Algebraic Geometry

Jacob Denson

June 19, 2017

Table Of Contents

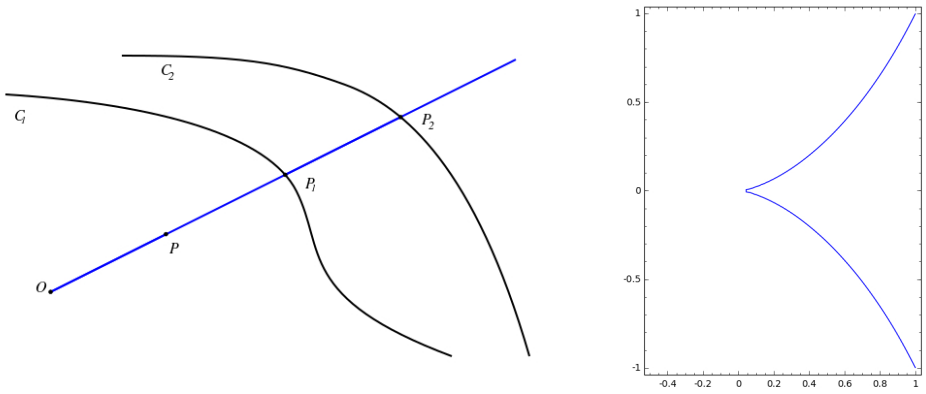
Chapter 1

Affine Algebraic Sets

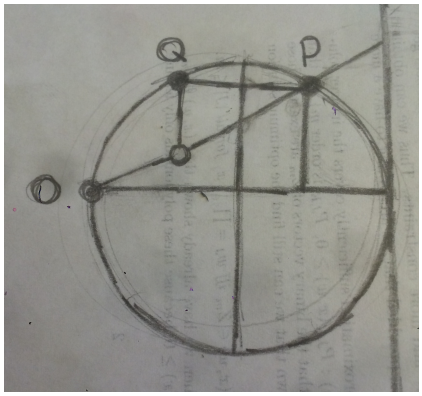
1.1 Planar Algebraic Curves

Classically, Euclidean geometry discusses the relations between the intersections of lines and circles. However, this geometry has proved inadequete for certain constructions. For instance, one cannot use lines and circles to construct a cube with twice the volume of another cube. Similarily, one cannot trisect general angles, nor . However, by introducing more complicated geometric shapes, one is able to solve these problems. The Greek mathematician Menaechmus introduced the parabola to square the circle. 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 geometric shape 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 obtained from two other given curves C and C' and a pole O. It is the locus of points Q for which there exists $P \in C$ and $P' \in C'$ such that O lies on PP', the distance between O and Q is equal to the distance between P and P', and OQ is oriented the same as PP'.*



The **cissoid of Diocles** *is obtained by taking C to be a circle, C' a line tangent to the circle, and O the point opposite the tangent line. In this case, we can also describe the cissoid as those points obtained by taking a point P on the circle, and considering the intersection of the line OP with the line parallel to the tangent at O through the point Q obtained by reflecting P about the line parallel to the tangent which bisects the circle. This claim is proved by a simple geometric argument (similarity of triangles) suggested by the diagram below.*



If we consider a coordinate system in which O lies at the origin, C is the unit circle with center $(1,0)$, and C' is the line $X = 2$, then in polar coordinates, we may write the points on C as the solutions to $r = 2\sec t$, and the points on C' by $r = 2\cos t$. The advantage of these coordinates is that the angles t parameterize the lines through O, and thus the possible pairs of intersection points P and P' on the individual curves. The distance between the line $r = 2\sec t$ and $r = 2\cos t$ is the difference $2\sec t - 2\cos t$ in radii, and therefore the polar

equation defining the cissoid of Diocles is

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

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

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

so the cissoid of Diocles is the locus of points which form the solution set to the polynomial $(X^2 + Y^2)X = 2Y^2$, and is therefore a planar curve.

Example. *One can also describe the cissoid of Diocles by a construction due to Newton. Let O be a point, and let P vary over a line not containing O. If we consider a point Q such that PQO forms a right triangle, and the distance between P and Q is the same as the distance between O and a line, then the midpoints between P and Q forms Diocles' cissoid. This means we have expressed Diocles' cissoid as a* **conchoid**, *that is, as a curve defined as a locus between a point O and a curve C, which are the set of points generated by the lines through O and a point varying over C which lie at a fixed distance from the curve. In this case, the curve is a single line, and the distance is half the distance between the point and the line, so the cissoid is only half of a conchoid. 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 $2X = a + 1$ and $2Y = b + p$. 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 created 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 family of conchoidal curves (again, actually only half a conchoid) which are deformations of the cissoid of Diocles, obtained by taking another point on the line between P and Q, rather than the midpoints. 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* $c > 1/2$*. This polynomial is quartic, but we can obtain cubic curves with much the same behaviour. 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* $a = -1$*.*

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 rearrange 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 we choose a coordinate system where P lies at $(a,0)$, Q lies at $(b,0)$, and the distance parameter is b, then the equation describing the conchoid is $(X-a)^2(X^2+Y^2) = b^2X^2$. 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 P is further away than the distance defining the conchoid, then we obtain two smooth curves. If P is closer than the distance defining the conchoid, then the conchoid appears to 'swing' around P, meeting P in two intersection points. If the distance to P is equal to the distance defining the conchoid, then we obtain a 'cusp' at P, where the curve appears to swing towards P, and then swing out giving a sharp point. An interesting feature of the conchoid is that the time for an object to reach the 'bottom' of the conchoid under the influence of gravity is independent of its original starting position. Another interesting feature of the conchoid is that it consists of two separate components lying on either side of the line $X = a$, with the line as an asymptote.*

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 circumpherence of the outer circle is a rational multiple of the circumpherence 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 circumpherence of the inner circle is equal to the circumpherence 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 circumpherence 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. Similarily, 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.

Example. *The Watt curves were discovered by James Watt in his work on the steam engine. The construction is as follows. Consider a quadrilateral $ABCD$, where A and D are fixed points in the plane, and B and C can be moved on circles around A and D such that the length BC is fixed. We obtain a Watt curve by watching the possible configurations of an arbitrary point M on the line BC. If one takes linkages of further roads, one can generate even more complicated curves. Every finite segment of an planar algebraic curve can be obtained by some linkage of rods.*

Example. *The Lissajous curves is the planar trajectory of a pendulum, or of a particle undergoing simple harmonic motion. Provided the harmonic motion is rational, so that the curves are closed, the curve obtained is algebraic.*

Algebraic curves are a rich source of geometric problems. For this reason, they inspired the general theory of algebraic geometry.

- Given an algebraic curve in the plane, is it possible to parameterize it'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

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

where f is a rational function in two variables, then we know that the parameterization $x = t^2 - 1$ gives $\sqrt{x^2+x^3} = t(t^2-1)$, so by a change of variables we have reduced the problem of integrating a complicated rational function with square roots with a rational function of a single variable of the form

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

This shows that the antiderivative of every rational function of x and $\sqrt{x^2+x^3}$ is expressible in terms of elementary functions. 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, and if $t^2 = 1$, then $t = \pm 1$ is 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 in algebraic geometry.

- It is an important problem in algebraic geometry to classify curves. 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 can separately quantify the degree one curves, which are are equivalent lines, the degree two curves, known as conic sections which, after discounting degenerate solutions, are classified into parabolas, hyperbolas, and ellipses. The cubics are a whole new world – Newton gave a classification into 72 kinds of conics, 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, and we can imagine the situtation with cubics simplifies considerably in projective space. 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**. 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.

An issue in the theory of algebraic plane curves, which does not appear in other areas of the theory of curves is the presence of **singular points**, which is a point on the curve where the curve is no longer 'smooth'. If an algebraic curve is defined with respect to a polynomial f, and p is a point on the curve where $f_Y(p) \neq 0$, then locally around p we can describe the algebraic curve as a differentiable function of X. Similarily, if $f_X(p) \neq 0$, then we can describe the curve as a differentiable function of Y. However,

it is possible on an algebraic curve to have $f_X(p)$ and $f_Y(p)$ equal to 0, in which case the algebraic curve does not behave as a 'smooth curve'. One reason this can occur is if the algebraic curve has a **node**, which occurs if p is the intersection point of two 'separate curves' of which the algebraic curve is composed. Another reason is if the function is the overlapping point of two differentiable curves which do not meet at a common tangent, 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 in the analysis of algebraic curves.

Another issue in algebraic plane curves, more avoidable in the last, is that certain 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 $\{(0,0)\}$ is an algebraic curve. This annoyance disappears when we study algebraic plane curves over the complex numbers, because 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 duality between the analytic and algebraic viewpoints.

On a related note, a short time after the methods of algebraic geometry were discovered, it was noticed that most techniques generalize naturally to studying algebraic plane curves defined over any field, not only because polynomial equations are definable over any field. In general, we shall assume some field K is fixed, and we then study 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 discrete or abstract field, 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.

Example. *If P is a point outside a circle, it is the intersection point of two tangent lines on the circle, and we define the line through these two points to be the polar line with respect to P. If we are considering the unit circle, then the tangent line at a point (x,y) is defined to be the set of points satisfying $xX + yY = 1$, so if P has coordinates (a,b), then we wish to find values of x and y such that $ax + by = 1$ and $x^2 + y^2 = 1$, and this reduces to finding values of y such that $(a^2 + b^2)y^2 - 2by + (1 - a^2) = 0$, which has two real valued solutions provided that $a^2 + b^2 > 1$. The polar line is then obviously defined by $aX + bY = 1$. If $a^2 + b^2 < 1$, so that P lies within the circle, we cannot find two real valued solutions to the equation, but we know that the equation $(a^2 + b^2)y^2 - 2by + (1 - a^2) = 0$ has two distinct complex solutions z and $\overline{z}$ which are complex conjugates of each other. This means that on the complex solution to the equation $X^2 + Y^2 = 1$, points in the (real) plane not lying on the circle are coincident to two tangent planes. If w is the unique solution to $aw + bz = 1$, then $\overline{w}$ is the unique solution to $a\overline{w} + b\overline{z} = 1$, so (w,z) and $(\overline{w},\overline{z})$ give the two points generating the tangent lines which intersect at (a,b). The unique complex plane between (w,z) and $(\overline{w},\overline{z})$ intersects the real plane in the line $aX + bY = 1$, so we can still define the polar line about the circle with respect to a point inside the circle. Geometrically, the points on this line are exactly the points whose polar line passes through P. Thus we have a duality property that would not be so visible were we to restrict ourselves to real space, and this leads to the theory of inversive geometry.*

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 these general tools. Moreover, they are historically the reason algebraic geometry was studied in the first place, and I think the best way to understand of modern terminology is from the historical development of a subject.

1.2 Affine Varieties

Given a polynomial $f \in K[X_1,\ldots,X_n]$, we can consider the zero set $V(f) = \{p \in \mathbf{A}^n : f(p) = 0\}$, which is the **hypersurface** defined by f, a generalization of planar curves to higher dimensions. However, in higher dimensional space, it is also natural consider the geometric object formed from

the common zeroes of two polynomials $V(f,g) = V(f) \cap V(g)$. More generally, given a set S of polynomials, we can consider the set $V(S)$, which consists of the points forming the set of common zeroes of all polynomials in S. These zero sets are called **affine varieties**, and they are the main object of study in algebraic geometry. The planar algebraic curves are the affine varieties in $\mathbf{A}^2$ defined by a single polynomial.

The class of affine varieties is interesting from the point of view of Euclidean geometry, because it is invariant under affine transformations, and many of the interesting shapes in Euclidean geometry can be identified with certain varieties, through the tools of analytic geometry. If T is an affine transformation on $\mathbf{A}^n$, and if we define the endomorphism T^* on $K[X_1,\ldots,X_n]$ by letting $T^*f = f \circ T$, then

$$\begin{aligned} T^{-1}(V(\mathfrak{a})) &= \{x : (\forall f \in \mathfrak{a} : f(Tx) = 0)\} \\ &= \{x : (\forall f \in \mathfrak{a} : (T^*f)(x) = 0)\} = V(T^*\mathfrak{a}) \end{aligned}$$

It is easy to prove that T^* preserves the degree of polynomials, which is the reason why the degree of polynomials is an *isomorphism-invariant* property of algebraic curves, when our isomorphisms are obtained by affine transformations.

Theorem 1.1. *A plane curve of degree n intersects a line in at most n places.*

Proof. Consider first the easy case where the line is just the X axis. In this case, the zeroes of a polynomial $f(X,Y) = \sum a_{ij}X^iY^j$ which lie on the X axis are in one to one correspondence with the set of solutions to the univariate polynomial $f(X,0) = \sum a_{i0}X^i$. If $f(X,0) = 0$, then $V(f)$ contains the X axis, and otherwise f can have at most $\deg f(X,0) \leqslant \deg f(X,Y) = n$ points, the intersection of $V(f)$ and the X axis can contain at most $\deg f(X,0) \leqslant \deg f(X,Y) = n$ separate points. Now in general, we can map the X axis to any line by some affine transformation T, and the points on the intersection of the line and $V(f)$ are in one to one correspondence with the intersections of the X axis and $V(T^*f)$. Since T^* preserves the degree of polynomials, the theorem is proved in general. □

More generally, the same techniques allow us to argue that if $\mathfrak{a}$ is an ideal containing a polynomial f of degree n, then $V(\mathfrak{a})$ either contains a certain line, or contains at most n points on the line. It is also important to notice that over an algebraically closed field, a plane curve of degree n

intersects a line in *exactly* n places, because a one dimensional polynomial of degree n over an algebraically closed field splits into exactly n linear factors, and the only reason we can have less than n separate intersection points is if these intersections points overlap. In this case, this fact still remains true if we count these intersection points by their *multiplicity*. A generalization of this theorem is due to Bézout, which says that if we count curves which are asymptotically parallel along certain branches as 'intersecting at infinity', then the number of points of intersection between a curve of degree n and a curve of degree m is exactly the product nm. This result takes quite a bit of machinery, but we will prove it in due time. For now, the most interesting result of this theorem is to prove that certain planar curves are *not* affine varieties.

