Asymptotic density of Catalan numbers modulo 3 and powers of 2

Rob Burns

3rd March 2022

Abstract

We establish the asymptotic density of the Catalan numbers modulo 3 and modulo $2^k$ for $k \in \mathbb{N}$ and $k \geq 1$.

1 Introduction

The Catalan numbers are defined by

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

There has been much work in recent years and also going back to Kummer [6] on analysing the Catalan numbers modulo primes and prime powers. Deutsch and Sagan [2] provided a complete characterisation of Catalan numbers modulo 3. A characterisation of the Catalan numbers modulo 2 dates back to Kummer. Eu, Liu and Yeh [3] provided a complete characterisation of Catalan numbers modulo 4 and 8. This was extended by Liu and Yeh [8] to a complete characterisation modulo 8, 16 and 64. This result was restated in a more compact form by Kauers, Krattenthaler and Müller in [4] by representing the generating function of $C_n$ as a polynomial involving a special function. The polynomial for $C_n$ modulo 4096 was also calculated. A method for extracting the coefficients of the generating function (i.e. $C_n$ modulo a prime power) was provided, though given the complexity of the polynomials (the polynomial for the 4096 case takes a page and a half to write down) this would need to be done by computer. Krattenthaler and Müller [5] used a similar method to examine $C_n$ modulo powers of 3. They wrote down the polynomial for the generating function of $C_n$ modulo 9 and 27 thereby generalised the mod 3 result of [2]. The article by Lin [7] discussed the possible values of the odd Catalan numbers modulo $2^k$ and Chen and Jiang [1] dealt with the possible values of the Catalan numbers modulo prime powers. Rowland and Yassawi [9] investigated $C_n$ in the general setting of
automatic sequences. The values of $C_n$ (as well as other sequences) modulo prime powers can be computed via automata. Rowland and Yassawi provided algorithms for creating the relevant automata. They established a full characterisation of $C_n$ modulo $\{2, 4, 8, 16, 3, 5\}$ in terms of automata. They also extended previous work by establishing forbidden residues for $C_n$ modulo $\{32, 64, 128, 256, 512\}$. In theory the automata can be constructed for any prime power but computing power and memory quickly becomes a barrier.

Some of this work can be used to determine the asymptotic densities of the Catalan numbers modulo $2^k$ and $3$.

In section 2 we will use a result of Liu and Yeh [8] to obtain some asymptotic densities of the Catalan numbers modulo powers of 2. Here, the asymptotic density of a subset $S$ of $\mathbb{N}$ is defined to be

$$\lim_{N \to \infty} \frac{1}{N} \#\{n \in S : n \leq N\}$$

if the limit exists, where $\#S$ is the number of elements in a set $S$. Since the Catalan numbers $C_n$ are highly multiplicative as $n$ increases, it is expected that for a fixed $m \in \mathbb{N}$ “almost all” Catalan numbers are divisible by $m$. We show that this is the case when $m = 3$ and $m = 2^k$ for $k \geq 1$.

## 2 Catalan numbers modulo $2^k$

Firstly, as in [8], let the $p$-adic order of a positive integer $n$ be defined by

$$\omega_p(n) := \max\{\alpha \in \mathbb{N} : p^\alpha | n\}$$

and the cofactor of $n$ with respect to $p^{\omega_p(n)}$ be defined by

$$CF_p(n) := \frac{n}{p^{\omega_p(n)}}.$$ 

In addition the function $\alpha : \mathbb{N} \to \mathbb{N}$ is defined by

$$\alpha(n) := \frac{CF_2(n + 1) - 1}{2}.$$ 

For a number $p$, we write the base $p$ expansion of a number $n$ as

$$[n]_p = \langle n_r n_{r-1} \ldots n_1 n_0 \rangle$$

where $n_i \in [0, p - 1]$ and

$$n = n_r p^r + n_{r-1} p^{r-1} + \ldots + n_1 p + n_0.$$
Then the function $d_p : \mathbb{N} \to \mathbb{N}$ is defined by
\[
d_p(n) := \#\{i : n_i = 1\}.
\] (1)

Here we will only be interested in the case $p = 2$ and will refer to $d_2$ as merely $d$.

When $p = 2$, $d(n)$ is the sum of the digits in the base 2 representation of $n$.

The following result appears in [8].

**Theorem 1** (Corollary 4.3 of [8]). In general, we have $\omega_2(C_n) = d(\alpha(n))$.

In particular, $C_n \equiv 0 \mod 2^k$ if and only if $d(\alpha(n)) \geq k$, and $C_n \equiv 2^{k-1} \mod 2^k$ if and only if $d(\alpha(n)) = k - 1$.

