SUMS OF PRIME DIVISORS AND MERSENNE NUMBERS

WILLIAM D. BANKS AND FLORIAN LUCA

Communicated by George Washington

Abstract. In this note, we study those positive integers $n$ with the property that the sum of the distinct prime factors of $n$ divides the $n$-th Mersenne number.

1. Introduction

For any integer $n \geq 1$, let $\beta(n)$ denote the sum of the prime divisors of $n$:

$$\beta(n) = \sum_{p | n} p.$$ 

The study of the function $\beta(n)$ originated in the paper of Nelson, Penney, and Pomerance [7], where the question was raised as to whether the set of Ruth-Aaron numbers (i.e., natural numbers $n$ for which $\beta(n) = \beta(n + 1)$) has zero density in the set of all positive integers. This question was answered in the affirmative by Erdős and Pomerance [5], and the main result of [5] was later improved by Pomerance [10].

More recently, De Koninck and Luca [3] studied positive composite integers $n$ which are divisible by the sum of their prime divisors; they showed that, for suitable positive constants $c_1$ and $c_2$, the inequalities

$$xe^{-c_1 \sqrt{\log x \log \log x}} \leq \# \{ \text{composite } n \leq x : \beta(n) | n \} \leq xe^{-c_2 \sqrt{\log x \log \log x}}$$

hold for all sufficiently large values of $x$.

2000 Mathematics Subject Classification. 11A41; 11A05.

Key words and phrases. Mersenne numbers, prime divisors.

The second author was supported in part by the grant SEP-CONACyT 46755 and by a Guggenheim Fellowship.
In this note, we extend the ideas of [3] and derive upper and lower bounds for the number of positive integers \( n \leq x \) for which \( \beta(n) \) is a divisor of \( 2^n - 1 \).

**Theorem 1.** There exist positive constants \( c_3 \) and \( c_4 \) such that the inequalities
\[
2^{1-c_3 \log \log \log x / \log \log x} \leq \# \{ n \leq x : \beta(n) \mid 2^n - 1 \} \leq c_4 \frac{x \log \log x}{\log x}
\]
hold for all sufficiently large values of \( x \).

Throughout the paper, the letter \( p \) is always used to denote a prime number, and we put \( \pi(x) = \# \{ p \leq x \} \) as usual. For an integer \( n > 1 \), we use \( P(n) \) to denote the largest prime factor of \( n \), \( \omega(n) \) the number of distinct prime divisors of \( n \), and \( \Omega(n) \) the total number of prime factors of \( n \), counted with multiplicity; we also put \( P(1) = 1 \), and \( \omega(1) = \Omega(1) = 0 \). The Euler function is denoted by \( \varphi(n) \). For any real number \( x > 0 \) and an integer \( k \geq 2 \), \( \log_k x \) denotes the \( k \)-th iterate of the function \( \log x = \max \{ \ln x, 1 \} \), where \( \ln(\cdot) \) is the natural logarithm.

Finally, we use the Vinogradov symbols \( \gg \) and \( \ll \), as well as the Landau symbols \( O \) and \( o \), with their usual meanings.

2. Proof of Theorem 1

2.1. The Upper Bound. For an odd prime \( p \), let \( t(p) \) denote the multiplicative order of \( 2 \) modulo \( p \). Let \( x \) be a large real number and put
\[
y = y(x) = \exp \left( \frac{\log x \log_3 x}{2 \log_2 x} \right)
\]
and
\[
u = u(x) = \frac{\log x}{\log y} = \frac{2 \log_2 x}{\log_3 x}.
\]
Let
\[\mathcal{B}(x) = \{ n \leq x : \beta(n) \mid 2^n - 1 \},\]
and consider the following subsets of \( \mathcal{B}(x) \):
\[
\mathcal{B}_1(x) = \{ n \in \mathcal{B}(x) : P(n) \leq y \},
\mathcal{B}_2(x) = \{ n \in \mathcal{B}(x) \setminus \mathcal{B}_1(x) : P(n)^2 \mid n \},
\mathcal{B}_3(x) = \{ n \in \mathcal{B}(x) \setminus (\mathcal{B}_1(x) \cup \mathcal{B}_2(x)) : \omega(n) > 10 \log_2 x \},
\mathcal{B}_4(x) = \{ n \in \mathcal{B}(x) \setminus (\bigcup_{j=1}^3 \mathcal{B}_j(x)) : \beta(n) \geq 2P(n) \},
\mathcal{B}_5(x) = \{ n \in \mathcal{B}(x) \setminus (\bigcup_{j=1}^4 \mathcal{B}_j(x)) : \Omega(\beta(n)) > 10 \log_2 x \},
\mathcal{B}_6(x) = \{ n \in \mathcal{B}(x) \setminus (\bigcup_{j=1}^5 \mathcal{B}_j(x)) : t(p) > \log^2 x \text{ for some } p \mid \beta(n) \},
\mathcal{B}_7(x) = \mathcal{B}(x) \setminus (\bigcup_{j=1}^6 \mathcal{B}_j(x)).
\]
As $B(x)$ is the union of the sets $B_j(x)$, $j = 1, \ldots, 7$, it suffices to find an appropriate bound for the cardinality of each set $B_j(x)$.

