Novi Sad J. Math.
Vol. 33, No. 2, 2003, 49–66

POSITIVE INTEGERS n SUCH THAT $n|a^{\sigma(n)} - 1$

Florian Luca[1]

Abstract. For a positive integer n let $\sigma(n)$ be the sum of divisors function of n. In this note, we fix a positive integer a and we investigate the positive integers n such that $n|a^{\sigma(n)} - 1$. We also show that under a plausible hypothesis related to the distribution of prime numbers there exist infinitely many positive integers n such that $n|a^{\sigma(n)} - 1$ holds for all integers a coprime to n.

AMS Mathematics Subject Classification (2000): 11A07, 11N25

Key words and phrases: pseudoprimes

1. Introduction

Throughout this paper, n is a positive integer and $\phi(n)$, $\sigma(n)$, $\tau(n)$, $\Omega(n)$, $\omega(n)$ denote the classical arithmetical functions of n, namely the Euler function, the sum of divisors function, the number of divisors function, and the number of prime divisors function (counted with or without multiplicity), respectively. Euler's Theorem asserts that the divisibility relation $n|a^{\phi(n)} - 1$ holds whenever $a > 1$ is an integer which is coprime to n. When n is prime, this reduces to Fermat's Little Theorem, namely that $n|a^{n-1} - 1$. When $a > 1$ is fixed and n satisfies the above divisibility relation without being a prime, the number n is called a *base a pseudoprime*. When n is a composite number which is a pseudoprime with respect to all bases $a > 1$ and coprime to n, then n is called a *Carmichael number*. The most celebrated result on Carmichael numbers is Theorem 1 from [1], which shows that for large values of the positive real number x there are more than $x^{2/7}$ Carmichael numbers $n < x$ (see also [6] for some related results). Several interesting results about pseudoprimes and Carmichael numbers can also be found in [11].

In this paper, we address a question raised in [11] (see also [12]), namely whether or not there exist infinitely many positive integers n so that $n|a^{\sigma(n)} - 1$ holds for all positive integers $a > 1$ which are coprime to n.

In order to formulate our results, we introduce some more notations. For any positive integer k and any positive real x we set $\log_k x := \max\{\log \log_{k-1} x, 1\}$, where log stands for the natural logarithm function. When $k = 1$, we simply write $\log_1 x = \log x$ and we understand that this number is always ≥ 1.

[1]IMATE, UNAM, Ap. Postal 61–3 (Xangari), CP 58 089, Morelia, Michoacán, Mexico, e–mail: fluca@matmor.unam.mx

We start by fixing the number $a > 1$. Write R_a for the set of all $n \geq 1$ so that $n | a^{\sigma(n)} - 1$. For any positive real x we write $R_a(x) := \{1 \leq n < x \,|\, n \in R_a\}$. Our first theorem gives a lower bound for $\#R_a(x)$.

Theorem 1. *There exists a constant $c_1 := c_1(a)$ depending on a so that for large values of x holds the inequality*

$$\#R_a(x) > \exp(c_1 \log_2 x \log_3 x). \tag{1}$$

Theorem 1 shows that the number of numbers $n < x$ which are members of R_a as a function of x exceeds any fixed power of the logarithm of x once x is sufficiently large. In particular, the series

$$\sum_{n \in R_a} \frac{1}{\log n} \tag{2}$$

is divergent for all $a > 1$.

Our next theorem gives an upper bound on $\#R_a(x)$.

Theorem 2. *There exists an absolute constant c_2 so that for large x the estimate*

$$\#R_a(x) < x \, \exp(-c_2 (\log x \log_2 x)^{1/2}) \tag{3}$$

holds uniformly in $a \leq 2 \log x$.

Let R be the set of all $n \geq 1$ so that $n \in R_a$ holds for all positive integers $a > 1$ which are coprime to n. Thus, Rotkiewicz's question asks if R is an infinite set. Since for large n there exists a prime number $p < 2 \log n$ which does not divide n, Theorem 2 immediately implies that the sum $\sum_{n \in R} \frac{1}{n}$ is finite. In particular, if R is an infinite set, then the above series is convergent. While we have not been able to prove unconditionally that R is an infinite set, we show that this is indeed so if we assume a certain conjecture on the distribution of primes in arithmetical progressions. For every positive coprime integers $1 \leq a < d$ and any positive real number x let $\pi(x; a, d)$ denote the number of primes $p < x$ so that $p \equiv a (\mathrm{mod}\, d)$, let $\pi(x)$ denote the total number of primes $p < x$, and let $R(x) := \{1 \leq n < x \,|\, n \in R\}$. Our result is:

Theorem 3. *Assume that there exists $\delta < 1/2$ and $x_\delta > 0$ so that the estimate*

$$\#\{p \leq x : p \equiv 1 (\mathrm{mod}\, d)\} \geq \frac{\pi(x)}{2\phi(d)} \tag{4}$$

holds for all coprime positive integers $1 \leq a < d < x^{1-\delta}$ once $x \geq x_\delta$. Then, for every $\varepsilon > 0$, there exists a number $x_{\delta,\varepsilon}$ so that the inequality $\#R(x) > x^{1-2\delta-\varepsilon}$ holds once $x > x_{\delta,\varepsilon}$. In particular, if such an x_δ exists for all $\delta \in (0, 1/2)$, then $\#R(x) = x^{1-o(1)}$.

A theorem similar to our Theorem 3 addressing the Carmichael numbers appears in [1] and also in [6]. We point out that the fact that inequality (4) holds with some rather large $\delta \in [1/2, 1]$ for almost all pairs of coprime integers $1 \le a < d < x^{1-\delta}$ follows from Theorem 2.1 in [1]. Specifically, that theorem shows that inequality (4) holds with $\delta \ge 7/12$ for all $1 \le a < d < x^{1-\delta}$ with a and d coprime, except possibility for the set of all d divisible by some member of $D_\delta(x)$, a finite set whose cardinality is bounded in terms of δ alone and independently on x, and all members of $D_\delta(x)$ are larger than $\log x$. While the fact that (4) holds for almost all choices of d with some $\delta < 1$ was sufficient in [1] to prove that there are infinitely many Carmichael numbers, our argument, while closely following the argument from [1], does work only under the assumption that $\delta < 1/2$.

Throughout the proofs we use $c_1, c_2, \ldots$ for computable constants which are either absolute or depend on the given data, like a in the case of Theorem 1, or δ and ε in the case of Theorem 3. We also use the Vinogradov symbols $\gg$ and $\ll$ and the Landau symbols O and o with their regular meanings.

2. The proof of Theorem 1

The line of attack here is as follows. We fix a large positive real number x and first construct some small number $m \in R_a$. We shall request that $\sigma(m) < (\log^{1/2} x)/\log a$, that $\tau(\sigma(m)) \gg \log_2 x$, and that $\sigma(m)$ has a number $\gg \log_3 x$ of odd divisors. We then take the relation $m|a^{\sigma(m)}-1$. By the primitive divisor theorem (see [5], [3]), we get that $\omega(a^{\sigma(m)}-1) \gg \log_2 x$. Since $\omega(m) \ll \log m/\log_2 m \ll \log_2 x/\log_3 x$, it follows that most of the prime factors of $a^{\sigma(m)}-1$ are not prime factors of m. Select $\lambda := \lfloor c_1 \log_2 x \rfloor$ such odd prime factors of $a^{\sigma(m)}-1$ which are not prime factors of m, where c_1 is some constant with $c_1 < 1/(2\log 2)$, and let M be the product of all these primes. Then we still have $mM|a^{\sigma(m)}-1$, and therefore $mM|a^{\sigma(m)\sigma(M)}-1$. From the way we have selected our numbers, we have that $2^\lambda|\sigma(M)$, and $2^\lambda\sigma(m) < \log x/\log a$. In particular, $a^{2^\lambda\sigma(m)}-1 < x$ is a multiple of mM. Since $\lambda \gg \log_2 x$ and $\sigma(m)$ has a number $\gg \log_3 x$ of odd divisors, we get that $\tau(2^\lambda m) \gg \log_2 x \log_3 x$. Applying the primitive divisor theorem one more time, we get that $\omega(a^{2^\lambda\sigma(m)}-1) \gg \log_2 x \log_3 x$. Since $\omega(mM) \ll \log_2 x$, it follows that $a^{2^\lambda\sigma(m)}-1$ has a number $\gg \log_2 x \log_3 x$ prime factors which do not divide mM, and therefore this number has at least $\exp(c_2 \log_2 x \log_3 x)$ divisors d coprime to mM. Clearly, $mMd < x$ and $mMd \in R_a$ holds for every such value of d, which achieves the conclusion of Theorem 1.

