Analytic Number Theory
Lecture 4

Mariusz Mirek
Rutgers University

Padova, March 18, 2025.

Supported by the NSF grant DMS-2154712,
and the CAREER grant DMS-2236493.

Summation methods

Proposition

Let $x \geqslant 1$ be a real number and f, g be two arithmetic functions. Then

$$\sum_{n \in [x]} (f \star g)(n) = \sum_{d \in [x]} f(d) \sum_{k \in [x/d]} g(k).$$

Here we use the convention $[x] := \mathbb{Z} \cap (0, x] := \{1, \ldots, \lfloor x \rfloor\}$ for any $x \in \mathbb{R}_+$.

Proof.

- Observe that

$$\sum_{n \in [x]} (f \star g)(n) = \sum_{n \in [x]} \sum_{d \mid n} f(d) g(n/d) = \sum_{d \in [x]} f(d) \sum_{\substack{n \in [x] \\ d \mid n}} g(n/d).$$

- Making the change of variable $n = kd$ in the inner sum and noticing that $n \in [x]$ and $d \mid n$ is equivalent to $k \in [x/d]$, we finally obtain

$$\sum_{n \in [x]} (f \star g)(n) = \sum_{d \in [x]} f(d) \sum_{k \in [x/d]} g(k). \qquad \square$$

Examples

- Since $\tau = \mathbf{1} \star \mathbf{1}$, then by the previous proposition and the easy equality $\lfloor x \rfloor = x + O(1)$ we obtain

$$\sum_{n \in [x]} \tau(n) = \sum_{d \in [x]} \sum_{k \in [x/d]} 1 = \sum_{d \in [x]} \left\lfloor \frac{x}{d} \right\rfloor = x \sum_{d \in [x]} \frac{1}{d} + O(x) = x \log x + O(x).$$

We shall see later how to improve on this result.

- Since $\sigma = \mathbf{1} \star \mathrm{Id}$, then by the previous proposition we have

$$\sum_{n \in [x]} \sigma(n) = \sum_{d \in [x]} \sum_{k \in [x/d]} k = \frac{1}{2} \sum_{d \in [x]} \left\lfloor \frac{x}{d} \right\rfloor \left(\left\lfloor \frac{x}{d} \right\rfloor + 1 \right)$$

and using $\lfloor x \rfloor = x + O(1)$, we obtain

$$\sum_{n \in [x]} \sigma(n) = \frac{x^2}{2} \sum_{d \in [x]} \frac{1}{d^2} + O\left(x \sum_{d \in [x]} \frac{1}{d} \right).$$

Since $\sum_{d \in [x]} \frac{1}{d^2} = \sum_{d=1}^{\infty} \frac{1}{d^2} - \sum_{d > x} \frac{1}{d^2} = \zeta(2) + O\left(\frac{1}{x}\right)$, we conclude

$$\sum_{n \in [x]} \sigma(n) = \frac{x^2 \pi^2}{12} + O(x \log x).$$

Examples

- Since $\varphi = \mu \star \mathrm{Id}$, then by the previous proposition we have

$$\sum_{n\in[x]} \varphi(n) = \sum_{d\in[x]} \mu(d) \sum_{k\in[x/d]} k = \frac{1}{2} \sum_{d\in[x]} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor \left(\left\lfloor \frac{x}{d} \right\rfloor + 1 \right)$$

and using $\lfloor x \rfloor = x + O(1)$, we obtain

$$\sum_{n\in[x]} \varphi(n) = \frac{x^2}{2} \sum_{d\in[x]} \frac{\mu(d)}{d^2} + O\left(x \sum_{d\in[x]} \frac{1}{d} \right).$$

Since $\sum_{d\in[x]} \frac{\mu(d)}{d^2} = \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} - \sum_{d>x} \frac{\mu(d)}{d^2} = \zeta(2)^{-1} + O\left(\frac{1}{x}\right)$, we conclude

$$\sum_{n\in[x]} \varphi(n) = \frac{3x^2}{\pi^2} + O(x \log x).$$

Examples

- Since $\mu \star \mathbf{1} = \delta$, by the previous proposition we obtain

$$1 = \sum_{n \in [x]} (\mu \star \mathbf{1})(n) = \sum_{d \in [x]} \mu(d) \sum_{k \in [x/d]} 1 = \sum_{d \in [x]} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor,$$

  which is a very important identity.
- Using this identity we obtain that

$$N \sum_{n \in [N]} \frac{\mu(n)}{n} = \sum_{n \in [N]} \mu(n) \left\lfloor \frac{N}{n} \right\rfloor + \sum_{n \in [N]} \mu(n) \left\{ \frac{N}{n} \right\} = 1 + \sum_{n \in [N-1]} \mu(n) \left\{ \frac{N}{n} \right\}$$

  so that

$$\left| \sum_{n \in [N]} \frac{\mu(n)}{n} \right| \leqslant \frac{1}{N} \left(1 + \sum_{n \in [N-1]} \mu(n) \left\{ \frac{N}{n} \right\} \right) \leqslant \frac{1 + N - 1}{N} = 1.$$

Dirichlet hyperbola principle

Proposition (Dirichlet hyperbola principle)

Let $1 \leqslant T \leqslant x$ *be real numbers and* f, g *be two arithmetic functions. Then*

$$\sum_{n\in[x]} (f\star g)(n) = \sum_{n\in[T]} f(n) \sum_{k\in[x/n]} g(k) + \sum_{k\in[x/T]} g(k) \sum_{n\in[x/k]} f(n) - \sum_{n\in[T]} f(n) \sum_{k\in[x/T]} g(k).$$

Proof.

