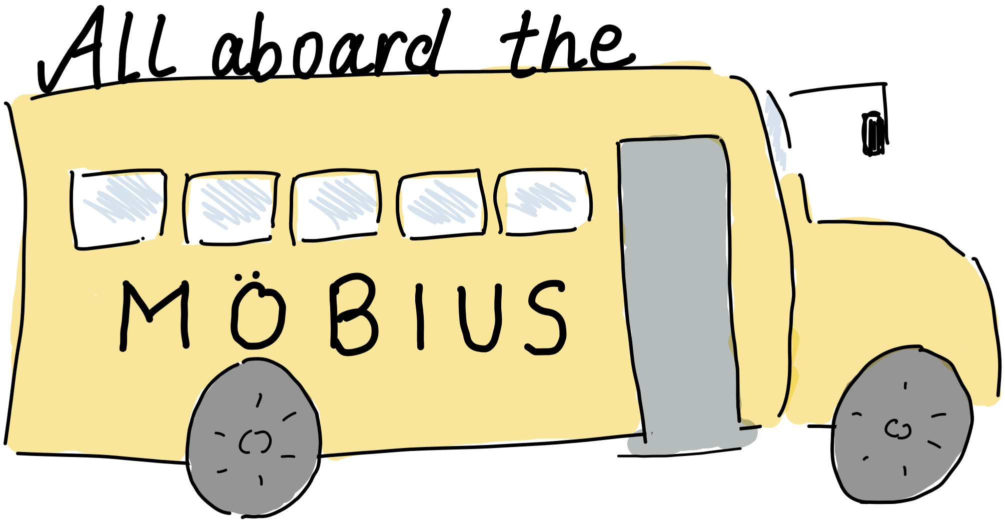


The Möbius function and her friends

Our star players are <u>ARITHMETIC FUNCTIONS</u>: functions from $\mathbb{N}$ to $\mathbb{C}$. Usually we only care about the ones that have nice number-theoretic properties. Let's start with the star of the show.

<u>Definition</u>: The <u>MÖBIUS FUNCTION</u> $\mu : \mathbb{N} \longrightarrow \mathbb{C}$ is

$$\mu(n) = \begin{cases} 1, & n = 1 \\ (-1)^k, & n = p_1 \cdots p_k \text{ is the product of } \underline{\underline{\text{distinct}}} \text{ primes} \\ 0, & \text{otherwise} \end{cases}$$

The first part of this class will focus on sums of the form $\sum_{d|n} f(d)$.

The Möbius function is fundamental to finding closed forms for these sums

$$\underset{\text{'d divides n'}}{\downarrow}$$

Proposition: $$\sum_{d|n} \mu(d) = \begin{cases} 1, & n = 1 \\ 0, & \text{otherwise} \end{cases}$$

Proof: Let n factor as $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$. Every divisor d of n has the form $d = p_1^{\beta_1} \cdots p_k^{\beta_k}$ for $0 \leq \beta_i \leq \alpha_i$. Of course, μ is only nonzero for squarefree integers, so

$$\sum_{d|n} \mu(d) = \sum_{I \subseteq \{p_1, \dots, p_k\}} (-1)^{|I|} = \begin{cases} 1, & \text{if } \{p_1, \dots, p_k\} = \emptyset \\ 0, & \text{otherwise} \end{cases}$$

every squarefree divisor comes from a subset of the prime factors

this uses a combinatorial fact that the number of even-sized subsets of a nonempty set is the same as the number of odd-sized subsets.

QED

Up next we have the <u>EULER TOTIENT FUNCTION</u> $\varphi(n)$, which counts the number of positive integers $\leq n$ that are coprime to n. For example, $\varphi(1) = 1$

$\varphi(2) = 1$

$\varphi(3) = 2$

$\varphi(4) = 2$

> Why do we care?
>
> If you've seen group theory, this is the order of the multiplicative group $\mathbb{Z}_n^\times$!

It turns out that $\varphi(n)$, like $\mu(n)$, can be computed using the prime factorization of n.

Step 1 : Find $\varphi(p^k)$ when p is prime and $k \geq 1$.

↳ The only positive integers not coprime to p^k are multiples of p. So,

$\varphi(p^k)$ = # integers between 1 and p^k that are not multiples of p

$= p^k$ − # integers that are multiples of p

$= p^k - p^{k-1}$.

Step 2: If m and n are coprime, $\varphi(mn) = \varphi(m)\varphi(n)$

↳ We will prove this by looking at modular arithmetic.

Define a function $F : \mathbb{Z}_{mn}^{\times} \longrightarrow \mathbb{Z}_m^{\times} \oplus \mathbb{Z}_n^{\times}$

$F : a \longrightarrow (a \bmod m,\ a \bmod n)$

(i) F is surjective

↳ Choose any $a \in \mathbb{Z}_m^{\times}$ and $b \in \mathbb{Z}_n^{\times}$. Let $c = (an + bm) \bmod mn$. Now, $an + bm$ is coprime to both m and n, so it is coprime to mn (since m and n are coprime). So, $F(c) = (a, b)$

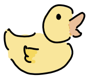
(ii) F is injective

↳ This is the tricky part. Suppose $F(a) = F(b)$, so $a \equiv b \bmod m$ and $a \equiv b \bmod n$. Then, $m \mid a-b$ and $n \mid a-b$. Since m and n are coprime, their product mn also divides $a - b \Rightarrow a \equiv b \bmod mn$.

Step 3 : Profit!

If $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$, by steps 1 and 2,

$\varphi(n) = (p_1^{\alpha_1} - p_1^{\alpha_1 - 1}) \cdots (p_k^{\alpha_k} - p_k^{\alpha_k - 1})$.



