\square$

The proposition has an interesting corollary:

6.6.9. Corollary. *The ideals of all hypersurfaces of degree d are isomorphic, when they are regarded as $\mathcal{O}$-modules.* $\square$

(6.6.10) twisting a module

6.6.11. Definition Let $\mathcal{M}$ be an $\mathcal{O}$-module on projective space $\mathbb{P}^d$, and let $\mathcal{O}(n)$ be the twisting module. The nth *twist* of $\mathcal{M}$ is defined to be the tensor product $\mathcal{M}(n) = \mathcal{M} \otimes_{\mathcal{O}} \mathcal{O}(n)$, and similarly, $\mathcal{M}(nH) = \mathcal{M} \otimes_{\mathcal{O}} \mathcal{O}(nH)$. Twisting is a functor on $\mathcal{O}$-modules.

If X is a closed subvariety of $\mathbb{P}^d$ and $\mathcal{M}$ is an $\mathcal{O}_X$-module, $\mathcal{M}(n)$ and $\mathcal{M}(nH)$ are obtained by twisting the extension of $\mathcal{M}$ by zero. (See the equivalence of categories (6.5.5)).

A section of $\mathcal{M}(n)$ on an open subset V of $\mathbb{U}^0$ can be written in the form $s = m \otimes fx_0^n$, where f is a regular function on V and m is a section of $\mathcal{M}$ on V (6.6.1). The function f can be moved over to m, so a section can be written in the form $s = m \otimes x_0^n$. This expression is unique.

6.6.12. The modules $\mathcal{O}(n)$ and $\mathcal{O}(nH)$ form directed sets that are related by a diagram

$$\begin{array}{ccccccc} \mathcal{O} & \xrightarrow{\subset} & \mathcal{O}(H) & \xrightarrow{\subset} & \mathcal{O}(2H) & \xrightarrow{\subset} & \cdots \\ \| & & {\scriptstyle x_0}\downarrow & & {\scriptstyle x_0^2}\downarrow & & \\ \mathcal{O} & \xrightarrow{x_0} & \mathcal{O}(1) & \xrightarrow{x_0} & \mathcal{O}(2) & \longrightarrow & \cdots \end{array} \tag{6.6.13}$$

In this diagram, the vertical arrows are bijections. The limit of the upper directed set is the module whose sections are allowed to have arbitrary poles on H. This is also the module $j_*\mathcal{O}_{\mathbb{U}}$, where j denotes the inclusion of the standard affine open set $\mathbb{U} = \mathbb{U}^0$ into X (see (6.5.8) **(iii)**):

$$\varinjlim \mathcal{O}(nH) = j_*\mathcal{O}_{\mathbb{U}} \tag{6.6.14}$$

The next diagram is obtained by tensoring Diagram 6.6.13 with $\mathcal{M}$.

$$\begin{array}{ccccccc} \mathcal{M} & \longrightarrow & \mathcal{M}(H) & \longrightarrow & \mathcal{M}(2H) & \longrightarrow & \cdots \\ \| & & {\scriptstyle x_0}\downarrow & & {\scriptstyle x_0^2}\downarrow & & \\ \mathcal{M} & \xrightarrow{1\otimes x_0} & \mathcal{M}(1) & \xrightarrow{1\otimes x_0} & \mathcal{M}(2) & \longrightarrow & \cdots \end{array} \tag{6.6.15}$$

Because $\mathcal{M}$ may have torsion, the horizontal maps in these two directed sets needn't be injective.

(6.6.16) generating an $\mathcal{O}$-module

A set $m = (m_1, ..., m_k)$ of global sections of an $\mathcal{O}$-module $\mathcal{M}$ defines a map

$$\mathcal{O}^k \xrightarrow{m} \mathcal{M} \tag{6.6.17}$$

that sends a section $(\alpha_1, ..., \alpha_k)$ of $\mathcal{O}^k$ on an open set to the combination $\sum \alpha_i m_i$. The set of global sections $\{m_1, ..., m_k\}$ *generates* $\mathcal{M}$ if this map is surjective. If the sections generate $\mathcal{M}$, then they (more precisely, their restrictions to U) generate the $\mathcal{O}(U)$-module $\mathcal{M}(U)$ for every affine open set U. They may fail to generate $\mathcal{M}(U)$ when U isn't affine.

6.6.18. Example. Let $X = \mathbb{P}^1$. For $n \geq 0$, the global sections of the twisting module $\mathcal{O}(n)$ are the polynomials of degree n in the coordinate variables x_0, x_1 (6.6.2). Consider the map $\mathcal{O}^2 \xrightarrow{(x_0^n, x_1^n)} \mathcal{O}(n)$. On $\mathbb{U}^0$, $\mathcal{O}(n)$ has basis x_0^n. Therefore this map is surjective on $\mathbb{U}^0$. Similarly, it is surjective on $\mathbb{U}^1$. So it is a surjective map on all of X (6.3.5). The global sections x_0^n, x_1^n generate $\mathcal{O}(n)$. However, the global sections of $\mathcal{O}(n)$ are the homogeneous polynomials of degree n. When $n > 1$, the two sections x_0^n, x_1^n don't span the space of global sections, and the map $\mathcal{O}^2 \xrightarrow{(x_0^n, x_1^n)} \mathcal{O}(n)$ isn't surjective. □

The next theorem explains the importance of the twisting operation.

6.6.19. Theorem. *Let $\mathcal{M}$ be a finite $\mathcal{O}$-module on a projective variety X. For large n, the twist $\mathcal{M}(n)$ is generated by global sections.*

The proof won't be long, once the notation is introduced.

We may assume that X is projective space $\mathbb{P}^n$. Let $\mathbb{U}^i$, denote the standard affine open subsets of $\mathbb{P}^n$, let and their intersections be $\mathbb{U}^{ij} = \mathbb{U}^i \cap \mathbb{U}^j$, and $\mathbb{U}^{ijk} = \mathbb{U}^i \cap \mathbb{U}^j \cap \mathbb{U}^k$.

Let A_i, A_{ij}, and A_{ijk} denote the coordinate algebras of $\mathbb{U}^i$, $\mathbb{U}^{ij}$, and $\mathbb{U}^{ijk}$, respectively, and let M_i^k, M_{ij}^k, and M_{ijk}^k denote the modules of sections of $\mathcal{M}(k)$ on $\mathbb{U}^i$, $\mathbb{U}^{ij}$ and $\mathbb{U}^{ijk}$. So $M_j^k = [\mathcal{M}(k)](\mathbb{U}^j)$. Similarly, let $A_j^k = [\mathcal{O}(k)](\mathbb{U}^j)$ and $M_j = \mathcal{M}(\mathbb{U}^j)$, etc. Recall that $\mathcal{M}(k) = \mathcal{M} \otimes_{\mathcal{O}} \mathcal{O}(k)$.

Multiplication by the global section x_i^k of $\mathcal{O}(k)$ maps $\mathcal{O}$ to $\mathcal{O}(k)$, and therefore it defines maps $A_j \to A_j^k$ and $M_j \to M_j^k$. Since A_j^k is a free A_j-module with basis x_j^k, multiplication by x_j^k defines bijective, and therefore invertible, maps $A_j \to A_j^k$ and $M_j \to M_j^k$.

When we compose the maps

$$M_j \xrightarrow{x_i} M(1)_j \xrightarrow{x_j^{-1}} M_j$$

we obtain the map $M_j \xrightarrow{u_{ij}} M_j$, where $u_{ij} = x_i/x_j$. The next lemma shows that, when coordinates are in general position, this is an injective map.

6.6.20. Lemma. *Let $\mathcal{M}$ be a finite module on $\mathbb{P}^n$.*
(i) *Let $w = c_0x_0 + \cdots + c_nx_n$ be a generic combination of the coordinate variables. Multiplication by w^k defines an injective map $\mathcal{M} \to \mathcal{M}(k)$ for all k.*
(ii) *If coordinates are in general position, then multiplication by x_i^k defines an injective map $\mathcal{M} \to \mathcal{M}(k)$ for all i.*

proof. **(i)** It suffices to show that multiplication by w^k is injective when we restrict to the standard affine open sets $\mathbb{U}^j$ (6.3.5). We work with the index 0. The rational function $z = w/x_0$ is a generic linear combination of $u_{10}, ..., u_{n0}$, and the map $M_0 \xrightarrow{w^k} M_0^k$ is injective if and only if $M_0 \xrightarrow{z^k} M_0$ is injective.

Let $P_1, ..., P_k$ be the maximal annihilators in A_0 of the finite module M_0. (See (**6.4.16**).) Scalar multiplication by z^k is injective on M_0 if and only if z isn't contained in any of those maximal annihilators (6.4.18). This will be true when w and z are generic.

(ii) When we make a generic change of coordinates, x_i become generic homogeneous linear polynomials. □

6.6.21. Lemma. *Let $\mathcal{M}$ be a finite $\mathcal{O}$-module on $\mathbb{P}^n$, and let coordinates $x_0, ..., x_n$ be in general position. If m_0 is an element of M_0, then when k is sufficiently large, the product $m_0x_0^k$ is the restriction of a global section of $\mathcal{M}(k)$.*

proof. s Let's use the notation $M[s^{-1}]$ to denote localization: $M[s^{-1}] = M \otimes_A A[s^{-1}]$. The module M_{01} is a localization of M_1 $\big($and a localization of $M_0\big)$. Let $u_{ij} = x_i/x_j$. Then $M_{ij} = M_j[u_{ij}^{-1}] = M_j[x_j/x_i]$ and in particular, $M_{01} = M_1[u_{01}^{-1}] = M_1[x_1/x_0]$. If m_0 is an element of M_0, its restriction to M_{01} can be written as $m_0 = m_1u_{01}^{-k} = m_1x_1^k/x_0^k$, for some m_1 in M_1 and some k. Multiplying by x_0^k, $m_0x_0^k = m_1x_1^k$ in M_{01}^k. Similarly, if we increase k as needed, there will be elements m_i in M_i such that $m_0x_0^k = m_ix_i^k$ in M_{0i}^k for all $i = 0, ..., n$. Then $m_ix_i^k = m_0x_0^k = m_jx_j^k$ is true in M_{0ij}^k for all i, j. Lemma 6.6.20 shows that, because coordinates are in general position, multiplication by x_0 is injective on $\mathcal{M}$. Therefore multiplication by $u_{0j} = x_0/x_j$ is injective on M_{ij} and on M_{ij}^k. It follows that the localization map $M_{ij}^k \to M_{0ij}^k$ is injective. The equation $m_ix_i^k = m_jx_j^k$, which is true in M_{kij}, is true in M_{ij}^k too. The sheaf property shows that $\mathcal{M}(k)$ has a global section w that restricts to $m_jx_j^k$ on U^j, and in particular, that restricts to $m_0x_0^k$ on $\mathbb{U}^0$. □

proof of Theorem 6.6.19. We are to show that the global sections generate $\mathcal{M}(k)$ when k is large, and it suffices to show that for each $i = 0, ..., n$, the restrictions of those global sections to the standard open set $\mathbb{U}^i$ generate the A_i-module M_i^k (6.3.5). We work with the index 0 as before.

Since $\mathcal{M}$ is a finite $\mathcal{O}$-module and $\mathbb{U}^0$ is affine, M_0 is a finite A_0-module. We choose a finite set $\{m^\nu\}$ that generates M_0. Then the set $\{m^\nu x_0^k\}$ generates M_0^k. It suffices to show $m^\nu x_0^k$ are the restrictions to $\mathbb{U}^0$ of global sections of $\mathcal{M}(k)$, when k is large. If so, then because the analogous statements are true for all indices $i = 0, ..., n$, $\mathcal{M}(k)$ will be generated by global sections. □

6.7 Proof of Theorem 6.3.2.

The statement to be proved is that an $\mathcal{O}$-module $\mathcal{M}$ on a variety X has a unique extension to a functor

$$\textit{(opens)} \xrightarrow{\widetilde{\mathcal{M}}} \textit{(modules)}$$

having the sheaf property, and that a homomorphism $\mathcal{M} \to \mathcal{N}$ of $\mathcal{O}$-modules has a unique extension to a homomorphism $\widetilde{\mathcal{M}} \to \widetilde{\mathcal{N}}$.

The proof has the following steps:

1. Verification of the sheaf property for a covering of an affine open set by localizations.

2. Extension of the functor $\mathcal{M}$ to all morphisms between affine open sets.

3. Definition of $\widetilde{\mathcal{M}}$.

4. Verification of the sheaf property for $\widetilde{\mathcal{M}}$ and for an affine covering of an open set Y.

Step 1. (*the sheaf property for a covering of an affine open set by localizations*)

Suppose that an affine open set $Y = \operatorname{Spec} A$ is covered by a family of localizations $\mathbb{U}_0 = \{U^{s_i}\}$. Then if $\mathcal{M}$ is an $\mathcal{O}_Y$-module and if we let $M = \mathcal{M}(Y)$, $M_i = \mathcal{M}(U^{s_i})$ and $M_{ij} = \mathcal{M}(U^{s_i s_j})$ the sequence (6.3.9) for the covering diagram $Y \longleftarrow \mathbb{U}_0 \leftleftarrows \mathbb{U}_1$ that we obtain is

$$0 \to M \xrightarrow{\alpha} \prod M_i \xrightarrow{\beta} \prod M_{ij} \tag{6.7.1}$$

In this sequence, the map α sends an element m of M to the vector $(m, ..., m)$ of its images in $\prod_i M_i$, and β sends a vector $(m_1, ..., m_k)$ in $\prod_i M_i$ to the matrix (z_{ij}), with $z_{ij} = m_j - m_i$ in M_{ij} To be precise, M_i and M_j map to M_{ij}, and z_{ij} is the difference of their images.

We must show that the sequence (6.7.1) is exact. Since U^i cover Y, the elements $s_1, ..., s_k$ generate the unit ideal.

exactness at M: Let m be an element of M that maps to zero in every M_i. Then there exists an n such that $s_i^n m = 0$, and we can use the same exponent n for all i. The elements s_i^n generate the unit ideal. Writing $\sum a_i s_i^n = 1$, we have $m = \sum a_i s_i^n m = \sum a_i 0 = 0$.

exactness at $\prod M_i$: Let m_i be elements of M_i such that $m_i = m_j$ in M_{ij} for all i, j. We must find an element w in M that maps to m_j in M_j for every j.

We write m_i as a fraction: $m_i = s_i^{-n} x_i$, or $x_i = s_i^n m_i$, with x_i in M, using the same integer n for all i. The equation $m_i = m_j$ in M_{ij} tells us that $s_j^n x_i = s_i^n x_j$ is true in M_{ij}, and then $(s_i s_j)^r s_j^n x_i = (s_i s_j)^r s_i^n x_j$ will be true in M, if r is large (see **5.1.21**).

We adjust the notation. Let $\widetilde{x}_i = s_i^r x_i$, and $\widetilde{s}_i = s_i^{r+n}$. Then in M, $\widetilde{x}_i = \widetilde{s}_i m_i$ and $\widetilde{s}_j \widetilde{x}_i = \widetilde{s}_i \widetilde{x}_j$. Since the elements s_i generate the unit ideal, so do their powers $\widetilde{s}_i$. There is an equation in A, of the form $\sum a_i \widetilde{s}_i = 1$. Let $w = \sum a_i \widetilde{x}_i$. This is an element of M, and

$$\widetilde{x}_j = \big(\sum_i a_i \widetilde{s}_i\big)\widetilde{x}_j = \sum_i a_i \widetilde{s}_j \widetilde{x}_i = \widetilde{s}_j w$$

Since $m_j = \widetilde{s}_j^{-1} \widetilde{x}_j$, $m_j = w$ is true in M_j. Since j is arbitrary, w is the required element of M. □

Step 2. (*extending an $\mathcal{O}$-module to all morphisms between affine open sets*)

The $\mathcal{O}$-module $\mathcal{M}$ comes with localization maps $\mathcal{M}(U) \to \mathcal{M}(U_s)$. It doesn't come with homomorphisms $\mathcal{M}(U) \to \mathcal{M}(V)$ when $V \to U$ is an arbitrary inclusion of affine open sets. We define those maps here.

Let $\mathcal{M}$ be an $\mathcal{O}$-module and let $V \to U$ be an inclusion of affine open sets. To describe the canonical homomorphism $\mathcal{M}(U) \to \mathcal{M}(V)$, we cover V by a family $\mathbb{V}_0 = \{V^i\}$ of open sets that are localizations of U and therefore also localizations of V, and we inspect the covering diagram $V \leftarrow \mathbb{V}_0 \leftleftarrows \mathbb{V}_1$ and the corresponding exact sequence $0 \to \mathcal{M}(V) \xrightarrow{\alpha} \mathcal{M}(\mathbb{V}_0) \xrightarrow{\beta} \mathcal{M}(\mathbb{V}_1)$. The two maps $\mathbb{V}_1 \to U$ obtained by composition from the maps

$$U \leftarrow V \leftarrow \mathbb{V}_0 \leftleftarrows \mathbb{V}_1$$

are equal. Since V^i are localizations of U and V^{ij} are localizations of V^i and of V^j, the $\mathcal{O}$-module $\mathcal{M}$ comes with maps $\mathcal{M}(U) \xrightarrow{\psi} \mathcal{M}(\mathbb{V}_0) \rightrightarrows \mathcal{M}(\mathbb{V}_1)$. The two composed maps $\mathcal{M}(U) \to \mathcal{M}(\mathbb{V}_1)$ are equal, so their difference $\beta\psi$ is zero. Therefore ψ maps $\mathcal{M}(U)$ to the kernel of β which, according to Step 1, is $\mathcal{M}(V)$. This defines a map $\mathcal{M}(U) \xrightarrow{\eta} \mathcal{M}(V)$ making a diagram

$$\begin{array}{ccc} \mathcal{M}(U) & \xrightarrow{\eta} & \mathcal{M}(V) \\ \| & & \downarrow{\scriptstyle\lambda} \\ \mathcal{M}(U) & \xrightarrow{\psi} & \mathcal{M}(\mathbb{V}_0) \end{array}$$

Both ψ and λ are compatible with multiplication by a regular function f on U, and λ is injective. So η is also compatible with multiplication by f.

We must check that η is independent of the covering $\mathbb{V}_0$. Let $\mathbb{V}_0' = \{V'^j\}$ be another covering of V by localizations of U. We cover each of the open sets $V^i \cap V'^j$ by localizations $W^{ij\nu}$ of U. Taken together, these open sets form a covering $\mathbb{W}_0$ of U, and we have a map $\mathbb{W}_0 \xrightarrow{\epsilon} \mathbb{V}_0$ that give us a diagram

$$\begin{array}{ccccccc} 0 & \longrightarrow & \mathcal{M}(V) & \longrightarrow & \mathcal{M}(\mathbb{V}_0) & \xrightarrow{\beta_{\mathbb{V}}} & \mathcal{M}(\mathbb{V}_1) \\ & & \| & & \downarrow & & \downarrow \\ 0 & \longrightarrow & \mathcal{M}(V) & \longrightarrow & \mathcal{M}(\mathbb{W}_0) & \xrightarrow{\beta_{\mathbb{W}}} & \mathcal{M}(\mathbb{W}_1) \end{array}$$

whose rows are exact sequences. In this diagram, $\mathcal{M}(U)$ maps to the kernels of $\beta_{\mathbb{V}}$ and $\beta_{\mathbb{W}}$, both of which are equal to $\mathcal{M}(V)$. Looking at the diagram, one sees that the map $\mathcal{M}(U) \to \mathcal{M}(\mathbb{W}_0)$ is the composition of the maps $\mathcal{M}(U) \to \mathcal{M}(\mathbb{V}_0) \to \mathcal{M}(\mathbb{W}_0)$. Therefore the two maps $\mathcal{M}(U) \to \mathcal{M}(V)$ are equal, and they are also equal to the map defined by the covering $\mathbb{V}_0'$ □

Step 3. (*definition of* $\widetilde{\mathcal{M}}$)

Let Y be an open subset of X. We can use the sheaf property to make a definition of $\widetilde{\mathcal{M}}(Y)$, but we need to choose a covering of Y by affine open sets. There is a canonical choice of covering, namely the set of *all* affine open subsets of Y. Let's denote this set by $\mathbb{A}_0(Y)$:

$$\mathbb{A}_0(Y) = \{U \,|\, U \text{ affine, and } U \subset Y\}$$

This is an infinite set, but that doesn't cause problems. Let $\mathit{aff}_1(Y)$ be the family of intersections of $\mathbb{A}_0(Y)$. We define $\widetilde{\mathcal{M}}(Y)$ to be the kernel $K_{\mathbb{A}}$ of the map $\mathbb{A}_0(Y) \xrightarrow{\beta_{\mathbb{A}}} \mathbb{A}_1(Y)$, where β is the map (6.3.9). This works well. If $Z \to Y$ are open sets, then $\mathbb{A}_0(Z) \subset \mathbb{A}_0(Y)$ and $\widetilde{\mathcal{M}}(Z)$ maps to $\widetilde{\mathcal{M}}(Y)$. This inclusion makes $\widetilde{\mathcal{M}}$ into a functor.

If Y is affine, it is the maximal element of $\mathbb{A}_0(Y$. Then $\mathcal{M}(Y)$ is a minimal element among $\mathcal{M}(U)$ with $U \in \mathbb{A}_0(Y)$, and so $\widetilde{\mathcal{M}}(Y) = K_{\mathbb{A}} = \mathcal{M}(Y)$. □

Step 4. (*the sheaf property of* $\widetilde{\mathcal{M}}$)

Let $\mathbb{U}_0 = \{U^i\}$ be an affine covering of an open set Y. We show that the kernel $K_{\mathbb{U}}$ of the map $\mathcal{M}(\mathbb{U}_0) \xrightarrow{\beta_{\mathbb{U}}} \mathcal{M}(\mathbb{U}_1)$ is canonically isomorphic to the kernel $K_{\mathbb{A}}$ of the analogous map $\mathcal{M}(\mathbb{A}_0) \xrightarrow{\beta_{\mathbb{A}}} \mathcal{M}(\mathbb{A}_1)$, which, by defnition, is $\widetilde{\mathcal{M}}(Y)$. Since the open sets U^i are included in $\mathbb{A}_0$, there are maps $\mathcal{M}(\mathbb{A}_i) \to \mathcal{M}(\mathbb{U}_i)$, and $\widetilde{\mathcal{M}}(Y) = K_{\mathbb{A}} \to K_{\mathbb{U}}$.

We consider a family $\mathbb{W}_0 = \{U^i, V\}$ obtained by adding one affine open subset V of Y to $\mathbb{U}_0$, and we let $\mathbb{W}_1$ be the family of intersections of pairs of elements of $\mathbb{W}_0$. Then we have a map $K_{\mathbb{W}} \to K_{\mathbb{U}}$. We show that, for any element (u_i) in the kernel $K_{\mathbb{U}}$, there is a unique element v in $\mathcal{M}(V)$ such that $((u_i), v)$ is in the kernel $K_{\mathbb{W}}$. This will show that $K_{\mathbb{W}} = K_{\mathbb{U}}$. Then if α is an element of $K_{\mathbb{A}}$ whose U_i component is u_i, its V-component must be v. Since V is an arbitrary affine subset of Y, it will follow that (u_i) is the image of a unique element of $K_{\mathbb{A}}$, namely the element whose V–component is v.

When the subsets in the family $\mathbb{W}_1$ are listed in the order

$$\mathbb{W}_1 = \{U^i \cap U^j\}, \{V \cap U^j\}, \{U^i \cap V\}, \{V \cap V\}$$

the map $\beta_{\mathbb{W}}$ sends a set $((u_i), v)$ of sections to $((u_j - u_i), (u_j - v), (v - u_i), 0)$, restricted appropriately.

Suppose that (u_i) is in the kernel of $\beta_{\mathbb{U}}$, i.e., that $u^j - u^i = 0$ on U^{ij}, and let $V^j = V \cap U^j$. The sets V^j form an affine covering $\mathbb{V}_0$ of the affine open set V. Let v_j denote the section obtained by restricting u_j to V^j. Since $u_j - u_i = 0$ on U^{ij}, it is also true that $v_j - v_i = 0$ on the smaller open set V^{ij}. So (v_i) is in the kernel of the map $\mathcal{M}(\mathbb{V}_0) \xrightarrow{\beta_{\mathbb{V}}} \mathcal{M}(\mathbb{V}_1)$. Since V is affine, Step 2 shows that the kernel is $\mathcal{M}(V)$. So there is a unique section v of $\mathcal{M}$ on V that restricts to v_j on V_j for all j. Then $((u_i), v)$ is in the kernel of $\beta_{\mathbb{W}}$, as required.

This completes the proof of Theorem 6.3.2. □

Chapter 7 COHOMOLOGY

7.1 Cohomology

To simplify the construction, we define cohomology only for $\mathcal{O}$-modules. Anyway, the Zariski topology has limited use for cohomology with other coefficients.

Let $\mathcal{M}$ be an $\mathcal{O}$-module on a variety X. The *zero-dimensional cohomology* of $\mathcal{M}$ is the space $\mathcal{M}(X)$ of its global sections. When speaking of cohomology, one denotes this space by $H^0(X,\mathcal{M})$.

The functor

$$(\mathcal{O}\text{-modules}) \xrightarrow{H^0} (\text{vector spaces})$$

that carries an $\mathcal{O}$-module $\mathcal{M}$ to $H^0(X,\mathcal{M})$ is left exact: If

(7.1.1) $$0 \to \mathcal{M} \to \mathcal{N} \to \mathcal{P} \to 0$$

is an exact sequence of $\mathcal{O}$-modules, the associated sequence of global sections

(7.1.2) $$0 \to H^0(X,\mathcal{M}) \to H^0(X,\mathcal{N}) \to H^0(X,\mathcal{P})$$

is exact. But unless X is affine, the map $H^0(X,\mathcal{N}) \to H^0(X,\mathcal{P})$ needn't be surjective. The *cohomology* of $\mathcal{M}$ is a sequence of functors *($\mathcal{O}$-modules)* $\xrightarrow{H^q}$ *(vector spaces)*,

$$H^0(X,\mathcal{M}),\ H^1(X,\mathcal{M}),\ H^2(X,\mathcal{M}),\ \dots$$

beginning with H^0, one for each *dimension*, that compensates for the lack of exactness in the following way: Every short exact sequence (7.1.1) of $\mathcal{O}$-modules has an associated long exact *cohomology sequence*

(7.1.3)
$$0 \to H^0(X,\mathcal{M}) \to H^0(X,\mathcal{N}) \to H^0(X,\mathcal{P}) \xrightarrow{\delta^0}$$
$$\xrightarrow{\delta^0} H^1(X,\mathcal{M}) \to H^1(X,\mathcal{N}) \to H^1(X,\mathcal{P}) \xrightarrow{\delta^1}$$
$$\cdots\cdots$$
$$\xrightarrow{\delta^{q-1}} H^q(X,\mathcal{M}) \to H^q(X,\mathcal{N}) \to H^q(X,\mathcal{P}) \xrightarrow{\delta^q} \cdots$$

And, given a diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathcal{M} & \longrightarrow & \mathcal{N} & \longrightarrow & \mathcal{P} & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & \mathcal{M}' & \longrightarrow & \mathcal{N}' & \longrightarrow & \mathcal{P}' & \longrightarrow & 0 \end{array}$$

whose rows are short exact sequences of $\mathcal{O}$-modules, the diagrams

(7.1.4)
$$\begin{array}{ccc} H^q(X,\mathcal{P}) & \xrightarrow{\delta^q} & H^{q+1}(X,\mathcal{M}) \\ \downarrow & & \downarrow \\ H^q(X,\mathcal{P}') & \xrightarrow{\delta^q} & H^{q+1}(X,\mathcal{M}') \end{array}$$

that are obtained from the map of cohomology sequences commute. The other diagrams commute because H^q are functors. Thus a map of short exact sequences induces a map of cohomology sequences. The maps δ^q are the *coboundary maps*.

