Counting shifted-prime divisors

Steve Fan

Joint work with Carl Pomerance (Dartmouth College)

February 27, 2024

Shifted primes

A *shifted prime* is a positive integer of the form $p+a$, where p is prime and $a \in \mathbb{Z} \setminus \{0\}$.

In this talk, we will concentrate on the case $a=-1$, i.e., shifted primes of the form $p-1$.

Shifted primes

A *shifted prime* is a positive integer of the form $p+a$, where p is prime and $a \in \mathbb{Z} \setminus \{0\}$.

In this talk, we will concentrate on the case $a=-1$, i.e., shifted primes of the form $p-1$.

We say that $p-1$ is a *shifted-prime divisor* of $n \in \mathbb{N}$ if $(p-1) \mid n$.

For each $n \in \mathbb{N}$, we denote by $\omega^*(n)$ the number of shifted-prime divisors of n, i.e.,

$$\omega^*(n) := \sum_{(p-1)\mid n} 1.$$

Example

Shifted-prime divisors of 24: 1, 2, 4, 6, 12. So $\omega^*(24)=5$.

Why shifted primes?

1. Open problems.
   - Twin primes: Are there infinitely many shifted primes $p+2$ that are prime?

Why shifted primes?

1. Open problems.
   - Twin primes: Are there infinitely many shifted primes $p+2$ that are prime?
   - Sophie Germain primes: Are there infinitely many primes p with $(p-1)/2$ also prime? (The prime $q=(p-1)/2$ is then called a Sophie Germain prime, and $p=2q+1$ is called a safe prime.)

Why shifted primes?

1. Open problems.
   - Twin primes: Are there infinitely many shifted primes $p+2$ that are prime?
   - Sophie Germain primes: Are there infinitely many primes p with $(p-1)/2$ also prime? (The prime $q=(p-1)/2$ is then called a Sophie Germain prime, and $p=2q+1$ is called a safe prime.)
   - A conjecture of Erdős and Pomerance on smooth shifted primes: For any fixed $a \in \mathbb{Z}\setminus\{0\}$ and $u \in [1,\infty)$, we have

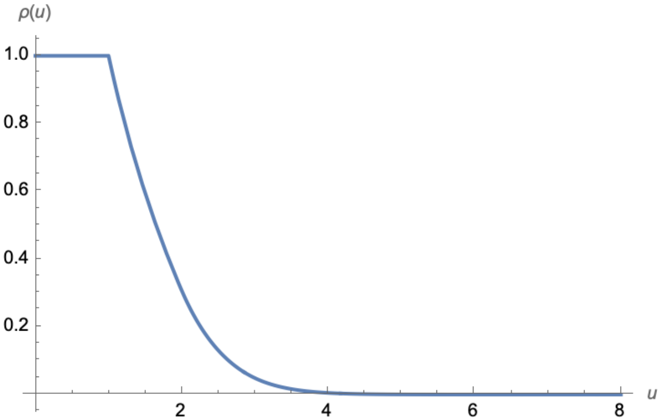
     $$\#\{p \le x \colon P^+(p+a) \le x^{1/u}\} \sim \rho(u)\pi(x) \quad \text{as } x \to \infty,$$

     where $P^+(p+a)$ denotes the largest prime factor of $p+a$, $\pi(x)$ is the prime counting function, and $\rho(u)$ is the Dickman–de Bruijn function, which is the unique continuous function on $[0,\infty)$ satisfying

     $$\begin{cases} \rho(u)=1 & \text{if } u\in[0,1],\\ u\rho'(u)+\rho(u-1)=0 & \text{if } u\ge 1.\end{cases}$$

Why shifted primes?

Figure 1: The Dickman–de Bruijn function $\rho(u)$ on $[1, 8]$



$$\#\{n \le x \colon P^+(n) \le x^{1/u}\} \sim \rho(u)x \quad \text{as } x \to \infty.$$

Why shifted primes?

2. Applications.
   - Carmichael numbers: A Carmichael number n is a composite number satisfying $b^n \equiv b \pmod{n}$ for all $b \in \mathbb{Z}$. Korselt showed in 1899 that $n \in \mathbb{N}$ is a Carmichael number if and only if n is square-free, and $p \mid n \Rightarrow p - 1 \mid n - 1$. Alford, Granville and Pomerance (1994) proved that for sufficiently large x, the interval $[1, x]$ contains at least $x^{2/7}$ Carmichael numbers. One of the key ingredients in their proof is a variant of a result of Prachar on the maximal order of ω^*.

Why shifted primes?

2. Applications.
   - Carmichael numbers: A Carmichael number n is a composite number satisfying $b^n \equiv b \pmod{n}$ for all $b \in \mathbb{Z}$. Korselt showed in 1899 that $n \in \mathbb{N}$ is a Carmichael number if and only if n is square-free, and $p \mid n \Rightarrow p - 1 \mid n - 1$. Alford, Granville and Pomerance (1994) proved that for sufficiently large x, the interval $[1, x]$ contains at least $x^{2/7}$ Carmichael numbers. One of the key ingredients in their proof is a variant of a result of Prachar on the maximal order of ω^*.
   - Bernoulli numbers: The von Staudt–Clausen theorem states that $B_n + \sum_{(p-1)|n} 1/p \in \mathbb{Z}$ for every $n \in 2\mathbb{N}$. By counting numbers with large shifted-prime divisors, Erdős and Wagstaff (1980) proved that for any $n \in 2\mathbb{N}$, the set of $m \in 2\mathbb{N}$ with $B_m \equiv B_n \pmod{1}$ has a positive natural density. Further study of these densities was carried out by Sunseri (1980) and Pomerance and Wagstaff (2023).

