1.10 07/01/24 - Intersection Multiplicity, the Projective Plane

Today, we'll finish the proof of Theorem 1.9.9, and start talking about the projective plane and projective curves.

1.10.1 Intersection Multiplicity

Let's proceed to the proof of Theorem 1.9.9. We need to show two things: existence and uniqueness of i. We'll start with uniqueness.

Proof of Uniqueness in Theorem 1.9.9. We will give an algorithm that takes as input (f, g, P) and returns $i_P(f, g)$ in finitely many steps, using only the axioms (1) - (7).

(a) By (6), we can reduce to the case $P = (0, 0)$.

(b) By (2) and (3), we are done if either f and g have a common component through P, or if $P \notin C_f \cap C_g$, so assume that we are not in either of these cases (we then say that C_f and C_g intersect properly at P).

(c) Consider the polynomials $f(x, 0), g(x, 0) \in k[x]$, and suppose they have degrees $d, e \geq 0$ respectively, where we use the convention that $\deg 0 = 0$. By (1), we may assume by switching f and g if needed that $0 \leq d \leq e$. Now we split into two cases:

Case 1. If $d > 0$, then we may perform the Euclidean algorithm to produce an integer $n \geq 1$ and polynomials $q_1, q_2, \ldots, q_{n+1}, r_1, \ldots, r_n, r_{n+1} \in k[x]$ such that for $i = 0, 1, \ldots, n$, we have

$$r_{i-1} = r_i \cdot q_{i+1} + r_{i+1},$$

and $\deg r_{i+1} < \deg r_i$, where $r_{-1} := g(x, 0)$, $r_0 := f(x, 0)$, $r_1 \cdots r_n \neq 0$, and $r_{n+1} = 0$; then $r_n = \gcd(f(x, 0), g(x, 0))$. Define polynomials $h_1, \ldots, h_n, h_{n+1} \in k[x, y]$ by

$$h_i = h_{i-2} - q_i \cdot h_{i-1}$$

for $i = 1, \ldots, n + 1$, where we set $h_{-1} := g$ and $h_0 = f$. We find inductively using (5) that

$$i_P(f, g) = i_P(h_1, f) = i_P(h_2, h_1) = \cdots = i_P(h_n, h_{n-1}) = i_P(h_{n+1}, h_n),$$

and $h_i(x, 0) = r_i(x)$ for each $i = 1, \ldots, n + 1$, and hence $h_{n+1}(x, 0) = 0$. We replace (f, g) by (h_{n+1}, h_n) and land on

Case 2. If $d = 0$, then $y \mid f$, and so we can write $f = y^N p$ for some $N \geq 1$ and $p \in k[x, y]$ such that $y \nmid p$. Then by (4) we have

$$i_P(f, g) = N \cdot i_P(y, g) + i_P(p, g).$$

By (5), we have

$$i_P(y, g) = i_P(y, g(x, 0)) = i_P(y, y - g(x, 0)) = m_0(g(x, 0)),$$

where in the last step we have used the computation in Example 1.9.12 (this uses (7)). By our assumption that $g|_P = 0$, we have $m_0(g(x, 0)) \geq 1$, and hence $i_P(y, g) \geq 1$, whence $i_P(p, g) < i_P(f, g)$. Either $i_P(p, g) = 0$, in which case we are done; else, return to the beginning of Step (c) with (f, g) replaced by (g, p).

■

It is clear that if such an i exists, then the above algorithm terminates in finitely many steps, and determines the function i uniquely. Let's work out an example in detail to see this in practice.

Example 1.10.1. Let's take

$$\begin{aligned} f(x,y) &= y^2 - x^3 + x, \\ g(x,y) &= (x^2+y^2-3x)^2 - 4x^2(2-x), \end{aligned}$$

and $P = (0,0)$. For simplicity, we work out the case when $\operatorname{ch} k \neq 2$, and leave this (easier) case to the reader. Note that C_f and C_g do not share a component because f is irreducible and $C_f \not\subseteq C_g$: and plugging in $y^2 = x^3 - x$ into g recovers nonzero polynomial

$$x^2(x-1)(x^3+3x^2-4x-8),$$

which has finitely many roots. Let's now apply Step (c).

