Chapter 1
Basic Notions

1 Algebraic Curves in the Plane

Chapter 1 discusses a number of the basic ideas of algebraic geometry; this first section treats some examples to prepare the ground for these ideas.

1.1 Plane Curves

An *algebraic plane curve* is a curve consisting of the points of the plane whose coordinates x, y satisfy an equation

$$f(x, y) = 0, \tag{1.1}$$

where f is a nonconstant polynomial. Here we fix a field k and assume that the coordinates x, y of points and the coefficients of f are elements of k. We write $\mathbb{A}^2$ for the *affine plane*, the set of points (a, b) with $a, b \in k$; because the affine plane $\mathbb{A}^2$ is not the only ambient space in which algebraic curves will be considered—we will be meeting others presently—an algebraic curve as just defined is called an *affine* plane curve.

The degree of (1.1), that is, the degree of the polynomial $f(x, y)$, is also called the *degree* of the curve. A curve of degree 2 is called a *conic*, and a curve of degree 3 a *cubic*.

It is well known that the polynomial ring $k[X, Y]$ is a unique factorisation domain (UFD), that is, any polynomial f has a unique factorisation $f = f_1^{k_1} \cdots f_r^{k_r}$ (up to constant multiples) as a product of irreducible factors f_i, where the irreducible f_i are nonproportional, that is, $f_i \neq \alpha f_j$ with $\alpha \in k$ if $i \neq j$. Then the algebraic curve X given by $f = 0$ is the union of the curves X_i given by $f_i = 0$. A curve is *irreducible* if its equation is an irreducible polynomial. The decomposition $X = X_1 \cup \cdots \cup X_r$ just obtained is called a decomposition of X into irreducible components.

In certain cases, the notions just introduced turn out not to be well defined, or to differ wildly from our intuition. This is due to the specific nature of the field k in

I.R. Shafarevich, *Basic Algebraic Geometry 1*, DOI 10.1007/978-3-642-37956-7_1,

which the coordinates of points of the curve are taken. For example if $k = \mathbb{R}$ then following the above terminology we should call the point $(0, 0)$ a "curve", since it is defined by the equation $x^2 + y^2 = 0$. Moreover, this "curve" should have "degree" 2, but also any other even number, since the same point $(0, 0)$ is also defined by the equation $x^{2n} + y^{2n} = 0$. The curve is irreducible if we take its equation to be $x^2 + y^2 = 0$, but reducible if we take it to be $x^6 + y^6 = 0$.

Problems of this kind do not arise if k is an algebraically closed field. This is based on the following simple fact.

Lemma *Let k be an arbitrary field, $f \in k[x, y]$ an irreducible polynomial, and $g \in k[x, y]$ an arbitrary polynomial. If g is not divisible by f then the system of equations $f(x, y) = g(x, y) = 0$ has only a finite number of solutions.*

Proof Suppose that x appears in f with positive degree. We view f and g as elements of $k(y)[x]$, that is, as polynomials in one variable x, whose coefficients are rational functions of y. It is easy to check that f remains irreducible in this ring: if f splits as a product of factors, then after multiplying each factor by the common denominator $a(y) \in k[y]$ of its coefficients, we obtain a relation that contradicts the irreducibility of f in $k[x, y]$. For the same reason, g is not divisible by f in the new ring $k(y)[x]$. Hence there exist two polynomials $\widetilde{u}, \widetilde{v} \in k(y)[x]$ such that $f\widetilde{u} + g\widetilde{v} = 1$. Multiplying this equality through by the common denominator $a \in k[y]$ of all the coefficients of $\widetilde{u}$ and $\widetilde{v}$ gives $fu + gv = a$, where $u = a\widetilde{u}$, $v = a\widetilde{v} \in k[x, y]$, and $0 \neq a \in k[y]$. It follows that if $f(\alpha, \beta) = g(\alpha, \beta) = 0$ then $a(\beta) = 0$, that is, there are only finitely many possible values for the second coordinate β. For each such value, the first coordinate α is a root of $f(x, \beta) = 0$. The polynomial $f(x, \beta)$ is not identically 0, since otherwise $f(x, y)$ would be divisible by $y - \beta$, and hence there are also only a finite number of possibilities for α. The lemma is proved. □

An algebraically closed field k is infinite; and if f is not a constant, the curve with equation $f(x, y) = 0$ has infinitely many points. Because of this, it follows from the lemma that an irreducible polynomial $f(x, y)$ is uniquely determined, up to a constant multiple, by the curve $f(x, y) = 0$. The same holds for an arbitrary polynomial, under the assumption that its factorisation into irreducible components has no multiple factors. We can always choose the equation of a curve to be a polynomial satisfying this condition. The notion of the degree of a curve, and of an irreducible curve, is then well defined.

Another reason why algebraic geometry only makes sense on passing to an algebraically closed field arises when we consider the number of points of intersection of curves. This phenomenon is already familiar from algebra: the theorem that the number of roots of a polynomial equals its degree is only valid if we consider roots in an algebraically closed field. A generalisation of this theorem is the so-called Bézout theorem: the number of points of intersection of two distinct irreducible algebraic curves equals the product of their degrees. The lemma shows that, in any

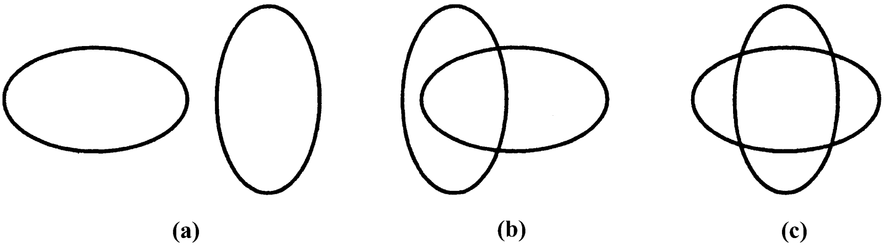


Figure 1 Intersections of conics

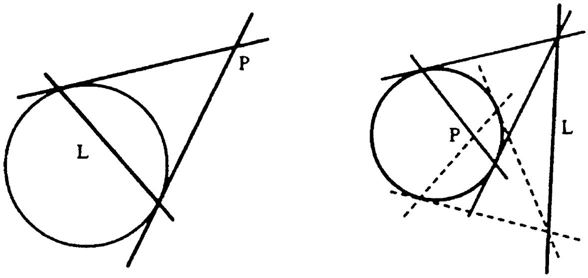
case, this number is finite. The theorem on the number of roots of a polynomial is a particular case, for the curves $y - f(x) = 0$ and $y = 0$.

Bézout's theorem holds only after certain amendments. The first of these is the requirement that we consider points with coordinates in an algebraically closed field. Thus Figure 1 shows three cases for the relative position of two curves of degree 2 (ellipses) in the real plane. Here Bézout's theorem holds in case (c), but not in cases (a) and (b).

We assume throughout what follows that k is algebraically closed; in the contrary case, we always say so. This does not mean that algebraic geometry does not apply to studying questions concerned with algebraically nonclosed fields k_0. However, applications of this kind most frequently involve passing to an algebraically closed field k containing k_0. In the case of $\mathbb{R}$, we pass to the complex number field $\mathbb{C}$. This often allows us to guess or to prove purely real relations. Here is the most elementary example of this nature. If P is a point outside a circle C then there are two tangent lines to C through P. The line joining their points of contact is called the *polar line* of P with respect to C (Figure 2, (a)). All these constructions can be expressed in terms of algebraic relations between the coordinates of P and the equation of C. Hence they are also applicable to the case that P lies inside C. Of course, the points of tangency of the lines now have complex coordinates, and can't be seen in the picture. But since the original data was real, the set of points obtained (that is, the two points of tangency) should be invariant on replacing all the numbers by their complex conjugates; that is, the two points of tangency are complex conjugates. Hence the line L joining them is real. This line is also called the polar line of P with respect to C. It is also easy to give a purely real definition of it: it is the locus of points outside the circle whose polar line passes through P (Figure 2, (b)).

Here are some other situations in which questions arise involving algebraic geometry over an algebraically nonclosed field, and whose study usually requires passing to an algebraically closed field.

(1) $\boldsymbol{k = \mathbb{Q}}$. The study of points of an algebraic curve $f(x, y) = 0$, where $f \in \mathbb{Q}[x, y]$, and the coordinates of the points are in $\mathbb{Q}$. This is one of the fundamental problems of number theory, the theory of indeterminate equations. For example, Fermat's last theorem requires us to describe points $(x, y) \in \mathbb{Q}^2$ of the curve $x^n + y^n = 1$.



(a) **(b)**

Figure 2 The polar line of a point with respect to a conic

(2) Finite fields. Let $k = \mathbb{F}_p$ be the field of residues modulo p. Studying the points with coordinates in k on the algebraic curve given by $f(x, y) = 0$ is another problem of number theory, on the solutions of the congruence $f(x, y) \equiv 0 \bmod p$.

(3) $\boldsymbol{k = \mathbb{C}(z)}$. Consider the algebraic surface in $\mathbb{A}^3$ given by $F(x, y, z) = 0$, with $F(x, y, z) \in \mathbb{C}[x, y, z]$. By putting z into the coefficients and thinking of F as a polynomial in x, y, we can consider our surface as a curve over the field $\mathbb{C}(z)$ of rational functions in z. This is an extremely fertile method in the study of algebraic surfaces.

1.2 Rational Curves

As is well known, the curve given by

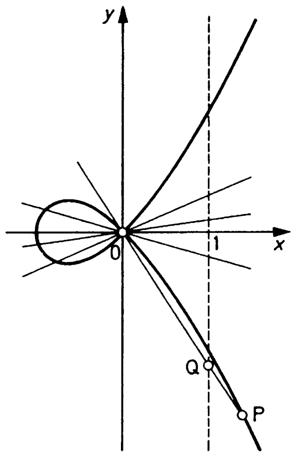
$$y^2 = x^2 + x^3 \tag{1.2}$$

has the property that the coordinates of its points can be expressed as rational functions of one parameter. To deduce these expressions, note that the line through the origin $y = tx$ intersects the curve (1.2) outside the origin in a single point. Indeed, substituting $y = tx$ in (1.2), we get $x^2(t^2 - x - 1) = 0$; the double root $x = 0$ corresponds to the origin $0 = (0, 0)$. In addition to this, we have another root $x = t^2 - 1$; the equation of the line gives $y = t(t^2 - 1)$. We thus get the required parametrisation

$$x = t^2 - 1, \qquad y = t\bigl(t^2 - 1\bigr), \tag{1.3}$$

and its geometric meaning is evident: t is the slope of the line through 0 and (x, y); and (x, y) are the coordinates of the point of intersection of the line $y = tx$ with the curve (1.2) outside 0. We can see this parametrisation even more intuitively by drawing another line, not passing through 0 (for example, the line $x = 1$) and projecting the curve from 0, by sending a point P of the curve to the point Q of intersection of the line $0P$ with this line (see Figure 3). Here the parameter t plays the role of coordinate on the given line. Either from this geometric description, or from (1.3), we see that t is uniquely determined by the point (x, y) (for $x \neq 0$).

Figure 3 Projection of a cubic



We now give a general definition of algebraic plane curves for which a representation in these terms is possible. We say that an irreducible algebraic curve X defined by $f(x,y)=0$ is *rational* if there exist two rational functions $\varphi(t)$ and $\psi(t)$, at least one nonconstant, such that

$$f\big(\varphi(t),\psi(t)\big)\equiv 0, \tag{1.4}$$

as an identity in t. Obviously if $t=t_0$ is a value of the parameter, and is not one of the finitely many values at which the denominator of φ or ψ vanishes, then $(\varphi(t_0),\psi(t_0))$ is a point of X. We will show subsequently that for a suitable choice of the parametrisation φ, ψ, the map $t_0\mapsto(\varphi(t_0),\psi(t_0))$ is a one-to-one correspondence between the values of t and the points of the curve, provided that we exclude certain finite sets from both the set of values of t and the points of the curve. Then conversely, the parameter t can be expressed as a rational function $t=\chi(x,y)$ of the coordinates x and y.

If the coefficients of the rational functions φ and ψ belong to some subfield k_0 of k and $t_0\in k_0$ then the coordinates of the point $(\varphi(t_0),\psi(t_0))$ also belong to k_0. This observation points to one possible application of the notion of rational curve. Suppose that $f(x,y)$ has rational coefficients. If we know that the curve given by (1.1) is rational, and that the coefficients of φ and ψ are in $\mathbb{Q}$, then the parametrisation $x=\varphi(t)$, $y=\psi(t)$ gives us all the rational points of this curve, except possibly a finite number, as t runs through all rational values. For example, all the rational solutions of the indeterminate equation (1.2) can be obtained from (1.3) as t runs through all rational values.

Another application of rational curves relates to integral calculus. We can view the equation of the curve (1.1) as determining y as an algebraic function of x. Then any rational function $g(x,y)$ is a (usually complicated) function of x. The rationality of the curve (1.1) implies the following important fact: for any rational function $g(x,y)$, the indefinite integral

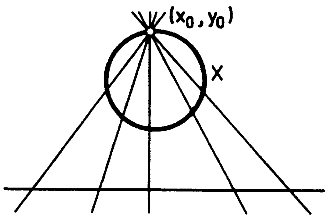


Figure 4 Projection of a conic

$$\int g(x, y)\mathrm{d}x \tag{1.5}$$

can be expressed in elementary functions. Indeed, since the curve is rational, it can be parametrised as $x = \varphi(t)$, $y = \psi(t)$ where φ, ψ are rational functions. Substituting these expressions in the integral (1.5), we reduce it to the form $\int g(\varphi(t), \psi(t))\varphi'(t)\mathrm{d}t$, which is an integral of a rational function. It is known that an integral of this form can be expressed in elementary functions. Substituting the expression $t = \chi(x, y)$ for the parameter in terms of the coordinates, we get an expression for the integral (1.5) as an elementary function of the coordinates.

We now give some examples of rational curves. Curves of degree 1, that is, lines, are obviously rational. Let us prove that an irreducible conic X is rational. Choose a point (x_0, y_0) on X. Consider the line through (x_0, y_0) with slope t. Its equation is

$$y - y_0 = t(x - x_0). \tag{1.6}$$

We find the points of intersection of X with this line; to do this, solve (1.6) for y and substitute this in the equation of X. We get the equation for x

$$f\bigl(x, y_0 + t(x - x_0)\bigr) = 0, \tag{1.7}$$

which has degree 2, as one sees easily. We know one root of this quadratic equation, namely $x = x_0$, since by assumption (x_0, y_0) is on the curve. Divide (1.7) by the coefficient of x^2, and write A for the coefficient of x in the resulting equation; the other root is then determined by $x + x_0 = -A$. Since t appears in the coefficients of (1.7), A is a rational function of t. Substituting the expression $x = -x_0 - A$ in (1.6), we get an expression for y also as a rational function of t. These expressions for x and y satisfy the equation of the curve, as can be seen from their derivation, and thus prove that the curve is rational.

The above parametrisation has an obvious geometric interpretation. A point (x, y) of X is sent to the slope of the line joining it to (x_0, y_0); and the parameter t is sent to the point of intersection of the curve with the line through (x_0, y_0) with slope t. This point is uniquely determined precisely because we are dealing with an irreducible curve of degree 2. In the same way as the parametrisation of the curve (1.2), this parametrisation can be interpreted as the projection of X from the point (x_0, y_0) to some line not passing through this point (Figure 4).

Note that in constructing the parametrisation we have used a point (x_0, y_0) of X. If the coefficients of the polynomial $f(x, y)$ and the coordinates of (x_0, y_0) are contained in some subfield k_0 of k, then so do the coefficients of the functions giving the parametrisation. Thus we can, for example, find the general form for the solution in rational numbers of an indeterminate equation of degree 2 if we know just one solution.

The question of whether there exists one solution is rather delicate. For the rational number field $\mathbb{Q}$ it is solved by Legendre's theorem (see for example Borevich and Shafarevich [15, Section 7.2, Chapter 1]).

We consider another application of the parametrisation we have found. The second degree equation $y^2 = ax^2 + bx + c$ defines a rational curve, as we have just seen. It follows from this that for any rational function $g(x, y)$, the integral $\int g(x, \sqrt{ax^2 + bx + c})\mathrm{d}x$ can be expressed in elementary functions. The parametrisation we have given provides an explicit form of the substitutions that reduce this integral to an integral of a rational function. It is easy to see that this leads to the well-known *Euler substitutions*.

The examples considered above lead us to the following general question: how can we determine whether an arbitrary algebraic plane curve is rational? This question relates to quite delicate ideas of algebraic geometry, as we will see later.

1.3 Relation with Field Theory

We now show how the question at the end of Section 1.2 can be formulated as a problem of field theory. To do this, we assign to every irreducible plane curve a certain field, by analogy with the way we assign to an irreducible polynomial in one variable the smallest field extension in which it has a root.

Let X be the irreducible curve given by (1.1). Consider rational functions $u(x, y) = p(x, y)/q(x, y)$, where p and q are polynomials with coefficients in k such that the denominator $q(x, y)$ is not divisible by $f(x, y)$. We say that such a function $u(x, y)$ is a *rational function* defined on X; and two rational functions $p(x, y)/q(x, y)$ and $p_1(x, y)/q_1(x, y)$ defined on X are *equal* on X if the polynomial $p(x, y)q_1(x, y) - q(x, y)p_1(x, y)$ is divisible by $f(x, y)$. It is easy to check that rational functions on X, up to equality on X, form a field. This field is called the *function field* or *field of rational functions* of X, and denoted by $k(X)$.

A rational function $u(x, y) = p(x, y)/q(x, y)$ is defined at all points of X where $q(x, y) \neq 0$. Since by assumption q is not divisible by f, by Lemma of Section 1.1, there are only finitely many points of X at which $u(x, y)$ is not defined. Hence we can also consider elements of $k(X)$ as functions on X, but defined everywhere except at a finite set. It can happen that a rational function u has two different expressions $u = p/q$ and $u = p_1/q_1$, and that for some point $(\alpha, \beta) \in X$ we have $q(\alpha, \beta) = 0$ but $q_1(\alpha, \beta) \neq 0$. For example, the function $u = (1 - y)/x$ on the circle $x^2 + y^2 = 1$ at the point $(0, 1)$ has an alternative expression $u = x/(1 + y)$ whose denominator does not vanish at $(0, 1)$. If u has an expression $u = p/q$ with $q(P) \neq 0$ then we say that u is *regular* at P.

Every element of $k(X)$ can obviously be written as a rational function of x and y; now x, y are algebraically dependent, since they are related by $f(x, y) = 0$. It is easy to check from this that $k(X)$ has transcendence degree 1 over k.

If X is a line, given say by $y = 0$, then every rational function $\varphi(x, y)$ on X is a rational function $\varphi(x, 0)$ of x only, and hence the function field of X equals the field of rational functions in one variable, $k(X) = k(x)$.

Now assume that the curve X is rational, say parametrised by $x = \varphi(t)$, $y = \psi(t)$. Consider the substitution $u(x, y) \mapsto u(\varphi(t), \psi(t))$ that takes any rational function $u = p(x, y)/q(x, y)$ on X into the rational function in t obtained by substituting $\varphi(t)$ for x and $\psi(t)$ for y. We check first that this substitution makes sense, that is, that the denominator $q(\varphi(t), \psi(t))$ is not identically 0 as a function of t. Assume that $q(\varphi(t), \psi(t)) = 0$, and compare this equality with (1.4). Recalling that the field k is algebraically closed, and therefore infinite, by making t take different values in k, we see that $f(x, y) = 0$ and $q(x, y) = 0$ have infinitely many common solutions. But by Lemma of Section 1.1, this is only possible if f and q have a common factor.

Thus our substitution sends any rational function $u(x, y)$ defined on X into a well-defined element of $k(t)$. Moreover, since φ and ψ satisfy the relation (1.4), the substitution takes rational functions u, u_1 that are equal on X to the same rational function in t. Thus every element of $k(X)$ goes to a well-defined element of $k(t)$. This map is obviously an isomorphism of $k(X)$ with some subfield of $k(t)$. It takes an element of k to itself.

At this point we make use of a theorem on rational functions. This is the result known as Lüroth's theorem, that asserts that a subfield of the field $k(t)$ of rational functions containing k is of the form $k(g(t))$, where $g(t)$ is some rational function; that is, the subfield consists of all the rational functions of $g(t)$. If $g(t)$ is not constant, then sending $f(u) \mapsto f(g(t))$ obviously gives an isomorphism of the field of rational functions $k(u)$ with $k(g(t))$. Thus Lüroth's theorem can be given the following statement: a subfield of the field of rational functions $k(t)$ that contains k and is not equal to k is itself isomorphic to the field of rational functions. Lüroth's theorem can be proved from simple properties of field extensions (see van der Waerden [76, 10.2 (Section 73)]). Applying it to our situation, we see that if X is a rational curve then $k(X)$ is isomorphic to the field of rational functions $k(t)$. Suppose, conversely, that for some curve X given by (1.1), the field $k(X)$ is isomorphic to the field of rational functions $k(t)$. Suppose that under this isomorphism x corresponds to $\varphi(t)$ and y to $\psi(t)$. The polynomial relation $f(x, y) = 0 \in k(X)$ is respected by the field isomorphism, and gives $f(\varphi(t), \psi(t)) = 0$; therefore X is rational.

It is easy to see that any field $K \supset k$ having transcendence degree 1 over k and generated by two elements x and y is isomorphic to a field $k(X)$, where X is some irreducible algebraic plane curve. Indeed, x and y must be connected by a polynomial relation, since K has transcendence degree 1 over k. If this dependence relation is $f(x, y) = 0$, with f an irreducible polynomial, then we can obviously take X to be the algebraic curve defined by this equation. It follows from this that the question on rational curves posed at the end of Section 1.2 is equivalent to the following question of field theory: when is a field $K \supset k$ with transcendence degree 1 over k and generated by two elements x and y isomorphic to the field of rational functions of one variable $k(t)$? The requirement that K is generated over k by two elements is not very natural from the algebraic point of view. It would be more natural to consider field extensions generated by an arbitrary finite number of elements. However, we will prove later that doing this does not give a more general notion (compare Theorem 1.8 and Proposition A.7).

In conclusion, we note that the preceding arguments allow us to solve the problem of obtaining a generically one-to-one parametrisation of a rational curve. Let X be a rational curve. By Lüroth's theorem, the field $k(X)$ is isomorphic to the field of rational functions $k(t)$. Suppose that this isomorphism takes x to $\varphi(t)$ and y to $\psi(t)$. This gives the parametrisation $x = \varphi(t)$, $y = \psi(t)$ of X.

Proposition *The parametrisation $x = \varphi(t)$, $y = \psi(t)$ has the following properties*:

(i) *Except possibly for a finite number of points*, *any $(x_0, y_0) \in X$ has a representation $(x_0, y_0) = (\varphi(t_0), \psi(t_0))$ for some t_0.*
(ii) *Except possibly for a finite number of points*, *this representation is unique*.

Proof Suppose that the function that maps to t under the isomorphism $k(X) \to k(t)$ is $\chi(x, y)$. Then the inverse isomorphism $k(t) \to k(X)$ is given by the formula $u(t) \mapsto u(\chi(x, y))$. Writing out the fact that the correspondences are inverse to one another gives

$$x = \varphi\bigl(\chi(x, y)\bigr), \qquad y = \psi\bigl(\chi(x, y)\bigr), \tag{1.8}$$

$$t = \chi\bigl(\varphi(t), \psi(t)\bigr). \tag{1.9}$$

Now (1.8) implies (i). Indeed, if $\chi(x, y) = p(x, y)/q(x, y)$ and $q(x_0, y_0) \neq 0$, we can take $t_0 = \chi(x_0, y_0)$; there are only finitely many points $(x_0, y_0) \in X$ at which $q(x_0, y_0) = 0$, since $q(x, y)$ and $f(x, y)$ are coprime. Suppose that (x_0, y_0) is such that $\chi(x_0, y_0)$ is distinct from the roots of the denominators of $\varphi(t)$ and $\psi(t)$; there are only finitely many points (x_0, y_0) for which this fails, for similar reasons. Then formula (1.8) gives the required representation of (x_0, y_0). In the same way, it follows from (1.9) that the value of the parameter t, if it exists, is uniquely determined by the point (x_0, y_0), except possibly for the finite number of points at which $q(x_0, y_0) = 0$. The proposition is proved. □

Note that we have proved (i) and (ii) not for any parametrisation of a rational curve, but for a specially constructed one. For an arbitrary parametrisation, (ii) can be false: for example, the curve (1.2) has, in addition to the parametrisation given by (1.3), another parametrisation $x = t^4 - 1$, $y = t^2(t^4 - 1)$, obtained from (1.3) on replacing t by t^2. Obviously here the values t and $-t$ of the parameter correspond to the same point of the curve.

1.4 Rational Maps

A rational parametrisation is a particular case of a more general notion. Let X and Y be two irreducible algebraic plane curves, and $u, v \in k(X)$. The map $\varphi(P) =$

$(u(P), v(P))$ is defined at all points P of X where both u and v are defined; it is called a *rational map* from X to Y if $\varphi(P) \in Y$ for every $P \in X$ at which φ is defined. If Y has the equation $g = 0$ then $g(u, v) \in k(X)$ must vanish at all but finitely many points of X, and therefore we must have $g(u, v) = 0 \in k(X)$.

For example, the projection from a point P considered in Section 1.2 is a rational map of X to the line. A rational parametrisation of a rational curve X is a rational map of the line to X.

A rational map $\varphi\colon X \to Y$ is *birational*, or is a *birational equivalence* of X to Y, if φ has a rational inverse, that is, if there exists a rational map $\psi : Y \to X$ such that $\varphi \circ \psi$ and $\psi \circ \varphi$ are the identity (at the points where they are defined). In this case, we say that X and Y are *birational*, or *birationally equivalent*.

A birational map is not constant, that is, at least one of the functions defining it is not an element of k. Indeed, a constant map is defined everywhere, and sends X to a single point $Q \in Y$. Taking any point $Q' \neq Q$ at which the inverse ψ of φ is defined contradicts the definition.

It follows that for any point $Q \in Y$ the inverse image $\varphi^{-1}(Q)$ of Q (the set of points $P \in X$ such that $\varphi(P) = Q$) is finite; this follows at once from Lemma of Section 1.1. Let S be the finite set of points of X at which a birational map $\varphi\colon X \to Y$ is not defined, $U = X \setminus S$ its complement, and T and V the same for $\psi : Y \to X$. It follows from what we said above that the complement in X of $\varphi^{-1}(V) \cap U$ and in Y of $\psi^{-1}(U) \cap V$ are finite, and φ establishes a one-to-one correspondence between $\varphi^{-1}(V) \cap U$ and $\psi^{-1}(U) \cap V$.

Birational equivalence is a fundamental equivalence relation in algebraic geometry, and we usually classify algebraic curves up to birational equivalence. We have seen that the rational curves are exactly the curves birational to the line.

Suppose that the equation $f(x, y)$ of an irreducible curve of degree n is a polynomial all of whose terms are monomials in x and y of degree $n - 1$ and n only. Then the projection from the origin defines a birational map of our curve and the line: this can be proved by a direct generalisation of the arguments for the curve (1.2).

Now suppose that the equation f has terms of degrees $n - 2$, $n - 1$ and n, that is, $f = u_{n-2} + u_{n-1} + u_n$, where u_i is homogeneous of degree i. Again we set $y = tx$ and cancel the factor of x^{n-2} from the equation, thus reducing it to the form $a(t)x^2 + b(t)x + c(t) = 0$, where $a(t) = u_n(1, t)$, $b(t) = u_{n-1}(1, t)$ and $c(t) = u_{n-2}(1, t)$. Setting $s = 2ax + b$ to complete the square (assuming that the ground field has characteristic $\neq 2$), we see that our curve is birational to the curve given by $s^2 = p(t)$, where $p = b^2 - 4ac$. A curve of this type is called a *hyperelliptic curve*. If $p(t)$ has even degree $2m$ then rewriting it in the form $p(t) = q(t)(t - \alpha)$ and dividing both sides of the equation through by $(t - \alpha)^{2m}$ shows that the curve is birational to the curve given by

$$\eta^2 = h(\xi), \quad \text{where } \xi = \frac{1}{t - \alpha},\ \eta = \frac{s}{(t - \alpha)^m} \text{ and } h(\xi) = \frac{q(t)}{(t - \alpha)^{2m-1}},$$

in which h is a polynomial of degree $\leq 2m - 1$ in ξ.

These ideas apply in particular to any cubic curve, if we take the origin to be any point of the curve. We see that, if $\operatorname{char} k \neq 2$, an irreducible cubic curve is birational to a curve given by $y^2 = f(x)$ where f is a polynomial of degree ≤ 3. If $f(x)$ has degree ≤ 2 then the cubic is rational. If it has degree 3 then we can assume that its leading coefficient is 1. Then the equation takes the form

$$y^2 = x^3 + ax^2 + bx + c.$$

This is called the *Weierstrass normal form* of the equation of a cubic. If $\operatorname{char} k \neq 3$ then after making a translation $x \mapsto x - a/3$ we can reduce the equation to the form

$$y^2 = x^3 + px + q. \tag{1.10}$$

Let X and Y be two irreducible algebraic plane curves that are birational, and suppose that the maps between them are given by

$$(u, v) = \bigl(\varphi(x, y), \psi(x, y)\bigr) \quad \text{and} \quad (x, y) = \bigl(\xi(u, v), \eta(u, v)\bigr).$$

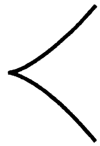
As in our study of rational curves, we can establish a relation between the function fields $k(X)$ and $k(Y)$ of these two curves. For this, we send a rational function $w(x, y) \in k(X)$ to $w(\xi(u, v), \eta(u, v))$, viewed as a rational function on Y. It is easy to check that this defines a map $k(X) \to k(Y)$ that is an isomorphism between these two fields. Conversely, if $k(X)$ and $k(Y)$ are isomorphic, then under this isomorphism $x, y \in k(X)$ correspond to functions $\xi(u, v), \eta(u, v) \in k(Y)$, and $u, v \in k(Y)$ to functions $\varphi(x, y), \psi(x, y) \in k(X)$, and it is again trivial to check that the pairs of functions φ, ψ and ξ, η define birational maps between the curves X and Y. Thus two curves are birational if and only if their rational function fields are isomorphic.

We see that the problem of classifying algebraic curves up to birational equivalence is a geometric aspect of the natural algebraic problem of classifying finitely generated extension fields of k of transcendence degree 1 up to isomorphism. In this problem, it is also natural not to restrict to fields of transcendence degree 1, but to consider fields of any finite transcendence degree. We will see later that this wider formulation of the problem also has a geometric interpretation. However, for this we have to leave the framework of the theory of algebraic curves, and consider algebraic varieties of any dimension.

1.5 Singular and Nonsingular Points

We borrow a definition from coordinate geometry: a point P is a *singular point* or *singularity* of the curve defined by $f(x, y) = 0$ if $f'_x(P) = f'_y(P) = f(P) = 0$, where f'_x denotes the partial derivative $\partial f/\partial x$. If we translate P to the origin, we can say that $(0, 0)$ is singular if f does not have constant or linear terms. A point

Figure 5 A cusp



is *nonsingular* if it is not singular, that is, if $f'_x(P)$ or $f'_y(P) \neq 0$. A curve all of whose points are nonsingular is *nonsingular* or *smooth*. It is well known that an irreducible conic is nonsingular; the simplest example of a singular curve is the curve of (1.2).

For an irreducible curve, either f'_x vanishes at only finitely many points of the curve, or f'_x is divisible by f. However, since f'_x has smaller degree than f, the latter is only possible if $f'_x = 0$. The same holds for f'_y. But $f'_x = f'_y = 0$ implies, if $\operatorname{char} k = 0$, that $f \in k$, and, if $\operatorname{char} k = p > 0$, that f involves x and y only as pth powers; in this last case, taking pth roots of the coefficients of f and using the well-known characteristic p identity $(\alpha + \beta)^p = \alpha^p + \beta^p$, we deduce that

$$f = \sum a_{ij} x^{pi} y^{pj} = \Bigl(\sum b_{ij} x^i y^j\Bigr)^p \quad \text{where } b_{ij}^p = a_{ij},$$

which contradicts the irreducibility of the curve. This shows that an irreducible curve has only a finite number of singular points.

