Compositions with the Euler and Carmichael Functions

William D. Banks
Department of Mathematics
University of Missouri
Columbia, MO 65211, USA
bbanks@math.missouri.edu

Florian Luca
Instituto de Matemáticas
Universidad Nacional Autónoma de México
C.P. 58089, Morelia, Michoacán, México
fluca@matmor.unam.mx

Filip Saidak
Department of Mathematics
University of Missouri
Columbia, MO 65211, USA
filip@math.missouri.edu

Pantelimon Stănică
Mathematics Department
Auburn University Montgomery
Montgomery, AL 36124, USA
pstanica@mail.aum.edu

September 1, 2005

Abstract

Let φ and λ be the Euler and Carmichael functions, respectively. In this paper, we establish lower and upper bounds for the number of positive integers $n \le x$ such that $\varphi(\lambda(n)) = \lambda(\varphi(n))$. We also study the normal order of the function $\varphi(\lambda(n))/\lambda(\varphi(n))$.

1 Introduction

The *Euler φ-function* (first introduced in [20] of 1760) and the *Carmichael λ-function* (first introduced in [9] of 1910) are two of the most interesting, useful and versatile arithmetic functions that have ever been studied. For a positive integer n, the value $\varphi(n)$ of the Euler function is defined to be the number of natural numbers less than or equal to n and coprime to n. Equivalently,

$$\varphi(n) = \#(\mathbb{Z}/n\mathbb{Z})^* = \prod_{p^\alpha \,\|\, n} p^{\alpha-1}(p-1).$$

For a positive integer n, the value $\lambda(n)$ of the Carmichael function is defined to be the maximal order of any element in the multiplicative group $(\mathbb{Z}/n\mathbb{Z})^*$. More explicitly, for a prime power p^α,

$$\lambda(p^\alpha) = \begin{cases} p^{\alpha-1}(p-1), & \text{if } p \ge 3 \text{ or } \alpha \le 2; \\ 2^{\alpha-2}, & \text{if } p = 2 \text{ and } \alpha \ge 3; \end{cases}$$

and for an arbitrary integer $n \ge 2$ with prime factorization $n = p_1^{\alpha_1} \dots p_k^{\alpha_k}$, one has

$$\lambda(n) = \operatorname{lcm}\big[\lambda(p_1^{\alpha_1}),\, \dots,\, \lambda(p_k^{\alpha_k})\big].$$

Also, $\lambda(1) = 1$. In contrast to φ, the function λ is *not* multiplicative.

Over the years, the Euler and Carmichael functions have been extensively researched in the literature, and quite a large number of results have been obtained concerning the rate of growth and the arithmetical properties of these functions. To gain insight into the nature of these and other arithmetic functions, many subsequent investigations considered compositions (or iterations) of such functions. Let us mention a few relevant examples that have motivated our present investigations.

(1) In 1929, Pillai [41] was the first to study properties of the iterates $\{\varphi^{(k)} : k \ge 1\}$ of the Euler function, where $\varphi^{(1)} = \varphi$, and $\varphi^{(k)} = \varphi \circ \varphi^{(k-1)}$

for $k \geq 2$. Pillai showed that if $W(n) = k$ is the least integer k for which $\varphi^{(k)}(n) = 1$, then

$$\left[\frac{\log 3n}{\log 3}\right] \;\leq\; W(n) \;\leq\; \left[\frac{\log 2n}{\log 2}\right].$$

Later, Shapiro [44], Mills [37], Erdős [14] and Erdős and Hall [16] investigated related questions. Extending some of Shapiro's work [44], the problem of finding integers n with the property that $\varphi^{(k)}(n) \mid n$ was first considered by Hausman [31] in 1982, and these results were later generalized by Halter-Koch and Steindl [28] and by Siva Rama Prasad and Fonseca [46].

(2) For a positive integer n, let $\Omega(n)$ and $\omega(n)$ denote the number of prime factors of n counted with and without multiplicity, respectively, and let $\Delta(n) = \Omega(n) - \omega(n)$.

Generalizing the fundamental theorem of Turán [48] from 1934, in 1984, Erdős and Pomerance [18] and Murty and Murty [39] proved, independently, that both compositions $\Omega(\varphi(n))$ and $\omega(\varphi(n))$ have *normal order* $\frac{1}{2}(\log_2 n)^2$, and they also gave analogues of the *Erdős-Kac theorem* (see [17]) for these functions.[1]

In 1999, Bassily, Kátai and Wijsmuller [7] gave proofs of similar theorems for $\varphi^{(k)}$. In particular, they showed that, as $x \to \infty$,

$$\Delta(\varphi^{(k)}(n)) \;=\; (1 + o(1))\frac{1}{k!}(\log_2 x)^k \log_4 x$$

holds for almost all $n \leq x$, and

$$\lim_{x\to\infty} \frac{1}{x}\#\left\{n \leq x\colon \frac{\Delta(\varphi^{(k)}(n)) - s(x)}{\sqrt{\log_2 x}\log_4 x} < z\right\} = \Phi(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-t^2/2}dt,$$

where $s(x) = (\log_4 x + c + o(1))\log_2 x$. In other words, the quantity

$$\frac{\Delta(\varphi^{(k)}(n)) - s(x)}{\sqrt{\log_2 x}\log_4 x}$$

is normally distributed.

Earlier, in 1990, Erdős, Granville, Pomerance and Spiro [15] proved that, under the *Elliott-Halberstam conjecture*, the normal order of Pillai's function

[1] $\log_k$ denotes the k-th iterate of the natural logarithm.

$W(n)$ is $\Theta \log n$ for some constant Θ, and for every positive integer k, the normal order of $\varphi^{(k)}(n)/\varphi^{(k+1)}(n)$ is $e^{\gamma} k \log_3 n$, where γ is the *Euler-Mascheroni constant*.

(3) In 1944, Alaoglu and Erdős [2] considered the compositions $\varphi \circ \sigma$ and $\sigma \circ \varphi$, where σ is the *sum of divisors function*, and they proved that

$$\liminf_{n\to\infty} \frac{\varphi(\sigma(n))}{n} = 0 \qquad \text{and} \qquad \limsup_{n\to\infty} \frac{\sigma(\varphi(n))}{n} = \infty.$$

Conversely, in 1964, Makowski and Schinzel [35] proved that

$$\limsup_{n\to\infty} \frac{\varphi(\sigma(n))}{n} = \infty \qquad \text{and} \qquad \liminf_{n\to\infty} \frac{\sigma(\varphi(n))}{n} \le \frac{1}{2} + \frac{1}{2^{34}-4},$$

and in 1989, Pomerance [43] showed that

$$\liminf_{n\to\infty} \sigma(\varphi(n))/n > 0.$$

In 1992, Golomb [25] observed that the value of $\sigma(\varphi(n)) - \varphi(\sigma(n))$ takes both positive and negative values infinitely often and asked for the proportion of each. De Koninck and Luca [11] have shown that this function is positive for almost all values of n.

(4) Recently, Martin and Pomerance [36] have studied iterates of the Carmichael function and have shown that the normal order of the function $\log(n/\lambda(\lambda(n)))$ is $(\log_2 n)^2 \log_3 n$. In other words,

$$\lambda(\lambda(n)) = n \exp\big(-(1+o(1))(\log_2 n)^2 \log_3 n\big) \tag{1}$$

as $n \to \infty$ through a set of integers of asymptotic density one.

For a variety of other results with a similar flavor, we refer the reader to [3, 4, 5, 6, 8, 10, 12, 13, 21, 22, 23, 24, 26, 33, 34, 40, 42, 45, 49] and the references contained therein.

In this paper, we initiate the study of the composite functions $\varphi \circ \lambda$ and $\lambda \circ \varphi$ by establishing lower and upper bounds for the counting function of the set

$$\mathcal{A}(x) = \big\{n \le x : \varphi(\lambda(n)) = \lambda(\varphi(n))\big\}.$$

Our main results are the following:

Theorem 1. *There exist positive constants C and x_0 such that the following bound holds for all $x \geq x_0$:*

$$\#\mathcal{A}(x) \geq \exp\left(C\,\frac{\log x}{\log\log\log x}\right).$$

Theorem 2. *The inequality*

$$\#\mathcal{A}(x) \leq \frac{x}{(\log x)^{3/2+o(1)}}$$

holds as $x \to \infty$.

Remark. If we denote $\mathcal{A} = \{n : \varphi(\lambda(n)) = \lambda(\varphi(n))\}$, then Theorem 2 implies that

$$\sum_{n\in\mathcal{A}} \frac{1}{n} < \infty.$$

It is natural to conjecture that the estimate

$$\#\mathcal{A}(x) = \frac{x}{(\log x)^{c+o(1)}} \tag{2}$$

holds for some positive constant c. Since the *Sophie-German primes* (i.e., primes p for which $q = (p-1)/2$ is also prime) are all contained in $\mathcal{A}$, the *Hardy-Littlewood conjectures* (see [29]) would suggest that $c \leq 2$. Taking into account the very special structure of the integers in $\mathcal{A}$, it is also natural to expect the Sophie-German primes to form a subset of $\mathcal{A}$ of positive relative asymptotic density, and we therefore conjecture that (2) holds with $c = 2$.

Although our focus in this paper is primarily on the set $\mathcal{A}$ of positive integers for which the values of $\varphi\circ\lambda$ and $\lambda\circ\varphi$ coincide, we have also been led to consider the related question: Which value is larger, $\varphi(\lambda(n))$ or $\lambda(\varphi(n))$, for a "typical" integer n? In the last section, we study the normal order of the function $\varphi(\lambda(n))/\lambda(\varphi(n))$; our result, which relies heavily on (1), is the following:

Theorem 3. *The estimate*

$$\frac{\varphi(\lambda(n))}{\lambda(\varphi(n))} = \exp\big((1+o(1))(\log\log n)^2 \log\log\log n\big)$$

holds on a set of positive integers n of asymptotic density one.

In particular, one sees that $\varphi(\lambda(n))$ is much larger than $\lambda(\varphi(n))$ for almost all positive integers n.

Acknowledgements. The authors would like to thank the anonymous referee for a careful reading of the manuscript and for useful suggestions. This paper was written during an enjoyable visit by F. L. and P. S. to the University of Missouri–Columbia; these authors wish to express their thanks to that institution for its hospitality and support. Research of W. B. was supported in part by NSF grant DMS-0070628, that of F. L. by grants SEP-CONACYT 37259-E and 37260-E, and that of P. S. by a grant from his institution.

2 Notation

Throughout the paper, we use the Landau symbols 'o' and 'O' and the Vinogradov symbols '$\ll$' and '$\gg$' with the understanding that the implied constants are absolute; we recall that, for positive functions U and V, the notations $U \ll V$, $V \gg U$, and $U = O(V)$ are each equivalent to the assertion that the inequality $U \le cV$ holds for some constant $c > 0$. As usual, $P(n)$ denotes the largest prime factor of $n > 1$, and $\omega(n)$ denotes the number of distinct prime factors of n. Throughout the paper, the letters p, q, and r always denote prime numbers. For a positive real number x, we use $\log x$ to denote maximum of 1 and the natural logarithm of x. For an integer $k \ge 2$, $\log_k x$ denotes the k-th iterate of the function $\log x$. For a positive real number x and a subset $\mathcal{B}$ of the positive integers, we write $\mathcal{B}(x) = \mathcal{B} \cap [1, x]$. Finally, we use $c_0, c_1, \ldots$ to represent positive constants that are absolute.

For the convenience of the reader, we have included a brief index at the end of the paper which contains, in particular, a list of notation for our proof of Theorem 2 below.

3 Coincidences between $\varphi \circ \lambda$ and $\lambda \circ \varphi$

Throughout this section, we focus our study on the set

$$\mathcal{A} = \{n \geq 1 : \varphi(\lambda(n)) = \lambda(\varphi(n))\}.$$

Our goal is establish lower and upper bounds for the counting function $\#\mathcal{A}(x) = \#(\mathcal{A} \cap [1, x])$, where x is a real parameter.

