Elective Subject Mathematics: Number Theory

Lecture notes

Graz University of Technology

Summer term 2020

Lecturer: Kostadinka Lapkova

Initial typesetting by Georg Zechner

January 21, 2021

Contents

Literature:

[1] J. Brüdern, Einführung in die analytische Zahlentheorie, Springer 1995

[2] E. Hlawka, J. Schoissengeier, R. Taschner, Geometric and Analytic Number Theory, Springer 1991

[3] H.L. Montgomery, R.C. Vaughan, Multiplicative number theory I. Classical Theory, Cambridge 2007

[4] M. Ram Murty, Problems in Analytic Number Theory, Springer 2008

[5] Дойчин Толев, Лекции по елементарна и аналитична теория на числата, Университетско издателство "Св.Климент Охридски", 2016 (in Bulgarian); D. Tolev, Lectures in elementary and analytic number theory, Sofia 2016

[6] R.C. Vaughan, The Hardy-Littlewood method, Cambridge 1981

[7] D. Zagier, Newman's short proof of the prime number theorem, Amer. Math. Monthly 104 (1997), no. 8, 705–708

Chapter 1

Introduction

This is an introductory course in analytic number theory for graduate students. The first main focus of the course is presenting the Newman's proof of the Prime number theorem. The material is based mostly on the chapters $4, 5, 6$ from [2]. The second main topic is application of the circle method in the ternary Goldbach's problem. The used literature for the circle method is mostly [1], also [5] and [6].

Some of the exercise problems are from [4]; few results are borrowed from [3], which is a fundamental reference for the subject.

Let us present two proofs of one classical result. The second proof of Euler presents one of the first instances of application of analytic methods in number theory.

Theorem 1.0.1. *There are infinitely many prime numbers.*

Euclid's proof. Assume there are only finitely many primes. Denote them by $p_1, ..., p_k$ and define $N = p_1 \cdots p_k + 1$. Then there does not exist a p_i for $i = 1, ..., k$ such that p_i is a divisor of N. Therefore N is either prime itself or contains a prime factor differing from $p_1, ...p_k$. This contradicts the assumption and therefore the theorem holds. $\square$

Euler's proof (1737). Consider the harmonic series $\sum_{n=1}^{\infty} \frac{1}{n^s}$ for $s > 1$. We see that for $s = 1$

$$\sum_{n=1}^{2^k} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{2^k} > 1 + \frac{k}{2}$$

by taking 2^i elements together and estimating them with the smallest one, e.g. $\frac{1}{3} + \frac{1}{4} > 2\frac{1}{4} = \frac{1}{2}$. But now we have

$$\sum_{n=1}^{\infty} \frac{1}{n} = \lim_{k\to\infty} \sum_{n=1}^{2^k} \frac{1}{n} > \lim_{k\to\infty} \left(1 + \frac{k}{2}\right) \to \infty.$$

On the other hand we know that by the main theorem of arithmetic every integer has a unique representation as a factor of prime powers and then

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \left(1 + \frac{1}{2^s} + \frac{1}{2^{2s}} + \cdots\right)\left(1 + \frac{1}{3^s} + \frac{1}{3^{2s}} + \cdots\right)\cdots$$
$$= \prod_{p\in\mathbb{P}} \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots\right) = \prod_{p\in\mathbb{P}} \frac{1}{1 - \frac{1}{p^s}}$$

Now, if there were only finitely many primes, then when $s \to 1$ the product $\prod_{p\in\mathbb{P}} \left(1 - \frac{1}{p^s}\right)^{-1}$ converges to, say, c, but then we should also have $\sum_{n\geq 1} \frac{1}{n} = c$, which is a contradiction. □

Chapter 2

Arithmetic functions

2.1 Basic properties

Definition 2.1.1.

*i) A function $f : \mathbb{N} \to \mathbb{C}$ is said to be **arithmetic** (or number theoretic).*

*ii) A function f is said to be **multiplicative** if f is not the zero function and for coprime $m, n \in \mathbb{N}$ we have $f(mn) = f(m)f(n)$.*

Some examples for multiplicative functions are:

- The Möbius-function $\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if a square is a divisor of n} \\ (-1)^k & \text{if } n = p_1 \cdots p_k \end{cases}$

- The Euler φ function $\varphi(n) = \sum\limits_{\substack{1 \leq k \leq n \\ \gcd(n,k)=1}} 1$

- The number of divisors function $\tau(n) = \sum_{d|n} 1$

- The sum of divisors function $\sigma(n) = \sum_{d|n} d$

- The prime counting function $\pi(n) = \sum_{p \leq n} 1$

Definition 2.1.2. *The **Sum function** of an arithmetic function f is defined to be $S_f(n) = \sum_{d|n} f(d)$.*

Lemma 2.1.1. *Let f be a multiplicative function and $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$. Then*

a) $S_f(n) = \sum_{d|n} f(d) = (1 + f(p_1) + \cdots + f(p_1^{\alpha_1})) \cdots (1 + f(p_k) + \cdots + f(p_k^{\alpha_k}))$

b) $S_{\mu f}(n) = \sum_{d|n} \mu(d)f(d) = \begin{cases} (1 - f(p_1)) \cdots (1 - f(p_k)) & \textit{if } n > 1 \\ 1 & \textit{if } n = 1 \end{cases}$

Proof.

a)

$$RHS = \prod_{\substack{p_i \\ 1\le i\le k}} \sum_{0\le\beta_i\le\alpha_i} f(p_i^{\beta_i}) = \sum_{0\le\beta_i\le\alpha_i} \prod_{\substack{p_i \\ 1\le i\le k}} f(p_i^{\beta_i}) = \sum_{0\le\beta_i\le\alpha_i} f(p_1^{\beta_1})\cdots f(p_k^{\beta_k}) =$$
$$= \sum_{0\le\beta_i\le\alpha_i} f(p_1^{\beta 1}\cdots p_k^{\beta_k}) = \sum_{d|n} f(d) = S_f(n)$$

b) Follows by a) □

Corollary. *Let again $n = p_1^{\alpha_1}\cdots p_k^{\alpha_k}$. Then*

i) $\sigma(n) = \sum_{d|n} d = (1+p_1+\cdots+p_1^{\alpha_1})\cdots(1+p_k+\cdots+p_k^{\alpha_k}) = \frac{p_1^{\alpha_1+1}-1}{p_1-1}\cdots\frac{p_k^{\alpha_k+1}-1}{p_k-1}$ *is multiplicative. To see this we use* $f(d)=d$.

ii) $\tau(n) = \sum_{d|n} 1 = (1+\cdots+1)\cdots(1+\cdots+1) = (\alpha_1+1)\cdots(\alpha_k+1)$ *is multiplicative. To see this we use* $f(d)=1$.

iii) $S_\mu(n) = \sum_{d|n}\mu(d) = \begin{cases} 0 & \text{if } n>1 \\ 1 & \text{if } n=1 \end{cases}$ *,this follows by Lemma 2.1.1 b).*

Lemma 2.1.2 (Vinogradov's lemma)**.** *Let S be a finite set, G a commutative group written additively f and g both functions from S to $\mathbb{N}$ or both from S to G. Then we have*

$$\sum_{\substack{s\in S \\ f(s)=1}} g(s) = \sum_{m=1}^{\infty}\mu(m)\sum_{\substack{s\in S \\ m|f(s)}} g(s).$$

Proof.

$$LHS = \sum_{\substack{s\in S \\ f(s)=1}} g(s) \overset{iii)}{=} \sum_{s\in S} g(s)\sum_{d|f(s)}\mu(d) = \sum_{m=1}^{\infty}\mu(m)\sum_{\substack{s\in S \\ m|f(s)}} g(s)$$

□

Theorem 2.1.1 (Möbius inversion formula)**.** *Let g be an arithmetic function, then it can be expressed in terms of its sum function.*

$$g(n) = \sum_{d|n}\mu(d)S_g\left(\frac{n}{d}\right) = \sum_{d|n}\mu(d)\sum_{l|\frac{n}{d}} g(l).$$

Proof. Define $f(s) = \frac{n}{s}$ thus we have $f(s)=1$ if and only if $n=s$ and $S = \{k\in\mathbb{N}|\frac{n}{k}\in\mathbb{N}\}$. Then we get

$$g(n) = \sum_{\substack{s\in S \\ f(s)=1}} g(s) = \sum_{m=1}^{\infty}\mu(m)\sum_{\substack{s\in S \\ m|f(s)}} g(s) = \sum_{m|n}\mu(m)\sum_{\substack{s\in S \\ s|\frac{n}{m}}} g(s)$$

□

Definition 2.1.3. *The* ***Dirichlet product*** *(or convolution) of two functions f, g is defined to be $f * g = \sum_{d|n} f(d)g(\frac{n}{d})$.*

Corollary. *Let again $n = p_1^{\alpha_1}\cdots p_k^{\alpha_k}$. Then* *i)* $\varphi(n) = n(1-\frac{1}{p_1})\cdots(1-\frac{1}{p_k}) = (p_1^{\alpha_1} - p_1^{\alpha_1-1})\cdots(p_k^{\alpha_k} - p_k^{\alpha_k-1})$

ii) φ *is a multiplicative function*

iii) $S_\varphi(n) = \sum_{d|n}\varphi(d) = n$

Proof.

