Introduction to Number Theory Chapter 8 The Distibution of Primes

Robert C. Vaughan

April 16, 2025

- Gauss suggested that a good approximation to $\pi(x)$, the number of primes not exceeding x, is

$$\mathrm{li}(x) = \int_2^x \frac{dt}{\log t}.$$

- Gauss suggested that a good approximation to $\pi(x)$, the number of primes not exceeding x, is

$$\mathrm{li}(x) = \int_2^x \frac{dt}{\log t}.$$

- He also carried out some calculations for $x \leq 1000$. Today we have much more extensive calculations.

x	$\pi(x)$	$\mathrm{li}(x)$
10^4	1229	1245
10^5	9592	9628
10^6	78498	78626
10^7	664579	664917
10^8	5761455	5762208
10^9	50847534	50849233
10^{10}	455052511	455055613
10^{11}	4118054813	4118066399
10^{12}	37607912018	37607950279
10^{13}	346065536839	346065645809
10^{14}	3204941750802	3204942065690
10^{15}	29844570422669	29844571475286
10^{16}	279238341033925	279238344248555
10^{17}	2623557157654233	2623557165610820
10^{18}	24739954287740860	24739954309690413
10^{19}	234057667276344607	234057667376222382
10^{20}	2220819602560918840	2220819602783663483
10^{21}	21127269486018731928	21127269486616126182
10^{22}	201467286689315906290	201467286691248261498

- This table has been extended out to at least 10^{27}. So is

$$\pi(x) < \mathrm{li}(x)$$

always true?

- This table has been extended out to at least 10^{27}. So is

$$\pi(x) < \mathrm{li}(x)$$

always true?

- No! Littlewood in 1914 showed that there are infinitely many values of x for which

$$\pi(x) > \mathrm{li}(x)$$

and now we believe that the first sign change occurs when

$$x \approx 1.387162 \times 10^{316}$$

well beyond what can be calculated directly.

- This table has been extended out to at least 10^{27}. So is

$$\pi(x) < \mathrm{li}(x)$$

always true?

- No! Littlewood in 1914 showed that there are infinitely many values of x for which

$$\pi(x) > \mathrm{li}(x)$$

and now we believe that the first sign change occurs when

$$x \approx 1.387162 \times 10^{316}$$

well beyond what can be calculated directly.

- For many years it was only known that the first sign change in $\pi(x) - \mathrm{li}(x)$ occurs for *some* x satisfying

$$x < 10^{10^{10^{964}}}.$$

- This table has been extended out to at least 10^{27}. So is

$$\pi(x) < \mathrm{li}(x)$$

always true?

- No! Littlewood in 1914 showed that there are infinitely many values of x for which

$$\pi(x) > \mathrm{li}(x)$$

and now we believe that the first sign change occurs when

$$x \approx 1.387162 \times 10^{316}$$

well beyond what can be calculated directly.

- For many years it was only known that the first sign change in $\pi(x) - \mathrm{li}(x)$ occurs for *some* x satisfying

$$x < 10^{10^{10^{964}}}.$$

- This number was computed by Skewes and G. H. Hardy once wrote that this is probably the largest number which has ever had any *practical* (my emphasis) value!

- The strongest results we know about the distribution of primes use complex analytic methods.

- The strongest results we know about the distribution of primes use complex analytic methods.
- However there are some very useful and basic results that can be established elementarily.

- The strongest results we know about the distribution of primes use complex analytic methods.
- However there are some very useful and basic results that can be established elementarily.
- Many expositions of the results we are going to describe use nothing more than properties of binomial coefficients, but it is good to start to get the flavour of more sophisticated methods even though here they could be interpreted as just properties of binomial coefficients.

- We start by introducing **The von Mangold function.** This is defined by

$$\Lambda(n) = \begin{cases} 0 & \text{if } p_1 p_2 | n \text{ with } p_1 \neq p_2, \\ \log p & \text{if } n = p^k. \end{cases}$$

- We start by introducing **The von Mangold function.** This is defined by

$$\Lambda(n) = \begin{cases} 0 & \text{if } p_1 p_2 | n \text{ with } p_1 \neq p_2, \\ \log p & \text{if } n = p^k. \end{cases}$$

- The interesting thing is that the support of Λ is on the prime powers, the higher powers are quite rare, at most $\sqrt{x}$ of them not exceeding x.

- We start by introducing **The von Mangold function.** This is defined by

$$\Lambda(n) = \begin{cases} 0 & \text{if } p_1 p_2 | n \text{ with } p_1 \neq p_2, \\ \log p & \text{if } n = p^k. \end{cases}$$

- The interesting thing is that the support of Λ is on the prime powers, the higher powers are quite rare, at most $\sqrt{x}$ of them not exceeding x.
- This function is definitely not multiplicative, since $\Lambda(1) = 0$.

- However the von Mangoldt function does satisfy some interesting relationships.

Lemma 1

Let $n \in \mathbb{N}$. *Then* $\sum_{m|n} \Lambda(m) = \log n$.

- However the von Mangoldt function does satisfy some interesting relationships.

Lemma 1

Let $n \in \mathbb{N}$. Then $\sum_{m|n} \Lambda(m) = \log n$.

- The proof is a simple counting argument.

Proof.

Write $n = p_1^{k_1} \dots p_r^{k_r}$ with the p_j distinct. Then for a non-zero contribution to the sum we have $m = p_s^{j_s}$ for some s with $1 \le s \le r$ and j_s with $1 \le j_s \le k_s$. Thus the sum is

$$\sum_{s=1}^{r} \sum_{j_s=1}^{k_s} \log p_s = \log n.$$

□

- We need to know something about the average of $\log n$.

Lemma 2 (Stirling)

Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then

$$\sum_{n \leq X} \log n = X(\log X - 1) + O(\log X).$$

- This can be thought of as the logarithm of Stirling's formula for $\lfloor X \rfloor !$.

Proof.