Hold on! What's the point of this class if we already have a formula for $\mu(n)$ AND $\varphi(n)$?

Great question! Both formulas require factorizing n, and ...

prime factorization is hard!

Instead of worrying about closed forms for arithmetic functions (often impractical), we will focus on finding relations between arithmetical functions.

Theorem. $\sum_{d|n} \varphi(d) = n.$

Proof. This is a very elegant combinatorial proof. For each divisor d of n, define

$A(d) = \{ k : \gcd(k,n) = d \text{ and } 1 \leq k \leq n \}$.

The sets $A(d)$ partition $\{1, \dots, n\}$. What is the size of $A(d)$?

Well, $\gcd(k,n) = d \iff \gcd(k/d, n/d) = 1$.

So, $|A(d)| = \varphi(n/d)$. Can you finish the proof from here?

QED

This is all well and good, but why are we trying to express n (a number we know) using $\varphi(d)$ (many numbers we don't know)?

Theorem: $\varphi(n) = \sum_{d|n} \mu(d) \, n/d$

Proof: Introduce a function on all REAL numbers,

$$I(x) = \begin{cases} 1, & x = 1 \\ 0, & \text{otherwise.} \end{cases}$$

We've seen this function before! $\sum_{d|n} \mu(d) = I(n)$.

Now we do a bunch of rewriting,

$$\varphi(n) = \sum_{k=1}^{n} I\left(\frac{1}{\gcd(n,k)}\right)$$

$$= \sum_{k=1}^{n} \sum_{d|\gcd(n,k)} \mu(d) = \sum_{k=1}^{n} \sum_{\substack{d|k \\ d|n}} \mu(d)$$

The condition '$d|k$ and $d|n$' is equivalent to '$k = qd$ for some $1 \leqslant q \leqslant n/d$'

$$= \sum_{d|n} \sum_{q=1}^{n/d} \mu(d)$$

$$= \sum_{d|n} \mu(d)\, {}^{n}\!/_{d}$$

These two theorems are a preview of the more general MÖBIUS INVERSION.

Day 2

Dirichlet convolution & Möbius inversion

We are going to see lots of sums that look like $\sum_{d|n} \mu(d)\, n/d$.

Definition. The DIRICHLET CONVOLUTION of two arithmetic functions f and g is

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

Think of convolution as a kind of multiplication. The intuition for this will become clear when we talk about Dirichlet series.

Examples For convenience, let $\mathbb{1}$ be the arithmetic function $\mathbb{1}(n) = 1 \;\; \forall n$, and N the function such that $N(n) = n \;\; \forall n$. We have already seen that

(1) $I = \mu * \mathbb{1}$

(2) $N = \varphi * \mathbb{1}$

(3) $\varphi = \mu * N$

It turns out that (2) and (3) are no coincidence : the $*$-operation is a group operation and the Möbius function helps invert it .

Exercise : Show that (i) $*$ is commutative : $f*g = g*f$

(ii) $*$ is associative : $(f*g)*h = f*(g*h)$

(iii) $*$ has an identity : $f * I = I * f = f$ for all f .

↓ parts (i) and (iii) are worth working through

All that remains is to show that $*$ has inverses. The catch here is that only arithmetic functions that satisfy $f(1) \neq 0$ will be $*$-invertible. You can actually find a formula for f^{-1} by recursion.

Exercise Show that if $f(1) \neq 0$, then the function f^{-1} defined by

$$f^{-1}(1) = 1/f(1) \quad \text{and}$$

$$f^{-1}(n) = \frac{-1}{f(1)} \sum_{\substack{d \mid n \\ d < n}} f\left(\frac{n}{d}\right) f^{-1}(d) \quad \text{for } n > 1$$

is a $\star$-inverse for f. That is, $f \star f^{-1} = f^{-1} \star f = I$

↓ optional! Only if you like Computations!

As you now know, we don't care for formulas! We want nice-looking relations!

Theorem (Möbius inversion formula) If $f(n) = \sum_{d \mid n} g(d)$, then $g(n) = \sum_{d \mid n} f(d)\mu(n/d)$.

Proof. This is now simple to show using the properties of Dirichlet convolution!

If $f = g \star \mathbb{1}$, then

$$\begin{aligned} f \star \mu &= (g \star \mathbb{1}) \star \mu \\ &= g \star (\mathbb{1} \star \mu) \\ &= g \end{aligned}$$

by associativity

since $\mathbb{1}$ and μ are inverses

QED

Multiplicative functions

Let's stay with the nice arithmetic functions for a minute. We had an easy time deriving a formula for $\varphi(n)$ once we knew its value at prime powers.

Definition. An arithmetic function $f(n)$ is MULTIPLICATIVE if

$$f(m)f(n) = f(mn) \quad \text{whenever } \gcd(m,n) = 1.$$

Exercise. (a) Show that the Möbius function $\mu(n)$ is multiplicative.

(b) Show that if f is multiplicative, then $f(1)=1$.

We like multiplicative functions because they preserve the multiplicative structure of the integers. They also behave well with Dirichlet convolution.

Exercise. (a) If f and g are multiplicative, so is $f\star g$.

(Trickier) (b) If f and $f\star g$ are multiplicative, so is g.

(Corollary) (c) If g is multiplicative, so is its Dirichlet inverse g^{-1}.

Of course, the nicest arithmetic functions are COMPLETELY MULTIPLICATIVE: $f(m)f(n)=f(mn)$ for ALL $m,n\in\mathbb{N}$. A silly example is the function $\mathbb{1}(n)$.

WARNING. If f and g are completely multiplicative, $f\star g$ need not be completely multiplicative.

