Algebraic Geometry

Jacob Denson

January 20, 2021

Table Of Contents

Part I

Classical Algebraic Geometry

Chapter 1

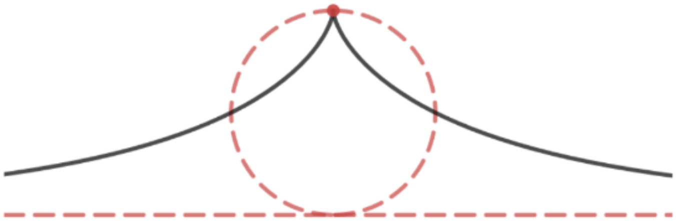
Algebraic Sets

1.1 A Brief History of Algebraic Curves

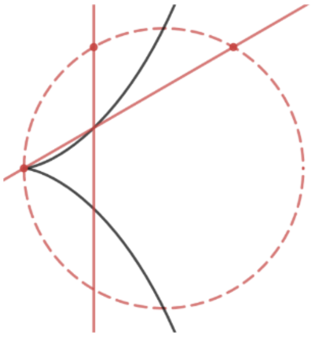
Classically, Euclidean geometry discusses the relations between lines, circles, and their intersections. The practical foundations for this were that the most reliable tools of the time, the compass and unmarked ruler, trivialized the precise constructions of these shapes. However, in the late 19th century, this geometry was proven insufficient for certain constructions. One cannot use a ruler and compass to split an arbitrary angle into three, nor construct a square with the same area as a given circle. The Greeks did not have the apparatus required to show that attempts of these impossible problems were doomed to fail, but they realized their ineptitude, and introduced more sophisticated tools to solve these problems. Menaechmus ‘doubled a cube‘ with the introduction of the parabola. Pappus used the hyperbola to trisect the angle. Thus the conic sections became a canonical part of Euclidean geometry.

Thanks to Descartes’ analytical geometry, we can identify the Euclidean plane with coordinates: tuples of two numbers. Furthermore, lines, circles, and conic sections are then described as the *locus* of points satisfying some algebraic relationship between the coordinates. The circle is the set of points satisfying $X^2 + Y^2 = 1$, the hyperbola $X^2 - Y^2 = 1$, and the parabola $Y = X^2$. An *algebraic plane curve* is a subset in the plane described as the locus of a polynomial equation. They are ubiquitous throughout geometry, and the techniques used to understand them formed the historical foundation for the study of algebraic geometry.

Example. *A* cissoid *is a curve constructed from two curves C_1, C_2, and a pole O. It is obtained by taking lines through O intersecting both C_1 and C_2 at points P_1 and P_2, and marking off the points Q such that OQ has the same length and orientation as P_1P_2. If we take C_1 to be a circle, C_2 a line tangent to the circle, and O the point on the circle opposite the tangent line, we obtain the cissoid of Diocles, introduced to solve the problem of doubling a cube.*



The name cissoid originates from the Greek κισσοειδης, meaning 'ivy shaped', probably because of the singular, or cuspal points common in the curves, like in the cissoid of Diocles. We can also describe the cissoid of Diocles by letting a point P range over all points on the circle, and then obtaining a point Q by first reflecting P across the line parallel to the fixed tangent bisecting the circle to obtain a point P', and then intersecting the line OP with the line through P' parallel to the tangent. The easiest way to see why this is true is to look at the diagram below, and use similarity of triangles.



If we consider a coordinate system in which O lies at the origin, the circle is the unit circle with center $(1,0)$, and the tangent is the linear $X = 2$, then in polar coordinates, we may write the points on the circle as the solutions to the equation $r = 2\sec\theta$, and the points on the tangent as solutions to $r = 2\cos\theta$. The distance between the tangent line and the circle for a fixed angle θ

is the difference in radii, which is just $2\sec\theta - 2\cos\theta$, and therefore the polar equation defining the cissoid of Diocles is

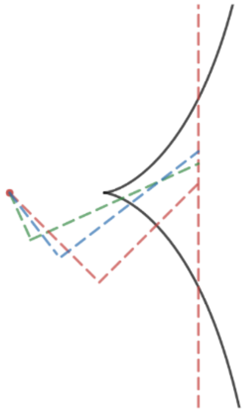
$$r = 2(\sec t - \cos t) = \frac{2 - 2\cos^2 t}{\cos t} = \frac{2\sin^2 t}{\cos t}$$

Since $X = r\cos t$, $Y = r\sin t$, and $r^2 = X^2 + Y^2$, in cartesian coordinates this equation becomes

$$X = \frac{2Y}{X^2 + Y^2}$$

so the cissoid of Diocles is the locus of points defined by the polynomial equation $(X^2 + Y^2)X = 2Y^2$, and is therefore an algebraic planar curve.

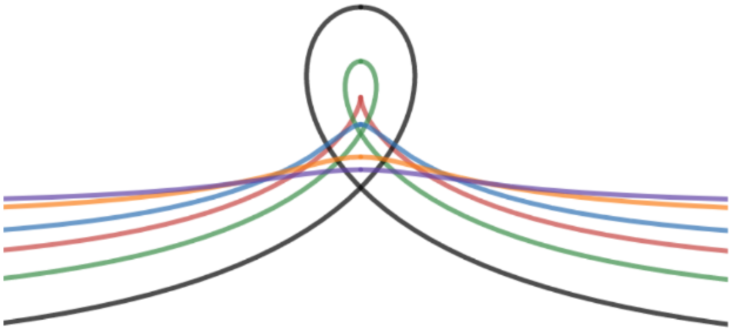
Example. *Another construction of Diocles' cissoid was discovered by Newton. Consider a rigid right angled joint forced to pass through a point O, and another point P lying on a line not passing through O, where the length of the joint from the bend to the point P is fixed, but the length through O is allowed to vary. As we move the point P along the line, the joint slides back and forth through the point O. If we take the midpoints Q between the point P and the bend in the joint, then we obtain the cissoid of Diocles. We have expressed Diocles' cissoid as a* conchoid, *which is constructed from a general point O and curve by taking points on lines through O lying at a fixed distance from the curve. In this case, the curve is a line, and the distance is half the distance between the point and the line.*



Suppose we choose a coordinate system in which $O = (-1, 0)$*, and the line over which* P *varies is* $X = 1$*. If* Q *has coordinates* (a, b)*, and* P *has coordinates* $(1, p)$*, then the fact that* OQP *is a right angle is equivalent to saying that* $\langle O - Q, P - Q \rangle = 0$*, which gives the equation* $a^2 + b^2 = 1 + bp$*. The condition that* PQ *has length 2 is equivalent to the algebraic equation* $a^2 + b^2 + p^2 = 3 + 2a + 2bp$*, which, assuming the first equation is satisfied, is equivalent to* $p^2 = 2 + 2a + bp$*. Introducing the midpoint* (X, Y) *of* PQ*, which is the point we want to exist in the first place, we find that*

$$2X = a + 1 \quad 2Y = b + p$$

The first equation allows us to eliminate a *from the first two equations, and the second allows us to eliminate* b*. We obtain that the values of* X *and* Y *which lie on the shape are exactly those such that there exists a value* p *such that* $(2X - 1)^2 + (2Y - p)^2 = 1 + (2Y - p)p$ *and* $p^2 = 2 + 2(2X - 1) + (2Y - p)p$*, which is simplified to the two equation* $4X^2 + 4Y^2 + 2p^2 = 6pY + 4X$ *and* $p^2 = 2X + pY$*. Substituting the second equation into the first gives* $X^2 + Y^2 = pY$*, and substituting this equation back into the second, after multiplying the equation by* Y^2 *on both sides gives* $(X^2 + Y^2)^2 = (2X + X^2 + Y^2)Y^2$*, which can be simplified to* $(X^2 + Y^2)X = 2Y^2$*, so this constructing describes exactly the cissoid of Diocles.*



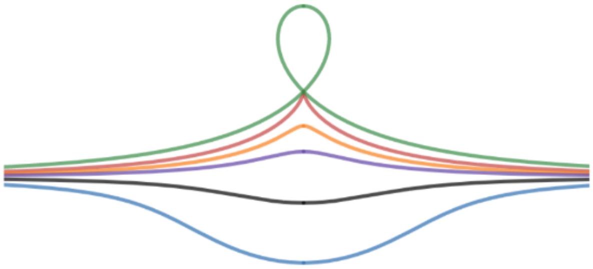
The advantage of Newton's construction is that we can obtain a one-parameter family of conchoidal curves which are deformations of the cissoid of Diocles, by taking points on the joint lying at a different ratio than the midpoint. Indeed, if $X = ac + (1 - c)$ *and* $Y = cb + (1 - c)p$*, then provided that* $c \neq 0$ *we can still use the first equation to eliminate* a*, and use the second to eliminate* b*, obtaining that*

$$X^2 + Y^2 + 2(c - 1)X + p(c - 2)Y + (1 - c)p^2 = 2c - 1 \quad p^2 = 2X + pY + 4c - 2$$

Substituting the second equation into the first gives $pY = X^2+Y^2-4c^2+4c-1$, and substituting the equation back into the second once multiplying by Y on both sides of the equation gives

$$X^4+X^2Y^2-2XY^2-2(2c-1)^2X^2+(1-4c^2)Y^2 = -16c^4+32c^3-24c^2+8c-1$$

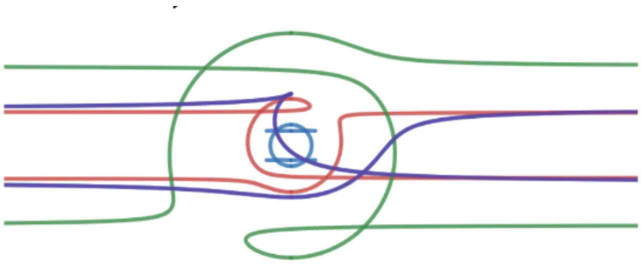
These are quartic curves, which for $c < 1/2$ have a 'loop' singularity, and are smooth for $c > 1/2$. The polynomial equations defining the quartic conchoids of Diocles are quartic, but we can obtain similar behavior in the cubic sense. These are the conchoids of de Sluze, described by the equation $(X-1)(X^2+Y^2) = aX^2$, which is equal to the conchoid of Diocles when $a = -1$.



Example. *Another example of an algebraic plane curve is the conchoid of Dürer, obtained by taking a pair of perpendicular lines intersecting at a point O, considering points Q and R moving on these lines such that the sum of the distances from O to Q and O to R is constant, and then taking the point on QR at a fixed distance from Q. If we take the perpendicular lines as the X and Y axis, with O the origin, take b as the sum of distances, and take a as the distance from Q, then each point (X,Y) lies on the curve if, first, it lies on a line PQ, where $Q = (x,0)$ and $P = (0,y)$, where $yX + xY = xy$, such that $x+y = b$, and $(X-x)^2 + Y^2 = a^2$. We can eliminate y from the equation since we can write $y = b-x$, so that $(b-x)X + xY = x(b-x)$. For a fixed X and Y, this equation is quadratic in x, which can be rewritten as $x^2 + bX = x(b+X-Y)$. The equation $(X-x)^2 + Y^2 = a^2$ gives $x^2 = a^2 - Y^2 - X^2 + 2xX$, hence $a^2-Y^2-X^2+bX = x(b-X-Y)$, which gives $(b-X-Y)x = a^2-Y^2-X^2+bX$. Finally, we obtain the constraints on X and Y by multiplying the equation $(b-x)X + xY = x(b-x)$ on both sides by $(b-X-Y)^2$ allows us to eliminate the remaining values of x, which can be rearranged to give the equation*

$$2y^2(x^2+y^2) + (b^2-3a^2)y^2 + 2a^2b(x+y) + a^2(a^2-b^2) = a^2x^2 + 2by^2(x+y)$$

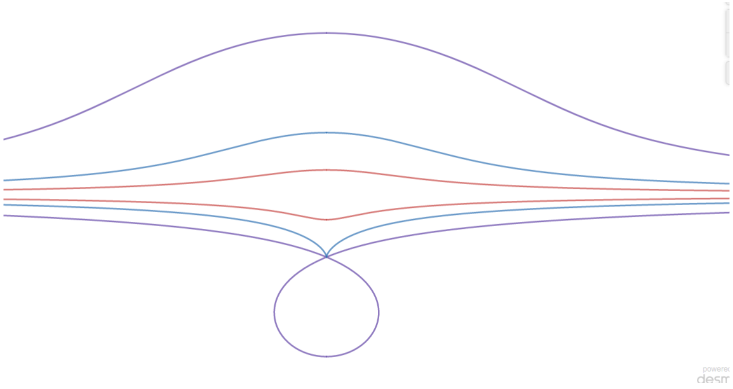
so the curve is a quartic curve. For $b = 0$, the curve becomes a pair of lines together with a circle, and for $a = 0$, we obtain two coincident straight lines.



Example. *The conchoid of Nicomedes is obtained by fixing a point P, and letting Q vary over a line not containing P. For each Q, the conchoid consists of the points on the line PQ at a fixed distance away from Q. If P lies at the origin, Q varies along the line $y = a$, and the distance parameter is b, then the equation describing the conchoid is*

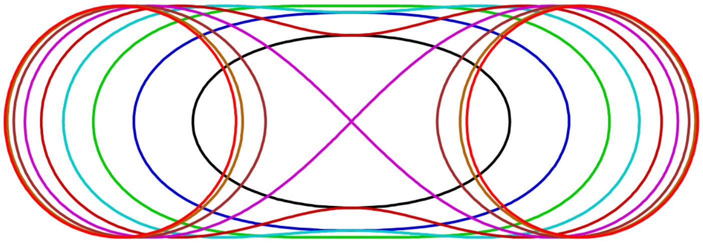
$$(Y - a)^2(X^2 + Y^2) = b^2Y^2.$$

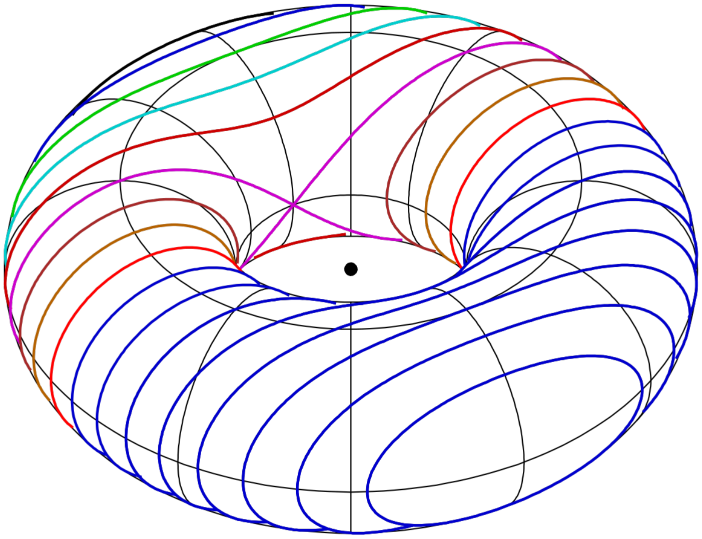
The conchoids appear to take three different forms depending on the relation between the distance between P and the line, and the distance defining the conchoid. If $a > b$, we obtain two smooth curves. If $a < b$, the conchoid appears to 'swing' around the point P, with P as a nodal point. If $a = b$, then we obtain a 'cusp' at P. An interesting feature of the conchoid of Nicomedes is that it is isochronous*; the time for an object to reach the 'bottom' of the conchoid under the influence of gravity starting from zero velocity is independent of it's original starting position on the curve.*



the time for an object to reach the 'bottom' of the conchoid under the influence of gravity is independent of its original starting position.

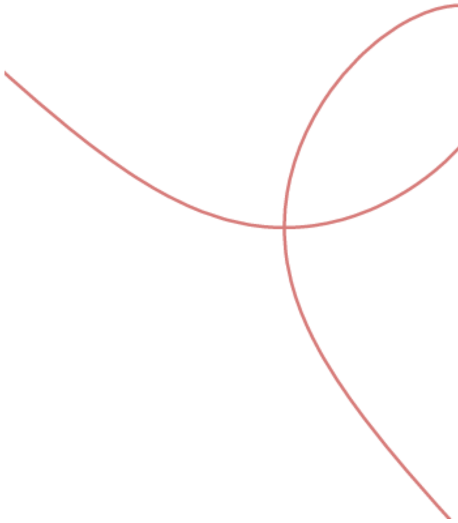
Example. *Just as we can obtain the conic sections by intersecting a cone with a parabola, the spiric sections of perseus are obtained by intersecting a plane with a torus. The general form of an equation describing a spiric section is of the form* $(X^2+Y^2)^2 = dX^2 + eY^2 + f$.





A particular family of spiric sections include the Cassini curves. These can be constructed by taking two focal points P and Q, and considering the locus of points such that the product of the distances to P and to Q are a fixed quantity. If we fix the focal points at $(-1,0)$ *and* $(1,0)$*, then an equation for the Cassini curve is* $(X^2+Y^2)^2 - 2(X^2-Y^2) + 1 = a^4$*, where a is the distance parameter to the curve. Cassini was an astronomer who believed the sun rotated around the earth according to these curves. The curves are smooth, except if we let* *the distance a be equal to* 1*, in which case the point* $(0,0)$ *is singular since the curve intersects twice here. This curve is the lemniscate of Bernoulli, described by the equation* $(X^2+Y^2)^2 = 2(X^2-Y^2)$*.*

Example. *The folium of Descartes is the algebraic curve defined by the equation* $X^3+Y^3 = 3XY$*. It's claim to fame is that was the curve that lead to the problem of implicit differentiation. In 1638, Descartes challenged Fermat to find the tangent line to the curve at any point on the circle, and with some primordial techniques of the calculus, Fermat was able to derive the tangent line at an arbitrary point.*



Example. *If we consider a circle lying on the outside of a circle, and we fix a point on the circle as it rotates around the circle, then provided that the circumference of the outer circle is a rational multiple of the circumference of the inner circle, we obtain an algebraic curve known as an epicycloid, and the rational multiple determined the period of rotation of the circle. If the inner circle has radius R, and the outer circle radius r, then we have a parameterization given by*

$$\left((r+R)\cos t - r\cos\left(\frac{r+R}{r}t\right), (r+R)\sin t - r\sin\left(\frac{r+R}{r}t\right)\right)$$

If the circumference of the inner circle is equal to the circumference of the outer circle, the outer circle completes a single rotation before returning to its original location, and this curve is known as a cardoid, because it has a single cusp *which looks like a heart. If the outer circle has half the circumference as the inner circle, the outer circle rotates twice before returning to its original position, and we obtain a function with a single cusp, and we call this shape a nephroid, since the shape looks like a kidney. Similarly, the hypocycloids are obtained by revolving a circle along the interior of the circle. If we revolve the outer circle around the inner circle, we obtain a pericycloid. If we fix a general point on these circles, we obtain the trochoids.*

Algebraic curves are a rich source of geometric problems. For this reason, they inspired the general theory of algebraic geometry. Here are some natural questions we can ask about algebraic curves:

- Is it possible to parameterize an algebraic curve's points by a rational function of a single argument? We call such curves *rational curves*. This is more difficult than it seems. For instance, the algebraic curve $Y^2 = X^2 + X^3$ has a rational parameterization. For each value of t, the line $Y = tX$ intersects the curve in a single position outside of the origin, because the solutions are given by nonzero values of X such that $(tX)^2 = X^2 + X^3$, so that $(t^2 - 1)X^2 = X^3$, so $X = t^2 - 1$, $Y = t^3 - t$ gives the unique point off the origin on the line $Y = tX$. But since every point on $Y^2 = X^2 + X^3$ lies on some line through the origin, we find that $(t^2 - 1, t^3 - t)$ gives a parameterization of the curve. This is not just a novel problem, because if we wish to perform an integration

  $$\int \varphi\left(x, \sqrt{x^2 + x^3}\right) dx \qquad \int \varphi\left(x, -\sqrt{x^2 + x^3}\right) dx$$

  where φ is a rational function in two variables, then we know that the substitution $x = t^2 - 1$ gives $\sqrt{x^2 + x^3} = t(t^2 - 1)$, so we are reduced to performing the integration

  $$\int \varphi(t^2 - 1, t(t^2 - 1)) 2t \, dt$$

  Thus the antiderivative of every rational function of x and $\sqrt{x^2 + x^3}$ is expressible in elementary terms. On the other hand, the cubic curve $Y^2 = X^3 + 1$ is not parameterized by a rational function of a single variable, and this is closely related to the fact that the integral

  $$\int \frac{dx}{\sqrt{x^3 + 1}}$$

is not expressible in terms of elementary functions.

- Another reason to study rational curves is to determine the points on a curve with rational coefficients. For instance, we can consider the rational solutions to the curve $Y^2 = X^2 + X^3$. We know that for each $t \in \mathbf{Q}$, $(t^2 - 1, t(t^2 - 1))$ gives a point on the curve with rational coordinates. Conversely, if $(t^2 - 1, t(t^2 - 1))$ is a rational coordinate, then t^2 is a rational number, and provided that $t^2 - 1 \neq 0$, we conclude that t is also a rational number. On the other hand, if $t^2 - 1 = 0$, then $t = \pm 1$ is obviously rational. Thus we can obtain all rational points on the curve by taking the function of a single rational number. Fermat's last theorem asks us to determine whether the equation $X^n + Y^n = Z^n$ has any integer solutions for $n > 2$, which is equivalent to the existence of rational solutions to the equation $X^n + Y^n = 1$, so the problem is very closely related to problems about algebraic curves.

- It is an important problem in geometry to classify geometric objects. The classical way to identify two curves is if they are equal to one another once we change our coordinate system. One invariant of this process is the *degree* of the curve, that is, the degree of the polynomial defining the curve. Thus we have degree one curves, which are lines, the degree two curves, the conics, which are after discounting degenerate solutions, classified into parabolas, hyperbolas, and ellipses. But the cubics give a whole new world; Newton gave a classification of the cubic curves into 72 families; Plücker into a more systematic 219 class system.

  To simplify the situation, we can 'weaken' the classification we use. If we view algebraic curves as lying in projective space rather than affine space, and identify curves by changing projective coordinates rather than affine coordinates, then the ellipse, hyperbola, and parabola are all identified as the same family of curves. We can imagine that the classification of cubics is also simplified considerably.

  Another way to simplify this situation is to identify two curves which are 'intrinsically the same', in the sense that we can map one curve onto another by a coordinate map given by polynomial equations, whose inverse can also be specified by polynomial equations. We call these isomorphisms *regular maps*. If we identify curves by rational functions rather than polynomials, we obtain the *birational maps* between curves. The curves birationally equivalent to a line are exactly the rational curves. These are the basic notions leading to the intrinsic theory of algebraic geometry.

Polynomials, being a rather restricted class of functions, defined a class of fairly well behaved curves. Aside from the class of smooth curves, however, they possess certain irregularities.

- Algebraic curves can have *singular points* where the curve is no longer 'smooth'. If an algebraic curve is defined with respect to a polynomial f, and $(\nabla f)(p) \neq 0$ where $f(p) = 0$, then we can locally describe one coordinate on the curve as a function of another curve, so the curve is smooth. But it is entirely possible for $(\nabla f)(p)$ to vanish, in which case the algebraic curve no longer behaves smoothly. One reason this can occur is if the algebraic curve has a *node*, which occurs if p is the intersection point of two smooth branches of the curve. Another reason is if the function rapidly changes direction, in which case we have a *cusp*. Unless we restrict the class of algebraic curves we are considering to the *non-singular* curves, then there is no way to avoid this issue, and we must face singular points head on.

- The zero sets of some polynomials do not have 'curve-like' solutions at all. For instance, the equation $X^2 + Y^2 = 0$ has only a single solution $(0,0)$, so given our present definition we have to agree that the set $\{(0,0)\}$ is an algebraic curve. This annoyance disappears when we study algebraic plane curves over the complex numbers, whose algebraic completeness means that $X^2 + Y^2$ factors into $(X + iY)(X - iY)$, so the solution set is the union of two planes, known as *complex lines*, through the origin, so the solution set behaves locally like a two dimensional space. A two-dimensional space is 'one-dimensional' over the complex numbers, so the solution set over the complex numbers behaves like a *complex curve*. This is where the theory of Riemann surfaces enter the picture, and we find an interesting interplay between analytic and algebraic viewpoints.

On the other hand, being of essentially algebraic character, most of the basic techniques can be formalized to study algebraic curves over any field.

In general, we shall assume some field k is fixed; the most elementary object of study will be the n dimensional affine space $\mathbf{A}^n$ over the field k, which can be identified with the space k^n of n tuples of field elements after a coordinate system is fixed. $\mathbf{A}^1$ is referred to colloquially as the affine line, and $\mathbf{A}^2$ as the affine plane. On a first glance, geometric intuition appears to break down over the finite fields, or other abstract fields, but surprisingly, the arguments which justify certain solutions to algebraic geometry over the complex numbers generalize to most other fields. The only specialization we may need to introduce is to assume that k is an algebraically closed field, but we can always obtain results over any field by embedding a field in its algebraic closure. Often, this even gives further geometric insight enabling us to prove relations entirely in the original field.

Example. *Here is an example where working over the complex numbers enables us to prove relations about geometry in the Euclidean plane. If p is a point outside a circle C, then it lies on precisely two tangent lines to the circle. We define the* polar line *to be the line through the two points on the circle whose tangents pass through p.*

If we assume without loss of generality that our circle is described by the equation $X^2 + Y^2 = 1$, then for each point $q = (x,y) \in C$, the tangent line L_q to C at q is described by the equation $xX + yY = 1$. If $p = (a,b)$, then the pair of points q such that $p \in L_q$ is precisely the family of points (x,y) such that $ax + by = 1$ and $x^2 + y^2 = 1$. Multiplying this second equation by a^2 and substituting $1 - by$ for ax, we conclude that

$$(a^2 + b^2)y^2 - 2by + (1 - a^2) = 0$$

The discriminant of this equation is $4a^2(a^2 + b^2 - 1)$, and so this equation has two distinct solutions provided that $a^2 + b^2 > 1$.

If the point p lies on the interior of the circle C, then $0 < a^2 + b^2 < 1$, then the discriminant is negative, which implies that p lies on no tangent to the circle. On the other hand, we can still solve these equations to determine two complex points on the 'complex circle'. The coordinates of the two complex points we find must be conjugates of each other, and so if (x,y) is one of these points then the line between these two points is given by

$$(\overline{x} - x)(Y - y) = (X - x)(\overline{y} - y)$$

Simplifying, the line is described by the equation

$$Im(x)Y - Im(y)X = Im(x)Re(y) - Im(y)Re(x)$$

a purely real line, which can still be interpreted in a polar sense. Geometrically, it is the locus of all points whose polar line passes through p. Thus we are lead to a construction which would not have seemed so simple would we have restricted ourselves to real space, and it leads naturally to an exploration of the theory of inversive geometry.

Later on, we will generalize our study from affine curves in $\mathbf{A}^n$ to *projective curves* in $\mathbf{P}^n$, where the projective space $\mathbf{P}^n$ is a form of 'compactification' of $\mathbf{A}^n$ which has various niceties when compared to $\mathbf{A}^n$. This compactness leads to many powerful algebraic consequences. For instance, in the study of algebraic curves, we obtain Bezout's theorem.

We shall return to the study of algebraic plane curves after we introduce some general tools from the framework of algebraic geometry. However, algebraic curves offer a nice source of nontrivial examples with which to try out the general tools of algebraic geometry. Moreover, they are historically the reason algebraic geometry was studied in the first place, and the aid us understand modern teminology from the historical development of the subject.

1.2 Affine Varieties

Given a polynomial $f \in k[X_1,\dots,X_n]$, we can consider the set

$$Z(f) = \{p \in \mathbf{A}^n : f(p) = 0\},$$

which is the *algebraic hypersurface* defined by f. We can also consider the geometric set obtained from the common zeroes of two polynomials $Z(f,g) = Z(f) \cap Z(g)$. More generally, given a set S of polynomials, we can consider the set $Z(S)$, which consists of the common zeroes of all polynomials in S. Sets formed from $\mathbf{A}^n$ by taking the common zeroes of a family of polynomials are refered to as *affine varieties*, and they are the main object of study in algebraic geometry. let us consider some elementary geometric properties.

Theorem 1.1. *A hypersurface Σ in $\mathbf{A}^n$ specified by a polynomial of degree d either contains a line, or intersects it in at most d places.*

Proof. Let Σ be specified as the locus of a polynomial f. Fix $x_0, t_0 \in k$. For any line L described by some equation $X = x_0 + t_0 Y$, the points on $\Sigma \cap L$

are in one to one correspondence with the zeroes of the polynomial $g(Y) = f(x_0+t_0Y, Y) \in k[Y]$, and $\deg(g) \leqslant \deg(f) = d$. thus, unless $g = 0$, in which case Σ contains L, the surface Σ intersects L in at most d places. □

Remark. A simple consequence of this theorem is that

$$\{(x,y) \in \mathbf{R}^2 : y = \sin(x)\}$$

is not an algebraic curve in $\mathbf{R}^2$, since it intersects the x-axis infinitely many times but does not contain the x-axis. Neither is the set

$$\{(z,w) \in \mathbf{C}^2 : |z|^2 + |w|^2 = 1\}$$

since the intersection of the complex sphere with the z axis is a circle, which has infinitely many points.

There are some elementary observations we can make on the construction $Z(S)$ from a set of polynomials S, which open the floodworks to reducing geometric problems on varieties to the ring theory of $k[X_1,\dots,X_n]$.

- If $S \subset T$, then $Z(T) \subset Z(S)$.
- If $\mathfrak{a}$ is the smallest ideal containing S, then $Z(S) = Z(\mathfrak{a})$, so every affine variety can be described as the common zeroes of some ideal in $k[X_1,\dots,X_n]$.
- If we have a family $\{\mathfrak{a}_\alpha\}$ of ideals, then $Z(\bigoplus \mathfrak{a}_\alpha) = \bigcap Z(\mathfrak{a}_\alpha)$, so the intersection of an arbitrary family of varieties forms a variety.
- For any two polynomials f and g, $Z(fg) = Z(f) \cup Z(g)$. More generally, if $\mathfrak{a}$ and $\mathfrak{b}$ are ideals, then $Z(\mathfrak{a}\mathfrak{b}) = Z(\mathfrak{a}) \cup Z(\mathfrak{b})$, so finite unions of varieties are varieties.
- $Z(0) = \mathbf{A}^n$, $Z(1) = \varnothing$, and for any $a \in k^n$, $Z(X_1 - a_1,\dots,X_n - a_n)$ is just the singleton set $\{a\}$. It follows from the last point that finite point sets are varieties.

The largest heuristic that guides the subject of algebraic geometry is that geometric properties of a variety are completely summarized in the algebraic structure of the ring of functions $k[X_1,\dots,X_n]$ acting on this variety, and other algebraic structure one can construct from this ring. The correspondence is strengthened tenfold when the field k is algebraically closed, since then Hilbert's nullstellensatz applies to give powerful results. Before we discuss the nullstellensatz, let us explore some algebraic properties of $k[X_1,\dots,X_n]$ and the relation of these properties to various geometric properties of affine varieties.

Recall that $k[X_1,\dots,X_n]$ is a *Noetherian ring*; any ideal in $k[X_1,\dots,X_n]$ is finitely generated. Since any algebraic variety is formed from the common zeroes of polynomials found in an ideal $\mathfrak{a}$ of $k[X_1,\dots,X_n]$, it follows that any algebraic variety is formed from the common zeroes of a *finite family* of polynomials. Thus the class of algebraic varieties has a combinatorial nature not found in other families of surfaces, like the space of smooth submanifolds of $\mathbf{A}^n$.

Example. *The subvarieties of $\mathbf{A}^1$ are exactly the finite point sets, aside from $\mathbf{A}^1$ itself. This follows precisely because the zero set of any polynomial $f \in k[X]$ is finite. This makes algebraic geometry in $\mathbf{A}^1$ essentially trivial.*

Example. *If k is finite, any subset of $\mathbf{A}^n$ is an algebraic variety, because all subsets of $\mathbf{A}^n$ are finite subsets. In order to study subsets of finite fields using polynomials, we need to place more specifications on the ideals we use in $\mathbf{A}^n$ so that interesting subsets of $\mathbf{A}^n$ are carved out.*

Proposition 1.2. *If k is algebraically closed, and $f \in k[X_1,\dots,X_n]$ is non-constant, then $\mathbf{A}^n - Z(f)$ contains infinitely many points. If $n \geqslant 2$, then $Z(f)$ consists of infinitely many points.*

Proof. First, recall that every algebraically closed field is infinite. It follows that $\mathbf{A}^1 - Z(f)$ is infinite for any polynomial $f \in k[X]$, because $Z(f)$ is finite. Given any non-constant polynomial $f \in k[X_1,\dots,X_n]$, there is a line L in $\mathbf{A}^n$ upon which f is not identically zero (for otherwise we would have $f(p) = 0$ for all $p \in \mathbf{A}^n$, which would imply $f = 0$). But reducing our argument to the one dimensional case shows that $L - Z(f)$ is infinite, hence $\mathbf{A}^n - Z(f)$ is infinite.

We can argue similarily to prove the second statement. Given any $f \in k[X_1,\dots,X_n]$, there is a plane upon which f does not vanish identically, so we are reduced to proving this theorem in $\mathbf{A}^2$. Given any nonconstant polynomial $f \in k[X,Y]$, we can write

$$f(X,Y) = \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} X^i Y^j.$$

Without loss of generality, we may assume that $Z(f)$ does not contain the origin. For each $t \in k$, we consider the line L_t through the origin described by the equation $Y = aX$. Since the family of sets $L_t \cap Z(f)$ is disjoint, it suffices to show that L_t is non-empty for infinitely many t. Now the points on $L_t \cap Z(f)$ are in one-to-one correspondence with the zeroes of the polynomial

$$f(X, tX) = \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} t^j X^{i+j}$$

This polynomial has a zero provided it is non-constant, and this polynomial is only constant if for all $1 \leqslant m \leqslant 2N$,

$$\sum a_{(m-k)k} t^k = 0.$$

But for each m, there are only finitely many t satisfying this equation, which implies there are only finitely many t such that $L_t \cap Z(f) = \varnothing$. □

If X is any subset of $\mathbf{A}^n$, then we shall let $I(X)$ be the subset of $k[X_1,\dots,X_n]$ of polynomials which vanish over X. The set forms an ideal, and it is clear that in the case where $X = Z(S)$, the ideal contains all elements of S, hence all elements of $\mathfrak{a}$. The generation of an ideal $I(X)$ from a set X is dual to the notion of generating a set $Z(\mathfrak{a})$ from an ideal $\mathfrak{a}$. We make a few elementary observations about this operator.

- If $X \subset Y$, then $I(Y) \subset I(X)$.

- $I(\varnothing) = k[X_1,\dots,X_n]$, and $I(\mathbf{A}^n) = (0)$.

- $S \subset I(Z(S))$ for any subset S of polynomials, and $X \subset Z(I(X))$. In fact, $Z(I(X))$ is the smallest variety in $\mathbf{A}^n$ containing X, the *Zariski closure* of X.

- The last point implies that for any variety V, $Z(I(V)) = V$. It is simple to see that $V \subset Z(I(Z))$. Conversely, if $V = Z(S)$ for some $S \subset k[X_1,\dots,X_n]$, then $S \subset I(V)$, so $Z(I(V)) \subset Z(S) = V$. Similarily, if $\mathfrak{a}$ is an ideal in $k[X_1,\dots,X_n]$ which is equal to $I(X)$ for some set X, then $I(Z(\mathfrak{a})) = \mathfrak{a}$.

- If $f \in k[X_1,\dots,X_n]$ and $f^k \in I(X)$ for some k, then $f \in I(X)$. This means exactly that $I(X)$ is a *radical ideal* of $k[X_1,\dots,X_n]$. The smallest radical ideal containing some ideal $\mathfrak{a}$ is denoted $\text{Rad}(\mathfrak{a})$.

Proposition 1.3. *If V and W are varieties, $I(V) = I(W)$ if and only if $V = W$.*

Proof. This follows because $Z(I(V)) = V$ and $Z(I(W)) = W$. □

This simple proposition is clearly not true if V and W are not algebraic sets, hinting that the class of varieties is well separated by polynomial functions. We will soon see that if we are working over an algebraically closed field, and $\mathfrak{a}$, $\mathfrak{b}$ are radical ideals, then $Z(\mathfrak{a}) = Z(\mathfrak{b})$ holds if and only if $\mathfrak{a} = \mathfrak{b}$, so there is a one two one correspondence between radical ideals and algebraic sets. This constitutes the theory of Hilbert's nullstellensatz, which we will come back to later in this chapter.

Corollary 1.4. *If V is an algebraic set in $\mathbf{A}^n$, and $p \notin V$, then there is a polynomial $f \in I(V)$ with $f(p) = 1$.*

Proof. Since V and $V \cup \{p\}$ are both algebraic sets, $I(V) - I(V \cup \{p\})$ is nonempty, so there must be $f \in I(V)$ which vanishes on V, but with $f(p) \neq 0$¡ it follows by normalizing that we can assume $f(p) = 1$. □

Similarly, by taking an algebraic set V, and n points $p_1, \dots, p_n \notin V$, we may apply this theorem to find polynomials $f_1, \dots, f_n \in I(V)$ with $f_i(p_j) = \delta_{ij}$ for each $1 \leqslant i, j \leqslant n$. By considering linear combinations of the f_i, for any $a_{ij} \in k$, we can find $f_1, \dots, f_n \in I(V)$ with $f_i(p_j) = a_{ij}$. This shows the space of polynomials which vanish over V has enough degrees of freedom to specify values on a finite set of points outside of V.

1.3 The Zariski Topology

The family of varieties is closed under infinite intersection, and finite unions. This means we can form a topology on $\mathbf{A}^n$ by declaring all varieties to be closed subsets of the topology. This topology is known as the Zariski topology, and a set will be called *Zariski open* or *Zariski closed* if it is open or closed with respect to this topology. Similarily, one takes the induced Zariski topology induced on any affine variety, whose closed subsets are precisely subvarieties of the given variety.

The Zariski topology is *not* Hausdorff, which is slightly paradoxical at first as a useful topological space, but is still T1. It can be used as a useful way to describe geometric / algebraic properties of a variety V. For instance, the previously described *Zariski closure* is precisely the closure operation in the Zariski topology.

Example. *Every variety in* $\mathbf{A}^1$ *consists of finitely many points, which means the Zariski topology* $\mathbf{A}^1$ *is the* cofinite topology*: a set is Zariski open if and only if it is the empty set, or it's complement is finite.*

For any $f \in k[X_1,\dots,X_n]$, the set $U_f = \{p \in \mathbf{A}^n : f(p) \neq 0\}$ is a Zariski open set. Conversely, if U is any Zariski open set, then there exists $f_1,\dots,f_n \in k[X_1,\dots,X_n]$ such that $U^c = Z(f_1,\dots,f_n)$. It thus follows that $U^c = U_{f_1} \cup \dots \cup U_{f_n}$, which implies the family of sets $\{U_f\}$ form a *basis* for the Zariski topology. These sets are known as *basic open sets*. One can perform a similar construction for the Zariski topology on any variety.

1.4 Reducibility

An algebraic variety V is said to be *reducible* if it can be written as the union of two proper algebraic subsets. Otherwise, we say V is *irreducible*. Ring theory allows us to characterize this criterion in terms of the ideal generating the ideal.

Proposition 1.5. *A variety V is irreducible if and only if $I(V)$ is prime.*

Proof. Suppose that $I(V)$ is not prime, so there is $f,g \in k[X_1,\dots,X_n]$ such that $fg \in I(V)$, but $g \notin I(V)$. It follows that f cannot be a scalar multiple of g, because $I(V)$ is a radical ideal. The fact that $f \notin I(V)$ and $g \notin I(V)$ means that f and g do not vanish on V, so $Z(f,I(V))$ and $Z(g,I(V))$ are proper *subvarieties* of V. But this means that V is reducible.

Conversely, if V is reducible, we can write $V = V_1 \cup V_2$ for two proper subvarieties V_1 and V_2 of V. Then $I(V)$ is a proper subset of $I(V_1)$ and $I(V_2)$, so we may select $f \in I(V_1) - I(V)$ and $g \in I(V_2) - I(V)$. Then $fg \in I(V)$, which shows $I(V)$ is not prime. □

Example. *The parabola $V = Z(Y - X^2)$ is an irreducible variety, provided we are working over an infinite field. It is easy to verify that $Y - X^2$ is irreducible, and since $k[X,Y]$ is a unique factorization domain, this implies $(Y - X^2)$ is prime. Provided that $I(V) = (Y - X^2)$, this verifies that V is irreducible.*

If $f \in I(V)$, then we may apply the division algorithm to write

$$f = g(Y - X^2) + h,$$

for some $g \in k[X,Y]$ and $h \in k[X]$. Since $f \in I(V)$, $f(x,x^2) = 0$ for all $x \in k$, which implies $h(x) = 0$ for all $x \in k$; provided k is infinite, this means $h = 0$, so f is divisible by $Y - X^2$. Thus $I(V) = (Y - X^2)$.

In most of mathematics, it is often a useful strategy to break objects into 'atomic' components, understand these components, then understand the more general objects by forming these components back together. Algebraic geometry is no different, where it is often a useful strategy to break down varieties into a finite union of irreducible varieties. The idea to break the variety down is simple. If a variety V is not irreducible, then we can break it apart into two proper algebraic subsets $V_1 \cup W_1$. If V_1 is not irreducible, we can break it apart into two proper subsets $V_2 \cup W_2$. If this process is guaranteed to terminate at some point (so that V_n is eventually irreducible), we can recursively break apart varieties into irreducible varieties. This is guaranteed by the fact that $k[X_1, \dots, X_n]$ is Noetherian.

Lemma 1.6. *If $\mathfrak{a}$ is an ideal of a Noetherian ring A, then out of the set of prime ideals containing $\mathfrak{a}$, there are only finitely many minimal ones.*

Proof. If there was an ideal $\mathfrak{a}$ forming a counterexample to this lemma, then using the Noetherian property of A, there would certainly be a maximal such counterexample. The ideal $\mathfrak{a}$ could not be prime, because then there would only be a single minimal prime ideal containing $\mathfrak{a}$. Thus there is $x, y \in A$ with $xy \in I$, but with $x, y \notin I$. Thus $(x) + \mathfrak{a}$ and $(y) + \mathfrak{a}$ are ideals bigger than I, and thus are not counterexamples to the theorem. Thus there are only finitely many prime ideals minimal with respect to contains $(x) + \mathfrak{a}$ and $(y) + \mathfrak{a}$ respectively. But if $\mathfrak{p}$ is a prime ideal containing $\mathfrak{a}$, then $xy \in \mathfrak{p}$, so either $x \in \mathfrak{p}$, or $y \in \mathfrak{p}$, implying that either $(x) + \mathfrak{a} \subset \mathfrak{p}$, or $(y) + \mathfrak{a} \subset \mathfrak{p}$. Thus a minimal prime ideal containing $\mathfrak{a}$ must be a minimal prime ideal in $(x) + \mathfrak{a}$ or $(y) + \mathfrak{a}$, which is a contradiction to the fact that there are infinitely many such prime ideals. □

Theorem 1.7. *Every variety can be written uniquely as the finite union of irreducible varieties, with no irreducible variety containing the other. These irreducible varieties are known as the* irreducible components *of the variety they form.*

Proof. If V is a variety, there are only finitely many minimal prime ideals containing $I(V)$, and thus there are only finitely many maximal irreducible varieties contained in V. Now if V can be written as the union of two families of irreducible varieties $V_1, \dots, V_N$ and $W_1, \dots, W_M$, then $V_n = \bigcup (V_n \cap W_m)$, so either $V_n \cap W_m = \varnothing$ or $V_n \cap W_m = V_n$, which implies $V_n \subset W_m$ for some m. Performing the same process in reverse gives

$W_m \subset V_{n'}$ for some n, and obviously $n = n'$. By matching up elements of the decomposition, we conclude that the decomposition is unique. $\square$

Example. *Consider the variety* $V = Z(Y^4 - X^2, Y^4 - X^2Y^2 + XY^2 - X^3)$ *in* $\mathbf{C}^2$. *Since* $Y^4 - X^2 = (Y^2 - X)(Y^2 + X)$, *we find*

$$Y^4 - X^2Y^2 + XY^2 - X^3 = (Y + iX)(Y - iX)(Y - X)(Y + X)$$

Considering the zeroes which satisfy these equations on a case by case basis, we find that V *is just a set of discrete points, each an irreducible factor in the decomposition of the variety.*

Example. *The polynomial* $Y^2 + X^2(X - 1)^2$ *is irreducible over* $\mathbf{R}[X, Y]$, *but factors into* $(Y + iX(X - 1))(Y - iX(X - 1))$ *over* $\mathbf{C}[X, Y]$. *The consequence is that even though* $Y^2 + X^2(X - 1)^2$ *is an irreducible polynomial, the variety it generates is not irreducible, consisting of the two points* $(0, 0)$ *and* $(1, 0)$. *This is a consequence of the fact that over the real numbers,*

$$I(Z(Y^2 + X^2(X - 1)^2) = (Y, X(X - 1)) \neq (Y^2 + X^2(X - 1)^2)$$

is not a prime ideal.

