Journal of Integer Sequences, Vol. 12 (2009),
Article 09.7.5

Generalized Catalan Numbers: Linear Recursion and Divisibility

B. Sury
Stat-Math Unit
Indian Statistical Institute
8th Mile Mysore Road
Bangalore 560059
India
sury@isibang.ac.in

Abstract

We prove a *linear* recursion for the generalized Catalan numbers $C_a(n) := \frac{1}{(a-1)n+1}\binom{an}{n}$ when $a \geq 2$. As a consequence, we show $p \mid C_p(n)$ if and only if $n \neq \frac{p^k-1}{p-1}$ for all integers $k \geq 0$. This is a generalization of the well-known result that the usual Catalan number $C_2(n)$ is odd if and only if n is a Mersenne number $2^k - 1$. Using certain beautiful results of Kummer and Legendre, we give a second proof of the divisibility result for $C_p(n)$. We also give suitably formulated inductive proofs of Kummer's and Legendre's formulae which are different from the standard proofs.

1 Introduction

The Catalan numbers $\frac{1}{n+1}\binom{2n}{n}$ arise in diverse situations like counting lattice paths, counting rooted trees etc. In this note, we consider for each natural number $a \geq 2$, generalized Catalan numbers (referred to henceforth as GCNs) $C_a(n) := \frac{1}{(a-1)n+1}\binom{an}{n}$ and give a *linear* recursion for them. Note that $a = 2$ corresponds to the Catalan numbers. The linear recursion seems to be a new observation. We prove the recursion by a suitably formulated induction. This new recursion also leads to a divisibility result for $C_p(n)$'s for a prime p and, thus also, to another proof of the well-known parity assertion for the usual Catalan numbers. The latter asserts $C_2(n)$ is odd if and only if n is a Mersenne number; that is, a number of the form $2^k - 1$ for some positive integer k. Using certain beautiful results of Kummer and Legendre, we give a second proof of the divisibility result for $C_p(n)$. We also give suitably formulated inductive proofs of Kummer's and Legendre's formulae mentioned below. This is different from the standard proofs [2] and [3]. In this paper, the letter p always denotes a prime number.

2 Linear recursion for GCNs

Lemma 1. *For any $a \geq 2$, the numbers $C_a(n) = \frac{1}{(a-1)n+1}\binom{an}{n}$ can be defined recursively by*

$$C_a(0) = 1$$

$$C_a(n) = \sum_{k=1}^{\lfloor \frac{(a-1)n+1}{a} \rfloor} (-1)^{k-1} \binom{(a-1)(n-k)+1}{k} C_a(n-k) \ , n \geq 1.$$

In particular, the usual Catalan numbers $C_2(n)$ satisfy the linear recursion

$$C_2(n) = \sum_{k=1}^{\lfloor \frac{n+1}{2} \rfloor} (-1)^{k-1} \binom{n-k+1}{k} C_2(n-k) \ , n \geq 1.$$

2.1 A definition and remarks

Before proceeding to prove the lemma, we recall a useful definition. One defines the *forward difference operator* Δ on the set of functions on $\mathbb{R}$ as follows. For any function f, the new function Δf is defined by

$$(\Delta f)(x) := f(x+1) - f(x).$$

Successively, one defines $\Delta^{k+1} f = \Delta(\Delta^k f)$ for each $k \geq 1$. It is easily proved by induction on n (see, for instance [1, pp. 102–103]) that

$$(\Delta^n f)(x) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} f(x+n-k).$$

We note that if f is a polynomial of degree d, then Δf is also a polynomial and has degree $d-1$. In particular, $\Delta^N f \equiv 0$, the zero function, when $N > d$. Therefore, $(\Delta^N f)(0) = 0$.

