On the largest prime factor of a Mersenne number

Leo Murata and Carl Pomerance

1. Introduction

Over two millennia ago Euclid demonstrated that a prime p of the form $2^n - 1$ gives rise to the perfect number $2^{n-1}p$, and he found four such primes. Presumably Euclid also knew that if $2^n - 1$ is prime, then so is n prime, and that the converse does not always hold. In the 18th century, Euler showed that Euclid's formula for perfect numbers gives rise to all even examples. (It is conjectured that there are no odd perfect numbers.) Probably because of the connection to perfect numbers, many mathematicians over the ages have been interested in finding primes of the form $2^n - 1$. In the early 17th century, the French monk and mathematician Marin Mersenne made an uncanny guess of which n in the range $29 \le n \le 257$ give such primes. He guessed that it was $n = 31$, 67, 127, and 257, while the truth is $n = 31$, 61, 89, 107, and 127. What is uncanny is that he got two right, and that there are indeed so few of them. In fact the sparsity of the Mersenne primes is hardly suggested by the evidence below 31: Seven of ten candidates are prime. In honor of Mersenne and his guess, primes of the form $2^n - 1$ are now called Mersenne primes. Currently (June, 2003) we know 39 Mersenne primes, the largest having exponent $n = 13466917$, and rumors of a 40-th are circulating. Nevertheless, it has never been proved that there are infinitely many Mersenne primes, nor has it been proved that there are infinitely many prime numbers n with $2^n - 1$ composite.

We shall call any number of the form $2^n - 1$ a *Mersenne number*, regardless if n is prime. For an integer $m > 1$, let $P(m)$ denote the largest prime factor of m, and let $P(1) = 1$. We are interested in this paper in studying $P(2^n - 1)$.

First, we shall review what is known about $P(2^n - 1)$. Schinzel [11] has shown that $P(2^n - 1) \ge 2n + 1$ for all $n > 12$. Remarkably, this result still stands as the best lower bound proved for $P(2^n - 1)$ for all sufficiently large n. It is perhaps reasonable to conjecture that $P(2^n - 1) > n^K$ for every fixed K and all sufficiently large n depending on the choice of K, or maybe even $P(2^n - 1) > 2^{n/\log n}$ for sufficiently large n, but clearly we are very far from proving such assertions.

It is natural, in light of Euclid, to consider the special case of $P(2^p - 1)$ with p prime. Here we have somewhat better results: Stewart [12] showed that $P(2^p - 1) > \frac{1}{2}p(\log p)^{1/4}$ for all primes p beyond an effectively computable constant. This was improved by Stewart [13], and independently by Erdős and Shorey [3], to $P(2^p - 1) > c_1 p \log p$ for all primes $p > c_2$, where c_1, c_2 are effectively computable.

In addition, people have studied $P(2^n - 1)$ ignoring a set of numbers n of asymptotic density 0. For example, consider that $2^n - 1$ is always divisible by a prime that is congruent to 1 mod n, when $n > 1$. Thus, those values of n where the least prime $p \equiv 1 \pmod n$ is large give rise to values of n where $P(2^n - 1)$ is large. In particular, it follows easily from the Brun–Titchmarsh inequality that if $\epsilon(n) > 0$ tends to zero monitonically and arbitrarily slowly, then the least prime $p \equiv 1 \pmod{n}$ satisfies $p > \epsilon(n) n \log n$ on a set of asymptotic density 1. (The proof: For a positive integer $j \le \epsilon(x) \log x$, the number of positive integers n with $n \le x$ and $jn + 1$ prime is $O(jx/\varphi(j) \log x)$, by the Brun–Titchmarsh inequality. Summing on j, we involve only $O(\epsilon(x)x)$ numbers n, and ignoring these, we have for the other numbers $n \le x$ and any prime $p \equiv 1 \pmod{n}$ that $(p-1)/j > \epsilon(x) \log x \ge \epsilon(n) \log n$.) Thus, ignoring a set of integers n of asymptotic density 0, we have $P(2^n - 1) > \epsilon(n) n \log n$. Improving on this, Stewart [13], showed that

$$P(2^n - 1) \ > \ \epsilon(n) n (\log n)^2 / \log\log n \tag{1}$$

ignoring a set of integers n of asymptotic density 0. Earlier, Erdős and Shorey [3] had shown that

$$P(2^p - 1) \ > \ cp(\log p)^2/(\log\log p \log\log\log p) \tag{2}$$

but for $o(\pi(x))$ primes $p \le x$.

