p^q-Catalan Numbers and Squarefree Binomial Coefficients

Pantelimon Stănică *
Auburn University Montgomery, Department of Mathematics
Montgomery, AL 36117, USA,
(334) 244-3321, e-mail: stanpan@strudel.aum.edu

July 10, 2004

Running Head:
p^q-Catalan Numbers

*On leave from the Institute of Mathematics of Romanian Academy, Bucharest, Romania

Abstract

In this paper we consider the generalized Catalan numbers $F(s,n) = \frac{1}{(s-1)n+1}\binom{sn}{n}$, which we call s-Catalan numbers. We find all natural numbers n such that for p prime, p^q divides $F(p^q, n)$, $q \geq 1$ and all distinct residues of $F(p^q, n) \pmod{p^q}$, $q = 1, 2$. As a byproduct we settle a question of Hough and the late Simion on the divisibility of the 4-Catalan numbers by 4. We also prove that $\binom{p^q n+1}{n}$, $p^q \leq 99999$, is squarefree for n sufficiently large (explicit), and with the help of the generalized Catalan numbers we find the set of possible exceptions. As consequences, we obtain that $\binom{4n+1}{n}$, $\binom{9n+1}{n}$ are squarefree for $n \geq 2^{1518}$, respectively $n \geq 3^{956}$, with at most $2^{18.2}$, respectively $3^{15.3}$ possible exceptions.

Keywords. Binomial Coefficients, Divisibility, Congruences, Residues

p^q-Catalan Numbers and Squarefree Binomial Coefficients

1 Introduction

Problems involving binomial coefficients were considered by many mathematicians for over two centuries. R.K. Guy in [6] mentions several problems on divisibility of binomial coefficients (see **B31, B33**). Erdös conjectured that for $n > 4$, $\binom{2n}{n}$ is never squarefree. This was proved by Sárközy in [12], for sufficiently large n, and by Granville and Ramaré in [5] for any $n > 4$.

Many people (see, for instance, [1, 2, 7, 8, 9, 11, 15]) proposed and studied the following generalization of classical Catalan numbers $\frac{1}{n+1}\binom{2n}{n}$, which we will call *s-Catalan numbers*, $F(s,n) = \frac{1}{(s-1)n+1}\binom{sn}{n}$. There are many interpretations of this sequence [2, 7, 9, 11, 15], for instance: the number of s-ary trees with n source-nodes, the number of ways of associating n applications of a given s-ary operator, the number of ways of dividing a convex polygon into n disjoint $(s+1)$-gons with nonintersecting diagonals, and the number of s-good paths (below the line $y = sx$) from $(0,-1)$ to $(n,(s-1)n-1)$.

Naturally, some of the questions proposed by Erdös on the classical Catalan numbers, may be asked here as well, as Hough and the late Simion proposed [8]: (a) When p is prime, for what values of n is $F(p,n)$ divisible by p? (b)* For what values of n is $F(4,n)$ divisible by 4? (c)* What can you say when s takes on the other composite values? There are no known answers for $(b),(c)$. In this paper we give a simple proof to (a), and we show that $F(p^2,n)$ is divisible by p^2, unless $(p^2-1)n+1$ is an even power of p, or a sum of odd powers of p (with the numbers of distinct powers summing to p), thereby proving (b), and (c) for $s = p^2$. We also prove that $\binom{p^q n+1}{n}$, $p^q \leq 99999$, is squarefree for n sufficiently large (explicit), and with the help of the generalized Catalan numbers we find the set of possible exceptions. As consequences, we obtain that $\binom{4n+1}{n}$, $\binom{9n+1}{n}$ are squarefree for $n \geq 2^{1518}$, respectively $n \geq 3^{956}$, with at most $2^{18.2}$, respectively $3^{15.3}$ possible exceptions.

2 Preliminary Results

Let $[x]$ be the largest integer smaller than x. In this section we state a few results which will be needed later. Lucas (1878) (see [3]) found a simple method to find $\binom{m}{n} \pmod{p}$.

Theorem 1 (Lucas). *If p is prime, then $\binom{m}{n} \equiv \binom{[m/p]}{[n/p]}\binom{m_0}{n_0} \pmod{p}$, where m_0, n_0 are the least non-negative residues modulo p of m, respectively n.*

Define $n!_p$ to be the product of all integers $\leq n$, that are not divisible by p. We see that $n!_p = \dfrac{n!}{[n/p]!p^{[n/p]}}$. Granville in [4] proves the following beautiful generalization of Lucas' Theorem.

Theorem 2 (Granville). *Suppose that the prime power p^q and positive integers $m = n+r$ are given. Let N_j be the least positive residue of $[n/p^j] \pmod{p^q}$ for each $j \geq 0$ (that is, $N_j = n_j + n_{j+1}p + \cdots + n_{j+q-1}p^{q-1}$): also make the corresponding definitions for m_j, M_j, r_j, R_j. Let e_j be the number of indices $i \geq j$ for which $m_i < n_i$ (that is, the number of* carries*, when adding n and r in base p, on or beyond the jth digit). Then*

$$\frac{1}{p^{e_0}}\binom{m}{n} \equiv (\pm 1)^{e_{q-1}} \frac{M_0!_p}{N_0!_p\, R_0!_p} \frac{M_1!_p}{N_1!_p\, R_1!_p} \cdots \frac{M_d!_p}{N_d!_p\, R_d!_p} \pmod{p^q},$$

where $(\pm)$ is (-1) except if $p = 2$ and $q \geq 3$.