We now give details. In the next lemma we explain how we choose the number m.

Lemma 1. *Let $k \ge 4$. Then there exists $d \in \{1, 3, 5, a^2-1\}$ so that $m := d \cdot (a^2+1) \cdot (a^4+1) \cdot \ldots \cdot (a^{2^{k-1}}+1) \in R_a$.*

Proof. Assume first that a is even. Then, each one of the numbers $a^2-1, a^2+1, \ldots, a^{2^{k-1}}+1$ is odd, they are coprime any two, and none of them is a perfect square. Thus, the sum of divisors of each one of these numbers is even, and therefore $2^k | \sigma(a^2-1)\cdot\sigma(a^2+1)\cdot\ldots\cdot\sigma(a^{2^{k-1}}+1) = \sigma(a^{2^k}-1)$. Hence, with $d := a^2-1$, and $m := (a^2-1)\cdot(a^2+1)\cdot\ldots\cdot(a^{2^{k-1}}+1) = a^{2^k}-1$, we have $m | a^{2^k}-1 | a^{\sigma(m)}-1$.

Assume now that a is odd. Then, each one of the numbers $(a^{2^i}+1)/2$ for $i = 2, \ldots, k-1$ is odd, and none of them is a perfect square. Indeed, if one of these numbers, say $(a^{2^i}+1)/2 = y^2$ is a perfect square with some $i \geq 2$, then with the substitution $x := a^{2^{i-2}}$ we would get a positive integer solution (x, y) with $x > 1$ for the Diophantine equation $x^4 + 1 = 2y^2$, and it is known that there is no such solution (see [13]). Thus, $2 | \sigma((a^{2^i}+1)/2)$ for $i = 2, \ldots, k-1$. We now write $(a^2+1)/2 = sy^2$, where $s \geq 1$ is a squarefree number.

Assume that $\omega(s) \geq 2$. Then $4 | \sigma((a^2+1)/2)$. This implies immediately that $2^k | \sigma(a^{2^k}-1)$, and so with $d := a^2-1$ we have that $m \in R_a$.

Assume that $s = 1$. Then $3 \not| a$, because otherwise we would get $1 \equiv 2y^2 (\mathrm{mod}\, 3)$, which is impossible. Clearly, $3 \not| (a^{2^i}+1)$ because -1 is not a quadratic residue modulo 3. Thus, with $d := 3$, we have

$$m := 3(a^2+1)\cdot\ldots\cdot(a^{2^{k-1}}+1) = 2^{k-1}\cdot 3\cdot\prod_{i=1}^{k-1}\Big(\frac{a^{2^i}+1}{2}\Big),$$

and

$$\sigma(m) = \sigma(2^{k-1})\cdot\sigma(3)\cdot\prod_{i=1}^{k-1}\sigma\Big(\frac{a^{2^i}+1}{2}\Big) = (2^k-1)\cdot 4\cdot\prod_{i=1}^{k-1}\sigma\Big(\frac{a^{2^i}+1}{2}\Big),$$

is a multiple of 2^k. Clearly, $a^{2^i}+1 | a^{2^k}-1$ holds for $i = 1, \ldots, k-1$, and $3 | a^{2^k}-1$ because 3 does not divide a. Thus, $m \in R_a$.

Assume that s is a prime. Thus, $2 | \sigma((a^2+1)/2)$. Write μ for the order at which 2 appears in the prime factorization of a^2-1. If $(a^2-1)/2^\mu$ is not a perfect square, it follows that we may set $d := a^2-1$, and then $2^k | \sigma(a^{2^k}-1)$, therefore $m \in R_a$. If $(a^2-1)/2^\mu$ is a perfect square, it follows that μ is odd, and therefore $a^2-1 = 2z^2$ holds with some positive integer z. Clearly, $5 \not| a$, because the congruence $-1 \equiv 2z^2 (\mathrm{mod}\, 5)$ is impossible. If $s \neq 5$, it follows that we may set $d := 5$, and then both primes 5 and s will appear at an odd power in the factorization of $5(a^2+1)$. Clearly, 5 is coprime to $a^{2^i}+1$ for $i \geq 2$, and thus it follows that $2^k | \sigma(5(a^2+1)\cdot\ldots\cdot(a^{2^{k-1}}+1))$, therefore $m \in R_a$. Finally, if $s = 5$, we get that $a^2+1 = 10y^2$, and therefore $a^4-1 = 20(yz)^2$. The only solution of this equation is for $a = 3$ (see [2]). In this case, $\sigma(3^{16}-1)$ is a multiple of 2^4, and by the previous arguments it follows that if we set $d := a^2-1$ and $m := a^{2^k}-1$, then $2^k | \sigma(m)$ holds for all $k \geq 4$. This completes the proof of Lemma 1. □

We now let x to be large. We want to construct such an m like in Lemma 1, so that the inequality

$$a^{\sigma(m)} < \exp(\sqrt{\log x}) \tag{5}$$

holds. Certainly, inequality (5) is equivalent to $\sigma(m) < (\log^{1/2} x)/\log a$. But $\sigma(m) \ll m \log_2 m < a^{2^k} \log_2 a^{2^k} \ll a^{2^k} k$ holds for large values of k, and, in particular, it follows that the inequality $\sigma(m) < a^{2^{k+1}}$ holds for large values of k as well. Thus, in order for (5) to hold, it suffices that $a^{2^{k+1}} < (\log^{1/2} x)/\log a$ holds, which is implied by

$$2^{k+1} < \frac{\log_2 x}{2\log a} - \frac{\log_2 a}{\log a}. \tag{6}$$

Thus, with c_2 any constant so that $c_2 < 1/(2\log a)$, we get that inequality (5) holds for large x provided that $2^{k+1} < c_2 \log_2 x$. We choose k to be the largest possible integer so that this last inequality holds. In particular, $2^k \geq c_3 \log_2 x$, where $c_3 := c_2/4$.

To continue, we need to know some lower bounds for $\tau(\sigma(m))$, where m is the number constructed in Lemma 1.

Lemma 2. *Let $m \in R_a$ be the number appearing in Lemma 1. There exist two constants $c_4 := c_4(a)$ and $c_5 := c_5(a)$, depending on a, so that if k is sufficiently large, then $\sigma(m)$ has at least $c_4 \cdot 2^k$ divisors of which at least $c_5 k$ of them are odd.*

Proof. We let c_6 to be a constant which is a positive integer to be fixed later, and we look at the numbers of the form

$$\sigma\Big(\frac{a^{2^{i_1}}+1}{2^\gamma}\Big)\cdot\sigma\Big(\frac{a^{2^{i_2}}+1}{2^\gamma}\Big)\cdot\ldots\cdot\sigma\Big(\frac{a^{2^{i_s}}+1}{2^\gamma}\Big), \tag{7}$$

where $k > i_1 > i_2 \ldots > i_s > c_6$ and $\gamma = 0, 1$ according to whether a is even or odd. It is clear that all these numbers are divisors of $\sigma(m)$. The question reduces therefore to showing that a positive proportion of those are distinct. Well, let us assume that

$$\begin{aligned}\sigma\Big(\frac{a^{2^{i_1}}+1}{2^\gamma}\Big)\cdot\sigma\Big(\frac{a^{2^{i_2}}+1}{2^\gamma}\Big)\cdot\ldots\cdot\sigma\Big(\frac{a^{2^{i_s}}+1}{2^\gamma}\Big)\\ = \sigma\Big(\frac{a^{2^{j_1}}+1}{2^\gamma}\Big)\cdot\sigma\Big(\frac{a^{2^{j_2}}+1}{2^\gamma}\Big)\cdot\ldots\cdot\sigma\Big(\frac{a^{2^{j_{s'}}}+1}{2^\gamma}\Big)\end{aligned} \tag{8}$$

holds where $k > i_1 > i_2 > \ldots > i_s > c_6$, $k > j_1 > j_2 > \ldots > j_{s'} > c_6$, and $(i_1, \ldots, i_s) \neq (j_1, \ldots, j_{s'})$. We shall first show that there exist a constant c_7, so that the number shown at (7) satisfies

$$\sigma\Big(\frac{a^{2^{i_1}}+1}{2^\gamma}\Big)\cdot\sigma\Big(\frac{a^{2^{i_2}}+1}{2^\gamma}\Big)\cdot\ldots\cdot\sigma\Big(\frac{a^{2^{i_s}}+1}{2^\gamma}\Big) < c_7\frac{a^{2^{i_1}+\ldots+2^{i_s}}}{2^{\gamma s}}. \tag{9}$$

