Introduction to Number Theory Chapter 7 Arithmetical Functions

Robert C. Vaughan

March 21, 2025

Introduction

- It is convenient to make the following definition.

Definition 1

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

Introduction

- It is convenient to make the following definition.

Definition 1

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

- The range of a function might be a subset of $\mathbb{C}$, e.g. $\mathbb{R}$ or $\mathbb{Z}$. There are several important arithmetical functions.

Definition 2 (The divisor function)

The number of positive divisors of n, $d(n) = \sum_{m|n} 1$.

Introduction

- It is convenient to make the following definition.

Definition 1

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

- The range of a function might be a subset of $\mathbb{C}$, e.g. $\mathbb{R}$ or $\mathbb{Z}$. There are several important arithmetical functions.

Definition 2 (The divisor function)

The number of positive divisors of n, $d(n) = \sum_{m|n} 1$.

Definition 3 (The Möbius function)

$$\mu(n) = \begin{cases} (-1)^k & \text{if } n \text{ is a product of } k \text{ distinct primes,} \\ 0 & \text{if there is a prime } p \text{ such that } p^2 | n. \end{cases}$$

-

- It is also convenient to introduce three simple functions.

- It is also convenient to introduce three simple functions.

Definition 4 (The Unit)

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

-

- It is also convenient to introduce three simple functions.

Definition 4 (The Unit)

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

-

Definition 5 (The One)

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

-

- It is also convenient to introduce three simple functions.

Definition 4 (The Unit)

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

-

Definition 5 (The One)

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

-

Definition 6 (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

Definition 7 (The primitive character modulo 4)

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

-

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

Definition 7 (The primitive character modulo 4)

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

-

- Similar functions we have already met are Euler's function ϕ, the Legendre symbol and its generalization the Jacobi symbol
$$\left(\frac{n}{m}\right)_J.$$
Here we think of it as a function of n, keeping m fixed, but we could also think of it as a function of m keeping n fixed.

Definition 8 (**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.

-

Definition 8 (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.

-
- This is the function $r_2(n)$ of the previous chapter.

Definition 8 (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.

-
- This is the function $r_2(n)$ of the previous chapter.

Example 9

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 $r(65) = 16$.

-

- The functions d, ϕ, e, $\mathbf{1}$, N, χ_1, $\left(\frac{\cdot}{m}\right)_J$ have an important property. That is that they are multiplicative. We already discussed this in connection with Euler's function and the Legendre and Jacobi symbols. Here is a reminder.

Definition 10

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

$$f(mn) = f(m)f(n) \tag{1.1}$$

whenever $(m, n) = 1$. Let $\mathcal{M}$ denote the set of multiplicative functions. If (1.1) holds for **all** m and n, then we say that f is **totally multiplicative**.

- The functions d, ϕ, e, $\mathbf{1}$, N, χ_1, $\left(\frac{\cdot}{m}\right)_J$ have an important property. That is that they are multiplicative. We already discussed this in connection with Euler's function and the Legendre and Jacobi symbols. Here is a reminder.

Definition 10

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

$$f(mn) = f(m)f(n) \tag{1.1}$$

whenever $(m,n) = 1$. Let $\mathcal{M}$ denote the set of multiplicative functions. If (1.1) holds for **all** m and n, then we say that f is **totally multiplicative**.

- The function $r(n)$ is not multiplicative, since $r(65) = 16$ but $r(5) = r(13) = 8$.

- The functions d, ϕ, e, $\mathbf{1}$, N, χ_1, $\left(\frac{\cdot}{m}\right)_J$ have an important property. That is that they are multiplicative. We already discussed this in connection with Euler's function and the Legendre and Jacobi symbols. Here is a reminder.

Definition 10

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

$$f(mn) = f(m)f(n) \tag{1.1}$$

whenever $(m, n) = 1$. Let $\mathcal{M}$ denote the set of multiplicative functions. If (1.1) holds for **all** m and n, then we say that f is **totally multiplicative**.

- 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.

- The following simple theorem is very helpful in singling out non-multiplicative functions.
  Theorem 7.1. *Suppose that* $f \in \mathcal{M}$*. Then* $f(1) = 1$*.*

- The following simple theorem is very helpful in singling out non-multiplicative functions.
  Theorem 7.1. *Suppose that* $f \in \mathcal{M}$*. Then* $f(1) = 1$*.*
- **Proof.** Since f is not identically 0 there is an n such that $f(n) \neq 0$.

- The following simple theorem is very helpful in singling out non-multiplicative functions.
  Theorem 7.1. *Suppose that* $f \in \mathcal{M}$. *Then* $f(1) = 1$.
- **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.

- The following simple theorem is very helpful in singling out non-multiplicative functions.
  Theorem 7.1. *Suppose that* $f \in \mathcal{M}$. *Then* $f(1) = 1$.
- **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.
- It is pretty obvious that e, $\mathbf{1}$ and N are in $\mathcal{M}$, and it is actually quite easy to show

- The following simple theorem is very helpful in singling out non-multiplicative functions.
  Theorem 7.1. *Suppose that $f \in \mathcal{M}$. Then $f(1) = 1$.*
- **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.
- It is pretty obvious that e, $\mathbf{1}$ and N are in $\mathcal{M}$, and it is actually quite easy to show
- **Theorem 7.2.** *We have $\mu \in \mathcal{M}$.*

- The following simple theorem is very helpful in singling out non-multiplicative functions.
  Theorem 7.1. *Suppose that $f \in \mathcal{M}$. Then $f(1) = 1$.*
- **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.
- It is pretty obvious that e, $\mathbf{1}$ and N are in $\mathcal{M}$, and it is actually quite easy to show
- **Theorem 7.2.** *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 $\mu(mn) = 0 = \mu(m)\mu(n)$.

- The following simple theorem is very helpful in singling out non-multiplicative functions.
  Theorem 7.1. *Suppose that $f \in \mathcal{M}$. Then $f(1) = 1$.*
- **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.
- It is pretty obvious that e, $\mathbf{1}$ and N are in $\mathcal{M}$, and it is actually quite easy to show
- **Theorem 7.2.** *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 7.3. Suppose $f \in \mathcal{M}$, $g \in \mathcal{M}$ and h is defined by $h(n) = \displaystyle\sum_{m|n} f(m)g(n/m)$. Then $h \in \mathcal{M}$.

- The following is very useful.
  Theorem 7.3. Suppose $f \in \mathcal{M}$, $g \in \mathcal{M}$ and h is defined 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 divisor m of $n_1 n_2$ is uniquely of the form $m_1 m_2$ with $m_1|n_1$, $m_2|n_2$. Hence

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

- The following is very useful.
  Theorem 7.3. Suppose $f \in \mathcal{M}$, $g \in \mathcal{M}$ and h is defined 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 divisor m of $n_1 n_2$ is uniquely of the form $m_1 m_2$ with $m_1 | n_1$, $m_2 | n_2$. Hence
$$\begin{aligned} h(n_1 n_2) &= \sum_{m_1|n_1} \sum_{m_2|n_2} f(m_1 m_2) g(n_1 n_2/(m_1 m_2)) \\ &= \sum_{m_1|n_1} f(m_1) g(n_1/m_1) \sum_{m_2|n_2} f(m_2) g(n_2/m_2). \end{aligned}$$
- This enables us to establish an interesting property of μ.
  Theorem 7.4. *We have* $\sum_{m|n} \mu(m) = e(n)$.

- The following is very useful.
  Theorem 7.3. Suppose $f \in \mathcal{M}$, $g \in \mathcal{M}$ and h is defined 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 divisor m of $n_1 n_2$ is uniquely of the form $m_1 m_2$ with $m_1 | n_1$, $m_2 | n_2$. Hence

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

- This enables us to establish an interesting property of μ.
  Theorem 7.4. *We have* $\sum_{m|n} \mu(m) = e(n)$.
- **Proof.** By the definition of $\mathbf{1}$ the sum here is $\sum_{m|n} \mu(m)\mathbf{1}(n/m)$ and by the previous theorem it is in $\mathcal{M}$.

- The following is very useful.
  Theorem 7.3. Suppose $f \in \mathcal{M}$, $g \in \mathcal{M}$ and h is defined 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 divisor m of $n_1 n_2$ is uniquely of the form $m_1 m_2$ with $m_1|n_1$, $m_2|n_2$. Hence
$$\begin{aligned} h(n_1 n_2) &= \sum_{m_1|n_1} \sum_{m_2|n_2} f(m_1 m_2) g(n_1 n_2/(m_1 m_2)) \\ &= \sum_{m_1|n_1} f(m_1)g(n_1/m_1) \sum_{m_2|n_2} f(m_2)g(n_2/m_2). \end{aligned}$$
- This enables us to establish an interesting property of μ.
  Theorem 7.4. *We have* $\sum_{m|n} \mu(m) = e(n)$.
- **Proof.** By the definition of $\mathbf{1}$ the sum here is $\sum_{m|n} \mu(m)\mathbf{1}(n/m)$ and by the previous theorem it is in $\mathcal{M}$.
- And if $k \geq 1$, then $\sum_{m|p^k} \mu(m) = \mu(1) + \mu(p) = 1 - 1 = 0$.