We have

$$\begin{aligned}\sum_{n\le X} \log n &= \sum_{n\le X} \left(\log X - \int_n^X \frac{dt}{t} \right) \\ &= \lfloor X \rfloor \log X - \int_1^X \frac{\lfloor t \rfloor}{t} dt \\ &= X(\log X - 1) + \int_1^X \frac{t - \lfloor t \rfloor}{t} dt + O(\log X).\end{aligned}$$

□

- Now we can say something about averages of the von Mangoldt function.

Theorem 3

Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then

$$\sum_{m \leq X} \Lambda(m) \left\lfloor \frac{X}{m} \right\rfloor = X(\log X - 1) + O(\log X).$$

- Now we can say something about averages of the von Mangoldt function.

Theorem 3

Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then

$$\sum_{m \leq X} \Lambda(m) \left\lfloor \frac{X}{m} \right\rfloor = X(\log X - 1) + O(\log X).$$

- This is easy

Proof.

We substitute from the first lemma into the second. Thus

$$\sum_{n \leq X} \sum_{m|n} \Lambda(m) = X(\log X - 1) + O(\log X).$$

Now we interchange the order in the double sum and count the number of multiples of m not exceeding X. □

- At this stage it is necessary to introduce some of the fundamental counting functions of prime number theory.

- At this stage it is necessary to introduce some of the fundamental counting functions of prime number theory.
- For $X \geq 0$ we define

$$\psi(X) = \sum_{n \leq X} \Lambda(n),$$

$$\vartheta(X) = \sum_{p \leq X} \log p,$$

$$\pi(X) = \sum_{p \leq X} 1.$$

- The following theorem shows the close relationship between these three functions.

Theorem 4

Suppose that $X \geq 2$. Then

$$\psi(X) = \sum_k \vartheta(X^{1/k}),$$

$$\vartheta(X) = \sum_k \mu(k)\psi(X^{1/k}),$$

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

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

Note that each of these functions are 0 when $X < 2$, so the sums are all finite.

- We prove

$$\psi(X) = \sum_k \vartheta(X^{1/k}),$$

$$\vartheta(X) = \sum_k \mu(k)\psi(X^{1/k}),$$

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

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

- We prove

$$\psi(X) = \sum_k \vartheta(X^{1/k}),$$

$$\vartheta(X) = \sum_k \mu(k)\psi(X^{1/k}),$$

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

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

- By the definition of Λ we have

$$\psi(X) = \sum_k \sum_{p \le X^{1/k}} \log p = \sum_k \vartheta(X^{1/k}).$$

- We prove

$$\psi(X) = \sum_k \vartheta(X^{1/k}),$$

$$\vartheta(X) = \sum_k \mu(k)\psi(X^{1/k}),$$

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

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

- By the definition of Λ we have

$$\psi(X) = \sum_k \sum_{p \le X^{1/k}} \log p = \sum_k \vartheta(X^{1/k}).$$

- Hence we have

$$\sum_k \mu(k)\psi(X^{1/k}) = \sum_k \mu(k) \sum_l \vartheta(X^{1/(kl)}).$$

- Collecting together the terms for which $kl = m$ for a given m this becomes

$$\sum_m \vartheta(X^{1/m}) \sum_{k|m} \mu(k) = \vartheta(X).$$

- Collecting together the terms for which $kl = m$ for a given m this becomes

$$\sum_m \vartheta(X^{1/m}) \sum_{k|m} \mu(k) = \vartheta(X).$$

- We also have

$$\begin{aligned}\pi(X) &= \sum_{p \le X} (\log p) \left(\frac{1}{\log X} + \int_p^X \frac{dt}{t \log^2 t} \right) \\ &= \frac{\vartheta(X)}{\log X} + \int_2^X \frac{\vartheta(t)}{t \log^2 t} dt.\end{aligned}$$

- Collecting together the terms for which $kl = m$ for a given m this becomes

$$\sum_m \vartheta(X^{1/m}) \sum_{k|m} \mu(k) = \vartheta(X).$$

- We also have

$$\begin{aligned}\pi(X) &= \sum_{p \le X} (\log p) \left(\frac{1}{\log X} + \int_p^X \frac{dt}{t \log^2 t} \right) \\ &= \frac{\vartheta(X)}{\log X} + \int_2^X \frac{\vartheta(t)}{t \log^2 t} dt.\end{aligned}$$

- The final identity is similar.

$$\vartheta(X) = \sum_{p \le X} \log X - \sum_{p \le X} \int_p^X \frac{dt}{t}$$

etcetera.

- Now we come to a series of theorems which are still used frequently.

Theorem 5 (Chebyshev)

There are positive constants C_1 and C_2 such that for each $X \in \mathbb{R}$ with $X \geq 2$ we have

$$C_1 X < \psi(X) < C_2 X.$$

- Now we come to a series of theorems which are still used frequently.

Theorem 5 (Chebyshev)

There are positive constants C_1 and C_2 such that for each $X \in \mathbb{R}$ with $X \geq 2$ we have

$$C_1 X < \psi(X) < C_2 X.$$

- Proof. For any $\theta \in \mathbb{R}$ let

$$f(\theta) = \lfloor \theta \rfloor - 2 \left\lfloor \frac{\theta}{2} \right\rfloor.$$

Then f is periodic with period 2 and

$$f(\theta) = \begin{cases} 0 & (0 \leq \theta < 1), \\ 1 & (1 \leq \theta < 2). \end{cases}$$

- Hence

$$\psi(X) \geq \sum_{n \leq X} \Lambda(n) f(X/n)$$

$$= \sum_{n \leq X} \Lambda(n) \left\lfloor \frac{X}{n} \right\rfloor - 2 \sum_{n \leq X/2} \Lambda(n) \left\lfloor \frac{X/2}{n} \right\rfloor.$$

- Hence

$$\psi(X) \geq \sum_{n \leq X} \Lambda(n) f(X/n)$$
$$= \sum_{n \leq X} \Lambda(n) \left\lfloor \frac{X}{n} \right\rfloor - 2 \sum_{n \leq X/2} \Lambda(n) \left\lfloor \frac{X/2}{n} \right\rfloor.$$

- Here we used the fact that there is no contribution to the second sum when $X/2 < n \leq X$.