In 1808 Legendre showed that the exact power of p dividing $n!$ is

$$[n/p] + [n/p^2] + [n/p^3] + \cdots . \qquad (1)$$

We define (see [4]) the *sum of digits function* $\sigma_p(n) = n_0 + n_1 + \cdots + n_d$, if $n = n_0 + n_1 p + \cdots + n_d p^d$. Then, using σ, (1) transforms into

$$\frac{n - \sigma_p(n)}{p-1}. \tag{2}$$

We will need the following result which belongs to Kummer

Theorem 3 (Kummer). *The power to which the prime p divides the binomial coefficient $\binom{m}{n}$, say $v_p(\binom{m}{n})$, is given by the number of carries when we add n and $m-n$ in base p.*

Our first result gives a complete answer to the first posed question (a), generalizing the well-known result on Catalan numbers, or equivalently, on middle binomial coefficients (see [6]), which states that $4 \,|\, \binom{2n}{n}$, unless $n = 2^k$.

Theorem 4. *Let p be a prime. Then, p divides $F(p,n)$, unless n is of the form $\frac{p^k-1}{p-1}$, $k \in \mathbf{N}$, in which case $F(p,n) \equiv 1 \pmod p$.*

Proof. We re-write $F(p,n) = \dfrac{1}{(p-1)n+1}\dbinom{pn}{n} = \dfrac{1}{pn+1}\dbinom{pn+1}{n}$. We shall find the values n such that $p \not| F(p,n)$. Since $p \not| F(p,0)$, we assume $n \neq 0$. Applying Lucas' Theorem repeatedly for the base p representations $(0 \le m_i, n_i \le p-1)$, $m = m_0 + m_1 p + \cdots + m_d p^d$ and $n = n_0 + n_1 p + \cdots + n_d p^d$, we obtain $\binom{m}{n} \equiv \binom{m_0}{n_0}\binom{m_1}{n_1}\cdots\binom{m_d}{n_d} \pmod p$. For $m = pn+1 \equiv 1 \pmod p$, we get

$$F(p,n) \equiv \binom{pn+1}{n} \equiv \binom{1}{n_0}\binom{n_0}{n_1}\cdots\binom{n_{d-1}}{n_d} \pmod p, \quad n_d \neq 0.$$

Let $n_{-1} = 1$. Since for any j, $n_j < p$, if $\binom{n_{i-1}}{n_i} \neq 0$, then $p \not| \binom{n_{i-1}}{n_i}$. Thus, the numbers n with $p \not| F(p,n)$ are those positive integers n with $\binom{n_{i-1}}{n_i} \neq 0, i = 0, 1, \ldots, d$, or $n = 0$. From $n_{i-1} \ge n_i$, $n_{-1} = 1, n_d \neq 0$, and $\binom{n_{i-1}}{n_i} \neq 0$, we obtain $n_{i-1} = n_i$. Thus, $n = 1 + p + \cdots + p^d = \frac{p^{d+1}-1}{p-1}$. Therefore, if p is prime, $p | F(p,n)$ for any number $n \neq \frac{p^k-1}{p-1}, k \in \mathbf{N}$.

Using Lucas' Theorem we easily see that for $n = \frac{p^k-1}{p-1}$, the least residue of $F(p,n)$ modulo p is 1. $\square$

The following lemma will be extensively used throughout the paper

Lemma 5. *We have*

$$v_p(F(p^q,n)) = \frac{\sigma_p((p^q-1)n+1)-1}{p-1}.$$

Proof. We use the identity $F(p^q,n) = \frac{1}{p^qn+1}\binom{p^qn+1}{n}$. Using (2) we get that the power of p dividing $\binom{m}{n}$ is

$$v_p\left(\binom{m}{n}\right) = \frac{\sigma_p(n)+\sigma_p(m-n)-\sigma_p(m)}{p-1}. \tag{3}$$

Let $m = p^qn+1$. Thus, (3) becomes

$$v_p(F(p^q,n)) = \frac{\sigma_p(n)+\sigma_p((p^q-1)n+1)-\sigma_p(p^qn+1)}{p-1} = \frac{\sigma_p((p^q-1)n+1)-1}{p-1},$$

since $\sigma_p(p^qn+1) = \sigma_p(n)+1$. $\square$

3 Scarce squarefree p^2-Catalan numbers

Denote by $n = (ab\dots)_p$ the base p representation of n, a being the most significant bit. Our next result refers to the third question of Hough and Simion.

Theorem 6. *Given a prime p, p^2 divides $F(p^2,n)$, unless n is of the form $\dfrac{p^{2t}-1}{p^2-1}$, $t \in \mathbf{N}$, in which case $F(p^2,n) \equiv 1 \pmod{p^2}$, or of the form $\dfrac{c_1p^{2i_1+1}+\cdots+c_sp^{2i_s+1}-1}{p^2-1}$, $i_1 < \cdots < i_s$, with $\sum_{i=1}^{s} c_i = p$, $s \in \mathbf{N}$, $0 \le c_i < p$, in which case $F(p^2,n) \equiv \binom{p}{c_1,c_2,\dots,c_s} \pmod{p^2}$ (the multinomial coefficient).*

Proof. As before $F(p^2,n)=\frac{1}{p^2n+1}\binom{p^2n+1}{n}$. A number n, which does not satisfy the divisibility, must satisfy (see Lemma 5)

$$v_p(F(p^2,n))=\frac{\sigma_p((p^2-1)n+1)-1}{p-1}\leq 1, \tag{4}$$

which implies $\sigma_p((p^2-1)n+1)\leq p$.

