Part III – Analytic Number Theory (Rough)

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

Lent 2019

Contents

0   Introduction

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 $a \mod q$ 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 $|f(x)| \le C|g(x)|$ 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: $1 \star f \ne f$.

- 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:
$$\tau(n) = |\{d \mid d \text{ a factor of } n\}|$$
$$\tau = 1 \star 1$$

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.

Fact (Möbius inversion)**.**
$$1 \star f = g \iff \mu \star g = f.$$

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

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 $\{d : d \text{ divides } n\} = \{1, p, \ldots, p^k\}$ so the left hand side is $1 - 1 + 0 + \ldots + 0 = 0$, 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. $\qquad \square$

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 $\pi(x) = \sum_{1 \leq n \leq x} 1_p(n)$. 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 $n = p_1^{k_1} \ldots p_k^{n_k}$,

$$= \sum_{i=1}^{r} \sum_{j=1}^{k_i} \Lambda(p_i^j)$$

$$= \sum_{i=1}^{r} \sum_{j=1}^{k_i} \log p_i$$

$$= \sum_{i=1}^{r} k_i \log p_i = \sum_{i=1}^{r} \log p_i^{k_i} = \log n. \qquad \square$$

Example. We can write

$$\Lambda(n) = \sum_{d|n} \mu(d) \log\left(\frac{n}{d}\right)$$

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

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

$$\sum_{1 \leq n \leq x} \Lambda(n) = -\sum_{1 \leq n \leq x} \sum_{d|n} \mu(d) \log d$$

$$= -\sum_{d \leq x} \mu(d) \log(d) \left(\sum_{\substack{1 \leq n \leq x \\ d|n}} 1 \right)$$

but $\sum_{\substack{1 \leq n \leq x \\ d|n}} 1 = \lfloor \frac{x}{d} \rfloor = \frac{x}{d} + \mathcal{O}(1)$, so

$$= -x \sum_{d \leq x} \mu(d) \frac{\log d}{d} + \mathcal{O}\left(\sum_{d \leq x} \mu(d) \log d \right).$$

1.2 Partial summation

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

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

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

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

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

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

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

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

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

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

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

$$\sum_{1 \leq n \leq N} a_n f(n) = \sum_{1 \leq n \leq N} f(n)\, (A(n) - A(n-1))$$

(note $A(0) = 0$)

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

Now

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

So

$$\sum_{1 \leq n \leq N} a_n f(n) = A(N)f(N) - \sum_{n=1}^{N-1} f'(t)\, dt$$

$$= A(N)f(N) - \int_1^N A(t)f'(t)\, dt$$

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

$$A(x)f(x) = A(N)f(x)$$

$$= A(N)\left(f(N) + \int_N^x f'(t)\, dt\right). \qquad \square$$

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

$$= 1 + \mathcal{O}\left(\tfrac{1}{x}\right) + \int_1^x \frac{1}{t}\, dt - \int_1^x \frac{\{t\}}{t^2}\, dt$$

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

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

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

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

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 $f(x) = \log x$, $a_n = 1$, $A(x) = \lfloor x \rfloor$.

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

$$= x \log x + \mathcal{O}(\log x) - \int_1^x 1\, dt + \mathcal{O}\left(\int_1^x \frac{1}{t}\, dt\right)$$

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

$$= x \log x - x + \mathcal{O}(\log x). \qquad\square$$

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.
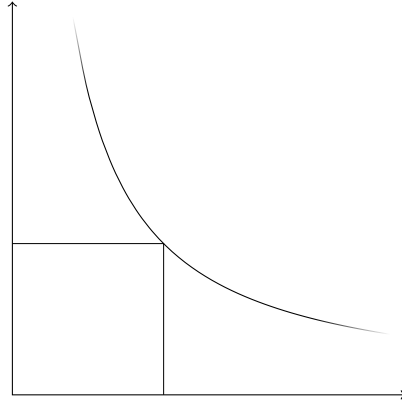
Instead, play around with the definition:

$$\sum_{1 \le n \le x} \tau(n) = \sum_{1 \le n \le x} \sum_{d | x} 1$$

$$= \sum_{1 \le d \le x} \sum_{\substack{1 \le n \le x \\ d | n}} 1$$

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

$$= \sum_{1 \le d \le x} \left\lfloor \frac{x}{d} \right\rfloor$$

$$= \sum_{1 \le d \le x} \frac{x}{d} + \mathcal{O}(x)$$

$$= x \sum_{1 \le d \le x} \frac{1}{d} + \mathcal{O}(x)$$

$$= x \log x + \gamma x + \mathcal{O}(x)$$

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

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

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

$$\sum_{1 \le n \le x} \tau(n) = \sum_{a \le x^{\frac{1}{2}}} \sum_{b \le \frac{x}{a}} 1 + \sum_{b \le x^{\frac{1}{2}}} \sum_{a \le \frac{x}{b}} 1 - \sum_{a, b \le x^{\frac{1}{2}}} 1$$

$$= 2 \sum_{a \le x^{\frac{1}{2}}} \left\lfloor \frac{x}{a} \right\rfloor - \underbrace{\left\lfloor x^{\frac{1}{2}} \right\rfloor^2}_{= \left(x^{\frac{1}{2}} + \mathcal{O}(1) \right)^2}$$

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

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

$$= x \log x + (2\gamma - 1)x + \mathcal{O}(x^{\frac{1}{2}}). \qquad \square$$

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.

Theorem 1.5. For any $n \geq 1$,

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

In particular,

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

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

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 $p \geq 2^{\frac{1}{\epsilon}}$, then $\frac{k+1}{p^{k\epsilon}} \leq \frac{k+1}{2^k} \leq 1$.

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

So

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

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,

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

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

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

1.4 Estimates for the primes

Recall

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

and

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

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 $\psi(x) \gg x$.

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

$x \log x$ is a trivial upper bound for this, (each summand is $\leq \log x$); 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 $\Sigma \leq 1$. We'll use the identity $\lfloor x \rfloor \leq 2\lfloor \frac{x}{2} \rfloor + 1$, valid for $x \geq 0$. (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

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

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

$$. = \sum_{n \leq x} \Lambda(n) \sum_{m \leq \frac{x}{n}} 1 - 2 \sum_{n \leq x} \Lambda(n) \sum_{m \leq \frac{x}{2n}} 1$$

$$= \sum_{mn \leq x} \Lambda(n) - 2 \sum_{nm \leq \frac{x}{2}} \Lambda(n)$$

$$= \sum_{d \leq x} 1 \star \Lambda(d) - 2 \sum_{d \leq \frac{x}{2}} 1 \star \Lambda(d)$$

$$= \sum_{d \leq x} \log d - 2 \sum_{d \leq \frac{x}{2}} \log d$$

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

$$= (\log 2)x + \mathcal{O}(\log x) \gg x.$$

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

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

Thus

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

$$\psi(x) = \left(\psi(x) - \psi\left(\tfrac{x}{2} \right) \right) + \left(\psi(\tfrac{x}{2}) - \psi(\tfrac{x}{4}) \right) + \cdots$$

$$\leq \log 2 \left(x + \tfrac{x}{2} + \tfrac{x}{4} + \cdots \right) + \mathcal{O}((\log x)^2)$$

$$= 2 \log 2x + \mathcal{O}((\log x)^2). \qquad \square$$

Lemma 1.7.
$$\sum_{p \leq x} \frac{\log p}{p} = \log x + \mathcal{O}(1).$$

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

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

$$= \sum_{a \leq x} \Lambda(a) \lfloor \tfrac{x}{a} \rfloor = x \sum_{a \leq x} \frac{\Lambda(a)}{a} + \mathcal{O}(\psi(x))$$

$$= x \sum_{a \leq x} \frac{\Lambda(a)}{a} + \mathcal{O}(x)$$

But from Lemma 1.3,

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

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

Remains to note

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

So

$$\sum_{n \leq x} \frac{\Lambda(n)}{n} = \sum_{p \leq x} \frac{\log p}{p} + \mathcal{O}(1). \qquad \square$$

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:

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

whereas

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

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

Thus,

$$\psi(x) = \pi(x) \log x + \mathcal{O}(x^{\frac{1}{2}} \log x) - \int_1^x \frac{\pi(t)}{t} \, dt$$

$$= \pi(x) \log x + \mathcal{O}(x^{\frac{1}{2}}) + \mathcal{O}\left(\int_1^x \frac{1}{\log t} \, dt\right)$$

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

---

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

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

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

Lemma 1.9.
$$\sum_{p \leq x} \frac{1}{p} = \log \log x + b + \mathcal{O}(\tfrac{1}{\log x})$$

where b is some constant.

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

$$\sum_{2 \leq p \leq x} \frac{1}{p} = \frac{A(x)}{\log x} + \int_2^x \frac{A(t)}{t(\log t)^2} \, dt$$

$$= 1 + \mathcal{O}(\tfrac{1}{\log x}) + \int_2^x \frac{1}{t \log t} \, dt + \int_2^x \frac{R(t)}{t(\log t)^2} \, dt$$

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

$$\sum_{2 \leq p \leq x} \frac{1}{p} = 1 + c + \mathcal{O}(\tfrac{1}{\log x}) + \log \log x - \log \log 2 + \mathcal{O}\left(\int_x^\infty \frac{1}{t(\log t)^2} \, dt \right)$$

$$= \log \log x + b + \mathcal{O}\left(\tfrac{1}{\log x} \right). \qquad\square$$

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 $\sum_{p \leq x} \frac{1}{p}$.

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

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

$$= \frac{c}{\log x} + o\left(\frac{1}{\log x} \right) + (c + o(1)) \int_1^x \frac{1}{t \log t} \, dt$$

$$= \mathcal{O}\left(\frac{1}{\log x} \right) + (c + o(1)) \log \log x.$$

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

Lemma 1.11.
$$\prod_{p \leq x} \left(1 - \frac{1}{p} \right)^{-1} = c \log x + \mathcal{O}(1)$$

where c is some constant.

Proof.

$$\log\left(\prod_{p\le x}\left(1-\frac{1}{p}\right)^{-1}\right) = -\sum_{p\le x}\log\left(1-\frac{1}{p}\right)$$

$$= \sum_{p\le x}\sum_{k}\frac{1}{kp^k}$$

$$= \sum_{p\le x}\frac{1}{p} + \sum_{k\ge 2}\sum_{p\le x}\frac{1}{kp^k}$$

$$= \log\log x + c' + \mathcal{O}\left(\frac{1}{\log x}\right).$$

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

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

$$= c\log x + \mathcal{O}(1). \qquad \square$$

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

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 $2e^{-\gamma} \approx 1.122\ldots$, 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

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

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

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

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

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

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

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

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

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

$$\sum_{d|n}(\Lambda(d)\log d + \Lambda \star \Lambda(d)) = \sum_{d|n}\Lambda(d)\log d + \sum_{ab|n}\Lambda(a)\Lambda(b)$$

$$= \sum_{d|n}\Lambda(d)\log d + \sum_{a|n}\Lambda(a)\left(\sum_{b|\frac{n}{a}}\Lambda(b)\right)$$

$$= \sum_{d|n}\Lambda(d)\log d + \sum_{d|n}\Lambda(d)\log\left(\frac{n}{d}\right)$$

$$= \log n \sum_{d|n}\Lambda(d) = (\log n)^2.$$

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. $\qquad\square$

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

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

Proof.

$$\sum_{n \leq x}\Lambda_2(n) = \sum_{n \leq x}\mu \star (\log)^2(n)$$

$$= \sum_{ab \leq x}\mu(a)(\log b)^2$$

$$= \sum_{a \leq x}\mu(a)\left(\sum_{b \leq \frac{x}{a}}(\log b)^2\right).$$

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 $A(t) = \sum_{n \leq t}\tau(n) = t\log t + Ct + \mathcal{O}(t^{\frac{1}{2}})$)

$$\sum_{m \leq x}\frac{\tau(m)}{m} = \frac{A(x)}{x} + \int_1^x \frac{A(t)}{t^2}\,dt$$

$$= \log x + C + \mathcal{O}(x^{-\frac{1}{2}}) + \int_1^x \frac{\log t}{t}\,dt + c\int_1^x \frac{1}{t}\,dt + \mathcal{O}\left(\int_1^x \frac{1}{t^{\frac{3}{2}}}\,dt\right)$$

$$= \frac{(\log x)^2}{2} + c_1\log x + c_2 + \mathcal{O}(x^{-\frac{1}{2}}).$$

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

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

Now, we show that the last three terms here are $\mathcal{O}(x)$ in reverse order: First, note that
$$x^{\frac{1}{2}} \sum_{a \leq x} \frac{1}{a^{\frac{1}{2}}} = \mathcal{O}(x).$$

