Topics In Algebra
Elementary Algebraic Geometry

David Marker

Spring 2003

Contents

Let F be a field and suppose $f_1, \ldots, f_m \in F[X_1, \ldots, X_n]$. A central problems of mathematics is to study the solutions to systems of polynomial equations:

$$\begin{aligned} f_1(X_1, \ldots, X_n) &= 0 \\ f_2(X_1, \ldots, X_n) &= 0 \\ &\vdots \\ f_m(X_1, \ldots, X_n) &= 0 \end{aligned}$$

where $f_1, \ldots, f_m \in F[X_1, \ldots, X_n]$. Of particular interest are the cases when F is the field $\mathbb{Q}$ of rational numbers, $\mathbb{R}$ of real numbers, $\mathbb{C}$ of complex numbers or a finite field like $\mathbb{Z}_p$. For example Fermat's Last Theorem, is the assertion that if $x, y, z \in \mathbb{Q}$, $n > 2$ and

$$x^n + y^n = z^n,$$

then at least one of x, y, z is zero.

When we look at the solution to systems of polynomials over $\mathbb{R}$ (or $\mathbb{C}$), we can consider the geometry of the solution set in $\mathbb{R}^n$ (or $\mathbb{C}^n$). For example the solutions to

$$X^2 - Y^2 = 1$$

is a hyperbola. There are many questions we can ask about the solution space. For example:

i) The circle $X^2 + Y^2 = 1$ is smooth, while the curve $Y^2 = X^3$ has a cusp at $(0, 0)$. How can we tell if the solution set is smooth?

ii) If $f, g \in \mathbb{C}[X, Y]$ how many solutions are there to the system

$$\begin{aligned} f(X, Y) &= 0 \\ g(X, Y) &= 0? \end{aligned}$$

The main theme of the course will be that there are deep connections between the geometry of the solution sets and algebraic properties of the polynomial rings.

1 Algebraically Closed Fields

We will primarily be considering solutions to $f(X, Y) = 0$ where f is a polynomial in two variables, but we start by looking at equations $f(X) = 0$ in a single variable. In general if $f \in F[X]$ there is no reason to believe that $f(X) = 0$ has a solution in F. For example, $X^2 - 2 = 0$ has no solution in $\mathbb{Q}$ and $X^2 + 1 = 0$ has no solution in $\mathbb{R}$. The fields where every nonconstant polynomial has a solution play an important role.

Definition 1.1 We say that a field F is *algebraically closed* if every nonconstant polynomial has a zero in F.

The Complex Numbers

Theorem 1.2 (Fundamental Theorem of Algebra) *The field $\mathbb{C}$ of complex numbers is algebraically closed.*

Although this is a purely algebraic statement. Most proofs of the Fundamental Theorem of Algebra use ideas from other areas of mathematics, such as Complex Analysis or Algebraic Topology. We will sketch one proof that relies on a central theorem of Complex Analysis.

Recall that if $z \in \mathbb{C}$ and $z = a + bi$ where $a, b \in \mathbb{R}$, we let $|z| = \sqrt{a^2 + b^2}$.

Theorem 1.3 (Liouville's Theorem) *Suppose $g : \mathbb{C} \to \mathbb{C}$ is a differentiable function and there is an M such that $|g(z)| < M$ for all $z \in \mathbb{C}$, then g is constant.*

Proof of Fundamental Theorem of Algebra

Let $f \in \mathbb{C}[X]$. Suppose $f(z) \neq 0$ for all $z \in \mathbb{C}$. We must show f is constant.

Let $g(z) = \frac{1}{f(z)}$. Then $g : \mathbb{C} \to \mathbb{C}$ is differentiable. Suppose f has degree n and

$$f(X) = \sum_{i=0}^{n} a_i X^i = a_n X^n \left(1 + \frac{a_{n-1}}{a_n}\frac{1}{X} + \ldots + \frac{a_0}{a_n}\frac{1}{X^n}\right).$$

Then $|f(z)| \to \infty$ as $|z| \to \infty$ and $|g(z)| \to 0$ as $|z| \to \infty$. Thus we can find r such that $|g(z)| < 1$ for $|z| > r$. The set $\{z : z \leq r\}$ is compact. Thus there is $M > 0$ such that $|g(z)| \leq M$ if $|z| \leq r$. Thus $|g|$ is bounded on $\mathbb{C}$. By Liouville's Theorem, g is constant and hence f is constant.

Existence of Algebraically Closed Fields

While the complex numbers is the most natural algebraically closed field, there are other examples. Indeed every field has an algebraically closed extension field. The key idea is that even though a nonconstant polynomial does not have a zero in a field F it will have one in an extension of F.

Theorem 1.4 (Fundamental Theorem of Field Theory) *If F is a field and $f \in F[X]$ is a nonconstant polynomial, there is an extension field $K \supseteq F$ containing a zero of F.*

Sketch of Proof Let $p \in F[X]$ be an irreducible factor of f. It suffices to find a zero of p. Since p is irreducible, $\langle p \rangle$ is a maximal ideal and $K = F[X]/\langle p \rangle$ is a field. By identifying $a \in F$ with $a + \langle p \rangle$, we can view F as a subfield of K. The element $\alpha = X + \langle p \rangle$ is a zero of p.

While this might seem artificial, this construction is really quite natural. Indeed if p is an irreducible polynomial of degree n and α is a zero of p in K, then

$$F(\alpha) = \left\{\sum_{i=0}^{n-1} a_i \alpha^i : a_0, \ldots, a_{n-1} \in F\right\}$$

is an extension field isomorphic to $F[X]/\langle p\rangle$.

Any proof that every field has an algebraically closed extension needs a little set theory. We will simplify the set theory involved by only considering the countable case.

Recall that a set A is *countable* if there is an onto function $f : \mathbb{N} \to A$. In this case $f(0), f(1), \ldots$ is a listing of A (possibly with repetitions).

The next lemma summarizes all we will need about countability.

Lemma 1.5 *i) If A is countable and $f : A \to B$ is onto, then B is countable.*
ii) $\mathbb{N} \times \mathbb{N}$ is countable.
iii) If A is a countable set, then A^n is countable.
iv) If $A_0, A_1, \ldots$ are countable then $\bigcup_{n=0}^{n} A_n$ is countable.

Proof

i) If $g : \mathbb{N} \to A$ is onto and $f : A \to B$ is onto, then $f \circ g$ is onto.

ii) Define $\phi : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ as follows if $x \in \mathbb{N}$ we can factor $x = 2^n 3^m y$ where neither 2 nor 3 divides y. Let $\phi(x) = (n, m)$. Then ϕ is onto.

iii) We prove this by induction on n. It is clearly true for $n = 1$. Suppose A^n is countable. There are onto functions $f : \mathbb{N} \to A$ and $g : A \to A^n$. Let $h : \mathbb{N} \times \mathbb{N} \to A^{n+1}$ be the function

$$h(n, m) = (f(n), g(n)).$$

Then h is onto. By i) and ii) A^{n+1} is countable.

iv) Suppose $A_0, A_1, \ldots$ are all countable sets. Let $f_n : \mathbb{N} \to A_n$ be onto. Let

$$g : \mathbb{N} \times \mathbb{N} \to \bigcup_{n=0}^{\infty} A_n$$

be the function $g(n, m) = f_n(m)$. Then g is onto and $\bigcup_{n=0}^{\infty} A_n$ is onto.

Corollary 1.6 *i) If F is a countable field and $f \in F[X]$ is nonconstant, we can find a countable field $K \supseteq F$ containing a zero of f.*
ii) If F is a countable field, then $F[X]$ is countable.

Proof i) Let p be an irreducible factor of F. Let α be a zero of p in an extension field. If p has degree n, then

$$F(\alpha) = \left\{ \sum_{i=0}^{n-1} a_i \alpha^i : a_0, \ldots, a_{n-1} \in F \right\}$$

and the map

$$(a_0, \ldots, a_{n-1}) \mapsto \sum_{i=0}^{n-1} a_i \alpha^i$$

is a function from F^n onto $F(\alpha)$.

ii) Let P_n be the polynomials in $F[X]$ of degree at most n. The map

$$(a_0, \ldots, a_n) \mapsto \sum_{i=0}^{n} a_i X^i$$

is an function from F^n onto P_n. Thus P_n is countable and $F[X] = \bigcup_{n=0}^{\infty} P_n$ is countable.

We need one more basic lemma.

Lemma 1.7 *Suppose $F_0 \subseteq F_1 \subseteq F_2 \subseteq \ldots$ are fields. Then $F = \bigcup_{n=0}^{\infty} F_n$ is a field.*

Sketch of Proof We first note that F is closed under addition and multiplication. If $a, b \in F$ we can find n_0, n_1 such that $a \in F_{n_0}$ and $b \in F_{n_1}$. Let $n = \max(n_0, n_1)$. Then $a, b \in F_n$ and $a + b, ab \in F_n \subseteq F$. Similarly if $a \in F$ and $a \neq 0$, there is an n such that $a \in F_n$ and $\frac{1}{a} \in F_n$.

It is easy to check that all of the field axioms hold. For example, if $a, b, c \in F$, there is an n such that $a, b, c \in F_n$. Since F_n is a field $a + (b + c) = (a + b) + c$. All of the field axioms have analogous proofs.

Lemma 1.8 *If F is a countable field, there is a countable field $K \supseteq F$ such that if $f \in F[X]$ is a nonconstant polynomial, there is $\alpha \in K$ such that $f(\alpha) = 0$.*

Proof Since $F[X]$ is countable, we can find $f_0, f_1, \ldots$ an enumeration of $F[X]$. We build a sequence of countable fields

$$F_0 \subseteq F_1 \subseteq F_2 \subseteq \ldots$$

as follows. Let $F_0 = F$. Given F_i if f_i is a constant polynomial let $F_{i+1} = F_i$, otherwise let $F_{i+1} \supseteq F_i$ be a countable extension field containing a zero of f_i. This is possible by Corollary 1.6. Let $K = \bigcup_{i=0}^{\infty} F_i$. Then K is a countable field extending F. If $f_i \in F[X]$, then f_i has a zero in $F_i \subseteq K_i$. Thus f_i has a zero in K.

Theorem 1.9 *If F is a field, then there is an algebraically closed $K \supseteq F$.*

Proof We will prove this only in case F is countable. We build fields

$$K_0 \subseteq K_1 \subseteq \dots$$

as follows. Let $K_0 = F$. Given K_n a countable field, we can find a countable field $K_{n+1} \supseteq K_n$ such that every nonconstant polynomial $f \in K_n[X]$ has a zero in K_{n+1}. Let $K = \bigcup_{n=0}^{\infty} K^n$. Suppose $f \in K[X]$. Let

$$f = \sum_{i=0}^{n} a_i X^i.$$

For each i we can find m_i such that $a_i \in K_{m_i}$. Let $m = \max(m_0, \dots, m_n)$. Then $f \in K_m[X]$. If f is nonconstant, f has a zero in $K_{m+1}[X]$.

Thus every nonconstant polynomial in $K[X]$ has a zero in K.

Solving Equations in Algebraically Closed Fields

Lemma 1.10 *If F is a field, $f \in F[X]$, $a \in F$ and $f(a) = 0$, then we can factor $f = (X - a)g$ for some $g \in F[X]$.*

Proof By the Division Algorithm there are g and $r \in F[X]$ such that

$$f = g(X - a) + r$$

and either $r = 0$ or deg $r < 1$. In either case, we see that $r \in F$. But

$$0 = f(a) = g(a)(a - a) + r = r.$$

Thus $f = (X - a)g$.

Corollary 1.11 *Suppose $f \in F[X]$ is nonconstant, then the number of zeros of F is at most* deg f.

Proof We prove this by induction on deg f. If deg $f = 1$, then $f(X) = aX + b$ for some $a, b \in F$ with $a \neq 0$ and the only solution is $-\frac{b}{a}$. Suppose deg $f > 1$.

case 1: f has no zeros in F.

In this case the number of zeros is less than deg f, as desired.

case 2: f has a zero $a \in F$.

By the previous lemma, there is $g \in F[X]$ such that $f = (X - a)g$ and deg $g =$ deg $f - 1$. If $f(x) = 0$ then either $x = a$ or $g(x) = 0$. By induction, g has at most deg $f - 1$ zeros. Thus f has at most deg f zeros.

In algebraically closed fields we can get more precise information. Suppose $f \in F[X]$ has degree n. We say that f *splits* over F if

$$f = b(X - a_1)(X - a_2) \cdots (X - a_n)$$

for some $a_1, \dots, a_n, b \in F$.

Proposition 1.12 *If K is an algebraically closed field, then every $f \in K[X]$ splits over K.*

Proof We prove this by induction on the degree of f. If deg $f \leq 1$, this is clear. Suppose deg $f > 1$. There is $a \in K$ such that $f(a) = 0$. Thus $f = (X - a)g$ for some $g \in F[X]$ with deg $g <$ deg f. By induction, we can factor

$$g = b(X - c_1) \cdots (X - c_{\text{deg } g}).$$

Hence

$$f = b(X - a)(X - c_1) \cdots (X - c_{\text{deg } g}).$$

If

$$f = b(X - a_1)(X - a_2) \cdots (X - a_n),$$

then the zeros of f are $\{a_1, \ldots, a_n\}$. There is no reason to believe that $a_1, \ldots, a_n$ are distinct as f might have repeated zeros.

If K is an algebraically closed field, $f \in K[X]$ and $a_1, \ldots, a_n$ are the distinct zeros of f, then we can factor

$$f = b(X - a_1)^{m_1} \cdots (X - a_n)^{m_n}.$$

Since $K[X]$ is a unique factorization domain, this factorization is unique, up to renumbering the a_i.

Definition 1.13 If F is a field, $f \in F[X]$ is a nonconstant polynomial, $a \in F$, we say that a is a *multiple zero* of f if $(X - a)^2$ divides f in $F[X]$.

We say that a has *multiplicity* m if we can factor $f = (X - a)^m g$ where $g(a) \neq 0$.

If

$$f = b(X - a_1)^{m_1} \cdots (X - a_n)^{m_n},$$

where $a_1, \ldots, a_m$ are distinct, then a_i has multiplicity m_i. The following Proposition is useful, but quite easy.

Proposition 1.14 *If K is an algebraically closed field, $f \in K[X]$ is a nonconstant polynomial, $a_1, \ldots, a_n$ are the distinct zeros of f and a_i has multiplicity m_i. Then $m_1 + \ldots + m_n =$ deg f.*

In other words, "counted correctly" f always has deg f zeros in K.

There is an easy test to see if a is a multiple zero of f.

Lemma 1.15 *Let F be a field, $a \in F$, $f \in F[X]$ nonconstant, then a is a multiple zero of f if and only if $f(a) = f'(a) = 0$.*

Proof
($\Rightarrow$) If $f = (X - a)^2 g$, then

$$f' = (X - a)^2 g' + 2(X - a)g$$

and $f'(a) = 0$.

($\Leftarrow$) Suppose $f(a) = 0$. Then $f = (X - a)g$ for some $g \in F[X]$. Then $f' = (X - a)g' + g$. If $f'(a) = 0$, then $g(a) = 0$. Thus $g = (X - a)h$ for some $h \in F[X]$, $f = (X - a)^2 h$, and a is a multiple zero of f.

Although a polynomial $f \in F[X]$ may have no zeros in F, the above idea also allows us to test if f has multiple zeros in an extension of F. We need one lemma about polynomial rings. This lemma is the analog that in $\mathbb{Z}$ we can find greatest common divisors and $\gcd(n, m) = ns + mt$ for some $s, t \in \mathbb{Z}$. The proof is essentially the same.

Lemma 1.16 *Suppose F is a field and $f, g \in F[X]$ are nonzero. There is a nonzero $h \in F[X]$ such that:*
i) h divides f and g;
ii) if $k \in F[X]$ divides f and g then, k divides h;
iii) there are $s, t \in F[X]$ such that $h = fs + gt$.

Proof Consider $A = \{fs + gt : s, t \in F[X]\}$. Let $h \in S$ be a nonzero polynomial of minimal degree. Using the Division Algorithm we can find $q, r \in F[X]$ such that $f = qh + r$ and either $r = 0$ or deg $r <$ deg h. If $h = fs + gt$, then

$$r = f(1 - qs) - gqt \in A.$$

By choice of h, we must have $r = 0$ and $f = qh$. Thus h divides f. An analogous argument proves that h divides g and i) holds. Clearly iii) holds.

To show ii), suppose k divides f and h. Let $f = uk$ and $g = vk$. Then

$$h = fs + gt = uks + vkt = (us + vt)k$$

and k divides h.

Corollary 1.17 *If $f, g \in F[X]$ are nonzero polynomials with no common nonconstant factor, then there are $s, t \in F[X]$ such that $fs + gt = 1$.*

Proof Let h be as in the previous lemma. Since f and g have no common nonconstant factor, h must be a constant polynomial. If $fs + gt = h$, then $f\frac{s}{h} + g\frac{t}{h} = 1$.

Corollary 1.18 *Suppose F is a field, $f \in F[X]$ is a nonconstant polynomial, and $K \supseteq F$ is an algebraically closed field. Then f has a multiple zero in K if and only if f and f' have a common nonconstant factor in $F[X]$.*

Proof

($\Rightarrow$) If f and f' have no common nonconstant factor, then we can find $s, t \in F[X]$ such that $fs + f't = 1$. Suppose a is a multiple zero in K, then $f(a) = f'(a) = 0$. But then

$$0 = f(a)s(a) + f'(a)t(a) = 1$$

a contradiction.

($\Leftarrow$) Suppose $g \in F[X]$ is a nonconstant polynomial dividing f and f'. In K we can find a such that $g(a) = 0$. But then $f(a) = f'(a) = 0$. Hence f has a multiple zero in K.

Corollary 1.19 *If $f \in F[X]$ is irreducible and f has a multiple root in K, then $f' = 0$.*

Proof If f has a multiple zero, then f and f' have a common nonconstant factor g. Since f is irreducible, we must have $f = cg$ for some constant c and we must have deg g = deg f. Since g also divides f' and deg $f' <$ deg f, we must have $f' = 0$.

How is this possible? If

$$f = \sum_{i=0}^{d} a_i X^i,$$

then

$$f' = \sum_{i=0}^{d-1} i a_i X^{i-1}.$$

The only way we can have $f' = 0$ is if $ia_i = 0$ for all i. In characteristic zero this is impossible.

Corollary 1.20 *Suppose F is a field of characteristic zero and $f \in F[X]$ is irreducible, if $K \supseteq F$ is algebraically closed, then f has no multiple zeros in F.*

Proof If f has degree n, and $a_n \neq 0$ is the coefficient of X^n, then the X^{n-1} coefficient in f' is $na_n \neq 0$. Thus $a_n \neq 0$.

