# ON THE EULER FUNCTION OF THE CATALAN NUMBERS

FLORIAN LUCA AND PANTELIMON STĂNICĂ

Abstract. We study the solutions of the equation $\phi(C_m)/\phi(C_n) = r$, where $r$ is a fixed rational number, $C_k$ is the $k$th Catalan number and $\phi$ is the Euler function. We note that the number $r = 4$ is special for this problem and for it we construct solutions $(m, n)$ to the above equation which are related to primes $p$ such that $2p - 1$ or $4p - 3$ is also prime.

## 1. An observation concerning $\phi(C_{n+1})/\phi(C_n)$

For a positive integer $n$, let

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

be the $n$-th Catalan number. For a positive integer $m$ we put $\phi(m)$ for the Euler function of $m$.

A Carmichael's conjecture [5], which is still open, states that for every $n$ it is possible to find an $m \neq n$ such that $\phi(m) = \phi(n)$. Since this problem seems to be currently out of reach, one would look at the behavior of the Euler's phi function $\phi(\bullet)$, or at quotients $\phi(\bullet)/\phi(\bullet)$, when the arguments belong to some smaller classes of integers, like the binomial coefficients, binary recurrent sequences, or even Catalan numbers. In fact, there is a growing literature on arithmetic functions with binomial coefficients [12], [15] and [17], or on arithmetic functions with members of binary recurrent sequences [3, 11, 13, 14, 16, 18, 19, 20, 21], etc.

At first, we wanted to test whether we could find distinct $m$ and $n$ such that $\phi(C_m) = \phi(C_n)$ but did not find solutions other than the trivial solution $\phi(C_1) = \phi(C_2) = 1$. So, we checked numerically for the values of the ratios $\phi(C_m)/\phi(C_n)$ for $m \neq n$. While computing such ratios for small values of $m$ and $n$, we first noted, then we proved, the following result.

**Theorem 1.1.** *The equality*

$$\phi(C_{n+1}) = 4\phi(C_n) \tag{2}$$

*holds in each of the following two instances:*

---

*Date*: February 26, 2012.
*Key words and phrases.* Catalan Numbers; factorials; Euler functions
*Mathematics Subject Classification*: 11A25, 11N37, 11Y35, 11B65.

(i) $n = 2p - 2$, *where* $p \geq 5$ *is a prime such that* $q = 4p - 3$ *is also a prime.*
(ii) $n = 3p - 2$, *where* $p > 5$ *is a prime such that* $q = 2p - 1$ *is also prime.*

*Proof.* We have

$$C_{n+1} = \frac{2(2n+1)}{n+2} C_n. \tag{3}$$

For (i), we use (3) with $n = 2p - 2$ where both $p$ and $q = 4p - 3$ are primes, getting

$$pC_{n+1} = qC_n. \tag{4}$$

Hence,

$$C_{n+1} = qC \qquad \text{and} \qquad C_n = pC \tag{5}$$

for some positive integer $C$. Since $q = 2n + 1 > 2n$, it follows that $q$ does not divide $C_n$, so in particular $q$ does not divide $C$. Since $p = (n + 2)/2$, it follows that $p^3 \| (2n)!$ and $p^2 \| n!(n + 1)!$, so $p \| C_n$. Here and in what follows, for a prime $p$ and positive integers $a$ and $m$ we write $p^a \| m$ when $p^a \mid m$ but $p^{a+1} \nmid m$. It follows that $p \nmid C$. Thus, $\gcd(pq, C) = 1$. Applying the Euler function to the equalities in (5), and taking the ratio of the resulting relations we get

$$\frac{\phi(C_{n+1})}{\phi(C_n)} = \frac{(q-1)\phi(C)}{(p-1)\phi(C)} = 4, \tag{6}$$

because $q - 1 = 2n = 4(p - 1)$, which is what we wanted. The argument for (ii) is similar. Namely, in this case $n + 2 = 3p$ and $2n + 1 = 6p - 3 = 3q$, so that instead of relation (4) we get

$$pC_{n+1} = 2qC_n. \tag{7}$$

Hence,

$$C_{n+1} = 2qC \qquad \text{and} \qquad C_n = pC \tag{8}$$

for some positive integer $C$. Let us first see that $C$ is even. If $C$ is odd, then $C_n$ is odd, therefore $n = 2^a - 1$ for some positive integer $a$ (see [1]). Thus, $p = (n + 2)/3 = (2^a + 1)/3$, is an integer, so $a$ is odd. Further, since $p > 5$, it follows that $a \geq 5$. Next,

$$q = 2p - 1 = \frac{2^{a+1} - 1}{3} = \frac{1}{3}(2^{(a+1)/2} + 1)(2^{(a+1)/2} - 1).$$

Since $a \geq 5$, it follows that both numbers $2^{(a+1)/2} + 1$ and $2^{(a+1)/2} - 1$ are larger than 3, and in particular, $q = (2^{a+1} - 1)/3$ cannot be prime, which is a contradiction. This shows that $C$ is even. Since $q = (2n + 1)/3$, it follows that $q^2 \| (2n)!$ and $q^2 \| n!(n + 1)!$, so $q \nmid C_n$. Thus, $q \nmid C$. Since $p = (n + 2)/3$, it follows that $p^5 \| (2n)!$ and $p^4 \| n!(n + 1)!$, so $p \| C_n$. In particular, $p \nmid C$.

Thus, we have that $\gcd(pq, C) = 1$. Taking the Euler function in relations (8) and dividing the resulting expressions we get

$$\frac{\phi(C_{n+1})}{\phi(C_n)} = \frac{2(q-1)\phi(C)}{(p-1)\phi(C)} = 4,$$

because $(q-1) = 2(n-1)/3 = 2(p-1)$, which is what we wanted. $\square$

## 2. The Main result

As we shall see later, there are many other solutions to (2) and we shall explain how to find some of them. We do not know if there are infinitely many primes $p$ such that $4p-3$ is a prime, or $2p-1$ is a prime. It follows by the Hardy and Littlewood conjectures (see [9]) that for large $x$ the number of such primes should be asymptotically $c_0\, x/(\log x)^2$ for some positive constant $c_0$. We asked ourselves whether it is likely for some positive integer $n$ to exist another positive integer $m$ such that $\phi(C_n)/\phi(C_m)$ is a fixed rational number $r$. In the above, we allow $r = 1$, but in this case we impose that $m \neq n$. More precisely, for a fixed $r \in \mathbb{Q}_+$, define the following set

$$\mathcal{N}_r = \left\{ n \ : \ \frac{\phi(C_m)}{\phi(C_n)} = r \text{ holds for some } m \neq n \right\}. \tag{9}$$

For a large real number $x$ we put $\mathcal{N}_r(x) = \mathcal{N}_r \cap [1, x]$. Computer experiments turned up lots of solutions for $r = 4$ and the symmetrical $r = 1/4$, but very few solutions for other values of $r$. We asked ourselves if $r = 4$ and $r = 1/4$ are special in this respect. Our main result below together with the above Hardy and Littlewood conjectures seem to indicate that this is indeed the case.

In what follows, we use the Landau symbols $O$ and $o$ as well as the Vinogradov symbols $\ll$, $\gg$ and $\asymp$ and $\sim$ with their usual meanings. The constant and speed of convergence implied by them might depend on our parameter $r$. Recall that $A = O(B)$, $A \ll B$ and $B \gg A$ are all equivalent and mean that the inequality $|A| < cB$ holds with some positive constant $c$. Further, $A \asymp B$ means that both $A \ll B$ and $B \ll A$ hold, $A = o(B)$ means that $A/B$ tends to zero, whereas $A \sim B$ means that $A/B$ tends to 1. We use $c_0, c_1, \ldots$ for positive constants which might depend on our parameter $r$. We write $P(m)$ and $p(m)$ for the largest and smallest prime factor of the positive integer $m$, respectively. We write $p$, $q$ and $\rho$ with or without subscripts for prime numbers. For a positive real number $x$ we write $\log x$ for the natural logarithm of $x$.

**Theorem 2.1.** *The estimate*

$$\#\mathcal{N}_r(x) \le \frac{x}{(\log x)^{3+o(1)}} \quad \textit{holds for} \quad r \notin \{4, 1/4\} \quad \textit{as} \quad x \to \infty. \tag{10}$$

*However,*

$$\#\mathcal{N}_r(x) \ll \frac{x}{(\log x)^2} \quad \textit{holds when} \quad r = 4,\ 1/4 \quad \textit{for all} \quad x > 10. \tag{11}$$

## 3. The proof of Theorem 2.1

Since $r$ is fixed, we write only $\mathcal{N}(x)$ and omit the dependence on $r$. We let $x$ be large and let $\mathcal{M}(x) = \mathcal{N} \cap (x/2, x]$. It is enough to prove that the upper bounds (10) and (11) hold on $\#\mathcal{M}(x)$, since afterwards the same upper bounds on $\#\mathcal{N}(x)$ will follow by replacing $x$ with $x/2$, then with $x/4$, and so on.