Most of these results are proved using theorems about linear forms in logarithms of algebraic numbers, in particular estimates of Baker, and subsequent improvements. In a recent paper, Murty and Wong [8] replace this deep tool with an even deeper unproved hypothesis, namely the ABC conjecture. This conjecture asserts that for each $\epsilon > 0$, there are at most finitely many coprime triples a, b, c of positive integers with $a + b = c$ and $c > m^{1+\epsilon}$, where m is the product of the distinct primes that divide abc. Assuming this hypothesis, Murty and Wong prove that

$$P(2^n - 1) \ > \ n^{2-\epsilon} \tag{3}$$

for each fixed $\epsilon > 0$ and all sufficiently large values of n, depending on the choice of ϵ. In the case when $n = p$, a prime, this result is easy to see: If m is the largest squarefree divisor of $2^p - 1$, then the ABC conjecture applied with $a = 2^p - 1, b = 1, c = 2^p$ implies that $2^p < (2m)^{1+\epsilon}$ for all sufficiently large primes p. But all the prime factors of m are 1 mod p, so if they are all smaller than $p^{2-\epsilon}$, then there are at most $p^{1-\epsilon}$ such primes, and so $m < p^{(2-\epsilon)p^{1-\epsilon}}$. Hence, assuming $\epsilon < 1$,

$$2^p < (2m)^{1+\epsilon} \ < \ 4p^{3p^{1-\epsilon}} \ = \ o(2^p),$$

a contradiction. In fact, a slight elaboration of this argument using the Brun–Titchmarsh inequality shows that $P(2^p - 1) > cp^2$ for some effectively computable positive constant c and all sufficiently large primes p.

It is our goal in this paper to also give some conditional results about $P(2^n - 1)$, but the condition we shall assume is not the ABC conjecture, but rather the Generalized Riemann Hypothesis (GRH) for various Kummerian fields of the form $\mathbf{Q}(2^{1/j}, e^{2\pi i/jk})$. In addition, we strongly use some results in Stewart [13] and Erdős and Shorey [3] (for our first 3 theorems) which depend on estimates for linear forms in logarithms of algebraic numbers. We shall prove the following four theorems:

Theorem 1. (GRH) *The set of natural numbers n with*

$$P(2^n - 1) \ > \ n^{4/3}/\log\log n$$

has asymptotic density 1*.*

Theorem 2. (GRH) *But for $o(\pi(x))$ primes $p \le x$ we have*

$$P(2^p - 1) \ > \ p^{4/3}/(\log p)^{2/3}\log\log p.$$

Let $\Phi_n(x)$ denote the nth cyclotomic polynomial.

Theorem 3. (GRH) *Uniformly for each $\epsilon > 0$ and $x \ge 2$, but for $O(\epsilon x + x/(\log x)^2)$ integers $n \le x$, every prime factor of $\Phi_n(2)$ exceeds $n^{1+\epsilon}$. Further, but for $O(\epsilon\pi(x))$ primes $p \le x$, every prime factor of $2^p - 1$ exceeds $p^{1+\epsilon}$.*

Theorem 4. (GRH) *Let $M(x)$ denote the number of primes $p \le x$ which divide some $2^q - 1$ with q prime. That is, $M(x)$ is the number of odd primes $p \le x$ such that the order of $2 \bmod p$ is a prime number. Then $M(x) = O(x/(\log x)^2)$.*

It should be noted that the conclusions of Theorems 1 and 2 are considerably weaker than the conclusion (3) of the Murty–Wong theorem. In particular, not only is their inequality stronger, we allow density-0 exceptions, while they allow only finitely many exceptions. However, they assume the ABC conjecture, while we assume the GRH. Of course, since neither conjecture is presently a theorem, it is unclear which is the stronger hypothesis.

Theorem 3 appears to be new with no known ABC analog. We mention that Theorem 4 is a strengthening of the GRH-conditional estimate $M(x) = O(x \log\log x/(\log x)^2)$ achieved by the second author in [10]. In the same paper, an unconditional estimate of $M(x) = O(\pi(x)\log\log\log x/\log\log x)$ is proved. We give below a heuristic argument that Theorem 4 is best possible, and we even suggest a candidate number c for which it may be true that $M(x) \sim cx/(\log x)^2$.