3.1 Lower Bound

Theorem 1. *There exists a positive absolute constant c_0 such that the following bound holds for all $x \geq 2$:*

$$\#\mathcal{A}(x) \geq \exp\left(c_0 \frac{\log x}{\log_3 x}\right).$$

Proof. Observe that, since $\#\mathcal{A}(2) = 2$, it suffices to establish the inequality for all sufficiently large values of x.

For a positive integer n, let $\varphi^{(0)}(n) = n$, and define $\varphi^{(k)}(n)$ inductively by $\varphi^{(k)}(n) = \varphi\left(\varphi^{(k-1)}(n)\right)$ for all $k \geq 1$. Let $\mathcal{P}(n)$ denote the set of *odd* prime factors of the integer $\prod_{k\geq 0} \varphi^{(k)}(n)$. By a result of Pillai [41], the equality $\varphi^{(k)}(n) = 1$ holds for some $k \leq K = \lceil (\log n)/\log 2 \rceil$; consequently,

$$\prod_{p \in \mathcal{P}(n)} p \leq n^{K+1} \leq \exp\left(2(\log n)^2\right). \tag{3}$$

We also have

$$\#\mathcal{P}(n) \leq (K+1) \max_{\ell \leq n}\{\omega(\ell)\} \ll \frac{(\log n)^2}{\log_2 n}. \tag{4}$$

Now let $n \geq 3$ be an odd square-free integer that is coprime to $\varphi^{(k)}(n)$ for all $k \geq 1$, and put

$$M = 2 \prod_{p \in \mathcal{P}(n)} p \qquad \text{and} \qquad N = 2nM\lambda(M).$$

Note that $\mathcal{P}(n) \cup \{2\}$ is precisely the set of primes that divide N. Denoting by $v_p(\cdot)$ the standard p-adic valuation, it is easy to check that

$$v_p(N) = \begin{cases} 2 + \max_{q \in \mathcal{P}(n)}\{v_2(q-1)\}, & \text{if } p = 2; \\ 1 + \max_{q \in \mathcal{P}(n)}\{v_p(q-1)\}, & \text{if } p \in \mathcal{P}(n) \text{ and } p \nmid n; \\ 2, & \text{if } p \mid n. \end{cases}$$

Put

$$\alpha_p = \begin{cases} \max_{q\in\mathcal{P}(n)}\{v_p(q-1)\}, & \text{if } p \in \mathcal{P}(n)\cup\{2\} \text{ and } p \nmid n; \\ 1, & \text{if } p \mid n. \end{cases}$$

Observe that $\alpha_p \geq 1$ for all primes $p \in \mathcal{P}(n)\cup\{2\}$. Since

$$N = 2^{\alpha_2+2}\prod_{p\in\mathcal{P}(n)} p^{\alpha_p+1},$$

we have

$$\begin{aligned} \lambda(N) &= \operatorname{lcm}\left[\lambda(2^{\alpha_2+2}), \lambda(p^{\alpha_p+1}) : p \in \mathcal{P}(n)\right] \\ &= \operatorname{lcm}\left[2^{\alpha_2}, p^{\alpha_p}(p-1) : p \in \mathcal{P}(n)\right] = 2^{\alpha_2}\prod_{p\in\mathcal{P}(n)} p^{\alpha_p}. \end{aligned}$$

Therefore,

$$\varphi\big(\lambda(N)\big) = 2^{\alpha_2-1}\prod_{p\in\mathcal{P}(n)} p^{\alpha_p-1}(p-1) = 2^{\alpha_2+\delta_2-1}\prod_{p\in\mathcal{P}(n)} p^{\alpha_p+\delta_p-1},$$

where we have factored

$$\prod_{p\in\mathcal{P}(n)}(p-1) = 2^{\delta_2}\prod_{p\in\mathcal{P}(n)} p^{\delta_p}.$$

Here, we have used the fact that for each $p \in \mathcal{P}(n)$, the odd prime factors of $p-1$ also lie in $\mathcal{P}(n)$.

On the other hand, we have

$$\varphi(N) = 2^{\alpha_2+1}\prod_{p\in\mathcal{P}(n)} p^{\alpha_p}(p-1) = 2^{\alpha_2+\delta_2+1}\prod_{p\in\mathcal{P}(n)} p^{\alpha_p+\delta_p}.$$

Since $\alpha_2,\ \delta_2 \geq 1$, and $\alpha_p \geq 1$ for all $p \in \mathcal{P}(n)$, it follows that

$$\begin{aligned} \lambda(\varphi(N)) &= \operatorname{lcm}\left[\lambda(2^{\alpha_2+\delta_2+1}), \lambda(p^{\alpha_p+\delta_p}) : p \in \mathcal{P}(n)\right] \\ &= \operatorname{lcm}\left[2^{\alpha_2+\delta_2-1}, p^{\alpha_p+\delta_p-1}(p-1) : p \in \mathcal{P}(n)\right] \\ &= 2^{\alpha_2+\delta_2-1}\prod_{p\in\mathcal{P}(n)} p^{\alpha_p+\delta_p-1} = \varphi(\lambda(N)). \end{aligned}$$

Thus, we have shown that the integer N lies in the set $\mathcal{A}$.

We now use the fact that there exists an absolute constant $c_1 > 0$ such that for all $y > 10$, there exists a positive integer $\ell_y \le y^2$ with the property that

$$\sum_{(p-1)\,|\,\ell_y} 1 \ge \exp\left(c_1 \frac{\log y}{\log_2 y}\right) \tag{5}$$

(see [1], for example). Let $\mathcal{S}$ be the set of odd primes p such that $p-1$ is a divisor of ℓ_y, but p does not lie in $\mathcal{P}(\ell_y)$. Using the inequalities (4) and (5), it follows that

$$\#\mathcal{S} \ge \sum_{\substack{(p-1)\,|\,\ell_y \\ p\ne 2}} 1 - \#\mathcal{P}(\ell_y) \ge s := \exp\left(c_2 \frac{\log y}{\log_2 y}\right) \tag{6}$$

holds with $c_2 = c_1/2$, provided that y is sufficiently large. Replacing $\mathcal{S}$ by one of its subsets, if necessary, we can assume that $\#\mathcal{S} = \lceil s \rceil$. We now set $t = \lfloor \sqrt{s} \rfloor$ and consider subsets $\mathcal{T} \subset \mathcal{S}$ of cardinality t. The number of these subsets is

$$\binom{\lceil s \rceil}{t} \ge \left(\frac{\lceil s \rceil - t}{t}\right)^t = \exp\big((0.5 + o(1)) s^{1/2} \log s\big) \ge \exp\left(c_3 \frac{s^{1/2} \log y}{\log_2 y}\right),$$

where $c_3 = c_2/3$, provided that y is large enough. For each subset $\mathcal{T}$, put $n_{\mathcal{T}} = \prod_{p\in\mathcal{T}} p$. Since

$$\varphi(n_{\mathcal{T}}) = \prod_{p\in\mathcal{T}} (p-1) \,\Big|\, \ell_y^t,$$

it follows that

$$\varphi^{(k)}(n_{\mathcal{T}}) \,\Big|\, \varphi^{(k-1)}(\ell_y^t), \qquad k \ge 1.$$

As $\mathcal{P}(\ell_y^t) = \mathcal{P}(\ell_y)$, we see that $n_{\mathcal{T}}$ is coprime to $\varphi^{(k)}(n_{\mathcal{T}})$ for all $k \ge 1$, and

$$\prod_{p\in\mathcal{P}(n_{\mathcal{T}})} p \qquad \text{divides} \qquad n_{\mathcal{T}} \prod_{p\in\mathcal{P}(\ell_y)} p. \tag{7}$$

The construction given at the beginning of the proof now shows that if

$$M_{\mathcal{T}} = 2 \prod_{p\in\mathcal{P}(n_{\mathcal{T}})} p,$$

then the positive integer

$$N_{\mathcal{T}} = 2 n_{\mathcal{T}} M_{\mathcal{T}} \lambda(M_{\mathcal{T}})$$

lies in the set $\mathcal{A}$. Moreover, it is clear that distinct subsets $\mathcal{T}$ give rise to distinct elements of $\mathcal{A}$ (for if $N_{\mathcal{T}_1} = N_{\mathcal{T}_2}$, then by comparing those parts of $N_{\mathcal{T}_1}$ and $N_{\mathcal{T}_2}$ composed of primes in $\mathcal{S}$, we obtain that $n_{\mathcal{T}_1}^2 = n_{\mathcal{T}_2}^2$, and by unique factorization this leads to $\mathcal{T}_1 = \mathcal{T}_2$). To bound the size of $N_{\mathcal{T}}$, we first use (3) and (7) to estimate

$$M_{\mathcal{T}} \le 2n_{\mathcal{T}} \prod_{p \in \mathcal{P}(\ell_y)} p \le n_{\mathcal{T}} \exp\left(O((\log y)^2)\right).$$

Since $\lambda(M_{\mathcal{T}}) \le M_{\mathcal{T}}$, we can use this bound for $\lambda(M_{\mathcal{T}})$ as well. Also,

$$n_{\mathcal{T}} = \prod_{p \in \mathcal{T}} p \le (y^2+1)^t = \exp\left((2+o(1))s^{1/2}\log y\right).$$

Therefore,

$$N_{\mathcal{T}} \ll n_{\mathcal{T}} M_{\mathcal{T}} \lambda(M_{\mathcal{T}}) \le \exp\left((6+o(1))s^{1/2}\log y + O((\log y)^2)\right)$$

Now, given a large real number x, let y be defined implicitly by the equation $x = \exp\left(7s^{1/2}\log y\right)$, where s is defined as in (6). Then $N_{\mathcal{T}} \le x$ holds for all such subsets $\mathcal{T}$, provided that x is sufficiently large. Since

$$\log_2 x = (0.5+o(1))\log s \gg \frac{\log y}{\log_2 y},$$

it follows that

$$\log_2 y \le (1+o(1))\log_3 x.$$

Therefore, if x is large enough, then

$$\#\mathcal{A}(x) \ge \exp\left(c_3 \frac{s^{1/2}\log y}{\log_2 y}\right) \ge \exp\left(c_0 \frac{\log x}{\log_3 x}\right),$$

where $c_0 = c_3/8$. This completes the proof. □

3.2 Upper Bound

We begin this subsection with a few technical lemmas that are used in our proof of Theorem 2 below.

The following result is a weakened and simplified version of a well-known result of Hildebrand [32] (see, for example, Chapter III.5 in the book by Tenenbaum [47]):

Lemma 1. *Uniformly for* $\exp((\log_2 x)^2) \le y \le x$, *the cardinality* $\Psi(x,y)$ *of the set of smooth numbers*

$$\mathcal{S}(x,y) = \{n \le x : P(n) \le y\}$$

is bounded by

$$\Psi(x,y) \le xu^{-u+o(u)},$$

where $u = (\log x)/(\log y)$.