Note that the Noetherian property of $k[X_1, \dots, X_n]$ also implies that if V is a variety containing a family of subvarieties $\{V_\alpha\}$, then there exists a minimal member of this family. Thus if $\{U_\alpha\}$ is a cover of V by Zariski open subsets, then we consider the family of Zariski closed sets $V_{\alpha_1 \dots \alpha_n} = V - U_{\alpha_1} - \dots - U_{\alpha_n}$. This family must have a minimal member, which must be the empty set because the family $\{U_\alpha\}$ covers V. This implies the open cover has a finite subcover. Thus V is compact in the Zariski topology. In algebraic geometry we often say V is *quasi-compact* instead of compact, because in topological spaces that are not Hausdorff, compact sets do not behave quite as nicely than in spaces that are Hausdorff (for instance, a quasi-compact set need not be closed).

Theorem 1.8. *Any nonempty Zariski open subset of an irreducible variety is Zariski dense.*

Proof. Let V be an irreducible variety, and U a Zariski open subset. Let W be the Zariski closure of U. Then W and $V - U$ are two Zariski closed subsets of V covering V. The irreducibility of V thus implies that either $W = V$, or $V - U = V$. The fact that U is non-empty implies that the former case is true. $\square$

1.5 Classification of Planar Algebraic Sets

It an interesting task to classify the algebraic subsets of $\mathbf{A}^2$, because it is the first nontrivial family of algebraic sets. Whereas the algebraic subsets of $\mathbf{A}^1$ are trivial, the plane contains numerous infinite families of varieties, such as parabolas, ellipses, hyperbolas, and elliptic curves. We begin with a simple observation.

Theorem 1.9. *If two polynomials $f, g \in k[X,Y]$ are relatively prime, then $Z(f,g)$ consists of finitely many points.*

Proof. If f and g have no common factor over $k[X,Y]$, then by the Gauss lemma they also have no common factor over $k(X)[Y]$, and because $k(X)[Y]$ is a Euclidean domain, we may write $af + bg = 1$ for some $a, b \in k(X)[Y]$. If $a(X,Y) = \sum a_i(X) \cdot Y^i$ and $b(X,Y) = \sum b_i(X) \cdot Y^i$, then we may find a nonzero $c \in k[X]$ such that $ca, cb \in k[X,Y]$. This implies that $(ac)f + (bc)g = c$. It follows that for any x for which there exists y with $f(x,y) = g(x,y) = 0$, $c(x) = 0$. Since c has only finitely many roots, we conclude that $f(x,y)$ and $g(x,y)$ equal zero at only finitely many values of x. By symmetry, they can also only simultaneously be zero at finitely many values of y. And so it follows that there are only finitely many pairs (x,y) such that $f(x,y) = g(x,y) = 0$. □

Remark. It follows from this that the Zariski topology induced on a planar curve is precisely the cofinite topology; a set is closed if and only if it is empty, or it's complement is finite. It follows that any bijective mapping from one planar curve to another is automatically a homeomorphism.

Corollary 1.10. *If $f \in K[X,Y]$ is irreducible, and $Z(f)$ is infinite, then*

$$I(Z(f)) = (f),$$

and $Z(f)$ is irreducible.

Proof. If $g \in I(Z(f))$, then $Z(f,g) = Z(f)$ is infinite, so f and g must have a common factor, hence f must divide g since f is irreducible. Thus we conclude that $I(Z(f))$ consists only of multiples of f. □

Corollary 1.11. *The irreducible algebraic planar sets over an infinite field are exactly $\mathbf{A}^2$, $\varnothing$, singletons, and irreducible plane curves $Z(f)$, where f is irreducible and $Z(f)$ is infinite.*

Proof. It is obvious that $\{p\}$ is an irreducible set, as is $\emptyset$. Since k is an infinite field, it is also obvious that $\mathbf{A}^2$ is irreducible, since $I(\mathbf{A}^2) = (0)$ is irreducible. Any other irreducible algebraic set must be of the form $Z(f)$ for some irreducible polynomial f, and these are the irreducible planar curves provided $Z(f)$ is infinite, by the last lemma. □

We have seen that if k is algebraically closed, then we have seen that for any non-constant polynomial $f \in k[X,Y]$, $Z(f)$ consists of infinitely many points. Thus $Z(f)$ is irreducible if f is irreducible.

Corollary 1.12. *If k is algebraically closed, and $f = f_1^{n_1} \dots f_m^{n_m}$ where each f_i is irreducible, then the irreducible components of $Z(f)$ are exactly the $Z(f_i)$, and $I(Z(f)) = (f_1, \dots, f_n)$.*

Proof. It is clear that

$$Z(f) = Z((f_1^{n_1}) \dots (f_m^{n_m})) = \bigcup Z(f_i^{n_i}) = \bigcup Z(f_i)$$

and that each f_i is irreducible. Since our field is algebraically closed, each $Z(f_i)$ is infinite, If $Z(f_i) \subset Z(f_j)$, then $Z(f_i, f_j) = Z(f_j)$, and since $Z(f_j)$ is infinite, this implies that f_j divides f_i, which is impossible. Thus the $Z(f_i)$ really are the decomposition of $Z(f)$. □

Example. *Over the real numbers, there is not a one-to-one correspondence between prime ideals and irreducible varieties. For instance, $X^2 + Y^2 + 1$ is an irreducible polynomial, but*

$$Z(X^2 + Y^2 + 1) = \emptyset = Z(\mathbf{R}[X,Y]).$$

This is the first of many algebraic deficiencies of non algebraically closed fields, which is one of the reasons we will soon switch to studying algebraically closed fields.

Example. $Y^2 - X(X^2 - 1)$ *is an irreducible polynomial over* $\mathbf{R}$ *and* $\mathbf{C}$, *and its solution set is infinite, so in both cases $Z(Y^2 - X(X^2 - 1))$ is an irreducible variety. But over the real numbers, $Z(Y^2 - X(X^2 - 1))$ has two connected components in the Euclidean topology, a union of a circle and a line. On the other hand, over the complex numbers the solution set of $Y^2 - X(X^2 - 1)$ is a connected set which can be written as the union of the two branches of the function $Y = \sqrt{X(X^2 - 1)}$, and since the Riemann surface corresponding to the square root operation is homeomorphic to* $\mathbf{C}$, *the solution set of this polynomial is also homeomorphic to* $\mathbf{C}$; *it has three singularities at* $(-1,0)$, $(0,0)$, *and* $(1,0)$, *and the solution set behaves like a cone around these solution sets.*

Example. *Over the real numbers, $X^3 + X - X^2Y - Y = (X - Y)(X^2 + 1)$ is just the line $X = Y$, and hence $Z(X^3 + X - X^2Y - Y)$ is irreducible. However, over the complex numbers, X is the union of the three lines $X = Y$, $X = i$, and $X = -i$, and is therefore reducible.*

1.6 The Nullstellensatz

In the last few sections, we have seen the duality between affine varieties and radical ideals over the ring $k[X_1,\dots,X_n]$. Over algebraically closed fields, the correspondence between radical ideals and algebraic sets becomes exact. This is the content of Hilbert's nullstellensatz theorem. A precursor to the nullstellensatz, known as Study's lemma will suffice for the study of planar algebraic curves.

Theorem 1.13 (Study)**.** *If k is algebraically closed, and $f, g \in k[X,Y]$ satisfy $Z(f) \subset Z(g)$, where f is irreducible, then f divides g.*

Proof. If f did not divide g, then f and g would be relatively prime, so $Z(f,g)$ consists of finitely many points. But $Z(f)$ is contained in $Z(f,g)$, which is impossible since $Z(f)$ contains infinitely many points. □

In terms of ideals, Study's lemma implies that if f is a irreducible polynomials, then $I(Z(f)) = (f)$. Hilbert's first generalizes this result to saying that if f is an irreducible polynomial in any dimension, then $I(Z(f)) = (f)$, and more generally, if f vanishes on a variety $Z(\mathfrak{a})$, then $f^n \in \mathfrak{a}$ for some integer n.

Lemma 1.14 (Weak Nullstellensatz)**.** *If k is an algebraically closed field, and if $\mathfrak{a}$ is a proper ideal of $k[X_1,\dots,X_n]$, then $Z(\mathfrak{a}) \neq \varnothing$.*

Proof. We shall actually prove that if $\mathfrak{a}$ is a maximal ideal, then $Z(\mathfrak{a})$ is a set containing a single point. Since we may always extend every ideal to a maximal ideal, this will prove the proposition. So we take $\mathfrak{a}$ to be any maximal ideal. Then $k[X_1,\dots,X_n]/\mathfrak{a} = L$ is a field, which can be viewed as a field extension of k because we can embed k as the set of constant polynomials in $k[X_1,\dots,X_n]$, and we then compose with the quotient homomorphism to obtain a map into L. We write $x_i \in L$ for the element of the field corresponding to X_i. If we know that the embedding of k in L gives an isomorphism between the two fields, then for each x_i there is

$a_i \in k$ with $x_i - a_i \in \mathfrak{a}$. But $(X_1 - a_1, \dots, X_n - a_n)$ is a maximal ideal in $k[X_1, \dots, X_n]$, hence $\mathfrak{a} = (X_1 - a_1, \dots, X_n - a_n)$. Now we can conclude that $Z(\mathfrak{a}) = \{(a_1, \dots, a_n)\}$. □

An important thing to note about this proof of the weak nullstellensatz is that it implies that the maximal ideals of $k[X_1, \dots, X_n]$ are in one to one correspondence with the points of $\mathbf{A}^n$. The fact that maximal ideals are in one to one correspondence with points in space occurs in other context of mathematics, for instance, in the ring theory of $C(X)$, where X is Hausdorff and locally compact. This point of view is often so useful that, when we study general rings A, we consider the set of maximal ideals of A as points in a space, and then viewing elements of A as functions on this space. This idea reoccurs later in our study of the local rings attached to varieties, and in greater generality in the study of schemes.

To finish off the proof of the weak nullstellensatz, it suffices to prove that if k is an algebraically closed field, then for every field L, if there is a surjective homomorphism from $k[X_1, \dots, X_n]$ to a field extension L of k fixing elements of k, then $k = L$. This is an easy consequence of Zariski's lemma, which we prove now.

Lemma 1.15. *If $k[x_1, \dots, x_n]$ is a field, then it is a finite extension of k.*

Proof. We prove this by induction on n. For $n = 1$, this is a classical argument in Galois theory. To continue the induction, suppose we have proved the theorem for all fields of the form $k[x_1, \dots, x_m]$, where $m < n$. We may then apply induction to $k[x_1, \dots, x_n] = k(x_1)[x_2, \dots, x_n]$ to conclude that $k[x_1, x_2, \dots, x_n]$ is a finite extension of $k(x_1)$. But this means that there are rational functions $a_0(x_1), \dots, a_{N-1}(x_1)$ of x_1 such that $x_1^N = \sum a_n(x_1) x_1^n$. But now there exists a polynomial $f \in k[x_1]$, such that $f(x_1)a_n(x_1) \in k[x_1]$ for all n, and we find

$$x_1^N = \sum f(x_1) a_n(x_1) x_1^n$$

Which means x_1 is algebraic over k, and therefore a finite extension of k. This completes the proof. □

Since all finite extensions of a field are algebraic over that field, we conclude that if k is algebraically closed, then every field of the form $k[x_1, \dots, x_n]$ is an algebraic extension of k, and therefore $k[x_1, \dots, x_n] = k$. This finishes our proof of the weak nullstellensatz. We now consider the extension to the full nullstellensatz.

Theorem 1.16. *Suppose k is algebraically closed. If $\mathfrak{a}$ is an ideal in $k[X_1,\dots,X_n]$, then $I(Z(\mathfrak{a})) = Rad(\mathfrak{a})$, where $Rad(\mathfrak{a})$ is the set of all f for which there exists n with $f^n \in \mathfrak{a}$.*

Proof. We may assume that $\mathfrak{a}$ is generated by finitely many polynomials, so $\mathfrak{a} = (f_1,\dots,f_m)$. Concretely, the nullstellensatz then says that if g vanishes on the common nullset of the $f_1,\dots,f_m$, then $g^N = \sum h_k f_k$ for some $h_m \in k[X_1,\dots,X_n]$. Suppose that $g \in I(Z(\mathfrak{a}))$. Consider the ideal $\mathfrak{b} = (f_1,\dots,f_m, X_{n+1}g - 1) \subset k[X_1,\dots,X_{n+1}]$. Then $Z(\mathfrak{b}) = \emptyset$, since if $f_1(x) = \dots = f_n(x) = 0$, then $g(x) = 0$, so $x_{n+1}g(x) - 1 = -1$. The weak nullstellensatz implies that there are $a_k \in k[X_1,\dots,X_{n+1}]$ such that

$$\sum a_k f_k + b(X_{n+1}g - 1) = 1$$

Introducing $Y = 1/X_{n+1}$, we may multiply the equation by Y^N for a large enough N to find that

$$Y^N = \sum Y^N a_k f_k + bY^{N-1}(g - Y)$$

Where the Y in $Y^N a_k$ can $Y^{N-1}b$ can be used to cancel out all instances of X_{n+1}. Setting $Y = g$ gives the required equation over $k[X_1,\dots,X_n]$. □

Corollary 1.17. *There is a one to one correspondence between radical ideals and algebraic sets in affine space over an algebraically closed field.*

Corollary 1.18. *If $\mathfrak{a}$ is a prime ideal, then it is also a radical non total ideal, so $Z(\mathfrak{a}) \neq \emptyset$ is an irreducible algebraic variety, and there is a one to one correspondence with such prime ideals and irreducible varieties. The maximal ideals correspond to points in $\mathbf{A}^n$.*

Example. *$Z(Y^2 - X(X-1)(X-\lambda))$ is an irreducible planar curve in $\mathbf{A}^2$ in any algebraically closed field, because $Y^2 - X(X-1)(X-\lambda)$ is an irreducible polynomial. If the polynomial does factor, it factors as $(Y + f(X))(Y - f(X))$ where $-f(X)^2 = X(X-1)(X-\lambda)$, but then this equation has no solution because $X(X-1)(X-\lambda)$ isn't a square of a polynomial in $k[X]$, which is a unique factorization domain.*

Corollary 1.19. *If $f \in k[X_1,\dots,X_n]$ has a decomposition as $f_1^{n_1}\dots f_m^{n_m}$, where k is algebraically closed, then $Z(f) = Z(f_1 \dots f_n) = \bigcup Z(f_i)$ is the decomposition of f into its irreducible factors. There is a one to one correspondence (up to scalar multiples) between irreducible hyperplanes and irreducible polynomials.*

It is clear that if k is not an algebraically closed field, then the weak nullstellensatz cannot hold, because in one dimension, the weak nullstellensatz is exactly the condition that implies k is algebraically closed. Since $k[X]$ is a principal ideal domain, the nullstellensatz states that if $(f) \neq k[X]$, then $Z(f) \neq 0$, which means that if f is a non constant polynomial, then f has a root. The reason we work over algebraically closed fields is so that we have enough points on our algebraic curves to ensure a correspondence between ideals and polynomials.

1.7 Projective Varieties

Projective geometry is a natural extension of affine geometry in which two lines always have a unique point of intersection. In the same sense, projective geometry plays an important role in the theory of curves, because the fact that two lines intersect in a unique position extends to the fact that two curves of degree n and m in the projective plane intersect in nm locations. More generally, projective space provides a 'completion' or 'compactification' of affine space which proves useful in a great many problems in algebraic geometry. In order to employ projective space, we must consider *projective varieties* rather than just affine varieties.

We may view projective space as a compactification of affine space, adding 'asymptotic' intersection points to varieties in affine space. In other words, we obtain a projective variety from an affine variety by taking the directions that the variety approaches asymptotically as additional points in the space. Recall that if V is a vector space, then we define the *projectivization* $\mathbf{P}V$ to be the space of all lines through the origin in V, which can also be identified as a quotient space of V modulo the group action over the multiplicative group of nonzero elements of k given by scalar multiplication. To obtain a form of algebraic geometry over projective spaces, we must consider a natural coordinate system on these vector spaces. If we consider $\mathbf{P}^n = \mathbf{P}(\mathbf{A}^{n+1})$, then the natural choice of coordinates are the homogenous coordinates $[X_1, \dots, X_n, X_{n+1}]$, for $X_1, \dots, X_{n+1} \in k$ not *all* zero, which stand for the line generated by the vector $(X_1, \dots, X_{n+1}) \in \mathbf{A}^{n+1}$, so that $[X_1, \dots, X_{n+1}] = [\lambda X_1, \dots, \lambda X_{n+1}]$, where $\lambda \neq 0$. Note that the values X_k are not actually *functions* on $\mathbf{P}^n$, because they are not invariant under scalar multiples. On the other hand, the ratios X_i/X_j are functions, at least on the subset of points of $\mathbf{P}^n$ where X_j doesn't vanish.

Now that we have homogenous coordinates, we can identify subsets of projective space defined by polynomials in the homogenous coordinates. We would like to consider the zero sets of polynomials f in the homogenous coordinates $X_0, \dots, X_n$ as subsets of projective space, but unfortunately $f(\lambda x)$ need not equal $f(x)$ for $\lambda \neq 0$, so an arbitrary polynomial does not descend to a map on projective space. However, if f is *homogenous* of some degree k, then $f(\lambda x) = \lambda^k f(x)$, and this in particular implies that the zero set of f in $\mathbf{A}^{n+1}$ is a union of lines through the origin, and thus the zero set of f can be considered a subset of projective space. Given a set S of homogenous polynomials, we let $Z(S)$ denote the projective variety which is the locus of the points S, and we call these sets *projective varieties*. The space $\mathbf{A}^n$ has a natural family of covering maps on $\mathbf{P}^n$. If, for each i, we define $U_i = \{x \in \mathbf{P}^n : x_i \neq 0\}$, then we have a bijection from U_i to $\mathbf{A}^n$ obtained by the map

$$[X_0, \dots, X_n] \mapsto (X_0/X_i, \dots, \widehat{X_i/X_i}, \dots, X_n/X_i)$$

Geometrically, this map is obtained by intersecting a line through the origin in $\mathbf{P}^n$ with the hyperplane $X_i = 1$, which is $\mathbf{A}^n$. We think of U_i as being *finite points* of projective plane, especially when $i = 0$. A subset V of $\mathbf{P}^n$ is a projective variety if and only if $V \cap U_i$ is an affine variety for all i.

Example. *The affine lines defined by equations of the form $Y = mX + b$ in $\mathbf{A}^2$ can be embedded into projective lines in $\mathbf{P}^2$ defined by the homogenous equation $Y = mX + bZ$. This line has exactly the same finite points as the affine lines, as well as a unique point $[1 : m : 0]$ at infinity. In particular, if $Y = mX + b$ and $Y = mX + c$ are two projective lines corresponding to two parallel affine lines, then they intersect at a unique point at infinity, and this is the only reason two lines will intersect at infinity. We shall find the method of homogenization, adding a new homogenous coordinate and using it to 'balance out' polynomial equations, is the most natural way to embed affine varieties in projective varieties.*

Example. *The affine hyperbola $Y^2 = X^2 + 1$ is most naturally embedded in a projective variety by considering the homogenized form $Y^2 = X^2 + Z^2$ of the polynomial. The projective hyperbola has two points at infinity, $[1 : -1 : 0]$ and $[1 : 1 : 0]$. These correspond to the intersections of the hyperbola with the lines $Y = X$ and $Y = -X$, and at least in the real case, this corresponds to the asymptotic behaviour of the hyperbola.*

Example. *The affine twisted cubic curve in $\mathbf{A}^3$ is the intersection of the two quadratic curves $Y = X^2$ defined by $Z = X^3$. To extend the curve to $\mathbf{P}^3$, we add a new coordinate W, and homogenize, considering the two equations $YW = X^2$ and $ZW = X^3$. However, the variety corresponding to these equations does not precisely describe the asymptotics of the twisted cubic, because the zero set contains an entire line at infinity. To fix this situation, we add the addition equation $Y^2 = XZ$ to our constraints. The resultant nullset is an irreducible projective curve which is naturally described as the projective cubic.*

Example. *A* linear subvariety *of $\mathbf{P}^n$ of dimension $n - m$ is the projective subvariety if it is the locus of m forms of degree one, which can be seen as hyperplanes in $\mathbf{P}^n$. The space of linear subvarieties is invariant under projective changes of coordinates, and all linear subvarieties of the same dimension are projectively equivalent to one another.*

The only problem with the definition of projective varieties is that the connection between the ideal theory of $k[X_1, \dots, X_{n+1}]$ is not emphasized. We say an ideal $\mathfrak{a}$ in $k[X_1, \dots, X_n]$ is a *homogenous ideal* if it is closed under projection onto the homogenous components of polynomials. Thus if f is in the ideal, then $f_0, f_1, \dots$ are also elements of the ideal.

Example. *If f is a homogenous polynomial, then (f) is a homogenous ideal, since if f is degree m, then*

$$(gf)_i = \begin{cases} 0 & i < m \\ g_{i-m} f & m \leqslant i \end{cases}$$

More generally, $(f_1, \dots, f_n)$ is a homogenous ideal if each f_i is homogenous, say, of degree n_i, because then

$$\left(\sum g_i f_i \right)_k = \sum (g_i f_i)_k = \sum (g_i)_{k-n_i} f_i$$

These are all examples of homogenous ideals, because if $\mathfrak{a}$ is any homogenous ideal, it is finitely generated by some set of polynomials, and considering the homogenous parts of each polynomial ideal, we find a finite generating set of $\mathfrak{a}$ by homogenous polynomials.

Given a homogenous ideal $\mathfrak{a}$, we define

$$Z(\mathfrak{a}) = \{ x \in \mathbf{P}^n : \text{for every homogenous } f \in \mathfrak{a}, f(x) = 0 \}$$

We know that $\mathfrak{a}$ is generated by a finite set of polynomials, and by taking the homogenous parts of the generating set, we obtain a generating set of $\mathfrak{a}$ consisting only of homogenous polynomials. If $\mathfrak{a} = (f_1, \dots, f_m)$, then $Z(\mathfrak{a}) = \bigcap Z(f_i)$, because if $g = \sum g_i f_i$ is homogenous, then $g(x) = 0$ is implied by the fact that $f_i(x) = 0$ for each i. Given a projective variety V, we can consider the ideal $I(V)$ generated by homogenous polynomials f vanishing on V.

Theorem 1.20. *For any projective variety V, $I(V)$ is radical.*

Proof. Suppose that $f^n \in I(V)$. We claim by induction that $f_m \in I(V)$ for each m. For each m we have $(f^n)_m$ vanishes on V. In particular, $(f^n)_0 = (f_0)^n$ vanishes on V, so f_0 vanishes on V. Assuming that $f_0, \dots, f_m$ vanishes on V, we calculate that, for points on V, $(f^n)_{m+1} = f_{m+1}^n$, hence f_{m+1} vanishes on V, so $f_{m+1} \in I(V)$. □

Most of the other properties of the corresponding operators for affine varieties remain true.

- If $\{\mathfrak{a}_\alpha\}$ is a family of homogenous ideals, then $\bigoplus \mathfrak{a}_\alpha$ is homogenous, and $Z(\bigoplus \mathfrak{a}_\alpha) = \bigcap Z(\mathfrak{a}_\alpha)$. Thus arbitrary intersections of projective varieties are projective varieties.

- If $\mathfrak{a}$ and $\mathfrak{b}$ are homogenous ideals, then $\mathfrak{a} \cap \mathfrak{b}$ is a homogenous ideal, and $Z(\mathfrak{a}) \cup Z(\mathfrak{b}) = Z(\mathfrak{a}\mathfrak{b})$. Thus finite unions of projective varieties are projective varieties.

- $Z(0) = \mathbf{P}^n$, $Z(1) = \varnothing$, and for any $a \in k^{n+1}$, with $a_i \neq 0$,

  $$Z(a_i X_1 - a_1 X_i, \dots, a_i X_n - a_n X_i) = \{[a_1 : \dots : a_{n+1}]\}$$

  It follows that finite point sets are projective varieties.

- $I(\varnothing) = k[X_1, \dots, X_{n+1}]$, and $I(\mathbf{A}^n) = (0)$.

- For any set S of homogenous polynomials, $S \subset I(Z(S))$, and $X \subset Z(I(X))$ for any set X. This implies $Z(I(Z(S)) = Z(S)$ for any set S of homogenous polynomials, and so if V and W are projective varieties, $I(V) = I(W)$ if and only if $V = W$.

A projective variety is *irreducible* if it is not the union of two proper subvarieties. As in the study of affine varieties, V is irreducible if and only if $I(V)$ is a prime ideal. The proof essentially mirrors the affine case.

Theorem 1.21. *V is irreducible if and only if $I(V)$ is a prime ideal.*

Proof. Suppose that $fg \in I(V)$, with $f,g \notin I(V)$. We may assume that f and g are homogenous polynomials, because if i is the smallest index such that $f_i \notin I(V)$, and j the smallest index such that $g_j \notin I(V)$, then $(fg)_{i+j}$ is equal to $f_i g_j$ on V, hence $f_i g_j$ vanishes on V. Since $I(V)$ is a radical ideal, it follows that f cannot be a scalar multiple of g. Furthermore, $Z(f,I(V))$ and $Z(g,I(V))$ are both proper subsets of $I(V)$, whose union is $Z(fg,I(V)) = Z(I(V)) = V$. Conversely, suppose $V = W \cup U$, where W and U are both proper projective subsets of V. Then from the properties above we conclude that $I(V)$ is a proper subset of $I(W)$ and $I(U)$, so we may select a homogenous polynomial f vanishing on W, but not on V, and a homogenous polynomial g vanishing on U, but not on V. It follows that fg is a homogenous polynomial vanishing on V, and therefore $fg \in I(V)$, with $f,g \notin I(V)$. □

To verify that a homogenous ideal $\mathfrak{a}$ is an ideal, it suffices to show that if $fg \in \mathfrak{a}$ for two forms f and g, then either $f \in \mathfrak{a}$ or $g \in \mathfrak{a}$. This is because if f and g are general functions with $fg \in \mathfrak{a}$, with $g \notin \mathfrak{a}$, then let i be the smallest index with $g_i \notin \mathfrak{a}$. Then $(fg)_i = \sum f_j g_{i-j}$ is congruent to f_0, hence $f_0 \in \mathfrak{a}$, and by induction, if $f_l \in \mathfrak{a}$ for all $l < k$, then $(fg)_{i+k} = \sum f_j g_{i+k-j}$ is congruent to $f_k g_i$, hence $f_k \in \mathfrak{a}$. This makes it much easier to verify a projective variety is irreducible.

The main way to reduce questions about projective varieties to questions about affine varieties is to associate an affine variety in $\mathbf{A}^{n+1}$ for each projective variety in $\mathbf{P}^n$. Given a projective variety V, we associate the *cone* variety

$$C(V) = \{x \in \mathbf{A}^{n+1} : [x] \in V \text{ or } x = 0\}$$

For instance, this gives an easy proof of a version of the nullstellensatz for projective varieties, once we note that if $V \neq \varnothing$ is a projective variety defined by some homogenous ideal $\mathfrak{a}$, then $C(V)$ is the affine variety defined by $\mathfrak{a}$.

Theorem 1.22. *Over an algebraically closed field, if $\mathfrak{a}$ is a homogenous ideal, then $Z(\mathfrak{a}) = \varnothing$ if and only if there is some n such that $\mathfrak{a}$ contains all forms of degree $\geqslant n$, and if $Z(\mathfrak{a}) \neq \varnothing$, then $I(Z(\mathfrak{a}))$ is the radical ideal generated by $\mathfrak{a}$.*

Proof. If the projective variety $Z(\mathfrak{a}) = \varnothing$, then over affine space, $Z(\mathfrak{a}) = \varnothing$ or $Z(\mathfrak{a}) = \{0\}$. In the first case, we can apply the nullstellensatz to conclude that $\mathfrak{a} = k[X_1, \dots, X_{n+1}]$. In the second case, we conclude that the radical ideal generated by $\mathfrak{a}$ is equal to (X, Y), and therefore there is n such that $(X, Y)^n \subset \mathfrak{a}$, and we obtain the statement above. For nonempty projective varieties, the map $V \mapsto C(V)$ maps the projective variety generated by $\mathfrak{a}$ to the affine variety generated by $\mathfrak{a}$, and since $I(C(V)) = I(V)$ when V is nonempty, we conclude that $I(V) = I(C(V))$ is the radical ideal generated by $\mathfrak{a}$. □

Corollary 1.23. *There is a one to one correspondence between projective hyperplanes and homogeneous forms $f \in k[X_0, \dots, X_n]$ containing no repeated factors. Irreducible hypersurfaces correspond to irreducible forms.*

1.8 Projective Closure

The whole reason we introduce the theory of projective geometry to the study of affine varieties is to simplify the situation. To see the correspondence between affine varieties and projective varieties, we begin by looking at the correspondences between ideals in the respective rings generating these varieties. If $\mathfrak{a}$ is an ideal in $k[X_1, \dots, X_n]$, we let $\mathfrak{a}^*$ denote the homogenous ideal generated by f^*, for $f \in \mathfrak{a}$. Conversely, if $\mathfrak{a}$ is an ideal in $k[X_1, \dots, X_{n+1}]$,, then $\mathfrak{a}_*$ is the ideal consisting of f_*, for $f \in \mathfrak{a}$. The relationship between these processes is reflected in the fact that $(\mathfrak{a}^*)_* = \mathfrak{a}$ and $\mathfrak{a} \subset (\mathfrak{a}_*)^*$, but we can have $\mathfrak{a}$ a proper subset even if $\mathfrak{a}$ is a homogenous ideal, i.e. if $\mathfrak{a} = (X_{n+1})$, in which case $\mathfrak{a}_* = (1)$, and $(\mathfrak{a}_*)^* = k[X_1, \dots, X_{n+1}]$. Given an affine variety V in $\mathbf{A}^n$, we let V^* be the projective variety in $\mathbf{P}^n$ generated by $I(V)^*$. This is the *projective closure* of V. Conversely, if V is a projective variety in $\mathbf{P}^n$, then V_* is the affine variety in $\mathbf{A}^n$ generated by $I(V)_*$, which can also be described as $V \cap \mathbf{A}^n$.

Theorem 1.24. *The following properties hold for the correspondence.*

(a) An affine variety consists of exactly the finite points in its affine closure.

(b) If $V \subset W$, then $V^* \subset W^*$ if V and W are affine varieties, and $V_* \subset W_*$ if V and W are projective varieties.

(c) If V is an irreducible affine variety, then V^* is irreducible.

(d) If $V = \bigcup V_i$ is the decomposition of an affine variety into irreducible components, then $V^* = \bigcup V_i^*$ is the irreducible decomposition of a projective variety.

(e) V^* is the smallest projective variety containing V.

(f) If V is a nonempty affine variety forming a proper subset of $\mathbf{A}^n$, then no component of V^* lies in or contains the plane at infinity.

(g) Over an algebraically complete field, if V is a projective variety, with no component lying in or containing the plane at infinity, then V_* is a proper subset of $\mathbf{A}^n$, and $(V_*)^* = V$.

Proof. Let $[x : 1]$ be a finite point in a variety $V^* \subset \mathbf{P}^n$, for $x \in k^n$. Since $f^*(x, 1) = f(x)$, this implies that $f(x) = 0$ for all $f \in I(V)$, so $x \in V$. This proves (a). (b) follows because $I(V) \subset I(W)$ implies $I(V)^* \subset I(W)^*$ if V and W are affine variety, and $I(V)_* \subset I(W)_*$ if V and W are projective varieties. If $\mathfrak{a}$ is prime, then $\mathfrak{a}^*$ is prime, because if $fg = \sum h_i f_i^*$, where f and g are homogenous, then $f_* g_* = \sum (h_i)_* f_i \in \mathfrak{a}$, hence either f_* or g_* is in $\mathfrak{a}$, and since $(f_*)^*$ divides f, and $(g_*)^*$ divides g, we conclude that either $f \in \mathfrak{a}^*$ or $g \in \mathfrak{a}^*$. This implies that if V is irreducible, V^* is irreducible. Provided that $\mathfrak{a}$ is a prime ideal containing (X_{n+1}), $\mathfrak{a}_*$ is a prime ideal, because if $f_* g_* = h_*$, then $(fg - h) \in \mathfrak{a}$, hence $fg \in \mathfrak{a}$, so $f \in \mathfrak{a}$ or $g \in \mathfrak{a}$, and this implies $f_* \in \mathfrak{a}_*$ or $g_* \in \mathfrak{a}_*$. We can prove (e) because if W is a projective variety containing an affine variety V, then any homogenous polynomial $f \in I(W)$, so $f_* \in I(V)$, and therefore $(f_*)^* \in I(V^*)$, and $(f_*)^*$ divides f. Thus $I(W) \subset I(V^*)$, hence $V^* \subset W$. (d) follows from the previous properties because $(\bigcup V_i)^* = \bigcup V_i^*$. To prove (f), we may assume that V is irreducible. If P_∞ is the plane at infinity, and $P_\infty \subset V^*$, then $I(V)^* \subset I(V^*) \subset I(P^\infty) = (X_{n+1})$. But if f is a nonzero polynomial in $I(V)$, then X_{n+1} does not divide f^*, so we conclude $I(V) = (0)$, hence $V = \mathbf{A}^n$. To prove (g), we may assume V is irreducible. Since $V_* \subset V$, $(V_*)^* \subset V^*$, so it suffices to show that $V \subset (V_*)^*$, which is equivalent to showing that $I((V_*)^*) \subset I(V)$. If $f \in I(V_*)$, then $f^n \in I(V)_*$ for some n, and so $X_{n+1}^t (f^*)^n \in I(V)$ for some t. Since $I(V)$ is prime, and $X_{n+1} \notin I(V)$, because V is not contained in the plane at infinity, then $f^* \in I(V)$, so $I(V_*)^* \subset I(V)$. □

1.9 Products and Bi-Projective Varieties

If V and W are varieties in $\mathbf{A}^n$ and $\mathbf{A}^m$, then the Cartesian product $V \times W$ is naturally a variety in $\mathbf{A}^{n+m}$. This is because $\mathbf{A}^n \times \mathbf{A}^m$ can be naturally identified with $\mathbf{A}^{n+m}$. On the other hand, the fact that $\mathbf{P}^n \times \mathbf{P}^m$ cannot be naturally identified with $\mathbf{P}^{n+m}$ makes the analysis of Cartesian products of projective varieties more complex.

One way to get around this is to identify $\mathbf{P}^n \times \mathbf{P}^m$ with a subvariety of $\mathbf{P}^{n+m+nm}$, via the *Segre embedding*. If $[X_0 : \dots : X_n]$ give coordinates for $\mathbf{P}^n$, and $[Y_0 : \dots : Y_m]$ give coordinates for $\mathbf{P}^m$, then the map

$$\varphi([X_0 : \dots : X_n], [Y_0 : \dots : Y_m]) = [X_i Y_j]$$

gives a one-to-one map. This gives an equivalent theory to the one we give here. Another approach, which we describe here, is to define a natural family of polynomials on $\mathbf{P}^{n+m}$ with which to define *biprojective varieties*.

Write $k[X,Y]$ for $k[X_1,\dots,X_n,Y_1,\dots,Y_m]$. A polynomial $f \in k[X,Y]$ is a *biform* of *bidegree* (p,q) if it is homogeneous of degree p viewed as a polynomial in $k[Y][X]$, and homogeneous of degree q viewed as a polynomial in $k[X][Y]$. Every polynomial can be written expanded as a sum of biforms of each degree.

One can naturally consider the zeroes of a biform in $\mathbf{P}^n \times \mathbf{P}^m$. If S is a set of biforms, we consider the set $Z(S)$ of all points in $\mathbf{P}^n \times \mathbf{P}^m$. Any set of this form is called algebraic, or a *biprojective variety*. One can define *bihomogeneous ideals*, and one obtains a form of the nullstellensatz in this domain by a reduction to a study of cones in $\mathbf{A}^{n+1} \times \mathbf{A}^{m+1}$.

A similar approach to this can be carried out to give a theory of algebraic varieties in any space of the form $\mathbf{P}^{n_1} \times \dots \times \mathbf{P}^{n_k} \times \mathbf{A}^m$, the advantage of this space being that affine and projective varieties occur in the same space. Moreover, we can take products of varieties, which will also be a variety.

Theorem 1.25. *Suppose V and W are irreducible varieties. Then $V \times W$ is irreducible.*

Proof. Suppose $V \times W$ is the union of two closed subvarieties Z_1 and Z_2. For each $y \in W$, $V \times \{y\}$ is irreducible and is contained in $Z_1 \cup Z_2$, we find that either $V \times \{y\} \subset Z_1$ or $V \times \{y\} \subset Z_2$. Set Z_i' to be the set of $y \in W$ such that $V \times \{y\}$ is a subset of Z_i. Then $Z_1' \cup Z_2' = W$. If we can prove that Z_i' is a subvariety of W, this would complete the proof.

Suppose Z_i is the locus of a family of polynomials $f_1, \dots, f_k$. We know that $y \in Z_i'$ if and only if $f_i(x, y) = 0$ for all $x \in V$ and $i \in \{1, \dots, k\}$. Thus if $y_0 \notin Z_i'$, then there is $i \in \{1, \dots, k\}$ and $x_0 \in V$ such that $f_i(x_0, y_0) = 0$. But then $\{y \in W : f_i(x_0, y) \neq 0\}$ is a Zariski open set containing y_0 and no points in Z_i'. Thus $W - Z_i'$ is Zariski open, proving that Z_i' is Zariski closed. $\square$

The subvarieties of $\mathbf{P}^{n_1} \times \cdots \times \mathbf{P}^{n_k} \times \mathbf{A}^m$ are essentially the most general family of varieties that will be considered in this text. But the technicality in carrying out the definition of these subvarieties should convince even the most practical reader that an *intrinsic theory* of varieties is important for removing this technicality, i.e. a theory of varieties which isolates geometric information about the variety from the ambient space it is contained in. This is carried out in the most modern generality in the theory of schemes.

Chapter 2

Intrinsic Properties of Algebraic Varieties

Let us now study some intrinsic properties of algebraic varieties. Unless otherwise specified, we assume throughout that the field we are working over is algebraically closed.

2.1 Coordinate Rings

Often, to study the structure of some space X, we look at the space of functions on X with some particular property reflecting the structure of X. For instance, if X is a topological space, we look at the space of continuous functions. If X is the complex plane, we look at the space of holomorphic functions. Normally, these spaces of functions will turn out to have an algebraic structure, like that of a ring, or an algebra over a field, and determining this algebraic structure up to isomorphism often classifies the geometric structure of X.

Let us begin by carrying out this approach for affine varieties. Given an affine variety V, the natural functions on V 'should' be the polynomial functions. We define the *coordinate ring* of a variety V, also known as the ring of *regular functions* on V, denoted $k[V]$, to be the ring $k[X_1,\dots,X_n]/I(V)$. We think of this as the family of functions obtained by restricting polynomial functions in ambient space to the variety V. When k is infinite, this correspondence can be made exact, since the ring $k[X_1,\dots,X_n]$ of polynomial *expressions* can be identified with the family of all polynomial *func-*

tions. Regardless, it is still possible to evaluate elements of $k[V]$ on points of V, though when k is finite the evaluations may not uniquely describe an element of $k[V]$.

A *subvariety* of a variety V is an algebraic set which occurs as a subset of V. The nullstellensatz tells us that in an algebraically complete field, the subvarieties of V are in one to one correspondence with the radical ideals containing $I(V)$. Applying the fourth isomorphism theorem, since the image of an ideal containing $I(V)$ is radical in $k[V]$ if and only if it is radical in $k[X_1,\dots,X_n]$, we find that the subvarieties of V are in one to one correspondence with the radical ideals in $k[V]$. The points in V are also in one to one correspondence with the maximal ideals containing $I(V)$. This is the first instance of the fact that we can view V as an 'algebraic space' independent of $\mathbf{A}^n$, in which 'being a subvariety' corresponds to 'being the locus of a radical ideal'. As another example, we note that if $I_Z(W)$ is the ideal of functions in $k[V]$ vanishing on W, then $k[W]$ is isomorphic to $k[V]/I_Z(W)$, so that quotienting by functions vanishing on a subvariety is a natural way to form a coordinate ring on a subvariety of an arbitrary variety, not just in an affine space. This is the first step in forming a 'coordinate independent' way of defining varieties, which gives rise to modern algebraic geometry, wherein varieties need not lie in an ambient space.

Proposition 2.1. *Let k be infinite. A variety V contains only finitely many points if and only if $k[V]$ is a finite dimensional vector space over k, and then*

$$\dim_k(k[V]) = \#(V).$$

Proof. Interpolation techniques shows that for any finite set of points $p_1,\dots,p_n \in V$, we can find polynomials $f_1,\dots,f_n \in k[V]$ such that $f_i(p_j) = \delta_{ij}$. If $V = \{p_1,\dots,p_n\}$, then taking linear combinations of the functions $f_1,\dots,f_n$ shows that $k[V]$ contains all functions from V to k, and so $k[V]$ is isomorphic to k^n. On the other hand, it is easy to see that $f_1,\dots,f_n$ are linearly independant, and so $\dim_k(k[V]) \geqslant n$; if $\#(V) = \infty$, then one can take n to be arbitrarily large, from which it follows that $k[V]$ is infinite dimensional. □

One can extend this result, using a technique which will become much more useful later in our study of planar curves.

Lemma 2.2. *Let k be algebraically closed. If $\mathfrak{a}$ is any ideal of $k[X_1,\dots,X_n]$ with the property that $Z(\mathfrak{a})$ is finite, then $k[X_1,\dots,X_n]/\mathfrak{a}$ is finite dimensional over k, and*

$$\#(Z(\mathfrak{a})) \leqslant \dim_k(k[X_1,\dots,X_n]/\mathfrak{a}).$$

Proof. This theorem follows by this nullstellensatz if $\mathfrak{a}$ is a radical ideal, for then $I(Z(\mathfrak{a})) = \mathfrak{a}$, and so $k[V]$ is isomorphic to $k[X_1,\dots,X_n]/\mathfrak{a}$. In general, if $\mathfrak{a}$ is not radical, let $\mathfrak{b}$ denote $I(Z(\mathfrak{a})) = \text{Rad}(\mathfrak{a})$ and let $R = k[X_1,\dots,X_n]/\mathfrak{b}$, and $S = k[X_1,\dots,X_n]/\mathfrak{a}$. The natural surjective morphism from S to R shows that $\dim_k(S) \geqslant \dim_k(R) = \#(V)$. Thus the only remaining content of the proof is to show taht S is finite dimensional over k.

Since $\mathfrak{b}$ is finitely generated, and each element of $\mathfrak{b}$ has a power which is in $\mathfrak{a}$, we conclude that there is an integer n such that $\mathfrak{b}^n \subset \mathfrak{a}$. Let $S_i = k[X_1,\dots,X_n]/\mathfrak{b}^i$. Then $R = S_1$ and $S = S_n$. Our goal is to show by induction that S_i is finite dimensional. For each i, we have a natural morphism from S_{i+1} to S_i, which has kernel $\mathfrak{b}^i / \mathfrak{b}^{i+1}$. The isomorphism theorem for vector spaces implies we only have to prove that $\dim_k(\mathfrak{b}^i / \mathfrak{b}^{i+1}) < \infty$. But this follows because if $\mathfrak{b}^i = (f_1,\dots,f_m)$, then $f_1,\dots,f_m$ span $\mathfrak{b}^i / \mathfrak{b}^{i+1}$ over k because $f_i f_j \in \mathfrak{b}^{i+1}$ for any i,j. Thus the induction is complete. □

Example. *Consider the locus V of the polynomials $X^2 - Y^2$ and $X^2 + Y^2$ over an algebraically closed field k not of characteristic 2. Since $(X^2 - Y^2, X^2 + Y^2) = (X^2, Y^2)$, the radical ideal of these polynomials is (X, Y), and so $k[V] = k[X,Y]/(X,Y) \cong k$ is a one dimensional vector space over k. This makes sense because $V = \{0\}$ in this case. On the other hand, if k is of characteristic two, then $X^2 - Y^2$ and $X^2 + Y^2$ are the same polynomial, and in fact $X^2 + Y^2 = (X+Y)^2$. In this case the radical ideal of $(X^2 - Y^2, X^2 + Y^2)$ is $(X + Y)$. The ring $k[V] = k[X,Y]/(X+Y) \cong k[X]$ is infinite dimensional, and V consists of infinitely many points.*

Since the basic notion of algebraic geometry is the set of polynomials, the natural structure preserving maps between varieties should be those maps $f : V \to W$ should be those maps induced by polynomial maps. These are the *regular maps*, also known as *polynomial maps*. To be specific, a map $f : \mathbf{A}^n \to \mathbf{A}^m$ is regular if each coordinate map $f_1,\dots,f_m$ is induced by a polynomial function. The regular maps between two varieties V and W are then exactly those induced by a restriction of a polynomial map between $\mathbf{A}^n$ and $\mathbf{A}^m$. If $f : X \to Y$ is a map between two sets, then it induces a 'pullback' map $f^* : k^Y \to k^X$ obtained by composition: $f^* g = g \circ f$. This map has many useful properties for our studies:

- If $g : Y \to Z$ is another polynomial map, then $(g \circ f)^* = f^* \circ g^*$.
- If $f(X_0)$ is a subset of Y_0, then f^* descends to a map from k^{X_0} to k^{Y_0}, and this function respects the restriction homomorphisms.