i) Define the functions $f(s) = \gcd(s,n)$ and $g(s) = 1$ and set $S = \{1,2,...,n\}$. Then we have

$$\sum_{\substack{s\in S\\ f(s)=1}} g(s) = \sum_{\substack{1\le s\le n\\ \gcd(s,n)=1}} 1 = \varphi(n) = \sum_{m=1}^{\infty}\mu(m)\sum_{\substack{1\le s\le n\\ m|\gcd(s,n)}} 1 =$$
$$\sum_{m|n}\mu(m)\sum_{\substack{1\le s\le n\\ m|s}} 1 = \sum_{m|n}\mu(m)\frac{n}{m} = n\sum_{m|n}\frac{\mu(m)}{m} = n(1-\frac{1}{p_1})\cdots(1-\frac{1}{p_k})$$

by using Lemma 2.1.1 part b) because the function $\frac{\mu(m)}{m}$ is multiplicative.

ii) By the representation we found in i) multiplicativity follows immediately.

iii) Since φ is multiplicative we can again use Lemma 2.1.1 and we get that

$$\begin{aligned} S_\varphi(n) &= (1+\varphi(p_1)+\cdots+\varphi(p_1^{\alpha_1}))\cdots(1+\varphi(p_k)+\cdots+\varphi(p_k^{\alpha_k})) = \\ &= (1+p_1-1+p_1^2-p_1+\cdots+p_1^{\alpha_1}-p_1^{\alpha_1-1})\cdots \\ &\cdots(1+p_k-1+p_k^2-p_k+\cdots+p_k^{\alpha_k}-p_1^{\alpha_k-1}) = p_1^{\alpha_1}\cdots p_k^{\alpha_k} = n. \end{aligned}$$

□

Lemma 2.1.3. *If f is a multiplicative function and* $\lim_{p^k\to\infty} f(p^k) = 0$ *then also* $\lim_{n\to\infty} f(n) = 0$ *where p runs through the primes.*

Proof. Let $n = p_1^{\alpha_1}\cdots p_k^{\alpha_k} = \prod_{p\in\mathbb{P}} p^{\nu(p,n)}$ the canonical representation of n as factors of prime powers. Then $f(n) = f(\prod_{p\in\mathbb{P}} p^{\nu(p,n)}) = \prod_{p\in\mathbb{P}} f(p^{\nu(p,n)})$ by multiplicativity of f. Now for each $\epsilon > 0$ there is some constant c such that for all powers $p^k \ge c$ we have that $|f(p^k)| < \epsilon$. If n is in such a fashion that every prime power is smaller than some constant c it follows that $n < c^c$, thus for every $n > c^c$ there is at least one prime factor such that $p^{\nu(p,n)} > c$ and therefore $f(p^{\nu(p,n)})$ gets arbitrarily small and so $f(n)$ itself also gets small.
More formally one would split the prime powers into three sets, one where $p^{\nu(p,n)}$ is smaller than some B, one where it is in between B and C and one where it is bigger than C. We know that there is some A such that $|f(p^{\nu(p,n)})| < A$ for all values of $p^{\nu(p,n)}$ and we choose B in a way that $|f(p^{\nu(p,n)})| < 1$ holds for all values in B, and C such that $|f(p^{\nu(p,n)})| < \epsilon$ holds for all values in C. This then gives $|f(n)| < A^B\epsilon \to 0$. □

Definition 2.1.4.

Landau symbols*: Let f, g, h be functions then we say that*

$f(x) = g(x) + \mathcal{O}(h(x))$ *if there exist some x_0 such that for all $x \ge x_0$ we have* $|f(x) - g(x)| \le c|h(x)|$ *for some* $c > 0$.

$f(x) = g(x) + o(h(x))$ *if* $\lim_{x\to\infty}\frac{f(x)-g(x)}{h(x)} \to 0$

*Two functions f, g are called **asymptotically equal**, written $f(x) \sim g(x)$, if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$.*

Example 2.1.1.
$\sin(x) = \mathcal{O}(1)$ *because* $|\sin(x)| \leq 1$.

$\sin(x) = x + \mathcal{O}(x^3)$ *as* $x \to \infty$ *because* $\sin(x) = \sum_{n=1}^{\infty}(-1)^{n-1}\frac{x^{2n-1}}{(2n-1)!}$.

$\sinh(x) = \frac{e^x - e^{-x}}{2} \sim \frac{e^x}{2}$.
$\pi(x) \sim \frac{x}{\log x}$.

Now we want to use Lemma 2.1.3 to get some direct estimates of arithmetic functions.

Claim 2.1.1. *We have that $\varphi(n) = \mathcal{O}(n)$ and for all $\epsilon > 0$ we get $n^{1-\epsilon} = o(\varphi(n))$.*

Proof. By the definition of the φ function we have $\varphi(n) = \sum_{\substack{1\leq k\leq n \\ \gcd(k,n)=1}} 1 \leq n$ and therefore $\varphi(n) = \mathcal{O}(n)$ holds.
For the second claim let $\epsilon > 0$ be arbitrary and consider the function $f(n) = \frac{n^{1-\epsilon}}{\varphi(n)}$. This way we get $f(p^m) = \frac{p^{m(1-\epsilon)}}{p^m - p^{m-1}} = \frac{p^{-m\epsilon}}{1-\frac{1}{p}} \leq 2p^{-m\epsilon}$. So clearly for every growing prime power we have $\lim_{p^m\to\infty} f(p^m) = 0$ and thus by Lemma 2.1.3 also $\lim_{n\to\infty} f(n) = 0$, which is what we wanted to show. □

Claim 2.1.2. *We have $n = \mathcal{O}(\sigma(n))$ and on the other hand for all $\epsilon > 0$ that $\sigma(n) = o(n^{1+\epsilon})$.*

Proof. By definition we get that $\sigma(n) = \sum_{d|n} d \geq n + 1$ and thus $n = \mathcal{O}(\sigma(n))$.
For the second part consider again a function $f(n) = \frac{\sigma(n)}{n^{1+\epsilon}}$. This way we get $f(p^m) = \frac{p^{m+1}-1}{p-1}\frac{1}{p^{m(1+\epsilon)}} = \frac{1-\frac{1}{p^{m+1}}}{p^{m\epsilon}(1-\frac{1}{p})} \leq 2\frac{1}{p^{m\epsilon}}$ and applying Lemma 2.1.3 gives the desired result. □

Claim 2.1.3. *A better upper bound is given by $\sigma(n) = \mathcal{O}(n \log n)$.*

This time we truly use analytic tools comparing the sum to a corresponding integral to avoid perturbations. Dirichlet himself compared $\frac{1}{N}\sum_{n=1}^{N} f(n)$ to $\frac{1}{T}\int_0^T f(t)dt$.

Proof.

$$\begin{aligned}
\sigma(n) &= \sum_{m|n} \frac{n}{m} = n\sum_{m|n}\frac{1}{m} \leq n\sum_{m\leq n}\frac{1}{m} = n + n\sum_{m=2}^{n}\frac{1}{m} = \\
&= n + n\sum_{m=2}^{n}\frac{1}{m}\int_{m-1}^{m} 1dt \leq n + n\sum_{m=2}^{n}\int_{m-1}^{m}\frac{1}{t}dt = \\
&= n + n\sum_{m=2}^{m}(\log(m) - \log(m-1)) = n + n\log n.
\end{aligned}$$

Since $n\log n$ is the dominating factor we get the desired result. This works because $\frac{1}{m}$ is smaller than or equal to the $\frac{1}{t}$ for t running from $m-1$ to m. □

2.2 Abel transformation

Remark 2.2.1 (Integration by parts). *For f, g functions we get $\int_a^b f(x)\,dg(x) = f(b)g(b) - f(a)g(a) - \int_a^b g(x)\,df(x)$.*

Proof. We know that $(f(x)g(x))' = f'(x)g(x) + g'(x)f(x)$ and therefore $\int_a^b f(x)\,dg(x) = \int_a^b (f(x)g(x))' - \int_a^b g(x)\,df(x)$ and by the fundamental theorem of calculus the statement follows. □

Theorem 2.2.1 (Abel Transformation 1). *Let f, g be arithmetic functions (or even only sequences). Then for $1 \leq P \leq Q$ we have that*

$$\sum_{n=P}^{Q} f(n)(g(n) - g(n-1)) = f(Q+1)g(Q) - f(P)g(P-1) - \sum_{n=P}^{Q} g(n)(f(n+1) - f(n)).$$

Proof.

