Part III – Analytic Number Theory (Ongoing course, rough)

Based on lectures by Dr T. Bloom
Notes taken by Bhavik Mehta

Lent 2019

Contents

0 Introduction

Lecture 1 Analytic Number Theory is the study of numbers using analysis. It is a fascinating field because because a number - in particular in this course an integer - is discrete, whilst analysis involves the real/complex numbers which are continuous.

In this course, we will ask quantitative questions things like 'how many' or 'how large', in reference to simple number-theoretic objects.

Example.

1. How many primes? We can define the prime-counting function
$$\pi(x) = |\{n : n \leq x \text{ and } n \text{ is prime}\}|.$$
Then the prime number theorem, which we will prove in this course, states
$$\pi(x) \sim \frac{x}{\log x}.$$
(We will always take 'numbers' to mean natural numbers, not including zero).

2. How many twin primes (p such that $p+2$ is also prime) are there? It is not known whether there are infinitely many but since 2014, there has been immense progress by Zhang, Maynard and a Polymath project which has determined there are infinitely many primes at most 246 apart. Guess: there are $\approx \frac{x}{(\log x)^2}$ many twin primes $\leq x$.

3. How many primes are there congruent to $a \bmod q$ where $(a,q) = 1$. We know, by Dirichlet's theorem proven in the 20th century, that there are infinitely many such. The guess for how many there are in the interval $[1,x]$ is
$$\frac{1}{\varphi(q)} \frac{x}{\log x}.$$
This is known for small q. Recall that $\varphi(n) := |\{1 \leq m \leq n : (m,n) = 1\}|$, Euler's totient function.

The course will be split up into 4 (roughly equal) parts

1. Elementary techniques (real analysis)
2. Sieve methods
3. Riemann Zeta function, Prime Number Theorem (complex analysis)
4. Primes in arithmetic progressions

1 Elementary Techniques

We begin with a review of asymptotic notations:

- $f(x) = \mathcal{O}(g(x))$ if there is $C > 0$ such that $|f(x)| \leq C|g(x)|$ for all large enough x. (Landau notation)
- $f \ll g$ is the same as $f = \mathcal{O}(g)$ (Vinogradov notation)
- $f \sim g$ if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$ (i.e. $f = (1 + o(1))g$).
- $f = o(g)$ if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$

1.1 Arithmetic Functions

Definition. An **arithmetic function** is a function $f : \mathbb{N} \to \mathbb{C}$.

Definition. An important operation for multiplicative number theory is the **multiplicative convolution**

$$f \star g(n) := \sum_{ab=n} f(a)g(b).$$

Example.

- $1(n) := 1 \ \ \forall n$. Caution: $1 \star f \neq f$.
- Möbius function:
$$\mu(n) = \begin{cases} (-1)^k & \text{if } n = p_1 \cdots p_k \\ 0 & \text{if } n \text{ not squarefree} \end{cases}$$
- Liouville function:
$$\lambda(n) = (-1)^k \text{ if } n = p_1 \cdots p_k, \text{ not necessarily distinct}$$
- Divisor function:
$$\begin{aligned} \tau(n) &= |\{d \mid d \text{ a factor of } n\}| \\ \tau &= 1 \star 1 \end{aligned}$$

Definition (Multiplicative function)**.** An arithmetic function is a **multiplicative function** if $f(nm) = f(n)f(m)$ for $(n, m) = 1$. In particular, a multiplicative function is determined by its values on prime powers $f(p^k)$.

Fact.

- If f, g are multiplicative, then so is $f \star g$.
- $\log n$ is not multiplicative. $1, \mu, \lambda, \tau$ are multiplicative.

Note almost all arithmetic functions are not multiplicative.

Fact (Möbius inversion)**.**

$$1 \star f = g \iff \mu \star g = f.$$

Proof. First show

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

We have $1, \mu$ are multiplicative, so $1 \star \mu$ is multiplicative. Hence it is enough to check the identity for prime powers: If $n = p^k$, then $\{d : d \text{ divides } n\} = \{1, p, \dots, p^k\}$ so the left hand side is $1 - 1 + 0 + \dots + 0 = 0$, unless $k = 0$ when the left hand side is $\mu(1) = 1$.

The right hand side here is the identity of convolution, and convolution is associative, giving the required result. □

Our ultimate goal is to study the primes. This would suggest that we should work with the indicator function of the primes:

$$1_p(n) = \begin{cases} 1 & \text{if } n \text{ prime} \\ 0 & \text{otherwise.} \end{cases}$$

For example $\pi(x) = \sum_{1 \le n \le x} 1_p(n)$. This is an awkward function to work with. Instead, define the **von Mangoldt function**

$$\Lambda(n) = \begin{cases} \log p & \text{if } n \text{ is a prime power} \\ 0 & \text{otherwise} \end{cases}$$

i.e. weight the prime powers. This function is easier to use. Why?

Lemma.

$$1 \star \Lambda = \log \quad \text{and} \quad \mu \star \log = \Lambda$$

Proof. The second part follows immediately by Möbius inversion from the first.

$$1 \star \Lambda(n) = \sum_{d|n} \Lambda(d)$$

so write $n = p_1^{k_1} \dots p_k^{n_k}$,

$$\begin{aligned} &= \sum_{i=1}^{r} \sum_{j=1}^{k_i} \Lambda(p_i^j) \\ &= \sum_{i=1}^{r} \sum_{j=1}^{k_i} \log p_i \\ &= \sum_{i=1}^{r} k_i \log p_i = \sum_{i=1}^{r} \log p_i^{k_i} = \log n. \end{aligned}$$

□

Example. We can write

$$\begin{aligned} \Lambda(n) &= \sum_{d|n} \mu(d) \log\left(\frac{n}{d}\right) \\ &= \log n \sum_{d|n} \mu(d) - \sum_{d|n} \mu(d) \log d \\ &= -\sum_{d|n} \mu(d) \log d. \end{aligned}$$

$$
\begin{aligned}
\sum_{1\le n\le x} \Lambda(n) &= -\sum_{1\le n\le x}\sum_{d\mid n} \mu(d)\log d \\
&= -\sum_{d\le x} \mu(d)\log(d)\Big(\sum_{\substack{1\le n\le x\\ d\mid n}} 1\Big)
\end{aligned}
$$

but $\sum_{\substack{1\le n\le x\\ d\mid n}} 1 = \left\lfloor \frac{x}{d}\right\rfloor = \frac{x}{d} + \mathcal{O}(1)$, so

$$
= -x\sum_{d\le x}\mu(d)\frac{\log d}{d} + \mathcal{O}\Big(\sum_{d\le x}\mu(d)\log d\Big).
$$

1.2 Partial summation

Lecture 2 Given an arithmetic function, we can ask for estimates of $\sum_{n \le x} f(n)$, which gives a rough idea of how large $f(n)$ is on average.

Definition. We say that f has **average order** g if

$$\sum_{1 \le n \le x} f(n) \sim xg(x).$$

Example. For example, if $f \equiv 1$,

$$\sum_{1 \le n \le x} f(n) = \lfloor x \rfloor = x + \mathcal{O}(1) \sim x$$

so average order of f is 1. Now take $f(n) = n$,

$$\sum_{1 \le n \le x} n \sim \frac{x^2}{2}$$

so the average order of n is $\frac{n}{2}$. The Prime Number Theorem is the statement that 1_p has average order $\frac{1}{\log x}$.

Lemma 1.1 (Partial summation)**.** If (a_n) is a sequence of complex numbers and f is such that f' is continuous, then

$$\sum_{1 \le n \le x} a_n f(n) = A(x)f(x) - \int_1^x A(t)f'(t)\,dt$$

where $A(x) = \sum_{1 \le n \le x} a_n$.

Proof. Suppose $x = N$ is an integer. Note that $a_n = A(n) - A(n-1)$. So

$$\sum_{1 \le n \le N} a_n f(n) = \sum_{1 \le n \le N} f(n)\left(A(n) - A(n-1)\right)$$

(note $A(0) = 0$)

$$= A(N)f(N) + \sum_{n=1}^{N-1} A(n)\left(f(n+1) - f(n)\right).$$

Now

$$f(n+1) - f(n) = \int_n^{n+1} f'(t)\,dt.$$

So

$$\begin{aligned} \sum_{1 \le n \le N} a_n f(n) &= A(N)f(N) - \sum_{n=1}^{N-1} f'(t)\,dt \\ &= A(N)f(N) - \int_1^N A(t)f'(t)\,dt \end{aligned}$$

where we set $A(n) = A(t)$ $\forall t \in [n, n+1)$. If $N > \lfloor x \rfloor$, i.e. x not an integer,

$$\begin{aligned} A(x)f(x) &= A(N)f(x) \\ &= A(N)\left(f(N) + \int_N^x f'(t)\,dt\right). \end{aligned}$$

□

Lemma 1.2.

$$\sum_{1\le n\le x}\frac{1}{n}=\log x+\gamma+\mathcal{O}\left(\tfrac{1}{x}\right)$$

Proof. Partial summation with $f(x)=\frac{1}{x}$ and $a_n=1$, so $A(x)=\lfloor x\rfloor$:

$$\sum_{1\le n\le x}\frac{1}{n}=\frac{\lfloor x\rfloor}{x}+\int_1^x\frac{\lfloor t\rfloor}{t^2}\,dt$$

recall $\lfloor t\rfloor=t-\{t\}$

$$\begin{aligned}
&=1+\mathcal{O}\left(\tfrac{1}{x}\right)+\int_1^x\frac{1}{t}\,dt-\int_1^x\frac{\{t\}}{t^2}\,dt\\
&=1+\mathcal{O}\left(\tfrac{1}{x}\right)+\log x-\int_1^\infty\frac{\{t\}}{t^2}\,dt+\underbrace{\int_x^\infty\frac{\{t\}}{t^2}\,dt}_{\le\int_x^\infty\frac{1}{t^2}\,dt\le\frac{1}{x}}\\
&=\gamma+\mathcal{O}\left(\tfrac{1}{x}\right)+\log x+\mathcal{O}\left(\tfrac{1}{x}\right)\\
&=\log x+\gamma+\mathcal{O}\left(\tfrac{1}{x}\right)
\end{aligned}$$

where $\gamma=1-\int_1^\infty\frac{\{t\}}{t^2}\,dt$. □

This γ is called Euler's constant (Euler-Mascheroni). $\gamma\approx 0.577\ldots$ but we don't know if γ is irrational or not.

Lemma 1.3.

$$\sum_{1\le n\le x}\log n=x\log x-x+\mathcal{O}(\log x).$$

Proof. Partial summation with $f(x)=\log x$, $a_n=1$, $A(x)=\lfloor x\rfloor$.

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

□

This is not really Number Theory - we haven't really used multiplication yet.

1.3 Divisor function

Recall that

$$\tau(n)=1\star 1(n)=\sum_{ab\mid n}1=\sum_{d\mid n}1$$

We will analyse how many divisors an integer has.

Theorem 1.4.

$$\sum_{1\le n\le x}\tau(n)=x\log x+(2\gamma-1)x+\mathcal{O}(x^{\frac{1}{2}})$$

So average order of τ is $\log x$.

Proof. Partial summation involves turning a sum $\sum a_n \rightsquigarrow \sum a_n f(n)$, but what does $\tau(\frac{1}{2})$ even mean? There is no continuous function to use.