- One function we have seen and used briefly without comment in Chapter 5 is the following

Definition 11

For real numbers α we define the **floor function** $\lfloor \alpha \rfloor$ to be the largest integer not exceeding α.

Example 12

Thus $\lfloor \frac{5}{2} \rfloor = 2$ and $\lfloor -\sqrt{2} \rfloor = -2$.

- One function we have seen and used briefly without comment in Chapter 5 is the following

Definition 11

For real numbers α we define the **floor function** $\lfloor\alpha\rfloor$ to be the largest integer not exceeding α.

Example 12

Thus $\lfloor\frac{5}{2}\rfloor = 2$ and $\lfloor-\sqrt{2}\rfloor = -2$.

- The only property we will use a little later in this chapter is that for any real number α and integer k we have $\lfloor\alpha - k\rfloor = \lfloor\alpha\rfloor - k$, which is easy to check, and otherwise it is just a useful shorthand.

- One function we have seen and used briefly without comment in Chapter 5 is the following

Definition 11

For real numbers α we define the **floor function** $\lfloor \alpha \rfloor$ to be the largest integer not exceeding α.

Example 12

Thus $\lfloor \frac{5}{2} \rfloor = 2$ and $\lfloor -\sqrt{2} \rfloor = -2$.

- The only property we will use a little later in this chapter is that for any real number α and integer k we have $\lfloor \alpha - k \rfloor = \lfloor \alpha \rfloor - k$, which is easy to check, and otherwise it is just a useful shorthand.
- We will explore it in more detail in Chapter 8.

- Theorem 7.3 suggests a way of defining new functions.

Definition 13

Given two arithmetical functions f and g we define the **Dirichlet convolution** $f * g$ to be the function defined by

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

- Theorem 7.3 suggests a way of defining new functions.

Definition 13

Given two arithmetical functions f and g we define the **Dirichlet convolution** $f * g$ to be the function defined by

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

- $m \leftrightarrow n/m$ is bijective, so $(f * g)(n) = \sum_{m|n} f(m)g(n/m) = \sum_{m|n} g(n/m)f(m) = \sum_{m|n} g(m)f(n/m)$, so $*$ commutes.

- Theorem 7.3 suggests a way of defining new functions.

Definition 13

Given two arithmetical functions f and g we define the **Dirichlet convolution** $f * g$ to be the function defined by

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

- $m \leftrightarrow n/m$ is bijective, so $(f * g)(n) = \sum_{m|n} f(m)g(n/m) = \sum_{m|n} g(n/m)f(m) = \sum_{m|n} g(m)f(n/m)$, so $*$ commutes.
- It is also associative $(f * g) * h = f * (g * h)$. Write the left hand side as $\sum_{m|n} \Big(\sum_{l|m} f(l)g(m/l) \Big) h(n/m)$, interchange the order and replace m by kl, so $kl|n$, i.e $l|n$, $k|n/l$. The above is $\sum_{l|n} f(l) \sum_{k|n/l} g(k)h\big((n/l)/k\big) = f * (g * h)(n)$.

- 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 μ 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 μ is the inverse of $\mathbf{1}$, and *vice versa*.
- 3. Theorem 7.3 tells us that if $f \in \mathcal{M}$ and $g \in \mathcal{M}$, then $f * g \in \mathcal{M}$.

- 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 μ is the inverse of $\mathbf{1}$, and *vice versa*.
- 3. Theorem 7.3 tells us that if $f \in \mathcal{M}$ and $g \in \mathcal{M}$, then $f * g \in \mathcal{M}$.
- 4. Theorem 3.2 says that $\phi * \mathbf{1} = N$.

- 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 μ is the inverse of $\mathbf{1}$, and *vice versa*.
- 3. Theorem 7.3 tells us that if $f \in \mathcal{M}$ and $g \in \mathcal{M}$, then $f * g \in \mathcal{M}$.
- 4. Theorem 3.2 says that $\phi * \mathbf{1} = N$.
- 5. $d = \mathbf{1} * \mathbf{1}$, 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 μ is the inverse of $\mathbf{1}$, and *vice versa*.
- 3. Theorem 7.3 tells us that if $f \in \mathcal{M}$ and $g \in \mathcal{M}$, then $f * g \in \mathcal{M}$.
- 4. Theorem 3.2 says that $\phi * \mathbf{1} = N$.
- 5. $d = \mathbf{1} * \mathbf{1}$, so $d \in \mathcal{M}$. Hence
- 6. $d(p^k) = k + 1$ and $d(p_1^{k_1} \dots p_r^{k_r}) = (k_1 + 1) \dots (k_r + 1)$.

- A remarkable property discovered by Möbius, although special cases were certainly known to Gauss and Riemann.
Theorem 7.5. [Möbius inversion I]. *Suppose that* $f \in \mathcal{A}$ *and* $g = f * \mathbf{1}$*. Then* $f = g * \mu$*.*

- A remarkable property discovered by Möbius, although special cases were certainly known to Gauss and Riemann. **Theorem 7.5. [Möbius inversion I].** *Suppose that* $f \in \mathcal{A}$ *and* $g = f * \mathbf{1}$. *Then* $f = g * \mu$.
- **Proof.** We have

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

- A remarkable property discovered by Möbius, although special cases were certainly known to Gauss and Riemann. **Theorem 7.5. [Möbius inversion I].** *Suppose that $f \in \mathcal{A}$ and $g = f * \mathbf{1}$. Then $f = g * \mu$.*
- **Proof.** We have

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

- **Theorem 7.6. [Möbius inversion II.]** *Suppose that $g \in \mathcal{A}$ and $f = g * \mu$, then $g = f * \mathbf{1}$.*

- A remarkable property discovered by Möbius, although special cases were certainly known to Gauss and Riemann. **Theorem 7.5. [Möbius inversion I].** *Suppose that* $f \in \mathcal{A}$ *and* $g = f * \mathbf{1}$*. Then* $f = g * \mu$*.*
- **Proof.** We have

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

- **Theorem 7.6. [Möbius inversion II.]** *Suppose that* $g \in \mathcal{A}$ *and* $f = g * \mu$*, then* $g = f * \mathbf{1}$*.*
- The proof is similar.

- A remarkable property discovered by Möbius, although special cases were certainly known to Gauss and Riemann. **Theorem 7.5. [Möbius inversion I].** *Suppose that* $f \in \mathcal{A}$ *and* $g = f * \mathbf{1}$. *Then* $f = g * \mu$.
- **Proof.** We have

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

- **Theorem 7.6. [Möbius inversion II.]** *Suppose that* $g \in \mathcal{A}$ *and* $f = g * \mu$, *then* $g = f * \mathbf{1}$.
- The proof is similar.
- **Theorem 7.7.** *We have* $\phi = \mu * N$ *and* $\phi \in \mathcal{M}$. *Moreover*

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

- A remarkable property discovered by Möbius, although special cases were certainly known to Gauss and Riemann. **Theorem 7.5. [Möbius inversion I].** *Suppose that* $f \in \mathcal{A}$ *and* $g = f * \mathbf{1}$*. Then* $f = g * \mu$*.*
- **Proof.** We have

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

- **Theorem 7.6. [Möbius inversion II.]** *Suppose that* $g \in \mathcal{A}$ *and* $f = g * \mu$*, then* $g = f * \mathbf{1}$*.*
- The proof is similar.
- **Theorem 7.7.** *We have* $\phi = \mu * N$ *and* $\phi \in \mathcal{M}$*. Moreover*

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

- **Proof.** By property 4. and Theorem 7.5 we have $\phi = N * \mu = \mu * N$.

- A remarkable property discovered by Möbius, although special cases were certainly known to Gauss and Riemann.
  Theorem 7.5. [Möbius inversion I]. *Suppose that* $f \in \mathcal{A}$ *and* $g = f * \mathbf{1}$*. Then* $f = g * \mu$*.*
- **Proof.** We have

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

- **Theorem 7.6. [Möbius inversion II.]** *Suppose that* $g \in \mathcal{A}$ *and* $f = g * \mu$*, then* $g = f * \mathbf{1}$*.*
- The proof is similar.
- **Theorem 7.7.** *We have* $\phi = \mu * N$ *and* $\phi \in \mathcal{M}$*. Moreover*

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

- **Proof.** By property 4. and Theorem 7.5 we have $\phi = N * \mu = \mu * N$.
- Therefore, by property 3 and Theorem 7.2, $\phi \in \mathcal{M}$, and $\phi(p^k) = p^k - p^{k-1}$ and we are done.