Using the following well known estimate for smooth numbers (see [2]):
\[
\# \{n \leq x : P(n) \leq y \} = x u^{-u+o(u)},
\]
which holds as $x \to \infty$ with our choice of $y$ and $u$, we derive that
\begin{equation}
\# B_1(x) \leq x \exp \left( -2(1 + o(1)) \log_2 x \right) = \frac{x}{(\log x)^{2+o(1)}} \quad (x \to \infty).
\end{equation}

Next, for every integer $n \in B_2(x)$ there exists a prime $p > y$ such that $p^2 | n$. For a fixed prime $p$, the number of positive integers $n \leq x$ with the latter property is precisely $\lfloor x/p^2 \rfloor$; therefore,
\begin{equation}
\# B_2(x) \leq \sum_{p > y} \frac{x}{p^2} \leq \sum_{k > y} \frac{1}{k^2} \ll \frac{x}{y \log x} = o \left( \frac{x}{\log x} \right) \quad (x \to \infty).
\end{equation}

To estimate $\# B_3(x)$, let $A = \{n : \Omega(n) > 10 \log_2 n\}$, and put $A(t) = A \cap [1, t]$ for all $t \geq 1$. A result of Nicolas [8] implies that the bound
\begin{equation}
\# A(t) \ll \frac{t \log t \log_2 t}{2^{10 \log_2 t}} = \frac{t}{(\log t)^{\alpha + o(1)}}
\end{equation}
holds as $t \to \infty$, where $\alpha = 10 \ln 2 - 1 > 3$. Since $B_3(x) \subset A(x)$, it follows that the inequality
\begin{equation}
\# B_3(x) \leq \frac{x}{\log^5 x}
\end{equation}
holds if $x$ is sufficiently large.

Next, let $n \in B_4(x)$. Write $n = Pm$ with $P = P(n)$, and put $Q = P(m)$; then $P > \max\{y, Q\}$. Since
\[ \beta(m) + P = \beta(n) \geq 2P, \]
we have the estimate
\[ P \leq \beta(m) \leq \omega(m)Q \leq 10Q \log_2 x, \]
or
\[ Q \geq \frac{P}{10 \log_2 x}. \]
Noting that $PQ | n$, for fixed values of $P$ and $Q$ the number of such integers $n \in B_4(x)$ does not exceed $\lfloor x/(PQ) \rfloor$. Summing up over all possible choices for $P$ and $Q$ and using Mertens’ estimate
\[ \sum_{p \leq x} \frac{1}{p} = \log_2 x + b_1 + O \left( \frac{1}{\log x} \right) \]
for some constant $b_1$, we derive that
\[
\#B_4(x) \leq \sum_{P > y} \frac{1}{P} \sum_{0.1P/\log_2 x \leq Q < P} \frac{x}{PQ} \\
= x \sum_{P > y} \frac{1}{P} \left( \log_2 P - \log_2 \left( \frac{P}{10 \log_2 x} \right) + O \left( \frac{1}{\log(0.1P/\log_2 x)} \right) \right) \\
= x \sum_{P > y} \frac{1}{P} \left( \log \left( 1 + \frac{\log(10 \log_2 x)}{\log P - \log(10 \log_2 x)} \right) + O \left( \frac{1}{\log P} \right) \right) \\
\ll x \sum_{P > y} \frac{\log_3 x}{P \log P} = x \log_3 x \int_{t > y} \frac{d\tau(t)}{t \log t},
\]
where we have used the fact that the estimate
\[
\log(0.1P/\log_2 x) = \log P \left( 1 - \frac{\log_3 x + O(1)}{\log P} \right) \\
= \log P \left( 1 + O \left( \frac{\log_3 x}{\log y} \right) \right) = (1 + o(1)) \log P
\]
holds uniformly for $P > y$ as $x \to \infty$.

