Topics in Analytic Number Theory

E. Assing

November 17, 2023

Abstract

These are some extended lecture notes for the course *Selected Topics in Analysis - Topics in Analytic Number Theory* taught in the winter term 2020/21 at the university of Bonn. **Attention**: This manuscript probably contains many misprints and inaccuracies. For personal use only!

This course is concerned with **randomness in the primes**. More precisely, we study the distribution of primes in long and short intervals covering some milestone results in the theory around prime numbers. We will not further motivate this endeavour and suggest anyone who feels the need for further motivation to take another course.

These lectures are inspired by essentially two courses on analytic number theory the author took at the university of Göttingen. The first being an introductory course to the topic taught by Prof. V. Blomer. The second was an advanced course on several aspects of prime numbers taught by Prof. J. Brüdern. The latter contained the infamous theorem on primes in short intervals due to Maier, which is the main highlight of this course.

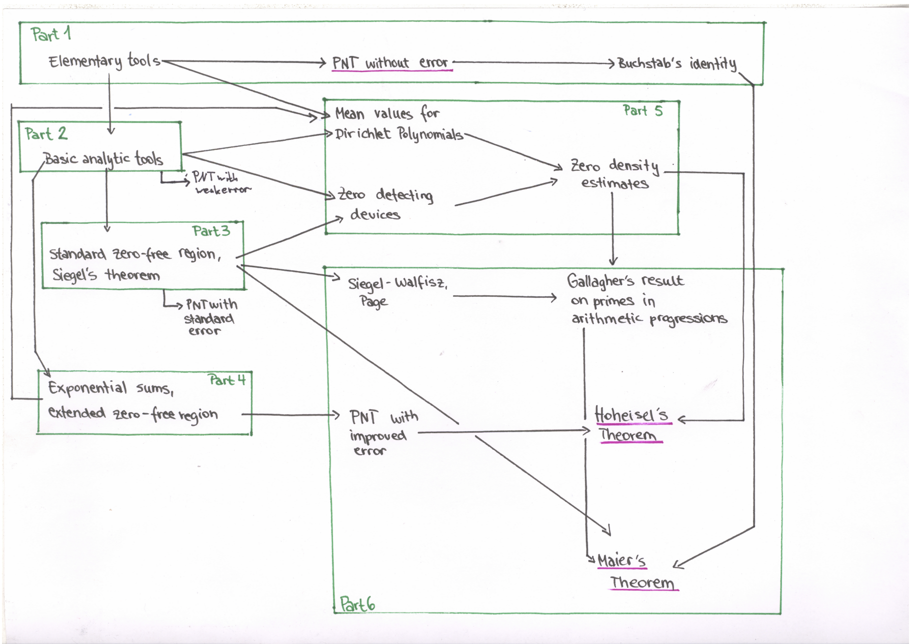
This course however, whose formal title is **Selected Topics in Analysis - Topics in Analytic Number Theory**, features only one lecture per week for approximately 13 weeks. To further complicate things the author wishes not to assume any number theoretic pre-requests. Thus the challenge is to teach a course culminating in Maier's theorem starting from the basics. The approach taken to this adventure can be seen as an experiment.

Note that we will only prove the main theorems, the highlights of the lecture so to speak, at the very end. We will however announce it as soon as we have developed enough theory to prove them and encourage the interested reader to perform the proof as an exercise. We must warn the reader that we expect a solid knowledge of standard complex analysis. We will use deep results such as the residue theorem, Cauchy's Integral Formula, Jensens Formula, Stirling's approximation and the Weierstrass Product Expansion without further explanations. We end this short introduction by providing a literature list as well as an overview over the content of this course.

Some literature suggestions:

- A good (german) introduction to complex analysis is provided in [2]
- A thorough (german) introduction to analytic number theory, which contains large parts of the material presented here is [1]
- The (classical) go-to reference for everything evolving around the Riemann zeta function is [10].

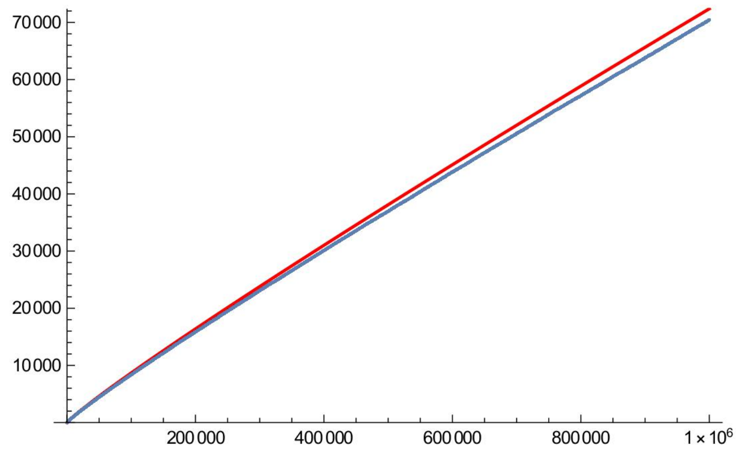
- A great overview over almost every aspect of analytic number theory is given in [6].
- A particular treat are the 10 lectures given in [9].
- The original treatment of Maier's theorem about primes in short intervals, [8], is impossible to improve.



Contents

Three Pictures

Before starting let us look at 3 pictures. The first picture contains the graph of $f(x) = \frac{x}{\log(x)}$ in red and the $\sharp\{x < p \leq 2x\}$ in blue:



Indeed the *Prime Number Theorem* (PNT) predicts that

$$\pi(x) = \sharp\{p \leq x\} = \frac{x}{\log(x)} + E(x)$$

with $E(x) = o(\frac{x}{\log(x)})$. This implies that

$$\sharp\{x < p \leq 2x\} = \pi(2x) - \pi(x) = \frac{2x}{\log(2x)} - \frac{x}{\log(x)} + E(2x) - E(x) \sim \frac{x}{\log(x)}.$$

This motivates the choice for $f(x)$ in the picture above. Note that Legendre (1798/1808) conjectured that $\pi(x) = \frac{x}{\log(x)+A(x)}$ or $\pi(x) = \frac{x}{A\log(x)+B}$. This was based on some numerical work. However already earlier (1792-93 unpublished) Gauß conjectured that

$$\sharp\{p \in [a, b]\} \sim \int_a^b \frac{dx}{\log(x)}.$$

This was included later in a letter to Encke. **Note that the tables of prime numbers Gauß produced throughout his live cover a wider range than the one pictured above!** The experiments made by Gauß and Legendre started the quest of proving the Prime Number Theorem. Let us mention some highlights of this journey. It was shown by Tschebyscheff (1851-52) that

$$C_1 \leq \liminf \frac{\pi(x)}{x/\log(x)} \leq 1 \leq \limsup \frac{\pi(x)}{x/\log(x)} \leq C_2.$$

Riemann (1960) linked the distribution of primes to the zeros of the famous Riemann zeta function $\zeta(s)$. Finally Hadamard and (independently) de la Valle Poussin (1896) proved the Prime Number Theorem with an error bound

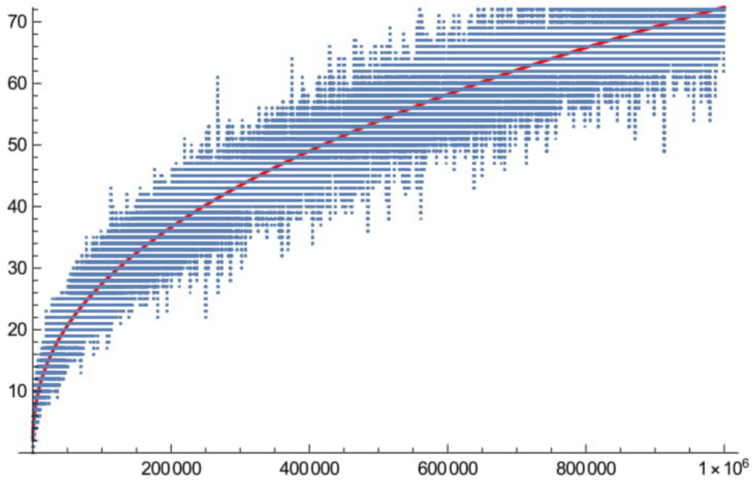
$$E(x) \ll xe^{-c\log(x)^{\frac{1}{14}}}.$$

A key ingredient in their proof is a zero-free region for the Riemann zeta function. Note that the error term can be slightly improved. The current record, based on the Vinogradov-Korobov zero-free region, is

$$E(x) \ll x \exp\left(-c' \frac{\log(x)^{\frac{3}{5}}}{\log\log(x)^{\frac{1}{5}}}\right).$$

Nowadays several proofs of the Prime Number Theorem are known including some elementary ones. The first such proof was found in 1948 by A. Selberg and P. Erdös.

One observes that better control on the error $E(x)$ in the Prime Number Theorem allows to predict the number of primes in shorter intervals. This leads us to the second picture. We plot $f(x) = \frac{\sqrt{x}}{\log(x)}$ in red and $\sharp\{x < p \leq x + \sqrt{x}\}$ in blue:

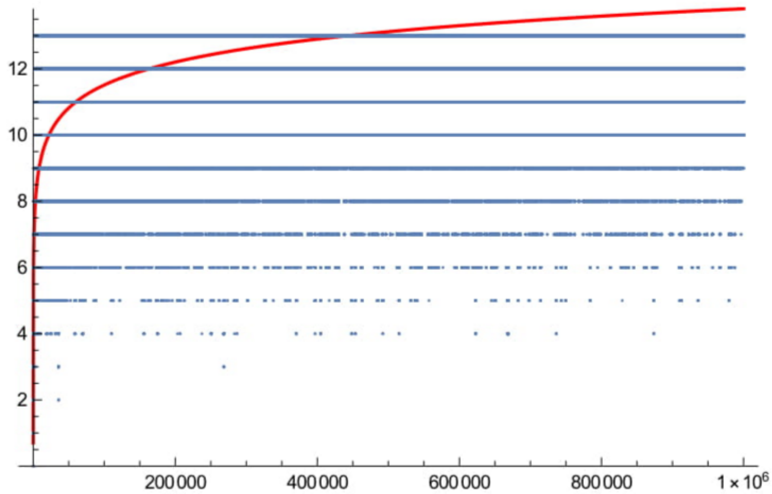


Indeed assuming the Riemann Hypothesis, which states that all *non-trivial zeros* of $\zeta(s)$ have real part $\frac{1}{2}$, one can show that $\pi(x + x^{\frac{1}{2}+\epsilon}) - \pi(x) \sim \frac{x^{\frac{1}{2}+\epsilon}}{\log(x)}$. This motivates the choices made in the picture. Note that unconditionally Huxley (based on work of Hoheisel and Ingham) showed that

$$\pi(x + x^{\frac{7}{12}+\epsilon}) - \pi(x) \sim \frac{x^{\frac{7}{12}+\epsilon}}{\log(x)}.$$

Allowing some exceptional x one can do even better. Indeed, one obtains $\pi(x + x^{\frac{1}{6}+\epsilon}) - \pi(x) \sim \frac{x^{\frac{1}{6}+\epsilon}}{\log(x)}$ for almost all x.

Assuming the Riemann Hypothesis Selberg pushed the so far mentioned results even further. He showed that $\pi(x + \Phi(x)) - \pi(x) \sim \frac{\Phi(x)}{\log(x)}$ for almost all x as long as $\frac{\Phi(x)}{\log(x)^2} \to \infty$. This brings us to the final picture which shows $f(x) = \log(x)$ in red and $\sharp\{x < p \leq x + \log(x)^2\}$ in blue:



Here it becomes obvious that the ranges we are considering are to small to produce informative pictures. Anyway one can ask if Selberg's just mentioned result is actually true for all x (sufficiently large) without any exceptions. This question was answered by Maier and we are aiming to present his argument at the end of this course.

1 Part 1: The basics

In this first part we lay the ground work for later chapters. We will start with an elementary proof of the prime number theorem (without error term) even though historically such a proof came much later. We further discuss some properties of the sieve of Eratosthenes. In the end we discuss some analytic preliminaries.

Notation: We write $f(x) = O(g(x))$ (formally correct would be $f(x) \in O(g(x))$) when there are $C, x_0 > 0$ such that $|f(x)| \leq C\,|g(x)|$ for all $x \geq x_0$. A stronger statement is $f(x) = o(g(x))$ (or $f(x) \in o(g(x))$), which means that for all $C > 0$ there is $x_0 > 0$ such that $|f(x)| \leq C\,|g(x)|$ for all $x \geq x_0$. This can be rephrased in terms of limits as follows:

$$f(x) = O(g(x)) \Leftrightarrow \limsup_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| < \infty \text{ and } f(x) = o(g(x)) \Leftrightarrow \lim_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| = 0.$$

We will often encounter expressions like $f(x) = g(x) + O(h(x))$ (resp. $f(x) = g(x) + o(h(x))$) meaning $f(x) - g(x) = O(h(x))$ (resp. $f(x) - g(x) = o(h(x))$). Further we use the standard abbreviation $f(x) \ll g(x)$ for $f(x) = O(g(x))$. Finally let us introduce $f(x) \asymp g(x)$ (resp. $f(x) \sim g(x)$), which means $f(x) \ll g(x) \ll f(x)$ (resp. $f(x) = g(x)(1 + o(1)) = g(x) + o(g(x))$.

The letter p is usually reserved for prime numbers and the set $\mathbb{N}$ of natural numbers does not contain 0. We write (n, m) for the greatest common divisor (gcd) of $m, n \in \mathbb{N}$.

1.1 Arithmetic functions, convolution and partial summation

A function $f \colon \mathbb{N} \to \mathbb{C}$ is called an **arithmetic function**. We write $\mathcal{A}$ for the set of all arithmetic functions. We equip $\mathcal{A}$ with point wise addition and define the (Dirichlet) convolution

$$[f \star g](n) = \sum_{m|n} f(m)g(\frac{n}{m}) \text{ for } f, g \in \mathcal{A}.$$

We obtain a commutative ring $(\mathcal{A}, +, \star)$ with identity element

$$\eta(n) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{else.} \end{cases}$$

One can check that $f \in \mathcal{A}$ is invertible if and only if $f(1) \neq 0$.

We call an arithmetic function $f \in \mathcal{A}$ **multiplicative** if it is non-trivial and satisfies $f(nm) = f(n)f(m)$ for all $n, m \in \mathbb{N}$ with $(n, m) = 1$. We say f is **completely multiplicative** if $f(nm) = f(n)f(m)$ for all $n, m \in \mathbb{N}$. A direct consequence of the definition is that a multiplicative arithmetic function f satisfies $f(1) = 1$. In particular such an f is invertible. Furthermore, the inverse of a multiplicative arithmetic function as well as the convolution of two multiplicative arithmetic functions are also multiplicative.

The identity η is a first obvious example of a completely multiplicative arithmetic function. Another straight forward one is the constant one function

$\epsilon(n) = 1$ for all $n \in \mathbb{N}$. The most common completely multiplicative arithmetic functions appearing in practice are Dirichlet characters. These are lifts of characters

$$\chi\colon (\mathbb{Z}/N\mathbb{Z})^\times \to S^1$$

to $\mathbb{Z}$ by $\chi(m) = \delta_{(m,N)=1} \cdot \chi(m + N\mathbb{Z})$. We call N the modulus of χ. As a consequence of elementary group theory one has the orthogonality relations:

$$\frac{1}{\varphi(q)} \sum_{\chi \bmod q} \overline{\chi(a)}\chi(n) = \delta_{n \equiv a \bmod q} \text{ for } (a,q) = 1.$$

Important examples for multiplicative arithmetic functions include the Möbius function $\mu = \epsilon^{-1}$. Recursively one obtains

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1, \\ (-1)^k & \text{if } n = p_1 \cdot \ldots \cdot p_k \text{ is square-free with } k \text{ distinct prime factors,} \\ 0 & \text{else.} \end{cases}$$

Other important multiplicative arithmetic functions are $\sigma_k(n) = \epsilon \star (\mathrm{id}^k)$, $d(n) = \sigma_0(n)$ and $\varphi(n) = \mu \star \mathrm{id}$. The von Mangoldt function given by

$$\Lambda(n) = \begin{cases} \log(p) & \text{if } n = p^k \text{ for } p \text{ prime and } k \in \mathbb{N}, \\ 0 & \text{else} \end{cases} \tag{1}$$

is very important but not multiplicative.

Note that in order to define a multiplicative function it is enough to specify its values on prime powers. Similar one can define completely multiplicative functions by defining their values on primes. We use this and define the completely multiplicative functions v_y and u_y by setting $v_y(p) = \mathbb{1}_{[1,y]}(p)$ and $u_y(p) = 1 - v_y(p)$. To get used to the convolution product let us state the following fundamental identities.

Lemma 1. *We have*

$$v_y \star u_y = \epsilon,\ (\mu \cdot v_y) \star \epsilon = u_y,\ (\mu \cdot v_y) \star (\mu \cdot u_y) = \mu,$$
$$\Lambda \star \mu = -(\mu \cdot \log), \text{ and } (v_y \cdot \Lambda) \star (v_y \cdot \mu) = -(v_y \cdot \mu \cdot \log).$$

Proof. The first equality is easily checked on prime powers:

$$[v_y \star u_y](p^r) = \sum_{0 \le k \le r} v_y(p^k) u_y(p^{r-k}) = \begin{cases} v_y(1)u_y(p^r) & \text{if } p > y, \\ v_y(p^r)u_y(1) & \text{if } p \le y \end{cases} = 1 = \epsilon(p^r).$$

For the second we write

$$[(\mu \cdot v_y) \star \epsilon](p^r) = \sum_{k=0}^{r} v_y(p^k)\mu(p^k) = 1 - v_y(p) = u_y(p) = u_y(p^r).$$

The third equation is easily seen to hold in the same fashion.

To see the first equality we observe that $\Lambda \star \epsilon = \log$. By Möbius inversion this yields

$$\Lambda = \mu \star \log = -\mu \log \star \epsilon.$$

The claimed equality follows since $\mu = \epsilon^{-1}$. The final equality follows directly from the fourth, because v_y is completely multiplicative. $\square$

One often encounters sums of arithmetic functions. We introduce the notation

$$S_f(y,x) = \sum_{y\leq n\leq x} f(n),$$

for $f \in \mathcal{A}$. Due to their importance we highlight the functions

$$\pi(x) = S_{\chi_{\mathcal{P}}}(1,x),\ \psi(x) = S_{\Lambda}(1,x) \text{ and } M(x) = S_{\mu}(1,x).$$

Here $\chi_{\mathcal{P}}$ is the indicator function on the primes.

The following two Lemmata are an essential piece in the toolkit of an analytic number theorist.[1]

Lemma 2 (Partial summation)**.** *Let $y \in \mathbb{N}$ and $x \in \mathbb{R}$ with $y < x$. For $g \in \mathcal{C}^1([y,x])$ we have*

$$\sum_{y\leq n\leq x} f(n)g(n) = S_f(y,x)g(x) - \int_y^x S_f(y,z)g'(z)dz.$$

Proof. Using the fundamental theorem of calculus we get

$$\begin{aligned} S_f(y,x)g(x) - \sum_{y\leq n\leq x} f(n)g(n) &= \sum_{y\leq n\leq x} f(n)(g(x)-g(n)) \\ &= \sum_{y\leq n\leq x} f(n)\int_n^x g'(\xi)d\xi = \int_y^x g'(\xi)\sum_{y\leq n\leq \xi} f(n)d\xi. \end{aligned}$$

□

Lemma 3 (Möbius inversion)**.** *Let $f\colon \mathbb{R}_+ \to \mathbb{C}$ be a complex valued function and $P \in \mathcal{A}$ be completely multiplicative. Then we have*

$$k(x) = \sum_{n\leq x} P(n)f(\frac{x}{n}) \text{ and } f(x) = \sum_{n\leq x}\mu(n)P(n)k(\frac{x}{n}).$$

Proof. We compute

$$\begin{aligned} \sum_{n\leq x}\mu(n)P(n)k(\frac{x}{n}) &= \sum_{n\leq x}\mu(n)P(n)\sum_{m\leq \frac{x}{n}} P(m)f(\frac{x}{mn}) \\ &= \sum_{mn\leq x}\mu(n)P(mn)f(\frac{x}{mn}) \\ &= \sum_{c\leq x} P(c)f(\frac{x}{c})\sum_{d|c}\mu(d) = f(x)P(1) = f(x). \end{aligned}$$

A similar computation shows how to recover k from f. □

Exercise 1. *Prove the asymptotic*

$$\sum_{p\leq x}\frac{\log(p)}{p} = \log(x) + O(1). \tag{2}$$

[1] Arguably partial summation is the most important tool after the Cauchy-Schwarz inequality.

Hint: One can use Stirling's approximation in the form

$$\log(n!) = n\log(n) - n + O(\log(n))$$

and the elementary observation

$$n! = \prod_{p\leq n} p^{e(p)} \implies e(p) = \sum_{k\geq 1}\left\lfloor \frac{n}{p^k}\right\rfloor.$$

Solution. We write

$$\log(n!) = \sum_{p\leq n} e(p)\log(p) = \sum_{p\leq n}\sum_{k\geq 1}\left\lfloor \frac{n}{p^k}\right\rfloor \log(p) = \sum_{p\leq n}\left\lfloor \frac{n}{p}\right\rfloor \log(p) + E_1.$$

The error can be estimated as follows

$$E_1 = \sum_{p\leq n}\sum_{k\geq 2}\left\lfloor \frac{n}{p^k}\right\rfloor \log(p) \leq n\sum_{p\leq n}\log(p)\sum_{k\geq 2}\frac{1}{p^k} = n\sum_{p\leq n}\frac{\log(p)}{p(p-1)} = O(n).$$

Using Stirling's approximation yields

$$\sum_{p\leq n}\left\lfloor \frac{n}{p}\right\rfloor \log(p) = n\log(n) + O(n).$$

We can remove the Gauß brackets and obtain

$$n\sum_{p\leq n}\frac{\log(p)}{p} = \sum_{p\leq n}\left(\left\lfloor \frac{n}{p}\right\rfloor + O(1)\right)\log(p) = n\log(n) + O(\tilde{\psi}(n) + n).$$

Thus we still need an estimate for $\tilde{\psi}(x) = \sum_{p\leq x}\log(p)$.[2] This is obtained as follows:

$$\begin{aligned}\tilde{\psi}(2n) - \tilde{\psi}(n) = \sum_{n<p\leq 2n}\log(p) &\leq \sum_{p\leq 2n}\left(\left\lfloor \frac{2n}{p}\right\rfloor - 2\left\lfloor \frac{n}{p}\right\rfloor\right)\log(p)\\ &= 2n\log(2n) - 2n\log(n) + O(n) = O(n).\end{aligned}$$

With this at hand we can estimate

$$\tilde{\psi}(x) \leq \sum_{i=1}^{\infty}\left|\tilde{\psi}(\frac{x}{2^{i-1}}) - \tilde{\psi}(\frac{x}{2^i})\right| \ll x\sum_{i=1}^{\infty}2^{-i} \ll x$$

and the proof is complete. □

[2]Note that $\psi(x) = \tilde{\psi}(x) + \sum_{\substack{r\geq 2,\\ p^r\leq x}}\log(p)$.

1.2 An elementary proof of the prime number theorem

Besides proving the prime number theorem, which is a mile stone in the field, we will learn a handful of tricks we will frequently use later on. The approach we take is termed the convolution method and was developed in [4, 3].

Lemma 4. *We have*

$$\lim_{x\to\infty} \frac{S_{u_y}(1,x)}{x} = \prod_{p\le y}\left(1-\frac{1}{p}\right).$$

Proof. We compute

$$\begin{aligned}
\frac{S_{u_y}(1,x)}{x} &= \frac{1}{x}\sum_{n\le x}[(v_y\cdot\mu)\star\epsilon](n)\\
&= \frac{1}{x}\sum_{n\le x}\sum_{m|n} v_y(m)\mu(m)\\
&= \sum_{n\le x}\frac{v_y(n)\mu(n)}{n}\cdot\frac{n}{x}\sum_{m\le\frac{x}{n}} 1\\
&= \sum_{n\le x}\frac{v_y(n)\mu(n)}{n}\cdot\frac{n}{x}\left\lfloor\frac{x}{n}\right\rfloor.
\end{aligned}$$

Taking the limit yields

$$\lim_{x\to\infty}\frac{S_{u_y}(1,x)}{x} = \sum_{n\ge 1}\frac{v_y(n)\mu(n)}{n}.$$

It is important to note that the remaining sum is finite. Indeed, if we set $\mathcal{S}(y) = \{n\colon \square-\text{free with } p\mid n \implies p\le y\}$, then

$$\lim_{x\to\infty}\frac{S_{u_y\cdot\mu}(1,x)}{x} = \sum_{n\in\mathcal{S}(y)}\mu(n)n^{-1} = \prod_{p\le y}\left(1-\frac{1}{p}\right).$$

□

Lemma 5. *We have*

$$\limsup_{x\to\infty}\frac{|M(x)|}{x} \le \prod_{p\le y}\left(1-\frac{1}{p}\right)\int_1^\infty \frac{\left|S_{v_y\cdot\mu}(1,t)\right|}{t^2}dt.$$

Proof. We start by observing

$$\begin{aligned}
M(x) &= \sum_{n\le x}[(\mu\cdot u_y)\star(\mu\cdot v_y)](n)\\
&= \sum_{n\le x}\sum_{m|n} u_y(m)\mu(m)v_y(n/m)\mu(n/m)\\
&= \sum_{n\le x} u_y(n)\mu(n)\sum_{m\le\frac{x}{n}} v_y(m)\mu(m)\\
&= \sum_{n\le x} u_y(n)\mu(n)S_{v_y\cdot\mu}(1,\frac{x}{n}).
\end{aligned}$$

We now write

$$\mathcal{S}(y) \cup \{\infty\} = \{1 = d_1 < d_2 < \ldots < d_k < d_{k+1} = \infty\}.$$

Note that if $\frac{x}{d_{j+1}} < n \leq \frac{x}{d_j}$, then $S_{v_y \cdot \mu}(1, \frac{x}{n}) = S_{v_y \cdot \mu}(1, d_j)$, for $j = 1, \ldots, k$. We get

$$M(x) = \sum_{j=1}^{k} S_{v_y \cdot \mu}(1, d_j) \sum_{\frac{x}{d_{j+1}} < n \leq \frac{x}{d_j}} u_y(n)\mu(n).$$

We get the estimate

$$\begin{aligned}
\limsup_{x \to \infty} \frac{|M(x)|}{x} &\leq \sum_{j=1}^{k} \left|S_{v_y \cdot \mu}(1, d_j)\right| \lim_{x \to \infty} \frac{1}{x} \sum_{\frac{x}{d_{j+1}} < n \leq \frac{x}{d_j}} u_y(n) \\
&= \prod_{p \leq y} \left(1 - \frac{1}{p}\right) \sum_{j=1}^{k} \left(\frac{1}{d_j} - \frac{1}{d_{j+1}}\right) \left|S_{v_y \cdot \mu}(1, d_j)\right| \\
&= \prod_{p \leq y} \left(1 - \frac{1}{p}\right) \int_1^{\infty} \frac{\left|S_{v_y \cdot \mu}(1, t)\right|}{t^2} dt.
\end{aligned}$$

□

Lemma 6. *For $y \geq 2$ and $t \geq 1$, we have*

$$\left|S_{v_y \cdot \mu}(1, t)\right| \log(t) \leq \sum_{n \leq t} v_y(n)\Lambda(n) \left|S_{v_y \cdot \mu}(1, \frac{t}{n})\right| + \sum_{n \leq t} v_y(n) \log(\frac{t}{n}).$$

Proof. The proof only uses a convolution identity and the triangle inequality. Indeed

$$\begin{aligned}
\left|S_{v_y \cdot \mu}(1, t)\right| \log(t) &\leq \left|\sum_{n \leq t} v_y(n)\mu(n) \log(n)\right| + \left|\sum_{n \leq t} v_y(n)\mu(n) \log(\frac{t}{n})\right| \\
&\leq \left|\sum_{n \leq t} [(v_y \cdot \Lambda) \star (v_y \cdot \mu)](n)\right| + \sum_{n \leq t} v_y(n) \log(\frac{t}{n}).
\end{aligned}$$

The claim follows by opening the convolution and interchanging summation:

$$\begin{aligned}
\sum_{n \leq t} [(v_y \cdot \Lambda) \star (v_y \cdot \mu)](n) = \sum_{n \leq t} v_y(n) \sum_{m | n} \Lambda(m)\mu(\frac{n}{m}) \\
= \sum_{m \leq t} v_y(m)\Lambda(m) \underbrace{\sum_{m | n} v_y(\frac{n}{m})\mu(\frac{n}{m})}_{= S_{v_y \cdot \mu}(1, \frac{t}{m})}.
\end{aligned}$$

□

We define an auxiliary function

$$k(u) = \int_0^{\infty} \exp\left(-ux + \int_0^{\infty} \frac{1 - e^{-t}}{t} dt\right) dx,$$

for $u > 0$. This function is positive, monotone decreasing and we make the following claim.
Claim: We have $uk(u) - \int_u^{u+1} k(x)dx = 1$ for all $u > 0$.
To see this we set $f(x) = \int_0^\infty \frac{1-e^{-t}}{t}dt$. Note that by the fundamental theorem of calculus we have $f'(x) = \frac{1-e^{-x}}{x}$. We compute

$$\begin{aligned}\int_u^{u+1} k(x)dx = \int_0^\infty e^{f(x)} \int_u^{u+1} e^{-xz}dzdx &= \int_0^\infty e^{-ux} f'(x)e^{f(x)}dx \\ &= -1 + u\int_0^\infty e^{-ux+f(x)}dx = -1 + uk(u).\end{aligned}$$

In the second to last step we applied by partial integration.

We rescale k by setting

$$H_y(t) = \frac{1}{\log(y)}k\left(\frac{\log(t)}{\log(y)}\right).$$

Note that after a change of variables our claim produces the identity

$$H_y(t)\log(t) - \int_t^{yt} \frac{H_y(x)}{x}dx = 1. \tag{3}$$

By partial summation and (2) we find the formulae

$$\sum_{p\leq y} H_y(pt)\frac{\log(p)}{p} = \int_t^{yt} \frac{H_y(v)}{v}dv + O(H_y(y)) \text{ for } t \geq y \text{ and} \tag{4}$$

$$\sum_{\frac{y}{t}<p\leq t} H_y(pt)\frac{\log(p)}{p} = \int_y^{yt} \frac{H_y(v)}{v}dv + O(H_y(y)) \text{ for } t > 1. \tag{5}$$

Let us present the details only for (4) as the second formula follows similarly. One starts from

$$\begin{aligned}&\sum_{p\leq y} H_y(pt)\frac{\log(p)}{p} = H_y(yt)\sum_{p\leq y}\frac{\log(p)}{p} - \int_1^y \frac{d}{dz}H_y(zt)\sum_{p\leq z}\frac{\log(p)}{p}dz \\ &= \log(y)H_y(yt) - \int_1^y \log(z)\frac{d}{dz}H_y(zt)dz + O\left(|H_y(yt)| + \int_1^y \left|\frac{d}{dz}H_y(zt)\right| dz\right).\end{aligned}$$

By partial integration and a change of variables we have

$$\log(y)H_y(yt) - \int_1^y \log(z)\frac{d}{dz}H_y(zt)dz = \int_t^{yt} \frac{H_y(v)}{v}dv.$$

Thus we have to treat the $O(\ldots)$-term. Obviously we have $H_y(yt) \leq H_y(y)$. The integral can be treated trivially after observing that

$$\frac{d}{dz}H_y(zt) = \frac{1}{z\log(y)^2}k'\left(\frac{\log(zt)}{\log(y)}\right) = -\frac{2k\left(\frac{\log(zt)}{\log(y)}\right) - k\left(\frac{\log(zt)}{\log(y)}+1\right)}{z\log(zt)\log(y)} \ll \frac{H_y(y)}{z\log(zt)}.$$

Here we used the differential equation $uk'(u) = k(u+1) - 2k(u)$ satisfied by k. This follows by differentiating the equality in the claim above.

Lemma 7 (Mertens+ϵ)**.** *We have*

$$\prod_{p<x}\left(1-\frac{1}{p}\right)\sim\frac{1}{C\log(x)}. \tag{6}$$

Furthermore,

$$\int_1^2 k(u)(2-u)du = C-1.$$

Proof. We start by showing that there is a constant C such that (6) holds. We write $\sum_{p<x}\frac{\log(p)}{p}=\log(x)+R(x)$ with $R(x)\ll 1$. By partial summation we derive

$$\begin{aligned}
\sum_{p<x}\frac{1}{p} &= 1+\underbrace{\frac{R(x)}{\log(x)}}_{O(\frac{1}{\log(x)})}+\underbrace{\int_2^x\frac{dx}{x\log(x)}}_{=\log\log(x)-\log\log(2)}+\int_2^x\frac{R(x)}{x\log(x)^2}dx \\
&= \log\log(x)+\underbrace{\int_2^\infty\frac{R(x)}{x\log(x)^2}dx-\log\log(2)}_{=M}-\underbrace{\int_x^\infty\frac{R(x)}{x\log(x)^2}dx}_{=O(\frac{1}{\log(x)})}+O(\frac{1}{\log(x)}) \\
&= \log\log(x)+M+O(\frac{1}{\log(x)}).
\end{aligned}$$

from (2). Now we put

$$B=\sum_p\sum_{k>1}\frac{p^{-k}}{k}\le\sum_p p^{-\frac{3}{2}}\sum_{k>1}\frac{p^{\frac{3}{2}-k}}{k}\le\sum_n n^{-\frac{3}{2}}\sum_{k>1}\frac{2^{\frac{3}{2}-k}}{k}<\infty.$$

Taylor expanding the logarithm at 1 we get

$$\log\prod_{p<x}(1-\frac{1}{p})^{-1}=\sum_{p<x}\frac{1}{p}+\sum_{p<x}\sum_{k>1}\frac{p^{-k}}{k}=\log\log(x)+M+B+O(\frac{1}{\log(x)}).$$

Exponentiating again gives the constant $C=e^{M+B}$. A similar argument shows

$$\sum_n\frac{v_y(n)}{n}=\prod_{p\le y}\left(1+\frac{1}{p}\right)\ll\log(y),$$

which we will use below.