If $P = (0,0)$ and the leading terms in the equation of the curve have degree r, then r is called the *multiplicity* of P, and we say that P is an *r-tuple point*, or *point of multiplicity r*. Thus a nonsingular point has multiplicity 1. If $P = (0,0)$ has multiplicity 2 and the terms of degree 2 in the equation of the curve are $ax^2 + bxy + cy^2$ then there are two possibilities: (a) $ax^2 + bxy + cy^2$ factorises into two distinct linear factors; or (b) $ax^2 + bxy + cy^2$ is a perfect square. In case (a) the singularity is called a *node* (see Figure 3), and in case (b) a *cusp* (Figure 5).

It follows from the definition that a curve of degree n cannot have a singularity of multiplicity $>n$. If a singular point has multiplicity n then the equation of the curve is a homogeneous polynomial in x and y of degree n, and therefore factorises as a product of linear factors, so that the curve is reducible. In Section 1.4 we proved that if an irreducible curve of degree n has a point of multiplicity $n-1$ it is rational, and if it has a point of multiplicity $n-2$ then it is hyperelliptic. The cubic curve written in Weierstrass normal form (1.10) is nonsingular if and only if the cubic polynomial on the right-hand side has no multiple roots, that is, $4p^3 + 27q^2 \neq 0$. In this case it is called an *elliptic curve*.

If $k = \mathbb{R}$ and P is a nonsingular point of the curve with equation $f(x,y) = 0$, and $f'_y(P) \neq 0$, say, then by the implicit function theorem we can write y as a function of x in some neighbourhood of P. Substituting this expression for y, this represents any rational function on the curve as a function of x near P.

When k is a general field, x can still be used to describe all the rational functions on the curve, admittedly to a more modest extent. For simplicity, set $P = (0,0)$. Then $f = \alpha x + \beta y + g$, where g contains only terms of degree ≥ 2 and $\beta \neq 0$. We distinguish the terms in f that involve x only, writing $f = x\varphi(x) + y\beta + yh$, with $h(0,0)=0$. Thus on the curve $f=0$ we have $y(\beta+h)=-x\varphi(x)$, or, in other words, $y=xv$, where $v=-\varphi(x)/(\beta+h)$ is a regular function at P (because $\beta+h(P)\neq 0$).

Let u be any rational function on our curve that is regular at P and has $u(P)=0$. Then $u=p/q$, where $p,q\in k[x,y]$ with $p(P)=0$ and $q(P)\neq 0$. Substituting our expression for y in this gives $p(x,y)=p(x,xv)=xr$ (because p has no constant term), where r is a regular function on the curve, and hence $u=xr/q=xu_1$. If $u_1(P)=0$ then we can repeat the argument, getting $u=x^2u_2$, and so on. We now prove that, provided u is not identically 0 on the curve, this process must stop after a finite number of steps.

For this, return to the expression $u=p/q$, in which, by assumption, p is not divisible by f. Hence there exist $\xi,\eta\in k[x,y]$ and a polynomial $a\in k[x]$ with $a\neq 0$ such that $f\xi+p\eta=a$ (we have already used this argument in the proof of Lemma of Section 1.1). Suppose $a=x^ka_0$ with $a_0(0)\neq 0$. Then $p\eta=a$ on the curve, and a representation $p=x^lw$ with $l>k$ would give a contradiction: $x^k(x^{l-k}w-a_0)=0$ on the curve, that is, $x^{l-k}w-a_0=0$. If $w=c/d$ with $c,d\in k[x,y]$ and $d(P)\neq 0$ then $x^{l-k}c-a_0d=0$ on the curve, that is, $x^{l-k}c-a_0d$ is divisible by f. But this is impossible, since x^{l-k} vanishes at P and a_0d does not. Since any rational function is a ratio of regular functions, we have proved the following theorem.

Theorem 1.1 *At any nonsingular point P of an irreducible algebraic curve, there exists a regular function t that vanishes at P and such that every rational function u that is not identically* 0 *on the curve can be written in the form*

$$u=t^kv, \tag{1.11}$$

with v regular at P and $v(P)\neq 0$. The function u is regular at P if and only if $k\geq 0$ in (1.11).

A function t with this property is called a *local parameter* on the curve at P. Obviously two different local parameters are related by $t'=tv$, where v is regular at P and $v(P)\neq 0$. We saw in the proof of the theorem that if $f'_y(P)\neq 0$ then x can be taken as a local parameter.

The number k in (1.11) is called the *multiplicity of the zero* of u at P. It is independent of the choice of the local parameter.

Let X and Y be algebraic curves with equations $f=0$ and $g=0$, and suppose that X is irreducible and not contained in Y, and that $P\in X\cap Y$ is a nonsingular point of X. Then g defines a function on X that is not identically zero; the multiplicity of the zero of g at P is called the *intersection multiplicity*[2] of X and Y at P. The notion of intersection multiplicity is one of the amendments needed in a correct

[2]This is discussed at length later in the book; see Section 1.1, Chapter 4 for the general definition of intersection multiplicity, which is symmetric in X and Y, and for the fact that it coincides with the simple notion used here.

statement of Bézout's theorem: for the theorem that the number of roots of a polynomial is equal to its degree is false unless we count roots with their multiplicities. Here we analyse intersection multiplicities in the case that X is a line.

Let $P = (\alpha, \beta) \in X$, and suppose that the equation of X is written in the form $f(x, y) = a(x - \alpha) + b(y - \beta) + g$, where the polynomial g expanded in powers of $x - \alpha$ and $y - \beta$ has only terms of degree ≥ 2. We write the equation of a line L through P in the form

$$x = \alpha + \lambda t, \qquad y = \beta + \mu t. \tag{1.12}$$

t is a local parameter on L at P. The restriction of f to L is of the form

$$f(\alpha + \lambda t, \beta + \mu t) = (a\lambda + b\mu)t + t^2\varphi(t).$$

From this we see that if P is singular, that is, if $a = b = 0$, then every line through P has intersection multiplicity >1 with X at P. On the other hand, if the curve is nonsingular, then there is only one such line, namely that for which $a\lambda + b\mu = 0$, with equation $a(x - \alpha) + b(y - \beta) = 0$. Obviously $a = f'_x(P)$, $b = f'_y(P)$, and hence this equation can we expressed

$$f'_x(P)(x - \alpha) + f'_y(P)(y - \beta) = 0. \tag{1.13}$$

The line given by this equation is called the *tangent line* to X at the nonsingular point P.

We now determine when a line has intersection multiplicity ≥ 3 with a curve at a nonsingular point $P = (\alpha, \beta)$. For this, we write the equation in the form

$$\begin{aligned} f(x, y) = {} & a(x - \alpha) + b(y - \beta) \\ & + c(x - \alpha)^2 + d(x - \alpha)(y - \beta) + e(y - \beta)^2 + h, \end{aligned} \tag{1.14}$$

where h is a polynomial which has only terms of degree ≥ 3 when expanded in power of $x - \alpha$ and $y - \beta$. Restricting f to the line L given by (1.12), we get that $f = (a\lambda + b\mu)t + (c\lambda^2 + d\lambda\mu + e\mu^2)t^2 + t^3\psi(t)$. Therefore the intersection multiplicity will be ≥ 3 if the two conditions $a\lambda + b\mu = c\lambda^2 + d\lambda\mu + e\mu^2 = 0$ hold. The first of these, as we have seen, means that L is the tangent line to X at P, and the second that moreover $cu^2 + duv + ev^2$ is divisible by $au + bv$ as a homogeneous polynomial in u, v. Together they show that $q = au + bv + cu^2 + duv + ev^2$ is reducible: it is divisible by $au + bv$. Conversely, if q is reducible, then $q = rs$, and r and s must have degree 1, and one of them, say r, must vanish when $u = v = 0$. But then r is proportional to $au + bv$ and $cu^2 + duv + ev^2$ is divisible by it. Thus the reducibility of the conic $q = au + bv + cu^2 + duv + ev^2$ is a necessary and sufficient condition for there to exist a line L through P with intersection multiplicity ≥ 3 at P. Such a point is called an *inflexion point* or *flex* of X.

We know from coordinate geometry the condition for a conic to be reducible. We assume that k has characteristic $\neq 2$; then recalling that $a = f'_x(P)$, $b = f'_y(P)$,

$c = (1/2) f''_x(P)$, $d = f''_{xy}(P)$ and $e = (1/2) f''_{yy}(P)$, we can write this condition in the form

$$\begin{vmatrix} f''_{xx} & f''_{xy} & f'_x \\ f''_{xy} & f''_{yy} & f'_y \\ f'_x & f'_y & 0 \end{vmatrix} (P) = 0. \tag{1.15}$$

1.6 The Projective Plane

We return to Bézout's theorem stated in Section 1.1. Even if we consider points with coordinates in an algebraically closed field and take account of multiplicities of intersections, this fails in very simple cases, and still needs one further amendment. This can already be seen in the example of two lines, which have no points of intersection if they are parallel. However, on the projective plane, parallel lines do intersect, in a point of the line at infinity.

In the same way, any two circles in the plane, although they are curves of degree 2, have at most 2 points of intersection, and never 4 as predicted by Bézout's theorem. This follows from the fact that the quadratic term in the equation of all circles is always the same, namely $x^2 + y^2$, so that subtracting the equation of one circle from that of the other gives a linear equation, and therefore the intersection of two circles is the same thing as the intersection of a circle and a line. Moreover, if the circles are not tangent, their multiplicity of intersection is 1 at each point of intersection.

To understand what lies behind this failure of Bézout's theorem, write the equation of the circle $(x-a)^2 + (y-b)^2 = r^2$ in homogeneous coordinates by setting $x = \xi/\zeta$ and $y = \eta/\zeta$. We get the equation $(\xi - a\zeta)^2 + (\eta - b\zeta)^2 = r^2\zeta^2$, from which we see that the circle intersects the line at infinity $\zeta = 0$ in the points $\xi^2 + \eta^2 = 0$, that is, in the two circular points at infinity $(1, \pm i, 0)$. Thus all circles have the two points $(1, \pm i, 0)$ at infinity in common. Taken together with the two finite points of intersection, we thus get 4 points of intersection, in agreement with Bézout's theorem. This type of phenomenon motivates passing from the affine to the projective plane.

Recall that a point of the projective plane $\mathbb{P}^2$ is determined by 3 elements (ξ, η, ζ) of the field k, not all simultaneously zero. Two triples (ξ, η, ζ) and (ξ', η', ζ') determine the same point if there exists $\lambda \in k$ with $\lambda \neq 0$ such that $\xi = \lambda\xi'$, $\eta = \lambda\eta'$ and $\zeta = \lambda\zeta'$. Any triple (ξ, η, ζ) defining a point P is called a set of *homogeneous coordinates* of P, and we write $P = (\xi : \eta : \zeta)$.

There is an inclusion $\mathbb{A}^2 \subset \mathbb{P}^2$ which sends $(x, y) \in \mathbb{A}^2$ to $(x : y : 1)$. We get in this way all points with $\zeta \neq 0$: a point $(\xi : \eta : \zeta) \in \mathbb{P}^2$ with $\zeta \neq 0$ corresponds to the point $(\xi/\zeta, \eta/\zeta) \in \mathbb{A}^2$. The points of the complementary set $\zeta = 0$ are called *points at infinity*. This notion is related to the choice of the coordinate ζ. In fact, $\mathbb{P}^2$ contains 3 sets that are copies of the affine plane in this way: $\mathbb{A}^2_1$ (given by $\xi \neq 0$), $\mathbb{A}^2_2$ (given by $\eta \neq 0$), and $\mathbb{A}^2_3$ (given by $\zeta \neq 0$). These intersect, of course: if a point

$P \in \mathbb{A}_3^2$ has coordinates $x = \xi/\zeta$, $y = \eta/\zeta$ and $\eta \neq 0$ then in $\mathbb{A}_2^2$ the same point has coordinates $x' = \xi/\eta$, $y' = \zeta/\eta$, so that $x' = x/y$, $y' = 1/y$; if $\xi \neq 0$ then in $\mathbb{A}_1^2$ it has coordinates $x'' = \eta/\xi$, $y'' = \zeta/\xi$, so that $x'' = y/x$, $y'' = 1/x$. Every point $P \in \mathbb{P}^2$ is contained in at least one of the pieces $\mathbb{A}_1^2$, $\mathbb{A}_2^2$ or $\mathbb{A}_3^2$, and can be written down in the affine coordinates of that piece.

An algebraic curve in $\mathbb{P}^2$, or a *projective algebraic plane curve* is defined in homogeneous coordinates by an equation $F(\xi, \eta, \zeta) = 0$, where F is a homogeneous polynomial. Then whether $F(\xi, \eta, \zeta) = 0$ holds or not is independent of the choice of the homogeneous coordinates of a point; that is, it is preserved on passing from ξ, η, ζ to $\xi' = \lambda\xi$, $\eta' = \lambda\eta$, $\zeta' = \lambda\zeta$ with $\lambda \neq 0$. A homogeneous polynomial is also called a *form*. An affine algebraic curve of degree n with equation $f(x, y) = 0$ defines a homogeneous polynomial $F(\xi, \eta, \zeta) = \zeta^n f(\xi/\zeta, \eta/\zeta)$, and hence a projective curve with equation $F(\xi, \eta, \zeta) = 0$. It is easy to see that intersecting this curve with the affine plane $\mathbb{A}_3^2$ gives us the original affine curve, to which it therefore only adds points at infinity with $\zeta = 0$. If the equation of the projective curve is $F(\xi, \eta, \zeta) = 0$, then that of the corresponding affine curve is $f(x, y) = 0$, where $f(x, y) = F(x, y, 1)$. Since every point $P \in \mathbb{P}^2$ is contained in one of the affine sets $\mathbb{A}_1^2$, $\mathbb{A}_2^2$ or $\mathbb{A}_3^2$, we can use this correspondence to write out the properties of curves, defined above for affine curves, in terms of homogeneous coordinates. We do this now for the notions of tangent line, singular point and inflexion point of an algebraic curve. We always assume that $P \in \mathbb{A}_3^2$.

In affine coordinates, the equation of the tangent is

$$\frac{\partial f}{\partial x}(P)(x - \alpha) + \frac{\partial f}{\partial y}(P)(y - \beta) = 0.$$

By assumption $f(x, y) = F(x, y, 1)$, where $F(\xi, \eta, \zeta) = 0$ is the homogeneous equation of our curve. Hence writing F'_x etc. for the partial derivatives, we get $\partial f/\partial x = F'_x(x, y, 1)$ and $\partial f/\partial y = F'_y(x, y, 1)$, and by the well-known theorem of Euler on homogeneous functions, we have

$$F'_\xi \xi + F'_\eta \eta + F'_\zeta \zeta = nF.$$

Since $P = (\alpha : \beta : 1)$ is a point of the curve, $F'_\xi(P)\alpha + F'_\eta(P)\beta + F'_\zeta(P) = 0$, so that the equation of the tangent is $F'_\xi(P)x + F'_\eta(P)y + F'_\zeta(P) = 0$, or in homogeneous coordinates

$$F'_\xi(P)\xi + F'_\eta(P)\eta + F'_\zeta(P)\zeta = 0.$$

The conditions in affine coordinates for a singular point are $f'_x = f'_y = f = 0$. Hence in homogeneous coordinates $F'_\xi = F'_\eta = F = 0$, and by Euler's theorem, since $\zeta = 1$, also $F'_\zeta = 0$. If the characteristic of the field k is 0 then it is enough to require the conditions $F'_\xi(P) = F'_\eta(P) = F'_\zeta(P) = 0$, since then also $F(P) = 0$.

The condition defining an inflexion point is given by the relation (1.15). Here again $f(x, y) = F(x, y, 1)$, so that $f'_x = F'_x$, $f'_y = F'_y$, $f''_{xx} = F''_{xx}$, $f''_{xy} = F''_{xy}$, $f''_{yy} = F''_{yy}$. From now on, in the homogeneous polynomial F we write ξ for x and η

for y. We substitute these expressions in the determinant of (1.15), and use Euler's theorem

$$\begin{aligned} F''_{\xi\xi}\xi + F''_{\xi\eta}\eta + F''_{\xi\zeta}\zeta &= (n-1)F'_{\xi}, \\ F''_{\xi\eta}\xi + F''_{\eta\eta}\eta + F''_{\zeta\eta}\zeta &= (n-1)F'_{\eta}, \\ F'_{\xi}\xi + F'_{\eta}\eta + F'_{\zeta}\zeta &= nF. \end{aligned}$$

Multiply the last column of our determinant by $(n-1)$, and subtract from it ξ times the first column and η times the second. Using the above identities and recalling that $F(P)=0$, we get the determinant

$$\begin{vmatrix} F''_{\xi\xi} & F''_{\xi\eta} & F''_{\xi\zeta} \\ F''_{\xi\eta} & F''_{\eta\eta} & F''_{\zeta\eta} \\ F'_{\xi} & F'_{\eta} & F'_{\zeta} \end{vmatrix}(P).$$

Now perform the same operation on the rows of the determinant. The condition for P to be an inflexion point then takes the form

$$\begin{vmatrix} F''_{\xi\xi} & F''_{\xi\eta} & F''_{\xi\zeta} \\ F''_{\eta\xi} & F''_{\eta\eta} & F''_{\eta\zeta} \\ F''_{\zeta\xi} & F''_{\zeta\eta} & F''_{\zeta\zeta} \end{vmatrix}(P) = 0. \tag{1.16}$$

The determinant on the left-hand side of (1.16) is called the *Hessian form* of F, and denoted by $H(F)$.

We now proceed to considering rational functions. Making the substitution $x=\xi/\zeta$, $y=\eta/\zeta$ and clearing denominators, we can rewrite a rational function $f=p(x,y)/q(x,y)$ on $\mathbb{A}^2_3$ in the form $P(\xi,\eta,\zeta)/Q(\xi,\eta,\zeta)$, where P and Q are homogeneous polynomials of the same degree. Hence its value at a point $(\xi:\eta:\zeta)$ does not change on multiplying the homogeneous coordinates through by a common multiple, and hence f can be viewed as a partially defined function on $\mathbb{P}^2$.

Given a rational map $\varphi\colon \mathbb{A}^2_3 \to \mathbb{A}^2_3$ defined by $(x,y)\mapsto (u(x,y), v(x,y))$, we first rewrite it, as just explained, in the form

$$\frac{U(\xi,\eta,\zeta)}{R(\xi,\eta,\zeta)}, \quad \frac{V(\xi,\eta,\zeta)}{S(\xi,\eta,\zeta)},$$

where U, V, R, S are homogeneous polynomials, with $\deg U = \deg R$ and $\deg V = \deg S$. Next we put the two components over a common denominator, that is, in the form $(A/C, B/C)$, with $\deg A = \deg B = \deg C$. Finally, introducing homogeneous coordinates $\xi'/\zeta' = A/C$, $\eta'/\zeta' = B/C$, we write the map in the form

$$(\xi:\eta:\zeta)\mapsto \bigl(A(\xi:\eta:\zeta):B(\xi:\eta:\zeta):C(\xi:\eta:\zeta)\bigr),$$

where A, B, C are homogeneous polynomials of the same degree. Now φ is naturally a rational map $\mathbb{P}^2 \to \mathbb{P}^2$. The map is regular at a point P if one of A, B, C does not vanish at P. Studying properties related to points P in the affine set $\mathbb{A}^2_3$, say, we can divide each of A, B, C by ζ^n, where n is their common degree, and write the map in the form $(x, y) \mapsto (u(x, y), v(x, y), w(x, y))$, where u, v and w are polynomials. This map is regular at P if the 3 polynomials do not vanish simultaneously at P.

As a first illustration we prove the following important result.

Theorem 1.2 *A rational map from a projective plane curve C to $\mathbb{P}^2$ is regular at every nonsingular point of C (see Section 1.5 for the definition).*

Proof Suppose that the nonsingular point P is in the affine piece $\mathbb{A}^2_3$ with coordinates denoted by x, y. We write the map as above in the form $(x, y) \mapsto (u_0 : u_1 : u_2)$ where u_0, u_1, u_2 are polynomials, and apply Theorem 1.1 to these. Restricting the u_i to C, we can write them in the form $u_i = t^{k_i} v_i$, where t is a local parameter, $v_i(P) \neq 0$ and $k_i \geq 0$ for $i = 0, 1, 2$. Suppose that k_0, say, is the smallest of the numbers k_0, k_1, k_2. Then the same map can be rewritten in the form $(x, y) \mapsto (v_0 : t^{k_1 - k_0} v_1 : t^{k_2 - k_0} v_2)$, with $k_1 - k_0 \geq 0$, $k_2 - k_0 \geq 0$, and $v_0(P) \neq 0$. It follows that it is regular at P. The theorem is proved. □

Corollary *A birational map between nonsingular projective plane curves is regular at every point, and is a one-to-one correspondence.*

As an example, consider a birational map of the projective line to itself. Just as with any rational map, this can be written as a rational function $x \mapsto p(x)/q(x)$, with $p(x), q(x) \in k[x]$ (here we assume that x is a coordinate on our line, for example the line given by $y = 0$). The points that map to a given point α are those for which $p(x)/q(x) = \alpha$, that is, $p(x) - \alpha q(x) = 0$. Hence from the fact that the map is birational, it follows that p and q are linear, that is, the map is of the form $x \mapsto (ax + b)/(cx + d)$ with $ad - bc \neq 0$. As a consequence, we get that *a birational map of the line to itself has at most two fixed points*, the roots of the equation $x(cx + d) = ax + b$.

Now consider the elliptic curve given by (1.10), and assume that $4p^3 + 27q^2 \neq 0$. All its finite points are nonsingular. Passing to homogeneous coordinates, we can write its equation in the form $\eta^2\zeta = \xi^3 + p\xi\zeta^2 + q\zeta^3$. Hence it has a unique point on the line at infinity $\zeta = 0$, namely the point $o = (0 : 1 : 0)$. Dividing through by η^3 we write the equation of the curve in the form $v = u^3 + puv^2 + qv^3$, in coordinates u, v, where $u = \xi/\eta$ and $v = \zeta/\eta$. The point $o = (0, 0)$ in these coordinates is also nonsingular. Hence our curve is nonsingular. The map $(x, y) \mapsto (x, -y)$ is obviously a birational map of the curve to itself. Its fixed points in the finite part of the plane are the points with $y = 0$, $x^3 + px + q = 0$, that is, there are 3 such points. The point o is also a fixed point, since $u = x/y$, $v = 1/y$, and in coordinates u, v, the map is written $(u, v) \mapsto (-u, -v)$. We have constructed on an elliptic curve an automorphism having 4 fixed points. It follows from this that *an elliptic curve is not birational to a line, that is, is not rational*. This shows that the problem of birational classification of curves is not trivial: not all curves are birational to one another.

Passing to projective curves is the final amendment required in the statement of Bézout's theorem. One version of this is as follows:

Theorem *Let X and Y be projective curves, with X nonsingular and not contained in Y. Then the sum of the multiplicities of intersection of X and Y at all points of $X \cap Y$ equals the product of the degrees of X and Y.*

We will prove this theorem and a series of generalisations in a later section (Section 2.2, Chapter 3 and Section 2.1, Chapter 4). Here we verify the two simplest cases, when X is a line or a conic.

Let X be a line. By Lemma of Section 1.1, X and Y have a finite number of points of intersection. We choose a convenient coordinate system, so that the line $\zeta = 0$ does not pass through the points of intersection, and is not equal to X, and $\eta = 0$ is the line X. Then the points of intersection of X and Y are contained in the affine plane with coordinates $x = \xi/\zeta$, $y = \eta/\zeta$, and the equation of X is $y = 0$. Let $f(x, y) = 0$ be the equation of the curve Y and $f = f_0 + f_1(x, y) + \cdots + f_n(x, y)$ its expression as a sum of homogeneous polynomials. The point $(1 : 0 : 0)$ is not contained in Y by the choice of the coordinate system, and hence $f_n(1, 0) \neq 0$, that is, f contains the term ax^n with $a \neq 0$. Hence $f(x, 0)$, the restriction of f to X, has degree n. The function $x - \alpha$ is a local parameter of X at the point $x = \alpha$, and the multiplicity of intersection of X and Y at this point equals the multiplicity of the root $x = \alpha$ of the polynomial $f(x, 0)$. Therefore the sum of these multiplicities equals n.

Let X be a conic. Take any point $P \in X$ with $P \notin Y$, and choose coordinates so that $\zeta = 0$ is the tangent line to X at P, and $\xi = 0$ some other line through P. An easy calculation in coordinates shows that X is a parabola in the affine plane with coordinates $x = \xi/\zeta$, $y = \eta/\zeta$ (since it touches the line at infinity), with equation $y = px^2 + qx + r$ and $p \neq 0$. As before, $f = f_0 + \cdots + f_n(x, y)$, and now $f_n(0, 1) \neq 0$, that is, $f(x, y)$ contains the term ay^n with $a \neq 0$. The conic X has no other points of intersection with the line $\zeta = 0$ except P, and hence all the points of intersection of X and Y are contained in the finite part of the plane. At any point with $x = \alpha$ the function $x - \alpha$ is a local parameter on X, and the multiplicity of intersection of X and Y at this point is equal to the multiplicity of the root $x = \alpha$ of the polynomial $f(x, px^2 + qx + r)$. Since $f(x, y)$ contains the term ay^n with $a \neq 0$, the degree of $f(x, px^2 + qx + r)$ is $2n$, so that the sum of multiplicities of all the points of intersection equals $2n$.

This proves the theorem in the case X is a line or conic.

Already this simple particular case of Bézout's theorem has beautiful geometric applications. One of these is the proof of Pascal's theorem, which asserts that for a hexagon inscribed in a conic, the 3 points of intersection of pairs of opposite sides are collinear. Let l_1 and m_1, l_2 and m_2, l_3 and m_3 be linear forms that are the equations of the opposite sides of a hexagon (see Figure 6). Consider the cubic with the equation $f_\lambda = l_1 l_2 l_3 + \lambda m_1 m_2 m_3$ where λ is an arbitrary parameter. This has six points of intersection with the conic, the vertexes of the hexagon. Moreover, we can choose the value of λ so that $f_\lambda(P) = 0$ for

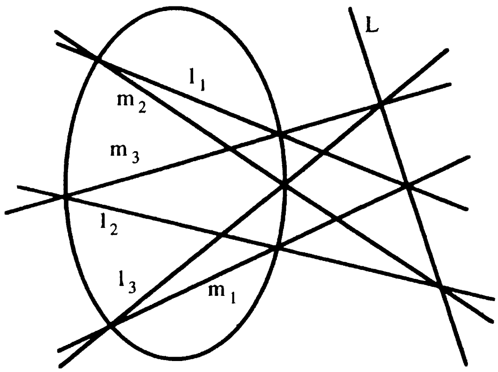


Figure 6 Pascal's theorem

any given point $P \in X$, distinct from these 6 points of intersection. We get a cubic f_λ having 7 points of intersection with a conic X, and by Bézout's theorem this must decompose as the conic X plus a line L. This line L must contain the points of intersection $l_1 \cap m_1$, $l_2 \cap m_2$ and $l_3 \cap m_3$. (This proof is due to Plücker.)

1.7 Exercises to Section 1

1 Find a characterisation in real terms of the line through the points of intersection of two circles in the case that both these points are complex. Prove that it is the locus of points having the same power with respect to both circles. (The power of a point with respect to a circle is the square of the distance between it and the points of tangency of the tangent lines to the circle.)

2 Which rational functions $p(x)/q(x)$ are regular at the point at infinity of $\mathbb{P}^1$? What order of zero do they have there?

3 Prove that an irreducible cubic curve has at most one singular point, and that the multiplicity of a singular point is 2. If the singularity is a node then the cubic is projectively equivalent to the curve in (1.2); and if a cusp then to the curve $y^2 = x^3$.

4 What is the maximum multiplicity of intersection of two nonsingular conics at a common point?

5 Prove that if the ground field has characteristic p then every line through the origin is a tangent line to the curve $y = x^{p+1}$. Prove that over a field of characteristic 0, there are at most a finite number of lines through a given point tangent to a given irreducible curve.

6 Prove that the sum of multiplicities of two singular points of an irreducible curve of degree n is at most n, and the sum of multiplicities of any 5 points is at most $2n$.

7 Prove that for any two distinct points of an irreducible curve there exists a rational function that is regular at both, and takes the value 0 at one and 1 at the other.

8 Prove that for any nonsingular points $P_1, \dots, P_r$ of an irreducible curve and numbers $m_1, \dots, m_r \geq 0$ there exists a rational function that is regular at all these points, and has a zero of multiplicity m_i at P_i.

9 For what values of m is the cubic $x_0^3 + x_1^3 + x_2^3 + m x_0 x_1 x_2 = 0$ in $\mathbb{P}^2$ nonsingular? Find its inflexion points.

10 Find all the automorphisms of the curve of (1.2).

11 Prove that on the projective line and on a conic of $\mathbb{P}^2$, a rational function that is regular at every point is a constant.

12 Give an interpretation of Pascal's theorem in the case that pairs of vertexes of the hexagon coincide, and the lines joining them become tangents.

2 Closed Subsets of Affine Space

Throughout what follows, we work with a fixed algebraically closed field k, which we call the *ground field*.

2.1 Definition of Closed Subsets

At different stages of the development of algebraic geometry, there have been changing views on the basic object of study, that is, on the question of what is the "natural definition" of an algebraic variety; the objects considered to be most basic have been projective or quasiprojective varieties, abstract algebraic varieties, schemes or algebraic spaces.

In this book, we consider algebraic geometry in a gradually increasing degree of generality. The most general notion considered in the first chapters, embracing all the algebraic varieties studied here, is that of quasiprojective variety. In the final chapters this role will be taken by schemes. At present we define a class of algebraic varieties that will play a foundational role in all the subsequent definitions. Since the word variety will be reserved for the more general notions, we use a different word here.

We write $\mathbb{A}^n$ for the n-dimensional affine space over the field k. Thus its points are of the form $\alpha = (\alpha_1, \dots, \alpha_n)$ with $\alpha_i \in k$.

Definition A *closed subset* of $\mathbb{A}^n$ is a subset $X \subset \mathbb{A}^n$ consisting of all common zeros of a finite number of polynomials with coefficients in k. We will sometimes say simply *closed set* for brevity.

From now on we will write $F(T)$ to denote a polynomial in n variables, allowing T to stand for the set of variables $T_1, \dots, T_n$. If a closed set X consists of all common zeros of polynomials $F_1(T), \dots, F_m(T)$, then we refer to $F_1(T) = \cdots = F_m(T) = 0$ as the *equations* of the set X.

A set X defined by an infinite system of equations $F_\alpha(T) = 0$ is also closed. Indeed, the ideal $\mathfrak{A}$ of the polynomial ring in $T_1, \dots, T_n$ generated by all the polynomials $F_\alpha(T)$ is finitely generated (the Hilbert Basis Theorem, see Atiyah and Macdonald [8, Theorem 7.5]), that is, $\mathfrak{A} = (G_1, \dots, G_m)$. One checks easily that X is defined by the system of equations $G_1 = \cdots = G_m = 0$.

It follows from this that the intersection of any number of closed sets is closed. Indeed, if X_α are closed sets, then to get a system of equations defining $X = \bigcap X_\alpha$, we need only take the union of the systems defining all the X_α.

The union of a finite number of closed sets is again closed. It is obviously enough to check this for two sets. If $X = X_1 \cup X_2$, where X_1 is defined by the system of equations $F_i(T) = 0$ for $i = 1, \dots, m$ and X_2 by $G_j(T) = 0$ for $j = 1, \dots, l$ then it is easy to check that X is defined by the system $F_i(T)G_j(T) = 0$ for $i = 1, \dots, m$ and $j = 1, \dots, l$.