Example. *The set of points (x, y) in the real affine plane satisfying $y = \sin(x)$ cannot form a planar curve, because the curve intersects the X axis infinitely often, yet the set of points does not contain the X axis. If the points did form an algebraic variety $V(\mathfrak{a})$, where $\mathfrak{a}$ contains some nonzero polynomial f, then $V(f) \supset V(\mathfrak{a})$ would intersect the X axis infinitely often, which is clearly impossible.*

Example. *The complex sphere is the set of points (z, w) in the complex affine plane satisfying $|z|^2 + |w|^2 = 1$, because the intersection of the complex sphere with the z axis forms a circle, which has infinitely many points. This justifies that the set cannot form a planar curve, and the same technique as in the last example shows the sphere cannot be an affine variety in general.*

Example. *The set of points $\{(\cos t, \sin t, t) : t \in \mathbf{A}^3\}$ over real affine space is not a variety, because it contains infinitely many points of the form $(1, 0, \pi n)$, which lie on the same line.*

1.3 The Structure of Ideals and Algebraic Sets

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

- If $S \subset T$, then $V(T) \subset V(S)$.
- If $\mathfrak{a}$ is the smallest ideal containing S, then $V(S) = V(\mathfrak{a})$, so every affine variety can be described as the common zeroes of some ideal.

- If we have a family $\{\mathfrak{a}_\alpha\}$ of ideals, then $V(\bigoplus \mathfrak{a}_\alpha) = \bigcap V(\mathfrak{a}_\alpha)$, so the intersection of an arbitrary family of varieties forms a variety.
- For any two polynomials f and g, $V(fg) = V(f) \cup V(g)$. More generally, if $\mathfrak{a}$ and $\mathfrak{b}$ are ideals, then $V(\mathfrak{ab}) = V(\mathfrak{a}) \cup V(\mathfrak{b})$, so finite unions of varieties are varieties.
- $V(0) = \mathbf{A}^n$, $V(1) = \varnothing$, and for any $a \in K^n$, $V(X_1 - a_1, \ldots, X_n - a_n)$ is just the singleton set $\{a\}$. It follows from the last point that finite point sets are varieties.

Because of these properties, we begin to see that the analysis of $K[X_1, \ldots, X_n]$ and its ideals is key to the study of algebraic geometry. This is the key idea used to attack problems in algebraic geometry, and by the end of this section we will have seen this correspondence strengthened tenfold.

The space of all curves in $\mathbf{A}^2$ is infinite dimensional, and can only be effectively analyzed using the methods of functional analysis. However, we are not worried with this process in algebraic geometry, because even if a variety is specified by an ideal $\mathfrak{a}$ consisting of infinitely many polynomials, the variety is still a finitary object.

Proposition 1.2. *Every variety can be specified as the common zeroes of a finite set of polynomials.*

Proof. Due to the reduction of the study of varieties to the study of ideals $K[X_1, \ldots, X_n]$, we rely on an important result in the ideal theory of $K[X_1, \ldots, X_n]$, known as Hilbert's basis theorem, which states that every ideal of polynomials is finite generated, or equivalently, that $K[X_1, \ldots, X_n]$ is *Noetherian*. If W is any variety, we can write $W = V(\mathfrak{a})$ for any ideal $\mathfrak{a}$, and if we use Hilbert's basis theorem to write $\mathfrak{a} = (f_1, \ldots, f_n)$ as being generated by finitely many polynomials, and then $W = V(\mathfrak{a}) = V(f_1, \ldots, f_n)$ is specified as the common zeroes of finitely many polynomials. □

Example. *The varieties of $\mathbf{A}^1$ are exactly the finite point sets (other than the trivial variety $V(0) = \mathbf{A}^1$. First, note that since $K[X]$ is a principal ideal domain, we may assume we are considering the varieties of the form $V(f)$ for some particular polynomial $f(X) = \sum a_i X^i$. We know that $f(a) = 0$ if and only if $X - a$ is one of the prime factors of f. Since f decomposes into finitely many prime factors, it follows that $V(f)$ can consist of at most $\deg(f)$ points, a*

finite quantity. Thus the theory of one dimensional algebraic geometry is essentially trivial. This example shows that the countable union of affine varieties need not be a variety, because the countable union of finite sets need not be finite.

Example. *If K is a finite field, then all subsets of $\mathbf{A}^n$ are varieties, because all subsets of $\mathbf{A}^n$ are finite subsets.*

Proposition 1.3. *If f is a non constant polynomial over an algebraically complete field, $\mathbf{A}^n - V(f)$ contains infinitely many points for $n \geqslant 1$, and $V(f)$ contains infinitely many points for $n \geqslant 2$.*

Proof. First, recall that every algebraically complete field K must have infinitely many points, because if K only contains $a_1, \ldots, a_n$, we are unable to factor $(X - a_1) \ldots (X - a_n) + 1$ into linear factors. It follows that $\mathbf{A}^1 - V(f)$ is infinite for any polynomial $f \in K[X]$, because $V(f)$ is finite. Given any polynomial $f \in K[X_1, \ldots, X_n]$, there is a line in $\mathbf{A}^n$ upon which f is not identically zero (for otherwise f is equal to zero everywhere), and reducing our argument to the one dimensional case, we see that infinitely many points on this line cannot be zeroes of f. Arguing similarily, given any $f \in K[X_1, \ldots, X_n]$, there is a plane upon which $f \neq 0$, and so we must show that any nonconstant $f(X, Y) = \sum a_{ij} X^i Y^j$ in the plane has infinitely many zeroes. Without loss of generality, we may assume that $V(f)$ does not contain the origin. Assuming this, the intersections of $V(f)$ with the lines through the origin break $V(f)$ into disjoint classes of points, and provided we can show infinitely many of these classes are nonempty, we can conclude that $V(f) = \varnothing$. For any line through the origin of the form $Y = aX$, the polynomial takes the form $f(X, aX) = \sum a_{ij} a^j X^{i+j}$, and we know this polynomial has a zero unless it is constant, and if this occurs, then for any $1 \leqslant m < \infty$, $\sum a_{(m-k)k} a^k = 0$. Each of these are polynomials in $K[a]$, and at least one of these polynomials is nonzero, so we conclude there can only be finitely many values a such that $[a : 1]$ does not have an intersection with $V(f)$, and it follows that $V(f)$ is infinite, because K is infinite. □

If X is any set, then we shall let $I(X)$ be the subset of $K[X_1, \ldots, X_n]$ of polynomials which vanish over X. The set forms an ideal, and it is clear that in the case where $X = V(S)$, 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 $V(\mathfrak{a})$ from an ideal $\mathfrak{a}$. We make a few elementary observations about this operator.

- It is clear that if $X \subset Y$, then $I(Y) \subset I(X)$.
- $I(\varnothing) = K[X_1, \ldots, X_n]$, and $I(\mathbf{A}^n) = (0)$.
- $S \subset I(V(S))$ for any subset S of polynomials, and $X \subset V(I(X))$.
- Combining the last two points, it follows that $V(I(V(S))) = V(S)$, because $V(S) \subset V(I(V(S)))$ follows from the second point of the last bullet, and $V(I(V(S)) \subset V(S)$ follows because $S \subset I(V(S))$. Similarily, we can argue that $I(V(I(X))) = I(X)$. Thus if X is an algebraic set, then $V(I(X)) = X$, and if $\mathfrak{a}$ is an ideal equal to $I(X)$ for some set X, then $I(V(\mathfrak{a})) = \mathfrak{a}$.
- If $f^n \in I(X)$, then $f^n(p) = 0$ for all $p \in X$, which implies $f(p) = 0$ because K is an integral domain, so that $f \in I(X)$. This means exactly that $I(X)$ is a *radical ideal* (a radical ideal $\mathfrak{a}$ is an ideal such that if $x^n \in \mathfrak{a}$ is in x. The smallest radical ideal containing some ideal $\mathfrak{a}$ is denoted $\mathrm{Rad}(\mathfrak{a})$.

Proposition 1.4. *For any two algebraic sets V and W, $I(V) = I(W)$ if and only if $V = W$.*

Proof. This follows because $V(I(V)) = V$, and $V(I(W)) = W$, so if $I(V) = I(W)$, then $V = V(I(V)) = V(I(W)) = W$. □

This is clearly not true if V and W are not algebraic sets. For instance, if Y is the closure of some open set X, then $I(X) = I(Y)$, because polynomials are continuous so if they vanish on X, they certainly vanish on Y. Thus the class of algebraic sets are 'separated' in some manner. The fact also shows that there is a certain correspondence between algebraic sets and radical ideals. If two algebraic sets determine the same radical ideal, they are equal. We will soon see that if we are working over an algebraically closed field, and $\mathfrak{a}$, $\mathfrak{b}$ are radical ideals, then $V(\mathfrak{a}) = V(\mathfrak{b})$ 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 1.5. *If V is an algebraic set in $\mathbf{A}^n$, and $p \notin V$, then there is a polynomial f which vanishes on V, but with $f(p) = 1$.*

Proof. Since $V \neq V \cup \{p\}$, and V and $V \cup \{p\}$ are both algebraic sets, $I(V) \neq I(V \cup \{p\})$, and since $I(V) \supset I(V \cup \{p\})$, there must be a polynomial f which vanishes on V, but with $f(p) \neq 0$. It follows by normalizing that we can assume $f(p) = 1$. □

Similarily, by taking an algebraic set V, and n points $p_1, \ldots, p_n \notin V$, we may apply this theorem to find polynomials $f_1, \ldots, f_n \in I(V)$ with $f_i(p_j) = \delta_{ij}$. By considering linear combinations of the f_i, for any $a_{ij} \in K$, we can find $f_1, \ldots, f_n \in I(V)$ 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 finitely many points in the set.

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 (these subsets need not be reducible). Otherwise, we say V is **irreducible**. Ring theory allows us to characterize this criterion in terms of the ideal generating the ideal.

Proposition 1.6. *A variety V is irreducible if and only if $I(V)$ is prime.*

Proof. Suppose that $I(V)$ is not prime, so $fg \in I(V)$, whereas $f \notin I(V)$, $g \notin I(V)$. 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 $V(f, I(V))$ and $V(g, I(V))$ are proper subsets of V. But $V(f, I(V)) \cup V(g, I(V)) = V(fg, I(V)) = V(I(V)) = V$, so V is reducible. Conversely, if $V = W \cup U$, where W and U are proper algebraic subsets of V, then $I(V)$ is a proper subset of both $I(W)$ and $I(U)$, so we may select f vanishing on W, but not on V, and g vanishing on U, but not on all of V. This means that fg vanishes on $W \cup U = V$. Thus we have found $f, g \notin I(V)$, but with $fg \in I(V)$, so $I(V)$ cannot be prime. □

Example. *The parabola $V(Y - X^2)$ is an irreducible variety over an infinite field. First, we must justify that $I(V(Y - X^2)) = (Y - X^2)$. If $f(X, Y) \in K[X, Y]$ is a polynomial, then we may apply the division algorithm, viewing $K[X, Y]$ as the one dimension polynomial ring $K[X][Y]$ with coefficients in $K[X]$, to obtain that $f(X, Y) = g(X, Y)(Y - X^2) + h(X)$. If $f(x, x^2) = 0$ for all $x \in K$, then $h(x) = 0$ for all $x \in K$, so if we are working over an infinite field we conclude that $h = 0$, and therefore f is divisible by $Y - X^2$. Now we prove that*

$(Y - X^2)$ is prime, and since $K[X_1,\ldots,X_n]$ is a unique factorization domain, it suffices to show that $Y - X^2$ is an irreducible polynomial. If $Y - X^2$ is the product of two polynomials, write these two polynomials as $Yf(X)+g(X)$ and $h(X)$ (if there are more Y's in the factorization, they clearly cannot multiply to $Y - X^2$), where f is monic. But then $(Yf + g)(h) = Yfh + gh$, so $fh = 1$, implying that h is a unit.

As in most of mathematics, irreducible varieties have a nice theory, and we can use this theory to understand the varieties obtainable from the union of irreducible varieties. The idea 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. The ring theoretic property we need to employ here is the fact that $K[X_1,\ldots,X_n]$ is a *Noetherian ring* – every ascending chain of ideals is guaranteed to terminate.