Indeed, for every odd prime number p which divides $a^{2^i}+1$ for some i, write $t(p) := 2^i$. Clearly, $t(p)$ is uniquely determined. Using the fact that the inequality $\sigma(n)/n < n/\phi(n)$ holds for all positive integers n, we get

$$\begin{aligned}(10)\qquad \sigma\Big(\frac{a^{2^{i_1}}+1}{2^\gamma}\Big)\cdot\sigma\Big(\frac{a^{2^{i_2}}+1}{2^\gamma}\Big)\cdot\ldots\cdot\sigma\Big(\frac{a^{2^{i_s}}+1}{2^\gamma}\Big)&\\ < \frac{a^{2^{i_1}+\ldots+2^{i_s}}}{2^{\gamma s}}\cdot\prod_{j=c_6}^{\infty}\Big(1+\frac{1}{a^{2^j}}\Big)\cdot\prod_{j=c_6}^{\infty}\frac{\phi((a^{2^j}+1)/2^\gamma)}{(a^{2^j}+1)/2^\gamma}&\\ \ll \frac{a^{2^{i_1}+\ldots+2^{i_s}}}{2^{\gamma s}}\cdot\prod_{\substack{t(p)=2^j\\ j\ge c_6}}\Big(1+\frac{1}{p-1}\Big).&\end{aligned}$$

In [7], it was shown that the sum of the reciprocals of all prime divisors of the Fermat numbers is convergent. The same argument from there can be used to lead to the conclusion that the sum of the reciprocals of the prime divisors of the numbers of the form $a^{2^i}+1$ for $i \ge 1$ is convergent as well. This and (10) imply (9). We now return to (8), assume that $\{i_1,\ldots,i_s\}$ is disjoint from $\{j_1,\ldots,j_{s'}\}$, and assume that $i_1 > j_1$. Write $S := 2^{i_1}+\ldots+2^{i_s}$ and $S' := 2^{j_1}+\ldots+2^{j_{s'}}$. Using the trivial lower bound $\sigma(n) > n$ on the left–hand side of (8), and using (9) to get an upper bound for the right–hand side of (8), we get

$$(11)\qquad a^{S-S'} \le \frac{c_7}{2^{\gamma(s'-s)}}.$$

Since $\max(s',s) \le i_1$, it follows that the estimate $S-S' = O(i_1)$ holds, where the constant understood in O above depends on a. It is now easy to see that since the binary digits of 1 of S and S' do not overlap (they are concentrated in different positions), there exists a constant c_8 so that $j_1 = i_1-1, j_2 = i_1-2,\ldots,j_l = i_1-l$ holds for $l < \log i_1/\log 2 - c_8$. But this shows that $s \ll \log i_1$ and $s' \gg i_1$. In particular, if i_1 is large enough, then (11) with $\gamma = 1$ shows that $a^{S-S'} < 1$, which is impossible because $S > S'$. Thus, the only possibility when i_1 is large enough (i.e., when $i_1 > c_6$ and c_6 is large enough) is when $\gamma = 0$, for which we get $S - S' = O(1)$. But since $i_s > c_6,\ j_{s'} > c_6$, we get that $S-S'$ is a multiple of 2^{c_6}. This together with the fact that $S-S' = O(1)$ leads to a contradiction once c_6 is chosen to be sufficiently large. Thus, we have shown that all the numbers shown at (7) are distinct. The number of such numbers is

$$\sum_{s=0}^{k-c_6-1}\binom{k-1-c_6}{s} = 2^{k-1-c_6} = c_4\cdot 2^k,$$

where $c_4 := 2^{-1-c_6}$ (the choice $s := 0$ above accounts for the divisor 1 of $\sigma(m)$).

It remains to find a lower bound for the number of odd divisors. For this, it suffices to find an upper bound for the exponent at which 2 divides $\sigma(m)$. Write λ_i for the exponent at which 2 divides $\sigma((a^{2^i}+1)/2^\gamma)$, where again

$\gamma = 0, 1$ according to whether a is even or odd. Let Ω_i be the number of prime divisors counted with multiplicities of $(a^{2^i}+1)/2^\gamma$. Since every prime divisor p of $(a^{2^i}+1)/2^\gamma$ is congruent to 1 modulo 2^{i+1} for $i \geq 1$, it follows that $a^{2^i} > (a^{2^i}+1)/2 > (2^{i+1})^{\Omega_i}$, and therefore $\Omega_i \ll 2^i/i$. Assume $p^\alpha || a^{2^i}+1$, where p is odd. Then, $\sigma(p^\alpha) = p^\alpha + \ldots + 1 \equiv \alpha + 1 (\mathrm{mod}\, 2^{i+1})$. Since $\alpha \leq \Omega_i \ll 2^i/i$, it follows that if $i > c_9$ is sufficiently large, then $\alpha + 1 < 2^{i+1}$. Thus, the order at which 2 divides $\sigma(p^\alpha)$ is at most $\log(\alpha+1)/\log 2 \ll \alpha$. This argument shows that $\lambda_i \ll \Omega_i \ll 2^i/i$. Thus, the order at which 2 can appear in $\sigma(m)$ is

$$\ll \sum_{i=1}^{k-1} \frac{2^i}{i} \ll \int_1^k \frac{2^t}{t} dt \ll \frac{2^k}{k}.$$

Since the total number of divisors of $\sigma(m)$ is $\gg 2^k$, we get that the number of odd divisors of $\sigma(m)$ is $\gg k$, which concludes the proof of Lemma 2. □

We now continue with our argument. Since for us we have $2^k > c_3 \log_2 x$, Lemma 2 shows that $\tau(\sigma(m)) > c_{10} \log_2 x$ and that at least $c_{11} \log_3 x$ of the divisors of $\sigma(m)$ are odd, where $c_{10} = c_3 \cdot c_4$, and c_{11} can be taken to be any constant slightly smaller than $c_5/\log 2$ provided that x is large. Let d be any divisor of $\sigma(m)$. By the primitive divisor theorem (see [5], [3]), there exists a prime number $p|a^d - 1$ (in particular, dividing $a^{\sigma(m)} - 1$) so that $p \not| a^t - 1$ for any positive integer $t < d$, except possibly if $d = 1, 2, 3, 6$. Thus, $a^{\sigma(m)} - 1$ has at least $\tau(\sigma(m)) - 4 > c_{10} \log_2 x - 4$ prime factors. Of course, some of these are prime factors of m. However, since $\omega(m) \ll \log m/\log_2 m \ll 2^k/k \ll (\log_2 x)/(\log_3 x)$, it follows that $a^{\sigma(m)} - 1$ has at least $c_{12} \log_2 x$ odd prime factors coprime to m, where c_{12} can be taken to be any fixed constant smaller than c_{10}. Choose $\lambda := \lfloor c_1 \log_2 x \rfloor$ such prime factors of $a^{\sigma(m)} - 1$, where $c_1 := \min\{c_{11}, 1/(2 \log 2)\}$, and let M be the product of all of these. Then m and M are coprime, $Mm | a^{\sigma(m)} - 1$ and $\sigma(mM) = \sigma(m) \cdot \sigma(M)$. Clearly, $\sigma(M)$ is a multiple of $2^\lambda < \log^{1/2} x$. From (5), we get

$$a^{2^\lambda \sigma(m)} - 1 < a^{2^\lambda \sigma(m)} < e^{\log x} = x. \tag{12}$$

