Polynomials over Power Series and their Applications to Limit Computations (tutorial version)

Marc Moreno Maza
University of Western Ontario
IBM Center for Advanced Studies

CASC 2018 Tutorial
Université de Lille
Bâtiment M^3
May 14, 2021

Plan

1. Polynomials over Power Series
   - Weierstrass Preparation Theorem

Plan

Plan

Weierstrass Polynomials (1/4)

Remark

Let $f \in \mathbb{K}[[X_1, \ldots, X_n]]$. We write $f = \sum_{j=0}^{\infty} f_j X_n^j$ with $f_j \in \mathbb{K}[[X_1, \ldots, X_{n-1}]]$ for $j \in \mathbb{N}$. Let $\rho = (\rho_1, \ldots, \rho_n) \in \mathbb{R}_{>0}^n$. We write $\rho' = (\rho_1, \ldots, \rho_{n-1})$. Then we have

$$\| f \|_\rho = \sum_{j=0}^{\infty} \| f_j \|_{\rho'} \rho_n^j.$$

Hence, if $f \in \mathbb{K}\langle X_1, \ldots, X_n \rangle$ holds, then so does $f_j \in \mathbb{K}\langle X_1, \ldots, X_{n-1} \rangle$ for all $j \in \mathbb{N}$.

Definition

Let $f \in \mathbb{K}[[X_1, \ldots, X_n]]$ with $f \neq 0$. We write $f(\underline{0}, X_n) = f(0, \ldots, 0, X_n)$. Let $k \in \mathbb{N}$. We say that f is

- *general* in X_n if $f(\underline{0}, X_n) \neq 0$ holds,
- *general* in X_n of order k if $\mathrm{ord}(f(\underline{0}, X_n)) = k$,

Clearly $\mathrm{ord}(f) \leq \mathrm{ord}(f(\underline{0}, X_n))$ holds. However, we have the following.

Weierstrass Polynomials (2/4)

Lemma 1

Let $f \in \mathbb{K}[[X_1, \ldots, X_n]]$ with $f \neq 0$ and $k := \mathrm{ord}(f)$. Then there is a shear:

$$\begin{aligned} X_i &= Y_i + c_i Y_n \quad i = 1, \ldots, n-1 \\ X_n &= Y_n \end{aligned}$$

such that $g(Y) = f(X(Y)) \in \mathbb{K}[[Y_1, \ldots, Y_n]]$ is general in Y_n of order k.

Proof (1/2)

- Let $d \in \mathbb{N}$. We write
$$f_{(d)} = \sum_{|e|=d} a_e X_1^{e_1} \cdots X_{n-1}^{e_{n-1}} X_n^{e_n}.$$
- Since the coordinate change is linear, we have
$$g_{(d)}(Y) = f_{(d)}(X(Y)).$$

Weierstrass Polynomials (3/4)

Proof (2/2)

- For $d = k$ in particular, we have
$$\begin{array}{rcl} g_{(k)}(Y) & = & \sum_{|e|=k} a_e (Y_1 + c_1 Y_n)^{e_1} \cdots (Y_{n-1} + c_{n-1} Y_n)^{e_{n-1}} Y_n^{e_n} \\ & = & \left(\sum_{|e|=k} a_e c_1^{e_1} \cdots c_{n-1}^{e_{n-1}} Y_n^k \right) + h(Y) \end{array}$$
where $h(Y)$ necessarily satisfies $h(\underline{0}, Y_n) = 0$.
- Observe also that the coefficient of Y_n^k is a polynomial in $c_1, \ldots, c_{n-1}$, which is not identically zero.
- Indeed, if it would, then all its coefficients would be, that is, $f_{(k)} = 0$ would hold, in contradiction to our assumption $k := \mathrm{ord}(f)$.
- Since this polynomial in $c_1, \ldots, c_{n-1}$ is not zero, the variables $c_1, \ldots, c_{n-1}$ can be specialized to values that ensure that $g_{(k)}(Y)$ has degree k in Y_n. Quod erat demonstrandum!