3.1. **An upper bound for $|m-n|$.** We use the asymptotic

$$C_n \sim c_1 \frac{2^{2n}}{n^{3/2}} \qquad \text{as} \quad n \to \infty, \tag{12}$$

where $c_1 = 1/\sqrt{\pi}$ (see Exercise 9.8 in [6]). We also use the fact that the bounds

$$\frac{\ell}{\log\log \ell} \ll \phi(\ell) \le \ell \tag{13}$$

hold for all positive integers $\ell \ge 3$ (see Theorem 328 in [10]). Using estimate (13) with $\ell = C_m$ and $\ell = C_n$, we get that

$$\frac{2^{2m}}{m^{3/2}\log m} \ll \phi(C_m) \ll \frac{2^{2m}}{m^{3/2}} \quad \text{and} \quad \frac{2^{2n}}{n^{3/2}\log n} \ll \phi(C_n) \ll \frac{2^{2n}}{n^{3/2}}. \tag{14}$$

Assume now that $n \in \mathcal{M}(x)$ and that $m \ne n$ is such that $r = \phi(C_m)/\phi(C_n)$. Taking logarithms and using estimates (14), we get

$$|\log r| = |\log(\phi(C_m)/\phi(C_n))| = 2|m-n|\log 2 + O(\log(m+n)).$$

The above estimate shows that $m = n + O(\log x)$. We return to (14) and observe that in fact it yields

$$\begin{aligned} \log \phi(C_m) &= 2m\log 2 - (3/2)\log m + O(\log\log m), \\ \log \phi(C_n) &= 2n\log 2 - (3/2)\log n + O(\log\log n). \end{aligned} \tag{15}$$

Applying estimate (15) with $n$ and $m$ and taking the difference of the resulting relations, we get that

$$\begin{aligned} \log\phi(C_m) - \log\phi(C_n) &= 2(m-n)\log 2 - (3/2)\log(m/n) + O(\log\log x) \\ &= 2(m-n)\log 2 - (3/2)\log\left(1 + O\left(\frac{\log x}{x}\right)\right) \\ &+ O(\log\log x) \\ &= 2(m-n)\log 2 + O(\log\log x). \end{aligned}$$

We thus get that

$$|\log r| = |\log(\phi(C_m)/\phi(C_n))| = 2|m-n|\log 2 + O(\log\log x),$$

which implies that $m = n + O(\log\log x)$.

Let $c_2$ be the constant implied by the previous $O$-symbol. We also let $K = c_2 \log\log x$. Thus, $m = n + k$, where $0 < |k| \le K$. We write $\mathcal{M}^{(k)}(x)$

for the set of $n \in \mathcal{M}(x)$ for which there exists $m$ with $m = n + k$ such that $\phi(C_m)/\phi(C_n) = r$. Clearly,

$$\#\mathcal{M}(x) \le \sum_{0<|k|\le K} \#\mathcal{M}^{(k)}(x).$$

It remains to estimate $\#\mathcal{M}^{(k)}(x)$. We treat only the case of the positive number $k$, since the case when $k$ is negative can be dealt with in a similar way. We fix the number $k \le K$.

3.2. **Deducing the $STMN$ equation.** We have

$$C_m = \frac{2^k(2n+1)(2n+3)\cdots(2n+2k-1)}{(n+2)(n+3)\cdots(n+k+1)} C_n, \tag{16}$$

so that

$$(n+2)\cdots(n+k+1)C_m = 2^k(2n+1)(2n+3)\cdots(2n+2k-1)C_n. \tag{17}$$

Observe that if $p \mid n+i+1$ and $p \mid 2n+2j-1$ holds for some $i, j \in \{1, \ldots, k\}$, then $p \mid 2j - 2i - 3$, and this last number is odd and has absolute value at most $2K + 1$. Thus, such primes $p$ are at most $2K + 1$. The same is true for prime factors $p$ common to $n + i + 1$ and $n + i_1 + 1$ for some $i \ne i_1$ in $\{1, 2, \ldots, k\}$, as well as for prime factors $p$ common to both $2n + 2j - 1$ and $2n + 2j_1 - 1$ for some $j \ne j_1$ also in the set $\{1, 2, \ldots, k\}$. The above relation (17) can be written as

$$UMC_m = VNC_n, \tag{18}$$

where $U, V$ are coprime integers with $P(UV) \le 2K + 1$, $M$ and $N$ are coprime integers with $p(MN) > 2K + 1$, and

$$(n+2)\cdots(n+k+1) = UMD \quad \text{and} \quad 2^k(2n+1)\cdots(2n+2k-1) = VND, \tag{19}$$

for some positive integer $D$ with $P(D) \le 2K + 1$, where

$$D = \gcd\Big((n+2)\cdots(n+k+1), 2^k(2n+1)\cdots(2n+2k-1)\Big).$$

Equation (18) gives

$$C_m = VNC \qquad \text{and} \qquad C_n = UMC \tag{20}$$

for some positive integer $C$. Write $C = \Gamma_U\Gamma_V\Gamma_M\Gamma_N\Gamma$, where

$$\Gamma_I = \prod_{\substack{p^{\alpha_p}\|C \\ p|I}} p^{\alpha_p} \qquad \text{for each of} \quad I \in \{U, V, M, N\},$$

and $\Gamma$ is the largest divisor of $C$ which is coprime to $UVMN$. We now apply the Euler function to the two relations (20) getting

$$\begin{aligned} \phi(C_m) &= \phi(V)\phi(N)\Gamma_V\Gamma_N\phi(\Gamma_U)\phi(\Gamma_M)\phi(\Gamma), \\ \phi(C_n) &= \phi(U)\phi(M)\Gamma_U\Gamma_M\phi(\Gamma_V)\phi(\Gamma_N)\phi(\Gamma). \end{aligned} \tag{21}$$

Write $r = u/v$ with coprime positive integers $u$ and $v$. Observe that the relation $r = \phi(C_m)/\phi(C_n)$, is the same as $u\phi(C_n) = v\phi(C_m)$, which, via the relations (21), leads to

$$S\,\phi(M)\,\Gamma_M\,\phi(\Gamma_N) = T\,\phi(N)\,\Gamma_N\,\phi(\Gamma_M), \tag{22}$$

where $S = u\phi(U)\Gamma_U\phi(\Gamma_V)$ and $T = v\phi(V)\Gamma_V\phi(\Gamma_U)$ have the property that $P(ST) \le 2K+1$ provided that $x$ is large enough, say large enough such that $2K+1 \ge \max\{u, v\}$. We refer to (22) as the $STMN$-*equation.*

3.3. **Large and very large primes.** Next let $c_3$ be some absolute constant to be determined later and put $y = (\log x)^{10}$ and $z = x^{1/(c_3 \log\log x)}$. We also put $\mathcal{J} = (y, z]$. We say that a prime $p$ is *large* if $p > y$ and *very large* if $p > z$. Hence, primes in $\mathcal{J}$ are large but not very large.

3.4. **The case when $(n+1)MN$ is divisible by the square of a large prime.** Let $\mathcal{M}_1^{(k)}(x)$ be the subset of $n \in \mathcal{M}^{(k)}(x)$ for which $p^2 \mid (n+1)MN$ for some large prime $p$. We assume that $x$ is sufficiently large such that $y > 2K+3$. It then follows that either there exists $i \in \{0, 1, 2, \ldots, k\}$ such that $p^2 \mid n+i+1$, or $j \in \{1, 2, \ldots, k\}$ such that $p^2 \mid 2n+2j-1$. Since

$$\max\{n+i+1, 2n+2j-1\} \le x + 2K - 1 < 2x,$$

for large $x$, it follows that the number of such positive integers $n \le x$ for a fixed $i$ (or $j$) is at most $2x/p^2$. Varying $i$ (or $j$) in $\{1, \ldots, k\}$, it follows that the number of such possibilities is $\le 4(k+1)x/p^2$. Summing this up over all the large primes $p$, we get that

$$\begin{aligned}\#\mathcal{M}_1^{(k)}(x) &\le 4(K+1)\sum_{p \ge y}\frac{x}{p^2} \le 8Kx\left(\frac{1}{y^2} + \int_y^\infty \frac{dt}{t^2}\right) \le \frac{16Kx}{y}\\ &\ll \frac{x(\log\log x)}{(\log x)^{10}}.\end{aligned}$$

Putting $\mathcal{M}_1(x) = \bigcup_{0<|k|\le K}\mathcal{M}_1^{(k)}(x)$, we get that

$$\#\mathcal{M}_1(x) \le \sum_{0<|k|\le K}\#\mathcal{M}_1^{(k)}(x) \ll \frac{x(\log\log x)^2}{(\log x)^{10}} = o\left(\frac{x}{(\log x)^3}\right), \tag{23}$$

as $x \to \infty$. The bound (23) is acceptable for us. From now on, we work under the assumption that $(n+1)MN$ is not divisible by squares of large primes.

