APPENDIX A

Projective Geometry

In this appendix we have tried to summarize the basic properties of the projective plane and projective curves that are used elsewhere in this book. For further reading about projective algebraic geometry, the reader might profitably consult Brieskorn-Knörrer [1], Fulton [1], or Reid [1]. More high-powered accounts of modern algebraic geometry are given in Hartshorne [1] and Griffiths and Harris [1].

1. Homogeneous Coordinates and the Projective Plane

There are many ways to approach the construction of the projective plane. We will describe two constructions, one algebraic and one geometric, since each in its own way provides enlightenment.

We begin with a famous question from number theory, namely the solution of the equation

$$x^N + y^N = 1 \qquad \text{(Fermat Equation \#1)}$$

in rational numbers x and y. Suppose that we have found a solution, say $x = a/c$, $y = b/d$, where we write the fractions in lowest terms and with positive denominators. Substituting and clearing denominators gives the equation

$$a^N d^N + b^N c^N = c^N d^N.$$

It follows that $c^N | a^N d^N$. But $\gcd(a, c) = 1$ by assumption, so we conclude that $c^N | d^N$, and hence $c | d$. Similarly $d^N | b^N c^N$ and $\gcd(b, d) = 1$, which implies that $d | c$. Therefore $c = \pm d$, and since we've assumed that c and d are positive, we find that $d = c$. Thus any solution to the Fermat Equation #1 in rational numbers has the form $(a/c, b/c)$, and thus gives a solution in integers (a, b, c) to the homogeneous equation

$$X^N + Y^N = Z^N. \qquad \text{(Fermat Equation \#2)}$$

Conversely, any integer solution (a, b, c) to the second Fermat Equation with $c \neq 0$ will give a rational solution $(a/c, b/c)$ to the first. However,

different integer solutions (a, b, c) may lead to the same rational solution. For example, if (a, b, c) is an integer solution to Fermat's Equation #2, then for any integer t the triple (ta, tb, tc) will also be a solution, and clearly (a, b, c) and (ta, tb, tc) give the same rational solution to Fermat's Equation #1. The moral is that in solving Fermat's Equation #2, we should really treat triples (a, b, c) and (ta, tb, tc) as being the same solution, at least for non-zero t. This leads to the notion of *homogeneous coordinates* which we will describe in more detail later.

There is one more observation we wish to make before leaving this example, namely the "problem" that Fermat's Equation #2 may have some integer solutions that do not correspond to rational solutions of Fermat's Equation #1. First, the point $(0, 0, 0)$ is always a solution of the second equation, but this solution is so trivial that we will just discard it. Second, and potentially more serious, is the fact that if N is odd, then Fermat's Equation #2 has the solutions $(1, -1, 0)$ and $(-1, 1, 0)$ which do not give solutions to the first Fermat Equation. To see what is happening, suppose that we take a sequence of solutions (a_i, b_i, c_i), $i = 1, 2, 3, \ldots$, such that $(a_i, b_i, c_i) \to (1, -1, 0)$ as $i \to \infty$. Of course, we cannot do this with integer solutions, so now we'll let the a_i, b_i, c_i's be real numbers. The corresponding solutions to the first Fermat Equation are $(a_i/c_i, b_i/c_i)$, and we see that these solutions approach (∞, ∞) as $(a_i, b_i, c_i) \to (1, -1, 0)$. In other words, the extra solutions $(1, -1, 0)$ and $(-1, 1, 0)$ to Fermat's Equation #2 somehow correspond to solutions of the first Fermat Equation which lie "at infinity." As we will see, the theory of solutions of polynomial equations becomes neater and clearer if we treat these extra points "at infinity" just like we treat all of the other points.

We are now ready for our first definition of the projective plane, which is essentially an algebraic definition. We define the *projective plane* to be the set of triples $[a, b, c]$, with a, b, c not all zero, such that two triples $[a, b, c]$ and $[a', b', c']$ are considered to be the same point if there is a non-zero t such that $a = ta'$, $b = tb'$, $c = tc'$. The numbers a, b, c are called *homogeneous coordinates* for the point $[a, b, c]$. We will denote the projective plane by $\mathbb{P}^2$. In other words, we define an equivalence relation $\sim$ on the set of triples $[a, b, c]$ by the rule

$$[a, b, c] \sim [a', b', c'] \quad \text{if there is a non-zero } t \text{ so that } a = ta',\ b = tb',\ c = tc'.$$

Then $\mathbb{P}^2$ consists of the set of equivalence classes of triples $[a, b, c]$ except that we exclude the triple $[0, 0, 0]$:

$$\mathbb{P}^2 = \frac{\{[a, b, c] \,:\, a, b, c \text{ are not all zero}\}}{\sim}.$$

More generally, for any integer $n \geq 1$ we define *projective n-space* to be the set of equivalence classes of homogeneous $n + 1$-tuples,

$$\mathbb{P}^n = \frac{\{[a_0, a_1, \ldots, a_n] \,:\, a_0, a_1, \ldots, a_n \text{ not all zero}\}}{\sim},$$

where $[a_0, a_1, \ldots, a_n] \sim [a_0', a_1', \ldots, a_n']$ if there is a non-zero t so that $a_0 = ta_0', \ldots, a_n = ta_n'$.

We eventually want to do geometry in the projective plane, so we need to define some geometric objects. In the next section we will study quite general curves, but for the moment we will be content to describe the lines in $\mathbb{P}^2$. We define a *line in* $\mathbb{P}^2$ to be the set of points $[a, b, c] \in \mathbb{P}^2$ whose coordinates satisfy an equation of the form

$$\alpha X + \beta Y + \gamma Z = 0$$

for some constants α, β, γ not all zero. Note that if $[a, b, c]$ satisfies such an equation, then so does $[ta, tb, tc]$ for any t, so to check if a point of $\mathbb{P}^2$ is on a given line, one can use any homogeneous coordinates for the point.

In order to motivate our second description of the projective plane, we consider a geometric question. It is well-known that two points in the usual (x, y) plane determine a unique line, namely the line which goes through them. Similarly, two lines in the plane determine a unique point, namely the point where they intersect, unless the two lines happen to be parallel. From both an aesthetic and a practical viewpoint, it would be nice to provide these poor parallel lines with an intersection point of their own. Since the plane itself doesn't contain the requisite points, we will add on the extra points by fiat. How many extra points do we need? For example, would it suffice to use one extra point P and decree that any two parallel lines intersect at the point P? The answer is no, and here's why.

Let L_1 and L_2 be parallel lines, and let P be the extra point where they are to intersect. Similarly, let L_1' and L_2' be parallel lines which intersect at the extra point P'. (See Figure A.1.) Suppose that L_1 and L_1' are not parallel. Then L_1 and L_1' intersect at some ordinary point, say $L_1 \cap L_1' = \{Q\}$. But two lines are allowed to have only one point in common, so it follows that the points $P \in L_1$ and $P' \in L_1'$ must be distinct. So we really need to add an extra point for each distinct direction in the ordinary plane, and then decree that a line L consists of its usual points together with the extra point determined by its direction.

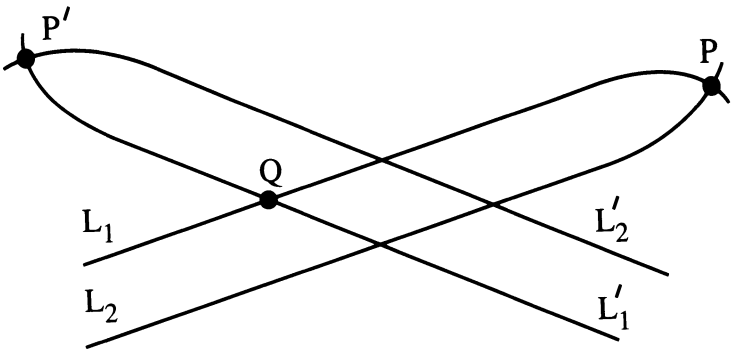
This leads to our second definition of the projective plane, this time in purely geometric terms. For simplicity, we will denote the usual *Euclidean* (also called *affine*) *plane* by

$$\mathbb{A}^2 = \{(x, y) : x \text{ and } y \text{ any numbers}\}.$$

Then we define the projective plane to be

$$\mathbb{P}^2 = \mathbb{A}^2 \cup \{\text{the set of directions in } \mathbb{A}^2\},$$

where *direction* is a non-oriented notion. Two lines have the same direction if and only if they are parallel. Logically we could define a direction in this sense to be an equivalence class of parallel lines, that is, a direction is a collection of all lines parallel to a given line. The extra points in $\mathbb{P}^2$



Parallel Lines With Intersection Points "At Infinity"

Figure A.1

associated to directions, that is the points in $\mathbb{P}^2$ that are not in $\mathbb{A}^2$, are often called *points at infinity.*

As indicated above, a line in $\mathbb{P}^2$ then consists of a line in $\mathbb{A}^2$ together with the point at infinity specified by its direction. The intersection of two parallel lines is the point at infinity corresponding to their common direction. Finally, the set of all points at infinity is itself considered to be a line L_∞, and the intersection of any other line L with L_∞ is the point at infinity corresponding to the direction of L. With these conventions, it is easy to see that there is a unique line going through any two distinct points of $\mathbb{P}^2$, and further any two distinct lines in $\mathbb{P}^2$ intersect in exactly one point. So the projective plane in this geometric incarnation eliminates the need to make a distinction between parallel and non-parallel lines. In fact, $\mathbb{P}^2$ has no parallel lines at all!

We now have two definitions of the projective plane, so it behooves us to show that they are equivalent. First we need a more analytic description of the set of directions in $\mathbb{A}^2$. One way to describe these directions is by the set of lines in $\mathbb{A}^2$ going through the origin, since every line in $\mathbb{A}^2$ is parallel to a unique line through the origin. Now the lines through the origin are given by equations

$$Ay = Bx$$

with A and B not both zero. However, it is possible for two pairs to give the same line. More precisely, the pairs (A, B) and (A', B') will give the same line if and only if there is a non-zero t such that $A = tA'$ and $B = tB'$. Thus the set of directions in $\mathbb{A}^2$ is naturally described by the points $[A, B]$ of the projective line $\mathbb{P}^1$. This allows us to write our second description of $\mathbb{P}^2$ in the form

$$\mathbb{P}^2 = \mathbb{A}^2 \cup \mathbb{P}^1;$$

a point $[A, B] \in \mathbb{P}^1 \subset \mathbb{P}^2$ corresponds to the direction of the line $Ay = Bx$.

How is this related to the definition of $\mathbb{P}^2$ in terms of homogeneous coordinates? Recall in our original example we associated a point $(x, y) \in \mathbb{A}^2$ with the point $[x, y, 1] \in \mathbb{P}^2$, and similarly a point $[a, b, c] \in \mathbb{P}^2$ with $c \neq 0$ was associated to the point $(a/c, b/c) \in \mathbb{A}^2$. But the remaining points in $\mathbb{P}^2$, namely those with $c = 0$, just give a copy of $\mathbb{P}^1$. In other words, the following maps show how to identify our two definitions of the projective plane:

$$\begin{array}{ccc} \dfrac{\{[a,b,c] \,:\, a,b,c \text{ not all zero}\}}{\sim} & \longleftrightarrow & \mathbb{A}^2 \cup \mathbb{P}^1 \\ \hline [a,b,c] & \longrightarrow & \begin{cases} \left(\dfrac{a}{c}, \dfrac{b}{c}\right) \in \mathbb{A}^2 & \text{if } c \neq 0 \\ [a,b] \in \mathbb{P}^1 & \text{if } c = 0 \end{cases} \\ \hline [x,y,1] & \longleftarrow & (x,y) \in \mathbb{A}^2 \\ [A,B,0] & \longleftarrow & [A,B] \in \mathbb{P}^1 \end{array}$$

It's easy to check that these two maps are inverses. For example, if $c \neq 0$ then

$$[a,b,c] \longmapsto \left(\frac{a}{c}, \frac{b}{c}\right) \longmapsto \left[\frac{a}{c}, \frac{b}{c}, 1\right] = [a,b,c].$$

We'll leave the remaining verifications to you.

Each of our definitions of the projective plane came with a description of what constitutes a line, so we should also check that the lines match up properly. For example, a line L in $\mathbb{P}^2$ using homogeneous coordinates is the set of solutions $[a, b, c]$ to an equation

$$\alpha X + \beta Y + \gamma Z = 0.$$

Suppose first that α and β are not both zero. Then any point $[a, b, c] \in L$ with $c \neq 0$ is sent to the point

$$\left(\frac{a}{c}, \frac{b}{c}\right) \quad \text{on the line} \quad \alpha x + \beta y + \gamma = 0 \quad \text{in } \mathbb{A}^2.$$

And the point $[-\beta, \alpha, 0] \in L$ is sent to the point $[-\beta, \alpha] \in \mathbb{P}^1$, which corresponds to the direction of the line $-\beta y = \alpha x$. This is exactly right, since the line $-\beta y = \alpha x$ is precisely the line going through the origin that is parallel to the line $\alpha x + \beta y + \gamma = 0$. This takes care of all of the lines except for the line $Z = 0$ in $\mathbb{P}^2$. But the line $Z = 0$ is sent to the line in $\mathbb{A}^2 \cup \mathbb{P}^1$ consisting of all of the points at infinity. So the lines in our two descriptions of $\mathbb{P}^2$ are consistent.

2. Curves in the Projective Plane

An *algebraic curve* in the affine plane $\mathbb{A}^2$ is defined to be the set of solutions to a polynomial equation in two variables,

$$f(x, y) = 0.$$

For example, the equation $x^2+y^2-1 = 0$ is a circle in $\mathbb{A}^2$, and $2x-3y^2+1 = 0$ is a parabola.

In order to define curves in the projective plane $\mathbb{P}^2$, we will need to use polynomials in three variables, since points in $\mathbb{P}^2$ are represented by homogeneous triples. But there is the further difficulty that each point in $\mathbb{P}^2$ can be represented by many different homogeneous triples. It thus makes sense to look only at polynomials $F(X, Y, Z)$ with the property that $F(a, b, c) = 0$ implies that $F(ta, tb, tc) = 0$ for all t. These turn out to be the homogeneous polynomials, and we use them to define curves in $\mathbb{P}^2$.

More formally, a polynomial $F(X, Y, Z)$ is called a *homogeneous polynomial of degree d* if it satisfies the identity

$$F(tX, tY, tZ) = t^d F(X, Y, Z).$$

This identity is equivalent to the statement that F is a linear combination of monomials $X^i Y^j Z^k$ with $i+j+k = d$. We define a *projective curve C* in the projective plane $\mathbb{P}^2$ to be the set of solutions to a polynomial equation

$$C : F(X, Y, Z) = 0,$$

where F is a non-constant homogeneous polynomial. We will also call C an *algebraic curve* or sometimes just a *curve* if it is clear we are working in $\mathbb{P}^2$. The *degree of the curve C* is the degree of the polynomial F. For example,

$$C_1 : X^2 + Y^2 - Z^2 = 0 \qquad \text{and} \qquad C_2 : Y^2Z - X^3 - XZ^2 = 0$$

are projective curves, where C_1 has degree 2 and C_2 has degree 3.

In order to check if a point $P \in \mathbb{P}^2$ is on the curve C, we can take any homogeneous coordinates $[a, b, c]$ for P and check if $F(a, b, c)$ is zero. This is true because any other homogeneous coordinates for P look like $[ta, tb, tc]$ for some non-zero t. Thus, $F(a, b, c)$ and $F(ta, tb, tc) = t^d F(a, b, c)$ are either both zero or both non-zero.

