Chapter 2
Curves in \( \mathbb{A}^2_k \) and in \( \mathbb{P}^2_k \)

In this chapter we introduce the first interesting class of planar curves, namely the conic sections. This leads to a first discussion of singular and non singular points. Closely tied to these concepts is the notion of the tangent at a point on a curve. We then move on to a discussion of curves of higher degrees, and introduce the concepts of tangent lines, the tangent cone and the multiplicity of a point on a curve which may have singularities. A number of important examples of higher order curves are discussed. Elliptic curves are briefly discussed, this class of curves (which are certainly not ellipses) played an important role for the fruitful interplay between geometry and function theory, so central in the pathbreaking work of Niels Henrik Abel.

2.1 Conic Sections

A conic section\(^1\) is a curve in the affine plane \( \mathbb{A}^2_k \) of degree 2, over a field which we shall assume to be of characteristic \( \neq 2 \). The general form of the equation is usually written as

\[
q(x, y) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0.
\]

In [27] these curves are treated with \( k = \mathbb{R} \) in some detail, and we refer the reader who is unfamiliar to the basics of this subject to the treatment there, as a suitable basis for reading the present chapter. In particular, this reference gives the proof that all such curves can be obtained as the intersection between a fixed double circular cone and a varying plane, in the case when \( k = \mathbb{R} \).

The non-degenerate conics are the ellipses, the parabolas and the hyperbolas. In addition to these, we have the degenerate cases. In the three non-degenerate cases the equation can be brought on one of the three canonical forms. We refer to [27] for more information on this.

\(^1\)We also say conic curve or just a conic.
In [27] we consider the following problem: There is given 5 distinct points in \( \mathbb{A}^2_k \), \((x_1, y_1), (x_2, y_2), \ldots, (x_5, y_5)\). If these points are in sufficiently general position, then there is a unique, non-degenerate conic curve, in other words a non-degenerate curve of degree 2, passing through them. The condition that the points be in sufficiently general position, here amounts to the requirement that no three of them be collinear. The equation for the conic in question is

\[
\begin{vmatrix}
    x^2 & xy & y^2 & x & y & 1 \\
    x_1^2 & x_1y_1 & y_1^2 & x_1 & y_1 & 1 \\
    x_2^2 & x_2y_2 & y_2^2 & x_2 & y_2 & 1 \\
    x_3^2 & x_3y_3 & y_3^2 & x_3 & y_3 & 1 \\
    x_4^2 & x_4y_4 & y_4^2 & x_4 & y_4 & 1 \\
    x_5^2 & x_5y_5 & y_5^2 & x_5 & y_5 & 1 \\
\end{vmatrix} = 0.
\]

\[2.2 \text{ Singular and Non-singular Points}\]

We need the notion of a non-singular point of a plane curve, we return to a refined treatment of this important concept in Sect. 2.8.

We define the derivative of a polynomial with coefficients form a field \( k \) formally as follows:

**Definition 2.1** Let \( P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \) be a polynomial in \( x \) with coefficients from \( k \). We define the derivative of \( P(x) \) with respect to \( x \) as

\[
P'(x) = \frac{dP}{dx} = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \cdots + a_1.
\]

If \( F(x, y, u, \ldots) \in k[x, y, u, \ldots] \), then the partial derivatives are defined analogously in a formal manner.

The basic properties of derivatives still hold: The derivative of a constant is zero, the formulas for the derivatives of sums and products of polynomials hold, as does the chain rule.

In characteristic 0 we still have the Taylor formula in its usual form in one and several variables, and may proceed as for \( k = \mathbb{R} \). In characteristic \( p > 0 \) the procedure is somewhat modified, this is omitted here. All these observations only apply to polynomials, of course.

We note the following result, which was shown in Sect. 12.7 in [27] for \( k = \mathbb{R} \). It will be given a different proof in Sect. 2.3.

**Theorem 2.1** The equation

\[
q(x, y) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0
\]
yields a non-degenerate conic curve if and only if the following determinantal criterion is satisfied:

\[
\begin{vmatrix}
A & B & D \\
B & C & E \\
D & E & F
\end{vmatrix} \neq 0.
\]

We now make the following important definition:

**Definition 2.2** Let \( Z \) be an affine plane curve given by the equation

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

Let \((x_0, y_0)\) be a point on the curve such that the two partial derivatives do not both vanish,

\[
\left( \frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0) \right) \neq (0, 0).
\]

Such a point is called a non-singular point on the curve. At all non-singular points we define the tangent line\(^2\) by the equation

\[
\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) = 0.
\]

A point which is not non-singular is called a singular point.

We next turn to the tangents of non-degenerate conics in \( \mathbb{A}_k^2 \), as well as the related concepts of *pole and polar line*.

Let \( P = (x_0, y_0) \) be a point on the non-singular conic curve given by the equation

\[ q(x, y) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0, \]

the tangent at \( P \) is given by

\[ (Ax_0 + By_0 + D)(x - x_0) + (Bx_0 + Cy_0 + E)(y - y_0) = 0, \]

which after a short calculation takes the form

\[ Ax_0x + B(y_0x + x_0y) + C y_0y + D(x + x_0) + E(y + y_0) + F = 0. \]

This equation is also of interest when the point is not on \( C \). We have the following result:

---

\(^2\)This concept will be explained in more detail in Sect. 2.9.
Proposition 2.2 Let $P = (x_0, y_0)$ be a point and let $\mathcal{C}$ be the conic given by the equation

$$ q(x, y) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0. $$

There are two tangent lines to $\mathcal{C}$ passing through $P$, coinciding if $P$ is on $\mathcal{C}$. Let the two points of tangency be $Q_1$ and $Q_2$. Then the line $p$ passing through $Q_1$ and $Q_2$ is given by the equation

$$ Ax_0x + B(y_0x + x_0y) + Cy_0y + D(x + x_0) + E(y + y_0) + F = 0. $$

Proof The situation is shown in Fig. 2.1.