3.5. **The instance when $n$ or $m$ has few large digits in a prime base $p \in \mathcal{J}$ dividing $MN$.** Now we assume that $MN$ is divisible by some prime $p \in \mathcal{J}$. Then $p$ divides either $n+i+1$ for some $i \in \{1, 2, \ldots, k\}$ or $2n+2j-1$ for some $j \in \{1, 2, \ldots, k\}$. The situation here is entirely symmetric so we only consider the case when $p$ divides $n+i+1$ for some $i \in \{1, 2, \ldots, k\}$. Let

$$n = n_0p^\lambda + n_1p^{\lambda-1} + \cdots + n_\lambda$$

be the base $p$ representation of $n$. Observe that $n_\lambda = p - i - 1$ is fixed for large $x$ (namely for $x$ so large that $y > K+1$), and so it is enough to investigate the number

$$n' = n_0 p^{\lambda-1} + \cdots + n_{\lambda-1}.$$

Similarly, we write

$$m = m_0 p^\gamma + m_1 p^{\gamma-1} + \cdots + m_\gamma$$

for the base $p$ representation of $m = n + k$. Its last digit is $k - i - 1$ if $i \le k-1$ and $p-1$ if $i = k$, so it is enough to investigate the number

$$m' = m_0 p^{\gamma-1} + \cdots + m_{\gamma-1}.$$

We now let $\mathcal{M}_2^{(k)}(x)$ be that subset of $n \in \mathcal{M}^{(k)}(x) \backslash \mathcal{M}_1^{(k)}(x)$ for which

$$\begin{array}{ccccc} \text{either} & s & = & \#\{1 \le j \le \lambda - 1 : n_j > p/2\} & < & \lambda/4, \\ \text{or} & t & = & \#\{1 \le j \le \gamma - 1 : m_j > p/2\} & < & \gamma/4. \end{array}$$

The situation is entirely symmetric when dealing with the digits of $n$ in base $p$, and with the digits of $m$ in base $p$, so we deal only with the number $s$. Fix the positive integer $s < \lambda/4$. The indices $\{i_1, \ldots, i_s\} \subset \{1, \ldots, \lambda-1\}$ for which $n_j > p/2$ can be fixed in $\binom{\lambda-1}{s} < \binom{\lambda}{s}$ ways, and summing up the number of such choices over $s < \lambda/4$, we get that the total number of such choices is at most

$$(\lambda/4)\binom{\lambda}{\lfloor \lambda/4 \rfloor} \ll \lambda \frac{\lambda^{\lfloor \lambda/4 \rfloor}}{\lfloor \lambda/4 \rfloor!} \ll \lambda \left( \frac{\lambda}{(\lambda/(4e))} \right)^{\lfloor \lambda/4 \rfloor} \ll \lambda (4e)^{\lambda/4}. \tag{24}$$

Here, we used the inequality $\ell! > (\ell/e)^\ell$ valid for positive integers $\ell$, together with the fact that

$$\lfloor u \rfloor! \ge (\lfloor u \rfloor/e)^{\lfloor u \rfloor} = (u/e)^{\lfloor u \rfloor} (1 + O(1/u))^{\lfloor u \rfloor} \gg (u/e)^{\lfloor u \rfloor}$$

valid for all real numbers $u > 1$ (in (24) we took $u = \lambda/4$).

There are $(p-1)/2$ possible digits larger than $p/2$ and $(p+1)/2$ possible digits smaller than $p/2$. Once the subset $\{i_1, \ldots, i_s\}$ of indices in $\{1, 2, \ldots, \lambda-1\}$ has been chosen, the number of choices for the digits $\{n_1, \ldots, n_{\lambda-1}\}$ such that $n_{i_j} > p/2$ for $j = 1, \ldots, s$ and the remaining $\lambda - 1 - s$ digits are smaller than $p/2$ is therefore

$$\left( \frac{p-1}{2} \right)^s \left( \frac{p+1}{2} \right)^{\lambda-1-s} < \frac{p^{\lambda-1}}{2^{\lambda-1}} \left( 1 + \frac{1}{p} \right)^\lambda. \tag{25}$$

Since certainly $\lambda = O(\log x)$, whereas $p > y$, we get that

$$\left( 1 + \frac{1}{p} \right)^\lambda = \exp\left( O\left( \frac{\log x}{y} \right) \right) = O(1). \tag{26}$$

Thus, the number of possibilities for the number

$$n'' = n_1 p^{\lambda-2} + \cdots + n_{\lambda-1}$$

is, after multiplying bounds (24) and (25) and using estimate (26), of order at most

$$\lambda(4e)^{\lambda/4}\frac{p^{\lambda-1}}{2^{\lambda-1}} \ll \frac{\lambda p^{\lambda-1}}{(4/e)^{\lambda/4}}. \tag{27}$$

Say $n''$ is fixed in one of the above ways. Since $x/2 < n \le x$, we have that

$$\frac{x/2-(p-i-1)}{p} < n' = n_0p^{\lambda-1} + n'' \le \frac{x-(p-i-1)}{p},$$

so that

$$\frac{x/2-(p-i-1)}{p} - n'' < n_0p^{\lambda-1} \le \frac{x-(p-i-1)}{p} - n''.$$

The number of multiples of $p^{\lambda-1}$ in the interval

$$\mathcal{I} = \left(\frac{x/2-(p-i-1)}{p} - n'', \frac{x-(p-i-1)}{p} - n''\right]$$

is at most

$$\left[\frac{x}{2p^{\lambda}}\right] + 1 \le \frac{3x}{2p^{\lambda}}, \tag{28}$$

because $p^{\lambda} \le n \le x$. In the above, we also used the fact that the length of $\mathcal{I}$ is $x/(2p)$. Thus, the number of ways of choosing $n_0$ is of order at most $x/p^{\lambda}$. In conclusion, the number of choices for $n$ is, after multiplying bounds (27) and (28), of order at most

$$\frac{x}{p^{\lambda}} \times \frac{\lambda p^{\lambda-1}}{(4/e)^{\lambda/4}} \ll \frac{x\lambda}{p(4/e)^{\lambda/4}}.$$

Observe that since $x/2 < n \le x$, it follows that $x/(2p) < p^{\lambda} \le x$, so that $\lambda = \log x/\log p + O(1)$. Since $p \le z$, we get that $\lambda \ge c_3 \log\log x + O(1)$. This is a lower bound on $\lambda$, while certainly $\lambda \le \log x$ is an upper bound for $\lambda$. Thus, putting $c_4 = (c_3/4)\log(4/e)$, we get that the number of choices for such $n$ is of order at most

$$\frac{x\log x}{p(\log x)^{c_4}}.$$

The same inequality applies to the cardinality of the subset consisting of those $n \in \mathcal{M}^{(k)}(x)\backslash\mathcal{M}_1^{(k)}(x)$ for which $t < \gamma/4$. Further, all this was for a fixed $i$ (or $j$) in $\{1, 2, \ldots, k\}$. Summing up over all possible values of $i$ (or $j$), and then over all the possible primes $p \in \mathcal{J}$, we get that

$$\#\mathcal{M}_2^{(k)}(x) \ll \frac{x(\log x)K}{(\log x)^{c_4}} \sum_{p\in\mathcal{J}} \frac{1}{p} \ll \frac{x(\log\log x)^2}{(\log x)^{c_4-1}},$$

where we used the fact that the estimate

$$\sum_{p\le u} \frac{1}{p} = \log\log u + O(1)$$

holds for all $u \ge 10$ (see Theorem 427 in [10]). Similarly, as before, we put $\mathcal{M}_2(x) = \bigcup_{0<|k|\le K} \mathcal{M}_2^{(k)}(x)$ and get that

$$\#\mathcal{M}_2(x) \le \sum_{0<|k|\le K} \#\mathcal{M}_2^{(k)}(x) \ll \frac{x(\log\log x)^3}{(\log x)^{c_4-1}} = o\left(\frac{x}{(\log x)^3}\right), \tag{29}$$

as $x \to \infty$, provided that $c_4 \ge 5$; that is, $c_3 \ge 20/\log(4/e)$, which we are assuming. In fact, we take $c_3 = 52$. The bound (29) is acceptable for us.