Secondly,

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

Thirdly,

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

So

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

Recall $\tau = 1 \star 1$ so $\mu \star \tau = \mu \star 1 \star 1 = 1$
$$\begin{aligned}
&= 2x \sum_{d \leq x} \frac{1}{d} + \mathcal{O}(x) \\
&= 2x \log x + \mathcal{O}(x). \qquad \square
\end{aligned}$$

1.6 *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

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

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

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

(3) Show

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

(4) Show

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

(5) Show

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

(6) Deduce

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

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

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

(8) Show that

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

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

2   Sieve Methods

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:

$$\textcircled{1} \quad \cancel{2} \quad \cancel{3} \quad \cancel{4} \quad \textcircled{5} \quad \cancel{6} \quad \textcircled{7} \quad \cancel{8} \quad \cancel{9} \quad \cancel{10}$$

$$\textcircled{11} \quad \cancel{12} \quad \textcircled{13} \quad \cancel{14} \quad \cancel{15} \quad \cancel{16} \quad \textcircled{17} \quad \cancel{18} \quad \textcircled{19} \quad \cancel{20}$$

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
$$|A_d| = \left\lfloor \frac{x+y}{d} \right\rfloor - \left\lfloor \frac{x}{d} \right\rfloor = \frac{x+y}{d} - \frac{x}{d} + \mathcal{O}(1)$$
$$= \frac{y}{d} + \mathcal{O}(1)$$

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
$$A = \{1 \le n \le y : n \equiv a \pmod{q}\}.$$
Then
$$A_d = \left\{1 \le m \le \frac{y}{d} : dm \equiv a \pmod{q}\right\}.$$
This congruence only has solutions if $(d, q) \mid a$, so
$$|A_d| = \begin{cases} \frac{(d,q)}{dq} y + \mathcal{O}((d,q)) & \text{if } (d,q) \mid a \\ \mathcal{O}((d,q)) & \text{otherwise} \end{cases}$$
$$f(d) = \begin{cases} (d,q) & \text{if } (d,q) \mid a \\ 0 & \text{otherwise,} \end{cases}$$

taking $X = \frac{y}{q}$. 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 $p \mid n(n+2) \iff n \equiv 0$ 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
$$S(A, P; x^{\frac{1}{2}}) = |\{1 \le p \le x : p, p+2 \text{ both prime}\}| + \mathcal{O}(x^{\frac{1}{2}})$$
$$= \pi_2(x) + \mathcal{O}(x^{\frac{1}{2}})$$

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)**.**

$$S(A, P; z) = X W_P(z) + \mathcal{O}\left(\sum_{d \mid P(z)} R_d\right).$$

Proof.

$$S(A, P; z) = \sum_{n \in A} 1_{(n,P(z))=1}$$

$$= \sum_{n \in A} \sum_{d \mid (n,P(z))} \mu(d)$$

$$= \sum_{n \in A} \sum_{\substack{d \mid n \\ d \mid P(z)}} \mu(d)$$

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

$$= \sum_{d \mid P(z)} \mu(d) |A_d|$$

$$= X \sum_{d \mid P(z)} \frac{\mu(d) f(d)}{d} + \sum_{d \mid P(z)} \mu(d) R_d$$

$$= X \prod_{\substack{p \in P \\ p < z}} \left(1 - \frac{f(p)}{p}\right) + \mathcal{O}\left(\sum_{d \mid P(z)} |R_d|\right). \qquad \square$$

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 $z = \log y$. $\qquad \square$

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)**.**

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

where

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

Recall

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

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

$$G(t, z) \to \sum_{d|P(z)} g(d)$$

$$= \prod_{p < z} (1 + g(p))$$

$$= \prod_{p < z} \left(1 - \frac{f(p)}{p} \right)^{-1} = \frac{1}{W}.$$

Corollary 2.4.

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

Compare this with Corollary 2.2.

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

$$G(z,z) = \sum_{\substack{d|P(z) \\ d<z}} \prod_{p|d} (p-1)^{-1}$$

$$= \sum_{d=p_1\cdots p_r<z} \prod_i \sum_{k\geq 1}^{\infty} \frac{1}{p_i^k}$$

$$= \sum_{p<z} \sum_{\substack{k_r\geq 1 \\ p_1\cdots p_r<z}} \frac{1}{p_1^{k_1}\cdots p_r^{k_r}}$$

$$= \sum_n \frac{1}{n} \text{ for } n \text{ where the square-free part of } n \text{ is } \leq t$$

$$\geq \sum_{d<z} \frac{1}{d}$$

$$\gg \log z.$$

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 $z = y^{\frac{1}{3}}$. $\qquad\square$

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

$$S(A, P; z) = \sum_{n\in A} 1_{(n,P(z))=1}$$

$$\leq \sum_{n\in A} \left(\sum_{d|(n,P(z))} \lambda_d \right)^2$$

$$= \sum_{d,e|P(z)} \lambda_d\lambda_e \sum_{n\in A} 1_{d|n,\ e|n}$$

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

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

$[d,e]$ 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

$$\left| \sum_{d,e|P(z)} \lambda_d\lambda_e R_{[d,e]} \right| \leq \sum_{\substack{d,e<t \\ d,e|P(z)}} |R_{[d,e]}|$$

$$\leq \sum_{\substack{n|P(z) \\ n<t^2}} |R_n| \sum_{d,e} 1_{[d,e]=n}$$

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

as n is squarefree.

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

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

$$V = \sum_{c \mid P(z)} \frac{f(c)}{c} \sum_{\substack{ab \mid P(z) \\ (a,b)=1}} \frac{f(a)f(b)}{ab} \lambda_{ac} \lambda_{bc}$$

$$= \sum_{c \mid P(z)} \frac{f(c)}{c} \sum_{ab \mid P(z)} \frac{f(a)}{a} \frac{f(b)}{b} \sum_{d \mid a, \, d \mid b} \mu(d) \lambda_{ac} \lambda_{bc}$$

$$= \sum_{c \mid P(z)} \frac{f(c)}{c} \sum_{d \mid P(z)} \mu(d) \left(\sum_{d \mid a \mid P(z)} \frac{f(a)}{a} \lambda_{ac} \right)^2$$

taking $ac = n$,

$$= \sum_{d \mid P(z)} \mu(d) \sum_{c \mid P(z)} \frac{c}{f(c)} \left(\sum_{cd \mid n \mid P(z)} \frac{f(n)}{n} \lambda_n \right)^2$$

$$= \sum_{d \mid P(z)} \mu(d) \sum_{c \mid P(z)} \frac{c}{f(c)} y_{cd}^2$$

$$= \sum_{k \mid P(z)} \left(\sum_{cd=k} \mu(d) \frac{c}{f(c)} \right) y_k^2$$

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

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

Want to choose V as small as possible.

What is the relationship between y_k and λ_d?
$$y_k = \sum_{k \mid n \mid P(z)} \frac{f(n)}{n} \lambda_n.$$

Fix d.

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

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

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 $\forall k$,

$$\frac{\mu(k)y_k}{g(k)^{\frac{1}{2}}} = cg(k)^{\frac{1}{2}}$$

$$\implies y_k = c\mu(k)g(k) \quad (k < t)$$

What is c? We know that

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

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

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

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

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

$$= \sum_{k|d} \sum_{\substack{n|P(z) \\ (m,d)=1 \\ m<\frac{t}{k}}} g(m)$$

$$\geq \left(\sum_{k|d} g(k) \right) \left(\sum_{\substack{m|P(z) \\ (m,d)=1 \\ m<\frac{t}{d}}} g(m) \right)$$

Note that for primes p,

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

So

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

so $|\lambda_d| \leq 1$. $\qquad\square$

Theorem 2.5 (Brun). Let $\pi_2(x) = \#\{1 \leq n \leq n : n \text{ and } n+2 \text{ are prime}\}$. 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 $A = \{n(n+2) : 1 \leq n \leq x\}$, and $P = $ all primes except 2. Then

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

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 $t = z = x^{\frac{1}{4}}$,

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

Focus on the error term first:

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

It remains to show

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

Note

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

so if d is odd and squarefree,

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

Thus,

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

since we added in

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

Now,

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

$$
\geq \sum_{d < z} \frac{2^{\omega(d)}}{d}.
$$

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

$$\sum_{d<z} 2^{\omega(d)} = \sum_{a<z} g(a) \sum_{b \le \frac{z}{a}} \tau(b)$$

$$= \sum_{a<z} g(a) \frac{z}{a} \log(\frac{t}{a}) + c \sum_{a<z} g(a) \frac{z}{a} = \mathcal{O}\left(z^{\frac{1}{2}} \underbrace{\sum_{a<z} \frac{1}{a^{\frac{1}{2}}}}_{\ll z} \right)$$

$$= \sum_{d<z^{\frac{1}{2}}} \mu(d) \frac{z}{d^2} \log z - 2 \underbrace{\sum_{d<z^{\frac{1}{2}}} \mu(d) \frac{z}{d^2} \log d}_{\ll z \sum_{d<z^{\frac{1}{2}}} \frac{\log d}{d^2} \ll z.}$$

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

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

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.

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

Recall that we defined

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

Corollary. For any $r \geq 1$,

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

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

$$(-1)^r \sum_{\substack{d|P(z) \\ \omega(d)=r}} \left(|A_d| - \sum_{\substack{p \in P \\ p < \lambda(d)}} S(A_{pd}, P; p)\right)$$

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

$\square$

In particular, note that if r is even

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

(and the inequality is reversed if r odd).

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

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

Compare this to Eratosthenes sieve:

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

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

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

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

Using the trivial bounds

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

this is

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

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

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

So

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

Error term:

$$\ll X \sum_{\substack{d|P(z) \\ \omega(d)=r}} \frac{f(d)}{d} + \sum_{\substack{d|P(z) \\ \omega(d)\leq r}} |R_d|$$

$$\leq \sum_{\substack{d|P(z) \\ d\leq z^r}} |R_d|$$

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

Remains to show

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

Note that

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

$$\leq \left(\frac{e\sum_{p|P(z)} \frac{f(p)}{p}}{r}\right)^r$$

Now

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

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

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

$\square$

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 $\gg \frac{x}{(\log z)^2}$.

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

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

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

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

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

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

$$S(A, P; z) = \#\{1 \leq n \leq x : p \mid n \Rightarrow p > z\}$$

$$= e^{-\gamma} \frac{x}{\log z} + o\left(\frac{x}{\log z}\right) + O\left(2^{-r} x + \sum_{\substack{d \mid P(z) \\ d < z^r}} |R_d|\right)$$

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 $o(\frac{x}{\log z})$. $\qquad \square$

3 Riemann Zeta function

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,

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

If $|R(u)| \leq \epsilon$ for all $u \geq M$ then

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

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. $\square$

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 $\sum_{n \leq N} \frac{f(n)}{n^\sigma} = 0$,

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

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

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.

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

$\square$

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:

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

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

as $y \to \infty$. $\square$

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,

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

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

$$= \frac{\lfloor x \rfloor}{x^s} + \frac{s}{s-1} \left[t^{-s+1} \right]_1^x - s \int_1^x \frac{\{t\}}{t^{s+1}} \, dt$$

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

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

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, $\ne 0$).

Proof.

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

$$0 < \int_1^\infty \frac{\{t\}}{t^{\sigma+1}} \, dt < \frac{1}{\sigma}. \qquad \square$$

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.

$$\zeta(s) - \frac{1}{s-1} = 1 - s \int_1^\infty \frac{\{u\}}{u^{s+1}} \, du$$

$$= \mathcal{O}(1) + \mathcal{O}\left(|s| \int_1^\infty \frac{1}{u^{\sigma+1}} \, du \right)$$

$$= \mathcal{O}(1) + \mathcal{O}_\delta(1). \qquad \square$$

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 $A(x) = \sum_{n \leq x} a_n$: If $\sigma > \max(0, \sigma_c)$ then

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

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

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

The converse is called Perron's formula:

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

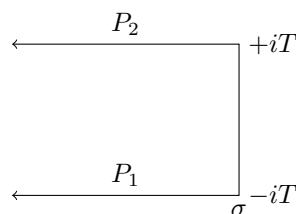Then, we can hope to understand $\psi(x)$ using analysis on $\frac{\zeta'}{\zeta}$. In particular, we get

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

'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 13   *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

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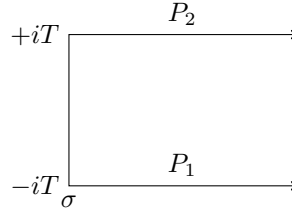
Similarly,

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

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

$$1_{n<x} = \frac{1}{2\pi i} \int_{\sigma_0-iT}^{\sigma_0+iT} \frac{(x/n)^s}{s} \, ds + \mathcal{O}\left(\frac{(x/n)^{\sigma_0}}{T \left|\log(\frac{x}{n})\right|}\right)$$

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

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

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

For the error term:

1. Contribution from $n \leq \frac{x}{2}$ or $n \geq 2x$, where $\left|\log(\frac{x}{n})\right| \gg 1$, 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

$$\frac{x^{\sigma_0}}{T} \sum_{\frac{x}{2} < n < 2x} \frac{|a_n|}{n^{\sigma_0}|\log(\frac{x}{n})|} \ll \frac{x^{\sigma_0}}{T} \sum_{\frac{x}{2} < n < 2x} \frac{|a_n|x}{n^{\sigma_0}|x-n|} \ll \frac{2^{\sigma_0}}{T} \sum_{\frac{x}{2} < n < 2x} \frac{|a_n|x}{|x-n|}.$$

$\square$

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 $|t| \ge \frac{7}{8}$ then $\zeta(s) \ne 0$ and $\frac{\zeta'(s)}{\zeta(s)} \ll \log(|t|+4)$.