This tells us what a projective curve is when we use the definition of $\mathbb{P}^2$ by homogeneous coordinates. It will be very illuminating to relate this to the description of $\mathbb{P}^2$ as $\mathbb{A}^2 \cup \mathbb{P}^1$, where $\mathbb{A}^2$ is the usual affine plane and the points at infinity (i.e. the points in $\mathbb{P}^1$) correspond to the directions in $\mathbb{A}^2$. Let $C \subset \mathbb{P}^2$ be a curve given by a homogeneous polynomial of degree d,

$$C : F(X, Y, Z) = 0.$$

If $P = [a,b,c] \in C$ is a point of C with $c \neq 0$, then according to the identification $\mathbb{P}^2 \leftrightarrow \mathbb{A}^2 \cup \mathbb{P}^1$ described in Section 1, the point $P \in C \subset \mathbb{P}^2$ corresponds to the point

$$\left(\frac{a}{c}, \frac{b}{c}\right) \in \mathbb{A}^2 \subset \mathbb{A}^2 \cup \mathbb{P}^1.$$

On the other hand, combining $F(a,b,c) = 0$ with the fact that F is homogeneous of degree d shows that

$$0 = \frac{1}{c^d} F(a,b,c) = F\left(\frac{a}{c}, \frac{b}{c}, 1\right).$$

In other words, if we define a new (non-homogeneous) polynomial $f(x,y)$ by

$$f(x,y) = F(x,y,1),$$

then we get a map

$$\begin{aligned} \{[a,b,c] \in C \,:\, c \neq 0\} &\longrightarrow \{(x,y) \in \mathbb{A}^2 \,:\, f(x,y) = 0\}. \\ [a,b,c] &\longmapsto \left(\frac{a}{c}, \frac{b}{c}\right) \end{aligned}$$

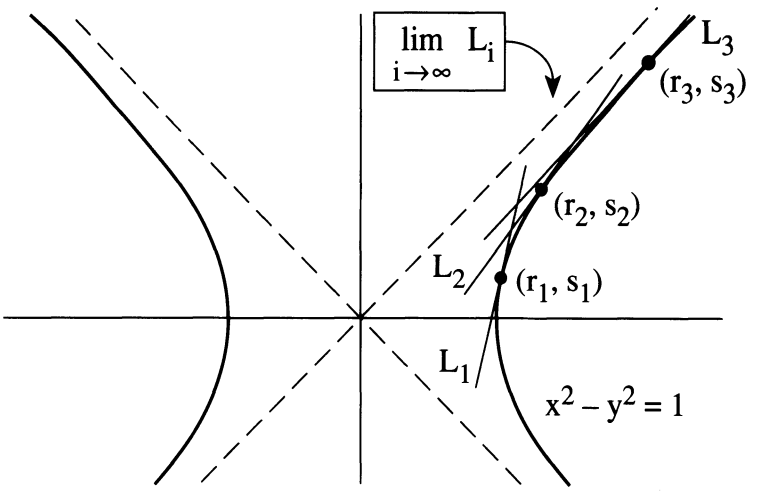
And it is easy to see that this map is one-to-one and onto, since if $(r,s) \in \mathbb{A}^2$ satisfies the equation $f(x,y) = 0$, then clearly $[r,s,1] \in C$. We call the curve $f(x,y) = 0$ in $\mathbb{A}^2$ the *affine part* of the projective curve C.

It remains to look at the points $[a,b,c]$ on C with $c = 0$ and describe them geometrically in terms of the affine part of C. The points $[a,b,0]$ on C satisfy the equation $F(X,Y,0) = 0$, and they are sent to the points at infinity $[a,b] \in \mathbb{P}^1$ in $\mathbb{A}^2 \cup \mathbb{P}^1$. We claim that these points, which recall are really directions in $\mathbb{A}^2$, correspond to the limiting tangent directions of the affine curve $f(x,y) = 0$ as we move along the affine curve out to infinity. In other words, and this is really the intuition to keep in mind, an affine curve $f(x,y) = 0$ is somehow "missing" some points which lie out at infinity, and the points that are missing are the (limiting) directions as one moves along the curve out towards infinity.

Rather than giving a general proof, we will illustrate with two examples. First, consider the line

$$L : \alpha X + \beta Y + \gamma Z = 0,$$

say with $\alpha \neq 0$. The affine part of L is the line $L_0 : \alpha x + \beta y + 1 = 0$ in $\mathbb{A}^2$. The points at infinity on L correspond to the points with $Z = 0$. There is only one such point, namely $[-\beta, \alpha, 0] \in L$, which corresponds to the point at infinity $[-\beta, \alpha] \in \mathbb{P}^1$, which in turn corresponds to the direction $-\beta y = \alpha x$ in $\mathbb{A}^2$. This direction is exactly the direction of the line L_0. Thus L consists of the affine line L_0 together with the single point at infinity corresponding to the direction of L_0.



Points At Infinity Are Limits Of Tangent Directions

Figure A.2

Next we look at the projective curve

$$C : X^2 - Y^2 - Z^2 = 0.$$

There are two points on C with $Z = 0$, namely $[1, 1, 0]$ and $[1, -1, 0]$. These two points correspond respectively to the points at infinity $[1, 1], [1, -1] \in \mathbb{P}^1$, or equivalently to the directions $y = x$ and $y = -x$ in $\mathbb{A}^2$. The affine part of C is the hyperbola

$$C_0 : x^2 - y^2 - 1 = 0.$$

Suppose we take a sequence of points $(r_1, s_1), (r_2, s_2), \ldots$ on C_0 such that these points tend towards infinity, say $|s_2| \to \infty$. If we rewrite $r_i^2 - s_i^2 - 1 = 0$ as

$$\left(\frac{r_i}{s_i} - 1\right)\left(\frac{r_i}{s_i} + 1\right) = \frac{1}{s_i^2},$$

then the right-hand side goes to 0 as $i \to \infty$, so we see that either

$$\lim_{i\to\infty} \frac{r_i}{s_i} = 1 \qquad \text{or} \qquad \lim_{i\to\infty} \frac{r_i}{s_i} = -1,$$

depending on which branch of the hyperbola we travel on. (See Figure A.2.)

Let L_i be the tangent line to C_0 at the point (r_i, s_i). We claim that as $i \to \infty$, the direction of the tangent line L_i approaches the direction of one of the lines $y = \pm x$. This is nothing more than the assertion that

the lines $y = \pm x$ are asymptotes for the curve C_0. To check this assertion analytically we implicitly differentiate the equation $x^2 - y^2 - 1 = 0$ to get

$$\frac{dy}{dx} = \frac{x}{y}, \quad \text{and so}$$
$$(\text{slope of } L_i) = \big(\text{slope of } C_0 \text{ at } (r_i, s_i)\big) = \frac{r_i}{s_i} \underset{i\to\infty}{\longrightarrow} \pm 1.$$

The preceding discussion shows that if we start with a projective curve $C : F(X, Y, Z) = 0$, then we can write C as the union of its affine part C_0 and its points at infinity. Here C_0 is the affine curve given by the equation

$$C_0 : f(x, y) = F(x, y, 1) = 0;$$

and the points at infinity on C are the points with $Z = 0$, which correspond to the limiting directions of the tangent lines to C_0. The process of replacing the homogeneous polynomial $F(X, Y, Z)$ by the inhomogeneous polynomial $f(x, y) = F(x, y, 1)$ is called *dehomogenization (with respect to the variable Z)*. We would now like to reverse this process.

Thus suppose we begin with an affine curve C_0 given by an equation $f(x, y) = 0$. We want to find a projective curve C whose affine part is C_0, or equivalently we want to find a homogeneous polynomial $F(X, Y, Z)$ so that $F(x, y, 1) = f(x, y)$. This is easy to do, although we want to be careful not to also include the line at infinity in our curve. If we write the polynomial $f(x, y)$ as $\sum a_{ij} x^i y^j$, then the *degree of f* is defined to be the largest value of $i + j$ for which the coefficient a_{ij} is not zero. For example,

$$\deg(x^2 + xy + x^2y^2 + y^3) = 4 \quad \text{and} \quad \deg(y^2 - x^3 - ax^2 - bx - c) = 3.$$

Then we define the *homogenization* of the polynomial $f(x, y) = \sum a_{ij} x^i y^j$ to be

$$F(X, Y, Z) = \sum_{i,j} a_{ij} X^i Y^j Z^{d-i-j}, \qquad \text{where } d = \deg(f).$$

It is clear from this definition that F is homogeneous of degree d and that $F(x, y, 1) = f(x, y)$. Further, our choice of d ensures that $F(X, Y, 0)$ is not identically zero, so the curve defined by $F(X, Y, Z) = 0$ does not contain the entire line at infinity. Thus by using homogenization and dehomogenization we obtain a one-to-one correspondence between affine curves and projective curves that do not contain the line at infinity.

We should also mention that there is nothing sacred about the variable Z. We could just as well dehomogenize a curve $F(X, Y, Z)$ with respect to one of the other variables, say Y, to get an affine curve $F(x, 1, z) = 0$ in the affine xz plane. It is sometimes convenient to do this if we are especially interested in one of the points at infinity on the projective curve C. In essence what we are doing is taking a different line, in this case the

line $Y = 0$, and making it into the "line at infinity." An example should make this clearer. Suppose we want to study the curve

$$C : Y^2Z - X^3 - Z^3 = 0 \qquad \text{and the point} \quad P = [0,1,0] \in C.$$

If we dehomogenize with respect to Z, then the point P becomes a point at infinity on the affine curve $y^2 - x^3 - 1 = 0$. So instead we dehomogenize with respect to Y, which means setting $Y = 1$. We then get the affine curve

$$z - x^3 - z^3 = 0, \qquad \text{and the point } P \text{ becomes the point } (x,z) = (0,0).$$

In general, by taking different lines to be the line at infinity, we can break a projective curve C up into a lot of overlapping affine parts, and then these affine parts can be "glued" together to form the entire projective curve.

Up to now we have been working with polynomials without worrying overmuch about what the coefficients of our polynomials look like, and similarly we've talked about solutions of polynomial equations without specifying what sorts of solutions we mean. Classical algebraic geometry is concerned with describing the complex solutions to systems of polynomial equations, but in order to study number theory we will be more interested in finding solutions whose coordinates are in non-algebraically closed fields like $\mathbb{Q}$, or even in rings like $\mathbb{Z}$. That being the case, it makes sense to look at curves given by polynomial equations with rational or integer coefficients.

We call a curve C *rational* if it is the set of zeros of a polynomial having rational coefficients.† Note that the solutions of the equation $F(X,Y,Z) = 0$ and the equation $cF(X,Y,Z) = 0$ are the same for any non-zero c. This allows us to clear the denominators of the coefficients, so a rational curve is in fact the set of zeros of a polynomial with integer coefficients. All of the examples given above are rational curves, since their equations have integer coefficients.

Let C be a projective curve that is rational, say C is given by the equation $F(X,Y,Z) = 0$ for a homogeneous polynomial F having rational coefficients. The *set of rational points on* C, which we denote by $C(\mathbb{Q})$, is the set of points of C having rational coordinates:

$$C(\mathbb{Q}) = \big\{[a,b,c] \in \mathbb{P}^2 \,:\, F(a,b,c) = 0 \quad \text{and} \quad a,b,c \in \mathbb{Q}\big\}.$$

Note that if $P = [a,b,c]$ is in $C(\mathbb{Q})$, it is not necessary that a,b,c themselves be rational, since a point P has many different homogeneous coordinates. All one can say is that $[a,b,c] \in C$ is a rational point of C (i.e. is in $C(\mathbb{Q})$) if and only if there is a non-zero number t so that $ta, tb, tc \in \mathbb{Q}$.

Similarly, if C_0 is an affine curve that is rational, say $C : f(x,y) = 0$, then the set of rational points on C_0 is denoted $C_0(\mathbb{Q})$ and consists of

† We should warn the reader that this terminology is non-standard. In the usual terminology of algebraic geometry, a curve is called rational if it is isomorphic to the projective line $\mathbb{P}^1$, and a curve given by polynomials with rational coefficients is said to be defined over $\mathbb{Q}$.

all $(r,s) \in C$ with $r,s \in \mathbb{Q}$. It is easy to see that if C_0 is the affine piece of a projective curve C, then $C(\mathbb{Q})$ consists of $C_0(\mathbb{Q})$ together with those points at infinity which happen to be rational. Some of the most famous questions in number theory involve the set of rational points $C(\mathbb{Q})$ on certain curves C. For example, the N^{th} Fermat curve C_N is the projective curve

$$C_N : X^N + Y^N = Z^N,$$

and Fermat's last theorem asserts that $C_N(\mathbb{Q})$ consists of only those points with one of X, Y, or Z equal to zero.

The theory of Diophantine equations also deals with integer solutions of polynomial equations. Let C_0 be an affine curve that is rational, say given by an equation $f(x,y) = 0$. We define the *set of integer points of* C_0, which we denote $C_0(\mathbb{Z})$, to be the set of points of C_0 having integer coordinates:

$$C_0(\mathbb{Z}) = \big\{(r,s) \in \mathbb{A}^2 \,:\, f(r,s) = 0 \quad \text{and} \quad r,s \in \mathbb{Z}\big\}.$$

Why do we only talk about integer points on affine curves and not on projective curves? The answer is that for a projective curve, the notion of integer point and rational point coincide. Here we might say that a point $[a,b,c] \in \mathbb{P}^2$ is an integer point if its coordinates are integers. But if $P \in \mathbb{P}^2$ is any point which is given by homogeneous coordinates $P = [a,b,c]$ that are rational, then we can find an integer t to clear the denominators of a,b,c, and so $P = [ta,tb,tc]$ also has homogenous coordinates which are integers. So for a projective curve C we would have $C(\mathbb{Q}) = C(\mathbb{Z})$.

It is also possible to look at polynomial equations and their solutions in rings and fields other than $\mathbb{Z}$ or $\mathbb{Q}$ or $\mathbb{R}$ or $\mathbb{C}$. For example, one might look at polynomials with coefficients in the finite field $\mathbb{F}_p$ with p elements and ask for solutions whose coordinates are also in the field $\mathbb{F}_p$. You may worry about your geometric intutions in situations like this. How can one visualize points and curves and directions in $\mathbb{A}^2$ when the points of $\mathbb{A}^2$ are pairs (x,y) with $x,y \in \mathbb{F}_p$? There are two answers to this question. The first and most reassuring is that you can continue to think of the usual Euclidean plane (i.e., $\mathbb{R}^2$) and most of your geometric intuitions concerning points and curves will still be true when you switch to coordinates in $\mathbb{F}_p$. The second and more practical answer is that the affine and projective planes and affine and projective curves are defined algebraically in terms of ordered pairs (r,s) or homogenous triples $[a,b,c]$ without any reference to geometry. So in proving things one can work algebraically using coordinates, without worrying at all about geometric intuitions. We might summarize this general philosophy as

Think Geometrically, Prove Algebraically.

One of the fundamental questions answered by the differential calculus is that of finding the tangent line to a curve. If $C : f(x,y) = 0$ is an affine curve, then implicit differentiation gives the relation $\dfrac{\partial f}{\partial x} + \dfrac{\partial f}{\partial y}\dfrac{dy}{dx} = 0$. So

if $P = (r, s)$ is a point on C, the tangent line to C at P is given by the equation

$$\frac{\partial f}{\partial x}(r,s)(x-r) + \frac{\partial f}{\partial y}(r,s)(y-s) = 0.$$

This is the answer as provided by elementary calculus. But we clearly have a problem if both of the partial derivatives are 0. For example, this happens for each of the curves

$$C_1 : y^2 = x^3 + x^2 \qquad \text{and} \qquad C_2 : y^2 = x^3$$

at the point $P = (0, 0)$. If we sketch these curves, we see that they look a bit strange at P. (See Figures 1.13 and 1.14 near the end of Section 3 of Chapter I.) The curve C_1 crosses over itself at P, so it has two distinct tangent directions there. The curve C_2, on the other hand, has a cusp at P, which means it comes to a sharp point at P. We will say that P is a *singular point* of the curve $C : f(x, y) = 0$ if

$$\frac{\partial f}{\partial x}(P) = \frac{\partial f}{\partial y}(P) = 0.$$

We call P a *non-singular point* if it is not singular, and we say that C is a *non-singular curve* (or *smooth curve*) if every point of C is non-singular. If P is a non-singular point of C, then we define the *tangent line to C at P* to be the line described above.

For a projective curve $C : F(X, Y, Z) = 0$ described by a homogeneous polynomial we make similar definitions. More precisely, if $P = [a, b, c]$ is a point on C with $c \neq 0$, then we go to the affine part of C and check whether or not the point

$$P_0 = \left(\frac{a}{c}, \frac{b}{c}\right) \quad \text{is singular on the affine curve} \quad C_0 : F(x, y, 1) = 0.$$

And if $c = 0$, then we can dehomogenize in some other way. For example, if $a \neq 0$, we check whether or not the point

$$P_0 = \left(\frac{b}{a}, \frac{c}{a}\right) \quad \text{is singular on the affine curve} \quad C_0 : F(1, y, z) = 0.$$

We say that C is non-singular or smooth if all of its points, including the points at infinity, are non-singular. If P is a non-singular point of C, we define the tangent line to C at P by dehomogenizing, finding the tangent line to the affine part of C at P, and then homogenizing the equation of that tangent line to get a line in $\mathbb{P}^2$. (An alternative method to check for singularity and find tangent lines on projective curves is described in the exercises.)

