

Chapter 1

Introduction

Algebraic geometry is a vast generalization of linear algebra and algebra. Recall that, in linear algebra, you studied the solutions of systems of linear equations

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= 0, \\ &\dots \\ a_{n1}x_1 + \dots + a_{nn}x_n &= 0, \end{aligned}$$

where the coefficients a_{ij} were taken from some field K . The set of solutions turned out to be a vector space, whose dimension does not change if we replace K by some bigger field.

In algebra (or, more exactly, in Galois theory), you looked at the solutions of polynomial equations

$$f(x) = a_nx^n + \dots + a_1x + a_0 = 0,$$

where the coefficients a_i were taken from some field K . This time, the set of solutions depended strongly on the choice of the field K : for example, the polynomial $x^2 + 1$ has no solution in \mathbb{R} , but two solutions in \mathbb{C} .

The natural generalization of these two problems is the following: given a field K and a set of polynomials $f_1, \dots, f_d \in K[X_1, \dots, X_n]$, investigate the set of solutions of the system of polynomial equations

$$\begin{aligned} f_1(X_1, \dots, X_n) &= 0, \\ &\dots \\ f_d(X_1, \dots, X_n) &= 0. \end{aligned}$$

On the one hand, this problem belongs to algebra, since the objects we are dealing with are polynomials over some field; on the other hand, we can think of the set of real solutions of $X^2 + Y^2 - 1 = 0$ as the unit circle, which is a geometric object. In algebraic geometry, the solutions of such systems of polynomial equations are studied using both algebra and geometry.

The simplest objects we will study are plane algebraic curves: these are the zero sets of polynomials $f(X, Y) \in K[X, Y]$ in two variables. Examples are lines (here $f(X, Y) = aX + bY + c = 0$ has degree 1), conics (curves of degree 2), cubics, etc. We say that a plane algebraic curve $\mathcal{C} : f(X, Y) = 0$ is defined over K if $f \in K[X, Y]$, and we call the zero set $\mathcal{C}(K) = \{(x, y) \in K \times K : f(x, y) = 0\}$ the set of K -rational points of \mathcal{C} .

1.1 The Unit Circle

Let me now explain how to use geometry to solve an algebraic problem, namely that of finding all rational solutions of $\mathcal{C} : X^2 + Y^2 = 1$. We start with an obvious solution, say $P = (-1, 0)$ (any rational point on the unit circle \mathcal{C} would do). A line through P with rational slope t will intersect the circle \mathcal{C} in two points, namely P and one other point, say P_t (see Figure 1.1). It is easy to see that P_t is a rational point if t is rational, and that in fact every rational point $\neq P$ on \mathcal{C} is one of these points P_t : in fact, if $Q = (x, y) \neq P$ is a rational point on \mathcal{C} , then $t = \frac{y}{x+1}$ (the slope of the line PQ) gives $Q = P_t$.

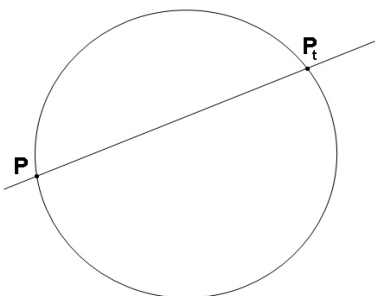


Figure 1.1: Parametrizing the Unit Circle

The actual calculation gives the formulas

$$x = \frac{1-t^2}{1+t^2}, \quad y = \frac{2t}{1+t^2}. \quad (1.1)$$

Note that the point $P = (-1, 0)$ does not correspond to a rational value of t (we could get P by admitting the value $t = \infty$; this will be made precise once we know about projective planes).

Note that, over the field of rational numbers, dividing through by $1+t^2$ is not a problem since $1+t^2 \neq 0$. When determining the solutions of $X^2 + Y^2 = 1$ in arbitrary fields K , one has to be careful if K contains a primitive fourth root of unity.