2. $\zeta(s) \ne 0$ for $\frac{8}{9} \le \sigma \le 1$, $|t| \le \frac{7}{8}$.

3. $\frac{\zeta'(s)}{\zeta(s)} = -\frac{1}{s-1} + \mathcal{O}(1)$ for $1 - \frac{c}{\log(|t|+4)} < \sigma \le 2$ for $|t| \le \frac{7}{8}$.

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 $1 < \sigma_0 \le 2$

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

In the error term,

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

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 $\sigma_1 = 1 - \frac{c}{\log T}$.

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

and

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

$$\ll x^{\sigma_1} \log T + \frac{x^{\sigma_1}}{1 - \sigma_1} \ll x^{\sigma_1}(\log T)$$

Now,

$$\psi(x) = x + \mathcal{O}\left(\frac{x}{T}(\log x)^2 + x^{1-\frac{c}{\log T}}(\log T)\right)$$

$$= x + \mathcal{O}\left(\frac{x}{\exp(c\sqrt{\log x})}\right)$$
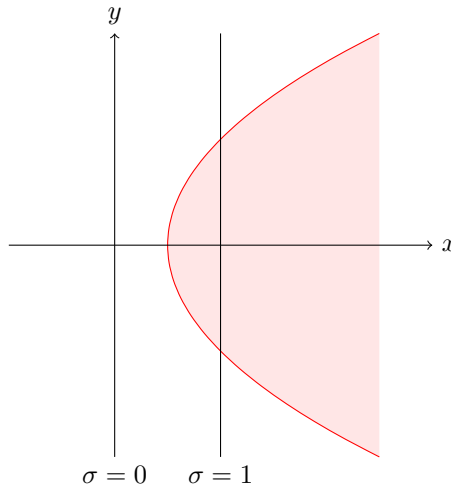
by choosing $T = \exp(c\sqrt{\log x})$. □

3.3   Zero-free region

Firstly, near $s = 1$, things are easy because of the pole.

Theorem. If $\sigma > (1 + t^2)/2$, then $\zeta(s) \neq 0$. In particular, $\zeta(s) \neq 0$ if $\frac{8}{9} \leq \sigma \leq 1$, $|t| \leq \frac{7}{8}$. Also,

$$\zeta(s) = \frac{1}{s-1} + \mathcal{O}(1) \quad -\frac{\zeta'}{\zeta}(s) = \frac{1}{s-1} + \mathcal{O}(1)$$

uniformly in $\frac{8}{9} \leq \sigma \leq 2$ and $|t| \leq \frac{7}{8}$.

Proof. Recall that

$$\zeta(s) = \frac{s}{s-1} + s\int_1^\infty \frac{\{u\}}{u^{s+1}}\,du,$$

so

$$\left|\zeta(s) - \frac{s}{s-1}\right| \leq |s|\int_1^\infty \frac{1}{u^{\sigma+1}}\,du \leq \frac{|s|}{\sigma}$$

so if $\sigma > |s - 1|$, $\zeta(s) \neq 0$, i.e. if $\sigma > \frac{1+t^2}{2}$ and $\frac{1+(\frac{7}{8})^2}{2} < \frac{8}{9}$. Also,

$$\left|\zeta(s) - \frac{1}{s-1}\right| \leq 1 + |s|\int_1^\infty \frac{1}{u^{\sigma+1}} = \mathcal{O}(1).$$

By general theory of holomorphic functions, the result holds for $-\frac{\zeta'}{\zeta}$ also. □

For $|t|$ large, need a different idea. How do we show that there aren't zeros on $\sigma = 1$? Suppose there is a zero, of order m, at $1 + it$. Then

$$-\frac{\zeta'}{\zeta}(1 + \delta + it) \sim \frac{m}{\delta}$$

$$-\sum \frac{\Lambda(n)}{n^{1+\delta+it}} \sim \frac{m}{\delta}$$

$$|\text{LHS}| \leq \sum \Lambda(n) n^{1+\delta} = -\frac{\zeta'}{\zeta}(1 + \delta) \sim \frac{1}{\delta}.$$

Then,

$$\sum_p \frac{\log p}{p^{1+\delta}} \cdot e^{it \log p} \sim -\sum \frac{\log p}{p^{1+\delta}}$$

$$\cos(t \log p) \approx -1 \text{ for almost all } p$$

so $p^{it} \approx -1$ for almost all p, i.e. $p^{2it} \approx 1$. Then there is a pole at $1 + 2it$.

Lemma (Borel-Carathéodory Lemma)**.** If f is holomorphic on $|z| \leq R$ and $f(0) = 0$. If $\operatorname{Re} f(z) \leq M$ for all $|z| \leq R$, for any $r < R$,

$$\sup_{|z| \leq r} (|f(z)|, |f'(z)|) \ll_{r,R} M$$

Proof. Let $g(z) = \frac{f(z)}{z(2M-f(z))}$. This is holomorphic in $|z| \leq R$. If $|z| = R$ then $|2M - f(z)| \geq |f(z)|$ and so

$$|g(z)| \leq \frac{|f(z)|}{R|f(z)|} \leq \frac{1}{R}.$$

So for all $|z| \leq r < R$, by maximum modulus,

$$|g(z)| = \frac{|f(z)|}{|z||2M - f(z)|} \leq \frac{1}{R}$$

so

$$R|f(z)| \leq r|2M - f(z)| \leq 2Mr + r|f(z)|.$$

So $|f(z)| \leq \frac{2Mr}{R-r} \ll M$. And for $f'(z)$, we use Cauchy's formula,

$$f'(z) = \frac{1}{2\pi i} \int_{|w|=r'} \frac{f(w)}{(z-w)^2} \, dw \quad r < r' < R$$

so $f'(z) \ll M$ also. $\qquad\square$

Lemma. If f is holomorphic on a domain including $|z| \leq 1$, $|f(z)| \leq M$ in this disc, and $f(0) \neq 0$. If $0 < r < R < 1$, then for $|z| \leq r$

$$\frac{f'}{f}(z) = \sum_{k=1}^{K} \frac{1}{z - z_k} + \mathcal{O}_{r,R}\left(\log \frac{M}{|f(0)|}\right)$$

where z_k ranges over all zeros of f in $|z| \leq R$.

Proof. Suppose that $f(0) = 1$ without loss of generality. Say first there are no zeros; consider $h(z) = \log f(z)$ and $\operatorname{Re} h(z) = \log |f(z)| \leq \log M$, so by B-C lemma,

$$|h'(z)| = \left| \frac{f'}{f}(z) \right| \ll \log M$$

as required.

In general, we define an auxiliary function g with no zeros:

$$g(z) := f(z) \prod_{k=1}^{K} \frac{R^2 - z\bar{z}_k}{(z - z_k)R}.$$

The kth factor has a pole at $z = z_k$ and on $|z| = R$ has modulus 1. So on $|z| \leq R$, $|g(z)| \leq M$, in particular

$$|g(0)| = \prod_{k=1}^{K} \frac{R}{|z_k|} \leq M.$$

Note also that $|g(0)| \geq 1$. Now, let

$$h(z) = \log \frac{g(z)}{g(0)}$$

and $\operatorname{Re} h(z) = \log |g(z)| - \log |g(0)| \leq M$ for $|z| \leq R$. By Borel-Carathéodory,

$$|h'(z)| = \left| \frac{f'}{f}(z) - \sum_{k=1}^{K} \frac{1}{z - z_k} + \sum_{k=1}^{K} \frac{1}{z - \frac{R^2}{\bar{z}_k}} \right| \ll \log M$$

so

$$\frac{f'}{f}(z) = \sum_{k=1}^{K} \frac{1}{z - z_k} - \sum_{k=1}^{K} \frac{1}{z - \frac{R^2}{\bar{z}_k}} + \mathcal{O}(\log M)$$

and if $|z| \leq r$,

$$\left| z - \frac{R^2}{\bar{z}_k} \right| \geq \frac{|R^2|}{|z_k|} - |z| \geq R - r \gg 1.$$

and $K \ll \log M$. $\qquad\square$

Corollary. If $|t| \geq \frac{7}{8}$ and $\frac{5}{6} \leq \sigma \leq 2$ then

$$\frac{\zeta'}{\zeta}(s) = \sum_{\rho} \frac{1}{s - \rho} + \mathcal{O}(\log |t|)$$

where ρ is over all zeros in $|\rho - (\frac{3}{2} + it)| \leq \frac{5}{6}$.

Lecture 15 **Theorem.** There is $c > 0$ such that $\zeta(s) \neq 0$ if $\sigma \geq 1 - \frac{c}{\log t}$.

Proof. Assume $\zeta(\rho) = 0$, for $\rho = \sigma + it$. Let $\delta > 0$ be chosen later. Then

$$\frac{\zeta'}{\zeta}(1 + \delta + it) = \frac{1}{1 + \delta + it - \rho} + \sum_{\rho' \neq \rho} \frac{1}{1 + \delta + it - \rho'} + \mathcal{O}(\log t),$$

(assume $\sigma \geq \frac{99}{100}$, say).

$$\operatorname{Re} \frac{\zeta'}{\zeta}(1 + \delta + it) = \operatorname{Re} \frac{1}{1 + \delta + it - \rho} + \operatorname{Re} \sum_{\rho' \neq \rho} \frac{1}{1 + \delta + it - \rho'} + \mathcal{O}(\log t)$$

$$> \frac{1}{1 + \delta - \sigma} + \mathcal{O}(\log t)$$

since $\operatorname{Re} \rho' \leq 1$ so $\operatorname{Re} \frac{1}{1 + \delta + it - \rho'} > 0$. Similarly,

$$\operatorname{Re} \frac{\zeta'}{\zeta}(1 + \delta + 2it) > \mathcal{O}(\log t).$$

Also,

$$\frac{\zeta'}{\zeta}(1 + \delta) = -\frac{1}{\delta} + \mathcal{O}(1).$$

Hence we can write

$$\frac{3}{\delta} - \frac{4}{1 + \delta - \sigma} + \mathcal{O}(\log t) > \operatorname{Re}\left(-3\frac{\zeta'}{\zeta}(1 + \delta) - 4\frac{\zeta'}{\zeta}(1 + \delta + it) - \frac{\zeta'}{\zeta}(1 + \delta + 2it)\right)$$

$$= \operatorname{Re}\left(3\sum_n \frac{\Lambda(n)}{n^{1+\delta}} + 4\sum_n \frac{\Lambda(n)}{n^{1+\delta+it}} - \sum_n \frac{\Lambda(n)}{n^{1+\delta+2it}}\right)$$

$$= \sum_n \frac{\Lambda(n)}{n^{1+\delta}}(3 + 4\cos(t\log n) + \cos(2t\log n)).$$

But $3 + 4\cos\theta + \cos 2\theta = 2(1 + \cos\theta)^2 \geq 0$, so

$$\frac{3}{\delta} \geq \frac{4}{1 + \delta - \sigma} + \mathcal{O}(\log t).$$

Choose $\delta = \frac{C}{\log t}$ for large enough C, so for contradiction, need

$$\frac{4}{1 + \delta - \sigma} < \frac{10}{\delta} \implies \sigma \geq 1 - \frac{c}{\log t}$$

for some $c > 0$. $\qquad\square$

Lemma. If $\sigma > 1 - \frac{c}{2\log t}$ and $|t| \geq \frac{7}{8}$ then

$$\left|\frac{\zeta'}{\zeta}(s)\right| \ll \log t$$

(with the same c as in the above theorem).

