

Bézout's Theorem

MATH CS 120AG Final Project

Harry Coleman

June 8, 2021

1 Introduction

Take K to be an algebraically closed ground field. The fundamental theorem of algebra tells us that for any nonconstant $f \in K[x]$, the number of roots of f in K , counted with multiplicity, equal to its degree. Geometrically, the roots of f are the intersections of its graph with the x -axis in the affine plane. The graph of f is precisely the zeros in \mathbb{A}^2 of the polynomial $f - y \in K[x, y]$, and the x -axis is the zeros of $y \in K[x, y]$. It is then rather uninformative to state that the number of intersections of these two sets, counted with multiplicity, is still $\deg(f - y) \cdot \deg y = \deg f$.

To begin generalizing this notion, we will ask of a pair of polynomials in $K[x, y]$, when it is true that their number of intersections, counted with multiplicity, is precisely the product of their degrees. There are more general questions to ask, along similar lines, such as when K is not algebraically closed, or if we were to consider polynomials in more variables. To keep things simple, we will focus on bivariate polynomials over an algebraically closed field. In order to state the result properly, towards the end, we will consider zeros of polynomials in the projective plane \mathbb{P}^2 , but most work will be done in the affine setting.

2 Background

We will take for granted some introductory knowledge of algebraic geometry, such as a few basic results about affine and projective varieties. Some new notation will be introduced for the ease of working on this particular problem. The first simplification we will make is the identification of polynomials and their zero loci.

Definition 1. An **(affine) curve** is a nonconstant polynomial $F \in K[x, y]$ modulo units, i.e., modulo the equivalence relation $F \sim G$ if $F = \lambda G$ for some $\lambda \in K^\times$. The **degree** of a curve is its degree as a polynomial, denoted $\deg F$. The **(irreducible) components** of a curve F are simply its irreducible components as a polynomial.

Taking the zero locus of a curve $F \in K[x, y]$ is well-defined with respect to the equivalence relation imposed on curves. That is $F(a) = 0$ if and only if $\lambda F(a) = 0$ for all $\lambda \in K^\times$. Therefore, one can regard the equivalence classes of curves to be the equivalence classes of

polynomials which induce the same zero locus. It may then seem natural to make explicit this correspondence.

Notation 2. We will sometimes write

- (a) $a \in F$ instead of $a \in V(F)$, i.e., $F(a) = 0$;
- (b) $F \cap G$ instead of $V(F, G)$ for the points that lie on both F and G ;
- (c) $F \cup G$ for the curve FG ;
- (d) $G \subseteq F$ instead of $G \mid F$.

From algebraic geometry, we know that the sheaf \mathcal{O}_X of an affine variety X , admits a stalk $\mathcal{O}_{X,a}$ at a point $a \in X$, which is isomorphic to the localization $A(X)_{I(a)}$ of the coordinate ring $A(X)$ at the maximal ideal $I(a) \subseteq A(X)$. The stalk is a local ring, in the sense that it has a unique maximal ideal $I_a \subseteq \mathcal{O}_{X,a}$.

This stalk will be a useful when looking at the intersections of curves, as we will be exclusively interested in the behavior around the point of intersection. However, the full generality of the sheaf of regular functions is not necessary for our purposes, so we will define a more lightweight version.

Definition 3. Let $a \in \mathbb{A}^2$ be a point. The **local ring** of \mathbb{A}^2 at a is defined as

$$\mathcal{O}_a = \left\{ \frac{g}{f} : f, g \in K[x, y] \text{ with } f(a) \neq 0 \right\} \subseteq K(x, y)$$

The evaluation map $\mathcal{O}_a \rightarrow K$, given by $\frac{g}{f} \mapsto \frac{g(a)}{f(a)}$ is a well-defined ring homomorphism. Its kernel will be denoted by

$$I_a = \{\varphi \in \mathcal{O}_a \mid \varphi(a) = 0\},$$

which is the unique maximal ideal in \mathcal{O}_a .

