ON THE OPPENHEIM’S "FACTORISATIO NUMERORUM" FUNCTION

FLORIAN LUCA, ANIRBAN MUKHOPADHYAY AND KOTYADA SRINIVAS

1. Introduction

Let \( f(n) \) denote the number of distinct unordered factorisations of the natural number \( n \) into factors larger than 1. For example, \( f(28) = 4 \) as 28 has the following factorisations

\[
28, \ 2 \cdot 14, \ 4 \cdot 7, \ 2 \cdot 2 \cdot 7.
\]

In this paper, we address three aspects of the function \( f(n) \). For the first aspect, in [1], Canfield, Erdős and Pomerance mention without proof that the number of values of \( f(n) \) that do not exceed \( x \) is \( x^{o(1)} \) as \( x \to \infty \). Our first theorem in this note makes this result explicit.

For a set \( A \) of positive integers we put \( A(x) = \{ n \in A : n \leq x \} \).

Theorem 1. Let \( A = \{ f(m) : m \in \mathbb{N} \} \). Then

\[
\# A(x) = x^{O(\log \log \log x / \log \log x)}, \quad \text{as } x \to \infty.
\]

Secondly, there is a large body of literature addressing average values of various arithmetic functions in short intervals. Our next result gives a lower bound for the average of \( f(n) \) over a short interval.

Theorem 2. Uniformly for \( x \geq 1 \) and \( y > e^{e^e} \), we have

\[
\frac{1}{y} \sum_{x \leq n \leq x+y} f(n) \geq \exp \left( \left( \frac{4}{\sqrt{2e}} + O \left( \frac{(\log \log y)^2}{\log \log y} \right) \right) \frac{\sqrt{\log y}}{\log \log y} \right).
\]

Finally, there are also several results addressing the behavior of positive integers \( n \) which are multiples of some other arithmetic function of \( n \). See, for example, [3], [5], [9] and [10] for problems related to counting positive integers \( n \) which are divisible by either \( \omega(n) \), \( \Omega(n) \) or \( \tau(n) \), where these functions are the number of distinct prime factors of \( n \), the number of total prime factors of \( n \), and number of divisors of \( n \), respectively. Our next and last result gives an upper bound on the counting function of the set of positive integers \( n \) which are multiples of \( f(n) \).
Theorem 3. Let $B = \{ n : f(n) | n \}$. Then
\[
\#B(x) = \frac{x}{(\log x)^{1+o(1)}}, \quad \text{as } x \to \infty.
\]

2. Preliminaries and lemmas

The function $f(n)$ is related to various partition functions. For example, $f(2^n) = p(n)$, where $p(n)$ is the number of partitions of $n$. Furthermore, $f(p_1 p_2 \cdots p_k) = B_k$, where $B_k$ is the $k$th Bell number which counts the number of partitions of a set with $k$ elements in nonempty disjoint subsets. In general, $f(p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k})$ is the number of partitions of a multiset consisting of $\alpha_i$ copies of $\{i\}$ for each $i = 1, \ldots, k$.

Throughout the paper, we put $\log x$ for the natural logarithm of $x$. We use $p$ and $q$ for prime numbers and $O$ and $o$ for the Landau symbols.

The following asymptotic formula for the $k$th Bell number is due to de Bruijn [4].

**Lemma 1.**
\[
\frac{\log B_k}{k} = \log k - \log \log k - 1 + \frac{\log \log k}{\log k} + \frac{1}{\log k} + O \left( \frac{(\log \log k)^2}{(\log k)^2} \right).
\]

We also need the Stirling numbers of the second kind $S(k, l)$ which count the number of partitions of a $k$ element set into $l$ nonempty disjoint subsets. Clearly,

\begin{equation}
B_k = \sum_{l=1}^{k} S(k, l).
\end{equation}

We now formulate and prove a few lemmas about the function $f(n)$ which will come in handy later on.

The first lemma is an easy observation, so we state it without proof.

**Lemma 2.** If $a | b$, then $f(a) \leq f(b)$.

We let $p_n$ denote the $n$th prime number and $\alpha_1(n)$ denote the maximal exponent of a prime appearing in the prime factorization of $n$. Let $n$ be a positive integer with prime factorization
\[n = q_1^{\alpha_1} q_2^{\alpha_2} \cdots q_k^{\alpha_k},\]
where $q_1, \ldots, q_k$ are distinct primes and $\alpha_1(n) := \alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_k$. We put $n_0 = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$, and observe that $f(n) = f(n_0)$. This observation will play a crucial role in the proof of Theorem [4].

