

On the average number of divisors of the Euler function

FLORIAN LUCA

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

CARL POMERANCE

Department of Mathematics
Dartmouth College
Hanover, NH 03755–3551, USA
carl.pomerance@dartmouth.edu

March 3, 2006

Abstract

We let $\varphi(\cdot)$ and $\tau(\cdot)$ denote the Euler function and the number-of-divisors function, respectively. In this paper, we study the average value of $\tau(\varphi(n))$ when n ranges in the interval $[1, x]$.

Mathematics Subject Classification: 11N37

Key Words: divisors, Euler function.

The first author was supported in part by UNAM grant PAPIIT IN104505. The second author was supported in part by NSF grant DMS-0401422. This work was completed while the first author visited Dartmouth College as a Shapiro Fellow.

1 Introduction

For a positive integer n , let $\varphi(n)$ denote the Euler function of n , and let $\tau(n)$, $\omega(n)$ and $\Omega(n)$ denote the number of divisors of n , the number of prime divisors of n , and the number of prime-power divisors of n , respectively. There have been a number of papers that have discussed arithmetic properties of $\varphi(n)$, many of these inspired by the seminal paper of Erdős [5] from 1935. In particular, in [7] (see also [6]), the normal number of prime factors of $\varphi(n)$ is considered. It has been known since Hardy and Ramanujan that the normal value of $\omega(n)$ (or $\Omega(n)$) is $\sim \log \log n$, and since Erdős and Kac that $(f(n) - \log \log n)/\sqrt{\log \log n}$ has a Gaussian distribution for $f = \omega$ or Ω . In [7], it is shown that $\varphi(n)$ normally has $\sim \frac{1}{2}(\log \log n)^2$ prime factors, counted with or without multiplicity. In addition, there is a Gaussian distribution for

$$\frac{f(\varphi(n)) - \frac{1}{2}(\log \log n)^2}{\frac{1}{\sqrt{3}}(\log \log n)^{3/2}}$$

for $f = \omega$ and $f = \Omega$. In [2], it is shown that the normal value of $\Omega(\varphi(n)) - \omega(\varphi(n))$ is $\sim \log \log n \log \log \log n$.

Note that it is an easy exercise to show that $\tau(n)$ is on average $\sim \log n$. That is,

$$\sum_{n \leq x} \tau(n) \sim \sum_{n \leq x} \log n.$$

However, from Hardy and Ramanujan, since $2^{\omega(n)} \leq \tau(n) \leq 2^{\Omega(n)}$, we know that for most numbers n , $\tau(n) = (\log n)^{\log 2 + o(1)}$, where $\log 2 = 0.693\dots$. Thus, $\tau(n)$ is on the average somewhat larger than what it is normally. Similarly, for most numbers n ,

$$\tau(\varphi(n)) = 2^{(\frac{1}{2} + o(1))(\log \log n)^2}.$$

One might suspect then that on average, $\tau(\varphi(n))$ is somewhat larger. It comes perhaps as a bit of a shock that the average order of $\tau(\varphi(n))$ is considerably larger.

Our main result is the following:

Theorem 1. *Let*

$$D_\varphi(x) := \frac{1}{x} \sum_{n \leq x} \tau(\varphi(n)).$$

Then, the estimate

$$D_\varphi(x) = \exp \left(c_\varphi(x) \left(\frac{\log x}{\log \log x} \right)^{1/2} \left(1 + O \left(\frac{\log \log \log x}{\log \log x} \right) \right) \right) \quad (1)$$

holds for large real numbers x where $c_\varphi(x)$ is a number in the interval

$$[7^{-1}e^{-\gamma/2}, 2^{3/2}e^{-\gamma/2}], \quad (2)$$

and γ is the Euler constant.

We point out that Theorem 1 above has already been used in the proof of Theorem 1 in [9] to give a sharp error term for a certain sum related to Artin's conjecture on average for composite moduli.

Recall that the *Carmichael function* of n , sometimes also referred to as the *universal exponent* of n and denoted by $\lambda(n)$, is the exponent of the multiplicative group of invertible elements modulo n . If $n = p_1^{\nu_1} \dots p_k^{\nu_k}$ is the factorization of n , then

$$\lambda(n) = \text{lcm}[\lambda(p_1^{\nu_1}), \dots, \lambda(p_k^{\nu_k})],$$

where if p^ν is a prime power then $\lambda(p^\nu) = p^{\nu-1}(p-1)$ except when $p = 2$ and $\nu \geq 3$ in which case, $\lambda(2^\nu) = 2^{\nu-2}$.

It is clear that $\lambda(n) | \varphi(n)$ and that $\omega(\lambda(n)) = \omega(\varphi(n))$. The function $\Omega(\varphi(n)/\lambda(n)) = \Omega(\varphi(n)) - \Omega(\lambda(n))$ was studied in [2]. In addition to the result on $\Omega(\varphi(n)) - \omega(\varphi(n))$ mentioned above, it is shown in [2] that

$$\Omega(\varphi(n)) - \Omega(\lambda(n)) \sim \log \log n \log \log \log n$$

on a set of n of asymptotic density 1.

In the recent paper [1], Arnold writes “it would be interesting to study experimentally how are distributed the different divisors of the number $\varphi(n)$ provided by the periods T of the geometric progressions of residues modulo n .” It is clear that the numbers T range only over the divisors of $\lambda(n)$. We have the following result.

Theorem 2. *Let*

$$D_\lambda(x) := \frac{1}{x} \sum_{n \leq x} \tau(\lambda(n)).$$

(i) *The estimate*

$$D_\lambda(x) = \exp \left(c_\lambda(x) \left(\frac{\log x}{\log \log x} \right)^{1/2} \left(1 + O \left(\frac{\log \log \log x}{\log \log x} \right) \right) \right)$$

holds for large real numbers x where $c_\lambda(x)$ is a number in the interval shown at (2).

(ii) *With $D_\varphi^*(x) = \max_{y \leq x} D_\varphi(y)$, the estimate*

$$D_\lambda(x) = o(D_\varphi^*(x))$$

holds as $x \rightarrow \infty$.

Concerning part (ii) of Theorem 2, we suspect that even the sharper estimate

$$D_\lambda(x) = o(D_\varphi(x))$$

holds as $x \rightarrow \infty$, but we were unable to prove this statement.

We mention that in [3], in the course of investigating sparse RSA exponents, it was shown that

$$\sum_{\substack{n \leq x \\ \Omega(n)=2}} \tau(\varphi(n)) \ll x \log x$$

(see [3], page 347). In particular, the average value of the function $\tau(\varphi(n))$ over those positive integers $n \leq x$ which are the product of two primes is bounded above by a constant multiple of $\log^2 x / \log \log x$.

Our methods can also be applied to study the average number of divisors of values of other multiplicative functions as well. For example, assume that $f : \mathbb{N} \rightarrow \mathbb{Z}$ is a multiplicative function with the property that there exists a polynomial $P_k \in \mathbb{Z}[X]$ of degree k with $P_1(0) \neq 0$ such that $f(p^k) = P_k(p)$ holds for all prime numbers p and all positive integers k . For any positive integer n we shall write $\tau(f(n))$ for the number of divisors of the nonnegative integer $|f(n)|$, with the convention that $\tau(0) = 1$. In this case, our methods show that there exist two positive constants α and β , depending only on the polynomial P_1 , such that the estimate

$$\frac{1}{x} \sum_{n \leq x} \tau(f(n)) = \exp \left(c(x) \left(\frac{\log x}{\log \log x} \right)^{1/2} \left(1 + O \left(\frac{\log \log \log x}{\log \log x} \right) \right) \right)$$

holds for large values of x with some number $c(x) \in [\alpha, \beta]$. In particular, the same estimate as (1) holds if we replace the function $\varphi(n)$ by the function $\sigma(n)$. Indeed, the lower bound follows exactly as in the proof of Theorem 1, while for the upper bound one only needs to slightly adapt our argument.

We close this section by pointing out that it could be very interesting to study the average value of the number of divisors of $f(n)$ for some other integer valued arithmetic functions f . We mention three instances.