To evaluate the integral will be a little harder. We start by using the convolution identity $v_y\log = v_y\cdot\Lambda\star v_y$ to get

$$\begin{aligned}
\sum_{n\le t}v_y(n)\log(n) &= \sum_{n\le t}\sum_{m|n}\Lambda(m)v_y(m)v_y(\frac{n}{m}) \\
&= \sum_{m\le t}v_y(m)\Lambda(m)\sum_{m|n\le t}v_y(\frac{n}{m})=\sum_{m\le t}v_y(m)\Lambda(m)S_{v_y}(1,\frac{t}{m}).
\end{aligned}$$

We rewrite this as

$$S_{v_y}(1,t)\log(t)=\sum_{\substack{p\le y,\\ p\le t}}\log(p)S_{v_y}(1,\frac{t}{p})+\sum_{\substack{p\le y,\\ p^r\le t,\, r\ge 2}}\log(p)S_{v_y}(1,\frac{t}{p^r})+\sum_{n\le t}v_y(n)\log(\frac{t}{n}).$$

Integrating this formula against $H_y(t)t^{-2}$ we get

$$\int_y^\infty \frac{S_{v_y}(1,t)}{t^2}\log(t)H_y(t)dt = \int_y^\infty \sum_{\substack{p\leq y,\\ p\leq t}} \log(p)S_{v_y}(1,\frac{t}{p})\frac{H_y(t)}{t^2}dt$$
$$+\underbrace{\int_y^\infty \sum_{\substack{p\leq y,\\ p^r\leq t,\, r\geq 2}} \log(p)S_{v_y}(1,\frac{t}{p^r})\frac{H_y(t)}{t^2}dt}_{=E_1} + \underbrace{\int_y^\infty \sum_{n\leq t} v_y(n)\log(\frac{t}{n})\frac{H_y(t)}{t^2}dt}_{=E_2}. \quad (7)$$

We now estimate the two errors (we will see similar computations below). First look at

$$E_1 \leq H_y(y)\sum_{p,\, r\geq 2}\log(p)\int_y^\infty S_{v_y}(1,\frac{t}{p^r})\frac{dt}{t^2} = H_y(y)\left(\sum_{p,\, r\geq 2}\frac{\log(p)}{p^r}\right)\left(\int_1^\infty \frac{S_{v_y}(1,\frac{t}{p^r})}{t^2}dt\right)$$
$$\ll H_y(y)\int_1^\infty \frac{S_{v_y}(1,\frac{t}{p^r})}{t^2}dt \ll H_y(y)\int_1^\infty \sum_{n\leq t} v_y(n)\frac{dt}{t^2} \ll H_y(y)\sum_n \frac{v_y(n)}{n} \ll 1.$$

since $H_y(y) \ll \log(y)^{-1}$ and the sum over p and r can be estimated trivially. The second error is treated similarly. Indeed

$$E_2 \leq H_y(y)\sum_n v_y(n)\int_n^\infty \log(\frac{t}{n})\frac{dt}{t^2} = H_y(y)\sum_n \frac{v_y(n)}{n}\int_1^\infty \frac{\log(t)}{t^2}dt \ll H_y(y)\log(y) \ll 1.$$

The remaining contribution will be manipulated as follows

$$\int_y^\infty \sum_{\substack{p\leq y,\\ p\leq t}} \log(p)S_{v_y}(1,\frac{t}{p})\frac{H_y(t)}{t^2}dt = \sum_{p\leq y}\frac{\log(p)}{p}\int_{\frac{y}{p}}^\infty \frac{S_{v_y}(1,t)}{t^2}H_y(pt)dt$$
$$= \int_1^y \frac{S_{v_y}(1,t)}{t^2}\sum_{\frac{y}{t}<p\leq y}\frac{\log(p)}{p}H_y(pt)dt + \int_y^\infty \frac{S_{v_y}(1,t)}{t^2}\sum_{p\leq y}\frac{\log(p)}{p}H_y(pt)dt.$$

In the last step we split the integral into the ranges $[\frac{y}{p}, y]$ and $[y,\infty]$ and interchanged summation and integration once again. We apply (4) and (5) to the p-sums and get

$$\int_y^\infty \sum_{\substack{p\leq y,\\ p\leq t}} \log(p)S_{v_y}(1,\frac{t}{p})\frac{H_y(t)}{t^2}dt$$
$$= \int_1^y \frac{S_{v_y}(1,t)}{t^2}\int_y^{yt}\frac{H_y(x)}{x}dxdt + \int_y^\infty \frac{S_{v_y}(1,t)}{t^2}\int_t^{yt}\frac{H_y(x)}{x}dxdt + O(1). \quad (8)$$

The error term comes from

$$H_y(y)\int_1^\infty \frac{S_{v_y}(1,t)}{t^2}dt = H_y(y)\sum_n \frac{v_y(n)}{n} \ll 1.$$

Combining (7) with (8) and the two error estimates yields

$$\int_y^\infty \frac{S_{v_y}(1,t)}{t^2}\log(t)H_y(t)dt = \int_1^y \frac{S_{v_y}(1,t)}{t^2}\int_y^{yt}\frac{H_y(x)}{x}dxdt + \int_y^\infty \frac{S_{v_y}(1,t)}{t^2}\int_t^{yt}\frac{H_y(x)}{x}dxdt + O(1). \quad (9)$$

With this at hand we get

$$\int_y^\infty \frac{S_{v_y}(1,t)}{t^2}dt = \int_y^\infty \frac{S_{v_y}(1,t)}{t^2}\underbrace{\left(H_y(t)\log(t) - \int_t^{yt}\frac{H_y(x)}{x}dx\right)}_{\overset{(3)}{=}1}dt$$

$$\overset{(9)}{=} \int_1^y \frac{S_{v_y}(1,t)}{t^2}\int_y^{yt}\frac{H_y(x)}{x}dxdt + O(1) = \int_1^y\int_y^{yt} H_y(x)\frac{dx}{x}\frac{dt}{t} + O(1).$$

Here we noted that for $t \le y$ we one simply has $S_{v_y}(1,t) = \lfloor t \rfloor = t + O(1)$. The so obtained error can be ignored since

$$\int_1^y\int_y^{yt} H_y(x)\frac{dx}{x}\frac{dt}{t^2} \le H_y(y)\int_1^\infty \log(t)\frac{dt}{t^2} = O(\frac{1}{\log(y)}).$$

One checks that

$$\begin{aligned}
&\int_1^y\int_y^{yt} H_y(x)\frac{dx}{x}\frac{dt}{t} = \int_1^y\int_1^{1+\frac{\log(t)}{\log(y)}} k(z)dz\frac{dt}{t}\\
&= \log(y)\int_1^2 k(z)dz - \int_1^y \frac{\log(t)}{\log(y)}k(1+\frac{\log(t)}{\log(y)})\frac{dt}{t}\\
&= \log(y)\int_1^2 k(z) - (z-1)k(z)dz = \log(y)\int_1^2 k(z)(2-z)dz.
\end{aligned}$$

Finally we compute that

$$\begin{aligned}
\int_y^\infty \frac{S_{v_y}(1,t)}{t^2}dt &= \sum_{n\in\mathbb{N}}\frac{v_y(n)}{n} - \log(y) + O(1) = \prod_{p\le y}\sum_{r\ge 0} p^{-r} - \log(y) + O(1)\\
&= \prod_{p\le y}(1-\frac{1}{p})^{-1} - \log(y) + O(1) = (C+o(1))\log(y) - \log(y) + O(1).
\end{aligned}$$

Thus we have shown that

$$\log(y)\int_1^2 k(z)(2-z)dz = (C-1+o(1))\log(y) + O(1)$$

and the claim follows by dividing by $\log(y)$ and choosing y sufficiently large. $\square$

By a simple substitution the lemma above yields

$$\int_y^{y^2}\frac{H_y(t)}{t}\log(\frac{y^2}{t})dt = (C-1)\log(y).$$

Lemma 8. *We have*

$$\int_y^\infty \frac{\left|S_{v_y\cdot\mu}(1,t)\right|}{t^2}dt \le \int_1^y \frac{\left|S_{v_y\cdot\mu}(1,t)\right|}{t^2}\left(\int_y^{yt}\frac{H_y(x)}{x}dx\right)dt + O(1).$$

The following proof has some overlaps with the previous proof.

Proof. We start from Lemma 6 and get

$$\int_y^\infty \left|S_{v_y\cdot\mu}(1,t)\right|\log(t)\frac{H_y(t)}{t^2}dt \le \underbrace{\int_y^\infty \sum_{n\le t} v_y(n)\Lambda(n)\left|S_{v_y\cdot\mu}(1,\frac{t}{n})\right|\frac{H_y(t)}{t^2}dt}_{=T}$$
$$+\underbrace{\sum_{n\le t}\int_y^\infty v_y(n)\int_y^\infty \log(\frac{t}{n})\frac{H_y(t)}{t^2}dt}_{E_2}.$$

Here E_2 is the same error as in the proof of the previous lemma. There we obtained $E_2 \ll 1$. To deal with T we use the definition of v_y and split the sum in two pieces:

$$T = \underbrace{\int_y^\infty \sum_{p\le y}\log(p)\left|S_{v_y\cdot\mu}(1,\frac{t}{p})\right|\frac{H_y(t)}{t^2}dt}_{=T_0} + \underbrace{\int_y^\infty \sum_{\substack{p\le y,\\ r\ge 2,\\ p^r\le t}}\log(p)\left|S_{v_y\cdot\mu}(1,\frac{t}{p^r})\right|\frac{H_y(t)}{t^2}dt}_{=E_1'}.$$

If E_1 is the same error as in the previous proof, then we obtain $E_1' \ll E_1 \ll 1$ simply by the triangle inequality. Turning to T_0 we exchange summation and integration and make a change of variables to find

$$T_0 = \sum_{p\le y}\frac{\log(p)}{p}\int_{\frac{y}{p}}^\infty \frac{\left|S_{v_y\cdot\mu}(1,t)\right|}{t^2}H_y(pt)dt.$$

We split the integral in two pieces, exchange summation and integration again and apply (4) and (5). This way we get

$$\begin{aligned}
T_0 &= \int_1^y \frac{\left|S_{v_y\cdot\mu}(1,t)\right|}{t^2}\sum_{\frac{y}{t}<p\le y}\frac{\log(p)}{p}H_y(pt)dt + \int_y^\infty \frac{\left|S_{v_y\cdot\mu}(1,t)\right|}{t^2}\sum_{p\le t}\frac{\log(p)}{p}H_y(pt)dt \\
&= \int_1^y \frac{\left|S_{v_y\cdot\mu}(1,t)\right|}{t^2}\int_y^{yt}\frac{H_y(v)}{v}dvdt + \int_y^\infty \frac{\left|S_{v_y\cdot\mu}(1,t)\right|}{t^2}\underbrace{\int_t^{yt}\frac{H_y(v)}{v}dv}_{=H_y(t)\log(t)-1}dt \\
&\quad + O\left(H_y(y)\int_1^\infty \frac{\left|S_{v_y\cdot\mu}(1,t)\right|}{t^2}dt\right).
\end{aligned}$$

The error can be treated as above and is $O(1)$. We use our integral identities for the v-integrals to get

$$T_0 = \int_1^y \frac{\left|S_{v_y\cdot\mu}(1,t)\right|}{t^2}\int_y^{yt}\frac{H_y(v)}{v}dvdt + O\left(\int_y^\infty \frac{\left|S_{v_y\cdot\mu}(1,t)\right|}{t^2}(1+H_y(t)\log(t))dt + 1\right).$$

The O-term is easily seen to be $\ll 1$ and this concludes the proof. □

Lemma 9. *For any* $\limsup_{x\to\infty} \frac{|M(x)|}{x} < \beta < 2$ *we have*

$$\int_y^\infty \frac{\left|S_{v_y\cdot\mu}(1,t)\right|}{t^2}dt \leq \beta(C-1)\log(y) + O(1).$$

Proof. We begin by fixing x_0 large enough such that

$$|M(x)| \leq \beta x \text{ for } x \geq x_0.$$

Further observe that for $n \leq y$ we have $v_y(n) = 1$, so that

$$S_{v_y\cdot\mu}(1,x) = M(x) \text{ if } x \leq y.$$

These observations together with the result of the previous lemma yield the upper bound

$$\int_y^\infty \frac{\left|S_{v_y\cdot\mu}(1,t)\right|}{t^2}dt \leq \beta \int_1^y t^{-1}\left(\int_y^{yt} \frac{H_y(x)}{x}dx\right)dt + O\left(1 + \int_1^{x_0} t^{-1}\left(\int_y^{yx_0} \frac{H_y(x)}{x}dx\right)dt\right).$$

The error term is $O(1)$ and we conclude the proof by observing that

$$\int_1^y t^{-1}\left(\int_y^{yt} \frac{H_y(x)}{x}dx\right)dt = \int_y^{y^2} \frac{H_y(t)}{t}\log(\frac{y^2}{t})dt = (C-1)\log(y).$$

□

Lemma 10. *There is some constant* $M > 0$ *such that*[3]

$$\left|\int_a^b \frac{S_\mu(1,x)}{x^2}dx\right| \leq M. \tag{10}$$

Proof. By partial summation we find

$$\int_a^b \frac{S_\mu(1,x)}{x^2}dx = \frac{1}{b}S_\mu(1,b) - \frac{1}{a}S_\mu(1,a) + \sum_{a\leq n\leq b}\frac{\mu(n)}{n} \ll 1 + \left|\sum_{a\leq n\leq b}\frac{\mu(n)}{n}\right|.$$

The latter sum can be bounded with elementary means.

We deduce this from the generalised Möbius inversion formula Lemma 3. Indeed applying this lemma with $f \equiv 1$ and $P = \epsilon$ we find that $k(x) = \lfloor x \rfloor$. The inversion formula then implies

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

The desired bound

$$\sum_{n\leq x}\frac{\mu(n)}{n} \ll 1$$

follows. □

[3] One could simply write

$$\int_a^b \frac{S_\mu(1,x)}{x^2}dx \ll 1,$$

but we will need a name for the implicit constant later.

Lemma 11. *There is δ such that*
$$\int_1^y \frac{\left|S_{v_y\cdot\mu}(1,t)\right|}{t^2}dt \le \frac{\beta}{\delta}\log(y) + o(\log(y)),$$
for all
$$\limsup_{x\to\infty}\frac{|M(x)|}{x} < \beta < 2.$$
Even more, if $\limsup_{x\to\infty}\frac{|M(x)|}{x} > 0$*, then* $\delta > 1$.

Proof. Let M be the constant in (10) and $\alpha = \limsup_{x\to\infty}\frac{|M(x)|}{x}$. We set
$$\delta = \min(2, 1+\frac{\alpha^2}{4M}).$$
Note that as also observed earlier we have
$$\int_1^y \frac{\left|S_{v_y\cdot\mu}(1,t)\right|}{t^2}dt = \int_1^y \frac{|M(t)|}{t^2}dt.$$
If $M(t)$ does not change sign in $[1,y]$ then the statement holds trivially as long as y is large enough. As before we fix x_0 such that $M(x) \le \beta x$ for all $x \ge x_0$. We set
$$I(a,b) = \int_a^b \frac{M(t)}{t^2}dt.$$
Let a and b be zeros of $M(x)$ and assume that $x_0 < a < b$. We claim that
$$|I(a,b)| \le \frac{\beta}{\delta}\log(\frac{b}{a}).$$
If this claim is established the statement follows by decomposing the full integral accordingly. We consider several cases.

If $\log(b/a) \ge \frac{\delta M}{\beta}$, we simply have
$$|I(a,b)| \le M \le \frac{\beta}{\delta}\log(b/a)$$
as desired.

If $\log(b/a) \le \frac{\delta M}{\beta}$ and $\frac{b}{a} \le \frac{1}{1-\frac{\beta}{2}}$, then
$$|M(t)| = |M(t)-M(a)| \le |t-a| = t(1-\frac{a}{t}) \le t(1-\frac{a}{b}) \le \frac{\beta}{2}t.$$
We conclude that
$$|I(a,b)| \le \frac{\beta}{2}\int_a^b t^{-1}dt = \frac{\beta}{2}\log(b/a).$$
Finally, if $\log(b/a) \le \frac{\delta M}{\beta}$ and $\frac{b}{a} \ge \frac{1}{1-\frac{\beta}{2}}$, then
$$\begin{aligned}
|I(a,b)| &\le \left|I(a,\frac{a}{1-\beta/2})\right| + \left|I(\frac{a}{1-\beta/2},b)\right| \\
&\le \frac{\beta}{2}\log(\frac{1}{1-\beta/2}) + \beta\log(\frac{b}{a}(1-\beta/2)) \\
&= \beta\log(\frac{b}{a}) + \frac{\beta}{2}\log(1-\frac{\beta}{2}) \le \beta\log(\frac{b}{a}) - \frac{\beta^2}{4} \\
&\le \beta\log(\frac{b}{a}) - M(\delta-1).
\end{aligned}$$

Here we first applied the previous case together with a trivial estimate. Further we used that the elementary inequality $\log(1-\frac{\beta}{2}) \leq -\frac{\beta}{2}$ holds for all $\beta < 2$. In the last step we used the definition of δ which implies $\delta \leq 1 + \frac{\beta^2}{4M}$ and thus $\frac{\beta^2}{4} \geq (\delta - 1)M$. The claim follows since $-M \leq -\frac{\beta}{\delta}\log(b/a)$ by assumption. □

Proposition 1. *We have*

$$\limsup_{x\to\infty} \frac{|M(x)|}{x} = 0.$$

Proof. Put $\alpha = \limsup_{x\to\infty} \frac{|M(x)|}{x}$ By Lemma 5,9 and 11 we have

$$\alpha \leq \prod_{p\leq y}\left(1 - \frac{1}{p}\right)\log(y)\left[\beta(C + \delta^{-1} - 1) + o(1)\right] \tag{11}$$

for all $\alpha < \beta < 2$. We take the limit $y \to \infty$ and using (6) we find

$$\alpha \leq \beta\left(1 + \frac{\delta^{-1} - 1}{C}\right).$$

Taking $\beta \to \alpha$ we get the inequality

$$0 \leq \frac{\delta^{-1} - 1}{C}.$$

Since $C > 0$ we deduce that

$$\delta \leq 1.$$

But now we deduce from Lemma 11 that $\alpha = 0$. □

Theorem 1 (Prime Number Theorem)**.** *We have*

$$\psi(x) \sim x.$$

Proof. We show that the statement follows from Proposition 1.[4] We will use the fact

$$S_d(1, x) = x\log(x) + Kx + O(\sqrt{x}) \text{ and } \sum_{n\leq x} \frac{d(n)}{n} \ll \sqrt{x}. \tag{12}$$

We start by observing the following three convolution identities.

$$\psi(x) = \sum_{n\leq x} \Lambda(n) = \sum_{n\leq x} [\mu \star \log](n) = \sum_{n\leq x}\sum_{d|n} \mu(d)\log(\frac{n}{d}) = \sum_{mn\leq x} \mu(n)\log(m),$$

$$\sum_{mn\leq x} \mu(n)d(m) = \sum_{mn\leq x} \mu(n)\sum_{d|m} 1 = \sum_{dr\leq x}\sum_{n|r} \mu(n) = \sum_{d\leq x} 1 = x + O(1) \text{ and}$$

$$\sum_{mn\leq x} \mu(n) = \sum_{c\leq x}\sum_{n|c} \mu(n) = 1.$$

[4]Actually they are equivalent, but we will not need this fact. The full equivalence is shown in [7].

We put $f_E(m) = \log(m) - d(m) + E$ and observe that combining our identities yields

$$\psi(x) - x + O(1) = \sum_{mn \leq x} \mu(n) f_E(m).$$

Thus it remains to show that the remaining double sum is $o(x)$ for a suitable constant E.

By partial summation one easily sees

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

Thus we find

$$F_E(x) = \sum_{n \leq x} f_E(m) = (E + 1 - K)x + O(\sqrt{x}).$$

We choose $E = K - 1$ and drop the subscript E. In particular we have $F(x) \ll \sqrt{x}$. By the triangle inequality we also have the easy estimate

$$\sum_{n \leq x} \frac{|f(n)|}{n} \ll \sqrt{x}.$$

Write $x = yz$ and write

$$\sum_{mn \leq x} \mu(n) f(n) = \sum_{m \leq y} f(m) M(\frac{x}{m}) + \sum_{n \leq z} \mu(n) F(\frac{x}{n}) - M(z)F(y).$$

Let $\epsilon > 0$ and suppose $M(s) \leq \epsilon s$ for all $s \geq x_0(\epsilon)$. If $x \geq x_0(\epsilon)y$, then $\frac{x}{m} \geq z = \frac{x}{y} \geq x_0(\epsilon)$ such that

$$\left| M(\frac{x}{m}) \right| \leq \epsilon \frac{x}{m} \text{ and } |M(z)| \ll \epsilon z,$$

for all $m \leq y$. This is possible by Proposition 1. We get

$$\left| \sum_{mn \leq x} \mu(n) f(n) \right| \ll \epsilon x \sqrt{y} + \frac{x}{\sqrt{y}} \ll \sqrt{\epsilon} x.$$

In the last inequality we chose $y = \epsilon^{-1}$. Note that our requirement on x now reads $x \geq x_0(\epsilon)\epsilon^{-1}$. The result follows by directly verifying the definition $o(x)$. $\square$

Exercise 2. *Show that* $\pi(x) \sim \frac{x}{\log(x)}$. *Further prove the inequalities* (12) *to complete the prove of Theorem 1.*

Solution. We trivially estimate

$$\sum_{\substack{1 < p^k \leq x, \\ k > 1}} \log(p) \ll \log(x)^2 \sharp\{1 \leq n^2 \leq x\} \leq \sqrt{x} \log(x)^2.$$

Thus we get

$$\tilde{\psi}(x) = \sum_{1 < p \leq x} \log(p) = \psi(x) + O(\sqrt{x} \log(x)^2) \sim x.$$

By partial summation we get

$$\pi(x) = \sum_{2\le p\le x} \frac{\log(p)}{\log(p)} = \frac{\tilde{\psi}(x)}{\log(x)} + \int_2^x \frac{\tilde{\psi}(t)}{t\log(t)^2} \sim \mathrm{li}(x).$$

We conclude by using $\mathrm{li}(x) \sim \frac{x}{\log(x)}$.

We turn towards the divisor estimates and start by deriving a preliminary estimate concerning the harmonic sum. Indeed by writing $\lfloor t\rfloor = t - \{t\}$ and using partial summation we get

$$\begin{aligned}\sum_{n\le z} \frac{1}{z} &= \frac{\lfloor z\rfloor}{z} + \int_1^z \lfloor t\rfloor t^{-2}dt \\ &= \log(z) + 1 - \int_1^\infty \{t\}t^{-2}dt + O(z^{-1}).\end{aligned}$$

We put[5]

$$c = 1 - \int_1^\infty \{t\}t^{-2}dt > \infty.$$

Now we can estimate

$$\begin{aligned}S_d(1,x) &= \sum_{n\le x} d(n) = \sharp\{(u,v)\in\mathbb{N}^2 \colon uv\le x\} \\ &= 2\sum_{\substack{uv\le x,\\ u\ge\sqrt{x}}} 1 + \lfloor\sqrt{x}\rfloor^2 \\ &= 2\sum_{v\le\sqrt{x}}\Big(\frac{x}{v} - \sqrt{x} + O(1)\Big) + x + O(\sqrt{x}) \\ &= 2x\sum_{v\le\sqrt{x}}\frac{1}{v} - x + O(\sqrt{x}) \\ &= x\log(x) + (2c-1)x + O(\sqrt{x}).\end{aligned}$$

The second estimate follows trivially from the bound $d(n)\le\sqrt{n}$ which is obvious. □

[5]Indeed $c=\gamma$ is the Euler-Mascheroni constant which is usually defined through the limit

$$\gamma = \lim_{z\to\infty}\left(\sum_{n\le z}\frac{1}{n} - \log(z)\right).$$

1.3 The sieve of Eratosthenes

Set $P(y) = \prod_{p \leq y} p$ and define

$$\Phi(x, y) = \sharp\{n \leq x \colon (n, P(y)) = 1\}.$$

Further put

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

We define the function $w \colon [1, \infty) \to \mathbb{R}$ recursively by setting

$$w(u) = \frac{1}{u} \text{ for } u \in [1, 2]$$

and requiring[6]

$$\frac{d}{du}(uw(u)) = w(u-1) \text{ for } u \geq 2.$$

Lemma 12. *We have* $\lim_{u \to \infty} w(u) = e^{-\gamma}$

The following proof is taken from [5, Section 4]

Proof. We start by defining the auxiliary function[7]

$$h(u) = \int_0^\infty \exp\left(-ux + \int_0^x \frac{e^{-t} - 1}{t} dt\right) dx. \tag{13}$$

This function is analytic for $u > 0$ and satisfies

$$uh'(u) + h(u+1) = 0.$$

This is easily seen by partial integration as follows. We set $l(x) = \int_0^x \frac{e^{-t}-1}{t} dt$. By the fundamental theorem of calculus we have $l'(x) = (e^{-x} - 1)/x$. In particular, $1 + xl'(x) = e^{-x}$. We compute

$$\begin{aligned} uh'(u) &= \int_0^\infty (-uk) e^{-ux} e^{l(x)} dx = -\int_0^\infty e^{-ux} \frac{d}{dx}[xe^{l(x)}] dx \\ &= -\int_0^\infty e^{-ux}(e^{l(x)} + xl'(x)e^{l(x)}) dx = -\int_0^\infty e^{-(u+1)x + l(x)} dx = -h(u+1). \end{aligned}$$

Similarly partial integration also yields $h(u) \sim \frac{1}{u}$ as $u \to \infty$. We define

$$f(a) = \int_a^{a+1} w(u-1)h(u) du + aw(a)h(a),$$

for $h(u)$ as in (13). Using the definition of w and the functional equation of h one checks that

$$f'(a) = w(a)h(a+1) - w(a-1)h(a) + \underbrace{\frac{d}{da}(aw(a))}_{=w(a-1)} + \underbrace{ah'(a)}_{-h(a+1)} w(a) = 0$$

[6]At $u = 2$ the right-hand derivative is taken.

[7]Note that this is very similar to the function $k(u)$ defined in the previous section!

for all $a > 0$. Thus f is constant. Since $h(a) \sim a^{-1}$ we conclude that the limit $\lim_{u\to\infty} w(u)$ must exist and

$$f(a) = \lim_{u\to\infty} w(u).$$

On the other hand we compute

$$f(2) = \int_2^3 w(u-1)h(u)du + h(2) = -\int_2^3 h'(u-1)du + h(2) = h(1).$$

We further set $F(x) = \exp(\log(x) + \int_0^x (e^{-t} - 1)\frac{dt}{t})$. We will use that $F(0) = 0$ and

$$\lim_{x\to\infty} F(x) = e^{-\gamma}. \tag{14}$$

Indeed that gives

$$h(1) = \lim_{u\to 0} uh'(u) = \lim_{u\to 0}\int_0^\infty (-u)e^{-ux}F(x)dx = \lim_{u\to 0}\int_0^\infty e^{-ux}F'(x)dx = \lim_{x\to\infty} F(x) = e^{-\gamma}.$$

The proof is complete if we can establish (14). To prove this we recall that the Euler-Mascheroni constant is defined by

$$\gamma = \lim_{z\to\infty}\left(\sum_{n\le z}\frac{1}{n} - \log(z)\right).$$

We write

$$H_n = 1 + \ldots + \frac{1}{n} = \int_0^1 1 + \ldots + x^{n-1}dx = \int_0^1 \frac{1-x^n}{1-x}dx.$$

On the other hand we have

$$\int_0^{1-\frac{1}{n}} \frac{1}{1-x}dx = \log(n).$$

Combining these two expressions yields

$$H_n - \log(n) = \int_0^n \frac{\mathbb{1}_{[1-\frac{1}{n},1]}(x) - x^n}{1-x}dx = \int_0^n \frac{\mathbb{1}_{[0,1]}(t) - (1-t/n)^n}{t}dt.$$

Taking the limit we get

$$\gamma = \int_0^\infty \frac{\mathbb{1}_{[0,1]}(t) - e^{-t}}{t}dt.$$

This little trick allows us to write

$$\int_0^x \frac{e^{-t}-1}{t}dt = -\log(x) - \text{Ei(-x)} - \gamma,$$

for $-\text{Ei}(-x) = \int_x^\infty e^{-t}\frac{dt}{t}$. In particular we found $F(x) = e^{-\gamma}\exp(-\text{Ei}(-x))$. We obtain (14), since $\lim_{x\to\infty}\exp(-\text{Ei}(-x)) = 1$. $\square$

Lemma 13 (Lemma 4, [8])**.** *The function* $w(u) - e^{-\gamma}$ *changes sign in any interval* $[a-1, a]$ *for* $a \geq 2$.

Proof. We define another auxiliary function

$$g(a) = \int_a^{a+1} h(u)du + ah(a) \text{ for } a > 0. \tag{15}$$

One can easily check that $\lim_{a\to\infty} g(a) = 1$ and $g'(a) = 0$ for $a > 0$. Indeed

$$\begin{aligned}
&\lim_{a\to\infty} g(a) = 1 + \log(a+1) - \log(a) = 1 \text{ and} \\
&g'(a) = h(a+1) + ah'(a) = 0.
\end{aligned}$$

This implies $g(a) = 1$.

Recall the function f from the previous lemma as well as g from (15). Since $f(a) \equiv e^{-\gamma}$ and $g(a) \equiv 1$ for $a \geq 2$ we have

$$0 = f(a) - e^{-\gamma} g(a) = \int_{a-1}^{a} (w(u) - e^{-\gamma})h(u)du + a(w(a) - e^{-\gamma})h(a-1).$$

Since $h(u) \geq 0$ and $w(u) \not\equiv e^{-\gamma}$ on $[1, \infty)$ the last equality implies the existence of sign changes in the interval $[a-1, a]$. $\square$

Claim: We even have $w(u) = e^{-\gamma} + O(\Gamma(u)^{-1})$.
To see this we note that by the mean value theorem and the differential recursion satisfied by w we get

$$|w'(u)| = \frac{|w(u-1) - w(u)|}{u} \leq u^{-1} \sup_{x\in[u-1,u]} |w'(x)|.$$

Putting $M(u) = \sup_{x\in[u,\infty)} |w'(x)|$ we get

$$M(u) \leq uM(u-1) \ll \Gamma(u+1)^{-1}.$$

With this at hand we can estimate

$$\left|w(u) - e^{-\gamma}\right| = \int_u^\infty |w'(u)| \leq \int_u^\infty \frac{dt}{\Gamma(t+1)} \leq \Gamma(u-1)^{-1} \int_u^\infty \frac{dt}{t(t-1)} \ll \Gamma(u)^{-1}$$

as desired.

The following lemma is of key importance to Maier's argument. It goes back to Buchstab (in Russian), but we follow the argument given in [5].

Lemma 14. *Let* $\lambda > 1$. *Then*

$$\lim_{z\to\infty} \frac{\Phi(z^\lambda, z)}{z^\lambda W(z)} = Cw(\lambda).$$

Proof. We start by observing that by the prime number theorem we can find $R(y)$ with $R(y) \to 0$ as $y \to \infty$ such that[8]

$$|\pi(y) - \mathrm{li}(y)| < \frac{y}{\log(y)} R(y).$$

[8]With our current knowledge this function can not be specified. But Later we will see some functions that are admissible.

The starting point of the argument is Buchstab's identity:

$$\Phi(x,y) = \sum_{y<p\leq y^h} \Phi(\frac{x}{p},p) + \Phi(x,y^h) + O(\frac{x}{y}).$$

To see this we first assume that there is exactly one prime p between y and y^h. In this case we have

$$\Phi(\frac{x}{p},p) + \Phi(x,y^h) = \sharp\{n \leq \frac{x}{p} \colon (n,pP(y)) = 1\} + \Phi(x,y^h)$$
$$= \sharp\{n \leq x \colon p \mid n,\, p^2 \nmid n \text{ and } (n,P(y)) = 1\} + \{n \leq x \colon (n,pP(y)) = 1\}$$
$$= \sharp\{n \leq x \colon p \mid n \text{ and } (n,P(y)) = 1\} + \{n \leq x \colon (n,pP(y)) = 1\} + O(\frac{x}{p^2}) = \Phi(x,y) + O(\frac{x}{p^2}).$$

The general case is for example obtained by induction.