We do have a very nice description of the $\star$- inverse.

Theorem. Let f be multiplicative. Then, f is completely multiplicative IFF $f^{-1}(n)=\mu(n)f(n)$.

Proof. If f is completely multiplicative, then

$$(f\mu\star f)(n)=\sum_{d\mid n}\mu(d)f(d)f(n/d)=f(n)\sum_{d\mid n}\mu(d)=I(n).$$

Conversely, if $f^{-1}(n)=\mu(n)f(n)$, then

$$\sum_{d\mid n}\mu(d)f(d)f(n/d)=0 \quad \text{for all } n>1.$$

In particular, if $n=p^{\alpha}$, $\mu(1)f(1)f(p^{\alpha})+\mu(p)f(p)f(p^{\alpha-1})=0$.

$$\Rightarrow f(p^{\alpha})=f(p)f(p^{\alpha-1}).$$

By induction on α, this tells us that $f(p^\alpha) = f(p)^\alpha$, which is enough to conclude that f is completely multiplicative.

QED

Example. This theorem actually helps us compute φ^{-1}.

We know that $\varphi = \mu * N$, so $\varphi^{-1} = \mu^{-1} * N^{-1}$. Since N is completely multiplicative, $N^{-1} = \mu N$. So,

$$\varphi^{-1}(n) = \sum_{d|n} d\mu(d).$$

Example. LIOUVILLE'S FUNCTION $\lambda(n)$

Liouville's function is kind of like the Möbius function: $\lambda(n) = (-1)^{\alpha_1 + \cdots + \alpha_k}$, if $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$. This immediately tells us that λ is completely multiplicative. Its divisor sum is an indicator function for squares, i.e.,

$$\sum_{d|n} \lambda(d) = \begin{cases} 1, & n \text{ is square} \\ 0, & \text{otherwise} \end{cases}$$

The proof goes by checking the identity for prime powers, since $\lambda * \mathbb{1}$ is multiplicative. If $n = p^\alpha$,

$$\begin{aligned}(\lambda * \mathbb{1})(p^\alpha) &= \sum_{d|p^\alpha} \lambda(d) \\ &= \lambda(1) + \lambda(p) + \cdots + \lambda(p^{\alpha-1}) + \lambda(p^\alpha) \\ &= \begin{cases} 1, & \alpha \text{ is even} \\ 0, & \alpha \text{ is odd} \end{cases}\end{aligned}$$

The $\star$-inverse of λ is the indicator function for squarefree numbers; $\lambda^{-1}(n) = |\mu(n)|$. Check this using the theorem!

Example. The DIVISOR FUNCTION $\sigma(n)$

Define $\sigma(n) = \sum\limits_{d|n} d$. One use of $\sigma(n)$ is to test whether n is a perfect number. It also turns out to be multiplicative, since $\sigma = \mathbb{1} \star N$ is the convolution of multiplicative functions. Since N is completely multiplicative, we can use the same trick as earlier to find σ^{-1}.

$$\sigma^{-1} = \mathbb{1}^{-1} \star N^{-1} = \mu \star \mu N$$

$$\Rightarrow \sigma^{-1}(n) = \sum_{d|n} d\mu(d)\mu\left(\frac{n}{d}\right)$$

Funkier sums

Another neat trick to study arithmetic functions uses formal power series.

Definition. For an arithmetic function f and prime p, the p^{th} BELL SERIES of f is $f_p(x) = \sum\limits_{n \geq 0} f(p^n) x^n$.

These are especially useful for multiplicative functions; If f and g are multiplicative, then $f = g \iff f_p(x) = g_p(x)$ for all primes p.

Example. The Bell series of μ.

Since $\mu(p^n) = 0$ if $n \geq 2$, $\mu_p(x) = 1 - x$. Coincidentally, $\mu^{-1}(p^n) = \mathbb{1}(p^n) = 1$, and $\mathbb{1}_p(x) = \sum_{n\geq 0} x^n = 1/1-x$.

Example. The Bell series of φ

$$\varphi_p(x) = 1 + \sum_{n\geq 1} (p^n - p^{n-1}) x^n$$

$$= \sum_{n\geq 0} p^n x^n - x \sum_{n \geq 0} p^n x^n$$

$$= (1-x) \sum_{n\geq 0} (px)^n = \frac{1-x}{1-px}.$$

Theorem. For arithmetic functions f and g, let $h = f \star g$. Then, $h_p(x) = f_p(x) g_p(x)$ for all primes p. As a corollary, if f^{-1} is the $\star$-inverse of f, then $f_p(x)$ is invertible and $[f_p(x)]^{-1} = f^{-1}_p(x)$.

Proof. We need to check the power series identity by looking at the coefficients.

$h(p^n) = \sum_{k=0}^{n} f(p^k) g(p^{n-k})$ which is the coefficient of x^n in the product $f_p(x) g_p(x)$.

QED

A quick example. Let $\nu(n)$ count the number of distinct prime divisors of n. The function $f(n) = 2^{\nu(n)}$ is multiplicative, and its p^{th} Bell series is $f_p(x) = 1 + \sum_{n\geq 1} 2x^n = 1 + \frac{2x}{1-x} = \frac{1+x}{1-x}$. Since $\mu^2{}_p(x) = 1 + x$ and $\mathbb{1}_p(x) = 1/1-x$, $f_p(x) = \mu^2_p(x) \cdot \mathbb{1}_p(x) \Rightarrow 2^{\nu(n)} = \sum_{d \mid n} \mu^2(d)$.