Let $a > 1$ be a fixed positive integer and let $f(n)$ be the multiplicative order of a modulo n if a is coprime to n and 0 otherwise. We recall that the functions $\omega(f(n))$ and $\Omega(f(n))$ were studied by Murty and Saidak in [11]. It would be interesting to study the average order of $\tau(f(n))$ in comparison with that of $\tau(\lambda(n))$.

Let \mathbf{E} be an elliptic curve defined over \mathbb{Q} . Let $f(n)$ be the multiplicative function which on prime powers p^k equals $p^k + 1 - a_{p^k}$, the number of points of \mathbf{E} defined over the finite field \mathbb{F}_{p^k} with p^k elements, including the point at infinity.

Let $f(n)$ be the Ramanujan “ τ function” which is the coefficient of q^n in the formal identity

$$\left(\prod_{k=1}^{\infty} (1 - q^k) \right)^{24} = 1 + \sum_{n \geq 1} f(n) q^n.$$

We believe that it should be interesting to study the average number of divisors of $f(n)$ for these functions $f(n)$ and other multiplicative functions that arise from modular forms. Perhaps the methods from this paper dealing with the “easy case” of $\varphi(n)$ will be of help. A relevant paper here is by Murty and Murty [10] in which, building on work of Serre [12, 13], the function $\omega(f(n))$ is analyzed, where $f(n)$ is the Ramanujan τ function.

Throughout this paper, we use c_1, c_2, \dots to denote computable positive constants and x to denote a positive real number. We also use the Landau symbols O and o , the Vinogradov symbols \gg and \ll , and the equal-order-of-magnitude symbol \asymp with their usual meanings. For a positive integer k we use $\log_k x$ for the recursively defined function $\log_1 x := \max\{\log x, 1\}$ and $\log_k x := \max\{\log(\log_{k-1}(x)), 1\}$ where \log denotes the natural logarithm function. When $k = 1$ we simply write $\log_1 x$ as $\log x$ and we therefore understand that $\log x \geq 1$ always. We write p and q for prime numbers. For

two positive integers a and b we write $[a, b]$ for the least common multiple of a and b .

Acknowledgements. The authors thank the anonymous referee for a careful reading of the manuscript and for suggestions that improved the quality of this paper.

2 Some Lemmas

Throughout this section, A, A_1, A_2, A_3, B and C are positive numbers. We write $z := z(x)$ for a function of the real positive variable x which tends to infinity with x in a way which will be made more precise below. We write $P_z := \prod_{p \leq z} p$. The results in this section hold probably in larger ranges than the ones indicated, but the present formulations are enough for our purposes.

For any integer $n \geq 2$ we write $p(n)$ and $P(n)$ for the smallest and largest prime factor of n , respectively, and we let $p(1) = +\infty, P(1) = 1$.

Lemma 3. *Assume that $z \leq \log x / \log_2 x$. (i) For any $A > 0$ there exists $B := B(A)$ such that if $QP_z < \frac{x}{\log^B x}$, we then have*

$$E_z(x) := \sum_{r|P_z} \mu(r) \sum_{\substack{n \leq Q \\ r|n}} \left(\pi(x; n, 1) - \frac{\pi(x)}{\varphi(n)} \right) \ll \frac{x}{\log^A x}. \quad (3)$$

The constant implied in \ll depends at most on A .

(ii) Let $A, A_1, A_2 > 0$ be arbitrary positive numbers. Assume that u is a positive integer with $p(u) > z, u < \log^{A_1} x$ and $\tau(u) < A_2$. There exists $B := B(A, A_1, A_2)$ such that if $QP_z < \frac{x}{\log^B x}$, then

$$E_{u,z}(x) := \sum_{r|P_z} \mu(r) \sum_{\substack{n \leq Q \\ r|n}} \left(\pi(x; [u, n], 1) - \frac{\pi(x)}{\varphi([u, n])} \right) \ll \frac{x}{\log^A x}. \quad (4)$$

The constant implied in \ll depends at most on A, A_1, A_2 .

Proof. Note that

$$\begin{aligned} \sum_{r|P_z} \mu(r) \sum_{\substack{n \leq Q \\ r|n}} \left(\psi(x; n, 1) - \frac{x}{\varphi(n)} \right) \\ = \sum_{n_1 \leq P_z} \sum_{n_2 \leq Q} \gamma_{n_1} \delta_{n_2} \left(\psi(x; n_1 n_2, 1) - \frac{x}{\varphi(n_1 n_2)} \right), \end{aligned} \quad (5)$$

where $\gamma_{n_1} := \mu(n_1)$ if $P(n_1) \leq z$ and it is zero otherwise, and $\delta_{n_2} := 1$ for all $n_2 \leq Q$. Similarly,

$$\begin{aligned} \sum_{r|P_z} \mu(r) \sum_{\substack{n \leq Q \\ r|n}} \left(\psi(x; [u, n], 1) - \frac{x}{\varphi([u, n])} \right) \\ = \sum_{n_1 \leq P_z} \sum_{n_2 \leq Q} \gamma'_{n_1} \delta'_{n_2} \left(\psi(x; n_1 n_2, 1) - \frac{x}{\varphi(n_1 n_2)} \right), \end{aligned} \quad (6)$$

where $\gamma'_{n_1} := \gamma_{n_1}$ and $\delta'_{n_2} := 0$ if $u \nmid n_2$, and it is the cardinality of the set $\{d \leq Q \mid [d, u] = n_2\}$ otherwise. Note that if $n_2 \leq Q$, then $\delta'_{n_2} = \tau(u) \ll 1$ is a constant (i.e., does not depend on n_2) provided that δ'_{n_2} is nonzero. The same argument as the one used in the proof of Theorem 9 in [4] leads to the conclusion that both (5) and (6) are of order of magnitude at most $x/\log^A x$ provided that B is suitably large (in terms of A and of A_1 and A_2 , respectively). Now (3) and (4) follow from (5) and (6) by partial summation and using the fact that these sums are of order of magnitude at most $x/\log^A x$. \square

From now on until the end of the paper we use c_1 for the constant $e^{-\gamma}$, where γ is the Euler constant.

Lemma 4. *Let $A > 0$ and $1 < z \leq (\log x)^A$. We have*

$$L_z(x) := \sum_{\substack{n \leq x \\ p(n) > z}} \frac{1}{n} = c_1 \frac{\log x}{\log z} + O\left(\frac{\log x}{\log^2 z}\right), \quad (7)$$

and

$$M_z(x) := \sum_{\substack{n \leq x \\ p(n) > z}} \frac{1}{\varphi(n)} = c_1 \frac{\log x}{\log z} + O\left(\frac{\log x}{\log^2 z}\right). \quad (8)$$

The constants implied by the above O 's depend only on A .

Proof. Write

$$K_z(x) := \sum_{\substack{n \leq x \\ p(n) > z}} 1.$$

By Brun's Sieve (see Theorems 2.2 on page 68 and 2.5 on page 82 in [8]), we have that

$$K_z(x) = c_1 \frac{x}{\log z} + O\left(\frac{x}{\log^2 z}\right) \quad \text{if } z < x^{1/\log_2 x}, \quad (9)$$

and

$$K_z(x) \ll \frac{x}{\log z} \quad \text{if } 1 \leq z \leq x. \quad (10)$$

We shall now assume that $z < x^{1/\log_2^3 x}$. Using partial summation, we have

$$L_z(x) = \int_1^x \frac{dK_z(t)}{t} = \frac{1}{x} K_z(x) + \int_1^x \frac{K_z(t)}{t^2} dt.$$

Clearly,

$$\frac{1}{x} K_z(x) = O\left(\frac{1}{\log z}\right) \quad (11)$$

by estimate (9). We break the integral at $x^{1/\log_2 x}$. By estimate (10), we get

$$\int_1^{x^{1/\log_2 x}} \frac{K_z(t)}{t^2} dt \ll \frac{1}{\log z} \int_1^{x^{1/\log_2 x}} \frac{dt}{t} \ll \frac{\log x}{\log z \log_2 x}. \quad (12)$$

For the second range we use estimate (9) to get

$$\begin{aligned} \int_{x^{1/\log_2 x}}^x \frac{K_z(t)}{t^2} dt &= \frac{c_1}{\log z} \left(1 + O\left(\frac{1}{\log z}\right)\right) \int_{x^{1/\log_2 x}}^x \frac{dt}{t} \\ &= \frac{c_1 \log x}{\log z} \left(1 + O\left(\frac{1}{\log z}\right)\right) \left(1 + O\left(\frac{1}{\log_2 x}\right)\right). \end{aligned} \quad (13)$$

Collecting together all estimates (11)–(13) we get

$$L_z(x) = \frac{c_1 \log x}{\log z} + O\left(\frac{\log x}{\log^2 z} + \frac{\log x}{\log z \log_2 x} + \log z\right),$$

and it is easy to see that the above error is bounded above as in (7) when $z \leq (\log x)^A$, as in the hypothesis of the lemma.