- Hence

$$\psi(X) \geq \sum_{n \leq X} \Lambda(n) f(X/n)$$
$$= \sum_{n \leq X} \Lambda(n) \left\lfloor \frac{X}{n} \right\rfloor - 2 \sum_{n \leq X/2} \Lambda(n) \left\lfloor \frac{X/2}{n} \right\rfloor.$$

- Here we used the fact that there is no contribution to the second sum when $X/2 < n \leq X$.
- Now we apply Theorem 3 and obtain for $x \geq 4$

$$X(\log X - 1) - 2\frac{X}{2}\left(\log \frac{X}{2} - 1\right) + O(\log X)$$
$$= X \log 2 + O(\log X).$$

- Hence

$$\psi(X) \geq \sum_{n \leq X} \Lambda(n) f(X/n)$$
$$= \sum_{n \leq X} \Lambda(n) \left\lfloor \frac{X}{n} \right\rfloor - 2 \sum_{n \leq X/2} \Lambda(n) \left\lfloor \frac{X/2}{n} \right\rfloor.$$

- Here we used the fact that there is no contribution to the second sum when $X/2 < n \leq X$.
- Now we apply Theorem 3 and obtain for $x \geq 4$

$$X(\log X - 1) - 2\frac{X}{2}\left(\log \frac{X}{2} - 1\right) + O(\log X)$$
$$= X \log 2 + O(\log X).$$

- This establishes the first inequality of the theorem for all $X > C$ for some positive constant C. Since $\psi(X) \geq \log 2$ for all $X \geq 2$ the conclusion follows if C_1 is small enough.

- We also have, for $X \geq 4$,

$$\psi(X) - \psi(X/2) \leq \sum_{n \leq X} \Lambda(n) f(X/n)$$

and we have already seen that this is

$$X \log 2 + O(\log X).$$

- We also have, for $X \geq 4$,

$$\psi(X) - \psi(X/2) \leq \sum_{n \leq X} \Lambda(n) f(X/n)$$

and we have already seen that this is

$$X \log 2 + O(\log X).$$

- Hence for some positive constant C we have, for all $X > 0$,

$$\psi(X) - \psi(X/2) \leq CX.$$

Hence, for any $k \geq 0$,

$$\psi(X2^{-k}) - \psi(X2^{-k-1}) < CX2^{-k}.$$

- We also have, for $X \geq 4$,

$$\psi(X) - \psi(X/2) \leq \sum_{n \leq X} \Lambda(n) f(X/n)$$

and we have already seen that this is

$$X \log 2 + O(\log X).$$

- Hence for some positive constant C we have, for all $X > 0$,

$$\psi(X) - \psi(X/2) \leq CX.$$

Hence, for any $k \geq 0$,

$$\psi(X2^{-k}) - \psi(X2^{-k-1}) < CX2^{-k}.$$

- Summing over all k gives the desired upper bound.

- The following now follow easily from the last couple of theorems.

Corollary 6 (Chebyshev)

There are positive constants C_3, C_4, C_5, C_6 such that for every $X \geq 2$ we have

$$C_3 X < \vartheta(X) < C_4 X,$$
$$\frac{C_5 X}{\log X} < \pi(X) < \frac{C_6 X}{\log X}.$$

- It is also possible to establish a more precise version of Euler's result on the primes.

Theorem 7 (Mertens)

There is a constant B such that whenever $X \geq 2$ we have

$$\sum_{n \leq X} \frac{\Lambda(n)}{n} = \log X + O(1),$$

$$\sum_{p \leq X} \frac{\log p}{p} = \log X + O(1),$$

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

- Proof By Theorem 3 we have

$$\sum_{m \le X} \Lambda(m) \left\lfloor \frac{X}{m} \right\rfloor = X(\log X - 1) + O(\log X).$$

- Proof By Theorem 3 we have

$$\sum_{m \le X} \Lambda(m) \left\lfloor \frac{X}{m} \right\rfloor = X(\log X - 1) + O(\log X).$$

- The left hand side is

$$X \sum_{m \le X} \frac{\Lambda(m)}{m} + O(\psi(X)).$$

- Proof By Theorem 3 we have

$$\sum_{m \le X} \Lambda(m) \left\lfloor \frac{X}{m} \right\rfloor = X(\log X - 1) + O(\log X).$$

- The left hand side is

$$X \sum_{m \le X} \frac{\Lambda(m)}{m} + O(\psi(X)).$$

- Hence by Cheyshev's theorem we have

$$X \sum_{m \le X} \frac{\Lambda(m)}{m} = X \log X + O(X).$$

- Proof By Theorem 3 we have

$$\sum_{m \le X} \Lambda(m) \left\lfloor \frac{X}{m} \right\rfloor = X(\log X - 1) + O(\log X).$$

- The left hand side is

$$X \sum_{m \le X} \frac{\Lambda(m)}{m} + O(\psi(X)).$$

- Hence by Cheyshev's theorem we have

$$X \sum_{m \le X} \frac{\Lambda(m)}{m} = X \log X + O(X).$$

- Dividing by X gives the first result.

- We also have

$$\sum_{m\leq X}\frac{\Lambda(m)}{m}=\sum_{k}\sum_{p^k\leq X}\frac{\log p}{p^k}.$$

- We also have

$$\sum_{m\le X}\frac{\Lambda(m)}{m}=\sum_{k}\sum_{p^k\le X}\frac{\log p}{p^k}.$$

- The terms with $k\ge 2$ contribute

$$\le\sum_{p}\sum_{k\ge 2}\frac{\log p}{p^k}\le\sum_{n=2}^{\infty}\frac{\log n}{n(n-1)}$$

which is convergent, and this gives the second expression.