Proof. Let $s_1 = 1 + \frac{1}{\log t} + it = \sigma_1 + it$. Here

$$\left|\frac{\zeta'}{\zeta}(s_1)\right| \ll \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{\sigma_1}} \ll \frac{1}{\sigma_1 - 1} \ll \log t.$$

From the corollary, we have

$$\frac{\zeta'}{\zeta}(s_1) = \sum_\rho \frac{1}{s_1 - \rho} + \mathcal{O}(\log t)$$

so therefore

$$\operatorname{Re}\sum_{\rho}\frac{1}{s_1-\rho}\ll\log t.$$

Now if $s=\sigma+it$ where $\sigma\geq 1-\frac{c}{2\log t}$, then

$$\frac{\zeta'}{\zeta}(s)-\frac{\zeta'}{\zeta}(s_1)=\sum_{\rho}\left(\frac{1}{s-\rho}-\frac{1}{s_1-\rho}\right)+\mathcal{O}(\log t)..$$

Also $|s-\rho|\asymp|s_1-\rho|$, so

$$\left|\frac{1}{s-\rho}-\frac{1}{s_1-\rho}\right|\ll\frac{1}{|s_1-\rho|^2\log t}\ll\operatorname{Re}\frac{1}{s_1-\rho}$$

(since $\operatorname{Re}\frac{1}{z}=\frac{\operatorname{Re}z}{|z|^2}$). So

$$\sum_{\rho}\left|\frac{1}{s-\rho}-\frac{1}{s_1-\rho}\right|\ll\operatorname{Re}\sum_{\rho}\frac{1}{s_1-\rho}\ll\log t. \qquad\square$$

Assuming the Riemann Hypothesis, we can show

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

(sheet 3). Using partial summation, we can deduce that

$$\pi(x)=\operatorname{Li}(x)+\mathcal{O}_\epsilon(x^{\frac{1}{2}+\epsilon})$$
$$\text{where }\operatorname{Li}(x)=\int_2^x\frac{1}{\log t}\,dt$$
$$=\frac{x}{\log x}+\mathcal{O}\left(\frac{x}{(\log x)^2}\right)$$

So if $\pi(x)=\frac{x}{\log x}+E(x)$, then $E(x)\gg\frac{x}{(\log x)^2}$.

3.4 Error terms

In this section, we will show that $|\psi(x)-x|\gg x^{\frac{1}{2}}$ 'often'. Actually, we will show that $\psi(x)=x+\Omega_\pm(x^{\frac{1}{2}})$, i.e.

$$\limsup_{x\to\infty}\left(\frac{\psi(x)-x}{x^{\frac{1}{2}}}\right)>0$$
$$\liminf_{x\to\infty}\left(\frac{\psi(x)-x}{x^{\frac{1}{2}}}\right)<0.$$

Lecture 16   Our goal is to show that $\psi(x)=x+\Omega_\pm(x^{\frac{1}{2}})$ i.e.

$$\limsup_{x\to\infty}\left(\frac{\psi(x)-x}{x^{\frac{1}{2}}}\right)\geq c>0$$
$$\liminf_{x\to\infty}\left(\frac{\psi(x)-x}{x^{\frac{1}{2}}}\right)\leq -c<0$$

For contradiction, suppose that $\psi(x)-x\leq cx^{\frac{1}{2}}$ for all large x. So

$$cx^{\frac{1}{2}}-\psi(x)+x\geq 0,$$

take Mellin transform of this.

Lemma (Landau). Let $A(x)$ be integrable and bounded on any finite interval and $A(x) \geq 0$ for all $x \geq X$. Let

$$\sigma_c = \inf \left\{ \sigma : \int_X^\infty A(x) x^{-\sigma} \, dx < \infty \right\}$$

Then if

$$F(s) = \int_1^\infty A(x) x^{-s} \, dx$$

then F is analytic for $\operatorname{Re} s > \sigma_c$ and not at $s = \sigma_c$.

Proof. Divide integral into $[1, X]$ and (X, ∞), and write the corresponding partition of $F = F_1 + F_2$. F_1 is entire. If $\operatorname{Re} s > \sigma_c$, the integral converges absolutely, so F_2 is analytic for $\operatorname{Re} s > \sigma_c$. By contradiction, suppose F_2 is analytic at $s = \sigma_c$. Write F_2 as a Taylor series around $\sigma_c + 1$,

$$F_2(s) = \sum_{k=0}^\infty c_k (s - \sigma_c - 1)^k$$

where

$$c_k = \frac{F_2^{(k)}(\sigma_c + 1)}{k!} = \frac{1}{k!} \int_X^\infty A(x)(-\log x)^k x^{-\sigma_c - 1} \, dx.$$

This power series has a radius of convergence, which must be $1 + \delta$ for some $\delta > 0$ (since any poles are isolated). Evaluate the series at $s = \sigma_c - \frac{\delta}{2}$.

$$F_2(s) = \sum_{k=0}^\infty \frac{(1 - \sigma_c - s)^k}{k!} \int_X^\infty A(x)(\log x)^k x^{-1-\sigma_c} \, dx$$

At $s = \sigma_c - \frac{\delta}{2}$, can interchange integral and summation, so

$$F_2(\sigma_c - \frac{\delta}{2}) = \int_X^\infty A(x) x^{-1-\sigma_c} \exp((1 + \sigma_c - s) \log x) \, dx$$

$$= \int_X^\infty A(x) x^{-s} \, dx$$

so the integral converges at $\sigma_c - \frac{\delta}{2}$ contradicts the definition of σ_c. □

Theorem (Landau). If σ_0 is the supremum of the real parts of $\{\rho : \zeta(\rho) = 0\}$, then

1. for any $\sigma < \sigma_0$, $\psi(x) - x = \Omega_\pm(x^\sigma)$

2. if there is a zero ρ with $\sigma = \sigma_0$, then

$$\psi(x) - x = \Omega_\pm(x^{\sigma_0})$$

Corollary. Assuming there is a zero with $\sigma = \frac{1}{2}$ (there is: $\rho = \frac{1}{2} + 14.13472 \ldots i$)

$$\psi(x) - x = \Omega_\pm(x^{\frac{1}{2}}).$$

Sketch proof. Use cases of RH true or false. □

Corollary. The Riemann hypothesis is equivalent to $\psi(x) = x + \mathcal{O}(x^{\frac{1}{2}+o(1)})$.

Proof of theorem. Let $c > 0$ be chosen later, and suppose that

$$\psi(x) - x \le cx^\sigma \text{ for all } x \ge X.$$

Consider

$$F(s) = \int_1^\infty (cx^\sigma - \psi(x) + x)x^{-s-1}\, dx.$$

Recall that by partial summation,

$$\frac{\zeta'}{\zeta}(s) = -s\int_1^\infty \psi(x)x^{-s-1}\, dx \text{ and } \int_1^\infty x^{-s}\, dx = \frac{1}{s-1}$$

both for $\operatorname{Re} s > 1$. So,

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

for $\operatorname{Re} s > 1$. This has a pole at $s = \sigma$ and is analytic for $\operatorname{Re} s > \sigma$. So by Landau's lemma, in fact this integral converges for all s with $\operatorname{Re} s > \sigma$. This proves 1., because if $\sigma < \sigma_0$, then there is a zero of ζ with $\sigma < \operatorname{Re} \rho \le \sigma_0$, and at ρ, F has a singularity ($\rho \notin \mathbb{R}$).

Suppose there is $\rho = \sigma_0 + it_0$. Repeat the above with $\sigma = \sigma_0$. Consider instead

$$G(s) = F(s) + \frac{e^{i\theta}F(s+it_0) + e^{-i\theta}F(s-it_0)}{2}$$

for $\theta \in \mathbb{R}$. $G(s)$ still is analytic for $\operatorname{Re} s > \sigma_0$, and has a pole at $s = \sigma_0$. From $F(s)$, have residue c. From $F(s+it_0)$, have residue $\frac{m}{\rho}$ where m is the order of ρ. From $F(s-it_0)$, have residue $\frac{m}{\bar{\rho}}$ so $G(s)$ has a pole at $s = \sigma_0$ with residue

$$c + \frac{e^{i\theta}m}{2\rho} + \frac{e^{-i\theta}m}{2\bar{\rho}} = c - \frac{m}{|\rho|}$$

if I choose θ such that $\frac{e^{i\theta}}{\rho} = -\frac{1}{|\rho|}$. If I choose $c < \frac{m}{|\rho|}$, then this residue is < 0. So as $s \to \sigma_0$ from the right, along $\mathbb{R}$, $G(s) \to -\infty$. But for $\operatorname{Re} s > \sigma_0$,

$$G(s) = \int_1^\infty (cx^{\sigma_0} - \psi(x) + x)x^{-s-1}\left(\underbrace{1 + \frac{e^{i\theta}x^{-it_0}}{2} + \frac{e^{-i\theta}x^{it_0}}{2}}_{=1+\cos(\theta - t_0 \log x) \ge 0}\right)\, dx$$

so

$$G(s) = G_1(s) + G_2(s)$$

where $G_1(s)$ is an integral over $[1, X]$, which is entire, and $G_2(s)$ is a non-negative integral, a contradiction. This proves $\psi(x) - x = \Omega_+(x^\sigma)$. Ω_- is the same, just $x - 1$. $\square$

3.5   Functional equation

Recall that $\sigma > 0$,

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

Let $f(t) = \frac{1}{2} - \{t\}$, so

$$\zeta(s) = \frac{1}{s-1} + \frac{1}{2} + s \int_1^\infty \frac{f(t)}{t^{s+1}}\, dt.$$

This integral actually converges when $\sigma > -1$: let

$$F(x) = \int_0^x f(t)\, dt$$

so
$$\int_X^Y \frac{f(t)}{t^{s+1}}\, dt = \left[\frac{F(t)}{t^{s+1}} \right]_X^Y + (s+1) \int_X^Y \frac{F(t)}{t^{s+2}}\, dt$$

and note that $F(t)$ is bounded

Therefore,

$$\int_1^\infty \frac{f(t)}{t^{s+1}}\, dt$$

converges when $\sigma > -1$. For instance, $\zeta(0) = -\frac{1}{2} = 1 + 1 + 1 + \cdots$. Note that for $-1 < \sigma < 0$,

$$s \int_0^1 \frac{f(t)}{t^{s+1}}\, dt = \frac{s}{2} \int_0^1 \frac{1}{t^{s+1}}\, dt - s \int_0^1 \frac{1}{t^s}\, dt = \frac{1}{2} + \frac{1}{s-1}.$$

So for $-1 < \sigma < 0$,

$$\zeta(s) = s \int_0^\infty \frac{f(t)}{t^{s+1}}\, dt.$$

By Fourier analysis, $f(t)$ has a Fourier series,

$$f(t) = \sum_{n=1}^\infty \frac{\sin(2n\pi t)}{n\pi} \quad \text{converges for } t \notin \mathbb{Z}$$

so where $-1 < \sigma < 0$,

$$\zeta(s) = s \int_0^\infty \frac{1}{t^{s+1}} \sum_{n=1}^\infty \frac{\sin(2n\pi t)}{n\pi}\, dt$$

$$= s \sum_{n=1}^\infty \frac{1}{n\pi} \int_0^\infty \frac{\sin(2n\pi t)}{t^{s+1}}\, dt \quad \text{take } y = 2n\pi t$$

$$= s \sum_{n=1}^\infty \frac{(2n\pi)^s}{n\pi} \left(\int_0^\infty \frac{\sin(y)}{y^{s+1}}\, dy \right)$$

Here,

$$\sum_{n=1}^\infty \frac{(2n\pi)^s}{n\pi} = 2^s \pi^{s-1} \zeta(1-s)$$

and

$$\int_0^\infty \frac{\sin(y)}{y^{s+1}}\, dy = \frac{1}{2i} \left(\int_0^\infty \frac{e^{iy}}{y^{s+1}}\, dy - \int_0^\infty \frac{e^{-iy}}{y^{s+1}}\, dy \right)$$
$$= -\sin\left(\frac{s\pi}{2}\right) \Gamma(-s).$$

---

by setting $u = iy$ and $u = -iy$ in the two integrals respectively, where

$$\Gamma(s) := \int_0^\infty t^{s-1} e^{-t} \, dt \quad \text{for } \sigma > 0.$$

Take a moment to derive some important properties of the Γ function:

$$\begin{aligned}
\Gamma(s+1) &= \int_0^\infty t^s e^{-t} \, dt \\
&= [-t^s e^{-t}]_0^\infty + s \int_0^\infty t^{s-1} e^{-t} \, dt \\
&= s\Gamma(s).
\end{aligned}$$

In particular, since $\Gamma(1) = \int_0^\infty e^{-t} \, dt = 1$, we have $\Gamma(n) = (n-1)!$. Also note $\Gamma(s+1) = s\Gamma(s)$ allows us to extend $\Gamma(s)$ to $\mathbb{C}$, with poles at $s = 0, -1, -2, \dots$. For the zeta function, this means for $-1 < \sigma < 0$,

$$\begin{aligned}
\zeta(s) &= s 2^s \pi^{s-1} \zeta(1-s)(-\sin(\tfrac{s\pi}{2})\Gamma(-s)) \\
&= 2^s \pi^{s-1} \sin(\tfrac{\pi s}{2})\Gamma(1-s)\zeta(1-s).
\end{aligned}$$

The right hand side is defined for all $\sigma < 0$. Thus, define

$$\zeta(s) = 2^s \pi^{s-1} \sin(\tfrac{\pi s}{2})\Gamma(1-s)\zeta(1-s).$$

This gives an analytic continuation of $\zeta(s)$ to $\mathbb{C}$.

