MATH 8450 – LECTURE 3 – JAN 25, 2023

ABDELMALEK ABDESSELAM

Formal power series continued:

The ring of formal power series (FPS):

Again, let R be a ring (commutative with unit). The ring of formal power series $R[[X_1,\ldots,X_n]]$ over R, i.e., with coefficients in R, is defined as follows. As a set, this is just the set of all multisequences $(a_\alpha)_{\alpha\in\mathbb{N}^n}$ of elements of R. Namely,

$$R[[X_1,\ldots,X_n]] := R^{\mathbb{N}^n}\ .$$

This is almost like the previous ring of polynomials, except we do not impose the almost-finite condition. We will use the notation

$$(a_\alpha)_{\alpha\in\mathbb{N}^n} =: \sum_{\alpha\in\mathbb{N}^n} a_\alpha X^\alpha$$

as we did for polynomials. For an element $A=\sum_{\alpha\in\mathbb{N}^n} a_\alpha X^\alpha$, or $A(X)$, or $A(X_1,\ldots,X_n)$ in $R[[X_1,\ldots,X_n]]$, and for a multiindex γ, we let $[X^\gamma]A$ denote the coefficient of the monomial X^γ in A, i.e., $[X^\gamma]A := a_\gamma$.

We also define addition and multiplication in the same manner as before. Namely, we let

$$\left(\sum_{\alpha\mathbb{N}^n} a_\alpha X^\alpha\right) + \left(\sum_{\alpha\mathbb{N}^n} b_\alpha X^\alpha\right) := \sum_{\alpha\mathbb{N}^n} (a_\alpha+b_\alpha) X^\alpha\ .$$

Alternatively, one could have said, for all A, B in the ring of formal power series $R[[X]]$ (shorthand for $R[[X_1,\ldots,X_n]]$ if clear from the context), the sum is defined by letting

$$[X^\gamma](A+B) = ([X^\gamma]A) + ([X^\gamma]B)$$

for all $\gamma\in\mathbb{N}^n$. Multiplication is also defined as we did for formal polynomials:

$$\left(\sum_{\alpha\in\mathbb{N}^n} a_\alpha X^\alpha\right) \times \left(\sum_{\alpha\in\mathbb{N}^n} b_\alpha X^\alpha\right) := \sum_{\alpha\in\mathbb{N}^n} c_\alpha X^\alpha$$

where, for all multiindices $\gamma\in\mathbb{N}^n$,

$$c_\gamma = \sum_{\alpha,\beta\in\mathbb{N}^n} \mathbb{1}\{\alpha+\beta=\gamma\}\ a_\alpha b_\beta\ . \tag{1}$$

If one were to adhere to the uber uptight Bourbaki standards of rigor, the sum would be written as

$$\sum_{(\alpha,\beta)\in\mathbb{N}^n\times\mathbb{N}^n}$$

because we are summing over ordered pairs (α, β) of multiindices with n components. We will allow ourselves to be a little sloppy and write instead

$$\sum_{\alpha,\beta\in\mathbb{N}^n}$$

which should be clear enough. The sum defining the value of c_γ looks like an infinite double sum, but it is really a finite sum in disguise because only finitely many pairs (α, β) can contribute a nonzero value (in the ground ring R). More precisely,

$$|\{(\alpha, \beta) \in \mathbb{N}^n \times \mathbb{N}^n \mid \alpha + \beta = \gamma\}| = (\gamma_1 + 1)\cdots(\gamma_n + 1) < \infty$$

because this amounts to counting α's such that $0 \le \alpha \le \gamma$ where the inequalities are component-wise, i.e., mean $\forall i \in [n], 0 \le \alpha_i \le \gamma_i$.

Proposition 1. $(R[[X]], +, \times)$ *is a ring (commutative with unit).*

Exercise 1. *Prove this proposition.*

The ring $R[[X]] = R^{\mathbb{N}^n} = \prod_{\alpha\in\mathbb{N}^n} R$, a countable product of copies of the set R, can also be seen as a topological space. Namely, put the discrete topology on R, and then take the product topology. This immediately gives us a notion of convergent sequences in $R[[X]]$. Let $(S_k)_{k\ge0}$ be a sequence of elements in $R[[X]]$ and let ℓ be an element in $R[[X]]$. The sequence converges to ℓ, i.e., $\lim_{k\to\infty} S_k = \ell$ iff

$$\forall\alpha \in \mathbb{N}^n, \exists K \ge 0, \forall k \ge K, [X^\alpha]S_k = [X^\alpha]\ell\ .$$

If you know about topology, prove the above characterization, and if not, take it as a definition of convergence for sequences of FPS's.