- At this stage you might want to look up the definition of a group.

  Theorem 7.8. *Let* $\mathcal{D} = \{f \in \mathcal{A} : f(1) \neq 0\}$. *Then* $\langle \mathcal{D}, * \rangle$ *is an abelian group.*

- At this stage you might want to look up the definition of a group.
  Theorem 7.8. *Let* $\mathcal{D} = \{f \in \mathcal{A} : f(1) \neq 0\}$*. Then* $\langle \mathcal{D}, * \rangle$ *is an abelian group.*
- **Proof.** Of course e is the unit, and closure is obvious. We already checked commutativity and associativity.

- At this stage you might want to look up the definition of a group.
  Theorem 7.8. *Let* $\mathcal{D} = \{f \in \mathcal{A} : f(1) \neq 0\}$. *Then* $\langle \mathcal{D}, * \rangle$ *is an abelian group.*
- **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.

- At this stage you might want to look up the definition of a group.
  Theorem 7.8. *Let $\mathcal{D} = \{f \in \mathcal{A} : f(1) \neq 0\}$. Then $\langle \mathcal{D}, * \rangle$ is an abelian group.*
- **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

$$g(1) = 1/f(1)$$
$$g(n) = -\sum_{\substack{m|n \\ m>1}} f(m)g(n/m)/f(1)$$

- At this stage you might want to look up the definition of a group.

  Theorem 7.8. *Let* $\mathcal{D} = \{f \in \mathcal{A} : f(1) \neq 0\}$. *Then* $\langle \mathcal{D}, * \rangle$ *is an abelian group.*
- **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 $f * g = e$.

- 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.
- For example if we know that $0 \le f(n) \le \sqrt{n}$ for every n and that
$$\sum_{n\le x} f(n) > \frac{1}{2}x^{3/2},$$

- One of the most powerful techniques we have is to take an average.
- For example if we know that $0 \le f(n) \le \sqrt{n}$ for every n and that
$$\sum_{n\le x} f(n) > \frac{1}{2}x^{3/2},$$
- then it would follow that when $x/2 < n \le x$ we have $1 \ge f(n)/\sqrt{x}$
$$\mathrm{card}\{x/2n \le x : f(n) > 0\} \ge \frac{1}{\sqrt{x}} \sum_{x/2<n\le x} f(n) \ge \frac{x}{2}$$

- One of the most powerful techniques we have is to take an average.
- For example if we know that $0 \le f(n) \le \sqrt{n}$ for every n and that
$$\sum_{n \le x} f(n) > \frac{1}{2} x^{3/2},$$
- then it would follow that when $x/2 < n \le x$ we have $1 \ge f(n)/\sqrt{x}$
$$\operatorname{card}\{x/2n \le x : f(n) > 0\} \ge \frac{1}{\sqrt{x}} \sum_{x/2 < n \le x} f(n) \ge \frac{x}{2}$$
- and so $f(n) > 0$ for at least half of all n.

- A more sophisticated version of this would be that if one could show that

$$\sum_{x<n\le 2x} \left(f(n) - n^{1/2}\right)^2 < x^{7/4},$$

then it would follow that for most n the function $f(n)$ is about $n^{1/2}$.

- A more sophisticated version of this would be that if one could show that

$$\sum_{x<n\le 2x}\left(f(n)-n^{1/2}\right)^2 < x^{7/4},$$

then it would follow that for most n the function $f(n)$ is about $n^{1/2}$.

- This technique has been used to show that “almost all” even numbers are the sum of two primes and “almost all” positive integers are the sum of four cubes. Later we will show that “almost all” positive integers n have about $\log\log n$ prime factors.

- We are going to need some notation which avoids the use of $C_1, C_2, \ldots$, etc., to denote unspecified constants.

- We are going to need some notation which avoids the use of $C_1, C_2, \ldots$, etc., to denote unspecified constants.
- Given functions f and g defined on some domain $\mathcal{X}$ with $g(x) \geq 0$ for all $x \in \mathcal{X}$ we write

$$f(x) = O\big(g(x)\big) \tag{3.2}$$

to mean that there is some constant C such that

$$|f(x)| \leq Cg(x) \text{ for every } x \in \mathcal{X}.$$

- We are going to need some notation which avoids the use of $C_1, C_2, \ldots$, etc., to denote unspecified constants.
- Given functions f and g defined on some domain $\mathcal{X}$ with $g(x) \geq 0$ for all $x \in \mathcal{X}$ we write
$$f(x) = O\big(g(x)\big) \tag{3.2}$$
to mean that there is some constant C such that
$$|f(x)| \leq Cg(x) \text{ for every } x \in \mathcal{X}.$$
- We also use $f(x) = o\big(g(x)\big)$ to mean that if there is some limiting operation, such as $x \to \infty$, or $x \to l$, then
$$\frac{f(x)}{g(x)} \to 0$$

- We are going to need some notation which avoids the use of $C_1, C_2, \ldots$, etc., to denote unspecified constants.
- Given functions f and g defined on some domain $\mathcal{X}$ with $g(x) \ge 0$ for all $x \in \mathcal{X}$ we write

$$f(x) = O\big(g(x)\big) \tag{3.2}$$

to mean that there is some constant C such that

$$|f(x)| \le Cg(x) \text{ for every } x \in \mathcal{X}.$$

- We also use $f(x) = o\big(g(x)\big)$ to mean that if there is some limiting operation, such as $x \to \infty$, or $x \to l$, then

$$\frac{f(x)}{g(x)} \to 0$$

- and $f(x) \sim g(x)$ to mean $\frac{f(x)}{g(x)} \to 1$.

- We are going to need some notation which avoids the use of $C_1, C_2, \ldots$, etc., to denote unspecified constants.
- Given functions f and g defined on some domain $\mathcal{X}$ with $g(x) \geq 0$ for all $x \in \mathcal{X}$ we write

$$f(x) = O\big(g(x)\big) \tag{3.2}$$

to mean that there is some constant C such that

$$|f(x)| \leq Cg(x) \text{ for every } x \in \mathcal{X}.$$

- We also use $f(x) = o\big(g(x)\big)$ to mean that if there is some limiting operation, such as $x \to \infty$, or $x \to l$, then

$$\frac{f(x)}{g(x)} \to 0$$

- and $f(x) \sim g(x)$ to mean $\frac{f(x)}{g(x)} \to 1$.
- The symbol O was introduced by Bachmann in 1894, and the symbol o by Landau in 1909.

- The O-symbol can be a bit clumsy for complicated expressions.

- The O-symbol can be a bit clumsy for complicated expressions.
- We will often instead use the Vinogradov symbols, which I. M. Vinogradov introduced about 1934.

- The O-symbol can be a bit clumsy for complicated expressions.
- We will often instead use the Vinogradov symbols, which I. M. Vinogradov introduced about 1934.
- Thus we will use

$$f \ll g \tag{3.3}$$

to mean (3.2).

- The O-symbol can be a bit clumsy for complicated expressions.
- We will often instead use the Vinogradov symbols, which I. M. Vinogradov introduced about 1934.
- Thus we will use
$$f \ll g \tag{3.3}$$
to mean (3.2).
- This also has the advantage that we can write strings of inequalities in the form
$$f_1 \ll f_2 \ll f_3 \ll \dots$$
and retain the meaning at each stage.

- The O-symbol can be a bit clumsy for complicated expressions.
- We will often instead use the Vinogradov symbols, which I. M. Vinogradov introduced about 1934.
- Thus we will use
$$f \ll g \tag{3.3}$$
to mean (3.2).
- This also has the advantage that we can write strings of inequalities in the form
$$f_1 \ll f_2 \ll f_3 \ll \dots$$
and retain the meaning at each stage.
- Also if f is also non-negative we may use
$$g \gg f$$
to mean (3.3).

- Our first theorem on averages concerns the function $r(n)$ and is due to Gauss. The proof illustrates a rather general principle.

Theorem. 7.9. [Gauss.] *Let $X \geq 1$ and $G(X)$ denote the number of lattice points in the disc centre 0 of radius $\sqrt{X}$, i.e. the number of ordered pairs of integers x, y with $x^2 + y^2 \leq X$. Then*

$$G(X) = \sum_{n \leq X} r(n)$$

and

$$G(X) = \pi X + O(X^{1/2}).$$

- Our first theorem on averages concerns the function $r(n)$ and is due to Gauss. The proof illustrates a rather general principle.