When one is faced with a complicated equation, it is natural to try to make a change of variables in order to simplify it. Probably the first

significant example of this that you have seen is the process of completing the square to solve a quadratic equation. Thus to solve $Ax^2 + Bx + C = 0$ we multiply through by $4A$ and rewrite the equation as

$$(2Ax + B)^2 + 4AC - B^2 = 0.$$

This suggests the substitution $x' = 2Ax + B$, and then we can solve $x'^2 + 4AC - B^2 = 0$ for $x' = \pm\sqrt{B^2 - 4AC}$. The crucial final step uses the fact that our substitution is invertible, so we can solve for x in terms of x' to obtain the usual quadratic formula

$$x = \frac{-B + x'}{2A} = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}.$$

More generally, suppose we are given a projective curve C of degree d, say defined by an equation $C : F(X, Y, Z) = 0$. In order to change coordinates on $\mathbb{P}^2$ we make a substitution

$$\begin{aligned} X &= m_{11}X' + m_{12}Y' + m_{13}Z', \\ Y &= m_{21}X' + m_{22}Y' + m_{23}Z', \\ Z &= m_{31}X' + m_{32}Y' + m_{33}Z'. \end{aligned} \qquad (*)$$

Then we get a new curve C' given by the equation $F'(X', Y', Z') = 0$, where F' is the polynomial

$$\begin{aligned} F'(X', Y', Z') = F(&m_{11}X' + m_{12}Y' + m_{13}Z', \\ &m_{21}X' + m_{22}Y' + m_{23}Z', m_{31}X' + m_{32}Y' + m_{33}Z'). \end{aligned}$$

The change of coordinates $(*)$ gives a map from C' to C; that is, given a point $[a', b', c'] \in C'$, we substitute $X' = a'$, $Y' = b'$, and $Z' = c'$ into $(*)$ to get a point $[a, b, c] \in C$. Further, this map $C' \to C$ will have an inverse provided that the matrix $M = (m_{ij})_{1 \le i,j \le 3}$ is invertible. More precisely, if $M^{-1} = N = (n_{ij})$, then the change of coordinates

$$\begin{aligned} X' &= n_{11}X + n_{12}Y + n_{13}Z, \\ Y' &= n_{21}X + n_{22}Y + n_{23}Z, \\ Z' &= n_{31}X + n_{32}Y + n_{33}Z, \end{aligned}$$

will map C to C'. We call a change of coordinates on $\mathbb{P}^2$ given by an invertible 3×3 matrix a *projective transformation.* Note that if the matrix has rational coefficients, then the corresponding projective transformation gives a one-to-one correspondence between $C(\mathbb{Q})$ and $C'(\mathbb{Q})$. So the number theoretic problem of finding the rational points on the curve C is equivalent to the problem of finding the rational points on C'.

3. Intersections of Projective Curves

Recall that our geometric construction of the projective plane was based on the desire that every pair of distinct lines should intersect in exactly one point. In this section we are going to discuss the intersection of curves of higher degree.

How many intersection points should two curves have? Let's begin with a thought experiment, and then we'll consider some examples and see to what extent our intuition is correct. Let C_1 be an affine curve of degree d_1 and let C_2 be an affine curve of degree d_2. Thus, C_1 and C_2 are given by polynomials

$$\begin{aligned} C_1 &: f_1(x,y) = 0 \qquad \text{with } \deg(f_1) = d_1, \text{ and} \\ C_2 &: f_2(x,y) = 0 \qquad \text{with } \deg(f_2) = d_2. \end{aligned}$$

The points in the intersection $C_1 \cap C_2$ are the solutions to the simultaneous equations $f_1(x,y) = f_2(x,y) = 0$. Suppose now that we consider f_1 as a polynomial in the variable y whose coefficients are polynomials in x. Then $f_1(x,y) = 0$, being a polynomial equation of degree d_1 in y, should in principle have d_1 roots $y_1, y_2, \ldots, y_{d_1}$. Now we substitute each of these roots into the second equation $f_2(x,y) = 0$ to find d_1 equations for x, namely

$$f_2(x,y_1) = 0, \qquad f_2(x,y_2) = 0, \qquad \ldots \qquad f_2(x,y_{d_1}) = 0.$$

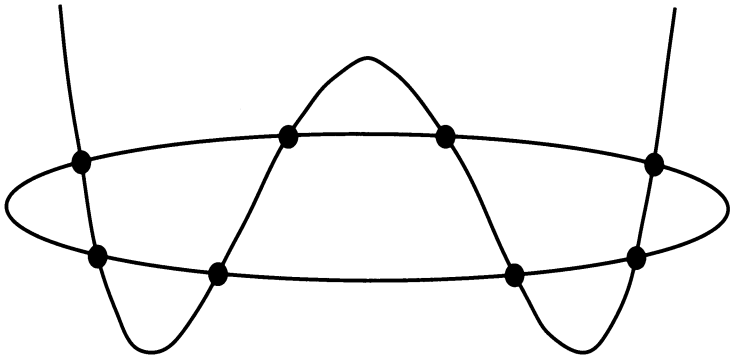
Each of these equations is a polynomial in x of degree d_2, so in principal each equation should yield d_2 values for x. Altogether we appear to get $d_1 d_2$ pairs (x,y) which satisfy $f_1(x,y) = f_2(x,y) = 0$, which seems to indicate we should have $\#(C_1 \cap C_2) = d_1 d_2$. For example, a curve of degree 2 and a curve of degree 4 should intersect in 8 points, as is illustrated in Figure A.3. This assertion, that curves of degree d_1 and d_2 intersect in $d_1 d_2$ point, is indeed true provided that it is interpreted properly. However, matters are considerably more complicated than they appear at first glance, as will be clear from the following examples. [Can you find all of the ways in which our plausibility argument fails to be a valid proof? For example, the "roots" $y_1, \ldots, y_{d_1}$ really depend on x, so we should write $f_2\big(x, y_i(x)\big) = 0$, and then it is not at all clear how many roots we should expect.]

Curves of degree one are lines, and curves of degree two are called *conics*. We already know that two lines in $\mathbb{P}^2$ intersect in a unique point, so the next simplest case is the intersection of a line and a conic. Our discussion above leads us to expect two intersection points, so we look at some examples to see what really happens. The (affine) line and conic

$$C_1 : x + y + 1 = 0 \qquad \text{and} \qquad C_2 : x^2 + y^2 = 1$$

intersect in the two points $(-1,0)$ and $(0,-1)$, as is easily seen by substituting $y = -x - 1$ into the equation for C_2 and solving the resulting quadratic equation for x. (See Figure A.4(a).) Similarly,

$$C_1 : x + y = 0 \qquad \text{and} \qquad C_2 : x^2 + y^2 = 1$$



Curves Of Degree Two And Degree Four Intersect In Eight Points

Figure A.3

intersect in the two points $\left(\frac{1}{2}\sqrt{2}, -\frac{1}{2}\sqrt{2}\right)$ and $\left(-\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2}\right)$. Note that we have to allow real coordinates for the intersection points, even though C_1 and C_2 are rational curves. (See Figure A.4(b).)

What about the intersection of the line and conic

$$C_1 : x + y + 2 = 0 \qquad \text{and} \qquad C_2 : x^2 + y^2 = 1?$$

They do not intersect at all in the usual Euclidean plane $\mathbb{R}^2$, but if we allow complex numbers then we again find two intersection points

$$\left(-1 + \frac{\sqrt{2}}{2}i, -1 - \frac{\sqrt{2}}{2}i\right) \qquad \text{and} \qquad \left(-1 - \frac{\sqrt{2}}{2}i, -1 + \frac{\sqrt{2}}{2}i\right).$$

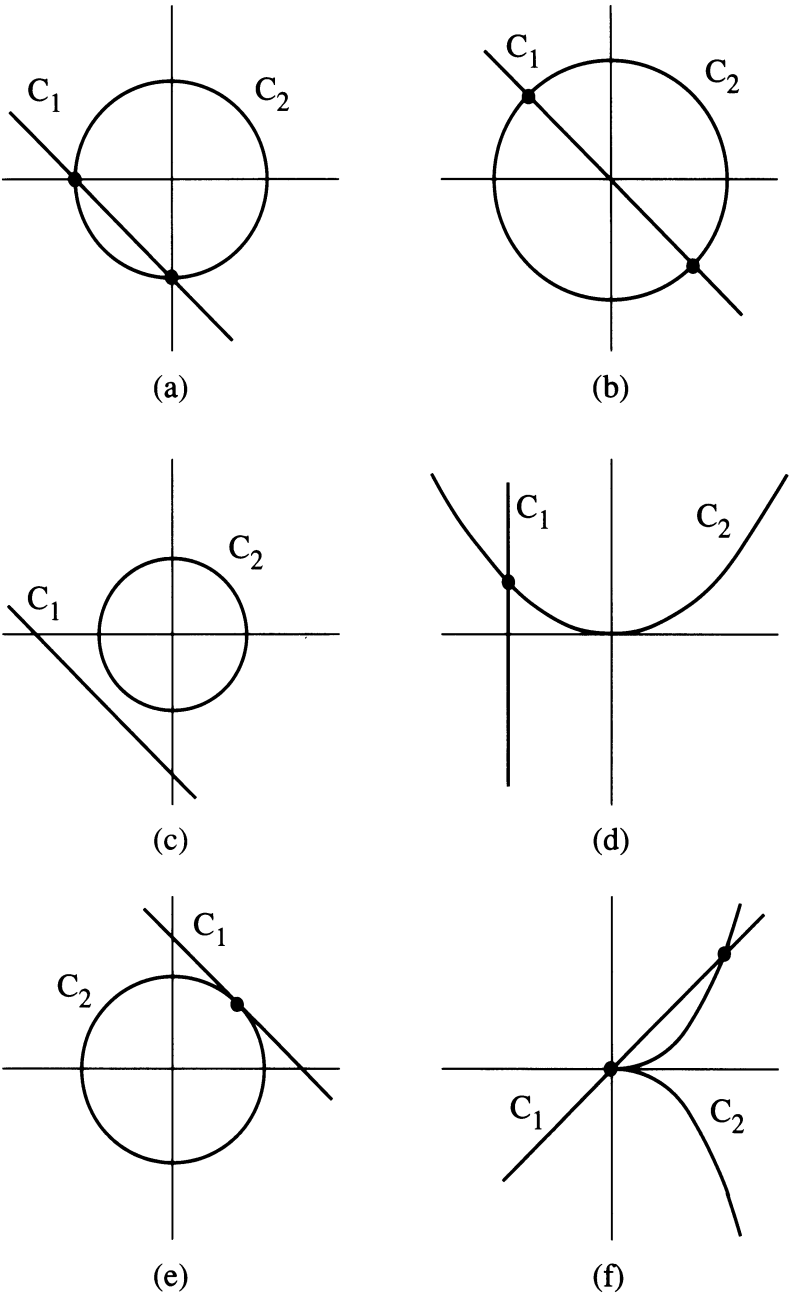
(See Figure A.4(c).) Of course, it is reasonable to allow complex coordinates, since even for polynomials in one variable we need to use complex numbers to ensure that a polynomial of degree d actually has d roots counting multiplicities.

Next we look at

$$C_1 : x + 1 = 0 \qquad \text{and} \qquad C_2 : x^2 - y = 0.$$

These curves appear to intersect in the single point $(-1, 1)$, but appearances can be deceiving. (See Figure A.4(d).) Remember even for two lines, we may need to also look at the points at infinity in $\mathbb{P}^2$. In our case the line C_1 is in the vertical direction, and the tangent lines to the parabola C_2 approach the vertical direction, so geometrically C_1 and C_2 should have a common point at infinity corresponding to the vertical direction. Following our maxim from Section 2, we now check this algebraically. First we homogenize the equations for C_1 and C_2 to get the corresponding projective curves

$$\bar{C}_1 : X + Z = 0 \qquad \text{and} \qquad \bar{C}_2 : X^2 - YZ = 0.$$



Some Of The Ways In Which Curves Can Intersect

Figure A.4

Then $\bar{C}_1 \cap \bar{C}_2$ consists of the two points $[-1,1,1]$ and $[0,1,0]$, as can be seen by substituting $X = -Z$ into the equation for $\bar{C}_2$. So we get the expected two points provided we work with projective curves.

All of this looks very good, but the next example illustrates another problem which may occur. Consider the intersection of the line and conic

$$C_1 : x + y = 2 \qquad \text{and} \qquad C_2 : x^2 + y^2 = 2.$$

(See Figure A.4(e).) Then $C_1 \cap C_2$ consists of the single point $(1,1)$, and even if we go to the projective curves

$$\bar{C}_1 : X + Y = 2Z \qquad \text{and} \qquad \bar{C}_2 : X^2 + Y^2 = 2Z^2$$

we still find the single intersection point $[1,1,1]$. What is wrong?

Geometrically we immediately see the problem, namely the line C_1 is tangent to the circle C_2 at the point $(1,1)$, so in some sense that point should count double. We can also see this algebraically. If we substitute the relation $y = 2 - x$ from C_1 into the equation for C_2 and simplify, we get the equation $2x^2 - 4x + 2 = 0$, or equivalently $2(x-1)^2 = 0$. So we do have a quadratic equation to solve for x, and normally we would expect to find two distinct roots, but in this case we happen to find one root repeated twice. This makes sense, since even a degree d polynomial of one variable can only be said to have d complex roots if we count multiple roots according to their multiplicities.

This multiplicity problem can also occur if one of the curves is singular at P, even if the two curves do not have the same tangent direction. For example, consider the intersection of the line and the degree three curve

$$C_1 : x - y = 0 \qquad \text{and} \qquad C_2 : x^3 - y^2 = 0.$$

(See Figure A.4(f).) Our intuition says that $C_1 \cap C_2$ should consist of three points. Substituting $y = x$ into the equation for C_2 gives $x^3 - x^2 = 0$. This is a cubic equation for x, but it only has two distinct roots, namely $x = 0$ and $x = 1$. Thus $C_1 \cap C_2$ only contains the two points $(0,0)$ and $(1,1)$, but the point $(0,0)$ needs to be counted twice, which gives the expected three points when we count points with their multiplicity.

Finally, we look at an example where things go spectacularly wrong. Consider the intersection of the line and conic

$$C_1 : x + y + 1 = 0 \qquad \text{and} \qquad C_2 : 2x^2 + xy - y^2 + 4x + y + 2 = 0.$$

When we substitute $y = -x - 1$ into the equation for C_2 we find that everything cancels out and we are left with $0 = 0$. This happens because the equation for C_2 factors as

$$2x^2 + xy - y^2 + 4x + y + 2 = (x + y + 1)(2x - y + 2),$$

so every point on C_1 lies on C_2. Notice that C_2 is the union of two curves, namely C_1 and the line $2x - y + 2 = 0$.

In general, if C is a curve given by an equation $C : f(x, y) = 0$, then we factor f into a product of irreducible polynomials

$$f(x, y) = p_1(x, y)p_2(x, y) \cdots p_n(x, y).$$

Note that $\mathbb{C}[x, y]$ is a unique factorization domain, so every polynomial has an essentially unique factorization into such a product. Then the *irreducible components of the curve* C are the curves

$$p_1(x, y) = 0, \qquad p_2(x, y) = 0, \qquad \cdots \qquad p_n(x, y) = 0.$$

We say that C is *irreducible* if it has only one irreducible component, or equivalently if $f(x, y)$ is an irreducible polynomial. Next, if C_1 and C_2 are two curves, we say that C_1 and C_2 *have no common components* if their irreducible components are distinct. It is not hard to prove that $C_1 \cap C_2$ consists of a finite set of points if and only if C_1 and C_2 have no common components. Finally, if we work instead with projective curves C, C_1, C_2, then we make the same definitions using factorizations into products of irreducible homogeneous polynomials in $\mathbb{C}[X, Y, Z]$.

We now consider the general case of projective curves C_1 and C_2, which we assume to have no common components. The intersection $C_1 \cap C_2$ is then a finite set of points with complex coordinates. To each point $P \in \mathbb{P}^2$ we assign a *multiplicity* or *intersection index* $I(C_1 \cap C_2, P)$. This is a non-negative integer reflecting the extent to which C_1 and C_2 are tangent to one another at P or are not smooth at P. We will give a formal definition in Section 4, but one can get a good feeling for the intersection index from the following properties:

(i) If $P \notin C_1 \cap C_2$, then $I(C_1 \cap C_2, P) = 0$.

(ii) If $P \in C_1 \cap C_2$, if P is a non-singular point of C_1 and C_2, and if C_1 and C_2 have different tangent directions at P, then $I(C_1 \cap C_2, P) = 1$. (One often says in this case that C_1 and C_2 intersect *transversally at* P.)

(iii) If $P \in C_1 \cap C_2$ and if C_1 and C_2 do not intersect transversally at P, then $I(C_1 \cap C_2, P) \geq 2$.

(For a proof of these properties, see the last part of Section 4.)

With these preliminaries, we are now ready to formally state the theorem which justifies the plausibility argument we gave at the beginning of this section.

Bezout's Theorem. *Let C_1 and C_2 be projective curves with no common components. Then*

$$\sum_{P \in C_1 \cap C_2} I(C_1 \cap C_2, P) = (\deg C_1)(\deg C_2),$$

where the sum is over all points of $C_1 \cap C_2$ having complex coordinates. In particular, if C_1 and C_2 are smooth curves with only transversal intersections, then $\#(C_1 \cap C_2) = (\deg C_1)(\deg C_2)$; and in all cases there is an inequality

$$\#(C_1 \cap C_2) \leq (\deg C_1)(\deg C_2).$$

PROOF. We will give the proof of Bezout's theorem in Section 4. □

It would be hard to overestimate the importance of Bezout's theorem in the study of projective geometry. We should stress how amazing a theorem it is. The projective plane was constructed so as to ensure that any two lines (i.e., curves of degree 1) intersect in exactly one point, so one could say that the projective plane is formed by taking the affine plane and adding just enough points to make Bezout's theorem true for curves of degree 1. It then turns out that the projective plane has enough points to make Bezout's theorem true for all projective curves!

