ANALYTIC NUMBER THEORY

ABSTRACT. Analytic Number Theory is a great course that covers many branches of math. I here compile the class's content and make some supplement. It's for reviewing as well as for later study.

CONTENTS

1. Introduction

We study primes in this class——everything about it as much as possible. Starting from the Prime Number Theorem. But let's do some basics first.

1.1. About primes.

Def 1.1. *Let $p \in \mathbb{N}$, $p \geq 2$ be called a **prime** if $d|p \Rightarrow d = 1$ or $d = p$. Let $\mathscr{P}$ denote the set of all prime numbers.*

Theorem 1.1. *(Fundamental Theorem of Arithmetic):*
If $a \in \mathbb{N}$, $a \geq 2$, then $\exists!\{p_1, p_2, \cdots, p_n\} \subset \mathcal{P}$ such that $a = \sum_{i=1}^{n} p_i^{\alpha_i}$ up to rearrangement.

1.2. Asymptotic Estimates.

We start with a few asymptotic symbols:

- $\{x\} = x - \lfloor x \rfloor$
- $f(x) = O(g(x))$ if $|f(x)| \leq c \cdot g(x)$ for some constant $c = c(f, g, I)$, where $x \in I$
- $f(x) \ll g(x)$ iff $f(x) = O(g(x))$ (Vinogradov's notation)
- $f(x) \asymp g(x)$ iff $f(x) = O(g(x))$ and $g(x) = O(f(x))$
- $f(x) \underset{x \to x_0}{\sim} g(x)$ iff $\lim_{x \to x_0} \frac{f(x)}{g(x)} = 1$
- $f(x) = o_{x \to x_0}(g(x))$ iff $\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0$

Let's look at one of the most useful formula in all mathematics, Integral by Parts:

$$\int u dv = uv - \int v du$$

Example 1.1. *Let $Li(x) = \int_2^x \frac{1}{\log y} dy$, then $Li(x) \sim \frac{x}{\log x}$ as $x \to \infty$*

Proof.

$$\begin{aligned}
Li(x) &= \int_2^x \frac{1}{\log y} dy = \frac{y}{\log y}\Big|_2^x - \int_2^x y d\frac{1}{\log y} \\
&= \frac{x}{\log x} - \frac{2}{\log 2} + \int_2^x \frac{1}{\log^2 y} dy \\
&= \frac{x}{\log x} - \frac{2}{\log 2} + \cdots + (n-1)! + \frac{x}{\log^n x} - (n-1)!\frac{2}{\log^n 2} + n! \int_2^x \frac{1}{\log^{n+1} y} dy \\
&= \frac{x}{\log x} + \frac{x}{\log^2 x} + \cdots + \frac{x}{\log^n x} + O_n(1) + n! \int_2^x \frac{1}{\log^{n+1} y} dy
\end{aligned}$$

where by L'Hopital's rule we have

$$\lim_{x\to\infty}\frac{\int_2^x \frac{1}{log^{n+1}y}dy}{\frac{x}{log^{n+1}x}} = \lim_{x\to\infty}\frac{\frac{1}{log^{n+1}x}}{\frac{log^{n+1}x - x\frac{1}{x}(n+1)log^n x}{(log^{n+1}x)^2}}$$

$$= \lim_{x\to\infty}\frac{log^{n+1}x-(n+1)log^n x}{log^{n+1}x} = 1 - \lim_{x\to\infty}\frac{(n+1)}{logx} = 1$$

$$\Rightarrow \int_2^x \frac{1}{log^{n+1}y}dy \sim \frac{x}{log^{n+1}x} \text{ as } x\to\infty$$

So for any n, we have :

$$Li(x) = \frac{x}{logx} + \frac{x}{log^2x} + \cdots + \frac{x}{log^n x} + O\left(\frac{x}{log^{n+1}x}\right) + O_n(1)$$

In particular, we get $Li(x) \sim \frac{x}{logx}$ as $x\to\infty$

□

Theorem 1.2. *(Euler Maclaurin's Formula) Suppose $f \in C^1[y,z]$, then*

$$\sum_{y<n\le z} f(n) = \int_y^z f(t)dt - \{t\}f(t)\Big|_{t=y}^z + \int_y^z \{t\}f'(t)dt$$

In particular, if $f \in C^1[1,\infty)$, then

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

Proof. To begin with, we note that the sum can be written as an integral:

$$\sum_{y<n\le z} f(n) = \int_y^z f(t)\Delta dt$$

where $\Delta = \sum_{n\in\mathbb{Z}} \delta_n$ is the sum of Dirac masses at each integer value. Now we want to write Δ as the derivative of another function, but note that it's integral is nothing but the step function $\lfloor t\rfloor$(left-continuous) since there's a jump of value 1 at each integer. Therefore we have:

$$\sum_{y<n\le z} f(n) = \int_y^z f(t)\Delta dt = \int_y^z f(t)\lfloor t\rfloor dt$$

$$= \lfloor t\rfloor f(t)\Big|_y^z - \int_y^z f'(t)\lfloor t\rfloor dt + \int_y^z tf'(t)dt - \int_y^z tf'(t)dt$$

$$= \lfloor t\rfloor f(t)\Big|_y^z - \int_y^z f'(t)\{t\}dt - tf(t)\Big|_y^z + \int_y^z f(t)dt$$

$$= \int_y^z f(t)dt - \{t\}f(t)\Big|_{t=y}^z + \int_y^z \{t\}f'(t)dt$$

□

Theorem 1.3. *(Sterling's Formula) For $n \in \mathbb{N}$, we have*

$$n! = \left(\frac{n}{e}\right)^n \sqrt{2\pi n}\left(1 + O\left(\frac{1}{n}\right)\right)$$

Proof. We take the log and use the Euler Maclaurin formula.

$$\log n! = \sum_{k=2}^{n} \log k = \sum_{1<k\leq n} \log k = \int_1^n \log x dx + \int_1^n \frac{\{x\}}{x} dx$$

by letting $f = \log x$ then use the Euler Maclaurin formula. The problem now is to deal with the second integral. The trick is transforming $\{x\}$ into $\{x\} - \frac{1}{2}$, thus making the summation of it uniformly bounded, since $\{x\} - \frac{1}{2}$ has integral 0 in its period. More rigorously,

$$\int_1^n \frac{\{x\}}{x} dx = \int_1^n \frac{\{x\} - \frac{1}{2}}{x} dx + \int_1^n \frac{1}{2t} dt = \int_1^n \frac{1}{x} dF(x) + \frac{\log n}{2}$$

where $F(x) = \int_0^x \{t\} - \frac{1}{2} dt \in [0, \frac{1}{8}]$ and F is 0 at all integers, so

$$\int_1^n \frac{\{x\}}{x} dx = \frac{F(t)}{t}\Big|_1^n + \int_1^n \frac{F(t)}{t^2} dt + \frac{\log n}{2} = \int_1^n \frac{F(t)}{t^2} dt + \frac{\log n}{2}$$

where the integral converges and is of order $O\left(\frac{1}{n}\right)$.

So putting back all we have we get:

$$\log n! = n \log n - n + 1 + O\left(\frac{1}{n}\right) + \frac{\log n}{2}$$

□

The Saddle point method.

The simple form we use here is also called the **Laplace Method**. The key point is that if a function attains a unique maximum, then we can expect that most masses is around the maximum point.

Let's say we have a integral that we want to estimate. In most cases it is possible to write the integral in the form $\int_a^b e^f$ for some new function f. Then, say that c is the unique maximum of f, we would get $f'(c) = 0$ and $f''(c)$ is negative-definite by the second-derivative method.

If we take f to be a single variable function we can say that since

$$f(x) \approx f(c) + \frac{f''(c)}{2}(x - c)^2$$

since $f''(c) < 0$,we sort of expect that

$$\int_a^b e^{f(x)}dx \approx \int_{x\approx c} e^{f(c)-\frac{|f''(c)|}{2}(x-c)^2}dx = e^{f(c)}\int_{x\approx c} e^{-\frac{|f''(c)|}{2}(x-c)^2}dx$$

But then we know that the Gaussian decays extremely fast as $(x-c)$ grows, so we further can expect

$$e^{f(c)}\int_{x\approx c} e^{-\frac{|f''(c)|}{2}(x-c)^2}dx \approx e^{f(c)}\int_{\mathbb{R}} e^{-\frac{|f''(c)|}{2}(x-c)^2}dx = e^{f(c)}\sqrt{\frac{2\pi}{f''(c)}}$$

Where the last equality is because of Gaussian.

Def 1.2. *(Euler's Gamma Function)* $\Gamma(s) = \int_0^\infty e^{-x}x^{s-1}dx \quad$ *for* $Re(s) > 0$

Remark 1.1. *The Gamma function extends the factorial function. This this because* $\Gamma(1) = 1$ *and* $\Gamma(s+1) = s\Gamma(s)$ *with integral by parts.*

To justify the above result

We apply the equation above n times and get the formula :$\Gamma(s) = \dfrac{\Gamma(s+n)}{s(s+1)\cdots(s+n-1)}$

We use this formula to extend $\Gamma(s)$ to $Re(s) > -n$, thereby the whole complex plane. This formula shows explicitly that all poles of Γ are in $\{0, -1, -2, \cdots\}$. Since all poles are of order 1 we compute the residue simply as

$$res_{s=-n}\Gamma(s) = \lim_{s\to -n}(s+n)\frac{\Gamma(s+n+1)}{s(s+1)\cdots(s+n)} = \frac{(-1)^n}{n!}$$

Example 1.2. *We use the saddle point method to give an estimate of* Γ*, which will give us the same leading term as Sterling's formula, but a slightly larger error*

1.3. **Basic Results from Complex Analysis.**

Theorem 1.4. *(Cauchy's integral theorem)*

Theorem 1.5. *(Residue Theorem)*

1.4. **Exercises.**

Exercise 1.1. *(1.1)*

Exercise 1.2. *(1.7)*

Exercise 1.3. *(1.11)*

2. Combinatorial Ways of Counting Primes

In this chapter we see a few combinatoriral ways of getting the bound of $\pi(x)$. In particular, Chebychev's estimate already gives us the exact order.

Def 2.1. *(prime counting function)* $\pi(x) = \#\{p \in \mathscr{P}, p \leq x\}$

Def 2.2. $\omega(m) = \#\{p \in \mathscr{P}; p|m\}$; $\Omega(m) = \#\{p \in \mathscr{P}; p|m\}$;

Def 2.3. *(Euler function)* $\varphi(n) = \#\{m \leq n : (m, n) = 1\}$

2.1. **Eratosthenes & Legendre.**

Theorem 2.1. $\#\{n \leq x : (n, m) = 1\} = x \prod_{p|m} \left(1 - \frac{1}{p}\right) + O\left(2^{\omega(m)}\right)$

Before we go into the proof of the theorem, let's see an application and some failed attempts of the proof of the Prime Number Theorem.

Application 2.1. $\pi(x) \leq \frac{x}{loglogx} + o(x)$

It's easy to see that $n \in (\sqrt{x}, x]$ is a prime iff it has no prime factor less than $\sqrt{x}$, hence $\pi(x) = \#\{n \leq x : (n, P(\sqrt{x}))\} + O(\sqrt{x})$ where $P(Y) = \prod_{p \leq Y, p \in \mathscr{P}} p$

Lemma 2.2. $\prod_{p \leq y}(1 - \frac{1}{p}) \leq \frac{1}{\log y}$

Proof. $\prod_{p \leq y} \frac{1}{1 - \frac{1}{p}} \geq \prod_{p \leq y}(\sum_{k \geq 0} \frac{1}{p^k}) \geq \sum_{n : y-smooth} \frac{1}{n} \geq \sum_{n \leq y} \int_n^{n+1} \frac{du}{u} = \log(y + 1) \geq \log(y)$

Taking the reciprocal we get $\prod_{p \leq y}(1 - \frac{1}{p}) \leq \frac{1}{\log y}$

□

Letting $m = P(\sqrt{x})$, we use Thm 2.1 and Lemma 2.2 to get

$$\begin{aligned}
\pi(x) &\leq x \prod_{p|P(\sqrt{x})} \left(1 - \frac{1}{p}\right) + O\left(2^{\omega(P(\sqrt{x}))}\right) + O\left(\sqrt{x}\right) \\
&\leq \frac{x}{\log \sqrt{x}} + O\left(2^{\omega(P(\sqrt{x}))}\right) + O\left(\sqrt{x}\right) \\
&= 2\frac{x}{\log x} + O\left(2^{\omega(P(\sqrt{x}))}\right) + O\left(\sqrt{x}\right)
\end{aligned}$$

At first glance we have proved Prime Number Theorem! But that is not true since $O\left(2^{\omega(P(\sqrt{x}))}\right)$ is in fact pretty bad since $P(\sqrt{x}) \geq \pi(x)$ and even $2^{\pi(x)} \asymp 2^{\frac{x}{logx}} \gg \frac{x}{logx}$, which means that the error is a lot greater than the leading term.

Legendre has a better idea on this issue. He tries the same method with a much smaller m than $P(\sqrt{x})$. For now let's assume this smaller m is y. Since for any y and any prime p, $(p, y) = 1$ we have

$$\pi(x) \leq \#\{n \leq x : (n, P(y)) = 1\} + O(y) = x \prod_{p \leq y} \left(1 - \frac{1}{p}\right) + O(2^{\omega(P(y))}) + O(y)$$

$$\leq \frac{x}{logy} + O(2^{\pi(y)}) + O(y)$$

Since y is on the exponential, we take $y = \log x$, so the equation

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

$$\leq \frac{x}{\log\log x}$$

Hence we conclude the application. Now we give a failed attempt of the proof of **Theorem 2.1**. Even though it fails, it's also interesting to know the method.

Theorem 2.3. *(Bezout's Lemma) If $(a, b) = 1$, then $\exists k, l \in \mathbb{Z}$ such that $ka + lb = 1$*

Theorem 2.4. *(Chinese Remainder Theorem) If $(n_1, n_2) = 1$, for any $a_1, a_2 \in \mathbb{Z}$, $\exists! m \in \{1, 2, \cdots, n_1 n_2\}$ such that* $\begin{cases} m \equiv a_1 (mod n_1) \\ m \equiv a_2 (mod n_2) \end{cases}$

Def 2.4. *An **arithmetic function** is a function $f : \mathbb{N} \to \mathbb{C}$. Let $\mathcal{A}$ be the set of all arithmetic functions.*

Def 2.5. *An arithmetic function $f : \mathbb{N} \to \mathbb{C}$ is **multiplicative** if $\forall (n_1, n_2) = 1$, $f(n_1 n_2) = f(n_1) f(n_2)$. It is **completely multiplicative** if $f(n_1 n_2) = f(n_1) f(n_2)$ for any n_1, n_2.*

Lemma 2.5. $\varphi(n) = n \prod_{p|n} (1 - \frac{1}{p})$

Proof. We first prove that φ is multiplicative.

If $(n_1, n_2) = 1$, by Chinese Remainder Theorem $(\mathbb{Z}/_{n_1 n_2 \mathbb{Z}})^* \simeq (\mathbb{Z}/_{n_1 \mathbb{Z}})^* \times (\mathbb{Z}/_{n_2 \mathbb{Z}})^*$, we count the order of the group and we get $\varphi(n_1 n_2) = \varphi(n_1)\varphi(n_2)$, so φ is multiplicative.

Then, since $\varphi(p^k) = p^k - p^{(k-1)} = p^k(1 - \frac{1}{p})$, we have

$$\varphi(n) = \varphi(p_1^{\alpha_1}) \cdots \varphi(p_k^{\alpha_k}) = p_1 \cdots p_k (1 - \frac{1}{p_1}) \cdots (1 - \frac{1}{p}) = n \prod_{p|n} (1 - \frac{1}{p})$$

□

In order to find $\#\{n \leq x : (n,m) = 1\}$, let's just assume $x \in (Nm, (N+1)m]$ for some $N \in \mathbb{Z}$. But we notice that there's exactly $\varphi(m)$ integers coprime to m on $(Nm, (N+1)m]$ for every N, hence $N\varphi(m) \leq \#\{n \leq x : (n,m) = 1\} \leq (N+1)\varphi(m)$, plugging in Lemma 2.4 we have

$$x\prod_{p|m}(1-\frac{1}{p}) \leq \#\{n \leq x : (n,m) = 1\} \leq x\prod_{p|m}(1-\frac{1}{p}) + O(\varphi(m))$$

We could conclude the theorem if $\varphi(m) = O(2^{\omega(m)})$. Unfortunately that is not true since for large $m \in \mathscr{P}$, $\varphi(m) = m - 1 \neq O(2^{\omega(m)}) = O(2)$

Proof. (of Theorem 2.1)
We want to show $\#\{n \leq x : (n,m) = 1\} = x\prod_{p|m}(1-\frac{1}{p}) + O(2^{\omega(m)})$

The idea here is to use the Inclusion-Exclusion principle:

$$\begin{aligned}
\#\{n \leq x : (n,m) = 1\} &= \#\bigcap_{p|m}\{n \leq x : p \nmid n\} \\
&= \#\{m \leq x\} - \#\sum_{p|m}\#\{n \leq x : p|n\} + \sum_{\substack{p_1p_2|m \\ p_1 \leq p_2}}\#\{n \leq x : p_1p_2|n\} - \cdots \\
&= \lfloor x \rfloor - \sum_{p|m}\lfloor \frac{x}{p} \rfloor + \sum_{\substack{p_1p_2|m \\ p_1 \leq p_2}}\lfloor \frac{x}{p_1p_2} \rfloor - \cdots + (-1)^{\omega(m)}\lfloor \frac{x}{p_1 \cdots p_{\omega(m)}} \rfloor \\
&= x - \sum_{p|m}\frac{x}{p} + \sum_{\substack{p_1p_2|m \\ p_1 \leq p_2}}\frac{x}{p_1p_2} - \cdots + (-1)^{\omega(m)}\frac{x}{p_1 \cdots p_{\omega(m)}} + O((1+1)^{\omega(m)}) \\
&= x\prod_{p|m}(1-\frac{1}{p}) + O(2^{\omega(m)})
\end{aligned}$$

By rearrangement.

□

2.2. **Chebychev's Estimate.** Chebychev has a genius idea which in fact gives us the exact order of Prime Number Theorem. However, this proof is too tricky that no one has came up with any improvement, or so professor Bourgade says.

The trick to Chebychev's proof is this: if $p \in \mathscr{P}\bigcap[n, 2n]$, then $p|\binom{2n}{n} = \frac{(2n)!}{n!n!}$. Hence $\prod_{n<p\leq 2n} p \leq \frac{(2n)!}{n!n!}$, which implies $(n)^{\#\{p\in(n,2n]\}} \leq \frac{(2n)!}{n!n!}$ since each p is greater than n.

Before the formal proof, we need a tool:

Def 2.6. $\nu_p(n)$ = *the largest power of p dividing n.*

Lemma 2.6. $\nu_p(n!) = \prod_{k\geq 1} \left\lfloor \frac{n}{p^k} \right\rfloor$

Proof.

$$\begin{aligned}
\nu_p(n!) &= 1 \cdot \#\{m < n, p|m, p^2 \nmid m\} + 2 \cdot \#\{m < n, p^2|m, p^3 \nmid m\} + \cdots \\
&= \sum_{j\leq 1} j \left(\left\lfloor \frac{n}{p^j} \right\rfloor - \left\lfloor \frac{n}{p^{j+1}} \right\rfloor \right) \\
&= \sum_{j\leq 1} j \left\lfloor \frac{n}{p^j} \right\rfloor - \sum_{k\leq 1} k \left\lfloor \frac{n}{p^{k+1}} \right\rfloor \\
&= \sum_{j\leq 1} j \left\lfloor \frac{n}{p^j} \right\rfloor - \sum_{k\leq 1} \left((k-1) \left\lfloor \frac{n}{p^k} \right\rfloor \right) \\
&= \sum_{j\leq 1} \left\lfloor \frac{n}{p^j} \right\rfloor
\end{aligned}$$

□

Theorem 2.7. *(Chebychev's Estimate)* $\pi(x) \asymp \frac{x}{logx}$ *for* $x \geq 2$

Proof. We show first the upper bound for π and then the lower bound.