Lemma 2. *Uniformly for* $x \ge y \ge 2$, *the cardinality of the set*

$$\mathcal{F}(x,y) = \{n \in x : q \mid \gcd(n, \varphi(n)) \text{ for some prime } q > y\}$$

is bounded by

$$\#\mathcal{F}(x,y) \ll \frac{x \log_2 x}{y \log y}.$$

Proof. If n lies in $\mathcal{F}(x,y)$, then either there exists a prime $q > y$ such that $q^2 \mid n$, or there exists a prime $q > y$ and a prime $p \equiv 1 \pmod q$ such that $pq \mid n$. In the first case, the number of such integers $n \le x$ is bounded by

$$\sum_{q>y} \left\lfloor \frac{x}{q^2} \right\rfloor \le x \sum_{q>y} \frac{1}{q^2} \ll \frac{x}{y \log y},$$

and in the second case, the number of such integers $n \le x$ is at most

$$\sum_{q>y} \sum_{\substack{p<x \\ p\equiv 1 \pmod q}} \left\lfloor \frac{x}{pq} \right\rfloor \le x \sum_{q>y} \frac{1}{q} \sum_{\substack{p\le x \\ p\equiv 1 \pmod q}} \frac{1}{p} \ll x \log_2 x \sum_{q>y} \frac{1}{q^2} \ll \frac{x \log_2 x}{y \log y},$$

where we have used the uniform bound

$$\sum_{\substack{p\le x \\ p\equiv 1 \pmod m}} \frac{1}{p} \ll \frac{\log_2 x}{\varphi(m)}, \tag{8}$$

which follows from the inequality (3.1) of [15] (see also Lemma 1 of [7]). The result follows. □

Lemma 3. *If $x \geq 2$ and $w \geq \lfloor 2e \log_2 x \rfloor + 1$, then the cardinality of the set*

$$\mathcal{G}(x, w) = \{n \in x : \omega(n) > w\}$$

is bounded by

$$\#\mathcal{G}(x, y) \leq \frac{x}{\log x} 2^{-w(1+o(1))}.$$

Proof. By results of Hardy and Ramanujan [30], the number of positive integers $n \leq x$ for which $\omega(n) = k$ is bounded above by

$$\frac{x}{\log x} \cdot \frac{1}{(k-1)!} \cdot (\log_2 x + O(1))^{k-1} \leq \frac{x}{\log x} \left(\frac{e \log_2 x + O(1)}{k-1} \right)^{k-1}.$$

In the above inequality (and in many others to follow), we have used Stirling's formula to conclude that $k! \geq (k/e)^k$ holds for all positive integers k. In particular, if $w \geq \lfloor 2e \log_2 x \rfloor + 1$, then

$$\#\mathcal{G}(x, w) \leq \frac{x}{\log x} \sum_{k>w} (0.5 + o(1))^k = \frac{x}{\log x} 2^{-w(1+o(1))},$$

which is the desired estimate. □

Theorem 2. *The inequality*

$$\#\mathcal{A}(x) \leq \frac{x}{(\log x)^{3/2+o(1)}}$$

holds as $x \to \infty$.

Proof. Our strategy is to express $\mathcal{A}(x)$ as a union of boundedly many subsets, each of which has a cardinality bounded above by $x(\log x)^{-3/2+o(1)}$.

Let x be a large positive real number. The first five subsets that we consider are the following:

(*i*) The subset

$$\mathcal{A}_1(x) = \{n \in \mathcal{A}(x) : P(n) \leq y_1\},$$

where

$$y_1 = \exp\left(\frac{\log x \log_3 x}{3 \log_2 x} \right).$$

Since $\mathcal{A}_1(x) \subset \mathcal{S}(x, y_1)$, Lemma 1 immediately implies that

$$\#\mathcal{A}_1(x) \le \frac{x}{(\log x)^{3+o(1)}}. \tag{9}$$

Note that if x is large, then $\mathcal{A}_1(x)$ contains all powers of 2 which are smaller than x.

(*ii*) The subset

$$\mathcal{A}_2(x) = \{n \le x : q \mid \gcd(n, \varphi(n)) \text{ or } q^2 \mid \varphi(n) \text{ for some } q > y_2 \text{ prime}\}$$

where $y_2 = (\log x)^2$. Write $\mathcal{A}_2(x) = \mathcal{A}_2'(x) \cup \mathcal{A}_2''(x)$, where $\mathcal{A}_2'(x)$ consists of those $n \in \mathcal{A}_2(x)$ such that $q \mid \gcd(n, \varphi(n))$ for some prime $q > y_2$, and $\mathcal{A}_2''(x)$ consists of the remaining $n \in \mathcal{A}_2(x)$. Since $\mathcal{A}_2'(x)$ is a subset of $\mathcal{F}(x, y_2)$, Lemma 2 provides the estimate

$$\#\mathcal{A}_2'(x) \ll \frac{x}{(\log x)^2}. \tag{10}$$

Now suppose that $n \in \mathcal{A}_2''(x)$. Then there exist a prime $q > y_2$ and either a prime factor p of n with $p \equiv 1 \pmod{q^2}$, or two prime factors $p_1 < p_2$ of n such that $q \mid \gcd(p_1 - 1, p_2 - 1)$. Fix q, and p in the first case, and p_1 and p_2 in the second case, respectively. Then, the number of possible multiples $n \le x$ of p is $\lfloor x/p \rfloor \le x/p$ in the first case, while the number of possible multiples $n \le x$ of $p_1 p_2$ is $\lfloor x/p_1 p_2 \rfloor \le x/p_1 p_2$ in the second case. Therefore,

$$\begin{aligned}
\#\mathcal{A}_2''(x) &\le x \left(\sum_{q > y_2} \sum_{\substack{p \le x \\ p \equiv 1 \pmod{q^2}}} \frac{1}{p} + \sum_{q > y_2} \sum_{\substack{p_1 < p_2 \le x \\ p_i \equiv 1 \pmod{q} \text{ for } i=1,2}} \frac{1}{p_1 p_2} \right) \\
&\le x \left(\sum_{q > y_2} \sum_{\substack{p \le x \\ p \equiv 1 \pmod{q^2}}} \frac{1}{p} + \sum_{q > y_2} \frac{1}{2} \left(\sum_{\substack{p \le x \\ p \equiv 1 \pmod{q}}} \frac{1}{p} \right)^2 \right) \\
&\ll x (\log_2 x)^2 \sum_{q > y_2} \frac{1}{q^2} \ll \frac{x (\log_2 x)^2}{y_2} = \frac{x (\log_2 x)^2}{(\log x)^2},
\end{aligned} \tag{11}$$

where we have used the estimate (8) twice. The estimates (10) and (11) immediately imply that

$$\#\mathcal{A}_2(x) \le \frac{x}{(\log x)^{2+o(1)}}. \tag{12}$$

(*iii*) The subset

$$\mathcal{A}_3(x) = \{n \in \mathcal{A}(x) : \omega(n) > w_1\},$$

where

$$c_1 = 2e \qquad \text{and} \qquad w_1 = \lfloor c_1 \log_2 x \rfloor + 1.$$

Since $\mathcal{A}_3(x) \subset \mathcal{G}(x, w_1)$, Lemma 3 provides the upper bound

$$\#\mathcal{A}_3(x) \le \frac{x}{\log x}\, 2^{-w_1(1+o(1))} = \frac{x}{(\log x)^{1+2e\log 2+o(1)}}. \tag{13}$$

(*iv*) The subset

$$\mathcal{A}_4(x) = \{n \in \mathcal{A}(x) : 2^{w_2} \mid n \text{ or } 2^{w_2} \mid (p-1) \text{ for some prime } p \mid n\},$$

where

$$c_2 = 2/\log 2 \qquad \text{and} \qquad w_2 = \lfloor c_2 \log_2 x \rfloor.$$

Clearly,

$$\begin{aligned} \#\mathcal{A}_4(x) &\le \left\lfloor \frac{x}{2^{w_2}} \right\rfloor + \sum_{\substack{p \le x \\ p \equiv 1 \pmod{2^{w_2}}}} \left\lfloor \frac{x}{p} \right\rfloor \\ &\le \frac{x}{2^{w_2}} + x \sum_{\substack{p \le x \\ p \equiv 1 \pmod{2^{w_2}}}} \frac{1}{p} \\ &\ll \frac{x \log_2 x}{2^{w_2}} = \frac{x}{(\log x)^{2+o(1)}}, \end{aligned} \tag{14}$$

where we have used (8) in the last step.

For the remaining subsets of $\mathcal{A}(x)$, our estimates are presented as a series of technical lemmas. We begin with the following:

Lemma 4. *Let*

$$\mathcal{A}_5(x) = \Big\{ n \in \mathcal{A}(x) \setminus \big(\cup_{j=1}^{4} \mathcal{A}_j(x)\big) : F(n) > w_3 \Big\},$$

where

$$F(n) = v_2(n) + \sum_{p \,|\, n} v_2(p-1) \tag{15}$$

for every positive integer n, and

$$w_3 = c_3 \log_2 x \log_3 x, \qquad c_3 = \frac{c_1 + 1}{\log 2}.$$

If x is sufficiently large, then

$$\#\mathcal{A}_5(x) \le \frac{x}{(\log x)^2}. \tag{16}$$

Proof. For each integer $n \in \mathcal{A}_5(x)$, we have:

- $P(n) > y_1$ and $P(n)^2 \nmid n$;
- $\omega(n) \le w_1$;
- $2^{w_2} \nmid n$, and $2^{w_2} \nmid (p-1)$ for every prime p dividing n;
- n is not a power of 2;
- $F(n) > w_3$.

Let $n = 2^\alpha \prod_{\ell=1}^{k} p_\ell^{\beta_\ell}$ be the prime factorization of n, where the primes p_ℓ are odd and distinct; reordering the odd primes, if necessary, we can assume that $v_2(p_\ell - 1)$ is a nondecreasing function of ℓ. Then there exists an integer $t \ge 1$, integers $1 \le \alpha_1 < \cdots < \alpha_t$, and integers $\kappa_1, \ldots, \kappa_t \ge 1$ with the following three properties:

(*i*) $2^{\alpha_j} \mid (p_\ell - 1)$ if $1 \le j \le t$ and $\kappa_1 + \cdots + \kappa_{j-1} < \ell \le \kappa_1 + \cdots + \kappa_j$;

(*ii*) The inequality

$$\sum_{j=1}^{t} \kappa_j = k = \omega(n) - \delta \le c_1 \log_2 x$$

holds, where $\delta = 0$ if $\alpha = 0$, and $\delta = 1$ otherwise;

(*iii*) The equality

$$\alpha + \sum_{j=1}^{t} \kappa_j \alpha_j = \lfloor c_3 \log_2 x \log_3 x \rfloor$$

holds.

Here, the integers α_i are all the possible values of $v_2(p-1)$ as p runs over the odd prime factors of n, arranged increasingly, and κ_i is the *multiplicity* with which α_i occurs (i.e., the number of prime factors p of n such that $\alpha_i = v_2(p-1)$).

Let $\mathcal{D}$ be the set of all $(2t+2)$-tuples $(t, \alpha, \alpha_1, \dots, \alpha_t, \kappa_1, \dots, \kappa_t)$ for which these properties hold for some $n \in \mathcal{A}_5(x)$. Clearly,

$$\begin{aligned} \#\mathcal{D} &\le \sum_{0 \le \alpha < N \le c_2 \log_2 x \log_3 x} p(N-\alpha) \\ &\le \exp\left(O\left((\log_2 x \log_3 x)^{1/2}\right)\right) = (\log x)^{o(1)}, \end{aligned}$$

where for a positive integer m, we have used $p(m)$ to denote the number of partitions of m.