Theorem (Functional Equation). For all $s \in \mathbb{C}$,

$$\zeta(s) = 2^s \pi^{s-1} \sin(\tfrac{\pi s}{2})\Gamma(1-s)\zeta(1-s).$$

What about $s = 1$?

$$\zeta(1) = 2 \times 1 \times 1 \times \Gamma(0) \times \zeta(0).$$

Γ has a pole at 0, and ζ has a pole at 1, so this makes sense. Are there any other poles? The first three factors are entire, so cannot give poles. So, we only need to worry about $\Gamma(1-s)\zeta(1-s)$. These are both entire for $\sigma < 0$. So $\zeta(s)$ is analytic everywhere in $\mathbb{C}$ except for a simple pole at $s = 1$.

What about zeros of ζ?

$$0 = \zeta(s) = (\text{non-zero factors}) \sin(\frac{\pi s}{2})\Gamma(1-s)\zeta(1-s).$$

If $\sigma < 0$, $\zeta(1-s) \neq 0$ and $\Gamma(1-s) \neq 0$. So, $\sin(\frac{\pi s}{2}) = 0$, so $s = -2n$ for $n \in \mathbb{N}$. In $\sigma < 0$, $\zeta(s)$ has zeros at $-2, -4, -6, \dots$

If $\sigma > 1$, $\zeta(s)$ has no zeros. Similarly for $\sigma = 1$, so by the functional equation, similarly for $\sigma = 0$.

Thus, except for the trivial zeros, $\zeta(s)$ only has zeros in $0 < \sigma < 1$. If $0 < \sigma < 1$,

$$0 = \zeta(s) = (\text{non-zero factors})\Gamma(1-s)\zeta(1-s)$$

but $\Gamma(1-s)$ is nonzero in this region, so $\zeta(1-s) = 0$. Since $\zeta(\bar{s}) = \overline{\zeta(s)}$, these zeros come in fours (except if $\sigma = \frac{1}{2}$, where they come in pairs).

Theorem. Riemann Hypothesis is equivalent to $\psi(x) = x + \mathcal{O}(x^{\frac{1}{2}+o(1)})$

Proof.

$\Rightarrow$ Contour integration: exercise.

$\Leftarrow$ If $\sigma_0 = \sup\{\operatorname{Re}\rho : \zeta(\rho) = 0\}$ then

$$\psi(x) = x + \Omega_{\pm}(x^{\sigma}) \quad \forall \sigma < \sigma_0.$$

If the Riemann Hypothesis fails, there is a zero ρ with $0 < \sigma < 1$ such that $\sigma \neq \frac{1}{2}$. Therefore by symmetry, $\sigma_0 \geq \max(\sigma, 1 - \sigma) > \frac{1}{2}$, so $\psi(x) = x + \Omega_{\pm}(x^{\sigma'})$ where $\frac{1}{2} < \sigma' < \sigma$.

$\square$

4 Primes in Progressions

4.1 Dirichlet characters and L-functions

Definition. Fix $q \in \mathbb{N}$. A **Dirichlet character** of modulus q is a homomorphism $\chi : (\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}^\times$.

$(\mathbb{Z}/q\mathbb{Z})^\times$ is a finite abelian group of order $\varphi(q)$, so the set of Dirichlet characters of modulus q forms a finite abelian group of order $\varphi(q)$.

We can also think of χ as defining a function $\chi : \mathbb{Z} \to \mathbb{C}$, given by

$$\begin{cases} \chi(a \bmod q), & \text{if } (a,q) = 1 \\ 0 & \text{otherwise} \end{cases}$$

Note this χ is periodic with period q, and totally multiplicative. If χ is the trivial homomorphism on $(\mathbb{Z}/q\mathbb{Z})^\times$, we call it the principal Dirichlet character of modulus q and normally denote it as χ_0.

Lemma 4.1.

(1) Let χ be a Dirichlet character of modulus q. Then

$$\sum_{a \in (\mathbb{Z}/q\mathbb{Z})^\times} \chi(a) = \sum_{1 \leq a \leq q} \chi(a) = \begin{cases} \varphi(q) & \chi = \chi_0 \\ 0 & \chi \neq \chi_0. \end{cases}$$

(2) Let $a \in (\mathbb{Z}/q\mathbb{Z})^\times$. Then

$$\sum_\chi \chi(a) = \begin{cases} \varphi(q) & a \equiv 1 \pmod{q} \\ 0 & a \not\equiv 1 \pmod{q}. \end{cases}$$

Proof. We treat (2). If $a \equiv 1 \pmod q$ then $\chi(a) = 1$ for all χ. So

$$\sum_\chi \chi(a) = \sum_\chi 1 = \varphi(q).$$

If $a \not\equiv 1 \pmod q$ then there exists $\psi : (\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}^\times$ such that $\psi(a) \neq 1$. The map $\chi \mapsto \chi\psi$ is a permutation of the set of Dirichlet characters mod q. Hence

$$\sum_\chi \chi(a) = \sum_\chi (\chi\psi)(a) = \psi(a) \sum_\chi \chi(a) \Rightarrow \sum_\chi \chi(a) = 0. \qquad \square$$

Definition. Let $\mathbb{1}_{x \equiv a \bmod q} : \mathbb{Z} \to \mathbb{C}$ be defined by

$$\mathbb{1}(x) = \begin{cases} 1 & \text{if } (x,q) = 1, x \equiv a \bmod q \\ 0 & \text{otherwise} \end{cases}$$

for $a \in \mathbb{Z}$, $(a,q) = 1$.

Then Lemma 4.1 says

$$\mathbb{1}_{x \equiv a \bmod q}(x) = \frac{1}{\varphi(q)} \sum_\chi \chi(a)^{-1} \chi(x).$$

It follows that

$$\sum_{\substack{p \le x \\ p \equiv a \bmod q}} 1 = \sum_{p \le x} \mathbb{1}_{x \equiv a \bmod q}(p) = \frac{1}{\varphi(q)} \sum_{p \le x} \sum_{\chi} \chi(a)^{-1} \chi(p).$$

Estimating this is closely related to estimating

$$\frac{1}{\varphi(q)} \sum_{n \le x} \sum_{\chi} \chi(a)^{-1} \chi(n) \Lambda(n)$$

$$= \sum_{\chi} \frac{\chi(a)^{-1}}{\varphi(q)} \sum_{n \le x} \chi(n) \Lambda(n).$$

Our strategy to prove Dirichlet's theorem is to consider the contribution of each character χ separately.

We will do this using Dirichlet L-functions

$$L(s, \chi) = \sum_{n \ge 1} \frac{\chi(n)}{n^s}.$$

This series converges absolutely in the region $\sigma > 1$, and defines an analytic function there.

Lemma 4.2. If $\chi \ne \chi_0$, then $\sum_{n \ge 1} \chi(n) n^{-s}$ converges in $\sigma > 0$.

Proof. Use Partial summation with $a_n = \chi(n)$, $f(t) = t^{-s}$. This gives

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

where $A(x) = \sum_{n \le x} \chi(n)$. Note Lemma 4.1 says $\sum_{1 \le n \le q} \chi(n) = 0$, as $\chi \ne \chi_0$. Hence $A(n)$ is periodic, and $|A(x)| \le \varphi(q)$ for all x. Hence

$$\sum_{n \le x} \chi(n) n^{-s} = \frac{A(x)}{x^s} - \int_1^x A(t) f'(t)\, dt.$$

Note $\left| \frac{A(x)}{x^s} \right| \le \varphi(q) x^{-\sigma}$ and the integral is absolutely convergent. $\square$

Consequence: $L(s, \chi)$ is analytic in the region $\sigma > 0$, and in particular does not have a pole at $s = 1$. Since $\chi(n)$ is multiplicative, we have an Euler product identity

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

valid in the region $\sigma > 1$.

Remark. When $\chi = \chi_0$,

$$L(s, \chi_0) = \zeta(s) \prod_{p \mid q} \left(1 - \frac{1}{p^s} \right).$$

Hence $L(s, \chi_0)$ has a meromorphic continuation to all $s \in \mathbb{C}$ and a simple pole at $s = 1$.

Remark. In the region $\sigma > 1$, and for any χ, we have a formula for the logarithmic derivative of $L(s, \chi)$ in the same way as for $\zeta(s)$. This gives an identity

$$-\frac{L'}{L}(s, \chi) = \sum_{n \geq 1} \frac{\chi(n)\Lambda(n)}{n^s}.$$

Lecture 19   The Euler product we saw earlier gives

$$\log L(s, \chi) = \sum_p \sum_k \frac{\chi(p)^k p^{-ks}}{k},$$

hence

$$\frac{L'(s, \chi)}{L(s, \chi)} = \sum_p \sum_{k \geq 1} \chi(p)^k (-\log p) p^{-ks} = -\sum_{n \geq 1} \chi(n)\Lambda(n) n^{-s},$$

valid in $\sigma > 1$. Fix $a \in \mathbb{N}$ with $(a, q) = 1$. We combine this with the identity valid for any $n \in \mathbb{N}$:

$$\mathbb{1}_{n \equiv a \bmod q}(n) = \frac{1}{\varphi(q)} \sum_\chi \chi(a)^{-1}\chi(n).$$

We get

$$\sum_{n \geq 1} \mathbb{1}_{n \equiv a \bmod q}\Lambda(n) n^{-s} = -\frac{1}{\varphi(q)} \sum_\chi \chi(a)^{-1}\frac{L'(s, \chi)}{L(s, \chi)}$$

valid in $\sigma > 1$.

4.2   Dirichlet's theorem

Theorem (Dirichlet). Take $(a, q) = 1$. There are infinitely many primes p such that $p \equiv a \pmod q$.

We have

$$L(s, \chi_0) = \zeta(s) \prod_{p | q}(1 - p^{-s})$$

so $L(s, \chi_0)$ has a simple pole at $s = 1$. Hence

$$\sum_{n \geq 1} \mathbb{1}_{n \equiv a \bmod q}(n)\Lambda(n) n^{-s} = \frac{1}{\varphi(q)}\frac{1}{s - 1} + \mathcal{O}(1) - \sum_{\chi \neq \chi_0} \frac{\chi(a)^{-1}}{\varphi(q)}\frac{L'(s, \chi)}{L(s, \chi)}.$$

If there are finitely many primes $p \equiv a \pmod q$, the LHS would be bounded as $s \to 1$. To show Dirichlet's theorem, it's enough to show $\forall \chi \neq \chi_0$, $\frac{L'(s,\chi)}{L(s,\chi)}$ is analytic at $s = 1$.

This is equivalent to showing that if $\chi \neq \chi_0$, then $L(1, \chi) \neq 0$.

Theorem. If $\chi \neq \chi_0$, then $L(1, \chi) \neq 0$.

Proof.

$$\prod_\chi L(s,\chi) = \exp(\sum_\chi \log L(s,\chi))$$

$$= \exp\left(\sum_\chi \sum_p \sum_{k\geq 1} \chi(p)^k \frac{p^{-ks}}{k}\right)$$

$$= \exp(\sum_\chi \sum_{n\geq 1} \frac{\chi(n)n^{-s}\Lambda(n)}{\log n})$$

$$= \exp\left(\sum_{n\geq 1} \frac{n^{-s}\Lambda(n)}{\log n} \sum_\chi \chi(n)\right)$$

We know

$$\sum_\chi \chi(n) = \begin{cases} 0 & \text{if } (q,n) > 1 \text{ or } (q,n) = 1 \text{ and } n \not\equiv 1 \pmod q \\ \varphi(q) & \text{if } n \equiv 1 \pmod q. \end{cases}$$

Hence

$$\prod_\chi L(s,\chi) = \exp\left(\underbrace{\sum_{\substack{n\geq 1 \\ n\equiv 1 \bmod q}} \frac{n^{-s}\Lambda(n)}{\log n}\varphi(q)}_{\text{for } s \text{ real, } s > 1, \text{ this is a non-negative real number}} \right)$$

valid in $\sigma > 1$. So for $s \in (1,\infty)$, $\prod_\chi L(s,\chi) \in [1,\infty)$.