For (8), note that

$$\begin{aligned}
M_z(x) &= \sum_{\substack{n \leq x \\ p(n) > z}} \frac{1}{\varphi(n)} = \sum_{\substack{n \leq x \\ p(n) > z}} \frac{1}{n} \sum_{d|n} \frac{\mu^2(d)}{\varphi(d)} \\
&= \sum_{\substack{d \leq x \\ p(d) > z}} \frac{\mu^2(d)}{d\varphi(d)} \sum_{\substack{m \leq x/d \\ p(m) > z}} \frac{1}{m} = \sum_{\substack{d \leq x \\ p(d) > z}} \frac{\mu^2(d)}{d\varphi(d)} L_z(x/d).
\end{aligned}$$

When $d = 1$, $\mu^2(d)/d\varphi(d) = 1$, while when $d > 1$, since $p(d) > z$, it follows that

$$\sum_{\substack{d > 1 \\ p(d) > z}} \frac{\mu^2(d)}{d\varphi(d)} \leq \sum_{d > z} \frac{1}{d\varphi(d)} \ll \frac{1}{z},$$

where the last estimate above is due to Landau. Thus,

$$\begin{aligned}
M_z(x) &= L_z(x) + \sum_{\substack{1 < d \leq x \\ z \leq p(d)}} \frac{\mu^2(d)}{d\varphi(d)} L_z(x/d) = L_z(x) + O\left(\frac{L_z(x)}{z}\right) \\
&= c_1 \frac{\log x}{\log z} + O\left(\frac{\log x}{\log^2 z}\right),
\end{aligned}$$

which completes the proof of the lemma. \square

For $x, z > 0$, let

$$\mathcal{D}_z(x) = \{n \leq x : p(n) > z\}$$

and let $\tau_z(m)$ be the number of divisors of m in $\mathcal{D}_z(m)$.

Lemma 5. *Let $A > 0$ and $1 \leq z \leq A \frac{\log x}{\log_2^4 x}$. We then have*

$$R_z(x) := \sum_{p \leq x} \tau_z(p-1) = c_1 \frac{x}{\log z} + O\left(\frac{x}{\log^2 z}\right) \tag{14}$$

and

$$S_z(x) := \sum_{p \leq x} \frac{\tau_z(p-1)}{p} = c_1 \frac{\log x}{\log z} + O\left(\frac{\log x}{\log^2 z}\right), \tag{15}$$

where the constants implied in O above depend only on A .

Proof. Let $y \leq x$ be any positive real number. Our plan is to estimate $R_z(y)$, so proving (14), and then use partial summation to prove (15).

Note that

$$R_z(y) = \sum_{p \leq y} \tau_z(p-1) \leq \sum_{p \leq y} \tau(p-1) \leq 2 \sum_{d \leq \sqrt{y}} \pi(y; d, 1) \ll y, \quad (16)$$

where the last estimate follows from the Brun–Titchmarsh inequality. (Corollary 1 in [4] gives a more precise estimate.) We shall use this estimate when y is relatively small. In general,

$$R_z(y) = \sum_{p \leq y} \tau_z(p-1) = \sum_{p \leq y} \sum_{\substack{d \in \mathcal{D}_z(y) \\ d|p-1}} 1 = \sum_{d \in \mathcal{D}_z(y)} \pi(y; d, 1).$$

Assume now that $y > e^{z \log^2 z}$. We write B for a constant to be determined later. If y is large, we then split the sum appearing in $R_z(y)$ at $Q := \frac{y}{P_z \log^B y}$. Then,

$$R_z(y) = \sum_{d \in \mathcal{D}_z(Q)} \pi(y; d, 1) + \sum_{\substack{Q < d \leq y \\ p(d) > z}} \pi(y; d, 1) := R_1 + R_2. \quad (17)$$

Note that if $d > Q$ and $p \leq y$ is a prime with $p \equiv 1 \pmod{d}$, then $p = 1 + du$ with $u < y/Q = P_z \log^B y$. Thus,

$$R_2 \leq \sum_{u \leq P_z \log^B y} \pi(y; u, 1).$$

By the Brun–Titchmarsh inequality, we get

$$\begin{aligned} R_2 &\ll \pi(y) \sum_{u \leq P_z \log^B y} \frac{1}{\varphi(u)} \ll \pi(y) \log(P_z \log^B y) \ll \frac{y \log_2 y}{\log y} + \frac{yz}{\log y} \\ &\ll \frac{y}{\log^2 z}, \end{aligned} \quad (18)$$

where the last inequality above holds because $y > e^{z \log^2 z}$.

We now deal with R_1 . We claim that

$$R_1 = \pi(y) \sum_{d \in \mathcal{D}_z(Q)} \frac{1}{\varphi(d)} + O\left(\frac{y}{\log^2 y}\right) \quad (19)$$

holds if B is suitably chosen.

Indeed, note that by the principle of inclusion and exclusion, we have

$$R_1 = \sum_{\substack{d \leq Q \\ p(d) > z}} \pi(y; d, 1) = \sum_{r|P_z} \mu(r) \sum_{\substack{n \leq Q \\ r|n}} \pi(y; n, 1).$$

Thus,

$$R_1 = E_z(y) + \sum_{r|P_z} \mu(r) \sum_{\substack{n \leq Q \\ r|n}} \frac{\pi(y)}{\varphi(n)},$$

where $E_z(y)$ has been defined in Lemma 3. By Lemma 3, the estimate

$$R_1 = \sum_{r|P_z} \mu(r) \sum_{\substack{n \leq Q \\ r|n}} \frac{\pi(y)}{\varphi(n)} + O\left(\frac{y}{\log^C y}\right) = \pi(y) \sum_{d \in \mathcal{D}_z(Q)} \frac{1}{\varphi(d)} + O\left(\frac{y}{\log^C y}\right)$$

holds with any value of $C > 0$ provided that B is chosen to be sufficiently large with respect to C . We set $C := 2$, and we obtain (19). Since

$$Q = \frac{y}{P_z \log^B y} > y^{1/2} > \exp\left(\frac{1}{2}z \log^2 z\right),$$

it follows that $z \ll Q/\log^2 Q$, and we are therefore entitled to apply Lemma 4 and conclude that

$$\begin{aligned} R_1 &= c_1 \frac{\pi(y) \log Q}{\log z} + O\left(\frac{\pi(y) \log Q}{\log^2 z}\right) + O\left(\frac{y}{\log^2 y}\right) \\ &= c_1 \frac{y}{\log z} + O\left(\frac{y}{\log^2 z}\right). \end{aligned} \tag{20}$$

Combining (17)–(20), we get that

$$R_z(y) = c_1 \frac{y}{\log z} + O\left(\frac{y}{\log^2 z}\right) \tag{21}$$

holds when $y > e^{z \log^2 z}$, which in particular proves estimate (14). To arrive at (15), we now simply use partial summation to get that

$$\begin{aligned} S_z(x) &= \int_1^x \frac{dR_z(t)}{t} = \frac{R_z(t)}{t} \Big|_{t=1}^{t=x} + \int_1^x \frac{R_z(t)}{t^2} dt \\ &= \int_1^{e^{z \log^2 z}} \frac{R_z(t)}{t^2} dt + \int_{e^{z \log^2 z}}^x \frac{R_z(t)}{t^2} dt + O(1). \end{aligned}$$