- Finally we can see that

$$\sum_{p\le X}\frac{1}{p}=\sum_{p\le X}\frac{\log p}{p}\left(\frac{1}{\log X}+\int_p^X\frac{dt}{t\log^2 t}\right)$$
$$=\frac{1}{\log X}\sum_{p\le X}\frac{\log p}{p}+\int_2^X\sum_{p\le t}\frac{\log p}{p}\frac{dt}{t\log^2 t}.$$

- Finally we can see that

$$\sum_{p\le X}\frac{1}{p}=\sum_{p\le X}\frac{\log p}{p}\left(\frac{1}{\log X}+\int_p^X\frac{dt}{t\log^2 t}\right)$$
$$=\frac{1}{\log X}\sum_{p\le X}\frac{\log p}{p}+\int_2^X\sum_{p\le t}\frac{\log p}{p}\frac{dt}{t\log^2 t}.$$

- $E(t)=\sum_{p\le t}\frac{\log p}{p}-\log t$ so that by the second part of the theorem we have $E(t)\ll 1$.

- Finally we can see that

$$\sum_{p\le X}\frac{1}{p}=\sum_{p\le X}\frac{\log p}{p}\left(\frac{1}{\log X}+\int_p^X\frac{dt}{t\log^2 t}\right)$$
$$=\frac{1}{\log X}\sum_{p\le X}\frac{\log p}{p}+\int_2^X\sum_{p\le t}\frac{\log p}{p}\frac{dt}{t\log^2 t}.$$

- $E(t)=\sum_{p\le t}\frac{\log p}{p}-\log t$ so that by the second part of the theorem we have $E(t)\ll 1$.
- Then the above is

$$=\frac{\log X+E(X)}{\log X}+\int_2^X\frac{\log t+E(t)}{t\log^2 t}dt$$
$$=\log\log X+1-\log\log 2+\int_2^\infty\frac{E(t)}{t\log^2 t}dt$$
$$+\frac{E(X)}{\log X}-\int_X^\infty\frac{E(t)}{t\log^2 t}dt.$$

- Finally we can see that

$$\sum_{p\le X}\frac{1}{p}=\sum_{p\le X}\frac{\log p}{p}\left(\frac{1}{\log X}+\int_p^X\frac{dt}{t\log^2 t}\right)$$
$$=\frac{1}{\log X}\sum_{p\le X}\frac{\log p}{p}+\int_2^X\sum_{p\le t}\frac{\log p}{p}\frac{dt}{t\log^2 t}.$$

- $E(t)=\sum_{p\le t}\frac{\log p}{p}-\log t$ so that by the second part of the theorem we have $E(t)\ll 1$.
- Then the above is

$$=\frac{\log X+E(X)}{\log X}+\int_2^X\frac{\log t+E(t)}{t\log^2 t}dt$$
$$=\log\log X+1-\log\log 2+\int_2^\infty\frac{E(t)}{t\log^2 t}dt$$
$$+\frac{E(X)}{\log X}-\int_X^\infty\frac{E(t)}{t\log^2 t}dt.$$

- The first integral converges and the last two terms are $\ll\frac{1}{\log X}$.

- Another theorem which can be deduced is the following.

Theorem 8 (Mertens)

We have

$$\prod_{p \le X} \left(1 - \frac{1}{p}\right)^{-1} = e^C \log X + O(1).$$

- There is an interesting application of the above which lead to some important developments.

- There is an interesting application of the above which lead to some important developments.
- As a companion to the definition of a multiplicative function we have **Definition.** An $f \in \mathcal{A}$ is **additive** when it satisfies $f(mn) = f(m) + f(n)$ whenever $(m, n) = 1$.
- Now we introduce two further functions. **Definition.** We define $\omega(n)$ to be the number of different prime factors of n and $\Omega(n)$ to be the total number of prime factors of n.

- There is an interesting application of the above which lead to some important developments.
- As a companion to the definition of a multiplicative function we have **Definition.** An $f \in \mathcal{A}$ is **additive** when it satisfies $f(mn) = f(m) + f(n)$ whenever $(m, n) = 1$.
- Now we introduce two further functions. **Definition.** We define $\omega(n)$ to be the number of different prime factors of n and $\Omega(n)$ to be the total number of prime factors of n.
- **Example.** We have $360 = 2^3 3^2 5$ so that $\omega(360) = 3$ and $\Omega(360) = 6$. Generally, if the p_j are distinct, $\omega(p_1^{k_1} \dots p_r^{k_r}) = r$ and $\Omega(p_1^{k_1} \dots p_r^{k_r}) = k_1 + \cdots k_r$.

- There is an interesting application of the above which lead to some important developments.
- As a companion to the definition of a multiplicative function we have **Definition.** An $f \in \mathcal{A}$ is **additive** when it satisfies $f(mn) = f(m) + f(n)$ whenever $(m, n) = 1$.
- Now we introduce two further functions. **Definition.** We define $\omega(n)$ to be the number of different prime factors of n and $\Omega(n)$ to be the total number of prime factors of n.
- **Example.** We have $360 = 2^3 3^2 5$ so that $\omega(360) = 3$ and $\Omega(360) = 6$. Generally, if the p_j are distinct, $\omega(p_1^{k_1} \dots p_r^{k_r}) = r$ and $\Omega(p_1^{k_1} \dots p_r^{k_r}) = k_1 + \cdots k_r$.
- One might expect that most of the time Ω is appreciably bigger than ω, but in fact this is not so.

