SHIFTED-PRIME DIVISORS

KAI (STEVE) FAN AND CARL POMERANCE

For Helmut Maier on his seventieth birthday

Abstract. Let $\omega^*(n)$ denote the number of divisors of n that are shifted primes, that is, the number of divisors of n of the form $p-1$, with p prime. Studied by Prachar in an influential paper from 70 years ago, the higher moments of $\omega^*(n)$ are still somewhat a mystery. This paper addresses these higher moments and considers other related problems.

1. Introduction

Let $\omega(n)$ denote the number of different primes that divide n. This function has been well-studied, and in particular we know that

$$\begin{aligned} &\frac{1}{x}\sum_{n\le x}\omega(n) = \log\log x + O(1),\\ (1)\qquad &\frac{1}{x}\sum_{n\le x}\omega(n)^2 = (\log\log x)^2 + O(\log\log x), \end{aligned}$$

after results of Hardy–Ramanujan and Turán. Further, $\omega(n)$ obeys a normal distribution as given by the Erdős–Kac theorem. For extreme values, we know that

$$\omega(n) \le (1+o(1))\log n/\log\log n,\ \ n\to\infty, \tag{2}$$

a best-possible result of Ramanujan.

Consider the analogous function $\omega^*(n)$ which counts the number of shifted prime divisors of n, that is, the number of divisors of n of the form $p-1$, with p prime. One might guess that assertions like (1) and (2) hold as well for ω^*. And in fact, it is easy to prove that

$$\frac{1}{x}\sum_{n\le x}\omega^*(n) = \log\log x + O(1). \tag{3}$$

However, the analogy stops here. As it turns out, the function ω^* is considerably wilder than ω. In some sense, ω^* is closer to the total number $\tau(n)$ of divisors of n. For example, after work of Prachar [15]

Date: January 25, 2024.

we have $\omega^*(n) \geq n^{c/(\log\log n)^2}$ for some positive constant c and infinitely many n. This was improved in [1, Proposition 10] to

$$\omega^*(n) \geq n^{c/\log\log n} \tag{4}$$

for a positive constant c and infinitely many n, a result which is clearly best possible, but for the choice of c, due to the upper bound

$$\omega^*(n) \leq \tau(n) \leq n^{(\log 2+o(1))/\log\log n}, \ n \to \infty,$$

a result due to Wigert. (Also see [2, Section 3].)

This paper deals with the moments

$$M_k(x) := \frac{1}{x}\sum_{n\leq x}\omega^*(n)^k,$$

for $k = 2$ and 3. Prachar [15] showed that

$$M_2(x) \ll (\log x)^2. \tag{5}$$

In a letter to the same journal, Erdős [5] proved that

$$S_2(x) := \frac{1}{x}\sum_{[p-1,q-1]\leq x} 1 \ll (\log\log x)^3, \tag{6}$$

and indicated how the exponent 3 could be replaced by 1, and possibly even by 0. Here, p, q run over prime numbers and $[a, b]$ denotes the least common multiple of a and b. The connection of these results on $S_2(x)$ to Prachar's theorem is as follows. We have

$$M_2(x) = \frac{1}{x}\sum_{[p-1,q-1]\leq x}\left\lfloor\frac{x}{[p-1,q-1]}\right\rfloor, \tag{7}$$

where $[a, b]$ denotes the least common multiple of a, b, so that (6) and a partial summation argument imply that

$$M_2(x) \ll \log x(\log\log x)^3, \tag{8}$$

with the same remark pertaining to the exponent 3.

In a recent paper Murty and Murty [11], completed the proof that

$$M_2(x) \ll \log x, \tag{9}$$

and they showed the lower bound

$$M_2(x) \gg (\log\log x)^3, \tag{10}$$

which improves on the trivial bound

$$M_2(x) \geq \left(\frac{1}{x}\sum_{n\leq x}1\right)^{-1}\left(\frac{1}{x}\sum_{n\leq x}\omega^*(n)\right)^2 \gg (\log\log x)^2$$

implied by (3) and the Cauchy–Schwarz inequality. Further, they made the conjecture that there is a positive constant C such that

$$M_2(x) \sim C \log x, \quad x \to \infty. \tag{11}$$

The topic was picked up again by Ding [3] who, using the claim [11, Equation (4.8)], showed that

$$M_2(x) \gg \log x. \tag{12}$$

Further in [4] he gave a heuristic argument for the Murty–Murty conjecture (11) based on the Elliott–Halberstam conjecture, with $C = 2\zeta(2)\zeta(3)/\zeta(6) \approx 3.88719$.

However, as it turns out, there is an error in the proof of [11, Equation (4.8)]. In particular, the error term $O(x)$ there, which results from removing the floor symbol in (7), is only valid for those pairs p, q with $[p-1, q-1] \le x$ and not for all pairs $p, q \le x$. We show below in Section 7 how a modified version of Ding's argument [3] can save the proof of (12). Further, we show that not only is the proof of [11, Equation (4.8)] in error, but the assertion is false, see Section 8. This unfortunately seems to invalidate the heuristic in [4]. We certainly agree that the Murty–Murty conjecture (11) holds, but we think the correct constant is closer to 3.1. We give the results of some calculations that support this.

We conjecture that $M_k(x) \asymp (\log x)^{2^k-k-1}$ and prove this for the third moment $M_3(x)$. The proof is considerably more involved than the second moment, but hopefully we have presented it in a manner that leaves open the possibility of getting analogous results for higher moments.

We also consider the level sets $\{n : \omega^*(n) = j\}$, showing that for each fixed positive integer j, the natural density exists and is positive, with the sum of these densities being 1.

It may be worth pointing out that our methods used to treat the moments of $\omega^*(n)$ can be used to deal with the natural generalization where $p-1$ is replaced with $p+a$ for a fixed integer $a \ne 0$.

Throughout we let p, q, r, s, ℓ run over prime numbers. We let (m, n) denote the greatest common divisor of m, n, and $[m, n]$ their least common multiple. We also use the standard order notations $\ll, \asymp, \gg$ from analytic number theory.

2. The constant C in (11)

In Section 8 we shall prove Theorem 5 which not only shows that the correction that we make to Ding's proof of the lower bound for $M_2(x)$ is necessary (see Section 7), but it also suggests that the constant $C =$

$2\zeta(2)\zeta(3)/\zeta(6)$ for the Murty–Murty conjecture shown by his heuristic argument in [4] is probably incorrect. So, what is the correct value of C? We leave this as an unsolved problem, but perhaps it is helpful to look at some actual numbers. We have numerical calculations of the values of $M_2(x) = \frac{1}{x}\sum_{n\le x}\omega^*(n)^2$ with $x = 10^k$ and $2 \le k \le 10$ using Mathematica. In view of the relation

$$M_2(x) = \int_1^x \frac{S_2(t)}{t}\,dt + O(1),$$

we also calculated the values of $S_2(x) := (1/x)\sum_{[p-1,q-1]\le x} 1$ for x in the same range. These values are recorded in the table below.

Table 1. Numerical values of $M_2(10^k)$ and $S_2(10^k)$

k	$M_2(10^k)$	$S_2(10^k)$
2	9.71	2.42
3	15.530	2.624
4	21.9128	2.8175
5	28.49311	2.88636
6	35.261891	2.950910
7	42.1296839	2.9923851
8	49.02181351	3.02166709
9	56.067311859	3.043042188
10	63.1033824202	3.0595625,181

The M_2 numbers in Table 1 seem to fit nicely with $3\log x - 6$, and the S_2 numbers may fit with $3.2(1 - 1/\log x)$. Perhaps $C \approx 3.1$?

3. The level sets of $\omega^*(n)$

For $x, y \ge 1$, let $N(x,y) := \#\{n \le x : \omega^*(n) \ge y\}$. The following theorem provides upper and lower bounds for $N(x,y)$.