A sequence H^q, $q = 0, 1, ...$ of functors from $\mathcal{O}$-modules to vector spaces that comes with long cohomology sequences for every short exact sequence (7.1.1) is called a *cohomological functor* .

Unfortunately, there is no really natural construction of the cohomology. Sometimes one needs to look at an explicit construction, but it is usually best to work with the characteristic properties that are described in the Section 7.3. We present a construction in Section 7.4, but it isn't canonical.

The one-dimensional cohomology H^1 has an interesting interpretation that you can read about if you like. We won't use it. The higher cohomology H^q has no useful direct interpretation.

7.2 Complexes

We need complexes because they are used in the definition of cohomology.

A *complex* $V^\bullet$ of vector spaces is a sequence of homomorphisms of vector spaces

(7.2.1)
$$\cdots \to V^{n-1} \xrightarrow{d^{n-1}} V^n \xrightarrow{d^n} V^{n+1} \xrightarrow{d^{n+1}} \cdots$$

indexed by the integers, such that the composition $d^n d^{n-1}$ of adjacent maps is zero – the image of d^{n-1} is contained in the kernel of d^n. The q-dimensional *cohomology* of a complex $V^\bullet$ is the quotient

(7.2.2)
$$\mathbf{C}^q(V^\bullet) = (\ker\, d^q)/(\mathrm{im}\, d^{q-1}).$$

An exact sequence is a complex whose cohomology is zero.

If a finite sequence of homomorphisms (7.2.1) is given, say $V^k \to V^{k+1} \cdots \to V^\ell$, it can be made into a complex by defining $V^n = 0$ for all other integers n. In our applications V^q will be zero when $q < 0$.

A homomorphism of vector spaces $V^0 \xrightarrow{d^0} V^1$ can be made into the complex

$$\cdots \to 0 \to V^0 \xrightarrow{d^0} V^1 \to 0 \to \cdots$$

For this complex, $\mathbf{C}^0 = \ker\, d^0$, $\mathbf{C}^1 = \mathrm{coker}\, d^0$, and $\mathbf{C}^q = 0$ for all other q.

A *map* $V^\bullet \xrightarrow{\varphi} V'^\bullet$ of complexes is a collection of homomorphisms $V^n \xrightarrow{\varphi^n} V'^n$ making a diagram

$$\begin{array}{ccccccccc} \longrightarrow & V^{n-1} & \xrightarrow{d^{n-1}} & V^n & \xrightarrow{d^n} & V^{n+1} & \longrightarrow & \cdots \\ & \downarrow{\scriptstyle \varphi^{n-1}} & & \downarrow{\scriptstyle \varphi^n} & & \downarrow{\scriptstyle \varphi^{n+1}} & & \\ \longrightarrow & V'^{n-1} & \xrightarrow{d'^{n-1}} & V'^n & \xrightarrow{d'^n} & V'^{n+1} & \longrightarrow & \cdots \end{array}$$

A map of complexes induces maps on the cohomology

$$\mathbf{C}^q(V^\bullet) \to \mathbf{C}^q(V'^\bullet)$$

because $\ker d^q$ maps to $\ker d'^q$ and $\operatorname{im} d^q$ maps to $\operatorname{im} d'^q$.

A sequence

(7.2.3) $$\cdots \to V^\bullet \xrightarrow{\varphi} V'^\bullet \xrightarrow{\psi} V''^\bullet \to \cdots$$

of maps of complexes is *exact* if the sequences

$$\cdots \to V^q \xrightarrow{\varphi^q} V'^q \xrightarrow{\psi^q} V''^q \to \cdots$$

are exact for every q.

7.2.4. Proposition.
Let $0 \to V^\bullet \to V'^\bullet \to V''^\bullet \to 0$ be a short exact sequence of complexes. For every q, there are maps $\mathbf{C}^q(V''^\bullet) \xrightarrow{\delta^q} \mathbf{C}^{q+1}(V^\bullet)$ such that the sequence

$$\to \mathbf{C}^0(V^\bullet) \to \mathbf{C}^0(V'^\bullet) \to \mathbf{C}^0(V''^\bullet) \xrightarrow{\delta^0} \mathbf{C}^1(V^\bullet) \to \mathbf{C}^1(V'^\bullet) \to \mathbf{C}^1(V''^\bullet) \xrightarrow{\delta^1} \mathbf{C}^2(V^\bullet) \to \quad \cdots$$

is exact.

The exact sequence displayed above is the *cohomology sequence* associated to the short exact sequence of complexes. This property makes the set of functors $\{\mathbf{C}^q\}$ into a *cohomological functor* on the category of complexes.

7.2.5. Example. We make the Snake Lemma into a cohomology sequence. Suppose given a diagram

$$\begin{array}{ccccccc} & & V & \xrightarrow{u} & V' & \longrightarrow & V'' & \longrightarrow 0 \\ & & {\scriptstyle f}\downarrow & & {\scriptstyle f'}\downarrow & & {\scriptstyle f''}\downarrow & \\ 0 & \longrightarrow & W & \longrightarrow & W' & \xrightarrow[v]{} & W'' & \end{array}$$

with exact rows. We form the complex $0 \to V \xrightarrow{f} W \to 0$ with V in degree zero, and we do the analogous thing for the maps f' and f'', so that the diagram becomes a short exact sequence of complexes. Its cohomology sequence is the one given by the Snake Lemma. □

proof of Proposition 7.2.4. Let

$$V^\bullet = \{\cdots \to V^{q-1} \xrightarrow{d^{q-1}} V^q \xrightarrow{d^q} V^{q+1} \xrightarrow{d^{q+1}} \cdots\}$$

be a complex, let B^q be the image of d^{q-1} in V^q, and let Z^q be the kernel of d^q. So $B^q \subset Z^q \subset V^q$, and the cohomology of the comples is $\mathbf{C}^q(V^\bullet) = Z^q/B^q$. Also, let D^q be the cokernel of d^{q-1}. So $D^q = V^q/B^q$, and there is an exact sequence

$$0 \to B^q \to V^q \to D^q \to 0$$

Again since $B^q \subset Z^q$, the map d^q can be written as the composition of three maps

$$V^q \xrightarrow{\pi^q} D^q \xrightarrow{f^q} Z^{q+1} \xrightarrow{i^{q+1}} V^{q+1}$$

where π^q is the projection from V^q to its quotient D^q and i^{q+1} is the inclusion of Z^{q+1} into V^{q+1}. Studying these maps, one sees that

(7.2.6) $$\mathbf{C}^q(V^\bullet) = \ker f^q \quad \text{and} \quad \mathbf{C}^{q+1}(V^\bullet) = \operatorname{coker} f^q.$$

Suppose given a short exact sequence of complexes $0 \to V^\bullet \to V'^\bullet \to V''^\bullet \to 0$ as in the proposition. In the diagram below, the rows are exact because cokernel is a right exact functor and kernel is a left exact functor.

$$\begin{array}{ccccccc} & & D^q & \longrightarrow & D'^q & \longrightarrow & D''^q & \longrightarrow 0 \\ & & {\scriptstyle f^q}\downarrow & & {\scriptstyle f'^q}\downarrow & & {\scriptstyle f''^q}\downarrow & \\ 0 & \longrightarrow & Z^{q+1} & \longrightarrow & Z'^{q+1} & \longrightarrow & Z''^{q+1} & \end{array}$$

When we apply (7.2.6) and the Snake Lemma to this diagram, we obtain an exact sequence

$$\mathbf{C}^q(V^\bullet) \to \mathbf{C}^q(V'^\bullet) \to \mathbf{C}^q(V''^\bullet) \xrightarrow{\delta^q} \mathbf{C}^{q+1}(V^\bullet) \to \mathbf{C}^{q+1}(V'^\bullet) \to \mathbf{C}^{q+1}(V''^\bullet)$$

The cohomology sequence is obtained by splicing these sequences together. □

The coboundary maps δ^q in cohomology sequences are related in a natural way. If

$$\begin{array}{ccccccccc} 0 & \longrightarrow & U^\bullet & \longrightarrow & U'^\bullet & \longrightarrow & U''^\bullet & \longrightarrow & 0 \\ & & \big\downarrow & & \big\downarrow & & \big\downarrow & & \\ 0 & \longrightarrow & V^\bullet & \longrightarrow & V'^\bullet & \longrightarrow & V''^\bullet & \longrightarrow & 0 \end{array}$$

is a diagram of maps of complexes whose rows are short exact sequences, the diagrams

$$\begin{array}{ccc} \mathbf{C}^q(U''^\bullet) & \xrightarrow{\delta^q} & \mathbf{C}^{q+1}(U^\bullet) \\ \big\downarrow & & \big\downarrow \\ \mathbf{C}^q(V''^\bullet) & \xrightarrow{\delta^q} & \mathbf{C}^{q+1}(V^\bullet) \end{array}$$

commute. Thus a map of short exact sequences induces a map of cohomology sequences.

7.3 Characteristic Properties of Cohomology

The cohomology $H^q(X, \cdot\,)$ of $\mathcal{O}$-modules, the sequence of functors $H^0(X, \cdot\,), H^1(X, \cdot\,), H^2(X, \cdot\,), \cdots$ from *($\mathcal{O}$-modules)* to (*vector spaces)*, is characterized by the three properties below, the first two of which have already been mentioned.

(7.3.1) characteristic properties

- $H^0(X, \mathcal{M})$ is the space $\mathcal{M}(X)$ of global sections of $\mathcal{M}$.
- The sequence $H^0, H^1, H^2, \cdots$ is a cohomological functor on $\mathcal{O}$-modules: A short exact sequence of $\mathcal{O}$-modules produces a long exact cohomology sequence.
- Let $Y \xrightarrow{f} X$ be the inclusion of an *affine* open subset Y into X, let $\mathcal{N}$ be an $\mathcal{O}_Y$-module, and let $f_*\mathcal{N}$ be its direct image on X. The cohomology $H^q(X, f_*\mathcal{N})$ is zero for all $q > 0$.

7.3.2. Example. Let j be the inclusion of the standard affine open set $\mathbb{U}^0$ into projective space X. The third property tells us that the cohomology $H^q(X, j_*\mathcal{O}_{\mathbb{U}^0})$ of the direct image $j_*\mathcal{O}\mathbb{U}^0$ is zero when $q > 0$. The direct image is isomorphic to the limit $\varinjlim \mathcal{O}_X(nH)$ (6.6.13). We will see below (7.4.28) that cohomology commutes with direct limits. Therefore the limits of $H^q(X, \mathcal{O}_X(nH))$ and of $H^q(X, \mathcal{O}_X(n))$ are zero when X is projective space and $q > 0$. This will be useful.

Intuitively, the third property tells us that allowing poles on the complement of an affine open set kills cohomology in positive dimension. □

7.3.3. Theorem. *There exists a cohomology theory with the properties (**7.3.1**), and it is unique up to unique isomorphism.*

The proof is in the next section.

7.3.4. Corollary. *If X is an affine variety, $H^q(X, \mathcal{M}) = 0$ for all $\mathcal{O}$-modules $\mathcal{M}$ and all $q > 0$.*

This follows when one applies the third characteristic property to the identity map $X \to X$. □

7.4 Existence of Cohomology

The proof of existence of cohomology and its uniqueness are based on tthe following facts:

• The intersection of two affine open subsets of a variety is an affine open set.

• A sequence $\cdots \to \mathcal{M} \to \mathcal{N} \to \mathcal{P} \to \cdots$ of $\mathcal{O}$-modules on a variety X is exact if and only if, for every affine open subset U, the sequence of sections $\cdots \to \mathcal{M}(U) \to \mathcal{N}(U) \to \mathcal{P}(U) \to \cdots$ is exact. (This is the definition of exactness of a sequence of $\mathcal{O}$-modules.)

We begin by choosing an arbitrary affine covering $\mathbb{U} = \{U^\nu\}$ of our variety X by finitely many affine open sets U^ν, and we use this covering to describe the cohomology. When we have shown that the cohomology is unique, we will know that it doesn't depend on our choice of covering.

Let $\mathbb{U}$ denote our chosen covering of X, and let $\mathbb{U} \xrightarrow{j} X$ denote the family of inclusions $U^\nu \xrightarrow{j^\nu} X$. If $\mathcal{M}$ is an $\mathcal{O}$-module, $\mathcal{R}_\mathcal{M}$ will denote the $\mathcal{O}$-module $j_*\mathcal{M}_\mathbb{U} = \prod j^\nu_* \mathcal{M}_{U^\nu}$, where $\mathcal{M}_{U^\nu}$ is the restriction of $\mathcal{M}$ to the open set U^ν. As has been noted (6.5.8), there is a canonical map $\mathcal{M} \to j^\nu_*\mathcal{M}_{U^\nu}$, and therefore a canonical map $\mathcal{M} \to \mathcal{R}_\mathcal{M}$.

7.4.1. Lemma. (i) *Let X' be an open subset of X. The module of sections $\mathcal{R}_\mathcal{M}(X')$ of $\mathcal{R}_\mathcal{M}$ on X' is is the product $\prod_\nu \mathcal{M}(X' \cap U^\nu)$. In particular, the space of global sections $\mathcal{R}_\mathcal{M}(X)$ is the product $\prod_\nu \mathcal{M}(U^\nu)$.*
(ii) *The canonical map $\mathcal{M} \to \mathcal{R}_\mathcal{M}$ is injective. Thus, if $\mathcal{S}_\mathcal{M}$ denotes the cokernel of that map, there is a short exact sequence of $\mathcal{O}$-modules*

(7.4.2)
$$0 \to \mathcal{M} \to \mathcal{R}_\mathcal{M} \to \mathcal{S}_\mathcal{M} \to 0$$

(iii) *For any cohomology theory with the characteristic properties and for any $q > 0$, $H^q(X, \mathcal{R}_\mathcal{M}) = 0$.*

proof. **(i)** This is seen by going through the definitions:

$$\mathcal{R}(X') \;=\; \prod\nolimits_\nu [j^\nu_*\mathcal{M}_{U^\nu}](X') = \prod\nolimits_\nu \mathcal{M}_{U^\nu}(X' \cap U^\nu) = \prod\nolimits_\nu \mathcal{M}(X' \cap U^\nu).$$

(ii) Let X' be an open subset of X. The map $\mathcal{M}(X') \to \mathcal{R}_\mathcal{M}(X')$ is the product of the restriction maps $\mathcal{M}(X') \to \mathcal{M}(X' \cap U^\nu)$. Because the open sets U^ν cover X, the intersections $X' \cap U^\nu$ cover X'. The sheaf property of $\mathcal{M}$ tells us that the map $\mathcal{M}(X') \to \prod_\nu \mathcal{M}(X' \cap U^\nu)$ is injective.

(iii) This follows from the third characteristic property. □

7.4.3. Lemma. (i) *A short exact sequence $0 \to \mathcal{M} \to \mathcal{N} \to \mathcal{P} \to 0$ of $\mathcal{O}$-modules embeds into a diagram*

(7.4.4)
$$\begin{array}{ccccc} \mathcal{M} & \longrightarrow & \mathcal{N} & \longrightarrow & \mathcal{P} \\ \downarrow & & \downarrow & & \downarrow \\ \mathcal{R}_\mathcal{M} & \longrightarrow & \mathcal{R}_\mathcal{N} & \longrightarrow & \mathcal{R}_\mathcal{P} \\ \downarrow & & \downarrow & & \downarrow \\ \mathcal{S}_\mathcal{M} & \longrightarrow & \mathcal{S}_\mathcal{N} & \longrightarrow & \mathcal{S}_\mathcal{P} \end{array}$$

whose rows and columns are short exact sequences. (We have suppressed the surrounding zeros.)
(ii) *The sequence of global sections $0 \to \mathcal{R}_\mathcal{M}(X) \to \mathcal{R}_\mathcal{N}(X) \to \mathcal{R}_\mathcal{P}(X) \to 0$ is exact.*

proof. **(i)** We are given that the top row of the diagram is a short exact sequence, and we have seen that the columns are short exact sequences. To show that the middle row

(7.4.5)
$$0 \to \mathcal{R}_\mathcal{M} \to \mathcal{R}_\mathcal{N} \to \mathcal{R}_\mathcal{P} \to 0$$

is exact, we must show that if X' is an affine open subset, the sections on X' form a short exact sequence. The sections are explained in Lemma 7.4.1 **(i)**. Since products of exact sequences are exact, we must show that the sequence

$$0 \to \mathcal{M}(X' \cap U^\nu) \to \mathcal{N}(X' \cap U^\nu) \to \mathcal{P}(X' \cap U^\nu) \to 0$$

is exact. This is true because $X' \cap U^\nu$ is an intersection of affine opens, and is therefore affine.

Now that we know that the first two rows of the diagram are short exact sequences, the Snake Lemma tells us that the bottom row of the diagram is a short exact sequence.

(ii) The sequence of of global sections is the product of the sequences

$$0 \to \mathcal{M}(U^\nu) \to \mathcal{N}(U^\nu) \to \mathcal{P}(U^\nu) \to 0$$

These sequences are exact because the open sets U^ν are affine. □

(7.4.6) uniqueness of cohomology

Suppose that a cohomology with the characteristic properties (**7.3.1**) is given, and let $\mathcal{M}$ be an $\mathcal{O}$-module. Then $H^q(X, \mathcal{R}_\mathcal{M}) = 0$ if $q > 0$ (Lemma 7.4.1 **(iii)**). The cohomology sequence associated to the sequence $0 \to \mathcal{M} \to \mathcal{R}_\mathcal{M} \to \mathcal{S}_\mathcal{M} \to 0$ is

$$0 \to H^0(X, \mathcal{M}) \to H^0(X, \mathcal{R}_\mathcal{M}) \to H^0(X, \mathcal{S}_\mathcal{M}) \xrightarrow{\delta^0} H^1(X, \mathcal{M}) \to H^1(X, \mathcal{R}_\mathcal{M}) \to \cdots$$

Since $H^q(X, \mathcal{R}_\mathcal{M}) = 0$ when $q > 0$, this sequence breaks up into an exact sequence

$$0 \to H^0(X, \mathcal{M}) \to H^0(X, \mathcal{R}_\mathcal{M}) \to H^0(X, \mathcal{S}_\mathcal{M}) \xrightarrow{\delta^0} H^1(X, \mathcal{M}) \to 0 \tag{7.4.7}$$

and isomorphisms

$$0 \to H^q(X, \mathcal{S}_\mathcal{M}) \xrightarrow{\delta^q} H^{q+1}(X, \mathcal{M}) \to 0 \tag{7.4.8}$$

for every $q > 0$. The first three terms of the sequence (7.4.7), and the arrows connecting them, depend on our choice of covering of X, but the important point is that they don't depend on the cohomology. So that sequence determines $H^1(X, \mathcal{M})$ up to unique isomorphism as the cokernel of a map that is independent of the cohomology, and this is true for every $\mathcal{O}$-module $\mathcal{M}$, including for the module $\mathcal{S}_\mathcal{M}$. Therefore it is also true that $H^1(X, \mathcal{S}_\mathcal{M})$ is determined uniquely. This being so, $H^2(X, \mathcal{M})$ is determined uniquely for every $\mathcal{M}$, by the isomorphism (7.4.8), with $q = 1$. The isomorphisms (7.4.8) determine the rest of the cohomology up to unique isomorphism by induction on q.

(7.4.9) construction of cohomology

One can use the sequence (7.4.2) and induction to construct cohomology as well as to prove uniqueness, but it will be clearer to proceed by iterating the construction of $\mathcal{R}_\mathcal{M}$.

Let $\mathcal{M}$ be an $\mathcal{O}$-module. We rewrite the exact sequence (7.4.2), labeling $\mathcal{R}_\mathcal{M}$ as $\mathcal{R}^0_\mathcal{M}$, and $\mathcal{S}_\mathcal{M}$ as $\mathcal{M}^1$:

$$0 \to \mathcal{M} \to \mathcal{R}^0_\mathcal{M} \to \mathcal{M}^1 \to 0 \tag{7.4.10}$$

and we repeat the construction with $\mathcal{M}^1$. Let $\mathcal{R}^1_\mathcal{M} = \mathcal{R}^0_{\mathcal{M}^1}$ $(= j_*\mathcal{M}^1_\mathbb{U})$, so that there is an exact sequence

$$0 \to \mathcal{M}^1 \to \mathcal{R}^1_\mathcal{M} \to \mathcal{M}^2 \to 0 \tag{7.4.11}$$

analogous to the sequence (7.4.10), with $\mathcal{M}^2 = \mathcal{R}^1_\mathcal{M}/\mathcal{M}^1$. We combine the sequences (7.4.10) and (7.4.11) into an exact sequence

$$0 \to \mathcal{M} \to \mathcal{R}^0_\mathcal{M} \to \mathcal{R}^1_\mathcal{M} \to \mathcal{M}^2 \to 0 \tag{7.4.12}$$

Then we let $\mathcal{R}^2_\mathcal{M} = \mathcal{R}^0_{\mathcal{M}^2}$. We continue in this way, to construct modules $\mathcal{R}^k_\mathcal{M}$ that form an exact sequence

$$0 \to \mathcal{M} \to \mathcal{R}^0_\mathcal{M} \to \mathcal{R}^1_\mathcal{M} \to \mathcal{R}^2_\mathcal{M} \to \cdots \tag{7.4.13}$$

The next lemma follows by induction from Lemmas 7.4.1 and 7.4.3.

7.4.14. Lemma.
(i) *Let* $0 \to \mathcal{M} \to \mathcal{N} \to \mathcal{P} \to 0$ *be a short exact sequence of* $\mathcal{O}$*-modules. For every* n*, the sequences*

$$0 \to \mathcal{R}^n_{\mathcal{M}} \to \mathcal{R}^n_{\mathcal{N}} \to \mathcal{R}^n_{\mathcal{P}} \to 0$$

are exact, and so are the sequences of global sections

$$0 \to \mathcal{R}^n_{\mathcal{M}}(X) \to \mathcal{R}^n_{\mathcal{N}}(X) \to \mathcal{R}^n_{\mathcal{P}}(X) \to 0$$

(ii) *If* H^q*,* $q = 0, 1, 2, ...$ *is a cohomology theory, then* $H^q(X, \mathcal{R}^n_{\mathcal{M}}) = 0$ *for all* n *and all* $q > 0$. □

An exact sequence such as (7.4.13) is called a *resolution* of $\mathcal{M}$, and because $H^q(X, \mathcal{R}^n_{\mathcal{M}}) = 0$ when $q > 0$, it is an *acyclic resolution*.

Continuing with the proof of existence, we consider the complex of $\mathcal{O}$-modules $\mathcal{R}^\bullet_{\mathcal{M}}$ that is obtained by omitting the first term from (7.4.13):

$$0 \to \mathcal{R}^0_{\mathcal{M}} \to \mathcal{R}^1_{\mathcal{M}} \to \mathcal{R}^2_{\mathcal{M}} \to \cdots \tag{7.4.15}$$

and the complex $\mathcal{R}^\bullet_{\mathcal{M}}(X)$ of its global sections:

$$0 \to \mathcal{R}^0_{\mathcal{M}}(X) \to \mathcal{R}^1_{\mathcal{M}}(X) \to \mathcal{R}^2_{\mathcal{M}}(X) \to \cdots \tag{7.4.16}$$

which we could also write as

$$0 \to H^0(X, \mathcal{R}^0_{\mathcal{M}}) \to H^0(X, \mathcal{R}^1_{\mathcal{M}}) \to H^0(X, \mathcal{R}^2_{\mathcal{M}}) \to \cdots$$

The sequence $\mathcal{R}^\bullet_{\mathcal{M}}$ becomes the resolution (7.4.13) when the module $\mathcal{M}$ is inserted. So the complex (7.4.15) is exact except at $\mathcal{R}^0_{\mathcal{M}}$, but because the global section functor is only left exact, the sequence (7.4.16) of global sections $\mathcal{R}^\bullet_{\mathcal{M}}(X)$ needn't be exact anywhere. However, $\mathcal{R}^\bullet_{\mathcal{M}}(X)$ is a complex because $\mathcal{R}^\bullet_{\mathcal{M}}$ is a complex. The composition of adjacent maps is zero.

Recall that the cohomology of a complex $0 \to V^0 \xrightarrow{d^0} V^1 \xrightarrow{d^1} \cdots$ of vector spaces is defined to be $\mathbf{C}^q(V^\bullet) = (\ker\, d^q)/(\mathrm{im}\, d^{q-1})$, and that $\{\mathbf{C}^q\}$ is a cohomological functor on complexes (7.2.4).

7.4.17. Definition. The cohomology of an $\mathcal{O}$-module $\mathcal{M}$ is the cohomology of the complex $\mathcal{R}^\bullet_{\mathcal{M}}(X)$: $H^q(X, \mathcal{M}) = \mathbf{C}^q(\mathcal{R}^\bullet_{\mathcal{M}}(X))$.

Thus if we denote the maps in the complex (7.4.16) by d^q,

$$0 \to \mathcal{R}^0_{\mathcal{M}}(X) \xrightarrow{d^0} \mathcal{R}^1_{\mathcal{M}}(X) \xrightarrow{d^1} \mathcal{R}^2_{\mathcal{M}}(X) \to \cdots$$

then $H^q(X, \mathcal{M}) = (\ker\, d^q)/(\mathrm{im}\, d^{q-1})$.

7.4.18. Lemma. *Let* X *be an affine variety. With cohomology defined as above,* $H^q(X, \mathcal{M}) = 0$ *for all* $\mathcal{O}$*-modules* $\mathcal{M}$ *and all* $q > 0$.

proof. When X is affine, the sequence of global sections of the exact sequence (7.4.13) is exact. □

To show that our definition gives the (unique) cohomology, we verify the characteristic properties. Since the sequence (7.4.13) is exact and since the global section functor is left exact, $\mathcal{M}(X)$ is the kernel of the map $\mathcal{R}^0_{\mathcal{M}}(X) \to \mathcal{R}^1_{\mathcal{M}}(X)$, and this kernel is also equal to $\mathbf{C}^0(\mathcal{R}^\bullet_{\mathcal{M}}(X))$. So our cohomology has the first property: $H^0(X, \mathcal{M}) = \mathcal{M}(X)$.