There is a natural embedding $K[x, y] \hookrightarrow \mathcal{O}_a$, which maps $f \mapsto \frac{f}{1}$, and we will generally consider $K[x, y] \subseteq \mathcal{O}_a$, under this embedding. With this, we can now define one of the key ingredients for Bézout's lemma, which is the multiplicity of the intersection of two curves. In the case of univariate polynomials, it was simply a matter of decomposing into linear factors. Although $K[x, y]$ is a unique factorization domain, it is not a Euclidean domain, so we should not expect an arbitrary polynomial to decompose into linear factors. Hence, the definition of intersection multiplicity in $K[x, y]$ is a less obvious construction.

Definition 4. The **intersection multiplicity** of two curves F and G at a point $a \in \mathbb{A}^2$ is defined to be

$$\mu_a(F, G) = \dim \mathcal{O}_a / \langle F, G \rangle \in \mathbb{N} \cup \{\infty\}.$$

(For any suitable space, $\dim X$ will always refer to the dimension of X as a K -vector space.)

Proposition 5. (Not proved here) Let F and G be curves and $a \in \mathbb{A}^2$.

- $\mu_a(F, G) \geq 1$ if and only if $a \in F \cap G$.
- If F and G are coprime, then $F \cap G$ and $\mu_a(F, G)$ are finite.

3 Proof

Lemma 6. Let F and G be coprime curves and $a \in F \cap G$. Then for $h \in K[x, y]$ such that $h(a) = 0$, we have $h^n \in \langle F, G \rangle \subseteq \mathcal{O}_a$ for some $n \in \mathbb{N}$.

Proof. As F and G are coprime, put $m = \mu_a(F, G) < \infty$, by Proposition 5. Then we have a descending chain of ideals in \mathcal{O}_a :

$$\mathcal{O}_a = \langle F, G, h^0 \rangle \supseteq \langle F, G, h^1 \rangle \supseteq \cdots \supseteq \langle F, G, h^m \rangle \supseteq \langle F, G \rangle.$$

Taking the quotient of the local ring \mathcal{O}_a by these ideal, we obtain the inequalities

$$0 = \dim \mathcal{O}_a / \mathcal{O}_a \leq \dim \mathcal{O}_a / \langle F, G, h^1 \rangle \leq \cdots \leq \dim \mathcal{O}_a / \langle F, G, h^m \rangle \leq \dim \mathcal{O} / \langle F, G \rangle = m.$$

Since there are $m + 2$ quotients over $m + 1$ possible dimensions, we must have some adjacent pair equal, i.e.,

$$\dim \mathcal{O}_a / \langle F, G, h^n \rangle = \dim \mathcal{O}_a / \langle F, G, h^{n+1} \rangle$$

for some $n \in \mathbb{N}$. In which case, $\langle F, G, h^n \rangle = \langle F, G, h^{n+1} \rangle$, so for some $\alpha, \beta, \gamma \in \mathcal{O}_a$,

$$h^n = \alpha F + \beta G + \gamma h^{n+1} \implies h^n(1 - \gamma h) = \alpha F + \beta G \in \langle F, G \rangle.$$

Since $h(a) = 0$, then $1 - \gamma(a)h(a) = 1$, so $1 - \gamma h \in \mathcal{O}_a^\times$. Hence, $h^n \in \langle F, G \rangle \subseteq \mathcal{O}_a$. □

Lemma 7. Let F and G be coprime curves and $a \in \mathbb{A}^2$ be any point. Then every element of $\mathcal{O}_a / \langle F, G \rangle$ has a polynomial representation.

Proof. If $a \notin F$, then $F(a) \neq 0$. That is, $F \in \mathcal{O}_a^\times$, implying $\langle F, G \rangle = \mathcal{O}_a$. Therefore,

$$\mathcal{O}_a / \langle F, G \rangle = \mathcal{O}_a / \mathcal{O}_a = \{0\},$$

where the result is trivially true of $0 \in K[x, y]$. The same is true of $a \notin G$.

Suppose now that $a \in F \cap G$. By definition, every element of $\mathcal{O}_a / \langle F, G \rangle$ is of the form $\frac{g}{f}$, where $g, f \in K[x, y]$ and $f(a) \neq 0$. Without loss of generality, we may assume $f(a) = 1$; otherwise, $f(a)^{-1} \in K^\times$ and we could write

$$\frac{g}{f} = \frac{f(a)^{-1}g}{f(a)^{-1}f},$$

where $f(a)^{-1}f(a) = 1$.