Let $Q_1 = (x_1, y_1)$ and $Q_2 = (x_2, y_2)$, then the two tangents in question will have equations

$$ Ax_1x + B(y_1x + x_1y) + Cy_1y + D(x + x_1) + E(y + y_1) + F = 0, $$

$$ Ax_2x + B(y_2x + x_2y) + Cy_2y + D(x + x_2) + E(y + y_2) + F = 0. $$

These lines pass through $P = (x_0, y_0)$, thus

$$ Ax_1x_0 + B(y_1x_0 + x_1y_0) + Cy_1y_0 + D(x_0 + x_1) + E(y_0 + y_1) + F = 0, $$

$$ Ax_2x_0 + B(y_2x_0 + x_2y_0) + Cy_2y_0 + D(x_0 + x_2) + E(y_0 + y_2) + F = 0. $$

But this demonstrates that the line whose equation is given in the assertion of the proposition, does indeed pass through the two points $Q_1$ and $Q_2$. Hence the claim follows. \(\square\)

Definition 2.3 The point $P$ and the line $p$ in Proposition 2.2 are called the pole and the polar line corresponding to each other.

In the case when the field $k$ is not algebraically closed, such as when $k = \mathbb{R}$, we encounter some apparently puzzling phenomena. For example, are the “conic sections” given by the equations $x^2 + y^2 + 1 = 0$ and $x^2 + y^2 = 0$
really curves? The former has no points in $\mathbb{A}^2_{\mathbb{R}}$, while the latter has only the origin as a point on it. According to Theorem 2.1 the former is non-degenerate while the latter is degenerate. An explanation for this apparent paradox is that we need to consider not only points over $k$, but also points over the algebraic closure $\overline{k}$ in order to understand an algebro-geometric object such as an affine curve.

Moreover, if we take a point inside an ellipse, then there will be no real points of tangency, even though we get a well defined polar line using the equation we have derived in Proposition 2.2. But if we compute the complex points of tangency, we find that corresponding coordinates are complex conjugates, and we get a real line joining them.

Finally, if we choose the center of the unit circle, say, then the “line” given by the formula is just $0 = 1$, which has no points on it. The explanation for this is that the polar of the center is the line at infinity. Thus we see here both the need for computing complex points as well as for considering points at infinity in order to understand algebraic curves over $\mathbb{R}$.

2.3 Conics in the Projective Plane

We shall now consider the so-called projective closure of the conics in $\mathbb{A}^2_k$. We substitute

$$x = \frac{X_1}{X_0}, \quad y = \frac{X_2}{X_0}$$

into the equation of the conic $\mathcal{C}$,

$$q(x, y) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0,$$

which yields the following homogeneous equation

$$Q(X_0, X_1, X_2) = AX_1^2 + 2BX_1X_2 + CX_2^2 + 2DX_0X_1 + 2EX_0X_2 + FX_0^2 = 0.$$ 

Hence we get the equation of a curve $\overline{\mathcal{C}}$ in $\mathbb{P}^2_k$, which we refer to as the projective closure of $\mathcal{C}$. When intersected with $D_+(X_0) = \mathbb{A}^2_k$ it gives back the original curve.

We first wish to determine its points at infinity. Those are the points $\overline{\mathcal{C}} \cap V_+(X_0)$. The point $(u : v : 0)$ is in $\overline{\mathcal{C}}$ if

$$Au^2 + 2Buv + Cv^2 = 0,$$

and we immediately get the following information:
Proposition 2.3 1. \( C \) has no real points at infinity if \( B^2 - AC < 0 \).

2. \( C \) has one real point at infinity if \( B^2 - AC = 0 \).

3. \( C \) has two points at infinity if \( B^2 - AC > 0 \).

Thus 1. corresponds to a possibly degenerate ellipse, 2. to a possibly degenerate parabola and 3. to a possibly degenerate hyperbola.

In general, let \( C \) be the curve in \( \mathbb{P}_k^2 \) defined by \( F(X_0, X_1, X_2) = 0 \).

Let \( P = (a_0 : a_1 : a_2) \) be a point on it. In Chap. 3, Sect. 3.4 we show that the equation

\[
\frac{\partial F}{\partial X_0}(a_0, a_1, a_2)X_0 + \frac{\partial F}{\partial X_1}(a_0, a_1, a_2)X_1 + \frac{\partial F}{\partial X_2}(a_0, a_1, a_2)X_2 = 0
\]

yields the tangent line to \( C \) at \( P \), provided that the coefficients involved do not all vanish.

Definition 2.4 If the partial derivatives involved in the equation above all vanish at some point on the curve, then the point is said to be a singular point. If they do not all vanish, the point is called non-singular.

The equation for the tangent to the conic curve in \( \mathbb{P}_k^2 \) given by the equation \( Q(X_0, X_1, X_2) = 0 \) at the point \( P = (x_0, x_1, x_2) \) is

\[
(Ax_1 + Bx_2 + Dx_0)X_1 + (Bx_1 + Cx_2 + Ex_0)X_2 + (Dx_1 + Ex_2 + Fx_0)X_0 = 0
\]

or written on a more appealing form

\[
Ax_1X_1 + B(x_1X_2 + x_2X_1) + Cx_2X_2 + D(x_0X_1 + x_1X_0) + E(x_0X_2 + x_2X_0) + Fx_0X_0 = 0.
\]

This is similar to what we found in the affine case.

If the point \( P \) is singular, then its projective coordinates constitute a non-trivial solution of the following homogeneous system of equations:

\[
Au + Bv + Dw = 0 \\
Bu + Cv + Ew = 0 \\
Du + Ev + Fw = 0
\]

and thus we have in this case

\[
\begin{vmatrix}
A & B & D \\
B & C & E \\
D & E & F
\end{vmatrix} = 0.
\]
But the argument works both ways, thus the determinant above vanishes if and only if the conic section has a singular point, at least over $\overline{k}$. On the other hand, if the conic section has such a singular point, then passing to $\overline{k}$ and switching to a suitable projective coordinate system, we may assume that the singular point is $(1 : 0 : 0)$. But then $D = E = F = 0$, thus the equation of the conic curve is $Ax^2 + Bxy + Cy^2$ in $D_+(X_0) = \mathbb{P}^2_k$. Since this polynomial splits as a product of linear forms in $x$ and $y$, we have proved the theorem stated below:

**Theorem 2.4** Assume that $k = \overline{k}$. The following are equivalent:

1. The equation
   $$q(x, y) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$
   yields a non-degenerate conic section.
2. The projective closure in $\mathbb{P}_k^2$ of the curve in $\mathbb{A}_k^2$ given by $q(x, y) = 0$ is non-singular.
3. $$\begin{vmatrix} A & B & D \\ B & C & E \\ D & E & F \end{vmatrix} \neq 0.$$ 

We finally note the following result:

**Theorem 2.5** Let $k = \overline{k}$. Assume that $AX_1^2 + 2BX_1X_2 + CX_2^2 + 2DX_0X_1 + 2EX_0X_2 + FX_0^2 = 0$ is the equation of a non-singular conic in $\mathbb{P}_k^2$. Then we may choose the projective coordinate system such that $B = D = E = 0$ and $A = C = E = 1$.

