Arithmetical Functions

Averages of arithmetical functions

Elementary Prime number theory

Orders of magnitude of arithmetical functions.

# Math 571 Chapter 1 Elementary Results

Robert C. Vaughan

January 11, 2023

- Since I are not sure of the number theory background of everyone in the class I will start by discussing some useful topics from elementary number theory.

- The set $\mathcal{A}$ of arithmetical functions is defined by

$$\mathcal{A} = \{f : \mathbb{N} \to \mathbb{C}\}.$$

- The set $\mathcal{A}$ of arithmetical functions is defined by

$$\mathcal{A} = \{f : \mathbb{N} \to \mathbb{C}\}.$$

- The set $\mathcal{A}$ of arithmetical functions is defined by

$$\mathcal{A} = \{f : \mathbb{N} \to \mathbb{C}\}.$$

- Of course the range of any particular function might well be a subset of $\mathbb{C}$.

- The set $\mathcal{A}$ of arithmetical functions is defined by

$$\mathcal{A} = \{f : \mathbb{N} \to \mathbb{C}\}.$$

- Of course the range of any particular function might well be a subset of $\mathbb{C}$.
- There are quite a number of important arithmetical functions.

- Some examples are

- Some examples are
- **The divisor function**. The number of positive divisors of $n$.

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

- Some examples are
- **The divisor function**. The number of positive divisors of $n$.

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

- **Euler's function.** The number $\phi(n)$ of integers $m$ with $1 \le m \le n$ and $(m, n) = 1$. This is important because it counts the number of units in $\mathbb{Z}/n\mathbb{Z}$.

- Euler's function satisfies an interesting relationship.

**Theorem 1**

*We have* $\sum_{m|n}\phi(m)=n$.

- Euler's function satisfies an interesting relationship.

**Theorem 1**

*We have* $\sum_{m|n}\phi(m)=n$.

- One way of seeing this is as follows. Consider the $n$ fractions
$$\frac{1}{n},\frac{2}{n},\ldots,\frac{n}{n}.$$

- Euler's function satisfies an interesting relationship.

**Theorem 1**

*We have* $\sum_{m|n} \phi(m) = n$.

- One way of seeing this is as follows. Consider the $n$ fractions
$$\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}.$$
- Then factor out any common factors between denominators and numerators. Then one will obtain each fraction of the form
$$\frac{l}{m}$$
with $m|n$, $1 \le l \le m$ and $(l, m) = 1$.

- Euler's function satisfies an interesting relationship.

**Theorem 1**

*We have* $\sum_{m|n} \phi(m) = n$.

- One way of seeing this is as follows. Consider the $n$ fractions
$$\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}.$$
- Then factor out any common factors between denominators and numerators. Then one will obtain each fraction of the form
$$\frac{l}{m}$$
with $m|n$, $1 \le l \le m$ and $(l, m) = 1$.
- The number of such fractions is
$$\sum_{m|n} \phi(m).$$

- **The Möbius function.** This is a more peculiar function. It is defined to be

$$\mu(n) = (-1)^k$$

when $n = p_1 \dots p_k$ and the $p_j$ are distinct, and is defined to be 0 otherwise.

- It is also convenient to introduce three less interesting functions.

- It is also convenient to introduce three less interesting functions.
- **The unit.**

$$e(n) = \begin{cases} 1 & (n = 1), \\ 0 & (n > 1). \end{cases}$$

- It is also convenient to introduce three less interesting functions.
- **The unit.**

$$e(n) = \begin{cases} 1 & (n = 1), \\ 0 & (n > 1). \end{cases}$$

- **The one.**

$$\mathbf{1}(n) = 1 \text{ for every } n.$$

- It is also convenient to introduce three less interesting functions.
- **The unit.**

$$e(n) = \begin{cases} 1 & (n = 1), \\ 0 & (n > 1). \end{cases}$$

- **The one.**

$$\mathbf{1}(n) = 1 \text{ for every } n.$$

- **The identity.**

$$N(n) = n.$$

- Two other functions which have interesting structures but which we will say less about at this stage are

- Two other functions which have interesting structures but which we will say less about at this stage are
- **The primitive character modulo** 4. We define

$$\chi_1(n) = \begin{cases} (-1)^{\frac{n-1}{2}} & 2 \nmid n, \\ 0 & 2|n. \end{cases}$$

- **Sums of two squares.** We define $r(n)$ to be the number of ways of writing $n$ as the sum of two squares of integers.