We remark that it is likely that our results can be generalized to the largest prime factor of expressions such as $a^n - b^n$, $(\alpha^n - \beta^n)/(\alpha - \beta)$ where α, β are conjugate algebraic numbers of degree 2, and to similar sequences, as was done in most of the referenced papers. We leave these details to another time, and perhaps to another person.

Acknowledgments. We thank Ram Murty, Tarlok Shorey, and Cam Stewart for their comments and for their interest in this note. We also acknowledge some helpful comments from an anonymous referee. The first author would like to thank the Number Theory Foundation and the hospitality of Bell Laboratories, and the second author would like to thank the hospitality of RIMS in Kyoto, Japan.

2. Preliminaries

In the following we often use the Vinogradov order notation $\ll$, which means the same as the big-O notation. That is, for positive functions $f(x), g(x)$ we write $f(x) \ll g(x)$ if and only if $f(x) = O(g(x))$. In addition, we write $f(x) \gg g(x)$ if and only if $g(x) = O(f(x))$. In what follows, all implied constants are absolute.

For an odd integer $m > 0$, let $l(m)$ denote the least positive integer l with $2^l \equiv 1 \pmod{m}$. We say a prime factor p of $\Phi_n(2)$ is *primitive* if $l(p) = n$, and otherwise we say that p is *intrinsic*.

Lemma 1. *The prime p is an intrinsic prime factor of $\Phi_n(2)$ if and only if $n = p^j l(p)$ for some integer $j > 0$. For each integer $n \neq 1, 6$ there is at least one primitive prime factor of $\Phi_n(2)$.*

These results are well-known. A reference for the first is Nagell [9], and the second, known as Bang's theorem, has been independently proved multiple times, an early English language reference being [1]. Note that as a corollary of the first assertion, an intrinsic prime factor of $\Phi_n(2)$ must be the largest prime factor of n.

Lemma 2. *Let j and k be natural numbers, and let $n_{j,k} = [\mathbf{Q}(2^{1/j}, e^{2\pi i/k}) : \mathbf{Q}]$. Then we have*

$$n_{j,k} = \begin{cases} \frac{1}{2} j\varphi(k), & \textit{if } 8|k \\ \\ j\varphi(k), & \textit{otherwise.} \end{cases}$$

Proof. Since $[\mathbf{Q}(2^{1/j}) : \mathbf{Q}] = j$, $[\mathbf{Q}(e^{2\pi i/k}) : \mathbf{Q}] = \varphi(k)$, and

$$n_{j,k} = \frac{[\mathbf{Q}(2^{1/j}) : \mathbf{Q}][\mathbf{Q}(e^{2\pi i/k}) : \mathbf{Q}]}{[\mathbf{Q}(2^{1/j}) \cap \mathbf{Q}(e^{2\pi i/k}) : \mathbf{Q}]},$$

we need only to compute $[\mathbf{Q}(2^{1/j}) \cap \mathbf{Q}(e^{2\pi i/k}) : \mathbf{Q}]$. We remark here that the maximal normal subfield which is contained in $\mathbf{Q}(2^{1/j})$ is $\mathbf{Q}$ if j is odd, and is $\mathbf{Q}(\sqrt{2})$ if j is even. Further, any cyclotomic field which contains $\mathbf{Q}(\sqrt{2})$ also contains $\mathbf{Q}(e^{2\pi i/8})$. Since the field $\mathbf{Q}(2^{1/j}) \cap \mathbf{Q}(e^{2\pi i/k})$ is a normal extension of $\mathbf{Q}$, it is equal to $\mathbf{Q}$ or $\mathbf{Q}(\sqrt{2})$, and the latter happens if and only if $\mathbf{Q}(\sqrt{2}) \subset \mathbf{Q}(e^{2\pi i/k})$ if and only if $\mathbf{Q}(e^{2\pi i/8}) \subset \mathbf{Q}(e^{2\pi i/k})$ if and only if $8|k$. This completes the proof of the lemma.

Let $\tau(n)$ denote the number of positive divisors of n, and let $\nu(n)$ denote the number of these divisors which are prime. In [3, Lemma 2], Erdős and Shorey showed that there is an effectively computable positive constant c_0 such that if p is an odd prime, then $P(2^p - 1) < p^2$ implies that $\nu(2^p - 1) > c_0 \log p / \log\log p$. We now consider an analogous result for arbitrary integers n, but here we must allow for some exceptions, and the conclusion is slightly weaker.