The following lemma gives upper bounds for $\alpha_1(n)$ and $\omega(n)$ when $f(n) \leq x$. 

Lemma 3. Let $n = q_1^{\alpha_1}q_2^{\alpha_2} \cdots q_k^{\alpha_k}$, where $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_k$ and $f(n) \leq x$. Then

(i) $\alpha_1 = O((\log x)^2)$;
(ii) $k = \omega(n) = O(\log x/\log \log x)$.

Proof. It follows from Lemma 2 that

$$f(n) \geq f(q_1^{\alpha_1}) = p(\alpha_1).$$

Using the following asymptotic formula for $p(n)$ due to Hardy and Ramanujan [6]

$$p(n) \sim \frac{1}{4n\sqrt{3}} \exp(\pi \sqrt{2n/3}),$$

we conclude that $\exp(c \sqrt{\alpha_1}) \leq x$ holds with some constant $c > 0$. Hence, (i) follows. In order to prove (ii), let again $n_0 = p_1p_2\cdots p_k$. By Lemma 2 we have $f(n_0) \leq f(n) \leq x$. Furthermore, $f(n_0) = B_k$. It now follows from Lemma 1 that

$$\exp((1 + o(1))k \log k) = B_k \leq x,$$

as $k \to \infty$, yielding

$$k = O\left(\frac{\log x}{\log \log x}\right),$$

which completes the proof of the lemma.

Recall that the Möbius function $\mu(m)$ of the positive integer $m$ is $(-1)^{\omega(m)}$ if $m$ is squarefree and 0 otherwise.

For a positive integer $k$ and positive real numbers $A \leq B$ we let

$$\mathcal{M}_{k,A,B} = \{m : \mu(m) \neq 0, \omega(m) = k, \text{ if } p | m \text{ then } p \in [A, B]\}.$$

We also put

$$S_{A,B} = \sum_{A \leq p \leq B} \frac{1}{p}.$$

Lemma 4. Uniformly in $A \geq 2$, $B \geq 3$ and $k \geq 2$, we have

$$\sum_{m \in \mathcal{M}_{k,A,B}} \frac{1}{m} \geq \left(1 + O\left(\frac{k^2}{S_{A,B}^2 A \log A}\right)\right) \frac{1}{k!} S_{A,B}^k.$$

Proof. We omit the dependence of the subscripts in order to simplify the presentation. It is not hard to see that

$$\sum_{m \in \mathcal{M}} \frac{1}{m} \geq \frac{1}{k!} \left(\sum_{A \leq p \leq B} \frac{1}{p}\right)^k - \sum_{A \leq p \leq B} \frac{1}{p^2} \left(\sum_{A \leq p \leq B} \frac{1}{p}\right)^{k-2}. $$

(3)
Indeed, if $m = q_1^{\alpha_1} \cdots q_s^{\alpha_s}$, with $\alpha_1 \geq 2$ and $\alpha_1 + \cdots + \alpha_s = k$, then, by unique factorization, in the first sum on the right hand side of inequality (3), the number $1/m$ appears with coefficient
\[
\frac{1}{k!} \left( \frac{k!}{\alpha_1! \cdots \alpha_s!} \right) = \frac{1}{\alpha_1! \cdots \alpha_s!},
\]
while in the second sum in the right hand side of the inequality (3), the number $1/m$ appears with coefficient
\[
\sum_{1 \leq i \leq s, \; \alpha_i \geq 2} \frac{1}{(k-2)!} \left( \frac{(k-2)!}{\alpha_1! \cdots (\alpha_i - 2)! \cdots \alpha_s!} \right) > \frac{1}{\alpha_1! \cdots \alpha_s!}.
\]
This establishes inequality (3). Using this inequality, we get
\[
\sum_{m \in M} \frac{1}{m} \geq \frac{S^k}{k!} \left(1 - \frac{k^2}{S^2} \sum_{p \geq A} \frac{1}{p^2} \right).
\]
An argument involving the Prime Number Theorem and partial summation gives
\[
\sum_{p \geq A} \frac{1}{p^2} = O \left( \frac{1}{A \log A} \right).
\]
Hence,
\[
\sum_{m \in M} \frac{1}{m} \geq \frac{S^k}{k!} \left(1 + O \left( \frac{k^2}{S^2 A \log A} \right) \right).
\]
This completes the proof of the lemma.

