Lecture 4. Three proofs that (C_n) is not a linear recurrence sequence

M. Klazar

March 15, 2024

I am not yet prepared to prove Corollary 5 in the previous lecture and therefore instead I survey (and, where possible, improve) the article [1]. In this article we (MK and RH) gave four proofs that the sequence $(C_n) = (1, 1, 2, 5, 14, 42, \dots)$ of Catalan numbers is not a *linear recurrence sequence* (LRS). In this lecture we go through three of them. The last fourth proof will be given in the next lecture. We begin with the definition of the class of sequences called LRS.

Definition 1 *Let K be a field. We say that a sequence $(a_n) = (a_1, a_2, \dots) \subset K$ is a linear recurrence sequence (over K), briefly* LRS*, if there exist $k \in \mathbb{N}_0$ constants $c_0, \dots, c_{k-1}$ in K such that for every $n \in \mathbb{N}$ the relation*

$$a_{n+k} = \sum_{j=0}^{k-1} c_j a_{n+k-1-j}$$

holds. For $k = 0$ we define the sum as 0.

Thus for $k = 0$ we have the *zero sequence* $(a_n) = (0_K, 0_K, \dots)$. A well known LRS (over $\mathbb{Q}$) is the sequence (F_n) of *Fibonacci numbers* given by $F_1 = F_2 = 1$ and $F_{n+2} = F_{n+1} + F_n$, thus

$$(F_n) = (1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ 89,\ 144,\ \dots)\,.$$

Note that in Definition 1, relaxing "for every $n \in \mathbb{N}$" to "for every $n \ge n_0$" does not yield new LRS: if (a_n) satisfies the recurrence for every $n \ge n_0$ then it satisfies for every $n \in \mathbb{N}$ the extended recurrence obtained by replacing k with $k + n_0 - 1$ and adding the dummy coefficients $c_k = c_{k+1} = \dots = c_{k-2+n_0} = 0_K$.

So we give four proofs of the following theorem; three in this lecture and the fourth one in the next lecture.

Theorem 2 *The sequence $(C_n) = (1, 1, 2, 5, 14, 42, \dots)$ is not* LRS *over $\mathbb{C}$.*

Proof 1 (by generating functions)

This proof is based on the following well known (? — not quite in this formulation) characterization of generating functions of LRS.

Proposition 3 *Let K be a field, $(a_n) \subset K$ and $A(x) = \sum_{n=1}^{\infty} a_n x^n$ in $K[[x]]$ be the* GF *of (a_n). Then*

$$(a_n) \text{ is LRS over } K \iff \exists\, p,\, q \in K[x]:\ q(0_K) = 1_K \wedge q(x)A(x) = p(x)\,.$$

Proof. The implication $\Rightarrow$. Let $a_{n+k} = \sum_{j=0}^{k-1} c_j a_{n+k-1-j}$ for every $n \in \mathbb{N}$ and some $k \in \mathbb{N}_0$ constants $c_0, \ldots, c_{k-1}$ in K. We set $q(x) = 1_K - \sum_{j=0}^{k-1} c_j x^{j+1}$ (again, empty sum is 0_K). Then $q(0_K) = 1_k$ and from the recurrence we see that for every $n \in \mathbb{N}$,

$$[x^{n+k}]\, q(x)A(x) = a_{n+k} - \sum_{j=0}^{k-1} c_j a_{n+k-1-j} = 0_K\,.$$

Hence $[x^n]\, q(x)A(x) = 0_K$ for every $n > k$, which means that $q(x)A(x) = p(x)$ for some polynomial $p \in K[x]$ with degree at most k (we define the degree of the zero polynomial as $-\infty$).

The implication $\Leftarrow$. Let $q(x)A(x) = p(x)$ where p and q are as stated. Let $q(x) = 1_K - \sum_{j=0}^{k-1} c_j x^{j+1}$ where $k \in \mathbb{N}_0$ and $c_j \in K$, and let $l = \deg p(x)$. Clearly,

$$n \geq n_0 = \max(\{1,\, l-k+1\}) \Rightarrow a_{n+k} - \sum_{j=0}^{k-1} c_j a_{n+k-1-j} = [x^{n+k}]\, q(x)A(x) = 0_K\,.$$

Thus the relation $a_{n+k} = \sum_{j=0}^{k-1} c_j a_{n+k-1-j}$ holds for every $n \geq n_0$ and by the above remark the sequence (a_n) is LRS over K. $\square$

Now suppose for the contrary that (C_n) is LRS over $\mathbb{C}$. By Proposition 3 there exist $p, q \in \mathbb{C}[x]$ such that $q(0) = 1$ and

$$q(x)C(x) = q(x) \cdot \frac{1 - \sqrt{1-4x}}{2} = p(x)\,.$$

Here $\sqrt{1-4x} \in \mathbb{C}[[x]]$ is determined by the conditions that $(\sqrt{1-4x})(0) = 1$ and $\left(\sqrt{1-4x}\right)^2 = 1 - 4x$; this suffices, we do not need the explicit formula for the coefficients. Thus

$$q(x) \cdot \sqrt{1-4x} = p_0(x)$$

where $q(x)$ is as before and $p_0(x) = q(x) - 2p(x) \in \mathbb{C}[x]$. Clearly, $p_0(x) \neq 0$. But then

$$q(x)^2 \cdot (1-4x) = p_0(x)^2$$

which is impossible because the polynomial on the left side has an odd degree but the polynomial on the right side has an even degree. $\square$

Proof 2 (by elementary number theory)

For this proof we need two auxiliary results. The first one characterizes odd Catalan numbers.