Corollary 1.21 *If $f \in \mathbb{Q}[X]$ is irreducible, then f has* deg f *distinct zeros in $\mathbb{C}$.*

In characteristic $p > 0$, it is possible to have $f' = 0$, but f nonconstant. For example the polynomial $f = X^4 + 1$ in $\mathbb{Z}_2[X]$ has $f' = 0$.

We need to work a little harder to get a counterexample to the Corollary in characteristic p. Suppose $F = \mathbb{Z}_2(t)$, the field of rational functions over $\mathbb{Z}_2$ in a single variable t. There is no square root of t in F. Thus $f = X^2 - t$ is irreducible but $f' = 0$.

Resultants

Suppose K is an algebraically closed field, $f, g \in K[X]$. Can we determine if f and g have a common solution?

Suppose $f = a_nX^n + a_{n-1}X^{n-1} + \ldots a_0$ and $g = b_mX^m + b_{m-1}X^{m-1} + \ldots + b_0$ where $a_n, b_m \neq 0$.

The *resultant* of f and g is the determinant of the following $(n+m) \times (n+m)$-matrix.

$$R_{f,g} = \begin{vmatrix} a_0 & a_1 & \ldots & \ldots & \ldots & a_n & 0 & \ldots & \ldots & 0 \\ 0 & a_0 & a_1 & \ldots & \ldots & \ldots & a_n & 0 & \ldots & 0 \\ & & & \ddots & \ddots & & & & & \\ 0 & \ldots & \ldots & 0 & a_0 & a_1 & \ldots & \ldots & \ldots & a_n \\ b_0 & b_1 & \ldots & \ldots & b_m & 0 & \ldots & \ldots & \ldots & 0 \\ 0 & b_0 & b_1 & \ldots & \ldots & b_m & 0 & \ldots & \ldots & 0 \\ & & & \ddots & \ddots & & & & & \\ 0 & \ldots & \ldots & \ldots & 0 & b_0 & b_1 & \ldots & \ldots & b_m \end{vmatrix}$$

where there are m rows of $a's$ and n rows of $b's$.

Theorem 1.22 *Let F be a field, $f, g \in F[X]$. Then the following are equivalent:*
i) f and g have a common nonconstant factor;
ii) $R_{f,g} = 0$.

Before giving the proof we recall some basic linear algebra. Consider the homogeneous system of linear equations

$$\begin{pmatrix} a_{1,1} & \ldots & a_{1,n} \\ & \vdots & \\ a_{n,1} & \ldots & a_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0.$$

This system always has $\vec{0}$ as a trivial solution.

Theorem 1.23 *If A is an $n \times n$ matrix over a field F, the following are equivalent:*
i) the homogeneous system

$$A\vec{x} = 0$$

has a nontrivial solution;
ii) the rows of A are linearly independent;
iii) det $A = 0$.

Proof of Theorem 1.22

i) $\Rightarrow$ ii) Suppose h is a common nonconstant factor $f = f_1h$ and $g = g_1h$. Note that $fg_1 = gf_1$. Let $f = \sum_{i=0}^{n} a_iX^i$ $g = \sum_{i=0}^{m} b_iX^i$ where $a_n, b_m \neq 0$. Since deg $f_1 \leq n-1$ and deg $g_1 \leq m-1$. Let

$$f_1 = \sum_{i=0}^{n-1} c_iX^i \text{ and } g_1 = \sum_{i=0}^{m-1} d_iX^i.$$

Then

$$fg_1 = \sum_{i=0}^{m+n-1} \sum_{j+k=i} a_j d_k X^i$$

and

$$gf_1 = \sum_{i=0}^{m+n-1} \sum_{j+k=i} b_j c_k X^i.$$

Since $fg_1 = gf_1$ we have the following system of equations

$$\begin{aligned} a_0d_0 &= b_0c_0 \\ a_0d_1 + a_1d_0 &= b_0c_1 + b_1c_0 \\ &\vdots \\ a_nc_{m-1} &= b_md_{n-1} \end{aligned}$$

If A is the matrix

$$\begin{pmatrix} a_0 & a_1 & \dots & \dots & \dots & a_n & 0 & \dots & \dots & 0 \\ 0 & a_0 & a_1 & \dots & \dots & \dots & a_n & 0 & \dots & 0 \\ & & & \ddots & \ddots & & & & & \\ 0 & \dots & \dots & 0 & a_0 & a_1 & \dots & \dots & \dots & a_n \\ -b_0 & -b_1 & \dots & \dots & -b_m & 0 & \dots & \dots & \dots & 0 \\ 0 & -b_0 & -b_1 & \dots & \dots & -b_m & 0 & \dots & \dots & 0 \\ & & & \ddots & \ddots & & & & & \\ 0 & \dots & \dots & \dots & 0 & -b_0 & -b_1 & \dots & \dots & -b_m \end{pmatrix},$$

Then $(d_0, \dots, d_{m-1}, c_0, \dots, c_{n-1})$ is a nontrivial solution to the homogenous linear system

$$(x_0, \dots, x_{m-1}, y_0, \dots, y_{n-1})A = 0.$$

This is a system of $n+m$ homogeneous linear equations in $n+m$ variables. If the rows of A are linearly independent, then the trivial solution is the unique solution. Thus the rows of A are linearly dependent and $\det A = 0$. But $\det A = (-1)^n \det R_{f,g}$.

$(\Leftarrow)$ Suppose $\det R = 0$. Then the system of equations

$$(x_0, \dots, x_{m-1}, y_0, \dots, y_{n-1})A = 0$$

has a nontrivial solution $(\alpha_0, \dots, \alpha_{m-1}, \beta_0, \dots, \beta_{n-1})$. Let

$$g_2 = \sum_{i=0}^{m-1} \alpha_i X^i \text{ and } f_2 = \sum_{i=0}^{m-1} \beta_i X^i.$$

Then $g_2 f = f_2 g$.

We now use unique factorization in $F[X]$. Factor $f = p_1 \ldots p_k$ and $f_2 = q_1 \ldots q_l$ where each p_i and q_i are irreducible. Renumber the p's and q's such that p_i and q_i are associates for $i \leq s$ and if $i, j > s$ then p_i and q_j are not associates. Then

$$f_2 = cp_1 \ldots p_s q_{s+1} \cdots q_l$$

where $c \in F$ and p_i and q_j are not associates for $s < i \leq k$, $s < j \leq l$. Since $\deg f > \deg f_2$ we have $s < k$. But p_{s+1} divides $\frac{f_2}{p_1 \cdots p_s} g$ and p_{s+1} does not divide $\frac{f_2}{p_1 \cdots p_s}$. Since p_{s+1} is irreducible, $\langle p_{s+1} \rangle$ is a prime ideal. Thus p_{s+1} divides g and f and g have a common factor.

Corollary 1.24 *If F is a field, $f, g \in F[X]$ and $K \supseteq F$ is an algebraically closed field, then f and g have a common zero in K if and only if $R_{f,g} = 0$.*

We return to the question of whether a polynomial has multiple roots in an algebraically closed extension.

Definition 1.25 If $f \in F[X]$ the *discriminant* of f is $R_{f,f'}$.

Corollary 1.26 *If F is a field of characteristic zero and $K \supseteq F$ is algebraically closed, then $f \in F[X]$ has a multiple zero in K if and only if the discriminant is zero.*

Computations in MAPLE

Suppose $f, g \in \mathbb{Q}[X]$. Theorem 1.22 gives us an easy way to decide if f and g have a common zero in $\mathbb{C}$, as we need only check if $R_{f,g} = 0$. This is easily done in MAPLE.

To calculate a resultant we need only now how to enter a matrix and take its determinant. Here are the steps you need to know.

i) Loading the linear algebra package.

```
> with(linalg);
```

ii) Enter an $n \times m$ matrix as `array([` $row_1 \ldots, row_m$`])` where row_i is `[` $x_1, \ldots, x_n$`]`. For example:

```
> A:=array([[1,2,3],[2,-1,-1],[0,1,1]]);
```

Enters the matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & -1 \\ 0 & 1 & 1 \end{pmatrix}.$$

```
> B:=array([[a,b,c],[1,0,-1]]);
```

Enters the matrix

$$\begin{pmatrix} a & b & c \\ 1 & 0 & -1 \end{pmatrix}.$$

iii) Computing the determinant.

```
> det(A);
```

Computes the determinant of a square matrix A.

Let $f(X) = X^4 + X^3 - X^2 + X - 2$ and $g(X) = X^3 + X^2 + X + 1$. The resultant $R_{f,g}$ is the determinant of a 9×9 matrix.

```
> A:=array([[1,1,-1,1,-2,0,0], [0,1,1,-1,1,-2,0], [0,0,1,1,-1,1,-2],
[1,1,1,1,0,0,0],[0,1,1,1,1,0,0],[0,0,1,1,1,1,0], [0,0,0,1,1,1,1]]);
```

$$A = \begin{pmatrix} 1 & 1 & -1 & 1 & -2 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & -2 & 0 \\ 0 & 0 & 1 & 1 & -1 & 1 & -2 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

```
> det(A);
```

$$\det\,(A) = 0$$

This “hands on” method works fine, but in fact MAPLE has a built in resultant function. If f and g are polynomials in variable v, then

`> resultant(f,g,v);` Computes the resultant.

For example

```
>resultant(X^4+X^3-X^2+X-2, X^3+X^2+X+1,X);
```

computes the resultant of $X^4 + X^3 - X^2 + X - 2$ and $X^3 + X^2 + X + 1$ and

```
>resultant(3*X^2+2*X+1, X^3+X^2+X+1,X);
```

computes the resultant of $3X^2 + 2X + 1$ and $X^3 + X^2 + X + 1$.

MAPLE also has factoring routines. Suppose $f \in \mathbb{Q}[X]$ and we want to find the number of zeros and their multiplicities in $\mathbb{C}$. For example: let $f = X^8 - X^6 - 2X^5 + 2X^3 + X^2 - 1$. Using MAPLE we can find an irreducible factorization of f in $\mathbb{Q}[X]$.

```
> factor(X^8-X^6-2*X^5+2*X^3+X^2-1);
```

Gives us the factorization $(X^2+X+1)^2(X-1)^3(X+1)$. Since polynomial $X^2 + X + 1$ is irreducible in $\mathbb{Q}[X]$, it has two distinct complex zeros α and β. The zeros of f are $\alpha, \beta, 1, -1$. The zeros α and β have multiplicity 2, while 1 has multiplicity 3 and -1 has multiplicity 1.

We can also factor over $\mathbb{Z}_p$. Let $g = X^7+2X^5+2X^4+X^3+4X^2+2 \in \mathbb{Z}_7[X]$.

```
> Factor(X^7+2*X^5+2*X^4+X^3+4X^2+2) mod 7;
```

Gives us the factorization $(X^2 + 1)^2(X^3 + 2)$. Suppose $K \supset \mathbb{Z}_7$ is an algebraically closed field. Since $(X^2 + 1)' = 2X \neq 0$, the polynomial $X^2 + 1$ has no multiple zeros. Similarly, $(X^3 + 2)$ has no multiple zeros. It follows that g has 5 zeros. Two of them have multiplicity 2, the other have multiplicity 1.

2 Affine Lines and Conics

Let k be a field.

Definition 2.1 *Affine n-space* over k is

$$\mathbb{A}_n(k) = \{(x_1, \ldots, x_n) : x_1, \ldots, x_n \in k\}.$$

Definition 2.2 We say that $V \subseteq \mathbb{A}_n(k)$ is an *algebraic set* if there are polynomials $f_1, \ldots, f_m \in k[X_1, \ldots, X_n]$ such that

$$V = \{x \in \mathbb{A}_n(k) : f_1(x) = f_2(x) = \ldots = f_m(x) = 0\}.$$

If $m = 1$, i.e. V is the solutions of a single polynomial, we call V a *hypersurface.* We will be particularly interested in the case where $n = 2$ and $m = 1$. In this case we call V a *plane algebraic curve.*

Suppose $f \in k[X_1, \ldots, X_n]$ is nonzero. We can write

$$f = \sum_{i_1=0}^{m_1} \ldots \sum_{i_n=0}^{m_n} a_{i_1,\ldots,i_n} X_1^{i_1} X_2^{i_2} \ldots X_n^{i_n}.$$

Definition 2.3 The *degree* of f is defined by

$$\deg f = \max\{i_1 + \ldots + i_n : i_1 \leq m_1, i_2 \leq m_2, \ldots, i_n \leq m_n, a_{i_1,\ldots,i_n} \neq 0\}.$$

For example, $X^4 + 3X^3YZ - X^2Y + YZ^3$ has degree 5 because of the $3X^3YZ$-term. If f is a nonzero constant, then f has degree 0.

We begin by carefully studying plane curves of degree 1 and 2. Our main tools will be high school algebra and some very elementary calculus and linear algebra.

Lines

Polynomials $f \in k[X, Y]$ of degree 1 are called *linear.* A linear polynomial is of the form

$$aX + bY + c$$

where at least one of a and b is nonzero. The zero set of a linear polynomial is called a *line.*

Of course, if $k = \mathbb{R}$ then lines have a clear geometric meaning. But if k is the field $\mathbb{Z}_3$ then the line $X + 2Y + 1 = 0$ is just the discrete set of points $L = \{(0, 1), (1, 2), (2, 0)\}$. We will see that the well-know geometric properties of lines hold in arbitrary fields even when there is no obvious geometry.

Suppose L is the line $aX + bY + c = 0$. We can easily find $\phi : k \to L$ a parametrization of L. If $b \neq 0$, let

$$\phi(t) = \left(t, \frac{-c - at}{b}\right)$$

while, if $b = 0$, then $a \neq 0$ and we let

$$\phi(t) = \left(\frac{-c}{a}, t\right).$$

We leave the proof of the following proposition as an exercise.

Proposition 2.4 *$\phi : k \to L$ is a bijection. In particular, if k is infinte, then so is L.*

Suppose $f_i = a_iX + b_iY + c_i$ for $i = 1, 2$. Let L_i be the line $f_i = 0$. If $(x, y) \in L_1 \cap L_2$, then

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

Linear algebra tells us exactly what the solutions look like.

If the matrix $A = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$ is invertible, then $A^{-1} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ is the unique solution.

But A is invertible if and only if the rows are linearly independent. Thus A is not invertible if and only if there is a λ such that $a_2 = \lambda a_1$ and $b_2 = \lambda b_1$. In this case there are two possibilities. If $c_2 = \lambda c_1$, then $f_2 = \lambda f_1$ and L_1 and L_2 are the same line. If $c_2 \neq \lambda c_1$, then the system has no solution and $L_1 \cap L_2 = \emptyset$.

We summarize these observations in the following proposition.

Proposition 2.5 *Suppose $f_1, f_2 \in k[X]$ are linear polynomials and L_i is the line $f_i = 0$ for $i = 1, 2$.*

i) $L_1 = L_2$ if and only if $f_2 = \lambda f_1$ for some $\lambda \in k$.

ii) If L_1 and L_2 are distinct lines, then $|L_1 \cap L_2| \leq 1$.

iii) If $L_1 \cap L_2 = \emptyset$ and $f_1 = a_1X + b_1Y + c_1$, then for some λ $f_2 = \lambda a_1X + \lambda b_1Y + d$ where $d \neq \lambda c_1$.

One sees that "usually" two distinct lines intersect in exactly one point. The fact that we can have parallel lines that do not intersect is one of the annoying features of affine space.

Proposition 2.6 *If (x_1, y_1) and (x_2, y_2) are distinct points in $\mathbb{A}_2(k)$, there is a unique line containing both points.*

Proof We are looking for $f = aX + bY + c$ such that

$$\begin{aligned} ax_1 + by_1 + c &= 0 \\ ax_2 + by_2 + c &= 0. \end{aligned}$$

This is a system of linear equations in the variables a, b, c. Since (x_1, y_1) and (x_2, y_2) are distinct, the rows of the matrix $\begin{pmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{pmatrix}$ are linearly independent. Since the matrix has rank 2, linear algebra tells us that we can find an nontrivial solution (a, b, c) and every other solution is of the form $(\lambda a, \lambda b, \lambda c)$. Thus there is a unique line through (x_1, y_1) and (x_2, y_2).

Affine Transformation

In $\mathbb{R}^2$ we can transform any line to any other line by rotating the plane and then translating it. We will show that this is possible for any field.

Definition 2.7 We say that $T : \mathbb{A}_2(k) \to \mathbb{A}_2(k)$ is an *affine transformation* if there are linear polynomials f and g such that $T(x, y) = (f(x, y), g(x, y))$. In this case there is a 2×2 matrix A with entries from k and a vector $\vec{b} \in k^2$ such that

$$\begin{pmatrix} f(x,y) \\ g(x,y) \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} + \vec{b}.$$

If $\vec{b} = 0$ we say that T is a *linear transformation.*

An affine transformation can be though of as a linear change of variables

$$\begin{aligned} U &= a_1X + b_1Y + c_1 \\ V &= a_2X + b_2X + c_2 \end{aligned}$$

Proposition 2.8 *If C is a line in $\mathbb{A}_2(k)$ there is an invertible affine transformation taking C to the line $X = 0$.*

Proof Suppose C is given by the equation $aX + bY + c = 0$.

If $a \neq 0$, consider the affine transformation

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} c \\ 0 \end{pmatrix}.$$

In other words we make the invertible change of variables $U = aX + bY + c$ and $V = Y$. This transforms C to the line $U = 0$.

If $a = 0$, we use the transformation

$$\begin{pmatrix} 0 & b \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} c \\ 0 \end{pmatrix},$$

i.e., the change of variables $U = bY + c$ and $V = X$, to transform C to $U = 0$.

Conics

We next look at solution sets to second degree equations in $\mathbb{A}_2(k)$. Our first goal is the following theorem.

Theorem 2.9 *Suppose k is a field and the characteristic of k is not 2. If $p(X, Y) \in k[X, Y]$ has degree 2, there is an affine transformation taking the curve $p(X, Y) = 0$ to one of the form $aX^2+bY^2+c = 0$ where $b \neq 0$, $aX^2+Y = 0$ or $X^2 + c = 0$.*

We will prove this by making a sequence of affine transformations. We begin with a polynomial

$$aX^2 + bY^2 + cXY + dX + eY + f = 0.$$

Claim 1 We may assume that $a \neq 0$.

If $a = 0$ and $b \neq 0$, we use the transformation $T(x, y) = (y, x)$.