Splitting the sum of the right-hand side of of the proposition gives

$$\sum_{n\in[x]} (f\star g)(n) = \sum_{d\in[T]} f(d) \sum_{k\in[x/d]} g(k) + \sum_{T<d\leqslant x} f(d) \sum_{k\in[x/d]} g(k),$$

and

$$\sum_{T<d\leqslant x} f(d) \sum_{k\in[x/d]} g(k) = \sum_{k\in[x/T]} g(k) \sum_{T<d\leqslant x/k} f(d) = \sum_{k\in[x/T]} g(k) \left(\sum_{d\in[x/k]} f(d) - \sum_{d\in[T]} f(d) \right). \qquad \square$$

Improved bounds for the divisor problem

Theorem (Dirichlet)

For $x \in \mathbb{R}_+$ sufficiently large, we have

$$\sum_{n \in [x]} \tau(n) = x(\log x + 2\gamma - 1) + O(\sqrt{x}).$$

Proof.

- Since $\tau = \mathbf{1} \star \mathbf{1}$, by the previous proposition with $T = \sqrt{x}$ and the estimate $\lfloor x \rfloor = x + O(1)$ we obtain

$$\begin{aligned}
\sum_{n \in [x]} \tau(n) &= 2 \sum_{m \in [\sqrt{x}]} \left\lfloor \frac{x}{m} \right\rfloor - \lfloor \sqrt{x} \rfloor^2 \\
&= 2x \sum_{m \in [\sqrt{x}]} \frac{1}{m} + O(\sqrt{x}) - (\sqrt{x} + O(1))^2 \\
&= 2x \left(\log \sqrt{x} + \gamma + O\big(x^{-1/2}\big) \right) - x + O(\sqrt{x}) \\
&= x \log x + x(2\gamma - 1) + O(\sqrt{x}). \qquad \square
\end{aligned}$$

Important arithmetic functions involvong primes

1. The von Mangoldt function Λ is defined by

$$\Lambda(n) := \begin{cases} \log p, & \text{if } n = p^k \text{ for some prime } p \text{ and } k \in \mathbb{Z}_+, \\ 0, & \text{otherwise} \end{cases}.$$

2. The first Chebyshev function ϑ is defined for $x \geqslant 2$ by

$$\vartheta(x) := \sum_{p \in \mathbb{P}_{\leq x}} \log p,$$

while it is convenient to set $\vartheta(x) := 0$ for $0 < x < 2$, where $\mathbb{P}_{\leq x} := \{p \in \mathbb{P} : p \leq x\} = \mathbb{P} \cap [0, x]$.

3. The second Chebyshev function ψ is defined for $x \geqslant 2$ by

$$\psi(x) := \sum_{n \in [x]} \Lambda(n),$$

while it is convenient to set $\psi(x) := 0$ for $0 < x < 2$.

4. The prime counting function π is defined by

$$\pi(x) := \sum_{p \in \mathbb{P}_{\leq x}} 1 = \#\mathbb{P}_{\leq x},$$

while it is convenient to set $\pi(x) := 0$ for $0 < x < 2$.

Simple relations

Theorem

For $x \geq 2$ we have

$$\vartheta(x) = \pi(x)\log x - \int_2^x \pi(t)\frac{dt}{t},$$

and

$$\pi(x) = \frac{\vartheta(x)}{\log x} + \int_2^x \frac{\vartheta(t)}{t\log^2 t}dt.$$

Proof.

- We have

  $$\pi(x) = \sum_{p\in\mathbb{P}_{\leq x}} 1 = \sum_{1<n\leq x} \mathbf{1}_{\mathbb{P}}(n),$$

  and

  $$\vartheta(x) = \sum_{p\in\mathbb{P}_{\leq x}} \log p = \sum_{1<n\leq x} \mathbf{1}_{\mathbb{P}}(n)\log n.$$

- If $x, y \in \mathbb{R}_+$ with $\lfloor y \rfloor < \lfloor x \rfloor$, and $g \in C^1([y,x])$, then we know

  $$\sum_{y<n\leq x} f(n)g(n) = F(x)g(x) - F(y)g(y) - \int_y^x F(t)g'(t)dt$$

  where $F(t) := \sum_{1\leq n\leq x} f(n)$.