  Theorem. 7.9. [Gauss.] *Let $X \ge 1$ and $G(X)$ denote the number of lattice points in the disc centre 0 of radius $\sqrt{X}$, i.e. the number of ordered pairs of integers x, y with $x^2 + y^2 \le X$. Then*

$$G(X) = \sum_{n \le X} r(n)$$

  and

$$G(X) = \pi X + O(X^{1/2}).$$

- Let

$$E(X) = G(X) - \pi X.$$

  The question of the actual size of $E(X)$ is one of the classic open problems of analytic number theory.

- **Theorem. 7.9. [Gauss.]** *Let* $X \ge 1$ *and* $G(X)$ *denote the number of lattice points in the disc centre* 0 *of radius* $\sqrt{X}$, *i.e. the number of ordered pairs of integers* x, y *with* $x^2 + y^2 \le X$. *Then* $G(X) = \pi X + O(X^{1/2})$.

- **Theorem. 7.9. [Gauss.]** *Let $X \ge 1$ and $G(X)$ denote the number of lattice points in the disc centre 0 of radius $\sqrt{X}$, i.e. the number of ordered pairs of integers x, y with $x^2 + y^2 \le X$. Then $G(X) = \pi X + O(X^{1/2})$.*
- **Proof.** We associate with each $(x, y) \in \mathbb{Z}^2$ the square $S(x, y) = [x, x + 1) \times [y, y + 1)$, giving a partition of $\mathbb{R}^2$.

- **Theorem. 7.9. [Gauss.]** *Let $X \ge 1$ and $G(X)$ denote the number of lattice points in the disc centre 0 of radius $\sqrt{X}$, i.e. the number of ordered pairs of integers x, y with $x^2 + y^2 \le X$. Then $G(X) = \pi X + O(X^{1/2})$.*
- **Proof.** We associate with each $(x, y) \in \mathbb{Z}^2$ the square $S(x, y) = [x, x + 1) \times [y, y + 1)$, giving a partition of $\mathbb{R}^2$.
- The squares with $x^2 + y^2 \le X$ are contained in the disc centred at 0 of radius $\sqrt{X} + \sqrt{2}$ (apply the triangle inequality and Pythagoras's theorem).

- **Theorem. 7.9. [Gauss.]** *Let $X \ge 1$ and $G(X)$ denote the number of lattice points in the disc centre 0 of radius $\sqrt{X}$, i.e. the number of ordered pairs of integers x, y with $x^2 + y^2 \le X$. Then $G(X) = \pi X + O(X^{1/2})$.*
- **Proof.** We associate with each $(x, y) \in \mathbb{Z}^2$ the square $S(x, y) = [x, x + 1) \times [y, y + 1)$, giving a partition of $\mathbb{R}^2$.
- The squares with $x^2 + y^2 \le X$ are contained in the disc centred at 0 of radius $\sqrt{X} + \sqrt{2}$ (apply the triangle inequality and Pythagoras's theorem).
- On the other hand their union contains the disc centered at 0 of radius $\sqrt{X} - \sqrt{2}$.

- **Theorem. 7.9. [Gauss.]** *Let $X \ge 1$ and $G(X)$ denote the number of lattice points in the disc centre 0 of radius $\sqrt{X}$, i.e. the number of ordered pairs of integers x, y with $x^2 + y^2 \le X$. Then $G(X) = \pi X + O(X^{1/2})$.*
- **Proof.** We associate with each $(x, y) \in \mathbb{Z}^2$ the square $S(x, y) = [x, x+1) \times [y, y+1)$, giving a partition of $\mathbb{R}^2$.
- The squares with $x^2 + y^2 \le X$ are contained in the disc centred at 0 of radius $\sqrt{X} + \sqrt{2}$ (apply the triangle inequality and Pythagoras's theorem).
- On the other hand their union contains the disc centered at 0 of radius $\sqrt{X} - \sqrt{2}$.
- Moreover their area is $G(X)$ and it lies between the areas of the two discs,

- **Theorem. 7.9. [Gauss.]** *Let $X \geq 1$ and $G(X)$ denote the number of lattice points in the disc centre 0 of radius $\sqrt{X}$, i.e. the number of ordered pairs of integers x, y with $x^2 + y^2 \leq X$. Then $G(X) = \pi X + O(X^{1/2})$.*
- **Proof.** We associate with each $(x, y) \in \mathbb{Z}^2$ the square $S(x, y) = [x, x+1) \times [y, y+1)$, giving a partition of $\mathbb{R}^2$.
- The squares with $x^2 + y^2 \leq X$ are contained in the disc centred at 0 of radius $\sqrt{X} + \sqrt{2}$ (apply the triangle inequality and Pythagoras's theorem).
- On the other hand their union contains the disc centered at 0 of radius $\sqrt{X} - \sqrt{2}$.
- Moreover their area is $G(X)$ and it lies between the areas of the two discs,
- so
$$\pi(\sqrt{X} - \sqrt{2})^2 \leq G(X) \leq \pi(\sqrt{X} + \sqrt{2})^2,$$

- **Theorem. 7.9. [Gauss.]** *Let $X \ge 1$ and $G(X)$ denote the number of lattice points in the disc centre 0 of radius $\sqrt{X}$, i.e. the number of ordered pairs of integers x, y with $x^2 + y^2 \le X$. Then $G(X) = \pi X + O(X^{1/2})$.*
- **Proof.** We associate with each $(x, y) \in \mathbb{Z}^2$ the square $S(x, y) = [x, x+1) \times [y, y+1)$, giving a partition of $\mathbb{R}^2$.
- The squares with $x^2 + y^2 \le X$ are contained in the disc centred at 0 of radius $\sqrt{X} + \sqrt{2}$ (apply the triangle inequality and Pythagoras's theorem).
- On the other hand their union contains the disc centered at 0 of radius $\sqrt{X} - \sqrt{2}$.
- Moreover their area is $G(X)$ and it lies between the areas of the two discs,
- so
$$\pi(\sqrt{X} - \sqrt{2})^2 \le G(X) \le \pi(\sqrt{X} + \sqrt{2})^2,$$
- $$\pi X - \pi 2\sqrt{2}\sqrt{X} + 2\pi < G(X) \le \pi X + \pi 2\sqrt{2}\sqrt{X} + 2\pi$$

- **Theorem. 7.9. [Gauss.]** *Let $X \ge 1$ and $G(X)$ denote the number of lattice points in the disc centre 0 of radius $\sqrt{X}$, i.e. the number of ordered pairs of integers x, y with $x^2 + y^2 \le X$. Then $G(X) = \pi X + O(X^{1/2})$.*
- **Proof.** We associate with each $(x, y) \in \mathbb{Z}^2$ the square $S(x, y) = [x, x+1) \times [y, y+1)$, giving a partition of $\mathbb{R}^2$.
- The squares with $x^2 + y^2 \le X$ are contained in the disc centred at 0 of radius $\sqrt{X} + \sqrt{2}$ (apply the triangle inequality and Pythagoras's theorem).
- On the other hand their union contains the disc centered at 0 of radius $\sqrt{X} - \sqrt{2}$.
- Moreover their area is $G(X)$ and it lies between the areas of the two discs,
- so
$$\pi(\sqrt{X} - \sqrt{2})^2 \le G(X) \le \pi(\sqrt{X} + \sqrt{2})^2,$$
- $$\pi X - \pi 2\sqrt{2}\sqrt{X} + 2\pi < G(X) \le \pi X + \pi 2\sqrt{2}\sqrt{X} + 2\pi$$
- Hence $|G(X) - \pi X| \le \pi 2\sqrt{2}\sqrt{X} + 2\pi \ll \sqrt{X}$.

- The general principle involved in the above proof is that if one has some finite convex region in the plane and one expands it homothetically, then the number of lattice points in the region is approximately the area of the region with an error of order the length of the boundary.

- The general principle involved in the above proof is that if one has some finite convex region in the plane and one expands it homothetically, then the number of lattice points in the region is approximately the area of the region with an error of order the length of the boundary.
- Thus in the theorem above the unit disc centered at the origin has its linear dimensions blown up by a factor of $\sqrt{X}$ (its radius)

- The general principle involved in the above proof is that if one has some finite convex region in the plane and one expands it homothetically, then the number of lattice points in the region is approximately the area of the region with an error of order the length of the boundary.
- Thus in the theorem above the unit disc centered at the origin has its linear dimensions blown up by a factor of $\sqrt{X}$ (its radius)
- and the number of lattice points is approximately its area, πX with an error of order the length of the boundary $2\pi\sqrt{X}$.

- Before proceeding to look further at some of the arithmetical functions we have defined above, consider for $X \geq 1$ the important sum

$$S(X) = \sum_{n \leq X} \frac{1}{n}. \tag{3.4}$$

- Before proceeding to look further at some of the arithmetical functions we have defined above, consider for $X \geq 1$ the important sum

$$S(X) = \sum_{n \leq X} \frac{1}{n}. \tag{3.4}$$

- This crops up in many places.