Proof of 1. The asserted recursion can be rewritten as

$$\sum_{k \geq 0} (-1)^k \binom{n}{k} \binom{a(n-k)}{n-1} = 0.$$

One natural way to prove such identities is to try and view the sum as $(\Delta^n f)(0)$ for a polynomial f of degree $< n$. In our case, we may take $f(x) = ax(ax-1)\cdots(ax-n+2)$ which is a polynomial of degree $< n$. Then,

$$(\Delta^n f)(x) = \sum_{k \geq 0} (-1)^k \binom{n}{k} f(x+n-k) \equiv 0.$$

This gives

$$(\Delta^n f)(0) = \sum_{k\geq 0}(-1)^k \binom{n}{k}\binom{a(n-k)}{n-1} = 0.$$

Thus the asserted recursion follows. ■

Using this lemma, we have the following:

Theorem 2. *The prime $p \,|\, C_p(n)$ if and only if $n \neq \frac{p^k-1}{p-1}$ for all integers $k \geq 0$. In particular, $C_2(n)$ is odd if and only if n is a Mersenne number $2^k - 1$.*

Proof. We shall apply induction on n. The result holds for $n = 1$ since $C_p(1) = 1$. Assume $n > 1$ and that the result holds for all $m < n$. Let $p^r \leq n \leq p^{r+1} - 1$. Let us read the right hand side of

$$C_p(n) = \sum_{k=1}^{\lfloor \frac{(p-1)n+1}{p} \rfloor} (-1)^{k-1}\binom{(p-1)(n-k)+1}{k} C_p(n-k)$$

modulo p. We use the induction hypothesis that for $m < n$, $C_p(m)$ is a multiple of p whenever $(p-1)m+1$ is not a power of p. Modulo p, the terms in the above sum which are non-zero are those for which $n-k$ is of the form $\frac{p^N-1}{p-1}$. But, since $p^r \leq n < p^{r+1}$, the only non-zero term modulo p is the one corresponding to the index k for which $(p-1)(n-k) = p^r - 1$ if $n \leq \frac{p^{r+1}-1}{p-1}$ (respectively, $(p-1)(n-k) = p^{r+1} - 1$ if $n > \frac{p^{r+1}-1}{p-1}$). This term is, to within sign, $\binom{p^r}{n-\frac{p^r-1}{p-1}} C_p(\frac{p^r-1}{p-1})$ if $n \leq \frac{p^{r+1}-1}{p-1}$ (respectively, $\binom{p^{r+1}}{n-\frac{p^{r+1}-1}{p-1}} C_p(\frac{p^{r+1}-1}{p-1})$ if $n > \frac{p^{r+1}-1}{p-1}$). As the binomial coefficient $\binom{p^r}{s}$ is a multiple of p if and only if $0 < s < p^r$, the above term is a multiple of p if and only if $0 < n - \frac{p^r-1}{p-1} < p^r$ if $n \leq \frac{p^{r+1}-1}{p-1}$ (respectively, $0 < n - \frac{p^{r+1}-1}{p-1} < p^{r+1}$ if $n > \frac{p^{r+1}-1}{p-1}$). This is equivalent to $p^r < (p-1)n+1 < p^{r+1}$ if $n \leq \frac{p^{r+1}-1}{p-1}$ (respectively, $p^{r+1} < (p-1)n+1 < p^{r+2}$ if $n > \frac{p^{r+1}-1}{p-1}$), which means that $(p-1)n+1$ is not a power of p. The theorem is proved. □

3 Another proof of Theorem using Kummer's algorithm

Kummer proved that, for $r \leq n$, the p-adic valuation $v_p(\binom{n}{r})$ is simply the number of carries when one adds r and $n-r$ in base-p. We give another proof of Theorem 2 now using Kummer's algorithm.