Instead, play around with the definition:

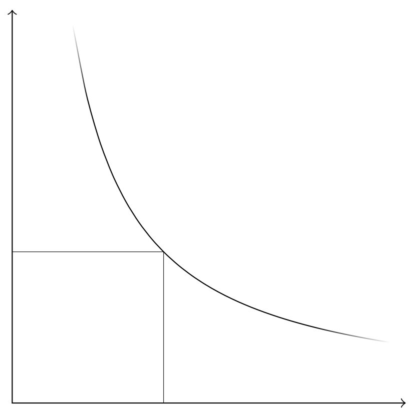
$$\begin{aligned}\sum_{1\leq n\leq x}\tau(n) &= \sum_{1\leq n\leq x}\sum_{d\mid x}1 \\ &= \sum_{1\leq d\leq x}\sum_{\substack{1\leq n\leq x\\ d\mid n}}1\end{aligned}$$

note that $\sum_{\substack{1\leq n\leq x\\ d\mid n}}1 = \lfloor\frac{x}{d}\rfloor$

$$\begin{aligned}&= \sum_{1\leq d\leq x}\left\lfloor\frac{x}{d}\right\rfloor \\ &= \sum_{1\leq d\leq x}\frac{x}{d} + \mathcal{O}(x) \\ &= x\sum_{1\leq d\leq x}\frac{1}{d} + \mathcal{O}(x) \\ &= x\log x + \gamma x + \mathcal{O}(x)\end{aligned}$$

using Lemma 1.2. To reduce the error term, we use (Dirichlet's) hyperbola trick.

$$\sum\tau(n) = \sum_{1\leq n\leq x}\sum_{ab=n}1 = \sum_{ab\leq x}1 = \sum_{a\leq x}\sum_{b\leq\frac{x}{a}}1$$



When summing over $ab\leq x$, we can sum over $a\leq x^{\frac{1}{2}}$, $b\leq x^{\frac{1}{2}}$ separately, and subtract the overlap.

$$\begin{aligned}\sum_{1\leq n\leq x}\tau(n) &= \sum_{a\leq x^{\frac{1}{2}}}\sum_{b\leq\frac{x}{a}}1 + \sum_{b\leq x^{\frac{1}{2}}}\sum_{a\leq\frac{x}{b}}1 - \sum_{a,b\leq x^{\frac{1}{2}}}1 \\ &= 2\sum_{a\leq x^{\frac{1}{2}}}\left\lfloor\frac{x}{a}\right\rfloor - \underbrace{\left\lfloor x^{\frac{1}{2}}\right\rfloor^2}_{=\left(x^{\frac{1}{2}}+\mathcal{O}(1)\right)^2} \\ &= 2\sum_{a\leq x^{\frac{1}{2}}}\frac{x}{a} + \mathcal{O}(x^{\frac{1}{2}}) - x + \mathcal{O}(x^{\frac{1}{2}}) \\ &= 2x\log x^{\frac{1}{2}} + 2\gamma x - x + \mathcal{O}(x^{\frac{1}{2}}) \\ &= x\log x + (2\gamma-1)x + \mathcal{O}(x^{\frac{1}{2}}).\end{aligned}$$

$\square$

Analytic Number Theory is mostly just controlling the error term.

Remark. Improving this $\mathcal{O}(x^{\frac{1}{2}})$ error term is a famous and hard problem! Probably, $\mathcal{O}(x^{\frac{1}{4}+\epsilon})$. The current best known is $\mathcal{O}(x^{0.3148})$.

This does not mean that $\tau(n) = \log n$: the average order does not give any information about specific values.

Lecture 3 **Theorem 1.5.** For any $n \ge 1$,

$$\tau(n) \le n^{\mathcal{O}\left(\frac{1}{\log\log n}\right)}.$$

In particular,

$$\tau(n) \ll_\epsilon n^\epsilon \ \forall \epsilon > 0$$

i.e. $\forall \epsilon > 0$, $\exists C(\epsilon) > 0$ such that $\tau(n) \le Cn^\epsilon$.

Proof. τ is multiplicative, so enough to calculate at prime powers. $\tau(p^k) = k+1$, so if $n = p_1^{k_1} \cdots p_r^{k_r}$ then

$$\tau(n) = \prod_{i=1}^r (k_i + 1).$$

Let $\epsilon > 0$ be chosen later and consider $\frac{\tau(n)}{n^\epsilon}$.

$$\frac{\tau(n)}{n^\epsilon} = \prod_{i=1}^r \frac{k_i+1}{p^{k_i\epsilon}}.$$

Note that as p is large, $\frac{k+1}{p^{k\epsilon}} \to 0$. In particular, if $p \ge 2^{\frac{1}{\epsilon}}$, then $\frac{k+1}{p^{k\epsilon}} \le \frac{k+1}{2^k} \le 1$.

What about small p? Can't do better than $p \ge 2$. In this case, $\frac{k+1}{p^{k\epsilon}} \le \frac{k+1}{2^{k\epsilon}} \le \frac{1}{\epsilon}$. Why? Rearrange to say $\epsilon k + \epsilon \le 2^{k\epsilon}$ (if $\epsilon \le \frac{1}{2}$), which follows from $x + \frac{1}{2} \le 2^x \ \forall x \ge 0$

So

$$\frac{\tau(n)}{n^\epsilon} \le \prod_{\substack{i=1 \\ p_i < 2^{\frac{1}{\epsilon}}}} \frac{k_i+1}{p^{k_i\epsilon}} \le \left(\frac{1}{\epsilon}\right)^{\pi(2^{\frac{1}{\epsilon}})} \le \left(\frac{1}{\epsilon}\right)^{2^{\frac{1}{\epsilon}}}.$$

Now choose optimal ϵ. (Trick: if you want to choose x to minimise $f(x) + g(x)$, choose x such that $f(x) = g(x)$).

So have,

$$\tau(n) \le n^\epsilon \epsilon^{-2^{\frac{1}{\epsilon}}} = \exp\left(\epsilon \log n + 2^{\frac{1}{\epsilon}} \log \frac{1}{\epsilon}\right).$$

Choose ϵ such that $\log n \approx 2^{\frac{1}{\epsilon}}$, i.e. $\epsilon \approx \frac{1}{\log\log n}$.

$$\tau(n) \le n^{\frac{1}{\log\log n}} (\log\log n)^{2^{\log\log n}} = n^{\frac{1}{\log\log n}} e^{(\log n)^{\log 2} \log\log\log n} \le n^{\mathcal{O}\left(\frac{1}{\log\log n}\right)}. \qquad \square$$

1.4 Estimates for the primes

Recall

$$\pi(x) = |\{ p \le x \}| = \sum_{1 \le n \le x} 1_p(n)$$

and

$$\psi(x) = \sum_{1 \le n \le x} \Lambda(n).$$

The Prime Number Theorem is $\pi(x) \sim \frac{x}{\log x}$ or equivalently $\psi(x) \sim x$. It was 1850 before the correct magnitude of $\pi(x)$ was proved. Chebyshev showed $\pi(x) \asymp \frac{x}{\log x}$, (where $f \asymp g$ means $g \ll f \ll g$).

Theorem 1.6 (Chebyshev)**.**

$$\psi(x) \asymp x$$

Proof. First we'll prove the lower bound, i.e. that $\psi(x) \gg x$.

$$\psi(x) = \sum_{n \le x} \Lambda(n).$$

$x \log x$ is a trivial upper bound for this, (each summand is $\le \log x$); we'd like to remove the factor of $\log x$. Recall $1 \star \Lambda = \log$, i.e.

$$\sum_{ab=n} \Lambda(a) = \log n.$$

The trick is to find a sum Σ such that $\Sigma \le 1$. We'll use the identity $\lfloor x \rfloor \le 2\lfloor \frac{x}{2} \rfloor + 1$, valid for $x \ge 0$. (Proof: Say $\frac{x}{2} = n + \theta$, with $\theta \in [0, 1)$, so $\lfloor \frac{x}{2} \rfloor = n$ then $x = 2n + 2\theta$ so $\lfloor x \rfloor = 2n$ or $2n + 1$.)

So

$$\psi(x) \ge \sum_{n \le x} \Lambda(n) \left(\lfloor \tfrac{x}{n} \rfloor - 2 \lfloor \tfrac{x}{2n} \rfloor \right).$$

Note $\lfloor \frac{x}{n} \rfloor = \sum_{m \le \frac{x}{n}} 1$

$$\begin{aligned}
. &= \sum_{n \le x} \Lambda(n) \sum_{m \le \frac{x}{n}} 1 - 2 \sum_{n \le x} \Lambda(n) \sum_{m \le \frac{x}{2n}} 1 \\
&= \sum_{mn \le x} \Lambda(n) - 2 \sum_{nm \le \frac{x}{2}} \Lambda(n) \\
&= \sum_{d \le x} 1 \star \Lambda(d) - 2 \sum_{d \le \frac{x}{2}} 1 \star \Lambda(d) \\
&= \sum_{d \le x} \log d - 2 \sum_{d \le \frac{x}{2}} \log d \\
&= x \log x - x + \mathcal{O}(\log x) - 2 \left(\frac{x}{2} \log \frac{x}{2} - \frac{x}{2} + \mathcal{O}(\log x) \right) \\
&= (\log 2)x + \mathcal{O}(\log x) \gg x.
\end{aligned}$$

For the upper bound, note $\lfloor x \rfloor = 2\lfloor \frac{x}{2} \rfloor + 1$ for $x \in (1, 2)$ so

$$\sum_{\frac{x}{2} < n < x} \Lambda(n) = \sum_{\frac{x}{2} < n < x} \Lambda(n) \left(\lfloor \tfrac{x}{n} \rfloor - 2 \lfloor \tfrac{x}{2n} \rfloor \right) \le \sum_{1 \le n \le x} \Lambda(n) \left(\lfloor \tfrac{x}{n} \rfloor - 2 \lfloor \tfrac{x}{2n} \rfloor \right)$$

Thus

$$\psi(x) - \psi\left(\frac{x}{2}\right) \le (\log 2)x + \mathcal{O}(\log x).$$

$$\begin{aligned}
\psi(x) &= \left(\psi(x) - \psi\left(\tfrac{x}{2}\right) \right) + \left(\psi(\tfrac{x}{2}) - \psi(\tfrac{x}{4}) \right) + \cdots \\
&\le \log 2 \left(x + \tfrac{x}{2} + \tfrac{x}{4} + \cdots \right) + \mathcal{O}((\log x)^2) \\
&= 2 \log 2x + \mathcal{O}((\log x)^2).
\end{aligned}$$

□

Lemma 1.7.

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

Proof. Recall $\log = 1 \star \Lambda$. So

$$\begin{aligned}\sum_{n\le x}\log n &= \sum_{ab\le x}\Lambda(a) = \sum_{a\le x}\Lambda(a)\sum_{b\le \frac{x}{a}} 1\\ &= \sum_{a\le x}\Lambda(a)\lfloor \tfrac{x}{a}\rfloor = x\sum_{a\le x}\frac{\Lambda(a)}{a} + \mathcal{O}(\psi(x))\\ &= x\sum_{a\le x}\frac{\Lambda(a)}{a} + \mathcal{O}(x)\end{aligned}$$

But from Lemma 1.3,

$$\sum_{n\le x}\log n = x\log x - x + \mathcal{O}(\log x)$$

$$\text{So } \sum_{n\le x}\frac{\Lambda(n)}{n} = \log x - 1 + \mathcal{O}(\tfrac{\log x}{x}) + \mathcal{O}(1) = \log x + \mathcal{O}(1).$$

Remains to note

$$\sum_{p\le x}\sum_{n=2}^{\infty}\frac{\log p}{p^k} = \sum_{p\le x}\log p\sum_{k=2}^{\infty}\frac{1}{p^k} = \sum_{p\le x}\frac{\log p}{p^2-p} \le \sum_{p=2}^{\infty}\frac{1}{p^{\frac{3}{2}}} = \mathcal{O}(1).$$

So

$$\sum_{n\le x}\frac{\Lambda(n)}{n} = \sum_{p\le x}\frac{\log p}{p} + \mathcal{O}(1).$$ □

Lecture 4 **Lemma 1.8.**

$$\pi(x) = \frac{\psi(x)}{\log x} + \mathcal{O}\left(\frac{x}{(\log x)^2}\right).$$

In particular, $\pi(x) \asymp \frac{x}{\log x}$ and the statement of the prime number theorem ($\pi(x) \sim \frac{x}{\log x}$) is equivalent to $\psi(x) \sim x$.

