Polynomials over Power Series and their Applications to Symbolic Analysis

Marc Moreno Maza

University of Western Ontario

May 12, 2021

Plan

1. From Formal to Convergent Power Series

Plan

Formal power series (1/4)

Notations

- $\mathbb{K}$ is a complete field, that is, every Cauchy sequence in $\mathbb{K}$ converges.
- $\mathbb{K}[[X_1, \ldots, X_n]]$ denotes the set of formal power series in $X_1, \ldots, X_n$ with coefficients in $\mathbb{K}$.
- These are expressions of the form $\Sigma_e\, a_e X^e$ where e is a multi-index with n coordinates $(e_1, \ldots, e_n)$, X^e stands for $X_1^{e_1} \cdots X_n^{e_n}$, $|e| = e_1 + \cdots + e_n$ and $a_e \in \mathbb{K}$ holds.
- For $f = \Sigma_e\, a_e X^e$ and $d \in \mathbb{N}$, we define

$$f_{(d)} = \sum_{|e|=d} a_e X^e \quad \text{and} \quad f^{(d)} = \sum_{k \leq d} f_{(k)},$$

which are the *homogeneous part* and *polynomial part* of f in degree d.

Addition and multiplication

For $f, g \in \mathbb{K}[[X_1, \ldots, X_n]]$, we define

$$f + g = \sum_{d \in \mathbb{N}} \left(f_{(d)} + g_{(d)}\right) \quad \text{and} \quad fg = \sum_{d \in \mathbb{N}} \left(\Sigma_{k+\ell=d} \left(f_{(k)} g_{(\ell)}\right)\right).$$

Formal power series (2/4)

Order of a formal power series

For $f \in \mathbb{K}[[X_1, \dots, X_n]]$, we define its *order* as

$$\operatorname{ord}(f) = \begin{cases} \min\{d \mid f_{(d)} \neq 0\} & \text{if } f \neq 0, \\ \infty & \text{if } f = 0. \end{cases}$$

Remarks

For $f, g \in \mathbb{K}[[X_1, \dots, X_n]]$, we have

$$\operatorname{ord}(f+g) \geq \min\{\operatorname{ord}(f), \operatorname{ord}(g)\} \quad \text{and} \quad \operatorname{ord}(fg) = \operatorname{ord}(f) + \operatorname{ord}(g).$$

Consequences

- $\mathbb{K}[[X_1, \dots, X_n]]$ is an integral domain.
- $\mathcal{M} = \{f \in \mathbb{K}[[X_1, \dots, X_n]] \mid \operatorname{ord}(f) \geq 1\}$ is the only maximal ideal of $\mathbb{K}[[X_1, \dots, X_n]]$.
- We have $\mathcal{M}^k = \{f \in \mathbb{K}[[X_1, \dots, X_n]] \mid \operatorname{ord}(f) \geq k\}$ for all $k \in \mathbb{N}$.

Krull Topology

Recall $\mathcal{M} = \{f \in \mathbb{K}[[X_1, \ldots, X_n]] \mid \mathrm{ord}(f) \geq 1\}$. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of elements of $\mathbb{K}[[\underline{X}]]$ and let $f \in \mathbb{K}[[\underline{X}]]$. We say that

- $(f_n)_{n\in\mathbb{N}}$ *converges* to f if for all $k \in \mathbb{N}$ there exists $N \in \mathbb{N}$ s.t. for all $n \in \mathbb{N}$ we have $n \geq N \quad \Rightarrow \quad f - f_n \in \mathcal{M}^k$,
- $(f_n)_{n\in\mathbb{N}}$ is a *Cauchy sequence* if for all $k \in \mathbb{N}$ there exists $N \in \mathbb{N}$ s.t. for all $n, m \in \mathbb{N}$ we have $n, m \geq N \quad \Rightarrow \quad f_m - f_n \in \mathcal{M}^k$.