We now count the number of divisors of $2^\lambda \sigma(m)$. The number of them is at least $\lambda + 1$ times the number of odd divisors of $\sigma(m)$, and so it is at least $c_{13} \log_2 x \log_3 x$, where $c_{13} := c_{11} \cdot c_{12}$. By the primitive divisor theorem, the number $a^{2^\lambda \sigma(m)} - 1$ will have at least $c_{13} \log_2 x \log_3 x - 4$ prime factors. We now show that most of these are coprime to mM. Indeed, $\omega(mM) \leq \omega(m) + \omega(M) = \lambda + O(\log_2 x/\log_3 x) \ll \log_2 x$, therefore indeed $a^{2^\lambda \sigma(m)} - 1$ has at least $c_{14} \log_2 x \log_3 x$ prime factors coprime to mM, where we can take c_{14} to be any constant strictly smaller than c_{13}. Let d be an arbitrary squarefree number built up with these primes. Then $n := mMd$ has $\sigma(n) = \sigma(m)\sigma(M)\sigma(d)$ a multiple of $2^\lambda \sigma(m)$, and $n | a^{2^\lambda \sigma(m)} - 1$. In particular, $n < x$. Thus, n is counted by $\# R_a(x)$, and the number of such numbers n is $> 2^{c_{14} \log_2 x \log_3 x} = \exp(c_{15} \log_2 x \log_3 x)$, where $c_{15} := c_{14} \log 2$. Theorem 1 is therefore proved.

3. The proof of Theorem 2

For every positive integer n, we write $P(n)$ for the largest prime factor of n with the convention that $P(1) := 1$. For $1 < y < x$ we write $\Psi(x, y) := \#\{1 \le n < x \,|\, P(n) < y\}$. For this proof, we shall need some estimates for $\Psi(x, y)$ once x is large and in a certain range of y versus x. The following result appears in [9].

Lemma 3. *Suppose that $\varepsilon > 0$ is arbitrarily small, but fixed. If y satisfies $\exp((\log x)^{\varepsilon}) < y < \exp((\log x)^{1-\varepsilon})$, then*

$$\Psi(x, y) = x \exp((-1 + o(1))u \log u), \qquad u := \frac{\log x}{\log y}. \tag{13}$$

We now let x be large, and set $y := \exp(c_1(\log x \log_2 x)^{1/2})$, where c_1 is a constant to be fixed later. Clearly, for large x our y satisfies the condition from Lemma 3 with $\varepsilon := 1/3$.

Let $A_1(x) := \{1 \le n < x \,|\, P(n) < y\}$. Then $u := \log x/\log y = 2/c_1 \cdot (\log x/\log_2 x)^{1/2}$, therefore $u \log u = \frac{1}{c_1} \cdot (1 + o(1))(\log x \log_2 x)^{1/2}$. Thus, with Lemma 3, we get that

$$\#A_1(x) := \Psi(x, y) = x \cdot \exp\Big(-\frac{1}{c_1}(1 + o(1))(\log x \log_2)^{1/2}\Big). \tag{14}$$

Let $A_2(x) := \{1 \le n < x \,|\, n \notin A_1(x) \text{ and } P(n)^2 | n\}$. Clearly,

$$\#A_2(x) < \sum_{p>y} \frac{x}{p^2} = o\Big(\frac{x}{y}\Big) < x \cdot \exp(-c_1(\log x \log_2 x)^{1/2}). \tag{15}$$

We now assume $n \in R_a(x)$ for some $a < 2 \log x$, and suppose that $n \notin A_1(x) \cup A_2(x)$. Then $n = mP$, where $P \ge y$, and $P(m) < P$. Fix a and for a prime p, let $t_a(p)$ be the order of apparition of p in the sequence $(a^n - 1)_{n \ge 1}$. That is, $t_a(p)$ is the smallest positive integer k so that $p | a^k - 1$ (and it is infinity if $p | a$). Let $\mathcal{P}_a := \{p \text{ prime} \,|\, t_a(p) < p^{1/3}\}$, and for any positive integer z let $\mathcal{P}_a(z) := \{p \in \mathcal{P}_a \,|\, p < z\}$. We claim that the inequality $\#\mathcal{P}_a(z) < 2z^{2/3} \log a$ holds uniformly in $a > 1$ and $z > 1$. Indeed, fix the number z. Then, every prime number counted by $\#\mathcal{P}_a(z)$ satisfies $p | a^k - 1$ for some positive integer $k < z^{1/3}$. In particular,

$$\prod_{\substack{p \in \mathcal{P}_a \\ p < z}} p \le \prod_{1 \le k < z^{1/3}} (a^k - 1) < \exp\Big(\log a \Big(\sum_{k < z^{1/3}} k \Big)\Big) < \exp(z^{2/3} \log a),$$

where the last inequality holds for all $z > 1$. Let $t := \#\mathcal{P}_a(z)$. Since the product on the left is at least $\prod_{i=1}^{t} p_i > 2^t = \exp(t \log 2)$, where $2 = p_1 < p_2 < \dots$ are all the prime numbers, it follows that

$$t < \frac{\log a}{\log 2} z^{2/3} < 2z^{2/3} \log a. \tag{16}$$

Let $A_3(x) := \{1 \le n < x \,|\, n \notin A_1(x) \cup A_2(x), P(n) \in \mathcal{P}_a \text{ holds with some } a < 2\log x\}$. Fix the number a and the number m. Then $n = Pm$, and $P < x/m$. By (16), it follows that the number of such numbers P when a and m are fixed is $< 2(x/m)^{2/3} \cdot \log a$. Summing up over all $a < 2\log x$, we get that the number of such numbers n with m fixed is $< 4\log x(\log(2\log x))(x/m)^{2/3} < 5\log x \log_2 x (x/m)^{2/3}$, with the last inequality holding for large values of x. Summing up over all the values of m, and keeping in mind that $m < x/y$, it follows that

$$\begin{aligned}(17)\qquad \#A_3(x) &< 5\log x\log_2 x \sum_{1\le m<x/y}\Big(\frac{x}{m}\Big)^{2/3} = 5x^{2/3}\log x\log_2 x \sum_{1\le m<x/y}\frac{1}{m^{2/3}}\\ &\ll x^{2/3}\log x\log_2 x\int_1^{x/y}\frac{dt}{t^{2/3}} \ll x^{2/3}\log x\log_2 x\cdot t^{1/3}\Big|_{t:=1}^{t:=x/y}\\ &\ll \frac{x\log x\log_2 x}{y^{1/3}} = x\exp\Big(-\frac{c_1}{3}(1+o(1))(\log x\log_2 x)^{1/2}\Big).\end{aligned}$$

Finally, let $A_4(x) := \{1 \le n < x \,|\, n \in R_a(x) \text{ for some } a < 2\log x \text{ and } n \notin A_1(x) \cup A_2(x) \cup A_3(x)\}$. Fix again the number m and the number a. Since $n \notin A_2(x)$, it follows that $\sigma(n) = \sigma(m)(P+1)$. Since $n \in R_a$, it follows that $P|a^{\sigma(mP)}-1$, therefore $P|a^{\sigma(m)(P+1)}-1$. However, from the definition of $t_a(P)$ and Fermat's Little Theorem, we have that $t_a(P)|(P-1)$. In particular, it follows that $t_a(P)|\gcd(P-1, \sigma(m)(P+1))$, therefore $t_a(P)|\gcd(2\sigma(m), P-1)$. Write $d := t_a(p)$. It follows that d is a divisor of $2\sigma(m)$, and $P \equiv 1 \pmod d$. Moreover, since $n \notin A_1(x) \cup A_3(x)$, it follows that $d = t_a(P) > P^{1/3} > y^{1/3}$. Since $P < x/m$ and $P \equiv 1 (\mathrm{mod}\, d)$, it follows that with m, a and d fixed, the number of such numbers $n < x$ is at most $\pi(x/m, 1, d) \ll x/(md)$, (note that $d < P$, therefore $md < mP < x$). Keeping m and a fixed and summing up over all $d > y^{1/3}$, we get that the number of such n is

$$(18)\qquad \ll \frac{x}{y^{1/3}m}\cdot\tau(2\sigma(m)) \ll \frac{x}{y^{1/3}}\cdot\frac{\tau(m)}{m}.$$

The upper bound (18) is independent on a, so that we get that the number of numbers $n \in A_4(x)$ for which m is fixed is

$$(19)\qquad \ll \frac{x\log x}{y^{1/3}}\cdot\frac{\tau(\sigma(m))}{m}.$$

Summing up over all $m < x$, we get that

$$(20)\qquad \#A_4(x) \ll \frac{x\log x}{y^{1/3}}\sum_{m<x}\frac{\tau(\sigma(m))}{m}.$$

The following lemma is an adaptation of a result from [8].