To show that we obtain a cohomological functor, we apply Lemma 7.4.14 to conclude that, for a short exact sequence $0 \to \mathcal{M} \to \mathcal{N} \to \mathcal{P} \to 0$, the global sections

$$0 \to \mathcal{R}^\bullet_{\mathcal{M}}(X) \to \mathcal{R}^\bullet_{\mathcal{N}}(X) \to \mathcal{R}^\bullet_{\mathcal{P}}(X) \to 0, \tag{7.4.19}$$

form an exact sequence of complexes. Cohomology $H^q(X, \cdot\,)$ is a cohomological functor because cohomology of complexes is a cohomological functor.

We make a digression before verifying the third characteristic property.

(**7.4.20**) **affine morphisms**

Let $Y \xrightarrow{f} X$ be a morphism of varieties. Let $U \xrightarrow{j} X$ be the inclusion of an open subvariety into X and let V be the inverse image $f^{-1}U$, which is an open subvariety of Y. These varieties and maps form a diagram

(7.4.21)
$$\begin{array}{ccc} V & \xrightarrow{i} & Y \\ {\scriptstyle g}\downarrow & & {\scriptstyle f}\downarrow \\ U & \xrightarrow{j} & X \end{array}$$

We use the notation $\mathcal{M}_U$ for the restriction of $\mathcal{M}$ to an open subset U of(**6.5.6**).

7.4.22. Lemma. *With notation as above, let $\mathcal{N}$ be an $\mathcal{O}_Y$-module. The $\mathcal{O}_U$-modules $g_*[\mathcal{N}_V]$ and $[f_*\mathcal{N}]_U$ are canonically isomorphic.*

proof. Let U' be an open subset of U, and let $V' = g^{-1}U'$. Then

$$[f_*\mathcal{N}]_U(U') = [f_*\mathcal{N}](U') = \mathcal{N}(V') = \mathcal{N}_V(V') = [g_*[\mathcal{N}_V]](U') \qquad \square$$

7.4.23. Definition. An *affine morphism* is a morphism $Y \xrightarrow{f} X$ of varieties with the property that the inverse image $f^{-1}(U)$ of every affine open subset U of X is an affine open subset of Y. $\square$

The following are examples of affine morphisms:

- the inclusion of an affine open subset Y into X,
- the inclusion of a closed subvariety Y into X,
- a finite morphism, or an integral morphism.

But, if Y is a closed subset of $\mathbb{P}^n \times X$, the projection $Y \to X$ will not be an affine morphism unless its fibres are finite, in which case Chevalley's Finitenss Theorem tells us that it is a finite morphism.

7.4.24. Lemma. *If $Y \xrightarrow{f} C$ is an affine morphism and if $\mathcal{N} \to \mathcal{N}' \to \mathcal{N}''$ is an exact sequence of $\mathcal{O}_Y$-modules, the sequence of direct images $f_*\mathcal{N} \to f_*\mathcal{N}' \to f_*\mathcal{N}''$ is exact.* $\square$

Let $Y \xrightarrow{f} X$ be an affine morphism, let j be the map from our chosen affine covering $\mathbb{U} = \{U^\nu\}$ to X, and let $\mathbb{V}$ denote the family $\{V^\nu\} = \{f^{-1}U^\nu\}$ of inverse images. Then $\mathbb{V}$ is an affine covering of Y, and there is a morphism $\mathbb{V} \xrightarrow{g} \mathbb{U}$. We form a diagram analogous to (7.4.21), in which $\mathbb{V}$ and $\mathbb{U}$ replace V and U, respectively:

$$\begin{array}{ccc} \mathbb{V} & \xrightarrow{i} & Y \\ {\scriptstyle g}\downarrow & & {\scriptstyle f}\downarrow \\ \mathbb{U} & \xrightarrow{j} & X \end{array}$$

7.4.25. Proposition. *Let $Y \xrightarrow{f} X$ be an affine morphism, and let $\mathcal{N}$ be an $\mathcal{O}_Y$-module. Let $H^q(X, \cdot\,)$ be cohomology defined in (7.4.17), and let $H^q(Y, \cdot\,)$ be cohomology defined in the analogous way, using the covering $\mathbb{V}$ of Y. Then $H^q(X, f_*\mathcal{N})$ is isomorphic to $H^q(Y, \mathcal{N})$.*

proof. To compute the cohomology of $f_*\mathcal{N}$ on X, we substitute $\mathcal{M} = f_*\mathcal{N}$ into (7.4.17):

$$H^q(X, f_*\mathcal{N}) = \mathbf{C}^q(\mathcal{R}^\bullet_{f_*\mathcal{N}}(X)).$$

To compute the cohomology of $\mathcal{N}$ on Y, we let

$$\mathcal{R}'^0_{\mathcal{N}} = i_*[\mathcal{N}_V]$$

and we continue, to construct a resolution $\mathcal{R}'^{\bullet}_{\mathcal{N}} \;=\; 0 \to \mathcal{N} \to \mathcal{R}'^{0}_{\mathcal{N}} \to \mathcal{R}'^{1}_{\mathcal{N}} \to \cdots$ and the complex of its global sections $\mathcal{R}'^{\bullet}_{\mathcal{N}}(Y)$. (The prime is there to remind us that $\mathcal{R}'$ is defined using the covering $\mathbb{V}$ of Y.) Then

$$H^q(Y,\mathcal{N}) = \mathbf{C}^q(\mathcal{R}'^{\bullet}_{\mathcal{N}}(Y)).$$

It suffices to show that the complexes $\mathcal{R}^{\bullet}_{f_*\mathcal{N}}(X)$ and $\mathcal{R}'^{\bullet}_{\mathcal{N}}(Y)$ are isomorphic, and because $\mathcal{R}'^{q}_{\mathcal{N}}(Y) = [f_*\mathcal{R}'^{q}_{\mathcal{N}}](X)$, it suffices to show that $\mathcal{R}^{q}_{f_*\mathcal{N}} \approx f_*\mathcal{R}'^{q}_{\mathcal{N}}$.

##reread this##

We look back at the definition (7.4.11) of the $\mathcal{O}_X$-modules $\mathcal{R}^0$. We have $\mathcal{R}']^{0}_{\mathcal{N}} = i_*\mathcal{N}_{\mathbb{V}}$. So the sequence for $\mathcal{N}$ analogous to (7.4.10) can be written as

$$0 \to \mathcal{N} \to i_*\mathcal{N}_{\mathbb{V}} \to \mathcal{N}^1 \to 0$$

and since $fi = jg$, its direct image can be written as

(7.4.26) $$0 \to f_*\mathcal{N} \to j_*g_*[\mathcal{N}_{\mathbb{V}}] \to f_*[\mathcal{N}^1] \to 0$$

The sequence for $f_*\mathcal{N}$ analogous to (7.4.10) is

$$0 \to f_*\mathcal{N} \to j_*[f_*\mathcal{N}]_{\mathbb{U}} \to [f_*\mathcal{N}]^1 \to 0$$

According to Lemma 7.4.22, $[f_*\mathcal{N}]_{\mathbb{U}}$ is isomorphic to $g_*[\mathcal{N}_{\mathbb{V}}]$. So this sequence can also be written as

(7.4.27) $$0 \to f_*\mathcal{N} \to j_*g_*[\mathcal{N}_{\mathbb{V}}] \to [f_*\mathcal{N}]^1 \to 0$$

Combining reffstarN) and (7.4.27), one sees that $\mathcal{R}^{0}_{f_*\mathcal{N}} \approx f_*\mathcal{R}'^{0}_{\mathcal{N}}$ and that $f_*[\mathcal{N}^1] \approx [f_*\mathcal{N}]^1$. Then induction applies. □

We go back to the proof of existence of cohomology to verify the third characteristic property, that when $Y \xrightarrow{f} X$ is the inclusion of an affine open subset, $H^q(X, f_*\mathcal{N}) = 0$ for all $\mathcal{O}_Y$-modules $\mathcal{N}$ and all $q > 0$. The inclusion of an affine open set is an affine morphism, so $H^q(Y,\mathcal{N}) = H^q(X, f_*\mathcal{N})$ (7.4.25), and since Y is affine, $H^q(Y,\mathcal{N}) = 0$ for all $q > 0$ (7.4.18). □

Proposition 7.4.25 is one of the places where a specific construction of cohomology is used. The characteristic properties don't apply directly. The next proposition is another such place.

7.4.28. Lemma. *Cohomology is compatible with limits of directed sets of $\mathcal{O}$-modules: $H^q(X, \varinjlim \mathcal{M}_\bullet) \approx \varinjlim H^q(X, \mathcal{M}_\bullet)$ for all q.*

proof. The direct and inverse image functors and the global section functor are all compatible with $\varinjlim$, and $\varinjlim$ is exact (**??**). So the module $\mathcal{R}^{q}_{\varinjlim \mathcal{M}_\bullet}$ that is used to compute the cohomology of $\varinjlim \mathcal{M}_\bullet$ is isomorphic to $\varinjlim [\mathcal{R}^{q}_{\mathcal{M}_\bullet}]$ and $\mathcal{R}^{q}_{\varinjlim \mathcal{M}_\bullet}(X)$ is isomorphic to $\varinjlim [\mathcal{R}^{q}_{\mathcal{M}_\bullet}](X)$. □

(7.4.29) uniqueness of the coboundary maps

We have constructed a cohomology $\{H^q\}$ that has the characteristic properties, and we have shown that the functors H^q are unique. We haven't shown that the coboundary maps δ^q that appear in the cohomology sequences (7.1.3) are unique. To make it clear that there is something to show, we note that the cohomology sequence (7.1.3) remains exact when some of the coboundary maps δ^q are multiplied by -1. Why can't we define a new collection of coboundary maps by changing some signs? The reason we can't do this is that we used the coboundary maps δ^q in (7.4.7) and (7.4.8) to identify $H^q(X,\mathcal{M})$. Having done that, we aren't allowed to change δ^q for the particular short exact sequences (7.4.2). We show that the coboundary maps for those particular sequences determine the coboundary maps for every short exact sequence of $\mathcal{O}$-modules

$$(A) \qquad 0 \to \mathcal{M} \longrightarrow \mathcal{N} \longrightarrow \mathcal{P} \to 0$$

The sequences (7.4.2) were rewritten as (7.4.10). We will use that form.

To show that the coboundaries for the sequence (A) are determined uniquely, we relate it to a sequence (B) for which the coboundary maps are fixed:

$$(B) \qquad 0 \to \mathcal{M} \longrightarrow \mathcal{R}^0_{\mathcal{M}} \longrightarrow \mathcal{M}^1 \to 0$$

We map (A) and (B) to a third short exact sequence

$$(C) \qquad 0 \to \mathcal{M} \xrightarrow{\psi} \mathcal{R}^0_{\mathcal{N}} \longrightarrow \mathcal{Q} \to 0$$

where ψ is the composition of the injective maps $\mathcal{M} \to \mathcal{R}^0_{\mathcal{M}} \to \mathcal{R}^0_{\mathcal{N}}$ and $\mathcal{Q}$ is the cokernel of ψ.

First, we inspect the diagram

$$\begin{array}{cccccc} (A) & \mathcal{M} & \longrightarrow & \mathcal{N} & \longrightarrow & \mathcal{P} \\ & \| & & \downarrow & & \downarrow \\ (C) & \mathcal{M} & \xrightarrow{\psi} & \mathcal{R}^0_{\mathcal{N}} & \longrightarrow & \mathcal{Q} \end{array}$$

and its diagram of coboundary maps

$$\begin{array}{cccc} (A) & H^q(X,\mathcal{P}) & \xrightarrow{\delta^q_A} & H^{q+1}(X,\mathcal{M}) \\ & \downarrow & & \| \\ (C) & H^q(X,\mathcal{Q}) & \xrightarrow{\delta^q_C} & H^{q+1}(X,\mathcal{M}) \end{array}$$

This diagram shows that the coboundary map δ^q_A for the sequence (A) is determined by the coboundary map δ^q_C for (C).

Next, we inspect the diagram

$$\begin{array}{cccccc} (B) & \mathcal{M} & \longrightarrow & \mathcal{R}^0_{\mathcal{M}} & \longrightarrow & \mathcal{M}^1 \\ & \| & & {\scriptstyle u}\downarrow & & {\scriptstyle v}\downarrow \\ (C) & \mathcal{M} & \xrightarrow{\psi} & \mathcal{R}^0_{\mathcal{N}} & \longrightarrow & \mathcal{Q} \end{array} \tag{7.4.30}$$

and its diagram of coboundary maps

$$\begin{array}{cccc} (B) & H^q(X,\mathcal{M}^1) & \xrightarrow{\delta^q_B} & H^{q+1}(X,\mathcal{M}) \\ & {\scriptstyle v}\downarrow & & \| \\ (C) & H^q(X,\mathcal{Q}) & \xrightarrow{\delta^q_C} & H^{q+1}(X,\mathcal{M}) \end{array}$$

When $q > 0$, δ^q_C and δ^q_B are bijective because the cohomology of $\mathcal{R}^0_{\mathcal{M}}$ and $\mathcal{R}^0_{\mathcal{N}}$ is zero in positive dimension. Then δ^q_C is uniquely determined by δ^q_B, and so is δ^q_A.

We have to look more closely to settle the case $q = 0$. The maps labeled u and v in (7.4.30) are injective, and the Snake Lemma shows that their cokernels are isomorphic. We write both of them as $\mathcal{R}^0_{\mathcal{P}}$. When we add the cokernels to the diagram, we obtain a cohomology diagram whose relevant part is

$$\begin{array}{cccccc} (B) & H^0(X,\mathcal{R}^0_{\mathcal{M}}) & \longrightarrow & H^0(X,\mathcal{M}^1) & \xrightarrow{\delta^0_B} & H^1(X,\mathcal{M}) \\ & {\scriptstyle u}\downarrow & & \downarrow{\scriptstyle v} & & \| \\ (C) & H^0(X,\mathcal{R}^0_{\mathcal{N}}) & \xrightarrow{\beta} & H^0(X,\mathcal{Q}) & \xrightarrow{\delta^0_C} & H^1(X,\mathcal{M}) \\ & \downarrow{\scriptstyle \gamma} & & \downarrow & & \\ & H^0(X,\mathcal{R}^0_{\mathcal{P}}) & = & H^0(X,\mathcal{R}^0_{\mathcal{P}}) & & \end{array}$$

The rows and columns in the diagram are exact. We want to show that the map δ^0_C is determined uniquely by δ^0_B. It is determined by δ^0_B on the image of v and it is zero on the image of β. To show that δ^0_C is determined by δ^0_B, it suffices to show that the images of v and β together span $H^0(X,\mathcal{Q})$. This follows from the fact that γ is surjective. Thus δ^0_C is determined uniqely by δ^0_B, and so is δ^0_A. $\square$

7.5 Cohomology of the Twisting Modules

We determine the cohomology of the twisting modules $\mathcal{O}(d)$ on $\mathbb{P}^n$ here. As we will see, $H^q(\mathbb{P}^n, \mathcal{O}(d))$ is zero for most values of q. This will help to determine the cohomology of other modules.

Lemma 7.4.18 about vanishing of cohomology on an affine variety, and Lemma 7.4.25 about the direct image via an affine morphism, were stated using a particular affine covering. Since we know that cohomology is unique, that particular covering is irrelevant. Though it isn't necessary, we restate the lemmas here as a corollary:

7.5.1. Corollary. (i) *On an affine variety X, $H^q(X, \mathcal{M}) = 0$ for all $\mathcal{O}$-modules $\mathcal{M}$ and all $q > 0$.*

(ii) *Let $Y \xrightarrow{f} X$ be an affine morphism. If $\mathcal{N}$ is an $\mathcal{O}_Y$-module, then $H^q(X, f_*\mathcal{N})$ and $H^q(Y, \mathcal{N})$ are isomorphic. If Y is an affine variety, $H^q(X, f_*\mathcal{N}) = 0$ for all $q > 0$.* □

One case in which **(ii)** applies is that f is the inclusion of a closed subvariety Y into X.

Let $\mathcal{M}$ be a finite $\mathcal{O}$-module on projective space $\mathbb{P}^n$. The twisting modules $\mathcal{O}(d)$ and the twists $\mathcal{M}(d) = \mathcal{M} \otimes_{\mathcal{O}} \mathcal{O}(d)$ are isomorphic to the modules $\mathcal{O}(dH)$ and $\mathcal{M}(dH) = \mathcal{M} \otimes_{\mathcal{O}} \mathcal{O}(dH)$, respectively. They form maps of directed sets

$$\begin{array}{ccccccc} \mathcal{O} & \xrightarrow{\subset} & \mathcal{O}(H) & \xrightarrow{\subset} & \mathcal{O}(2H) & \xrightarrow{\subset} & \cdots \\ {\scriptstyle 1}\downarrow & & {\scriptstyle x_0}\downarrow & & {\scriptstyle x_0^2}\downarrow & & \\ \mathcal{O} & \xrightarrow{x_0} & \mathcal{O}(1) & \xrightarrow{x_0} & \mathcal{O}(2) & \xrightarrow{x_0} & \cdots \end{array} \quad , \quad \begin{array}{ccccccc} \mathcal{M} & \longrightarrow & \mathcal{M}(H) & \longrightarrow & \mathcal{M}(2H) & \longrightarrow & \cdots \\ {\scriptstyle 1}\downarrow & & {\scriptstyle x_0}\downarrow & & {\scriptstyle x_0^2}\downarrow & & \\ \mathcal{M} & \xrightarrow{x_0} & \mathcal{M}(1) & \xrightarrow{x_0} & \mathcal{M}(2) & \xrightarrow{x_0} & \cdots \end{array}$$

(See (**??**)). The second diagram is obtained from the first one by tensoring with $\mathcal{M}$. Let $\mathbb{U}$ denote the stanard affine open subset $\mathbb{U}^0$ of $\mathbb{P}^n$, and let j be the inclusion of $\mathbb{U}$ into $\mathbb{P}^n$. Then $\varinjlim \mathcal{O}(dH) \approx j_*\mathcal{O}_{\mathbb{U}}$ (**??**) and because $\varinjlim$ is compatible with tensor products, $\varinjlim \mathcal{M}(dH) \approx j_*\mathcal{M}_{\mathbb{U}}$. Since j is an affine morphism and $\mathbb{U}^0$ is an affine open set, $H^q(\mathbb{P}^n, j_*\mathcal{O}_{\mathbb{U}}) = 0$ and $H^q(\mathbb{P}^n, j_*\mathcal{M}_{\mathbb{U}}) = 0$ for all $q > 0$.

The next corollary follows from the facts that $\mathcal{M}(d)$ is isomorphic to $\mathcal{M}(dH)$, and that cohomology is compatible with direct limits (7.4.28).

7.5.2. Corollary. *For all projective varieties X and all $\mathcal{O}$-modules $\mathcal{M}$, $\varinjlim H^q(X, \mathcal{O}(d)) = 0$ and $\varinjlim H^q(X, \mathcal{M}(d)) = 0$ when $q > 0$.* □

7.5.3. Notation. If $\mathcal{M}$ is an $\mathcal{O}$-module, we denote the dimension of $H^q(X, \mathcal{M})$ by $\mathbf{h}^q(\mathcal{M})$ or by $\mathbf{h}^q(X, \mathcal{M})$. We can write $\mathbf{h}^q(\mathcal{M}) = \infty$ if the dimension is infinite. However, in Section 7.7, we will see that when $\mathcal{M}$ is a finite $\mathcal{O}$-module on a projective variety X, $H^q(X, \mathcal{M})$ has finite dimension for every q.

7.5.4. Theorem.
(i) *For $d \geq 0$, $\mathbf{h}^0(\mathbb{P}^n, \mathcal{O}(d)) = \binom{d+n}{n}$ and $\mathbf{h}^q(\mathbb{P}^n, \mathcal{O}(d)) = 0$ if $q \neq 0$.*
(ii) *For $r > 0$, $\mathbf{h}^n(\mathbb{P}^n, \mathcal{O}(-r)) = \binom{r-1}{n}$ and $\mathbf{h}^q(\mathbb{P}^n, \mathcal{O}(-r)) = 0$ if $q \neq n$.*

In particular, part **(ii)** implies that $\mathbf{h}^q(\mathbb{P}^n, \mathcal{O}(-1)) = 0$ for all q.

proof. We have described the global sections of $\mathcal{O}(d)$ before: If $d \geq 0$, $H^0(X, \mathcal{O}(d))$ is the space of homogeneous polynomials of degree d in the coordinate variables, and if $d < 0$, $H^0(X, \mathcal{O}(d)) = 0$ (see (6.6.2)).

(i) (*the case $d \geq 0$*)

Let $X = \mathbb{P}^n$, and let $Y \xrightarrow{i} X$ be the inclusion of the hyperplane at infinity into X. By induction on n, we may assume that the theorem has been proved for Y, which is a projective space of dimension $n-1$. We consider the exact sequence

(7.5.5) $$0 \to \mathcal{O}_X(-1) \xrightarrow{x_0} \mathcal{O}_X \to i_*\mathcal{O}_Y \to 0$$

and its twists

(7.5.6) $$0 \to \mathcal{O}_X(d-1) \xrightarrow{x_0} \mathcal{O}_X(d) \to i_*\mathcal{O}_Y(d) \to 0$$

The twisted sequences are exact because they are obtained by tensoring (7.5.5) with the invertible $\mathcal{O}$-modules $\mathcal{O}(d)$. Because the inclusion i of Y into X is an affine morphism, $H^q(X, i_*\mathcal{O}_Y(d)) \approx H^q(Y, \mathcal{O}_Y(d))$.

The monomials of degree d in $n+1$ variables form a basis of the space of global sections of $\mathcal{O}_X(d)$. Setting $x_0 = 0$ and deleting terms that become zero gives us a basis of $\mathcal{O}_Y(d)$. Therefore every global section of $\mathcal{O}_Y(d)$ is the restriction of a global section of $\mathcal{O}_X(d)$. The sequence of global sections

$$0 \to H^0(X, \mathcal{O}_X(d-1)) \xrightarrow{x_0} H^0(X, \mathcal{O}_X(d)) \to H^0(Y, \mathcal{O}_Y(d))$$

is exact, and it remains exact when a zero is added on the right. This tells us that the map

$$H^1(X, \mathcal{O}_X(d-1)) \longrightarrow H^1(X, \mathcal{O}_X(d))$$

is injective. By induction on n, $H^q(Y, \mathcal{O}_Y(d)) = 0$ for $d \geq 0$ and $q > 0$. When combined with the injectivity noted above, the cohomology sequence of (7.5.6) gives us bijections $H^q(X, \mathcal{O}_X(d-1)) \to H^q(X, \mathcal{O}_X(d))$ for every $q > 0$. Since the limits are zero (7.5.2), $H^q(X, \mathcal{O}_X(d)) = 0$ for all $d \geq 0$ and all $q > 0$.

(ii) (*the case* $d < 0$)
We use induction on the integers r and n. We suppose the theorem proved for r, and we substitute $d = -r$ into the sequence (7.5.6):

(7.5.7)
$$0 \to \mathcal{O}_X(-(r+1)) \xrightarrow{x_0} \mathcal{O}_X(-r) \to i_*\mathcal{O}_Y(-r) \to 0$$

The base case $r = 0$ is the exact sequence (7.5.5). In the cohomology sequence associated to that sequence, the terms $H^q(X, \mathcal{O}_X)$ and $H^q(Y, \mathcal{O}_Y)$ are zero when $q > 0$, and $H^0(X, \mathcal{O}_X) = H^0(Y, \mathcal{O}_Y) = \mathbb{C}$. Therefore

(7.5.8)
$$H^q(X, \mathcal{O}_X(-1)) = 0 \ \text{ for every } \ q.$$

This proves **(ii)** for $r = 1$.

Our induction hypothesis is that, $\mathbf{h}^n(\mathbb{P}^n, \mathcal{O}(-r)) = \binom{r-1}{n}$ and $\mathbf{h}^q = 0$ if $q \neq n$. By induction on n, we may suppose that $\mathbf{h}^{n-1}(\mathbb{P}^{n-1}, \mathcal{O}(-r)) = \binom{r-1}{n-1}$ and $\mathbf{h}^q = 0$ if $q \neq n-1$. Instead of displaying the cohomology sequence associated to (7.5.7), we assemble the dimensions of cohomology into a table in which the asterisks stand for entries that are to be determined:

(7.5.9)

	$\mathcal{O}_X(-(r+1))$	$\mathcal{O}_X(-r)$	$i_*\mathcal{O}_Y(-r)$
$\mathbf{h}^0$:	$*$	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathbf{h}^{n-2}$:	$*$	0	0
$\mathbf{h}^{n-1}$:	$*$	0	$\binom{r-1}{n-1}$
$\mathbf{h}^n$:	$\mathbf{h}^n(\mathcal{O}(-(r+1))$	$\binom{r-1}{n}$	0

The second column is determined by induction on r and the third by induction on n. The cohomology sequence shows that the entries labeled with an asterisk are zero, and that

$$\mathbf{h}^n(\mathbb{P}^n, \mathcal{O}(-(r+1))) = \binom{r-1}{n-1} + \binom{r-1}{n}$$

The right side of this equation is equal to $\binom{r}{n}$. □

7.6 Cohomology of Hypersurfaces

We determine the cohomology of a plane projective curve first. Let $X = \mathbb{P}^2$ and let $C \xrightarrow{i} X$ denote the inclusion of a plane curve of degree k. The ideal $\mathcal{I}$ of functions that vanish on C is isomorphic to the twisting module $\mathcal{O}_X(-k)$ (6.6.8), so one has an exact sequence

(7.6.1)
$$0 \to \mathcal{O}_X(-k) \to \mathcal{O}_X \to i_*\mathcal{O}_C \to 0$$

We form a table showing dimensions of the cohomology. Theorem 7.5.4 determines the first two columns, and the cohomology sequence determines the last column.

(7.6.2)

	$\mathcal{O}_X(-k)$	$\mathcal{O}_X$	$i_*\mathcal{O}_C$
$\mathbf{h}^0$:	0	1	1
$\mathbf{h}^1$:	0	0	$\binom{k-1}{2}$
$\mathbf{h}^2$:	$\binom{k-1}{2}$	0	0

Since the inclusion of the curve C into the projective plane X is an affine morphism, $\mathbf{h}^q(X, i_*\mathcal{O}_C) = \mathbf{h}^q(C, \mathcal{O}_C)$. Therefore

$$\mathbf{h}^0(C, \mathcal{O}_C) = 1, \ \ \mathbf{h}^1(C, \mathcal{O}_C) = \binom{k-1}{2}, \ \ \text{and} \ \ \mathbf{h}^q = 0 \ \ \text{when} \ \ q > 1.$$

The dimension $\mathbf{h}^1(C, \mathcal{O}_C)$, which is $\binom{k-1}{2}$, is called the *arithmetic genus* of C. It is denoted by p_a. We will see later (8.9.2) that when C is a smooth curve, its arithmetic genus is equal to its topological genus: $p_a = g$, but the arithmetic genus of a plane curve of degree k is $\binom{k-1}{2}$ also when C is singular.