The first integral above is, by (16),

$$\int_1^{e^{z \log^2 z}} \frac{R_z(t)}{t^2} dt \ll \int_1^{e^{z \log^2 z}} \frac{1}{t} dt \leq z \log^2 z \ll \frac{\log x}{\log^2 z}, \quad (22)$$

while the second integral above is, by (21),

$$\begin{aligned} \int_{e^{z \log^2 z}}^x \frac{R_z(t)}{t^2} dt &= \frac{c_1}{\log z} \int_{e^{z \log^2 z}}^x \frac{1}{t} dt + O\left(\frac{1}{\log^2 z} \int_{e^{z \log^2 z}}^x \frac{1}{t} dt\right) \\ &= c_1 \frac{\log x}{\log z} - c_1 z \log z + O\left(\frac{\log x}{\log^2 z}\right) \\ &= c_1 \frac{\log x}{\log z} + O\left(\frac{\log x}{\log^2 z}\right), \end{aligned} \quad (23)$$

and (15) now follows from (22) and (23). \square

Lemma 6. (i) Let A and z be as in Lemma 5 and $1 \leq u \leq x$ be any positive integer with $p(u) > z$. Then

$$S_{u,z}(x) := \sum_{\substack{p \leq x \\ p \equiv 1 \pmod{u}}} \frac{\tau_z(p-1)}{p} \ll \frac{\tau(u)}{u} S_z(x) \log x. \quad (24)$$

(ii) Let $A_1 > 0$, $0 < A_2 < 1/2$, $u < \log^{A_1} x$ and $\log^{A_2} x < z \leq \frac{\sqrt{\log x}}{\log_2^6 x}$.

Assume that $p(u) > z$. Then

$$R_{u,z}(x) := \sum_{\substack{p \leq x \\ p \equiv 1 \pmod{u}}} \tau_z(p-1) = c_1 \frac{\tau(u)}{u} \frac{x}{\log z} + O\left(\frac{x}{u \log^2 z}\right) \quad (25)$$

and

$$S_{u,z}(x) = \frac{\tau(u)}{u} S_z(x) \left(1 + O\left(\frac{1}{\log z}\right)\right). \quad (26)$$

The implied constants depend at most on A and A_1, A_2 , respectively.

Proof. To see inequality (24), we replace the prime summand p with an integer summand n , so that

$$S_{u,z}(x) \leq \sum_{\substack{n \leq x \\ n \equiv 1 \pmod{u}}} \frac{\tau_z(n-1)}{n-1} = \sum_{\substack{n \leq x \\ n \equiv 1 \pmod{u}}} \frac{1}{n-1} \sum_{\substack{d \in \mathcal{D}_z(x) \\ d|n-1}} 1.$$

Thus,

$$\begin{aligned}
S_{u,z}(x) &\leq \sum_{d \in \mathcal{D}_z(x)} \sum_{\substack{n \equiv 1 \pmod{[u,d]} \\ n \leq x}} \frac{1}{n-1} = \sum_{d \in \mathcal{D}_z(x)} \frac{1}{[u,d]} \sum_{m \leq x/[u,d]} \frac{1}{m} \\
&\ll \log x \sum_{d \in \mathcal{D}_z(x)} \frac{1}{[u,d]} \leq \log x \frac{\tau(u)}{u} \sum_{d \in \mathcal{D}_z(x)} \frac{1}{d} \\
&= \frac{\tau(u)}{u} L_z(x) \log x \ll \frac{\tau(u)}{u} S_z(x) \log x,
\end{aligned}$$

where in the above inequalities we used Lemmas 4 and 5.

For inequality (26), let us first notice that under the conditions (ii), we have that $\Omega(u) \ll 1$; hence, $\tau(u) \ll 1$, and also that

$$\frac{u\varphi(d)}{\varphi(ud)} = 1 + O\left(\frac{1}{z}\right) \quad (27)$$

holds uniformly in such positive integers u and all positive integers d .

The proof of (26) now closely follows the method of proof of (15). That is, let x be large, assume that z is fixed, and for $y \leq x$ write

$$R_{u,z}(y) := \sum_{\substack{p \leq y \\ p \equiv 1 \pmod{u}}} \tau_z(p-1) = \sum_{d \in \mathcal{D}_z(y)} \pi(y; [u, d], 1).$$

Let $w := \exp\left(\frac{\sqrt{\log x}}{\log_2 x}\right)$. Note that for large x the inequality $z < \frac{\log y}{(\log_2 y)^4}$ holds whenever $y > w$. For $y \leq w$, we use the trivial inequality

$$R_{u,z}(y) \ll \frac{y \log y}{u}. \quad (28)$$

Assume now that $y > w$. Since $\log y > \log^{1/3} x$ holds for large x , and $u < \log^{A_1} x$, we get that $u < \log^{3A_1} y$. We write B for a constant to be determined later and we split the sum appearing in $R_{u,z}(y)$ at $Q := \frac{y}{P_z \log^B y}$.

Thus,

$$R_{u,z}(y) = \sum_{d \in \mathcal{D}_z(Q)} \pi(y; [u, d], 1) + \sum_{\substack{Q < d \leq y \\ p(d) > z}} \pi(y; [u, d], 1) := R_1 + R_2.$$

It is easy to see that

$$R_2 \leq \sum_{d \leq P_z \log^B y} \pi(y; [u, d], 1).$$

Thus, by the Brun–Titchmarsh inequality,

$$\begin{aligned} R_2 &\ll \pi(y) \sum_{d \leq P_z \log^B y} \frac{1}{\varphi([u, d])} \ll \pi(y) \frac{\tau(u)}{u} \log(P_z \log^{3A_1+B} y) \\ &\ll \frac{y \log_2 y}{u \log y} + \frac{yz}{u \log y} \ll \frac{y}{u \log^2 z}, \end{aligned} \quad (29)$$

where we used $\tau(u) \ll 1$ together with (27).

We now deal with R_1 . We claim, as in the proof of Lemma 5, that

$$R_1 = \pi(y) \sum_{d \in \mathcal{D}_z(Q)} \frac{1}{\varphi([u, d])} + O\left(\frac{y}{u \log^2 y}\right) \quad (30)$$

holds if B is suitably chosen.

Indeed, note that since u and P_z are coprime, by the principle of inclusion and exclusion, we have

$$R_1 = \sum_{\substack{d \leq Q \\ p(d) > z}} \pi(y; [u, d], 1) = \sum_{r|P_z} \mu(r) \sum_{\substack{n \leq Q \\ r|n}} \pi(y; [u, n], 1).$$

Thus,

$$R_1 = E_{u,z}(y) + \sum_{r|P_z} \mu(r) \sum_{\substack{n \leq Q \\ r|n}} \frac{\pi(y)}{\varphi([u, n])} = E_{u,z}(y) + \pi(y) \sum_{d \in \mathcal{D}_z(Q)} \frac{1}{\varphi([u, d])}, \quad (31)$$

where $E_{u,z}(y)$ is the sum appearing in Lemma 3. Estimate (30) now follows from (4). Since

$$\frac{Q}{u} > \frac{y}{P_z u \log^B x} > \frac{y}{P_z \log^{3A_1+B}} > y^{1/2} > \exp\left(\frac{1}{2} z \log^4 z\right),$$

it follows that $z \ll Q/(u \log^4(Q/u))$, and we are therefore entitled to apply Lemmas 4 and 5 together with estimate (27) and conclude that

$$\begin{aligned}
R_1 &= \pi(y) \sum_{d \in \mathcal{D}_z(Q)} \frac{1}{\varphi([u, d])} + O\left(\frac{y}{u \log^2 y}\right) \\
&= \pi(y) \sum_{d_1 \in \mathcal{D}_z(Q/u)} \frac{\tau(u)}{\varphi(ud_1)} + O\left(\pi(y) \sum_{\substack{Q/u < d_1 \leq Q \\ p(d) > z}} \frac{1}{\varphi(ud_1)}\right) + O\left(\frac{y}{u \log^2 y}\right) \\
&= \pi(y) \frac{\tau(u)}{u} M_z(Q/u) + O\left(\frac{\pi(y)}{uz} M_z(Q)\right) \\
&\quad + O\left(\frac{\pi(y)}{u} (M_z(Q) - M_z(Q/u))\right) + O\left(\frac{y}{u \log^2 y}\right) \\
&= c_1 \pi(y) \frac{\tau(u)}{u} \frac{\log(Q/u)}{\log z} + O\left(\frac{\pi(y) \log u}{u \log z}\right) + O\left(\frac{\pi(y) \log Q}{u \log^2 z}\right) \\
&= c_1 \frac{\tau(u)}{u} \frac{y}{\log z} + O\left(\frac{y}{u \log^2 z}\right). \tag{32}
\end{aligned}$$

Combining (28)–(32), we get that

$$R_{u,z}(y) = c_1 \frac{\tau(u)}{u} \frac{y}{\log z} + O\left(\frac{y}{u \log^2 z}\right) \tag{33}$$

holds when $y > w$, which proves estimate (25).