Theorem 1. *There exists a suitable constant $c > 0$ such that*

$$\left\lfloor \frac{x}{y^{c\log\log y}} \right\rfloor \le N(x,y) \ll \frac{x\log y}{y}$$

for all $x \ge 1$ and all sufficiently large y.

Proof. The lower bound follows immediately from [1, Proposition 10], which asserts that there exists some constant $c_0 > 0$ such that for all $z > 100$, there is some positive integer $m_z < z$ with $\omega^*(m_z) >$

$e^{c_0 \log z/\log\log z}$. Taking $z = y^{c\log\log y}$ with some suitable constant $c > 0$, we have $\omega^*(m_z) > y$ and hence

$$N(x, y) \geq \left\lfloor \frac{x}{m_z} \right\rfloor \geq \left\lfloor \frac{x}{y^{c\log\log y}} \right\rfloor.$$

To prove the upper bound, we first note that since the average of $\omega^*(n)$ for $n \leq x$ is $\log\log x + O(1)$, it follows that $N(x, y) \ll x\log\log x/y$. So we have the desired upper bound when $y > (\log x)^{.05}$, say. Assume now that $y \leq (\log x)^{.05}$, and let $z = \exp(y^{19})$, so that $z = \exp((\log x)^{0.95}) = x^{o(1)}$. There are two possibilities for n counted by $N(x, y)$:

(1) n is divisible by a shifted prime $p - 1 > z$,
(2) n is divisible by at least y shifted primes $p - 1 \leq z$.

By [9, Theorem 1.2], the count of the numbers in (1) is $\ll x/(\log z)^{\beta+o(1)}$, where $\beta := 1 - (1 + \log\log 2)/\log 2$ is the Erdős–Ford–Tenenbaum constant. Since $19\beta > 1$, the count in this case is $\ll x\log y/y$. For (2), let $\omega_z^*(n)$ denote the number of shifted primes $p-1 \leq z$ with $(p-1) \mid n$. It is easily seen that the average value of $\omega_z^*(n)$ for $n \leq x$ is $\log\log z + O(1)$. Thus, the count in this case is $\ll x\log\log z/y \ll x\log y/y$. Adding up the bounds for the counts in both cases yields the desired upper bound for $N(x, y)$. □

Now we study the k-level set $\mathcal{L}_k := \{n \in \mathbb{N} \colon \omega^*(n) = k\}$ for each $k \in \mathbb{N}$. It is clear that

$$N(x, y) = \sum_{k \geq y} \#(\mathcal{L}_k \cap [1, x]).$$

We shall show that each $\mathcal{L}_k$ has a positive natural density δ_k, which is defined by

$$\delta_k := \lim_{x\to\infty} \frac{\#(\mathcal{L}_k \cap [1, x])}{x}. \tag{13}$$

Theorem 2. *For every $k \in \mathbb{N}$, the k-level set $\mathcal{L}_k$ admits a positive natural density δ_k. Moreover, we have $\sum_k \delta_k = 1$.*

We first show that each $\mathcal{L}_k$ is nonempty.

Lemma 1. *For every $k \in \mathbb{N}$, we have $\mathcal{L}_k \neq \emptyset$.*

Proof. Note that $\mathcal{L}_1 = \mathbb{N} \setminus 2\mathbb{N}$ and $2 \in \mathcal{L}_2$. So we may assume that $k \geq 2$, so that $\mathcal{L}_k \subseteq 2\mathbb{N}$. We shall show that for any $n \in 2\mathbb{N}$, there exists an integral multiple $m \in \mathbb{N}$ of n such that $\omega^*(m) = \omega^*(n) + 1$. The lemma would then follow from this result in an inductive manner.

To prove this, we fix $n \in 2\mathbb{N}$ and consider

$$\mathcal{P}_2(x) := \{2 < p \le x \colon \Omega((p-1)/2) \le 2 \text{ and } P^-((p-1)/2) > x^{3/11}\}.$$

(The notation here is standard, signifying that $(p-1)/2$ is either prime or the product of two primes, and this prime or primes are $> x^{3/11}$.) By [7, Theorem 25.11], we have $\#\mathcal{P}_2(x) \gg x/(\log x)^2$ for all sufficiently large x. We wish to find some large $p \in \mathcal{P}_2(x)$ with $\omega^*(n(p-1)/2) = \omega^*(n)+1$. To this end, we shall show that the number of those $p \in \mathcal{P}_2(x)$ which do not have this property is $O(x \log\log x/(\log x)^3)$. Note that if $p \in \mathcal{P}_2(x)$ does not possess this property, then we can find $a \mid n$ and $b \mid (p-1)/2$ with $a, b > 1$ such that $ab+1$ is a prime not equal to p.

There are two possibilities: (i) $b = (p-1)/2$ and $ab+1$ is prime with $a > 2$ and (ii) $p-1 = 2qr$ with q, r primes in $(x^{3/11}, x^{8/11}/2)$ and $aq+1$ is prime.

Case (i) is simple. Fix $a \mid n$ with $a > 2$. The number of integers $b \le x$ with $P^-(b) > x^{3/11}$ and both $2b+1$ and $ab+1$ are prime is $\ll x/(\log x)^3$. (The implied constant depends on a but there is a bounded number of choices for a.)

Now we consider Case (ii). Again, let us fix $a \mid n$. For any prime $q \in (x^{3/11}, x^{8/11}/2)$, the number of primes $b < x/2q$ such that both $ab+1$ and $2qb+1$ are prime is

$$\ll \frac{x}{q(\log x)^3} \prod_{r \mid (2q-a)} \left(1 - \frac{1}{r}\right)^{-1} \ll \frac{\log\log q}{q} \cdot \frac{x}{(\log x)^3}.$$

Summing this bound over all $q \in (x^{3/11}, x^{8/11}/2)$ and $a \mid n$, we see that the number of choices of p with b in Case (ii) is $\ll x \log\log x/(\log x)^3$. This completes the proof. $\square$

We are now ready to prove Theorem 2.

Proof of Theorem 2. As we have pointed out, it suffices to demonstrate the existence and positivity of δ_k as defined by (13). The case $k = 1$ is obvious, since the level set $\mathcal{L}_1$ consists of the odd numbers, so that $\delta_1 = 1/2$. Now let us fix $k \ge 2$. Then $\mathcal{L}_k \subseteq 2\mathbb{N}$. We define an equivalence relation $\simeq$ on $\mathcal{L}_k$ by declaring that $m \simeq n$ if and only if m and n have exactly the same set of shifted prime divisors. Let $\mathcal{C}_k$ be the set of all equivalence classes[1] $\langle n \rangle$ of $\mathcal{L}_k$ under $\simeq$. Then

$$\mathcal{L}_k = \bigcup_{\langle n \rangle \in \mathcal{C}_k} \langle n \rangle. \tag{14}$$

[1]In [14], the class containing n is denoted $\mathcal{S}_n$.

It is known [6, Theorem 3] that each $\langle n\rangle$ has a positive natural density. Thus, if natural density were countably additive, then we would be able to conclude that δ_k exists and equals the sum of the natural densities of the sets $\langle n\rangle \in \mathcal{C}_k$. Since Lemma 1 implies that $\mathcal{C}_k \neq \emptyset$, we would also have $\delta_k > 0$. Unfortunately, $\#\mathcal{C}_k$ may be infinite and natural density is only finitely additive.

To overcome this issue we appeal to the following elementary result.