- If $f : V \to W$ is a polynomial function, then f^* maps functions in $k[W]$ to functions in $k[V]$. A polynomial map f maps elements of V into elements of W if and only if f^* maps $I(W)$ into $I(V)$.
- If $f : V \to W$ is a surjective map, then $f^* : k[W] \to k[V]$ is injective.

In fact, the algebra structure of $k[V]$ classifies V as a variety, up to an application of a polynomial map.

Proposition 2.3. *There is a one to one correspondence between regular maps between V and W and algebra homomorphisms from $k[W]$ to $k[V]$.*

Proof. Given a homomorphism $T : k[W] \to k[V]$, we can define a polynomial map $f : V \to W$ by letting $f = (TX_1, \dots, TX_n)$, which is well defined over V. We claim that $Tg = g \circ f$ for all polynomials $g \in k[W]$. It is clear that the set of polynomials satisfying this equation include 1 and $X_1, \dots, X_n$, and if $Tg_0 = g_0 \circ f$ and $Tg_1 = g_1 \circ f$, then $Tg_0 g_1 = (g_0 \circ f)(g_1 \circ f) = (g_0 g_1 \circ f)$. Since $1, X_1, \dots, X_n$ generate $k[V]$ as an algebra, we conclude that the equation is satisfied by all g. If f is any polynomial map between two varieties, then $f^*(g) = g \circ f$, and the construction above reconstructs the function f, so we know there is a one to one correspondence. □

A polynomial map is a *regular isomorphism* if it is bijection, and its inverse is also a polynomial map. Thus the intrinsic study of curves is characterized up to a regular isomorphism; the theory attempts to study the properties of varieties which are invariant under polynomial isomorphisms. We have argued that $k[V]$ is an isomorphism invariant of the variety V: two varieties V and W are isomorphic if and only if $k[V]$ and $k[W]$ are isomorphic. This means that the coordinate rings have sufficient expressive power to consider the structure of V. This is the same as how the coordinate ring $C(X)$ classifies the topological structure of a compact Hausdorff space X up to isomorhpism. Over an algebraically closed field, every *reduced* finitely generated k algebra R is the coordinate ring of some variety, the map $V \mapsto k[V]$ is an contravariant equivalence between the category of affine varieties and the category of reduced finitely generated k algebras, when k is algebraically closed. This is the reason the nullstellensatz enables us to view geometric problems algebraically, and conversely, view algebraic problems geometrically.

Proposition 2.4. *If $f : V \to W$ is a surjective polynomial map between two varieties V and W, and V is irreducible, then W is irreducible.*

Proof. Suppose that W is reducible, so that we may write $W = W_1 \cup W_2$. Then we have a decomposition of V as $V_1 = f^{-1}(W_1)$ and $V_2 = f^{-1}(W_2)$, and $V_1, V_2 \neq V$ because otherwise this would imply that either $W_1 = W$ or $W_2 = W$. Alternatively, we note that $k[V]$ is an integral domain, and $f^* : k[W] \to k[V]$ is injective, so that $k[W]$ can be identified as a subring of $k[V]$, and is therefore an integral domain. □

Example. *We have seen that $\{(t,t^2,t^3) : t \in k\}$ is an affine variety, because it is the locus of the polynomials $X^2 = Y$ and $X^3 = Z$. Another way to see this is to note that the variety is the image of the polynomial map from $\mathbf{A}^1$ to $\mathbf{A}^3$ defined by $t \mapsto (t,t^2,t^3)$. It is irreducible because it is the image of $\mathbf{A}^1$, which is an irreducible variety. What's more, the variety is isomorphic to $\mathbf{A}^1$, because the embedding has a polynomial inverse $(x,y,z) \mapsto x$.*

Example. *The variety V defined by the equation $XY - 1$ is a curve not isomorphic to $\mathbf{A}^1$, because $k[V]$ is not isomorphic to $k[X]$. To see this, we note that an isomorphism $f : k[X] \to k[V]$ would induces a map on the group of units in each ring. But the units in $k[X]$ are exactly the elements of $k - \{0\}$, so the isomorphism, since it fixed k, allows us to conclude that $k[V]$ cannot have any other units than $k - \{0\}$. But X and Y are both units in $k[V]$.*

Example. *The locus V of polynomials of the polynomials $XZ = Y^2$, $YZ = X^3$, and $Z^2 = X^2Y$ forms an irreducible variety over $\mathbf{C}$. Note that $Y^3 - X^4$ is in the ideal $(XZ - Y^2, YZ - X^3, Z^2 - X^2Y)$, and if $x, y \in k$ are picked such that $x^4 = y^3$, there is a unique $z \in k$ with $z = y^2/x = x^3/y$, unless $x = y = 0$. In this case, we conclude that $z = 0$ because $z^2 = x^2y$. Otherwise $z^2 = x^2y$ follows automatically because $z^2 = (y^2/x)(x^3/y)$. The polynomial map $t \mapsto (t^3,t^4,t^5)$ is therefore a surjective map from $\mathbf{A}^1$ onto V. For any $y \neq 0$, there are exactly four values of t such that $t^4 = y$, and if t is any solution then it, $-it$, and $-t$ form the other three solutions to the equation. Now if $x^4 = y^3$, then $x^4 = t^{12}$, and*

$$x^4 - t^{12} = (x - t^3)(x + t^3)(x - it^3)(x + it^3)$$

This implies that either $x = t^3$, $x = -t^3$, $x = it^3$, or $x = -it^3$. But by replacing t with any of the other roots of the equation $t^4 = y$, we find that there is a unique value of t such that $t^4 = y$ and $t^3 = x$. We conclude that the map $t \mapsto (t^3,t^4,t^5)$ is actually a bijection. The same argument essentially shows that V is irreducible in any algebraically closed field: one must just take a bit of extra care when we are doing computations over a field of characteristic two.

Example. *For any $f \in k[V]$, where V is some variety in $\mathbf{A}^n$, define the* graph *$G(f)$ of f to be the set of tuples $(a_1,\dots,a_{n+1}) \in \mathbf{A}^{n+1}$, where $(a_1,\dots,a_n) \in V$ and $a_{n+1} = f(a_1,\dots,a_n)$. $G(f)$ is isomorphic to V under the projection map $(a_1,\dots,a_{n+1}) \mapsto (a_1,\dots,a_n)$, because for each $a_1,\dots,a_n$ the number a_{n+1} is uniquely determined.*

Example. *A bijective polynomial map need not be an isomorphism. Consider the polynomial map from $\mathbf{A}^1$ to $Z(Y^2 - X^3)$ defined by letting $f(t) = (t^3, t^2)$. Then f is a bijection, but f^* is not surjective, for it maps X onto t^3, and Y onto t^2, so $f^*(X^2) = f^*(Y^3)$, and the image of the map is therefore $k[t^3, t^2]$, which is a proper subset of $k[t]$.*

As should be expected by a geometer, the isomorphisms of $\mathbf{A}^n$ contain the family of affine translations $x \mapsto Mx + b$, where $b \in \mathbf{A}^n$ and $GL_n(k)$. This is exactly the reason why we work in a space denoted by $\mathbf{A}^n$ rather than k^n, because the choice of isomorphisms mean that the particular choice of affine coordinates used to define varieties is of no real consequence to the geometry of varieties - we don't care where the origin is.

Example. *The affine subplanes of $\mathbf{A}^n$ are varieties known as* linear subvarieties. *Any variety of the form $Z(f_1,\dots,f_m)$, where each f_i is of degree one, is a linear subvariety, in which case the subplane has dimension $n - m$. These subplanes are all isomorphic to $\mathbf{A}^{n-m}$. This can easily be seen by a projection, but can also be seen because a linear subvariety of dimension n has coordinate ring isomorphic to $k[X_1,\dots,X_n]$.*

2.2 The Function Field of a Variety

The ring $k[V]$ is an isomorphism invariant of V, but it is often difficult to work with. However, when V is an irreducible variety, then $k[V]$ is an integral domain, because it is the quotient of $k[X_1,\dots,X_n]$ by a prime ideal. This means we can form the field of fractions, which we denote $k(V)$. The elements of $k(V)$ correspond to functions on V defined except at certain singularity sets, known as the set of *poles* of the function. Given $f \in k(V)$, we say f is *defined*, or *regular* at $p \in V$ if we may write $f = g/h$, where $h(p) \neq 0$. Then the quantity $f(p) = g(p)/h(p)$ is well defined, because if $g_0/h_0 = g_1/h_1$, then $h_1 g_0 = h_0 g_1$, and so $h_1(p)g_0(p) = h_0(p)g_1(p)$, and so $g_0(p)/g_1(p) = h_0(p)/h_1(p)$. Though $k(V)$ is a weaker invariant than $k[V]$, an isomorphism between $k(V)$ and $k(W)$ still provides strong relations between two varieties V and W. To see one such relation, recall that a set $X \subset V$ is *Zariski dense* in V if no proper subvariety of V contains X. This, in particular, implies any polynomial $f \in k[V]$ which vanishes on X also vanishes on V.

Theorem 2.5. *The k-algebra homomorphisms $T : k(W) \to k(V)$ are in one to one correspondence with partially-defined maps $f : V \to W$ expressed as rational functions, whose image is Zariski dense in W.*

Proof. Suppose that $T : k(W) \to k(V)$ is a homomorphism. If $V \subset \mathbf{A}^n$ is definable in the coordinates $(X_1, \dots, X_n)$, and $W \subset \mathbf{A}^m$ is definable in the coordinate $(Y_1, \dots, Y_m)$, then let $T(Y_i) = f_i$. We claim that $f = (f_1, \dots, f_m)$ gives a surjective map from V to W, where defined. Fix $x \in V$ such that $f_1, \dots, f_m$ are all defined at x; we shall show that $y = f(x)$ is an element of W. If y was not an element of W, then there would be $g \in I(W)$ with $g(y) = 1$. But then $T(g) = 0$, which is impossible since $T(g)(x) = g(f(x)) = g(y) = 1$. On the other hand, if $y \in W$ was not in the Zariski closure of $f(V)$, then we could define a function $g \in k[W]$ vanishing on $f(V)$ but with $g(y) = 1$. Then $Tg = 0$, which implies $g = 0$, which is impossible.

Conversely, if $f : V \to W$ is definable by rational functions in the coordinates whose image is Zariski dense in W, we can define $T : k[W] \to k(V)$ by letting $T(Y_i) = f_i$, because if $\sum a_\alpha Y^\alpha \in I(W)$, then $\sum a_\alpha f^\alpha \in I(V)$. If $T(\sum a_\alpha Y^\alpha) = \sum a_\alpha f^\alpha = 0 \in k(V)$, then $\sum a_\alpha Y^\alpha$ must vanish on $f(V)$, and this implies that $\sum a_\alpha Y^\alpha = 0$. This implies the map descends to a map from $k(V)$ to $k(W)$. □

The association of f with T is a contravariant functor from the category of irreducible algebraic varieties to the category of fields, so in particular, if we find T^{-1} for some isomorphism T, then the rational functions $(f_1, \dots, f_m)$ corresponding to T and the rational functions $(g_1, \dots, g_n)$ corresponding to T^{-1} are inverse functions of one another, viewed as maps from V to W. A map f specifiable by rational functions with an inverse of the form g is known as a *birational* map.

Example. *Let $f \in K[X,Y]$ be an arbitrary irreducible quadratic, defining a planar curve C. We claim that $k(V)$ is isomorphic to $k(t)$, which implies there is a birational map $g : k \to V$ parameterizing V. We can already guess such a birational map, since if $(x_0, y_0) \in V$, then there is at most one point in V on* *any line through (x_0, y_0), since V is quadratic, implying we can parameterize the line by slope. If we define*

$$T = \frac{Y - y_0}{X - x_0} \in k(V)$$

In $k(V)$, we know $f(X, y_0 + T(X - x_0)) = f(X, Y) = 0$. This is a polynomial relation in $k(T)[X]$, and we know setting $X = x_0$ causes the polynomial to vanish, so $f(X, y_0 + T(X - x_0)) = (X - x_0)(a(T)X - b(T))$, for some $a, b \in k(T)$. Since $X - x_0 \neq 0$ in $k(V)$, $X = b(T)/a(T)$ on V. But this means $X \in k(T)$, and therefore all $Y = T(X - x_0) + y_0 \in k(T)$, so $k(V) = k(T)$. Finally, we know $k(T)$ is isomorphic to the field of rational functions in a single variable. An interesting application of this is to compute the indefinite integral of

$$\int \varphi\left(x, \sqrt{ax^2 + bx + c}\right)$$

where φ is a rational function. Since the map $x \mapsto (x, \sqrt{ax^2 + bx + c})$ maps onto the curve V defined by $y^2 = ax^2 + bx + c$. Thus we have

$$\varphi\left(x, \sqrt{ax^2 + bx + c}\right) = \psi\left(x, \sqrt{ax^2 + bx + c}\right)$$

for any rational function ψ for which ψ and φ are equal on V, so we can interpret $\varphi \in \mathbf{C}(V)$. In particular, since there is a rational function $t(x, y)$ such that $\mathbf{C}(V) \cong \mathbf{C}(t)$, we know that there is a rational function ψ such that

$$\varphi\left(x, \sqrt{ax^2 + bx + c}\right) = \psi\left(t\left(x, \sqrt{ax^2 + bx + c}\right)\right)$$

As well as the existence of a rational function η such that $\eta(t(x, \sqrt{ax^2 + bx + c})) = x$. Thus making the change of variables to t, we find

$$\begin{aligned} \int \varphi\left(x, \sqrt{ax^2 + bx + c}\right)\, dx &= \int \psi\left(t\left(x, \sqrt{ax^2 + bx + c}\right)\right)\, dx \\ &= \int \psi(t)\eta'(t)\, dt \end{aligned}$$

The right hand side is a rational function of a single variable, which we know how to integrate. This is the technique of Euler substitution. For the same reason, the fact that $\mathbf{C}(C) \cong \mathbf{C}(t)$ when C is the circle explains why every *rational function in sines and cosines has an indefinite integral. On the other hand, the function field of the curve $y^2 = x^3 + 1$ is no longer isomorphic to the rational functions in a single variable, and this is very related to the fact that the indefinite integral of the function*

$$f(x) = \frac{1}{\sqrt{x^3 + 1}}$$

is not expressible in terms of the elementary functions.

Proposition 2.6. *The pole set of any $f \in k(V)$ is a subvariety of V. If k is algebraically closed, then the only functions in $k(V)$ without poles are regular.*

Proof. For any $f \in k(V)$, let $\mathfrak{a}$ be the ideal of all $h \in k[X_1, \dots, X_n]$ such that $hf \in k[V]$. If $f = g/h$, then $h \in \mathfrak{a}$. Conversely, if $hf = g$, and h is nonzero, then $f = g/h$, so $\mathfrak{a} - \{0\}$ is exactly the set of possible denominators for fractional expressions of f, and so $Z(\mathfrak{a})$ gives the set of poles of f. If f has no poles, then $Z(\mathfrak{a}) = \varnothing$, so applying the nullstellensatz, we conclude that $\text{Rad}(\mathfrak{a}) = k[X_1, \dots, X_n]$, so that $1 = 1^n \in \mathfrak{a}$, so that $f \in k[V]$, because we can express f as a fraction with denominator 1. □

An element of $k(V)$ can be considered a function on the complement of its pole set. If V is an infinite variety, and two functions $f, g \in k(V)$ share the same pole set, and agree as functions on the complement of their pole set, then $f = g$. This follows because if we write $f = f_0/f_1$, and $g = g_0/g_1$, then $f_0(x)g_1(x) = g_0(x)f_1(x)$ holds for all $x \in V$, hence $f_0 g_1 = g_0 f_1$ in $k[V]$, and this implies $f = g$ in $k(V)$. This is good news, because it means we can analyze elements of $k(V)$ as functions on a subset of V. The only bad side of this is that the elements of $k(V)$ may not be defined on subvarieties of V, but instead the difference of two varieties.

Example. *Consider the solution set V to the polynomial $XW - YZ$ in $\mathbf{A}^4$. Then for each x and y, the set of w and z satisfying $xw - yz$ forms a line through the origin, except when $x = y = 0$. This implies that $Y, W \neq 0$ in $k[V]$, and so we may consider the rational function $f = X/Y = W/Z \in k(V)$, which is defined at all points except at the points where $y = 0$ and $w = 0$. This is because the ideal of all possible denominators is equal to (Y, W), because if there is a polynomial f such that $(X/Y)f \in k[V]$, then we can write $fX = Yg + [XW - YZ]h$ for some polynomials g and h. Rearranging, we find $X[f - Wh] = Y[g - Zh]$, so $g - Zh$ is divisible by X, and we can write $g = Zh + Xg_1$*

for some polynomial g_1. The equation then reads $fX = X[Yg_1 + Wh]$, hence $f = Yg_1 + Wh \in (Y, W)$.

Example. *Let V be the locus of $Y^2 = X^2(X+1)$. Let us see where the function Y/X is defined. The ideal of denominators of the function include X and Y, because $Y(Y/X) = Y^2/X = X^2(X+1)/X = X(X+1)$, so $Y/X = X(X+1)/Y$. No element of k can be a denominator, for if we have an equality of polynomials of the form $tY = Xg(X,Y) + [Y^2 - X^2(X+1)]h(X,Y)$ in $k[X_1, \dots, X_n]$, then $Y[t - Yh(X,Y)] = X[g(X,Y) - X(X+1)h(X,Y)]$, hence $g(X,Y) - X(X+1)h(X,Y)$ divides Y, and we can write $g(X,Y) = X(X+1)h(X,Y) + Yg_1(X,Y)$, hence $t = Xg_1(X,Y) + Yh(X,Y)$, which is impossible unless $t = 0$. Thus the pole set of Y/X is exactly $X = Y = 0$. You might imagine that Y^2/X^2 has a smaller pole set than Y/X, but since $Y^2 = X^2(X+1)$ we can rewrite the function as $X^2(X+1)/X^2 = X+1$, so the function is defined everywhere!*

2.3 Localization at a Point

Given an irreducible variety V, we define the local ring $\mathcal{O}_p(V)$ at p to be the subring of rational functions on V which are defined at p. We shall find the ring represents the 'local structure' of the variety V around p. More generally, if V is an arbitrary (non irreducible) variety, then we can still define $\mathcal{O}_p(V)$ as the localization of $k[V]$ by the set $S_p(V)$ of $f \in k[V]$ such that $f(p) \neq 0$. However, unlike in $k(V)$, we cannot in general represent elements of $\mathcal{O}_p(V)$ as functions on V in a natural way – the elements of $\mathcal{O}_p(V)$ are only well defined at p. This makes sense from the topological sense of locality – the family of continuous functions locally equal around a point p do not necessarily share any values in common except their value at p. We shall find that the ring $\mathcal{O}_p(V)$ models the 'local' properties of the variety around the point p. In this case, it makes sense that we cannot necessarily define the values of functions in $\mathcal{O}_p(V)$ at points $q \neq p$, because q is not 'local' enough to p, whereas we can define the function at p because the value at p is a 'local' property. When V is irreducible, $k[V]$ is an integral domain, so the global definition of functions in $\mathcal{O}_p(V)$ is a 'local' property, which tells us that irreducible varieties will have more powerful results when moving from local properties to global properties. This is certainly true in the localization of other rings of functions around points, like in complex analysis, when we study $\mathcal{O}_p(D)$, which is the localization of the space of holomorphic functions on some connected set D by functions not vanishing at p; the theory of power series of such functions shows that the structure of $\mathcal{O}_p(D)$ is determined by successive derivatives of the function at the point p.

Example. *Consider the reducible curve V defined by the equation $XY = 0$. Then $\mathbf{C}[V] \cong \mathbf{C}[X,Y]/(XY)$, because (XY) is a radical ideal. Let us consider the local ring $\mathcal{O}_0(V)$. If f is a zero divisor and $f(0,0) \neq 0$, then there is g such that $fg = 0$ on V. In particular, this means fg vanishes on the X axis, and since the variety $Z(X)$ is irreducible, we conclude that g vanishes on the X axis. Similarily, we conclude that since $Z(Y)$ is irreducible, g vanishes on the Y axis. Thus g vanishes on V, so f cannot be a zero divisor. Elements of $\mathcal{O}_0(V)$ can be written in the form*

$$\frac{a + Xf_1(X) + Yf_2(Y)}{b + Xg_1(X) + Yg_2(Y)}$$

where $b \neq 0$. Since no zero divisors are inverted in the localization, two polynomials f_1/g_1 and f_2/g_2 are identified in $\mathcal{O}_0(V)$ if and only if $f_1g_2 = f_1g_2$ on the whole of V. On the other hand, if $p = (a,0)$, where $a \neq 0$, then $\mathcal{O}_p(V)$ has a slightly stranger structure. If $f(p) \neq 0$, then f can still be a zero divisor if it vanishes on the Y axis, because then any $g(X,Y) = Yg_1(Y)$, where $g_1 \neq 0$, satisfies $fg = 0$ on V. Thus in the ring $\mathcal{O}_p(V)$, two rational functions f_1/g_1 and f_2/g_2 are identified if $f_1g_2 = f_2g_1$ on the X axis, which, in a sense, means the rational functions agree with one another on the X axis. Similarly, on $p = (0,b)$ functions are identified in $\mathcal{O}_p(XY)$ if they agree on the Y axis. This makes sense, because the points on one axis are not 'local' to points on the other axis.

Example. *The most extreme example of locality occurs if V contains finitely many points $p_1, \dots, p_n$. It then follows that $k[V]$ is isomorphic to k^V, and in particular two rational functions in $\mathcal{O}_p(V)$ are identified if they have the same value at p. This implies that $k[V]$ is isomorphic to the direct product of $\mathcal{O}_p(V)$, as p ranges over all points in V.*

There is a more general result along this line, that will be more useful in further studies of local rings. Note that it is essentially a generalization of the last example.

Theorem 2.7. *If k is algebraically closed, and $\mathfrak{a}$ is an ideal of $k[X_1,\dots,X_n]$ such that $Z(\mathfrak{a}) = \{p_1,\dots,p_n\}$, then*

$$k[X_1,\dots,X_n]/\mathfrak{a} \cong \prod_{i=1}^n \mathcal{O}_{p_i}(\mathbf{A}^n)/S_{p_i}^{-1}\mathfrak{a}.$$

Remark. This theorem follows from general commutative algebra, i.e. by showing that if $k[X_1,\dots,X_n]/\mathfrak{a}$ is an Artinian ring. To see this, we note that any prime ideal in this ring corresponds to a prime ideal in $k[X_1,\dots,X_n]$ containing $\mathfrak{a}$. Applying the nullstellensatz, these prime ideals correspond precisely to irreducible subvarieties of $Z(\mathfrak{a})$, i.e. to each of the finite points in this set. But this means that any prime ideal in $k[X_1,\dots,X_n]/\mathfrak{a}$ is maximal. Since this ring is also Noetherian, we conclude the ring is Artinian. But the direct sum decomposition above holds for general Artinian rings (by localizing over maximal ideals in the ring).

Proof. First, note that since localization commutes with quotienting, the ring $\mathcal{O}_{p_i}(\mathbf{A}^n)/S_{p_i}^{-1}\mathfrak{a}$ is isomorphic to the ring obtained by localization of the form $\mathcal{O}_i = (S_{p_i}/\mathfrak{a})^{-1}(k[X]/\mathfrak{a})$. We will let T_i denote the canonical embedding of $k[X]/\mathfrak{a}$ in $\mathcal{O}_i$. If $T_i(f) = 0$, this means that there is a function g with $g(p_i) \neq 0$ such that $gf \in \mathfrak{a}$. We will prove that there is a set of functions e_i such that if $g(p_i) \neq 0$, then there is t such that $tg \equiv e_i$ modulo $\mathfrak{a}$, $\sum e_i \equiv 1$ modulo $\mathfrak{a}$, $T_i(e_i) = 1$, and $e_ie_j \in \mathfrak{a}$. This implies that if $T_if = 0$ for all i, then $e_if \in \mathfrak{a}$, and therefore that $f \equiv (\sum e_i)f = \sum e_if \in \mathfrak{a}$, which implies that $f \in \mathfrak{a}$, so that the product map $(T_1,\dots,T_n)$ is injective. To prove surjectivity, we consider an arbitrary point $(a_1/s_1,\dots,a_n/s_n) \in \mathcal{O}_i$. Since $s_i(p_i) \neq 0$, we can write $t_is_i = e_i$, which implies that $a_i/s_i = a_it_i/e_i = a_it_i$. Since $e_ie_j \in \mathfrak{a}$, $T_i(e_j) = T_i(e_ie_j) = 0$, so the image of $\sum a_it_ie_i$ by the map T is $(a_1/s_1,\dots,a_n/s_n)$. This proves that T is an isomorphism.

To finish the proof, we construct the values e_i. First, find $f_i \in k[V]$ with $f_i(p_j) = \delta_{ij}$. If $\mathfrak{b}$ is the radical ideal obtained from $\mathfrak{a}$, then the nullstellensatz implies that $\mathfrak{b} = Z(I(\mathfrak{a}))$, and since $\mathfrak{b}$ is finitely generated we can choose n such that $\mathfrak{b}^n \subset \mathfrak{a}$. Let $\mathfrak{m}_1,\dots,\mathfrak{m}_n$ be the maximal ideals of $k[X]$ corresponding to the points $p_1,\dots,p_n$. These are exactly the maximal ideals containing $\mathfrak{a}$, and their intersection is $\mathfrak{b}$. If we define $e_i = 1-(1-f_i^n)^n$, then f_i^n divides e_i, and therefore $e_i \in \mathfrak{m}_j^n$ for each $j \neq i$. We now verify the required properties of the e_i.

- $1 - \sum e_i = (1 - e_k) - \sum_{i \neq k} e_i \in \mathfrak{m}_k^n$, because each $e_i \in \mathfrak{m}_k^n$ for $i \neq k$, and $1 - e_i = (1 - f_i)^{n_i} \in \mathfrak{m}_k^n$. This implies that $1 - \sum e_i \in \bigcap \mathfrak{m}_k^n$, which is equal to $(\bigcap \mathfrak{m}_k)^n$ because the $\mathfrak{m}_k$ are comaximal, and this is a subset of $\mathfrak{b}^n$, which is a subset of $\mathfrak{a}$, so $1 - \sum e_i \in \mathfrak{a}$.

- $e_i - e_i^2 = e_i(1 - f_i^n)^n$, which is the product of an element of $\mathfrak{m}_i^n$ with an element of $\bigcap_{j \neq i} \mathfrak{m}_j^n$, and since these two are comaximal, this is equal to $\bigcap \mathfrak{m}_j^n = \mathfrak{b}^n \subset \mathfrak{a}$.

- For similar reasons, $e_i e_j \in (\bigcap_{k \neq i} \mathfrak{m}_k^n)\, \mathfrak{m}_i^n \subset \mathfrak{a}$.

- If $g(p_i) \neq 0$ (we may assume that $g(p_i) = 1$), then $1 - g \in \mathfrak{m}_i$, so $(1 - g)^n e_i \in \mathfrak{a}$. But then

$$\begin{aligned} e_i g(1{+}(1-g) &+ \cdots + (1-g)^{n-1}) \\ &= e_i(1 - (1-g))(1 + (1-g) + \cdots + (1-g)^{n-1}) \\ &= e_i - e_i(1-g)^n \end{aligned}$$

  so if $t = e_i g(1 + (1-g) + \cdots + (1-g)^{n-1})$, then $tg - e_i = e_i(1-g)^n \in \mathfrak{a}$.

- The fact that $T_i(e_i) = 1$ follows because e_i is a unit in $\mathcal{O}_i$, and $T_i(e_i e_j) = T_i(0) = 0$, hence $T_i(e_j) = 0$. Now $\sum e_i$ is congruent to 1 modulo $\mathfrak{a}$, so $T_i(e_i) = T_i(\sum e_i) = T(1) = 1$.

This completes the proof. □

Corollary 2.8. *For any ideal $\mathfrak{a}$ with $Z(\mathfrak{a}) = \{p_1, \ldots, p_n\}$,*

$$dim_k(k[X]/\mathfrak{a}) = \sum_{i=1}^{n} dim_k(\mathcal{O}_{p_i}(\mathbf{A}^n))/S_{p_i}^{-1}\mathfrak{a}$$

The invertible elements of $\mathcal{O}_p(V)$ are exactly those functions f with $f(p) \neq 0$. The complement of this set is the maximal ideal at p, denoted $\mathfrak{m}_p(V)$, which consists of all functions which vanish at p. Because the set of non-invertible elements in $\mathcal{O}_p(V)$ forms an ideal, the space has a unique maximal ideal. Such a ring is called *local*. This means that $\mathcal{O}_p(V)/\mathfrak{m}_p(V)$ is a field, and an isomorphism between this set and the field k is induced by the evaluation map $\mathrm{ev}_p : \mathcal{O}_p(V) \to k$. As a subring of a field, it is an integral domain. What's more, as a localization of a Noetherian ring, it is

Noetherian as well. The following propositions begin to hint at how the ring theoretic structure of $\mathcal{O}_p(V)$ tells us about the local properties of the variety V around p.

Proposition 2.9. *The irreducible varieties passing through p are in one to one correspondence with the proper radical ideals of* $\mathcal{O}_p(V)$.

Proof. By general properties of localization, there is a one to one correspondence between proper prime ideals of $\mathcal{O}_p(V)$ and ideals in $k[V]$ disjoint from the multiplicative set defining the localization, in this case, ideals consisting of functions vanishing at p. The correspondence is obtained from the projection map $k[V] \to \mathcal{O}_p(V)$. □

Proposition 2.10. *Every polynomial map* $f : V \to W$ *with* $f(p) = q$ *induces a unique homomorphism* $f^* : \mathcal{O}_q(W) \to \mathcal{O}_p(V)$, *with* $\mathfrak{m}_q(W)$ *being mapped into* $\mathfrak{m}_p(V)$.

Proof. Each polynomial map f induces $f^* : k[W] \to k[V]$, which we may view as a map from $k[W]$ to $\mathcal{O}_p(V)$. If $g \in k[W]$ has $g(q) \neq 0$, then $(f^*g)(p) = (g \circ f)(p) = g(q) \neq 0$. This implies that f^* induces a unique homomorphism from $\mathcal{O}_p(V)$ to $\mathcal{O}_q(V)$ agreeing with f^*. f^* must map $\mathfrak{m}_q(W)$ into $\mathfrak{m}_p(V)$, because if we consider any g/h with $g(q) = 0$, then $f^*(g/h) = f^*(g)f^*(h)^{-1} = (g \circ f)/(h \circ f)$, and $(g \circ f)(p) = g(q) = 0$. □

If $T : \mathbf{A}^n \to \mathbf{A}^n$ is an affine isomorphism with $T(p) = q$, then it induces $T^* : \mathcal{O}_q(\mathbf{A}^n) \to \mathcal{O}_p(\mathbf{A}^n)$. T^* is an isomorphism from $k[W]$ to $k[V]$, mapping the set of functions not vanishing at q to the set of functions not vanishing at p, so in particular the induces isomorphism between the ring of fractions of the two rings by the corresponding multiplicative subset. More importantly, T^* induces an isomorphism from $\mathcal{O}_q(W)$ to $\mathcal{O}_p(V)$ if V and W are arbitrary varieties containing p and q respectively.

2.4 Projective Varieties

Let us try and carry out some analogous constructions to the affine varieties on projective varieties. The hardest one to begin with is the coordinate ring. Given a projective variety $V \subset \mathbf{P}^n$, we can form the *homogeneous coordinate ring* of the variety by taking $k_h[V] = k[C(V)] = k[X_0, \dots, X_n]/I(V)$. However, there is no way to evaluate these functions at points on V since a general element of $k_h[V]$ is not scale invariant. However, if f is a homogeneous function in $k_h[V]$ (the image of a homogeneous element in $k[X_0,\dots,X_n]$) then we can consider it's zeroes on V. Every function in $k_h[V]$ is the sum of such homogeneous functions.

Proposition 2.11. *Every element $f \in k_h[V]$ can be uniquely expanded as $\sum f_i$, where f_i is a form in $k_h[V]$. In other words, $k_h[V]$ is naturally a graded ring.*

Proof. Consider the residue of $f \in k[X_1,\dots,X_{n+1}]$ in $k_h[V]$, and write $f = \sum f_i$. If $f = g$ in $k[X_1,\dots,X_{n+1}]$, then $f - g \in I(V)$, and so $f_i - g_i \in I(V)$ for each i, hence $f_i = g_i$ in $k_h[V]$, implying the expansion like above is unique. □

If V is irreducible, then $k_h[V]$ is an integral domain, and we can form the *homogenous function field* $k_h(V)$. General elements of $k_h(V)$ cannot be viewed as functions on V. However, if f and g are both homogeneous forms of the same degree, then f/g *can* be identified as a function on V, since if both forms have degree m, then

$$f(\lambda x)/g(\lambda x) = \lambda^m f(x)/\lambda^m g(x) = f(x)/g(x).$$

The set of all expressions formed by taking quotients of forms in this way is called the *rational function field* of V, denoted $k(V)$. These functions have poles, just like on affine varieties, and these pole sets form projective varieties, because the representations of a function have denominators forming a class of homogenous functions, and the intersection of the zero sets of these denominators forms the pole set.

From the function field $k(V)$, we can form the localizations $\mathcal{O}_p(V)$ at a point $p \in \mathbf{P}^n$ as the set of elements $f/g \in k(V)$ with $g(p) \neq 0$. If V is not an irreducible variety, $k(V)$ is not well defined, but we can still form the localization of $k_h[V]$ by the set of forms which do not vanish at p, and then take the subring of quotients of the form f/g, where f and g are forms of the same degree. This subring is the local ring $\mathcal{O}_p(V)$. Just as in the case of affine rings, it has a unique maximal ideal $\mathfrak{m}_p(V)$, consisting of functions vanishing at p. Alternatively, once we define $\mathcal{O}_p(\mathbf{P}^n)$, we can define $\mathcal{O}_p(V)$ as the quotient of $\mathcal{O}_p(\mathbf{P}^n)$ by $I(V)$.

Example. *Consider $\mathbf{P}^1$ over an infinite field. Then $k_h[\mathbf{P}^1] \cong k[X,Y]$. Thus $k_h(\mathbf{P}^1)$ is just $k(X,Y)$. The elements of $k(\mathbf{P}^1)$ are therefore the form f/g, where $f,g \in k[X,Y]$ are both homogenous of the same degree. If we write $t = X/Y$,* *then $aX+bY = Y(at+b)$. Provided we are working over an algebraically closed field, every homogenous polynomial breaks down into linear factors of the form $aX + bY$, and if we consider f/g, where f breaks down into factors of the form $a_iX + b_iY$, and g breaks down into $c_iX + d_iY$, then*

$$\frac{f(X,Y)}{g(X,Y)} = \frac{(a_1t + b_1)\dots(a_nt + b_n)}{(c_1t + b_1)\dots(c_nt + b_n)}$$

In particular, this implies that $k(\mathbf{P}^1)$ is isomorphic to $k(t)$. For each point $p = [a : 1] \in \mathbf{P}^1$, we obtain the local ring

$$\mathcal{O}_p(\mathbf{P}^1) = \mathcal{O}_a(\mathbf{P}^1) = \{f(t)/g(t) : g(a) \neq 0\}.$$

and for $p = [1 : 0]$, we obtain the local ring

$$\mathcal{O}_p(\mathbf{P}^1) = \mathcal{O}_\infty(\mathbf{P}^1) = \{f(t)/g(t) : \deg(g) \geqslant \deg(f)\}$$

This is because if $u \in k(V)$ is defined at p, then we can write $u = f/g$, where $f, g \in k[X,Y]$ are degree m homogeneous forms with $g(1,0) \neq 0$. Then we can write

$$f(X,Y) = a_0X^m + a_1X^{m-1}Y + \dots + a_mY^m$$

and

$$g(X,Y) = b_0X^m + b_1X^{m-1}Y + \dots + b_mY^m,$$

where $b_0 \neq 0$. Then

$$u(X,Y) = \frac{a_0t^m + a_1t^{m-1} + \dots + a_m}{b_0t^m + b_1t^{m-1} + \dots + b_m} = \frac{f_0(t)}{g_0(t)}$$

and clearly $\deg(g_0) \geqslant \deg(f_0)$*. Conversely, for a general pair of polynomials $f_0(t) = a_0t^m + \dots + a_m$ and $g_0 = b_0t^n + \dots + b_n$, where $a_0, b_0 \neq 0$. Then we can write*

$$\frac{a_0t^m + \dots + a_m}{b_0t^n + \dots + b_n} = Y^{n-m}\frac{a_0X^m + a_1X^{m-1}Y + \dots + a_mY^m}{b_0X^n + b_1X^{n-1}Y + \dots + b_nY^n}$$

If this rational function is equal to $f(X,Y)/g(X,Y)$, where f and g are degree k forms with $g(1,0) \neq 0$, then

$$Y^n(a_0X^m + \dots + a_mY^m)g(X,Y) = Y^m(b_0X^n + \dots + b_nY^n)f(X,Y)$$

in $k[X,Y]$. If $n < m$, then

$$(a_0X^m + \dots + a_mY^m)g(X,Y) = Y^{m-n}(b_0X^n + \dots + b_nY^n)f(X,Y).$$

Evaluating these polynomials at $(1,0)$*, we conclude that* $a_0 g(1,0) = 0$*, which gives a contradiction since* a_0 *and* $g(1,0)$ *are both nonzero. Thus we conclude* $n \geqslant m$*.*

These local rings actually describe all *discrete valuation domains* A*, which are subrings of* $k(t)$ *containing* k *with quotient field is equal to* $k(t)$*. To see this, we state a simple lemma; suppose* A *contains another discrete valuation domain* B *whose quotient field is equal to* $k(t)$*, and suppose* $\mathfrak{m}_B \subset \mathfrak{m}_A$*. We claim* $A = B$*. Let us see how this is used to prove the theorem. If* A *is a discrete valuation ring, then either* $t \in A$ *or* $1/t \in A$*. In the first case,* $k[t]$ *is a subring of* A*, and* $\mathfrak{m} \cap k[t]$ *is a prime ideal of* $k[X]$*; algebraic closure implies this prime ideal is maximal, and is therefore equal to* $(t-a)$ *for some* $a \in k$*. This implies* $\mathcal{O}_a(\mathbf{P}^1)$ *is contained in* A*, because any polynomial* $f \in k[t]$ *with* $f(a) \neq 0$ *is invertible in* A *since it does not lie in* $\mathfrak{m}$*. Moreover,* $\mathfrak{m}_a \subset \mathfrak{m}$*, which implies that* $A = \mathcal{O}_a(\mathbf{P}^1)$*. The case where* $1/t \in A$ *adds the only additional ring* $\mathcal{O}_\infty(\mathbf{P}^1)$ *as a possibility.*

To prove the lemma, since

$$B = \{u \in k(t) : ord_B(u) \geqslant 0\},$$

it suffices to show that A *contains no elements* u *of* $k(t)$ *with* $ord_B(u) < 0$*. If such an element* u *existed, then* $ord_B(1/u) > 0$*, which would imply* $1/u \in B$*, and is therefore in* A*. Thus* u *and* $1/u$ *are units in* A*, which implies* u *and* $1/u$ *are not elements of* $\mathfrak{m}_A$*. But* $ord_B(1/u) > 0$*, implying that* $1/u \in \mathfrak{m}_B$*, and thus* $1/u \in \mathfrak{m}_A$*, which gives a contradiction. Thus* $A = B$*.*

As might be expected, at finite points local rings in affine and projective space correspond.

Theorem 2.12. *Let* V *be a projective variety, and let* p *be a finite point. Then* $\mathcal{O}_p(V) \cong \mathcal{O}_p(V_*)$*. Conversely, if* V *is an affine variety containing a point* p*, then* $\mathcal{O}_p(V) \cong \mathcal{O}_p(V^*)$*.*

Proof. We treat the affine to projective case, and leave the projective to affine case an exercise. Let V be an affine variety containing a point p. The map $p \mapsto (p,1)$ is a polynomial map embedding V in $C(V^*)$, and this induces a homomorphism from $k[C(V^*)] \to k[V]$, which is just the map $f \mapsto f_*$. This descends to a map from $\mathcal{O}_{(p,1)}(C(V^*))$ to $\mathcal{O}_p(V)$, because if $f(p,1) \neq 0$, then $f_*(p) \neq 0$. Thus we obtain a homomorphism from $\mathcal{O}_p(V^*)$ to $\mathcal{O}_p(V)$. Now suppose f and g are degree m forms, and $f_*/g_* = 0$ in

$\mathcal{O}_p(V)$, then there is $h \in k[V]$ with $h(p) \neq 0$, such that $f_* = g_* h$ in $k[V]$. But this means that $(f_*)^* = (g_*)^* h^*$ in $k[C(V^*)]$, and $h^*(p,1) = h(p) \neq 0$, so $(f_*)^*/(g_*)^* = 0$ in $\mathcal{O}_{(p,1)}(C(V^*))$. Since $(f_*)^*/(g_*)^*$ differs from f/g by a factor of X_{n+1}, which is a unit in $\mathcal{O}_{(p,1)}(C(V^*))$, we conclude that $f/g = 0$ in $\mathcal{O}_{(p,1)}(C(V^*))$, and thus is equal to zero in $\mathcal{O}_p(V^*)$. Conversely, we have a right inverse map given by mapping $f/g \in \mathcal{O}_p(V)$ to f^*/g^*, since $(f^*)_*/(g^*)_* = f/g$, which implies that the map is surjective. Thus we have shown the correspondence is an isomorphism. □

Since affine transformations can map a point to any other point, the easiest way to understand the local rings of a projective variety at points at infinity is to consider a projective transformation mapping that point to a finite point, which preserves the structure of the variety, and then to reduce the analysis to an affine variety as above.

It does not take much work to generalize this process from projective varieties to *biprojective varieties*. If $V \subset \mathbf{P}^n \times \mathbf{P}^m$ is a biprojective variety, we can consider $k_h[V] = k[X,Y]/I(C(V))$, where $C(V) \subset \mathbf{A}^{n+1} \times \mathbf{A}^{m+1}$ is the cone of V. Just as above, $k_h[V]$ is naturally graded, but $\mathbf{Z}^2$ graded rather than $\mathbf{Z}$ graded by the biforms of particular degrees. If V is irreducible, then one can consider $k_h(V)$, and we can define the ring $k(V)$ as the subring of this ring as the family of expressions of the form f/g, where f and g are bihomogeneous forms of the same degree. By taking a subring of $k(V)$ in the irreducible case, or taking a subring of the localization of $k_h[V]$ in the non-irreducible case, one can define the local rings $\mathcal{O}_p(V)$ around any point on the variety V. The localizations are isomorphic to the localizations of the variety $V_* \subset \mathbf{A}^n \times \mathbf{A}^m$ at pairs of finite points in $\mathbf{P}^n \times \mathbf{P}^m$.

Note that in this section we have not given a natural way to define a *morphism* between projective and biprojective varieties. This is not as easy as in the affine case, where we have algebraic and geometric correspondence between polynomial maps and morphisms of coordinate ring, since in the case of a projective variety we have no coordinate ring to work with. Projective transformations in ambient space do provide a simple family of linear maps through which we can identify projective varieties, but do not seem general enough to define all morphisms. To define morphisms between these families, it is useful to note that projective and biprojective varieties are *locally* affine varieties, and so it is natural to define a morphism between such varieties as those maps which are *locally* given as polynomial maps when working in an affine portion of the overall variety.

We carry this out in a greater generality in the next section.

2.5 Quasi-Affine Varieties and Regular Functions

Let us continue to study the intrinsic study of varieties, which we now consider as a general locus of points in $\mathbf{P}^{n_1} \times \cdots \times \mathbf{P}^{n_k} \times \mathbf{A}^m$. By taking varieties as closed sets, we obtain the Zariski topology on this space, and by restriction, a Zariski topology on any subvariety. Before we begin, we prove a small lemma about the basic open sets in the Zariski topology.

Lemma 2.13. *Suppose k is algebraically closed. Fix $f, g \in k[X_1, \dots, X_n]$. If $U_f \subset U_g$, then g divides f^k for some k.*

Proof. It follows that $f \in Z(g)$, so by the nullstellensatz, $f \in \text{Rad}(g)$, which immediately implies the claim. □

A key trick to obtaining more geometric information about a variety is to equip any open subset with an algebraic structure. For a variety V, and any Zariski open set U, we will introduce a ring $k[U]$, consisting of *regular functions* on U. An element of $k[U]$ is a function $u : U \to k$ such that each $p \in U$ has a Zariski open neighbourhood $U' \subset U$ and there exists $f, g \in k[V]$ such that $g(p) \neq 0$ for all $p \in V$, and $u(p) = f(p)/g(p)$ for all $p \in U'$. Algebraically, $k[U]$ has the structure of a ring. We will very quickly see that this agrees with the classical definition in the case of the coordinate ring of a variety.