**Proof** By Proposition 1.1 we may assume that the following two points lie on the conic:

$$(0 : i : 1) \quad \text{and} \quad (0 : -i : 1)$$

where $i = \sqrt{-1} \in k$, as $k = \overline{k}$. Thus

$$-A + 2Bi + C = 0 \quad \text{and} \quad -A - 2Bi + C = 0$$

from which it follows that $B = 0$, hence $A = C$. Since the conic is non-degenerate, we must have $A = C \neq 0$ so we may assume $A = C = 1$, and the equation becomes

$$X_1^2 + X_2^2 + 2DX_0X_1 + 2EX_0X_2 + FX_0^2 = 0.$$

Evidently this is transformed as follows by completing two squares:

$$(X_1 + DX_0)^2 + (X_2 + EX_0)^2 + (F - D^2 - E^2)X_0^2 = 0.$$
Changing projective coordinate system again if necessary we obtain
\[ X_1^2 + X_2^2 + GX_0^2 = 0 \]

where \( G \neq 0 \) since otherwise the conic would be degenerate. A final change of projective coordinate system yields
\[ X_1^2 + X_2^2 + X_0^2 = 0 \]

and the proof is complete.

For more on conics, including elementary proofs of the theorems of Pappus and Pascal, which we will not include here, we refer to Sects. 12.8 and 12.9 in [27] or Sects. 13.8 and 13.9 in [28].

2.4 The Cubic Curves in \( \mathbb{A}^2_k \)

The simplest curve of higher degree, by which we mean degree higher than 2, is the curve known as the cubic parabola. The parabola has the equation \( y = x^2 \), after a suitable change of coordinate system in \( \mathbb{A}^2_k \). Classically the term parabola was used in a wider sense, as the name of a curve whose graph would lie “parallel” to the \( y \)-axis.

Thus curves with an equation of the form \( y = x^m \), \( m \) being a positive integer or a rational number, would be called parabolas as well. Accordingly, a curve which may be brought on the from \( y = x^3 \) is referred to as a cubic parabola.

The next step in complexity is a curve which may be brought on the form \( y^2 = x^3 \). It is called a semi-cubic parabola. It has the graph displayed in Fig. 2.2.

The concept of degenerate curves and the related process of degeneration of a family of curves are important.

A curve is said to be degenerate if it decomposes into the union of two or more curves of lower degrees. For a cubic curve this means that it is a union of a conic curve and a line, or of three lines (some possibly coinciding).

Planar curves of degree 3 already constitute a much richer and interesting group of geometrical objects than the ones of degree 2.

The simplest example of a degenerate cubic curve would be the \( y \)-axis with multiplicity 3. Its equation is \( x^3 = 0 \). We have not yet made the notion of curves with multiplicity precise, this comes in Sect. 2.7. But we may already at this point consider a family of semi-cubic parabolas, degenerating to the triple \( y \)-axis. Namely, consider the curves depending on the parameter \( t \), as \( t \to 0: ty^2 = x^3 \).

We show some members of this family in Fig. 2.2. The values of \( t \) in the plots are \( t = 10, 4, 1, 0.1 \).
2.4 The Cubic Curves in $\mathbb{A}^2_k$

Fig. 2.2 The semi-cubic parabola given by $y^2 = x^3$ to the left, to the right we show the degeneration $ty^2 = x^3$ of the semi-cubic parabola to the triple y-axis

Fig. 2.3 The Folium of Descartes, with $a = 1$, the curve is generally given by the equation $x^3 + y^3 = 3axy$, it then turns out that the curve approaches the line $x + y + a = 0$ as an asymptote

Note also that when $t \to \infty$, then the limit is the $x$-axis with multiplicity 2, since this degeneration is equivalent to letting $u$ tend to 0 for the family given by the equation $y^2 = ux^3$.

The term degeneration is used rather loosely, without a formal definition. The idea we intend to convey by this, is to have one curve, say the semi-cubic parabola $y^2 = x^3$, be a member of a family of curves depending on a parameter, all but a finite number of which are of the same type. Then the exceptional members are understood as degenerate cases. This is, of course, the way we may view two intersecting lines as a degenerate hyperbola, or a double line as a degenerate hyperbola or a degenerate parabola, and so on.

Two more types of non-degenerate curves of degree three exist, up to a projective change of coordinate system. We will explain this projective equivalence for curves in $\mathbb{P}^2_k$ (and in $\mathbb{A}^2_k$) later, in Sect. 3.5. The simple affine
equivalence for two curves means that one may be obtained from the other by a suitable affine transformation, or a change of coordinate system in \( \mathbb{A}^2_k \).

This kind of equivalence is more complicated than the projective equivalence, there are more equivalence classes of affine cubic curves under this affine equivalence. But from our point of view, the projective equivalence is more interesting than the affine one.

The first of the remaining classes of cubic curves is represented by the Folium of Descartes. The French mathematician René Descartes, 1596–1650, is credited by some historians of mathematics as being the founder of algebraic geometry. However, this is disputed by others.