Weierstrass Polynomials (4/4)

Remark

- Let $f \in \mathbb{K}[[X_1, \ldots, X_n]]$ such that $f \in \mathbb{K}\langle X_1, \ldots, X_{n-1}\rangle[X_n]$ holds and $k := \deg(f, X_n)$. Assume (just for this remark) that $\mathbb{K} = \mathbb{C}$.
- Hence, we write $f = \sum_{j=0}^{k} f_j X_n^j$ with $f_j \in \mathbb{K}\langle X_1, \ldots, X_{n-1}\rangle$ for all $j = 0 \cdots k$.
- In this case, the power series $f_0, \ldots, f_k$ have a common radius of convergence $\rho' \in \mathbb{R}_{>0}^{n-1}$ so that they are holomorphic in the polydisk $D' := \{x \in \mathbb{K}^{n-1} \mid |x_i| < \rho_i\}$.
- Consequently f is holomorphic in $D' \times \mathbb{K}$.

Definition

Let $k \in \mathbb{N}$. Let $f = \sum_{j=0}^{k} f_j X_n^j \in \mathbb{K}[[X_1, \ldots, X_{n-1}]][X_n]$ with $f_j \in \mathbb{K}\langle X_1, \ldots, X_{n-1}\rangle$ for $j = 0 \cdots k$ and with $f_k \neq 0$. We say that f is a *Weierstrass polynomial* if we have

$$f_0(\underline{0}) = \cdots = f_{k-1}(\underline{0}) = 0 \quad \text{and} \quad f_k = 1.$$

Theorem 3

Let $g \in \mathbb{K}\langle X_1, \ldots, X_n \rangle$ be general of order k. Then, there is a unique pair (α, p) with $\alpha \in \mathbb{K}\langle X_1, \ldots, X_n \rangle$ and $p \in \mathbb{K}\langle X_1, \ldots, X_{n-1} \rangle [X_n]$ such that

1. α is a unit,
2. p is a Weierstrass polynomial of degree k,
3. we have $g = \alpha\, p$.

Thus we have

$$g = \alpha(\underline{X}) \left(X_n^k + a_1(X_1, \ldots, X_{n-1}) X_n^{k-1} + \cdots + a_k(X_1, \ldots, X_{n-1}) \right),$$

with $a_1(\underline{0}) = \cdots = a_k(\underline{0}) = 0$. Moreover, if $g \in \mathbb{K}\langle X_1, \ldots, X_{n-1} \rangle [X_n]$ then $\alpha \in \mathbb{K}\langle X_1, \ldots, X_{n-1} \rangle [X_n]$ also holds.

Remark

The above theorem implies that in some neighborhood of the origin, the zeros of g are the same as those of the Weierstrass polynomial p.

Weierstrass division theorem

Theorem 4

Let $f, g \in \mathbb{K}\langle X_1, \ldots, X_n \rangle$ with g general in X_n of order k. Then, there exists a unique pair (q, r) with $q \in \mathbb{K}\langle X_1, \ldots, X_n \rangle$ and $r \in \mathbb{K}\langle X_1, \ldots, X_{n-1} \rangle [X_n]$ such that we have

1. $\deg(r, X_n) \leq k - 1$,
2. $f = qg + r$.

Moreover, if $f, g \in \mathbb{K}\langle X_1, \ldots, X_{n-1} \rangle [X_n]$ with

$$g = g_0 + g_1 X_n + \cdots + g_k X_n^k \quad \text{and} \quad g_k(0) \neq 0,$$

then g_k is a unit in the ring $\mathbb{K}\langle X_1, \ldots, X_{n-1} \rangle$ and the classical division theorem (in polynomial rings) gives $q \in \mathbb{K}\langle X_1, \ldots, X_{n-1} \rangle [X_n]$.