Proposition 1.7. *Every variety is the finite union of irreducible varieties.*

Proof. If this theorem did not hold, we have justified that we can find an infinite sequence $V_1 \supsetneq V_2 \supsetneq V_3 \ldots$ of descending algebraic subsets. This implies that $I(V_1) \subsetneq I(V_2) \subsetneq I(V_3)$, an infinite ascending chain of ideals. Because $K[X_1,\ldots,X_n]$ is Noetherian, this situation cannot occur, so V_n must eventually be an irreducible variety, and this implies that the process of breaking reducible varieties in a decomposition must eventually yield a set of irreducible varieties. □

The last proposition guarantees the existence of a decomposition of an arbitrary variety V into a finite union $\bigcup V_i$ of irreducible varieties. If $V_j \subset V_k$ for $j \neq k$, then we may remove V_j from the union, and we still obtain a decomposition of V. We may therefore assume that no element of the decomposition is a subset of any other. Once we assume this, we obtain a unique decomposiiton. Suppose we have $\bigcup V_i = \bigcup W_i$, for two families of irreducible varieties, where none of the V_i is a subset of the V_j, and none of the W_i is a subset of the W_j. Then for each i, j, $V_i = (W_j \cap V_i) \cup (\bigcup_{k\neq j} W_k \cap V_i)$, and these finite intersections form varieties, so either $W_j \cap V_i = \varnothing$, or $W_j \cap V_i = V_i$. If $W_j \cap V_i = \varnothing$ for all j, then $V_i = V_i \cap \bigcup V_i = V_i \cap \bigcup W_i = \bigcup V_i \cap W_i = \varnothing$, which we assumed was impossible. Thus $V_i \subset W_j$ for some

j. Similarily, we may apply this technique to conclude that $W_j \subset V_k$ for some k, and by assumption, we must have $k = i$, so $W_j = V_i$. By matching up elements of the decomposition, we conclude that the decomposition is unique. The elements of this decomposition are known as the **irreducible components** of V.

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

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

Considering the zeroes which satisfy these equations on a case by case basis, we find that the variety is just the set of discrete points

$$\{(0,0),(1,1),(1,-1),(-1,-1),(-1,1),(-1,-i),(-1,i),(1,i),(1,-i)\}$$

and so the variety is a union of finitely many points, and this is the decomposition into irreducible factors.

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 $\{(0,0),(1,0)\}$. This is the consequence of the fact that $I(V(Y^2 + X^2(X-1)^2) = (Y, X(X-1))$ is not a prime ideal.*

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 1.8. *If two polynomials $f, g \in K[X,Y]$ are relatively prime, then $V(f,g)$ consists of finitely many points.*

Proof. If f and g have no common factor over $K[X,Y]$, then 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 $a = \sum a_i(X)Y^i$

and $b = \sum b_i(X)Y^i$, then we may find $c \in K[X]$ such that $ca, cb \in K[X,Y]$. This implies that $(ac)f + (bc)g = c$, and the values of x such that there is a y with $f(x,y) = g(x,y) = 0$ are contained within the roots of the polynomial c, because if $f(x,y) = g(x,y) = 0$, then the equation $(ac)f + (bc)g = c$ gives $c(x) = 0$. By symmetry, there can also only be finitely many values of y such that there is an x with $f(x,y) = g(x,y) = 0$, so in conclusion we find there can only be finitely many intersection points. □

Corollary 1.9. *If $f(X,Y)$ is irreducible, and $V(f)$ is infinite, then $I(V(f)) = (f)$, and $V(f)$ is irreducible.*

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

Corollary 1.10. *The irreducible algebraic planar sets over an infinite field are exactly $\mathbf{A}^2$, $\varnothing$, singletons, and irreducible plane curves $V(f)$, where f is irreducible and $V(f)$ is infinite.*

Proof. It is obvious that $\{p\}$ is an irreducible set, as is $\varnothing$. 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 $V(f)$ for some irreducible polynomial f, and these are the irreducible planar curves provided $V(f)$ is infinite, by the last lemma. □

Corollary 1.11. *If we are working over an algebraically closed field, and f is not irreducible, so we can write $f = f_1^{n_1} \dots f_m^{n_m}$ where each f_i is irreducible, then the irreducible components of $V(f)$ are exactly the $V(f_i)$. We find that $I(V(f)) = (f_1, \dots, f_n)$.*

Proof. It is clear that

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

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

Example. *Over the real numbers, $X^2 + Y^2 + 1$ is irreducible, yet no points in the real plane satisfy the equation $X^2 + Y^2 = 1$, so $I(V(X^2 + Y^2 + 1)) = \mathbf{R}[X,Y]$,*

which is not equal to (X^2+Y^2+1)*. Conversely,* X^2+Y^2+1 *is also irreducible over the complex numbers, but the solution set to the polynomial forms an irreducible complex curve.*

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. *As another example, note that the variety over the real numbers corresponding to* $Y^2 - XY - X^2Y + X^3 = (Y-X)(Y-X^2)$ *is the union of the line* $Y = X$ *and the parabola* $Y = X^2$*, and this is the decomposition into irreducible elements. The same is true for the decomposition over the complex numbers.*

Example. $Y^2 - X(X^2-1)$ *is an irreducible polynomial, and its solution set is infinite, so* $V(Y^2 - X(X^2-1))$ *is an irreducible variety both over the real and complex numbers. In the topology of the Euclidean plane,* $V(Y^2 - X(X^2-1))$ *is disconnected though, the union of a shape isomorphic to the disjoint 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* $Y = \sqrt{X(X^2-1)}$*, 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* $X \in \{-1, 0, 1\}$*, 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* $V(X^3 + X - X^2Y - Y)$ *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

We have seen the duality between affine varieties and radical ideals over the ring $K[X_1, \ldots, X_n]$. 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, but the classical proof of this result is annoying, so we will use it as motivation for the Nullstellensatz.

Theorem 1.12 (Study)**.** *If f and g are polynomials in the affine plane over an algebraically closed field with $V(f) \subset V(g)$, and f is irreducible, then f divides g.*

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

Lemma 1.13 (Weak Nullstellensatz)**.** *If K is an algebraically closed field, and if $\mathfrak{a}$ is a proper ideal of $K[X_1,\ldots,X_n]$, then $V(\mathfrak{a}) \neq \varnothing$.*

Proof. We shall actually prove that if $\mathfrak{a}$ is a maximal ideal, then $V(\mathfrak{a})$ 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 $K[X_1,\ldots,X_n]/\mathfrak{a} = L$ 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 $K[X_1,\ldots,X_n]$, 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 $(X_1 - a_1,\ldots,X_n - a_n)$ is a maximal ideal in $K[X_1,\ldots,X_n]$, because every polynomial in $K[X_1,\ldots,X_n]/(X_1 - a_1,\ldots,X_n - a_n)$ is congruent to an element of K, hence $\mathfrak{a} = (X_1 - a_1,\ldots,X_n - a_n)$. Now we can conclude that $V(\mathfrak{a}) = \{(a_1,\ldots,a_n)\}$. □

An important thing to note about this proof of the weak Nullstellensatz is that it implies that the maximal ideals of $K[X_1,\ldots,X_n]$ 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, that in a general ring A, we consider the set of maximal ideals of A as points in a space. This idea reoccurs later in our study of the local rings attached to varieties. 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 $K[X_1,\ldots,X_n]$ 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 1.14. *If $K[S_1,\dots,S_n]$ 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 $K[S_1,\dots,S_m]$, where $m < n$. We may then apply induction to $K[S_1,\dots,S_n] = K(S_1)[S_2,\dots,S_n]$ to conclude that $K[S_1,S_2,\dots,S_n]$ is a finite extension of $K(S_1)$. This means that for every S_i there are polynomials $a_{ij} \in K(S_1)$ such that $a_{0j} + a_{1j}S_j + \cdots + S_j^{n_j} = 0$. If we consider a polynomial $b \in K[S_1]$ large enough such that $a_{ij}b^j \in K[S_1]$ for all i, j, then we find that for each S_j, bS_j is integral over $K[S_1]$, and it follows from the theory of integral extensions (in particular, that the set of integral elements form a subring of the ring) that for any $f \in K[X_1,\dots,X_n]$, bf is integral over $K[S_1]$, and in particular bS_1 is integral over $K[S_1]$, implying that S_1 is algebraic over K, and therefore that $K(S_1)$ is a finite extension of K, hence $K[S_1,\dots,S_n]$ is a finite extension of K. □

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 $K[S_1,\dots,S_n]$ is an algebraic extension of K, and therefore $K[S_1,\dots,S_n] = K$. This finishes our proof of the weak nullstellensatz. We now consider the extension to the full nullstellensatz.

Theorem 1.15. *If $\mathfrak{a}$ is an ideal in $K[X_1,\dots,X_n]$, where K is algebraically closed, then $I(V(\mathfrak{a})) = Rad(\mathfrak{a})$.*

Proof. We may assume that $\mathfrak{a}$ is generated by finitely many polynomials, so $\mathfrak{a} = (f_1,\dots,f_m)$. Concretely, the nullstellensatz says that if $g \in I(V(f_1,\dots,f_m))$, then $g^n = \sum h_i f_i$ for some $h_i \in K[X_1,\dots,X_n]$. Suppose that $g \in I(V(\mathfrak{a}))$. Consider the ideal $\mathfrak{b} = (f_1,\dots,f_m, X_{n+1}g-1) \subset K[X_1,\dots,X_{n+1}]$. Then $V(\mathfrak{b}) = \varnothing$, since if $f_1(x) = \cdots = f_n(x) = 0$, then $g(x) = 0$, so $x_{n+1}g(x) - 1 = -1$. The weak nullstellensatz implies that there are $a_i \in K[X_1,\dots,X_{n+1}]$ such that $\sum a_i f_i + b(X_{n+1}g - 1) = 1$. Introducing $Y = 1/X_{n+1}$, we may multiply the equation by Y^N for a large enough N to find that

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

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

Corollary 1.16. *There is a one to one correspondence between radical ideals and algebraic sets in affine space over an algebraically closed field.*

Corollary 1.17. *If $\mathfrak{a}$ is a prime ideal, then it is also a radical non total ideal, so $V(\mathfrak{a}) \neq \varnothing$ 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. *$V(Y^2 - X(X-1)(X-\lambda))$ 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 1.18. *If $f \in K[X_1, \ldots, X_n]$ has a decomposition as $f_1^{n_1} \ldots f_m^{n_m}$, where K is algebraically closed, then $V(f) = V(f_1 \ldots f_n) = \bigcup V(f_i)$ 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 gives that K is an algebraically closed field. Since $K[X]$ is a principal ideal domain, the theorem states that if $(f) \neq K[X]$, then $V(f) \neq 0$, which means that if f is a non constant polynomial, then f has a root. Correspondingly, none of the corollaries to the weak nullstellensatz hold in non algebraically closed fields either. Suppose f is a polynomial without a root in K. Then $V(f) = V(K[X])$, yet f and $K[X]$ are both radical ideals, so we don't obtain a one to one correspondence. f and $K[X]$ are also prime, so the correspondence between irreducible varieties and prime ideals is not one to one.

Example. *If q is a prime element of a unique factorization domain, then every prime ideal $\mathfrak{a} \subset (q)$ is either trivial or equal to (p). To see this, assuming $\mathfrak{a} \neq (0)$ find a nonzero $p_1^{n_1} \ldots p_m^{n_m} q^k \in \mathfrak{a}$ minimizing $k + \sum n_i$. Then we must have $k > 0$, so $q(p_1^{n_1} \ldots p_m^{n_m} q^{k-1}) \in \mathfrak{a}$, and because of our minimization, $p_1^{n_1} \ldots p_m^{n_m} q^{k-1} \notin \mathfrak{a}$, so $q \in \mathfrak{a}$. This implies that if V is an irreducible hyperplane, there is no irreducible variety containing V, except for $\mathbf{A}^n$ itself.*

Example. *Sometimes, we have to be a bit clever to determine if an ideal is reducible. Consider the ideal* $(X^2 - Y^3, Y^2 - Z^3)$ *in* $K[X,Y,Z]$*, where* K *is algebraically closed. Consider the homomorphism* f *from* $K[X,Y,Z]$ *to* $K[T]$ *preserving elements of* K*, and with* $X \mapsto T^9$*,* $Y \mapsto T^6$*, and* $Z \mapsto T^4$*. Then certainly* $(X^2 - Y^3, Y^2 - Z^3)$ *is contained in the kernel of* f*. But an arbitrary element of* $K[X,Y,Z]/(X^2 - Y^3, Y^2 - Z^3)$ *can be denoted* $a + bX + cY + dXY$*, with* $a,b,c,d \in K[Z]$*, and if* $a = \sum a_i Z^i$*,* $b = \sum b_i Z^i$*,* $c = \sum c_i Z^i$*,* $d = \sum d_i Z^i$*, then* $a + bX + cY + dXY$ *maps to*

$$\sum a_i T^{4i} + \sum b_i T^{9+4i} + \sum c_i T^{6+4i} + \sum d_i T^{15+4i}$$

and since the terms of each sum occur over different residues mod four, $f(a + bX + cY + dXY) = 0$ *if and only if* $a + bX + cY + dXY = 0$*, hence the kernel of* f *is exactly* $(X^2 - Y^3, Y^2 - Z^3)$*. Since* $K[T]$ *is an integral domain, this shows that* $(X^2 - Y^3, Y^2 - Z^3)$ *is prime, and therefore that* $V(X^2 - Y^3, Y^2 - Z^3)$ *is an irreducible variety.*

Chapter 2

Intrinsic Properties of Algebraic Varieties

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 spatial structure of X. Viewing an algebraic variety V as a space, it seems difficult to think of which functions on V are the natural ones. Studying continuous functions on V can give us certain topological information about the variety, such as connectedness, compactness, and so on, but this information doesn't seem very related to the definition of V in terms of $K[X_1,\ldots,X_n]$. To make the connection between V and $K[X_1,\ldots,X_n]$, we make a decision that the natural functions 'should' be the polynomial functions. We define the **coordinate ring** of V, also known as the ring of **regular functions** on V, denoted $K[V]$, to be the space of all functions which are the restriction of some polynomial function on $\mathbf{A}^n$.