$$\begin{aligned}\sum_{n=P}^{Q} f(n)(g(n) - g(n-1)) &= \sum_{n=P}^{Q} f(n)g(n) - \sum_{n=P}^{Q} f(n)g(n-1) = \\ &= \sum_{n=P}^{Q} f(n)g(n) - \sum_{n=P-1}^{Q-1} f(n+1)g(n) = \\ &= \sum_{n=P}^{Q-1} (f(n) - f(n+1))g(n) + f(Q)g(Q) - f(P)g(P-1) = \\ &= \sum_{n=P}^{Q} (f(n) - f(n+1))g(n) + \\ &+ f(Q)\,g(Q) - f(P)g(P-1) - f(Q)g(Q) + f(Q+1)g(Q) = \\ &= f(Q+1)g(Q) - f(P)g(P-1) - \sum_{n=P}^{Q} g(n)(f(n+1) - f(n))\end{aligned}$$

□

Theorem 2.2.2 (Abel Transformation 2). *Define a function $g(n) = \sum_{m=P}^{n} h(m)$ for some function h and set $g(P-1) = 0$. Then we get*

$$\sum_{n=P}^{Q} f(n)h(n) = f(Q+1)\sum_{n=P}^{Q} h(n) - \sum_{n=P}^{Q} (f(n+1) - f(n)) \sum_{m=P}^{n} h(m).$$

Proof. Follows directly by applying Theorem 2.2.1. □

Theorem 2.2.3 (Abel Transformation 3). *Let f be continuously differentiable on $[1, \infty)$ and h be an arithmetic function. Define $g(x) = \sum_{m=1}^{[x]} h(m)$, then we have*

$$\sum_{n=1}^{[x]} f(n)h(n) = f(x)g(x) - \int_1^x g(t)f'(t)dt.$$

Here $[x]$ denotes the integer part of x and $\{x\}$ is the fractional part of x.

Proof. Apply Theorem 2.2.2 with $P = 1$ and $Q = [x]$ and use the fact that $f(n) - f(n+1) = -\int_n^{n+1} f'(t)dt$.

$$\sum_{n=1}^{[x]} f(n)h(n) = f([x]+1)\sum_{n=1}^{[x]} h(n) + \sum_{n=1}^{[x]}(f(n) - f(n+1))\sum_{m=1}^{n} h(m) =$$
$$f([x]+1)g(x) + S.$$

$$\begin{aligned} S &= -\sum_{n=1}^{[x]}\left(\int_n^{n+1} f'(t)dt\right) g(n) = -\sum_{n=1}^{[x]}\int_n^{n+1} f'(t)g(t)dt = \\ &= -\int_1^{[x]+1} g(t)f'(t)dt = -\int_1^x g(t)f'(t)dt - \int_x^{[x]+1} g(t)f'(t)dt = \\ &= -\int_1^x g(t)f'(t)dt - g([x])\int_x^{[x]+1} f'(t)dt = \\ &= -\int_1^x g(t)f'(t)dt - g([x])(f([x]+1) - f(x)) = \\ &= -\int_1^x g(t)f'(t)dt - g(x)f([x]+1) + f(x)g(x). \end{aligned}$$

Therefore the statement follows. □

Remark 2.2.2. *We can extend arithmetic functions $g : \mathbb{N} \to \mathbb{C}$ to $g : [1,\infty) \to \mathbb{C}$ via $g(x) = g([x])$. Then instead of $\sum_{n=1}^{[x]}$ one writes $\sum_{n\le x}$ and we have again the* ***Abel summation formula***

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

for $g(x) = \sum_{m\le x} h(m)$.

Theorem 2.2.4 (Euler summation formula)**.** *Let $a \in \mathbb{N}$ and $f : [a,\infty) \to \mathbb{C}$ be a continuously differentiable function. Then we have*

$$\sum_{a\le n\le x} f(n) = \int_a^x f(t)dt + R$$

with the remainder term of the form

$$R = \int_a^x \{t\}f'(t)dt + f(a) - \{x\}f(x).$$

In particular if f is monotone and $f \ge 0$ we have

$$R = \begin{cases} \mathcal{O}(f(x)) & \text{if } f \text{ is incresing} \\ \mathcal{O}(f(a)) & \text{if } f \text{ is decreasing} \end{cases}$$

Proof. Recall that we can write the fractional part of x as $\{x\} = x - [x]$ and Theorem 2.2.3 applied to $g(x) = \sum_{m \leq x} h(m)$ gives

$$\sum_{n \leq x} f(n)h(n) = f(x)g(x) - \int_1^x g(t)f'(t)dt.$$

Define now $h(m) = \begin{cases} 0 & \text{if } m = 1, 2, ..., a-1 \\ 1 & \text{if } m \geq a \end{cases}$ thus $g(t) = \begin{cases} 0 & \text{if } t \leq a-1 \\ [t] - a + 1 & m \geq a \end{cases}$.
Therefore we get

$$\begin{aligned} \sum_{n \leq x} f(n)h(n) = \sum_{n=a}^{x} f(n) = f(x)([x] - a + 1) - \int_a^x ([t] - a + 1)f'(t)dt = \\ = [x]f(x) - (a-1)f(x) - \int_a^x [t]f'(t)dt + (a-1)\int_a^x f'(t)dt = \\ = [x]f(x) - (a-1)f(x) - \int_a^x [t]f'(t)dt + (a-1)(f(x) - f(a)) = \\ = [x]f(x) - (a-1)f(a) - \int_a^x [t]f'(t)dt. \end{aligned}$$

Now integration by party gives $\int_a^x f(t)dt = xf(x) - af(a) - \int_a^x tf'(t)dt$, and we arrive at

$$\begin{aligned} \sum_{n=a}^{x} f(n) - \int_a^x f(t)dt = ([x] - x)f(x) + f(a) - \int_a^x ([t] - t)f'(t)dt = \\ = \int_a^x \{t\}f'(t)dt + f(a) - \{x\}f(x) = R \end{aligned}$$

Now we get that $|\int_a^x \{t\}f'(t)dt| \leq \int_a^x |\{t\}f'(t)|dt \leq \int_a^x |f'(t)|dt = |\int_a^x f'(t)dt| = |f(x) - f(a)|$ because we consider f to be monotone, which implies that f' does not change its sign.
In conclusion this yields $|R| \leq |f(x) - f(a)| + |f(x)| + |f(a)|$ and by triangular inequality $R = \mathcal{O}(|f(x)| + |f(a)|)$. Finally a case distinction gives

$$R = \begin{cases} \mathcal{O}(f(x)) & \text{if f is incresing} \\ \mathcal{O}(f(a)) & \text{if f is decreasing} \end{cases}.$$

□

Claim 2.2.1. *For $P, Q \in \mathbb{N}$ witch $P < Q$ and $s > 1 \in \mathbb{R}$ we get*

$$\sum_{n=P}^{Q} \frac{1}{n^s} = \frac{1}{s-1}\left(\frac{1}{P^{s-1}} - \frac{1}{Q^{s-1}}\right) + \mathcal{O}\left(\frac{1}{p^s}\right).$$

In particular

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \frac{1}{s-1} + \mathcal{O}(1) \quad \text{and} \quad \sum_{n=N}^{\infty} \frac{1}{n^s} = \frac{1}{s-1}\frac{1}{N^{s-1}} + \mathcal{O}\left(\frac{1}{N^s}\right).$$

Proof. For $s > 1$ the function $f(t) = t^{-s}$ is monotonically decreasing and $f(t) > 0$ for all $t > 0$. Then by Theorem 2.2.4 we get

$$\sum_{n=P}^{Q} f(n) = \int_P^Q f(t)dt + R \quad \text{where} \quad R = \mathcal{O}(f(P)).$$

Therefore we have

$$\sum_{n=P}^{Q} \frac{1}{n^s} = \int_P^Q t^{-s}dt + \mathcal{O}(P^{-s}) = -\frac{1}{s-1}(Q^{-(s-1)} - P^{-(s-1)}) + \mathcal{O}(P^{-s}) = \frac{1}{s-1}\left(\frac{1}{P^{s-1}} - \frac{1}{Q^{s-1}}\right) + \mathcal{O}\left(\frac{1}{P^s}\right).$$

The special cases $P = 1, Q = \infty$ and $P = N, Q = \infty$ follow directly. □

Claim 2.2.2. *We have*

$$\sum_{n=1}^{N} \frac{1}{n} = \log(N) + \gamma + \mathcal{O}\left(\frac{1}{N}\right)$$

where

$$\gamma = 1 - \int_1^\infty \frac{\{t\}}{t^2}dt = \lim_{N\to\infty}\left(\sum_{n=1}^{N}\frac{1}{n} - \log(N)\right).$$

denotes the ***Euler-Mascheroni constant****.*

Proof. Set $f(t) = \frac{1}{t}$ and apply Theorem 2.2.4. Then this yields

$$\begin{aligned}\sum_{n=1}^{N} \frac{1}{n} &= \int_1^N \frac{1}{t}dt + f(1) - \{N\}f(N) + \int_1^N \{t\}f'(t)dt = \\ &= \int_1^N \frac{1}{t}dt + 1 - \frac{\{N\}}{N} - \int_1^N \frac{\{t\}}{t^2}dt = \\ &= \log(N) + 1 - \int_1^\infty \frac{\{t\}}{t^2}dt + \int_N^\infty \frac{\{t\}}{t^2}dt - \frac{\{N\}}{N}.\end{aligned}$$

Now we have that

$$\int_N^\infty \frac{\{t\}}{t^2}dt \leq \int_N^\infty \frac{1}{t^2}dt = \frac{1}{N} \text{ and so } \sum_{n=1}^{\infty}\frac{1}{n} = \log(N) + \gamma + \mathcal{O}\left(\frac{1}{N}\right).$$

□

Remark 2.2.3. *The Euler-Mascheroni constant $\gamma \sim 0.5772156$ is* ***conjectured*** *to be irrational, but it is still open. It even is not known if it is algebraic or transcendental.*

2.3 Average estimates of arithmetic functions

Recall that we have already shown that $\varphi(n) = \mathcal{O}(n)$ and for all $\epsilon > 0$ we have $n^{1-\epsilon} = \mathcal{O}(\varphi(n))$. Even though the φ function itself is not behaving nicely we will see that averaging has a smoothing effect on the estimates.

