

the Newton-Puiseux method

The main reference for this note is [2, IV§1-3]. We first list properties of formal power series, consider substitution, parametrizations of algebraic curves, and then fractional power (or Puiseux) series. Considering polynomials with power series as coefficients, the theorem of Puiseux generalizes the fundamental theorem of algebra. The constructive proof involves the Newton polygon.

1 Normal Forms

We consider a domain D (for example \mathbb{Z}) and polynomials $p(x) \in D[x]$. While polynomials are finite sums, formal power series are infinite sums and we denote the set of formal power series as $D[[x]]$.

Lemma 1.1 *The series $a_0 + a_1x + a_2x^2 + \dots$ has an inverse in $D[[x]] \Leftrightarrow a_0$ has an inverse in D .*

Proof. We compute the inverse of $a_0 + a_1x + a_2x^2 + \dots$ denoting it as $b_0 + b_1x + b_2x^2 + \dots$. The coefficients of the inverse must satisfy

$$(a_0 + a_1x + a_2x^2 + \dots)(b_0 + b_1x + b_2x^2 + \dots) = 1. \quad (1)$$

Expanding the product above, we solve linear equations in the coefficients of the inverse:

$$a_0b_0 = 1 \Rightarrow b_0 = a_0^{-1}, \text{ because } a_0 \text{ has an inverse} \quad (2)$$

$$a_0b_1 + a_1b_0 = 0 \Rightarrow b_1 = -a_1b_0a_0^{-1} \quad (3)$$

$$a_0b_2 + a_1b_1 + a_2b_0 = 0 \Rightarrow b_2 = -(a_1b_1 + a_2b_0)a_0^{-1}, \text{ etc.} \quad (4)$$

The computation shows \Rightarrow . For \Leftarrow to hold, we take $b_0 + b_1x + b_2x^2 + \dots$ as the inverse of $a_0 + a_1x + a_2x^2 + \dots$ and the product of the two leads to $a_0b_0 = 1$, so b_0 is the inverse of a_0 . \square

We apply Lemma 1.1 to show that every formal power series with coefficients in a field has a normal form.

Theorem 1.1 *Let K be any field. Every element in $K[[x]]$ is of the form*

$$x^{-h}(a_0 + a_1x + a_2x^2 + \dots), \quad h \geq 0. \quad (5)$$

Proof. Any $f \in K[[x]]$ is of the form

$$f = \frac{b_0 + b_1x + b_2x^2 + \dots}{c_0 + c_1x + c_2x^2 + \dots}. \quad (6)$$

Let c_h be the smallest nonzero coefficient. Because K is a field, c_h has an inverse and by Lemma 1.1, $c_h + c_{h+1}x + c_{h+2}x^2 + \dots$ has an inverse $d_0 + d_1x + d_2x^2 + \dots$, so we simplify f as

$$f = \frac{(b_0 + b_1x + b_2x^2 + \dots)(d_0 + d_1x + d_2x^2 + \dots)}{x^h(c_h + c_{h+1}x + c_{h+2}x^2 + \dots)(d_0 + d_1x + d_2x^2 + \dots)} = \frac{a_0 + a_1x + a_2x^2 + \dots}{x^h}. \quad (7)$$

So this gives the form in (5). \square

Theorem 1.1 justifies the definition of the order of a series.

Definition 1.1 For $f \in K[[x]]$, $f(x) = x^k(a_0 + O(x))$ or $f(x) = a_0x^k(1 + O(x))$, k is the order of f , denoted by $O(f)$.

For $f, g \in K[[x]]$, we have $O(fg) = O(f) + O(g)$ and $O(f \pm g) \geq \min(O(f), O(g))$. For convenience, we take $O(0) = \infty$.

Another application of the Lemma 1.1 concerns factorization: $f, g \in K[[x]] : f|g \Leftrightarrow O(f) \leq O(g)$.

2 Substitution

Evaluation for power series may only have meaning for $x = 0$ and then $f(0)$ reflects only what the order and first coefficient of f are. Substitution of one series into another is a more important operation than evaluation.

Definition 2.1 $f, g \in K[[x]]$ are congruent modulo x^m , denoted by $f \equiv g \pmod{x^m}$, if $f - g$ is divisible by x^m , or equivalently $O(f - g) \geq m$, or the first m coefficients of f and g are equal.