Proof of the division theorem (1/7)

Proof of existence (1/5)

- We write $f = \sum_{j=0}^{\infty} f_j X_n^j$ with $f_j \in \mathbb{K}\langle X_1, \ldots X_{n-1}\rangle$ for $j \in \mathbb{N}$.
- We write $f = \hat{f} + \tilde{f} X_n^k$ with
$$\hat{f} = \sum_{j=0}^{k-1} f_j X_n^j \quad \text{and} \quad \tilde{f} = \sum_{j=k}^{\infty} f_j X_n^{j-k}.$$
- Let $\rho = (\rho_1, \ldots, \rho_n) \in \mathbb{R}_{>0}^n$. We have $\parallel f \parallel_\rho = \parallel \hat{f} \parallel_\rho + \parallel \tilde{f} \parallel_\rho \rho_n^k$. In particular
$$\parallel \tilde{f} \parallel_\rho \leq \rho_n^{-k} \parallel f \parallel_\rho. \qquad (1)$$
- Similarly, we write $g = \hat{g} + \tilde{g} X_n^k$.
- Since g is general in X_n at order k, it follows that $\tilde{g}$ is a unit.
- Let ρ be chosen such that all of $f, g, \tilde{g}^{-1}$ are in B_ρ.
- We consider the auxiliary function h defined as
$$h \;=\; X_n^k - g\tilde{g}^{-1} \;=\; -\hat{g}\tilde{g}^{-1}.$$

Proof of the division theorem (2/7)

Proof of existence (2/5)

- We claim that for all $\nu \in \mathbb{R}$, with $0 < \nu < 1$, we can choose ρ such that we have
$$\parallel h \parallel_\rho \leq \nu \rho_n^k. \tag{2}$$
- Recall that we have $h = X_n^k - g\tilde{g}^{-1}$ and $\tilde{g}^{-1}(0_1, \ldots, 0_n) \neq 0$.
- More precisely, since $g = \hat{g} + \tilde{g}X_n^k$ holds, we have
$$h = X_n^k - g\tilde{g}^{-1} = X_n^k - \left(\hat{g} + \tilde{g}X_n^k\right)\tilde{g}^{-1} = -\tilde{g}^{-1}\left(\sum_{j=0}^{k-1} g_j X_n^j\right),$$
with $g_j \in \mathbb{K}\langle X_1, \ldots X_{n-1}\rangle$ and $g_j(0_1, \ldots, 0_{n-1}) = 0$ for $j = 0, \ldots, k-1$. Therefore $h(0_1, \ldots, 0_{n-1}, X_n)$ is identically zero.
- Writing $h = \hat{h} + \tilde{h}X_n^k$ with $\hat{h} = \sum_{j=0}^{k-1} h_j X_n^j$ and $h_j \in \mathbb{K}\langle X_1, \ldots X_{n-1}\rangle$, we deduce $\tilde{h}(0_1, \ldots, 0_n) = 0$.

Proof of the division theorem (3/7)

Proof of existence (3/5)

- Since $\tilde{h}(0_1, \ldots, 0_n) = 0$, we can decrease ρ such that we have

$$\| \tilde{h} \|_\rho \leq \frac{\nu}{2}, \text{ thus } \quad \| \tilde{h} X_n^k \|_\rho \leq \frac{\nu}{2} \rho_n^k. \tag{3}$$

- With $\rho' = (\rho_1, \ldots, \rho_{n-1})$, and writing $\hat{h} = \sum_{j=0}^{k-1} h_j X_n^j$, we have
$$\| \hat{h} \|_\rho \leq \sum_{j=0}^{k-1} \| h_j \|_\rho \rho_n^j.$$
- Since $h_0(\underline{0}) = \cdots = h_{k-1}(\underline{0}) = 0$ holds, we can decrease ρ (actually ρ') while holding ρ_n fixed such that for $j = 0, \ldots, k-1$, we have

$$\| h_j \|_{\rho'} \leq \frac{\nu}{2} \rho_n^{k-j}, \text{ thus } \quad \| \hat{h} \|_\rho \leq \frac{\nu}{2} \rho_n^k. \tag{4}$$

- Finally, the claim of (2) follows from (3) and (4).