3. Proofs of the theorems

3.1. **Proof of Theorem [1]** For a positive integer $n$, we let again $n_0$ and $\alpha_1(n)$ be the functions defined earlier. We let $\mathcal{A}(x) = \{m_1, \ldots, m_t\}$ be such that $m_1 < m_2 < \cdots < m_t$ and let $\mathcal{N} = \{n_1, \ldots, n_t\}$ be positive integers such that $n_i$ is minimal among all positive integers $n$ with $f(n) = m_i$ for all $i = 1, \ldots, t$. It is clear that if $n \in \mathcal{N}$, then $n = n_0$. Since $\# \mathcal{A}(x) = t = \# \mathcal{N}$, it suffices to bound the cardinality of $\mathcal{N}$.

We partition this set as $\mathcal{N} = \mathcal{N}_1 \cup \mathcal{N}_2 \cup \mathcal{N}_3$, where
\[
\mathcal{N}_1 = \{n \in \mathcal{N} : \alpha_1(n) \leq \log \log x\},
\]
\[
\mathcal{N}_2 = \left\{n \in \mathcal{N} : \omega(n) \leq \frac{\log x}{(\log \log x)^2} \right\},
\]
and
\[
\mathcal{N}_3 = \mathcal{N} \setminus \mathcal{N}_1 \setminus \mathcal{N}_2.
\]
and
\[ N_3 = N \setminus N_1 \cup N_2. \]
If \( n \in N_1 \), then \( n \) has at most \( O(\log x / \log \log x) \) prime factors (by Lemma \( \ref{lem:prime_factors} \), each one appearing at an exponent less than \( \log \log x \).

Therefore,
\[
\#N_1 = (\log \log x)^{O(\log x / \log \log x)} = \exp \left( O \left( \frac{\log x \log \log x}{\log \log x} \right) \right)
\]
(4)
\[ = x^{O \left( \frac{\log \log x}{\log \log x} \right)} \]
as \( x \to \infty \).

Next, we observe that an integer in \( N_2 \) has at most \( \log x / (\log \log x)^2 \) prime factors, each appearing at an exponent \( O((\log x)^2) \) (by Lemma \( \ref{lem:prime_factors} \)). Thus,
\[
\#N_2 \leq (O(\log x)^2 \log x / (\log \log x)^2)^{\log x / (\log \log x)^2} = \exp \left( \frac{(2 + o(1)) \log x}{\log \log x} \right)
\]
(5)
\[ = x^{O \left( \frac{\log x}{\log \log x} \right)} \]
as \( x \to \infty \).

Finally, let \( n \in N_3 \), and write it as
\[ n = p_1^{\alpha_1} \cdots p_i^{\alpha_i} p_{i+1}^{\alpha_{i+1}} \cdots p_k^{\alpha_k}, \]
where we put
\[ i := \max\{ j \leq k : \alpha_j \geq y \}, \]
where \( y = \log \log x / \log \log \log x \).

Observe that the divisor \( p_{i+1}^{\alpha_{i+1}} \cdots p_i^{\alpha_i} \) of \( n \) can be chosen in at most
\[
(y + 1)^k = (y + 1)^{O(\log x / \log \log x)} = \exp \left( O \left( \frac{\log x \log \log x}{\log \log x} \right) \right)
\]
(6)
ways. Furthermore, by Lemma \( \ref{lem:prime_factors} \) we trivially have that \( n' = p_1^{\alpha_1} \cdots p_i^{\alpha_i} \) can be chosen in at most
\[
(O((\log x)^2))^i = \exp (O(i \log \log x)).
\]

Thus, putting \( N_4 \) for the subset of \( N_3 \) such that \( i \leq \log x / (\log \log x)^2 \), we get that
\[
\#N_4 \leq \exp \left( O \left( \frac{\log x}{\log \log x} \right) \right).
\]
(7)
From now on, we look at \( n \in N_5 = N_3 \setminus N_4 \).
For each \( t \), we let \( k_t \) be such that \( S(t, k_t) \) is maximal among the numbers \( S(t, k) \) for \( k = 1, \ldots, t \). By formula (I), the definition of \( k_t \), and Lemma I we have that

\[
S(t, k_t) \geq \frac{B_t}{t} = \frac{\exp((1 + o(1))t \log t)}{t} = \exp((1 + o(1))t \log t)
\]
as \( t \to \infty \). We now claim that

\[
f(n) \geq f(n') \geq f((p_1 \cdots p_i)^y) \geq \frac{S(i, k_i)^y}{(yk_i)!}.
\]
The first three inequalities follow immediately from Lemma 2 so let us prove the last one.