We rewrite this as

$$\psi(x,y) = \sum_{y<p\leq y^h} \frac{1}{p} \prod_{p<p'\leq y} \left(1-\frac{1}{p'}\right)\psi(\frac{x}{p},p) + \prod_{y<p\leq y^h}\left(1-\frac{1}{p}\right)\psi(x,y^h) + O(\frac{\log(y)}{y}).$$

For notational convenience we define

$$G(h) = \sum_{y<p\leq y^h} \frac{1}{p} \prod_{p<p'\leq y}\left(1-\frac{1}{p'}\right) = 1 - \frac{W(y^h)}{W(y)}.$$

With the help of Stieltjes integral[9] we write

$$\psi(z^\lambda,z) = \int_1^h \psi(z^{\lambda-\sigma},z^\sigma)dG(\sigma) + \psi(z^\lambda,z^h)(1-G(h)) + O(\frac{\log(z)}{z}).$$

By Merten's estimate for the product $W(x)$ we find $\frac{W(y^\sigma)}{W(y)} = \frac{1}{\sigma} + o(\frac{1}{\sigma})$. If necessary we replace our function R with something larger that still tends to 0 at infinity and find[10]

$$\left|G(\sigma) - 1 + \frac{1}{\sigma}\right| < \frac{R(y)}{\sigma} \leq R(y) \text{ for } \sigma \geq 1.$$

We now define

$$\theta(z^\lambda,z) = C\log(z)\int_1^\lambda z^{t-\lambda}w(t)dt.$$

[9]The Riemann-Stieltjes integral is defined similarly to the common Riemann integral by

$$\int_a^b f(x)dh(x) = \lim_{x_{i+1}-x_i\to 0}\sum_i \sup_{x\in[x_i,x_{i+1}]} f(x)[h(x_{i+1})-h(x_i)] = \lim_{x_{i+1}-x_i\to 0}\sum_i \inf_{x\in[x_i,x_{i+1}]} f(x)[h(x_{i+1})-h(x_i)].$$

The following two properties are essentially everything one should remember

$$\int_a^b f(x)dh(x) = \int_a^b f(x)h'(x) \text{ if } h \in \mathcal{C}^1((a,b))$$

as well as the partial integration formula

$$\int_a^b f(x)dh(x) = f(b)h(b) - f(a)h(a) - \int_a^b h(x)df(x).$$

[10]One can actually show this asymptotic using the prime number theorem so that the following estimate holds with the same R. But this is not of interest to us.

The upcoming slightly involved argument intends to show that $\theta(z^\lambda, z)$ is a good approximation to $\psi(z^\lambda, z)$.

We will first treat several auxiliary expressions. First look at

$$E_1(\lambda, z, h) = \theta(z^\lambda, z) - \int_1^h \theta(z^{\lambda-\sigma}, z^\sigma)\frac{d\sigma}{\sigma^2} - \frac{\theta(z^\lambda, z^h)}{h}.$$

Of course $E_1(\lambda, z, 1) = 0$. On the other hand we can compute the derivative using the differential recursion satisfied by w. We get

$$\frac{\partial}{\partial h}E_1(\lambda, z, h) = -\theta(z^{\lambda-h}, z^h)h^{-2} + \theta(z^\lambda, z^h)h^{-2} - h^{-1}\frac{\partial}{\partial h}\theta(z^\lambda, z^h)$$
$$= -\theta(z^{\lambda-h}, z^h)h^{-2} + \theta(z^\lambda, z^h)h^{-2} - \frac{C\log(z)}{h}z^{h-\lambda} + \frac{C\log(z)}{h}\int_1^{\frac{\lambda}{h}} w(t-1)z^{ht-\lambda}dt$$
$$= -C\log(z)\frac{z^{h-\lambda}}{h}.$$

We get

$$|E_1(\lambda, z, h)| = C\log(z)\int_1^h \frac{z^{t-\lambda}}{t}dt \leq Cz^{h-\lambda}.$$

Further put

$$E_2(\lambda, z, h) = -\int_1^h \theta(z^{\lambda-\sigma}, z^\sigma)d(G(\sigma) - 1 + \frac{1}{\sigma}) - \theta(z^\lambda, z^h)(1 - \frac{1}{h} - G(h)).$$

Using our bound for $G(\sigma)$ we obtain by partial integration that

$$|E_2(\lambda, z, h)| \leq R(z)\left(\left|\theta(z^\lambda, z^h) - \theta(z^{\lambda-h}, z^h)\right| + \int_1^h \left|\frac{d}{d\sigma}\theta(z^{\lambda-\sigma}, z^\sigma)\right| d\sigma\right)$$

We estimate E_2 trivially using

$$\left|\theta(z^\lambda, z) - Ce^{-\gamma}\right| \ll \Gamma(\lambda-1)^{-1} + z^{-1}, \qquad (16)$$

which simply follows by inserting $w(\lambda) = e^{-\gamma} + O(\Gamma(\lambda)^{-1})$ into the definition of $\theta(z^\lambda, z)$. This yields

$$E_2(\lambda, z, h) \ll \frac{R(z)}{\lambda^2},$$

for suitable h. Finally we put

$$E_3(\lambda, z, h) = \theta(z^\lambda, z) - \int_1^h \theta(z^{\lambda-\sigma}, z^\sigma)dG(\sigma) - \theta(z^\lambda, z^h)(1 - G(h)) = E_1(\lambda, z, h) - E_2(\lambda, z, h).$$

From the estimates for E_1 and E_2 we conclude that

$$\left|E_3(\lambda, z, \frac{\lambda}{k})\right| \ll \frac{R(z)}{k^2} \text{ for } k \leq \lambda < k+1 \text{ and } z \geq 2. \qquad (17)$$

We are now ready to consider the difference

$$\eta(z^\lambda, z) = \psi(z^\lambda, z) - \theta(z^\lambda, z).$$

This difference satisfies

$$\eta(z^\lambda, z) = \int_1^h \eta(z^{\lambda-\sigma}, z^\sigma) dG(\sigma) + \eta(z^\lambda, z^h)(1 - G(h)) - E_3(\lambda, z, h) + O(\frac{\log(z)}{z}).$$

Note that for $1 \leq \lambda \leq 2$ we simply have

$$\psi(z^\lambda, z) = \frac{\pi(z^\lambda) - \pi(z)}{z^\lambda P(z)} \tag{18}$$

We obtain

$$\eta(z^\lambda, z) = \frac{\mathrm{li}(z^\lambda) - \mathrm{li}(z) - \int_1^\lambda z^t \frac{dt}{t}}{z^\lambda P(z)} + O(R(z)) = O(R(z)).$$

in this region. Here we used

$$\int_1^\lambda z^t \frac{dt}{t} = \mathrm{Ei}(\lambda \log(z)) - \mathrm{Ei}(\log(z)) = \mathrm{li}(z^\lambda) - \mathrm{li}(z).$$

For $k \geq 1$ we now set

$$s_k(z) = \sup_{\substack{k \leq u < k+1, \\ x \geq z}} |\eta(x^u, x)| .$$

We have just see that $s_1(z) \ll R(y)$ and by (17) we get

$$s_k(z) < s_{k-1}(z) + O(\frac{R(z)}{k^2}).$$

A direct consequence is the estimate $s_k(z) \ll R(z)$ and we deduce that

$$\psi(z^\lambda, z) = \theta(z^\lambda, z) + O(R(z)).$$

However, partially integrating twice shows that[11]

$$\theta(z^\lambda, z) = Cw(\lambda)(1 + o(1)) \text{ as } z \to \infty.$$

This concludes the proof. $\square$

Note that we have actually proved

$$\left|\psi(z^\lambda, z) - \theta(z^\lambda, z)\right| \ll R(z)$$

uniformly in λ. Combining this with (16) we even get

$$\left|\psi(z^\lambda, z) - Ce^{-\gamma}\right| \ll R(z) + \Gamma(u)^{-1}. \tag{19}$$

Remark 1. *It can be shown that $C = e^\gamma$. Indeed one argument goes as follows. In the exercise below it is shown that $\psi(z^\lambda, z) = 1 + O(\frac{2^z}{z^\lambda W(z)})$. Thus for fixed z we have $\lim_{\lambda\to\infty} \psi(z^\lambda, z) = 1$. On the other hand we have seen that*

$$\lim_{x\to\infty} \psi(z^\lambda, z) = Cw(\lambda).$$

[11] The $o(1)$ term appearing below is not uniform in λ!

The proof did actually also produce the uniform estimate (19)*, so that the following interchange of limits is justified:*

$$
\begin{aligned}
1 = \lim_{z\to\infty} 1 = \lim_{z\to\infty} \lim_{\lambda\to\infty} \frac{\Phi(z^\lambda, z)}{z^\lambda W(z)} \\
= \lim_{\lambda\to\infty} \lim_{z\to\infty} \frac{\Phi(z^\lambda, z)}{z^\lambda W(z)} = C \lim_{\lambda\to\infty} w(\lambda) = Ce^{-\gamma}.
\end{aligned}
$$

In particular we obtain $C = e^\gamma$ as claimed. Of course one can also prove Merten's formula directly with the correctly identified constant.

Exercise 3. *Show Legendre's formula*

$$\Phi(x, y) = \sum_{d|P(y)} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor$$

and deduce the estimate

$$|\Phi(x, y) - xW(y)| < 2^y.$$

for $y \geq 2$.

Solution. Let us first show Legendre's identity for $y = 2$. Here we obviously have

$$\Phi(x, y) = \sharp\{n \leq x\colon \text{odd}\} = \lfloor x \rfloor - \lfloor \frac{x}{2} \rfloor,$$

which is the desired identity since $P(2) = 2$ has divisors 1 and 2. The general formula follows by writing

$$\{n \leq x\colon (n, P(y)) = 1\} = \{n \leq x\} \setminus \bigcup_{p\leq y} \{n \leq x\colon p \mid n\}.$$

By inclusion exclusion we get

$$
\begin{aligned}
\Phi(x, y) &= \sharp\{n \leq x\} - \sum_{p\leq y} \sharp\{n \leq x\colon p \mid n\} + \sum_{\substack{p_1,p_2\leq y,\\ p_1\neq p_2}} \sharp\{n \leq x\colon p_1p_2 \mid n\} - \dots \\
&= \lfloor x \rfloor - \sum_{p\leq y} \lfloor \frac{x}{p} \rfloor + \sum_{\substack{p_1,p_2\leq y,\\ p_1\neq p_2}} \lfloor \frac{x}{p_1p_2} \rfloor - \dots.
\end{aligned}
$$

This completes the proof.

We now turn to the desired estimate. Note that

$$xW(y) = x \prod_{p\leq y} \left(1 - \frac{1}{p}\right) = \sum_{d|P(y)} \mu(d) \frac{x}{d}.$$

This is a similar computation we have seen before. (Indeed we have $\mathcal{S}(y) = \{d \mid P(y)\}$.) The estimate is now essentially trivial:

$$|\Phi(x, y) - xW(y)| = \left| \sum_{d|P(y)} \mu(d) \left(\frac{x}{d} - \lfloor \frac{x}{d} \rfloor\right) \right| \leq \sharp\{d \mid P(y)\} < 2^y.$$

$\square$

1.4 Dirichlet series and their basic properties

Given an arithmetic function $f \in \mathcal{A}$ we associate (at the moment formally) the **Dirichlet series**

$$D_f(s) = \sum_{n\in\mathbb{N}} f(n)n^{-s}.$$

Important Dirichlet series with special names are $\zeta(s) = D_\epsilon(s)$, $L(s,\chi) = D_\chi(s)$. By formally multiplying such series one finds that

$$D_f(s)D_g(s) = D_{f\star g}(s).$$

One can show that $\frac{1}{\zeta(s)} = D_\mu(s)$ and $-\frac{\zeta'(s)}{\zeta(s)} = D_\Lambda(s)$. Similarly there are many more interesting relations to play with.

Lemma 15. *Let $f \in \mathcal{A}$ be multiplicative. If $D_f(s_0)$ converges absolutely for some $s_0 \in \mathbb{C}$, then $D_f(s_0)$ has the* ***Euler product****:*

$$D_f(s_0) = \prod_p \sum_{k\in\mathbb{Z}_{\geq 0}} f(p^k)p^{-ks_0}.$$

Furthermore, if f is completely multiplicative we have

$$\sum_{k\in\mathbb{Z}_{\geq 0}} f(p^k)p^{-ks_0} = \frac{1}{1-f(p)p^{-s_0}}.$$

Proof. Given $\epsilon > 0$ we take K large such that $\sum_{n>K} |f(n)n^{-s_0}| < \epsilon$. Let

$$\mathcal{P}_K = \{p \text{ prime}\colon p \leq K\}.$$

The fundamental theorem of arithmetic implies

$$\prod_{\mathcal{P}_K} \sum_{k\in\mathbb{Z}_{\geq 0}} f(p^k)p^{-ks_0} = \sum_{\substack{n\in\mathbb{N},\\ p|n \implies p\in\mathcal{P}_K}} f(n)n^{-s_0}.$$

Of course this implies

$$\left|\prod_{\mathcal{P}_K} \sum_{k\in\mathbb{Z}_{\geq 0}} f(p^k)p^{-ks_0} - D_f(s_0)\right| \leq \sum_{n>K} \left|f(n)n^{-s_0}\right| < \epsilon.$$

We complete the proof by taking $\epsilon \to 0$ and $K \to \infty$. $\square$

We now turn towards analytic properties of Dirichlet series. We start by addressing some convergence issues.

Lemma 16. *Let $f \in \mathcal{A}$ and take $s_0 \in \mathbb{C}$ such that $D_f(s_0)$ converges. Then, for every $\delta > 0$, the series $D_f(s)$ converges uniformly in the region*

$$G_\delta = \{s \in \mathbb{C}\colon\ |\arg(s-s_0)| \leq \frac{\pi}{2} - \delta\}.$$

Proof. By partial summation we have

$$\sum_{M\le n\le N} f(n)n^{-s} = N^{s_0-s}\sum_{M\le n\le N} a(n)n^{-s_0}+(s-s_0)\int_M^N\left(\sum_{M\le n\le x} f(n)n^{-s_0}\right)x^{s_0-s-1}dx.$$

Since $D_f(s_0)$ converges there is $M=M(\epsilon)$ such that

$$\left|\sum_{M\le n\le N} a(n)n^{-s_0}\right| < \epsilon \text{ for all } N>M.$$

Estimating everything trivially yields

$$\left|\sum_{M\le n\le N} f(n)n^{-s}\right| < \epsilon N^{\sigma_0-\sigma}+\epsilon\frac{|s_0-s|}{\sigma_0-\sigma}(N^{\sigma_0-\sigma}-M^{\sigma_0-\sigma}) < \epsilon\left(1+\frac{2}{\sin(\delta)}\right).$$

Here $\sigma=\Re(s)$, $\sigma_0=\Re(s_0)$ and we used that $\sigma_0-\sigma<0$ as well as $\frac{|s_0-s|}{|\sigma_0-\sigma|}<\frac{1}{\sin(\delta)}$. This concludes the proof. □

Corollary 1. *Let $f\in\mathcal{A}$ and suppose that $D_f(s_0)$ converges for $s_0\in\mathbb{C}$. Then $D_f(s)$ is holomorphic in the half plane $\Re(s)>\Re(s_0)$.*

Proof. This is a simple consequence of Morera's theorem and Lemma 16. □

For $D_f(s)$ we define the **abscissa of convergence** by

$$\inf\{\alpha\in\mathbb{R}\colon D_f(\alpha) \text{ converges}\}.$$

The following result is analogue to the identity theorem for power series.

Lemma 17. *Let $f,g\in\mathcal{A}$. Suppose that for all sufficiently large $\alpha\in\mathbb{R}$ the Dirichlet series $D_f(\alpha)$ and $D_g(\alpha)$ converge and satisfy $D_f(\alpha)=D_g(\alpha)$. Then we must have $f=g$.*

Proof. We have

$$f(1)=\lim_{\alpha\to\infty} D_f(s)=\lim_{\alpha\to\infty} D_g(s)=g(1).$$

The interchange of sum and limit is justified by Lemma 16. One argues inductively using the identity

$$f(n)=\lim_{\alpha\to\infty}\left[n^{\alpha}D_f(\alpha)-\sum_{m\le n-1} f(m)n^{\alpha}m^{-\alpha}\right].$$

□

Lemma 18. *Define the indicator like function*

$$\delta(x)=\begin{cases}0 & \text{if } 0<x<1,\\ \frac{1}{2} & \text{if } x=1,\\ 1 & \text{if } 1<x\end{cases}$$

and the integral

$$I(x,T) = \frac{1}{2\pi i}\int_{(c)_T} x^s \frac{ds}{s},$$

here $(c)_T$ is the vertical line from $c-iT$ to $c+iT$. For $c,x,T>0$ we have

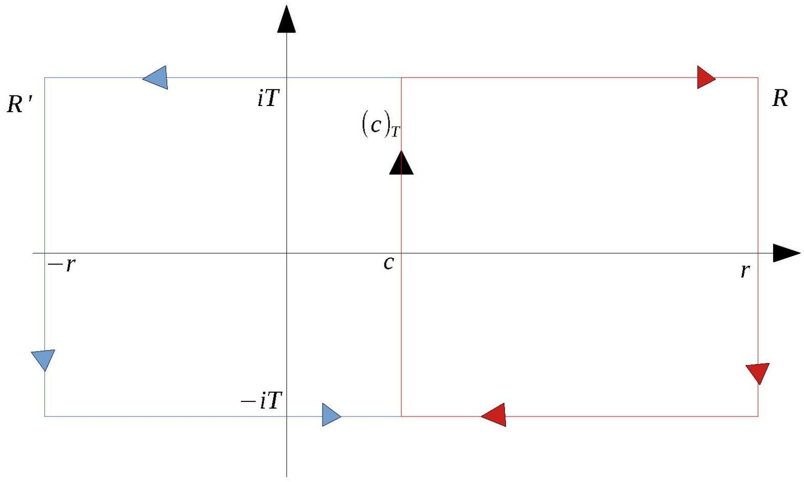
$$|I(x,T)-\delta(x)| < \begin{cases} x^c \max(1, T\,|\log(x)|)^{-1} & \text{if } x \neq 1, \\ \frac{c}{T} & \text{if } x = 1. \end{cases}$$

Proof. We treat the different cases separately. First, consider $x=1$. In this case we simply compute

$$I(1,T) = \frac{1}{2\pi}\int_{-T}^{T} \frac{dt}{c+it} = \frac{1}{\pi}\int_0^T \frac{c}{c^2+t^2}dt = \frac{1}{2} - \frac{1}{\pi}\int_{T/c}^{\infty}\frac{dx}{1+x^2}.$$

The estimate follows directly.

Second we look at $0<x<1$. Define the rectangle R with corners $\{c \pm iT, r \pm iT\}$ for some $r>c$.



By Cauchy's integral formula we have

$$\frac{1}{2\pi i}\int_R x^s\frac{ds}{s} = 0.$$

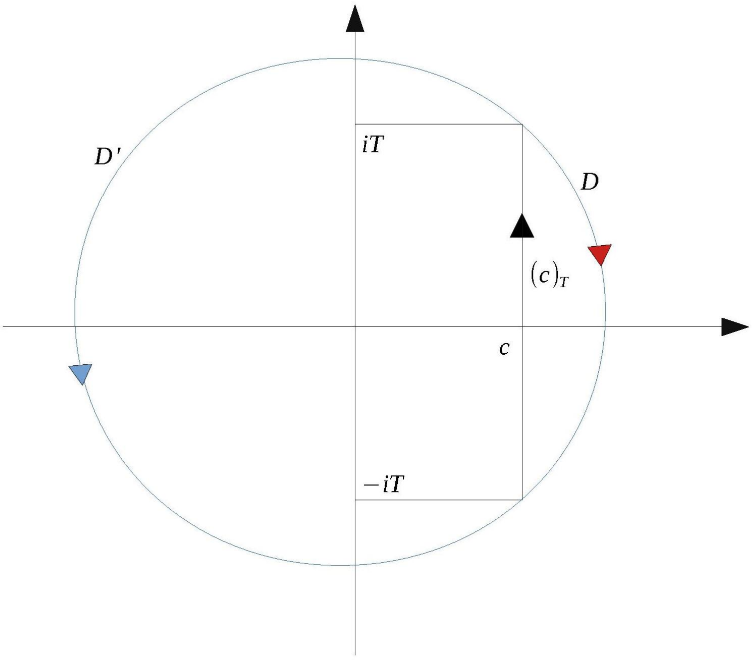
Taking the limit $r\to\infty$ we are left with

$$I(x,T) = -\frac{1}{2\pi i}\int_{c+iT}^{\infty+iT}\frac{x^s}{s}ds + \frac{1}{2\pi i}\int_{c-iT}^{\infty-iT}\frac{x^s}{s}ds.$$

Estimating trivially yields

$$|I(x,T)| \leq \frac{1}{\pi T}\int_c^{\infty} x^{\sigma}d\sigma < \frac{x^c}{T\,|\log(x)|}. \tag{20}$$

The secondary estimate $|I(x,T)| < x^c$ is obtained by replacing the curve R by the curve that supplements the line segment $[c-iT, c+iT]$ with a circle segment.



The rest of the argument works similarly. We get

$$0 = \frac{1}{2\pi i}\int_D x^s \frac{ds}{s} = I(x,T) + \frac{1}{2\pi i}\int_{D\setminus(c)_T} x^s \frac{ds}{s}.$$

The part of the circle integral can be estimated by

$$\left|\frac{1}{2\pi i}\int_{D\setminus(c)_T} x^s \frac{ds}{s}\right| \le \frac{x^\sigma}{2\pi}\int_{\theta_1}^{\theta_2} d\theta \le x^\sigma.$$

Here θ_1 and θ_2 are simply the two angles that define the circle segment $D\setminus(c)_T$.

Finally consider $x > 1$. In this case we take the rectangle R' with the corners $\{-r \pm iT, c \pm iT\}$. Applying the residual theorem and taking $r \to \infty$ we end up with

$$I(x,T) = 1 + \frac{1}{2\pi i}\int_{-\infty+iT}^{c+iT} \frac{x^s}{s} ds - \frac{1}{2\pi i}\int_{-\infty-iT}^{c-iT} \frac{x^s}{s} ds.$$

One estimates as above. The complementary estimate also works analogously by using the curve D' which is analogue to D. ☐

Theorem 2 (Perron's formula (in explicit form))**.** *Let $c > 0$ and $x, T \ge 2$. Suppose $f \in \mathcal{A}$ such that $D_f(c)$ converges absolutely. Then we have*

$$S_f(1,x) - \frac{f(x)}{2}\delta_{x\in\mathbb{N}} = \frac{1}{2\pi i}\int_{c-iT}^{c+iT} D_f(s) x^s \frac{ds}{s} + O\left(\frac{x^c}{T} D_{|f|}(c) + A_x(1 + \frac{x\log(x)}{T})\right),$$

for $A_x = \max_{\frac{3}{4}x \le n \le \frac{5}{4}x} |f(n)|$.

Proof. Using Lemma 18 we immediately get

$$S_f(1,x) - \frac{f(x)}{2}\delta_{x\in\mathbb{N}} = \frac{1}{2\pi i}\int_{c-iT}^{c+iT} D_f(s) x^s \frac{ds}{s} + E$$

with

$$|E| \leq x^c \sum_{n=1}^{\infty} |f(n)|\, n^{-c} \min(1, T^{-1} \left|\log(\frac{x}{n})\right|^{-1}).$$

We split the n-sum in 3 parts. The first part is $|n-x| < 2$, this part can be estimated trivially and contributes at most $\ll A_x$. The second part is estimated as follows:

$$x^c \sum_{2\leq|n-x|<\frac{1}{4}x} |f(n)|\, n^{-c} \min(1, T^{-1} \left|\log(\frac{x}{n})\right|^{-1}) \ll A_x \frac{x^c}{T} \sum_{2\leq|n-x|<\frac{1}{4}x} n^{-c} \left|\log(\frac{x}{n})\right|^{-1}$$
$$\ll A_x \frac{x\log x}{T}.$$

Finally, for $|n-x| \geq \frac{1}{4}x$ we have $\left|\log(\frac{x}{n})\right|^{-1} \ll 1$. Thus, this part of the sum yields the first part of the error term and we are left with estimating the rest. We estimate the terms $|n-x| \leq 2$ trivially by A_x $\square$

Exercise 4. *Show that the constant $C = e^{\gamma}$, where C is the constant given in* (6).

Solution. Since ϵ is completely multiplicative and $\zeta(s) = D_\epsilon(s)$ we have the famous Euler product

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

Since the harmonic series $\sum_n \frac{1}{n} = \zeta(1)$ diverges we have a pole of order 1 and one obtains the estimate[12]

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

Taking logarithms yields

$$\sum_p \log(1-p^{-s}) = \log(s-1) + O(|s-1|).$$

We now insert $s = 1 + \frac{\epsilon}{\log(x)}$. Reading the equation above backwards yields

$$-\log(\frac{1}{\epsilon}) - \log\log(x) + o(1) = \sum_p \log(1 - \frac{1}{p^{1+\frac{\epsilon}{\log(x)}}}).$$

We will treat the sum on the right using the following claim which we proof at the end:
Claim I: For a compactly supported Riemann-integrable function independent of x we have

$$\sum_p \frac{\log(p)}{p} F\left(\frac{\log(p)}{\log(x)}\right) = \log(x) \int_0^\infty F(t)dt + o(\log(x)).$$

[12] For $\Re(s) > 1$ we have $n^{-s} = \int_n^{n+1} t^{-s}dt + O(n^{-2})$. Summing this up yields $\zeta(s) = \int_1^\infty t^{-s}dt + O(\sum_n n^{-2}) = \frac{1}{s-1} + O(1)$.

We put $F = \frac{\exp(-\epsilon t)}{t}$ using the Taylor expansion of the logarithm we find

$$\sum_{p>x} \log(1 - \frac{1}{p^{1+\frac{\epsilon}{\log(x)}}}) = -\sum_{p>x} \frac{1}{p^{1+\frac{\epsilon}{\log(x)}}} + o(1)$$
$$= -\log(x)^{-1} \sum_{p>x} \frac{\log(p)}{p} F(\frac{\log(p)}{\log(x)}) + o(1)$$

We now use our Claim I to write

$$-\log(x)^{-1} \sum_{x<p<x^N} \frac{\log(p)}{p} F_N(\frac{\log(p)}{\log(x)}) = -\int_1^N F_N(x)dx + o(1).$$

for fixed N. The tail can be bounded using (2) and taking $N \to \infty$ we get

$$\sum_{p>x} \log(1 - \frac{1}{p^{1+\frac{\epsilon}{\log(x)}}}) = -\int_\epsilon^\infty e^{-t} \frac{dt}{t}.$$

We now make a second claim:
Claim II: We have

$$\int_\epsilon^\infty e^{-t} \frac{dt}{t} = \log(\frac{1}{\epsilon}) - \gamma + O(\epsilon).$$

Thus we obtain

$$-\log\log(x) - \gamma + O(\epsilon) + o(1) = \sum_{p \le x} \log(1 - \frac{1}{p^{1+\frac{\epsilon}{\log(x)}}})$$
$$= \sum_{p \le x} \left(\log(1 - p^{-1}) + O(\frac{\epsilon \log(p)}{p \log(x)}) \right)$$
$$= \sum_{p \le x} \log(1 - \frac{1}{p}) + O(\epsilon).$$

Thus taking ϵ arbitrarily small and exponentiating this asymptotic we find

$$\prod_{p \le x} \left(1 - \frac{1}{p}\right) = \frac{e^{o(1)}}{e^\gamma \log(x)} = \frac{(1 + o(1))}{e^\gamma \log(x)}.$$

This solves the exercise modulo our two claims.

Proof of Claim I: Without loss of generality we can assume that F is smooth. By Fourier inversion we have

$$F(t) = \int_{\mathbb{R}} e^{-(1+iu)t} f(u) du$$

for a Schwartz function f. We thus write

$$\sum_p \frac{\log(p)}{p} F\left(\frac{\log(p)}{\log(x)}\right) = \int_{\mathbb{R}} f(u) \left(\sum_p \frac{\log(p)}{p^{1+\frac{1+iu}{\log(x)}}} \right) du.$$

We can estimate

$$\sum_p \frac{\log(p)}{p^{1+\frac{1+iu}{\log(x)}}} \ll \log(x)$$

easily. However, we can also look at

$$\frac{1}{s-1} + O(1) = \frac{\zeta'(s)}{\zeta(s)} = \sum_p \frac{\log(p)}{p^s - 1} = \sum_p \frac{\log(p)}{p^s} + O(1).$$

With this hat hand we derive

$$\sum_p \frac{\log(p)}{p} F\left(\frac{\log(p)}{\log(x)}\right) = \log(x) \int_{\mathbb{R}} \frac{f(u)}{1+iu} du + O(1) \tag{21}$$

The claim follows by Fourier inversion.

Proof of Claim II: We write

$$H_n = 1 + \ldots + \frac{1}{n} = \int_0^1 1 + \ldots + x^{n-1} dx = \int_0^1 \frac{1-x^n}{1-x} dx.$$

We also have

$$\int_0^{1-\frac{1}{n}} \frac{1}{1-x} = \log(n).$$

Thus we get

$$H_n - \log(n) = \int_0^n \frac{\mathbb{1}_{[1-\frac{1}{n},1]}(x) - x^n}{1-x} dx = \int_0^n \frac{\mathbb{1}_{[0,1]}(t) - (1-t/n)^n}{t} dt.$$

Taking the limit (and justifying the interchange) we get

$$\gamma = \int_0^\infty \frac{\mathbb{1}_{[0,1]}(t) - e^{-t}}{t} dt.$$

We derive

$$\int_0^\infty \frac{e^{-t} - \mathbb{1}_{[0,\epsilon]}(t)}{t} dt = \log(\frac{1}{\epsilon}) - \gamma.$$

From this one easily derives the claim. Note that this also implies (14), which we used earlier. □

2 Part 2: Basic properties of $\zeta(s)$ and $L(s,\chi)$

We start by a simple consequence of Poisson-summation. Recall that for $f \in \mathcal{S}(\mathbb{R})$ we have[13]

$$\sum_{n\in\mathbb{Z}} f(n+\alpha) = \sum_{n\in\mathbb{Z}} \hat{f}(n)e(n\alpha). \tag{22}$$

Lemma 19. *The function* $\Theta_\alpha(x) = \sum_{n\in\mathbb{Z}} e^{-\pi x(n+\alpha)^2}$ *satisfies*

$$\Theta_\alpha\left(\frac{1}{x}\right) = \sqrt{x}\sum_{n\in\mathbb{Z}} e^{-\pi x n^2}e(n\alpha),$$

for $\alpha \in (-1,1)$.

Proof. Put $f_x(\xi) = e^{-\pi\frac{\xi^2}{x}}$. It is well known that

$$\hat{f}_x(\xi) = \sqrt{x}e^{-\pi x\xi^2}.$$

The statement follows after applying (22) to $\Theta_\alpha(\frac{1}{x}) = \sum_{n\in\mathbb{Z}} f_x(n+\alpha)$. □

With this at hand we can deduce some crucial properties of the **Riemann-zeta function**.

Theorem 3. *The function $\zeta(s)$ has a meromorphic continuation to $\mathbb{C}$ with a unique simple pole with residue 1 at $s=1$. Furthermore it satisfies the functional equation*

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

Proof. In the region of uniform and absolute convergence $\Re(s) > 1$ we write

$$\pi^{-\frac{2}{s}}\Gamma(\frac{s}{2})\zeta(s) = \int_0^\infty \sum_{n=1}^\infty e^{-\pi n^2 y}y^{\frac{s}{2}-1}dy. \tag{23}$$

This follows simply from the definition of the Gamma-function and a change of variables. Expressing the sum in terms of $\Theta_0(y)$ we get

$$\begin{aligned}\pi^{-\frac{2}{s}}\Gamma(\frac{s}{2})\zeta(s) &= \int_0^\infty \frac{\Theta_0(y)-1}{2}y^{\frac{s}{2}-1}dy \\ &= \int_1^\infty \frac{\Theta_0(y)-1}{2}y^{\frac{s}{2}-1}dy + \int_1^\infty \frac{\Theta_0(\frac{1}{y})-1}{2}y^{-\frac{s}{2}-1}dy.\end{aligned}$$

Applying Lemma 19 with $\alpha = 0$ yields

$$\pi^{-\frac{2}{s}}\Gamma(\frac{s}{2})\zeta(s) = \int_1^\infty \frac{\Theta_0(y)-1}{2}[y^{\frac{s}{2}-1} + y^{-\frac{s}{2}-\frac{1}{2}}]dy + \frac{1}{2}\int_1^\infty [[y^{-\frac{s}{2}-\frac{1}{2}} - y^{\frac{s}{2}-1}]]dy.$$

Computing the second y-integral yields the all important formula

$$\pi^{-\frac{2}{s}}\Gamma(\frac{s}{2})\zeta(s) = -\frac{1}{s} + \frac{1}{s-1} + \int_1^\infty \frac{\Theta_0(y)-1}{2}[y^{\frac{s}{2}-1} + y^{-\frac{s}{2}-\frac{1}{2}}]dy.$$

[13] One can weaken the assumption on f considerable. For example $f \in \mathcal{C}^2(\mathbb{R})$ and $f(x) \ll |x|^2$ suffices.

