INTEGERS: ELECTRONIC JOURNAL OF COMBINATORIAL NUMBER THEORY x (200x), #Axx

Coincidences of Catalan and q-Catalan numbers

Florian Luca
Instituto de Matemáticas, Universidad Nacional Autonoma de México, C.P. 58089, Morelia, Michoacán, México
fluca@matmor.unam.mx

Paul Thomas Young
Department of Mathematics, College of Charleston, Charleston, SC 29424, USA
paul@math.cofc.edu

Received: , Revised: , Accepted: , Published:

To Professor Carl Pomerance on his 65th birthday

Abstract

Let C_n and $C_n(q)$ be the nth Catalan number and the nth q-Catalan number, respectively. In this paper, we show that the Diophantine equation $C_n = C_m(q)$ has only finitely many integer solutions (m, n, q) with $m > 1,\ n > 1,\ q > 1$. Moreover, they are all effectively computable.

1. Introduction

For a positive integer n, let

$$C_n = \frac{1}{n+1}\binom{2n}{n}$$

be the nth Catalan number. For a positive integer $q > 1$ we put $[k]_q = q^{k-1} + q^{k-2} + \cdots + 1 = (q^k - 1)/(q-1)$, $[n]_q! = \prod_{1\le k\le n}[k]_q$ and

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{[n]_q}{[k]_q[n-k]_q}.$$

The nth q-Catalan number is

$$C_n(q) = \frac{1}{[n+1]_q}\begin{bmatrix} 2n \\ n \end{bmatrix}_q.$$

There are many papers in the literature treating Diophantine equations involving binomial coefficients. For example, all the solutions of the Diophantine equation

$$\binom{n}{k} = \binom{m}{l} \qquad \text{with} \quad n \geq 2k \geq 4, \quad m \geq 2l \geq 4 \quad \text{and} \quad (m,k) \neq (n,l), \tag{1}$$

are still not known, although partial results on it appear in [1] and [8]. On the other hand, in the recent paper [6] it was shown that the analogous q-Diophantine equation

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{bmatrix} m \\ l \end{bmatrix}_q \qquad \text{with} \quad n \geq 2k \geq 4, \quad m \geq 2l \geq 4 \quad \text{and} \quad (n,k) \neq (m,l), \tag{2}$$

has no positive integer solutions (m,k,n,l,q) with $q > 1$. In this paper, we look at whether there could be common terms in the sequences of Catalan and q-Catalan numbers. Our theorem is the following.

Theorem 1. *The Diophantine equation*

$$C_n = C_m(q) \tag{3}$$

has only finitely many positive integer solutions (m,n,q) *with* $m > 1, \ n > 1, \ q > 1$.

Observe that $C_1 = C_1(q) = 1$, which is why we imposed the restrictions that $m > 1$ and $n > 1$ in the statement of Theorem 1. Observe also that $5 = C_3 = C_2(2)$. While we cannot compute all the finitely many solutions of the Diophantine equation (3), we will show in the third section of the paper that there are no other solutions to equation (3) with $q = 2$. Some remarks concerning the effectiveness of the proof of Theorem 1 appear in the last section of the paper. Our method uses results from analytic number theory and the specific tools that we use will be revealed when needed. Throughout the paper, we use the Landau symbols O and o as well as the Vinogradov symbols $\gg$, $\ll$, $\asymp$ with their usual meanings. Constants implied by them are absolute.

2. The proof of Theorem 1

Assume that (m,n,q) is a solution to equation (3) with large n. We shall first show that m is bounded, and later that n is also bounded. We assume therefore for the first part that m is large.