- Before proceeding to look further at some of the arithmetical functions we have defined above, consider for $X \geq 1$ the important sum

$$S(X) = \sum_{n \leq X} \frac{1}{n}. \tag{3.4}$$

- This crops up in many places.
- You might guess that if there were only a finite number of primes, then this sum would converge as $X \to \infty$,

- Before proceeding to look further at some of the arithmetical functions we have defined above, consider for $X \geq 1$ the important sum

$$S(X) = \sum_{n \leq X} \frac{1}{n}. \tag{3.4}$$

- This crops up in many places.
- You might guess that if there were only a finite number of primes, then this sum would converge as $X \to \infty$,
- and one could see this by multiplying out

$$\prod_p \left(1 - \frac{1}{p}\right)^{-1} = \prod_p \left(1 + \frac{1}{p} + \cdots + \frac{1}{p^k} + \cdots\right).$$

- Before proceeding to look further at some of the arithmetical functions we have defined above, consider for $X \geq 1$ the important sum

$$S(X) = \sum_{n \le X} \frac{1}{n}. \tag{3.4}$$

- This crops up in many places.
- You might guess that if there were only a finite number of primes, then this sum would converge as $X \to \infty$,
- and one could see this by multiplying out

$$\prod_p \left(1 - \frac{1}{p}\right)^{-1} = \prod_p \left(1 + \frac{1}{p} + \cdots + \frac{1}{p^k} + \cdots\right).$$

- But, of course, the sum $S(X)$ behaves a like the integral so is a bit like $\log X$ which tends to infinity with X.

- Before proceeding to look further at some of the arithmetical functions we have defined above, consider for $X \geq 1$ the important sum

$$S(X) = \sum_{n \leq X} \frac{1}{n}. \tag{3.4}$$

- This crops up in many places.
- You might guess that if there were only a finite number of primes, then this sum would converge as $X \to \infty$,
- and one could see this by multiplying out

$$\prod_p \left(1 - \frac{1}{p}\right)^{-1} = \prod_p \left(1 + \frac{1}{p} + \cdots + \frac{1}{p^k} + \cdots\right).$$

- But, of course, the sum $S(X)$ behaves a like the integral so is a bit like $\log X$ which tends to infinity with X.
- In fact there is something more precise which one can say, which was discovered by Euler.

- Recall that $\lfloor * \rfloor$ is defined in Definition 11.

Theorem 7.10. [Euler.] *When $X \geq 1$ the sum $S(X)$ satisfies*

$$S(X) = \log X + C_0 + O\left(\frac{1}{X}\right)$$

where $C_0 = 0.577\ldots$ is Euler's constant

$$C_0 = 1 - \int_1^\infty \frac{t - \lfloor t \rfloor}{t^2} dt.$$

- Recall that $\lfloor * \rfloor$ is defined in Definition 11.
  Theorem 7.10. [Euler.] *When $X \ge 1$ the sum $S(X)$ satisfies*
  $$S(X) = \log X + C_0 + O\left(\frac{1}{X}\right)$$
  where $C_0 = 0.577\ldots$ is Euler's constant
  $$C_0 = 1 - \int_1^\infty \frac{t - \lfloor t \rfloor}{t^2} dt.$$
- **Proof.** We have
  $$S(X) = \sum_{n \le X} \left(\frac{1}{X} + \int_n^X \frac{dt}{t^2}\right) = \frac{\lfloor X \rfloor}{X} + \int_1^X \frac{\lfloor t \rfloor}{t^2} dt$$
  $$= \int_1^X \frac{dt}{t} + 1 - \int_1^X \frac{t - \lfloor t \rfloor}{t^2} dt - \frac{X - \lfloor X \rfloor}{X}$$
  $$= \log X + C_0 + \int_X^\infty \frac{t - \lfloor t \rfloor}{t^2} dt - \frac{X - \lfloor X \rfloor}{X}.$$

- Recall that $\lfloor * \rfloor$ is defined in Definition 11.
  Theorem 7.10. [Euler.] *When $X \ge 1$ the sum $S(X)$ satisfies*
  $$S(X) = \log X + C_0 + O\left(\frac{1}{X}\right)$$
  where $C_0 = 0.577\ldots$ is Euler's constant
  $$C_0 = 1 - \int_1^\infty \frac{t - \lfloor t \rfloor}{t^2} dt.$$
- **Proof.** We have
  $$S(X) = \sum_{n \le X} \left(\frac{1}{X} + \int_n^X \frac{dt}{t^2}\right) = \frac{\lfloor X \rfloor}{X} + \int_1^X \frac{\lfloor t \rfloor}{t^2} dt$$
  $$= \int_1^X \frac{dt}{t} + 1 - \int_1^X \frac{t - \lfloor t \rfloor}{t^2} dt - \frac{X - \lfloor X \rfloor}{X}$$
  $$= \log X + C_0 + \int_X^\infty \frac{t - \lfloor t \rfloor}{t^2} dt - \frac{X - \lfloor X \rfloor}{X}.$$
- Euler computed C_0 to 19 decimal places (by hand!).

- Recall that $\lfloor * \rfloor$ is defined in Definition 11.
  Theorem 7.10. [Euler.] *When $X \geq 1$ the sum $S(X)$ satisfies*
  $$S(X) = \log X + C_0 + O\left(\frac{1}{X}\right)$$
  where $C_0 = 0.577\ldots$ is Euler's constant
  $$C_0 = 1 - \int_1^\infty \frac{t - \lfloor t \rfloor}{t^2} dt.$$
- **Proof.** We have
  $$\begin{aligned} S(X) &= \sum_{n \leq X} \left(\frac{1}{X} + \int_n^X \frac{dt}{t^2}\right) = \frac{\lfloor X \rfloor}{X} + \int_1^X \frac{\lfloor t \rfloor}{t^2} dt \\ &= \int_1^X \frac{dt}{t} + 1 - \int_1^X \frac{t - \lfloor t \rfloor}{t^2} dt - \frac{X - \lfloor X \rfloor}{X} \\ &= \log X + C_0 + \int_X^\infty \frac{t - \lfloor t \rfloor}{t^2} dt - \frac{X - \lfloor X \rfloor}{X}. \end{aligned}$$
- Euler computed C_0 to 19 decimal places (by hand!).
- Actually that is not so hard.

- One of the more famous theorems concerning averages is

Theorem 7.11. [Dirichlet.] *Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then*

$$\sum_{n \leq X} d(n) = X \log X + (2C_0 - 1)X + O(X^{1/2}).$$

- One of the more famous theorems concerning averages is **Theorem 7.11. [Dirichlet.]** *Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then*
$$\sum_{n \leq X} d(n) = X \log X + (2C_0 - 1)X + O(X^{1/2}).$$
- Let $\Delta(X) = \sum_{n \leq X} d(n) - X \log X - (2C_0 - 1)X$.

- One of the more famous theorems concerning averages is **Theorem 7.11. [Dirichlet.]** *Suppose that* $X \in \mathbb{R}$ *and* $X \geq 2$. *Then*

$$\sum_{n \leq X} d(n) = X \log X + (2C_0 - 1)X + O(X^{1/2}).$$

- Let $\Delta(X) = \sum_{n \leq X} d(n) - X \log X - (2C_0 - 1)X$.
- As with the similar question for the Gauss lattice point problem one can ask "how does $\Delta(X)$ really behave?"

- One of the more famous theorems concerning averages is **Theorem 7.11. [Dirichlet.]** *Suppose that* $X \in \mathbb{R}$ *and* $X \geq 2$. *Then*

$$\sum_{n \leq X} d(n) = X \log X + (2C_0 - 1)X + O(X^{1/2}).$$

- Let $\Delta(X) = \sum_{n \leq X} d(n) - X \log X - (2C_0 - 1)X$.
- As with the similar question for the Gauss lattice point problem one can ask "how does $\Delta(X)$ really behave?"
- The divisor function $d(n)$ can be thought of as the number of ordered pairs of positive integers m, l such that $ml = n$.

- One of the more famous theorems concerning averages is **Theorem 7.11. [Dirichlet.]** *Suppose that* $X \in \mathbb{R}$ *and* $X \geq 2$*. Then*
$$\sum_{n \leq X} d(n) = X \log X + (2C_0 - 1)X + O(X^{1/2}).$$
- Let $\Delta(X) = \sum_{n \leq X} d(n) - X \log X - (2C_0 - 1)X$.
- As with the similar question for the Gauss lattice point problem one can ask "how does $\Delta(X)$ really behave?"
- 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 $d(n)$ over $n \leq X$ we are just counting the number of ordered pairs m, ℓ such that $m\ell \leq X$.