- Two other functions which have interesting structures but which we will say less about at this stage are
- **The primitive character modulo 4.** We define

$$\chi_1(n) = \begin{cases} (-1)^{\frac{n-1}{2}} & 2 \nmid n, \\ 0 & 2|n. \end{cases}$$

- **Sums of two squares.** We define $r(n)$ to be the number of ways of writing $n$ as the sum of two squares of integers.
- For example, $1 = 0^2 + (\pm 1)^2 = (\pm 1)^2 + 0^2$, so $r(1) = 4$, $r(3) = r(6) = r(7) = 0$, $r(9) = 4$,
$65 = (\pm 1)^2 + (\pm 8)^2 = (\pm 4)^2 + (\pm 7)^2$ so $r(65) = 16$.

- Two other functions which have interesting structures but which we will say less about at this stage are
- **The primitive character modulo 4.** We define

$$\chi_1(n) = \begin{cases} (-1)^{\frac{n-1}{2}} & 2 \nmid n, \\ 0 & 2|n. \end{cases}$$

- **Sums of two squares.** We define $r(n)$ to be the number of ways of writing $n$ as the sum of two squares of integers.
- For example, $1 = 0^2 + (\pm1)^2 = (\pm1)^2 + 0^2$, so $r(1) = 4$, $r(3) = r(6) = r(7) = 0$, $r(9) = 4$,
$65 = (\pm1)^2 + (\pm8)^2 = (\pm4)^2 + (\pm7)^2$ so $r(65) = 16$.
- $d$, $\phi$, $e$, $\mathbf{1}$, $N$, $\chi_1$ have an interesting property. That is they are multiplicative.

- **Definition** An arithmetical function $f$ which is not identically 0 is **multiplicative** when it satisfies

$$f(mn) = f(m)f(n)$$

whenever $(m, n) = 1$.

- **Definition** An arithmetical function $f$ which is not identically 0 is **multiplicative** when it satisfies

$$f(mn) = f(m)f(n)$$

whenever $(m, n) = 1$.
- Let $\mathcal{M}$ denote the set of multiplicative functions.

- **Definition** An arithmetical function $f$ which is not identically 0 is **multiplicative** when it satisfies

$$f(mn) = f(m)f(n)$$

whenever $(m, n) = 1$.

- Let $\mathcal{M}$ denote the set of multiplicative functions.
- The function $r(n)$ is not multiplicative, since $r(65) = 16$ but $r(5) = r(13) = 8$.

- **Definition** An arithmetical function $f$ which is not identically 0 is **multiplicative** when it satisfies

$$f(mn) = f(m)f(n)$$

whenever $(m, n) = 1$.

- Let $\mathcal{M}$ denote the set of multiplicative functions.
- The function $r(n)$ is not multiplicative, since $r(65) = 16$ but $r(5) = r(13) = 8$.
- Indeed the fact that $r(1) \neq 1$ would contradict the next theorem.

- **Definition** An arithmetical function $f$ which is not identically 0 is **multiplicative** when it satisfies

$$f(mn) = f(m)f(n)$$

whenever $(m, n) = 1$.

- Let $\mathcal{M}$ denote the set of multiplicative functions.
- The function $r(n)$ is not multiplicative, since $r(65) = 16$ but $r(5) = r(13) = 8$.
- Indeed the fact that $r(1) \neq 1$ would contradict the next theorem.
- However it is true that $r(n)/4$ is multiplicative, but this is a little trickier to prove.

- We have

**Theorem 2**

*Suppose that $f \in \mathcal{M}$. Then $f(1) = 1$.*

- We have

**Theorem 2**

*Suppose that $f \in \mathcal{M}$. Then $f(1) = 1$.*

- The proof is easy.

**Proof.**

Since $f$ is not identically 0 there is an $n$ such that $f(n) \neq 0$. Hence $f(n) = f(n \times 1) = f(n)f(1)$, and the conclusion follows. $\square$

- It is pretty obvious that $e$, $\mathbf{1}$ and $N$ are in $\mathcal{M}$, and it is actually quite easy to show

**Theorem 3**

*We have* $\mu \in \mathcal{M}$.