Day 3

Averages of arithmetic functions

Though sums of the form $\sum_{d|n} f(d)$ are nice to study, we are often more interested in sums like $\sum_{n \leqslant x} f(n)$. For example, if $p(n)$ is the PRIME INDICATOR FUNCTION — $p(n) = 1$ if n is prime and $p(n) = 0$ otherwise — then the Prime Number Theorem estimates $\sum_{n \leqslant x} p(n)$. Today we're going to work through some estimates and introduce the VON MANGOLDT function

NOTATION :

→ $f(x) = O(g(x))$ if $|f(x)| \leqslant M g(x)$ for some $M > 0$ $\forall x$

→ $f(x) \sim g(x)$ if $\lim_{x \to \infty} f(x)/g(x) = 1$

→ $f(x) = o(g(x))$ if $\lim_{x \to \infty} f(x)/g(x) = 0$

→ For $s > 1$, the Riemann zeta function is $\zeta(s) = \sum_{n \geq 1} 1/n^s$

Theorem. (A bunch of facts)

① $$\sum_{n \leqslant x} \frac{1}{n} = \log x + C + O\left(\tfrac{1}{x}\right)$$

② $$\sum_{n \leqslant x} \frac{1}{n^s} = \frac{x^{1-s}}{1-s} + \zeta(s) + O\left(\frac{1}{x^s}\right) \quad \text{if } s > 1$$

③ $$\sum_{\boxed{n > x}} \frac{1}{n^s} = O(x^{1-s}) \quad \text{if } s > 1$$

note the change here!

④ $$\sum_{n \leqslant x} n^\alpha = \frac{x^{\alpha+1}}{\alpha+1} + O(x^\alpha) \quad \text{if } \alpha > 0$$

The average order of $\sigma(n)$

<u>Goal</u> $\sum_{n \leqslant x} \sigma(n) = \sum_{n \leqslant x} \sum_{d \mid n} d$

$d \mid n$ and $n \leqslant x$

$\Longleftrightarrow n = qd \leqslant x$

$$= \sum_{\substack{d,q \\ qd \leqslant x}} d$$

$qd \leqslant x$

$\Longleftrightarrow q \leqslant x$ and $d \leqslant x/q$

$$= \sum_{q \leqslant x} \sum_{d \leqslant \frac{x}{q}} d$$

estimate $\sum_{d \leqslant x/q} d$

$$= \sum_{q \leqslant x} \left[\tfrac{1}{2} \left(x/q \right)^2 + O\left(x/q \right) \right]$$

$$= x^2/2 \sum_{q \leqslant x} 1/q^2 + O\left(x \sum_{q \leqslant x} 1/q \right)$$

apply the theorem to each sum

$$= x^2/2 \left[\frac{-1}{x} + \zeta(2) + O(x^{-2}) \right] + O\left(x \log x + Cx + 1 \right)$$

simplify (eat constants)

$$= -\frac{x}{2} + \zeta(2)\, x^2/2 + O(x \log x)$$

$x \log x$ eats $-\frac{x}{2}$

$$= \frac{\zeta(2)}{2} x^2 + O(x \log x)$$

<u>Theorem.</u> $\sigma(n) \sim \frac{\zeta(2)}{2} x^2$

The same trick can be used to show that $\sum_{n \leqslant x} d(n) \sim x \log x$,

where $d(n)$ is the DIVISOR COUNTING FUNCTION $d(n) = \sum_{d \mid n} 1$.

For the averages of $\varphi(n)$, we need to assume that

$$\sum_{n\geq 1}\frac{\mu(n)}{n^2} = 1/\zeta(2), \quad \text{so} \quad \sum_{n\leqslant x}\mu(n)/n^2 = 1/\zeta(2) + O(1/x)$$

Then, $\displaystyle\sum_{n\leqslant x}\varphi(n) = \frac{1}{2\zeta(2)}x^2 + O(x\log x)$

Partial sums and Dirichlet convolution

Let $h = f\star g$. What is the relationship between their partial sums?

Defn. If α is an arithmetic function and $F:\mathbb{R}\longrightarrow\mathbb{C}$ is any function,

$$(\alpha\circ F)(x) = \sum_{n\leqslant x}\alpha(n)F(x/n)$$

Lemma. If α and β are arithmetic functions and $F:\mathbb{R}\longrightarrow\mathbb{C}$,

$$(\alpha\star\beta)\circ F = \alpha\circ(\beta\circ F)$$

proof.

$$\begin{aligned}
(\alpha\star\beta)\circ F(x) &= \sum_{n\leqslant x}(\alpha\star\beta)(n)F(x/n)\\
&= \sum_{n\leqslant x}\sum_{d\mid n}\alpha(d)\beta(n/d)F(x/n)\\
&= \sum_{dk\leqslant x}\alpha(d)\beta(k)F(x/dk)\\
&= \sum_{d\leqslant x}\alpha(d)\sum_{k\leqslant x/d}\beta(k)F\left(\frac{x/d}{k}\right)\\
&= \alpha\circ(\beta\circ F)(x)
\end{aligned}$$

QED

Theorem. Let $h = f * g$, and $F(x) = \sum_{n \leqslant x} f(n)$, $G(x) = \sum_{n \leqslant x} g(n)$ and $H(x) = \sum_{n \leqslant x} h(n)$. Then,

$$H(x) = \sum_{n \leqslant x} f(n)\, G\left(x/n\right) = \sum_{n \leqslant x} g(n)\, F(x/n).$$

Proof. Define $U(x) = \begin{cases} 0, & 0 \leqslant x < 1 \\ 1, & x \geqslant 1 \end{cases}$.

Then, $F = f \circ U$, $G = g \circ U$, and $H = h \circ U$, so using the lemma,
$H = h \circ U = (f * g) \circ U = f \circ (g \circ U) = f \circ G$.