Assume first that $\sigma_p((p^2-1)n+1)=1$. Therefore, $(p^2-1)n+1=p^k\equiv(-1)^k \pmod{p^2-1}$, therefore k must be even, say $k=2t$, so $n=\dfrac{p^{2t}-1}{p^2-1}$.

Assume now that $\sigma_p((p^2-1)n+1)=l$, and $1<l\leq p$. It follows that $(p^2-1)n+1=p^{\alpha_1}+\cdots+p^{\alpha_l}$, $\alpha_1\leq\alpha_2\leq\cdots\leq\alpha_l$. Therefore, $(p^2-1)n+1\equiv l\equiv 1 \pmod{p-1}$, and since $1<l\leq p$, we get that $l=p$. Then, $(p^2-1)n+1=p^{\alpha_1}+\cdots+p^{\alpha_p}$, $\alpha_1\leq\alpha_2\leq\cdots\leq\alpha_p$. It follows that

$$\begin{aligned}(p^2-1)n &\equiv -1+p\sum_{\alpha_i \text{ odd}}1+\sum_{\alpha_i \text{ even}}1\\ &\equiv -1+p^2-(p-1)\sum_{\alpha_i \text{ even}}1\\ &\equiv -(p-1)\sum_{\alpha_i \text{ even}}1 \pmod{p^2-1},\end{aligned}$$

so $\sum_{\alpha_i \text{ even}}1$ must be divisible by $p+1$. Since $0\leq\sum_{\alpha_i \text{ even}}1\leq p$, we see that $\sum_{\alpha_i \text{ even}}1$ must be an empty sum. Therefore, all $\alpha_j=2i_j+1$. We obtain $n=\dfrac{p^{2i_1+1}+\cdots+p^{2i_p+1}-1}{p^2-1}$, $i_1\leq i_2\leq\cdots\leq i_p$, and the first claim is proved.

Let $n_{-1}=0$. Consider $n=\dfrac{p^{2t}-1}{p^2-1}$. It follows that $n=(1010\cdots101)_p$ and p^2n+1 attaches to this string the block 01 to the right, so it is of the same form. Since $M_i=N_{i-2}$ and $R_i=1$, except for $R_{2t-1}=p$, we get, using Granville's theorem,

$$\frac{1}{(p^2-1)n+1}\binom{p^2n}{n}\equiv\frac{1}{p^2n+1}\binom{p^2n+1}{n}\equiv p^{e_0}(-1)^{e_1}\frac{M_0!_pM_1!_p}{R_{2t-1}!_p} \pmod{p^2}. \tag{5}$$

Now, $M_0 = m_0 + m_1 p = m_0 + n_{-1} p = 1$, $M_1 = m_1 + m_2 p = n_{-1} + n_0 p = p$ and $R_{2t-1} = p$, implies $M_0!_p = 1, M_1!_p = p!_p = (p-1)!$ and $R_{2t-1}!_p = (p-1)!$. Thus, (5) becomes $F(p^2, n) \equiv p^{e_0}(-1)^{e_1} \equiv 1 \pmod{p^2}$, since $e_0 = e_1 = 0$.

Consider $n = \dfrac{c_1 p^{2i_1+1} + \cdots + c_s p^{2i_s+1} - 1}{p^2-1}$, $i_1 < \cdots < i_s$ and $c_1 + c_2 + \cdots + c_s = p$. Observe that $s \geq 2$. It follows that

$$\begin{aligned} n &= \frac{c_s(p^{2i_s+1} - p) + \cdots + c_1(p^{2i_1+1} - p) + p^2 - 1}{p^2-1} && (6) \\ &= c_s(p^{2i_s-1} + p^{2i_s-3} + \cdots + 1) + \cdots + c_1(p^{2i_1-1} + p^{2i_1-3} + \cdots + 1) + 1 && (7) \\ &= c_s p^{2i_s-1} + c_s p^{2i_s-3} + \cdots + (c_s + c_{s-1}) p^{2i_{s-1}-1} + \cdots + 1. && (8) \end{aligned}$$

But $p^2 n + 1$ attaches the block 01 to the right of the base p representation of n, and since there is a carry in this case, we get $e_0 = e_1 = 1$. Also, $n_0 = 1$, $M_0!_p = 1, M_1!_p = (p-1)!$, $R_i!_p = 1$ except for $R_{2i_k}!_p = (c_k p)!_p = \dfrac{(c_k p)!}{c_k! p^{c_k}}$ and $R_{2i_k+1}!_p = (c_k)!_p = c_k!$, for $k = 1, 2, \ldots, s$. Now, applying Granville's theorem we get

$$\begin{aligned} F(p^2, n) &\equiv p^{e_0}(-1)^{e_1} \frac{M_0!_p M_1!_p}{R_0!_p \cdots R_{d+2}!_p} \equiv (-1)p \, \frac{(p-1)!}{\prod_k (c_k p)!_p \, c_k!} \\ &\equiv (-1)p \, \frac{(p-1)!}{\prod_k \frac{(c_k p)!}{p^{c_k}}} \equiv \frac{(-1)p^{p+1}(p-1)!}{\prod_k (c_k p)!} \pmod{p^2}, \end{aligned} \quad (9)$$