Claim 2.3.1.

i) $\frac{1}{N}\sum_{n=1}^{N}\varphi(n) = \frac{3}{\pi^2}N + \mathcal{O}(\log(N))$

ii) $\frac{1}{N}\sum_{n=1}^{N}\frac{\varphi(n)}{n} = \frac{6}{\pi^2} + \mathcal{O}\left(\frac{\log^2(N)}{N}\right)$.

Proof.

i) Here we use that $\sum_{d|n}\varphi(d) = S_\varphi(n) = n$ and the Möbius-inversion formula which gives $\varphi(n) = \sum_{d|n}\mu(d)S_\varphi(\frac{n}{d}) = \sum_{d|n}\mu(d)\frac{n}{d}$. This way we get

$$\begin{aligned}\sum_{n=1}^{N}\varphi(n) &= \sum_{n=1}^{N}\sum_{d|n}\mu(d)\frac{n}{d} = \sum_{k}\sum_{\substack{d\\ kd=n}} k\mu(d) = \sum_{d=1}^{N}\mu(d)\sum_{kd\le N}k = \\ &= \sum_{d=1}^{N}\mu(d)\sum_{k\le\frac{N}{d}}k = \sum_{d=1}^{N}\mu(d)\frac{\left[\frac{N}{d}\right]\left(\left[\frac{N}{d}\right]+1\right)}{2} = \\ &= \frac{1}{2}\sum_{d=1}^{N}\mu(d)\left(\frac{N}{d}-\left\{\frac{N}{d}\right\}\right)\left(\frac{N}{d}-\left\{\frac{N}{d}\right\}+1\right) = \\ &= \frac{1}{2}\sum_{d=1}^{N}\mu(d)\left(\frac{N^2}{d^2}-2\frac{N}{d}\left\{\frac{N}{d}\right\}+\left\{\frac{N}{d}\right\}^2+\frac{N}{d}-\left\{\frac{N}{d}\right\}\right) = \\ &= \frac{N^2}{2}\underbrace{\sum_{d=1}^{N}\frac{\mu(d)}{d^2}}_{\Sigma_1} + \underbrace{\frac{N}{2}\sum_{d=1}^{N}\frac{\mu(d)}{d}\left(1-2\left\{\frac{N}{d}\right\}\right)}_{\Sigma_2} + \underbrace{\frac{1}{2}\sum_{d=1}^{N}\mu(d)\left\{\frac{N}{d}\right\}\left(\left\{\frac{N}{d}\right\}-1\right)}_{\Sigma_3}.\end{aligned}$$

By the last Claim we get that $\sum_{m=1}^{N}\frac{1}{m} = \log(N) + \gamma + \mathcal{O}(\frac{1}{N})$ and with that $\Sigma_2 = \mathcal{O}(\sum_{d=1}^{N}\frac{1}{d}) = \mathcal{O}(\log(N))$. For the last part we obviously get $\Sigma_3 = \mathcal{O}(\sum_{d=1}^{N}1) = \mathcal{O}(N)$. An estimate for Σ_1 goes as follows

$$\Sigma_1 = \sum_{m=1}^{N}\frac{\mu(m)}{m^2} = \sum_{m=1}^{\infty}\frac{\mu(m)}{m^2} - \sum_{m=N+1}^{\infty}\frac{\mu(m)}{m^2}.$$

Here $\sum_{m=N+1}^{\infty}\frac{\mu(m)}{m^2}$ is absolutely convergent because $|\frac{\mu(m)}{m^2}| \le \frac{1}{m^2}$ and thus

$$\begin{aligned}\sum_{m=N+1}^{\infty}\frac{\mu(m)}{m^2} &= \mathcal{O}\left(\sum_{m=N+1}^{\infty}\frac{1}{m^2}\right) = \mathcal{O}\left(\frac{1}{2-1}\frac{1}{(N+1)^{2-1}} + \mathcal{O}\left(\frac{1}{(N+1)^2}\right)\right) = \\ &= \mathcal{O}\left(\frac{1}{N}\right) + \mathcal{O}\left(\frac{1}{N^2}\right) = \mathcal{O}\left(\frac{1}{N}\right).\end{aligned}$$

This follows again by one of the above Claims with $s = 2$ and now
$\Sigma_1 = \sum_{m=1}^{\infty}\frac{\mu(m)}{m^2} + \mathcal{O}(\frac{1}{N})$.
In order to find an estimate for the sum we multiply $\zeta(s) = \sum_{k=1}^{\infty}\frac{1}{k^s}$ with the infinite series $\sum_{m=1}^{\infty}\frac{\mu(m)}{m^s}$, that is

$$\begin{aligned}\sum_{k=1}^{\infty}\frac{1}{k^s}\sum_{m=1}^{\infty}\frac{\mu(m)}{m^s} &= \sum_{k\ge1}\sum_{m\ge1}\frac{\mu(m)}{(km)^s} = \sum_{n=1}^{\infty}\sum_{d|n}\frac{\mu(d)}{n^s} = \\ &= \sum_{n=1}^{\infty}\frac{1}{n^s}\sum_{d|n}\mu(d) = 1.\end{aligned}$$

Now we have that $\sum_{m=1}^{\infty} \frac{\mu(m)}{m^s} = \frac{1}{\zeta(s)}$ for every $s > 1$. In our case we have $s = 2$ and $\zeta(2) = \frac{\pi^2}{6}$, so it follows that

$$\Sigma_1 = \frac{6}{\pi^2} + \mathcal{O}\left(\frac{1}{N}\right).$$

Plugging together this yields

$$\sum_{n=1}^{N} \varphi(n) = \frac{N^2}{2}\left(\frac{6}{\pi^2} + \mathcal{O}\left(\frac{1}{N}\right)\right) + \mathcal{O}(N\log(N)) = \frac{3}{\pi^2}N^2 + \mathcal{O}(N\log(N))$$

and

$$\frac{1}{N}\sum_{n=1}^{N} \varphi(n) = \frac{3}{\pi^2}N + \mathcal{O}(\log(N)).$$

ii) Applying Theorem 2.2.3 to $g(x) = \sum_{m \leq x} h(m)$ we get

$$\sum_{n \leq x} f(n)h(n) = f(x)g(x) - \int_1^x g(t)f'(t)dt.$$

By setting $f(n) = \frac{1}{n}$ and $h(n) = \varphi(n)$ this yields

$$\begin{aligned}
\sum_{m=1}^{N} \frac{\varphi(n)}{n} &= f(N)g(N) - \int_1^N g(t)f'(t)dt = \frac{1}{N}\sum_{n=1}^{N}\varphi(n) + \int_1^N \left(\sum_{n=1}^{t}\varphi(n)\right)\frac{1}{t^2}dt = \\
&= \frac{3}{\pi^2}N + \mathcal{O}(\log(N)) + \int_1^N \left(\frac{3}{\pi^2}t^2 + \mathcal{O}(t\log(t))\right)\frac{1}{t^2}dt = \\
&= \frac{3}{\pi^2}N + \frac{3}{\pi^2}(N-1) + \mathcal{O}(\log(N)) + \mathcal{O}\left(\int_1^N \log(t)d\log(t)\right) = \\
&= \frac{6}{\pi^2}N + \mathcal{O}(\log(N)) + \mathcal{O}\left((\log(N))^2 - (\log(1))^2\right) = \frac{6}{\pi^2}N + \mathcal{O}(\log^2(N)).
\end{aligned}$$

Therefore

$$\frac{1}{N}\sum_{n=1}^{N}\frac{\varphi(n)}{n} = \frac{6}{\pi^2} + \mathcal{O}\left(\frac{\log^2(N)}{N}\right) \sim \frac{6}{\pi^2}.$$

□

2.4 Density of k-free and square-free numbers

Definition 2.4.1. *A natural number $n \in \mathbb{N}$ is called **k-free** if there is no $m > 1 \in \mathbb{N}$ such that m^k is a divisor of n, that is that no k-th power is a divisor of n. **Square-free** is the special case $k = 2$. Define a function*

$$\mu_k(n) = \begin{cases} 1 & \text{if } n \text{ is k-free} \\ 0 & \text{else} \end{cases}.$$

Remark 2.4.1. *Let n be a natural number such that $n = h^k l$ where l is k-free and $h \geq 1$. If m^k is a divisor of n then it is also a divisor of h^k and thus m itself is a divisor of h.*

The other way around if m is a divisor of h then clearly also m^k is a divisor of h^k. With that we have

$$\sum_{m^k|h} \mu(m) = \sum_{m|h} \mu(m) = \begin{cases} 1 & \text{if } h = 1 \\ 0 & \text{if } h > 1 \end{cases}.$$

Therefore $\mu_k(n) = \sum_{m^k|n} \mu(m)$.