We now use partial summation to get that

$$\begin{aligned}
S_{u,z}(x) &= \int_1^x \frac{dR_{u,z}(t)}{t} = \frac{R_{u,z}(t)}{t} \Big|_{t=1}^{t=x} + \int_1^x \frac{R_{u,z}(t)}{t^2} dt \\
&= \int_1^w \frac{R_{u,z}(t)}{t^2} dt + \int_w^x \frac{R_{u,z}(t)}{t^2} dt + O\left(\frac{1}{u \log z}\right).
\end{aligned}$$

The first integral above is, by (28),

$$\int_1^w \frac{R_{u,z}(t)}{t^2} dt \ll \frac{1}{u} \int_1^w \frac{\log t}{t} dt \ll \frac{\log^2 w}{u} \ll \frac{\log x}{u \log^2 x} \ll \frac{\log x}{u \log^2 z}. \tag{34}$$

Finally, the second integral above is, by (33),

$$\begin{aligned}
\int_w^x \frac{R_{u,z}(t)}{t^2} dt &= c_1 \frac{\tau(u)}{u} \frac{1}{\log z} \int_w^x \frac{1}{t} dt + O\left(\frac{1}{u \log^2 z} \int_w^x \frac{1}{t} dt\right) \\
&= c_1 \frac{\tau(u)}{u} \frac{1}{\log z} (\log x - \log w) + O\left(\frac{\log x}{u \log^2 z}\right) \\
&= c_1 \frac{\tau(u)}{u} \frac{\log x}{\log z} + O\left(\frac{\log x}{u \log^2 z}\right) \\
&= c_1 \frac{\tau(u)}{u} S_z(x) \left(1 + O\left(\frac{1}{\log z}\right)\right),
\end{aligned}$$

which completes the proof of Lemma 6. \square

Lemma 7. *Let $z = z(x) := \frac{\sqrt{\log x}}{\log_2^6 x}$. Let $I(x) := (z, z^5]$. Let $\mathcal{Q}(x)$ be the set of all prime numbers p with $z < p \leq x$ such that $p - 1$ is not divisible by the square of any prime $q > z$, and $p - 1$ has at most 7 prime factors in the interval $I(x)$. Then for large x we have*

$$S'_z(x) := \sum_{p \in \mathcal{Q}(x)} \frac{\tau_z(p-1)}{p} > 0.7 S_z(x).$$

Proof. Let $\mathcal{Q}_1(x)$ be the set of those primes p such that $q^2 | p - 1$ for some $q > z$. Fix q . Assume first that $q > \log^A x$, where A is a constant to be determined later. Then, by (24),

$$S_{q^2,z}(x) \ll \frac{\log x}{q^2} S_z(x),$$

and therefore

$$\sum_{q > \log^A x} S_{q^2,z}(x) \ll (\log x) S_z(x) \sum_{q > \log^A x} \frac{1}{q^2} \ll \frac{S_z(x)}{\log^{A-1} x \log_2 x}.$$

Choosing $A = 1$, we see that

$$\sum_{q > \log x} S_{q^2,z}(x) \ll \frac{S_z(x)}{\log_2 x}.$$

Assume now that $q \in (z, \log x]$. By (26), it follows that $S_{q^2, z}(x) \ll S_z(x)/q^2$, and therefore

$$\sum_{z < q \leq \log x} S_{q^2, z}(x) \ll S_z(x) \sum_{q > z} \frac{1}{q^2} \ll \frac{S_z(x)}{z \log z} \ll \frac{S_z(x)}{\log_2 x}.$$

In particular, we have

$$\sum_{q > z} S_{q^2, z}(x) = O\left(\frac{\log x}{\log_2^2 x}\right). \quad (35)$$

We now let B be a positive integer to be fixed later, and assume that u is a squarefree number having $\omega(u) = B$, and such that all its prime factors are in the interval $I(x)$. Let \mathcal{U}_B be the set of such numbers u . Since B is fixed, we have $u < z^{5B} < \log^{5B/2} x$, and therefore, by (26), we have

$$S_{u, z}(x) = \frac{2^B}{u} S_z(x) \left(1 + O\left(\frac{1}{\log_2 x}\right)\right).$$

Summing up over all possible values of $u \in \mathcal{U}_B$, we get

$$\sum_{u \in \mathcal{U}_B} S_{u, z}(x) = S_z(x) 2^B \left(\sum_{u \in \mathcal{U}_B} \frac{1}{u}\right) \left(1 + O\left(\frac{1}{\log_2 x}\right)\right).$$

Clearly,

$$\begin{aligned} \sum_{u \in \mathcal{U}_B} \frac{1}{u} &\leq \frac{1}{B!} \left(\sum_{p \in I(x)} \frac{1}{p}\right)^B = \frac{1}{B!} \left(\log_2(z^5) - \log_2 z + O\left(\frac{1}{\log z}\right)\right)^B \\ &= \frac{(\log 5)^B}{B!} \left(1 + O\left(\frac{1}{\log_2 x}\right)\right). \end{aligned}$$

Hence,

$$\sum_{u \in \mathcal{U}_B} S_{u, z}(x) \leq \frac{(2 \log 5)^B}{B!} S_z(x) \left(1 + O\left(\frac{1}{\log_2 x}\right)\right).$$

Since $(2 \log 5)^8/8! < 0.286$, we have for $B = 8$ and large x that

$$\sum_{u \in \mathcal{U}_B} S_{u, z}(x) < 0.29 S_z(x). \quad (36)$$

Thus, with (35) and (36), and using $\sum_{p \leq z} \tau_z(p-1)/p = \sum_{p \leq z} 1/p \ll \log_2 z$,

$$\begin{aligned} S'_z(x) &= \sum_{p \in \mathcal{Q}(x)} \frac{\tau_z(p-1)}{p-1} \\ &\geq S_z(x) - \sum_{q > z} S_{q^2, z}(x) - \sum_{u \in \mathcal{U}_B} S_{u, z}(x) - \sum_{p \leq z} \frac{\tau_z(p-1)}{p} > 0.7S_z(x) \end{aligned}$$

for large x , which completes the proof of Lemma 7. \square

3 The Proof of Theorem 1

We shall analyze the expression

$$T(x) := \sum_{n \leq x} \frac{\tau(\varphi(n))}{n}.$$

3.1 The upper bound

For every positive integer n we write $\beta(n) := \prod_{p|n} p$. Then n can be written as $n = \beta(n)m$ where all prime factors of m are among the prime factors of $\beta(n)$. Moreover, $\varphi(n) = m\varphi(\beta(n))$, and therefore $\tau(\varphi(n)) \leq \tau(m)\tau(\varphi(\beta(n)))$. Thus,

$$\begin{aligned} T(x) &\leq \sum_{\substack{k \leq x \\ \mu(k) \neq 0}} \sum_{m \leq x/k} \frac{\tau(m)\tau(\varphi(k))}{mk} = \sum_{\substack{k \leq x \\ \mu(k) \neq 0}} \frac{\tau(\varphi(k))}{k} \sum_{m \leq x/k} \frac{\tau(m)}{m} \\ &\ll U(x) \log^2 x, \end{aligned} \tag{37}$$

where

$$U(x) := \sum_{\substack{n \leq x \\ \mu(n) \neq 0}} \frac{\tau(\varphi(n))}{n}.$$