We restate the results as a corollary.

7.6.3. Corollary. *Let C be a plane curve of degree k. Then $\mathbf{h}^0(C, \mathcal{O}_C) = 1$, $\mathbf{h}^1(C, \mathcal{O}_C) = \binom{k-1}{2} = p_a$, and $\mathbf{h}^q = 0$ if $q \neq 0, 1$.* □

The fact that $\mathbf{h}^0(C, \mathcal{O}_C) = 1$ tells us that the only rational functions that are regular everywhere on C are the constants. This reflects a fact that will be proved later: A plane curve is compact and connected in the classical topology. However, it isn't a proof of that fact.

We will need more technique in order to compute cohomology for curves in higher dimensional projective spaces. In the next section we will see that the cohomology on any projective curve is zero except in dimensions 0 and 1. Cohomology of projective curves is the topic of Chapter 8.

One can make a similar computation for the hypersurface Y in $X = \mathbb{P}^n$ defined by an irreducible homogeneous polynomial f of degree k. The ideal of Y is isomorphic to $\mathcal{O}_X(-k)$, and there is an exact sequence

$$0 \to \mathcal{O}_X(-k) \xrightarrow{f} \mathcal{O}_X \to i_*\mathcal{O}_Y \to 0$$

Since we know the cohomology of $\mathcal{O}_X(-k)$ and $\mathcal{O}_X$, and since $H^q(X, i_*\mathcal{O}_Y) \approx H^q(Y, \mathcal{O}_Y)$, we can use this sequence to compute the dimensions of the cohomology of $\mathcal{O}_Y$.

7.6.4. Corollary. *Let Y be a hypersurface of dimension d and degree k in a projective space of dimension $d+1$. Then $\mathbf{h}^0(Y, \mathcal{O}_Y) = 1$, $\mathbf{h}^d(Y, \mathcal{O}_Y) = \binom{k-1}{d+1}$, and $\mathbf{h}^q(Y, \mathcal{O}_Y) = 0$ for all other q.* □

If S is a surface in $\mathbb{P}^3$ defined by an irreducible polynomial of degree k, then $\mathbf{h}^0(S, \mathcal{O}_S) = 1$, $\mathbf{h}^1(S, \mathcal{O}_S) = 0$, $\mathbf{h}^2(S, \mathcal{O}_S) = \binom{k-1}{3}$, and $\mathbf{h}^q = 0$ if $q > 2$. When a projective surface S isn't embedded into $\mathbb{P}^3$, it is still true that $\mathbf{h}^q = 0$ when $q > 2$, but $\mathbf{h}^1(S, \mathcal{O}_S)$ may be nonzero. The dimensions $\mathbf{h}^1(S, \mathcal{O}_S)$ and $\mathbf{h}^2(S, \mathcal{O}_S)$ are invariants of the surface somewhat analogous to the genus of a curve. In classical terminology, $\mathbf{h}^2(S, \mathcal{O}_S)$ is the *geometric genus* p_g and $\mathbf{h}^1(S, \mathcal{O}_S)$ is the *irregularity* q . The *arithmetic genus* p_a is

(7.6.5)
$$p_a \ = \ \mathbf{h}^2(S, \mathcal{O}_S) - \mathbf{h}^1(S, \mathcal{O}_S) = p_g - q$$

Therefore the irregularity is $q = p_g - p_a$. When S is a surface in $\mathbb{P}^3$, $q = 0$ and $p_g = p_a$.

In modern terminology, it is more natural to replace the arithmetic genus by the Euler characteristic $\chi(\mathcal{O}_S) = \sum_q (-1)^q \mathbf{h}^q(\mathcal{O}_S)$. The Euler characteristic of a curve is

$$\chi(\mathcal{O}_C) = \mathbf{h}^0(C, \mathcal{O}_C) - \mathbf{h}^1(C, \mathcal{O}_C) = 1 - p_a$$

and the Euler characteristic of a surface S is

$$\chi(\mathcal{O}_S) = \mathbf{h}^0(S, \mathcal{O}_S) - \mathbf{h}^1(S, \mathcal{O}_S) + \mathbf{h}^2(S, \mathcal{O}_S) = 1 + p_a$$

But because of tradition, the arithmetic genus is still used quite often.

7.7 Three Theorems about Cohomology

We will use the concept of the *support* of an $\mathcal{O}$-module $\mathcal{M}$, the zero set of its annihilator.

7.7.1. Theorem. *Let X be a projective variety, and let $\mathcal{M}$ be a finite $\mathcal{O}_X$-module.*

(i) *If the support of $\mathcal{M}$ has dimension k, then $H^q(X,\mathcal{M}) = 0$ for all $q > k$. In particular, if the dimension of X is n, then $H^q(X,\mathcal{M}) = 0$ for all $q > n$.*

(ii) *Let $\mathcal{M}(d)$ be the twist of the finite $\mathcal{O}_X$-module $\mathcal{M}$. For sufficiently large d, $H^q(X,\mathcal{M}(d)) = 0$ for all $q > 0$.*

(iii) *For every q, the cohomology $H^q(X,\mathcal{M})$ is a finite-dimensional vector space.*

7.7.2. Notes. (a) The structure of the proofs is interesting. The first part allows us to use *descending* induction to prove the second and third parts, beginning with the fact that $\mathbf{h}^k(\mathcal{M}) = 0$ when k is larger $\dim X$. The descending induction step is to prove that if a statement S_k is true when $k = r+1$, then it is true when $k = r$.

The third part of the theorem tells us that, when $\mathcal{M}$ is a finite $\mathcal{O}$-module, the space $H^0(X,\mathcal{M})$ of global sections is finite-dimensional. This is one of the most important consequences of the theorem, and it isn't easy to prove directly.

(b) Let X be a projective variety. The highest dimension in which cohomology of an $\mathcal{O}_X$-module can be nonzero is called the *cohomological dimension* of X. Theorem 7.7.1 shows that its cohomological dimension is at most its algebraic dimension. In fact, it is equal to the algebraic dimension. On the other hand, X has dimension $2n$ in the classical topology, and the constant coefficient cohomology $H^{2n}_{class}(X,\mathbb{Z})$ in the classical topology will be nonzero. In the classical topology, the cohomological dimension of a projective variety X is its topological dimension $2n$. In the Zariski topology, $H^q(X,\mathbb{Z})$ is zero for every $q > 0$. □

In the theorem, we are given that X is a closed subvariety of a projective space $\mathbb{P}^n$. We can replace an $\mathcal{O}_X$-module by its extension by zero (7.5.1). This doesn't change the cohomology. So we may assume that X is a projective space.

The proofs are based on the cohomology of the twisting modules (7.5.4), the vanishing of the limit $\varinjlim H^q(X,\mathcal{M}(d))$ for $q > 0$ (7.5.2), and on two exact sequences. As we know, $\mathcal{M}(r)$ is generated by global sections if r is sufficiently large (6.6.19). Choosing generators gives us a surjective map $\mathcal{O}^m \to \mathcal{M}(r)$. Let $\mathcal{N}$ be the kernel of this map. When we twist the sequence $0 \to \mathcal{N} \to \mathcal{O}^m \to \mathcal{M}(r) \to 0$, we obtain short exact sequences

(7.7.3) $$0 \to \mathcal{N}(d) \to \mathcal{O}(d)^m \to \mathcal{M}(d{+}r) \to 0$$

for every $d \geq 0$. These sequences are useful because we know that $H^q(X,\mathcal{O}(d)) = 0$ when $q > 0$.

Next, Lemma 6.6.20 tells us that, with coordinates in general position, there will be an exact sequence $0 \to \mathcal{M}(-1) \xrightarrow{x_0} \mathcal{M} \to \overline{\mathcal{M}} \to 0$, where $\overline{\mathcal{M}}$ is the quotient $\mathcal{M}/x_0\mathcal{M}(-1)$. Twisting this sequence gives us exact sequences

(7.7.4) $$0 \to \mathcal{M}(d{-}1) \xrightarrow{x_0} \mathcal{M}(d) \to \overline{\mathcal{M}}(d) \to 0$$

Since the zero locus of x_0 is the hyperplane H at infinity, the support S of $\overline{\mathcal{M}}$ will be contained in $S \cap H$. If S has dimension k and x_0 is generic, the support of $\overline{\mathcal{M}}$ will have dimension less than k. This will allow us to use induction on k and d.

proof of Theorem 7.7.1 **(i)** (*vanishing in large dimension*)

We inspect the sequence (7.7.4). Let k be the dimension of the support of $\mathcal{M}$. If $k = 0$, then $\overline{\mathcal{M}} = 0$ and $H^q(X,\overline{\mathcal{M}}(d)) = 0$ for all q. With coordinates in general position, the support of $\overline{\mathcal{M}}$ will have dimension at most $k{-}1$, if $k > 0$. So by induction on k, we may assume that $H^q(X,\overline{\mathcal{M}}(d)) = 0$ for all $q > k{-}1$ and all d.

The cohomology sequence associated to the sequence (7.7.4) is

(7.7.5) $$\cdots \to H^{q-1}(X,\overline{\mathcal{M}}(d)) \xrightarrow{\delta^{q-1}} H^q(X,\mathcal{M}(d{-}1)) \xrightarrow{x_0} H^q(X,\mathcal{M}(d)) \to H^q(X,\overline{\mathcal{M}}(d)) \xrightarrow{\delta^q} \cdots$$

When $q > k$, the terms on the left and right of this display are zero, and therefore the map

$$H^q(X,\mathcal{M}(d-1)) \xrightarrow{x_0} H^q(X,\mathcal{M}(d))$$

is an isomorphism. According to (7.5.2), $\varinjlim H^q(X,\mathcal{M}(d)) = 0$. It follows that $H^q(X,\mathcal{M}(d)) = 0$ for all d, and in particular, $H^q(X,\mathcal{M}) = 0$ when $q > k$.

proof of Theorem 7.7.1 **(ii)** (*vanishing for a large twist*)

We must show this:

(*) Let $\mathcal{M}$ be a finite $\mathcal{O}$-module. For every $q > 0$ and for sufficiently large d, $H^q(X,\mathcal{M}(d)) = 0$.

By part **(i)**, we know that (*) is true for every $q > n = \dim X$, because all cohomology in dimension q is zero when $q > n$. This leaves a finite set of integers $q = 1, ..., n$ to consider, and it suffices to consider them one at a time. If (*) is true for each individual q there will be a single d such that it is true for $q = 1, ..., n$, and therefore for all positive integers q, as the theorem asserts.

We use descending induction on q, the base case being $q = n+1$, for which (*) is true. We suppose that (*) is true for every finite $\mathcal{O}$-module $\mathcal{M}$ when $q = p+1$, and that $p > 0$, and we show that (*) is true for every finite $\mathcal{O}$-module $\mathcal{M}$ when $q = p$.

We substitute $q = p$ into the cohomology sequence associated to the sequence (7.7.3). The relevant part of that sequence is

$$H^p(X,\mathcal{O}(d))^m \to H^p(X,\mathcal{M}(d+r)) \xrightarrow{\delta^p} H^{p+1}(X,\mathcal{N}(d))$$

Since p is positive, $H^p(X,\mathcal{O}(d)) = 0$ for all $d \geq 0$, and therefore the map δ^p is injective. Our induction hypothesis, applied to the $\mathcal{O}$-module $\mathcal{N}$, shows that $H^{p+1}(X,\mathcal{N}(d)) = 0$ for large d, and then

$$H^p(X,\mathcal{M}(d+r)) = 0$$

The particular integer $d+r$ isn't useful. Our conclusion is that, for every finite $\mathcal{O}$-module $\mathcal{M}$, $H^p(X,\mathcal{M}(k)) = 0$ when k is large enough. □

proof of Theorem 7.7.1 **(iii)** (*finiteness of cohomology*)

This proof also uses descending induction on q. As was mentioned above, it isn't easy to prove directly that the space $H^0(X,\mathcal{M})$ of global sections is finite-dimensional.

We go back to the sequence (7.7.4) and its cohomology sequence (7.7.5). Induction on the dimension of the support of $\mathcal{M}$ allows us to assume that $H^r(X,\overline{\mathcal{M}}(d))$ is finite-dimensional for all r. So, in the part of the cohomology sequence that is depicted in (7.7.5), the terms on the left and right are finite-dimensional. Therefore $H^q(X,\mathcal{M}(d\text{–}1))$ and $H^q(X,\mathcal{M}(d))$ are either both finite-dimensional, or else they are both infinite-dimensional, and this is true for every d.

Suppose that $q > 0$. Then $H^q(X,\mathcal{M}(d)) = 0$ when d is large enough. Since the zero space is finite-dimensional, we can use the sequence together with descending induction, to conclude that $H^q(X,\mathcal{M}(d))$ is finite-dimensional for every finite module $\mathcal{M}$ and every d. In particular, $H^q(X,\mathcal{M})$ is finite-dimensional.

This leaves the case that $q = 0$. To prove that $H^0(X,\mathcal{M})$ is finite-dimensional, we set $d = -r$ in the sequence (7.7.3):

$$0 \to \mathcal{N}(-r) \to \mathcal{O}(-r)^m \to \mathcal{M} \to 0$$

The corresponding cohomology sequence is

$$0 \to H^0(X,\mathcal{N}(-r)) \to H^0(X,\mathcal{O}(-r))^m \to H^0(X,\mathcal{M}) \xrightarrow{\delta^0} H^1(X,\mathcal{N}(-r)) \to \cdots .$$

Here $H^0(X,\mathcal{O}(-r))^m = 0$, and we've shown that $H^1(X,\mathcal{N}(-r))$ is finite-dimensional. It follows that $H^0(X,\mathcal{M})$ is finite-dimensional, and this completes the proof. □

Notice that the finiteness of H^0 comes out only at the end. The higher cohomology is essential for the proof.

(7.7.6) Euler characteristic

Theorem 7.7.1 allows us to define the Euler characteristic of a finite module on projective variety.

7.7.7. Definition. Let X be a projective variety. The *Euler characteristic* of a finite $\mathcal{O}$-module $\mathcal{M}$ is the alternating sum of the dimensions of the cohomology:

$$\chi(\mathcal{M}) = \sum (-1)^q \mathbf{h}^q(X, \mathcal{M}). \tag{7.7.8}$$

This makes sense because $\mathbf{h}^q(X, \mathcal{M})$ finite for every q, and is zero when q is large.

7.7.9. Proposition. (i) *If $0 \to \mathcal{M} \to \mathcal{N} \to \mathcal{P} \to 0$ is a short exact sequence of finite $\mathcal{O}$-modules on a projective variety X, then $\chi(\mathcal{M}) - \chi(\mathcal{N}) + \chi(\mathcal{P}) = 0$.*
(ii) *If $0 \to \mathcal{M}_0 \to \mathcal{M}_1 \to \cdots \to \mathcal{M}_n \to 0$ is an exact sequence of finite $\mathcal{O}$-modules on X, the alternating sum $\sum (-1)^i \chi(\mathcal{M}_i)$ is zero.*

7.7.10. Lemma. *Let $0 \to V^0 \to V^1 \to \cdots \to V^n \to 0$ be an exact sequence of finite dimensional vector spaces. The alternating sum $\sum (-1)^q \dim V^q$ is zero.* □

proof of Proposition 7.7.9. **(i)** Since the cohomology sequence associated to the given sequence is exact, the lemma tells us that the alternating sum of its dimensions is zero. That alternating sum is also equal to $\chi(\mathcal{M}) - \chi(\mathcal{N}) + \chi(\mathcal{P})$.
(ii) Let 's denote the given sequence by $\mathbb{S}_0$ and the alternating sum $\sum_i \chi(\mathcal{M}_i)$ by $\chi(\mathbb{S}_0)$.
Let $\mathcal{N} = \mathcal{M}_1/\mathcal{M}_0$. The sequence $\mathbb{S}_0$ decomposes into the two exact sequences

$$\mathbb{S}_1: \quad 0 \to \mathcal{M}_0 \to \mathcal{M}_1 \to \mathcal{N} \to 0 \quad \text{and} \quad \mathbb{S}_2: \quad 0 \to \mathcal{N} \to \mathcal{M}_2 \to \cdots \to \mathcal{M}_k \to 0 \to$$

Then $\chi(\mathbb{S}_0) = \chi(\mathbb{S}_1) - \chi(\mathbb{S}_2)$, so the assertion follows from **(i)** by induction on n. □

7.8 Bézout's Theorem

As an application of cohomology, we use it to prove Bézout's Theorem.

Recall that, if $f(x) = p_1(x)^{e_1} \cdots p_k(x)^{e_k}$ is a factorization of a homogeneous polyomial in $x = x_0, x_1, x_2$ into irreducible polynomials, the *divisor* of f is defined to be the integer combination $e_1C_1 + \cdots + e_kC_k$, where C_i is the curve of zeros of p_i.

We restate the theorem to be proved.

7.8.1. Bézout's Theorem. *Let Y and Z be the divisors in the projective plane X defined by relatively prime homogeneous polynomials f and g of degrees m and n, respectively. The number of intersection points $Y \cap Z$, counted with an appropriate multiplicity, is equal to mn. Moreover, the multiplicity is 1 at a point at which Y and Z intersect transversally.*

The definition of the multiplicity will emerge during the proof.

7.8.2. Example. Suppose that f and g are products of linear polynomials, so that Y is the union of m lines and Z is the union of n lines, and suppose that those lines are distinct. Since two distinct lines intersect transversally in a single point, there are mn intersection points of multiplicity 1. □

proof of Bézout's Theorem. We will suppress the notation for the extension by zero from a closed subset.

Multiplication by f defines a short exact sequence

$$0 \to \mathcal{O}_X(-m) \xrightarrow{f} \mathcal{O}_X \to \mathcal{O}_Y \to 0$$

where $\mathcal{O}_Y$ stands for $i_*\mathcal{O}_Y$, i being the inclusion $Y \to X$. This sequence describes $\mathcal{O}_X(-m)$ as the ideal $\mathcal{I}$ of Y, and there is a similar sequence describing the module $\mathcal{O}_X(-n)$ as the ideal $\mathcal{J}$ of Z. The zero locus of the ideal $\mathcal{I}+\mathcal{J}$ is the intersection $Y \cap Z$.

We denote the quotient $\mathcal{O}_X/(\mathcal{I}+\mathcal{J})$ by $\overline{\mathcal{O}}$. Since f and g have no common factor, $Y \cap Z$ is a finite set of points $\{p_1, ..., p_k\}$, and $\overline{\mathcal{O}}$ is isomorphic to a direct sum $\bigoplus \overline{\mathcal{O}}_i$, where $\overline{\mathcal{O}}_i$ is a finite-dimensional algebra whose support is p_i (**6.4.20**). We define the *intersection multiplicity* of Y and Z at p_i to be the dimension of $\overline{\mathcal{O}}_i$, which is also equal to $\mathbf{h}^0(X, \overline{\mathcal{O}}_i)$, and we denote the multiplicity by μ_i. The dimension of $H^0(X, \overline{\mathcal{O}})$ is the sum $\mu_1+\cdots+\mu_k$, and $H^q(X,\overline{\mathcal{O}})=0$ for all $q>0$ (Theorem 7.7.1 **(i)**). So the Euler characteristic $\chi(\overline{\mathcal{O}})$ is equal to $\mathbf{h}^0(X,\overline{\mathcal{O}})$. We'll show that $\chi(\overline{\mathcal{O}}) \;=\; mn$, and therefore that $\mu_1+\cdots+\mu_k \;=mn$. This will prove Bézout's Theorem.

We form an exact sequence of $\mathcal{O}$-modules, in which $\mathcal{O}=\mathcal{O}_X$:

(7.8.3)
$$0\to\mathcal{O}(-m-n)\xrightarrow{(g,f)^t}\mathcal{O}(-m)\times\mathcal{O}(-n)\xrightarrow{(f,-g)}\mathcal{O}\xrightarrow{\pi}\overline{\mathcal{O}}\to 0$$

In order to interpret the maps in this sequence as matrix multiplication with homomorphisms acting on the left, sections of $\mathcal{O}(-m)\times\mathcal{O}(-n)$ should be represented as column vectors $(u,v)^t$, u and v being sections of $\mathcal{O}(-m)$ and $\mathcal{O}(-n)$, respectively.

7.8.4. Lemma. *The sequence (7.8.3) is exact.*

proof. To prove exactness, it suffices to show that the sequence of sections on each of the standard affine open sets is exact. We look at $\mathbb{U}^0$, as usual. Let's suppose s coordinates are chosen so that none of the points making up $Y\cap Z$ lie on the coordinate axes. Let A be the algebra of regular functions on $\mathbb{U}^0$, the polynomial algebra $\mathbb{C}[u_1,u_2]$, with $u_i=x_i/x_0$. We identify $\mathcal{O}(k)$ with $\mathcal{O}(kH)$, H being the hyperplane at infinity. The restriction of the module $\mathcal{O}(kH)$ to $\mathbb{U}^0$ is isomorphic to $\mathcal{O}_{\mathbb{U}^0}$. Its sections on $\mathbb{U}^0$ are the elements of A. Let $\overline{A}$ be the algebra of sections of $\overline{\mathcal{O}}$ on $\mathbb{U}^0$. Since f and g are relatively prime, so are their dehomogenizations $F=f(1,u_1,u_2)$ and $G=g(1,u_1,u_2)$. The sequence of sections of (7.8.3) on $\mathbb{U}^0$ is

$$0\to A\xrightarrow{(G,F)^t}A\times A\xrightarrow{(F,-G)}A\to\overline{A}\;\to 0$$

and the only place at which exactness of this sequence isn't obvious is at $A\times A$. Suppose that $(u,v)^t$ is in the kernel of the map $(F,-G)$, i.e., that $Fu=Gv$ Since F and G are relatively prime, F divides v, G divides u, and $v/F=u/G$. Let $w=v/F=u/G$. Then $(u,v)^t=(G,F)^tw$. □

Since cohomology is compatible with products, $\chi(\mathcal{M}\times\mathcal{N})=\chi(\mathcal{M})+\chi(\mathcal{N})$. Proposition 7.7.9**(ii)**, applied to the exact sequence (7.8.3), tells us that the alternating sum

(7.8.5)
$$\chi(\mathcal{O}(-m-n))\;-\;\big(\chi(\mathcal{O}(-m))+\chi(\mathcal{O}(-n))\big)\;+\;\chi(\mathcal{O})-\chi(\overline{\mathcal{O}})$$

is zero. Solving for $\chi(\overline{\mathcal{O}})$ and applying Theorem 7.5.4,

$$\chi(\overline{\mathcal{O}})\;=\;\binom{n+m-1}{2}-\binom{m-1}{2}-\binom{n-1}{2}+1$$

This equation shows that the term $\chi(\overline{\mathcal{O}})$ depends only on the integers m and n. Since we know that the answer is mn when Y and Z are unions of distinct lines, it is mn in every case. This completes the proof.

If you are suspicious of this reasoning, you can evaluate the right side of the equation. □

We still need to explain the assertion that the mutiplicity at a transversal intersection p is equal to 1. This will be true if and only if $\mathcal{I}+\mathcal{J}$ generates the maximal ideal at p locally, and it is obvious when Y and Z are lines. In that case we may choose affine coordinates so that p is the origin in $\mathbb{A}^2\;=\;\operatorname{Spec}A$, $A=\mathbb{C}[y,z]$ and the curves are the coordinate axes $\{z=0\}$ and $\{y=0\}$. The variables u,v generate the maximal ideal at the origin, so the quotient algebra $A/(y,z)$ has dimension 1.

Suppose that Y and Z intersect transverally at p, but that they aren't lines. We choose affine coordinates so that p is the origin and that the tangent directions are the coordinate axes. The affine equations of Y and Z will have the form $y'=0$ and $z'=0$, where $y'=y+g(y,z)$ and $z'=z+h(y,z)$, g and h being polynomials all of whose terms have degree at least 2. Because Y and Z may intersect at points other than p, the elements y' and z' may not generate the maximal ideal at p. However, it suffices to show that they generate the maximal ideal locally.

Let $\widetilde{A}$ be the local ring of the polynomial ring $\mathbb{C}[y,z]$ at the origin, and let $\widetilde{\mathfrak{m}}$ and $\widetilde{k}$ be the maximal ideal and residue field of $\widetilde{A}$, respectively. To show that y',z' generate $\widetilde{\mathfrak{m}}\;=\;(y,z)\widetilde{A}$, the Local Nakayama Lemma 5.1.24 tells us that it suffices to show that their images generate $\widetilde{\mathfrak{m}}/\widetilde{\mathfrak{m}}^2$. The images of g and h in $\widetilde{\mathfrak{m}}^2$ are zero, so y' and z' are congruent to y and z modulo $\widetilde{\mathfrak{m}}^2$. They do generate $\widetilde{\mathfrak{m}}/\widetilde{\mathfrak{m}}^2$, so they generate $\widetilde{\mathfrak{m}}$. □

Chapter 8 THE RIEMANN-ROCH THEOREM FOR CURVES

The topic of this chapter is a classical problem of algebraic geometry, to determine the rational functions on a smooth projective curve with given poles. This can be difficult, and one is usually happy if one can determine the dimension of the space of such functions. The most important tool for this is the Riemann-Roch Theorem.

8.1 Branched Coverings

Smooth affine curves were discussed in Chapter 5. An affine curve is smooth if its local rings are valuation rings, or if its coordinate ring is a normal domain. An arbitrary curve is smooth if it has an open covering by smooth affine curves.

An integral morphism $Y \xrightarrow{\pi} X$ of smooth curves will be called a *branched covering*. It follows from Chevalley's Finiteness Theorem that every nonconstant morphism of smooth projective curves is a branched covering.

If $Y \to X$ is a branched covering, the function field K of Y will be a finite extension of the function field F of X. he *degree* of the covering is the degree $[K:F]$ of the field extension. It will be denoted by $[Y:X]$

If a branched covering $Y \to X$ is given, and if $X' = \operatorname{Spec} A$ is an affine open subset of X, its inverse image Y' will be a smooth affine curve, $Y' = \operatorname{Spec} B$, and if the degree $[Y:X]$ of the covering is n, B will be a locally free A-module of rank $[Y:X]$.