Consider for example the finite fields $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ for odd primes p . The formulas (1.1) still give points on the unit circle over \mathbb{F}_p , one for each value $t \in \mathbb{F}_p$, except for those values of t for which $1+t^2 = 0$. We know from

elementary number theory that the congruence $t^2 \equiv -1 \pmod{p}$ has a solution (in fact exactly two) if and only if $p \equiv 1 \pmod{4}$. Thus the parametrization (1.1) provides us with p points on \mathcal{C} if $p \equiv 3 \pmod{4}$, and with $p - 2$ points if $p \equiv 1 \pmod{4}$. The same argument as over \mathbb{Q} shows that every point $\neq P$ is actually parametrized by (1.1), which finally shows that the set $\mathcal{C}(\mathbb{F}_p)$ of points on the unit circle with coordinates in \mathbb{F}_p has cardinality

$$\#\mathcal{C}(\mathbb{F}_p) = \begin{cases} p + 1 & \text{if } p \equiv 3 \pmod{4}, \\ p - 1 & \text{if } p \equiv 1 \pmod{4}. \end{cases}$$

The above example is typical for algebraic geometry: we found parametrization of the unit circle using a simple geometric idea (valid over the base field $K = \mathbb{R}$), but the proof itself was algebraic and turned out to be valid for general fields. The motto of algebraic geometry is: think geometrically, prove algebraically!

Unfortunately, not every curve admits a parametrization such as the one above; we will see, for example, that the Fermat curve $X^n + Y^n = 1$ does not have any parametrization whenever $n > 2$. The proof we will give uses Mason's Theorem, whose proof in turn exploits unique factorization in the polynomial ring $K[X]$.

Remark. We have not been very precise in the discussion above; for example, we have not defined what a parametrization is. If we replace t by t^2 in the parametrization of the unit circle, then we get

$$x = \frac{1 - t^4}{1 + t^4}, \quad y = \frac{2t^2}{1 + t^4}.$$

This also parametrizes the unit circle in the sense that we get infinitely many rational points, but we will get fewer.

1.2 Unique Factorization Domains

Let us briefly recapitulate some basic results from algebra. A ring (for us, a ring is always commutative and has a unit element 1) is called a domain if it has no zero divisors, that is, if $ab = 0$ for $a, b \in R$ implies that $a = 0$ or $b = 0$. For example, $\mathbb{Z}/6\mathbb{Z}$ is not an integral domain because $2 \cdot 3 = 0$, but $2 \neq 0$ and $3 \neq 0$ in this ring. Domains R have the cancellation property: if $ab = ac$ for $a, b, c \in R \setminus \{0\}$, then $b = c$; this can be proved by constructing the field of quotients and then multiplying through by a^{-1} . Rings with zero divisors do not have the cancellation property: for example, $3 \cdot 2 = 3 \cdot 4 = 0$, but $2 \neq 4$ in $\mathbb{Z}/6\mathbb{Z}$.

Now let R be a domain. We say that $b \mid a$ for elements $a, b \in R$ if there is a $c \in R$ such that $a = bc$. Elements dividing 1 are called units and form a group R^\times , the unit group of R . A simple observation is that if $a \mid b$ and $b \mid a$, then $a = bu$ for some unit u : in fact, $a \mid b$ implies $b = ac$ for some $c \in R$, and $b \mid a$ shows $a = bd$ for some $d \in R$. Thus $a = bd = acd$

An element $p \in R \setminus R^\times$ is called

- irreducible if every factorization $p = ab$ in R is trivial, that is, if $p = ab$ implies that a or b is a unit;
- prime if it has the property that, whenever $p \mid ab$ divides a product, it divides one of the factors: $p \mid a$ or $p \mid b$.

It is easy to see that primes are always irreducible: in fact, if a prime p has the factorization $p = ab$, then $p \mid ab$; since p is prime, it divides one of the factors, say a ; but then we have $p \mid a$, i.e. $a = pc$ for some $c \in R$. Thus $a = pc = abc$, and from the cancellation law we conclude that $bc = 1$, hence b is a unit and any factorization of p is trivial.

The converse, namely that irreducibles are prime, is not true in general, but it holds for unique factorization domains (UFDs). These are domains in which every nonunit has a factorization into primes that is unique up to units and the order of the factors.

It is known that any Euclidean ring is a UFD; a Euclidean ring is a ring with a Euclidean algorithm, that is, a function $f : R \rightarrow \mathbb{N}$ with the property that

- $f(a) = 0$ if and only if $a = 0$;
- for any pair $a, b \in R \setminus \{0\}$, there is a $q \in R$ such that $f(a - bq) < f(b)$.