3.6. **The case when $\Gamma_M\Gamma_N$ is divisible by some large prime: Set Up.** In this and the next section, we suppose that there exists a large prime $p$ dividing $\Gamma_M\Gamma_N$. We start by noticing that this is always the case when $p \mid MN$ for some prime $p \in \mathcal{J}$ and $n \notin \mathcal{M}_1^{(k)}(x) \cup \mathcal{M}_2^{(k)}(x)$. Let us justify this observation. For a prime $q$ and a positive integer $u$, we put $\nu_q(u)$ for the exponent of $q$ in the factorization of $u$. By Kummer's theory relating the number of digits of $n$ and $m$ in base $p$ which exceed $p/2$ with the exponent of $p$ in $\binom{2m}{m}$ and $\binom{2n}{n}$ (see [7]), we get, from the fact that $n \notin \mathcal{M}_2^{(k)}(x)$ that

$$\min\left\{\nu_p\left(\binom{2n}{n}\right), \nu_p\left(\binom{2m}{m}\right)\right\} \ge \lambda/4 > 13\log\log x + O(1).$$

Since $p^2$ does not divide either $n+1$ or $m+1 = n+k+1$ (this is because $n \notin \mathcal{M}_1^{(k)}(x)$), it follows that for large $x$ we have

$$\min\{\nu_p(C_n), \nu_p(C_m)\} \ge \lambda/4 > 13\log\log x + O(1) > 2. \tag{30}$$

Since $MN$ is not divisible by squares of primes in $\mathcal{J}$ (again because we have $n \notin \mathcal{M}_1^{(k)}$), it follows by inequality (30) and the $STMN$-equation (22), that $\Gamma_M$ and $\Gamma_N$ are divisible by all the prime factors of $M$ and $N$, respectively, which belong to $\mathcal{J}$. In particular, there exists a large prime dividing $\Gamma_M\Gamma_N$, which is what we wanted.

3.7. **The case when $\Gamma_M\Gamma_N$ is divisible by some large prime: Sieves.** Assume, for example, that $p^{\alpha_p}\|\Gamma_M$ for some large prime $p$, where $\alpha_p \ge 1$. From (22), we read that $p \mid \phi(MN)$. The same conclusion holds, namely that $p \mid \phi(MN)$, when $p^{\alpha_p}\|\Gamma_N$. Since $p^2 \nmid MN$ (because $n \notin \mathcal{M}_1^{(k)}(x)$), we conclude that there exist $i$ (or $j$) in $\{1, 2, \ldots, k\}$ such that $p \mid n+i+1$ (or $p \mid 2n+2j-1$), and also $i_1$ (or $j_1$) in the same set $\{1, 2, \ldots, k\}$ and a prime $q$ congruent to 1 modulo $p$ dividing $n+i_1+1$ (or $2n+2j_1-1$). In all cases, we get that $n$ is in a certain arithmetic progression modulo $p$ and in another arithmetic progression modulo $q$, so by the Chinese Remainder Lemma, $n$ is in a fixed arithmetic progression modulo $pq$. The number of such $n \le x$ is $\le x/(pq)+1$.

We consider first the case when $pq \le 10x$. Then

$$\frac{x}{pq} + 1 \le \frac{11x}{pq}.$$

Keeping $i$ (or $j$) fixed and $i_1$ (or $j_1$) fixed, and summing first over all the primes $q \le x + 2K - 1 \le 3x$ with $q \equiv 1 \pmod p$, then over all $p \in \mathcal{J}$, we get a bound of

$$11 \sum_{p \in \mathcal{J}} \frac{x}{p} \sum_{\substack{q \equiv 1 \pmod p \\ q \le 3x}} \frac{1}{q} \ll x \log\log x \sum_{p \in \mathcal{J}} \frac{1}{p^2} \ll \frac{x(\log\log x)^2}{y} \tag{31}$$

on the number of such possibilities $n$. In the above, we used the Brun-Titchmarsh bound

$$\sum_{\substack{q \equiv 1 \pmod d \\ q \le u}} \frac{1}{q} \ll \frac{\log\log u}{\phi(d)} \tag{32}$$

valid for all real numbers $u \ge 10$ and all positive integers $d$ (see Lemma 1 in [2]).

Summing up the above bound (31) over all pairs $i$ (or $j$) and $i_1$ (or $j_1$), we get a bound of order

$$\frac{x(\log\log x)^2 K^2}{y} \ll \frac{x(\log\log x)^4}{(\log x)^{10}} \tag{33}$$

on the number of such possibilities for $n$. To organize ideas, we write $\mathcal{M}_3^{(k)}(x)$ for the set of positive integers under scrutiny. Recall that this set is the set of $n \in \mathcal{M}^{(k)}(x) \backslash \left(\mathcal{M}_1^{(k)}(x) \bigcup \mathcal{M}_2^{(k)}(x)\right)$ for which there exists a large prime $p \mid MN$ and $q \mid MN$ such that $p \mid q-1$ and such that furthermore $pq \le 10x$. Bound (33) gives

$$\#\mathcal{M}_3^{(k)}(x) \ll \frac{x(\log\log x)^4}{(\log x)^{10}}.$$

Putting $\mathcal{M}_3(x) = \bigcup_{0<|k|\le K} \mathcal{M}_3^{(k)}(x)$, we get that

$$\#\mathcal{M}_3(x) \le \sum_{0<|k|\le K} \#\mathcal{M}_2^{(k)}(x) \ll \frac{x(\log\log x)^5}{(\log x)^{10}} = o\left(\frac{x}{(\log x)^3}\right) \tag{34}$$

as $x \to \infty$. The bound (34) is acceptable for us.

We now take a look at the case when $pq > 10x$.

**Case 1.** *The case when $p \mid N$.*

Say $2n + 2j - 1 \equiv 0 \pmod p$ for some $j \in \{1, 2, \ldots, k\}$. We now write $2n+2j-1 = pa$. Suppose first that $n+i_1+1 = qb$ for some $i_1 \in \{1, 2, \ldots, k\}$. Then $2qb = (2n + 2j - 1) + (2i_1 + 3 - 2j)$. Reducing the above equation modulo $p$, we get that $2b \equiv 2i_1 + 3 - 2j \pmod p$. However, observe that $b \le (n + K + 1)/q < 2x/q < p/4$ for large $x$ (because $pq > 10x$), so that

$$|2b - (2i_1 + 3 - 2j)| \le 2b + (2K + 1) < \frac{p}{2} + 2K + 1 < p$$

for large $x$, because $p > y > 4K + 2$. Since $2b - (2i_1 + 3 - 2j)$ is a multiple of $p$ smaller than $p$ in absolute value, it should be the number zero, but this is impossible because it is an odd number. A similar argument deals with the case when $2n+2j_1-1=qb$ for some $j_1\in\{1,2,\ldots,k\}$. In this case, $j_1\neq j$, for otherwise we would get that $p\mid b$, therefore $pq\le 2n+2K-1<3x$ for large $x$, which contradicts the fact that $pq>10x$. Further, we have $b\le (2n+2K-1)/q<3x/q<p/3$ for large $x$ (again, because $pq>10x$). Thus,

$$qb=(2n+2j-1)+2(j_1-j).$$

Reducing the above equation modulo $p$ we get $b\equiv 2(j_1-j)\pmod p$. However, the inequality

$$|b-2(j_1-j)|<\frac{p}{3}+2K<p$$

holds for large $x$, and since $b-2(j_1-j)$ is a multiple of $p$, it should be the number zero, which is again impossible since this number is in fact odd, because $b$ is odd. This takes care of the case when $p$ divides $N$.

**Case 2.** *The case when $p\mid M$.*

Assume that $p\mid n+i+1$ for some $i\in\{1,2,\ldots,k\}$. As in Case 1, we write $n+i+1=pa$. Suppose first that $q\mid n+i_1+1$ for some $i_1\in\{1,2,\ldots,k\}$ and write $n+i_1+1=qb$. If $i=i_1$, then $p\mid b$, so $pq\mid n+i+1$, so $pq\le x+K+1<2x$ for large $x$, contradicting the fact that $pq>10x$. Thus, $i\neq i_1$. Clearly, $b\le (n+K+1)/q<2x/q<p/4$ for large $x$. Then $qb=n+i+1+(i_1-i)$, and reducing the above relation modulo $p$ we get that $b\equiv i_1-i\pmod p$. However, for large $x$ we have

$$|b-(i_1-i)|\le b+K<\frac{p}{4}+K<p.$$

Thus, the number $b-(i_1-i)$ is zero, showing that $b=i_1-i$. In particular, $i_1>i$, and we get the equation

$$n+i_1+1=q(i_1-i). \tag{35}$$

Another possible case is when $q\mid 2n+2j_1-1$ for some $j_1\in\{1,2,\ldots,k\}$. In this case, $2n+2j_1-1=qb$, so, as before, $b\le(2n+2K-1)/q<3x/q<p/3$ for large $x$. Further, $qb=2n+2j_1-1=(2n+2i+2)+(2j_1-2i-3)$. Reducing the above relation modulo $p$, we get $b\equiv 2j_1-2i-3\pmod p$. Since the inequality

$$|b-(2j_1-2i-3)|\le b+2K+1<\frac{p}{3}+2K+1<p$$

holds for large $x$, we must have $b=2j_1-2i-3$. In particular, $j_1\ge i+2$ and

$$2n+2j_1-1=q(2j_1-2i-3). \tag{36}$$

So far, we learned that if there is a large prime factor $p$ of $\Gamma_M\Gamma_N$ and $n\in\mathcal{M}^{(k)}(x)\backslash(\mathcal{M}_1(x)\bigcup\mathcal{M}_2(x)\bigcup\mathcal{M}_3(x))$, then $p\mid M$, so $p\mid n+i+1$ for some $i\in\{1,2,\ldots,k\}$, and either relation (35) holds with some $i_1>i$ in $\{1,2,\ldots,k\}$ and some prime $q$, or relation (36) holds for some $j_1\ge i+2$ in