Now fix one such $(2t+2)$-tuple in $\mathcal{D}$, and suppose that $n \in \mathcal{A}_5(x)$ is of this type. Since the prime $P = P(n)$ is one of the primes p_j, and $P^2 \nmid n$, we can write $n = Pm$, where m satisfies the analogue of (i) with the data

$$(t-1, \alpha, \alpha_1, \dots, \widehat{\alpha}_j, \dots, \alpha_t, \kappa_1, \dots, \widehat{\kappa}_j, \dots, \kappa_t). \tag{17}$$

Here, the *hat* symbol indicates that the entry has been omitted. Clearly, $y_1 < P \le x/m$, and P lies in the arithmetic progression $1 \pmod{2^{\alpha_j}}$; hence, the number of such primes is

$$\ll \frac{x}{2^{\alpha_j} m \log(x/(2^{\alpha_j} m))}$$

(see [38]). Since $2^{\alpha_j} < 2^{w_2} \le (\log x)^2$, the inequalities

$$x/(2^{\alpha_j} m) > y_1/2^{\alpha_j} > y_1^{1/2}$$

hold, and therefore

$$\log(x/(2^{\alpha_j} m)) \gg \log y_1 \gg \frac{\log x \log_3 x}{\log_2 x}.$$

Consequently, for a fixed value of m, the number of such primes P is

$$\ll \frac{x\log_2 x}{\log x}\cdot\frac{1}{2^{\alpha_j}}\cdot\frac{1}{m} = \frac{x\log_2 x}{\log x}\cdot\frac{1}{2^{\alpha+\alpha_j}}\prod_{\substack{1\le \ell\le k\\ \ell\neq \kappa_1+\cdots+\kappa_{j-1}+1}}\frac{1}{p_\ell^{\beta_\ell}}.$$

We now sum up the above inequality over all possible integers m of type (17) (with j fixed) and deduce that the corresponding contribution to $\#\mathcal{A}_5(x)$ is

$$\begin{aligned}
&\ll \frac{x\log_2 x}{\log x}\cdot\frac{1}{2^{\alpha+\alpha_j}}\prod_{\substack{1\le i\le t\\ i\neq j}}\frac{1}{\kappa_i!}\left(\sum_{\substack{p\le x\\ p\equiv 1 \pmod{2^{\alpha_i}}}}\sum_{\beta\ge 1}\frac{1}{p^\beta}\right)^{\kappa_i}\\
&\le \frac{x\log_2 x}{\log x}\cdot\frac{1}{2^{\alpha+\sum_{j=1}^t\alpha_j\kappa_j}}\cdot(c_4\log_2 x)^{\sum_{j=1}^t\kappa_j}. \qquad (18)
\end{aligned}$$

Here, c_4 is an absolute constant for which the inequality

$$\sum_{\substack{1\le p\le x\\ p\equiv 1 \pmod{2^\alpha}}}\sum_{\beta\ge 1}\frac{1}{p^\beta}\le\frac{c_4\log_2 x}{2^\alpha}$$

holds for all sufficiently large x and uniformly in $\alpha\le w_2$ (see again (8)). Since

$$\begin{aligned}
(c_4\log_2 x)^{\sum_{j=1}^t\kappa_j} &\le \exp\left(c_1\log_2 x\log(c_4\log_2 x)\right)\\
&= \exp\left((c_1+o(1))\log_2 x\log_3 x\right),
\end{aligned}$$

and

$$\begin{aligned}
2^{\alpha+\sum_{j=1}^t\alpha_j\kappa_j} &= \exp\left(\lfloor c_3\log_2 x\log_3 x\rfloor\log 2\right)\\
&= \exp\left((c_1+1+o(1))\log_2 x\log_3 x\right),
\end{aligned}$$

we find that the expression (18) is bounded above by

$$\frac{x}{\exp\left((1+o(1))\log_2 x\log_3 x\right)}.$$

Summing this over all possible choices of $j\in\{1,\dots,t\}$, then over all possible $(2t+2)$-tuples in $\mathcal{D}$, we derive the estimate

$$\#\mathcal{A}_5(x)\ll\log_2 x\cdot\#\mathcal{D}\cdot\frac{x}{\exp\left((1+o(1))\log_2 x\log_3 x\right)}\le\frac{x}{(\log x)^2}$$

for all sufficiently large x. □

Now fix $n \in \mathcal{A}(x) \setminus \big(\cup_{j=1}^5 \mathcal{A}_j(x)\big)$, and let $n = 2^\alpha \prod_{\ell=1}^k p_\ell^{\beta_\ell}$ be its prime factorization. Let $\mathcal{P} = \{p_1, \dots, p_k\}$ be the set of odd primes that divide n, $\mathcal{P}_1 = \{p_1, \dots, p_{k_1}\}$ be the subset of $\mathcal{P}$ consisting of the odd primes p that divide $\gcd(n, \varphi(n))$, and $\mathcal{Q}$ be the set of odd prime factors of $\varphi(n)$ that do not divide n. Since $n \notin \mathcal{A}_2(x)$, every prime $p \in \mathcal{P}_1$ satisfies the bound $p \le y_2 = (\log x)^2$; thus, if $q > y_2$ and $q \mid \varphi(n)$, then $q \in \mathcal{Q}$. Let $q_1, \dots, q_s$ be the primes in $\mathcal{Q}$. Factoring

$$p_\ell - 1 = 2^{\alpha_\ell} \prod_{i=1}^{k_1} p_i^{\gamma_{i,\ell}} \prod_{j=1}^{s} q_j^{\delta_{j,\ell}}, \qquad \ell = 1, \dots, k,$$

we derive that

$$\varphi(n) = \varphi(2^\alpha) \cdot 2^{\sum_{\ell=1}^k \alpha_\ell} \prod_{i=1}^{k_1} p_i^{\beta_i - 1 + \sum_{\ell=1}^k \gamma_{i,\ell}} \prod_{j=1}^{s} q_j^{\sum_{\ell=1}^k \delta_{j,\ell}}.$$

Defining $\mu_j = v_2(q_j - 1)$ for $j = 1, \dots, s$, it follows that

$$\begin{aligned} v_2\left(\lambda(\varphi(n))\right) &= \max\Big\{ v_2\left(\lambda\left(\varphi(2^\alpha) \cdot 2^{\sum_{\ell=1}^k \alpha_\ell}\right)\right), \max_{1 \le i \le k_1} \{\alpha_i - 1\}, \max_{1 \le j \le s} \{\mu_j\} \Big\} \\ &= \max\Big\{ \alpha + \sum_{\ell=1}^k \alpha_\ell, \max_{1 \le j \le s} \{\mu_j\} \Big\} + O(1), \end{aligned}$$

The preceding formula may be rewritten in the form

$$v_2\left(\lambda(\varphi(n))\right) = \max\Big\{ F(n), \max_{q \in \mathcal{Q}} \{v_2(q-1)\} \Big\} + O(1). \tag{19}$$

On the other hand, we have

$$\lambda(n) = \max\Big\{ \lambda(2^\alpha), 2^{\max_\ell\{\alpha_\ell\}} \Big\} \prod_{i=1}^{k_1} p_i^{\max\{\beta_i - 1, \max_\ell\{\gamma_{i,\ell}\}\}} \prod_{j=1}^{s} q_j^{\max_\ell\{\delta_{j,\ell}\}},$$

where the maxima in the exponents are taken over $\ell \in \{1, \dots, k\}$. From the preceding relation, we see that $v_2\left(\varphi(\lambda(n))\right)$ is equal to

$$\begin{aligned} &v_2\Big(\max\Big\{\varphi(\lambda(2^\alpha)), 2^{\max_\ell\{\alpha_\ell - 1\}}\Big\}\Big) + \sum_{p \in \mathcal{P}_1} v_2(p-1) + \sum_{q \in \mathcal{Q}} v_2(q-1) \\ &\quad = \max\{v_2(n), \max_{p \mid n} \{v_2(p-1)\}\} + \sum_{p \in \mathcal{P}_1} v_2(p-1) + \sum_{q \in \mathcal{Q}} v_2(q-1) + O(1). \end{aligned}$$

Combining this result with (19) and the fact that $v_2(\lambda(\varphi(n))) = v_2(\varphi(\lambda(n)))$, we obtain:

$$\begin{aligned}\max\{v_2(n), \max_{p\,|\,n}\{v_2(p-1)\}\} + \sum_{p\in\mathcal{P}_1} v_2(p-1) + \sum_{q\in\mathcal{Q}} v_2(q-1) \\ = \max\Big\{F(n), \max_{q\in\mathcal{Q}}\{v_2(q-1)\}\Big\} + O(1).\end{aligned} \tag{20}$$

We now define $\mathcal{A}_6(x)$ be the set of those integers $n \in \mathcal{A}(x)\setminus\big(\cup_{j=1}^5 \mathcal{A}_j(x)\big)$ for which the maximum on the right hand side of (20) is *not* achieved with the term $F(n)$.

Lemma 5. *We have*

$$\#\mathcal{A}_6(x) \ll \frac{x}{(\log x)^{2+o(1)}}. \tag{21}$$

Proof. If $n \in \mathcal{A}_6(x)$, there exists a prime $q' \in \mathcal{Q}$ such that the maximum on the right hand side of (20) is achieved with the term $v(q'-1)$, and it follows that

$$\max\{v_2(n), \max_{p\,|\,n}\{v_2(p-1)\}\} + \sum_{p\in\mathcal{P}_1} v_2(p-1) + \sum_{\substack{q\in\mathcal{Q}\\ q\neq q'}} v_2(q-1) \ll 1,$$

which implies, in particular, that $\#\mathcal{P}_1 + \#\mathcal{Q} \ll 1$. Hence, there exists an absolute constant $c_5 > 0$ such that for every $n \in \mathcal{A}_6(x)$, the inequality $\omega(p-1) \le c_5$ holds for all prime factors p of n.

Let $\mathcal{R}_1 = \{p : \omega(p-1) \le c_5\}$. Our first step is to establish the estimate:

$$\#\mathcal{R}_1(x) \ll \frac{x(\log_2 x)^{c_5+1}}{(\log x)^2} \tag{22}$$

For this, let $p \in \mathcal{R}_1(x)$. We may assume that $P(p-1) > y_1$, for the number of primes $p \le x$ with $P(p-1) \le y_1$ is no more than

$$\Psi(x, y_1) \le x\exp\left(-(1+o(1))u\log u\right) \ll \frac{x}{(\log x)^2}. \tag{23}$$

We may also assume that there does not exist a prime $q > y_3 = \log x$ such that $q^2 \mid p-1$. Indeed, the number of primes $p \le x$ for which $q^2 \mid p-1$ for some $q > y_3$ is no more than

$$\sum_{y_3<q<x^{1/2}} \pi(x; q^2, 1) \le \sum_{y_3\le x\le x^{1/3}} \pi(x; q^2, 1) + \sum_{x^{1/3}<q<x^{1/2}} \pi(x; q^2, 1) = S_1 + S_2,$$

say. For the sum S_1, we apply the Montgomery-Vaughan upper bound on the number of primes in an arithmetical progression (see [38]) to conclude that

$$S_1 \le \sum_{y_3 < q \le x^{1/3}} \frac{2x}{q^2 \log(x/q^2)} \le \frac{6x}{\log x} \sum_{q > y_3} \frac{1}{q^2} \ll \frac{x}{\log x} \cdot \frac{1}{y_3} \ll \frac{x}{(\log x)^2}, \quad (24)$$

and for the sum S_2 we need only the trivial fact that $\pi(x, q^2, 1) \le x/q^2$ to derive the bound:

$$S_2 \le x \sum_{q > x^{1/3}} \frac{1}{q^2} \ll \frac{x^{2/3}}{\log x} \ll \frac{x}{(\log x)^2}. \quad (25)$$

Thus, we may assume that $p - 1 = Pm$, where $P > \max\{P(m), y_1\}$. Since $\omega(p-1) \le c_5$, it follows that $\omega(m) \le c_5 - 1$. Fixing one such number m, we apply Brun's method (see [27]) to deduce that the number of possibilities for the prime p is

$$\ll \frac{x}{\varphi(m)(\log(x/m))^2}.$$

Since $x/m > y_1$, we have

$$\log(x/m) > \log y_1 \gg \frac{\log x}{\log_2 x}. \quad (26)$$

Hence, the number of possibilities for $p \le x$ when m is fixed, is

$$\ll \frac{x(\log_2 x)^2}{(\log x)^2} \cdot \frac{1}{\varphi(m)}.$$

Summing this inequality over all m with $\omega(m) \le c_5 - 1$, and using the inequalities (23), (24), and (25), we obtain that

$$\#\mathcal{R}_1(x) \le \frac{x(\log_2 x)^2}{(\log x)^2} \sum_{\substack{m < x/y \\ \omega(m) \le c_5 - 1}} \frac{1}{\varphi(m)} + O\left(\frac{x}{(\log x)^2}\right). \quad (27)$$

Now, since

$$\begin{aligned} \sum_{\substack{m < x/y \\ \omega(m) \le c_5 - 1}} \frac{1}{\varphi(m)} &\le 1 + \sum_{k=1}^{c_5 - 1} \frac{1}{k!} \left(\sum_{q \le x} \sum_{\gamma \ge 1} \frac{1}{q^{\gamma - 1}(q-1)} \right)^k \\ &\ll 1 + \sum_{k=1}^{c_5 - 1} \frac{1}{k!} \cdot (\log_2 x + O(1))^k \ll (\log_2 x)^{c_5 - 1}, \end{aligned}$$

the inequality (27) implies the estimate (22).