(1) We have
$$f(x,0) = -x(x-1)(x+1) \text{ and } g(x,0) = x^2(x-1)^2,$$
so that $d = 3$ and $e = 4$. Applying the Euclidean algorithm gives us $n = 1$ with
$$\begin{aligned} q_1(x) &= -x+2, & r_1(x) &= 2x(x-1), \\ q_2(x) &= -\frac{1}{2}(x+1), & r_2(x) &= 0. \end{aligned}$$
Then
$$\begin{aligned} h_1 &= y^4 + (2x^2-5x-2)y^2 + 2x(x-1) \text{ and} \\ h_2 &= \frac{1}{2}y^2\left((1+x)y^2 + x(2x^2-3x-7)\right). \end{aligned}$$
Setting $(f_1, g_1) := (h_2, h_1)$, we are now in Case 2.

(2) Here $N = 2$ and
$$p_1 = \frac{1}{2}\left((1+x)y^2 + x(2x^2-3x-7)\right).$$
Then
$$i_P(f_1,g_1) = 2\cdot m_0(g_1(x,0)) + i_P(p_1,g_1) = 2 + i_P(p_1,g_1).$$
Setting $(f_2, g_2) := (g_1, p_1)$ (switching for degree reasons), we are again in Case 1.

(3) We have
$$f_2(x,0) = 2x(x-1) \text{ and } g_2(x,0) = \frac{1}{2}x(2x^2-3x-7),$$
so that $d = 2$ and $e = 3$. Again, we get $n = 1$ with
$$\begin{aligned} q_1(x) &= \frac{1}{2}x - \frac{1}{4}, & r_1(x) &= -4x, \\ q_2(x) &= -\frac{1}{2}x + \frac{1}{2}, & r_2(x) &= 0. \end{aligned}$$
Then
$$\begin{aligned} h_1 &= \left(-\frac{1}{2}x + \frac{1}{4}\right)y^4 + x\left(-x^2+3x+\frac{1}{4}\right)y^2 - 4x \\ h_2 &= -\frac{1}{8}y^2\left((2x^2-3x-7)y^2 + (4x^4-16x^3-5x^2+41x+16)\right). \end{aligned}$$
Setting $(f_3, g_3) := (h_2, h_1)$, we are now in Case 2.

(4) Here again $N = 2$ and

$$p_3 = -\frac{1}{8}\left((2x^2 - 3x - 7)y^2 + (4x^4 - 16x^3 - 5x^2 + 41x + 16)\right).$$

Then

$$i_P(f_3, g_3) = 2 \cdot m_0(g_3(x, 0)) + i_P(p_3, g_3) = 2 + i_P(p_3, g_3).$$

At this point, we have $i_P(p_3, g_3) = 0$, and the algorithm terminates.

We conclude that $i_P(f, g) = 4$. Get Desmos to draw some pictures to make sure you believe this!

To show existence, we first define the local ring of $\mathbb{A}^2_k$ at a point $P \in \mathbb{A}^2_k$.

Definition 1.10.2. Given a $P \in \mathbb{A}^2_k$, the local ring of $\mathbb{A}^2_k$ at P, denoted $\mathcal{O}_P$, is the ring

$$\mathcal{O}_P := \{r \in k(x, y) : \text{there are } s, t \in k[x, y] \text{ s.t. } r = s/t \text{ and } t|_P \neq 0.\} \subset k(x, y).$$

Since $k[x, y]$ is a UFD and $k(x, y) = \operatorname{Frac} k[x, y]$, this ring can equivalently be defined as the set of $r \in k(x, y)$, which, when written in lowest terms as $r = s/t$ with $s, t \in k[x, y]$ and $t \neq 0$ satisfy $t|_P \neq 0$. We are now ready to sketch the proof of existence.

Proof Sketch of Existence in Theorem 1.9.9. Define

$$i_P(f, g) := \dim_k \mathcal{O}_P/(f, g)\mathcal{O}_P.$$

Properties (1), (5), and (6) are reasonably clear. To show (7), note that for $P = O = (0, 0)$, there is an evaluation map

$$\operatorname{eval}_P : \mathcal{O}_P \to k;$$

this is clearly surjective, and it is easy to see that its kernel is generated by x and y, whence we get an isomorphism

$$\mathcal{O}_P/(x, y)\mathcal{O}_P \xrightarrow{\sim} k$$

and so $i_P(x, y) = 1$. To show (3), note that if $f|_P \neq 0$, then $f \in \mathcal{O}_P^\times$, and so $(f, g)\mathcal{O}_P = \mathcal{O}_P$, and similarly if $g|_P \neq 0$. Conversely, if $f|_P = g|_P = 0$, then $(f, g)\mathcal{O}_P \subset \ker \operatorname{eval}_P$, so

$$\mathcal{O}_P/(f, g)\mathcal{O}_P \twoheadrightarrow \mathcal{O}_P/\ker \operatorname{eval}_P \cong k \text{ implies that } i_P(f, g) \geq 1.$$