Let $X \subset \mathbb{A}^n$ be a closed subset of affine space. We say that a set $U \subset X$ is *open* if its complement $X \setminus U$ is closed. Any open set $U \ni x$ is called a *neighbourhood* of x. The intersection of all the closed subsets of X containing a given subset $M \subset X$ is closed. It is called the *closure* of M and denoted by $\overline{M}$. A subset is *dense* in X if $\overline{M} = X$. This means that M is not contained in any closed subset $Y \subsetneqq X$.

Example 1.1 The whole affine space $\mathbb{A}^n$ is closed, since it is defined by the empty set of equations, or by $0 = 0$.

Example 1.2 The subset $X \subset \mathbb{A}^1$ consisting of all points except 0 is not closed: every polynomial $F(T)$ that vanishes at all $T \neq 0$ must be identically 0.

Example 1.3 Let us determine all the closed subsets $X \subset \mathbb{A}^1$. Such a set is given by a system of equations $F_1(T) = \cdots = F_m(T) = 0$ in one variable T. If all the F_i are identically 0 then $X = \mathbb{A}^1$. If the F_i don't have any common factor, then they don't have any common roots, and X does not contain any points. If the highest common factors of all the F_i is $D(T)$ then $D(T) = (T - \alpha_1) \cdots (T - \alpha_n)$ and X consists of the finitely many points $T = \alpha_1, \dots, T = \alpha_n$.

Example 1.4 Let us determine all the closed subsets $X \subset \mathbb{A}^2$. A closed subset is given by a system of equations

$$F_1(T) = \cdots = F_m(T) = 0, \tag{1.17}$$

where now $T = (T_1, T_2)$. If all the F_i are identically 0 then $X = \mathbb{A}^2$. Suppose this is not the case. If the polynomials $F_1, \dots, F_m$ do not have a common factor then, as follows from Lemma of Section 1.1, the system (1.17) has only a finite set of solutions (possibly empty). Finally, suppose that the highest common factor of all the $F_i(T)$ is $D(T)$. Then $F_i(T) = D(T)G_i(T)$, where now the polynomials $G_i(T)$ do not have a common factor. Obviously then $X = X_1 \cup X_2$ where X_1 is given by $G_1(T) = \cdots = G_m(T) = 0$ and X_2 is given by the single equations $D(T) = 0$. As we have seen, X_1 is a finite set. The closed sets defined in $\mathbb{A}^2$ by one equation are the algebraic plane curves. Thus a closed set $X \subset \mathbb{A}^2$ either consists of a finite set of points (possibly empty), or the union of an algebraic plane curve and a finite set of points, or the whole of $\mathbb{A}^2$.

Example 1.5 If $\alpha \in \mathbb{A}^r$ is the point with coordinates $(\alpha_1, \ldots, \alpha_r)$ and $\beta \in \mathbb{A}^s$ the point with coordinates $(\beta_1, \ldots, \beta_s)$, we take α, β into the point $(\alpha, \beta) \in \mathbb{A}^{r+s}$ with coordinates $(\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_s)$. Thus we identify $\mathbb{A}^{r+s}$ as the set of pairs (α, β) with $\alpha \in \mathbb{A}^r$ and $\beta \in \mathbb{A}^s$. Let $X \subset \mathbb{A}^r$ and $Y \subset \mathbb{A}^s$ be closed sets. The set of pairs $(x, y) \in \mathbb{A}^{r+s}$ with $x \in X$ and $y \in Y$ is called the *product* of X and Y, and denoted by $X \times Y$. This is again a closed set. Indeed, if X is given by $F_i(T) = 0$ and Y by $G_j(U) = 0$ then $X \times Y \subset \mathbb{A}^{r+s}$ is defined by $F_i(T) = G_j(U) = 0$.

Example 1.6 A set $X \subset \mathbb{A}^n$ defined by one equation $F(T_1, \ldots, T_n) = 0$ is called a *hypersurface*.

2.2 Regular Functions on a Closed Subset

Let X be a closed set in the affine space $\mathbb{A}^n$ over the ground field k.

Definition A function f defined on X with values in k is *regular* if there exists a polynomial $F(T)$ with coefficients in k such that $f(x) = F(x)$ for all $x \in X$.

If f is a given function, the polynomial F is in general not uniquely determined. We can add to F any polynomial entering in the system of equations of X without altering f. The set of all regular functions on a given closed set X forms a ring and an algebra over k; the operations of addition, multiplication and scalar multiplication by elements of k, are defined as in analysis, by performing the operations on the value of the functions at each point $x \in X$.The ring obtained in this way is denoted by $k[X]$ and is called the *coordinate ring* of X.

We write $k[T]$ for the polynomial ring with coefficients in k in variables $T_1, \ldots, T_n$. We can obviously associate with each polynomial $F \in k[T]$ a function $f \in k[X]$, by viewing F as a function on the set of points of X; in this way we get a homomorphism from $k[T]$ to $k[X]$. The kernel of this homomorphism consists of all polynomials $F \in k[T]$ that take the value 0 at every point $x \in X$. This is an ideal of $k[T]$, just as the kernel of any ring homomorphism; it is called the *ideal of the closed set* X, and denoted by $\mathfrak{A}_X$. Obviously

$$k[X] = k[T]/\mathfrak{A}_X.$$

Thus $k[X]$ is determined by the ideal $\mathfrak{A}_X \subset k[T]$.

Example 1.7 If X is a point then $k[X]=k$.

Example 1.8 If $X=\mathbb{A}^n$ then $\mathfrak{A}_X=0$ and $k[X]=k[T]$.

Example 1.9 Let $X \subset \mathbb{A}^2$ be given by the equation $T_1T_2=1$. Then $k[X]=k[T_1,T_1^{-1}]$, and it consists of all the rational functions in T_1 of the form $G(T_1)/T_1^n$ with $G(T_1)$ a polynomial and $n \geq 0$.

Example 1.10 We prove that if X and Y are any closed sets then $k[X\times Y]=k[X]\otimes_k k[Y]$. Define a homomorphism $\varphi\colon k[X]\otimes_k k[Y]\to k[X\times Y]$ by the condition

$$\varphi\left(\sum_i f_i\otimes g_i\right)(x,y)=\sum_i f_i(x)g_i(y).$$

The right-hand side is obviously a regular functions on $X\times Y$, and it is clear that φ is onto, since, in the notation of Example 1.5, the functions α_i and β_j are contained in the image of φ, and these generate $k[X\times Y]$. To prove that φ is one-to-one, it is enough to check that if $\{f_i\}$ are linearly independent in $k[X]$ and $\{g_j\}$ in $k[Y]$ then $\{f_i\otimes g_j\}$ are linearly independent in $k[X\times Y]$. Now an equality

$$\sum_{i,j} c_{ij} f_i(x)g_j(y)=0$$

implies the relation $\sum_j c_{ij}g_j(y)=0$ for any fixed y, and in turn that $c_{ij}=0$.

Since $k[X]$ is a homomorphic image of the polynomial ring $k[T]$, it satisfies the Hilbert basis theorem: any ideal of $k[X]$ is finitely generated. It also satisfies the following analogue of the Nullstellensatz (Proposition A.9): if a function $f\in k[X]$ is zero at every point $x\in X$ at which functions $g_1,\dots,g_m$ vanish then $f^r\in(g_1,\dots,g_m)$ for some $r>0$. Indeed, suppose that f is given by a polynomial $F(T)$, the g_i by polynomials $G_i(T)$, and let $F_j=0$ for $j=1,\dots,l$ be the equations of X. Then $F(T)$ vanishes at all points $\alpha\in\mathbb{A}^n$ at which all the polynomials $G_1,\dots,G_m,F_1,\dots,F_l$ vanish; for since $F_j(\alpha)=0$ it follows that $\alpha\in X$, and then by assumption $F(\alpha)=0$. Applying the Nullstellensatz in the polynomial ring we deduce that $F^r\in(G_1,\dots,G_m,F_1,\dots,F_l)$ for some $r>0$, and hence $f^r\in(g_1,\dots,g_m)$ in $k[X]$.

How is the ideal $\mathfrak{A}_X$ of a closed set X related to a system $F_1=\cdots=F_m=0$ of defining equations of X? Clearly $F_i\in\mathfrak{A}_X$ by definition of $\mathfrak{A}_X$, and hence $(F_1,\dots,F_m)\subset\mathfrak{A}_X$; however, it's not always true that $(F_1,\dots,F_m)=\mathfrak{A}_X$. For example, if $X\subset\mathbb{A}^1$ is defined by the equation T^2 then it consists just of the point $T=0$, so that $\mathfrak{A}_X$ consists of all polynomials with no constant term. That is, $\mathfrak{A}_X=(T)$, whereas $(F_1,\dots,F_m)=(T^2)$. We can however always define the same closed set X by a system of equation $G_1=\cdots=G_l=0$ in such a way that $\mathfrak{A}_X=(G_1,\dots,G_l)$. For this it is enough to recall that any ideal of $k[T]$ is finitely generated. Let $G_1,\dots,G_l$ be a basis of the ideal $\mathfrak{A}_X$, that is, $\mathfrak{A}_X=(G_1,\dots,G_l)$.

Then obviously the equations $G_1 = \cdots = G_l = 0$ define the same set X and have the required property. It is sometimes even convenient to consider a closed set as defined by the infinite system of equations $F = 0$ for all polynomials $F \in \mathfrak{A}_X$. Indeed, if $(F_1, \ldots, F_m) = \mathfrak{A}_X$ then these equations are all consequences of $F_1 = \cdots = F_m = 0$.

Relations between closed subsets are often reflected in their ideals. For example, if X and Y are closed sets in the affine space $\mathbb{A}^n$ then $X \supset Y$ if and only if $\mathfrak{A}_X \subset \mathfrak{A}_Y$. It follows from this that with any closed subset Y contained in X we can associate the ideal $\mathfrak{a}_Y$ of $k[X]$, consisting of the images under the homomorphism $k[T] \to k[X]$ of polynomials $F \in \mathfrak{A}_Y$. Conversely, any ideal $\mathfrak{a}$ of $k[X]$ defines an ideal $\mathfrak{A}$ in $k[T]$, consisting of all inverse images under $k[T] \to k[X]$ of elements of $\mathfrak{a}$. Clearly $\mathfrak{A} \supset \mathfrak{A}_X$. The equations $F = 0$ for all $F \in \mathfrak{A}$ define the closed set $Y \subset X$.

It follows from the Nullstellensatz that Y is the empty set if and only if $\mathfrak{a}_Y = k[X]$. The ideal $\mathfrak{a}_Y \subset k[X]$ can alternatively be described as the set of all functions $f \in k[X]$ that vanish at all points of the subset Y.

In particular, each point $x \in X$ is a closed subset, and hence defines an ideal $\mathfrak{m}_x \subset k[X]$. By definition this ideal is the kernel of the homomorphism $k[X] \to k$ that takes a function $f \in k[X]$ to its value $f(x)$ at x. Since $k[X]/\mathfrak{m}_x = k$ is a field, the ideal $\mathfrak{m}_x$ is maximal. Conversely, every maximal ideal $\mathfrak{m} \subset k[X]$ corresponds in this way to some point $x \in X$. Indeed, it defines a closed subset $Y \subset X$; for any point $y \in Y$ we have $\mathfrak{m}_y \supset \mathfrak{m}$, and then $\mathfrak{m}_y = \mathfrak{m}$ since $\mathfrak{m}$ is maximal. For $u \in k[X]$ the set of points $x \in X$ at which $u(x) = 0$ is closed; it is denoted by $V(u)$, and called a *hypersurface* in X.

2.3 Regular Maps

Let $X \subset \mathbb{A}^n$ and $Y \subset \mathbb{A}^m$ be closed subsets.

Definition A map $f : X \to Y$ is *regular* if there exist m regular functions $f_1, \ldots, f_m$ on X such that $f(x) = (f_1(x), \ldots, f_m(x))$ for all $x \in X$.

Thus any regular map $f : X \to \mathbb{A}^m$ is given by m functions $f_1, \ldots, f_m \in k[X]$; in order to know that this maps into the closed subset $Y \subset \mathbb{A}^m$, it is obviously enough to check that $f_1, \ldots, f_m$ as elements of $k[X]$ satisfy the equations of Y, that is

$$G(f_1, \ldots, f_m) = 0 \in k[X] \quad \text{for all } G \in \mathfrak{A}_Y.$$

Example 1.11 A regular function on X is exactly the same thing as a regular map $X \to \mathbb{A}^1$.

Example 1.12 A linear map $\mathbb{A}^n \to \mathbb{A}^m$ is a regular map.

Example 1.13 The projection map $(x, y) \mapsto x$ defines a regular map of the curve defined by $xy = 1$ to $\mathbb{A}^1$.

Example 1.14 The preceding example can be generalised as follow: let $X \subset \mathbb{A}^n$ be a closed subset and F a regular function on X. Consider the subset $X' \subset X \times \mathbb{A}^1$ defined by the equation $T_{n+1}F(T_1, \dots, T_n) = 1$. The projection $\varphi(x_1, \dots, x_{n+1}) = (x_1, \dots, x_n)$ defines a regular map $\varphi\colon X' \to X$.

Example 1.15 The map $f(t) = (t^2, t^3)$ is a regular map of the line $\mathbb{A}^1$ to the curve given by $y^2 = x^3$.

Example 1.16 (The zeta function of a variety over $\mathbb{F}_p$) We give an example that is very important for number theory. Suppose that the coefficients of the equations $F_i(T)$ of a closed subset $X \subset \mathbb{A}^n$ belong to the field $\mathbb{F}_p$ with p elements, where p is a prime number.

As we said in Section 1.1, the points of X with coordinates in $\mathbb{F}_p$ correspond to solutions of the system of congruences $F_i(T) \equiv 0 \bmod p$. Consider the map $\varphi\colon \mathbb{A}^n \to \mathbb{A}^n$ defined by

$$\varphi(\alpha_1, \dots, \alpha_n) = \bigl(\alpha_1^p, \dots, \alpha_n^p\bigr).$$

This is obviously a regular map. The important thing is that it takes $X \subset \mathbb{A}^n$ to itself. Indeed, if $\alpha \in X$, that is, $F_i(\alpha) = 0$, then since $F_i(T) \in \mathbb{F}_p[T]$, it follows from properties of fields of characteristic p that $F_i(\alpha_1^p, \dots, \alpha_n^p) = (F_i(\alpha_1, \dots, \alpha_n))^p = 0$. The map $\varphi\colon X \to X$ obtained in this way is called the *Frobenius map*. Its significance is that the points of X with coordinates in $\mathbb{F}_p$ are characterised among all points of X as the fixed points of φ. Indeed, the solutions of the equation $\alpha_i^p = \alpha_i$ are exactly all the elements of $\mathbb{F}_p$.

In exactly the same way, the elements $\alpha \in \mathbb{F}_{p^r}$ of the field with p^r elements are characterised as the solutions of $\alpha^{p^r} = \alpha$, and hence the points $x \in X$ with coordinates in $\mathbb{F}_{p^r}$ are the fixed points of the map φ^r. For each r, write ν_r for the number of points $x \in X$ with coordinates in $\mathbb{F}_{p^r}$. To get a better overall view of the set of numbers ν_r, we consider the generating function

$$P_X(t) = \sum_{r=1}^{\infty} \nu_r t^r.$$

A deep general theorem asserts that this function is always a rational function of t (for a fairly elementary proof, see Koblitz [49, Chapter V]). In this way the function $P_X(t)$ gives an expression in finite terms for the infinite sequence of numbers ν_r.

The function $P_X(t)$ associated with the closed set X has some properties analogous to those of the Riemann zeta function. To express these, note that if $x \in X$ is a point whose coordinates are in $\mathbb{F}_{p^r}$ and generate this field, then X contains all the points $\varphi^i(x)$ for $i = 1, \dots, r$, and these are all distinct. We call a set $\xi = \{\varphi^i(x)\}$ of this form a *cycle*, and the number r of points of ξ the *degree* of ξ, denoted $\deg \xi$. Now we can group together all the ν_r points $x \in X$ with coordinates in $\mathbb{F}_{p^r}$ into cycles. The coordinates of any of these points generate some subfield $\mathbb{F}_{p^d} \subset \mathbb{F}_{p^r}$, and it is known that $d \mid r$ (see for example van der Waerden [76, Ex. 6.23 of Section 6.7 (Ex. 1 of Section 43)]). We get a formula

$$\nu_r = \sum_{d|r} d\mu_d,$$

where μ_d is the number of cycles of degree d, hence

$$P_X(t) = \sum_{r=1}^{\infty} \sum_{d|r} d\mu_d t^r = \sum_{d=1}^{\infty} d\mu_d \sum_{m=1}^{\infty} t^{md} = \sum_{d=1}^{\infty} \mu_d \frac{dt^d}{1-t^d}. \tag{1.18}$$

We introduce the function

$$Z_X(t) = \prod_{\xi} \frac{1}{1-t^{\deg \xi}}, \tag{1.19}$$

where the product runs over all cycles ξ. Then the formula (1.18) can obviously be rewritten as

$$P_X(t) = \frac{Z'_X(t)}{Z_X(t)} t.$$

Equation (1.19) is analogous to the Euler product for the Riemann zeta function. To emphasise this analogy we set $p^{\deg \xi} = N(\xi)$ and $t = p^{-s}$. Then (1.19) takes the form

$$Z_X(t) = \zeta_X(s) = \prod_{\xi} \frac{1}{1-N(\xi)^{-s}}.$$

This function (either $Z_X(t)$ or $\zeta_X(s)$) is called the *zeta function* of X.

We now find out how a regular map acts on the ring of regular functions on a closed set. We start with a remark concerning arbitrary maps between sets. If $f: X \to Y$ is a map from a set X to a set Y then we can associate with every function u on Y (taking values in an arbitrary set Z) a function v on X by setting $v(x) = u(f(x))$. Obviously the map $v: X \to Z$ is the composite of $f: X \to Y$ and $u: Y \to Z$. We set $v = f^*(u)$, and call it the *pullback* of u. We get in this way a map f^* from functions on Y to functions on X. Now suppose that $f\ colon X \to Y$ is a regular map; then f^* takes regular functions on Y into regular functions on X. Indeed, if u is given by a polynomial function $G(T_1, \dots, T_n)$ and f by polynomials $F_1, \dots, F_m$ then $v = f^*(u)$ is obtained simply by substituting F_i for T_i in G, so that v is given by the polynomial $G(F_1, \dots, F_m)$. Moreover, regular maps can be characterised as the maps that take regular functions into regular functions. Indeed, suppose that a map $f: X \to Y$ of closed set has the property that for any regular function u on Y the function $f^*(u)$ on X is again regular. Then this applies in particular to the functions t_i defined by the coordinates T_i on Y for $i = 1, \dots, m$; thus the functions $f^*(t_i)$ are regular on X. But this just means that f is a regular map.

We have seen that if f is regular then the pullback of functions defines a map $f^*\colon k[Y] \to k[X]$. It follows easily from the definition of f^* that it is a homomorphism of k-algebras. We show that, conversely, every algebra homomorphism $\varphi\colon k[Y] \to k[X]$ is of the form $\varphi = f^*$ for some regular map $f\colon X \to Y$. Let $t_1, \dots, t_m$ be coordinates in the ambient space $\mathbb{A}^m$ of Y, viewed as functions on Y. Obviously $t_i \in k[Y]$, and hence $\varphi(t_i) \in k[X]$. Set $\varphi(t_i) = s_i$ and consider the map f given by the formula $f(x) = (s_1(x), \dots, s_m(x))$. This is of course a regular map. We prove that $f(X) \subset Y$. Indeed, if $H \in \mathfrak{A}_Y$ then $H(t_1, \dots, t_m) = 0$ in $k[Y]$, hence also $\varphi(H) = 0$ on X. Let $x \in X$; then $H(f(x)) = \varphi(H)(x) = 0$, and therefore $f(x) \in Y$.

Definition A regular map $f\colon X \to Y$ of closed sets is an *isomorphism* if it has an inverse, that is, if there exists a regular map $g\colon Y \to X$ such that $f \circ g = 1$ and $g \circ f = 1$. In this case we say that X and Y are *isomorphic*.

An isomorphism is obviously a one-to-one correspondence. It follows from what we have said that if f is an isomorphism then $f^*\colon k[Y] \to k[X]$ is an isomorphism of algebras. It is easy to see that the converse is also true; in other words, closed sets are isomorphic if and only if their rings of regular functions are isomorphic over k.

The facts we have just proved show that $X \mapsto k[X]$ defines an equivalence of categories between closed subsets of affine spaces (with regular maps between them) and a certain subcategory of the category of commutative algebras over k (with algebra homomorphisms). What is this subcategory, that is, which algebras are of the form $k[X]$?

Theorem 1.3 *An algebra A over a field k is isomorphic to a coordinate ring $k[X]$ of some closed subset X if and only if A has no nilpotents* (*that is $f^m = 0$ implies that $f = 0$ for $f \in A$*) *and is finitely generated as an algebra over k.*

Proof These conditions are all obviously necessary. If an algebra A is generated by finitely many elements $t_1, \dots, t_n$ then $A \cong k[T_1, \dots, T_n]/\mathfrak{A}$, where $\mathfrak{A}$ is an ideal of the polynomial ring $k[T_1, \dots, T_n]$. Suppose that $\mathfrak{A} = (F_1, \dots, F_m)$, and consider the closed set $X \subset \mathbb{A}^n$ defined by the equations $F_1 = \dots = F_m = 0$; we prove that $\mathfrak{A}_X = \mathfrak{A}$, from which it will follow that $k[X] \cong k[T_1, \dots, T_n]/\mathfrak{A} \cong A$.

If $F \in \mathfrak{A}_X$ then $F^r \in \mathfrak{A}$ for some $r > 0$ by the Nullstellensatz. Since A has no nilpotents, also $F \in \mathfrak{A}$. Thus $\mathfrak{A}_X \subset \mathfrak{A}$, and since obviously $\mathfrak{A} \subset \mathfrak{A}_X$, we have $\mathfrak{A}_X = \mathfrak{A}$. The theorem is proved. □

Example 1.17 The generalised parabola, defined by the equation $y = x^k$ is isomorphic to the line, and the maps $f(x, y) = x$ and $g(t) = (t, t^k)$ define an isomorphism.

Example 1.18 The projection $f(x, y) = x$ of the hyperbola $xy = 1$ to the x-axis is not an isomorphism, since the map is not a one-to-one correspondence: the hyperbola does not contain any point (x, y) for which $f(x, y) = 0$. Compare also Exercise 4.

Example 1.19 The map $f(t) = (t^2, t^3)$ of the line to the curve defined by $y^2 = x^3$ is easily seen to be a one-to-one correspondence. However, it is not an isomorphism, since the inverse map is of the form $g(x, y) = y/x$, and the function y/x is not regular at the origin. (See Exercise 5.)

Example 1.20 Let X and $Y \subset \mathbb{A}^n$ be closed sets. Consider $X \times Y \subset \mathbb{A}^{2n}$ as in Example 1.5, and the linear subspace $\Delta \subset \mathbb{A}^{2n}$ defined by equations $t_1 = u_1, \dots, t_n = u_n$, called the *diagonal*. Consider the map that sends each point $z \in X \cap Y$ to $\varphi(z) = (z, z) \in \mathbb{A}^{2n}$, which is obviously a point of $X \times Y \cap \Delta$. It is easy to check that the map $\varphi\colon X \cap Y \to X \times Y \cap \Delta$ obtained in this way is an isomorphism from $X \cap Y$ to $X \times Y \cap \Delta$. Using this, we can always reduce the study of the intersection of two closed sets to considering the intersection of a different closed set with a linear subspace.

Example 1.21 Let X be a closed set and G a finite group of automorphisms of X. Suppose that the characteristic of the field k does not divide the order N of G. Set $A = k[X]$, and let A^G be the subalgebra of invariants of G in A, that is, $A^G = \{f \in A \mid g^*(f) = f \text{ for all} g \in G\}$. According to Proposition A.6, the algebra A^G is finitely generated over k. From Theorem 1.3 it follows that there exists a closed set Y such that $A^G \cong k[Y]$, and a regular map $\varphi\colon X \to Y$ such that $\varphi^*(k[Y]) = A^G$. This set Y is called the *quotient variety* or *quotient space* of X by the action of G, and is written X/G.

Given two points $x_1, x_2 \in X$, there exists $g \in G$ such that $x_2 = g(x_1)$ if and only if $\varphi(x_1) = \varphi(x_2)$. Indeed, if $x_2 = g(x_1)$ then $f(x_2) = f(x_1)$ for every $f \in k[X]^G = k[Y]$, and hence $\varphi(x_1) = \varphi(x_2)$. Conversely, if $x_2 \neq g(x_1)$ then we must take a function $f \in k[X]$ such that $f(g(x_2)) = 1$, $f(g(x_1)) = 0$ for all $g \in G$. Then the symmetrised function $S(f)$ (see Section 4, Appendix) is G-invariant and satisfies $S(f)(x_2) = 1$ and $S(f)(x_1) = 0$, and hence $\varphi(x_2) \neq \varphi(x_1)$. Thus X/G parametrises the orbits $\{g(x) \mid g \in G\}$ of G acting on X.

In what follows we will mainly be interested in notions and properties of closed sets invariant under isomorphism. The system of equations defining a set is clearly not a notion of this kind; two sets X and Y can be isomorphic although given by different systems of equations in different spaces $\mathbb{A}^n$. Thus it would be natural to try to give an intrinsic definition of a closed set independent of its realisation in some affine space; a definition of this kind will be given in Chapters 5–6 in connection with the notion of a scheme.

Now we determine when a homomorphism $f^*\colon k[Y] \to k[X]$ corresponding to a regular map $f\colon X \to Y$ has no kernel, that is, when f^* is an isomorphic inclusion $k[Y] \hookrightarrow k[X]$. For $u \in k[Y]$, let's see when $f^*(u) = 0$. This means that $u(f(x)) = 0$ for all $x \in X$. In other words, u vanishes at all points of the image $f(X)$ of X. The points $y \in Y$ for which $u(y) = 0$ obviously form a closed set, and hence if this contains $f(X)$, it also contains the closure $\overline{f(X)}$. Repeating the same arguments backwards, we see that $f^*(u) = 0$ if and only if u vanishes on $\overline{f(X)}$, or equivalently,

$u \in \mathfrak{a}_{\overline{f(X)}}$. It follows in particular that the kernel of f^* is zero if and only if $\overline{f(X)} = Y$, that is, $f(X)$ is dense in Y.

This is certainly the case if $f(X) = Y$, but cases with $\overline{f(X)} = Y$ but $f(X) \neq Y$ are possible (see Example 1.13).

In what follows we will be concerned mainly with algebraic varieties in projective space. But closed subsets of affine space have a geometry with a specific flavour, which is often quite nontrivial. As an example we give the following theorem due to Abhyankar and Moh:

Theorem *A curve $X \subset \mathbb{A}^2$ is isomorphic to $\mathbb{A}^1$ if and only if there exists an automorphism of $\mathbb{A}^2$ that takes X to a line.* (*Here an automorphism is an isomorphism from $\mathbb{A}^2$ to itself.*)

The group $\operatorname{Aut} \mathbb{A}^2$ of automorphisms of the plane is an extremely interesting object. Some examples of automorphisms are very simply to construct: the affine linear maps, and maps of the form

$$x' = \alpha x, \qquad y' = \beta y + f(x), \tag{1.20}$$

where $\alpha, \beta \neq 0$ are constants, and f a polynomial. It is known that the whole group $\operatorname{Aut} \mathbb{A}^2$ is generated by these automorphisms. Moreover, the expression of an element $g \in \operatorname{Aut} \mathbb{A}^2$ as a word in affine linear maps and maps of the form (1.20) is almost unique: the only relations in $\operatorname{Aut} \mathbb{A}^2$ between maps of these two classes are those expressing the fact that the two classes have a subset in common, namely maps of the form (1.20) with f a linear polynomial. In the language of abstract group theory, $\operatorname{Aut} \mathbb{A}^2$ is the free product (or amalgamation) of two subgroups, the maps of the form (1.20) and the affine maps, over their common subgroup (see Kurosh [53, Section 35, Chapter IX, Vol. II and Ex. 10]).

A famous unsolved problem related to automorphisms of $\mathbb{A}^2$ is the Jacobian conjecture. This asserts that, if the ground field k has characteristic 0, a map given by

$$x' = f(x, y), \qquad y' = g(x, y)$$

with $f, g \in k[x, y]$ is an automorphism of $\mathbb{A}^2$ if and only if the Jacobian determinant $\frac{\partial(f,g)}{\partial(x,y)}$ is a nonzero constant. At present this conjecture is proved when the degrees of f and g are not too large (the order of 100). There is a similar conjecture for the n-dimensional affine space $\mathbb{A}^n$.

2.4 *Exercises to Section 2*

1 The set $X \subset \mathbb{A}^2$ is defined by the equation $f\colon x^2 + y^2 = 1$ and $g\colon x = 1$. Find the ideal $\mathfrak{A}_X$. Is it true that $\mathfrak{A}_X = (f, g)$?

2 Let $X \subset \mathbb{A}^2$ be the algebraic plane curve defined by $y^2 = x^3$. Prove that an element of $k[X]$ can be written uniquely in the form $P(x) + Q(x)y$ with $P(x)$, $Q(x)$ polynomials.

3 Let X be the curve of Exercise 2 and $f(t) = (t^2, t^3)$ the regular map $\mathbb{A}^1 \to X$. Prove that f is not an isomorphism. [Hint: Try to construct the inverse of f as a regular map, using the result of Exercise 2.]

4 Let X be the curve defined by the equation $y^2 = x^2 + x^3$ and $f\colon \mathbb{A}^1 \to X$ the map defined by $f(t) = (t^2 - 1, t(t^2 - 1))$. Prove that the corresponding homomorphism f^* maps $k[X]$ isomorphically to the subring of the polynomial ring $k[t]$ consisting of polynomials $g(t)$ such that $g(1) = g(-1)$. (Assume that $\operatorname{char} k \neq 2$.)

5 Prove that the hyperbola defined by $xy = 1$ and the line $\mathbb{A}^1$ are not isomorphic.

6 Consider the regular map $f\colon \mathbb{A}^2 \to \mathbb{A}^2$ defined by $f(x, y) = (x, xy)$. Find the image $f(\mathbb{A}^2)$; is it open in $\mathbb{A}^2$? Is it dense? Is it closed?

7 The same question as in Exercise 6 for the map $f\colon \mathbb{A}^3 \to \mathbb{A}^3$ defined by $f(x, y, z) = (x, xy, xyz)$.

8 An isomorphism $f\colon X \to X$ of a closed set X to itself is called an *automorphism*. Prove that all automorphisms of the line $\mathbb{A}^1$ are of the form $f(x) = ax + b$ with $a \neq 0$.

9 Prove that the map $f(x, y) = (\alpha x, \beta y + P(x))$ is an automorphism of $\mathbb{A}^2$, where $\alpha, \beta \in k$ are nonzero elements, and $P(x)$ is a polynomial. Prove that maps of this type form a group B.

10 Prove that if $f(x_1, \dots, x_n) = (P_1(x_1, \dots, x_n), \dots, P_n(x_1, \dots, x_n))$ is an automorphism of $\mathbb{A}^n$ then the Jacobian $J(f) = \det|\frac{\partial P_i}{\partial x_j}| \in k$. Prove that $f \mapsto J(f)$ is a homomorphism from the group of automorphisms of $\mathbb{A}^n$ into the multiplicative group of nonzero elements of k.

11 Suppose that X consists of two points. Prove that the coordinate ring $k[X]$ is isomorphic to the direct sum of two copies of k.

12 Let $f\colon X \to Y$ be a regular map. The subset $\Gamma_f \subset X \times Y$ consisting of all points of the form $(x, f(x))$ is called the *graph* of f. Prove that (a) $\Gamma_f \subset X \times Y$ is a closed subset, and (b) Γ_f is isomorphic to X.

13 The map $p_Y\colon X \times Y \to Y$ defined by $p_Y(x, y) = y$ is called the *projection* to Y or the *second projection*. Prove that if $Z \subset X$ and $f\colon X \to Y$ is a regular map then

$f(Z) = p_Y((Z \times Y) \cap \Gamma_f)$, where Γ_f is the graph of f and $Z \times Y \subset X \times Y$ is the subset of (z, y) with $z \in Z$.