Note that \( S(i, k_i) \) counts the number of factorizations of \( p_1 p_2 \cdots p_i \) in precisely \( k_i \) factors. Therefore, \( (S(i, k_i))^y \) counts the number of factorizations of \( (p_1 p_2 \cdots p_i)^y \) into \( k_i y \) square-free factors, where we count each such factorization at most \( (k_i y)! \) times. This establishes the claim.

Since \( i \) tends to infinity for \( n \in \mathcal{N}_5 \), we get that

\[
S(i, k_i)^y \geq \exp((1 + o(1))yi \log i).
\]
Furthermore, we trivially have

\[
(k_i y)! \leq (k_i y)^{k_i y} = \exp(k_i y \log(k_i y)).
\]
Thus,

\[
f(n) \geq \frac{S(i, k_i)^y}{(k_i y)!} \geq \exp((1 + o(1))yi \log i - k_i y \log(k_i y))
\]
as \( x \to \infty \). We next show that for our choices of \( y \) and \( i \) we have

\[
k_i y \log(k_i y) = o(yi \log i) \quad \text{as} \quad x \to \infty.
\]
Indeed, using the fact

\[
k_i = (1 + o(1)) \frac{i}{\log i}
\]
(see, for example, [2]), we see that the above condition is equivalent to

\[
\log y = o((\log i)^2),
\]
which holds as \( x \to \infty \) because \( y = \log \log x / \log \log \log x \) and \( i \geq \log x / (\log \log x)^2 \). Now the inequality \( f(n) \leq x \) together with [3] and the fact that \( \log i \geq (1 + o(1)) \log \log x \) implies that

\[
i \leq (1 + o(1)) \frac{\log x}{y \log \log x} \quad \text{as} \quad x \to \infty,
\]
therefore $n'$ can be chosen in at most

$$\left( O \left( \left( \log x \right)^2 \right) \right)^i \leq \left( O \left( \left( \log x \right)^2 \right) \right)^{(1 + o(1))} \frac{\log y}{\log \log x} = \exp \left( \left( 2 + o(1) \right) \frac{\log y}{y} \right)$$

ways. As we have already seen at (6), the complementary divisor $n/n' = \prod_{i=1}^{t} p_i^{a_i}$ of $n$ can be chosen in at most

$$\exp \left( O \left( \frac{y \log \log \log \log x}{\log \log x} \right) \right)$$

ways. Thus, the total number of choices for $n$ in $\mathcal{N}_5$ is

$$\# \mathcal{N}_5 \leq \exp \left( O \left( \frac{\log x \log \log \log \log x}{\log \log \log \log \log x} \right) \right).$$

From estimates (7) and (12) we get

$$\# \mathcal{N}_4 \leq \# \mathcal{N}_1 + \# \mathcal{N}_2 + \# \mathcal{N}_3 \leq \exp \left( O \left( \frac{\log \log \log \log x}{\log \log \log \log \log x} \right) \right),$$

which completes the proof of the theorem.

3.2. Proof of Theorem 2. We assume that $y$ is as large as we wish otherwise there is nothing to prove. Let $s = \lfloor 3 \log \log y \rfloor$. Let

$$\mathcal{N} = \{ n \in (x, x+y) : n \text{ has } k + j \text{ prime factors in } [A, B], \ 0 \leq j \leq s - 1 \},$$