QED

The point of all this is that divisor sums help us estimate partial sums!

Notation! The FLOOR of a real number x, $\lfloor x \rfloor$ is the greatest integer $\leqslant x$. The FRACTIONAL PART of x is $\{x\} = x - \lfloor x \rfloor$.

Corollary. $$\sum_{n \leqslant x} F\left(\frac{x}{n}\right) = \sum_{n \leqslant x} \sum_{d \mid n} f(d) = \sum_{n \leqslant x} f(n) \left\lfloor \frac{x}{n} \right\rfloor$$

Prove this as an exercise!

An estimate for $\mu(n)$

It turns out that estimating the partial sums of $\mu(n)$ is hard – Riemann Hypothesis hard. We can still get enough information to prove the Prime Number Theorem. The Prime Number Theorem is equivalent to the statement that $\sum_{n \leqslant x} \frac{\mu(n)}{n} \longrightarrow 0$ as $x \longrightarrow \infty$.

Using Dirichlet convolution, we can show that the sequence of partial sums is at least bounded: From the corollary,

$$\sum_{n \leqslant x} \mu(n) \lfloor x/n \rfloor = \sum_{n \leqslant x} \sum_{d \mid n} \mu(d) = 1.$$

Now,

$$x \left| \sum_{n \leqslant x} \frac{\mu(n)}{n} = \right| = \left| \sum_{n \leqslant x} \mu(n) \lfloor x/n \rfloor + \sum_{n \leqslant x} \mu(n) \{ x/n \} \right|$$

$$= \left| 1 + \sum_{n \leqslant x} \mu(n) \left\{ \frac{x}{n} \right\} \right|$$

$$\leqslant 1 + \{x\} + \sum_{2 \leqslant n \leqslant x} \left\{ \frac{x}{n} \right\}$$

$$\leqslant 1 + \{x\} + \lfloor x \rfloor - 1 = x$$

We just showed that $\left| \sum_{n \leqslant x} \frac{\mu(n)}{n} \right| \leqslant 1$ for all x.

The von Mangoldt function

An arithmetic function we'll use to derive estimates (like the Prime Number Theorem) is the NATURAL LOGARITHM $\log(n)$. (This is the base-e logarithm!)

Can we find an arithmetic function $g(n)$ so that $\log(n) = \sum_{d|n} g(d)$?

Well, with our nifty Möbius inversion,

$$g(n) = \sum_{d|n} \log d \, \mu(n/d)$$

Changing variables

$$= \sum_{d|n} \log(n/d) \, \mu(d)$$

$\log(n/d) = \log(n) - \log(d)$

$$= \log(n) \sum_{d|n} \mu(d) - \sum_{d|n} \log(d) \mu(d)$$

if $n = 1$, $\log(n) = 0$

if $n \neq 1$, $\sum_{d|n} \mu(d) = 0$

$$= -\sum_{d|n} \log(d) \mu(d)$$

Combinatorics comes to our rescue once again to simplify this sum:

If $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$, then

$$\sum_{d|n} \log(d) \mu(d) = \sum_{I \subseteq [k]} (-1)^{|I|} \log\left(\prod_{p \in I} p\right)$$

$$= \sum_{I \subseteq [k]} (-1)^{|I|} \sum_{p \in I} \log(p)$$

$$= \sum_{i=1}^{k} \log(p_i) \sum_{I \ni p_i} (-1)^{|I|}$$

The subsets of $\{p_1, \dots, p_k\}$ containing p_i are in bijection with subsets of $\{p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_k\}$. If $k \geq 2$, this becomes a nontrivial sum over the subsets of a nonempty set, so $\sum_{I \ni p_i} (-1)^{|I|} = 0 \quad \forall i.$

What if $k = 1$?

Then, $n = p^\alpha$ for some prime p, and

$$g(n) = \sum_{d|n} -\mu(d)\log(d) = -\mu(p)\log(p) = \log(p)$$

<u>Definition</u>: The VON MANGOLDT function Λ is

$$\Lambda(n) = \begin{cases} \log(p), & \text{if } n = p^\alpha \\ 0, & \text{otherwise} \end{cases}$$

We derived this function using Möbius inversion on $\log(n)$, but it turns out to be useful for many reasons! For example, the derivative of the RIEMANN-ZETA function can be expressed using log:

$$d/ds\, \zeta(s) = \sum_{n\geq 1} \log(n)/n^s.$$

The von Mangoldt function plays the same role for the LOGARITHMIC DERIVATIVE of $\zeta(s)$.

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n\geq 1} \Lambda(n)/n^s.$$

We can use this to estimate $\sum\limits_{n \leq x} \Lambda(n) \lfloor x/n \rfloor = \sum\limits_{n \leq x} \sum\limits_{d|n} \Lambda(d)$

$$= \sum_{n \leq x} \log(n)$$

$$= \log(\lfloor x \rfloor !)$$

Tomorrow, we'll derive the more complicated SELBERG IDENTITY involving $\Lambda(n)$ and handwave how PNT follows from it.

Day 4

Today we will focus on the **Prime Number Theorem**

If $\pi(x) = \#$ of primes $\leq x$, then $\pi(x) \sim x/\log x$

Stop 1: introducing our key players

We have our arithmetic functions $\mu(n)$ and $\Lambda(n)$.

Definition. CHEBYSHEV'S ψ- FUNCTION is $\psi(x) = \sum_{n \leq x} \Lambda(n)$.