If $a = b = 0$, then since the polynomial has degree 2 we must have $c \neq 0$. We make the change of variables

$$\begin{aligned} X &= X \\ V &= Y - X \end{aligned}$$

Then $XV = XY - X^2$ and our curve is transformed to

$$cX^2 + cXV + (d + e)X + eV + f = 0.$$

Claim 2 We may assume that $c = 0$.

This is the old algebra trick of "completing the square". We make a change of variables $U = X + \alpha Y$ so that $aU^2 = aX^2 + cXY + \beta Y^2$ for some appropriate β. To get this to work we would need $2a\alpha = c$. So we make the change of variables $U = X + \frac{c}{2a}Y$. Note at this point we have to divide by 2. This is one reason we had to assume that the characteristic of k is not 2.

This change of variables transforms the curve to:

$$aU^2 + \left(b - \frac{c^2}{4a^2}\right)Y^2 + dU + \left(e - \frac{dc}{2a}\right)Y + f = 0.$$

Thus, by affine transformations, we may assume that our curve is given by

$$aX^2 + bY^2 + cX + dY + e = 0.$$

There are two cases to consider.

case 1 $b \neq 0$.

We must do two more applications of completing the square. First we make a change of variables $U = X + \alpha$ so that $aU^2 = aX^2 + cX + \beta$. We need $2a\alpha = c$ so $\alpha = \frac{c}{2a}$. Similarly we let $V = Y + \frac{d}{2a}$. The transformed curve is given by

$$aU^2 + bV^2 + e - \frac{c^2 + d^2}{4a^2} = 0.$$

case 2 $b = 0$.

As above we complete the square by taking $U = X + \frac{c}{2a}$. That transforms the curve to

$$aU^2 + dY + e - \frac{c^2}{4a^2} = 0.$$

If $d \neq 0$, the change of variables $V = dY + e - \frac{c^2}{4a^2}$ gives us

$$aU^2 + V = 0.$$

If $d = 0$, then the curve is already given by the equation

$$X^2 + \frac{e - \frac{c^2}{4a^2}}{a} = 0.$$

Conics in $\mathbb{A}_2(\mathbb{R})$

When our field is the field $\mathbb{R}$ of real numbers, we can give even more precise information.

Theorem 2.10 *If $p \in \mathbb{R}[X, Y]$ has degree 2, then there is an affine tranformation taking the curve $p = 0$ to one of the following curves:*
i) (parabola) $Y = X^2$;
ii) (circle) $X^2 + Y^2 = 1$;
iii) (hyperbola) $X^2 - Y^2 = 1$;
iv) (point) $X^2 + Y^2 = 0$;
v) (crossed-lines) $X^2 - Y^2 = 0$;
vi) (double line) $X^2 = 0$;
vii) (parallel lines) $X^2 = 1$
viii) (empty set) $X^2 = -1$
ix) (empty set) $X^2 + Y^2 = -1$.

Proof
If we can transform p to $aX^2 + Y = 0$, then the tranformation $V = \frac{-Y}{a}$ gives the parabola $V = X^2$.

Suppose we can transform p to $X^2 + c = 0$. If $c < 0$, the transformation $U = \frac{X}{\sqrt{-c}}$, transforms p to $X^2 = 1$. While if $c > 0$, the transformation $U = \frac{X}{\sqrt{c}}$, transforms p to $X^2 = -1$.

Suppose we have transformed p to $aX^2 + bY^2 + c = 0$.

case 1 $c \neq 0$

Our curve is the same as the curve $\frac{-a}{c}X^2 + \frac{-b}{c}Y^2 = 1$. Thus we may assume our curve is $aX^2 + bY^2 = 1$.

case 1.1 $a > 0$ and $b > 0$

The transformation $U = \frac{X}{\sqrt{a}}$, $V = \frac{Y}{\sqrt{b}}$ transforms the curve to the circle $U^2 + V^2 = 1$.

case 1.2 $a > 0$ and $b < 0$

The transformation $U = \frac{X}{\sqrt{a}}$, $V = \frac{Y}{\sqrt{-b}}$ transforms the curve to the hyperbola $U^2 - V^2 = 1$.

case 1.3 $a < 0$ and $b > 0$

We make the transformation $T(x, y) = (y, x)$ which transforms the curve to case 1.2.

case 1.4 $a < 0$ and $b < 0$

The transformation $U = \frac{X}{\sqrt{-a}}$, $V = \frac{Y}{\sqrt{-b}}$ transforms the curve to $U^2 + V^2 = -1$, which has no solutions in $\mathbb{R}^2$.

case 2 $c = 0$.

If $b = 0$, then we have curve $aX^2 = 0$ which is the same as $X^2 = 0$.

Suppose $b \neq 0$. We may assume $a > 0$. If $b > 0$, then the change of variables $U = \frac{X}{\sqrt{a}}$, $V = \frac{Y}{\sqrt{b}}$ transforms the curve to $U^2 + V^2 = 0$ which has a single solution $\{(0, 0)\}$.

If $b < 0$, then the change of variables $U = \frac{X}{\sqrt{a}}$, $V = \frac{Y}{\sqrt{-b}}$ transforms the curve to $U^2 - V^2 = 0$. The solution set is pair of lines $U + V = 0$ and $U - V = 0$.

Cases i)–iii) are considered nondegenerate. In §4 we will see how we make this distinction.

The classification we have given is optimal for $\mathbb{A}_2(\mathbb{R})$ as no affine transformation can take one of these curves to another one. For example, if C is a circle and T is an affine transformation, since C is compact $T(C)$ is also compact. This means there is no affine transformation of $\mathbb{A}_2(k)$ taking C to a parabola or a hyperbola. Since parabolas are connected and the continuous image of a connected set is connected, an affine transformation can't take a parabola to a hyperbola.

Conics in $\mathbb{A}_2(\mathbb{C})$

Over the complex field, we can simplify the classification.

Theorem 2.11 *If $p \in \mathbb{C}[X, Y]$ has degree 2, then there is an affine tranformation taking the curve $p = 0$ to one of the following curves:*
i) (parabola) $Y = X^2$;
ii) (circle) $X^2 + Y^2 = 1$;
iii) (crossed-lines) $X^2 - Y^2 = 0$;
iv) (double line) $X^2 = 0$;
v) (parallel lines) $X^2 = 1$.

Proof We have gotten rid of four cases.

The change of variable $V = iY$ transforms the hyperbola $X^2 - Y^2 = 1$ to the circle $X^2 + V^2 = 1$.

The same change of variables transforms the $X^2 + Y^2 = 0$ to the double line $X^2 - V^2 = 0$ (Note: in $\mathbb{A}_2(\mathbb{C})$ $X^2 + Y^2 = 0$ is the two lines $(X + iY) = 0$ and $(X - iY) = 0$.)

Finally the transformations $U = iX$, $V = iY$ transform $X^2 + Y^2 = -1$ to the circle $U^2 + V^2 = 1$ and transform $X^2 = -1$ to $X^2 = 1$.

Exercise 2.12 Show that the same classification works in $\mathbb{A}_2(K)$ for any algebraically closed field K with characteristic different from 2.

Intersecting Lines and Circles

For the moment we will restrict our attention to $\mathbb{A}_2(\mathbb{R})$.

What happens when we intersect the circle C with the line L given by equation $Y = aX + b$? If $(x, y) \in C \cap L$ then

$$\begin{aligned} y &= ax + b \\ x^2 + y^2 &= 1 \end{aligned}$$

Thus $x^2 + (ax + b)^2 = 1$ and $(a^2 + 1)x^2 + 2abx + b^2 = 1$. The polynomial

$$g(X) = (a^2 + 1)X^2 + 2abX + (b^2 - 1)$$

is a nonzero polynomial of degree at most 2 (Note: if $a^2 = -1$, then g has degree at most 1), and hence has at most two solutions. Thus $|C \cap L| \leq 2$. It is easy to see that all of these possibilites occur. Let L_a be the line $Y = a$. Then $C \cap L_0 = \{(\pm 1, 0)\}$, $C \cap L_1 = \{(0, 1)\}$ and $C \cap L_2 = \emptyset$.

A separate but similar argument is needed to show that lines $x = a$ will also intersect C in at most two places. The same ideas work just as well for nondegenerate conics.

Proposition 2.13 *If C is any parabola, circle or hyperbola in $\mathbb{A}_2(\mathbb{R})$ and L is a line, then $|L \cap C| \leq 2$.*

Note that the proposition is not true for degenerate conics in $\mathbb{A}_2(\mathbb{R})$. For example, the conic $X^2 = 1$ has infinite intersection with the line $X = 1$.

Let's work in $\mathbb{A}_2(\mathbb{C})$ instead of $\mathbb{A}_2(\mathbb{R})$. There are several cases to consider.

Case 1 $a = \pm i$ and $b = 0$.

In this case g is the constant polynomial -1 and there are no solutions.

Case 2 $a = \pm i$ and $b \neq 0$.

In this case g is a linear polynomial and there is a single solution.

Case 3 $a \neq \pm i$ and $b^2 - a^2 = 1$.

In this case

$$g = \frac{1}{a^2 + 1}((a^2 + 1)X + ab)^2$$

and the unique point of intersection is $(\frac{-ab}{a^2+1}, \frac{b}{a^2+1})$.

If (c, d) is a point on the circle, then a little calculus shows that the tangent line has slope $-\frac{c}{d}$ and equation

$$Y = -\frac{c}{d}X + \frac{c^2}{d} + d.$$

Thus the tangent line at $(\frac{-ab}{a^2+1}, \frac{b}{a^2+1})$ is $Y = aX + b$. In other words, case 3 arises when the line $Y = aX + b$ is tangent to the curve.

Case 4 $a \neq \pm i$ and $b^2 - a^2 \neq 1$.

In this case g has two distinct zeros in $\mathbb{C}$ and $|C \cap L| = 2$.

Case 4 is the general case. We understand why there is only one solution in case 3 as the line is tangent. Cases 1 and 2 are annoying. In the next section we will see that they occur because the lines have intersections "at infinity".

Parameterizing Circles

Consider the circle $C \subset \mathbb{A}_2(\mathbb{R})$ given by the equation $X^2 + Y^2 = 1$. It is well known that we can parameterize the curve by taking $f(t) = (\cos t, \sin t)$. There are two problems with this parameterization. First, we are using transcendental functions. Second, the parameterization is not one-to-one. For each $(x, y) \in C$ if $f(\theta) = (x, y)$, then $f(\theta + 2\pi) = (x, y)$. We will show how to construct a better parameterization.

Note that the point $(0, 1) \in C$. Let L_λ be the line $\lambda X + Y = 1$. Then L_λ is the line throught the point $(0, 1)$ with slope $-\lambda$. Consider $L_\lambda \cap C$. If $(x, y) \in L_\lambda \cap C$, then

$$\begin{aligned} x^2 + (1 - \lambda x)^2 &= 1 \\ (\lambda^2 + 1)x^2 - 2\lambda x &= 0 \\ x((\lambda^2 + 1)x - 2\lambda x) &= 0 \end{aligned}$$

Thus there are two solutions $x = 0$ and $x = \frac{2\lambda}{\lambda^2+1}$. The solution $x = 0$ corresponds to the point we already know $(0, 1)$. So we have one additional point

$$x = \frac{2\lambda}{\lambda^2 + 1}, y = \frac{1 - \lambda^2}{\lambda^2 + 1}.$$

We consider the parameterization $\rho : \mathbb{R} \to C$ given by

$$\rho(\lambda) = \left(\frac{2\lambda}{\lambda^2 + 1}, \frac{1 - \lambda^2}{\lambda^2 + 1}\right).$$

Suppose $(x, y) \in C$ and $x \neq 0$. Let $\lambda = \frac{1-y}{x}$. Then

$$\begin{align*} \rho(\lambda) &= \left(\frac{2\frac{1-y}{x}}{\frac{1-y}{x}^2 + 1}, \frac{1 - \frac{1-y}{x}^2}{\frac{1-y}{x}^2 + 1}\right) \tag{1} \\ &= \left(\frac{2(1-y)x}{1 - 2y + y^2 + x^2}, \frac{x^2 - 1 - 2y + y^2}{(1 - 2y + y^2) + x^2}\right) \tag{2} \\ &= \left(\frac{2(1-y)x}{2 - 2y}, \frac{2y - 2y^2}{2 - 2y}\right) \tag{3} \\ &= (x, y) \tag{4} \end{align*}$$

Where we get (2) from (3) since $x^2 + y^2 = 1$ and get (4) since $y \neq 1$. The following proposition is now clear.

Proposition 2.14 *The parameterization $\rho : \mathbb{R} \to C$ is one-to-one. The image is $C \setminus (0, -1)$.*

It is annoying that the image misses one point on the circle. If we wanted to get the point $(0, -1)$ we would need to use the line $x = 0$ with infinite slope.

This construction is very general and could be used for any nondegenerate conic once we know one point.

Exercise 2.15 Find a rational parameterization of the hyperbola $X^2 - 2Y^2 = 1$. [Hint: Start with the point $(1, 0)$.]

Connections to Number Theory

It is well known that the equation $X^2 + Y^2 = Z^2$ has many solutions in $\mathbb{Z}$. For example $(1,0,1)$, $(3,4,5)$ and $(5,12,13)$. If (a,b,c) is a solution and n is any integer then (na,nb,nc) is also a solution. Thus there are infinitely many integer solutions. Can we find infinitely many solutions where (a,b,c) have no common factor greater than 1?

If (a,b,c) is a solution to $X^2+Y^2=Z^2$ other than $(0,0,0)$, then $(a/c, b/c)$ is a solution to

$$X^2+Y^2=1.$$

Thus we are interested in studying points on the circle C in $\mathbb{A}_2(\mathbb{Q})$. Let $C(\mathbb{Q}) = C \cap \mathbb{A}_2(\mathbb{Q})$. Let $\rho(\lambda) = \left(\frac{2\lambda}{\lambda^2+1}, \frac{1-\lambda^2}{\lambda^2+1}\right)$ be the parameterization of C. Note that if $\lambda \in \mathbb{Q}$ then $\rho(\lambda) \in \mathbb{Q}$ and if $(x,y) \in C(\mathbb{Q})$, then $\lambda = \frac{1-y}{x} \in \mathbb{Q}$. Hence $\rho : \mathbb{Q} \to C(\mathbb{Q}) \setminus \{(0,-1)\}$ is a bijection.

Suppose $p_1, p_2 \in C(\mathbb{Q})$ with $p_1 \neq p_2$. We can write $p_i = (a_i/c_i, b_i/c_i)$ where a_i, b_i, c_i have no common factor. In particular there is no $\lambda \in \mathbb{Z}$ such that that $(a_2,b_2,c_2) = \lambda(a_1,b_1,c_1)$, Thus distinct points on $C(\mathbb{Q})$ give rise to distinct relatively prime solutions to $X^2+Y^2=Z^2$. Thus there are infinitely many solutions (a,b,c) where a,b,c have no common factors.

If m and n are relatively prime integers, then

$$\rho(\frac{m}{n}) = \left(\frac{2\frac{m}{n}}{\frac{m^2}{n^2}+1}, \frac{1-\frac{m^2}{n^2}}{\frac{m^2}{n^2}+1}\right)$$

and $(2mn, n^2-m^2, n^2+m^2)$ is a solution to $X^2+Y^2=Z^2$. If one of m and n are even, then $2mn, n^2-m^2, n^2+m^2$ have no common factor . If both m and n are odd, then mn, $\frac{n^2-m^2}{2}$, $\frac{n^2+m^2}{2}$ have no common factor. We find all integral solutions this way.

Similar ideas work for any nondegenerate conic defined over the integers. Once we have a point in $\mathbb{A}_2(\mathbb{Q})$.

Exercise 2.16 Use the parameterization from Exercise 2.15 to find all integral solutions to $X^2-2Y^2=Z^2$.

Of course in order for this to work we need to find at least one point in $\mathbb{A}_2(\mathbb{Q})$.

Proposition 2.17 *The only integral solution to* $X^2+2Y^2=5Z^2$ *is* $(0,0,0)$. *Thus the ellipse* $X^2+2Y^2=5$ *has no points in* $\mathbb{A}_2(\mathbb{Q})$.

Proof Suppose $(a,b,c) \in \mathbb{Z}^3$ is a nonzero solution to $X^2+2Y^2=5Z^2$. By dividing by the greatest common divisor, we may assume that a,b,c have no common factor. Note that if $a \neq 0$ or $b \neq 0$, then $c \neq 0$. Thus we may assume that $c \neq 0$.

Since $a^2+2b^2=5c^2$, we also have

$$\begin{aligned} a^2+2b^2 &= 5c^2 \bmod 5 \\ a^2+2b^2 &= 0 \quad \bmod 5 \end{aligned}$$

The only squares mod 5 are $0, 1, 4$. If $a^2 + 2b^2 = 0$ mod 5, we must have $a = b = 0$ mod 5. But then a and b are both divisible by 5 and $c^2 = a^2 + 2b^2$ is divisible by 25. Considering the factors of c we see that c is divisible by 5 contradicting our assumtion that a, b, c have no common factors.

If $(a_1/n_1, a_2/n_2)$ is a rational solution to $X^2+2Y^2 = 5$, then (n_2a_1, n_1a_2, n_1n_2) is an integral solution to $X^2+2Y^2 = 5Z^2$. Thus there are no points in $\mathbb{A}_2(\mathbb{Q})$.

3 Projective Space

One important problem in algebraic geometry is understanding $|L \cap C|$ where L is a line and C is a curve of degree d. The basic idea is that if L is given by $Y = aX + b$ and C is given by $f(X, Y) = 0$, we substitute and must solve the equation $f(X, aX + b) = 0$. Usually this is a polynomial of degree d and there are at most d solutions. We would like to say that there are exactly d solutions, but we have already seen examples where this is not true.

1) In $\mathbb{A}_2(\mathbb{R})$ the line $Y = X$ is a subset of the solution set to $X^2 - Y^2 = 0$. In this case $L \cap C$ is infinite.

2) In $\mathbb{A}_2(\mathbb{C})$ the line $Y = 1$ is tangent to the circle $X^2 + Y^2 = 1$ and $|L \cap C| = 1$.

3) In $\mathbb{A}_2(\mathbb{R})$ the line $Y = 2$ does not intersect the circle $X^2+Y^2 = 1$, because there are no real solutions to the equation $X^2 + 3 = 0$.

4) In $\mathbb{A}_2(\mathbb{C})$ the point $(0, 1)$ is the only point of intersection of the line $Y = iX + 1$ and the circle $X^2 + Y^2 = 1$, because $X^2 + (iX + 1)^2 = 1$ if and only if $X = 0$.