- There is an interesting application of the above which lead to some important developments.
- As a companion to the definition of a multiplicative function we have **Definition.** An $f \in \mathcal{A}$ is **additive** when it satisfies $f(mn) = f(m) + f(n)$ whenever $(m, n) = 1$.
- Now we introduce two further functions. **Definition.** We define $\omega(n)$ to be the number of different prime factors of n and $\Omega(n)$ to be the total number of prime factors of n.
- **Example.** We have $360 = 2^3 3^2 5$ so that $\omega(360) = 3$ and $\Omega(360) = 6$. Generally, if the p_j are distinct, $\omega(p_1^{k_1} \dots p_r^{k_r}) = r$ and $\Omega(p_1^{k_1} \dots p_r^{k_r}) = k_1 + \cdots k_r$.
- One might expect that most of the time Ω is appreciably bigger than ω, but in fact this is not so.
- By the way, there is some connection with the divisor function. It is not hard to show that $2^{\omega(n)} \le d(n) \le 2^{\Omega(n)}$.

- There is an interesting application of the above which lead to some important developments.
- As a companion to the definition of a multiplicative function we have **Definition.** An $f \in \mathcal{A}$ is **additive** when it satisfies $f(mn) = f(m) + f(n)$ whenever $(m, n) = 1$.
- Now we introduce two further functions. **Definition.** We define $\omega(n)$ to be the number of different prime factors of n and $\Omega(n)$ to be the total number of prime factors of n.
- **Example.** We have $360 = 2^3 3^2 5$ so that $\omega(360) = 3$ and $\Omega(360) = 6$. Generally, if the p_j are distinct, $\omega(p_1^{k_1} \dots p_r^{k_r}) = r$ and $\Omega(p_1^{k_1} \dots p_r^{k_r}) = k_1 + \cdots k_r$.
- One might expect that most of the time Ω is appreciably bigger than ω, but in fact this is not so.
- By the way, there is some connection with the divisor function. It is not hard to show that $2^{\omega(n)} \le d(n) \le 2^{\Omega(n)}$.
- In fact this is a simple consequence of the chain of inequalities $2 \le k + 1 \le 2^k$.

- We can now establish that the average number of prime divisors of a number n is $\log\log n$.

Theorem 9

Suppose that $X \geq 2$. Then

$$\sum_{n\leq X} \omega(n) = X\log\log X + BX + O\left(\frac{X}{\log X}\right)$$

where B is the constant of Theorem 7, and

$$\sum_{n\leq X} \Omega(n) =$$

$$X\log\log X + \left(B + \sum_{p} \frac{1}{p(p-1)}\right) X + O\left(\frac{X}{\log X}\right).$$

- Here is the proof for ω.

Proof.

We have

$$\sum_{n\le X}\omega(n)=\sum_{n\le X}\sum_{p\mid n}1=\sum_{p\le X}\left\lfloor\frac{X}{p}\right\rfloor$$
$$=X\sum_{p\le X}\frac{1}{p}+O\big(\pi(x)\big)$$

and the result follows by combining Corollary 6 and Theorem 7. The case of Ω is similar. $\square$

- Hardy and Ramanujan made the remarkable discovery that $\log\log n$ is not just the average of $\omega(n)$, but is its normal order.

- Hardy and Ramanujan made the remarkable discovery that $\log\log n$ is not just the average of $\omega(n)$, but is its normal order.
- Later Turán found a simple proof of this.

Theorem 10 (Hardy & Ramanujan)

Suppose that $X \geq 2$. Then

$$\sum_{n\leq X}\left(\omega(n)-\sum_{p\leq X}\frac{1}{p}\right)^2 \ll X\sum_{p\leq X}\frac{1}{p},$$

$$\sum_{n\leq X}(\omega(n)-\log\log X)^2 \ll X\log\log X$$

and

$$\sum_{2\leq n\leq X}(\omega(n)-\log\log n)^2 \ll X\log\log X$$

- Here is Turán's proof. It is easily seen that

$$\sum_{n\le X}\left(\sum_{p\le X}\frac{1}{p}-\log\log X)\right)^2\ll X$$

and (generally if $Y\ge 1$ we have $\log Y\le 2Y^{1/2}$)

$$\begin{aligned}\sum_{2\le n\le X}(\log\log X-\log\log n)^2&=\sum_{2\le n\le X}\left(\log\frac{\log X}{\log n}\right)^2\\&\ll\sum_{n\le X}\frac{\log X}{\log n}\\&=\sum_{n\le X}\int_n^X\frac{dt}{t}\\&=\int_1^X\frac{\lfloor t\rfloor}{t}dt\\&\le X.\end{aligned}$$

- Thus it suffices to prove the second statement in the theorem.

- Thus it suffices to prove the second statement in the theorem.
- We have

$$\sum_{n\le X}\omega(n)^2=\sum_{p_1\le X}\sum_{\substack{p_2\le X\\ p_2\ne p_1}}\left\lfloor\frac{X}{p_1p_2}\right\rfloor+\sum_{p\le X}\left\lfloor\frac{X}{p}\right\rfloor$$
$$\le X(\log\log X)^2+O(X\log\log X).$$

- Thus it suffices to prove the second statement in the theorem.
- We have

$$\sum_{n\le X}\omega(n)^2=\sum_{p_1\le X}\sum_{\substack{p_2\le X\\ p_2\ne p_1}}\left\lfloor\frac{X}{p_1p_2}\right\rfloor+\sum_{p\le X}\left\lfloor\frac{X}{p}\right\rfloor$$
$$\le X(\log\log X)^2+O(X\log\log X).$$

- Hence

$$\sum_{n\le X}(\omega(n)-\log\log X)^2\le 2X(\log\log X)^2$$
$$-2(\log\log X)\sum_{n\le X}\omega(n)+O(X\log\log X)$$

and this is $\ll X\log\log X$.

- One way of interpreting this theorem is to think of it probabilistically.

- One way of interpreting this theorem is to think of it probabilistically.
- It is saying that the events $p|n$ are approximately independent and occur with probability $\frac{1}{p}$.