Proposition 1

- We have $\bigcap_{k\in\mathbb{N}} \mathcal{M}^k \;=\; \langle 0 \rangle$,
- If every Cauchy sequence in $\mathbb{K}$ converges, then every Cauchy sequence of $\mathbb{K}[[\underline{X}]]$ converges too.

Formal power series (4/4)

Proposition 2

For all $f \in \mathbb{K}[[X_1, \ldots, X_n]]$, the following properties are equivalent:

(i) f is a unit,

(ii) $\mathrm{ord}(f) = 0$,

(iii) $f \notin \mathcal{M}$.

Sketch of proof

This follows from the classical observation that for $g \in \mathbb{K}[[X_1, \ldots, X_n]]$, with $\mathrm{ord}(g) > 0$, the following holds in $\mathbb{K}[[X_1, \ldots, X_n]]$

$$(1 - g)(1 + g + g^2 + \cdots) = 1$$

Since $(1 + g + g^2 + \cdots)$ is in fact a sequence of elements in $\mathbb{K}[[X_1, \ldots, X_n]]$, proving the above relation formally requires the use of Krull Topology.

Abel's Lemma (1/2)

Geometric series

From now on, the field $\mathbb{K}$ is equipped with an absolute value. The *geometric series* $\Sigma_e X^e$ is absolutely convergent provided that $|x_1| < 1, \ldots, |x_n| < 1$ all hold. Then we have

$$\Sigma_e\, x_1^{e_1} \cdots x_n^{e_n} = \frac{1}{(1-x_1)\cdots(1-x_n)}.$$

Abel's Lemma

Let $f = \Sigma_e\, a_e X^e \in \mathbb{K}[[\underline{X}]]$, let $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$, let $M \in \mathbb{R}_{>0}$ and Let $\rho_1, \ldots, \rho_n$ be real numbers such that

(i) $|a_e x^e| \leq M$ holds for all $e \in \mathbb{N}^n$,

(ii) $0 < \rho_j < |x_j|$ holds for all $j = 1 \cdots n$.

Then f is uniformly and absolutely convergent in the polydisk

$$D = \{z \in \mathbb{K}^n \ \mid \ |z_j| < \rho_j\}.$$

In particular, the limit of the sum is independent of the summand order.

Abel's Lemma (2/2)

Corollary 1

Let $f = \Sigma_e \, a_e X^e \in \mathbb{K}[[\underline{X}]]$. Then, the following properties are equivalent:

(i) There exists $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$, with $x_j \neq 0$ for all $j = 1 \cdots n$, s.t. $\Sigma_e \, a_e x^e$ converges.

(ii) There exists $\rho = (\rho_1, \ldots, \rho_n) \in {\mathbb{R}_{>0}}^n$ s.t. $\Sigma_e \, a_e \rho^e$ converges.

(iii) There exists $\sigma = (\sigma_1, \ldots, \sigma_n) \in {\mathbb{R}_{>0}}^n$ s.t. $\Sigma_e \, |a_e| \sigma^e$ converges.

Definition

A power series $f \in \mathbb{K}[[\underline{X}]]$ is said *convergent* if it satisfies one of the conditions of the above corollary. The set of the convergent power series of $\mathbb{K}[[\underline{X}]]$ is denoted by $\mathbb{K}\langle \underline{X} \rangle$.

Remark

It can be shown that, within its domain of convergence, a formal power series is a multivariate holomorphic function. Conversely, any multivariate holomorphic function can be expressed locally as the sum of a power series.

ρ-norm of a power series

Notation

Let $\rho = (\rho_1, \ldots, \rho_n) \in {\mathbb{R}_{>0}}^n$. For all $f = \Sigma_e\, a_e X^e \in \mathbb{K}[[\underline{X}]]$, we define

$$\parallel f \parallel_\rho = \Sigma_e\, |a_e| \rho^e.$$

Proposition 3

For all $f, g \in \mathbb{K}[[\underline{X}]]$ and all $\lambda \in \mathbb{K}$, we have