3.1 Another proof of Theorem 2

Write the base-p expansion of $(p-1)n+1$ as

$$(p-1)n+1 = a_s \cdots a_{r+1} 0 \cdots 0$$

say, where $a_{r+1} \neq 0, s \geq r+1$ and $r \geq 0$. Evidently, $v_p((p-1)n+1) = r$. Thus, unless $(p-1)n+1$ is a power of p, the base-p expansion of $(p-1)n$ will have the same number of digits as above. It is of the form

$$(p-1)n = * \cdots * \, (a_{r+1}-1) \, \underbrace{(p-1) \cdots (p-1)}_{r \text{ times}}$$

where $a_{r+1}-1$ is between 0 and $p-2$. So, the base-p expansion of n itself looks like

$$n = * \cdots * \, 1 \cdots 1$$

with r ones at the right end. Also, there are at least r carries coming from the right end while adding the base-p expansions of n and $(p-1)n$. Moreover, unless $(p-1)n+1$ is a power of p, consider the first non-zero digit to the left of the string of $(p-1)$'s at the end of the expansion of $(p-1)n$. If it is denoted by u, and the corresponding digit for n is v, then $(p-1)v \equiv u \pmod{p}$; that is, $u+v$ is a non-zero multiple of p (and therefore $\geq p$). Thus, there are at least $r+1$ carries coming from adding the base-p expansions of n and $(p-1)n$ unless $(p-1)n+1$ is a power of p. This proves Theorem 2 again. ■

4 Kummer and Legendre's formulae inductively

Legendre observed that $v_p(n!)$ is $\frac{n-s(n)}{p-1}$ where $s(n)$ is the sum of the digits in the base-p expansion of n. In [2], Honsberger deduces Kummer's theorem (used in the previous section) from Legendre's result and refers to Ribenboim's book [3] for a proof of the latter. Ribenboim's proof is by verifying that Legendre's base-p formula agrees with the standard formula

$$v_p\,(n!) = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \cdots . \tag{1}$$

Surprisingly, it is possible to prove Legendre's formula without recourse to the above formula and that the standard formula follows from such a proof. What is more, Kummer's formula also follows without having to use Legendre's result.

4.1 Legendre's formula:

Lemma 3. *Let $n = (a_k \cdots a_1 a_0)_p$ and $s(n) = \sum_{r=0}^{k} a_r$. Then,*

$$v_p\,(n!) = \frac{n-s(n)}{p-1} \tag{2}$$

Proof. The formulae are evidently valid for $n=1$. We shall show that if Legendre's formula $v_p\,(n!) = \frac{n-s(n)}{p-1}$ holds for n, then it also holds for $pn+r$ for any $0 \leq r < p$. Note that the base-p expansion of $pn+r$ is

$$a_k \cdots a_1 a_0 \; r.$$

Let $f(m) = \frac{m-s(m)}{p-1}$, where $m \geq 1$. Evidently,

$$f(pn+r) = \frac{pn - \sum_{i=0}^{k} a_i}{p-1} = n + f(n).$$

On the other hand, it follows by induction on n that

$$v_p\left((pn+r)!\right) = n + v_p(n!). \tag{3}$$

For, if it holds for all $n < m$, then

$$\begin{aligned} v_p((pm+r)!) &= v_p(pm) + v_p((pm-p)!) \\ &= 1 + v_p(m) + m - 1 + v_p((m-1)!) \ = \ m + v_p(m!). \end{aligned}$$

Since it is evident that $f(m) = 0 = v_p(m!)$ for all $m < p$, it follows that $f(n) = v_p(n!)$ for all n. This proves Legendre's formula.

Note also that the formula

$$v_p(n!) = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \cdots$$

follows inductively using Legendre's result. □

4.2 Kummer's algorithm:

Lemma 4. *For $r, s \geq 0$, let $g(r,s)$ be the number of carries when the base-p expansions of r and s are added. Then, for $k \leq n$,*

$$v_p\left(\binom{n}{k}\right) = g(k, n-k). \tag{4}$$

Proof. Once again, this is clear if $n < p$, as both sides are then zero. We shall show that if the formula holds for all integers $0 \leq j \leq n$ (and *every* $0 \leq k \leq j$), it does so for $pn + r$ for $0 \leq r < p$ (and any $k \leq pn + r$). This would prove the result for all natural numbers.

Consider a binomial coefficient of the form $\binom{pn+r}{pm+a}$, where $0 \leq a < p$.

First, suppose $a \leq r$.