Proof of the division theorem (4/7)

Proof of existence (4/5)

- The function h is used as follows. For every $\phi \in \mathbb{K}\langle X_1, \ldots, X_n \rangle$, we define $h(\phi) = h\tilde{\phi}$ where $\tilde{\phi}, \hat{\phi}$ are defined as $\tilde{f}, \hat{f}$.
- By combining (1) and (2), we deduce
$$\| h(\phi) \|_\rho \leq \| h \|_\rho \| \tilde{\phi} \|_\rho \leq \nu \rho_n^k \rho_n^{-k} \| \phi \|_\rho = \nu \| \phi \|_\rho .$$
- This lets us write an iteration process
$$\phi_0 := f, \ \phi_{i+1} := h(\phi_i) = h\tilde{\phi}_i .$$
- Observe that the series $\phi := \sum_{i=0}^{\infty} \phi_i$ converges for the metric topology of B_ρ since
$$\| \phi \|_\rho \leq \sum_{i=0}^{\infty} \| \phi_i \|_\rho \leq \sum_{i=0}^{\infty} \nu^i \| f \|_\rho = \| f \|_\rho \tfrac{\nu}{1-\nu} .$$
We define
$$q := \tilde{\phi}\tilde{g}^{-1} \ \text{ and } \ r := \hat{\phi}.$$
- Observe that $q \in B_\rho$ and $r \in B_{\rho'}[X_n]$ hold.

Proof of the division theorem (5/7)

Proof of existence (5/5)

- Clearly we have

$$\tilde{\phi} = \sum_{i=0}^{\infty} \tilde{\phi}_i \quad \text{and} \quad \hat{\phi} = \sum_{i=0}^{\infty} \hat{\phi}_i.$$

- Observe also that we have

$$\begin{array}{rcl} \phi_i - \phi_{i+1} & = & \phi_i - h\tilde{\phi}_i \\ & = & \hat{\phi}_i + X_n^k \tilde{\phi}_i - \left(X_n^k - g\tilde{g}^{-1}\right) \tilde{\phi}_i \\ & = & \hat{\phi}_i + g\tilde{g}^{-1}\tilde{\phi}_i. \end{array}$$

- Putting everything together

$$\begin{array}{rcl} f & = & \phi_0 \\ & = & \sum_{i=0}^{\infty} \left(\phi_i - \phi_{i+1}\right) \\ & = & \sum_{i=0}^{\infty} \hat{\phi}_i + g\tilde{g}^{-1} \sum_{i=0}^{\infty} \tilde{\phi}_i \\ & = & r + gq. \end{array}$$

- This proves existence.

Proof of the division theorem (6/7)

Proof of uniqueness (1/2)

- Proving the uniqueness is equivalent to prove that for all q, r satisfying $\deg(r, X_n) < k$ and $0 = qg + r$ we have $q = r = 0$.
- So let $q \in \mathbb{K}\langle \underline{X} \rangle$ and $r \in \mathbb{K}\langle X_1, \ldots, X_{n-1} \rangle [X_n]$ $\deg(r, X_n) < k$ and $0 = qg + r$.
- We have seen that there exists $\rho \in \mathbb{R}^n_{>0}$ such that $g, q, r, \tilde{g}^{-1} \in B_\rho$ holds.
- For $h = X_n^k - g\tilde{g}^{-1}$ as above, we have

$$q\tilde{g}h \;=\; q\tilde{g}X_n^k - q\tilde{g}g\tilde{g}^{-1} \;=\; q\tilde{g}X_n^k + r.$$

Proof of the division theorem (7/7)

Proof of uniqueness (2/2)

- We assume that ρ is chosen such that (2) holds, that is, $\parallel h \parallel_\rho \leq \nu \rho_n^k$. Defining $M = \parallel q\tilde{g} \parallel_\rho \rho_n^k$, and using $\deg(r, X_n) < k$, we have:

$$
\begin{array}{rcl}
M & = & \parallel q\tilde{g}X_n^k \parallel_\rho \\
 & \leq & \parallel q\tilde{g}X_n^k + r \parallel_\rho \\
 & = & \parallel q\tilde{g}h \parallel_\rho \\
 & \leq & \parallel q\tilde{g} \parallel_\rho \parallel h \parallel_\rho \\
 & \leq & \parallel q\tilde{g} \parallel_\rho \nu \rho_n^k \\
 & = & \nu M.
\end{array}
$$

- Since $0 < \nu < 1$, we deduce $M = 0$.
- Since $\rho_n \neq 0$, we have $\parallel q\tilde{g} \parallel_\rho = 0$.
- Since $\tilde{g} \neq 0$, we finally have $q = 0$, and thus $r = 0$.