- $\parallel f \parallel_\rho = 0 \quad \Longleftrightarrow \quad f = 0,$
- $\parallel \lambda f \parallel_\rho = |\lambda| \parallel f \parallel_\rho,$
- $\parallel f + g \parallel_\rho \leq \parallel f \parallel_\rho + \parallel g \parallel_\rho,$
- If $f = \Sigma_{k \leq d}\, f_{(d)}$ is the decomposition of f into homogeneous parts, then $\parallel f \parallel_\rho = \Sigma_{k \leq d} \parallel f_{(d)} \parallel_\rho$ holds.
- If f, g are polynomials, then $\parallel fg \parallel_\rho \leq \parallel f \parallel_\rho \parallel f \parallel_\rho,$
- $\lim_{\rho \to 0} \parallel f \parallel_\rho \; = \; |f(0)|.$

Convergent power series form a ring (1/5)

Notation

Let $\rho = (\rho_1, \ldots, \rho_n) \in {\mathbb{R}_{>0}}^n$. We define

$$B_\rho = \{f \in \mathbb{K}[[\underline{X}]] \ \mid \ \| f \|_\rho < \infty\}$$

Cauchy's estimate

Observe that for all $f = \Sigma_e \, a_e X^e \in \mathbb{K}[[\underline{X}]]$, we have for all $e \in \mathbb{N}^e$

$$|a_e| \leq \frac{\|f\|_\rho}{\rho^e}.$$

Theorem 1

The set B_ρ is a Banach algebra. Moreover,

1. if $\rho \leq \rho'$ holds then we have $B_{\rho'} \subseteq B_\rho$,
2. we have $\bigcup_\rho B_\rho = \mathbb{K}\langle \underline{X} \rangle$.

Proof (1/3)

- From Proposition 3, we know that B_ρ is a normed vector space.
- Proving that $\parallel fg \parallel_\rho \leq \parallel f \parallel_\rho \parallel g \parallel_\rho$ holds for all $f, g \in \mathbb{K}[[\underline{X}]]$ is routine. Thus, B_ρ is a normed algebra.
- It remains to show that B_ρ is complete.
- Let $(f_j)_{j \in \mathbb{N}}$ be a Cauchy sequence in B_ρ. We write $f_j = \Sigma_e \, a_e^{(j)} X^e$.
- From Cauchy's estimate, for each $e \in \mathbb{N}^n$, for all $i, j \in \mathbb{N}$ we have $|a_e^{(j)} - a_e^{(i)}| \leq \dfrac{\parallel f_j - f_i \parallel_\rho}{\rho^e}$.

Proof (2/3)

- Since $\mathbb{K}$ is complete, for each $e \in \mathbb{N}^n$, the sequence $(a_e^{(j)})_{j \in \mathbb{N}}$ converges to an element $a_e \in \mathbb{K}$.
- We define $f = \Sigma_e \, a_e X^e$. It must be shown that
  - (i) $f \in B_\rho$ holds and
  - (ii) $\lim_{j \to \infty} f_j = f$ holds in the metric topology induced by the ρ-norm of the normed vector space B_ρ.
- Hence we must show that
  - (i) $\parallel f \parallel_\rho < \infty$ holds, and
  - (ii) for all $\varepsilon > 0$ there exists $j_0 \in \mathbb{N}$ s.t. for all $j \in \mathbb{N}$ we have $j \geq j_0 \quad \Rightarrow \quad \parallel f - f_j \parallel_\rho \leq \varepsilon$.
- Let $\varepsilon > 0$. Since $(f_j)_{j \in \mathbb{N}}$ is a Cauchy sequence in B_ρ, there exists $j_0 \in \mathbb{N}$ s.t. for all $j \geq j_0$ and all $i \geq 0$ we have

$$\sum_e |a_e^{(j+i)} - a_e^{(j)}|\rho^e \;=\; \parallel f_{j+i} - f_j \parallel_\rho \;<\; \tfrac{\varepsilon}{2}.$$

Convergent power series form a ring (4/5)

Proof (3/3)