Example. *Suppose V is an irreducible variety, and let U be a Zariski open subset. Any element u of $k(V)$ whose domain contains U can be interpreted as a regular function. If A is the subring of $k(V)$ consisting of rational functions whose domains contain U, we thus have a natural morphism frim A to $k[U]$. We claim this actually gives an isomorphism of the two rings.*

First, let us see that the morphism is injective, since if $u \in k(V)$ satisfies $u(p) = 0$ for each $p \in U$, then if we find $f, g \in k[V]$ such that $u = f/g$, and write $U' = \{p \in U : g(p) \neq 0\}$, then U' is Zariski dense in V, and $f(p) = 0$ for all $p \in U'$, which implies $f(p) = 0$ for all $p \in V$, so $f = 0$, hence $u = 0$.

The morphism is also surjective. If $u : U \to k$ is a regular function, then there exists an open set $U' \subset U$ and $f, g \in k[V]$ such that $g(p) \neq 0$ for all $p \in V$, and $u(p) = f(p)/g(p)$ for all $p \in U'$. If we choose $U'' \subset U$ and find $f', g' \in k[V]$ such that $g'(p) \neq 0$ for all $p \in U''$ such that $u(p) = f'(p)/g'(p)$ for all $p \in U''$, *then $f'(p)g(p) = f(p)g'(p)$ for all $p \in U' \cap U''$. Since $U' \cap U''$ is Zariski dense in V, this implies $f'g = fg'$ in $k[V]$, and so $f'/g' = f/g$ in $k(V)$. Thus the map $u \mapsto f/g$ is invariant of the choice of open set U' we choose, which gives a right inverse for the morphism we were studying, proving surjectivity.*

Example. *Let V be a variety. We claim the ring $k[V]$ of regular functions is equal to the coordinate ring $k[V]$ in the classical sense. Temporarily denote the new definition by A, and use $K[V]$ to denote the classical condition Using quasicompactness, consider a finite family of Zariski open sets $\{U_i\}$ covering V, where without loss of generality we may assume that $U_i = U_{a_i}$ for some $a_i \in k[V]$. Fix $u \in A$, and for each i, find $f_i, g_i \in k[V]$ such that $g_i(p) \neq 0$ for any $p \in U_i$, and such that $u(p) = f_i(p)/g_i(p)$ for all $p \in U_i$. Since $U_{a_i} \subset U_{g_i}$, this means that g_i divides $a_i^{k_i}$ in $k[V]$ for some integer k_i. Replacing a_i with $a_i^{k_i}$, we may assume that g_i divides a_i for each i, so we may write $a_i = b_i g_i$. But this means that for each $p \in U_i$,*

$$u(p) = f_i(p)/g_i(p) = b_i(p)f_i(p)/b_i(p)g_i(p) = b_i(p)f_i(p)/a_i(p).$$

Thus we may actually assume without loss of generality that $a_i = g_i$. Now $V = \bigcup U_{g_i^2}$, which implies that $Z(g_1^2, \dots, g_n^2) = \varnothing$. It therefore follows from the nullstellensatz that $(g_1^2, \dots, g_n^2) = k[V]$, i.e. there are $c_1, \dots, c_n \in k[V]$ such that $1 = c_1 g_1^2 + \dots + c_n g_n^2$ in $k[V]$. We claim that for all $p \in V$,

$$u(p) = c_1(p)f_1(p)g_1(p) + \dots + c_n(p)f_n(p)g_n(p).$$

For each i and j, $g_i g_j (f_i g_j - f_j g_i) = 0$ in $k[V]$, i.e. $f_i g_i (g_j)^2 = f_j g_j (g_i)^2$. It therefore follows that on $k[V]$, for each i,

$$g_i^2 \sum_{k=1}^n c_k f_k g_k = \sum_{k=1}^n c_k f_k g_k (g_i)^2 = \sum_{k=1}^n c_k f_i g_i (g_k)^2 = f_i g_i.$$

Thus whenever $g_i(p) \neq 0$,

$$\sum_{k=1}^n c_k(p) f_k(p) g_k(p) = f_i(p)/g_i(p) = u(p).$$

Since i was arbitrary, this completes the proof.

Example. *Let V be a variety, let $f \in k[V]$ be reduced, and let $U = U_f$. Then $k[U]$ is isomorphic to the localization of $k[V]$ obtained by adding inverse to f^n for each $n > 0$. Let A be this localization. Every element of A induces a well-defined regular function on U, for if $g/f^n = g'/f^m$ in A, then there exists k such that $gf^{m+k} - g'f^{n+k} = 0$ in $k[V]$, which means that $g(p)/f(p)^n = g'(p)/f(p)^m$ for each $p \in V$ with $f(p) \neq 0$. Moreover, this map is injective, for if $g(p)/f(p)^n = 0$ for each $p \in V$ with $f(p) \neq 0$, then $gf = 0$ in $k[V]$, which implies $g/f^n = 0$ in A. The surjectivity is proved in a very similar way to the example above and is left as an exercise.*

Theorem 2.14. *Any regular map $f : U \to k$ is Zariski continuous.*

Proof. Since Zariski closed sets in k consist precisely of finite point sets, it suffices to show that $f^{-1}(s)$ is Zariski closed in U for each $s \in k$. Since each component of V is closed in V, we may assume without loss of generality that V is irreducible. But then for each $p \in U$ with $f(p) = s$, there exists $u, v \in k_h[V]$ with the right homogeneity conditions such that $f = u/v$ and $v(p) \neq 0$. The set U' of points of U where $v(p) \neq 0$ is a Zariski open set containing p. And $f^{-1}(s) \cap U' = Z(u) \cap U'$ is a Zariski closed subset of U'. But this implies that we have a family of open sets $\{U_\alpha\}$ covering U such that $f^{-1}(s) \cap U_\alpha$ is Zariski closed in U_α for each α. And this implies $f^{-1}(s)$ is closed, because for any $p \notin f^{-1}(s)$, there exists an open set U_α containing p, and then $U_\alpha - (f^{-1}(s) \cap U_\alpha)$ is Zariski open in U_α, hence Zariski open in U, and so p has a neighbourhood not containing any points in $f^{-1}(s)$. □

An open subset of a variety will be known as a *quasi-variety*. We will find that the theory of quasi-varieties has a theory analogous to that of varieties. In particular, we can consider a category of such sets. The family of varieties embed in the family of quasi-varieties, which will lead to a nice relation between these two families.

2.6 Homomorphisms of Quasi-Varieties

A *morphism* between two quasi-varieties X and Y is a Zariski-continuous map $\phi : X \to Y$ such that for each open set V of Y with $\phi^{-1}(V) = U$, we have $\phi^*(k[V]) \subset k[U]$, i.e. if u is a regular map on V, then $u \circ f$ is a regular map on X.

Example. *Let V and W be affine varieties. If $\phi : V \to W$ is a morphism of quasi-varieties, then we obtain a ring morphism $\phi^* : k[W] \to k[V]$. This is just a ring morphism from $k[W]$ to $k[V]$, and in previous sections we have seen this corresponds precisely to a polynomial map from V to W. Since a polynomial map is automatically Zariski continuous, every polynomial map from V to W is a morphism of quasi-varieties. Thus the family of affine varieties and polynomial maps between them form a* full subcategory *of the category of quasi-varieties.*

Since in this chapter we are trying to work as intrinsically as possible with algebraic varieties, we now redefine a *affine variety* as any quasi-variety which is isomorphic to the affine varieties we have defined before. What we referred to previously as an affine variety will now be referred to as an affine subvariety of $\mathbf{A}^n$. Similarily, we redefine the notion of a *projective variety* and a *biprojective variety*. Since isomorphic quasi-varieties are homeomorphic, these spaces carry the same topological information. Moreover, the induced ring morphisms ϕ^* will be isomorphisms, so they also carry the same algebraic information through their regular maps.

Example. *Using this new terminology, it makes sense to say that every projective / biprojective variety is a finite union of affine subvarieties. For simplicity, we consider only the case of an irreducible projective variety. Let V be an irreducible projective variety in $\mathbf{P}^n$ with coordinates $[X_0 : \dots : X_n]$; then let us consider the ring of regular functions on the Zariski open set $U = \{x \in V : x_0 \neq 0\}$. Assuming U is nonempty, we conclude X_0 is nonzero in $k_h[V]$. Thus the forms $Z_1 = X_1/X_0, \dots, Z_n = X_n/X_0$ are well defined elements of $k(V)$. Moreover, these forms are all regular on U, and are thus elements of $k[U]$. We claim that $k[U] = k[Z_1, \dots, Z_n]$. Suppose u lies in $k[U]$, and let $\mathfrak{a}$ be the homogeneous ideal of $k_h[V]$ generated by all forms b for which there exists a form a of the same degree with $u = a/b$. Since u is well defined on U, $Z(\mathfrak{a}) \subset Z(X_0)$, which implies by the nullstellensatz that $(X_0) \subset Rad(\mathfrak{a})$. Thus there exists some k such that $X_0^k \in \mathfrak{a}$. It is simple to verify that this means there exists a homogeneous form $a \in k_h[V]$ of degree k such that $u = a/X_0^k$. But if $a(X) = \sum_{i_0 + \dots + i_n = k} a_i X_0^{i_0} \dots X_n^{i_n}$, then*

$$u(X) = a(X)/X_0^k = \sum_{i_0 + \dots + i_n} a_i Z_1^{i_1} \dots Z_n^{i_n},$$

which completes the proof that $k[U] = k[Z_1, \dots, Z_n]$. It is then simple to argue from this that the embedding map $f : \mathbf{A}^n \to U_0$ is an isomorphism, where U_0

is the finite points of $\mathbf{P}^n$; it then follows that f restricts to an isomorphism between the closed affine variety $f^{-1}(U)$ and U.

Theorem 2.15. *Let $\phi : X \to Y$ be a map between quasivarieties, and let X be the union of Zariski-open sets $\{W_\alpha\}$. Suppose that for each α, the restriction maps $\phi_\alpha : W_\alpha \to Y$ are morphisms of quasi-varieties. Then ϕ is a morphism of quasi-varieties.*

Proof. Since ϕ_α is Zariski continuous for each α, the map ϕ is Zariski continuous. Let V be a Zariski open subset of Y, and let $U = \phi^{-1}(V)$. For each α, let $U_\alpha = \phi^{-1}(V) \cap W_\alpha$. For $g \in k[V]$, let $f = \phi^*(g)$. Then for each α, we let $f_\alpha \in k[U_\alpha]$ be equal to $\phi_\alpha^*(f_\alpha)$. Then each f_α is regular, and for each $p \in W_\alpha$, $f(p) = f_\alpha(p)$. But it is simple to verify that this implies the map f is regular, since each point has a neighbourhood contained in the domains on one of the f_α, upon which it behaves like a rational map. □

Remark. Being an isomorphism is also a local property of ϕ, in the sense that if $\phi : X \to Y$ is a bijective map, and we can find an open cover $\{U_\alpha\}$ of X and an open cover $\{V_\alpha\}$ of Y such that the restriction map $\phi : U_\alpha \to V_\alpha$ is an isomorphism of quasi-varieties, then ϕ is an isomorphism.

Using the Segre embedding mentioned earlier, we can embed $\mathbf{P}^n \times \mathbf{P}^m$ into $\mathbf{P}^{n+m+nm}$, from which it follows that biprojective varieties are just projective varieties in disguise.

Theorem 2.16. *Any closed subvariety of $\mathbf{P}^{n_1} \times \cdots \times \mathbf{P}^{n_k}$ is a projective variety.*

Proof. Since every closed subvariety of a projective variety is a projective variety, it suffices to prove that $\mathbf{P}^{n_1} \times \mathbf{P}^{n_k}$ is a projective variety. Using induction, it suffices to prove that the Segre embedding $S : \mathbf{P}^n \times \mathbf{P}^m \to \mathbf{P}^{n+m+nm}$ has as an image a projective variety V, and that S restricts to an isomorphism between these two quasi-varieties.

First, let us identify V. Recall that the Segre embedding is given by the map $(X, Y) \mapsto [X_i Y_j]$, where $X = [X_0 : \cdots : X_n]$ and $Y = [Y_0 : \cdots : Y_m]$. Let us write the homogeneous coordinates of $\mathbf{P}^{n+m+nm}$ as $[Z_{ij}]$, where $Z_{ij} = X_i Y_j$. Any element of the image of the embedding satisfies $Z_{i_1 j_1} Z_{i_2 j_2} - Z_{i_1 j_2} Z_{i_2 j_1} = 0$. We claim that the image consists of precisely the points satisfying all such conditions. Suppose a point $p \in \mathbf{P}^{n+m+nm}$ satisfies all these conditions, and has coordinates $[z_{ij}]$. Let $S = \{(i,j) : z_{ij} \neq 0\}$. Then S is a Cartesian product, for if $z_{i_1 j_1}$ and $z_{i_2 j_2}$ are nonzero, then so too are $z_{i_1 j_2}$

and $z_{i_2 j_1}$. If we write $S = I \times J$, then setting coordinates outside of I and J to be zero, it is easy to find a pair of points x and y such that $S(x,y) = z$. Thus the image of $\mathbf{P}^n \times \mathbf{P}^m$ is certainly a variety. If W is a projective subvariety of V, then $S^{-1}(W)$ is a biprojective variety in $\mathbf{P}^n \times \mathbf{P}^m$. Since $\mathbf{P}^n \times \mathbf{P}^m$ is irreducible, it follows that V is an irreducible projective variety.

It is simple to prove that S is a bijective map between $\mathbf{P}^n \times \mathbf{P}^m$ and V, so all that remains is to show that the resulting maps S^* are isomorphisms. In fact, by symmetry, it suffices to show that the map $S^* : k[U_0 \cap V] \to k[U_1 \times U_2]$ is an isomorphism, where U_0, U_1, and U_2 are the set of finite points in their respect projective spaces, i.e. with $Z_{00} = 0$, $X_0 = 0$, and $Y_0 = 0$ respectively. We may identify $k[U_0 \cap V]$ with $k[V_*]$, and $k[U_1 \times U_2]$ with $k[X_1, \dots, X_n, Y_1, \dots, Y_m]$. But then it is easy to see that S, restricted to a map from $U_1 \times U_2$ to $U_0 \cap V$, is a polynomial map. Conversely, we have an inverse map given by the map $Z \mapsto ((Z_{10}, \dots, Z_{n0}), (Z_{01}, \dots, Z_{0n}))$, and this is also clearly a polynomial map, so the map is an isomorphism. □

Example. *A product of affine varieties is an affine variety. The last theorem shows that a product of projective varieties is a projective variety.*

Thus, intrinsically, the theory of biprojective varieties is identical to the theory of projective varieties, which means we can avoid the kind of technicalities we were worried about in previous sections. In particular, any *variety*, in the sense that we defined earlier, occurs as a quasi-variety in projective space or one of it's closed subvarieties.

However, the theory we have developed has some strange consequences that we might not have expected; a closed subset of an affine variety is certainly an affine variety, but an open subset of an affine variety can *also* be an affine variety.

Example. *Let V be an affine variety, let $f \in k[V]$ be reduced, and let U be the Zariski open subset of V consisting of points $p \in V$ where $f(p) \neq 0$. We have seen that $k[U]$ can be identified with the localization of $k[V]$ obtained by adding inverses to f^n for each $n > 0$. If V is irreducible, and we write $k[V] = k[X_1, \dots, X_n]/I(V)$, then this localization is isomorphic to $k[X_1, \dots, X_n, Y]/(I(V) + (f(X_1, \dots, X_n)Y - 1))$, and this is the coordinate ring of some variety W.*

We have a natural polynomial map $\phi : W \to U$ given by projection. Conversely, we have an inverse map from U to W mapping a point x to $(x, 1/f(x))$, and this map is Zariski continuous because if C is a closed subset of W given *by the zeroes of a family of polynomials $t_1,\dots,t_k \in k[W]$, then $\phi(C)$ is given by the set of points x such that $t_1(x,1/f(x)) = \cdots = t_k(x,1/f(x)) = 0$. If $t_i(x,y) = \sum a_{ij}X^iY^j$, then this equation really says that $\sum a_{ij}x^i/f(x)^j = 0$, and if the maximum value of j is r, the this equation holds in U if and only if $\sum a_{ij}x^i f(x)^{k-j} = 0$, which is a polynomial equation in x. Thus $\phi(C)$ is Zariski closed in U and this proves continuity of the inverse map to ϕ. Thus ϕ is a homeomorphism between W and U.*

To complete the proof, the only remaining fact is to prove that ϕ is an isomorphism between $k[U_1]$ to $k[U_2]$ for any open set U_1 in W and $U_2 = \phi(U_1)$. Injectivity of this morphism is clear, so it suffices to prove surjectivity. If $u \in k[U_2]$, then around each $p \in U_2$, we may find $f,g \in k[V]$ such that $u = f/g$, and $g(p) \neq 0$. Let $U' = \{p \in V : g(p) \neq 0\}$. Then $\phi^{-1}(U')$ is a Zariski open subset of W, and $\phi^*(f/g) = u|_{U'}$. The composition property of regular maps gives the required surjectivity.

The open sets given by the non-vanishing points of a polynomial are the largest Zariski open sets in a classical affine variety. As a result of the above example, we obtain a simple consequence.

Theorem 2.17. *Let U be a quasi-variety. Then every point p has an open neighborhood isomorphic to an affine variety.*

Proof. Without loss of generality, we may assume U is an open subset of an affine variety V. Since $V - U$ is a subvariety of V not containing p, there is $f \in k[V]$ such that $f(p) \neq 0$, but vanishes on $V - U$. Then $U' = \{p \in V : f(p) \neq 0\}$ is a Zariski open subset of p in U, and the last example showed it is an affine variety. □

Remark. More generally, for any finitely family of points $p_1,\dots,p_n \in U$ in a quasivariety U, there exists an affine subvariety of U containing $p_1,\dots,p_n$.

As a result of this theorem, to construct morphisms between quasi-varieties it suffices to choose affine varieties locally, then show that in the appropriate affine coordinates, the map is a polynomial.

Theorem 2.18. *Suppose X and Y are quasi-varieties, and X' is a closed subvariety of X. Let $f : X \to Y$ be a morphism, and suppose that $f(X') \subset Y'$ for some open/closed quasi-variety Y' contained in Y. Then the induced map $\tilde{f} : X' \to Y'$ is a morphism.*

Proof. By reducing to an open cover in each space, we may assume that X and Y are affine varieties, which we may assume are subsets of $\mathbf{A}^n$ and $\mathbf{A}^m$ respectively. Then f acts as a polynomial map from X to Y. Since X' is a closed subvariety of X, X' is also an affine variety, and so $f : X' \to Y$ is also a polynomial map, hence a morphism. This completes the proof in the case where Y' is an open subvariety of Y since the quasi-variety morphism descends to open subsets quite easily. Similarily if Y' is a closed subvariety, then Y' is also an affine variety, and $f : X' \to Y'$ a polynomial map, which completes the proof. □

2.7 Algebraic Groups

We can now use our intrinsic theory to build up a theory of *algebraic groups*, i.e. a variety that has a law of composition which is *regular*. More precisely, an algebraic group is a quasi-variety V together with a group operation such that the maps $(x,y) \mapsto xy$ and $x \mapsto x^{-1}$ are *regular* from $V \times V$ and V to V respectively.

Example. *For any n, $\mathbf{A}^n$ is an algebraic group with the operation of addition, since the maps $(x,y) \mapsto x+y$ and $x \mapsto -x$ are all regular maps.*

Example. *The quasi-variety $\mathbf{A}^1 - \{0\}$ is an open affine subvariety of $\mathbf{A}^1$, which forms a group under multiplication. To see the group operations are regular, we embed $\mathbf{A}^1 - \{0\}$ as a closed variety of $\mathbf{A}^2$ corresponding to the closed variety V which is the locus of points satisfying $XY = 1$. The multiplication map on $\mathbf{A}^1 - \{0\}$ corresponds to the map $f((x_1,y_1),(x_2,y_2)) = (x_1x_2,y_1y_2)$, and the inversion map corresponds to the map $f(x,y) = (y,x)$, which are easily seen to be polynomial maps, and thus regular.*

Example. *A more general form of the last example is the group $GL_n(k)$, which can be identified with the open subvariety of $\mathbf{A}^{n^2}$ consisting of matrices with non-zero determinant. In fact, $GL_n(k)$ is a basic open set, and is therefore an affine subvariety; we can identify $GL_n(k)$ with the closed subvariety of $\mathbf{A}^{n^2+1}$ which, if we give coordinates $(M_{11},\dots,M_{nn},Y)$ to $\mathbf{A}^{n^2+1}$, satisfy the equation $\det(M)Y = 1$. The multiplication map takes the form $f((M_1,y_1),(M_2,y_2)) = (M_1M_2,y_1y_2)$, which is easily seen to be a polynomial map. The inversion map can be seen to be a polynomial using Cramer's rule. Recalling that the adjoint* *matrix $Adj(M)$ is a polynomial in the entries of M, and that*

$$M^{-1} = Adj(M)/\det(M),$$

we conclude that the inversion map is given by $(M, y) \mapsto (y \cdot Adj(M), \det(M))$, which is a polynomial in M and y.

Example. *Recall that if C is a cubic curve in the projective plane, then the Zariski open set of simple points on C form a commutative group by fixing a simple point o, considering the map $\phi : C \times C \to C$, where $\phi(p, q) = r$ if there is a line L intersecting p, q, and r, and then defining $p \oplus q = \phi(o, \phi(p, q))$. If $r = \phi(o, o)$, then the inverse map is given by $p \mapsto \phi(p, r)$. Thus to show that the simple points form an algebraic group, it suffices to show that ϕ is a regular map. Performing a projective transformation, in all coordinates, we may assume without loss of generality that C has a simple point at $[0 : 1 : 0]$. Then C is the locus of some polynomial of the form $Y^2Z = X^3 + aX^2Z + bXZ^2 + cZ^3$, for $a, b, c \in k$. It suffices to show that ϕ is regular when it's domain is restricted to the affine points. So let $p = [x_1 : y_1 : 1]$ and let $q = [x_2 : y_2 : 1]$. If $x_1 \neq x_2$, then set $\lambda = (y_2 - y_1)/(x_2 - x_1)$. Then the line L between p and q is given as the locus of points satisfying $Y = \lambda(X - x_1Z) + y_1Z = \lambda X + (y_1 - \lambda x_1)Z$. On this line, the equation $Y^2Z = X^3 + aX^2Z + bXZ^2 + cZ^3$ is equivalent to the equation*

$$X^3 + (a - \lambda^2)X^2Z + (b - 2\lambda(y_1 - \lambda x_1))XZ^2 + (c - (y_1 - \lambda x_1)^2)Z^3 = 0.$$

This equation factors as

$$(X - x_1Z)(X - x_2Z)(X - (\lambda^2 - a - x_1 - x_2)Z).$$

Thus we conclude that

$$\phi(p, q) = [\lambda^2 - a - x_1 - x_2 : \lambda(\lambda^2 - a - 2x_1 - x_2) + y_1 : 1].$$

This is a polynomial map in x_1, x_2 and λ. Performing a similar construction when $y_1 \neq y_2$ gives a polynomial map in y_1, y_2 and $1/\lambda$. A similar approach works when $p = q$.

Example. *Let C be the family of simple points on the cubic curve satisfying $Y^2Z = X^3$, i.e. all points except for $[0 : 0 : 1]$. We have a morphism $f : \mathbf{A}^1 \to C$ given by $f(t) = [t : 1 : t^3]$. This morphism is surjective, since any point* *satisfying $Y^2Z = X^3$ with a non-zero y coordinate can be written as $[t : 1 : t^3]$. Consider C as an algebraic group by letting $[0 : 1 : 0]$ be the unit element. Given $p = [t_1 : 1 : t_1^3]$ and $q = [t_2 : 1 : t_2^3]$, the line L between p and q is given by the locus of points satisfying $Z = \lambda X + (t_1^3 - \lambda t_1)Y$, where*

$$\lambda = (t_2^3 - t_1^3)/(t_2 - t_1) = t_2^2 + t_1 t_2 + t_1^2.$$

The intersections of C with L are in one to one correspondence with the zeroes of the equation

$$X^3 - \lambda X Y^2 + t_1 t_2 (t_1 + t_2) Y^3 = 0,$$

and this factors as

$$(X - t_1)(X - t_2)(X + (t_1 + t_2)).$$

Thus we conclude the x coordinate of $\phi(p,q)$ is $-(t_1 + t_2)$. But this means $p \oplus q$ is equal to $t_1 + t_2$. Thus f is actually an isomorphism of groups. To show that the inverse is regular, and thus that the two algebraic groups are isomorphic, it suffices to note that the inverse can be written as $[x : y : z] \mapsto x/y$, and is therefore regular.

2.8 Dimensions of Varieties

We now give a method for determining the *dimension* of a variety intrinsically. For instance, a hypersurface in n dimensional space should be a variety of dimension $n-1$, and a 'generic' variety specified as the locus of m polynomials should have dimension $n - m$. The key to measuring this dimension is via a concept known as the *transcendence degree*.

Given a field K containing a subfield k, the *transcendence degree* of K over k, denoted $\text{tr.deg}_k(K)$, is the smallest number of elements $x_1, \dots, x_n$ such that K is algebraic over $k(x_1, \dots, x_n)$. The elements $\{x_1, \dots, x_n\}$ are then algebraically independant over k and are the largest such set. We say that K is an *algebraic function field in n variables over k*.

If X is an irreducible quasi-variety, then $k[X]$ is an integral domain, and one can form the field of fractions $k(X)$ - in this case, this ring will be identical to $k(V)$, if X is Zariski dense in V. Then $k(X)$ is a finitely generated extension of k, and we define the *dimension* of X to be $\text{tr.deg}_k(K(X))$. A variety with dimension one or two is called a *curve* or *surface* respectively.

Lemma 2.19. *Let K be an extension of an algebraically closed field k with $tr.deg_k(K) = 1$. Fix $x \in K - k$. Then K is algebraic over $k(x)$, and if k has characteristic zero, there exists $y \in K$ such that $K = k(x,y)$. If A is a subring of K containing k, with quotient field K, and $\mathfrak{p}$ is a prime ideal of A, then the natural map from $k \to A/\mathfrak{p}$ is an isomorphism (so every prime ideal is maximal).*

Proof. Find $t \in K$ such that K is algebraic over $k(t)$. Then there exists a polynomial $f \in k[X,T]$ such that $f(x,t) = 0$. Since x is not algebraic over k, this shows t is algebraic over $k(x)$. Thus $k(x,t)$ is algebraic over $k(x)$, which implies K is algebraic over $k(x)$ by transitivity. The fact that $K = k(x,y)$ for some $y \in K$ follows from the primitive element theorem.

The morphism from k to $A/\mathfrak{p}$ is trivially injective, so it suffices to show it is surjective. Fix $a \in A$ and a nonzero element $x \in \mathfrak{p}$. Now K is algebraic over $k(x)$ since $x \notin k$, which implies there is a polynomial $f \in k[A,X]$ such that $f(a,x) = 0$. If we write $f(A,X) = \sum_i a_i(A)X^i$, then by choosing f to be of smallest degree, it follows that $a_0(A) \neq 0$. But then since $f(a,x) = 0$, $a_0(a) \in \mathfrak{p}$, which means $a + \mathfrak{p}$ is algebraic over k in $A/\mathfrak{p}$. Since k is algebraically closed, this implies that $a + \mathfrak{p} \in k$, i.e. $A/\mathfrak{p} = k$. □

It is simple to see that if X is an irreducible quasi-variety which is Zariski dense in a variety V, then $\dim(U) = \dim(V)$. Similarily, if V is an irreducible affine variety, then $k(V)$ is isomorphic to $k(V^*)$ and so $\dim(V) = \dim(V^*)$. Here are some other basic properties of dimension.

Theorem 2.20. *Here are some other basic properties of dimension:*

- *An irreducible variety of dimension zero is a point.*
- *Every irreducible closed subvariety of a curve is a point.*
- *An irreducible closed subvariety of $\mathbf{A}^2$ or $\mathbf{P}^2$ has dimension one if and only if it is the zero set of a single irreducible polynomial.*

Proof. Let V be an irreducible variety of dimension zero. Since k is algebraically closed, it follows that $k(V) = k$, and thus $k[V] = k$. But this can only be true if $I(V)$ is a maximal ideal of $k[X_1,\dots,X_n]$, and thus by the Nullstellensatz, it follows that V is a point.

Next, if V is a curve, then $A = k[V]$ is a subring of $k(V)$; it follows from the last lemma that if W is an irreducible subvariety of V, then $I(W) \subset A$ is a prime ideal, hence a maximal ideal, and so W corresponds to a point.

If V is an irreducible closed subvariety of $\mathbf{A}^2$, then it has dimension zero, one, or two. We know that $Z(f,g)$ consists of finitely many points if f and g are relatively prime, so since V is irreducible, unless V is point (and therefore has dimension zero) or $V = \mathbf{A}^2$, we have $V = Z(f)$ for a single irreducible polynomial f. It suffices to show $\dim(V) = 1$ in this case. Since $k(V)$ is generated by the image of X and Y in $k[V]$, it suffices to show that Y is algebraic over $k(X)$. But this follows immediately from the fact that $f(X,Y) = 0$.

If V is an irreducible closed subvariety of $\mathbf{P}^2$, then without loss of generality by applying a projective transformation we may assume that V consists of finitely many points. Since the infinite points of V form a closed subvariety, it follows from irreduciblity that $(V_*)^* = V$. But this means that $\dim(V) = \dim(V_*)$. Thus if $\dim(V) = 0$, then V is a point, and if $\dim(V) = 2$, then $V = \mathbf{P}^2$. Neither of these sets are the zero set of a single polynomial by the nullstellensatz. Thus it suffices to show that $\dim(V) = 1$ if V is the zero set of a single, irreducible homogeneous polynomial $f \in k[X,Y,Z]$. But $k(V)$ is generated by the image of X/Y and Y/Z, and f has degree m, then

$$f(X/Z, Y/Z, 1) = Z^m f(X,Y,Z) = 0,$$

so Y/Z is algebraic over $k(X/Z)$. □

If K is any field which is an algebraic extension of $k(x_1, \dots, x_n)$,

2.9 Rational Maps

Let us try and define *rational maps* between two quasi-varieties. We want to think of a rational map between two quasi-varieties X and Y as a morphism from a Zariski open subset U of X to Y. But we also want to identify two rational maps $f : U \to Y$ and $g : V \to Y$ if they agree when restricted to maps on $U \cap V$. So a *rational map* will be an equivalence class of such maps. The *domain* of a rational map will be the largest Zariski open set upon which the rational map can be defined. For a rational map $f : X \to Y$ with domain U, $f(p)$ is well defined for each $p \in U$. A rational map $f : X \to Y$ with domain U will be called *dominating* if $f(U)$ is dense in X.

Lemma 2.21. *Let X and Y be irreducible. Every dominating rational map $f : X \to Y$ induces a homomorphism $f^* : k(Y) \to k(X)$.*

Proof. Let U be the domain of f, and consider the induced morphism $f^* : k[Y] \to k[U]$. It suffices to show this map is injective. If u is a regular function on Y with $f^*(u) = 0$, then u vanishes on the image of f. But this image is dense, which implies (since u is continuous) that u vanishes everywhere, i.e. $u = 0$. Thus f^* induces a unique map from $k(Y)$ to $k(U)$.

Since $\mathcal{O}_q(Y)$ is a local ring, so too is $f^*(\mathcal{O}_q(Y))$. It's maximal ideal consists precisely of the image of any regular function $u \in \mathcal{O}_q(Y)$ with $u(q) = 0$. But it is easy to verify that $u \circ f \in \mathcal{O}_p(X)$. □

Remark. It follows from this that if $f : X \to Y$ is a dominating rational map, then $\dim(Y) \leqslant \dim(X)$.

Let us analyze the homomorphism $f^* : k(Y) \to k(X)$. If p lies in the domain of f, with $f(p) = q$. It is simple to verify that $f^*(\mathcal{O}_q(Y))$ is a subring of $\mathcal{O}_p(X)$, since if $u \in k(Y)$ is defined at q, f^*u is defined at p. Now $f^*(\mathcal{O}_q(Y))$ is a local ring, which is a subring of $\mathcal{O}_p(X)$. If A and B are local rings, where A is a subring of B, then we say B *dominates* A if $\mathfrak{m}_A \subset \mathfrak{m}_B$.

Theorem 2.22. *If p belongs to the domain of a dominating rational map f, with $f(p) = q$, then $\mathcal{O}_p(X)$ dominates $f^*(\mathcal{O}_q(Y))$. Conversely, if $p \in X$, $q \in Y$, $f^*(\mathcal{O}_q(Y)) \subset \mathcal{O}_p(X)$, and $\mathcal{O}_p(X)$ dominates $f^*(\mathcal{O}_q(Y))$, then p lies in the domain of f, and $f(p) = q$.*

Proof. It is simple to see that $\mathcal{O}_p(X)$ dominates $f^*(\mathcal{O}_q(Y))$ in the first case, for if $u \in k(Y)$ is defined at q with $u(q) = 0$, then $f^*u(p) = u(f(p)) = 0$. To prove the converse, consider neighborhoods $U \subset X$ and $V \subset Y$ of p and q respectively, each isomorphic to an irreducible affine variety. If V is isomorphic to a subset of $\mathbf{A}^n$, then we may write $k[V] = k[y_1, \dots, y_n]$, where $Z(y_1, \dots, y_n) = \{q\}$. Since $f^*(\mathcal{O}_q(Y))$ is a subring of $\mathcal{O}_p(X)$, there exists polynomials $a_i, b_i \in k[U]$ such that $f^*(y_i) = a_i/b_i$, where $b_i(p) \neq 0$ and $a_i(p) = 0$. Let $b = b_1 \dots b_n$ and consider the basic neighborhood U_b of q, which we may assume without loss of generality is a subset of U. Thus f induces a homomorphism f^* from $k[V]$ to $k[U_b]$, which corresponds to a unique morphism $g : U_b \to V$ since these sets are affine. This morphism satisfies $g(p) = q$ precisely because $f^*(\mathfrak{m}_q) \subset \mathfrak{m}_p$ (if $g(p)$ equalled some other element r than q, we could find a regular function u vanishing at q

but not r, and then $f^*(u)$ would vanish at p, which is impossible). Moreover, if U_0 is the domain of f, then f agrees with g on $U_0 \cap U_b$ because both induce the same morphism from $k[V]$ to $k[U_b]$. Thus f is defined on U_b and is equal to g there. □

Theorem 2.23. *Any homomorphism from $k(Y)$ into $k(X)$ is induced by a unique dominating rational map from X to Y.*

Proof. Without loss of generality, we may assume X and Y are closed affine subvarieties of $\mathbf{A}^n$ and $\mathbf{A}^m$. If $\phi : k(Y) \to k(X)$ is any homomorphism, with $k[Y] = k[y_1,\ldots,y_m]$, then there are $a_i, b_i \in k[X]$ such that $\phi(y_i) = a_i/b_i$. If $b = b_1 \ldots b_m$ and $U = U_b$, then $\phi(k[Y]) \subset k[U_b]$. Since Y and U_b are affine, this means that there is a morphism from U_b and Y which induces the homomorphism ϕ. Since ϕ is injective from $k[Y]$ to $k[U_b]$, this means $f(U_b)$ is dense in Y. □

A rational map $f : X \to Y$ is *birational* if there are open sets $U \subset X$ and $V \subset Y$ with U contained in the domain of X and $f(U) \subset V$, such that $f : U \to V$ is an isomorphism. The quasi-varieties X and Y are then referred to as *birationally equivalent*.

Example. *Any quasi-variety is birationally equivalent to an open subvariety of itself. In particular, $\mathbf{A}^n$ and $\mathbf{P}^n$ are birationally equivalent. Similarily, $\mathbf{P}^n \times \mathbf{P}^m$ is birationally equivalent to $\mathbf{P}^{n+m}$. Note that $\mathbf{P}^n \times \mathbf{P}^m$ is certainly not isomorphic to $\mathbf{P}^{n+m}$. For instance, any two curves in $\mathbf{P}^2$ must intersect, whereas $\mathbf{P}^1 \times \mathbf{P}^1$ contains curves that do not intersect.*

Theorem 2.24. *Let X and Y be irreducible quasivarieties. Then X and Y are birationally equivalent if and only if $k(X)$ is isomorphic to $k(Y)$.*

Proof. Suppose U and V are open subsets of X and Y and $f : U \to V$ is an isomorphism. Then the induced map $f^* : k[V] \to k[U]$ is an isomorphism, and thus so too is the map from $k(V)$ to $k(U)$. But $k(V)$ is isomorphic to $k(Y)$, and $k(U)$ is isomorphic to $k(X)$, so we conclude that $k(X)$ and $k(Y)$ are isomorphic.

Conversely, suppose $k(X)$ and $k(Y)$ are isomorphic. If $\phi : k(Y) \to k(X)$ is an isomorphism, it is induced by a unique rational map f from X to Y. Now ϕ restricts to a morphism from $k[Y]$ to $k[U_b]$ for some basic open set U_b in X, and conversely, ϕ^{-1} restricts to a morphism from $k[X]$ to $k[V_c]$ for some basic open set V_c in Y. The homomorphism from $k[Y]$ to

$k[U_b]$ corresponds to a morphism $f : U_b \to Y$, and the homomorphism from $k[X]$ to $k[V_c]$ corresponds to a morphism $g : V_c \to X$, and for any $p \in U_b \cap f^{-1}(V_c)$, $f(p) \in f(U_b) \cap V_c$ and $g(f(p)) = p$. Thus we obtain an isomorphism from $U_b \cap f^{-1}(V_c)$ and $f(U_b) \cap V_c$. □

Corollary 2.25. *Every irreducible curve is birationally equivalent to an irreducible planar curve.*

Proof. If C is an irreducible curve, we have seen that $k(C) = k(x,y)$ for two $x,y \in k(C)$. Let $\mathfrak{a}$ be the kernel from the map from $k[x,y]$ to $k(x,y)$. Then $\mathfrak{a}$ is prime, so $V = Z(\mathfrak{a})$ is an irreducible variety. The induced map from $k[V]$ to $k(x,y)$ is an embedding, and therefore induced an embedding of $k(V)$ in $k(x,y)$. But this implies V and C are birationally equivalent. □

A *rational curve* is a curve birationally equivalent to $\mathbf{A}^1$. More generally, a variety is *rational* if it is is birationally equivalent to $\mathbf{A}^n$ for some n. As we saw in the introduction to these notes, any conic in $\mathbf{A}^2$ is a rational curve.

Example. *Any planar cubic with a multiple point is rational. First, we note that any planar cubic with a multiple point is projectively equivalent to the curve C_1 described by $Y^2 = X^3$ or the curve C_2 described by $Y^2 = X^3 + X^2$. We claim that $k(C_i) = k(Y/X)$ for each i. Indeed, in $k(C_1)$, we find that $(Y/X)^2 = Y^2/X^2 = X^3/X^2 = X$, and $(Y/X)^3 = X(Y/X) = Y$. Thus $k(C_1) = k(X,Y)$ is contained $k(Y/X)$, which gives equality. In C_2, we have $(Y/X)^2 = X + 1$, so that $X \in k(Y/X)$, and thus $Y \in k(Y/X)$ as well, so that $K(C_2)$ is contained in $k(Y/X)$. If*

$$a_0 + a_1(Y/X) + \cdots + a_n(Y/X)^n = 0$$

in $k(C_i)$, then $a_0X^n + a_1YX^{n-1} + \cdots + a_nY^n = 0$ in $k[C_i]$, which implies that $a_0X^n + \cdots + a_nY^n$ is either divisible by $Y^2 - X^3$ or $Y^2 - X^3 - X^2$. But this is impossible since every factor of a homogeneous polynomial is homogeneous. Thus Y/X is algebraically independent in $k(C_i)$, which implies $k(C_i)$ is isomorphic to $k(X)$, which means C_i is birationally equivalent to $\mathbf{A}^1$.

2.10 Blowing Up a Point

A *simple point* on a curve C will be a point p such that $\mathcal{O}_p(C)$ is a discrete valuation domain, and are otherwise called *multiple points*. Since all curves C are birationally equivalent to a planar curve, and all planar curves have only finitely many multiple points, it follows that C has only finitely many multiple points.

Let K be a field extension of k. We say a ring A is a *local ring* of K if A is a subring of K containing k, which is local, and whose quotient field is K. If A is a discrete valuation ring we say A is a *discrete valuation ring* of K.

Theorem 2.26. *Let C be an irreducible projective curve with $K = K(C)$. Let L be any field containing K, and let R be a discrete valuation ring of L over k, where K is* not *a subset of R. Then there is a unique point $p \in C$ such that R dominates $\mathcal{O}_p(C)$.*

Proof. Let us show that if p exists, it is unique. Suppose R dominates $\mathcal{O}_p(C)$ and $\mathcal{O}_q(C)$. Find $f \in k(C)$ defined at p and q, with $f(p) = 0$ and $f(q) \neq 0$. Then $f \in \mathfrak{m}_p \subset \mathfrak{m}_R$, and $1/f \in \mathcal{O}_q(C) \subset R$. This implies f is a unit in R contained in $\mathfrak{m}_R$, which is impossible. This gives uniqueness.

To prove that p exists, suppose C is a projective variety in $\mathbf{P}^n$. By applying a projective transformation, assume without loss of generality that C intersects each finite affine plane $U_0, \dots, U_{n+1}$. Then in $k_h[C]$, $X_1, \dots, X_{n+1}$ are all non-zero. Let $N = \max(\mathrm{ord}_R(x_i/x_j))$. Without loss of generality, permuting variables if necessary, assume that $N = \mathrm{ord}_R(x_n/x_{n+1})$. Then for each i,

$$\mathrm{ord}_R(x_i/x_{n+1}) = N - \mathrm{ord}_R(x_n/x_i) \geqslant 0.$$

Thus $x_i/x_{n+1} \in R$ for each i. If C_0 is the affine variety of finite points in C, then we have seen that the ring $k[C_0]$ of regular functions on C_0 can be identified with the subring $k[X_1/X_{n+1}, \dots, X_n/X_{n+1}]$. Thus $k[C_0]$ is a subring of R.

Let $\mathfrak{a} = \mathfrak{m}_R \cap k[C_0]$. Then $\mathfrak{a}$ is prime, so corresponds to a closed, irreducible subvariety V of C_0. We cannot have $\mathfrak{a} = (0)$, since then $k[C_0] \cap \mathfrak{m}_R = (0)$, so all non-zero elements of $k[C_0]$ are invertible in R, implying $K = k(C_0)$ is a subring of R, which we assumed was not the case. But this means that V is a proper irreducible subvariety, and is therefore just a single point, i.e. there is $p \in C_0$ such that $V = \{p\}$. Thus if $f \in k[C_0]$ and $f(p) = 0$, then $f \in \mathfrak{a}$, and if $f(p) \neq 0$, $f \notin \mathfrak{a}$, implying f is a unit in R. This means that R contains $\mathcal{O}_p(C_0)$, and that $\mathfrak{m}_p(C_0)$ is contained in $\mathfrak{m}_R$. Thus R dominates $\mathcal{O}_p(C_0)$. □

We will use this lemma to understand rational maps between two curves.

Lemma 2.27. *Let $f : C_1 \to C_2$ be a non-constant rational map between two irreducible curves. Then f is dominating.*

Proof. Without loss of generality, working with affine neighborhoods, we may assume f is a morphism, and that C_1 and C_2 are both affine. Thus we have a homomorphism $f^* : k[C_2] \to k[C_1]$. Irreducibility implies $k[C_1]$ is an integral domain, which implies the kernel $\mathfrak{a}$ of f^* is a prime ideal of $k[C_2]$. If $\mathfrak{a} = (0)$, then f^* is injective, implying f is dominating. On the other hand, if $\mathfrak{a}$ is nonzero, then $Z(\mathfrak{a})$ is a proper non-empty irreducible subvariety of C_2, i.e. $Z(\mathfrak{a}) = \{q\}$ for some $q \in C_2$. Thus $f^*(u) = 0$ if and only if $u(q) = 0$, which can be used quite simply to show $f(p) = q$ for all $p \in C_1$. □

Corollary 2.28. *Let $f : C_1 \to C_2$ be a rational map between two curves, where C_2 is projective. If p is a simple point on C_1, then p is in the domain of f. In particular, if C_1 is nonsingular, then f is a morphism between the two curves.*

Proof. Let U be the domain of f. Then either $f : U \to C_2$ is domainting or constant. If f is constant, then f obviously extends to a constant map on all points of C_1. If f is dominating, it induces a map $f^* : k(C_2) \to k(C_1)$. Let $K = f^*(k(C_2))$ and let $L = k(C_1)$. If p is a simple point on C_1, set $R = \mathcal{O}_p(C_1)$. Provided R does not contain K, the hypotheses of the previous lemma implies that there exists a unique $q \in C$ such that R dominates $f^*(\mathcal{O}_q(C_1))$. But we have seen this means that p is in the domain of f and $f(p) = q$. Thus all that remains is to establish that R does not contain K.