Definition 2.4.2. *The* ***natural density*** *of a set $A \subset \mathbb{N}$ is defined to be $\frac{\#\{n \le N | n \in A\}}{N} \to \alpha$ as $N \to \infty$. Clearly here we have $0 \le \alpha \le 1$.*

Claim 2.4.1. *We have*

$$\frac{1}{N} \sum_{n=1}^{N} \mu_k(n) = \frac{1}{\zeta(k)} + \mathcal{O}\left(N^{\frac{1}{k}-1}\right).$$

That is that de density of the k-free numbers equals $\frac{1}{\zeta(k)}$, therefore in particular the density of the square-free numbers equals $\frac{6}{\pi^2}$.

Proof.

$$\begin{aligned}
\frac{1}{N} \sum_{n=1}^{N} \mu_k(n) &= \frac{1}{N} \sum_{k \le N} \sum_{m^k|n} \mu(m) = \frac{1}{N} \sum_{m,l} \sum_{m^k l \le N} \mu(m) = \frac{1}{N} \sum_{m \le N^{\frac{1}{k}}} \mu(m) \sum_{l \le \frac{N}{m^k}} 1 = \\
&= \frac{1}{N} \sum_{m \le N^{\frac{1}{k}}} \mu(m) \left[\frac{N}{m^k}\right] = \frac{1}{N} \sum_{m \le N^{\frac{1}{k}}} \mu(m) \left(\frac{N}{m^k} - \left\{\frac{N}{m^k}\right\}\right) = \\
&= \sum_{m=1}^{\infty} \frac{\mu(m)}{m^k} - \sum_{m > N^{\frac{1}{k}}} \frac{\mu(m)}{m^k} + \frac{1}{N}\mathcal{O}\left(\sum_{m \le N^{\frac{1}{k}}} 1\right) = \\
&= \frac{1}{\zeta(k)} + \mathcal{O}\left(\sum_{m > N^{\frac{1}{k}}} \frac{1}{m^k}\right) + \mathcal{O}\left(N^{\frac{1}{k}-1}\right) = \\
&= \frac{1}{\zeta(k)} + \mathcal{O}\left(\frac{1}{k-1} \frac{1}{N^{\frac{1}{k}(k-1)}}\right) + \mathcal{O}\left(N^{\frac{1}{k}-1}\right) = \frac{1}{\zeta(k)} + \mathcal{O}\left(N^{\frac{1}{k}-1}\right).
\end{aligned}$$

□

Claim 2.4.2. *For the* ***Dirichlet divisor problem*** *we get that*

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

Here we are going to use the so called **Dirichlet hyperbola method** to count the lattice points under the hyperbola $xy = N$ in order to get the sum of the divisors.

Proof. First recall that $\sum_{n=1}^{N} \tau(n) = \sum_{n=1}^{N} \sum_{mk=n} 1$ and observe that this summands can be represented by integer points under the hyperbola $xy = N$. A further observation (dating back to Dirichlet) is, that the square with coordinated $(\sqrt{N}, \sqrt{N})$ divides the area under the hyperbola in two parts with equally many integer points, here denoted by $A \cup B$ and $B \cup C$ (see figure below). Thus it follows that we can express the sum as

$$\begin{aligned}
\sum_{n=1}^{N} \tau(n) &= 2\sum_{m=1}^{\sqrt{N}} \sum_{k \le \frac{N}{m}} 1 - [\sqrt{N}]^2 = 2\sum_{n=1}^{\sqrt{N}} \left[\frac{N}{n}\right] - [\sqrt{N}]^2 = \\
&= 2\sum_{n=1}^{\sqrt{N}} \left(\frac{N}{n} - \left\{\frac{N}{n}\right\}\right) - (\sqrt{N} - \{\sqrt{N}\})^2 = \\
&= 2N\sum_{n=1}^{\sqrt{N}} \frac{1}{n} - 2\sum_{n=1}^{\sqrt{N}} \left\{\frac{N}{n}\right\} - (N - 2\sqrt{N}\{\sqrt{N}\} + \{\sqrt{N}\}^2) = \\
&= 2N(\log([\sqrt{N}] + \gamma + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) + \mathcal{O}\left(\sqrt{N}\right) - N + 2\sqrt{N}\mathcal{O}(1) + \mathcal{O}(1) = \\
&= 2N\log(\sqrt{N} - \{\sqrt{N}\}) + 2N\gamma + \mathcal{O}\left(\sqrt{N}\right) - N.
\end{aligned}$$

Now we claim that $\log(\sqrt{N} - \{\sqrt{N}\}) = \log(\sqrt{N}) + \mathcal{O}(\frac{1}{\sqrt{N}})$, and indeed by the mean value theorem there exists some $c \in (\sqrt{N} - \{\sqrt{N}\}, \sqrt{N})$ such that

$$(\log x)'_c = \frac{1}{c} = \frac{\log(\sqrt{N}) - \log(\sqrt{N} - \{\sqrt{N}\})}{\{\sqrt{N}\}}$$

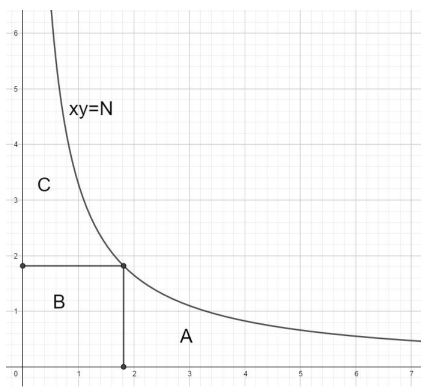
and therefore

$$\log(\sqrt{N} - \{\sqrt{N}\}) = \log(\sqrt{N}) - \frac{\{\sqrt{N}\}}{c}.$$

Since $\frac{1}{c} < \frac{1}{\sqrt{N} - \{\sqrt{N}\}} < \frac{2}{\sqrt{N}}$ holds for large enough N we see that $\frac{1}{c} = \mathcal{O}(\frac{1}{\sqrt{N}})$. In conclusion we get that

$$\begin{aligned}
\sum_{n=1}^{N} \tau(n) &= 2N(\log(\sqrt{N} + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) + N(2\gamma - 1) + \mathcal{O}\left(\sqrt{N}\right) = \\
&= N\log(N) + (2\gamma - 1)N + \mathcal{O}\left(\sqrt{N}\right).
\end{aligned}$$

$\square$



Definition 2.4.3. *The* ***Dirichlet series*** *of an arithmetic function f at some point $s \in \mathbb{C}$ is defined to be*

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

Claim 2.4.3. *If the Dirichlet series $D_f(s)$ and $D_g(s)$ are absolutely convergent then we have the identity*

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

Proof.

$$D_f(s)D_g(s) = \sum_{m=1}^{\infty} \frac{f(m)}{m^s} \sum_{k=1}^{\infty} \frac{g(k)}{k^s} = \sum_{m=1}^{\infty} \sum_{k=1}^{\infty} \frac{f(m)g(k)}{(mk)^s} =$$
$$= \sum_{n=1}^{\infty} \sum_{\substack{m,k \\ mk=n}} \frac{f(m)g(k)}{n^s} = \sum_{n=1}^{\infty} \frac{1}{n^s} \sum_{m|n} f(n)g(\frac{n}{m}) = \sum_{n=1}^{\infty} \frac{f*g(n)}{n^s}.$$

□

Corollary 2.4.1. *We have that*

$$\sum_{m=1}^{\infty} \frac{\mu(m)}{m^s} = \frac{1}{\zeta(s)} \qquad \textit{for } s > 1.$$

We have already seen a proof of this statement in Claim 2.3.1, but here we show it again using the Dirichlet product.