Sometimes Bezout's theorem is used to determine if two curves are the same, or at least have a common component. For example, if C_1 and C_2 are conics, and if C_1 and C_2 have five points in common, then Bezout's theorem tells us that they have a common component. Since the degree of a component can be no larger than the degree of the curve, it follows that either there is some line L contained in both C_1 and C_2, or else $C_1 = C_2$. Thus there is only one conic going through any given five points as long as no three of the points are collinear. This is analogous to the fact that there is a unique line going through two given points. More generally, one sees from Bezout's theorem if C_1 and C_2 are irreducible curves of degree d with d^2+1 points in common, then $C_1 = C_2$. Note, however, that for $d \geq 3$ there is in general no curve of degree d going through $d^2 + 1$ preassigned points. This is because the number d^2+1 of conditions to be met is greater than the number $(d+1)(d+1)/2$ unknown coefficients of a homogeneous polynomial of degree d.

We now want to consider a slightly more complicated situation. Suppose that C_1 and C_2 are two cubic curves, that is curves of degree 3, which intersect in 9 distinct points $P_1, \ldots, P_9$. Suppose further that D is another cubic curve which happens to go through the first 8 points $P_1, \ldots, P_8$. We claim that D also goes through the ninth point P_9. To see why this is true, we consider the collection of all cubic curves in $\mathbb{P}^2$, which we will denote by $\mathcal{C}^{(3)}$. An element $C \in \mathcal{C}^{(3)}$ is given by a homogeneous equation

$$C : aX^3 + bX^2Y + cXY^2 + dY^3 + eX^2Z + fXZ^2 + gY^2Z + hYZ^2 + iZ^3 + jXYZ = 0,$$

so C is determined by the ten coefficients $a, b, \ldots, j$. Of course, if we multiply the equation for C by any non-zero constant, then we get the same curve, so really C is determined by the homogeneous 10-tuple $[a, b, \ldots, j]$.

Conversely, if two 10-tuples give the same curve, then they differ by multiplication by a constant. In other words, the set of cubic curves $\mathcal{C}^{(3)}$ is in a very natural way isomorphic to the projective space $\mathbb{P}^9$.

Suppose we are given a point $P \in \mathbb{P}^2$ and ask for all cubic curves that go through P. This describes a certain subset of $\mathcal{C}^{(3)} \cong \mathbb{P}^9$, and it is easy to see what this subset is. If P has homogeneous coordinates $P = [X_0, Y_0, Z_0]$, then substituting P into the equation for C shows that C will contain P if and only if the 10-tuple $[a, b, \ldots, j]$ satisfies the homogeneous linear equation

$$(X_0^3)a + (X_0^2Y_0)b + (X_0Y_0^2)c + (Y_0^3)d + (X_0^2Z_0)e + (X_0Z_0^2)f + (Y_0^2Z_0)g + (Y_0Z_0^2)h + (Z_0^3)i + (X_0Y_0Z_0)j = 0.$$

[N.B. This is a linear equation in the 10 variables $a, b, \ldots, j$.] In other words, for a given point $P \in \mathbb{P}^2$, the set of cubic curves $C \in \mathcal{C}^{(3)}$ which contain P corresponds to the zeros of a homogeneous linear equation in $\mathbb{P}^9$.

Similarly, if we fix two points $P, Q \in \mathbb{P}^2$, then the set of cubic curves $C \in \mathcal{C}^{(3)}$ containing both P and Q is given by the common solutions of two linear equation in $\mathbb{P}^9$, where one linear equation is specified by P and the other by Q. Continuing in this fashion, we find that for a collection of n points $P_1, P_2, \ldots, P_n \in \mathbb{P}^2$ there is a one-to-one correpsondence between the two sets

$$\{C \in \mathcal{C}^{(3)} : P_1, \ldots, P_n \in C\} \quad \text{and} \quad \left\{ \begin{array}{c} \text{simultaneous solutions of a} \\ \text{certain system of } n \text{ homo-} \\ \text{geneous linear equations in } \mathbb{P}^9 \end{array} \right\}.$$

For example, suppose we take $n = 9$. The solutions to a system of 9 homogeneous linear equations in 10 variables generally consists of the multiples of a single solution. In other words, if $\mathbf{v_0}$ is a non-zero solution, then every solution will have the form $\lambda \mathbf{v_0}$ for some constant λ. Now let $C_1 : F_1(X, Y, Z) = 0$ and $C_2 : F_2(X, Y, Z) = 0$ be cubic curves in $\mathbb{P}^2$, each going through the given 9 points. The coefficients of F_1 and F_2 are then 10-tuples which are solutions to the given system of linear equations, so we conclude that $F_1 = \lambda F_2$, and hence $C_1 = C_2$. Thus, we find that, in general, there is exactly one cubic curve in $\mathbb{P}^2$ that passes through 9 given points. Note, however, that for special sets of nine points it is possible to have a one parameter family of cubic curves going through them.

That is the situation of our original problem, to which we now return. Namely, we take two cubic curves C_1 and C_2 in $\mathbb{P}^2$ that intersect in nine distinct points $P_1, \ldots, P_9$. Let C_1 and C_2 be given by the equations

$$C_1 : F_1(X, Y, Z) = 0 \qquad \text{and} \qquad C_2 : F_2(X, Y, Z) = 0.$$

We consider the set of all cubic curves $C \in \mathcal{C}^{(3)}$ which pass through the first 8 points $P_1, \ldots, P_8$. This set corresponds to the simultaneous solution of 8 homogeneous linear equations in 10 variables. The set of solutions

of this system consists of all linear combinations of two linearly independent 10-tuples; in other words, if $\mathbf{v}_1$ and $\mathbf{v}_2$ are independent solutions, then every solution has the form $\lambda_1\mathbf{v}_1+\lambda_2\mathbf{v}_2$ for some constants λ_1, λ_2. (In principle, the set of solutions might have dimension greater than two. We leave it as an exercise for you to check that because the eight points $P_1, \dots, P_8$ are distinct, the corresponding linear equations will be independent.)

But we already know two cubic curves passing through $P_1, \dots, P_8$, namely C_1 and C_2. The coefficients of their equations F_1 and F_2 will thus give two 10-tuples solving the system of 8 homogeneous linear equations, so they will span the complete solution set. This means that if D is any other cubic curve in $\mathbb{P}^2$ that contains the 8 points $P_1, \dots, P_8$, then the equation for D has the form

$$D : \lambda_1 F_1(X,Y,Z) + \lambda_2 F_2(X,Y,Z) = 0 \qquad \text{for some constants } \lambda_1, \lambda_2.$$

But the ninth point P_9 is on both C_1 and C_2, so $F_1(P_9) = F_2(P_9) = 0$. It follows from the equation for D that D also contains the point P_9, which is exactly what we have been trying to demonstrate.

More generally, the following theorem is true.

Cayley-Bacharach Theorem. *Let C_1 and C_2 be curves in $\mathbb{P}^2$ without common components of respective degrees d_1 and d_2, and suppose that C_1 and C_2 intersect in d_1d_2 distinct points. Let D be a curve in $\mathbb{P}^2$ of degree $d_1 + d_2 - 3$. If D passes through all but one of the points of $C_1 \cap C_2$, then D must pass through the remaining point also.*

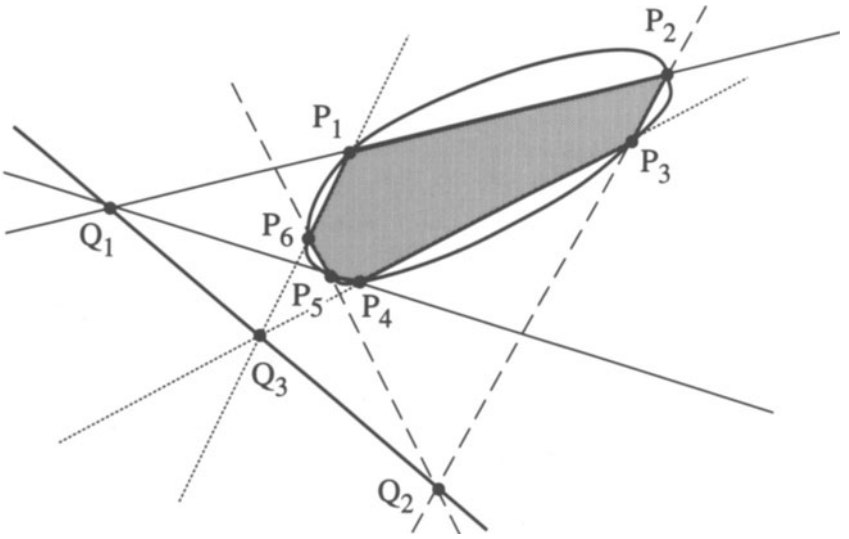
It is not actually necessary that C_1 and C_2 intersect in distinct points. For example, if $P \in C_1 \cap C_2$ is a point of multiplicity 2, say because C_1 and C_2 have the same tangent direction at P, then one needs to require that D also have the same tangent direction at P. The most general result is somewhat difficult to state, so we will content ourselves with the following version:

Cubic Cayley-Bacharach Theorem. *Let C_1 and C_2 be cubic curves in $\mathbb{P}^2$ without common components, and assume that C_1 is smooth. Suppose that D is another cubic curve that contains 8 of the intersection points of $C_1 \cap C_2$ counting multiplicities. This means that if $C_1 \cap C_2 = \{P_1, \dots, P_r\}$, then*

$$\begin{aligned} &I(C_1 \cap D, P_i) \ge I(C_1 \cap C_2, P_i) \quad \text{for } 1 \le i < r, \text{ and} \\ &I(C_1 \cap D, P_r) \ge I(C_1 \cap C_2, P_r) - 1. \end{aligned}$$

Then D goes through the ninth point of $C_1 \cap C_2$. In terms of multiplicities, this means that $I(C_1 \cap D, P_r) = I(C_1 \cap C_2, P_r)$.

We will conclude this section of the appendix by applying the Cayley-Bacharach theorem to prove a beautiful geometric result of Pascal. Let C be a smooth conic, for example a hyperbola, a parabola, or an ellipse. Choose



Pascal's Theorem

Figure A.5

any six points lying on the conic, say labeled consecutively as $P_1, P_2, \ldots, P_6$, and play connect-the-dots to draw a hexagon. Now take the lines through opposite sides of the hexagon and extend them to find the intersection points as illustrated in Figure A.5, say

$$\overleftrightarrow{P_1P_2} \cap \overleftrightarrow{P_4P_5} = \{Q_1\}, \quad \overleftrightarrow{P_2P_3} \cap \overleftrightarrow{P_5P_6} = \{Q_2\}, \quad \overleftrightarrow{P_3P_4} \cap \overleftrightarrow{P_6P_1} = \{Q_3\}.$$

Pascal's Theorem. *The three points Q_1, Q_2, Q_3 described above lie on a line.*

To prove Pascal's theorem, we consider the two cubic curves

$$C_1 = \overleftrightarrow{P_1P_2} \cup \overleftrightarrow{P_3P_4} \cup \overleftrightarrow{P_5P_6} \qquad \text{and} \qquad C_2 = \overleftrightarrow{P_2P_3} \cup \overleftrightarrow{P_4P_5} \cup \overleftrightarrow{P_6P_1}.$$

Why do we call C_1 and C_2 cubic curves? The answer is that if we choose an equation for the line $\overleftrightarrow{P_iP_j}$, say $\alpha_{ij}X + \beta_{ij}Y + \gamma_{ij}Z = 0$, then C_1 is given by the homogeneous cubic equation

$$C_1 : (\alpha_{12}X + \beta_{12}Y + \gamma_{12}Z)(\alpha_{34}X + \beta_{34}Y + \gamma_{34}Z)(\alpha_{56}X + \beta_{56}Y + \gamma_{56}Z) = 0,$$

and similarly for C_2.

Notice that all nine of the points

$$P_1, P_2, P_3, P_4, P_5, P_6, Q_1, Q_2, Q_3 \quad \text{are on both } C_1 \text{ and } C_2.$$

This sets us up to use the Cayley-Bacharach theorem. We take D to be the cubic curve that is the union of our original conic C with the line through Q_1 and Q_2,

$$D = C \cup \overleftrightarrow{Q_1 Q_2}.$$

Clearly D contains the eight points $P_1, P_2, P_3, P_4, P_5, P_6, Q_1, Q_2$. The Cayley-Bacharach theorem then tells us that D contains the ninth point in $C_1 \cap C_2$, namely Q_3. Now Q_3 does not lie on C, since otherwise the line $\overleftrightarrow{P_6 P_1}$ would intersect the conic C in the three points P_6, P_1, Q_3, contradicting Bezout's theorem. Therefore Q_3 must be on the line $\overleftrightarrow{Q_1 Q_2}$. In other words, Q_1, Q_2, and Q_3 are collinear, which completes the proof of Pascal's theorem.

4. Intersection Multiplicities and a Proof of Bezout's Theorem

We give the proof of Bezout's theorem in the form of a long exercise with hints. It is quite elementary. For the first weak inequality (which is all that is needed in many important applications of the theorem) we use only linear algebra and the notion of the dimension of a vector space. After that we need the concepts of commutative ring, ideal, and quotient ring, and the fact that unique factorization holds in polynomial rings, but that is about all.

Let C_1 and C_2 be curves in $\mathbb{P}^2$ of respective degrees n_1 and n_2, without common components. Until the last step of the proof we assume that the line at infinity is not a component of either curve, and we work with affine coordinates x, y. Let

$$f_1(x, y) = 0 \qquad \text{and} \qquad f_2(x, y) = 0$$

be the equations for the two curves in the affine plane $\mathbb{A}^2$. The assumptions we have made mean that the polynomials f_1 and f_2 have no common factor and are of degree n_1 and n_2 respectively.

The proof is pure algebra (though the geometric ideas behind it should be apparent) and works over any algebraically closed ground field k. The reader is welcome to take $k = \mathbb{C}$, but k could also be an algebraic closure of the finite field $\mathbb{F}_p$, for example.

Let $R = k[x, y]$ be a polynomial ring in two variables and let $(f_1, f_2) = Rf_1 + Rf_2$ be the ideal in R generated by the polynomials f_1 and f_2. The steps in the proof of Bezout's theorem are as follows:

(1) We prove the following two inequalities which, on eliminating the middle term, show that the number of intersection points of C_1 and C_2 in $\mathbb{A}^2$ is at most $n_1 n_2$:

$$\#(C_1 \cap C_2 \cap \mathbb{A}^2) \overset{(A)}{\leq} \dim\left(\frac{R}{(f_1, f_2)}\right) \overset{(B)}{\leq} n_1 n_2.$$

Note: In this section, dim means the dimension as a k-vector space.
(2) We show that (B) is an equality if C_1 and C_2 do not meet at infinity.
(3) We strengthen (A) to get

$$\sum_{P \in C_1 \cap C_2 \cap \mathbb{A}^2} I(C_1 \cap C_2, P) \overset{(A^+)}{\leq} \dim\left(\frac{R}{(f_1, f_2)}\right),$$

where $I(C_1 \cap C_2, P)$ is a suitably defined *intersection multiplicity* of C_1 and C_2 at P.
(4) We show that (A^+) is in fact an equality.

The fact that k is algebraically closed is not needed for the inequalities (1) and (3), but is essential for the equalities (2) and (4). Taken together, (2) and (4) give Bezout's theorem in case C_1 and C_2 do not meet at infinity. To get it in general there is one more step.
(5) We show that the definition of intersection multiplicity does not change when we make a projective transformation, and that there is a line L in $\mathbb{P}^2$ not meeting any intersection point. Changing coordinates so that the line L is the line at infinity, we then get Bezout in general.

To round out the argument we include one more segment:
(6) We prove some basic properties satisfied by the intersection multiplicity $I(C_1 \cap C_2, P)$ and show that it depends only on the initial part of the Taylor expansions of f_1 and f_2 at P.

Now we sketch the proof as a series of exercises with hints, breaking each of the segments (1)–(5) into smaller steps.
(1.1) Let $P_1, P_2, \ldots, P_m$ be m different points in the (x, y) plane. Show that for each i there is a polynomial $h_i = h_i(x, y)$ such that $h_i(P_i) = 1$ and $h_i(P_j) = 0$ for $j \neq i$. (*Idea.* Construct h_i as a product of linear polynomials, using the fact that for each $j \neq i$ there is a line through P_j not meeting P_i.)
(1.2) Suppose that the m points P_i from (1.1) lie in $C_1 \cap C_2$. Prove that the polynomials h_i are linearly independent modulo (f_1, f_2), and consequently that

$$m \leq \dim\left(\frac{R}{(f_1, f_2)}\right).$$

This proves inequality (A). (*Idea.* Consider a possible dependence

$$c_1 h_1 + c_2 h_2 + \cdots + c_m h_m = g_1 f_1 + g_2 f_2 \in (f_1, f_2)$$

with $c_i \in k$. Substitute P_i into the equation to show that every $c_i = 0$.)