14 Prove that for any regular map $f\colon X \to Y$ there exists a regular map $g\colon X \to X \times Y$ that is an isomorphism of X with a closed subset of $X \times Y$ and such that $f = p_Y \circ g$. In other words, any map is the composite of an embedding and a projection.

15 Prove that if $X = \bigcup U_\alpha$ is any covering of a closed set X by open subsets U_α then there exists a finite number $U_{\alpha_1}, \dots, U_{\alpha_r}$ of the U_α such that $X = U_{\alpha_1} \cup \cdots \cup U_{\alpha_r}$.

16 Prove that the Frobenius map φ (Example 1.16) is a one-to-one correspondence. Is it an isomorphism, for example if $X = \mathbb{A}^1$?

17 Find the zeta function $Z_X(t)$ for $X = \mathbb{A}^n$.

18 Determine $Z_X(t)$ for X a nonsingular conic in $\mathbb{A}^2$.

3 Rational Functions

3.1 Irreducible Algebraic Subsets

In Section 1.1 we introduced the notion of an irreducible algebraic curve in the plane. Here we formulate the analogous notion in general.

Definition A closed algebraic set X is *reducible* if there exist proper closed subsets $X_1, X_2 \subsetneqq X$ such that $X = X_1 \cup X_2$. Otherwise X is *irreducible*.

Theorem 1.4 *Any closed set X is a finite union of irreducible closed sets.*

Proof Suppose that the theorem fails for a set X. Then X is reducible, $X = X_1 \cup X_1'$, and the theorem must fail either for X_1 or for X_1'. If X_1, then it is reducible, and again it is made up of closed sets one of which is reducible. In this way we construct an infinite strictly decreasing chain of closed subsets $X \supsetneqq X_1 \supsetneqq X_2 \cdots$. We prove that there cannot be such a chain. Indeed, the ideals corresponding to the X_i would form an increasing chain $\mathfrak{A}_X \subsetneqq \mathfrak{A}_{X_1} \subsetneqq \mathfrak{A}_{X_2} \subsetneqq \cdots$. But such an infinite strictly increasing chain cannot exist, since every ideal of the polynomial ring has a finite basis, and hence an increasing chain of ideals terminates. The theorem is proved. □

If $X = \bigcup X_i$ is an expression of X as a finite union of irreducible closed sets, and if $X_i \subset X_j$ for some $i \neq j$ then we can delete X_i from the expression. Repeating this several times, we arrive at a representation $X = \bigcup X_i$ in which $X_i \not\subset X_j$

for all $i \neq j$. We say that such a representation is *irredundant*, and the X_i are the *irreducible components* of X.

Theorem 1.5 *The irredundant representation of X as a finite union of irreducible closed sets is unique.*

Proof Let $X = \bigcup_i X_i = \bigcup_j Y_j$ be two irredundant representations. Then

$$X_i = X_i \cap X = X_i \cap \bigcup_j Y_j = \bigcup_j (X_i \cap Y_j).$$

Since by assumption X_i is irreducible, we have $X_i \cap Y_j = X_i$ for some j, that is, $X_i \subset Y_j$. Repeating the argument with the X_i and Y_j interchanged gives $Y_j \subset X_{i'}$ for some i'. Hence $X_i \subset Y_j \subset X_{i'}$, so that by the irredundancy of the representation, $i = i'$ and $Y_j = X_i$. The theorem is proved. □

We now restate the condition that a closed set X is irreducible in terms of its coordinate ring $k[X]$. If $X = X_1 \cup X_2$ is reducible then since $X \supsetneqq X_1$ there exists a polynomial F_1 that is 0 on X_1 but not 0 on X, and a similar polynomial F_2 for X_2. Then the product $F_1 F_2$ is 0 on both X_1 and X_2, hence on X. The corresponding regular functions $f_1, f_2 \in k[X]$ have the property that $f_1, f_2 \neq 0$, but $f_1 f_2 = 0$. In other words, f_1 and f_2 are zerodivisors in $k[X]$. Conversely, suppose that $k[X]$ has zerodivisors $f_1, f_2 \neq 0$, with $f_1 f_2 = 0$. Write X_1, X_2 for the closed subsets of X corresponding to the ideals (f_1) and (f_2) of $k[X]$. In other words, X_i consists of the points $x \in X$ such that $f_i(x) = 0$, for $i = 1$ or 2. Obviously both $X_i \subsetneqq X$, since $f_i \neq 0$ on X, and $X = X_1 \cup X_2$ since $f_1 f_2 = 0$ on X, so that at each point $x \in X$ either $f_1(x) = 0$ or $f_2(x) = 0$. Therefore, a closed set X is irreducible if and only if its coordinate ring $k[X]$ has no zerodivisors. This in turn is equivalent to $\mathfrak{A}_X$ being a prime ideal.

If a closed subset Y is contained in X then obviously so are its irreducible components. In terms of the ring $k[X]$ the irreducibility of a closed subset $Y \subset X$ is reflected in $\mathfrak{a}_Y \subset k[X]$ being a prime ideal.

A hypersurface $X \subset \mathbb{A}^n$ with equation $f = 0$ is irreducible if and only if the polynomial f is irreducible. Thus our terminology is compatible with that used in Section 1 in the case of plane curves.

Theorem 1.6 *A product of irreducible closed sets is irreducible.*

Proof Suppose that X and Y are irreducible, but $X \times Y = Z_1 \cup Z_2$, with $Z_i \subsetneqq X \times Y$ for $i = 1, 2$. For any point $x \in X$, the closed set $x \times Y$, consisting of points (x, y) with $y \in Y$, is isomorphic to Y, and is therefore irreducible. Since

$$x \times Y = \bigl((x \times Y) \cap Z_1\bigr) \cup \bigl((x \times Y) \cap Z_2\bigr),$$

either $x \times Y \subset Z_1$ or $x \times Y \subset Z_2$. Consider the subset $X_1 \subset X$ consisting of points $x \in X$ such that $x \times Y \subset Z_1$; we now prove that X_1 is a closed set. Indeed, for any point $y \in Y$, the set X_y of points $x \in X$ such that $x \times y \in Z_1$ is closed: it is characterised by $(X \times y) \cap Z_1 = X_y \times y$, and the left-hand side is closed as an intersection of closed sets; now $X_1 = \bigcap_{y \in Y} X_y$ is closed. In the same way, the set X_2 consisting of all points $x \in X$ such that $x \times Y \subset Z_2$ is also closed. We see that $X_1 \cup X_2 = X$, and since X is irreducible it follows from this that $X_1 = X$ or $X_2 = X$. In the first case $X \times Y = Z_1$, and in the second $X \times Y = Z_2$. This contradiction proves the theorem. □

3.2 Rational Functions

It is known that any ring without zerodivisors can be embedded into a field, its field of fractions.

Definition If a closed set X is irreducible then the field of fractions of the coordinate ring $k[X]$ is the *function field* or *field of rational functions* of X; it is denoted by $k(X)$.

Recalling the definition of the field of fractions, we can say that the function field $k(X)$ consists of rational functions $F(T)/G(T)$ such that $G(T) \notin \mathfrak{A}_X$, and $F/G = F_1/G_1$ if $FG_1 - F_1G \in \mathfrak{A}_X$. This means that the field $k(X)$ can be constructed as follows. Consider the subring $\mathcal{O}_X \subset k(T_1, \dots, T_n)$ of rational functions $f = P/Q$ with $P, Q \in k[T]$ and $Q \notin \mathfrak{A}_X$. The functions f with $P \in \mathfrak{A}_X$ form an ideal M_X and $k(X) = \mathcal{O}_X/M_X$.

In contrast to regular functions, a rational function on a closed set X does not necessarily have well-defined values at every point of X; for example, the function $1/x$ at $x = 0$ or x/y at $(0, 0)$. We now find out when this is possible.

Definition A rational function $\varphi \in k(X)$ is *regular* at $x \in X$ if it can be written in the form $\varphi = f/g$ with $f, g \in k[X]$ and $g(x) \neq 0$. In this case we say that the element $f(x)/g(x) \in k$ is the *value* of φ at x, and denote it by $\varphi(x)$.

Theorem 1.7 *A rational function φ that is regular at all points of a closed subset X is a regular function on X.*

Proof Suppose $\varphi \in k(X)$ is regular at every point $x \in X$. This means that for every $x \in X$ there exists f_x, $g_x \in k[X]$ with $g_x(x) \neq 0$ such that $\varphi = f_x/g_x$. Consider the ideal $\mathfrak{a}$ generated by all the functions g_x for $x \in X$. This has a finite basis, so that there are a finite number of points $x_1, \dots, x_N$ such that $\mathfrak{a} = (g_{x_1}, \dots, g_{x_N})$. The functions g_{x_i} do not have a common zero $x \in X$, since then all functions in $\mathfrak{a}$ would vanish at x, but $g_x(x) \neq 0$. From the analogue of the Nullstellensatz it follows that $\mathfrak{a} = (1)$, and hence there exist functions $u_1, \dots, u_N \in k[X]$ such that $\sum_{i=1}^{N} u_i g_{x_i} = 1$. Multiplying both sides of this equality by φ and using the fact that $\varphi = f_{x_i}/g_{x_i}$, we get that $\varphi = \sum_{i=1}^{N} u_i f_{x_i}$, that is, $\varphi \in k[X]$. The theorem is proved. □

If φ is a rational function on a closed set X, the set of points at which φ is regular is nonempty and open. The first assertion follows since φ can be written $\varphi = f/g$ with $f, g \in k[X]$ and $g \neq 0$; hence $g(x) \neq 0$ for some $x \in X$, and obviously φ is regular at this point. To prove the second assertion, consider all possible representations $\varphi = f_i/g_i$. For any regular function g_i the set $Y_i \subset X$ of points $x \in X$ for which $g_i(x) = 0$ is obviously closed, and hence $U_i = X \setminus Y_i$ is open. The set U of points at which φ is regular is by definition $U = \bigcup U_i$, and is therefore open. This open set is called the *domain of definition* of φ. For any finite system $\varphi_1, \dots, \varphi_m$ of rational functions, the set of points $x \in X$ at which they are all regular is again open and nonempty. The first assertion follows since the intersection of a finite number of open sets is open, and the second from the following useful proposition: the intersection of a finite number of nonempty open sets of an irreducible closed set is nonempty. Indeed, let $U_i = X \setminus Y_i$ for $i = 1, \dots, m$ be such that $\bigcap U_i = \emptyset$. Then $Y_i \neq X$ and $\bigcup Y_i = X$; but the Y_i are closed sets, and this contradicts the irreducibility of X.

Thus for any finite set of rational functions, there is some nonempty open set on which they are all defined and can be compared. This remark is useful because a rational function $\varphi \in k(X)$ is uniquely determined if it is specified on some nonempty open subset $U \subset X$. Indeed, if $\varphi(x) = 0$ for all $x \in U$ and $\varphi \neq 0$ on X then any expression $\varphi = f/g$ with $f, g \in k[X]$ gives a representation of X as a union $X = X_1 \cup X_2$ of two closed sets, where $X_1 = X - U$ and X_2 is defined by $f = 0$. This contradicts the irreducibility of X.

3.3 Rational Maps

Let $X \subset \mathbb{A}^n$ be an irreducible closed set. A rational map $\varphi\colon X \to \mathbb{A}^m$ is a map given by an arbitrary m-tuple of rational functions $\varphi_1, \dots, \varphi_m \in k(X)$. Thus a rational map φ is not a map defined on the whole set X to the set $\mathbb{A}^m$, but it clearly defines a map of some nonempty open set $U \subset X$ to $\mathbb{A}^m$. Working with functions and maps that are not defined at all points is an essential difference between algebraic geometry and other branches of geometry, for example, topology.

We now define the notion of rational map $\varphi\colon X \to Y$ to a closed subset $Y \subset \mathbb{A}^m$.

Definition A *rational map* $\varphi\colon X \to Y \subset \mathbb{A}^m$ is an m-tuple of rational functions $\varphi_1, \dots, \varphi_m \in k(X)$ such that, for all points $x \in X$ at which all the φ_i are regular, $\varphi(x) = (\varphi_1(x), \dots, \varphi_m(x)) \in Y$; we say that φ is *regular* at such a point x, and $\varphi(x) \in Y$ is the *image* of x. The *image* of X under a rational map φ is the set of points

$$\varphi(X) = \bigl\{\varphi(x) \mid x \in X \text{ and } \varphi \text{ is regular at } x\bigr\}.$$

As we proved at the end of Section 3.2, there exists a nonempty open set $U \subset X$ on which all the rational functions φ_i are defined, hence also the rational map $\varphi = (\varphi_1, \dots, \varphi_m)$. Thus we can view rational maps as maps defined on open subsets; but we have to bear in mind that different maps may have different domains of definition. The same of course also applies to rational functions.

To check that rational functions $\varphi_1, \dots, \varphi_m \in k(X)$ define a rational map $\varphi\colon X \to Y$ we need to check that $\varphi_1, \dots, \varphi_m$, as elements of $k(X)$, satisfy all the equations of Y. Indeed, if this property holds then for any polynomial $u(T_1, \dots, T_m) \in \mathfrak{A}_Y$ the function $u(\varphi_1, \dots, \varphi_m) = 0$ on X. Then at each point x at which all the φ_i are regular, we have $u(\varphi_1(x), \dots, \varphi_m(x)) = 0$ for all $u \in \mathfrak{A}_Y$, that is, $(\varphi_1(x), \dots, \varphi_m(x)) \in Y$. Conversely, if $\varphi\colon X \to Y$ is a rational map, then for every $u \in \mathfrak{A}_Y$ the function $u(\varphi_1, \dots, \varphi_m) \in k(X)$ vanishes on some nonempty open set $U \subset X$, and so is 0 on the whole of X. It follows from this that $u(\varphi_1, \dots, \varphi_m) = 0$ in $k(X)$.

We now study how rational maps act on rational functions on a closed set. Let $\varphi\colon X \to Y$ be a rational map and assume that $\varphi(X)$ is dense in Y. Consider φ as a map $U \to \varphi(X) \subset Y$, where U is the domain of definition of φ, and construct the map φ^* on functions corresponding to it. For any function $f \in k[Y]$ the function $\varphi^*(f)$ is a rational function on X. Indeed, if $Y \subset \mathbb{A}^m$, and f is given by a polynomial $u(T_1, \dots, T_m)$, then $\varphi^*(f)$ is given by the rational function $u(\varphi_1, \dots, \varphi_m)$. Thus we have a map $\varphi^*\colon k[Y] \to k(X)$ which is obviously a ring homomorphism of the ring $k[Y]$ to the field $k(X)$. This homomorphism is even an isomorphic inclusion $k[Y] \hookrightarrow k(X)$. Indeed, if $\varphi^*(u) = 0$ for $u \in k[Y]$ then $u = 0$ on $\varphi(X)$. But if $u \neq 0$ on Y then the equality $u = 0$ defines a closed subset $V(u) \subsetneqq Y$. Then $\varphi(X) \subset V(u)$, but this contradicts the assumption that $\varphi(X)$ is dense in Y. The inclusion $\varphi^*\colon k[Y] \hookrightarrow k(X)$ can be extended in an obvious way to an isomorphic inclusion of the field of fractions $k(Y)$ into $k(X)$. Thus if $\varphi(X)$ is dense in Y, the rational map φ defines an isomorphic inclusion $\varphi^*\colon k(Y) \hookrightarrow k(X)$.

Given two rational maps $\varphi\colon X \to Y$ and $\psi\colon Y \to Z$ such that $\varphi(X)$ is dense in Y then it is easy to see that we can define a composite $\psi \circ \varphi\colon X \to Z$; if in addition $\psi(Y)$ is dense in Z then so is $(\psi \circ \varphi)(X)$. Then the inclusions of fields satisfy the relation $(\psi \circ \varphi)^* = \varphi^* \circ \psi^*$.

Definition A rational map $\varphi\colon X \to Y$ is *birational* or is a *birational equivalence* if φ has an inverse rational map $\psi\colon Y \to X$, that is, $\varphi(X)$ is dense in Y and $\psi(Y)$ in X, and $\psi \circ \varphi = 1$, $\varphi \circ \psi = 1$ (where defined). In this case we say that X and Y are *birational* or *birationally equivalent*.

Obviously if $\varphi\colon X \to Y$ is a birational map then the inclusion of fields $\varphi^*\colon k(Y) \to k(X)$ is an isomorphism. It is easy to see that the converse is also true (for algebraic plane curves this was done in Section 1.4). Thus closed sets X and Y are birational if and only if the fields $k(X)$ and $k(Y)$ are isomorphic over k.

Examples In Section 1 we treated a series of examples of birational maps between algebraic plane curves. Isomorphic closed sets are obviously birational. The regular maps in Examples 1.18–1.19, although not isomorphisms, are birational maps.

A closed set that is birational to an affine space $\mathbb{A}^n$ is said to be *rational*. Rational algebraic curves were discussed in Section 1. We now give some other examples of rational closed sets.

Example 1.22 An irreducible quadric $X \subset \mathbb{A}^n$ defined by a quadratic equation $F(T_1, \dots, T_n) = 0$ is rational. The proof given in Section 1.2 for the case $n = 2$ works in general. The corresponding map can once again be interpreted as the projection of X from some point $x \in X$ to a hyperplane $L \subset \mathbb{A}^n$ not passing through x (*stereographic projection*). We need only choose x so that it is not a vertex of X, that is, so that $\partial F/\partial T_i(x) \neq 0$ for at least one value of $i = 1, \dots, n$.

Example 1.23 Consider the hypersurface $X \subset \mathbb{A}^3$ defined by the 3rd degree equation $x^3 + y^3 + z^3 = 1$. We suppose that the characteristic of the ground field k is different from 3. The surface X contains several lines, for example the two skew lines L_1 and L_2 defined by

$$L_1\colon x + y = 0, z = 1, \quad \text{and} \quad L_2\colon x + \varepsilon y = 0, \quad z = \varepsilon,$$

where $\varepsilon \neq 1$ is a cube root of 1.

We give a geometric description of a rational map of X to the plane, and leave the reader to write out the formulas, and also to check that it is birational. Choose some plane $E \subset \mathbb{A}^3$ not containing L_1 or L_2. For $x \in X \setminus (L_1 \cup L_2)$, it is easy to verify that there is a unique line L passing through x and intersecting L_1 and L_2. Write $f(x)$ for the point of intersection $L \cap E$; then $x \mapsto f(x)$ is the required rational map $X \to E$.

This argument obviously applies to any cubic surface in $\mathbb{A}^3$ containing two skew lines.

In algebraic geometry we work with two different equivalence relations between closed sets, isomorphism and birational equivalence. Birational equivalence is clearly a coarser equivalence relation than isomorphism; in other words, two closed sets can be birational without being isomorphic. Thus it often turns out that the classification of closed sets up to birational equivalence is simpler and more transparent than the classification up to isomorphism. Since it is defined at every point, isomorphism is closer to geometric notions such as homeomorphism and diffeomorphism, and so more convenient. Understanding the relation between these two equivalence relations is an important problem; the question is to understand how much coarser birational equivalence is compared to isomorphism, or in other words, how many closed sets are distinct from the point of view of isomorphism but the same from that of birational equivalence. This problem will reappear frequently later in this book.

We conclude this section by proving one result that illustrates the notion of birational equivalence.

Theorem 1.8 *Any irreducible closed set X is birational to a hypersurface of some affine space $\mathbb{A}^m$.*

Proof $k(X)$ is generated over k by a finite number of elements, for example the coordinates $t_1, \dots, t_n$ in $\mathbb{A}^n$, viewed as functions on X.

Suppose that d is the maximal number of the t_i that are algebraically independent over k. According to Proposition A.7, the field $k(X)$ can be written in the form $k(z_1, \dots, z_{d+1})$, where $z_1, \dots, z_d$ are algebraically independent over k and

$$f(z_1, \dots, z_{d+1}) = 0, \tag{1.21}$$

with f irreducible over k and $f'_{T_{d+1}} \neq 0$. The function field $k(Y)$ of the closed set Y defined by (1.21) is obviously isomorphic to $k(X)$. This means that X and Y are birational. The theorem is proved. □

Remark 1.1 According to Proposition A.7, the element z_{d+1} is separable over the field $k(z_1, \dots, z_d)$. Hence the $k(z_1, \dots, z_d) \subset k(X)$ is a finite separable field extension.

Remark 1.2 It follows from the proof of Proposition A.7 and the primitive element theorem of Galois theory that $z_1, \dots, z_{d+1}$ can be chosen as linear combinations of the original coordinates $x_1, \dots, x_n$, that is, of the form $z_i = \sum_{j=1}^n c_{ij} x_j$ for $i = 1, \dots, d+1$. The map $(x_1, \dots, x_n) \mapsto (z_1, \dots, z_{d+1})$ given by these formulas is a projection of the space $\mathbb{A}^n$ parallel to the linear subspace defined by $\sum_{j=1}^n c_{ij} x_j = 0$ for $i = 1, \dots, d+1$. This shows the geometric meaning of the birational map whose existence is established in Theorem 1.8.

3.4 Exercises to Section 3

1 Suppose that k is a field of characteristic $\neq 2$. Decompose into irreducible components the closed set $X \subset \mathbb{A}^3$ defined by $x^2 + y^2 + z^2 = 0$, $x^2 - y^2 - z^2 + 1 = 0$.

2 Prove that if X is the closed set of Exercise 4 of Section 2.4 then the elements of the field $k(X)$ can be expressed in a unique way in the form $u(x) + v(x)y$ where $u(x)$ and $v(x)$ are arbitrary rational functions of x.

3 Prove that the maps f of Exercises 3, 4 and 6 of Section 2.4 are birational.

4 Decompose into irreducible components the closed set $X \subset \mathbb{A}^3$ defined by $y^2 = xz$, $z^2 = y^3$. Prove that all its components are birational to $\mathbb{A}^1$.

5 Let $X \subset \mathbb{A}^n$ be the hypersurface defined by an equation $f_{n-1}(T_1, \dots, T_n) + f_n(T_1, \dots, T_n) = 0$, where f_{n-1} and f_n are homogeneous polynomials of degrees $n-1$ and n. (A hypersurface of this form is called a *monoid*.) Prove that if X is irreducible then it is birational to $\mathbb{A}^{n-1}$. (Compare the case of plane curves treated in Section 1.4.)

6 At what points of the circle given by $x^2 + y^2 = 1$ is the rational function $(1-y)/x$ regular?

7 At which points of the curve X defined by $y^2 = x^2 + x^3$ is the rational function $t = y/x$ regular? Prove that $y/x \notin k[X]$.

4 Quasiprojective Varieties

4.1 Closed Subsets of Projective Space

Let V be a vector space of dimension $n+1$ over the field k. The set of lines (that is, 1-dimensional vector subspaces) of V is called the *n-dimensional projective space*, and denoted by $\mathbb{P}(V)$ or $\mathbb{P}^n$. If we introduce coordinates $\xi_0, \dots, \xi_n$ in V then a point $\xi \in \mathbb{P}^n$ is given by $n+1$ elements $(\xi_0 : \cdots : \xi_n)$ of the field k, not all equal to 0; and two points $(\xi_0 : \cdots : \xi_n)$ and $(\eta_0 : \cdots : \eta_n)$ are considered to be equal in $\mathbb{P}^n$ if and only if there exists $\lambda \neq 0$ such that $\eta_i = \lambda \xi_i$ for $i = 0, \dots, n$. Any set $(\xi_0 : \cdots : \xi_n)$ defining the point ξ is called a set of *homogeneous coordinates* for ξ (compare Section 1.6).

We say that a polynomial $f(S) \in k[S_0, \dots, S_n]$ vanishes at $\xi \in \mathbb{P}^n$ if $f(\xi_0, \dots, \xi_n) = 0$ for any choice of the coordinates $(\xi_0, \dots, \xi_n)$ of ξ. Obviously, then also $f(\lambda\xi_0, \dots, \lambda\xi_n) = 0$ for all $\lambda \in k$ with $\lambda \neq 0$. Write f in the form $f = f_0 + f_1 + \cdots + f_r$, where f_i is the sum of all terms of degree i in f. Then

$$
\begin{aligned}
f(\lambda\xi_0, \dots, \lambda\xi_n) = f_0(\xi_0, \dots, \xi_n) \\
\quad + \lambda f_1(\xi_0, \dots, \xi_n) + \cdots + \lambda^r f_r(\xi_0, \dots, \xi_n).
\end{aligned}
$$

Since k is an infinite field, the equality $f(\lambda\xi_0, \dots, \lambda\xi_n) = 0$ for all $\lambda \in k$ with $\lambda \neq 0$ implies that $f_i(\lambda\xi_0, \dots, \lambda\xi_n) = 0$. Thus if f vanishes at a point ξ then all of its homogeneous components f_i also vanish at ξ.

Definition $X \subset \mathbb{P}^n$ is a *closed subset* if it consists of all points at which a finite number of polynomials with coefficients in k vanish. A closed subset defined by one homogeneous equation $F = 0$ is called a *hypersurface*, as in the affine case. The degree of the polynomial is the *degree* of the hypersurface. A hypersurface of degree 2 is called a *quadric*.

The set of all polynomials $f \in k[S_0, \dots, S_n]$ that vanish at all points $x \in X$ forms an ideal of $k[S]$, called the ideal of the closed set X, and denoted by $\mathfrak{A}_X$. By what we said above, the ideal $\mathfrak{A}_X$ has the property that whenever it contains an element f it also contains all the homogeneous components of f. An ideal with this property is said to be *homogeneous* or *graded*. Thus the ideal of a closed set X of projective space is homogeneous. It follows from this that it has a basis consisting of homogeneous polynomials: we need only start from any basis and take the system of homogeneous components of polynomials of the basis. In particular, any closed set can be defined by a system of homogeneous equations.

Thus to each closed subset $X \subset \mathbb{P}^n$ there is a corresponding homogeneous ideal $\mathfrak{A}_X \subset k[S_0, \dots, S_n]$. Conversely, any homogeneous ideal $\mathfrak{A} \subset k[S]$ defines a closed subset $X \subset \mathbb{P}^n$. That is, if $F_1, \dots, F_m$ is a homogeneous basis of $\mathfrak{A}$ then X is defined by the system of equation $F_1 = \dots = F_m = 0$. If this system of equations has no other solutions in the vector space V other than 0 then it is natural to take X to be the empty set.

Examples of Closed Subsets of Projective Space

Example 1.24 (The Grassmannian) The projective space $\mathbb{P}(V)$ parametrises the 1-dimensional vector subspaces $L^1 \subset V$ of a vector space V. The *Grassmannian* or *Grassmann variety* $\operatorname{Grass}(r, V)$ plays the same role for r-dimensional vector subspaces $L^r \subset V$. To define this, consider the rth exterior power $\bigwedge^r V$ of V, and send a basis $f_1, \dots, f_r$ of a vector subspace L into the element $f_1 \wedge \dots \wedge f_r \in \bigwedge^r V$. On passing to another basis of the same vector subspace this element is multiplied by a nonzero element $\alpha \in k$, the determinant of the matrix of the coordinate change, and hence the corresponding point of the projective space $\mathbb{P}(\bigwedge^r V)$ is uniquely determined by the subspace L. Write $P(L)$ for this point. It is easy to see that it determines the subspace L uniquely. If $\{e_i\}$ is a basis of V then $\{e_{i_1} \wedge \dots \wedge e_{i_r}\}$ is a basis of $\bigwedge^r V$ and $P(L) = \sum_{i_1 < \dots < i_r} p_{i_1 \dots i_r}(e_{i_1} \wedge \dots \wedge e_{i_r})$. The homogeneous coordinates $p_{i_1 \dots i_r}$ of $P(L)$ are called the *Plücker coordinates* of L.

Except for the trivial cases of subspaces having dimension or codimension 1, not every point $P \in \mathbb{P}(\bigwedge^r V)$ is of the form $P(L)$, or in other words, not every element $x \in \bigwedge^r V$ is of the form $f_1 \wedge \dots \wedge f_r$ with $f_i \in V$. The necessary and sufficient condition for this to hold uses the notion of *convolution*. Let $u \in V^*$ be a vector of the dual vector space. For $x \in \bigwedge^1 V = V$ the convolution $u \lrcorner x$ is an element of k, and is just the scalar product (u, x) or the value $u(x)$. For $x \in \bigwedge^0 V = k$ we set $u \lrcorner x = 0$. For any $x \in \bigwedge^r V$ the convolution $u \lrcorner x = 0$ can be extended in a unique way from $x \in \bigwedge^1 V$ if we require the property

$$u \lrcorner (x \wedge y) = (u \lrcorner x) \wedge y + (-1)^a \bigl(x \wedge (u \lrcorner y)\bigr) \quad \text{for } x \in \bigwedge^a V. \tag{1.22}$$

Here $u \lrcorner \bigwedge^r V \subset \bigwedge^{r-1} V$. The element $u \lrcorner x$ for $u \in V^*$ and $x \in \bigwedge^r V$ is called the *convolution* of u and x. Finally, for $u_1, \dots, u_s \in V^*$ the element $u_1 \lrcorner (u_2 \lrcorner \dots \lrcorner (u_s \lrcorner x) \cdots)$ depends only on x and $y = u_1 \wedge \dots \wedge u_s \in \bigwedge^s V^*$, and is denoted by $y \lrcorner x$. Here $y \lrcorner x \in \bigwedge^{r-s} V$ if $r \geq s$ and $y \lrcorner x = 0$ if $r < s$.

Necessary and sufficient conditions for $x \in \bigwedge^r V$ to be of the form $x = f_1 \wedge \dots \wedge f_r$ are given by

$$(y \lrcorner x) \wedge x = 0 \quad \text{for all } y \in \bigwedge^{r-1} V^*. \tag{1.23}$$

It is obviously enough to check the conditions (1.23) for $y = u_{i_1} \wedge \cdots \wedge u_{i_{r-1}}$, where $\{u_i\}$ is a basis of V^*; in particular, if we take $\{u_i\}$ to be the basis dual to the basis $\{e_i\}$ of V then (1.23) can be written in coordinates. They take the form

$$\sum_{t=1}^{r+1} (-1)^t p_{i_1 \dots i_{r-1} j_t} p_{j_1 \dots \widehat{j_t} \dots j_{r+1}} = 0 \tag{1.24}$$

for all sequences $i_1, \dots, i_{r-1}$ and $j_1, \dots, j_{r+1}$.

The variety defined in $\mathbb{P}(\bigwedge^r V)$ by the relations (1.23) or (1.24) is called the *Grassmannian*, and denoted by $\mathrm{Grass}(r, V)$ or $\mathrm{Grass}(r, n)$ where $n = \dim V$.

We need a method of reconstructing a vector subspace L explicitly from its Plücker coordinates $p_{i_1 \dots i_r}$ satisfying (1.24). Suppose for example that $p_{1 \dots r} \neq 0$. If $p = (p_{i_1 \dots i_r}) = \mathbb{P}(L)$ then L has a basis of the form

$$f_i = e_i + \sum_{k > r} a_{ik} e_k \quad \text{for } i = 1, \dots, r.$$

It follows easily from this that $p_{1 \dots \widehat{i} \dots rk} = (-1)^k a_{ik}$, from which we get $a_{ik} = (-1)^k p_{1 \dots \widehat{i} \dots rk}$, where we have set $p_{1 \dots r} = 1$ for convenience.

Thus the open affine sets $p_{i_1 \dots i_r} \neq 0$ of $\mathrm{Grass}(r, V)$ are all isomorphic to the affine space $\mathbb{A}^{r(n-r)}$ with coordinates a_{ik} (for $i = 1, \dots, r$ and $k = r+1, \dots, n$). We can see, for example, that in the open set $p_{1 \dots r} \neq 0$ (1.24) can be solved explicitly with the coordinates $p_{1 \dots r} \neq 0$ and $p_{1 \dots \widehat{i} \dots rk}$ as free parameters. That is, if $m \geq 2$ of the subscripts $i_1, \dots, i_r$ are $> r$ then

$$p_{i_1 \dots i_r} = \frac{F(\dots, p_{1 \dots \widehat{i} \dots rk}, \dots)}{(p_{1 \dots r})^m},$$

where F is a form of degree m in $p_{1 \dots r} \neq 0$ and $p_{1 \dots \widehat{i} \dots rk}$ with $i \leq r$ and $k > r$. A detailed treatment of Grassmannians is contained, for example, in the survey article Kleiman and Laksov [47].