Next, we turn our attention to integers $n \in B_5(x)$. As before, write $n = Pm$ with $P = P(n)$ and $m < x/y$. For a fixed choice of $m < x/y$, the prime $P = \beta(n) - \beta(m)$ is determined uniquely by $\beta(n) \in A$. We also have
\[
\beta(n) < 2P \leq 2x/m.
\]
Using estimate (3) again, it follows that the inequality
\[
\#A(t) \leq \frac{t}{\log^4 t}
\]
holds whenever $t > t_0$, for some constant $t_0$. Now, assuming that $x$ is large enough, we have $2x/m > P > y > t_0$, therefore the number of possibilities for $P$ (or $\beta(n)$) for each fixed choice of $m$ is at most
\[
\#A(2x/m) \leq \frac{2x}{m \log^3(2x/m)} \ll \frac{x}{m \log^3 y} \ll \frac{x \log_2^3 x}{m \log^2 x}.
\]
Summing over all the possible choices for $m$, we get that
\[
\#B_5(x) \ll \frac{x \log_2^3 x}{\log^3 x} \sum_{m < x/y} \frac{1}{m} \ll \frac{x \log_2^3 x}{\log^2 x}.
\]
Next, we study integers \( n \in \mathcal{B}_6(x) \). Again, write \( n = Pm \) with \( P = P(n) \); then \( x/m \geq P > y \). Fix \( m \). For every prime \( p \mid \beta(n) \mid 2^n - 1 \), it is clear that \( t(p) \mid n \). If \( P \mid t(p) \), then since \( t(p) \mid p - 1 \), we see that \( p \equiv 1 \pmod{P} \). Since \( P > y > 2 \) is odd, it follows that \( p > 2P + 1 \); thus, \( \beta(n) \geq p > 2P + 1 \); but this contradicts the fact that \( n \not\in \mathcal{B}_4(x) \). Hence, \( P \nmid t(p) \), and therefore \( t(p) \mid m \).

Now, let \( p \) be a prime factor of \( \beta(n) \) for which \( d = t(p) > \log^2 x \); note that \( p \equiv 1 \pmod{d} \). We claim that \( p \neq P \). Indeed, if \( p = P \), then \( P \mid \beta(n) \), and since \( \beta(n) < 2P \), it follows that \( \beta(n) = P \). But this implies that \( n = P \); hence, \( P | 2P - 1 \), which is clearly impossible since \( 2P - 1 \equiv 1 \pmod{P} \).

The congruence \( \beta(n) \equiv 0 \pmod{p} \) leads to \( P \equiv -\beta(m) \equiv 0 \pmod{p} \). Since \( P \leq x/m \), it follows that (for fixed \( m \)) the number of possibilities for \( P \) cannot exceed \( x/m \); since \( p \leq \beta(n) < 2P \leq 2x/m \), we see that the number of such possibilities is \( \leq x/(mp) + 1 \leq 3x/(mp) \). Now, summing over all primes \( p \equiv 1 \pmod{d} \), then over positive multiples \( m \leq x \) of \( d \), and finally, over all integers \( d > \log^2 x \), it follows that

\[
\#\mathcal{B}_0(x) \leq 3 \sum_{d > \log^2 x} \sum_{m \leq x \pmod{d}} \sum_{p \leq x \pmod{d}} \frac{x}{mp} \ll x \sum_{d > \log^2 x} \frac{\log x \log x}{d \varphi(d)} \ll x \frac{\log x \log x}{\log x}.
\]