Since K and L are both fields with transcendence dimension one, and K is a subfield of L, it follows that L is algebraic over K. If R contained K, then R would be a ring between K and L, and therefore a field, which is impossible since R is local. □

Corollary 2.29. *If C_1 is a nonsingular curve, and C_2 is a projective curve, then there is a one to one correspondence between morphisms $k(C_2) \to k(C_1)$ and morphisms from C_1 to C_2. If C_1 and C_2 are both nonsingular projective curves, C_1 is isomorphic to C_2 if and only if $k(C_1)$ is isomorphic to $k(C_2)$.*

Corollary 2.30. *If C is nonsingular and projective, then the points of C are in one-to-one correspondence with the discrete valuation rings of $k(C)$.*

Proof. Setting $K = L$ in the lemma above, we see that if R is a local ring of L, then R dominates $\mathcal{O}_p(C)$ for a unique p. We claim $R = \mathcal{O}_p(C)$. If $f \in K$, then either $f \in R$, or $1/f \in \mathfrak{m}_R$, but not both. The same is true of $\mathcal{O}_p(C)$ in place of R. If R is not equal to $\mathcal{O}_p(C)$, then there is $f \in R$ which is not in $\mathcal{O}_p(C)$. This means that $1/f \in \mathfrak{m}_p$, which means $1/f \in \mathfrak{m}_R$, which gives a contradiction. Thus $R = \mathcal{O}_p(C)$. □

Since every field of finite transcendence degree over k is the function field of some irreducible variety TODO, it follows that the study of irreducible nonsingular curves can be identified with the study of a certain family of fields of transcendence degree one over k. Given such a field K, we can recover the curve C to which it corresponds by identifying points with discrete valuation rings of K. We give the set $\mathcal{C}$ of discrete valuation rings of K the cofinite topology. If $U \subset \mathcal{C}$ is open, we can then define $k[U] = \bigcap U$, i.e. the set of $x \in K$ which lie in every discrete valuation ring contained in U. Provided we can give a geometric understanding of the collection of such rings (i.e. through scheme theory), we can understand $\mathcal{C}$ as being isomorphic to the curve it represents.

Our goal now, for each irreducible curve C, is to come up with a nonsingular curve C' which is birationally equivalent to C. To do this, we take a multiple point $p \in C$ and *blow it up*, replacing this point with multiple points which are less singular than the original point p. We consider this first in the case of a single point in the affine plane.

The idea is to consider the polynomial map $\psi : \mathbf{A}^2 \to \mathbf{A}^2$ given by $\psi(x,z) = (x,xz)$. Then ψ is a birational equivalence. Indeed, if $U = \{(x,z) \in \mathbf{A}^2 : x \neq 0\}$, then ψ restricts to a isomorphism from U to itself, with inverse map $\psi(x,y) = (x,y/x)$. The idea behind this map is that if C is a curve with a multiple point at the origin, then C has branches that point in multiple directions corresponding to the different tangent lines at C. Thus as these branches approach the origin, their slopes will diverge, and so in $\psi^{-1}(C - \{0\})$, these branches will be separated, converging to different points on the line $L = \{(x,z) : x = 0\}$ so we will have 'separated' the singularities. On the other hand, $C - \{0\}$ is a Zariski open subset of C, and ψ restricts to a birational equivalence from $\psi^{-1}(C - \{0\})$. If we let C' denote the closure of this set, then C and C' are birationally equivalent. Let us study this correspondence more carefully.

Lemma 2.31. *Let $C = V(f)$, where $f(X,Y) = f_r(X,Y) + \dots + f_n(X,Y)$, where $f_i \in k[X,Y]$ is homogeneous of degree i. Then $C' = V(f')$, where $f'(X,Z) = f_r(1,Z) + Xf_{r+1}(1,Z) + \dots + X^{n-r} f_n(1,Z)$.*

Proof. The set $\psi^{-1}(C - \{0\})$ is the set of tuples (x,z) where $x \neq 0$ and $f(x,xz) = 0$. Now

$$f(x,xz) = x^r f_r(1,z) + \dots + x^n f_n(1,z).$$

Since $x \neq 0$, $\psi^{-1}(C - \{0\})$ can also be described as the set of (x,z) with $x \neq 0$ such that $f'(x,z) = f_r(1,z) + \cdots + x^{n-r} f_n(1,z)$. Thus f' lies in $I(C')$, and it suffices to show that f' is irreducible. If $f'(X,Z) = g(X,Z)h(X,Z)$ for $g,h \in k[X,Z]$, then $f(X,Y) = X^n g(X,Y/X)h(X,Y/X)$. If g has degree m and h has degree l, it follows from the irreducibility of f that either $X^m g(X,Y/X)$ or $X^l h(X,Y/X)$ is a constant. But this means that g or h is a constant, so f' is irreducible. □

Let us assume that the Y-axis is not tangent to C at the origin. Then we may write

$$f_r(X,Y) = (Y - t_1 X)^{r_1} \dots (Y - t_s X)^{r_s}.$$

Let us more rigorously establish the claim that C' separates the tangents of C.

Lemma 2.32. *We have $f^{-1}(0) = \{p_1, \dots, p_s\}$, where*

$$m_{p_i}(C') \leqslant I_{p_i}(f', X).$$

In particular, if 0 is an ordinary multiple point of C, then each point $p_1, \dots, p_s$ is a simple point on C' with $ord_{p_i}(X) = 1$.

Proof. We have

$$f^{-1}(0) = \{(0,z) : f'(0,z) = 0\} = \{(0,z) : f_r(1,z) = 0\}.$$

But

$$f_r(1,Z) = (Z - t_1)^{r_1} \dots (Z - t_s)^{r_s}$$

so $f^{-1}(0) = \{(0,t_1), \dots, (0,t_s)\} = \{p_1, \dots, p_s\}$. Now

$$m_{p_i}(C') \leqslant I_{p_i}(f', X) = I_{p_i}(f_r, X) = I_{p_i}((Z - t_i)^{r_i}, X) \leqslant r_i. \qquad □$$

Theorem 2.33. *There is an affine neighborhood W of 0 on C such that $W' = \psi^{-1}(W) \cap C'$ is an affine open subvariety of C', $\psi(W') = W$, and $k[W']$ is a finitely generated $k[W]$ module, with $X^{r-1} k[W'] \subset \psi^*(k[W])$.*

Proof. Write $f(X,Y) = \sum_{i+j \geqslant r} a_{ij} X^i Y^j$ and let $h(Y) = \sum_{j \geqslant r} a_{0j} Y^{j-r}$. Then $h(0) = a_{0r} \neq 0$ since X is not a tangent line for C, so the basic open set W in C corresponding to h contains 0. Then $W' = \psi^{-1}(W_h)$ is the basic open set corresponding to $\tilde{h}$, where $\tilde{h}(X,Z) = h(XZ) = \sum a_{0j} X^{j-r} Z^{j-r}$. To see that $\psi(W') = W$, it suffices to show that the resulting map $\psi^* : k[W] \to k[W']$ is injective. If $u = g/h^n$ is an element of $k[W]$, then $f^*(u)$ is the polynomial $g(X, XZ)/h(XZ)^n$. If this polynomial vanishes on W', then it vanishes on C', so that $g(X, XZ)$ is divisible by f', i.e. $g(X, XZ) = k_1(X, Z)f'(X, Z)$. But then

$$g(X, Y) = k_1(X, Y/X)f'(X, Y/X) = k_1(X, Y/X)f(X, Y)/X^r$$

which implies g vanishes on C.

Since $k[W']$ is generated by X and Z, to show it is finitely generated over $k[W]$ it suffices to find $b_1, \dots, b_r \in k[W]$ such that

$$Z^r + b_1(X, XZ)Z^{r-1} + \cdots + b_r(X, XZ) = 0$$

in $k[W']$. Now

$$f'(X, Z) = \sum a_{ij} X^{i+j-r} Z^j = \sum a_{ij} \psi^*(Y^{i+j-r}) Z^{r-i} = 0$$

on $k[W']$, which implies that we have an equation of the above sort if we set $b_i(Y) = \sum a_{ij} Y^{i+j-r}/h(Y)$ for $i < r$, and set $b_r(Y) = \sum_{i \geqslant r} a_{ij} X^{i-r} Y^j / h$. Thus $k[W']$ is generated by $\{1, \dots, Z^{r-1}\}$ as a module over $K[W]$. For each $0 \leqslant i \leqslant r-1$, $X^{r-1}Z^i = \psi^*(Y^i X^{r-1-i}) \in \psi^*(k[W])$, which gives the required inclusion. □

Chapter 3

Algebraic Curves

We now apply the algebraic tools we have developed to the study of planar algebraic curves. This is one of the classical areas of algebraic geometry, which is still a wide source of research material today. We already know that the interesting varieties in the plane are those specified as the nullset of a single polynomial f, or the nullset of a principal ideal. The varieties corresponding to nonprincipal ideals consist of finitely many points. In this situation it is natural to want to determine exactly *how many* points are in this variety, especially if we are considering $Z(f,g)$, which is the topic of the field of *intersection theory* (i.e. the intersection theory of two curves), the most basic concept in the theory of curves.

In the scenario of intersection theory, it is naturally to consider a more general family of geometric objects not necessarily corresponding to an algebraic set. Even if $f = f_1^{n_1} \dots f_m^{n_m}$ has the same locus at $g = f_1 \dots f_m$, we would like to think of f as defining a different algebraic curve to f than g, one which has 'n_1 copies' of the irreducible planar curve defined by f_1, 'n_2 copies' of f_2, and so on. Thus in this chapter we define an *affine planar algebraic curve* C to be a principal ideal $\mathfrak{a}$ in $k[X,Y]$ not equal to (0). If $\mathfrak{a} = (f)$, we say that C is *defined* by f. The geometric correspondence is slightly lost by this generalization, but can be recovered by appealing to the more modern machinery of scheme theory, which is a topic for another time. As for varieties, we introduce the coordinate ring $k[C] = k[X_1,\dots,X_n]/\mathfrak{a}$, $I(C) = \mathfrak{a}$, $Z(C) = Z(\mathfrak{a})$, and $\mathcal{O}_p(C)$ the localization of $k[C]$ by the set $S_p(C)$ of functions $f \in k[C]$ with $f(p) \neq 0$.

3.1 Differentials

Let f be a polynomial defining a planar curve C. We say a point $p \in C$ is a *simple point* of f if $f_X(p)$ and $f_Y(p)$ are not both zero. In this case, we can define the tangent line T_pC by the equation $f_X(p)(X-a) + f_Y(p)(Y-b) = 0$. Over the real numbers, we can locally parameterize C by a smooth map around any simple point. Over other fields, we have a more modest equivalence.

Theorem 3.1. *Suppose p is a simple point on a curve C. Then $\mathcal{O}_p(C)$ is a discrete valuation ring.*

Proof. Let C be defined by $f \in k[X,Y]$. By symmetry, without loss of generality we may assume that $p = 0$, and that T_pC is the y axis, i.e. that $f_X(0) = 1$ and that $f_Y(0) = 0$. Write $f = Xf_1 - Yf_2$ for $f_1, f_2 \in k[X,Y]$, where $f_1(0) = 1$. In the coordinate ring $k[C]$, $Xf_1 = Yf_2$, and since $f_1(0) \neq 0$, in $\mathcal{O}_0(C)$ we can write $X = Yf_2/f_1$. A consequence of this is that $\mathfrak{m}_0 = (X,Y) = (Y)$ is a principal ideal. Thus all that remains to show that $\mathcal{O}_0(C)$ is a discrete valuation ring is to show that $\mathcal{O}_0(C)$ is a domain. Since 0 is a simple point of C, we can write $f = g_1 g_2^{n_2} \dots g_m^{n_m}$, where each g_i is irreducible, $g_1(0) = 0$, and $g_2(0), \dots, g_m(0) \neq 0$. Since localizations commute with quotients, $\mathcal{O}_0(C)$ is obtained by taking the quotient of $S_p^{-1}(\mathbf{A}^n)$ by $S_p^{-1}(f) = (g_1)$. This ideal is prime in $S_p^{-1}(\mathbf{A}^n)$ since g_1 is irreducible, so we conclude $\mathcal{O}_0(C)$ is a Noetherian integral domain (essentially, we've argued that the local ring is isomorphic to the local ring of the unique variety containing the point, which makes sense since the space can only model local information about p). □

Remark. More general, this argument shows that if the line described by the linear equation $aX + bY$ is not tangent to a curve C at a simple point p, then $aX + bY$ acts as a uniformizing parameters for $\mathcal{O}_p(C)$.

The converse to this statement is also true (that p is simple if $\mathcal{O}_p(C)$ is a discrete valuation domain), which we prove shortly in a more general scenario.

If p is a simple point on a curve C, we write ord_p for the order function at p over $\mathcal{O}_p(C)$, and if our curve is irreducible, the function over $k(C)$. If the curve isn't clear, we denote the order function by ord_p^C.

Example. *If $p_1, \dots, p_m$ are simple points on a curve C, and the m_i are nonnegative integers, then we can find a function $f \in k[C]$ with $ord_p(f) = m_i$. We just take $f = (a_1X + b_1Y)^{n_1} \dots (a_mX + b_mY)^{n_m}$, $a_iX + b_iY$ is a line not tangent to f at p, but passing through p, and not passing though any other point.*

Example. *A simple point p is called a* flex *if the equation describing its tangent line has order greater than 2 at p. We say the flex is* ordinary *if the order of the tangent is exactly three, and a* higher flex *otherwise. The curve $f(X,Y) = Y - X^n$ has a tangent line $Y = 0$ at the origin, and since $f_Y(0) = 1 \neq 0$, X is a local parameter in $\mathcal{O}_0(f)$. Since $Y = X^n$, the order of the tangent line is n, so the origin is a flex for $n \geqslant 3$, and an ordinary flex for $n = 3$.*

In general, if we take a curve whose tangent at the origin (which is a simple point of the curve) is the X axis, we may write the equation defining the curve as $Y = aX^2 + X^3g + Yh$, where $h(0) = 0$. This implies that in $\mathcal{O}_0(C)$,

$$Y = \frac{aX^2 + X^3g}{1-h} = X^2 \cdot \frac{a + Xg}{1-h}$$

We split our analysis into three cases:

- *If $a \neq 0$, then $ord_0(Y) = 2$ since $(a + Xg)/(1-h)$ is a unit in $\mathcal{O}_0(C)$.*
- *If $a = 0$ and $g = 0$, then the equation describing C is given by $Y(1-h) = 0$, which implies $Y = 0$ in $\mathcal{O}_0(C)$. Thus $ord_0(Y) = \infty$.*
- *If $a = 0$ and $g \neq 0$, we may write*

  $$Y = X^3 \cdot \frac{g}{1-h}.$$

  Thus $ord_0(Y) = 3 + ord_0(g) \geqslant 3$.

In particular, the origin is a flex point if and only if $a = 0$.

Since $\mathcal{O}_p(C)$ contains a copy of the field isomorphic to the field obtained by quotienting by a maximal idea, this ring is precisely the type of ring where functions can be uniquely expanded in power series in X over k. This corresponds in some sense to the fact that Y is 'parameterizable' in terms of X, because we can rewrite arbitrary functions $g \in \mathcal{O}_p(f)$, which we think of as two variables, as one dimensional power series in X. On the other hand, if $f_X(p) = f_Y(p) = 0$, we know from calculus that there is no hope of parameterizing the curve in terms of X and Y, and p is known as a *multiple* or *singular point*. A curve with no singular points is called a *nonsingular curve*.

Example. *The polynomial $f = Y - X^2$ defines a nonsingular curve, for $f_Y = 1$ is constant, and therefore never zero. The only time $f_X = 0$ is at the origin, in which case we cannot parameterize the function in terms of Y because the function branches to the left and right.*

Example. *If $f(X,Y) = Y^2 - X^3 + X$ is a polynomial over a field not of characteristic 2 or 3, then $f_Y = 2Y$ vanishes for $Y = 0$, and $f_X = 1 - 3X^2$ vanishes for $X = \pm\sqrt{1/3}$, yet the points $(\sqrt{1/3}, 0)$ and $(-\sqrt{1/3}, 0)$ are not on $Z(f)$, so the curve itself is nonsingular. If we are working over a field of characteristic 3, then $f_X = 1$ never vanishes, so the curve is nonsingular. On the other hand, the polynomial can be singular over a field of characteristic 2, because $f_Y = 0$, and $f_X = 1 + X^2 = (1+X)^2$ vanishes for $X = 1$, so the curve has a single singularity at $(1,0)$.*

Example. *The polynomial $f(X,Y) = Y^2 - X^3$ has a single singularity over a field of any characteristic, for $f_X = -3X^2$ vanishes for $X = 0$, and $f_Y = 2Y$ vanishes for $Y = 0$, and since $(0,0)$ lies on the curve this is where the polynomial has a singularity.*

Example. *The polynomial $f(X,Y) = Y^2 - X^3 - X^2$ has a single singularity. The derivative $f_Y = 2Y$ vanishes for $Y = 0$, and $f_X = -3X^2 - 2X$ vanishes for $X = 0$ and $X = -2/3$, yet only the point $(0,0)$ lies on $Z(f)$ and has all derivatives of the polynomial vanishing. Over a field of characteristic 3, $f_X = X$ vanishes only for $X = 0$, so there is only a single singularity, and over a field of characteristic 2, $f_X = X^2$ vanishes for $X = 0$, and there is only a single point on $Z(f)$ whose X coordinate is equal to zero, so $(0,0)$ is the only singularity.*

Example. *The polynomial $f(X,Y) = (X^2 + Y^2)^2 + 3X^2Y - Y^3$ has*

$$f_X = 4X(X^2 + Y^2) + 6XY = X(4X^2 + 4Y^2 + 6Y)$$

and

$$f_Y = 4Y(X^2 + Y^2) + 3X^2 - 3Y^2$$

If $X = 0$, then $f_Y = 4Y^3 - 3Y^2$ vanishes only for $Y = 0$ and $Y = 3/4$, and only the point $(0,0)$ lies on $Z(f)$, so this is a singularity point. Otherwise, the only reason f_X vanishes is if $4X^2 + 4Y^2 + 6Y = 0$, so $f_Y = -(3/2)(4Y^2 + 4Y + 3)$. If this vanishes also, then $f(X,Y) = Y/4 - 21/16$, which can only vanish for $Y = 21/4$, which doesn't satisfy $4Y^2 + 4Y + 3 = 0$, so there are no other singularities.

Example. *For the polynomial $f = (X^2 + Y^2)^3 - 4X^2Y^2$, we find*

$$f_X = 6X(X^2 + Y^2)^2 - 8XY^2 \qquad f_Y = 6Y(X^2 + Y^2)^2 - 8X^2Y$$

f_X vanishes for $X = 0$, or $6(X^2 + Y^2)^2 = 8Y^2$, and f_Y vanishes for $Y = 0$ and $6(X^2 + Y^2)^2 = 8X^2$. If both X and Y are nonzero, then we conclude $Y^2 = X^2$, and this implies $f = 8X^2 - 4X^4$ can only vanish for $X = \pm\sqrt{2}$, but in a similar fashion we find that f_X can only vanish for $X = \pm\sqrt{1/3}$.

Unless we want to restrict ourselves to nonsingular curves, singularity points are a natural part of a study of algebraic curves, and we have to face them head on. The trick is to note that, though singular points will not necessarily have a unique tangent line, there are still lines through the point which behave have tangent lines 'should' behave, and in the singular case we may end up with multiple tangent lines. First, note that if $f(X,Y)$ is a polynomial, the local behavior of the polynomial around the origin is determined solely by the lowest order terms. If we collect the lowest degree terms into a homgeneous polynomial, the zeroes of this polynomial will form a union of lines, which we can take as a family of tangents to f at the origin.

Lemma 3.2. *Over an algebraically closed field, every homogeneous polynomial decomposes into a product of linear forms.*

Proof. Given a homogeneous polynomial $f \in K[X,Y]$ of degree m, consider the polynomial $f(X,1) \in k[X]$, which has degree at most m. Since k is algebraically closed, we may write

$$f(X,1) = C \cdot (X - a_1)^{n_1} \dots (X - a_l)^{n_l}$$

for some $a_1, \dots, a_l \in k$. But then working in $K(X,Y)$, we conclude that

$$\begin{aligned} f(X,Y) &= C \cdot Y^m f(X/Y, 1) \\ &= C \cdot Y^m (X/Y - a_1)^{n_1} \dots (X/Y - a_l)^{n_l} \\ &= C \cdot Y^{m - n_1 - \dots - n_l} (X - a_1 Y)^{n_1} \dots (X - a_l Y)^{n_l} \end{aligned}$$

and the same identity continues to hold in $k[X,Y]$. □

Thus we have a way of constructing m tangent lines at a point of degree m on a curve. The point is called *ordinary* if all the tangent lines are distinct. A *double point* is a point of degree two, a *triple point* is a point of degree three, and so on and so forth. A *node* is a point is an ordinary point of degree two. By applying a linear transformation to move a point to the origin we can construct tangent lines at any point.

Before we move on, it is nice to note that on an irreducible curve, even if singularities occur, there can only be finitely many. This is because if f is an irreducible polynomial, then $Z(f, f_X, f_Y)$ contains only finitely many points.

Proposition 3.3. *If $T : \mathbf{A}^2 \to \mathbf{A}^2$ is a polynomial map with $T(p) = q$, then for any polynomial $g \in k[X_1, \dots, X_m]$, $m_q(g) \leqslant m_p(f^* g)$. If the two by two Jacobian matrix $(\partial T^i / \partial X^j)$ is invertible at p, then $m_q(g) = m_p(f^* g)$.*

Proof. Since multiplicity is preserved under translation, we may assume p and q both lie at the origin. Then we may write

$$T(X,Y) = (aX + bY + T_0, cX + dY + T_1)$$

for some polynomials T_0, T_1 with no monomial terms of degree less than two. If $g = g_0 + \dots + g_n$, then $f^* g = f^* g_0 + \dots + f^* g_n$, and if $g_k = \prod h_i^{k_i}$, then

$$\begin{aligned}(f^* g_k)(X,Y) &= g_k(aX + bY + T_0, cX + dY + T_1)\\ &= \prod h_i^{k_i}(aX + bY + T_0, cX + dY + T_1)\end{aligned}$$

If $h_i(X,Y) = \lambda_i X + \gamma_i Y$, then computing modulo terms of order $k+1$ or greater, we find that

$$(f^* g_k)(X,Y) \equiv \prod [(\lambda_i a + \gamma_i c)X + (\lambda_i b + \gamma_i d)Y]^{k_i}$$

If this value vanishes, then in particular $\prod (\lambda_i a + \gamma_i c)^{k_i} = 0$, which implies that (a, c) lies on the line h_i for some i. Looking at the expansion of $X^{k-k_i} Y^{k_i}$, we conclude that

$$\prod_{j \neq i} (\lambda_j a + \gamma_j c)^{k_j} (\lambda_i b + \gamma_i d)^{k_i} = 0$$

so either $\lambda_i b + \gamma_i d$, implying that (a, c) lies on the same line through the origin as (b, d), or (a, c) lies on more than one line through the origin, implying $a = c = 0$. In both cases, $ad - bc = 0$. □

3.2 Singularity Analysis Through Local Rings

The multiplicity $m_p(C)$ of a point p on a curve C should clearly be a local property. Since we have built up a heuristic that the algebraic structure of $\mathcal{O}_p(C)$ gives all local geometric information about the curve C at the point, we should be able to tease out the multiplicity of a singularity purely through an analysis of $\mathcal{O}_p(C)$, which we shall now see.

Proposition 3.4. *Let p be a point on a curve C. Then, for sufficiently large n,*

$$m_p(C) = \dim(\mathfrak{m}_p(C)^n / \mathfrak{m}_p(C)^{n+1})$$

In particular, $\mathcal{O}_p(C)$ uniquely determines the multiplicity $m_p(C)$, and as such is preserved by isomorphisms of curves.

Proof. Write $\mathfrak{m} = \mathfrak{m}_p(C)$, and $\mathcal{O} = \mathcal{O}_p(C)$. We have an exact sequence

$$0 \to \mathfrak{m}^n / \mathfrak{m}^{n+1} \to \mathcal{O}/\mathfrak{m}^{n+1} \to \mathcal{O}/\mathfrak{m}^n \to 0$$

So it suffices to show that for sufficiently large n,

$$\dim \mathcal{O}/\mathfrak{m}^{n+1} - \dim \mathcal{O}/\mathfrak{m}^n = m_p(C)$$

In fact, we will show that there is a constant s such that $\dim \mathcal{O}/\mathfrak{m}^n = n m_p(C) + s$ for all $n \geqslant m_p(C)$. Assume that $p = 0$. Then $\mathfrak{m} = (X, Y)$. But

$$\mathcal{O}/\mathfrak{m}^n \cong \mathcal{O}_p(\mathbf{A}^n)/(f, \mathfrak{m}^n) \cong k[X,Y]/(f, \mathfrak{m}^n)$$

where the last isomorphism follows because $Z(f, \mathfrak{m}^n) = \{p\}$, as we proved in the last chapter. Thus we need only calculate the dimension of the last ring. We find that we have a short exact sequence

$$0 \to k[X,Y]/\mathfrak{m}^{n-m} \to k[X,Y]/\mathfrak{m}^n \to k[X,Y]/(\mathfrak{m}^n, f) \to 0$$

where the first map is the linear map $g \mapsto fg$, not a ring homomorphism, and the second map is the natural one. Since as a vector space $k[X,Y]/\mathfrak{m}^k$ is isomorphic to the set of polynomials of degree less than k, we find that the space has dimension

$$\sum_{i=0}^{k-1} (i+1) = \frac{k(k+1)}{2}$$

so the dimension of $k[X,Y]/\mathfrak{m}^n$ is

$$\frac{n(n+1)}{2} - \frac{(n-m)(n-m+1)}{2} = nm - \frac{m(m+1)}{2}$$

and this was the required result. $\square$

It is interesting to calculate the dimension of $\mathfrak{m}^n/\mathfrak{m}^{n+1}$ even for $n < m_p(C)$, just to analyze this process in full. In this case, the map

$$k[X,Y]/\mathfrak{m}^n \to k[X,Y]/(\mathfrak{m}^n, f)$$

is an isomorphism, because $f \in \mathfrak{m}^n$, which implies the dimension of $\mathcal{O}/\mathfrak{m}^n$ is $n(n+1)/2$. This implies that

$$\dim \mathfrak{m}^n/\mathfrak{m}^{n+1} = \dim \mathcal{O}/\mathfrak{m}^{n+1} - \dim \mathcal{O}/\mathfrak{m}^n = \frac{(n+1)(n+2)}{2} - \frac{n(n+1)}{2} = n+1$$

In particular, we find that a point p is simple if and only if $\dim(\mathfrak{m}/\mathfrak{m}^2) = 1$, for otherwise $\dim_k(\mathfrak{m}/\mathfrak{m}^2) = 2$.

Corollary 3.5. *A point p on a planar curve C is simple if and only if $\mathcal{O}_p(C)$.*

Proof. We have already proved $\mathcal{O}_p(C)$ is a discrete valuation domain if p is simple. Conversely, if $\mathcal{O}_p(C)$ is a discrete valuation domain, then $\dim_k(\mathfrak{m}/\mathfrak{m}^2) = 1$, which from the last discussion shows that p is a simple point on a curve. $\square$

We note that the function $\chi(n) = \dim_k(\mathcal{O}/\mathfrak{m}^n)$, which is a polynomial in n for large values of n, is called the *Hilbert-Samuel polynomial*. It plays an important role in the modern study of multiplicities of local rings. Here we see how the results above interact with more general local rings than algebraic curves.

Example. *Over the ring $\mathcal{O}_p(\mathbf{A}^m)$ (where we assume without loss of generality $p = 0$), we find $\mathcal{O}_p(\mathbf{A}^m)/\mathfrak{m}^n$ is isomorphic to the vector space of polynomials of degree less than n over $k[X_1,\dots,X_m]$, because in $\mathcal{O}_p(\mathbf{A}^m)/\mathfrak{m}^n$ we find that if f has no constant coefficient, then $(1+f)^{-1} = \sum(-1)^k f^k$, where the series is finite modulo $\mathfrak{m}^n$ since $f^k = 0$ for $k \geqslant n$. This implies we can move denominators into numerators in $\mathcal{O}_p(\mathbf{A}^m)/\mathfrak{m}^n$, so this ring is isomorphic to $k[X_1,\dots,X_n]/\mathfrak{m}^n$*

(provided $I(\mathbf{A}^m) = (0)$), and this ring has dimension equal to the number of monomials of degree less than n, which we can write as the formula

$$\chi_m(n) = \sum_{k=0}^{n-1} \#\{(r_1,\dots,r_m) : \sum r_i = k\} = \sum_{k=0}^{n-1} g(m,k)$$

We prove that this is a polynomial in n by induction on m. To see this, we apply the recurrence relation, then $g(m,n) = \sum_{k=0}^n g(m-1,n-k) = \sum_{k=0}^n g(m-1,k)$, and $g(1,n) = 1$. This implies that

$$\chi_m(n) = \sum_{k=0}^{n-1} g(m,k) = \sum_{k=0}^{n-1}\sum_{l=0}^{k} g(m-1,l) = \sum_{k=1}^{n} \chi_{m-1}(k)$$

For $m = 1$, we find $\chi(n) = n$, which is a polynomial. If $\chi_{m-1}(n) = \sum a_i n^i$, then we find

$$\chi_m(n) = \sum_{k=1}^{n} \sum a_i k^i = \sum a_i \sum_{k=1}^{n} k^i$$

Since $\sum k^i$ is a polynomial in n for each i, we find that χ_m is also a polynomial in n. We also claim the leading coefficient of these polynomials is $1/m!$, which has degree m. This is true for $\chi_1(n) = n$, and if the statement is true for χ_{m-1}, the leading coefficient for χ_m is the highest order coefficient in

$$\frac{1}{(m-1)!}\sum_{k=1}^{n} k^{m-1} = \frac{1}{(m-1)!}\left(\frac{n^m}{m} + O(n^{m-1})\right) = \frac{n^m}{m!} + O(n^{m-1})$$

Example. *Consider a polynomial $f \in k[X_1,\dots,X_d]$ defining a hypersurface H. By writing $f = f_m + \dots$, where f_m is nonzero, we can generalize the notion of multiplicity to such hypersurfaces by letting $m_0(H) = m$. In this case, we find that the Hilbert polynomial $\chi(n)$ is a polynomial of degree $m-1$ for sufficiently large n, with leading coefficient $m_p(f)/(n-1)!$. To see this, we first apply the isomorphism of $\mathcal{O}_p(f)/\mathfrak{m}^n$ with $k[X_1,\dots,X_d]/(\mathfrak{m}^n, f)$. Then, if $n \geqslant m$, we can consider the analogous exact sequence*

$$0 \to k[X_1,\dots,X_d]/\mathfrak{m}^{n-m} \to k[X_1,\dots,X_d]/\mathfrak{m}^n \to k[X_1,\dots,X_d]/(\mathfrak{m}^n, f) \to 0$$

to the one considered in the case of curves. Thus if we write

$$\chi_d(n) = n^d/d! + an^{d-1} + O(n^{d-2})$$

then the dimension of the right hand ring, which is $\chi(n)$, is equal to

$$\begin{aligned}\dim_k(k[X_1,\dots,X_m]/\mathfrak{m}^n) &- \dim_k(k[X_1,\dots,X_m]/\mathfrak{m}^{n-m})\\ &= \chi_d(n) - \chi_d(n-m)\\ &= n^d/d! - (n-m)^d/d! + an^{d-1} - a(n-m)^{d-1} + O(n^{d-2})\\ &= \frac{m}{(d-1)!}\cdot n^{d-1} + O(n^{m-2}).\end{aligned}$$

Thus we conclude that the leading coefficient of the Hilbert polynomial, multiplied by $(d-1)!$, is equal to $m_p(H)$.

It should be clear for this that an analysis of the Hilbert polynomial defined about is key to a more in depth analysis of the multiplicities of singularities. If we have a 'dimension r' variety V containing a point p, then we might expect that the leading coefficient of the Hilbert polynomial at p, multiplied by $r!$, is the multiplicity of the point p on V. Let us consider an example where we can reason out the dimension of a variety by intuition.

Example. *This property does not hold for curves in higher dimensional space. Let us look at the variety V specified as the locus of $X^2 - Y^3$ and $Y^2 - Z^3$. The map $X \mapsto t^9$, $Y \mapsto t^6$, $Z \mapsto t^4$ from $k[X,Y,Z]$ to $k[t]$ has kernel $(X^2 - Y^3, Y^2 - Z^3)$, and since $k[t]$ is an integral domain we conclude that $(X^2 - Y^3, Y^2 - Z^3)$ is a prime ideal. For an arbitrary n, the ideal $\mathfrak{m}^n$ is generated by monomials of the form $X^iY^jZ^k$, where $i+j+k=n$. However, if $i > 1$, then $X^iY^jZ^k = X^{i-2}Y^{j+3}Z^k \in \mathfrak{m}^{n+1}$, and if $j > 1$, then $X^iY^jZ^k = X^iY^{j-2}Z^{k+3} \in \mathfrak{m}^{n+1}$. Thus the $\mathcal{O}_0(C)$ module $\mathfrak{m}^n/\mathfrak{m}^{n+1}$ is generated by monomials of the form $X^iY^jZ^k$, where $i \leqslant 1$ and $j \leqslant 1$, and $i+j+k=n$. But there are only four such monomials for $n \geqslant 3$, namely XYZ^{n-2}, XZ^{n-1}, YZ^{n-1}, and Z^n, and these are linearly independent over k. Thus for $n \geqslant 3$,*

$$\dim_k(\mathfrak{m}^n/\mathfrak{m}^{n+1}) = 4.$$

Thus we might think of V as having a singularity of order 4 at the origin.

3.3 Intersection Numbers

Analogous to the multiplicity of a point on a curve is the multiplicity of an intersection between two curves at a point. We will build up the definition of multiplicity axiomatically, detailing properties that the concept should follow. Then we will show that such properties uniquely define the definition, and it remains to construct a function which has these properties. Given two curves C_0 and C_1, we denote the *intersection number* between them at a point p, as the value $I_p(C_0, C_1)$, or $I_p(f, g)$ if C_0 is defined by a polynomial f, and C_1 by a polynomial g. First, the intersection number should satisfy two properties that should be obvious:

- $I_p(C_0, C_1) = I_p(C_1, C_0)$.
- The intersection number is 'multiplicatively additive' on the space of algebraic curves, with $I_p(f_0 f_1, g) = I_p(f_0, g) + I_p(f_1, g)$.
- $I_p(C_0, C_1) = 0$ if and only if $p \notin C_0 \cap C_1$.
- The intersection number is invariant of affine coordinates.

We say two curves C_0 and C_1 *intersect properly* at p if C_0 and C_1 have no common component containing p. The second property says that the intersection number explodes if C_0 and C_1 overlap around p.

- If C_0 and C_1 intersect properly at p, $I_p(C_0, C_1)$ is a non negative integer, and $I_p(C_0, C_1) = \infty$ if C_0 and C_1 don't intersect properly at p.

These two properties are consistent, because if $p \notin C_0 \cap C_1$, then C_0 and C_1 cannot contain a common component at p. Two curves C_0 and C_1 *intersect transversally* at p if p is a simple point on C_0 and C_1, and if the tangent line to C_0 at p is different to the tangent line to C_1 at p. This is a formal way of saying the two curves 'intersect once'. More generally, we require that

- $I_p(C_0, C_1) \geqslant m_p(C_0) m_p(C_1)$, with equality occurring if and only if C_0 and C_1 have no tangent lines in common at p.

The last property is least intuitive. It says that one should be able to determine the multiplicity of intersection of two curves C_0 and C_1 is determined by how C_0 'looks like' in C_1. In other words, the intersection number should be determined by the image of the polynomial f which defines C_0 in $k[C_1]$.

- For any f, g, k, $I_p(f, g) = I_p(f, g + fk)$.

In other words, $I_p(f,g)$ is well defined when we interpret g as an element of $k[X,Y]/(f)$. Surprisingly, these properties uniquely define the intersection number.

Theorem 3.6. *$I_p(f,g)$ is uniquely determined by the given properties.*

Proof. We will give a constructive procedure for calculating $I_p(f,g)$ from the properties above. Because $I_p(f,g)$ is invariant to affine translations, we may assume p lies at the origin, and that $I_p(f,g)$ is finite. We may then proceed by induction on the value of $I_p(f,g)$. Since $I_p(f,g) = 0$ holds if and only if p does not lie on $Z(f) \cap Z(g)$, this quantity is certainly uniquely determined in this case. Now we prove by induction on the value $I_p(f,g)$. So assume that the intersection numbers are uniquely determined for pairs of polynomials with multiplicity is less than n, annd suppose $I_p(f,g) = n$. Suppose $\deg(f(X,0)) = r$ and $\deg(g(X,0)) = s$. Without loss of generality, we may assume $r \leqslant s$. If $f(X,0)$ is constant, then $f(X,Y) = Yf_0(X,Y)$, then $I_p(f,g) = I_p(Y,g) + I_p(f_0,g)$, and we can now apply induction. On the other hand, if $r > 0$, then we may assume without loss of generality that $f(X,0)$ and $g(X,0)$ are monic. Let $h(X,Y) = g(X,Y) - X^{s-r}f(X,Y)$. Then $\deg(h(X,0)) < \deg(g(X,0))$, and $I_p(f,g) = I_p(f,h)$. Thus repeating this process ala the Euclidean algorithm, we may decrease the required values until we get $r = 0$, in which case the value is obvious, which gives an explicit algorithm for computing the intersection number using the properties above. □

The proof of uniqueness shows that it is very easy to compute the intersection number of two polynomials at the origin. In fact, there is an algorithm that runs essentially in time proportional to the sum of the degrees of the polynomials. Another fact about the proof of uniqueness is that is shows that some of the axioms are essentially redundant. We only need to show that $I_0(X,Y) = 1$, rather then the more general bound $I_p(f,g) \geqslant m_p(f)m_p(g)$. To prove that the algorithm is 'formally' correct, we must prove that a function $I_p(f,g)$ exists with the required properties. We formally define the *intersection number* of two algebraic curves f and g to be the dimension of the localization of the ring $k[X,Y]/(f,g)$ at p, as a vector space over k.

Theorem 3.7. *Setting*

$$I_p(f,g) = \dim_k \left(\mathcal{O}_p(\mathbf{A}^2)/(f,g)\right)$$

satisfies the required properties of an intersection number.

Proof. We shall write $\mathcal{O}$ for $\mathcal{O}_p(\mathbf{A}^2)$. We will repeatedly use the fact that, since localization commutes with quotients, the localization of $k[X,Y]/(f,g)$ at p is isomorphic to $\mathcal{O}/(f,g)$. Let us verify the properties one by one:

- It is clear that $I_p(f,g) = I_p(g,f)$, since $(f,g) = (g,f)$.
- It is clear that $I_p(f,g) = I_p(f,g+kf)$, since $(f,g) = (f,g+kf)$.
- It is clear $I_p(f,g)$ is invariant under affine changes of coordinates, since $\mathcal{O}$ is preserved under isomorphisms of $\mathbf{A}^2$.
- If p isn't in $Z(f) \cap Z(g)$, then (f,g) contains a unit in $\mathcal{O}$, and consequently, $\mathcal{O}/(f,g) = (0)$. Thus this space is dimension zero, so $I_p(f,g) = 0$. Conversely, the space $\mathcal{O}/(f,g)$ can only have dimension zero if (f,g) contains a unit, which implies $f(p) \neq 0$ and $g(p) \neq 0$.
- If f and g contain some common irreducilbe component h vanishing at p, then $(f,g) \subset (h)$, so we have a surjective map from $\mathcal{O}/(f,g)$ to $\mathcal{O}/(h)$ at p, which is infinite dimensional because $\mathcal{O}_p(Z(h))$ contains $k[Z(h)]$ as a subring. But this means that $I_p(f,g) = \dim_k(\mathcal{O}/(f,g)) = \infty$.

  Conversely, if f and g contain no irreducible component, then we have seen that $Z(f,g)$ is finite, which implies that $\mathcal{O}/(f,g)$ is finite dimensional.
- Let us now prove that $I_p(f_1f_2,g) = I_p(f_1,g) + I_p(f_2,g)$. Without loss of generality, we may assume that g shares no components with f_1 nor f_2. We then consider an exact sequence of the form

  $$0 \to \mathcal{O}/(f_1,g) \to \mathcal{O}/(f_1f_2,g) \to \mathcal{O}/(f_2,g) \to 0$$

  where the first map is given by $k \mapsto kf_2$, and the second as the natural quotient. It is obvious that this second map is surjective; to see that the first map is injective, we suppose $kf_2 = h_1f_1f_2 + h_2g$ for two polynomials h_1 and h_2. Then $(k - h_1f_1)f_2 = h_2g$. Since g shares no components with f_2, this means h_2 is divisible by f_2, so we can write $h_2 = f_2h_2'g$. Thus $(k - h_1f_1 - h_2'g)f_2 = 0$, hence $k = h_1f_1 + h_2'g$ since $f_2 \neq 0$. But this means we gave orived that if $kf_2 \in (f_1f_2,g)$, then

$k \in (f_1, g)$. Thus the first map is injective. The exactness on the interior of the diagram is obvious. But then the isomorphism theorem gives the additive property.

- To prove that $I_p(f,g) \geqslant m_p(f)m_p(g)$, assume $p = 0$, write $m = m_p(f)$ and $n = m_p(g)$. Let $\mathfrak{a}$ denote the ideal $\mathfrak{m}_p(f)$. Consider the diagram

$$\begin{tikzcd}
k[X,Y]/\mathfrak{a}^n \times k[X,Y]/\mathfrak{a}^m \arrow[d] & \\
k[X,Y]/\mathfrak{a}^{n+m} \arrow[d] & \mathcal{O}/(f,g) \arrow[d] \\
k[X,Y]/(\mathfrak{a}^{n+m},f,g) \arrow[r] \arrow[d] & \mathcal{O}/(\mathfrak{a}^{n+m},f,g) \arrow[d] \\
0 & 0
\end{tikzcd}$$

Where the map from the direct product is given by $(k_0,k_1) \mapsto k_0 f + k_1 g$, the map from left to right is given by the embedding of $k[X,Y]$ in $\mathcal{O}$, and the rest of the maps are given by a quotient. It is clear that the quotient maps are surjective. Furthermore, the map from left to right is an isomorphism, since $Z(\mathfrak{a}^{n+m},f,g) = \{p\}$. Finally, the kernel of the quotient map is $(f,g)/\mathfrak{a}^{n+m}$, which is exactly the image of the multiplication map, so the left column is exact. This implies that

$$\begin{aligned}
I_p(f,g) &= \dim \mathcal{O}/(f,g) \geqslant \dim \mathcal{O}/(\mathfrak{a}^{n+m},f,g) \\
&= \dim k[X,Y]/(\mathfrak{a}^{n+m},f,g) \\
&\geqslant \dim k[X,Y]/\mathfrak{a}^{n+m} - \dim k[X,Y]/\mathfrak{a}^n - \dim k[X,Y]/\mathfrak{a}^m \\
&= \frac{(n+m)(n+m+1)}{2} - \frac{n(n+1)}{2} - \frac{m(m+1)}{2} = nm
\end{aligned}$$

This is only an equality if the quotient map on the right is an isomorphism, so that $\mathfrak{a}^{n+m} \subset (f,g)$, and if the multiplication map is injective. We will prove this is true if and only if f and g have distinct tangents at p, in a series of lemmas following this proof.