Using (22), it follows that

$$\sum_{p\in\mathcal{R}_1}\sum_{\gamma\ge 1}\frac{1}{p^\gamma}=c_6$$

is a constant; thus, writing $\mathcal{M}_1$ for the set of positive integers m composed from the primes in $\mathcal{R}_1$, we have

$$\sum_{m\in\mathcal{M}_1}\frac{1}{m}=\prod_{p\in\mathcal{R}_1}\left(1+\sum_{\gamma\ge 1}\frac{1}{p^\gamma}\right)<\exp\left(\sum_{p\in\mathcal{R}_1}\sum_{\gamma\ge 1}\frac{1}{p^\gamma}\right)=\exp(c_6).$$

For any $n\in\mathcal{A}_6(x)$, we can write $n=Pm$, where $P\in\mathcal{R}_1(x/m)$ satisfies $P>\max\{P(m),y_1\}$, and $m\in\mathcal{M}_1(x)$. Let m be fixed. According to (22), the prime $P\in\mathcal{R}_1(x/m)$ can be chosen in at most

$$\ll\frac{x\,(\log_2(x/m))^{c_5+1}}{m(\log(x/m))^2}$$

different ways. By inequality (26), the number of possibilities for P is

$$\ll\frac{x(\log_2 x)^{c_5+3}}{(\log x)^2}\cdot\frac{1}{m}.$$

Summing this inequality over all $m\in\mathcal{M}_1(x)$, we derive that

$$\begin{aligned}\#\mathcal{A}_6(x)&\ll\frac{x(\log_2 x)^{c_5+3}}{(\log x)^2}\sum_{m\in\mathcal{M}_1(x)}\frac{1}{m}\\&\ll\frac{x(\log_2 x)^{c_5+3}}{(\log x)^2}=\frac{x}{(\log x)^{2+o(1)}},\end{aligned}$$

which finishes the proof. □

Now let n be an integer in $\mathcal{A}(x)\setminus\big(\cup_{j=1}^6\mathcal{A}_j(x)\big)$. Then, by (20), we have

$$\begin{aligned}\max\{v_2(n),\max_{p\,|\,n}\{v_2(p-1)\}\}&+\sum_{p\in\mathcal{P}_1}v_2(p-1)+\sum_{q\in\mathcal{Q}}v_2(q-1)\\&=F(n)+O(1).\end{aligned}$$

In particular, it follows that

$$\sum_{q\in\mathcal{Q}} v_2(q-1) \le \sum_{p\in\mathcal{P}} v_2(p-1) + O(1). \tag{28}$$

Since, for every prime factor $q > y_2$ of $\varphi(n)$, there exists a unique $p \mid n$ such that $q \mid p-1$ (because $n \notin \mathcal{A}_2(x)$), from (28) we deduce that

$$\sum_{p\,|\,n} \sum_{\substack{q>y_2\\ q\,|\,p-1}} v_2(q-1) \le \sum_{p\,|\,n} v_2(p-1) + O(1). \tag{29}$$

For any real number $z > 0$, let

$$f(z) = \exp\left(\exp\left(\frac{\log_2 z}{(\log_3 z)^2}\right)\right).$$

For any real number z and positive integer n, we write $\omega_{\le z}(n)$ and $\omega_{>z}(n)$ for the number of distinct prime factors of n that are $\le z$ and $> z$, respectively. Consider the following set of prime numbers:

$$\mathcal{R}_2 = \{p : \omega_{>f(p)}(p-1) \le \log_2 p/\log_3 p\}.$$

We claim that the estimate

$$\#\mathcal{R}_2(x) \le \frac{x}{(\log x)^{2+o(1)}} \tag{30}$$

holds. Indeed, let $p \in \mathcal{R}_2(x)$ be fixed. Since $\pi(x/\log x) \ll x/(\log x)^2$, we can assume that $p > x/\log x$. As in the above estimate for the counting function $\#\mathcal{R}_1(x)$ of $\mathcal{R}_1$, we can assume that $P = P(p-1) > y_1$ and that q^2 does not divide $p-1$ for any prime $q > y_3 = \log x$, since the size of the exceptional set is $\ll x/(\log x)^2$. Finally, we can assume that $\omega(p-1) \le w_1$, for otherwise $p-1 \in \mathcal{A}_3(x)$, and we have already seen that $\#\mathcal{A}_3(x) \ll x/(\log x)^2$. Now, write $p-1 = Pm$, where $P = P(p-1) > \max\{P(m), y_1\}$, m satisfies further $\omega(m) \le w_1 = c_1 \log_2 x$, and let $\mathcal{M}_2(x)$ be the set of all integers m obtained in this way. For fixed $m \in \mathcal{M}_2(x)$, we apply Brun's method once again, which shows that the number of possibilities for p is

$$\ll \frac{x}{\varphi(m)(\log(x/m))^2} \ll \frac{x(\log_2 x)^3}{(\log x)^2}\cdot\frac{1}{m}$$

(for the second estimate, we use (26) and the bound $\varphi(m) \gg m/\log_2 m$). Clearly, for $m \in \mathcal{M}_2(x)$, we have $\omega_{>f(x)}(m) < \log_2 x/\log_3 x$. Thus, if k_2 and k_3 denote the number of prime factors of m which are $\le f(x)$ and $> f(x)$, respectively, then $k_2 \le w_1$, and $k_3 \le w_4 = \lfloor \log_2 x/\log_3 x \rfloor$. Therefore,

$$\#\mathcal{R}_2(x) \le \frac{x(\log_2 x)^3}{(\log x)^2} \sum_{m \in \mathcal{M}_2(x)} \frac{1}{m} + O\left(\frac{x}{(\log x)^2}\right), \tag{31}$$

and

$$\sum_{m \in \mathcal{M}_2(x)} \frac{1}{m} \le \sum_{\substack{k_2 \le w_1 \\ k_3 \le w_4}} \frac{1}{k_2!} \left(\sum_{p \le f(x)} \sum_{\beta \ge 1} \frac{1}{p^\beta}\right)^{k_2} \cdot \frac{1}{k_3!} \left(\sum_{p \le x} \sum_{\beta \ge 1} \frac{1}{p^\beta}\right)^{k_3}.$$

Note that for $k_2 \le w_1$, using Stirling's formula, we have

$$\begin{aligned} \frac{1}{k_2!} \left(\sum_{p \le f(x)} \sum_{\beta \ge 1} \frac{1}{p^\beta}\right)^{k_2} &\ll \left(\frac{e \log_2 f(x) + O(1)}{k_2}\right)^{k_2} \\ &= \left(\frac{e \log_2 x/(\log_3 x)^2 + O(1)}{k_2}\right)^{k_2} \\ &\le \exp\left(\frac{e \log_2 x}{(\log_3 x)^2} + O(1)\right) = (\log x)^{o(1)}, \end{aligned}$$

uniformly in $k_2 \le w_1$. Similarly, for $k_3 \le w_4$, we have

$$\begin{aligned} \frac{1}{k_3!} \left(\sum_{p \le x} \sum_{\beta \ge 1} \frac{1}{p^\beta}\right)^{k_3} &\le \left(\frac{e \log_2 x + O(1)}{k_3}\right)^{k_3} \ll \left(\frac{e \log_2 x + O(1)}{w_4}\right)^{w_4} \\ &\ll \exp\left(\frac{\log_2 x \log_4 x}{\log_3 x}\right) = (\log x)^{o(1)}. \end{aligned}$$

Since the pair $(k_2,\ k_3)$ can be chosen in at most $O((\log_2 x)^2)$ different ways, we get that

$$\sum_{m \in \mathcal{M}_2(x)} \frac{1}{m} \ll (\log x)^{o(1)}, \tag{32}$$

which together with (31) implies our claim (30).

In particular, it follows from (30) that

$$\sum_{p\in\mathcal{R}_2}\frac{1}{p}=c_7.$$

Now, let $n\in\mathcal{A}(x)\setminus\left(\cup_{j=1}^{6}\mathcal{A}_j(x)\right)$. Using (29) together with the fact that $n\notin\mathcal{A}_5(x)$, we see that

$$\begin{aligned}\sum_{p\,|\,n}\omega_{>y_2}(p-1) &\le \sum_{p\,|\,n}\sum_{\substack{q>y_2\\ q\,|\,p-1}} v_2(q-1)\le\sum_{p\,|\,n}v_2(p-1)+O(1)\\ &\le w_3+O(1)\ll\log_2 x\log_3 x. \qquad (33)\end{aligned}$$

Put

$$y_4=\exp\left(\exp\left((\log_3 x)^3\right)\right).$$

Clearly, every prime factor p of n satisfies precisely one of the following properties:

(i) $p\in\mathcal{R}_2$ and $p\ge y_4$;

(ii) $p<y_4$;

(iii) p lies in the interval $[y_4,x]$ but not in $\mathcal{R}_2$.

Suppose that n has k_4 primes of type (i), k_5 primes of type (ii), and k_6 primes of type (iii). Note that

$$f(y_4)=\exp\left(\exp\left(\frac{\log_2 y_4}{(\log_3 y_4)^2}\right)\right)=\exp\left(\exp\left(\frac{(\log_3 x)^3}{9(\log_4 x)^2}\right)\right)>y_2$$

if x is sufficiently large. Thus, primes p of type (iii) have the property that

$$\omega_{>y_2}(p-1)\ge\omega_{>f(p)}(p-1)>\frac{\log_2 p}{\log_3 p}\ge\frac{\log_2 y_4}{\log_3 y_4}=\frac{(\log_3 x)^3}{9(\log_4 x)^2},$$

From (33), it follows that

$$k_6\cdot\frac{(\log_3 x)^3}{(\log_4 x)^2}\ll\log_2 x\log_3 x,$$

and therefore

$$k_6\ll\frac{\log_2 x\ (\log_4 x)^2}{(\log_3 x)^2}.$$

Let c_8 be the constant implied in the preceding inequality. Put

$$w_5 = \frac{\log_2 x}{\log_3 x}, \qquad w_6 = c_8 \frac{\log_2 x \ (\log_4 x)^2}{(\log_3 x)^2}, \qquad w_7 = w_5 - w_6,$$

and consider the set

$$\mathcal{A}_7(x) = \left\{ n \in \mathcal{A}(x) \backslash \left(\cup_{j=1}^{6} \mathcal{A}_j(x) \right) : \omega(n) > w_5 \right\}.$$

Lemma 6. *The following estimate holds:*

$$\#\mathcal{A}_7(x) \ll \frac{x}{(\log x)^{2+o(1)}}. \tag{34}$$

Proof. Let $\mathcal{R}_3(x)$ denote the set of primes $p \le y_4$ together with the set of primes $p \in \mathcal{R}_2(x)$. Clearly,

$$\sum_{p \in \mathcal{R}_3(x)} \sum_{\beta \ge 1} \frac{1}{p^\beta} \le \sum_{p \le y_4} \frac{1}{p} + O(1) \le \log_2 y_4 + O(1) = (\log_3 x)^3 + O(1). \tag{35}$$

As before, write $n = Pm$, where $P > \max\{P(m), y_1\}$, and denote by $\mathcal{M}_3(x)$ the set of all integers m obtained in this way. Since $n \in \mathcal{A}_7(x)$, we know that m has $k_6 \le w_6$ primes $p \le x$ that are not in $\mathcal{R}_3(x)$, and

$$k_7 := k_4 + k_5 \ge w_5 - k_6 \ge w_5 - w_6 \ge w_7$$

primes in $\mathcal{R}_3(x)$. Further, $k_7 \le w_1$ since $n \notin \mathcal{A}_3(x)$. For fixed $m \in \mathcal{M}_3(x)$, the prime number P can be chosen in at most

$$\pi(x/m) \ll \frac{x}{\varphi(m)(\log(x/m))^2} \ll \frac{x(\log_2 x)^3}{(\log x)^2} \cdot \frac{1}{m}$$

different ways, where we have again used the inequality (26) together with the fact that $\varphi(m) \gg m/\log_2 m$. Summing this estimate over all $m \in \mathcal{M}_3(x)$, we derive that

$$\#\mathcal{A}_7(x) \ll \frac{x(\log_2 x)^3}{(\log x)^2} \sum_{m \in \mathcal{M}_3(x)} \frac{1}{m}, \tag{36}$$

where

$$\sum_{m \in \mathcal{M}_3(x)} \frac{1}{m} \le \sum_{\substack{k_6 \le w_6 \\ w_7 \le k_7 \le w_1}} \frac{1}{k_6!} \left(\sum_{p<x} \sum_{\beta \ge 1} \frac{1}{p^\beta} \right)^{k_6} \cdot \frac{1}{k_7!} \left(\sum_{p \in \mathcal{R}_3(x)} \sum_{\beta \ge 1} \frac{1}{p^\beta} \right)^{k_7}.$$

For fixed $k_6 \le w_6$, by Stirling's formula again, it follows easily that

$$\begin{aligned}\frac{1}{k_6!}\left(\sum_{p<x}\sum_{\beta\ge 1}\frac{1}{p^\beta}\right)^{k_6} &\ll \left(\frac{e\log_2 x+O(1)}{k_6}\right)^{k_6} \le \left(\frac{e\log_2 x+O(1)}{w_6}\right)^{w_6}\\ &= \exp\left(O\left(\frac{\log_2 x\ (\log_4 x)^2}{\log_3 x}\right)\right) = (\log x)^{o(1)}.\end{aligned}$$

Here, we used the fact that if $A>1$ is fixed, then the function $t\mapsto (A/t)^t$ is increasing for $t\le A/e$ and decreasing for $t>A/e$. For example, above we used this argument with $A=e\log_2 x+O(1)$ and $t=k_6\le w_6<A/e$, once x is sufficiently large.