Why shifted primes?

2. Applications.
    - Carmichael numbers: A Carmichael number n is a composite number satisfying $b^n \equiv b \pmod{n}$ for all $b \in \mathbb{Z}$. Korselt showed in 1899 that $n \in \mathbb{N}$ is a Carmichael number if and only if n is square-free, and $p \mid n \Rightarrow p-1 \mid n-1$. Alford, Granville and Pomerance (1994) proved that for sufficiently large x, the interval $[1, x]$ contains at least $x^{2/7}$ Carmichael numbers. One of the key ingredients in their proof is a variant of a result of Prachar on the maximal order of ω^*.
    - Bernoulli numbers: The von Staudt–Clausen theorem states that $B_n + \sum_{(p-1)|n} 1/p \in \mathbb{Z}$ for every $n \in 2\mathbb{N}$. By counting numbers with large shifted-prime divisors, Erdős and Wagstaff (1980) proved that for any $n \in 2\mathbb{N}$, the set of $m \in 2\mathbb{N}$ with $B_m \equiv B_n \pmod{1}$ has a positive natural density. Further study of these densities was carried out by Sunseri (1980) and Pomerance and Wagstaff (2023).
    - Fermat's Last Theorem, public key cryptography, primality testing.

The function ω^*

Recall that

$$\omega^*(n) := \sum_{(p-1)|n} 1.$$

How are the values of $\omega^*(n)$ distributed?

The function ω^*

Recall that

$$\omega^*(n) := \sum_{(p-1)|n} 1.$$

How are the values of $\omega^*(n)$ distributed?

It is interesting to compare $\omega^*(n)$ with $\omega(n)$ and $\tau(n)$, where

$$\omega(n) := \sum_{p|n} 1,$$

$$\tau(n) := \sum_{d|n} 1.$$

It is clear that $1 \leq 2^{\omega(n)}, \omega^*(n) \leq \tau(n)$.

ω, τ, and ω^*: extremal orders

The minimal orders of ω, τ and ω^* are 1, 2, 1, respectively.

ω, τ, and ω^*: extremal orders

The minimal orders of ω, τ and ω^* are 1, 2, 1, respectively.

For the maximal orders, we have

$$\limsup_{x\to\infty} \frac{\omega(n)}{\log n/\log\log n} = 1,$$

$$\limsup_{x\to\infty} \frac{\log \tau(n)}{\log n/\log\log n} = \log 2. \quad \text{(Wigert, 1907)}$$

ω, τ, and ω^*: extremal orders

The minimal orders of ω, τ and ω^* are 1, 2, 1, respectively.

For the maximal orders, we have

$$\limsup_{x\to\infty} \frac{\omega(n)}{\log n/\log\log n} = 1,$$

$$\limsup_{x\to\infty} \frac{\log \tau(n)}{\log n/\log\log n} = \log 2. \quad \text{(Wigert, 1907)}$$

Prachar (1955) showed that for infinitely many n,

$$\omega^*(n) > \exp\left(c_1 \frac{\log n}{(\log\log n)^2}\right) \quad \text{(unconditionally)},$$

$$\omega^*(n) > \exp\left((\log\sqrt{2} - \epsilon)\frac{\log n}{\log\log n}\right) \quad \text{(under GRH)},$$

where $c_1 > 0$ is some absolute constant, and $\epsilon > 0$ is fixed but otherwise arbitrary.

ω, τ, and ω^*: extremal orders

Adleman, Pomerance and Rumely (1983) removed one $\log\log n$ factor from Prachar's unconditional bound, obtaining

$$\omega^*(n) > \exp\left(c_2 \frac{\log n}{\log\log n}\right)$$

for infinitely many n, where $c_2 > 0$ is some absolute constant. Combining this with Wigert's result, we have

$$0 < \limsup_{x\to\infty} \frac{\log \omega^*(n)}{\log n / \log\log n} \leq \log 2.$$

Prachar's conditional result implies that this limsup is $\geq \log \sqrt{2}$.

So, $\omega^*(n)$ behaves more like $\tau(n)$ than $\omega(n)$ at the extreme end of the spectrum.

ω, τ, and ω^*: densities

For any arithmetic function f, we denote by $\delta_k(f)$ the *natural density* of the level set $\{n \in \mathbb{N} \colon f(n) = k\}$ for each $k \in \mathbb{N}$, namely,

$$\delta_k(f) := \lim_{x\to\infty} \frac{\#\{n \le x \colon f(n) = k\}}{x},$$

provided that this limit exists.

ω, τ, and ω^*: densities

For any arithmetic function f, we denote by $\delta_k(f)$ the *natural density* of the level set $\{n \in \mathbb{N} \colon f(n) = k\}$ for each $k \in \mathbb{N}$, namely,

$$\delta_k(f) := \lim_{x \to \infty} \frac{\#\{n \leq x \colon f(n) = k\}}{x},$$

provided that this limit exists. Landau (1900) showed that for every fixed $k \in \mathbb{N}$,

$$\#\{n \leq x \colon \omega(n) = k\} \sim \frac{1}{(k-1)!} \cdot \frac{x(\log\log x)^{k-1}}{\log x}$$

as $x \to \infty$. So $\delta_k(\omega) = 0$. Since $\tau(n) \geq 2^{\omega(n)}$, we also have $\delta_k(\tau) = 0$ for every $k \in \mathbb{N}$.

