Lecture 3. $C_n = O(c^n)$ from first principles. M. Artin's Approximation Theorem (no proof)

M. Klazar

March 8, 2024

Enumerative combinatorics literature takes for granted that any algebraic FPS $f = f(x) = \sum_{n=0}^{\infty} a_n x^n \in \mathbb{C}[[x]]$ is convergent, meaning that $a_n = O(c^n)$ for all $n \in \mathbb{N}_0$ and some constant $c > 1$. Rigorous justification of this fact from first principles by a formal algebraic argument, without recourse to theories of analytic functions and/or algebraic curves, is seldom given. We plan to fill this gap too, like the one concerning the justification of the equality $\sqrt{1-4x} = (1-4x)^{1/2}$ in $\mathbb{R}[[x]]$. We start, of course, with our running example of Catalan numbers.

Proposition 1 *If the* FPS $f = f(x) = \sum_{n=0}^{\infty} a_n x^n \in \mathbb{R}[[x]]$ *solves the equation* $f^2 - f + x = 0$ *then* $|a_1| = 1$ *and for* $n \geq 2$,

$$|a_n| = \sum_{k=1}^{n-1} |a_k| \cdot |a_{n-k}| \ .$$

Proof. We know that the two solutions of this equation are $f_1(x) = \sum_{n\geq 1} C_n x^n$ and $f_2(x) = 1 + \sum_{n\geq 1}(-C_n)x^n$. The recurrence follows. $\square$

The generic solution of the equation $f^2 - f + x = 0$ satisfies for $n \geq 2$ the relation

$$a_n = \sum_{k=0}^{n} a_k \cdot a_{n-k} \ .$$

In the first solution $a_0 = 0$ and $a_n = C_n$ for $n > 0$; in the second solution $a_0 = 1$ and $a_n = -C_n$ for $n > 0$. Both solutions yield the same recurrence $C_1 = 1$ and $C_n = \sum_{k=1}^{n-1} C_k C_{n-k}$ if $n \geq 2$.

Proposition 2 *If* $a_1 = 1$ *and* $a_n = \sum_{k=1}^{n-1} a_k a_{n-k}$ *for* $n \geq 2$ *then for every* $n \in \mathbb{N}$,

$$a_n \leq \frac{8^{n-1}}{n^2} \ .$$

For the proof we need (again) a convolution result.

Lemma 3 *For every $n \geq 2$,*

$$S(n) = \sum_{k=1}^{n-1} \frac{1}{k^2(n-k)^2} < \frac{8}{n^2} \,.$$

Proof. For $k, n \in \mathbb{N}$ and $k < n$,

$$\frac{1}{k(n-k)} = \frac{1}{n}\left(\frac{1}{k} + \frac{1}{n-k}\right) .$$

Substituting it for the summand in $S(n)$ we get that

$$S(n) = \frac{2}{n^2}\sum_{k=1}^{n-1}\frac{1}{k^2} + \frac{2}{n^2}\sum_{k=1}^{n-1}\frac{1}{k(n-k)} = \frac{2}{n^2}\underbrace{\sum_{k=1}^{n-1}\frac{1}{k^2}}_{<2} + \frac{4}{n^3}\underbrace{\sum_{k=1}^{n-1}\frac{1}{k}}_{<n} < \frac{8}{n^2}$$

because $\sum_{k=2}^{\infty}\frac{1}{k^2} < \sum_{k=2}^{\infty}\left(\frac{1}{k-1} - \frac{1}{k}\right) = 1$. □

Proof of Proposition 2. We proceed by induction on n. For $n = 1$ the bound holds as an equality. For $n \geq 2$ we get by induction and with the help of Lemma 3 that

$$a_n = \sum_{k=1}^{n-1} a_k \cdot a_{n-k} \leq \sum_{k=1}^{n-1}\frac{8^{k-1}}{k^2}\cdot\frac{8^{n-k-1}}{(n-2)^2} = 8^{n-2}\sum_{k=1}^{n-1}\frac{1}{k^2(n-k)^2} < \frac{8^{n-1}}{n^2} \,.$$

□

Thus we have this corollary.

Corollary 4 *If $f \in \mathbb{C}[[x]]$ satisfies the equation $f^2 - f + x = 0$ then for all $n \in \mathbb{N}_0$,*

$$[x^n]f(x) = O(8^n) \,.$$

We aim to greatly generalize this simple bound, in two directions. First we replace univariate FPS $K[[x]]$, where K is $\mathbb{R}$ or $\mathbb{C}$, with multivariate FPS

$$K[[X]] = \left\{ \textstyle\sum_{\overline{k}\in\mathbb{N}_0^m} a_{\overline{k}} X^{\overline{k}} \;\middle|\; a_{\overline{k}} \in K \right\}$$

where $X = x_1, \dots, x_m$, $m \in \mathbb{N}$, $\overline{k} = (k_1, \dots, k_m)$ and $X^{\overline{k}} = x_1^{k_1} \dots x_m^{k_m}$. A multivariate FPS $f = f(X) \in K[X]$ is *convergent* if there is a $c > 0$ such that the series $f(c, c, \dots, c)$ absolutely converges; in terms of coefficients it means that

$$a_{\overline{k}} = O(c^{|\overline{k}|}) = O(c^{k_1+\cdots+k_m})$$

for all $\overline{k} \in \mathbb{N}_0^m$ and some $c > 1$. Second, we extend the domain of coefficients in the polynomial equation for f from the polynomial ring $K[X] = K[x_1, \dots, x_m]$ to the ring $K\{X\} \subset K[[X]]$ of convergent multivariate FPS. Our goal is to prove the following general exponential upper bound.