Proof. Idea is to use Partial summation:

$$\theta(x) \coloneqq \sum_{p\le x}\log p = \pi(x)\log x - \int_1^x \frac{\pi(t)}{t}\,dt$$

whereas

$$\psi(x) = \sum_{n\le x}\Lambda(n) = \sum_{p^k\le x}\log p.$$

$$\psi(x) - \theta(x) = \sum_{k=2}^{\infty}\sum_{p^k\le x}\log p = \sum_{k=2}^{\infty}\theta(x^{\frac{1}{k}}) \le \sum_{k=2}^{\log x}\psi(x^{\frac{1}{k}}) \ll \sum_{k=2}^{\log x} x^{\frac{1}{k}} \ll x^{\frac{1}{2}}\log x$$

Thus,

$$\begin{aligned}\psi(x) &= \pi(x)\log x + \mathcal{O}(x^{\frac{1}{2}}\log x) - \int_1^x\frac{\pi(t)}{t}\,dt\\ &= \pi(x)\log x + \mathcal{O}(x^{\frac{1}{2}}) + \mathcal{O}\left(\int_1^x \frac{1}{\log t}\,dt\right)\\ &= \pi(x)\log x + \mathcal{O}\left(\frac{x}{\log x}\right)\end{aligned}$$

where we used the fact that $\pi(t) \ll \frac{t}{\log t}$: Trivially, $\pi(t) \le t$, so

$$\psi(x) = \pi(x)\log x + \mathcal{O}(x^{\frac{1}{2}}\log x) + \mathcal{O}(x)$$

so $\pi(x)\log x = \mathcal{O}(x)$. $\square$

Lemma 1.9.

$$\sum_{p \le x} \frac{1}{p} = \log\log x + b + \mathcal{O}(\tfrac{1}{\log x})$$

where b is some constant.

Proof. We use partial summation. Let $A(x) = \sum_{p \le x} \frac{\log p}{p} = \log x + R(x)$ (and $R(x) \ll 1$).

$$\begin{aligned}\sum_{2 \le p \le x} \frac{1}{p} &= \frac{A(x)}{\log x} + \int_2^x \frac{A(t)}{t(\log t)^2}\,dt \\ &= 1 + \mathcal{O}(\tfrac{1}{\log x}) + \int_2^x \frac{1}{t\log t}\,dt + \int_2^x \frac{R(t)}{t(\log t)^2}\,dt\end{aligned}$$

Note $\int_2^\infty \frac{R(t)}{t(\log t)^2}\,dt$ exists, say it is c.

$$\begin{aligned}\sum_{2 \le p \le x} \frac{1}{p} &= 1 + c + \mathcal{O}(\tfrac{1}{\log x}) + \log\log x - \log\log 2 + \mathcal{O}\left(\int_x^\infty \frac{1}{t(\log t)^2}\right) \\ &= \log\log x + b + \mathcal{O}\left(\tfrac{1}{\log x}\right).\end{aligned}$$

$\square$

Theorem 1.10 (Chebyshev)**.** If

$$\pi(x) \sim c\frac{x}{\log x}$$

then $c = 1$.

Chebyshev also showed if $\pi(x) \sim \frac{x}{\log x - A(x)}$ then $A \sim 1$, which was a surprise since it was believed $A \sim 1.08\ldots$

Proof. Partial summation on $\sum_{p \le x} \frac{1}{p}$.

$$\sum_{p \le x} \frac{1}{p} = \frac{\pi(x)}{x} + \int_1^x \frac{\pi(t)}{t^2}\,dt.$$

If $\pi(x) = (c + o(1))\frac{x}{\log x}$ then

$$\begin{aligned}&= \frac{c}{\log x} + o\left(\frac{1}{\log x}\right) + (c + o(1))\int_1^x \frac{1}{t\log t}\,dt \\ &= \mathcal{O}\left(\frac{1}{\log x}\right) + (c + o(1))\log\log x.\end{aligned}$$

But $\sum_{p \le x} \frac{1}{p} = (1 + o(1))\log\log x$ by Lemma 1.9. Hence $c = 1$. $\square$

Lemma 1.11.

$$\prod_{p \le x} \left(1 - \frac{1}{p}\right)^{-1} = c\log x + \mathcal{O}(1)$$

where c is some constant.

Proof.

$$\begin{aligned}\log\left(\prod_{p\le x}\left(1-\frac{1}{p}\right)^{-1}\right) &= -\sum_{p\le x}\log\left(1-\frac{1}{p}\right)\\ &= \sum_{p\le x}\sum_{k}\frac{1}{kp^k}\\ &= \sum_{p\le x}\frac{1}{p}+\sum_{k\ge 2}\sum_{p\le x}\frac{1}{kp^k}\\ &= \log\log x + c' + \mathcal{O}\left(\frac{1}{\log x}\right).\end{aligned}$$

Now note that $e^x = 1 + \mathcal{O}(x)$ for $|x| \le 1$. So

$$\begin{aligned}\prod_{p\le x}\left(1-\frac{1}{p}\right)^{-1} = c\log x\ e^{\mathcal{O}(\frac{1}{\log x})} &= c\log x\ (1+\mathcal{O}(\tfrac{1}{\log x}))\\ &= c\log x + \mathcal{O}(1).\end{aligned}$$ ◻

It turns out that $c = e^\gamma = 1.78\dots$

1.4.1 Why is the Prime Number Theorem hard?

Let's try a probabilistic heuristic for the PNT: the 'probability' that $p \mid n$ is $\frac{1}{p}$. What is the 'probability' that n is prime?

$$n \text{ is prime} \iff n \text{ has no prime divisors } p \le n^{\frac{1}{2}}.$$

Make the guess that the events 'divisible by p' are independent, so $\mathbb{P}(p \nmid n) = 1 - \frac{1}{p}$.

$$\mathbb{P}(n \text{ is prime}) \approx \prod_{p\le n^{\frac{1}{2}}}\left(1-\frac{1}{p}\right) \approx \frac{1}{c\log n^{\frac{1}{2}}} = \frac{2}{c}\frac{1}{\log n}.$$

So

$$\pi(x) = \sum_{n\le x} 1_{n\text{ prime}} \approx \frac{2}{c}\sum_{n\le x}\frac{1}{\log n} \approx \frac{2}{c}\frac{x}{\log x} \approx 2e^{-\gamma}\frac{x}{\log x}.$$

But $2e^{-\gamma} \approx 1.122\dots$, so this heuristic says there are around 12% more primes than there are. This shows that heuristics might be good for order of magnitude estimates, but the constants may not be accurate.

Let's try another approach: Recall that $1 \star \Lambda = \log$ so $\mu \star \log = \Lambda$. So

$$\psi(x) = \sum_{n\le x}\Lambda(n) = \sum_{ab\le x}\mu(a)\log b = \sum_{a\le x}\mu(a)\left(\sum_{b\le \frac{x}{a}}\log b\right).$$

Recall that

$$\begin{aligned}\sum_{m\le x}\log m &= x\log x - x + \mathcal{O}(\log x)\\ \sum_{m\le x}\tau(m) &= x\log x + (2\gamma-1)x + \mathcal{O}(x^{\frac{1}{2}})\end{aligned}$$

Thus

$$\psi(x) = \sum_{a \leq x} \mu(a) \left(\sum_{b \leq \frac{x}{a}} \tau(b) - 2\gamma \frac{x}{a} + \mathcal{O}\left(\frac{x^{\frac{1}{2}}}{a^{\frac{1}{2}}} \right) \right)$$

Consider the first term, which has highest order

$$\begin{aligned} \sum_{ab \leq x} \mu(a)\tau(b) = \sum_{abc \leq x} \mu(a) &= \sum_{b \leq x} \sum_{ac \leq \frac{x}{b}} \mu(a) = \sum_{b \leq x} \sum_{d \leq \frac{x}{b}} \mu \star 1(d) \\ &= \lfloor x \rfloor = x + \mathcal{O}(1). \end{aligned}$$

This leaves an error term of

$$-2\gamma \sum_{a \leq x} \mu(a) \frac{x}{a} = \mathcal{O}\left(x \sum_{a \leq x} \frac{\mu(a)}{a} \right)$$

so we still need to show that $\sum_{a \leq x} \frac{\mu(a)}{a} = o(1)$. But this is in fact equivalent to the PNT.

1.5 Selberg's identity and an elementary proof of the PNT

Lecture 5 Recall that the statement of the prime number theorem is

$$\psi(x) = \sum_{n \leq x} \Lambda(n) = x + o(x).$$

Let

$$\Lambda_2(n) \coloneqq \mu \star \log^2(n) = \sum_{ab=n} \mu(a)(\log b)^2.$$

called **Selberg's function**. (To see why this is denoted Λ_2, recall that $\Lambda = \mu \star \log$). The idea is to prove a 'Prime Number Theorem for Λ_2' with elementary methods. In particular, we will try the same method as just before, but the leading order term will be larger, so the error term can safely be ignored.

Lemma 1.12.

(1) $\Lambda_2(n) = \Lambda(n) \log n + \Lambda \star \Lambda(n)$

(2) $0 \leq \Lambda_2(n) \leq (\log n)^2$

(3) If $\Lambda_2(n) \neq 0$ then n has at most 2 distinct prime divisors.

Proof. For (1), we use Möbius inversion, so it is enough to show that

$$\sum_{d \mid n} (\Lambda(d) \log d + \Lambda \star \Lambda(d)) = (\log n)^2.$$

Recall that $1 \star \Lambda = \log$, so

$$\begin{aligned}
\sum_{d|n}(\Lambda(d)\log d + \Lambda \star \Lambda(d)) &= \sum_{d|n} \Lambda(d)\log d + \sum_{ab|n} \Lambda(a)\Lambda(b) \\
&= \sum_{d|n} \Lambda(d)\log d + \sum_{a|n} \Lambda(a) \left(\sum_{b|\frac{n}{a}} \Lambda(b) \right) \\
&= \sum_{d|n} \Lambda(d)\log d + \sum_{d|n} \Lambda(d) \log\left(\frac{n}{d}\right) \\
&= \log n \sum_{d|n} \Lambda(d) = (\log n)^2.
\end{aligned}$$

For (2), $\Lambda_2(n) \geq 0$ since both terms on the RHS in (1) are ≥ 0 and since $\sum_{d|n} \Lambda_2(d) = (\log n)^2$ we get $\Lambda_2(n) \leq (\log n)^2$.