Algebraically, $K[V]$ forms an algebra over K, where the functions in $K[V]$ corresponding to elements of K correspond to the constant functions on V. There is an alternate description of the coordinate ring which

is more amenable to algebraic manipulation. The homomorphism from $K[X_1,\ldots,X_n]$ to $K^{\mathbf{A}^n}$ obtained by mapping a polynomial to the function it defines can be composed with the restriction homomorphism from $K^{\mathbf{A}^n}$ to K^V, and the image of this composition is exactly $K[V]$. Applying the first isomorphism theorem, we find that $K[V]$ is isomorphic to $K[X_1,\ldots,X_n]$ modulo the kernel of the homomorphism. This kernel is exactly the space of polynomials which vanish over V, which we previously denoted $I(V)$, so we find that $K[V]$ is isomorphic to $K[X_1,\ldots,X_n]/I(V)$. This is useful for proving things formally about the coordinate ring of a variety.

Example. *If we are working over an infinite field, then the coordinate ring $K[\mathbf{A}^n]$ is equal to $K[X_1,\ldots,X_n]$, because $I(\mathbf{A}^n) = 0$. This is well known in the univariate case. In general, if we have a nonzero polynomial $f(X_1,\ldots,X_n,Y)$, then viewing the polynomial as a univariate polynomial in Y with coefficients in $K(X_1,\ldots,X_n)$, we conclude that there must be $y \in K$ with $f(X_1,\ldots,X_n,y) \neq 0$, and by induction we conclude there are $x_1,\ldots,x_n \in K$ with $f(x_1,\ldots,x_n,y) \neq 0$. In particular, in an algebraically closed field this will always be true.*

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 $K[X_1,\ldots,X_n]$, 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 $I_V(W)$ is the ideal of functions in $K[V]$ vanishing on W, then $K[W]$ is isomorphic to $K[V]/I_V(W)$, 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.

Proposition 2.1. *V is a finite variety if and only if $K[V]$ is a finite vector space over K, and in this case the dimension is equal to the number of points in the variety.*

Proof. If $p_1,\dots,p_n \in V$, we have seen that we can choose $f_1,\dots,f_n \in K[V]$ with $f_i(p_j) = \delta_{ij}$. Then $f_1,\dots,f_n$ are linearly independent in $K[V]$, because if $g = \sum \lambda_i f_i$ vanishes on V, then $g(p_j) = \sum \lambda_i f_i(p_j) = \lambda_j = 0$. In particular, if V has infinitely many points, then $K[V]$ is an infinite dimensional vector space over K. Conversely, if $p_1,\dots,p_n$ are the only points in V, then $K[V]$ is spanned by the f_i, so $K[V]$ is finite dimensional. □

One can extend this result, using a technique which will become more useful later in our analysis of the ring theory of $K[X_1,\dots,X_n]$.

Lemma 2.2. *Over an algebraically complete field, for any ideal $\mathfrak{a}$ of $K[X_1,\dots,X_n]$ such that $V(\mathfrak{a})$ is finite, $K[X_1,\dots,X_n]/\mathfrak{a}$ is finite dimensional over K. The number of points in $V(\mathfrak{a})$ is bounded by the dimension of the vector space.*

Proof. We have already seen this theorem if $\mathfrak{a}$ is a radical ideal, for then $I(V(\mathfrak{a})) = \mathfrak{a}$, and so $K[V]$ is isomorphic to $K[X_1,\dots,X_n]/\mathfrak{a}$. In general, $I(V(\mathfrak{a})) = \mathrm{Rad}(\mathfrak{a})$, which we shall denote by $\mathfrak{b}$. 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 $\mathfrak{b}^n \subset \mathfrak{a} \subset \mathfrak{b}$. We then have the following sequence of surjective maps

$$K[X_1,\dots,X_n]/\mathfrak{b}^n \to \cdots \to K[X_1,\dots,X_n]/\mathfrak{b}^2 \to K[X_1,\dots,X_n]/\mathfrak{b}$$

The kernel of the map from $K[X_1,\dots,X_n]/\mathfrak{b}^{i+1}$ to $K[X_1,\dots,X-n]/\mathfrak{b}^i$ is $\mathfrak{b}^i/\mathfrak{b}^{i+1}$, which is finite dimensional, because if $\mathfrak{b}^i = (f_1,\dots,f_m)$, then the f_i span $\mathfrak{b}^i/\mathfrak{b}^{i+1}$ because $f_i f_j \in \mathfrak{b}^{i+1}$. Using the rank nullity theorem, we can use induction to prove that $K[X_1,\dots,X_n]/\mathfrak{b}^i$ is a finite dimensional vector space over K for each i. The base case is that $K[X_1,\dots,X_n]/\mathfrak{b}$ is finite dimensional, and the general case is that

$$\dim K[X_1,\dots,X_n]/\mathfrak{b}^{i+1} = \dim \mathfrak{b}^{i+1}/\mathfrak{b}^i + \dim K[X_1,\dots,X_n]/\mathfrak{b}^i$$

and the sum of two finite values is finite. Since $\mathfrak{b}^n \subset \mathfrak{a}$, we have a surjective map from $K[X_1,\dots,X_n]/\mathfrak{b}^n$ to $K[X_1,\dots,X_n]/\mathfrak{a}$, so in particular $K[X_1,\dots,X_n]/\mathfrak{a}$ is finite dimensional. □

Example. *Consider the locus V of the polynomials $Y^2 - X^2$ and $Y^2 + X^2$ over an algebraically closed field K. Since $(Y^2-X^2, Y^2+X^2) = (Y^2, X^2)$, the radical ideal of these polynomials is (Y, X), and so $K[V] = K[X,Y]/(Y,X) \cong K$ is a one dimensional vector space over K. This makes sense, because $Y^2 - X^2 = (Y-X)(Y+X)$, so on V we find $Y = X$ or $Y = -X$, and in either of these cases $Y^2 + X^2 = 0$ if and only if $X = Y = 0$.*

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 $f_1, \ldots, f_m$ 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 $f = (TX_1, \ldots, TX_n)$, 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 $X_1, \ldots, X_n$, 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 $1, X_1, \ldots, X_n$ 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, just like how $C(X)$ classifies X when X is a compact Hausdorff space, though this is much more difficult to prove.

Proposition 2.4. *The image of an irreducible variety under a polynomial map is an irreducible variety.*

Proof. Let $f : V \to W$ be a surjective polynomial map between two varieties, and 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 $W_2 = W$. □

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 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 $(a_1,\ldots,a_{n+1}) \in \mathbf{A}^{n+1}$, where $(a_1,\ldots,a_n) \in V$ and $a_{n+1} = f(a_1,\ldots,a_n)$. $G(f)$ is isomorphic to V under the projection map $(a_1,\ldots,a_{n+1}) \mapsto (a_1,\ldots,a_n)$, because for each $a_1,\ldots,a_n$ 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 $V(Y^2 - X^3)$ 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 by $\mathbf{A}^n$ rather than K^n, because the isomorphisms mean that the particular choice of affine coordinates used to define varieties is of no real consequence to the affine geometry.

Example. *The affine subplanes of $\mathbf{A}^n$ are varieties known as* **linear subvarieties**. *Any variety of the form $V(f_1,\ldots,f_m)$, 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 $K[X_1,\ldots,X_n]$.*

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 $K[X_1,\ldots,X_n]$ 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 $g(p)/h(p)$ is defined irrespective of the choice of g and h, for if $g_0/h_0 = g_1/h_1$, then $h_1 g_0 = h_0 g_1$, and so $h_1(p)g_0(p) = h_0(p)g_1(p)$, hence $g_0(p)/g_1(p) = h_0(p)/h_1(p)$.

Though $K(V)$ does not classify the isomorphism classes of algebraic varieties, an isomorphism between $K(V)$ and $K(W)$ still provides strong

relations between the varieties V and W. To consider this relation, we say that a set $X \subset V$ is **Zariski dense** in V if V is the only subvariety of V to contain all the points in X. This, in particular, implies that any function $f \in K[V]$ which vanishes on X also vanishes on V, for the set of zeroes of f form a subvariety of V containing X. In the more modern context of algebraic geometry, this density can be interpreted as a topological density of a set in the **Zariski topology** of a variety.

Theorem 2.5. *The K homomorphisms $T : K(W) \to K(V)$ are in one to one correspondence with maps $f : V \to W$ definable by rational functions of the coordinates (so f is really a partial map) 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 $(X_1,\ldots,X_n)$, and $W \subset \mathbf{A}^m$ is definable in the coordinate $(Y_1,\ldots,Y_m)$, then let $T(Y_i) = f_i$. We claim that $(f_1,\ldots,f_m)$ gives a surjective map from V to W, where defined. Suppose that $y = (f_1(x),\ldots,f_m(x))$ is not an element of W. Then there is $g \in I(W)$ with $g(y) = 1$. If $g(Y) = \sum a_\alpha Y^\alpha$, then since g is equal to 0 in $K(W)$, we find that $T(g) = \sum a_\alpha f^\alpha = 0$ in V, so in particular $\sum a_\alpha f^\alpha(x) = g(y) = 0$, which is impossible. Similarily, 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 $g(y) = 1$, and if $g = \sum a_\alpha Y^\alpha$, 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 algebraic varieties to the category of fields, so in particular, if we find T^{-1} for some isomorphism T, then the rational functions $(f_1,\ldots,f_m)$ corresponding to T and the rational functions $(g_1,\ldots,g_n)$ 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.

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 $h \in K[X_1, \ldots, X_n]$ 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 theset of possible denominators for fractional expressions of f, and so $V(\mathfrak{a})$ gives the set of poles of f. If f has no poles, then $V(\mathfrak{a}) = \varnothing$, so applying the nullstellensatz, we conclude that $\mathrm{Rad}(\mathfrak{a}) = K[X_1, \ldots, X_n]$, 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.

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 $X = Y = 0$. Now $K[V] = K[X, Y, W, Z]/(XW - YZ)$, and so $K(V)$ contains the function $f = X/Y = Z/W$, which is defined at all points except where $Y = W = 0$. The ideal of 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 $K[X_1, \ldots, X_n]$, 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 Local Rings

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 of functions f 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.

Example. *Consider the reducible curve $XY = 0$. Then $\mathbf{C}[XY] \cong \mathbf{C}[X,Y]/(XY)$, because (XY) is a radical ideal. Every element of $\mathbf{C}[XY]$ is equivalent to a unique polynomial of the form $a + Xf(X) + Yg(Y)$, where f and g are univariate polynomials. Consider the local ring $\mathcal{O}_0(XY) = S^{-1}\mathbf{C}[XY]$, where S is the set of functions in $\mathbf{C}[XY]$ not equal to zero at the origin. Since localization commutes with taking quotients, we find that $\mathcal{O}_0(XY)$ is isomorphic to $\mathcal{O}_0(\mathbf{A}^n)/(XY)$ (and this isomorphism in fact preserves the evaluation of polynomials at zero on $V(XY)$). We find that elements of $\mathcal{O}_0(XY)$ can be written in the form $f(X,Y)/g(X,Y)$, where $g(0,0) \neq 0$, and where f and g contain no mixed term monomials, and where two rational functions are identified if*

they agree with one another on an axis. On the other hand, if $p = (a, 0)$, then $\mathcal{O}_p(XY)$ is the set of functions of the form $f(X,Y)/g(X,Y)$, where $g(a,0) \neq 0$ and where f and g have no mixed term monomials, and where two rational functions are identified if they agree with each other on the X axis. Similarily, 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 $p_1, \ldots, p_n$. 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 such that $V(\mathfrak{a})$ consists of finitely many points $p_1, \ldots, p_n$, then $K[X]/\mathfrak{a}$ is isomorphic to the direct product of $\mathcal{O}_{p_i}(\mathbf{A}^n)/S_{p_i}^{-1}\mathfrak{a}$.*

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 $\mathfrak{a}$, $\sum e_i \equiv$ modulo $\mathfrak{a}$, $T_i(e_i) = 1$, and $e_i e_j \in \mathfrak{a}$. This implies that if $T_i f = 0$ for all i, then $e_i f \in \mathfrak{a}$, and therefore that $f \equiv (\sum e_i) f = \sum e_i f \in \mathfrak{a}$, which implies that $f \in \mathfrak{a}$, so that the product map $(T_1, \ldots, T_n)$ is injective. To prove surjectivity, we consider an arbitrary point $(a_1/s_1, \ldots, a_n/s_n) \in \mathcal{O}_i$. Since $s_i(p_i) \neq 0$, we can write $t_i s_i = e_i$, which implies that $a_i/s_i = a_i t_i/e_i = a_i t_i$. Since $e_i e_j \in \mathfrak{a}$, $T_i(e_j) = T_i(e_i e_j) = 0$, so the image of $\sum a_i t_i e_i$ by the map T is $(a_1/s_1, \ldots, a_n/s_n)$. 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 $\mathfrak{b} = V(I(\mathfrak{a}))$, and since $\mathfrak{b}$ is finitely generated we can choose n such that $\mathfrak{b}^n \subset \mathfrak{a}$. Let $\mathfrak{m}_1, \ldots, \mathfrak{m}_n$ be the maximal ideals of $K[X]$

corresponding to the points $p_1, \ldots, p_n$. 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. *The dimension of $K[X]/\mathfrak{a}$ over K is the sum of the dimensions of $\mathcal{O}_i$ over K.*

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, and we call these types of rings local rings. 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 maps $f : V \to W$ 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.