Lemma 2. *Let $\mathcal{A}_1, \mathcal{A}_2, \ldots$ be an infinite sequence of pairwise disjoint subsets of $\mathbb{N}$, such that each $\mathcal{A}_i$ has a natural density $\delta(\mathcal{A}_i)$. If the upper asymptotic density of $\bigcup_{i>j} \mathcal{A}_i \to 0$ as $j \to \infty$, then the density of $\bigcup_{i\ge 1} \mathcal{A}_i$ exists and*

$$\delta\Big(\bigcup_{i\ge 1} \mathcal{A}_i\Big) = \sum_{i\ge 1} \delta(\mathcal{A}_i).$$

This result can be applied to the sets in $\mathcal{C}_k$, since if there are infinitely many, then all but finitely many have their elements divisible by a shifted prime $p-1 > y$, for any fixed y. Appealing to [6, Theorem 2] (or to Theorem 1 above), the union of these sets has upper density tending to 0 as $y \to \infty$. Thus, to complete the proof, we now have

$$\delta_k = \sum_{\langle n\rangle \in \mathcal{C}_k} \delta(\langle n\rangle) \quad \text{and} \quad \sum_k \delta_k = 1.$$

□

Here are some exact counts of the level sets $\mathcal{L}_k$ for $k \le 11$.

TABLE 2. Exact counts of level sets for $k < 12$

k	10^4	10^6	10^8	10^{10}	$\approx \delta_k$
1	5,000	500,000	50,000,000	5,000,000,000	.5
2	834	77,696	7,436,825	720,726,912	.070
3	965	91,602	8,826,498	859,002,140	.084
4	877	79,986	7,691,971	748,412,490	.074
5	612	59,518	5,684,323	555,900,984	.055
6	456	40,641	4,031,009	401,146,301	.040
7	287	29,565	3,016,881	300,330,932	.030
8	202	23,190	2,324,769	233,611,502	.023
9	153	17,914	1,800,298	182,793,491	.018
10	159	13,899	1,401,307	144,740,573	.015
11	103	10,487	1,131,836	118,302,267	.012
≥ 12	352	55,682	6,654,283	735,032,408	

The largest values of k encountered here up to the various bounds: 10^4: 28, 10^6: 86, 10^8: 247, 10^{10}: 618.

Perhaps the densities δ_k are monotone for $k \geq 3$. Note that in [14] it is shown that the largest density of some $\langle n \rangle$ for n even is given for $n = 2$, which improves slightly on an earlier result of Sunseri. It would be good to have some sort of asymptotic inequalities for these densities, and a result in this direction is produced in the next section.

4. A lower bound for $\delta(\langle n \rangle)$

Let $n \in 2\mathbb{N}$ and consider the equivalence class $\langle n \rangle$ of $\mathbb{N}$ under the same relation $\simeq$ as introduced in the proof of Theorem 2 above. Suppose that $n = \min\langle n \rangle$. In other words, n is the least common multiple of all shifted prime divisors of n. We clearly have $\delta(\langle n \rangle) < 1/n$. Erdős and Wagstaff [6] asked what a positive lower bound could be for $\delta(\langle n \rangle)$. The following theorem provides such a lower bound.

Theorem 3. *Let $n \in 2\mathbb{N}$ be such that $n = \min\langle n \rangle$. Then*

$$\delta(\langle n \rangle) \geq \frac{1}{n^{O(\tau(n))}}.$$

Proof. We follow the proof of [6, Theorem 3] on the existence and positivity of $\delta(\langle n \rangle)$. For any $a_1, ..., a_r \in \mathbb{N}$, denote by $T_n(a_1, ..., a_r)$ the natural density of the set of multiples of n which are not divisible by any a_i for $1 \leq i \leq r$. Explicitly, we have

$$T_n(a_1, ..., a_r) = \sum_{j=0}^{r} (-1)^j \sum_{1 \leq i_1 < \cdots < i_j \leq r} \frac{1}{[n, a_{i_1}, ..., a_{i_j}]}.$$

By [6, Eq. (2), p. 110], we have

$$\frac{1}{n} T_n(a_1, ..., a_{r+s}) \geq T_n(a_1, ..., a_r) T_n(a_{r+1}, ..., a_{r+s})$$

for any integers $r, s \geq 0$ and any $a_1, ..., a_{r+s} \in \mathbb{N}$. From this inequality with $s = 1$ it follows immediately by induction that

$$T_n(a_1, ..., a_r) \geq \frac{1}{n} \prod_{i=1}^{r} \left(1 - \frac{n}{[n, a_i]}\right)^{1/m_i}, \tag{15}$$

where $m_i := \#\{1 \leq j \leq r \colon a_j = a_i\}$.

It suffices to prove the theorem for large values of n. Let $y \geq 1$ be a parameter depending on n. Let $\mathscr{A}_1 := \{[p-1, n] \colon p-1 \leq y \text{ and } (p-1) \nmid n\}$ and $\mathscr{A}_2 := \{[p-1, n] \colon p-1 > y \text{ and } (p-1) \nmid n\}$. Denote by $\mathscr{B}(\mathscr{A}_2)$ the set of multiples of elements of $\mathscr{A}_2$. We arrange the elements of $\mathscr{A}_1$

as a strictly increasing sequence $\{a_i\}_{i=1}^r$. The proof of [6, Theorem 3] shows that

$$\frac{1}{n}\delta(\langle n\rangle) \geq T_n(a_1, ..., a_r)\left(\frac{1}{n} - \delta(\mathscr{B}(\mathscr{A}_2))\right) > 0,$$

provided that y is sufficiently large in terms of n. To get a positive lower bound for $\delta(\langle n\rangle)$, it suffices to obtain a positive lower bound for $T_n(a_1, ..., a_r)$ and an upper bound $< 1/n$ for $\delta(\mathscr{B}(\mathscr{A}_2))$. Take $y = e^{n^{1/\beta}}$. By [9, Theorem 1.2] we have

$$\delta(\mathscr{B}(\mathscr{A}_2)) \ll \frac{1}{(\log y)^\beta\sqrt{\log\log y}} \ll \frac{1}{n\sqrt{\log n}}.$$

To handle $T_n(a_1, ..., a_r)$, we appeal to (15) to obtain

$$\begin{aligned}
T_n(a_1, ..., a_r) &\geq \frac{1}{n}\prod_{d|n}\prod_{\substack{p\leq y+1\\(p-1,n)=d\\(p-1)\nmid n}}\left(1 - \frac{d}{p-1}\right)\\
&\geq \frac{1}{n}\exp\left(-\sum_{d\,|\,n}\sum_{\substack{p\leq y+1\\(p-1,n)=d\\(p-1)\nmid n}}\frac{d}{p-1}\right)\\
&= \frac{e^{\omega^*(n)}}{n}\exp\left(-\sum_{d\,|\,n}\sum_{\substack{p\leq y+1\\(p-1,n)=d}}\frac{d}{p-1}\right).
\end{aligned}$$

If we replace $d/(p-1)$ with d/p, the error created in the double sum is $\ll \sigma(n)/n$, where σ is the sum-of-divisors function. Thus,

$$T_n(a_1, ..., a_r) \geq \frac{e^{\omega^*(n)}}{n}\exp\left(-\sum_{d\,|\,n}\sum_{\substack{p\leq y+1\\(p-1,n)=d}}\frac{d}{p} + O(\log\log n)\right). \tag{16}$$

Lemma 3. *For each number $A > 0$ there is a positive constant κ such that for all large x and $d < (\log x)^A$, we have*

$$\sum_{\substack{d<p\leq x\\p\equiv a \pmod d}}\frac{1}{p} = \frac{\log\log x}{\varphi(d)} + E(d) + O(\exp(-\kappa(\log x)^{1/2})),$$

for all a coprime to d. The number $E(d)$ satisfies $|E(d)| \ll \log(2d)/\varphi(d)$.