ω, τ, and ω^*: densities

For any arithmetic function f, we denote by $\delta_k(f)$ the *natural density* of the level set $\{n \in \mathbb{N}\colon f(n) = k\}$ for each $k \in \mathbb{N}$, namely,

$$\delta_k(f) := \lim_{x \to \infty} \frac{\#\{n \le x\colon f(n) = k\}}{x},$$

provided that this limit exists. Landau (1900) showed that for every fixed $k \in \mathbb{N}$,

$$\#\{n \le x\colon \omega(n) = k\} \sim \frac{1}{(k-1)!} \cdot \frac{x(\log\log x)^{k-1}}{\log x}$$

as $x \to \infty$. So $\delta_k(\omega) = 0$. Since $\tau(n) \ge 2^{\omega(n)}$, we also have $\delta_k(\tau) = 0$ for every $k \in \mathbb{N}$.

We shall see that $\delta_k(\omega^*) > 0$ for every $k \in \mathbb{N}$!

ω, τ, and ω^*: normal orders

For any arithmetic function f, we say that the nonnegative function g (usually simple and nice) is a *normal order* of f if for every $\epsilon > 0$,

$$|f(n) - g(n)| \le \epsilon g(n)$$

holds for all but $o(x)$ values of $n \in \mathbb{N} \cap [1, x]$.

ω, τ, and ω^*: normal orders

For any arithmetic function f, we say that the nonnegative function g (usually simple and nice) is a *normal order* of f if for every $\epsilon > 0$,

$$|f(n) - g(n)| \leq \epsilon g(n)$$

holds for all but $o(x)$ values of $n \in \mathbb{N} \cap [1, x]$.

Hardy and Ramanujan (1917) showed that $\log\log n$ is a normal order of $\omega(n)$.

ω, τ, and ω^*: normal orders

For any arithmetic function f, we say that the nonnegative function g (usually simple and nice) is a *normal order* of f if for every $\epsilon > 0$,

$$|f(n) - g(n)| \leq \epsilon g(n)$$

holds for all but $o(x)$ values of $n \in \mathbb{N} \cap [1, x]$.

Hardy and Ramanujan (1917) showed that $\log\log n$ is a normal order of $\omega(n)$.

For $\tau(n)$, it is more convenient to study $\log_2 \tau(n) = \log \tau(n)/\log 2$. It can be shown that just like $\omega(n)$, $\log_2 \tau(n)$ has normal order $\log\log n$. One may say that $(\log n)^{\log 2}$ is a "normal order" of $\tau(n)$.

ω, τ, and ω^*: normal orders

For any arithmetic function f, we say that the nonnegative function g (usually simple and nice) is a *normal order* of f if for every $\epsilon > 0$,

$$|f(n) - g(n)| \leq \epsilon g(n)$$

holds for all but $o(x)$ values of $n \in \mathbb{N} \cap [1, x]$.

Hardy and Ramanujan (1917) showed that $\log \log n$ is a normal order of $\omega(n)$.

For $\tau(n)$, it is more convenient to study $\log_2 \tau(n) = \log \tau(n) / \log 2$. It can be shown that just like $\omega(n)$, $\log_2 \tau(n)$ has normal order $\log \log n$. One may say that $(\log n)^{\log 2}$ is a "normal order" of $\tau(n)$.

What about $\omega^*(n)$ (or $\log \omega^*(n)$)?

ω, τ, and ω^*: normal orders

For any arithmetic function f, we say that the nonnegative function g (usually simple and nice) is a *normal order* of f if for every $\epsilon > 0$,

$$|f(n) - g(n)| \le \epsilon g(n)$$

holds for all but $o(x)$ values of $n \in \mathbb{N} \cap [1, x]$.

Hardy and Ramanujan (1917) showed that $\log\log n$ is a normal order of $\omega(n)$.

For $\tau(n)$, it is more convenient to study $\log_2 \tau(n) = \log \tau(n)/\log 2$. It can be shown that just like $\omega(n)$, $\log_2 \tau(n)$ has normal order $\log\log n$. One may say that $(\log n)^{\log 2}$ is a "normal order" of $\tau(n)$.

What about $\omega^*(n)$ (or $\log \omega^*(n)$)? No nice normal orders.

ω, τ, and ω^*: moments and distributions

For any arithmetic function f, we denote by $M_k(x; f)$ the *kth moment* of f for each $k \in \mathbb{N}$. That is,

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

ω, τ, and ω^*: moments and distributions

For any arithmetic function f, we denote by $M_k(x; f)$ the k*th moment* of f for each $k \in \mathbb{N}$. That is,

$$M_k(x; f) := \frac{1}{x} \sum_{n \leq x} f(n)^k.$$

For every fixed $k \in \mathbb{N}$, we have

$$M_k(x; \omega) \sim (\log \log x)^k,$$
$$M_k(x; \tau) \sim a_k (\log x)^{2^k - 1},$$

where

$$a_k := \frac{1}{(2^k - 1)!} \prod_p \left(1 - \frac{1}{p}\right)^{2^k} \sum_{\nu \geq 0} \frac{(\nu + 1)^k}{p^\nu}.$$

ω, τ, and ω^*: moments and distributions

In fact, Delange (1953) showed that

$$\frac{1}{x}\sum_{n\le x}(\omega(n)-\log\log n)^k=(1_{2\mathbb{N}}(k)+o(1))(k-1)!!(\log\log x)^{\frac{k}{2}}, \qquad (1)$$

which implies that

$$\lim_{x\to\infty}\frac{1}{x}\cdot\#\left\{n\le x\colon \frac{\omega(n)-\log\log n}{\sqrt{\log\log n}}\le V\right\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{V}e^{-v^2/2}\,dv \qquad (2)$$

for any given $V\in\mathbb{R}$. This is the celebrated Erdős–Kac theorem, first established by Erdős and Kac in 1940. Delange's result (1) was generalized by Halberstam (1954) to general additive functions with bounded values on primes. Particularly, Halberstam's result implies that (1) and (2) continue to hold with ω replaced by $\log_2\tau$.