The standard examples of Euclidean rings are \mathbb{Z} , where f can be taken to be the absolute value (although there exist other possible choices as well), and the polynomial ring $K[X]$ in one variable over a field K , where f is the function $f(r) = 2^{\deg r}$ for polynomials $r \in K[X]$:

Proposition 1.2.1. *Polynomial rings $K[X]$ in one variable over fields K are Euclidean; in particular, they are principal ideal domains and unique factorization domains.*

From what we have seen above, any irreducible element in $K[X]$ is prime; for example, $X - 4$ and $X^2 + 1$ are primes in $\mathbb{Q}[X]$, and $X^2 + X + 1$ is irreducible in $\mathbb{F}_2[X]$.

The reason we have inserted this section on UFDs is that we will need the following result below:

Proposition 1.2.2. *Let R be a UFD. If a, b, m are nonzero elements of R such that $a \mid m$ and $b \mid m$, and if $\gcd(a, b) = 1$, then $ab \mid m$.*

Proof. Just look at the prime factorizations: we know $m = ac$ for some c , and that $b \mid m = ac$. Since the prime factors of b do not occur in the prime factorization of a , they must all occur in the prime factorization of c , which implies $b \mid c$ and therefore $m = abd$, that is, $ab \mid m$. \square

For Euclidean domains R there is a simpler proof using only the fact that if $\gcd(a, b) = d$, then there are $r, s \in R$ with $d = ar + bs$ (Bezout's Lemma). Now assume that $m = ac = bd$, and write $ar + bs = 1$ (a and b are coprime). Then $bd = mr = arc = (1 - bs)c = c - bsc$, which immediately implies that $b \mid c$, and we are done.

1.3 The Theorem of Stothers-Mason

So far we have seen that the rational points on the unit circle can be described easily using the techniques of sweeping lines; except for one point, we get a bijection between the rational points on the unit circle and the rational numbers. Is it possible to do this for any algebraic curve? As a matter of fact, it is not; below we will give a few examples of curves that cannot be parametrized.

A few years ago, the high school student (now Harvard undergraduate) Noah Snyder [*An alternate proof of Mason's theorem*, Elem. Math. **55** (2000), 93–94] came up with a ‘proof from the Book’ for Mason’s ABC theorem (actually it is very close to a proof posted by Bill Dubuque to sci.math a few years earlier). The following version of his proof is lifted from an article by Dan Bernstein.

Since $K[T]$ is a UFD, we can factor every nonzero polynomial f into irreducible factors: $f = cp_1^{a_1} \cdots p_r^{a_r}$, where the p_i are polynomials in T , and where $c \in K^\times$ is a constant. The product of the distinct prime factors of f is called the radical of f : $\text{rad } f = p_1 \cdots p_r$.

We will also need two well known results:

Lemma 1.3.1. *Let R be a UFD (unique factorization domain).*

1. *If $m \mid ab$ and $\gcd(m, a) = 1$, then $m \mid b$.*
2. *If $a \mid m$, $b \mid m$ and $\gcd(a, b) = 1$, then $ab \mid m$.*

Both claims follow immediately from by looking at the prime factorizations of a , b and m .

The basic lemma in Snyder’s proof is the following:

Lemma 1.3.2. *For nonzero polynomials $f \in K[X]$ we have $\frac{f}{\text{rad } f} \mid f'$.*

Proof. In fact, if $f(T) = p_1(T)^{a_1} \cdots p_r(T)^{a_r}$ is the prime factorization of f , then $\text{rad } f = p_1(T) \cdots p_r(T)$ and $\frac{f}{\text{rad } f} = p_1(T)^{a_1-1} \cdots p_r(T)^{a_r-1}$. On the other hand, $f'(T) = a_1 p_1(T)^{a_1-1} p_2(T)^{a_2} \cdots p_r(T)^{a_r} + p_1(T)^{a_1} \frac{d}{dx} p_2(T)^{a_2} \cdots p_r(T)^{a_r}$, hence $p_1(T)^{a_1-1} \mid f'(T)$. Similarly $p_j(T)^{a_j} \mid f'(T)$, and since the p_j are coprime, the result follows. \square

We need one more trivial statement (note that $\deg 0 = -\infty$ by convention):

Lemma 1.3.3. *Assume that $f \mid g$ for polynomials $f, g \in K[T]$ with $\deg f > \deg g$. Then $g = 0$.*

In fact, $f \mid g$ implies $g = fh$ for some $h \in K[T]$. Then $\deg g = \deg f + \deg h$ is impossible if $\deg g$ is an integer; thus we necessarily have $g = h = 0$.