Lemma 4. *There exists an absolute constant c_3 so that for x sufficiently large holds the inequality*

$$(21)\qquad \sum_{n<x}\tau(\sigma(n))/n < \exp(c_3(\log x)^{1/2}).$$

We postpone for the time being the proof of this lemma and we complete the proof of Theorem 2. With (20) and Lemma 4, it follows that

$$\#A_4(x) < x\,\exp\Big(-\frac{c_1}{3}(\log x\log_2 x)^{1/2}+\log_2 x+c_3(\log x)^{1/2}\Big) = x\,\exp\Big(-\frac{c_1}{3}(1+o(1))(\log x\log_2 x)^{1/2}\Big). \tag{22}$$

It is now clear that the sum of $\#A_i(x)$ for $i=1,\dots,4$ is an upper bound for the number of $n<x$ which belong to $R_a(x)$ for some $a<2\log x$, and comparing (14), (15), (17), and (22), it follows that the estimate $\#\cup_{1<a<2\log x} R_a(x) < x\exp(-c_3(\log x\log_2 x)^{1/2})$ holds for large x with any constant c_3 so that $c_3<\min\{1/c_1, c_1/3\}$. The best cut point for our constant c_3 is, of course, achieved when $1/c_1=c_1/3$, i.e. $c_1=\sqrt{3}$, and therefore the constant c_2 from the statement of Theorem 2 can be chosen to be any constant strictly smaller than $1/\sqrt{3}$, and then estimate (3) will hold once x is sufficiently large.

The proof of Lemma 4. This is a simplified version of the method proving the main Theorem in [8], where we gave lower and upper bounds for the mean value of the function $\tau(\phi(n))$ in the interval $(1,x)$. We use the fact that the inequality $\tau(mn)\le\tau(m)\tau(n)$ holds for all positive integers mn. For every $n:=\prod_{p^{\alpha_p}||n} p^{\alpha_p}$, where the numbers p denote distinct primes and α_p denote positive integers, we use the above inequality to say that

$$\tau(\sigma(n))\le\prod_{p^{\alpha_p}||n}\tau(\sigma(p^{\alpha_p})). \tag{23}$$

We let x be a large real number, $s>0$ to be a parameter depending on x to be chosen later, and write $U(x):=\sum_{n<x}\tau(\sigma(n))/n$. We use (23), to say that

$$\begin{aligned} U(x) &\le \sum_{k\ge0}\sum_{\substack{p_1^{\alpha_1},\dots,p_k^{\alpha_k}\\ p_1^{\alpha_1}\cdot\dots\cdot p_k^{\alpha_k}<x}}\prod_{i=1}^{k}\frac{\tau(\sigma(p_i^{\alpha_i}))}{p_i^{\alpha_i}}\\ &\le \sum_{k\ge0}\sum_{\substack{p_1^{\alpha_1},\dots,p_k^{\alpha_k}\\ p_1^{\alpha_1}\cdot\dots\cdot p_k^{\alpha_k}<x}}\Big(\frac{x}{p_1^{\alpha_1}\cdot\ldots\cdot p_k^{\alpha_k}}\Big)^s\cdot\prod_{i=1}^{k}\frac{\tau(\sigma(p_i^{\alpha_i}))}{p_i^{\alpha_i}}\\ &\le x^s\sum_{k\ge0}\sum_{p_1^{\alpha_1},\dots,p_k^{\alpha_k}<x}\prod_{i=1}^{k}\frac{\tau(\sigma(p_i^{\alpha_i}))}{p_i^{\alpha_i(1+s)}} = x^s\prod_{2\le p<x}\Big(1+\sum_{\alpha\ge1}\frac{\tau(\sigma(p^\alpha))}{p^{\alpha(1+s)}}\Big)\\ &= \exp\Big(s\log x+\sum_{2\le p<x}\frac{\tau(p+1)}{p^{1+s}}+\sum_{p\ge2}\sum_{\alpha\ge2}\frac{\tau(\sigma(p^\alpha))}{p^{\alpha(1+s)}}\Big), \end{aligned} \tag{24}$$

where the indices of summation $p_i^{\alpha_i}$ in (24) stand for distinct prime powers >1, and in the inequality (24) we used the fact that the inequality $1+y<\exp(y)$

holds for all $y>0$. And so, with Rankin's method, it suffices to find a value of s, depending on x, so that the expression appearing in the right–hand side of (24) is as small as possible. Let us first notice that the double sum appearing in the right–hand side of (24) is $O(1)$. Indeed, for every $\varepsilon>0$, there exists n_ε, so that if $n>n_\varepsilon$ then the inequality $\tau(\sigma(n))<n^\varepsilon$ holds. Taking $\varepsilon:=1/3$, we get

$$\sum_{p\geq 2}\sum_{\alpha\geq 2}\frac{\tau(\sigma(p^\alpha))}{p^{\alpha(1+s)}}<O(1)+\sum_{p\geq 2}\sum_{\alpha\geq 2}\frac{1}{p^{2\alpha/3}} \tag{25}$$

$$=O(1)+\sum_{p\geq 2}\frac{1}{p^{4/3}}\sum_{\beta\geq 0}\frac{1}{p^{2\beta/3}}\ll 1+\sum_{p\geq 2}\frac{1}{p^{4/3}}=O(1).$$

So, we shall now concentrate on finding, for a given large value of x, a parameter $s>0$ such that

$$f(s):=s\log x+\sum_{2\leq p<x}\frac{\tau(p+1)}{p^{1+s}} \tag{26}$$

is as small as possible function of x. For any positive real number y, let $C(y):=\sum_{2\leq p<y}\tau(p+1)$. From [4], we know that there exists an absolute constant c_4 such that

$$C(y):=c_4y+O\Big(\frac{y\log_2 y}{\log y}\Big) \tag{27}$$

holds for large values of y. We use partial integration and (27) to get that

$$\begin{aligned}\sum_{2\leq p<x}\frac{\tau(p+1)}{p^{1+s}} &= \frac{C(x)}{x^{1+s}}+\int_2^x\frac{1+s}{t^{2+s}}C(t)\,\mathrm{d}t\\ &= O(1)+\int_2^x\frac{(1+s)c_4}{t^{1+s}}+O\left(\frac{\log\log t}{t^{1+s}\log t}\right)\,\mathrm{d}t\\ &= \frac{c_4}{s}(1+s)(2^{-s}-x^{-s})+O\left(\log\log x\int_2^x\frac{\mathrm{d}t}{t^{1+s}\log t}\right).\end{aligned}$$

We assume that $1/\log x<s<1$. The last integral may be broken at $e^{1/s}$. The integrand for t smaller than this bound is $\ll 1/t\log t$, and in the remaining range the integrand is $\ll 1/st\log^2 t$. So the integral in the first range is $\ll \log(1/s)\leq\log\log x$, and in the second range is $\ll 1$. Hence,

$$\sum_{p\geq 2}\frac{\tau(p+1)}{p^{1+s}}=\frac{c_4}{s}(2^{-s}-x^{-s})+O((\log\log x)^2),$$

so that

$$f(s)\leq s\log x+\frac{c_4}{s}+O((\log\log x)^2). \tag{28}$$

Setting $g(s):=s\log x+c_4/s$, we have $g'(s)=\log x-c_4/s^2$, and therefore $g(s)$ is minimal when $s=s_x:=\sqrt{c_4/\log x}$. Substituting this value of s in (28) and

setting $c_5 := 2\sqrt{c}_4$, we get $f(s_x) = c_5\sqrt{\log x} + O((\log\log x)^2)$, and putting this together with (25) into (24) we get

$$U(x) < \exp(c_5\sqrt{\log x} + O((\log\log x)^2)), \tag{29}$$

which proves (21) for large x, where the constant c_3 appearing in (21) can be taken to be any constant strictly larger than our c_5.