To describe the fibre of a branched covering $Y \xrightarrow{\pi} X$ over a point p of X, we may localize. So we may assume that X and Y are affine, say $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} B$, and that the maximal ideal $\mathfrak{m}_p$ of A at a point p is a principal ideal, generated by an element x of A. If a point q of Y lies over p, the *ramification index* at q is defined to be $e = \mathrm{v}_q(x)$, where v_q is the valuation of the function field K corresponding to q. Then, if y is a local generator for the maximal ideal $\mathfrak{m}_q$ of B at q, we will have

$$x = uy^e$$

where u is a local unit.

The next lemma follows from Lemma 8.2.2 and the Chinese Remainder Theorem.

8.1.1. Lemma. *Let $Y \xrightarrow{\pi} X$ be a branched covering, with $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} B$. Let $q_1, ..., q_k$ be the points of Y that lie over a point p of X, let x be a generator for the the maximal ideal $\mathfrak{m}_p$ at p, and let $\mathfrak{m}_i$ and e_i be the maximal ideal and ramification index at q_i, respectively.*

(i) *The extended ideal $\mathfrak{m}_p B = xB$ is the product ideal $\mathfrak{m}_1^{e_1} \cdots \mathfrak{m}_k^{e_k}$ of B.*
(ii) *Let $\overline{B}_i = B/\mathfrak{m}_i^{e_i}$. The quotient $\overline{B} = B/xB$ is isomorphic to the product $\overline{B}_1 \times \cdots \times \overline{B}_k$.*
(iii) *The degree $[Y:X]$ of the covering is the sum $e_1 + \cdots + e_k$ of the ramification indices at the points q_i in the fibre over p.* □

Points q of Y whose ramification indices are greater than one are called *branch points*. We also call a point p of X a *branch point* of the covering if there is a branch point q that lies over p.

8.1.2. Lemma. *A branched covering $Y \to X$ has finitely many branch points. If a point p is not a branch point, the fibre over p consists of $n = [Y:X]$ points with ramification indices equal to* 1.

proof. We can delete finite sets of points, so we may suppose that X and Y are affine, $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} B$. Then B is a finite A-module of rank n. Let F and K be the fraction fields of A and B, respectively, and let β be an element of B that generates the field extension K/F. Then $A[\beta] \subset B$, and since these two rings have the same fraction field, there will be a nonzero element $s \in A$ such that $A_s[\beta] = B$. We may suppose that $B = A[\beta]$. Let g be the monic irreducible polynomial for β over A. The discriminant of g is nonzero (1.10.1), so for all but finitely many points p of X, there will be n points of Y over p with ramification indices equal to 1. □

8.1.3. Corollary. *A branched covering $Y \xrightarrow{\pi} X$ of degree one is an isomorphism.*

proof. When $[Y:X] = 1$, the function fields of Y and X are equal. Then, because $Y \to X$ is an integral morphism and X is normal, $Y = X$. □

figure: a branched covering

(8.1.4) local analytic structure

The local analytic structure of a branched covering $Y \xrightarrow{\pi} X$ in the classical topology is very simple. We e explain it there because it is useful and helpful for intuition.

Let q be a point of Y, let p be its image in X, and let x be a local generator for the maximal ideal of X at p. Also, let $e = \mathrm{v}_q(x)$ be the ramification index at q.

8.1.5. Proposition. *In the classical topology, Y is locally isomorphic to the e-th root covering $y^e = x$.*

proof. Let z be a local generator for the maximal ideal $\mathfrak{m}_q$ of $\mathcal{O}_Y$. If the ramification index is e, then x has the form uz^e, where u is a local unit at q. In a neighborhood of q in the classical topology, u will have an analytic e-th root w. Then $y = wz$ also generates $\mathfrak{m}_q$ locally, and $x = y^e$. It follows from the implicit function theorem that x and y are local analytic coordinate functions on X and Y (see (**??**)). □

8.1.6. Corollary. *Let $Y \xrightarrow{\pi} X$ be a branched covering, let $\{q_1, ..., q_k\}$ be the fibre over a point p of X, and let e_i be the ramification index at q_i. As a point p' of X approaches p, e_i points of the fibre over p' approach q_i.* □

8.2 Modules on a Smooth Curve

A *torsion element* of a module M over a domain A is an element that is annihilated by some nonzero element a of A: $am = 0$. The set of torsion elements of M is its *torsion submodule*, and a module whose torsion submodule is zero is *torsion-free*. These definitions are extended to $\mathcal{O}$-modules by applying them to affine open sets.

8.2.1. Lemma. *Let Y be a smooth curve.*
(i) *A finite $\mathcal{O}$-module $\mathcal{M}$ is locally free if and only if it is torsion-free.*
(ii) *If an $\mathcal{O}$-module $\mathcal{M}$ swe isn't torsion-free, it has a nonzero global section.*

proof. **(i)** We may assume that Y is affine, $Y = \operatorname{Spec} B$, and that $\mathcal{M}$ is the $\mathcal{O}$-module associated to a B-module M. Let $\widetilde{B}$ and $\widetilde{M}$ be the localizations of B and M at a point q, respectively. Then $\widetilde{M}$ is a finite, torsion-free module over the local ring $\widetilde{B}$. It suffices to show that, for every point q of Y, $\widetilde{M}$ is a free $\widetilde{B}$-module (**5.1.17**). The local ring $\widetilde{B}$ is a valuation ring. A valuation ring is a principal ideal domain because the nonzero ideals of $\widetilde{B}$ are powers of the maximal ideal $\widetilde{\mathfrak{m}}$, which is a principal ideal. Every finite, torsion-free module over a principal ideal domain is free.

(ii) If the torsion submodule of $\mathcal{M}$ isn't zero, there will be an affine open set U, and there will be nonzero elements m in $\mathcal{M}(U)$ and a in $\mathcal{O}(U)$, such that $am = 0$. Let C be the finite set of zeros of a in U, and let $V = Y - C$ be the complement of C in Y. Then a is invertible on the intersection $W = U \cap V$, and since $am = 0$, the restriction of m to W is zero.

The open sets U and V cover Y, and the sheaf property for this covering can be written as an exact sequence

$$0 \to \mathcal{M}(Y) \to \mathcal{M}(U)\times\mathcal{M}(V) \xrightarrow{+,-} \mathcal{M}(W)$$

(Lemma 6.3.10). In this sequence, the section $(m, 0)$ of $\mathcal{M}(U)\times\mathcal{M}(V)$ is mapped to zero in $\mathcal{M}(W)$. Therefore it is the image of a nonzero global section of $\mathcal{M}$. □

8.2.2. Lemma. *Let Y be a smooth curve. Every nonzero ideal $\mathcal{I}$ of $\mathcal{O}_Y$ is a product of powers of maximal ideals of $\mathcal{O}_Y$: $\mathcal{I} = \mathfrak{m}_1^{e_1} \cdots \mathfrak{m}_k^{e_k}$.*

proof. This follows for any smooth curve from the case that Y is affine, which is Proposition 5.3.7. □

8.2.3. Notation. When considering a branched covering $Y \xrightarrow{\pi} X$ of smooth curves, we will often pass between an $\mathcal{O}_Y$-module $\mathcal{M}$ and its direct image $\pi_*\mathcal{M}$, and it will be convenient to work primarily on X. Recall that if Y' is the inverse image of an open subset X' of X, then

$$[\pi_*\mathcal{M}](X') = \mathcal{M}(Y')$$

One can think of the direct image $\pi_*\mathcal{M}$ as working with $\mathcal{M}$, but looking only at open subsets Y' of Y that are inverse images of open subsets X' of X. If we look only at such open subsets, the only significant difference between $\mathcal{M}$ and its direct image will be that the $\mathcal{O}_Y(Y')$-module $\mathcal{M}(Y')$ is made into an $\mathcal{O}_X(X')$-module by restriction of scalars. To simplify notation, we will often drop the symbol π_*, and write $\mathcal{M}$ instead of $\pi_*\mathcal{M}$. If X' is an open subset of X, $\mathcal{M}(X')$ will stand for $\mathcal{M}(\pi^{-1}X')$. When thinking of an $\mathcal{O}_Y$-module $\mathcal{M}$ as the direct image, we may refer to it as an $\mathcal{O}_X$-module. In accordance with this convention, we may also write $\mathcal{O}_Y$ for $\pi_*\mathcal{O}_Y$, but we must be careful to include the subscript Y.

If you find this abbreviation confusing, you can put the symbol π_* into the text where appropriate. □

8.2.4. Lemma. *Let $Y \to X$ be a branched covering of smooth curves.*
(i) *A finite $\mathcal{O}_Y$-module $\mathcal{N}$ is a torsion $\mathcal{O}_Y$-module if and only if it is a torsion $\mathcal{O}_X$-module.*
(ii) *A finite $\mathcal{O}_Y$-module $\mathcal{N}$ is a locally free $\mathcal{O}_Y$-module if and only if it is a locally free $\mathcal{O}_X$-module. If $\mathcal{N}$ is a locally free $\mathcal{O}_Y$-module of rank r, then it is a locally free $\mathcal{O}_X$-module of rank nr, where n is the degree $[Y:X]$ of the covering.* □

(8.2.5) the module Hom

Let M and N be modules over a ring A. We are going to need the A-module that is usually denoted by $\text{Hom}_A(M, N)$, of homomorphisms $M \to N$. The set of such homomorphisms becomes an A-module with some fairly obvious laws of composition: If φ and ψ are homomorphisms and a is an element of A, then $\varphi{+}\psi$ and $a\varphi$ are defined by

(8.2.6) $$[\varphi{+}\psi](m) = \varphi(m) + \psi(m) \quad \text{and} \quad [a\varphi](m) = a\varphi(m)$$

If φ is a module homomorphism, we also have $\varphi(m_1{+}m_2) = \varphi(m_1) + \varphi(m_2)$, and $a\varphi(m) = \varphi(am)$.

8.2.7. Lemma. *An A-module N is canonically isomorphic to $\text{Hom}_A(A, N)$. The homomorphism $A \xrightarrow{\varphi} N$ that corresponds to an element n of N is multiplication by n: $\varphi(a) = an$. Conversely, the element of N that corresponds to a homomorphism $A \xrightarrow{\varphi} N$ is $n = \varphi(1)$.*

Similarly, an $\mathcal{O}$-module $\mathcal{M}$ on a smooth curve Y is isomorphic to $\underline{\text{Hom}}_{\mathcal{O}}(\mathcal{O}, \mathcal{M})$. □

Thus $\mathrm{Hom}_A(A^k, N)$ is isomorphic to N^k, and $\mathrm{Hom}_A(A^\ell, A^k)$ is isomorphic to the module $A^{\ell \times k}$ of $k \times \ell$ A-matrices.

8.2.8. Lemma. *Let A be a noetherian ring.*
(i) *For every finite A-module M, there is an exact sequence $A^\ell \to A^k \to M \to 0$.*
(ii) *If M and N are finite A-modules, then ${}_A(M, N)$ is a finite A-module.* □

An exact sequence of the form $A^\ell \to A^k \to M \to 0$ is called a *presentation* of module M.

The module Hom is compatible with localization:

8.2.9. Lemma. *Let M and N be modules over a noetherian domain A, and suppose that M is a finite module. Let S be a multiplicative system in A. The localization $S^{-1}\mathrm{Hom}_A(M, N)$ is canonically isomorphic to $\mathrm{Hom}_{S^{-1}A}(S^{-1}M, S^{-1}N)$.*

proof. We choose a presentation $A^\ell \to A^k \to M \to 0$. Its localization $(S^{-1}A)^\ell \to (S^{-1}A)^k \to S^{-1}M \to 0$ is a presentation of $S^{-1}M$. Because $\mathrm{Hom}_A(\,\cdot\,,\,\cdot\,)$ is a left exact, contravariant functor of the first variable, the sequence

$$0 \to \mathrm{Hom}_A(M, N) \to \mathrm{Hom}_A(A^k, N) \to \mathrm{Hom}_A(A^\ell, N)$$

is exact, as is its localization. This it suffices to prove the lemma in the case that $M = A$. It is true in that case. □

This lemma shows that when $\mathcal{M}$ and $\mathcal{N}$ are finite $\mathcal{O}$-modules on a variety X, there is an $\mathcal{O}$-module of homomorphisms $\mathcal{M} \to \mathcal{N}$, which will be denoted by $\underline{\mathrm{Hom}}_{\mathcal{O}}(\mathcal{M}, \mathcal{N})$. If $U = \operatorname{Spec} A$ is an affine open set, $M = \mathcal{M}(U)$ and $N = \mathcal{N}(U)$, the module of sections of $\underline{\mathrm{Hom}}_{\mathcal{O}}(\mathcal{M}, \mathcal{N})$ on U is $\mathrm{Hom}_A(M, N)$. We use a new symbol $\underline{\mathrm{Hom}}$ here because the vector space of homomorphisms $\mathcal{M} \to \mathcal{N}$ defined on all of X, which is the space of global sections of $\underline{\mathrm{Hom}}_{\mathcal{O}}(\mathcal{M}, \mathcal{N})$, is customarily denoted by $\mathrm{Hom}_{\mathcal{O}}(\mathcal{M}, \mathcal{N})$.

8.2.10. Notation. The notation $\mathrm{Hom}_A(M, N)$ is cumbersome. It seems permissible to drop the symbol Hom, and to write ${}_A(M, N)$ for $\mathrm{Hom}_A(M, N)$. Similarly, if $\mathcal{M}$ and $\mathcal{M}$ are $\mathcal{O}$-modules on a variety X, we will write ${}_{\mathcal{O}}(\mathcal{M}, \mathcal{N})$ or ${}_X(\mathcal{M}, \mathcal{N})$ for $\underline{\mathrm{Hom}}_{\mathcal{O}}(\mathcal{M}, \mathcal{N})$.

8.2.11. Lemma. *Let $A \subset B$ be rings, let M be an A-module, and let N be a B-module. Then ${}_A(M, N)$ becomes a B-module.*

When we write ${}_A(M, N)$, we are interpreting the B-module N as an A-module by restriction of scalars.

proof. Let $M \xrightarrow{\varphi} N$ be a homomorphism of A-modules, and let b be an element of B. Then multiplication by b is defined by the rule $[b\varphi](m) = \varphi(bm)$. There are several things to check. We list here as a reminder:

The map $[b\varphi]$ is a homomorphism of A-modules $M \to N$:

$$[b\varphi](m_1 + m_2) = [b\varphi](m_1) + [b\varphi](m_2) \text{ and } [b\varphi](am) = a[b\varphi](m)$$

The A-module ${}_A(M, N)$ has the structure of a B-module:

$$[b(\varphi_1 + \varphi_2]) = [b\varphi_1] + [b\varphi_2]\ ,\ [(b_1 + b_2)\varphi] = [b_1\varphi] + [b_2\varphi]\ ,\ [1\varphi] = \varphi\ ,\ \text{and } [b_1[b_2\varphi]] = [b_1 b_2\varphi]$$ □

8.2.12. Lemma.
(i) *The functor Hom_A is a left exact and **contravariant** in the first variable. An exact sequence $M_1 \to M_2 \to M_3 \to 0$ of A-modules induces, for any A-module N, an exact sequence*

$$0 \to {}_A(M_3, N) \to {}_A(M_2, N) \to {}_A(M_1, N)$$

(ii) *The funcctor Hom_A is a left exact and **covariant** in the second variable. An exact sequence $0 \to N_1 \to N_1 \to N_3$ of A-modules induces, for any A-module M, an exact squence*

$$0 \to {}_A(M, N_1) \to {}_A(M, N_2) \to {}_A(M, N_3)$$

The analogous statements are true for $\underline{\mathrm{Hom}}_{\mathcal{O}}$. □

(8.2.13) the dual module

The *dual module* $\mathcal{M}^*$ of a locally free $\mathcal{O}$-module $\mathcal{M}$ is the $\mathcal{O}$-module ${}_{\mathcal{O}}(\mathcal{M}, \mathcal{O})$. A section of $\mathcal{M}^*$ on an open set U is a homomorphism $\mathcal{M}(U) \to \mathcal{O}(U)$. The dual is contravariant. A homomorphism $\mathcal{M} \to \mathcal{N}$ of locally free $\mathcal{O}$-modules induces a homomorphism $\mathcal{M}^* \leftarrow \mathcal{N}^*$.

If $\mathcal{M}$ is a free module with basis $v_1, ..., v_k$, then $\mathcal{M}^*$ will also be free, with the dual basis v_i^* defined by $v_i^*(v_i) = 1$ and $v_i^*(v_j) = 0$ if $i \neq j$. Therefore, when $\mathcal{M}$ is locally free, $\mathcal{M}^*$ is also locally free. The dual $\mathcal{O}^*$ of the structure sheaf $\mathcal{O}$ is $\mathcal{O}$ itself. If $\mathcal{M}$ and $\mathcal{N}$ are locally free $\mathcal{O}$-modules, the dual $(\mathcal{M} \otimes_{\mathcal{O}} \mathcal{N})^*$ is isomorphic to the tensor product $\mathcal{M}^* \otimes_{\mathcal{O}} \mathcal{N}^*$.

There is a canonical $\mathcal{O}$-bilinear map $\mathcal{M}^* \times \mathcal{M} \to \mathcal{O}$. If α and m are sections of $\mathcal{M}^*$ and $\mathcal{M}$, respectively, the bilinear map evaluates α at m: $\langle \alpha, m \rangle = \alpha(m)$.

8.2.14. Corollary. *A locally free $\mathcal{O}$-module $\mathcal{M}$ is canonically isomorphic to its bidual: $(\mathcal{M}^*)^* \approx \mathcal{M}$.*

□

8.2.15. Proposition. *Let $0 \to \mathcal{M} \to \mathcal{N} \to \mathcal{P} \to 0$ be an exact sequence of $\mathcal{O}$-modules on a variety X.*
(i) *If $\mathcal{P}$ is a free $\mathcal{O}$-module, and if the map of global sections $H^0(\mathcal{N}) \to H^0(\mathcal{P})$ is surjective, the sequence splits: $\mathcal{N}$ is isomorphic to the direct sum $\mathcal{M} \oplus \mathcal{P}$.*
(ii) *If $\mathcal{P}$ is locally free, the dual modules form an exact sequence $0 \to \mathcal{P}^* \to \mathcal{N}^* \to \mathcal{M}^* \to 0$.*

proof. **(i)** Let $\{p_i\}$ be a basis of global sections of $\mathcal{P}$, let p_i' be global sections of $\mathcal{N}$ that map to p_i, and let $\mathcal{P}'$ be the free $\mathcal{O}$-submodule of $\mathcal{N}$ spanned by $\{p_i'\}$. So $\mathcal{P}'$ is isomorphic to $\mathcal{P}$, and $\mathcal{N} \approx \mathcal{M} \oplus \mathcal{P}'$.

(ii) The sequence $0 \to \mathcal{P}^* \to \mathcal{M}^* \to \mathcal{N}^*$ is exact whether or not the modules are locally free (8.2.12**(ii)**). The zero on the right comes from the fact that, when $\mathcal{P}$ is locally free, it is free on some affine covering. Thus the given sequence splits locally. □

invertmod **(8.2.16) invertible modules**

An *invertible* $\mathcal{O}$-module is a locally free module of rank one – a module that is isomorphic to the free module $\mathcal{O}$ in a neighborhood of any point.

The tensor product $\mathcal{L} \otimes_{\mathcal{O}} \mathcal{M}$ of invertible modules is invertible. The dual $\mathcal{L}^*$ of an invertible module $\mathcal{L}$ is invertible. Part **(i)** of the next lemma explains the adjective 'invertible'.

8.2.17. Lemma. *Let $\mathcal{L}$ be an invertible $\mathcal{O}$-module.*
(i) *Let $\mathcal{L}^*$ be the dual module. The canonical map $\mathcal{L}^* \otimes_{\mathcal{O}} \mathcal{L} \to \mathcal{O}$ defined by $\gamma \otimes \alpha \mapsto \gamma(\alpha)$ is an isomorphism.*
(ii) *The map $\mathcal{O} \to {}_{\mathcal{O}}(\mathcal{L}, \mathcal{L})$ that sends a regular function α to multiplication by α is an isomorphism.*
(iii) *Every nonzero homomorphism $\mathcal{L} \xrightarrow{\varphi} \mathcal{M}$ to a locally free module $\mathcal{M}$ is injective.*

proof. **(i),(ii)** It is enough to verify these assertions in the case that $\mathcal{L}$ is free, isomorphic to $\mathcal{O}$, in which case they are clear.

(iii) The problem is local, so we may assume that the variety is affine, say $Y = \operatorname{Spec} A$, and that $\mathcal{L}$ and $\mathcal{M}$ are free. Then φ becomes a nonzero homomorphism $A^1 \to A^k$. Such a homomorphism is injective because A is a domain. □

8.3 Divisors

A *divisor* on a smooth curve Y is a finite integer combination of points:

$$D = r_1 q_1 + \cdots + r_k q_k$$

with $r_i \in \mathbb{Z}$. The terms $r_i q_i$ whose integer coefficients r_i are zero can be omitted or not, as desired.

The *support* of D is the set of points q_i of Y such that $r_i \neq 0$. The *degree* of D is the sum $r_1 + \cdots + r_k$ of the coefficients.

Let Y' be an open subset of Y. The *restriction* of a divisor $D = r_1 p_1 + \cdots + r_k p_k$ to Y' is the divisor on Y' obtained from D by deleting points that aren't in Y'. Thus, if $D = q$, the restriction of D to Y' is q when $q \in Y'$, and is zero whan $q \notin Y'$.

A divisor $D = \sum r_i q_i$ is *effective* if all of its coefficients r_i are non-negative, and D is *effective on an open subset* Y' if its restriction to Y' is effective – if $r_i \ge 0$ for every i such that q_i is a point of Y'.

(8.3.1) the divisor of a function divfn

Let f be a rational function on a smooth curve Y. The *divisor of* f is

$$\operatorname{div}(f) = \sum_{q\in Y} \mathrm{v}_q(f)\, q$$

where v_q denotes the valuation of K that corresponds to the point q of Y.
This divisor is written here as a sum over all points q, but it becomes a finite sum when we disregard terms with coefficient zero, because f has finitely many zeros and poles. The coefficients will be zero at all other points.

The map

$$K^\times \to (divisors)^+$$

that sends a rational function to its divisor is a homomorphism from the multiplicative group $K^\times$ of nonzero elements of K to the additive group of divisors:

$$\operatorname{div}(fg) = \operatorname{div}(f) + \operatorname{div}(g)$$

As before, a rational function f has a *zero* of order r at q if $\mathrm{v}_q(f) = r$ with $r > 0$, and it has a *pole* of order r at q if $\mathrm{v}_q(f) = -r$. Thus the divisor of f is the difference of two effective divisors:

$$\operatorname{div}(f) = \; zeros(f) - \; poles(f)$$

A rational function f is regular on Y if and only if its divisor is effective – if and only if $poles(f) = 0$.

The divisor of a rational function is called a *principal divisor*, and two divisors D and E are *linearly equivalent* if their difference $D - E$ is a principal divisor. For instance, the divisors $zeros(f)$ and $poles(f)$ of a rational function f are linearly equivalent.

8.3.2. Lemma. *Let f be a rational function on a smooth curve Y. For all complex numbers c, the divisors of zeros of $f - c$, the level sets of f, are linearly equivalent.*

proof. The functions $f - c$ have the same poles as f. □

(8.3.3) the module $\mathcal{O}(D)$

To analyze the space of functions with given poles on a smooth curve Y, we associate an $\mathcal{O}$-module $\mathcal{O}(D)$ to a divisor D. The nonzero sections of $\mathcal{O}(D)$ on an open subset V of Y are the rational functions f such that the the divisor $\operatorname{div}(f) + D$ is effective on V – such that its restriction to V is effective.

(8.3.4) $$[\mathcal{O}(D)](V) = \{f \mid \operatorname{div}(f) + D \text{ is effective on } V\} \cup \{0\}$$

Points that aren't in the open set V impose no conditions on the sections on V.
When D is an effective divisor, a rational function f is a global section of $\mathcal{O}(D)$ if $poles(f) \le D$.

Say that $D = \sum r_i q_i$. If q_i is a point of an open set V and if $r_i > 0$, a section of $\mathcal{O}(D)$ on V may have a pole of order at most r_i at q_i, and if $r_i < 0$ a section must have a zero of order at least $-r_i$ at q_i. For example, the module $\mathcal{O}(-q)$ is the maximal ideal $\mathfrak{m}_q$. The sections of $\mathcal{O}(-q)$ on an open set V that contains q are the regular functions on V that are zero at q. Similarly, the sections of $\mathcal{O}(q)$ on an open set V that contains q are the rational functions that have a pole of order at most 1 at q and are regular at every other point of V. The sections of $\mathcal{O}(-q)$ and of $\mathcal{O}(q)$ on an open set V that doesn't contain p are the regular functions on V. For any D, sections of $\mathcal{O}(D)$ on V can have arbitrary zeros or poles at points that aren't in V.

The fact that a section of $\mathcal{O}(D)$ is allowed to have a pole at q_i if $r_i > 0$ contrasts with the divisor of a function. If $\operatorname{div}(f) = \sum r_i q_i$, then $r_i > 0$ means that f has a zero at q_i. If $\operatorname{div}(f) = D$, then f will be a global section of $\mathcal{O}(-D)$.

8.3.5. Lemma. **(i)** *If D and E are divisors and if $E-D$ is effective, then $\mathcal{O}(D) \subset \mathcal{O}(E)$.*
(ii) *The function field module $\mathcal{K}$ of a smooth curve Y is the union of the modules $\mathcal{O}(D)$.* □

The next lemma follows from Lemma 8.2.2.

8.3.6. Lemma. *Let Y be a smooth curve. The product ideal $\mathcal{I} = \mathfrak{m}_1^{r_1} \cdots \mathfrak{m}_k^{r_k}$ of $\mathcal{O}_Y$ is isomorphic to the $\mathcal{O}$-module $\mathcal{O}_Y(-D)$, where D is the effective divisor $\sum r_i p_i$. Thus nonzero ideals of $\mathcal{O}_Y$ correspond bijectively to divisors $-D$, where D is effective.* □

8.3.7. Proposition. *Let D and E be divisors on a smooth curve Y.*
(i) *The $\mathcal{O}$-module $\mathcal{O}(D)$ is invertible.*
(ii) *The map $\mathcal{O}(D)\otimes_{\mathcal{O}}\mathcal{O}(E) \to \mathcal{O}(D+E)$ that sends $f\otimes g$ to the product fg is an isomorphism.*
(iii) *The dual module $\mathcal{O}(D)^*$ is $\mathcal{O}(-D)$.*
(iv) *Every invertible $\mathcal{O}$-module $\mathcal{L}$ is isomorphic to a module of the form $\mathcal{O}(D)$.*

The only difference between an invertible module $\mathcal{L}$ and a module of the form $\mathcal{O}(D)$ is that $\mathcal{O}(D)\otimes_{\mathcal{O}}\mathcal{K}$ is equal to $\mathcal{K}$, whereas $\mathcal{L}_{\mathcal{K}}$ can be a one-dimensional K-vector space without chosen basis.