Proof

- Taking $f(n) = \mathbf{1}_{\mathbb{P}}(n)$ and $g(x) = \log x$ with $y = 1$ we obtain

$$\vartheta(x) = \sum_{1<n\leq x} \mathbf{1}_{\mathbb{P}}(n) \log n = \pi(x) \log x - \pi(1) \log 1 - \int_1^x \pi(t) \frac{dt}{t},$$

  which proves the first identity since $\pi(t) = 0$ for $t < 2$.
- Next, let $f(n) = \mathbf{1}_{\mathbb{P}}(n) \log n$ and $g(x) = 1/\log x$ and write

$$\pi(x) = \sum_{3/2<n\leq x} f(n) \frac{1}{\log n}, \quad \vartheta(x) = \sum_{1<n\leq x} f(n)$$

- Using the summation by parts formula with $y = 3/2$ we obtain

$$\pi(x) = \frac{\vartheta(x)}{\log x} - \frac{\vartheta(3/2)}{\log 3/2} + \int_{3/2}^x \frac{\vartheta(t)}{t \log^2 t} dt$$

  which proves the second identity, since $\vartheta(t) = 0$ if $t < 2$. □

Useful pointwise bounds

Lemma

(i) *For all $x \in \mathbb{R}_+$, we have*

$$\vartheta(x) \leqslant \psi(x) \leqslant \vartheta(x) + \pi(\sqrt{x}) \log x.$$

(ii) *For all $x \geqslant 2$ and all $a > 1$, we have*

$$\frac{\vartheta(x)}{\log x} \leqslant \pi(x) \leqslant \frac{a\vartheta(x)}{\log x} + \pi\left(x^{1/a}\right).$$

Proof of (i).

- One may suppose $x \geqslant 2$. We first have

$$\psi(x) - \vartheta(x) = \sum_{p^k \leqslant x} \log p - \sum_{p \leqslant x} \log p = \sum_{p \leqslant \sqrt{x}} \sum_{k=2}^{\lfloor \frac{\log x}{\log p} \rfloor} \log p,$$

so that

$$\psi(x) \geqslant \vartheta(x).$$

Proof

- On the other hand, we have

$$\psi(x) - \vartheta(x) = \sum_{p \leqslant \sqrt{x}} \sum_{k=2}^{\lfloor \frac{\log x}{\log p} \rfloor} \log p \leqslant \sum_{p \leqslant \sqrt{x}} \log p \left\lfloor \frac{\log x}{\log p} \right\rfloor \leqslant \sum_{p \leqslant \sqrt{x}} \log x$$
$$= \pi(\sqrt{x}) \log x. \qquad \square$$

Proof of (ii).

- We have

$$\pi(x) = \sum_{p \leqslant x} 1 = \sum_{p \leqslant x} \frac{\log p}{\log p} \geqslant \frac{1}{\log x} \sum_{p \leqslant x} \log p = \frac{\vartheta(x)}{\log x}.$$

- For $2 \leqslant T < x$, we also have

$$\pi(x) = \sum_{p \leqslant T} 1 + \sum_{T < p \leqslant x} 1 = \pi(T) + \sum_{T < p \leqslant x} \frac{\log p}{\log p} \leqslant \pi(T) + \frac{\vartheta(x)}{\log T}.$$

and the choice of $T = x^{1/a}$ implies the asserted estimate. $\square$

Equivalent forms of the prime number theorem

Theorem

The following relations are logically equivalent:

$$\lim_{x\to\infty} \frac{\pi(x)\log x}{x} = 1. \tag{A}$$

$$\lim_{x\to\infty} \frac{\vartheta(x)}{x} = 1. \tag{B}$$

$$\lim_{x\to\infty} \frac{\psi(x)}{x} = 1. \tag{C}$$

Proof.

- We know that

$$\frac{\vartheta(x)}{x} \leqslant \frac{\psi(x)}{x} \leqslant \frac{\vartheta(x)}{x} + \frac{\pi(\sqrt{x})\log x}{x}.$$

- Also

$$\lim_{x\to\infty} \frac{\pi(\sqrt{x})\log x}{x} = 0$$

- Hence the equivalence between (B) and (C) follows.

Proof

- For every $a > 1$ know that

$$\frac{\vartheta(x)}{x} \leqslant \frac{\pi(x)\log x}{x} \leqslant \frac{a\vartheta(x)}{x} + \frac{\pi\left(x^{1/a}\right)\log x}{x}.$$

- For every $a > 1$ we also know that

$$\lim_{x\to\infty} \frac{\pi\left(x^{1/a}\right)\log x}{x} = 0.$$

- Then

$$\lim_{x\to\infty} \frac{\vartheta(x)}{x} \leqslant \lim_{x\to\infty} \frac{\pi(x)\log x}{x} \leqslant \lim_{x\to\infty} \frac{a\vartheta(x)}{x}.$$

- Since $a > 1$ is arbitrary, we obtain equivalence between (A) and (B).

This completes the proof of the theorem. □

Dirichlet theorem, upper bound

Theorem

For all $x \geqslant 1$, we have

$$\vartheta(x) \leqslant x \log 4.$$

Proof.