Proof. Define the function $I(n) = 1$ for all $n \in \mathbb{N}$, then we have that $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = D_I(s)$. Now we use the Möbius-inversion formula $\mu * S_f = f$ and get that

$$\mu * I = \mu * S_e = e = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n > 1 \end{cases}$$

because $S_e(n) = \sum_{d|n} e(\frac{n}{d}) = 1 = I(n)$. Therefore we get

$$D_I(s)D_\mu(s) = D_{I*\mu}(s) = D_e(s) = \sum_{n=1}^{\infty} \frac{e(n)}{n^s} = 1$$

and with that $\zeta(s)D_\mu(s) = 1$ concluding the proof. □

Claim 2.4.4. *Assume that $D_f(s)$ is absolutely convergent. If f is a multiplicative function then we get*

$$D_f(s) = \prod_{p \in \mathbb{P}} \sum_{\nu \geq 0} \frac{f(p^\nu)}{p^{\nu s}}.$$

If f is even strongly multiplicative, that is $f(mn) = f(m)f(n)$ for all m, n then we get the ***Euler product*** *representation*

$$D_f(s) = \prod_{p \in \mathbb{P}} \frac{1}{1 - \frac{f(p)}{p^s}}.$$

Proof. First note that $\sum_{\nu\geq 0}\frac{f(p^\nu)}{p^{\nu s}}$ is absolutely convergent for any prime p as a sub-series of $\sum_{n=1}^{\infty}\frac{f(n)}{n^s}$. Therefore

$$\prod_{p\leq k}\sum_{\nu\geq 0}\frac{f(p^\nu)}{p^{\nu s}} = \sum_{\substack{n\\ p|n\Rightarrow p\leq k}}\frac{f(n)}{n^s}.$$

Let $k\to\infty$ then

$$\prod_{p}\sum_{\nu\geq 0}\frac{f(p^\nu)}{p^{\nu s}} = \sum_{n=1}^{\infty}\frac{f(n)}{n^s} = D_f(S).$$

If f is strongly multiplicative then $f(p^k)=f(p)^k$ and so

$$\sum_{\nu\geq 0}\frac{f(p^\nu)}{p^{\nu s}} = \sum_{\nu\geq 0}\left(\frac{f(p)}{p^s}\right)^\nu = \frac{1}{1-\frac{f(p)}{p^s}}$$

because $|\frac{f(p)}{p^s}|<1$ since else the series would not be convergent. □

2.5 Analytic properties of the Dirichlet series

Claim 2.5.1. *If the Dirichlet series $D_f(s)=\sum_{n=1}^{\infty}\frac{f(n)}{n^s}$ converges for some $s_0\in\mathbb{C}$ then it converges uniformly for every $s\in\mathbb{C}$ for which $-\alpha\leq arg(s-s_0)\leq\alpha$ for any $\alpha<\frac{\pi}{2}$.*

Let us first recall the definition of **uniform convergence**. Let S be a set and $f_n: S\to\mathbb{C}$ a sequence of functions. We say that $\{f_n\}$ converges uniformly to a limit f if we have that for every $\epsilon>0$ there exists some index $N\in\mathbb{N}$ such that for every $x\in S$ and $n\geq N$ $|f_n(x)-f(x)|<\epsilon$. Therefore we see that the series $D_f(s)=\sum_{n=1}^{\infty}\frac{f(n)}{n^s}$ converges uniformly if the sequence $\{D_{f,m}(s)\}_{m=1}^{\infty}=\{\sum_{n=1}^{m}\frac{f(n)}{n^s}\}_{m=1}^{\infty}$ converges uniformly for every $s\in\mathbb{C}$.

Proof. Without loss of generality we can assume $s_0=0$ because else we can transform the series by $D_f(s_0)=D_{\frac{f(n)}{n^{s_0}}}(0)$. Since $D_f(0)$ is convergent by assumption we get that the partial sums from some point on get arbitrarily small, that is for every $\epsilon>0$ and fixed $\alpha<\frac{\pi}{2}$ there is some index K depending on ϵ and α such that $|\sum_{n=M}^{N}f(n)|<\epsilon cos(\alpha)$ whenever $N\geq M\geq K$.
We want to show that for every $s\in\mathbb{C}$ with $|arg(s)|<\alpha$ we have that $|\sum_{n=M}^{N}\frac{f(n)}{n^s}|<\epsilon$ and use Theorem 2.2.2 for that matter.

$$\left|\sum_{n=M}^{N}\frac{f(n)}{n^s}\right| = \left|(N+1)^{-s}\sum_{n=M}^{N}f(n)+\sum_{n=M}^{N}(n^{-s}-(n+1)^{-s})\sum_{m=M}^{n}f(m)\right|\leq$$
$$\leq |(N+1)^{-s}|\left|\sum_{n=M}^{N}f(n)\right|+\sum_{n=M}^{N}|n^{-s}-(n+1)^{-s}|\left|\sum_{m=M}^{n}f(m)\right|.$$

Recall now the **de Moivre's formula**. Let $x\in\mathbb{R}$ then we have that $e^{ix}=\cos(x)+i\ \sin(x)$ and for $s\in\mathbb{C}$ we have that $x^s=x^\sigma x^{it}=x^\sigma e^{\log x it}=x^\sigma(\cos(t\log x+i\sin(t\log x)$. Since the absolute value of the latter part is 1 we get that $|x^s|=|x^\sigma|=|x^{Re(s)}|$.
Now we use this fact and the estimate for the series in 0 and get that

$$\left|\sum_{n=M}^{N}\frac{f(n)}{n^s}\right|\leq (N+1)^{-Re(s)}\epsilon\cos(\alpha)+\epsilon\cos(\alpha)\sum_{n=M}^{N}|n^{-s}-(n+1)^{-s}|.$$

As an estimate for the summands we get that

$$\left|\frac{1}{n^s} - \frac{1}{(n+1)^s}\right| = |s| \left|\int_n^{n+1} \frac{1}{t^{s+1}} dt\right| \leq |s| \int_n^{n+1} \frac{1}{|t^{s+1}|} dt = |s| \int_n^{n+1} t^{-Re(s)-1} dt =$$
$$= \frac{|s|}{Re(s)}(n^{-Re(s)} - (n+1)^{-Re(s)}) \leq \frac{1}{cos(\alpha)}(n^{-Re(s)} - (n+1)^{-Re(s)}).$$

This last inequality follows because for $|arg(s)| < \alpha$ we have that $cos(arg(s)) = \frac{Re(s)}{|s|} \geq \alpha$ and therefore $\frac{|s|}{Re(s)} \leq \frac{1}{cos(\alpha)}$.
Plugging things together yields

$$\left|\sum_{n=M}^{N} \frac{f(n)}{n^s}\right| \leq (N+1)^{-Re(s)} \epsilon \cos(\alpha) + \frac{\epsilon \cos(\alpha)}{cos(\alpha)} \sum_{n=M}^{N} (n^{-Re(s)} - (n+1)^{-Re(s)}) =$$
$$= (N+1)^{-Re(s)} \epsilon \cos(\alpha) + \epsilon(M^{-Re(s)} - (N+1)^{-Re(s)}) =$$
$$= \epsilon M^{-Re(s)} + \epsilon(N+1)^{-Re(s)}(cos(\alpha) - 1) \leq \epsilon M^{-Re(s)} \leq \epsilon.$$

□

Definition 2.5.1. *The* ***convergence abscissa*** *of the Dirichlet series $D_f(s)$ is defined as*

$$\sigma_0 = inf\{\sigma = Re(s) : D_f(s) \text{ is convergent}\}.$$

Definition 2.5.2. *The* ***abscissa of absolute convergence*** *is defined to be*

$$\sigma_0' = inf\{\sigma = Re(s) : D_{|f|}(s) \text{ is convergent}\}.$$

Example 2.5.1.

i) For $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ we have that $\sigma_0 = 1$.

ii) For $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^s}$ we have that $\sigma_0 = 0$ and $\sigma_0' = 1$.

Claim 2.5.2. *For the abscissa we have that $\sigma_0' \leq \sigma_0 + 1$.*

Proof. Let $s = \sigma + it \in \mathbb{C}$ and $\sigma > \sigma_0$. Then we know that $D_f(s)$ is convergent and each of its summands is bounded, that is $|f(n)n^{-s}| = |f(n)||n^{-s}| = |f(n)n^{-\sigma}| \leq K$ for all $\sigma > \sigma_0$ and some K constant. Therefore we have that

$$\sum_{n=1}^{\infty} \left|\frac{f(n)}{n^{s+1+\epsilon}}\right| = \sum_{n=1}^{\infty} \frac{|f(n)|}{|n^{\sigma+1+\epsilon}|} \leq K \sum_{n=1}^{\infty} \frac{1}{n^{1+\epsilon}} < \infty$$

for any $\epsilon > 0$. Thus we see that $D_{|f|}(s + 1 + \epsilon)$ is absolutely convergent and so σ_0' is at most $\sigma + 1$. □

Claim 2.5.3. *Let the Dirichlet series $D_f(s)$ and $D_g(s)$ be convergent for $Re(s) > \sigma_0$. If for all s with $Re(s) > \sigma_0$ the equation $D_f(s) = D_g(s)$ holds, then we have that $f(n) = g(n)$ holds for all $n \in \mathbb{N}$.*

Proof. Consider the difference Dirichlet series

$$D_h(s) = D_f(s) - D_g(s) = \sum_{n=1}^{\infty} \frac{f(n) - g(n)}{n^s} = \sum_{n=1}^{\infty} \frac{h(n)}{n^s} = 0.$$

We now want to show that $h(n) = 0$ for all $n \in \mathbb{N}$. Suppose that there exists some $N \in \mathbb{N}$ such that $h(N) \neq 0$ and take this N to be minimal. Then we get that

$$D_h(s) = \sum_{n=N}^{\infty} \frac{h(n)}{n^s} = 0 \text{ and so } N^s \sum_{n=N}^{\infty} \frac{h(n)}{n^s} = 0.$$