Interlude: ω and τ on shifted primes

The distribution of ω on shifted primes is similar to its distribution on natural numbers. Erdős (1935) showed that $\log\log p$ is a normal order of $\omega(p-1)$.

Interlude: ω and τ on shifted primes

The distribution of ω on shifted primes is similar to its distribution on natural numbers. Erdős (1935) showed that $\log\log p$ is a normal order of $\omega(p-1)$. This was greatly improved by Halberstam (1955) who obtained in particular an Erdős–Kac theorem for ω on shifted primes $p+a$:

$$\lim_{x\to\infty}\frac{1}{\pi(x)}\cdot\#\left\{p\le x\colon \frac{\omega(p+a)-\log\log p}{\sqrt{\log\log p}}\le V\right\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{V}e^{-v^2/2}\,dv.$$

Similarly, his results also imply that the same holds with ω replaced by $\log_2\tau$.

Interlude: ω and τ on shifted primes

The distribution of ω on shifted primes is similar to its distribution on natural numbers. Erdős (1935) showed that $\log\log p$ is a normal order of $\omega(p-1)$. This was greatly improved by Halberstam (1955) who obtained in particular an Erdős–Kac theorem for ω on shifted primes $p+a$:

$$\lim_{x\to\infty} \frac{1}{\pi(x)} \cdot \#\left\{p \le x \colon \frac{\omega(p+a) - \log\log p}{\sqrt{\log\log p}} \le V\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{V} e^{-v^2/2}\, dv.$$

Similarly, his results also imply that the same holds with ω replaced by $\log_2 \tau$.

For τ, Titchmarsh (1931) proved, conditionally on GRH, that

$$\frac{1}{\pi(x)} \sum_{p\le x} \tau(p-1) \sim \frac{\zeta(2)\zeta(3)}{\zeta(6)} \log x.$$

Linnik (1961) gave an unconditional proof based on his complicated dispersion method. Independently, Rodriguez (1965) and Halberstam (1967) obtained quick proofs based on the Bombieri–Vinogradov theorem which came out in 1965.

ω, τ, and ω^*: moments and distributions

Prachar (1955) showed $M_1(x;\omega^*) \sim \log\log x$, by observing that

$$\frac{1}{x}\sum_{n\le x}\omega^*(n) = \frac{1}{x}\sum_{n\le x}\sum_{p-1|n} 1 = \frac{1}{x}\sum_{p\le x+1}\left\lfloor \frac{x}{p-1}\right\rfloor$$

and applying Mertens' second theorem. Since $M_1(x;\omega) \sim \log\log x$, perhaps $M_2(x;\omega^*) \asymp (\log\log x)^2$ just like $M_2(x;\omega)$?

ω, τ, and ω^*: moments and distributions

Prachar (1955) showed $M_1(x;\omega^*) \sim \log\log x$, by observing that

$$\frac{1}{x}\sum_{n\le x}\omega^*(n) = \frac{1}{x}\sum_{n\le x}\sum_{p-1|n} 1 = \frac{1}{x}\sum_{p\le x+1}\left\lfloor \frac{x}{p-1}\right\rfloor$$

and applying Mertens' second theorem. Since $M_1(x;\omega) \sim \log\log x$, perhaps $M_2(x;\omega^*) \asymp (\log\log x)^2$ just like $M_2(x;\omega)$?

Prachar proved $M_2(x;\omega^*) = O((\log x)^2)$. This was improved to $O(\log x)$ by Murty and Murty (2021) who also showed $M_2(x;\omega^*) \gg (\log\log x)^3$. They also conjectured $M_2(x;\omega^*) \sim C\log x$ for some constant $C > 0$.

ω, τ, and ω^*: moments and distributions

Prachar (1955) showed $M_1(x;\omega^*) \sim \log\log x$, by observing that

$$\frac{1}{x}\sum_{n\le x}\omega^*(n) = \frac{1}{x}\sum_{n\le x}\sum_{p-1|n} 1 = \frac{1}{x}\sum_{p\le x+1}\left\lfloor \frac{x}{p-1}\right\rfloor$$

and applying Mertens' second theorem. Since $M_1(x;\omega) \sim \log\log x$, perhaps $M_2(x;\omega^*) \asymp (\log\log x)^2$ just like $M_2(x;\omega)$?

Prachar proved $M_2(x;\omega^*) = O((\log x)^2)$. This was improved to $O(\log x)$ by Murty and Murty (2021) who also showed $M_2(x;\omega^*) \gg (\log\log x)^3$. They also conjectured $M_2(x;\omega^*) \sim C\log x$ for some constant $C > 0$.

Via a simple application of the Bombieri–Vinogradov theorem, Ding (2023) obtained the stronger lower bound $M_2(x;\omega^*) \gg \log x$, matching the order of the upper bound of Murty and Murty.