Chapter 3

Algebraic Curves

We now focus on a particular area of algebraic geometry, the theory of planar algebraic curves. This is one of the classical areas of algebraic geometry, which is still a wide source of research material today. Though we think of a planar algebraic curve as a variety in the plane, it is often more elegant to extend the definition of algebraic curves to a more general object. 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. However, we think of f defining the same curve as af, for any $a \in K$, because scaling doesn't change the weight of a zero set. Thus we generalize the definition of a planar algebraic curve to be an equivalence class of curves in the plane, where f is identified with af for each nonzero $a \in K$. That is, an algebraic curve is a non-constant element of $\mathbf{P}(K[X_1, \dots, X_n])$. One cannot add algebraic curves like in the theory of polynomials, but we can still multiply them. Though we still talk of planar curves as algebraic subsets of the plane, one should always think of such curves as being defined by a particular polynomial.

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 TC_p by the equation $f_X(p)(X - a) + f_Y(p)(Y - b)$. These are the set of points which lie on $V(f)$ 'up to first order'. Over

the real numbers, this implies that we can locally parameterize C by a differentiable map with respect to one of the variables. Over other fields, we have a more modest equivalence to this parameterization.

Theorem 3.1. *If f is nonsingular at p, then $\mathcal{O}_p(f)$ is a discrete valuation ring.*

Proof. Suppose, without loss of generality, that $p = 0$, and $f_Y(p) \neq 0$. Write

$$f(X,Y) = Xg(X) - Yh(X,Y)$$

If $f_Y(p) \neq 0$, then $h(0) = f_Y(0) \neq 0$. In the coordinate ring $K[f]$, $Xg(X) = Yh$, and since $h(0) \neq 0$, we can write $Y = Xg/h = Xk$ in $\mathcal{O}_p(f)$. If u/v is any rational function in $\mathcal{O}_p(f)$ with $u(0) = 0$, then we can write $u(X,Y) = Xu_0 + Yu_1 = X[u_0 + ku_1]$, so X divides u/v and as such the maximal ideal of $\mathcal{O}_p(f)$ is principal, generated by X. It remains to show that $\mathcal{O}_p(f)$ is a domain if p is nonsingular. To prove this, we write $f = g_1^{n_1} \dots g_m^{n_m}$, where g_i is irreducible. Then $g_i(p) = 0$ for some p, and $n_i = 1$ here, because otherwise the product rule implies the derivative vanishes here. The same reason implies that $g_j(p) \neq 0$ for any $i \neq j$. The Nullstellensatz implies that $I(f) = (g_1 g_2 \dots g_n)$, and since only one of these results vanishes at p, the set of elements of $K[X_1, \dots, X_n]$ which have null divisors is generated by g_1, because if g has a null divisor h with $h(p) \neq 0$, then g_i does not divide h, yet $g_1 \dots g_n$ must divide gh, so g_i must divide g. When we form $\mathcal{O}_p(f)$, we can first quotient $K[X,Y]$ by the smallest ideal containing $I(V)$ and all zeroes divisors, and then form the ring of fractions without having to form zero divisors. In this case, the smallest ideal is (g_i), which is a prime ideal because g_i is irreducible, hence the quotient is an integral domain, and as a result of the ring of fractions is also an 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. Since $\mathcal{O}_p(f)$ is a domain, and as a localization of a Noetherian ring, is Noetherian, we conclude that $\mathcal{O}_p(f)$ is a discrete valuation ring. □

Example. *More generally, this argument shows that if $aX + bY = 0$ is not tangent to p, then it is a uniformizing parameter for $\mathcal{O}_p(X)$, because the coordinate system is arbitrary. We shall show that if $aX + bY$ is not tangent to the curve at the origin, then it has order one, and if it is the tangent at the origin, it*

has order greater than one. Assume without loss of generality that $f_Y(p) \neq 0$. Then, using the notation from the last proof,

$$aX + bY = X[c + dk] = X\frac{ch + dg}{h}$$

and $ch(0) + dg(0) = cf_Y(0) + df_X(0) = 0$ holds only when $cX + dY$ is the tangent line at the origin, so if the line isn't the tangent, it has order one.

Example. *If $p_1, \ldots, p_m$ are simple points on a curve C, and the m_i are non-negative integers, then we can find a function $f \in K[C]$ with $ord_p(f) = m_i$. We just take $f = L_1^{n_1} \ldots L_m^{n_m}$, where L_i is a line not tangent to f at p, but passing through p, and not passing though any other point. More generally, if C is a irreducible curve, we can find $f \in K(C)$ with $ord_p(f) = m_i$ for any integers m_i, by the exact same construction.*

Example. *A simple point p is called a* **flex** *if its tangent line has order $\geqslant 3$. 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 Y axis, we may write the equation defining the curve as $Y = aX^2 + X^3 g(X) + Yh(X, Y)$, where $h(0) = 0$. This implies that in $\mathcal{O}_p(C)$,*

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

If $a \neq 0$, this implies that the tangent line to the curve has order two. If $a = 0$, then we could have an order one tangent if $g(X) = 0$, so the equation defining the curve is $Y = Yh(X, Y)$, so the irreducible curve containing the origin is $Y = 0$, in which case the tangent has infinite order. Otherwise, we may write $Y[1 - h] - X^3 g(X) = (X - b_1)^{n_1} \ldots (X - b_m)^{n_m}[Yh_0 - X^3 g_0]$, where $Yh_0 - X^3 g_0$ contains no common factors. If $Xg(X) = 0$, then there is k such that $Xg(X) = k[Yh_0 - X^3 g_0]$, so we may write $k = Xk_0$, and $g = k_0[Yh_0 - X^3 g_0]$. But this implies k_0 and $Yh_0 - X^3 g_0$ is a function of X, which is impossible unless $g = 0$ (the case we have already considered), or if $h_0 = 0$, in which case

$$Y[1 - h] - X^3 g(X) = -X^3 (X - b_1)^{n_1} \ldots (X - b_m)^{n_m} g_0(X)$$

But this is impossible because Y occurs on the left hand side, but not on the right hand side. Thus if $a = 0$, and we are not considering $Y = 0$, then $Xg(X) \neq 0$, and it has an extra zero at the origin so ord$(Y) > 2$.

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. Since $\mathcal{O}_p(f)$ 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 $V(f)$, 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 $V(f)$ 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 $V(f)$ 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 $V(f)$, 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 nozero, 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 behaviour of the polynomial around the origin is determined by the lowest order nonzero terms x^4y is negligable compared to x^4. Thus, if we write a polynomial f as the sum of homogenous polynomials, the smallest nonzero homogenous polynomial will determine the local behaviour of the function. At least with respect to tangents, this correspondence is 'roughly exact', because the locus of solutions to a homogenous polynomial is just a union of lines, which we can

view as tangent lines to the polynomial. Over an algebraically closed field, Study's lemma implies that the linear form defining each line through the origin divides f, so the algebraic curve can only be a union of finitely many lines.

Recalling the theory of homogenous polynomials, we note that the polynomial ring $K[X_1,\ldots,X_n]$ has a K vector space decomposition into the subspaces $K_m[X_1,\ldots,X_n]$ of homogenous polynomials of degree m. If f is an arbitrary polynomial, then we denote f_m the monomials of order m in the decomposition of f. If f is a homogenous polynomial of degree m, and g homogenous of degree n, then fg is homogenous of degree $m+n$, so this decomposition turns $K[X_1,\ldots,X_n]$ into a graded algebra over K. The homomorphism from $K[X_1,\ldots,X_n]$ to $K[X_1,\ldots,X_{n-1}]$ given by mapping $f(X_1,\ldots,X_n)$ to $f(X_1,\ldots,X_{n-1},1)$ is a bijection when restricting the domain to $K_m[X_1,\ldots,X_n]$, and this map is known as the **dehomogenization**, and on this domain we denote the map by $f \mapsto f_*$. The reverse of this map takes a polynomial $f(X_1,\ldots,X_{n-1}) = \sum a_\alpha X^\alpha$ of degree n, and returns the homogenous polynomial $f^*(X_1,\ldots,X_n) = \sum a_\alpha X^\alpha X_n^{n-|\alpha|}$ of degree n, which is the form of smallest degree such that $(f^*)_* = f$. This process is known as **homogenization**. The map $f \mapsto f^*$ is 'almost' a homomorphism, in the sense that $(fg)^* = f^*g^*$, and $(f+g)^* = f^* + g^*$ if f and g have the same degree, but if $\deg(f) < \deg(g)$, say $\deg(f) + m = \deg(g)$, then $(f+g)^* = X_{n+1}^m f^* + g^*$. If f is a form of degree n and f_* has degree m, then $f = X_{n+1}^{n-m}(f_*)^*$. The process of homogenization will become much more useful later in the study of projective varieties, but for now, we use it to factorize forms in the plane. For any homogenous polynomial $f(X,Y)$, $f_*(X) = f(X,1)$ is a polynomial in a single variable, so it factorizes into linear variables. This implies that we can write $f(X,1) = a(X-\lambda_1)^{k_1}\ldots(X-\lambda_l)^{k_l}$, with $\sum k_i = n$, and homogenizing, we find that f is the product of factors of the form $aX+bY$, which are linear forms defining lines through the origin.

Now suppose that f is an arbitrary polynomial in the plane. Given an arbitrary point $p=(a,b)$, we can write f uniquely as a sum of monomials in the variables $X-a$ and $Y-b$. If the lowest order monomials of this sum are of degree m, then we say that f has multiplicity m at p, and we denote this value by $m_p(f)$. In this case, f_m is a nonzero homogenous polynomial in the variables $X-a$ and $Y-b$, and by our analysis above we find that it factors into linear forms in the variables X and Y, $f_m(X,Y) = g_1^{k_1}\ldots g_n^{k_n}$,

with $\sum k_i = m$. We call the lines defined by the equations g_i the **tangent lines** to f at p, where the line defined g_i has multiplicity k_i. These properties are preserved under affine translations, so we often assume when dealing theoretically with these polynomials that $p = 0$. We say a point p is a **double point** if it has multiplicity two, a **triple point** if it has multiplicity three, and so on and so forth. A point is an **ordinary multiple point** if its tangent lines are distinct, and an ordinary double point is often called a **node**.

Example. *A point $p = (a,b)$ lies on $V(f)$ if and only if $m_p(f) > 0$, for if*

$$f(X,Y) = f_0 + f_1(X-a, Y-b) + \cdots + f_n(X-a, Y-b)$$

and if $f_0 \neq 0$, then $f(a,b) = f_0 \neq 0$. Similarily, p is a nonsingular point on $V(f)$ if and only if $m_p(f) = 1$, because if

$$f(X,Y) = \lambda(X-a) + \gamma(Y-b) + f_2(X-a, Y_b) + \cdots + f_n(X-a, Y-b)$$

then $f_X(a,b) = \lambda$ and $f_Y(a,b) = \gamma$, so p is singular at p if and only if $\lambda = \gamma = 0$.

It is nice to know that, even though singularities occur, we can only have finitely many. This is because if f is an irreducible polynomial in the plane with $V(f_X) \cap V(f) \cap V(f_Y)$ containing infinitely many points, then in particular we conclude that f divides f_X and f_Y. Over a field of characteristic zero, this implies that $f_X = f_Y = 0$, so f takes a constant value. Over a field of characteristic p, it implies that $f(X,Y) = \sum a_{ij} X^{pi} Y^{pj} = (\sum a_{ij} X^i Y^j)^p$, and therefore f is not irreducible, contrary to the assumption.

Proposition 3.2. *If $T : \mathbf{A}^2 \to \mathbf{A}^2$ is a polynomial map with $T(p) = q$, then for any polynomial $g \in K[X_1, \ldots, X_m]$, $m_q(g) \leqslant m_p(f^*g)$. 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 $g = g_0 + \cdots + g_n$, then $f^*g = f^*g_0 + \cdots + f^*g_n$, 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, $ad - bc = 0$. □

Proposition 3.3. *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

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

where the last isomorphism follows because $V(f, \mathfrak{m}^n) = \{p\}$, as we proved in the last chapter. Thus we need only calculate the dimension of the last ring. If $m = m_p(C)$, then for any $g \in \mathfrak{m}^{n-m}$, $fg \in \mathfrak{m}^n$. 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 a ring homomorphism. 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

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

and this was the required result. □

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 the dimension of the space is two. In particular, this implies that if $\mathcal{O}$ is a discrete valuation domain, then $\mathfrak{m}^n/\mathfrak{m}^{n+1} = (t^n)/(t^{n+1})$, which is isomorphic to K, and is therefore one dimensional. This implies, and so $m_p(C) = 1$. This completes the proof that $\mathcal{O}$ is a discrete valuation domain if and only if p is a nonsingular point.