- It is pretty obvious that $e$, $\mathbf{1}$ and $N$ are in $\mathcal{M}$, and it is actually quite easy to show

## Theorem 3

*We have $\mu \in \mathcal{M}$.*

## Proof.

Suppose that $(m, n) = 1$. If $p^2|mn$, then $p^2|m$ or $p^2|n$, so $\mu(mn) = 0 = \mu(m)\mu(n)$. If

$$m = p_1 \dots p_k, \quad n = p_1' \dots p_l'$$

with the $p_i, p_j'$ distinct, then

$$\mu(mn) = (-1)^{k+l} = (-1)^k(-1)^l = \mu(m)\mu(n).$$

□

-

- The following is very useful.

**Theorem 4**

*Suppose the* $f \in \mathcal{M}$, $g \in \mathcal{M}$ *and* $h$ *is defined for each* $n$ *by* $h(n) = \sum_{m|n} f(m)g(n/m)$. *Then* $h \in \mathcal{M}$.

- The following is very useful.

## Theorem 4

*Suppose the $f \in \mathcal{M}$, $g \in \mathcal{M}$ and $h$ is defined for each $n$ by $h(n) = \sum_{m|n} f(m)g(n/m)$. Then $h \in \mathcal{M}$.*

## Proof.

Suppose $(n_1, n_2) = 1$. Then a typical divisor $m$ of $n_1 n_2$ is uniquely of the form $m_1 m_2$ with $m_1|n_1$ and $m_2|n_2$. Hence

$$\begin{aligned} h(n_1 n_2) &= \sum_{m_1|n_1} \sum_{m_2|n_2} f(m_1 m_2) g(n_1 n_2/(m_1 m_2)) \\ &= \sum_{m_1|n_1} f(m_1) g(n_1/m_1) \sum_{m_2|n_2} f(m_2) g(n_2/m_2). \end{aligned}$$

$\square$

-

- This enables is to establish an interesting property of the Möbius function.

## Theorem 5

*We have*

$$\sum_{m|n} \mu(m) = e(n).$$

- This enables is to establish an interesting property of the Möbius function.

## Theorem 5

*We have*

$$\sum_{m|n}\mu(m)=e(n).$$

## Proof.

By the previous theorem the sum here is $\sum_{m|n}\mu(m)\mathbf{1}(n/m)$ is in $\mathcal{M}$. Moreover if $k\ge 1$, then

$$\sum_{m|p^k}\mu(m)=\mu(1)+\mu(p)=1-1=0$$

□

-

- This suggests a general way of defining new functions.
  **Definition.** Given two arithmetical functions $f$ and $g$ we define the **Dirichlet convolution** $f * g$ to be the function defined by

$$(f * g)(n) = \sum_{m|n} f(m)g(n/m).$$

- This suggests a general way of defining new functions.
  **Definition.** Given two arithmetical functions $f$ and $g$ we define the **Dirichlet convolution** $f * g$ to be the function defined by

$$(f * g)(n) = \sum_{m|n} f(m)g(n/m).$$

- Note that this operation is commutative - simply replace $m$ by $n/m$.

- This suggests a general way of defining new functions.
  **Definition.** Given two arithmetical functions $f$ and $g$ we define the **Dirichlet convolution** $f * g$ to be the function defined by
  $$(f * g)(n) = \sum_{m|n} f(m)g(n/m).$$
- Note that this operation is commutative - simply replace $m$ by $n/m$.
- It is also quite easy to see that
  $$(f * g) * h = f * (g * h).$$

- This suggests a general way of defining new functions.
**Definition.** Given two arithmetical functions $f$ and $g$ we define the **Dirichlet convolution** $f * g$ to be the function defined by

$$(f * g)(n) = \sum_{m|n} f(m)g(n/m).$$

- Note that this operation is commutative - simply replace $m$ by $n/m$.
- It is also quite easy to see that

$$(f * g) * h = f * (g * h).$$

- Write the left hand side as

$$\sum_{m|n} \left( \sum_{l|m} f(l)g(m/l) \right) h(n/m)$$

and interchange the order of summation and replace $m$ by $kl$.

- Dirichlet convolution has some interesting properties

- Dirichlet convolution has some interesting properties
- 1. $f * e = e * f = f$ for any $f \in \mathcal{A}$, so $e$ is really acting as a unit.