On the other hand, for fixed k_7 in the interval $[w_7,w_1]$, using Stirling's formula again and the estimate (35), we have

$$\begin{aligned}\frac{1}{k_7!}\left(\sum_{p\in\mathcal{R}_3(x)}\sum_{\beta\ge 1}\frac{1}{p^\beta}\right)^{k_7} &\le \left(\frac{e(\log_3 x)^3+O(1)}{k_7}\right)^{k_7}\\ &\le \left(\frac{e(\log_3 x)^3+O(1)}{w_7}\right)^{w_7}\\ &= \exp\left(-(1+o(1))\log_2 x\right) = (\log x)^{-1+o(1)},\end{aligned}$$

where now we have used the fact that $(B/t)^t$ is decreasing for the fixed $B=e(\log_3 x)^3+O(1)$ and $t\ge w_7>B/e$, once x is sufficiently large.

Since the pair $(k_6,\ k_7)$ can be chosen in at most $O((\log_2 x)^2)$ distinct ways, we obtain that

$$\sum_{m\in\mathcal{M}_3(x)}\frac{1}{m}\le\frac{1}{(\log x)^{1+o(1)}},$$

which together with (36) leads to the proof of (34). □

From now on, we consider only integers $n\in\mathcal{A}(x)\setminus\left(\cup_{j=1}^7\mathcal{A}_j(x)\right)$. If we again write every such n in the form $n=Pm$, and let $\mathcal{M}_4(x)$ be the set of integers m that arise in this way, then it follows that

$$\sum_{m\in\mathcal{M}_4(x)}\frac{1}{m}\le(\log x)^{o(1)}.$$

Indeed,

$$\begin{aligned}\sum_{m\in\mathcal{M}_4(x)}\frac{1}{m} &\ll \sum_{k\le w_5}\frac{1}{k!}\left(\sum_{p\le x}\sum_{\beta\ge 1}\frac{1}{p^\beta}\right)^k \ll \frac{\log_2 x}{\log_3 x}\left(\frac{e\log_2 x+O(1)}{w_5}\right)^{w_5}\\ &\le \exp\left(O\left(\frac{\log_2 x\log_4 x}{\log_3 x}\right)\right)=(\log x)^{o(1)}. \qquad (37)\end{aligned}$$

We now put

$$\mathcal{A}_8(x)=\{n\in\mathcal{A}(x)\setminus\left(\cup_{j=1}^7\mathcal{A}_j(x)\right): F(n)>w_2\},$$

where $F(n)$ is defined by (15).

Lemma 7. *The following estimate holds:*

$$\#\mathcal{A}_8(x)\ll\frac{x}{(\log x)^{2+o(1)}}. \qquad (38)$$

Proof. To prove (38), we follow the same arguments used to bound $\#\mathcal{A}_5(x)$. Suppose that we are given an element $(t,\alpha,\alpha_1,\dots,\alpha_t,\kappa_1,\dots,\kappa_t)$ of $\mathcal{D}$ and that this element encodes the powers of 2 in n and in $p-1$ for the odd primes $p\mid n$, as in the proof of the upper bound (16) for $\#\mathcal{A}_5(x)$. Assume further that $P=p_i$ for some $i\in\{1,\dots,t\}$. When all these data are fixed, the number of corresponding integers n is bounded above (see inequality (18)) by

$$\frac{x\log_2 x}{\log x}\cdot\frac{1}{2^{\alpha+\sum_{j=1}^t\alpha_j\kappa_j}}\cdot(c_4\log_2 x)^{\sum_{j=1}^t\kappa_j}.$$

Since $\alpha+\sum_{j=1}^t\alpha_j\kappa_j>w_2=c_2\log_2 x$, and $\sum_{j=1}^t\kappa_j=\omega(m)\le w_5$, it follows that

$$\begin{aligned}\frac{1}{2^{\alpha+\sum_{j=1}^t\alpha_j\kappa_j}}\cdot(c_4\log_2 x)^{\sum_{j=1}^t\kappa_j} &\le \exp\left(w_5\log(c_4\log_2 x)-w_2\log 2\right)\\ &= \exp(-(1+o(1))\log_2 x)=\frac{1}{(\log x)^{1+o(1)}}.\end{aligned}$$

Therefore,

$$\#\mathcal{A}_8(x)\ll\#\mathcal{D}\cdot\log_2 x\cdot\frac{x}{(\log x)^{2+o(1)}}=\frac{x}{(\log x)^{2+o(1)}},$$

which is the desired result. □

In order to continue our argument, we shall need the following technical result:

Lemma 8. *Uniformly for $2 \leq d \leq x$, the following estimate holds:*

$$\sum_{\substack{p \in \mathcal{R}_2(x) \\ p \equiv 1 \pmod{d}}} \frac{1}{p} \ll \frac{\log_2 d}{\varphi(d)}.$$

Proof. Let $d \geq 2$. For $x \geq d$, we first determine an upper bound on the counting function of the set $\mathcal{R}_2(x; d, 1)$ of primes $p \in \mathcal{R}_2(x)$ in the arithmetic progression $p \equiv 1 \pmod{d}$. Without loss of generality, we can assume that $x \geq \exp\left((\log d)^2\right)$, for otherwise the inequality asserted by the lemma follows from the fact that

$$\sum_{\substack{p \leq y \\ p \equiv 1 \pmod{d}}} \frac{1}{p} \ll \frac{\log_2 y}{\varphi(d)},$$

which is valid for all $1 \leq d \leq y$ (see again (8)). We can also assume that d is large enough for our purposes at hand. Now, put

$$y = \exp\left(\frac{\log x \log_3 x}{6 \log_2 x}\right).$$

Then, if d is sufficiently large,

$$y \geq \exp\left(\frac{(\log d)^2 \left(\log_3 d + \log 2\right)}{12 \log_2 d}\right) > d.$$

For each prime $p \in \mathcal{R}(x; d, 1)$, write $p - 1 = dm$. We first remark that the number of primes p for which $P(m) \leq y$ (and thus, $P(p-1) \leq y$) cannot exceed

$$\Psi\left(\frac{x}{d}, y\right) \leq \frac{x}{d} \exp\left(-(1 + o(1)) v \log v\right),$$

where

$$v = \frac{\log(x/d)}{\log y} > \frac{1}{2} \cdot \frac{\log x}{\log y} = \frac{3 \log_2 x}{\log_3 x}.$$

Hence, $v \log v \geq (3 + o(1)) \log_2 x$, therefore the number of such primes p is

$$\ll \frac{x}{d(\log x)^{3+o(1)}} \leq \frac{x}{d(\log x)^2}$$

if x is sufficiently large.

Now, consider primes $p \in \mathcal{R}(x; d, 1)$ such that $P(m) > y$. From this subset, we discard those primes p for which there exists a prime $q > y$ such that $q^2 \mid p - 1$; it is easy to see, using an argument similar the one used to analyze the counting function $\#\mathcal{R}_1(x)$, that the number of discarded primes is

$$\ll \frac{x}{dy \log y} \ll \frac{x}{d(\log x)^2}.$$

For the remaining primes, write $p - 1 = dm_1 P$, where p is a prime satisfying $P = P(p-1) > \max\{P(m_1), y\}$. Let $\mathcal{M}(x)$ be the set of integers m_1 that occur in this way. By Brun's method, for every fixed value of m_1, the number of possibilities for the prime P is

$$\ll \frac{x}{\varphi(dm_1)(\log(x/dm_1))^2} \ll \frac{x \log_2 x}{\varphi(d)(\log(x/dm_1))^2} \cdot \frac{1}{m_1}.$$

Since $x/(dm_1) \geq P(p-1) > y$, we get that

$$\log(x/dm_1) > \log y \gg \frac{\log x \log_3 x}{\log_2 x},$$

therefore the number of possibilities for P is

$$\ll \frac{x(\log_2 x)^3}{\varphi(d)(\log x)^2} \cdot \frac{1}{m_1}. \tag{39}$$

Recalling that $\mathcal{M}(x)$ is contained in the set of those positive integers that have at most $\log_2 x / \log_3 x$ primes $> f(x)$, the argument used in the analysis of the counting function $\#\mathcal{R}_2(x)$ (see estimate (32)) shows that

$$\sum_{m_1 \in \mathcal{M}(x)} \frac{1}{m_1} \leq (\log x)^{o(1)}.$$

Therefore, summing up the inequality (39) over all the possible values for $m_1 \in \mathcal{M}(x)$, and combining the result with our previous estimates, we obtain the bound:

$$\#\mathcal{R}_2(x; d, 1) \leq \frac{x}{\varphi(d)(\log x)^{2+o(1)}} < \frac{x}{\varphi(d)(\log x)^{3/2}},$$

if $x \geq \exp((\log d)^2)$ and d is sufficiently large. The desired inequality now follows by partial summation. □

Next, defining $F(n)$ as usual by (15), we show that the following estimate holds:

Lemma 9.

$$\#\mathcal{A}(x) \le \frac{x}{(\log x)^{1+o(1)}} \sum_{\ell \in F\left(\mathcal{A}(x)\setminus\left(\cup_{j=1}^{8}\mathcal{A}_j(x)\right)\right)} \frac{1}{2^\ell} + O\left(\frac{x}{(\log x)^{2+o(1)}}\right). \tag{40}$$

Proof. To prove this lemma, we apply a modification of the argument used to bound $\#\mathcal{A}_5(x)$ and $\#\mathcal{A}_8(x)$. We let $n \in \mathcal{A}(x)\setminus\left(\cup_{j=1}^{8}\mathcal{A}_j(x)\right)$. As before, we write $n = Pm$, and we note that $m = m' \cdot m''$, where

$$m' = \prod_{\substack{p^{\beta_p} \parallel m \\ p \in \mathcal{R}_3(x)}} p^{\beta_p} \qquad \text{and} \qquad m'' = \prod_{\substack{p^{\beta_p} \parallel m \\ p \notin \mathcal{R}_3(x)}} p^{\beta_p}.$$

Recall that $\omega(m') = k_7$ and $\omega(m'') = k_6 \le w_6$. Now, let us suppose that $(t', \alpha', \alpha'_1, \dots, \alpha'_{t'}, \kappa'_1, \dots, \kappa'_{t'})$ and $(t'', \alpha'', \alpha''_1, \dots, \alpha''_{t''}, \kappa''_1, \dots, \kappa''_{t''})$ encode the powers of 2 in m' and $p'_i - 1$ for $i = 1, \dots, k_7$ (where $p'_i \mid m'$) and in m'' and $p''_j - 1$ for $j = 1, \dots, k_6$ (where $p''_j \mid m''$), respectively, as in the analysis of $\mathcal{A}_5(x)$. Suppose further that $2^{\alpha_P} \parallel P - 1$. Then,

$$F(n) = \alpha_P + \alpha' + \alpha'' + \sum_{i=1}^{t'} \alpha'_i \kappa'_i + \sum_{j=1}^{t''} \alpha''_j \kappa''_j.$$