5) Even if we only consider the intersection of two lines L_1 and L_2 we might have no intersection points if the lines are parallel.

We can avoid problem 1) by only looking at the cases where $L \not\subseteq C$. For example in §4 we will see that this holds if C is irreducible and $d > 1$. Problem 2) is unavoidable. We will eventually get around this by carefully assigning multiplicities to points of intersection. Tangent lines will intersect with multiplicity at least 2. This will allow us to prove results saying that "counted correctly" there are d points of intersection. In arbitrary fields k we will always run into problems like 3) where there are fewer than d points of intersection because there are polynomials with no zeros. We can avoid this by restricting our attention to algebraically closed fields.

In this section we will try to avoid problems 4) and 5) by working in projective space rather than affine space. In $\mathbb{P}_2(\mathbb{C})$ we will find extra intersection points "at infinity". There are a few other problems that working in projective space will solve.

6) The parameterization we gave of the circle missed the point $(0, -1)$ because we needed to use the line $X = 0$ with "infinite slope".

7) When we consider the hyperbola $X^2 - Y^1 = 1$ we see that curve is asymptotic to the lines $Y = \pm X$. Can we define "asymptote" so that the

concept makes sense in arbitrary fields k?

$\mathbb{P}_n(k)$

Let k be a field. We define an equivalence relation $\sim$ on $k^{n+1} \setminus \{(0, \dots, 0)\}$ by $(x_1, \dots, x_{n+1}) \sim (y_1, \dots, y_{n+1})$ if and only if there is $\lambda \in k$ such that

$$(y_1, \dots, y_{n+1}) = (\lambda x_1, \dots, \lambda x_{n+1}).$$

We let

$$[(x_1, \dots, x_{n+1})] = \{(y_1, \dots, y_{n+1}) \in k^{n+1} \backslash \{0\} : (x_1, \dots, x_{n+1}) \sim (y_1, \dots, y_{n+1})\}$$

denote the equivalence class of $(x_1, \dots, x_{n+1})$.

Definition 3.1 *Projective n-space* over a field k is

$$\mathbb{P}_n(k) = \{[(x_1, \dots, x_{n+1})] : (x_1, \dots, x_{n+1}) \in k^{n+1} \setminus \{0\}\}.$$

We say that $(x_1, \dots, x_{n+1})$ is a set of *homogeneous coordinates* for the $\sim$-equivalence class $[(x_1, \dots, x_{n+1})]$.

For any point $p \in \mathbb{P}_n(k)$ we have a number of choices for homogeneous coordinates. If $(x_1, \dots, x_{n+1})$ is one choice of homogeneous coordinates for p. Then $[(x_1, \dots, x_{n+1})]$ is exactly the line with parametric equation

$$f(t) = \begin{pmatrix} x_1 t \\ x_2 t \\ \vdots \\ x_{n+1} t \end{pmatrix}.$$

Thus the $\sim$-equivalence classes are exactly the lines through $(0, \dots, 0)$ in k^{n+1}. This gives alternative characterization of $\mathbb{P}_n(k)$.

Proposition 3.2 *There is a bijection between $\mathbb{P}_n(k)$ and the set of lines through 0 in k^{n+1}.*

Let $U = \{p \in \mathbb{P}_n(k) : p$ has homogeneous coordinates $(x_1, \dots, x_{n+1})$ where $x_{n+1} \neq 0\}$. If $[(x_1, \dots, x_{n+1}] \in U$, then

$$(x_1, \dots, x_{n+1}) \sim \left(\frac{x_1}{x_{n+1}}, \dots, \frac{x_n}{x_{n+1}}, 1\right)$$

and if $(y_1, \dots, y_{n+1}, 1)$ are also homogeneous coordinates for p, then $y_1 = \frac{x_1}{x_{n+1}}, \dots, y_n = \frac{x_n}{x_{n+1}}$.

Proposition 3.3 *The map $(x_1, \dots, x_n) \mapsto [(x_1, \dots, x_n, 1)]$ is a bijection between $\mathbb{A}_n(k)$ and U*

In this way we view $\mathbb{A}_n(k)$ as a subset of $\mathbb{P}_n(k)$. Note that we had a great deal of freedom in this choice. If $U_i = \{p \in \mathbb{P}_n(k) : p$ has homogeneous coordinates $(x_1, \ldots, x_{n+1})$ where $x_i \neq 0\}$. Then we could also identify $\mathbb{A}_n(k)$ with U_i.

We think of the points of $\mathbb{P}_n(k) \setminus U$ as being "points at infinity". The points in $\mathbb{P}_n(k) \setminus U$ are those with homogeneous coordinates $(x_1, \ldots, x_n, 0)$ where not all $x_i = 0$. Note that $(x_1, \ldots, x_n, 0) \sim (y_1, \ldots, y_n, 0)$ if and only if $(x_1, \ldots, x_n) \sim (y_1, \ldots, y_n)$. This proves:

Proposition 3.4 *The map* $[(x_1, \ldots, x_n, 0)] \mapsto [(x_1, \ldots, x_n)]$ *is a bijection between* $\mathbb{P}_n(k) \setminus U$ *and* $\mathbb{P}_{n-1}(k)$.

We look more carefully at $\mathbb{P}_n(k)$ for $n = 0, 1, 2$.

For $n = 0$, if $x, y \in k \setminus \{0\}$, then $(y) = \frac{y}{x}(x)$. Thus $(y) \sim (x)$ and $\mathbb{P}_0(k)$ is a single point.

For $n = 1$, we have $U = \{p \in \mathbb{P}_1(k) : p$ has homogeneous coordinates $(x, 1)\}$ that we identify with $\mathbb{A}_1(k)$. There is a unique point $p \in \mathbb{P}_1(k) \setminus U$ and p has homogeneous coordinates $(1, 0)$. We call $\mathbb{P}_1(k)$ the *projective line* over k.

Here is another way to think about $\mathbb{P}_1(\mathbb{R})$. Consider the upper semi-circle $X = \{(x, y) : x^2 + y^2 = 1, x, y \geq 0\}$ in $\mathbb{A}_2(\mathbb{R})$. The line $Y = 0$ intersects X in two point $(\pm 1, 0)$, while all other lines through $(0, 0)$ intersect X in exactly one point. When we identify $(1, 0)$ and $(-1, 0)$, we see that $\mathbb{P}_1(\mathbb{R})$ topologically looks like a circle.

For $n = 2$, we have $U = \{p \in \mathbb{P}_2(k) : p$ has homogeneous coordinates $(x, y, 1)\}$ that we identify with $\mathbb{A}_2(k)$. The points in $\mathbb{P}_2 \setminus U$ have homogeneous coordinates $(x, y, 0)$. We can either think of $\mathbb{P}_2 \setminus U$ as a projective line or divide into two pieces $V = \{p : p$ has homogeneous coordinates $(x, 1, 0)\}$ that looks like an affine line and the remaining point with homogeneous coordinates $(1, 1, 0)$

A similar argument to the one above shows that $\mathbb{P}_2(\mathbb{R})$ looks like the topological space obtained by identifying antipodal points on a sphere in $\mathbb{R}^3$.

Projective Algebraic Sets in $\mathbb{P}_2(k)$

We will restrict our attention to the projective plane $\mathbb{P}_2(k)$, though the ideas we present generalize immediately to $\mathbb{P}_n(k)$.

How do we talk about solutions to polynomial equations in $\mathbb{P}_2(k)$? Some care is needed. For example, let $f(X, Y, Z) = X^2 + Y^2 + Z$. If $p = \in \mathbb{P}_2(\mathbb{R})$ has homogeneous coordinates $(1, 1, -2)$ then $f(1, 1, -2) = 0$. But p also has homogeneous coordinates $(3, 3, -6)$ and $f(3, 3, -6) = 12$.

Definition 3.5 A *monomial of degree* d in $k[X, Y, Z]$ is a polynomial $aX^iY^jZ^k$ where $a \in k$ and $i + j + k = d$.

We say that a polynomial $f \in k[X, Y, Z]$ is *homogeneous of degree* d if it is a sum of monomials of degree d.

We say f is homogeneous if it is homogeneous of degree d for some d.

For example $f(X, Y, Z) = X^2 + Y^2 - Z^2$ is homogeneous of degree 2.

Proposition 3.6 *If f is homogeneous of degree d, then*

$$f(tx,ty,tz) = t^d f(x,y,z)$$

for all $t,x,y,z \in k$.

Proof If f is homogeneous of degree d, then

$$f(X,Y,Z) = \sum_{n=1}^{m} a_n X^{i_n} Y^{j_n} Z^{k_n}$$

where each $a_n \in k$ and $i_n + j_n + k_n = d$ for each $n = 1, \ldots, m$. Then

$$\begin{aligned} f(tx,ty,tz) &= \sum_{n=1}^{m} a_n (tx)^{i_n} (ty)^{j_n} (tz)^{k_n} \\ &= \sum_{n=1}^{m} a_n t^{i_n+j_n+k_n} x^{i_n} y^{j_n} z^{k_n} \\ &= \sum_{n=1}^{m} a_n t^d x^{i_n} y^{j_n} z^{k_n} \\ &= t^d f(x,y,z). \end{aligned}$$

Corollary 3.7 *If $f \in k[X,Y,Z]$ is homogeneous, and $f(x,y,z) = 0$, then $f(\lambda x, \lambda y, \lambda z) = 0$ for all $\lambda \in k$.*

Thus if (x_1,y_1,z_1) and (x_2,y_2,z_2) are different homogeneous coordinates for $p \in \mathbb{P}_2(k)$ and f is homogeneous, then $f(x_1,y_1,z_1) = 0$ if and only if $f(x_2,y_2,z_2) = 0$.

Definition 3.8 An *algebraic set* in $\mathbb{P}_2(k)$ is the set of

$$\{[(x,y,z)] \in \mathbb{P}_2(k) : f_1(x,y,z) = \ldots = f_m(x,y,z) = 0\}$$

where $f_1, \ldots, f_m$ are homogeneous polynomials.

A *projective curve* in $\mathbb{P}_2(k)$ is

$$\{[(x,y,z)] \in \mathbb{P}_2(k) : f(x,y,z) = 0\}$$

where f is a nonzero homogeneous polynomial.

Lines in $\mathbb{P}_2(k)$

A homogeneous polynomial of degree 1 in $k[X,Y,Z]$ is of the form $aX+bY+cZ$ where at least one of $a,b,c \neq 0$. The zero set of such a polynomial is a *projective line*.

We can now demonstrate the first nice property of projective space.

Proposition 3.9 *If L_1 and L_2 are projective distinct projective lines, then* $|L_1 \cap L_2| = 1$.

Proof Suppose L_i is the line $a_iX + b_iY + c_iZ = 0$. Points $p \in L_1 \cap L_2$ have homogeneous coordinates (x, y, z) such that

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

a homogeneous system of 2 linear equations in 3 unknowns.

If $(a_2, b_2, c_2) = \lambda(a_1, b_1, c_1)$ for some λ, then L_1 and L_2 are the same line. Thus we may assume that (a_1, b_1, c_1) and (a_2, b_2, c_2) are linearly independent. Thus the matrix

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}$$

has rank 2. It follows that the homogeneous system has a one dimensional solution space. Let (x, y, z) be a nonzero solution. Then all solutions are of the form $(\lambda x, \lambda y, \lambda z)$ for some $\lambda \in k$. In other words, $L_1 \cap L_2 = \{[(x, y, z)]\}$.

Suppose L is the line in $\mathbb{P}_2(k)$ given by the equation $aX + bY + cZ = 0$ a point p with homogeneous coordinates $(x, y, 1)$ is on L if and only if $aX + bY + c = 0$. Thus when we identify affine space $\mathbb{A}_2(k)$ with $U = \{p \in \mathbb{P}_2(k) : p$ has homogeneous coordinates (x, y, z) with $z \neq 0\}$. Then the points on $L \cap \mathbb{A}_2(k)$ are exactly the points on the affine line $aX + bY + c = 0$.

Start with an affine line $aX + bY + c = 0$ where at least one of $a, b \neq 0$ and let L be the projective line $aX + bY + cZ = 0$. If $ax + by = 0$, then $(x, y, 0)$ is also a point on L. It is easy to see that any such point is of the form $(\lambda b, -\lambda a, 0)$. Thus L contains a unique point in $\mathbb{P}_2(k) \setminus \mathbb{A}_2(k)$. We consider this the *point at infinity* on L.

Note that $[(b, -a, 0)]$ is also a point on the line $aX + bY + dZ = 0$ for any d. Thus we have shown that "parallel" affine line intersect at the point at infinity.

What happens when $a = b = 0$. In this case we just have the line $Z = 0$. As this line contains no points of U, we think of it as the line at infinity.

Proposition 3.10 *If $p_1, p_2 \in \mathbb{P}_2(k)$ are distinct points, there is a unique line L with $p_1, p_2 \in L$.*

Proof Suppose $p_i = [(x_i, y_i, z_i)]$. We look for a line L with equation $aX + bY + cZ = 0$ such that

$$\begin{aligned} ax_1 + by_1 + cz_1 &= 0 \\ ax_2 + by_2 + cz_2 &= 0 \end{aligned}$$

Thus we are looking for a, b, c that are nontrivial solutions to the homogeneous system of equations

$$\begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Since $(x_1, y_1, z_1) \not\sim (x_2, y_2, z_2)$, the matrix has rank 2. Thus there is a nontrivial solution (a, b, c) and every solution is of the form $(\lambda a, \lambda b, \lambda c)$ for some $\lambda \in k$. Any two equations of the form $\lambda aX + \lambda bY + \lambda cZ = 0$ define the same line. Thus $aX + bY + cZ = 0$ is the unique line through p_1 and p_2.

The previous proofs leads to an interesting observation.

Proposition 3.11 *Let L_i be the line $a_iX + b_iY + c_iZ = 0$ for $i = 1, 2$. Then $L_1 = L_2$ if and only if $(a_1, b_1, c_1) \sim (a_2, b_2, c_2)$.*

Proof Clearly if $(a_2, b_2, c_2) = \lambda(a_1, b_1, c_1)$, then $L_1 = L_2$. On the other hand, if $(a_1, b_1, c_1) \not\sim (a_2, b_2, c_2)$, then the matrix

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}$$

has rank 2, and $|L_1 \cap L_2| = 1$. In particular the lines are distinct.

Corollary 3.12 *Let $L_{a,b,c}$ be the line with equation $aX + bY + cZ = 0$. The map $[(a, b, c)] \mapsto L_{a,b,c}$ is a bijection between $\mathbb{P}_2(k)$ and $\{L : L \subset \mathbb{P}_2(k)$ is a line$\}$.*

Projective Transformations of $\mathbb{P}_n(k)$

Recall that $T : k^n \to k^n$ is a *linear transformation* if $T(a\vec{x}+b\vec{y}) = aT(\vec{x})+bT(\vec{y})$ for all $a, b \in k$ and $\vec{x}, \vec{y} \in k^n$. Let $\mathrm{GL}_n(k)$ be the set of invertible $n \times n$ matrices with entries from k. If T is a linear transformation of k^n, then there is an $n \times n$ matrix A such that $T(\vec{x}) = A\vec{x}$ for all $\vec{x} \in k^n$. If T is invertible, then $A \in \mathrm{GL}_n(k)$.

If $T : k^{n+1} \to k^{n+1}$ is a linear transformation, then $T(0) = 0$ and $T(\lambda\vec{x}) = \lambda(T(\vec{x}))$. Thus if $\vec{x} \sim \vec{y}$, then $T(\vec{x}) \sim T(\vec{y})$. Moreover, if T is invertible and $T(\vec{y}) = \lambda T(\vec{x})$, then $\vec{y} = \lambda\vec{x}$. Thus

$$\vec{x} \sim \vec{y} \Leftrightarrow T(\vec{x}) = T(\vec{y})$$

for invertible linear T. In particular if $T : k^{n+1} \to k^{n+1}$ is an invertible linear transformation, then there map

$$[\vec{x}] \mapsto [T(\vec{x})]$$

is a well-defined function from $\mathbb{P}_n(k)$ to $\mathbb{P}_n(k)$. We call such functions *projective transformations.*

We first look at the case $n = 1$.

Suppose

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is invertible, then

$$T([(x, y)]) = [(ax + by, cx + dy)]$$

is a projective transformation of $\mathbb{P}_1(k)$. We can think of $\mathbb{P}_1(k)$ as k (i.e. the points with homogeneous coordinates $[(x,1)]$ together with an additional point $\infty = [(1,0)]$.

How does T act on k?

$$T([(x,1)]) = [(ax+b, cx+d)] = \left[\left(\frac{ax+b}{cx+d}\right), 1\right]$$

if $cx+d \neq 0$. Thus we can view T as extending the function $x \mapsto \frac{ax+b}{cx+d}$ on $k \setminus \{\frac{-d}{c}\}$. Since A is invertible, $a(\frac{-d}{c}) + b \neq 0$. Thus

$$T\left(\left[\left(\frac{-d}{c}, 1\right)\right]\right) = \left[\left(\frac{-da}{c} + b, 0\right)\right] = [(1,0)].$$

Also

$$T([(1,0)] = [a,c].$$

There are two cases to consider.

<u>case 1</u>: $c = 0$

In this case $T(x) = \frac{ax+b}{d}$ for $x \in k$ and $T(\infty) = \infty$.

<u>case 2</u>: $c \neq 0$

In this case $T(x) = \frac{ax+b}{cx+d}$ for $x \in k \setminus \{\frac{-d}{c}\}$, $T(\frac{-d}{c}) = \infty$ and $T(\infty) = [(a,c)] = [(\frac{a}{c})]$.

In case 1, T extends an affine transformation of $\mathbb{A}_(k)$. This is not true in case 2.

We prove two results about projective transformations of $\mathbb{P}_2(k)$.

Proposition 3.13 *If $L \subseteq \mathbb{P}_2(k)$ is a line, then $T(L)$ is a line.*

Proof Let L be the line $aX + bY + cZ = 0$. Let $A \in \mathrm{GL}_3(k)$ be a matrix such that $T([\vec{x}]) = [A\vec{x}]$. Then L is the set of $[(x,y,z)]$ where x,y,z is a solution to

$$\begin{pmatrix} a & b & c \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Let

$$\begin{pmatrix} \alpha & \beta & \gamma \end{pmatrix} = \begin{pmatrix} a & b & c \end{pmatrix} A^{-1}.$$

Then

$$\begin{pmatrix} \alpha & \beta & \gamma \end{pmatrix} A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Thus $T([(x,y,z)])$ is a point on the line L_1 with equation $\alpha X + \beta Y + \gamma Z = 0$.