since $\sum_{k=1}^{s} c_k = p$. We prove that the last expression is the multinomial coefficient (this was observed by one of our referees, whom we thank). First, assume $p = 2$. Since $s = 2$ in this case, we get $c_1 = c_2 = 1$ and the claim is trivially satisfied. Let $p > 2$. We observe that

$$\frac{(mp+1)\cdots(mp+p-1)}{(p-1)!} = \prod_{j=1}^{p-1}\left(1 + \frac{mp}{j}\right) \equiv 1 + mp \sum_{j=1}^{p-1} \frac{1}{j} \pmod{p^2} \equiv 1 \pmod{p^2},$$

since in the last sum $\frac{1}{j} + \frac{1}{p-j} \equiv 0 \pmod{p}$ for $p > 2$. Therefore,

$$\frac{(c_k p)!}{p^{c_k}} \equiv c_k! \, (p-1)! \pmod{p^2}.$$

Taking the product $\prod_{k=1}^{s}\frac{(c_kp)!}{p^{c_k}}\equiv(p-1)!^p\prod_{k=1}^{s}c_k!\equiv-\prod_{k=1}^{s}c_k!$, since $(p-1)!^p\equiv-1 \pmod p$, which replaced in (9) produces the claim. The theorem is proved. □

The following corollary gives a complete answer to the second question of Hough and Simion.

Corollary 7. *$F(4,n)$ is divisible by 4, unless n is of the form $\frac{2^{2t}-1}{3}$, in which case $F(4,n)\equiv 1 \pmod 4$, or of the form $\frac{2^{2t+1}+2^{2j+1}-1}{3}$, for $t>j$, in which case $F(4,n)\equiv 2$ $\pmod 4$.*

We include here the base 2 representations of the above numbers ≤ 60, namely $1,3$, $5,11$, $13,21$, $43,45,53$ since it suggests a recursive construction of the sequence,

$$1_2, 11_2,$$

$$101_2, 1011_2, 1101_2,$$

$$10101_2, 101011_2, 101101_2, 110101_2.$$

We see that on each row, we start with $1010\ldots1$, obtaining the rest of the strings by inserting the bit 1 to the right of an already existent bit 1, starting with the rightmost one.

What are the possible *least* distinct residues of the p^2-Catalan numbers modulo p^2? We give the following table, with the least residues of $\frac{1}{(p^2-1)n+1}\binom{p^2n}{n} \pmod{p^2}$, for $p=2,3,5,7$, computed easily by hand, using Theorem 6. We listed the partitions of p, and we computed the residue modulo p^2 of the multinomial coefficient corresponding to each partition, eliminating duplicates. For instance, if $p=5$, the partitions of 5 are: $\{\{5\},\{4,1\},\{3,2\}, \{3,1,1\},\{2,2,1\}, \{2,1,1,1\},\{1,1,1,1,1\}\}$, so by Theorem 6 the least residues of $F(5^2,n)$ modulo 5^2 are: 0 and $\binom{5}{5}=1$, $\binom{5}{4,1}=5$, $\binom{5}{3,2}=10$, $\binom{5}{3,1,1}=20$,

$\binom{5}{2,2,1} = 30 \equiv 5 \pmod{25}$, $\binom{5}{2,1,1,1} = 60 \equiv 10 \pmod{25}$, $\binom{5}{1,1,1,1,1} = 120 \equiv 20 \pmod{25}$

p	least residues modulo p^2 of $F(p^2, n)$
2	$0, 1, 2$
3	$0, 1, 3, 6$
5	$0, 1, 5, 10, 20$
7	$0, 1, 7, 14, 21, 35, 42$

We provide here the following weak bound.

Corollary 8. *The number of distinct residues of p^2-Catalan numbers $\pmod{p^2}$ is less than or equal to $\pi(p) + 1$, where $\pi(p)$ is the number of partitions of p.*

Proof. Straightforward. □

Remark 9. *If we denote by a_s the number of distinct residues of s^2-Catalan numbers $\pmod{s^2}$, then $\{a_s\}_s$ is the sequence* A053991 *in [14].*

4 Divisibility of p^q-Catalan numbers

Now we attempt to find all natural numbers for which p^q divides $F(p^q, n)$, $q \geq 3$. Denote by j_i the least non negative residue of $\alpha_i \pmod{q}$. We prove the result

Theorem 10. *When p is an odd prime and $q \geq 3$, then p^q divides $F(p^q, n)$, unless n is of the form $\dfrac{p^{tq} - 1}{p^q - 1}$, for some $t \in \mathbf{N}$, or of the form $\dfrac{p^{qt_1 + j_1} + \cdots + p^{qt_{m(p-1)+1} + j_{m(p-1)+1}} - 1}{p^q - 1}$, for some $t_i \in \mathbf{N}$, $1 \leq m \leq q - 1$, $0 \leq j_i \leq q - 1$, and $\sum_i p^{j_i} \equiv 1 \pmod{p^q - 1}$.*

Proof. By Lemma 5, if $p^q \not| F(p^q, n)$, then

$$v_p(F(p^q, n)) = \frac{\sigma_p((p^q - 1)n + 1) - 1}{p - 1} \leq q - 1,$$

so $\sigma_p((p^q - 1)n + 1) \leq (p - 1)(q - 1) + 1$. If $\sigma_p((p^q - 1)n + 1) = 1$, then $(p^q - 1)n + 1 = p^{tq+i}$, for some $0 \leq i \leq q - 1$. Working modulo $p^q - 1$ implies $i = 0$. Thus, $n = \dfrac{p^{tq} - 1}{p^q - 1}$. Assume