- Dirichlet convolution has some interesting properties
- 1. $f * e = e * f = f$ for any $f \in \mathcal{A}$, so $e$ is really acting as a unit.
- 2. $\mu * \mathbf{1} = \mathbf{1} * \mu = e$, so $\mu$ is the inverse of $\mathbf{1}$, and *vice versa*.

- Dirichlet convolution has some interesting properties
- 1. $f * e = e * f = f$ for any $f \in \mathcal{A}$, so $e$ is really acting as a unit.
- 2. $\mu * \mathbf{1} = \mathbf{1} * \mu = e$, so $\mu$ is the inverse of $\mathbf{1}$, and *vice versa*.
- 3. $d = \mathbf{1} * \mathbf{1}$, so $d \in \mathcal{M}$. Hence

- Dirichlet convolution has some interesting properties
- 1. $f * e = e * f = f$ for any $f \in \mathcal{A}$, so $e$ is really acting as a unit.
- 2. $\mu * \mathbf{1} = \mathbf{1} * \mu = e$, so $\mu$ is the inverse of $\mathbf{1}$, and *vice versa*.
- 3. $d = \mathbf{1} * \mathbf{1}$, so $d \in \mathcal{M}$. Hence
- 4. $d(p^k) = k + 1$ and $d(p_1^{k_1} \dots p_r^{k_r}) = (k_1 + 1) \dots (k_r + 1)$.

- The Möbius function has other interesting properties.

**Theorem 6 (Möbius inversion I)**

*Suppose that $f \in \mathcal{A}$ and $g = f * \mathbf{1}$. Then $f = g * \mu$.*

- The Möbius function has other interesting properties.

**Theorem 6 (Möbius inversion I)**

*Suppose that $f \in \mathcal{A}$ and $g = f * \mathbf{1}$. Then $f = g * \mu$.*

- Using Dirichlet convolution the proof is easy.

**Proof.**

We have

$$g * \mu = (f * \mathbf{1}) * \mu = f * (\mathbf{1} * \mu) = f * e = f.$$

$\square$

- The Möbius function has other interesting properties.

**Theorem 6 (Möbius inversion I)**

*Suppose that $f \in \mathcal{A}$ and $g = f * \mathbf{1}$. Then $f = g * \mu$.*

- Using Dirichlet convolution the proof is easy.

**Proof.**

We have

$$g * \mu = (f * \mathbf{1}) * \mu = f * (\mathbf{1} * \mu) = f * e = f.$$

□

- There is a converse theorem

**Theorem 7 (Möbius inversion II)**

*Suppose that $g \in \mathcal{A}$ and $f = g * \mu$, then $g = f * \mathbf{1}$.*

- The Möbius function has other interesting properties.

**Theorem 6 (Möbius inversion I)**

*Suppose that $f \in \mathcal{A}$ and $g = f * \mathbf{1}$. Then $f = g * \mu$.*

- Using Dirichlet convolution the proof is easy.

**Proof.**

We have

$$g * \mu = (f * \mathbf{1}) * \mu = f * (\mathbf{1} * \mu) = f * e = f.$$

□

- There is a converse theorem

**Theorem 7 (Möbius inversion II)**

*Suppose that $g \in \mathcal{A}$ and $f = g * \mu$, then $g = f * \mathbf{1}$.*

- The proof is similar.

- There are some interesting consequences

## Theorem 8

*We have $\phi = \mu * N$ and $\phi \in \mathcal{M}$. Moreover*

$$\phi(n) = n \sum_{m|n} \frac{\mu(m)}{m} = n \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

- There are some interesting consequences

## Theorem 8

*We have $\phi = \mu * N$ and $\phi \in \mathcal{M}$. Moreover*

$$\phi(n) = n \sum_{m|n} \frac{\mu(m)}{m} = n \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

- Again the proof is easy.

## Proof.

We saw in Theorem 1 that $\phi * \mathbf{1} = N$. Hence by the previous theorem we have $\phi = N * \mu = \mu * N$. Therefore, by Theorem 4, $\phi \in \mathcal{M}$. Moreover $\phi(p^k) = p^k - p^{k-1}$ and we are done. $\square$

- A structure theorem.

**Theorem 9**

*Let* $\mathcal{D} = \{f \in \mathcal{A} : f(1) \neq 0\}$. *Then* $\langle \mathcal{D}, * \rangle$ *is an abelian group.*

- A structure theorem.