- We first prove by induction that, for all $n \in \mathbb{Z}_+$, we have

$$\vartheta(n) \leqslant n \log 4.$$

- This inequality is clearly true for $n \in [3]$. If $n \geqslant 4$ is even, we have

$$\vartheta(n) = \vartheta(n-1) \leqslant (n-1) \log 4 < n \log 4^n.$$

- Suppose now that $n \geqslant 5$ is odd and set $n = 2m + 1$ with $m \in \mathbb{Z}_+$. The idea is to use the fact that the product

$$\prod_{m+1 < p \leqslant 2m+1} p$$

divides the binomial coefficient $\binom{2m+1}{m}$.

Proof

- To see this, observe that $p \in \mathbb{P}$ such that $m+1 < p \leqslant 2m+1$ divides $(2m+1)!$ because of $p \leqslant 2m+1$, but does not divide $m!(m+1)!$ because of $p > m+1$, so that

$$\prod_{m+1<p\leqslant 2m+1} p \quad \text{divides} \quad (2m+1)! = m!(m+1)!\binom{2m+1}{m}$$

  and since the product is coprime to $m!(m+1)!$ the claim follows.
- Taking logarithms, we then obtain

$$\vartheta(2m+1) - \vartheta(m+1) = \sum_{m+1<p\leqslant 2m+1} \log p \leqslant \log\binom{2m+1}{m}.$$

- Using Stirling's formula we have $\binom{2m+1}{m} \leq \frac{2m+1}{m+1}\frac{4^m}{\sqrt{\pi m}} \leq 4^m$, thus

$$\vartheta(2m+1) \leq m\log 4 + \vartheta(m+1) \leq m\log 4 + (m+1)\log 4 = (2m+1)\log 4,$$

  where we have used the induction hypothesis applied to $\vartheta(m+1)$.
- The lemma follows from

$$\vartheta(x) = \vartheta(\lfloor x \rfloor) \leq \lfloor x \rfloor \log 4 \leqslant x \log 4.$$

  The proof is complete. □

Dirichlet theorem, lower bound

Theorem

For all $x \geqslant 1537$, we have

$$\vartheta(x) > \frac{x}{\log 4}.$$

Proof.

- We first notice that the function f defined by

$$f(x) = \lfloor x \rfloor - \left\lfloor \frac{x}{2} \right\rfloor - \left\lfloor \frac{x}{3} \right\rfloor - \left\lfloor \frac{x}{5} \right\rfloor + \left\lfloor \frac{x}{30} \right\rfloor,$$

  is periodic of period 30, since $\lfloor x + n \rfloor = \lfloor x \rfloor + n$ for any $n \in \mathbb{Z}$.
- Moreover, for $x \notin \mathbb{Z}$, we have

$$f(30 - x) = 1 - f(x),$$

  since $\lfloor -x \rfloor = -\lfloor x \rfloor - 1$ for $x \notin \mathbb{Z}$.
- An inspection of its values when $x \in [1, 15)$ allows us to infer that $f(x)$ only takes the values 0 or 1 if $x \notin \mathbb{Z}$.
- Since f is continuous on the right, we also have $f(x) = 0$ or 1 when $x \in \mathbb{Z}$. By periodicity, we infer that $f(x) = 0$ or 1 for all $x \in \mathbb{R}$.

Proof

- Hence we obtain

$$\psi(x) \geqslant \sum_{n\leqslant x} \Lambda(n) f\left(\frac{x}{n}\right) = \sum_{n\leqslant x} \Lambda(n)\left(\left\lfloor\frac{x}{n}\right\rfloor - \left\lfloor\frac{x}{2n}\right\rfloor - \left\lfloor\frac{x}{3n}\right\rfloor - \left\lfloor\frac{x}{5n}\right\rfloor + \left\lfloor\frac{x}{30n}\right\rfloor\right)$$
$$= \sum_{n\leqslant x} \Lambda(n)\left\lfloor\frac{x}{n}\right\rfloor - \sum_{n\leqslant x/2} \Lambda(n)\left\lfloor\frac{x}{2n}\right\rfloor - \sum_{n\leqslant x/3} \Lambda(n)\left\lfloor\frac{x}{3n}\right\rfloor$$
$$- \sum_{n\leqslant x/5} \Lambda(n)\left\lfloor\frac{x}{5n}\right\rfloor + \sum_{n\leqslant x/30} \Lambda(n)\left\lfloor\frac{x}{30n}\right\rfloor,$$

  where we used the fact that $\lfloor x/kn \rfloor = 0$ whenever $n > x/k$ for $k \in \{2, 3, 5, 30\}$.
- By a simplified form of Striling's formula we know for all $x \geq 1$ that

$$\sum_{n\in[x]} \log n = x \log x - x + 1 + R(x), \quad \text{where} \quad 0 \leq R(x) \leq \log x.$$

- Since $\Lambda \star \mathbf{1} = \log$, then we conclude that

$$\sum_{n\in[x]} \log n = \sum_{n\in[x]} \sum_{d|n} \Lambda(d) = \sum_{d\in[x]} \Lambda(d) \sum_{k\in[x/d]} = \sum_{d\in[x]} \Lambda(d)\left\lfloor\frac{x}{d}\right\rfloor.$$