- One of the more famous theorems concerning averages is
Theorem 7.11. [Dirichlet.] *Suppose that* $X \in \mathbb{R}$ *and* $X \geq 2$. *Then*
$$\sum_{n \leq X} d(n) = X \log X + (2C_0 - 1)X + O(X^{1/2}).$$
- Let $\Delta(X) = \sum_{n \leq X} d(n) - X \log X - (2C_0 - 1)X$.
- As with the similar question for the Gauss lattice point problem one can ask "how does $\Delta(X)$ really behave?"
- 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 $d(n)$ over $n \leq X$ we are just counting the number of ordered pairs m, ℓ such that $m\ell \leq X$.
- In other words we are counting the number of *lattice points* m, l under the rectangular hyperbola $xy = X$.

- One of the more famous theorems concerning averages is **Theorem 7.11. [Dirichlet.]** *Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then*

$$\sum_{n \leq X} d(n) = X \log X + (2C_0 - 1)X + O(X^{1/2}).$$

- Let $\Delta(X) = \sum_{n \leq X} d(n) - X \log X - (2C_0 - 1)X$.
- As with the similar question for the Gauss lattice point problem one can ask "how does $\Delta(X)$ really behave?"
- 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 $d(n)$ over $n \leq X$ we are just counting the number of ordered pairs m, ℓ such that $m\ell \leq X$.
- In other words we are counting the number of *lattice points* m, l under the rectangular hyperbola $xy = X$.
- The method that Gauss employed for his lattice point problem fails here, because the area under the rectangular hyperbola is infinite, and so is the boundary.

- One of the more famous theorems concerning averages is
Theorem 7.11. [Dirichlet.] *Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then*
$$\sum_{n \leq X} d(n) = X \log X + (2C_0 - 1)X + O(X^{1/2}).$$
- Let $\Delta(X) = \sum_{n\leq X} d(n) - X \log X - (2C_0 - 1)X$.
- As with the similar question for the Gauss lattice point problem one can ask "how does $\Delta(X)$ really behave?"
- 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 $d(n)$ over $n \leq X$ we are just counting the number of ordered pairs m, ℓ such that $m\ell \leq X$.
- In other words we are counting the number of *lattice points* m, l under the rectangular hyperbola $xy = X$.
- The method that Gauss employed for his lattice point problem fails here, because the area under the rectangular hyperbola is infinite, and so is the boundary.
- Nevertheless the number of lattice points under the curve is finite.

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

- We follow Dirichlet's ingenious proof method, which has become known as the *method of the hyperbola*.
- We *could* just crudely count, given $m \le X$, the number of choices for ℓ, namely

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

- We follow Dirichlet's ingenious proof method, which has become known as the *method of the hyperbola*.
- 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
$$\sum_{m\le X} \frac{X}{m} + O(X)$$

- We follow Dirichlet's ingenious proof method, which has become known as the *method of the hyperbola*.
- 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
$$\sum_{m \le X} \frac{X}{m} + O(X)$$
- and then apply Euler's estimate for $S(X)$, but this gives a much weaker result.

- Dirichlet's idea is to divide the region under the hyperbola into two parts.

- Dirichlet's idea is to divide the region under the hyperbola into two parts.
- That with
$$m \le \sqrt{X},\ \ell \le \frac{X}{m}$$
and that with
$$\ell \le \sqrt{X},\ m \le \frac{X}{l}.$$

- Dirichlet's idea is to divide the region under the hyperbola into two parts.
- That with
$$m \le \sqrt{X}, \ell \le \frac{X}{m}$$
and that with
$$\ell \le \sqrt{X}, m \le \frac{X}{l}.$$
- Clearly each region has the same number of lattice points.

- Dirichlet's idea is to divide the region under the hyperbola into two parts.
- That with
$$m \le \sqrt{X},\ \ell \le \frac{X}{m}$$
and that with
$$\ell \le \sqrt{X},\ m \le \frac{X}{l}.$$
- Clearly each region has the same number of lattice points.
- However the points m, ℓ with $m \le \sqrt{X}$ and $\ell \le \sqrt{X}$ are counted in both regions.

- Dirichlet's idea is to divide the region under the hyperbola into two parts.
- That with
$$m \le \sqrt{X},\ \ell \le \frac{X}{m}$$
and that with
$$\ell \le \sqrt{X},\ m \le \frac{X}{l}.$$
- Clearly each region has the same number of lattice points.
- However the points m, ℓ with $m \le \sqrt{X}$ and $\ell \le \sqrt{X}$ are counted in both regions.
- Thus we obtain
$$\begin{aligned}\sum_{n\le X} d(n) &= 2\sum_{m\le\sqrt{X}} \left\lfloor \frac{X}{m} \right\rfloor - \lfloor \sqrt{X} \rfloor^2 \\ &= 2\sum_{m\le\sqrt{X}} \frac{X}{m} - X + O(X^{1/2}) \\ &= 2X\big(\log(\sqrt{X}) + C_0\big) - X + O(X^{1/2}).\end{aligned}$$
where in the last line we used Euler's estimate.

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

Theorem 7.12. *Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then*

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

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

Theorem 7.12. *Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then*

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

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

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

Theorem 7.12. *Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then*

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

- The infinite series here is "well known" to be $\frac{6}{\pi^2}$.
- **Proof.** We have $\phi = \mu * N$. Thus

$$\sum_{n \leq X} \phi(n) = \sum_{n \leq X} n \sum_{m|n} \frac{\mu(m)}{m} = \sum_{m \leq X} \mu(m) \sum_{\ell \leq X/m} \ell.$$

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

Theorem 7.12. *Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then*

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

- The infinite series here is "well known" to be $\frac{6}{\pi^2}$.
- **Proof.** We have $\phi = \mu * N$. Thus

$$\sum_{n \leq X} \phi(n) = \sum_{n \leq X} n \sum_{m|n} \frac{\mu(m)}{m} = \sum_{m \leq X} \mu(m) \sum_{\ell \leq X/m} \ell.$$

- We want a good approximation to the inner sum.

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

Theorem 7.12. *Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then*

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

- The infinite series here is "well known" to be $\frac{6}{\pi^2}$.
- **Proof.** We have $\phi = \mu * N$. Thus

$$\sum_{n \leq X} \phi(n) = \sum_{n \leq X} n \sum_{m|n} \frac{\mu(m)}{m} = \sum_{m \leq X} \mu(m) \sum_{\ell \leq X/m} \ell.$$

- We want a good approximation to the inner sum.
- This is just the sum of an arithmetic progression of $\lfloor X/m \rfloor$ terms with first term 1 and last term $\lfloor X/m \rfloor$.

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

Theorem 7.12. *Suppose that $X \in \mathbb{R}$ and $X \ge 2$. Then*

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

- The infinite series here is "well known" to be $\frac{6}{\pi^2}$.
- **Proof.** We have $\phi = \mu * N$. Thus

$$\sum_{n\le X} \phi(n) = \sum_{n\le X} n \sum_{m|n} \frac{\mu(m)}{m} = \sum_{m\le X} \mu(m) \sum_{\ell \le X/m} \ell.$$

- We want a good approximation to the inner sum.
- This is just the sum of an arithmetic progression of $\lfloor X/m \rfloor$ terms with first term 1 and last term $\lfloor X/m \rfloor$.
- Thus the sum is

$$\frac{1}{2}\left\lfloor \frac{X}{m} \right\rfloor \left(1 + \left\lfloor \frac{X}{m} \right\rfloor\right) = \frac{1}{2}\left(\frac{X}{m}\right)^2 + O\left(\frac{X}{m}\right).$$

- **Theorem 7.12.** *Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then*

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

- **Theorem 7.12.** *Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then*
$$\sum_{n \le X} \phi(n) = \frac{X^2}{2} \sum_{m=1}^{\infty} \frac{\mu(m)}{m^2} + O(X \log X).$$
- $$\sum_{n \le X} \phi(n) = \sum_{n \le X} n \sum_{m|n} \frac{\mu(m)}{m} = \sum_{m \le X} \mu(m) \sum_{\ell \le X/m} \ell.$$

- **Theorem 7.12.** *Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then*
$$\sum_{n \leq X} \phi(n) = \frac{X^2}{2} \sum_{m=1}^{\infty} \frac{\mu(m)}{m^2} + O(X \log X).$$
- $$\sum_{n\leq X} \phi(n) = \sum_{n \leq X} n \sum_{m|n} \frac{\mu(m)}{m} = \sum_{m \leq X} \mu(m) \sum_{\ell \leq X/m} \ell.$$
- $\sum_{l\leq X/m} \ell = \frac{1}{2}\left(\frac{X}{m}\right)^2 + O\left(\frac{X}{m}\right)$.