On the other hand if $\alpha x + \beta y + \gamma z = 0$, then

$$\begin{pmatrix} a & b & c \end{pmatrix} A^{-1} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \alpha & \beta & \gamma \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Thus $T^{-1}([x,y,z])$ is on L. Thus $T(L) = L_1$ and $T^{-1}(L_1) = L$.

Proposition 3.14 *If $L_1, L_2 \subseteq \mathbb{P}_2(k)$ are projective lines, there is a projective transformation taking L_1 to L_2.*

Proof By composing projective transformations and their inverses, it's enough to show that if L is the line $aX + bY + cZ = 0$, then there is a projective transformation taking L to the line $X = 0$.

The idea is similar to Proposition 2.8. We let $U = aX + bY + cZ$ and define $V = a_2X + b_2Y + c_2Z$, $W = a_3X + b_3Y + c_3Z$ in such a way that we insure the transformation is invertible. Choose (a_2, b_2, c_2) and (a_3, b_3, c_3) such that (a, b, c),(a_2, b_2, c_2) and (a_3, b_3, c_3) are linearly independent and let

$$A = \begin{pmatrix} a & b & c \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}.$$

Then $(a\ b\ c)A^{-1} = (1\ 0\ 0)$ and, arguing as above, we have transformed the line L to $X = 0$.

Homogenizing Affine Equations

Recall that we identify $\mathbb{A}_2(k)$ with the subset $U = \{[(x, y, 1)] : x, y \in k\}$ of $\mathbb{P}_2(k)$.

If C is a projective curve given by $F(X, Y, Z) = 0$, then

$$C \cap \mathbb{A}_2(k) = \{(x, y) : f(x, y) = 0\}$$

where $f(X, Y) = F(X, Y, 1)$. Thus $C \cap \mathbb{A}_2(k)$ is an affine curve. On the other hand, suppose $f(X, Y) \in k[X, Y]$ has degree d. For $i \leq d$, let f_i be the sum of all monomials in f of degree i. Then f_i is homogeneous of degree i and $f = \sum_{i=0}^{d} f_i$. Let

$$F(X, Y, Z) = \sum_{i=0}^{d} f_i(X, Y)Z^{d-i}.$$

Then $f(X, Y) = F(X, Y, 1)$, F is homogeneous of degree d, and the affine curve $f(X, Y) = 0$ is the affine part of the projective curve $F(X, Y, Z) = 0$.

For example if $f(X, Y) = Y - X^2$, then $F(X, Y, Z) = YZ - X^2$. This trick will allow us to use projective methods to study affine equations.

Solving equations in $\mathbb{P}_1(k)$

Suppose $F(X, Y) \in k[X, Y]$ is a homogeneous polynomial of degree d. We would like to study the zeros of F in $\mathbb{P}_1(k)$. Recall that we can view $\mathbb{P}_1(k)$ as $\mathbb{A}_1(k)$ with an additional point ∞. Points in $\mathbb{A}_1(k)$ have homogeneous coordinates $(x, 1)$, while infinity has homogeneous coordinates $(1, 0)$.

Let

$$F(X, Y) = \sum_{i=0}^{d} a_i X^i Y^{d-i}$$

and let

$$f(X) = F(X, 1) = \sum_{i=0}^{d} a_i X^i.$$

The solutions to $F(X, Y) = 0$ in $\mathbb{A}_1(k)$ are points $p = [(x, 1)]$ where x is zero of f. We let m_p be the multiplicity of f at x.

The point $\infty = [(1, 0)]$ is a solution if and only if $a_d = 0$ and Y divides $F(X, Y)$. Let k be the maximal such that Y^k divides $F(X, Y)$. We call $k = m_\infty$ the *multiplicity* at ∞.

Proposition 3.15 *If $F(X, Y)$ is homogeneous of degree d, $p_1, \ldots, p_k$ are the distinct zeros of F in $\mathbb{P}_1(k)$ and m_{p_i} is the multiplicity at p_i, then $\sum m_{p_i} \leq d$. If k is algebraically closed, then $\sum m_{p_i} = d$.*

Proof We can write $F(X, Y) = Y^m G(X, Y)$ where G is homogeneous of degree $d - m$ and Y does not divide G. If $m > 0$, then ∞ is a zero of F of multiplicity m. The affine zeros of F are the zeros of $g(X) = G(X, 1)$ a polynomial of degree $d - m$. We know that the sum of the multiplicities of zeros of g is at most $d - m$, with equality holding if k is algebraically closed.

Intersecting Projective Lines and Curves

At the beginning of this section we said that one of our reasons for passing to projective space was to show that lines intersect curves of degree d in exactly d points, when we count points correctly. Our next goal is to show that we haven't introduced too many points of intersection.

Theorem 3.16 *Suppose $F(X, Y, Z)$ is a homogeneous polynomial of degree d and C is the projective curve $F = 0$. Let $L \subseteq \mathbb{P}_2(k)$ be a projective line such that $L \not\subseteq C$. Then $|L \cap C| \leq d$.*

Proof Let L be the line $aX + bY + cZ = 0$. We will assume that $c \neq 0$ (the other cases are similar). If (x, y, z) are the homogeneous coordinates for a point of $L \cap C$, then $F(x, y, \frac{-ax-by}{c}) = 0$. Let $G(X, Y)$ be the polynomial obtained when we substitute $Z = \frac{-aX-bY}{c}$ into $F(X, Y, Z)$. If $F(x, y, z) = 0$ and $aX + bY + cZ = 0$, then $G(x, y) = 0$. Moreover, if $G(x, y) = 0$, then $F(x, y, \frac{-ax-by}{c}) = 0$.

Claim Either $G(X, Y) = 0$ or $G(X, Y)$ is homogeneous of degree d.

Suppose $F(X, Y, Z) = \sum_{n=0}^{m} a_n X i_n Y_{j_n} Z^{k_n}$ where $i_n + j_n + k_n = d$ for all n. When we expand $(-aX - bY)^{k_n}$ we get a homogeneous polynomial of degree k_n. Thus each monomial occurring in $a_n X^{i_n} Y^{j_n} (-aX - bY)^{k_n}$ has degree d. When we add up all of these terms, either they all cancel out and we get $F = 0$ or they don't and F is homogeneous of degree d.

Suppose $G(X, Y) = 0$. If $ax + by + cz = 0$, $F(x, y, z) = 0$. Thus the line L is contained in C.

Suppose $G(X,Y)$ is homogeneous of degree d. If $[(x,y,z)] \in C \cap L$, then $[(x,y)]$ is a zero of G in $\mathbb{P}_1(k)$. On the other hand, if $[(x,y)]$ is a zero of G in $\mathbb{P}_1(k)$, then $[(x,y,\frac{-ax-by}{c})]$ is a point on $C \cap L$. The map

$$[(x,y)] \mapsto \left[\left(x,y,\frac{-ax-by}{c}\right)\right]$$

is well-defined and one-to-one. Thus there is a bijection between points on $C \cap L$ and zeros of G. We know that there are at most d distinct zeros to G in $\mathbb{P}_1(k)$. Thus $|C \cap L| \leq d$.

The proof also gives us one idea of how we could assign multiplicities to make sure that we get the right number of points of intersection. Suppose $p = [(x,y,z)] \in C \cap L$. We let m_p be the multiplicity of $[(x,y)]$ as a zero of G. If k is algebraically closed we know that $\sum m_p = d$.

We will return to this idea later. For now let's look at two examples. Let C be the curve $X^2 + Y^2 - Z^2 = 0$ in $\mathbb{P}_2(\mathbb{C})$. Then $C \cap \mathbb{A}_2(\mathbb{C})$ is the affine circle $X^2 + Y^2 = 1$.

Let L_1 be the line $iX - Y + Z = 0$. This is the projective version of the line $Y = iX + 1$. We have already noticed that $\mathbb{A}_2(\mathbb{C}) \cap C \cap L_1 = \{(0,1)\}$.

Let

$$G(X,Y) = X^2 + Y^2 - (Y - iX)^2 = 2X^2 + 2iXY = 2X(X + iY).$$

Then G has two zeros in $\mathbb{P}_2(\mathbb{C})$, $[(0,1)]$ and $[(-i,1)]$, each with multiplicity 1. Thus $C \cap L_1$ has two points of intersection $[(0,1,1)]$ and $[(-i,1,0)]$. Each point of intersection has multiplicity one. Note that $[(0,1,1)]$ is the one point in $\mathbb{A}_2(\mathbb{C})$, while $[(-i,1,0)]$ is a point at infinity.

Now let's consider the intersection of C and L_2 given by $iX - Y = 0$. This corresponds to intersecting the affine circle $X^2 + Y^2 = 1$ with the line $Y = iX$. We argued above that there are no points of intersection in $\mathbb{A}_2(\mathbb{C})$. In this case, since the Z-coordinate is zero, we consider

$$G_1(X,Z) = F(X,iX,Z) = Z^2.$$

The equation $G_1(X,Z) = 0$ has a unique zero $[(1,0)]$ in $\mathbb{P}_1(k)$ of multiplicity 2. Thus $C \cap L_2$ has a unique point $[(1,i,0)]$ of multiplicity 2. We will return to this example later and argue that the line L_2 is tangent to C at this point.

Conics in $\mathbb{P}_2(k)$

Suppose $F(X,Y,Z)$ has degree 2. Then

$$F(X,Y,Z) = aX^2 + bY^2 + cZ^2 + dXY + eXZ + fYZ.$$

Let C be the curve $F = 0$.

Proposition 3.17 *Suppose k is a field of characteristic different from 2. There is a projective transformation that transforms C to a curve $X^2 + bY^2 + cZ^2 = 0$.*

Proof We will essentially repeat quickly the main ideas from the Proof of Theorem 2.10.

We first claim that we may assume $a = 1$. If $a = 0$ and either $b, c \neq 0$, then by permuting the variables we may assume $a \neq 0$. If $a = b = c = 0$, then, by permuting the variables, we may assume that $d \neq 0$ and make the transformation $V = Y - X$. This transforms C to the curve

$$dX^2 + dXV + (e - f)XZ + fVZ = 0$$

in the variables X, Y, Z. Now divide through by a to get an equation of the form

$$X^2 + bY^2 + cZ^2 + dXY + eXZ + fYZ = 0.$$

We next claim that we may assume $d = e = 0$. If not we complete the square. Let $U = X + \frac{d}{2}Y$. This transforms C to

$$U^2 + \left(b - \frac{d^2}{4}\right)Y^2 + cZ^2 + eUZ + \left(f - \frac{de}{2}\right)YZ = 0$$

in the variables U, Y, Z. We next let $V = U + \frac{e}{2}$. This transforms the curve above to

$$V^2 + \left(b - \frac{d^2}{4}\right)Y^2 + \left(c - \frac{e^2}{4}\right)Z^2 + \left(f - \frac{de}{2}\right)YZ = 0.$$

Thus we assume our curve has the form

$$X^2 + bY^2 + cZ^2 + dYZ = 0.$$

case 1: $b \neq 0$.

We let $V = Y + \frac{d}{2b}Z$ to complete the square and transform to an equation of the desired form.

case 2: $b = 0$ and $c \neq 0$.

We permute Y and Z to get to case 1.

case 3: $b = c = d = 0$

We are done.

case 4: $b = c = 0$, $d \neq 0$.

We let $V = Z - Y$ and transform to an equation of the form $X^2 + bV^2 + cVZ = 0$ where $b \neq 0$ and back in case 1.

Corollary 3.18 *If k is a field of characteristic different from 2, and C is a conic curve then there is a projective transformation transforming C to one of the following curves:*

i) $X^2 + bY^2 + cZ^2 = 0$ where $b, c \neq 0$.

ii) $X^2 + bY^2 = 0$ where $b \neq 0$.

iii) $X^2 = 0$.

We call conics of type i) nondegenerate.

We can now give a very simple classification of conics in $\mathbb{P}_2(K)$ for algebraically closed K.

Corollary 3.19 *Let K be an algebraically closed field of characteristic different from 2. If C is a conic curve in $\mathbb{P}_2(K)$ there is a projective transformation transforming C to one of the curves:*
i) (circle) $X^2+Y^2-Z^2=0$;
ii) (crossing lines) $X^2-Y^2=0$;
iii) (double line) $X^2=0$.

Proof If we can transform C to $X^2+bY^2+cZ^2=0$ where $a,b,c\neq 0$ we let $V=\sqrt{b}Y$, $W=i\sqrt{Z}$ to transform the curve to $X^2+V^2-W^2=0$.

Similar ideas work for $X^2+bY^2=0$.

By working in $\mathbb{P}_2(K)$ we eliminate the need to look at parallel lines and parabolas. Note that, up to projective transformations, there is only one nondegenerate conic in $\mathbb{P}_2(K)$.

Using this classification we can better understand what happens when a conic C intersects a line L. If C is degenerate, we can of course have $L\subseteq C$. This is impossible if C is nondegenerate.

Corollary 3.20 *Let K be algebraically closed if $C\subseteq\mathbb{P}_2(K)$ is a nondegenerate conic and $L\subseteq\mathbb{P}_2(K)$ is a line, then $L\not\subseteq C$.*

Proof By an projective transformation we may assume that C is given by $X^2+Y^2+Z^2=0$. Let L be the line $aX+bY+cZ=0$. Without loss of generality suppose $c=0$. If (x,y,z) are the homogeneous coordinates for a point on $C\cap L$, then $z=\frac{-ax-by}{c}$ and $x^2+y^2+\frac{(ax+by)^2}{c^2}=0$ and (x,y) is a zero of

$$(a^2+c^2)X^2+(b^2+c^2)Y^2+2abXY=0.$$

This polynomial is not identically zero (in order to have $2ab=0$ we must have $a=0$ or $b=0$, but $c\neq 0$, so we will either get $a^2+c^c\neq 0$ or $b^2+c^2\neq 0$). Since the polynomial has at most two solutions in $\mathbb{P}_2(K)$, $|C\cap L|\leq 2$.

The proof we gave of Proposition 3.17 was very "low tech". There is another way to look at this for those of you who know some more linear algebra.

Suppose F has degree 2. Since our field doesn't have characteristic 2 we can write F as

$$aX^2+bY^2+cZ^2+2dXY+2eXZ+2fYZ$$

Thus we can think of $F(X,Y,Z)=0$ as the matrix equations

$$\begin{pmatrix} X & Y & Z \end{pmatrix}\begin{pmatrix} a & d & e \\ e & b & f \\ d & f & c \end{pmatrix}\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}.$$

We would like to find a projective transformation making to transform this system of equations to one of the form

$$\begin{pmatrix} X & Y & Z \end{pmatrix} \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}.$$

Note that the matrix of coefficients is symmetric. If B is an $n \times m$ matrix with entries $b_{i,j}$, then the *transpose* of B is B^t the $m \times n$ matrix $(c_{i,j})$ where $c_{i,j} = b_{j,i}$. We say that a matrix B is symmetric if and only if $B = B^t$. If our transformation is given by

$$T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = A \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

then

$$T((x\ y\ z)) = \left(A \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right)^t = (x\ y\ z)A^t.$$

Thus we are looking for $A \in \mathrm{GL}_3(k)$ such that

$$A^t \begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix} A$$

is a diagonal matrix. The existence of such a matrix is a standard theorem from linear algebra.

Theorem 3.21 *Suppose k is a field with characteristic different from 2. If B is a symmetric $n \times n$ matrix, there is an invertible $n \times n$ matrix A such that A^tBA is a diagonal matrix.*

The argument we gave above can be used to give an elementary proof of Theorem 3.21.

Projective Parameterizations

We assume that k is a field of characteristic different from 2. In §2 we showed that

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

is a rational parameterization of the circle in $\mathbb{A}_2(k)$. There were a couple of problems with this parameterization:

i) it misses the point $(0, -1)$;

ii) it is undefined when $t^2 = -1$.

Both of these problems disappear in $\mathbb{P}_2(k)$.

Let C be the circle in $\mathbb{P}_2(k)$ with equation $X^2+Y^2-Z^2=0$.

We think of t as being $\frac{u}{v}$. Then

$$\rho(t)=\left(\frac{2uv}{v^2+u^2},\frac{v^2-u^2}{v^2+u^2}\right).$$

This gives us an idea of how to define a projective transformation. We define $f:\mathbb{P}_1(k)\to C$ by

$$[(x,y)]\mapsto[(2xy,y^2-x^2,y^2+x^2)].$$

We first argue that f is well-defined. First we see that f preserves $\sim$-equivalence classes since

$$[(\lambda x,\lambda y)]\mapsto[(2\lambda^2xy,\lambda^2(y^2-x^2),\lambda^2(y^2+x^2)]=[(2xy,y^2-x^2,y^2+x^2)].$$

Moreover if $(2xy,y^2-x^2,y^2+x^2)=(0,0,0)$, then $x=y=0$. Thus f is a well-defined function $f:\mathbb{P}_1(k)\to\mathbb{P}_2(k)$.

Also notice that

$$(2xy)^2+(y^2-x^2)^2-(y^2+x^2)^2=0.$$

Thus the image of f is contained in C.

We will argue that f is onto, by considering it's inverse. We want our inverse function to take $[(2xy,y^2-x^2,y^2+x^2)]$ to $[(x,y)]$. One possibility is to take

$$g_1([(a,b,c)])=[(c-b,a)]).$$

Then

$$g_1([(2xy,y^2-x^2,y^2+x^2)]=[(2x^2,2xy)]=[(x,y)].$$

It is easy to see that if $(a_1,b_1,c_1)\sim(a,b,c)$, then $(c-b,a)\sim(c_1-b_1,a)$. Thus g_1 is a well-defined function as long as we don't have $a=0$ and $b=c$. In particular g_1 maps $C\setminus\{[(0,1,1)]\}$ to $\mathbb{P}_1(k)$.

We argue that this is the inverse to f. If $[(a,b,c)]\in C\setminus[(0,1,1)]$, then

$$\begin{aligned}f([(c-b,a)] &= [(2(c-b)a,a^2-(c-b)^2,a^2+(c-b)^2)]\\ &= [(2(c-b)a,a^2-b^2-c^2+2cb,a^2+c^2+b^2-2cb)]\\ &= [(2(c-b)a,2cb-2b^2,2c^2-2cb)],\text{ since }a^2=c^2-b^2\\ &= [(a,b,c)],\text{ factoring out }c-b\end{aligned}$$

If we extend g_1 to $g:C\to\mathbb{P}_1(k)$ by

$$g(p)=\begin{cases}g_1(p) & \text{if } p\neq\{[(0,1,1)]\\ [(0,1)] & p=\{[(0,1,1)]\end{cases},$$

we see that f and g are inverses. Thus $f:\mathbb{P}_1(k)\to C$ is a bijection between $\mathbb{P}_1(k)$ and C.