The first nontrivial case of this theory is when $r = 2$. Then by (1.22)

$$(u \lrcorner x) \wedge x = \frac{1}{2}\bigl(u \lrcorner (x \wedge x)\bigr) \quad \text{for } u \in V^* \text{ and } x \in \bigwedge^2 V.$$

Hence (1.23) reduces to $u \lrcorner (x \wedge x) = 0$ for all $u \in V^*$, that is, simply

$$x \wedge x = 0. \tag{1.25}$$

Finally, when $n = 4$ we have $\dim \bigwedge^4 V = 1$, so that (1.25) reduces to a single equation in the Plücker coordinates $p_{12}, p_{13}, p_{14}, p_{23}, p_{24}, p_{34}$:

$$p_{12} p_{34} - p_{13} p_{24} + p_{14} p_{23} = 0. \tag{1.26}$$

Planes $L \subset V$ in a 4-dimensional vector space V correspond to lines $\ell \subset \mathbb{P}(V)$ in projective 3-space. In this case, coordinates in V are denoted by $x_0, x_1, x_2,$

x_3 and the Plücker coordinates p_{01}, p_{02}, p_{03}, p_{12}, p_{13}, p_{23}, and (1.26) takes the form

$$p_{01}p_{23} - p_{02}p_{13} + p_{03}p_{12} = 0. \tag{1.27}$$

This is a quadric in projective 5-space $\mathbb{P}(\bigwedge^2 V)$.

Example 1.25 (The variety of associative algebras) Let A be an associative algebra over a field k of rank n. Then after a choice of basis, A is determined by its multiplication table

$$e_i e_j = \sum c_{ij}^l e_l$$

with structure constants $c_{ij}^l \in k$. The associative condition for multiplication in A takes the form

$$\sum_l c_{ij}^l c_{lk}^m = \sum_l c_{il}^m c_{jk}^l \quad \text{for } i, j, k, m = 1, \dots, n. \tag{1.28}$$

this is again a system of quadratic equation in the structure constants c_{ij}^l. Multiplying all the basis elements e_i by a nonzero element $\alpha^{-1} \in k$ has the effect of multiplying all the c_{ij}^l by α. Thus if we discard the algebra with zero multiplication, all algebras are described by points of the closed set in the projective space $\mathbb{P}^{n^3-1}$ defined by (1.28).

To be more precise, points of this set correspond to associative multiplication laws written out in terms of a chosen basis $e_1, \dots, e_n$. The change to a different basis is given by a nondegenerate $n \times n$ matrix. Thus the set of associative algebras of rank n over a field k, up to isomorphism, is parametrised by the quotient of the set defined by (1.28) by the group of nondegenerate $n \times n$ matrixes. The extent to which this type of quotient can be identified with an algebraic variety is an extremely delicate question.

Example 1.26 (Determinantal varieties) Quadratic forms in n variables form a vector space V of dimension $\binom{n+1}{2} = (1/2)n(n+1)$. Quadrics in an $(n-1)$-dimensional projective space are parametrised by points of the projective space $\mathbb{P}(V)$. Among these, the degenerate quadrics are characterised by $\det(f) = 0$, where f is the corresponding quadratic form. This is a hypersurface $X_1 \subset \mathbb{P}(V)$. The quadrics of rank $\leq n-k$ correspond to points of a set X_k defined by setting all $(n-k+1) \times (n-k+1)$ minors of the matrix of f to 0. A set of this type is called a *determinantal variety*. Another type of determinantal variety M_k is defined in the space $\mathbb{P}(V)$, where V is the space of $n \times m$ matrixes, by the condition that a matrix has rank $\leq k$.

In the case of closed subsets of affine space, an ideal $\mathfrak{A} \subset k[T]$ defines the empty set only if $\mathfrak{A} = (1)$; this is the assertion of the Nullstellensatz. For closed subsets of projective spaces this is not the case: for example, the ideal $(S_0, \dots, S_n)$ also defines the empty set. Write I_s for the ideal of $k[S]$ consisting of polynomials having only terms of degree $\geq s$. Obviously I_s also defines the empty set: it contains, for example, the monomials S_i^s for $i = 0, \dots, n$, which have a common zero only at the origin.

Lemma 1.1 *A homogeneous ideal $\mathfrak{A} \subset k[S]$ defines the empty set if and only if it contains the ideal I_s for some $s > 0$.*

Proof We have already seen that the ideal I_s defines the empty set, and the same holds a fortiori for any ideal containing I_s. Suppose that a homogeneous ideal $\mathfrak{A} \subset k[S]$ defines the empty set. Let $F_1, \dots, F_r$ be a homogeneous basis of the ideal $\mathfrak{A}$ and set $\deg F_i = m_i$. Then from the assumption, it follows that the polynomials $F_i(1, T_1, \dots, T_n)$ have no common root, where $T_j = S_j/S_0$. Indeed, a common root $(\alpha_1, \dots, \alpha_n)$ would give a common root $(1, \alpha_1, \dots, \alpha_n)$ of $F_1, \dots, F_r$. By the Nullstellensatz there must exist polynomials $G_i(T_1, \dots, T_n)$ such that $\sum_i F_i(1, T_1, \dots, T_n) G_i(T_1, \dots, T_n) = 1$. Setting $T_j = S_j/S_0$ in this equality and multiplying through by a common denominator of the form $S_0^{l_0}$ we get $S_0^{l_0} \in \mathfrak{A}$. In the same way, for each $i = 1, \dots, n$ there exists a number $l_i > 0$ such that $S_i^{l_i} \in \mathfrak{A}$. If now $l = \max(l_0, \dots, l_n)$ and $s = (l-1)(n+1)+1$ then in any term $S_0^{a_0} \cdots S_n^{a_n}$ of degree $a_0 + \cdots + a_n \geq s$ we must have at least one term S_i with exponent $a_i \geq l \geq l_i$, and since $S_i^{l_i} \in \mathfrak{A}$, this term is contained in $\mathfrak{A}$. This proves that $I_s \subset \mathfrak{A}$. The lemma is proved. □

From now on we consider closed subsets of affine and projective spaces at one and the same time. We again call these affine and projective closed sets. For projective closed sets, we use the same terminology as for affine sets; that is, if $Y \subset X$ are two closed sets then we say that $X \setminus Y$ is an *open set* in X. As before, a union of an arbitrary number of open sets, and an intersection of finitely many open sets is again open. The set $\mathbb{A}_0^n \subset \mathbb{P}^n$ of points $\xi = (\xi_0 : \cdots : \xi_n)$ for which $\xi_0 \neq 0$ is obviously open. Its points can be put in one-to-one correspondence with the points of an n-dimensional affine space by setting $\alpha_i = \xi_i/\xi_0$ for $i = 1, \dots, n$, and sending $\xi \in \mathbb{A}_0^n$ to $(\alpha_1, \dots, \alpha_n) \in \mathbb{A}^n$. Thus we call the set $\mathbb{A}_0^n$ an *affine piece* of $\mathbb{P}^n$. In the same way, for $i = 0, \dots, n$, the set $\mathbb{A}_i^n$ consists of points for which $\xi_i \neq 0$. Obviously $\mathbb{P}^n = \bigcup_i \mathbb{A}_i^n$.

For any projective closed set $X \subset \mathbb{P}^n$, and any $i = 0, \dots, n$, the set $U_i = X \cap \mathbb{A}_i^n$ is open in X. It is closed as a subset of $\mathbb{A}_i^n$. Indeed, if X is given by a system of homogeneous equations $F_1 = \cdots = F_m = 0$ and $\deg F_j = n_j$ then, for example, U_0 is given by the system

$$S_0^{-n_j} F_j = F_j(1, T_1, \dots, T_n) = 0 \quad \text{for } j = 1, \dots, m,$$

where $T_i = S_i/S_0$ for $i = 1, \dots, n$. We call U_i the *affine pieces* of X; obviously $X = \bigcup U_i$. A closed subset $U \subset \mathbb{A}_0^n$ defines a closed projective set $\overline{U}$ called its *projective completion*; $\overline{U}$ is the intersection of all projective closed sets containing U. It is easy to check that the homogeneous equations of $\overline{U}$ are obtained by a process inverse to that just described. If $F(T_1, \dots, T_n)$ is any polynomial in the ideal $\mathfrak{A}$ of U of degree $\deg F = k$, then the equations of $\overline{U}$ are of the form $S_0^k F(S_1/S_0, \dots, S_n/S_0)$. It follows from this that

$$U = \overline{U} \cap \mathbb{A}_0^n. \tag{1.29}$$

Up to now we have considered two classes of objects that could claim to be called algebraic varieties; affine and projective closed sets. It is natural to try to introduce a unified notion of which both of these types will be particular cases. This will be done most systematically in Chapters 5–6 in connection with the notion of scheme. For the moment we introduce a more particular notion, that unifies projective and affine closed sets.

Definition A *quasiprojective variety* is an open subset of a closed projective set.

A closed projective set is obviously a quasiprojective variety. For affine closed sets this follows from (1.29). A *closed subset* of a quasiprojective variety is its intersection with a closed set of projective space. Open set and neighbourhood of a point are defined similarly. The notion of irreducible variety and the theorem on decomposing a variety as a union of irreducible components carries over word-for-word from the case of affine sets.

From now on we use *subvariety* Y of a quasiprojective variety $X \subset \mathbb{P}^n$ to mean any subset $Y \subset X$ which is itself a quasiprojective variety in $\mathbb{P}^n$. This is obviously equivalent to saying that $Y = Z \setminus Z_1$ with Z and $Z_1 \subset X$ closed subsets.

4.2 Regular Functions

We proceed to considering functions on quasiprojective varieties, and start with the projective space $\mathbb{P}^n$ itself. Here we meet an important distinction between functions of homogeneous and inhomogeneous coordinates: a rational function of the homogeneous coordinates

$$f(S_0, \dots, S_n) = \frac{P(S_0, \dots, S_n)}{Q(S_0, \dots, S_n)} \tag{1.30}$$

cannot be viewed as a function of $x \in \mathbb{P}^n$, even when $Q(x) \neq 0$, since the value $f(\alpha_0, \dots, \alpha_n)$ in general changes when all the α_i are multiplied through by a common factor. However, when f is a homogeneous function of degree 0, that is, when P and Q are homogeneous of the same degree, then f can be viewed as a function of $x \in \mathbb{P}^n$.

If $X \subset \mathbb{P}^n$ is a quasiprojective variety, $x \in X$ and $f = P/Q$ is a homogeneous function of degree 0 with $Q(x) \neq 0$, then f defines a function on a neighbourhood of x in X with values in k. We say that f is *regular* in a neighbourhood of x, or simply at x. A function on X that is regular at all points $x \in X$ is a *regular function* on X. All regular functions on X form a ring, that we denote by $k[X]$.

Let's prove that for a closed subset X of an affine space, our definition of regular function here is the same as that in Section 2.2 (after an obvious passage to inhomogeneous coordinates). For X irreducible, this is stated in Theorem 1.7. In the general case we only need to change slightly the arguments used to prove this theorem. In this proof we let f be a regular function in the affine sense of Section 2.2.

By assumption, each point $x \in X$ has a neighbourhood U_x with $q_x \neq 0$ on U_x in which $f = p_x/q_x$, where p_x, q_x are regular functions on X and $q_x \neq 0$ on U_x. Hence

$$q_x f = p_x \tag{1.31}$$

on U_x. But we can assume that (1.31) holds over the whole of X. To achieve this, we multiply both p_x and q_x by a regular function equal to 0 on $X \setminus U_x$ and nonzero at x; then (1.31) holds also on $X \setminus U_x$, since both sides are 0 there. As in the proof of Theorem 1.7, we can find points $x_1, \dots, x_N \in X$ and regular functions $h_1, \dots, h_N$ such that $\sum_{i=1}^N q_{x_i} h_i = 1$. Multiply (1.31) for $x = x_i$ by h_i and add, to get

$$f = \sum_{i=1}^N p_{x_i} h_i,$$

that is, f is a regular function.

In contrast to the case of closed affine sets, the ring $k[X]$ may consist only of constants. We will prove later (Theorem 1.11, Corollary 1.1) that this is always the case if X is an irreducible closed projective set. This is easy to prove directly if $X = \mathbb{P}^n$: indeed, if $f = P/Q$, with P and Q forms of the same degree, we can assume that P and Q have no common factors; then f is not regular at points x where $Q(x) = 0$. On the other hand, when X is only quasiprojective, $k[X]$ may turn out to be an unexpectedly large ring. If X is an affine closed set then as we have seen $k[X]$ is finitely generated as an algebra over k. However, Rees and Nagata constructed examples of quasiprojective varieties for which $k[X]$ is not finitely generated. This shows that $k[X]$ is only a reasonable invariant when X is an affine closed set.

We pass to maps. Any map of a quasiprojective variety X to an affine space $\mathbb{A}^n$ is given by n functions on X with values in k. If these functions are regular then we say the map is *regular*.

Definition Let $f\colon X \to Y$ be a map between quasiprojective varieties, with $Y \subset \mathbb{P}^m$. This map is *regular* if for every point $x \in X$ and for some affine piece $\mathbb{A}_i^m$ containing $f(x)$ there exists a neighbourhood $U \ni x$ such that $f(U) \subset \mathbb{A}_i^m$ and the map $f\colon U \to \mathbb{A}_i^m$ is regular.

We check that the regularity property is independent of the choice of affine piece $\mathbb{A}_i^m$ containing $f(x)$. If $f(x) = (y_0, \dots, \widehat{1}, \dots, y_m) \in \mathbb{A}_i^m$ (where $\widehat{1}$ in the ith place means that this coordinate is discarded) is also contained in $\mathbb{A}_j^m$, then $y_j \neq 0$, and the coordinates of this point in $\mathbb{A}_j^m$ are $(y_0/y_j, \dots, 1/y_j, \dots, \widehat{1}, \dots y_m/y_j)$, with $1/y_j$

in the ith place and $\widehat{1}$ discarded from the jth. Therefore if $f\colon U \to \mathbb{A}^m_i$ is given by functions $(f_0, \dots, \widehat{1}, \dots, f_m)$, the map f to $\mathbb{A}^m_j$ is given by

$$(f_0/f_j, \dots, 1/f_j, \dots, \widehat{1}, \dots f_m/f_j).$$

By assumption $f_j(x) \neq 0$, and the subset $U' \subset U$ of points at which $f_j \neq 0$ is open. The functions $f_0/f_j, \dots, 1/f_j, \dots, f_m/f_j$ are regular on U', and hence $f\colon U' \to \mathbb{A}^m_j$ is regular.

In the same way as for affine closed sets, a regular map $f\colon X \to Y$ defines a homomorphism $f^*\colon k[Y] \to k[X]$.

The question of how to write down formulas defining a regular map on an irreducible variety is solved in complete analogy with the case $n = 2$ treated in Section 1.6. Suppose for example that $f(x) \in \mathbb{A}^m_0$, and the map $f\colon U \to \mathbb{A}^m_0$ is given by regular functions $f_1, \dots, f_m$. By definition $f_i = P_i/Q_i$ where P_i, Q_i are forms of the same degree in the homogeneous coordinates of x and $Q_i(x) \neq 0$. Putting these fractions over a common denominator gives $f_i = F_i/F_0$, where $F_0, \dots, F_m$ are forms of the same degree and $F_0(x) \neq 0$. In other words, $f(x) = (F_0(x) : \cdots : F_m(x)) \in \mathbb{P}^m$. In this process, we must bear in mind that the representation of a regular function as a ratio of two forms is not unique. Hence two different formulas

$$f(x) = \bigl(F_0(x) : \cdots : F_m(x)\bigr) \quad \text{and} \quad g(x) = \bigl(G_0(x) : \cdots : G_m(x)\bigr) \tag{1.32}$$

may define the same map; this happens if and only if

$$F_i G_j = F_j G_i \quad \text{on } X \quad \text{for } 0 \le i, j \le m. \tag{1.33}$$

This brings us to a second form of the definition of a regular map:

Definition A *regular map* $f\colon X \to \mathbb{P}^m$ of an irreducible quasiprojective variety X to projective space $\mathbb{P}^m$ is given by an $(m + 1)$-tuple of forms

$$(F_0 : \cdots : F_m) \tag{1.34}$$

of the same degree in the homogeneous coordinates of $x \in \mathbb{P}^n$. We require that for every $x \in X$ there exists an expression (1.34) for f such that $F_i(x) \neq 0$ for at least one i; then we write $f(x)$ to denote the point $(F_0(x) : \cdots : F_m(x))$. Two maps (1.32) are considered equal if (1.33) holds.

Now we have a definition of regular maps between quasiprojective varieties, it is natural to define an *isomorphism* to be a regular map having an inverse regular map.

A quasiprojective variety X' isomorphic to a closed subset of an affine space will be called an *affine variety*. It can happen that X is given as a subset $X \subset \mathbb{A}^n$, but is not closed in $\mathbb{A}^n$. For example, the set $X = \mathbb{A}^1 \setminus 0$ is not closed in $\mathbb{A}^1$, although it is quasiprojective, and is isomorphic to the hyperbola $xy = 1$ (Example 1.13), which is a closed set of $\mathbb{A}^2$. Thus the notion of a closed affine set is not invariant under isomorphism, while that of affine variety is invariant by definition.

In the same way, a quasiprojective variety isomorphic to a closed projective set will be called a *projective variety*. We will prove in Theorem 1.10 that if $X \subset \mathbb{P}^n$ is a projective variety then it is closed in $\mathbb{P}^n$, so that the notions of closed projective set and projective variety coincide and are both invariant under isomorphism.

There are quasiprojective varieties that are neither affine nor projective (see Exercise 5 of Section 4.5 and Exercises 4–6 of Section 5.5).

In what follows, we will meet some properties of varieties X that need only be verified for some neighbourhood U of any point $x \in X$. In other words, if $X = \bigcup U_\alpha$, with U_α any open sets, then it is enough to verify the property for each of the U_α. We say that properties of this type are *local properties*. We give some example of local properties.

Lemma 1.2 *The property that a subset $Y \subset X$ is closed in a quasiprojective variety X is a local property.*

Proof The assertion means that if $X = \bigcup U_\alpha$ with open sets U_α, and $Y \cap U_\alpha$ is closed in each U_α then Y is closed in X. By definition of open sets, $U_\alpha = X \setminus Z_\alpha$ where the Z_α are closed, and by definition of closed sets, $U_\alpha \cap Y = U_\alpha \cap T_\alpha$ where the $T_\alpha \subset X$ are closed.

We check that $Y = \bigcap (Z_\alpha \cup T_\alpha)$, from which it follows of course that Y is closed. If $y \in Y$ and $y \in U_\alpha$ then $y \in U_\alpha \cap Y \subset T_\alpha$, and if $y \notin U_\alpha$ then $y \in X \setminus U_\alpha = Z_\alpha$, so that $y \in Z_\alpha \cup T_\alpha$ for every α. Conversely, suppose that $x \in Z_\alpha \cup T_\alpha$ for every α. Since $X = \bigcup U_\alpha$ it follows that $x \in U_\beta$ for some β. Then $x \notin Z_\beta$, and hence $x \in T_\beta$, so that $x \in T_\beta \cap U_\beta \subset Y$. The lemma is proved. □

In studying local properties we can restrict ourselves to affine varieties in view of the following result.

Lemma 1.3 *Every point $x \in X$ has a neighbourhood isomorphic to an affine variety.*

Proof By assumption $X \subset \mathbb{P}^n$. If $x \in \mathbb{A}^n_0$ (that is, if the coordinate u_0 of x is nonzero) then $x \in X \cap \mathbb{A}^n_0$, and by definition of a quasiprojective variety $X \cap \mathbb{A}^n_0 = Y \setminus Y_1$ where Y and $Y_1 \subset Y$ are closed subsets of $\mathbb{A}^n_0$. Since $x \in Y \setminus Y_1$, there exists a polynomial F of the coordinates of $\mathbb{A}^n_0$ such that $F = 0$ on Y_1 and $F(x) \neq 0$. Write $V(F)$ for the set of points of Y where $F = 0$. Obviously $D(F) = Y \setminus V(F)$ is a neighbourhood of x. We prove that this neighbourhood is isomorphic to an affine variety. Suppose that $G_1 = \cdots = G_m = 0$ are the equations of Y in $\mathbb{A}^n_0$. Define a variety $Z \subset \mathbb{A}^{n+1}$ by the equations

$$\begin{gathered} G_1(T_1, \ldots, T_n) = \cdots = G_m(T_1, \ldots, T_n) = 0, \\ F(T_1, \ldots, T_n) \cdot T_{n+1} = 1. \end{gathered} \tag{1.35}$$

The map $\varphi \colon (x_1, \ldots, x_{n+1}) \mapsto (x_1, \ldots, x_n)$ obviously defines a regular map $Z \to D(F)$ and $\psi \colon (x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_n, F(x_1, \ldots, x_n)^{-1})$ a regular map $D(F) \to Z$ inverse to φ. This proves the lemma. □

If $Y = \mathbb{A}^1$, $F = T$ then the isomorphism just constructed is the map considered in Example 1.13.

Definition An open set $D(f) = X \setminus V(f)$ consisting of the points of an affine variety X such that $f(x) \neq 0$ is called a *principal open set*.

The significance of these sets is that they are affine, as we have seen, and the ring $k[D(f)]$ of regular function on them can be easily determined. Namely, by construction $f \neq 0$ on $D(f)$, so that $f^{-1} \in k[D(f)]$, and Theorem 1.7 together with (1.35) shows that $k[D(f)] = k[X][f^{-1}]$.

Lemmas 1.2–1.3 show for example that closed subsets map to closed subsets under isomorphisms. We prove in addition that the inverse image $f^{-1}(Z)$ under any regular map $f\colon X \to Y$ of any closed subset $Z \subset Y$ is closed in X.

By definition of a regular map $f colon X \to Y$, for any point $x \in X$ there are neighbourhoods U of x in X and V of $f(x)$ in Y such that $f(U) \subset V \subset \mathbb{A}^m$ and the map $f\colon U \to V$ is regular. By Lemma 1.3 we can assume that U is an affine variety. By Lemma 1.2, it is enough to check that $f^{-1}(Z) \cap U = f^{-1}(Z \cap V)$ is closed in U. Since $Z \cap V$ is closed in V, it is defined by equations $g_1 = \cdots = g_m = 0$, where the g_i are regular functions on V. But then $f^{-1}(Z \cap V)$ is defined by the equations $f^*(g_1) = \cdots = f^*(g_m) = 0$, and is hence also closed.

It follows also from what we have just proved that the inverse image of an open set is again open. It is easy to check that a regular map can be defined as a map $f\colon X \to Y$ such that the inverse image of any open set is open (that is, f is "continuous"), and for any point $x \in X$ and any function φ regular in a neighbourhood of $f(x) \in Y$, the function $f^*(\varphi)$ is regular in a neighbourhood of x.

4.3 Rational Functions

In discussing the definition of rational functions on quasiprojective varieties, we met a distinction of substance between the case of affine varieties and the general case. Namely, we defined rational functions on an affine variety X as ratios of functions that are regular on the whole of X. But in the general case, as we have said, it can happen that there are no everywhere regular functions except for the constants, so that if we used the same definition there would also be no rational functions except for the constants. For this reason we define rational functions on a quasiprojective variety $X \subset \mathbb{P}^n$ to be functions defined on X by homogeneous functions on $\mathbb{P}^n$ (as in Section 1.6 for $n = 2$).

More precisely, consider an irreducible quasiprojective variety $X \subset \mathbb{P}^n$ and (by analogy with Section 3.2) write $\mathcal{O}_X$ for the set of rational functions $f = P/Q$ in the homogeneous coordinates $S_0, \ldots, S_n$ such that P, Q are forms of the same degree and $Q \notin \mathfrak{A}_X$. As for affine varieties, from the fact that X is irreducible it follows that $\mathcal{O}_X$ is a ring. Write M_X for the set of functions $f \in \mathcal{O}_X$ with $P \in \mathfrak{A}_X$. Obviously the quotient ring $\mathcal{O}_X / M_X$ is a field, called the *function field* of X, and denoted by $k(X)$.

If U is an open subset of an irreducible quasiprojective variety X then, since a form vanishes on X if and only if it vanishes on U, we have $k(X) = k(U)$. In particular, $k(X) = k(\overline{X})$, where $\overline{X}$ is the projective closure of X in $\mathbb{P}^n$. Thus in discussing function fields we can restrict to affine or projective varieties if we want to.

It is easy to check that if X is an affine variety then the definition just given coincides with that given in Section 3.2. Indeed, dividing the numerator and denominator of a rational function $f = P/Q$ with $\deg P = \deg Q = m$ by S_0^m, we can write it as a rational function in $T_i = S_i/S_0$ for $i = 1, \dots, n$. By doing this, we establish an isomorphism of the field of homogeneous rational functions of degree 0 in $S_0, \dots, S_n$ with the field $k(T_1, \dots, T_n)$. An obvious verification shows that the subring and ideal of $k(T_1, \dots, T_n)$ denoted in Section 3.2 by $\mathcal{O}_X$ and M_X correspond to the objects denoted here by the same letters.

In Section 4.2 we have already used rational functions on $\mathbb{P}^n$ to define regular functions. As there, we say that $f \in k(X)$ is *regular* at a point $x \in X$ if it can be written in the form $f = F/G$, with F and G homogeneous of the same degree and $G(x) \neq 0$. Then $f(x) = F(x)/G(x)$ is the *value* of f at x. As in the case of affine varieties, the set of points at which a given rational function f is regular is a nonempty open set U of X, called the *domain of definition* of f. Obviously a rational function can also be defined as a function regular on some open set $U \subset X$.

A rational map $f \colon X \to \mathbb{P}^m$ is defined (as in the second definition of regular map in Section 4.2) by giving $m + 1$ forms $(F_0 : \cdots : F_m)$ of the same degree in the $n + 1$ homogeneous coordinates of $\mathbb{P}^n$ containing X. Here at least one of the forms must not vanish on X. Two maps $(F_0 : \cdots : F_m)$ and $(G_0 : \cdots : G_m)$ are equal if $F_i G_j = F_j G_i$ on X for all i, j. If we divide through all the forms F_i by one of them (nonzero on X), we can define a rational map by $m + 1$ rational functions on X, with the same notion of equality of maps. If a rational map f can be defined by functions $(f_0 : \cdots : f_m)$ such that all the f_i are regular at $x \in X$ and not all zero at x, then f is regular at x. It then defines a regular map of some neighbourhood of the point x to $\mathbb{P}^m$.

The set of points at which a rational map is regular is open. Hence we can also define a rational map to be a regular map of some open set $U \subset X$. If $Y \subset \mathbb{P}^m$ is a quasiprojective variety and $f \colon X \to \mathbb{P}^m$ a rational map, we say that f maps X to Y if there exists an open set $U \subset X$ on which f is regular and $f(U) \subset Y$. The union $\widetilde{U}$ of all such open sets is called the *domain of definition* of f, and $f(\widetilde{U}) \subset Y$ the *image* of X in Y.

As in the case of affine varieties, if the image of a rational map $f \colon X \to Y$ is dense in Y then f defines an inclusion of fields $f^* \colon k(Y) \hookrightarrow k(X)$. If a rational map $f \colon X \to Y$ has an inverse rational map then f is *birational* or is a *birational equivalence*, and X and Y are birational. In this case the inclusion of fields $f^* \colon k(Y) \hookrightarrow k(X)$ is an isomorphism.

We can now clarify the relation between the notions of isomorphism and birational equivalence.

Proposition 1.1 *Two irreducible varieties X and Y are birational if and only if they contain isomorphic open subsets $U \subset X$ and $V \subset Y$.*

Proof Indeed, suppose that $f: X \to Y$ is birational, and let $g = f^{-1}: Y \to X$ be the inverse rational map. Write $U_1 \subset X$ and $V_1 \subset Y$ for the domain of definition of f and g. Then by assumption $f(U_1)$ is dense in Y, so that $f^{-1}(V_1) \cap U_1$ is nonempty, and as proved in Section 4.2, is open. Set $U = f^{-1}(V_1) \cap U_1$ and $V = g^{-1}(U_1) \cap V_1$. A simple check shows that $f(U) = V$, $g(V) = U$ and $fg = 1$, $gf = 1$, that is, U and V are isomorphic. □

4.4 Examples of Regular Maps

Example 1.27 (Projection) Let E be a d-dimensional linear subspace of $\mathbb{P}^n$ defined by $n - d$ linearly independent linear equations $L_1 = \cdots = L_{n-d} = 0$, with L_i linear forms. The *projection* with centre E is the rational map $\pi(x) = (L_1(x) : \cdots : L_{n-d}(x))$. This map is regular on $\mathbb{P}^n \setminus E$, since at every point of this set one of the forms L_i does not vanish. Hence if X is any closed subvariety of $\mathbb{P}^n$ disjoint from E, the restriction of π defines a regular map $\pi: X \to \mathbb{P}^{n-d-1}$. The geometric meaning of projection is as follows: as a model of $\mathbb{P}^{n-d-1}$ take any $(n - d - 1)$-dimensional linear subspace $H \subset \mathbb{P}^n$ disjoint from E. Then there is a unique $(d + 1)$-dimensional linear subspace $\langle E, x\rangle$ passing through E and any point $x \in \mathbb{P}^n \setminus E$. This subspace intersects H in a unique point, which is $\pi(x)$. If X intersects E, but is not contained in it, then projection from E is a rational map on X. The case $d = 0$, a projection from a point, has already appeared several times.

Example 1.28 (The Veronese embedding) Consider all the homogeneous polynomials F of degree m in variables $S_0, \dots, S_n$. These form a vector space, whose dimension is easy to compute: it is the binomial coefficient $\binom{n+m}{m}$.

Consider the hypersurfaces of degree m in $\mathbb{P}^n$. Since polynomials define the same hypersurface if and only if they are proportional, hypersurfaces correspond to points of the projective space $\mathbb{P}^N$ of dimension $N = \nu_{n,m} = \binom{n+m}{m} - 1$. Write $v_{i_0 \dots i_n}$ for homogeneous coordinates of $\mathbb{P}^N$, where $i_0, \dots, i_n \geq 0$ are any nonnegative integers such that $i_0 + \cdots + i_n = m$. Consider the map $v_m: \mathbb{P}^n \to \mathbb{P}^N$ defined by

$$v_{i_0 \dots i_n} = u_0^{i_0} \cdots u_n^{i_n} \quad \text{for } i_0 + \cdots + i_n = m. \tag{1.36}$$

This is obviously a regular map, since the monomials on the right-hand side of (1.36) include in particular the elements u_i^m, which vanish only if all $u_i = 0$. The map v_m is called the mth *Veronese embedding* of $\mathbb{P}^n$, and the image $v_m(\mathbb{P}^n) \subset \mathbb{P}^N$ the *Veronese variety*. It follows from (1.36) that the relations

$$v_{i_0 \dots i_n} v_{j_0 \dots j_n} = v_{k_0 \dots k_n} v_{l_0 \dots l_n} \tag{1.37}$$

hold on $v_m(\mathbb{P}^n)$ whenever $i_0 + j_0 = k_0 + l_0, \dots, i_n + j_n = k_n + l_n$. Conversely, it's easy to deduce from (1.37) that at least one of the coordinates $v_{0 \dots m \dots 0}$ corresponding to the monomial u_i^m is nonzero, and that, for example, on the open set $v_{m0\dots0} \neq 0$, the map

$$u_0 = v_{m0\dots0}, \qquad u_i = v_{m-1,0\dots1\dots0} \quad \text{for } i \geq 1$$

is a regular inverse of v_m. Hence $v_m(\mathbb{P}^n)$ is defined by (1.37), and v_m is an isomorphic embedding $\mathbb{P}^n \hookrightarrow \mathbb{P}^N$.