**Theorem 9**

*Let $\mathcal{D} = \{f \in \mathcal{A} : f(1) \neq 0\}$. Then $\langle \mathcal{D}, * \rangle$ is an abelian group.*

- The proof is constructive.

**Proof.**

Of course $e$ is the unit, and closure is obvious. We already checked commutativity and associativity. It remains, given $f \in \mathcal{D}$, to construct an inverse. Define $g$ iteratively by $g(1) = 1/f(1)$, $g(n) = -\sum_{\substack{m|n\\ m>1}} f(m)g(n/m)/f(1)$ and it is clear that $f * g = e$. $\square$

- One of the most powerful techniques we have is to take an average.

- One of the most powerful techniques we have is to take an average.
- One of the more famous theorems of this kind is

**Theorem 10 (Dirichlet)**

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

$$\sum_{n \leq X} d(n) = X \log X + (2C - 1)X + O(X^{1/2}).$$

- We follow Dirichlet's proof method, which has become known as the *method of the parabola*.

- We follow Dirichlet's proof method, which has become known as the *method of the parabola*.
- The divisor function $d(n)$ can be thought of as the number of ordered pairs of positive integers $m, l$ such that $ml = n$.

- We follow Dirichlet's proof method, which has become known as the *method of the parabola*.
- The divisor function $d(n)$ can be thought of as the number of ordered pairs of positive integers $m, l$ such that $ml = n$.
- Thus when we sum over $n \le X$ we are just counting the number of ordered pairs $m, l$ such that $ml \le X$.

- We follow Dirichlet's proof method, which has become known as the *method of the parabola*.
- The divisor function $d(n)$ can be thought of as the number of ordered pairs of positive integers $m, l$ such that $ml = n$.
- Thus when we sum over $n \le X$ we are just counting the number of ordered pairs $m, l$ such that $ml \le X$.
- In other words we are counting the number of *lattice points* $m, l$ under the rectangular hyperbola

$$xy = X.$$

- We follow Dirichlet's proof method, which has become known as the *method of the parabola*.
- The divisor function $d(n)$ can be thought of as the number of ordered pairs of positive integers $m, l$ such that $ml = n$.
- Thus when we sum over $n \le X$ we are just counting the number of ordered pairs $m, l$ such that $ml \le X$.
- In other words we are counting the number of *lattice points* $m, l$ under the rectangular hyperbola

$$xy = X.$$

- We could just crudely count, given $m \le X$, the number of choices for $l$, namely

$$\left\lfloor \frac{X}{m} \right\rfloor$$

and obtain

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

but this gives a much weaker error term.

- Dirichlet's idea is divide the region under the hyperbola into two parts using its symmetry in the line $y = x$.

- Dirichlet's idea is divide the region under the hyperbola into two parts using its symmetry in the line $y = x$.
- That two regions are the part with

$$m \le \sqrt{X},\ l \le \frac{X}{m}$$

and that with

$$l \le \sqrt{X},\ m \le \frac{X}{l}.$$

- Dirichlet's idea is divide the region under the hyperbola into two parts using its symmetry in the line $y = x$.
- That two regions are the part with

$$m \le \sqrt{X},\ l \le \frac{X}{m}$$

and that with

$$l \le \sqrt{X},\ m \le \frac{X}{l}.$$

- Clearly each region has the same number of lattice points. However the points $m, l$ with $m \le \sqrt{X}$ and $l \le \sqrt{X}$ are counted in both regions.

- Thus we obtain

$$\begin{aligned}\sum_{n\le X} d(n) &= 2\sum_{m\le\sqrt{X}}\left\lfloor\frac{X}{m}\right\rfloor - \lfloor\sqrt{X}\rfloor^2\\ &= 2\sum_{m\le\sqrt{X}}\frac{X}{m} - X + O(X^{1/2})\\ &= 2X\big(\log(\sqrt{X}) + C\big) - X + O(X^{1/2}).\end{aligned}$$

where in the last line we used Euler's estimate for $S(x)$.