We first find a polynomial representation for $\frac{1}{f} \in \mathcal{O}_a / \langle F, G \rangle$. Define $h = 1 - f \in K[x, y]$, so

$$\frac{1}{f} = \frac{1}{1 - h}.$$

And $h(a) = 1 - f(a) = 0$, so Lemma 6 tells us that $h^n \in \langle F, G \rangle$ for some $n \in \mathbb{N}$. In other words, $h^n = 0$ in $\mathcal{O}_a/\langle F, G \rangle$, where we now have

$$\frac{1}{f} = \frac{1}{1-h} = \frac{1+h+h^2+\cdots+h^{n-1}}{1-h^n} = 1+h+h^2+\cdots+h^{n-1}.$$

Therefore, we obtain, in $\mathcal{O}_a/\langle F, G \rangle$, the polynomial representation

$$\frac{g}{f} = g(1+h+h^2+\cdots+h^{n-1}).$$

□

Lemma 8. Let F and G be coprime curves and consider the natural ring homomorphism

$$\Phi : K[x, y]/\langle F, G \rangle \longrightarrow \prod_{a \in F \cap G} \mathcal{O}_a/\langle F, G \rangle,$$

which sends the equivalence class of a polynomial $f \in K[x, y]$ to the class of $f \in \mathcal{O}_a$, in each factor $\mathcal{O}_a/\langle F, G \rangle$. Then Φ is an isomorphism

In particular, we have

$$\dim K[x, y]/\langle F, G \rangle = \sum_{a \in F \cap G} \mu_a(F, G) < \infty.$$

Proof. We first prove surjectivity. Suppose $F \cap G = \{a_0, \dots, a_m\}$ and put $a_j = (x_j, y_j)$. Define the polynomial

$$f = \prod_{x_j \neq x_0} \frac{x - x_j}{x_0 - x_j} \cdot \prod_{y_j \neq y_0} \frac{y - y_j}{y_0 - y_j} \in K[x, y].$$

This is similar to a Lagrange basis polynomial, with $f(a_0) = 1$ and $f(a_j) = 0$ for $j = 1, \dots, m$. In particular, $f \in \mathcal{O}_{a_0}^\times$ but $f \in I_{a_j}$ for $j \neq 0$. By Lemma 6, we can choose $n \in \mathbb{N}$ large enough so that $f^n \in \langle F, G \rangle$ in all the stalks $\mathcal{O}_{a_1}, \dots, \mathcal{O}_{a_m}$, i.e., $f^n = 0$ in $\mathcal{O}_{a_j}/\langle F, G \rangle$ for $j \neq 0$.

For an arbitrary $\varphi \in \mathcal{O}_{a_0}$, Lemma 7 gives us a polynomial representation $g \in K[x, y]$ of the element $\varphi \frac{1}{f^n} \in \mathcal{O}_{a_0}/\langle F, G \rangle$. Then $gf^n \in K[x, y]$ with the a_0 -component of $\Phi(gf^n)$ given by

$$g \cdot f^n = \varphi \frac{1}{f^n} \cdot f^n = \varphi \in \mathcal{O}_{a_0}/\langle F, G \rangle.$$

Moreover, for $j \neq 0$, the a_j -component of $\Phi(f^n)$ is given by

$$g \cdot f^n = g \cdot 0 = 0 \in \mathcal{O}_{a_j}/\langle F, G \rangle.$$

Hence, $\Phi(gf^n) = (\varphi, 0, \dots, 0)$. That is, Φ to be surjective on the subset of $\prod_a \mathcal{O}_a/\langle F, G \rangle$ with zeros in all but the a_0 -component. Repeating the same argument for each of components in the codomain, we deduce Φ is surjective on elements with zeros in all but one component.