$\{1,2,\ldots,k\}$ and some prime $q$. In both cases, $k \geq 2$ (in fact, in the second case we must have $k \geq 3$). Consider the forms

$$2n+1 \qquad \text{and} \qquad 2n+3.$$

Rewriting them in terms of the prime $q$, they become

$$\begin{array}{lcccc} 2(i_1-i)q-(2i_1+1) & \text{and} & 2(i_1-i)q-(2i_1-1) & \text{if} & (35); \\ (2j_1-2i-3)q-2(j_1-1) & \text{and} & (2j_1-2i-3)q-2(j_1-2) & \text{if} & (36). \end{array} \tag{37}$$

Since $i_1 > i \geq 1$ (so $i_1 \geq 2$), and $j_1 \geq i+2$ (so $j_1 \geq 3$), in both cases, we obtain two non–proportional linear forms in the prime $q$. Also, none of the two forms is proportional to $q$ itself (since the constant coefficients are not zero). Observe that $2n+1$ and $2n+3$ are free of primes in $\mathcal{J}$, otherwise, we are already in the case when $n \in \mathcal{M}_3^{(k)}(x)$ by the deduction from Subsection 3.6, and Case 1 of Subsection 3.7. Hence, we have two non–proportional linear forms with nonzero coefficients in the prime $q$ which are free of primes from $\mathcal{J}$. It follows from the sieve (see Theorem 5.8 in [8]), that the number of possibilities for $q \leq 2x$ (hence, for $n \leq x$) is of order at most

$$\frac{x}{\log x}\prod_{p\in\mathcal{J}}\left(1-\frac{2}{p}\right)^2 \ll \frac{x}{\log x}\left(\frac{\log y}{\log z}\right)^2 \ll \frac{x(\log\log x)^4}{(\log x)^3}.$$

In the above application of the sieve, we implicitly used the fact that for large $x$ we have $y > 2K+1$, and, in particular, the two pairs of linear forms in $q$ shown at (37) are nonproportional modulo $p$ for all primes $p \in \mathcal{J}$. Of course, this was for fixed $i$ and $i_1$, or $i$ and $j_1$. Summing over all the possibilities for $i$ and $i_1$ or $j_1$, we get that if we put $\mathcal{M}_4^{(k)}(x)$ for the subset of $n$ under consideration, we get that

$$\#\mathcal{M}_4^{(k)}(x) \ll \frac{x(\log\log x)^4K^2}{(\log x)^3} \ll \frac{x(\log\log x)^6}{(\log x)^3}.$$

Putting $\mathcal{M}_4(x) = \bigcup_{0<|k|\leq K}\mathcal{M}_4^{(k)}(x)$, we get that

$$\#\mathcal{M}_4(x) \leq \sum_{0<|k|\leq K}\#\mathcal{M}_4^{(k)}(x) \ll \frac{x(\log\log x)^7}{(\log x)^3} = \frac{x}{(\log x)^{3+o(1)}}, \tag{38}$$

as $x \to \infty$. The bound (38) is acceptable for us. This completes the analysis of the case when $\Gamma_M\Gamma_N$ is a multiple of some large prime.

3.8. **The Case when $\Gamma_M\Gamma_N$ is free of large primes and $k \geq 2$.** By the results from Section 3.6, it follows that $MN$ is free of primes from $\mathcal{J}$. Assume that $k \geq 2$. Then each of

$$n+2, \quad n+3, \quad 2n+1, \quad 2n+3,$$

are free of primes $p \in \mathcal{J}$. These four linear forms in $n$ are non–proportional. Thus, by the sieve (see Theorem 5.7 in [8]), the number of such $n \leq x$ is of order at most

$$x\prod_{p\in\mathcal{J}}\left(1-\frac{4}{p}\right)\ll x\left(\frac{\log z}{\log y}\right)^4\ll\frac{x(\log\log x)^8}{(\log x)^4}.$$

Hence, putting $\mathcal{M}_5(x)$ for the set of such $n\le x$, we get that

$$\#\mathcal{M}_5(x)\ll\frac{x(\log\log x)^8}{(\log x)^4}=o\left(\frac{x}{(\log x)^3}\right),\tag{39}$$

as $x\to\infty$. The bound (39) is acceptable for us.

3.9. **The Case when $\Gamma_M\Gamma_N$ is free of large primes and $k=1$.** Going back to the results from Subsection 3.2, we see that $D=1,3$ and $P(UV)\le 3$. Now equation (22) tells us that the relation

$$S_1\phi(M)=T_1\phi(N)$$

holds with some positive integers $S_1$ and $T_1$ with $P(S_1T_1)\le y$. Replacing $M$ and $N$ by $(n+2)/UD$ and $(2n+1)/VD$, respectively (see relation (19)), we get that the relation

$$S_2\phi(n+2)=T_2\phi(2n+1)\tag{40}$$

holds with some positive integers $S_2$ and $T_2$ with $P(S_2T_2)\le y$. We also have the additional information that $(n+2)(2n+1)$ is free of primes from $\mathcal{J}$.

3.10. **The structure of solutions to equation** (40)**.** To handle such positive integers $n$, we recall that if we put

$$\Psi(t,w)=\{n\le t: P(n)\le w\},\tag{41}$$

then the inequality

$$\#\Psi(t,w)\ll\frac{t}{\exp(u/2)}\qquad\text{where}\qquad u=\frac{\log t}{\log w}$$

holds for all $2\le w\le t$ (see [22, Theorem 1, p. 359]). Better estimates for $\#\Psi(t,w)$ are known (see [4]), but we shall not need them. Let $\mathcal{M}_6(x)$ be the set of $n\in\mathcal{M}(x)\backslash\bigcup_{i=1}^5\mathcal{M}_i(x)$ such that $P(n)\le y$ or $P(2n+1)\le y$. By estimate (41), it follows that

$$\begin{aligned}\#\mathcal{M}_6(x)&\le 2\#\Psi(x,y)\ll\frac{x}{\exp\left((\log x)/(2\log y)\right)}\\&=\frac{x}{\exp\left((\log x)/(20\log\log x)\right)}=o\left(\frac{x}{(\log x)^3}\right),\end{aligned}\tag{42}$$

as $x\to\infty$. This is acceptable for us. From now on, we work with the remaining numbers $n$ in $\mathcal{M}(x)$.

We next write

$$n+2=p_1\cdots p_r a\qquad\text{and}\qquad 2n+1=q_1\cdots q_s b,\tag{43}$$

where $r\ge 1,\ s\ge 1,\ z<p_1<\cdots<p_r,\ z<q_1<\cdots<q_s$ are primes and $P(ab)\le y$. Such representations for $n$ and $2n+1$ exist because $n\notin\mathcal{M}_6(x)$, so there exist prime factors of both $n+2$ and $2n+1$ exceeding $y$; hence,

exceeding $z$ because $n+2$ and $2n+1$ are coprime with the primes from $\mathcal{J}$. Further, $n+2$ and $2n+1$ are not divisible by squares of large primes because $n \notin \mathcal{M}_1(x)$. Put $L = \log y = 10 \log\log x$. We let $\mathcal{M}_7(x)$ be the set of $n$ such that either $a \ge \exp(L^2)$, or $b \ge \exp(L^2)$. Fixing $a$, the number of such $n \le x$ is at most $(x+2)/a \le 2x/a$, while fixing $b$, the number of such $n$ is at most $(2x+1)/b \le 3x/b$. Thus,

$$\#\mathcal{M}_7(x) \le \sum_{\substack{a \ge \exp(L^2) \\ P(a) \le y}} \frac{2x}{a} + \sum_{\substack{b \ge \exp(L^2) \\ P(b) \le y}} \frac{3x}{b} \le 5x \sum_{\substack{a \ge \exp(L^2) \\ P(a) < y}} \frac{1}{a} := 5x\mathcal{S}. \tag{44}$$

Putting

$$\mathcal{A}(t) = \{\exp(L^2) \le a < t : P(a) \le y\},$$

we get, by estimate (41) and the fact that for $t \ge \exp(L^2)$ we have

$$\frac{\log t}{\log y} \ge \frac{L^2}{\log y} = 10 \log\log x,$$

that

$$\#\mathcal{A}(t) \le \#\Psi(t,y) \ll \frac{t}{\exp\left((\log t)/(2\log y)\right)} \le \frac{t}{\exp(5\log\log x)} = \frac{t}{(\log x)^5}. \tag{45}$$