Upper Bound: We know from above discussion that

$$\prod_{n<p\leq 2n} p \leq \binom{2n}{n} \leq \sum_{j=1}^{2n} \binom{2n}{j} = (1+1)^{2n} = 2^{2n}$$

Which implies $n^{\#\{p\in(n,2n]\}} \leq 2^{2n} \Rightarrow \pi(2n) - \pi(n) = \#\{p \in (n, 2n]\} \leq \frac{2n \log 2}{\log n}$

Yet if $2^n \leq x \leq 2^{n+1}$, we have

$$\pi(x) \leq \pi(2^{n+1}) = \sum_{k=0}^{n} (\pi(2^{k+1}) - \pi(2^k)) \leq \sum_{k=1}^{n} \frac{2^{k+1} \log 2}{\log 2^k}$$

$$\leq \sum_{k=1}^{n} 2\frac{2^k}{k} \leq \tilde{c}\frac{2^n}{n} \leq c' \frac{x}{logx}$$

Where $c \sum_{k=1}^{n} \frac{2^k}{k} \leq \tilde{c}\frac{2^n}{n}$ is because

$$\sum_{k=1}^{n} \frac{2^k}{k} \leq \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \frac{2^k}{k} + \sum_{k=\lfloor \frac{n}{2} \rfloor + 1}^{n} \frac{2^k}{k} \leq 2^{\lfloor \frac{n}{2} \rfloor + 1} + \lfloor \frac{2}{n} \rfloor 2^{n+1}$$

And $c2^{\lfloor \frac{n}{2} \rfloor} \leq \frac{2^n}{n}$ when $0 < c \leq \frac{2^{3/2}}{3}$

Lower Bound: By Lemma 2.6

$$\binom{2n}{n} = \frac{(2n)!}{n!n!} = \prod_{p\in\mathscr{P}} p^{\nu_p((2n)!)-2\nu_p(n!)} = \prod_{p\in\mathscr{P}} p^{\sum_{k\geq 1}\left(\left\lfloor\frac{2n}{p^k}\right\rfloor - 2\left\lfloor\frac{n}{p^k}\right\rfloor\right)}$$

Since $\lfloor 2x\rfloor - 2\lfloor x\rfloor \in \{0,1\}$ and when $\frac{2n}{p^k} \leq 1 \iff k \geq \frac{\log 2n}{\log p}$, the summand is 0, the above equation

$$\leq \prod_{p\in\mathscr{P}, p\leq 2n} p^{\frac{\log 2n}{\log p}} = \prod_{p\in\mathscr{P}, p\leq 2n} 2n = (2n)^{\pi(2n)}$$

We offer two ways to go from $\binom{2n}{n} \leq (2n)^{\pi(2n)}$ to $\pi(x) \geq c\cdot\frac{x}{logx}$

- 1st way:
  Since $\binom{2n}{n} = \max_{0\leq k\leq 2n}\binom{2n}{k}$, so $\binom{2n}{n}$ is larger than the average of $\binom{2n}{k}$, which is $\frac{1}{2n+1}\sum_{0\leq k\leq 2n}\binom{2n}{k} \geq \frac{(1+1)^{2n}}{2n+1}$, we get

  $$(2n)^{\pi(2n)} \geq \frac{2^{2n}}{2n+1} \Rightarrow \pi(2n) \geq c\cdot\frac{2n}{log2n} \Rightarrow \pi(x) \geq c\cdot\frac{x}{logx}$$

- 2nd way:
  By Sterling's Formula(Theorem 1.3)

  $$\frac{(2n)!}{n!n!} \sim \frac{\sqrt{2\pi 2n}(\frac{2n}{e})^{2n}}{(\sqrt{2\pi n}(\frac{2n}{e})^n)^2} = \frac{2^{2n}}{\sqrt{\pi n}}$$

  We conclude the same as in the 1st way.

Hence both the upper and lower bound is proved, we conclude the theorem.

2.3. **Exercises.**

Exercise 2.1. *(2.3)*

Exercise 2.2. *(2.5)*

Exercise 2.3. *(2.11)*

□

3. Dirichlet Convolution

Def 3.1. *For any $f, g \in \mathcal{A}$, let the* ***Dirichlet convolution*** *of them* $(f * g)(n) = \sum_{d|n} f(d)g(\frac{n}{d})$, *or equivalently* $(f * g)(n) = \sum_{ab=n} f(a)g(b)$

3.1. basic properties.

Note that we are not using the normal convolution $(f * g)(x) = \int f(y)g(x-y)dy$ on integrable functions in this class.

A few properties for the Dirichlet convolution is listed here:

- Commutative: $f * g = g * f$
- Distributive: $(f * g) * h = f * (g * h)$
- $(\mathcal{A}, +, *)$ forms a commutative unitary ring with multiplicative identity δ, where $\delta(n) = \begin{cases} 1 & n = 1 \\ 0 & \text{otherwise} \end{cases}$
- f has an inverse in $(\mathcal{A}, +, *)$ iff $f(1) \neq 0$. in particular, multiplicative functions have inverses since $f(1) = 1$

We introduce now an extremely important function that will be seen everywhere in this class:

Def 3.2. *The* ***Mobius function*** $\mu(n) = \begin{cases} (-1)^k & n \text{ is square-free and has } k \text{ divisors} \\ 0 & n \text{ not square-free} \end{cases}$, *where n is square-free means that $p^2 \nmid n$ for any prime. We also let $\mu(1) = 1$*

Let's list some interesting arithmetic functions:

- $\mu(n)$ as defined above in Def 3.2
- $\delta(n) = \begin{cases} 1 & n = 1 \\ 0 & \text{otherwise} \end{cases}$
- $\mathbb{1}(n) = 1$
- $id(n) = n$
- $\varphi(n) = \#\{m \leq n : (m, n) = 1\}$ (Euler's function)
- $\tau(n) = \#\{d : d|n\}$
- $\sigma(n) = \sum_{d|n} d$
- $\sigma_k(n) = \sum_{d|n} d^k$
- $\Lambda(n) = \begin{cases} \log p & n = p^k \text{ for some } k > 1 \\ 0 & \text{otherwise} \end{cases}$. (Von-Mangoldt function)

We now show how the first three functions are connected:

Lemma 3.1. *μ is the inverse of $\mathbb{1}$, which means that* $\sum_{d|n} \mu(d)\mathbb{1}(\frac{n}{d}) = \sum_{d|n} \mu(d) = \delta(n)$

Proof. We know that for any $n \geq 2$ can be written uniquely as $n = \prod_{i}^{k} p_i^{\alpha_i}$, and $\sum_{d|n} \mu(d) = \sum_{d|p_1 \cdots p_k} \mu(d)$ since otherwise d is not square-free and $\mu(d) = 0$, further,

$$\sum_{d|p_1 \cdots p_k} \mu(d) = \sum_{I \subseteq \{1,2 \cdots k\}} (-1)^{|I|} = \sum_{m=0}^{k} \binom{k}{m} (-1)^m = (1-1)^k = 0 = \delta(n)$$

If $n = 1$, then $\sum_{d|1} \mu(d) = 1 = \delta(1)$

So we conclude that μ is the inverse of $\mathbb{1}$ □

3.2. **Dirichlet Hyperbolic Method.**

We are interested in the partial sum of arithmetic functions. It it often useful to write out the Dirichlet convolution of f to calculate the average. i.e, if $f = h * g$, then

$$\sum_{n \leq x} f(n) = \sum_{n \leq x} \sum_{ab=n} h(a)g(b) = \sum_{ab \leq x} h(a)g(b)$$

Example 3.1. *We use Dirichlet Hyperbolic Method to count τ*

Notice $\tau(n) = \#\{d : d|n\} = \sum_{d|n} 1 = \mathbb{1} * \mathbb{1}(n)$, so $\sum_{n \leq x} \tau(n) = \sum_{ab \leq x} 1$ which is the number of lattice point in the first below the graph of $y = \frac{1}{x}$. To count this, we note that there are around the the size of area number of lattice points, and with error of the size of length of the perimeter. This conclusion holds for convex sets for sure since any curve completely "outside" the convex set has a larger perimeter. Another way to see it is because if we take finer grids then we get a better approximation of the area (Jordan Measure). Anyway, prof Bourgade did not give a full-round proof of this and neither will I.

We have

$$\sum_{n \leq x} \tau(n) = \sum_{ab \leq x} 1 = \iint_{ab \leq x} da\, db + O(x) \underset{x \to x_0}{\sim} x \log x + O(x)$$

So

$$\frac{1}{x} \sum_{n \leq x} \tau(n) \underset{x \to x_0}{\sim} \log x$$

Theorem 3.2. $\sum_{n \leq x} \varphi(n) \underset{x \to x_0}{\sim} \frac{3}{\pi^2} x^2$

Proof. We proceed step by step.

Step 1: write φ as a convolution.

We'll show that $\varphi = \mu * id$ in this step. We'll prove that this identity holds for powers of primes. Then, we'll show that the convolution is multiplicative. Since the two sides are equal at power of primes and they are both multiplicative, they'll be equal.

For $n = p^k$, $\varphi(n) = p^k - p^{k-1} = \sum_{d|n} d\,\mu(\frac{n}{d})$ because $\mu \neq 0$ only when d is p^k or p^{k-1}.

We claim that $\mu * id$ is multiplicative since μ and id are. Since μ and id are multiplicative are obvious, we only need to prove that if f and g are multiplicative, then so is $f * g$. This is because for $(a, b) = 1$,

$$\begin{aligned}(f * g)(ab) &= \sum_{d|ab} f(d)g(\frac{ab}{d}) = \sum_{d_1|a,\, d_2|b} f(d_1d_2)g(\frac{ab}{d_1d_2}) = \sum_{d_1|a,\, d_2|b} f(d_1)f(d_2)g(\frac{a}{d_1})g(\frac{b}{d_2}) \\ &= (\sum_{d_1|a} f(d_1)g(\frac{a}{d_1}))(\sum_{d_2|b} f(d_2)g(\frac{b}{d_2})) = ((f * g)(a)) \cdot ((f * g)(b))\end{aligned}$$

So we've shown $\varphi = \mu * id$. This might seem a little unintuitive for now, but it will look better after we've get to Dirichlet series.

Step 2: We use the Dirichlet Hyperbolic Method.

$$\sum_{n\le x} \varphi(n) = \sum_{n\le x} (\mu * id)(n) = \sum_{ab\le x} \mu(a)b$$

The trick here is that we fix a and sum over b(or the other way around).

$$\begin{aligned}\sum_{n\le x} \varphi(n) &= \sum_{a\le x} \mu(a) \sum_{b\le \frac{x}{a}} b = \sum_{a\le x} \mu(a)\Big(\big(\frac{(\frac{x}{a})^2}{2}\big) + O(\frac{x}{a})\Big) \\ &= \sum_{a\le x} \mu(a)\frac{x^2}{2a^2} + O(\sum_{a\le x} \frac{|\mu(a)|x}{a}) = x^2 \sum_{a\le x} \frac{\mu(a)}{2a^2} + O(x\log x)\end{aligned}$$

So if $\sum_{a=1}^{\infty} \frac{\mu(a)}{2a^2} = \frac{3}{\pi^2}$, we are done. So let's prove it.

Since $\mu * \mathbb{1} = \delta$, $\forall s \ge 2$,

$$\sum_{n\ge 1} \frac{\mu(n)}{n^s} \sum_{n\ge 1} \frac{1}{n^s} = \sum_{n\ge 1} \frac{\delta(n)}{n^s} = 1$$

Since the denominator of the n-th term in the sum is exactly $\sum_{ab=n} \mu(a) = \delta(n)$, the above equation holds. So

$$\sum_{n\geq 1}\frac{\mu(n)}{n^s}=\frac{1}{\sum_{n\geq 1}\frac{1}{n^s}}=\frac{6}{\pi^2}\Rightarrow\sum_{a=1}^{\infty}\frac{\mu(a)}{2a^2}=\frac{3}{\pi^2}$$

Hence we conclude the theorem. **Remark**: If we pick a,b uniformly at random in $[1,x]^2$, then the probability that they are coprime is

$$\frac{1}{x^2}\sum_{1\leq a,b\leq x}\mathbb{1}_{(a,b)=1}=\frac{1}{\frac{x(x-1)}{2}}\sum_{1\leq a<b\leq x}\mathbb{1}_{(a,b)=1]}$$

since we are counting in the second term the ordered pair (a,b). We do the same trick of fixing b and sum over a again (so really we sum on $\varphi(b)$) and get

$$\mathbb{P}=\frac{2}{x(x-1)}\sum_{1\leq b\leq x}\varphi(b)\sim\frac{6}{\pi^2}$$

Where the last step is by Theorem 3.2 □

Theorem 3.3. *(Merten's Estimates)*

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

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

(c) $\sum_{p\leq x}\left(1-\frac{1}{p}\right)=\frac{e^{-\gamma}}{\log x}(1+O(\frac{1}{\log x}))$ *where* $\gamma=\lim_{n\to\infty}(\sum_{k=1}^{n}\frac{1}{k}-\log n)=\int_1^{\infty}-\frac{1}{x}+\frac{1}{\lfloor x\rfloor}dx$ *the Euler constant*

Before we see the proof, let's make two remarks.

Remark 1: At the level of leading term, all three estimates are consequences of PNT by integral by parts. For instance,

$$\sum_{p\leq x}\frac{\log p}{p}=\int_1^x\frac{\log u}{u}d(\pi(u))\sim\int_1^x\frac{\log u}{u^2}\pi(u)du\underset{PNT}{\sim}\int_1^x\frac{du}{u}=\log x$$

Note that only PNT is not enough for the last asymptotic estimate since u is not always large (it goes from 1 to x, and only x is large). However, since PNT tells us that $\forall\delta>0$, $\exists N$ such that when $\gamma>N$,

$$\left|\pi(\gamma)-\frac{\gamma}{\log\gamma}\right|<\delta\frac{\gamma}{\log\gamma}$$

So for $x>2N$,

$$\left|\int_N^x\frac{\log\gamma}{\gamma^2}\pi(\gamma)d\gamma-\int_N^x\frac{d\gamma}{\gamma}\right|<\delta\int_N^x\frac{d\gamma}{\gamma}$$

But then $\sup_{\gamma \le N}(\pi(\gamma) - \frac{\gamma}{\log \gamma}) = c$ exists. So

$$\Big| \int_1^N \frac{\log \gamma}{\gamma^2} \pi(\gamma) d\gamma - \int_1^N \frac{d\gamma}{\gamma} \Big| \le c \int_1^N \frac{\log \gamma}{\gamma^2} d\gamma = c(-\frac{\log N - N - 1}{N}) = O(1)$$

So

$$\Big| \int_1^x \frac{\log \gamma}{\gamma^2} \pi(\gamma) d\gamma - \int_1^x \frac{d\gamma}{\gamma} \Big| \ll_\delta \delta \int_1^x \frac{d\gamma}{\gamma}$$

$$\Rightarrow \Big| \frac{\int_1^x \frac{\log \gamma}{\gamma^2} \pi(\gamma) d\gamma}{\int_1^x \frac{d\gamma}{\gamma}} - 1 \Big| \ll_\delta \delta$$

for any δ. So the asymptotic approximation is justified.

Remark 2: We cannot go from Merten's estimate to PNT since we cannot put the asymptotic $\ll$ back to $<$ due to the fact that the constant is not uniform in N.

Proof. We only do a detailed proof of (a) here. The proof is based on Dirichlet convolution. Recall that the Von-Mangoldt function is defined as $\Lambda(n) = \begin{cases} \log p & n = p^k \text{ for some } k > 1 \\ 0 & \text{otherwise} \end{cases}$, we claim that

$$\log = \mathbb{1} * \Lambda$$

Indeed, for $n = \prod_{p|n} p^{\alpha_p}$,

$$\log n = \sum_{p|n} \alpha_p \log p = \sum_{\substack{p^l | n \\ l \in \{1, 2 \cdot \alpha_p\}}} \Lambda(p^l) = \sum_{m|n} \Lambda(m) = \mathbb{1} * \Lambda(n)$$

Using the usual trick for Dirichlet Hyperbolic method,

$$\sum_{n \le x} \log n = \sum_{n \le x} \Lambda(a) \sum_{b \le \frac{x}{a}} 1 = x \sum_{a \le x} \frac{\Lambda(a)}{a} + O(\sum_{a \le x} \Lambda(a))$$

But $\sum_{a \le x} \Lambda(a) \le \pi(x) \log x = O(x)$ by Chebyvhev's estimate of π, so

$$\sum_{n \le x} \log n = x \sum_{a \le x} \frac{\Lambda(a)}{a} + O(x)$$

Now, on the one hand we have $\sum_{n \le x} \log n \sim x \log x$ by integration,

and on the other hand,

$$\sum_{a \le x} \frac{\Lambda(a)}{a} = \sum_{p \le x} \frac{\Lambda(p)}{p} + \sum_{\substack{p^l \le x \\ l \ge 2}} \frac{\Lambda(p^l)}{p^l} = \sum_{p \le x} \frac{\log p}{p} + \sum_{\substack{p^l \le x \\ l \ge 2}} \frac{\log p}{p^l}$$

For the term $\sum_{\substack{p^l \le x \\ l \ge 2}} \frac{\log p}{p^l}$, we see that if it is an error then we are done. So we first try to bound it like this:

$$\sum_{\substack{p^l \le x \\ l \ge 2}} \frac{\log p}{p^l} \le \log x \sum_{\substack{p^l \le x \\ l \ge 2}} \frac{1}{p^l} \le \log x \sum_{\substack{p \in \mathscr{P} \\ l \ge 2}} \frac{1}{p^l} \le \log x \sum_{\substack{n \ge 1 \\ l \ge 2}} \frac{1}{n^l} = \log n \sum_{n \ge 1} \frac{1}{n^2} \frac{1}{1 - \frac{1}{n}} = O(\log x)$$

which is too big for our purpose. But since we're only off by $\log x$, let's try not pulling it out in the start:

$$\sum_{\substack{p^l \le x \\ l \ge 2}} \frac{\log p}{p^l} \le \sum_{\substack{n \ge 1 \\ l \ge 2}} \frac{\log n}{n^l} = \sum_{n \ge 1} \frac{\log n}{n^2} \frac{1}{1 - \frac{1}{n}} = O(1)$$

Which is what we need.

In conclusion:

$$x \log x \sim \sum_{n \le x} \log n = x \sum_{a \le x} \frac{\Lambda(a)}{a} + O(x) = x \sum_{p \le x} \frac{\log p}{p} + xO(1) + O(x)$$

$$\Rightarrow \sum_{p \le x} \frac{\log p}{p} \sim \log x$$

For the remaining, we can get (b) and (c) from (a), but it is left to be added(not in class) □

3.3. **Exercises.**

Exercise 3.1. *(3.8)*

Exercise 3.2. *(3.9)*

4. Dirichlet Series

Def 4.1. *For an arithmetic function f, the **Dirichlet Series** of f is* $F(s) = \sum_{n\geq 1} \frac{f(n)}{n^s}$

4.1. Examples of Dirichlet Series, Riemann-Zeta Function.

Let's have some preliminary results or remarks:

- We often use capital letters to denote the Dirichlet series.
- We need $Re(s)$ to be large enough for F to be absolutely converge. In particular, when f is $\mathbb{1}$, we need $Re(s) > 1$
- $F(s)G(s) = \sum_{ab\geq 1} \frac{f(a)g(b)}{(ab)^s} = \sum_{n\geq 1} \frac{1}{n^s} \sum_{ab=n} f(a)g(b) = \sum_{n\geq 1} \frac{1}{n^s}(f * g)(n)$
- if g is multiplicative, i.e. $g(\prod p^{\alpha_p}) = \prod g(p^{\alpha_p})$, then

$$G(s) = \sum_{n\geq 1} \frac{g(p_1^{\alpha_1})g(p_2^{\alpha_2})\cdots g(p_r^{\alpha_r})}{(p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_r^{\alpha_r})^s} = \prod_{p\in\mathscr{P}}(1 + \frac{g(p)}{p^s} + \frac{g(p^2)}{p^{2s}} + \cdots)$$

Now let's see a few examples of Dirichlet series that shed light to a few Dirichlet convolutions.

Example 4.1. ζ *- Riemann-zeta function*

Here $f = \mathbb{1}$. So $\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s}$. By the last point above, we get the **Euler Product**:

$$\zeta(s) = \prod_{p\in\mathscr{P}}(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots) = \prod_{p\in\mathscr{P}} \frac{1}{1 - \frac{1}{p^s}}$$

Note that

$$\frac{1}{\zeta(s)} = \prod_{p\in\mathscr{P}}(1 - \frac{1}{p^s}) = \sum_{n\geq 1} \frac{\mu(n)}{n^s}$$

This really means that $\mu * \mathbb{1} = \delta$. We see from this that Dirichlet series is a way to figure "out of nowhere" Dirichlet convolutions.