Exercise 3.22 Alternatively we could have defined $g_2:C\setminus\{[(0,-1,1)]\}\to\mathbb{P}_1(k)$ by $g_2([(a,b,c)])=(a,b+c)$.

a) Prove that $f \circ g_2$ is the identity on $C \setminus \{[(0,-1,1)]\}$.
b) Prove that $g_1 = g_2$ on $C \setminus \{[(0,1,1)],[(0,-1,1)]\}$.

We began this section mentioning two problems with the parameterization we had found for the affine circle. Have these problems been fixed?

i) We worried that $(0,-1)$ was not in the image. We now have

$$f(\infty) = f([(1,0)]) = [(0,-1,0)].$$

So the affine point $(0,-1)$ is in the image.

ii) We worried about $\rho(t)$ if $t^2 = -1$. If $i \in k$ with $i^2 = -1$, then

$$f([i,1]) = [(2i,2,0)] \text{ and } f([-i,1]) = [(-2i,2,0)].$$

This is good since $[(i,1,0)]$ and $[(-i,1,0)]$ are the two points at infinity on C (if k has a square root of -1).

Suppose K is an algebraically closed field and C_1 is a nondegenerate conic. There is a projective transformation T of $\mathbb{P}_2(K)$ taking the circle C to C_1. Then $T \circ f$ is a parameterization of C_1. We will later need to consider the exact form of this parameterization. Suppose $A \in \mathrm{GL}_3(k)$ and $T([\vec{x}]) = [A\vec{x}]$. Then our parameterization is given by

$$(x,y,z) \mapsto A \begin{pmatrix} 2xy \\ y^2 - x^2 \\ y^2 + x^2 \end{pmatrix} = \begin{pmatrix} \alpha(x,y) \\ \beta(x,y) \\ \gamma(x,y) \end{pmatrix}$$

where α, β, γ are homogeneous of degree 2.

We can also find parameterizations directly.

Exercise 3.23 Suppose k is a field of characteristic different from 2. Suppose C is a nondegenerate conic over k and there is $p \in C \cap \mathbb{P}_2(k)$. Find a parameterization of C as follows:

i) Find a bijection $f : \mathbb{P}_1(k)$ to the set of lines through p.

ii) If L is a line through p let

$$g(L) = \begin{cases} q & \text{if } q \in L \cap C \setminus \{p\} \\ p & \text{if } L \text{ is tangent at } p \end{cases}.$$

Find an equations for $g \circ f$ and prove this parameterizes C.

Intersecting Conics with Curves

Theorem 3.24 *Suppose K is an algebraically closed field of characteristic different from 2, $C \subseteq \mathbb{P}_2(K)$ is a nondegenerate conic and $D \subseteq \mathbb{P}_2(K)$ is a curve of degree D. Either $C \subseteq D$ or $|C \cap D| \leq 2d$.*

Proof We have a projective parameterization of C. By the arguments above, there are homogeneous polynomials a, b, c of degree 2 such that

$$([x,y]) \mapsto [(a(x,y), b(x,y), c(x,y))]$$

is a parameterization of C.

Suppose D is the set of solutions to the degree d polynomial $F(X,Y,Z)=0$. Then points of $C \cap D$ are in bijective correspondence with points $[(x,y)] \in \mathbb{P}_1(K)$ such that

$$F(a(x,y),b(x,y),c(x,y))=0.$$

Let $g(X,Y)=F(a(X,Y),b(X,Y),c(X,Y))$. Since F is homogeneous of degree d and a,b,c are homogeneous of degree 2, g is either 0 or of degree $2d$.

If $g=0$, then $[(a(x,y),b(x,y),c(x,y))] \in D$ for all $[(x,y)] \in \mathbb{P}_1(K)$ and $C \subseteq D$. If g has degree $2d$, then there are at most $2d$ points $p \in \mathbb{P}_1(K)$ such that $g(p)=0$ and $[(a(x,y),b(x,y),c(x,y))] \in D$. Thus $|C \cap D| \leq 2d$.

Once again we could assign multiplicities by letting m_p be the multiplicity of p as a zero of g. Then we would have $\sum m_p = 2d$. One of the most important result in the subject is that if we have curves C and D of degrees d_1 and d_2 such that $C \cap D$ is finite, then if we assign multiplicites correctly we have $d_1 d_2$ points of intersection. We will return to this idea later.

While two points determine a line, 5 points (in general position) determine a conic.

Corollary 3.25 *If $p_1,\ldots,p_5 \in \mathbb{P}_2(k)$, there is a conic C with $p_1,\ldots,p_5 \in C$. If no four of $p_1,\ldots,p_5$ are colinear, then C is unique. If no three are colinear C is nondegenerate.*

Proof We first show that there is at least one conic. Let $(x_{i,1},\ldots,x_{i,5})$ be homogeneous coordinates for p_i. We are looking for $a_1,\ldots,a_6$ such that

$$\begin{pmatrix} x_{1,1}^2 & x_{1,2}^2 & x_{1,3}^2 & x_{1,1}x_{1,2} & x_{1,1}x_{1,3} & x_{1,2}x_{1,3} \\ & & & \vdots & & \\ x_{5,1}^2 & x_{5,2}^2 & x_{5,3}^2 & x_{5,1}x_{5,2} & x_{5,1}x_{5,3} & x_{5,2}x_{5,3} \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_6 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Since this is a homogeneous system of 5 linear equations in 6 unknowns it has a nonzero solution $(a_1,\ldots,a_6)$. Let $F=a_1X^2+a_2Y^2+a_3Z^2+a_4XY+a_5XZ+a_6YZ$. Then $F(p_i)=0$ for $i=1,\ldots,6$.

Suppose no four of $p_1,\ldots,p_5$ are collinear. We next must show uniqueness. Suppose $K \subseteq k$ is algebraically closed. It is enough to show that there is a unique conic in K containing $p_1,\ldots,p_5$. Suppose C_1 and C_2 are conics such that $p_1,\ldots,p_5 \in C_1 \cap C_2$.

case 1 C_1 and C_2 are non-degenerate.

Then by the previous theorem $C_1 \subseteq C_2$ and $C_2 \subseteq C_1$.

case 2 One C_i is degenerate and the other is not.

Suppose C_1 in nondegenerate and C_2 is degenerate. We know that C_1 does not contain a line and that C_2 is either two crossing lines or a single line. Since $p_1,\ldots,p_5$ are non collinear, C_2 must be two crossing lines.

case 3 C_1 and C_2 are degenerate.

Since the points $p_1, \dots, p_5$ are not collinear, $C_1 = L_0 \cup L_1$ and $C_2 = L_2 \cup L_3$ where each L_i is a line $L_0 \neq L_1$ and $L_2 \neq L_3$.

Since $\{p_1, \dots, p_5\} \subseteq L_0 \cup L_1$. At least one of the lines must contain three of the points. Suppose $p_1, p_2, p_3 \in L_0$. Since $p_1, p_2, p_3 \in L_2 \cup L_3$, at least one of the lines must contain at least two of those points. Say $p_1, p_2 \in L_3$. But then $p_1, p_2 \in L_0 \cap L_3$. Since two points determine a line we must have $L_0 = L_2$.

Thus we may assume that $C_1 = L_0 \cap L_1$, $C_2 = L_0 \cap L_2$ and $p_1, p_2, p_3 \in L_0$. Since no four points are collinear we must have $p_4, p_5 \in L_1$ and $p_4, p_5 \in L_2$. But then $L_1 = L_2$ and $C_1 = C_2$. Thus there is a unique conic through $p_1, \dots, p_5$.

If C is degenerate then C is either a line or the union of two lines. In either case at least three of the points are on a line. Thus if no three points are collinear C is nondegenerate.

The first part of the proof has a simple generalization.

Exercise 3.26 Suppose $p_1, \dots, p_N \in \mathbb{P}_2(k)$ where $N \leq \frac{d^2+3d}{2}$. Then there is a curve C of degree d with $p_1, \dots, p_N \in C$.

4 Irreducible Components

Let K be an algebraically closed field.

Recall that $C \subseteq \mathbb{A}_2(K)$ is a *curve* if there is a polynomial $f \in K[X,Y]$ such that

$$C = V(f) = \{(x,y) \in \mathbb{A}_2(K) : f(x,y) = 0\}.$$

Definition 4.1 We say that a curve $C \subseteq \mathbb{A}_2(k)$ is *reducible* if there are curves $C_1, C_2 \subseteq C$ such that $C_1, C_2 \neq C$ and $C = C_1 \cup C_2$. Otherwise we say that C is *irreducible.*

For example the curve $X^2 - Y^2 = 0$ is reducible since

$$X^2 - Y^2 = (X - Y)(X + Y)$$

and $V(f)$ is the union of the two lines $X = Y$ and $X = -Y$. In general, if $f = \prod_{i=1}^n g_i$, then $V(f) = V(g_1) \cup \ldots \cup V(g_n)$. But the $V(g_i)$ may not be distinct. For example $V(X^2) = V(X) \cup V(X) = V(X)$.

One of our goals of this section is to prove that any curve can be written as a finite union of irreducible curves in, essentially, a unique way. Clearly, to fully understand how to decompose a curve into irreducible components, we will need to understand factoring in $K[X,Y]$.

Unique Factorization Domains

We will prove that if k is a field, then the n-variable polynomial ring $k[X_1, \ldots, X_n]$ is a unique factorization domain.

We begin by recalling some of basic definitions and results from Math 330.

Definition 4.2 Let D be an integral domain.

We say that $u \in D$ is a *unit* if there is $v \in D$ such that $uv = 1$. We write $v = \frac{1}{u}$.

If $a \in D$ is nonzero and not a unit, we say that a is *irreducible* if whenever $b, c \in D$ and $a = bc$ then either a is a unit or b is a unit.

We say that $a, b \in D$ are *associates* if there is a unit u such that $a = ub$.

For example, if k is a field and $D = k[X]$, then the units are the nonzero elements of k, $f \in k[X]$ is irreducible if and only if deg $f > 0$ and f can not be factored as the product of two polynomials of lower degree, and f, g are associates if and only if there is $c \in k$ such that $f = cg$.

Lemma 4.3 *If D is an integral domain and $p \in D$ is irreducible, then $(p) = \{a \in D : p$ divides $a\}$ is a prime ideal (i.e., if $a, b \in D$ and p divides ab then p divides a or p divides b.*

Definition 4.4 If D is a domain, we say that D is a *Unique Factorization Domain* (or UFD) if:

i) if $f \in D$ is nonzero and not a unit, then there are irreducible elements $g_1, \ldots, g_n \in D$ such that $f = g_1 g_2 \cdots g_n$, and

ii) if $p_1, \ldots, p_n, q_1, \ldots, q_m \in D$ are irreducible, and $p_1 \cdots p_n = q_1 \cdots q_m$, then $n = m$ and there is $\sigma \in S_n$ such that p_i is an associate of $q_{\sigma(i)}$ for $i = 1, \ldots, n$. In other words, we can reorder the q_i so that p_i and q_i are associates for all i.

There are two important examples of UFDs that you have encountered before.

Theorem 4.5 *i) $\mathbb{Z}$ is a UFD.*
ii) If k is a field, then $k[X]$ is a UFD.

The following theorem will allow us to construct more complicated UFDs.

Theorem 4.6 *If D is a UFD, then the polynomial ring $D[X]$ is a UFD.*

We can always identify the polynomial ring $D[X_1, \ldots, X_n]$ with $D[X_1, \ldots, X_{n-1}][X_n]$. Thus if D is a UFD, the polynomial ring in n-variables over D is as well.

Corollary 4.7 *i) $\mathbb{Z}[X_1, \ldots, X_n]$ is a UFD.*
ii) If k is a field, then $k[X_1, \ldots, X_n]$ is a UFD.

Definition 4.8 A nonconstant polynomial $f(X) = \sum_{n=0}^{d} a_n X^n \in D[X]$ is *primitive* if the only common divisors of $a_0, \ldots, a_n$ are units.

For example, let $D = k[X]$ and consider polynomials in $D[Y] = k[X, Y]$. The polynomial

$$X^2Y + 2XY^2$$

is not primitive since X is a common nonunit divisor of X^2 and $2X^2$. The polynomial

$$X^2 + 2XY + (X+1)Y^2$$

is primitive since $X^2, 2X$ and $(X+1)$ have no common nonunit divisors. The polynomial $2X^2Y + 4Y^2$ is primitive since the only common divisors of $2X^2$ and 4 are units.

If $f = \sum_{n=0}^{d} a_n X^n$ is not primitive, we can find a nonunit c such that for each n, $a_n = cb_n$ for some $b_n \in D$. Thus

$$f = c \sum_{n=0}^{d} b_n X^n.$$

Since c is a nonunit, f is reducible.

Note that $2X + 4Y$ is primitive in $\mathbb{Q}[X][Y]$, but not primitive in $\mathbb{Z}[X, Y]$, since 2 is a unit in $\mathbb{Q}$ but not in $\mathbb{Z}$.

Proposition 4.9 *If $a \in D$ and a divides a nonconstant $g \in D[X]$, then a divides all of the coefficients of D. In particular, if g is primitive and $a \in D$ is a nonunit, then a does not divide g.*

Proof Suppose

$$g = a\left(\sum_{n=0}^{d} b_n X^n\right) = \sum_{n=0}^{d} ab_n X^n.$$

Then a divides each coefficient of g.

Lemma 4.10 *For any nonconstant $f \in D[X]$ we have $f(X) = cg(X)$ where $c \in D$ and $g(X)$ is primitive. Moreover if we also have $f(X) = dh(X)$, where $d \in D$ and h is primitive, then there is a unit $u \in D$ such that $h = gu$ and $d = \frac{c}{u}$. Thus c and g are "unique up to units".*

We call c the content *of g.*

Proof

(existence) Suppose

$$f = \sum_{n=0}^{d} a_n X^n.$$

Since D is a unique factorization domain, we can factor each nonzero nonunit a_n into a product of irreducibles. Let M_f be the maximal number of irreducible factors occuring in the factorization of some a_n. We prove the lemma by induction on M_f.

If $M_f = 0$, then f is primitive.

If f is primitive, then we are done. Suppose f is not primitive. Then $a_0, \ldots, a_n$ have a common irreducible factor p. Let

$$h = \sum_{n_0}^{d} \frac{a_n}{p} X^n.$$

Since $M_h = M_f - 1$, by induction, there is $c \in D$ and a primitive g such that $h = cg$. But then $f = (pc)g$.

(uniqueness) Suppose $f = cg = dh$ where $c, d \in D$ and g, h are primitive. We claim that $h = ug$ for some unit u. We can factor c as the product of M_c irreducible factors. We will prove this by induction on M_c.

Suppose $p \in D$ is an irreducible factor of c. Since h is primitive, p does not divide all of the coefficients of h. Since it divides all the coefficients of f it must divide d. Similarly, if $p \in D$ is an irreducible factor of d, then p divides c.

If $M_c = 0$, then c is a unit and the above argument shows that d is also a unit. Thus $h = \frac{c}{d}g$ and $\frac{c}{d}$ is a unit.

If $M_c > 0$, then c is not a unit. If p is an irreducible factor of c, then p is also an irreducible factor of d and $\frac{c}{p}g = \frac{d}{p}h$. By induction $h = ug$ for some unit u.

For example, suppose k is a field, $D = k[X]$ and $f = X^2Y + 2XY^2 \in D[Y]$, then $f = X(XY + 2Y^2)$. Since $XY + 2Y^2$ is primitive in $D[Y]$, X is the content of f in $D[Y]$.[1]

[1] Note that we could also consider $D^* = k[Y]$ and $f \in D^*[X]$. In that case $f = Y(X^2 + 2XY)$ so the content is Y. is Y.

Lemma 4.11 (Gauss' Lemma) *If D is a UFD and $f, g \in D[X]$ are primitive, then fg is primitive. Indeed, if $f_1, \ldots, f_m$ are primitive, then so is $\prod f_i$.*

Proof Let

$$f = \sum_{n=0}^{d} a_n X^n$$

and

$$g = \sum_{n=0}^{d_1} b_n X^n$$

be primitive in $D[X]$. Suppose gf is not primitive. There is an irreducible p such that p divides every coefficient of fg. Let s be least such that p does not divide a_s and let t be least such that p does not divide b_t. The X^{s+t}-coefficient of fg is

$$\sum_{i+j=s+t} a_i b_j = \sum_{i<s} a_i b_{s+t-i} + a_s b_t + \sum_{j<t} a_{s+t-j} b_j.$$

By choice of s and t, p divides each a_i in the first sum on the right-hand side and each b_j in the second. Since p divides the whole sum, it must also divide $a_s b_t$. Thus p divides a_s or b_t, a contradiction.

The second claim is now a simple induction.

Exercise 4.12 Suppose D is a UFD and $f, g \in D[X]$. Prove that the content of f times the content of g is (a unit times) the content of fg.

Suppose D is a unique factorization domain. Let

$$K = \left\{ \frac{a}{b} : a, b \in D, b \neq 0 \right\}$$

be the fraction field of D. Then $D[X]$ is a subring of $K[X]$. Of course $F[X]$ is a UFD. We want to consider the relationship between factoring in $D[X]$ and factoring in $K[X]$. The first thing to ask is which elements of $D[X]$ are irreducible in $K[X]$.

Lemma 4.13 *Suppose D is a UFD, K is its fraction field and $f \in D[X]$. Then f is irreducible in $D[X]$ if and only if f is primitive and f is irreducible in $K[X]$.*

Proof

($\Leftarrow$) Suppose there are nonunit $g, h \in D[X]$ such $f = gh$. If $g \in D$ or $h \in D$, then f is not primitive. If neither $g, h \in D$ then deg $f >$ deg g, deg h. Thus f is reducible in $K[X]$.

($\Rightarrow$) We argued above that every irreducible $f \in D[X]$ is primitive, so we need only show it is irreducible in $K[X]$.

Suppose $f = gh$ where deg g, deg $h <$ deg f. We must show that f factors in $D[X]$. Each coefficient of g and h is a quotient of elements of D, by clearing denominators we can find $a \in D$, $g_1, h_1 \in D[X]$ such that $af = g_1 h_1$ and deg $f >$ deg g_1, deg h_1. By Lemma 4.10 we can find $b, c, d \in D$ and primitive polynomials f_1, g_2, h_2 such that

$$f = bf_1, \;\; g_1 = cg_2, \;\text{ and } h_1 = dh_2.$$

Then

$$abf_1 = cdg_2h_2$$

and, by Gauss' Lemma g_2h_2 is primitive. Thus, by Lemma 4.10, there is a unit u, such that

$$abu = cd.$$

But then $bu = \frac{cd}{a}$ and

$$f = bf_1 = \frac{cd}{a} g_2h_2 = bug_2h_2.$$

Thus f factors into $(bug_2)h_2$ in $D[X]$.