Proof of the first point of the preparation theorem

Proof of the existence

- We apply the division theorem and divide $f = X_n^k$ by g leading to
$$X_n^k \ = \ qg + \sum_{i=1}^{k} a_i X_n^{k-i} \ \text{ with } \ a_i \in \mathbb{K}\langle X_1, \ldots, X_{n-1}\rangle.$$
- That is,
$$qg \ = \ X_n^k - \sum_{i=1}^{k} a_i X_n^{k-i}.$$
- We substitute $X_1 = \cdots = X_{n-1} = 0$ leading to
$$q(\underline{0}, X_n)(cX_n^k + \cdots) \ = \ X_n^k - \sum_{i=1}^{k} a_i(\underline{0}) X_n^{k-i}.$$
with $c \in \mathbb{K}$ and $c \neq 0$.
- Comparing the coefficients of X_n^{ℓ} for all $\ell \in \mathbb{N}$ shows that $q(0,0) = \frac{1}{c} \neq 0$ and $a_1(0) = \cdots = a_k(0) = 0$
- Thus q is a unit and setting $\alpha = q^{-1}$ completes the proof of the existence statement.

Proof of the uniqueness

Follows immediately from the uniqueness of the division theorem.

Proof of the second point of the preparation theorem

Proving $g \in \mathbb{K}\langle X_1, \ldots, X_{n-1}\rangle[X_n] \quad \Rightarrow \alpha \in \mathbb{K}\langle X_1, \ldots, X_{n-1}\rangle[X_n]$

- Let (α, p) be given by the first point of the preparation theorem, thus, $g = \alpha\, p$ and p is a Weierstrass polynomial of degree k,
- We further assume $g \in \mathbb{K}\langle X_1, \ldots, X_{n-1}\rangle[X_n]$.
- Since p is a monic polynomial in X_n, we can divide g by p in $\mathbb{K}\langle X_1, \ldots, X_{n-1}\rangle[X_n]$ yielding $q, r \in \mathbb{K}\langle X_1, \ldots, X_{n-1}\rangle[X_n]$ such that
$$g = qp + r \ \text{ and } \ \deg(r, X_n) < k.$$
- Applying the uniqueness of the Weierstrass preparation theorem, we deduce
$$\alpha = q \ \text{ and } \ r = 0.$$

Quod erat demonstrandum!

Implicit Function Theorem (1/3)

Remark

An important special case of the Weierstrass preparation theorem is when the polynomial f has order $k = 1$ in X_n. In this case, we change the notations for convenience.

Notations and assumptions

- Let $f = \sum_{j=0}^{\infty} f_j Y^j$ with $f_j \in \mathbb{K}\langle X_1, \ldots, X_n \rangle$, $f(0) = 0$ and $\frac{\partial(f)}{\partial(Y)}(0) \neq 0$. Then f is general in Y of order 1.
- By the preparation theorem, there exists a unit $\alpha \in \mathbb{K}\langle X_1, \ldots, X_n, Y \rangle$ and $\phi \in \mathbb{K}\langle X_1, \ldots, X_n \rangle$ such that
$$f = \alpha(Y - \phi) \;\; \text{and} \;\; \phi(0) = 0.$$
- In this section on the *Implicit Function Theorem* we also assume that $\mathbb{K} = \mathbb{C}$ holds.