This takes care of inequality (A). To prove (B) we define for each integer $d \geq 0$,

$$\begin{aligned}
\phi(d) &= \frac{1}{2}(d+1)(d+2) = \frac{1}{2}d^2 + \frac{3}{2}d + 1, \\
R_d &= (\text{vector space of polynomials } f(x, y) \text{ of degree} \leq d), \\
W_d &= R_{d-n_1} f_1 + R_{d-n_2} f_2.
\end{aligned}$$

Thus W_d is the k-vector space of polynomials of the form

$$f = g_1 f_1 + g_2 f_2 \qquad \text{with} \quad \deg g_i \leq d - n_i \quad \text{for} \quad i = 1, 2.$$

Notice that $W_d = 0$ if $d < \max\{n_1, n_2\}$, and in any case $V_d \subset (f_1, f_2)$.
(1.3) Show that $\dim R_d = \phi(d)$. (*Idea.* One way to see this is to note that

$$\phi(d) - \phi(d-1) = (\text{number of monomial } x^i y^j \text{ of degree } d) = d + 1$$

and use induction on d.)
(1.4) For $d \geq n_1 + n_2$, show that

$$R_{d-n_1} f_1 \cap R_{d-n_2} f_2 = R_{d-n_1-n_2} f_1 f_2.$$

(Here we use the hypothesis that f_1 and f_2 have no common factor.)
(1.5) Prove that for $d \geq n_1 + n_2$,

$$\dim R_d - \dim W_d = \phi(d) - \phi(d-n_1) - \phi(d-n_2) + \phi(d-n_1-n_2) = n_1 n_2.$$

(*Idea.* If f is a non-zero polynomial, then $g \mapsto gf$ defines an isomorphism $R_{d-j} \xrightarrow{\sim} R_{d-j} f$; hence $\dim R_{d-j} f = \phi(d-j)$. Now use the lemma from linear algebra which says that

$$\dim(U + V) = \dim(U) + \dim(V) - \dim(U \cap V)$$

for subspaces U, V of a finite dimensional vector space.)
(1.6) Prove inequality (B) by showing that if g_j, $1 \leq j \leq n_1 n_2 + 1$, are elements of R, then they are linearly dependent modulo (f_1, f_2). (*Idea.* Take d so large that the g_j are in R_d and so that (1.5) holds. Then use (1.5) to show that there is a non-trivial linear combination $g = \sum c_j g_j$ such that $g \in W_d \subset (f_1, f_2)$.

This finishes segment (1). For segment (2) we begin by recalling how one computes the intersections of an affine curve $f(x, y) = 0$ with the line at infinity.
(2.1) For each non-zero polynomial $f = f(x, y)$, let f^* denote the homogeneous part of f of highest degree. In other words, if

$$f = \sum_{i,j} c_{ij} x^i y^j \quad \text{has degree } n, \text{ then} \quad f^* = \sum_{i+j=n} c_{ij} x^i y^j.$$

Because k is algebraically closed, we can factor f^* into linear factors,

$$f^*(x, y) = \prod_{i=1}^{n} (a_i x + b_i y), \qquad a_i, b_i \in k, \quad n = \deg f = \deg f^*.$$

Show that the points at infinity on the curve $f(x, y) = 0$ are the points with homogeneous coordinates

$$[X, Y, Z] = [b_i, -a_i, 0].$$

(*Idea.* Put $x = X/Z$, $y = Y/Z$, etc.)

An example should make this clearer. Consider the polynomials

$$f(x,y) = x^4 - x^2y^2 + 3x^3 + xy^2 + 2y^3 + 2y^2 + 8x + 3,$$
$$f^*(x,y) = x^4 - x^2y^2 = x^2(x+y)(x-y),$$

each of which has degree 4. The quartic curve $f(x,y) = 0$ thus meets the line at infinity in the points $(0,1,0)$, $(1,-1,0)$, $(1,1,0)$. The fact that x^2 divides $f^*(x,y)$ means that the curve is tangent to the line at infinity at the point $(0,1,0)$.

The remaining steps in segment 2 are as follows:

(2.2) If C_1 and C_2 do not meet at infinity, show that f_1^* and f_2^* have no common factor.

(2.3) If f_1^* and f_2^* have no common factor, show that $(f_1, f_2) \cap R_d = W_d$ for all $d \geq n_1 + n_2$.

(2.4) If $(f_1, f_2) \cap R_d = W_d$ and $d \geq n_1 + n_2$, show that

$$\dim\left(\frac{R}{(f_1, f_2)}\right) \geq n_1 n_2.$$

(*Idea.* (2.2) is an easy consequence of (2.1). To do (2.3) we suppose that $f \in (f_1, f_2) \cap R_d$ is written in the form $f = g_1 f_1 + g_2 f_2$ with g_1 and g_2 of smallest possible degree. If $\deg g_1 > d - n_1$, then looking at the terms of highest degree shows that $g_1^* f_1^* + g_2^* f_2^* = 0$. Then use the fact that f_1^* and f_2^* are relatively prime to show that there is an h such that

$$\deg(g_1 + hf_2) < \deg(g_1) \qquad \text{and} \qquad \deg(g_2 - hf_1) < \deg(g_2).$$

Deduce that $\deg g_i \leq d - n_i$, and hence that $f \in W_d$. For (2.4) note that by (1.5) there are $n_1 n_2$ element in R_d which are linearly independent modulo W_d, and that if $(f_1, f_2) \cap R_d = W_d$, then they are linearly independent as elements of R modulo (f_1, f_2). Hence, $\dim R/(f_1, f_2) \geq n_1 n_2$.)

To define intersection multiplicity we introduce the important notion of the *local ring* $\mathcal{O}_P$ of a point $P \in \mathbb{A}^2$. Let $K = k(x,y)$ be the fraction field of $R = k[x,y]$, that is, K is the field of rational functions of x and y. For a point $P = (a,b)$ in the (x,y) plane and a rational function $\phi = f(x,y)/g(x,y) \in K$, we say that ϕ *is defined at* P if $g(a,b) \neq 0$, and then we put

$$\phi(P) = \frac{f(a,b)}{g(a,b)} = \frac{f(P)}{g(P)}.$$

For a given point P we define the *local ring of* P to be the set of all $\phi \in K$ which are defined at P. We leave the following basic properties of $\mathcal{O}_P$ as exercises. First, $\mathcal{O}_P$ is a subring of K, and the map $\phi \mapsto \phi(P)$ is a homomorphism of $\mathcal{O}_P$ onto k which is the identity on k. Let

$$M_P = \{\phi \in \mathcal{O}_P : \phi(P) = 0\}$$

be the kernel of that homomorphism. Then $\mathcal{O}_P$ is equal to a direct sum $\mathcal{O}_P = k + M_P$, and $\mathcal{O}_P/M_P \cong k$. An element $\phi \in \mathcal{O}_P$ has a multiplicative inverse in $\mathcal{O}_P$ if and only if $\phi \notin M_P$. Every ideal of $\mathcal{O}_P$ other than $\mathcal{O}_P$ itself is contained in M_P, and so M_P is the unique maximal ideal of $\mathcal{O}_P$. (A ring having a unique maximal ideal is called a *local ring*. We used another local ring $R_p \subset \mathbb{Q}$ in Chapter II, Section 4. See also Exercise 2.7.)

Now let $(f_1, f_2)_P = \mathcal{O}_P f_1 + \mathcal{O}_P f_2$ denote the ideal in $\mathcal{O}_P$ generated by f_1 and f_2. Our definition of the *intersection multiplicity* (also called the *intersection index*) *of* C_1 *and* C_2 *at* P is

$$I(C_1 \cap C_2, P) = \dim\left(\frac{\mathcal{O}_P}{(f_1, f_2)_P}\right).$$

We are now ready to do segment (3), which means taking inequality (A) and strengthening it to inequality (A^+).

(3.1) Show that

$$\dim\left(\frac{\mathcal{O}_P}{(f_1, f_2)_P}\right) \le \dim\left(\frac{R}{(f_1, f_2)}\right).$$

Deduce from inequality (B) that the intersection multiplicity $I(C_1 \cap C_2, P)$ is finite. (*Idea.* Note that any finite set of elements in $\mathcal{O}_P$ can be written over a common denominator. Show that if $g_1/h, g_2/h, \ldots, g_r/h$ are elements of $\mathcal{O}_P$ which are linearly independent modulo $(f_1, f_2)_P$, then $g_1, g_2, \ldots g_r$ are elements of R which are independent modulo (f_1, f_2).)

(3.2) Show that $\mathcal{O}_P = R + (f_1, f_2)_P$. (*Idea.* By (3.1) we may suppose that the elements g_i/h span $\mathcal{O}_P$ modulo $(f_1, f_2)_P$, and because $h^{-1} \in \mathcal{O}_P$, it follows that the polynomials g_i span $\mathcal{O}_P$ modulo $(f_1, f_2)_P$.)

(3.3) Show that if $P \notin C_1 \cap C_2$, then $I(C_1 \cap C_2, P) = 0$. Show that if $P \in C_1 \cap C_2$, then

$$(f_1, f_2)_P \subset M_P \qquad \text{and} \qquad I(C_1 \cap C_2, P) = 1 + \dim\left(\frac{M_P}{(f_1, f_2)_P}\right).$$

Conclude that if $P \in C_1 \cap C_2$, then $I(C_1 \cap C_2, P) \ge 1$, with equality if and only if $(f_1, f_2)_P = M_P$.

(3.4) Suppose that $P \in C_1 \cap C_2$. Let r satisfy $r \ge \dim(\mathcal{O}_P/(f_1, f_2)_P)$. Show that $M_P^r \subset (f_1, f_2)_P$. (*Idea.* We are to prove that, given any collection of r elements $t_1, t_2, \ldots, t_r$ in M_P, their product $t_1 t_2 \cdots t_r$ is in $(f_1, f_2)_P$. Define a sequence of ideals J_i in $\mathcal{O}_P$ by

$$J_i = t_1 t_2 \cdots t_i \mathcal{O}_P + (f_1, f_2)_P \quad \text{for } 1 \le i \le r, \text{ and } \quad J_{r+1} = (f_1, f_2)_P.$$

Then

$$M_P \supset J_1 \supset J_2 \supset \cdots \supset J_r \supset J_{r+1} = (f_1, f_2)_P.$$

Since $r > \dim\big(M_P/(f_1,f_2)_P\big)$, it follows that $J_i = J_{i+1}$ for some i with $1 \le i \le r$. If $i = r$, then $t_1t_2\cdots t_r \in (f_1,f_2)_P$ and we are done. If $i < r$, then we have

$$t_1t_2\cdots t_i = t_1t_2\cdots t_{i+1}\phi + \psi \qquad \text{with } \phi \in \mathcal{O}_P \text{ and } \psi \in (f_1,f_2)_P,$$

so $t_1t_2\cdots t_i(1 - t_{i+1}\phi) = \psi \in \mathcal{O}_P$. But $(1 - t_{i+1}\phi)(P) = 1$, so we have that $(1-t_{i+1}\phi)^{-1} \in \mathcal{O}_P$. Hence $t_1t_2\ldots t_r = \psi t_{i+1}\cdots t_r(1-t_{i+1}\phi)^{-1} \in \mathcal{O}_P$ as claimed.)

(3.5) Let $P \in C_1 \cap C_2 \cap \mathbb{A}^2$, and let $\phi \in \mathcal{O}_P$. Show that there exists a polynomial $g \in R$ such that

$$\begin{aligned} &g \equiv \phi \,(\mathrm{mod}\ (f_1,f_2)_P) \quad \text{and} \\ &g \equiv 0 \,(\mathrm{mod}\ (f_1,f_2)_Q) \quad \text{for all } Q \ne P,\ Q \in C_1 \cap C_2 \cap \mathbb{A}^2. \end{aligned}$$

(*Idea.* The inequalities (A) and (B) already proved show that ony a finite number of points are involved here (at most n_1n_2 in fact). Hence, by (1.1) there is a polynomial $h = h(x,y) \in R$ such that $h(P) = 1$ and $h(Q) = 0$ for $Q \ne P$, $Q \in C_1 \cap C_2 \cap \mathbb{A}^2$. This means $h^{-1} \in \mathcal{O}_P$ and $h \in M_Q$ for each of the other points Q. For integers $r \ge 1$ we have $h^{-r} \in \mathcal{O}_P$, and if r is sufficiently large, then, by (3.4), we will have $h^r \in (f_1,f_2)_Q$ for the other points Q. By (3.2) there is a polynomial $f \in R$ such that $f \equiv \phi h^{-r} \,(\mathrm{mod}\ (f_1,f_2)_P)$. Then $g = fh^r$ solves the problem.)

(3.6) Show that the natural map

$$R \longrightarrow \prod_{P \in C_1\cap C_2\cap \mathbb{A}^2} \frac{\mathcal{O}_P}{(f_1,f_2)_P} \qquad (*)$$

given by

$$f \longmapsto \big(\cdots, f \bmod (f_1,f_2)_P, \cdots\big)_{P\in C_1\cap C_2\cap \mathbb{A}^2}$$

is surjective, and conclude that the inequality (A^+) holds. (*Idea.* Let J be the kernel of the map. Then $(f_1,f_2) \subset J$, so $\dim R/(f_1,f_2) \ge \dim(R/J)$. The surjectivity of the map follows easily from (3.5) and implies that

$$\begin{aligned} \dim \frac{R}{J} &= (\text{dimension of the target space}) \\ &= \sum_P \dim \frac{\mathcal{O}_P}{(f_1,f_2)_P} = \sum_P I(C_1 \cap C_2, P).\ ) \end{aligned}$$

To prove that (A^+) is an equality is now seen to be the same as showing that the kernel J of the map $(*)$ is equal to (f_1,f_2). So we must show that $J \subset (f_1,f_2)$, the other inclusion being obvious. Let $f \in J$. Our strategy for showing $f \in (f_1,f_2)$ is to consider the set

$$L = \{g \in R : gf \in (f_1,f_2)\}$$

and to prove that $1 \in L$.
(4.1) Show that L is an ideal in R and that $(f_1, f_2) \subset L \subset R$.
(4.2) Show that L has the following property:

For every $P \in \mathbb{A}^2$ there is a polynomial $g \in L$ such that $g(P) \neq 0$. (**)

In fact, property (**) alone implies that $1 \in L$ by the famous "Nullstellensatz" of Hilbert. But we don't need the Nullstellensatz in full generality, because we have an additional piece of information about L, namely that $(f_1, f_2) \subset L$, and hence $\dim(R/L)$ is finite. Using this, and assuming that $1 \notin L$ in order to prove a contradiction, verify the following assertion.
(4.3) There is an $a \in k$ such that $1 \notin L + R(x-a)$. (*Idea.* The powers of x cannot all be linearly indpendent modulo L, so there are constants $c_i \in k$ and an integer n such that $x^n + c_1 x^{n-1} + \cdots + c_n \in L$. Since k is algebraically closed, we can write this as $(x-a_1)(x-a_2)\cdots(x-a_n) \in L$ with suitable $a_i \in k$. Show that if $1 \in L + R(x-a_i)$ for all $i = 1, \ldots, n$, then we get a contradiction to the assumption that $1 \notin L$.)
(4.4) There is a $b \in k$ such that $1 \notin L + R(x-a) + R(y-b)$. (*Idea.* Replace L by $L + R(x-a)$ and x by y and repeat the argument of (4.3).)
(4.5) Let $P = (a, b)$ and show that $g(P) = 0$ for all $g \in L$. This contradicts (4.2) and shows that $1 \in L$. (*Idea.* Write

$$g(x,y) = g\big(a+(x-a), b+(y-b)\big) = g(a,b) + g_1(x,y)(x-a) + g_2(x,y)(y-b)$$

and conclude that $g(a, b) \in L$.)

Our next job is to describe $K, \mathcal{O}_P, M_P, (f_1, f_2)_P$ in terms of homogeneous coordinates, so that they make sense also for points P at infinity. This will allow us to check that they are invariant under an arbitrary projective coordinate change in $\mathbb{P}^2$. To see what to do we put as usual $x = X/Z$, $y = Y/Z$, and we view $R = k[x, y] = k[X/Z, Y/Z]$ as a subring of the field $k(X, Y, Z)$ of rational functions of X, Y, Z. Then $K = k(x, y)$ becomes identified with the set of all rational functions $\Phi = F/G$ of X, Y, Z which are *homogeneous of degree* 0 in the sense that F and G are homogeneous polynomials of the same degree. Indeed, for $\phi \in K$ we have

$$\phi(x,y) = \frac{f(x,y)}{g(x,y)} = \frac{Z^n f(X/Z, Y/Z)}{Z^n g(X/Z, Y/Z)} = \frac{F(X,Y,Z)}{G(X,Y,Z)} = \Phi(X,Y,Z), \text{ say,}$$

where F and G are homogeneous of the same degree $n = \max\{\deg f, \deg g\}$. On the other hand, if $\Phi = F/G$ is a quotient of forms of the same degree, then $\Phi(tX, tY, tZ) = \Phi(X, Y, Z)$, and

$$\Phi(X,Y,Z) = \Phi(x,y,1) = \frac{F(x,y,1)}{G(x,y,1)} \in K.$$

If $P = [A, B, C]$ is a point in $\mathbb{P}^2$ and $\Phi = F/G \in K$, we say that Φ is *defined at* P if $G(A, B, C) \neq 0$, i.e., if P is not on the curve $G(X, Y, Z) = 0$.