Here, we have used the bound

\[
\sum_{p \leq x \pmod{d}} \frac{1}{p} \ll \frac{\log x}{\varphi(d)},
\]

which holds uniformly for \( 2 \leq d \leq x \) (see formula (3.1) of [4], or Lemma 1 of [1]), together with Landau’s bound [6]:

\[
\sum_{d > t} \frac{1}{d \varphi(d)} \ll \frac{1}{t}.
\]

Finally, we consider integers \( n = Pm \) in the set \( \mathcal{B}_7(x) \). For every prime \( p \mid \beta(n) \), we have \( p \mid 2^d - 1 \), where \( d = t(p) \leq \log^2 x \), and \( d \mid m \) as before. Thus, every
prime divisor of $\beta(n)$ also divides

$$M = \prod_{d \leq \log^2 x} (2^d - 1) \leq \exp \left( \sum_{d \leq \log^2 x} d \right) = \exp(O(\log^4 x)).$$

If $W$ denotes the set of distinct prime factors of $M$, then

$$\#W = \omega(M) \ll \frac{\log M}{\log_2 M} \ll \frac{\log^4 x}{\log_2 x}.$$ 

Since $n \notin B_5(x)$, we have $\Omega(\beta(n)) \leq 10 \log_2 x$. Hence, if $m$ is fixed and $E_m = \{\beta(Pm) : n = Pm \in B_7(x)\}$, then

$$\#E_m \leq \sum_{k \leq 10 \log_2 x} \binom{\#W + k - 1}{k} \leq 10 \log_2 x (\#W + 10 \log_2 x)^{10 \log_2 x} = \exp(O((\log_2 x)^2)).$$

Since $m \leq x/y$, we obtain that

$$\#B_7(x) \leq \sum_{m \leq x/y} \#E_m \leq \frac{x}{y} \exp(O((\log_2 x)^2)) = x \exp(O((\log_2 x)^2) - \frac{2 \log x \log_3 x}{\log_2 x}),$$

which implies that

$$\#B_7(x) = o\left(\frac{x}{\log x}\right) \quad (x \to \infty). \quad (8)$$

Combining the estimates (1), (2), (4), (5), (6), (7), and (8), we obtain the upper bound stated in Theorem 1.

2.2. The Lower Bound. Let $x$ be a large real number. Let $y = y(x)$ be a function that tends to infinity with $x$ (to be determined later), and put

$$z = \exp\left(10\sqrt{\log y \log_2 y}\right) \quad \text{and} \quad v = \frac{\log y}{\log z} = \frac{1}{10} \sqrt{\frac{\log y}{\log_2 y}}$$

Let $P$ denote the set of primes $p$ in the interval $[z/2, z]$ with the property that $p - 1$ is square-free; it is known (see [9]) that $\#P = 0.5\alpha(1 + o(1))\pi(z)$ as $x \to \infty$, where

$$\alpha = \prod_{p \geq 2} \left(1 - \frac{1}{p(p - 1)}\right) = 0.37395 \ldots.$$
is the Artin constant. In particular,
\[ \#P > \frac{z}{10 \log z} > |v| \]
if \( x \) is sufficiently large. Let \( M \) be a fixed square-free positive integer obtained by multiplying together \( |v| \) distinct elements of \( P \). Then,
\[ y = z^v \geq z^{[v]} \geq M \geq \left( \frac{z}{2} \right)^{[v]} \geq \frac{y}{2^v z} > \frac{y}{z^2} \]
if \( x \) is large enough. We now put
\[ R = \sum_{p \leq z^2} p \quad \text{and} \quad N = M - R. \]
Clearly, \( R = o(M) \) as \( x \to \infty \); hence, the inequalities
\[ y > N > \frac{y}{2z^2} \]
hold if \( x \) is sufficiently large.

Next, let \( K \) be an integer of size
\[ K = \frac{z}{\log z} + O(1) \]
such that \( K \) satisfies the parity condition \( K \equiv N \pmod{2} \). Let \( \mathcal{I} \) be the interval \([N/(3K), N/(2K)]\). If \( x \) is sufficiently large, then, by (9), we have
\[ \frac{N}{K} > \frac{y}{2z^2} \cdot \frac{\log z}{2z} = \frac{y \log z}{4z^3}. \]
In particular, \( N/(3K) > z \), which shows every prime in \( \mathcal{I} \) is larger than \( z \). Using (9) and (11), we also see that \( \pi(\mathcal{I}) \), the number of primes in the interval \( \mathcal{I} \), satisfies the lower bound
\[ \pi(\mathcal{I}) = \pi(N/(2K)) - \pi(N/(3K)) \]
\[ \geq \frac{1}{7 K \log(N/K)} > \frac{y \log z}{28z^3 \log y} > \frac{y}{z^3} \]
if \( x \) is large enough, and therefore \( \pi(\mathcal{I}) > 2K \).