Since the remaining y-integral converges for all $s \in \mathbb{C}$ we obtain the desired analytic properties of $\zeta(s)$. Further, the right hand side of the final equation is invariant under the transformation $s \leftrightarrow 1-s$, which implies the functional equation. $\square$

This leads us to define the **completed zeta-function** by

$$\Lambda(s) = \pi^{-\frac{2}{s}}\Gamma(\frac{s}{2})\zeta(s).$$

Theorem 4. *The completed zeta function has no zeros outside the strip* $0 \leq \Re(s) \leq 1$ *and satisfies*

$$s(s-1)\Lambda(s) = e^{Bs}\prod_{\rho\in\mathcal{N}}\left(1-\frac{s}{\rho}\right)e^{\frac{s}{\rho}}.$$

Proof. For $\Re(s) > 1$ there can not be any zeros of $\zeta(s)$. This can be easily seen by looking at the Euler product

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

Now we note that also $\Gamma(\frac{s}{2})$ and $\pi^{-\frac{s}{2}}$ do not have any zeros in that region. To exclude zeros with $\Re(s) < 0$ one uses the functional equation.

Now the product expansion follows from Hadamard's product theorem after we establish that Λ has order 1. The latter follows from

$$(s-1)\zeta(s) \ll |s|^2,$$

for $\Re(s) \geq \frac{1}{2}$ which can be seen by elementary means. The pre-factor is originally given by e^{A+Bs}, but we can compute

$$e^A = \lim_{s\to 0} s(s-1)\Lambda(s) = 1.$$

$\square$

Lemma 20. *In the region* $|t| \geq 8$ *and* $1 - \frac{1}{2\log(|t|)} \leq \sigma \leq 2$ *we have*

$$\zeta(\sigma+it) \ll \log(|t|) \text{ and } \zeta'(\sigma+it) \ll \log(|t|)^2.$$

Proof. We apply partial summation to

$$\sum_{n>N}^{\infty} n^{-s} = -N^{-s} + s\int_N^\infty [x]x^{-s-1}dx = -N^{-s} + \frac{s}{s-1}N^{1-s} - s\int_N^\infty \{x\}x^{-s-1}dx.$$

Note that the remaining integral converges for all $\Re(s) > 0$. Thus by the principle of analytic continuation we find that

$$\begin{aligned}\zeta(s) &= \sum_{n\leq N} n^{-s} + \frac{s}{s-1}N^{1-s} - N^{-s} - s\int_N^\infty \{x\}x^{-s-1}dx \\ &= \sum_{n\leq N} n^{-s} + \frac{s}{s-1}N^{1-s} + O(N^{-\sigma}(1+\frac{|s|}{\sigma})), \qquad (24)\end{aligned}$$

for all s with $\Re(s) > 0$. Where the integral has been estimated trivially.

Of course for $\sigma \geq 1 - \frac{1}{\log(|t|)}$ we have $n^{-\sigma} \leq en^{-1}$ as long as $n \leq |t|$. Thus choosing $N = [|t|]$ in (24) we get

$$\zeta(s) \ll \sum_{n<|t|} n^{-1} + \frac{|t|^{1-\sigma}}{\sigma} \ll \log(|t|).$$

For $r = (2\log(|t|))^{-1}$ we use Chauchy's integral formula to estimate

$$\zeta'(s) = \frac{1}{2\pi i} \int_{\partial B_r(s)} \frac{\zeta(z)}{(z-s)^2} dz \ll \log(|t|)^2.$$

□

Lemma 21. *There is a $\delta > 0$ such that in the region $s = \sigma + it$ with $\sigma \geq 1 - \delta \min(1, \log(|t|)^{-9})$ we have $\zeta(s) \neq 0$. In the same region with the additional assumption $|s-1| \geq 2$ we have*

$$\frac{\zeta'}{\zeta}(s) \ll \log(|t|)^9.$$

Proof. Using the Euler product of $\zeta(s)$ with the Taylor expansion of the logarithm at 1 we find

$$\zeta(s) = \exp\left(-\sum_p \log(1-p^{-s})\right) = \exp\left(\sum_p \sum_{k=1}^{\infty} \frac{1}{k} p^{-ks}\right).$$

Taking absolute values, which amounts to taking the real part in the exponential, we obtain

$$|\zeta(\sigma+it)| = \exp\left(\sum_p \sum_{k=1}^{\infty} \frac{1}{k} p^{-k\sigma} \cos(kt\log(p))\right).$$

The key observation is the following elementary inequality

$$3 + 4\cos(x) + \cos(2x) = 2(1+\cos(x))^2 \geq 0.$$

Applying this with $x = kt\log(p)$ yields

$$\begin{aligned}\zeta(\sigma)^3 |\zeta(\sigma+it)|^4 |\zeta(\sigma+2it)| &= \exp\left(\sum_p \sum_{k=1}^{\infty} \frac{1}{k} p^{-k\sigma} \left[3 + 4\cos(kt\log(p)) + \cos(2kt\log(p))\right]\right) \\ &\geq \exp(0) = 1.\end{aligned}$$

We conclude that for $t > 8$ and $1 < \sigma < 2$ we have

$$|\zeta(\sigma+it)|^{-1} \leq \zeta(\sigma)^{\frac{3}{4}} |\zeta(\sigma+2it)|^{\frac{1}{4}} \ll (\sigma-1)^{-\frac{3}{4}} \log(t)^{\frac{1}{4}}.$$

We will now use the estimate

$$\zeta(1+it) - \zeta(\sigma+it) = -\int_1^{\sigma} \zeta'(x+it)dx \ll |\sigma-1| \log(t)^2.$$

In particular this yields

$$|\zeta(1+it)| \geq C_1(\sigma-1)^{\frac{3}{4}} \log(t)^{-\frac{1}{4}} - C_2(\sigma-1)\log(t)^2.$$

By choosing σ appropriately we obtain the estimate

$$|\zeta(1+it)| \gg (\log(t))^{-7}.$$

Now taking $\sigma > 1 - \delta \log(t)^{-9}$ we again use the same trick to find

$$|\zeta(\sigma+it)| \geq (C - C'\delta)\log(t)^{-7}.$$

With this estimates at hand it is easy to complete the proof of the lemma. $\square$

We now turn to a slightly different argument and aim to investigate the slightly more difficult **Hurwitz-zeta function**

$$\zeta(s,\alpha) = \sum_{n=1}^{\infty} (n+\alpha)^{-s}.$$

Lemma 22. *For $0 < \alpha \leq 1$ the function $\zeta(s,\alpha)$ has meromorphic continuation to $\mathbb{C}$. The only pole appears at $s = 1$, is of order 1 and has residue 1.*

Proof. We start by the integral representation

$$\Gamma(z) = \frac{1}{e(z)-1} \int_H t^{z-1} e^{-t} dt,$$

where H is a path from $i+\infty$ to $-1+i$ to $-1-i$ to $-i+\infty$. We get

$$\begin{aligned}(e(s)-1)\Gamma(s)\zeta(s,\alpha) &= \sum_{n=1}^{\infty} \int_H \left(\frac{t}{n+\alpha}\right)^s e^{-t} \frac{dt}{t} \\ &= \int_H \frac{t^{s-1}}{e^{\alpha t}(1-e^{-t})} dt = \int_H \frac{t^{s-1} e^{\alpha t}}{e^t - 1} dt.\end{aligned}$$

The right hand side of this formula is holomorphic for all $s \in \mathbb{C}$. Looking at the left hand side reveals that $\frac{1}{e(s)-1}$ cancels exactly all the poles of $\Gamma(s)$ when $s \in -\mathbb{N}$. Thus the claim follows by investigating the behaviour at $s = 1$. But by computing residues the integral on the right is easily seen to equal $2\pi i$. Since $e(s) - 1$ has residue $\frac{1}{2\pi i}$ this concludes the proof. $\square$

Lemma 23. *For $s < 0$ and $0 < \alpha \leq 1$ we have*

$$\zeta(s,\alpha) = \frac{2\Gamma(1-s)}{(2\pi)^{1-s}} \left(\sin(\frac{\pi s}{2}) \sum_{n=1}^{\infty} \frac{\cos(2\pi n\alpha)}{n^{1-s}} + \cos(\frac{\pi s}{2}) \sum_{n=1}^{\infty} \frac{\sin(2\pi n\alpha)}{n^{1-s}} \right).$$

Proof. The statement follows by computing the residues, using

$$\Gamma(s)^{-1} = \frac{\sin(\pi s)}{\pi} \Gamma(1-s)$$

and some well known relations between sine and cosine. We leave the details to the reader. $\square$

We define the **conductor** of a Dirichlet character modulo q to be the smallest (positive integer) $q_1 = \mathrm{cond}(\chi)$ such that $\chi(n + q_1 m) = \chi(n)$ for all n, m. of course this implies $q_1 \mid q$. We call the character $\chi_0(n) = \delta_{(n,q)=1}$ the **principal character** modulo q. Further we call χ **primitive** if $\chi \neq \chi_0$ and $\mathrm{cond}(\chi) = q$. We define the Gauß sum

$$\tau(\chi) = \sum_{a=1}^{q} \chi(a) e(\frac{a}{q}).$$

Lemma 24. *Let χ be a primitive character modulo q. Then we have*

$$\chi(n)\tau(\overline{\chi}) = \sum_{p=1}^{q} \overline{\chi}(p) e(\frac{np}{q}).$$

Furthermore there is $\epsilon(\chi) \in S^1$ such that $\tau(\chi) = \epsilon(\chi)\sqrt{q}$.

Proof. For $(n,q) = 1$ this follows simply by writing

$$\chi(n)\tau(\overline{\chi}) = \sum_{p=1}^{q} \overline{\chi}(\overline{n}p) e(\frac{p}{q}) = \sum_{p=1}^{q} \overline{\chi}(p) e(\frac{np}{q}). \tag{25}$$

Suppose $\frac{n}{q} = \frac{n_1}{q_1}$ for $q_1 q_2 = q$ with $q_2 \neq 1$. Then we compute

$$\sum_{p=1}^{q} \overline{\chi}(p) e(\frac{np}{q}) = \sum_{p=1}^{q} \overline{\chi}(p) e(\frac{n_1 p}{q_1}) = \sum_{y=1}^{q_1} \left(\underbrace{\sum_{u=1}^{q_2} \chi(xq_1 + y)}_{=0} \right) e(\frac{yn_1}{q_1}) = 0.$$

Thus, we are done since $\chi(n) = 0$ for $(n,q) = q_2 \neq 1$.

To prove the second statement we observe that

$$\begin{aligned} \varphi(q) \left|\tau(\chi)\right|^2 &= \sum_{n=1}^{q} \left|\chi(n)\right|^2 \left|\tau(\chi)\right|^2 = \sum_{n=1}^{q} \sum_{p_1=1}^{q} \sum_{p_2=1}^{q} \chi(p_2)\overline{\chi}(p_1) e(\frac{n(p_1 - p_2)}{q}) \\ &= \sum_{p_1=1}^{q} \sum_{p_2=1}^{q} \chi(p_2)\overline{\chi}(p_1) \underbrace{\sum_{n=1}^{q} e(\frac{n(p_1 - p_2)}{q})}_{=q\delta_{p_1 \equiv p_2 \bmod q}} = q \sum_{p=1}^{q} \left|\chi(p)\right|^2 = \varphi(q) q. \end{aligned}$$

This implies $|\tau(\chi)| = \sqrt{q}$ and we are done. ☐

We call χ **even** (resp. **odd**) if $\chi(-1) = 1$ (resp. $\chi(-1) = -1$). Note that the lemma above obviously implies

$$\begin{aligned} &\sum_{p=1}^{q} \chi(p) \sin(2\pi \frac{np}{q}) = \begin{cases} -i\overline{\chi}(n)\tau(\chi) & \text{if } \chi \text{ is odd,} \\ 0 & \text{else} \end{cases} \\ &\qquad\qquad \text{and } \sum_{p=1}^{q} \chi(p) \cos(2\pi \frac{np}{q}) = \begin{cases} \overline{\chi}(n)\tau(\chi) & \text{if } \chi \text{ is even,} \\ 0 & \text{else.} \end{cases} \end{aligned} \tag{26}$$

Theorem 5. *Let χ be a primitive Dirichlet character modulo q, then $L(s,\chi)$ has analytic continuation to $\mathbb{C}$ and satisfies the functional equation*

$$\Lambda(s,\chi) = (-i)^{\rho} \epsilon(\chi) \Lambda(1-s, \overline{\chi}), \text{ for } \Lambda(s,\chi) = \pi^{-\frac{1+\rho}{2}} q^{\frac{1+\rho}{2}} \Gamma(\frac{1+\rho}{2}) L(s,\chi).$$

Here $\rho \in \{0,1\}$ such that $\chi(-1) = (-1)^{\rho}$.

Proof. We observe that

$$\begin{aligned} L(s,\chi) = \sum_{\substack{p=1,\\(p,q)=1}}^{q-1} \chi(p) \sum_{n\in\mathbb{N}, n\equiv p \bmod q} n^{-s} \\ = \sum_{\substack{p=1,\\(p,q)=1}}^{q-1} \chi(p) \sum_{n\in\mathbb{N}} (p+qn)^{-s} = q^{-s} \sum_{\substack{p=1,\\(p,q)=1}}^{q-1} \chi(p)\zeta(s,\frac{p}{q}). \end{aligned}$$

Thus Lemma 22 implies an analytic continuation of $L(s,\chi)$ to $\mathbb{C}\setminus\{1\}$. However, applying character orthogonality and using the fact that χ is non principal we find that the singularity at $s=1$ is liftable.

It remains to prove the functional equation. The principle of analytic continuation allows us to prove the functional equation for s with $\Re(s) < 0$ and then extend it to the full complex plane. We show the argument for χ even as the odd case is analogous. Note that combining (26) with Lemma 23 yields

$$\Lambda(s,\chi) = \underbrace{2^s \sin(\frac{\pi s}{2}) \frac{\Gamma(1-s)\Gamma(\frac{s}{2})}{\sqrt{\pi}\Gamma(\frac{1-s}{2})}}_{=1} \epsilon(\chi)\Lambda(s,\overline{\chi}).$$

The Γ-factors are removed by the well known doubling formulae. □

In analogy to the Theorem 4 we have the following product expansion for completed Dirichlet L-functions.

Theorem 6. *The completed Dirichlet L-function $\Lambda(s,\chi)$ has no zeros outside the strip $0 \leq \Re(s) \leq 1$ and satisfies*

$$\Lambda(s,\chi) = e^{A(\chi)+B(\chi)s} \prod_{\rho\in\mathcal{N}(\chi)} \left(1-\frac{s}{\rho}\right) e^{\frac{s}{\rho}}.$$

The following preliminary bound will be important later on.

Lemma 25. *For $\Re(s) > 1 - \frac{1}{2\log(q)}$ we have*

$$L(s,\chi) \ll |s|\log(q) \text{ and } L'(s,\chi) \ll |s|\log(q)^2.$$

Proof. Partial summation shows that

$$L(s,\chi) = s\int_1^\infty x^{-s-1} S_\chi(x)dx,$$

with $S_\chi(x) = \sum_{n\leq x} \chi(n)$. Estimating trivially yields

$$\int_1^q x^{-s-1} S_\chi(x)dx \ll \int_1^q x^{-\sigma}dx \ll \log(q),$$

for $\sigma \geq 1-\log(q)^{-1}$. To estimate the other part we use $S_\chi(x) \leq q$, which follows from character orthogonality. Indeed we get

$$\int_q^\infty x^{-s-1} S_\chi(x)dx \ll q\int_q^\infty x^{-\sigma-1}dx \ll 1.$$

This shows the first claim. The second follows by estimating the Cauchy integral formula.

$\square$

Exercise 5. *Use the tools of this section to proof the* ***prime number theorem*** *in the form*

$$\pi(x) = \sharp\{p \leq x\colon \; \textit{prime} \; \} = \int_2^x \frac{du}{\log u} + O(E(x))$$

and specify the error $E(x)$. A complete solution to this exercise will be given in Part 6.

3 Part 3: Zeros of $\zeta(s)$ and $L(s,\chi)$

Recall that $\mathcal{N}$ is the set of all zeros $s(s-1)\Lambda(s)$. This agrees with the (multi)-set of zeros of $\zeta(s)$ in the critical strip $0 < \Re(s) < 1$.[14] Similarly the (multi)-set $\mathcal{N}(\chi)$ is defined.

We now define

$$N(T) = \sharp\{\rho \in \mathcal{N}\colon\ |\Im(\rho)| \leq T\} \text{ and } N_\chi(T) = \sharp\{\rho \in \mathcal{N}(\chi)\colon\ |\Im(\rho)| \leq T\}.$$

Lemma 26. *We have*

$$N(T+1) - N(T-1) \ll \log(T) \text{ and } N_\chi(T+1) - N_\chi(T-1) \ll \log(qT),$$

for a primitive Dirichlet character modulo q.

Proof. We apply Jensen's formula to get

$$\int_0^1 \log(|\zeta(2+iT+re(\theta))|)d\theta = \log(|\zeta(2+iT)|) + \sum_{\substack{\rho\in\mathcal{N},\\ \rho\in B_r(2+iT)}} \log(\frac{r}{|\rho|}).$$

Here we choose $r \in [3,4]$ such that $\mathcal{N} \cap \partial B_r(2+iT) = \emptyset$. Using a simple bounds of the form $\zeta(s) \ll |\Im(s)|^k$ for some $k > 0$ allows us to estimate

$$\sum_{\substack{\rho\in\mathcal{N},\\ \rho\in B_r(2+iT)}} \log(\frac{r}{|\rho|}) \ll \log(T).$$

The proof is complete in combination with the estimate

$$N(T+1)-N(T-1) \leq \sharp\{\rho \in \mathcal{N}\colon\ |\Im(\rho) - T| \leq 1\} \leq \log(\frac{r}{\sqrt{5}}) \sum_{\substack{\rho\in\mathcal{N},\\ \rho\in B_r(2+iT)}} \log(\frac{r}{|\rho|}).$$

The proof for Dirichlet L-functions is similar, the only difference being that the estimate $L(s,\chi) \ll |q\Im(s)|^k$ is used for some $k > 0$. $\square$

Corollary 2. *We have*

$$N(T) \ll T\log(T) \text{ and } N_\chi(T) \ll T\log(qT).$$

Lemma 27. *We have*

$$\frac{\zeta'}{\zeta}(s) - \delta_{\substack{\Re(s)\in[-1,2],\\ |\Im(s)|\geq 1}} \sum_{\substack{\rho\in\mathcal{N},\\ |\Im(s-\rho)|<1}} \frac{1}{s-\rho} \ll \begin{cases} 1 & \text{if } \Re(s) \geq 2, \\ \log(|s|) & \text{if } \Re(s) \leq 1, |s+2m| > \frac{1}{4} \text{ for all } m \in \mathbb{N}, \\ 1+\log(|s|) & \text{if } -1 \leq \Re(s) \leq 2, |\Im(s)| \geq 1. \end{cases}$$

Proof. We do the different cases separately. First, for $\Re(s) \geq 2$ we have

$$\frac{\zeta'}{\zeta}(s) \leq \sum_{n\geq 1} \log(n)n^{-2} \ll 1.$$

[14] If ρ is a zero with multiplicity m, then it will appear m times in $\mathcal{N}$.

If $\Re(s) \leq -1$, we put $s = 1 - s'$ and apply functional equation to get

$$\frac{\zeta'}{\zeta}(s) = -\log(2\pi) - \frac{\pi}{2}\tan(\frac{\pi s'}{2}) + \frac{\Gamma'}{\Gamma}(s') + \frac{\zeta'}{\zeta}(s').$$

The claimed bound follows by using Stirling's formula.

We turn towards the final range. Using the product formula and Stitling's formula reveals

$$\frac{\zeta'}{\zeta}(s) = B + \sum_{\rho\in\mathcal{N}}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right) + O(\log(|\Im(s)|)).$$

From this we derive

$$\frac{\zeta'}{\zeta}(s) = \sum_{\rho\in\mathcal{N}}\left(\frac{1}{s-\rho} - \frac{1}{2+it-\rho}\right) + \frac{\zeta'}{\zeta}(2+it) + O(\log(|\Im(s)|)).$$

We write $\rho' = 2 + it - \rho$. First we trivially estimate $\frac{\zeta'}{\zeta}(2+it) \ll 1$ and

$$\sum_{|\Im(\rho'-s)|\leq 1} |\rho'|^{-1} \ll \log(|\Im(s)|).$$

Finally we assume $|\Im(\rho - s)| > 1$. Then $|s-\rho| \asymp |\Im(s-\rho)|$ and

$$\frac{1}{s-\rho} - \frac{1}{\rho'} = \frac{2-\sigma}{(s-\rho)(2+it-\rho)} \ll |\Im(s-\rho)|^{-2}.$$

We now finally estimate

$$\begin{aligned}\sum_{|\Im(\rho-s)|>1}\left(\frac{1}{s-\rho} - \frac{1}{\rho'}\right) &= \sum_{n\in\mathbb{N}}\sum_{|\Im(s)|+n\leq|\Im(\rho)|<|\Im(s)|+n+1}\left(\frac{1}{s-\rho} - \frac{1}{\rho'}\right) \\ &\ll \sum_{n\in\mathbb{N}}[N(|\Im(s)|+n+1) - N(|\Im(s)|+n)]n^{-2} \ll \sum_{n\in\mathbb{N}} n^{-2}\log(|\Im(s)|+n) \\ &\ll \log(|\Im(s)|).\end{aligned}$$

This completes the proof. □

Remark 2. *Similarly one can prove the following estimate for the logarithmic derivative of Dirichlet L-functions.*

$$\begin{aligned}&\frac{L'}{L}(s,\chi) - \delta_{\substack{|s|\leq\frac{1}{2},\\ \chi(-1)=1}}\frac{1}{s} - \delta_{\substack{\Re(s)\in[-1,2],\\ |\Im(s)|\geq 1}}\sum_{\substack{\rho\in\mathcal{N}(\chi),\\ |\Im(s-\rho)|<1}}\frac{1}{s-\rho} \\ &\qquad \ll \begin{cases}1 & \text{if } \Re(s)\geq 2, \\ \log(q\,|s|) & \text{if } \Re(s)\leq 1, |s+2m| > \frac{1}{4} \text{ for all } m\in\mathbb{N}, \\ \log(q(2+|\Im(s)|)) & \text{if } -1\leq\Re(s)\leq 2, |\Im(s)|\geq 1.\end{cases} \qquad (27)\end{aligned}$$

Here χ is a primitive Dirichlet character modulo q.

Theorem 7. *Let $x > 2$ and define*

$$\langle x \rangle = \min\{\left|x - p^k\right| : p \text{ prim }, k \in \mathbb{N}, x \neq p^k\}.$$

We have

$$\psi(x) - \delta_{x\in\mathbb{N}}\Lambda(n) = x - \sum_{\substack{\rho\in\mathcal{N}, \\ |\Im(\rho)|<T}} \frac{x^\rho}{\rho} - \frac{\zeta'}{\zeta}(0) - \frac{1}{2}\log(1 - x^{-2})$$
$$+ O\left(\frac{x}{T}\log(xT)^2 + \log(x)\min(1, \frac{x}{T\langle x\rangle})\right).$$

Proof. Put $c = 1 + \log(x)^{-1}$ and apply Lemma 18 to get

$$\left|\psi(x) - \delta_{x\in\mathbb{N}}\Lambda(n) + \frac{1}{2\pi i}\int_{c-iT}^{c+it} \frac{\zeta'}{\zeta}(s)\frac{x^s}{s}ds\right| < \sum_{n\neq x} \Lambda(n)\left(\frac{x}{n}\right)^c \min(1, (T\left|\log(\frac{x}{n})\right|)^{-1}) + \delta_{x\in\mathbb{N}} c\frac{\Lambda(x)}{T}.$$

We have to estimate this error carefully. We start by considering $x \notin [\frac{1}{2}x, 2x]$, such that $\left|\log(\frac{x}{n})\right| \gg 1$. This terms can at most contribute

$$\sum_{n\notin[\frac{1}{2}x,2x]} \Lambda(n)\left(\frac{x}{n}\right)^c \ll \frac{x}{T}\left|\frac{\zeta'}{\zeta}(1 + \log(x)^{-1})\right| \ll \frac{x}{T}\log(x).$$

This absorbs of course the term $\delta_{x\in\mathbb{N}} c\frac{\Lambda(x)}{T}$. To estimate $x \in (\frac{1}{2}x, x)$ we set $x_1 = \max\{p^k \colon p^k < x\}$. Without loss of generality we can assume $x_1 > \frac{1}{2}x$. We compute

$$\log(\frac{x}{x_1}) = -\log(1 - \frac{x - x_1}{x}) \geq \frac{x - x_1}{x} \geq \frac{\langle x\rangle}{x}.$$

With this at hand we get

$$\Lambda(x_1)\left(\frac{x}{x_1}\right)^c \min(1, (T\left|\log(\frac{x}{x_1})\right|)^{-1}) \ll \log(x)\min(1, \frac{x}{T\langle x\rangle}).$$

For other $n \in (\frac{1}{2}x, x_1)$ we write $x = x_1 - v$. One checks that $\log(\frac{x}{n}) \geq \frac{v}{x_1}$. This yields

$$\sum_{\frac{1}{2}x<n<x_1} \Lambda(n)\left(\frac{x}{n}\right)^c \min(1, (T\left|\log(\frac{x}{n})\right|)^{-1}) \ll \sum_{v\leq\frac{1}{2}x} \Lambda(x_1 - v)\frac{x_1}{Tv} \ll \frac{x}{T}(\log(x)).$$

Since $\Lambda(n) = 0$ for all $x_1 < n < x$ by construction of x_1 this completes the estimation to the contribution from $(\frac{1}{2}x, x)$. The part $(x, 2x)$ is done similarly. Altogether we have seen that

$$\psi(x) - \delta_{x\in\mathbb{N}}\Lambda(x) = -\frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{\zeta'}{\zeta}(s)\frac{x^s}{s}ds + O\left(\frac{x}{T}\log(xT)^2 + \log(x)\min(1, \frac{x}{T\langle x\rangle})\right).$$

It remains to appropriately manipulate the contour integral. We start by computing residues. If ρ is a zero of $\zeta(s)$, then

$$\operatorname*{res}_{s=\rho} \frac{\zeta'}{\zeta}(s)\frac{x^s}{s} = \frac{x^\rho}{\rho}.$$

Note that ρ is in $\mathcal{N}$ or an even negative integer.

$$\operatorname*{res}_{s=0} \frac{\zeta'}{\zeta}(s)\frac{x^s}{s} = \frac{\zeta'}{\zeta}(0) \text{ and } \operatorname*{res}_{s=0} \frac{\zeta'}{\zeta}(s)\frac{x^s}{s} = -x.$$

Without loss of generality we assume that there is no $\rho \in \mathcal{N}$ with $\Im(\rho) = T$. Applying the residual theorem we find

$$-\frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{\zeta'}{\zeta}(s)\frac{x^s}{s}ds = x - \frac{\zeta'}{\zeta}(0) - \sum_{\substack{\rho\in\mathcal{N},\\ |\Im(\rho)|<T}} \frac{x^\rho}{\rho} + \sum_{m\leq\frac{1}{2}U} \frac{x^{-2m}}{2m}$$

$$-\frac{1}{2\pi i}\int_{-U-iT}^{-U+iT} \frac{\zeta'}{\zeta}(s)\frac{x^s}{s}ds + \frac{1}{2\pi i}\int_{-U-iT}^{c-iT} \frac{\zeta'}{\zeta}(s)\frac{x^s}{s}ds - \frac{1}{2\pi i}\int_{-U+iT}^{c+iT} \frac{\zeta'}{\zeta}(s)\frac{x^s}{s}ds.$$

It remains to estimate the contribution of the remaining integrals. We will then take $U \to \infty$ through odd integers. Trivial bounds suffice to get

$$\frac{1}{2\pi i}\int_{-U-iT}^{-U+iT} \frac{\zeta'}{\zeta}(s)\frac{x^s}{s}ds \ll \frac{\log(U)}{U}x^{-U}T.$$

Let us consider the remaining two integrals for $T' \in [T-1, T+1]$ with the property that $|\Im(\rho) - T| \gg \log(T)^{-1}$. Such a T' always exists due to Lemma 26. Of course we have the estimate

$$\frac{\zeta'}{\zeta}(\sigma + iT') \ll \sum_{|\Im(\rho)-T'|<1} \frac{1}{|s-\rho|} + \log(T') \ll \log(T')^2.$$

Thus we find that

$$\frac{1}{2\pi i}\int_{-U\pm iT}^{c\pm iT} \frac{\zeta'}{\zeta}(s)\frac{x^s}{s}ds \ll \int_{-1}^{c} \frac{x^\sigma}{|\sigma \pm iT'|}\log(T')^2 d\sigma + \int_{-U}^{-1} \log(|\sigma \pm iT'|)\frac{x^\sigma}{|\sigma \pm iT'|}d\sigma$$
$$\ll \frac{\log(T')^2}{T'}\int_{-\infty}^{c} x^\sigma d\sigma \ll \frac{\log(T)^2}{T}\frac{x}{\log(x)}.$$

By shifting the lines of integration from T to T' one picks up other residues which contribute

$$\sum_{|\Im(\rho)|\in[T,T']} \frac{x^\rho}{\rho} \ll \frac{x}{T}\log(T).$$

Collecting all the errors together the theorem follows. □

The same argument works for Dirichlet L-functions. However, the residue computations are slightly different. The result is the following.

Theorem 8. *Let $\psi_\chi(x) = \sum_{n\leq x}\chi(n)\Lambda(n)$. Then, for $x > 2$ and a primitive Dirichlet character χ modulo q, one has*

$$\psi_\chi(x) - \delta_{x\in\mathbb{N}}\chi(n)\Lambda(n) = -\sum_{\substack{\rho\in\mathcal{N}(\chi),\\ |\Im(\rho)|<T}} \frac{x^\rho}{\rho} - \Delta(x) + O\left(\frac{x}{T}\log(xqT)^2 + \log(x)\min(1, \frac{x}{T\langle x\rangle})\right).$$

Where

$$\Delta(x) = \begin{cases} -\log(x) - B(\chi) + \sum_{m=1}^{\infty}\frac{x^{-2m}}{2m} & \text{if } \chi \text{ is even,}\\ -\frac{L'}{L}(0,\chi) + \sum_{m=1}^{\infty}\frac{x^{1-2m}}{2m-1} & \text{if } \chi \text{ is odd.}\end{cases}$$

After having seen some preliminary results on zeros in boxes we now turn to the distribution of zeros close to the line $\Re(s) = 1$.

Theorem 9 (Standard zero-free region)**.** *There is a constant $C > 0$ such that for all $\rho \in \mathcal{N}$ we have*

$$\Re(\rho) < 1 - C \min(1, \log(\Im(\rho))^{-1}).$$

Proof. We start by the same inequality used in the proof of lemma 21 to find

$$\begin{aligned}\Re\left(-3\frac{\zeta'}{\zeta}(\sigma) - 4\frac{\zeta'}{\zeta}(\sigma+it) - \frac{\zeta'}{\zeta}(\sigma+2it)\right)&\\ = \sum_{n=1}^{\infty} \Lambda(n) n^{-\sigma}(3 + 4\cos(t\log(n)) + \cos(2t\log(n))) \geq 0.& \quad (28)\end{aligned}$$

First we observe that the existence of the pole ensures that

$$-\frac{\zeta'}{\zeta}(\sigma) < \frac{1}{\sigma - 1} + c,$$

for some large c. By making c larger if necessary we find

$$\Re(-\frac{\zeta'}{\zeta}(s)) = -\sum \frac{1}{\Re(s-\rho)} + O(\log(|s|)) < -\Re\left(\frac{1}{s-\rho_0}\right) + c\log(|s|) < c\log(|s|).$$

in the region $1 < \Re(s) < 2$ and $t \geq 1$. Here we used that since $\Re(s) > 1$ we must have $\Re(s-\rho) > 0$ for all zeros ρ.

Zeros with $0 \leq \Im(\rho) \leq 1$ are easily excluded as in the proof of Lemma 21. Exploiting symmetry we can assume that $\Im(\rho) > 1$. From the considerations above we get

$$0 \leq \frac{3}{\sigma - 1} - \frac{4}{\sigma - \Re(\rho)} + C\log(t+2).$$

This is valid for some suitable constant C and all zeros $\rho = \beta + it$. Choosing $\sigma = 1 + \frac{\delta}{\log(t+2)}$ we get

$$\Re(\rho) < 1 + \frac{\delta}{\log(t+2)} - \frac{4\delta}{(3 + C\delta)\log(t+2)}.$$

Choosing $\delta = \frac{1}{3C}$ completes the proof. □

We now turn towards the slightly more difficult situation of Dirichlet L-functions. Here the result is as follows.

Theorem 10. *Let $c > 0$ be sufficiently small and $\chi \neq \chi_0$ be a character modulo q. Suppose $L(s,\chi)$ has a zero s_0 with*

$$\Re(s_0) \geq 1 - \frac{c}{\log(q(|\Im(s_0)| + 2))},$$

then χ is a real character, the zero is simple, real and unique.