This follows from the Siegel–Walfisz theorem, where the estimation for $E(d)$ appears in works of Norton and Pomerance, see [11, Lemma 2.1].

Consider the double sum in (16). It follows that

$$\begin{aligned}\sum_{d|n} d \sum_{\substack{p\le y+1\\(p-1,n)=d}} \frac{1}{p} &= \sum_{cd|n} \mu(c) d \sum_{\substack{p\le y+1\\ p\equiv 1 \pmod{cd}}} \frac{1}{p}\\
&= \sum_{cd|n} \mu(c) d \left(\frac{\log\log y}{\varphi(cd)} + E(cd) + O(\exp(-\kappa(\log y)^{1/2})) \right)\\
&= \tau(n)\log\log y + \sum_{m\,|\,n} \varphi(m)E(m) + O(n^2 \exp(-\kappa(\log y)^{1/2}))\\
&\le \tau(n)(\log\log y + O(\log n)) = O(\tau(n)\log n).\end{aligned}$$

Hence, we have

$$T_n(a_1, ..., a_r) \ge \frac{e^{\omega^*(n)}}{n^{O(\tau(n))}} = \frac{1}{n^{O(\tau(n))}},$$

the last estimate coming from $\omega^*(n) \le \tau(n)$. Combining the above estimate with that for $\delta(\mathscr{B}(\mathscr{A}_2))$ completes the proof. $\square$

5. Higher moments of $\omega^*(n)$

For every $k \in \mathbb{N}$, we define the kth moment of $\omega^*(n)$ by

$$M_k(x) := \frac{1}{x} \sum_{n \le x} \omega^*(n)^k.$$

Then we have

$$M_k(x) = \frac{1}{x} \sum_{[p_1-1,...,p_k-1]\le x} \left\lfloor \frac{x}{[p_1-1, ..., p_k-1]} \right\rfloor. \tag{17}$$

This shows that $M_k(x)$ is intimately related to

$$S_k(x) := \frac{1}{x} \sum_{[p_1-1,...,p_k-1]\le x} 1.$$

In fact, if $S_k(x) \ll (\log x)^{c_k}$, then a partial summation argument applied to the upper bound in (17) afforded by removing the floor function shows that $M_k(x) \ll (\log x)^{c_k+1}$. A similar argument shows that a lower bound for $S_k(x)$ implies one for $M_k(x)$.

For $k \ge 2$, it is natural to relate the function $\omega^*(n)^k$ to $\tau(n)^k$. It is well-known that

$$\frac{1}{x} \sum_{n \le x} \tau(n)^k \sim \frac{1}{(2^k-1)!} \prod_p \left(1 - \frac{1}{p}\right)^{2^k} \sum_{\nu \ge 0} \frac{(\nu+1)^k}{p^\nu} (\log x)^{2^k-1}$$

for every $k \geq 1$. Comparing ω^* with τ and taking the primality conditions into account, one may conjecture that $M_k(x) \sim \mu_k(\log x)^{2^k-k-1}$ for every $k \geq 2$, where $\mu_k > 0$ is a constant depending on k. Similarly, one may also conjecture that $S_k(x) \sim (2^k - k - 1)\mu_k(\log x)^{2^k-k-2}$ for every $k \geq 2$ with the same constant μ_k. As in the case $k = 2$, we have the upper and lower bounds for $M_3(x)$ of the conjectured magnitude.

Theorem 4. *We have $M_3(x) \asymp (\log x)^4$ for all $x \geq 2$.*

The upper and lower bounds will be proved by using different types of arguments. The rest of this section will be devoted to proving the upper bound $M_3(x) \ll (\log x)^4$, with the proof of the lower bound $M_3(x) \gg (\log x)^4$ given in Section 6.

We begin with some lemmas. The first is a variant of [11, Lemma 2.7].

Lemma 4. *Uniformly for coprime integers e, f in $[1, x]$,*

$$\sum_{\substack{a,b \leq x \\ (ae,bf)=1 \\ ae \neq bf}} \frac{1}{ab} \prod_{p \mid ab(ae-bf)} \left(1 + \frac{1}{p}\right) \ll (\log x)^2. \tag{18}$$

Proof. First note that the product contributes at most a factor of magnitude $\log\log x$ to the sum, so the result holds trivially if either a or b is bounded by $x^{1/\log\log x}$. Hence, we may assume that $a, b > x^{1/\log\log x}$. Further, every integer $n \leq x$ has $< \log x$ prime divisors, so that

$$\prod_{\substack{p \mid n \\ p > (\log x)^{1/2}}} \left(1 + \frac{1}{p}\right) \ll 1,$$

uniformly. Let u be the product of all primes $p \leq (\log x)^{1/2}$. Thus, we may restrict the primes p in the product in the lemma to those that also divide u. We have the expression in (18) is

$$\ll \sum_{\substack{x^{1/\log\log x} < a,b \leq x \\ (ae,bf)=1 \\ ae \neq bf}} \frac{1}{ab} \sum_{\substack{j \mid u \\ j \mid ab(ae-bf)}} \frac{1}{j} \leq \sum_{j \mid u} \frac{1}{j} \sum_{\substack{j < a,b \leq x \\ j \mid ae-bf \\ (ae,bf)=1 \\ ae \neq bf}} \frac{1}{ab}. \tag{19}$$

(Note that we assume here that $a, b > j$, since they are $> x^{1/\log\log x}$ and $j \leq u \leq \exp((1+o(1))(\log x)^{1/2})$.) For $p \mid j$ with $j \mid ab(ae-bf)$, we have either $a \equiv 0 \pmod{p}$, $b \equiv 0 \pmod{p}$, or $a \equiv bfe^{-1} \pmod{p}$ (if $p \mid e$ then $p \nmid ae-bf$). Since j is squarefree, there are at most $3^{\omega(j)}j$ pairs $a, b \pmod{j}$ with $j \mid ab(ae - bf)$. For a fixed pair of residues $\pmod{j}$ that we have here, the sum of $1/ab$ in this class is $\ll (\log x)^2/j^2$ uniformly, so the total contribution in the last sum in (19) is $\ll 3^{\omega(j)}(\log x)^2/j$. We thus have the last double sum in (19) is

$$\ll \sum_{j|u} \frac{3^{\omega(j)}(\log x)^2}{j^2} = (\log x)^2 \prod_{p\le(\log x)^{1/2}} \left(1+\frac{3}{p^2}\right) \ll (\log x)^2,$$

which completes the proof of the lemma. $\square$

Lemma 5. *Uniformly for* $1 \le u < x$ *we have*

$$\sum_{\substack{q\le x\\ q\equiv 1 \,(\mathrm{mod}\ u)}} \frac{\tau((q-1)/u)}{\varphi((q-1)/u)} \ll \frac{u}{\varphi(u)} \log x.$$

Proof. The result holds trivially when $x \le 2$ or $u \ge x/2$, so assume that $x > 2$ and $u < x/2$. (In fact, the lemma follows from a trivial argument if $u > x/\exp((\log x)^{1/2})$, but we won't use this.) We first consider

$$T = \sum_{\substack{q\le x\\ q\equiv 1 \,(\mathrm{mod}\ u)}} \frac{\tau((q-1)/u)(q-1)/u}{\varphi((q-1)/u)}.$$

Using that

$$\frac{n}{\varphi(n)} = \sum_{d|n} \frac{\mu^2(d)}{\varphi(d)},$$

we have

$$T \le \sum_{d<x} \frac{1}{\varphi(d)} \sum_{\substack{q\le x\\ q\equiv 1 \,(\mathrm{mod}\ du)}} \tau((q-1)/u).$$