- Let $s \in \mathbb{N}$ be fixed. Since for each $e \in \mathbb{N}^n$ the sequence $(a_e - a_e^{(i)})_{i \in \mathbb{N}}$ converges to 0 in $\mathbb{K}$, there exists $i_0 \in \mathbb{N}$ s.t. for all $j \geq j_0$ and all $i \geq i_0$ we have
$$\sum_{|e|=0}^{s} |a_e - a_e^{(j+i)}|\rho^e \ < \ \tfrac{\varepsilon}{2}.$$
- Therefore, for all $j \geq j_0$ and all $i \geq i_0$ we have
$$\sum_{|e|=0}^{s} |a_e - a_e^{(j)}|\rho^e \leq$$
$$\sum_{|e|=0}^{s} |a_e - a_e^{(j+i)}|\rho^e + \sum_{e} |a_e^{(j+i)} - a_e^{(i)}|\rho^e < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon.$$
- Since the above holds for all s, we deduce that for all $j \geq j_0$
$$\| f - f_j \|_\rho = \sum_{e} |a_e - a_e^{(j)}|\rho^e \leq \varepsilon,$$
- which proves (ii). Finally, (i) follows from
$$\| f \|_\rho \leq \ \| f - f_{j_0} \|_\rho + \ \| f_{j_0} \|_\rho \leq \varepsilon + \ \| f_{j_0} \|_\rho < \infty.$$

Convergent power series form a ring (5/5)

Corollary 2

$\mathbb{K}\langle \underline{X} \rangle$ is a subring of $\mathbb{K}[[\underline{X}]]$.

Proof

For $f, g \in \mathbb{K}\langle \underline{X} \rangle$, there exists $\rho \in {\mathbb{R}_{>o}}^n$ s.t. $f, g \in B_\rho$. While proving the previous theorem we proved $fg \in B_\rho$. Moreover, $f + g \in B_\rho$ clearly holds.

Corollary 3

Let $f \in \mathbb{K}\langle \underline{X} \rangle$. If f is a unit in $\mathbb{K}[[\underline{X}]]$, then f is also a unit in $\mathbb{K}\langle \underline{X} \rangle$.

Sketch of Proof

W.l.o.g. we can assume $f(0) = 1$ and we define $g = 1 - f$. We know that f^{-1} is the limit of the sequence $1 + g + g^2 + \cdots$ in Krull's topology. Since $g(0) = 0$, there exists $\rho \in {\mathbb{R}_{>o}}^n$ s.t. $\Theta := \parallel g \parallel_\rho < 1$. It follows that $\parallel f^{-1} \parallel_\rho \leq \sum_{k \in \mathbb{N}} \Theta^k = \frac{1}{1-\Theta}$ holds, thus we have $f^{-1} \in B_\rho$.

Substitution of power series (1/4)

Remark

If $g_1, \ldots, g_n \in \mathbb{K}[\underline{Y}]$ then $\Phi_g : \begin{array}{ccc} \mathbb{K}[\underline{X}] & \longrightarrow & \mathbb{K}[\underline{Y}] \\ f & \longmapsto & f(g_1(\underline{Y}), \ldots, g_n(\underline{Y})) \end{array}$ defines a homomorphism of $\mathbb{K}$-algebras. This is not always true of convergent power series, e.g. $\mathbb{K}[[\underline{X}]] \longrightarrow \mathbb{K}[[\underline{Y}]], X_1, \ldots, X_n \longmapsto 1$.

Theorem 2

For $g_1, \ldots, g_n \in \mathbb{K}[[\underline{Y}]]$, with $\mathrm{ord}(g_i) \geq 1$, there is a $\mathbb{K}$-algebra homomorphism

$$\overline{\Phi_g} : \begin{array}{ccc} \mathbb{K}[[\underline{X}]] & \longrightarrow & \mathbb{K}[[\underline{Y}]] \\ f & \longmapsto & f(g_1(\underline{Y}), \ldots, g_n(\underline{Y})) \end{array}$$

with the following properties

1. If $g_1, \ldots, g_n$ are polynomials, then $\overline{\Phi_g}$ is an extension of Φ_g
2. If $g_1, \ldots, g_n$ are convergent power series, then we have $\overline{\Phi_g}(\mathbb{K}\langle \underline{X} \rangle) \subseteq \mathbb{K}\langle \underline{Y} \rangle$.