Let $\epsilon > 0$ be arbitrary, then since $N^s D_h(s)$ is uniformly convergent for all s where $Re(s) > \sigma_0$ there exist some M such that $|\sum_{n=M+1}^{\infty} \frac{h(n)N^s}{n^s}| < \frac{\epsilon}{2}$.
We have that

$$0 = \sum_{n=N}^{\infty} \frac{N^s h(n)}{n^s} = h(N) + \sum_{n=N+1}^{\infty} \frac{N^s h(n)}{n^s}$$

and therefore

$$-h(N) = \sum_{n=N+1}^{M} \frac{N^s h(n)}{n^s} + \sum_{n=M+1}^{\infty} \frac{N^s h(n)}{n^s}.$$

For the absolute value we see that

$$|h(N)| \leq \left| \sum_{n=N+1}^{M} \frac{N^s h(n)}{n^s} \right| + \left| \sum_{n=M+1}^{\infty} \frac{N^s h(n)}{n^s} \right| < 2\frac{\epsilon}{2} = \epsilon$$

because once ϵ is fixed so are both M and N and then $\left(\frac{N}{n}\right)^s$ gets small when $n \in \{N+1, ...M\}$. Thus we get that $h(N) = 0$ which is a contradiction. □

Remark 2.5.1. *Since $D_f(s)$ is uniformly convergent on any compact subregion of $Re(s) > \sigma_0$ and each partial sum $\sum_{n=1}^{N} \frac{f(n)}{n^s}$ is analytic, it follows that $D_f(s)$ itself is analytic and may be differentiated term by term.*

2.6 Connection between arithmetic functions and Dirichlet series

Definition 2.6.1. *The **Liouville function** is defined to be*

$$\lambda(n) = \begin{cases} 1 & \text{if } n = 1 \\ (-1)^{\alpha_1 + \cdots + \alpha_k} & \text{if } n = p_1^{\alpha_1} \cdots p_k^{\alpha_k} \end{cases}.$$

We have that $\lambda(n)$ is a strongly multiplicative function.

Definition 2.6.2. *The **von Mangoldt function** is defined to be*

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

Claim 2.6.1. *The following identities hold:*

i) $\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}$ *if* $Re(s) > 1$

ii) $\sum_{n=1}^{\infty} \frac{\tau(n)}{n^s} = \zeta(s)^2$ *if* $Re(s) > 1$

iii) $\sum_{n=1}^{\infty} \frac{\sigma(n)}{n^s} = \zeta(s)\,\zeta(s-1)$ *if* $Re(s) > 2$

iv) $\sum_{n=1}^{\infty} \frac{\varphi(n)}{n^s} = \frac{\zeta(s-1)}{\zeta(s)}$ *if* $Re(s) > 2$

v) $\sum_{n=1}^{\infty} \frac{\lambda(n)}{n^s} = \frac{\zeta(2s)}{\zeta(s)}$ *if* $Re(s) > 1$

vi) $\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)}$ *if* $Re(s) > 1$.

Proof.

i) We have already seen two proofs of this, one in 2.3.1.

ii) Note again that $\zeta(s) = D_I(s)$ and so by claim 2.4.3 it is enough to show that $I * I = \tau$.

$$I * I(N) = \sum_{d|n} I(d)I(\frac{n}{d}) = \sum_{d|n} 1 = \tau(n).$$

iii) We can write

$$D_\sigma(s) = \sum_{n=1}^{\infty} \frac{\sigma(n)}{n^s} = \sum_{n=1}^{\infty} \frac{\sigma(n)/n}{n^{s-1}} \overset{?}{=} \sum_{n=1}^{\infty} \frac{1/n}{n^{s-1}} \sum_{n=1}^{\infty} \frac{1}{n^{s-1}} = \zeta(s)\ \zeta(s-1).$$

Clearly here we have that the convergence abscissa is $\sigma_0 = 2$, and set $f(n) = \frac{1}{n}$. With this choice of f we have that $D_\sigma(s) = D_{\sigma \cdot f}(s-1)$ and we want to show that $D_{\sigma \cdot f}(s-1) = D_f(s-1)D_I(s-1) = D_{f*I}(s-1)$. Now we see

$$f * I(n) = \sum_{d|n} I(d)f(\frac{n}{d}) = \sum_{d|n} 1\frac{d}{n} = \frac{1}{n}\sum_{d|n} d = \frac{1}{n}\sigma(n) = (\sigma \cdot f)(n).$$

iv) Use again the function $f(n) = \frac{1}{n}$. We have

$$\sum_{n=1}^{\infty} \frac{1/n}{n^{s-1}} \sum_{n=1}^{\infty} \frac{\varphi(n)/n}{n^{s-1}} \overset{?}{=} \sum_{n=1}^{\infty} \frac{1}{n^{s-1}}.$$

To verify this we have to prove that $D_f(s-1)D_{\varphi \cdot f}(s-1) = D_I(s-1)$.

$$f * (\varphi \cdot f)(n) = \sum_{d|n} f(d)\ \varphi(\frac{n}{d})\ f(\frac{n}{d}) = \sum_{d|n} \frac{1}{d}\varphi(\frac{n}{d})\frac{d}{n} = \frac{1}{n}\sum_{d|n} \varphi(d) = \frac{1}{n}n = 1 = I(n).$$

v) Since we noted that λ is a strongly multiplicative function we get by claim 2.4.4 an Euler product representation of the form

$$D_\lambda(s) = \prod_p \frac{1}{1 - \frac{\lambda(p)}{p^s}} = \prod_p \frac{1}{1 + \frac{1}{p^s}} = \prod_p \frac{1 - \frac{1}{p^s}}{1 - \frac{1}{p^{2s}}} = \prod_p \frac{1}{1 - \frac{1}{p^{2s}}} \left(\prod_p \frac{1}{1 - \frac{1}{p^s}} \right)^{-1} = \frac{\zeta(2s)}{\zeta(s)}.$$

Here we used that $(a+b)(a-b) = a^2 - b^2$ and therefore also $(a+b) = \frac{a^2-b^2}{a-b}$.

vi) Recall that

$$D_f'(s) \sum_{n=1}^{\infty} \frac{-\log n\ f(n)}{n^s} \text{ and so } -\zeta'(s) = \sum_{n=1}^{\infty} \frac{\log n}{n^s} = D_{\log}(s).$$

We want to show that $D_\Lambda(s)\ \zeta(s) = D_{\log}(s)$, that is $\Lambda * I = S_\Lambda = \log$. Indeed we have that

$$S_\Lambda(n) = \sum_{d|n} \Lambda(d) = \sum_{p_i^{\alpha_i}|n} \log(p_i) = \sum_{i=1}^{k} \alpha_i \log(p_i) = \sum_{i=1}^{k} \log(p_i^{\alpha_i}) = \log n.$$

□

Remark 2.6.1. *In the last step of the proof we have just seen that $S_\Lambda(n) = \log n$.*

Claim 2.6.2. *For the Liouville function we get the identity*

$$\lambda(n) = \sum_{m^2|n} \mu\left(\frac{n}{m^2}\right).$$

Proof. Define a function $q(m) = \begin{cases} 1 & \text{if } m = n^2 \\ 0 & \text{else} \end{cases}$, with that we get

$$\zeta(2s) = \sum_{n=1}^{\infty} \frac{1}{n^{2s}} = \sum_{m=1}^{\infty} \frac{q(m)}{m^s} = D_q(s).$$

In the last claim we have proved that $D_\lambda(s)\,\zeta(s) = D_q(s)$ and thus by claim 2.5.3 we get that $\lambda * I = S_\lambda = q$. By the Möbius inversion formula we get that $\lambda = \mu * S_\lambda = \mu * q$, which gives

$$\lambda(n) = \sum_{d|n} \mu(\frac{n}{d})\; q(d) = \sum_{m^2|n} \mu\left(\frac{n}{m^2}\right).$$

□

2.7 Analytic continuity of Dirichlet series

Definition 2.7.1. *Let $U, V \subset \mathbb{C}$ be open subsets such that $U \subset V$. Let $f : U \to \mathbb{C}$ be an analytic function and $F : V \to \mathbb{C}$ an analytic function such that $F|_U = f$, that is $F(z) = f(z)$ for all $z \in U$. In this setting we call F an **analytic continuation** of f.*

Remark 2.7.1. *Set $A(n) = (-1)^{n-1}$ for all $n \geq 1$, then we get*

$$(1 - 2^{1-s})\zeta(s) = (1 - 2^{1-s}) \sum_{n=1}^{\infty} \frac{1}{n^s} = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \cdots - 2\left(\frac{1}{2^s} + \frac{1}{4^s} + \cdots\right) =$$

$$= 1 - \frac{1}{2^s} + \frac{1}{3^s} - \frac{1}{4^s} \cdots = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^s} = D_A(s).$$

Here we see that $D_A(s)$ has a convergence abscissa $\sigma_0 = 0$ while $\zeta(s)$ is only convergent for $Re(s) > 1$.