Now we can state Mason’s theorem (also called the Mason-Stothers Theorem since it was discovered independently by Stothers (1981) and Mason (1984)):

Theorem 1.3.4. *Let K be a field and A, B, C nonzero elements of $K[T]$ with $A + B + C = 0$ and $\gcd(A, B, C) = 1$. If $\deg A \geq \deg \text{rad } ABC$, then $A' = B' = C' = 0$.*

What this theorem says is that if $A + B = C$, then the number of prime factors of ABC cannot be too small (unless the derivatives of the factors are 0).

If the field K has characteristic 0, we can say more. In fact, in this case $A' = 0$ implies that A is constant (not so in characteristic p , where the polynomial T^p has derivative $pT^{p-1} = 0$). Thus if none of A , B or C are constant, then $\deg A \leq \deg \text{rad } ABC - 1$, and by symmetry we have

Corollary 1.3.5. *Let K be a field of characteristic 0. If A, B, C are nonzero polynomials in $K[T]$ with $A + B + C = 0$ and $\gcd(A, B, C) = 1$, then*

$$\max\{\deg A, \deg B, \deg C\} \leq \deg \text{rad } ABC - 1. \quad (1.2)$$

Note that this is best possible: if $A = 1$, $B = T^n$ and $C = 1 + T^n$, then $\text{rad } ABC = T(T^n + 1)$ (note that $T^n + 1$ is squarefree since it is coprime to its derivative) and $n = \max\{\deg A, \deg B, \deg C\} = \deg \text{rad } ABC - 1$.

If K has characteristic p (where p is odd), the example $(1 - T)^p + T^p = 1$ has $\text{rad } ABC = T(1 - T)$ and $\max\{\deg A, \deg B, \deg C\} = p$, so the inequality (1.2) is not satisfied.

Proof of Theorem 1.3.4. First observe that $\gcd(A, B) = 1$ since any common divisor of A and B will also divide C , and this would imply $\gcd(A, B, C) \neq 1$. Similarly, $\gcd(B, C) = \gcd(C, A) = 1$.

Another ingredient is the equation

$$C'B - CB' = AB' - A'B. \quad (1.3)$$

Its proof is easy: $C'B - CB' = (-A' - B')B + (A + B)B' = AB' - A'B$.

Now let us start with the actual proof. By Lemma 1.3.2, $\frac{C}{\text{rad } C}$ divides both C and C' , hence their linear combination $C'B - CB'$. Similarly, $\frac{B}{\text{rad } B} \mid (C'B - CB')$. Finally, (1.3) shows that $\frac{A}{\text{rad } A} \mid AB' - A'B = C'B - CB'$. Since the A , B and C are coprime, so are the quotients $\frac{A}{\text{rad } A}$, $\frac{B}{\text{rad } B}$ and $\frac{C}{\text{rad } C}$. Thus we conclude that

$$\frac{A}{\text{rad } A} \frac{B}{\text{rad } B} \frac{C}{\text{rad } C} \mid (C'B - CB').$$

Since A, B, C are pairwise coprime, we have $(\text{rad } A)(\text{rad } B)(\text{rad } C) = \text{rad}(ABC)$, hence

$$\frac{ABC}{\text{rad}(ABC)} \mid (C'B - CB').$$

Our assumption $\deg A \geq \deg \text{rad } ABC$ now implies

$$\begin{aligned} \deg \frac{ABC}{\text{rad}(ABC)} &= \deg ABC - \deg \text{rad } ABC \\ &\geq \deg ABC - \deg A = \deg BC \\ &> \deg(C'B - CB'). \end{aligned}$$

Lemma 1.3.3 implies that $0 = C'B - CB' = AB' - A'B$. But then $A \mid A'B$, hence $A \mid A'$ since $(A, B) = 1$. Since $\deg A > \deg A'$, this implies $A' = 0$. Similarly, $BC' = CB'$ implies $B \mid B'$, and thus we find $A' = B' = C' = 0$. \square

1.4 Fermat's Last Theorem for Polynomials

As you all know, Fermat's Last Theorem states that there are no nonzero integers such that $X^n + Y^n = Z^n$ for $n > 2$. Equivalently, after dividing through by Z^n it claims that there are no nontrivial rational points on the Fermat curve $x^n + y^n = 1$.