- One way of interpreting this theorem is to think of it probabilistically.
- It is saying that the events $p|n$ are approximately independent and occur with probability $\frac{1}{p}$.
- One might guess that the distribution is normal, and this indeed is true and was established by Erdős and Kac about 1941.

- One way of interpreting this theorem is to think of it probabilistically.
- It is saying that the events $p|n$ are approximately independent and occur with probability $\frac{1}{p}$.
- One might guess that the distribution is normal, and this indeed is true and was established by Erdős and Kac about 1941.
- Let

$$\Phi(a,b) = \lim_{x\to\infty} \frac{1}{x}\mathrm{card}\{n \le x : a < \frac{\omega(n) - \log\log n}{\sqrt{\log\log n}} \le b\}.$$

Then

$$\Phi(a,b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} dt.$$

- One way of interpreting this theorem is to think of it probabilistically.
- It is saying that the events $p|n$ are approximately independent and occur with probability $\frac{1}{p}$.
- One might guess that the distribution is normal, and this indeed is true and was established by Erdős and Kac about 1941.
- Let

$$\Phi(a,b) = \lim_{x\to\infty} \frac{1}{x}\mathrm{card}\{n \le x : a < \frac{\omega(n) - \log\log n}{\sqrt{\log\log n}} \le b\}.$$

Then

$$\Phi(a,b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} dt.$$

- The proof uses sieve theory, which we might explore later.

- Let $k \in \mathbb{N}$ and $\Phi_k(z)$ be the k-th cyclotomic polynomial.

$$\Phi_k(z) = \prod_{\substack{\ell=1 \\ (\ell,k)=1}}^{k} (z - \varpi^\ell) \text{ where } \varpi = e^{2\pi i/k}.$$

- Let $k \in \mathbb{N}$ and $\Phi_k(z)$ be the k-th cyclotomic polynomial.

$$\Phi_k(z) = \prod_{\substack{\ell=1 \\ (\ell,k)=1}}^{k} (z - \varpi^\ell) \text{ where } \varpi = e^{2\pi i/k}.$$

- The roots of Φ_k are the primitive k-th roots of unity.

- Let $k \in \mathbb{N}$ and $\Phi_k(z)$ be the k-th cyclotomic polynomial.

$$\Phi_k(z) = \prod_{\substack{\ell=1 \\ (\ell,k)=1}}^{k} (z - \varpi^\ell) \text{ where } \varpi = e^{2\pi i/k}.$$

- The roots of Φ_k are the primitive k-th roots of unity.
- Note that $\Phi_k(z)$ is a (polynomial) factor of $z^k - 1$.

- Let $k \in \mathbb{N}$ and $\Phi_k(z)$ be the k-th cyclotomic polynomial.

$$\Phi_k(z) = \prod_{\substack{\ell=1 \\ (\ell,k)=1}}^{k} (z - \varpi^\ell) \text{ where } \varpi = e^{2\pi i/k}.$$

- The roots of Φ_k are the primitive k-th roots of unity.
- Note that $\Phi_k(z)$ is a (polynomial) factor of $z^k - 1$.
- The Möbius function will remove the condition $(\ell, k) = 1$.

- Let $k \in \mathbb{N}$ and $\Phi_k(z)$ be the k-th cyclotomic polynomial.

$$\Phi_k(z) = \prod_{\substack{\ell=1 \\ (\ell,k)=1}}^{k} (z - \varpi^\ell) \text{ where } \varpi = e^{2\pi i/k}.$$

- The roots of Φ_k are the primitive k-th roots of unity.
- Note that $\Phi_k(z)$ is a (polynomial) factor of $z^k - 1$.
- The Möbius function will remove the condition $(\ell, k) = 1$.
- Thus $\Phi_k(z) = \prod_{\ell=1}^{k} (z - \varpi^l)^{\sum_{m|(\ell,k)} \mu(m)} =$

$$\prod_{\ell=1}^{k} \prod_{m|(\ell,k)} (z - \varpi^\ell)^{\mu(m)} = \prod_{m|k} \left(\prod_{n=1}^{k/m} (z - \varpi^{nm}) \right)^{\mu(m)}.$$

- Thus $\Phi_k(z) = \prod_{m|k} \left(\prod_{n=1}^{k/m} (z - \varpi^{nm}) \right)^{\mu(m)}$.

- Thus $\Phi_k(z) = \prod_{m|k} \left(\prod_{n=1}^{k/m} (z - \varpi^{nm}) \right)^{\mu(m)}$.
- Therefore
$$\Phi_k(z) = \prod_{m|k} (z^{k/m} - 1)^{\mu(m)}.$$

- Thus $\Phi_k(z) = \prod_{m|k} \left(\prod_{n=1}^{k/m} (z - \varpi^{nm}) \right)^{\mu(m)}$.
- Therefore
$$\Phi_k(z) = \prod_{m|k} (z^{k/m} - 1)^{\mu(m)}.$$
- **Example.** The cases $k = 4$ and 6 are
$$\Phi_4(z) = (z - i)(z + i) = z^2 + 1 = \frac{z^4 - 1}{z^2 - 1}$$
and
$$\Phi_6(z) = (z - \varpi)(z - \varpi^5) = z^2 - z + 1 = \frac{(z^6 - 1)(z - 1)}{(z^3 - 1)(z^2 - 1)}.$$
For any prime p
$$\Phi_p(z) = z^{p-1} + z^{p-2} + \cdots + z + 1.$$

- **Theorem 8.14.** *$\Phi_k(z)$ has integer coefficients.*

- **Theorem 8.14.** *$\Phi_k(z)$ has integer coefficients.*
- **Proof.** The case $k = 1$ is clear.