with the parameters $A = k^2, B = y^{1/(k+s+1)}$, where we take $k \in [c_1 \sqrt{\log y}, c_2 \sqrt{\log y}]$, and $0 < c_1 < c_2$ are two constants to be made more precise later. We will spend some time getting a lower bound on the cardinality of $\mathcal{N}$. For this, observe that for each $n \in \mathcal{N}$ there is a squarefree number $m$ with exactly $k$ distinct prime factors in $[A, B]$ such that $m \mid n$. Clearly, $m \leq y^{k/(k+s+1)}$. Fix such an $m$ and put $\mathcal{N}_m$ for the set of multiples of $m$ in $\mathcal{N}$. To get a lower bound on $\# \mathcal{N}_m$, observe first that the number of multiples of $m$ in $(x, x+y)$ is

$$\geq \left\lfloor \frac{y}{m} \right\rfloor \geq \frac{y}{m} - 1 = \frac{y}{m} \left( 1 + O \left( \frac{m}{y} \right) \right) = \frac{y}{m} \left( 1 + O \left( \frac{1}{\log y} \right) \right).$$

Of course, not all such numbers are in $\mathcal{N}_m$ since some of them might have more than $k + s - 1$ distinct prime factors in $[A, B]$. We next get an upper bound for the number of such “bad” multiples $n$ of $m$. For each such bad $n$, there exists a number $m_1$ having $s$ prime factors in $[A, B]$ and coprime to $m$ such that $mm_1 \mid n$. Note that $mm_1 \leq
For fixed $m$ and $m_1$, the number of such positive integers in $(x, x + y)$ is

$$\leq \left\lfloor \frac{y}{mm_1} \right\rfloor + 1 \leq \frac{2y}{mm_1}.$$ 

Summing up over all possibilities for $m_1$, we get that the number of such $n$ is

$$\leq \frac{2y}{m} \sum_{m_1 \in M} \frac{1}{m_1} \leq \frac{2y}{ms!} \left( \sum_{A \leq p \leq B} \frac{1}{p} \right) = \frac{2yS^s}{ms!},$$

where we put

$$S := \sum_{A \leq p \leq B} \frac{1}{p}.$$ 

Observe that, by Mertens’s formula, we have

$$S = \left( \frac{1}{2} + o(1) \right) \log \log y + O(1)$$

as $y \to \infty$, therefore for $y > y_0$ we have that $S < s/3$. We record that

$$S = \log \log y - \log k - \log \log k - \log 2 + O \left( \frac{1}{\log k} \right).$$

As far as errors go, note that since $s = 3 \log \log y + O(1)$ and $k \asymp \sqrt{\log y}$, we have that

$$\frac{s}{k} \ll \frac{\log k}{k} \ll \frac{1}{\log k}.$$ 

Furthermore, $S = (1/2 + o(1)) \log \log y$ as $y \to \infty$, therefore for $y > y_0$ we have that $S < s/3$. We record that

$$S = \log \log y - \log k - \log \log k - \log 2 + O \left( \frac{1}{\log k} \right).$$

The above arguments show that

$$\#N_m \geq \frac{y}{m} \left( 1 - 2S^s \right) + O \left( \frac{1}{\log y} \right).$$ 