- One can also compute an average for Euler's function

**Theorem 11**

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

$$\sum_{n \le x} \phi(n) = \frac{x^2}{2} \sum_{m=1}^{\infty} \frac{\mu(m)}{m^2} + O(x \log x).$$

- One can also compute an average for Euler's function

## Theorem 11

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

$$\sum_{n \leq x} \phi(n) = \frac{x^2}{2} \sum_{m=1}^{\infty} \frac{\mu(m)}{m^2} + O(x \log x).$$

- We remark that the infinite series here is "well known" to be $\frac{6}{\pi^2}$.

- One can also compute an average for Euler's function

**Theorem 11**

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

$$\sum_{n \leq x} \phi(n) = \frac{x^2}{2} \sum_{m=1}^{\infty} \frac{\mu(m)}{m^2} + O(x \log x).$$

- We remark that the infinite series here is "well known" to be $\frac{6}{\pi^2}$.
- We leave the proof largely to the class as homework.
- Hint: Use $\phi = \mu * N$ to obtain

$$\sum_{n \leq x} \phi(n) = \sum_{n \leq x} n \sum_{m|n} \frac{\mu(m)}{m} = \sum_{m \leq x} \mu(m) \sum_{l \leq x/m} l$$

and use a good approximation to the inner sum.

- Likewise the sum of two squares function

**Theorem 12 (Gauss)**

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

$$\sum_{n \leq X} r(n) = \pi X + O(X^{1/2}).$$

- Likewise the sum of two squares function

**Theorem 12 (Gauss)**

*Suppose that $x \in \mathbb{R}$ and $x \ge 2$. Then*

$$\sum_{n \le X} r(n) = \pi X + O(X^{1/2}).$$

- Again we leave the proof as an exercise. As a hint, there is a general principal which is easy to prove in this case that the number of lattice points in a convex region is equal to the area of the region with an error proportional to the length of the boundary.

- 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 $x \le 1000$. 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 $\Lambda$ 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 $\Lambda$ 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 13**

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

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

## Lemma 13

*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

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

$\square$

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

**Lemma 14 (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

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

□

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

## Theorem 15

*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 15

*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 $X \ge 0$ we define

$$\psi(X) = \sum_{n\le X} \Lambda(n),$$
$$\vartheta(X) = \sum_{p\le X} \log p,$$
$$\pi(X) = \sum_{p\le X} 1.$$

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

### Theorem 16

*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 $\Lambda$ 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 $\Lambda$ 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 17 (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 17 (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

$$\begin{aligned}\psi(X) &\ge \sum_{n\le X} \Lambda(n) f(X/n)\\ &= \sum_{n\le X} \Lambda(n) \left\lfloor \frac{X}{n} \right\rfloor - 2 \sum_{n\le X/2} \Lambda(n) \left\lfloor \frac{X/2}{n} \right\rfloor.\end{aligned}$$

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

- Hence

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

- Here we used the fact that there is no contribution to the second sum when $X/2 < n \leq X$.
- Now we apply Theorem 15 and obtain for $x \geq 4$

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

- 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 15 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 $X \ge 4$,

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

and we have already seen that this is

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

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

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

and we have already seen that this is

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

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

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

Hence, for any $k \ge 0$,

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

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

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

and we have already seen that this is

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

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

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

Hence, for any $k \ge 0$,

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

- Summing over all $k$ gives the desired upper bound.

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

## Corollary 18 (Chebyshev)

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

$$\begin{aligned} C_3X <&\vartheta(X) < C_4X, \\ \frac{C_5X}{\log X} <&\pi(X) < \frac{C_6X}{\log X}. \end{aligned}$$

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

**Theorem 19 (Mertens)**

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

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

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

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

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

**Theorem 19 (Mertens)**

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

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

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

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

- I don't want to spend time on the proof, but it is given below and you can see it in the files if you are interested.

- Proof By Theorem 15 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 15 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 15 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 15 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

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

- 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 20 (Mertens)**

*We have*

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

- Another theorem which can be deduced is the following.

**Theorem 20 (Mertens)**

*We have*

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

- I do not give the proof here. In practice the third estimate in the previous theorem is usually adequate.

- 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 $\Omega$ is appreciably bigger than $\omega$, 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 $\Omega$ is appreciably bigger than $\omega$, 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 $\Omega$ is appreciably bigger than $\omega$, 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 21

*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 19, and*

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

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

- We skip the proof.

**Proof.**

We have

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

and the result follows by combining Corollary 18 and Theorem 19.