For each fixed set of data, the number of positive integers counted is

$$\begin{aligned} &\ll \frac{x \log_2 x}{2^{\alpha'+\alpha''+\alpha_P} \log x} \prod_{i=1}^{t'} \frac{1}{\kappa'_i!} \left(\sum_{\substack{p'_i \in \mathcal{R}_3(x) \\ p'_i \equiv 1 \pmod{2^{\alpha'_i}}}} \sum_{\beta'_i \ge 1} \frac{1}{{p'_i}^{\beta'_i}} \right)^{\kappa'_i} \\ &\quad \times \prod_{j=1}^{t''} \frac{1}{\kappa''_j!} \left(\sum_{\substack{p''_j \notin \mathcal{R}_3(x) \\ p''_j \equiv 1 \pmod{2^{\alpha''_j}}}} \sum_{\beta''_j \ge 1} \frac{1}{{p''_j}^{\beta''_j}} \right)^{\kappa''_j}. \end{aligned} \tag{41}$$

Clearly,

$$\sum_{\substack{p_i' \in \mathcal{R}_3(x) \\ p_i' \equiv 1 \pmod{2^{\alpha_i'}}}} \sum_{\beta_i' \geq 1} \frac{1}{{p_i'}^{\beta_i'}} \ll \sum_{\substack{p_i' \in \mathcal{R}_3(x) \\ p_i' \equiv 1 \pmod{2^{\alpha_i'}}}} \frac{1}{p_i'} \ll S_1' + S_2',$$

where

$$S_1' = \sum_{\substack{p_i' \in \mathcal{R}_2(x) \\ p_i' \equiv 1 \pmod{2^{\alpha_i'}}}} \frac{1}{p_i'} \qquad \text{and} \qquad S_2' = \sum_{\substack{p_i' \leq y_4 \\ p_i' \equiv 1 \pmod{2^{\alpha_i'}}}} \frac{1}{p_i'}.$$

Using Lemma 8 together with the fact that $\alpha_i' \leq w_2$, it follows immediately that $S_1' \ll (\log_3 x)/2^{\alpha_i'}$. We also have

$$S_2' \ll \frac{\log_2 y_4}{2^{\alpha_i'}} \ll \frac{(\log_3 x)^3}{2^{\alpha_i'}}.$$

Since $\sum_{i=1}^{t'} \kappa_i' \ll \log_2 x / \log_3 x$, the estimates above imply that

$$\begin{aligned} \prod_{i=1}^{t'} \frac{1}{\kappa_i'!} \left(\sum_{\substack{p_i' \in \mathcal{R}_3(x) \\ p_i' \equiv 1 \pmod{2^{\alpha_i'}}}} \sum_{\beta_i' \geq 1} \frac{1}{{p_i'}^{\beta_i'}} \right)^{\kappa_i'} &= \frac{(\log_3 x)^{O\left(\frac{\log_2 x}{\log_3 x}\right)}}{2^{\sum_{i=1}^{t'} \alpha_i' \kappa_i'}} \\ &= \frac{(\log x)^{o(1)}}{2^{\sum_{i=1}^{t'} \alpha_i' \kappa_i'}}. \end{aligned} \tag{42}$$

Furthermore, since

$$\sum_{\substack{p_j'' \leq x \\ p_j'' \equiv 1 \pmod{2^{\alpha_j''}}}} \sum_{\beta_j'' \geq 1} \frac{1}{{p_j''}^{\beta_j''}} \ll \sum_{\substack{p_j'' \notin \mathcal{R}_3(x) \\ p_j'' \equiv 1 \pmod{2^{\alpha_j''}}}} \frac{1}{p_j''} \ll \frac{\log_2 x}{2^{\alpha_j''}},$$

and

$$\sum_{j=1}^{t''} \kappa_j'' \leq k_6 \ll \frac{\log_2 x \ (\log_4 x)^2}{(\log_3 x)^2},$$

we also see that

$$\prod_{j=1}^{t''} \frac{1}{\kappa_j''!} \left(\sum_{\substack{p_j'' \le x \\ p_j'' \equiv 1 \pmod{2^{\alpha_j''}}}} \sum_{\beta_j'' \ge 1} \frac{1}{p_j''^{\beta_j''}} \right)^{\kappa_j''} = \frac{(\log_2 x)^{O(\log_2 x \ (\log_4 x)^2/(\log_3 x)^2)}}{2^{\sum_{j=1}^{t''} \alpha_j'' \kappa_j''}}$$

$$= \frac{(\log x)^{o(1)}}{2^{\sum_{j=1}^{t''} \alpha_j'' \kappa_j''}}. \tag{43}$$

Summing up the preceding estimates (42) and (43) over all possible data sets $(t', \alpha', \alpha_1', \dots, \alpha_{t'}', \kappa_1', \dots, \kappa_{t'}')$ and $(t'', \alpha'', \alpha_1'', \dots, \alpha_{t''}'', \kappa_1'', \dots, \kappa_{t''}'')$ (there are at most $(\log x)^{o(1)}$ possibilities), and using the estimates we have obtained above for $\#\mathcal{A}_j(x)$, $j = 1, \dots, 8$, we obtain the desired estimate (40). □

In particular, if we put $w_8 = \log_2 x$, and

$$\mathcal{A}_9(x) = \{n \in \mathcal{A} \setminus \left(\cup_{j=1}^{8} \mathcal{A}_j(x)\right) : F(n) > w_8\},$$

then the above argument and the estimate (40) immediately implies that

$$\#\mathcal{A}_9(x) \le \frac{x}{(\log x)^{1+\log 2+o(1)}}. \tag{44}$$

Finally, we come to the last set $\mathcal{A}_{10}(x) = \mathcal{A}(x) \setminus \left(\cup_{j=1}^{9} \mathcal{A}_j(x)\right)$. From now on, we consider only integers $n \in \mathcal{A}_{10}(x)$.

Lemma 10. *The following estimate holds:*

$$\#\mathcal{A}_{10}(x) \le \frac{x}{(\log x)^{3/2+o(1)}}. \tag{45}$$

Proof. For $n \in \mathcal{A}_{10}(x)$, write $n = Pm$, where $P = P(n)$ and $m \in \mathcal{M}_4(x)$. From the inequality (29), we see that

$$\begin{aligned} \omega_{>y_2}(P-1) \le \sum_{p \,|\, n} \omega_{>y_2}(p-1) &\le v_2(n) + \sum_{p \,|\, n} v_2(p-1) + O(1) \\ &= F(n) + O(1). \end{aligned}$$

Let c_9 be the constant implied by $O(1)$ in this estimate. Let $\nu = F(n) \le w_8$, and fix the integer ν. Put $g(p) = (\log p)^2$, let $\alpha \le w_2$ be a positive integer, and let

$$\mathcal{R}_{4,\alpha,\nu} = \{p : 2^\alpha \,\|\, p-1 \text{ and } \omega_{>g(p)}(p-1) \le \nu + c_9\}.$$

Then, $P = P(n) \in \mathcal{R}_{4,\alpha,\nu}(x/m)$ for some positive integer α. We now show that the estimate

$$\#\mathcal{R}_{4,\alpha,\nu}(x/m) \le \frac{x}{2^\alpha m(\log x)^{h(\delta)+o(1)}} + \frac{x}{m(\log x)^{2+o(1)}} \tag{46}$$

holds uniformly for all $\nu \le w_8$, $\alpha \le w_2$, and $m \in \mathcal{M}_4(x)$, where we put $h(\delta) = 2 - \delta\log(e/\delta)$, with $\delta = \nu/\log_2 x$. To do this, we first observe that it suffices to prove only the weaker assertion that

$$\#\mathcal{R}_{4,\alpha,\nu}(x) \le \frac{x}{2^\alpha(\log x)^{h(\delta)+o(1)}} + \frac{x}{(\log x)^2} \tag{47}$$

holds uniformly for all $\nu \le w_8$ and $\alpha \le w_2$. Indeed, since $x/m > y_1$, it follows that $\log(x/m) = (\log x)^{1+o(1)}$ and that $\delta_m = \nu/\log_2(x/m) = \delta + o(1)$, thus (47) implies (46) uniformly for all $m \in \mathcal{M}_4(x)$. To establish (47), let p be a prime in $\mathcal{R}_{4,\alpha,\nu}(x)$. As in the estimation of $\mathcal{R}_1(x)$ and $\mathcal{R}_2(x)$, we can assume that $P(p-1) > y_1$, that there does not exist a prime $q > y_3$ for which $q^2 \mid p-1$, and that $\omega(p-1) < w_1$. Let us write $p - 1 = 2^\alpha Pm$, where m is odd and $P > \max\{P(m), y_1\}$, and let $\mathcal{M}_5(x)$ be the set of all integers m obtained in this way. Using Brun's method again, we see that for each fixed value of $m \in \mathcal{M}_5(x)$, the number of possibilities for the prime p is

$$\ll \frac{x}{2^\alpha\varphi(m)(\log(x/m))^2} \ll \frac{x(\log_2 x)^3}{2^\alpha(\log x)^2}\cdot\frac{1}{m}$$

(see inequality (26) again). Every $m \in \mathcal{M}_5(x)$ has $k_8 \le w_9 = \delta\log_2 x + c_9$ prime factors larger than $y_5 = (\log x)^3$, and $k_9 \le w_1$ remaining primes which are $\le y_5$; therefore,

$$\sum_{m\in\mathcal{M}_5(x)}\frac{1}{m} \le \sum_{\substack{k_8\le w_9\\ k_9\le w_1}}\frac{1}{k_8!}\left(\sum_{p\le x}\sum_{\beta\ge1}\frac{1}{p^\beta}\right)^{k_8}\cdot\frac{1}{k_9!}\left(\sum_{p\le y_5}\sum_{\beta\ge1}\frac{1}{p^\beta}\right)^{k_9}.$$

For fixed $k_9 \le w_1$, we have by Stirling's formula:

$$\begin{aligned}\frac{1}{k_9!}\left(\sum_{p\le y_5}\sum_{\beta\ge1}\frac{1}{p^\beta}\right)^{k_9} &\le \left(\frac{e\log_2 y_5 + O(1)}{k_9}\right)^{k_9}\\ &\le \left(\frac{e\log_3 x + O(1)}{k_9}\right)^{k_9}\\ &\le \exp\left(\log_3 x + O(1)\right) \ll \log_2 x.\end{aligned}$$

Similarly, for fixed $k_8 \le w_9$, we have:

$$\begin{aligned}\frac{1}{k_8!}\left(\sum_{p\le x}\sum_{\beta\ge 1}\frac{1}{p^\beta}\right)^{k_8} &\le \left(\frac{e\log_2 x + O(1)}{k_8}\right)^{k_8} \ll \left(\frac{e\log_2 x + O(1)}{w_9}\right)^{w_9} \\ &\ll \exp\left(z(\delta)\log_2 x\right) = (\log x)^{z(\delta)},\end{aligned}$$

where $z(\delta) = \delta\log(e/\delta)$. Thus, since the pair $(k_8,\ k_9)$ can be chosen in at most $O((\log_2 x)^2)$ ways, we get that

$$\sum_{m\in\mathcal{M}_5(x)}\frac{1}{m} \ll (\log x)^{z(\delta)}(\log_2 x)^3.$$

Consequently,

$$\begin{aligned}\mathcal{R}_{4,\alpha,\nu}(x) &\ll \frac{x(\log_2 x)^3}{2^\alpha(\log x)^2}\sum_{m\in\mathcal{M}_5(x)}\frac{1}{m} + \frac{x}{(\log x)^2} \\ &\le \frac{x(\log_2 x)^6}{2^\alpha(\log x)^{2-z(\delta)+o(1)}} + \frac{x}{(\log x)^2} \\ &= \frac{x}{2^\alpha(\log x)^{h(\delta)+o(1)}} + \frac{x}{(\log x)^2},\end{aligned}$$

which establishes (47).