Proof of Theorem 4.6

(existence) Suppose D is a UFD. We want show that $D[X]$ is a UFD. We first prove that every nonzero nonunit $f \in D[X]$ can be factored into a product of irreducibles. If deg $f = 0$, then $f \in D$ and we are done since D is a UFD.

Suppose deg $f > 0$. We may view f as an element of $K[X]$, where K is the fraction field of D. Since $K[X]$ is a UFD, we can factor

$$f = \prod_{i=1}^{n}$$

where $g_1, \ldots, g_n$ are irreducible factors in $K[X]$. Let $a_i \in D$ be the product of all of the denominators occuring in coefficients in g_i, let $a = \prod a_i$ and let $h_i = a_i g_i \in D[X]$. Then

$$af = \prod h_i.$$

Since $a_i \in D$ is a unit in K, $h_i = a_i g_i$ is irreducible in $K[X]$. By Lemma 4.10, there are $c, d_1, \ldots, d_n \in D$ and primitive f^*,$h_1^*, \ldots, h_n^* \in D[X]$ such that

$$f = cf^*, h_1 = d_1h_1^*, \ldots, h_n = d_nh_n^*.$$

Note that each h_i^* is irreducible in $K[X]$ and primitive in $D[X]$. Thus by Lemma 4.13, each h_i^* is irreducible in $D[X]$.

But

$$acf^* = \prod_{i=1}^{n} d_i \prod_{i=1}^{n} h_i^*$$

and, by Gauss' Lemma, $\prod h_i^*$ is primitive. Thus, by 4.10 there is a unit $u \in D$ such that

$$acu = \prod_{i=1}^{n} d_i.$$

Then

$$acf^* = acu\prod_{i=1}^{n} h_i^*$$

and

$$f = cf^* = cu\prod_{i=1}^{n} h_i^*.$$

Since D is a UFD, we can factor cu into a product of irreducibles. This gives a factorization of f into a product of irreducibles.

(uniqueness) If deg $f = 0$, then $f \in D$ and we have unique factorization since D is a UFD. Suppose deg $f > 0$. Suppose

$$f = a_1 \cdots a_s p_1 \cdots p_m = b_1 \cdots b_t q_1 \cdots q_n$$

are two irreducible factorizations of $f \in D[X]$ where a_i, b_j are irreducible elements of D and p_i, q_j are irreducible nonconstant elements of $D[X]$. By Lemma 4.13, each p_i and q_j is irreducible in $K[X]$. In $K[X]$, $\prod p_i$ is a unit times $\prod q_i$. Thus, since $K[X]$ is a UFD, $m = n$ and we can reorder the q_i such that $p_i = \frac{c_i}{d_i} q_i$ for some $c_i, d_i \in D$. But then $d_i p_i = c_i q_i$ and each p_i and q_j is irreducible in $D[X]$ and, hence, primitive. Thus $\frac{c_i}{d_i}$ is a unit in D.

Thus there is a unit $u \in D$ such that

$$\prod_{i=1}^{s} a_i \prod_{i=1}^{m} p_i = u\prod_{i=1}^{n} b_i \prod_{i=1}^{m} p_i.$$

Since D is a UFD, $s = t$ and we can reorder the b_i such that $a_i = w_i b_i$ for some unit. Thus the factorization is unique up to permuting the factors and multiplication by units.

Study's Lemma

We return to trying to understand algebraic curves in $\mathbb{A}_2(K)$.

For the remainder of this section K is an algebraically closed field

Our first goal is to relate divisibility in $K[X, Y]$ with containment of curves. If $f, g \in K[X, Y]$ and f divides g, then there is $h \in K[X, Y]$ such that $g = fh$ and $V(g) = V(f) \cup V(h)$. Thus $V(f) \subseteq V(g)$. We will argue that the converse is also true for irreducible f.

Theorem 4.14 (Study's Lemma) *If $f, g \in K[X, Y]$, f is irreducible and nonconstant, and $V(f) \subseteq V(g)$, then f divides g.*

Note that Study's Lemma may fail if K is not algebraically closed. For example $X^2 + Y^2$ is irreducible in $\mathbb{R}[X, Y]$ and $V(X^2 + Y^2) \subset V(X)$ in $\mathbb{A}_2(\mathbb{R})$, but $X^2 + Y^2$ does not divide X.

Before proving Study's Lemma, we derive some consequences.

Corollary 4.15 *If $f \in K[X, Y]$ is nonconstant, then $V(f) \neq \emptyset$.*

Proof If $V(f) = \emptyset \subseteq V(1)$, then, by Study's Lemma, f divides 1 and f is constant.

Exercise 4.16 Show that $V(f)$ is infinite.

Corollary 4.17 *Let $f \in K[X, Y]$ be nonconstant. The curve $V(f) \subseteq \mathbb{A}_2(K)$ is irreducible, if and only if there is an irreducible polynomial g and $k > 0$ such that $f = g^k$.*

Proof

($\Rightarrow$) Since $K[X, Y]$ is a UFD, we can factor $f = \prod_{i=1}^{n} g_i^{k_i}$ where $g_1, \dots, g_n$ are irreducible and relatively prime. If $n > 1$, then $V(f) = \bigcup V(g_i)$ and $V(g_i) \neq V(f)$, contradicting the irreducibility of f.

($\Leftarrow$) Suppose $f = g^k$ where g is irreducible and $V(f) = V(h_1) \cup V(h_2)$ where $V(h_1) \neq V(h_2)$. Let p be an irreducible factor of h_1. Then $V(p) \subseteq V(h_1) \subset V(f)$. By Study's Lemma, p divides g^k. Since g is irreducible $p = cg$ for some $c \in K$ and $V(p) = V(g) = V(f)$.

Theorem 4.18 *If $f \in K[X, Y]$ is nonconstant and $C = V(f)$, then there are irreducible curves $C_1, \dots, C_n$ such that $C = C_1 \cup \dots \cup C_n$ and $C_i \not\subseteq C_j$ for $i \neq j$. Moreover, if $D \subseteq C$ is an irreducible curve, then $D = C_i$ for some i.*

We call $C_1, \dots, C_n$ the irreducible components *of C.*

Proof We can factor

$$f = \prod_{i=1}^{n} g_i^{k_i}$$

where each g_i is irreducible and g_i and g_j are relatively prime for $i \neq j$. Then

$$C = V(g_1) \cup \dots V(g_n)$$

and, by Corollary 4.17, each $V(g_i)$ is irreducible. By Study's Lemma, $C_i \not\subseteq C_j$ for $i \neq j$.

If $D \subseteq C$ is an irreducible curve, there is an irreducible $h \in K[X, Y]$ such that $D = V(h)$. Since $V(h) \subseteq V(f)$, h divides f. Since $K[X, Y]$ is a unique factorization domain, $h = cg_i$ for some $i \leq n$ and $c \in K$. Thus $V(h) = V(g_i)$.

We still must prove Study's Lemma. The proof will use a generalization of Theorem 1.22 on resultants. Suppose D is an integral domain and $f, g \in D[X]$ are nonzero polynomials. If $f = a_n X^n + a_{n-1} X^{n-1} + \dots a_0$ and $g = b_m X^m + b_{m-1} X^{m-1} + \dots + b_0$ where $a_n, b_m \neq 0$, we can still form the resultant

$$R_{f,g} = \begin{vmatrix} a_0 & a_1 & \dots & \dots & \dots & a_n & 0 & \dots & \dots & 0 \\ 0 & a_0 & a_1 & \dots & \dots & \dots & a_n & 0 & \dots & 0 \\ & & & \ddots & \ddots & & & & & \\ 0 & \dots & \dots & 0 & a_0 & a_1 & \dots & \dots & \dots & a_n \\ b_0 & b_1 & \dots & \dots & b_m & 0 & \dots & \dots & \dots & 0 \\ 0 & b_0 & b_1 & \dots & \dots & b_m & 0 & \dots & \dots & 0 \\ & & & \ddots & \ddots & & & & & \\ 0 & \dots & \dots & \dots & 0 & b_0 & b_1 & \dots & \dots & b_m \end{vmatrix}.$$

We will have $R_{f,g} \in D$.

We need a mild generalization of Theorem 1.22

Theorem 4.19 *Suppose D is a unique factorization domain and $f, g \in D[X]$ are nonzero polynomials. The following are equivalent:*

i) f and g have a common nonconstant factor in $D[X]$;

ii) There are nonzero $f_1, g_1 \in D[X]$ such that $\deg f_1 < \deg f$, $\deg g_1 < \deg g$ *and*

$$fg_1 = f_1 g.$$

iii) $R_{f,g} = 0$.

Exercise 4.20 Show that the proof of Theorem 1.22 can be modified to prove Theorem 4.19. [Note: The proof of ii) ⇔ iii) should work for any integral domain D.]

Proof of Study's Lemma

We have $f, g \in K[X,Y]$, f is irreducible and $V(f) \subseteq V(g)$. Without loss of generality assume that Y occurs in some monomial of f (otherwise we work with X instead of Y). We can write

$$f = \sum_{i=0}^{n} a_i Y^i$$

and

$$g = \sum_{i=0}^{m} b_i Y^i$$

where $a_i, b_i \in K[X]$ for all i, and $a_n, b_m \neq 0$.

Suppose for purposes of contradiction, that f does not divide g. Since f is irreducible, f and g have no common nonconstant factors. Thus by Theorem 1.22, $R_{f,g} \in K[X]$ is nonzero. Since K is algebraically closed K is infinite. The polynomials $a_n, b_m, R_{f,g} \in K[X]$ are nonzero, thus we can find $c \in K$ such that $a_n(c) \neq 0$, $b_m(c) \neq 0$ and $R_{f,g}(c) \neq 0$.

Consider the polynomials

$$p(Y) = f(c,Y) = \sum_{i=0}^{n} a_i(c) Y^i$$

and

$$q(Y) = g(c,Y) = \sum_{i=0}^{m} b_i(c) Y^i.$$

Note that p has degree $n > 0$, q has degree m and

$$R_{p,q} = \begin{vmatrix} a_0(c) & a_1(c) & \dots & \dots & \dots & a_n(c) & 0 & \dots & \dots & 0 \\ 0 & a_0(c) & a_1(c) & \dots & \dots & \dots & a_n(c) & 0 & \dots & 0 \\ & & & \ddots & \ddots & & & & & \\ 0 & \dots & \dots & 0 & a_0(c) & a_1(c) & \dots & \dots & \dots & a_n(c) \\ b_0(c) & b_1(c) & \dots & \dots & b_m(c) & 0 & \dots & \dots & \dots & 0 \\ 0 & b_0(c) & b_1(c) & \dots & \dots & b_m(c) & 0 & \dots & \dots & 0 \\ & & & \ddots & \ddots & & & & & \\ 0 & \dots & \dots & \dots & 0 & b_0(c) & b_1(c) & \dots & \dots & b_m(c) \end{vmatrix}$$

Thus $R_{p,q} = R_{f,g}(c) \neq 0$. Since $\deg p > 0$ there is $d \in K$ such that $p(d) = 0$ and, by Theorem 1.22, $q(d) \neq 0$. But then $f(c, d) = 0$ and $g(c, d) \neq 0$ contradicting the fact that $V(f) \subseteq V(g)$.

Irreducible Components in Projective Space

We next consider irreducible components in $\mathbb{P}_2(K)$. If $F \in K[X, Y, Z]$ is homogeneous, then, because $K[X, Y, Z]$ is a UFD, we can factor F into irreducible factors. We first need to notice that the irreducible factors are homogeneous.

Lemma 4.21 *Suppose $F, G, H \in K[X_1, \dots, X_n]$, F is homogeneous and $F = GH$. Then G and H are homogeneous.*

Proof We can write $G = \sum_{i=0}^{n} G_i$ and $H = \sum_{i=0}^{m} H_i$ where G_i and H_i are homogeneous of degree i, $G_n \neq 0$ and $H_m \neq 0$. Let i_0 be least such that $G_{i_0} \neq 0$ and let j_0 be least such that $H_{j_0} \neq 0$. Then

$$F = G_n H_m + \sum_{k=i_0+j_0+1}^{m+n-1} \sum_{i+j=k} G_i H_j + G_{i_0} H_{j_0}$$

and

$$\sum_{i+j=k} G_i H_j$$

is homogeneous of degree k. Since F is homogeneous we must have $n = i_0$ and $m = j_0$. Thus $G = G_n$ and $H = H_m$ are homogeneous.

Our next goal is to prove the projective version of Study's Lemma. We must first investigate the relationship between factoring polynomials in $K[X, Y]$ and factoring their homogenizations in $K[X, Y, Z]$.

Recall that if $f \in K[X, Y]$ has degree d. We can write $f = \sum_{i=0}^{d} f_i$ where $f_i \in K[X, Y]$ is homogeneous of degree i and let

$$F(X, Y, Z) = \sum_{i=0}^{d} f_i Z^{d-i}.$$

Then F is homogeneous of degree d and $f(X, Y) = F(X, Y, 1)$. We call F the *homogenization* of f. Note that Z is not a factor of the homogenization.

Lemma 4.22 *Suppose $f, g, h \in K[X, Y]$ are nonconstant polynomials and $f = gh$. Suppose F, G, H in $K[X, Y, Z]$ are the homogenizations of f, g and h. Then $F = GH$.*

Proof Let

$$f = \sum_{i=0}^{n+m} f_i, \ g = \sum_{i=0}^{n} g_i, \text{ and } h = \sum_{i=0}^{m} h_i$$

where f_i, g_i, h_i are homogeneous of degree i, $g_n, h_m, f_{n+m} \neq 0$. Then

$$f_k = \sum_{i+j=k} g_i h_j.$$

But

$$G = \sum_{i=0}^{n} g_i Z^{n-1}, \ H = \sum_{i=0}^{m} h_i Z^{m-i}$$

and

$$GH = \sum_{k=0}^{m+n} \left(\sum_{i+j} g_i Z^{n-i} h_j Z^{m-j} \right) = \sum_{k=0}^{m+n} f_k Z^{m+n-k} = F.$$

Corollary 4.23 *Suppose $f \in K[X, Y]$ and $F \in K[X, Y, Z]$ is its homogenization. Then f is irreducible in $K[X, Y]$ if and only if F is irreducible in $K[X, Y, Z]$.*

Proof

($\Leftarrow$) Clear from the Lemma.

($\Rightarrow$) Suppose $F = GH$. If $g(X, Y) = G(X, Y, 1)$ and $h(X, Y) = H(X, Y, 1)$, then $f = gh$. Since f is irreducible, one of g or h is constant. Suppose g is constant. Since Z is not a factor of F, Z is not a factor of G. Thus deg $g =$ deg G and G is constant. Thus F is irreducible.

We can now prove the projective version of Study's Lemma. If $F \in K[X, Y, Z]$ is homogeneous, we let

$$V_{\mathbb{P}}(F) = \{p \in \mathbb{P}_2(K) : F(p) = 0\}.$$

Proposition 4.24 *Suppose K is an algebraically closed field. If $F, G \in K[X, Y, Z]$ are homogeneous, F is irreducible and $V_{\mathbb{P}}(F) \subseteq V_{\mathbb{P}}(G)$, then F divides G.*

Proof There are two cases to consider.

<u>case 1</u>: $F = aZ$ for some $a \in K \setminus \{0\}$.

We will suppose F does not divide G and show that $V_{\mathbb{P}}(F) \not\subseteq V_{\mathbb{P}}(G)$. Let

$$G(X, Y) = \sum_{i=0}^{n} g_i(X, Y) Z^i$$

where $g_i \in K[X, Y]$ is homogeneous of degree $n - i$. Since Z does not divide G, $g_0 \neq 0$. Thus there are $x, y \in K$ such that $x \neq 0$, $y \neq 0$ and $g_0(x, y) \neq 0$. But then $[x, y, 0] \in V_{\mathbb{P}}(F) \setminus V_{\mathbb{P}}(G)$ and $V_{\mathbb{P}}(F) \not\subseteq V_{\mathbb{P}}(G)$.

case 2: Z does not divide F.

We can factor $G = Z^d H$ where Z does not divide H. Let $f(X, Y) = F(X, Y, 1)$, and $h(X, Y) = H(X, Y, 1)$. If $f(x, y) = 0$, then $h(x, y) = 0$. Thus $V(f) \subseteq V(h)$ and by Study's Lemma, f divides h. By Lemma 4.22, F divides H. Thus F divides G.

We can now follow the arguments given above for affine curves. We say that a projective curve C is irreducible, if there are no projective curves $D_0, D_1 \subset C$ with $C = D_0 \cup D_1$, $C \neq D_0$, and $C \neq D_1$.

Exercise 4.25 Suppose K is an algebraically closed field and $F \in K[X, Y, Z]$ is nonconstant and homogeneous.

a) Show that $V_{\mathbb{P}}(F)$ is irreducible if and only if $F = G^k$ for some irreducible $G \in K[X, Y, Z]$.

b) Let C be a projective curve. Show that there are irreducible projective curves $C_1, \ldots, C_n$ such that $C = C_1 \cup \ldots \cup C_n$ and $C_i \not\subseteq C_j$ for $i \neq j$. Moreover, if $D \subseteq C$ is an irreducible curve, then $D = C_i$ for some i. We call $C_1, \ldots, C_n$ the irreducible components of C.

c) Suppose $f = F(X, Y, 1)$ and Z does not divide F. Show that if $C_1, \ldots, C_n$ are the irreducible components of $V_{\mathbb{P}}(F)$, then $C_1 \cap \mathbb{A}_2(K), \ldots, C_n \cap \mathbb{A}_2(K)$ are the irreducible components of the affine curve $V(f)$.

d) What happens in c) if Z divides F?

5 Bézout's Theorem

Throughout this section we will assume that K is an algebraically closed field.

Suppose $f, g \in K[X, Y]$ are nonconstant. Our goal in this section is to analyze $|V(f) \cap V(g)|$. One possibility is that f and g have a common nonconstant factor h. In that case $V(h) \subseteq V(f) \cap V(g)$ and $V(f) \cap V(g)$ is infinite. In case f and g have no common nonconstant factor we will prove that $V(f) \cap V(g)$ is finite and

$$|V(f) \cap V(g)| \leq \deg f \deg g.$$

We begin by describing the main idea of the proof. Suppose $f, g \in K[X, Y]$ are nonconstant polynomials with no common nonconstant factors, $\deg f = n$ and $\deg g = m$. By applying an affine transformation if necessary, we may assume that $f(0,0) \neq 0$ and $g(0,0) \neq 0$. The following proposition is the key to the proof.