Lemma 3. *Let $\epsilon(n)$ be a positive function that tends to 0 arbitrarily slowly. There is a set of natural numbers of asymptotic density 1 such that if n is in this set and $P(\Phi_n(2)) < n^2$ then $\nu(\Phi_n(2)) > \epsilon(n) \log n / \log\log n$.*

Proof. We sketch the proof following the argument given in Stewart [13]. We may assume that $\epsilon(n)$ tends to 0 monotonically. Let $\mathcal{S}$ denote the set of natural numbers n with $\tau(n) < \log n$ and n has two consecutive divisors $d_0 < d_1$ with $n/d_0, n/d_1$ both squarefree and $d_1/d_0 > n^{\sqrt{\epsilon(n)}}$. Since the normal value of $\tau(n)$ is $(\log n)^{\log 2 + o(1)}$, the condition that $\tau(n) < \log n$ removes only a density-0 set of numbers n that otherwise would be in $\mathcal{S}$. That $\mathcal{S}$ has asymptotic density 1 now follows from Lemma 11 in [13] plus a short additional argument based on the easy fact that most numbers n are nearly squarefree, that is, they are not divisible by a large square.

Suppose $n \in \mathcal{S}$ and write

$$R_0 = \prod_{d|n,\, d\le d_0} (1-2^{-d})^{\mu(n/d)}, \quad R_1 = \prod_{d|n,\, d\ge d_1} (1-2^{-d})^{\mu(n/d)}.$$

Note that

$$\Phi_n(2) = 2^{\varphi(n)} R_0 R_1$$

and that

$$|\log R_1| \le -\sum_{d|n,\, d\ge d_1} \log(1-2^{-d}) < 2\sum_{d\ge d_1} 2^{-d} = 2^{2-d_1}.$$

Also note that from the second assertion of Lemma 1, it is easy to see that the rational number R_1 is not 1. It is clear that the rational number R_0 has height at most $2^{d_0\tau(n)} < 2^{d_0 \log n}$.

Let $\nu(\Phi_n(2)) = k$. We apply Lemma 9 in [13] (linear forms in logarithms of algebraic numbers) to the identity $R_1 = 2^{-\varphi(n)}\Phi_n(2)R_0^{-1}$. With this result and the assumption that $P(\Phi_n(2)) < n^2$, there is an absolute constant $c > 0$ such that

$$\log|\log R_1| \ge -(2k^c \log n)^{k+2} d_0.$$

But we above argued that

$$\log|\log R_1| < -d_1 \log 2 + 2\log 2.$$

Putting these last two assertions together with $d_1/d_0 > n^{\sqrt{\epsilon(n)}}$, we have

$$k \gg \sqrt{\epsilon(n)} \log n / \log\log n,$$

from which the lemma immediately follows.

Remark. The condition $P(\Phi_n(2)) < n^2$ may be replaced with $P(\Phi_n(2)) \le n^{O(1)}$.

Lemma 4. (GRH) *For all natural numbers j, d and all $x \ge 2$, let $N_{j,d}(x)$ denote the number of primes $p \le jx$ which split completely in the field $\mathbf{Q}(2^{1/j}, e^{2\pi i/jd})$. Then, using the notation of Lemma 2, we have uniformly*

$$N_{j,d}(x) = \frac{1}{n_{j,jd}} \mathrm{li}(jx) + O\big((jx)^{1/2}\log(jx) + j^2 d \log(jd)\big). \tag{4}$$

This result follows from Theorem 1.1 in Lagarias and Odlyzko [7] and a calculation for the discriminant of the field $\mathbf{Q}(2^{1/j}, e^{2\pi i/jd})$. One might also compare this theorem with (28) in Hooley [6].