Theorem 2.1 *The following statements are true:*

1. The congruence modulo x^m is an equivalence relation.
2. If $f_1 \equiv f_2 \pmod{x^m}$ and $g_1 \equiv g_2 \pmod{x^m}$, then $f_1 \pm g_1 \equiv f_2 \pm g_2 \pmod{x^m}$ and $f_1 g_1 \equiv f_2 g_2 \pmod{x^m}$.
3. If $f \equiv g \pmod{x^m}$ for arbitrarily large m , then $f = g$.
4. If f_1 and f_2 are polynomials, if $O(g_1) > 0$, $O(g_2) > 0$, and if $f_1 \equiv f_2 \pmod{x^m}$, $g_1 \equiv g_2 \pmod{x^m}$, then $f_1(g_1) \equiv f_2(g_2) \pmod{x^m}$.
5. If $f_1, f_2, \dots \in K[[x]]$: $f_{m+1} \equiv f_m \pmod{x^m}$, $m = 1, 2, \dots$, then there exists a unique $f \in K[[x]]$: $f_m \equiv f \pmod{x^m}$, $m = 1, 2, \dots$

Proof. Items 1, 2, and 3 follow from knowing that the first m coefficients are equal.

4. Note that $O(g_1) > 0$ and $O(g_2) > 0$ means $g_1(0) = 0 = g_2(0)$. If f_1, f_2, g_1 and g_2 all agree for the first m coefficients, then we have to prove that the results of substituting g_1 into f_1 and g_2 into f_2 agree for the first m coefficients. Since $f_1 \equiv f_2 \pmod{x^m}$, we can write $f_1 = f_2 + x^m f_3$, where f_3 makes up for all the different coefficients. By 2, adding and multiplying congruent series produces congruent series, applied to $f_1(g_1)$ leads to $f_1(g_1) \equiv f_2(g_2) + g_2^m f_3(g_2) \pmod{x^m}$. Then we compute the order

$$O(g_2^m f_3(g_2)) = m \underbrace{O(g_2)}_{\geq 1} + \underbrace{O(f_3(g_2))}_{\geq 0} \geq m. \quad (8)$$

So the term $g_2^m f_3(g_2)$ is of order m and thus $f_1(g_1) \equiv f_2(g_2) \pmod{x^m}$.

5. We have series $f_1, f_2, f_3, f_4, \dots$ with increasing similarities: $f_2 \equiv f_1 \pmod{x}$, $f_3 \equiv f_2 \pmod{x^2}$, $f_4 \equiv f_3 \pmod{x^3}$, \dots . Denote the coefficient with x^j in f_i as a_{ij} , then:

$$f_1 = a_{10} + a_{11}x + a_{12}x^2 + a_{13}x^3 + \dots \quad (9)$$

$$f_2 = a_{10} + a_{21}x + a_{22}x^2 + a_{23}x^3 + \dots \quad (10)$$

$$f_3 = a_{10} + a_{31}x + a_{32}x^2 + a_{33}x^3 + \dots \quad (11)$$

$$f_4 = a_{10} + a_{41}x + a_{42}x^2 + a_{43}x^3 + \dots \quad (12)$$

Selecting coefficients from the equalities above along the diagonal leads to

$$f = a_{10} + a_{21}x + a_{32}x^2 + a_{43}x^3 + \dots + a_{m+1m}x^m + \dots, \quad (13)$$

so we know f exists, we still have to prove f is unique. Suppose $g \in K[[x]]$: $f_m \equiv g \pmod{x^m}$. Because \equiv is an equivalence relation, we have $f \equiv g \pmod{x^m}$. Since m can be arbitrarily large, we apply 3 to arrive at $f = g$. \square

The relevance of 5 of the theorem above is that we may compute the substitution of the power series one term after the other.

Theorem 2.2 *The following statements are true:*

1. For fixed g , $f \rightarrow f(g)$ is a homomorphism of $K[[x]]$ into itself.
2. If $f(g) \neq 0$, $O(f(g)) = O(f)O(g)$.
3. If $O(g) > 0$, $O(h) > 0$: substituting h in $f(g)$ is the same as substituting $g(h)$ in f .

Proof. 1 and 2 follow from the definition of substitution. To prove 3, let $k = f(g)$ and $\ell = g(h)$, let f_m, g_m, h_m by polynomials congruent modulo x^m to the corresponding series f, g , and h ; note that $k_m = f_m(g_m)$ and $\ell_m = g_m(h_m)$ are polynomials.