Proposition 5.1 *There are at most mn lines L through $(0,0)$ such that $L \cap V(f) \cap V(g) \neq \emptyset$.*

We first argue that $V(f) \cap V(g)$ is finite. Let $L_1, \ldots, L_s$ be the lines through $(0,0)$ that intersect $V(f) \cap V(g)$. If $p \in V(f) \cap V(g)$, then there is a unique line L containing p and $(0,0)$ and L must be one of the L_i. Thus

$$V(f) \cap V(g) = \bigcup_{i=1}^{s} (V(f) \cap V(g) \cap L_i).$$

If L is a line and $V(f) \cap V(g) \cap L$ is infinite, then $V(f) \cap L$ is infinite and, by Theorem 3.16 $L \subseteq V(f)$. Similarly, $L \subseteq V(g)$. If $h = 0$ is the linear equation for L, then, by Study's Lemma, h is a common factor of f and g. Thus each L_i intersects $V(f) \cap V(g)$ in at most finitely many points and $V(f) \cap V(g)$ is finite.

Suppose $|V(f) \cap V(g)| = N$. For each pair of distinct points $p, q \in V(f) \cap V(g)$ let $L_{p,q}$ be the unique line containing p and q. Note that

$$|\{L_{p,q} : p, q \in V(f) \cap V(g) \text{ distinct}\}| = \frac{N(N-1)}{2}.$$

By doing a second affine transformation we may assume that $(0,0)$ is not on $V(f)$, $V(g)$ or any of the lines $L_{p,q}$. Let $L_1, \ldots, L_s$ be all lines through $(0,0)$ containing a point of $V(f) \cap V(g)$. Since $(0,0) \in L_i$, $L_i \neq L_{p,q}$ for any distinct $p, q \in V(f) \cap V(g)$. Thus $|L_i \cap V(f) \cap V(g)| = 1$ and

$$|V(f) \cap V(g)| = s \leq mn.$$

In fact, we will not prove the proposition in the form we have stated it. It is somewhat easier to work in projective space rather than affine space and by working in projective space we will be able to prove the following stronger result.

For $F \in K[X, Y, Z]$ homogeneous, we let

$$V_{\mathbb{P}}(F) = \{p \in \mathbb{P}_2(K) : F(p) = 0\}.$$

Theorem 5.2 (Bézout's Theorem) *Let $F, G \in K[X,Y,Z]$ be nonconstant homogeneous of degree m and n respectively. Either F and G have a common nonconstant factor or $|V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)| \leq mn$.*

Moreover, if F and G have no common nonconstant factor, there is a natural way to assign intersection multiplicities $m_p(F,G)$ for each $p \in V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)$ such that

$$\sum_{p \in V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)} m_p(F,G) = mn.$$

Resultants of Homogeneous Polynomials

The key to proving Bézout's Theorem is a result about resultants of homogeneous polynomials. We need one basic fact about homogeneous polynomials.

Exercise 5.3 Suppose $F \in K[X,Y,Z]$ is nonzero. Consider the polynomial $F(TX,TY,TZ) \in K[X,Y,Z,T]$. Then F is homogeneous of degree d if and only if

$$F(TX,TY,TZ) = T^d F(X,Y,Z).$$

[Hint: See 3.6]

Theorem 5.4 *If $F, G \in K[X,Y,Z]$ are nonconstant homogeneous polynomials with no common nonconstant factors such that $F(0,0,1) \neq 0$ and $G(0,0,1) \neq 0$, then $R_{F,G}$ is homogeneous of degree* $\deg F \cdot \deg G$.

Proof By Theorem 4.19, if F and G have no common factor, then $R_{F,G}$ is a nonzero polynomial. We will prove that $R_{F,G}$ is homogeneous of degree d by showing that

$$R_{F,G}(TX,TY) = T^{nm} R_{F,G}(X,Y).$$

Let

$$F = \sum_{i=0}^{n} a_i Z^i \text{ and } G = \sum_{i=0}^{m} b_i Z^i$$

where $a_i, b_i \in K[X,Y]$, a_i is homogeneous of degree $n-i$, b_i is homogeneous of degree $m-i$. Since $F(0,0,1) \neq 0$, $a_n \neq 0$. Similarly, $b_m \neq 0$. Thus $R_{F,G}(TX,TY) =$

$$\begin{vmatrix}
a_0T^n & a_1T^{n-1} & \dots & \dots & \dots & a_n & 0 & \dots & \dots & 0 \\
0 & a_0T^n & a_1T^{n-1} & \dots & \dots & \dots & a_n & 0 & \dots & 0 \\
 & & & \ddots & \ddots & & & & & \\
0 & \dots & \dots & 0 & a_0T^n & a_1T^{n-1} & \dots & \dots & \dots & a_n \\
b_0T^m & b_1T^{m-1} & \dots & \dots & b_m & 0 & \dots & \dots & \dots & 0 \\
0 & b_0T^m & b_1T^{m-1} & \dots & \dots & b_m & 0 & \dots & \dots & 0 \\
 & & & \ddots & \ddots & & & & & \\
0 & \dots & \dots & \dots & 0 & b_0T^m & b_1T^{m-1} & \dots & \dots & b_m
\end{vmatrix}.$$

We modify the determinant by multiplying the ith row by T^{m+1-i} for $i = 1...m$ and the $m+i$th row by T^{n+1-i} for $i = 1...n$.

The first m lines now look like:

$$\begin{array}{cccccccccc}
a_0T^{n+m} & a_1T^{n+m-1} & \dots & \dots & \dots & a_nT^m & 0 & \dots & \dots & 0 \\
0 & a_0T^{n+m-1} & a_1T^{n+m-2} & \dots & \dots & \dots & a_nT^{m-1} & 0 & \dots & 0 \\
 & & & \ddots & \ddots & & & & & \\
0 & \dots & \dots & 0 & a_0T^{n+1} & a_1T^n & \dots & \dots & \dots & a_nT
\end{array}$$

while the last n lines look like:

$$\begin{array}{cccccccccc}
b_0T^{m+n} & b_1T^{m+n-1} & \dots & \dots & b_mT^n & 0 & \dots & \dots & \dots & 0 \\
0 & b_0T^{m+n-1} & b_1T^{m+n-2} & \dots & \dots & b_mT^{n-1} & 0 & \dots & \dots & 0 \\
 & & & \ddots & \ddots & & & & & \\
0 & \dots & \dots & \dots & 0 & b_0T^{m+1} & b_1T^m & \dots & \dots & b_mT
\end{array}.$$

Recall that if we multiply one row of a matrix A by λ, then the determinant of the new matrix is λdet A. Thus the determinant above is equal to

$$T^{(\sum_{i=1}^n i + \sum_{j=1}^m j)} R_{F,G}(TX, TY).$$

To finish the proof we want to show in every element of the jth column the power of T occuring is $T^{m+n+1-j}$. Suppose $i = 1, \dots, m$ in the matrix to compute $R_{F,G}(TX, TY)$ the element in the ith row and jth column is 0 if $j < i$ or $j > i + n + 1$. Otherwise it is

$$a_{j-i}T^{n+i-j}.$$

When we modify the matrix to make the second determinant if $1 \leq i \leq m$ and $i \leq j \leq n + i + 1$, the element in the ith row and jth column is

$$a_{j-i}T^{n+i-j}T^{m+1-i} = a_{j-1}T^{n+m+1-j}.$$

Similarly if $1 \leq i \leq n$ and $i \leq j \leq m + i + 1$ then element in the $(m + i, j)$ position of the first determinant is

$$b_{j-i}T^{m+i-j}$$

while in the second matrix it is

$$b_{j-i}T^{m+i-j}T^{n+1-i} = b_{j-i}T^{n+m+1-j}.$$

Notice that all nonzero entries of the jth column of the second matrix have a $T^{n+m+1-j}$ term. It follows that we could have gotten the second determinant starting with the matrix to compute $R_{F,G}$ and multiplying the first column by T^{n+m}, the second by T^{n+m-1},. . . ,the last by T. Since multiplying a column of a matrix by λ multiplies the determinant by λ. This shows that the second determinant is equal to

$$T^{\sum_{j=1}^{n+m} j} R_{F,G}.$$

Thus

$$T^{\sum_{j=1}^{n+m} j} R_{F,G} = T^{(\sum_{i=1}^{n} i + \sum_{j=1}^{m} j)} R_{F,G}(TX, TY)$$

and

$$R_{F,G}(TX, TY) = T^{(\sum_{j=1}^{n+m} j - \sum_{i=1}^{n} i - \sum_{k=1}^{m} k)} R_{F,G}.$$

But

$$\sum_{i=1}^{s} i = \frac{s(s+1)}{2}.$$

Thus

$$\sum_{j=1}^{n+m} j - \sum_{i=1}^{n} i - \sum_{j=1}^{m} j = \frac{(n+m)^2 + (n+m) - n^2 - n - m^2 - m}{2} = nm$$

and

$$R_{F,G}(TX, TY) = T^{nm} R_{F,G}$$

as desired.

Proof of Bézout's Theorem

We now state and prove the projective version of Proposition 5.1

Proposition 5.5 *Suppose $F, G \in K[X, Y, Z]$ are nonconstant homogeneous polynomials with* deg $F = n$ *and* deg $G = m$ *such that* $F(0,0,1) \neq 0$, $G(0,0,1) \neq 0$, *and F and G have no common nonconstant factors. Then there are at most mn lines in $\mathbb{P}_2(K)$ through $[0,0,1]$ containing a point of $V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)$.*

Proof Since F and G have no common factor, by Theorem 5.4, $R_{F,G}$ is a homogeneous polynomial of degree mn.

In general, projective lines have equations $aX + bY + cZ = 0$, but lines through $[0,0,1]$ have equations $aX + bY = 0$.

Claim Let L be the line $aX + bY = 0$. Then $V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G) \cap L \neq \emptyset$ if and only if $R_{F,G}(b, -a) = 0$.

If $[x, y, z]$ are homogeneous coordinates for a point on L where $x \neq 0$, then $y = \frac{-a}{b} x$ and

$$F(x, y, z) = 0 \Leftrightarrow F(x, \frac{-a}{b} x, z) = 0 \Leftrightarrow F(b, -a, \frac{bz}{x}) = 0.$$

It follows that $V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G) \cap L \neq \emptyset$ if and only if there is a w such that

$$F(b, -a, w) = G(b, -a, w) = 0.$$

Let $f(X) = F(b, -a, X)$ and $g(X) = G(b, -a, X)$. By Theorem 1.22, f and g have a common zero if and only if $R_{f,g} = 0$. But, as in the proof of Study's Lemma,

$$R_{f,g} = R_{F,G}(b, -a).$$

Thus $V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G) \cap L \neq \emptyset$ if and only if $R_{F,G}(b, -a) = 0$.

Since $\lambda aX + \lambda bY = 0$ is the same line as $aX + bY = 0$, lines that contain points of $V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)$ correspond to points of $\mathbb{P}_1(K)$ where $R_{F,G} = 0$. But $R_{F,G} = 0$ has degree mn and at most mn zeros in $\mathbb{P}_1(K)$. Thus there are at most mn lines through $[0, 0, 1]$ intersecting $V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)$.

The proof of Bézout's Theorem now follows the outline at the beginning of the section.

Proof of Bézout's Theorem

Suppose $F, G \in K[X, Y, Z]$ are homogeneous of degree n and m respectively with no common factor. By making a projective transformation we may assume that $F(0, 0, 1) \neq 0$ and $G(0, 0, 1) \neq 0$. Suppose L is a line through $[0, 0, 1]$. If $L \cap V_{\mathbb{P}}(F)$ is infinite, then by Theorem 3.16, $L \subseteq V_{\mathbb{P}}(F)$. By the projective version of Study's Lemma, if $H = 0$ is the homogeneous linear equation for L, then H divides F. Thus if $V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G) \cap L$ is infinite, then H divides F and G, a contradiction. Thus $L \cap V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)$ is finite. Since only finitely many lines through $[0, 0, 1]$ intersect $V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)$, $V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)$ is finite.

Let $C_1, \ldots, C_N$ be all lines containing two or more points of $V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)$. By making a second projective transformation, we may, in addition, assume that $[0, 0, 1] \notin C_i$ for $i = 1, \ldots, N$. Thus if $L_1, \ldots, L_s$ are the lines through $[0, 0, 1]$ intersecting $V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)$, then $|L \cap V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)| = 1$. Thus

$$|V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)| = s \leq mn.$$

It remains to show how to define the intersection multiplicities. Assume, via projective transformations, that we are in the setting where $F(0, 0, 1) \neq 0$, $G(0, 0, 1) \neq 0$ and no line through $[0, 0, 1]$ contains more than one point of $V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)$. There is a one-to-one correspondence between:

i) points of $V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)$;

ii) lines through $[0, 0, 1]$ intersecting $V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)$;

iii) zeros of $R_{F,G}$ in $\mathbb{P}_1(K)$.

Indeed if $p = [a, b, c] \in V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)$, then $R_{F,G}(a, b) = 0$ and c is the unique z such that $[a, b, z] \in V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)$. We let $m_p(F, G)$ be the multiplicity of $[a, b]$ as a zero of $R_{F,G}$. By the remarks after the proof of Theorem 3.16, we see that

$$\sum_{p \in V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)} m_p(F, G) = mn$$

as desired.

Suppose $C = V_{\mathbb{P}}(F)$ is a projective curve. We can factor $F = F_1^{m_1} \cdots F_k^{m_k}$ where $F_1, \ldots, F_k$ are relatively prime irreducible polynomials. Note that

$$C = V(F_1 \cdots F_k).$$

We say that F is a *minimal polynomial for C* if $F = V_{\mathbb{P}}(C)$ and F has no multiple irreducible factors. The previous paragraph shows that every curve has a minimal polynomial.

Exercise 5.6 Suppose F and G are minimal polynomials for a projective curve C. Prove that $F = aG$ for some $a \in K$.

Definition 5.7 If C is a projective curve, the *degree of* C is the degree of a minimal polynomial for C.

Corollary 5.8 *If C and D are projective curves with no common component, then $|C \cap D| \leq \deg C \cdot \deg D$.*

Three Example

Example 1 Let $F(X,Y,Z) = Z^3 - XY^2$ and $G(X,Y,Z) = Z^3 + XY^2$. Using the MAPLE command

```
   resultant(F,G,Z);
```

We find that

$$R_{F,G} = 8X^3Y^6.$$

Suppose $[x,y,z] \in V_{\mathbb{P}}(F) \cap V_{\mathbb{P}}(G)$ then $x = 0$ or $y = 0$. If $x = 0$, then $z = 0$. While if $y = 0$, $z = 0$. Thus $[0,1,0]$ and $[1,0,0]$ are the unique points of intersection. We have

$$m_{[0,1,0]}(F,G) = 3 \text{ and } m_{[1,0,0]}(F,G) = 6.$$

What does this mean in $\mathbb{A}_2(\mathbb{C})$. Let $f(X,Y) = F(X,Y,1) = 1 - XY^2$ and $g(X,Y) = G(X,Y,1) = 1 + XY^2$. The affine curves $V(f)$ and $V(g)$ have no points of intersection. But there are two points of intersection "at infinity".

Example 2 Let $F(X,Y,Z) = X^2 - 2XZ - YZ + Z^2$ and $G(X,Y,Z) = X^2 - 4XZ - YZ + 4Z^2$. In this case the resultant is

$$R_{F,G} = -X^3(6Y - X).$$

Since $R_{F,G}(0,1) = 0$, there must be a point of intersection with homongeneous coordinates $[0,1,z]$. But then $-z + z^2 = 0$ and $-z + 4z^2 = 0$. Thus $z = 0$. Thus $[0,1,0]$ is the unique point of intersection on the line $X = 9$ and this point has multiplicity 3.

We also need to look for a point of intersection on the line $X = 6Y$. We look for a point with homogeneous coordinates $[6,1,z]$. Then

$$0 = 36 - 12z - z + z^2 = (z-9)(z-4)$$

and

$$0 = 36 - 24z - z + 4z^2 = (4z-9)(z-4).$$

Thus $z = 4$ and $[6,1,4]$ is the point of intersection. This point has multiplicity 1.

Let's look at what this means in $\mathbb{A}_2(\mathbb{C})$. Let $f(X,Y) = F(X,Y,1) = X^2 - 2X - Y + 1$ and $g(X,Y) = G(X,Y,1) = X^2 - 4X - Y + 4$. Then these two parabolas have a unique point of interesection in $\mathbb{A}_2(K)$. Since $(6, 1, 4) \sim (\frac{3}{2}, \frac{1}{4}, 1)$. The point of intersection is $(\frac{3}{2}, \frac{1}{4})$. There is an additional point of intersection "at infinity".

We still must address the question of what the intersection multiplicity means. The next example begins to shed some light.

Example 3 Consider the affine curves $Y = X^2 + 1$ and $Y = 1$. We investigate their intersection by first homogenizing them to

$$F(X, Y, Z) = X^2 - YZ + Z^2 \text{ and } G(X, Y, Z) = Y - Z.$$

Then $R_{F,G} = X^2$. If $x = 0$ and $y = 1$, then $z = 1$. Thus $[0, 1, 1]$ is the unique point of intersection and it has multiplicity 2.

Considering the affine equations this is not surprising since the paraboloa $Y = X^2 + 1$ and the line $Y = 1$ intersect at (0,1) and the line is tangent at this point.

Suppose we change the previous problem by taking the line $Y = a$ for any $a \neq 1$. Then

$$F(X, Y, Z) = X^2 - YZ + Z^2 \text{ and } G(X, Y, Z) = Y - aZ.$$

Then

$$R_{F,G} = aX^2 - (a - 1)Y^2$$

If $a \neq 0$ and $\alpha^2 = \frac{a-1}{a}$, then

$$R_{F,G} = a(X - \alpha Y)(X + \alpha Y).$$

Since $Y = aZ$, there are two distinct solutions $[\alpha, a, 1]$ and $[-\alpha, a, 1]$.

Thus is we move the line $Y = 1$ to the line $Y = 1 \pm \epsilon$ for small $\epsilon > 0$, we get two points of intersection. This is the right intuition. If two curves point of intersection of multiplicity > 1 and we perturbe the curves slightly, than we get p distinct points of intesection.