$\sigma_p((p^q-1)n+1) = l, 1 < l \leq (p-1)(q-1)+1$. We obtain $(p^q-1)n+1 = p^{\alpha_1}+\cdots+p^{\alpha_l}, \alpha_1 \leq \cdots \leq \alpha_l$. Modulo $(p-1)$, this transforms into

$$(p^q-1)n \equiv -1 + l \equiv 0 \pmod{p-1},$$

which will imply $l = m(p-1)+1$, for some m. Since $1 < l \leq (q-1)(p-1)+1$, we get $0 < m \leq q-1$. We obtain, for $\alpha_i = qt_i + j_i$, $0 \leq j_i \leq q-1$,

$$n = \frac{p^{qt_1+j_1} + \cdots + p^{qt_{m(p-1)+1}+j_{m(p-1)+1}} - 1}{p^q - 1}, \ m \in \mathbf{N},$$

with the condition $\sum_i p^{j_i} \equiv 1 \pmod{p^q-1}$. □

We use in the next section the following

Corollary 11. $p^q \mid \frac{1}{(p^q-1)n+1}\binom{p^q n}{n}$ *if and only if* $p^q \mid \binom{p^q n+1}{n}$.

5 Squarefree Binomial Coefficients

In this section we study squarefree binomial coefficients of the form $\binom{p^q n+1}{n}$, with the help of generalized Catalan numbers. Thus, in order to study these squarefree binomial coefficients, it suffices to consider only n of the form $\frac{p^{qt_1+j_1} + \cdots + p^{qt_{m(p-1)+1}+j_{m(p-1)+1}} - 1}{p^q-1}$, $t_i \in \mathbf{N}$, $1 \leq m \leq q-1$, $0 \leq j_i \leq q-1$, such that $\sum_i p^{j_i} \equiv 1 \pmod{p^q-1}$.

In [5], the authors proved that if $\binom{n}{k}$ is squarefree, then n or $n-k$ must be small. Finding explicit bounds is a much more difficult task. They showed that $\binom{2n}{n}$ is squarefree for $n > 2^{1617}$, and used some arguments to simplify the computer's work, in checking the possible exceptions $n = 2^r$. However, our job is not as hard; we rely on [5] and use some estimates on the Chebyshev's function $\psi(x) = \sum_{d \leq x} \Lambda(d)$, where $\Lambda(d)$ is the Von Mangoldt's function, $\Lambda(d) = \log r$, if $d = r^s$, r prime and $\Lambda(d) = 0$, otherwise, to show our results. Define $e(x) = e^x$ and $\psi(x) = 0$, if x is an integer, and $\psi(x) = \{x\} - \frac{1}{2}$, otherwise, where $\{x\}$ is the fractional part of x.

The following lemma proves to be very useful

Lemma 12. *If $p^q \leq 99999$, the inequality*

$$0.9999975\sqrt{p^q n + 1} - 1.0000025\sqrt{(p^q - 1)n + 1} > \\ 21.683\, p^{\frac{23q}{48}}\, n^{\frac{23}{48}}\, \left(\log\left(256((p^q-1)n+1)\right)\right)^{\frac{11}{4}} + \frac{11}{8}(3\log n + 2q\log p). \tag{10}$$

is true for $n \geq \tau_0$ sufficiently large.

Proof. First we prove that $\sqrt{1+x} + \sqrt{1+x-n} \leq 2\sqrt{x}$, $2 \leq n \leq x+1$. By squaring we get $2x + 2 - n + 2\sqrt{(x+1)(x+1-n)} \leq 4x$, which is equivalent to $4(x+1)(x+1-n) \leq (n+2x-2)^2$. The last inequality is equivalent to $n^2 - 16x + 8nx \geq 0$, which is certainly true if $n \geq 2$. Now, let $x' = 1 + x$. We evaluate

$$(1-\alpha)\sqrt{1+x} - (1+\alpha)\sqrt{1+x-n} = \frac{(1-\alpha)^2 x' - (1+\alpha)^2(x'-n)}{(1-\alpha)\sqrt{x'} + (1+\alpha)\sqrt{x'-n}}$$
$$= \frac{n(1+\alpha)^2 - 4\alpha x'}{(1-\alpha)\sqrt{x'} + (1+\alpha)\sqrt{x'-n}} \geq \frac{n(1+\alpha)^2 - 4\alpha x'}{(1+\alpha)(\sqrt{x'} + \sqrt{x'-n})} \geq \frac{n\left(\frac{(1-\alpha)^2}{1+\alpha} - \frac{4\alpha}{n(1+\alpha)}x\right)}{2\sqrt{x}}.$$

Therefore,

$$(1-\alpha)\sqrt{1+x} - (1+\alpha)\sqrt{1+x-n} \geq \left(\frac{(1-\alpha)^2}{1+\alpha} - \frac{4\alpha}{1+\alpha}\frac{x}{n}\right)\frac{n}{2\sqrt{x}}. \tag{11}$$

Taking $x = p^q n$, $\alpha = \dfrac{1}{4\cdot 10^5}$, in (11), we get

$$0.9999975\sqrt{p^q n + 1} - 1.0000025\sqrt{(p^q-1)n+1} \\ \geq \left(\frac{0.9999975^2}{1.0000025} - \frac{1}{100000.25}\,p^q\right)\frac{1}{2\sqrt{p^q}}\,n^{\frac{1}{2}}. \tag{12}$$