- **Theorem 7.12.** *Suppose that $X \in \mathbb{R}$ and $X \ge 2$. Then*

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

- $$\sum_{n\le X} \phi(n) = \sum_{n\le X} n \sum_{m|n} \frac{\mu(m)}{m} = \sum_{m\le X} \mu(m) \sum_{\ell \le X/m} \ell.$$
- $\sum_{l\le X/m} \ell = \frac{1}{2}\left(\frac{X}{m}\right)^2 + O\left(\frac{X}{m}\right)$.
- Inserting this in the formula above gives

$$\sum_{n\le X} \phi(n) = \frac{X^2}{2} \sum_{m\le X} \frac{\mu(m)}{m^2} + O\left(\sum_{m\le X} \frac{X}{m}\right).$$

- **Theorem 7.12.** *Suppose that $X \in \mathbb{R}$ and $X \ge 2$. Then*
$$\sum_{n \le X} \phi(n) = \frac{X^2}{2} \sum_{m=1}^{\infty} \frac{\mu(m)}{m^2} + O(X \log X).$$
- $$\sum_{n \le X} \phi(n) = \sum_{n \le X} n \sum_{m|n} \frac{\mu(m)}{m} = \sum_{m \le X} \mu(m) \sum_{\ell \le X/m} \ell.$$
- $\sum_{l \le X/m} \ell = \frac{1}{2} \left(\frac{X}{m}\right)^2 + O\left(\frac{X}{m}\right).$
- Inserting this in the formula above gives
$$\sum_{n \le X} \phi(n) = \frac{X^2}{2} \sum_{m \le X} \frac{\mu(m)}{m^2} + O\left(\sum_{m \le X} \frac{X}{m}\right).$$
- The error term is $\ll X \log X$ by Euler's theorem on $S(X)$.

- **Theorem 7.12.** *Suppose that $X \in \mathbb{R}$ and $X \ge 2$. Then*

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

- $\displaystyle\sum_{n\le X} \phi(n) = \sum_{n \le X} n \sum_{m|n} \frac{\mu(m)}{m} = \sum_{m \le X} \mu(m) \sum_{\ell \le X/m} \ell.$
- $\sum_{l \le X/m} \ell = \frac{1}{2}\left(\frac{X}{m}\right)^2 + O\left(\frac{X}{m}\right).$
- Inserting this in the formula above gives

$$\sum_{n \le X} \phi(n) = \frac{X^2}{2} \sum_{m \le X} \frac{\mu(m)}{m^2} + O\left(\sum_{m \le X} \frac{X}{m} \right).$$

- The error term is $\ll X \log X$ by Euler's theorem on $S(X)$.
- The main term is $\displaystyle\frac{X^2}{2} \sum_{m=1}^{\infty} \frac{\mu(n)}{m^2} + O\left(\sum_{m > X} \frac{X^2}{m^2} \right)$

- **Theorem 7.12.** *Suppose that $X \in \mathbb{R}$ and $X \ge 2$. Then*
$$\sum_{n \le X} \phi(n) = \frac{X^2}{2} \sum_{m=1}^{\infty} \frac{\mu(m)}{m^2} + O(X \log X).$$
- $\displaystyle\sum_{n \le X} \phi(n) = \sum_{n \le X} n \sum_{m|n} \frac{\mu(m)}{m} = \sum_{m \le X} \mu(m) \sum_{\ell \le X/m} \ell.$
- $\sum_{l \le X/m} \ell = \frac{1}{2}\left(\frac{X}{m}\right)^2 + O\left(\frac{X}{m}\right).$
- Inserting this in the formula above gives
$$\sum_{n \le X} \phi(n) = \frac{X^2}{2} \sum_{m \le X} \frac{\mu(m)}{m^2} + O\left(\sum_{m \le X} \frac{X}{m}\right).$$
- The error term is $\ll X \log X$ by Euler's theorem on $S(X)$.
- The main term is $\displaystyle\frac{X^2}{2} \sum_{m=1}^{\infty} \frac{\mu(n)}{m^2} + O\left(\sum_{m > X} \frac{X^2}{m^2}\right)$
- By the monotonicity the error term is $\displaystyle\ll X^2 \int_X^{\infty} \frac{dy}{y^2} \ll X.$

- **Theorem 7.12.** *Suppose that $X \in \mathbb{R}$ and $X \ge 2$. Then*
$$\sum_{n\le X}\phi(n) = \frac{X^2}{2}\sum_{m=1}^{\infty}\frac{\mu(m)}{m^2} + O(X\log X).$$
- $\displaystyle\sum_{n\le X}\phi(n) = \sum_{n\le X} n\sum_{m|n}\frac{\mu(m)}{m} = \sum_{m\le X}\mu(m)\sum_{\ell\le X/m}\ell.$
- $\sum_{l\le X/m}\ell = \frac{1}{2}\left(\frac{X}{m}\right)^2 + O\left(\frac{X}{m}\right).$
- Inserting this in the formula above gives
$$\sum_{n\le X}\phi(n) = \frac{X^2}{2}\sum_{m\le X}\frac{\mu(m)}{m^2} + O\left(\sum_{m\le X}\frac{X}{m}\right).$$
- The error term is $\ll X\log X$ by Euler's theorem on $S(X)$.
- The main term is $\displaystyle\frac{X^2}{2}\sum_{m=1}^{\infty}\frac{\mu(n)}{m^2} + O\left(\sum_{m>X}\frac{X^2}{m^2}\right)$
- By the monotonicity the error term is $\displaystyle\ll X^2\int_X^\infty\frac{dy}{y^2} \ll X.$
- Collecting together our bounds gives the theorem.

- **Theorem 7.12.** *Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then*

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

- **Theorem 7.12.** *Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then*
$$\sum_{n \leq X} \phi(n) = \frac{X^2}{2} \sum_{m=1}^{\infty} \frac{\mu(m)}{m^2} + O(X \log X).$$
- There is a curious application of this.
Theorem 7.13. *The probability that $(m, n) = 1$ is $\frac{6}{\pi^2}$.*

- **Theorem 7.12.** *Suppose that $X \in \mathbb{R}$ and $X \ge 2$. Then*

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

- There is a curious application of this.
Theorem 7.13. *The probability that $(m, n) = 1$ is $\frac{6}{\pi^2}$.*
- i.e. $\dfrac{\text{card}\{m, n : m, n \le X, (m, n) = 1\}}{X^2} \to \dfrac{6}{\pi^2}$ as $X \to \infty$.

- **Theorem 7.12.** *Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then*
$$\sum_{n \leq X} \phi(n) = \frac{X^2}{2} \sum_{m=1}^{\infty} \frac{\mu(m)}{m^2} + O(X \log X).$$
- There is a curious application of this.
Theorem 7.13. *The probability that $(m, n) = 1$ is $\frac{6}{\pi^2}$.*
- i.e. $\dfrac{\operatorname{card}\{m, n : m, n \leq X, (m, n) = 1\}}{X^2} \to \dfrac{6}{\pi^2}$ as $X \to \infty$.
- **Proof.** $\displaystyle\sum_{n \leq X} \phi(n) = \sum_{n \leq X} \sum_{\substack{m \leq n \\ (m,n)=1}} 1$
$= \operatorname{card}\{m, n : m \leq n \leq X, (m, n) = 1\}$
$= \dfrac{1}{2} \operatorname{card}\{m, n : m, n \leq X, (m, n) = 1\} + \dfrac{1}{2}$ since if $m > 1$, then $(m, m) = m > 1$.

- **Theorem 7.12.** *Suppose that $X \in \mathbb{R}$ and $X \geq 2$. Then*
$$\sum_{n\le X}\phi(n)=\frac{X^2}{2}\sum_{m=1}^{\infty}\frac{\mu(m)}{m^2}+O(X\log X).$$
- There is a curious application of this.
Theorem 7.13. *The probability that $(m,n)=1$ is $\frac{6}{\pi^2}$.*
- i.e. $\dfrac{\operatorname{card}\{m,n : m,n\le X, (m,n)=1\}}{X^2}\to\dfrac{6}{\pi^2}$ as $X\to\infty$.
- **Proof.** $\displaystyle\sum_{n\le X}\phi(n)=\sum_{n\le X}\sum_{\substack{m\le n\\(m,n)=1}}1$
$= \operatorname{card}\{m,n : m\le n\le X, (m,n)=1\}$
$= \dfrac{1}{2}\operatorname{card}\{m,n : m,n\le X, (m,n)=1\}+\dfrac{1}{2}$ since if $m>1$, then $(m,m)=m>1$.
- Thus $\dfrac{1}{X^2}\operatorname{card}\{m,n : m,n\le X, (m,n)=1\}=$
$-\dfrac{1}{X^2}+\dfrac{2}{X^2}\displaystyle\sum_{n\le X}\phi(n)$.