Infinite sums in the ring of FPS's:

Let I be a set (possibly uncountable, be we will only use countable examples), and let $(S_i)_{i\in I}$ be a family of elements of $R[[X]]$ indexed by I. We will say that sum

$$\sum_{i\in I} S_i$$

converges and is equal to some element $S \in R[[X]]$ iff for all $\alpha \in \mathbb{N}^n$, there exists a finite subset J of I such that for all K finite set satisfying $J \subset K \subset I$, we have

$$[X^\alpha]\sum_{i\in K} S_i = [X^\alpha]S\ .$$

The definition might seem strange, but it is written that way so one does not need to invoke some ordering (order of summation) of the index set I. When I is countable and convergence holds, the sum does not depend on any order of summation.

Remark 1. *A cautionary tale is provided by Riemann's Series Rearrangement Theorem (see, e.g.,* [1]*) which says that if a series of real numbers is conditionally convergent (converges but is not absolutely convergent) then one can permute the terms in order to make the sum be equal to any real number we want! The above definition of convergent sums in $R[[X]]$ is a robust notion of convergence called unconditional convergence (often used as synonymous for convergence regardless of summation order).*

An equivalent way to phrase $\sum_{i\in I} S_i = S$ is to say that for all multiindex α, the sum $\sum_{i\in I}[X^\alpha]S_i$ has only finitely many nonzero terms and is equal to $[X^\alpha]S$.

For a multiindex α, we will see the monomial X^α as a particular FPS, namely, the one given by the collection of coefficients

$$(\mathbb{1}\{\gamma = \alpha\})_{\gamma\in\mathbb{N}^n} \ .$$

We can also define a scalar multiplication $\cdot : R \times R[[X]] \to R[[X]]$, in a coefficient-wise manner,

$$c \cdot \sum_{\alpha\in\mathbb{N}^n} a_\alpha X^\alpha := \sum_{\alpha\in\mathbb{N}^n} (ca_\alpha) X^\alpha \ .$$

With these last definitions, the expression

$$\sum_{\alpha\in\mathbb{N}^n} a_\alpha X^\alpha$$

can now also be understood as a series in $R[[X]]$. It is easy to see, that this series converges to the element of $R[[X]]$ given by $(a_\alpha)_{\alpha\in\mathbb{N}^n}$. In other words, what earlier was merely some funny notation can now be taken seriously as a sum, both interpretations being consistent.

A few more properties of FPS's: Contrary to polynomials, one does not have a way to plug elements of the ground ring R (later this will be $\mathbb{R}$ or $\mathbb{C}$) inside FPS's in order to get functions/maps $R^n \to R$. A notable exception is to plug zero, i.e., doing the substitutions $X_1 := 0_R, \ldots, X_n := 0_R$. This is just the evaluation at zero map $R[[X]] \to R$ given by

$$\sum_{\alpha\in\mathbb{N}^n} a_\alpha X^\alpha \longrightarrow a_{(0\ldots,0)} \ ,$$

i.e., the extraction of the constant term.

One also has a well defined notion of differentiation. For any $i \in [n]$, and for any $A = \sum_{\alpha\in\mathbb{N}^n} a_\alpha X^\alpha$ in $R[[X]]$, we let

$$\frac{\partial A}{\partial X_i} := \sum_{\substack{\alpha\in\mathbb{N}^n \\ |\alpha|\ge 1}} \alpha_i a_\alpha X^{\alpha - e_i}$$

where the last sum is a convergent series in the previous sense. Here e_i is notation for the canonical basis of $\mathbb{R}^n$ which contains $\mathbb{N}^n$.

An element a in a ring R is called invertible iff it has a multiplicative inverse, i.e., $\exists b \in R, ab = 1$. One has the following result about rings of FPS's.

Proposition 2. *An element A in the ring of FPS's $R[[X]] = R[[X_1, \ldots, X_n]]$ is invertible iff its constant term $[X^0]A$ is an invertible element of R.*

Exercise 2. *Prove the last proposition. Hint: use the formula $(1+x)^{-1} = 1-x+x^2-x^3+\cdots$. If this is too hard, read the next subsection on composition and then come back to finish this exercise.*

Composition:

We now arrive at the main core subtopic regarding FPS's which is about how to substitute FPS's inside other FPS's. This is a purely formal algebraic analogue of the notion of composition of multivariate maps. From now on $R = \mathbb{C}$ so we will be only talking about FPS's with complex (or real) coefficients.