Proof. We start by showing the normal zero-free region for complex characters $\chi^2 \neq \chi_0$. The argument starts by writing

$$\Re(\chi(n)e^{-it\log(n)}) = \cos(\alpha_n) \text{ and } \Re(\chi^2(n)e^{-2it\log(n)}) = \cos(2\alpha_n).$$

The by now familiar argument shows

$$-3\frac{L'}{L}(\sigma,\chi_0) - 4\Re(\frac{L'}{L}(\sigma+it,\chi)) - \Re(\frac{L'}{L}(\sigma+2it,\chi^2)) \geq 0. \tag{29}$$

Of course we have the estimate

$$-\frac{L'}{L}(\sigma,\chi_0) \leq -\frac{\zeta'}{\zeta}(\sigma) < \frac{1}{\sigma-1} + A.$$

Now let $\rho \in \mathcal{N}(\chi)$ and set $t = \Im(\rho)$. For $1 < \sigma < 2$ we use (27) and drop all the irrelevant terms to get

$$-\Re(\frac{L'}{L}(\sigma+it,\chi)) < A\log(q(|t|+2)) - \frac{1}{\sigma - \Re(\rho)}.$$

For the final case we use the bound

$$-\Re(\frac{L'}{L}(\sigma+2it,\chi^2)) < A\log(q(|t|+2)).$$

Note that if χ^2 is not primitive one needs to artificially insert the missing Euler-factors. However, their contribution is easily handled. Combining our 3 bounds we get

$$\frac{4}{\sigma-\Re(\rho)} < \frac{3}{\sigma-1} + 8A\log(q(|t|+2)).$$

We choose $\sigma = 1 + \delta\log(q(|t|+2))^{-1}$ to find

$$1 - \Re(\rho) > \frac{\delta}{\log(q(|t|+2))}\left(\frac{4}{8A\delta+3} - 1\right).$$

By choosing δ accordingly we conclude this case.

We turn towards $\chi^2 = \chi_0$. If we want to use the same argument as above we need to replace the bound for $-\frac{L'}{L}(\sigma+2it,\chi^2)$ in (29). After handling the missing Euler factors we find

$$\begin{aligned} -\Re(\frac{L'}{L}(\sigma+2it,\chi_0)) &= -\Re(\frac{\zeta'}{\zeta}(\sigma+2it)) + O(\log(q)) \\ &< \Re\left(\frac{1}{\sigma-1+sit}\right) + A\log(q(|t|+2)). \end{aligned}$$

Thus, the usual argument provides the bound

$$\frac{4}{\sigma-\Re(\rho)} < \frac{3}{\sigma-1} + \Re\left(\frac{1}{\sigma-1-2it}\right) + 8A\log(q(|t|+2)),$$

where $\sigma > 1$ and $t = \Im(\rho)$. Suppose that $|t| > \frac{\delta}{\log(q(|t|+2))}$, then

$$\Re\left(\frac{1}{\sigma-1-2it}\right) = \frac{\sigma-1}{(\sigma-1)^2+4t^2} \leq \frac{1}{5\delta}\log(q(|t|+2)).$$

With this at hand we find

$$1 - \Re(\rho) > \frac{4 - 40A\delta}{16 + 40A\delta} \cdot \frac{\delta}{\log(q(|t| + 2))}.$$

Therefore, up to choosing suitable δ, we established the zero-free region for real characters as long as $\Im(\rho) > \delta \log(q)^{-1}$.

It remains to show that $L(s, \chi)$ has at most a unique, real zero in the region $\Im(s) \leq \delta \log(q)^{-1}$, $\Re(s) > 1 - \frac{\delta}{\log(q)}$. To do so we consider

$$-\frac{L'}{L}(\sigma, \chi) < A \log(q) - \sum_{\Im(\rho)<1} \frac{1}{\sigma - \rho}.$$

Note that this inequality only makes sense because both sides are real. Indeed this is obvious for $-\frac{L'}{L}(\sigma, \chi)$ and for the sum over zeros this follows since $\rho \in \mathcal{N}(\chi)$ implies $\overline{\rho} \in \mathcal{N}(\chi)$. If we write $\rho = \beta + i\gamma$. If $\gamma > 0$, then

$$\frac{1}{\sigma - \rho} + \frac{1}{\sigma - \overline{\rho}} = \frac{2(\sigma - \beta)}{(\sigma - \beta)^2 + \gamma^2} > 0.$$

We get the upper bound

$$-\frac{L'}{L}(\sigma, \chi) < A \log(q) - \frac{2(\sigma - \beta)}{(\sigma - \beta)^2 + \gamma^2}.$$

To complement this we compute

$$-\frac{L'}{L}(\sigma, \chi) = \sum_n \chi(n)\Lambda(n)n^{-\sigma} \geq -\sum_n \Lambda(n)n^{-\sigma} = \frac{\zeta'}{\zeta}(\sigma) > -\frac{1}{\sigma - 1} - A.$$

We conclude that

$$-\frac{1}{\sigma - 1} < 2A \log(q) - \frac{2(\sigma - \beta)}{(\sigma - \beta)^2 + \gamma^2} < 2A \log(q) - \frac{8}{5(\sigma - \beta)}.$$

Here the last inequality holds for $\delta < 1$, $\gamma \leq \delta \log(q)^{-1}$ and $\sigma = 1 + 2\delta \log(q)^{-1}$. One easily deduces $\beta < 1 - \delta \log(q)^{-1}$. It remains the possibility that there are real zeros in the region. Suppose $\beta_1 \leq \beta_2$ are two such zeros. The same argument as above now shows the inequality

$$-\frac{1}{\sigma - 1} < 2A \log(q) - \frac{2\sigma - \beta_1 - \beta_2}{(\sigma - \beta_1)(\sigma - \beta_2)}.$$

From this it follows that at most one of the zeros β_1 and β_2 lies in the desired region. $\square$

If a real character χ has such zero in the region described in the theorem, we call it **exceptional character** and refer to the (unique) zero as **exceptional zero**.

Theorem 11 (Landau). *Suppose χ_1 and χ_2 are two exceptional characters and with corresponding exceptional zeros β_1 and β_2, then*

$$\min(\beta_1, \beta_2) < 1 - \frac{C}{\log(q_1 q_2)},$$

for sufficiently small $C > 0$.

Proof. By looking at the corresponding Dirichlet series we get

$$-\frac{\zeta'}{\zeta}(\sigma)-\frac{L'}{L}(\sigma,\chi_1)-\frac{L'}{L}(\sigma,\chi_2)-\frac{L'}{L}(\sigma,\chi_1\chi_2)=\sum_{n=1}^{\infty}\Lambda(n)n^{-\sigma}(1+\chi_1(n))(1+\chi_2(n))\geq 0.$$

Since $\chi_1\chi_2$ is not the principal character we can estimate

$$-\frac{L'}{L}(\sigma,\chi_1\chi_2)<A\log(q_1q_2).$$

For the other logarithmic derivatives we have the usual estimates and derive the inequality

$$\frac{1}{\sigma-\beta_1}+\frac{1}{\sigma-\beta_2}<\frac{1}{\sigma-1}+3A\log(q_1q_2)$$

in the usual way. Put $\sigma=1+\delta\log(q_1q_2)^{-1}$ for sufficiently small δ and assume $\beta_1\leq\beta_2$. Then it follows that $\beta_1\leq 1-\frac{C}{\log(q_1q_2)}$. □

Corollary 3 (Page)**.** *Let C be as in Theorem 11 and $Q\geq 3$. Then there is at most one real primitive Dirichlet character χ modulo q with $q\leq Q$ such that $L(s,\chi)$ has a real zero $\beta\in(1-\frac{C}{2\log(Q)},1]$.*

Proof. Suppose χ_1,χ_2 are two such characters. Then we have

$$\beta_i>1-\frac{C}{2\log(Q)}\geq 1-\frac{C}{\log(q_1q_2)}.$$

However, this is a contradiction to Theorem 11. □

We say $q>1$ is a **good modulus** if $L(s,\chi)\neq 0$ for all characters χ mod q and all $s=\sigma+it$ such that

$$\sigma>1-\frac{C}{\log(q(|t|+1))}.$$

Lemma 28. *There is a constant $C>0$ such that there are arbitrarily large $z\in\mathbb{R}$ such that $P(z)=\prod_{p<z}p$ is a good modulus.*

Proof. Let C_1 be the constant given by Corollary 3. Given x_1 w will construct $x\geq x_1$ with $P(x)$ good. This goes as follows. First, if $P(x_1)$ is good, then we are done. Otherwise there is exactly one exceptional character χ_0 with exceptional zero $\beta>1-\frac{C_1}{\log(P(x_1))}$. Note that by the prime number theorem we have $\log(P(x))\sim x$, so that we can find x satisfying

$$\frac{C_1/2}{\log(P(x))}<1-\beta<\frac{C_1}{\log(P(x))}.$$

Since $P(x_1)\mid P(x)$ we can view χ_0 as character modulo $P(x)$ and it is still the only exceptional character (with respect to C_1). Thus, $P(x)$ must be good for $C=\frac{C_1}{2}$. □

Theorem 12 (Siegel)**.** *For every $\epsilon>0$ there is a constant $C_1=C_1(\epsilon)$ such that the following holds. Suppose β is a real zero of $L(s,\chi)$ for a real character χ modulo q, then $\beta<1-C_1q^{-\epsilon}$.*

Proof. By Lemma 25 it suffices to show that $L(1,\chi) > C(\epsilon)q^{-\epsilon}$ for primitive real characters χ modulo q.

Take two distinct real characters χ_1, χ_2 and define

$$F(s) = \zeta(s)L(s,\chi_1)L(s,\chi_2)L(s,\chi_1\chi_2).$$

Note that $F(s) = D_f(s)$ for $f = \epsilon\star\chi_1\star\chi_2\star\chi_1\chi_2$. Multiplying the Euler-products we find that

$$\log(F(s)) = \sum_p \sum_{k=1}^{\infty} \frac{p^{-ks}}{k}(1+\chi_1(p^k))(1+\chi_2(p^k)).$$

In particular we observe that $f(1) = 1$ and $f(n) \geq 0$ for all $n \in \mathbb{N}$.

From the analytic properties of $\zeta(s)$ and Dirichlet L-functions we conclude that $F(s)$ is holomorph in $\mathbb{C}\setminus\{1\}$ with

$$\lambda = \operatorname*{res}_{s=1} F(s) = L(1,\chi_1)L(1,\chi_2)L(1,\chi_1\chi_2).$$

Let us expand

$$F(s) = \sum_{k=0}^{\infty} b_k(2-s)^k.$$

in a convergent Taylor expansion at $s = 2$. Due to the pole at $s = 1$ the radius of convergence is 1. We compute the coefficients

$$b_k = \frac{(-1)^k}{k!}\sum_{n=1}^{\infty}(-\log(n))^k\frac{f(n)}{n^2} \geq 0.$$

We can write $(s-1)^{-1} = (1-(2-s))^{-1}$ and use the geometric series to find

$$F(s) - \frac{\lambda}{s-1} = \sum_{k=0}^{\infty}(b_k-\lambda)(2-s)^k.$$

The left hand side is holomorphic in $\mathbb{C}$. This is implies that the right hand side must converge everywhere. We obtain the useful estimate

$$(-1)^k(b_k-\lambda) = \frac{1}{2\pi i}\int_{\partial B_{\frac{3}{2}}(2)} \frac{F(s)-\lambda(s-1)^{-1}}{(s-2)^{k+1}}ds \ll \left(\frac{2}{3}\right)^k (q_1q_2)^2.$$

With this at hand we can truncate the Taylor expansion at some large constant K. For real $\frac{7}{8} < s < 1$ this yields

$$F(s) - \frac{\lambda}{s-1} \geq 1 - \lambda\sum_{k=0}^{K-1}(2-s)^k - 4A(q_1q_2)^2\left(\frac{3}{4}\right)^K.$$

We choose K such that $\frac{3}{8} \leq 4A(q_1q_2)^2(\frac{3}{4})^K < \frac{1}{2}$. In particular this implies $K \leq 8\log(q_1q_2) + c_1$. Computing the geometric series shows

$$F(s) \geq \frac{1}{2} - \lambda\frac{(2-s)^K}{1-s} \geq \frac{1}{2} - \frac{e^{c_1}\lambda}{1-s}(q_1q_2)^{8(1-s)}. \tag{30}$$

This is the key inequality that drives the proof.

If there is a real character χ with a real zero $\beta \in [1-\frac{\epsilon}{16}, 1)$ then we put $\chi_1 = \chi$ and $\beta_1 = \beta$. In this case we have $F(\beta_1) = 0$ for all χ_2. Otherwise we choose $\beta_1 \in [1 - \frac{\epsilon}{16}, 1)$ and χ_1 real arbitrary. Observe that $\zeta(\beta) < 0$ but $L(\beta_1, \chi_1)$, $L(\beta_1, \chi_2)$ and $L(\beta_1, \chi_1\chi_2)$ are positive by construction (they are positive at 1 and there is no zero between β_1 and 1). Thus in this case $F(\beta_1) < 0$. In any case we conclude that

$$F(\beta_1) \leq 0.$$

Inserting this in (30) we get

$$c\lambda > \frac{1}{2}(1 - \beta_1)(q_1 q_2)^{-8(1-\beta_1)}.$$

For fixed χ_1 and β_1 we have

$$\lambda \leq A \log(q_1) \log(q_1 q_2) L(1, \chi_2).$$

We deduce that

$$L(1, \chi) > C q_2^{-8(1-\beta_1)} \log(q_2)^{-1}.$$

The claimed inequality follows from $8(1 - \beta_1) < \frac{1}{2}\epsilon$. □

Remark 3. *How strange these possible exceptional zeros are can be seen from the so called exceptional-zero-repulsion, which states the following. Suppose* $\beta = 1 - \delta$ *is the unique exceptional zero in* $\sigma > 1 - \frac{c_1}{\log(T)}$ *belonging to a Dirichlet* L*-function of conductor* $q \leq T$*, then the zero free region can be widened to*

$$\sigma > 1 - \frac{c_2}{\log(T)} \log(\frac{ec_1}{\delta \log(T)}),\ |t| \leq T, \tag{31}$$

for $\delta \log(T) \to 0$ *and* β *is still the only exception.*

Exercise 6. *Use the tools developed in this section to prove the Siegel theorem about primes in arithmetic progressions, which states*

$$\pi(x; a, q) = \sharp\{p \leq x \colon p \equiv a \ mod\ q,\ p\ prime\ \} \sim \frac{1}{\varphi(q)} \int_2^x \frac{du}{\log(u)}.$$

What can you say about the error term? This exercise will be resolved in Part 6.

4 Part 4: Exponential sums and refined properties of $\zeta(s)$

In this section we will get to know two methods to treat exponential sums. We will apply the first one to derive an approximation to the Riemann zeta function. The second one will be used to derive an extended zero free region for $\zeta(s)$.

Lemma 29. *Let $F \in \mathcal{C}^1([a,b])$ and $G \in \mathcal{C}([a,b])$. Suppose that $F'(x) \neq 0$ for $x \in [a,b]$. If $\frac{G}{F'}$ is monotone, then we have*

$$\left| \int_a^b e^{iF(x)} G(x) dx \right| \leq 4 \left| \frac{G(a)}{F'(a)} \right| + 4 \left| \frac{G(b)}{F'(b)} \right|.$$

Proof. Without loss of generality we can assume $F'(x) > 0$. If Φ is the inverse of F, then we have $\Phi'(y) = F'(\Phi(y))^{-1}$. Substituting $x = \Phi(y)$ we get

$$\int_a^b e^{iF(x)} G(x) dx = \int_{F(a)}^{F(b)} e^{iy} \underbrace{\frac{G(\Phi(y))}{F'(\Phi(y))}}_{H(y)} dy.$$

Note that H is monotone, say increasing. Put $c = F(a)$ and $d = F(b)$. Using the mean value theorem we get

$$\Re \left(\int_c^d e^{iy} H(y) dy \right) = H(c) \int_c^x \cos(y) dy + H(d) \int_x^d \cos(y) dy.$$

Using the analogous equality for the imaginary part and estimating trivially gives the result. $\square$

Another useful integral estimate goes as follows.

Lemma 30. *Let $F \in \mathcal{C}^2([a,b])$ and suppose $F''(x) \leq -r < 0$. Then F' has at most one zero $c \in [a,b]$. Take another function $G \in \mathcal{C}^1([a,b])$ such that $|G(x)| \leq M$ and $\frac{G(t)}{F'(t)}$ is monotone for $t \neq c$. We get*

$$\left| \int_a^b e^{iF(x)} G(x) dx \right| \leq \frac{12M}{\sqrt{r}}.$$

Proof. Suppose c exists and fix $\delta > 0$. We define $K_1 = [a, c-\delta]$, $K_2 = [c-\delta, c+\delta] \cap [a,b]$ and $K_3 = [c+\delta, b]$. We decompose the integral in the pieces

$$I_j = \int_{K_j} e^{iF(x)} G(x) dx.$$

We first estimate

$$|F'(x)| = \left| \int_x^c F''(t) dt \right| \geq r(c-x) \geq r\delta,$$

for $x \in K_1$. Thus, by the lemma above we estimate $|I_1| \leq \frac{8M}{r\delta}$. The same method can be applied to I_3. We estimate I_2 trivially by

$$|I_2| \leq 2\delta M.$$

The claim follows by $\delta = r^{-\frac{1}{2}}$. The case when c does not exist in $[a,b]$ is even easier. $\square$

Lemma 31 (Euler's summation formula)**.** *We have*

$$\sum_{a<n\le b} F(n) = \int_a^b F(\xi)d\xi + \int_a^b (\{\xi\} - \frac{1}{2})F'(\xi)d\xi$$
$$+ (\{a\} - \frac{1}{2})F(a) - (\{b\} - \frac{1}{2})F(b).$$

Proof. By partial summation we get

$$\sum_{a<n\le b} F(n) = \lfloor b\rfloor F(b) - \lfloor a\rfloor F(a) - \int_a^b F'(\xi)\lfloor \xi\rfloor d\xi$$
$$= \lfloor b\rfloor F(b) - \lfloor a\rfloor F(a) + \int_a^b F'(\xi)(\xi - \lfloor \xi\rfloor - \frac{1}{2})d\xi - \int_a^b (\xi - \frac{1}{s})F'(\xi)d\xi.$$

The result follows by partially integrating the last integral. □

Lemma 32. *For $\alpha \notin \mathbb{Z}$ we have*

$$\{\alpha\} - \frac{1}{2} = \sum_{0\ne|m|\le M} \frac{e(-m\alpha)}{2\pi i m} + O(\frac{1}{M\|\alpha\|}).$$

Proof. Without loss of generality we assume $0 < \alpha \le \frac{1}{2}$. We observe that, for $m \ne 0$, we have

$$\int_\alpha^{\frac{1}{2}} e(-mt)dt = \frac{(-1)^{m+1}}{2\pi i m} + \frac{e(-\alpha m)}{2\pi i m}.$$

Summing up this identity and completing the geometric series in the integral yields

$$\sum_{0\ne|m|\le M} \frac{e(-m\alpha)}{2\pi i m} - \alpha + \frac{1}{2} = \int_\alpha^{\frac{1}{2}} \sum_{|m|\le M} e(mt)dt = \int_\alpha^{\frac{1}{2}} \frac{\sin((2M+1)\pi t)}{\sin(\pi t)}dt.$$

By the mean value theorem we get

$$\sum_{0\ne|m|\le M} \frac{e(-m\alpha)}{2\pi i m} - \alpha + \frac{1}{2} = \int_\alpha^{\xi} \frac{\sin((2M+1)\pi t)}{\sin(\pi \alpha)}dt.$$

This implies the result by estimating the integral trivially and using the bound $\sin(\pi\alpha)^{-1} \le \|\alpha\|^{-1}$. □

This lemma shows that we have the Fourier series

$$\sum_{m\ne 0} \frac{e(-m\alpha)}{2\pi i m} = \begin{cases} \{\alpha\} - \frac{1}{2} & \text{if } \alpha \notin \mathbb{Z}, \\ 0 & \text{else,} \end{cases} \tag{32}$$

when the terms of opposite sign are always summed together. One can further refine the bound

$$\sum_{0\ne|m|\le M} \frac{e(-m\alpha)}{2\pi i m} - \alpha + \frac{1}{2} = \int_\alpha^{\frac{1}{2}} \frac{\sin((2M+1)\pi t)}{\sin(\pi t)}dt$$
$$= \int_\alpha^{\frac{1}{2}} \frac{\sin((2M+1)\pi t)}{\pi t}dt + \int_\alpha^{\frac{1}{2}} \sin((2M+1)\pi t)\left(\frac{1}{\sin(\pi t)} - \frac{1}{\pi t}\right)dt$$

and see that the partial sums are uniformly bounded by treating the integrals trivially.

Proposition 2 (Van der Corput summation)**.** *Let $\eta > 0$. Further take $f, g \in \mathcal{C}^1([a,b])$ such that f', g and $|g'|$ are monotone decreasing and $g \geq 0$. We have*

$$\sum_{a<n\leq b} g(n)e(f(n)) = \sum_{f'(b)-\eta<h<f'(a)+\eta} \int_a^b g(x)e(f(x)-hx)dx \\ + O_\eta(|g'(a)| + g(a)\log(|f'(a)| + |f'(b)| + 2)).$$

Proof. We apply Euler's summation formula and insert the Fourier series for $\{\alpha\} - \frac{1}{2}$. This leads to

$$\begin{aligned}\sum_{a<n\leq b} g(n)e(f(n)) &= \int_a^b g(x)e(f(x))dx + \int_a^b \left(\sum_{m\neq 0} \frac{e(-m\alpha)}{2\pi i m}\right) \frac{d}{dx}(g(x)d(f(x)))dx + O(g(a)) \\ &= \int_a^b g(x)e(f(x))dx + \sum_{m\neq 0} \frac{1}{m}\Bigg(\underbrace{\int_a^b f'(x)g(x)e(f(x)-mx)dx}_{=I_1(m)} \\ &\qquad + \frac{1}{2\pi i}\underbrace{\int_a^b g'(x)e(f(x)-mx)dx}_{=I_2(m)}\Bigg) + O(g(a)).\end{aligned}$$

Here interchanging summation and integration is justified due to the convergence properties discussed above. We partially integrate I_1 and find

$$I_1(m) = -\frac{1}{2\pi i}I_2(m) + m\int_a^b f(x)e(f(x)-mx)dx + O(g(a)).$$

The integral $\int_a^b g(x)e(f(x))dx$ can either be included in the sum if $f'(b) - \eta \leq 0 \leq f'(a) + \eta$ or it can simply be estimated by $O(g(a))$. Thus we get

$$\sum_{a<n\leq b} g(n)e(f(n)) = \sum_{f'(b)-\eta<h<f'(a)+\eta} \int_a^b g(x)e(f(x)-hx)dx$$
$$+O\Bigg(g(a) + \sum_{\substack{f'(b)-\eta<h<f'(a)+\eta,\\ h\neq 0}} \frac{g(a)}{h} + \sum_{\substack{h\in\mathbb{Z}\setminus\{0\},\\ h\notin(f'(b)-\eta,f'(a)+\eta)}} \frac{|I_1(h)|}{|h|} + \sum_{\substack{h\in\mathbb{Z}\setminus\{0\},\\ h\notin(f'(b)-\eta,f'(a)+\eta)}} \frac{|I_2(h)|}{|h|}\Bigg).$$

We only have to treat the error term. Of course only the terms involving the integrals $I_1(h)$ and $I_2(h)$ are non-trivial.Without loss of generality we assume $h \geq f'(a) + \eta$ and $f'(b) > 0$ (the other case works analogous). We start by estimating $I_1(h)$. By Lemma 29 we get

$$I_1(h) \ll \left|\frac{g(a)f'(a)}{f'(a)-h}\right|.$$

This implies

$$\sum_{\substack{h\geq f'(a)+\eta,\\ h\neq 0}} \frac{|I_1(h)|}{|h|} \ll g(a) \sum_{0<h\leq 2|f'(h)|} \frac{1}{h} + g(a) \sum_{h>|f'(a)|} \frac{|f'(a)|}{h^2}$$
$$\ll g(a)\log(|f'(a)|+|f'(b)|+2).$$

Finally we deal with the $I_2(h)$ terms. here we only show how to treat $\Re(I_2(h))$ as the imaginary part works similar. We start by observing that

$$\Re(I_2(h)) = -\int_a^b |g'(x)|\cos(2\pi(f(x)-hx))dx = g'(a)\int_a^\xi \cos(2\pi(f(x)-hx))dx$$
$$\ll \frac{1}{|h|} + \frac{1}{|h|}\cdot\frac{|f'(a)|}{|f'(a)-m|}.$$

The final estimate can be seen by partial integration applied to the remaining integral. We conclude that

$$\sum_{\substack{h\geq f'(a)+\eta,\\ h\neq 0}} \frac{|I_1(h)|}{|h|} \ll |g'(a)| \sum_{m\neq 0} \frac{1}{|h|^2} \ll |g'(a)|\,.$$

□

Corollary 4. *We have*

$$\sum_{n=1}^{N} n^{-it} \ll \frac{N}{t} + t^{\frac{1}{2}}\log(t).$$

Proof. The claimed bound follows from

$$\sum_{M\leq n\leq 2M} m^{it} \ll \begin{cases} \frac{M}{t} & \text{if } 1\leq t<5M,\\ \sqrt{t} & \text{if } g\geq 5M. \end{cases}$$

We use the just developed van der Corput summation formula to find

$$\sum_{M\leq n\leq 2M} m^{it} = \sum_{\frac{t}{4\pi M}-\eta<k<\frac{t}{2\pi M}+\eta} \int_M^{2M} e(f(x)-kx)dx + O(\log(2+\frac{t}{M})),$$

for $f(x) = \frac{t}{2\pi}\log(x)$ and $\eta = \frac{1}{20}$.

If $t\leq 5M$, then the k-sum only contains the term $k=0$. This term can be bounded by applying Lemma 29 and $f'(x) \gg \frac{t}{M}$.

If $t\geq 5M$, we observe

$$\frac{d^2}{dx^2}(f(x)-kx) = f''(x) < 0 \text{ and } |f''(x)| \gg tM^{-2}.$$

Thus according to Lemma 30 we can estimate the integral by $\ll Mt^{-\frac{1}{2}}$. The bound follows by summing over admissible k. □

We now turn towards the so called approximate functional equation. This is a tool similar to (24) but in practice it works much better.

Proposition 3 (Approximate functional equation)**.** *Let $s = \sigma + it$ with $0 < \sigma < 1$. For $xy = \frac{t}{2\pi}$ we have*

$$\zeta(s) = \sum_{n \le x} n^{-s} + \frac{(2\pi)^s}{2\cos(\frac{\pi s}{2})\Gamma(s)} \sum_{n \le y} n^{s-1} + O((x^{-\sigma} + t^{\frac{1}{2}-\sigma} y^{\sigma-1}) \log(t)).$$

This formula exists in many forms. Normally a smooth version is stated which holds for a wide class of L-functions including all $L(s,\chi)$. For our purposes this formula will suffice.

Proof. We apply van der Corput's summation formula with $g(z) = z^{-\sigma}$ and $f(z) = -\frac{t}{2\pi}\log(z)$ and get

$$\sum_{x<n\le N} n^{-s} = \sum_{\frac{t}{2\pi N} - \eta < h < y+\eta} \int_x^N e(\xi h)\xi^{-s} d\xi + O(x^{-\sigma}\log(\frac{t}{x} + \frac{t}{N} + 2)).$$

We take $\eta = \frac{1}{4}$, $N > t$ and assume y to be a half-integer. Then the h-sum ranges from 0 to $y - \frac{1}{2}$. Note that the term $h = 0$ can be computed explicitly. With (24) we have

$$\zeta(s) = \sum_{n \le x} n^{-s} + \sum_{h=1}^{y-\frac{1}{2}} \int_x^N e(\xi h)\xi^{-s} d\xi - \frac{x^{1-s}}{1-s} + O(\frac{t}{\sigma}N^{-\sigma} + x^{-\sigma}\log(t)).$$

We make several observations. First,

$$\left|\frac{x^{1-s}}{1-s}\right| \ll \frac{x^{1-\sigma}}{t} \ll x^{-\sigma}$$

can be absorbed in the error. Second, due to the lemma above, we have

$$\int_N^\infty \xi^{-s} e(\xi h) d\xi \ll \frac{N^{-\sigma}}{h - \frac{t}{2\pi N}} \ll N^{-\sigma}h^{-1}.$$

Finally we get

$$\int_0^x \xi^{-s} e(\xi h) d\xi = \frac{x^{1-s}}{1-s}e(xh) - \frac{2\pi i h}{1-s}\int_0^x \xi^{1-s} e(\xi h) d\xi \ll x^{1-\sigma}t^{-1} + \frac{h}{t}\cdot\frac{x^{1-\sigma}}{h - \frac{t}{2\pi x}}.$$

With this we can enlarge the Fourier integrals from $[x, N]$ to $[0, \infty)$. Now they can be evaluated. Indeed we have

$$\int_0^\infty e(\xi h)\xi^{-s} d\xi = (-2\pi i h)^{s-1}\Gamma(1-s).$$

Putting all of these together and treating the error trivially we get

$$\zeta(s) = \sum_{n \le x} n^{-s} + (-2\pi i)^{s-1}\Gamma(1-s)\sum_{h \le y} h^{s-1} + O(x^{-\sigma}\log(t) + \frac{t}{\sigma}N^{-\sigma}).$$

Here we can take $N \to \infty$ and simplify the error term accordingly. Using a standard functional equation for the Gamma function we also find

$$(-2\pi i)^{s-1}\Gamma(1-s) = \frac{(2\pi)^s}{s\cos(\frac{\pi s}{2})\Gamma(s)}(1 + O(e^{-\pi t}))$$

this gives the formula stated under the assumption that y is a half-integer.

To remove this final restriction we argue as follows. First, if $y \geq x$ we can simply replace y by $\lfloor y \rfloor + \frac{1}{2}$ and adjust x according to $xy = \frac{t}{2\pi}$. The error we make with this replacements is $\ll x^{-\sigma}$. If $y < x$ We can exchange the roles of x and y an apply the just established formula for $\zeta(1-s)$. The claim then follows from the functional equation. $\square$

We will now refine our machinery to estimate sums of the form

$$S_f = \sum_{n=a+1}^{b} e(f(n)).$$

Lemma 33. *Suppose $f \in \mathcal{C}^2([a,b])$ and $0 < \lambda_2 \leq |f''(x)| \leq h\lambda_2$. Then*

$$\sum_{a<n\leq b} e(f(n)) \ll h(b-a)\lambda_2^{\frac{1}{2}} + \lambda_2^{-\frac{1}{2}}.$$

Proof. Without loss of generality we can assume $f''(x) < 0$. Then the result follows from the van der Corput summation formula and Lemma 30. $\square$

Lemma 34. *For $q \leq b-a$ we have*

$$\sum_{a<n\leq b} e(f(n)) \ll \frac{b-a}{q^{\frac{1}{2}}} + \left(\frac{b-a}{q} \sum_{r=1}^{q-1} \left| \sum_{a<n\leq b-r} e(f(n+r) - f(n) \right| \right)^{\frac{1}{2}}.$$

Proof. We have

$$\begin{aligned}
\left| \sum_{a<n\leq b} e(f(n)) \right| &= \frac{1}{q} \left| \sum_{a-q<n\leq b-1} \sum_{m=1}^{q} e(f(m+n)) \right| \\
&\leq \frac{1}{q} \sum_{a-q<n\leq b-1} \left| \sum_{m=1}^{q} e(f(m+n)) \right| \\
&\leq \frac{1}{q} \left(\underbrace{(b-a+q)}_{2(b-a)} \sum_{a-q<n\leq b-1} \left| \sum_{m=1}^{q} e(f(m+n)) \right|^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

By expanding the square we get

$$\left| \sum_{m=1}^{q} e(f(m+n)) \right|^2 = q + \sum_{\mu<m} e(f(m+n) - f(\mu+n)) + \sum_{m<\mu} e(f(\mu+n) - f(m+n)).$$

Thus, summing this over n yields

$$\begin{aligned}\left|\sum_{m=1}^{q} e(f(m+n))\right|^2 &= (b-a+q)q + 2\left|\sum_{n}\sum_{\mu<m} e(f(m+n)-f(\mu+n))\right| \\ &= (b-a+q)q + 2\left|\sum_{r=1}^{q-1}(q-r)\sum_{a<\nu\le b-r} e(f(\nu+r)-f(\nu))\right| \\ &\le 2(b-a)q + 2q\sum_{r=1}^{q-1}\left|\sum_{a<\nu\le b-r} e(f(\nu+r)-f(\nu))\right|.\end{aligned}$$

In the middle step we put $\nu=\mu+n$ and $r=m-\nu$. Using this in our first inequality yields

$$\left|\sum_{n} e(f(n))\right| \le \frac{1}{q}\left(4(b-a)^2q + 4(b-a)q\sum_{r=1}^{q-1}\left|\sum_{a<\nu\le b-r} e(f(\nu+r)-f(\nu))\right|\right)^{\frac{1}{2}}.$$

This easily implies the statement. $\square$

Proposition 4 (kth Derivative Test)**.** *Let $f\in\mathcal{C}^k([a,b],\mathbb{R})$ for $k\ge 2$. Suppose $\lambda_k \le \left|f^{(k)}(x)\right| \le h\lambda_k$. Then, for $b-a\ge 1$ and $K=2^{k-1}$ we have*

$$\sum_{a<n\le b} e(f(n)) \ll h^{\frac{2}{K}}(b-a)\lambda_k^{\frac{1}{2K-2}} + (b-a)^{1-\frac{2}{K}}\lambda_k^{-\frac{1}{2K-2}}.$$

The implicit constant is independent of k.