By Abel's summation formula and estimate (45) together with the observation that $\max\{a,b\} \le 3x$, we get

$$\begin{aligned} \mathcal{S} = \sum_{a \in \mathcal{A}(3x)} \frac{1}{a} &= \int_L^{3x} \frac{d\#\mathcal{A}(t)}{t} \le \left( \frac{\#\mathcal{A}(t)}{t}\Big|_{t=L}^{t=3x} \right) + \int_L^{3x} \frac{\#\mathcal{A}(t)dt}{t^2} \\ &\ll \frac{\#\mathcal{A}(3x)}{x} + \frac{1}{(\log x)^5} \int_L^{3x} \frac{dt}{t} \ll \frac{1}{(\log x)^4}. \end{aligned} \tag{46}$$

Inserting estimate (46) into estimate (44), we get

$$\#\mathcal{M}_7(x) \ll \frac{x}{(\log x)^4}. \tag{47}$$

This is acceptable for us. From now on, assume that $\max\{a,b\} < \exp(L^2)$. Observe next that since $p_1 > y$ and $q_1 > y$, and for large $x$, we have that $z^{10L} = x^{100/52} > 2x+1$, it follows that $\max\{r,s\} < 10L$. Now write

$$p_i - 1 = A_i a_i, \text{ where } \quad p(A_i) > y \quad \text{and} \quad P(a_i) \le y, \quad \text{for all} \quad i = 1, \ldots, r. \tag{48}$$

Similarly, we write

$$q_j - 1 = B_j b_j, \text{ where } \quad p(B_j) > y \quad \text{and} \quad P(b_j) \le y, \quad \text{for all} \quad j = 1, \ldots, s. \tag{49}$$

We let $\mathcal{M}_8(x)$ to be the subset of $n \in \mathcal{M}(x) \setminus \bigcup_{\ell=1}^7 \mathcal{M}_\ell(x)$ such that either the inequality $a_i \ge \exp(L^2)$ holds for some $i = 1, \ldots, r$, or the inequality $b_j \ge \exp(L^2)$ holds for some $j = 1, \ldots, s$. Assume that $a_i \ge \exp(L^2)$ for some $i = 1, \ldots, r$. Then there exists a prime $p$ dividing $n+2$ and a divisor $a$ of $p-1$ with $a \ge \exp(L^2)$ and $P(a) \le y$. Fixing such a positive integer

$a$ and then the prime $p \equiv 1 \pmod a$, the number of such $n \le x$ is at most $(x+2)/p \le 2x/p$. Summing first over all $p \equiv 1 \pmod a$, then over all the suitable values for $a \le 2x$, we get that the contribution of such $n \le x$ is of order at most

$$\begin{aligned}\sum_{\substack{a \ge \exp(L^2)\\ P(a)\le y}} \sum_{\substack{p\equiv 1 \pmod a\\ p\le 2x}} \frac{x}{p} &\ll x\log\log x \sum_{\substack{a\ge\exp(L^2)\\ P(a)\le y}} \frac{1}{\phi(a)}\\ &\ll x(\log\log x)^2 \sum_{\substack{a\ge\exp(L^2)\\ P(a)\le y}} \frac{1}{a} \qquad (50)\\ &= x(\log\log x)^2\mathcal{S} \ll \frac{x(\log\log x)^2}{(\log x)^4}.\end{aligned}$$

In the above inequalities we used a variety of inequalities such as the Brun-Titchmarsh inequality (32) to estimate the inner sum over the reciprocals of the primes $p \le x$ congruent to 1 modulo $a$, the minimal order of the Euler function (13) to deduce that $1/\phi(a) \ll (\log\log x)/a$ for all $a \le 2x$, as well as the estimate (46). The case when $b_j \ge \exp(L^2)$ holds for some $j = 1, \ldots, s$ is analogous. Namely, in this case we get that there exists $b \ge \exp(L^2)$ with $P(b) \le y$ dividing $q-1$ for some prime factor $q$ of $2n+1$. Fixing $b$ and $q$, the number of such $n \le x$ is at most $(2x+1)/q \le 3x/q$. Summing up the above bound over all $q \equiv 1 \pmod b$ with $q \le 3x$ and then over all $b \le 3x$ with $b \ge \exp(L^2)$ and $P(b) \le y$, we get an estimate of the same order as (50). Hence,

$$\#\mathcal{M}_8(x) \ll \frac{x(\log\log x)^2}{(\log x)^4}. \qquad (51)$$

This is acceptable for us.

We now work with the remaining set of $n \in \mathcal{M}(x)$. Equation (40) implies that

$$\prod_{i=1}^{r} A_i = \prod_{j=1}^{s} B_j. \qquad (52)$$

Put $\mathcal{M}_9(x)$ for the set of such $n$ with $\min\{r,s\} \ge 2$. Let $n$ be such a number. We certainly know that the primes $p_1, \ldots, p_r, q_1, \ldots, q_s$ are all distinct. Assume that $p_r > q_s$, since the remaining case can be handled similarly. Then equation (52) shows that there exists $j \in \{1, 2, \ldots, s\}$ such that

$$D_j = \gcd(A_1, B_j) > A_1^{1/s} = \left(\frac{p_1-1}{a_1}\right)^{1/s} > \left(\frac{z-1}{\exp(L^2)}\right)^{1/(10L)} > x^{1/L^3},$$

where the last inequality holds for all sufficiently large values of $x$. Observe that $D_j \mid p_1 - 1$ and $D_j \mid q_j - 1$. Further, the congruences $n+2 \equiv 0 \pmod{p_1}$ and $2n+2 \equiv 0 \pmod{q_j}$ put $n \le x$ into a certain arithmetic progression modulo $p_1 q_j$ by the Chinese Remainder Lemma. Since $r \ge 2$ and $p_r > q_s$, we have that $p_1 q_j \le p_1 q_s < p_1 p_r \le n+2 \le 2x$ (in case $q_s > p_r$, the last member of the corresponding inequality is $q_1 q_s \le 2n+1 \le 3x$, which is good enough for the purposes of the subsequent argument). The number of $n \le x$ in the above arithmetic progression modulo $p_1 q_j$ is of order $x/(p_1 q_j)$. We now vary $p_1$ and $q_j$ through the set of all primes not exceeding $3x$ and which are congruent to 1 modulo $d$, while keeping $d = D_j$ fixed, and then over all $d > x^{1/L^3}$, getting a contribution of order

$$
\begin{aligned}
\sum_{d > x^{1/L^3}} \sum_{\substack{p_1 < q_j \le 3x \\ p_1 \equiv q_j \equiv 1 \pmod d}} \frac{2x}{p_1 q_j} & \ll x \sum_{d > x^{1/L^3}} \left( \sum_{\substack{p \le 3x \\ p \equiv 1 \pmod d}} \frac{1}{p} \right)^2 \\
& \ll x(\log\log x)^2 \sum_{d > x^{1/L^3}} \frac{1}{\phi(d)^2} \\
& \ll x(\log\log x)^4 \sum_{d > x^{1/L^3}} \frac{1}{d^2} \\
& \ll \frac{x(\log\log x)^2}{x^{1/L^3}}.
\end{aligned} \tag{53}
$$

A similar argument applies to the case when $q_s > p_r$, and the number of such $n$ is of the same order as shown in (53) above. We thus get that

$$
\#\mathcal{M}_9(x) \ll \frac{x(\log\log x)^4}{x^{1/L^3}} = o\left( \frac{x}{(\log x)^3} \right), \tag{54}
$$

as $x \to \infty$. This is acceptable for us.

So, from now on, we assume that $r = 1$ or $s = 1$. We show that $r = 1$ implies $s = 1$. The reciprocal is also true and the details are similar. For $r = 1$, we get that $A_1 = B_1 \cdots B_s$. Using this relation together with equations (48) and (49) into (43), we get

$$
2(A_1 a_1 + 1)a - 3 = 2(n+2) - 3 = 2n+1 = (B_1 b_1 + 1) \cdots (B_s b_s + 1) b, \tag{55}
$$

leading to

$$
|B_1 B_2 \cdots B_s (2a_1 a - b_1 b_2 \cdots b_s b)| \le 2a + 3 + \frac{2^s B_1 B_2 \cdots B_s b_1 \cdots b_s b}{\min\{B_i : 1 \le i \le s\}}.
$$

Dividing across by $B_1 \cdots B_s$ and using the bound $s < 10L$ together with the bound $\max\{a, b, b_1, \ldots, b_s\} < \exp(L^2)$, and assuming that $j \in \{1, 2, \ldots, s\}$ is such that $B_j$ is minimal, we get that

$$
\begin{aligned}
|2a_1 a - b_1 b_2 \cdots b_s b| & < \frac{2a + 3 + 2^s b_1 \cdots b_s b}{B_j} \le \frac{5a 2^s b_1 \cdots b_s b}{(p_j - 1)/b_j} \\
& < \frac{10 \times 2^s \exp\left(L^2(s+3)\right)}{p_j} \\
& < \frac{10 \exp\left(L^2(10L+3) + 10L\right)}{z} < 1,
\end{aligned}
$$

where the last inequality holds provided that $x$ is sufficiently large. Thus, for large $x$ we have $2aa_1 = b_1 \cdots b_s b$. Using this information in equation (55), we get

$$2a - 3 = (B_1 b_1 + 1) \cdots (B_s b_s + 1) b - B_1 \cdots B_s b_1 \cdots b_s b. \tag{56}$$

The right–hand side above is positive so $a \geq 2$. If $s \geq 2$, in the right–hand side above we have a sum of $2^s - 1$ terms one of them being

$$B_1 \cdots B_{s-1} b_1 b_2 \cdots b_{s-1} b \geq B_1 = \frac{p_1 - 1}{b_1} > \frac{p_1}{2b_1}.$$

Comparing this with (56), we get that

$$z < p_1 < 4ab_1 < 4\exp(2L^2),$$

which is false for large $x$. Hence, $s = 1$. As we have already said, a similar argument shows that $s = 1$ implies that $r = 1$.