For (3), note that if n is divisible by 3 distinct primes, then $\Lambda(n) = 0$, and $\Lambda \star \Lambda(n) = \sum_{ab=n} \Lambda(a)\Lambda(b) = 0$ since at least one of a or b has ≥ 2 distinct prime divisors. □

Theorem 1.13 (Selberg's identity)**.**

$$\sum_{n \leq x} \Lambda_2(n) = 2x \log x + \mathcal{O}(x).$$

Proof.

$$\begin{aligned}
\sum_{n \leq x} \Lambda_2(n) &= \sum_{n \leq x} \mu \star (\log)^2(n) \\
&= \sum_{ab \leq x} \mu(a)(\log b)^2 \\
&= \sum_{a \leq x} \mu(a) \left(\sum_{(b \leq \frac{x}{a})} (\log b)^2 \right).
\end{aligned}$$

By Partial summation,

$$\sum_{m \leq x} (\log m)^2 = x(\log x)^2 - 2x\log x + 2x + \mathcal{O}((\log x)^2).$$

By Partial summation again, (with $A(t) = \sum_{n \leq t} \tau(n) = t\log t + Ct + \mathcal{O}(t^{\frac{1}{2}})$)

$$\begin{aligned}
\sum_{m \leq x} \frac{\tau(m)}{m} &= \frac{A(x)}{x} + \int_1^x \frac{A(t)}{t^2}\, dt \\
&= \log x + C + \mathcal{O}(x^{-\frac{1}{2}}) + \int_1^x \frac{\log t}{t}\, dt + c\int_1^x \frac{1}{t}\, dt + \mathcal{O}\left(\int_1^x \frac{1}{t^{\frac{3}{2}}}\, dt \right) \\
&= \frac{(\log x)^2}{2} + c_1 \log x + c_2 + \mathcal{O}(x^{-\frac{1}{2}}).
\end{aligned}$$

So

$$\frac{x(\log x)^2}{2} = \sum_{m \leq x} \tau(m)\frac{x}{m} + c_1' \sum_{m \leq x} \tau(m) + c_2' x + \mathcal{O}(x^{\frac{1}{2}})$$

so

$$\sum_{m \leq x} (\log m)^2 = 2\sum_{m \leq x} \tau(m)\frac{x}{m} + c_3 \sum_{m \leq x} \tau(m) + c_4 + \mathcal{O}(x^{\frac{1}{2}})$$

so

$$\sum_{n\le x}\Lambda_2(n) = 2\sum_{a\le x}\mu(a)\sum_{b\le\frac{x}{a}}\frac{\tau(b)x}{ab} + c_5\sum_{a\le x}\mu(a)\sum_{b\le\frac{x}{a}}\tau(b) + c_6\sum_{a\le x}\mu(a)\frac{x}{a} + \mathcal{O}\left(\sum_{a\le x}\frac{x^{\frac{1}{2}}}{a^{\frac{1}{2}}}\right).$$

Now, we show that the last three terms here are $\mathcal{O}(x)$: First, note that

$$x^{\frac{1}{2}}\sum_{a\le x}\frac{1}{a^{\frac{1}{2}}} = \mathcal{O}(x).$$

Secondly,

$$\begin{aligned}
x\sum_{a\le x}\frac{\mu(a)}{a} &= \sum_{a\le x}\left\lfloor\frac{x}{a}\right\rfloor + \mathcal{O}(x) \\
&= \sum_{a\le x}\sum_{b\le\frac{x}{a}} 1 + \mathcal{O}(x) \\
&= \sum_{d\le x}\mu\star 1(d) + \mathcal{O}(x) \\
&= 1 + \mathcal{O}(x) = \mathcal{O}(x).
\end{aligned}$$

Thirdly,

$$\begin{aligned}
\sum_{a\le x}\mu(a)\sum_{b\le\frac{x}{a}}\tau(b) &= \sum_{a\le x}\mu(a)\sum_{b\le\frac{x}{a}}\sum_{cd=b}1 \\
&= \sum_{a\le x}\mu(a)\sum_{cd\le\frac{x}{a}}1 \\
&= \sum_{acd\le x}\mu(a) \\
&= \sum_{d\le x}\sum_{ac\le\frac{x}{d}}\mu(a) \\
&= \sum_{d\le x}\sum_{e\le\frac{x}{d}}\mu\star 1(e) \\
&= \sum_{d\le x}1 = \mathcal{O}(x).
\end{aligned}$$

So

$$\begin{aligned}
\sum_{n\le x}\Lambda_2(n) &= 2\sum_{a\le x}\mu(a)\sum_{b\le\frac{x}{a}}\frac{\tau(b)x}{ab} + \mathcal{O}(x) \\
&= 2x\sum_{d\le x}\frac{1}{d}\mu\star\tau(d) + \mathcal{O}(x)
\end{aligned}$$

Recall $\tau = 1\star 1$ so $\mu\star\tau = \mu\star 1\star 1 = 1$

$$\begin{aligned}
&= 2x\sum_{d\le x}\frac{1}{d} + \mathcal{O}(x) \\
&= 2x\log x + \mathcal{O}(x).
\end{aligned}$$

□

*A 14-point plan to prove PNT from Selberg's identity

Let $r(x) = \frac{\psi(x)}{x} - 1$, so PNT is equivalent to $\lim_{x\to\infty} |r(x)| = 0$.

(1) Show that Selberg's identity gives

$$r(x)\log x = -\sum_{n\le x} \frac{\Lambda(n)}{n} r(\tfrac{x}{n}) + \mathcal{O}(1).$$

(2) Considering (1) with x replaced by $\frac{x}{m}$, summing over m, show

$$|r(x)|(\log x)^2 \le \sum_{n\le x} \frac{\Lambda_2(n)}{n} \left|r\left(\tfrac{x}{n}\right)\right| + \mathcal{O}(\log x).$$

(3) Show

$$\sum_{n\le x} \Lambda_2(n) = 2\int_1^{\lfloor x\rfloor} \log t\, dt + \mathcal{O}(x).$$

(4) Show

$$\sum_{n\le x} \frac{\Lambda_2(n)}{n} \left|r(\tfrac{x}{n})\right| = 2\sum_{2\le n\le x} \frac{r(\frac{x}{n})}{n} \int_{n-1}^{n} \log t\, dt + \mathcal{O}(x\log x).$$

(5) Show

$$\sum_{2\le n\le x} \frac{r(\frac{x}{n})}{n} \int_{n-1}^{n} \log t\, dt + \mathcal{O}(x\log x) = \int_1^x \frac{\left|r(\frac{x}{t})\right|}{t\log t}\, dt + \mathcal{O}(x\log x).$$

(6) Deduce

$$\sum_{n\le x} \frac{\Lambda_2(n)}{n} \left|r(\tfrac{x}{n})\right| = 2\int_1^x \frac{\left|r(\frac{x}{t})\right|}{t\log t}\, dt + \mathcal{O}(x\log x).$$

(7) Let $V(u) = r(e^u)$. Show that

$$u^2|V(u)| \le 2\int_0^u \int_0^v |V(t)|\, dt\, dv + \mathcal{O}(u)$$

(8) Show that

$$\alpha := \limsup |V(u)| \le \limsup \frac{1}{u}\int_0^u |V(t)|\, dt =: \beta$$

(9)-(14) If $\alpha > 0$, then can show from (7) that $\beta < \alpha$, contradiction, so $\alpha = 0$ and PNT.

2 Sieve Methods

Lecture 6 In the Sieve of Eratosthenes, we write out the numbers up to a given bound, then remove multiples of small primes. For example, crossing out multiples of 2 first, then multiples of 3, we are left with:

$$\begin{array}{cccccccccc} 1 & \cancel{2} & \cancel{3} & \cancel{4} & 5 & \cancel{6} & 7 & \cancel{8} & \cancel{9} & \cancel{10} \\ 11 & \cancel{12} & 13 & \cancel{14} & \cancel{15} & \cancel{16} & 17 & \cancel{18} & 19 & \cancel{20} \end{array}$$

We are left with all the primes above 3, and 1. Alternatively, we can use the inclusion-exclusion principle to count how much is left. Our interest is in using the sieve to count things: how many numbers are left?

$$\pi(20) + 1 - \pi(\sqrt{20}) = 20 - \left\lfloor \frac{20}{2} \right\rfloor - \left\lfloor \frac{20}{3} \right\rfloor + \left\lfloor \frac{20}{6} \right\rfloor.$$

This is the general idea: We get an expression relating some quantity we are interested in - the number of primes below a certain limit - in terms of how much we 'sieved' out at each stage.

2.1 Setup

We generally use:

- a finite set $A \subset \mathbb{N}$ (the set to be sifted)
- a set of primes P (the set of primes we sift out by, usually all primes).
- a sifting limit z (sift with all primes in $P < z$)
- a sifting function
$$S(A, P; z) = \sum_{n \in A} 1_{(n, P(z))=1}$$
where
$$P(z) := \prod_{\substack{p \in P \\ p < z}} p.$$
The goal is to estimate $S(A, P; z)$.
- For d, let
$$A_d = \{n \in A : d \mid n\}.$$
We write
$$|A_d| = \frac{f(d)}{d} X + R_d$$
where f is completely multiplicative ($f(mn) = f(m)f(n)\ \forall m, n$) and $0 \leq f(d)\ \forall d$. Note many textbooks write ω for f.
- Note that
$$|A| = \frac{f(1)}{1} X + R_1 = X + R_1$$
so we think of R_d as an error term
- We choose f so that $f(p) = 0$ if $p \notin P$ (so $R_p = |A_p|$)

- Let
$$W_P(z) := \prod_{\substack{p<z \\ p\in P}} \left(1 - \frac{f(p)}{p}\right).$$

Example.

(1) Take $A = (x, x+y] \cap \mathbb{N}$, and P the set of all primes, so
$$\begin{aligned} |A_d| &= \left\lfloor \frac{x+y}{d} \right\rfloor - \left\lfloor \frac{x}{d} \right\rfloor = \frac{x+y}{d} - \frac{x}{d} + \mathcal{O}(1) \\ &= \frac{y}{d} + \mathcal{O}(1) \end{aligned}$$
so $f(d) \equiv 1$ and $R_d = \mathcal{O}(1)$. So
$$S(A, P; z) = |\{x < n \le x + y : \text{if } p \mid n \text{ then } p \ge z\}|$$
e.g. if $z \approx (x+y)^{\frac{1}{2}}$ then
$$S(A, P; z) = \pi(x+y) - \pi(x) + \mathcal{O}((x+y)^{\frac{1}{2}})$$

(2) Take
$$A = \{1 \le n \le q : n \equiv a \pmod q\}.$$
Then
$$A_d = \left\{1 \le m \le \frac{x}{d} : dm \equiv a \pmod q\right\}.$$
This congruence only has solutions if $(d, q) \mid a$, so
$$|A_d| = \begin{cases} \frac{(d,q)}{dq} y + \mathcal{O}((d,q)) & \text{if } (d,q) \mid a \\ \mathcal{O}((d,q)) & \text{otherwise} \end{cases}$$
$$f(d) = \begin{cases} (d,q) & \text{if } (d,q) \mid a \\ 0 & \text{otherwise.} \end{cases}$$
We will do this example in more detail later, but it shows how f can be more complicated, and that we can use sieve methods to count primes congruent to a $\pmod q$.

(3) What about twin primes? Take $A = \{n(n+2) : 1 \le n \le x\}$, and P as all primes except 2. So $p \mid n(n+2) \iff n \equiv 0 \text{ or } -2 \pmod p$. Now,
$$|A_p| = 2\frac{x}{p} + \mathcal{O}(1).$$
So $f(p) = 2$, so $f(d) = 2^{\omega(d)}$. Then
$$\begin{aligned} S(A, P; x^{\frac{1}{2}}) &= |\{1 \le p \le x : p, p+2 \text{ both prime}\}| + \mathcal{O}(x^{\frac{1}{2}}) \\ &= \pi_2(x) + \mathcal{O}(x^{\frac{1}{2}}) \end{aligned}$$
We expect $\pi_2(x) \approx \frac{x}{(\log x)^2}$. We cannot prove the lower bound, but we can prove the upper bound using this sieve soon.

Theorem 2.1 (Sieve of Eratosthenes-Legendre)**.**
$$S(A, P; z) = XW_p(z) + \mathcal{O}\left(\sum_{d \mid p(z)} R_d\right).$$

Proof.

$$
\begin{aligned}
S(A,P;z) &= \sum_{n\in A} 1_{(n,P(z))=1} \\
&= \sum_{n\in A}\sum_{d\mid(n,P(z))} \mu(d) \\
&= \sum_{n\in A}\sum_{\substack{d\mid n\\ d\mid P(z)}} \mu(d) \\
&= \sum_{d\mid P(z)} \mu(d) \sum_{n\in A} 1_{d\mid n} \\
&= \sum_{d\mid P(z)} \mu(d)|A_d| \\
&= X\sum_{d\mid P(z)} \frac{\mu(d)f(d)}{d} + \sum_{d\mid P(z)} \mu(d)R_d \\
&= X\prod_{\substack{p\in P\\ p<z}} \left(1-\frac{f(p)}{p}\right) + \mathcal{O}\left(\sum_{d\mid P(z)} |R_d|\right). \qquad \square
\end{aligned}
$$

Corollary 2.2.

$$\pi(x+y)-\pi(x) \ll \frac{y}{\log\log y}.$$

Proof. In Example 1, recall $f\equiv 1$ and $|R_d| \ll 1$, $X = y$. So

$$W_P(z) = \prod_{p\le z}\left(1-\frac{1}{p}\right) \ll (\log z)^{-1}$$

and

$$\sum_{d\mid P(z)} |R_d| \ll \sum_{d\mid P(z)} 1 \le 2^z.$$

So $\pi(x+y)-\pi(x) \ll \frac{y}{\log z} + 2^z \ll \frac{y}{\log\log y}$ by choosing $z = \log y$. $\square$

2.2 Selberg's sieve

Lecture 7 From Sieve of Eratosthenes-Legendre, we got

$$S(A,P;z) \le XW + \mathcal{O}\left(\sum_{d\mid P(z)} |R_d|\right).$$

The problem here is that we have to consider 2^z many divisors of $P(z)$, so get 2^z many error terms. We can do a different sieve, and only consider those divisors of $P(z)$ which are small, say $\le D$.