Lemma 5. (GRH) *Fix a positive integer j. Let $M_j(x)$ denote the number of primes $p \le jx$ with $p \equiv 1 \pmod{j}$, $2^{(p-1)/j} \equiv 1 \pmod{p}$, and $(p-1)/j$ is prime. Then uniformly for $j \le x^{1/3}/2$,*

$$M_j(x) \ \ll \ \frac{j}{\varphi(j)^2} \frac{x}{\log x \, \log\left(x^{1/3}/j\right)} \ + \ \frac{x^{5/6}}{j^{1/2}} \big(\log(x^{1/3}/j)\big)^2 \log x.$$

Proof. We use the form of Selberg's sieve as presented in Theorem 5.2, page 153 of Halberstam and Richert [4]. Let $z = \left(x^{1/3}/j\right)^{1/2}$ and let $\mathcal{A}$ denote the set of numbers $(p-1)/j$ where $p \le jx$ is prime and $2^{(p-1)/j} \equiv 1 \pmod{p}$. Further, let $\mathcal{P}$ denote the set of primes that do not divide j. Then clearly,

$$M_j(x) \ \le \ \mathcal{S}(\mathcal{A}, \mathcal{P}, z) + \pi(z),$$

where $\mathcal{S}(\mathcal{A}, \mathcal{P}, z)$ denotes the number of members of $\mathcal{A}$ which are not divisible by any member of $\mathcal{P} \cap [1, z]$. Note that from (4) we have $X := |\mathcal{A}| = N_{j,1}(x)$ well estimated. Suppose d is squarefree, relatively prime to j, and $d \le z^2$. To apply the sieve we need to compute $\omega(d)$, and to estimate $R_d := |\mathcal{A}_d| - (\omega(d)/d)|\mathcal{A}|$, where $\mathcal{A}_d$ is the set of members of $\mathcal{A}$ which are divisible by d.

Note that we may assume that j is even, since the only prime of the form $(p-1)/j$ with j odd is 2. For a prime $p \le jx$, $(p-1)/j$ is in $\mathcal{A}_d$ if and only if p splits completely in the field $\mathbf{Q}\left(2^{1/j}, e^{2\pi i/jd}\right)$. Hence $|\mathcal{A}_d| = N_{j,d}(x)$. Since d is coprime to the even number j, Lemma 2 implies that $n_{j,jd} = \varphi(d) n_{j,j}$. Further, by Lemma 4, and taking $\omega(d) = d/\varphi(d)$, we have

$$\begin{aligned} |\mathcal{A}_d| \ &= \ \frac{1}{\varphi(d) n_{j,j}} \mathrm{li}(jx) + O\left((jx)^{1/2} \log(jdx) + j^2 d \log(jd)\right) \\ &= \ \frac{\omega(d)}{d} X + O\left((jx)^{1/2} \log x + j^2 d \log x\right). \end{aligned}$$

Note that $j^2 d \le j^2 z^2 = j x^{1/3} \le (jx)^{1/2}$ in our range for j. Thus we may take $R_d \ll (jx)^{1/2} \log x$ for each d, so that by the sieve theorem mentioned, we have

$$\begin{aligned} \mathcal{S}(\mathcal{A}, \mathcal{P}, z) \ &\ll \ \frac{j}{\varphi(j)} \frac{X}{\log z} + \sum_{d \le z^2} \mu^2(d) 3^{\nu(d)} (jx)^{1/2} \log x \\ &\ll \ \frac{j}{\varphi(j)} \frac{X}{\log z} + z^2 (\log z)^2 (jx)^{1/2} \log x. \end{aligned}$$

The lemma now follows from an estimate for X afforded by Lemma 4, and the choice of z.

Remark. The proof supports the same inequality for $M_j'(x)$, which is defined exactly as $M_j(x)$, but with the condition that $(p-1)/j$ is prime being replaced with the condition that $(p-1)/j$ has each of its prime factors exceeding $(x^{1/3}/j)^{1/2}$.

3. Proofs of the theorems

Proof of Theorem 1. By Lemma 4, if $J \le x$, we have

$$\begin{aligned}\sum_{j\le J} N_{j,1}(x) &\ll \sum_{j\le J}\left(\frac{1}{\varphi(j)}\frac{x}{\log x} + (Jx)^{1/2}\log x + J^2\log J\right)\\ &\ll \frac{x\log J}{\log x} + J^{3/2}x^{1/2}\log x + J^3\log x.\end{aligned}$$

Applying this result with $J = x^{1/3}/\log\log x$, we get that

$$\sum_{j\le x^{1/3}/\log\log x} N_{j,1}(x) \ll x\log x/(\log\log x)^{3/2}. \tag{5}$$