We start with asymptotic estimates of both sides of (3). On the one hand, by Stirling's formula to approximate the factorial, we know that

$$C_n = \frac{2^{2n}}{\pi^{1/2} n^{3/2}} \left(1 + O\left(\frac{1}{n}\right)\right). \tag{4}$$

On the other hand, we have

$$\begin{aligned} C_m(q) &= \prod_{m+2\le k\le 2m}(q^k-1)\left(\prod_{2\le j\le m}(q^j-1)\right)^{-1} \\ &= q^M\prod_{m+2\le k\le 2m}\left(1-\frac{1}{q^k}\right)\left(\prod_{2\le j\le m}\left(1-\frac{1}{q^j}\right)\right)^{-1} \\ &= q^M\eta(q)\prod_{m+2\le k\le m}\left(1-\frac{1}{q^k}\right)\prod_{m+1\le j}\left(1-\frac{1}{q^j}\right), \end{aligned} \tag{5}$$

where

$$M=\sum_{m+2\le k\le 2m}k-\sum_{2\le j\le m}j=m^2-m, \qquad \text{and} \qquad \eta(q)=\prod_{j\ge 2}\left(1-\frac{1}{q^j}\right)^{-1}.$$

It is clear that $1\le\eta(q)=O(1)$ uniformly for $q\ge 2$. Furthermore, since

$$1-\frac{1}{q^k}=\exp\left(O\left(\frac{1}{q^k}\right)\right),$$

we get that

$$\begin{aligned} \prod_{m+2\le k\le m}\left(1-\frac{1}{q^k}\right)\prod_{m+1\le j}\left(1-\frac{1}{q^j}\right) &= \exp\left(O\left(\sum_{m\le j}\frac{1}{q^j}\right)\right) \\ &= \exp\left(O\left(\frac{1}{q^m}\right)\right) \\ &= 1+O\left(\frac{1}{q^m}\right). \end{aligned}$$

We thus get that

$$C_m(q)=q^{m^2-m}\eta(q)\left(1+O\left(\frac{1}{q^m}\right)\right). \tag{6}$$

Comparing (4) and (6) we deduce that

$$\frac{2^{2n}}{\pi^{1/2}n^{3/2}}\left(1+O\left(\frac{1}{n}\right)\right)=q^{m^2-m}\eta(q)\left(1+O\left(\frac{1}{q^m}\right)\right),$$

and taking logarithms, we obtain that

$$2n\log 2-1.5\log n=(m^2-m)\log q+O(1),$$

which implies that

$$n\asymp m^2\log q. \tag{7}$$

Next, we use the arithmetic properties of the Catalan number and q-Catalan number, respectively. For a positive integer $s > 1$ let $P(s)$ be the largest prime factor of s. For a real number $x > 1$ write

$$\mathcal{P}(x) := \{x < p \le 2x \ : \ P(p-1) > p^{2/3}\}. \tag{8}$$

A result of Fouvry [4] asserts that $\#\mathcal{P}(x) \gg x/\log x$ provided that $x \gg 1$. Assume now that n is sufficiently large. Then all primes $p \in (n, 2n)$ divide C_n with at most one exception, and this exception occurs only when $n+1$ is prime. Thus, putting $\mathcal{Q} := \{p \in \mathcal{P}(n) \ : \ p \mid C_n\}$, we have that

$$\sum_{p\in\mathcal{Q}} \log p \ge (\log n)\#\mathcal{Q} \ge (\log n)(\#\mathcal{P}(n) - 1) \gg n \qquad (n \gg 1). \tag{9}$$

Let $p \in \mathcal{Q}$. Since $p \mid C_m(q)$, it follows that $p \mid q^k - 1$ for some $k \in [m+2,\ 2m]$. Let t_p be the multiplicative order of q modulo p. Then $t_p \mid k$, therefore $t_p \le 2m$. However, t_p also divides $p-1$. If $P(p-1) \mid t_p$, then $n^{2/3} < p^{2/3} \le P(p-1) \le t_p \le k \le 2m$, therefore

$$m^2 \log q \ge m^2 \log 2 \gg n^{4/3}.$$

Comparing the above inequality with (7), we get $n \gg n^{4/3}$, therefore $n \ll 1$, which is what we wanted to prove. Thus, assuming that n is sufficiently large, we conclude that t_p is coprime to $P(p-1)$, therefore $t_p \mid (p-1)/P(p-1)) < p^{1/3} \ll n^{1/3}$ and this inequality holds for all primes $p \in \mathcal{Q}$. We now split $\mathcal{Q}$ in two subsets as follows:

$$\mathcal{R} := \{p \in \mathcal{Q} \ : \ t_p \le m/\log m\}, \qquad \text{and} \qquad \mathcal{S} := \mathcal{Q}\backslash\mathcal{R}.$$