If $p^q \leq 99999$, then (12) implies our claim that the inequality (10) is true for $n \geq \tau_0$ sufficiently large, since, by (12), the left side is $O(n^{\frac{1}{2}})$ and the right side is $O(n^{\frac{23}{48}})$. $\square$

Our main result of this section is stated in the next

Theorem 13. *Assume* $p^q \leq 99999$. *Then,* $\binom{p^q n+1}{n}$ *is not squarefree for* $n \geq \tau_1 = \max\left(\frac{e^{60}-1}{p^q-1}, 5^{10}p^{5q}, \tau_0\right)$. *Moreover, the exceptions, for* $n < \tau_1$, *if they exist, are of the form* $\frac{p^{tq}-1}{p^q-1}$, *for any* $t \in \mathbf{N}$, *or of the form* $\frac{p^{qt_1+j_1}+\cdots+p^{qt_{m(p-1)+1}+j_{m(p-1)+1}}-1}{p^q-1}$, *for any* $t_i \in \mathbf{N}$, $1 \leq m \leq q-1$, $0 \leq j_i \leq q-1$, *and* $\sum_i p^{j_i} \equiv 1 \pmod{p^q-1}$.

We proceed to the proof of the theorem. Let $P = n(p^q n - n + 1)(p^q n + 1)$. Corollary 3.2 (p. 82) of [5] implies

Lemma 14. *Suppose that* $\binom{p^q n+1}{n}$ *is squarefree. Then,*

$$\begin{aligned} &\left|\sum_{d\in I}\psi\left(\frac{p^q n+1}{d}\right)\Lambda(d)\right| + \left|\sum_{d\in I}\psi\left(\frac{n}{d}\right)\Lambda(d)\right| \\ &+\left|\sum_{d\in I}\psi\left(\frac{(p^q-1)n+1}{d}\right)\Lambda(d)\right| \geq \frac{1}{2}\sum_{d\in I,(d,P)=1}\Lambda(d), \end{aligned} \tag{13}$$

where I *is the set of integers* d *in the range* $\sqrt{(p^q-1)n+1} < d \leq \sqrt{p^q n+1}$.

An immediate consequence of Lemma 7.1 of [5] (see also [16]) is

$$\left|\sum_{d\in I}\psi\left(\frac{X}{d}\right)\Lambda(d)\right| \leq \frac{1}{2R+2}\sum_{d\in I}\Lambda(d) + \left(\sum_{0<|r|\leq R}|a_r^{\pm}|\right)\max_{X\leq x\leq XR}\left|\sum_{d\in I}e\left(\frac{x}{d}\right)\Lambda(d)\right|,$$

where $a_r^{\pm} = \frac{i}{2\pi(R+1)}\left(\pi\left(1-\frac{|r|}{R+1}\right)\cot\left(\frac{\pi r}{R+1}\right)+\frac{|r|}{r}\right) \pm \frac{1}{2R+2}\left(1-\frac{|r|}{R+1}\right)$.

Taking $R = 10$ and using *Mathematica*[1] we obtained $\sum_{0<|r|\leq 10}|a_r^{\pm}| \sim 0.868 \leq \frac{86}{99}$, which implies

Lemma 15.

$$\left|\sum_{d\in I}\psi\left(\frac{X}{d}\right)\Lambda(d)\right| \leq \frac{1}{22}\sum_{d\in I}\Lambda(d) + \frac{86}{99}\max_{X\leq x\leq 10X}\left|\sum_{d\in I}e\left(\frac{x}{d}\right)\Lambda(d)\right|.$$

[1] A Trademark of Wolfram Research

Using (13) and the previous lemma we get

$$\begin{aligned}
\frac{1}{2}\sum_{d\in I,\,(d,P)=1}\Lambda(d) &\le \left|\sum_{d\in I}\psi\left(\frac{p^q n+1}{d}\right)\Lambda(d)\right| + \left|\sum_{d\in I}\psi\left(\frac{n}{d}\right)\Lambda(d)\right| \\
&+ \left|\sum_{d\in I}\psi\left(\frac{(p^q-1)n+1}{d}\right)\Lambda(d)\right| \le \frac{3}{22}\sum_{d\in I}\Lambda(d) \\
&+ \frac{86}{99}\max_{p^q n+1\le x\le 10(p^q n+1)}\left|\sum_{d\in I}e\left(\frac{x}{d}\right)\Lambda(d)\right| \\
&+ \frac{86}{99}\max_{(p^q-1)n+1\le x\le 10((p^q-1)n+1)}\left|\sum_{d\in I}e\left(\frac{x}{d}\right)\Lambda(d)\right| \\
&+ \frac{86}{99}\max_{n\le x\le 10n}\left|\sum_{d\in I}e\left(\frac{x}{d}\right)\Lambda(d)\right| \\
&\le \frac{3}{22}\sum_{d\in I}\Lambda(d) + \frac{86}{33}\max_{n\le x\le 10(p^q n+1)}\left|\sum_{d\in I}e\left(\frac{x}{d}\right)\Lambda(d)\right|.
\end{aligned}$$

Since, $\sum_{d\in I,\,(d,P)>1}\Lambda(d) \le \log n + \log((p^q-1)n+1) + \log(p^q n+1) \le 3\log n + 2q\log p$, for $n\ge 2$, we obtain

$$\sum_{d\in I}\Lambda(d) \le \frac{43}{6}\max_{n\le x\le 10(p^q n+1)}\left|\sum_{d\in I}e\left(\frac{x}{d}\right)\Lambda(d)\right| + \frac{11}{8}(3\log n + 2q\log p). \tag{14}$$