Proof

- Hence for $x \geq 1$, we have

$$\sum_{d\in[x]} \Lambda(d)\left\lfloor \frac{x}{d} \right\rfloor = x\log x - x + 1 + R(x), \quad \text{where} \quad 0 \leq R(x) \leq \log x.$$

- Inserting this bounds to the previous formula we obtain

$$\begin{aligned}
\psi(x) \geqslant & x\log x - x + 1 - \left(\frac{x}{2}\log\frac{x}{2} - \frac{x}{2} + 1 + \log\frac{x}{2}\right) \\
& - \left(\frac{x}{3}\log\frac{x}{3} - \frac{x}{3} + 1 + \log\frac{x}{3}\right) - \left(\frac{x}{5}\log\frac{x}{5} - \frac{x}{5} + 1 + \log\frac{x}{5}\right) \\
& + \left(\frac{x}{30}\log\frac{x}{30} - \frac{x}{30} + 1\right) \\
= & x\log\left(2^{7/15}3^{3/10}5^{1/6}\right) - 3\log x + \log 30 - 1.
\end{aligned}$$

- Using the estimate $\psi(x) \leq \vartheta(x) + \pi(\sqrt{x})\log x$ we obtain the desired lower bound

$$\vartheta(x) \geq x\log\left(2^{7/15}3^{3/10}5^{1/6}\right) - (\sqrt{x} + 3)\log x \geq \frac{x}{\log 4},$$

whenever $x \in \mathbb{R}_+$ is sufficiently large. This completes the proof. □

Bertrand's postulate

Note that

$$\log\left(2^{7/15}3^{3/10}5^{1/6}\right) \approx 0.92129\ldots,$$

which is a very good lower bound. It was sufficient to allow Chebyshev to prove Bertrand's famous postulate.

Theorem

Let $n \in \mathbb{Z}_+$. Then the interval $(n, 2n]$ contains a prime number.

Proof.

- We check numerically the result for $n \in [768]$ and we suppose $n \geqslant 769$.
- Using Chebyshev's upper and lower bounds we deduce

$$\sum_{n<p\leqslant 2n} \log p = \vartheta(2n) - \vartheta(n) > n\left(\frac{2}{\log 4} - \log 4\right) > 0.$$

- This shows that $(n, 2n] \cap \mathbb{P} \neq \emptyset$, which implies the desired result.

□

Chebyshev's estimates for the prime counting function

Theorem

For all $x \geqslant 5$, we have

$$\frac{1}{\log 4}\frac{x}{\log x} < \pi(x) < \left(2 + \frac{1}{\log 4}\right)\frac{x}{\log x}.$$

Proof.

- For all integers $n \in \{5, \ldots, 1537\}$ one can verify that

$$\frac{1}{\log 4}\frac{n+1}{\log n} < \pi(n) < \left(2 + \frac{1}{\log 4}\right)\frac{n}{\log(n+1)}$$

  which implies the desired inequalities for all $x \in [5, 1537]$.
- Suppose that $x \geqslant 1537$. Using the upper bound for $\pi(x)$ with $a = 3/2$ we obtain

$$\pi(x) \leq \frac{3\vartheta(x)}{2\log x} + \pi(x^{2/3}) \leq \frac{3x\log 4}{2\log x} + x^{2/3}.$$

- The inequality $\log x < 3e^{-1.55}x^{1/3}$, valid for all $x \geqslant 1537$, implies that

$$\pi(x) < \frac{3x\log 4}{2\log x} + \frac{3e^{-1.55}x}{\log x} < \left(2 + \frac{1}{\log 4}\right)\frac{x}{\log x},$$

  as required. □

More refined bounds

Theorem

(i) *For all $x \geqslant 25$, we have*

$$\psi(x) < \vartheta(x) + \left(4 + \frac{1}{\log 2}\right)\sqrt{x},$$

$$\pi(x) < \frac{3}{2\log x}\left(\vartheta(x) + \left(2 + \frac{1}{\log 4}\right)x^{2/3}\right).$$

(ii) *For all $x \geqslant 1$, we have*

$$\psi(x) < 2x.$$

(iii) *Let p_n be the nth prime number. Then, for all integers $n \geqslant 3$, we have*

$$\frac{1}{2}n\log n < p_n < 2n\log n.$$

The lower bound for p_n immediately implies that

$$\sum_{p\in\mathbb{P}}\frac{1}{p} = \sum_{n\in\mathbb{Z}_+}\frac{1}{p_n} = \infty.$$

Proof

Proof of (i) and (ii).

