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

\[
f(x, y) = 0,
\]

(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, © Springer-Verlag Berlin Heidelberg 2013
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 \( \tilde{u}, \tilde{v} \in k(y)[x] \) such that \( f\tilde{u} + g\tilde{v} = 1 \). Multiplying this equality through by the common denominator \( a \in k(y) \) of all the coefficients of \( \tilde{u} \) and \( \tilde{v} \), we obtain an equation 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 \( \tilde{u}, \tilde{v} \in k(y)[x] \) such that \( f\tilde{u} + g\tilde{v} = 1 \). Multiplying this equality through by the common denominator \( a \in k(y) \) of all the coefficients of \( \tilde{u} \) and \( \tilde{v} \) gives \( fu + gv = a \), where \( u = a\tilde{u}, v = a\tilde{v} \in k[x, y], \) 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 \( \beta \). For each such value, the first coordinate \( \alpha \) 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 \( \alpha \). The lemma is proved. \( \square \)

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
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. $k = \mathbb{Q}$. 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 $x^n + y^n = 1$. 

rajas@math.tamu.edu
(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 \( f(x, y) \equiv 0 \mod p \).

(3) \( k = \mathbb{C}(z) \). 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

\[
y^2 = x^2 + x^3
\]  

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

\[
x = t^2 - 1, \quad y = t(t^2 - 1),
\]  

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 \( x \neq 0 \)).
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

$$f(\varphi(t), \psi(t)) \equiv 0,$$

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 $\varphi$ or $\psi$ 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 $\varphi$, $\psi$, 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 $\varphi$ and $\psi$ 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 $\varphi$ and $\psi$ 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
can be expressed in elementary functions. Indeed, since the curve is rational, it can be parametrised as \( x = \varphi(t), \ y = \psi(t) \) where \( \varphi, \ \psi \) are rational functions. Substituting these expressions in the integral (1.5), we reduce it to the form
\[
\int g(\varphi(t), \psi(t))\varphi'(t)dt,
\]
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
\[
y - y_0 = t(x - x_0).
\]
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 \( x \)
\[
f(x, y_0 + t(x - x_0)) = 0,
\]
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 $\int g(x, \sqrt{ax^2 + bx + c}) \, dx$ 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 $P$. 

rojas@math.tamu.edu
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 $\varphi$ and $\psi$ 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 $k(t)$. 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

\[
x = \varphi(\chi(x, y)), \quad y = \psi(\chi(x, y)),
\]

\[
t = \chi(\varphi(t), \psi(t)).
\]

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. $\square$

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 φ(P) ∈ Y for every P ∈ X at which φ is defined. If Y has the equation g = 0 then g(u, v) ∈ k(X) must vanish at all but finitely many points of X, and therefore we must have g(u, v) = 0 ∈ 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 φ : X → 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 ψ : Y → X such that φ ◦ ψ and ψ ◦ φ 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 ∈ Y. Taking any point Q′ ̸= Q at which the inverse ψ of φ is defined contradicts the definition.

It follows that for any point Q ∈ Y the inverse image φ−1(Q) of Q (the set of points P ∈ X such that φ(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 φ : X → Y is not defined, U = X \ S its complement, and T and V the same for ψ : Y → X. It follows from what we said above that the complement in X of φ−1(V) ∩ U and in Y of ψ−1(U) ∩ V are finite, and φ establishes a one-to-one correspondence between φ−1(V) ∩ U and ψ−1(U) ∩ 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 = un−2 + un−1 + un, 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) = un(1, t), b(t) = un−1(1, t) and c(t) = un−2(1, t). Setting s = 2ax + b to complete the square (assuming that the ground field has characteristic ̸= 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 − α) and dividing both sides of the equation through by (t − α)^2m shows that the curve is birational to the curve given by

\[ η^2 = h(ξ), \quad \text{where} \quad ξ = \frac{1}{t − α}, \quad η = \frac{s}{(t − α)^m} \quad \text{and} \quad h(ξ) = \frac{q(t)}{(t − α)^{2m−1}}, \]

in which h is a polynomial of degree ≤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 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 \( \leq 3 \). If \( f(x) \) has degree \( \leq 2 \) 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 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

\[
(u, v) = (\varphi(x, y), \psi(x, y)) \quad \text{and} \quad (x, y) = (\xi(u, v), \eta(u, v)).
\]