So, from now on, we have $r = s = 1$, so $A_1 = B_1$, and $2aa_1 = b_1 b$. Then equation (43) is

$$2(a_1 A_1 + 1)a - 3 = (B_1 b_1 + 1)b,$$

which together with the fact that $2aa_1 A_1 = B_1 b_1$ yields $2a - 3 = b$.

3.11. **Bounding $\#\mathcal{N}_r(x)$.** Since $n + 2 = pa$ and $2n + 1 = q(2a - 3)$, we get that

$$(2a)p - (2a - 3)q = 3.$$

Clearly, the greatest common divisor between $a$ and $2a-3$ is 1 or 3 according to whether $a$ is coprime to 3 or not. Further, the smallest positive integer solution $(u, v)$ of equation $2au - (2a - 3)v = 3$ is $(u, v) = (1, 1)$. Hence, we get that

$$\begin{cases} p &= 1 + (2a - 3)\lambda; \\ q &= 1 + 2a\lambda, \end{cases} \quad \text{if} \quad 3 \nmid a; \quad \begin{cases} p &= 1 + (2a/3 - 1)\lambda; \\ q &= 1 + (2a/3)\lambda, \end{cases} \quad \text{if} \quad 3 \mid a.$$

Since $ap = n + 2 \leq 2x$, it follows that $p \leq 2x/a$, so that $\lambda \leq 2x/(a(2a - 3))$ if $3 \nmid a$, whereas $\lambda \leq 6x/(a(2a - 3))$ if $3 \mid a$. In both cases, we have that $\lambda \leq 12x/a^2$, because $2a - 3 \geq a/2$ for all $a \geq 2$. Thus, fixing $a < \exp(L^2)$, we have that $\lambda \leq 12x/a^2$. Further, $\lambda$ has the property that a pair of non-proportional linear forms in $\lambda$ are both primes. By the sieve (see Theorem 5.7 in [8]), the number of such $\lambda \leq 12x/a^2$ is of order at most

$$\frac{x}{a^2 \log(x/a^2)} \left( \frac{E}{\phi(E)} \right)^2, \tag{57}$$

where we can take $E = 2a(2a - 3)$. Since $E < 4a^2$, by inequality (13), we deduce that the estimate $E/\phi(E) \ll \log\log(a + 1)$ holds for all $a \geq 2$. Since $a^2 < \exp(2L^2) < x^{1/2}$ for large $x$, we get that the expression (57) is of order

$$\frac{x}{(\log x)^2} \left( \frac{(\log\log(a + 1))^2}{a^2} \right).$$

Summing up the last bound above for all possible values of $a$, it follows that the remaining set of $n \le x$, call it $\mathcal{M}_{10}(x)$, has cardinality satisfying the inequality

$$\#\mathcal{M}_{10}(x) \ll \frac{x}{(\log x)^2} \sum_{2\le a<\exp(L^2)} \frac{(\log\log(a+1))^2}{a^2} \ll \frac{x}{(\log x)^2}. \tag{58}$$

3.12. **The values of $r$.** To conclude the proof of Theorem 2.1 it suffices to show that the numbers $n \in \mathcal{M}_{10}(x)$ lead to a solution $(m,n) = (n+1,n)$ of equation $\phi(C_m)/\phi(C_n) = r$ with $r = 4$. We go back through the argument from Section 3.2 keeping track of all the parameters. We have

$$C_m = C_{n+1} = \frac{2(2n+1)}{n+2} C_n = \frac{2bq}{ap} C_n,$$

so that

$$apC_{n+1} = 2bqC_n.$$

Let $d = \gcd(a, 2b) = \gcd(a, 4a-6b)$. Clearly, $d \in \{1,2,3,6\}$. Put $a_1 = a/d$ and $b_1 = 2b/d$. We then have

$$a_1pC_{n+1} = b_1qC_n.$$

Thus, there exists some positive integer $C$ such that

$$C_{n+1} = b_1qC \qquad \text{and} \qquad C_n = a_1pC. \tag{59}$$

Observe that $n = ap-2 = (a-1)p+(p-2)$, and $a-1 < \exp(L^2) < z/2 < p/2$ for large $x$, so that $\nu_p\left(\binom{2n}{n}\right) = 1$. Since $p \mid n+2$, we have that $p \nmid n+1$, so we get that $p \| C_n$. This shows that $p \nmid C$. Further, $n = (b-1)/2q + (q-1)/2$, and $(b-1)/2 < \exp(L^2) < z/2 < q/2$ for large $x$, therefore $\nu_q\left(\binom{2n}{n}\right) = 0$. Hence, $q \nmid C_n$, showing that $q \nmid C$.

We now study the exponents of the small primes in $\binom{2n}{n}$. Let $\rho \le y$ be any small prime. Write, as in Section 3.5,

$$n = n_0\rho^{\lambda} + n_1\rho^{\lambda-1} + \cdots + n_\rho$$

for the base $\rho$ representation of $n$. Let us count the number of $n \le x$ such that for some $\rho \le y$, we have that

$$s = \#\{1 \le i \le \lambda : n_i < \rho/2\} < L^3.$$

There are $\lambda+1$ possible locations $\{0,1,\ldots,\lambda\}$. There are

$$\sum_{s\le L^3} \binom{\lambda+1}{s} < L^3(\lambda+1)^{L^3} < \exp(L^4) \qquad \text{for large } x \tag{60}$$

possibilities of choosing subsets $\{i_1,\ldots,i_s\}$ with at most $L^3$ elements, where the digits smaller than $\rho/2$ are located. Once these positions are chosen, the number of possibilities of actually assigning digits $n_j < \rho/2$ whenever

$j \in \{i_1, \ldots, i_s\}$, and $n_j > \rho/2$ whenever $j \in \{0, 1, \ldots, \lambda\} \backslash \{i_1, \ldots, i_s\}$, is at most

$$\left(\frac{\rho}{2}\right)^s \left(\frac{\rho+1}{2}\right)^{\lambda+1-s} \ll \rho \frac{\rho^\lambda}{2^\lambda} \left(1 + \frac{1}{\rho}\right)^{\lambda+1} \ll \rho \frac{x}{1.3^\lambda}, \tag{61}$$

and this estimate is uniform in $2 \le \rho \le y$. Note that

$$\lambda = \frac{\log x}{\log \rho} + O(1) \ge \frac{\log x}{\log y} + O(1) = \frac{\log x}{10 \log\log x} + O(1) > \frac{\log x}{4L \log(1.3)} \tag{62}$$

for large $x$ uniformly in $\rho \le y$.

Multiplying bounds (60) and (61) and using the lower bound (62) on $\lambda$, we get that the number of such $n \le x$ is, for a given $\rho$, at most

$$\frac{xy \exp(L^4)}{1.3^{\log x/(4L \log(1.3))}} = \frac{xy \exp(L^4)}{x^{1/(4L)}}.$$

Summing the above bound over all primes $\rho \le y$, we get a bound of

$$\frac{xy^2 \exp(L^4)}{x^{1/(4L)}}.$$

Putting $\mathcal{M}_{11}(x)$ for the set of such $n \le x$, we get

$$\#\mathcal{M}_{11}(x) \ll \frac{xy^2 \exp(L^4)}{x^{1/(4L)}} = o\left(\frac{x}{(\log x)^3}\right), \tag{63}$$

as $x \to \infty$. This is acceptable for us.