Substitution of power series (2/4)

Proof (1/3)

- Let $f \in \mathbb{K}[[\underline{X}]]$. To define $\overline{\Phi_g}(f)$, we consider the polynomial part $f^{(k)}$ of f, for all $k \in \mathbb{N}$.
- Since $\mathbb{K}[[\underline{Y}]]$ is a ring, we observe that $f^{(k)}(g_1, \ldots, g_n) \in \mathbb{K}[[\underline{Y}]]$ holds.
- Let $k, \ell \in \mathbb{N}$ with $k < \ell$. Observe that we have $\mathrm{ord}(f^{(\ell)} - f^{(k)}) \geq k + 1$.
- Since $\mathrm{ord}(g_i) \geq 1$ holds, we deduce $\mathrm{ord}(f^{(\ell)}(g) - f^{(k)}(g)) \geq k + 1$.
- It follows that $(f^{(k)}(g))_{k \in \mathbb{N}}$ is a Cauchy sequence in Krull Topology and thus converges to an element $f(g) \in \mathbb{K}[[\underline{X}]]$. Therefore, $\overline{\Phi_g}(f)$ is well defined.
- Of the properties asserted for the map $\overline{\Phi_g}$ only the second one requires some care.

Proof (2/3)

- Let $\rho = (\rho_1, \ldots, \rho_n) \in \mathbb{R}_{>0}^n$.
- It suffices to prove the following: there exists $\sigma = (\sigma_1, \ldots, \sigma_n) \in {\mathbb{R}_{>0}}^n$ such that we have
$$\overline{\Phi_g}(B_\rho) \subseteq B_\sigma.$$
- Since $g_j(0) = 0$ for all $j = 1 \cdots n$, there exists $\sigma_j \in {\mathbb{R}_{>0}}^n$ such that we have $\parallel g_j \parallel_{\sigma_j} \leq \rho_j$ for all $j = 1 \cdots n$.
- Taking the "component-wise min" of these $\sigma_j \in {\mathbb{R}_{>0}}^n$, we deduce the existence of a $\sigma \in {\mathbb{R}_{>0}}^n$ such that we have
$$\parallel g_j \parallel_\sigma \leq \rho_j$$
for all $j = 1 \cdots n$.
- It turns out that this σ has the desired property.

Substitution of power series (4/4)

Proof (3/3)

- Indeed, writing $f = \Sigma_e\, a_e X^e$, we have

$$\begin{array}{rcl} \| \, f^{(k)}(g) \, \|_\sigma & = & \| \sum_{d \leq k} f_{(k)}(g) \, \|_\sigma \\ & \leq & \sum_{d \leq k} \, \| \, f_{(k)}(g) \, \|_\sigma \\ & \leq & \sum_{d \leq k} \sum_{|e|=k} |a_e| \, \| \, g_1 \, \|_\sigma{}^{e_1} \cdots \, \| \, g_n \, \|_\sigma{}^{e_n} \\ & \leq & \sum_{d \leq k} \sum_{|e|=k} |a_e| {\rho_1}^{e_1} \cdots {\rho_n}^{e_n} \\ & = & \| \, f^{(k)} \, \|_\rho . \end{array}$$

- Thus, we have

$$\begin{array}{rcl} \| \, f(g) \, \|_\sigma & = & \lim_{k \to \infty} \, \| \, f^{(k)}(g) \, \|_\sigma \\ & \leq & \lim_{k \to \infty} \, \| \, f^{(k)} \, \|_\rho \\ & \leq & \| \, f \, \|_\rho . \end{array}$$

- Finally, we have

$$f \in B_\rho \quad \Rightarrow \quad f(g) \in B_\sigma .$$