From the elementary estimate $s! \geq (s/e)^s$, we get

$$\frac{2S^s}{s!} \ll \left( \frac{Se}{s} \right)^s \ll \left( \frac{e}{3} \right)^s.$$
and the last number above is $< 1/3$ if $y$ is sufficiently large. Hence, the inequality
\[ \#N_m \geq \frac{y}{2m} \]
holds uniformly in squarefree integers $m$ having $k$ distinct prime factors all in $[A, B]$. We now sum over $m$ and use Lemma 4 to get that
\[ (15) \sum_{m \in M_{k,A,B}} \#N_m \geq \frac{y}{2} \sum_{m \in M_{k,A,B}} \frac{1}{m} \gg \frac{yS^k}{k!} \left( 1 + O \left( \frac{k^2}{S^2 A \log A} \right) \right) \gg \frac{yS^k}{k!} \]
for large $y$, because $A = k^2$, therefore the expression $k^2/(S^2 A \log A)$ is arbitrarily small if $y$ is large. Next let us note that if $n \in N$, then $n$ has $k + j$ distinct prime factors in $[A, B]$ for some $j = 0, 1, \ldots, s - 1$. Thus, the number of possibilities for $m \mid n$ in $M_{k,A,B}$ is
\[ \binom{k + j}{k} \leq \binom{k + s}{s} < \left( e \frac{ek}{s} \right)^s = \exp(O((\log \log y)^2)). \]
Here, we used again the fact that $s! \geq (s/e)^s$. In particular, the sum on the left of (15) counts numbers $n \in N$ and each number is counted at most $\exp(O((\log \log y)^2))$ times. Hence, dividing by this number we get a lower bound on $\#N$ which is
\[ \#N \geq \frac{yS^k}{k!} \exp(O((\log \log y)^2)). \]
If $n \in N$, then there is an $m \in M$ such that $m \mid n$. It now follows, from Lemma 3 that $f(n) \geq f(m) \geq B_k$. Thus,
\[ \frac{1}{y} \sum_{x \leq n \leq x + y} f(n) \geq \frac{1}{y} \sum_{n \in N} f(n) \geq \frac{1}{y} B_k \#N \geq \frac{B_k S^k}{k!} \exp(O((\log \log y)^2)). \]
We now maximize $B_k S^k/k!$ by choosing $k$ appropriately versus $y$. Using Stirling’s formula
\[ k! \sim \left( \frac{k}{e} \right)^k (2\pi k)^{1/2} \]
to estimate $k!$, Lemma 1 as well as estimate (14), we get
\[ (16) \frac{B_k S^k}{k!} \exp(O((\log \log y)^2)) = \exp \left( h(k) + O \left( \frac{k(\log \log k)^2}{(\log k)^2} \right) \right), \]
where the function $h(k)$ is
\[ h(k) = k \log(\log y - \log k - \log \log k - \log 2) - k \log \log k + k \frac{\log \log k}{\log k} + \frac{k}{\log k}. \]
The error term under the exponential in formula \((16)\) comes from the estimate given by Lemma \([1]\) on \(B_k\), estimate \((14)\) which tells us that

\[
k \log S = k \log \left( \log \log y - \log k - \log \log k - \log 2 + O \left( \frac{1}{\log k} \right) \right)
\]

\[
= k \log(\log \log y - \log k - \log \log k - \log 2) + O \left( \frac{k}{(\log k)^2} \right),
\]

because \(\log \log y - \log k - \log \log k - \log 2 \propto \log k\) for our choice of \(k\) versus \(y\), as well as the fact that \((\log \log y)^2 \ll (\log \log k)^2/(\log k)^2\), again by our choice of \(k\) versus \(y\).

We now choose

\[
k = \left\lfloor \frac{1}{\sqrt{2e}} (\log y)^{1/2} \right\rfloor.
\]

Note that with \(c_1 = 1/4\) and \(c_2 = 1/2\) we indeed have that \(k \in [c_1 (\log y)^{1/2}, c_2 (\log y)^{1/2}]\), as promised. Then,

\[
k = \frac{1}{\sqrt{2e}} (\log y)^{1/2} + O(1);
\]

\[
\log k = \frac{1}{2} \log \log y - \log(\sqrt{2e}) + O \left( \frac{1}{\log \log y} \right);
\]

\[
\frac{1}{\log k} = \frac{2}{\log \log y} + O \left( \frac{1}{(\log \log y)^2} \right).
\]

In particular,

\[
\log \log y - \log k - \log \log k - \log 2
\]

\[
= \frac{1}{2} \log \log y + \log(\sqrt{2e}/2) - \log \log k + O \left( \frac{1}{\log \log y} \right)
\]

\[
= \left( \frac{1}{2} \log \log y - \log(\sqrt{2e}) \right) - \log \log k + 1 + O \left( \frac{1}{\sqrt{\log y}} \right)
\]

\[
= \log k - (\log \log k - 1) + O \left( \frac{1}{\sqrt{\log y}} \right),
\]

so that

\[
\log(\log \log y - \log k - \log \log k - \log 2)
\]

\[
= \log \left( \log k - (\log \log k - 1) + O \left( \frac{1}{\sqrt{\log y}} \right) \right)
\]

\[
= \log(\log k - (\log \log k - 1)) + O \left( \frac{1}{k \sqrt{\log y}} \right)
\]

\[
= \log(\log k - (\log \log k - 1)) + O \left( \frac{1}{\log y} \right).
\]
Thus,

\[ k \log \log y - \log k - \log k - \log 2 - k \log \log k = k \log \left( \frac{\log k - (\log \log k - 1)}{\log k} \right) \left( 1 + O \left( \frac{1}{\log y} \right) \right) \]

\[ = k \log \left( 1 - \frac{\log \log k - 1}{\log k} \right) + O \left( \frac{1}{\sqrt{\log y}} \right) \]

\[ = -\frac{k(\log \log k - 1)}{\log k} + O \left( \frac{k(\log \log k)^2}{(\log k)^2} + \frac{1}{k} \right) \]

\[ = -\frac{k \log \log k}{\log k} + \frac{k}{\log k} + O \left( \frac{k(\log \log k)^2}{(\log k)^2} \right). \]