Descartes was the first to systematically introduce coordinates and equations into geometry, and our usual coordinate system in the plane is named after him, a Cartesian coordinate system. His name was originally Cartes, and when he was knighted it changed into Des Cartes. Descartes was for some time engaged in a bitter feud with another great French mathematician, Pierre de Fermat, 1601–1665. One of the issues they could not agree on was the proper way to define the tangent to a curve at a given point.
2.4 The Cubic Curves in $\mathbb{A}^2_k$

Fig. 2.6 The usual nodal cubic

We also give another curve, belonging to the same class as the Folium under projective equivalence, but to a separate class under the affine equivalence. It is often referred to as the *usual nodal cubic*. It is given by the equation $y^2 - x^3 - x^2 = 0$. It looks somewhat similar to the semi-cubic parabola. In fact, the latter may be obtained by deforming the former. At this time the tools from calculus needed for what we regard as the proper solution to this question had not yet been sufficiently developed, and to us the methods of both Descartes and Fermat would look strange and clumsy. The curve was given as an example by Descartes in this argument with Fermat.

We also give another curve, belonging to the same class as the Folium under projective equivalence, but to a separate class under the affine equivalence. It looks somewhat similar to the semi-cubic parabola. In fact, the latter may be obtained by deforming the former. This is the simplest and most used example of a *nodal cubic curve* in $\mathbb{A}^2_k$. It is shown in Fig. 2.6. The deformation referred to is obtained from the family $y^2 - x^3 - tx^2 = 0$.

To the left in Fig. 2.7 we see some of the corresponding plots, for $t = 0, 0.5, 2$, degenerating the usual nodal cubic given by $y^2 - x^3 - x^2 = 0$ to the *semi-cubic parabola* with equation $y^2 - x^3 = 0$. To the right an unusual “nodal cubic” given by $y^2 - x^3 + x^2 = 0$. Actually, the origin is on the curve, but that point appears to be isolated from the main part of it.

But there are complex points, invisible in $\mathbb{A}^2_\mathbb{R}$, which establish the connection.

We have now come to a very interesting class of curves. These curves are tied to a real leap forward in mathematics which occurred in the 19th century, and is tied to such mathematical giants as Niels Henrik Abel and Carl Gustav Jacob Jacobi. The ground had been prepared by mathematicians like Leonhard Euler and Adrien-Marie Legendre, who had studied the mysterious so called *elliptic integrals*, occurring when one wanted to compute arc lengths of segments of ellipses and of the *lemniscate*. We have arrived at the concept of an *elliptic curve*. 
2.5 Elliptic Integrals and the Elliptic Transcendentals

The reason for the name elliptic curve is that such curves come up when one attempts to compute arc length for ellipses. The corresponding problem for a circle is quite simple: We represent the circle by the equation $x^2 + y^2 = R^2$. We then have to compute the integral $L = \int_{x_1}^{x_2} \sqrt{1 + y'^2} \, dx$.

Then as $y = \sqrt{R^2 - x^2}$ we find $y' = -\frac{x}{\sqrt{R^2 - x^2}}$ and thus to find the arc length of the circle between two given points corresponding to $x_1$ and $x_2$ in
2.5 Elliptic Integrals and the Elliptic Transcendentals

The first or second quadrant, say, we have to compute the integral

\[ L = \int_{x_1}^{x_2} \sqrt{1 + y'^2} \, dx = \int_{x_1}^{x_2} \frac{R \, dx}{\sqrt{R^2 - x^2}}. \]

In this case we may introduce polar coordinates, \( x = R \cos(\varphi) \) and \( y = R \sin(\varphi) \). Then the integral becomes

\[ L = R \int_{\varphi_1}^{\varphi_2} \frac{-\sin(\varphi)}{\sqrt{1 - \cos^2(\varphi)}} \, d\varphi = -R \int_{\varphi_1}^{\varphi_2} d\varphi = R(\varphi_2 - \varphi_1). \]

However, consider the corresponding problem for the ellipse

\[ \left( \frac{x}{a} \right)^2 + \left( \frac{y}{b} \right)^2 = 1 \]

where \( a > b \). Then the same method applied to \( x = a \cos(\varphi) \), \( y = b \sin(\varphi) \) leads to the integral

\[ L = a \int \sqrt{1 - k^2 \cos^2(\varphi)} \, d\varphi, \]

where \( k = \sqrt{\frac{a^2 - b^2}{a^2}} \) is the eccentricity of the ellipse. Putting \( t = \cos(\varphi) \), this integral is reduced to

\[ I_2 = \int_0^x \frac{\sqrt{1 - k^2 t^2}}{\sqrt{1 - t^2}} \, dt, \]

referred to as an elliptic integral of the second kind. An elliptic integral of the first kind is of the form

\[ I_1 = \int_0^x \frac{1}{\sqrt{(1 - t^2)(1 - k^2 t^2)}} \, dt, \]

while an elliptic integral of the third kind is

\[ I_3 = \int_0^x \frac{1}{(1 + nt^2)\sqrt{(1 - xt^2)(1 - k^2 t^2)}} \, dt. \]

These three forms are referred to as Legendre’s standard forms for elliptic integrals. Before Abel’s (and Jacobi’s) time these integrals, as functions of the upper limit \( x \), were considered as the elliptic functions, the so-called elliptic transcendentals. Abel, and later on Jacobi, turned this around and defined the elliptic functions as the inverse of these integral-functions. Thus, as for instance

\[ I_2 = \int_0^x \frac{dt}{\sqrt{1 - t^2}} = \text{Arccos}(t), \]

...
we find the elliptic function associated to an elliptic integral of the second kind with \( k = 0 \) to be the function \( y = \sin(x) \). So elliptic functions are vast generalizations of the trigonometric functions.

In general an elliptic integral is an integral of the form \( \int_a^x \frac{dt}{f(t, \sqrt{R})} \) where \( f \) is a rational expression in two variables and \( R \) is a cubic or biquadratic expression in \( t \). Legendre succeeded in expressing all such integrals in terms of his normal forms above.

Today one uses the Weierstrass Normal Form,

\[
u = \int_a^x \frac{dt}{\sqrt{4t^3 - g_1t - g_2}}
\]

and we note that the denominator with \( g_1 = 4 \) and \( g_2 = 0 \) gives rise to the equation

\[
y^2 = 4x(x^2 - 1),
\]

giving a curve which is equivalent to the elliptic curve displayed in Fig. 2.9. We explain this in more detail in Sect. 4.12.

### 2.6 More Curves in \( \mathbb{A}_R^2 \)

Before proceeding with the general theory, we shall look at some other examples of curves in \( \mathbb{A}_R^2 \). Some of them have interesting histories, here we shall just present a curve which is due to Colin Maclaurin.