Let $\epsilon(x)$ decrease to 0 arbitrarily slowly. It follows from Lemma 3 that but for $o(x)$ choices of $n \in (x/2, x]$, if $P(2^n - 1) \le n^{4/3}/\log\log n$, then

$$\nu(\Phi_n(2)) > 1 + \epsilon(x)\log x/\log\log x. \tag{6}$$

Say there are E numbers $n \in (x/2, x]$ with $P(\Phi_n(2)) \le n^{4/3}/\log\log n$ and (6) holding. Consider such a number n, and let the distinct *primitive* prime factors of $\Phi_n(2)$ be denoted $p_1, \dots, p_k$ where, by Lemma 1 and (6), we have $k > \epsilon(x)\log x/\log\log x$. Then each $p_i = j_i n + 1$ for some $j_i \le J$ and $2^{(p_i-1)/j_i} \equiv 1 \pmod{p_i}$. That is, p_i is counted by $N_{j_i,1}(x)$. Since these primitive prime factors of $\Phi_n(2)$ are not primitive prime factors of any other $\Phi_{n'}(2)$ it follows that

$$E\,\epsilon(x)\log x/\log\log x \le \sum_{j\le J} N_{j,1}(x),$$

so that from (5),

$$E \ll \frac{x}{\epsilon(x)(\log\log x)^{1/2}}.$$

Since we may take $\epsilon(x) = 1/\log\log\log x$, say, the theorem follows.

Remark. The proof allows a slightly stronger result: For any positive function $\epsilon(n)$ which decreases to 0, the set of natural numbers n such that $P(\Phi_n(2)) > \epsilon(n)n^{4/3}/(\log\log n)^{2/3}$ has asymptotic density 1.

Proof of Theorem 2. Let E denote the number of primes $p \in (x/2, x]$ with

$$P(2^p - 1) \le p^{4/3}/(\log p)^{2/3}\log\log p.$$

As remarked before Lemma 3, it follows from [3] that for such a prime p, $\nu(2^p - 1) \gg \log x/\log\log x$. Let $p_1, \dots, p_k$ be the distinct prime factors of $2^p - 1$ for one of these exceptional primes p, so that $k \gg \log x/\log\log x$. Then each $p_i = j_i p + 1$ for some integer

$j_i \le J := x^{1/3}/(\log x)^{2/3}\log\log x$. Hence $p = (p_i - 1)/j_i$ is counted by the expression $M_{j_i}(x)$ defined in Lemma 5, so that

$$E \,\log x/\log\log x \;\ll\; \sum_{j\le J} M_j(x). \tag{7}$$

We have, by Lemma 5, that

$$\sum_{j\le x^{1/3}/(\log x)^3} M_j(x) \;\ll\; \frac{x\log\log x}{\log x} + \frac{x(\log\log x)^2}{(\log x)^{1/2}},$$

and by (4), that

$$\sum_{x^{1/3}/(\log x)^3<j\le J} M_j(x) \;\le\; \sum_{x^{1/3}/(\log x)^3<j\le J} N_{j,1}(x) \;\ll\; \frac{x\log\log x}{\log x} + \frac{x}{(\log\log x)^{3/2}}.$$

Hence,

$$\sum_{j\le J} M_j(x) \;\ll\; \frac{x}{(\log\log x)^{3/2}},$$

so that (7) implies that $E = o(\pi(x))$. This completes the proof of the theorem.

Remark. The proof gives a slightly stronger result: For almost all primes p (that is, but for a set of primes with relative density 0 in the set of all primes) we have $P(2^p - 1) > \epsilon(p)p^{4/3}/(\log p\log\log\log p)^{2/3}$.

Proof of Theorem 3. By Lemma 1, $\Phi_n(2)$ has an intrinsic prime factor p if and only if $n = p^a l(p)$ for some positive integer a. Since $l(p) \gg \log p$ we have that the number of integers $n \le x$ of the form $p^a l(p)$ for some prime p is $O(x/(\log x)^2)$. Hence it suffices to consider only those numbers n where every prime factor of $\Phi_n(2)$ is primitive. In particular, it is sufficient to show that but for $O(\epsilon x)$ integers $n \le x$, every *primitive* prime factor of $\Phi_n(2)$ exceeds $n^{1+\epsilon}$.