Observe that

$$\prod_{p\in\mathcal{R}} p \mid \prod_{k\le m/\log m} (q^k - 1),$$

therefore

$$\sum_{p\in\mathcal{R}} \log p \le \sum_{k\le m/\log m} \log(q^k - 1) \le (\log q) \sum_{k\le m/\log m} k \ll \frac{m^2 \log q}{(\log m)^2} \ll \frac{n}{(\log m)^2}. \tag{10}$$

Assume now that $\mathcal{S} \ne \emptyset$. Then for $p \in \mathcal{S}$, we have $m/\log m \le t_p \ll n^{1/3}$, therefore $m \ll n^{1/3}\log n$. For large n, the above inequality implies that $m \le n^{1/2}$. Observe now that primes $p \in \mathcal{S}$ have the property that $p \le 2n$ and p has a divisor t_p in the interval $[m/\log m, m]$, where $m \le n^{1/2}$. Let $H(x, y, z, \mathcal{P}_{-1})$ be the counting function of the number of primes $p \le x$ such that $p - 1$ has a divisor in the interval $[y, z]$. A result from [2] says that uniformly for $2y \le z \le y^2$ and $3 \le y \le x^{1/2}$, we have the estimate

$$H(x, y, z, \mathcal{P}_{-1}) \ll \frac{x}{\log x} u^{\delta} (\log(2/u))^{-3/2},$$

where

$$y^{1+u} = z \quad \text{and} \quad \delta = 1 - \frac{1 + \log\log 2}{\log 2} = 0.086071\ldots.$$

For $m \geq e^2$ and large n, setting $x := 2n,\ y = m/\log m$ and $z = m$, all the hypothesis needed to apply the result from [2] are fulfilled. Since

$$u = \frac{\log z}{\log y} - 1 = \frac{\log m}{\log m - \log\log m} - 1 = \frac{\log\log m}{\log m - \log\log m},$$

we deduce that

$$\#\mathcal{S} \ll \left(\frac{n}{\log n}\right) \frac{1}{(\log m)^{\delta}(\log\log m)^{3/2+\delta}}.$$

Hence,

$$\sum_{p\in\mathcal{S}} \log p \leq (\log(2n))\#\mathcal{S} \ll \frac{n}{(\log m)^{\delta}(\log\log m)^{3/2+\delta}}. \tag{11}$$

Hence, from inequalities (9), (10) and (11), we get that

$$n \ll \sum_{p\in\mathcal{Q}} \log p \leq \sum_{p\in\mathcal{R}} \log p + \sum_{p\in\mathcal{S}} \log p \ll n\left(\frac{1}{(\log m)^2} + \frac{1}{(\log m)^{\delta}(\log\log m)^{3/2+\delta}}\right),$$

which implies that

$$(\log m)^{\delta}(\log\log m)^{3/2+\delta} \ll 1,$$

yielding that $m \ll 1$. Thus, m is bounded, as promised.

Assume now that $m \geq 2$ is fixed. Then $C_m(q)$ is a polynomial of degree $m^2 - m$ in the variable q. Observe that all roots of $C_m(q)$ are roots of unity of degree at most $2m$. Furthermore, it is easy to see that $C_m(1) = C_m$ and $C_m(-1) = \binom{m}{\lfloor m/2\rfloor}$. In particular, $C_m(q)$ has no real roots. Using the fact that

$$q^k - 1 = \prod_{d|k} \Phi_d(q),$$

where $\Phi_d(X) \in \mathbb{Z}[X]$ is the dth cyclotomic polynomial, we get that