Note: $L(s,\chi_0)$ has a simple pole at $s = 1$. If there are at least two distinct characters ψ, ψ' of modulus q such that $L(1,\psi) = L(1,\psi') = 0$ then $\prod_\chi L(s,\chi)$ would be analytic in a neighbourhood of $s = 1$ and vanish at $s = 1$.

This can't happen, so there's at <u>most</u> one character ψ such that $L(1,\psi) = 0$.

Note: for any χ, $L(1,\overline{\chi}) = \overline{L(1,\chi)}$. If $L(1,\chi) = 0$, then $L(1,\overline{\chi}) = 0$. Hence if $L(1,\chi) = 0$, then $\chi = \overline{\chi}$. In other words, χ takes values in $\{\pm 1\}$ (we call such characters **quadratic** as then $\chi^2 = 1$.)

Suppose for contradiction there exists a non-principal quadratic character

$$\psi : (\mathbb{Z}/q\mathbb{Z})^\times \to \{\pm 1\}$$

such that $L(1,\psi) = 0$.

We consider the product $L(s,\phi)\zeta(s)$. This function is analytic in $\sigma > 0$. In $\sigma > 1$, we have the expression

$$L(s,\psi)\zeta(s) = \left(\sum_{n\geq 1} \psi(n)n^{-s}\right)\left(\sum_{n\geq 1} n^{-s}\right) = \sum_{n\geq 1} r(n)n^{-s} \tag{1}$$

where $r(n) = \sum_{d|n} \psi(d)$. Note $r(n)$ is multiplicative. Note also that $r(n) \geq 0$. It suffices to check this for prime powers p^k:

$$r(p^k) = \psi(1) + \psi(p) + \cdots + \psi(p^k) = \begin{cases} k+1 & \psi(p) = 1 \\ 1 & \psi(p) = 0 \text{ or } \psi(p) = -1, \ k \text{ is even} \\ 0 & \psi(p) = -1, k \text{ is odd}. \end{cases}$$

Note also that $r(n^2) \geq 1$, by the same argument. We now use Landau's lemma:

Lemma. Let $f(s) = \sum_{n \geq 1} a_n n^{-s}$, where a_n are non-negative real numbers. Suppose given $\sigma_0 \in \mathbb{R}$ such that $f(s)$ is convergent in $\sigma > \sigma_0$. Suppose that $f(s)$ admits an analytic continuation to the disk $\{|s - \sigma_0| < \epsilon\}$. Then $f(s)$ is convergent in $\sigma > \sigma_0 - \epsilon$.

Let $f(s) = L(\psi, s)\zeta(s) = \sum_{n \geq 1} r(n)n^{-s}$, valid in $\sigma > 1$. Then we can use Landau's lemma, together with the fact that $f(s)$ is analytic in $\sigma > 0$ to conclude that $f(s)$ is convergent in $\sigma > 0$. But

$$f(\frac{1}{2}) = \sum_{n \geq 1} r(n)n^{-\frac{1}{2}} \geq \sum_{n \geq 1} \frac{r(n^2)}{n} \geq \sum_{n \geq 1} \frac{1}{n},$$

and this series is divergent. This is a contradiction, so we must have $L(1, \psi) \neq 0$. $\square$

4.3 Zero-free region

We have proved there are ∞-many primes $\equiv a \bmod q$ for fixed a, q with $(a, q) = 1$ (Dirichlet's theorem), using

$$-\frac{L'}{L}(s, \chi) = \sum_{n=1}^{\infty} \frac{\Lambda(n)\chi(n)}{n^s}.$$

We want to prove a PNT for such primes. To do this, we'll use Perron's formula. We need information about the zeros of $L(s, \chi)$: we've already seen $L(1, \chi) \neq 0$.

Similarities to zero-free region for $\zeta(s)$, but important difference: $\zeta(s)$ has a pole at $s = 1$, $L(s, \chi)$ has no poles for $\sigma > 0$ (for $\chi \neq \chi_0$). The pole helps us to get zero-free regions for ζ, so it is harder for $L(s, \chi)$.

Let $\tau = |t| + 4$. Recall

Lemma. If $f(z)$ is analytic on a region containing $|z| \leq 1$ and $f(0) \neq 0$ and $|f(z)| \leq M$ for $|z| \leq 1$. Then for $0 < r < R < 1$, for $|z| \leq r$,

$$\frac{f'}{f}(z) = \sum \frac{1}{z - z_k} + \mathcal{O}\left(\log \frac{M}{|f(0)|}\right)$$

where the sum is over zeros of f in $|z_k| \leq R$.

Lemma 4.3. If $\chi \neq \chi_0$ and $\frac{5}{6} \leq \sigma \leq 2$ then

$$\frac{L'}{L}(s, \chi) = \sum_{\rho} \frac{1}{s - \rho} + \mathcal{O}(\log q\tau)$$

over ρ with $|\rho - (\frac{3}{2} + it)| \leq \frac{5}{6}$ (where q came from the Dirichlet character).

Proof. Follows from the lemma with $f(z) = L(z + \frac{3}{2} + it, \chi)$ with $R = \frac{5}{6}$ and $r = \frac{2}{3}$. Note

$$|f(0)| = \left|L\left(\frac{3}{2} + it, \chi\right)\right| = \prod_p \left|1 - \frac{\chi(p)}{p^{\frac{3}{2}+it}}\right| \geq \prod_p (1 + \frac{1}{p^{\frac{3}{2}}}) \gg 1.$$

By [Partial summation], if $F(t) = \sum_{1 \leq n \leq t} \chi(n)$, for $\sigma > 0$,

$$L(s, \chi) = s \int_1^{\infty} \frac{F(t)}{t^{s+1}} dt$$

so

$$|L(s, \chi)| \ll |s|q \int_1^{\infty} \frac{1}{t^{\sigma+1}} dt \ll q\tau. \qquad \square$$

Theorem. Let χ be a non-quadratic character. There is an absolute constant $c > 0$ such that

$$L(s, \chi) \neq 0 \quad \text{if} \quad \sigma > 1 - \frac{c}{\log(q\tau)}.$$

Proof. Since

$$L(s, \chi_0) = \zeta(s) \prod_{p|q} (1 - p^{-s})$$

in this region $\sigma > 0$, zeros of $L(s, \chi_0)$ are the same as $\zeta(s)$, so done. Let $\rho = \sigma + it$ such that $L(\rho, \chi) = 0$. Idea is to compare (as $\delta \to 0$),

$$\frac{L'}{L}(1 + \delta + it, \chi), \frac{L'}{L}(1 + \delta + 2it, \chi^2) \text{ and } \frac{L'}{L}(1 + \delta, \chi_0)$$

Note that

$$\text{Re}\left(-3\frac{L'}{L}(1 + \delta, \chi_0) - 4\frac{L'}{L}(1 + \delta + it, \chi) - \frac{L'}{L}(1 + \delta + 2it, \chi^2)\right)$$

$$= \sum_{\substack{n=1 \\ (n,q)=1}} \frac{\Lambda(n)}{n^{1+\delta}} \text{Re}(3 + 4\chi(n)n^{-it} + n^{-2it})$$

and $\forall \theta$, $3 + 4\cos\theta + \cos(2\theta) \geq 0$. But the left hand side is $\text{Re}(3 + 4e^{i\theta} + e^{2i\theta})$, i.e. $\forall z$, $|z| = 1$, $\text{Re}(3 + 4z + z^2) \geq 0$.

By lemma,

$$-\text{Re}\frac{L'}{L}(1 + \delta, \chi_0) = \frac{1}{\delta} + \mathcal{O}(\log q)$$

$$-\text{Re}\frac{L'}{L}(1 + \delta + it, \chi) \leq -\frac{1}{1 + \delta - \sigma} + \mathcal{O}(\log q\tau)$$

$$\text{and } \text{Re}\frac{L'}{L}(1 + \delta + 2it, \chi^2) \ll \log(q\tau)$$

since $\chi^2 \neq \chi_0$ so

$$\frac{3}{\delta} - \frac{4}{1 + \delta - \sigma} + \mathcal{O}(\log q\tau) \geq 0,$$

contradiction if $\delta \approx \frac{c'}{\log q\tau}$ and $\sigma \geq 1 - \frac{c}{\log q\tau}$. $\qquad\square$

Theorem 4.4. If χ is a quadratic character, $\exists c > 0$ such that

$$L(s, \chi) \neq 0 \quad \text{if} \quad \sigma > 1 - \frac{c}{\log q\tau} \quad \text{and } t \neq 0.$$

We cannot rule out a zero ρ of $L(s, \chi)$ with $\rho \in \mathbb{R}$ close to 1.

Theorem 4.5. Let χ be a quadratic character. There is an absolute constant $c > 0$ such that $L(s, \chi)$ has at most one zero $\rho \in (0, 1)$ such that $\rho \geq 1 - \frac{c}{\log q}$.

Such zeros are called 'exceptional zeros' or 'Siegel zeros'.
First, we need a lemma for $L(s, \chi_0)$.

Lemma 4.6. If $\frac{5}{6} \leq \sigma \leq 2$ then

$$-\frac{L'}{L}(s, \chi_0) = \frac{1}{s-1} - \sum_{\rho} \frac{1}{s-\rho} + \mathcal{O}(\log q\tau)$$

over zeros ρ with $|\rho - (\frac{3}{2} + it)| \leq \frac{5}{6}$.

Proof. Follows from

$$-\frac{\zeta'}{\zeta}(s) = -\sum_{\rho} \frac{1}{s-\rho} + \mathcal{O}(\log \tau) + \frac{1}{s-1}$$

since

   (a) for $\sigma > 0$, zeros of ζ are exactly the zeros of $L(s, \chi_0)$

   (b) by the Euler product,

$$\frac{L'}{L}(s, \chi_0) = \frac{\zeta'}{\zeta} + \sum_{p|q} \frac{\log p}{p^s - 1} \ll \omega(q) \ll \log q.$$

$\square$

Theorem 4.4 proof sketch. For t large, same as previous proof ($\chi^2 = \chi_0$, but no pole). For t small, $0 < |t| \ll \frac{1}{\log q\tau}$, instead of comparing χ_0, χ, χ^2, compare $\rho, \bar{\rho}$. $\square$

Theorem 4.5 proof sketch. Compare two such real zeros. $\square$

Lecture 21   *Proof of Theorem 4.4.* As before, let $\rho = \sigma + it$ be a zero of $L(s, \chi)$. Let $\delta > 0$. By Lemma 4.3,

$$-\frac{L'}{L}(1 + \delta + it, \chi) = -\sum_{\rho'} \frac{1}{1 + \delta + it - \rho} + \mathcal{O}(\log q\tau)$$

$$-\operatorname{Re} \frac{L'}{L}(1 + \delta + it, \chi) \leq -\frac{1}{1 + \delta + it - \rho} + \mathcal{O}(\log q\tau)$$

$$= -\frac{1}{1 + \delta - \sigma} + \mathcal{O}(\log q\tau).$$

By Lemma 4.6,

$$-\operatorname{Re} \frac{L'}{L}(1 + \delta, \chi_0) \leq \frac{1}{\delta} + \mathcal{O}(\log q\tau).$$

First suppose $\tau \geq C(1 - \sigma)$. Here,

$$-\operatorname{Re} \frac{L'}{L}(1 + \delta + 2it, \chi^2) = -\operatorname{Re} \frac{L'}{L}(1 + \delta + 2it, \chi_0)$$

$$\leq \operatorname{Re} \frac{1}{\delta + 2it} + \mathcal{O}(\log q\tau)$$

$$\leq \frac{\delta}{\delta^2 + 4t^2} + \mathcal{O}(\log q\tau).$$

As before,

$$\operatorname{Re}\left(-3\frac{L'}{L}(1 + \delta, \chi_0) - 4\frac{L'}{L}(1 + \delta + it, \chi) - \frac{L'}{L}(1 + \delta + 2it, \chi^2)\right) \geq 0$$

---

but also

$$\leq \frac{3}{\delta} - \frac{4}{1 + \delta - \sigma} + \frac{\delta}{\delta^2 + 4t^2} + \mathcal{O}(\log q\tau).$$

If $\sigma = 1$, contradiction as $\delta \to 0$. For $\sigma \neq 1$, if we choose $\delta = c(1 - \sigma)$, then this is

$$0 \leq \frac{3}{c(1 - \delta)} - \frac{4}{(c+1)(1-\sigma)} + \frac{c'}{1 - \sigma} + \mathcal{O}(\log q\tau)$$

$(\tau \gg 1 - \sigma \gg \delta)$. Can choose c, C, hence c' such that this is $\leq \frac{c''}{1-\sigma} + \mathcal{O}(\log q\tau)$ and so $\sigma \leq 1 - \frac{c'''}{\log q\tau}$.