CHEBYSHEV'S ϑ- FUNCTION is $\vartheta(x) = \sum_{\substack{p \leq x \\ p \text{ prime}}} \log(p)$

Of course, we have $$\mathbb{p}(n) = \begin{cases} 1, & n \text{ is prime} \\ 0, & \text{otherwise} \end{cases}$$

and $\pi(x) = \sum_{n \leq x} \mathbb{p}(n)$. All three functions are very closely related

Theorem. A bunch of facts

① $\pi(x) = \sum_{n \leq x} \mathbb{p}(n)$, $\vartheta(x) = \sum_{n \leq x} \mathbb{p}(n) \log n$, $\psi(x) = \sum_{m \leq \log_2 x} \vartheta(x^{1/m})$

② $\lim_{x \to \infty} \left| \frac{\psi(x)}{x} - \frac{\vartheta(x)}{x} \right| = 0$

③ $\pi(x) \sim \frac{x}{\log x}$ IFF $\vartheta(x) \sim x$ IFF $\psi(x) \sim x$

Proof (with some details omitted)

① The formula for $\psi(x)$ is the only nontrivial one

$$\psi(x) = \sum_{n \leqslant x} \Lambda(n) = \sum_{m \geq 1} \sum_{\substack{p^m \leqslant x \\ p \text{ prime}}} \log p$$

$$= \sum_{m \leqslant \log_2 x} \sum_{\substack{p \leqslant x^{1/m} \\ p \text{ prime}}} \log p$$

$$= \sum_{m \leqslant \log_2 x} \vartheta(x^{1/m}) .$$

② From ①, $0 \leqslant \psi(x) - \vartheta(x) \leqslant \sum_{2 \leqslant m \leqslant \log_2 x} \vartheta(x^{1/m})$

Now we use the silly bound $\vartheta(x) \leqslant x \log x$,

$$\Rightarrow \quad 0 \leqslant \psi(x) - \vartheta(x) \leqslant \sum_{2 \leqslant m \leqslant \log_2 x} x^{1/m} \log x^{1/m}$$

$$\leqslant \log_2 x \cdot x^{1/2} \log x^{1/2}$$

Now divide by x :)

③ (A sketch!) Since $\pi(x) = \sum_{n \leqslant x} \mathbb{P}(n)$ and $\vartheta(x) = \sum_{n \leqslant x} \mathbb{P}(n) \log n$, a secret result from analysis tells us that

$$\vartheta(x) = \pi(x) \log x - \pi(1) \log(1) - \int_1^x \frac{\pi(t)}{t} \, dt$$

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

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

$$= \frac{\vartheta(x)}{\log x} + \int_{2}^{x} \frac{\vartheta(t)}{t \log^2 t}\, dt$$

Part ② tells us that $\vartheta(x) \sim x$ IFF $\psi(x) \sim x$. To show the equivalence for $\pi(x)$ and $\vartheta(x)$ involves showing that the integral error terms we derived disappear <u>if</u> any one of the asymptotics is true

Stop 2 : Investigating $\Lambda(n)$

Now that we've established that the Prime Number Theorem is equivalent to the statement that $\sum_{n \leq x} \Lambda(n) \sim x$, let's investigate $\Lambda(n)$.

We already saw that $\sum_{n \leq x} \Lambda(n) \lfloor x/n \rfloor = \log(\lfloor x \rfloor !)$. Using the Stirling approximation to the factorial, we can turn this into an estimate $\sum_{n \leq x} \Lambda(n) \lfloor x/n \rfloor = x \log x - x + O(\log x)$. We also have our theorem that

$$\sum_{n \leq x} f(n) \lfloor x/n \rfloor = \sum_{n \leq x} F(x/n), \text{ so } \sum_{n \leq x} \psi(x/n) = x \log x - x + O(\log x)$$

Our big result to estimate $\Lambda(n)$ has a much longer proof.

Notation : $f(x) = \Theta(g(x))$ if $f(x) = O(g(x))$ AND $g(x) = O(f(x))$.
For example $x = \Theta(2x)$.

Theorem. $\sum_{n \leq x} \frac{\Lambda(n)}{n} = \log x + O(1)$ and $\psi(x) = \Theta(x)$

Proof sketch The bound $x = O(\psi(x))$ is tricky to show. Here's a sketch of how the other bound works.

Let $T(x) = \sum_{n \leq x} \Lambda(n) \lfloor x/n \rfloor$. We know that

$$T(x) = x \log x - x + O(\log x) = x \log x + O(x)$$

Some rearranging of terms tells us that $T(x) - 2T(x/2) \geq \psi(x) - \psi(x/2)$.

We use our crude estimate for $T(x)$ to conclude that $\psi(x) - \psi(x/2) = O(x)$, and recursion to show that $\psi(x) = O(x)$.

Now, we use this to estimate $\sum_{n \leq x} \frac{\Lambda(n)}{n}$.

$$x \sum_{n \leq x} \frac{\Lambda(n)}{n} = \sum_{n \leq x} \left(\left\lfloor \frac{x}{n} \right\rfloor + O(1) \right) \Lambda(n)$$

$$= \sum_{n \leq x} \Lambda(n) \lfloor x/n \rfloor + O\left(\sum_{n \leq x} \Lambda(n) \right)$$

$$= T(x) + O(\psi(x))$$

$$= x \log x + O(x)$$

And dividing by x completes the proof.

Stop 3 : Selberg's asymptotic formula

This is the last meaty thing we will see before the 'proof' of PNT

We will derive Selberg's asymptotic formula using another kind of inversion result.