Proposition 4 *For every $n \in \mathbb{N}$,*

$$C_n \text{ is odd} \iff n = 2^l \text{ for some } l \in \mathbb{N}_0\,.$$

Proof. This follows from the recurrence that $C_1 = 1$ and $C_n = \sum_{j=1}^{n-1} C_j C_{n-j}$ for $n \ge 2$, and from expressing any $n \in \mathbb{N}$ as $n = 2^l m$ where $l \in \mathbb{N}_0$ and $m \in \mathbb{N}$ is odd. If $n > 1$ is odd then $C_n = 2\sum_{j=1}^{(n-1)/2} C_j C_{n-j}$ is even; $C_1 = 1$ is odd. If n is even then

$$C_n = C_{n/2}^2 + 2\sum_{j=1}^{n/2-1} C_j C_{n-j}$$

and, modulo 2, $C_n \equiv C_{n/2} \equiv \cdots \equiv C_m$. Let n be even. We see that if $n = 2^l$ then $C_n \equiv C_1 \equiv 1$, and if $n = 2^l m$ with odd $m > 1$ then $C_n \equiv C_m \equiv 0$. □

The second auxiliary result is an important theoretical result about LRS in general. The proof is a nice application of linear algebra, see [1] for it.

Proposition 5 *Let $K \subset L$ be an extension of fields and $(a_n) \subset K$ be* LRS *over L. Then (a_n) is* LRS *over K.*

Suppose again for the contrary that (C_n) is LRS over $\mathbb{C}$. By Proposition 5, (C_n) is LRS over $\mathbb{Q}$. Thus we have for some $k \in \mathbb{N}_0$ fractions $c_0 = 1$, c_1, …, c_k and every $n \in \mathbb{N}$ that

$$\sum_{j=0}^{k} c_j C_{n+k-j} = 0\,.$$

We multiply this equation by an $m \in \mathbb{N}$ such that then all $c_j := mc_j \in \mathbb{Z}$ and are mutually coprime (if $d \in \mathbb{N}$ divides all c_j then $d = 1$). In particular at least one of them, let us call it c_l, is odd. We use Proposition 4 and select $N \in \mathbb{N}$ such that for every $j \in \{0, 1, \ldots, k\}$, the number C_{N+k-j} is odd iff $j = l$. But then

$$\sum_{j=0}^{k} c_j C_{N+k-j} = 0$$

is an impossible equality because the left side is odd (exactly one summand, for $j = l$, is odd) but the right side is even. □

Proof 3 (by polynomials)

This proof is based on the next well known result in algebra.

Proposition 6 *Let R be an integral domain and let $p(x)$ in $R[x]$ be a nonzero polynomial. Then*

$$|Z(p)| := |\{a \in R \mid p(a) = 0_R\}| \le \deg p(x)\,.$$

Proof. We proceed by induction on the degree $d = \deg p(x)$. If $d = 0$ then $Z(p) = \emptyset$ and the claim holds. Suppose that $d > 0$. If still $Z(p) = \emptyset$, the claim holds. Else we take a root of $p(x)$ $a \in Z(p)$ and by dividing $p(x)$ by $x - a = 1_R x - a$ with remainder we express $p(x)$ as

$$p(x) = (x - a)q(x) + r$$

where $q(x) \in R[x]$ and $r \in R$. By setting $x = a$ we see that $r = 0_R$ and so $p(x) = (x - a)q(x)$. Since R is an integral domain,

$$Z(p) = \{a\} \cup Z(q)\,.$$

But $\deg q(x) = d - 1$ and thus induction gives that

$$|Z(p)| \le 1 + |Z(q)| \le 1 + d - 1 = d\,.$$

$\square$

The idea of the third proof is to take the relation

$$\sum_{j=0}^{k} a_j C_{n+j} = 0, \quad (n \in \mathbb{N})\,,$$

where $k \in \mathbb{N}_0$, $a_k = 1$ and all $a_j \in \mathbb{C}$, substitute for the Catalan numbers in it the expressions $C_m = \frac{1}{m}\binom{2m-2}{m-1}$, and by clearing denominators and common factors obtain a relation saying that a complex polynomial has infinitely many roots $n \in \mathbb{N}$. If the polynomial is non-zero, we get a contradiction.

We use the polynomials $(x)_k = x(x-1)\dots(x-k+1)$ for $k \in \mathbb{N}$, and $(x)_0 = 1$. In the above displayed relation we set

$$C_{n+j} = \frac{1}{n+j} \cdot \frac{(2n+2j-2)!}{(n+j-1)!^2}\,,$$

multiply the result by

$$(n+k)_{k+1} \cdot \frac{(n+k-1)!^2}{(2n-2)!}$$

and get that for every $n \in \mathbb{N}$,

$$p(n) := \sum_{j=0}^{k} a_j \cdot \frac{(n+k)_{k+1}}{n+j} \cdot (n+k-1)_{k-j}^2 \cdot (2n+2j-2)_{2j} = 0\,.$$

If we regard n as a formal variable then $p(n)$ is a polynomial in $\mathbb{C}[n]$. It is nonzero because

$$p(-k) = 1 \cdot (-1)_k \cdot 1 \cdot (-2)_{2k} \neq 0\,.$$

Thus we have a contradiction with Proposition 6 because $\deg p(n) \le 3k$ and thus may have only at most $3k$ roots. $\square$

References

[1] M. Klazar and R. Horský, Are the Catalan numbers a linear recurrence sequence?, *Amer. Math. Monthly* **129** (2022), 166–171