□

Lemma 3.8. *If $L_1, L_2, \dots$ and $M_1, M_2, \dots$ is a series of linear forms in $k[X,Y]$, where we cannot write $L_i = \lambda M_j$ for any i, j, and if we let $A_{ij} = L_1 \dots L_i M_1 \dots M_j$, then the set $\{A_{ij} : i + j = n\}$ forms a basis for the set of forms of degree n in $k[X,Y]$.*

Proof. We prove the theorem by induction. For $n = 1$, we note that the set of linear forms has dimension two, and since $A_{10} = L_1$ and $A_{01} = M_1$ are linearly independent because they don't differ by a scalar, they span the space. If $\sum \lambda_i A_{i(n-i)} = 0$, then $\lambda_0 M_1 \dots M_n = -L_1(\sum \lambda_i L_2 \dots L_i M_1 \dots M_{n-i})$. If $\lambda_0 \neq 0$, this implies that L_1 divides M_i for some i, and this can only happen if L_1 and M_i differ by a constant. Thus $\lambda_0 = 0$, which implies $\sum \lambda_i L_2 \dots L_i M_1 \dots M_{n-i}$. Since $L_2, L_3, \dots$ and $M_1, M_2, \dots$ satisfy the conditions of the theorem, by induction we find that the set of products in the sum form a basis for the set of monomials of degree $n-1$, so $\lambda_2, \dots, \lambda_n$ are all zero. □

Lemma 3.9. *If f and g have no common tangents at p, then $\mathfrak{a}^{m+n-1} \subset (f,g)$.*

Proof. Let $L_1, \dots, L_m$ be the tangents to f at p, and $M_1, \dots, M_n$ the tangents to g at p. If we define $A_{ij} = L_1 \dots L_i M_1 \dots M_j$, with $M_j = M_n$ if $j > n$, $L_i = L_m$ if $i > m$, then $\{A_{ij} : i + j = n\}$ forms a basis for the set of forms of degree n because of the last lemma. It therefore suffices to show that $A_{ij} \in (f,g)$ if $i + j \geqslant m + n - 1$. If $i + j \geqslant m + n - 1$, then either $i \geqslant m$ or $j \geqslant n$. Without loss of generality, we may assume $i \geqslant m$, so that $A_{ij} = A_{m0} B$ for some form B of degree $i + j - m$. Write $f = A_{m0} + f_0$, where f_0 only has terms of degree greater than m. Then $A_{ij} = Bf - Bf_0$, where each term of Bf_0 has degree greater than $i + j + 1$. It clearly suffices to prove that $Bf_0 \in (f,g)$, and therefore we need only prove that $A_{ij} \in (f,g)$ for $i + j \geqslant m + n$. We may reapply this technique repeatedly to pump up the value of $i + j$ as large as desired, and therefore we need only prove that $\mathfrak{a}^t \subset (f,g)$ for sufficiently large t. This can be proved from the nullstellensatz. Because we are working in the local ring, we may assume that f and g have no common components at all, so $Z(f,g) = \{p, q_1, \dots, q_s\}$. We may then choose a polynomial h which vanishes on all q_i, but with $h(p) = 1$. Then $hX, hY \in I(Z(f,g))$, so $(hX)^t, (hY)^t \in (f,g)$ for sufficiently large t. Since h is a unit in $\mathcal{O}$, we conclude that $X^t, Y^t \in (f,g)$, hence $\mathfrak{a}^{2t} \in (f,g)$. □

Lemma 3.10. *The map $(k_0, k_1) \mapsto k_0 f + k_1 g$ from $k[X,Y]/\mathfrak{a}^n \times k[X,Y]/\mathfrak{a}^m$ is injective if and only if f and g have distinct tangents at p.*

Proof. Suppose that f and g have distinct tangents, and $k_0 f + k_1 g \in \mathfrak{a}^{n+m}$. We must conclude that $k_0 \in \mathfrak{a}^n$ and $k_1 \in \mathfrak{a}^m$, so assume otherwise. That is, take k_0 to have multiplicity r, and k_1 to have multiplicity s, and assume $r < n$ or $s < m$. Then $k_0 f + k_1 g = (k_0)_r f_m + (k_1)_s g_n +$ higher terms. We must have $r + m = s + n$, and that these terms cancel out, or else $k_0 f + k_1 g \notin \mathfrak{a}^{n+m}$. Thus $(k_0)_r f_m = -(k_1)_s g_n$. Since f and g have no common tangent, f_m and g_n have no common factors, so g_n divides $(k_0)_r$ and f_m divides $(k_1)_s$. But this means that $r \geqslant n$ and $m \geqslant s$ by contradiction. Conversely, if L was a common tangent to f and g, write $f_m = Lf'$ and $g_n = Lg'$, then $g' f = f' g$, so $(g', -f')$ maps to zero under the product map, yet the degree of f' is less than m, and the degree of g' is less than n. □

Example. *Let us calculate the intersection number of the two polynomials*

$$f = (X^2 + Y^2)^2 + 3X^2 Y - Y^3 \qquad g = (X^2 + Y^2)^3 - 4X^2 Y^2$$

at the origin. First, we replace g with $g - (X^2 + Y^2) f$, which is equal to

$$\begin{aligned} g_0 &= (X^2 + Y^2) Y^3 - 4X^2 Y^2 - 3(X^2 + Y^2) X^2 Y \\ &= Y[(X^2 + Y^2)(Y^2 - 3X^2) - 4X^2 Y] = Y g_1. \end{aligned}$$

Now $I_p(f, g) = I_p(f, Y) + I_p(f, g_1)$, and if we let

$$\begin{aligned} g_2 = g_1 + 3f &= 4Y^2(X^2 + Y^2) + 5X^2 Y - 3Y^3 \\ &= Y(4Y(X^2 + Y^2) + 5X^2 - 3Y^2) = Y g_3 \end{aligned}$$

So $I_p(f, g_1) = I_p(f, Y) + I_p(f, g_3)$. But f has tangent lines $Y = 0$ and $Y = \sqrt{3}X$ and $Y = -\sqrt{3}X$, and g_3 has tangent lines $Y = \sqrt{5/3}X$, $Y = -\sqrt{5/3}X$, which are distinct, hence $I_p(f, g_3) = m_p(f) m_p(g_3) = 6$, and $I_p(f, Y) = I_p(X^4, Y) = 4I_p(X, Y) = 4$. Thus $I_p(f, g) = 2I_p(f, Y) + I_p(f, g_3) = 14$.

Example. *A double point p is a* cusp *on f if it has a single tangent line L. Then $I_p(f, L) \geqslant 3$, because if we assume $L = Y$, and p the origin, and if we write $f = Y^2 + f'(X, Y)$, then $I_p(f, Y) = I_p(f', Y) \geqslant m_p(f') \geqslant 3$. A point is a simple cusp if the intersection number I_p is exactly 3. In the case $L = Y$ specifically, we find that f has a simple cusp if and only if $f_{XXX}(0) \neq 0$, because if we write $f(X, Y) = Y^2 + aX^3 + bX^2 Y + cXY^2 + dY^3 + \dots$, then $f_{XXX}(0) = 6a$, and $I_p(f, Y) = 3$ only when Y does not divide $aX^3 + bX^2 Y + cXY^2 + dY^3$, i.e. when $a \neq 0$. If a curve has a cusp at p, then there is only a single component of* *the curve which passes through p. To see this, suppose $f_0 f_1$ have a cusp p, and f_0 and f_1 both pass through p. Since p is a double point, p must be a simple point on f_0 and f_1, and they must have the same tangent L. But then*

$$I_p(f_0 f_1, L) = I_p(f_0, L) + I_p(f_1, L) \geqslant 2 + 2 = 4$$

So p cannot be a simple cusp. More generally, p is a hypercusp if $m_p(f) > 1$, f has a unique tangent L at p, and $I_p(f,L) = m_p(f) + 1$. Essentially the same arguments show that the origin which has a unique tangent Y on f is a hypercusp if $f_{X^{n+1}}(0) \neq 0$, and a hypercusp can only lie on a single component of a curve.

Before we move on, we would like to note two more properties of the intersection number that will become more useful later.

Theorem 3.11. *If p is simple on a curve f, then $I_p(f,g) = ord_p(g)$ on f.*

Proof. We may assume that f is an irreducible curve. From the general properties of discrete valuation rings, $\text{ord}_p(g)$ is equal to the dimension of $\mathcal{O}_p(f)/(g)$. But since localization commutes with quotients, we find that $\mathcal{O}_p(f)/(g)$ is isomorphic to $\mathcal{O}_p(\mathbf{A}^2)/(f,g)$, and this is the definition of the intersection number. Alternatively, we can prove this by using the axiomatic properties of intersection properties. We may assume p is the origin, and f has a tangent $Y = 0$ at the origin. Then X is a uniformizing parameter for the local ring at the origin. This means that for any polynomial g, $g = X^n h/k$, where $h(0), k(0) \neq 0$. Then $kg = X^n h$. Now $I_p(f,kg) = I_p(f,k) + I_p(f,g) = I_p(f,g)$, because p does not lie on the curve defined by k. But $I_p(f,kg) = I_p(f,X^n) + I_p(f,h) = I_p(f,X^n)$, and the fact that $I_p(f,X^n) = n$ follows because X^n does not have $Y = 0$ as a tangent. □

This in particular implies that if p is simple on f, then

$$I_p(f, g+h) \geqslant \min(I_p(f,g), I_p(f,h))$$

This need not be true if p is not simple on f, for instance, if $L_1 = X + Y$ and $L_2 = X - Y$ are the two tangents to the cubic $Y^2 - X^2 - X^3$ at the origin, then

$$I_0(Y^2 - X^2 - X^3, L_1 + L_2) = I_0(Y^2 - X^2 - X^3, X) = I_0(Y^2, X) = 2$$

But

$$I_0(Y^2 - X^2 - X^3, X + Y) = I_0(Y^2 - X^2 - X^3, X - Y) = 3$$

Theorem 3.12. *If f and g have no common components, then*

$$\sum_p I_p(f,g) = \dim k[X,Y]/(f,g)$$

Proof. If f and g are relatively prime, then $Z(f,g)$ contains finitely many points, and we know that $k[X,Y]/(f,g)$ is isomorphic to the direct product of the $\mathcal{O}_p(\mathbf{A}^2)/(f,g)$, as p ranges over $Z(f,g)$. But taking the dimension of both of these spaces gives the sum formula above. □

We will at some point need a nice criterion to determine if a point p on an irreducible curve is an ordinary multiple point solely through the analysis of $\mathcal{O}_p(f)$, with $m_p(f) = m > 1$. Suppose that p is the origin. The embedding of $k_1[X,Y]$ in $\mathfrak{m}/\mathfrak{m}^2$ is an isomorphism, because both spaces are vectors spaces of dimension 2, and if $aX + bY \in \mathfrak{m}^2$, then $aX + bY + gf = hk$, with $h(0) = k(0) = 0$, then $m_p(hk) \geqslant 2$, and $m_p(aX + bY + gf)$ can only be greater than or equal to 2 if $a = b = 0$. If p is an ordinary multiple point with tangents $L_1, \dots, L_m$, then $I_p(f, L_i) > m$, and L_i is not congruent to λL_j modulo $\mathfrak{m}^2$ for any λ and $i \neq j$, because if L_i and L_j form a basis for $k_1[X,Y]$. Conversely, if there are $g_1, \dots, g_m \in k[X,Y]$ with g_i not congruent to λg_j in $\mathfrak{m}/\mathfrak{m}^2$ with $I_p(f, g_i) > m$, then in particular $m_p(g_i) = 1$, the g_i are congruent to their tangents modulo $\mathfrak{m}^2$, and the g_i have distinct tangents. The fact that $I_p(f, g_i) > m$ implies that the tangent corresponding to g_i is also a tangent of f, and since f has only tangents up to multiplicity m, we conclude that f has distinct tangents. Since $I_p(f, g_i)$ is defined with respect to $\mathcal{O}_p(f)$, this gives us the required criterion. More generally, this is true provided that f is not divisible by multiple factors of the same irreducible polynomials, because $\mathcal{O}_p(f_1^{n_1} \dots f_m^{n_m}) \cong \mathcal{O}_p(f_1 \dots f_m)$ is unable to detect powers of polynomials.

3.4 Projective Planar Curves

If we work with projective varieties instead of affine varieties, we obtain the most powerful global results. Set theoretically, a curve in the projective plane $\mathbf{P}^2$ is defined to be the locus of a single homogenous polynomial in $k[X,Y,Z]$. As with affine plane curves, however, we will find it more elegant to allow curves to have 'multiplicities', so that projective planar curves will be defined to be an equivalence class of homogenous polynomials which are identified under scalar multiplication. All the notation we used for affine planar curves carries over, without much modification, to this projective case.

Just as in the affine case, given $p \in \mathbf{P}^2$ and a projective curve C, we define the multiplicity $m_p(C)$ to be the dimension of

$$\dim_k(\mathfrak{m}_p(C)^n/\mathfrak{m}_p(C)^{n+1})$$

for large n. This multiplicity is invariant under projective transformations, and if p is a finite point, we know that $\mathcal{O}_p(f)$ is isomorphic to $\mathcal{O}_p(f_*)$, so $m_p(f) = m_p(f_*)$.

Example. *The polynomial $XY^4 + YZ^4 + XZ^4$ is irreducible, because, dehomogenizing, elementary arguments prove that $f = XY^4 + Y + X$ is irreducible. Now the plane curve has no finite singular points, because if $f_X = Y^4 + 1$, $f_Y = 4XY^3 + 1$, then $f_X = 0$ holds if and only if $Y^4 = -1$, which implies $XY^4 + Y + X = Y$, which vanishes if and only if $Y = 0$, and then f vanishes only when $X = 0$, but then $f_Y(0) = 1$ is nonzero. However, the curve does have two points at infinity, $[0 : 1 : 0]$ and $[1 : 0 : 0]$. To understand the first point qualitatively, we dehomogenize with respect to Y, looking at $X + Z^4 + XZ^4$, which is simple at the origin, hence simple at $[0 : 1 : 0]$. However, dehomogenizing with respect to X to obtain qualitative properties of the second point, we obtain $Y^4 + YZ^4 + Z^4$, which has a multiple point of degree four at the origin, with tangent lines $Y + \omega Z$, $Y + i\omega Z$, $Y - \omega Z$, and $Y - i\omega Z$, so $[0 : 1 : 0]$ is an ordinary multiple point of degree four, with the tangents $Y + \omega Z$, $Y + i\omega Z$, $Y - \omega Z$, and $Y - i\omega Z$, which are four lines parallel to the Y axis.*

Example. *The polynomial $X^2Y^3 + X^2Z^3 + Y^2Z^3$ defines an irreducible curve. Dehomogenizing, we find that $X^2Y^3 + X^2 + Y^2$ only has a singular point at the origin, where the two tangents are $X = iY$ and $X = -iY$. The curve has two points at infinity, $[1 : 0 : 0]$ and $[0 : 1 : 0]$. Dehomogenizing with respect to each variable tells us that the first point has three tangents $Y - \omega Z$, $Y - \omega\nu Z$, and $Y - \omega\nu^2 Z$, where ν is a third root of unity, and $\omega^3 = -1$, and that the second point has a single double tangent X.*

Since we cannot decompose homogenous polynomials into tangents in the same way that we can in affine space, we must identify tangents to algebraic curves in a more abstract manner. We say a line L is tangent to a projective curve C at a point p if $I_p(C,L) > m_p(C)$. If p is a finite point, this is equivalent to saying that L is tangent in affine coordinates. A point p is ordinary if it has distinct tangents.

Example. *Let C be a curve defined by a polynomial f containing a point p. Then p is a multiple point if and only if $f(p) = 0$ and $f_X(p) = f_Y(p) = f_Z(p) = 0$. Without loss of generality, we may assume that $p = [0:0:1]$ and that f is homogeneous of degree m. Then the fact that $f(p) = 0$ implies that*

$$f(X,Y,Z) = aXZ^{m-1} + bYZ^{m-1} + f_0(X,Y,Z),$$

where f_0 only has terms up to degree two in X and Y. Then $f_X(p) = f_Y(p) = f_Z(p) = 0$ if and only if $a = b = 0$, and this holds if and only if $m_p(C) > 1$.

Let us now use some of these properties to classify some low degree projective curves.

Example. *There is only a single irreducible conic up to projective equivalence and it is nonsingular. Let C be an irreducible conic defined by a polynomial f. By a projective transformation, we may assume that C has a simple point at $[0:1:0]$ with tangent $Z = 0$. Thus we can write*

$$f = YZ - aX^2 - bXZ - cZ^2$$

for some ocnstants a, b, and c. Since f is irreducible, Z does not divide f, and as such a must be nonzero. But dehomogenizing in the Z variable shows this is just the equation for a parabola, and all parabolas are affinely equivalent. Thus all irreducible conics are projectively equivalent to the conic described by the equation $YZ = X^2$.

Example. *There is only a single projective cubic with a cusp, up to projective equivalence. Without loss of generality, we may assume the cusp occurs at the origin with tangent $Y = 0$. The equation for such a curve must be of hte form*

$$Y^2Z = aX^3 + bX^2Y + cXY^2 + dY^3.$$

Scaling X, we may assume that $a = 1$, and applying a Tschirnhäus transformation to X, we can also set $b = 0$. Finally, replacing Z with $Z + cX + dY$ shows that the cubic we were looking at is equivalent to the cubic $Y^2Z = X^3$. Thus, up to projective equivalence, there is a unique cubic with a cusp.

Example. *There is a single irreducible cubic curve in the plane with an ordinary double point. By similar techniques to the last example, if we assume the double point occurs at the origin, with the two tangents $X = 0$ and $Y = 0$, then one can apply projective transformations to reduce the general equation obtained down to the form $XYX = X^3 + Y^3$. This describes the folium of Descartes.*

Remark. The last two examples shows that any irreducible cubic is either nonsingular, or projectively equivalent to the curve described by the equation $Y^2Z = X^3$ or the curve described by the equation $XYZ = X^3 + Y^3$. In particular, an irreducible cubic curve can have only a single singularity.

If p is a simple point on a projective curve, in the sense that $m_p(C) = 1$, then $\mathcal{O}_p(C)$ is a discrete valuation domain. We extend the order function on $\mathcal{O}_p(C)$ to any form $g \in k[X,Y,Z]$, by letting $\text{ord}_p(g) = \text{ord}_p(g/h)$, where h is a form of the same degree as g, with $h(p) \neq 0$.

To define the intersection multiplicities of two projective curves at a finite point p, we let

$$I_p(f,g) = I_p(f_*,g_*) = \dim_k(\mathcal{O}_p(\mathbf{P}^2)/(f_*,g_*)).$$

Then the intersection number is invariant of projective transformations that map p to a finite point, and because of this, we may define the intersection number at infinite points by first mapping p to a finite point, and then calculating the intersection number there. This new definition of intersection numbers satisfies all the axioms of intersection numbers, except that $I_p(f,g) = I_p(f,g+kf)$ only when $g+kf$ is a homogenous polynomial, which only happens if k is homogenous of degree $\deg g - \deg f$. Generalizing properties of affine curves, we say a line L is tangent to f at p if $I_p(f,L) > m_p(f)$. We say p is ordinary if it has $m_p(f)$ distinct tangents.

3.5 Linear Systems of Curves

We often focus on a the class of all projective planar curves of a fixed degree n. The space of homogenous polynomials of degree n in $K[X,Y,Z]$ is a vector space of dimension $\sum_{i=0}^n (i+1) = (n+1)(n+2)/2 = n(n+3)/2 + 1$. However, two polynomials which are scalar multiples of each other define the same curve, so the space of planar curves of degree n is naturally identified with $\mathbf{P}^{n(n+3)/2}$. This is a basic example of a *moduli space*, i.e. a geometric space formed from a class of objects.

Example. *The space of lines in the plane is naturally identified with $\mathbf{P}^2$, which is a manifestation of the duality between points and lines in projective geometry.*

Example. *The space of projective conics in the plane is in one to one correspondence with $\mathbf{P}^5$, the cubics $\mathbf{P}^9$, and the quartics with $\mathbf{P}^{14}$.*

If a subset of curves of a particular degree form a linear subvariety of projective space, we call that family a *linear system of curves*. We note that if $T : \mathbf{P}^2 \to \mathbf{P}^2$ is a projective transformation, then the induced map $T^* : \mathbf{P}^{n(n+3)/2} \to \mathbf{P}^{n(n+3)/2}$ on the family of degree n curves is also a projective transformation.

Example. *For any point $p \in \mathbf{P}^2$, the space of degree n curves through p forms a hyperplane in $\mathbf{P}^{n(n+3)/2}$. If $p = [x : y : z]$, and if a curve is described by the equation $\sum a_{ij} X^i Y^j Z^{n-i-j}$, then this curve contains p if and only if $\sum a_{ij} x^i y^j z^{n-i-j} = 0$, which is a linear equation in the coefficients $\{a_{ij}\}$. More generally, the set of degree n curves passing through m points forms a linear system of curves. Since the intersection of n hyperplanes on $\mathbf{P}^n$ is always nonempty, there is always a curve of degree n passing through any given set of $n(n+3)/2$ points.*

Example. *Fix $m > 0$ and a point p. Then the set of curves f of degree d such that $m_p(f) \geqslant m$ forms a linear subvariety of dimension $d(d+3)/2 - m(m+1)/2$. Because of the fact that projective transformations translate the coordinates of the moduli of curves projectively, we may assume $p = [0 : 0 : 1]$. Write $f = \sum f_i(X,Y) Z^{n-i}$. Then $m_p(f) \geqslant m$ if and only if $f_0 = f_1 = \cdots = f_{m-1} = 0$.*

More generally, for any points $p_1, \dots, p_n$ and integers $m_1, \dots, m_n$, the family of curves f of degree n such that $m_{p_i}(f_i) \geqslant m_i$ for each i is a linear family. We let $Z(d; p_1, n_1, \dots, p_k, m_k)$ be the family of degree d curves of this form. This is a space with dimension at least $d(d+3)/2 - \sum n_i(n_i+1)/2$, because we are placing $\sum n_i(n_i+1)/2$ such constraints on this space. If $d \geqslant \sum n_i - 1$, then this is an exact equality, which we prove by induction.

We prove this result by induction. First, suppose that $n_i = 1$ for each i. Let $V_i = Z(d; p_1, \dots, p_i)$. It suffices to show that $V_n \neq V_{n+1}$. For this, we choose lines L_i which pass through p_i, but not through p_j for $i \neq j$, and a line L_0 not passing through any points p_i. Then $f = L_1 \dots L_{n-1} L_0^{d-n+1}$ is in V_{n-1}, *but not in V_n. Next, let $n_i > 1$, and for simplicity in notation let $i = 1$. Let $V_0 = Z(d; p_1, n_1 - 1, p_2, n_2, \dots, p_m, n_m)$. For $f \in V_0$ let*

$$f_* = \sum a_i X^i Y^{n_1 - 1 - i} + \textit{higher terms}$$

Let $V_i = \{f \in V_0 : a_j = 0 \text{ for } j < i\}$. It is again, enough to show that $V_i \neq V_{i+1}$. Let $W_0 = Z(d - 1; p_1, n_1 - 2, p_2, n_2, \dots, p_m, n_m)$, and

$$W_i = \{f \in W_0 : a_j = 0 \text{ for } j < i\}$$

By induction,

$$W_0 \supsetneq W_1 \supsetneq \cdots \supsetneq W_{n_1} = Z(p_1, n_1 - 1, p_2, n_2, \dots, p_m, n_m)$$

If $f_i \in W_i$, $f_i \notin W_{i+1}$, then $Y f_i \in V_i$, $Y f_i \notin V_{i+1}$ and $X f_{n_1 - 2} \in V_{n_1 - 1}$, $X f_{n_1 - 2} \notin V_{n_1}$. This shows $V_i \neq V_{i+1}$ for all $0 \leqslant i \leqslant n_1 - 1$, completing the proof.

On the other hand, for more complicated arrangements of places the dimension of the resulting linear family can highly depend on the arrangement of the points in question.

Example. *Let $p_1, p_2, p_3, p_4 \in \mathbf{P}^2$ be four points. Let V be the linear system of conics passing through the four points. Then V has dimension 2 if $p_1, \dots, p_4$ lie on a line, and V has dimension 1 otherwise. If the four points are non colinear, then three of the points aren't colinear, and by a projective transformation we may assume they occur at $[0:0:1]$, $[0:1:0]$, $[1:0:0]$, and $[x:y:z]$. Any projective curve that vanishes on the first three points is of the form $aXY + bYZ + cZX$, and the set of projective curves which pass through the final point satisfy the equation $axy + byz + czx = 0$. This is a nondegenerate linear functional on the space, because if $xy = yz = zx = 0$, then two or more of x, y, and z are equal to 0, in which case find that $[x:y:z]$ is equal to one of the first three points. Thus this linear equation reduces the dimension of solutions by a single dimension, and since the space of $aXY + bYZ + cZX$ is two dimensional, we conclude the space of solutions is one dimensional. If the four points are colinear, we may assume the first two are $[0:0:1]$ and $[1:0:0]$, and that the other two are of the form $[x:0:1]$ and $[y:0:1]$. The set of conics passing through the first two points are of the form $aY^2 + bXY + cYZ + dZX$, and the conditions guaranteeing that the other two points lie on this conic are that $dx = dy = 0$. Since $x, y \neq 0$, this is equivalent to saying that $d = 0$, so the space of curves are exactly $aY^2 + bXY + cYZ$, a two dimensional projective linear subvariety.*

3.6 Bezout's Theorem

We have finally come to the famous theorem of Bezout, which says that projective planar curves intersect in 'just the right amount' of places.

Theorem 3.13. *If f and g are curves of degree n and m, then $\sum_p I_p(f,g) = mn$.*

Proof. We may assume that f and g do not intersect on the line at infinity by applying a projective transformation. Then for each finite point p, $I_p(f,g)$ is equal to the dimension of $\mathcal{O}_p(\mathbf{A}^2)/(f_*,g_*)$. But since $Z(f_*,g_*)$ contains only finitely many points, we actually have

$$k[X,Y]/(f_*,g_*) \cong \prod_p \mathcal{O}_p(\mathbf{A}^2)/(f_*,g_*).$$

This implies that

$$\sum_p I_p(f,g) = \sum_p I_p(f_*,g_*) = \dim k[X,Y]/(f_*,g_*)$$

Thus the theorem is proven if we can show that $\dim_k k[X,Y]/(f_*,g_*) = nm$.

Let $A = k[X,Y,Z]$. For each integer k, let A_k denote the set of forms of degree k, let B denote $k[X,Y,Z]/(f,g)$, and let C denote $k[X,Y]/(f_*,g_*)$. We have an exact sequence

$$0 \to A \to A \times A \to A \to B \to 0$$

where the first map is $h \mapsto (gh, -fh)$, the second map is the map given by $(h_0,h_1) \mapsto h_0 f + h_1 g$, and the third map is the quotient map. By restricting the degrees of the considered polynomials, we find that for $d \geqslant n+m$ we have an exact sequence

$$0 \to A_{d-m-n} \to A_{d-m} \times A_{d-n} \to A_d \to B_d \to 0$$

and it follows by dimension counting that for $d \geqslant n+m$, $\dim B_d = nm$.

Next, we prove the map from B to itself given by mapping h to Zh is injective. It suffices to show that if $Zh = af + bg$ for some $a,b \in k[X,Y,Z]$, then $h = a'f + b'g$ for some a',b'. For any polynomial $k(X,Y,Z)$, denote by $k_0(X,Y)$ the polynomial $K(X,Y,0)$. Since f,g, and Z have no common zeroes, f_0 and g_0 are relatively prime in $k[X,Y]$, for if f_0 and g_0 shared a common nonconstant factor k, then we would find $k(x,y,0) = 0$ for some points x and y, and then $(x,y,0)$ lies on f,g, and Z. Note that $a_0f_0 = -b_0g_0$, so $b_0 = f_0c$ and $a_0 = -g_0c$ for some c. Since this means $(a+cg)_0 = (b-cf)_0 = 0$, $a+cg = a'Z$, and $b-cf = b'Z$. Then since $Zh = (a+cg)f + (b-cf)g = Za'f + Zb'g$, we find that $h = a'f + b'g$.

Finally, let $d \geqslant m+n$, and choose $a_1,\dots,a_{mn} \in A_d$ which form a basis in B_d. Let $(a_i)_* = a_i(X,Y,1)$. Then we claim that the $(a_i)_*$ form a basis for C. The map $h \mapsto Zh$ is an isomorphism of B_d onto B_{d+1}, because an injective map between vector spaces of the same dimension must be an isomorphism. It follows that $Z^ra_1,\dots,Z^ra_{mn}$ form a basis for $k_{d+r}[X,Y,Z]/(f,g)$ for all $r \geqslant 0$. If $h \in k[X,Y]$ is arbitrary, we must show it is congruent to a unique sum of the $(a_i)_*$. Now we can choose t such that Z^th^* is a form of degree $d+r$, so $Z^th^* = \sum_{i=1}^{mn} \lambda_i Z^r a_i + bf + cg$ for some $\lambda_i \in k$, but then $h = (Z^th^*)_* = \sum_{i=1}^{mn} \lambda_i (a_i)_* + bf_* + cg_*$. This shows that $(a_i)_*$ form a spanning set. To show independence, suppose $\sum \lambda_i (a_i)_* = bf_* + cg_*$. Then $Z^t \sum \lambda_i a_i = Z^s b^* f + Z^u c^* g$, in which case $\sum \lambda_i (Z^t a_i) = 0$ in B_{d+r}, and the Z^ta_i form a basis in B_{d+r}, so the λ_i are identically zero. This finishes the proof. □

There are many geometrical applications of Bezout's theorem. Let us begin by using it to count multiple points on a curve. A simple corollary of Bezout's theorem is that if two curves meet in mn distinct points, then each such point in the intersection is simple on both curves; more generally for two curves C_1 and C_2 of degree n and m not sharing any common components,

$$\sum_p m_p(C_1) m_p(C_2) \leqslant nm.$$

In particular, applying this to an irreducible curve C described by a polynomial f, we conclude that

$$\sum_p m_p(f) m_p(f_{X_1}) \leqslant n(n-1)$$

But

$$m_p(f) m_p(f_{X_1}) = m_p(f)(m_p(f) - 1)$$

so we conclude that

$$\sum m_p(f)(m_p(f)-1) \leqslant n(n-1)$$

In particular, f has at most $n(n-1)/2$ multiple points. If we are slightly more careful, we can obtain a much tighter result.

Theorem 3.14. *Let C be an irreducible curve of degree n. Then*

$$\sum_p m_p(C)(m_p(C)-1) \leqslant (n-1)(n-2).$$

In particular, C has at most $(n-1)(n-2)/2$ multiple points.

Proof. Choose

$$r = \frac{(n-1)(n+2)}{2} - \sum \frac{m_p(C)(m_p(C)-1)}{2}.$$

The last paragraph showed that $r \geqslant 0$. Pick r simple points $p_1, \dots, p_r$ on C. Then our theory of linear systems of curves implies there exists a curve C' of degree $n-1$ with $m_p(C') \geqslant m_p(C) - 1$ for all p, and $m_{p_i}(C') \geqslant 1$ for each i. Applying Bezout's theorem, we conclude that

$$n(n-1) \geqslant \sum m_p(C) m_p(C') \geqslant r + \sum m_p(C)(m_p(C)-1).$$

Rearranging proves the result. □

Remark. Modifying this proof shows that if C is a plane curve of degree n consisting of m simple components (and no multiple components), then

$$\sum m_p(C)(m_p(C)-1) \leqslant (n-1)(n-2) + 2(m-1),$$

which is a slightly improved estimate to that given

Another simple corollary is that a nonsingular projective planar curve is irreducible, because if the curve contains two or more components, they must intersect somewhere by Bezout's theorem, and thus the curve contains a multiple point.

We can also use Bezout's theorem to count the number of flex points on a particular curve. Given a projective curve f in the plane, form the 3 by 3 symmetric *Hessian matrix* H with $H_{ij} = f_{X_i X_j}$. Since each element of this matrix has degree $n-2$, the Hessian h, which is the determinant of this matrix, is a polynomial of degree $3(n-2)$.

Theorem 3.15. *Over a field of characteristic zero, if p is a flex point on a curve C, then $H(p) = 0$. If $p \in C$ and $H(p) = 0$, then p is either a flex point or a multiple point. Moreover, $I_p(C, H) = 1$ if and only if p is an ordinary flex point.*

Proof. Since a projective transformation T only changes the Hessian by a constant scalar value $\det(T)^2$, this theorem is invariant under projective transformations, so we may assuming that we are analyzing the point p at the origin. Write $f_1(X,Y) = f(X,Y,1)$, and $h_1(X,Y) = h(X,Y,1)$. Applying Euler's theorem for homogenous polynomials, if f has degree n, we find that

$$(n-1)f_X = Xf_{XX} + Yf_{XY} + Zf_{XZ},$$
$$(n-1)f_Y = Xf_{XY} + Yf_{YY} + Zf_{YZ},$$

and

$$(n-1)f_Z = Xf_{XZ} + Yf_{YZ} + Zf_{ZZ}.$$

Next, if we add X times the first row and Y times the second row to the third row in H, we find that

$$\begin{aligned} &H(X,Y,Z)\\ &\quad = \det\begin{pmatrix} f_{XX} & f_{XY} & f_{XZ} \\ f_{XY} & f_{YY} & f_{YZ} \\ Xf_{XX} + Yf_{XY} + f_{XZ} & Xf_{XY} + Yf_{YY} + f_{YZ} & Xf_{XZ} + Yf_{YZ} + f_{ZZ} \end{pmatrix}\\ &\quad = \det\begin{pmatrix} f_{XX} & f_{XY} & f_{XZ} \\ f_{XY} & f_{YY} & f_{YZ} \\ (n-1)f_X + (1-Z)f_{XZ} & (n-1)f_Y + (1-Z)f_{YZ} + f_{YZ} & (n-1)f_Z + (1-Z)f_{ZZ} \end{pmatrix} \end{aligned}$$

In particular,

$$H(X,Y,1) = (n-1)\det\begin{pmatrix} f_{XX} & f_{XY} & f_{XZ} \\ f_{YX} & f_{YY} & f_{YZ} \\ f_X & f_Y & f_Z \end{pmatrix}$$

Next, adding X times the first column and Y times the second column to the third column, we find that

$$\begin{aligned} \det\begin{pmatrix} f_{XX} & f_{XY} & f_{XZ} \\ f_{YX} & f_{YY} & f_{YZ} \\ f_X & f_Y & f_Z \end{pmatrix} &= \det\begin{pmatrix} f_{XX} & f_{XY} & Xf_{XX} + Yf_{YX} + f_{XZ} \\ f_{YX} & f_{YY} & Xf_{XY} + Yf_{YY} + f_{ZY} \\ f_X & f_Y & Xf_X + Yf_Y + f_Z \end{pmatrix}\\ &= \det\begin{pmatrix} f_{XX} & f_{XY} & (n-1)f_X + (1-Z)f_{XZ} \\ f_{YX} & f_{YY} & (n-1)f_Y + (1-Z)f_{YZ} \\ f_X & f_Y & (n-1)f + (1-Z)f_Z \end{pmatrix} \end{aligned}$$

Thus

$$\begin{aligned} H(X,Y,1) &= (n-1)^2 \det \begin{pmatrix} f_{XX} & f_{XY} & f_X \\ f_{XY} & f_{YY} & f_Y \\ f_X & f_Y & f \end{pmatrix} \\ &= (n-1)^2 (f f_{XX} f_{YY} - f_{XX} f_Y^2 - f f_{XY}^2 + 2 f_X f_Y f_{XY} - f_{YY} f_X^2) \end{aligned}$$

In particular, modulo (f), we find that

$$I_p(f,H) = I_p(f, f_X^2 f_{YY} + f_Y^2 f_{XX} - 2 f_X f_Y f_{XY}).$$

Set $g(X,Y,Z) = f_X^2 f_{YY} + f_Y^2 f_{XX} - 2 f_X f_Y f_{XY}$. It is simple to check that if p is a multiple point on f, then p is a multiple point on g. On the other hand, if p is a simple point on f, and we assume $Y = 0$ is the tangent line to f at p, so we can write

$$f_1(X,Y) = Y + aX^2 + bXY + cY^2 + dX^3 + eX^2Y + \dots.$$

We have seen f is flex at p if and only if $a = 0$, and this flex is ordinary if and only if $d \neq 0$. We calculate that working only in linear terms,

$$\begin{aligned} g(X,Y) &= (2aX + bY)^2(2c) + (1 + bX + 2cY)^2(2a + 6dX + 2eY) + \mathfrak{m}^2 \\ &= -2(2aX + bY)(1 + bX + 2cY)(b + 2eX) + \mathfrak{m}^2 \\ &= 2a + 6dX + (8ac - 2b^2 + 2e)Y + \mathfrak{m}^2. \end{aligned}$$

Thus $g(p) = 0$ if and only if $a = 0$, identifying the flexes, and $I_p(f,g) = 1$ if and only if $d \neq 0$, because otherwise g and f have the same tangent at the origin. □

Corollary 3.16. *A nonsingular curve of degree > 2 over a field of characteristic zero always has a flex. A nonsingular cubic has nine flexes, all ordinary.*

Proof. If f has degree $n > 2$, and defines a nonsingular projective curve, then H is a homogenous polynomial of degree $3(n-2)$, and so Bezout's theorem implies that

$$\sum I_p(f,H) = 3n(n-2).$$

In particular, this implies f and h intersect. Since f has no multiple points, this point must be a flex point.

If f is a cubic, then f and H have intersection multiplicity 9. We will prove that each intersection point is transversal. Without loss of generality, assume f and H have an intersection point at the origin, and that f has tangent $Y = 0$ at the origin, so that since f is a flex, we may write

$$f(X,Y) = YZ^2 + aXYZ + bY^2Z + cX^3 + dX^2Y + eXY^2$$

Then

$$H = \det \begin{pmatrix} 6cX + 2dY & aZ + 2dX + 2eY & aY \\ aZ + 2dX & 2bZ + 2eX & 2Z + aX + 2bY \\ aY & 2Z + aX + 2bY & 2Y \end{pmatrix}$$

We find

$$\begin{aligned} H_X(0) &= \det \begin{pmatrix} 6c & 2d & 0 \\ a & 2b & 2 \\ 0 & 2 & 0 \end{pmatrix} + \det \begin{pmatrix} 0 & a & 0 \\ 2d & 2e & a \\ 0 & 2 & 0 \end{pmatrix} + \det \begin{pmatrix} 0 & a & 0 \\ a & 2b & 2 \\ 0 & a & 0 \end{pmatrix} \\ &= -24c \end{aligned}$$

$$\begin{aligned} H_Y(0) &= \det \begin{pmatrix} 2d & 2e & a \\ a & 2b & 2 \\ 0 & 2 & 0 \end{pmatrix} + \det \begin{pmatrix} 0 & a & 0 \\ 0 & 0 & 2b \\ 0 & 2 & 0 \end{pmatrix} + \det \begin{pmatrix} 0 & a & 0 \\ a & 2b & 2 \\ a & 2b & 2 \end{pmatrix} \\ &= 2a^2 - 8d \end{aligned}$$

If H has p as a simple point, but has $Y = 0$ as a tangent, or if H has a multiple point at the origin, then $c = 0$. But this implies Y divides f, which implies that f is not irreducible, and therefore cannot be nonsingular. We conclude that f and H must intersect transversally wherever they touch, and therefore f has 9 distinct ordinary flexes on the projective plane. □

Let us conclude this section with a final argument. It is a general heuristic that transversality is a 'generic' property. In other words, a generic line through should be transversal to a given curve. This remains true, in some senses, even if we restrict the line to pass through a multiple point.

Theorem 3.17. *Let p be a point on an irreducible planar curve C of degree n, with $m_p(C) = m$. Then for all but finitely many lines L through p, L intersects C in $n - m$ distinct points other than p.*

Proof. Assume without loss of generality that $p = [0 : 1 : 0]$. For each λ, write L_λ for the line described by the equation $X - \lambda Z = 0$. The family of all such lines describes all lines through p except for the line $Z = 0$. Thus it suffices to show that for all but finitely many λ, L_λ passes through $n - m$ distinct points other than p. Since $m_p(C) = m$, if f describes the curve C, then we can write

$$f(X,Y,Z) = A_m(X,Z)Y^{n-m} + \cdots + A_n(X,Z),$$

where A_i is homogeneous of degree i.

We claim that if $A_m(\lambda,1) \neq 0$ and $Z(f) \cap Z(f_Y) \cap Z(L_\lambda) = \{p\}$, then L_λ passes through $n - m$ other distinct points. Indeed, Bezout's theorem implies that

$$\sum_{q \neq p} I_q(f, L_\lambda) = n - m$$

and so it suffices to prove that $I_q(f, L_\lambda) \leqslant 1$ for all $q \neq p$. By symmetry it suffices to show this when $\lambda = 0$, $q = [0 : 0 : 1]$, and $q \in Z(f)$. Now

$$I_q(f, L_0) = I_q(f(0,Y,1), X).$$

Now $f(0,Y,1)$ is a degree $n - m$ polynomial in Y, and it's nullset in $\mathbf{A}^2$ is a union of horizontal lines. On the other hand, X is a vertical line, so these two curves are transversal to one another. But now note that $m_0(f(0,Y,1)) = 1$ because $f_Y(0,Y,1) = 0$. Thus $I_p(f,X) = 1$.

Finally, it suffices to note that $A_m(\lambda,1) \neq 0$ for all but finitely many λ, because $A_m(\lambda,1)$ is a non-zero polynomial in λ. And $Z(f) \cap Z(f_Y) \cap Z(L_\lambda) = \{p\}$ for all but finitely many λ, because f is irreducible so $Z(f) \cap Z(f_Y)$ meet in only finitely many points, and only finitely many such values λ will pass through any of these points. □

Remark. This theorem continues to hold for non-irreducible curves provided these curves have no repeated components. Indeed, if $C = C_1 \cup \cdots \cup C_k$ is a curve of degree n containing a point p with $m_p(C) = m$, and $n = n_1 + \cdots + n_k$, where C_i is an irreducible curve of degree n_i. Then $m = m_1 + \cdots + m_n$ where $m_p(C_i) = m_i$. If $m_i > 0$, then for all but finitely many lines L_λ, L_λ passes through $n_i - m_i$ distinct points on C_i. These are only finitely many multiple points on C and so only finitely many of the lines L_λ pass through these multiple points. Thus all that remains to be shown is that if $m_i = 0$, then all but finitely many lines L_λ pass through finitely many points of C_i. But a similar approach as above works because this condition holds if $Z(f_i) \cap Z((f_i)_Y) \cap Z(L_\lambda) = \emptyset$, and this condition holds at all but finitely many points.

3.7 Max Noether's Fundamental Theorem

We now consider Bezóut's theorem from the zero dimensional point of view. If we let X denote the class of all projective plane curves, then the space of all curves can be put into one to one correspondence of the set of all finite positive abelian sums of elements of X, that is, with the positive elements of the free abelian group $\mathbf{Z}\langle X \rangle$. The elements of X can be seen as the 'one dimensional' algebraic subsets of the plane. Similarly, we can view the irreducible zero dimensional subsets of the projective plane are precisely the points, and we can define a general 'zero dimensional planar variety' as a positive element of $\mathbf{Z}\langle \mathbf{P}^2 \rangle$. We shall call the elements of $\mathbf{Z}\langle \mathbf{P}^2 \rangle$ *zero cycles*, and we define the degree of a cycle $\sum n_p p$ to be $\sum n_p$.

Let f and g be projective plane curves of dimension m and n, with no common components. We define the *intersection cycle* $f \cdot g$ to be the positive zero cycle $\sum I_p(f,g)p$. Bezóut's theorem says that $f \cdot g$ is always a cycle of order nm. The properties of intersection cycles tells us that

- $f \cdot g = g \cdot f$.
- $f \cdot gh = f \cdot g + f \cdot h$.
- $f \cdot (g + af) = f \cdot g$ if a is a form with g and af the same degree.

Max Noether concerned himself with the following situation. Suppose that f, g, and h are curves with $h \cdot f \geqslant g \cdot f$, so that h intersects f at every point that g intersects f, and with a higher intersection multiplicity at each of these points. We want to determine when there is a curve k with $k \cdot f = h \cdot f - g \cdot f$. This is easy if $h = gk$, or more generally if $h = gk + fl$. We shall find more general conditions for which we can find k.

Let p be a finite point in the projective plane, and f and g curves with no common component through p, and h another curve. We say that *Noether's conditions are satisfied at p* if $h_* \in (f_*, g_*) \subset \mathcal{O}_p(\mathbf{P}^2)$. This property is a local property around p which is invariant under projective transformations, because if we consider a projective transformation T with

$T^*Z = L$, then $T_i[x:y:z] = a_ix + b_iy + c_iz$, then if h has degree m, then $h(X,Y) = \sum d_{ij}X^iY^jZ^{m-i-j}$, then

$$\begin{aligned}(T^*h)_*(X,Y) &= (T^*h)(X,Y,1)\\ &= h(a_1X + b_1Y + c_1, a_2X + b_2Y + c_3, a_3X + b_3Y + c_3)\\ &= \sum d_{ij}(a_1X + b_1Y + c_1)^i(a_2X + b_2Y + c_2)^j(a_3X + b_3Y + c_3)^{m-i-j}\\ &= (a_3X + b_3Y + c_3)^m \sum d_{ij}\left(\frac{a_1X + b_1Y + c_1}{a_3X + b_3Y + c_3}\right)^i\left(\frac{a_2X + b_2Y + c_2}{a_3X + b_3Y + c_3}\right)^j\\ &= (a_3X + b_3Y + c_3)^m T^*h_*(X,Y)\end{aligned}$$

If T does not map the line at infinity to a line through p, and $T(q) = p$, then $a_3p_1 + b_3p_2 + c_3 \neq 0$, and so $(T^*h)_*$ and T^*h_* differ by a unit in $\mathcal{O}_q(\mathbf{P}^2)$. This implies that $h_* \in (f_*, g_*)$ if and only if $T^*h_* \in (T^*f_*, T^*g_*)$, which is equivalent to saying that $(T^*h)_* \in ((T^*f)_*, (T^*g)_*)$. Thus Noether's condition also makes sense for points at infinity.