We may assume that $\epsilon < 1/4$. If $n \le x$ is such that $\Phi_n(2)$ has a primitive prime factor p with $p \le n^{1+\epsilon}$, then $p = jn + 1$ for some integer $j < p/n \le n^\epsilon \le x^\epsilon$, and $2^{(p-1)/j} \equiv 1 \pmod p$. Thus, the number of such numbers $n \le x$ is at most, using (4),

$$\sum_{j\le x^\epsilon} N_{j,1}(x) \;\ll\; \epsilon x.$$

The proof for $2^p - 1$, with p prime, is similar. In this case every prime factor of $2^p - 1 = \Phi_p(2)$ is primitive. If $p \le x$ and $q|2^p - 1$ is prime with $q \le p^{1+\epsilon}$, then p is counted by some $M_j(x)$ (the notation coming from Lemma 5), with $j \le \epsilon$. The theorem now follows since by Lemma 5, the number of such primes p is at most

$$\sum_{j\le x^\epsilon} M_j(x) \;\ll\; \epsilon x/\log x.$$

Proof of Theorem 4. Suppose q is prime and $\log x < q \le x^{1/5}$. If $l(p)$ is prime and $q|p-1$, then either $l(p) = q$ or $2^{(p-1)/q} \equiv 1 \pmod{p}$. The number of primes p with $l(p) = q$ is trivially bounded by q (since their product divides $2^q - 1$), and the number of primes $p \le x$ in the second category is $\ll x/q^2 \log x$, using Lemma 4, with "x, j, d" of that lemma replaced with $x/q, q, 1$, respectively. Thus, the number of primes $p \le x$ with $l(p)$ prime and $q|p-1$ for some prime q in $(\log x, x^{1/5}]$, is $o(x/(\log x)^2)$. Hence we may consider only primes $p \le x$ with $p-1$ not divisible by any prime in $(\log x, x^{1/5}]$. Write $p-1 = mr$, where every prime factor of m is at most $\log x$ and every prime factor of r is $> x^{1/5}$. For $m \le x^{1/5}$, the number of primes $p \le x$, with $p \equiv 1 \pmod{m}$, $2^{(p-1)/m} \equiv 1 \pmod{p}$, and $(p-1)/m$ divisible only by primes bigger than $x^{1/5}$, is $\ll x/\varphi(m)^2(\log x)^2$ by the remark following Lemma 5. Since $\sum 1/\varphi(m)^2 \ll 1$, we thus may assume that $m > x^{1/5}$. But, using Theorem 1 in de Bruijn [2], we have that for $t > x^{1/5}$ the number of integers $m \le t$ with $P(m) \le \log x$ is at most $t^{c/\log\log x}$ for some absolute constant c. Hence, by partial summation, we have

$$\sum_{m > x^{1/5},\, P(m) \,\le\, \log x} \frac{1}{m} \ \ll\ x^{-1/5+o(1)},$$

so that the number of integers $mr \le x$ with $m > x^{1/5}, P(m) \le \log x$ is negligible. This completes the proof of the theorem.

We remark that Theorem 4 improves the estimate $M(x) = O(x \log\log x/(\log x)^2)$ attained in [10]. The new element in Theorem 4 that allows the elimination of the factor $\log\log x$ is the use of Lemma 5. We also remark that using Lemma 1 in Heath-Brown [5], one has $\gg x/(\log x)^2$ primes p up to x with $l(p)$ either a prime or the product of two primes both larger than $x^{1/4}$. (Indeed, from this lemma we obtain $\gg x/(\log x)^2$ primes $p \le x$ with $p \equiv 7 \pmod{8}$ and either $p-1 = 2q$ with q prime or $p-1 = 2rs$ where r, s are primes $> x^{1/4}$. For such a prime p we have $l(p) = q, r, s$ or rs.) Note that it is trivial that there are at least infinitely many primes p with $l(p)$ prime, in fact there are $\gg \log x/\log\log x$ such primes p below x.

It is probably the case that the estimate of Theorem 4 is best possible. In particular, using the Hardy–Littlewood version of the prime k-tuples conjecture, the number of primes $p \le x$ with $p \equiv 7 \pmod{8}$ and $(p-1)/2$ prime is of order of magnitude $x/(\log x)^2$. But each of these primes p has $l(p) = (p-1)/2$, a prime number. One might conjecture that

$$M(x) \ \sim\ cx/(\log x)^2,$$

where the value of c may be evaluated as follows. First, by Hardy–Littlewood, for a fixed even number k, the number of primes $q \le x/k$ with $kq+1$ prime is $\sim c_k x/(\log x)^2$, where