The key part of Sieve of Eratosthenes-Legendre was

$$1_{(n,P(z))=1} = \sum_{d\mid(n,P(z))} \mu(d).$$

For an upper bound, we note that it is enough to use *any* function F in place of μ such that

$$F(n) \ge \begin{cases} 1 & n=1 \\ 0 & \text{otherwise} \end{cases}$$

(we used $F = \mu$ in the proof of Sieve of Eratosthenes-Legendre)

Selberg's observation was that if λ_i is an sequence of reals with $\lambda_1 = 1$ then

$$F(n) = \left(\sum_{d|n} \lambda_d\right)^2$$

works:

$$F(1) = \left(\sum_{d|1} \lambda_d\right)^2 = \lambda_1^2 = 1.$$

We make the additional assumption on f that $0 < f(p) < p$ if $p \in P$. Recall that $|A_p| = \frac{f(p)}{p}X + R_p$, so these are reasonable restrictions to have on a sieve.

This lets us define a new multiplicative function g such that

$$g(p) = \left(1 - \frac{f(p)}{p}\right)^{-1} - 1 = \frac{f(p)}{p - f(p)}$$

Theorem 2.3 (Selberg's sieve)**.**

$$\forall t \quad S(A, P; z) \le \frac{X}{G(t,z)} + \sum_{\substack{d|P(z)\\ d<t^2}} 3^{\omega(d)}|R_d|$$

where

$$G(t,z) = \sum_{\substack{d|P(z)\\ d<t}} g(d).$$

Recall

$$W = \prod_{\substack{p\in P\\ p\le z}} \left(1 - \frac{f(p)}{p}\right)$$

so the expected size of $S(A, P; z)$ is XW. Note that as $t \to \infty$,

$$\begin{aligned} G(t,z) &\to \sum_{d|P(z)} g(d) \\ &= \prod_{p<z} (1 + g(p)) \\ &= \prod_{p<z} \left(1 - \frac{f(p)}{p}\right)^{-1} = \frac{1}{W}. \end{aligned}$$

Corollary 2.4.

$$\pi(x+y) - \pi(x) \ll \frac{y}{\log y}.$$

Compare this with Corollary 2.2.

Proof. Take $A = \{x < n \le x + y\}$, $f(p) = 1$, $R_d = \mathcal{O}(1)$, $X = y$. Since $g(p) = \frac{1}{p-1} =$

$\frac{1}{\varphi(p)}$, so $g(d) = \frac{1}{\varphi(d)}$, The main term from Theorem 2.3 gives

$$\begin{aligned}
G(z,z) &= \sum_{\substack{d|P(z)\\ d<z}} \prod_{p|d} (p-1)^{-1} \\
&= \sum_{d=p_1\cdots p_r<z} \prod_i \sum_{k\geq 1}^{\infty} \frac{1}{p_i^k} \\
&= \sum_{p<z} \sum_{\substack{k_r\geq 1\\ p_1\cdots p_r<z}} \frac{1}{p_1^{k_1}\cdots p_r^{k_r}} \\
&= \sum_n \frac{1}{n} \text{ for } n \text{ where the square-free part of } n \text{ is } \leq t \\
&\geq \sum_{d<z} \frac{1}{d} \\
&\gg \log z.
\end{aligned}$$

So the main term is $\ll \frac{y}{\log z}$. Note that $3^{\omega(d)} \leq \tau_3(d) \ll_\epsilon d^\epsilon$. So the error term is

$$\ll_\epsilon t^\epsilon \sum_{d<t^2} 1 \ll t^{2+\epsilon} = z^{2+\epsilon}$$

since we are taking $t = z$. So

$$S(A,P;z) \ll \frac{y}{\log z} + z^{2+\epsilon} \ll \frac{y}{\log y}$$

by taking $z = y^{\frac{1}{3}}$. □

Proof of Theorem 2.3. Let (λ_i) be a sequence of reals, with $\lambda_1 = 1$, to be chosen later. Then

$$\begin{aligned}
S(A,P;z) &= \sum_{n\in A} 1_{(n,P(z))=1} \\
&\leq \sum_{n\in A} \left(\sum_{d|(n,P(z))} \lambda_d \right)^2 \\
&= \sum_{d,e|P(z)} \lambda_d\lambda_e \sum_{n\in A} 1_{d|n,\ e|n} \\
&= \sum_{d,e|P(z)} \lambda_d\lambda_e |A_{[d,e]}| \\
&= X \sum_{d,e|P(z)} \lambda_d\lambda_e \frac{f([d,e])}{[d,e]} + \sum_{d,e|P(z)} \lambda_d\lambda_e R_{[d,e]}.
\end{aligned}$$

$[d,e]$ denotes the least common multiple of d and e. We will choose λ_d such that $|\lambda_d| \leq 1$ and $\lambda_d = 0$ if $d \geq t$. Then

$$\begin{aligned}
\left| \sum_{d,e|P(z)} \lambda_d\lambda_e R_{[d,e]} \right| &\leq \sum_{\substack{d,e<t\\ d,e|P(z)}} |R_{[d,e]}| \\
&\leq \sum_{\substack{n|P(z)\\ n<t^2}} |R_n| \sum_{d,e} 1_{[d,e]=n}
\end{aligned}$$

and

$$\sum_{d,e} 1_{[d,e]=n} = 3^{\omega(n)}$$

as n is squarefree.

Let

$$V = \sum_{d,e|P(z)} \lambda_d \lambda_e \frac{f([d,e])}{[d,e]}$$

Write $[d,e] = abc$ where $d = ab$, $e = bc$ and $(a,b) = (b,c) = (a,c) = 1$, which we can do since $\lambda_d = 0$ if d is not square-free.

Lecture 8

$$\begin{aligned}
V &= \sum_{c|P(z)} \frac{f(c)}{c} \sum_{\substack{ab|P(z)\\(a,b)=1}} \frac{f(a)f(b)}{ab} \lambda_{ac}\lambda_{bc} \\
&= \sum_{c|P(z)} \frac{f(c)}{c} \sum_{ab|P(z)} \frac{f(a)}{a}\frac{f(b)}{b} \sum_{d|a,\, d|b} \mu(d)\lambda_{ac}\lambda_{bc} \\
&= \sum_{c|P(z)} \frac{f(c)}{c} \sum_{d|P(z)} \mu(d) \left(\sum_{d|a|P(z)} \frac{f(a)}{a}\lambda_{ac} \right)^2
\end{aligned}$$

taking $ac = n$,

$$\begin{aligned}
&= \sum_{d|P(z)} \mu(d) \sum_{c|P(z)} \frac{c}{f(c)} \left(\sum_{cd|n|P(z)} \frac{f(n)}{n}\lambda_n \right)^2 \\
&= \sum_{d|P(z)} \mu(d) \sum_{c|P(z)} \frac{c}{f(c)} y_{cd}^2 \\
&= \sum_{k|P(z)} \left(\sum_{cd=k} \mu(d)\frac{c}{f(c)} \right) y_k^2
\end{aligned}$$

For primes p,

$$\sum_{cd=p} \mu(d)\frac{c}{f(c)} = -1 + \frac{p}{f(p)} = \frac{p - f(p)}{f(p)} = \frac{1}{g(p)}.$$

Therefore $\forall h \mid P(z)$

$$\sum_{cd=k} \mu(d)\frac{c}{f(c)} = \frac{1}{g(k)}.$$

Note that if $k \geq t$ then

$$y_k = \sum_{\substack{k|n|P(z)\\h\geq t}} \frac{f(n)}{n}\lambda_n = 0$$

So

$$V = \sum_{\substack{k|P(z)\\k<t}} \frac{y_k^2}{g(k)}$$

Want to choose V as small as possible.

What is the relationship between y_k and λ_d?

$$y_k = \sum_{k|n|P(z)} \frac{f(n)}{n}\lambda_n.$$

Fix d.

$$\sum_{d|k|P(z)} \mu(k) y_k = \sum_{h|P(z)} \mu(k) \sum_{n|P(z)} \frac{f(n)}{n} \lambda_n 1_{d|k} 1_{k|n}$$
$$= \sum_{n|P(z)} \frac{f(n)}{n} \lambda_k 1_{d|n} \sum_{d \ midk|n} \mu(k)$$

Considering this innermost sum, write $k = de$, so we have

$$\mu(d) \sum_{e|\frac{n}{d}} \mu(e) = \begin{cases} \mu(d) & n = d \\ 0 & n > d \end{cases}$$

Thus

$$\sum_{d|k|P(z)} \mu(k) y_k = \mu(d) \frac{f(d)}{d} \lambda_d.$$

Recall $\lambda_1 = 1$, so must have

$$1 = \sum_{k|P(z)} \mu(k) y_k$$

$$1 = \left(\sum_{\substack{k|P(z) \\ k<t}} \mu(k) y_k g(k)^{\frac{1}{2}} \times \frac{1}{g(k)^{\frac{1}{2}}} \right)^2 \le \left(\sum_{\substack{k|P(z) \\ k<t}} g(k) \right) \left(\sum_{\substack{k|P(z) \\ k<t}} \frac{y_k^2}{g(k)} \right) = GV$$

So $V \ge \frac{1}{G}$; but equality holds iff $\exists c$ such that $\forall k$,

$$\frac{\mu(k) y_k}{g(k)^{\frac{1}{2}}} = c g(k)^{\frac{1}{2}}$$
$$\implies y_k = c \mu(k) g(k) \quad (k < t)$$

What is c? We know that

$$1 = c \sum_{k|P(z)} \mu(k)^2 g(k) = cG$$

so choose $c = \frac{1}{G}$. Check:

1. $\lambda_1 = 1$ ✓
2. $\lambda_d = 0$ if $d \ge t$ ✓
3. $|\lambda_d| \le 1$:

$$\lambda_d = \mu(d) \frac{d}{f(d)} \sum_{d|k|P(z)} \mu(k) y_k$$

so

$$|\lambda_d| = \frac{d}{f(d)} \frac{1}{G} \sum_{d|k|P(z)} g(k).$$

$$
\begin{aligned}
G &= \sum_{\substack{e \mid P(z) \\ e<t}} g(e) \\
&= \sum_{k \mid d} \sum_{\substack{e \mid P(z) \\ e<t \\ (d,e)=k}} g(e) \\
&= \sum_{k \mid d} \sum_{\substack{n \mid P(z) \\ (m,d)=1 \\ m<\frac{t}{k}}} g(m) \\
&\ge \left(\sum_{k \mid d} g(k)\right) \left(\sum_{\substack{m \mid P(z) \\ (m,d)=1 \\ m<\frac{t}{d}}} g(m)\right)
\end{aligned}
$$

Note that for primes p,

$$\sum_{k \mid p} g(k) = 1 + \frac{f(p)}{p - f(p)} = \frac{p}{p - f(p)} = \frac{p}{f(p)} g(p).$$

So

$$G \ge \frac{d}{f(d)} g(d) \left(\sum_{\substack{m \mid P(z) \\ (m,d)=1 \\ m<\frac{t}{d}}} g(m)\right) = \frac{d}{f(d)} \sum_{d \mid k \mid P(z)} g(k) = |\lambda_d| G$$

so $|\lambda_d| \le 1$. □

Theorem 2.5 (Brun)**.** Let $\pi_2(x) = \#\{1 \le n \le n : n \text{ and } n+2 \text{ are prime}\}$. Then

$$\pi_2(x) \ll \frac{x}{(\log x)^2}$$

We can reasonably expect $\pi_2(x) \asymp \frac{x}{(\log x)^2}$, but proving the lower bound would mean there are infinitely many twin primes.