It now follows immediately that

\[ h(k) = k \log \log y - \log k - \log \log k - \log 2 - k \log \log k \]

\[ + \frac{k \log \log k}{\log k} + \frac{k}{\log k} = 2k \frac{\log \log k}{\log k} + O \left( \frac{k(\log \log k)^2}{(\log k)^2} \right). \]

One can in fact check that the above estimate is the maximum of \( h(k) \) as a function of \( k \) when \( y \) is fixed. We will not drag the reader through this computation. Comparing the above estimate with (16), we get that

\[ \frac{B_k S^k}{k!} \exp(\log \log y) \geq \exp \left( 2k \frac{\log \log k}{\log k} + O \left( \frac{k(\log \log k)^2}{(\log k)^2} \right) \right) \]

\[ = \exp \left( \frac{4}{\sqrt{2e}} \left( \log \log y \right)^{1/2} \left( 1 + O \left( \frac{(\log \log y)^2}{\log \log y} \right) \right) \right). \]

We thus get that

\[ \frac{1}{y} \sum_{x \leq n \leq x+y} f(n) \geq \frac{B_k S^k}{k!} \exp(\log \log y) \]

\[ \geq \exp \left( \left( \frac{4}{\sqrt{2e}} \left( \log \log y \right)^{1/2} + O \left( \frac{(\log \log y)^2}{\log \log y} \right) \right) \sqrt{\log y} \log \log y \right), \]

which is what we wanted.

3.3. Proof of Theorem 3. We observe that primes are in \( \mathcal{A} \) as \( f(p) = 1 \) for all prime \( p \). Thus,

\[ \#\mathcal{A}(x) \geq \frac{x}{\log x}. \]

This completes the lower bound part of the theorem. To obtain the upper bound, we cover the set \( \mathcal{A}(x) \) by three subsets \( \mathcal{A}_1, \mathcal{A}_2 \) and \( \mathcal{A}_3 \).
as follows:

\[ A_1 = \{ n \leq x : \Omega(n) > 10 \log \log x \}, \]
\[ A_2 = \left\{ n \leq x : \omega(n) < \frac{\log \log x}{\log \log \log x} \right\}, \]

and

\[ A_3 = \{ n \leq x : n \equiv 0 \pmod{f(n)}, n \notin A_1 \cup A_2 \}. \]

We recall the following bound

\[ \# \{ n \leq x : \Omega(n) = k \} \ll \frac{kx}{2^k} \]

valid uniformly in \( k \) (see, for example, Lemma 13 in [8]). Using the above estimate, we get

\[ (17) \quad \# A_1 \leq x \sum_{k>10 \log \log x} \frac{k}{2^k} \ll \frac{x \log \log x}{2^{10 \log \log x}} = o\left(\frac{x}{\log x}\right) \]

as \( x \to \infty \). To find an upper bound for \( A_2 \), we use the Hardy-Ramanujan bounds (see [6])

\[ \# \{ n \leq x : \omega(n) = k \} \ll \frac{x (\log \log x + c_1)^{k-1}}{\log x(k-1)!} \]

with some positive constant \( c_1 \). Using the elementary estimate \( m! \geq (m/e)^m \) with \( m = k - 1 \), we get

\[ \# \{ n \leq x : \omega(n) = k \} \ll \frac{x (\log \log x + c_2)^{k-1}}{\log x(k-1)!} \]

where \( c_2 = ec_1 \). The right hand side is an increasing function of \( k \) in our range for \( k \) versus \( x \) when \( x \) is large. Since \( k < (\log \log x)/(\log \log \log x) \), we deduce that

\[ (18) \quad \# A_2 \ll \frac{x}{\log x} \left( O(\log \log x) \right)^{\log \log x/\log \log \log x} = \frac{x}{(\log x)^{1+o(1)}} \]

as \( x \to \infty \).