Example 4.2. *Dirichlet series of* φ

On the one hand we have:

$$\begin{aligned} F(s) &= \sum_{n\geq 1} \frac{\varphi(n)}{n^s} = \prod_{p\in\mathscr{P}}(1 + \frac{\varphi(p)}{p^s} + \frac{\varphi(p^2)}{p^{2s}} + \cdots) = \prod_{p\in\mathscr{P}}(1 + \frac{p-1}{p^s} + \frac{p^2-p}{p^{2s}} + \cdots) \\ &= \prod_{p\in\mathscr{P}}(1 + \frac{p-1}{p}\frac{p}{p^s} + \frac{p-1}{p}\frac{p^2}{p^{2s}} + \cdots) = \prod_{p\in\mathscr{P}}[1 + \frac{p-1}{p}\sum_{k\geq 1}(\frac{1}{p^{s-1}})^k] \\ &= \prod_{p\in\mathscr{P}}(1 + \frac{p-1}{p}\frac{1}{p^{s-1}}\frac{1}{1-\frac{1}{p^{s-1}}}) = \prod_{p\in\mathscr{P}}(1 + \frac{p-1}{p}\frac{1}{p^s-p}) = \prod_{p\in\mathscr{P}} \frac{p^s-1}{p^s-p} \end{aligned}$$

Whereas on the other hand since $\zeta(s-1) = \sum_{n\geq 1} \frac{n}{n^s}$ is the Dirichlet series of id, we have:

$$\sum_{n\geq 1} \frac{id(n)}{n^s} \sum_{n\geq 1} \frac{\mu(n)}{n^s} = \zeta(s-1)\frac{1}{\zeta(s)} = \prod_{p\in\mathscr{P}} \left[\frac{1-\frac{1}{p^s}}{1-\frac{1}{p^{s-1}}}\right] = \prod_{p\in\mathscr{P}} \frac{p^s-1}{p^s-p}$$

Therefore $\sum_{n\geq 1} \frac{\varphi(n)}{n^s} = \sum_{n\geq 1} \frac{id(n)}{n^s} \sum_{n\geq 1} \frac{\mu(n)}{n^s} \Rightarrow \varphi = \mu * id$, which is step 1 in theorem 3.2.

Example 4.3. *Dirichlet series of* Λ*, the Von-Mangoldt function*

We deal with ζ first and we'll see how Λ comes into play very soon.

$$\log\zeta(s) = \log\prod_{p\in\mathscr{P}} \frac{1}{1-\frac{1}{p^s}} = -\sum_{p\in\mathscr{P}} (1-\frac{1}{p^s}) = -\sum_{p\in\mathscr{P}} \sum_{k\geq 1} \frac{1}{k}(\frac{1}{p^s})^k$$

Differentiate both side in s yields

$$\frac{\zeta'(s)}{\zeta(s)} = -\sum_{\substack{k\geq 1\\ p\in\mathscr{P}}} \frac{\log p}{p^{ks}} = -\sum_{n\geq 1} \frac{\Lambda(n)}{n^s}$$

Since $\zeta'(n) = -\sum_{n\geq 1} \frac{\log n}{n^s}$ and $\frac{1}{\zeta(n)} = -\sum_{n\geq 1} \frac{\mu(n)}{n^s}$

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

We get $\Lambda = \mu * \log$, or equivalently, $\log = \mathbb{1} * \Lambda$ since μ is the inverse of $\mathbb{1}$. This explains the convolution in the proof of Merten's Estimates.

4.2. **Analytic Properties of Dirichlet Series.**

Assume $f(n) = O(n^\theta)$ for some $\theta > 0$, i.e. f is of polynomial growth. Then f converges absolutely when $Re(s) > \theta + 1$ since $\sum_{n\geq 1} \left|\frac{f(n)}{n^s}\right| \leq c \sum_{n\geq 1} \frac{n^\theta}{Re(s)}$, which converges when $Re(s) > \theta + 1$. In fact, the Dirichlet series not only converges, it is even holomorphic on the half plane, which is why we don't care about conditional convergent of Dirichlet series at all.

Theorem 4.1. *Assume F converge absolutely at $s = s_0$, then it converges absolutely for any $Re(s) \geq Re(s_0)$ and it is holomorphic on the half plane.*

Proof. If $Re(s) \geq Re(s_0)$, $|\frac{f(n)}{n^s}| \leq |\frac{f(n)}{n^{s_0}}|$, so $F(s)$ converges absolutely as well.

Now let γ be any closed contour in the half plane $Re(s) \geq Re(s_0)$. If we can show that for any such γ, $\oint_\gamma F(z)dz = 0$, we know that F is holomorphic on the half plane. So let's prove it.

We know that $n^{-z} = e^{-z \log n}$ is holomorphic in $\mathbb{C}$, so the partial sums $F_k(z) = \sum_{n \leq k} \frac{f(n)}{n^z}$ is holomorphic. Which means $\oint_\gamma F_k(z) dz = 0$.

If we can show that $F_k \to F$ uniformly, then since uniform convergence keeps the integral, i.e. $|\int f - \int g| \leq \int |f - g|$, we are done.

The convergence is uniform since $\forall Re(z) \leq Re(s_0)$,

$$|F(z) - F_k(z)| = \Big| \sum_{n>k} \frac{f(n)}{n^z} \Big| \leq \sum_{n>k} \frac{|f(n)|}{n^{Re(z)}} \leq \sum_{n>k} \frac{|f(n)|}{n^{Re(s_0)}} \underset{k \to \infty}{\to} 0$$

Since f converges absolutely on the right half-plane. □

4.3. **Exercise.**

Exercise 4.1. *(4.8)*

Exercise 4.2. *(4.9)*

5. The Prime Number Theorem

Our goal is to prove the Prime Number Theorem:

Theorem 5.1. *(The Prime Number Theorem)* $\pi(x) \sim \dfrac{x}{\log x}$

PNT is sometimes written in other forms, such as $\pi(x) = Li(x) + O(xe^{-c\sqrt{\log x}})$ for some $c > 0$.

5.1. Zeta function and its analytic continuation.

Remember from last chapter $\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s} = \prod_{p\in\mathscr{P}} \frac{1}{1-\frac{1}{p^s}}$ for $Re(s) > 1$. Even though ζ does diverge at $s = 1$, we notice that it actually converges as $s \to 1 + it$ where $t \neq 0$. Therefore it's tempting to do an analytic continuation to it.

We first see how we can extend ζ to $Re(s) > 0$. The intuition here is to use integral by parts since $\sum \frac{1}{n^{s+1}}$ converges for the area. Indeed, we make it rigorous by the **Euler-Maclaurin Formula**(Theorem 1.2): for $f \in C^1$

$$\sum_{y<n\leq z} f(n) = \int_y^z f(t)dt - \{t\}f(t)\Big|_{t=y}^{z} + \int_y^z \{t\}f'(t)dt$$

Theorem 5.2. $\zeta(s) = \dfrac{s}{s-1} - s\int \dfrac{\{t\}}{t^{n+1}}$ *on* $Re(s) > 0$

Proof. Let's take $y = 1$, $z = \infty$, and $f(t) = \frac{1}{t^s}$ in the Euler-Maclaurin Forumla and get:

$$\zeta(s) - 1 = \sum_{n\geq 2} \frac{1}{n^s} = \int_1^\infty \frac{dt}{t^s} - s\int_1^\infty \frac{\{t\}}{t^{s+1}} = \frac{1}{s-1} - s\int_1^\infty \frac{\{t\}}{t^{s+1}}dt$$

Which is well defined on $Re(s) > 0$ except when $s = 1$ □

Remark 1: The same method extends to $Re(s) > -k$ for any natural number k, thanks to a higher order Euler-Maclaurin formula, which is deduced basically by integrate by part a lot of times while adjusting the height of the function in each step. The formula eventually gives us

$$\zeta(s) = \frac{s}{s-1} + \sum_{l=1}^{k} \frac{B_l}{l!}\prod_{j=o}^{l-2}(s+j) - \frac{1}{k!}\prod_{i=0}^{k-1}(s+j)\int_1^\infty \frac{B_k(\{x\})}{x^{s+k}}dx$$

where B_l are the Bernoulli numbers (constants) and $B_k(s)$ are polynomials of k-th degree. This formula does explicitly extends ζ to the whole plane as a meromorphic function, yet as you can see, it's just to complicated to be of any use for us. In fact, we rarely refer to the specific expression of ζ later.

Remark 2: Let ζ_k be the analytic continuation of ζ by the above method for $Re(s) > -k$, similar for ζ_l for $l < k$. Then they coincide on the part where their domain coincides. This is because they're the same on $Re(s) > 1$ and thus exists open balls in which they coincides. Basic complex analysis tells that they must be the same since they are both analytic.

5.2. **Functional Equation.**

We've shown that ζ can be analytically continued to the whole complex plane. But for most part of the plane we're not interested since we know well how ζ behaves. This is due to the functional equation ξ in this chapter.

Def 5.1. $\xi(s) = \pi^{-\frac{s}{2}}\Gamma(\frac{s}{2})\zeta(s)$ *where* Γ *is as in Def 1.2*

Theorem 5.3. $\forall s \in \mathbb{C} - \{1\}$, $\xi(s) = \xi(1-s)$

Remark: WOW!

Let's first do some preparations of this. We'll use the Poisson summation formula.

Def 5.2. *The* ***Fourier transform*** *of* $f : \mathbb{C} \to \mathbb{R}$ *is* $\hat{f}(\xi) = \int_{\mathbb{R}} f(x)e^{-2\pi i\xi x}dx$

Theorem 5.4. *(Poisson Summation Formula)* $\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \hat{f}(n)$ *for* f *good enough (say moderately decrease or Schwartz-Class).*

Proof. (of Poisson Summation Formula)

Let $g(x) = \sum_{n\in\mathbb{Z}} f(x+n)$, then g is periodic 1 and g converges since f is Schwartz Class. From Fourier series we get

$$g(x) = \sum_{n\in\mathbb{Z}} a_n e^{2\pi inx}$$

Where

$$\begin{aligned} a_n &= \int_0^1 g(x)e^{-2\pi inx}dx = \sum_{n\in\mathbb{Z}} \int_0^1 f(x+n)e^{-2\pi inx}dx \\ &\underset{y=n+x}{=} \sum_{n\in\mathbb{Z}} \int_n^{n+1} f(y)e^{-2\pi iny}e^{-2\pi in^2}dy = \sum_{n\in\mathbb{Z}} \int_n^{n+1} f(y)e^{-2\pi iny}dy \\ &= \hat{f}(n) \end{aligned}$$

Therefore

$$\sum_{n\in\mathbb{Z}} f(x+n) = g(x) = \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{2\pi inx}$$

Take $x = 0$

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \hat{f}(n)$$

□

Corollary 5.5. *For $x > 0$, let $f(u) = e^{-u^2\pi x}$, $\theta(x) = \sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} e^{-n^2\pi x}$. Then $\theta(x) = \frac{1}{\sqrt{x}}\theta(\frac{1}{x})$*

Proof.

$$\hat{f}(\xi) = \int_{\mathbb{R}} e^{-u^2\pi x} e^{-2\pi i\xi u} du \underset{\frac{y^2}{2}=u^2\pi x}{=} \int_{\mathbb{R}} e^{-\frac{y^2}{2}} e^{-i\xi\sqrt{\frac{2\pi}{x}}y} \frac{dy}{\sqrt{2\pi x}} = \frac{1}{\sqrt{x}} e^{-\frac{\pi}{x}\xi^2}$$

since we know from Gaussian that $\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{y^2}{2}} e^{iay} = e^{-\frac{a^2}{2}}$. By Possion summation formula:

$$\theta(x) = \sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \hat{f}(n) = \sum_{n\in\mathbb{Z}} \frac{1}{\sqrt{x}} e^{-\frac{\pi}{x}n^2} = \frac{1}{\sqrt{x}}\theta(\frac{1}{x})$$

□

Proof. (of **theorem 5.3**):

Remember that $\Gamma(s) = \int_0^\infty e^{-t} t^{s-1} dt$. Doing the change of variable $t = \pi n^2 x$ yields

$$\begin{aligned}
\Gamma(\frac{s}{2}) = \int_0^\infty e^{-\pi n^2 x} (\pi n^2 x)^{\frac{s}{2}-1} \pi n^2 dx = \pi^{\frac{s}{2}} n^s \int_0^\infty x^{\frac{s}{2}-1} e^{-n^2\pi x} dx \\
\iff \pi^{-\frac{s}{2}} \Gamma(\frac{s}{2}) n^{-s} = \int_0^\infty x^{\frac{s}{2}-1} e^{-n^2\pi x} dx \\
\iff \sum_{n\geq 1} \pi^{-\frac{s}{2}} \Gamma(\frac{s}{2}) n^{-s} = \sum_{n\geq 1} \int_0^\infty x^{\frac{s}{2}-1} e^{-n^2\pi x} dx
\end{aligned}$$

since $\sum_{n\in\mathbb{Z}} e^{-n^2\pi x}$ converges absolutely, letting $\omega(x) = \sum_{n\geq 1} e^{-n^2\pi x}$ we have

$$\pi^{-\frac{s}{2}} \Gamma(\frac{s}{2}) \zeta(s) = \int_0^\infty x^{\frac{s}{2}-1} \omega(x) dx$$

where the left hand side is exactly the definition of $\xi(s)$.

We note that $2\omega(x) + 1 = (\sum_{n\leq 1} + \sum_{n\geq 1} + \sum_{n=0}) e^{-n^2\pi x} = \theta(x)$, by corollary 5.5 we get

$$2\omega(x) + 1 = \theta(x) = \frac{1}{\sqrt{x}}\theta(\frac{1}{x}) = \frac{1}{\sqrt{x}}[2\omega(\frac{1}{x}) + 1]$$

So

$$\begin{aligned}
\xi(s) &= \int_1^\infty x^{\frac{s}{2}-1}\omega(x)dx + \int_0^1 x^{\frac{s}{2}-1}\omega(x)dx \\
&= \int_1^\infty x^{\frac{s}{2}-1}\omega(x)dx + \int_1^\infty (\frac{1}{x})^{\frac{s}{2}-1}\omega(\frac{1}{x})\frac{dx}{x^2} \\
&= \int_1^\infty x^{\frac{s}{2}-1}\omega(x)dx + \int_1^\infty x^{-(\frac{s}{2}-1)}(-\frac{1}{2}+\frac{\sqrt{x}}{2}+\sqrt{x}\omega(x))\frac{dx}{x^2} \\
&= \int_1^\infty x^{-\frac{s}{2}-1}(-\frac{1}{2}+\frac{\sqrt{x}}{2})dx + \int_1^\infty (x^{\frac{s}{2}-1}+x^{-\frac{s}{2}-1})\omega(x)dx \\
&= \frac{1}{s(s-1)} + \int_1^\infty (x^{\frac{s}{2}-1}+x^{-\frac{s}{2}-1})\omega(x)dx
\end{aligned}$$

Since the last expression is invariant if we change s to $1-s$, so $\xi(s) = \xi(1-s)$. □

Since Γ is extended to the whole plane and ζ on $Re(s) > 0$, the ξ function's symmetry offers us an analytic continuation of ζ to the whole plane once and for all.

This continuation of ζ offers us a very easy way to characterize the zeros of ζ, which we characterize as the following:

Remark 1: If $Re(s) > 1$, then $\zeta(s) \neq 0$. We give 2 proofs of this statement.

Proof. 1 of the statement:

Since μ is the inverse of $\mathbb{1}$, we've seen multiple times in chapter 4 that $\zeta(s)\sum_{n\geq 1}\frac{\mu(n)}{n^s} = 1$ for $Re(s) > 1$. So $\zeta(s)$ is zero at some point means that $\sum_{n\geq 1}\frac{\mu(n)}{n^s}$ approaches ∞ at that point. Yet we know $|\sum_{n\geq 1}\frac{\mu(n)}{n^s}|$ is finite. This concludes the proof. □

Proof. 2 of the statement:

$$\zeta(s) = \prod_{p\in\mathscr{P}}\frac{1}{1-\frac{1}{p^s}} \iff \frac{1}{|\zeta(s)|} = \prod_{p\in\mathscr{P}}|1-\frac{1}{p^s}| \leq \prod_{p\in\mathscr{P}}(1+\frac{1}{p^s}) < \infty$$

where the last step is because

$$\log(\prod_{p\in\mathscr{P}}(1+\frac{1}{p^s})) = \sum_{p\in\mathscr{P}}\log(1+\frac{1}{p^s}) \leq \sum_{p\in\mathscr{P}}\frac{1}{p^s} < \infty$$

□

Remark 2: By Theorem 1.14 in Book, we know that $\forall s \in \mathbb{C}, \Gamma(s) = \frac{e^{-\gamma s}}{s}\prod_{n=1}^{\infty}\frac{e^{\frac{s}{n}}}{1+\frac{s}{n}}$ where γ is the Euler Constant. This tells us that Γ has no zero and simple poles at $s \in \{0, -1, -2\cdots\}$.

Applying the result to $\xi(s) = \pi^{-\frac{s}{2}}\Gamma(\frac{s}{2})\zeta(s)$, we see that $\xi(s) \neq 0$ for any $Re(s) > 1$ by remark 1. But the same applies for $Re(s) < 0$ due to the symmetry of ξ with respect to $Re(s) = \frac{1}{2}$. Since ζ and Γ are analytic on $Re(s) > 0$, so is ξ.

In particular, for $\frac{1-s}{2} \in \{-1, -2 \cdots\}$, $\xi(s) = \xi(1-s) = \pi^{-\frac{1-s}{2}}\Gamma(\frac{1-s}{2})\zeta(1-s)$ is bounded and non-zero, so $\zeta(1-s)$ must have a simple zero. At $\frac{1-s}{2} = 0$, $s = 1$, and we know already that $\Gamma(0)$ and $\zeta(1)$ are two simple poles.

More explicitly, we summarize the above remarks and conclude $\zeta(s) = 0$ if s is an negative even integer out side of the strip $0 \leq Re(s) \leq 1$. We call this strip **the Critical Strip** of ζ and all zeros outside of it **trivial zeros**. (It's not really trivial, but we know really well about it.)

Just for fun, we can calculate ζ at a few more values. (Fill this up later)

We now observe that if $\zeta(s) = 0$ for $s = \sigma + it$ in the critical strip, then so is $s = (1-\sigma)+it$, $s = \sigma - it$, and $s = (1-\sigma) - it$. Taking the conjugate we see that $\zeta(\sigma - it) = 0$, and using the functional equation we easily see that the other two are zeros too.

This is where we can get too now. Our goal is captured by Riemann Hypothesis:

Hypothesis 5.6. *(Riemann Hypothesis)* $\begin{cases} \zeta(s) = 0 \\ 0 \leq Re(s) \leq 1 \end{cases} \Rightarrow Re(s) = \frac{1}{2}$

5.3. **Primes & Riemann Hypothesis: Heuristics.**

We study the heuristics of the relation between Riemann Hypothesis and the distribution of primes in this section. Therefore, this section will be largely informal.

Def 5.3. $\psi(x) = \sum_{n \leq x} \Lambda(n)$

We will in later sections prove a theorem that is informally represented as:

Informal Theorem: $\psi(s) \approx x - \sum_{\rho} \frac{x^\rho}{\rho}$, where the sum is over non-trivial zeroes of ζ.

Admitting the informal theorem we can see why Riemann Hypothesis is related to the proof of PNT.

Theorem 5.7. *Riemann Hypothesis* $\iff \psi(x) = x + O(x^{\frac{1}{2}+\epsilon}) \iff \pi(x) = \frac{x}{\log x} + O(x^{\frac{1}{2}+\epsilon})$ *for any* ϵ.

Proof. (informal)

Riemann Hypothesis $\iff \psi(x) = x + O(x^{\frac{1}{2}+\epsilon})$:

$\Rightarrow$: This direction is obvious. If all non-trivial zeroes falls on the line $Re(s) = \frac{1}{2}$, then as long as $\sum_\rho \frac{1}{\rho}$ is summable, we get the right hand side. Why it is summable will be explained in section 5.4.

$\Leftarrow$: Assume $\psi(x) = \sum_{n\le x} \Lambda(x) = x + O(x^{\frac{1}{2}+\epsilon})$, then $\psi(x) - x = O(x^{\frac{1}{2}+\epsilon})$. Example 4.3 gives us that

$$\frac{\zeta'}{\zeta}(s) = -\sum_{n\ge 1} \frac{\Lambda(n)}{n^s} = -\int_1^\infty \frac{1}{x^s} d\psi(x) = \frac{s}{s-1} + s\int_1^\infty \frac{\psi(s)-x}{x^{s+1}} dx$$

is well-defined for $\mathrm{Re}(s) > \frac{1}{2}$, i.e., $\frac{\zeta'}{\zeta}$ is finite for the same half-plane.