Proof. Take $A = \{n(n+2) : 1 \le n \le x\}$, and $P =$ all primes except 2. Then

$$|A_d| = \#\{1 \le n \le x : d \mid n(n+2)\}$$

if $d = p_1 \cdots p_r$ odd and squarefree.

$$d \mid n(n+2) \iff p_i \mid n(n+2)\ \forall i \iff n \equiv 0 \text{ or } -2 \pmod{p_i}\ \forall i$$

By CRT, true iff n lies in one of $2^{\omega(d)}$ many residue classes mod d. So

$$|A_d| = \frac{2^{\omega(d)}}{d} x + \mathcal{O}(2^{\omega(d)})$$

so $f(d) = 2^{\omega(d)}$ for d odd, square-free, and $R_d \ll 2^{\omega(d)}$.

By Selberg's sieve, with $t = z = x^{\frac{1}{4}}$,

$$\begin{aligned}\pi_2(x) &\le \#\{1 \le n \le x : p \mid n(n+2) \Rightarrow p = 2 \text{ or } p > x^{\frac{1}{4}}\} + \mathcal{O}(x^{\frac{1}{4}}) \\ &= S(A, P; x^{\frac{1}{4}}) + \mathcal{O}(x^{\frac{1}{4}}) \\ &\le \frac{x}{G(z,z)} + \mathcal{O}(\sum_{\substack{d \mid P(z) \\ d < z^2}} 6^{\omega(d)})\end{aligned}$$

Focus on the error term first:

$$\sum_{d < z^2} 6^{\omega(d)} \le z^{2+o(1)} = x^{\frac{1}{2}+o(1)}.$$

Lecture 9 It remains to show

$$G(z,z) \gg (\log z)^2.$$

Note

$$g(p) = \frac{f(p)}{p - f(p))} = \frac{2}{p-2} \ge \frac{2}{p-1}$$

so if d is odd and squarefree,

$$g(d) \ge \frac{2^{\omega(d)}}{\varphi(d)}.$$

Thus,

$$G(z,z) \ge \sum_{\substack{d \mid P(z) \\ d < z}} \frac{2^{\omega(d)}}{\varphi(d)} \gg \sum_{\substack{d < z \\ d \text{ squarefree}}} \frac{2^{\omega(d)}}{\varphi(d)}$$

since we added in

$$\sum_{\substack{d < z \\ d \text{ squarefree} \\ 2 \mid d}} \frac{2^{\omega(d)}}{\varphi(d)} = 2 \sum_{\substack{e < \frac{z}{2} \\ e \text{ squarefree} \\ e \text{ odd}}} \frac{2^{\omega(e)}}{\varphi(e)} \le 2\epsilon_1$$

Now,

$$\begin{aligned}\sum_{\substack{d < z \\ d \text{ squarefree}}} \frac{2^{\omega(d)}}{\varphi(d)} &= \sum_{\substack{d < z \\ d \text{ squarefree} \\ d = p_1 \cdots p_r}} 2^{\omega(d)} \prod_{i=1}^{r} \left(\frac{1}{p_i} + \frac{1}{p_i^2} + \dots\right) = \sum_{\substack{e < z \\ d = em^2 \\ e \text{ squarefree}}} \frac{2^{\omega(d)}}{d} \\ &\ge \sum_{d < z} \frac{2^{\omega(d)}}{d}.\end{aligned}$$

By Partial summation, it's enough to show $\sum_{d<z} 2^{\omega(d)} \gg z \log z$. Recall that to show $\sum_{d<z} \tau(d) \gg z \log z$ we used $\tau = 1 \star 1$. We want to write $2^{\omega(n)} = \sum_{d \mid n} f(d) g(\frac{n}{d})$.

If we try $f = \tau$, it turns out that

$$g(n) = \begin{cases} 0 & \text{if } n \text{ not a square} \\ \mu(d) & \text{if } n = d^2 \end{cases}$$

works, and $2^{\omega(n)} = \tau \star g(n)$. So

$$\begin{aligned}
\sum_{d<z} 2^{\omega(d)} &= \sum_{a<z} g(a) \sum_{b \le \frac{z}{a}} \tau(b) \\
&= \sum_{a<z} g(a)\frac{z}{a}\log(\frac{t}{a}) + c\sum_{a<z} g(a)\frac{z}{a} = \mathcal{O}\left(\underbrace{z^{\frac{1}{2}}\sum_{a<z}\frac{1}{a^{\frac{1}{2}}}}_{\ll z}\right) \\
&= \sum_{d<z^{\frac{1}{2}}} \mu(d)\frac{z}{d^2}\log z - 2\underbrace{\sum_{d<z^{\frac{1}{2}}} \mu(d)\frac{z}{d^2}\log d}_{\ll z\sum_{d<z^{\frac{1}{2}}}\frac{\log d}{d^2} \ll z.}
\end{aligned}$$

Note

$$\sum_{d<z^{\frac{1}{2}}} \frac{\mu(d)}{d^2} = c + \mathcal{O}\left(\sum_{d<z^{\frac{1}{2}}}\frac{1}{d^2}\right) = c + \mathcal{O}\left(\frac{1}{z^{\frac{1}{2}}}\right)$$

so

$$\sum_{d<z} 2^{\omega(d)} = cz\log z + \mathcal{O}(z) \gg z\log z.$$

Remains to show that $c > 0$: either

- Note LHS can't be $\mathcal{O}(z)$
- Calculate the first couple of terms in the series
- Note that $c = \frac{6}{\pi^2} > 0$. □

2.3 Combinatorial sieve

The sieve of Eratosthenes-Legendre gave a sieve with a large error bound, and Selberg just gave an upper bound sieve.

$$S(A, P; z) = |A| - \sum_p |A_p| + \sum_{p,q} |A_{p,q}| + \cdots$$

The idea of a combinatorial sieve is to 'truncate' the sieve process.

Lemma (Buchstab Formula)**.**

$$S(A, P; z) = |A| - \sum_{p|P(z)} S(A_p, P; p).$$

Proof. Aim to show

$$|A| = S(A, P; z) + \sum_{p|P(z)} S(A_p, P; p)$$

Write

$$\begin{aligned}
S_1 &= \{n \in A : p \mid n, p \in P \Rightarrow p \ge z\} \\
S_p &= \{n \in A : n = mp, q \mid n, q \in P \Rightarrow q \ge p\}
\end{aligned}$$

and note $S(A, P; z) = \#S_1$ and $S(A_p, P; p) = \#S_p$. Every $n \in A$ is either in S_1, or has some prime divisors from $P(z)$. If p is the least such prime divisor, then $n \in S_p$. □

Similarly,

Lemma.

$$W(z) = 1 - \sum_{p \mid P(z)} \frac{f(p)}{p} W(p).$$

Recall that we defined

$$W(z) = \prod_{p \mid P(z)} \left(1 - \frac{f(p)}{p}\right)$$

Corollary. For any $r \geq 1$,

$$S(A, P; z) = \sum_{\substack{d \mid P(z) \\ \omega(d) < r}} \mu(d) |A_d| + (-1)^r \sum_{\substack{d \mid P(z) \\ \omega(d) = r}} S(A_d, P; l(d))$$

where $l(d)$ is the least prime divisor of d.

Proof. Induction on r. $r = 1$ is the Buchstab formula. For the inductive step, use

$$S(A_d, P; l(d)) = |A_d| - \sum_{\substack{p \in P \\ p < l(d)}} S(A_{dp}, P; p).$$

and

$$\begin{aligned} &(-1)^r \sum_{\substack{d \mid P(z) \\ \omega(d) = r}} \left(|A_d| - \sum_{\substack{p \in P \\ p < \lambda(d)}} S(A_{pd}, P; p) \right) \\ &= \sum_{\substack{d \mid P(z) \\ \omega(d) = r}} \mu(d) |A_d| + (-1)^{r+1} \sum_{\substack{e \mid P(z) \\ \omega(e) = r+1}} S(A_e, P; l(e)). \end{aligned}$$

□

In particular, note that if r is even

$$S(A, P; z) \geq \sum_{\substack{d \mid P(z) \\ \omega(d) < r}} \mu(d) |A_d|$$

(and the inequality is reversed if r odd).

Lecture 10 **Theorem** (Brun's Pure Sieve)**.** For any $r \geq 6 |\log W(z)|$,

$$S(A, P; z) = XW(z) + \mathcal{O}\left(2^{-r} X + \sum_{\substack{d \mid P(z) \\ d \leq z^r}} |R_d|\right)$$

Compare this to Eratosthenes sieve:

$$S(A, P; z) + XW(z) + \mathcal{O}\left(\sum_{d \mid P(z)} |R_d|\right)$$

Proof. Recall that from iterating Buchstab's formula, for any $r \geq 1$,

$$\begin{aligned} S(A,P;z) &= \sum_{\substack{d \mid P(z) \\ \omega(d)<r}} \mu(d)|A_d| + (-1)^r \sum_{\substack{d \mid P(z) \\ \omega(d)=r}} S(A_d,P;l(d)) \\ &= X \sum_{\substack{d \mid P(z) \\ \omega(d)<r}} \mu(d)\frac{f(d)}{d} + \sum_{\substack{d \mid P(z) \\ \omega(d)<r}} \mu(d)R_d + (-1)^r \sum_{\substack{d \mid P(z) \\ \omega(d)=r}} S(A_d,P;l(d)). \end{aligned}$$

Using the trivial bounds

$$0 \leq S(A_d,P;l(d)) \leq |A_d| = X\frac{f(d)}{d} + R_d,$$

this is

$$S(A,P;z) = X \sum_{\substack{d \mid P(z) \\ \omega(d)<r}} \mu(d)\frac{f(d)}{d} + \mathcal{O}\left(\sum_{\substack{d \mid P(z) \\ \omega(d)<r}} |R_d| + \sum_{\substack{d \mid P(z) \\ \omega(d)=r}} |A_d| \right)$$

By Buchstab again, applied to $W(z)$,

$$W(z) = \sum_{\substack{d \mid P(z) \\ \omega(d)<r}} \mu(d)\frac{f(d)}{d} + (-1)^r \sum_{\substack{d \mid P(z) \\ \omega(d)=r}} \mu(d)\frac{f(d)}{d} W(l(d))$$

So

$$S(A,P;z) = XW(z) + \mathcal{O}\left(\sum_{\substack{d \mid P(z) \\ \omega(d)<r}} |R_d| + \sum_{\substack{d \mid P(z) \\ \omega(d)=r}} |A_d| + X \sum_{\substack{d \mid P(z) \\ \omega(d)=r}} \frac{f(d)}{d} \right).$$

Error term:

$$\begin{aligned} &\ll X \sum_{\substack{d \mid P(z) \\ \omega(d)=r}} \frac{f(d)}{d} + \sum_{\substack{d \mid P(z) \\ \omega(d)\leq r}} |R_d| \\ &\leq \sum_{\substack{d \mid P(z) \\ d \leq z^r}} |R_d| \end{aligned}$$

because $d \mid P(z) = \prod_{\substack{p \in P \\ p<z}} p$.