For small τ we need a different argument. Since $L(\rho, \chi) = 0$, also $L(\bar{\rho}, \chi) = 0$ (and $\rho \neq \bar{\rho}$), since

$$L(s, \chi) = s \int_1^\infty \frac{\sum_{1 \leq n \leq t} \chi(n)}{t^{s+1}} \, dt.$$

It follows that

$$-\operatorname{Re} \frac{L'}{L}(1 + \delta + it, \chi) \leq -\operatorname{Re} \frac{1}{1 + \delta - \rho} - \operatorname{Re} \frac{1}{1 + \delta - \bar{\rho}} + \mathcal{O}(\log q\tau)$$

(assuming that $|t| \leq c(1 - \sigma)$, in particular $|t| \leq c'$ small constant). The RHS is

$$\frac{-2(1 + \delta - \sigma)}{(1 + \delta - \sigma)^2 + t^2} + \mathcal{O}(\log q\tau).$$

As before,

$$-\operatorname{Re} \frac{L'}{L}(1 + \delta, \chi_0) \leq \frac{1}{\delta} + \mathcal{O}(\log q\tau)..$$

Now,

$$-\operatorname{Re} \frac{L'}{L}(1 + \delta, \chi_0) - \operatorname{Re} \frac{L'}{L}(1 + \delta + it, \chi) = \sum_{\substack{n=1 \\ (n,q)=1}} \frac{\Lambda(n)}{n^{1+\delta}}(1 + \operatorname{Re}(\chi(n)n^{it})) \geq 0$$

since the absolute value of $\chi(n)n^{it}$ is 1. So

$$\frac{1}{\delta} - \frac{2(1 + \delta - \sigma)}{(1 + \delta - \sigma)^2 + t^2} + \mathcal{O}(\log \tau) \geq 0..$$

If we choose $\delta = c(1 - \sigma)$, the HHS is $\leq \frac{c'}{1-\sigma} + \mathcal{O}(\log q\tau)$. So $\sigma \leq 1 - \frac{c'}{\log q\tau}$. $\qquad \square$

Proof. Suppose $\rho_0 < \rho_1 \leq 1$ are zeros of $L(s, \chi)$. Then for $\sigma \in (0, 1)$,

$$-\operatorname{Re} \frac{L'}{L}(\sigma, \chi) \leq -\operatorname{Re} \frac{1}{\sigma - \rho_0} - \operatorname{Re} \frac{1}{\sigma - \rho_1} + \mathcal{O}(\log q)$$

for $\sigma \geq 1 - \frac{1}{1000000}$, say

$$\leq -\frac{2}{\sigma - \rho_0} + \mathcal{O}(\log q).$$

So

$$\frac{1}{\sigma - 1} - \frac{2}{\sigma - \rho_0} + \mathcal{O}(\log q) \geq \left(-\operatorname{Re} \frac{L'}{L}(\sigma, \chi_0) - \operatorname{Re} \frac{L'}{L}(\sigma, \chi)\right) \geq 0.$$

Hence

$$\rho_0 \leq 1 - \frac{c}{\log q}.$$

$\qquad \square$

Lemma. If $\chi \neq \chi_0$ and $\sigma \geq 1 - \frac{c}{\log q\tau}$ for some absolute $c > 0$, then if either χ has no exceptional zero, or χ has an exceptional zero at β, but $|s - \beta| \geq \frac{1}{\log q}$, then

$$\frac{L'}{L}(s, \chi) \ll \log q\tau.$$

Proof. If $\sigma > 1$, note

$$\left|\frac{L'}{L}(s, \chi)\right| \leq \sum_{\substack{n=1 \\ (n,q)=1}}^{\infty} \frac{\Lambda(n)}{n^\sigma} \ll \frac{1}{\sigma - 1}.$$

In particular, if $s = \sigma + it$ and $s_1 = 1 + \frac{1}{\log q\tau} + it$, then

$$\left|\frac{L'}{L}(s, \chi)\right| \ll \log q\tau.$$

By [Lemma 4.3](),

$$\frac{L'}{L}(s, \chi) = \sum_\rho \frac{1}{s - \rho} + \mathcal{O}(\log q\tau)$$

for all zeros ρ, $|s - \rho| \asymp |s_1 - \rho|$. So

$$\left|\frac{L'}{L}(s, \chi) - \frac{L'}{L}(s, \chi)\right| \ll \left|\sum_\rho \frac{1}{s - \rho} - \frac{1}{s_1 - \rho}\right| + \mathcal{O}(\log q\tau)$$

$$\ll \operatorname{Re} \sum_\rho \frac{1}{s - \rho} + \mathcal{O}(\log q\tau)$$

$$\ll \mathcal{O}(\log q\tau). \qquad \square$$

4.4   Prime Number Theorem for Arithmetic Progressions

Recall that

$$\sum_{\substack{1 \leq n \leq x \\ n \equiv a \bmod q}} \Lambda(n) = \frac{1}{\varphi(q)} \sum_\chi \overline{\chi(a)} \sum_{1 \leq n \leq x} \Lambda(n)\chi(n)$$

$$= \frac{1}{\varphi(q)} \sum_\chi \overline{\chi(a)} \psi(x, \chi).$$

Theorem. If $q \leq \exp(\mathcal{O}(\sqrt{\log x}))$, then

1. $\psi(x, \chi_0) = x + \mathcal{O}(x \exp(-c\sqrt{\log x}))$

2. If $\chi \neq \chi_0$ and χ has no exceptional zero, then

$$\psi(x, \chi) = \mathcal{O}(x \exp(-c\sqrt{\log x})).$$

3. If $\chi \neq \chi_0$ and χ has an exceptional zero at β then

$$\psi(x, \chi) = -\frac{x^\beta}{\beta} + \mathcal{O}(x \exp(-c\sqrt{\log x})).$$

Theorem. If χ_1, χ_2 are distinct quadratic characters modulo q, then $L(s, \chi_1)L(s, \chi_2)$ has at most one real zero β with $1 - \frac{c}{\log q} < \beta < 1$.

Proof. Say β_i is a real zero of $L(s_i, \chi_i)$ for $i = 1, 2$. Without loss of generality $\frac{5}{6}\beta_1 \leq \beta_2 < 1$. Fix $\delta \geq 0$.

1.
$$-\operatorname{Re} \frac{L'}{L}(1 + \delta, \chi_i) \leq -\frac{1}{1 + \delta - \beta_i} + \mathcal{O}(\log q) \quad (i = 1, 2)$$

2.
$$-\operatorname{Re} \frac{L'}{L}(1 + \delta, \chi_1\chi_2) \leq \mathcal{O}(\log q) \quad (\chi_1\chi_2 \neq \chi_0)$$

3.
$$-\frac{\zeta'}{\zeta}(1 + \delta) \leq \frac{1}{\delta} + \mathcal{O}(1) \quad \left(-\frac{\zeta'}{\zeta}(s) = \frac{1}{s - 1} + \mathcal{O}(1)\right)$$

(could use $-\operatorname{Re}\frac{L'}{L}(1 + \delta, \chi_0)$)

Therefore
$$-\frac{\zeta'}{\zeta}(1 + \delta) - \frac{L'}{L}(1 + \delta, \chi_1) - \frac{L'}{L}(1 + \delta, \chi_2) - \frac{L'}{L}(1 + \delta, \chi_1\chi_2)$$
$$\leq \frac{1}{\delta} - \frac{2}{1 + \delta - \beta_1} + \mathcal{O}(\log q)$$

(since the LHS is real anyway). But the LHS is
$$\sum \frac{\Lambda(n)}{n^{1+\delta}} \operatorname{Re}(1 + \chi_1(n) + \chi_2(n) + \chi_1\chi_2(n)) \geq 0$$

since $(1 + \chi_1(n))(1 + \chi_2(n)) \geq 0$. Choose $\delta = c(1 - \beta_1)$ and therefore $\beta_1 \leq 1 - \frac{c}{\log q}$. $\square$

Recall that
$$1_{n \equiv a \bmod q} = \frac{1}{\varphi(q)} \sum_{\chi} \overline{\chi(a)}\chi(n).$$

$$\psi(x; q, a) = \sum_{\substack{n \leq x \\ n \equiv a \bmod q}} \Lambda(n) = \frac{1}{\varphi(q)} \sum_{\chi} \overline{\chi(a)}\psi(x, \chi).$$

Corollary (to Theorem before this one)**.** If $(a, q) = 1$ then (for $q \leq \exp(\mathcal{O}(\sqrt{\log x}))$),
$$\psi(x, q, a) = \frac{x}{\varphi(q)} + \mathcal{O}(x\exp(-c\sqrt{\log x}))$$

assuming there is no exceptional zero. If q has an exceptional zero at β and χ_1, then
$$\psi(x; q, a) = \frac{x}{\varphi(q)} - \frac{\chi_1(a)}{\varphi(q)}\frac{x^\beta}{\beta} + \mathcal{O}(x\exp(-c\sqrt{\log x}))$$

Proof of Theorem before this one. By Perron's formula, $(\sigma_0 > 1, T \geq 1)$
$$\psi(x, \chi) = -\frac{1}{2\pi i}\int_{\sigma_0 - iT}^{\sigma_0 + iT} \frac{L'}{L}(s, \chi)\frac{x^s}{s}\,ds + \mathcal{O}\left(\frac{x}{T}\sum_{\frac{x}{2} < n < 2x} \frac{\Lambda(n)}{|x - n|} + \frac{x^{\sigma_0}}{T}\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{\sigma_0}}\right)$$

By the same argument as for $\zeta(s)$, error term is $\ll \frac{x(\log x)^2}{T}$ (using $\sigma_0 = 1 + \frac{1}{\log x}$). Take C to be a rectangular contour with corners at $\sigma_0 \pm iT$, $\sigma_1 \pm iT$. So
$$\psi(x, \chi) = \frac{1}{2\pi i}\int_C + \mathcal{O}\left(\int_{\sigma_1 \pm iT} + \int_{\sigma_0 + iT}^{\sigma_1 + iT} + \int_{\sigma_0 - iT}^{\sigma_1 - iT} + \frac{x(\log x)^2}{T}\right)$$

Error terms we bound as for $\zeta(s)$, so in total,

$$\psi(x,\chi) = -\frac{1}{2\pi i}\int_C \frac{L'}{L}(s,\chi)\frac{x^s}{s}\,ds + \mathcal{O}\underbrace{\left(\frac{x(\log x)^2}{T} + x^{1-q\sigma_1}\right)}_{\ll\exp(-c\sqrt{\log x}),\ T=\exp(\mathcal{O}(\sqrt{\log x}))}$$

We have $\sigma_1 = 1 - \frac{c}{\log qT}$ so $x^{\sigma_1} \ll x\exp(-c\sqrt{\log x})$ if $q \ll T \approx \exp(\mathcal{O}(\sqrt{\log x}))$.
Main term?

1. If $\chi = \chi_0$, then take σ_1 as above, so no zeros of $L(s,\chi_0)$. So $\frac{L'}{L}$ has just a simple pole at $s = 1$, and $\frac{1}{2\pi i}\int_C = x$.

2. If $\chi \neq \chi_0$, and no exceptional zero, then there are no zeros of $L(s,\chi)$ with $\sigma \geq \sigma_1$, so no poles of $\frac{L'}{L}(s,\chi)$, so $\int_C = 0$.

3. If χ has an exceptional zero at β, then inside C, $\frac{L'}{L}$ has a pole at β. So $\frac{L'}{L}(s,\chi)\frac{x^s}{s}$ has residue $\frac{x^\beta}{\beta}$ at this pole, so

$$\frac{1}{2\pi i}\int_C \frac{L'}{L}(s,\chi)\frac{x^s}{s}\,ds = \frac{x^\beta}{\beta}. \qquad \square$$

4.5   Siegel-Walfisz Theorem

Theorem (Siegel-Walfisz). $\forall A \geq 0$, if $q \leq (\log x)^A$ and $(a,q) = 1$, then

$$\psi(x;q,a) = \frac{x}{\varphi(q)} + \mathcal{O}_A(x\exp(-c\sqrt{\log x}))$$

This follows from

Theorem. If $q \leq (\log x)^A$ and x is large enough (depending on A) then

$$\psi(x,\chi) = \mathcal{O}_A(\exp(-c\sqrt{\log x})) \quad \forall \chi \neq \chi_0$$

This in turn follows from:

Theorem (Siegel). $\forall \epsilon > 0$, $\exists c_\epsilon$ such that if χ is a quadratic character modulo q, and β is a real zero, then

$$\beta < 1 - c_\epsilon q^{-\epsilon}.$$

Proof. Omitted. $\qquad \square$

The constant c_ϵ is ineffective: the proof gives no way to calculate c_ϵ.