Proof. Note that the statement is trivial if $\lambda_k\ge 1$. Thus we assume the contrary. Note that for $k=2$ we already established the result above. Thus we argue by induction on k.

Put $g(x)=f(x+r)-f(x)$. Then we have

$$g^{(k-1)}(x) = f^{(k-1)}(x+r) - f^{(k-1)}(x) = rf^{(k)}(\xi),$$

for some $\xi\in[x,x+r]$. Thus, by induction hypothesis, we have

$$\left|\sum_{a<n\le b-r} e(f(n))\right| \le A_1 h^{\frac{4}{K}}(b-a)(r\lambda_k)^{\frac{1}{K-2}} + A_2(b-a)^{1-\frac{4}{K}}(r\lambda_k)^{-\frac{1}{K-2}}.$$

Note that

$$\sum_{r=1}^{q} r^{-\frac{1}{K-2}} < \int_0^q r^{-\frac{1}{K-2}}\,dr = \frac{q^{1-\frac{1}{K-2}}}{1-\frac{1}{K-2}} \le 2q^{1-\frac{1}{K-2}},$$

for $K\ge 4$. Thus we get

$$\sum_{r=1}^{q-1}\left|\sum_{a<n\le b-r} e(f(n))\right| \le A_1 h^{\frac{4}{K}}(b-a)q^{1+\frac{1}{K-2}}\lambda_k^{\frac{1}{K-2}} + 2A_2(b-a)^{1-\frac{4}{K}}q^{1-\frac{1}{K-2}}\lambda_k^{-\frac{1}{K-2}}.$$

The previous lemma thus yields

$$\left|\sum_n e(f(n))\right| \leq A_3(b-a)q^{-\frac{1}{2}}+A_4A_1^{\frac{1}{2}}h^{\frac{2}{K}}(b-a)q^{\frac{1}{2K-4}}\lambda_k^{\frac{1}{2K-4}}+A_4(2A_2)^{\frac{1}{2}}(b-a)^{1-\frac{2}{K}}q^{-\frac{1}{2K-4}}\lambda_k^{-\frac{1}{2K-4}}.$$

We equalise the second and third term by picking

$$\lambda_k^{-\frac{1}{K-1}} \leq q \leq 2\lambda_k^{-\frac{1}{K-1}}.$$

Here we implicitly assume $2\lambda_k^{-\frac{1}{K-1}} \leq b-a$. We get

$$\left|\sum_n e(f(n))\right| \leq (A_3+2A_4A_1^{\frac{1}{2}})h^{\frac{2}{K}}(b-a)\lambda_k^{\frac{1}{2K-2}} + A_4(2A_2)^{\frac{1}{2}}(b-a)^{1-\frac{2}{K}}\lambda_k^{-\frac{1}{2K-2}} \tag{33}$$

The result follows since for large enough A_1, A_2 we have

$$A_3+2A_4A_1^{\frac{1}{2}} \leq A_1 \text{ and } A_4(2A_2)^{\frac{1}{2}} \leq A_2.$$

It remains to treat the case $2\lambda_k^{-\frac{1}{K-1}} \geq b-a$. However, in the latter case the trivial estimate $\sum_n e(f(n)) \ll (b-a)$ suffices. □

Theorem 13. *For $s = 1-\frac{l}{2L-2}+it$ with $L = 2^{l-1}$ we have*

$$\zeta(s) \ll t^{\frac{1}{2L-2}}\log(t).$$

Furthermore, we have

$$\zeta(1+it) \ll \frac{\log(t)}{\log\log(t)}.$$

Proof. We start with the first estimate. We want to apply the kth derivative test with $f(x) = -\frac{t}{2\pi}\log(x)$. We compute $f^{(k)}(x) = (-1)^k\frac{(k-1)!t}{2\pi x^k}$. In particular, for $a \leq x \leq b \leq 2a$ we have

$$\frac{(k-1)!t}{2\pi(2a)^k} \leq f^{(k)}(x) \leq \frac{(k-1)!t}{2\pi a^k}.$$

Thus we can use $\lambda_k = \frac{(k-1)!t}{2\pi(2a)^k}$ and $h = 2^k$. Thus we get

$$\sum_{a<n\leq b} e(f(n)) \ll a^{1-\frac{k}{2K-2}}t^{\frac{1}{2K-2}} + a^{1-\frac{2}{K}+\frac{k}{2K-2}}t^{-\frac{1}{2K-2}}.$$

It is easy to check that the implicit constants do not depend on k. Note that if $a < t^{\frac{K}{kK-2K+2}}$, then the second term is dominated by the first. In this case partial summation yields

$$\sum_{a<n\leq b} n^{-s} \ll a^{\frac{l}{2L-2}-\frac{k}{2K-2}}t^{\frac{1}{2K-2}}.$$

Using dyadic dissections we find

$$\sum_{1\leq\leq t^{\frac{L}{lL-2L+2}}} n^{-s} \ll t^{\frac{1}{2L-2}}\log(t) \text{ and}$$

$$\sum_{2^{-m}t<n\leq 2^{-m+1}t} n^{-s} \ll t^{(\frac{l}{2L-2}-\frac{k}{2K-2})\frac{K}{(k+1)K-2K+1}+\frac{1}{2K-2}}, \text{ for } t^{\frac{K}{(k+1)K-2K+1}} \leq 2^{-m} \leq t^{\frac{K}{kK-2K+2}}.$$

Note that for $2 \le k < l$ we have $L - K = (2^{l-k} - 1)K \ge (l-k)K$ so that

$$(\frac{l}{2L-2} - \frac{k}{2K-2})\frac{K}{(k+1)K - 2K + 1} + \frac{1}{2K-2} \le \frac{1}{2L-2}.$$

Thus we get the bound

$$\sum_{t^{\frac{L}{lL-2L+2}} < n \le t} n^{-s} \ll t^{\frac{1}{2L-2}} \log(t).$$

The desired bound follows from the approximate functional equation.

We turn towards the second bound. First observe that arguing as above and using partial summation once again we get

$$\sum_{a<n\le b} n^{-1-it} \ll a^{-\frac{k}{2K-2}} t^{\frac{1}{2K-2}} \ll t^{-\frac{1}{2(k-1)K+2}}$$

for $a < b \le 2a$ and $t^{\frac{K}{(k+1)K-2K+1}} < a \le t^{\frac{K}{kK-2K+2}}$ and the implicit constant does not depend on k. We fix $r = \lfloor \log\log(t) \rfloor$ and write $R = 2^{r-1} \le 2^{\log\log(t)-1} = \frac{1}{2}\log(t)^{\log(2)}$. Using our discussion for $2 \le k \le r$ combined with a dyadic decomposition we get

$$\sum_{t^{\frac{R}{(r-1)R+1}} < n \le t} n^{-1-it} \ll t^{-\frac{1}{2(r-1)R+2}} \log(t).$$

Since

$$t^{\frac{1}{2(r-1)R+2}} > \exp(\log(t)^{\epsilon}) > \epsilon \log(t)$$

the sum above is bounded. Thus, from (24) we infer that

$$\zeta(1+it) \ll 1 + \underbrace{\left| \sum_{1 \le n \le t^{\frac{2}{r+2}} \log(t)^{-r}} n^{-1-it} \right|}_{\ll \frac{\log(t)}{r} = \frac{\log(t)}{\log\log(t)}}$$

and we are done. $\square$

Corollary 5 (Hardy-Littlewood zero-free region)**.** *There is a constant $A > 0$ such that $\zeta(s)$ has no zeros in the region $\sigma \ge 1 - A\frac{\log\log(t)}{\log(t)}$ and $t \gg 1$. Further one has the bounds*

$$\zeta(s) \ll \log(t)^5 \text{ and } \frac{1}{\zeta(s)}, \frac{\zeta'}{\zeta}(s) \ll \frac{\log(t)}{\log\log(t)}$$

in the same region.

Proof. First we pick $l = \lfloor \log(2)^{-1} \log(\frac{\log(t)}{\log\log(t)^2}) \rfloor$. In particular $\frac{l}{2L-2} \ge \frac{1}{2}\frac{\log\log(t)^2}{\log(t)}$. Then we have

$$\zeta(1 - \frac{l}{2L-2} + it) \ll t^{\frac{1}{2L-2}} \log(t) \ll \log(t)^5.$$

Combining this with the estimate for $\zeta(1+it)$ and the Pharagmen-Lindelöf principle we get $\zeta(s) \ll \log(t)^5$ uniformly in the region $1 - \frac{1}{2}\frac{\log(t)}{\log\log(t)^2} \le \sigma \le 1$ for $t \gg 1$.

The zero free region and the remaining estimates follow now from the general framework developed in Exercise 7. □

This is a slight improvement of the standard zero-free and it is not obvious that it justifies all the work we put into it. However, we will shortly see that it has some interesting applications.

Theorem 14 (Vinogradov-Korobov). *There is $c > 0$ such that $\zeta(s)$ has no zeros in the region $\sigma \ge 1 - c\log(t)^{-\frac{2}{3}}\log\log(t)^{-\frac{1}{3}}$, $t \ge 3$. Furthermore one has the bounds*

$$\zeta(s) \ll \log(t)^{\frac{2}{3}} \text{ and } \frac{1}{\zeta(s)}, \frac{\zeta'}{\zeta}(s) \ll \log(t)^{\frac{2}{3}}\log\log(t)^{\frac{1}{3}}.$$

in this region.

Exercise 7 (Landau). *Show the following general result, which completes the proof of Corollary 5. Assume $\zeta(s) \ll e^{\phi(t)}$ in $1 - \theta(t) \le \sigma \le 2$ for positive non-decreasing functions ϕ, θ^{-1}. Further assume $\theta(t) \le 1$ and $\frac{\phi(t)}{\theta(t)} = o(e^{\phi(t)})$. Then there is a constant A_1 such that $\zeta(s)$ has no-zeros in*

$$\sigma \ge 1 - A_1\frac{\theta(2t+1)}{\phi(2t+1)}.$$

Hint: *One can use the following statement which can be derived from basic principles of complex analysis: Suppose $f(s)$ is regular and satisfies $\left|\frac{f(s)}{f(s_0)}\right| < e^M$ for $s \in B_r(s_0)$. If there is a zero ρ_0 between $s_0 - \frac{r}{2}$ ans s_0, then*

$$-\Re(\frac{f'(s_0)}{f(s_0)}) < \frac{AM}{r} - \frac{1}{s_0 - \rho_0}.$$

If there is no such zero, the last term can be dropped.

Solution. Let $\beta + i\gamma$ be a zero of $\zeta(s)$. Without loss of generality we assume $\gamma > 0$. Pick $1 + e^{-\phi(2\gamma+1)} \le \sigma_0 \le 2$ and set $s_0 = \sigma_0 + i\gamma$, $s_0' = \sigma_0 + 2i\gamma$ and $r = \theta(2\gamma+1)$. We have

$$|\zeta(s_0)| < \frac{A}{\sigma_0 - 1} < Ae^{\phi(2\gamma+1)}$$

and a similar estimate holds for s_0'. By assumption we get

$$\left|\frac{\zeta(s)}{\zeta(s_0)}\right|, \left|\frac{\zeta(s')}{\zeta(s_0')}\right| < e^{A_2\phi(2\gamma+1)}$$

in $|s - s_0|, |s' - s_0'| \le r$. In the case $\beta \le \sigma_0 - \frac{1}{2}r$ one can easily check that the statement holds. Thus we assume $\beta > \sigma_0 - \frac{1}{2}r$, so that the hint implies

$$-\Re(\frac{\zeta'(\sigma_0 + i\gamma)}{\zeta(\sigma_0 + i\gamma)}) < \frac{A_3\phi(2\gamma+1)}{\theta(2\gamma+1)} - \frac{1}{\sigma_0 - \beta}.$$

Further, if σ_0 is close enough to 1, we get $-\frac{\zeta'(\sigma_0)}{\zeta(\sigma_0)} < \frac{a}{\sigma_0 - 1}$. Inserting these observations in (28) we find

$$\frac{3a}{\sigma_0 - 1} + \frac{5A_3\phi(2\gamma+1)}{\theta(2\gamma+1)} - \frac{4}{\sigma - \beta}.$$

This yields

$$1 - \beta \geq \left(\frac{3a}{4\sigma - 1} + \frac{5A_3}{4} \cdot \frac{\phi(2\gamma+1)}{\theta(2\gamma+1)}\right)^{-1} - (\sigma_0 - 1).$$

By making σ_0 close enough to 1 we can take $a = \frac{5}{4}$. For γ large enough one can then use the condition $\frac{\phi(t)}{\theta(t)} = o(e^{\phi(t)})$ and deduce

$$1 - \beta \geq A_4 \frac{\theta(2\gamma+1)}{\phi(2\gamma+1)}.$$

□

5 Part 5: Zero-density estimates

In this part we ultimately discuss zero-density estimates. These are important supplements to the available zero-free regions.

5.1 Mean value estimates for Dirichlet polynomials

Lemma 35 (Duality)**.** *Given a matrix* $A = (a_{nm}) \in M_{m\times n}(\mathbb{C})$*. Then the following two assertions are equivalent*

- *For all* $\mathbf{y} \in \mathbb{C}^N$ *we have*

$$\sum_{m=1}^{M} \left| \sum_{n=1}^{N} a_{mn} y_n \right|^2 \leq c^2 \|\mathbf{y}\|_2^2.$$

- *For all* $\mathbf{x} \in \mathbb{C}^M$ *we have*

$$\sum_{n=1}^{N} \left| \sum_{m=1}^{M} a_{mn} x_m \right|^2 \leq c^2 \|\mathbf{x}\|_2^2.$$

Proof. We introduce an intermediate step. Using the first bullet-point and Cauchy-Schwarz we get

$$\left| \sum_{m=1}^{M} x_m \sum_{n=1}^{N} a_{mn} y_m \right|^2 \leq \|\mathbf{x}\|_2^2 \sum_{m=1}^{M} \left| \sum_{n=1}^{N} a_{mn} y_n \right|^2 \leq c^2 \|\mathbf{x}\|_2^2 \|\mathbf{y}\|_2^2. \tag{34}$$

We now show that (34) implies the first bullet-point. By duality this completes the proof of the theorem. To see the required implication we set $x_m = \sum_{n=1}^{N} a_{mn} y_n$. Then

$$\|\mathbf{x}\|_2^2 = \sum_{m=1}^{M} \overline{x}_m \sum_{n=1}^{N} a_{mn} y_n \leq c \|\mathbf{x}\|_2 \|\mathbf{y}\|_2,$$

where we applied (34). Now if $\|\mathbf{x}\|_2 = 0$ we are done. Otherwise one concludes by dividing both sides by $\|\mathbf{x}\|_2 = 0$. ☐

Let us record the following inequality. Put

$$R = \max_m \sum_n |a_{mn}| \text{ and } C = \max_n \sum_m |a_{mn}|.$$

Via Cauchy-Schwarz we get

$$\left| \sum_{n,m} a_{mn} x_m y_n \right|^2 \leq \left(\sum_{m,n} |a_{mn}| \, |x_m|^2 \right) \left(\sum_{m,n} |a_{mn}| \, |x_m|^2 \right) \left(\sum_{m,n} |a_{mn}| \, |y_n|^2 \right) \leq RC \|\mathbf{x}\|_2^2 \|\mathbf{y}\|_2^2. \tag{35}$$

We introduce the **Dirichlet polynomials**

$$D(s) = \sum_{n=1}^{N} a_n n^{-s}.$$

Proposition 5. *Suppose* $0 \leq t_1 \leq t_2 < \ldots < t_R \leq T$ *satisfying* $t_{r+1} - t_r \geq 1$. *Then*

$$\sum_{r=1}^{R} |D(it_r)|^2 \ll (N + RT^{\frac{1}{2}}) \log(T) \sum_{n=1}^{N} |a_n|^2 .$$

Proof. By Lemma 35 it is enough to show that

$$\sum_{n=1}^{N} \left| \sum_{r=1}^{R} y_r n^{-it_r} \right| \ll (N + RT^{\frac{1}{2}}) \log(T) \sum_{r=1}^{R} |y_r|^2 .$$

Opening the square and interchanging the order of summation gives us

$$\sum_{n=1}^{N} \left| \sum_{r=1}^{R} y_r n^{-it_r} \right| = N |y_r|^2 + \sum_{\substack{1 \leq q,r \leq R, \\ q \neq r}} Z_N(i(t_q - t_r)) y_q \overline{y}_r,$$

for $Z_N(s) = \sum_{n=1}^{N} n^{-s}$. According to (35) we have

$$\sum_{n=1}^{N} \left| \sum_{r=1}^{R} y_r n^{-it_r} \right| \leq \left(N + \max_{1 \leq q \leq R} \sum_{\substack{1 \leq r \leq R, \\ r \neq q}} |Z(i(t_q - t_r))| \right) \sum_{r=1}^{R} |y_r|^2 .$$

The result follows from

$$Z(it) \ll \frac{N}{t} + t^{\frac{1}{2}} \log(t),$$

which we showed in Corollary 4. □

Proposition 6. *We have*

$$\int_0^T |D(it)|^2 \, dt \ll (T + N) \sum_{n=1}^{N} |a_n|^2 .$$

Proof. We define the function

$$f(t) = \begin{cases} 1 + \frac{t}{N} & \text{if } -N < t \leq T, \\ 1 & \text{if } 0 < t \leq T, \\ 1 - \frac{t-T}{N} & \text{if } T < t \leq T + N, \\ 0 & \text{else.} \end{cases}$$

By positivity we have

$$\int_0^T |D(it)|^2 \, dt \leq \int_{\mathbb{R}} f(t) \, |D(it)|^2 = \sum_{m,n} a_m \overline{a_n} \int_R f(t) \left(\frac{m}{n} \right)^{it} dt.$$

We use that $\int_{\mathbb{R}} f(t) dt = N + T$ and the bound

$$\int_{\mathbb{R}} f(t) x^{it} dt \ll N^{-1} \log(x)^{-2} \text{ for } x \neq 1.$$

Indeed this yields

$$\int_{R} f(t) \left(\frac{m}{n}\right)^{it} dt \ll \frac{1}{N} \log(\frac{m}{n})^{-2} \ll \frac{1}{N} \left(\frac{m+n}{m-n}\right)^2 \ll \frac{N}{(n-m)^2} \text{ for } n \neq m.$$

We obtain the estimate

$$\int_0^T |D(it)|^2 \, dt \ll (T+N) \sum_n |a_n|^2 + N \sum_{n \neq m} |a_n a_m| \, (m-n)^2.$$

We are done since $2 |a_n a_m| \leq |a_n|^2 + |a_m|^2$. □

We can use this continuous mean value theorem into a discrete one by applying the following interesting inequality.

Lemma 36 (Gallagher)**.** *Let $\mathcal{T} = \{t_1, \ldots, t_R\}$ be a set of points with $\frac{1}{2} \leq t_r \leq T - \frac{1}{2}$ and $|t_{r_1} - t_{r_2}| \geq 1$ for all $r_1 \neq r_2$. For a smooth function $F\colon [0,T] \to \mathbb{C}$ we have*

$$\begin{aligned} \sum_r |F(t_r)|^2 &\leq \int_0^T (|F(t)|^2 + |F(t)F'(t)|) dt \\ &\leq \int_0^T (|F(t)|^2 \, dt + \left(\int_0^T (|F(t)|^2 \, dt\right)^{\frac{1}{2}} \left(\int_0^T (|F'(t)|^2 \, dt\right)^{\frac{1}{2}}. \end{aligned}$$

Proof. By partial integration we have the identity

$$f(x) = \int_0^1 f(t)dt + \int_0^x t f'(t) dt + \int_x^1 (t-1) f'(t) dt.$$

Taking $x = \frac{1}{2}$ yields

$$\left| f(\frac{1}{2}) \right| \leq \int_0^1 (|f(t)| + \frac{1}{2} |f'(t)|) dt.$$

Replacing f by f^2 gives

$$\left| f(\frac{1}{2}) \right|^2 \leq \int_0^1 (|f(t)|^2 + \frac{1}{2} |f(t) f'(t)|) dt.$$

We get

$$|F(t_r)|^2 \leq \int_{t_r - \frac{1}{2}}^{t_r + \frac{1}{2}} (|F(t)|^2 + \frac{1}{2} |F(t) F'(t)|) dt.$$

The result follows by summing over r and combining the integrals. This is possible because the integrals do not overlap. □

Corollary 6. *Suppose $0 \leq t_1 \leq t_2 < \ldots < t_R \leq T$ satisfying $t_{r+1} - t_r \geq 1$. Then*

$$\sum_{r=1}^R |D(it_r)|^2 \ll (N+T) \log(2N) \sum_{n=1}^N |a_n|^2.$$

Proof. The proof follows directly by combining Lemma 36 with Proposition 6 after observing that $D'(s)$ is also a Dirichlet Polynomial. □

Corollary 7. *Suppose* $0 \le t_1 \le t_2 < \ldots < t_R \le T$ *satisfying* $t_{r+1} - t_r \ge 1$. *Then*

$$\sum_{r=1}^{R} \left|\zeta(\frac{1}{2} + it_r)\right|^4 \ll T \log(T)^5.$$

Proof. We start from the approximate functional equation (Proposition 3) with $x \asymp \sqrt{t}$ to get

$$\left|\zeta(\frac{1}{2} + it)\right| \ll \left|Z_x(\frac{1}{2} + it)\right| + \left|Z_{\frac{t}{2\pi x}}(\frac{1}{2} - it)\right| + t^{-\frac{1}{4}} \log(t).$$

Thus, for $t \asymp T$ we get

$$\begin{aligned}\left|\zeta(\frac{1}{2} + it)\right|^4 &\ll \frac{1}{\sqrt{T}} \int_{\sqrt{T}}^{2\sqrt{T}} \left|Z_x(\frac{1}{2} + it)\right|^4 + \left|Z_{\frac{t}{2\pi x}}(\frac{1}{2} - it)\right|^4 dx + T^{-1} \log(T)^4 \\ &= \frac{1}{\sqrt{T}} \int_{\sqrt{T}}^{2\sqrt{T}} \left|Z_x(\frac{1}{2} + it)\right|^4 dt + \frac{1}{\sqrt{T}} \int_{\frac{t}{2\pi\sqrt{T}}}^{\frac{t}{4\pi\sqrt{T}}} \left|Z_y(\frac{1}{2} - it)\right|^4 \frac{t}{2\pi y^2} dy + T^{-1} \log(T)^4 \\ &\ll \frac{1}{\sqrt{T}} \int_{A\sqrt{T}}^{B\sqrt{T}} \left|Z_x(\frac{1}{2} + it)\right|^4 dt + T^{-1} \log(T)^4.\end{aligned}$$

By considering dyadic decompositions it is enough to consider $t_1, \ldots, t_R \asymp T$. In particular, since they are well spaced, we have $R \ll T$. We conclude

$$\sum_{r=1}^{R} \left|\zeta(\frac{1}{2} + it_r)\right|^4 \ll \frac{1}{\sqrt{T}} \int_{A\sqrt{T}}^{B\sqrt{T}} \sum_{r=1}^{R} \left|Z_x(\frac{1}{2} + it)\right|^4 dt + \log(T)^4.$$

Next we write

$$Z_x(\frac{1}{2} + it)^2 = \sum_{n \le x^2} \underbrace{\left(\sum_{\substack{de=n,\\ d,e \le x}} 1 \right)}_{=a_x(n)} n^{-\frac{1}{2} - it}.$$

Of course we have the trivial bound $a_x(n) \le d(n)$. Applying Corollary 6 yields

$$\sum_{r=1}^{R} \left|Z_x(\frac{1}{2} + it)\right|^4 dt \ll (x^2 + T) \log(2x^2) \sum_{n=1}^{T} \frac{d(n)^2}{n}.$$

It can be shown by elementary means that $\sum_{n=1}^{T} \frac{d(n)^2}{n} \ll \log(T)^4$.Thus we get

$$\sum_{r=1}^{R} \left|\zeta(\frac{1}{2} + it_r)\right|^4 \ll T \log(T)^5.$$

□

We will also need the following estimates on how often a Dirichlet Polynomial can assume large values.

Theorem 15. *Suppose $\{t_1, \ldots, t_R\}$ is a well spaced set of points in $[0, T]$ with*

$$|D(it_r)| \geq V.$$

Then we have

$$R \ll \frac{T+N}{V^2} \log(2N) \sum_{n=1}^{N} |a_n|^2 . \tag{36}$$

Suppose further that $V \geq T^{\frac{1}{4}} \log(2T) \left(\sum_n |a_n|^2\right)^{\frac{1}{2}}$, then

$$R \ll \frac{N}{V^2} \log(2T) \sum_n |a_n|^2 .$$

Finally we have the estimate

$$R \ll \left(\frac{N}{V^2} + \frac{NT}{V^6} \left(\sum_n |a_n|^2 \right)^2 \right) \log(2T)^6 \sum_n |a_n|^2 .$$

again in the general case.

Proof. The first estimate follows directly from

$$R \leq \frac{1}{V^2} \sum_r |D(it_r)|^2 .$$

Turning to the second statement one can assume without loss of generality that $T \gg N$. Now the result follows by the same trick using the estimate

$$RV^2 \ll (N + RT^{\frac{1}{2}}) \log(2T) \sum_n |a_n|^2 .$$

We turn to the final result. Write $G = \sum_n |a_n|^2$ and assume that $V \geq \log(2T) G^{\frac{1}{2}}$. Define

$$T_0 = \min(T, G^{-2} V^4 \log(2T)^{-4}).$$

We have $1 \leq T_0 \leq T$. We now write $\mathcal{T}_l = \{t_r \colon t_r \in [lT_0, (l+1)T_0]\}$ with $l \leq TT_0^{-1}$. We also put $R_l = \sharp \mathcal{T}_l$. By the second statement of the theorem we have

$$R_l \ll GNV^{-2} \log(2T).$$

The result follows after observing that $R = \sum_l R_l$. □

We now turn towards some mean-value estimates for exponential sums

$$S(t) = \sum_{v \geq 0} c(v) e(vt).$$

We assume at least that the sum defining $S(t)$ is absolutely convergent.

Lemma 37. *Let $\delta = \frac{\theta}{T}$, with $0 < \theta < 1$. Then*

$$\int_{-T}^{T} |S(t)|\, dt \ll_\theta \int_{-\infty}^{\infty} \left| \delta^{-1} \sum_{x \le v \le x+\delta} c(v) \right|^2 dx.$$

If $S(t) = \sum_{n \in \mathbb{N}} a_n n^{it}$, then

$$\int_{-T}^{T} |S(t)|\, dt \ll_\theta T^2 \int_{0}^{\infty} \left| \sum_{y \le v \le y e^{\frac{1}{T}}} a_n \right|^2 \frac{dy}{y}.$$

Proof. We start by proving the first part. Put $F_\delta(x) = \delta^{-1} \mathbb{1}_{B_0(\frac{\delta}{2})}(x)$ and

$$C_\delta(x) = \sum_v c(v) F_\delta(x - v) = \sum_{|v - x| \le \frac{\delta}{2}} c(v).$$

Taking Fourier transforms yields $\hat{C}_\delta = S \cdot \hat{f}_\delta$. By the Plancherel theorem we have

$$\int_{-\infty}^{\infty} \left| \delta^{-1} \sum_{x \le v \le x+\delta} c(v) \right|^2 dx = \int_{\mathbb{R}} |C_\delta(x)|^2\, dx = \int_{\mathbb{R}} \left| S(t) \hat{F}_\delta(t) \right|^2 dt.$$

The claim follows since

$$\hat{F}_\delta(t) = \frac{\sin(\pi \delta t)}{\pi \delta t} \gg 1 \text{ for } |t| \le T.$$

The second part follows from the first by taking $\theta = \frac{1}{2\pi}$ and making the change of variables $\log(y) = 2\pi x$. ∎

Lemma 38. *We have*

$$\sum_{\chi \bmod q} \left| \sum_{y \le n \le y+z} a_n \chi(n) \right|^2 \le (q + z) \sum_{y \le n \le y+z} |a_n|^2.$$

Proof. This follows by character orthogonality and Cauchy-Schwarz as follows:

$$\begin{aligned}
\sum_{\chi \bmod q} \left| \sum_{y \le n \le y+z} a_n \chi(n) \right|^2 &= \sum_{n,m} a_n \overline{a_m} \sum_{\chi \bmod q} \chi(n) \overline{\chi(m)} \\
&= \varphi(q) \sum_{\substack{n,m \\ n \equiv m \bmod q}} a_n \overline{a_m} \\
&= \varphi(q) \sum_{\substack{h=1, \\ (h,q)=1}}^{q} \left| \sum_{n \equiv h \bmod q} a_n \right|^2 \\
&\le \varphi(q) \left(\frac{z}{q} + 1 \right) \sum_{\substack{h=1, \\ (h,q)=1}}^{q} \sum_{n \equiv h \bmod q} |a_n|^2 \\
&\le (q + z) \sum_{y \le n \le y+z} |a_n|^2.
\end{aligned}$$

□

Proposition 7. *For $T \geq 1$ we have*

$$\sum_{\chi \bmod q} \int_{-T}^{T} |S(\chi,t)|^2 \, dt \ll \sum_n (qT+n) \, |a_n|^2 \, .$$

Proof. Applying the previous two lemmata we find

$$\sum_{\chi \text{ mod } q} \int_{-T}^{T} |S(\chi,t)|^2 \, dt \ll T^2 \int_0^\infty \sum_{\chi \text{ mod } q} \left| \sum_{y \leq n \leq y e^{\frac{1}{T}}} a_n \right|^2 \frac{dy}{y}$$

$$\ll T^2 \int_0^\infty (q + y(e^{\frac{1}{T}} - 1)) \sum_{y \leq n \leq y e^{\frac{1}{T}}} |a_n|^2 \, \frac{dy}{y} \tag{37}$$

Interchanging summation and integration yields

$$\sum_{\chi \text{ mod } q} \int_{-T}^{T} |S(\chi,t)|^2 \, dt \ll \sum_n |a_n|^2 \int_{ne^{-\frac{1}{T}}}^{n} \frac{T^2 q}{y} +_T^2 (e^{\frac{1}{T}} - 1) dy$$

$$= \sum_n (Tq + T^2(e^{\frac{1}{T}} - 1)(1 - e^{-\frac{1}{T}})n) \, |a_n|^2 \, .$$

The result follows by estimating

$$T^2(e^{\frac{1}{T}} - 1)(1 - e^{-\frac{1}{T}}) = (2T \sinh(\frac{1}{2T}))^2 \ll 1.$$

□

Lemma 39. *We have*

$$\sum_{q \leq Q} \sum_{\substack{a=1, \\ (a,q)=1}}^{q} \left| \sum_{x<n\leq x+z} a_n e(\frac{na}{q}) \right|^2 \leq (z + Q^2) \sum_{x<n\leq x+z} |a_n|^2 \, .$$

Proof. Set $S(\alpha) = \sum_{x<n\leq x+z} a_n e(\alpha n)$.

We start by showing that

$$\left| \sum_{\substack{1 \leq r,s \leq R, \\ r \neq s}} \frac{w_r \overline{w_s}}{\sin(\pi(\alpha_r - \alpha_2))} \right| \leq \delta^{-1} \sum_r |w_r|^2 \, .$$

To see this we use the expansion

$$\frac{\pi}{\sin(\pi\alpha)} = \lim_{J\to\infty} \sum_{|k|\leq J} (1 - \frac{|k|}{J}) \frac{(-1)^k}{\alpha + k}.$$

Thus it is enough to show that

$$\left| \sum_{\substack{1 \leq r,s \leq R, \\ r \neq s}} \sum_{|k|\leq J} (1 - \frac{|k|}{J})(-1)^k \frac{w_r \overline{w_s}}{\alpha_r - \alpha_s + k} \right| \leq \frac{\pi}{\delta} \sum_{r \leq R} |w_r|^2 \, .$$

This can be derived from the following general inequality[15]

$$\sum_{i,j\in I, i\neq j} \frac{v_i\overline{v}_j}{\lambda_i-\lambda_j} \leq \frac{\pi}{\min_{i\neq j}|\lambda_i-\lambda_j|}\sum_{i\in I}|v_i|^2. \tag{38}$$

We let $I=\{1,\dots,R\}\times\{1,\dots,J\}$ and put $v_{(r,m)}=(-1)^m w_r$ and $\lambda_{(r,m)}=\alpha_r+m$. We get

$$\left|\sum_{(r,m)\neq(s,n)}(-1)^{m+n}\frac{w_r\overline{w}_s}{\alpha_r-\alpha_s+m-n}\right| \leq J\pi\delta^{-1}\sum_r |w_r|^2.$$

Since

$$\sum_{\substack{1\leq m,n\leq J\\ m\neq n}}\sum_{r,s}(-1)^{m+n}\frac{|w_r|^2}{m-n}=0$$

we can replace the condition $(r,m)\neq(s,n)$ by $r\neq s$. Further we put $k=m-n$ and get

$$\left|\sum_{|k|\leq J}\sum_{\substack{1\leq n\leq J,\\ 1\leq n+k\leq J}}\sum_{r\neq s}(-1)^k\frac{w_r\overline{w}_s}{\alpha_r-\alpha_s+k}\right| \leq J\pi\delta^{-1}\sum_r|w_r|^2.$$

The n-sum is trivial to evaluate and dividing by J gives the required inequality.