- **Theorem 7.12.** *Suppose that $X \in \mathbb{R}$ and $X \ge 2$. Then*
$$\sum_{n\le X}\phi(n) = \frac{X^2}{2}\sum_{m=1}^{\infty}\frac{\mu(m)}{m^2} + O(X\log X).$$
- There is a curious application of this.
Theorem 7.13. *The probability that $(m,n)=1$ is $\frac{6}{\pi^2}$.*
- i.e. $\dfrac{\mathrm{card}\{m,n : m,n\le X,(m,n)=1\}}{X^2} \to \dfrac{6}{\pi^2}$ as $X\to\infty$.
- **Proof.** $\displaystyle\sum_{n\le X}\phi(n) = \sum_{n\le X}\sum_{\substack{m\le n\\(m,n)=1}} 1$
$= \mathrm{card}\{m,n : m\le n\le X,(m,n)=1\}$
$= \dfrac{1}{2}\mathrm{card}\{m,n : m,n\le X,(m,n)=1\} + \dfrac{1}{2}$ since if $m>1$, then $(m,m)=m>1$.
- Thus $\dfrac{1}{X^2}\mathrm{card}\{m,n : m,n\le X,(m,n)=1\} =$
$-\dfrac{1}{X^2} + \dfrac{2}{X^2}\displaystyle\sum_{n\le X}\phi(n).$
- and the result follows from the previous theorem.

- It is sometimes useful to know something about the way that an arithmetical function grows.

- It is sometimes useful to know something about the way that an arithmetical function grows.
- Multiplicative functions tend to oscillate quite a bit in size.

- It is sometimes useful to know something about the way that an arithmetical function grows.
- Multiplicative functions tend to oscillate quite a bit in size.
- For example $d(p) = 2$ but if we take n to be the product of the first k primes where k is large, then

$$d(n) = 2^k.$$

- It is sometimes useful to know something about the way that an arithmetical function grows.
- Multiplicative functions tend to oscillate quite a bit in size.
- For example $d(p) = 2$ but if we take n to be the product of the first k primes where k is large, then

$$d(n) = 2^k.$$

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

- It is sometimes useful to know something about the way that an arithmetical function grows.
- Multiplicative functions tend to oscillate quite a bit in size.
- For example $d(p) = 2$ but if we take n to be the product of the first k primes where k is large, then

$$d(n) = 2^k.$$

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

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

- It is sometimes useful to know something about the way that an arithmetical function grows.
- Multiplicative functions tend to oscillate quite a bit in size.
- For example $d(p) = 2$ but if we take n to be the product of the first k primes where k is large, then

$$d(n) = 2^k.$$

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

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

- Note, such a statement is often written as

$$d(n) = O_{\varepsilon}(n^{\varepsilon}) \text{ or } d(n) \ll_{\varepsilon} n^{\varepsilon}.$$

- **Theorem 7.14.** *Let $\varepsilon > 0$. Then there is a positive number C which depends at most on ε such that for every $n \in \mathbb{N}$ we have $d(n) < Cn^{\varepsilon}$.*

- **Theorem 7.14.** *Let $\varepsilon > 0$. Then there is a positive number C which depends at most on ε such that for every $n \in \mathbb{N}$ we have $d(n) < Cn^{\varepsilon}$.*
- It suffices to prove the theorem when $\varepsilon \le \frac{1}{\log 2}$.

- **Theorem 7.14.** *Let $\varepsilon > 0$. Then there is a positive number C which depends at most on ε such that for every $n \in \mathbb{N}$ we have $d(n) < Cn^{\varepsilon}$.*
- It suffices to prove the theorem when $\varepsilon \le \frac{1}{\log 2}$.
- Write $n = p_1^{k_1} \dots p_r^{k_r}$ where the p_j are distinct.

- **Theorem 7.14.** *Let $\varepsilon > 0$. Then there is a positive number C which depends at most on ε such that for every $n \in \mathbb{N}$ we have $d(n) < Cn^{\varepsilon}$.*
- It suffices to prove the theorem when $\varepsilon \le \frac{1}{\log 2}$.
- Write $n = p_1^{k_1} \dots p_r^{k_r}$ where the p_j are distinct.
- Recall $d(n) = (k_1 + 1) \dots (k_r + 1)$, so $\dfrac{d(n)}{n^{\varepsilon}} = \displaystyle\prod_{j=1}^{r} \frac{k_j + 1}{p_j^{\varepsilon k_j}}$.

- **Theorem 7.14.** *Let $\varepsilon > 0$. Then there is a positive number C which depends at most on ε such that for every $n \in \mathbb{N}$ we have $d(n) < Cn^{\varepsilon}$.*
- 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 $d(n) = (k_1 + 1) \dots (k_r + 1)$, so $\dfrac{d(n)}{n^{\varepsilon}} = \displaystyle\prod_{j=1}^{r} \frac{k_j + 1}{p_j^{\varepsilon k_j}}$.
- Since we are only interested in an upper bound the terms for which $p_j^{\varepsilon} > 2$ can be thrown away since $2^k \geq k + 1$.

- **Theorem 7.14.** *Let $\varepsilon > 0$. Then there is a positive number C which depends at most on ε such that for every $n \in \mathbb{N}$ we have $d(n) < Cn^{\varepsilon}$.*
- It suffices to prove the theorem when $\varepsilon \le \frac{1}{\log 2}$.
- Write $n = p_1^{k_1} \dots p_r^{k_r}$ where the p_j are distinct.
- Recall $d(n) = (k_1 + 1) \dots (k_r + 1)$, so $\dfrac{d(n)}{n^{\varepsilon}} = \displaystyle\prod_{j=1}^{r} \frac{k_j + 1}{p_j^{\varepsilon k_j}}$.
- Since we are only interested in an upper bound the terms for which $p_j^{\varepsilon} > 2$ can be thrown away since $2^k \ge k + 1$.
- However there are only $\le 2^{1/\varepsilon}$ primes p_j for which $p_j^{\varepsilon} \le 2$.

- **Theorem 7.14.** *Let $\varepsilon > 0$. Then there is a positive number C which depends at most on ε such that for every $n \in \mathbb{N}$ we have $d(n) < Cn^{\varepsilon}$.*
- It suffices to prove the theorem when $\varepsilon \le \frac{1}{\log 2}$.
- Write $n = p_1^{k_1} \dots p_r^{k_r}$ where the p_j are distinct.
- Recall $d(n) = (k_1+1)\dots(k_r+1)$, so $\dfrac{d(n)}{n^{\varepsilon}} = \displaystyle\prod_{j=1}^{r} \frac{k_j+1}{p_j^{\varepsilon k_j}}$.
- Since we are only interested in an upper bound the terms for which $p_j^{\varepsilon} > 2$ can be thrown away since $2^k \ge k+1$.
- However there are only $\le 2^{1/\varepsilon}$ primes p_j for which $p_j^{\varepsilon} \le 2$.
- Moreover for any such prime we have

$$p_j^{\varepsilon k_j} \ge 2^{\varepsilon k_j} = \exp(\varepsilon k_j \log 2) \ge 1 + \varepsilon k_j \log 2 \ge (k_j+1)\varepsilon \log 2.$$

- **Theorem 7.14.** *Let $\varepsilon > 0$. Then there is a positive number C which depends at most on ε such that for every $n \in \mathbb{N}$ we have $d(n) < Cn^{\varepsilon}$.*
- It suffices to prove the theorem when $\varepsilon \le \frac{1}{\log 2}$.
- Write $n = p_1^{k_1} \dots p_r^{k_r}$ where the p_j are distinct.
- Recall $d(n) = (k_1 + 1) \dots (k_r + 1)$, so $\dfrac{d(n)}{n^{\varepsilon}} = \displaystyle\prod_{j=1}^{r} \frac{k_j + 1}{p_j^{\varepsilon k_j}}$.
- Since we are only interested in an upper bound the terms for which $p_j^{\varepsilon} > 2$ can be thrown away since $2^k \ge k + 1$.
- However there are only $\le 2^{1/\varepsilon}$ primes p_j for which $p_j^{\varepsilon} \le 2$.
- Moreover for any such prime we have

$$p_j^{\varepsilon k_j} \ge 2^{\varepsilon k_j} = \exp(\varepsilon k_j \log 2) \ge 1 + \varepsilon k_j \log 2 \ge (k_j + 1)\varepsilon \log 2.$$

- Thus $\dfrac{d(n)}{n^{\varepsilon}} \le \left(\dfrac{1}{\varepsilon \log 2}\right)^{2^{1/\varepsilon}}$.