$$C_m(q) = \prod_{3\leq k\leq 2m} \Phi_k(q)^{a_k},$$

with some integer multiplicities $a_k \geq 0$. Now it is well-known that a prime number p dividing $\Phi_k(q)$ for some positive integer q has the property that either $p \mid k$, or $p \equiv 1 \pmod{k}$. Take now $P = 4\prod_{3\leq r\leq 2m} r$, where in the above product r ranges over primes. By known results on the distribution of primes in arithmetic progressions, for large n, the interval $(n, 2n)$ contains $\geq n/(2\phi(P)\log n)$ primes $p \equiv -1 \pmod{P}$. In particular, if $n > 2m$ is sufficiently large, there exist at least two such primes. At least one of them call it p divides C_n, therefore it also divides $C_m(q)$. Thus,

$$p \mid \Phi_l(q)$$

for some $3 \leq l \leq 2m$, and $p > 2m$, therefore $p \equiv 1 \pmod{l}$. Since $l \geq 3$, it follows that either $4 \mid l$, or l is divisible by some odd prime. Thus, either $p \equiv 1 \pmod{4}$, or $p \equiv 1 \pmod{r}$ for some odd prime $r \leq 2m$, which is false because $p \equiv -1 \pmod{P}$. Hence, n is also bounded. The proof of the Theorem 1 is therefore complete.

3. The case $q = 2$

For $q = 2$, the only solution is $(n, m) = (3, 2)$, which can be proved in the following way. Observe first that the number $C_m(2)$ is odd. Hence, C_n is odd, therefore $n = 2^k - 1$ for some positive integer k. It was shown in Lemma 1 in [7] that in this case

$$C_n = \frac{2^N}{2^{\delta/2}\pi^{1/2}}\left(1 + \zeta_n\right), \qquad \text{where} \qquad \frac{1}{2^{k+2}} < \zeta_n < \frac{1}{2^{k+1}},$$

where $N := n - \lfloor 3k/2 \rfloor$ and $\delta \in \{0, 1\}$ is such that $k \equiv \delta \pmod 2$. In particular, if $k \geq 10$, then the first four binary significant digits of C_n are the same as the first four binary significant digits either of

$$\frac{1}{\sqrt{\pi}} = 0.1001000001\ldots, \qquad \text{or of} \qquad \frac{1}{\sqrt{2\pi}} = 0.01100110001\ldots,$$

respectively, according to whether k is even or odd. Hence, the first four significant binary digits of C_n are either 1001 or 1100. On the other hand, let us make estimate (5) explicit. Suppose that $m \geq 10$. Then using the fact that $1 + x < e^x < 1 + x + x^2$ and $1 + x > e^{2x}$ for $x \in (-1/2, 0)$, we have

$$\begin{aligned}
\prod_{m+2\leq k\leq m}\left(1-\frac{1}{2^k}\right)\prod_{m+1\leq j}\left(1-\frac{1}{2^j}\right) &< \exp\left(-\sum_{m+2\leq k\leq 2m}\frac{1}{2^j}-\sum_{m+1\leq j}\frac{1}{2^j}\right)\\
&= \exp\left(-\frac{1}{2^m}-\frac{1}{2^{m+1}}+\frac{1}{2^{2m}}\right)\\
&< 1-\frac{1}{2^m}-\frac{1}{2^{m+1}}+\frac{1}{2^{2m}}+\frac{9}{2^{2m+2}}\\
&< 1-\frac{1}{2^m},
\end{aligned}$$

while

$$\begin{aligned}
\prod_{m+2\leq k\leq m}\left(1-\frac{1}{2^k}\right)\prod_{m+1\leq j}\left(1-\frac{1}{2^j}\right) &> \exp\left(-2\sum_{m+2\leq k\leq 2m}\frac{1}{2^j}-2\sum_{m+1\leq j}\frac{1}{2^j}\right)\\
&> \exp\left(-\frac{1}{2^{m-1}}-\frac{1}{2^m}+\frac{1}{2^{2m-1}}\right)\\
&> 1-\frac{1}{2^{m-1}}-\frac{1}{2^m}\\
&> 1-\frac{1}{2^{m-2}}.
\end{aligned}$$

Thus,

$$C_m(2) = 2^M\eta(2)(1-\zeta_m), \qquad \text{where} \qquad \frac{1}{2^m} < \zeta_m < \frac{1}{2^{m-2}},$$

for $m \geq 10$. Now since the binary expansion of $\eta(2)$ is

$$1.0101001000000000110\ldots$$

it follows that for $m \geq 10$ the first four binary significant digits of $C_m(2)$ are 1010. Thus, we have just showed that if $C_n = C_m(2)$, then $n = 2^k - 1$ with $\min\{m,k\} < 10$. An immediate calculation reveals no other solutions to $C_n = C_m(2)$.