Is Fermat's Last Theorem true for polynomials? According to Ribenboim's books, it was Liouville (not the J. Liouville famous for his theorems on elliptic functions) in 1879 who proved that it actually does hold, even for $n = 2$. I don't understand his proof, which is not too bad since we actually know a counterexample: our parametrization of the unit circle showed that

$$(1 - t^2)^2 + (2t)^2 = (1 + t^2)^2.$$

If we substitute $t = T^n$, we get

$$(1 - T^{2n})^2 + (2T^n)^2 = (1 + T^{2n})^2.$$

This is just what is allowed by Mason's theorem: we have $\deg A = 4n$ and $\deg \text{rad } ABC = \deg T(1 - T^{2n})(1 + T^{2n}) = 4n + 1$.

The correct version of Fermat's Last Theorem for polynomials states

Theorem 1.4.1. *The Fermat curve $x^n + y^n + 1 = 0$ does not have a nontrivial $\mathbb{C}(t)$ -rational point for $n > 2$.*

Proof. Assume that the curve can be parametrized by rational functions (i.e., quotients of polynomials). Clearing denominators we find polynomials with $X(T)^n + Y(T)^n + Z(T)^n = 0$. Hence $\deg X(T)^n \leq h(XYZ) \leq s - 1$, where s is the number of distinct roots of XYZ . Thus $s - 1 < \deg XYZ$, and therefore $n \deg X = \deg X^n \leq \deg X + \deg Y + \deg Z - 1$. The same inequality holds for Y and Z ; adding them gives $n(\deg X + \deg Y + \deg Z) \leq 3(\deg X + \deg Y + \deg Z) - 3$, which implies that $n < 3$. \square

Note that if the Fermat curve does not have nontrivial $\mathbb{C}(t)$ -rational points, then the same is true for $\mathbb{Q}(t)$ -rational points.

Proposition 1.4.2. *The only coprime solutions of $X^4 + Y^4 = Z^2$ in $\mathbb{C}[T]$ are constants.*

Proof. Mason's Theorem. \square

This can easily be generalized:

Theorem 1.4.3. *Let $p, q, r \in \mathbb{N}$ integers. If $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} \leq 1$, then the coprime solutions $\mathbb{C}[t]$ of the generalized Fermat equation*

$$X^p + Y^q + Z^r = 0$$

satisfy $XYZ \in \mathbb{C}$.

Proof. Mason's theorem gives $p \cdot \deg X < \deg XYZ$; now divide through by $p \cdot \deg XYZ$: this shows that $\frac{\deg X}{\deg XYZ} < \frac{1}{p}$. Similarly, $\frac{\deg Y}{\deg XYZ} < \frac{1}{q}$ and $\frac{\deg Z}{\deg XYZ} < \frac{1}{r}$. Adding these three inequalities gives $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} > 1$. \square

1.5 ABC Conjecture.

The result for integers analogous to Mason's ABC theorem is a conjecture not likely to be proved in the near future. Define $h(A, B, C) = \max\{|A|, |B|, |C|\}$ and let $\text{rad}(A)$ denote the product of the distinct prime factors of A , that is, the largest squarefree divisor of A .

ABC Conjecture. *If A, B, C are integers such that $\gcd(A, B, C) = 1$ and $A + B = C$, then for every $\varepsilon > 0$ there is a constant $\mu(\varepsilon)$ such that*

$$h(A, B, C) \leq \mu(\varepsilon) \cdot \text{rad}(ABC)^{1+\varepsilon}.$$

This would be completely analogous to Mason's ABC theorem if we could choose $\varepsilon = 0$. Unfortunately, we cannot: it can be proved rather easily that the choice $\varepsilon = 0$ leads to counterexamples, no matter how large the constant μ is chosen.

For coprime natural numbers a, b, c with $a + b = c$ define

$$P(a, b, c) = \frac{\log \max\{a, b, c\}}{\log \text{rad}(abc)}.$$

Then the ABC conjecture can be stated as

ABC Conjecture. For any $\eta > 1$ there are only finitely many integers a, b, c with $P(a, b, c) > \eta$.