ω, τ, and ω^*: moments and distributions

Prachar (1955) showed $M_1(x;\omega^*) \sim \log\log x$, by observing that

$$\frac{1}{x}\sum_{n\le x}\omega^*(n) = \frac{1}{x}\sum_{n\le x}\sum_{p-1|n} 1 = \frac{1}{x}\sum_{p\le x+1}\left\lfloor\frac{x}{p-1}\right\rfloor$$

and applying Mertens' second theorem. Since $M_1(x;\omega) \sim \log\log x$, perhaps $M_2(x;\omega^*) \asymp (\log\log x)^2$ just like $M_2(x;\omega)$?

Prachar proved $M_2(x;\omega^*) = O((\log x)^2)$. This was improved to $O(\log x)$ by Murty and Murty (2021) who also showed $M_2(x;\omega^*) \gg (\log\log x)^3$. They also conjectured $M_2(x;\omega^*) \sim C\log x$ for some constant $C > 0$.

Via a simple application of the Bombieri–Vinogradov theorem, Ding (2023) obtained the stronger lower bound $M_2(x;\omega^*) \gg \log x$, matching the order of the upper bound of Murty and Murty. So $M_2(x;\omega^*)$ grows more like $M_2(x;\tau) \asymp (\log x)^3$ with an additional primality constraint placed.

ω, τ, and ω^*: moments and distributions

Murty and Murty observed that

$$M_2(x;\omega^*) = \frac{1}{x}\sum_{n\le x}\left(\sum_{p-1\mid n} 1\right)^2 = \frac{1}{x}\sum_{[p-1,q-1]\le x}\left\lfloor \frac{x}{[p-1,q-1]}\right\rfloor.$$

An old result of Erdős and Prachar (1955) states that the number of prime pairs (p,q) with $[p-1,q-1]\le x$ is $O(x)$. Using this we arrive at

$$M_2(x;\omega^*) = \sum_{[p-1,q-1]\le x}\frac{1}{[p-1,q-1]} + O(1).$$

The upper bound $M_2(x;\omega^*) = O(\log x)$ follows now from the theorem of Erdős and Prachar and partial summation. Murty and Murty went on to conclude that

$$M_2(x;\omega^*) = \sum_{p,q\le x}\frac{1}{[p-1,q-1]} + O(1).$$

ω, τ, and ω^*: moments and distributions

They proved $M_2(x;\omega^*) \gg (\log\log x)^3$ by bounding the last sum above. This last equation above is also the starting point of Ding's proof that $M_2(x;\omega^*) \gg \log x$.

ω, τ, and ω^*: moments and distributions

They proved $M_2(x;\omega^*) \gg (\log\log x)^3$ by bounding the last sum above. This last equation above is also the starting point of Ding's proof that $M_2(x;\omega^*) \gg \log x$.

Based on the same equation, Ding also argued, assuming the Elliott–Halberstam conjecture, that $M_2(x;\omega^*) \sim C\log x$, where $C = 2\zeta(2)\zeta(3)/\zeta(6) \approx 3.88719$.

ω, τ, and ω^*: moments and distributions

They proved $M_2(x;\omega^*) \gg (\log\log x)^3$ by bounding the last sum above. This last equation above is also the starting point of Ding's proof that $M_2(x;\omega^*) \gg \log x$.

Based on the same equation, Ding also argued, assuming the Elliott–Halberstam conjecture, that $M_2(x;\omega^*) \sim C\log x$, where $C = 2\zeta(2)\zeta(3)/\zeta(6) \approx 3.88719$.

However, there is a problem with the last equation: Murty and Murty concluded

$$M_2(x;\omega^*) = \sum_{[p-1,q-1]\le x} \frac{1}{[p-1,q-1]} + O(1) = \sum_{p,q\le x} \frac{1}{[p-1,q-1]} + O(1).$$

For the second equality to hold, they assumed implicitly that

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

But is this really true?

ω, τ, and ω^*: moments and distributions

They proved $M_2(x;\omega^*) \gg (\log\log x)^3$ by bounding the last sum above. This last equation above is also the starting point of Ding's proof that $M_2(x;\omega^*) \gg \log x$.

Based on the same equation, Ding also argued, assuming the Elliott–Halberstam conjecture, that $M_2(x;\omega^*) \sim C\log x$, where $C = 2\zeta(2)\zeta(3)/\zeta(6) \approx 3.88719$.

However, there is a problem with the last equation: Murty and Murty concluded

$$M_2(x;\omega^*) = \sum_{[p-1,q-1]\le x} \frac{1}{[p-1,q-1]} + O(1) = \sum_{p,q\le x} \frac{1}{[p-1,q-1]} + O(1).$$

For the second equality to hold, they assumed implicitly that

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

But is this really true? The answer is no.

Our goals

Our research addresses the following:

1. correcting the error in Ding's proof of $M_2(x;\omega^*) \gg \log x$;
2. studying the density $\delta_k(\omega^*)$ of the level set $\{n \in \mathbb{N} \colon \omega^*(n) = k\}$;
3. investigating higher moments of ω^*, starting with $M_3(x;\omega^*)$.

Confirming the error

Recall our question:

$$\sum_{\substack{p,q\le x\\ [p-1,q-1]>x}} \frac{1}{[p-1,q-1]} = O(1)?$$

The following theorem disproves this.

Theorem 1 (F., Pomerance, 2024)

We have

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

for sufficiently large x.