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 \( \varphi, \psi \) and \( \xi, \eta \) 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 char \( k = 0 \), that \( f \in k \), and, if char \( k = p > 0 \), that \( f \) involves \( x \) and \( y \) only as \( p \)th powers; in this last case, taking \( p \)th roots of the coefficients of \( f \) and using the well-known characteristic \( p \) identity \((\alpha + \beta)^p = \alpha^p + \beta^p\), we deduce that

\[
f = \sum a_{ij}x^{p_i}y^{p_j} = \left( \sum b_{ij}x^{i}y^{j} \right)^p \quad \text{where} \quad b_{ij}^p = a_{ij},
\]

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 \( f = ax + \beta y + g \), where \( g \) contains only terms of degree \( \geq 2 \) 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 \( p(x, y) = p(x, xv) = x^r \) (because \( p \) has no constant term), where \( r \) is a regular function on the curve, and hence \( u = x^r/q = xu_1 \). 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: \( x^k(x^{1-k}w - a_0) = 0 \) 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 \( f_\gamma'(P) \neq 0 \) 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 poly-
nomial is equal to its degree is false unless we count roots with their multiplicities.
Here we analyse intersection multiplicities in the case that \( \mathcal{X} \) is a line.

Let \( P = (\alpha, \beta) \in \mathcal{X} \), and suppose that the equation of \( \mathcal{X} \) 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 \( \geq 2 \). We write the equation of a line \( L \) through \( P \) in the form
\[
x = \alpha + \lambda t, \quad y = \beta + \mu t.
\]
(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 \( \geq 1 \) with \( \mathcal{X} \) 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 \( a = \frac{\partial f}{\partial x}(P) \), \( b = \frac{\partial f}{\partial y}(P) \), and
hence this equation can we expressed
\[
f_x(P)(x - \alpha) + f_y(P)(y - \beta) = 0.
\]
(1.13)
The line given by this equation is called the tangent line to \( \mathcal{X} \) at the nonsingular
point \( P \).

We now determine when a line has intersection multiplicity \( \geq 3 \) with a curve at a
nonsingular point \( P = (\alpha, \beta) \). For this, we write the equation in the form
\[
f(x, y) = a(x - \alpha) + b(y - \beta) + c(x - \alpha)^2 + d(x - \alpha)(y - \beta) + e(y - \beta)^2 + h,
\]
(1.14)
where \( h \) is a polynomial which has only terms of degree \( \geq 3 \) when expanded in
power of \( x - \alpha \) and \( y - \beta \). Restricting \( f \) to the line \( L \) given by (1.12), we get
that \( f = (a\lambda + b\mu)t + (c\lambda^2 + d\lambda\mu + e\mu^2)t^2 + t^3 \varphi(t) \). Therefore the intersection
multiplicity will be \( \geq 3 \) 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 \( \mathcal{X} \) 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 \( \geq 3 \) at \( P \).
Such a point is called an inflexion point or flex of \( \mathcal{X} \).

We know from coordinate geometry the condition for a conic to be reducible. We assume that \( k \) has characteristic \( \neq 2 \); then recalling that \( a = f'_x(P) \), \( b = f'_y(P) \),
\( f_x(P), f_y(P) \) are not zero. If the condition \( c\lambda^2 + d\lambda\mu + e\mu^2 = 0 \) are not zero. If the condition \( c\lambda^2 + d\lambda\mu + e\mu^2 = 0 \) holds, then the
conic \( q = au + bv + cu^2 + duv + ev^2 \) is reducible, and one of the factors, 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 \( \geq 3 \) at \( P \).

rojas@math.tamu.edu
\[ c = (1/2) f''_x(P), \quad d = f''_{xy}(P) \quad \text{and} \quad e = (1/2) f''_{yy}(P), \]
we can write this condition in the form
\[
\begin{vmatrix}
  f''_x & f''_{xy} & f'_x \\
  f''_{xy} & f''_{yy} & f'_y \\
  f'_x & f'_y & 0
\end{vmatrix}(P) = 0. \tag{1.15}
\]

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 \((\xi, \eta, \zeta)\) of the field \(k\), not all simultaneously zero. Two triples \((\xi, \eta, \zeta)\) and \((\xi', \eta', \zeta')\) 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 \((\xi, \eta, \zeta)\) 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 \(\zeta\). 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

rojas@math.tamu.edu
defines a homogeneous polynomial $F(\xi, \eta, \zeta)$ requiring the conditions

coordinates also called a form it has coordinates $x \in P^1$, $y \in P^1$, $z \in P^1$. Every point $P \in P^2$ is contained in at least one of the pieces $A^2_1$, $A^2_2$ or $A^2_3$, and can be written down in the affine coordinates of that piece.