Finally, we estimate \( A_3 \). Each \( n \in A_3 \) can be written as

\[ n = q_1^{\alpha_1} q_2^{\alpha_2} \cdots q_k^{\alpha_k}, \]

where \( q_1, \ldots, q_k \) are distinct primes, \( \alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_k \), \( \alpha_1 + \alpha_2 + \cdots + \alpha_k \leq 10 \log \log x \) and \( k > K := [\log \log x/\log \log \log x] \). Let \( T \) be the set of all such tuples \( (k, \alpha_1, \ldots, \alpha_k) \). For each such \( n \), we have that

\[ f(n) \geq B_K \geq \exp((1 + o(1))K \log K) \geq \exp((1 + o(1)) \log \log x) = (\log x)^{1+o(1)}. \]
The number of tuples \((k, \alpha_1, \ldots, \alpha_k)\) satisfying the above conditions is at most

\[
\# T \ll \log \log x \sum_{n \leq 10 \log \log x} p(n),
\]

where again \(p(n)\) is the partition function of \(n\). Using estimate (2), we get that the cardinality of \(T\) is at most

\[
\# T \ll (\log \log x)^2 \exp(O(\sqrt{\log \log x})) = (\log x)^o(1) \quad \text{as } x \to \infty.
\]

Thus,

\[
(19) \quad \# A_3 \ll \sum_{(k, \alpha_1, \ldots, \alpha_k) \in T} \frac{x}{f(p_1^{\alpha_1} \cdots p_k^{\alpha_k})} \ll \frac{x\# T}{B_K} = \frac{x}{(\log x)^{1+o(1)}}
\]

as \(x \to \infty\). Now inequalities (17), (18) and (19) yield the desired upper bound and complete the proof.

References

[1] R. E. Canfield, P. Erdős and C. Pomerance: On a problem of Oppenheim concerning "factorisatio numerorum". J. Number Theory 17 (1983), 1–28.
[2] R. E. Canfield and C. Pomerance: On the problem of uniqueness for the maximum Stirling number(s) of the second kind. Integers 2 (2002), A1, 13 pp. (electronic).
[3] C. N. Cooper and R. E. Kennedy: Chebyshev’s inequality and natural density. Amer. Math. Monthly 96 (1989), 118–124.
[4] N. G. de Bruijn: Asymptotic methods in analysis. North-Holland Publishing Co., Amsterdam; P. Noordhoff Ltd., Groningen; Interscience Publishers Inc., New York, 1958.
[5] P. Erdős and C. Pomerance: On a theorem of Besicovitch: values of arithmetic functions that divide their arguments. Indian J. Math. 32 (1990), 279–287.
[6] G. H. Hardy and S. Ramanujan: Asymptotic Formulae in Combinatory Analysis. Proc. London Math. Soc. 17 (1918), 75–115.
[7] G. H. Hardy and S. Ramanujan: The normal number of prime factors of an integer. Quart. Journ. Math. (Oxford) 48 (1917), 76–92.
[8] F. Luca and C. Pomerance: Irreducible radical extensions and Euler-function chains, in Combinatorial Number Theory (eds. Landman, Nathanson, Nešetřil, Nowakowski, Pomerance), Proceedings of the "INTEGERS" Conference in honor of R. Graham’s 70th birthday, de Gruyter, 2007, 351–362.
[9] C. A. Spiro: How often does the number of divisors of an integer divide its successor. J. London Math. Soc. (2) 31 (1985), 30–40.
[10] C. Spiro: How often is the number of divisors of \(n\) a divisor of \(n\). J. Number Theory 21 (1985), 81–100.
[11] G. Tenenbaum: Introduction to analytic and probabilistic number theory. Cambridge Studies in Advanced Mathematics, 46. Cambridge University Press, Cambridge, 1995.
Mathematical Institute, UNAM, Ap. Postal 61-3 (Xangari), CP 58089, Morelia, Michoacán, Mexico

E-mail address, Florian Luca: fluca@matmor.unam.mx

Institute of Mathematical Sciences, CIT Campus, Taramani, Chennai 600113, India

E-mail address, Anirban Mukhopadhyay: anirban@imsc.res.in
E-mail address, Kotyada Srinivas: srini@imsc.res.in