Remains to show

$$\sum_{\substack{d \mid P(z) \\ \omega(d)=r}} \frac{f(d)}{d} \ll 2^{-r}.$$

Note that

$$\begin{aligned} \sum_{\substack{d \mid P(z) \\ \omega(d)=r}} \frac{f(d)}{d} = \sum_{\substack{p_1 \cdots p_r \\ p_i \in P \\ p_i < z}} \frac{f(p_1)\cdots f(p_r)}{p_1 \cdots p_r} &\leq \frac{\left(\sum_{p \mid P(z)} \frac{f(p)}{p} \right)^r}{r!} \\ &\leq \left(\frac{e \sum_{p \mid P(z)} \frac{f(p)}{p}}{r} \right)^r \end{aligned}$$

Now

$$\sum_{p|P(z)} \frac{f(p)}{p} \le \sum_{p|P(z)} -\log\left(1 - \frac{f(p)}{p}\right) = -\log W(z).$$

So if $r \ge 2e|\log W(z)|$ then

$$\sum_{\substack{d|P(z)\\ \omega(d)=r}} \frac{f(d)}{d} \le \left(\frac{e|\log W(z)|}{r}\right)^r \le 2^r.$$

□

Recall Selberg's sieve shows $\pi_2(x) \ll \frac{x}{(\log x)^2}$. In the twin prime sieve setting, recall that

$$W(z) \asymp \frac{1}{(\log z)^2}$$

So in Brun's sieve, need to take $r \gg 2\log\log z$. If $r = C\log\log z$ for C large enough, then $2^r X \ll \frac{X}{(\log z)^{100}}$. The main term is $\gg \frac{x}{(\log z)^2}$.

$$|R_d| \ll 2^{\omega(d)} = d^{o(1)}$$

$$\sum_{\substack{d|P(z)\\ d \le z^r}} |R_d| \ll z^{r+o(1)} = z^{2\log\log z + o(1)}$$

For this to be $o(\frac{x}{(\log z)^2})$, need to choose $z \approx \exp((\log x)^{\frac{1}{2}})$. We need to relate

$$S(A, P; z) \leftrightarrow \pi_2(x)$$

but $S(A, P; z) = \#\{1 \le n \le x : p \mid n(n+2) \text{ then } p \gg z = \exp((\log x)^{\frac{1}{4}})\}$ which includes many non-twin-primes.

Corollary. For any $z \le \exp(o((\frac{\log x}{\log\log x})))$,

$$\#\{1 \le n \le x : p \mid n \Rightarrow p \ge z\} \sim e^{-\gamma} \frac{x}{\log z}.$$

Remark.

(1) In particular, $z = (\log x)^A$ is allowed for any A but $z = x^c$ for any $c > 0$ is not allowed.

(2) In particular, can't count primes like this ($z = x^{\frac{1}{2}}$). Recall heuristic from before says if this asymptotic here correct for primes, then

$$\pi(x) \sim 2e^{-\gamma} \frac{x}{\log x}$$

which contradicts PNT.

Proof. Again, use $A = \{1 \leq n \leq x\}$ so $f(d) = 1$ and $|R_d| \ll 1$. Then

$$W(z) = \prod_{p<z} \left(1 - \frac{1}{p}\right) \sim \frac{e^{-\gamma}}{\log z} + o\left(\frac{1}{\log z}\right)$$

so

$$\begin{aligned} S(A, P; z) &= \#\{1 \leq n \leq x : p \mid n \Rightarrow p > z\} \\ &= e^{-\gamma} \frac{x}{\log z} + o\left(\frac{x}{\log z}\right) + O\left(2^{-r}x + \sum_{\substack{d \mid P(z) \\ d < z^r}} |R_d|\right) \end{aligned}$$

If $r \geq 6|\log W(z)|$, so $r \geq 100 \log\log z$ is fine.

$$2^{-r}x \leq (\log z)^{-(\log 2)100} x = o\left(\frac{x}{\log z}\right)$$

and (choose $r = \lceil 100 \log\log z \rceil$),

$$\sum_{\substack{d \mid P(z) \\ d \leq z^r}} |R_d| \ll \sum_{d \leq z^r} 1 \ll z^r \ll 2^{500(\log\log z)(\log z)}$$

Remains to note that if

$$\log z = o\left(\frac{\log x}{\log\log x}\right) = \frac{\log x}{\log\log x} F(x)$$

then this is

$$\log z \log\log z = o\left(\frac{\log x}{\log\log x} \cdot \log\log x\right) = o(\log x)$$

so $2^{500 \log\log z \log z} \leq x^{\frac{1}{10}}$ if x is large enough, which is $o(\frac{x}{\log z})$. $\square$

3 Riemann Zeta function

Lecture 11 First, a trivial remark (writing $s = \sigma + it$ throughout): If $n \in \mathbb{N}$, $n^s = e^{s \log n} = n^\sigma e^{it \log n}$.

Definition. The **Riemann zeta function** is defined for $\sigma > 1$ by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

3.1 Dirichlet series

For any arithmetic function $f : \mathbb{N} \to \mathbb{C}$, we have a **Dirichlet series**

$$L_f(s) := \sum_{n=1}^{\infty} \frac{f(n)}{n^s}.$$

Lemma 3.1. For any f, there is an abscissa of convergence σ_c such that

(1) $\sigma < \sigma_c \Rightarrow L_f(s)$ diverges

(2) $\sigma > \sigma_c \Rightarrow L_f(s)$ converges uniformly in some neighbourhood of s (in particular $L_f(s)$ is holomorphic at s).

Proof. It is enough to show if $L_f(s)$ converges at s_0 and $\sigma > \sigma_0$ then there is a neighbourhood of s on which L_f converges uniformly ($\sigma_c = \inf\{\sigma : L_f(s) \text{ converges}\}$). Let $R(u) = \sum_{n>u} f(n) n^{-s_0}$. By Partial summation,

$$\sum_{M<n\le N} f(n)n^{-s} = R(M)M^{s_0-s} - R(N)N^{s_0-s} + (s_0 - s)\int_M^N R(u)u^{s_0-s-1}\,du.$$

If $|R(u)| \le \epsilon$ for all $u \ge M$ then

$$\left| \sum_{M<n\le N} f(n)n^{-s} \right| \le 2\epsilon + \epsilon|s_0 - s| \int_M^N u^{\sigma_0-\sigma-1}\,du \le \left(2 + \frac{|s_0 - s|}{|\sigma_0 - \sigma|}\right)\epsilon$$

Note there is a neighbourhood of s in which $\frac{|s-s_0|}{|\sigma-\sigma_0|} \ll_s 1$. So $\sum \frac{f(n)}{n^s}$ converges uniformly here. □

Lemma 3.2. If

$$\sum \frac{f(n)}{n^s} = \sum \frac{g(n)}{n^s}$$

for all s in some halfplane $\sigma > \sigma_0 \in \mathbb{R}$ then $f(n) = g(n)\ \forall n$.

Proof. Enough to consider $\sum \frac{f(n)}{n^s} \equiv 0$ for $\forall \sigma > \sigma_0$. Suppose $\exists n\ f(n) = 0$. Let N be the least such that $f(N) \neq 0$. Since $\sum_{n\le N} \frac{f(n)}{n^\sigma} = 0$,

$$f(N) = -N^\sigma \sum_{n\ge N} \frac{f(n)}{n^\sigma}$$

So $|f(n)| \ll n^\sigma$ and so the series $\sum_{n\ge N} \frac{f(n)}{n^{\sigma+1+\epsilon}}$ is absolutely convergent. So since $\frac{f(n)}{n^\sigma} \to 0$ as $\sigma \to \infty$, the RHS $\to 0$ so $f(N) = 0$. □

Lemma 3.3. If $L_f(s)$ and $L_g(s)$ are absolutely convergent at s, then

$$L_{f\star g}(s) = \sum_{n=1}^{\infty} \frac{f \star g(n)}{n^s}$$

is also absolutely convergent at s and is equal to $L_f(s)L_g(s)$.

Proof.

$$\left(\sum_{n=1}^{\infty} \frac{f(n)}{n^s}\right)\left(\sum_{m=1}^{\infty} \frac{g(m)}{m^s}\right) = \sum_{n,m=1}^{\infty} \frac{f(n)g(m)}{(nm)^s} = \sum_{k=1}^{\infty} \frac{1}{k^s}\left(\sum_{\substack{n,m \\ nm=k}} f(n)g(m)\right).$$

□

Lemma 3.4 (Euler product)**.** If f is multiplicative then (if $L_f(s)$ is absolutely convergent at s)

$$L_f(s) = \prod_p \left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \cdots\right).$$

Proof. Let y be arbitrary:

$$\prod_{p<y}\left(1 + \frac{f(p)}{p^s} + \cdots\right) = \sum_{\substack{n \\ p|n \Rightarrow p<y}} \frac{f(n)}{n^s}$$

$$\left|\prod_{p<y}\left(1 + \frac{f(p)}{p^s} + \cdots\right) - \sum_{n=1}^{\infty} \frac{f(n)}{n^s}\right| \le \sum_{\substack{n \\ \exists p|n, p\ge y}} \frac{|f(n)|}{n^\sigma} \le \sum_{\substack{n \\ n\ge y}} \frac{|f(n)|}{n^\sigma} \to 0$$

as $y \to \infty$. □

Definition. For $\sigma > 1$,

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

defines a holomorphic function and converges absolutely for $\sigma > 1$.

Note that

$$\zeta'(s) = \sum \left(\frac{1}{n^s}\right)' = -\sum \frac{\log n}{n^s}.$$

Since 1 is completely multiplicative,

$$1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots = \left(1 - \frac{1}{p^s}\right)^{-1}$$

so $\zeta(s) = \prod_p (1 - \frac{1}{p^s})^{-1}$. So

$$\frac{1}{\zeta(s)} = \prod_p \left(1 - \frac{1}{p^s}\right) = \sum_n \frac{\mu(n)}{n^s}$$

$$\log \zeta(s) = -\sum_p \log\left(1 - \frac{1}{p^s}\right) = \sum_p \sum_k \frac{1}{kp^{ks}}$$

$$= \sum \frac{\Lambda(n)}{\log n} \frac{1}{n^s}$$

$$\frac{\zeta'(s)}{\zeta(s)} = -\sum \frac{\Lambda(n)}{n^s}$$

so e.g. $\frac{\zeta'(s)}{\zeta(s)} \times \zeta(s) = \zeta'(s)$, thus $\Lambda \star 1 = \log$. Similarly if $f \star 1 = g$, then $L_f \times \zeta = L_g$ so $L_f = \frac{1}{\zeta} \times L_g$ so $f = \mu \star g$.

Lemma. For $\sigma > 1$,

$$\zeta(s) = 1 + \frac{1}{s-1} - s \int_1^\infty \frac{\{t\}}{t^{s+1}}\, dt.$$

Proof. By Partial summation,

$$\begin{aligned}
\sum_{1 \le n \le x} \frac{1}{n^s} &= \frac{\lfloor x \rfloor}{x^s} + s \int_1^x \frac{\lfloor t \rfloor}{t^{s+1}}\, dt \\
&= \frac{\lfloor x \rfloor}{x^s} + s \int_1^x \frac{1}{t^s}\, dt - s \int_1^x \frac{\{t\}}{t^{s+1}}\, dt \\
&= \frac{\lfloor x \rfloor}{x^s} + \frac{s}{s-1} [t^{-s+1}]_1^x - s \int_1^x \frac{\{t\}}{t^{s+1}}\, dt
\end{aligned}$$

Now taking the limit as $x \to \infty$:

$$= \frac{s}{s-1} - s \int_1^x \frac{\{t\}}{t^{s+1}}\, dt.$$

□

The integral converges absolutely for $\sigma > 0$, so this gives

$$\zeta(s) = \frac{1}{s-1} + F(s)$$

where $F(s)$ is holomorphic in $\sigma > 0$. We define

$$\zeta(s) = 1 + \frac{1}{s-1} - s \int_1^\infty \frac{\{t\}}{t^{s+1}}\, dt \text{ for } \sigma > 0.$$

$\zeta(s)$ is meromorphic in $\sigma > 0$, with only a simple pole at $s = 1$.