Schoenfeld [13], obtained, for $x \ge e^{30}$, (see also [10])

$$\left|\sum_{d\le x}\Lambda(d) - x\right| < \frac{1}{4\cdot 10^5}x.$$

Since $\sum_{d\in I}\Lambda(d) = \sum_{d\le\sqrt{p^q n+1}}\Lambda(d) - \sum_{d\le\sqrt{(p^q-1)n+1}}\Lambda(d)$, we obtain

$$\begin{aligned}
&\sqrt{p^q n+1} - \frac{1}{4\cdot 10^5}\sqrt{p^q n+1} - \sqrt{(p^q-1)n+1} - \frac{1}{4\cdot 10^5}\sqrt{(p^q-1)n+1} \\
&-\frac{11}{8}(3\log n + 2q\log p) = 0.9999975\sqrt{p^q n+1} - \\
&1.0000025\sqrt{(p^q-1)n+1} - \frac{11}{8}(3\log n + 2q\log p) \\
&< \frac{43}{6}\max_{n\le x\le 10(p^q n+1)}\left|\sum_{d\in I}e\left(\frac{x}{d}\right)\Lambda(d)\right|,
\end{aligned} \tag{15}$$

for $n \geq \dfrac{e^{60}-1}{p^q-1}$.

Now, we apply Theorem 9 of [5], a consequence of some very important bounds on exponential sums.

Theorem 16 (Granville-Ramaré). *If $k > 0$ integer and $y \leq \frac{1}{5}x^{3/5}$, then*

$$\left|\sum_{y\leq d\leq y'} e\left(\frac{x}{d}\right)\Lambda(d)\right| \leq \frac{50}{3}\, y \left(\frac{x}{y^{\frac{k+3}{2}}}\right)^{\frac{1}{4(2^k-1)}} (\log 16y)^{\frac{11}{4}},$$

for any $y \leq y' \leq 2y$.

Since $\sqrt{p^q n+1} \leq 2\sqrt{(p^q-1)n+1}$, using the above theorem of Granville and Ramaré we get, for $n > 5^{10}p^{5q}$ (to have the bound $y \leq \frac{1}{5}x^{3/5}$),

$$\max_{n\leq x\leq 10(p^q n+1)} \left|\sum_{d\in I} e\left(\frac{x}{d}\right)\Lambda(d)\right| \leq \max_{n\leq x\leq 10(p^q n+1)} \frac{50}{3}\sqrt{(p^q-1)n+1}\,\cdot$$
$$\left(\frac{x}{((p^q-1)n+1)^{\frac{k+3}{4}}}\right)^{\frac{1}{4(2^k-1)}} \left(\log\left(16\sqrt{(p^q-1)n+1}\right)\right)^{\frac{11}{4}}$$
$$= \frac{50}{3}\sqrt{(p^q-1)n+1}\left(\frac{10(p^q n+1)}{((p^q-1)n+1)^{\frac{k+3}{4}}}\right)^{\frac{1}{4(2^k-1)}}\cdot$$
$$\left(\log\left(16\sqrt{(p^q-1)n+1}\right)\right)^{\frac{11}{4}} \leq \frac{50}{3}\, 11^{\frac{1}{4(2^k-1)}}\, p^{\frac{q}{4(2^k-1)}}\,\cdot$$
$$n^{\frac{1}{4(2^k-1)}}\,((p^q-1)n+1)^{\frac{1}{2}-\frac{k+3}{4^2(2^k-1)}}\; 2^{-\frac{11}{4}}\,(\log(256(p^q-1)n+1))^{\frac{11}{4}} \leq$$
$$\frac{50}{3}\,2^{-\frac{11}{4}}\,11^{\frac{1}{4(2^k-1)}}\,p^{q\left(\frac{1}{2}-\frac{k-1}{4^2(2^k-1)}\right)}\,n^{\frac{1}{2}-\frac{k-1}{4^2(2^k-1)}}\,(\log(256((p^q-1)n+1)))^{\frac{11}{4}}.$$

We obtain (by taking $k=2$ - that will suffice for our purpose)

$$\max_{n\leq x\leq 10(p^q n+1)} \left|\sum_{d\in I} e\left(\frac{x}{d}\right)\Lambda(d)\right| \leq \frac{50}{3}\,2^{-\frac{11}{4}}\,11^{\frac{1}{12}}\,p^{\frac{23q}{48}}\,n^{\frac{23}{48}}\,(\log(256((p^q-1)n+1)))^{\frac{11}{4}}$$

By combining (15) and the previous inequality, we get

$$0.9999975\sqrt{p^q n+1} - 1.0000025\sqrt{(p^q-1)n+1} \leq$$
$$21.683\,p^{\frac{23q}{48}}\,n^{\frac{23}{48}}\,(\log(256((p^q-1)n+1)))^{\frac{11}{4}} + \frac{11}{8}(3\log n + 2q\log p), \tag{16}$$

which is certainly false for n sufficiently large by Lemma 12. Thus the assumption in Lemma 14 is false, which implies our Theorem 13.