We now let $z = z(x)$ be as in Lemma 7. For every positive integer n we write $\tau'_z(n)$ for the number of divisors of the largest divisor of n composed only of primes $p \leq z$. Clearly, $\tau(n) = \tau_z(n)\tau'_z(n)$. If $n \leq x$ and $p^\alpha \parallel n$, then $\alpha + 1 < 2 \log x$. This shows that

$$\tau'_z(n) \leq (2 \log x)^{\pi(z)} < \exp\left(10 \frac{\sqrt{\log x}}{\log_2^5 x}\right). \tag{38}$$

Using also the fact that $\tau_z(ab) \leq \tau_z(a)\tau_z(b)$ holds for all positive integers a and b , together with (37) and (38), we get that the inequality

$$T(x) \leq V_z(x) \exp\left(O\left(\frac{\sqrt{\log x}}{\log^5 x}\right)\right)$$

holds, where

$$V_z(x) := \sum_{\substack{n \leq x \\ \mu(n) \neq 0}} \prod_{p|n} \frac{\tau_z(p-1)}{p}.$$

To find an upper bound on the last expression, we use Rankin's method. Let $s = s(x) < 1$ be a small positive real number depending on x to be determined later, and note that

$$\begin{aligned} V_z(x) &\leq x^s \sum_{\substack{n \leq x \\ \mu(n) \neq 0}} \frac{1}{n^s} \prod_{p|n} \frac{\tau_z(p-1)}{p} = x^s \sum_{\substack{n \leq x \\ \mu(n) \neq 0}} \prod_{p|n} \frac{\tau_z(p-1)}{p^{1+s}} \\ &\leq x^s \prod_{p \leq x} \left(1 + \frac{\tau_z(p-1)}{p^{1+s}}\right) \leq \exp\left(s \log x + \sum_{p \leq x} \frac{\tau_z(p-1)}{p^{1+s}}\right). \end{aligned}$$

We now find s in such a way as to minimize

$$f_s(x) := s \log x + \sum_{p \leq x} \frac{\tau_z(p-1)}{p^{1+s}}.$$

For this, recall that from the proof of Lemma 5, we have

$$\begin{aligned} \sum_{p \leq x} \frac{\tau_z(p-1)}{p^{1+s}} &= \int_2^x \frac{dR_z(t)}{t^{1+s}} = \frac{R_z(t)}{t^{1+s}} \Big|_{t=2}^{t=x} + (1+s) \int_2^x \frac{R_z(t)}{t^{2+s}} dt \\ &= (1+s) \int_2^x \frac{R_z(t)}{t^{2+s}} dt + O\left(\frac{1}{x^s \log z}\right). \end{aligned} \quad (39)$$

We shall later choose $s := \frac{c_1^{1/2}}{\sqrt{\log x \log z}}$. In order to compute the above integral (39), we split it at $x_0 := e^{1/(s \log^2 z)}$. In the first (smaller) range, we use the fact that $R_z(t) \ll t$ and that $t^s \geq 1$ to get

$$\int_2^{x_0} \frac{R_z(t)}{t^{2+s}} dt \ll \int_2^{x_0} \frac{1}{t} dt \leq \log x_0 = \frac{1}{s \log^2 z}.$$

Note that $x_0 \geq e^{z \log^2 z}$ for x sufficiently large. Thus, from the estimate of $R_z(t)$ from Lemma 5, we have

$$\begin{aligned} \int_{x_0}^x \frac{R_z(t)}{t^{2+s}} dt &= \frac{c_1}{\log z} \int_{x_0}^x \frac{dt}{t^{1+s}} + O\left(\frac{1}{\log^2 z} \int_{x_0}^x \frac{dt}{t^{1+s}}\right) \\ &= \frac{c_1}{s \log z} (x_0^{-s} - x^{-s}) + O\left(\frac{1}{s \log^2 z}\right) \\ &= \frac{c_1}{s \log z} + O\left(\frac{1}{s \log^2 z}\right), \end{aligned}$$

where we used the fact that $x_0^s = \exp(1/\log^2 z) = 1 + O(1/\log^2 z)$. Thus,

$$f_s(x) = s \log x + \frac{c_1(1+s)}{s \log z} + O\left(\frac{1}{s \log^2 z}\right).$$

With our choice for s , we have

$$f_s(x) = 2c_1^{1/2} \left(\frac{\log x}{\log z}\right)^{1/2} + O\left(\frac{\log^{1/2} x}{\log^{3/2} z}\right).$$

Since $z = \sqrt{\log x}/\log_2^6 x$, we get

$$f_s(x) = 2^{3/2} c_1^{1/2} \left(\frac{\log x}{\log_2 x}\right)^{1/2} \left(1 + O\left(\frac{\log_3 x}{\log_2 x}\right)\right).$$

Thus, we have obtained the upper bound

$$T(x) \leq \exp\left(2^{3/2} e^{-\gamma/2} \left(\frac{\log x}{\log_2 x}\right)^{1/2} \left(1 + O\left(\frac{\log_3 x}{\log_2 x}\right)\right)\right). \quad (40)$$

Since $D_\varphi(x) \leq xT(x)$, we have the upper bound in Theorem 1.

3.2 The lower bound

We write c_2 for a constant to be computed later and we set

$$v := \left\lfloor c_2 \left(\frac{\log x}{\log_2 x}\right)^{1/2} \right\rfloor - 2.$$

We write

$$y := \exp\left(\frac{1}{c_2} (\log x \log_2 x)^{1/2}\right).$$

We now write $z := z(y)$, where the function z is the one appearing in Lemma 7. We write $\mathcal{Q} := \mathcal{Q}(y)$ for the set of primes defined in Lemma 7. Recall that $\mathcal{Q}(y)$ is the set of primes $p \leq y$ such that $p - 1$ is not divisible by the square of any prime $q > z$ and $p - 1$ has at most 7 distinct prime factors in $(z, z^5]$. Consider squarefree numbers n having $\omega(n) = v$ and such that all their prime factors are in \mathcal{Q} . Let \mathcal{N} be the set of those numbers. It is clear that if $n \in \mathcal{N}$ then $n \leq x/y^2$. Let

$$V_{\mathcal{N},z}(x) := \sum_{n \in \mathcal{N}} \frac{\tau_z(\varphi(n))}{n}. \quad (41)$$

For a number $n \in \mathcal{N}$, we write $\tau''(n) := \prod_{p|n} \tau_z(p-1)$ and we look at the sum

$$W_{\mathcal{N}} := \sum_{n \in \mathcal{N}} \frac{\tau''(n)}{n}.$$

By the binomial formula, and Stirling's formula, it follows that

$$W_{\mathcal{N}} \leq \frac{1}{v!} \left(\sum_{p \in \mathcal{Q}} \frac{\tau_z(p-1)}{p} \right)^v \asymp \frac{1}{\sqrt{v}} \left(\frac{e}{v} \sum_{p \in \mathcal{Q}} \frac{\tau_z(p-1)}{p} \right)^v.$$

A simple calculation based on Lemmas 5 and 7 shows that

$$0.7e \frac{S_z(y)}{v} \leq \frac{e}{v} \sum_{p \in \mathcal{Q}} \frac{\tau_z(p-1)}{p} \leq e \frac{S_z(y)}{v}$$

and that

$$\frac{S_z(y)}{v} = c_3 + O\left(\frac{\log_3 x}{\log_2 x}\right),$$

where $c_3 := \frac{4c_1}{c_2^2}$. We now observe that

$$\begin{aligned} & \frac{1}{v!} \left(\sum_{p \in \mathcal{Q}} \frac{\tau_z(p-1)}{p} \right)^v \\ & \leq W_{\mathcal{N}} + \frac{1}{(v-2)!} \left(\sum_{p \in \mathcal{Q}} \frac{\tau_z(p-1)}{p} \right)^{v-2} \frac{1}{2} \left(\sum_{p \in \mathcal{Q}} \frac{\tau_z(p-1)^2}{p^2} \right) \\ & = W_{\mathcal{N}} + O \left(\frac{1}{v!} \left(\sum_{p \in \mathcal{Q}} \frac{\tau_z(p-1)}{p} \right)^v \left(\sum_{p \in \mathcal{Q}} \frac{\tau_z(p-1)^2}{p^2} \right) \right). \end{aligned}$$