Corollary. For $0 < \sigma < 1$,

$$\frac{1}{\sigma - 1} < \zeta(\sigma) < \frac{\sigma}{\sigma - 1}.$$

In particular, $\zeta(\sigma) < 0$ for $0 < \sigma < 1$ (in particular, $\neq 0$).

Proof.

$$\zeta(\sigma) = 1 + \frac{1}{\sigma - 1} - \sigma \int_1^\infty \frac{\{t\}}{t^{\sigma+1}}\, dt . 0 < \int_1^\infty \frac{\{t\}}{t^{\sigma+1}}\, dt < \frac{1}{\sigma}.$$

□

Corollary. For $0 < \delta \le \sigma \le 2$ and $|t| \le 1$,

$$\zeta(s) = \frac{1}{s-1} + \mathcal{O}_\delta(1) \text{ uniformly.}$$

Proof.

$$\begin{aligned}
\zeta(s) - \frac{1}{s-1} &= 1 - s \int_1^\infty \frac{\{t\}}{t^{s+1}}\, dt \\
&= \mathcal{O}(1) + \mathcal{O}\left(\int_1^\infty \frac{1}{t^{\sigma+1}}\, dt \right) \\
&= \mathcal{O}(1) + \mathcal{O}_\delta(1).
\end{aligned}$$

□

Lemma. $\zeta(s) \neq 0$ for $\sigma > 1$.

Proof. For $\sigma > 1$,

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}$$

and the infinite product converges, and no factors are zero. □

Conjecture (Riemann Hypothesis)**.** If $\zeta(s) = 0$ and $\sigma > 0$, then $\sigma = \frac{1}{2}$.

3.2 Prime Number Theorem

Let $\alpha(s) = \sum \frac{a_n}{n^s}$. Partial summation lets us write $\alpha(s)$ in terms of $A(x) = \sum_{n \le x} a_n$. If $\sigma > \max(0, \sigma_c)$ then $\alpha(s) = s \int_1^\infty \frac{A(t)}{t^{s+1}}\, dt$. This is often called the Mellin transform.

What about the converse? Note if $\alpha(s) = \frac{\zeta'(s)}{\zeta(s)}$ then $a_n = \Lambda(n)$ so

$$\begin{aligned} A(x) &= \sum_{n \le x} \Lambda(n) \\ &= \psi(x). \end{aligned}$$

The converse is called Perron's formula:

$$A(x) = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} \alpha(s) \frac{x^s}{s}\, ds \quad \sigma > \max(0, \sigma_c).$$

In particular, we get

$$\psi(x) = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} -\frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s}\, ds \quad \sigma > 1.$$

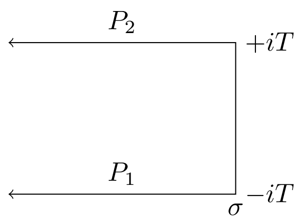
'Prime Number Theorem is equivalent to no zeros on $\sigma = 1$'

Lemma (Pre-Perron's formula)**.** If $\sigma > 0$, then (for $y \neq 1$)

$$\frac{1}{2\pi i} \int_{\sigma - iT}^{\sigma + iT} \frac{y^s}{s}\, ds = \begin{cases} 1 & y > 1 \\ 0 & y < 1 \end{cases} + \mathcal{O}\left(\frac{y^\sigma}{T|\log y|}\right).$$

Lecture 12

Proof. For $y > 1$, we use the contour C:



Since $\frac{y^s}{s}$ has a single pole at $s = 0$ with residue 1, by the residue theorem,

$$\frac{1}{2\pi i} \int_C \frac{y^s}{s}\, ds = 1.$$

Now we bound

$$\int_{P_1} \frac{y^s}{s}\,ds = \int_{-\infty}^{\sigma} \frac{y^{u+iT}}{u+iT}\,du \ll \frac{1}{T}\int_{-\infty}^{\sigma} y^u\,du = \frac{y^\sigma}{T\log y}.$$

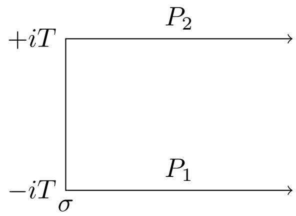
Similarly,

$$\int_{P_2} \frac{y^s}{s}\,ds \ll \frac{y^\sigma}{T\log y},$$

so

$$\int_C \frac{y^s}{s}\,ds = \int_{\sigma-iT}^{\sigma+iT} \frac{y^s}{s}\,ds + \mathcal{O}\left(\frac{y^\sigma}{T\log y}\right).$$

For $y < 1$, use the same argument with



□

Theorem (Perron's formula)**.** Suppose $\alpha(s) = \sum \frac{a_n}{n^s}$ is absolutely convergent for $\sigma > \sigma_a$. If $\sigma_0 > \max(0, \sigma_a)$ and x is not an integer, then

$$\sum_{n<x} a_n = \frac{1}{2\pi i}\int_{\sigma_0-iT}^{\sigma_0+iT} \alpha(s)\frac{x^s}{s}\,ds + \mathcal{O}\left(\frac{2^{\sigma_0}x}{T}\sum_{\frac{x}{2}<n<2x}\frac{a_n}{x-n} + \frac{x^{\sigma_0}}{T}\sum_{n=1}^{\infty}\frac{|a_n|}{n^{\sigma_0}}\right).$$

Proof. Since $\sigma_0 > 0$, we can write

$$\begin{aligned}
1_{n<x} &= \frac{1}{2\pi i}\int_{\sigma_0-iT}^{\sigma_0+iT}\frac{(x/n)^s}{s}\,ds + \mathcal{O}\left(\frac{(x/n)^{\sigma_0}}{T\left|\log(\frac{x}{n})\right|}\right)\\
\sum_{n<x} a_n &= \frac{1}{2\pi i}\sum_n a_n\int_{\sigma_0-iT}^{\sigma_0+iT}\frac{x^s}{n^s s}\,ds + \mathcal{O}\left(\frac{x^{\sigma_0}}{T}\sum_n\frac{|a_n|}{n^{\sigma_0}\left|\log(\frac{x}{n})\right|}\right)\\
&= \frac{1}{2\pi i}\int_{\sigma_0-iT}^{\sigma_0+iT}\frac{x^s}{s}\sum_n\frac{a_n}{n^s}\,ds + \mathcal{O}\left(\frac{x^{\sigma_0}}{T}\sum_n\frac{|a_n|}{n^{\sigma_0}\left|\log(\frac{x}{n})\right|}\right)\\
&= \frac{1}{2\pi i}\int_{\sigma_0-iT}^{\sigma_0+iT}\alpha(s)\frac{x^s}{s}\,ds + \mathcal{O}\left(\frac{x^{\sigma_0}}{T}\sum_n\frac{|a_n|}{n^{\sigma_0}\left|\log(\frac{x}{n})\right|}\right).
\end{aligned}$$

For the error term:

1. Contribution from $n \le \frac{x}{2}$ or $n \ge 2x$, where $|\log(\frac{x}{n})| \gg 1$, is
$$\ll \frac{x^{\sigma_0}}{T}\sum\frac{|a_n|}{n^{\sigma_0}}.$$

2. Contribution from $\frac{x}{2} < n < 2x$, we write $|\log(\frac{x}{n})| = |\log(1+\frac{n-x}{x})|$ and $|\log(1+\delta)| \asymp |\delta|$ uniformly for $-\frac{1}{2} \le \delta \le 1$. So
$$\frac{x^{\sigma_0}}{T}\sum_{\frac{x}{2}<n<2x}\frac{|a_n|}{n^{\sigma_0}|\log(\frac{x}{n})|} \ll \frac{x^{\sigma_0}}{T}\sum_{\frac{x}{2}<n<2x}\frac{|a_n|x}{n^{\sigma_0}|x-n|} \ll \frac{2^{\sigma_0}}{T}\sum_{\frac{x}{2}<n<2x}\frac{|a_n|x}{|x-n|}.$$

□

We will now prove a strong form of the PNT, under the assumptions

1. $\exists c > 0$, such that if $\sigma > 1 - \frac{c}{\log(|t|+4)}$ and $|t| \geq \frac{7}{8}$ then $\zeta(s) \neq 0$ and $\frac{\zeta' s}{\zeta s} \ll \log(|t|+4)$.
2. $\zeta(s) \neq 0$ for $\frac{8}{9} \leq \sigma \leq 1$, $|t| \leq \frac{7}{8}$.
3. $\frac{\zeta'(s)}{\zeta(s)} = -\frac{1}{s-1} + \mathcal{O}(1)$ for $1 - \frac{c}{\log(|t|+4)} < \sigma \leq 2$ for $|t| \leq \frac{7}{8}$.

We will come back and prove these soon.

Theorem (Prime Number Theorem)**.** There exists $c > 0$ such that

$$\psi(x) = x + \mathcal{O}\left(\frac{x}{\exp(c\sqrt{\log x})}\right)$$

In particular, $\psi(x) \sim x$.

Proof. Assume that $x = N + \frac{1}{2}$. By Perron's formula, for any $1 < \sigma_0 \leq 2$

$$\psi(x) = \sum_{n \leq x} \Lambda(n) = \frac{1}{2\pi i} \int_{\sigma_0 - iT}^{\sigma_0 + iT} -\frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s} \, ds + \mathcal{O}\left(\frac{x}{T} \sum_{\frac{x}{2} < n < 2x} \frac{\Lambda(n)}{|x-n|} + \frac{x^{\sigma_0}}{T} \sum \frac{\lambda(n)}{n^{\sigma_0}}\right)$$

In the error term,

$$R_1 \ll \log x \cdot \frac{x}{T} \cdot \sum_{\frac{x}{2} < n < 2x} \frac{1}{|x-n|} \ll \log x \cdot \frac{x}{T} \sum_{1 \leq m \leq 4x} \frac{1}{m} \ll \frac{x}{T} (\log x)^2$$

and

$$R_2 \ll \frac{x^{\sigma_0}}{T} \frac{1}{|\sigma_0 - 1|} \ll \frac{x}{T} \log x \quad \text{if } \sigma_0 = 1 + \frac{1}{\log x}.$$

where the bound on R_2 used assumption 3. Let C be the contour **missing picture** with $\sigma_1 < 1$. Then

$$\frac{1}{2\pi i} \int_C -\frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s} \, ds = x$$

by the residue theorem and assumptions 1 and 2.

Take $\sigma - 1 = 1 - \frac{c}{\log T}$.

$$\int_{\sigma_0 + iT}^{\sigma_1 + iT} -\frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s} \, ds \ll \log T \int_{\sigma_0}^{\sigma_1} \frac{x^u}{T} \, du \ll \frac{\log T}{T} x^{\sigma_1} (\sigma_1 - \sigma_0) \ll \frac{x}{T}$$

and

$$\begin{aligned}
\int_{\sigma_1 - iT}^{\sigma_1 + iT} -\frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s} \, ds &\ll (\log T) \left| \int_{\sigma_1 \pm iT}^{\sigma_1 \pm i} \frac{x^u}{u} \, du \right| + \left(\int_{\sigma_1 - i}^{\sigma_1 + i} x^{\sigma_1} \frac{1}{\sigma_1 - 1} \right) \\
&\ll x^{\sigma_1} \log T + \frac{x^{\sigma_1}}{1 - \sigma_1} \ll x^{\sigma_1} (\log T)
\end{aligned}$$

Now,

$$\begin{aligned}
\psi(x) &= x + \mathcal{O}\left(\frac{x}{T} (\log x)^2 + x^{1 - \frac{c}{\log T}} (\log T)\right) \\
&= x + \mathcal{O}\left(\frac{x}{\exp(c\sqrt{\log x})}\right)
\end{aligned}$$

by choosing $T = \exp(c\sqrt{\log x})$. □

Index of Notation

Index