From now on, we work with the remaining numbers $n$. For them, the inequality $\nu_\rho(\binom{2n}{n}) \ge L^3$ holds for all small primes $\rho$. If $\rho$ divides $n+2$, then $\rho$ divides $a$. Since $a < \exp(L^2)$, it follows that $\nu_\rho(a) \le (L^2)/(\log 2) < 2L^2$. Hence, $\nu_\rho(C_n) > L^3 - 2L^2$. Otherwise, that is if $\rho$ does not divide $n+2$, then $\nu_\rho(C_n) = \nu_\rho(\binom{2n}{n}) > L^3$. Hence, at any rate, if $\rho$ divides $a_1 b_1$, then

$$\nu_\rho(C_n) > L^3 - 2L^2 > 2L^2 > \nu_\rho(a_1 b_1).$$

Going back to equation (59), it follows that all primes dividing $a_1 b_1$ divide in fact $C$. Thus, write

$$C = \Gamma_{a_1} \Gamma_{b_1} \Gamma,$$

where

$$\Gamma_{a_1} = \prod_{\substack{\rho^{\alpha_\rho} \| C \\ \rho | a_1}} \rho^{\alpha_\rho} \qquad \text{and} \qquad \Gamma_{b_1} = \prod_{\substack{\rho^{\alpha_\rho} \| C \\ \rho | b_1}} \rho^{\alpha_\rho}.$$

Equations (59) become

$$C_{n+1} = b_1 q \, \Gamma_{a_1} \Gamma_{b_1} \Gamma \qquad \text{and} \qquad C_n = a_1 p \, \Gamma_{a_1} \Gamma_{b_1} \Gamma.$$

Taking the Euler functions, we get

$$\phi(C_{n+1}) = \phi(b_1)(q-1)\phi(\Gamma_{a_1})\Gamma_{b_1}\phi(\Gamma); \quad \phi(C_n) = \phi(a_1)(p-1)\Gamma_{a_1}\phi(\Gamma_{b_1})\phi(\Gamma).$$

Put $\operatorname{rad}(k) = \prod_{p|k} p$ for the *radical* of $k$. Using

$$\phi(a_1) = a_1 \frac{\phi(\operatorname{rad}(a_1))}{\operatorname{rad}(a_1)} \quad \text{and} \quad \phi(\Gamma_{a_1}) = \Gamma_{a_1} \frac{\phi(\operatorname{rad}(a_1))}{\operatorname{rad}(a_1)},$$

the similar relations with $a_1$ replaced by $b_1$, and the fact that

$$b(q-1) = 2a(p-1), \quad \text{therefore} \quad \frac{q-1}{p-1} = \frac{2a}{b} = \frac{4(a/d)}{2b/d} = \frac{4a_1}{b_1},$$

we get that

$$r = \frac{\phi(C_{n+1})}{\phi(C_n)} = \frac{b_1 \frac{\phi(\mathrm{rad}(b_1))}{\mathrm{rad}(b_1)}}{a_1 \frac{\phi(\mathrm{rad}(a_1))}{\mathrm{rad}(a_1)}} \times \frac{4a_1}{b_1} \times \frac{\Gamma_{a_1} \frac{\phi(\mathrm{rad}(a_1))}{\mathrm{rad}(a_1)}}{\Gamma_{a_1}} \times \frac{\Gamma_{b_1}}{\Gamma_{b_1} \frac{\phi(\mathrm{rad}(b_1))}{\mathrm{rad}(b_1)}} = 4.$$

The case $r = 1/4$ comes from the case when $m < n$ (in particular, when $m = n-1$). We now conclude that the conclusions of Theorem 2.1 hold. Indeed, estimates (23), (29), (34), (38), (39), (42), (47), (51) and (54) show that the cardinalities of $\mathcal{M}_j(x)$ for $j = 1, \ldots, 9$ are bounded as shown in (10), while if $n$ has made it to $\mathcal{M}_{10}(x)$, whose cardinality is bounded as in (58), but $r \neq 4,\ 1/4$, then it must be the case that in fact $n \in \mathcal{M}_{11}(x)$, a set whose cardinality is bounded, from inequality (63), by the right–hand side of (10). Thus, $r = 4,\ 1/4$ remain the only options for $n \in \mathcal{M}_{10}(x) \backslash \mathcal{M}_{11}(x)$, and this is bounded as shown in (11) by estimate (58).

## 4. Open questions

Numerically, it seems that $\{\phi(C_n)\}_{n \geq 2}$ is an increasing sequence. We leave this as a research problem for the reader. It would be interesting to study the Carmichael $\lambda$-function of the Catalan numbers. We conjecture that for all $k \geq 1$, there are infinitely many positive integers $n$ such that

$$\lambda(C_{n+1}) = \lambda(C_{n+2}) = \cdots = \lambda(C_{n+k}).$$

*Acknowledgement.* We thank the referee and the editor for helpful remarks. P. S. worked on this project while he visited the Mathematical Center of the UNAM in Morelia, Mexico in September, 2011. He thanks the members of that institution for their hospitality. F. L. was supported in part by project SEP-CONACyT 79685.

## 5. Appendix

We display in Figure 1 a graph showing the number of solutions of the equation $4 = \phi(C_{n+1})/\phi(C_n)$ in the range $n \in [0, 3000]$.

## References

[1] R. Alter and K. Kubota, "Prime and prime power divisibility of Catalan numbers", *J. Combin. Theory Ser. A* **15** (1973), 243–256.
[2] N. L. Bassily, I. Kátai and M. Wijsmuller, "On the prime power divisors of the iterates of the Euler-$\phi$ function", *Publ. Math. Debrecen* **55** (1999) 17–32.
[3] K. A. Broughan, M. J. González, R. H. Luis, F. Luca, V. J. Mejía Huguet and A. Togbé, "There are no multiply–perfect Fibonacci numbers", *INTEGERS* **11A** (2011), A7.
[4] R. Canfield, P. Erdős and C. Pomerance, "On a problem of Oppenheim concerning "Factorisatio Numerorum"", *J. Number Theory* **17** (1983), 1–28.

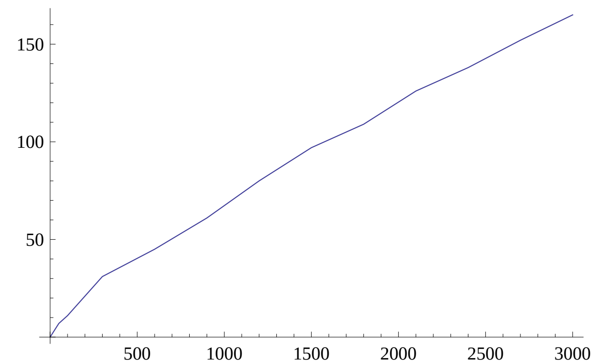


FIGURE 1. Number of solutions for $\phi(C_{n+1}) = 4\phi(C_n), n \leq 3000$


[5] R. D. Carmichael, "Note on Euler's $\phi$ function", *Bull. Amer. Math. Soc.* **8** (1922), 109–110.

[6] R. L. Graham, D. E. Knuth and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science*, 2nd ed. Reading, MA, Addison-Wesley, 1994.

[7] A. Granville, "Arithmetic Properties of Binomial Coefficients. I. Binomial Coefficients Modulo Prime Powers", in *Organic Mathematics. Proceedings of the Workshop Held in Burnaby, BC, December 12-14, 1995* (Ed. J. Borwein, P. Borwein, L. Jorgenson and R. Corless). Providence, RI, Amer. Math. Soc., pp. 253–276, 1997.

[8] H. Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press, London, 1974.

[9] G. H. Hardy and J. E. Littlewood, "Some problems of Partition Numerorum, III: On the expression of a number as a sum of primes", *Acta Math.* **44** (1923), 1–70.

[10] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford University Press, 1979.

[11] S. V. Konyagin, F. Luca and P. Stănică, "Sum of divisors of Fibonacci numbers", *Unif. Distrib. Theory* **4** (2009), 1–8.

[12] F. Luca, "Pascal's triangle and constructible polygons", *Util. Math.* **58** (2000), 209–214.

[13] F. Luca,"Perfect Fibonacci and Lucas numbers", *Rend. Circ. Mat. Palermo (2)* **49** (2000), 313–318.

[14] F. Luca, "Multiply perfect numbers in Lucas sequences with odd parameters", *Publ. Math. Debrecen* **58** (2001), 121–155.

[15] F. Luca, "On the sum of divisors of binomial coefficients", *Int. Math. Res. Not. IMRN* **2007**, Art. ID rnm101.

[16] F. Luca, "Fibonacci numbers with the Lehmer property", *Bull. Pol. Acad. Sci. Math.* **55** (2007), 7–15.

[17] F. Luca and J. L. Varona, "Multiperfect numbers on lines of the Pascal triangle", *J. Number Theory* **129** (2009), 1136–1148.

[18] F. Luca and B. de Weger, "$\sigma_k(F_m) = F_n$", *New Zealand J. Math.* **40** (2010), 1–13.

[19] F. Luca and P. Pollack, "Multiperfect numbers with identical digits", *J. Number Theory* **131** (2011), 260–284.

[20] F. Luca and H. te Riele, "Aliquot cycles of repdigits", *INTEGERS* **11** (2011), A53.

[21] P. Pollack, "Perfect numbers with identical digits", *INTEGERS* **11A** (2011), A18.

[22] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge U. Press, 1995.

Florian Luca
Centro de Ciencias Matemáticas
Universidad Nacional Autónoma de México
Ap. Postal 61–3 (Xangari)
C.P. 58089, Morelia, Michoacán, México
fluca@matmor.unam.mx

Pantelimon Stănică
Naval Postgraduate School
Applied Mathematics Department
Monterey, CA 93943, USA
pstanica@nps.edu