The Trisectrix of Maclaurin is given by the equation

\[
x^3 + xy^2 + y^2 - 3x^2 = 0.
\]
Suppose there is given an angle $u = \angle ABC$. Then the two lines AB and BC are extended to $s$ and $t$, respectively, as shown in Fig. 2.11. A Cartesian coordinate system is introduced, so that $t$ becomes the $x$-axis and the origin is located on $t$ to the left of $B$ at a distance of 2. We then plot the curve given by $x^3 + xy^2 + y^2 - 3x^2 = 0$. This curve intersects the line $s$ in the point $P$, and we draw the line $OP$ between the origin $O$ and $P$. We claim that if $v = \angle POC$, then $u = 3v$. Indeed, it suffices to show that $\sin(u) = \sin(3v)$, in other words that $\sin(u) = 3\sin(v) - 4\sin^3(v)$. Now we have $\sin(u) = \frac{PD}{PB}$ and $\sin(v) = \frac{PD}{PO}$. Moreover, $PD = y$, $PO = \sqrt{x^2 + y^2}$ and $PB = \sqrt{(x-2)^2 + y^2}$. Thus we need to verify the following identity:

$$\frac{y}{\sqrt{(x-2)^2 + y^2}} = 3 \frac{y}{\sqrt{x^2 + y^2}} - 4 \left( \frac{y}{\sqrt{x^2 + y^2}} \right)^3$$
The Clover Leaf Curve has equation \((x^2 + y^2)^2 + 3x^2y - y^3 = 0\) in the presence of the relation \(x^3 + xy^2 + y^2 - 3x^2 = 0\), or equivalently

\[
\sqrt{\frac{x^2 + y^2}{(x-2)^2 + y^2}} = 3 - 4 \frac{y^2}{x^2 + y^2}
\]

i.e.

\[(x^2 + y^2)^3 = ((x - 2)^2 + y^2)(3x^2 - y^2)^2.
\]

An evaluation finally yields

\[
(x^2 + y^2)^3 - ((x - 2)^2 + y^2)(3x^2 - y^2)^2 = -4(-2y^2x + y^2 - 3x^2 + 2x^3)(y^2x + y^2 - 3x^2 + x^3)
\]

from which the claim follows.

Another curve looks like a clover leaf. It has equation \((x^2 + y^2)^2 + 3x^2y - y^3 = 0\) and is shown in Fig. 2.12. According to the picture, the curve is smooth except at the origin. There this curve displays a more complicated behavior, and gives the appearance of being the shadow, or the projection of a knot-like space curve. We shall make these features precise later.

Our first curve of degree higher than three was the Clover Leaf Curve above. Another such interesting curve is the famous Airplane Wing Curve. Amazingly it looks very similar to a section through the wing of an airplane.

We now have a sufficient base of examples to appreciate some more general theory. We already mentioned the need to incorporate complex points in connection with elliptic cubic curves above. In addition to this, a curve may consist of several components. That is to say, it may consist of several curves taken together. And some of these components may also occur with a multiplicity. Thus for instance, a double line is a different curve from a single line. We now take a closer look.
2.7 General Affine Algebraic Curves

Let the curve \( C \) be given by the equation \( f(x,y) = 0 \). We then study the set of pairs \((u,v)\) of complex numbers such that \( f(u,v) = 0 \). So we consider the zero locus of \( f(x,y) = 0 \) in \( \mathbb{A}^2_C \). We may denote this set by \( C(\mathbb{C}) \), and the curve considered as a subset of \( \mathbb{A}^2_{\mathbb{R}} \) we may denote by \( C(\mathbb{R}) \). If we identify \( \mathbb{A}^2_C \) with \( \mathbb{A}^4_{\mathbb{R}} \), this locus is identified with a surface defined by two equations. Namely, writing

\[
\begin{align*}
u &= x_1 + ix_2, \\
v &= x_3 + ix_4
\end{align*}
\]

and

\[
f(u,v) = f_1(x_1,x_2,x_3,x_4) + if_2(x_1,x_2,x_3,x_4)
\]

then \( f_1 \) and \( f_2 \) are polynomials with real coefficients in four variables, and the set of all complex points on the curve is given as

\[
C(\mathbb{C}) = \left\{ (a_1,a_2,a_3,a_4) \in \mathbb{A}^4_{\mathbb{R}} \mid f_1(a_1,a_2,a_3,a_4) = 0, f_2(a_1,a_2,a_3,a_4) = 0 \right\}.
\]

This is a surface in four-space, in \( \mathbb{A}^4_{\mathbb{R}} \), defined by two polynomials. In many situations we really need to include all complex points of a curve, although we usually still confine ourselves to sketch the real points only. And even if the complex points form a surface in \( \mathbb{A}^4_{\mathbb{R}} \), it is important to keep in mind that we really are studying a curve in the plane, and not a surface in four space. Indeed, of we switch to regard our object under study as a surface in \( \mathbb{A}^4_{\mathbb{R}} \), then it will also have complex points, thus yielding a fourfold in \( \mathbb{A}^8_{\mathbb{R}} \), and so on. Thus we have to remember that we are studying complex points on a curve in the plane, rather than the real points of a surface in four space.

The further important extension is to include the points at infinity of a curve. This is a somewhat more technical matter, which we come to in
Chap. 3, where we study projective curves. But first we give more details on the affine case.

We are given a curve in the plane \( \mathbb{R}^2 \) as the set of zeroes of the equation

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

where \( f(x, y) \) is a polynomial in the variables \( x \) and \( y \):

\[
f(x, y) = a_{0, 0} + x a_{1, 0} + y a_{0, 1} + x^2 a_{2, 0} + x y a_{1, 1} + y^2 a_{0, 2} + \cdots + x^d a_{d, 0} + x^{d-1} y a_{d-1, 1} + \cdots + y^d a_{0, d}.
\]

Some, but not all, of the coefficients may be zero. The largest integer \( d \) such that not all \( a_{d-i,i} \) are zero is the degree of the polynomial, and this is by definition the degree of the curve. But here we have a problem, best elucidated by an example.