Returning to the integers $n \in \mathcal{A}_{10}(x)$, we see that each one has the form $n = Pm$, where $P > \max\{P(m), y_1\}$ belongs to the set $\mathcal{R}_{4,\alpha,\nu}(x/m)$, and $m \in \mathcal{M}_4(x)$. For fixed values of m, $\nu \le w_8$, and $\alpha \le w_2$, the number of possibilities for P is at most

$$\#\mathcal{R}_{4,\alpha,\nu}(x/m) \ll \left(\frac{x}{2^\alpha(\log x)^{h(\delta)+o(1)}} + \frac{x}{(\log x)^2}\right)\frac{1}{m}.$$

Now an argument similar to the one used to prove estimate (40) (and similar to the one used to find upper bounds on $\mathcal{A}_5(x)$ and $\mathcal{A}_8(x)$), leads easily to the conclusion that in formula (40) we may replace the first exponent $1+o(1)$ by $h(\delta)+o(1)$ once $\delta < 1$, in particular, when we are counting numbers in $\mathcal{A}_{10,\nu}(x) = \{n \in \mathcal{A}_{10}(x) : F(n) = \nu\}$. Thus, summing up over all possible values of α and m, we obtain that

$$\#\mathcal{A}_{10,\nu}(x) \le \frac{x}{(\log x)^{h(\delta)+\delta\log 2+o(1)}}, \tag{48}$$

since $h(\delta) < 2$ for $\delta < 1$. Finally, summing up over all possible values of ν (at most $O(\log_2 x)$ of them), and noticing that the minimum of the function

$$h(\delta) + \delta \log 2 = 2 - \delta \log(e/\delta) + \delta \log 2$$

occurs at $\delta = 1/2$ with a value of $3/2$, we obtain the stated result. □

Theorem 2 now follows at once from the estimates (9), (12), (13), (14), (16), (21), (34), (38), (44), and (45). □

4 The normal order of $\varphi(\lambda(n))/\lambda(\varphi(n))$

Theorem 3. *The estimate*

$$\frac{\varphi(\lambda(n))}{\lambda(\varphi(n))} = \exp\big((1 + o(1))(\log_2 n)^2 \log_3 n\big)$$

holds on a set of positive integers n of asymptotic density one.

Proof. Clearly,

$$\varphi(\lambda(n)) = \frac{\varphi(\lambda(n))}{\lambda(n)} \cdot \lambda(n).$$

Since the inequalities

$$1 \geq \frac{\varphi(\lambda(n))}{\lambda(n)} \gg \frac{1}{\log_2 \lambda(n)} \geq \frac{1}{\log_2 n}$$

hold for all n, and the estimate

$$\lambda(n) = n \exp(-(1 + o(1)) \log_2 n \log_3 n) \tag{49}$$

holds for almost all n (see [19]), it follows that

$$\varphi(\lambda(n)) = n \exp(-(1 + o(1)) \log_2 n \log_3 n) \tag{50}$$

holds for almost all positive integers n.

We also have:

$$\lambda(\lambda(n)) \leq \lambda(\varphi(n)) = \lambda\left(\frac{\varphi(n)}{\lambda(n)} \cdot \lambda(n)\right) \leq \lambda(\lambda(n)) \cdot \frac{\varphi(n)}{\lambda(n)}.$$

Here, we used the fact that the prime factors of $\lambda(n)$ and $\varphi(n)$ are the same, together with the (easily proved) fact that if $m = ab$ and every prime factor of b divides a, then $\lambda(ab) \mid \lambda(a)b$. Now, writing

$$\frac{\varphi(n)}{\lambda(n)} = \frac{\varphi(n)}{n} \cdot \frac{n}{\lambda(n)}$$

and using the estimate $1 \geq \varphi(n)/n \gg 1/\log_2 n$ together with (49), we see that

$$\lambda(\varphi(n)) = \lambda(\lambda(n)) \exp((1 + o(1)) \log_2 n \log_3 n)$$

holds for almost all n. Applying the result (1) of Martin and Pomerance [36], it follows that the estimate

$$\lambda(\varphi(n)) = n \exp\left(-(1 + o(1))(\log_2 n)^2 \log_3 n\right) \tag{51}$$

holds for almost all positive integers n. The result now follows from (50) and (51). □

References

[1] L. M. Adleman, C. Pomerance and R. S. Rumely, 'On distinguishing prime numbers from composite numbers', *Ann. of Math. (2)* **117** (1983), 173–206.

[2] L. Alaoglu and P. Erdős, 'A conjecture in elementary number theory', *Bull. Amer. Math. Soc.* **50** (1944), 881–882.

[3] W. Banks, J. B. Friedlander, C. Pomerance and I. E. Shparlinski, Multiplicative structure of values of the Euler function, *Proc. Conf. in Number Theory in Honour of Prof. H.C. Williams*, AMS (2004), 29–48.

[4] W. Banks, F. Luca and I. E. Shparlinski, 'Arithmetic properties of $\varphi(n)/\lambda(n)$ and the structure of the multiplicative group modulo n', *Comm. Math. Helvetici*, to appear.

[5] W. Banks, K. Ford, F. Luca, F. Pappalardi and I. E. Shparlinski, 'Values of the Euler function in various sequences', *Monatsh. Math.*, to appear.

[6] W. Banks, F. Luca, F. Saidak and I. E. Shparlinski, 'Values of arithmetical functions equal to a sum of two squares', *Quart. J. Math. (Oxford)* **56** (2005), 123–139.

[7] N. L. Bassily, I. Kátai and M. Wijsmuller, 'On the prime power divisors of the iterates of the Euler φ-function', *Publ. Math. Debrecen* **55** (1999), 17–32.

[8] L. Carlitz, 'A note on the composition of arithmetic functions', *Duke Math. J.* **33** (1966), 629–632.

[9] R. D. Carmichael, 'Note of a new number theory function', *Bull. Amer. Math. Soc.* **16** (1910), pp. 232–238.

[10] J.-M. De Koninck, F. Luca and A. Sankaranarayanan, 'Positive integers n whose Euler function is a power of the kernel function', *Rocky Mtn. J. Math.*, to appear.

[11] J.-M. De Koninck and F. Luca, 'On the compositions of Euler's function and the sum of divisors function', *Colloq. Math.*, to appear.

[12] T. Dence and C. Pomerance, 'Euler's function in residue classes', *The Ramanujan J.* **2** (1998), 7–20.

[13] P. Erdős, 'On the normal order of prime factors of $p-1$ and some related problems concerning Euler's φ-function', *Quart. Journ. Math. (Oxford)* **6** (1935), 205 - 213.

[14] P. Erdős, 'Some remarks on number theory. II.', *Mat. Lapok* **12** (1961), 161–169.

[15] P. Erdős, A. Granville, C. Pomerance and C. Spiro, 'On the normal behavior of the iterates of some arithmetic functions', *Analytic Number Theory*, Birkhäuser, Boston, 1990, 165–204.

[16] P. Erdős and R. R. Hall, 'Euler's function and its iterate', *Mathematika* **24** (1977), 173–177.

[17] P. Erdős and M. Kac, 'The Gaussian law of errors in the theory of additive number theoretic functions', *Amer. J. Math.* **62** (1940), 738-742.

[18] P. Erdős and C. Pomerance, 'On the normal number of prime factors of $\varphi(n)$', *Rocky Mountain J. Math.* **15** (1985), 343–352.

[19] P. Erdős, C. Pomerance and E. Schmutz, 'Carmichael's lambda function', *Acta Arith.* **58** (1991), 363–385.

[20] L. Euler, *Novi Comm. Acad. Petrop.* **8** (1760/61), 74.

[21] K. Ford, 'The distribution of totients', *The Ramanujan J.* **2** (1998), 67–151.

[22] K. Ford, 'The number of solutions of $\varphi(x) = m$', *Annals of Math.* **150** (1999), 283–311.

[23] K. Ford, S. Konyagin and C. Pomerance, 'Residue classes free of values of Euler's function', *Proc. Number Theory in Progress*, Walter de Gruyter, Berlin, 1999, 805–812.

[24] J. B. Friedlander, C. Pomerance and I. E. Shparlinski, 'Period of the power generator and small values of Carmichael's function', *Math. Comp.* **70** (2001), 1591–1605.

[25] S. W. Golomb, 'Equality among number-theoretic functions', *Abstracts AMS* **14** (1993), 415–416.

[26] R. K. Guy, *Unsolved Problems in Number Theory*, Springer-Verlag, 1994.

[27] H. Halberstam and H.-E. Richert, *Sieve methods*, Academic Press, London, UK, 1974.

[28] F. Halter-Koch and W. Steindl, 'Teilbarkeitseigenschaften der iterierten Eulerschen Phi-Funktion', *Arch. Math. (Basel)* **42** (1984), no. 4, 362–365.

[29] G. H. Hardy and J. E. Littlewood, 'Some problems on partitio numerorum III. On the expression of a number as a sum of primes', *Acta Math.* **44** (1923), 1–70.

[30] G. H. Hardy and S. Ramanujan, 'The normal number of prime factors of an integer', *Quart. J. Math. (Oxford)* **48** (1917), 76-92.

[31] M. Hausman, 'The solution of a special arithmetic equation', *Canad. Math. Bull.* **25** (1982), 114–117.

[32] A. Hildebrand, 'On the number of positive integers $\leq x$ and free of prime factors $> y$,' *J. Number Theory* **22** (1986), 289–307.

[33] F. Luca and C. Pomerance, 'On some problems of Mąkowski–Schinzel and Erdős concerning the arithmetical functions φ and σ', *Colloq. Math.* **92** (2002), 111–130.

[34] H. Maier and C. Pomerance, 'On the number of distinct values of Euler's φ function,' *Acta Arith.* **49** (1988), 263–275.

[35] A. Makowski and A. Schinzel, 'On the functions φ and $\sigma(n)$,' *Colloq. Math.* **13** (1964-65), 95–99.

[36] G. Martin and C. Pomerance, 'The iterated Carmichael λ-function and the number of cycles of the power generator', *Acta Arith.* **118** (2005), 303–335.

[37] W. H. Mills, 'Iteration of the $\varphi(n)$-function', *American. Math. Monthly* **50** (1943), 547–549.

[38] H. L. Montgomery and R. C. Vaughan,'The large sieve', *Mathematika* **20** (1973), 119–134.

[39] M. R. Murty and V. K. Murty, 'Prime divisors of Fourier coefficients of modular forms', *Duke Math. J.* **51** (1984), 57–76.

[40] I. Niven, 'The iteration of certain arithmetic functions', *Canadian J. Math.* **2** (1950), 406–408.

[41] S. S. Pillai, 'On a function connected with $\varphi(n)$', *Bull. Amer. Math. Soc.* **35** (1929), 837–841.

[42] C. Pomerance, 'Popular values of Euler's function', *Mathematika* **27** (1980), 84–89.

[43] C. Pomerance, 'On the composition of the arithmetic functions σ and φ', *Colloq. Math.* **58** (1989), 11–15.

[44] H. N. Shapiro, 'Iterates of arithmetic functions and a property of the sequence of primes', *Amer. Math. Monthly* **50** (1943), 18–30.

[45] H. N. Shapiro, 'Iterates of arithmetic functions and a property of the sequence of primes', *Pacific J. Math.* **3** (1953), 647–655.

[46] V. Siva Rama Prasad and P. Fonseca, 'On the iterates of arithmetic functions in a class', *Indian J. Pure Appl. Math.* **23** (1992), 7–14.

[47] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, University Press, Cambridge, UK, 1995.

[48] P. Turán, 'On a theorem of Hardy and Ramanujan', *J. London Math. Soc.* **9** (1934), 274–276.

[49] R. Warlimont, 'On the iterates of Euler's function', *Arch. Math. (Basel)* **76** (2001), 345–349.

Index