Write $m = b_k \cdots b_0$ and $n - m = c_k \cdots c_0$ in base-p. Then the base-p expansions of $pm + a$ and $p(n-m) + (r-a)$ are, respectively,

$$\begin{aligned} pm + a &= b_k \cdots b_0 \ a \\ p(n-m) + (r-a) &= c_k \cdots c_0 \ r - a. \end{aligned}$$

Evidently, the corresponding number of carries is

$$g(pm + a, p(n-m) + (r-a)) = g(m, n-m).$$

By the induction hypothesis, $g(m, n-m) = v_p(\binom{n}{m})$. Now $v_p\left(\binom{pn+r}{pm+a}\right)$ is equal to

$$\begin{aligned} & v_p((pn+r)!) - v_p((pm+a)!) - v_p((p(n-m)+r-a)!) \\ = \;& n + v_p(n!) - m - v_p(m!) - (n-m) - v_p((n-m)!) \;=\; v_p\left(\binom{n}{m}\right). \end{aligned}$$

Thus, the result is true when $a \leq r$.

Now suppose that $r < a$. Then $v_p\left(\binom{pn+r}{pm+a}\right)$ is equal to

$$\begin{aligned} & v_p((pn+r)!) - v_p((pm+a)!) - v_p((p(n-m-1)+(p+r-a))!) \\ = \;& n + v_p(n!) - m - v_p(m!) - (n-m-1) - v_p((n-m-1)!) \\ = \;& 1 + v_p(n) + v_p((n-1)!) - v_p(m!) - v_p((n-m-1)!) \\ = \;& 1 + v_p(n) + v_p\left(\binom{n-1}{m}\right). \end{aligned}$$

We need to show that

$$g(pm+a, p(n-m-1)+(p+r-a)) = 1 + v_p(n) + g(m, n-m-1). \tag{5}$$

Note that $m < n$. Write $n = a_k \cdots a_0$, $m = b_k \cdots b_0$ and $n-m-1 = c_k \cdots c_0$ in base-p. If $v_p(n) = d$, then $a_i = 0$ for $i < d$ and $a_d \neq 0$. In base-p, we have

$$n = a_k \cdots a_d \, 0 \cdots 0$$

and, therefore,

$$n - 1 = a_k \cdots a_{d+1}(a_d - 1)\ (p-1) \cdots\ (p-1).$$

Now, the addition $m + (n-m-1) = n-1$ gives $b_i + c_i = p-1$ for $i < d$ (since they must be $< 2p-1$). Moreover, $b_d + c_d = a_d - 1$ or $p + a_d - 1$.

Note the base-p expansions

$$\begin{aligned} pm + a &= b_k \cdots b_0 \ a \\ p(n-m-1) + (p+r-a) &= c_k \cdots c_0 \ (p+r-a). \end{aligned}$$

We add these using that fact that there is a carry-over in the beginning and that $1+b_i+c_i = p$ for $i < d$. Since there is a carry-over at the first step as well as at the next d steps, we have

$$pn + r = * \ * \ \cdots \ a_d \underbrace{0 \cdots 0}_{d \text{ times}} \ r$$

and

$$g(pm+a, p(n-m-1)+(p+r-a)) = 1 + d + g(m, n-m-1).$$

This proves Kummer's assertion also. $\square$

5 Acknowledgment

It is a real pleasure to thank the referee who read through the paper carefully and made a number of constructive suggestions and a few corrections as well.

References

[1] M. Aigner, *Combinatorial Theory*, Springer-Verlag, 1997.

[2] R. Honsberger, *In Polya's Footsteps, Miscellaneous Problems and Essays*, The Dolciani Mathematical Expositions, No. 19, Mathematical Association of America, 1997.

[3] P.Ribenboim, *The New Book of Prime Number Records*, Springer-Verlag, 1996.

2000 *Mathematics Subject Classification*: Primary 05A10; Secondary 11B83.
Keywords: generalized Catalan numbers, linear recursion, divisibility.

(Concerned with sequence A000108.)

Received May 21 2009; revised version received October 28 2009. Published in *Journal of Integer Sequences*, November 4 2009.

Return to Journal of Integer Sequences home page.