- We know that
$$\vartheta(x) \leqslant \psi(x) \leqslant \vartheta(x) + \pi(\sqrt{x}) \log x.$$
- From the previous theorem
$$\pi(x) < \left(2 + \frac{1}{\log 4}\right) \frac{x}{\log x}.$$
- Hence
$$\psi(x) \leqslant \vartheta(x) + \pi(\sqrt{x}) \log x \leq \vartheta(x) + 2\left(2 + \frac{1}{\log 4}\right)\sqrt{x}.$$
- The inequality $\psi(x) < 2x$ is first numerically checked for all integers $n \in [1, 100]$ and for $x > 100$ we invoke the previous bound.
- We know that
$$\frac{\vartheta(x)}{\log x} \leqslant \pi(x) \leqslant \frac{3\vartheta(x)}{2\log x} + \pi\left(x^{2/3}\right)$$
- Again by the previous theorem we conclude that
$$\pi(x) < \frac{3}{2\log x}\left(\vartheta(x) + \left(2 + \frac{1}{\log 4}\right) x^{2/3}\right). \qquad \square$$

Proof

Proof of (iii).

- We first check the inequalities for $n \in \{3, \ldots, 337\}$. Suppose $n \geqslant 338$ so that $p_n \geqslant 2273$. The proof rests on the fact that $\pi(p_n) = n$.
- For the lower bound, we use the second upper bound of (i) above, the inequality $x^{2/3} < x/13$ valid for all $x \geqslant 2273$ and the trivial inequality $p_n > n$, which imply the desired lower bound

$$n = \pi(p_n) < \frac{3p_n}{2\log p_n}\left\{\log 4 + \frac{1}{13}\left(2 + \frac{1}{\log 4}\right)\right\} < \frac{2p_n}{\log p_n} < \frac{2p_n}{\log n}.$$

- For the upper bound, we use the inequality $\frac{\log x}{x^{1-\log 2}} < \frac{1}{\log 4}$ valid for all $x \geqslant 2273$ (applied with $x = p_n$), and the lower bound for $\pi(x)$ giving

$$\frac{\log p_n}{p_n^{1-\log 2}} < \frac{1}{\log 4} < \frac{\pi(p_n)\log p_n}{p_n} = \frac{n\log p_n}{p_n}$$

- Hence $p_n < n^{1/\log 2}$, and $\log p_n < \frac{\log n}{\log 2}$. Thus, the upper bound follows

$$n = \pi(p_n) > \frac{1}{\log 4}\frac{p_n}{\log p_n} > \frac{p_n}{2\log n}. \qquad \square$$

Mertens first theorem

Theorem

For all $x \geqslant 2$, we have

$$\sum_{p \leqslant x} \frac{\log p}{p} = \log x + O(1).$$

Proof.

- We first notice that

$$\sum_{d \leqslant x} \frac{\Lambda(d)}{d} = \sum_{p \leqslant x} \frac{\log p}{p} + \sum_{\substack{p^k \leqslant x \\ k \geqslant 2}} \frac{\log p}{p^k}.$$

- The second sum is

$$\sum_{\substack{p^k \leqslant x \\ k \geqslant 2}} \frac{\log p}{p^k} \leqslant \sum_{p \leqslant \sqrt{x}} \log p \sum_{k=2}^{\infty} \frac{1}{p^k} = \sum_{p \leqslant \sqrt{x}} \frac{\log p}{p(p-1)} = O(1).$$

Proof

- We have proved that
$$\sum_{p\leqslant x}\frac{\log p}{p}=\sum_{d\leqslant x}\frac{\Lambda(d)}{d}+O(1).$$
- We know that for $x\geq 1$, we have
$$\sum_{d\in[x]}\Lambda(d)\left\lfloor\frac{x}{d}\right\rfloor=x\log x-x+1+R(x),\quad\text{where}\quad 0\leq R(x)\leq\log x.$$
- Hence, we have
$$x\log x+O(x)=\sum_{d\in[x]}\Lambda(d)\left\lfloor\frac{x}{d}\right\rfloor=x\sum_{d\in[x]}\frac{\Lambda(d)}{d}+O(\psi(x)).$$
- Since $\psi(x)<2x$, we obtain
$$x\sum_{d\in[x]}\frac{\Lambda(d)}{d}=x\log x+O(x)$$
- Therefore the desired conclusion follows
$$\sum_{d\leqslant x}\frac{\Lambda(d)}{d}=\log x+O(1).\qquad\square$$

Mertens second theorem

Theorem

There exists a constant $B \in \mathbb{R}_+$ such that for all $x \geq 2$ we have

$$\sum_{p \leq x} \frac{1}{p} = \log\log x + B + O\left(\frac{1}{\log x}\right).$$

Proof.

- We can write

$$\sum_{p \leq x} \frac{1}{p} = \sum_{p \leq x} \frac{\log p}{p} \frac{1}{\log p} = \sum_{2 \leq n \leq x} f(n) g(n),$$

  where $f(n) = \mathbf{1}_{\mathbb{P}}(n) \frac{\log n}{n}$ and $g(t) = \frac{1}{\log t}$
- Let

$$F(t) = \sum_{n \leq t} f(n) = \sum_{p \leq t} \frac{\log p}{p}.$$

- Then $F(t) = 0$ for $t < 2$. By the first Mertens theorem we have

$$F(t) = \log t + r(t), \quad \text{where} \quad r(t) = O(1).$$

Proof

- Therefore, the integral $\int_2^\infty \frac{r(t)}{t(\log t)^2}dt$ converges absolutely, and