Now, let \( S \) be an arbitrary set of \( K - 3 \) distinct primes in \( \mathcal{I} \), and put \( S = \sum_{p \in S} p \). It is clear that
\[ S \leq (K - 3) \frac{N}{2K} < \frac{N}{2}, \]
and that the numbers \( S \) and \( N \) have opposite parities. Thus, \( N - S > N/2 \) is an odd number. We now apply Vinogradov’s Three Primes Theorem (see, for
example, [11]), to conclude that the number $V(S)$ of representations of $N - S$ of the form

$$N - S = p_1 + p_2 + p_3, \quad p_1 < p_2 < p_3,$$

where $p_1$, $p_2$, and $p_3$ are prime numbers, satisfies the lower bound

$$V(S) > c_5 \frac{(N - S)^2}{(\log(N - S))^3} > c_6 \frac{N^2}{(\log N)^3}$$

for suitable (absolute) positive constants $c_5$ and $c_6$ once $x$ is large enough. Moreover, we can assume that each prime $p_1$ that appears in (13) satisfies the bound

$$p_1 > c_7(N - S) > c_8N,$$

where $c_7$ is a positive absolute constant, and $c_8 = c_7/2$. Since $K$ tends to infinity with $x$, it follows that $c_8N > N/(2K)$ if $x$ is large enough; therefore, $p_j \not\in S$ for $j = 1, 2, 3$.

The above argument shows that if $W$ is the set of ordered $K$-tuples of primes $(q_1, \ldots, q_K)$ satisfying the conditions

- $q_1 < \cdots < q_K$;
- $q_j \in I$ for $j = 1, \ldots, K - 3$;
- $q_{K-2} > N/(2K)$;
- $N = q_1 + \cdots + q_K$;

then the cardinality of $W$ is at least

$$\#W \geq \sum_{S \subset I} V(S) \geq c_6 \frac{N^2}{(\log N)^3} \left( \frac{\pi(I)}{K - 3} \right).$$

Let $Q = (q_1, \ldots, q_K)$ be such a $K$-tuple in $W$, and put

$$n_Q = \prod_{p \leq z} p \prod_{j=1}^K q_j.$$

Since every prime in $I$ is larger than $z$, it follows that $n_Q$ is square-free. Moreover, by unique factorization, the map $Q \mapsto n_Q$ is one-to-one. We claim that each integer $n_Q$ satisfies the desired property, namely, $\beta(n_Q) \mid 2^{n_Q} - 1$. To see this, we first observe that

$$\beta(n_Q) = \sum_{p \leq z} p + \sum_{j=1}^k q_j = R + N = M.$$
Let $\lambda(M)$ be the value of the Carmichael function at $M$; i.e., the exponent of the multiplicative group $(\mathbb{Z}/M\mathbb{Z})^*$. Since $M$ is odd, square-free, and composed solely of primes $q$ from $I$, we have

$$\lambda(M) = \text{lcm}_{q|M}(q-1) \prod_{p \leq z} p^{\alpha_p}. $$

Here, we have used the fact that $q - 1$ is square-free and smaller than $z$ for all primes $q \mid M$. Since $2 \nmid M$, it follows that $2 \in (\mathbb{Z}/M\mathbb{Z})^*$, and since $\lambda(M) \mid n_Q$, we have

$$2^{n_Q} \equiv 1 \pmod{\lambda(M)}. $$

Then, in view of (15), we have $\beta(n_Q) \mid 2^{n_Q} - 1$, as claimed.