The case of $\Omega$ 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 22 (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.

Arithmetical Functions

Averages of arithmetical functions

Elementary Prime number theory

Orders of magnitude of arithmetical functions.

- Multiplicative functions oscillate quite a bit.

- Multiplicative functions oscillate quite a bit.
- For example $d(p)=2$ but if $n$ is the product of the first $k$ primes $n=\prod_{p\le X} p$, then $\log n = \vartheta(X)$ so that $X \ll \log n \ll X$ by Chebyshev.

- Multiplicative functions oscillate quite a bit.
- For example $d(p)=2$ but if $n$ is the product of the first $k$ primes $n=\prod_{p\le X} p$, then $\log n=\vartheta(X)$ so that $X \ll \log n \ll X$ by Chebyshev.
- Thus $\log X \sim \log\log n$, but $d(n)=2^{\pi(X)}$ so that

$$\begin{aligned}\log d(n) = (\log 2)\pi(X) &\ge (\log 2)\frac{\vartheta(X)}{\log X} \\ &\sim (\log 2)\frac{\log n}{\log\log n}.\end{aligned}$$

- We have

**Theorem 23**

*For every $\varepsilon > 0$ there are infinitely many $n$ such that*

$$d(n) > \exp\left(\frac{(\log 2 - \varepsilon)\log n}{\log\log n}\right).$$

- We have

**Theorem 23**

*For every $\varepsilon > 0$ there are infinitely many $n$ such that*

$$d(n) > \exp\left(\frac{(\log 2 - \varepsilon)\log n}{\log\log n}\right).$$

- The function $d(n)$ also arises in comparisons, for example in deciding the convergence of certain important series.

- Thus it is useful to have a simple universal upper bound.

**Theorem 24**

*Let $\varepsilon > 0$. Then there is a positive number $C$ which depends at most on $\varepsilon$ such that for every $n \in \mathbb{N}$ we have*

$$d(n) < Cn^{\varepsilon}.$$

- Thus it is useful to have a simple universal upper bound.

## Theorem 24

*Let $\varepsilon > 0$. Then there is a positive number $C$ which depends at most on $\varepsilon$ such that for every $n \in \mathbb{N}$ we have*

$$d(n) < Cn^{\varepsilon}.$$

- Note, such a statement is often written as

$$d(n) = O_{\varepsilon}(n^{\varepsilon})$$

or

$$d(n) \ll_{\varepsilon} n^{\varepsilon}.$$

- Thus it is useful to have a simple universal upper bound.

## Theorem 24

*Let $\varepsilon > 0$. Then there is a positive number $C$ which depends at most on $\varepsilon$ such that for every $n \in \mathbb{N}$ we have*

$$d(n) < Cn^{\varepsilon}.$$

- Note, such a statement is often written as

$$d(n) = O_{\varepsilon}(n^{\varepsilon})$$

or

$$d(n) \ll_{\varepsilon} n^{\varepsilon}.$$

- It suffices to prove the theorem when

$$\varepsilon \le \frac{1}{\log 2}.$$

- Thus it is useful to have a simple universal upper bound.

## Theorem 24

*Let $\varepsilon > 0$. Then there is a positive number $C$ which depends at most on $\varepsilon$ such that for every $n \in \mathbb{N}$ we have*

$$d(n) < Cn^{\varepsilon}.$$

- Note, such a statement is often written as

$$d(n) = O_{\varepsilon}(n^{\varepsilon})$$

or

$$d(n) \ll_{\varepsilon} n^{\varepsilon}.$$

- It suffices to prove the theorem when

$$\varepsilon \leq \frac{1}{\log 2}.$$

- Write $n = p_1^{k_1} \dots p_r^{k_r}$ where the $p_j$ are distinct.

- Recall that $d(n) = (k_1 + 1) \dots (k_r + 1)$.

- Recall that $d(n) = (k_1 + 1) \dots (k_r + 1)$.
- Thus

$$\frac{d(n)}{n^{\varepsilon}} = \prod_{j=1}^{r} \frac{k_j + 1}{p_j^{\varepsilon k_j}}.$$

- Recall that $d(n) = (k_1 + 1) \dots (k_r + 1)$.
- Thus
$$\frac{d(n)}{n^\varepsilon} = \prod_{j=1}^{r} \frac{k_j + 1}{p_j^{\varepsilon k_j}}.$$
- Since we are only interested in an upper bound, the terms for which $p_j^\varepsilon > 2$ can be thrown away since $2^k \ge k + 1$.