Given any $\varphi \in \prod_a \mathcal{O}_a/\langle F, G \rangle$, we can write it as $\varphi = \varphi_0 + \cdots + \varphi_m$, where each φ_j has zeros in all but the a_j -component (these are essentially the coordinates of φ). Then, for $j = 0, \dots, m$,

there exists $f_j \in K[x, y]/\langle F, G \rangle$ such that the a_j -component of $\Phi(f_j)$ is precisely φ_j . Hence, we have an element $f = f_0 + \cdots + f_m$ in the domain of Φ , with

$$\Phi(f) = \Phi(f_0 + \cdots + f_m) = \varphi_0 + \cdots + \varphi_m = \varphi.$$

This proves Φ is surjective

To see that Φ is injective, we will show its kernel is trivial. Suppose $f \in \ker \Phi$, then define

$$J = \{g \in K[x, y] : gf \in \langle F, G \rangle\}.$$

We claim that $V(J) = \emptyset$. Suppose, for contradiction, that there exists a point $a \in V(J)$. In particular, $F(a) = G(a) = 0$ so $a \in F \cap G$, corresponding to some component in the codomain of Φ . Then $f \in \ker \Phi$ implies $f = 0 \in \mathcal{O}_a/\langle F, G \rangle$. That is, for some $p_j, q_j \in K[x, y]$ with $q_j(a) \neq 0$, $j = 1, 2$, we have

$$f = \frac{p_1}{q_1}F + \frac{p_2}{q_2}G \in \mathcal{O}_a$$

Put $q = q_1q_2 \in K[x, y]$, then we must have $q(a) = q_1(a)q_2(a) \neq 0$. However,

$$qf = p_1q_2F + p_2q_1G \in \langle F, G \rangle,$$

which implies $q \in J$. By assumption, $a \in V(J)$, so we must have $q(a) = 0$, which is a contradiction. Hence, $V(J) = \emptyset$.

By the affine Nullstellensatz, we have

$$\sqrt{J} = I(V(J)) = I(\emptyset) = K[x, y].$$

In particular, $1 \in K[x, y] = \sqrt{J}$, implying $1 = 1^n \in J$ for some $n \in \mathbb{N}$. By the construction of J , this means $f = 1f \in \langle F, G \rangle$. That is, $f = 0 \in K[x, y]/\langle F, G \rangle$. Hence, $\ker \Phi = 0$, so Φ is injective, and therefore an isomorphism.

Since F and G are coprime, Proposition 5 tells us that $\mu_a(F, G) < \infty$ for all $a \in F \cap G$, which contains only finitely many points. Hence, as a finite sum of finite terms,

$$\dim K[x, y]/\langle F, G \rangle = \sum_{a \in F \cap G} \mu_a(F, G) < \infty.$$

□

For a polynomial f of degree d , we write $\text{LP}(f) = f_d$ to denote the leading part of f , i.e., the degree- d part of f .

Lemma 9. Let F and G be curves, such that their leading parts $\text{LP}(F)$ and $\text{LP}(G)$ are coprime. Then every $f \in \langle F, G \rangle$ can be written as $f = hF + kG$, for some $h, k \in K[x, y]$ with

$$\deg h \leq \deg f - \deg F \quad \text{and} \quad \deg k \leq \deg f - \deg G.$$

Proof. Choose $h, k \in K[x, y]$ such that $f = hF + kG$ and $\deg h$ is minimized. If it is the case that $\text{LP}(hF) + \text{LP}(kG) \neq 0$, then we have

$$d = \deg f = \max\{\deg(hF), \deg(kG)\} = \max\{\deg h + m, \deg k + n\},$$

from which the result immediately follows.