Corollary 5 (M. Artin) *Let $m, p \in \mathbb{N}$, $X = x_1, \ldots, x_m$, K be $\mathbb{R}$ or $\mathbb{C}$, $a_0(X)$, $\ldots$, $a_p(X)$ be in $K\{X\}$, $a_0 \neq 0$ and let $f(X) \in K[[X]]$ be such that*

$$\sum_{j=0}^{p} a_{p-j}(X) \cdot f(X)^j = 0 \ .$$

Then $f(X) \in K\{X\}$ as well, $f(X)$ is convergent.

This result follows from M. Artin's ([3]) *Approximation Theorem* [1] which we quote here partially from the textbook [5, Chapter V.3] and also from the original article [1]. The theorem approximately solves not a single polynomial equation but a system of several analytic equations. Again K is $\mathbb{R}$ or $\mathbb{C}$ (or a field with nontrivial valuation). For any nonzero $f = f(x) = \sum_{\overline{k} \in \mathbb{N}_0^m} a_{\overline{k}} X^{\overline{k}} \in K[[X]]$ we set

$$\mathrm{ord}(f) = \min(\{|\overline{k}| \mid \overline{k} \in \mathbb{N}_0^m \wedge a_{\overline{k}} \neq 0\}) \in \mathbb{N}_0$$

($|\overline{k}| = k_1 + \cdots + k_m$) and set $\mathrm{ord}(0) := +\infty$.

Theorem 6 (M. Artin) *Let $n, m, k, l \in \mathbb{N}$, $X = x_1, \ldots, x_n$, $Y = y_1, \ldots, y_m$, let $F_1, \ldots, F_k$ be in $K\{X, Y\}$, and let $f_1, \ldots, f_m$ in $K[[X]]$ have zero constant terms and be such that*

$$F_1(X, f_1(X), \ldots, f_m(X)) = 0, \ldots, F_k(X, f_1(X), \ldots, f_m(X)) = 0 \ .$$

Then there exist $g_1, \ldots, g_m$ in $K\{X\}$ such that $\mathrm{ord}(f_1(X) - g_1(X)) \geq l$, $\ldots$, $\mathrm{ord}(f_m(X) - g_m(X)) \geq l$ *and also*

$$F_1(X, g_1(X), \ldots, g_m(X)) = 0, \ldots, F_k(X, g_1(X), \ldots, g_m(X)) = 0 \ .$$

The requirement that constant terms of f_i are zero means that we substitute f_i for y_i in $F_1, \ldots, F_k$ formally, coefficients resulting in the substitution are given by finite expressions. Note that the F_j have to have zero constant terms as well. The theorem says that any formal solution f_i, $i \in [m]$, of any system $F_j = 0$, $j \in [k]$, of analytic equations can be approximated to any order l by a convergent solution g_i.

How does Corollary 5 follow from Theorem 6? This is explained in [5, p. 106]. A problem is that the FPS $f(X)$ in Corollary 5 may have a nonzero constant term, like the solution $1 + \sum_{n \geq 1}(-C_n)x^n$ of $f^2 - f + x = 0$, and then Theorem 6 does not apply. But this is easy to fix. We change the variable y to $f(X) = g(X) + c$ with $c \in K$ such that $g(X)$ has zero constant term; then $g(x)$ solves a new polynomial equation. Suppose that $g(X)$ is not convergent. Using Theorem 6 we can get a convergent solution of the new polynomial equation that approximates $g(x)$ to any given precision but is (necessarily) distinct from $g(x)$. Increasing the precision we could generate infinitely many distinct (convergent) solutions of the new polynomial equation. But this contradict the bound on the number of roots of a polynomial over a ring. Thus $g(X)$ and $f(X)$ are convergent.

This is all very nice but it would be nice to have a proof of Corollary 5 via a direct and elementary argument without using the machinery of M. Artin's Approximation Theorem, something in the spirit of the proof of Corollary 4. In our lectures we will attempt to present such a proof. It was given by R. Pierzchała in [4] and is based on a strengthening of a lemma in [2].

References

[1] M. Artin, On the solutions of analytic equations, *Invent. Math.* **5** (1968), 277–291

[2] A. M. Gabrielov, Formal relations among analytic functions, *Mathematics of the USSR-Izvestiya* **7** (1973), 1056–1088

[3] Michael Artin, article in Wikipedia, `https://en.wikipedia.org/wiki/Michael_Artin`

[4] R. Pierzchała, On roots of polynomials with power series coefficients, *Annales Polonici Mathematici* **80** (2003), 211–217

[5] J. M. Ruiz, *The Basic Theory of Power Series*, Vieweg, Braunschweig 1993