The significance of the Veronese embedding is that if

$$F = \sum a_{i_0\dots i_n} u_0^{i_0} \cdots u_n^{i_n}$$

is a form of degree m in the homogeneous coordinates of $\mathbb{P}^n$ and $H \subset \mathbb{P}^n$ is the hypersurface defined by $F = 0$, then $v_m(H) \subset v_m(\mathbb{P}^n) \subset \mathbb{P}^N$ is the intersection of $v_m(\mathbb{P}^n)$ with the hyperplane of $\mathbb{P}^N$ with equation $\sum a_{i_0\dots i_n} v_{i_0\dots i_n}$. Thus the Veronese embedding allows us to reduce the study of some problems concerning hypersurfaces of degree m to the case of hyperplanes.

The mth Veronese image of the projective line $v_m(\mathbb{P}^1) \subset \mathbb{P}^m$ is called the Veronese curve, the twisted m-ic curve, or the *rational normal curve* of degree m.

4.5 Exercises to Section 4

1 Prove that an affine variety U is irreducible if and only if its projective closure $\overline{U}$ is irreducible.

2 Associate with any affine variety $U \subset \mathbb{A}_0^n$ its projective closure $\overline{U}$ in $\mathbb{P}^n$. Prove that this defines a one-to-one correspondence between the affine subvarieties of $\mathbb{A}_0^n$ and the projective subvarieties of $\mathbb{P}^n$ with no components contained in the hyperplane $S_0 = 0$.

3 Prove that the variety $X = \mathbb{A}^2 \setminus (0,0)$ is not isomorphic to an affine variety. [Hint: Compute the ring $k[X]$ of regular functions on X, and use the fact that if Y is an affine variety, every proper ideal $\mathfrak{A} \subsetneqq k[Y]$ defines a nonempty set.]

4 Prove that any quasiprojective variety is open in its projective closure.

5 Prove that every rational map $\varphi\colon \mathbb{P}^1 \to \mathbb{P}^n$ is regular.

6 Prove that any regular map $\varphi\colon \mathbb{P}^1 \to \mathbb{A}^n$ maps $\mathbb{P}^1$ to a point.

7 Define a birational map f from an irreducible quadric hypersurface $X \subset \mathbb{P}^3$ to the plane $\mathbb{P}^2$ by analogy with the stereographic projection of Example 1.22. At which points is f not regular? At which points is f^{-1} not regular?

8 In Exercise 7, find the open subsets $U \subset X$ and $V \subset \mathbb{P}^2$ that are isomorphic.

9 Prove that the map $y_0 = x_1x_2$, $y_1 = x_0x_2$, $y_2 = x_0x_1$ defines a birational map of $\mathbb{P}^2$ to itself. At which points are f and f^{-1} not regular? What are the open sets mapped isomorphically by f? (Compare Section 3.5, Chapter 4.)

10 Prove that the Veronese image $v_m(\mathbb{P}^n) \subset \mathbb{P}^N$ is not contained in any linear subspace of $\mathbb{P}^N$.

11 Prove that the variety $\mathbb{P}^2 \setminus X$, where X is a plane conic, is affine. [Hint: Use the Veronese embedding.]

5 Products and Maps of Quasiprojective Varieties

5.1 Products

The definition of the product of affine varieties (Example 1.5) was so natural as not to require any comment. For general quasiprojective varieties, things are somewhat more complicated. Because of this, we first consider quasiprojective subvarieties of affine spaces. If $X \subset \mathbb{A}^n$ and $Y \subset \mathbb{A}^m$ are varieties of this type then $X \times Y = \{(x, y) \mid x \in X, y \in Y\}$ is a quasiprojective variety in $\mathbb{A}^n \times \mathbb{A}^m$. Indeed, if $X = X_1 \setminus X_0$ and $Y = Y_1 \setminus Y_0$ where $X_1, X_0 \subset \mathbb{A}^n$, and $Y_1, Y_0 \subset \mathbb{A}^m$ are closed subvarieties, then writing

$$X \times Y = X_1 \times Y_1 \setminus \big((X_1 \times Y_0) \cup (X_0 \times Y_1)\big)$$

shows that $X \times Y$ is quasiprojective. This quasiprojective variety is the *product* of X and Y. At this point, we should check that if X and Y are replaced by isomorphic varieties then so is $X \times Y$. This is easy to see. Suppose that $\varphi\colon X \to X' \subset \mathbb{A}^p$ and $\psi\colon Y \to Y' \subset \mathbb{A}^q$ are isomorphisms. Then $\varphi \times \psi\colon X \times Y \to X' \times Y'$ defined by $(\varphi \times \psi)(x, y) = (\varphi(x), \psi(y))$ is a regular map, with regular inverse $\varphi^{-1} \times \psi^{-1}$.

We return to quasiprojective varieties, and decide what properties we want the notion of product to have. Let $X \subset \mathbb{P}^n$ and $Y \subset \mathbb{P}^m$ be two quasiprojective varieties. Write $X \times Y$ for the set of pairs (x, y) with $x \in X$ and $y \in Y$. We want to consider this set as a quasiprojective variety, and for this, we have to produce an embedding φ of $X \times Y$ into a projective space $\mathbb{P}^N$ in such a way that the image $\varphi(X \times Y) \subset \mathbb{P}^N$ is a quasiprojective subvariety. At the same time, it is reasonable to require that the definition is local, in the sense that for any points $x \in X$ and $y \in Y$ there exist affine neighbourhoods $X \supset U \ni x$ and $Y \supset V \ni y$ such that $\varphi(U \times V)$ is open in $\varphi(X \times Y)$, and φ defines an isomorphism of the product of the affine varieties U and V, whose definition we already know, to the subvariety $\varphi(U \times V) \subset \varphi(X \times Y)$.

It is easy to see that the local property of φ determines it uniquely; more precisely, if $\psi\colon X \times Y \to \mathbb{P}^M$ is another embedding of the same kind, then $\psi \circ \varphi^{-1}$ defines an isomorphism between $\varphi(X \times Y)$ and $\psi(X \times Y)$. Indeed, for this, it is enough to prove that for any $x \in X$ and $y \in Y$, there exist neighbourhoods $\varphi(X \times Y) \supset W_1 \ni \varphi(x, y)$ and $\psi(X \times Y) \supset W_2 \ni \psi(x, y)$ such that $\psi \circ \varphi^{-1}\colon W_1 \to W_2$ is an isomorphism. Consider affine neighbourhoods $X \supset U \ni x$ and $Y \supset V \ni y$ the existence of which is provided by the local property; passing if necessary to smaller affine neighbourhoods, we can assume that $U \times V$ is isomorphic to both $\varphi(U \times V)$ and $\psi(U \times V)$. Then $\varphi(U \times V) = W_1$ and $\psi(U \times V) = W_2$ are the affine neighbourhoods we need, since both are isomorphic to the product $U \times V$ of the affine varieties U and V.

We now proceed to construct an embedding φ with the required properties. For this, we can at once restrict to the case $X = \mathbb{P}^n$, $Y = \mathbb{P}^m$; for once an embedding $\varphi\colon \mathbb{P}^n \times \mathbb{P}^m \hookrightarrow \mathbb{P}^N$ is constructed, it is easy to check that its restriction to $X \times Y \subset \mathbb{P}^n \times \mathbb{P}^m$ has all the required properties.

To construct the embedding φ, consider the projective space $\mathbb{P}^N$ with homogeneous coordinates w_{ij} having two subscripts $i = 0, \dots, n$ and $j = 0, \dots, m$; thus $N = (n+1)(m+1) - 1$. If $x = (u_0 : \cdots : u_n) \in \mathbb{P}^n$ and $y = (v_0 : \cdots : v_m) \in \mathbb{P}^m$ then we set

$$\varphi(x, y) = (w_{ij}), \quad \text{with } w_{ij} = u_i v_j \quad \text{for } 0 \le i \le n \text{ and } 0 \le j \le m. \tag{1.38}$$

Multiplying the homogeneous coordinates of x or y by a common scalar obviously does not change the point $\varphi(x, y) \in \mathbb{P}^N$. To prove that $\varphi(\mathbb{P}^n \times \mathbb{P}^m)$ is a closed set of $\mathbb{P}^N$, we write out its defining equations:

$$w_{ij} w_{kl} = w_{kj} w_{il} \quad \text{for } 0 \le i, k \le n \text{ and } 0 \le j, l \le m. \tag{1.39}$$

Substituting the w_{ij} given by (1.38) shows at once that they satisfy (1.39). Conversely, if w_{ij} satisfy (1.39), and, say, $w_{00} \ne 0$, then setting $k, l = 0$ in (1.39) gives that $(w_{ij}) = \varphi(x, y)$, where

$$x = (w_{00} : \cdots : w_{n0}) \quad \text{and} \quad y = (w_{00} : \cdots : w_{0m}).$$

This argument proves at the same time that $\varphi(x, y)$ determines x and y uniquely, that is, φ is an embedding $\mathbb{P}^n \times \mathbb{P}^m \hookrightarrow \mathbb{P}^N$ with image the subvariety $W \subset \mathbb{P}^N$ defined by (1.39). Consider the open sets $\mathbb{A}^n_0 \subset \mathbb{P}^n$ given by $u_0 \ne 0$, $\mathbb{A}^m_0 \subset \mathbb{P}^m$ by $v_0 \ne 0$, and $\mathbb{A}^N_{00} \subset \mathbb{P}^N$ by $w_{00} \ne 0$, having inhomogeneous coordinates $x_i = u_i/u_0$, $y_j = v_j/v_0$ and $z_{ij} = w_{ij}/w_{00}$ respectively. Then obviously $\varphi(\mathbb{A}^n_0 \times \mathbb{A}^m_0) = W \cap \mathbb{A}^N_{00} = W_{00}$. As we have just seen, on W_{00} we have $z_{i0} = x_i$, $z_{0j} = y_j$ and $z_{ij} = x_i y_j = z_{i0} z_{0j}$ for $i, j > 0$. It follows from this that $\varphi(\mathbb{P}^n \times \mathbb{P}^m) \cap \mathbb{A}^N_{00} = W_{00}$ is isomorphic to $\mathbb{A}^{n+m}$ with coordinates $(x_1, \dots, x_n, y_1, \dots, y_m)$, and φ defines an isomorphism $\mathbb{A}^n_0 \times \mathbb{A}^m_0 \to W_{00}$. This proves that φ satisfies the local requirement of our construction. The embedding $\varphi\colon \mathbb{P}^n \times \mathbb{P}^m \hookrightarrow \mathbb{P}^N$ with $N = (n+1)(m+1) - 1$ just constructed is called the *Segre embedding*, and the image $\mathbb{P}^n \times \mathbb{P}^m \subset \mathbb{P}^N$ the Segre variety.

Remark 1.3 The point (w_{ij}) can be interpreted as an $(n+1) \times (m+1)$ matrix, and (1.39) express the vanishing of the 2×2 minors:

$$\det \begin{vmatrix} w_{ij} & w_{il} \\ w_{kj} & w_{kl} \end{vmatrix} = 0.$$

That is, they express the condition that the matrix (w_{ij}) has rank 1, and (1.38) shows that such a matrix is a product of a $1 \times (n+1)$ column matrix and a $(m+1) \times 1$ row matrix. Thus $\varphi(\mathbb{P}^n \times \mathbb{P}^m)$ is a determinantal variety (see Example 1.26).

Remark 1.4 The simplest case $n = m = 1$ has a simple geometric interpretation: in this case, (1.39) is the single equation $w_{11}w_{00} = w_{01}w_{10}$, so that $\varphi(\mathbb{P}^1 \times \mathbb{P}^1)$ is just a nondegenerate quadric surface $Q \subset \mathbb{P}^3$. For $\alpha = (\alpha_0, \alpha_1) \in \mathbb{P}^1$, the set $\varphi(\alpha \times \mathbb{P}^1)$ is the line in $\mathbb{P}^3$ given by $\alpha_1 w_{00} = \alpha_0 w_{10}$, $\alpha_1 w_{01} = \alpha_0 w_{11}$. As α runs through $\mathbb{P}^1$, these lines give all the generators of one of the two families of lines of Q. Similarly the set $\varphi(\mathbb{P}^1 \times \beta)$ is a line of $\mathbb{P}^3$, and as β runs through $\mathbb{P}^1$, these lines give the generators of the other family.

It is convenient, now that we have defined the product $X \times Y$ of quasiprojective varieties using the embedding $\varphi \colon \mathbb{P}^n \times \mathbb{P}^m \hookrightarrow \mathbb{P}^N$, with $N = (n+1)(m+1) - 1$, to explain some ideas of algebraic geometry that are originally defined in terms of $\mathbb{P}^n \times \mathbb{P}^m$ and of this embedding.

Let us determine, for example, what are the subsets of $\mathbb{P}^n \times \mathbb{P}^m$ that are mapped by φ to algebraic subvarieties of $\mathbb{P}^N$; these will then be the closed algebraic subvarieties of the product $\mathbb{P}^n \times \mathbb{P}^m$. A subvariety $V \subset \mathbb{P}^N$ is defined by equations $F_k(w_{00} : \cdots : w_{nm}) = 0$, where the F_k are homogeneous polynomials in the w_{ij}. After making the substitution (1.38), we can write these in the coordinates u_i and v_j as equations

$$G_k(u_0 : \cdots : u_n; v_0 : \cdots : v_m) = 0,$$

where the G_k are homogeneous in each set of variables $u_0, \dots, u_n$ and $v_0, \dots, v_m$, and of the same degree in both. Conversely, it is easy to see that a polynomial with this bihomogeneity property can always be written as a polynomial in the products $u_i v_j$. However, equations that are bihomogeneous in u_i and v_j always define an algebraic subvariety of $\mathbb{P}^n \times \mathbb{P}^m$ even if the degrees of homogeneity in the two sets of variables are different. For if $G(u_0 : \cdots : u_n; v_0 : \cdots : v_m)$ has degree r in u_i and s in v_j, and, say, $r > s$, then $G = 0$ is equivalent to the system of equations $v_i^{r-s} G = 0$ for $i = 0, \dots, m$, and we know that these define an algebraic variety.

In what follows, we also need to answer the same question for the product $\mathbb{P}^n \times \mathbb{A}^m$. Suppose that $\mathbb{A}^m = \mathbb{A}_0^m \subset \mathbb{P}^m$ is given by $v_0 \neq 0$. The equations of a closed subset of $\mathbb{P}^n \times \mathbb{P}^m$ are $G_k(u_0 : \cdots : u_n; v_0 : \cdots : v_m) = 0$. Suppose that G_k is homogeneous of degree r_k in $v_0 : \cdots : v_m$. Dividing the equation by $v_0^{r_k}$ and setting $y_j = v_j / v_0$ gives equations $g_k(u_0 : \cdots : u_n; y_1 : \cdots : y_m) = 0$ that are homogeneous in the u_i, and (in general) inhomogeneous in the y_j. This proves the following result:

Theorem 1.9 *A subset $X \subset \mathbb{P}^n \times \mathbb{P}^m$ is a closed algebraic subvariety if and only if it is given by a system of equations*

$$G_k(u_0 : \cdots : u_n; v_0 : \cdots : v_m) = 0 \quad \textit{for } k = 1, \dots, t,$$

homogeneous separately in each set of variables u_i and v_j. Every closed algebraic subvariety of $\mathbb{P}^n \times \mathbb{A}^m$ is given by a system of equations

$$g_k(u_0 : \cdots : u_n; y_1 : \cdots : y_m) = 0 \quad \text{for } k = 1, \ldots, t \tag{1.40}$$

that are homogeneous in $(u_0, \ldots, u_n)$.

Of course, the same kind of thing holds for a product of any number of spaces. For example, a subvariety of $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_k}$ is given by a system of equations homogeneous in each of the k sets of variables.

5.2 The Image of a Projective Variety is Closed

The image of an affine variety under a regular map does not have to be a closed set; this is illustrated in Examples 1.13–1.14 for a map from an affine variety to an affine variety. For maps from an affine variety to a projective variety it is even more obvious: an example is given by the embedding of $\mathbb{A}^n$ into $\mathbb{P}^n$ as the open subset $\mathbb{A}^n_0$. In this respect, projective varieties are fundamentally different from affine varieties.

Theorem 1.10 *The image of a projective variety under a regular map is closed.*

The proof uses a notion that will occur later. Let $f : X \to Y$ be a regular map between arbitrary quasiprojective varieties. The subset Γ_f of $X \times Y$ consisting of pairs $(x, f(x))$ is called the *graph* of f.

Lemma 1.4 *The graph of a regular map is closed in $X \times Y$.*

Proof First of all, it is enough to assume that Y is projective space. Indeed, if $Y \subset \mathbb{P}^m$ then $X \times Y \subset X \times \mathbb{P}^m$, and f defines a map $\overline{f} : X \to \mathbb{P}^m$ with $\Gamma_f = \Gamma_{\overline{f}} \subset X \times Y \subset X \times \mathbb{P}^m$. Thus set $Y = \mathbb{P}^m$. Let ι be the identity map from $\mathbb{P}^m$ to itself. Consider the regular map $(f, \iota) : X \times \mathbb{P}^m \to \mathbb{P}^m \times \mathbb{P}^m$ given by $(f, \iota)(x, y) = (f(x), y)$. Obviously Γ_f is the inverse image under the regular map (f, ι) of the graph Γ_ι of ι. We proved in Section 4.2 that the inverse image of a closed set under a regular map is closed. Hence everything reduces to proving that $\Gamma_\iota \subset \mathbb{P}^m \times \mathbb{P}^m$ is closed. But Γ_ι consists of points $(x, y) \in \mathbb{P}^m \times \mathbb{P}^m$ such that $x = y$. If $x = (u_0 : \cdots : u_m)$ and $y = (v_0 : \cdots : v_m)$ then the condition is that $(u_0 : \cdots : u_m)$ and $(v_0 : \cdots : v_m)$ are proportional; this condition can be expressed $u_i v_j = u_j v_i$, that is, $w_{ij} = w_{ji}$ for $i, j = 0, \ldots, m$. This proves that Γ_ι is closed, and therefore the lemma. □

We return to the proof of the theorem. Let Γ_f be the graph of f, and $p : X \times Y \to Y$ the second projection, defined by $p(x, y) = y$. Obviously $f(X) = p(\Gamma_f)$. In view of Lemma 1.4, Theorem 1.10 follows from the following more general assertion.

Theorem 1.11 *If X is a projective variety, and Y a quasiprojective variety, the second projection $p\colon X \times Y \to Y$ takes closed sets to closed sets.*

Proof The proof of this theorem can be reduced to a simple particular case. First of all, if $X \subset \mathbb{P}^n$ is a closed subset then the theorem for X follows from the theorem for $\mathbb{P}^n$: for $X \times Y$ is closed in $\mathbb{P}^n \times Y$, so that if Z is closed in $X \times Y$, it is also closed in $\mathbb{P}^n \times Y$. Thus we can assume that $X = \mathbb{P}^n$. Secondly, since closed is a local property, it is enough to cover Y by affine open sets U_i and prove the theorem for each of these. Hence we can assume that Y is an affine variety. Finally if $Y \subset \mathbb{A}^m$ then $\mathbb{P}^n \times Y$ is closed in $\mathbb{P}^n \times \mathbb{A}^m$, and hence it is enough to prove the theorem in the particular case $X = \mathbb{P}^n$ and $Y = \mathbb{A}^m$.

Let's see what the theorem means in this case. According to Theorem 1.9, any closed subvariety $Z \subset \mathbb{P}^n \times \mathbb{A}^m$ is defined by (1.40), that we write in the form $g_i(u; y) = 0$ for $i = 1, \dots, t$. Write $p\colon Z \to \mathbb{A}^m$ for the restriction of the second projection. Obviously the inverse image $p^{-1}(y_0)$ of $y_0 \in \mathbb{A}^m$ consists of all nonzero solutions of the system $g_i(u, y_0) = 0$, and hence $y_0 \in p(Z)$ if and only if the system of equations $g_i(u; y_0) = 0$ has a nonzero solution in $(u_0, \dots, u_n)$. Thus Theorem 1.11 asserts that for any system of (1.40), the subset T of $y_0 \in \mathbb{A}^m$ for which $g_i(u; y_0) = 0$ has a nonzero solution is closed.

Now in view of Lemma 1.1, $g_i(u; y_0) = 0$ has a nonzero solution if and only if

$$\bigl(g_1(u, y_0), \dots, g_t(u, y_0)\bigr) \not\supset I_s \quad \text{for all } s = 1, 2, \dots.$$

We now show that for given $s \geq 1$, the set of points $y_0 \in \mathbb{A}^m$ for which $(g_1(u, y_0), \dots, g_t(u, y_0)) \not\supset I_s$ is a closed set T_s. Then $T = \bigcap T_s$, and T is also closed. Write k_i for the degree of the homogeneous polynomial $g_i(u, y)$ in the variables $u_0, \dots, u_n$. Let $\{M^\alpha\}_\alpha$ be the monomials of degree s in $u_0, \dots, u_n$ written out in some order. The condition $(g_1(u, y_0), \dots, g_t(u, y_0)) \supset I_s$ means that each monomial M^α can be expressed in the form

$$M^\alpha = \sum_{i=1}^{t} g_i(u, y_0) F_{i,\alpha}(u). \tag{1.41}$$

Comparing the homogeneous components of degree s shows that there must also be an expression (1.41) for M^α with $\deg F_{i,\alpha} = s - k_i$, or $F_{i,\alpha} = 0$ if $k_i > s$. Let $\{N_i^\beta\}_\beta$ be the monomials of degree $s - k_i$ written out in some order. We see that the conditions (1.41) hold if and only if every monomial M^α is a linear combination of the polynomials $g_i(u, y_0)N_i^\beta$. This, in turn, is equivalent to the condition that the polynomials $g_i(u, y_0)N_i^\beta$ span the entire vector space S of homogeneous polynomials of degree s in $u_0, \dots, u_n$. Conversely, $(g_1(u, y_0), \dots, g_t(u, y_0)) \not\supset I_s$ means that $g_i(u, y_0)N_i^\beta$ do not span S. To turn this condition into equations for T_s, write out the coefficients of the M^α appearing in all the polynomials $g_i(u, y_0)N_i^\beta$ as a rectangular matrix $\{a_\alpha\beta\}$, and set to zero all of its $\sigma \times \sigma$ minors, where $\sigma = \dim S$. These minors are obviously polynomials in the coefficients of the polynomials $g_i(u, y_0)$, and are therefore polynomials in the coordinates of the point y_0; they give the equations of the set T_s. Theorem 1.11 is proved, and with it Theorem 1.10. $\square$

Remark One sees from the proof that Theorem 1.10 generalises to a wider class of maps $f: X \to Y$ between quasiprojective varieties, namely those that factor as a composite of a *closed embedding* $\iota: X \hookrightarrow \mathbb{P}^n \times Y$ (that is, an isomorphism of X with a closed subvariety) and the projection $p: \mathbb{P}^n \times Y \to Y$. Such maps are said to be *proper*. For example, if $f: X \to Y$ is a regular map of projective varieties then the restriction $f: f^{-1}(U) \to U$ to an open subset $U \subset Y$ is proper. Obviously if $f: X \to Y$ is a proper map the inverse image $f^{-1}(y)$ of a point $y \in Y$ is a projective variety.

Corollary 1.1 *If φ is a regular function on an irreducible projective variety then $\varphi \in k$, that is, φ is constant.*

Proof We can view φ as a map $f: X \to \mathbb{A}^1$, and hence as a map $\overline{f}: X \to \mathbb{P}^1$. Since φ is a regular function, f is a regular map, and hence so is $\overline{f}$; by Theorem 1.10 its image $f(X) \subset \mathbb{P}^1$ is closed. But since f itself is regular, $f(X) = \overline{f}(X)$, and therefore $\overline{f}(X)$ is a closed subset of $\mathbb{P}^1$ and is contained in $\mathbb{A}^1$, that is, it does not contain the point at infinity $x_\infty \in \mathbb{P}^1$. It follows from this that either $f(X) = \mathbb{A}^1$ or $f(X)$ is a finite set $S \subset \mathbb{A}^1$ (see Example 1.3). The first case is impossible, since $f(X)$ is also supposed to be closed in $\mathbb{P}^1$, and $\mathbb{A}^1$ is not. Hence $f(X) = S$. If S consists of finitely many points $\alpha_1, \dots, \alpha_t$ then $X = \bigcup f^{-1}(\alpha_i)$, and $t > 1$ would contradict the irreducibility of X. Hence S consists of one point only, and so φ is constant. The corollary is proved. $\square$

Corollary 1.1 and Theorem 1.7 provide an example of affine and projective varieties having diametrically opposite properties. On an affine variety there is a host of regular functions (they make up the whole coordinate ring $k[X]$), but on an irreducible projective variety, only the constants. The next result is a second example of affine and projective varieties being opposites.

Corollary 1.2 *A regular map $f: X \to Y$ from an irreducible projective variety X to an affine variety Y maps X to a point.*

Proof Suppose that $Y \subset \mathbb{A}^m$. Then f is given by m functions $f(x) = (\varphi_1(x), \dots, \varphi_m(x))$. Each of the functions φ_i is constant by Corollary 1.1, that is $\varphi_i = \alpha_i \in k$. Hence $f(X) = (\alpha_1, \dots, \alpha_m)$. The corollary is proved. $\square$

We give another example of an application of Theorem 1.10. For this, we use the representation of forms of degree m in $n+1$ variables by points of the projective space $\mathbb{P}^N$ with $N = \nu_{n,m} = \binom{m+n}{m} - 1$, as in Example 1.28.

Proposition *Points $\xi \in \mathbb{P}^N$ corresponding to reducible homogeneous polynomials F form a closed set.*

Remark 1.5 The proposition asserts that the condition for a homogeneous polynomial to be reducible can be written as polynomial conditions on its coefficients. For curves of degree 2, that is, the case $m = n = 2$, this relation is well known from coordinate geometry: if $F = \sum_{i=0}^{2} a_{ij} U_i U_j$ then F is irreducible if and only if $\det |a_{ij}| \neq 0$.

Remark 1.6 Passing to inhomogeneous coordinates, we see that in the vector space of all polynomial of degree $\leq m$, the reducible polynomials together with the polynomials of degree $<m$ form a closed set.

Proof Proceeding to the proof of the proposition, we write $X \subset \mathbb{P}^N$ for the set of points ξ corresponding to reducible polynomials, and X_k for the set of points corresponding to polynomials F that split as a product of two polynomials of degrees k and $m-k$ (for $k = 1, \dots, m$). Obviously $X = \bigcup X_k$, and we need only prove that each X_k is closed.

Consider the projective space $\mathbb{P}^{\nu_{n,k}}$ and $\mathbb{P}^{\nu_{n,m-k}}$ of forms of degree k and $m-k$, where $\nu_{n,k} = \binom{n+k}{k} - 1$ is as in Example 1.28. Multiplying polynomials of degree k and $m-k$ defines a map $f \colon \mathbb{P}^{\nu_{n,k}} \times \mathbb{P}^{\nu_{n,m-k}} \to \mathbb{P}^N$, and it is easy to see that f is regular. Obviously $X_k = f(\mathbb{P}^{\nu_{n,k}} \times \mathbb{P}^{\nu_{n,m-k}})$. We saw in Section 5.1 that the product of two projective spaces is a projective variety, and hence X_k closed follows by Theorem 1.10. The proposition is proved. $\square$

5.3 *Finite Maps*

The projection map introduced in Section 4.4 has an important property; in order to state this, we first recall some notions from algebra. Let B be a ring, and A a subring containing the identity element 1_B. We say that an element $b \in B$ is *integral* over A if it satisfies an equation

$$b^k + a_1 b^{k-1} + \cdots + a_k = 0 \quad \text{with } a_i \in A.$$

B is integral over A if every element $b \in B$ is integral over A. It is easy to prove (see for example Atiyah and Macdonald [8, Proposition 5.1 and Corollary 5.2 of Chapter 5]) that a ring B that is finitely generated as an A-algebra is integral over A if and only if it is finite as a module over A.

Let X and Y be affine varieties and $f \colon X \to Y$ a regular map such that $f(X)$ is dense in Y. Then f^* defines an isomorphic inclusion $k[Y] \hookrightarrow k[X]$. We view $k[Y]$ as a subring of $k[X]$ by means of f^*.

Definition 1.1 f is a *finite map* if $k[X]$ is integral over $k[Y]$.

From the properties of integral rings recalled above it follows that the composite of two finite maps is again finite. A typical example of a map that is not finite is Example 1.13.

Example 1.29 Let X be an affine algebraic variety, G a finite group of automorphisms of X and $Y = X/G$ the quotient space (see Example 1.21). Then the map $\varphi\colon X \to Y$ is finite. Indeed, the proof of Proposition A.6 shows that the generators u_i of the algebra $k[X]$ are integral over the algebra $k[X]^G = k[Y]$. It follows from this that $k[X]$ is integral over $k[Y]$.

If f is a finite map then any point $y \in Y$ has at most a finite number of inverse images. Indeed, suppose that $X \subset \mathbb{A}^n$ and let $t_1, \dots, t_n$ be the coordinates of $\mathbb{A}^n$ viewed as functions on X. It is enough to prove that any coordinate t_i takes only a finite number of values on the set $f^{-1}(y)$. By definition t_i satisfies an equation $t_i^k + a_1 t_i^{k-1} + \cdots + a_k = 0$ with $a_i \in k[Y]$. For $y \in Y$ and $x \in f^{-1}(y)$, we get an equation

$$t_i(x)^k + a_1(y) t_i(x)^{k-1} + \cdots + a_k(y) = 0, \tag{1.42}$$

which has only a finite number of roots.

The meaning of the finite condition is that as y moves in Y, none of the roots of (1.42) tends to infinity, since the coefficient 1 of the leading term does not vanish on Y. Thus as y moves in Y, points of $f^{-1}(y)$ can merge together, but cannot disappear. We make this remark more precise in the following result.

Theorem 1.12 *A finite map is surjective.*

Proof Let X and Y be affine varieties, $f\colon X \to Y$ a finite map, and $y \in Y$. Write $\mathfrak{m}_y$ for the ideal of $k[Y]$ consisting of functions that take the value 0 at y. If $t_1, \dots, t_n$ are the coordinate functions on Y and $y = (\alpha_1, \dots, \alpha_n)$ then $\mathfrak{m}_y = (t_1 - \alpha_1, \dots, t_n - \alpha_n)$. The equations of the variety $f^{-1}(y)$ then have the form $f^*(t_1) = \alpha_1, \dots, f^*(t_n) = \alpha_n$, and by the Nullstellensatz $f^{-1}(y) = \emptyset$ if and only if the elements $f^*(t_i) - \alpha_i$ generate the trivial ideal:

$$\bigl(f^*(t_1) - \alpha_1, \dots, f^*(t_n) - \alpha_n\bigr) = k[X].$$

From now on we view $k[Y]$ as a subring of $k[X]$, and do not distinguish between a function $u \in k[Y]$ and $f^*(u) \in k[X]$. Then the above condition is of the form $(t_1 - \alpha_1, \dots, t_n - \alpha_n) = k[X]$, that is, $\mathfrak{m}_y k[X] = k[X]$. Since $k[X]$ is integral over $k[Y]$ it follows that it is a finite $k[Y]$-module; Theorem 1.12 follows from this and the following purely algebraic assertion:

Lemma *If a ring B is a finite A-module where $A \subset B$ is a subring containing 1_B, then for an ideal $\mathfrak{a}$ of A,*

$$\mathfrak{a} \subsetneqq A \implies \mathfrak{a}B \subsetneqq B.$$

See Proposition A.11, Corollary A.1 for the proof.