The equation

\[
y = 0
\]

defines the \( x \)-axis. But so does the equation

\[
y^2 = 0
\]

at least as a point-set. But algebraically we need to distinguish between these two cases. The former equation defines the \( x \)-axis as a line, whereas the latter defines a double line along the \( x \)-axis: Informally speaking, it defines twice the \( x \)-axis.

The situation becomes even more difficult when we consider complicated polynomials. Thus for example we may consider the curve defined by the equation

\[
(y^2 - x^3 - x^2)(y^2 - x^2) = 0.
\]

When we are given the equation on this partly factored form, it is not difficult to see what we get: It is the nodal cubic curve displayed in Fig. 2.6 together with the two lines defined by \( y = \pm x \). But suppose that we are given the following equation, on expanded form

\[
3y^2 x^4 - 3y^4 x^2 + y^6 - x^7 + 2x^5 y^2 - x^3 y^4 - x^6 = 0
\]

then it is not so easy to understand the situation. Using some PC-program to plot this curve, we should get the same picture as above. But this result is quite deceptive. Indeed, if we factor the left hand side of the equation, say again by some PC-program, we find that the equation becomes

\[
(x^3 + x^2 - y^2)(x + y)^2(x - y)^2 = 0
\]

which certainly defines the same point set, but reveals that this time the two lines occurring should be counted with multiplicity 2.
Recall that an *irreducible polynomial* in $x$ and $y$ is a polynomial $p(x, y)$ which may not be factored as a product of two polynomials, both non-constants. Thus for instance $p(x, y) = x^3 + x^2 - y^2$ is irreducible, as is $r(x, y) = x + y$ and $s(x, y) = x - y$. A special case of an important theorem is the following:

**Theorem 2.6 (Unique Factorization of Polynomials)** Any polynomial in $x$ and $y$ with real (respectively complex) coefficients, may be factored as a product of powers of irreducible polynomials with real (respectively complex) coefficients. These irreducible polynomials are unique except for possibly being proportional by constant factors.

We make the following definition:

**Definition 2.5 (The factorization in irreducible polynomials)** The irreducible factorization of $f(x, y)$ is defined as an expression

$$f(x, y) = p_1(x, y)^{n_1} \cdots p_r(x, y)^{n_r}$$

where $n_i$ are positive integers and all $p_i(x, y)$ are irreducible and no two are proportional by a constant factor.

This factorization is unique up to constant factors, by the theorem.

**Corollary 2.7** Theorem 2.6 also holds for a polynomial in a number of variables, 1 up to any $N$. Definition 2.5 is also unchanged in the general case.

A polynomial may be irreducible as a polynomial with real coefficients, but *reducible* when considered as a polynomial with complex coefficients. This is the case for the polynomial

$$g(x, y) = x^2 + y^2,$$

which may not be factored as a polynomial with real coefficients, while

$$x^2 + y^2 = (x + iy)(x - iy).$$

The curve given by this polynomial has another interesting feature: As a curve in $\mathbb{A}_R^2$ it consists only of the origin, while it consists of two (complex) lines in $\mathbb{A}_C^2$, with equations $y = \pm ix$. They have only one real point on them, namely their point of intersection which is the origin. We would consider this as a degenerate case, say as a member of a family of circles, where the radius has shrunk to zero.

**Definition 2.6 (Real Affine Curve)** A real affine plane curve $C$ is the set of points $(a, b) \in \mathbb{A}_R^2$ which are zeroes of a polynomial $f(x, y)$ with real coefficients. The irreducible polynomials $p_i(x, y)$ occurring in the irreducible
factorization of \( f(x, y) \) referred to in Definition 2.5 define subsets \( C_i \) of \( C \) called the irreducible components of \( C \). The exponent \( n_i \) of \( p_i(x, y) \) in the factorization of \( f(x, y) \) is called the multiplicity of the irreducible component.

In other words, \( C_i \) occurs with multiplicity \( n_i \) in \( C \).

**Remark** This definition suffices as a first approximation, but it should not be concealed that it does represent a simplification. Indeed, according to the definition the “real affine curve” defined by \( x^2 + y^2 = 0 \) is the same as the one defined by \( x^2 + 2y^2 = 0 \). For a variety of reasons this is undesirable. One solution is to simply define a curve in \( \mathbb{A}^2_\mathbb{R} \) as being an equivalence class of polynomials, two polynomials being regarded as equivalent if one is a non-zero constant multiple of the other. This is mathematically sound, but only applies to a special geometric situation, where one geometric object, here the curve, is contained in another geometric object of one dimension higher, here the plane, and is defined by one “equation”. The final clarification of this concept will come when we explain the notion of a scheme, which was introduced by Alexander Grothendieck.

After a change of variables, which corresponds to a change of coordinate system,

\[
\begin{align*}
\bar{x} &= a + \alpha_{1,1}x + \alpha_{1,2}y \\
\bar{y} &= b + \alpha_{2,1}x + \alpha_{2,2}y
\end{align*}
\]

the curve given by \( f(x, y) = 0 \) is expressed by the equation \( \bar{f}(\bar{x}, \bar{y}) = 0 \), where \( \bar{f}(\bar{x}, \bar{y}) \) is obtained by substituting the expressions obtained by solving for \( x \) and \( y \),

\[
\begin{align*}
x &= \bar{a} + \beta_{1,1}\bar{x} + \beta_{1,2}\bar{y} \\
y &= \bar{b} + \beta_{2,1}\bar{x} + \beta_{2,2}\bar{y}
\end{align*}
\]

into \( f(x, y) \).

There are curves in the affine plane \( \mathbb{A}^2_\mathbb{R} \) which are not affine algebraic, but nevertheless form an important subject in geometry. The Archimedean spiral and the quadratrix of Hippias are such curves. They both were invented to solve some of the Classical Problems. They are not defined by a polynomial equation. Some other simple examples are the curves defined by \( y = \sin(x) \) or by \( y = e^x \). This class of curves is called the Transcendental Curves. In this book we will confine the general theory to treating the algebraic curves, that is to say the ones defined by a polynomial equation.