It is important to note that, though every invertible module $\mathcal{M}$ is isomorphic to one of the form $\mathcal{O}(D)$, the divisor D isn't uniquely determined by $\mathcal{M}$. (See (8.3.11) below.)

8.3.8. Definition. Let $\mathcal{L}$ be an invertible $\mathcal{O}$-module on a smooth projective curve Y. If $\mathcal{L}$ is isomorphic to $\mathcal{O}(D)$, we call the degree of D the *degree* of $\mathcal{L}$. With this definition, $\chi(\mathcal{L}) = \deg\, \mathcal{L} + 1 - p_a$.

proof of Proposition 8.3.7. **(i)** We may assume that Y is affine and that the support of D contains at most one point: $D = rp$. We may also assume that the maximal ideal at p is a principal ideal, generated by an element x. In that case, $\mathcal{O}(D)$ will be the free module with basis x^r.

(ii),(iii) Proceeding as in the proof of **(i)**, we may assume that $D = rp$ and $E = sp$. Then $\mathcal{O}(D)$, $\mathcal{O}(-D)$, $\mathcal{O}(E)$, and $\mathcal{O}(D+E)$ have bases x^r, x^{-r}, x^s and x^{r+s}, respectively.

(iv) Let K be the function field of Y, and let $\mathcal{K}$ be the function field module. When $\mathcal{L}$ is an invertible $\mathcal{O}$-module, $\mathcal{L}_K = \mathcal{L}\otimes_{\mathcal{O}}\mathcal{K}$ will be a one-dimensional K-vector space (see (**??**)). Since the function field module $\mathcal{K}$ of Y is the union $\mathcal{K} = \bigcup \mathcal{O}(D)$, we also have $\mathcal{L}_K = \bigcup \mathcal{L}(D)$, where $\mathcal{L}(D)$ denotes the tensor product $\mathcal{L}\otimes_{\mathcal{O}}\mathcal{O}(D)$. A nonzero global section α of $\mathcal{L}_K$ will be a global section of $\mathcal{L}(D)$ for some D. It will define a map $\mathcal{O} \xrightarrow{\alpha} \mathcal{L}(D)$. Passing to duals, $\mathcal{L}(D)^* = \mathcal{L}^*\otimes_{\mathcal{O}}\mathcal{O}(D)^* \approx \mathcal{L}^*(-D)$. The dual of the map α is a nonzero and therefore injective map $\mathcal{L}^*(-D) \to \mathcal{O}$ whose image is an ideal of $\mathcal{O}$. So $\mathcal{L}^*(-D)$ is isomorphic to $\mathcal{O}(-E)$ for some effective divisor E, and therefore $\mathcal{L}^*$ is isomorphic to $\mathcal{O}(D-E)$. Dualizing once more, $\mathcal{L}$ is isomorphic to $\mathcal{O}(E-D)$. □

If CL is an invertible module, we denote by $\mathcal{L}(D)$ the invertible module $\mathcal{L}\otimes_{\mathcal{O}}\mathcal{O}(D)$.

8.3.9. Proposition. *Let $\mathcal{L} \subset \mathcal{M}$ be an inclusion of invertible $\mathcal{O}$-modules. Then $\mathcal{M} = \mathcal{L}(E)$ for some effective divisor E.*

proof. $\mathcal{L}$ is isomorphic to $\mathcal{O}(D)$ for some D. Then $\mathcal{L}(-D) \subset \mathcal{O}$ so $\mathcal{L}(-D)$ is an ideal, isomorphic to $\mathcal{O}(-E)$ for some effective divisor E. Then $\mathcal{L}(E) \approx \mathcal{O}(D) \approx \mathcal{M}$. □

If D and E are divisors, $\mathcal{O}(D)$ is a submodule of $\mathcal{O}(E)$ only when $E-D$ is effective. But as the next proposition explains. there may be homomorphisms from $\mathcal{O}(D)$ to $\mathcal{O}(E)$ that aren't inclusions.

8.3.10. Proposition. *Let D and E be divisors on a smooth curve Y. Multiplication by a rational function f such that $\operatorname{div}(f)+E-D \geq 0$ defines a homomorphism of $\mathcal{O}$-modules $\mathcal{O}(D) \to \mathcal{O}(E)$, and every homomorphism $\mathcal{O}(D) \to \mathcal{O}(E)$ is multiplication by such a function.*

proof. For any $\mathcal{O}$-module $\mathcal{M}$, a homomorphism $\mathcal{O} \to \mathcal{M}$ is multiplication by a global section of $\mathcal{M}$ (6.4.4). Then a homomorphism $\mathcal{O} \to \mathcal{O}(E–D)$ will be multiplication by a rational function f such that $\operatorname{div}(f)+E–D \geq 0$. If f is such a function, one obtains a homomorphism $\mathcal{O}(D) \longrightarrow \mathcal{O}(E)$ by tensoring with $\mathcal{O}(D)$. □

8.3.11. Corollary.
(i) *The modules $\mathcal{O}(D)$ and $\mathcal{O}(E)$ are isomorphic if and only if the divisors D and E are linearly equivalent.*
(ii) *Let f be a rational function on Y, and let $D = \operatorname{div}(f)$. Multiplication by f defines an isomorphism $\mathcal{O}(D) \to \mathcal{O}$.* □

8.4 The Riemann-Roch Theorem I

Let Y be a smooth projective curve. In Chapter 7, we learned that when $\mathcal{M}$ is a finite $\mathcal{O}_Y$-module, the cohomology $H^q(Y,\mathcal{M})$ is a finite-dimensional vector space for all q, and is zero if $q \neq 0,1$. As before, we denote the dimension of the space $H^q(Y,\mathcal{M})$ by $\mathbf{h}^q(\mathcal{M})$ or, if there is ambiguity about the variety, by $\mathbf{h}^q(Y,\mathcal{M})$.

The *Euler characteristic* (7.6.5) of a finite $\mathcal{O}$-module $\mathcal{M}$ is

(8.4.1) $$\chi(\mathcal{M}) = \mathbf{h}^0(\mathcal{M}) - \mathbf{h}^1(\mathcal{M})$$

In particular,

$$\chi(\mathcal{O}_Y) \;=\; \mathbf{h}^0(\mathcal{O}_Y) - \mathbf{h}^1(\mathcal{O}_Y)$$

The dimension $\mathbf{h}^1(\mathcal{O}_Y)$ is the *arithmetic genus* of Y. It denoted by p_a. We will see below, in (8.4.8)**(iv)**, that $\mathbf{h}^0(\mathcal{O}_Y) = 1$. So

(8.4.2) $$\chi(\mathcal{O}_Y) \;=\; 1 - p_a$$

8.4.3. Riemann-Roch Theorem (version 1). *Let $D = \sum r_i p_i$ be a divisor on a smooth projective curve Y. Then*

$$\chi(\mathcal{O}(D)) \;=\; \chi(\mathcal{O}) + \deg\, D \qquad \big(= \deg\, D + 1 - p_a\big)$$

proof. We analyze the effect on cohomology when a divisor is changed by adding or subtracting a point by inspecting the inclusion $\mathcal{O}(D-p) \subset \mathcal{O}(D)$. Let ϵ be the cokernel of the inclusion map, so that there is a short exact sequence in which ϵ is a one-dimensional vector space supported at p, with $\mathbf{h}^0(\epsilon) = 1$, and $\mathbf{h}^1(\epsilon) = 0$.

(8.4.4) $$0 \to \mathcal{O}(D-p) \to \mathcal{O}(D) \to \epsilon \to 0$$

Since $\mathfrak{m}_p$ is isomorphic to $\mathcal{O}(-p)$, this sequence can be obtained by tensoring the sequence

(8.4.5) $$0 \to \mathfrak{m}_p \to \mathcal{O} \to \kappa_p \to 0$$

with the invertible module $\mathcal{O}(D)$.

Let's denote the one-dimensional vector space $H^0(Y,\epsilon)$ by [1]. The cohomology sequence associated to (8.4.4) is

(8.4.6) $$0 \to H^0(Y,\mathcal{O}(D-p)) \to H^0(Y,\mathcal{O}(D)) \xrightarrow{\gamma} [1] \xrightarrow{\delta} H^1(Y,\mathcal{O}(D-p)) \to H^1(Y,\mathcal{O}(D)) \to 0$$

In this exact sequence, one of the two maps, γ or δ, must be zero. Either

(1) γ is zero and δ is injective. In this case

$$\mathbf{h}^0(\mathcal{O}(D-p)) = \mathbf{h}^0(\mathcal{O}(D)) \quad \text{and} \quad \mathbf{h}^1(\mathcal{O}(D-p)) = \mathbf{h}^1(\mathcal{O}(D)) + 1, \quad \text{or}$$

(2) δ is zero and γ is surjective. In this case

$$\mathbf{h}^0(\mathcal{O}(D-p)) = \mathbf{h}^0(\mathcal{O}(D)) - 1 \quad \text{and} \quad \mathbf{h}^1(\mathcal{O}(D-p)) = \mathbf{h}^1(\mathcal{O}(D))$$

In either case,

(8.4.7) $$\chi(\mathcal{O}(D)) = \chi(\mathcal{O}(D-p)) + 1$$

The Riemann-Roch theorem follows from this, because we can get from $\mathcal{O}$ to $\mathcal{O}(D)$ by a finite number of operations, each of which changes the divisor by adding or subtracting a point. □

Because $\mathbf{h}^0 \geq \mathbf{h}^0 - \mathbf{h}^1 = \chi$, this version of the Riemann-Roch Theorem gives reasonably good control of H^0. It is less useful for controlling H^1. To do that, one wants the full Riemann-Roch Theorem. That theorem requires some preparation, so we have put it into Section 8.8. However, version 1 has important consequences:

8.4.8. Corollary. *Let Y be a smooth projective curve.*
(i) *The divisor of a rational function has degree zero: The number of zeros is equal to the number of poles.*
(ii) *Linearly equivalent divisors have equal degrees.*
(iii) *A nonconstant rational function takes every value, including infinity, the same number of times.*
(iv) *A rational function that is regular at every point is a constant: $H^0(Y, \mathcal{O}) = \mathbb{C}$.*
(v) *Let D be a divisor. If $\deg D \geq p_a$, then $\mathbf{h}^0(\mathcal{O}(D)) > 0$.*
(vi) *If $\mathbf{h}^0(\mathcal{O}(D)) > 0$, then $\deg D \geq 0$.*

proof. **(i)** Let $D = \text{div}(f)$. Multiplication by the rational function f defines an isomorphism $\mathcal{O}(D) \to \mathcal{O}$, so $\chi(\mathcal{O}(D)) = \chi(\mathcal{O})$. On the other hand, by Riemann-Roch, $\chi(\mathcal{O}(D)) = \chi(\mathcal{O}) + \deg D$. Therefore $\deg D = 0$.

(ii) If two divisors D and E are linearly equivalent, say $D - E = \text{div}(f)$, then $D - E$ has degree zero, and $\deg D = \deg E$.

(iii) The zeros of the functions $f - c$ are linearly equivalent to the poles of f (8.3.2).

(iv) According to **(iii)**, a nonconstant function must have a pole.

(v) $\mathbf{h}^0 \geq \mathbf{h}^0 - \mathbf{h}^1 = \chi = \deg D + 1 - p_a$.

(vi) Suppose that $\mathcal{O}(D)$ has a nonzero global section f, a rational function such that $\text{div}(f) + D = E$ is effective. Then $\deg E \geq 0$. Since the degree of $\text{div}(f)$ is zero, $\deg D \geq 0$. □

8.4.9. Theorem. *With its classical topology, a smooth projective curve Y is a connected, compact, orientable two-dimensional manifold.*

proof. All points except connectedness have been discussed before (Theorem 1.11.1). A nonempty topological space is connected if it isn't the union of two disjoint, nonempty, closed subsets. We argue by contradiction. Suppose that, in the classical topology, Y is the union of disjoint, nonempty closed subsets Y_1 and Y_2. Both Y_1 and Y_2 will be compact manifolds. Let q be a point of of Y_1. Part **(v)** of Corollary 8.4.8 shows that $\mathbf{h}^0(\mathcal{O}(nq)) > 1$ when n is large. A nonconstant global section f of $\mathcal{O}(nq)$ will be a regular function on the complement $Y - q$ of q. Then f is analytic, and it has no pole on the compact manifold Y_2. It will map Y_2 to a compact subset of the complex plane. A nonconstant analytic function maps open sets to open sets. So if f weren't constant on Y_2, its image would be open. A compact subset of $\mathbb{C}$ can't be open, so f must be constant on Y_2. When we subtract that constant from f, we obtain a nonconstant rational function g that is zero on Y_2. But since Y has dimension 1, the zero locus of a rational function is finite. This is a contradiction. □

8.5 The Birkhoff-Grothendieck Theorem

This theorem describes finite, torsion-free modules on the projective line.

8.5.1. Birkhoff-Grothendieck Theorem. *A finite, torsion-free $\mathcal{O}$-module $\mathcal{M}$ on the projective line $\mathbb{P}^1$ is isomorphic to a direct sum of twisting modules: $\mathcal{M} \approx \bigoplus \mathcal{O}(n_i)$.*

We recall the cohomology of the twisting modules on $\mathbb{P}^1$: If $n \geq 0$, then $\mathbf{h}^0(\mathcal{O}(n)) = n+1$ and $\mathbf{h}^1(\mathcal{O}(n)) = 0$, and if $r > 0$, then $\mathbf{h}^0(\mathcal{O}(-r)) = 0$ and $\mathbf{h}^1(\mathcal{O}(-r)) = r-1$ (Theorem 7.5.4).

8.5.2. Lemma. *Let X denote the projective line, and let $\mathcal{M}$ be a finite, torsion-free $\mathcal{O}$-module on X.*
(i) *The integers r for which there exists a nonzero map $\mathcal{O}(r) \to \mathcal{M}$ are bounded above.*
(ii) *For large r, $\mathbf{h}^0(X, \mathcal{M}(-r)) = 0$.*

proof. **(i)** Since $\mathcal{M}$ is torsion-free, any nonzero map $\mathcal{O} \to \mathcal{M}$, which is multiplication by a global section of $\mathcal{M}$, will be injective. Since $\mathcal{O}(r)$ is locally isomorphic to $\mathcal{O}$, a nonzero map $\mathcal{O}(r) \to \mathcal{M}$ will be injective too, and the associated map $H^0(X, \mathcal{O}(r)) \to H^0(X, \mathcal{M})$ will be injective. Then $\mathbf{h}^0(X, \mathcal{O}(r)) \leq \mathbf{h}^0(X, \mathcal{M})$. Since $\mathbf{h}^0(X, \mathcal{O}(r)) = r+1$ and $\mathbf{h}^0(X, \mathcal{M})$ is finite, r is bounded.

(ii) A global section of $\mathcal{M}(-r)$ defines a map $\mathcal{O} \to \mathcal{M}(-r)$. Its twist by r will be a map $\mathcal{O}(r) \to \mathcal{M}$. □

By the way, the conclusions of the lemma are true for any projective variety X.

proof of the Birkhoff-Grothendieck Theorem. This is Grothendieck's proof. A version of Birkhoff's proof, which uses matrices, is suggested as an exercise.

Lemma 8.2.1 tells us that $\mathcal{M}$ is locally free. We use induction on the rank of $\mathcal{M}$. We suppose that the theorem has been proved for locally free $\mathcal{O}$-modules of rank less than r, that $\mathcal{M}$ has rank r, and that $r > 0$. The plan is to show that $\mathcal{M}$ has a twisting module as a direct summand, so that $\mathcal{M} = \mathcal{W} \oplus \mathcal{O}(n)$ for some $\mathcal{W}$. Then we can apply induction on the rank to $\mathcal{W}$.

Since twisting is compatible with direct sums, we may replace $\mathcal{M}$ by a twist $\mathcal{M}(n)$. Instead of showing that $\mathcal{M}$ has a twisting module $\mathcal{O}(n)$ as a direct summand, we show that, after we replace $\mathcal{M}$ by a suitable twist, the structure sheaf $\mathcal{O}$ will be a direct summand.

As we know (6.6.19), the twist $\mathcal{M}(n)$ will have a nonzero global section when n is sufficiently large, and by Lemma 8.5.2 **(ii)**, it will have no nonzero global section when n is sufficiently negative. Therefore, when we replace $\mathcal{M}$ by a suitable twist, we will have $H^0(X, \mathcal{M}) \neq 0$ but $H^0(X, \mathcal{M}(-1)) = 0$. We assume that this is true for $\mathcal{M}$.

We choose a nonzero global section s of $\mathcal{M}$ and consider the injective multiplication map $\mathcal{O} \xrightarrow{s} \mathcal{M}$. Let $\mathcal{W}$ be the cokernel of this map, so that we have a short exact sequence

(8.5.3) $$0 \to \mathcal{O} \xrightarrow{s} \mathcal{M} \to \mathcal{W} \to 0$$

8.5.4. Lemma. *Let $\mathcal{W}$ be the $\mathcal{O}$-module that appears in the sequence (8.5.3).*
(i) $H^0(X, \mathcal{W}(-1)) = 0$.
(ii) *$\mathcal{W}$ is torsion-free, and therefore locally free.*
(iii) *$\mathcal{W}$ is a direct sum $\bigoplus_{i=1}^{r-1} \mathcal{O}(n_i)$ of twisting modules on $\mathbb{P}^1$, with $n_i \leq 0$.*

proof. **(i)** This follows from the cohomology sequence associated to the twisted sequence

$$0 \to \mathcal{O}(-1) \to \mathcal{M}(-1) \to \mathcal{W}(-1) \to 0$$

because $H^0(X, \mathcal{M}(-1)) = 0$ and $H^1(X, \mathcal{O}(-1)) = 0$.

(ii) If $\mathcal{W}$ had a nonzero torsion submodule, so would $\mathcal{W}(-1)$, and then $\mathcal{W}(-1)$ would have a nonzero global section (8.2.1).

(iii) The fact that $\mathcal{W}$ is a direct sum of twisting modules follows by induction on the rank: $\mathcal{W} \approx \bigoplus \mathcal{O}(n_i)$. Since $H^0(X, \mathcal{W}(-1)) = 0$, we must have $H^0(X, \mathcal{O}(n_i - 1)) = 0$ too. Therefore $n_i - 1 < 0$, and $n_i \leq 0$. □

We go back to the proof of Theorem 8.5.1. Lemma 8.2.15 tells us that the dual of the sequence (8.5.3) is an exact sequence

$$0 \to \mathcal{W}^* \longrightarrow \mathcal{M}^* \longrightarrow \mathcal{O}^* \to 0$$

and $\mathcal{W}^* \approx \bigoplus \mathcal{O}(-n_i)$ with $-n_i \geq 0$. Therefore $\mathbf{h}^1(\mathcal{W}^*) = 0$. The map $H^0(\mathcal{M}) \to H^0(\mathcal{O}^*)$ is surjective. Lemma 8.2.15 tells us that $\mathcal{M}^*$ is isomorphic to $\mathcal{W}^* \oplus \mathcal{O}^*$. Then $\mathcal{M}$ is isomorphic to $\mathcal{W} \oplus \mathcal{O}$. □

8.6 Differentials

Why differentials enter into the Riemann-Roch Theorem is a mystery, but they do, so we introduce them here.

Let A be an algebra and let M be an A-module. A *derivation* $A \xrightarrow{\delta} M$ is a $\mathbb{C}$-linear map that satisfies the product rule for differentiation – a map with these properties:

(8.6.1) $$\delta(ab) = a\,\delta b + b\,\delta a\ , \quad \delta(a+b) = \delta a + \delta b\ , \quad \text{and} \quad \delta c = 0$$

for all a, b in A and all c in $\mathbb{C}$. The fact that δ is $\mathbb{C}$-linear, i.e., that $\delta(cb) = c\,\delta b$, follows.

For example, differentiation $\frac{d}{dt}$ is a derivation $\mathbb{C}[t] \to \mathbb{C}[t]$.

The *module of differentials* Ω_A of an algebra A is an A-module generated by elements denoted by da, one for each element a of A. Its elements are (finite) combinations $\sum b_i\, da_i$, with a_i and b_i in A. The defining relations among the generators da are the ones that make the map $A \xrightarrow{d} \Omega_A$ that sends a to da a derivation: For all a, b in A and all c in $\mathbb{C}$,

$$d(ab) = a\,db + b\,da\ , \quad d(a{+}b) = da + db\ , \quad \text{and} \quad dc = 0 \tag{8.6.2}$$

The elements of Ω_A are called *differentials*.

8.6.3. Lemma.
(i) *Let $\Omega_A \xrightarrow{\varphi} M$ be a homomorphism of $\mathcal{O}$-modules. When we compose φ with the derivation $A \overset{d}{\to} \Omega_A$, we obtain a derivation $A \xrightarrow{\varphi\circ d} M$. Composition with d defines a bijection between homomorphisms $\Omega_A \to M$ and derivations $A \xrightarrow{\delta} M$.*
(ii) *Ω is a functor: An algebra homomorphism $A \xrightarrow{u} B$ induces a homomorphism $\Omega_A \xrightarrow{v} \Omega_B$ that is compatible with the ring homomorphism u, and that makes a diagram*

$$\begin{array}{ccc} B & \xrightarrow{d} & \Omega_B \\ {\scriptstyle u}\uparrow & & \uparrow{\scriptstyle v} \\ A & \xrightarrow{d} & \Omega_A \end{array}$$

proof. **(i)** When we compose the derivation d with a homomorphism φ, we do get a derivation $A \xrightarrow{\delta} M$. In the other direction, given a derivation $A \xrightarrow{\delta} M$, we define a map $\Omega_A \xrightarrow{\varphi} M$ by $\varphi(r\,da) = r\delta(a)$. It follows from the defining relations for Ω_A that φ is a homomorphism of A-modules.

(ii) When Ω_B is made into an A-module by restriction of scalars, the composed map $A \overset{u}{\to} B \overset{d}{\to} \Omega_B$ will be a derivation to which **(i)** applies. □

8.6.4. Lemma. *Let R be the polynomial ring $\mathbb{C}[x_1, ..., x_n]$. The R-module of differentials Ω_R is free, with basis $dx_1, ..., dx_n$.*

proof. The formula $df = \sum \frac{df}{dx_i} dx_i$ follows from the defining relations. It shows that the elements $dx_1, ..., dx_n$ generate the R-module Ω_R.

Let V be a free R-module with basis $v_1, ..., v_n$. The map $R \xrightarrow{\delta} V$ defined by $\delta(f) = \sum \frac{\partial f}{\partial x_i} v_i$ is a derivation. It induces a surjective module homomorphism $\Omega_R \xrightarrow{\varphi} V$ that sends dx_i to v_i. Since $dx_1, ..., dx_n$ generate Ω_R and since $v_1, ..., v_n$ is a basis, φ is an isomorphism. □

8.6.5. Proposition. *Let I be an ideal of an algebra R, let A be the quotient algebra R/I, and let dI denote the set of differentials df with f in I. The subset $N = dI + I\Omega_R$ is a submodule of Ω_R, and Ω_A is isomorphic to the quotient module Ω_R/N.*

The proposition can be interpreted this way: Suppose that the ideal I is generated by elements $f_1, ..., f_r$ of R. Then Ω_A is the quotient of Ω_R obtained from Ω_R by introducing these two rules:

- $df_i = 0$, and
- multiplication by f_i is zero.

For example, let A be the quotient $\mathbb{C}[x]/(x^n)$ of a polynomial ring in one variable and let $\overline{x}$ be the residue of x in A. Then Ω_A is generated by an element $d\overline{x}$, with the relation $n\overline{x}^{n-1}d\overline{x} = 0$.

proof of Proposition 8.6.5. First, $I\Omega_R$ is a submodule of Ω_R, and dI is an additive subgroup of Ω_R. To show that N is a submodule, we must show that scalar multiplication by an element of R carries dI to N, i.e., that if g is in R and f is in I, then $g\,df$ is in N. By the product rule, $g\,df = d(fg) - f\,dg$. Since I is an ideal, fg is in I. Then $d(fg)$ is in dI and $f\,dg$ is in $I\Omega_R$. So $g\,df$ is in N.

The two rules shown above hold in Ω_A because the generators f_i of I are zero in A. Therefore N is in the kernel of the surjective map $\Omega_R \xrightarrow{v} \Omega_A$ defined by the homomorphism $R \to A$. The quotient module

$\overline{\Omega} = \Omega_R/N$, is an A-module, and v defines a surjective map of A-modules $\overline{\Omega} \xrightarrow{\overline{v}} \Omega_A$. We show that $\overline{v}$ is bijective. Let x be an element of R, let a be its image in A, and let $\overline{dx}$ be the image of dx in $\overline{\Omega}$. The composed map $R \xrightarrow{d} \Omega_R \to \overline{\Omega}$ is a derivation that sends x to $\overline{dx}$, and I is in its kernel. It defines a derivation $R/I = A \xrightarrow{\delta} \overline{\Omega}$ that sends a to $\overline{dx}$. This derivation corresponds to a homomorphism of A-modules $\Omega_A \to \overline{\Omega}$ that sends da to $\overline{dx}$, and that inverts $\overline{v}$ (8.6.3). □

8.6.6. Corollary. *If A is a finite-type algebra, then Ω_A is a finite A-module.*

This follows from Proposition 8.6.5 because the module of differentials on the polynomial ring $\mathbb{C}[x_1, ..., x_n]$ is a finite module. □

8.6.7. Lemma. *Let S be a multiplicative system in a domain A, and let $S^{-1}\Omega_A$ be the module of fractions of Ω_A. The modules $S^{-1}\Omega_A$ and $\Omega_{S^{-1}A}$ are canonically isomorphic. In particular, if K is the field of fractions of A, then $K \otimes_A \Omega_A \approx \Omega_K$.*

We have moved the symbol S^{-1} to the left for clarity.

proof of Lemma 8.6.7. The composition $A \to S^{-1}A \xrightarrow{d} \Omega_{S^{-1}A}$ is a derivation that defines an A-module homomorphism $\Omega_A \to \Omega_{S^{-1}A}$. This map extends to an $S^{-1}A$-homomorphism $S^{-1}\Omega_A \xrightarrow{\varphi} \Omega_{S^{-1}A}$ because scalar multiplication by the elements of S is invertible in $\Omega_{S^{-1}A}$. The relation $ds^{-k} = -ks^{k-1}ds$ follows from the definition of a differential, and it shows that φ is surjective. We use the quotient rule

$$\delta(s^{-k}a) = -ks^{-k-1}a\,ds + s^{-k}da$$

to define a derivation $S^{-1}A \xrightarrow{\delta} S^{-1}\Omega_A$. That derivation will correspond to a homomorphism $\Omega_{S^{-1}A} \to S^{-1}\Omega_A$ that inverts φ. However, we must show that δ is well-defined, that $\delta(s_1^{-k}a_1) = \delta(s_2^{-\ell}a_2)$ if $s_1^{-\ell}a_1 = s_2^{-k}a_2$, and that δ is a derivation. You will be able to do this. □

Lemma 8.6.7 shows that a finite $\mathcal{O}$-module Ω_Y of differentials on a variety Y is defined, such that, when $U = \operatorname{Spec} A$ is an affine open subset of Y, $\Omega_Y(U) = \Omega_A$.