Using the maximal order of the divisor function and that q is an integer that is 1 (mod ud) and $> ud$, the contribution to T from a particular number d is $\ll (x/du)(x/u)^{\epsilon}$, so the contribution to T from numbers $d > (x/u)^{1/4}$ is $\ll (x/u)^{3/4+\epsilon} < (x/u)^{4/5}$, say. We also use that $\tau((q-1)/u)$ is at most twice the number of divisors $j \mid (q-1)/u$ with $j \le (x/u)^{1/2}$. Thus,

$$T \ll \sum_{d\le(x/u)^{1/4}} \frac{1}{\varphi(d)} \sum_{j\le(x/u)^{1/2}} \sum_{\substack{q\le x\\ q\equiv 1 \,(\mathrm{mod}\ [j,d]u)}} 1 + (x/u)^{4/5}.$$

We have $[j,d] \le (x/u)^{3/4}$ and so $x/([j,d]u) \ge (x/u)^{1/4}$ and the inner sum here is $\ll x/(\varphi([j,d]u)\log(x/u))$. Now $[j,d] = jd/i$, where $i = (j,d)$, so

$$T \ll \sum_{d\le(x/u)^{1/4}} \frac{1}{\varphi(d)} \sum_{i|d} \sum_{k\le(x/u)^{1/2}/i} \frac{x}{\varphi(d)\varphi(u)\varphi(k)\log(x/u)} + (x/u)^{4/5}.$$

The sum of $1/\varphi(k)$ in the indicated range is $\ll \log(x/u)$, so

$$T \ll \sum_{d \le (x/u)^{1/4}} \frac{x\tau(d)}{\varphi(u)\varphi(d)^2} + (x/u)^{4/5} \ll \frac{x}{\varphi(u)}.$$

It immediately follows that

$$\sum_{\substack{q \le x \\ q \equiv 1 \,(\mathrm{mod}\ u)}} \frac{\tau((q-1)/u)(q-1)}{\varphi((q-1)/u)} \ll \frac{ux}{\varphi(u)}$$

and the lemma follows by partial summation. □

At this point we find it convenient to reprise the upper bound proof for $k = 2$ from [11] since we take a slightly different perspective, the proof is short, and the case $k = 3$ follows with similar tools. We show that

$$S_2(x) \ll 1. \tag{20}$$

We are to count pairs of primes p, q with $[p-1, q-1] \le x$. Let

$$d = (p-1, q-1),\ \ p-1 = ad,\ \ q-1 = bd.$$

The case $p = q$ has the count $O(x/\log x)$, so we may assume that $a \ne b$. So, we are counting triples a, b, d with $(a, b) = 1$, $a \ne b$, $abd \le x$, with $ad+1, bd+1$ both prime. First suppose that $d = \max\{a, b, d\}$. Since $abd \le x$, we have $ab \le x^{2/3}$. For a given choice of a, b, the number of choices for $d \le x/ab$ with $ad+1, bd+1$ both prime is

$$\ll \frac{x}{ab(\log x)^2} \prod_{\ell \,|\, ab(a-b)} \left(1 + \frac{1}{\ell}\right),$$

where ℓ runs over primes. This follows from the upper bound in either Brun's or Selberg's sieve. Lemma 4 in the case $e = f = 1$ completes the proof of (20) in this case.

Now assume that $a = \max\{a, b, d\}$. Then $bd = q - 1 \le x^{2/3}$. For a given prime $q \le x^{2/3} + 1$ and a divisor d of $q - 1$, we count values of $a \le x/(q-1)$ with $ad + 1$ prime. By the Brun–Titchmarsh inequality, the number of such values of a is

$$\ll \frac{d}{\varphi(d)} \frac{x}{(q-1)\log x} \le \frac{x}{\varphi(q-1)\log x}.$$

So, in all, there are $\ll \tau(q-1)x/(\varphi(q-1)\log x)$ choices for a. Lemma 5 in the case $u = 1$ completes the proof of (20) in this case. The last case $b = \max\{a, b, d\}$ is completely symmetric with the case just considered, so we are done.

Now we prove the upper bound $M_3(x) \ll (\log x)^4$ asserted in Theorem 4. For this it is sufficient to prove that

$$S_3(x) = \frac{1}{x} \sum_{[p-1,\, q-1,\, r-1] \le x} 1 \ll (\log x)^3, \tag{21}$$

where p, q, r run over prime numbers. From the case of the second moment, we may assume that p, q, r are distinct. Note that

$$[p-1, q-1, r-1] = abcdefg,$$

where

$$g = \gcd(p-1, q-1, r-1),$$
$$dg = \gcd(p-1, q-1),\ eg = \gcd(p-1, r-1),\ fg = \gcd(q-1, r-1),$$
$$a = (p-1)/deg,\ b = (q-1)/dfg,\ c = (r-1)/efg.$$

Note that we have a, b, c pairwise coprime, as well as d, e, f. Also,

$$\gcd(ae, bf) = 1,\ \gcd(ad, ce) = 1,\ \gcd(bd, cf) = 1.$$

To prove (21), we consider 7 cases depending on the largest of $a, \ldots, g$. By symmetry this collapses to 3 cases:

$$\max\{a, \ldots, g\} = a, d, \text{ or } g.$$

Beginning with the max being a, first choose a prime r and a factorization of $r-1$ as $cefg$. Next choose a prime q with $q \equiv 1 \pmod{fg}$ and take a factorization of $(q-1)/fg$ as bd. Finally, let $a \le x/bcdefg$ with $adeg + 1$ prime. The number of choices for a is

$$\ll \frac{x}{bcdefg \log x} \frac{deg}{\varphi(deg)}.$$

The number of choices for c, e, f, g is $\tau_4(r-1)$. Given an ordered factorization $cefg$ of $r-1$, let $u = u_{r-1} = fg$. The number of choices for b, d is $\tau((q-1)/u)$. Thus, the total number of choices in this case is

$$\ll \sum_{r<x} \frac{\tau_4(r-1)}{\varphi(r-1)} \sum_{\substack{q<x \\ q \equiv 1 \pmod{u_{r-1}}}} \frac{\tau((q-1)/u)}{\varphi((q-1)/u)} \frac{x}{\log x}.$$

Using Lemma 5, we have the number of choices

$$\ll x \sum_{r<x} \frac{\tau_4(r-1)}{\varphi(r-1)} \frac{u}{\varphi(u)} \le x \sum_{r<x} \frac{\tau_4(r-1)(r-1)}{\varphi(r-1)^2}. \tag{22}$$

We now appeal to [12, Theorem 1.2] or [13, Corollary 1.2] from which we see this last sum is $\ll (\log x)^3$. This completes the proof when $a = \max\{a, \ldots, g\}$.

Now assume that $d = \max\{a, \dots, g\}$. We choose a prime $r < x$ and a factorization $cefg$ of $r-1$. We then choose a, b with $ab(r-1) \le x^{6/7}$. We now let d run up to $x/(ab(r-1)$ with $adeg+1$ and $bdfg+1$ prime. The number of choices is

$$\ll \frac{x\tau_4(r-1)}{ab(r-1)(\log x)^2} \prod_{\ell \mid abef(bf-ae)} \Big(1+\frac{1}{\ell}\Big) \prod_{\ell \mid g} \Big(1+\frac{1}{\ell}\Big),$$

where ℓ runs over prime numbers. We can absorb the part of the product coming from $\ell \mid ef$ and $\ell \mid g$ into the main term, getting

$$\frac{x\tau_4(r-1)(r-1)}{\varphi(r-1)^2 ab(\log x)^2} \prod_{\ell \mid ab(bf-ae)} \Big(1+\frac{1}{\ell}\Big).$$

Note that this final product is finite, since $(ae, bf) = 1$ and $ae \ne bf$ (If $a = b = e = f = 1$, then one has $p = q$, a possibility we ruled out). Lemma 4 and then the argument as in (22) completes the proof of the case when d is the maximum of $a, \dots, g$.