As we want to substitute h in $f(g)$, consider $k_m(h_m)$ and as we want to substitute $g(h)$ in f , consider $f_m(\ell_m)$. All of these are polynomials, so $k_m(h_m) = f_m(\ell_m)$. We have: $k_m \equiv k \pmod{x^m}$, $h_m \equiv h \pmod{x^m}$, $f_m \equiv f \pmod{x^m}$, $\ell_m \equiv \ell \pmod{x^m}$, and also: $k_m(h_m) \equiv k(h) \pmod{x^m}$, $f_m(\ell_m) \equiv f(\ell) \pmod{x^m}$. Then by equivalence it follows: $k(h) \equiv f(\ell) \pmod{x^m}$ for arbitrary m , so finally: $k(h) = f(\ell)$. \square

Theorem 2.3 *If $O(g_1) = 1$ and $f_2 = f_1(g_1)$, then*

1. $O(f_2) = O(f_1)$.
2. There is a g_2 , $O(g_2) = 1$, such that $f_1 = f_2(g_2)$, for every $f_1 \in K[[x]]$.

Proof. To show 1, apply 2 of Theorem 2.2.

Let $g_1 = b_1x + b_2x^2 + \dots$, $b_1 \neq 0$ and $g_2 = c_1x + c_2x^2 + \dots$ with coefficients to be determined so that $g_1(g_2) = x$, then $f_2 = f_1(g_1)$, apply 3 of Theorem 2.2: $f_2(g_2) = f_1(g_1(g_2)) = f_1$. Now the proof is reduced to finding the coefficients c_i of g_2 .

The condition $g_1(g_2) = x$ leads to a recurrence on c_i :

$$g_1(g_2) = b_1g_2 + b_2g_2^2 + b_3g_2^3 + \dots \quad (14)$$

$$= b_1(c_1x + c_2x^2 + c_3x^3 \dots) + b_2(c_1x + c_2x^2 + c_3x^3 \dots)^2 + b_3(c_1x + c_2x^2 + c_3x^3 \dots)^3 + \dots \quad (15)$$

$$= b_1c_1x + (b_1c_2 + b_2c_1^2)x^2 + (b_1c_3 + 2b_2c_1c_2 + b_3c_1^3)x^3 + \dots \quad (16)$$

$$+ (b_1c_n + P_n(b_2, \dots, b_n, c_1, \dots, c_{n-1}))x^n + \dots \quad (17)$$

where P_n is a polynomial, and for the coefficient with x^n to be zero: $c_n = -b_1^{-1}P_n(b_2, \dots, b_n, c_1, \dots, c_{n-1})$. \square

3 Parametrizations

We consider series in an auxiliary variable t .

Definition 3.1 Let $F(\mathbf{x}) = 0$ be the equation of an algebraic curve C in the projective plane. Then $x_0, x_1, x_2 \in K[[t]]$ are the coordinates of a parametrization of C if

1. $F([x_0 : x_1 : x_2]) = 0$; and
2. there is no $e \in K[[t]]$ such that $ex_i \in K$, for $i = 1, 2, 3$.

Definition 3.2 For any parametrization x_1, x_2, x_3 of C , let $h = -\min(O(x_i))$ and if $y_i = t^h x_i \in K[[t]]$ is the same parametrization, then $y_i(0) = a_i$ exists and at least one $a_i \neq 0$. The point $[a_0 : a_1 : a_2]$ is the center of the parametrization.

Definition 3.3 Let x_0, x_1, x_2 be a parameterization and $s \in K[[t]]$, $s \neq 0$, and $O(s) > 0$, then $y_i = x_i(t)$ is a parametrization with the same center. If $O(s) = 1$, then $[x_0 : x_1 : x_2]$ and $[y_0 : y_1 : y_2]$ are equivalent.

We call the class of all equivalent parametrizations the place of a curve.

Definition 3.4 If $x \in K[[t^r]]$ for some $r > 1$, then x is reducible as we can simplify x by replacing t^r by s .

The following theorem gives a criterion for when a parametrization is reducible.