$$c_k \ =\ \frac{2\alpha}{k} \prod_{p>2,\, p|k} \frac{p-1}{p-2},$$

and where α is the twin prime constant $\prod_{p>2}(1 - 1/(p-1)^2)$. If $2\|k$, then heuristically, the further chance that 2 is a k-th power modulo the prime $kq+1$ is $1/k$. If $2^2\|k$, then the further chance that 2 is a k-th power modulo the prime $kq+1$ is 0, since 2 is a quadratic nonresidue modulo $kq+1$. And if $2^3|k$, then the further chance that 2 is a k-th power modulo the prime $kq+1$ is $2/k$. Thus, a perhaps reasonable guess for the number c is

$$c \;=\; \frac{7}{12}\prod_{p>2}\left(1-\frac{1}{(p+1)(p-1)^2}\right),$$

as per the following calculation:

$$\begin{aligned} c \;&=\; \sum_{2\|k}\frac{c_k}{k} \;+\; \sum_{2^3|k}\frac{2c_k}{k} \\ &=\; \frac{\alpha}{2}\sum_{k\ \mathrm{odd}}\frac{1}{k^2}\prod_{p|k}\frac{p-1}{p-2} \;+\; \frac{\alpha}{16}\sum_{k}\frac{1}{k^2}\prod_{p|k,\,p>2}\frac{p-1}{p-2} \\ &=\; \frac{\alpha}{2}\prod_{p>2}\frac{p^2-p-1}{p^2-p-2} \;+\; \frac{\alpha}{16}\,\frac{4}{3}\prod_{p>2}\frac{p^2-p-1}{p^2-p-2} \\ &=\; \frac{7\alpha}{12}\prod_{p>2}\frac{p^2-p-1}{p^2-p-2} \\ &=\; \frac{7}{12}\prod_{p>2}\left(1-\frac{1}{(p+1)(p-1)^2}\right). \end{aligned}$$

Thus, $c \approx 0.53824$.

References

[1] G. D. Birkhoff and H. S. Vandiver, On the integral divisors of $a^n - b^n$, *Ann. Math.* **5** (1904), 173–180.

[2] N. G. de Bruijn, On the number of positive integers $\le x$ and free of prime factors $> y$, II, *Nederl. Akad. Wetensch. Proc. Ser. A* **69** =*Indag. Math.* **28** (1966), 239–247.

[3] P. Erdős and T. N. Shorey, On the greatest prime factor of $2^p - 1$ for a prime p and other expressions, *Acta Arith.* **30** (1976), 257–265.

[4] H. Halberstam and H.-E. Richert, *Sieve methods*, London Mathematical Society Monographs, No. 4, Academic Press, London–New York, 1974.

[5] D. R. Heath-Brown, Artin's conjecture for primitive roots, *Quart. J. Math. Oxford Ser.* (2) **37** (1986), 27–38.

[6] C. Hooley, On Artin's conjecture, *J. Reine Angew. Math.* **225** (1967), 209–220.

[7] J. C. Lagarias and A. M. Odlyzko, Effective versions of the Chebotarev density theorem, *Algebraic number fields: L-functions and Galois properties (Proc. Sympos., Univ. Durham, Durham, 1975)*, pp. 409–464, Academic Press, London, 1977.

[8] M. R. Murty and S. Wong, The ABC conjecture and prime divisors of the Lucas and Lehmer sequences, *Proc. Millennial Conference on Number Theory*, to appear.

[9] T. Nagell, *Introduction to number theory*, Second edition, Chelsea, New York, 1964.

[10] C. Pomerance On primitive divisors of Mersenne numbers, *Acta Arith.* **46** (1986), 355–367.

[11] A. Schinzel, On primitive prime factors of $a^n - b^n$, *Proc. Cambridge Philos. Soc.* **58** (1962), 555–562.

[12] C. L. Stewart, The greatest prime factor of $a^n - b^n$, *Acta Arith.* **26** (1974/75), 427–433.

[13] C. L. Stewart, On divisors of Fermat, Fibonacci, Lucas, and Lehmer numbers, *Proc. London Math. Soc.* (3) **35** (1977), 425–447.

Leo Murata
Department of Mathematics
Faculty of Economics
Meiji Gakuin University
1-2-37 Shirokanedai, Minato-ku
Tokyo 108-8636, Japan
leo@eco.meijigakuin.ac.jp

Carl Pomerance
Department of Mathematics
Dartmouth College
Hanover, NH 03755, USA
carlp@math.dartmouth.edu