An easy fix

We start with

$$M_2(x;\omega^*) = \frac{1}{x} \sum_{[p-1,q-1]\le x} \left\lfloor \frac{x}{[p-1,q-1]} \right\rfloor .$$

Note that if $p, q \le \sqrt{x}$, then $[p-1, q-1] \le (p-1)(q-1) < x$. Thus,

$$M_2(x;\omega^*) \ge \frac{1}{x} \sum_{p,q\le\sqrt{x}} \left\lfloor \frac{x}{[p-1,q-1]} \right\rfloor = \sum_{p,q\le\sqrt{x}} \frac{1}{[p-1,q-1]} + O\left(\frac{1}{\log x}\right).$$

What Ding actually proved is

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

Applying this lower bound with $\sqrt{x}$ in place of x yields $M_2(x;\omega^*) \gg \log x$. We also have a new, quick proof of this lower bound independent of Ding's.

The constant C

The constant $C = 2\zeta(2)\zeta(3)/\zeta(6) \approx 3.88719$ that Ding got for the Murty–Murty conjecture $M_2(x;\omega^*) \sim C \log x$ is probably incorrect. So, what is the correct value of C?

The constant C

The constant $C = 2\zeta(2)\zeta(3)/\zeta(6) \approx 3.88719$ that Ding got for the Murty–Murty conjecture $M_2(x;\omega^*) \sim C \log x$ is probably incorrect. So, what is the correct value of C?

Let

$$S_2(x;\omega^*) := \frac{1}{x} \cdot \#\{(p,q) \colon [p-1, q-1] \le x\}.$$

The result of Erdős and Prachar is equivalent to $S_2(x;\omega^*) = O(1)$. Partial summation gives the connection between $M_2(x;\omega^*)$ and $S_2(x;\omega^*)$:

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

So, the conjecture $S_2(x;\omega^*) \sim C$ implies the Murty–Murty conjecture.

The constant C

Table 1: Numerical values of $M_2(10^k;\omega^*)$ and $S_2(10^k;\omega^*)$

k	$M_2(10^k;\omega^*)$	$S_2(10^k;\omega^*)$	$3\log 10^k-6$	$3.2\left(1-\frac{1}{\log 10^k}\right)$
2	9.71	2.42	7.82	2.51
3	15.530	2.624	14.723	2.737
4	21.9128	2.8175	21.6310	2.8526
5	28.49311	2.88636	28.53878	2.92205
6	35.261891	2.950910	35.446532	2.968376
7	42.1296839	2.9923851	42.3542870	3.0014654
8	49.02181351	3.02166709	49.26204223	3.02628221
9	56.067311859	3.043042188	56.169797511	3.045584184
10	63.1033824202	3.0595625181	63.0775527898	3.0610257658

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

The densities $\delta_k(\omega^*)$

We have seen that $\delta_k(\omega) = \delta_k(\tau) = 0$ for every fixed $k \in \mathbb{N}$. Consequently, the densities of the tails $\{n \in \mathbb{N} \colon \omega(n) > k\}$ and $\{n \in \mathbb{N} \colon \tau(n) > k\}$ are both equal to 1.

The densities $\delta_k(\omega^*)$

We have seen that $\delta_k(\omega) = \delta_k(\tau) = 0$ for every fixed $k \in \mathbb{N}$. Consequently, the densities of the tails $\{n \in \mathbb{N} \colon \omega(n) > k\}$ and $\{n \in \mathbb{N} \colon \tau(n) > k\}$ are both equal to 1. But this is not the case for ω^*.

Theorem 2 (F., Pomerance, 2024)

For $x, y \geq 1$, let $N(x,y) := \#\{n \leq x \colon \omega^*(n) \geq y\}$. Then there exists a suitable constant $c > 0$ such that for all $x \geq 1$ and all sufficiently large y,

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

The lower bound follows from the result of Adleman, Pomerance and Rumely (1983) on the maximal order of ω^*, while the proof of the upper bound makes use of a theorem due to McNew, Pollack and Pomerance (2017), which asserts that the number of $n \leq x$ with a shifted prime divisor $> y$ is $O(x/(\log y)^{\beta+o(1)})$, where $\beta = 1 - (1 + \log\log 2)/\log 2$ is the Erdős–Ford–Tenenbaum constant.

The densities $\delta_k(\omega^*)$

Now we turn to the k-level set $\mathcal{L}_k := \{n \in \mathbb{N} \colon \omega^*(n) = k\}$.

Theorem 3 (F., Pomerance, 2024)

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

In order to establish Theorem 3, one should at least be to able to verify that $\mathcal{L}_k \neq \emptyset$. This is the key step in our proof of Theorem 3.

The densities $\delta_k(\omega^*)$

Now we turn to the k-level set $\mathcal{L}_k := \{n \in \mathbb{N} \colon \omega^*(n) = k\}$.

Theorem 3 (F., Pomerance, 2024)

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

In order to establish Theorem 3, one should at least be to able to verify that $\mathcal{L}_k \neq \emptyset$. This is the key step in our proof of Theorem 3.

Our strategy: Since $\mathcal{L}_1 = \mathbb{N} \setminus 2\mathbb{N}$, we may suppose $k \geq 2$, so that $\mathcal{L}_k \subseteq 2\mathbb{N}$. The idea is to show that there exists a prime p such that $\omega^*(n(p-1)/2) = \omega^*(n) + 1$, from which the claim that $\mathcal{L}_k \neq \emptyset$ follows by induction. To find such a prime, we appeal to Chen's theorem which asserts that the number of primes $p \leq x$ for which $(p-1)/2$ is the product of at most two prime factors, each of which is $> x^{3/11}$, is $\gg x/(\log x)^2$. We then show that the number of those unqualified p's is negligible, completing the proof of our claim.