If Φ is defined at P, we put $\Phi(P) = F(A,B,C)/G(A,B,C)$, this ratio being independent of the choice of homogeneous coordinate triple for P. Clearly we should put

$$\mathcal{O}_P = \{\Phi \in K : \Phi \text{ is defined at } P\} \quad \text{and} \quad M_P = \{\Phi \in \mathcal{O}_P : \Phi(P) = 0\}.$$

We leave it to the conscientious reader to check the following assertion.
(5.1) If $P = (a,b) = [a,b,1] \in \mathbb{A}^2$, then these definitions of $\mathcal{O}_P$, of $\Phi(P)$ for $\Phi \in \mathcal{O}_P$, and of M_P coincide with our earlier definitions.

Now let $C_1 : F_1 = 0$ and $C_2 : F_2 = 0$ be two curves in $\mathbb{P}^2$ without any common components. Let $f_1(x,y) = F_1(x,y,1)$ and $f_2(x,y) = F_2(x,y,1)$ be the polynomials defining their affine parts. Define

$$(F_1,F_2)_P = \left\{ \frac{F}{G} \in \mathcal{O}_P : F \text{ is of the form } F = H_1F_1 + H_2F_2 \right\}.$$

(Do you see why we cannot just say that $(F_1,F_2)_P$ is the ideal in $\mathcal{O}_P$ generated by F_1 and F_2?)
(5.2) Check that if $P \in \mathbb{A}^2$, then $(F_1,F_2)_P = (f_1,f_2)_P$ is the ideal in $\mathcal{O}_P$ generated by f_1 and f_2.

Of course, we now define the intersection multiplicity of C_1 and C_2 at every point $P \in \mathbb{P}^2$ by

$$I(C_1 \cap C_2, P) = \dim \frac{\mathcal{O}_P}{(F_1,F_2)_P}.$$

We know from (5.2) this coincides with our earlier definition for $P \in \mathbb{A}^2$.
(5.3) Check that the definitions of $\mathcal{O}_P$ and $(F_1,F_2)_P$, and hence also of the intersection multiplicity $I(C_1 \cap C_2, P)$, are independent of our choice of homogeneous coordinates in $\mathbb{P}^2$, i.e., they are are invariant under a linear change of the coordinates X,Y,Z.

To finally complete our proof of Bezout's theorem, we must show that there is a line L in $\mathbb{P}^2$ which does not meet $C_1 \cap C_2$. Then we can take a new coordinate system in which L is the line at infinity, and thereby reduce to the case already proved. To show that L exists, we use the following:
(5.4) Prove that, given any finite set S of points in $\mathbb{P}^2$, there is a line L not meeting S. (*Idea.* Use that an algebraically closed field k is not finite.)

Finally, the next result allows us to apply (5.4).
(5.5) Prove that $C_1 \cap C_2$ is finite. (*Idea.* Use the fact that for every line L which is not a component of either C_1 or C_2, we know (by putting L at infinity and using part (1) of this proof) that $C_1 \cap C_2$ contains a finite number of points not on L.)

That completes our proof of Bezout's Theorem in all its gory detail. To study more closely the properties of the intersection multiplicity $I(C_1 \cap C_2, P)$ at one point P, we can without loss of generality choose coordinates so that $P = (0,0) = [0,0,1]$ is the origin in the affine plane, and we can work with affine coordinates x,y. Let $R = k[x,y]$ as before, and let

$$M = \{f = f(x,y) \in R : f(P) = f(0,0) = 0\}.$$

(6.1) Prove that $M = (x, y) = Rx + Ry$ and $M_P = \mathcal{O}_P x + \mathcal{O}_P y$.

It follows that for each $n \geq 1$, M^n is the ideal in R generated by the monomials $x^n, x^{n-1}y, \ldots, xy^{n-1}, y^n$. Hence every polynomial $f \in R$ can be written uniquely as polynomial of degree at most n plus a remainder polynomial $r \in M^{n+1}$:

$$f(x,y) = c_{00} + c_{10}x + c_{01}y + \cdots + c_{ij}x^i y^j + \cdots + c_{n0}x^n + c_{n-1,1}x^{n-1}y + \cdots + c_{0n}y^n + r. \qquad (*)$$

(6.2) Prove that every $\phi = f/g \in \mathcal{O}_P$ can be written uniquely in the form $(*)$ with $c_{ij} \in k$ and $r \in M_P^{n+1}$. In other words, the inclusion $R \subset \mathcal{O}_P$ induces an isomorphism $R/M^{n+1} \cong \mathcal{O}_P/M_P^{n+1}$ for every $n \geq 0$. (*Idea.* We must show that $\mathcal{O}_P = R + M_P^{n+1}$ and that $R \cap M_P^{n+1} = M^{n+1}$. For the first, show that every $\phi \in \mathcal{O}_P$ can be written in the form $\phi = f/(1-h)$ with $f \in R$ and $h \in M$. Hence

$$\phi = \frac{f}{1-h} = f \cdot (1 + h + \cdots + h^n) + \frac{fh^{n+1}}{1-h} \in R + M_P^{n+1}.$$

The second reduces to showing that if $gf \in M^n$ and $g(P) \neq 0$, then $f \in M^n$. This can be done by considering the terms of *lowest degree* in g and f and gf.)

Now we can already compute some intersection indices to see if our definition gives answers which are geometrically reasonable. As a matter of notation we introduce the symbol

$$I(f_1, f_2) = \dim\left(\frac{\mathcal{O}_P}{(f_1, f_2)_P}\right)$$

for the intersection multiplicity of two curves $f_1 = 0$ and $f_2 = 0$ at the origin.

(6.3) Check that the curve $y = x^n$ and the x axis intersect with multiplicity n at the origin, i.e., show that $I(y - x^n, y) = n$. (*Idea.* Note first that the ideals $(y - x^n, y)$ and (x^n, y) are equal, and that this ideal contains M^n. Then, using what we know (6.2) about $\mathcal{O}_P/M_P^n$, show that $1, x, \ldots, x^{n-1}$ is a basis for the vector space $\mathcal{O}_P/(x^n, y)\mathcal{O}_P$.

(6.4) (Nakayama's Lemma) Suppose J is an ideal of $\mathcal{O}_P$ contained in a finitely generated ideal $\Phi = (\phi_1, \phi_2, \ldots, \phi_m)\mathcal{O}_P$. Suppose some elements of J generate Φ modulo $M_P\Phi$, i.e., $\Phi = J + M_P\Phi$. Then $J = \Phi$. (*Idea.* The case $\Phi = (\phi_1, \phi_2)\mathcal{O}_P$ is all we need. To prove that case, write

$$\phi_1 = j_1 + \alpha\phi_1 + \beta\phi_2, \qquad \phi_2 = j_2 + \gamma\phi_1 + \delta\phi_2,$$

with $j_1, j_2 \in J$ and $\alpha, \beta, \gamma, \delta \in M_P$. Then use the fact that the determinant of the matrix $\begin{pmatrix} 1-\alpha & \beta \\ \gamma & 1-\delta \end{pmatrix}$ is non-zero in order to express the ϕ's in terms of the j's.)

(6.5) Suppose that

$$f_1 = ax + by + (\text{higher terms}), \qquad f_2 = cx + dy + (\text{higher terms}),$$

where "higher terms" means elements of M^2. Show that the following are equivalent.

(i) The curves $f_1 = 0$ and $f_2 = 0$ meet transversally at the origin, i.e., are smooth with distinct tangent directions there.
(ii) The determinant $ad - bc$ is not equal to zero.
(iii) $(f_1, f_2)_P = M_P$, i.e., $I(f_1, f_2) = 1$.

(*Idea.* (i) $\iff$ (ii) follows directly from the definitions. One way to do (ii) $\Rightarrow$ (iii) is to use (6.4) with $\phi_1 = x$, $\phi_2 = y$, and $J = (f_1, f_2)_P$. To do (iii) $\Rightarrow$ (ii) note that if $ad - bc = 0$, then

$$\dim\left(\frac{(f_1, f_2)_P + M_P^2}{M_P^2}\right) \leq 1,$$

whereas, by (6.2), $\dim(M_P/M_P^2) = 2$.
(6.6) Let $f(x, y) \in R$. Show that $I\big(f(x, y), y\big) = m$, where x^m is the highest power of x dividing $f(x, 0)$. (*Idea.* Use the fact that the ideal $\big(f(x, y), y\big)$ is the same as the ideal $\big(f(x, 0), y\big)$. Then argue as in (6.3).)
(6.7) Let $C : F(X, Y, Z) = 0$ be a curve in $\mathbb{P}^2$ that does not contain the line $L_\infty : Z = 0$. Show that for each point $Q = [a, b, 0] \in L_\infty$, we have $I(C \cap L_\infty, Q) = m$, where $(bX - aY)^m$ is the highest power of $bX - aY$ dividing $F(X, Y, 0)$. (*Idea.* Make a suitable coordinate change to reduce to (6.6).)

5. Reduction Modulo p

Let $\mathbb{P}^2(\mathbb{Q})$ denote the set of rational points in $\mathbb{P}^2$. We say that a homogeneous coordinate triple $[A, B, C]$ is *normalized* if A, B, C are integers with no common factors. Each point $P \in \mathbb{P}^2(\mathbb{Q})$ has a normalized coordinate triple which is unique up to sign. To obtain it we start with any triple of rational coordinates, multiply through by a common denominator, and then divide the resulting triple of integers by their greatest common divisor. For exampe

$$\left[\frac{4}{5}, -\frac{2}{3}, 2\right] = [12, -10, 30] = [6, -5, 15].$$

The other normalized coordinate triple for this point is $[-6, 5, -15]$.

Let p be a fixed prime number and for each integer $m \in \mathbb{Z}$, let $\tilde{m} \in \mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ denote its residue modulo p. If $[l, m, n]$ is a normalized coordinate triple for a point $P \in \mathbb{P}^2(\mathbb{Q})$, then the triple $[\tilde{l}, \tilde{m}, \tilde{n}]$ defines a point $\tilde{P}$ in $\mathbb{P}^2(\mathbb{F}_p)$ because at least one of the three numbers l, m, and n is not

divisible by p. Since P determines the triple $[l,m,n]$ up to sign, the point $\tilde{P}$ depends only on P, not on the choice of coordinates for P. Thus, $P \mapsto \tilde{P}$ gives a well-defined map

$$\mathbb{P}^2(\mathbb{Q}) \longrightarrow \mathbb{P}^2(\mathbb{F}_p),$$

called for obvious reasons the *reduction mod p map*. Note that reduction mod p does not map $\mathbb{A}^2(\mathbb{Q})$ into $\mathbb{A}^2(\mathbb{F}_p)$. For example,

$$P = \left(\frac{1}{p}, 0\right) = \left[\frac{1}{p}, 0, 1\right] = [1,0,p] \longmapsto [1,0,0] \notin \mathbb{A}^2(\mathbb{F}_p).$$

In fact, if $P = (a,b) = [a,b,1] \in \mathbb{A}^2(\mathbb{Q})$, then its reduction $\tilde{P}$ is in $\mathbb{A}^2(\mathbb{F}_p)$ if and only if the rational numbers a and b are "p-integral," i.e., have denominators prime to p.

Let $C : F(X,Y,Z) = 0$ be a rational curve in $\mathbb{P}^2$. By rational we mean as usual that the coefficients of F are rational numbers. Clearing the denominators of the coefficients and then dividing by the greatest common divisor of their numerators, we can suppose that the coefficients of F are integers with greatest common divisor equal to one. Call such an F *normalized*. Then $\tilde{F}$, the polynomial we obtain by reducing the coefficients of F modulo p, is non-zero and defines a curve $\tilde{C}$ in characteristic p. If $[l,m,n]$ is a normalized coordinate triple and if $F(l,m,n) = 0$, then $\tilde{F}(\tilde{l},\tilde{m},\tilde{n}) = 0$, because $x \to \tilde{x}$ is a homomorphism. In other words, if P is a rational point on C, then $\tilde{P}$ is a point on $\tilde{C}$; reduction mod p takes $C(\mathbb{Q})$ and maps it into $\tilde{C}(\mathbb{F}_p)$.

If C_1 and C_2 are two curves, it follows that

$$\widetilde{\big(C_1(\mathbb{Q}) \cap C_2(\mathbb{Q})\big)} \subset \tilde{C}_1(\mathbb{F}_p) \cap \tilde{C}_2(\mathbb{F}_p).$$

Is there some sense in which $\widetilde{(C_1 \cap C_2)} = \tilde{C}_1 \cap \tilde{C}_2$ if we count multiplicities? After all, the degrees of the reduced curves $\tilde{C}_i$ are the same as those of the C_i, so by Bezout's theorem the intersection before and after reduction has the same number of points if we count multiplicities. But Bezout's theorem requires that the ground field be algebraically closed, and we don't have the machinery to extend our reduction mod p map to that case. However, if we assume that all of the complex intersection points are rational, then everything is okay. We treat only the special case in which one of the curves is a line. This case suffices for the application to elliptic curves we are after, and it is easy to prove.

Proposition. *Suppose C is a rational curve and L is a rational line in $\mathbb{P}^2$. Suppose that all of the complex intersection points of C and L are rational. Let $C \cap L = \{P_1, P_2, \ldots, P_d\}$, where $d = \deg(C)$ and each point P_i is repeated in the list as many times as its multiplicity. Assume*

that $\tilde{L}$ is not a component of $\tilde{C}$. Then $\tilde{C} \cap \tilde{L} = \{\tilde{P}_1, \tilde{P}_2, \ldots, \tilde{P}_d\}$ with the correct multiplicities.

PROOF. Suppose first that L is the line at infinity $Z = 0$. Let $F(X, Y, Z) = 0$ be a normalized equation for C. The assumption that $\tilde{L}$ is not a component of $\tilde{C}$ means that $\tilde{F}(X, Y, 0) \neq 0$, i.e., that some coefficient of $F(X, Y, 0)$ is not divisible by p. For each intersection point P_i, let $P_i = [l_i, m_i, 0]$ in normalized coordinates. Then

$$F(X, Y, 0) = c \prod_{i=1}^{d} (m_i X - l_i Y) \qquad (*)$$

for some constant c. This is true because the intersection points of a curve $F = 0$ with the line $Z = 0$ correspond, with the correct multiplicities, to the linear factors of $F(X, Y, 0)$. Since each of the linear polynomials on the right of $(*)$ is normalized and some coefficient of F is not divisible by p, we see that c must be an integer not divisible by p. Therefore we can reduce $(*)$ modulo p to obtain

$$\tilde{F}(X, Y, 0) = \tilde{c} \prod_{i=1}^{d} (\tilde{m}_i X - \tilde{l}_i Y), \qquad (\tilde{*})$$

which shows that $\tilde{C} \cap \tilde{L} = \{\tilde{P}_1, \tilde{P}_2, \ldots, \tilde{P}_d\}$ as claimed.

What if the line L is not the line $Z = 0$? Then we just make a linear change of coordinates

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \begin{pmatrix} n_{11} & n_{12} & n_{13} \\ n_{21} & n_{22} & n_{23} \\ n_{31} & n_{32} & n_{33} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

so that L is the line $Z' = 0$ in the new coordinate system.

Is that all there is to it? No, we must be careful to make sure our change of coordinates is compatible with reduction modulo p. This is not true for general linear changes with $n_{ij} \in \mathbb{Q}$. However, if we change using a matrix (n_{ij}) with integer entries and determinant 1, then the inverse matrix (m_{ij}) will have integer entries, and the reduced matrices $(\tilde{n}_{ij})$ and $(\tilde{m}_{ij})$ are inverses giving a corresponding coordinate change in characteristic p. And clearly, if we change coordinates with (n_{ij}) and reduce mod p, the result will be the same as if we first reduce mod p and then change coordinates with $(\tilde{n}_{ij})$.

Thus, to complete our proof we must show that for every rational line L in $\mathbb{P}^2$ there is an "integral" coordinate change such that in the new coordinates L is the line at infinity. To do this, we let

$$L : aX + bY + cZ = 0$$

be a normalized equation for the line L and use the following result.