We now consider the case that $g = \max\{a, \dots, g\}$, which is quite similar to the previous case. For a given choice of $a, \dots, f$, we have $a \dots f \le x^{6/7}$, so the number of values of $g \le x/a \dots f$ with $adeg+1, bdfg+1, cefg+1$ all prime is

$$\ll \frac{x}{A(\log x)^3} \prod_{\ell \mid AE} \Big(1+\frac{2}{\ell}\Big),$$

where

$$A = abcdef, \quad E = (ae-bf)(ad-cf)(bd-ce).$$

Without the product, the sum of $1/A$ is $O((\log x)^6)$. We would like to show the same estimate holds with the product included. Note however that the product is in the worst case $O((\log\log x)^2)$, so our result holds trivially if any of $a, \dots, f$ is $\le x^{1/(\log\log x)^2}$. We thus assume they are all $> x^{1/(\log\log x)^2}$. Further, as in the proof of Lemma 4, let u be the product of all primes $\ell \le (\log x)^{1/2}$. We may restrict primes ℓ in the product to such primes. We wish to estimate

$$\sum_{j \mid u} \frac{2^{\omega(j)}}{j} \sum_{\substack{a,\dots,f<x \\ j \mid AE}} \frac{1}{a \dots f}.$$

Note that AE is the product of 9 expressions, so that in the inner sum, the 6-tuple $(a, \dots, f)$ lies in $\le 9^{\omega(j)} j^5$ residue classes mod j. For each one of these classes the inner sum is $\ll (\log x)^6/j^6$, so we have

$$\sum_{j|u} \frac{2^{\omega(j)}}{j}\frac{9^{\omega(j)}}{j}(\log x)^6 = (\log x)^6 \sum_{j|u} \frac{18^{\omega(j)}}{j^2} \ll (\log x)^6.$$

This completes our proof of the upper bound in Theorem 4.

6. A lower bound for the third moment

By (17), we have

$$M_3(x) \geq \frac{1}{2} \sum_{[p-1,q-1,r-1]\leq x} \frac{1}{[p-1,q-1,r-1]}.$$

Thus, we wish to show that

$$\sum_{[p-1,q-1,r-1]\leq x} \frac{1}{[p-1,q-1,r-1]} \gg (\log x)^4. \tag{23}$$

We restrict to the case that p, q, r are distinct primes, noting that the complementary case is negligible. We use the identity

$$\frac{1}{[p-1,q-1,r-1]} = \sum_{\substack{u\,|\,r-1\\ u\,|\,[p-1,q-1]}} \frac{\varphi(u)}{[p-1,q-1](r-1)}.$$

Let

$$M_2(x;u) := \sum_{\substack{[p-1,q-1]\leq x\\ u\,|\,[p-1,q-1]}} \frac{1}{[p-1,q-1]}. \tag{24}$$

We thus have that

$$M_3(x) \geq \frac{1}{2} \sum_{u\leq x^{1/3}} \sum_{\substack{r\leq x^{1/3}\\ u\,|\,r-1}} \frac{1}{r-1} M_2(x^{2/3};u). \tag{25}$$

Our goal then is to obtain a lower bound for $M_2(x^{2/3},u)$ and use that in (25).

Helpful will be a tool from [2], namely Theorem 2.1: *For each $\varepsilon > 0$ there are numbers $\delta > 0$ and x_0, such that if $x > x_0$, $k < x^\delta$, and $(a,k) = 1$, then*

$$\left| \sum_{\substack{p\leq y\\ p\equiv a \pmod k}} \log p - \frac{x}{\varphi(k)} \right| \leq \varepsilon \frac{y}{\varphi(k)}, \tag{26}$$

for all $y \geq x$, except possibly for those k divisible by a certain number $k_0(x) > \log x$.

If $k_0(x)$ should exist and is divisible by a prime $> (1/3)\log\log x$, let $s = s(x)$ be the largest such prime. Otherwise, let $s = s(x)$ be the least prime $> (1/3)\log\log x$. Note that if x is sufficiently large and $k_0(x)$ exists and is not divisible by any prime $> (1/3)\log\log x$, then $k_0(x)$ must be divisible by the cube of some prime. Indeed

$$\prod_{\ell \le (1/3)\log\log x} \ell^2 = (\log x)^{2/3+o(1)},$$

so this product is smaller than $k_0(x)$. In particular, if x is large and k is cube-free, then (26) holds whenever $s \nmid k$.

Corollary 1. *Suppose that $\varepsilon = 1/4$ and we have the corresponding number δ as above. If x is sufficiently large, $k < x^\delta$ is cube-free and not divisible by $s(x)$, then*

$$\sum_{\substack{x<p\le x^2 \\ p \equiv a \pmod k}} \frac{1}{p} \gg \frac{1}{\varphi(k)},$$

uniformly.

This follows instantly from the above theorem, namely [2, Theorem 2.1], by either partial summation or a dyadic summation.

Let $\varphi_2(n)$ be the multiplicative function with value at a prime power ℓ^j equal to $\ell^j(1-2/\ell)$.

Proposition 1. *Let δ be the corresponding constant for $\varepsilon = 1/4$. Suppose that x is large and $u < x^{\delta/12}$ is squarefree and not divisible by $s = s(x^{1/6})$. Then*

$$M_2(x^{2/3}; u) \gg \log x \sum_{\substack{u=u_1u_2u_3 \\ u_1u_2 \text{ odd}}} \frac{u_3\varphi_2(u_1u_2)}{\varphi(u)^2},$$

uniformly.

Proof. We have

$$M_2(x^{2/3}; u) \ge \sum_{\substack{u=u_1u_2u_3 \\ u_1u_2 \text{ odd}}} \sum_{\substack{p \le x^{1/3} \\ p \equiv 1 \pmod{u_1u_3} \\ (p-1,u_2s)=1}} \sum_{\substack{q \le x^{1/3} \\ q \equiv 1 \pmod{u_2u_3} \\ (q-1,u_1s)=1}} \frac{1}{[p-1,q-1]}. \tag{27}$$

We write $1/[p-1,q-1]$ as $(p-1,q-1)/(p-1)(q-1)$ and note that $u_3 \mid (p-1,q-1)$. Thus,

$$1/[p-1,q-1] = u_3 \sum_{d \mid (p-1,q-1)/u_3} \varphi(d)/(p-1)(q-1).$$

Thus, for a given choice of u_1, u_2, u_3, the contribution in (27) is at least

$$\sum_{\substack{d<x^{1/12}\\ d\,\text{squarefree}\\ (d,s)=1}} u_3\varphi(d) \sum_{\substack{p\le x^{1/3}\\ p\equiv 1 \pmod{du_1u_3}\\ (p-1,u_2s)=1}} \frac{1}{p-1} \sum_{\substack{q\le x^{1/3}\\ q\equiv 1 \pmod{du_2u_3}\\ (q-1,u_1s)=1}} \frac{1}{q-1}. \tag{28}$$

For the sum over p we temporarily ignore the condition that $s \nmid p-1$. Then p runs over $\varphi_2(u_2)$ residue classes mod du. In each of these classes, the sum of $1/(p-1)$ is $\gg 1/\varphi(du)$ by Corollary 1. So, the sum over p appears to be $\gg \varphi_2(u_2)/\varphi(du)$. But, we also need to take into account the condition $s \nmid p-1$. For this, we compute an upper bound for the sum where $s \mid p-1$. An upper bound sieve result shows that the contribution is $\ll \varphi_2(u_2)/(\varphi(du)\log\log x)$, which justifies ignoring the condition $s \nmid p-1$. For the sum over q, the analogous argument shows that it is $\gg \varphi_2(u_1)/\varphi(du)$.