It remains to get a lower bound for the number of integers $n_Q$ constructed in this way. First, note that the largest such $n_Q$ satisfies the inequality

$$n_Q \leq \left( \prod_{p \leq z} p \right)^K \left( \frac{y}{2K} \right)^K < \exp (2z + K \log y) \leq \exp \left( 2z + \frac{z \log y}{\log z} + \log y \right), $$

since we can choose

$$K \in \left\{ \left\lfloor \frac{z}{\log z} \right\rfloor, \left\lfloor \frac{z}{\log z} \right\rfloor + 1 \right\}$$

to satisfy (10) as well as the parity condition $K \equiv N \pmod{2}$. Therefore, given $x$, let us now define $y$ (hence also $z$) implicitly via the relation

$$(16) \quad x = \exp \left( 2z + \frac{z \log y}{\log z} + \log y \right). $$

We remark that, if $y$ is large enough, the right side of (16) is a strictly increasing function of $y$; therefore, if $x$ is sufficiently large, the value of $y$ is uniquely determined by (16). Using (14), we have for all sufficiently large $x$:

$$\#B(x) \geq \#W \geq \frac{N^2}{(\log N)^3} \left( \frac{\pi(I)}{K - 3} \right)^K \geq \frac{N^2}{(\log N)^3} \left( \frac{\pi(I)}{K - 3} \right)^{K-3} \geq \frac{N^2}{(\log N)^3} \left( \frac{\pi(I)}{K - 3} \right)^K \exp \left( K \log \left( \frac{\pi(I)}{K} \right) \right). $$(17)

By (9) and the fact that $N = y^{1 + o(1)}$ as $x \to \infty$, it follows that

$$\frac{N^2}{(\log N)^3} \geq \frac{y^2}{z^4(\log y)^3}. $$
Using (9) and (11), we also have the bound

\[ \pi(I) \leq \pi(N/K) \ll \frac{N}{K \log(N/K)} \ll \frac{y}{K \log y} \]

Applying the last two estimates and (12) to the bound (17), we derive that

\[ \#B(x) \gg K^6 \frac{y}{y^2 z^4} \exp \left( K \log \left( \frac{y}{K z^3} \right) \right) \]
\[ = \exp \left( K \log y - K \log(K z^3) + \log \left( K^6/(y^2 z^4) \right) \right) \]
\[ = \exp \left( \log x \left( 1 - \frac{E}{\log x} \right) \right) \]

where

\[ E = \log x - K \log y + K \log(K z^3) - \log \left( K^6/(y^2 z^4) \right). \]

Using (10) and (16), we immediately deduce that

\[ E = 6 z \left( 1 + O \left( \frac{\log y}{\log z} \right) \right) = (6 + o(1)) z, \]

and therefore

\[ \#B(x) \geq \exp \left( \log x \left( 1 - \frac{(6 + o(1)) z}{\log x} \right) \right) \quad (x \to \infty). \]

Using (16) and the definition of \( z \), we also have

\[ \log x = (1 + o(1)) \frac{\log y}{\log z} = (0.1 + o(1)) z \sqrt{\frac{\log y}{\log_2 y}}, \]

which implies that

\[ \log_2 x = (1 + o(1)) \log z = (10 + o(1)) \sqrt{\log y \log_2 y}, \]

or

\[ \log y = \left( \frac{1}{200} + o(1) \right) \frac{(\log_2 x)^2}{\log_3 x} \quad (x \to \infty). \]

Inserting this expression into (19), it follows that

\[ \frac{z}{\log x} = (200 + o(1)) \frac{\log_4 x}{\log_2 x} \quad (x \to \infty). \]

Finally, substituting this expression into (18), we obtain a lower bound of the form stated in Theorem 1, where we can take \( c_3 = 1201 \).
3. Remarks and Comments

Theorem 1 shows that the number of $n \leq x$ for which $\beta(n) \mid 2^n - 1$ is $x^{1+o(1)}$ as $x \to \infty$. On the other hand, the upper bound is rather weak; in particular, we cannot say anything about the convergence or the divergence of the sum of the reciprocals of integers $n$ with this property, and we would like to pose this as an open problem for the reader.

References


(William D. Banks) Department of Mathematics, University of Missouri, Columbia, MO 65211, USA
E-mail address: bbanks@math.missouri.edu

(Florian Luca) Instituto de Matemáticas, Universidad Nacional Autónoma de México, C.P. 58089, Morelia, Michoacán, México
E-mail address: fluca@matmor.unam.mx