$$\int_x^\infty \frac{r(t)dt}{t(\log t)^2} = O\left(\frac{1}{\log x}\right).$$

- By partial summation, we obtain

$$\begin{aligned}
\sum_{p\leq x}\frac{1}{p} =& \sum_{n\leq x} f(n)g(n) = F(x)g(x) - \int_2^x F(t)g'(t)dt \\
=& \frac{\log x + r(x)}{\log x} + \int_2^x \frac{\log t + r(t)}{t(\log t)^2}dt \\
=& 1 + O\left(\frac{1}{\log x}\right) + \int_2^x \frac{dt}{t\log t} + \int_2^x \frac{r(t)}{t(\log t)^2}dt \\
=& \log\log x + 1 - \log\log 2 + \int_2^\infty \frac{r(t)}{t(\log t)^2}dt - \int_x^\infty \frac{r(t)}{t(\log t)^2}dt + O\left(\frac{1}{\log x}\right) \\
=& \log\log x + B + O\left(\frac{1}{\log x}\right),
\end{aligned}$$

where

$$B = 1 - \log\log 2 + \int_2^\infty \frac{r(t)}{t(\log t)^2}dt. \qquad \square$$

Mertens third theorem

Theorem

For all $x \geqslant e$, we have

$$\prod_{p \leqslant x} \left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}}{\log x} \left(1 + O\left(\frac{1}{\log x}\right)\right).$$

Remark: (Exercise)

In the second Mertens theorem one can show that

$$B := \gamma + \sum_{p \in \mathbb{P}} \left(\log\left(1 - \frac{1}{p}\right) + \frac{1}{p}\right) \approx 0.2614972128\ldots,$$

which is called the Mertens constant. We will use this fact below.

Proof.

- We have

$$\begin{aligned}\log \prod_{p \leqslant x} \left(1 - \frac{1}{p}\right)^{-1} &= \sum_{p \leqslant x} \log\left(1 - \frac{1}{p}\right)^{-1} \\ &= \sum_{p \leqslant x} \frac{1}{p} + \sum_{p \leqslant x} \left(\log\left(1 - \frac{1}{p}\right)^{-1} - \frac{1}{p}\right).\end{aligned}$$

Proof

- Using the Taylor series expansion we conclude that

$$\log\left(1-\frac{1}{p}\right)^{-1}-\frac{1}{p}\leqslant\frac{1}{2p(p-1)}\leqslant\frac{1}{p^2}.$$

- By the last bound we have $\sum_{p>x}\frac{1}{p^2}=O\left(\frac{1}{x}\right)$.
- By the second Mertens theorem we may write

$$\begin{aligned}\log\prod_{p\leqslant x}\left(1-\frac{1}{p}\right)^{-1}&=\sum_{p\leqslant x}\frac{1}{p}-\sum_{p\in\mathbb{P}}\left(\log\left(1-\frac{1}{p}\right)+\frac{1}{p}\right)+O\left(\frac{1}{x}\right)\\&=\log\log x+\gamma+O\left(\frac{1}{\log x}\right).\end{aligned}$$

- Using $e^h=1+O(h)$ for all $h\to 0$ finally gives

$$\prod_{p\leqslant x}\left(1-\frac{1}{p}\right)^{-1}=e^{\gamma}\log x\left(1+O\left(\frac{1}{\log x}\right)\right),$$

which is easily seen to be equivalent to the desired formula. □

Improved estimates for the Euler totient function

Theorem

Let $\varepsilon > 0$ be given. Then there exists an $N_\varepsilon \in \mathbb{Z}_+$ such that

$$\varphi(n) \geq (1-\varepsilon)\frac{e^{-\gamma}n}{\log\log n} \quad \textit{for all} \quad n \geq N_\varepsilon.$$

Proof.

- Take any $n > 1$ and write
$$\frac{\varphi(n)}{n} = \prod_{p|n}\left(1-\frac{1}{p}\right) = P_1(n)P_2(n),$$
where
$$P_1(n) = \prod_{\substack{p|n \\ p\leq \log n}}\left(1-\frac{1}{p}\right), \quad \text{and} \quad P_2(n) = \prod_{\substack{p|n \\ p>\log n}}\left(1-\frac{1}{p}\right).$$
- Then
$$P_2(n) > \prod_{\substack{p|n \\ p>\log n}}\left(1-\frac{1}{\log n}\right) = \left(1-\frac{1}{\log n}\right)^{f(n)},$$
where $f(n)$ is the number of primes which divide n and exceed $\log n$.