This completes the proof of Theorem 1.12. $\square$

Corollary *A finite map takes closed sets to closed sets.*

Proof It is enough to check this for an irreducible closed set $Z \subset X$. We apply Theorem 1.12 to the restriction of f to Z, that is $\overline{f} \colon Z \to \overline{f(Z)}$. This is clearly a finite map between affine varieties, hence $f(Z) = \overline{f(Z)}$ by Theorem 1.12, that is, $f(Z)$ is closed. The corollary is proved. $\square$

Finiteness is a local property:

Theorem 1.13 *If $f \colon X \to Y$ is a regular map of affine varieties, and every point $x \in Y$ has an affine neighbourhood $U \ni x$ such that $V = f^{-1}(U)$ is affine and $f \colon V \to U$ is finite, then f itself is finite.*

Proof Set $k[X] = B$, $k[Y] = A$. Principal open sets were defined in Section 4.2. We can take a neighbourhood U of any point of Y such that U is a principal open set and satisfies the assumption of the theorem (see Exercise 11 of Section 5.5). Let $D(g_\alpha)$ be a family of such open sets, which we can take to be finite. Then $Y = \bigcup D(g_\alpha)$, that is, the ideal generated by the g_α is the whole of A. In our case $V_\alpha = f^{-1}(D(g_\alpha)) = D(f^*(g_\alpha))$ and $k[D(g_\alpha)] = A[1/g_\alpha]$, $k[V_\alpha] = B[1/g_\alpha]$. By assumption $B[1/g_\alpha]$ has a finite basis $\omega_{i,\alpha}$ over $A[1/g_\alpha]$. We can assume that $\omega_{i,\alpha} \in B$, since if the basis consisted of elements $\omega_{i,\alpha}/g_\alpha^{m_i}$ with $\omega_{i,\alpha} \in B$ then the elements $\omega_{i,\alpha}$ would also be a basis. We take the union of all the bases $\omega_{i,\alpha}$ and prove that they form a basis of B over A.

An element $b \in B$ has an expression

$$b = \sum_i \frac{a_{i,\alpha}}{g_\alpha^{n_\alpha}} \omega_{i,\alpha}$$

for each α. Since the $g_\alpha^{n_\alpha}$ generate the unit ideal of A, there exist $h_\alpha \in A$ such that $\sum_\alpha g_\alpha^{n_\alpha} h_\alpha = 1$. Hence

$$b = b \sum_\alpha g_\alpha^{n_\alpha} h_\alpha = \sum_i \sum_\alpha a_{i,\alpha} h_\alpha \omega_{i,\alpha},$$

which proves the theorem. $\square$

Definition 1.2 A regular map $f \colon X \to Y$ of quasiprojective varieties is *finite* if any point $y \in Y$ has an affine neighbourhood V such that the set $U = f^{-1}V$ is affine and $f \colon U \to V$ is a finite map between affine varieties.

Obviously, for a finite map f the set $f^{-1}(y)$ is finite for every $y \in Y$. It follows from Theorem 1.12 that any finite map is surjective. This property has important consequences, that relate to arbitrary maps.

Theorem 1.14 *If $f \colon X \to Y$ is a regular map and $f(X)$ is dense in Y then $f(X)$ contains an open set of Y.*

Proof The assertion of the theorem reduces at once to the case that both X and Y are irreducible and affine, and we assume this in what follows. Then $k[Y] \subset k[X]$. We write r for the transcendence degree of the field extension $k(X)/k(Y)$, and choose r elements $u_1, \dots, u_r \in k[X]$ that are algebraically independent over $k(Y)$. Then

$$k[X] \supset k[Y][u_1, \dots, u_r] \supset k[Y] \quad \text{and} \quad k[Y][u_1, \dots, u_r] = k\big[Y \times \mathbb{A}^r\big].$$

This represents f as the composite $f = g \circ h$ of two maps $h\colon X \to Y \times \mathbb{A}^r$ and $g\colon Y \times \mathbb{A}^r \to Y$, where g is simply the projection to the first factor. Any element $v \in k[X]$ is algebraic over $k[Y \times \mathbb{A}^r]$, hence there exists an element $a \in k[Y \times \mathbb{A}^r]$ such that av is integral over $k[Y \times \mathbb{A}^r]$. Let $v_1,\dots v_m$ be a system of generators of $k[X]$, and $a_1, \dots, a_m \in k[Y \times \mathbb{A}^r]$ elements such that each $a_i v_i$ is integral over $k[Y \times \mathbb{A}^r]$, and set $F = a_1 \cdots a_m$. Since all the functions a_i are invertible on the principal open set $D(F) \subset Y \times \mathbb{A}^r$, the functions v_i on $D(h^*(F)) \subset X$ are integral over $k[Y \times \mathbb{A}^r][1/F]$, that is, the restricted map

$$h\colon D\big(h^*(F)\big) \to D(F)$$

is finite. Thus $h(D(h^*(F))) = D(F)$ by Theorem 1.12, so that $D(F) \subset h(X)$. It remains to prove that $g(D(F))$ contains an open set of Y. Suppose that

$$F = F(y, T) = \sum F_\alpha(y) T^\alpha,$$

where T^α are monomials in the variables $T_1, \dots, T_r$, the coordinates of $\mathbb{A}^r$. For points $y \in Y$ at which not all $F_\alpha(y) = 0$, there exist values $T_i = \tau_i$ for which $F(y, \tau) \neq 0$. Hence $g(D(F)) \supset \bigcup D(F_\alpha)$. This proves Theorem 1.14. □

Theorem 1.14 shows one respect in which regular maps of algebraic varieties are simpler than continuous or differentiable maps. The famous example of an everywhere dense line in the torus $T = \mathbb{R}^2/\mathbb{Z}^2$, a map such as

$$f: \mathbb{R}^1 \to T \quad \text{given by} \quad f(x) = (x, \sqrt{2}x) \bmod \mathbb{Z}^2$$

is an example of a situation that cannot happen for algebraic varieties, by Theorem 1.14.

Theorem 1.15 *If $X \subset \mathbb{P}^n$ is a closed subvariety disjoint from a d-dimensional linear subspace $E \subset \mathbb{P}^n$ then the projection $\pi: X \to \mathbb{P}^{n-d-1}$ with centre E (see Example 1.27) defines a finite map $X \to \pi(X)$.*

Proof Let $y_0, \dots, y_{n-d-1}$ be homogeneous coordinates on $\mathbb{P}^{n-d-1}$, and suppose that π is given by $y_j = L_j(x)$ for $j = 0, \dots, n-d-1$, where $x \in X$. Obviously $U_i = \pi^{-1}(\mathbb{A}_i^{n-d-1}) \cap X$ is given by the condition $L_i(x) \neq 0$, and is an affine open subset of X. We prove that $\pi: U_i \to \mathbb{A}_i^{n-d-1} \cap \pi(X)$ is a finite map. Any function $g \in k[U_i]$ is of the form $g = G_i(x_0, \dots, x_n)/L_i^m$, where G_i is a form of degree m.

Consider the map $\pi_1\colon X \to \mathbb{P}^{n-d}$ given by $z_j = L_j^m(x)$ for $j = 0, \dots, n-d-1$ and $z_{n-d} = G_i(x)$, where $z_0, \dots, z_{n-d}$ are homogeneous coordinates in $\mathbb{P}^{n-d}$. This is a regular map, and its image $\pi_1(X) \subset \mathbb{P}^{n-d}$ is closed by Theorem 1.10. Suppose that $\pi_1(X)$ is given by equations $F_1 = \cdots = F_s = 0$.

Since X is disjoint from E, the forms L_i for $i = 0, \dots, n-d-1$ have no common zeros on X. Hence the point $0 = (0 : \cdots : 0 : 1) \in \mathbb{P}^{n-d}$ is not contained in $\pi_1(X)$, or in other words, the equations $z_0 = \cdots = z_{n-d-1} = F_1 = \cdots = F_s = 0$ do not have solutions in $\mathbb{P}^{n-d}$. By Lemma 1.1, it follows from this that $(z_0, \dots, z_{n-d-1}, F_1, \dots, F_s) \supset I_k$ for some $k > 0$. In particular, $(z_0, \dots, z_{n-d-1}, F_1, \dots, F_s) \ni z_{n-d}^k$. This means that we can write

$$z_{n-d}^k = \sum_{j=0}^{n-d-1} z_j H_j + \sum_{j=1}^{s} F_j P_j,$$

where H_j and P_j are polynomials. Writing $H^{(q)}$ for the homogeneous component of H of degree q, we deduce from this that

$$\Phi(z_0, \dots, z_{n-d}) = z_{n-d}^k - \sum z_j H_j^{(k-1)} = 0 \quad \text{on } \pi_1(X). \tag{1.43}$$

The homogeneous polynomial Φ is of degree k and as a polynomial in z_{n-d} it has leading coefficient 1:

$$\Phi = z_{n-d}^k - \sum_{j=0}^{k-1} A_{k-j}(z_0, \dots, z_{n-d-1}) z_{n-d}^j. \tag{1.44}$$

If we substitute in (1.43) the formulas defining the map $\pi_1\colon X \to \mathbb{P}^{n-d}$, we get that $\Phi(L_0^m, \dots, L_{n-d-1}^m, G_i) = 0$ on X, with Φ of the form (1.44). Dividing this relation by L_i^{mk} we get the required relation

$$g^k - \sum_{j=0}^{k-1} A_{k-j}\bigl(x_0^m, \dots, 1, \dots, x_{n-d-1}^m\bigr) g^j = 0,$$

where $x_r = y_r / y_i$ are coordinates on $\mathbb{A}_i^{n-d-1}$. The theorem is proved. $\square$

Using the Veronese embedding (Example 1.28) allows the following substantial generalisation of Theorem 1.15.

Theorem 1.16 *Suppose that $F_0, \dots, F_s$ are forms of degree m on $\mathbb{P}^n$ having no common zeros on a closed variety $X \subset \mathbb{P}^n$. Then*

$$\varphi(x) = \bigl(F_0(x) : \cdots : F_s(x)\bigr)$$

defines a finite map $\varphi\colon X \to \varphi(X)$.

Proof Let $v_m\colon \mathbb{P}^n \to \mathbb{P}^N$ be the Veronese embedding (with $N = \binom{n+m}{m} - 1$) and L_i the linear forms on $\mathbb{P}^N$ corresponding to the forms F_i on $\mathbb{P}^n$. Then obviously $\varphi = \pi \circ v_m$ where π is the projection defined by the linear forms $L_0, \dots, L_s$. Since $v_m\colon X \to v_m(X)$ is an isomorphism, Theorem 1.16 follows from Theorem 1.15. □

5.4 Noether Normalisation

Consider an irreducible projective variety $X \subset \mathbb{P}^n$ distinct from the whole of $\mathbb{P}^n$. Then there exists a point $x \in \mathbb{P}^n \setminus X$, and the map φ obtained by projecting X away from x will be regular. The image $\varphi(X) \subset \mathbb{P}^{n-1}$ is projective by Theorem 1.10, and the map $\varphi\colon X \to \varphi(X)$ is finite by Theorem 1.15. If $\varphi(X) \neq \mathbb{P}^{n-1}$ then we can repeat the same argument for it. We finally arrive at a map $X \to \mathbb{P}^m$, which is finite, since it is a composite of finite maps. The result we have proved is called the Noether normalisation theorem:

Theorem 1.17 *For an irreducible projective variety X there exists a finite map $\varphi\colon X \to \mathbb{P}^m$ to a projective space.*

The analogous result also holds for affine varieties. To prove this, consider an affine variety $X \subset \mathbb{A}^n$. Embed $\mathbb{A}^n$ as an open $\mathbb{A}^n \subset \mathbb{P}^n$, and write $\overline{X}$ for the projective closure of X in $\mathbb{P}^n$. Suppose that $X \neq \mathbb{A}^n$. We choose a point at infinity $x \in \mathbb{P}^n \setminus \mathbb{A}^n$ with $x \notin \overline{X}$, and consider the projection $\varphi\colon \overline{X} \to \mathbb{P}^{n-1}$ from this point. Here X will map to points in the finite part of $\mathbb{P}^{n-1}$, that is, to points of $\mathbb{A}^{n-1} = \mathbb{P}^{n-1} \cap \mathbb{A}^n$. We can repeat this process as long as $X \neq \mathbb{A}^n$, and as a result we arrive at a projection $\varphi\colon \overline{X} \to \mathbb{P}^m$ for which $\varphi(X) = \mathbb{A}^m$. This proves the following result.

Theorem 1.18 *For an irreducible affine variety X there exists a finite map $\varphi\colon X \to \mathbb{A}^m$ to an affine space.*

Theorems 1.17–1.18 allow us to reduce the study of certain (very coarse) properties of projective and affine varieties to the case of projective and affine spaces. When $m = 1$ this point of view is due to Riemann, who considered algebraic curves as coverings of the Riemann sphere ($\mathbb{P}^1$ over the complex number field $\mathbb{C}$).

Theorem 1.18 means that an integral domains A that is finitely generated over the field k is integral over a subring isomorphic to a polynomial ring. This result can also easily be proved directly.

5.5 Exercises to Section 5

1 Prove that the Segre variety $\varphi(\mathbb{P}^n \times \mathbb{P}^m) \subset \mathbb{P}^N$ (where $N = (n+1)(m+1) - 1$) is not contained in any linear subspace strictly smaller than the whole of $\mathbb{P}^N$.

2 Consider the two maps of varieties $\mathbb{P}^1 \times \mathbb{P}^1 \to \mathbb{P}^1$ given by $p_1(x, y) = x$ and $p_2(x, y) = y$. Prove that $p_1(X) = p_2(X) = \mathbb{P}^1$ for any closed irreducible subset $X \subset \mathbb{P}^1 \times \mathbb{P}^1$, unless X is of one of the following types: (a) a point $(x_0, y_0) \in \mathbb{P}^1 \times \mathbb{P}^1$; (b) a line $x_0 \times \mathbb{P}^1$ for $x_0 \in \mathbb{P}^1$ a fixed point; (c) a line $\mathbb{P}^1 \times y_0$.

3 Verify Theorem 1.10, Corollary 1.1 directly for the case $X = \mathbb{P}^n$.

4 Let $X = \mathbb{A}^2 \setminus x$ where x is a point. Prove that X is not isomorphic to an affine nor a projective variety (compare Exercise 3 of Section 4.5).

5 The same question as Exercise 4, for $X = \mathbb{P}^2 \setminus x$.

6 The same question as Exercise 4, for $X = \mathbb{P}^1 \times \mathbb{A}^1$.

7 Is the map $f\colon \mathbb{A}^1 \to X$ finite, where X is given by $y^2 = x^3$, and f by $f(t) = (t^2, t^3)$.

8 Let $X \subset \mathbb{A}^r$ be a hypersurface of $\mathbb{A}^r$ and L a line of $\mathbb{A}^r$ through the origin. Let φ_L be the map projecting X parallel to L to an $(r-1)$-dimensional subspace not containing L. Write S for the set of all lines L such that φ_L is not finite. Prove that S is an algebraic variety. [Hint: Prove that $S = \overline{X} \cap \mathbb{P}^{r-1}_\infty$.] Find S if $r = 2$ and X is given by $xy = 1$.

9 Prove that any intersection of affine open subsets is affine. [Hint: Use Example 1.20.]

10 Prove that forms of degree $m = kl$ in $n+1$ variables that are lth powers of forms correspond to the points of a closed subset of $\mathbb{P}^N$, where $N = \binom{n+m}{m} - 1 = \nu_{n,m}$.

11 Let $f\colon X \to Y$ be a regular map of affine varieties. Prove that the inverse image of a principal affine open set is a principal affine open set.

6 Dimension

6.1 Definition of Dimension

In Section 2 we saw that closed algebraic subvarieties $X \subset \mathbb{A}^2$ are finite sets of points, algebraic plane curves, and $\mathbb{A}^2$ itself. This division into three cases corresponds to the intuitive notion of dimension, with varieties of dimension 0, 1 and 2. Here we give the definition of the dimension of an arbitrary algebraic variety.

How could we arrive at this definition? First, of course, we take the dimension of $\mathbb{P}^n$ and $\mathbb{A}^n$ to be n. Secondly, if there exists a finite map $X \to Y$ then it is natural to suppose that X and Y have the same dimension. Since by Noether normalisation

(Theorems 1.17–1.18), any projective or affine variety X has a finite map to some $\mathbb{P}^m$ or $\mathbb{A}^m$, it is natural to take m as the definition of the dimension of X. However, the question then arises as to whether this is well defined: might there not exist two finite maps $f\colon X \to \mathbb{A}^m$ and $g\colon X \to \mathbb{A}^n$ with $m \neq n$? Suppose that X is irreducible. Then the finiteness of a regular map $f\colon X \to \mathbb{A}^m$ implies that the rational function field $k(X)$ is a finite extension of the field $f^*(k(\mathbb{A}^m))$, which is in turn isomorphic to $k(t_1, \dots, t_m)$. Hence $k(X)$ has transcendence degree m over k; this gives a characterisation of the number m independent of the choice of the finite map $f\colon X \to \mathbb{A}^m$. This gives some motivation for the definition of dimension.

Definition The *dimension* of an irreducible quasiprojective variety X is the transcendence degree of the function field $k(X)$; it is denoted by $\dim X$. The dimension of a reducible variety is the maximum of the dimension of its irreducible components. If $Y \subset X$ is a closed subvariety of X then the number $\dim X - \dim Y$ is called the *codimension* of Y in X, and written $\operatorname{codim} Y$ or $\operatorname{codim}_X Y$. Algebraic varieties of dimension 1 and 2 are called *curves* and *surfaces*.[3]

Note that if X is an irreducible variety and $U \subset X$ is open then $k(U) = k(X)$, and hence $\dim U = \dim X$.

Example 1.30 $\dim \mathbb{A}^n = \dim \mathbb{P}^n = n$, because the field $k(\mathbb{A}^n)$ is the field of rational functions in n variables. Since dimension is by definition invariant under birational equivalence, we see that $\mathbb{A}^n$ and $\mathbb{A}^m$ are not birational if $n \neq m$.

Example 1.31 An irreducible plane curve is 1-dimensional, as we saw in Section 1.3.

Example 1.32 If X consists of a single point then obviously $\dim X = 0$, and thus the same holds if X is a finite set. Conversely, if $\dim X = 0$ then X is a finite set. It is enough to prove this for an irreducible affine variety X. Let $X \subset \mathbb{A}^n$, and write $t_1, \dots, t_n$ for the coordinates on $\mathbb{A}^n$ as functions on X, that is, as elements of $k[X]$. By assumption the t_i are algebraic over k, and can hence only take finitely many values. It follows from this that X is finite.

Example 1.33 We prove that if X and Y are irreducible varieties then

$$\dim(X \times Y) = \dim X + \dim Y.$$

We need only consider the case that $X \subset \mathbb{A}^N$ and $Y \subset \mathbb{A}^M$ are affine varieties. Suppose that $\dim X = n$, $\dim Y = m$, and let $t_1, \dots, t_N$ and $u_1, \dots, u_M$ be coordinates of $\mathbb{A}^N$ and $\mathbb{A}^M$ considered as functions on X and Y respectively, such that

[3] n-dimensional varieties are often called *n-folds*, for example 3-folds, 4-folds (or threefolds, fourfolds).

$t_1, \dots, t_n$ are algebraically independent in $k(X)$ and $u_1, \dots, u_m$ in $k(Y)$. By definition $k[X \times Y]$ is generated by the elements $t_1, \dots, t_N, u_1, \dots, u_M$, and under the current assumptions all of these are algebraically dependent on $t_1, \dots, t_n, u_1, \dots, u_m$. Hence it is enough to prove that these elements are algebraically independent. Suppose that there is a relation $F(T, U) = F(T_1, \dots, T_n, U_1, \dots, U_m) = 0$ on $X \times Y$. Then for any point $x \in X$ we have $F(x, U_1, \dots, U_m) = 0$ on Y. Since $u_1, \dots, u_m$ are algebraically independent in $k(Y)$, every coefficient $a_i(x)$ of the polynomial $F(x, U)$ is zero; this means that the corresponding polynomial $a_i(T_1, \dots, T_n)$ is 0 on X. Now we use the fact that $t_1, \dots, t_n$ are algebraically independent in $k(X)$ and deduce from this that $a_i(T_1, \dots, T_n) = 0$, and hence $F(T, U)$ is identically 0.

Example 1.34 The Grassmannian $\mathrm{Grass}(r, n)$ (see Example 1.24) is covered by open sets $p_{i_1 \dots i_r} \neq 0$ isomorphic to the affine space $\mathbb{A}^{r(n-r)}$. Thus $\dim \mathrm{Grass}(r, n) = r(n-r)$. It also follows from this that $\mathrm{Grass}(r, n)$ is rational.

Theorem 1.19 *If $X \subset Y$ then $\dim X \le \dim Y$. If Y is irreducible and $X \subset Y$ is a closed subvariety with $\dim X = \dim Y$ then $X = Y$.*

Proof It is enough to prove the assertions for X and Y irreducible affine varieties.

Suppose $X \subset Y \subset \mathbb{A}^N$ with $\dim Y = n$. Then any $n+1$ of the coordinate functions $t_1, \dots, t_N$ are algebraically dependent as elements of $k[Y]$, that is, are connected by a relation $F(t_{i_1}, \dots, t_{i_{n+1}}) = 0$ on Y. A fortiori this holds on X. This means that the transcendence degree of $k(X)$ is at most n, so that $\dim X \le \dim Y$.

Now suppose that $\dim X = \dim Y = n$. Then some n of the coordinates $t_1, \dots, t_N$ are algebraically independent on X; suppose that these are $t_1, \dots, t_n$. Then a fortiori they are algebraically independent on Y. Let $u \in k[Y]$ with $u \neq 0$ on Y. Then u on Y is algebraically dependent on $t_1, \dots, t_n$, that is, there is a polynomial $a(t, U) \in k[t_1, \dots, t_n][U]$ such that the relation

$$a_0(t_1, \dots, t_n)u^k + \dots + a_k(t_1, \dots, t_n) = 0 \qquad (1.45)$$

holds on Y. We can choose $a(t, U)$ to be irreducible, and then $a_k(t_1, \dots, t_n) \neq 0$ on Y. Relation (1.45) holds a fortiori on X. Suppose that $u = 0$ on X. Then (1.45) implies that $a_k(t_1, \dots, t_n) = 0$ on X. Since by assumption $t_1, \dots, t_n$ are independent on X, it follows that $a_k(t_1, \dots, t_n) = 0$ on the whole of $\mathbb{A}^N$. This contradicts $a_k(t_1, \dots, t_n) \neq 0$ on Y. Thus if $u = 0$ on X then also $u = 0$ on Y, and therefore $X = Y$. The theorem is proved. □

We have seen that an irreducible algebraic plane curve is 1-dimensional. The following result is a generalisation.

Theorem 1.20 *Every irreducible component of a hypersurface in $\mathbb{A}^n$ or $\mathbb{P}^n$ has codimension* 1.

Proof It is enough to consider the case of a hypersurface in $\mathbb{A}^n$. Suppose that a variety $X \subset \mathbb{A}^n$ is given by an equation $F(T) = 0$. The factorisation $F = F_1 \dots F_k$

of F into irreducible factors corresponds to an expression $X = X_1 \cup \cdots \cup X_k$, where X_i is defined by $F_i = 0$. It is obviously sufficient to prove the theorem for each variety X_i. Let us prove that X_i is irreducible: if X_i were reducible, there would exist polynomials G and H such that $GH = 0$ on X_i but $G, H \neq 0$ on X_i. From the Nullstellensatz it follows that $F_i \mid (GH)^l$ for some $l > 0$. Since F_i is irreducible it follows from this that $F_i \mid G$ or $F_i \mid H$, and this contradicts $G \neq 0$, $H \neq 0$ on X_i.

Suppose that the variable T_n actually appears in the polynomial $F_i(T)$, and prove that the coordinates $t_1, \ldots, t_{n-1}$ are algebraically independent on X_i. Indeed, a relation $G(t_1, \ldots, t_{n-1}) = 0$ on X_i would imply that $F_i \mid G^l$ for some $l > 0$, which is impossible since G does not involve T_n. Thus $\dim X_i \geq n - 1$; since $X \neq \mathbb{A}^n$, it follows from Theorem 1.19 that $\dim X_i = n - 1$. Theorem 1.20 is proved. □

Theorem 1.21 *Let $X \subset \mathbb{A}^n$ be a variety, and suppose that all the components of X have dimension $n - 1$. Then X is a hypersurface and the ideal $\mathfrak{A}_X$ is principal.*

Proof We only need consider the case that X is irreducible. Since $X \neq \mathbb{A}^n$ (because $\dim X = n - 1$), there exists a nonzero polynomial F which is zero on X. Since X is irreducible, some irreducible factor H of F is also zero on X. Write $Y \subset \mathbb{A}^n$ for the hypersurface defined by $H = 0$; we saw in the proof of Theorem 1.20 that Y is irreducible. Then $X \subset Y$, so that $X = Y$ by Theorem 1.19. If $G \in \mathfrak{A}_X$ then by the Nullstellensatz $H \mid G^l$ for some $l > 0$, and then $G \in (H)$ by the irreducibility of H, that is $\mathfrak{A}_X = (H)$.

Theorem 1.21 is proved. □

The following analogue of Theorem 1.21 is proved similarly:

Theorem 1.21′ *Let $X \subset \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_k}$ be a variety, and suppose that all the components of X have dimension $n_1 + \cdots + n_k - 1$. Then X is defined by one equation that is homogeneous in each of the k sets of variables.*

Proof We need only replace the unique factorisation of polynomials used in the proof of Theorem 1.21 by the unique factorisation of polynomials that are homogeneous in each of the k groups of variables into irreducible polynomials of the same type. This comes from the fact that if $F(x_0, \ldots, x_{n_1}, y_0, \ldots, y_{n_2}, \ldots, u_0, \ldots, u_{n_k})$ is homogeneous in each of the k sets of variables $\{x_0, \ldots, x_{n_1}\}, \ldots, \{u_0, \ldots, u_{n_k}\}$ and F factorises as $F = G \cdot H$, then G and H have the same homogeneity property. Theorem 1.21′ is proved. □

6.2 Dimension of Intersection with a Hypersurface

If we try to study varieties defined by more than one equation, we come up at once against the question of the dimension of intersection of a variety with a hypersurface. We study this question first for projective varieties. If X is closed in $\mathbb{P}^N$ and a form

F is not zero on X then we write X_F for the closed subvariety of X defined by $F = 0$.

For any projective variety $X \subset \mathbb{P}^N$ we can find a form $G(U_0, \dots, U_N)$ of any specified degree m which does not vanish on any components X_i of X. For this, it is enough to choose one point $x_i \in X_i$ in each irreducible component of X, and find a linear form L not vanishing on any of these; then we can take $G = L^m$ to be the appropriate power of L. Suppose that $X \subset \mathbb{P}^N$ is closed, and that a form F is not zero on any component of X. By Theorem 1.19 we have $\dim X_F < \dim X$. Set $X_F = X^{(1)}$ and apply the same argument to $X^{(1)}$, finding a form F_1 with $\deg F_1 = \deg F$ not vanishing on any component of $X^{(1)}$. We get a chain of varieties $X^{(i)}$ and forms F_i such that

$$X = X^{(0)} \supset X^{(1)} \supset \cdots, \quad \text{with } X^{(i+1)} = X^{(i)}_{F_i} \text{ and } F_0 = F. \tag{1.46}$$

By Theorem 1.19, $\dim X^{(i+1)} < \dim X^{(i)}$. Hence if $\dim X = n$, then $X^{(n+1)}$ is empty. In other words, the forms $F_0 = F, F_1, \dots, F_n$ have no common zeros on X.

Suppose now that X is irreducible. Consider the map $\varphi \colon X \to \mathbb{P}^n$ given by

$$\varphi(x) = \bigl(F_0(x) : \cdots : F_n(x)\bigr). \tag{1.47}$$

This map satisfies the assumptions of Theorem 1.16, and by this theorem the map $X \to \varphi(X)$ is finite. But if $X \to Y$ is a finite map then, as we have seen, $\dim X = \dim Y$. Hence $\dim \varphi(X) = \dim X = n$, and since $\varphi(X) \subset \mathbb{P}^n$ is closed by Theorem 1.10, we get $\varphi(X) = \mathbb{P}^n$ by Theorem 1.19. Suppose now that $\dim X^{(1)} = \dim X_F < n - 1$. Then in (1.46), already $X^{(n)}$ is empty. In other words, the forms $F_0, \dots, F_{n-1}$ have no common zeros on X. This means that the point $(0 : \cdots : 0 : 1)$ is not contained in $\varphi(X)$, which contradicts $\varphi(X) = \mathbb{P}^n$. Thus we have proved the following result.

Theorem 1.22 *If a form F is not 0 on an irreducible projective variety X then* $\dim X_F = \dim X - 1$.

Recall that this means that X_F contains one or more irreducible components of dimension $\dim X - 1$.

Corollary 1.3 *A projective variety X contains subvarieties of any dimension $s < dim X$.*

Corollary 1.4 (Inductive definition of dimension) *If X is an irreducible projective variety then* $\dim X = 1 + \sup \dim Y$, *where Y runs through all proper subvarieties of X.*

Corollary 1.5 *The dimension of a projective variety X can be defined as the maximal integer n for which there exists a strictly decreasing chain $Y_0 \supsetneqq Y_1 \supsetneqq \cdots \supsetneqq Y_n \supsetneqq \emptyset$ of length n of irreducible subvarieties $Y_i \subset X$.*

Corollary 1.6 *The dimension n of a projective variety $X \subset \mathbb{P}^N$ can be defined as $N - s - 1$, where s is the maximum dimension of a linear subspace of $\mathbb{P}^N$ disjoint from X.*

Proof Let $E \subset \mathbb{P}^N$ be a linear subspace of dimension s. If $s \geq N - n$ then E can be defined by $\leq n$ equations, and successive application of Theorem 1.22 proves that $\dim(X \cap E) \geq 0$, and hence $X \cap E$ is nonempty (the dimension of the empty set is -1!). Setting $m = 1$ in the construction of the chain (1.46) gives $n + 1$ linear forms $L_0, \dots, L_n$ with no common zeros on X. If E is the linear subspace defined by these, then $\dim E = N - n - 1$ and $X \cap E$ is empty. Corollary 1.6 is proved. □

Corollary 1.7 *The variety of common zeros of r forms $F_1, \dots, F_r$ on an n-dimensional projective variety has dimension $\geq n - r$.*

The proof is by $r - 1$ applications of Theorem 1.22. Corollary 1.7 provides a rather strong existence theorem.

Proposition *If $r \leq n$ then r forms have a common zero on an n-dimensional projective variety. For example, in the case $X = \mathbb{P}^n$, this says that n homogeneous equations in $n + 1$ variables have a nonzero solution.*

This existence theorem allows us to make a number of important deductions.

Corollary 1.8 *Any two curves of $\mathbb{P}^2$ intersect.*

This is clear, since a curve is given by a single homogeneous equation. However, there exist nonintersecting curves on a nonsingular quadric surface $Q \subset \mathbb{P}^3$, for example the lines of one family of generators. Therefore $\mathbb{P}^2$ and Q are not isomorphic. Since they are birational (Example 1.22) we get an example of two varieties that are birational but not isomorphic. This example will appear again later (Sections 4.1 and 4.5, Chapter 2, Example 4.11 of Section 2.3, Chapter 4).

Corollary 1.9 *Theorem 1.21 fails already for the curves on a nonsingular quadric surface Q: there exist curves $C \subset Q$ that cannot be defined by setting to zero a single form on $\mathbb{P}^3$.*

Indeed, if we assume that each of the disjoint curves C_1 and C_2 which we found on Q is defined by one equation $F_1 = 0$ and $F_2 = 0$, we get a contradiction to Corollary 1.7, according to which the system of equations $G = F_1 = F_2 = 0$ have a common solution (where G is the equation of Q).