To show (2), we may assume $P = O = (0, 0)$. First suppose that we have such a q; then $(f, g)\mathcal{O}_P \subset (q)\mathcal{O}_P$, and we get $\mathcal{O}_P/(f, g)\mathcal{O}_P \twoheadrightarrow \mathcal{O}_P/(q)\mathcal{O}_P$, so it suffices to show that $\mathcal{O}_P/(q)\mathcal{O}_P$ is not finite dimensional over k. To do this, we may assume by a linear change of coordinates that $y \nmid q$; we show that the classes of $1, y, y^2, \ldots$ in $\mathcal{O}_P/(q)\mathcal{O}_P$ are linearly independent. If they were not, then there would be a nonzero $p \in k[y]$ of least degree such that $p \in (q)\mathcal{O}_P$, which is to say that $p = qs/t$ for some nonzero $s, t \in k[x, y]$ with $t|_P \neq 0$. Then $p|_P = 0$ implies $y \mid p$, so if $y \nmid q$, then $y \mid s$, and we may cancel a y from both sides, contradicting our choice of p. Conversely, suppose that f and g have no common components through P. Since irreducible factors of f and g not through P are units in $\mathcal{O}_P$, we may assume by dividing by these factors that f and g are relatively prime in $k[x, y]$. Then, as in Example 1.6.4, Lemma 1.6.2 tells us that there are nonzero $p \in k[x]$ and $q \in k[y]$ such that $p, q \in (f, g)k[x, y] \subset (f, g)\mathcal{O}_P$. Now if we write $p = x^m p_0$ for some $m \geq 0$ and $p_0 \in k[x]$ with $p_0(0) \neq 0$, then $m \geq 1$ because $p \in \ker \operatorname{eval}_P$, and $p_0 \in \mathcal{O}_P^\times$, so that $x^m \in (f, g)\mathcal{O}_P$. Similarly, from q we get an integer $n \geq 1$ such that $y^n \in (f, g)\mathcal{O}_P$. Then it follows that any rational function of the form $1/t$ with $t|_P \neq 0$ can be expanded in $\mathcal{O}_P/(f, g)\mathcal{O}_P$ as $\sum_{i \geq 0}(1 - t)^i$, where all but finitely many terms are zero because of

$[x^n] = [y^m] = 0$. It is then easy to see that the classes of the monomials $x^i y^j$ with $0 \le i \le m-1$ and $0 \le j \le n-1$ span $\mathcal{O}_P/(f,g)\mathcal{O}_P$ as a k-vector space. Finally, to show (4), the result boils down to showing that there is a short exact sequence of the form

$$0 \to \mathcal{O}_P/(f_1,g)\mathcal{O}_P \xrightarrow{\cdot f_2} \mathcal{O}_P/(f_1 f_2, g)\mathcal{O}_P \to \mathcal{O}_P/(f_1,g)\mathcal{O}_P \to 0,$$

and the rank-nullity theorem. For full details, see [3, §3.3, Theorem 3] or [4, Chapter 2]. ■

1.10.2 The Projective Plane

As we have observed before, to count intersection points of curves properly, we have the need for a systematic way to study intersection points "at infinity". One way to do this is to note that every collection of parallel lines has a unique representative through the origin, and so points at infinity should correspond to lines through the origin–which are determined by their slope. Therefore, one approach would be to parametrize points at infinity via a parameter $t \in k$, where t corresponds to the point at infinity along the line $y - tx = 0$. However, this misses exactly one line: namely the vertical line $x = 0$, for which the value of t "would be" ∞.