Triples with $P > 1.4$ are called good abc triples. The current record holder is Eric Reyssat's example $A = 2, B = 3^{10} \cdot 109, C = 23^5$ with $P(A, B, C) \approx 1.62991$. Any triple with $P(A, B, C) > 1.53$ is in the current top ten.

The analogue of the Generalized Fermat Conjecture (Theorem 1.4.3) is believed to be true for integers:

Fermat-Catalan Conjecture. *If p, q, r are natural numbers with $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} < 1$, then there exist only finitely many coprime integers x, y, z such that $x^p + y^q = z^r$.*

This can be deduced from the ABC conjecture. Here's the list of known solutions:

$$\begin{array}{ll} 1 + 2^3 = 3^2 & 17^7 + 76271^3 = 21063928^2 \\ 2^5 + 7^2 = 3^4 & 1414^3 + 2213459^2 = 65^7 \\ 7^3 + 13^2 = 2^9 & 9262^3 + 15312283^2 = 113^7 \\ 2^7 + 17^3 = 71^2 & 43^8 + 96222^3 = 30042907^2 \\ 3^5 + 11^4 = 122^2 & 33^8 + 1549034^2 = 15613^3 \end{array}$$

The first entry here is the only solution to Catalan's equation $x^p + 1 = y^q$, as was proved by Mihailescu in 2002.

Exercises

- 5.1 Find all rational points on the circle $X^2 + Y^2 = 2$.
- 5.2 Determine the rational points on the hyperbola $X^2 - 2Y^2 = 1$.
- 5.3 Use the sweeping line technique to parametrize the conic $X^2 - Y^2 + 2X + 1 = 0$ using
1. $P = (0, 1)$
 2. $Q = (-1, 0)$
- as your starting point. Explain your observations (if you can't, use sing surf to sketch the curve).
- 5.4 Show that the curve $X^{1/3} + Y^{1/3} = 1$ is a plane algebraic curve.
- 5.5 Show that the curve $Y = \sin X$ is not a plane algebraic curve.
- 5.6 Use sing surf to sketch the following curves:
1. $Y^2 - X^3 - X^2 = 0$;
 2. $Y^3 + Y^2X - X^2 = 0$;
 3. Folium of Descartes: $X^3 + Y^3 - 3XY = 0$ (make sure you leave a blank space between X and Y).
 4. 5-leaved rose: $(X^2 + Y^2)^3 - 5X^4Y + 10X^2Y^3 - Y^5 = 0$.
- All these curves have the property that lines through $(0, 0)$ intersect the curve in exactly one other point. Use the sweeping lines technique to find all rational points on these curves.
- 5.7 Find all $\mathbb{Q}(T)$ -rational points on the conic $X^2 - (T^4 + T^3)Y^2 = 1$.
- 5.8 Show that $X^2 - (T^4 + T^3)Y^2 = 1$ does not have any nontrivial solutions in $\mathbb{Q}[T]$.
Hint: Mason's theorem.
- 5.9 Find a solution of $X^2 - (T^4 + T^3)Y^2 = 1$ in $\mathbb{F}_5[T]$.
Hint: solve $X^2 - (T^2 + T)Y^2 = 1$ first and then compute the powers of the corresponding unit $X + Y\sqrt{T^2 + T}$ in $\mathbb{F}_q(X)[\sqrt{X^2 + X}]$.
- 5.10 Describe all solutions $X, Y, Z \in \mathbb{F}_p[T]$ of the Fermat equation $X^p + Y^p = Z^p$.
- 5.11 Does $X^4 + Y^2 = Z^2$ have any nontrivial solutions in $\mathbb{C}[T]$?
- 5.12 Consider the equation $X^2 - DY^2 = 1$, where $D \in \mathbb{C}[T]$ is a nonconstant polynomial of degree $\deg D > 0$. Let $n(D)$ denote the number of distinct zeros of D . Show that the equation does not have any solutions $X, Y \in \mathbb{C}[T]$ except $(\pm 1, 0)$ if $2n(D) \leq \deg D$.
- 5.13 Let $X, Y \in \mathbb{C}(t)$ be polynomials. Show that $Y^2 - X^3$ is either 0 or has degree $> \frac{1}{2} \deg x$.