Example. *This property does not hold for curves in higher dimensional space. Let us look at the variety 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 $(X^2 - Y^3, Y^2 - Z^3)$, hence the quotient is an integral domain, and so the ideal is prime. In the local ring at the origin, we find that $\mathfrak{m}/\mathfrak{m}^2$ has a basis X, Y, and Z, because polynomials not vanishing at the origin have inverses specified by a power series which is finite since all degree two polynomials vanish, so the denominators in this quotient ring vanish. Thus the quotient has dimension 3.*

We also 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 $K[X_1,\ldots,X_m]$, 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 $K[X_1,\ldots,X_n]/\mathfrak{m}^n$ (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*

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

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 $f \in K[X_1,\dots,X_m]$, and suppose we can write $f = f_n + f_{n+1} + \dots$, in which case we generalize the notion of multiplicity, and say that f has multiplicity n at the origin, denoted $m_p(f)$. In this case, we find that $\chi(k)$ is a polynomial of degree $m-1$ for sufficiently large m, with leading coefficient $m_p(f)/(m-1)!$, so that the local ring $\mathcal{O}_p(f)$ can also calculate multiplicities on hypersurfaces. To see this, we first apply the isomorphism of $\mathcal{O}_p(f)/\mathfrak{m}^k$ with $K[X_1,\dots,X_m]/(\mathfrak{m}^k, f)$. Then, if $k \geqslant n$, we can consider the exact sequence*

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

Hence the dimension of the right hand ring, which is $\chi(k)$, if we write $\chi_m(k) = k^m/m! + ak^{m-1} + O(k^{m-2})$, is

$$\begin{aligned}\dim K[X_1,\dots,X_m]/\mathfrak{m}^k &- K[X_1,\dots,X_m]/\mathfrak{m}^{k-n} \\ &= \chi_m(k) - \chi_m(k-n) \\ &= k^m/m! - (k-n)^m/m! + ak^{m-1} - a(k-n)^{m-1} + O(k^{m-2}) \\ &= \frac{nk^{m-1}}{(m-1)!} + O(k^{m-2})\end{aligned}$$

and so the local ring determines the multiplicity of the hypersurface.

3.2 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 $I_p(fg, h) = I_p(f, h) + I_p(g, h)$.

- $I_p(C_0, C_1) = 0$ 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 consistant, 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 occuring 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 $K[C_1]$.

- $I_p(f, g) = I_p(f, g + kf)$ for any $k \in K[X_1, \ldots, X_n]$.

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

Theorem 3.4. *$I_p(f, g)$ 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 $I_p(f, g)$. If $I_p(f, g) = 0$, we can calculate this value just from the fact that this implies that p does not lie on $V(f) \cap V(g)$. Assume that if $I_p(f, g) < n$, we can calculate this value just from the properties of n. Suppose that $f(X, 0)$ and $g(X, 0)$ are of degree r

and s, which are equal to zero if the polynomial vanishes completely on the X axis. We may assume $r \leqslant s$. If $r = 0$, then $f(X,0) = a$ for some a, so $f(X,Y) = Yf_0(X,Y)$ (for it cannot be constant), and $I_p(f,g) = I_p(Y,g) + I_p(f_0,g)$. Now $0 < I_p(Y,g), I_p(f_0,g) < I_p(f,g)$, since p lies on the curve defined by Y, so by induction we can calculate these values uniquely from the properties above. In general, if $r > 0$, then we may assume that, via multiplying by scalars, $f(X,0)$ and $g(X,0)$ are monic. Let $h = g - X^{s-r}f$. Then $I_p(f,g) = I_p(f,g - X^{s-r}h)$, and the degree of $h(X,0)$ is smaller than the degree of g. Repeating this process, we decrease the value of one of the polynomials under consideration under we obtain that $r = 0$, in which case we can compute the values required. Thus we have an algorithm for computing the intersection number. □

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 propertial 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 $\mathcal{O}_p(\mathbf{A}^2)/(f,g)$ over K, or the localization of the ring $K[X,Y]/(f,g)$ at p.

Theorem 3.5. *$I_p(f,g) = \dim \mathcal{O}_p(\mathbf{A}^2)/(f,g)$ satisfies the required properties of an intersection number.*

Proof. We shall write $\mathcal{O}$ for $\mathcal{O}_p(\mathbf{A}^2)$. It is clear that $I_p(f,g) = I_p(g,f)$, since $(f,g) = (g,f)$, as well as that $I_p(f,g) = I_p(f,g+kf)$, because $(f,g) = (f,g+kf)$. Similarily, since the local ring $\mathcal{O}$ is preserved under isomorphisms, the intersection number is preserved under changes in affine coordinates. If p isn't in $V(f) \cap V(g)$, then (f,g) contains a unit in $\mathcal{O}$, and consequently, $\mathcal{O}/(f,g) = (0)$, and so this space has dimension zero, so $I_p(f,g) = 0$. Conversely, the space $\mathcal{O}_p(\mathbf{A}^2)/(f,g)$ can only have dimension zero if (f,g) contains a unit, which implies that either $f(p) \neq 0$ or $g(p) \neq 0$. If f and g contain some common irreducible component h which vanishes at p, and $(f,g) \subset (h)$, so we have a surjective map from $\mathcal{O}/(f,g)$ to $\mathcal{O}/(h)$. Since $\mathcal{O}/(h)$ is the localization of $K[X,Y]/(h)$ at p, which

is an infinite dimensional integral domain because $V(h)$ is infinite, we conclude that $\mathcal{O}/(h)$ is infinite dimensional. Conversely, we know that if f and g contain no irreducible factors which vanish at p, then as an ideal in $\mathcal{O}$, $(f,g) = (f',g')$, where f' and g' are relatively prime polynomials. Therefore we may assume f and g have no common factors from the beginning. This means that $V(f,g)$ contains only finitely many points $p_1,\ldots,p_m$, so $K[X,Y]/(f,g)$ is isomorphic to $\bigoplus \mathcal{O}_{p_i}(\mathbf{A}^n)/(f,g)$, and in particular $\dim \mathcal{O}/(f,g) \leqslant \dim K[X,Y]/(f,g)$, and we have proven that $K[X,Y]/(f,g)$ is of dimension $\leqslant m$. To prove that the intersection number is multiplicatively additive, it suffices to prove that we can find an exact sequence

$$0 \to \mathcal{O}_p(\mathbf{A}^2)/(f,h) \to \mathcal{O}_p(\mathbf{A}^2)/(fg,h) \to \mathcal{O}_p(\mathbf{A}^2)/(g,h) \to 0$$

We can take the first map as the one induced by mapping k to kg, because it maps $(f,h) = (fk,hk) \subset (fg,h)$. Conversely, if $kg = k_0fg + k_1h$, then $(k - k_0f)g = k_1h$. Since h does not divide g, we find that $k - k_0f \in (h)$, hence $k \in (h,f)$. Thus the map is injective. We take the second map as the quotient map, which is surjective. The kernel of this map is exactly $(g,h)/(fg,h) = (g)/(fg,h)$, which is exactly the image of the first map. To prove the multiplicity lower bound, we assume that $p = 0$. Let $m = m_p(f)$, and $n = m_p(g)$. Let $\mathfrak{a}$ denote the ideal $\mathfrak{m}_p(f)$. Consider the diagram

$$\begin{array}{ccc}
K[X,Y]/\mathfrak{a}^n \times K[X,Y]/\mathfrak{a}^m & & \\
\downarrow & & \\
K[X,Y]/\mathfrak{a}^{n+m} & & \mathcal{O}/(f,g) \\
\downarrow & & \downarrow \\
K[X,Y]/(\mathfrak{a}^{n+m},f,g) & \longrightarrow & \mathcal{O}/(\mathfrak{a}^{n+m},f,g) \\
\downarrow & & \downarrow \\
0 & & 0
\end{array}$$

Where the map from the direct product is given by $(k_0,k_1) \mapsto k_0f + k_1g$, 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 $V(\mathfrak{a}^{n+m},f,g) = \{p\}$. 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

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

We only have equality here only 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 3.6. *If $L_1, L_2, \ldots$ and $M_1, M_2, \ldots$ 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 $A_{ij} = L_1 \ldots L_i M_1 \ldots M_j$, 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 $\lambda_0 M_1 \ldots M_n = -L_1(\sum \lambda_i L_2 \ldots L_i M_1 \ldots M_{n-i})$. 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 $\sum \lambda_i L_2 \ldots L_i M_1 \ldots M_{n-i}$. Since $L_2, L_3, \ldots$ and $M_1, M_2, \ldots$ 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 $\lambda_2 = \cdots = \lambda_n = 0$. □

Lemma 3.7. *If f and g have no common tangents at p, then $\mathfrak{a}^{m+n-1} \subset (f,g)$, where (f,g) is the ideal in $\mathcal{O}$.*

Proof. Let $L_1, \ldots, L_m$ be the tangents to f at p, and $M_1, \ldots, M_n$ the tangents to g at p. If we define $A_{ij} = L_1 \ldots L_i M_1 \ldots M_j$, 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 $V(f,g) = \{p, q_1, \ldots, q_s\}$. We may then choose a polynomial h which vanishes on all q_i, but with $h(p) = 1$. Then $hX, hY \in I(V(f,g))$, 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 3.8. *The map $(A,B) \mapsto Af + Bg$ is injective if and only if f and g have distinct tangents at p.*

Proof. Suppose that f and g have distinct tangents, and $Af + Bg \in \mathfrak{a}^{n+m}$. We must conclude that $A \in \mathfrak{a}^n$ and $B \in \mathfrak{a}^m$, so assume otherwise. That is, take A to have multiplicity r, and B to have muliplicity s, and assume $r < n$ or $s < m$. Then $Af + Bg = A_r f_m + B_s g_n +$ higher terms. We must have $r+m = s+n$, and that these terms cancel out, or else $Af + Bg \notin \mathfrak{a}^{n+m}$. Thus $A_r f_m = -B_s g_n$. Since f and g have no common tangent, f_m and g_n have no common factors, so g_n divides A_r and f_m divides B_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*

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

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

$$\begin{aligned} g_0 &= (X^2+Y^2)Y^3 - 4X^2Y^2 - 3(X^2+Y^2)X^2Y \\ &= Y[(X^2+Y^2)(Y^2-3X^2) - 4X^2Y] = Yg_1. \end{aligned}$$

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

$$\begin{aligned} g_2 = g_1 + 3f &= 4Y^2(X^2+Y^2) + 5X^2Y - 3Y^3 \\ &= Y(4Y(X^2+Y^2) + 5X^2 - 3Y^2) = Yg_3 \end{aligned}$$

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 $f(X,Y) = Y^2 + aX^3 + bX^2Y + cXY^2 + dY^3 + \ldots$, then $f_{XXX}(0) = 6a$, and $I_p(f,Y) = 3$ only when Y does not divide $aX^3 + bX^2Y + cXY^2 + dY^3$, 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 3.9. *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, $\mathrm{ord}_p(g)$ 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 3.10. *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 $V(f,g)$ 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 $V(f,g)$. 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 $L_1, \ldots, L_m$, 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 $g_1, \ldots, g_m \in K[X,Y]$ 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.

Chapter 4

Projective Varieties

The first big result of the theory of projective curves is that two planar curves of degree n and m interect in exactly nm locations. However, there are numerous caveats to this statement. The simplest form of this statement is that two lines in the plane intersect in a unique position. This is clearly not true in the affine plane, because two parallel lines need not intersect. Thus we must expand our base space from affine space to projective space, and provided we are working over an algebraically closed field, this is the context in which Bezout's theorem holds.

To obtain a form of algebraic geometry over projective space, we must consider a coordinate system on the space,. The natural choice for coordinates on projective space are the homogenous coordinates obtained by identifying $\mathbf{P}^n$ with the set of lines through the origin in $\mathbf{A}^{n+1}$. The natural embedding of $\mathbf{A}^n$ in $\mathbf{P}^n$ is obtained by appending a coordinate 1, so that $(x_1,\ldots,x_n) \mapsto [x_1 : \cdots : x_n : 1]$. We would like to consider the zero sets of polynomials $f \in K[X_1,\ldots,X_{n+1}]$ 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*, then $f(\lambda x) = \lambda^k f(x)$, where k is the degree of f, and in particular this implies that even if f does not descend to a map on projective space, the zero set of f is a union of lines through the origin, and thus the zero set of f can be considered as a subset of projective space. Given a set S of homogenous polynomials, we let $V(S)$ denote the projective variety which is the locus of the points S, and we call these sets **projective varieties**.

Example. *The affine lines defined by equations of the form $Y = mX + b$ in $\mathbf{P}^2$*

can be embedded in the projective lines described by the homogenous equation $Y = mX + bZ$. This projective line has 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 is the most natural way to embed affine varieties in projective varieties.

Example. *Consider the affine hyperbola $Y^2 = X^2 + 1$. This 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 $Y = -X$.*

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 that 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 $K[X_1, \ldots, X_{n+1}]$ is not emphasized. We say an ideal $\mathfrak{a}$ in $K[X_1, \ldots, X_n]$ is a **homogenous ideal** if, whenever $f \in \mathfrak{a}$, then $f_0, f_1, \ldots, f_m \in \mathfrak{a}$, so the ideal is closed under the projections onto the homogenous parts of each polynomial.