Next we compute that

$$\begin{aligned}\sum_{M<n\leq M+N}\left|\sum_{r\leq R}b_r e(n\alpha_r)\right|^2 &= \sum_{r,s}b_r\bar{b}_s\sum_n e(n(\alpha_r-\alpha_s)) \\ &= N\sum_r|b_r|^2+\sum_{r\neq s}u_r\overline{u}_s\frac{\sin(\pi N(\alpha_r-\alpha_s))}{\sin(\pi(\alpha_r-\alpha_s))},\end{aligned}$$

for $u_r=e((M+N/2+1/2)\alpha_r)b_r$. Using the inequality proved above we get

$$\left|\sum_{r\neq s}u_r\overline{u}_s\frac{\sin(\pi N(\alpha_r-\alpha_s))}{\sin(\pi(\alpha_r-\alpha_s))}\right|\leq\sum_r|u_r|^2\delta^{-1}.$$

Combining this with the diagonal and dualising yields

$$\sum_{r\leq R}\left|S(\alpha_r)^2\right|\leq(N+\delta^{-1})\sum_{M<n\leq M+N}|a_n|^2.$$

Now we observe that for $1\leq a\leq q\leq Q$ with $(a,q)=1$ we have

$$\|\frac{a}{q}-\frac{a'}{q'}\|\geq\frac{1}{qq'}\geq\frac{1}{qQ}\geq\frac{1}{Q^2}.$$

In particular we have $\delta\geq Q^2$ and the result follows straight away. □

[15] This purely analytic inequality is a refinement of Hilbert's inequality by Vaughan and Montgomery and we omit the proof. Full details are given in [1][Satz 4.4.1]

Lemma 40. *Assume $a_n = 0$ if n has a prime factor p with $p \leq Q$. Then, for $T \geq 1$ we have*

$$\sum_{q \leq Q} \log(\frac{Q}{q}) \sum_{\chi \bmod q}^{*} \left| \sum_{y \leq n \leq y+z} a_n \chi(n) \right|^2 \ll (Q^2 + z) \sum_{y \leq n \leq y+z} |a_n|^2 .$$

Proof. We define

$$S(\chi) = \sum_{y \leq n \leq y+z} a_n \chi(n) \text{ and } S(\alpha) = \sum_{y \leq n \leq y+z} a_n e(n\alpha).$$

By the assumption and (25) we get

$$\tau(\overline{\chi}) S(\chi) = \sum_{a=1}^{q} \chi(a) S(\frac{a}{q}).$$

The orthogonality relations imply that

$$\sum_{\chi \bmod q} |\tau(\overline{\chi}) S(\chi)|^2 = \varphi(q) \sum_{\substack{a=1, \\ (a,q)=1}}^{q} \left| S(\frac{a}{q}) \right|^2 .$$

Applying Lemma 39 we get

$$\sum_{q \leq Q} \frac{1}{\varphi(q)} \sum_{\chi \bmod q} |\tau(\overline{\chi}) S(\chi)|^2 \ll (Q^2 + z) \sum_{y \leq n \leq y+z} |a_n|^2 .$$

Note that if f is the conductor of χ, then $fr = q$ and

$$|\tau(\overline{\chi})|^2 = \begin{cases} f & \text{if} (f, r) = 1 \text{ and } r \text{ is } \square\text{-free}, \\ 0 & \text{else.} \end{cases}$$

By assumption we also have $S(\chi) = S(\psi)$, where ψ is the character of conductor f underlying χ. We find

$$\sum_{q \leq Q} \frac{1}{\varphi(q)} \sum_{\chi \bmod q} |\tau(\overline{\chi}) S(\chi)|^2 = \sum_{f \leq Q} \frac{f}{\varphi(f)} \left(\sum_{\substack{r \leq \frac{Q}{f}, \\ (r,f)=1}} \frac{\mu(r)^2}{\varphi(r)} \right) \left(\sum_{\psi \bmod f}^{*} |S(\psi)|^2 \right) .$$

We conclude by the estimate

$$\sum_{\substack{r \leq \frac{Q}{f}, \\ (r,f)=1}} \frac{\mu(r)^2}{\varphi(r)} \geq \frac{\varphi(f)}{f} \log(\frac{Q}{f}).$$

$\square$

Proposition 8. *Assume $a_n = 0$ if n has a prime factor p with $p \leq Q$. Then for $T \geq 1$, we have*

$$\sum_{q \leq Q} \log(\frac{Q}{q}) \sum_{\chi \bmod q}^{*} \int_{-T}^{T} |S(\chi, t)|^2 \, dt \ll \sum_{n} (Q^2 T + n) |a_n|^2 .$$

Proof. The proof proceeds exactly as before. □

Exercise 8. *Show*

$$\sum_{n\leq x} d(n)^2 \ll x\log(x)^3 \text{ and } \sum_{n\leq x} \frac{d(n)^2}{n} \ll \log(x)^4$$

to complete the proof of Corollary 7.

Further prove

$$S_f(x) = \sum_{\substack{r\leq x,\\ (r,f)=1}} \frac{\mu(r)^2}{\varphi(r)} \geq \frac{\varphi(f)}{f}\log(x)$$

to complete the proof of Lemma 40.

Solution. For the first part we write

$$d(n)^2 = \sum_{d|n} g(d),$$

where g is a multiplicative function defined by $g(p^l) = 2l+1$. The estimate is derived as follows:

$$\sum_{n\leq x} d(n) = \sum_{d\leq x} g(d) \sum_{n\leq \frac{x}{d}} 1 \leq x\sum_{d\leq x}\frac{g(d)}{d} \leq x\prod_{p\leq x}\sum_{l=0}^{\infty}\frac{2l+1}{p^l} \ll x\log(x)^3.$$

The second bound follows by partial summation.

For the second part we proceed as follows. Let $q(n)$ be the greatest square-free divisor of n. We first estimate $S_1(x)$ by

$$S_1(x) = \sum_{1\leq n\leq x}\frac{\mu^2(n)}{n}\prod_{p|n}(1-\frac{1}{p})^{-1} = \sum_{1\leq q(n)\leq x}\frac{1}{n} \geq \sum_{1\leq n\leq x}\frac{1}{n} \geq \log(x).$$

On the other hand we have

$$S_1(x) = \sum_{d|f}\frac{\mu^2(d)}{\varphi(d)} S_f(\frac{x}{d}) \leq S_f(x)\sum_{d|f}\frac{\mu^2(d)}{\varphi(d)}.$$

One concludes the proof by observing

$$\sum_{d|f}\frac{\mu^2(d)}{\varphi(d)} = \prod_{p|f}\frac{p}{p-1} = \frac{f}{\varphi(f)}.$$

□

5.2 Zero detecting devices

We put $M_X(s) = \sum_{n \leq X} \mu(n) n^{-s}$. We write

$$\zeta(s) M_X(s) = \sum_{k=1}^{\infty} a(k) k^{-s} \text{ for } a(k) = \sum_{\substack{d|k,\\ d \leq X}} \mu(d).$$

Note that $a(1) = 1$ and $a(k) = 0$ for $2 \leq k \leq X$.

Theorem 16. *Let $0 < c_1 < 1 < c_2$ be fixed constants and pick $1 \ll T^{c_1} \leq X \leq Y \leq T^{c_2}$ and $\rho > \frac{1}{2}$. Then there are subsets*

$$\mathcal{R}_A, \mathcal{R}_B \subset \{\rho \in \mathcal{N} \colon \beta > \rho,\, \log(T)^2 < \gamma < T\}$$

and $N \in [X, Y \log(Y)^2]$ such that

$$N(\sigma, T) \ll \log(T)^6 (1 + \sharp \mathcal{R}_A + \sharp \mathcal{R}_B).$$

Furthermore we have $\rho \neq \rho' \in \mathcal{R}_A \cup \mathcal{R}_B$ then $|\Im(\rho - \rho')| > \log(T)^4$. The zeros in $\mathcal{R}_A$ are of type A and thus satisfy

$$\left| \sum_{N < n \leq 2N} a(n) n^{-\rho} e^{-\frac{n}{Y}} \right| \geq \frac{1}{4 \log(Y)}. \tag{39}$$

The zeros in $\mathcal{R}_B$ are of type B and thus satisfy

$$\left| \int_{\log(T)^2}^{\log(T)^2} \zeta(\frac{1}{2} + i\gamma + it) M_X(\frac{1}{2} + i\gamma + it) Y^{\frac{1}{2} - \beta + it} \Gamma(\frac{1}{2} - \beta + it) dt \right| \geq \frac{1}{2}. \tag{40}$$

Proof. By Mellin inversion we obtain

$$\frac{1}{2\pi i} \int_{(2)} \zeta(s+w) M_X(s+w) Y^s \Gamma(s) ds = e^{-\frac{1}{Y}} + \sum_{k > X} a(k) k^{-w} e^{-\frac{k}{Y}}.$$

We take $w = \rho = \beta + i\gamma$ to be a zero of ζ. Then we can shift the contour of integration to $\Re(s) = \frac{1}{2} - \beta$. Picking up the pole of $\zeta(s + \rho)$ at $1 - \rho$ we find

$$\frac{1}{2\pi i} \int_{(\frac{1}{2} - \beta)} \zeta(s+\rho) M_X(s+\rho) Y^s \Gamma(s) ds = e^{-\frac{1}{Y}} + \sum_{k > X} a(k) k^{-\rho} e^{-\frac{k}{Y}} - M_X(1) Y^{1-\rho} \Gamma(1-\rho).$$

We specialise ρ to a zero lying in the region specified in the statement of the theorem. Since $\gamma \geq \log(T)^2$ we have

$$M_X(1) Y^{1-\rho} \Gamma(1 - \rho) = o(1)$$

by Stirling's approximation. Truncating summation and integration appropriately yields

$$\frac{1}{2\pi} \int_{-\log(T)^2}^{\log(T)^2} \zeta(\frac{1}{2} + i\gamma + it) M_X(\frac{1}{2} + i\gamma + it) Y^{\frac{1}{2} - \beta + it} \Gamma(\frac{1}{2} - \beta + it) dt$$
$$= e^{-\frac{1}{Y}} + \sum_{X < k \leq Y \log(Y)^2} a(k) k^{-\rho} e^{-\frac{k}{Y}} + o(1).$$

Now we take T large enough so that $e^{-\frac{1}{Y}} \geq \frac{9}{10}$ and $o(1) \leq \frac{1}{4}$.[16] Thus at least one of the inequalities

$$\left| \frac{1}{2\pi} \int_{-\log(T)^2}^{\log(T)^2} \zeta(\frac{1}{2} + i\gamma + it) M_X(\frac{1}{2} + i\gamma + it) Y^{\frac{1}{2}-\beta+it} \Gamma(\frac{1}{2} - \beta + it) dt \right| > \frac{1}{2},$$

$$\left| \sum_{X < k \leq Y \log(Y)^2} a(k) k^{-\rho} e^{-\frac{k}{Y}} \right| > \frac{1}{2}$$

must hold.

Zeros that satisfy the first inequality are of type B. Recall that each strip $y \leq \Im(s) \leq y+1$ contains at most $O(\log(y))$ zeros. Thus, for $H < T$, each strip $H \leq \Im(s) \leq H + \log(T)^4$ contains at most $O(\log(T)^5)$ zeros. We decompose all zeros of type B in $\ll \log(T)^5$ disjoint subsets such that in each subset we have $|\Im(\rho - \rho')| > \log(T)^4$. We choose the subset with the most zeros and call it $\mathcal{R}_B$. Thus the total number of zeros of type B is $\ll \log(T)^5 \sharp \mathcal{R}_B$.

The same argument gives a set $\mathcal{R}_A^*$ such that the number of zeros of type A is $\ll \log(T)^5 \sharp \mathcal{R}_A^*$. Now we write $N = 2^j X$ with $j = 0, 1, \dots$ and $N \leq Y \log(Y)^2$. In particular we have $j \leq 2\log(Y) \ll \log(T)$ for sufficiently large Y. For each $\rho \in \mathcal{R}_A^*$ there is some such N such that

$$\left| \sum_{N < k \leq 2N} a(k) k^{-\rho} e^{-\frac{k}{Y}} \right| > \frac{1}{4\log(Y)}.$$

The set $\mathcal{R}_A$ is now constructed by choosing a suitable subset of $\mathcal{R}_A^*$.

We conclude the proof by including the zeros contributing to $N(\sigma, T)$ with $\gamma \leq \log(T)^2$ artificially using the estimate

$$N(\log(T)) \ll \log(T)^3.$$

□

We turn towards a second device, which will need some preparation.

Let $s_v = \sum_{n=1}^{N} b_n z_n^v$. We have the following nice theorem providing lower bounds for such power-sums.

Theorem 17 (Turán's Second Main Theorem)**.** *Suppose* $1 = |z_1| \geq |z_2| \geq \ldots \geq |z_N|$*. Then, for any non-negative integer* M *there is an integer* v*,* $M+1 \leq v \leq M+N$ *such that*

$$|s_v| \geq 2 \left(\frac{N}{8e(M+N)} \right)^N \min_{1 \leq j \leq N} \sum_{n=1}^{j} b_n.$$

Proof. Without loss of generality we assume that the z_n are distinct. We start with the little computation

$$\sum_{v=0}^{N-1} a_v s_{M+1+v} = \sum_{n=1}^{N} b_n z_n^{M+1} \sum_{v=0}^{N-1} a_v z_n^v = \sum_{n=1}^{N} b_n z_n^{M+1} p(z_n)$$

[16] With the obvious abuse of notation!

for $p(z) = \sum_{v=0}^{N-1} a_v z_n^v$. Suppose $p(z_n) = z_n^{-M-1}$ for $1 \le n \le j$ and $p(z_n) = 0$ for $j < n \le N$. This conditions determine p uniquely (for fixed j). In particular we obtain

$$\left|\sum_{n=1}^{j}\right| \le \left(\sum_{v=0}^{N-1} |a_v|\right) \max_{M+1\le n\le M+N} |s_v|. \tag{41}$$

It is now enough to find j such that

$$\sum_{v=0}^{N-1} |a_v| \le \frac{1}{2}\left(\frac{8e(M+N)}{N}\right)^N.$$

We write

$$p(z) = \sum_{k=0}^{N-1} c_k \prod_{n=1}^{k} (z - z_n).$$

Ir $R > 1$, then

$$\begin{aligned} c_k = \frac{1}{2\pi i}\int_{\partial B_R(0)} \frac{p(z)}{\prod_{n=1}^{k+1}(z-z_n)} dz &= \frac{1}{2\pi i}\int_{\partial B_R(0)} \frac{p(z) - z^{-M-1}}{\prod_{n=1}^{k+1}(z-z_n)} dz \\ = \frac{1}{2\pi i}\int_{\partial B_r(0)} \frac{p(z) - z^{-M-1}}{\prod_{n=1}^{k+1}(z-z_n)} dz &= \frac{-1}{2\pi i}\int_{\partial B_r(0)} \frac{z^{-M-1}}{\prod_{n=1}^{k+1}(z-z_n)} dz, \end{aligned}$$

for $r < 1$ such that $|z_j| < r < |z_{j+1}|$. With this at hand we estimate

$$|c_k| \le \frac{1}{r^M \prod_{n=1}^{k+1} |r - |z_n||} \le \frac{1}{r^M \prod_{n=1}^{N} |r - |z_n||}.$$

We put $F(z) = \prod_{n=1}^{N}(z - |z_n|)$ and put $I = [r_0, 1]$ for $0 < r_0 < 1$. Then, according to Exercise 9, there is $r_0 \le r \le 1$ such that

$$F(r) \ge 2(\frac{1-r_0}{4})^N.$$

In particular we get

$$|c_k| \le \frac{1}{2} r_0^{-M} \left(\frac{4}{1-r_0}\right)^N.$$

We now pick $r_0 = \frac{M}{M+N}$ and observe

$$r_0^{-M} = (1 + N/M)^M < e^N.$$

Thus we get $|c_k| \le C$ for

$$C = \frac{1}{2}\left(\frac{4e(M+N)}{N}\right)^N.$$

Now let $P(z) = C\sum_{k=0}^{N-1}(z+1)^k = \sum_{v}^{N-1} A_v z^v$. We complete the proof by observing that

$$\sum_{v=0}^{N-1} |a_v| \le \sum_{v=0}^{N-1} A_v = P(1) = C(2^N - 1) < 2^N C.$$

□

We define

$$S_{x,y}(\chi, v) = \sum_{x<p\leq y} \frac{\chi(p)\log(p)}{p^{1+iv}}.$$

Theorem 18. *Let $w = 1 + iv$. If $L(s,\chi)$ has a zero in $B_r(w)$ for $\log(T)^{-1} \leq r \leq r_0$, then*

$$\int_x^{x^B} |S_{x,y}(\chi, v)| \frac{dy}{y} \ggg x^{-Cr}\log(x)^2$$

for each $x \geq T^A$.

Proof. Recall that by (27) and $q \leq T$ we have

$$\frac{L'}{L}(s,\chi) = \sum_{\rho\in B_1(w)} \frac{1}{s-\rho} + O(\log(T)) \text{ for } |s-w| \leq \frac{1}{2}.$$

By Cauchy's inequality we get

$$\frac{d^k}{k!ds}\frac{L'}{L}(s,\chi) = (-1)^k \sum_{\rho\in B_1(w)} \frac{1}{(s-\rho)^{k+1}} + O(4^k\log(T)) \text{ for } |s-w| \leq \frac{1}{4}.$$

We pick $s = w + r$ and choose a parameter $r \leq \lambda \leq \frac{1}{4}$. Estimating the contribution of the zeros with $|\rho - w| > \lambda$ trivially ends up with

$$\frac{d^k}{k!ds}\frac{L'}{L}(s,\chi) = (-1)^k \sum_{\rho\in B_\lambda(w)} \frac{1}{(s-\rho)^{k+1}} + O(\lambda^{-k}\log(T)).$$

Now the sum has $\leq C_1\lambda\log(T)$ terms and by assumption $\min|\rho - s| \leq 2r$. Thus we can apply Theorem 17 and obtain

$$\sum_{\rho\in B_1(w)}{}^{'} \frac{1}{(s-\rho)^{k+1}} \geq (Dr)^{-k-1}$$

for some $k \in [K, 2K]$ with $K \geq C_1\lambda\log(T)$. We choose $\lambda = C_2Dr$ for C_2 large enough. Note that then $k \geq C_1C_2Dr\log(T) \geq C_1C_2D$, so that $C_2^k \geq C_0Dr\log(T)$ for any C_0. In particular, for $K \geq Er\log(T)$ and $r \leq r_0$ we have

$$\frac{d^k}{k!ds}\frac{L'}{L}(s,\chi) \gg (Dr)^{-k-1}.$$

We rewrite this as

$$\sum_n \frac{\Lambda(n)\chi(n)}{n^w} p_k(r\log(n)) \gg D^{-k}r^{-1},$$

for $p_k(u) = e^{-u}\frac{u^k}{k!}$. We can find constants B_1 and B_2 such that

$$p_k(u) \leq (2D)^{-k} \text{ for } u \leq B_1k \text{ and } p_k(u) \leq (2D)^{-k}e^{-\frac{u}{2}} \text{ for } u \geq B_2k.$$

We will use this to truncate the n-sum above. Indeed let $A = B_1E$, put $K = B_1^{-1}r\log(x)$ such that $K \geq Er\log(T)$. Given $x \geq T^A$ and $k \in [K, 2K]$ we have

$$\sum_{n\leq x} \frac{\Lambda(n)\chi(n)}{n^w} p_k(r\log(n)) \ll (2D)^{-k}\frac{k}{r}$$

and

$$\sum_{n \geq x^B} \frac{\Lambda(n)\chi(n)}{n^w} p_k(r\log(n)) \ll (2D)^{-k}\frac{1}{r}$$

for $B = \frac{2B_2}{B_1}$. We have now set up things such that

$$\sum_{n=x}^{x^B} \frac{\Lambda(n)\chi(n)}{n^w} p_k(r\log(n)) \gg \frac{D^{-k}}{r}.$$

The contribution from p^l with $l \geq 2$ can be ignored so that

$$\sum_{n=x}^{x^B} \frac{\Lambda(n)\chi(n)}{n^w} p_k(r\log(n)) = \int_x^{x^B} p_k(r\log(y))dS(y) \text{ for } S(y) = \sum_{x \leq p \leq y} \log(p)\chi(p)p^{-w}.$$

Partial integration (for Stieltjes integrals) yield

$$\int_x^{x^B} p_k(r\log(y))dS(y) = p_k(rB\log(x))S(x^B) - \int_x^{x^B} S(y)p_k'(r\log(y))r\frac{dy}{y}.$$

Treating the first term on the right side trivially (using the prime number theorem) and applying our lower bound gives

$$\int_x^{x^B} |S(y)|\frac{dy}{y} \gg \frac{D^{-k}}{r^2} \gg x^{-Cr}\log(x)^2.$$

□

Exercise 9 (Chebyshev)**.** *Let F be a monic polynomial of degree d and let $I \subset \mathbb{R}$ be an interval of length L. Then*

$$\max_{z \in I} |F(z)| \geq 2\left(\frac{L}{4}\right)^d.$$

This is the missing ingredient in the proof of Theorem 17.

5.3 Zero density estimates

We start by proving a basic zero density estimate which nicely illustrates the method.

Theorem 19 (Carlson)**.** *For any* $\frac{1}{2} \leq \alpha \leq 1$ *and* $T \geq 2$ *we have*

$$N(\alpha, T) \ll T^{4\alpha(1-\alpha)} \log(T)^{13}.$$

Proof. We start by some preliminary discussion. For s with $\Re(s) \geq \alpha$ and $T \leq \Im(s) \leq 2T$ we define

$$M_X(s) = \sum_{m \leq X} \frac{\mu(m)}{m^s}.$$

Note that in this region we have

$$\zeta(s) = \sum_{n \leq T} n^{-s} + O(T^{-\alpha}).$$

We find that

$$\zeta(s) M_X(s) = \sum_{n \leq TX} a_n n^{-s} + O(T^{-\alpha} X^{1-\alpha} \log(2X)) \tag{42}$$

with

$$a_n = \sum_{\substack{dm=n, \\ m \leq X, \\ d \leq T}} \mu(m).$$

Here we used the trivial bound $M_X(s) \ll X^{1-\alpha} \log(2X)$. We assume $X \leq T$ from now on. Note that $a_n = 0$ for $1 < n \leq X$ and $a_1 = 1$. For $N = 2^l X$ we define

$$D_l(s) = \sum_{N < n \leq 2N} a_n n^{-s} \text{ and } L = \frac{\log(T)}{\log(2)}.$$

We get

$$\zeta(s) M(s) = 1 + \sum_{0 \leq l < L} D_l(s) + O(T^{-\alpha} X^{1-\alpha} \log(2X)).$$

For $M^{1-\alpha} \leq T^{\alpha} \log(T)^{-2}$ the error is smaller than $\frac{1}{2}$. Now let ρ be a zero contributing to the count of $N(T, \alpha)$. Then $\zeta(\rho) M(\rho) = 0$ so that we must have

$$\left| \sum_{0 \leq l < L} D_l(\rho) \right| \geq \frac{1}{2}.$$

By the pigeon hole principle we have $|D_l(\rho)| \geq (2L)^{-1}$ for some $0 \leq l < L$.

Let R_l be the number of zeros ρ contributing to the count $N(T, \alpha)$ with $|D_l(\rho)| \geq (2L)^{-1}$. We say they are detected by D_l. By (36) we get

$$R_l \ll (T + N) N^{1-2\alpha} \log(T)^{11} \ll (T X^{1-2\alpha} + (TX)^{2-2\alpha}) \log(T)^{11}.$$

The result follows by summing over l, choosing $X = T^{2\alpha-1}$ and removing the condition $T \leq \Re(s) \leq 2T$. $\square$

Close to the critical line this estimate can be improved as follows.

Theorem 20 (Ingham)**.** *Let* $\frac{1}{2} < \sigma \leq 1$. *Then*

$$N(\sigma,T) \ll T^{\frac{3(1-\sigma)}{2-\sigma}} \log(T)^{15} \ll T^{3(1-\sigma)} \log(T)^{15}.$$

Proof. We start by estimating $\sharp\mathcal{R}_A$. To do so we put $f_\rho(n) = a(n)n^{-\sigma-i\gamma}e^{i\frac{n}{Y}}$. By squaring (39) and summing it over $\mathcal{R}_A$ we get

$$\sharp\mathcal{R}_A \ll \log(Y)^2 \sum_{\rho\in\mathcal{R}_A} \left| \sum_{N<n\leq 2N} f_\rho(n)n^{i\gamma} \right|^2.$$

By partial summation one has

$$\begin{aligned}\left| \sum_{N<n\leq 2N} f_\rho(n)n^{i\gamma} \right|^2 &\leq 2\left| \sum_{N<n\leq 2N}^{2} f_\rho(n) \right| + 2\left(N^{-1}\int_N^{2N} \left| \sum_{N<n\leq\lambda} f_\rho(n) \right| d\lambda \right)^2 \\ &\leq 4 \sup_{N<\lambda\leq 2N} \left| \sum_{N<n\leq\lambda} f_\rho(n) \right|^2 = 4\left| \sum_{N<n\leq 2N} b(n)n^{-\sigma-i\gamma}e^{i\frac{n}{Y}} \right|^2.\end{aligned}$$

Here we fix λ at which the supremum is attained and define $b(n) = a(n)$ for all $n \leq \lambda$ and $b(n) = 0$ for all $n > \lambda$. We end up with

$$\sharp\mathcal{R}_A \ll \log(Y)^2 \sum_{\rho\in\mathcal{R}_A} \left| \sum_{N<n\leq 2N} b(n)n^{-\sigma-i\gamma}e^{i\frac{n}{Y}} \right|^2.$$

Applying Corollary 6 and estimating $|b(n)|^2 \leq d(n)^2$ we get

$$\sharp\mathcal{R}_A \ll \log(Y)^4 e^{2\frac{N}{Y}}(T+N)N^{-2\sigma} \sum_{n\leq 2N} d(n)^2.$$

Estimating $\sum_{n\leq 2N} d(n)^2 \ll N\log(N)^3$ and taking $X \leq N \leq Y\log(Y)^2$ into account yields

$$\sharp\mathcal{R}_A \ll \log(Y)^9(TX^{1-2\sigma} + Y^{2-2\sigma}).$$

We turn towards an estimate for zeros of type B. Raising (40) to $\frac{4}{3}$ and summing it over $\mathcal{R}_B$ yield

$$\sharp\mathcal{R}_B \leq 4\left| \int_{\log(T)^2}^{\log(T)^2} \zeta(\frac{1}{2}+i\gamma+it)M_X(\frac{1}{2}+i\gamma+it)Y^{\frac{1}{2}-\beta+it}\Gamma(\frac{1}{2}-\beta+it)dt \right|^{\frac{4}{3}}$$

We set $I(\rho) = [\gamma - \log(T)^2, \gamma + \log(T)^2]$ and get by the Hölder inequality that

$$\begin{aligned}&\int_{\log(T)^2}^{\log(T)^2} \zeta(\frac{1}{2}+i\gamma+it)M_X(\frac{1}{2}+i\gamma+it)Y^{\frac{1}{2}-\beta+it}\Gamma(\frac{1}{2}-\beta+it)dt \\ &\leq \left(\int_{I(\rho)} \left|\zeta(\frac{1}{2}+it)\right|^4 dt \right)^{\frac{1}{4}} \left(\int_{I(\rho)} \left|M_X(\frac{1}{2}+it)\right|^2 dt \right)^{\frac{1}{2}} \left(\int_{\infty}^{\infty} \left|\Gamma(\frac{1}{2}-\beta+it)\right|^4 dt \right)^{\frac{1}{4}} Y^{\frac{1}{2}-\sigma}.\end{aligned}$$

It is easy to get (Stirling's approximation) the estimate

$$\left(\int_{\infty}^{\infty}\left|\Gamma(\frac{1}{2}-\beta+it)\right|^4 dt\right) \ll 1.$$

Note that the intervals $I(\rho)$ are disjoint and contained in $[-2T, 2T]$. Thus we obtain

$$\sharp\mathcal{R}_B \ll Y^{\frac{2}{3}-\frac{4\sigma}{3}}\left(\log(T)^2\sum_{\rho\in\mathcal{R}_B}\sup_{t\in I(\rho)}\left|\zeta(\frac{1}{2}+it)\right|^4\right)^{\frac{1}{3}}\left(\int_{-2T}^{2T}\left|M_X(\frac{1}{2}+it)\right|^2 dt\right)^{\frac{2}{3}}.$$

Using Proposition 6 we get

$$\int_{-2T}^{2T}\left|M_X(\frac{1}{2}+it)\right|^2 dt \ll (T+X)\log(X).$$

On the other hand we can use Corollary 7 to obtain

$$\sum_{\rho\in\mathcal{R}_B}\sup_{t\in I(\rho)}\left|\zeta(\frac{1}{2}+it)\right|^4 \ll T\log(T)^7.$$

We end up with

$$\sharp\mathcal{R}_B \ll Y^{\frac{2}{3}-\frac{4\sigma}{3}}(T+T^{\frac{1}{3}}X^{\frac{2}{3}})\log(T)^4.$$

Combining the estimate for $\mathcal{R}_A$ and $\mathcal{R}_B$ and choosing $X=T$ and $Y=T^{\frac{3}{4-2\sigma}}$ yields the desired estimate. □

Theorem 21 (Huxley)**.** *For $\sigma \geq \frac{2}{3}$ we have*

$$N(\sigma,T) \ll T^{\frac{(5\sigma-3)(1-\sigma)}{\sigma^2+\sigma-1}}\log(T)^{25}.$$

Proof. We start by estimating $\mathcal{R}_A$. To do so we pick $1 \leq Q$ and $T_0 \leq T$ and define

$$\mathcal{R}_A(Q) = \{\gamma \in \mathcal{R}_A \colon Q \leq \gamma \leq Q+T_0\}.$$

We estimate elements in this segment using the following trick (Halasz-Montgomery-inequality). Let $V=\mathbb{C}^N$ with standard scalar product. We consider the elements $v=(b(n)n^{-\sigma}e^{-\frac{n}{Y}})_{N<n\leq 2N}$ and $\varphi_\rho=(n^{-i\gamma})_{N<n\leq 2N}$. Here the notation is as in the proof of Ingham's zero density estimate. We observe

$$\langle v,\varphi_\rho\rangle = \sum_{N<n\leq 2N} b(n)n^{-\sigma+i\gamma}e^{-\frac{n}{Y}}$$

and

$$\langle \varphi_{\rho_1},\varphi_{\rho_2}\rangle = \sum_{N<n\leq 2N} n^{i(\gamma_2-\gamma_1)} = Z(\gamma_2-\gamma_1).$$

As earlier we have

$$\sharp\mathcal{R}_A(Q) \leq \log(Y)\sum_{\rho\in\mathcal{R}_A(Q)}\left|\langle v,\varphi_\rho\rangle\right|.$$

Let us write $|\langle v, \varphi_\rho\rangle| = c_\rho \langle v, \varphi_\rho\rangle$. Then we can use Cauchy-Schwarz to see

$$\begin{aligned}\sharp\mathcal{R}_A(Q) \le \log(Y)\left\langle v, \sum_{\rho\in\mathcal{R}_A(Q)} \overline{c}_\rho\varphi_\rho\right\rangle &\le \|v\|\left(\sum_{\rho_1,\rho_2\in\mathcal{R}_A(Q)} c_{\rho_1}\overline{c}_{\rho_2}\langle\varphi_{\rho_1},\varphi_{\rho_2}\rangle\right)^{\frac{1}{2}}\\ &\le \|v\|\left(\sum_{\rho_1,\rho_2\in\mathcal{R}_A(Q)} |\langle\varphi_{\rho_1},\varphi_{\rho_2}\rangle|\right)^{\frac{1}{2}}.\end{aligned}$$

With this at hand we observe that $Z(0) = N$. Further we estimate $\|v\|^2$ to get

$$(\sharp\mathcal{R}_A(Q))^2 \ll N^{1-2\sigma}\log(T)^6\Big(N\sharp\mathcal{R}_A(Q) + \sum_{\substack{\rho_1,\rho_2\in\mathcal{R}_A(Q),\\ \rho_1\neq\rho_2}} |Z(\gamma_2-\gamma_1)|\Big).$$

We use Corollary 4 to estimate

$$\begin{aligned}\sum_{\substack{\rho_1,\rho_2\in\mathcal{R}_A(Q),\\ \rho_1\neq\rho_2}} |Z(\gamma_2-\gamma_1)| &\ll \sum_{\substack{\rho_1,\rho_2\in\mathcal{R}_A(Q),\\ \rho_1\neq\rho_2}} \left(\frac{N}{|\gamma_2-\gamma_1|} - |\gamma_2-\gamma_1|^{\frac{1}{2}}\right)\\ &\ll N\sharp\mathcal{R}_A(Q)\log(T) + T_0^{\frac{1}{2}}(\sharp\mathcal{R}_A(Q))^2.\end{aligned}$$