Since $p \in \mathcal{Q}$ implies that $p > z$, and since for large x the inequality $\tau_z(p-1) \leq \tau(p-1) < p^{1/4}$ holds for all $p > z$, we get

$$\sum_{p \in \mathcal{Q}} \frac{\tau_z(p-1)^2}{p^2} \leq \sum_{p > z} \frac{1}{p^{3/2}} \ll \frac{1}{z^{1/2} \log z}.$$

Hence, the above argument shows that

$$\begin{aligned} W_{\mathcal{N}} & = \frac{1}{v!} \left(\sum_{p \in \mathcal{Q}} \frac{\tau_z(p-1)}{p} \right)^v \left(1 + O \left(\frac{1}{z^{1/2} \log z} \right) \right) \\ & \gg \frac{1}{\sqrt{v}} \left(c_4 + O \left(\frac{\log_3 x}{\log_2 x} \right) \right)^v, \end{aligned}$$

where $c_4 := 0.7c_3$.

We now select the subset \mathcal{M} of \mathcal{N} formed only by those n such that there is no prime number $q > z^5$ such that $q|p_1 - 1$ and $q|p_2 - 1$ holds for two distinct primes p_1 and p_2 dividing n . Note that this is equivalent to the fact that $\varphi(n)$ is not a multiple of a square of a prime $q > z^5$. To understand the sum $W_{\mathcal{M}}$ restricted only to those $n \in \mathcal{M}$, let us fix a prime number q . Then, summing up $\tau''(n)/n$ only over those n such that $q^2|\varphi(n)$ for the fixed prime $q > z^5$, we get

$$\begin{aligned} W_{q,\mathcal{N}} & := \sum_{\substack{n \in \mathcal{N} \\ q^2|\varphi(n)}} \frac{\tau''(n)}{n} \ll \frac{1}{(v-2)!} \left(\sum_{p \in \mathcal{Q}} \frac{\tau_z(p-1)}{p} \right)^{v-2} S_{q,z}(y)^2 \\ & \ll W_{\mathcal{N}} S_{q,z}(y)^2. \end{aligned}$$

Assume first that $q > \log^A y$, where A is a constant to be determined later. In this case, by Lemmas 6 and 5,

$$S_{q,z}(y) \ll \frac{S_z(y) \log y}{q} \ll \frac{\log^2 y}{q \log z},$$

and therefore

$$W_{q,\mathcal{N}} \ll W_{\mathcal{N}} \frac{\log^4 y}{q^2 \log^2 z}. \quad (42)$$

Summing up inequalities (42) for all $q \geq \log^A y$, we get

$$\sum_{q \geq \log^A y} W_{q,\mathcal{N}} \ll W_{\mathcal{N}} \frac{\log^4 y}{\log^2 z} \sum_{q > \log^A y} \frac{1}{q^2} \ll \frac{W_{\mathcal{N}}}{(\log y)^{A-4} \log^2 y} \ll \frac{W_{\mathcal{N}}}{\log y},$$

provided that we choose $A := 5$.

When $q < \log^5 y$, then the same argument based again on Lemma 6, shows that

$$W_{q,\mathcal{N}} \ll W_{\mathcal{N}} \frac{\log^2 y}{q^2 \log^2 z},$$

and therefore

$$\sum_{z^5 < q < \log^5 y} W_{q,\mathcal{N}} \ll W_{\mathcal{N}} \frac{\log^2 y}{\log^2 z} \sum_{q > z^5} \frac{1}{q^2} \ll W_{\mathcal{N}} \frac{\log^2 y}{z^5 \log^3 z} \ll W_{\mathcal{N}} \frac{\log_2^{27} y}{\log^{1/2} y}.$$

This shows that

$$\begin{aligned} W_{\mathcal{M}} &:= \sum_{n \in \mathcal{M}} \frac{\tau''(n)}{n} \geq W_{\mathcal{N}} - \sum_{q > z^5} W_{q,\mathcal{N}} = W_{\mathcal{N}}(1 + o(1)) \\ &\gg \frac{1}{\sqrt{v}} \left(c_4 + O\left(\frac{\log_3 x}{\log_2 x}\right) \right)^v. \end{aligned} \quad (43)$$

Notice now that if $n \in \mathcal{M}$ then

$$\tau_z(\varphi(n)) \geq \tau_{z^5}(\varphi(n)) = \prod_{p|n} \tau_{z^5}(p-1) \geq \prod_{p|n} \left(\frac{\tau_z(p-1)}{2^7} \right) \geq \frac{1}{2^{7v}} \tau''(n). \quad (44)$$

Thus, with (41) and (43), we get

$$V_{\mathcal{N},z}(x) \geq \frac{1}{2^{7v}} W_{\mathcal{M}} \gg \frac{1}{\sqrt{v}} \left(c_5 + O\left(\frac{\log_3 x}{\log_2 x}\right) \right)^v,$$

where $c_5 := 2^{-7}c_4$. So, we see that

$$V_{\mathcal{N},z}(x) \geq \exp \left(c_6 \left(\frac{\log x}{\log_2 x} \right)^{1/2} \left(1 + O \left(\frac{\log_3 x}{\log_2 x} \right) \right) \right), \quad (45)$$

holds for large x with

$$c_6 := c_2 \log \left(\frac{0.7 \cdot 4e^{1-\gamma}}{c_2^2 2^7} \right). \quad (46)$$

To see the lower bound in Theorem 1, we look at integers np , where $n \in \mathcal{M}$ and p is prime with $y < p \leq x/n$. Each such integer np arises in a unique way, and the number of primes p corresponding to a particular n is $\pi(x/n) - \pi(y) \geq \frac{1}{2}\pi(x/n) \gg x/(n \log x)$. Further, $\tau(\varphi(np)) > \tau_z(\varphi(n))$. Thus,

$$D_\varphi(x) \gg \frac{x}{\log x} V_{\mathcal{N},z}(x). \quad (47)$$

So, letting $c_2 = (2.8/2^7)^{1/2} e^{-\gamma/2} > \frac{1}{7} e^{-\gamma/2}$, we have $c_6 = c_2$, and from (45) and (47), we have the lower bound in Theorem 1 for all large x .

4 The proof of Theorem 2

Part (i) follows immediately from the proof of Theorem 1. Indeed, $\lambda(n) | \varphi(n)$, therefore $\tau(\lambda(n)) \leq \tau(\varphi(n))$ holds for all positive integers n . In particular, $D_\lambda(x) \leq D_\varphi(x)$. For the lower bound, it suffices to note that if \mathcal{M} is the set of integers constructed in the proof of the lower bound for $D_\varphi(x)$, then $\varphi(n)$ is not divisible by the square of any prime $p > z^5$. In particular, the inequality $\tau_z(\lambda(n)) = \tau''(n)/2^{7v}$ also holds (compare with (44)). Thus, the lower bound on $D_\lambda(x)$ follows from the proof of the lower bound for $D_\varphi(x)$.

To see (ii), we put

$$\kappa = \lfloor 10 \log_2 x \rfloor \quad \text{and} \quad w = \frac{\sqrt{\log x}}{\log_2^2 x},$$

and let $\mathcal{E}_1(x)$ be the set of $n \leq x$ such that either $2^\kappa | n$ or there exists a prime $p | n$ with $p \equiv 1 \pmod{2^\kappa}$, and $\mathcal{E}_2(x)$ be the set of $n \leq x$ with $\omega(n) \leq w$.