Example. *If f is a homogenous polynomial, then (f) is a homogenous ideal, since if f is degree m, then $(gf)_i = g_{i-m} f$ if $m \leqslant i$, or $(gf)_i = 0$. More generally, $(f_1, \ldots, f_n)$ 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

$$V(\mathfrak{a}) = \{x \in \mathbf{P}^n : (\forall \text{ homogenous } f \in \mathfrak{a}) : f(x) = 0\}$$

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 $\mathfrak{a} = (f_1, \ldots, f_m)$, then $V(\mathfrak{a}) = \bigcap V(f_i)$, 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, the ideal $I(V)$ of polynomials f in $K[X_1, \ldots, X_{n+1}]$ such that f_i vanishes on V for each i is a homogenous ideal, and is the projective version of affine variety. This ideal is also a radical ideal, because

$$f_m^n = \sum f_1^{j_1} \ldots f_m^{j_m}$$

where j_i range over all choices of indices with $\sum j_k = n$, $\sum k j_k = m$. In particular, this means that $f_0^n = (f_0)^n$ vanishes on V, hence f_0 vanishes on V. But this means that $f_n^n = (f_1)^n + f_0 g$, for some polynomial g, and since f_0 vanishes on V, we conclude that f_1^n vanishes on V, and therefore f_1 vanishes on V. By induction, assuming that f_j vanishes on V for each $j < k$, we conclude that $f_{kn}^n = (f_k)^n + \sum_{j<i} f_k g_j$ for some polynomials g_j, because if $\sum j_i = n$, $\sum i j_i = kn$, and $j_k = 0$ for $j < k$, and $j_k \neq n$, then $\sum i j_i > k \sum j_i = kn$, which means the index is not in the sum. We conclude that $(f_k)^n$ vanishes on V, and thus f_k vanishes on V. Thus $I(V)$ is a radical ideal. 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 a homogenous ideal, and $V(\bigoplus \mathfrak{a}_\alpha) = \bigcap V(\mathfrak{a}_\alpha)$. 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 $V(\mathfrak{a}) \cup V(\mathfrak{b}) = V(\mathfrak{a}\mathfrak{b})$. Thus finite unions of projective varieties are projective varieties.

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

  $$V(X_1 - (a_1/a_{n+1})X_{n+1}, \ldots, X_n - (a_n/a_{n+1})X_{n+1}) = \{[a_1 : \cdots : a_{n+1}]\}$$

  and also for $a \in K^n$ with $a_i \neq 0$

  $$V(X_1 - (a_1/a_i)X_i, \ldots, X_n - (a_n/a_i)X_i, X_{n+1}) = \{[a_1 : \cdots : a_n : 0]\}$$

  It follows that finite point sets are projective varieties.

- $I(\varnothing) = K[X_1, \ldots, X_{n+1}]$, and $I(\mathbf{A}^n) = (0)$.
- For any set S of homogenous polynomials, $S \subset I(V(S))$, and $X \subset V(I(X))$ for any set X. This implies $V(I(V(S)) = V(S)$ 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 4.1. *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, $V(f, I(V))$ and $V(g, I(V))$ are both proper subsets of $I(V)$, whose union is $V(fg, I(V)) = V(I(V)) = V$. 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 homogeonus 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 4.2. *Over an algebraically closed field, if $\mathfrak{a}$ is a homogenous ideal, then $V(\mathfrak{a}) = \varnothing$ if and only if there is some n such that $\mathfrak{a}$ contains all forms of degree $\geqslant n$, and if $V(\mathfrak{a}) \neq \varnothing$, then $I(V(\mathfrak{a}))$ is the radical ideal generated by $\mathfrak{a}$.*

Proof. If the projective variety $V(\mathfrak{a}) = \varnothing$, then over affine space, $V(\mathfrak{a}) = \varnothing$ or $V(\mathfrak{a}) = \{0\}$. In the first case, we can apply the Nullstellensatz to conclude that $\mathfrak{a} = K[X_1, \ldots, X_{n+1}]$. 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 4.3. *There is a one to one correspondence between projective hyperplanes and forms f containing no repeated factors. Irreducible hypersurfaces correspond to irreducible forms.*

Given a projective variety V, we can form the **homogenous coordinate ring** of the variety by taking $K_h[V] = K[X_1, \ldots, X_{n+1}]/I(V)$. Unlike the coordinate ring on an affine variety, these elements *cannot* be interpreted as functions on V. If $V \neq \varnothing$, we can view them as functions on $C(V)$, because $I(V) = I(C(V))$, but only the scale invariant functions can be viewed as functions on V. We call a residue of $K_h[V]$ a form if it contains a form.

Proposition 4.4. *Every element $f \in K_h[V]$ can be uniquely expanded as $\sum f_i$, where f_i is a form in $K_h[V]$.*

Proof. Consider the residue of $f \in K[X_1, \ldots, X_{n+1}]$ in $K_h[V]$, and write $f = \sum f_i$. If $f = g$ in $K[X_1, \ldots, X_{n+1}]$, then $f - g \in I(V)$, and so $f_i - g_i \in I(V)$ for each i, hence $f_i = g_i$, 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** $K_h(V)$. Again, there is no natural way to think of these functions as V, except in the case where we consider a fraction f/g, where f and g are both forms of the same degree, for if $f(\lambda x) =$

$\lambda^k x$, and $g(\lambda x) = \lambda^k x$, then $(f/g)(\lambda x) = (\lambda^k/\lambda^k)(f/g)(x) = (f/g)(x)$, so the functions is only defined up to lines through the origin. The set of all such functions forms 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, 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 $g(p) \neq 0$. Alternatively, if V is not an irreducible variety, we can localize $K_h[V]$ by the multiplicative subset of forms on V 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, and this will form the local ring $\mathcal{O}_p(V)$. It has a unique maximal ideal $\mathfrak{m}_p(V)$ of functions which vanish at p, so the ring is local.

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

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

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

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

Because if $\deg(f) = m$, then

$$\frac{a_0 + a_1t + \dots + a_mt^m}{b_0 + b_1t + \dots + b_nt^n} = \frac{a_0t^{-n} + a_1t^{1-n} + \dots + a_mt^{m-n}}{b_0t^{-n} + b_1t^{1-n} + \dots + b_n}$$

and though t is not defined at $[0:1]$, t^{-1} is defined at $[0:1]$, so provided that $m-n \leqslant 0$ this function is well defined at the point. The local rings describe the set of all local rings whose quotient field is $K(t)$.

The analogy of an affine transformation in $\mathbf{A}^n$ is a projective transformation in $\mathbf{P}^n$. Projective transformations are the natural isomorphisms of $\mathbf{P}^n$ in projective geometry, and also in the study of projective varieties. Such a map induces an isomorphism of the coordinate ring, and the local rings at points, in the same way that affine transformations induce isomorphisms on affine varieties.

4.1 Affine and Projective Varieties

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 $K[X_1,\ldots,X_n]$, we let $\mathfrak{a}^*$ denote the homogenous ideal generated by f^*, for $f \in \mathfrak{a}$. Conversely, if $\mathfrak{a}$ is an ideal in $K[X_1,\ldots,X_{n+1}]$,, 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 $(\mathfrak{a}_*)^* = K[X_1,\ldots,X_{n+1}]$. 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 4.5. *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 $I(V_*)^* \subset I(V)$. □

If $f \in K_h[V^*]$, we can define $f_* \in K[V]$, because the dehomogenization process is independent of the choice of f, and if $f \in K[V]$, then $f^* \in K_h[V^*]$ is also well defined because V^* is defined such that $I(V)^* \subset I(V^*)$. This gives a homomorphism from $K_h[V]$ to $K[V]$. Since $K_h[V^*]$ is just $K[C(V^*)]$, one can also obtain this morphism from the embedding

$x \mapsto (x,1)$. This map decends to a map from $\mathcal{O}_{(p,1)}(CV^*)$ to $\mathcal{O}_p(V)$, obtained by mapping f/g to f_*/g_*. If $f_*/g_* = 0$ in $\mathcal{O}_p(V)$, then there is h with $h(p) \neq 0$ with $f_* = g_* h$, hence $(f_*)^* = (g_*)^* h^*$, and $h^*(p,1) = h(p) \neq 0$, so $(f_*)^*/(g_*)^* = 0$, and since $(f_*)^*/(g_*)^*$ differs from f/g by a unit (a power of X_{n+1}), we conclude that $f/g = 0$, hence the map from $\mathcal{O}_p(V^*)$ to $\mathcal{O}_p(V)$ is injective. Given any $f/g \in \mathcal{O}_p(V)$, with $\deg f + k = \deg g$, then $X_{n+1}^k f^*/g^* \in \mathcal{O}_p(V^*)$ maps to f/g, so the map is an isomorphism. This makes sense, because the finite points of V^* are exactly the finite points of V, and the infinite points of V^* should not affect the local properties of finite points. 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, and then to reduce the analysis to an affine variety as above.

Chapter 5

Projective Planar Curves

We now unite our theory of projective varieties with our theory of affine plane curves, giving us the most powerful results in the theory of curves. 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. The degree of a curve is the degree of the homogenous polynomial that defines it. This is an invariant of projective coordinates, by essentially the same proof as in the affine case.

Given a point $p \in \mathbf{P}^2$, we define the multiplicity $m_p(f)$ of a curve defined by f to be the eventual constant value of the dimension of $\mathfrak{m}_p(f)^n/\mathfrak{m}_p(f)^{n+1}$ in $\mathcal{O}_p(f)$, a purely algebraic quantity. The 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 $m_p(f) = m_p(f_*)$. To define the intersection multiplicities of two projective curves at a finite point p, we let $I_p(f,g) = I_p(f_*,g_*) = \dim \mathcal{O}_p(\mathbf{A}^2)/(f_*,g_*)$. 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 $\deg g - \deg f$.

Example. *The polynomial $XY^4 + YZ^4 + XZ^4$, 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 tangets 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.*

Example. *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. *A point p is a multiple point of a polynomial f if and only if $f(p) = 0$ and $f_X(p) = f_Y(p) = f_Z(p) = 0$. The way to see this is a projective change of coordinates acts as a linear transformation on the derivatives, so we may assume that $p = [0 : 0 : 1]$, in which case $f(p) = 0$ implies that $f(X, Y, Z) = aX + bY + f_0(X, Y, Z)$, where f_0 only has terms of degree two or higher. It then follows that $f_X(p) = f_Y(p) = f_Z(p) = 0$ if and only if $a = b = 0$, and this holds if and only if $m_p(f) = m_p(f_*) > 1$.*

Example. *The above points provide us with a simple way to classify the irreducible projective conics up to equivalence. Let f be a projective irreducible*

conic with a simple point at $[0:1:0]$ *with tangent* $Z=0$ *there. Then we can write* $f = YZ - aX^2 - bXZ - cZ^2$. *Since* f *is irreducible, we cannot have* $a = 0$, *else* Z *divides* f. *Dehomogenizing with respect to* Z, *we find that the equation is* $Y = aX^2 + bX + c$, *which describes a parabola. All parabolas are affinely equivalent, because if we complete the square, replacing* X *with* $a^{-1/2}(X - b/2)$, *we obtain the equation* $Y = X^2 + b^2/4 + c$, *and by replacing* Y *with* $Y - b^2/4 - c$, *we obtain* $Y = X^2$. *Thus all irreducible projective conics are equivalent to the standard parabola.*

Example. *There is only a single projective cubic with a cusp, up to projective equivalence. Without loss of generality, let the cusp occur at the origin, with* $Y = 0$ *as the origin. It is easy to see that the equation for such a curve must be of the form*

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

Scaling X, *we can set* $a = 1$, *and applying a Tschirnhouse transformation to* X, *we can also set* $b = 0$. *Finally, translating* Z *to* $Z + cX + dY$, *we conclude the cubic is equivalent to the cubic* $Y^2Z = X^3$, *so there is a unique cubic with a cusp.*

Example. *There is a single projective cubic 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* $ZXY = X^3 + Y^3$. *This describes the folium of Descartes. This equation has no other singularities.*

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 $\operatorname{ord}_p(g) = \operatorname{ord}_p(g/h)$, where h is a form of the same degree as g, with $h(p) \neq 0$. This is equivalent to letting $\operatorname{ord}_p(g) = \operatorname{ord}_p(g_*)$, as we just defined g_* in the last paragraph.

Given two projective plane curves f and g, we define the **intersection number** $I_p(f,g)$ to be the dimension of $\mathcal{O}_p(\mathbf{P}^2)/(f_*, g_*)$. This definition satisfies all the properties of the intersection numbers we considered previously, except that it is invariant not only under affine transformations, but also under projective transformations, and that $I_p(f,g) = I_p(f, g + kf)$ only when 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.

5.1 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 has dimension $\sum_{i=0}^{n}(i+1) = (n+1)(n+2)/2$. However, we identify homogenous polynomials which differ by a scalar factor, so the space of planar curves of degree n is naturally identified with $\mathbf{P}^{n(n+3)/2}$. This is an incredibly basic example of a moduli space, a geometric space formed from a class of objects. An easy first result along these lines is that if T is a projective change of coordinates, then T^* is a projective change of coordinates in $\mathbf{P}^{n(n+3)/2}$.

Example. *The duality theory of projective geometry implies that the space of lines forms a projective space in of itself, and in particular, if we set $n = 1$, we find the space of projective lines in the plane can be identified with the projective plane $\mathbf{P}^2$.*

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 $\mathbf{P}^{14}$.*

Under this identification, we obtain a certain duality theory, because we can consider certain subsets of planar curves as projective varieties. If this subset forms a linear subvariety, we call it a **linear system of curves**.