But for any meromorphic function f, if $f = 0$ at some point x, then $(\log f)'$ has a pole at x since $(\log f)' = \frac{1}{(z-x)^m} + \frac{g'}{g}$ where g is analytic. This means that ζ has no zero for any $Re(s) \ge \frac{1}{2} + \epsilon$. Since ϵ is arbitrary, we get RH.

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

$\Rightarrow$: (To be completed: he actually didn't show this in class!)

$\Leftarrow$: (To be completed: he actually didn't show this in class!)

□

So we can say that Riemann Hypothesis offers us a strict bound on the distribution of primes, which is believed but yet to be proven. Note that the role of ϵ is simply to balance out the $\log x$ in the term.

Now we give an informal proof of the informal theorem, which will be a prelude of what is going on in the next few sections.

Informal Proof of The Theorem:

The key point the is Perron Inversion Formula: $(\alpha > 0, \forall y > 0, y \ne 1)$

$$\frac{1}{2\pi i}\int_{Re(s)=\alpha} \frac{y^{-s}}{s} ds = \mathbb{1}_{0<y<1}$$

where the integral on a vertical line is defined as

$$\frac{1}{2\pi i}\int_{Re(s)=\alpha} \frac{y^{-s}}{s} ds := \lim_{T\to\infty} \frac{1}{2\pi i}\int_{\substack{Re(s)=\alpha \\ Im(s)<T}} \frac{y^{-s}}{s} ds$$

Going back to the question, what we care about is really how to estimate $\sum_{n\le x} f(n)$ from its Dirichlet series $F(s) = \sum_{n\ge 1} \frac{f(n)}{n^s}$. This is done by the trick due to Perron's formula:

$$\sum_{n\le x} f(n) = \sum_{n\ge 1} f(n)\mathbb{1}_{\frac{n}{x}\in(0,1]} = \sum_{n\ge 1} \frac{f(n)}{2\pi i}\int_{Re(s)=\alpha} \frac{(\frac{n}{x})^{-}s}{s} ds = \frac{1}{2\pi i}\int_{Re(s)=\alpha} \frac{x^{-s}}{s} F(s) ds$$

Plugging in $f = \Lambda$, we get

$$\psi(x) = \frac{1}{2\pi i}\int_{Re(s)=\alpha} \frac{x^{-s}}{s} \cdot (-\frac{\zeta'}{\zeta}(s)) ds$$

From here we see that we only need to bound ζ.

5.4. **A Priori Estimates on Zeta-Function.**

We have seen from last chapter's discussion that we are interested in bounding $\zeta(\sigma + it)$ as $t \to \infty$, since we are doing the line integral in some step. Also, we need to show that $\sum \frac{1}{\rho}$ is summable as well as the bounds on $\frac{\zeta'}{\zeta}$. We'll do that in this section.

Theorem 5.8. *Let $\varepsilon, C > 0$, $s = \sigma + it$, then $\exists A_{\varepsilon,C}$ such that*

$$|\zeta(s)| \le A_{\varepsilon,C} \begin{cases} 1 & \text{if } \sigma \ge 1+\varepsilon \\ |t|^{\frac{1-\sigma+\varepsilon}{2}} & \text{if } -\varepsilon \le \sigma \le 1+\varepsilon \\ |t|^{\frac{1}{2}-\sigma} & \text{if } -C \le \sigma \le -\varepsilon \end{cases}$$

We proof first the complex analysis tool for the proof:

Lemma 5.9. *Let D be the strip $-\frac{\pi}{2} \le Re(s) \le \frac{\pi}{2}$. If $|h(z)| \le \exp(c \cdot e^{at})$ where $a \in (0,1)$, and $|h| \le 1$ on ∂D, we have $|h| \le 1$ on D.*

We see that this is an extended version of the maximum modules principle.

Proof. (of lemma)

The only problem of using the Maximum Modulues Principle is that D is unbounded. So let's truncate D into a tall rectangle and deal with the three parts separately.

Let $a < b < 1$, $h_\varepsilon(z) = \exp(-\varepsilon(e^{ibz} - e^{-ibz}))$.

Notice that $h \cdot h_\varepsilon \underset{z\to\pm i\infty}{\to} 0$ because

$$|h \cdot h_\varepsilon(z)| \le \exp(ce^{at}) \cdot \exp(-\varepsilon e^{bt}) \underset{|t|\to\infty}{\to} 0$$

as $b > a$.

Hence, $\forall 0 < k < 1, \exists C_k$ such that $|h \cdot h_\varepsilon(z)| \le k$ for $|t| > C_k$.

We put our cutoff line at $|t| = C_k$ and get the rectangle $R := \{D \bigcap |Im(s)| < C_k\}$. We have just proven $|h \cdot h_\varepsilon| \leq 1$ on $D \backslash R$, and we use maximum principle to get $|h \cdot h_\varepsilon| \leq 1$ on R. So $|h \cdot h_\varepsilon| \leq 1$ on D.

But as $\varepsilon \to 0$, $|h_\varepsilon| \to 1$ and we get $|h| \leq 1$ on D.

□

Theorem 5.10. *(Phragmen-Lindelöf Theorem):*

On the strip $\alpha_1 \leq Re(s) \leq \alpha_2$, under these conditions,

- *f is analytic*
- *$f(s) \leq c \cdot e^{|s|^c}$*
- *$|f(\alpha_1 + it)| \leq c(1 + |t|)^{\theta_1}$*
- *$|f(\alpha_2 + it)| \leq c(1 + |t|)^{\theta_2}$*

we have $f(\sigma + it) \leq \tilde{C}(1 + |t|)^{u\theta_1 + (1-u)\theta_2}$ for $\sigma = u\alpha_1 + (1-u)\alpha_2 \in (\alpha_1, \alpha_2)$.

Proof. (of Phragmen-Lindelof) We apply Lemma 5.9.

WLOG let $\alpha_1 = 0$ and $\alpha_2 = 1$, $\theta_1, \theta_2 \geq 0$. Also, define

$$h_t(z) := \frac{f(z + it)}{c(1 + |t|)^{(1-z)\theta_1 + z\theta_2} \cdot (1 + z)^{max(\lceil \theta_1 \rceil, \lceil \theta_2 \rceil)}}$$

One might ask what is the reason behind this definition, so we list a few:

- $Re(1 - z)\theta_1 + Re(z)\theta$ needs to appear somewhere in the proof. But that expression kills analyticity, so we take off the problematic $Re(\ )$.
- We shift z vertically by $i \cdot t$ because when taking off $Re(\ )$ we know that we want to look at $z \in \mathbb{R}$.
- $(1+z)^{max(\lceil \theta_1 \rceil, \lceil \theta_2 \rceil)}$ is analytic and it is inserted to bound h_t on the two sides of the strip by 1.

  For example, when $\sigma = 0$,

  $$|h_t(iy)| = \left| \frac{f(it + iy)}{c(1 + |t|)^{\theta_1} \cdot (1 + iy)^{max(\lceil \theta_1 \rceil, \lceil \theta_2 \rceil)}} \right| \leq \frac{c(1 + |y + t|)^{\theta_1}}{c(1 + |t|)^{\theta_1}(1 + |y|)^{\theta_1}} \leq 1$$

  due to the inequality

  $$(1 + |y + t|)^{\theta_1} \leq (1 + |t|)^{\theta_1}(1 + |y|)^{\theta_1}$$

  A similar argument says the same for $\sigma = 1$.

But now we have $|h_t(z)| \leq 1$ on the strip, by Lemma 5.9 $|h_t(\sigma)| \leq 1$ for any $\sigma \in [0, 1]$, which further says $f(\sigma + it) \leq \tilde{C}(1 + |t|)^{u\theta_1 + (1-u)\theta_2}$ for $\sigma = u\alpha_1 + (1-u)\alpha_2 \in (\alpha_1, \alpha_2)$.

□

Proof. (of theorem 5.8)

$\sigma \geq 1+\varepsilon$:

$$\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s} \Rightarrow |\zeta(s)| \leq \sum_{n\geq 1} \frac{1}{n^\sigma} \leq \sum_{n\geq 1} \frac{1}{n^{1+\varepsilon}}$$

is a fixed constant in ε

$-C \leq \sigma \leq -\varepsilon$:

Using the functional equation ξ we have $\pi^{-\frac{s}{2}}\Gamma(\frac{s}{2})\zeta(s) = \pi^{-\frac{1-s}{2}}\Gamma(\frac{1-s}{2})\zeta(1-s)$. Looking at the terms we notice:

- $\zeta(1-s) = O(1)$ is just a constant in this region.
- $|\pi^{-\frac{s}{2}}| = |\pi^{-\frac{\sigma}{2}}|$ and $\sigma \geq -c$ tells us $|\pi^{-\frac{s}{2}}| = O(1)$, also a constant.

Therefore

$$|\zeta(s)| = O\left(\left|\frac{\Gamma(\frac{1-s}{2})}{\Gamma(\frac{s}{2})}\right|\right)$$

Fact: for $|\sigma| \leq c$, $\Gamma(s) \asymp |t|^{\sigma-\frac{1}{2}}e^{-\pi\frac{|t|}{2}}$

Reason: By Sterling's Formula, $k! \sim \sqrt{2\pi k}(\frac{k}{e})^k$, which means $|\Gamma(s)| \approx |s|^s \approx |t|^\sigma$ There's some gap here that is to be Filled.

Then

$$|\zeta(s)| = O\left(\left|\frac{\Gamma(\frac{1-s}{2})}{\Gamma(\frac{s}{2})}\right|\right) = O\left(\frac{t^{\frac{1-\sigma}{2}-\frac{1}{2}}}{t^{\frac{\sigma}{2}-\frac{1}{2}}}\right) = O(t^{\frac{1}{2}-\sigma})$$

$-\varepsilon \leq \sigma \leq 1+\varepsilon$:

We use Phragmen-Lindelof Theorem here. Let $f(s) = \zeta(s)(s-1)$, then f is analytic between $-\varepsilon \leq \sigma \leq 1+\varepsilon$. By the explicit formula of ζ attained by Euler-Maclaurin formula, we know that ζ grows in t at most polynomially, so condition 2 is satisfied. Since the term (s-1) provides +1 in θ, we have $\theta_1 = \frac{3}{2}+\varepsilon$ and $\theta_2 = 1$ due to the previous two cases.

Say $\sigma = u(-\varepsilon) + (1-u)(1+\varepsilon) = 1+\varepsilon-u-2u\varepsilon$, then

$$u\theta_1 + (1-u)\theta_2 = \frac{3}{2}u + u\varepsilon + 1 - u = (u\varepsilon + \frac{1}{2}u) + 1 = -\frac{\sigma}{2} + \frac{1}{2} + \frac{\varepsilon}{2} + 1 = \frac{3-\sigma+\varepsilon}{2}$$

Therefore,

$$f(\sigma+it)(s-1) = O\big((1+|t|)^{\frac{3-\sigma+\varepsilon}{2}}\big) \Rightarrow f(\sigma+it) = O\big((1+|t|)^{\frac{1-\sigma+\varepsilon}{2}}\big)$$

which is what we want. □

With the bound on ζ, let's look at the distribution of zeros of ζ and an esitmate of $\frac{\zeta'}{\zeta}$:

Theorem 5.11. *Let $\rho = \beta + i\gamma$ be the non-trivial zeros of ζ, then we have:*

(a) $\#\{\rho : |\gamma - t| \leq 1\} = O_{t\to\infty}(\log t)$

(b) $\dfrac{\zeta'}{\zeta}(s) = -\dfrac{1}{s-1} + \sum_{|\gamma - t|<1} \dfrac{1}{s-\rho} + O(\log s)$

Again, to prove this theorem we need 2 lemmas.

Lemma 5.12. *Let g be analytic, $g(0) \neq 0$, $|g(z)| \leq e^M |g(0)|$ on $\mathscr{C}_{2R}(0) := \{|z| = 2R\}$ (the circle), then $\#\{|z_i| \leq R\} \leq 2M$, where $|z_i| < 2R$ are the zeros of g.*

Proof. (of Lemma 5.12)

Let $h(z) = g(z) \cdot \prod_{l=1}^{k} \dfrac{(2R - \frac{z\bar{z}_l}{2R})}{z - z_l}$, then $|h(z)| = |g(z)|$ on $\mathscr{C}_{2R}(0)$ since

$$|\frac{z}{2R}| = 1 \Rightarrow \frac{2R}{\bar{z}} = \frac{z}{2R}$$

and thus

$$|2R - \frac{z\bar{z}_l}{2R}| = |2R - \frac{2R}{\bar{z}}\bar{z}_l| \cdot |\frac{\bar{z}}{2R}| = |\bar{z} - \bar{z}_l| = |z - z_l|$$

Therefore,

$$e^M |g(0)| \geq \max_{\mathscr{C}_{2R}(0)} |g(z)| = \max_{\mathscr{C}_{2R}(0)} |h(z)| \underset{\substack{\text{Max Modules} \\ \text{Principle}}}{\geq} |h(0)| = g(0) \prod_{l=1}^{k} |\frac{2R}{z_l}| \geq g(0) \cdot 2^k$$

$$\Rightarrow k \leq \frac{M}{\log 2} \leq 2M$$

Note: we can assume there's only finitely many zeros since the disk is compact & complex analysis. □

Lemma 5.13. *Let g be analytic, $g(0) \neq 0$, $|g(z)| \leq e^M |g(0)|$ on $\mathscr{C}_{4R}(0)$. Also, let $z_1, z_2, \cdots, z_k$ be zeros of g on the disk $\mathbb{D}_{2R}(0)$. Then $\forall |z| < R$,*

$$\left| \frac{g'(z)}{g(z)} - \sum_{i=1}^{k} \frac{1}{z - z_i} \right| \leq \frac{16M}{R}$$

Proof. (of Lemma 5.13)

Let $G(z) = \dfrac{g(z)}{\prod_{i=1}^{k}(z - z_i)}$, then G is analytic on $|z| < 2R$ and has no zero on $|z| < 2R$.

Let $f(z) = \log \dfrac{G(z)}{G(0)}$, which is analytic on $|z| < 2R$ (since 0 not in the area and the area is connected). Notice that f' is exactly what we're looking for.

Now to make connection with M we use maximum modules principle to get

$$\max_{|z|=2R} \left| \frac{G(z)}{G(0)} \right| \leq \max_{|z|=4R} \left| \frac{G(z)}{G(0)} \right| = \max_{|z|=4R} \left| \frac{g(z)}{g(0)} \prod_{l=1}^{k} \frac{z_l}{z - z_l} \right| \leq e^M$$

since $|z_l| < 2R < |z - z_l|$ on the circle.

$\forall |z| < R$, $f'(z) = -\frac{1}{2\pi i}\int_\gamma \frac{f(w)}{(z-w)^2}dw$ by Cauchy's formula, so

$$f'(z) \le c\frac{\max\limits_{|\omega|=2R}|f(\omega)|}{R^2}R = \frac{CM}{R}$$

where $C = 16$ is enough.

□

Proof. (of Theorem 5.11)

(a)$\#\{\rho : |\gamma - t| \le 1\} = O_{t\to\infty}(\log t)$:

Let $g(z) = \zeta(2 + it + z)$. Since we want to apply Lemma 5.12 and not want $s = 1$ to be on the circle $\mathscr{C}_{2R}(2 + it + z)$, we can be generous and choose $|t| \ge 100$, $R = 10$. By Lemma 5.12 we have:$(a_t = 2 + it)$

$$\#\{\rho : |\gamma - t| \le 1\} \le \#\{\rho : |\rho - a_t| < 10\} \le 2 \cdot \log\left(\frac{\max\limits_{\mathscr{C}_{20}(a_t)}|\zeta|}{\zeta(a^t)}\right)$$

We have

- $\max\limits_{\mathscr{C}_{20}(a_t)}|\zeta| \le |t|^{10}$ due to theorem 5.8.
- $|\zeta(a_t)| \ge 1 - \sum\limits_{n\ge1}\frac{1}{n^2} > \frac{1}{10}$ since ζ's series expression converges at a_t, and each term after 1 we pick the minimum.

Therefore

$$\#\{\rho : |\gamma - t| \le 1\} \le 2 \cdot \log\left(\frac{|t|^{10}}{1/10}\right) \le c \cdot \log(|t|)$$

(b) $\frac{\zeta'}{\zeta}(s) = -\frac{1}{s-1} + \sum\limits_{|\gamma-t|<1}\frac{1}{s-\rho} + O(\log s)$:

For $|t| > 100, -1 < \sigma < 2$,

$$\left|\frac{\zeta'}{\zeta}(s) - \sum_{i=1}^{k}\frac{1}{z - z_i}\right| \le c\log t$$

by Lemma 5.13 and Theorem 5.8 since $\frac{\zeta(\sigma+it)}{\zeta(0)} \le \frac{|t|^{A_\sigma}}{-\frac{1}{2}}$, taking the log we see we're only left with $c\log t$.

For s close to 1, we apply the lemma and method to $g(s) = \zeta(s)\cdot(s-1)$ and $-\frac{1}{s-1}$ jumps out of the calculation. (Full proof in book.)

□

5.5. **Perron Inversion Formula.**

In this chapter we justify some results mentioned in section 5.3.

Lemma 5.14. *Uniformly in $y > 0, \alpha > 0, T \geq 1$ we have*

$$\frac{1}{2\pi i}\int_{\substack{Re(s)=\alpha \\ |Im(s)|\leq T}} \frac{y^{-s}}{s} ds = \begin{cases} \mathbb{1}_{0<y<1} + O(\frac{1}{y^\alpha \max\{1, T|\log y|\}}) & y \neq 1 \\ \frac{1}{2} + O(\frac{\alpha}{T}) & y = 1 \end{cases}$$

Proof.

Case 1: $0 < y < 1$

Let $\gamma_1 = [\alpha + iT, \alpha - l + iT], \gamma_2 = [\alpha - l + iT, \alpha - l - iT], \gamma_3 = [\alpha - l - iT, \alpha - iT], \gamma_4 = [\alpha - iT, \alpha + iT]$ trace the rectangle. Then

$$\int_{\gamma_1 \bigcup \gamma_2 \bigcup \gamma_3 \bigcup \gamma_4} \frac{1}{2\pi i}\frac{y^{-s}}{s} ds = Res(s = 0) = 1$$

Bounding each sides we have:

- $\int_{\gamma_2} \frac{y^{-s}}{s} ds \leq 2T\frac{y^{l-\alpha}}{l-\alpha} \to 0$ as $l \to \infty$
- $\int_{\gamma_1} \frac{y^{-s}}{s} ds \underset{s=u+it}{\leq} \frac{1}{T}\int_{\alpha-l}^{\alpha} y^{-u} du \underset{l\to\infty}{\to} \frac{1}{T}\int_{-\infty}^{\alpha} e^{-u\log y} \leq \frac{1}{T}\left[-\frac{e^{-u\log y}}{\log y}\right]_{-\infty}^{\alpha} \leq \frac{1}{Ty^\alpha|\log y|}$, which is the error term. Note that we can always improve as $y \to 1$ by taking the maximum between $T|\log y|$ and 1, which yields the error term exactly.
- Integral on γ_3 is exactly the same as on γ_1

So we conclude that the integral in question is $\int_{\gamma_4} \frac{1}{2\pi i}\frac{y^{-s}}{s} ds = 1 + O(\frac{1}{y^\alpha \max\{1, T|\log y|\}})$

Case 2: $y > 1$

For this case we use the another rectangle that goes to positive infinity, i.e. the rectangle with verticies $[\alpha + iT, \alpha + l + iT, \alpha + l - iT, \alpha - iT]$. We see that the boundaries yields the exact same result and there is no residue, so we're done with this case. We flip the side of integration due to the fact that y^{-s} need to go to 0.

Case 3: $y = 1$

Adding conjugate to $\frac{1}{s} = \frac{1}{\alpha + il}$ we get:

$$\begin{aligned} \frac{1}{2\pi i}\int_{Re(s)=\alpha; |Im(s)|\leq T} \frac{1}{s} ds &= \frac{1}{4\pi}\int_{-T}^{T} (\frac{1}{\alpha + il} + \frac{1}{\alpha - il})\frac{idl}{i} \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \frac{\alpha}{\alpha^2 + l^2} dl + O(\int_{T}^{\infty} \frac{\alpha}{l^2}) dl \\ &=_{l=\alpha v} \frac{1}{2\pi}\int_{\mathbb{R}} \frac{1}{1+v^2} dv + O(\frac{\alpha}{T}) = \frac{1}{2} + O(\frac{\alpha}{T}) \end{aligned}$$

□

Let's now prove two corollaries of the lemma.