- **Theorem 8.14.** $\Phi_k(z)$ *has integer coefficients.*
- **Proof.** The case $k = 1$ is clear.
- By the formula $\Phi_k(z) = \prod_{m|k}(z^{k/m} - 1)^{\mu(m)}$, when $|z| < 1$ and $k > 1$, we have $\Phi_k(z) =$

$$(-1)^{\sum_{m|k}\mu(m)} \prod_{m|k}(1 - z^{k/m})^{\mu(m)}$$

$$= \prod_{m|k}(1 - z^{k/m})^{\mu(m)}$$

$$= \prod_{\substack{m|k \\ \mu(m)=1}} (1 - z^{k/m}) \prod_{\substack{m|k \\ \mu(m)=-1}} (1 + z^{k/m} + z^{2k/m} + \cdots).$$

- **Theorem 8.14.** *$\Phi_k(z)$ has integer coefficients.*
- **Proof.** The case $k = 1$ is clear.
- By the formula $\Phi_k(z) = \prod_{m|k}(z^{k/m} - 1)^{\mu(m)}$, when $|z| < 1$ and $k > 1$, we have $\Phi_k(z) =$

$$(-1)^{\sum_{m|k}\mu(m)} \prod_{m|k}(1 - z^{k/m})^{\mu(m)}$$

$$= \prod_{m|k}(1 - z^{k/m})^{\mu(m)}$$

$$= \prod_{\substack{m|k \\ \mu(m)=1}} (1 - z^{k/m}) \prod_{\substack{m|k \\ \mu(m)=-1}} (1 + z^{k/m} + z^{2k/m} + \cdots).$$

- A finite product of absolutely convergent series with integer coefficients.

- **Theorem 8.14.** $\Phi_k(z)$ *has integer coefficients.*
- **Proof.** The case $k = 1$ is clear.
- By the formula $\Phi_k(z) = \prod_{m|k}(z^{k/m} - 1)^{\mu(m)}$, when $|z| < 1$ and $k > 1$, we have $\Phi_k(z) =$

$$(-1)^{\sum_{m|k}\mu(m)} \prod_{m|k}(1 - z^{k/m})^{\mu(m)}$$

$$= \prod_{m|k}(1 - z^{k/m})^{\mu(m)}$$

$$= \prod_{\substack{m|k \\ \mu(m)=1}} (1 - z^{k/m}) \prod_{\substack{m|k \\ \mu(m)=-1}} (1 + z^{k/m} + z^{2k/m} + \cdots).$$

- A finite product of absolutely convergent series with integer coefficients.
- The constant term of $\Phi_k(z)$ is $\prod_{\substack{\ell=1 \\ (\ell,k)=1}}^{k} (-\varpi^l)$ which has modulus 1. Thus it is ± 1.

- We can use these polynomials to prove the next theorem

- We can use these polynomials to prove the next theorem
- **Theorem 8.15.** *Suppose that $k \in \mathbb{N}$. Then there are infinitely many primes of the form $kx + 1$.*

- We can use these polynomials to prove the next theorem
- **Theorem 8.15.** *Suppose that $k \in \mathbb{N}$. Then there are infinitely many primes of the form $kx + 1$.*
- **Proof.** Suppose that $r \in \mathbb{N}$, $r > 1$ and p is a prime with $p \nmid k$ and $p | \Phi_k(r)$.

- We can use these polynomials to prove the next theorem
- **Theorem 8.15.** *Suppose that $k \in \mathbb{N}$. Then there are infinitely many primes of the form $kx + 1$.*
- **Proof.** Suppose that $r \in \mathbb{N}$, $r > 1$ and p is a prime with $p \nmid k$ and $p | \Phi_k(r)$.
- Then $p | r^k - 1$ and $p \nmid r$.

- We can use these polynomials to prove the next theorem
- **Theorem 8.15.** *Suppose that $k \in \mathbb{N}$. Then there are infinitely many primes of the form $kx + 1$.*
- **Proof.** Suppose that $r \in \mathbb{N}$, $r > 1$ and p is a prime with $p \nmid k$ and $p|\Phi_k(r)$.
- Then $p|r^k - 1$ and $p \nmid r$.
- Thus $e = \operatorname{ord}_p r|k$, and if $m|k$ and $p|r^m - 1$, then $e|m$.

- We can use these polynomials to prove the next theorem
- **Theorem 8.15.** *Suppose that $k \in \mathbb{N}$. Then there are infinitely many primes of the form $kx + 1$.*
- **Proof.** Suppose that $r \in \mathbb{N}$, $r > 1$ and p is a prime with $p \nmid k$ and $p | \Phi_k(r)$.
- Then $p | r^k - 1$ and $p \nmid r$.
- Thus $e = \operatorname{ord}_p r | k$, and if $m | k$ and $p | r^m - 1$, then $e | m$.
- Write $r^e = 1 + up^v$ for some positive integers u and v with $p \nmid u$.

- We can use these polynomials to prove the next theorem
- **Theorem 8.15.** *Suppose that $k \in \mathbb{N}$. Then there are infinitely many primes of the form $kx + 1$.*
- **Proof.** Suppose that $r \in \mathbb{N}$, $r > 1$ and p is a prime with $p \nmid k$ and $p | \Phi_k(r)$.
- Then $p | r^k - 1$ and $p \nmid r$.
- Thus $e = \operatorname{ord}_p r | k$, and if $m | k$ and $p | r^m - 1$, then $e | m$.
- Write $r^e = 1 + up^v$ for some positive integers u and v with $p \nmid u$.
- Then $r^{e\ell} - 1 = (1 + up^v)^\ell - 1 \equiv \ell up^v \pmod{p^{2v}}$.

- We can use these polynomials to prove the next theorem
- **Theorem 8.15.** *Suppose that $k \in \mathbb{N}$. Then there are infinitely many primes of the form $kx + 1$.*
- **Proof.** Suppose that $r \in \mathbb{N}$, $r > 1$ and p is a prime with $p \nmid k$ and $p|\Phi_k(r)$.
- Then $p|r^k - 1$ and $p \nmid r$.
- Thus $e = \operatorname{ord}_p r|k$, and if $m|k$ and $p|r^m - 1$, then $e|m$.
- Write $r^e = 1 + up^v$ for some positive integers u and v with $p \nmid u$.
- Then $r^{e\ell} - 1 = (1 + up^v)^\ell - 1 \equiv \ell up^v \pmod{p^{2v}}$.
- Thus if $\ell|k$, so that $p \nmid l$, then p^v is the exact power of p dividing $r^{e\ell} - 1$.