**2.8 Singularities and Multiplicities**

We now return to some general concepts introduced in Sect. 2.2, where we needed it to understand the degeneracy of conics. Consider an algebraic affine
curve $K$ with equation
\[ f(x, y) = 0. \]

Furthermore, let $(a, b)$ be a point on the curve, i.e., $f(a, b) = 0$. We note that the following definition relies heavily on the equation of the curve, not just the curve as a subset of $\mathbb{A}_R^2$:

**Definition 2.7** $(a, b)$ is said to be a smooth, or a non-singular, point on $K$ if
\[ \left( \frac{\partial f}{\partial x}(a, b), \frac{\partial f}{\partial y}(a, b) \right) \neq (0, 0). \]

Otherwise $(a, b)$ is called a singular point on $K$. A curve all of whose points are non-singular is referred to as a non-singular curve.

The vector $(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$ is referred to as the Jacobian vector (for short, the *Jacobian*) of the polynomial $f(x, y)$. Thus by definition a singular point is a point on the curve at which the Jacobian evaluates to the zero vector.

In Sect. 2.2 we saw that a non-degenerate conic curve is a non-singular curve. We look at the situation in more detail by the examples below.

It is time to turn to some examples.

(1) We first look at simple *conics*, and start out with a circle of radius $R > 0$, which has the equation
\[ x^2 + y^2 = R^2. \]

Here $f(x, y) = x^2 + y^2 - R^2$, and
\[ \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = (2x, 2y). \]

Evidently no point outside the origin can be a singular point of the circle, and as $R > 0$, every point on the circle is therefore smooth. We note that the same proof shows that an ellipse on standard form,
\[ \left( \frac{x}{a} \right)^2 + \left( \frac{y}{b} \right)^2 = 1 \]
is smooth everywhere as well.

A (non-degenerate) hyperbola on standard form, which is given as
\[ \left( \frac{x}{a} \right)^2 - \left( \frac{y}{b} \right)^2 = 1 \]
similarly has Jacobian
\[ \left( \frac{2}{a^2} x, -\frac{2}{b^2} y \right). \]
which also does not vanish outside the origin, showing that a hyperbola is smooth.

A degenerate hyperbola is one which has collapsed to the asymptotes, hence a curve with equation

\[
\left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 = 0.
\]

This curve has the same Jacobian as in the non-degenerate case, but now the origin actually lies on the curve, which therefore has the origin as its only singular point. Of course this degenerate hyperbola consists of two irreducible components which are lines intersecting at the origin, and that point is singular.

Our final conic curve is the parabola with equation

\[
ay - x^2 = 0
\]

where \(a \neq 0\). The Jacobian is \((-2x, a)\), so the only possibility of getting the zero vector at a point would be to have \(x = 0\) and \(a = 0\). For \(a \neq 0\) we therefore have no singular points. If \(a = 0\), then the equation yields the \(y\)-axis with multiplicity 2, and we see that then all points on the curve are singular.

(2) We next turn to the nodal cubic curve with equation

\[
y^2 - x^3 - x^2 = 0
\]

which is plotted in Fig. 2.6. The Jacobian is

\[
(-3x^2 - 2x, 2y)
\]

and thus \((x, y)\) is a singular point if and only if the two additional equations below are satisfied:

\[
-3x^2 - 2x = 0 \\
2y = 0.
\]

Thus \((x, y) = (0, 0)\) or \((x, y) = (-\frac{2}{3}, 0)\), and only the former lies on the curve, so the only singular point is \((0, 0)\).

(3) If \(f(x, y)\) is a polynomial, then all points on the curve given by \(f(x, y)^n = 0\) for \(n\) an integer greater than 1, will have all its points singular. This follows at once, since the Jacobian is

\[
\left(n f(x, y)^{n-1} \frac{\partial f}{\partial x}, n f(x, y)^{n-1} \frac{\partial f}{\partial y}\right).
\]

It is highly recommended that the reader examines the curves plotted in Sect. 2.4, and determines their singular points.
Let \((a, b)\) be a smooth point on the curve \(K\). Then we may find the equation for the tangent line at that point as follows. We first consider the parametric form for a line through \((a, b)\) with direction given by the vector \((u, v)\):

\[
L = \left\{(x, y) \mid x = a + ut, \quad y = b + vt \quad \text{where } t \in \mathbb{R}\right\}.
\]

This line will have the point \((a, b)\) in common with \(K\). We wish to determine other points of intersection. To do so we substitute the expressions for \(x\) and \(y\) in the parametric form for \(L\) into the equation for \(K\), and get

\[
f(a + ut, b + vt) = 0.
\]

Expanding the left hand side in a Taylor series we obtain

\[
f(a, b) + t \left( u \frac{\partial f}{\partial x}(a, b) + v \frac{\partial f}{\partial y}(a, b) \right) + t^2 \left( u^2 \frac{\partial^2 f}{\partial x^2}(a, b) + 2uv \frac{\partial^2 f}{\partial x \partial y}(a, b) + v^2 \frac{\partial^2 f}{\partial y^2}(a, b) \right) + \cdots = 0.
\]

which since \(f(a, b) = 0\) gives the following equation for \(t\):

\[
t \left( u \frac{\partial f}{\partial x}(a, b) + v \frac{\partial f}{\partial y}(a, b) \right) + t^2 \left( u^2 \frac{\partial^2 f}{\partial x^2}(a, b) + 2uv \frac{\partial^2 f}{\partial x \partial y}(a, b) + v^2 \frac{\partial^2 f}{\partial y^2}(a, b) \right) + \cdots = 0. \quad (2.1)
\]

The points of intersection between the curve and the line are found by solving this equation for \(t\). Of course we have \(t = 0\) as one solution, and we see that this solution will occur with multiplicity 1 if and only if

\[
u \frac{\partial f}{\partial x}(a, b) + v \frac{\partial f}{\partial y}(a, b) \neq 0.
\]

Such values of \(u, v\) exist if and only if \((a, b)\) is a smooth point on the curve. In that case there is exactly one line through \(P = (a, b)\) which does not intersect the curve with multiplicity 1, namely the line corresponding to \(u\) and \(v\) such that

\[
u \frac{\partial f}{\partial x}(a, b) + v \frac{\partial f}{\partial y}(a, b) = 0.
\]
By substituting
\[ ut = x - a \]
\[ vt = y - b \]
in this equation, we recover the equation for the tangent line to the curve at the point \((a, b)\)
\[ (x - a) \frac{\partial f}{\partial x}(a, b) + (y - b) \frac{\partial f}{\partial x}(a, b) = 0. \]

Earlier we used this equation to define the tangent line to a curve \(C\) at a point \(P = (a, b)\), but at that stage without a geometric justification. Now we see the geometric meaning of this definition.