Thus, the expression in (28) is at least of magnitude

$$\sum_{\substack{d<x^{1/12}\\ d\,\text{squarefree}\\ (d,s)=1}} \frac{u_3\varphi(d)\varphi_2(u_1u_2)}{\varphi(du)^2} \ge \sum_{\substack{d<x^{1/12}\\ d\,\text{squarefree}\\ (d,s)=1}} \frac{u_3\varphi_2(u_1u_2)}{d\varphi(u)^2} \gg \frac{u_3\varphi_2(u_1u_2)}{\varphi(u)^2}\log x.$$

Thus, the proposition now follows from (27). $\square$

We are now ready to complete the proof of the lower bound in Theorem 4, that is, $M_3(x) \gg (\log x)^4$. From (25) and Proposition 1 we have

$$M_3(x) \gg \sum_{\substack{u\le x^{\delta/12}\\ u\,\text{squarefree}\\ s\nmid u}} \varphi(u) \sum_{\substack{r\le x^{1/3}\\ u\mid r-1}} \frac{1}{r-1} \sum_{\substack{u=u_1u_2u_3\\ u_1u_2\ \text{odd}}} \frac{u_3\varphi_2(u_1u_2)}{\varphi(u)^2}\log x.$$

It thus follows from Corollary 1 that

$$M_3(x) \gg \sum_{\substack{u\le x^{\delta/12}\\ u\,\text{squarefree}\\ s\nmid u}} \sum_{\substack{u=u_1u_2u_3\\ u_1u_2\ \text{odd}}} \frac{u_3\varphi_2(u_1u_2)}{\varphi(u)^2}\log x.$$

We factor the u-expression as

$$\frac{u_3\varphi_2(u_1u_2)}{\varphi(u)^2} = \frac{\varphi_2(u_1)}{\varphi(u_1)^2}\frac{\varphi_2(u_2)}{\varphi(u_2)^2}\frac{u_3}{\varphi(u_3)^2}.$$

Note that for n odd we have $\varphi_2(n)/\varphi(n)^2 \gg 1/n$, so that

$$M_3(x) \gg \log x \Bigg(\sum_{\substack{u_1 \le x^{\delta/36} \\ u_1 \text{ odd, squarefree} \\ s \nmid u_1}} \frac{1}{u_1} \Bigg)^2 \sum_{\substack{u_3 \le x^{\delta/36} \\ u_3 \text{ squarefree} \\ s \nmid u_3}} \frac{1}{u_3} \gg (\log x)^4.$$

This completes the proof of (23).

7. The lower bound for the second moment

As mentioned, the proof in Ding [3] that $M_2(x) \gg \log x$ is not complete since it relies on an incorrect statement from [11]. The proof is easily correctable, and we give the few details here.

The goal is to show that

$$\sum_{[p-1,q-1] \le x} \frac{1}{[p-1,q-1]} \gg \log x. \tag{29}$$

We may assume here that $p \ne q$. As before,

$$\frac{1}{[p-1,q-1]} = \sum_{d \,|\, (p-1,q-1)} \frac{\varphi(d)}{(p-1)(q-1)},$$

so that

$$\sum_{[p-1,q-1] \le x} \frac{1}{[p-1,q-1]} = \sum_{d \le x} \varphi(d) \sum_{\substack{[p-1,q-1] \le x \\ d \,|\, (p-1,q-1)}} \frac{1}{(p-1)(q-1)}.$$

By placing additional restrictions on d, p, q the expression here only gets smaller. We do this as follows. Consider Corollary 1 from the previous section with $\varepsilon = 1/4$. We assume that d is squarefree, $d \le x^{\delta/4}$, and that $s(x^{1/4}) \nmid d$. We further assume that $p, q \in (x^{1/4}, x^{1/2}]$. So, Corollary 1 implies that $\sum_p 1/(p-1) \gg 1/\varphi(d)$, and the same for the sum over q. Thus,

$$\sum_{[p-1,q-1] \le x} \frac{1}{[p-1,q-1]} \gg \sum_d \frac{1}{\varphi(d)} \ge \sum_d \frac{1}{d} \gg \log x.$$

This completes the proof of (29).

A similar proof can show that $S_2(x) \gg 1$. Note that the claim that $S_2(x) \asymp 1$ was asserted without proof in [8]. Concerning $S_3(x)$, we have a proof that it is $\gg (\log x)^3$ (and so $S_3(x) \asymp (\log x)^3$ after the result in Section 5), but we do not present the details here.

8. A tail estimate

In this section we prove the following theorem.

Theorem 5. *We have*

$$\sum_{\substack{p,\, q\le x\\ [p-1,q-1]>x}} \frac{1}{[p-1,q-1]} \gg \log x.$$

In [11] it is claimed that

$$\sum_{p,q\le x} \frac{1}{[p-1,q-1]} = \sum_{[p-1,q-1]\le x} \frac{1}{[p-1,q-1]} + O(1), \tag{30}$$

see the discussion in [11] at the start of Section 4. However, the difference between the two sums in (30) is the sum in Theorem 5, so it cannot be $O(1)$.

Proof. We use the full strength of [2, Theorem 2.1] instead of the simplified version used in Section 6. Let $\mathcal{D} = \mathcal{D}(x) = \{d \le x^{1/20} : d \text{ even}, \gcd(15, d) = 1\}$ with x being sufficiently large. Let $\varepsilon = \delta = .01$, and let $\mathcal{D}_{\varepsilon,\delta} = \mathcal{D}_{\varepsilon,\delta}(x)$ be the possible set of exceptional moduli as described in [2, Theorem 2.1]. The set $\mathcal{D}_{\varepsilon,\delta}$ has cardinality $O_{\varepsilon,\delta}(1)$, and the members are all $> \log x$. Let $\mathcal{D}' = \mathcal{D}'(x)$ denote the subset of $\mathcal{D}$ of elements d with $30d$ not divisible by any member of $\mathcal{D}_{\varepsilon,\delta}(x)$.

For each $d \in \mathcal{D}'$ let $\mathcal{P} = \mathcal{P}(x, d)$ denote the set of primes p with

- $p \equiv 1 \pmod d$,
- $p \le x$,
- $\gcd(30, (p-1)/d) = 1$.

Since $\varphi(30d)/\varphi(d) = 16$, it follows from the conditions above that $\mathcal{P}$ consists of primes $p \le x$ in precisely 3 of the $16\varphi(d)$ reduced residue classes modulo $30d$. Indeed, if $2^a \,\|\, d$, with $a \ge 1$, then $p \equiv 2^a + 1 \pmod{2^{a+1}}$. Also, $p \equiv 2 \pmod 3$ and $p \equiv 2, 3, \text{or } 4 \pmod 5$.