- We can use these polynomials to prove the next theorem
- **Theorem 8.15.** *Suppose that $k \in \mathbb{N}$. Then there are infinitely many primes of the form $kx + 1$.*
- **Proof.** Suppose that $r \in \mathbb{N}$, $r > 1$ and p is a prime with $p \nmid k$ and $p | \Phi_k(r)$.
- Then $p | r^k - 1$ and $p \nmid r$.
- Thus $e = \operatorname{ord}_p r | k$, and if $m | k$ and $p | r^m - 1$, then $e | m$.
- Write $r^e = 1 + up^v$ for some positive integers u and v with $p \nmid u$.
- Then $r^{e\ell} - 1 = (1 + up^v)^\ell - 1 \equiv \ell up^v \pmod{p^{2v}}$.
- Thus if $\ell | k$, so that $p \nmid l$, then p^v is the exact power of p dividing $r^{e\ell} - 1$.
- Thus the exact power of p dividing $\Phi_k(r)$ is

$$\prod_{\substack{m|k \\ e|m}} (p^v)^{\mu(m)} = p^{v \sum_{l|(k/e)} \mu((k/e)/l)}.$$

- Thus the exact power of p dividing $\Phi_k(r)$ is

$$\prod_{\substack{m|k \\ e|m}} (p^v)^{\mu(m)} = p^{v\sum_{l|(k/e)}\mu((k/e)/l)}$$

and the exponent is 0 unless $e = k$.

- Thus the exact power of p dividing $\Phi_k(r)$ is

$$\prod_{\substack{m|k \\ e|m}} (p^v)^{\mu(m)} = p^{v \sum_{l|(k/e)} \mu((k/e)/l)}$$

  and the exponent is 0 unless $e = k$.
- Thus we have shown that if $p \nmid k$ and $p|\Phi_k(r)$, then r has order k modulo p.

- Thus the exact power of p dividing $\Phi_k(r)$ is

$$\prod_{\substack{m|k \\ e|m}} (p^v)^{\mu(m)} = p^{v\sum_{l|(k/e)}\mu((k/e)/l)}$$

  and the exponent is 0 unless $e = k$.
- Thus we have shown that if $p \nmid k$ and $p|\Phi_k(r)$, then r has order k modulo p.
- Thus $k = \operatorname{ord} p(r)|p-1$.

- Thus the exact power of p dividing $\Phi_k(r)$ is

$$\prod_{\substack{m|k \\ e|m}} (p^v)^{\mu(m)} = p^{v\sum_{l|(k/e)}\mu((k/e)/l)}$$

  and the exponent is 0 unless $e = k$.
- Thus we have shown that if $p \nmid k$ and $p|\Phi_k(r)$, then r has order k modulo p.
- Thus $k = \operatorname{ord} p(r)|p-1$.
- Now suppose there are only a finite number of primes $p_1, \dots, p_j$ in the residue class 1 modulo k and let $r = kyp_1 \dots p_j$ where y is chosen to ensure that $\Phi_k(r) > 1$.

- Thus the exact power of p dividing $\Phi_k(r)$ is

$$\prod_{\substack{m|k \\ e|m}} (p^v)^{\mu(m)} = p^{v\sum_{l|(k/e)}\mu((k/e)/l)}$$

  and the exponent is 0 unless $e = k$.
- Thus we have shown that if $p \nmid k$ and $p|\Phi_k(r)$, then r has order k modulo p.
- Thus $k = \operatorname{ord} p(r)|p-1$.
- Now suppose there are only a finite number of primes $p_1, \dots, p_j$ in the residue class 1 modulo k and let $r = kyp_1 \dots p_j$ where y is chosen to ensure that $\Phi_k(r) > 1$.
- Then there is at least one prime with $p|\Phi_k(r)$ and from above $p \equiv 1 \pmod{k}$.

- Thus the exact power of p dividing $\Phi_k(r)$ is

$$\prod_{\substack{m|k \\ e|m}} (p^v)^{\mu(m)} = p^{v\sum_{l|(k/e)}\mu((k/e)/l)}$$

  and the exponent is 0 unless $e = k$.
- Thus we have shown that if $p \nmid k$ and $p|\Phi_k(r)$, then r has order k modulo p.
- Thus $k = \operatorname{ord} p(r)|p-1$.
- Now suppose there are only a finite number of primes $p_1, \dots, p_j$ in the residue class 1 modulo k and let $r = kyp_1 \dots p_j$ where y is chosen to ensure that $\Phi_k(r) > 1$.
- Then there is at least one prime with $p|\Phi_k(r)$ and from above $p \equiv 1 \pmod{k}$.
- Thus $p|r$ also.

- Thus the exact power of p dividing $\Phi_k(r)$ is

$$\prod_{\substack{m|k\\ e|m}} (p^v)^{\mu(m)} = p^{v\sum_{l|(k/e)}\mu((k/e)/l)}$$

  and the exponent is 0 unless $e = k$.
- Thus we have shown that if $p \nmid k$ and $p|\Phi_k(r)$, then r has order k modulo p.
- Thus $k = \operatorname{ord} p(r)|p-1$.
- Now suppose there are only a finite number of primes $p_1, \dots, p_j$ in the residue class 1 modulo k and let $r = kyp_1 \dots p_j$ where y is chosen to ensure that $\Phi_k(r) > 1$.
- Then there is at least one prime with $p|\Phi_k(r)$ and from above $p \equiv 1 \pmod{k}$.
- Thus $p|r$ also.
- Hence p divides the constant term of $\Phi_k(z) = \pm 1$ which is absurd.