An algebraic curve in $P^2$, 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 $\xi, \eta, \zeta$ 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 $F(\xi, \eta, \zeta) = \zeta^n f(\frac{\xi}{\zeta}, \frac{\eta}{\zeta})$, 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 $A^2_3$ 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 $P \in P^2$ is contained in one of the affine sets $A^2_1$, $A^2_2$ or $A^2_3$, 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 $P \in A^2_3$.

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 $F'\xi$ etc. for the partial derivatives, we get $\frac{\partial f}{\partial x} = F'\xi(x, y, 1)$ and $\frac{\partial f}{\partial y} = F'\eta(x, y, 1)$, 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 $f'_x = f'_y = 0$. 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 $f'_x = F'_x, f'_y = F'_y, f''_x = F''_x, f''_y = F''_y, f'''_x = F'''_x, f'''_y = F'''_y$. From now on, in the homogeneous polynomial $F$ we write $\xi$ for $x$ and $\eta$
for \( y \). We substitute these expressions in the determinant of (1.15), and use Euler’s theorem

\[
F''_{\xi \xi} \xi + F''_{\xi \eta} \eta + F''_{\xi \zeta} \zeta = (n - 1) F'_\xi,
\]

\[
F''_{\xi \eta} \xi + F''_{\eta \eta} \eta + F''_{\eta \zeta} \zeta = (n - 1) F'_\eta,
\]

\[
F'_\xi \xi + F'_\eta \eta + F'_\zeta \zeta = nF.
\]

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

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

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

\[
\begin{vmatrix}
F''_{\xi \xi} & F''_{\xi \eta} & F''_{\xi \zeta} \\
F''_{\xi \eta} & F''_{\eta \eta} & F''_{\eta \zeta} \\
F'_\xi & F'_\eta & F'_\zeta
\end{vmatrix}
(P) = 0.
\]

(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 \( \mathbb{A}^2 \) 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 \( \varphi : \mathbb{A}^2 \to \mathbb{A}^2 \) defined by \((x, y) \mapsto (u(x, y), v(x, y))\), we first rewrite it, as just explained, in the form

\[
\begin{vmatrix}
U(\xi, \eta, \zeta) & V(\xi, \eta, \zeta) \\
R(\xi, \eta, \zeta) & S(\xi, \eta, \zeta)
\end{vmatrix},
\]

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

\[
(\xi : \eta : \zeta) \mapsto \left( A(\xi : \eta : \zeta) : B(\xi : \eta : \zeta) : C(\xi : \eta : \zeta) \right),
\]

where \( A, B, C \) are homogeneous polynomials of the same degree. Now \( \varphi \) 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 \( A^2 \), say, we can divide each of \( A, B, C \) by \( \zeta^n \), 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 \( A^2 \) 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 \( u_i = r^i v_i \), where \( r \) 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 \((x, y) \mapsto (v_0 : r^{k_1-k_0} v_1 : r^{k_2-k_0} v_2)\), 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. \( \square \)

**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 \( \alpha \) 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 \( \eta^3 \) we write the equation of the curve in the form \( \nu = u^3 + pu v^2 + q v^3 \), 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 $\xi = 0$ 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 $x = \xi/\xi$, $y = \eta/\xi$, 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 $\xi = 0$ 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 $x = \xi/\xi$, $y = \eta/\xi$ (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 $\xi = 0$ 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 $f_\lambda = l_1l_2l_3 + \lambda m_1m_2m_3$ where $\lambda$ 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 $\lambda$ so that $f_\lambda(P) = 0$ for
any given point \( P \in X \), distinct from these 6 points of intersection. We get a cubic \( f_\lambda \) 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 $P_1, \ldots, P_r$ of an irreducible curve and numbers $m_1, \ldots, m_r \geq 0$ 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 $x_0^3 + x_1^3 + x_2^3 + mx_0x_1x_2 = 0$ 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 $\alpha = (\alpha_1, \ldots, \alpha_n)$ 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.