This completes the proof of Lemma 4, and the proof of Theorem 2. □

4. The proof of Theorem 3

We let $\delta < 1/2$ be fixed so that inequality (4) holds for all $x > x_\delta$. We let $\varepsilon > 0$ be some sufficiently small number which we shall later on make more explicit in terms of δ. We write $\mathcal{D}$ for the set of all numbers $1 \le d < x^\alpha$, where $\alpha := ((1-\delta)^2 - \varepsilon)/\delta$ having $P(d) < y$, where $y := x^\varepsilon$. Write $z := x^{1-\delta}$. For every odd prime number $q < y$ we write $\mathcal{D}_q$ for subset of those $d \in \mathcal{D}$ which are coprime to q. Fix such an odd prime number q. For any $d \in \mathcal{D}_q$, write $a_{q,d}$ for the uniquely defined positive integer $< dq$ which is congruent to 1 modulo d and -1 modulo q. Such an integer exists and is unique by the Chinese Remainder Lemma, and is coprime to both d and q. Note that the inequality $dq < (dz)^{1-\delta}$ holds, because this inequality is equivalent to $d^\delta q < z^{1-\delta} = x^{(1-\delta)^2}$, and this last inequality is satisfied because $d^\delta q < d^\delta y < x^{\alpha\delta+\varepsilon} = x^{(1-\delta)^2}$ holds by our choice for α. With our assumption, for large x (that is x so large so that the inequality $z = x^{1-\delta} > x_\delta$ holds) we have that

$$\pi(dz; a_{q,d}, dq) \ge \frac{\pi(dz)}{2\phi(dq)} \ge \frac{\pi(dz)}{2dq} \ge \frac{dz}{2dq\log dz} > \frac{2z\delta}{(1-\delta)q\log x}, \tag{30}$$

where in the above estimates we used the fact that $\pi(t) \ge t/\log t$ holds for all $t \ge 17$ (see [10]), in particular, for $t := dz$ with x sufficiently large, together with the fact that

$$\begin{aligned} \log dz = \log d + \log z &< \alpha\log x + (1-\delta)\log x \\ &< \Big(\frac{(1-\delta)^2}{\delta} + (1-\delta)\Big)\log x = \frac{(1-\delta)}{\delta}\log x. \end{aligned} \tag{31}$$

Summing up (30) over all $d \in \mathcal{D}_q$, it follows that

$$\sum_{d\in\mathcal{D}_q} \pi(dz; a_{q,d}, dq) \ge \frac{2z\delta}{(1-\delta)\log xq} \cdot \#\mathcal{D}_q. \tag{32}$$

We claim that there exists a constant $c_1 := c_1(\delta,\varepsilon)$ such that for large x the estimate

$$\#\mathcal{D}_q > c_1 x^\alpha \tag{33}$$

holds. Indeed, it is clear that

$$\#\mathcal{D}_q = \Psi(x^\alpha, y) - \Psi\Big(\frac{x^\alpha}{q}, y\Big). \tag{34}$$

Write $u := \log(x^\alpha)/\log y = \alpha/\varepsilon$, and $u_q := \log(x^\alpha/q)/\log y = (\alpha - \log q/\log x)/\varepsilon$. By Theorem 6 on page 367 in [14], we know that

$$\Psi(x^\alpha, y) = \rho(u)x^\alpha + O\Big(\frac{x^\alpha}{\log y}\Big) = \rho\Big(\frac{\alpha}{\varepsilon}\Big)x^\alpha + O_\varepsilon\Big(\frac{x^\alpha}{\log x}\Big), \tag{35}$$

where ρ stands for Dickman's function. In particular, if $q > \log x$, then $\Psi(x^\alpha/q, y) \le x^\alpha/q \le x^\alpha/\log x$, and therefore with (34) and (35), we get that

$$\#\mathcal{D}_q = \rho\Big(\frac{\alpha}{\varepsilon}\Big)x^\alpha + O_\varepsilon\Big(\frac{x^\alpha}{\log x}\Big) > \frac{1}{2}\cdot\rho\Big(\frac{\alpha}{\varepsilon}\Big)\cdot x^\alpha, \tag{36}$$

with the last inequality holding for $x > x_{\delta,\varepsilon}$. If, on the other hand, $q < \log x$, we then get

$$\begin{aligned}
\#\mathcal{D}_q &= \rho\Big(\frac{\alpha}{\varepsilon}\Big)\cdot x^\alpha - \rho\Big(\frac{\alpha - \frac{\log q}{\log x}}{\varepsilon}\Big)\cdot\frac{x^\alpha}{q} + O_\varepsilon\Big(\frac{x^\alpha}{\log x}\Big) \\
&= \rho\Big(\frac{\alpha}{\varepsilon}\Big)\cdot\Big(1-\frac{1}{q}\Big)\cdot x^\alpha + \Big(\rho\Big(\frac{\alpha}{\varepsilon}\Big) - \rho\Big(\frac{\alpha - \frac{\log q}{\log x}}{\varepsilon}\Big)\Big)\cdot\frac{x^\alpha}{q} + O_\varepsilon\Big(\frac{x^\alpha}{\log x}\Big) \\
&= \rho\Big(\frac{\alpha}{\varepsilon}\Big)\cdot\Big(1-\frac{1}{q}\Big)\cdot x^\alpha + O_{\varepsilon,\delta}\Big(\frac{x^\alpha\log_2 x}{\log x}\Big) > \frac{1}{2}\cdot\rho\Big(\frac{\alpha}{\varepsilon}\Big)\cdot x^\alpha
\end{aligned}$$

holds for sufficiently large values of x, where in the above estimates we used the intermediate value theorem for the function ρ in the vicinity of α/ε for large values of x, together with the fact that $3 \le q < \log x$. This proves (33) with $c_1 := \rho(\alpha/\varepsilon)$. Writing $c_2 := 2\delta c_1/(1-\delta)$, we get that

$$\sum_{d\in\mathcal{D}_q}\pi(dz; a_{q,d}, dq) > c_2\frac{zx^\alpha}{q\log x}. \tag{37}$$

The left–hand side of (37) clearly counts all pairs (d,p), with $d \in \mathcal{D}_q$, p prime, $p < dz$ so that $p-1 = dm$ and $p+1 = qn$ hold for some integers m and n with $m < z$. Since m can assume at most z values, it follows that there exists a value of m, let us call it $m := m_q$, so that the number $md+1$ is a prime number p with $p+1 \equiv 0 (\mathrm{mod}\, q)$ for a subset of numbers $d \in \mathcal{D}_q$ of cardinality $> c_1x^\alpha/(q\log x)$. Fix such a value of m_q and let $\mathcal{P}_q$ be the set of such resulting primes p with $(p-1)/d = m_q$ for some $d \in \mathcal{D}_q$, and also $p+1 \equiv 0 (\mathrm{mod}\, q)$. We do this for all the prime numbers $q < y$ which are odd, and we obtain pairs $(m_q, \mathcal{P}_q)$ formed of positive integers $m_q < z = x^{1-\delta}$, and primes $\mathcal{P}_q$. Of course, we have lower bounds on the cardinality of $\mathcal{P}_q$ which are uniform in q, but for our purposes we will need to select subsets of primes $\mathcal{P}'_q$ of $\mathcal{P}_q$, so that the cardinalities of

these are still sufficiently large, but so that moreover all such are disjoint for distinct values of the parameter q. We show that if x is sufficiently large in a way depending on δ and ε, we can then choose subsets $\mathcal{P}'_q$ of $\mathcal{P}_q$, so that the inequality $\#\mathcal{P}'_q > 3x^{\alpha-3\varepsilon}$ holds for all odd $q < y$, and so that all these subsets are disjoint for distinct values of q. To see that we can do this, let us assume the contrary, that is that there is no way of choosing such subsets. Write k for the largest positive integer, $k < \pi(y) - 1 :=$ the number of odd primes below y, so that there exists k odd primes $q_1, \ldots, q_k$ smaller than y and for each one of these a subset $\mathcal{P}'_{q_i}$ of $\mathcal{P}_{q_i}$ of cardinality $\lfloor 3x^{\alpha-3\varepsilon} \rfloor + 1$ but so that for any other $q < y$ distinct from $q_i, i := 1, \ldots, k$, there are less than $3x^{\alpha-3\varepsilon}$ elements in $\mathcal{P}_q$ which are not already in $\cup_{1\le i\le k}\mathcal{P}'_{q_i}$. Well, note that if this is so, we then have

$$\#\cup_{1\le i\le k}\mathcal{P}'_{q_i} < \pi(y)(\lfloor 3x^{\alpha-3\varepsilon}\rfloor + 1) < 6x^{\alpha-2\varepsilon}. \tag{38}$$