Corollary 5.15. *If* $|f(n)| \leq cn^{\theta}(1+\log n)^{A}$ *for* $A, c > 0, \theta \geq -1$, F *is the Dirichlet series of* f *and* $x, T \geq 2$ *uniformly,* $\alpha \geq \theta + 1 + \frac{1}{\log x}$, *then*

$$\sum_{n \leq x} f(n) = \int_{\substack{Re(s)=\alpha \\ |Im(s)| \leq T}} F(s)\frac{x^s}{s} ds + O\Big(\frac{x^{\alpha}(\log x)^{A+1}}{T} + x^{\alpha-1}(\log x)^{A}\Big)$$

Proof.

The proof is an exercise. Nonetheless, a sketch is offered:

pick $y = \dfrac{n}{x}$ in lemma 5.14, then the error becomes

$$\sum_{n \leq T} |f(n)| \frac{1}{(\frac{n}{x})^{\alpha} \max\{1, T\log\frac{n}{x}\}} \leq n^{\theta}(1+\log n)^{A}$$

using the method of "forgot the logs." □

Corollary 5.16. $\psi(x) = x - \sum\limits_{|\gamma| \leq T} \dfrac{x^{\rho}}{\rho} + O\Big(\dfrac{x\log^2(xT)}{T} + \log x\Big)$, *where* ρ *are the non-trivial zeros of* ζ *as usual.*

Proof.

From Corollary 5.15, taking $\alpha = 1 + \dfrac{1}{\log x}$ and $A = 1$ we have

$$\psi(x) = \frac{1}{2\pi i} \int_{\substack{Re(s)=1+\frac{1}{\log x} \\ |Im(s)| \leq T}} -\frac{\zeta'}{\zeta}(s)\frac{x^s}{s} ds + O\left(\frac{x}{T}\log^2 x + \log x\right)$$

Now we do the contour integration again with the Rectangle parameterized by

- $\gamma = [1 + \dfrac{1}{\log x} - iT, 1 + \dfrac{1}{\log x} + iT]$
- $L_1 = [1 + \dfrac{1}{\log x} + iT, -2N - 1 + iT]$
- $L_0 = [-2N - 1 + iT, -2N - 1 - iT]$
- $L_0 = [-2N - 1 - iT, 1 + \dfrac{1}{\log x} - iT]$

Then the integral is converted into:

$$\psi(x) = \sum_{\omega : poles} Res(s = \omega, -\frac{\zeta'}{\zeta}(s)\frac{x^s}{s}) - \int_{L_1 \bigcup L_0 \bigcup L_{-1}} -\frac{\zeta'}{\zeta}(s)\frac{x^s}{s} ds + O(error)$$

We will show that the residues contributs all the leading terms and the remaing integral is the error.

Now let's look at the poles, which are either $s = 1$, the trivial zeros or the non-trivial zeros.

- Around $s = 1$, $\zeta \sim \frac{1}{s-1}$, so $\log \zeta \sim -\log(s-1)$, taking the derivative we get $-\frac{\zeta'}{\zeta}(s) \sim \frac{1}{s-1} \Rightarrow -\frac{\zeta'}{\zeta}(s)\frac{x^s}{s} \sim \frac{x}{s-1}$, so the residue is x.
- Close to the non-trivial zeros: we know that the zeros can be viewed as simple zeros since we count them according to multiplicity. So $\zeta(s) \sim c(s-\rho)$. Therefore by the same method as above we get
$$-\frac{\zeta'}{\zeta}(s) \sim -\frac{1}{s-\rho} \Rightarrow -\frac{\zeta'}{\zeta}(s)\frac{x^s}{s} \sim -\frac{x^s}{s(s-\rho)} \Rightarrow Res = -\frac{x^\rho}{\rho}$$
- Close to trivial zeros: computing the residue balance out $\frac{\zeta'}{\zeta}$ and so $-\frac{\zeta'}{\zeta}(s)\frac{x^s}{s} \sim \frac{x^{-2n}}{2n}$. Which is summable in x, so the order is O(1).

So $\sum_{\omega : poles} Res(s = \omega, -\frac{\zeta'}{\zeta}(s)\frac{x^s}{s}) \sim x - \sum_{|\gamma| \le T} \frac{x^\rho}{\rho}$.

We now deal with the integral.

On L_0: (By theorem 5.11 (b) and (a))

$$\frac{\zeta'}{\zeta}(s) = -\frac{1}{s-1} + \sum_{|\gamma - t| \le 1} \frac{1}{s-\rho} + O(\log|s|) \le \#\{|\gamma - t| \le 1\}\frac{1}{|s|} + O(\log|s|)$$
$$= \log|t|\frac{1}{|s|} + O(\log|s|) \le O(\log(2N+1))$$
$$\Rightarrow \int_{L_0} \left|\frac{\zeta'}{\zeta}\frac{x^s}{s}\right| \le O\left(\int_{L_0} \log(2N+1)\frac{x^{-2N-1}}{|s|}\right) = O(1) \text{ or } =_{N\to\infty} o(1)$$

On L_1 (same for L_{-1}):

We want no zeros to be too close to our contour, so let's assume we choose T such that no ρ is of distance $\frac{1}{C\log T}$ to L_1. This is in fact achievable since there are only $O(\log T)$ zeros such that $|\gamma - T| \le 1$ and by pigeonhole principle we can choose such a t for some C. Let's do that, then

$$\frac{\zeta'}{\zeta} \ll (\log T)^2 + \log T \ll (\log T)^2$$

where the square is made of $\log T$ summands and each summand is bounded by $O(\log T)$.

Then we further have

$$\int_{L_1} -\frac{\zeta'}{\zeta}(s)\frac{x^s}{s}ds \ll (\log T)^2\left(\int_{-1}^{1+\frac{1}{\log x}} \frac{x^\sigma}{T}d\sigma + \int_{-2N-1}^{-1} \frac{x^\sigma}{T}d\sigma\right)$$

where

$$\int_{-1}^{1+\frac{1}{\log x}} x^\sigma d\sigma = \frac{x^\sigma}{\log x}\Big|_{-1}^{1+\frac{1}{\log x}} \asymp \frac{x}{\log x}$$

and

$$\int_{-\infty}^{-1} x^{\sigma} d\sigma = \int_{1}^{\infty} e^{-\sigma \log x} d\sigma = \frac{1}{logx} \left[e^{-\sigma \log x} \right]_{1}^{\infty} \asymp \frac{x}{\log x}$$

Hence, the integral is of order $O\left(\frac{x}{\log x} \frac{\log^2 T}{T}\right)$, which is what we want.

Putting together every piece we get exactly our corollary. □

5.6. **The Prime Number Theorem.**

The Prime Number Theorem $\pi(x) \sim \frac{x}{\log x}$ is equivalent to $\psi = x + o(x)$ for a similar argument as we did in Theorem 5.7 second half. But then corollary 5.16 tells us if $\sum_{\rho} \frac{x^{\rho}}{\rho} = o(x)$, then we are done.

So what we need is only that the zeros of ζ never get too close to $Re(s) = 1$. So RH implies the theorem as shown in Theorem 5.7. But we cannot prove RH. Instead, we use a much weaker, but proved, zero-free region of ζ:

Theorem 5.17. *(Zero-free region by Hadamand & Lavallee Poussin)*
$\exists c > 0$ *such that* $\forall \sigma \geq 1 - \frac{c}{\log(|t| + 2)}$, $\zeta(\sigma + it) \neq 0$

We postpone the proof of the zero-free region and first see how we can use it to proof the Prime Number Theorem in detail.

Proof. (of Theorem 5.1)

The zero-free region of *zeta* plus corollary 5. 16 gives us the estimate:

$$\psi = x + O\left(\sum_{|\gamma \leq T|} \frac{x^{1 - \frac{c}{\log T}}}{|\rho|} + \frac{x(\log x)^2}{T} \right)$$

where we can choose T to go to infinity at an order of x, so the error is not too large.

Since for the most interesting zeros we have $|\rho| = |\gamma|$,

$$\sum_{|\gamma| \leq T} \frac{1}{|\rho|} \ll \sum_{|\gamma| \leq T} \frac{1}{|\gamma|} \ll \sum_{k \leq T} \sum_{|\gamma| \in (k, k+1]} \frac{1}{|\gamma|} \ll \sum_{k \leq T} \frac{\log k}{|k|} \ll (\log T)^2$$

therefore

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

To make it smallest we choose $T = e^{\sqrt{\log x}}$, then the order becomes

$$O\left(\frac{x(\log x)}{e^{\sqrt{\log x}}} + \frac{x(\log x)^2}{e^{\sqrt{\log x}}} \right) = O\left(\frac{x(\log x)^2}{e^{\sqrt{\log x}}} \right) = o(x)$$

This tells us $\psi(x) \sim x$, which is equivalent to $\pi(x) \sim \frac{x}{\log x}$, and we're done. □

Since the proof of Theorem 5.17 is rather technical, we first give a formal proof to make it intuitive, then complete the details.

Formal Proof of Theorem 5.17:

The heuristic is to see what is happening around the line $Re(s) = 1$.

Assume $\zeta(1 + i\gamma_0) = 0$, then $\zeta(\sigma + i\gamma_0) \underset{\sigma \to 1}{\sim} c(\sigma - 1)$ since it would be a simple zero.

Yet on the other hand $\zeta(\sigma + i\gamma_0) = \prod_{p \in \mathscr{P}} \left(1 - p^{-(\sigma + i\gamma_0)}\right)^{-1}$.

So by the analyticity of ζ it implies that as $\sigma \to 1$, $p^{i\gamma_0} \simeq -1$ for most primes, so that $\prod_{p \in \mathscr{P}} \left(1 - p^{-(\sigma + i\gamma_0)}\right) \to \infty$. ("most" compared to some kind of cancellation around the circle).

But then $(p^{i\gamma_0})^2 \simeq 1 \Rightarrow p^{-2\gamma_0} \simeq 1$, which in turn means

$$\zeta(\sigma + 2i\gamma_0) \simeq \prod_{p \in \mathscr{P}} \frac{1}{1 - \frac{1}{p^{\sigma + 2i\gamma_0}}} \simeq \prod_{p \in \mathscr{P}} \frac{1}{1 - \frac{1}{p}} \to \infty$$

which contradicts the fact that ζ has only one pole at $s = 1$.

Now to make it rigorous, we capture the notion of "mostly -1" by a distance.

Proof. (of Theorem 15.17)

Define the distance

$$D_\sigma^2(f, g) = \frac{1}{2} \sum_{\substack{p \in \mathscr{P} \\ m \geq 1}} \frac{|f(p^m) - g(p^m)|^2}{p^{\sigma m}} \log p$$

the fact that D_σ is indeed a metric is left out.

Observations of this metric:

- This is a metric on the space of arithmetic functions.
- When $m \geq 2$, the terms are summable, so the really important part is when $m = 1$.
- By the fact that f and g are both in an abs. value and the triangle inequality,

$$D_\sigma(\mathbb{1}, n^{2i\gamma_0}) = D_\sigma(n^{-i\gamma_0}, n^{i\gamma_0}) \leq D_\sigma(n^{-i\gamma_0}, \mu(n)) + D_\sigma(\mu(n), n^{i\gamma_0})$$

and by taking the conjugate we have

$$D_\sigma(\mathbb{1}, n^{2i\gamma_0}) \leq 2D_\sigma(\mu(n), n^{i\gamma_0})$$

This expression captures what is heuristically "mostly -1" and "mostly 1."

Since $|1 - p^{it}|^2 = 2 - Re(p^{-it})$ by direct computation, the LHS of the above equation can be written as

$$\begin{aligned} D_\sigma^2(\mathbb{1}, n^{2i\gamma_0}) &= \sum_{\substack{p\in\mathscr{P}\\ m\geq 1}} \frac{1 - Re(p^{-2i\gamma_0 m})}{p^{\sigma m}} \log p = \sum_{n\geq 1} \frac{\Lambda(n)}{n^\sigma} - Re\left(\sum_{n\geq 1} \frac{\Lambda(n)}{n^{\sigma+2i\gamma_0}}\right) \\ &= -\frac{\zeta'}{\zeta}(\sigma) + Re\left(\frac{\zeta'}{\zeta}(\sigma + 2i\gamma_0)\right) \end{aligned}$$

Whereas since $|\mu(p) - p^{it}|^2 = 2 + 2Re(p^{-it})$, the RHS turns into

$$D_\sigma^2(\mu(n), n^{i\gamma_0}) = \sum_{\substack{p\in\mathscr{P}\\ m=1}} \frac{1 + Re(p^{-i\gamma_0 m})}{p^{\sigma m}} \log p + O(1) = -\frac{\zeta'}{\zeta}(\sigma) - Re\left(\frac{\zeta'}{\zeta}(\sigma + i\gamma_0)\right) + O(1)$$

since we know the only diverging term is caused by $m = 1$. Therefore,

$$\begin{aligned} & D_\sigma(\mathbb{1}, n^{2i\gamma_0}) \leq 2D_\sigma(\mu(n), n^{i\gamma_0}) \\ \iff & -\frac{\zeta'}{\zeta}(\sigma) + Re\left(\frac{\zeta'}{\zeta}(\sigma + 2i\gamma_0)\right) \leq 4\left(-\frac{\zeta'}{\zeta}(\sigma) - Re\left(\frac{\zeta'}{\zeta}(\sigma + i\gamma_0)\right)\right) + O(1) \end{aligned}$$

Now we assume $\rho_0 = \beta_0 + i\gamma_0$ is a zero of ζ that is close to $1 + i\gamma_0$, where $\beta \leq 1$, $\sigma > 1$.

By Theorem 5.11 (b) we have

$$Re\left(\frac{\zeta'}{\zeta}(\sigma + 2i\gamma_0)\right) = Re\left(\sum_{|2\gamma_0 - \gamma|\leq 1} \frac{1}{\sigma + 2i\gamma_0 - \rho}\right) + O(\log|2\gamma_0|)$$

from

$$Re\left(\frac{1}{\sigma + 2i\gamma_0 - \rho}\right) = Re\left(\frac{\sigma - 2i\gamma_0 - \bar{\rho}}{|\sigma + 2i\gamma_0 - \rho|^2}\right) > 0$$

we have

$$Re\left(\frac{\zeta'}{\zeta}(\sigma + 2i\gamma_0)\right) \geq -O(\log|\gamma_0|)$$

In the same manner we have

$$\begin{aligned} -Re\left(\frac{\zeta'}{\zeta}(\sigma + i\gamma_0)\right) &= -Re\left(\sum_{|\gamma_0 - \gamma|\leq 1} \frac{1}{\sigma + i\gamma_0 - \rho}\right) + O(\log|\gamma_0|) \\ &= \sum_{|\gamma_0 - \gamma|\leq 1} \frac{-\sigma - \beta}{(\sigma - \beta)^2 + (\gamma_0 - \gamma)^2} + O(\log|\gamma_0|) \\ \text{leaving only } \rho_1 \text{ in sum} &\leq \frac{-\sigma - \beta_1}{(\sigma - \beta_0)^2 + (\gamma_0 - \gamma)^2} + O(\log|\gamma_0|) \\ &= -\frac{1}{\alpha - \beta_0} + O(\log|\gamma_0|) \end{aligned}$$

Since $\zeta(\sigma + i\gamma_0) = \frac{1}{\sigma-1} + O(\log|\gamma_0|)$ because σ is close to 1 and theorem 5.11 (b), we get:

$$-\frac{\zeta'}{\zeta}(\sigma) \leq -4\frac{\zeta'}{\zeta}(\sigma) - \frac{4}{\alpha - \beta_0} + O(\log|\gamma_0|)$$
$$\iff -\frac{3.1}{\sigma - 1} \leq -\frac{4}{\alpha - \beta_0} + O(\log|\gamma_0|)$$

Now if $\beta_0 \to 1$, RHS goes to $-\infty$ faster than LHS due to the constant, but that is a contradiction! So β_0 must stay some distance away from 1, but how far?

The last evaluation (or trick) is to take $\sigma = 1 + \frac{c}{\log|\gamma_0|}$, which will yield

$$\frac{4}{\alpha - \beta_0} \leq c\log|\gamma_0| \Rightarrow \beta_0 \geq 1 + \frac{\tilde{c}}{\log|\gamma_0|}$$

which completes the proof. □

This ends our discussion and proof on the Prime Number Theorem.

5.7. **Another use of Perron's Formula.**

Def 5.4. *A positive integer n is* ***square full*** *if* $\forall p|n,\ p^2|n$.

Theorem 5.18. $\#\{n \leq x, n \text{ is square full}\} \underset{x\to\infty}{\sim} \frac{\zeta(\frac{3}{2})}{\zeta(3)}\sqrt{x}$

Proof.

First, write out the Dirichlet Series of $\mathbb{1}_{\text{n-square full}}$:

$$F(s) = \prod_{p\in\mathscr{P}}\left(1 + \frac{1}{p^{2s}} + \frac{1}{p^{3s}} + \cdots\right) = \prod_{p\in\mathscr{P}}\left(1 + \frac{1}{p^{3s}}\right)\left(1 + \frac{1}{p^{2s}} + \frac{1}{p^{4s} + \cdots}\right)$$
$$= \prod_{p\in\mathscr{P}}\frac{1 + p^{-3s}}{1 - p^{-2s}} = \prod_{p\in\mathscr{P}}\frac{1 - p^{-6s}}{(1 - p^{-2s})(1 - p^{-3s})} = \frac{\zeta(2s)\zeta(3s)}{\zeta(6s)}$$

So the pole with the largest real part is at $s = \frac{1}{2}$, taking the residue we get the leading term due to Perron's inversion formula, which is:

$$x^{\frac{1}{2}}\frac{\zeta(3\cdot\frac{1}{2})}{\zeta(6\cdot\frac{1}{2})} = \frac{\zeta(\frac{3}{2})}{\zeta(3)}\sqrt{x}$$

(details in book; the $\frac{1}{2}$ is balanced by $\Phi(\frac{1}{2})$)

□

5.8. **Exercises.**

Exercise 5.1. *(1.10)(a)(b)*

Exercise 5.2. *(5.1)*

Exercise 5.3. *(5.3)*

Exercise 5.4. *(6.6)*

Exercise 5.5. *(5.2)*

Exercise 5.6. *(6.2)*

Exercise 5.7. *(7.7)*

Exercise 5.8. *(7.5)*

Exercise 5.9. *(8.5)(a)(b)*

Exercise 5.10. *(8.8)*

6. Dirichlet Characters

Our Final goal here in the following 3 chapters is to prove the Prime Number Theorem for arithmetic progressions. So we list the theorem here and explain a little bit.

Theorem 6.1. *(PNT for arithmetic progression) For fixed* $(q, a) = 1$,

$$\pi(x, q, a) = \#\{p \in \mathscr{P} : p \leq x, p \equiv a(mod\, q)\} \sim \frac{1}{\varphi(q)} \frac{x}{\log x}$$

What does it say? It really says that the primes are evenly distributed with the only constraints being that of modules. For example:

Example 6.1. *Theorem 6.1 tells us that for* $q = 4$,

$$\#\{p \in \mathscr{P} : p \leq x, p \equiv 1(mod 4)\} \sim \#\{p \in \mathscr{P} : p \leq x, p \equiv 3(mod 4)\} \sim \frac{1}{2} \frac{x}{\log x}$$

Intuition for this example: we can show this by showing

$$\pi(x, 4, 1) - \pi(x, 4, 3) = o\left(\frac{x}{\log x}\right)$$

where we can also write

$$\pi(x, 4, 1) - \pi(x, 4, 3) = \sum_{p \leq x} \varepsilon(p)$$

where $\varepsilon(p) = \begin{cases} 1 & p \equiv 1(mod 4) \\ -1 & p \equiv 3(mod 4) \\ 0 & \text{otherwise} \end{cases}$

It's also easy to see that ε is multiplicative. So we investigate the Euler product of its Dirichlet series and get

$$\sum_{n \geq 1} \frac{\varepsilon(n)}{n^s} = \prod_{p \in \mathscr{P}} \frac{1}{1 - \frac{\varepsilon(p)}{p^s}} \Rightarrow -\log \sum_{n \geq 1} \frac{\varepsilon(n)}{n^s} \approx \sum_{p \in \mathscr{P}} \frac{\varepsilon(p)}{p^s}$$

We'll also show that the LHS is analytic with no poles, so the equation holds. (I don't know what this all is about)

Now for mod q, we form linear combinations:

$$\sum_{p \leq x} f(p) := \sum_{a \in (\mathbb{Z}/\mathbb{Z}_q)^*} C_a \pi(x, q, a)$$

where $f(n) = \begin{cases} C_a & n \equiv a(mod\, q), (a, q) = 1 \\ 0 & otherwise \end{cases}$

and we can find C_a such that $\sum_{p \leq x} f(p) = o(\frac{x}{\log x})$. For that we would want $\begin{cases} \sum C_a = 0 \\ f\text{-multiplicative} \end{cases}$

Anyways, all above are just some intuitions. We'll start from the basics now.