Corollary 1.10 *Any curve of degree ≥ 3 has an inflexion point.*

Proof We have seen in Section 1.6 that the inflexion points of an algebraic plane curve with equation $F = 0$ is defined by $H(F) = 0$, where $H(F)$ is the Hessian form of F. If F has degree n then $H(F)$ has degree $3(n-2)$. Therefore for $n \geq 3$ the system of equations $F = H(F) = 0$ has a nonzero solution; that is, the curve $F = 0$ has an inflexion point. Corollary 1.10 is proved. □

The simplest case is when $n = 3$. We see that every cubic curve in $\mathbb{P}^2$ has an inflexion point. Choose a coordinate system (ξ_0, ξ_1, ξ_2) so that the inflexion point is $(0, 0, 1)$, and the inflexional tangent is the line $\xi_1 = 0$. Setting $u = \xi_0/\xi_2$, $v = \xi_1/\xi_2$, we see easily that our assumption is equivalent to saying that the equation $\varphi(u, v)$ of the curve has no constant term, or term in u or u^2. Changing to coordinates $x = \xi_0/\xi_1$, $y = \xi_2/\xi_1$, so that the inflexion point is at infinity, we find that the equation of our cubic has no term in y^3, y^2x or yx^2, that is, it is of the form $ay^2 + (bx + c)y + g(x) = 0$, where g is a polynomial of degree ≤ 3. If $a = 0$ then the inflexion point is singular. If $a \neq 0$ we can assume that $a = 1$. Assuming that $\operatorname{char} k \neq 2$, we can complete the square by setting $y_1 = y + (1/2)(bx + c)$ and reduce the equation to the form $y_1^2 = g_1(x)$, where $g_1(x)$ has degree ≤ 3, and $= 3$ if the cubic curve is nonsingular. Thus *the equation of a nonsingular cubic has Weierstrass normal form in some coordinate system*. In Section 1.4 we proved only the weaker statement that a cubic is *birational* to a curve with equation in Weierstrass normal form.

Corollary 1.11 (Tsen's theorem) *Let $F(x_1, \ldots, x_n)$ be a form in n variables of degree $m < n$ whose coefficients are polynomials in one variable t. Then the equation $F(x_1, \ldots, x_n) = 0$ has a solution in polynomials $x_i = p_i(t)$.*

Proof We look for x_i of the form $x_i = \sum_{j=0}^{l} u_{ij} t^j$ with unknown coefficients u_{ij}. Substituting these expressions in the equation $F(x_1, \ldots, x_n) = 0$, we get a polynomial in t all of whose coefficients must be set to 0. If the maximum of the degrees of the coefficients of a polynomial F equals k then the number of equations is at most $ml + k + 1$. The number of indeterminates is $n(l + 1)$. Since by assumption $n > m$, for l sufficiently large, the number of unknowns is greater than the number of equations, and hence the system has a nonzero solution. □

Example 1.35 An important particular case of Tsen's theorem is when $n = 3$ and F is a quadratic form. It can be given the following geometric interpretation: suppose that a surface $X \subset \mathbb{P}^2 \times \mathbb{A}^1$ is defined by the equation

$$q(x_0 : x_1 : x_2; t) = \sum_{i,j=0}^{2} a_{ij}(t) x_i x_j \quad \text{with } a_{ij}(t) \in k[t],$$

where $(x_0 : x_1 : x_2)$ are coordinates in $\mathbb{P}^2$ and t a coordinate on $\mathbb{A}^1$. The fibres of the map $X \to \mathbb{A}^1$ are the conics $q(x_0 : x_1 : x_2; a) = 0$ for $a \in \mathbb{A}^1$, and the surface is called a *conic bundle* or *pencil of conics*. Tsen's theorem proves that a pencil of conics has a section, that is, there exists a regular map $\varphi\colon \mathbb{A}^1 \to X$ such that $\varphi(a)$ is a point of the fibre over a for every $a \in \mathbb{A}^1$.

Another interpretation of this result is as follows. Consider our surface X as the conic C with equation $q(x_0 : x_1 : x_2; t) = \sum_{i,j=0}^{2} a_{ij} x_i x_j = 0$ in $\mathbb{P}^2$ over the algebraically nonclosed field $K = k(t)$. Obviously $K(C) = k(X)$. Then C has a point with coordinates in K.

We assume that the curve C is irreducible for a general point $t \in \mathbb{A}^1$, that is, that $\det|a_{ij}(t)|$ is not identically 0; we say that the pencil of conics is *nondegenerate* in this case. In Section 1.2 we saw that the conic is then rational, with the birational map to $\mathbb{P}^1$ defined over $K = k(t)$. In other words, the field $K(C)$ is isomorphic over K to the field $K(x)$, and since $K(C) = k(X)$ it follows that $k(X)$ is isomorphic to $K(x) = k(t, x)$. We have proved the next result.

Corollary 1.12 *A nondegenerate pencil of conics over $\mathbb{A}^1$ is a rational surface.*

Theorem 1.23 *Under the assumptions of Theorem 1.22, every component of X_F has dimension* $\dim X - 1$.

Proof Consider the finite map $\varphi\colon X \to \mathbb{P}^n$ (with $n = \dim X$) constructed in the proof of Theorem 1.22, and let $\mathbb{A}^n_i \subset \mathbb{P}^n$ for $i = 0, \dots, n$ be the affine open sets covering $\mathbb{P}^n$. Then using the Veronese embedding with $m = \deg F$, it is easy to see that $\varphi^{-1}(\mathbb{A}^n_i) = U_i$ are affine open sets of X. It is obviously enough to prove that each component of the affine variety $X_F \cap U_i$ has dimension $n - 1$ for each i. From now on our arguments apply to some fixed U_i, which we denote by U. Obviously $X_F \cap U = V(f)$, where $f = F/F_i$, that is, X_F coincides on U with the set of zeros of the regular function $f \in k[U]$. We constructed above a finite map $\varphi\colon U \to \mathbb{A}^n$, given by n regular functions $f_1, \dots, f_n$, with $f = f_1$.

To prove that each component of $V(f)$ has dimension $n - 1$, we only need to prove that it has dimension $\geq n - 1$. We prove that the functions $f_2, \dots, f_n$ are algebraically independent on each component. Let $P \in k[T_2, \dots, T_n]$. To prove that $R = P(f_2, \dots, f_n)$ does not vanish on any component of $V(f)$ it is enough to prove that for $Q \in k[U]$,

$$RQ = 0 \quad \text{on } V(f) \quad \Longrightarrow \quad Q = 0 \quad \text{on } V(f).$$

Indeed, if $V(f) = U^{(1)} \cup \dots \cup U^{(t)}$ is an irredundant decomposition into irreducible components, and $R = 0$ on $U^{(1)}$, then take Q to be any function that vanishes on $U^{(2)} \cup \dots \cup U^{(t)}$ but not on $U^{(1)}$. Then $RQ = 0$ on $V(f)$ but $Q \neq 0$ on $V(f)$.

By the Nullstellensatz our assertion can be restated as follows: if $f \mid (RQ)^l$ for some $l > 0$ then $f \mid Q^k$ for some $k > 0$. Thus Theorem 1.23 follows from the following purely algebraic fact:

Lemma *Set $B = k[T_1, \dots, T_n]$, and let $A \supset B$ be an integral domain that is integral over B; write $x = T_1$, and let $y = P(T_2, \dots, T_n) \neq 0$. Then for any $u \in A$,*

$$x \mid (yu)^l \quad \textit{in } A \textit{ for some } l > 0 \quad \Longrightarrow \quad x \mid u^k \quad \textit{for some } k > 0.$$

Proof of the Lemma The only property of x and y that we use is that they are relatively prime in the UFD $k[T_1, \dots, T_n]$. Note that we can replace y^l by z and u^l by

v, and then it is enough to prove that if x and z are relatively prime in $k[T_1, \dots, T_n]$ then $x \mid zv$ in A implies that $x \mid v^k$ for some $k > 0$. Thus the lemma asserts that the property of polynomials $x, z \in B$ being relatively prime is in a certain sense preserved on passing to a ring A that is integral over B.

Write K for the field of fractions of B. If $t \in A$ is integral over B then it is algebraic over K. Let $F(T) \in K[T]$ be the minimal polynomial of t over K, that is, the polynomial of least degree with leading coefficient 1 such that $F(t) = 0$. Division with remainder shows that any polynomial $G(T) \in K[T]$ with $G(t) = 0$ is divisible by $F(T)$ in $K[T]$. Now from this it follows that t is integral over B if and only if $F[T] \in B[T]$. Indeed, if t is integral and $G(t) = 0$ for $G \in B[T]$ with leading coefficient 1, then $G(T) = F(T)H(T)$ in $K[T]$. But $B = k[T_1, \dots, T_n]$ is a UFD, so a simple application of Gauss' lemma shows that $F(T), H(T) \in B[T]$.

It is now easy to complete the proof of the lemma. Suppose that $zv = xw$ with $v, w \in A$ and let $F(T) = T^k + b_1T^{k-1} + \dots + b_k$ be the minimal polynomial of w. Since w is integral over B, the coefficients b_i of F satisfy $b_i \in B$. It is easy to see that the minimal polynomial $G(T)$ of $v = xw/z$ is given by $(x/z)^k F(zT/x)$. Therefore

$$\begin{aligned} G(T) &= T^k + \frac{xb_1}{z} T^{k-1} + \dots + \frac{x^k b_k}{z^k}, \\ \text{and} \quad v^k &+ \frac{xb_1}{z} v^{k-1} + \dots + \frac{x^k b_k}{z^k} = 0. \end{aligned} \tag{1.48}$$

Since v is integral over B, also $x^i b_i / z^i \in B$, and because x and z are relatively prime it follows that $z^i \mid b_i$. It then follows from (1.48) that $x \mid v^k$. The lemma is proved, and with it Theorem 1.23. □

Corollary 1.13 *If $X \subset \mathbb{P}^N$ is an irreducible quasiprojective variety and F a form that is not identically* 0 *on X, then every* (*nonempty*) *component of X_F has codimension* 1. ($X_F = \emptyset$ is of course possible for quasiprojective varieties.)

Proof By definition X is open in some closed subset $\overline{X} \subset \mathbb{P}^N$. Since X is irreducible, so is $\overline{X}$, and hence $\dim \overline{X} = \dim X$. By Theorem 1.23, $(\overline{X})_F = \bigcup Y_i$ with $\dim Y_i = \dim X - 1$. But it is easy to see that $X_F = (\overline{X})_F \cap X$; it follows that $X_F = \bigcup (Y_i \cap X)$, and $Y_i \cap X$ is either empty or is open in Y_i, so that $\dim(Y_i \cap X) = \dim X - 1$. This proves Corollary 1.13. □

The particular case of this lemma that usually turns up is when $X \subset \mathbb{A}^n$ is an affine variety. Let $\mathbb{A}^n \subset \mathbb{P}^n$ be the subset $\mathbb{A}^n_0$ given by $u_0 \neq 0$, and write $m = \deg F$ and $f = F/u_0^m$; then $X_F = V(f)$. In other words, X_F is just the set of zeros of some regular function $f \in k[X]$.

Corollary 1.14 *Let $X \subset \mathbb{P}^N$ be an irreducible n-dimensional quasiprojective variety, and $Y \subset X$ the set of zeros of m forms on X. Then every* (*nonempty*) *component of Y has dimension $\geq n - m$.*

Proof The proof is by an obvious induction on m. In the case of an affine variety X we can again say that Y is the set of zeros of m regular functions on X. If X is projective and $m \le n$ then by the proposition after Theorem 1.22, Corollary 1.7 we can assert that $Y \neq \emptyset$. Corollary 1.14 is proved. □

Theorem 1.24 *Let $X, Y \subset \mathbb{P}^N$ be irreducible quasiprojective varieties with* $\dim X = n$ *and* $\dim Y = m$. *Then any (nonempty) component Z of $X \cap Y$ has* $\dim Z \ge n + m - N$.

Moreover, if X and Y are projective and $n + m \ge N$ then $X \cap Y \neq \emptyset$.

Proof The theorem is obviously local in nature, and we therefore only need to prove it in the case of affine varieties. Suppose that $X, Y \subset \mathbb{A}^N$. Write $\Delta \subset \mathbb{A}^N \times \mathbb{A}^N = \mathbb{A}^{2N}$ for the diagonal (see Example 1.20). Then $X \cap Y$ is isomorphic to $(X \times Y) \cap \Delta \subset \mathbb{A}^{2N}$. The theorem follows from Corollary 1.14, since $\Delta \subset \mathbb{A}^{2N}$ is defined by N equations.

For the final sentence, apply the first part to the affine cone over X and Y. The theorem is proved. □

Theorem 1.24 can be stated in a more symmetric form, in which it generalises at once to the intersection of any number of subvarieties:

$$\operatorname{codim}_X \bigcap_{i=1}^{r} Y_i \;\le\; \sum_{i=1}^{r} \operatorname{codim}_X Y_i. \tag{1.49}$$

6.3 The Theorem on the Dimension of Fibres

For a given regular map $f: X \to Y$ of quasiprojective varieties, and $y \in Y$, the set $f^{-1}(y)$ is called the *fibre* of f over y. It is obviously a closed subvariety of X. The idea behind the terminology is that f fibres X as the disjoint union of the fibres over the different points $y \in f(X)$.

Theorem 1.25 *Let $f: X \to Y$ be a regular map between irreducible varieties. Suppose that f is surjective: $f(X) = Y$, and that* $\dim X = n$, $\dim Y = m$. *Then $m \le n$, and*

(i) $\dim F \ge n - m$ *for any $y \in Y$ and for any component F of the fibre $f^{-1}(y)$;*
(ii) *there exists a nonempty open subset $U \subset Y$ such that* $\dim f^{-1}(y) = n - m$ *for $y \in U$.*

Proof of (i) This property is obviously local over Y, and it is enough to prove it after replacing Y by any open set $U \subset Y$ with $U \ni y$ and X by $f^{-1}(U)$. Hence we can assume that Y is affine. Suppose that $Y \subset \mathbb{A}^N$. In the chain of subvarieties of Y given by (1.46), $Y^{(m)}$ is a finite set $Y^{(m)} = Y \cap Z$, where Z is defined by m

equations and $y \in Z$. The open set U can be chosen such that $Z \cap Y \cap U = \{y\}$, and so we can assume that $Z \cap Y = \{y\}$. The subspace Z is defined by m equations $g_1 = \cdots = g_m = 0$. Thus in Y the system of equations $g_1 = \cdots = g_m = 0$ defines the point y. This means that in X the system of equations $f^*(g_1) = \cdots = f^*(g_m) = 0$ defines the subvariety $f^{-1}(y)$. Assertion (i) now follows from Theorem 1.23, Corollary 1.14 (the affine case). $\square$

Proof of (ii) We can replace Y by an affine open subset W and X by an open affine set $V \subset f^{-1}(W)$. Since V is dense in $f^{-1}(W)$ and f is surjective, $f(V)$ is dense in W. Hence f defines an inclusion $f^*\colon k[W] \hookrightarrow k[V]$. From now on we take $k[W] \subset k[V]$, therefore $k(W) \subset k(V)$. Write $k[W] = k[w_1, \dots, w_M]$ and $k[V] = k[v_1, \dots, v_N]$. Since $\dim W = m$ and $\dim V = n$, the field $k(V)$ has transcendence degree $n - m$ over $k(W)$. Suppose that $v_1, \dots, v_{n-m}$ are algebraically independent over $k(W)$, and the remaining v_i algebraic over $k(W)[v_1, \dots, v_{n-m}]$, with relations

$$F_i(v_i; v_1, \dots, v_{n-m}; w_1, \dots, w_M) = 0 \quad \text{for } i = n - m + 1, \dots, N.$$

Write $\overline{v}_i$ for the function v_i restricted to $f^{-1}(y) \cap V$. Then

$$k\bigl[f^{-1}(y) \cap V\bigr] = k[\overline{v}_1, \dots, \overline{v}_N]. \tag{1.50}$$

We now view F_i as a polynomial in v_i, $v_1, \dots, v_{n-m}$, with coefficients functions of $w_1, \dots, w_M$, and define Y_i to be the subvariety of W given by the vanishing of the leading term of F_i. Set $E = \bigcup Y_i$ and $U = W \setminus E$. Obviously U is open and nonempty. By construction of E, if $y \in U$ then none of the polynomials $F_i(T_i; T_1, \dots, T_{n-m}; w_1(y), \dots, w_M(y))$ is identically zero, and therefore all the $\overline{v}_i$ are algebraically dependent on $\overline{v}_1, \dots, \overline{v}_{n-m}$. Together with formula (1.50) this proves that $\dim f^{-1}(y) \le n - m$, so that (ii) of the proposition follows from (i). The theorem is proved. $\square$

It is easy to give examples where (ii) does not hold for every $y \in Y$; (see for example Exercise 6 of Section 2.4, and the end of Section 6.4). That is, the dimension of fibres may jump up.

Corollary *The sets $Y_k = \{y \in Y \mid \dim f^{-1}(y) \ge k\}$ are closed in Y.*

Proof By Theorem 1.25, $Y_{n-m} = Y$, and there exists a closed subset $Y' \subsetneqq Y$ such that $Y_k \subset Y'$ if $k > n - m$. If Z_i are the irreducible components of Y' and $f_i\colon f^{-1}(Z_i) \to Z_i$ the restrictions of f, then $\dim Z_i < \dim Y$, and we can prove the corollary by induction on $\dim Y$. The corollary is proved. $\square$

Theorem 1.25 implies a criterion for a variety to be irreducible which is often useful.

Theorem 1.26 *Let $f\colon X \to Y$ be a regular map between projective varieties, with $f(X) = Y$. Suppose that Y is irreducible, and that all the fibres $f^{-1}(y)$ for $y \in Y$ are irreducible and of the same dimension. Then X is irreducible.*

Proof Let $X = \bigcup X_i$ be an irreducible decomposition. By Theorem 1.10, each $f(X_i)$ is closed. Since $Y = \bigcup f(X_i)$ and Y is irreducible, $Y = f(X_i)$ for some i.

Set $\dim f^{-1}(y) = n$. For each i such that $Y = f(X_i)$, by Theorem 1.25, (ii), there exists a dense open set $U_i \subset Y$ and an integer n_i such that $\dim(f_i^{-1}(y)) = n_i$ for all $y \in U_i$. Extend the definition of U_i to i such that $f(X_i) \neq Y$ by setting $U_i = Y \setminus f(X_i)$. Consider $y \in \bigcap U_i$. Then since $f^{-1}(y)$ is irreducible, we must have $f^{-1}(y) \subset X_i$ for some i, say $i = 0$. Write $f_0 \colon X_0 \to Y$ for the restriction of f. Then $f^{-1}(y) \subset f_0^{-1}(y)$; but the opposite inclusion is trivial, so that $f^{-1}(y) = f_0^{-1}(y)$ and $n = n_0$.

Now since f_0 is surjective, we know that $f_0^{-1}(y) \subset f^{-1}(y)$ is nonempty for every $y \in Y$, and it has dimension $\geq n_0$ by Theorem 1.25, (i), so that $f_0^{-1}(y) = f^{-1}(y)$. Therefore $X_0 = X$. The theorem is proved. □

A very special case of Theorem 1.26 is the irreducibility of a product of irreducible projective varieties; see Theorem 1.6.

6.4 Lines on Surfaces

It is only natural, after the effort spent on the proof of Theorems 1.22–1.24 on the dimension of intersections, to look for some applications of these results. As an example, we now treat a simple question on lines on surfaces in $\mathbb{P}^3$.

As a general rule, the notion of dimension is useful in cases when we need to give rigorous meaning to a statement that some set depends on a given number of parameters. For this, we must identify the set with some algebraic variety, and apply the notion of dimension we have introduced.

For example, we have seen in Example 1.28 that hypersurfaces of $\mathbb{P}^n$, defined by equations of degree m, are in one-to-one correspondence with points of a projective space

$$\mathbb{P}^N, \quad \text{where } N = \nu_{n,m} = \binom{n+m}{m} - 1.$$

We proceed to subvarieties that are not hypersurfaces, the simplest of which are lines in $\mathbb{P}^3$. In Example 1.24, we saw that lines $l \subset \mathbb{P}^3$ are in one-to-one correspondence with points of the quadric hypersurface of $\Pi \subset \mathbb{P}^5$ defined by $p_{01}p_{23} - p_{02}p_{13} + p_{03}p_{12} = 0$. Obviously $\dim \Pi = 4$.

To study lines lying on surfaces, the following result is important.

Lemma *The conditions that the line l with Plücker coordinates p_{ij} be contained in the surface X with equation $F = 0$ are algebraic relations between the p_{ij} and the coefficients of F, homogeneous in both the p_{ij} and the coefficients of F.*

Proof We can write a parametric representation of l it terms of its Plücker coordinates: let x and y be a basis of a plane $\mathcal{L} \subset V$, with $\dim \mathcal{L} = 2$, $\dim V = 4$. Then it is easy to check that as f runs through the space of all linear forms on V, the set of vectors of the form

$$xf(y) - yf(x) \tag{1.51}$$

coincides with $\mathcal{L}$. If f has coordinates $(\alpha_0, \alpha_1, \alpha_2, \alpha_3)$, that is, if $f(x) = \sum \alpha_i x_i$, then the vector (1.51) has coordinates $z_i = \sum_j \alpha_j p_{ij}$, where $p_{ij} = x_i y_j - x_j y_i$. Hence if l is the line with Plücker coordinates p_{ij}, the points of l are the points with coordinates $\sum_j \alpha_j p_{ij}$ for $j = 0, \dots, 3$.

On substituting these expressions into the equation $F(u_0, u_1, u_2, u_3) = 0$ and equating to zero the coefficients of all the monomials in α_i, we get the condition that $l \subset X$, as a set of algebraic relations between the coefficients of F and the Plücker coordinates p_{ij}. The lemma is proved. □

We proceed to the question we are interested in, the lines lying on surfaces in $\mathbb{P}^3$. For given m, consider the projective space $\mathbb{P}^N$ with $N = \nu_{3,m} = \binom{m+3}{3} - 1$, whose points parametrise surfaces in $\mathbb{P}^3$ of degree m, that is, given by a homogeneous equation of degree m. Write $\Gamma_m \subset \mathbb{P}^N \times \Pi$ for the set of pairs $(\xi, \eta) \in \mathbb{P}^N \times \Pi$ such that the line l corresponding to $\eta \in \Pi$ is contained in the surface X corresponding to $\xi \in \mathbb{P}^N$. By the lemma, Γ_m is a projective variety. Let us determine the dimension of Γ_m. For this, consider the projection maps $\varphi\colon \mathbb{P}^N \times \Pi \to \mathbb{P}^N$ and $\psi\colon \mathbb{P}^N \times \Pi \to \Pi$ given by $\varphi(\xi, \eta) = \xi$ and $\psi(\xi, \eta) = \eta$. Obviously φ and ψ are regular maps. From now on, we only consider their restrictions to Γ_m. Note that $\psi(\Gamma_m) = \Pi$. This simply means that for every line l there is at least one surface of degree m passing through l, possibly reducible.

We determine the dimension of the fibres $\psi^{-1}(\eta)$ of ψ. By a projective transformation we can assume that the line corresponding to η is given by $u_0 = u_1 = 0$. Points $\xi \in \mathbb{P}^N$ such that $(\xi, \eta) \in \psi^{-1}(\eta) \subset \Gamma_m$ correspond to surfaces of degree m passing through this line. Such a surface is given by $F = 0$, where $F = u_0 G + u_1 H$, with G and H arbitrary forms of degree $m - 1$. The set of such forms is of course a linear subspace of $\mathbb{P}^N$ whose dimension is easy to calculate. It is equal to

$$\mu = \frac{m(m+1)(m+5)}{6} - 1. \tag{1.52}$$

Thus

$$\dim \psi^{-1}(\eta) = \frac{m(m+1)(m+5)}{6} - 1 = N - (m+1).$$

It follows from Theorem 1.26 that Γ_m is irreducible. Applying Theorem 1.25 we get that

$$\begin{aligned} \dim \Gamma_m &= \dim \psi(\Gamma_m) + \dim \psi^{-1}(\eta) \\ &= \frac{m(m+1)(m+5)}{6} + 3 \\ &= N + 3 - m. \end{aligned} \tag{1.53}$$

Consider now the other projection $\varphi\colon \Gamma_m \to \mathbb{P}^N$. Its image is a closed subset of $\mathbb{P}^N$, by Theorem 1.10. Obviously $\dim \varphi(\Gamma_m) \le \Gamma_m$. Thus if $\dim \Gamma_m < N$ then $\varphi(\Gamma_m) \ne \mathbb{P}^N$, or in other words, not every surface of degree m contains a line. By (1.53), the inequality $\dim \Gamma_m < N$ reduces to $m > 3$. We have obtained the following result.

Theorem 1.27 *For any $m > 3$, there exist surfaces of degree m that do not contain any lines. Moreover, such surfaces correspond to an open set of $\mathbb{P}^N$.*

Thus there exist nontrivial algebraic relations between the coefficients of a form $F(u_0, u_1, u_2, u_3)$ of degree $m > 3$ that are necessary and sufficient for the surface given by $F = 0$ to contain a line.

Of the remaining cases $m = 1, 2, 3$, the case $m = 1$ is trivial. We consider the case $m = 2$, although we already know the answer from 3-dimensional coordinate geometry. When $m = 2$ we have $N = 9$ and $\dim \Gamma_m = 10$. It follows from Theorem 1.25 that $\dim \varphi^{-1}(\xi) \ge 1$. This is the well-known fact that any quadric surface contains infinitely many lines.

We remark in passing, and without details of the proof, that this already provides an example of the phenomenon mentioned in Section 6.3 of the dimension of fibres jumping up: if the quadric surface corresponding to a point ξ is irreducible then $\dim \varphi^{-1}(\xi) = 1$, whereas if it splits as a pair of planes then of course $\dim \varphi^{-1}(\xi) = 2$.

Now consider the case $m = 3$. In this case, $\dim \Gamma_m = N = 19$. It is easy to construct a cubic surface $X \subset \mathbb{P}^3$ which contains only a finite number of lines. For example, if X is given in inhomogeneous coordinates by

$$T_1 T_2 T_3 = 1, \tag{1.54}$$

then X does not have a single line contained in $\mathbb{A}^3$. Indeed, if we write the equation of an affine line in the form $T_i = a_i t + b_i$ for $i = 1, 2, 3$ and substitute in (1.54), we get a contradiction; whereas the intersection of X with the plane at infinity contains 3 lines. Thus there exists a point of $\mathbb{P}^{19}$ for which $\varphi^{-1}(\xi)$ is nonempty and $\dim \varphi^{-1}(\xi) = 0$. By Theorem 1.25, this is only possible if $\dim \varphi(\Gamma_3) = 19$. Using Theorem 1.19, we see that $\varphi(\Gamma_3) = \mathbb{P}^{19}$. We have proved the following result.

Theorem 1.28 *Every cubic surface contains at least one line. There exists an open subset U of the space $\mathbb{P}^{19}$ parametrising all cubic surfaces such that a surface corresponding to a point of U contains only finitely many lines.*

Cubic surfaces that contain infinitely many lines do exist, for example cubic cones. Thus again the dimension of fibres can jump up. We will see later that most cubic surfaces contain only finitely many lines, and we will determine the number of these.

6.5 Exercises to Section 6

1 Let $L \subset \mathbb{P}^n$ be an $(n-1)$-dimensional linear subspace, $X \subset L$ an irreducible closed variety and y a point in $\mathbb{P}^n \setminus L$. Join y to all points $x \in X$ by lines, and denote by Y the set of points lying on all these lines, that is, the cone over X with vertex y. Prove that Y is an irreducible projective variety and $\dim Y = \dim X + 1$.

2 Let $X \subset \mathbb{A}^3$ be the reducible curve whose components are the 3 coordinate axes. Prove that the ideal $\mathfrak{A}_X$ cannot be generated by 2 elements.

3 Let $X \subset \mathbb{P}^2$ be the reducible 0-dimensional variety consisting of 3 points not lying on a line. Prove that the ideal $\mathfrak{A}_X$ cannot be generated by 2 elements.

4 Prove that any finite set $S \subset \mathbb{A}^2$ can be defined by two equations. [Hint: Choose the coordinates x, y in $\mathbb{A}^2$ in such a way that all points of S have different x coordinates; then show how to define S by the two equations $y = f(x)$, $\prod(x - \alpha_i) = 0$, where $f(x)$ is a polynomial.]

5 Prove that any finite set of points $S \subset \mathbb{P}^2$ can be defined by two equations.

6 Let $X \subset \mathbb{A}^3$ be an algebraic curve, and x, y, z coordinates in $\mathbb{A}^3$; suppose that X does not contain a line parallel to the z-axis. Prove that there exists a nonzero polynomial $f(x, y)$ vanishing at all points of X. Prove that all such polynomials form a principal ideal $(g(x, y))$, and that the curve $g(x, y) = 0$ in $\mathbb{A}^2$ is the closure of the projection of X onto the (x, y)-plane parallel to the z-axis.

7 We use the notation of Exercise 6. Suppose that $h(x, y, z) = g_0(x, y)z^n + \cdots + g_n(x, y)$ is the irreducible polynomial of smallest positive degree in z contained in the ideal $\mathfrak{A}_X$. Prove that if $f \in \mathfrak{A}_X$ has degree m as a polynomial in z, then we can write $fg_0^m = hU + v(x, y)$, where $v(x, y)$ is divisible by $g(x, y)$. Deduce that the equation $h = g = 0$ defines a reducible curve consisting of X together with a finite number of lines parallel to the x-axis, defined by $g_0(x, y) = g(x, y) = 0$.

8 Use Exercises 6–7 to prove that any curve $X \subset \mathbb{A}^3$ can be defined by 3 equations.

9 By analogy with Exercises 6–8, prove that any curve $X \subset \mathbb{P}^3$ can be defined by 3 equations.

10 Let $F_0(x_0, \ldots, x_n), \ldots, F_n(x_0, \ldots, x_n)$ be forms of degree $m_0, \ldots, m_n$ and consider the system of $n+1$ equations in $n+1$ variables $F_0(x) = \cdots = F_n(x) = 0$. Write Γ for the subset of $\prod_{i=0}^n \mathbb{P}^{\nu_{n,m_i}} \times \mathbb{P}^n$ (where $\nu_{n,m} = \binom{n+m}{m} - 1$) defined by

$$\Gamma = \big\{(F_0, \ldots, F_n, x) \mid F_0(x) = \cdots = F_n(x) = 0\big\}.$$

By considering the two projection maps $\varphi\colon \varGamma \to \prod_i \mathbb{P}^{\nu_{n,m_i}}$ and $\psi\colon \varGamma \to \mathbb{P}^n$, prove that $\dim \varGamma = \dim \varphi(\varGamma) = \sum_i \nu_{n,m_i} - 1$. Deduce from this that there exists a polynomial $R = R(F_0, \dots, F_n)$ in the coefficients of the forms $F_0, \dots, F_n$ such that $R = 0$ is a necessary and sufficient condition for the system of $n+1$ equations in $n+1$ variables to have a nonzero solution. What is the polynomial R if the forms $F_0, \dots, F_n$ are linear?

11 Prove that the Plücker hypersurface $\varPi \subset \mathbb{P}^5$ contains two systems of 2-dimensional linear subspaces. A plane of the first system is defined by a point $\xi \in \mathbb{P}^3$ and consists of all points of $\varPi$ corresponding to lines $l \subset \mathbb{P}^3$ through ξ. A plane of the second system is defined by a plane $\varXi \subset \mathbb{P}^3$ and consists of all points of $\varPi$ corresponding to lines $l \subset \mathbb{P}^3$ contained in $\varXi$. There are no other planes contained in $\varPi$.

12 Let $F(x_0, x_1, x_2, x_3)$ be an arbitrary form of degree 4. Prove that there exists a polynomial $\varPhi$ in the coefficients of F such that $\varPhi(F) = 0$ is a necessary and sufficient condition for the surface $F = 0$ to contain a line.

13 Let $Q \subset \mathbb{P}^3$ be an irreducible quadric surface and $\varLambda_X \subset \varPi$ the set of points on the Plücker hypersurface $\varPi \subset \mathbb{P}^5$ corresponding to lines contained in Q. Prove that $\varLambda_X$ consists of two disjoint conics.

http://www.springer.com/978-3-642-37955-0

Basic Algebraic Geometry 1
Varieties in Projective Space
Shafarevich, I.R.
2013, XVIII, 310 p., Hardcover
ISBN: 978-3-642-37955-0