**Definition 2.8** We denote the multiplicity of the solution \(t = 0\) of (2.1) by \(m_{K, P}(L)\). This number is referred to as the multiplicity with which the line \(L\) intersects the curve \(K\) in the point \(P\).

To sum up what we have so far, the point \(P\) is a smooth point on \(K\) provided there is exactly one line \(L\) intersecting \(K\) in \(P\) with multiplicity \(> 1\), and \(L\) is then called the tangent to \(K\) in \(P\). The normal situation is that the multiplicity is 2, if it is \(\geq 3\) then \(P\) is referred to as a flex (or an inflection point), if \(m_{K, P}(L) = 3\) the flex is said to be an ordinary flex. The term inflection point is also used for curves in \(A^2_R\) as a point where the sign of the curvature changes. A smooth point with this property is an inflection point in our sense, but not conversely. The tangent line at a flex is called an inflectional tangent.

We next turn to the question of what happens at a singular point. So let \(P = (a, b)\) be a singular point on the curve \(K\). Since the situation is more complicated than in the case when \(P\) is smooth, we introduce new variables by
\[ \bar{x} = x - a, \quad \bar{y} = y - b. \]
In other words, we shift the variables so that the new origin falls in \(P, P = (0, 0)\). We then find a new polynomial \(g\) such that
\[ f(x, y) = g(\bar{x}, \bar{y}) \]
by substituting \(x = \bar{x} + a\) and \(y = \bar{y} + b\) into \(f(x, y)\). The curve is also given by the equation
\[ g(\bar{x}, \bar{y}) = 0. \]

Since the origin is a point on the curve given by \(g(\bar{x}, \bar{y}) = 0\), it is clear that the polynomial \(g(\bar{x}, \bar{y})\) has no constant term. We now collect the terms of \(g(\bar{x}, \bar{y})\) which are of lowest total degree, and denote the sum of those terms by \(h(\bar{x}, \bar{y})\).
Thus for example, if
\[ g(x, y) = 2xy^2 - 5x^2y + 10x^9y^2 + 15x^{12}y, \]
then
\[ h(x, y) = 2xy^2 - 5x^2y. \]
The sum of all terms of lowest total degree of the polynomial \( g \) is called the initial part of the polynomial, and denoted by \( \text{in}(g) \). If the point \( P = (a, b) \) is smooth, then the Taylor expansion around the point \( (a, b) \) immediately shows that the polynomial \( h(x, y) \) is nothing but
\[ \frac{\partial g}{\partial x}(0, 0)x + \frac{\partial g}{\partial y}(0, 0)y = \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b). \]
Thus the concept introduced below generalizes the tangent at a smooth point, to a concept which applies to singular points as well.

With notations as above the polynomial \( h(x, y) \) defines a curve which is a finite union of lines through the point \((0, 0)\). In terms of \( x \) and \( y \), the equation
\[ h(x - a, y - b) = 0 \]
defines a finite union of lines through \( P = (a, b) \), some of them occurring with multiplicity \( >1 \). Indeed, we have
\[ h(x, y) = a_0x^m + a_1x^{m-1}y + \cdots + a_i x^{m-i}y^i + \cdots + a_m y^m \]
where not all \( a_i \) vanish. If \((\alpha_0, \beta_0)\) satisfies \( h(\alpha_0, \beta_0) = 0 \), then we also have \( h(s\alpha_0, s\beta_0) = 0 \) for all real numbers \( s \), as one immediately verifies since all the monomials of \( h \) are of the same total degree \( m \).

These lines are called the lines of tangency at the point \( P = (a, b) \). If \( P \) happens to be smooth, then there is only one line, occurring with multiplicity \( 1 \).

**Definition 2.9** The curve given by \( h(x - a, y - b) = 0 \) is referred to as the (affine) tangent cone of \( K \) at \( P \).

Any line through \( P = (a, b) \) may, as we have seen, be written on parametric form as
\[ x - a = ut, \quad y - b = vt \]
and its intersections with the curve is determined by the equation
\[ f(a + ut, b + vt) = g(ut, vt) = 0. \]
The multiplicity of the root \( t = 0 \) in this equation is referred to as the multiplicity of intersection between the curve and the line at the point \( P = (a, b) \).
All lines through \( P = (a, b) \) which do not coincide with one of the lines of tangency, intersect the curve with multiplicity equal to the number \( m \). This number \( m \) is of course only dependent upon the polynomial \( f(x, y) \) and the point \( P = (a, b) \).

In fact, we may assume that \( P = (0, 0) \). An arbitrary line through \( (0, 0) \) has the parametric form

\[
L = \left\{ (x, y) \mid x = ut, y = vt \text{ where } t \in \mathbb{R} \right\}.
\]

To find all points of intersection between this line and the curve \( K \), we substitute the expressions for \( x \) and \( y \) into \( f(x, y) \) and get

\[
f(a + ut, b + vt) = 0.
\]

This gives

\[
h(ut, vt) + R(ut, vt) = 0
\]

where \( R(x, y) \) denotes \( f(x, y) - h(x, y) \). Thus the points of intersection are given by the roots of the equation

\[
t^m(h(u, v) + t\varphi(t)) = 0.
\]

One of the roots is \( t = 0 \), and this solution will occur with multiplicity \( \geq m \), where equality holds if and only if

\[
h(u, v) \neq 0
\]

thus if and only if \( L \) is not one of the lines of tangency.

We conclude with the

**Definition 2.10** (Multiplicity of a point on a curve) The number \( m \) referred to above is called the multiplicity of the point \( P \) at \( K \).

We thus have the observation

**Proposition 2.8** A point on an affine algebraic curve is smooth if and only if it has multiplicity 1.
A Royal Road to Algebraic Geometry
Holme, A.
2012, XIV, 366 p., Hardcover
ISBN: 978-3-642-19224-1