Theorem 3.18. *If f, g, and h are projective plane curves, and f and g have no common components. Then $h \in (f,g) \subset k[X,Y,Z]$ if and only if Noether's conditions are satisfied for every point of intersection between f and g.*

Proof. If $h = kf + k'g$ in $k[X,Y,Z]$, then $h_* = k_*f_* + k'_*g_*$ at all points p. To prove the converse, we assume that f and g only intersect at finite points, which we conclude by a projective transformation. Noether's theorem implies that h_* is congruent to zero in $\mathcal{O}_p(\mathbf{P}^2)/(f_*, g_*)$ for each point of intersection between f and g. It follows from our discussion that h_* is congruent to zero in $k[X,Y]/(f_*, g_*)$, so $h_* = kf_* + k'g_*$. Then $Z^th = kf + k'g$ for some value t. We have seen in the proof of Bezóut's theorem that multiplication by Z is injective in $k[X,Y,Z]/(f,g)$, so $h = lf + l'g$ for some polynomials l and l'. We conclude that if f is degree m, g has degree n, and h has degree k, then $h = l_{k-m}f + l'_{k-n}g$. □

The power of Max Noether's theorem is that there are several easily verifiable and frequently occuring conditions under which Noether's condition is satisfied. Here are some of these criteria. Of course, Noether's condition is automatically satisfied at any point p with $p \notin Z(f) \cap Z(g)$ so we need only worry about the finitely many points in $Z(f) \cap Z(g)$.

Theorem 3.19. *If f, g, and h are two projective plane curves, then Noether's conditions are satisfied at p if any of the following are true:*

- *f and g meet transversally at p, and p lies on h.*
- *p is simple on f, and $I_p(h,f) \geqslant I_p(g,f)$.*
- *f and g have distinct tangents at p, and $m_p(h) \geqslant m_p(f) + m_p(g) - 1$.*

Proof. The first property is trivially implied by the third property, so it requires no proof. If p is simple on f, then

$$\text{ord}_p^f(h) \geqslant \text{ord}_p^f(g)$$

and this implies $(h_*) \subset (g_*) \subset \mathcal{O}_p(f_*)$. But since $\mathcal{O}_p(f_*)$ is isomorphic to $\mathcal{O}_p(\mathbf{P}^2)/(f_*)$ in the canonical fashion, this implies that

$$(h_*) \subset (g_*, f_*) \subset \mathcal{O}_p(\mathbf{P}^2)$$

For the third case, assume p is the origin, so that $m_p(h_*) \geqslant m_p(f_*)+m_p(g_*)-1$. We showed that this implied h_* was in $(X,Y)^t$ for large enough t, and $(X,Y)^t \subset (f_*, g_*)$ for large enough t, in our discussion of the properties of intersection numbers. □

Corollary 3.20. *If f and g meet in $(\deg f)(\deg g)$ distinct points, and h passes through these points, or if all intersections of f and g are simple on both curves, then there is a curve k such that $k \cdot f = h \cdot f - g \cdot f$.*

We now mention many geometric applications of Noether's theorem. They will not be required in later parts of this writing, but are certainly novel and interesting. To begin with, we consider a theorem which says 'entire functions' on a projective curve are constant.

Theorem 3.21. *Suppose C is an irreducible projective plane curve, and $u \in k(C)$ is defined at every point in C. Then $u \in k$.*

Proof. Suppose C is a degree n curve defined by an irreducible polynomial f. Write $u = v/w$, where v and w are each homogeneous polynomials of degree m in $k[X,Y,Z]$. For each p on C, there exists homogeneous polynomials a_p, b_p of degree m_p such that $vb_p - a_p w \in (f)$ with $b_p(p) \neq 0$. This means that in $\mathcal{O}_p$, $v \in (f, w)$. But this means Noether's condition is satisfied at p. If f does not divide w (which would imply $u = 0$), then Noether's theorem implies that we can write $v = c_1 f + c_2 w$ for two polynomials c_1 and c_2. But this implies that $u = v/w = c_2$. Since v and w have the same degree, c_2 is a constant, so u is a constant. □

Theorem 3.22. *Suppose $C{\cdot}C' = \sum_{i=1}^{9} P_i$. If Q is a conic meeting C at $p_1,\ldots,p_6$, which are simple points on C, then p_7, p_8, and p_9 are colinear.*

Proof. Let f describe C, let g describe Q, and let h describe C'. Then Noether's condition is satisfied at $p_1,\ldots,p_6$ for f and g, so we conclude that $h = af + bg$ for some polynomials a and b. Degree considerations imply that b is a linear form describing a line L, and that p_7, p_8, and p_9 are contained on L. □

Corollary 3.23 (Pascal's Theorem)**.** *If a hexagon is inscribed in an irreducible conic, then the opposite sides meet at collinear points.*

Proof. Let Q be the conic, and pick 6 points $p_1,\ldots,p_6$ on this conic. Let

$$p_7 = (p_1 \times p_2) \times (p_4 \times p_5),$$

$$p_8 = (p_2 \times p_3) \times (p_5 \times p_6),$$

and let

$$p_9 = (p_3 \times p_4) \times (p_6 \times p_1).$$

Let C be the union of the lines $p_1 \times p_2$, $p_3 \times p_4$, and $p_5 \times p_6$, and let C' be the union of the lines $p_2 \times p_3$, $p_4 \times p_5$, and $p_6 \times p_1$. Then $C \cdot C' = \sum_{i=1}^{9} p_i$. Since Q is irreducible, $p_1,\ldots,p_6$ are all simple on Q. The last theorem thus implies that p_7, p_8, and p_9 are all colinear. □

Corollary 3.24 (Pappus' Theorem)**.** *Let L_1 and L_2 be two lines, and choose three points p_1, p_2, p_3 on $L_1 - L_2$ and q_1, q_2, q_3 on $L_2 - L_1$.Let L_{ij} be the line form P_i to Q_j. For any permutation (i, j, k) of $(1, 2, 3)$, let r_k be the point on the intersection of L_{ij} and L_{ji}. Then r_1, r_2 and r_3 are colinear.*

Proof. Let Q be the degenerate conic formed from the union of L_1 and L_2. Then p_1, p_2, p_3, q_1, q_2, and q_3 are all simple points on Q. Let C be the cubic formed from the union of L_{12}, L_{23}, and L_{31}, and let C' be the union of L_{13}, L_{21}, and L_{32}. Then the last theorem implies that r_1, r_2, and r_3 are colinear. □

Theorem 3.25. *Let C, C', and C'' be cubics, with C irreducible. Let $C \cdot C' = \sum_{i=1}^{9} p_i$, where the p_i are simple (not necesssarily distinct) points on C. If $C \cdot C'' = \sum_{i=1}^{8} p_i + q$, then $q = p_9$.*

Proof. Suppose $q \neq p_9$. Choose a line L passing through p_9 but not passing through q. Then $C \cdot L = p_9 + r + s$ for some points r and s. Thus $C \cdot (L \cup C'') = \sum_{i=1}^{9} p_i + q + r + s$. We want to apply Max Noether's theorem to C, C', and $L \cup C''$. The conditions we have specified above (the simplicity of the p_i on C) allow us to do this, so we find there is a line L' such that $C \cdot (L' \cup C') = \sum_{i=1}^{9} p_i + q + r + s$. In particular, L' passes through q, r, and s. But this means L and L' share two points, which implies they are equal, contradicting the fact that L does not pass through q. □

Let us use this theorem to show that one can define an *abelian group structure* on any nonsingular cubic curve C. For any two points p and q on C, the line L between p and q intersects C at a unique third position r. We define a map $\phi : C \times C \to C$ by setting $\phi(p,q) = r$. To add a unit, we fix a point o, and define

$$p \oplus q = \phi(o, \phi(p,q)).$$

This operation is certainly commutative. It is also easy to check that o is an identity. If $o' = \phi(o,o)$, the additive inverse of p is given by $\phi(o',p)$, since

$$\phi(o, \phi(p, \phi(o',p))) = \phi(o,o') = o \oplus o = o.$$

The only hard property to check is that addition is associative. Fix p, q, r on a common nonsingular cubic C. Write

$$\begin{aligned} L_1 \cdot C &= p + q + s' \\ M_1 \cdot C &= o + s' + (p \oplus q) \\ L_2 \cdot C &= (p \oplus q) + r + t' \\ M_2 \cdot C &= q + r + u' \\ L_3 \cdot C &= o + u' + (q \oplus r). \\ M_3 \cdot C &= p + u + t''. \end{aligned}$$

Since $(p \oplus q) \oplus r = \phi(0, t')$, and $p \oplus (q \oplus r) = \phi(0, t'')$, it suffices to prove that $t' = t''$. Let $C = L_1 \cup L_2 \cup L_3$ and let $C' = M_1 \cup M_2 \cup M_3$. Applying the previous theorem shows that $t' = t''$.

Remark. Let C be any irreducible cubic. If we choose o to be a simple point, then the set C° of simple points on C can be made into a group in the same way as a nonsingular curve, by defining $p \oplus q = \phi(o, \phi(p,q))$. The only thing to notice here is that $\phi(C^\circ \times C^\circ) \subset C^\circ$ because if p and q are simple points on a curve, and $C \cdot L = p + q + r$, then $m_p(C) + m_q(C) + m_r(C) \leqslant 3$, and since $m_p(C) = m_q(C) = 1$ this implies that $m_r(C) = 1$, so r is simple.

Chapter 4

Where to Put It?

4.1 Birational Classification of Algebraic Curves

We briefly mentioned the idea of a rational curve in the introductory chapter, that is, an algebraic curve which can be parameterized by a rational function of a single argument. We now define this precisely. A rational curve is a curve C whose field of functions is isomorphic to the field of fractions in a single variable, i.e. that $k(C)$ is isomorphic to $k(t)$. Our previous discussion implies that this means precisely that there is a set of rational functions $f_1(t),\dots,f_n(t)$ in a single variable, whose image is Zariski dense in C, which have an inverse set of rational functions. However, the existence of the parameterization in one direction implies the rational parameterization in the other, because a map $f : \mathbf{A}^1 \to C$ induces a homomorphism $T : k(C) \to k(t)$. There is a theorem (which we won't rely upon for future except in examples) due to Lüroth, which says that every field between k and $k(t)$ is either isomorphic to k, or isomorphic to $k(t)$, so that the homomorphism T implies that either C is a discrete set of points, or C is a rational curve, and we needn't check that the parameterization $f_1(t),\dots,f_n(t)$ has an inverse, because the existence of any map implies the existence of a birational parameterization.

Example. *We shall now prove the circle is a rational curve. For convenience, we consider the circle defined by the equation* $(X-1)^2 + Y^2 = 1$*. Then, other than the origin, the lines* $Y = tX$ *also touch the circle at a unique position, and each point on the circle other than the origin lies on one such line. For each* t*, this unique point corresponds to the nonzero solutions of* $(X-1)^2 + (tX)^2 = 1$, *and since this is equivalent to $(1+t^2)X^2 = 2X$, we find that we can let*

$$X = \frac{2}{1+t^2} \qquad Y = \frac{2t}{1+t^2}$$

which gives a rational parameterization of the circle, because the image contains all points on the circle but the origin, and if a set contains all but finitely many points of a variety, it is Zariski dense in that variety. Essentially, the same technique of parameterizing a curve by slope works in any planar curve of degree two, showing that the corresponding field of fractions is isomorphic to $k(t)$, and that the curve is rational. In terms of our introductory exposition, this implies that for any curve of the form $Y^2 = aX^2 + bX + c$, we can find indefinite integrals of rational functions of the form $f(x, \sqrt{ax^2+bx+c})$, where f is a rational function.

Example. *If f defines an irreducible curve of degree n, which is composed of monomials of degree $n-1$ and n. Then projection form the origin gives a birational parameterization of $Z(f)$. This follows because the equation $Y = tX$ intersects the curve in a unique position. If $f(X,Y) = \sum a_i X^i Y^{n-1-i} - b_i X^i Y^{n-i}$, then $a(t)X^{n-1} - b(t)X^n = 0$, where $a(t) = \sum a_i t^{n-1-i}$ and $b(t) = \sum b_i t^{n-i}$ holds only when*

$$X = \frac{a(t)}{b(t)} \qquad Y = \frac{ta(t)}{b(t)}$$

and this gives a birational map with inverse Y/X. This generalizes the example of the circle.

Example. *Suppose that f is an irreducible curve composed instead of monomials of degree $n-2$, $n-1$, and n. If we write*

$$f(X,Y) = \sum a_i X^i Y^{n-i} + \sum b_i X^i Y^{n-1-i} + \sum c_i X^i Y^{n-2-i}$$

Then applying the strategy by setting $Y = tX$ and finding intersections gives $a(t)X^n + b(t)X^{n-1} + c(t)X^{n-2} = 0$, which is equivalent to $a(t) + b(t)X + c(t)X^2$. We can rewrite this as

$$(2aX+b)^2 = b^2 - 4ac$$

If we set $s = 2aX + b$, then $X = (s-b)/2c$ and $Y = t(s-b)/2c$, so we find that our curve is birationally equivalent to the curve $s^2 = b^2 - 4ac$ in the (s,t) plane. A curve of this form is called a hyperelliptic curve. *If k is an algebraically* *closed field, and $b^2 - 4ac$ has degree $2m$, then $b^2 - 4ac = g(t)(t-a)$ for some $a \in k$. Dividing both sides of the equation by $(t-a)^{2m}$ gives*

$$\left(\frac{s}{(t-a)^m}\right)^2 = \frac{g(t)}{(t-a)^{2m-1}}$$

Writing $\eta = s/(t-a)^m$ and $\xi = (t-a)^{-1}$, we obtain a birational equivalence with the hyperelliptic curve and the curve in the (η, ξ) plane given by $\eta^2 = h(\xi)$, where h is a polynomial of degree $< 2m$. As an example of this technique, over a field of characteristic $\neq 2$, an irreducible cubic cube is birationally equivalent to a curve of the form $y^2 = f(x)$, where f is a polynomial of degree less than or equal to 4, and the additional technique shows that over an algebraically closed field, we can assume f has degree less than or equal to 3. If it has degree 3 we can assume the leading coefficient has degree one, so the curve is defined by an equation of the form $y^2 = x^3 + ax^2 + bx + c$, which is called the Weirstrass normal form *of the cubic. If k does not have characteristic 3, then $x \mapsto x - a/3$ shows we can assume the curve is defined by the equation $y^2 = x^3 + bx + c$. This begins the classification of cubic curves.*

Example. *The map $f : t \mapsto (\cos t, \sin t)$ is a surjective map from $\mathbf{A}^1$ to S^1, and the induced map f^* gives an isomorphism between the coordinate ring $\mathbf{C}[S^1]$ and the algebra of functions $\mathbf{C}[\cos t, \sin t]$, obtained by mapping X to the function $\cos t$, and Y to the function $\sin t$. Correspondingly, this implies that the field $\mathbf{C}(S^1)$ of rational functions on S^1 is isomorphic to the ring $\mathbf{C}(\cos t, \sin t)$ of rational functions of the cosine and sine functions. This explains why the analysis of the functions $\cos t$ and $\sin t$ is often reduced to analysis of certain equations of algebra.*

Theorem 4.1. *If $f : \mathbf{A}^1 \to C$ is any nonconstant rational map, then the inverse image of any point $x = (x_1, \dots, x_n) \in C$ is finite.*

Proof. If $f_i = g_i/h_i$, then for a fixed value of x, $g_i(t)x_i = h_i(t)$ has finitely many solutions unless $g_i x_i = h_i$, in which case we conclude that if $x_i = 0$, then $h_i = 0$, which is impossible, and if $x_i \neq 0$, then $g_i/h_i = 1/x_i$ is a constant map. This cannot hold for all i, because f is nonconstant, so the inverse image of x must be finite. □

Part II

Modern Algebraic Geometry

We now enter a modern study of algebraic geometry, which aims to try and 'abstract' the general properties that varieties in affine and projective space possess, so that we can best concentrate on the properties of that variety up to isomorphism. Like the study of differential manifolds, this introduces many technical questions which will take a while to fix. But armed with the intuition from a classical study of varieties, we can proceed to define the necessary algebraic properties of isolated varieties. With these technologies, we can determine the structure of additional spaces which 'look' like varieties, for instance, like the structure of a variety with a single point removed, or a surface with a finite number of curves on the surface removed. Thus the study of these abstract spaces is worth the effort to understand.

4.2 The Zariski Topology

The first trick to isolating what it means for a subset of an abstract 'variety space' to be a subvariety. On varieties, this is provided by the Zariski topology. Since a finite union of varieties is a variety, and arbitrary intersections of varieties are varieties (though every intersection is really a finite intersection in disguise, by the Noetherian property of the coordinate ring of affine space), and thus the family of varieties forms all the structures of a family of closed sets on a space. On $\mathbf{A}^n$, we define the Zariski topology by this structure, so a subset is closed if and only if it is a variety. The Zariski topology on a variety $V \subset \mathbf{A}^n$ is obtained by taking the subspace topology, and since V is a closed subspace, the closed subsets in the relative topology are precisely the subvarieties of V.

Example. *The subvarieties of $\mathbf{A}^1$ are precisely $\mathbf{A}^1$ itself, $\emptyset$, and finite point sets. Thus the open sets of $\mathbf{A}^1$, under the Zariski topology, are the emptyset, $\mathbf{A}^1$, and sets which complement consists of a finite number of points (the finite complement topology). Notice this topology is not Hausdorff.*

A topological space is called *irreducible* if it cannot be written as the union of two proper nonempty closed subsets X_1 and X_2. In particular, a variety is irreducible if and only if it is irreducible in the Zariski topology.

Theorem 4.2. *Any non-empty open subset of an irreducible space X is dense and irreducible.*

Proof. If U is open, then $\overline{U} = X$, for otherwise $X = \overline{U} \cup U^c$. If $U = A \cup B$, where A and B are nonempty closed proper subsets of U, then $X = \overline{A} \cup \overline{B}$, where $\overline{A}$ and $\overline{B}$ are proper subsets because $\overline{A} \cap U = A$ and $\overline{B} \cap U = B$, which is impossible. □

Corollary 4.3. *If Y is irreducible in some space X, then $\overline{Y}$ is irreducible.*

Proof. If $\overline{Y} = A \cup B$, where A and B are proper and closed in $\overline{Y}$, then $Y = (A \cap Y) \cup (B \cap Y)$, and $A \cap Y$ and $B \cap Y$ are both proper for if $Y \subset A$, then $\overline{Y} \subset A$. □

An open subset of an affine variety is known as a *quasi affine variety*. Breaking our variety up into irreducible subvarieties, we see the intersection of a quasi affine variety is dense in each subvariety, hence dense in the entire variety, so really quasi affine varieties contain all the algebraic information that a normal variety posesses.

A *Noetherian* topological space is a space for which there exists no infinite decreasing family of closed subsets

$$A_0 \supsetneq A_1 \supsetneq A_2 \supsetneq \dots$$

Under the Zariski topology, $\mathbf{A}^n$ is a Noetherian topological space, for a decreasing chain of closed subsets corresponds to an increasing chain of radical ideals in $k[X_1, \dots, X_n]$, which cannot be infinite.

Theorem 4.4. *Every nonempty closed subset of a Noetherian topological space can be uniquely decomposed into the finite union of irreducible closed subsets, no one containing the other.*

Proof. If such a decomposition does not exist, it implies an infinite decreasing chain of closed subsets, which is impossible. To prove uniqueness, suppose that

$$A_1 \cup \dots \cup A_N = B_1 \cup \dots \cup B_M$$

For each n, $A_n = \bigcup (B_m \cap A_n)$, and since $B_m \cap A_n$ is closed, we must have $B_m = A_n$ for some n and m, since otherwise all subsets are proper. Proceeding by induction, we may remove A_n and B_m from the union and still have equality, so that the decomposition is unique. □

These arguments are simple reformulations of the fact that varieties can be broken up uniquely into irreducible subvarieties.

4.3 The Dimension of a Variety

If X is a topological space, we define the *dimension* of X as the supremum over all n such that there exists a chain $X_0 \subsetneq X_1 \subsetneq \dots \subsetneq X_n$ of closed irreducible subsets of X. Suprisingly, this notion of dimension well measures the dimension of varieties because of their discreteness – intuitively, a proper irreducible variety contained in some irreducible variety drops the dimension of the space down at least by one, so a curve drops down to a point, and has dimension one, a surface drops down to a curve, and then to a point, hence having dimension two, and so on and so forth.

Example. *The space* $\mathbf{A}^1$ *is one dimensional, because the only irreducible subsets of* $\mathbf{A}^1$ *are* $\mathbf{A}^1$ *itself, and points.*

The dimension of a variety is best described in terms of it's coordinate ring. If A is a ring, it's *Krull dimension* is the supremum over all n of an increasing chain of nonempty prime ideals $\mathfrak{a}_0 \subsetneq \dots \subsetneq \mathfrak{a}_n$. In particular, we find that the dimension of V is *precisely* the krull dimension of $k[V]$.

Theorem 4.5. *Under the Zariski topology,* $\dim V = \dim k[V]$.

Proof. Since all prime ideals are radical, the nullstellensatz implies that a increasing chain of nontrivial prime ideals corresponds to a decreasing family of nonempty irreducible varieties. This completes the proof. □

To calculate dimensions, we will rely on the following calculational tricks from commutative algebra. Let A be an integral domain which is a finitely generated k algebra for some field k. Then $\dim A$ is the trancendence degree of the quotient field $k(A)$ over k, and for any prime ideal $\mathfrak{p}$, $\dim A = \dim \mathfrak{p} + \dim A/\mathfrak{p}$ (this result says that when forming prime ideals, we can never find suboptimal chains – the dimension of A is the number of elements in any maximal chain of prime ideals).

Theorem 4.6. *Under the Zariski topology,* $\dim \mathbf{A}^n = n$.

Proof. The field $k(\mathbf{A}^n) = k(X_1, \dots, X_n)$ has a trancendence base $X_1, \dots, X_n$, which immediately implies the result because $k[X_1, \dots, X_n]$ is obviously a finitely generated k algebra. □

Theorem 4.7. *If X is a quasiaffine variety, then* $\dim X = \dim \overline{X}$.

Proof. Let $\varnothing \subsetneq X_0 \subsetneq X_1 \subsetneq X_2 \subsetneq \cdots \subsetneq X_N$ be a sequence of closed irreducible subsets of X. Then $\overline{X_0} \subsetneq \overline{X_1} \subsetneq \overline{X_2} \subsetneq \cdots \subsetneq \overline{X_N}$ are closed irreducible subsets of $\overline{X}$, giving us the bound $\dim X \leqslant \dim \overline{X}$, which implies that in particular, $\dim X$ is finite. If the X_n are chosen to be maximal, then in particular we know X_0 is just a point. We claim this chain is maximal. The increasing family of closed sets corresponds to a decreasing sequence of prime ideals

$$\mathfrak{p}_0 \supsetneq \cdots \supsetneq \mathfrak{p}_N$$

in $k[V]$. We know $\mathfrak{p}_0$ is maximal, TODO: Figure out later. □

Another trick to use, known as krull's hauptidealsatz, says that if A is Noetherian, and $f \in A$ is neither a unit nor a zero divisor, then the smallest prime ideal containing f has dimension 1. In addition, if A is a Noetherian integral domain, then A is a factorial ring if and only if every prime ideal with dimension 1 is principal.

Theorem 4.8. *An irreducible variety $V \subset \mathbf{A}^n$ has dimension $n-1$ if and only if it is the zero set of a single irreducible polynomial.*

Proof. krull's hauptidealsatz tells us that (f) has dimension one if f is irreducible, which implies that $n = \dim k[X_1, \dots, X_n] = \dim(f) + \dim k[V] = 1 + \dim k[V]$, hence $\dim k[V] = n-1$. Conversely, if V has dimension $n-1$, then $\dim k[X_1, \dots, X_n] = \dim I(V) + \dim k[V]$, implying $\dim I(V) = 1$. But since $k[\mathbf{A}^n]$ is a factorial ring, and $I(V)$ is prime since V is irreducible, we conclude that $I(V) = (f)$ for some irreducible f. □

Chapter 5

Schemes

In almost all aspects of modern geometry, functions play a central role.

- In topology, each topological space X we associate the ring $C(X)$ of complex-valued continuous functions.
- In differential geometry, each smooth manifold M is equipped with a ring $C^\infty(M)$ of infinitely differentiable complex-valued functions. A complex manifold is equipped with a further ring of holomorphic functions, which forms a subring of the C^∞ functions.
- In algebraic geometry, each affine variety V has an associated coordinate ring $k[V]$ of regular functions.

Given a geometric morphism between two spaces, we can naturally associate a ring morphism between the functions on these spaces; often, this process is invertible, so the structure of the ring of functions gives all the required geometric properties of space. One can think of the theory of schemes as an inversion of this relationship. Rather than associating a commutative ring with a geometric space, we associate a geometric space, known as an *affine scheme*, with each commutative ring, upon which the commutative ring operates as functions. Suprisingly, this process is useful even if we start with a space X, since the affine scheme associated to the ring of functions on X may contain more points than those we started with, and these points may describe useful facts about the geometry of the space X we started with!

From the point of view of algebraic geometry, working over schemes rather than just varieties offers numerous technical advantages:

- Manifolds are defined by gluing interlinking open balls together in a topologically compatible way. Smooth manifolds are obtain by gluing balls in a smoothly compatible way. This makes it easy to glue together two manifolds. On the other hand, the category of varieties seems very rigid, and it isn't natural to glue two varieties together, or what way we should do this to obtain a variety. The theory of schemes is much more fluid.

- The theory of schemes enables us to more easily describe 'multiplicities' which are not so easy to describe with ordinary varieties. For instance, the variety $\{0\}$ in $\mathbf{A}^1$ has coordinate ring $k[X]/(X)$. The ideal (X^2) in $k[X]$ is not radical, and therefore does not correspond to a variety, but we can think of the ring $k[X]/(X^2)$ as functions describing the geometric object which consists of a 'fat point' $\{0\}$ in $\mathbf{A}^1$, which we can think of as a double point rather than a single point.

Since Grothendieck introduced schemes in 1960, schemes have become the standard way to discuss most of modern algebraic geometry.

5.1 Spectra of Rings

One of the first main ideas of scheme theory is to become blind to the geometry of a space, and see if you can hear the geometry purely through the algebraic structure of the functions defined in terms of this geometry. For instance, if X is a compact Hausdorff space, then the maximal ideals of $C(X)$ are in one to one correspondence with the points in X, so one can recover X purely from the ring structure of $C(X)$. The nullstellensatz says that the points of a variety V are in one to one correspondence with the maximal ideals of the coordinate ring $k[V]$. In this case, we see that the geometry of the situation is captured perfectly by the algebra of functions on the space. But things get funkier in other situations; If one considers an arbitrary topological space X, then the maximal ideals of $C(X)$ are in one to one correspondence with the maximal ideals of $C(\beta X)$, where βX is the *Stone-Cech compactification* of X, a Hausdorff compact space whose points are in one to one correspondence with these maximal ideals. Though we start with only the points in X, the continuous functions on X have qualitative features that are unable to be discussed by the individual values of the functions on X, but which *are* captured by their values on βX. Thus, if we care about continuous functions on a space X, we may be naturally lead through the algebra of continuous functions to look at the Stone-Cech compactification βX. Algebra leads to a completely different perspective on the geometry of the situation. This is one of the main reasons to take a look at the maximal ideals of a ring as points in a space.

In scheme theory, we not only introduce maximal ideals for points in a space, but also 'generic points' corresponding to collections of points. If A is a commutative ring, then we define the *spectrum* $\text{Spec}(A)$, which we also denote by X_A, or X if the ring is obvious, which is a space whose points are *prime* ideals in A (we will justify why we don't only consider maximal ideals later). Given a prime ideal $\mathfrak{p}$, the ring $A/\mathfrak{p}$ is an integral domain, and can therefore be localized into it's field of fractions, which we will denote by $k(\mathfrak{p})$. Each element $f \in A$ operates as a 'function' on X_A, mapping an ideal $\mathfrak{p}$ to $f(\mathfrak{p}) = f + \mathfrak{p} \in k(\mathfrak{p})$. We therefore call f a *regular function* on X. For each subset $S \subset A$, we have a 'zero set'

$$Z(S) = \{\mathfrak{p} : S \subset \mathfrak{p}\} = \{\mathfrak{p} \in : (\forall f \in S : f(\mathfrak{p}) = 0)\}$$

which is of course equal to the zero set of the ideal generated by S, and we define the *Zariski topology* on X_A by declaring the sets $Z(S)$ to be all closed sets. We now consider examples, and then explain the reasonings for this choice of definitions.

Example. *Let V be a variety over an algebraically closed field. Then the nullstellensatz tells us that the prime ideals of $k[V]$ are in one to one correspondence with the irreducible subvarieties of V, so the spectrum consists of all points, which correspond to maximal ideals, in addition to points corresponding to irreducible curves, surfaces, and higher dimensional irreducible varieties contained in V. Restricted to the maximal ideals of $k[V]$, the Zariski topology is the same Zariski topology encountered in the study of varieties. However, the topology of $X_{k[V]}$ also has the property that any closed set in the original Zariski topology containing all points on an irreducible variety also contains the variety itself as a 'generic point'. Over the maximal ideals $\mathfrak{m}$ corresponding to some point $p \in V$, $k(\mathfrak{m})$ is isomorphic to k, and the map $f \mapsto f(\mathfrak{m})$ is essentially the same as the map $f \mapsto f(p)$. If $\mathfrak{p}$ is a prime ideal corresponding to an irreducible subvariety W of V, then $k[V]/\mathfrak{p}$ is isomorphic to $k[W]$ in such a way that $f(\mathfrak{p})$ can be viewed as the restriction of f to W.*

Example. *If X is a compact Hausdorff space, then the closed (under the supremum norm) prime ideals of $C(X)$ are in one to one correspondence with the* *closed subsets of X. The Zariski topology restricted to the maximal ideals, which correspond to points in X, give us precisely the same topology on X. This is a nice exercise in topology, using the fact that every compact Hausdorff space is normal, so that we can apply Urysohn's lemma. Similar to the case of varieties, if $\mathfrak{p}$ is a prime ideal corresponding to a closed subset Y, then for a function $f \in C(X)$, $f(\mathfrak{p}) \in C(X)/\mathfrak{p}$ corresponds to the restriction of f to $C(Y)$, and so a closed set in the Zariski topology of Spec$(C(X))$ also contains points corresponding to closed sets contained within the space.*

The reason for the study of maximal ideals as an underlying geometry to ring theory has already been justified. The reason we add the prime ideals is so that we can understand collections of points as if they were a single, unified point. In classical algebraic geometry, geometers attempted to prove things about 'generic points' on a variety by applying properties of the points which did not depend on the particular point in question. Often, these results turned out to fail at certain extreme points, but held true on a dense, open subset of the variety. Using the theory of schemes, we can formalize this process by thinking of a prime ideal $\mathfrak{p}$ as a 'generic point' representing the collection of maximal ideals $\mathfrak{p} \subset \mathfrak{m}$, or the points in the closure of the point set $\{\mathfrak{p}\}$ in the Zariski topology. As an example, the prime ideal in $k[V]$ corresponding to a subvariety W of V is a generic point for all points in W, and understanding this ideal will give us results about most of the points within W. Similarily, the 'generic point' corresponding to a closed subset Y of X allows us to prove things about most points in $C(Y)$ generically.

Every closed set of X is the intersection of zero sets of the form $Z(f)$ for a single regular function f. It follows that every open set is the union of sets of the form $Z(f)^c$, often denoted X_f and called the *distinguished* open subset associated with f. Points in this set consist of prime ideals not containing f, and therefore are in one-to-one correspondence with prime ideals in the localization A_f obtained by adding an inverse of f to the ring. The finite intersection of distinguished open sets is distinguished, because a prime ideal doesn't contain $g = f_1 \dots f_n$ if and only if it doesn't contain $f_1, \dots, f_n$. If X is compact, in the spectrum of $C(X)$ all open sets are distinguished, but in general this need not be the case.

5.2 Sheaf Theory

We now consider sheaves, which are abstract ways to understand local geometric structure through locally defined functions. The main example of a sheaf is the sheaf of continuous functions, which associates to each open subset U of a topological space the algebra $C(U)$ of complex-valued functions, and for each $U \subset V$, a restriction map $f \mapsto f|_V$ from $C(V)$ to $C(U)$. One can add structure to a geometric space by distinguishing a special 'subsheaf' of functions. For instance, one can see the C^∞ structure of a manifold M as declaring a subsheaf of smooth functions in the space of all topological functions. The choice of sheaf must be made such that the space is locally diffeomorphic to affine space. Analogously, we will define a *scheme* as a space equipped with a sheaf which makes it locally isomorphic to the spectrum of a ring.

A *presheaf* $\mathcal{F}$ on a topological space X associates with each open set $U \subset X$ an object $\mathcal{F}(U)$, also denoted $\Gamma(U, \mathcal{F})$, known as the *sections* of the sheaf over U, and for each $U \subset V$, a morphism $\rho_{VU} : \mathcal{F}(V) \to \mathcal{F}(U)$, which are mutually commuting in the sense that $\rho_{WV} \circ \rho_{UW} = \rho_{UV}$ for all $V \subset W \subset U$, and ρ_{UU} is the identity morphism. We often denote $\rho_{UV}(f)$ as $f|_U$, to reflect the fact that we view these as models of restriction maps of functions onto smaller domains. If we have a presheaf of sets (or a presheaf in any category whose objects can be interpreted as sets), we can introduce the sheaf axiom, which consists of two parts; let U be an open set covered by a family of open sets $\{U_\alpha\}$:

- If $f, g \in \mathcal{F}(U)$, and $f|_{U_\alpha} = g|_{U_\alpha}$ for all α, then $f = g$.
- If we have a family of functions $f_\alpha \in \mathcal{F}(U_\alpha)$ which are *compatible*, so that $f_\alpha|_{U_\alpha \cap U_\beta} = f_\beta|_{U_\alpha \cap U_\beta}$ for all α and β, then there is a $f \in \mathcal{F}(U)$ for which $f|_{U_\alpha} = f_\alpha$ for each α.

Note that the second point implies the first if we can show the f constructed is unique. A presheaf is a *sheaf* if it satisfies the sheaf axioms. Intuitively, the sheaf axiom algebraically represents the geometric property of functions which locally satisfy some property – a function is continuous / smooth / holomorphic / etc if and only if it is continuous in a cover of open neighbourhoods. In almost all the sheaves we will be interested in, each $\mathcal{F}(U)$ will have the structure of a commutative ring, and each restriction will be a ring homomorphism. The next result is confusing, but logically valid.

Lemma 5.1. *For any sheaf $\mathcal{F}$, $\mathcal{F}(\varnothing)$ is a singleton.*

Proof. $\varnothing$ is covered by the family consisting of no sets at all. In the sheaf axiom, specifying no element is therefore equivalent to specifying a function for each element of the family covering $\varnothing$, and therefore the sheaf axiom implies that $\mathcal{F}(\varnothing)$ consists of a single element (because the condition when the family is empty implies there exists a unique element of $\mathcal{F}(\varnothing)$ satisfying no conditions at all). □

A morphism between two sheaves $\mathcal{F}$ and $\mathcal{G}$ on the same topological space X is a family of morphisms $f : \mathcal{F}(U) \to \mathcal{G}(U)$ for each open set U commuting with the restriction maps in the obvious way. This gives us a category to work on the family of schemes over a particular topological space. We automatically get some constructions which are very useful when working with sheaves.

Example. *If $\mathcal{F}$ is a sheaf on X, and U is an open subset of X, we can define a sheaf $\mathcal{F}|_U$ on U by restricting the domain of definition for the sheaf. If $f : X \to Y$ is a continuous map, then we can define the pushforward sheaf $f_*\mathcal{F}$ on Y by setting $(f_*\mathcal{F})(U) = \mathcal{F}(f^{-1}(U))$. This is functorial in the sense that if φ is a morphism between $\mathcal{F}$ and $\mathcal{G}$, then it induces a morphism from $f_*\mathcal{F}$ and $f_*\mathcal{G}$.*

Example. *If $\mathcal{F}$ and $\mathcal{G}$ are presheaves of abelian groups, we can take direct sums $\mathcal{F} \oplus \mathcal{G}$ by taking the pointwise direct sum (and the direct sum of sheaves will also be a sheaf).*

Example. *If $\mathcal{F}$ and $\mathcal{G}$ are presheaves, we can define the tensor product pointwise as well, setting $(\mathcal{F} \otimes \mathcal{G})(U) = \mathcal{F}(U) \otimes \mathcal{G}(U)$. Note that in this definition, the tensor product of sheaves might not be a sheaf. An example is provided by letting $\mathcal{F}$ be the sheaf consisting of all locally constant integer valued functions (functions constant in a neighbourhood of each point) on some space X. If X has two connected components U and V, then the groups $(\mathcal{F} \otimes \mathcal{F})(U)$ and $(\mathcal{F} \otimes \mathcal{F})(V)$ are both isomorphic to $\mathcal{F}(U)$ and $\mathcal{F}(V)$ respectively, since they are one dimensional.*

if $f_U : U \to \mathbf{Z}$ and $f_V : V \to \mathbf{Z}$ are given by $f_U(x) = f_Z(y) = 1$, then there are infinitely many $f \in (\mathcal{F} \otimes \mathcal{F})(U)$ with $f|_U = (1 \otimes f_U)$ and $f|_V = (1 \otimes f_V)$.

the presheaf $\mathcal{F} \otimes \mathcal{F}$

$(\mathcal{F} \otimes \mathcal{F})(U)$ can be seen as the space of all functions on $U \times U$ locally constant in each variable, because such a function f assigns to each pair of connected components U_i and U_j of U, an integer a_{ij}, and then $f = \sum a_{ij}(\chi_{U_i} \otimes \chi_{U_j})$. The reason $\mathcal{F} \otimes \mathcal{F}$ doesn't have the sheaf property is that if X has two connected components U and V, then $(\mathcal{F} \otimes \mathcal{F})(U)$ and $(\mathcal{F} \otimes \mathcal{F})(V)$ are both isomorphic to $\mathbf{Z}$, whereas $(\mathcal{F} \otimes \mathcal{F})(X) \equiv \mathbf{Z}^2 \otimes \mathbf{Z}^2$, which is a

since $U \times U$ and $V \times V$ do not cover $X \times X$, we cannot uniquely extend locally constant functions to the entire space. Later on, we will fix this by modifying the definition of the tensor product of two sheaves to produce a sheaf, by sheafifying the definition of the tensor product here. In this new definition, for the sheaf $\mathcal{F}$ we discussed here, $\mathcal{F} \otimes \mathcal{F}$ will be isomorphic to $\mathcal{F}$.

A *subsheaf* of a sheaf $\mathcal{F}$ is a choice of subgroups $\mathcal{G}(U)$ of $\mathcal{F}(U)$ for each open set U, which is closed under restriction and satisfies the sheaf axiom. The kernel of a morphism of sheafs is a subsheaf, whereas the image and cokernel of the morphism may not satisfy the sheaf axiom.

Example. *If X is a topological space, we have a sheaf of continuous functions on X, obtained by the map $U \mapsto C(U)$, and $f \mapsto f|_U$ the standard restriction morphism. The family of functions which are locally constant forms a subsheaf, known as the constant sheaf on X.*

Example. *Let $f : Y \to X$ be a continuous map. A section of f on $U \subset X$ is a map $s : U \to Y$, such that $f \circ s$ is the identity map. One can think of this as a generalization of the theory of vector fields on manifolds, which are the sections where X is a manifold M, Y the tangent space TM, and f the projection of TM onto M.*

Example. *If V is a variety over an algebraically closed field, then the association of regular functions $\mathcal{O}(U)$ (those functions locally definable as rational polynomials) to each Zariski open set $U \subset V$ is a sheaf. To verify the sheaf axiom, we note that if f and g are two polynomials which agree on an open set U, then they must be equal, because they agree on infinitely many points. Thus if we have a family of locally defined rational functions f_α on U_α which agree on the intersections, the polynomials must all be equal to one another, and thus extend to a unique polynomial map on the whole space.*

Since a presheaf describes local objects, we are able to obtain infinitisimal objects by taking limits locally around points. If $\mathcal{F}$ is a presheaf on

X, and $p \in X$, we define the *stalk* at p to be the colimit of the groups $\mathcal{F}(U)$, for $p \in U$, as $U \to p$, denoted $\mathcal{F}_p$. One can view the stalk as the ring of germs locally defined in a neighbourhood of p. Indeed, if we consider the sheaf of analytic functions on some analytic manifold, then the presheaf at a point can be described as the ring of power series at that point. If $\mathcal{F}$ satisfies the sheaf axiom, then a section $s \in \mathcal{F}(U)$ is determined by it's image $s_p \in \mathcal{F}_p$, for each $p \in U$. This is because if $s_p = t_p$, then there is a neighbourhood V containing p such that $s|_V = t|_V$, and by letting p range over all points in U, we conclude there is a cover of U by neighbourhoods upon which s and t agree, hence $s = t$ by the uniqueness of extension. If $f : \mathcal{F} \to \mathcal{G}$ is a morphism of schemes, then it induces a map $f_p : \mathcal{F}_p \to \mathcal{G}_p$ for each p. The next theorem shows that the sheaf axiom reduces the analysis of presheafs to a sheaf's behaviour at a stalk.

Theorem 5.2. *If $\mathcal{F}$ and $\mathcal{G}$ are both sheaves, then f is an isomorphism if and only if each f_p is an isomorphism.*

Proof. If f is an isomorphism, then the isomorphisms $\mathcal{F}(U) \equiv \mathcal{G}(U)$ descend to an isomorphism between $\mathcal{F}_p$ and $\mathcal{G}_p$ as $U \to p$. Conversely, suppose each f_p is an isomorphism. If there is $s \in \mathcal{F}(U)$ with $f(s) = 0$. Then $f_p(s_p) = f(s)_p = 0$, hence $s_p = 0$ for all $p \in U$, and the sheaf axiom on $\mathcal{F}$ implies $s = 0$, so f is injective. Conversely, if $t \in \mathcal{G}(U)$ is a section, then for each $p \in U$, because f_p is an isomorphism, each p has a neighbourhood V_p and a section $s^p \in \mathcal{F}(V_p)$ such that $f(s^p)_p = t_p$, so there is a neighbourhood of p (which by restriction, we may assume to be V_p) upon which $f(s^p) = t|_{V_p}$. But then since f is injective, the s^p are a compatible family on sets covering U, so they extend to a unique $s \in \mathcal{F}(U)$ with $s|_{V_\alpha} = s^p$ by the sheaf axiom on $\mathcal{F}$, and then the sheaf axiom on $\mathcal{G}$ implies $f(s) = t$. □

Thus associated with each $s \in \mathcal{F}(U)$, we have an element of $\prod_{p \in U} \mathcal{F}_p$. Conversely, given an element $s \in \prod_{p \in U} \mathcal{F}_p$, we can construct an element $t \in \mathcal{F}(U)$ with these elements of stalks precisely when the element s are *compatible*, in the sense that we can locally choose t in the sheaf for which $t_p = s(p)$. We now show that this can be given a topological interpretation. The disjoint of union of the stalks $\overline{\mathcal{F}}$ over X can be given a topological structure by considering the strongest topology such that for every open $U \subset X$ and for every section $s \in \mathcal{F}(U)$, the map $\overline{s}(p) = s_p$ is continuous. The projection map $\pi : \overline{\mathcal{F}} \to X$ is then continuous, because for any open set U

and any section s on an open set $V \subset U$, $s^{-1}(\pi^{-1}(V)) = V$, so $\pi^{-1}(V)$ is open. A basis for this topology is given by the family of sets $V_{U,s} = \{(x, s_x) : x \in U\}$, because if t is any section on $W \subset U$, then $t^{-1}(V_{U,s}) = \{x \in V : s_x = t_x\}$, which is open because if $s_x = t_x$, then s and t agree in some neighbourhood of x, hence $s_y = t_y$ in a neighbourhood of x. Conversely, if U is any open set of $\overline{\mathcal{F}}$ containing a value s_x for some section $s \in \mathcal{F}(W)$, and if $W' = \overline{s}^{-1}(U)$, then U contains $V_{W',s}$, which contains s_x. If $s : U \to \overline{\mathcal{F}}$ is a continuous section, and $s(x) = t_x$ for some t defined at V, then $s(x) = t_x$ in a neighbourhood of x, since $s^{-1}(V_{W,t})$ is an open set containing x. But now using the sheaf axiom over all x, we can put these sections to gether to find a section $s' \in \mathcal{F}(U)$ such that $s(x) = s'_x$ for all $x \in U$. In other words, we have shown that if $\mathcal{F}$ is a sheaf, then the sheaf of sections over $\overline{\mathcal{F}}$ is isomorphic to $\mathcal{F}$ as a sheaf. The space $\overline{\mathcal{F}}$ is known as the *espace étale* of the sheaf, and shows we can view any sheaf as a sheaf of sections of some topological space. If $\mathcal{F}$ is not a sheaf, we can still construct the germs at points in the topological space. The resulting espace étale sheaf is known as the '*sheafification*' of $\mathcal{F}$, which has a useful universal property, which characterizes $\overline{\mathcal{F}}$ up to isomorphism.