Implicit Function Theorem (2/3)

Observations

- We have
$$f(\underline{X}, \phi(\underline{X})) = \alpha(\underline{X}, \phi(\underline{X}))\,(\phi(\underline{X}) - \phi(\underline{X})) = 0.$$
- Now consider an arbitrary series $\psi(\underline{X}) \in \mathbb{K}\langle \underline{X} \rangle$ such that $\psi(0) = 0$ and $f(\underline{X}, \psi(\underline{X})) = 0$ hold.
- From $f(\underline{X}, \psi(\underline{X})) = 0$, we deduce
$$0 = f(\underline{X}, \psi(\underline{X})) = \alpha(\underline{X}, \psi(\underline{X}))\,(\psi(\underline{X}) - \phi(\underline{X})) = 0.$$
- Since $\psi(0) = 0$ and $\alpha(0,0) \neq 0$, we have $\alpha(0, \psi(0)) \neq 0$.
- Since α and ψ are continuous, there exists a neighborhood of $\underline{0} \in \mathbb{K}^n$ in which $\alpha(x, \psi(x)) \neq 0$.
- It follows that $\psi(x) = \phi(x)$ holds in this neighborhood.
- Therefore, we have proved the following.

Implicit Function Theorem (3/3)

Theorem 5

Let $f \in \mathbb{C}\langle X_1, \ldots, X_n, Y\rangle$ such that

$$f(0) = 0 \text{ and } \frac{\partial(f)}{\partial(Y)}(0) \neq 0.$$

Then, there exists exactly one series $\psi \in \mathbb{C}\langle X_1, \ldots, X_n\rangle$ such that we have

$$\psi(0) = 0 \text{ and } f(X_1, \ldots, X_n, \psi(X_1, \ldots, X_n)) = 0.$$

Hensel Lemma (1/3)

Notations

- Let $f = a_0 Y^k + a_1 Y^{k-1} + \cdots + a_k$ with $a_k, \ldots, a_0 \in \mathbb{K}\langle X_1, \ldots, X_n \rangle$.
- We define $\overline{f} = f(0_1, \ldots, 0_n, Y) \in \mathbb{K}[Y]$.

Assumptions

1. f is monic in Y, that is, $a_0 = 1$.
2. $\mathbb{K}$ is algebraically closed. Thus, there exist positive integers $k_1, \ldots, k_r$ and pairwise distinct elements $c_1, \ldots, c_r \in \mathbb{K}$ such that we have
$$\overline{f} = (Y - c_1)^{k_1}(Y - c_2)^{k_2} \cdots (Y - c_r)^{k_r}.$$

Theorem 6

There exist $f_1, \ldots, f_r \in \mathbb{K}\langle X_1, \ldots, X_n \rangle[Y]$ all monic in Y s.t. we have

1. $f = f_1 \cdots f_r$,
2. $\deg(f_j, Y) = k_j$, for all $j = 1, \ldots, r$,
3. $\overline{f_j} = (Y - c_j)^{k_j}$, for all $j = 1, \ldots, r$.

Proof of Hensel Lemma (1/2)

- The proof is by induction on r.
- Assume first $r = 1$. Observe that $k = k_1$ necessarily holds. Now define $f_1 := f$. Clearly f_1 has all the required properties.
- Assume next $r > 1$. We apply a change of coordinates sending c_r to 0

$$\begin{array}{rcl} g(\underline{X}, Y) & = & f(\underline{X}, Y + c_r) \\ & = & (Y + c_r)^k + a_1(Y + c_r)^{k-1} + \cdots + a_k \end{array}$$

- By definition of $\overline{f}$ and c_r, we deduce that $g(\underline{X}, Y)$ is general in Y of order k_r.
- By the preparation theorem, there exist $\alpha, p \in \mathbb{K}\langle X_1, \ldots, X_n \rangle [Y]$ such that α is a unit, p is a Weierstrass polynomial of degree k_r and we have $g = \alpha p$.

Hensel Lemma (3/3)

Proof of Hensel Lemma (1/2)

- Then, we set $f_r(Y) = p(Y - c_r)$ and $f^* = \alpha(Y - c_r)$.
- Thus f_r is monic in Y and we have $f = f^* f_r$.
- Moreover, we have
$$\overline{f^*} = (Y - c_1)^{k_1}(Y - c_2)^{k_2} \cdots (Y - c_{r-1})^{k_{r-1}}.$$
- The existence of $f_1, \ldots, f_{r-1}$ follows by applying the induction hypothesis on f^*.