Since we are looking at the Catalan numbers $C_n \mod 2^k$ it will be convenient to first consider numbers $n < 2^t$ for a fixed but arbitrary $t \in \mathbb{N}$. This produces some interesting and simple formulae. We have the following results:

**Theorem 2.** $\#\{n < 2^t : d(\alpha(n)) = k\} = \left(\begin{array}{c} t \\ k+1 \end{array}\right)$.

**Proof.** Let
\[
\alpha = \alpha(n) = \frac{CF_2(n + 1) - 1}{2}.
\]

Then $n + 1 = 2^s(2\alpha + 1)$ for some arbitrary $s \in \mathbb{N} : s \geq 0$. Writing $n + 1$ in base 2 we have
\[
[n + 1]_2 = \langle [\alpha]_2, 0...0, 0 \rangle
\]
where there are $s$ 0’s at the end and $s \geq 0$ is arbitrary. So,
\[
[n]_2 = \langle [\alpha]_2, 0...1, 1 \rangle
\]
where there are $s$ 1’s at the end and $s \geq 0$ is arbitrary. It can be seen that $d(n) = d(\alpha) + s$. Since $n < 2^t$ and $d(\alpha(n)) = k$ the possible base 2 representations of $n$ are
\[
[n]_2 = \langle [\alpha]_2, 0 : \alpha < 2^{t-1} \rangle
\]
\[
[n]_2 = \langle [\alpha]_2, 0, 1 : \alpha < 2^{t-2} \rangle
\]
\[
\vdots
\]
\[
[n]_2 = \langle [\alpha]_2, 0, 1, ..., 1 : \alpha < 2^k \rangle
\]
and there are $(t - k - 1)$ 1’s at the end of the last representation.

Therefore, counting each of these possibilities and using the fact that $d(\alpha) = k$ gives
\[
\#\{n < 2^t : d(\alpha(n)) = k\} = \sum_{i=k}^{t-1} \left(\begin{array}{c} i \\ k \end{array}\right) = \left(\begin{array}{c} t \\ k+1 \end{array}\right).
\]

\[\square\]
Corollary 3. \(#\{n < 2^t : C_n \equiv 0 \mod 2^k\} = \sum_{i=k+1}^{t} \binom{t}{i}\).

Proof. The corollary follows from Theorem 1 and Theorem 2 above since
\[
\#\{n < 2^t : C_n \equiv 0 \mod 2^k\} = \#\{n < 2^t : d(\alpha(n)) \geq k\} \text{ from Theorem 1}
\]
\[
= 2^t - 1 - \sum_{i=0}^{k-1} \#\{n < 2^t : d(\alpha(n)) = i\}
\]
\[
= 2^t - 1 - \sum_{i=0}^{k-1} \binom{t}{i+1} \text{ from Theorem 2}
\]
\[
= 2^t - \sum_{i=0}^{k} \binom{t}{i}
\]
\[
= \sum_{i=k+1}^{t} \binom{t}{i} \text{ since } 2^t = \sum_{i=0}^{t} \binom{t}{i}
\]
\[\square\]

Corollary 4. \(#\{n < 2^t : C_n \equiv 2^{k-1} \mod 2^k\} = \binom{t}{k}\).

Proof. From Theorem 1,
\[
\#\{n < 2^t : C_n \equiv 2^{k-1} \mod 2^k\} = \#\{n < 2^t : d(\alpha(n)) = k - 1\}
\]
\[
= \binom{t}{k} \text{ from Theorem 2.}
\]
\[\square\]

Theorem 2 can be used to establish the asymptotic density of the set
\[
\{n < N : C_n \equiv 0 \mod 2^k\}
\]

Theorem 5. For \(k \in \mathbb{N}\) with \(k \geq 1\) we have
\[
\lim_{N \to \infty} \frac{1}{N} \#\{n < N : C_n \equiv 0 \mod 2^k\} = 1.
\]

4
Proof. Let \( N \in \mathbb{N} \) and choose \( r \in \mathbb{N} \) such that

\[
2^r \leq N < 2^{r+1}
\]

Then \( r = \lfloor \log_2(N) \rfloor \) and

\[
\#\{n < N : d(\alpha(n)) = k\} \\
\leq \#\{n < 2^{r+1} : d(\alpha(n)) = k\} \\
= \binom{r+1}{k+1}.
\]