Theorem 3.1 The parametrization $(x, y) \in K[[t]]$, with $x = t^n$, $n > 0$, and $y = a_1 t^{n_1} + a_2 t^{n_2} + \dots$, $0 < n_1 < n_2 < \dots$, $a_i \neq 0$ is reducible $\Leftrightarrow \gcd(n, n_1, n_2, \dots) > 1$.

Proof. \Leftarrow If $\gcd(n, n_1, n_2, \dots) = r > 1$, then replace t^r by s .

\Rightarrow If reducible, there is an $s \in K[[t]]$, with $O(s) = 1$ such that $x(s), y(s) \in K[[t^r]]$, with $r > 1$.

Suppose $s/t \notin K[[t^r]]$, then s is of the form

$$s = t(b_0 + b_1 t^{h_1} + \dots + b_k t^{h_k} + \dots) \quad (18)$$

where $b_0 b_k \neq 0$ and furthermore: $r \nmid h_k$, assuming what we want to prove were not true.

$$x(s) = s^n = t^n (b_0 + b_1 t^{h_1} + \dots + b_k t^{h_k} + \dots)^n \quad (19)$$

$$= t^n (b_0 + b_1 t^{h_1} + \dots)^n + n b_k t^{n+h_k} (b_0 + b_1 t^{h_1} + \dots)^{n-1} + \dots \quad (20)$$

Since $x(s)$ is reducible, $x(s) \in K[[t^r]]$, so $r|n$ and because $x(s)$ starts with $t^n b_0$:

$$x(s) - t^n (b_0 + b_1 t^{h_1} + \dots)^n = n t^{n+h_k} b_k b_0^{n-1} + \dots \quad (21)$$

As the left of the equation above belongs to $K[[t^r]]$, also the right hand side belongs to $K[[t^r]]$ and therefore $r|n + h_k$ which contradicts $r \nmid h_k$. Hence $s = tz$, $z \in K[[t^r]]$.

Let us now look at y . Suppose at least one of the n_1, n_2, \dots is not divisible by r and let n_k be the first k such that $r \nmid n_k$. Consider $y(s)$, substituting t in y by tz :

$$y(s) - (a_1 t^{n_1} z^{n_1} + a_2 t^{n_2} z^{n_2} + \dots) = a_{n_k} t^{n_k} (b_0 + b_1 t^{h_1} + \dots)^{n_k} + \dots \quad (22)$$

$$= a_{n_k} b_0^{n_k} t^{n_k} + \dots \quad (23)$$

As the left of the equation above belongs to $K[[t^r]]$, the right cannot belong to $K[[t^r]]$ unless the assumption $r \nmid n_k$ is wrong. \square

Theorem 3.2 In a suitable coordinate system, any given parametrization is equivalent to one of the type

$$\begin{cases} x = t^n, & 0 < n \\ y = a_1 t^{n_1} + a_2 t^{n_2} + \dots, & 0 < n_1 < n_2 < \dots \end{cases} \quad (24)$$

Proof. Choose the center as the origin of the coordinate system, then

$$x_1 = t^n (b_0 + b_1 t + \dots), n > 0, \quad y_1 = t^{n_1} (c_0 + c_1 t + \dots), n_1 > 0. \quad (25)$$

We may assume $b_0 \neq 0$, otherwise we relabel, increasing n . Let $s = t(d_1 + d_2 t + \dots)$, $d_1 \neq 0$ and consider $x = x_1(s)$, $y = y_1(s)$, replacing t by s in x :

$$x = t^n (d_1 + d_2 t + \dots)^n (b_0 + b_1 (d_1 t + \dots) + \dots) \quad (26)$$

$$= t^n (d_1^n b_0 + n(d_1^{n-1} d_2 b_0 + d_1^{n+1} b_1) t + \dots + n(d_1^{n-1} d_i b_0 + P_i(b_1, \dots, b_i, d_1, \dots, d_{i-1})) t^i + \dots). \quad (27)$$

We compute d_1, d_2, \dots , so that $x = t^n$:

$$d_1^n = b_0^{-1}, d_2 = -(n d_1^{n-1} b_0)^{-1} d_1^{n+1} b_1, d_i = -(n d_1^{n-1} b_0)^{-1} P_i, i = 3, 4, \dots \quad (28)$$

Then we have $x = t^n$ and (x, y) is of the required type. \square

4 Fractional Power Series

Instead of writing the parametrizations as $(t^n, y(t))$, we write (t, y) , where $y \in K[[x^{1/n}]]$. Symbolically, we manipulate $x^{1/n}$ along the following rules: $x^{1/1} = x$, $x^{m/n} = (x^{1/n})^m$, $(x^{1/(kn)})^k = x^{1/n}$, $x^{km/(kn)} = x^{m/n}$. Since $(x^{1/(kn)})^k = x^{1/n}$, we have $K[[x^{1/n}]] \subset K[[x^{1/(kn)}]]$.