Lemma. *Let (a,b,c) be a triple of integers satisfying $\gcd(a,b,c) = 1$. Then there exists a 3×3 matrix with integer coefficients, determinant 1, and bottom line (a,b,c).*

PROOF. Let $d = \gcd(b,c)$, choose integers r and s such that $rc - sb = d$, and note for later use that r and s are necessarily relatively prime. Now $\gcd(a,d) = 1$, so we can choose t and u such that $td + ua = 1$. Finally, since $\gcd(r,s) = 1$, we can choose v and w such that $vs - wr = u$. Then the matrix

$$\begin{pmatrix} t & v & w \\ 0 & r & s \\ a & b & c \end{pmatrix}$$

has the desired properties. □

Finally, we apply the proposition to show that the reduction mod p map respects the group law on a cubic curve.

Corollary. *Let C be a non-singular rational cubic curve in $\mathbb{P}^2$ and let $\mathcal{O}$ be a rational point on C, which we take as the origin for the group law on C. Suppose that $\tilde{C}$ is non-singular and take $\tilde{\mathcal{O}}$ as the origin for the group law on $\tilde{C}$. Then the reduction mod p map $P \to \tilde{P}$ is a group homomorphism $C(\mathbb{Q}) \to \tilde{C}(\mathbb{F}_p)$.*

PROOF. Let $P, Q \in C(\mathbb{Q})$, and let $R = P + Q$. This means that there are lines L_1 and L_2 and a rational point $S \in C(\mathbb{Q})$ such that, in the notation of the proposition,

$$C \cap L_1 = \{P, Q, S\} \qquad \text{and} \qquad C \cap L_2 = \{S, \mathcal{O}, R\}.$$

Putting tildes on everything, as is allowed by the proposition, we conclude that $\tilde{P} + \tilde{Q} = \tilde{R}$. □

EXERCISES

A.1. Let $\mathbb{P}^2$ be the set of homogeneous triples $[a,b,c]$ as usual, and recall that with this definition a line in $\mathbb{P}^2$ is defined to be the set of solutions of an equation of the form

$$\alpha X + \beta Y + \gamma Z = 0$$

for some numbers α, β, γ not all zero.
(a) Prove directly from this definition that any two distinct points in $\mathbb{P}^2$ are contained in a unique line.
(b) Similarly, prove that any two distinct lines in $\mathbb{P}^2$ intersect in a unique point.

A.2. Let K be a field, for example K might be the rational numbers or the real numbers or a finite field. Define a relation $\sim$ on $(n+1)$-tuples $[a_0, a_1, \ldots, a_n]$ of elements of K by the following rule:

$$[a_0, a_1, \ldots, a_n] \sim [a'_0, a'_1, \ldots, a'_n] \quad \text{if there is a non-zero } t \in K \text{ so that } a_0 = ta'_0,\ a_1 = ta'_1, \ldots, a_n = ta'_n.$$

(a) Prove that $\sim$ is an equivalence relation. That is, prove that for any $(n+1)$-tuples $\mathbf{a} = [a_0, a_1, \ldots, a_n]$, $\mathbf{b} = [b_0, b_1, \ldots, b_n]$, and $\mathbf{c} = [c_0, c_1, \ldots, c_n]$, the relation $\sim$ satisfies the following three conditions:

(i) $\mathbf{a} \sim \mathbf{a}$ (Reflexive)
(ii) $\mathbf{a} \sim \mathbf{b} \Longrightarrow \mathbf{b} \sim \mathbf{a}$ (Symmetric)
(iii) $\mathbf{a} \sim \mathbf{b}$ and $\mathbf{b} \sim \mathbf{c} \Longrightarrow \mathbf{a} \sim \mathbf{c}$ (Transitive)

(b) Which of these properties (i), (ii), (iii) fails to be true if K is replaced by a ring R that is not a field? (There are several answers to this question, depending on what the ring R looks like.)

A.3. We saw in Section 1 that the directions in the affine plane $\mathbb{A}^2$ correspond to the points of the projective line $\mathbb{P}^1$. In other words, $\mathbb{P}^1$ can be described as the set of lines in $\mathbb{A}^2$ going through the origin.

(a) Prove similarly that $\mathbb{P}^2$ can be described as the set of lines in $\mathbb{A}^3$ going through the origin.

(b) Let $\Pi \subset \mathbb{A}^3$ be a plane in $\mathbb{A}^3$ that goes through the origin, and let S_Π be the collection of lines in $\mathbb{A}^3$ going through the origin and contained in Π. From (a), S_Π defines a subset L_Π of $\mathbb{P}^2$. Prove that L_Π is a line in $\mathbb{P}^2$, and conversely that every line in $\mathbb{P}^2$ can be constructed in this way.

(c) Generalize (a) by showing that $\mathbb{P}^n$ can be described as the set of lines in $\mathbb{A}^{n+1}$ going through the origin.

A.4. Let $F(X, Y, Z) \in \mathbb{C}[X, Y, Z]$ be a homogeneous polynomial of degree d.

(a) Prove that the three partial derivatives of F are homogeneous polynomials of degree $d-1$.

(b) Prove that

$$X\frac{\partial F}{\partial X} + Y\frac{\partial F}{\partial Y} + Z\frac{\partial F}{\partial Z} = d \cdot F(X, Y, Z).$$

(*Hint.* Differentiate $F(tX, tY, tZ) = t^d F(X, Y, Z)$ with respect to t.)

A.5. Let $C : F(X, Y, Z) = 0$ be a projective curve given by a homogeneous polynomial $F \in \mathbb{C}[X, Y, Z]$, and let $P \in \mathbb{P}^2$ be a point.

(a) Prove that P is a singular point of C if and only if

$$\frac{\partial F}{\partial X}(P) = \frac{\partial F}{\partial Y}(P) = \frac{\partial F}{\partial Z}(P) = 0.$$

(b) If P is a non-singular point of C, prove that the tangent line to C at P is given by the equation

$$\frac{\partial F}{\partial X}(P)X + \frac{\partial F}{\partial Y}(P)Y + \frac{\partial F}{\partial Z}(P)Z = 0.$$

A.6. Let C be the projective curve given by the equation

$$C : Y^2Z - X^3 - Z^3 = 0.$$

(a) Show that C has only one point at infinity, namely the point $[0,1,0]$ corresponding to the vertical direction $x = 0$.
(b) Let $C_0 : y^2 - x^3 - 1 = 0$ be the affine part of C, and let (r_i, s_i) be a sequence of point on C_0 with $r_i \to \infty$. Let L_i be the tangent line to C_0 at the point (r_i, s_i). Prove that as $i \to \infty$, the slopes of the lines L_i approach infinity, i.e., they approach the slope of the line $x = 0$.

A.7. Let $f(x,y)$ be a polynomial.
(a) Expand $f(tx,ty)$ as a polynomial in t whose coefficients are polynomials in x and y. Prove that the degree of $f(tx,ty)$, considered as a polynomial in the variable t, is equal to the degree of the polynomial $f(x,y)$.
(b) Prove that the homogenization $F(X,Y,Z)$ of $f(x,y)$ is given by

$$F(X,Y,Z) = Z^d f\left(\frac{X}{Z}, \frac{Y}{Z}\right), \qquad \text{where } d = \deg(f).$$

A.8. For each of the given affine curves C_0, find a projective curve C whose affine part is C_0. Then find all of the points at infinity on the projective curve C.
(a) $C_0 : 3x - 7y + 5 = 0.$
(b) $C_0 : x^2 + xy - 2y^2 + x - 5y + 7 = 0.$
(c) $C_0 : x^3 + x^2y - 3xy^2 - 3y^3 + 2x^2 - x + 5 = 0.$

A.9. For each of the following curves C and points P, either find the tangent line to C at P or else verify that C is singular at P.

(a)	$C : y^2 = x^3 - x,$	$P = (1,0).$
(b)	$C : X^2 + Y^2 = Z^2,$	$P = [3,4,5].$
(c)	$C : x^2 + y^4 + 2xy + 2x + 2y + 1 = 0,$	$P = (-1,0).$
(d)	$C : X^3 + Y^3 + Z^3 = XYZ,$	$P = [1,-1,0].$

A.10. (a) Prove that a projective transformation of $\mathbb{P}^2$ sends lines to lines.
(b) More generally, prove that a projective transformation of $\mathbb{P}^2$ sends curves of degree d to curves of degree d.

A.11. Let P, P_1, P_2, P_3 be points in $\mathbb{P}^2$ and let L be a line in $\mathbb{P}^2$.
(a) If P_1, P_2, and P_3 do not lie on a line, prove that there is a projective transformation of $\mathbb{P}^2$ so that

$$P_1 \longmapsto [0,0,1], \qquad P_2 \longmapsto [0,1,0], \qquad P_3 \longmapsto [1,0,0].$$

(b) If no three of P_1, P_2, P_3 and P lie on a line, prove that there is a unique projective transformation as in (a) which also sends P to $[1,1,1]$.
(c) Prove that there is a projective transformation of $\mathbb{P}^2$ so that L is sent to the line $Z' = 0$.
(d) More generally, if P does not lie on L, prove that there is a projective transformation of $\mathbb{P}^2$ so that L is sent to the line $Z' = 0$ and P is sent to the point $[0,0,1]$.

A.12. For each of the pairs of curves C_1, C_2, find all of the points in the intersection $C_1 \cap C_2$. Be sure to include points with complex coordinates and points at infinity.

(a) $C_1 : x - y = 0$, $C_2 : x^2 - y = 0$.
(b) $C_1 : x - y - 1 = 0$, $C_2 : x^2 - y^2 + 2 = 0$.
(c) $C_1 : x - y - 1 = 0$, $C_2 : x^2 - 2y^2 - 5 = 0$.
(d) $C_1 : x - 2 = 0$, $C_2 : y^2 - x^3 + 2x = 0$.

A.13. For each of the pairs of curves C_1, C_2, compute the intersection index $I(C_1 \cap C_2, P)$ at the indicated point P. Also sketch the curves and the point in $\mathbb{R}^2$.

(a) $C_1 : x - y = 0$, $C_2 : x^2 - y = 0$, $P = (0,0)$.
(b) $C_1 : y = 0$, $C_2 : x^2 - y = 0$, $P = (0,0)$.
(c) $C_1 : x - y = 0$, $C_2 : x^3 - y^2 = 0$, $P = (0,0)$.
(d) $C_1 : x^2 - y = 0$, $C_2 : x^3 - y = 0$, $P = (0,0)$.
(e) $C_1 : x + y = 2$, $C_2 : x^2 + y^2 = 2$, $P = (1,1)$.

A.14. Let $\mathcal{C}^{(d)}$ be the collection of curves of degree d in $\mathbb{P}^2$.

(a) Show that $\mathcal{C}^{(d)}$ is naturally isomorphic to the projective space $\mathbb{P}^N$ for a certain value of N, and find N explicitly in terms of d.

(b) In Section 3 we gave a plausibility arguement for why the Cayley-Bacharach theorem is true for curves of degree 3. Give a similar argument for general curves C_1, C_2, and D of degrees d_1, d_2, and $d_1 + d_2 - 3$ respectively.

A.15. Let $P \in \mathbb{A}^2$. In this exercise we ask you to verify various properties of $\mathcal{O}_P$, the local ring of P, as defined in Section 4.

(a) Prove that $\mathcal{O}_P$ is a subring of $K = k(x,y)$.

(b) Prove that the map $\phi \mapsto \phi(P)$ is a homomorphism of $\mathcal{O}_P$ onto k. Let M_P be the kernel of this homomorphism.

(c) Prove that $\mathcal{O}_P$ equals the direct sum $k + M_P$.

(d) Prove that $\phi \in \mathcal{O}_P$ is a unit if and only if $\phi \notin M_P$.

(e) Let $I \subset \mathcal{O}_P$ be an ideal of $\mathcal{O}_P$. Prove that either $I = \mathcal{O}_P$, or else $I \subset M_P$. Deduce that M_P is the unique maximal ideal of $\mathcal{O}_P$.

A.16. Let P_1, P_2, P_3, P_4, P_5 be five distinct point in $\mathbb{P}^2$.

(a) Show that there exists a conic C (i.e., a curve of degree two) passing through the five points.

(b) Show that C is unique if and only if no four of the five points lie on a line.

(c) Show that C is irreducible if and only if no three of the five points lie on a line.

A.17. In this exercise we guide you in proving the cubic Cayley-Bacharach theorem in the case that the eight points are distinct. Let $C_1 : F_1 = 0$ and $C_2 : F_2 = 0$ be cubic curves in $\mathbb{P}^2$ without common component which have eight distinct points $P_1, P_2, \ldots, P_8$ in common. Suppose that $C_3 : F_3 = 0$ is a third cubic curve passing through these same eight points. Prove that C_3 is on the "line of cubics" joining C_1

and C_2, i.e., prove that there are constants λ_1 and λ_2 such that

$$F_3 = \lambda_1 F_1 + \lambda_2 F_2.$$

In order to prove this result, assume that no such λ_1, λ_2 exist and derive a contradiction as follows:
(i) Show that F_1, F_2, and F_3 are linearly independent.
(ii) Let P' and P'' be any two points in $\mathbb{P}^2$ different from each other and different from the P_i. Show that there is a cubic curve C passing through all ten points $P_1, \ldots, P_8, P', P''$. (*Hint.* Show that there exist constants $\lambda_1, \lambda_2, \lambda_3$ such that $F = \lambda_1 F_1 + \lambda_2 F_2 + \lambda_3 F_3$ is not identically 0 and such that the curve $F = 0$ does the job.)
(iii) Show that no four of the eight points P_i are collinear, and no seven of them lie on a conic. (*Hint.* Use the fact that C_1 and C_2 have no common component.)
(iv) Use the previous exercise to observe that there is a unique conic Q going through any five of the eight points $P_1, \ldots, P_8$.
(v) Show that no three of the eight points P_i are collinear. (*Hint.* If three are on a line L, let Q be the unique conic going through the other five, choose P' on L and P'' not on Q. Then use (ii) to get a cubic C which has L as a component, so is of the form $C = L \cup Q'$ for some conic Q'. This contradicts the fact that Q is unique.)
(vi) To get the final contradiction, let Q be the conic through the five points $P_1, P_2, \ldots, P_5$. By (iii), at least one (in fact two) of the remaining three points is not on Q. Call it P_6, and let L be the line joining P_7 and P_8. Choose P' and P'' on L so that again the cubic C through the ten points has L as a component. Show that this gives a contradiction.

A.18. Show that if C_1 and C_2 are both singular at the point P, then their intersection index satisfies $I(C_1 \cap C_2, P) \geq 3$.

A.19. Consider the affine curve $C : y^4 - xy - x^3 = 0$. Show that at the origin $(x, y) = (0, 0)$, C meets the y axis four times, C meets the x axis three times, and C meets every other line through the origin twice.

A.20. Show that the separation of real conics into hyperbolas, parabolas, and ellipses is an affine business and has no meaning projectively, by giving an example of a quadratic homogeneous polynomial $F(X, Y, Z)$ with real coefficients such that

$F(x, y, 1) = 0$ is a hyperbola in the real (x, y) plane,
$F(x, 1, z) = 0$ is a parabola in the real (x, z) plane,
$F(1, y, z) = 0$ is an ellipse in the real (y, z) plane.

Bibliography

Artin [1]
Artin, M., *Algebra,* Prentice Hall, Englewood Cliffs, N.J., 1991.

Baker [1]
Baker, A., Contributions to the theory of Diophantine equations (II). The Diophantine equation $y^2 = x^3 + k$. *Philos. Trans. Roy. Soc. London* **263** (1967/68), 193–208.

Baker [2]
———, *Transcendental Number Theory,* Cambridge University Press, Cambridge, 1975.

Billing-Mahler [1]
Billing, G., Mahler, K., On exceptional points on cubic curves. *J. London Math. Soc.* **15** (1940), 32–43.

Bremner-Cassels [1]
Bremner, A., Cassels, J.W.S., On the equation $Y^2 = X(X^2+p)$. *Math. Comp.* **42** (1984), 257–264.

Brieskorn-Knörrer [1]
Brieskorn, E., Knörrer, H., *Plane Algebraic Curves,* transl. by J. Stillwell, Birkhäuser, Basel, 1986.

Chahal [1]
Chahal, J.S., *Topics in Number Theory,* Plenum Press, New York-London, 1988.

Deligne [1]
Deligne, P., La conjecture de Weil I. *Publ. Math. IHES* **43** (1974), 273–307.

Faltings [1]
Faltings, G., Diophantine approximation on abelian varieties. *Annals of Math.* **133** (1991), 549–576.

Fueter [1]
Fueter, R., Über kubische diophantische Gleichungen. *Comm. Math. Helv.* **2** (1930), 69–89.

Fulton [1]
Fulton, W., *Algebraic Curves,* Benjamin, 1969.

Griffiths and Harris [1]
Griffiths, P., Harris, J., *Principles of Algebraic Geometry,* John Wiley & Sons, New York, 1978.

Hartshorne [1]
Hartshorne, R., *Algebraic Geometry,* Springer-Verlag, New York, 1977.