Proof

- Since
$$n \geq \prod_{p|n} p > \prod_{\substack{p|n \\ p>\log n}} p \geq (\log n)^{f(n)}$$
we find $\log n > f(n) \log\log n$, so $f(n) < \log n / \log\log n$.
- Since $1 - (1/\log n) < 1$, we obtain
$$P_2(n) > \left(1 - \frac{1}{\log n}\right)^{\log n/\log\log n} = \left(\left(1 - \frac{1}{\log n}\right)^{\log n}\right)^{1/\log\log n}.$$
- Now $\lim_{u\to\infty} \left(1 - \frac{1}{u}\right)^u = e^{-1}$, so we can conclude that
$$P_2(n) > 1 + o(1) \quad \text{as} \quad n \to \infty.$$
- Therefore, there exists $N_\varepsilon \in \mathbb{Z}_+$ such that for every $n \geq N(\varepsilon)$ we have
$$\frac{\varphi(n)}{n} = P_1(n)P_2(n) > (1 + o(1)) \prod_{\substack{p|n \\ p\leq\log n}} \left(1 - \frac{1}{p}\right)$$
$$\geq (1 + o(1)) \prod_{p\leq\log n} \left(1 - \frac{1}{p}\right) = (1 + o(1)) \frac{e^{-\gamma}}{\log\log n} \geq (1 - \varepsilon) \frac{e^{-\gamma}}{\log\log n}.$$
- This completes the proof. □

Improved bounds for the divisor function

Theorem

Let $\varepsilon > 0$ be given. Then there exists an $N_\varepsilon \in \mathbb{Z}_+$ such that

$$\tau(n) < 2^{(1+\varepsilon)\frac{\log n}{\log\log n}} = n^{\frac{(1+\varepsilon)\log 2}{\log\log n}} \quad \textit{for all} \quad n \geq N_\varepsilon.$$

In particular, for every $\delta > 0$, we obtain

$$\tau(n) = o\left(n^\delta\right).$$

Proof.

- Write $n = p_1^{a_i} \cdots p_k^{a_k}$, so that $\tau(n) = \prod_{i=1}^{k} (a_i + 1)$.
- We split the product into two parts, separating those prime divisors $< f(n)$ from those $\geq f(n)$, where $f(n)$ will be specified later.
- Then $\tau(n) = P_1(n) P_2(n)$ where

$$P_1(n) = \prod_{p_i < f(n)} (a_i + 1),$$

$$P_2(n) = \prod_{p_i \geq f(n)} (a_i + 1).$$

Proof

- In the product $P_2(n)$ we use the inequality $(a+1) \leq 2^a$ to obtain $P_2(n) \leq 2^{S(n)}$, where
$$S(n) = \sum_{\substack{i=1 \\ p_i \geq f(n)}}^{k} a_i.$$
- Now
$$n = \prod_{i=1}^{k} p_i^{a_i} \geq \prod_{p_i \geq f(n)} p_i^{a_i} \geq \prod_{p_i \geq f(n)} f(n)^{a_i} = f(n)^{S(n)},$$
hence
$$\log n \geq S(n) \log f(n) \quad \Longleftrightarrow \quad S(n) \leq \frac{\log n}{\log f(n)}.$$
- This gives us
$$P_2(n) \leq 2^{\log n / \log f(n)}.$$

Proof

- To estimate $P_1(n)$ we write

$$P_1(n) = \exp\left\{\sum_{p_i<f(n)} \log(a_i+1)\right\}$$

- We have $n \geq p_i^{a_i} \geq 2^{a_i}$, hence

$$\log n \geq a_i \log 2 \quad \Longleftrightarrow \quad a_i \leq \log n / \log 2.$$

  Therefore, for sufficiently large $n \in \mathbb{Z}_+$, we have

$$1 + a_i \leq 1 + \frac{\log n}{\log 2} < (\log n)^2$$

- Thus $n \geq N_0$ for some $N_0 \in \mathbb{Z}_+$ implies

$$\log(1+a_i) < \log(\log n)^2 = 2\log\log n.$$

- This gives us

$$P_1(n) < \exp\left(2\log\log n \sum_{p_i<f(n)} 1\right) \leq \exp(2\pi(f(n))\log\log n).$$

Proof

- Setting $c = 6/\log 2$ and using $\pi(x) < 3x/\log x$, we obtain

$$P_1(n) < \exp\left(\frac{6f(n)\log\log n}{\log f(n)}\right) = 2^{\frac{cf(n)\log\log n}{\log f(n)}}.$$

- Then we deduce $\tau(n) = P_1(n)P_2(n) < 2^{g(n)}$, where

$$g(n) = \frac{\log n + cf(n)\log\log n}{\log f(n)} = \frac{\log n}{\log\log n}\frac{1 + \frac{cf(n)\log\log n}{\log n}}{\frac{\log f(n)}{\log\log n}}.$$

- Now we choose $f(n)$ to make $f(n)\log\log n/\log n \to 0$ and also to make $\log f(n)/\log\log n \to 1$ as $n \to \infty$. For this it suffices to take

$$f(n) = \frac{\log n}{(\log\log n)^2}.$$

- Therefore, there exists $N_\varepsilon \in \mathbb{Z}_+$ such that for every $n \geq N(\varepsilon)$ we have

$$g(n) = \frac{\log n}{\log\log n}\frac{1 + o(1)}{1 + o(1)} = \frac{\log n}{\log\log n}(1 + o(1)) < (1 + \varepsilon)\frac{\log n}{\log\log n}.$$

- This proves the theorem. □