Thus, we have seen that there is some $C > 0$ such that

$$(\sharp\mathcal{R}_A(Q))^2 \le CN^{2-2\sigma}\log(T)^7\sharp\mathcal{R}_A(Q) + CN^{1-2\sigma}T_0^{\frac{1}{2}}\log(T)^6(\sharp\mathcal{R}_A(Q))^2.$$

We set $T_0 = \min(T, \frac{N^{4\sigma-2}}{16C^2\log(T)^{12}})$ and get

$$\sharp\mathcal{R}_A(Q) \ll N^{2-2\sigma}\log(T)^7.$$

Since this estimate is independent of Q we simply have

$$\sharp\mathcal{R}_A(Q) \ll \frac{T}{T_0}N^{2-2\sigma}\log(T)^7 \ll N^{2-2\sigma}\log(T)^7 + TN^{4-6\sigma}\log(T)^{19}.$$

Using $X \le N \le Y\log(Y)^2$ and $\sigma > \frac{2}{3}$ we end up with

$$\sharp\mathcal{R}_A(Q) \ll Y^{2-2\sigma}\log(T)^9 + TX^{4-6\sigma}\log(T)^{19}.$$

We now turn towards the type B zeros. Using Stirling's approximation we can estimate

$$\begin{aligned}&\left|\int_{\log(T)^2}^{\log(T)^2} \zeta(\frac{1}{2}+i\gamma+it)M_X(\frac{1}{2}+i\gamma+it)Y^{\frac{1}{2}-\beta+it}\Gamma(\frac{1}{2}-\beta+it)dt\right|\\ &\qquad\le CY^{\frac{1}{2}-\beta}\int_{I(\rho)}\left|\zeta(\frac{1}{2}+it)M_X(\frac{1}{2}+it)\right|dt.\end{aligned}$$

Using the lower bound characterising type B zeros and $\beta > \sigma$ we find $\tau = \tau(\rho) \in I(\rho)$ such that

$$\left|\zeta(\frac{1}{2}+it)M_X(\frac{1}{2}+it)\right| \ge \frac{Y^{\sigma-\frac{1}{2}}}{4C\log(T)^2}.$$

We choose a parameter U, which will be specified later, and decompose

$$\mathcal{U} = \{\rho \in \mathcal{R}_B \colon \left|\zeta(\frac{1}{2}+it)\right| > U\} \text{ and } \mathcal{V} = \mathcal{R}_B \setminus \mathcal{U}.$$

We can use Corollary 7 to estimate

$$\sharp\mathcal{U} \leq U^{-4}\sum_{\rho\in\mathcal{U}}\left|\zeta(\frac{1}{2}+i\tau)\right|^4 \ll U^{-4}T\log(T)^9.$$

For $\rho \in \mathcal{V}$ we have $\left|M_X(\frac{1}{2}+it)\right| \geq V$ and $\left|\zeta(\frac{1}{2}+it)\right| \leq U$, for $V = \frac{Y^{\sigma-\frac{1}{2}}}{4C\log(T)^2U}$. Put $T_0 = V^4\log(T)^{-5}$. Then, for $1 \leq Q \leq T$, we define

$$\mathcal{V}(Q) = \{\rho \in \mathcal{V} \colon Q \leq \tau(\rho) \leq Q + T_0\}.$$

We now apply Proposition 5 with $a_n = \mu(n)n^{-\frac{1}{2}+iQ}$ and $t_{rho} = \tau(\rho - Q)$. We obtain

$$\sharp\mathcal{V}(Q) \leq V^{-2}\sum_{\rho\in\mathcal{V}(Q)}\left|M_X(\frac{1}{2}+i\tau)\right| \ll V^{-2}(T_0^{\frac{1}{2}}\sharp\mathcal{V}(Q)+X)\log(T)^2 \ll \frac{X\log(T)^2}{V^2}.$$

Covering $\mathcal{V}$ with $\frac{T}{T_0}+1$ of such intervals we get

$$\mathcal{V} \ll \frac{TX\log(T)^7}{V^6} + \frac{X\log(T)^2}{V^2}.$$

Writing $\log(T) = L$ we get

$$\begin{aligned}\sharp\mathcal{R}_A + \mathcal{R}_B &= \sharp\mathcal{R}_A + \sharp\mathcal{U} + \sharp\mathcal{V}\\ &\ll Y^{2-2\sigma}L^9 + TX^{4-6\sigma}L^{19} + TU^{-4}L^9 + TV^{-6}XL^7 + V^{-2}XL^2.\end{aligned}$$

We equalise the third and fourth term by choosing

$$U = X^{-\frac{1}{10}}Y^{\frac{3}{10}(2\sigma-1)}L^{-1}.$$

Recall that this determines V. The claimed bound now follows by picking

$$X = T^{\frac{2\sigma-1}{2(\sigma^2+\sigma-1)}} \text{ and } T^{\frac{5\sigma-3}{2(\sigma^2+\sigma-1)}}.$$

□

Corollary 8. *For $T \geq 2$ and $\frac{1}{2} \leq \alpha \leq 1$ we have*

$$N(\alpha, T) \ll T^{\frac{12}{5}(1-\sigma)}\log(T)^B,$$

for some large B.

Proof. This follows directly by combining the density estimate of Ingham and Huxley. □

We will also require a so called log-free density estimate. Here one sacrifices a bit of the constant in the exponent of T for the sake of removing the logarithms from the estimate.

Theorem 22 (log-free zero density estimate)**.** *We have*

$$\sum_{q\le T}\sideset{}{^*}\sum_{\chi \bmod q} N_\chi(\sigma,T) \ll T^{c(1-\sigma)}.$$

Proof. Let α be sufficiently small. For $1-\sigma\ge\alpha$ we can estimate

$$\sum_{q\le T}\sideset{}{^*}\sum_{\chi \bmod q} N_\chi(\sigma,T) \ll \sum_{q\le T}\sideset{}{^*}\sum_{\chi \bmod q} T\log(qT) \ll \sum_{q\le T}\varphi(q)T\log(T) \ll T^{3+\epsilon} \ll T^{\frac{3+\epsilon}{\alpha}(1-\sigma)}.$$

Thus by making $c=c(\alpha)$ larger we can assume $1-\sigma\le\alpha$ to be small enough. On the other hand $T^{c(1-\sigma)}$ is constant if $1-\sigma\ll\log(T)^{-1}$. Thus, using standard zero-free regions we can assume $\log(T)^{-1}\ll 1-\sigma\le\alpha$. We can also ignore the finitely many zeros of $\zeta(s)$ in the box $0\le\beta\le1$ and $|\gamma|\le 2$.

Let $w=1+iv$ with $|v|\le T$ and $|v|\ge 2$ if $\chi=\chi_0$. Put $r=2(1-\sigma)$ and assume $\log(T)^{-1}\le r\le r_0$. We apply Theorem 18 with $x=T^{\max(A,5)}$ and get

$$T^{c(1-\sigma)}\log(T)^{-3}\int_x^{x^B}|S_{x,y}(\chi,v)|^2\frac{dy}{y}\gg 1 \text{ for } c=4C\max(A,5),$$

if $L(s,\chi)$ has a zero in $B_r(w)$.

In each disc $B_r(w)$ there are at $\ll r\log(T)^{-1}$ zeros of $L(s,\chi)$ we get

$$N_\chi(\sigma,T)\ll T^{c(1-\sigma)}\log(T)^{-2}\int_x^{x^B}\int_{-T}^{T}|S_{x,y}(\chi,v)|^2\,dv\frac{dy}{y}.$$

Summing over q and χ and handling the y-integral trivially yields

$$\sum_{q\le T}\sideset{}{^*}\sum_{\chi \bmod q} N_\chi(\sigma,T)\ll T^{c(1-\sigma)}\log(T)^{-1}\sum_{q\le T}\sideset{}{^*}\sum_{\chi \bmod q}\int_{-T}^{T}|S_{x,y}(\chi,v)|^2\,dv.$$

for some $y\in[x,x^B]$. Note that since $x\ge T^5$ we can apply Proposition 8. We estimate

$$\begin{aligned}\sum_{q\le T}\sideset{}{^*}\sum_{\chi \bmod q} N_\chi(\sigma,T) &\ll T^{c(1-\sigma)}\log(T)^{-2}\sum_{q\le T}\log(\frac{T^2}{q})\sideset{}{^*}\sum_{\chi \bmod q}\int_{-T}^{T}|S_{x,y}(\chi,v)|^2\,dv\\ &\ll T^{c(1-\sigma)}\log(T)^{-2}\sum_{q\le T^2}\log(\frac{T^2}{q})\sideset{}{^*}\sum_{\chi \bmod q}\int_{-T}^{T}|S_{x,y}(\chi,v)|^2\,dv\\ &\ll T^{c(1-\sigma)}\log(T)^{-2}\sum_{x<p\le y}(T^5+p)\frac{\log(p)^2}{p^2}\ll T^{c(1-\sigma)}.\end{aligned}$$

□

Exercise 10. *Show that almost all zeros of $\zeta(s)$ lie arbitrarily close to the critical line. (Hint: Figure out in which sense 'almost all' is to be understood.)*

6 Part 6: Maier's theorem and other highlights

We finally arrive at the first highlight of this course. We start this section by proving the prime number theorem.

Theorem 23 (Prime number theorem I)**.** *We have*

$$\pi(x) \sim \int_2^x \frac{du}{\log u}.$$

Proof. Applying Perron's formula we have

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

$$= x + I_1 + I_2 + I_3.$$

Here we shifted the contour across the simple pole $s = 1$ with residue 1 and obtain the error integrals

$$I_j = \frac{1}{2\pi i}\int_{\gamma_j} \frac{\zeta'}{\zeta}(s)x^s\frac{ds}{s},$$

for $\gamma_1 = [c+iT, c'+iT]$, $\gamma_2 = [c'+iT, c'-iT]$ and $\gamma_3 = [c'-iT, c-iT]$. According to Lemma 21 we can choose $c' = 1 - \lambda$ and $c = 1 + \lambda$ for $\lambda = \delta\log(T)^{-9}$. Estimating trivially yields

$$I_1 = \frac{1}{2\pi i}\int_{c+iT}^{c'+iT} \frac{\zeta'}{\zeta}(s)x^s\frac{ds}{s} \ll x^{1+\lambda}T^{-1}$$

and the same estimate holds for I_3. Similarly one gets

$$I_2 = \frac{1}{2\pi i}\int_{-T}^{T} \frac{\zeta'}{\zeta}(c'+it)x^{c'+it}\frac{dt}{c'+it} \ll x^{1-\lambda}\lambda^{-1}\int_{-T}^{T}\frac{dt}{1+|t|} \ll x^{1-\lambda}\log(T)^{10}.$$

By putting things together we get

$$\psi(x) - x \ll \frac{x^{1+\lambda}}{T\lambda} + \log(x) + \frac{x\log(x)}{T} + x^{1-\lambda}\log(T)^{10}.$$

Choosing a suitable T we get $\psi(x) \sim x$ and the theorem follows from partial summation. $\square$

For many applications one would like to get a good handle on the error term of this asymptotic. In other words, one needs quantitative estimates for

$$E_\pi(x) = \pi(x) - \int_2^x \frac{du}{\log(u)} \text{ or } E_\psi(x) = \psi(x) - x.$$

Note that the proof of the prime number theorem given above yields

$$E_\pi(x) \ll x\exp(-C(\log(x)^{\frac{1}{10}}))$$

for some $C > 0$.

Theorem 24 (Prime number Theorem II)**.** *We have*

$$E(x) \ll x \exp(-c\sqrt{\log(x)}),$$

for some positive constant c.

Proof. We start by applying Lemma 26 to get

$$\sum_{|\Im(\rho)|\leq T} |\rho|^{-1} \ll \sum_{m\leq T} \frac{\log(m)}{m} \ll \log(T)^2.$$

Further using Theorem 9 we get

$$\sum_{|\Im(\rho)|\leq T} x^{\Re(\rho)} |\rho|^{-1} \ll x^{1-C\log(T)^{-1}} \log(T)^2.$$

With this at hand the (truncated) explicit formula (Theorem 7) yields the estimate

$$E_\psi(x) = \psi(x) - x \ll x^{1-C\log(T)^{-1}} \log(T)^2 + \frac{x}{T}\log(x)^2.$$

Choosing $T = \exp(\sqrt{\log(x)})$ yields

$$E_\psi(x) \ll x \exp(-c\sqrt{\log(x)}).$$

The result for $E_\pi(x)$ follows by partial summation. □

Remark 4. *Using the extended zero-free region due to Vinogradov-Korobov one can establish the bound*

$$E_\psi(x) \ll x \exp(-c\log(x)^{\frac{3}{5}} \log\log(x)^{-\frac{1}{5}}).$$

The proof is the same as the proof given above, only that one replaces the standard zero free region by the extended one. We omit the details.

In particular these error estimates allow us to deduce

$$\psi(x+y) - \psi(x) \sim y$$

for $x \to \infty$ as long as $y = y(x)$ stays a bit larger as $E_\psi(x)$. In other words, there are always primes in intervals $(x, x+y]$ as long the length of the interval stays larger than $E_\psi(x)$. Unfortunately none of our error bounds is good enough to answer the question if there are primes in intervals of the form $(x, x+x^\theta]$ for $\theta < 1$. This makes the following result very satisfying.

Theorem 25 (Hoheisel)**.** *Suppose one has a zero-free region for $\zeta(s)$ of the type*

$$|t| \leq T,\, \sigma \geq 1 - B\frac{\log\log(T)}{\log(T)}.$$

Further we assume that a zero-density estimate of the form

$$N(\alpha, T) \ll T^{c(1-\alpha)} \log(T)^A \tag{43}$$

for all $\frac{1}{12} \leq \alpha \leq 1$ is available for some $A \geq 1$ and $c \geq 2$. We put $\theta = 1 - (c + \frac{A+1}{B})^{-1}$. Then

$$\psi(x+y) - \psi(x) = y + O(\frac{y}{\log(x)})$$

for $x^\theta \log(x)^3 \leq y \leq x$.

Proof. Using the (truncated) explicit formula (Theorem 7) we get

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

with $T = x^{1-\theta}$. We deduce

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

Due to the mean value theorem the following estimate suffices

$$\begin{aligned}
\sum_{|\gamma|\leq T} x^{\beta-1} &\leq 2\int_{\frac{1}{2}}^{1} x^{\alpha-1} dN(\alpha,T) \\
&\ll x^{-\frac{1}{2}} N(\frac{1}{2},T) + \log(x)\int_{\frac{1}{2}}^{1} x^{\alpha-1} N(\alpha,T)d\alpha \\
&\ll x^{-\frac{1}{2}} T\log T + \log(x)\log(T)^A \int_{\eta}^{\frac{1}{2}} \left(\frac{T^c}{x}\right)^{\alpha} d\alpha \\
&\ll x^{-\frac{1}{2}} T\log T + \left(\frac{T^c}{x}\right)^{\eta} \log(T)^A.
\end{aligned}$$

Here $\eta = B\log\log(T)\log(T)^{-1}$ is allowed due to the zero-free region. The parameters are chosen such that

$$\left(\frac{T^c}{x}\right)^{\eta} = \log(T)^{-A-1}.$$

We conclude that

$$\sum_{|\gamma|\leq T} \frac{(x+y)^\rho - x^\rho}{\rho y} \ll \sum_{|\gamma|\leq T} x^{\beta-1} \ll \log(x)^{-1}.$$

This completes the proof. $\square$

Remark 5. *With the results that were proved in this lecture we get some $\theta < 1$ which we can not further specify, since we did not work out the exponent of $\log(T)$ in Huxley's zero density estimate (Corollary 8) and we also did not work out the small constant appearing in the Hardy-Littlewood zero-free region. However, using the zero free region of Vinogradov-Korobov one can take B as large as needed, so that the constant A is irrelevant. Thus one obtains the existence of primes in intervals $(x, x+x^{\frac{7}{12}+\epsilon}]$, which is quite amazing.*

Of course one can still enquire about the distribution statistics of primes in even shorter intervals. The remaining lectures work towards proving an exciting result showing that for intervals of length $\log(x)^A$ strange things can happen.

We conclude this section by working out some results on primes in arithmetic progressions.

Theorem 26 (Siegel-Walfisz)**.** *Let $A > 0$ be a constant and let $q \leq \log(x)^A$ and $(a,q) = 1$. Then there is $C = C(A)$ with*

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

Proof. Let χ be a primitive Dirichlet character modulo q. We write

$$b(\chi) = \begin{cases} B(\chi) & \text{if } \chi \text{ is even}, \\ \frac{L'}{L}(0,\chi) & \text{if } \chi \text{ is odd}. \end{cases}$$

For $2 \leq T \leq x$ the explicit formula (Theorem 8) reads

$$\psi_\chi(x) = -\sum_{\substack{\rho \in \mathcal{N}(\chi), \\ |\Im(\rho)| < T}} \frac{x^\rho}{\rho} - b(\chi) + O(\frac{x}{T} \log(qx)^2).$$

We apply (27) to get

$$b(\chi) = -\sum_{|\Im(\rho)|<1} \frac{1}{\rho} + O(\log(q)).$$

Here we observe that $B(\chi) = \lim_{s\to 0} \left(\frac{L'}{L}(s,\chi) - \frac{1}{s}\right)$ for even χ.

According to Theorem 10 there is at most one real zero ρ with $|\Im(\rho)| < 1$ and $\Re(\rho) > 1 - c\log(q)^{-1}$. This zero will be denoted by β and must be real. Further $1-\beta$ is also a zero. Note that $\beta^{-1} \ll 1$, so that

$$\begin{aligned} b(\chi) &= \frac{1}{1-\beta} + \sum_{\substack{|\Im(\rho)|<1, \\ \rho \neq \beta, \beta-1}} \frac{1}{\rho} + O(\log(q)) \\ &= \frac{1}{1-\beta} + O(\log(q)^2). \end{aligned}$$

Further we can estimate

$$\frac{x^{1-\beta} - 1}{1-\beta} = x^\xi \log(x) \ll x^{\frac{1}{4}}$$

by the mean value theorem. To summarise this we find that

$$\psi_\chi(x) = -\delta_{\chi \text{ ex.}} \frac{x^\beta}{\beta} - \sum_{\substack{\rho \neq \beta, \beta-1, \\ |\Im(\rho)|<T}} \frac{x^\rho}{\rho} + O(\frac{x}{T}\log(qx)^2 + x^{\frac{1}{4}}). \tag{44}$$

An almost trivial argument shows that this also holds for non-primitive χ.

We now want to translate this into bounds for primes in arithmetic progressions. In this direction we observe that

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

We compute that

$$|\psi_{\chi_0}(x) - \psi(x)| \leq \sum_{\substack{n \leq x, \\ (n,q)>1}} \Lambda(n) \ll \log(x)\log(q).$$

By the prime number theorem we get that

$$\psi(x;a,q) = \frac{x}{\varphi(q)} + \frac{1}{\varphi(q)} \sum_{\chi \neq \chi_0} \overline{\chi}(a)\psi_\chi(x) + O(\frac{x}{\varphi(q)} \exp(-c\sqrt{\log(x)}) + \log(qx)^2).$$

By Landau's theorem there is at most one exceptional character modulo q (Theorem 11), which we denote by χ_1. Let β be the corresponding exceptional zero. We obtain

$$\sum_{\chi\neq\chi_0}\overline{\chi}(a)\psi_\chi(x)=\overline{\chi_1}(a)\frac{x^\beta}{\beta}+\sum_{\chi\neq\chi_0}\sum_{\substack{\rho\in\mathcal{N}(\chi),\\ |\Im(\rho)|<T,\\ \rho\neq\beta,\beta-1}}\overline{\chi}(a)\frac{x^\rho}{\rho}+O(\varphi(q)(xT^{-1}\log(qx)^2+x^{\frac{1}{4}}).$$

The remaining sum over the zeros can be estimated with Lemma 26 and Theorem 10 as in the proof of the prime number theorem above. All together we find

$$\psi(x;a,q)=\frac{x}{\varphi(q)}-\frac{\overline{\chi_1}(a)x^\beta}{\varphi(q)\beta}+O\bigg(\frac{1}{\varphi(q)}\exp(-c\sqrt{\log(x)})+xT^{-1}\log(qx)^2$$
$$+x^{\frac{1}{4}}+x\log(qx)^2\exp(-c_1\frac{\log(x)}{\log(qT)})\bigg).$$

If $q\leq\exp(C\sqrt{\log(x)})$, then we choose $T=\exp(C\sqrt{\log(x)})$ and get

$$\psi(x;a,q)=\frac{x}{\varphi(q)}-\frac{\overline{\chi_1}(a)x^\beta}{\varphi(q)\beta}+O(xe^{-C_1\sqrt{\log(x)}}).\tag{45}$$

Using Siegel's theorem (Theorem 12) we get

$$x^\beta\leq x\exp(-C_1(\epsilon)\log(x)q^{-\epsilon}).$$

Thus, for $q\leq\log(x)^A$ the term $\frac{\overline{\chi_1}(a)x^\beta}{\varphi(q)\beta}$ can be absorbed into the error term by after adjusting the constant C_1 in the exponential appropriately. □

Theorem 27 (Page)**.** *Let $C>0$ and $x\geq 10$ be given. Then there is $C_1=C_1(C)$ and $q_1=q_1(x)$ such that for all q with $q_1\nmid q$ and $q\leq\exp(C\sqrt{\log(x)})$ we have*

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

as long as $(a,q)=1$.

Proof. We start from (45) and only need to deal with the contribution of β. To do so we set $Q=\exp(C\sqrt{\log(x)})$. By Corollary 3 there is at most one exceptional modulus $q_1\leq Q$ with exceptional character χ_1 and exceptional zero β_1. All other L-series with exceptional zero belong to characters modulo q with $q_1\mid q$. For all other q we can estimate

$$\beta\leq 1-\frac{c}{\log(Q)}=1-\frac{c}{C\sqrt{\log(x)}}.$$

Inserting this estimate above completes the proof. □

Theorem 28 (Gallagher)**.** *Let q be a good modulus. Then, for $(a,q)=1$, $x\geq q^D$ and $\frac{x}{2}\leq h\leq x$ with $\log(q)\geq D\geq D_0$, we have*[17]

$$\pi(x+h,q,a)-\pi(x,q,a)=\frac{li(x+h)-li(x)}{\varphi(x)}\left(1+O(e^{-cD}+e^{-\sqrt{\log(x)}})\right).$$

[17]The constant D_0 as well as the implicit constant only depend upon the constant in the definition of the term good modulus.

Proof. By (44) we get

$$\psi_\chi(x) = -\sum_{\substack{\rho\in\mathcal{N}(\chi),\\ |\Im(\rho)|<T}} \frac{x^\rho}{\rho} + O(\frac{x\log(x)^2}{T}).$$

for a non-exceptional character χ with conductor $1 < k \leq Q$. We compute

$$\frac{(x+h)^\rho}{\rho} - \frac{x^\rho}{\rho} = \int_x^{x+h} y^{\rho-1}dy \ll hx^{\beta-1}.$$

Assume $\frac{x}{Q} \leq h \leq x$ and $\exp(\log(x)^{\frac{1}{2}}) \leq Q \leq x^b$ for the moment. Thus $\log(x) \leq \log(Q)^2$ and $x \leq hQ$. After treating the contribution of p^l with $l > 1$ trivially we get

$$\sum_{x<p\leq x+h} \chi(p)\log(p) \ll h\left(\sum x^{\beta-1} + \frac{Q^2}{T}\right).$$

Summing this over $1 < k \leq q$ and all primitive characters modulo k different from the possible exceptional character χ_e yields

$$\sum_{2\leq k\leq Q}\ \sideset{}{^*}\sum_{\substack{\chi \bmod k,\\ \chi\neq\chi_e}} \left|\sum_{x<p\leq x+h} \chi(p)\log(p)\right| \ll h\left(\sum_{2\leq k\leq Q}\ \sideset{}{^*}\sum_{\substack{\chi \bmod k,\\ \chi\neq\chi_e}} \sum x^{\beta-1} + \frac{Q^4}{T}\right).$$

Write $N(\sigma,T,Q)$ for $\sum_{1\leq k\leq Q}\sum_{\chi \bmod k}^{*} N(\sigma,T)$ and $N'(\sigma,T,Q) = N(\sigma,T,Q) - N_{\chi_e}(\sigma,T)$. Then we have

$$\begin{aligned}\sum_{2\leq k\leq Q}\ \sideset{}{^*}\sum_{\chi \bmod k} \sum x^{\beta-1} &\leq -\int_0^1 x^{\sigma-1}dN'(\sigma,T,Q)\\ &= \int_0^1 x^{\sigma-1}\log(x)N'(\sigma,T,Q)d\sigma + x^{-1}N(0,T,Q).\end{aligned}$$

We put $\eta(T) = -\frac{c_1}{\log(T)}$ and assume $T^c \leq x^{\frac{1}{2}}$ and apply Theorem 22. This yields

$$\begin{aligned}\int_0^1 x^{\sigma-1}\log(x)N'(\sigma,T,Q)d\sigma + x^{-1}N(0,T,Q) &\ll \int_0^{1-\eta(T)} x^{\sigma-1}\log(x)N(\sigma,T,Q)d\sigma + x^{-\frac{1}{2}}\\ &\ll \int_0^{1-\eta(T)} x^{\frac{1}{2}(\sigma-1)}\log(x)d\sigma + x^{-\frac{1}{2}}\\ &\ll x^{-\frac{1}{2}\eta(T)}.\end{aligned}$$

Choosing $T = Q^5$ yields

$$\sum_{2\leq k\leq Q}\ \sideset{}{^*}\sum_{\substack{\chi \bmod k,\\ \chi\neq\chi_e}} \left|\sum_{x<p\leq x+h} \chi(p)\log(p)\right| \ll h\left(\exp(-c'\frac{\log(x)}{\log(Q)}) + Q^{-1}\right) \ll h\exp(-c''\frac{\log(x)}{\log(Q)}),$$

since $\log(x) \leq \log(Q)^2$.

We write

$$\sum_{\substack{x<p\le x+h,\\ p\equiv a \bmod q}} \log(p) = \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \chi(a)^{-1} \sum_{x<p\le x+h} \chi(p)\log(p)$$
$$= \frac{\psi(x+h)-\psi(x)}{\varphi(q)} + \sum_{1\neq f|q} \sideset{}{^*}\sum_{\chi \bmod f} \sum_{x<p\le x+h} \chi(p)\log(p).$$

The result is now easy to derive. We uses a suitable version of the prime number theorem to deal with $\frac{\psi(x+h)-\psi(x)}{\varphi(q)}$. Since $q \le Q = x^{\frac{1}{D}}$ is a good modulus we can estimate

$$\left|\sum_{1\neq f|q} \sideset{}{^*}\sum_{\chi \bmod f} \chi(a)^{-1} \sum_{x<p\le x+h} \chi(p)\log(p)\right| \le \sum_{2\le k\le Q} \sideset{}{^*}\sum_{\substack{\chi \bmod k,\\ \chi\neq\chi_e}} \left|\sum_{x<p\le x+h} \chi(p)\log(p)\right|$$
$$\ll \exp(-c''\frac{\log(x)}{\log(Q)}) = \exp(-c''D).$$

The statement follows by partial summation. □

Remark 6. *The previous theorem can be modified to work for arbitrary moduli by using the exceptional-zero-repulsion phenomenon one can even improve the error in the presence of an exceptional zero.*

Theorem 29. *Let $\Phi(x) = \log(x)^{\lambda_0}$ with $\lambda_0 > 1$. Then*

$$\limsup_{x\to\infty} \frac{\pi(x+\Phi(x))-\pi(x)}{\phi(x)/\log(x)} > 1 \text{ and } \liminf_{x\to\infty} \frac{\pi(x+\Phi(x))-\pi(x)}{\phi(x)/\log(x)} < 1.$$

For the range $1 < \lambda_0 < e^\gamma$ we even have

$$\limsup_{x\to\infty} \frac{\pi(x+\Phi(x))-\pi(x)}{\phi(x)/\log(x)} \ge \frac{e^\gamma}{\lambda_0},$$

where γ is Euler's constant.

Proof. We fix $D \ge D_0$ depending only on $\epsilon > 0$ appearing later. Further we consider $z \to \infty$ through the set

$$\mathcal{G}_D = \{z \ge e^{cD} \colon P(z) \text{ is a good mdulus }\}.$$

Finally we choose $U = U(z) \le P(z)$.

We define the matrix $\mathfrak{M} = (a_{rs})$ with $a_{rs} = s + rP(z)$ for $1 \le s \le U$ and $P(z)^{D-1} < r \le 2P(z)^{D-1}$. Let $\pi(\mathfrak{M})$ be the number of primes contained in the matrix. Note that only columns with $(s, P(z)) = 1$ can contain primes. Such columns are called admissible. Note that the rows of $\mathfrak{M}$ are intervals of length U, while the columns are arithmetic progressions with modulus $P(z)$.

The number of admissible columns is $\Phi(U, z)$. In each admissible column we apply Theorem 28 with $q = P(z)$, $x = P(z)^D + s$ and $h = P(z)^D$. This yields the estimate

$$\pi(\mathfrak{M}) = \Phi(U,z) \cdot \frac{P(z)^D}{\varphi(P(z))\log(P(z)^D)} \left(1 + O\left(e^{-cD}\right)\right)$$
$$= U \cdot \frac{\Phi(U,z)}{UW(z)} \cdot \frac{P(z)^{D-1}}{\log(P(z)^D)} \left(1 + O\left(e^{-cD}\right)\right). \tag{46}$$

Note that by construction all the requirements of Theorem 28 are satisfied and $e^{-\sqrt{\log(x)}} \gg e^{cD}$.

By Lemma 13 we can pick $\lambda_1 > \lambda_0$ such that $w(\lambda_1) > e^{-\gamma}$. We put $U = z^{\lambda_1}$ and observe that by (46) and Lemma 14 there must be a row $\mathfrak{R}$ of $\mathfrak{M}$ with at least

$$e^{\gamma} w(\lambda_1) \cdot \frac{U}{\log(P(z)^D)} \left(1 + O\left(e^{-cD}\right)\right)$$

primes. Further put $l_0 = \log(P(z)^d)^{\lambda_0}$ and $K_0 = \lfloor \frac{U}{l_0} \rfloor + 1$. We divide the interval given by the row $\mathfrak{R}$ in K_0 subintervals of length $l_0 + o(l_0)$. At least one of these subintervals, say $(a, b]$, contains at least

$$e^{\gamma} w(\lambda_1) \cdot \frac{U}{K_0 \log(P(z)^D)} \left(1 + O\left(e^{-cD}\right)\right).$$

We put $x = a$, so that $(a, b] \subset (x, x + \Phi(x)]$. The latter integral is now constructed such that it contains at least

$$e^{\gamma} w(\lambda_1) \cdot \frac{\Phi(x)}{\log(x)} \left(1 + O\left(e^{-cD}\right)\right)$$

primes. This gives the first part as well as the final part of the theorem.

The lim inf-part is very similar. Here we choose $\lambda_2 > \lambda_0$ such that $w(\lambda_2) < e^{-\gamma}$ and put $U = z^{\lambda_2}$. It is left as an exercise to adapt the rest of the proof. $\square$

Exercise 11. *The* ***Riemann hypothesis*** *is the deep conjecture that all non-trivial zeros ρ of $\zeta(s)$ satisfy $\Re(\rho) = \frac{1}{2}$. What can be said about the error term in the prime number theory under assumption of the Riemann hypothesis? What can you say about primes in short intervals under the assumption of the Riemann hypothesis?*

References

[1] Jörg Brüdern. *Einführung in die analytische Zahlentheorie.* Springer-Verlag, 2013.

[2] Rolf Busam and Eberhard Freitag. *Funktionentheorie 1.* Springer-Verlag Berlin Heidelberg, 2006.

[3] H. Daboussi. On the prime number theorem for arithmetic progressions. *J. Number Theory*, 31(3):243–254, 1989.

[4] Hédi Daboussi. Sur le théorème des nombres premiers. *C. R. Acad. Sci. Paris Sér. I Math.*, 298(8):161–164, 1984.

[5] N. G. De Bruijn. On the number of uncancelled elements in the sieve of Eratosthenes. *Nederl. Akad. Wetensch., Proc.*, 53:803–812 = Indagationes Math. 12, 247–256 (1950), 1950.

[6] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*, volume 53. American Mathematical Soc., 2004.

[7] E. Landau. Über die äquivalenz zweir Hauptsatze der analytischen Zahlentheorie. *Wien. Sitzungsberichte*, (120):973–987, 1911.

[8] Helmut Maier et al. Primes in short intervals. *The Michigan Mathematical Journal*, 32(2):221–225, 1985.

[9] Hugh L Montgomery. *Ten lectures on the interface between analytic number theory and harmonic analysis*. Number 84. American Mathematical Soc., 1994.

[10] Edward Charles Titchmarsh, Edward Charles Titchmarsh Titchmarsh, DR Heath-Brown, et al. *The theory of the Riemann zeta-function*. Oxford University Press, 1986.