Example. *For any point p, the space of degree n curves through $p = [x : y : z]$ forms a hyperplane in $\mathbf{P}^{n(n+3)/2}$, because it is specified by the single homogenous equation $\sum a_{ijk}x^i y^j z^k = 0$. In particular, the set of degree n curves passing through m points $p_1, \ldots, p_m$ is a linear system of curves. Since the intersection of n hyperplanes on $\mathbf{P}^n$ is nonempty, there is a curve of degree n passing through any given set of $n(n+3)/2$ points.*

Example. *The set of curves f of degree n such that $m_p(f) \geqslant m$ form a linear subvariety of dimension $n(n+3)/2 - m(m+1)/2$. 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 $f_0 = f_1 = \cdots = f_{m-1} = 0$, and this gives the dimension above.*

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 $p_1, \ldots, p_4$ 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.*

Given a finite set $p_1, \ldots, p_m$ of points on $\mathbf{P}^2$, and nonnegative integers $n_1, \ldots, n_m$, set $V(d; p_1, n_1, \ldots, p_m, n_m)$ to be the set of curves of degree d which have multiplicity greater than n_i at each point p_i. It is a linear subvariety, with dimension lower bounded by $d(d+3)/2 - \sum n_i(n_i+1)/2$, because we have specified the variety by $\sum n_i(n_i+1)/2$ (some possibly redundant) constraints. If $d \geqslant \sum n_i - 1$, then we actually have an exact equality

$$\dim V(d; p_1, n_1, \ldots, p_m, n_m) = \frac{d(d+3)}{2} - \sum \frac{n_i(n_i+1)}{2}$$

We prove this result by induction. First, suppose that each $n_i = 1$. Let $V_i = V(d; p_1, \ldots, p_i)$. 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 $f = L_1 \ldots L_{n-1} L_0^{d-n+1}$ is in V_{n-1}, but not in V_n. Next, let $n_i > 1$, and for simplicity in notation let $i = 1$. Let $V_0 = V(d; p_1, n_1 - 1, p_2, n_2, \ldots, p_m, n_m)$. For $f \in V_0$ let

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

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 $W_0 = V(d-1; p_1, n_1 - 2, p_2, n_2, \ldots, p_m, n_m)$, and

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

By induction,

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

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

5.2 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 5.1. *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. It then follows that

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

The theorem will be proved if we can show that $\dim K[X,Y]/(f_*, g_*)$ has the same dimension as some subspace of forms of a fixed degree in the space $K[X,Y,Z]/(f,g)$, and that this space of forms has dimension nm. We shall start with this last part. Let $A = K[X,Y,Z]$, let A_k denote the space 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

$$0 \to A \to A^2 \to A \to B \to 0$$

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

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

Thus this space has the right dimension for large enough d.

Next, we prove the endomorphism on B given by mapping h to Zh is injective. It suffices to show that if $Zh = Af + Bg$, then $h = A'f + B'g$ for some A', B'. Temporarily denote the polynomial $k(X, Y, 0)$ by $k_0(X, Y)$ for any polynomial k. 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 $A_0 f_0 = -B_0 g_0$, so $B_0 = f_0 C$ and $A_0 = -g_0 C$ 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 $A_1, \ldots, A_{mn} \in A_d$ 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 $Z^r A_1, \ldots, Z^r A_{mn}$ 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 $Z^t h^*$ is a form of degree $d + r$, so $Z^t h^* = \sum_{i=1}^{mn} \lambda_i Z^r A_i + Bf + Cg$ for some $\lambda_i \in K$, but then $h = (Z^t h^*)_* = \sum_{i=1}^{mn} \lambda_i (A_i)_* + Bf_* + Cg_*$. 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 $Z^t A_i$ form a basis in B_{d+r}, so the λ_i are identically zero. This finishes the proof. □

Corollary 5.2. *If f and g meet in mn distinct points, then these points are all simple on both f and g.*

Corollary 5.3. *If f and g meet in more than mn points, then they share a common component.*

This theorem can be used in so many different geometric applications, that we will dwell on for the remainder of the chapter. For instance, it

is easy to prove on this that a nonsingular projective planar curve is irreducible, because if $f = gh$, then g and h meet in $(\deg g)(\deg h) \geqslant 0$ places, and these points will have multiplicity greater than one.

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 $3(n-2)$. This polynomial suffices to determine the flex points on a projective curve.

Theorem 5.4. *Over a field of characteristic zero, the points on a planar curve f for which the Hessian vanishes are precisely the multiple points and flex points.*

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, we find that

$$2f_{X_j} = Xf_{XX_j} + Yf_{YX_j} + Zf_{ZX_j}$$

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

$$\begin{aligned} h(X,Y,Z) &= \det \begin{pmatrix} f_{XX} & f_{XY} & f_{XZ} \\ f_{YX} & f_{YY} & f_{YZ} \\ Xf_{XX} + Yf_{YX} + f_{XZ} & Xf_{XY} + Yf_{YY} + f_{ZY} & Xf_{XZ} + Yf_{YZ} + f_{ZZ} \end{pmatrix} \\ &= \det \begin{pmatrix} f_{XX} & f_{XY} & f_{XZ} \\ f_{YX} & f_{YY} & f_{YZ} \\ 2f_X + (1-Z)f_{XZ} & 2f_Y + (1-Z)f_{YZ} + f_{ZY} & f_Z + (1-Z)f_{ZZ} \end{pmatrix} \end{aligned}$$

In particular, modulo scalars

$$h_0 \equiv \det \begin{pmatrix} f_{XX} & f_{XY} & f_{XZ} \\ f_{YX} & f_{YY} & f_{YZ} \\ f_X & f_Y & f_Z \end{pmatrix}$$

Next, adding X times the first column and Y times the second column to

the third column, we find that

$$\det\begin{pmatrix} f_{XX} & f_{XY} & f_{XZ} \\ f_{YX} & f_{YY} & f_{YZ} \\ f_X & f_Y & f_Z \end{pmatrix} = \det\begin{pmatrix} f_{XX} & f_{XY} & Xf_{XX}+Yf_{YX}+f_{XZ} \\ f_{YX} & f_{YY} & Xf_{XY}+Yf_{YY}+f_{ZY} \\ f_X & f_Y & Xf_X+Yf_Y+f_Z \end{pmatrix}$$
$$= \det\begin{pmatrix} f_{XX} & f_{XY} & 2f_X+(1-Z)f_{XZ} \\ f_{YX} & f_{YY} & 2f_Y+(1-Z)f_{YZ} \\ f_X & f_Y & 2f+(1-Z)f_Z \end{pmatrix}$$

So, working modulo scalars again, we find

$$h_0 \equiv \det\begin{pmatrix} f_{XX} & f_{XY} & f_X \\ f_{XY} & f_{YY} & f_Y \\ f_X & f_Y & f \end{pmatrix} = f f_{XX} f_{YY} - f_{XX} f_Y^2 - f f_{XY}^2 + 2 f_X f_Y f_{XY} - f_{YY} f_X^2$$

In particular, that working modulo (f), we find that

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

If p is a multiple point on f, then p is a multiple point on g, because if $f_X(p) = f_Y(p) = 0$, then using the chain rule we find $g_X(p) = 0$ and $g_Y(p) = 0$, and obviously $g(p) = 0$, so $m_p(g) > 1$. 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, and if we write $f_0(X,Y) = Y + aX^2 + bXY + cY^2 + dX^3 + eX^2Y + \ldots$, then we know f is a flex 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,

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

And so $g(p) = 0$ 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 5.5. *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 f and h have

intersections summing up to multiplicity $3n(n-2) > 0$. In particular, this implies f and h intersect, and since f has no multiple points, this point must be a flex by the theorem above. 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. □

5.3 Multiple Point Combinatorics

Bezóut's theorem has a simple corollary. If we sum the $m_p(f)m_p(g) \leqslant I_p(f,g)$ over all points p, we obtain that

$$\sum m_p(f)m_p(g) \leqslant \sum_p I_p(f,g) \leqslant \deg(f)\deg(g)$$

This means that if f has degree n, $\sum m_p(f)m_p(f_X) \leqslant n(n-1)$, and since any multiple point p for f has $m_p(f_X) = 1$, if we have k multiple points, then $2k \leqslant \sum m_p(f)m_p(f_X) \leqslant n(n-1)$. We conclude that a curve of degree n can have at most $n(n-1)/2$ multiple points. For irreducible curves, this can be improved fairly significantly.

Theorem 5.6. *If f is irreducible of degree n, then*

$$\sum m_p(f)[m_p(f)-1] \leqslant (n-1)(n-2)$$

Proof. Define

$$\begin{aligned} r &= \frac{(n-1)(n+2)}{2} - \sum \frac{m_p(f)[m_p(f)-1]}{2} \\ &\geqslant \frac{n(n-1)}{2} - \sum \frac{m_p(f)[m_p(f)-1]}{2} \geqslant 0 \end{aligned}$$

Then since r is non-negative, we may choose r simple points $q_1, \ldots, q_r$ on f. What's more, r is chosen such that

$$\frac{(n-1)(n+2)}{2} - \sum \frac{m_p(f)[m_p(f)-1]}{2} - r = 0$$

The maximal integer for which this equation is non-negative. It follows from our discussion on linear systems of curves that we can choose a curve g of degree $n-1$ passing through the points p_i, such that $m_p(g) \geqslant m_p(f)-1$ for all mutiple point p on f, and thus this equation holds for all points p since the equation is trivial if p is not a multiple point. If we apply Beźout's theorem to f and g, since f is irreducible and therefore cannot share any component with g, we conclude that

$$n(n-1) \geqslant \sum m_p(f)m_p(g) \geqslant \sum m_p(f)[m_p(f)-1] + r$$

But then, rearranging, we find

$$\sum \frac{m_p(f)[m_p(f)-1]}{2} \leqslant n(n-1) - \frac{(n-1)(n+2)}{2} = \frac{(n-1)(n-2)}{2}$$

and this gives the required bound. This inequality is tight, because the curve $X^n + Y^{n-1}Z$ has a point of multiplicity $n-1$ at the origin. □

This shows that all irreducible lines and conics are nonsingular. An irreducible cubic has

$$\sum m_p(f)[m_p(f)-1] \leqslant 2$$

Hence an irreducible cubic can have at most one double point, and no multiple points of higher degree. For $n = 4$, we find that the sum above is bounded by 6, so an irreducible quartic can have at most three double points, or a single triple point.

If f is a curve of degree n, but not an irreducible curve, but has no multiple components, we can still obtain an interesting bound. Write $f = f_1 \dots f_m$, with f_i of degree m_i. We have

$$\begin{aligned}\sum m_p(f)[m_p(f)-1] &= \sum_p \left(\sum_i m_p(f_i)\right)\left[\sum_i m_p(f_i) - 1\right] \\ &= \sum_p m_p(f_i)[m_p(f_i)-1] + \sum_{i\neq j} m_p(f_i)m_p(f_j)\end{aligned}$$

We can then apply the inequality we just proved, along with Bezóut's theorem on f_i and f_j to conclude that this sum is less than or equal to

$$\begin{aligned}\sum_i (m_i-1)(m_i-2) + \sum_{i\neq j} m_i m_j &= \left(\sum m_i^2\right) - 3n + 2m + \left(\sum_i m_i(n-m_i)\right) \\ &= n^2 - 3n + 2m \leqslant n^2 - n = n(n-1)\end{aligned}$$

5.4 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. Similarily, 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 $T_i[x : y : z] = a_i x + b_i y + c_i z$, then if h has degree m, then $h(X,Y) = \sum d_{ij} X^i Y^j Z^{m-i-j}$, 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 5.7. *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 $Z^t h = kf + k'g$ 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 $h = l_{k-m} f + l'_{k-n} g$. □

To use Max Noether's theorem, we require easy properties to check that imply Noether's condition. We now verify some of these criteria.

Theorem 5.8. *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

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

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 5.9. *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.

Theorem 5.10. *Let C and C' be two cubics meeting in 9 distinct points $p_1, \ldots, p_9$. 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 lie on a straight line.*

Proof. Applying Noether's theorem, since all intersections of Q and C are simple on C and Q (since all points on a conic) □

Let C be a nonsingular cubic curve. For any two points p and q on the curve, the line between p and q is not tangent to C at p nor at q, and because of this the line intersects C at a unique third position r. This follows because there is a unique line L such that $L \cdot C = p + q - r$ for some point r on the curve. Define $\varphi : C \times C \to C$ by setting $\varphi(p, q) = r$. Then φ is abelian, but has no identity. This can be fixed by fixing an arbitrary point O on the cubic, and defining addition as $p \oplus q = \varphi(O, \varphi(p, q))$. Then $\oplus$ gives C an abelian group structure, and the operation is called the **group law** on the curve. TODO: ELABORATE THIS SECTION.

Chapter 6

Where to Put It?

6.1 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 $f_1(t),\ldots,f_n(t)$ 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 $f_1(t),\ldots,f_n(t)$ 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*

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

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 paramaterization of* $V(f)$*. 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*

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

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 monimals 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 6.1. *If $f : \mathbf{A}^1 \to C$ is any nonconstant rational map, then the inverse image of any point $x = (x_1, \ldots, x_n) \in C$ 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. □

6.2 Product Varieties

If $\mathfrak{a} \subset K[X_1, \ldots, X_n]$ and $\mathfrak{b} \subset K[Y_1, \ldots, Y_m]$, then these ideals generate ideals $\mathfrak{a}'$ and $\mathfrak{b}'$ in $K[X_1, \ldots, X_n, Y_1, \ldots, Y_m]$. The intersection $V(\mathfrak{a}') \cap V(\mathfrak{b}')$ corresponds to the set of points (x, y) with $x \in V(\mathfrak{a})$ and $y \in V(\mathfrak{b})$, and we call this the **product variety** $V(\mathfrak{a}) \times V(\mathfrak{b})$.