For contradiction, assume $\text{LP}(hF) + \text{LP}(kG) = 0$, then in particular

$$\text{LP}(h)F_m = \text{LP}(hF) = -\text{LP}(kG) = -\text{LP}(k)G_n.$$

Since F_m and G_n have no common component, must have $G_n \mid \text{LP}(h)$ and $F_m \mid \text{LP}(k)$. Then there is a homogeneous polynomial $\ell \in K[x, y]$ such that

$$\text{LP}(h) = \ell G_n \quad \text{and} \quad \text{LP}(k) = -\ell F_m.$$

Define $h' = h - \ell G$ and $k' = k + \ell G$, which gives us the alternative representation

$$f = hF + kG + (\ell FG - \ell FG) = (h - \ell G)F + (k + \ell F)G = h'F + k'G.$$

However, since $\text{LP}(h) = \ell G_n = \text{LP}(\ell G)$, then we also have $\deg h' < \deg h$, a contradiction. □

Proposition 10. If the sequence of linear maps

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D \longrightarrow 0$$

is exact, then $\dim A - \dim B + \dim C - \dim D = 0$. □

Proof. By definition, we have

$$0 = \ker f, \quad \text{im } f = \ker g, \quad \text{im } g = \ker h, \quad \text{im } h = D.$$

With the rank-nullity theorem from linear algebra, we compute

$$\begin{aligned} \dim A &= \dim \text{im } f + \dim \ker f \\ &= \dim \ker g + \dim 0 \\ &= \dim B - \dim \text{im } g \\ &= \dim B - \dim \ker h \\ &= \dim B - (\dim C - \dim \text{im } h) \\ &= \dim B - \dim C + \dim D. \end{aligned}$$

□

Lemma 11. If F and G are coprime curves such that $\text{LP}(F)$ and $\text{LP}(G)$ are also coprime,

$$\dim K[x, y]/\langle F, G \rangle = \deg F \cdot \deg G.$$

Proof. Put $m = \deg F$ and $n = \deg G$.

For $r \in \mathbb{N}$, let $\mathbf{K}_r = \{f \in K[x, y] : \deg f \leq r\}$ be the K -linear subspace of $K[x, y]$, consisting of polynomials with degree at most r . The dimension of this space is $\dim \mathbf{K}_r = \binom{r+2}{2}$.

For an integer $d \geq m + n$, consider the sequence of K -linear maps

$$0 \longrightarrow \mathbf{K}_{d-m-n} \xrightarrow{\psi} \mathbf{K}_{d-m} \times \mathbf{K}_{d-n} \xrightarrow{\varphi} \mathbf{K}_d \xrightarrow{\pi} K[x, y]/\langle F, G \rangle \longrightarrow 0,$$

where $\psi : \ell \mapsto (\ell G, -\ell F)$, $\varphi : (h, k) \mapsto hF + kG$, and π is a restriction of the natural projection. We claim that this sequence is exact, for sufficiently large $d \in \mathbb{N}$.

It is easy to see that $\ker \psi = 0$. If $\ell \in \ker \psi$, then $(\ell G, -\ell F) = (0, 0)$. In particular, $\ell G = 0$ implies $\ell = 0$, since G is nonzero and $\mathbf{K}_{d-m} \subseteq K[x, y]$ is an integral domain.

We have $\text{im } \psi \subseteq \ker \varphi$, since $\varphi(\ell G, -\ell F) = \ell GF - \ell FG = 0$. To show the opposite inclusion, let $(h, k) \in \ker \varphi$, so $hF = -kG$. Since F and G have no common component, then we must have $G \mid h$ and $F \mid k$. Then there exists $\ell \in K[x, y]$ such that $h = \ell G$ and $k = -\ell F$. In particular, $h \in \mathbf{K}_{d-m}$ gives us

$$\deg \ell = \deg h - \deg G = \deg h - n \leq d - m - n,$$

which implies $\ell \in \mathbf{K}_{d-m-n}$ with $\psi(\ell) = (h, k)$. Hence, $\text{im } \psi = \ker \varphi$.

We have $\text{im } \varphi \subseteq \ker \pi$, since $\varphi(h, k) = hF + kG \in \langle F, G \rangle$. Given any $f \in \ker \pi$, we must have $f \in \langle F, G \rangle$ with $\deg f \leq d$. Since the leading terms of F and G have no common component, then Lemma 9 applies. Hence, $f = hF + kG$ for some $h, k \in K[x, y]$ such that $\deg h \leq d - m$ and $\deg k \leq d - n$. That is, $(h, k) \in \mathbf{K}_{d-m} \times \mathbf{K}_{d-n}$ with $\varphi(h, k) = f$. Hence, $\text{im } \varphi = \ker \pi$.