Remark 17. *The inequality (10) provides explicit bounds for n for any choice of p and q, with $p^q \leq 99999$. We can increase the bound for p^q, by using a weaker result of Schoenfeld [13]. However, in doing that we increase the bound on n as well, so we preferred a better bound on n.*

Theorem 18. $\binom{4n+1}{n}$ *is squarefree for $n \geq 2^{1518}$, and if $n < 2^{1518}$ there are at most $289,179 = \binom{761}{2} \sim 2^{18.2}$ possible exceptions.*

Proof. If $(p,q) = (2,2)$, the inequality (10), valid for $n \geq \max\left(\frac{e^{60}-1}{3}, 5^{10}2^{10}\right)$ changes into

$$\begin{aligned} &0.9999975\sqrt{4n+1} - 1.0000025\sqrt{3n+1} > \\ &42.1311\, n^{\frac{23}{48}} \left(\log\left(768n+1\right)\right)^{\frac{11}{4}} + \frac{33}{8}\log n + 1.65566, \end{aligned} \tag{17}$$

which is certainly true for $n \geq 2^{1518}$. Theorem 6 and Corollary 11 imply that the exceptions (if they exist) are of the form $\dfrac{2^{2t+1}+2^{2j+1}-1}{3}$, $j \leq t$. Therefore, since the number of pairs (j,t), giving different numbers of the above form, is less than $\binom{761}{2} \sim 2^{18.2}$, we get the theorem. □

Theorem 19. $\binom{4n+1}{n}$ *is squarefree for $n \geq 3^{956}$, and if $n < 3^{956}$, there are at most $18088476 = \binom{478}{3} \sim 3^{15.3}$ possible exceptions.*

Proof. If $(p,q) = (3,2)$, the inequality (10), valid for $n \geq \max\left(\frac{e^{60}-1}{8}, 5^{10}3^{10}\right)$, changes into

$$\begin{aligned} &0.9999975\sqrt{9n+1} - 1.0000025\sqrt{8n+1} > \\ &26.04\, n^{\frac{23}{48}} \left(\log\left(2048n+1\right)\right)^{\frac{11}{4}} + \frac{33}{8}\log n + 2.62417, \end{aligned} \tag{18}$$

which is true for $n \geq 3^{956}$. As in the previous proof, we get that the exceptions (if they exist) are of the form $\dfrac{2^{2t+1} + 2^{2j+1} + 2^{2i+1} - 1}{8}$, $i \leq j \leq t$. Therefore, since the number of triples (i, j, t), giving different numbers of the above form, is less than $\binom{478}{3} \sim 3^{15.3}$, we get the result. □

Although, there are not that many possible exceptions, because of their size, to check each value in acceptable time, is beyond any computer's capability at this moment, and we were not able to decrease the complexity. However, we conjecture

Conjecture 20. *Except for* 1*,* 3 *and* 45*,* $\binom{4n+1}{n}$ *is not squarefree. Except for* 1*,* 4 *and* 10*,* $\binom{9n+1}{n}$ *is not squarefree.*

Acknowledgements. The author would like to thank the anonymous referee for his helpful comments, which improved significantly the presentation of the paper.

References

[1] M. Bernstein, N.J.A. Sloane, *Some Canonical Sequences of Integers*, Linear Algebra and Its Applications, **226-228** (1995), pp. 57-72.

[2] W. Chu, *A new combinatorial interpretation for generalized Catalan numbers*, Discrete Mathematics **65** (1987), pp. 91-94.

[3] L.E. Dickson, *History of the Theory of Numbers*, vol. I, Chap. XI, Chelsea, New York, 1919.

[4] A. Granville, *Binomial coefficients modulo prime powers*, Canadian Mathematical Soc. Conf. Proceed. **20** (1997), pp. 253-275.

[5] A. Granville, O. Ramaré, *Explicit bounds on the exponential sums and the scarcity of squarefree binomial coefficients*, Mathematika **43** (1996), pp. 73-107.

[6] R.K. Guy, *Unsolved Problems in Number Theory*, Springer-Verlag, New York, pp. 47-50, 1981.

[7] P. Hilton, J. Pederson, *Catalan numbers, their generalization, and their uses*, Math. Intelligencer **13** (1991), pp. 64-75.

[8] D.S. Hough, R.E. Simion, *Problem P10794*, The American Mathematical Monthly, vol.107, Number 3, March 2000.

[9] D.A. Klarner, *Correspondences between plane trees and binary sequences*, J. Combin. Theory Ser. A, **9** (1970), pp. 401-411.

[10] N.C. Pereira, *Estimates for the Chebyshev Function* $\psi(x) - \theta(x)$, Math. of Comput. **440** (1985), pp. 211-221.

[11] A.D. Sands, *On generalized Catalan numbers*, Discrete Mathematics **22** (1978), pp. 218-221.

[12] A. Sàrközy, *On divisors of binomial coefficients I*, Journal of Number Theory **20** (1985), pp. 70-80.

[13] L. Schoenfeld, *Sharper Bounds for the Chebyshev Functions* $\theta(x)$ *and* $\psi(x)$, Math. of Comput. **30** (1976), pp. 337-360.

[14] N.J.A. Sloane, Online Encyclopedia of Integer Sequences, AT&T Research Labs, *http://www.research.att.com/˜njas/sequences/index.html.*

[15] R. Stanley, *Enumerative Combinatorics*, vol. 2, Cambridge University Press, 1999.

[16] J.D. Vaaler, *Some extremal functions in Fourier analysis*, Bull. Amer. Math. Soc. **12** (1985), pp. 183-216.