However, since $q < x^{\varepsilon}$, and since the inequality $x^{\varepsilon/2} > \log x/c_2$ holds for all sufficiently large values of x, we get by (37) that

$$\#\mathcal{P}_q > x^{\alpha-3\varepsilon/2} > 6x^{\alpha-2\varepsilon} + 3x^{\alpha-3\varepsilon} + 1. \tag{39}$$

Thus, with (38) and (39), we do conclude that such a maximal value of $k < \pi(y) - 1$ does not exist, so that we have indeed proved the existence of pairs $(m_q, \mathcal{P}'_q)_{3\le q<y}$ formed by numbers $m_q < z$ together with sets of primes $\mathcal{P}'_q$, so that every prime $p \in \mathcal{P}'_q$ has $p - 1 \equiv 0 (\mathrm{mod}\, m_q)$, $P((p-1)/m_q) < x^{\varepsilon}$, $p + 1 \equiv 0 (\mathrm{mod}\, q)$, and the sets $\mathcal{P}'_q$ have the properties that $\#\mathcal{P}'_q > 3x^{\alpha-3\varepsilon}$, and moreover $\mathcal{P}_q \cap \mathcal{P}_{q'} = \emptyset$ holds for all $3 \le q < q' < y$. For every odd $q < y$, we eliminate all the prime numbers $p < y$ which might belong to $\mathcal{P}'_q$ and write $\mathcal{P}''_q$ for the remaining subset of $\mathcal{P}'_q$. Clearly, the cardinality of $\mathcal{P}''_q$ satisfies $\#\mathcal{P}''_q > \#\mathcal{P}'_q - \pi(y) > 3x^{\alpha-3\varepsilon} - x^{\varepsilon} > 2x^{\alpha-3\varepsilon}$, if ε is chosen sufficiently small, say such that $4\varepsilon < \alpha$ holds.

We now look at the numbers m_q for odd values of $q < y$. Clearly, m_q and q are coprime because $m_q | p - 1$ and $q | p + 1$ for $p \in \mathcal{P}''_q$. We fix such a prime number q, and choose a number $l(q)$ which satisfies the following properties:

i. $m_q | (q^{l(q)} - 1)/(q - 1)$;

ii. $l(q)$ satisfies the inequality $z/2 \le l(q) < z$.

It clearly suffices to prove the existence of a positive integer $l(q) < z$ which satisfies i above. Indeed, if a positive integer $l(q) < z$ satisfies i above, then either $l(q) \ge z/2$, in which case $l(q)$ satisfies ii above as well, or $l(q) < z/2$. In this last case, any multiple of $l(q)$ will satisfy i above as well, and the interval $(z/2, z)$ will contain a multiple of $l(q)$. Thus, by replacing $l(q)$ by some multiple of it which lives in the interval $(z/2, z)$, we have produced a number $l(q)$ which satisfies both i and ii above. To produce a positive integer $l(q) < z$ satisfying i above, write

$$m_q := \prod_{p^{\alpha_{p,q}} || m_q} p^{\alpha_{p,q}} \tag{40}$$

and let

$$l(q) := \mathrm{lcm}[u(p^{\alpha_{p,q}}) \,|\, p^{\alpha_{p,q}} || m_q], \tag{41}$$

where

$$u(p^{\alpha_{p,q}}) := \begin{cases} \phi(p^{\alpha_{p,q}}), & \text{if } p \not| q-1; \\ p^{\alpha_{p,q}}, & \text{if } p | q-1. \end{cases} \tag{42}$$

It is easy to see, say by Euler's Theorem together with the standard divisibility properties of the Lucas sequences, that if the prime factorization of m_q is given by (40), then the number $l(q)$ constructed at (41) and (42) satisfies i above, and clearly $l(q) \le m_q < z$. Thus, there exists a number $l(q)$ satisfying both i and ii above. We also set $l(2)$ to be any number in the interval $(z/2, z)$.

We write:

$$n_0 := \prod_{2 \le q < y} q^{l(q)-1}. \tag{43}$$

We also write $t := \lfloor z \rfloor$, and consider numbers n of the form

$$n := n_0 \cdot \prod_{2 < q < y} M_q, \tag{44}$$

where each one of the numbers M_q is squarefree, has precisely t prime factors, and all of them belong to $\mathcal{P}_q''$. From now on, the proof of Theorem 3 proceeds in the following way. We first show that if x is sufficiently large with respect to ε and δ, then every number of the form (44) belongs to R. We next find an upper bound, call it T, for all the numbers of the form (44). Finally, we find a lower bound for the count of all the numbers of the form (44), and we complete the proof of the theorem.

4.1. The numbers shown at (44) are in R.

It is clear that $n \in R$ if and only if $\lambda(n)|\sigma(n)$, where $\lambda(n)$ stands for the maximal order of elements in the multiplicative group modulo n. The function $\lambda(n)$, which is also called the Carmichael function, has the property that if $n := \prod_{p^{\alpha_p} || n} p^{\alpha_p}$ then $\lambda(n) = \mathrm{lcm}[\lambda(p^{\alpha_p}) \mid p^{\alpha_p} || n]$, where $\lambda(p^{\alpha_p}) = p^{\alpha_p - 1}(p-1)$ holds whenever $p > 2$ or $\alpha_p \le 2$, and $\lambda(2^{\alpha_2}) = 2^{\alpha_2 - 2}$ holds when $\alpha_2 \ge 3$.

We now compute $\lambda(n)$ where n is as shown in (44). Since $\mathcal{P}_q''$ are disjoint and free of primes $p < y$, we get

$$\begin{aligned} \lambda(n) &\Big| \mathrm{lcm}[\phi(n_0),\ \phi(M_q) \mid q \in \mathcal{P}_q'] \\ &\Big| \mathrm{lcm}\Big[q^{l(q)-2}(q-1),\ m_q,\ \Big(\frac{p_q - 1}{m_q}\Big) \mid 2 < q < y,\ p_q \in \mathcal{P}_q''\Big]. \end{aligned} \tag{45}$$

We claim that (45) implies that

$$\lambda(n) | n_0 \cdot \mathrm{lcm}[m_q \mid 2 < q < y]. \tag{46}$$

Indeed, the point is that if $p_q \in \mathcal{P}_q''$, then $(p_q-1)/m_q$ is a number $d < x^\alpha$ whose largest prime factor is at most y. Thus, the power at which any fixed prime $q < y$ can appear in the factorization of $(p_q-1)/m_q$ is at most $\log(x^\alpha)/\log 2 = \alpha \log x/\log 2 < z/2 = x^{1-\delta}/2$, with the last inequality holding for x sufficiently large with respect to δ and ε. In particular, the exponent at which every fixed prime can appear in the prime factor factorization of $(p_q-1)/m_q$ is smaller than $l(q)$. Moreover, since $q < y$, it follows that all prime factors of $q-1$ are also smaller than y, and the same argument as above shows that if $q_1 | q-1$, then $q_1 \neq q$ and the exponent at which q_1 can appear in the prime factor factorization of $q-1$ is less than $l(q_1)$. These remarks show that (45) implies (46).

We now show that the number appearing in the right–hand side of (46) is a divisor of $\sigma(n)$, where n is given by (44). Clearly, since $\mathcal{P}_q''$ are disjoint and free of primes $q < y$, it follows that

$$\sigma(n) = \sigma(n_0) \prod_{2<q<y} \sigma(M_q) = \prod_{2\le q<y} \frac{q^{l(q)}-1}{q-1} \cdot \prod_{2<q<y} \prod_{p|M_q} (p+1). \tag{47}$$

From the way we have chosen the numbers l_q, it is clear that the first product appearing in (47) is a multiple of m_q for all odd $q < y$. Thus, it suffices to show that $q^{l(q)}$ divides the number appearing in the right–hand side of (47) for all $q < y$. Fix such an odd prime q. Since M_q contains $t = \lfloor z \rfloor \ge l(q)$ prime factors in $\mathcal{P}_q''$ and all such are congruent to -1 modulo q, it follows that q^t divides the number appearing in the right–hand side of (47). In particular, $q^{l(q)}$ divides this number. This is true for all odd primes $q < y$. For $q = 2$, this is also true because all primes p dividing M_{q_1} for some odd $q_1 < y$ are odd (in fact, they are larger than y, and so of course larger than 2), so 2^t divides the number appearing in the right–hand side of (47) as well. Thus, every number n given by (44) is indeed in R.