Since $\tau(\phi(ab)) \leq \tau(\phi(a))\tau(\phi(b))$, we have

$$\begin{aligned} \sum_{n \in \mathcal{E}_1(x)} \frac{\tau(\lambda(n))}{n} &\leq \sum_{n \in \mathcal{E}_1(x)} \frac{\tau(\varphi(n))}{n} \\ &\leq \frac{\tau(2^\kappa)}{2^\kappa} \sum_{m \leq x/2^\kappa} \frac{\tau(\varphi(m))}{m} + \sum_{\substack{p \leq x \\ p \equiv 1 \pmod{2^\kappa}}} \frac{\tau(p-1)}{p} \sum_{m \leq x/p} \frac{\tau(\varphi(m))}{m}. \end{aligned}$$

We majorize the inner sums with $T(x)$, so that

$$\begin{aligned} \sum_{n \in \mathcal{E}_1(x)} \frac{\tau(\lambda(n))}{n} &\leq \left(\frac{\kappa+1}{2^\kappa} + S_{2^\kappa,1}(x) \right) T(x) \ll \frac{\kappa \log x}{2^\kappa} S_1(x) T(x) \\ &\ll \frac{\log^2 x \log_2 x T(x)}{(\log x)^{10 \log 2}} \ll \frac{T(x)}{\log^2 x}, \end{aligned} \quad (48)$$

where in the above estimates we used Lemmas 5 and 6 to estimate $S_{2^\kappa,1}(x)$ and $S_1(x)$, respectively, and the fact that $10 \log 2 > 4$.

Furthermore, by the multinomial formula and the Stirling formula,

$$\begin{aligned} \sum_{n \in \mathcal{E}_2(x)} \frac{\tau(\lambda(n))}{n} &\leq \sum_{n \in \mathcal{E}_2(x)} \frac{\tau(\varphi(n))}{n} \leq \sum_{k \leq w} \sum_{\substack{n \leq x \\ \omega(n)=k}} \frac{\tau(\varphi(n))}{n} \\ &\leq \sum_{k \leq w} \frac{1}{k!} \left(\sum_{p^\alpha \leq x} \frac{\tau(\varphi(p^\alpha))}{p^\alpha} \right)^k \\ &\leq \sum_{k \leq w} \frac{1}{k!} \left(\sum_{p \leq x} \frac{\tau(\varphi(p-1))}{p} + \sum_{\substack{2 \leq p \\ 2 \leq \alpha}} \frac{\tau(\varphi(p^\alpha))}{p^\alpha} \right)^k \\ &\leq \sum_{k \leq w} \frac{1}{k!} (S_1(x) + O(1))^k \leq \sum_{k \leq w} \left(\frac{eS_1(x) + O(1)}{k} \right)^k. \end{aligned}$$

Let c_7 be the constant implied by the last $O(1)$. Since $S_1(x) \gg \log x$, the

function $k \mapsto \left(\frac{eS_1(x) + c_7}{k}\right)^k$ is increasing for $k \leq w$ once x is large. Thus,

$$\begin{aligned} \sum_{n \in \mathcal{E}_2(x)} \frac{\tau(\lambda(n))}{n} &\leq w \left(\frac{eS_1(x) + c_7}{w}\right)^w \\ &= \exp\left(w \log\left(\frac{eS_1(x) + c_7}{w}\right) + \log w\right) \\ &= \exp\left(\left(\frac{1}{2} + o(1)\right) \frac{\sqrt{\log x}}{\log_2 x}\right) = o\left(\frac{T(x)}{\log^2 x}\right), \end{aligned} \quad (49)$$

where the last estimate above follows from estimate (45) and $T(x) \geq V_{\mathcal{N},z}(x)$.

Finally, if we set $\mathcal{E}_3(x)$ for the set of all positive integers $n \leq x$ not in $\mathcal{E}_1(x) \cup \mathcal{E}_2(x)$, we then notice that if $n \in \mathcal{E}_2(x)$, then $2^\alpha \|\varphi(n)$, where $\alpha \geq \omega(n) - 1 \geq w - 1$, and $2^\beta \|\lambda(n)$, where $\beta < \kappa$. Hence,

$$\frac{\tau(\varphi(n))}{\tau(\lambda(n))} \geq \frac{(\alpha + 1)}{\beta + 1} \gg \frac{\sqrt{\log x}}{\log_2^3 x}.$$

Thus,

$$\frac{1}{x} \sum_{n \in \mathcal{E}_3(x)} \tau(\lambda(n)) \ll \frac{\log_2^3 x}{\sqrt{\log x}} D_\varphi(x), \quad (50)$$

while estimates (48), (49) and partial summation show that

$$\begin{aligned} \frac{\log^2 x}{x} \sum_{n \in \mathcal{E}_1(x) \cup \mathcal{E}_2(x)} \tau(\lambda(n)) &\leq \log^2 x \sum_{n \in \mathcal{E}_1(x) \cup \mathcal{E}_2(x)} \frac{\tau(\lambda(n))}{n} \\ &\ll T(x) = \int_1^x \frac{d(tD_\varphi(t))}{t} \leq D_\varphi(x) + \int_1^x \frac{D_\varphi(t)}{t} dt \\ &\ll D_\varphi^*(x) \left(1 + \int_1^x \frac{dt}{t}\right) \ll D_\varphi^*(x) \log x. \end{aligned}$$

Therefore

$$\frac{1}{x} \sum_{n \in \mathcal{E}_1(x) \cup \mathcal{E}_2(x)} \tau(\lambda(n)) \ll \frac{D_\varphi^*(x)}{\log x}. \quad (51)$$

Clearly, summing up estimates (50) and (51) we get

$$D_\lambda(x) \ll \frac{\log_2^3 x}{\sqrt{\log x}} D_\varphi^*(x),$$

and the proof of Theorem 2 is complete.

References

- [1] V. I. Arnold, ‘Number-theoretical turbulence in Fermat-Euler arithmetics and large Young diagrams geometry statistics’, *J. Math. Fluid Mech.* **7** (2005), suppl. 1, S4–S50.
- [2] W. D. Banks, F. Luca and I. E. Shparlinski, ‘Arithmetic properties of $\varphi(n)/\lambda(n)$ and the structure of the multiplicative group modulo n ’, *Comment. Math. Helvetici* **81** (2006), 1–22.
- [3] W. D. Banks and I. E. Shparlinski, ‘On the number of sparse RSA exponents’, *J. Number Theory* **95** no. 2 (2002), 340–350.
- [4] E. Bombieri, J. B. Friedlander and H. Iwaniec, ‘Primes in arithmetic progressions to large moduli’, *Acta Math.* **156** no. 3-4 (1986), 203–251.
- [5] P. Erdős, ‘On the normal number of prime factors of $p - 1$ and some other related problems concerning Euler’s φ function’, *Quart. J. Math. (Oxford Ser.)* **6** (1935), 205–213.
- [6] P. Erdős, A. Granville, C. Pomerance and C. Spiro, ‘On the normal behaviour of the iterates of some arithmetic functions’, in: *Analytic Number Theory, Proc. of a Conference in Honour of P.T. Bateman*, Birkhäuser; Boston, Inc. 1990, 165–204.
- [7] P. Erdős and C. Pomerance, ‘On the normal number of prime factors of $\varphi(n)$ ’, *Rocky Mtn. J. of Math.* **15** (1985), 343–352.
- [8] H. Halberstam and H.-E. Richert, *Sieve methods*, Academic Press, 1974.
- [9] S. Li, ‘An improvement of Artin’s conjecture on average for composite moduli’, *Mathematika*, to appear.
- [10] M. R. Murty and V. K. Murty, ‘Prime divisors of Fourier coefficients of modular forms’, *Duke Mathematical Journal* **51** (1984), 57–76.
- [11] M. R. Murty and F. Saidak, ‘Non-abelian generalizations of the Erdős-Kac theorem’, *Canad. J. Math.* **56** (2004), 356–372.
- [12] J.-P. Serre, ‘Divisibilité des certaines fonctions arithmétiques’, *L’Ens. Math.* **22** (1976), 227–260.

- [13] J.-P. Serre, ‘Quelques applications du théorème de densité de Chebotarev’, *Inst. Hautes Études Sci. Publ. Math.* **54** (1981), 323–401.