6.1. **Heuristics.**

Def 6.1. $\chi : \mathbb{Z} \to \mathbb{C}$ *is a* ***Dirichlet Character mod q*** *if* $\begin{cases} \chi \text{ is } q\text{-periodic} \\ \chi(m) \neq 0 \iff (m,q) = 1 \\ \chi \text{ is completely multiplicative over } \mathbb{Z} \end{cases}$

Remark 6.1. *χ must have modules 1 since otherwise we can use completely multiplicative to make one value of χ goes to infinity, which is contradictory to the periodicity.*

For given χ we associate $\tilde{\chi}$ be the group of homomorphism $\tilde{\chi} : (\mathbb{Z}/\mathbb{Z}_q)^* \to \mathbb{C}$

Define the space of functions with $f : (\mathbb{Z}/\mathbb{Z}_q)^* \to \mathbb{C}$ equipped with the inner product

$$\langle f, g \rangle = \frac{1}{\varphi(q)} \sum_{a \in (\mathbb{Z}/\mathbb{Z}_q)^*} f(a)\bar{g}(a)$$

A fact that we will prove: there are $\varphi(q)$ Dirichlet characters mod q and they form an orthonormal basis for $\langle\,,\rangle$. With previous notations we choose $C_a = \chi(a)$ for different characters χ, i.e.

$$\sum_{a \in (\mathbb{Z}/\mathbb{Z}_q)^*} \chi(s)\pi(x,q,a) = \sum_{p \leq x} \chi(p)$$

by an inversion with respect to the basis (we'll cover this later),

$$\pi(x,q,a) = \frac{1}{\varphi(a)} \sum_{\chi - mod\ q} \bar{\chi}(a) \sum_{\substack{p \leq x \\ p \in \mathscr{P}}} \chi(p)$$

Def 6.2. *Let the character constant to 1 on $(\mathbb{Z}/\mathbb{Z}_q)^*$ be called* ***the principal character****, denoted χ_0, i.e.* $\chi_0 = \begin{cases} 1 & (n,q) = 1 \\ 0 & \text{otherwise} \end{cases}$

We note that in the expression of $\pi(x,q,a)$ represented above the contribution from χ_0 is

$$\frac{1}{\varphi(q)} \sum_{\substack{p \leq x \\ p \in \mathscr{P}}} \chi_0(p) = \frac{\pi(x) + O(\log q)}{\varphi(q)} = \frac{1}{\varphi(q)} \frac{x}{\log x} + o(\frac{x}{\log x})$$

since the only $\neq 1$ terms in the summand are the divisors of q, and q has at most $\dfrac{\log q}{\log 2}$ divisors.

Let's look back to our example of $q = 4$. the only characters are in fact $\varepsilon(p)$ and χ_0.

We introduce $L(s,\chi) = \sum_{n \geq 1} \dfrac{\chi(n)}{n^s} = \prod_{p \in \mathscr{P}} = \dfrac{1}{1 - \frac{\chi(p)}{p^s}}$.

By the Perron Inversion Formula, we know that $\sum_{x \leq p} \chi(p)$ will be related to the analytic properties of $L(s,\chi)$.

Some of the results we'll eventually prove are:

Let's assume $q \in \mathscr{P}$ to simplify. Let χ be a character mod q, $\chi \neq \chi_0$, then

$$\begin{cases} L(s,\chi) \text{ has an analytic continuation to the full}\mathbb{C} \text{ with no pole at 1} \\ \text{there's no zero in the same zero-free region of } \zeta \end{cases}$$

A comment on RH:

The result of Perron's formula & the contour shift will be

$$\sum_{n\leq x} \chi(n)\Lambda(n) = -\sum_{|\gamma|\leq T} \frac{x^\rho - 1}{\rho} + O\left(\frac{x\log^2 x}{T}\right)$$

So if the RH holds for $L(s,\chi)$ rather than ζ, we will have

$$\pi(x,q,a) = \frac{1}{\varphi(q)}\frac{x}{\log x} + O(x^{\frac{1}{2}+\epsilon})$$

6.2. **Fourier Analysis on Finite Abelian Groups.**

Def 6.3. *For G-finite abelian, $\chi : G \to \mathbb{C}^*$ is a character of G if it is a group homomorphism. Let $\widehat{G}$ denote the set of characters.*

Note:

1. Still $\chi_0 \equiv 1$ is the principal character.
2. $\widehat{G}$ is a subset of functions on G, so it's in $L^2(G)$.

The inner product is defined as:

$$\langle \alpha, \beta \rangle_G = \frac{1}{|G|}\sum_{g\in G} \alpha(g)\bar{\beta}(g)$$

Lemma 6.2. *Let G_1, G_2 be abelian groups with $G_1 \times G_2 = G$. Then the function $\phi : \widehat{G_1} \times \widehat{G_2} \to \widehat{G}$ associating the pair (χ_1, χ_2) to the character $\chi(g_1, g_2) = \chi_1(g_1)\chi_2(g_2)$ is a group isomorphism.*

Theorem 6.3. *$\widehat{G}$ is an orthonormal basis of $L_2(G)$.*

Proof.

Step 1: prove that $\widehat{G}$ is orthonormal.

We know by Lagrange theorem that $\forall g \in G$, $g^{|G|} = 1$ since $|G| < \infty$, denote $n = |G|$. Then since χ is a group homomorphism,

$$\chi(g)^n = \chi(g^n) = \chi(1) = 1$$

which implies that $\chi(g)$ is an n-th root of unity, which further gives

$$\chi(g) \in \{1, e^{i\frac{2\pi}{n}}, e^{i\frac{4\pi}{n}}, \cdots, e^{i\frac{2\pi(n-1)}{n}}\}$$

so $\widehat{G}$ is finite.

We have that $\forall \chi, \phi \in \widehat{G}$, since $hG = G$

$$\begin{aligned}\langle \chi, \phi \rangle_G &= \frac{1}{|G|} \sum_{g \in G} (\chi \bar{\phi})(g) = \frac{1}{|G|} \sum_{g \in G} (\chi \bar{\phi})(gh) \\ &= (\chi \bar{\phi})(h) \frac{1}{|G|} \sum_{g \in G} (\chi \bar{\phi})(g)\end{aligned}$$

If $\chi \neq \phi$, then there exist some h such that $\chi(h) \neq \phi(h) \Rightarrow \chi \bar{\phi} \neq 1 \Rightarrow \langle \chi, \phi \rangle_G = 0$. On the other hand, direct computation tells us $\langle \chi, \chi \rangle_G = 1$.

Therefore,$\widehat{G}$ is an orthonormal set in $L_2(G)$.

Step 2: show that $|\widehat{G}| = |G|$

This is obvious when G is cyclic since in that case each n-th root of unity gives rise to one character and there's no other characters since the generator must be mapped to some root of unity.

If G is not cyclic, then it must be direct sums of cyclic groups, Lemma 6.2 helps us finish the proof.

Since the dimention of $L_2 G$ is $|G|$ and $|\widehat{G}|$ is a set of $|G|$ orthonormal functions, $|\widehat{G}|$ is actually the complete basis of $L_2(G)$.

□

Since we have a basis now, we can decompose any function $f : G \to \mathbb{C}$ into a linear combination of the basis. Let

$$\hat{f}(\chi) = \langle f, \chi \rangle_G$$

then

$$f = \sum_{\chi \in \widehat{G}} \hat{f}(\chi) \chi$$

is the Fourier transform in $L_2(G)$.

In particular, for $f(g) = \mathbb{1}_{g=h}$ where $g \in G$, $\hat{f}(\chi) = \langle f, \chi \rangle_G = \frac{1}{|G|} \bar{\chi}(h)$, and thus

$$\mathbb{1}_{g=h}(g) = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \chi(g) \bar{\chi}(h)$$

which implies that $\chi \in \widehat{G}$ are also orthonormal.

Theorem 6.4. *(Parseval's equality)*

$$\int |f|^2 = \int |\hat{f}|^2 \quad or \quad \frac{1}{|G|} \sum_{g \in G} |f(g)|^2 = \sum_{\chi \in \widehat{G}} |\hat{f}(\chi)|^2$$

Proof.

$$\sum_{\chi\in\widehat{G}} |\hat{f}(\chi)|^2 = \frac{1}{|G|^2}\sum_{\chi\in\widehat{G}}\left|\sum_{g\in G} f(g)\bar{\chi}(g)\right|^2$$

$$\text{(since } |z|^2 = z\bar{z}) \quad = \frac{1}{|G|^2}\sum_{g_1,g_2\in G} f(g_1)\bar{f}(g_2)\sum_{\chi\in\widehat{G}}\bar{\chi}(g_1)\chi(g_2)$$

$$= \frac{1}{|G|^2}\sum_{g_1\in G}|f(g_1)|^2 = \frac{1}{|G|^2}\sum_{g\in G}|f(g)|^2$$

□

Now let's specify some G and see what we get. In particular, we deal with

$$G = (\mathbb{Z}/q\mathbb{Z},+) \quad \text{and} \quad G' = ((\mathbb{Z}/\mathbb{Z}_q)^*,\cdot)$$

$G = (\mathbb{Z}/q\mathbb{Z},+)$:

G is cyclic, so we define $e(x) = e^{i2\pi x}$ and where χ is of type $\chi(n) = e(\frac{an}{q})$. In fact, it is enough to specify $\chi(1) = e(\frac{a}{q})$ since its a generator.

Therefore, instead of writing $\hat{f}(\chi)$, we can write

$$\hat{f}(a) = \frac{1}{q}\sum_{n\in\mathbb{Z}/q\mathbb{Z}} f(n)e(\frac{-an}{q})$$

where $\hat{f}$ can be thought of as a q-periodic function on $\mathbb{Z}$. We will use this result in the second case.

$G' = ((\mathbb{Z}/\mathbb{Z}_q)^*,\cdot)$:

Def 6.4. *For $\chi\in\widehat{G'}$, let*

$$\mathcal{G}(\chi) := q\widehat{\chi}(-1) = \sum_{1\le n\le q}\chi(n)e(\frac{n}{q})$$

Theorem 6.5. *If $(a,q)=1$, then $\mathcal{G}(\chi)\cdot\bar{\chi}(a) = \sum_{1\le n\le q}\chi(n)e(\frac{an}{q})$*

Remark: we will later see that the condition $(a,q) = 1$ can be taken out for primitive characters, after we've defined it.

Proof. (of theorem 6.5)

$(a,q) = 1$ means $f(g) = ag$ for $g\in(\mathbb{Z}/\mathbb{Z}_q)^*$ is a bijection. So let $m\equiv na(mod\,q)$ we get

$$\sum_{1\le n\le q}\chi(n)e(\frac{an}{q}) = \sum_{1\le m\le q}\chi(\frac{m}{a})e(\frac{m}{q}) = \chi(a^{-1})\sum_{1\le m\le q}\chi(m)e(\frac{m}{q}) = \mathcal{G}(\chi)\bar{\chi}(a)$$

□

Def 6.5. *For $q|m$, ξ-mod q and χ-mod m are multiplicative characters, then χ is **a lift** of ξ (or χ is **induced by** ξ) if*

$$\chi(n) = \mathbb{1}_{(m,n)=1}\xi(n')$$

for $n \equiv n'(mod\, q)$.

Def 6.6. *χ is a **primitive character** is there is no ξ such that χ is a lift of ξ.*

For instance, principal characters are never primitive (or imprimitive) since for χ_0 mod q, we always have

$$\chi_0(n) = 1 = \mathbb{1}_{(n,q)=1}\xi_0$$

for ξ_0 be the primitive (in fact the only) character mod 1.

We'll now prove a generalized version of Theorem 6.5 for primitive characters.

Theorem 6.6. *For primitive character χ,*

$$\mathcal{G}(\chi) \cdot \bar{\chi}(a) = \sum_{1\le n\le q} \chi(n)e(\frac{an}{q})$$

or equivalently

$$\chi(n) = \frac{1}{\mathcal{G}(\bar{\chi})} \sum_{1\le a\le q} \bar{\chi}(a)e(\frac{an}{q})$$

To prove theorem 6.6, we first prove three lemmas.

Def 6.7. *d is called a **period** of χ-(mod q) if*

$$\begin{cases} m \equiv n(mod\, d) \\ (mn, q) = 1 \end{cases} \Rightarrow \chi(m) = \chi(n)$$

Lemma 6.7. *If $(a, d) = 1$, $d|q$, then $\exists$ k such that $(a + kd, q) = 1$*

Proof. We construct $k = \prod_{p|\frac{q}{d}, p\nmid a} p$. We want to show that $(a + kd, q) = 1$, so we show that $\forall p|q, p \nmid a + kd$.

Since $d|q$, either $p|d$ or not. By enumeration we have

- If $p|d$, then $p \nmid a$ since $(a, d) = 1$ and thus $p \nmid a + kd$.
- If $p \nmid d$, then $p|\frac{q}{d}$. In this case either $p|a$ or not. If $p|a$, then by the construction of k we know that $p \nmid k$, therefore $p \nmid a + kq$; if $p \nmid a$, then by the construction of k we know that $p|k$, therefore $p \nmid a + kq$

We are done. □

Lemma 6.8. *χ is a character mod q, then*

$$\chi \text{ is imprimitive} \iff \exists d|q \text{ such that } d \text{ is a period of } \chi$$

.

Proof.

$\Rightarrow$:

This direction is easy. Let ξ mod d be the character inducing χ. Since

$$(mn, q) = 1 \Rightarrow \mathbb{1}_{(m,q)=1}(m) = 1 \quad \& \quad \mathbb{1}_{(n,q)=1}(n) = 1$$

and

$$m \equiv n(mod\, d) \Rightarrow \xi(m) = \xi(n)$$

we have

$$\chi(m) = \xi(m)\mathbb{1}_{(m,q)=1}(m) = \xi(n)\mathbb{1}_{(m,q)=1}(n) = \chi(n)$$

which is what we want.

$\Leftarrow$:

Define $\phi : \mathbb{Z} \to \mathbb{C}$ as follows: $\begin{cases} \phi(a) = 0 & (a, d) = 1 \\ \phi(a) = \chi(a + kd) & (a, d) > 1 \end{cases}$

where $(a + kd, q) = 1$. The existence of such k is by Lemma 6.7. Also, the value of $\phi(a)$ will not depend on the choice of k due to our premise. So ϕ is well-defined.

But then ϕ induces χ if ϕ is a character. By the above definition and explanation we know that ϕ is q-periodic and $\phi(m) \neq 0 \iff (m, d) = 1$, so we only need to prove that it is completely multiplicative: for $(a, d) \neq 1$,

$$\phi(a)\phi(a') = \chi(a + kd)\chi(a' + k'd) = \chi(aa' + d \cdot (ak' + a'k + d))$$
$$= \chi(aa' + cd) = \chi(aa') = \chi(aa' + k''d) = \phi(aa')$$

for constants chosen according to Lemma 6.7.

□

Lemma 6.9. *For χ mod q,*

$$d \text{ is a period of } \chi \iff \left(\begin{cases} (n, q) = 1 \\ n \equiv 1(mod\, d) \end{cases} \Rightarrow \chi(n) = 1 \right)$$

Proof. $\Rightarrow$:

By definition of periodicity, $\chi(n) = \chi(1) = 1$ since χ multiplicative.

$\Leftarrow$:

Assume $(mn, q) = 1$ and $m = n(mod\, d)$ where $d|q$, write $m = ld + n$. Since χ is multiplicative and $\chi(kd + 1) = 1$ by assumption, we have $\chi(m) = \chi((ld + n)(kd + 1))$, if we can show that $\exists k$ such that $(ld + n)(kd + 1) \equiv n(mod\, q)$, we're done.

But

$$(ld + n)(kd + 1) = n(mod\, q) \iff lkd^2 + ld + knd = 0(mod\, q)$$
$$\iff q|lkd^2 + ld + knd \iff \frac{q}{d}|k(ld + n) + l \iff \frac{q}{d}|km + l$$

since $(m, \frac{q}{d}) = 1$, for any l, $\exists k$ such that $km = l$, and we are done. □

Together with Lemma 6.8, we have

$$\chi \text{ is imprimitive } \iff \left(\begin{cases} (n,q) = 1 \\ n \equiv 1 (mod\, d) \end{cases} \Rightarrow \chi(n) = 1 \right)$$

Proof. (of Theorem 6.6):

We note that in the case of $(n, q) = 1$, by Theorem 6.5 we are done. Hence we only focus on the case when $(n, q) \neq 1$, but then $\chi(n) = 1$ by definition of a character. So we only need to prove $\sum_{1 \leq a \leq q} \bar{\chi}(a) e(\frac{an}{q}) = 0$.

In this case we can write $n = ml$ and $q = md$ for $m > 1$ such that $(d, l) = 1$. Then we have

$$\sum_{1 \leq a \leq q} \bar{\chi}(a) e\left(\frac{an}{q}\right) = \sum_{\substack{b+dj \\ 1 \leq j \leq m \\ 1 \leq b \leq d}} \bar{\chi}(b+dj) e\left(\frac{(b+dj)ml}{md}\right) = \sum_{1 \leq b \leq d} e\left(\frac{bl}{d}\right) \sum_{1 \leq j \leq m} \bar{\chi}(b+dj)$$

Since χ is primitive, by Lemma 6.8 and Lemma 6.9 we know that $\exists j_0 \in \mathbb{Z}$ such that $r := 1 + j_0 d$ satisfies $\begin{cases} (r,q) = 1 \\ \chi(r) \neq 1 \end{cases}$

Since $(r, q) = 1$, $f(g) = rg$ is a group permutation on $(\mathbb{Z}/\mathbb{Z}_q)^*$, which means the image of $f(b + dj)$ for different j is different under modules q. But $f(b + dj) = b (mod\, d)$, so f is also a group permutation of $\{b + dj : 1 \leq j \leq m\} (mod\, q)$ i.e.

$$\{r(b+dj) : 1 \leq j \leq m\} = \{b + dj : 1 \leq j \leq m\} (mod\, q)$$

which implies

$$\sum_{1 \leq j \leq m} \chi(b+dj) = \sum_{1 \leq j \leq m} \chi(r(b+dj)) = \chi(r) \sum_{1 \leq j \leq m} \chi(b+dj)$$

since $\chi(r) \neq 1$,

$$\sum_{1 \leq j \leq m} \chi(b+dj) = 0$$

which ends the proof.

□

A consequence is

Corollary 6.10. $|\mathcal{G}(\chi)| = \sqrt{q}$

Proof. (Need proofread)The Fourier transform of χ is

$$\hat{\chi}(n) = \frac{1}{q} \sum_{1 \leq n \leq q} \chi(n) e(-\frac{an}{q})$$

applying the finite Parseval identity with $f = \chi$ we get the equation

$$\sum_{1\le n\le q} |\hat{\chi}(n)|^2 = \frac{1}{q}\sum_{1\le n\le q} |\chi(n)|^2 = \frac{1}{q}\cdot\varphi(q)$$

But note

$$|\hat{\chi}(n)| = \left|\overline{\hat{\chi}(n)}\right| = \left|\frac{1}{q}\sum_{1\le n\le q}\bar{\chi}(n)e\left(\frac{an}{q}\right)\right| = |\chi(-1)|\left|\frac{1}{q}\sum_{1\le n\le q}\chi(n)e\left(\frac{an}{q}\right)\right| = \frac{1}{q}|\mathcal{G}(\chi)\bar{\chi}(n)|$$

so

$$\varphi(q)\left(\frac{1}{q}|\mathcal{G}(\chi)|\right)^2 = \sum_{1\le n\le q}\left(\frac{1}{q}|\mathcal{G}(\chi)\bar{\chi}(n)|\right)^2 = \frac{1}{q}\cdot\varphi(q)$$
$$\Rightarrow |\mathcal{G}(\chi)|^2 = q \iff |\mathcal{G}(\chi)| = \sqrt{q}$$

□

6.3. **The Twisted Poisson Summation Formula.**

The main goal here is to do an analogue of the Poisson Summation formula so that we can eventually get the functional equation of $L(s,\chi)$, and hence getting some bound on $L(s,\chi)$.

Theorem 6.11. *Let χ be a primitive character mod q, $N \in \mathbb{N}^*$ and $f \in C^2$ be such that $f'(x), f''(x), f'''(x) \ll \frac{1}{x^2}$. From Fourier analysis we know that this implies $|\hat{f}(\xi)| = O(\frac{1}{\xi^2})$. We have*

$$\sum_{n\in\mathbb{Z}}\chi(n)f(\frac{n}{N}) = \frac{\chi(-1)N}{\mathcal{G}(\bar{\chi})}\sum_{n\in\mathbb{Z}}\bar{\chi}(n)\hat{f}(\frac{nN}{q})$$

Proof.