Let $\mathbb{C}[[X]] = \mathbb{C}[[X_1, \ldots, X_n]]$ be the set of formal power series with complex coefficients with respect to the collection of variables $X_1, \ldots, X_n$. Let $\mathbb{C}[[Y]] = \mathbb{C}[[Y_1, \ldots, Y_p]]$ be the set of formal power series with complex coefficients with respect to another collection of variables $Y_1, \ldots, Y_p$. Let

$$f(X_1, \ldots, X_n) = f = \sum_{\alpha \in \mathbb{N}^n} f_\alpha X^\alpha \in \mathbb{C}[[X]]$$

and let $g_1, \ldots, g_n \in \mathbb{C}[[Y]]$, or to fix notations

$$g_i(Y) = g_i(Y_1, \ldots, Y_p) = \sum_{\beta \in \mathbb{N}^p} g_{i,\beta} Y^\beta$$

for all $i \in [n]$. Our goal is to define the FPS in $\mathbb{C}[[Y]]$ given by

$$h(Y_1, \ldots, Y_p) = f(g_1(Y_1, \ldots, Y_p), \ldots, g_n(Y_1, \ldots, Y_p)) \ .$$

The essential condition we will need is that the g's have no constant term:

$$\forall i \in [n], g_{i,0} = 0 \ .$$

The zero next to i is of course the zero multiindex $(0, \ldots, 0)$ with n components.

As suggested by the formula for f and the idea that X_i is replaced by $g_i(Y)$, we let

$$h := \sum_{\alpha \in \mathbb{N}^n} f_\alpha g_1(Y)^{\alpha_1} \cdots g_n(Y)^{\alpha_n} \ .$$

This is a series $\sum_{\alpha \in \mathbb{N}^n} S_\alpha$ in the ring of FPS's $\mathbb{C}[[Y]]$ with $S_\alpha := f_\alpha g_1(Y)^{\alpha_1} \cdots g_n(Y)^{\alpha_n}$. To show this is well defined, we need to show the following claim.

Claim: For all $\gamma \in \mathbb{N}^p$, only finitely many α's are such that $[Y^\gamma] S_\alpha \neq 0$.

Proof: We start by writing the product in full, i.e., writing $g_1(Y)^{\alpha_1} = g_1(Y) \cdots g_1(Y)$, α_1 times, etc. Then we insert the sum defining each such individual factor, making sure the (dummy) summation indices have different names for each factors.

$$\begin{aligned}
S_\alpha &= f_\alpha \left(\sum_{\beta(1,1) \in \mathbb{N}^p} g_{1,\beta(1,1)} Y^{\beta(1,1)} \right) \cdots \left(\sum_{\beta(1,\alpha_1) \in \mathbb{N}^p} g_{1,\beta(1,\alpha_1)} Y^{\beta(1,\alpha)} \right) \\
&\quad \times \left(\sum_{\beta(2,1) \in \mathbb{N}^p} g_{2,\beta(2,1)} Y^{\beta(2,1)} \right) \cdots \left(\sum_{\beta(2,\alpha_2) \in \mathbb{N}^p} g_{2,\beta(2,\alpha_2)} Y^{\beta(2,\alpha_2)} \right) \\
&\quad \vdots \\
&\quad \times \left(\sum_{\beta(n,1) \in \mathbb{N}^p} g_{n,\beta(n,1)} Y^{\beta(n,1)} \right) \cdots \left(\sum_{\beta(n,\alpha_n) \in \mathbb{N}^p} g_{n,\beta(n,\alpha_n)} Y^{\beta(n,\alpha_n)} \right) \\
&= f_\alpha \prod_{i=1}^{n} \left[\prod_{j=1}^{\alpha_i} \left(\sum_{\beta(i,j) \in \mathbb{N}^p} g_{i,\beta(i,j)} Y^{\beta(i,j)} \right) \right] \ .
\end{aligned}$$

Note that for each chosen values of i and j, $\beta(i,j)$ is a full multiindex with p components

$$\beta(i,j) = (\beta(i,j)_1, \ldots, \beta(i,j)_p)$$

where the $\beta(i,j)_k$ are nonnegative integers.

If you did Exercise 1, in particular the associativity of multiplication, you know how to expand a product of three sums, as a consequence of the definition (1). Go ahead and generalize that to a product of $\alpha_1 + \cdots + \alpha_n = |\alpha|$ sums, as in the expression for S_α. After the dust settles, we get

$$S_\alpha = f_\alpha \sum_{\mathbb{B}} \left[\prod_{i=1}^{n} \left(\prod_{j=1}^{\alpha_i} g_{i,\beta(i,j)} \right) \right] \times Y^{\sum_{i=1}^{n} \left(\sum_{j=1}^{\alpha_i} \beta(i,j) \right)}$$

where $\mathbb{B}$ is shorthand for a rather complicated data structure. Namely,

$$\mathbb{B} = (\beta(i,j))_{\substack{1 \le i \le n \\ 1 \le j \le \alpha_i}} \in (\mathbb{N}^p)^{\alpha_1 + \cdots + \alpha_n}$$

or equivalently

$$\mathbb{B} = (\beta(1,1), \ldots, \beta(1,\alpha_1); \ldots ; \beta(n,1), \ldots, \beta(n,\alpha_n)) \ .$$