Hasse [1]
Hasse, H., Beweis des Analogons der Riemannschen Vermutung für die Artinschen und F. K. Schmidtschen Kongruenzzetafunktionen in gewissen elliptischen Fällen. *Nachr. Ges. Wiss. Göttingen, Math.-Phys. K.* (1933), 253–262.

Heath-Brown and Patterson [1]
Heath-Brown, D.R., Patterson, S.J., The distribution of Kummer sums at prime arguments. *J. Reine Angew. Math.* **310** (1979), 111–130.

Herstein [1]
Herstein, I.N., *Topics in Algebra,* Xerox College Publishing, Lexington, MA, 1975.

Husemöller [1]
Husemöller, D., *Elliptic Curves,* Springer-Verlag, New York, 1987.

Jacobson [1]
Jacobson, N., *Basic Algebra I,II,* W.H. Freeman & Co., New York, 1985.

Koblitz [1]
Koblitz, N., *Introduction to Elliptic Curves and Modular Forms,* Springer-Verlag, New York, 1984.

Koblitz [2]
———, *A Course in Number Theory and Cryptography,* Springer-Verlag, New York, 1987.

Kummer [1]
Kummer, E.E., De residuis cubicis disquisitiones nonnullae analyticae. *J. Reine Angew. Math.* **32** (1846), 341–359.

Lang [1]
Lang, S., *Elliptic Functions,* Addison-Wesley, Reading, MA, 1973.

Lang [2]
———, *Elliptic Curves: Diophantine Analysis,* Springer-Verlag, Berlin, 1978.

Lang [3]
———, *Fundamentals of Diophantine Geometry,* Springer-Verlag, New York, 1983.

Lenstra [1]
Lenstra, H.W., Factoring integers with elliptic curves. *Annals of Math.* **126** (1987), 649–673.

Luck-Moussa-Waldschmidt [1]
Luck, J.M., Moussa, P., Waldschmidt, M., eds., *Number Theory and Physics, Proc. in Physics* **47**, Springer-Verlag, Berlin, 1990.

Lutz [1]
Lutz, E., Sur l'equation $y^2 = x^3 - Ax - B$ dans les corps p-adic. *J. Reine Angew. Math.* **177** (1937), 237–247.

Mazur [1]
Mazur, B., Modular curves and the Eisenstein ideal. *IHES Publ. Math.* **47** (1977), 33–186.

Mazur [2]
———, Rational isogenies of prime degree. *Invent. Math.* **44** (1978), 129–162.

Mestre [1]
Mestre, J.-F., Formules explicites et minorations de conducteurs de variétiés algebriques. *Compos. Math.* **58** (1986), 209-232.

Mestre [2]
———, Private communication, January, 1992.

Mordell [1]
Mordell, L.J., On the rational solutions of the indeterminate equations of the third and fourth degrees. *Proc. Camb. Philos. Soc.* **21** (1922), 179–192.

Nagell [1]
Nagell, T., Solution de quelque problémes dans la théorie arithmétique des cubiques planes du premier genre. *Wid. Akad. Skrifter Oslo I* (1935), Nr. 1.

Pollard [1]
Pollard, J.M., Theorems on factorization and primality testing. *Proc. Camb. Philos. Soc.* **76** (1974), 521–528.

Reid [1]
Reid, M., *Undergraduate Algebraic Geometry,* London Math. Soc. Student Texts 12, Cambridge University Press, Cambridge, 1988.

Robert [1]
Robert, A., *Elliptic Curves,* Lecture Notes in Math. 326, Springer-Verlag, Berlin, 1973.

Schmidt [1]
Schmidt, W., Simultaneous approximation to algebraic numbers by rationals. *Acta Math.* **125** (1970), 189–201.

Schmidt [2]
———, *Diophantine Approximation,* Lecture Notes in Math. 785, Springer-Verlag, Berlin, 1980.

Selmer [1]
Selmer, E., The diophantine equation $ax^3 + by^3 + cz^3 = 0$. *Acta Math.* **85** (1951), 203–362.

Serre [1]
Serre, J.-P., *A Course in Arithmetic,* Springer-Verlag, New York, 1973.

Serre [2]
———, *Linear Representations of Finite Groups,* Springer-Verlag, New York, 1973.

Serre [3]
———, *Abelian ℓ-adic Representations,* Benjamin, New York, 1968.

Serre [4]
———, Propriétés galoisiennes des points d'ordre fini des courbes elliptiques. *Invent. Math.* **15** (1972), 259–331.

Siegel [1]
Siegel, C.L., The integer solutions of the equation $y^2 = ax^n + bx^{n-1} + \cdots + k$. *J. London Math. Soc.* **1** (1926), 66–68.

Siegel [2]
———, Über einige Anwendungen diophantischer Approximationen (1929). In *Collected Works,* Springer-Verlag, Berlin, 1966, 209–266.

Silverman [1]
Silverman, J.H., Integer points and the rank of Thue elliptic curves. *Invent. Math.* **66** (1982), 395–404.

Silverman [2]
———, *The Arithmetic of Elliptic Curves,* Springer-Verlag, New York, 1986.

Skolem [1]
Skolem, Th., *Diophantische Gleichungen,* Springer-Verlag, Berlin, 1938.

Tate [1]
Tate, J., The arithmetic of elliptic curves. *Invent. Math.* **23** (1974), 171–206.

Thue [1]
Thue, A., Über Annäherungswerte Algebraischer Zahlen. *J. Reine Angew. Math.* **135** (1909), 284–305.

Vojta [1]
Vojta, P., Siegel's theorem in the compact case. *Annals of Math.* **133** (1991), 509–548.

Walker [1]
Walker, R.J., *Algebraic Curves,* Dover, 1962.

Weil [1]
Weil, A., *Sur les Courbes Algébriques et les Variétés qui s'en Déduisent,* Hermann, Paris, 1948.

Weil [2]
———, Number of solutions of equations over finite fields. *Bull. Amer. Math. Soc.* **55** (1949), 497–508.

List of Notation

$P * Q$	composition law on a cubic curve, 15
$P + Q$	addition law on a cubic curve, 18
$\mathcal{O}$	point at infinity on a cubic curve, 28
$C(\mathbb{Q})$	rational points on a cubic curve, 42
$C(\mathbb{R})$	real points on a cubic curve, 42
$C(\mathbb{C})$	complex points on a cubic curve, 42
g_2, g_3	Weierstrass coefficients for a lattice, 43
$\wp$	Weierstrass $\wp$ function, 43
D	discriminant of $f(x)$, 47
$\mathbb{Z}[x]$	polynomials with integer coefficients, 48
ord	order of a rational number, 49
$C(p^\nu)$	rational points with p^ν in the denominator, 50
R, R_p	rational numbers with no p in the denominator, 51
A_m	points of order m in an abelian group, 58
S, S_p	rational numbers with denominator a p power, 61
$H(x)$	height of a rational number, 63
$H(P)$	height of a point on a cubic curve, 64
$h(P)$	logarithm of $H(P)$, 64
$2C(\mathbb{Q})$	points which are twice some point, 65
mG	image of multiplication-by-m map, 65
Γ	$= C(\mathbb{Q})$, group of rational points on a cubic, 76
T	$= (0,0)$, a point of order two, 76
$\overline{C}$	a curve that C maps to, 77
ϕ	homomorphism from C to $\overline{C}$, 77
$\phi(\Gamma)$	image of the group of rational points, 83
$\mathbb{Q}^*$	multiplicative group of rational numbers, 85
$\mathbb{Q}^{*2}$	squares of non-zero rational numbers, 85
α	map from Γ to $\mathbb{Q}^*/\mathbb{Q}^{*2}$, 85
$\mathbb{Z}$	additive group of integers, 89
$\mathbb{Z}_m$	cyclic group of order m, 89
$\Gamma[2]$	elements of order two in Γ, 90
C_p	the curve $y^2 = x^3 + px$, 97
C_{ns}	non-singular points on a cubic curve, 99
$C_{\text{ns}}(\mathbb{Q})$	non-singular rational points on a cubic curve, 99
$\mathbb{F}_p$	field of integers modulo p, 107
$\mathbb{F}_q$	field with q elements, 107
$C(\mathbb{F}_p)$	set of points on C rational over $\mathbb{F}_p$, 108
M_p	number of points on a curve over $\mathbb{F}_p$, 111
$\mathbb{F}_p^*$	multiplicative group of a finite field, 111
R	group of cubic residues in $\mathbb{F}_p$, 112
S, T	cosets of R in $\mathbb{F}_p^*$, 112
$[X, Y, Z]$, $[XYZ]$	set of triples adding to zero, 112

$\alpha_1, \alpha_2, \alpha_3$	cubic Gauss sums, 114
θ_p	angle related to number of points modulo p, 120
$\pi(X)$	number of primes less than X, 120
$z \to \tilde{z}$	reduction modulo p, 121
$\tilde{C}$	reduction of C modulo p, 121
$\tilde{P}$	reduction of a point modulo p, 121
Φ	points of finite order in $C(\mathbb{Q})$, 122
$\mathcal{H}$	upper half plane, 141
$\| \ \|$	size of a vector, 158
$F^{(k)}(X,Y)$	k^{th} X-derivative, divided by $k!$, 161
$F(X,Y)$	auxiliary polynomial, 163
$P(X)+Q(X)Y$	auxiliary polynomial, 163
$W(X)$	Wronskian polynomial, 168
$\tau(d)$	exponent for Diophantine approximation, 174
$[K:F]$	the degree of K over F, 180
$\mathrm{Aut}(K)$	group of automorphisms $\sigma : K \to K$, 180
$\mathrm{Gal}(K/\mathbb{Q})$	Galois group of $K/\mathbb{Q}$, 181
$\mathrm{Aut}_F(K)$	group of automorphisms of K fixing F, 183
$\mathrm{Gal}(K/F)$	Galois group of K/F, 183
$C(K)$	K rational points on C, 185
$\sigma(P)$	action of Galois group on a point, 186
λ_n	n^{th}-power homomorphism on $\mathbb{C}^*$, 188
λ_n	multiplication-by-n map on C, 188
$C[n]$	kernel of multiplication-by-n map, 188
$\mathbb{Q}(C[n])$	field of definition of $C[n]$ over $\mathbb{Q}$, 193
$F(C[n])$	field of definition of $C[n]$ over F, 193
$\mathrm{GL}_r(R)$	general linear group, 195
ρ_n	Galois representation on points of order n, 195
t	cyclotomic representation, 196
i	$=\sqrt{-1}$, 197
ψ_n	division polynomial, 214
ϕ_n	multiplication-by-n polynomial, 214
ω_n	multiplication-by-n polynomial, 214
R_L	endomorphism ring of the lattice L, 217
$[a,b,c]$	homogeneous triple, a point in $\mathbb{P}^2$, 221
$\mathbb{P}^2$	the projective plane, 221
$\sim$	equivalence relation on triples $[a,b,c]$, 221
$\mathbb{P}^n$	projective n-space, 221
$[a_0, \ldots, a_n]$	homogeneous $n+1$-tuple, a point in $\mathbb{P}^n$, 221
$\sim$	equivalence relation on $n+1$-tuples, 221
$\mathbb{A}^2$	Euclidean (or affine) plane, 222
L_∞	line at infinity in $\mathbb{P}^2$, 223
$C(\mathbb{Q})$	set of rational points on a projective curve C, 229
$C_0(\mathbb{Q})$	set of rational points on an affine curve C_0, 230
$C_0(\mathbb{Z})$	set of integer points on an affine curve C_0, 230
$I(C_1 \cap C_2, P)$	intersection index of C_1 and C_2 at P, 237

$\mathcal{C}^{(3)}$	the collection of cubic curves in $\mathbb{P}^2$, 238
dim	dimension of a vector space, 242
$\mathcal{O}_P$	local ring of P, 245
$\phi(P)$	value of a rational function defined at P, 245
M_P	the maximal ideal of the local ring of P, 245
$(f_1, f_2)_P$	ideal in $\mathcal{O}_P$ generated by f_1 and f_2, 246
$I(C_1 \cap C_2, P)$	intersection index of C_1 and C_2 at P, 246
$\mathcal{O}_P$	homogeneous local ring of P, 249
M_P	maximal ideal of homogeneous local ring, 249
$(F_1, F_2)_P$	homogeneous ideal associated to F_1 and F_2, 249
$I(f_1, f_2)$	intersection index of two curves at $(0,0)$, 250
$\tilde{C}$	reduction modulo p of a curve C, 252
$\mathcal{C}^{(d)}$	the collection of curves of degree d in $\mathbb{P}^2$, 257

Index

Undergraduate Texts in Mathematics

(continued from page ii)

Gamelin: Complex Analysis.
Gordon: Discrete Probability.
Hairer/Wanner: Analysis by Its History. *Readings in Mathematics.*
Halmos: Finite-Dimensional Vector Spaces. Second edition.
Halmos: Naive Set Theory.
Hämmerlin/Hoffmann: Numerical Mathematics. *Readings in Mathematics.*
Harris/Hirst/Mossinghoff: Combinatorics and Graph Theory.
Hartshorne: Geometry: Euclid and Beyond.
Hijab: Introduction to Calculus and Classical Analysis.
Hilton/Holton/Pedersen: Mathematical Reflections: In a Room with Many Mirrors.
Hilton/Holton/Pedersen: Mathematical Vistas: From a Room with Many Windows.
Iooss/Joseph: Elementary Stability and Bifurcation Theory. Second edition.
Isaac: The Pleasures of Probability. *Readings in Mathematics.*
James: Topological and Uniform Spaces.
Jänich: Linear Algebra.
Jänich: Topology.
Jänich: Vector Analysis.
Kemeny/Snell: Finite Markov Chains.
Kinsey: Topology of Surfaces.
Klambauer: Aspects of Calculus.
Lang: A First Course in Calculus. Fifth edition.
Lang: Calculus of Several Variables. Third edition.
Lang: Introduction to Linear Algebra. Second edition.
Lang: Linear Algebra. Third edition.
Lang: Short Calculus: The Original Edition of "A First Course in Calculus."
Lang: Undergraduate Algebra. Second edition.
Lang: Undergraduate Analysis.
Laubenbacher/Pengelley: Mathematical Expeditions.
Lax/Burstein/Lax: Calculus with Applications and Computing. Volume 1.
LeCuyer: College Mathematics with APL.
Lidl/Pilz: Applied Abstract Algebra. Second edition.
Logan: Applied Partial Differential Equations.
Lovász/Pelikán/Vesztergombi: Discrete Mathematics.
Macki-Strauss: Introduction to Optimal Control Theory.
Malitz: Introduction to Mathematical Logic.
Marsden/Weinstein: Calculus I, II, III. Second edition.
Martin: Counting: The Art of Enumerative Combinatorics.
Martin: The Foundations of Geometry and the Non-Euclidean Plane.
Martin: Geometric Constructions.
Martin: Transformation Geometry: An Introduction to Symmetry.
Millman/Parker: Geometry: A Metric Approach with Models. Second edition.
Moschovakis: Notes on Set Theory.
Owen: A First Course in the Mathematical Foundations of Thermodynamics.
Palka: An Introduction to Complex Function Theory.
Pedrick: A First Course in Analysis.
Peressini/Sullivan/Uhl: The Mathematics of Nonlinear Programming.
Prenowitz/Jantosciak: Join Geometries.
Priestley: Calculus: A Liberal Art. Second edition.

Protter/Morrey: A First Course in Real Analysis. Second edition.
Protter/Morrey: Intermediate Calculus. Second edition.
Pugh: Real Mathematical Analysis.
Roman: An Introduction to Coding and Information Theory.
Ross: Elementary Analysis: The Theory of Calculus.
Samuel: Projective Geometry. *Readings in Mathematics.*
Saxe: Beginning Functional Analysis
Scharlau/Opolka: From Fermat to Minkowski.
Schiff: The Laplace Transform: Theory and Applications.
Sethuraman: Rings, Fields, and Vector Spaces: An Approach to Geometric Constructability.
Sigler: Algebra.
Silverman/Tate: Rational Points on Elliptic Curves.
Simmonds: A Brief on Tensor Analysis. Second edition.
Singer: Geometry: Plane and Fancy.
Singer/Thorpe: Lecture Notes on Elementary Topology and Geometry.
Smith: Linear Algebra. Third edition.
Smith: Primer of Modern Analysis. Second edition.
Stanton/White: Constructive Combinatorics.
Stillwell: Elements of Algebra: Geometry, Numbers, Equations.
Stillwell: Elements of Number Theory.
Stillwell: Mathematics and Its History. Second edition.
Stillwell: Numbers and Geometry. *Readings in Mathematics.*
Strayer: Linear Programming and Its Applications.
Toth: Glimpses of Algebra and Geometry. Second Edition. *Readings in Mathematics.*
Troutman: Variational Calculus and Optimal Control. Second edition.
Valenza: Linear Algebra: An Introduction to Abstract Mathematics.
Whyburn/Duda: Dynamic Topology.
Wilson: Much Ado About Calculus.