- Recall that $d(n) = (k_1 + 1) \dots (k_r + 1)$.
- Thus

$$\frac{d(n)}{n^\varepsilon} = \prod_{j=1}^{r} \frac{k_j + 1}{p_j^{\varepsilon k_j}}.$$

- Since we are only interested in an upper bound, the terms for which $p_j^\varepsilon > 2$ can be thrown away since $2^k \geq k + 1$.
- However there are only $\leq 2^{1/\varepsilon}$ primes $p_j$ for which

$$p_j^\varepsilon \leq 2.$$

- Recall that $d(n) = (k_1 + 1) \dots (k_r + 1)$.
- Thus
$$\frac{d(n)}{n^\varepsilon} = \prod_{j=1}^{r} \frac{k_j + 1}{p_j^{\varepsilon k_j}}.$$
- Since we are only interested in an upper bound, the terms for which $p_j^\varepsilon > 2$ can be thrown away since $2^k \geq k + 1$.
- However there are only $\leq 2^{1/\varepsilon}$ primes $p_j$ for which
$$p_j^\varepsilon \leq 2.$$
- Morever for any such prime we have
$$\begin{aligned} p_j^{\varepsilon k_j} &\geq 2^{\varepsilon k_j} = \exp(\varepsilon k_j \log 2) \\ &\geq 1 + \varepsilon k_j \log 2 \geq (k_j + 1)\varepsilon \log 2. \end{aligned}$$

- Recall that $d(n) = (k_1 + 1) \dots (k_r + 1)$.
- Thus
$$\frac{d(n)}{n^\varepsilon} = \prod_{j=1}^{r} \frac{k_j + 1}{p_j^{\varepsilon k_j}}.$$
- Since we are only interested in an upper bound, the terms for which $p_j^\varepsilon > 2$ can be thrown away since $2^k \geq k + 1$.
- However there are only $\leq 2^{1/\varepsilon}$ primes $p_j$ for which
$$p_j^\varepsilon \leq 2.$$
- Morever for any such prime we have
$$\begin{aligned} p_j^{\varepsilon k_j} &\geq 2^{\varepsilon k_j} = \exp(\varepsilon k_j \log 2) \\ &\geq 1 + \varepsilon k_j \log 2 \geq (k_j + 1)\varepsilon \log 2. \end{aligned}$$
- Thus
$$\frac{d(n)}{n^\varepsilon} \leq \left( \frac{1}{\varepsilon \log 2} \right)^{2^{1/\varepsilon}}.$$

- The above proof can be refined to give a companion to Theorem 23

**Theorem 25**

*Let $\varepsilon > 0$. Then for all $n > n_0$ we have*

$$d(n) < \exp\left(\frac{(\log 2 + \varepsilon)\log n}{\log\log n}\right).$$

- The above proof can be refined to give a companion to Theorem 23

## Theorem 25

*Let $\varepsilon > 0$. Then for all $n > n_0$ we have*

$$d(n) < \exp\left(\frac{(\log 2 + \varepsilon)\log n}{\log\log n}\right).$$

- We follow the proof of the previous theorem until the final inequality. Then replace the $\varepsilon$ there with

$$\frac{(1+\varepsilon/2)\log 2}{\log\log n}$$

which for large $n$ certainly meets the requirement of being no larger than $1/\log 2$.

- Now

$$\left(\frac{1}{\varepsilon\log 2}\right)^{2^{1/\varepsilon}} = \exp\left(\exp\left(\frac{\log\log n}{1+\varepsilon/2}\right)\log\frac{\log\log n}{(1+\varepsilon/2)\log 2}\right) < \exp\left(\frac{\varepsilon(\log n)\log 2}{2\log\log n}\right)$$

for sufficiently large $n$. Hence

$$d(n) < n^{\frac{(1+\varepsilon/2)\log 2}{\log\log n}} \exp\left(\frac{\varepsilon(\log n)\log 2}{2\log\log n}\right) = \exp\left(\frac{(1+\varepsilon)(\log n)\log 2}{\log\log n}\right) < \exp\left(\frac{(\log 2+\varepsilon)(\log n)}{\log\log n}\right).$$