Note that via [2, Theorem 2.1] and partial summation, if $x^{9/10} < t \le x$ with x sufficiently large, then

$$\left| \sum_{\substack{p\le t\\ p\in\mathcal{P}}} 1 - \frac{3t}{16\varphi(d)\log t} \right| \le \frac{6\varepsilon t}{16\varphi(d)\log t}. \tag{31}$$

With r running over primes, let

$$f(n) = f(n, x) = \sum_{\substack{7\le r\le x^{1/20}\\ r\,|\,n}} \frac{1}{r-1}.$$

Note that

$$\sum_{\substack{p\in\mathcal{P}\\ p\le t}} f((p-1)/d) = \sum_{7\le r\le x^{1/20}} \sum_{\substack{p\in\mathcal{P},\, p\le t\\ r\,|\,(p-1)/d}} \frac{1}{r-1}$$
$$< \frac{2t}{(16/3)\varphi(d)\log(t^{8/9}/30)} \sum_{r\ge 7} \frac{1}{(r-1)^2},$$

using the explicit version of the Brun–Titchmarsh inequality due to Montgomery–Vaughan [10, Theorem 2]. Since the final sum here is a constant smaller than .063, it follows from (31) that

$$\sum_{\substack{p\in\mathcal{P}\\ p\le t}} f((p-1)/d) \le \frac{3}{20} \sum_{\substack{p\in\mathcal{P}\\ p\le t}} 1 \tag{32}$$

for $x^{9/10} < t \le x$ and x sufficiently large. Let

$$\mathcal{P}' = \mathcal{P}'(x,d) = \{p \in \mathcal{P} : f((p-1)/d) \le 1/5\},$$

so that from (32) we see that

$$\sum_{\substack{p\in\mathcal{P}'\\ p\le t}} 1 \ge \frac{1}{4} \sum_{\substack{p\in\mathcal{P}\\ p\le t}} 1$$

for x sufficiently large and $x^{9/10} < t \le x$. Combining this with (31) and applying partial summation we obtain

$$\sum_{\substack{p\in\mathcal{P}'\\ x^{9/10}<p\le x}} \frac{1}{p} > \frac{.0048}{\varphi(d)} \ge \frac{.0096}{d} \tag{33}$$

for all $d \in \mathcal{D}'$ and x beyond some uniform bound.

For each $d \in \mathcal{D}'(x)$ let

$$\mathcal{Q} = \mathcal{Q}(x,d) = \{q \le x : q \equiv 1 \pmod{d}\},$$

so that for $x^{9/10} < t \le x$, we have

$$\left| \sum_{\substack{q\in\mathcal{Q}\\ q\le t}} 1 - \frac{t}{\varphi(d)\log t} \right| = \left| \pi(t;d,1) - \frac{t}{\varphi(d)\log t} \right| \le \frac{2\varepsilon t}{\varphi(d)\log t} \tag{34}$$

for x sufficiently large.

Next, for $d \in \mathcal{D}'(x)$ and $p \in \mathcal{P}'(x,d)$, let

$$\mathcal{Q}' = \mathcal{Q}'(x,d,p) = \{q \in \mathcal{Q} : \gcd(q-1,p-1) = d\}.$$

If $\gcd(q-1,p-1)>d$, then $rd \mid q-1$ for some prime $r \mid (p-1)/d$ with $r\ge 7$ (since $(p-1)/d$ is coprime to 30). For $x^{9/10}<t\le x$ we have (using $d\le x^{1/20}$ and $\pi(t;rd,1)\le t/rd$),

$$\begin{aligned}\sum_{r\,|\,(p-1)/d}\pi(t;rd,1) &= \sum_{\substack{r\,|\,(p-1)/d\\ r\le x^{1/20}}}\pi(t;rd,1)+\sum_{\substack{r\,|\,p-1\\ r>x^{1/20}}}\pi(t;rd,1)\\ &\le \sum_{\substack{r\,|\,(p-1)/d\\ r\le x^{1/20}}}\frac{2t}{\varphi(d)(r-1)\log(t/rd)}+\sum_{\substack{r\,|\,p-1\\ r>x^{1/20}}}\frac{t}{rd}\\ &\le \frac{9}{4}f((p-1)/d)\frac{t}{\varphi(d)\log t}+O\Big(\frac{t}{dx^{1/20}}\Big).\end{aligned}$$

Since $f((p-1)/d)\le 1/5$, we conclude that

$$\sum_{\substack{q\le t\\ q\in\mathcal{Q}\setminus\mathcal{Q}'}}1\le\frac{.46t}{\varphi(d)\log t}$$

for x sufficiently large and $x^{9/10}<t\le x$. Thus, from (34),

$$\sum_{\substack{q\le t\\ q\in\mathcal{Q}'}}1\ge\frac{t}{2\varphi(d)\log t},$$

so that

$$\sum_{\substack{q\in\mathcal{Q}'\\ x^{9/10}<q\le x}}\frac{1}{q}>\frac{.0525}{\varphi(d)}\ge\frac{.105}{d}. \tag{35}$$

Now for each pair p,q with $p\in\mathcal{P}'(x,d)$ and $q\in\mathcal{Q}'(x,d,p)$ with $p,q>x^{9/10}$, we have $[p-1,q-1]=(p-1)(q-1)/d>x^{1.75}$. Further, from (33) and (35),

$$\sum_{p,q}\frac{1}{[p-1,q-1]}=d\sum_{p}\frac{1}{p-1}\sum_{q}\frac{1}{q-1}>\frac{.001d}{d^2}=\frac{.001}{d}.$$

It remains to note that $\sum_{d\in\mathcal{D}'}1/d\gg\log x$. In fact, since every member of $\mathcal{D}_{\varepsilon,\delta}(x)$ exceeds $\log x$, we have

$$\sum_{a\in\mathcal{D}_{\varepsilon,\delta}(x)}\sum_{\substack{d\in\mathcal{D}\\ a|30d}}\frac{1}{d}=O_{\varepsilon,\delta}(1),$$

so that

$$\sum_{\substack{p,\,q\le x\\ [p-1,q-1]>x}} \frac{1}{[p-1,q-1]} > .001 \sum_{d\in\mathcal{D}} \frac{1}{d} + O(1) = \frac{1}{75000}\log x + O(1)$$

for all sufficiently large x. This completes the proof. $\square$

Acknowledgments

We thank Nathan McNew and Paul Pollack for some helpful comments.

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] W. R. Alford, A. Granville, and C. Pomerance, There are infinitely many Carmichael numbers, Ann. of Math. (2) **139** (1994), 703–722.

[3] Y. Ding, On a conjecture of M. R. Murty and V. K. Murty, Canad. Math. Bull. **66** (2023), 679–681.

[4] Y. Ding, On a conjecture of M. R. Murty and V. K. Murty II, arXiv:2209.01087.

[5] P. Erdős, Über die Anzahl der Lösungen von $[p-1,q-1]\le x$, (Aus einen Brief von P. Erdős an K. Prachar), Monatsh. Math. **59** (1955), 318–319.

[6] P. Erdős and S. S. Wagstaff, Jr., The fractional parts of the Bernoulli numbers, Illinois J. Math. **24** (1980), 104–112.

[7] J. Friedlander and H. Iwaniec, Opera de cribro, Amer. Math. Soc. Colloq. Pub. **57**, 2010.

[8] F. Luca and C. Pomerance, The range of Carmichael's universal exponent function, Acta Arith. **162** (2014), 289–308.

[9] N. McNew, P. Pollack, and C. Pomerance, Numbers divisible by a large shifted prime and large torsion subgroups of CM elliptic curves, Int. Math. Res. Not. IMRN (2017), no.18, 5525–5553.

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

[11] M. R. Murty and V. K. Murty, A variant of the Hardy–Ramanujan theorem, Hardy–Ramanujan J. **44** (2021), 32–40.

[12] P. Pollack, A remark on divisor weighted sums, Ramanujan J. **40** (2016), 63–69.

[13] P. Pollack, Nonnegative multiplicative functions on sifted sets, and the square roots of -1 modulo shifted primes, Glasg. Math. J. **62** (2020), no.1, 187–199.

[14] C. Pomerance and S. S. Wagstaff, Jr., The denominators of the Bernoulli numbers, Acta Arith. **209** (2023), 1–15.

[15] K. Prachar, Über die Anzahl der Teiler einer natürlichen Zahl, welche die Form $p-1$ haben, Monatsh. Math. **59** (1955), 91–97.

Max-Planck-Institut für Mathematik, 53111 Bonn, Germany
E-mail address: steve.fan@mpim-bonn.mpg.de

Mathematics Department, Dartmouth College, Hanover, NH 03755, USA
E-mail address: carlp@math.dartmouth.edu