The densities $\delta_k(\omega^*)$

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 $k \geq 3$.

The densities $\delta_k(\omega^*)$

In our proof of Theorem 3, we used a result of Erdős and Wagstaff (1980) concerning the density $\delta(\langle n \rangle)$ of $\langle n \rangle$ for a given $n \in \mathbb{N}$, where

$$\langle n \rangle := \#\{m \in \mathbb{N} \colon (p-1) \mid m \Leftrightarrow (p-1) \mid n\}.$$

Thus, $B_m \equiv B_n \pmod 1 \Leftrightarrow m \in \langle n \rangle$.

The densities $\delta_k(\omega^*)$

In our proof of Theorem 3, we used a result of Erdős and Wagstaff (1980) concerning the density $\delta(\langle n\rangle)$ of $\langle n\rangle$ for a given $n \in \mathbb{N}$, where

$$\langle n\rangle := \#\{m \in \mathbb{N} \colon (p-1) \mid m \Leftrightarrow (p-1) \mid n\}.$$

Thus, $B_m \equiv B_n \pmod 1 \Leftrightarrow m \in \langle n\rangle$.

Note that $\langle 1\rangle = \mathcal{L}_1 = \mathbb{N} \setminus 2\mathbb{N}$, so that $\delta(\langle n\rangle) = 1/2$ for odd n. Erdős and Wagstaff showed that $\delta(\langle n\rangle)$ exists and is positive for every $n \in \mathbb{N}$. They also observed that if $n = \min\langle n\rangle$, then $\delta(\langle n\rangle) < 1/n$. In this case, they asked for a positive lower bound for $\delta(\langle n\rangle)$.

The densities $\delta_k(\omega^*)$

In our proof of Theorem 3, we used a result of Erdős and Wagstaff (1980) concerning the density $\delta(\langle n\rangle)$ of $\langle n\rangle$ for a given $n \in \mathbb{N}$, where

$$\langle n\rangle := \#\{m \in \mathbb{N}\colon (p-1) \mid m \Leftrightarrow (p-1) \mid n\}.$$

Thus, $B_m \equiv B_n \pmod 1 \Leftrightarrow m \in \langle n\rangle$.

Note that $\langle 1\rangle = \mathcal{L}_1 = \mathbb{N} \setminus 2\mathbb{N}$, so that $\delta(\langle n\rangle) = 1/2$ for odd n. Erdős and Wagstaff showed that $\delta(\langle n\rangle)$ exists and is positive for every $n \in \mathbb{N}$. They also observed that if $n = \min\langle n\rangle$, then $\delta(\langle n\rangle) < 1/n$. In this case, they asked for a positive lower bound for $\delta(\langle n\rangle)$.

Theorem 4 (F., Pomerance, 2024)

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

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

The moments $M_k(x;\omega^*)$

For every $k \in \mathbb{N}$, we consider

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

Then we have

$$M_k(x;\omega^*) = \frac{1}{x}\sum_{[p_1-1,\ldots,p_k-1]\le x}\left\lfloor\frac{x}{[p_1-1,\ldots,p_k-1]}\right\rfloor.$$

This shows that $M_k(x;\omega^*)$ is intimately related to

$$S_k(x;\omega^*) := \frac{1}{x}\cdot\#\{(p_1,\ldots,p_k)\colon [p_1-1,\ldots,p_k-1]\le x\}.$$

Again, it can be shown by partial summation that if $S_k(x;\omega^*)(x) \asymp_k (\log x)^{c_k}$ for some absolute constant $c_k>0$, then $M_k(x;\omega^*) \asymp_k (\log x)^{c_k+1}$.

The moments $M_k(x;\omega^*)$

For $k \geq 2$, it is natural to relate the function $\omega^*(n)^k$ to $\tau(n)^k$. Recall that

$$M_k(x;\tau) = \frac{1}{x}\sum_{n\leq x}\tau(n)^k \sim a_k(\log x)^{2^k-1}$$

for every $k \geq 1$.

The moments $M_k(x;\omega^*)$

For $k \geq 2$, it is natural to relate the function $\omega^*(n)^k$ to $\tau(n)^k$. Recall that

$$M_k(x;\tau) = \frac{1}{x}\sum_{n\leq x}\tau(n)^k \sim a_k(\log x)^{2^k-1}$$

for every $k \geq 1$. Comparing ω^* with τ and taking the primality conditions into account, one may conjecture that

$$M_k(x;\omega^*) \sim \mu_k(\log x)^{2^k-k-1},$$
$$S_k(x;\omega^*) \sim (2^k-k-1)\mu_k(\log x)^{2^k-k-2},$$

for every $k \geq 2$, where $\mu_k > 0$ is a constant depending on k.

The moments $M_k(x;\omega^*)$

For $k \geq 2$, it is natural to relate the function $\omega^*(n)^k$ to $\tau(n)^k$. Recall that

$$M_k(x;\tau) = \frac{1}{x}\sum_{n\leq x}\tau(n)^k \sim a_k(\log x)^{2^k-1}$$

for every $k \geq 1$. Comparing ω^* with τ and taking the primality conditions into account, one may conjecture that

$$M_k(x;\omega^*) \sim \mu_k(\log x)^{2^k-k-1},$$
$$S_k(x;\omega^*) \sim (2^k-k-1)\mu_k(\log x)^{2^k-k-2},$$

for every $k \geq 2$, where $\mu_k > 0$ is a constant depending on k.