WLOG, we can assume $N = 1$ since by a change of variables we know from Fourier Analysis that $f(\frac{x}{l}) = l\hat{f}(xl)$. Using theorem 6.6, we have

$$\begin{aligned}\sum_{n\in\mathbb{Z}}\chi(n)f(n) &= \frac{1}{\mathcal{G}(\bar{\chi})}\sum_{n\in\mathbb{Z}}f(n)\sum_{1\le a\le q}\bar{\chi}(a)e(\frac{an}{q})\\ &= \frac{1}{\mathcal{G}(\bar{\chi})}\sum_{1\le a\le q}\bar{\chi}(a)\sum_{n\in\mathbb{Z}}f(n)e(\frac{an}{q})\end{aligned}$$

Take $g(x) = f(x)e(\frac{ax}{q})$, then $\hat{g}(\xi) = \hat{f}(\xi - \frac{a}{q})$, and by Poisson summation on g we get

$$\begin{aligned}\sum_{n\in\mathbb{Z}}\chi(n)f(n) &= \frac{1}{\mathcal{G}(\bar{\chi})}\sum_{1\le a\le q}\bar{\chi}(a)\sum_{\xi\in\mathbb{Z}}\hat{f}(\xi-\frac{a}{q}) = \frac{\bar{\chi}(-1)}{\mathcal{G}(\bar{\chi})}\sum_{1\le a\le q}\bar{\chi}(a)\sum_{\xi\in\mathbb{Z}}\hat{f}(\xi-\frac{a}{q})\\ &= \frac{\bar{\chi}(-1)}{\mathcal{G}(\bar{\chi})}\sum_{\substack{1\le a\le q\\ n\in\mathbb{Z}}}\bar{\chi}(nq-a)\hat{f}(\frac{nq-a}{q}) = \frac{\bar{\chi}(-1)}{\mathcal{G}(\bar{\chi})}\sum_{n\in\mathbb{Z}}\bar{\chi}(n)\hat{f}(\frac{n}{q})\end{aligned}$$

□

Remark 6.2.

We can bound the RHS of theorem 6.11 to understand the diffusive size of LHS, i.e. since

$$\sum_{n\in\mathbb{Z}}\left|\hat{f}(\frac{nN}{q})\right| \ll \sum_{|n|\leq\frac{q}{N}} 1 + \sum_{|n|\geq\frac{q}{N}} \frac{1}{(\frac{nN}{q})^2} = O(\frac{q}{N})$$

$$\sum_{n\in\mathbb{Z}} \chi(n) f(\frac{n}{N}) \ll \frac{N}{\sqrt{q}} \cdot O(\frac{q}{N}) = O(\sqrt{q})$$

by corollary 6.10 and thm 6.11.

We've seen what the estimate looks like when f is moderate decreasing, but in the case when $f = \mathbb{1}_I$ for some domain I we cannot apply the same method due to the fact that $\hat{f}$ would then be very bad for integration. But in that case the transform's error is about $\int_a^\infty \frac{1}{x}dx$, which is "barely infinite." This mean's there's still hope. Indeed, we offer one way to estimate it now.

Theorem 6.12. *(Polya-Vinogradov inequality) For χ- non-principal character mod q and $\forall M \in \mathbb{R}$, $N > 0$, we have*

$$\sum_{M\leq n\leq M+N} \chi(n) \ll \sqrt{q}\log q$$

Proof.

We prove the theorem by first proving for χ primitive, then extend to all characters.

For primitive characters:

By theorem 6.6

$$\chi(n) = \frac{1}{\mathcal{G}(\bar{\chi})} \sum_{-\frac{q}{2}\leq a\leq\frac{q}{2}} \bar{\chi}(a) e(\frac{an}{q})$$

which means

$$\sum_{M\leq n\leq M+N} \chi(n) = \frac{1}{\mathcal{G}(\bar{\chi})} \sum_{-\frac{q}{2}\leq a\leq\frac{q}{2}} \bar{\chi}(a) \sum_{M\leq n\leq M+N} e(\frac{an}{q})$$

Hence we only need to bound $\left|\sum_{M\leq n\leq M+N} e(\frac{an}{q})\right|$, in which form we see that the value of M does not matter at all:

$$\left|\sum_{M\leq n\leq M+N} e(\frac{an}{q})\right| \leq \left|\frac{1-e(\frac{aN}{q})}{1-e(\frac{a}{q})}\right| \cdot |1|$$

Notice that $|1-e(x)| = |e(-\frac{x}{2}) - e(\frac{x}{2})| = 2|\sin(\pi x)| \geq 4|x|$ and the nominator is less than 2, so $\left|\sum_{M\leq n\leq M+N} e(\frac{an}{q})\right| \leq \frac{1}{2|\frac{a}{q}|}$, which yields

$$\sum_{M\leq n\leq M+N} \chi(n) \ll \frac{1}{\sqrt{q}} \sum_{-\frac{q}{2}\leq a\leq \frac{q}{2}} \frac{q}{2a} \ll \sqrt{q}\log q$$

For non-primitive characters:

Assume χ is induced by ϕ where ϕ is primitive character mod d. Then

$$\chi(n) = \phi(n)\mathbb{1}_{(n,q)=1} = \phi(n)\mathbb{1}_{(n,\frac{q}{d})=1}$$

since the only difference between $\mathbb{1}_{(n,q)=1}$ and $\mathbb{1}_{(n,\frac{q}{d})=1}$ is when $(n,d) \neq 1$, but in that case $\phi(n) = 0$.

Thus

$$\sum_{M\leq n\leq M+N} \chi(n) = \sum_{\substack{M\leq n\leq M+N \\ (n,\frac{q}{d})=1}} \phi(n) = \sum_{M\leq n\leq M+N} \phi(n)\cdot \sum_{a|(n,\frac{q}{d})} \mu(a) = \sum_{a|\frac{q}{d}} \mu(a) \sum_{\substack{M\leq n\leq M+N \\ a|n}} \phi(n)$$

We write $n = ma \Rightarrow \phi(n) = \phi(m)\phi(a)$ and so since ϕ primitive

$$\begin{aligned}\sum_{M\leq n\leq M+N} \chi(n) &\ll \sum_{a|\frac{q}{d}} 1 \left|\sum_{\frac{M}{a}\leq n\leq \frac{M+N}{a}} \phi(m)\cdot 1\right| \\ &\ll \sum_{a|\frac{q}{d}} \sqrt{d}\log d \ll \log q\left(\sum_{a|\frac{q}{d}} \sqrt{d}\right) \\ &\ll \log q\sqrt{\frac{q}{d}}\sqrt{d} = \sqrt{q}\log q\end{aligned}$$

where the last step is due to the fact that at least one of a and $\frac{q}{da}$ is smaller than $\sqrt{\frac{q}{d}}$. □

6.4. **Exercises.**

Example 6.2. *(9.1)*

Example 6.3. *(9.3)*

Example 6.4. *(10.1)*

Example 6.5. *(9.6)*

Example 6.6. *(10.3)*

Example 6.7. *(10.4)*

7. Dirichlet L-function

In this chapter we study some basic properties of the Dirichlet L-function $L(s,\chi)$, it's analytic continuation, estimate, and explain the formulas.

Remember that DIrichlet functions are defined as $L(s,\chi)=\sum_{n\geq 1}\frac{\chi(n)}{n^s}$, so it is analytic on $Re(s)>1$. On this domain, due to the multiplicity of χ, we also have

$$L(s,\chi)=\prod_{p\in\mathcal{P}}\left(1+\frac{\chi(p)}{p^s}+\frac{\chi(p^2)}{p^{2s}}+\cdots\right)=\prod_{p\in\mathcal{P}}\frac{1}{1-\frac{\chi(p)}{p^s}}$$

Remark 7.1.

We will only work with primitive χ since that will be sufficient in practice because if χ is induced by primitive character ϕ, i.e. $\chi(n)=\phi(n)\cdot\mathbb{1}_{(n,q)=1}$, then

$$L(s,\chi)=\prod_{p\in\mathcal{P}}\frac{1}{1-\frac{\chi(p)}{p^s}}=\prod_{p|q}\frac{1}{1-\frac{\chi(p)}{p^s}}\prod_{p\nmid q}\frac{1}{1-\frac{\phi(p)}{p^s}}$$

but if $p|q$, $\chi(p)=0$ so we have

$$L(s,\chi)=\prod_{p\nmid q}\frac{1}{1-\frac{\phi(p)}{p^s}}=\prod_{p\in\mathcal{P}}\frac{1}{1-\frac{\phi(p)}{p^s}}\prod_{p|q}\left(1-\frac{\phi(p)}{p^s}\right)=L(s,\phi)\prod_{p|q}\left(1-\frac{\phi(p)}{p^s}\right)$$

where the product is finite. So it really is enough only to evaluate properties of ϕ.

If $\prod_{p|q}\left(1-\frac{\phi(p)}{p^s}\right)=0$, then

$$\frac{\phi(p)}{p^s}=1\Rightarrow\frac{|\phi(p)|}{|p^s|}=1\Rightarrow|p^s|=1\Rightarrow Re(s)=0$$

More explicitly, the only extra zeroes that χ has and ϕ does not are on the imaginary axis.

The Extended Riemann Hypothesis: For any χ, we have

$$\begin{cases}L(s,\chi)=0\\0<Re(s)<1\end{cases}\Rightarrow Re(s)=\frac{1}{2}$$

As you might noticed, all above are based on the fact that $L(s,\chi)$ has an analytic continuation to $\mathbb{C}$, so we'll prove that in this chapter. In particular, we'll prove a functional equation with a "twisted symmetry" similar to what we've done in section 5.2.

7.1. Dirichlet L functions at 1.

One very useful heuristic here is that $L(1,\chi)\neq 0$, which also implies that $L(1,\chi)$ is entire. So we'll prove it and see why it is important.

Theorem 7.1. *$L(1,\chi)$ converges and $L(1,\chi) \gg \dfrac{1}{\sqrt{q}\log^2 q}$.*

Proof.

Convergence: (try direct computation)

$$L(1,\chi) = \sum_{k\geq 1}\left|\sum_{kq<n\leq(k+1)q}\frac{\chi(n)}{n}\right|$$

where since (by basic ring theory)

$$\sum_{kq<n\leq(k+1)q}\chi(n) = 0$$

we have

$$\sum_{kq<n\leq(k+1)q}\frac{\chi(n)}{n} = \sum_{kq<n\leq(k+1)q}\frac{\chi(n)}{n} - \frac{1}{kq}\sum_{kq<n\leq(k+1)q}\chi(n) = \sum_{kq<n\leq(k+1)q}\chi(n)\left(\frac{1}{n}-\frac{1}{kq}\right)$$

note that $kq < n \leq (k+1)q$ implies $\left(\dfrac{1}{n}-\dfrac{1}{kq}\right) = O\left(\dfrac{1}{n^2}\right)$ is absolutely summable in n, so the whole thing turns into

$$L(x,\chi) = \sum_{k\geq 1}\left|\sum_{kq<n\leq(k+1)q}\chi(n)O\left(\frac{1}{n^2}\right)\right| = \sum_{k\geq 1}O\left(\sum_{kq<n\leq(k+1)q}\frac{1}{n^2}\right) = O\left(\sum_{n\geq 1}\frac{1}{n^2}\right) = O(1)$$

converges.

Non-zero:

We show this by showing that $L(s,\chi)$ is a limit of positive numbers that's bounded below. More explicitly, $\displaystyle\sum_{a\leq x}\frac{\chi(a)}{a} \to L(s,\chi)$.

We will start from the equation

$$\frac{1}{x}\sum_{a\leq x}\left\lfloor\frac{x}{a}\right\rfloor\chi(a)$$

We can sort of see what we did is to multiply $\dfrac{x}{x}$ to $\displaystyle\sum_{a\leq x}\frac{\chi(a)}{a}$ and take integer value. Of course this step induces an error, but let's deal with it in the end.

By the Dirichlet hyperbolic method, we know $\displaystyle\sum_{n\leq x}(f*g)(n) = \sum_{ab\leq x}f(a)g(b)$, so

$$\sum_{a\leq x}\left\lfloor\frac{x}{a}\right\rfloor\chi(a) = \sum_{n\leq x}(\mathbb{1}*\chi)(n)$$

since each $\chi(a)$ is added $\left\lfloor\frac{x}{a}\right\rfloor$ times (there are that many multiples of a).

In Theorem 3.2(a) we have shown that if f, g are multiplicative, then so is $f * g$. So this gives us some clue to first consider $\mathbb{1}\chi$ at prime powers.

$$(\mathbb{1} * \chi)(p^m) = \begin{cases} m+1 & \chi(p) = 1 \\ \mathbb{1}_{2|m} & \chi(p) = -1 \\ 1 & \chi(p) = 0 \end{cases}$$

since

$$(\mathbb{1} * \chi)(p^m) = \sum_{d|p^m} \chi(d) = \sum_{0 \le k \le m} \chi(p^k)$$

and

- if $\chi(p) = 0$, then we only count when $k = 0$;
- if $\chi(p) = 1$, then all terms are 1;
- if $\chi(p) = -1$, then the terms are $1, -1, 1, -1, \ldots$.

So only two obstacles are left:

(1) Since there is chance that $(\mathbb{1} * \chi)(p^m) = 0$, when such p^m occurs very often, there seems to be the danger of the sum is 0.
(2) We still haven't shown the validity of the very first expression of this process.

As for (2), the idea is to smoothen the $\lfloor x \rfloor$ function since the error $x - \lfloor x \rfloor \in (0, 1)$ is too big. This is done by

$$T := \sum_{n \le x} (\mathbb{1} * \chi)(n) \left(1 - \frac{n}{x}\right) = \sum_{a \le x} \chi(a) \sum_{b \le \frac{x}{a}} \left(1 - \frac{b}{\frac{x}{a}}\right)$$

the rest is by Euler-Maclaurin and Polya-Vinogradov. (Both two obstacles remain obstacles, but it is not in the test, so will fill in the holes later.)

□

Theorem 7.2. *For any χ-mod q, $(l, q) = 1$, there are infinitely many primes $\equiv l \pmod q$.*

Remark 7.2.

We set out to prove that in each partition of modules mod q there's $\left(\frac{1}{\varphi(q)} \frac{x}{\log x}\right)$ many primes. Theorem 7.2 already proves that each partition has infinite members. Even though it's not our target, it already offers great insight, and is by itself very amazingly simple.

Proof.

By Theorem 7.1 $L(s, \chi) \neq 0$.

By Theorem 6.3 we know that χ mod q are orthogonal and thus

$$\mathbb{1}_{l\equiv m(mod\, q)} = \frac{1}{\varphi(q)} \sum_{\chi mod\, q} \bar{\chi}(l)\chi(m)$$

$$\underset{m=p}{\Rightarrow} \sum_{\substack{p\equiv l(mod\, q)\\(l,q)=1}} \frac{1}{p} = \frac{1}{\varphi(q)} \sum_{\chi mod\, q} \bar{\chi}(l) \sum_{\substack{p=l(mod\, q)\\(l,q)=1}} \frac{\chi(p)}{p}$$

where we are simply summing up $\frac{1}{p}$ for $p \equiv l(mod\, q)$, and letting m=p. Moving on, we pick out the principal character $\chi_0$and get

$$\frac{1}{\varphi(q)} \sum_{\chi mod\, q} \bar{\chi}(l) \sum_{\substack{p=l(mod\, q)\\(l,q)=1}} \frac{\chi(p)}{p} = \frac{1}{\varphi(q)} \sum_{p\nmid q} \frac{1}{p} + \frac{1}{\varphi(q)} \sum_{\chi\neq\chi_0} \bar{\chi}(l) \sum_{\substack{p=l(mod\, q)\\(l,q)=1}} \frac{\chi(p)}{p}$$

We want to show that for $\chi \neq \chi_0$ the sum is just $O(1)$, which agrees with our earlier claim that most of the contribution is from the principal character. For $\chi \neq \chi_0$:

$$L(1,\chi) \neq 0 \iff \prod_{p\in\mathscr{P}} \frac{1}{1-\frac{\chi(p)}{p}} \neq 0 \Rightarrow \left| \sum_{p\in\mathscr{P}} \log\left(1-\frac{\chi(p)}{p}\right) \right| < \infty$$

and by Taylor

$$\left| \sum_{p\in\mathscr{P}} \frac{\chi(p)}{p} \right| < \infty$$

so the error term (the first sum is just finite sum)

$$\frac{1}{\varphi(q)} \sum_{\chi\neq\chi_0} \bar{\chi}(l) \sum_{\substack{p=l(mod\, q)\\(l,q)=1}} \frac{\chi(p)}{p} = O(1)$$

(Question: how do we go from $\left| \sum_{p\in\mathscr{P}} \frac{\chi(p)}{p} \right| < \infty$ to $\left| \sum_{\substack{p=l(mod\, q)\\(l,q)=1}} \frac{\chi(p)}{p} \right| < \infty$?)

Combining what we have up there,

$$\sum_{\substack{p=l(mod\, q)\\(l,q)=1}} \frac{1}{p} = \frac{1}{\varphi(q)} \sum_{p\nmid q} \frac{1}{q} + O(1) \Rightarrow \sum_{\substack{p=l(mod\, q)\\(l,q)=1}} \frac{1}{p} \to \infty$$

which means that there are infinitely many primes $\equiv l$ (mod q). □

7.2. **Functional equation and analytic continuation of Dirichlet L function.**

As we've discussed earlier, we only consider χ primitive in this section.

Def 7.1. $\begin{cases} \textit{If } \chi(-1) = 1, \textit{ then } \chi \textit{ is } \boldsymbol{even}; \\ \textit{If } \chi(-1) = -1, \textit{ then } \chi \textit{ is } \boldsymbol{odd}. \end{cases}$

Note that $\chi(1) = 1$, so χ is either even or odd. Also, this even and odd actually means what we think it means since $\chi(-n) = \chi(-1)\chi(n)$.

Def 7.2. $\xi(s,\chi) = \left(\frac{q}{\pi}\right)^{\frac{s+a}{2}} \Gamma(\frac{s+a}{2}) L(s,\chi) \qquad \textit{and} \qquad \varepsilon(\chi) := \frac{\mathcal{G}(\chi)}{i^a\sqrt{q}}$

$\textit{where } a_\chi = \begin{cases} 0 & \chi \textit{ even} \\ 1 & \chi \textit{ odd} \end{cases}$

Note since $|\varepsilon(\chi)| = 1$, we have $\frac{1}{\varepsilon(\bar{\chi})} = \varepsilon(\chi)$.

Theorem 7.3. *Assume χ is non-principal and primitive. Then $L(s,\chi)$ and $\xi(s,\chi)$ admit holomorphic extensions to $\mathbb{C}$ which satisfies $\xi(s,\chi) = \varepsilon(\chi)\cdot\xi(1-s,\chi)$.*

Proof.

We first extend the functions to $Re(s) > 0$ then realize that it is symmetric with respect to $\frac{1}{2}$, and then we are done.

Let $f(x) = 2x^a e^{-\pi x^2}$ (which implies $f(-x) = \chi(-1)f(x)$), then by a change of variable $x = \pi y^2$ we have

$$\begin{aligned} \pi^{-\frac{s+a}{2}}\Gamma(\frac{s+a}{2}) &= \pi^{-\frac{s+a}{2}} \int_0^\infty e^{-x} x^{\frac{s+a}{2}-1} dx = \pi^{-\frac{s+a}{2}} \int_0^\infty e^{-\pi y^2} (\pi y^2)^{\frac{s+a}{2}-1} 2\pi y dy \\ &= \int_0^\infty e^{-\pi y^2} y^{s+a-1} dy = \int_0^\infty f(y) y^{s-1} dy \end{aligned}$$

hence

$$q^{\frac{s}{2}} L(s,\chi)\pi^{-\frac{s+a}{2}}\Gamma(\frac{s+a}{2}) = q^{\frac{s}{2}} \sum_{n\geq 1} \int_0^\infty f(y) \frac{\chi(n)}{n^s} y^{s-1} dy$$

by change of variables $y = \frac{ny}{\sqrt{q}}$,

$$= \sum_{n\geq 1} \int_0^\infty \chi(n) f\left(\frac{ny}{\sqrt{q}}\right) y^{s-1} dy = \int_0^\infty \sum_{n\geq 1} \chi(n) f\left(\frac{ny}{\sqrt{q}}\right) y^{s-1} dy$$

for $Re(s) > 1$.