4.2. An upper bound for the size of the numbers shown at (44)

Note that

$$n_0 < \exp\Big(z \sum_{q<y} \log q\Big) = \exp\Big(zy(1+o(1))\Big) = \exp(x^{1-\delta+\varepsilon}(1+o(1)), \tag{48}$$

with the estimate (48) following from the Prime Number Theorem.

Moreover, if $p \in \mathcal{P}_q''$ for some odd $q < y$, then $p-1 = dm_q$ for some $d < x^\alpha$ and $m_q < z$, therefore $p < x^\alpha z = x^{\alpha+1-\delta}$. Thus,

$$\prod_{2<q<y} M_q < \exp\Big((\pi(y)-1)t \log(x^{\alpha+1-\delta})\Big) \tag{49}$$

$$= \exp\Big(\frac{y}{\log y}\, z\,(\alpha+1-\delta)\log x \cdot (1+o(1))\Big) = \exp\Big(\frac{\alpha+1-\delta}{\varepsilon} x^{1-\delta+\varepsilon}(1+o(1))\Big),$$

with estimate (49) following by the Prime Number Theorem. With (48) and (49), we get

$$n < \exp\Big(\frac{(\alpha+1-\delta+\varepsilon)}{\varepsilon}\cdot x^{1-\delta+\varepsilon}(1+o(1))\Big) \tag{50}$$
$$< \exp\Big(\frac{(\alpha+1-\delta+\varepsilon)(1+\varepsilon)}{\varepsilon}\cdot x^{1-\delta+\varepsilon}\Big) := T,$$

with the right most inequality in (50) holding for all x sufficiently large with respect to δ and ε.

4.3. A lower bound for the number of numbers of the form (44)

While we did not say this up to now, it is implicit that at least one number of the form (44) exists only when $\delta < 1/2$. Indeed, in order to be able to construct a number M_q, we will definitely need that $\#\mathcal{P}_q'' > t$ holds. We show that $x^{\alpha-3\varepsilon} > z \geq t$ holds with ε sufficiently small with respect to δ when $\delta < 1/2$. Indeed, in order for the above inequality to hold, it suffices that $\alpha - 3\varepsilon > 1-\delta$ holds. With the formula for α, this last inequality is equivalent to $((1-\delta)^2-\varepsilon)/\delta - 3\varepsilon > 1-\delta$, which is equivalent to $(1-\delta)(1-2\delta) > (3\delta+1)\varepsilon$, and this inequality does hold when $\delta < 1/2$, and ε is chosen so that

$$\varepsilon < \frac{(1-\delta)(1-2\delta)}{3\delta+1}. \tag{51}$$

In particular, with such ε, and since $\#\mathcal{P}_q'' > 2x^{\alpha-3\varepsilon}$, we get that $\#\mathcal{P}_q'' - t > x^{\alpha-3\varepsilon} > t$. We now note that the number of possibilities of choosing numbers of the form (44) is precisely

$$\prod_{2<q<y}\binom{\#\mathcal{P}_q''}{t} > \prod_{2<q<y}\Big(\frac{\#\mathcal{P}_q''-t}{t}\Big)^t > \Big(\frac{x^{\alpha-3\varepsilon}}{z}\Big)^{t(\pi(y)-1)} \tag{52}$$
$$= (x^{\alpha-(1-\delta)-3\varepsilon})^{t(\pi(y)-1)} = \exp\Big(t(\pi(y)-1)\log(x^{\alpha-(1-\delta)-3\varepsilon})\Big)$$
$$= \exp\Big(\frac{zy}{\log y}\cdot(\alpha-(1-\delta)-3\varepsilon)\log x\cdot(1+o(1))\Big)$$
$$= \exp\Big(\frac{(\alpha-(1-\delta)-3\varepsilon)}{\varepsilon}\cdot x^{1-\delta+\varepsilon}(1+o(1))\Big).$$

In particular, for large x, the number of numbers of the form (44) is at least

$$\exp\Big(\frac{(\alpha-(1-\delta)-3\varepsilon)(1-\varepsilon)}{\varepsilon}\cdot x^{1-\delta+\varepsilon}\Big) = T^\beta,$$

where the exponent β is given by the formula

$$\beta := \frac{(\alpha-(1-\delta)-3\varepsilon)(1-\varepsilon)}{(\alpha+(1-\delta)+\varepsilon)(1+\varepsilon)}. \tag{53}$$

The limit of β as a function of ε when $\varepsilon \to 0$ is

$$\frac{(1-\delta)^2/\delta-(1-\delta)}{(1-\delta)^2/\delta+1-\delta} = \frac{(1-\delta)(1-2\delta)}{(1-\delta)} = 1-2\delta.$$

In particular, if $\varepsilon_1 > 0$ is arbitrary, there for every $\varepsilon > 0$ and bounded above by some function depending on both δ and ε_1 has the property that both the inequality (51) and the inequality $\beta > 1 - 2\delta - \varepsilon_1$ are satisfied with such a value of ε. Replacing now ε_1 with ε, we get the assertion of Theorem 3.

Acknowledgements. The work was partly supported by grants SEP–CONACYT 37259–E, SEP–CONACYT 37260–E and PAPIIT IN104602 from the UNAM.

References

[1] Alford, W. R., Granville, A., Pomerance, C., There are infinitely many Carmichael numbers. Ann. of Math. 139(3) (1994), 703–722.

[2] Bennett, M. A., Walsh, P. G., The Diophantine equation $b^2X^4 - dY^2 = 1$. Proc. Amer. Math. Soc. 127(12) (1999), 3481–3491.

[3] Bilu, Y., Hanrot, G., Voutier, P. M., Existence of primitive divisors of Lucas and Lehmer numbers. With an appendix by M. Mignotte. J. Reine Angew. Math. 539 (2001), 75–122.

[4] Bombieri, E., Friedlander, J. B., Iwaniec, H., Primes in arithmetic progressions to large moduli. Acta Math. 156(3–4) (1986), 203–251.

[5] Carmichael, R. D., On the numerical factors of the arithmetic forms $\alpha^n \pm \beta^n$. Ann. Math. 15 (1913), 30–70.

[6] Granville, A., Pomerance, C., Two contradictory conjectures concerning Carmichael numbers. Math. Comp. 71(238) (2002), 883–908.

[7] Křížek, M., Luca, F., Somer, L., On the Convergence of Series of Reciprocals of Primes Related to the Fermat Numbers. J. Number Theory 97(1) (2002), 95–112.

[8] Luca, F., Pomerance, C., On the average value of $\tau(\phi(n))$. Preprint, 2003.

[9] Pomerance, C., Two methods in elementary analytic number theory. In: Number theory and applications (Banff, AB, 1988), pp. 135–161, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., 265, Kluwer Acad. Publ., Dordrecht, 1989.

[10] Rosser, J. B., Schoenfeld, L., Approximate formulas for some functions of prime numbers. Ill. J. of Math. 6 (1962), 64–94.

[11] Rotkiewicz, A., Pseudoprime numbers and their generalizations. Student Association of Faculty of Sciences, Univ. of Novi Sad, 1972.

[12] Rotkiewicz, A., Solved and unsolved problems on pseudoprime numbers and their generalizations. In: Applications of Fibonacci numbers. Vol. 8 (Rochester, NY, 1998), pp. 293–306, Kluwer Acad. Publ., Dordrecht, 1999.

[13] Sierpiński, W., Elementary Theory of Numbers. (A. Schinzel, ed.) Amsterdam: North–Holland 1987.

[14] Tenenbaum, G., Introduction to analytic and probabilistic number theory. Cambridge: Cambridge University Press 1995.

Received by the editors September 25, 2003