Theorem. If $G(x) = \log x \sum_{n \leq x} F(x/n)$, then

$$F(x)\log x + \sum_{n \leq x} F(x/n)\,\Lambda(n) = \sum_{d \leq x} \mu(d)\, G(x/d)$$

Proof. We look at each term on the left-hand side:

Since $I(n)$ is mostly zero ◡̈

$$F(x)\log x = \sum_{n \leq x} I(n)\, F(x/n) \log(x/n)$$

$$= \sum_{n \leq x} F(x/n) \log(x/n) \sum_{d \mid n} \mu(d)$$

Möbius inversion

and, $$\sum_{n \leq x} F(x/n)\,\Lambda(n) \overset{\text{Möbius inversion!}}{=} \sum_{n \leq x} F(x/n) \sum_{d \mid n} \mu(d) \log(n/d)$$

So, $$F(x)\log x + \sum_{n \leq x} F(x/n)\,\Lambda(n) = \sum_{n \leq x} F(x/n) \sum_{d \mid n} \mu(d) \left[\log(x/n) + \log(n/d)\right]$$

$$= \sum_{n \leq x} F(x/n) \sum_{d \mid n} \mu(d) \log x/d$$

$$= \sum_{d \leq x} \sum_{q \leq x/d} \mu(d) \log x/d \; F(x/qd)$$

we love rearranging divisor sums ◡̈

$$= \sum_{d \leq x} \mu(d)\, G(x/d)$$

QED

Finally,

Theorem. (SELBERG'S ASYMPTOTIC FORMULA)

$$\Psi(x)\log x + \sum_{n\leq x}\Lambda(n)\,\psi(x/n) = 2x\log x + O(x)$$

Proof. We are going to apply the previous theorem twice, with

$F_1(x) = \Psi(x)$ and $F_2(x) = x - C - 1$, where C is the constant from estimating $\sum_{n\leq x}\frac{1}{n} = \log x + C + O\left(\frac{1}{x}\right)$. So,

$$G_1(x) = \log x\sum_{n\leq x}\psi(x/n) = x\log^2 x - x\log x + O(\log^2 x)$$

from our Stirling estimate in step 2

$$G_2(x) = \log x\sum_{n\leq x}\left(\frac{x}{n} - C - 1\right) = x\log x\sum_{n\leq x}\frac{1}{n} - (C+1)\,x\log x$$

$$= x\log x\left[\log x + C + O\left(\frac{1}{x}\right)\right] - (C+1)\,x\log x$$

$$= x\log^2 x - x\log x + O(\log x).$$

Comparing the two, $G_1(x) - G_2(x) = O(\log^2 x)$, but we only need $G_1(x) - G_2(x) = O(x^{1/2})$

Now, we apply our inversion-type theorem

$$\left[F_1(x) - F_2(x)\right]\log x + \sum_{n \le x}\left[F_1(x/n) - F_2(x/n)\right]\Lambda(n)$$

$$= \sum_{d \le x} \mu(d)\left[G_1\left(x/d\right) - G_2\left(x/d\right)\right]$$

$$= \sum_{d \le x} \mu(d)\, O\left(\sqrt{x/d}\right)$$

$$= \sqrt{x}\, O\left(\sum_{d \le x} 1/\sqrt{d}\right)$$

from our estimate for $\sum_{n \le x} 1/n^{1/2}$

$$= O(x)$$

Of course, F_1 and F_2 need to be expressed in terms of ψ

$$\left[F_1(x) - F_2(x)\right]\log x + \sum_{n \le x}\left[F_1(x/n) - F_2(x/n)\right]\Lambda(n)$$

$$= \left(\psi(x) - x + C + 1\right)\log x + \sum_{n \le x}\left[\psi\left(x/n\right) - x/n + C + 1\right]\Lambda(n)$$

$$= \psi(x)\log x + \sum_{n \le x}\psi\left(x/n\right)\Lambda(n) - (x - C - 1)\log x - \sum_{n \le x}\left(\frac{x}{n} - C - 1\right)\Lambda(n)$$

Rearranging,

$$\psi(x)\log x + \sum_{n \le x}\psi\left(x/n\right)\Lambda(n) = (x \underbrace{- C - 1)\log x}_{\text{gets eaten}} + \sum_{n \le x}\left(\frac{x}{n} - C - 1\right)\Lambda(n) + O(x)$$

$$= x\log x + x\underbrace{\sum_{n \le x}\frac{\Lambda(n)}{n}} - (C+1)\underbrace{\sum_{n \le x}\Lambda(n)} + O(x)$$

$$= x\log x + x\left(\log x + \overset{\substack{\text{from our}\\ \text{estimate}}}{\overset{\downarrow}{O(1)}}\right) + \overset{\substack{\text{gets eaten since}\\ \psi(x) = O(x)}}{\overset{\downarrow}{O(x)}}$$

$$= 2x\log x + O(x)$$

QED

Stop 4 : Cleaning up

Here is how the prime number theorem ($\psi(x) \sim x$) follows.

Define a new function,

$$\tau(x) = e^{-x}\,\psi(e^{x}) - 1$$

So $\psi(x) \sim x$ IFF $\tau(x) \longrightarrow 0$ as $x \to \infty$

Some integration turns Selberg's formula into

$$|\tau(x)|\,x^2 \leqslant 2 \int_0^x \int_0^y |\tau(u)|\,du\,dy + O(x)$$

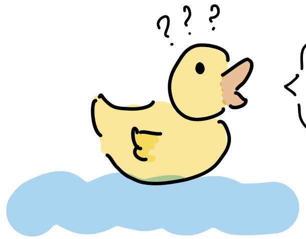
We look at the 'error term' of $\tau(x)$:

$$|\tau(x)| \leqslant \limsup_{x} |\tau(x)| + g(x), \text{ for some function } g(x) \longrightarrow 0.$$

Here is the hard part. If $\limsup_{x} |\tau(x)| > 0$, we manipulate the integral inequality to find some $0 < c_1 < \limsup_x |\tau(x)|$ so that

$|\tau(x)| \leqslant c_1 + g(x)$ still holds for all x.