A more symmetrical approach is to note that lines through the origin can be written as $\lambda x - \mu y = 0$, where $\lambda, \mu \in k$ are not both zero, and the pair (λ, μ) determines the same line as $(c\lambda, c\mu)$ for every $c \in k \setminus \{0\}$, so when $\mu \neq 0$, this corresponds to the above with $t = \lambda/\mu$, but when $\mu = 0$, this adds the line $x = 0$. In this case, we denote the "coordinates" of the line by $[\lambda : \mu]$ to emphasize that only the ratio between the coordinates matters. This gives us a way to think of the "projective plane" $\mathbb{P}^2_k$ as the disjoint union of points $(p,q) \in \mathbb{A}^2_k$ and the directions $[\lambda : \mu]$, but in fact there is a more symmetric way to do it. This leads us to

Definition 1.10.3. The projective plane over k, denoted $\mathbb{P}^2_k$, is the set of equivalence classes of ordered triples of elements (X, Y, Z) of k, not all zero, subject to the equivalence relation that $(X,Y,Z) \sim (cX, cY, cZ)$ for all $c \in k \setminus \{0\} = k^*$, i.e.

$$\mathbb{P}^2_k = \frac{\left\{(X,Y,Z) \in k^3 \setminus \{(0,0,0)\}\right\}}{(X,Y,Z) \sim (cX,cY,cZ)\,\forall\, c \in k^*}.$$

The class of a triple (X,Y,Z) in $\mathbb{P}^2_k$ is usually denoted by $[X : Y : Z]$, and X, Y, Z are called the homogeneous coordinates on $\mathbb{P}^2_k$.

Note that the homogeneous coordinates are not well-defined funcitons on $\mathbb{P}^2_k$–only their ratios are, and those too only away from the loci where the denominator vanishes. Note also that $[0 : 0 : 0]$ is not a well-defined point in $\mathbb{P}^2_k$. Homogeneous coordinates were introduced by Möbius in his 1827 treatise *Der Barycentrische Calcül*. This way of thinking about $\mathbb{P}^2_k$ is in a sense the same as that from before: if $Z \neq 0$, then the point $[X : Y : Z]$ has a unique representative of the form $[x : y : 1]$ where $x := X/Z$ and $y := Y/Z$, and these are the points that compose the $\mathbb{A}^2_k \subset \mathbb{P}^2_k$. When $Z = 0$, however, we get points of the form $[X : Y : 0]$, and these are exactly the points at ∞. One way to think about them is to think of them as the points that are limits of affine the form $[X/\varepsilon : Y/\varepsilon : 1]$ as $\varepsilon \to 0$. The advantage of this formulation is that it makes some additional symmetry–namely that between X, Y, and Z, obvious–which we will leverage to great effect.

Note that in the case of the projective plane, the distinction between polynomials and polynomial functions becomes even more crucial: an arbitrary polynomial $F \in k[X,Y,Z]$ does not even define a well-defined function $F : \mathbb{P}^2_k \to k$ because picking a different representatives (X,Y,Z) of a point $P = [X : Y : Z]$ will in general (i.e. for nonconstant F) yield different values under the polynomial function (on $\mathbb{A}^3_k$) arising from F. However, if F is homogenous of degree $d \geq 0$, then we see that for any $c \in k^\times$ we have

$$F(cX, cY, cZ) = c^d F(X, Y, Z),$$

whence the locus of points $P = [X : Y : Z] \in \mathbb{P}^2_k$ where $F|_P = 0$ still makes sense. This leads us to

Definition 1.10.4. A projective plane algebraic curve is the vanishing locus of a nonconstant homogeneous polynomial F in the projective plane, i.e. a subset $C \subset \mathbb{P}^2_k$ of the form

$$C = C_F := \{P \in \mathbb{P}^2_k : F|_P = 0\}$$

for a nonconstant homogeneous polynomial $F(X, Y, Z) \in k[X, Y, Z]$.

Next time, we'll define the homogenization of a polynomial and the projective closure of algebraic curves in more detail. Today, I want to end with one example.

Example 1.10.5. Consider the hyperbola C_f defined by $f(x, y) = xy - 1 \in k[x, y]$. Then the homogenization of f is $F = f^h = XY - Z^2 \in k[X, Y, Z]$, and the projective closure of C is the curve

$$\overline{C_f} = C_F = \{P = [X : Y : Z] \in \mathbb{P}^2 : XY - Z^2 = 0\}.$$

The intersection $C_F \cap \mathbb{A}^2_k$ is exactly C_f; on the other hand, the new points at infinity correspond to solutions to $XY - Z^2 = Z = 0$, which are the two points $[1 : 0 : 0]$ and $[0 : 1 : 0]$. These are the two points corresponding to the two asymptotes of C_f, namely the lines $x = 0$ and $y = 0$. In particular, over $k = \mathbb{R}$, the two branches which are disjoint in $\mathbb{A}^2_{\mathbb{R}}$ connect up to form one "continuous loop" in $\mathbb{P}^2_{\mathbb{R}}$.