Theorem 5.3. *If $\mathcal{F}$ is a presheaf, and $f : \mathcal{F} \to \mathcal{G}$ is a morphism into some sheaf $\mathcal{G}$, then there is a unique morphism $g : \overline{\mathcal{F}} \to \mathcal{G}$ such that if $i : \mathcal{F} \to \overline{\mathcal{F}}$ is the embedding of $\mathcal{F}$ in $\overline{\mathcal{F}}$, then $f = g \circ i$.*

Proof. If $s : U \to \overline{\mathcal{F}}$ is a continuous section in the espace étale, with $s(x) = t_x$, then there is a neighbourhood V of x with $s(y) = t_y$ for $y \in V$. We must clearly define $g(s)|_V = g(s|_V) = f(t|_V) = f(t)|_V$. Because $\mathcal{G}$ satisfies the scheme axiom, this allows us to unique define $g(s)$ such that this equality holds. □

When we discuss the tensor product, cokernel, and image in the contexts of schemes satisfying the scheme axiom, we will assume the presheafs obtained from the construction are sheafified afterwards, so that the objects we obtain are always sheafs. Note that we *need* to sheafify the image to get a sheaf, which means that for a morphism $f : \mathcal{F} \to \mathcal{G}$, we may have $\text{Im}(f) = \mathcal{G}$ whereas $f : \mathcal{F}(U) \to \mathcal{G}(U)$ may not be surjective for all U. We shall use the language of abelian categories to prevent confusion; we say a morphism f is a *monomomorphism* if $\ker(f) = 0$, and an epimorphism if $\text{Im}(f) = \mathcal{G}$. A monomorphic map is precisely an injective map, whereas an epimorphic map need not be surjective. However, we can conclude f is an epimorphism if and only if f_p is surjective for each p, and more generally, a sequence of morphisms is exact if and only if it is exact on each stalk.

Example. *Let $\mathcal{F}$ be the sheaf of multiplicative abelian groups given by the continuous, nowhere vanishing complex-valued functions on the topological space $\mathbf{C}^\times$, and consider the sheaf endomorphism φ defined by $\varphi(f) = f^2$. Then φ is clearly a monomorphism. Furthermore, φ is an isomorphism restricted to the sections of any simply connected open set φ, because on this set we can find a branch of the squareroot function, a continuous $s : U \to \mathbf{C}^\times$ with $s(z)^2 = z$. This means that for any non-vanishing function $f : U \to \mathbf{C}^\times$, $\varphi(s \circ f) = (s \circ f)^2 = f$. In particular, this means that each f_p is an isomorphism, so f itself is an isomorphism of sheaves. On the other hand, f is not an isomorphism on the global sections, from $\mathcal{F}(\mathbf{C}^\times)$ to itself, since there exists no global section s with s^2 the identity map. The reason why the map is still an isomorphism is that the presheaf image of this homomorphism is not a sheaf, since we cannot locally patch together maps which are locally the squares of functions into global squares of functions, but when we extend it to a sheaf, we obtain all functions because all functions are locally the squares of functions.*

Example. *Consider the sheaf $\mathcal{F}$ of analytic functions on the Riemann sphere $\mathbf{CP}^1$, and consider the subsheaf $\mathcal{G}$ of analytic functions which can be extended to analytic functions vanishing at the origin, and $\mathcal{H}$ the subsheaf of functions extendable to vanish at ∞. Then consider the addition map f from the direct sum sheaf $\mathcal{G} \oplus \mathcal{H}$ to $\mathcal{F}$. The maps f_p are surjective because either $\mathcal{F}_p$ or $\mathcal{G}_p$, or both, is equal to $\mathcal{H}_p$. On the other hand, the addition map isn't surjective when restricted to global sections, because $\mathcal{G}(\mathbf{CP}^1) = \mathcal{H}(\mathbf{CP}^1) = 0$, yet $\mathcal{F}(\mathbf{CP}^1) = \mathbf{C}$. This follows for similar reasons to the example above, because, while analytic functions are locally the sum of functions vanishing at the origin and infinity, one cannot globally patch together the sum of functions vanishing at the origin and infinity to get every function.*

Theorem 5.4. *A morphism $\varphi : \mathcal{F} \to \mathcal{G}$ of schemes is surjective if and only if, for every $g \in \mathcal{G}(U)$, U has a cover $\{U_\alpha\}$, and there are $f \in \mathcal{F}(U_\alpha)$ with $\varphi(f) = g|_{U_\alpha}$.*

Proof. The gist of this theorem is that we can identify the sheaf image $\text{Im}(\varphi)$ with the subset of all functions $f \in \mathcal{G}(U)$ for which there is $\{U_\alpha\}$ and $g_\alpha \in \mathcal{F}(U_\alpha)$ with $f|_{U_\alpha} = \varphi(g_\alpha)$ for each α. But if $\mathcal{H}$ is any other sheaf, and we have a morphism

$$\nu : \text{Im}'(\varphi) \to \mathcal{H}$$

Then if $f|_{U_\alpha} = \varphi(g_\alpha)$, we can define $\nu(f|_{U_\alpha})$ to be the function obtained from extending the $\nu(\varphi(g_\alpha))$ to be defined on U, and this is forced by definition, hence the extension is unique. □

Theorem 5.5. *For any morphism of sheaves $\varphi : \mathcal{F} \to \mathcal{G}$, show that for each point $p \in X$, $(\ker\varphi)_p = \ker\varphi_p$ and $(Im\varphi)_p = Im\varphi_p$.*

Proof. If f_p is a germ of a function f with $\varphi(f) = 0$, then

$$\varphi_p(f_p) = \varphi(f)_p = 0_p = 0$$

Conversely, if $\varphi_p(g_p) = \varphi(g)_p = 0$, then there is some neighbourhood U of p with $0 = \varphi(g)|_U = \varphi(g|_U)$, so $g_p = (g|_U)_p$, and $(g|_U)_p$ is in $\ker\varphi$. Conversely, every element of $(\text{Im}\varphi)_p$ is $\varphi(f)_p$ for some f defined on an open set U, which is also equal to $\varphi_p(f_p)$, showing the two families are equal. □

Recall that we define the quotient of two sheaves to be the sheafification of the quotient. The cokernel $\text{Coker}(\varphi)$ of a morphism $\varphi : \mathcal{F} \to \mathcal{G}$ can be defined as the sheafification of the sheaf associating $\text{Coker}(\varphi|_U)$ with each neighbourhood U.

Theorem 5.6. *If $\varphi : \mathcal{F} \to \mathcal{G}$ is a morphism of sheaves, then $\mathcal{F}/\ker\mathcal{F}$ is isomorphic to Im(φ), and Coker(φ) is isomorphic to $\mathcal{G}$/Im(φ).*

Proof. Consider $f \in \mathcal{F}/\ker\mathcal{F}$. Then f is locally made up of functions $f_\alpha \in \mathcal{F}(U_\alpha)/\ker\mathcal{F}(U_\alpha)$, and $\varphi(f)$ is obtained from composing the $\varphi(f_\alpha)$. In particular, if $\varphi(f) = 0$, then $\varphi(f_\alpha) = 0$ for all α, implying $f_\alpha = 0$, so $f = 0$. Thus the induced morphism is injective. The map is obviously surjective onto $\text{Im}(\varphi)$, since $\mathcal{F}$ factors through the map. Thus the map is an isomorphism. Now we have a surjective map from $\mathcal{G}$ to $\text{Coker}(\varphi)$ obtained by taking the quotient on each fibre, and the kernel of this map is precisely $\text{Im}(\varphi)$, which proves the property of the cokernel. □

The espace étale viewpoint of sheaves makes it much easier to show how we can construct a unique sheaf from the specification of the sections on a basis and the restrictions of the sheaf between open sets in the basis. Indeed, given a '$\mathcal{B}$ sheaf' $\mathcal{F}$, which is only defined on elements of $\mathcal{B}$, we can define the espace étale by forming the stalks of the sheaf, and these are put together to form the unique sheaf on all open sets of the space. Similarily, a morphism of '$\mathcal{B}$' sheafs extends unique to the sheafs extending these sheafs. This will be important when we construct the sheaf of regular functions on a scheme. An application of this shows that if we have a covering over a topological space X by a cover $\{U_\alpha\}$, and a sheaf $\mathcal{F}_\alpha$ on each U_α with the restricted topology which commutes with the other sheafs $\mathcal{F}_\beta$ in the sense that $\mathcal{F}_\alpha(V) = \mathcal{F}_\beta(V)$ for all $V \in U_\alpha \cap U_\beta$, and the restriction maps are the same on each space, then these sheafs extend uniquely to a sheaf on the entire space.

A sheaf is *flasque* if every restriction map is surjective, so every function on an open set can be extended to the entire space. This property is possessed by very fluid sheafs, like the sheaf of continuous complex-valued functions on a topological space, but not be more rigid space like the ring of holomorphic functions on the complex plane, or the ring of regular functions on an algebraic variety. If X is an irreducible space, so that any two open sets have a nonempty intersection, then the sheaf of locally constant functions on X into any set is flasque, because if U is an open set, then U is irreducible, so every locally constant function on U must be constant, and is therefore easily extendable to the whole space.

5.3 The Structure Sheaf of a Scheme

We now define the sheaf of regular functions over open subsets of a scheme. The aim is to generalize the relationship between an algebraic variety and it's sheaf of regular functions definable in terms of polynomials. Recall that a basis for the open sets of a scheme X formed from a ring A are given by the distinguished sets of the form $X_f = Z(f)^c = \{\mathfrak{a} : f \notin \mathfrak{a}\}$. These prime ideals are in one to one correspondence with the prime ideals in the localization A_f of A, obtained by adding an inverse to f. Thus, it seems that the *natural* choice for the sections of the sheaf on the set X_f is A_f. If $X_f \supset X_g$, then every prime ideal containing f also contains g, so $g \in \sqrt{(f)}$, and so $g^n = af$ for some integer n and $a \in A$. In particular, if g is invertible, then f is invertible, so we can thing of A_f as a subset of A_g, and so we choose the restriction map from A_f to A_g to be the embedding. These restriction maps are easily shown to commute, and therefore define a presheaf on the basis of distinguished sets of X_A. Provided we show this presheaf satisfies the sheaf axiom, we can extend it to a unique sheaf on all open sets of X_A. In order to prove the sheaf axiom for these sheaves, we require knowledge of how coverings work on a scheme.

Lemma 5.7. *Let X be the spectrum of A, and suppose $f_\alpha \in A$. Then the X_{f_α} covers X if and only if the f_α generate the unit ideal in A. Thus X is compact.*

Proof. The X_{f_α} cover X if and only if no prime ideal in A contains every element of f_α. But this means that $(f_\alpha) = A$, for every proper ideal is contained in a maximal ideal, which is prime. To prove X is compact, we note that it suffices to prove the condition by looking at covers by distinguished sets, since every open set is obtained as the union of distinguished sets. But if some family f_α generates the unit ideal in A, then some finite subfamily generates the unit ideal, and this corresponds precisely to compactness. □

If A is a Noetherian ring, then every closed subset of X is the finite intersection of sets of the form $Z(f)$, so every open subset of X is the finite union of distinguished sets. But this means that every cover has a finite subcover covering the same set, so we have proved that if A is Noetherian ring, then every subset of A is compact.

Theorem 5.8. *We now verify the local uniqueness and gluing aspects of the sheaf axiom. Let X_f be covered by open sets $X_{f_\alpha} \subset X_f$.*

1. *If $g, h \in A_f$ are equal in each X_{f_α}, then $g = h$.*

2. *If there are $g_\alpha \in A_{f_\alpha}$ such that g_α is equal to g_β in $A_{f_\alpha f_\beta}$, then there is $g \in A_f$ whose image in A_{f_α} is g_α for each α.*

Proof. Suppose first that $f = 1$.

1. Since $X_f = X$ is compact, we may assume that the X_{f_α} is a finite cover by elements $f_1, \dots, f_N$, and so we find $(f_1, \dots, f_N) = 1$. If g and h are equal in X_{f_α}, then there is M large enough that uniformly for each n, $f_n^M (g - h) = 0$. But this means

$$(1) = (1)^{NM} = (f_1, \dots, f_N)^{NM} \subset (f_1^M, \dots, f_n^M) \subset \text{Ann}(g - h)$$

and so $g - h = 0$.

2. We employ an algebraic formulation of a partition of unity argument. Choose M large enough that we can write $g_n = h_n / f_n^M$, with

$h_n \in A$. Since g_n agrees with g_m on $A_{f_n f_m}$, we can also make M large enough that $(f_n f_m)^M (g_n - g_m) = 0$. But

$$(f_n f_m)^M (g_n - g_m) = f_m^M h_n - f_n^M h_m = 0$$

so $f_m^M h_n = f_n^M h_m$. To form our 'partition of unity', we use the fact that

$$(f_1, \dots, f_N)^{NM} = 1$$

to find $a_n \in A$ such that $\sum a_n f_n^N = 1$ We then set $g = \sum a_n h_n$. Since

$$g - g_m = \sum a_n (h_n - g_m f_n^N)$$

and

$$f_m^N (h_n - g_m f_n^N) = f_m^N h_n - f_m^N g_m f_n^N = f_n^N h_m - h_m f_n^N = 0$$

we conclude g is equal to g_m in A_{f_m}, completing the proof.

The remainder of the proof, for a general f, reduces to our first case, since X_f is really just the spectrum of $\text{Spec}(A_f)$, and we can reduce our problem to the original problem by swapping each $f_\alpha \in A$ with $f f_\alpha \in A_f$, and the proof works in the same way. □

Because this sheaf over a basis satisfies the sheaf axiom, it extends uniquely to a sheaf on all open subsets of a scheme, and we define this to be the *structure sheaf* on the scheme. These are the functions of study in scheme theory, enabling us with an additional tool to geometrically visualize a scheme. Recall that the sheaf construction is done by proceeding down to the stalks of the sheaf on the distinguished sets, and considering the sheaf on all subsets as a family of sections.

Theorem 5.9. *For any $\mathfrak{a} \in Spec(A)$, $\mathcal{O}_\mathfrak{a}$ is isomorphic to $A_\mathfrak{a}$.*

Proof. If $f \in A_g$, and $\mathfrak{a} \in X_g$, then g is not contained in $\mathfrak{a}$, and so it makes sense to consider the homomorphism from A_g to $A_\mathfrak{a}$. This homomorphism commutes with restrictions because of the universal property of localization, and therefore the map descends to a map from $\mathcal{O}_\mathfrak{a}$ to $A_\mathfrak{a}$. It is easy to see this map is surjective, because if $f = g/h$, with $h \in \mathfrak{a}$, then $\mathfrak{a} \in X_h$, and $f \in A_h$. Now suppose that f and g are equal to one another in $A_\mathfrak{a}$, then there is $h \notin \mathfrak{a}$ for which $h(g - f) = 0$, hence $g = f$ in A_h, and $\mathfrak{a} \in X_h$, so g and f are locally identified around $\mathfrak{a}$. This completes the proof. □

Thus each stalk $\mathcal{O}_{\mathfrak{a}}$ over a sheaf is a local ring, whose unique maximal ideal is precisely the ideal generated by $\mathfrak{a}$ in $A_{\mathfrak{a}}$. Recall that the espace étale is obtained by considering a topology on the disjoint union of stalks. Thus we find $\mathcal{O}_X(U)$ can be viewed as the set of continuous functions

$$s : U \to \coprod_{\mathfrak{a} \in U} \mathcal{O}_{X,\mathfrak{a}} = \coprod_{\mathfrak{a} \in U} A_{\mathfrak{a}}$$

with $s(\mathfrak{a}) \in A_{\mathfrak{a}}$. The topology forces s to locally be given by quotients, i.e. for each $\mathfrak{a}$, there is an open X_f such that $s(\mathfrak{b}) = a/f$ for all $\mathfrak{b} \in X_f$.

Example. *Let $A = \mathbf{C}[x]/(x) \cong \mathbf{C}$. Then A has only a single prime ideal (0), and so the spectrum of A is $X = \{(0)\}$. Now if $\mathcal{O}$ is the structure sheaf, then since $X = X_1$, we find*

$$\mathcal{O}(X) = \mathcal{O}(X_1) = A_1 \cong \mathbf{C}$$

Thus X_A is really just a model for understanding functions about a single point. On the other hand, consider the sheaf corresponding to the ring $B = \mathbf{C}[x]/(x^2)$. The only prime ideal in this ring is (x), because these prime ideals are in one to one correspondence with prime ideals of $\mathbf{C}[x]$ containing x^2, and the only irreducible polynomials dividing x^2 are x. Note that (0) is not a prime ideal, since A has nilpotent elements. Thus the spectrum of B is $X = \{(x)\}$. If $\mathcal{O}$ is the structure sheaf, we now instead find that

$$\mathcal{O}(X) = \mathcal{O}(X_1) = B$$

and these 'functions', which can be uniquely expressed as $aX + b$, now not only possess a value b at the origin, but also include an 'infinitisimal' change a. Thus X_B is a space modelling the first order behaviour around a point, which is why we refer to B as corresponding to a 'fattened point' of order one at the origin.

5.4 Categorical Constructions

We list here some constructions one can do in the category of sheafs which will turn out to be useful later. We have already discussed two such constructions, the kernel and image sheafs of a morphism.

Example. *If $f : X \to Y$ is a continuous map between topological spaces, then for a presheaf $\mathcal{F}$ on X, we can define a* direct image *presheaf $f_*\mathcal{F}$ on Y by setting $(f_*\mathcal{F})(U) = \mathcal{F}(f^{-1}(U))$. If $\mathcal{F}$ is a sheaf, then $f_*\mathcal{F}$ is also a sheaf. This is a functorial assignment, in the sense that if we have a morphism $\varphi : \mathcal{F} \to \mathcal{G}$, then we get an induced morphism $f_*\varphi : f_*\mathcal{F} \to f_*\mathcal{G}$.*

Suppose $\mathcal{F}$ and $\mathcal{G}$ are two sheaves on a set X. We can then define the sheaf of morphisms between the two sheaves, denoted $\text{Hom}(\mathcal{F},\mathcal{G})$, by setting $\text{Hom}(\mathcal{F},\mathcal{G})(U)$ to be the set of morphisms between $\mathcal{F}|_U$ and $\mathcal{G}|_U$. If we have a family of morphisms φ_α defined on U_α, and we had a morphism φ, then for any $f \in \mathcal{F}(V)$, we would have

$$\varphi(f)|_{U_\alpha} = \varphi\left(f|_{U_\alpha}\right) = \varphi_\alpha\left(f|_{U_\alpha}\right)$$

and the sheaf axiom shows this uniquely defines $\varphi(f)$, and the resulting φ really is a morphism.

Example. *Unfortunately, taking the Hom sheaf does not commute with taking stalks. We may not be able to identify $Hom(\mathcal{F},\mathcal{G})_p$ with $Hom(\mathcal{F}_p,\mathcal{G}_p)$. We have a natural homomorphism $Hom(\mathcal{F},\mathcal{G})_p \to Hom(\mathcal{F}_p,\mathcal{G}_p)$, but this is in general neither injective nor surjective. Let $\mathcal{G}$ be the constant sheaf, and let $\mathcal{F}$ be the skyscraper sheaf at some closed point p in a topological space containing more points than just p, with respect to some common nonzero topological group G. Then there are no nontrivial homomorphisms φ from $\mathcal{F}$ to $\mathcal{G}$, for if $f \in \mathcal{F}$, then $f|_V = 0$ for each open set V with $p \notin V$, implying $\varphi(f)|_V = 0$. But since restrictions are the identity map in constant sheafs, this implies $\varphi(f) = 0$. This implies $Hom(\mathcal{F},\mathcal{G})_p = (0)$, whereas $\mathcal{F}_p = \mathcal{G}_p = G$, so $Hom(\mathcal{F}_p,\mathcal{G}_p) = Hom(G,G)$. Even if we sheafify, and let $\mathcal{G}$ be the sheaf of locally constant functions, then this argument still works.*

5.5 General Schemes

The schemes we have just been defining are what we now refer to as *affine schemes*. The advantage of the more generalized definition of a scheme we now introduce is that it is a local definition as a space which locally looks like an affine scheme, just like a differentiable manifold is a space that locally looks like a Euclidean space. The advantage of this is we are able to 'glue' two schemes together in a completely natural way. This is best introduced by operating in the category of ringed spaces. A *ringed space* is just a topological space X with a fixed sheaf $\mathcal{O}_X$ of rings. A morphism between two ringed spaces X and Y is a continuous map $f : X \to Y$ together with a morphism $\varphi : \mathcal{O}_Y \to f_* \mathcal{O}_X$, i.e. a family of morphisms $\varphi_U : \mathcal{O}_Y(U) \to \mathcal{O}_X(f^{-1}(U))$ for each open U in Y commuting with the restriction maps. We will define a general scheme to be a ringed space which is locally isomorphic to an affine scheme. In order to do this, we must first identify the main properties of the structure sheaf of an affine scheme.

So let's suppose that X is a ringed space, which is the affine scheme corresponding to some ring A. Then we can recover A from this ringed space, since $A = \mathcal{O}(X)$. For any $f \in A$, we can let $U_f \subset X$ denote the points $x \in X$ such that $f_x \in \mathcal{O}_x$ is a unit. If $x \in U_f$, then f_x is a unit, so there is a neighbourhood V containing x and $g \in \mathcal{O}(V)$ such that $g \cdot f|_V = 1$, and so all elements of V are in U_f, hence U_f is an open neighbourhood of x. We claim that $X_f = U_f$. Since f is invertible on X_f, $X_f \subset U_f$. Conversely, if $f \in \mathfrak{a}$, then we cannot possible invert f locally around $\mathfrak{a}$, because if there was $h \in A - \mathfrak{a}$ with $fg = 1$ in A_h, then $h^n(1 - fg) = 0 \in \mathfrak{a}$, hence $1 - fg \in \mathfrak{a}$ because $\mathfrak{a}$ is prime, which would imply $1 \in \mathfrak{a}$ since $fg \in \mathfrak{a}$, contradicting the fact that $\mathfrak{a}$ is prime. Thus $X_f = U_f$. Thus in any ringed space X, a necessary condition for X to be isomorphic to an affine scheme is that $\mathcal{O}(U_f)$ is isomorphic to A_f, such that the restriction map from A to $\mathcal{O}(U_f)$ is just the canonical map of A into A_f. Unfortunately, this isn't sufficient to show X is isomorphic to $\text{Spec}(A)$, as the next example shows.

Example. *Consider the ringed space X consisting of $\mathbf{C}$, equipped with the Zariski topology, and with rational functions as the functions, which is the affine scheme of the ring $\mathbf{C}[x]$. Let Y be the topological space obtained by gluing 0 and 1 together, and use the gluing map $\varphi : \mathbf{C} \to Y$ to obtain a sheaf $\mathcal{O}_Y = \varphi_*(X)$ on Y. For $[p] \in Y$ with $p \notin \{0,1\}$, $(\mathcal{O}_Y)_p$ is precisely the ring of rational functions defined at p, whereas $(\mathcal{O}_Y)_{\{0,1\}}$ is the ring of rational functions defined at 0 and 1.*

According to math stack overflow, if we let Y be a 'nodal curve', with the corresponding scheme, then this statement is true, but otherwise it's not a counterexample. TODO: When I know more, elaborate on this counterexample.

Consider the nodal curve V, which is the locus of solutions to the equation $Y = X(X-1)$. By embedding $\mathbf{C}$ in V by the map $i(X) = (X, X(X-1))$, and then projecting V into $\mathbf{C}$ by the map $\pi(X,Y) = Y$, we obtain a map $f(X) = X(X-1)$ from $\mathbf{C}$ to itself. If $f(X) = f(Y)$, then $X(X-1) = Y(Y-1)$, which can be *rewritten as $X^2 - Y^2 = X - Y$, hence $(X - Y)(X + Y) = X - Y$, which implies that if $X \neq Y$, then $X + Y = 1$.*

Consider the nodal curve V, which is the locus of solutions to the equation $y^2 = x^2 + x^3$. We can parameterize this curve by $x = t^2 - 1$, $y = t(t^2 - 1)$.

To obtain this map from a ringed space to an affine scheme, we now add the additional condition that the stalks $\mathcal{O}_x$ are local rings. If X is a ringed space such that the stalks $\mathcal{O}_x$ are all local rings, we say X is a *locally ringed space*. We let $\mathfrak{m}_x$ denote the maximal ideal of $\mathcal{O}_x$. Note that all affine schemes are locally ringed, since we have already seen $\mathcal{O}_{\mathfrak{a}}$ is isomorphic to $A_{\mathfrak{a}}$. A morphism $\varphi : X \to Y$ of locally ringed spaces is a morphism of ringed spaces such that the induced local maps $\varphi : (\mathcal{O}_Y)_{\varphi(x)} \to (\mathcal{O}_X)_x$ are *local homomorphisms*, in the sense that $\varphi^{-1}(\mathfrak{m}_{f(x)}) = \mathfrak{m}_x$. If X is a locally ringed space with $\mathcal{O}_X(U_f) = A_f$, then we can define a map $\varphi : X \to \text{Spec}(A)$ by setting $\varphi(x)$ to be the prime ideal consisting of all $f \in A$ such that $f_x \in \mathfrak{m}_x$. If $x \in U_f$, then f_x is invertible, hence it can't possibly be in $\mathfrak{m}_x$, so $f \notin \varphi(x)$, which means exactly that $\varphi(x) \in \text{Spec}(A)_f$. Conversely, if $f \notin \varphi(x)$, then $f_x \notin \mathfrak{m}_x$, so f_x is invertible in $\mathcal{O}_x$.

If φ is a homeomorphism between X and $\text{Spec}(A)$, then we can finally claim we have generated an isomorphism between ringed spaces, by setting φ_{U_f} to be the identity map

Theorem 5.10. *If $\varphi : A \to B$ is a homomorphism of rings, then φ induces a natural morphism $f : Spec(B) \to Spec(A)$ of locally ringed spaces, and every morphism of locally ringed spaces is induced by a homomorphism of rings.*

Proof. Given φ, define f by setting $f(\mathfrak{b}) = \varphi^{-1}(\mathfrak{b})$. Then f is continuous, since $f^{-1}(Z(\mathfrak{a})) = Z(\varphi(\mathfrak{a})$. For each $\mathfrak{a}$, we can localize φ to obtain a local homomorphism $\varphi_{\mathfrak{a}} : A_{\varphi^{-1}(\mathfrak{a})} \to B_{\mathfrak{a}}$, which we can put together to get a sheaf map on open subsets of the space, and since the stalk maps were local homomorphisms this really is a homomorphism of locally ringed spaces. Conversely, ... □

Chapter 6

Hartshorne Exercises

Theorem 6.1. *If H is a hypersurface in $\mathbf{P}^n$, then $\mathbf{P}^n - H$ is isomorphic to an affine variety.*

Proof. Let H be the nullset of some homogenous polynomial f of degree d, so that

$$f = \sum a_i M_i$$

where the M_i are homogenous polynomials of degree d. The d-uple embedding $T : \mathbf{P}^n \to \mathbf{P}^N$ is a regular embedding, because it is given by polynomial equations, and the inverse can also be given locally by polynomial equations. Thus $T(\mathbf{P}^n)$ and $T(H)$ are subvarieties of $\mathbf{P}^N$. In particular, $T(H)$ is precisely a hyperplane defined by the linear equation $\sum a_i M_i = 0$, and thus $\mathbf{P}^N - T(H)$ is isomorphic to $\mathbf{A}^N$. Composing this with the original map identifies $\mathbf{P}^n - T(H)$ as a closed subset of $\mathbf{A}^N$. □

A subset of a topological space is *locally closed* if it an open subset of it's closure, or equivalently, if it is an intersection of an open set with a closed set. If X is a quasi-affine or quasi-projective variety, and Y is an irreducible locally closed subset, then Y is also a quasi-affine (respectively, quasi-projective) variety, by virtue of it being a locally closed subset of the same affine or projective space. If $Y = A \cap B$, where A is open in X and B is closd in X, then A is open in the ambient space, and $B = X \cap B'$ where B' is closed in the ambient space, and then $Y = A \cap B'$. We call this the *induced structure* on Y, and we call Y a *subvariety* of X.

Theorem 6.2. *Now let $\varphi : X \to Y$ be a morphism. Let $X' \subset X$ and $Y' \subset Y$ be irreducible locally closed subsets such that $\varphi(X') \subset Y'$. Show that $\varphi|_X : X' \to Y'$ is a morphism.*

Proof. Let $x \in X'$. Then there is a neighbourhood U of x in X such that φ is given by a rational map on U, with poles lying outside of U. But then $U \cap X'$ is a neighbourhood of x in X' and φ is obivuously given by a rational map here as well. □

If X and Y are affine varieties in $\mathbf{A}^n$ and $\mathbf{A}^m$, then $X \times Y$ in $\mathbf{A}^{n+m}$ is irreducible. If $X = Z(\mathfrak{a})$, where $\mathfrak{a}$ is prime, and $Y = Z(\mathfrak{b})$, where $\mathfrak{b}$ is prime, then $X \times Y = Z(\mathfrak{a} \times \mathfrak{b})$, where

$$\mathfrak{a} \times \mathfrak{b} = \{(f,g) : f \in \mathfrak{a}, g \in \mathfrak{b}\}$$

s

$$k[X \times Y] = k[X] \otimes_k k[Y]$$

Consider the embedding $(f,g) \mapsto fg$ into $k[X \times Y]$. If T is a bilinear map from $k[X] \times k[Y]$ to some vector space V, then

Consider the basic case $X = \mathbf{A}^n$ and $Y = \mathbf{A}^m$. Then it suffices to verify that any bilinear map T from $k[X] \times k[Y]$ into some vector space V factors uniquely through $k[X, Y]$. It is clear that we can write any polynomial f uniquely as

$$f(X,Y) = \sum C_{\alpha\beta} X^\alpha Y^\beta$$

It is clear that we must map $X^\alpha Y^\beta$ to $T(X^\alpha, Y^\beta)$, and this uniquely defines the map since the $X^\alpha Y^\beta$ are a basis on the space. But then

$$T(f(X), g(Y)) = \sum a_\alpha b_\beta T(X^\alpha, Y^\beta) = \sum a_\alpha b_\beta T^*(X^\alpha Y^\beta) = T^* \left(\sum a_\alpha b_\beta X^\alpha Y^\beta \right)$$

This is well defined by bilinearity. But now $k[V] = k[X]/I(V)$ and $k[W] = k[Y]/I(W)$, and so if we have a bilinear map $T : k[V] \times k[W]$, we obtain a unique map on $k[X, Y]$ vanishing on $I(V) \times I(W)$, hence decending to a unique map on $k[V \times W]$, which shows the tensor product identity holds.

Theorem 6.3. *A morphism of sheaves is an isomorphism if and only if it is injective and surjective.*

Proof. Let $f : \mathcal{F} \to \mathcal{G}$ be a morphism with inverse $f^{-1} : \mathcal{G} \to \mathcal{F}$. Then we know each $f^U : \mathcal{F}(U) \to \mathcal{G}(U)$ is an isomorphism, so $\ker(f^U) = 0$ and $\text{Im}(f^U) = \mathcal{G}(U)$, and taking unions, we conclude $\ker(f) = 0$ and $\text{Im}(f) = \mathcal{G}$, hence f is injective and surjective. Conversely, if $\ker(f) = 0$, then, if we let $\mathcal{H}$ denote the *presheaf* image of $\mathcal{F}$ in $\mathcal{G}$, then by taking inverses over each fibre, we conclude that f has an inverse $f^{-1} : \mathcal{H} \to \mathcal{F}$. But then can quite boringly verify that $\mathcal{H}$ satisfies the sheaf property by pulling back onto $\mathcal{G}$, which implies that, since $\text{Im}(f)$ is the sheafification of $\mathcal{H}$, which we know to be a sheaf, we conclude $\mathcal{H} = \text{Im}(f) = \mathcal{G}$, finishing the proof. □

Theorem 6.4. *Let $(\mathcal{F}_\alpha, f_{\alpha\beta})$ be a direct system of sheaves on a Noetherian topological space X. We define the direct limit $\lim \mathcal{F}_\alpha$ pointwise over open sets, i.e. $(\lim \mathcal{F}_\alpha)(U) = \lim \mathcal{F}(U)$. This is a direct limit in the category of sheaves.*

Proof. We first verify that the direct limit, define pointwise over open sets, satisfies the sheaf axiom. The uniqueness of the sheaf axiom is clear, since if f and g are in the direct limit, they are in some common $\mathcal{F}_\alpha$, which satisfies the scheme axiom, and so they are equal if they agree on a partition of open subsets. The converse also isn't that hard. If f and g are elements of the limit agreeing on a common domain, they lie in some common sheaf in the sequence that defines the limit, and we can use the sheaf axiom there to uniquely extend f and g to an element on the union of their domain. This implies that we can uniquely extend a finite sequence $f_1, \dots, f_n$ of compatible maps. But this completes the proof in general, for a space is Noetherian if and only if every subspace is compact, so we know that if we have a potentially infinite compatible family $f_\alpha \in \lim \mathcal{F}(U_\alpha)$ with $\bigcup U_\alpha = U$, then U is compact, hence the U_α has a finite subcover, and if we extend the finite family of maps corresponding to the finite subcover, this map will be the extension we needed.

□

Theorem 6.5. *Let $\mathcal{F}, \mathcal{G}$ be sheaves of abelian groups on X. Then for an open set U, the set of morphisms from $\mathcal{F}|_U$ to $\mathcal{G}|_U$ forms a group, and the presheaf $U \mapsto Hom(\mathcal{F}|_U, \mathcal{G}|_U)$ is actually a sheaf, known as the sheaf of local morphisms.*

Proof. Let U be an open set covered by V_α, together with sections s_α which commute on the restriction. The sheaf axiom means that we can define s by setting $s(t)$ to be the unique sheaf such that $s(t)|_{V_\alpha} = s_\alpha(t|_{V_\alpha})$, and then this map obviously commutes with all the restrictions. Any such s must have these values on the restriction, so it a unique homomorphism on U extending the s_α. □

Theorem 6.6. *Let $\mathcal{G}$ be a subsheaf of a sheaf $\mathcal{F}$. Show that the natural map of $\mathcal{F}$ to $\mathcal{F}/\mathcal{G}$ is surjective, and has kernel $\mathcal{G}$. Thus there is an exact sequence*

$$0 \to \mathcal{G} \to \mathcal{F} \to \mathcal{F}/\mathcal{G} \to 0$$

Proof. This follows because, at each point p, we have an exact sequence

$$0 \to \mathcal{G}_p \to \mathcal{F}_p \to \mathcal{F}_p/\mathcal{G}_p \to 0$$

which holds precisely because of the first isomorphism theorem for abelian groups / rings / etc. □

Theorem 6.7. *Let $(\mathcal{F}_\bullet, F_{\bullet,\bullet})$ be an inverse system of sheaves on X. Show that the pre-sheaf $U \mapsto \lim \mathcal{F}_t(U)$ is a sheaf. It is called the inverse limit of the system $\mathcal{F}_t$, and is denoted by $\lim \mathcal{F}_t$. Show that it has the universal property of an inverse limit in the category of sheaves.*

Proof. The construction of the inverse limit is obvious since inverse limits exist in the category of abelian groups, and since a morphism of sheafs is essentially just defined 'pointwise' on each open set, the morphism property is obvious. All that remains is to show that the inverse limit of sheaves actually satisfies the sheaf axiom. Suppose that we have open sets U_α covering U and $f_\alpha \in \mathcal{F}_\infty(U_\alpha)$ such that $f_\alpha|_{U_\alpha \cap U_\beta} = f_\beta|_{U_\alpha \cap U_\beta}$. Then, using the fact that the projections $\pi_t : \mathcal{F}_\infty \to \mathcal{F}_t$ are sheaf morphisms, we conclude

$$\pi_t(f_\alpha)|_{U_\alpha \cap U_\beta} = \pi_t(f_\alpha|_{U_\alpha \cap U_\beta}) = \pi_t(f_\beta|_{U_\alpha \cap U_\beta}) = \pi_t(f_\beta)|_{U_\alpha \cap U_\beta}$$

Hence, for each t, using the sheaf axiom on $\mathcal{F}_t$, we can construct a map $f_t \in \mathcal{F}_t(U)$ with $f_t|_{U_\alpha} = \pi_t(f_\alpha)$. All that remains is to show that $F_{tu}(f_u) = f_t$, from which it follows from the inverse limit property of the presheaf that there is $f \in \mathcal{F}_\infty(U)$ with $\pi_t(f) = f_t$, hence $f|_{U_\alpha} = f_\alpha$ since for all t,

$$\pi_t(f|_{U_\alpha}) = \pi_t(f)|_{U_\alpha} = f_t|_{U_\alpha} = \pi_t(f_\alpha)$$

and an element of the group at a particular open set on the sheaf is determined by it's images under the projection maps π_t, hence the two elements must be equal. To show $F_{tu}(f_u) = f_t$, we calculate that for each α,

$$F_{tu}(f_u)|_{U_\alpha} = F_{tu}(\pi_u(f_\alpha)) = \pi_t(f_\alpha) = f_t|_{U_\alpha}$$

and using the sheaf axiom on $\mathcal{F}_t$, we conclude $f_u = f_t$ since they agree on a covering of open sets. □

Theorem 6.8. *Let $f : X \to Y$ be a continuous map of topological spaces. Show that for any sheaf $\mathcal{F}$ on X there is a natural map $f^{-1}f_*\mathcal{F} \to \mathcal{F}$ and for any sheaf $\mathcal{G}$ on Y there is a natural map $\mathcal{G} \to f_*f^{-1}\mathcal{G}$. Use these maps to show that there is a natural bijection of sets, for any sheaves $\mathcal{F}$ on X and $\mathcal{G}$ on Y,*

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

Hence we say that f^{-1} is a left adjoint of f_*, and that f_* is a right adjoint of f^{-1}.

Proof. Recall $f_*\mathcal{F}$ is the sheaf on Y defined by $(f_*\mathcal{F})(U) = \mathcal{F}(f^{-1}(U))$. Similarily, if $\mathcal{G}$ is any sheaf on Y, we can define a sheaf $f^{-1}\mathcal{G}$ on X by setting $(f^{-1}\mathcal{G})(U) = \lim_{V \downarrow f(U)} \mathcal{G}(V)$. If $F : \mathcal{F} \to \mathcal{G}$ is a homomorphism, we can define a homomorphism $F_* : f_*\mathcal{F} \to f_*\mathcal{G}$ by setting $F_*(f) = F(f)$, and if $F : \mathcal{G} \to \mathcal{F}$ is a homomorphism, we can define a homomorphism $f^{-1}F : f^{-1}\mathcal{G} \to f^{-1}\mathcal{F}$ such that $\pi_Z((f^{-1}F)(g)) = F(\pi_V(g))$. Now

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

If $g^V \in \mathcal{F}(f^{-1}V)$, then $g^V|_U$ makes sense, and commutes with the restriction maps, so, taking limits, we obtain a morphism from $\lim_{V \downarrow f(U)} \mathcal{F}(f^{-1}V)$ to $\mathcal{F}(U)$, and collecting these together gives us a morphism from $f^{-1}f_*\mathcal{F}$ to $\mathcal{F}$ that was required. This map is natural, because any morphism commutes with the restrictions. Similarily, we have

$$(f_*f^{-1}\mathcal{G})(U) = \lim_{V \downarrow f(f^{-1}(U))} \mathcal{G}(V)$$

If g is defined on U, then since $f(f^{-1}(U)) \subset U$, we can certainly restrict g to suitably small open sets V defined in the limit of f_*f^{-1}, giving us the natural map from $\mathcal{G}$ to $f_*f^{-1}\mathcal{G}$, which is natural, again, because it is defined using restrictions, which commutes with sheaf morphisms. But now we have a natural map from

$$H_X(f^{-1}\mathcal{G}, \mathcal{F}) \to H_Y(f_*f^{-1}\mathcal{G}, f_*\mathcal{F}) \to H_Y(\mathcal{G}, f_*\mathcal{F})$$

Just draw some commutative diagrams and be happy this is a natural bijection? □

Theorem 6.9. *Let X be a topological space, P a point, and A an abelian group. Define a sheaf $\mathcal{F}$ by setting $\mathcal{F}(U) = A$ if $P \in U$, and 0 otherwise. Verify that the stalk $\mathcal{F}_Q = A$ for all Q in the closure of the point P. Show this sheaf could also be described as $i_*(A)$ where A denotes the constant sheaf A on the closure $\overline{P}$ of P, and $i : \overline{P} \to X$ is the inclusion map.*

Proof. s □

A sheaf is flasque if restriction is always surjective.

Theorem 6.10. *A constant sheaf on an irreducible topological space X is flasque.*

Proof. We have to prove every locally constant map f on an open $U \subset X$ is extendable to a locally constant map on the whole space. But if X is irreducible, then U is irreducible, hence connected, and so f is constant on U, hence f is a constant map that can easily be extended globally. □

Theorem 6.11. *If $0 \to F' \to F \to F'' \to 0$ is exact, and F' is flasque, then for any U, $0 \to F'(U) \to F(U) \to F''(U) \to 0$ is exact.*

Proof. $0 \to F'(U) \to F(U)$ is easily exact. If $\text{im}(f_1) = \ker(f_2)$, then $\text{im}(f_1(U)) \subset \ker(f_2(U))$. Suppose $f_2(h) = 0$. Then $f_2(h_p) = 0$, so $h_p = f_1(k_p)$ for some k. Then $h = f_1(k)$ on a small enough neighbourhood

Every germ can be extended globally? □

Theorem 6.12. *Let A be a ring and $(X, \mathcal{O}_X)$ a scheme. Given a morphism $f : X \to Spec\ A$, we have an associated map $f^\sharp : \mathcal{O}_{Spec A} \to f_* \mathcal{O}_X$ on sheaves. Taking global sections we obtain a homomorphism $A \to \mathcal{O}_X(X)$. Thus there is a natural map*

$$\alpha : Hom_{Schemes}(X, Spec\ A) \to Hom_{Rings}(A, \mathcal{O}_X(X)))$$

Show that α is bijective.

Proof. Let $\varphi : A \to \mathcal{O}(X)$ be a homorphism of rings. Suppose that X is covered by charts $u_\alpha : U_\alpha \to \text{Spec}(R_\alpha)$. We define a family of ring homomorphisms $\varphi_\alpha : A \to R_\alpha$ by setting $\varphi_\alpha(a) = u_\alpha(\varphi(a)|_{U_\alpha})$. By Proposition 2.3, there are ringed space maps $f_\alpha : U_\alpha \to \text{Spec}(A)$ defined by setting $f_\alpha(\mathfrak{p}) = \varphi^{-1}(\mathfrak{p})$, for $\mathfrak{p} \in U_\alpha$, and with $(f_\alpha^\sharp)_\mathfrak{p} : A_{\varphi^{-1}(\mathfrak{p})} \to \mathcal{O}_p$ the local maps put together to form the sheaf morphism. The maps f_α are clearly independent of α, as are the maps $(f_\alpha^\sharp)_\mathfrak{p}$, so they can be glued together to form a sheaf morphism $f : X \to \text{Spec}(A)$. We claim that $f^\sharp = \varphi$, which would show α is surjective. But we know $f_\mathfrak{p}^\sharp : A_{\varphi^{-1}(\mathfrak{p})} \to \mathcal{O}_p$ is just the localization map obtained from φ, and $\varphi_\mathfrak{p}$ is also just exactly this localization, which shows $f_\mathfrak{p}^\sharp = \varphi_\mathfrak{p}$ for each $\mathfrak{p}$, and then, by the sheaf axiom, $f^\sharp = \varphi$. But conversely, if $f : X \to \text{Spec}(A)$ is a scheme morphism, then $f^\sharp : A \to \mathcal{O}(X)$ is a ring morphism, and if we form the scheme morphism from $g : X \to \text{Spec}(A)$ from $f^\sharp$, then we find $g = f$, since BLAH. This shows α is injective. □

Theorem 6.13. *Describe Spec*$(\mathbf{R}[x])$.

Proof. We of course have the maximal ideals $(x-t)$ corresponding to $t \in \mathbf{R}$, and the topology restricted to this set is precisely the finite closed topology, i.e. just like the Zariski one on $\mathbf{C}$. The other elements of the spectrum correspond to irreducibly real quadratic polynomials, which can also be viewed as pairs $\{z, \overline{z}\}$ of complex numbers with nonzero imaginary part which are conjugates of one another, which corresponds to the irreducible polynomial $x^2 - 2\mathfrak{Re}(z)x + |z|^2$. Thus we can think of the spectrum X as $\mathbf{C}$ folded in half about the real axis. If $f \in \mathbf{R}[x]$, then $Z(f)$ is then precisely those $z \in X$ with $f(z) = 0$ (note that since $f(\overline{z}) = \overline{f(z)}$, $Z(f)$ is invariant under reflection about the real axis, and is therefore well defined on X). Thus the closed sets are precisely the finite sets, so the topology on the spectrum is just the finite closed set topology. The sheaf at some open set obtained by removing finitely many points $\{p_1, \dots, p_n\}$ is precisely the set of rational functions which don't vanish at the p_n. □

Theorem 6.14. *Let X be a scheme, let $f \in \mathcal{O}(X)$, and set X_f*