But letting $x \to \infty$ in this new equation, $\limsup_x |\tau(x)| \leqslant c_1$, contradiction!



What just happened?

Let's take a step back. Our goal was the PRIME NUMBER THEOREM: that $\pi(x) \sim x/\log x$. Why is this equivalent to the statement that $\psi(x) \sim x$?

$$\text{We have } \psi(x) = \sum_{n \leqslant x} \Lambda(n) = \sum_{\substack{p^\alpha \leqslant x \\ p \text{ prime}}} \log p = \sum_{\substack{p \leqslant x \\ p \text{ prime}}} \left\lfloor \frac{\log x}{\log p} \right\rfloor \log p$$

$$\sim \log x \sum_{\substack{p \leqslant x \\ p \text{ prime}}} 1$$

$$= \log x \cdot \pi(x)$$

Looking at powers of $p \leqslant x$ and taking logarithms is about the same as counting primes $p \leqslant x$. The logarithm is just a nicer function to estimate than the prime indicator function $\mathbb{p}(n)$.

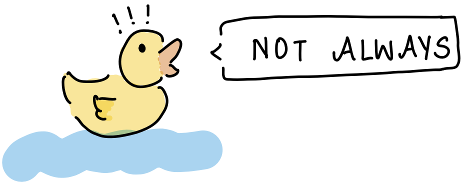
Since we only care about asymptotics, we can look at $\psi(e^x)$ to get rid of the logarithm. Our inversion-type results move around functions so we can get the bound in Selberg's asymptotic formula. Deriving the Prime Number Theorem just becomes some messy analysis to say that for the bound to make sense, $\psi(x)$ must be $\sim x$.

Day 5

Primes in arithmetic progression

How evenly are the primes distributed among the natural numbers? More specifically, can I find infinitely many primes in any arithmetic progression?

Let's reframe this more number-theoretically: an arithmetic progression $\{a + id : i \in \mathbb{N}\}$ is just the set of integers congruent to a mod d. So are there always infinitely many primes congruent to a mod d?



There are only finitely many primes congruent to 0 mod 2. Or 0 mod 4. Or 3 mod 6. In general, if $\gcd(a, d) > 1$, then only finitely many primes can live in that congruence class.

DIRICHLET'S THEOREM ON PRIMES IN ARITHMETIC PROGRESSIONS

Theorem. If $\gcd(k, n) = 1$, then there are infinitely many primes congruent to k mod n.

Let's prove this for a special case.

Claim. There are infinitely many primes congruent to 3 mod 4.

Proof. Suppose not. Let $p_1, \cdots, p_n$ be all the primes congruent to 3 mod 4.

Set $N = 4p_1 \cdots p_n + 3$. $N \equiv 3 \mod 4$, so it must have a prime divisor $p \equiv 3 \mod 4$, but p cannot be any of $p_1, \cdots, p_n$. Contradiction!

The argument for 1 mod 4 cannot use the same trick, but can still be proved with elementary number theory. The trick we just saw is the easiest way to show that there are infinitely many primes. What are other ways to show this?

Euler's famous result that $\sum_p \frac{1}{p}$ diverges is an extravagant way to show that there are infinitely many primes. But this is how we need to think to prove Dirichlet's theorem.

Goal: Show that $\sum_{\substack{p \equiv k \mod n \\ p \text{ prime}}} 1/p$ diverges.

Actually, it doesn't matter what the numerator is. As long as we pick any nonnegative function f, $\sum_{p \equiv k \mod n} f(p)/p$ diverges => there are infinitely many primes congruent to k mod n.

Theorem (Still Dirichlet) If $\gcd(n, k) = 1$, then

$$\sum_{\substack{p \leq x \\ p \equiv k \mod n}} \frac{\log p}{p} = \frac{1}{\varphi(n)} \log x + O(1)$$

This tells us something even stronger : the primes are evenly distributed among the congruence classes mod n. This is because on one hand,

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

and on the other,

$$\sum_{p \leq x} \frac{\log p}{p} = \underset{\substack{\downarrow \\ \text{there are} \\ \varphi(n) \text{ terms}}}{\sum_{\gcd(k,n)=1}} \underbrace{\sum_{\substack{p \leq x \\ p \equiv k \bmod n}} \frac{\log p}{p}}_{\hookrightarrow \text{ each term is } \sim \frac{1}{\varphi(n)} \log x}$$

This proof actually does involve some complex analysis.

<u>Definition</u>. A CHARACTER mod n is an arithmetic function $\chi : \mathbb{N} \to \mathbb{C}$ such that

① χ is completely multiplicative

② $\chi(m) = 0$ if $\gcd(m,n) > 1$.

③ $\chi(a) = \chi(a+n)$ for all integers a.

The characters mod n are determined entirely by their values mod n. In particular, they are only nonzero on the $\varphi(n)$ residue classes coprime to n. It turns out there are exactly $\varphi(n)$ distinct characters mod n, and they all map $\mathbb{Z}_n^\times : \{k \bmod n : \gcd(k,n) = 1\}$ to the $\varphi(n)^{th}$ roots of unity in $\mathbb{C}$.

Definition. If χ is a Dirichlet character mod n,

$$L(1,\chi) = \sum_{n\geq 1} \frac{\chi(n)}{n}.$$

The proof of Dirichlet's theorem hinges on showing that when χ is not identically 1, $L(1,\chi)$ converges BUT $L(1,\chi) \neq 0$.