From Lemma 8, $\dim K[x, y]/\langle F, G \rangle < \infty$, so there exists a finite subset $S \subseteq K[x, y]$, whose equivalence classes form a generating set of $K[x, y]/\langle F, G \rangle$ over K . Then every element in the quotient has a representative as a K -linear combination of the polynomials in S . If we assume $d \geq \max_{f \in S} \deg f$, then \mathbf{K}_d contains S and, therefore, the K -linear span of S . This says that every element of $K[x, y]/\langle F, G \rangle$ has a representative in \mathbf{K}_d , i.e., π is surjective.

This proves that the sequence is exact, so Proposition 10 gives us

$$\begin{aligned} \dim K[x, y]/\langle F, G \rangle &= \dim \mathbf{K}_d - \dim(\mathbf{K}_{d-m} \times \mathbf{K}_{d-n}) + \dim \mathbf{K}_{d-m-n} \\ &= \binom{d+2}{2} - \binom{d-m+2}{2} - \binom{d-n+2}{2} + \binom{d-m-n+2}{2} \\ &= mn. \end{aligned}$$

□

Proposition 12. Every homogeneous polynomial in $K[x, y]$ can be written as a product of linear polynomials.

Proof. Let $f \in K[x, y]$ and $d = \deg f$. Then

$$f = \sum_{i=0}^d c_i x^i y^{d-i} = y^d \sum_{i=0}^d c_i \left(\frac{x}{y}\right)^i,$$

where $c_i \in K$. The summation can be considered as a univariate inhomogeneous polynomial in $K[\frac{x}{y}]$. As K is algebraically closed, there is a decomposition into d linear factors:

$$f = y^d \prod_{i=1}^d \left(a_i \frac{x}{y} + b_i\right) = \prod_{i=1}^d (a_i x + b_i y).$$

□

Theorem 13. (Bézout's Theorem) If F and G are coprime projective curves, then

$$\sum_{a \in F \cap G} \mu_a(F, G) = \deg F \cdot \deg G.$$

Proof. Since K is infinite, we can find some line $L \subseteq \mathbb{P}^2 \setminus (F \cap G)$. Then we can perform a projective coordinate transformation on \mathbb{P}^2 , which sends L to the line $V_p(z)$ at infinity. Afterwards, $F \cap G$ is contained in the affine part of \mathbb{P}^2 , i.e., z is not a factor of F or G , so $\deg F = \deg F^i$ and $\deg G = \deg G^i$. Then F^i and G^i are affine curves with no common component; otherwise the homogenization of that common component would be a common component of F and G .

Since z does not divide F nor G , the points at infinity contained in these curves are given by the leading parts $\text{LP}(F^i)$ and $\text{LP}(G^i)$, respectively. But since $F \cap G$ is entirely contained in the affine part of \mathbb{P}^2 , the leading parts are homogenous polynomials in $K[x, y]$ with no common roots. By Proposition 12, we can decompose $\text{LP}(F^i)$ and $\text{LP}(G^i)$ into linear factors. However, any common linear factor would correspond to a common root, hence $\text{LP}(F^i)$ and $\text{LP}(G^i)$ have no common components.

From Lemmas 8 and 11, applied to F^i and G^i , we obtain

$$\sum_{a \in F^i \cap G^i} \mu_a(F^i, G^i) = \dim K[x, y] / \langle F^i, G^i \rangle = \deg F^i \cdot \deg G^i = \deg F \cdot \deg G.$$

Since all the points of $F \cap G$ are in the affine part of \mathbb{P}^2 , where F and G agree with F^i and G^i , then $F^i \cap G^i$ are simply the affine coordinates of the points in $F \cap G$. And for points in the affine part of \mathbb{P}^2 , we have $\mu_{[a_1:a_2:1]}(F, G) = \mu_{(a_1, a_2)}(F^i, G^i)$. So in fact,

$$\sum_{a \in F \cap G} \mu_a(F, G) = \sum_{a \in F^i \cap G^i} \mu_a(F^i, G^i) = \deg F \cdot \deg G.$$

□