8.6.8. Proposition. *The module Ω_Y of differentials on a smooth curve Y is invertible. If y is a local generator for the maximal ideal at a point q, then in a suitable neighborhood of q, Ω_Y will be a free $\mathcal{O}$-module with basis dy.*

proof. We may assume that Y is affine, say $Y = \operatorname{Spec} B$. Let q be a point of Y, and let y be an element of B with $\mathrm{v}_q(y) = 1$. To show that dy generates Ω_B locally, we may localize, so we may suppose that y generates the maximal ideal $\mathfrak{m}$ at q. We must show that after we localize B once more, every differential df with f in B will be a multiple of dy. Let c be the value of the function f at q: Then $f = c + yg$ for some g in B, and because $dc = 0$, $df = g\,dy + y\,dg$. Here $g\,dy$ is in $B\,dy$ and $y\,dy$ is in $\mathfrak{m}\Omega_B$. So

$$\Omega_B = B\,dy + \mathfrak{m}\Omega_B$$

Let M denote the quotient module $\Omega_B/(B\,dy)$. Then $M = \mathfrak{m}M$. The Nakayama Lemma applies. It tells us that there is an element z in $\mathfrak{m}$ such that $s = 1 - z$ annihilates M. When we replace B by its localization B_s, we will have $M = 0$ and $\Omega_B = B\,dy$, as required.

###ugh##

We must still verify that dy isn't a torsion element. If it were, say $b\,dy = 0$, then because dy is a local generator, Ω_B would be the zero module except at the finite set of zeros of b. Since we can take for q an arbitrary point of Y, it suffices to show that the local generator dy for Ω_B isn't zero. Let $R = \mathbb{C}[y]$ and $A = \mathbb{C}[y]/(y^2)$. The module Ω_R is free, with basis dy, and as noted above, if $\overline{y}$ is the residue of y in A, the A-module Ω_A is generated by $d\overline{y}$, with the relation $2\overline{y}\,d\overline{y} = 0$. It isn't the zero module. Proposition 5.3.7 tells us that, at our point q, the algebra $B/\mathfrak{m}_q^2$ is isomorphic to A, and Proposition 8.6.5 tells us that Ω_A is a quotient of Ω_B. Since Ω_A isn't zero, neither is Ω_B. □

8.7 Trace

(8.7.1) trace of a function

Let $Y \xrightarrow{\pi} X$ be a branched covering of smooth curves, and let F and K be the function fields of X and Y, respectively.

The trace map $K \xrightarrow{\text{tr}} F$ for a field extension of finite degree has been defined before (4.5.9). If α is an element of K, multiplication by α on the F-vector space K is an F-linear operator, and $\text{tr}(k)$ is the trace of that operator. The trace is F-linear: If f_i are in F and α_i are in K, then $\text{tr}(\sum f_i\alpha_i) = \sum f_i \,\text{tr}(\alpha_i)$. Moreover, the trace carries regular functions to regular functions: If $X' = \operatorname{Spec} A'$ is an affine open subset of X whose inverse image is $Y' = \operatorname{Spec} B'$, then because A' is a normal algebra, the trace of an element of B' will be in A' (4.5.5). Using our abbreviated notation $\mathcal{O}_Y$ for $\pi_*\mathcal{O}_Y$, the trace defines a homomorphism of $\mathcal{O}_X$-modules

(8.7.2)
$$\mathcal{O}_Y \xrightarrow{\text{tr}} \mathcal{O}_X$$

Analytically, the trace can be described as a sum over the sheets of the covering. Let $n = [Y:X]$. Over a point p of X that isn't a branch point, there will be n points $q_1, ..., q_n$ of Y. If U is a small neighborhood of p in X in the classical topology, its inverse image V will consist of disjoint neighborhoods V_i of q_i, each of which maps bijectively to U. On V_i, the ring $\mathcal{B}$ of analytic functions will be isomorphic to the ring $\mathcal{A}$ of analytic functions on U. So $\mathcal{B}$ is the direct sum $\mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n$ of n copies of $\mathcal{A}$. If a rational function g on Y is regular on V, its restriction to V can be written as $g = g_1 \oplus \cdots \oplus g_n$, with g_i in $\mathcal{A}_i$. The matrix of left multiplication by g on $\mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n$ is the diagonal matrix with entries g_i, so

(8.7.3)
$$\text{tr}(g) = g_1 + \cdots + g_n$$

8.7.4. Lemma. *Let $Y \xrightarrow{\pi} X$ be a branched covering of smooth curves, let p be a point of X, let $q_1, ..., q_k$ be the fibre over p, and let e_i be the ramification index at q_i. If a rational function g on Y is regular at the points $q_1, ..., q_k$, its trace is regular at p, and its value at p is* $[\text{tr}(g)](p) = e_1 g(q_1) + \cdots + e_k g(q_k)$.

proof. The regularity was discussed above. If p isn't a branch point, we will have $k = n$ and $e_i = 1$ for all i. In this case, the lemma follows by evaluating (8.7.3). It follows by continuity for any point p. As a point p' approaches p, e_i points q' of Y approach q_i (8.1.6). For each such point, the limit of $g(q')$ will be $g(q_i)$. □

(8.7.5) trace of a differential

The structure sheaf is naturally contravariant. A branched covering $Y \xrightarrow{\pi} X$ gives us an $\mathcal{O}_X$-module homomorphism $\mathcal{O}_X \to \mathcal{O}_Y$. The trace map for functions is a homomorphism of $\mathcal{O}_X$-modules in the opposite direction: $\mathcal{O}_Y \xrightarrow{\text{tr}} \mathcal{O}_X$.

Differentials are also naturally contravariant. A morphism $Y \to X$ induces an $\mathcal{O}_X$-module homomorphism $\Omega_X \to \Omega_Y$ that sends a differential dx on X to a differential on Y that we denote by dx too (8.6.3) **(ii)**. As is true for functions, there is a trace map for differentials in the opposite direction. It is defined below, in (8.7.7), and will be denoted by $\tau : \Omega_Y \xrightarrow{\tau} \Omega_X$.

First, a lemma about the contravariant map $\Omega_X \to \Omega_Y$:

8.7.6. Lemma. **(i)** *Let p be the image in X of a point q of Y, let x and y be local generators for the maximal ideals of X and Y at p and q, respectively, and let e be the ramification index of the covering at q. Then $dx = vy^{e-1}dy$, where v is a local unit at q.*
(ii) *The canonical homomorphism $\Omega_X \to \Omega_Y$ is injective.*

proof. **(i)** As we have noted before, x has the form uy^e, where u is a local unit. Since dy generates Ω_Y locally, there is a rational function z that is regular at q such that $du = zdy$. Let $v = yz + eu$. Since eu is a local unit and yz is zero at q, v is a local unit, and

$$dx \;=\; d(uy^e) \;=\; y^e z\,dy + ey^{e-1}u\,dy \;=\; vy^{e-1}dy$$

(ii) See (8.2.17). □

To define the trace for differentials, we begin with differentials of the functions fields. Let F and K be the function fields of X and Y, respectively. Because the $\mathcal{O}_Y$-module Ω_Y is invertible, the module Ω_K of

K-differentials, which is the localization $\Omega_Y \otimes_{\mathcal{O}} K$, is a free K-module of rank one. Any nonzero differential will form a K-basis. We choose a nonzero F-differential α. Its image in Ω_K, which we also denote by α, will be a K-basis for Ω_K. We can, for example, take $\alpha = dx$, where x is a local coordinate function on X.

An element β of Ω_K can be written uniquely in the form

$$\beta = g\alpha$$

where g is an element of K. The *trace* $\Omega_K \xrightarrow{\tau} \Omega_F$ is defined by

$$\tau(\beta) = \operatorname{tr}(g)\alpha \tag{8.7.7}$$

where $\operatorname{tr}(g)$ is the trace of the function g. Since the trace for functions is F-linear, τ is also an F-linear map.

We need to check that τ is independent of the choice of α. If α' is another nonzero F-differential, then $f\alpha' = \alpha$ for some nonzero element f of F, and $g\alpha = gf\alpha'$. Since tr is F-linear,

$$\operatorname{tr}(gf)\alpha' = \operatorname{tr}(g)f\alpha' = \operatorname{tr}(g)\alpha$$

Using α' in place of α gives the same value for the trace.

A differenial of the function field K will be called a *rational differential*. A rational differential β is *regular* at a point q of Y if there is an affine open neighborhood $Y' = \operatorname{Spec} B$ of q such that β is an element of Ω_B. If y is a local generator for the maximal ideal $\mathfrak{m}_q$ and $\beta = g\,dy$, then β is regular at q if the rational function g is regular at q.

Let p be a point of X. Working locally at p, we may suppose that X and Y are affine, $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} B$, that the maximal ideal at p is a principal ideal, generated by an element x of A, and that the differential dx generates Ω_A. Let $q_1, ..., q_k$ be the points of Y that lie over p, and let e_i be the ramification index at q_i.

8.7.8. Corollary. (i) *When viewed as a differential on Y, dx has zeros of orders $e_i - 1$ at q_i.*
(ii) *If a differential β on Y is regular at the points $q_i, ..., q_k$, it will have the form $\beta = g\,dx$, where g is a rational function with poles of orders at most $e_i - 1$ at q_i.*

This follows from Lemma 8.7.6 **(i)**. □

8.7.9. Main Lemma. *Let $Y \xrightarrow{\pi} X$ be a branched covering. Let p be a point of X, let $q_1, ..., q_k$ be the points of Y that lie over p, and let β be a rational differential on Y.*
(i) *If β is regular at the points $q_1, ..., q_k$, then its trace $\tau(\beta)$ is regular at p.*
(ii) *If β has a simple pole at q_i and is regular at q_j when $j \neq i$, then $\tau(\beta)$ is not regular at p.*

proof. **(i)** Corollary 8.7.8 tells us that $\beta = g\,dx$, where g has poles of orders at most $e_i - 1$ at the points q_i. Since x has a zero of order e_i at q_i, the function xg is regular at q_i, and its value there is zero. Then $\operatorname{tr}(xg)$ is regular at p, and its value at p is zero (8.7.4). So $x^{-1}\operatorname{tr}(xg)$ is a regular function at p. Since tr is F-linear and x is in F, $x^{-1}\operatorname{tr}(xg) = \operatorname{tr}(g)$. Therefore $\operatorname{tr}(g)$ and $\tau(\beta) = \operatorname{tr}(g)dx$ are regular at p.

(ii) With $\beta = g\,dx$, the function xg will be regular at p. Its value at q_j will be zero when $j \neq i$, and not zero when $j = i$. Then $\operatorname{tr}(xg)$ will be regular at p, but not zero there (8.7.4). Therefore $\tau(\beta) = x^{-1}\operatorname{tr}(xg)dx$ won't be regular at p. □

8.7.10. Corollary. *The trace map defines a homomorphism of $\mathcal{O}_X$-modules $\Omega_Y \xrightarrow{\tau} \Omega_X$.* □

8.7.11. Example. Let Y be the locus $y^e = x$ in $\mathbb{A}^2_{x,y}$. Multiplication by $\zeta = e^{2\pi i/e}$ permutes the sheets of Y over X. The trace of a power y^k is

$$\operatorname{tr}(y^k) = \sum_j \zeta^{kj} y^k \tag{8.7.12}$$

The sum $\sum \zeta^{kj}$ is zero unless $k \equiv 0$ modulo e. Then $\tau(y^r dy) = \tau(y^{r+1-e})dx = 0$ if $r \not\equiv -1$ modulo e, but $\tau(y^{-1}dy) = \operatorname{tr}(e^{-1}x^{-1})dx = x^{-1}dx$ isn't regular at $x = 0$. □

Let $Y \to X$ be a branched covering, and suppose that $Y = \operatorname{Spec} B$ and $X = \operatorname{Spec} A$ are affine. Both B and Ω_B are torsion-free, and therefore locally free A-modules. Let's assume that they are free A-modules, that the maximal ideal of A at p is generated by an element x, and that Ω_A is a free module of rank one with basis dx. Then ${}_A(B, \Omega_A)$ will be a free A-module too.

As is true for any B-module, Ω_B is isomorphic to ${}_B(B, \Omega_B)$. The map $B \to \Omega_B$ that corresponds to an element β of Ω_B is multiplication by β. It sends an element z of B to $z\beta$.

rethink wording##

If β is a B-linear map $B \to \Omega_B$, then because τ is A-linear, the composed map $B \xrightarrow{\beta} \Omega_B \xrightarrow{\tau} \Omega_A$ will be A-linear – a homomorphism of A-modules. Thus composition with the trace τ defines a map

(8.7.13) $$\Omega_B \approx {}_B(B, \Omega_B) \xrightarrow{\tau} {}_A(B, \Omega_A)$$

8.7.14. Theorem. *The map (8.7.13) is an isomorphism of B-modules.*

proof. This theorem follows from the Main Lemma 8.7.9, when one looks closely.

Let's denote ${}_A(B, \Omega_A)$ by $\mathcal{H}$. This is an A-module, but it becomes a B-module because B is a B-module (8.2.11). Scalar multiplication by an element b of B is defined as follows: Let $B \xrightarrow{u} \Omega_A$ be an A-linear map. Then bu is the map $[bu](z) = u(zb)$ for z in B.

Next, because B and Ω_A are locally free A-modules, $\mathcal{H}$ is a locally free A-module and a locally free B-module. Since Ω_A has A-rank 1, the A-rank of $\mathcal{H}$ is the same as the A-rank of B. Therefore the B-rank of $\mathcal{H}$ is 1 (8.2.4)**(ii)**. So $\mathcal{H}$ is an invertible B-module.

The trace map $\Omega_B \xrightarrow{\tau} \mathcal{H}$ isn't the zero map because $\tau\, dx \neq 0$. Since domain and range are invertible B-modules, τ is an injective homomorphism. Its image, which is isomorphic to Ω_B, is an invertible submodule of the B-module $\mathcal{H}$. Therefore $\mathcal{H}$ is isomorphic to the invertible module $\Omega_B(D)$ for some effective divisor D (8.3.7). To complete the proof of the theorem, we show that the divisor D is zero.

Suppose that $D > 0$ and let q be a point in the support of D. We may suppose that q lies over our chosen point p. Then $\Omega_B(q) \subset \Omega_B(D) \approx \mathcal{H}$. We choose a rational differential β in Ω_K that has a simple pole at q, and is regular at the other points of Y in the fibre over p. The Chinese Remainder Theorem allows us to do this. According to Lemma 8.7.9 **(ii)**, the trace $\tau(\beta)$ isn't regular at p. It isn't in $\mathcal{H}$. □

Note. This is a subtle theorem, and I don't like the proof. It is understandable, but it doesn't give much insight as to why the theorem is true. To get more insight, we would need a better understanding of differentials. My father Emil Artin said "One doesn't really understand differentials, but one can learn to work with them."

8.7.15. Theorem. *Let $Y \to X$ be a branched covering of affine varieties $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} B$, and let M be a finite B-module. Composition with the trace $\Omega_B \xrightarrow{\tau} \Omega_A$ defines a bijection*

(8.7.16) $${}_B(M, \Omega_B) \xrightarrow{\tau\circ} {}_A(M, \Omega_A)$$

proof. We choose a resolution

$$B^m \to B^n \to M \to 0$$

of M and form a diagram

$$\begin{array}{ccccccc} 0 & \longrightarrow & {}_B(M, \Omega_B) & \longrightarrow & {}_B(B, \Omega_B)^n & \longrightarrow & {}_B(B, \Omega_B)^m \\ & & \downarrow{\scriptstyle a} & & \downarrow{\scriptstyle b} & & \downarrow{\scriptstyle c} \\ 0 & \longrightarrow & {}_A(M, \Omega_A) & \longrightarrow & {}_A(A, \Omega_A)^n & \longrightarrow & {}_A(A, \Omega_A)^m \end{array}$$

in which the maps a, b, c are the compositions with the trace τ, as was described above. Because the functor Hom is left exact and contravariant in the first variable, the rows of this diagram are exact. Theorem 8.7.14 Shows that b and c are bijective. Therefore a is bijective too. □

Extension of this theorem to branched coverings $Y \to X$ in which Y and X aren't affine presents no problem.

8.7.17. Corollary. *Let $Y \xrightarrow{\pi} X$ be a branched covering of smooth curves, and let $\mathcal{M}$ be a finite $\mathcal{O}_Y$-module. The map ${}_Y(\mathcal{M}, \Omega_Y) \xrightarrow{\tau\circ} {}_X(\mathcal{M}, \Omega_X)$ obtained by composition with the trace $\Omega_Y \xrightarrow{\tau} \Omega_X$ is an isomorphism of $\mathcal{O}_X$-modules.* □

When written without dropping the symbol Hom or suppressing the notation for the direct image, this isomorphism becomes an isomorphism

$$\pi_*\big(\underline{\mathrm{Hom}}_{\mathcal{O}_Y}(\mathcal{M},\Omega_Y)\big) \xrightarrow{\tau\circ} \underline{\mathrm{Hom}}_{\mathcal{O}_X}(\pi_*\mathcal{M},\Omega_X)$$

8.8 The Riemann-Roch Theorem II

(8.8.1) the Serre dual

Let Y be a smooth projective curve, and let $\mathcal{M}$ be a locally free $\mathcal{O}_Y$-module. The *Serre dual* of $\mathcal{M}$, which we will denote by $\mathcal{M}^{\#}$, is the module

(8.8.2)
$$\mathcal{M}^{\#} \;=\; {}_Y(\mathcal{M},\Omega_Y) \;=\; \underline{\mathrm{Hom}}_{\mathcal{O}_Y}(\mathcal{M},\Omega_Y)$$

For example, $\mathcal{O}_Y^{\#} = \Omega_Y$ and $\Omega_Y^{\#} = \mathcal{O}_Y$.

Since the invertible module Ω_Y is locally isomorphic to $\mathcal{O}_Y$, the Serre dual $\mathcal{M}^{\#}$ will be locally isomorphic to the ordinary dual $\mathcal{M}^*$. It will be a locally free module with the same rank as $\mathcal{M}$, and the bidual $(\mathcal{M}^{\#})^{\#}$ will be isomorphic to $\mathcal{M}$.

8.8.3. Riemann-Roch Theorem, version 2. *Let $\mathcal{M}$ be a locally free $\mathcal{O}_Y$-module on a smooth projective curve Y, and let $\mathcal{M}^{\#}$ be its Serre dual. Then $\mathbf{h}^0(\mathcal{M}) = \mathbf{h}^1(\mathcal{M}^{\#})$ and $\mathbf{h}^1(\mathcal{M}) = \mathbf{h}^0(\mathcal{M}^{\#})$.*

Because $\mathcal{M}$ and $(\mathcal{M}^{\#})^{\#}$ are isomorphic, the two assertions of the theorem are equivalent.

For example, $\mathbf{h}^1(\Omega_Y) = \mathbf{h}^0(\mathcal{O}_Y) = 1$ and $\mathbf{h}^0(\Omega_Y) = \mathbf{h}^1(\mathcal{O}_Y) = p_a$.

If $\mathcal{M}$ is a locally free $\mathcal{O}_Y$-module on a smooth projective curve Y, then

(8.8.4)
$$\chi(\mathcal{M}) = \mathbf{h}^0(\mathcal{M}) - \mathbf{h}^0(\mathcal{M}^{\#})$$

A more precise statement of the Riemann-Roch Theorem is that $H^1(Y,\mathcal{M})$ and $H^0(Y,\mathcal{M}^{\#})$ are dual vector spaces in a canonical way. We omit the proof of this. The fact that their dimensions are equal is enough for many applications. The canonical isomorphism becomes important only when one wants to apply the theorem to a cohomology sequence. And of course, any complex vector spaces V and W whose dimensions are equal can be made into dual spaces by the choice of a nondegenerate bilinear form $V\times W \to \mathbb{C}$.

Our plan is to prove Theorem 8.8.3 directly for the projective line. The structure of locally free modules on $\mathbb{P}^1$ is very simple, so this will be easy. Following Grothendieck, swe derive it for an arbitrary smooth projective curve Y by projection to $\mathbb{P}^1$.

Let Y be a smooth projective curve, let $X = \mathbb{P}^1$, and let $Y \xrightarrow{\pi} X$ be a branched covering. Let $\mathcal{M}$ be a locally free $\mathcal{O}_Y$-module, and let the Serre dual of $\mathcal{M}$, as defined in (8.8.2), be $\mathcal{M}_1^{\#}$:

$$\mathcal{M}_1^{\#} \;=\; {}_Y(\mathcal{M},\Omega_Y)$$

The direct image of $\mathcal{M}$ is a locally free $\mathcal{O}_X$-module that we are denoting by $\mathcal{M}$ too, and we can form the Serre dual on X. Let

$$\mathcal{M}_2^{D} \;=\; {}_X(\mathcal{M},\Omega_X)$$

8.8.5. Corollary. *The direct image $\pi_*\mathcal{M}_1^{\#}$, which we also denote by $\mathcal{M}_1^{\#}$, is isomorphic to $\mathcal{M}_2^{\#}$.*

proof. This is Theorem 8.7.17. □

The corollary allows us to drop the subscripts from $\mathcal{M}^{\#}$. Because a branched covering is an affine morphism, the cohomology of $\mathcal{M}$ and of its Serre dual $\mathcal{M}^{\#}$ can be computed, either on Y or on X. (See (7.4.25).)

If $Y \xrightarrow{\pi} X$ is a branched covering of projective curves and $\mathcal{M}$ is a locally free $\mathcal{O}_Y$-module, then $H^q(Y,\mathcal{M}) \approx H^q(X,\mathcal{M})$ and $H^q(Y,\mathcal{M}^{\#}) \approx H^q(X,\mathcal{M}^{\#})$.

Thus it is enough to prove Riemann-Roch for the projective line.

(8.8.6) Riemann-Roch for the projective line

The Riemann-Roch Theorem for the projective line $X = \mathbb{P}^1$ is a simple consequence of the Birkhoff-Grothendieck Theorem, which tells us that every locally free $\mathcal{O}_X$-module $\mathcal{M}$ on X is a direct sum of twisting modules $\mathcal{O}_X(k)$. To prove Riemann-Roch for the projective line $X = \mathbb{P}^1$, it suffices to the theorem for the twisting modules.

8.8.7. Lemma. *The module of differentials Ω_X on X is isomorphic to the twisting module $\mathcal{O}_X(-2)$.*

proof. Since Ω_X is invertible, the Birkhoff-Grothendieck Theorem tells us that it is a twisting module $\mathcal{O}_X(k)$. We only need to identify the integer k. On the standard open subset $\mathbb{U}^0 = \operatorname{Spec}\mathbb{C}[x]$, the module of differentials is free, with basis dx, and $z = x^{-1}$ is the coordinate on $\mathbb{U}^1 = \operatorname{Spec}\mathbb{C}[z]$. Then $dx = d(z^{-1}) = -z^{-2}dz$ describes the differential dx on $\mathbb{U}^1$. Since the point p at infinity is $\{z = 0\}$, dx has a pole of order 2 there. It is a global section of $\Omega_X(2p)$, and as a section of that module, it isn't zero anywhere. So multiplication by dx defines an isomorphism $\mathcal{O} \to \Omega_X(2p)$ that sends 1 to dx. Tensoring with $\mathcal{O}(-2p)$, we find that Ω_X is isomorphic to $\mathcal{O}(-2p)$. □

8.8.8. Lemma. *Let let $\mathcal{M}$ and $\mathcal{N}$ be locally free $\mathcal{O}$-modules on the projective line X. Then $_X(\mathcal{M}(r),\mathcal{N})$ is canonically isomorphic to $_X(\mathcal{M},\mathcal{N}(-r))$.*

proof. When we tensor a homomorphism $\mathcal{M}(r) \xrightarrow{\varphi} \mathcal{N}$ with $\mathcal{O}(-r)$, we obtain a homomorphism $\mathcal{M} \to \mathcal{N}(-r)$, and tensoring with $\mathcal{O}(r)$ is the inverse operation. □

The Serre dual $\mathcal{O}(n)^\#$ of $\mathcal{O}(n)$ is therefore

$$\mathcal{O}(n)^\# = {}_X(\mathcal{O}(n), \Omega_X) \approx \mathcal{O}(-2-n)$$

To prove Riemann-Roch for $X = \mathbb{P}^1$, we must show that

$$\mathbf{h}^0(X, \mathcal{O}(n)) = \mathbf{h}^1(X, \mathcal{O}(-2-n)) \quad \text{and} \quad \mathbf{h}^1(X, \mathcal{O}(n)) = \mathbf{h}^0(X, \mathcal{O}(-2-n))$$

This follows from the computation of cohomology of the twisting modules (Theorem 7.5.4). □

8.9 Using Riemann-Roch

(8.9.1) genus

Three closely related numbers associated to a smooth projective curve Y are: its *topological genus* g, its *arithmetic genus* $p_a = \mathbf{h}^1(\mathcal{O}_Y)$, and the *degree* δ of the module of differentials Ω_Y.