So,

\[
\#\{n < N : C_n \equiv 0 \pmod{2^k}\} \\
= N - \#\{n < N : d(\alpha(n)) < k\} \\
= N - \sum_{i=0}^{k-1} \#\{n < N : d(\alpha(n)) = i\} \\
\geq N - \sum_{i=0}^{k-1} \binom{r+1}{i+1}.
\]

And so,

\[
\frac{1}{N} \#\{n < N : C_n \equiv 0 \pmod{2^k}\} \\
\geq 1 - \frac{1}{N} P_k(r) \\
= 1 - \frac{1}{N} P_k(\lfloor \log_2(N) \rfloor)
\]

(2)

where \( P_k(r) \) is a polynomial in \( r \) of degree \( k \) with coefficients depending on \( k \). Since \( k \) is fixed and

\[
\lim_{N \to \infty} \frac{1}{N} (\lfloor \log_2(N) \rfloor)^k = 0
\]

for fixed \( k \), the second term in (2) is zero in the limit and

\[
\lim_{N \to \infty} \frac{1}{N} \#\{n < N : C_n \equiv 0 \pmod{2^k}\} = 1.
\]

\( \square \)
3 Catalan numbers modulo 3

Deutsch and Sagan [2] provided a characterisation of the Catalan numbers modulo 3. They defined a set \( T^*(01) \) of natural numbers using the base 3 representation. If \( [n]_3 = (n_i) \) is the base 3 representation of \( n \) then

\[ T^*(01) = \{ n \geq 0 : n_i = 0 \text{ or } 1 \text{ for all } i \geq 1 \} . \]

They then defined the function \( d^*_3(n) \) similar to \( d_p(n) \) in (1) as

\[ d^*_3(n) = \#\{ n_i : i \geq 1, n_i = 1 \} \]

**Theorem 6.** The asymptotic density of the set \( T^*(01) \) is zero.

**Proof.** Let \( N \in \mathbb{N} \) and choose \( k \in \mathbb{N} : 3^k \leq N < 3^{k+1} \). Then \( k = \lfloor \log_3(N) \rfloor \) and

\[
\frac{1}{N} \# \{ n \leq N : n \in T^*(01) \} \\
\leq \frac{1}{N} 3 \times 2^k \\
\leq 3^{1-k} \times 2^k \to 0 \text{ as } N \to \infty.
\]

The following theorem comes from [2].

**Theorem 7.** (Theorem 5.2 of [2]) The Catalan numbers satisfy

\[
C_n \equiv (-1)^{d^*_3(n+1)} \mod 3 \text{ if } n \in T^*(01) - 1 \\
C_n \equiv 0 \mod 3 \text{ otherwise}.
\]

**Corollary 8.**

\[
\lim_{N \to \infty} \frac{1}{N} \# \{ n < N : C_n \equiv 0 \mod 3 \} = 1.
\]

**Proof.** Combining theorems 6 and 7, the set of \( n \) such that \( C_n \) is not congruent to \( 0 \mod 3 \) has asymptotic density 0.
References

[1] Yong-Gao Chen and Wen Jiang. Catalan numbers modulo a prime power. *Integers*, 13(A36), 2013. 1

[2] E. Deutsch and B.E. Sagan. Congruences for Catalan and Motzkin numbers and related sequences. *Journal of Number Theory*, 117(1):191–215, 2006. 1, 6

[3] Sen-Peng Eu, Shu-Chung Liu, and Yeong-Nan Yeh. Catalan and Motzkin numbers modulo 4 and 8. *European Journal of Combinatorics*, 29:1449–1466, 2008. 1

[4] Manuel Kauers, Christian Krattenthaler, and Thomas W. Müller. A method for determining the mod-$2^k$ behaviour of recursive sequences, with applications to subgroup counting. *The Electronic Journal of Combinatorics*, 18(2), 2012. 1

[5] Christian Krattenthaler and Thomas W. Müller. A method for determining the mod-$3^k$ behaviour of recursive sequences. *ArXiv*, http://arxiv.org/abs/1308.2856, 2013. 1

[6] E. E. Kummer. Über die Ergänzungssätze zu den allgemeinen Reciprocitatsgesetzen. *J. Reine Angew. Math.*, 44:93–146, 1852. 1

[7] Hsueh-Yung Lin. The odd Catalan numbers modulo $2^k$. *Integers*, 12(2):161–165, 2012. 1

[8] S.-C. Liu and J. Yeh. Catalan numbers modulo $2^k$. *J Integer Sequences*, 13, 2010. 1, 2, 3

[9] Eric Rowland and Reem Yassawi. Automatic congruences for diagonals of rational functions. *ArXiv*, https://arxiv.org/abs/1310.8635, 2013. 1