Definition 4.1 $K[[x]]^* = \bigcup_{n=1}^{\infty} K[[x^{1/n}]]$ is the set of fractional power series.

For $x \in K[[x^{1/n}]]$ and $y \in K[[x^{1/m}]]$, then $x, y \in K[[x^{1/(mn)}]]$ and so are their sum, product, and quotient. Thus, $K[[x]]^*$ is a field.

Theorem 4.1 (Puiseux) If K is algebraically closed, then $K[[x]]^*$ is algebraically closed.

Corollary 4.1 For any $f \in K[[x]]^*[y]$, there is a $z \in K[[x]]^*$: $f(z) = 0$, or more explicitly:

$$f(y) = \sum_{i=0}^n a_i y^i, \quad a_i \in K[[x]]^* \quad (29)$$

$$= a_n \prod_{i=1}^n (y - z_i), \quad f(z_i) = 0. \quad (30)$$

Proof of Theorem 4.1. First we consider necessary conditions, what does a root of $f \in K[[x]]^*[y]$ look like? As the trivial root $z = 0$ is not interesting, we assume $a_0 \neq 0$. A general solution has the form

$$y = c_1 x^{\gamma_1} + c_2 x^{\gamma_1 + \gamma_2} + c_3 x^{\gamma_1 + \gamma_2 + \gamma_3} + \dots, \quad c_i \neq 0, \gamma_1 = O(y), \gamma_2 > 0, \gamma_3 > 0, \dots \quad (31)$$

Abbreviating y to: $c = c_1$, $\gamma = \gamma_1$: $y = x^\gamma(c + y_1)$, $y_1 \in K[[x]]^*$ and substituting in f leads to

$$f(y) = a_0 + a_1 x^\gamma(c + y_1) + a_2 x^{2\gamma}(c + y_1)^2 + \dots + a_n x^{n\gamma}(c + y_1)^n \quad (32)$$

$$= a_0 + a_1 c x^\gamma + a_2 c^2 x^{2\gamma} + \dots + a_n c^n x^{n\gamma} + g(y_1). \quad (33)$$

As $y_1 = c_2 x^{\gamma_2}$, $O(y_1) = \gamma_2 > 0 \Rightarrow O(g) \geq \gamma_2$ and g collects terms of order higher than some of the $a_i c^i x^{i\gamma}$. Necessary conditions for $f(y) = 0$ are then:

1. At least two of $a_i c^i x^{i\gamma}$ have the same order, say j and k , and that order is less than any other:

$$O(a_j c^j x^{j\gamma}) = O(a_k c^k x^{k\gamma}) \leq O(a_i c^i x^{i\gamma}), \quad i \neq j, i \neq k \quad (34)$$

$$O(a_j) + j\gamma = O(a_k) + k\gamma \leq O(a_i) + i\gamma. \quad (35)$$

2. The coefficients of lowest order must cancel:

$$\sum_{\ell} a_\ell c^\ell = 0. \quad (36)$$

$$O(a_\ell) + \ell\gamma = O(a_j) + j\gamma$$

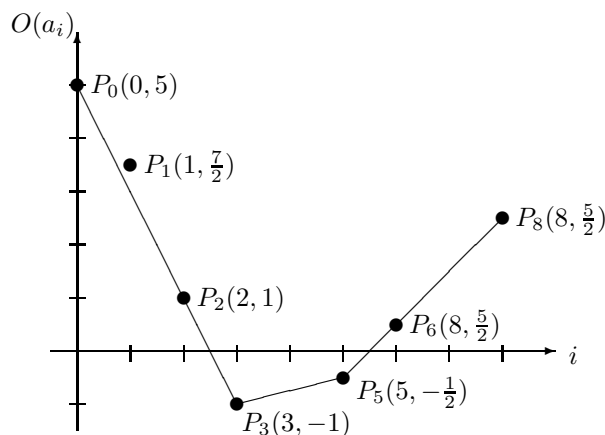
To determine values for γ we use the Newton polygon. To define the Newton polygon we use the points $P_i(i, O(a_i))$ where the coordinates consist of i the exponent of y and the order of the coefficient $a_i \in K[[x]]^*$. If $a_i = 0$, then there is no point with first coordinate i .