8.9.2. Theorem. *Let Y be a smooth projective curve. The topological genus g and the arithmetic genus p_a of Y are equal, and the degree δ of the module Ω_Y is $2p_a - 2$, which is equal to $2g - 2$.*

proof. Let $Y \xrightarrow{\pi} X$ be a branched covering of $X = \mathbb{P}^1$. The topological Euler characteristic $e(Y)$, which is $2-2g$, can be computed in terms of the branching data for the covering (see (1.11.4)). Let q_i be the ramification points in Y, and let e_i be the ramification index at q_i. Then e_i sheets of the covering come together at q_i. If the degree of Y over X is n, then since $e(X) = 2$,

(8.9.3) $$2 - 2g = e(Y) = ne(X) - \sum(e_i - 1) = 2n - \sum(e_i - 1)$$

We compute the degree δ of Ω_Y in two ways. First, the Riemann-Roch Theorem tells us that $\mathbf{h}^0(\Omega_Y) = \mathbf{h}^1(\mathcal{O}_Y) = p_a$ and $\mathbf{h}^1(\Omega_Y) = \mathbf{h}^0(\mathcal{O}_Y) = 1$. So $\chi(\Omega_Y) = -\chi(\mathcal{O}_Y) = p_a - 1$. The Riemann-Roch Theorem also tells us that $\chi(\Omega_Y) = \delta + 1 - p_a$ (8.3.8). Therefore

(8.9.4) $$\delta = 2p_a - 2$$

Next, we compute δ by computing the divisor of the differential dx on Y, x being a coordinate in X. Let q_i be one of the ramification points in Y, and let e_i be the ramification index at q_i. Then dx has a zero of order e_i-1 at q_i. On X, dx has a pole of order 2 at ∞. Let's suppose that the point at infinity isn't a branch point. Then there will be n points of Y at which dx has a pole of order 2, n being the degree of Y over X. The degree of Ω_Y is therefore

(8.9.5)
$$\delta = \mathit{zeros} - \mathit{poles} = \sum(e_i - 1) - 2n$$

Combining (8.9.5) with (8.9.3), one sees that $\delta = 2g - 2$. Since we also have $\delta = 2p_a - 2,\ g = p_a$. □

8.9.6. Corollary. *Let D be a divisor on a smooth projective curve Y of genus g. If* $\deg D > 2g - 2$ *then* $\mathbf{h}^1(\mathcal{O}(D)) = 0$*. If* $\deg D \le g - 2$*, then* $\mathbf{h}^1(\mathcal{O}(D)) > 0$.

proof. This follows from Corollary 8.4.8 **(v)** and **(vi)**. □

(8.9.7) curves of genus zero

Let Y be a smooth projective curve Y of genus zero, and let p be a point of Y. The exact sequence

$$0 \to \mathcal{O}_Y \to \mathcal{O}_Y(p) \to \epsilon \to 0$$

where ϵ is a one-dimensional module supported at p (8.4.6), gives us an exact cohomology sequence

$$0 \to H^0(Y, \mathcal{O}_Y) \to H^0(Y, \mathcal{O}_Y(p)) \to H^0(Y, \epsilon) \to 0$$

The zero on the right is due to the fact that, because $p_a = 0,\ H^1(Y, \mathcal{O}_Y) = 0$. We also have $\mathbf{h}^0(\mathcal{O}_Y) = \mathbf{h}^0(\epsilon) = 1$, so $\mathbf{h}^0(\mathcal{O}_Y(p)) = 2$. We choose a basis $(1, x)$ for $H^0(Y, \mathcal{O}_Y(p))$, 1 being the constant function and x being a nonconstant function with a single pole of order 1 at p. This basis defines a point of $\mathbb{P}^1$ with values in the function field K of Y, and therefore a morphism $Y \xrightarrow{\varphi} \mathbb{P}^1$. Because x has just one pole of order 1, it takes every value exactly once. Therefore φ is bijective. It is a map of degree 1, and therefore an isomorphism (8.1.3).

8.9.8. Corollary. *Every smooth projective curve of genus zero is isomorphic to the projective line* $\mathbb{P}^1$. □

A *rational curve* is a curve (smooth or not) whose function field is isomorphic to the field $\mathbb{C}(t)$ of rational functions in one variable. A smooth projective curve of genus zero is a rational curve.

(8.9.9) curves of genus one

A smooth projective curve of genus 1 is called an *elliptic curve*. The Riemann-Roch Theorem tells us that on an elliptic curve Y,

$$\chi(\mathcal{O}(D)) = \deg D$$

Since $\mathbf{h}^0(\Omega_Y) = \mathbf{h}^1(\mathcal{O}_Y) = 1$, Ω_Y has a nonzero global section ω. Since Ω_Y has degree zero (8.9.2), ω doesn't vanish anywhere. Multiplication by ω defines an isomorphism $\mathcal{O} \to \Omega_Y$. So Ω_Y is a free module of rank one. It follows that the Serre dual $\mathcal{M}^{\#}$ of an $\mathcal{O}$-module $\mathcal{M}$ is isomorphic to the ordinary dual $\mathcal{M}^*$.

The next lemma follows from Riemann-Roch.

8.9.10. Lemma. *Let p be a point of an elliptic curve Y. For any $r > 0$,* $\mathbf{h}^0(\mathcal{O}(rp) = r$*, and* $\mathbf{h}^1(\mathcal{O}(rp)) = 0$. □

Since $H^0(Y, \mathcal{O}_Y) \subset H^0(Y, \mathcal{O}_Y(p))$, and since both spaces have dimension one, they are equal. So (1) is a basis for $H^0(Y, \mathcal{O}_Y(p))$. We choose a basis $(1, x)$ for the two-dimensional space $H^1(Y, \mathcal{O}_Y(2p))$. Then x isn't a section of $\mathcal{O}(p)$. It has a pole of order precisely 2 at p. Next, we choose a basis $(1, x, y)$ for $H^1(Y, \mathcal{O}_Y(3p))$. So x and y are functions with poles of orders 2 and 3, respectively, at p, and no other poles. The point $(1, x, y)$ of $\mathbb{P}^2$ with values in K determines a morphism $Y \xrightarrow{\varphi} \mathbb{P}^2$. Let u, v, w be coordinates in $\mathbb{P}^2$. The map φ sends a point q distinct from p to $(u, v, w) = (1, x(q), y(q))$. Since Y has dimension one, φ is a finite morphism. Its image will be a closed subvariety of $\mathbb{P}^2$ of dimension one. Since $(1, x, y)$ are independent, the image isn't contained in a line.

To determine the image of the point p, we multiply $(1, x, y)$ by $\lambda = y^{-1}$ to normalize the second coordinate to 1, obtaining the equivalent vector $(y^{-1}, xy^{-1}, 1)$. The rational function y^{-1} has a zero of order 3 at p, and xy^{-1} has a simple zero there. Evaluating at p, we see that the image of p is the point $(0, 0, 1)$.

Let Y' be the image of Y, which is a curve in $\mathbb{P}^2$. The map $Y \to \mathbb{P}^2$ restricts to a finite morphism $Y \to Y'$. Let ℓ be a generic line $\{au+bv+cw = 0\}$ in $\mathbb{P}^2$. The rational function $a+bx+cy$ on Y has a pole of order 3 at p and no other pole. It takes every value, including zero, three times, and the set of points q of Y at which $a + bx + cy$ is zero is the inverse image of the intersection $Y' \cap \ell$. The only possibilities for the degree of Y' are 1 and 3. Since $1, x, y$ are independent, Y' isn't a line. So the image Y' is a cubic curve (Corollary 1.3.9).

To determine the image, we look for a cubic relation among the functions $1, x, y$ on Y. The seven monomials $1, x, y, x^2, xy, x^3, y^2$ have poles at p of orders $0, 2, 3, 4, 5, 6, 6$, respectively, and no other poles. They are sections of $\mathcal{O}_Y(6p)$. Riemann-Roch tells us that $\mathbf{h}^0(\mathcal{O}_Y(6p)) = 6$. So those seven functions are dependent. The linear dependency relation gives us a cubic equation among x and y, which we write in the form

$$cy^2 + (a_1x + a_3)y + (a_0x^3 + a_2x^2 + a_4x + a_6) = 0$$

There can be no linear relation among functions whose orders of pole at p are distinct. So when we delete either x^3 or y^2 from the list of monomials, we obtain an independent set of six functions that form a basis for the six-dimensional space $H^0(Y, \mathcal{O}(6p))$. In the cubic relation, the coefficients c and a_0 aren't zero. We can scale y and x to normalize c and a_0 to 1. We eliminate the linear term in y from this relation by substituting $y - \frac{1}{2}(a_1x + a_3)$ for y. Next, we eliminate the quadratic term in x. by substituting $x - \frac{1}{3}a_2$ for x. Bringing the terms in x to the other side of the equation, we are left with a cubic relation

$$y^2 = x^3 + a_4x + a_6$$

The coefficients a_4 and a_6 have changed, of course.

The cubic curve Y' defined by the homogenized equation $y^2z = x^3 + a_4xz^2 + a_6z^3$ is the image of Y. This curve Y' meets a generic line $ax + by + cz = 0$ in three points and, as we saw above, its inverse image in Y consists of three points too. Therefore the morphism $Y \xrightarrow{\varphi} Y'$ is generically injective, and Y is the normalization of Y'. Corollary 7.6.3 computes the cohomology of Y': $\mathbf{h}^0(\mathcal{O}_{Y'}) = \mathbf{h}^1(\mathcal{O}_{Y'}) = 1$. This tells us that $\mathbf{h}^q(\mathcal{O}_{Y'}) = \mathbf{h}^q(\mathcal{O}_Y)$ for all q. Let's denote the direct image of $\mathcal{O}_Y$ by the same symbol $\mathcal{O}_Y$. The quotient $\mathcal{F} = \mathcal{O}_Y/\mathcal{O}_{Y'}$ is a torsion module with no global sections, so it is zero (8.2.1) **(ii)**.

8.9.11. Corollary. *Every elliptic curve is isomorphic to a cubic curve in* $\mathbb{P}^2$. □

(8.9.12) the group law on an elliptic curve

The points of an elliptic curve form an abelian group, once one chooses a point to be the identity element.

We choose a point of Y, and label it o. We'll write the law of composition in the group as $p \oplus q$, using the symbol $\oplus$ to distinguish this sum, which is a point of Y, from the divisor $p + q$.

Let p and q be points of Y. To define $p \oplus q$, we compute the cohomology of $\mathcal{O}_Y(p+q-o)$. It follows from Riemann-Roch that $\mathbf{h}^0(\mathcal{O}_Y(p + q - o)) = 1$ and that $\mathbf{h}^1(\mathcal{O}_Y(p + q - o)) = 0$. There is a nonzero function f, unique up to scalar factor, with simple poles at p and q and a zero at o. This function has exactly one other zero. That zero is defined to be the sum $p \oplus q$ in the group. In terms of linearly equivalent divisors, $s = p \oplus q$ is the unique point such that s is linearly equivalent to $p + q - o$, or such that $p + q$ is linearly equivalent to $o + s$.

8.9.13. Proposition. *The law of composition* $\oplus$ *defined above makes an ellipic curve into an abelian group.*

The proof is an exercise. □

(8.9.14) interlude: maps to projective space

Let Y be a smooth projective curve. We have seen that any set $(f_0, ..., f_n)$ of rational functions on Y defines a morphism $Y \xrightarrow{\varphi} \mathbb{P}^n$ (5.3.3). As a reminder, let q be a point of Y and let $g_j = f_j/f_i$, where i is an index such that f_i has the minimum value $\mathrm{v}_q(f_i)$. Then g_j are regular at q for all j, and the morphism φ sends the point q to is $(g_0(q), ..., g_n(q))$. For example, the inverse image $Y^0 = \varphi^{-1}(\mathbb{U}^0)$ of the standard open set $\mathbb{U}^0$ is the set of points at which the functions $g_j = f_j/f_0$ are regular. If q is such a point, then $\varphi(q) = (1, g_1(q), ..., g_n(q))$.

8.9.15. Lemma. *Let $Y \xrightarrow{\varphi} \mathbb{P}^n$ be the morphism of a smooth projective curve Y to projective space that is defined by a set $(f_0, ..., f_n)$ of rational functions on Y.*
(i) *If the space spanned by $\{f_0, ..., f_n\}$ has dimension at least two, then φ is not a constant function.*
(ii) *If $\{f_0, ..., f_n\}$ are linearly independent, the image isn't contained in any hyperplane.* □

The *degree* d of a nonconstant morphism φ from a projective curve Y, smooth or not, to projective space $\mathbb{P}^n$, is the number of points of the inverse image $\varphi^{-1}(H)$ of a generic hyperplane H in $\mathbb{P}^n$. We check that this number is well-defined. Say that H is the locus $h(x) = 0$, where $h = \sum a_i x_i$, and that another generic hyperplane G is the locus $g(x) = 0$, where $g =: \sum b_i x_i$. Let $f(x) = h/g$. The divisor of the rational function $\widetilde{f} = f \circ \varphi$ on Y is $\varphi^{-1}H - \varphi^{-1}H'$.

(8.9.16) base points

If D is a divisor on the smooth projective curve Y, a basis $(f_0, ..., f_k)$ of global sections of $\mathcal{O}(D)$ defines a morphism $Y \to \mathbb{P}^{k-1}$. This is the most common way to construct such a morphism, though one could use any set of rational functions.

If a global section of $\mathcal{O}(D)$ vanishes at a point p of Y, it is a section of $\mathcal{O}(D-p)$. A point p is a *base point* of $\mathcal{O}(D)$ if every global section of $\mathcal{O}(D)$ vanishes at p. A base point can be described in terms of the usual exact sequence

$$0 \to \mathcal{O}(D-p) \to \mathcal{O}(D) \to \epsilon \to 0$$

The point p is a base point if $\mathbf{h}^0(\mathcal{O}(D-p)) = \mathbf{h}^0(\mathcal{O}(D))$, or if $\mathbf{h}^1(\mathcal{O}(D-p)) = \mathbf{h}^1(\mathcal{O}(D)) - 1$.

Let $Y \xrightarrow{\pi} \mathbb{P}^n$ is a morphism. The *degree* . of π is the number of points in the inverse image of a generic hyperplane.

8.9.17. Lemma. *Let D be a divisor on a smooth projective curve Y, and suppose that $H^0(\mathcal{O}(D)) \neq 0$. Let $Y \xrightarrow{\varphi} \mathbb{P}^n$ be the morphism defined by a basis of global sections.*
(i) *The image of φ isn't contained in any hyperplane.*
(ii) *If $\mathcal{O}(D)$ has no base points, the degree r of the morphism φ is equal to degree of D. If there are base points, the degree is lower.* □

(8.9.18) canonical divisors

Because the module Ω_Y of differentials on a smooth curve Y is invertible, it is isomorphic to $\mathcal{O}(K)$ for some divisor K. Such a divisor K is called a *canonical divisor*. It is often convenient to represent Ω_Y as a module $\mathcal{O}(K)$, though the canonical divisor K isn't unique. It is determined only up to linear equivalence (see (8.3.11)).

When written in terms of a canonical divisor K, the Serre dual of an invertible module $\mathcal{O}(D)$ will be $\mathcal{O}(D)^\# = {}_{\mathcal{O}}(\mathcal{O}(D), \mathcal{O}(K)) \approx \mathcal{O}(K{-}D)$. With this notation, the Riemann-Roch Theorem for $\mathcal{O}(D)$ becomes

(8.9.19) $$\mathbf{h}^0(\mathcal{O}(D)) = \mathbf{h}^1(\mathcal{O}(K{-}D)) \quad \text{and} \quad \mathbf{h}^1(\mathcal{O}(D)) = \mathbf{h}^0(\mathcal{O}(K{-}D))$$ □

8.9.20. Proposition. *Let K be a canonical divisor on a smooth projective curve Y of genus $g > 0$.*
(i) *$\mathcal{O}(K)$ has no base point.*
(ii) *Every point p of Y is a base point of $\mathcal{O}(K+p)$.*

proof. **(i)** Let p be a point of Y. We apply Riemann-Roch to the exact sequence

$$0 \to \mathcal{O}(K-p) \to \mathcal{O}(K) \to \epsilon_1 \to 0$$

where ϵ_1 denotes a one-dimensional module supported on a point p. The Serre duals $\mathcal{O}$ and $\mathcal{O}(p)$ of $\mathcal{O}(K)$ and $\mathcal{O}(K-p)$, respectively, form an exact sequence

$$0 \to \mathcal{O} \to \mathcal{O}(p) \to \epsilon_2 \to 0$$

When Y has positive genus, there is no rational function on Y with just one simple pole. So $\mathbf{h}^0(\mathcal{O}(p)) = \mathbf{h}^0(\mathcal{O}) = 1$. Riemann-Roch tells us that $\mathbf{h}^1(\mathcal{O}(K-p)) = \mathbf{h}^1(\mathcal{O}(K)) = 1$. The cohomology sequence

$$0 \to H^0(\mathcal{O}(K-p)) \to H^0(\mathcal{O}(K)) \to [1] \to H^1(\mathcal{O}(K-p)) \to H^1(\mathcal{O}(K)) \to 0$$

shows that $\mathbf{h}^0(\mathcal{O}(K-p)) = \mathbf{h}^0(\mathcal{O}(K)) - 1$. So p is not a base point.

(ii) Here, the relevant sequence is

$$0 \to \mathcal{O}(K) \to \mathcal{O}(K+p) \to \epsilon_3 \to 0$$

The Serre dual of $\mathcal{O}(K+p)$ is $\mathcal{O}(-p)$, which has no global section. Therefore $\mathbf{h}^1(\mathcal{O}(K+p)) = 0$, while $\mathbf{h}^1(\mathcal{O}(K)) = \mathbf{h}^0(\mathcal{O}) = 1$. The cohomology sequence

$$0 \to \mathbf{h}^0(\mathcal{O}(K)) \to \mathbf{h}^0(\mathcal{O}(k+p)) \to [1] \to \mathbf{h}^1(\mathcal{O}(K)) \to \mathbf{h}^1(\mathcal{O}(k+p)) \to 0$$

shows that $H^0(\mathcal{O}(K)) = H^0(\mathcal{O}(K+p))$. So p is a base point of $\mathcal{O}(K+p)$. $\square$

(8.9.21) hyperelliptic curves

A *hyperelliptic curve* Y is a smooth projective curve of genus $g > 1$ that can be represented as a branched double covering of the projective line. So Y is hyperelliptic if there is a morphism $Y \xrightarrow{\pi} X$ of degree two, with $X = \mathbb{P}^1$.

The topological Euler characteristic of a hyperelliptic curve Y can be computed in terms of the covering $Y \to X$, which will be branched at a finite set $p_1, ..., p_n$ of n points. Since π has degree two, the multiplicity of a branch point will be $e = 2$. The Euler characteristic is therefore $e(Y) = 2e(X) - n = 4 - n$. Since $e(Y) = 2 - 2g$, the number of branch points is $n = 2g + 2$. So when $g = 3$, $n = 8$.

It would take some experimentation to guess that the next remarkable theorem might be true, and some time to find a proof.

8.9.22. Theorem. *Let K be a canonical divisor on a hyperelliptic curve Y, and let $Y \xrightarrow{\pi} X = \mathbb{P}^1$ be the associated branched covering of degree 2. The morphism $Y \xrightarrow{\kappa} \mathbb{P}^{g-1}$ defined by the global sections of $\Omega_Y = \mathcal{O}(K)$ factors through X. There is a morphism $X \xrightarrow{u} \mathbb{P}^{g-1}$ such that $\pi = \kappa \circ u$: $Y \xrightarrow{\pi} X \xrightarrow{u} \mathbb{P}^{g-1}$.*

8.9.23. Corollary. *A curve of genus $g \geq 2$ can be presented as a branched covering of $\mathbb{P}^1$ of degree 2 in at most one way.* $\square$

proof of Theorem 8.9.22.
Let x be an affine coordinate in X, so that the standard affine open subset $\mathbb{U}^0$ of X is $\operatorname{Spec} \mathbb{C}[x]$. We may suppose that the point p_∞ at infinity isn't a branch point of the covering. Let $Y^0 = \pi^{-1}\mathbb{U}^0$. Then Y^0 will have an equation of the form

$$y^2 = f(x)$$

where f is a polynomial with $n = 2g + 2$ simple roots. There will be two points of Y above the point p_∞. They are interchanged by the automorphism $y \to -y$. Let's call those points q_1 and q_2.

We start with the differential dx, which we view as a differential on Y. Then $2y\,dy = f'(x)dx$. Since f has simple roots, f' doesn't vanish at any of them. Therefore dx has simple zeros on Y above the roots of f. We also have a regular function on Y^0 with simple roots at those points, namely the function y. Therefore the differential $\omega = \frac{dx}{y}$ is regular and nowhere zero on y^0. Because the degree of a differential on Y is $2g-2$ (**??**), ω has a total of $2g-2$ zeros at infinity. By symmetry, ω has zeros of order $g-1$ at the eah of two points q_1 and q_2. So $K = (g-1)q_1 + (g-1)q_2$ is a canonical divisor on Y, i.e., $\Omega_Y \approx \mathcal{O}_Y(K)$.

Since K has zeros of order $g-1$ at infinity, the rational functions $1, x, x^2, ..., x^{g-1}$, viewed as functions on Y, are among the global sections of $\mathcal{O}_Y(K)$. They are independent, and there are g of them. Since $\mathbf{h}^0(\mathcal{O}_Y(K)) = g$, they form a basis of $H^0(\mathcal{O}_Y(K))$. The map $Y \to \mathbb{P}^{g-1}$ defined by the global sections of $\mathcal{O}_Y(K)$ evaluates these powers of x, so it factors through the double covering $Y \xrightarrow{\pi} X$. □

(8.9.24) canonical embedding

Let Y be a smooth projective curve of genus $g \geq 2$, and let K be a canonical divisor on Y. Since $\mathcal{O}(K)$ has no base point (**??**), its global sections define a morphism $Y \xrightarrow{\kappa} \mathbb{P}^{g-1}$, the *canonical map* whose degree is equal to the degree $2g-2$ of the canonical divisor.

8.9.25. *Let Y be a smooth projective curve of genus g at least two. If Y is not hyperelliptic, the canonical map embeds Y as a closed subvariety of projective space $\mathbb{P}^{g-1}$.*

proof. We show first that, if κ isn't an injective map, then Y is hyperelliptic. Let p and q be distinct points such that $\kappa(p) = \kappa(q)$. We may assume that the canonical divisor K is effective, and that p and q are not in its support. We inspect the global sections of $\mathcal{O}(K-p-q)$. Since $\kappa(p) = \kappa(q)$, any global section of $\mathcal{O}(K)$ that vanishes at p vanishes at q too. Therefore $\mathcal{O}(K-p)$ and $\mathcal{O}(K-p-q)$ have the same global sections, and q is a base point of $\mathcal{O}(K-p)$. We've computed the cohomology of $\mathcal{O}(K-p)$: $\mathbf{h}^0(\mathcal{O}(K-p)) = g-1$ and $\mathbf{h}^1(\mathcal{O}(K-p)) = 1$. Then $\mathbf{h}^0(\mathcal{O}(K-p-q)) = g-1$ and $\mathbf{h}^1\mathcal{O}(K-p-q)) = 2$. The Serre dual of $\mathcal{O}(K-p-q)$ is $\mathcal{O}(p+q)$, so by Riemann-Roch, $\mathbf{h}^0(\mathcal{O}(p+q)) = 2$. If D is a divisor of degree one on a curve of positive genus, then $\mathbf{h}^0(\mathcal{O}(D)) \leq 1$ (Proposition **??**). Therefore $\mathcal{O}(p+q)$ has no base point. Its global sections define a morphism $Y \to \mathbb{P}^1$ of degree 2. So Y is hyperelliptic. Conversely, if Y is hyperelliptic, Theorem 8.9.22 shows that κ has degree 2.

If Y isn't hyperelliptic, the canonical map is injective, so Y is mapped bijectively to its image Y' in $\mathbb{P}^{g-1}$. This almost proves the theorem, but: Can Y' have a cusp? We must show that the bijective map $Y \xrightarrow{\kappa} Y'$ is an isomorphism.

We go over the computation made above for a pair of points p, q, this time taking $q = p$. The computation is the same. It shows that, since Y isn't hyperelliptic, p isn't a base point of $\mathcal{O}_Y(K-p)$. Therefore $\mathbf{h}^0(\mathcal{O}_Y(K-2p)) = \mathbf{h}^0(\mathcal{O}_Y(K-p)) - 1$. This tells us that there is a global section f of $\mathcal{O}_Y(K)$ that has a zero of order exactly 1 at p. When properly interpreted, this fact shows that κ doesn't collapse any tangent vectors to Y, and therefore that κ is an isomorphism. Since we haven't discussed tangent vectors, we prove this directly.

Since κ is a bijective, finite morphism, it is an integral morphism. The function fields of Y and its image Y' are equal, and Y is the normalization of Y'. Moreover, κ is an isomorphism except on a finite set.

We work locally at a point p of Y'. When we restrict the global section f of $\mathcal{O}_Y(K)$ found above to the image Y', we obtain an element of the maximal ideal $\mathfrak{m}'$ of $\mathcal{O}_{Y'}$ at p, that we denote by x. On Y, this element has a zero of order one at p, and therefore it is a local generator fot the maximal ideal $\mathfrak{m}_p$ of $\mathcal{O}_Y$. We may assume that x generates $\mathfrak{m}_p$, and that the quotient $\mathcal{F} = \mathcal{O}_Y/\mathcal{O}_{Y'}$ is a finite-dimensional vector space supported at p.

We multiply the short exact sequence $0 \to \mathcal{O}_{Y'} \xrightarrow{i} \mathcal{O}_Y \xrightarrow{\pi} \mathcal{F} \to 0$ by x. The cokernels of the multiplication maps form an exact sequence

$$\mathcal{O}_{Y'}/x\mathcal{O}_{Y'} \xrightarrow{\bar{i}} \mathcal{O}_Y/x\mathcal{O}_Y \xrightarrow{\bar{\pi}} \mathcal{F}/x\mathcal{F} \to 0$$

Since x generates $\mathcal{M}$, $\mathcal{O}_Y/x\mathcal{O}_Y$ is the residue field $\kappa(p)$, which has dimension one. The map $\bar{i}$ isn't zero because it sends the residue of 1 in $\mathcal{O}_{Y'}/x\mathcal{O}_{Y'}$ to the residue of 1 in $\kappa(p)$. Therefore $\bar{i}$ is surjective. This shows that $\mathcal{F}/x\mathcal{F} = 0$. But since $\mathcal{F}$ is a finite $\mathcal{O}_{Y'}$-module and x is in the maximal ideal $\mathfrak{m}'$, the quotient $\mathcal{F}/x\mathcal{F}$ can't be zero unless $\mathcal{F}$ is zero (5.1.24). Therefore $\mathcal{O}_{Y'} = \mathcal{O}_Y$. □

8.9.26. Examples. Let Y be a smooth projective curve of genus g.

(i) When $g = 2$, the canonical morphism κ is a map of degree $2g - 2 = 2$ from Y to $\mathbb{P}^1$. Every smooth projective curve of genus 2 is hyperelliptic.

(ii) When $g = 3$, κ is a morphism of degree 4 from Y to $\mathbb{P}^2$. If Y isn't hyperelliptic, its image will be a plane curve of degree 4, isomorphic to Y. The genus of a smooth projective curve of degree 4 is $\binom{3}{2} = 3$ (1.11.6), which checks.

The number of moduli of curves of degree 3 (the number of essential parameters) is obtained this way: There are 15 monomials of degree 4 in three variables. The group GL_3 of dimension 9 operates by coordinate changes. So the number of moduli is $15 - 9 = 6$. When a hyperelliptic curve of genus 3 is represented as a branched double covering of $\mathbb{P}^1$, there will be 8 branch points. The group GL_2 of dimension 4 operates on the branch points, but scalars don't move them. So the number of moduli of hyperelliptic curves of genus three is $8 - 3 = 5$. Since $5 < 6$, this agrees with the fact that not all curves of genus three are hyperelliptic.

(iii) When $g = 4$, κ is a morphism of degree 6 from Y to $\mathbb{P}^3$, and it becomes harder to count moduli. It is a fact that the number of moduli of curves of any genus g is $3g - 3$. The number of moduli of hyperelliptic curves of genus g is easy to compute. It is $2g - 1$. □

Index