Remember in the functional equation of ζ, the pole at 1 is caused by the additional term $n = 0$ when we try to flip θ; here, since $\chi(0) = 0$, that term goes away. More explicitly,

$$S_\chi(y) := \sum_{n\geq 1} \chi(n) f\left(\frac{ny}{\sqrt{q}}\right) = \frac{1}{2}\sum_{n\in\mathbb{Z}} \chi(n) f\left(\frac{ny}{\sqrt{q}}\right)$$

since χf is even (because they are both even or odd).

Now we use the twisted Poisson summation formula (Theorem 6.11) on f with $N = \frac{\sqrt{q}}{y}$ and get

$$S_\chi(y) = \frac{\chi(-1)\sqrt{q}}{2y\mathscr{G}(\bar{\chi})} \sum_{n\in\mathbb{Z}} \chi(n)\hat{f}\left(\frac{n}{q}\frac{\sqrt{q}}{y}\right)$$

On the other hand, due to the fact that f is Gaussian, $\hat{f} = \chi(-1)i^a f$, we have

$$S_\chi(y) = \varepsilon(\chi)\frac{S_{\bar{\chi}}(\frac{1}{y})}{y}$$

Therefore,

$$\begin{aligned} q^{-\frac{a}{2}}\xi(s,\chi) &= \int_0^1 \frac{\varepsilon(\chi)S_{\bar{\chi}}\left(\frac{1}{y}\right)}{y}\cdot y^{s-1}dy + \int_1^\infty S_\chi(y)\cdot y^{s-1}dy \\ &= \int_1^\infty \left[\varepsilon(\chi)S_{\bar{\chi}}(y)\cdot y^{-s} + S_\chi(y)\cdot y^{s-1}\right]dy \end{aligned}$$

Since the integral starts at 1, there's no problem with the value of y^{-s}, y^{s-1} and the Gaussian decay of S_χ means that the integral is holomorphic, which implies that $L(s,\chi)$ is holomorphic. (some justification of continuation first to Re(s)>0.)

Also, since $\varepsilon(\chi)\varepsilon(\bar{\chi}) = 1$, we have by plugging in

$$\xi(s,\chi) = \varepsilon(\chi)\xi(1-s,\bar{\chi})$$

□

Conjecture 1. *For any function with this kind of functional equation with symmetry, the extended RH holds, i.e., all zeros lies on the symmetry line.*

7.3. **Exercises.**

Example 7.1. *(11.2)*

8. PNT ON ARITHMETIC PROGRESSION

In this chapter we prove the Prime Number Theorem on Arithmetic Progressions. We have already shown in Chapter 7 that there are infinitely many primes $\equiv l \pmod q$ for $(l, q) = 1$. Let's restate our goal: (theorem 6.1)

For fixed $(q, a) = 1$,

$$\pi(x, q, a) = \#\{p \in \mathscr{P} : p \leq x, p \equiv a(mod\, q)\} \sim \frac{1}{\varphi(q)} \frac{x}{\log x}$$

Since the proof is not in the exam, I will leave it blank here. The idea is the same as the proof of PNT.

9. The Erdos-Kac Theorem

Theorem 9.1. *(Erdos-Kac Theorem)* $\forall a, b \in \mathbb{R}$,

$$\frac{1}{x}\#\left\{ n \leq x : \frac{\omega(n) - \log\log x}{\sqrt{\log\log x}} \in [a,b] \right\} \underset{x\to\infty}{\to} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{u^2}{2}} du$$

Remark: the the right hand side is just a normal distribution since

$$\mathbb{E}(\mathcal{N}(0,1) \in [a,b]) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{u^2}{2}} du$$

9.1. **Heuristics.**

To prove the theorem, we need to use some basic probabilistic facts and the Central Limit Theorem(CLT):

For $\mathcal{A} \subset \{n \leq x\}$, we write the uniform probability as:

$$\mathbb{P}_{n\leq x}(\mathcal{A}) = \frac{|\mathcal{A}|}{\lfloor x \rfloor}$$

Theorem 9.2. *(Central Limit Theorem) Assume $X_1, X_2, \ldots$ are independent variables, $\mu_i := \mathbb{E}(X_i)$, $\sigma^2 := Var(X_i)$; also, assume $\exists c > 0 : \forall i, |X_i| < c, \sigma^2 = o(\sum_{j=1}^{i} \sigma_j^2)$ (there are no dominant variable). Then $\forall a < b$*

$$\mathbb{P}\left(\frac{(X_1 + X_2 + \ldots + X_n) - (\mu_1 + \mu_2 + \ldots + \mu_n)}{(\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2)^{\frac{1}{2}}} \in [a,b] \right) \underset{x\to\infty}{\to} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{u^2}{2}} du$$

We leave the proof of CLT in the last part of this chapter. For now, let's see how we can apply it.

Before we go into the actual proof, let's have a look at 2 remarks that motivates our intuition.

Remark 9.1.

We know from Merten's estimate that $\sum_{p\leq x} \frac{1}{p} = \log\log x + O(1)$. So since $\sum_{p\leq x} \frac{1}{p^2} = O(1)$ we have

$$\sum_{p\leq x} \left(\frac{1}{p} - \frac{1}{p^2} \right) = \log\log x + O(1)$$

Let $X_2, X_3, X_5 \ldots$ be independent random variables (indexed by primes) that satisfies:

$$\mathbb{P}(X_p = 1) = \frac{1}{p} = 1 - \mathbb{P}(X_p = 0)$$

in other words, they are an "independent model" of prime distribution.

Then we have

- $\mathbb{E}(X_p) = \frac{1}{p}$
- $\mathbb{E}(X_p^2) = \frac{1}{p}$
- $Var(X_p) = \frac{1}{p} - \frac{1}{p^2}$

By CLT we know:

$$\mathbb{P}\left(\frac{\sum_{p\le x} X_p - \sum_{p\le x}\frac{1}{p}}{\left(\sum_{p\le x}\left(\frac{1}{p}-\frac{1}{p^2}\right)\right)^{\frac{1}{2}}} \in [a,b]\right) \underset{x\to\infty}{\to} \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{u^2}{2}}du$$

$$\Rightarrow \mathbb{P}\left(\frac{\sum_{p\le x} X_p - \log\log x}{\sqrt{\log\log x}} \in [a,b]\right) \underset{x\to\infty}{\to} \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{u^2}{2}}du$$

which is the exact expression in the theorem. However, the actual prime distribution is not independent for reasons like if $p > 2$ is a prime, then $p + 1$ is not a prime for sure. We will solve this problem in the next section.

Remark 9.2.

Pick n uniformly between 1 to x and write $\mathbb{P}_{n\le x}$ for the associated probability. The key step here is

$$\sum_{p\le x} X_p \approx \omega(n) = \sum_{p\le x} \mathbb{1}_{p|n}$$

But why?:

$$\mathbb{P}_{n\le x}(p_1|n,\dots,p_l|n) = \mathbb{P}_{n\le x}(p_1p_2\cdots p_l|n) = \frac{1}{\lfloor x\rfloor}\left\lfloor\frac{x}{p_1\cdots p_l}\right\rfloor \to \frac{1}{p_1\cdots p_l}$$

Where on the other hand

$$\mathbb{P}_{n\le x}(p_1|n)\cdots\mathbb{P}_{n\le x}(p_l|n) \to \frac{1}{p_1\cdots p_l}$$

So

$$\mathbb{P}_{n\le x}(\mathbb{1}_{p_1|n}=1,\dots,\mathbb{1}_{p_l|n}=1) \underset{x\to\infty}{\sim} \mathbb{P}_{n\le x}(X_{p_1}=1)\cdots\mathbb{P}_{n\le x}(X_{p_l}=1)$$

9.2. **The Kubilius Model.**

Let $y > 0, S(y) = n \in \mathbb{N} : p|n \Rightarrow p \leq y$. We call all numbers in $S(y)$ "y-smooth" and we assign a measure to $S(y)$.

For any $\mathcal{A} \subset S(y)$, define

$$\mathbb{P}_{S(y)}(\mathcal{A}) = \prod_{p \leq y}(1 - \frac{1}{p}) \sum_{n \in \mathcal{A}} \frac{1}{n}$$

we can see that it really is a probability measure with support $S(y)$ since

$$\mathbb{P}_{S(y)}(S(y)) = \prod_{p \leq y}(1 - \frac{1}{p}) \sum_{n \in S(y)} \frac{1}{n} = 1$$

by rearrangement.

Theorem 9.3. *Under probability measure $\mathbb{P}_{S(y)}$, the random variables $B_p := \mathbb{1}_{p|x}$ are independent.*

Proof.

Indeed,

$$\begin{aligned}
\mathbb{E}_{S(y)}[B_{p_1} \cdot B_{p_2} \cdots B_{p_l}] &= \prod_{p \leq y}(1 - \frac{1}{p}) \sum_{\substack{n \in S(y) \\ p_1 p_2 \cdots p_l | n}} \frac{1}{n} \\
&= \prod_{p \leq y}(1 - \frac{1}{p}) \sum_{m \in S(y)} \frac{1}{p_1 p_2 \cdots p_l} \frac{1}{m} \\
&= \frac{1}{p_1 p_2 \cdots p_l} = \mathbb{E}_{S(y)}[B_{p_1}] \cdot \mathbb{E}_{S(y)}[B_{p_2}] \cdots \mathbb{E}_{S(y)}[B_{p_l}]
\end{aligned}$$

where the middle step is because

$$\{n \in S(y) : p_1 p_2 \cdots p_l | n\} = \{(p_1 p_2 \cdots p_l)m : m \in S(y)\}$$

□

9.3. **Proof of the Erdos-Kac Theorem.**

To prove the Erdos-Kac Theorem, we use two lemmas.

Lemma 9.4. *(The method of moments) For random variables X_j with values in $\mathbb{R}$,*

(a) If

$$\mathbb{E}(X_j^k) \underset{j \to \infty}{\to} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} u^k e^{-\frac{u^2}{2}} du$$

for any $k \geq 0$, then $\forall a < b$,

$$\mathbb{P}(X_j \in [a, b]) \underset{j \to \infty}{\to} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{u^2}{2}} du$$

(b) If $\forall a < b$,

$$\mathbb{P}(X_j \in [a,b]) \underset{j\to\infty}{\to} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{u^2}{2}} du$$

and $\forall k \geq 0$, $\sup_{j\geq 1} \mathbb{E}[|X_j|^k] < +\infty$, *then*

$$\mathbb{E}(X_j^k) \underset{j\to\infty}{\to} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} u^k e^{-\frac{u^2}{2}} du$$

for any $k \geq 0$.

Remark: Note that the random variables in Lemma 9.4 can be of type

$$\frac{\omega(n) - \log\log x}{\sqrt{\log\log x}} \quad \text{or} \quad \frac{\sum_{i\leq y} B_i - \log\log x}{\sqrt{\log\log x}}$$

we can use it to apply and proof our theorem.

Lemma 9.5. $\forall k \geq 0$, $\sup_{j\geq 1} \mathbb{E}[|X_j|^k] < +\infty$ *is true for the Kubilius Model, i.e.,*

$$\forall k \geq 0, \sup_{j\geq 1} \mathbb{E}\left[\left|\frac{\omega(n) - \log\log x}{\sqrt{\log\log x}}\right|^k\right] < +\infty$$

Now we prove the theorem:

Proof. (of theorem 9.1)

Step 1: By Theorem 9.3 we have that under $\mathbb{P}_{S(y)}$, B_p are independent variables.

Also, since $\mathbb{E}(B_p) = \frac{1}{p}$ and $Var(B_p) = \frac{1}{p} - \frac{1}{p^2}$. Thus by CLT we have: $\forall a < b$,

$$\mathbb{P}_{S(y)}\left(\frac{\omega(n) - \log\log x}{\sqrt{\log\log x}} \in [a,b]\right) \underset{x\to\infty}{\to} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{u^2}{2}} du$$

By Lemma 9.5 and Lemma 9.4 (b) we know the convergence of the moments:

$$\mathbb{E}_{S(y)}\left[\left(\frac{\omega(n) - \log\log x}{\sqrt{\log\log x}} \in [a,b]\right)^k\right] \underset{x\to\infty}{\to} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} u^k e^{-\frac{u^2}{2}}$$

Now assume for some y, we have

$$\mathbb{E}_{S(y)}\left[\left(\frac{\omega(n) - \log\log x}{\sqrt{\log\log x}} \in [a,b]\right)^k\right] = \mathbb{E}_{n\leq x}\left[\left(\frac{\omega(n) - \log\log x}{\sqrt{\log\log x}} \in [a,b]\right)^k\right] + o(1) \quad (9.1)$$

then we're done.

Just to simplify notations, let $\lambda_x = \log\log x$, and for equation (9.1), let's just choose $y = x^{\frac{1}{\log\log\log x}}$. As for why, we'll explain later.

Now, by timing both sides with $\lambda_x^{\frac{k}{2}}$,

$$(9.1) \iff \mathbb{E}_{S(y)}[(\omega(n)-\lambda_y)^k] = \mathbb{E}_{n\le x}[(\omega(n)-\lambda_y)^k] + o(\lambda_x^{\frac{k}{2}})$$

since our choice of y guarantees $\log\log y \sim \log\log x$,

$$\iff \sum_{j=0}^{k}\binom{k}{j}\mathbb{E}_{S(y)}[\omega(n)^j]\lambda_x^{k-j} = \sum_{j=0}^{k}\binom{k}{j}\mathbb{E}_{n\le x}[\omega(n)^j]\lambda_x^{k-j} + o(\lambda_x^{\frac{k}{2}})$$

which means that we only need to show

$$\mathbb{E}_{S(y)}[\omega(n)^j] = \mathbb{E}_{n\le x}[\omega(n)^j] + o(\lambda_x^{j-\frac{k}{2}}) \tag{9.2}$$

for all $j \in [0,k]$.

Now let $\omega(n,y) = \#\{p \le y : p|n\}$, then

$$\omega(n) = \omega(n,y) + O(\log\log\log x)$$

since $\forall n \le x$, n has at most $\dfrac{\log x}{\log y}$ prime factors larger than y because:

Let $p_1,\dots,p_l$ be prime divisors of n that are larger than y, then

$$p_1\cdots p_l | n \le x \Rightarrow y^l \le p_1\cdots p_l \le x \Rightarrow l \le \frac{\log x}{\log y}$$

This means that (9.1) is equivalent to

$$\mathbb{E}_{S(y)}\left[\left(\frac{\omega(n,y)-\log\log x}{\sqrt{\log\log x}} \in [a,b]\right)^k\right] = \mathbb{E}_{n\le x}\left[\left(\frac{\omega(n)-\log\log x}{\sqrt{\log\log x}} \in [a,b]\right)^k\right] + o(1) \tag{9.3}$$

Remark: This explains why we've chosen $y = x^{\frac{1}{\log\log\log x}}$, that the error is $o(1)$. So in fact we can choose anything $\ll \sqrt{\log\log x}$, and our choice of y is just for convenience.

Now we prove (9.2) with the help of $\omega(n,y)$.

By definition, $\omega(n,y) = \sum_{p\le y} B_p$, so

$$\begin{aligned}\mathbb{E}_{n\le x}[\omega(n,y)^j] &= \mathbb{E}_{n\le x}\left[\sum_{p_1,\dots,p_j\le y} B_{p_1}\cdots B_{p_j}\right] \\ &= \sum_{p_1,\dots,p_j\le y}\mathbb{P}_{n\le x}\left([p_1,p_2,\dots,p_j]|n\right) = \sum_{p_1,\dots,p_j\le y}\left\lfloor\frac{x}{[p_1,p_2,\dots,p_j]}\right\rfloor\frac{1}{\lfloor x\rfloor}\end{aligned}$$

where the square bracket is means the least common multiple, we take that since it's possible that $p_i = p_k$.

So Let's bound the error if we take out $\lfloor\ \rfloor$:

$$\sum_{p_1,\dots,p_j\le y}\left\lfloor\frac{x}{[p_1,p_2,\dots,p_j]}\right\rfloor\frac{1}{\lfloor x\rfloor}=\sum_{p_1,\dots,p_j\le y}\frac{1}{[p_1,p_2,\dots,p_j]}+\frac{O(\pi(y)^j)}{\lfloor x\rfloor}$$

where by a very rough bound

$$\pi(y)^j\le y^j=x^{\frac{j}{\log\log\log x}}$$

On the other hand, $\forall A,c>0$,

$$\frac{x^{\frac{c}{\log\log\log x}}}{x}\ll(\log\log x)^A=\lambda^A$$

since by taking log we have

$$\frac{c\log x}{\log\log\log x}-\log x\ll -A\log\log\log x$$

So we have

$$\frac{O(\pi(y)^j)}{\lfloor x\rfloor}=o(\lambda_x^{j-\frac{k}{2}})$$

which means

$$\mathbb{E}_{n\le x}[\omega(n,y)^j]=\sum_{p_1,\dots,p_j\le y}\frac{1}{[p_1,p_2,\dots,p_j]}+o(\lambda_x^{j-\frac{k}{2}})$$

This reduces our problem into proving only

$$\mathbb{E}_{S(y)}[\omega(n)^j]=\sum_{p_1,\dots,p_j\le y}\frac{1}{[p_1,p_2,\dots,p_j]}$$

Since that for $S(y)$, B_i are independent random variables, we have

$$\mathbb{E}_{S(y)}[\omega(n)^j]=\mathbb{E}_{S(y)}\left[\sum_{p_1,\dots,p_j\le y}B_{p_1}\cdots B_{p_j}\right]$$

But we need to note that the independence is just for $p_i\neq p_k$, whereas when they are equal we can just leave only one of them. So let's say there are j' distinct primes in $\{p_1,\dots,p_j\}$ and we get

$$\mathbb{E}_{S(y)}[\omega(n)^j]=\sum_{p_1,\dots,p'_j\le y}\prod_{m=1}^{j'}\mathbb{E}_{S(y)}[B_{p_m}]=\sum_{p_1,\dots,p_j\le y}\frac{1}{[p_1,p_2,\dots,p_j]}$$

which justifies (9.3) and hence (9.1). So we are done.

□

9.4. **Proof of CLT and lemmas.**

I will prove CLT in this chapter, and offer some instinc on what the method of moments is about. As for Lemma 9.5, the proof is in book and I'll put it here after the exam.

9.5. **Exercises.**

Example 9.1. *(15.1)*

Example 9.2. *(15.4)*

Example 9.3. *(15.5)*

10. The Selbeg-Delange Method

There's some inner relations between $\sum_{n\le x}\frac{f(n)}{n^s}$ and $\sum_{p\le x}\frac{f(p)}{p^s}$ when f is multiplicative by methods similar to the Euler's product. In the proof of PNT, we goes from left to right. In this chapter, we will go the other direction, i.e., if $\sum_{p\le x}\frac{f(p)}{p^s}$ is understood, then so will be $\sum_{n\le x}\frac{f(n)}{n^s}$.

By "understood," in this chapter we only consider these settings: for $\kappa\in\mathbb{C}$ and $A>0$,

$$\sum_{p\le x} f(p)\log p = \kappa x + O_A\left(\frac{x}{(\log x)^A}\right) \tag{10.1}$$

$$|f| \le \tau_k \text{ for some fixed } k\ge 1 \tag{10.2}$$

10.1. **The Coefficients of Powers of Zeta.**

10.2. **An alternative for the Residue Theorem.**

Lemma 10.1. *(Hankel's formula)*

$$\frac{1}{2\pi i}\int_{(\alpha)}\frac{x^s}{s(s-1)^\kappa}ds = \frac{1}{\Gamma(\kappa)}\int_1^x(\log y)^{\kappa-1}dy$$

Proof. The key step is the change of variable $\frac{x^s}{x}\to y^s$ □

10.3. **The Selberg-Delange Theorem.**

Theorem 10.2. *(The Selberg-Delange Theorem)(only leading term): If (10.1) and (10.2) are satisfied, then*

$$\sum_{n\le x} f(n) = x\tilde{c_0}\frac{(\log x)^{\kappa-1}}{\Gamma(\kappa)}(1+o(1))$$

where

$$\tilde{c_0} = \prod_{p\in\mathcal{P}}\left[\left(1+\frac{f(p)}{p}+\frac{f(p^2)}{p^2}+\cdots\right)\left(1-\frac{1}{p}\right)^\kappa\right]$$

10.4. **Applications of the Selberg-Delange Theorem.**

10.5. **Exercises.**

Example 10.1. *(13.1)*

Example 10.2. *(13.4)*

11. Twin Pirmes and the Sieve Method

The only thing in exam here is to explain Brun's sieve, which is but a truncation of Legendre's sieve. Fill up later

Acknowledgements.