4. Remarks

It would be interesting to find all the solutions of the Diophantine equation (3) or at least an effective upper bound for $\max\{m,n,q\}$ when these three positive integers are such that equation (3) holds. A close analysis of our arguments shows that in order to do so we need effective versions of the result of Fouvry from [4] regarding a lower bound of the same order of magnitude as $\pi(x)$ on the cardinality of the set $\mathcal{P}(x)$ appearing in (8), as well as an effective version of the result of Ford and Shparlinski from [2] concerning an upper bound for the count of primes $p \leq x$ with $p-1$ having a divisor in $[y,z]$. While the result from [2] is effective, the proof of the result of Fouvry from [4] uses the Bombieri–Vinogradov theorem and as such is ineffective. However, our argument goes through provided that for some $\alpha > 0$ there exists $\beta > 0$ and x_α such that if we put

$$\mathcal{P}_\alpha(x) := \{x < p \leq 2x \ : \ P(p-1) > x^{1/2+\alpha}\},$$

then the inequality

$$\#\mathcal{P}_\alpha(x) > \beta\pi(x) \qquad \text{holds for all} \quad x > x_\alpha.$$

In the recent preprint [3], Fouvry gives an effective version of the Bombieri-Vinogradov theorem and as an application he shows that the above inequality holds with $\alpha = 1/10$ and some effectively computable constants β and x_α. An effective variant of the Bombieri-Vinogradov theorem also appears in the recent preprint [5]. Hence, all solutions of equation (3) are effectively computable. In order to write down a bound on the largest solution however, one will need to compute the appropriate effective constants from the corresponding results from [2] and [3] which might not be an easy task. Perhaps an easier research problem is to find the complete list of solutions to the equation (3) under the ERH.

Acknowledgements. We thank the referee for remarks which improved the quality of the paper. We also thank Professors Kevin Ford, Étienne Fouvry and Carl Pomerance for helpful discussions. This work started during a visit of F. L. at the Mathematics Department of the College of Charleston in September of 2008 and was finalized during a visit of P. T. Y. at the Mathematical Institute of the UNAM in Morelia, Mexico in November of 2009. Both authors thank the respective institutions for their hospitality and support.

References

[1] Y. Bugeaud, M. Mignotte, S. Siksek, M. Stoll and Sz. Tengely, 'Integral points on hyperelliptic curves', *Algebra Number Theory* **2** (2008), 859–885.

[2] K. Ford and I. E. Shparlinski, ‘On curves over finite fields with Jacobians of small exponent’, *Int. J. Number Theory* **4** (2008), 819–826.

[3] E. Fouvry, ‘A variant of the Bombieri-Vinogradov theorem with application’, *Preprint*, 2010.

[4] E. Fouvry, ‘Théorème de Brun-Titchmarsh: application au théorème de Fermat’, *Invent. Math.* **79** (1985), 383–407.

[5] H. W. Lenstra and C. Pomerance, ‘Primality testing with Gaussian periods’, *Preprint*, 2009.

[6] F. Luca, D. Marques and P. Stănică, ‘On the spacing between C-nomial coefficients’, *J. Number Theory* **130** (2010), 82–100.

[7] F. Luca and P. T. Young, ‘On the binary expansion of odd Catalan numbers’, *Proceedings of the XIVth International Conference on Fibonacci Numbers, Morelia, Mexico, 2010. eds. Luca and Stănică*, to appear.

[8] R. J. Stroeker and B. M. M. de Weger, ‘Elliptic binomial Diophantine equations’, *Math. Comp.* **68** (1999), 1257–1281.