To visualize the first necessary condition for $f(y) = 0$ we connect those points with a line for which there are no other points below that line. This corresponds to taking the lower convex hull of the set of points P_i .

We consider the following example:

$$f(y) = a_0x^5 + a_1x^{7/2}y + a_2xy^2 + a_3x^{-1}y^3 + a_5x^{-1/2}y^5 + a_6x^{1/2}y^6 + a_7x^{10/3}y^7 + a_8x^{5/2}y^8. \quad (37)$$

The Newton polygon for f is below:



For points $P_i(i, O(a_i))$, the order γ of the solution defines a slope $(\gamma, 1)$. With $(\gamma, 1)$ we can rank the points via the inner product: $\langle (i, O(a_i)), (\gamma, 1) \rangle = i\gamma + O(a_i)$. This slope condition gives a geometric interpretation to the first necessary condition on the orders of the coefficients a_i . In particular, the condition on γ is

$$O(a_j) + j\gamma = O(a_k) + k\gamma \quad \Rightarrow \quad \gamma = \frac{O(a_j) - O(a_k)}{k - j}. \quad (38)$$

We have three slopes and compute the two necessary conditions as follows:

$$P_0P_2P_3 : \quad \gamma = \frac{O(a_0) - O(a_2)}{2 - 0} = \frac{5 - 1}{2} = 2, \quad a_0c^0 + a_2c^2 + a_3c^3 = 0 \quad (39)$$

$$P_3P_5 : \quad \gamma = \frac{O(a_3) - O(a_5)}{5 - 3} = \frac{-1 - (-1/2)}{3} = -1/4, \quad a_3c^3 + a_5c^5 = 0 \quad (40)$$

$$P_5P_6P_8 : \quad \gamma = \frac{O(a_5) - O(a_6)}{6 - 5} = \frac{-1/2 - 1/2}{1} = -1, \quad a_5c^5 + a_6c^6 + c_8c^8 = 0. \quad (41)$$

Note that if we solve the polynomials in c , we find $3 + 2 + 3 = 8$ nonzero values for c as roots of the polynomials. Because the solutions are of the form $x^\gamma(c + y_1)$, negative values for γ correspond to solutions at infinity as $x \rightarrow 0$.

Three issues still need careful consideration: (1) There are $\deg(f)$ values for c . We look at consecutive points on the lower hull of the Newton polygon. The conditions on the coefficients of lowest order are polynomials with exponents as points on the slopes of the Newton polygon. (2) The Computation of other terms in the expansion. Consider $y = x^{\gamma_1}(c_1 + c_2x^{\gamma_2})$. Either $y = c_1x^{\gamma_1}$ and we are done, or we substitute and look for positive slopes, as $\gamma_2 > 0$. (3) The powers γ 's have bounded denominators.

As pointed in [2, §3.3], for $K = \mathbb{C}$, the proof of the theorem of Puiseux is much shorter, but then does not involve the Newton polygon. The discussion on Puiseux series in [1, §14] relies on the Hensel lemma to show the factorization of a polynomial with coefficients as power series.

5 Exercises

1. Similar to deriving polynomials, we can take derivatives of power series. Show that the product rule for the derivation of power series holds.
2. Maple offers the procedure `algebraiccurves[puiseux]` to determine the Puiseux expansion of an algebraic function. Apply this procedure to the example (37) used to illustrate the Newton polygon.
3. Take one root of the $f(y) = 0$ in (37) and compute the second term in its series expansion. Draw the Newton polygon of $f(y_1)$ after substitution of $y = x^{\gamma_1}(c_1 + c_2y_1^{\gamma_2})$ and verify there is a positive choice for γ_2 . Is this choice unique?
4. Continue the previous exercise by computing more terms of the series expansion. Verify that the denominators of the γ 's are bounded.

References

- [1] B.L. Van Der Waerden. *Einführung in die Algebraische Geometrie*. Dover Publications, 1945.
- [2] R.J. Walker. *Algebraic Curves*. Princeton University Press, 1950.