Thm 3 $\Rightarrow$ Thm 2. If an exceptional zero exists, $\beta < 1 - c_\epsilon q^\epsilon$ for every $\epsilon > 0$, then

$$\psi(x,\chi) = \mathcal{O}\left(\frac{x^\beta}{\beta} + x\exp(-c\sqrt{\log x})\right)$$
$$= x\,o\left(\exp(-c_\epsilon q^\epsilon \log x) + \exp(-c\sqrt{\log x})\right).$$

Since $q \leq (\log x)^A$, this is $\mathcal{O}(\exp(-c'_\epsilon\sqrt{\log x}))$ choosing $\epsilon = \frac{1}{3A}$, say. $\qquad \square$

Corollary. If $(a, q) = 1$ then

$$\pi(x; q, a) = \frac{\text{Li}(x)}{\varphi(q)} + \mathcal{O}\left(x \exp(-c\sqrt{\log x})\right)$$

if $q \leq (\log x)^A$ (unconditionally) or if $q \leq \exp\left(\mathcal{O}(\sqrt{\log x})\right)$ and q has no exceptional zero.

Note that assuming GRH, the bound on q when q has no exceptional zero can be improved to $q \leq x^{\frac{1}{2} - o(1)}$.

Proof. Let

$$F(x) = \sum_{\substack{p \leq x \\ p = a \bmod q}} \log p = \psi(x; q, a) + \mathcal{O}(x^{1/2})$$

and so

$$\pi(x; q, a) = \sum_{\substack{p \leq x \\ p = a \bmod q}} 1$$

$$= \frac{F(x)}{\log x} + \int_x^x \frac{F(t)}{t(\log t)^2} dt$$

$$= \frac{1}{\varphi(q)} \left(\frac{x}{\log x} + \underbrace{\int_2^x \frac{1}{(\log t)^2} dt}_{=\text{Li}(x)} \right) + \mathcal{O}(x \exp(-c\sqrt{\log x})). \qquad \square$$

Two applications of Siegel-Walfisz:

1. For fixed $(a, q) = 1$, how large is the smallest prime $\equiv a \pmod q$? Call this prime $p_{a,q}$.

   Corollary. $\forall \epsilon > 0$, $p_{a,q} \ll_\epsilon \exp(q^\epsilon)$.

   Proof. Let $x < p_{a,q}$, so $\psi(x; q, a) = 0$, so if $q \leq (\log x)^A$ then

   $$\frac{x}{\varphi(q)} \mathcal{O}_A(x \exp(-c\sqrt{\log x})),$$

   so $\exp(c\sqrt{\log x}) = \mathcal{O}_A(q)$, so $\log x \leq (\log q)^2 + \mathcal{O}_A(1)$, contradiction to $q \leq (\log x)^A$ i.e. for any A, if q is large enough, $q \leq (\log p_{a,q})^A$. $\square$

   Similarly,

   Corollary. If q has no exceptional zeros, then $p_{a,q} \leq q^{\mathcal{O}(\log q)}$.

   Conjecture. $p_{a,q} \leq q^{1 + o(1)}$ where the exponent $\to 1$ as $q \to \infty$.

   On GRH, $p_{a,q} \leq q^{2 + o(1)}$.

   Theorem (Linnik). There is a constant L such that $p_{a,q} \ll q^l$.

   (Xylouris 2011 gave $L = 5$).

   Proof. Non-examinable. See Iwaniec-Kowalski. $\square$

2.

Theorem (Walfisz). For any n,

$$r(n) = \#\text{ways of writing } n = p + (\square\text{-free}) \sim c_n \operatorname{li}(n)$$

where

$$c_n = \left(\prod_{p|n} \left(1 + \frac{1}{p^2 - p - 1} \right) \right) \left(\prod_p \left(1 - \frac{1}{p(p-1)} \right) \right)$$

Proof. Note that $1_{\square\text{-free}}(n) = \sum_{d^2|n} \mu(d)$, easily checked since both sides are multiplicative, enough to check for prime powers. So

$$r(n) = \sum_{p<n} 1_{\square\text{-free}}(n-p)$$

$$= \sum_{p<n} \sum_{d^2|(n-p)} \mu(d)$$

$$= \sum_{d<\sqrt{n}} \mu(d) \sum_{\substack{p<n \\ p \equiv n \bmod d^2}} 1$$

$$= \sum_{d<\sqrt{n}} \mu(d) \pi(n-1; d^2, n)$$

If $(n, d) > 1$, then $\pi(n-1; d^2, n) = \mathcal{O}(1)$, so in total this contributes $\mathcal{O}(n^{\frac{1}{2}})$ to $r(n)$. When $(d, n) = 1$ and $d \leq (\log n)^A$,

$$\pi(n-1; d^2, n) = \frac{\operatorname{li}(n)}{\varphi(d^2)} + \mathcal{O}(n \exp(-c\sqrt{\log n}))$$

So

$$\sum_{d<(\log n)^A} \mu(d) \pi(n-1; d^2, n) = \operatorname{li}(n) \sum_{\substack{d<(\log n)^A \\ (d,n)=1}} \frac{\mu(d)}{\varphi(d^2)} + \mathcal{O}(n \exp(-c\sqrt{\log n}))$$

Note that $\varphi(d^2) = d\varphi(d)$, and so

$$\sum_{(d,n)=1} \frac{\mu(d)}{d\varphi(d)} = \prod_{p \nmid n} \left(1 - \frac{1}{p(p-1)} \right) = c_n.$$

Since

$$\sum_{\substack{d>(\log n)^A \\ (d,n)}} \frac{\mu(d)}{d\varphi(d)} \leq \sum_{d>(\log n)^A} \frac{1}{d^{\frac{3}{2}}} \ll \frac{1}{(\log n)^{\frac{A}{2}}} = o(1)$$

as $n \to \infty$. For $d > (\log n)^A$, use (the cheap trick that) $\pi(x; q, a) \ll 1 + \frac{x}{q}$. So

$$\sum_{\substack{n^{\frac{1}{2}} > d > (\log n)^A \\ (d,n)=1}} \mu(d) \pi(n-1; d^2, n) \ll \sum_{(\log n)^A < d < n^{\frac{1}{2}}} \left(1 + \frac{n}{d^2} \right)$$

$$\ll n^{\frac{1}{2}} + n \sum_{d>(\log n)^A} \frac{1}{d^2} \ll \frac{n}{(\log n)^A}.$$

That is,

$$r(n) = c_n \operatorname{li}(n) + \mathcal{O}\left(n^{\frac{1}{2}} + \frac{\operatorname{li}(n)}{(\log n)^{\frac{A}{2}}} + \frac{n}{(\log n)^A} + n \exp(-c\sqrt{\log n})\right)$$

$$= (1 + o(1)) c_n \operatorname{li}(n). \qquad \Box$$

5 *Some highlights of analytic number theory

5.1 Gaps between primes

Note that the prime number theorem gives $p_n \sim n \log n$, so

$$p_{n+1} - p_n \sim (n+1)\log(n+1) - n \log n \sim n \log n.$$

5.1.1 Small gaps

The twin prime conjecture says $p_{n+1} - p_n$ infinitely often.

Goldston-Pintz-Yildimm (2005) showed

$$\liminf_{n \to \infty} \frac{p_{n+1} - p_n}{\log n} = 0.$$

Zhang (2013) showed that $p_{n+1} - p_n = \mathcal{O}(1)$ infinitely often, in particular that $p_{n+1} - p_n \leq 70,000,000$ infinitely often. The Polymath project aimed to improve this, and Maynard reduced the constant to ≤ 600, and further progress by the polymath project reduced it to 246.

5.1.2 Large gaps

We have that

$$n! + 2, \ldots, n! + n$$

are all composite, so at scale $\approx n^n$, the gap is $\approx n$, so that $p_{N+1} - p_N \gg \frac{\log N}{\log \log N}$. 1931 Westzynthius showed

$$\limsup_{n \to \infty} \frac{p_{n+1} - p_n}{\log n} = \infty.$$

and Rankin (1938) showed

$$p_{n+1} - p_n \gg \log n \left(\frac{\log \log n \log \log \log \log n}{(\log \log \log n)^2} \right)$$

infinitely often. Erdős offered a \$10,000 prize to improve this, which was claimed in 2014 by Ford-Green-Konyagin-Maynard-Tao to

$$p_{n+1} - p_n \gg \log n \left(\frac{\log \log n \log \log \log \log n}{\log \log \log n} \right)$$

infinitely often. The conjecture is $\gg (\log n)^2$, but current bounds are far off.

The best upper bound is $\forall n \; p_{n+1} - p_n \ll n^{0.525}$, and assuming GRH $\ll n^{\frac{1}{2} + o(1)}$.

5.2 Digits of primes

Manduit-Rivat looked at the sum of binary digits of primes, and showed it was even $\frac{1}{2}$ the time, and odd $\frac{1}{2}$ the time.

Maynard (2016) shows there are infinitely many primes without, say, a 1 in base 10, and the method works for larger bases too, but not for smaller bases. For base 2? It is open to ask if there are infinitely many primes without any 0s (these are primes of the form $2^n - 1$)

5.3   Arithmetic progressions

Vinogradov, Estermann in the 1930s showed there are infinitely many arithmetic progressions of primes of length 3.

Theorem (Green-Tao, 2004)**.** For any k, there are infinitely many kAPs of primes.

These methods were inspired by additive combinatorics,

Theorem (Szemerédi, 1975)**.** If $\liminf_{N \to \infty} \frac{|A \cap \{1,\ldots,N\}|}{N} > 0$, then A has infinitely many 4APs.

5.4   Sieve theory success

Theorem (Chen 1973)**.** There are infinitely many primes p such that $p + 2$ is prime or the product of 2 primes.

While sieve theory can show strong results, it typically fails to distinguish between numbers with one prime factor and two prime factors, so improving this using sieve methods seems unlikely. Similarly,

Theorem (Iwaniec 1978)**.** There are infinitely many n such that $n^2 + 1$ is prime or the product of two primes.

5.5   Number theory without zeta zeroes

Granville-Soundararajan attempted analytic number theory without focus on the zeta function, in contrast to the methods of the last 150 years. Instead of using Perron's formula, they use a theorem of Halasz (1960s): 'if $\sum_{n \le x} a_n$ behaves non-randomly' then a_n behaves like $\chi(n)$ or n^{it}. This is called pretentious number theory, and studies functions 'pretending' to be $\chi(n)$.

5.6   Circle method

1920s: Hardy-Littlewood This involves the study of additive number theory, such as:

1. Goldbach: every even number is the sum of two primes

2. the study partition function $p(n)$

3. Waring's problem: We know all integers are the sum of 4 squares. Conjecture that all large integers are the sum of 4 cubes, but the best bound is 7. Let $G(k) =$ minimum s such that every large n is the sum of $x_1^k + \cdots + x_s^k = n$. We know $G(2) = 4$, and that $4 \le G(3) \le 7$, and $G(4) = 16$.

Theorem (Wooley, 1990)**.**

$$G(k) \le (1 + o(1))k \log k.$$

The conjecture is that $G(k) \ll k$.

The method uses generating functions, writing the answer as a contour integral around a circle, and using analysis.

Index of Notation

$1(n)$ constant 1 function, 4

$\mathcal{O}$ Big $\mathcal{O}$ notation; Landau notation, 4

$\Gamma(s)$ Gamma function, 46

$\Lambda(n)$ von Mangoldt function, 5

$\lambda(n)$ Liouville function, 4

$\Lambda_2(n)$ Selberg's function, 14

$\ll$ Vinogradov notation, 4

$\mu(n)$ Möbius function, 4

$\pi(x)$ prime-counting function, 3

$\psi(x)$ summatory von Mangoldt function, 9

$\sim$ asymptotic equality, 4

$\star$ convolution, 4

τ divisor function, 4

$\varphi(x)$ Euler's totient function, 3

o Little o notation, 4

$S(A, P; z)$ sifting function, 18

Index