We proved the upper and lower bounds for $M_3(x;\omega^*)$ of the conjectured magnitude.

The third moment $M_3(x;\omega^*)$

We have the following theorem concerning $M_3(x;\omega^*)$.

Theorem 5 (F., Pomerance, 2024)

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

The third moment $M_3(x;\omega^*)$

We have the following theorem concerning $M_3(x;\omega^*)$.

Theorem 5 (F., Pomerance, 2024)

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

Proof ideas:

- To prove the upper bound, we show

$$S_3(x;\omega^*) = \frac{1}{x} \cdot \#\{(p,q,r) \colon [p-1,q-1,r-1] \leq x\} \ll (\log x)^3.$$

To do so, we write

$$\begin{aligned} p-1 &= adeg, & dg &= \gcd(p-1,q-1), \\ q-1 &= bdfg, & eg &= \gcd(p-1,r-1), \\ r-1 &= cefg, & fg &= \gcd(q-1,r-1), \end{aligned}$$

and $g = \gcd(p-1,q-1,r-1)$.

The third moment $M_3(x;\omega^*)$

The dictionary on the previous slide:

$$\begin{aligned} p-1 &= adeg, & dg &= \gcd(p-1,q-1), \\ q-1 &= bdfg, & eg &= \gcd(p-1,r-1), \\ r-1 &= cefg, & fg &= \gcd(q-1,r-1), \end{aligned}$$

and $g = \gcd(p-1,q-1,r-1)$. Then $[p-1,q-1,r-1] \le x$ becomes $abcdefg \le x$, subject to the condition that $adeg+1$, $bdfg+1$ and $cefg+1$ are simultaneously prime.

The third moment $M_3(x;\omega^*)$

The dictionary on the previous slide:

$$\begin{aligned} p-1 &= adeg, & dg &= \gcd(p-1,q-1), \\ q-1 &= bdfg, & eg &= \gcd(p-1,r-1), \\ r-1 &= cefg, & fg &= \gcd(q-1,r-1), \end{aligned}$$

and $g = \gcd(p-1,q-1,r-1)$. Then $[p-1,q-1,r-1] \le x$ becomes $abcdefg \le x$, subject to the condition that $adeg+1$, $bdfg+1$ and $cefg+1$ are simultaneously prime.

With this set-up, we see by symmetry that there are three possible cases:

$$m := \max\{a,b,c,d,e,f,g\} = a, d, \text{ or } g.$$

In each case, we sum over m with the other variables fixed and use sieve bounds to estimate the sum with the above primality constraints. Then we sum the result over the rest of variables in a convenient order and handle the average of certain nonnegative multiplicative functions over shifted primes.

The third moment $M_3(x;\omega^*)$

- To prove the lower bound, we start with

$$M_k(x;\omega^*) \geq \frac{1}{2} \sum_{[p-1,q-1,r-1]\leq x/2} \frac{1}{[p-1,q-1,r-1]}.$$

Using the convolution identity $\mathrm{id} = 1 * \varphi$, we may write

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

Then we have

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

The third moment $M_3(x;\omega^*)$

By considering only the squarefree u's, we arrive at

$$M_k(x;\omega^*) \geq \frac{1}{2}\sum_{r\leq z}\frac{1}{r-1}\sum_{u|r-1}\mu(u)^2\varphi(u)M(y;u),$$

where $y \geq z$ are suitable powers of x satisfying $yz \leq x$, and

$$M(y;u) := \sum_{\substack{[p-1,q-1]\leq y\\ u|[p-1,q-1]}}\frac{1}{[p-1,q-1]}.$$

The third moment $M_3(x;\omega^*)$

By considering only the squarefree u's, we arrive at

$$M_k(x;\omega^*) \geq \frac{1}{2}\sum_{r\leq z}\frac{1}{r-1}\sum_{u\mid r-1}\mu(u)^2\varphi(u)M(y;u),$$

where $y \geq z$ are suitable powers of x satisfying $yz \leq x$, and

$$M(y;u) := \sum_{\substack{[p-1,q-1]\leq y\\ u\mid[p-1,q-1]}}\frac{1}{[p-1,q-1]}.$$

The key to handling $M(y;u)$ is the following result due to Alford, Granville and Pomerance (1994): $\forall\epsilon > 0$, there exist $\delta \in (0,1)$ and $x_0 \geq 2$, such that

$$\left|\pi(y;k,a) - \frac{y}{\varphi(k)\log y}\right| \leq \epsilon\frac{y}{\varphi(k)\log y}$$

for all $y \geq x \geq x_0$, all $k \in \mathbb{N}\cap[1,x^\delta]$ and all $a \in \mathbb{Z}$ with $\gcd(a,k) = 1$, except possibly for those k divisible by a certain number $k_0(x) > \log x$.

Future research

We plan to investigate the following questions:

1. Can we prove good upper and lower bounds for the densities $\delta_k(\omega^*)$?
2. Can we improve the lower bound for $\delta(\langle n \rangle)$ supplied by Theorem 4?
3. What is the true value of C in the Murty–Murty conjecture $M_2(x;\omega^*) \sim C \log x$?
4. Can we prove upper and lower bounds of the conjectured magnitude for $M_k(x;\omega^*)$ when $k \geq 4$?
5. What is the value of
$$\limsup_{n\to\infty} \frac{\log \omega^*(n)}{n/\log\log n}?$$
6. What is the distribution of ω^* (or $\log \omega^*$)? What about $\omega^*(p-1)$?

Thank you!