This is like a two dimensional array/spreadsheet, where the rows can have different lengths and the cells are filled with mutiindices. A better handling of these data structures can be done as follows. For a given multiindex $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ we define a set of points with integer coordinates in the plane, or diagram $\mathbb{D}_\alpha$ by

$$\mathbb{D}_\alpha := \{(i,j) \in [n] \times \mathbb{Z}_{>0} \mid 1 \le j \le \alpha_i\} \ .$$

As in matrix algebra/Excel, it is good to draw the i coordinate on a vertical axis numbered from 1 to n from top to bottom, and j on a horizontal axis numbered from 1 and increasing from left to right. Elements of $\mathbb{D}_\alpha$ indicate location of cells in a spreadsheet. These cells will then be filled with contents or values which are multiindices. An object previously denoted by $\mathbb{B}$ is thus the same thing as a map $\mathbb{D}_\alpha \to \mathbb{N}^p$. Hence,

$$S_\alpha = f_\alpha \sum_{\mathbb{B} \in (\mathbb{N}^p)^{\mathbb{D}_\alpha}} \left[\prod_{i=1}^{n} \left(\prod_{j=1}^{\alpha_i} g_{i,\beta(i,j)} \right) \right] \times Y^{\sum_{i=1}^{n} \left(\sum_{j=1}^{\alpha_i} \beta(i,j) \right)} \ .$$

Remark 2. *We have seen in Lecture 2 the notion of set partition. There is also the notion of integer partition which is an important concept featuring in number theory, combinatorics and representation theory. It is the same thing as a multiindex α with the extra condition that the components are nonincreasing $\alpha_1 \ge \alpha_2 \ge \cdots$. If α is such a partition, then $\mathbb{D}_\alpha$ is called the Ferrers diagram of that integer partition, where locations (i,j) are represented by square boxes. A filling of such a diagram by integers, i.e., a map like $\mathbb{B}$ with $p = 1$, is called a Young tableau. If you studied QCD (Quantum Chromodynamics, the theory of strong interactions between quarks) which involves the representation theory of the group $SU(3)$, you have probably seen such diagrams and tableaux.*

We can now express, for a given Y monomial, the corresponding coefficient in S_α. Namely, for any $\gamma \in \mathbb{N}^p$ and any $\alpha \in \mathbb{N}^n$, we have

$$[Y^\gamma] S_\alpha = f_\alpha \sum_{\mathbb{B} \in (\mathbb{N}^p)^{\mathbb{D}_\alpha}} \left[\prod_{i=1}^{n} \left(\prod_{j=1}^{\alpha_i} g_{i,\beta(i,j)} \right) \right] \times \mathbb{1}\left\{ \sum_{i=1}^{n} \left(\sum_{j=1}^{\alpha_i} \beta(i,j) \right) = \gamma \right\} \ .$$

Suppose $[Y^\gamma] S_\alpha \neq 0$. Then, there exists $\mathbb{B}$, for which the corresponding term/summand is nonzero. This implies that for all $(i,j) \in \mathbb{D}_\alpha$, $g_{i,\beta(i,j)} \neq 0$, and therefore $|\beta(i,j)| \ge 1$ (this is the crux of the proof with the use of the zero constant term hypothesis for the g's). Thus, using the additivity of lengths of multiindices,

$$|\gamma| = \left| \sum_{i=1}^{n} \left(\sum_{j=1}^{\alpha_i} \beta(i,j) \right) \right| = \sum_{i=1}^{n} \left(\sum_{j=1}^{\alpha_i} |\beta(i,j)| \right) \geq \sum_{i=1}^{n} \left(\sum_{j=1}^{\alpha_i} 1 \right) = |\mathbb{D}_\alpha| = |\alpha| \ .$$

Since ony finitely many α's satisfy $|\alpha| \leq |\gamma|$, we proved that only finitely many α's are such that $[Y^\gamma]S_\alpha \neq 0$. □

This concludes the proof that composition is well defined. Next lecture, we will prove the Inverse Function Theorem for formal power series.

References

[1] S. Galanor, Riemann's rearrangement theorem. Mathematics Teacher **80** (1987), no. 8, 675–681. Available at `https://sites.math.washington.edu/~morrow/335_16/history%20of%20rearrangements.pdf` , last retrieved Jan 26, 2023.