Question: Does $\zeta(s)$, or other Dirichlet series, have an analytic continuation to the left of the convergence half-plane?

Theorem 2.7.1 (Landau). *If $f : \mathbb{N} \to \mathbb{R}$ is such that $f(n) \geq 0$ for all $n \in \mathbb{N}$ and the Dirichlet series $D_f(s)$ has a finite convergence abscissa σ_0, then $D_f(s)$ is holomorphic in $Re(s) > \sigma_0$ but can not be analytically continued past $Re(s) = \sigma_0$ to a region including the point $s = \sigma_0$.*

Proof. Recall that the Dirichlet series $D_f(s)$ is uniformly convergent for $Re(s) > \sigma_0$, so we can differentiate it term by term which gives:

$$D'_f(s) = \sum_{n=1}^{\infty} -\frac{f(n)\log n}{n^s};\; D''_f(s) = \sum_{n=1}^{\infty} \frac{f(n)\log^2(n)}{n^s}; \ldots;\; D_f^{(k)}(s) = \sum_{n=1}^{\infty} (-1)^k \frac{f(n)\log^k(n)}{n^s}.$$

Since $D_f(s)$ is holomorphic in an open disc around $\sigma = \sigma_0 + 1$ we can get a Taylor series expansion of $D_f(s)$ with center σ. Then

$$F(s) = \sum_{k=0}^{\infty} \frac{D_f^{(k)}(\sigma)}{k!}(s-\sigma)^k = \sum_{k=0}^{\infty}\sum_{n=1}^{\infty} \frac{(-1)^k}{k!}\frac{f(n)\log^k(n)}{n^\sigma}(s-\sigma)^k = \sum_{k=0}^{\infty}\sum_{n=1}^{\infty} \frac{(\sigma-s)^k}{k!}\frac{f(n)\log^k(n)}{n^\sigma}$$

is a power series with $D_f(s) = F(s)$ for each point of the disc centered in σ with radius one. Assume now that the power series $F(s)$ is convergent past the point $Re(s) = \sigma_0$, that it is has a radius of convergence bigger than one. Then it also converges for some $s < \sigma_0$ on the real axis and since we have that $f(n) \geq 0$ for all n we get that it is even absolutely convergent and so all permutations of the coefficients are permitted. With that we get that

$$F(s) = \sum_{n=1} \frac{f(n)}{n^\sigma}\sum_{k=0}^{\infty} \frac{((\sigma-s)\log n)^k}{k!} = \sum_{n=1}^{\infty} \frac{f(n)}{n^\sigma}e^{(\sigma-s)\log n} = \sum_{n=1}^{\infty} \frac{f(n)}{n^\sigma}n^{\sigma-s} = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} = D_f(s).$$

But with this our σ_0 would not be the convergence abscissa of $D_f(s)$ which gives a contradiction. Therefore $F(s)$ can not be convergent past $Re(s) = \sigma_0$ and so $D_f(s)$ is not analytic to the left of $Re(s) = \sigma_0$, including σ_0. □

Claim 2.7.1. *The Riemann ζ-function satisfies the integral formula*

$$\zeta(s) = s\int_1^\infty [x]x^{-s-1}dx \text{ for } Re(s) > 1.$$

Furthermore it satisfies

$$\zeta(s) = \frac{1}{s-1} + 1 - s\int_1^\infty \frac{\{x\}}{x^{s+1}}dx \text{ for } Re(s) > 0$$

and therefore can be analytically continued to a holomorphic function in $Re(s) > 0$ with a single pole at $s = 1$ with residue 1.

Proof. For the first integral formula we have

$$\begin{aligned}\zeta(s) &= \sum_{n=1}^{\infty}\frac{1}{n^s} = \sum_{n=1}^{\infty}\frac{n-(n-1)}{n^s} = \sum_{n=1}^{\infty}\frac{n}{n^s} - \sum_{n=1}^{\infty}\frac{n-1}{n^s} = \sum_{n=1}^{\infty}\frac{n}{n^s} - \sum_{n=0}^{\infty}\frac{n}{(n+1)^s} = \\ &= \sum_{n=1}^{\infty}\frac{n}{n^s} - \sum_{n=1}^{\infty}\frac{n}{(n+1)^s} = \sum_{n=1}^{\infty} n(n^{-s}-(n+1)^{-s}) = \sum_{n=1}^{\infty} n(-\int_n^{n+1}(x^{-s})'dx) = \\ &= \sum_{n=1}^{\infty} n\ s\int_n^{n+1} x^{-s-1}dx = s\sum_{n=1}^{\infty} n\int_n^{n+1} x^{-s-1}dx = s\sum_{n=1}^{\infty}\int_n^{n+1}[x]x^{-s-1}dx = \\ &= s\int_1^\infty [x]x^{-s-1}dx.\end{aligned}$$

This works since for $x \in [n, n+1)$ the integer part $[x] = n$. For the second part consider the integral

$$\begin{aligned}&s\int_1^\infty x\ x^{-s-1}dx = s\int_1^\infty x^{-s}dx = \frac{s}{-s+1}\int_1^\infty (-s+1)x^{-s}dx = -\frac{s}{s-1}\int_1^\infty (x^{-s+1})'dx = \\ &= -\frac{s}{s-1}\left(\frac{1}{p^{s-1}} - \frac{1}{1^{s-1}}\right)\Big|_{p\to\infty} = \frac{s}{s-1}\end{aligned}$$

for $Re(s) > 1$. So we see that $\zeta(s) - s\int_1^\infty x\, x^{-s-1}dx = \zeta(s) - 1 - \frac{1}{s-1}$ as well as

$$\zeta(s) - s\int_1^\infty x\, x^{-s-1}dx = s\int_1^\infty ([x]-x)x^{-s-1}dx = -s\int_1^\infty \frac{\{x\}}{x^{s+1}}dx.$$

The integral $\int_1^\infty \frac{\{x\}}{x^{s+1}}dx$ is absolutely convergent for $Re(s) > 0$ and $\zeta(s) = 1 + \frac{1}{s-1} - s\int_1^\infty \frac{\{x\}}{x^{s+1}}dx$ is holomorphic for $Re(s) > 0$ with an exception at $s = 1$ because the integral is holomorphic there. Clearly we have a simple pole at $s = 1$ with residue $a_{-1} = 1$. □

Remark 2.7.2. *The integral formula in the above claim is closely connected to the* ***Mellin transform*** *of* $[x]$

$$\{\mathcal{M}f\}(s) = \int_0^\infty f(x)\, x^{s-1}dx.$$

Since $\int_0^1 [x]x^{-s-1}dx = 0$ *(the integer part is always* 0*) we see that*

$$\zeta(s) = s\, \{\mathcal{M}[\ \cdot\ ]\}(-s) \quad \textit{for } Re(s) > 1.$$

Claim 2.7.2. *If* $f : \mathbb{N} \to \mathbb{C}$ *and* $g(x) = \sum_{n\le x} f(n)$ *are functions satisfying* $\frac{g(x)}{x^s} \to 0$ *for* $Re(s)$ *big enough and* x *going to infinity and* $\int_1^\infty g(x)\, x^{-s-1}dx$ *is convergent, then*

$$D_f(s) = s\int_1^\infty g(x)\, x^{-s-1}dx.$$

In particular for the ψ***-function of Chebyshev*** $\psi(x) = \sum_{n\le x}\Lambda(n)$ *we get the integral formula*

$$-\frac{\zeta'(s)}{\zeta(s)} = s\int_1^\infty \frac{\psi(x)}{s^{s+1}}dx \qquad \textit{for } Re(s) > 1.$$

Proof. Using Theorem 2.2.3 we have

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

If s is such that $\frac{g(x)}{x^s} \to 0$ as x grows and the integral is converging we get the integral formula for the Dirichlet series.
Recall that $\sum_{n=1}^\infty \frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)}$. Now we have $\psi(x) = \sum_{n\le x}\Lambda(n) = \sum_{p^\nu\le x}\log(p) \le x\log x$ and

$$|\frac{\psi(x)}{x^s}| \le \frac{x\log x}{x^\sigma} = \frac{\log x}{x^{\sigma-1}} \le \frac{1}{x^{\sigma-1-\epsilon}} \qquad \text{for all } \epsilon > 0 \text{ and } x \text{ large enough.}$$

If $\sigma > 1$, $\sigma - 1 - \epsilon > 0$, $|\frac{\psi(x)}{x^s}| \to 0$ as x grows and $|\frac{\psi(t)}{t^{s+1}}| \le \frac{1}{t^{\sigma-\epsilon}}$ then the integral $\int_1^\infty \frac{\psi(t)}{t^{s+1}}dt$ is convergent. □

Chapter 3

The Prime Number Theorem

Claim 3.0.1. *There exists some constant $c > 0$ such that for the prime counting function $\pi(x) = \sum_{p \leq x} 1$ we get $\pi(x) \geq c \log(\log x)$.*

Proof. First note that the claim also follows immediately by Euclid's theorem since $lim_{x \to \infty} \pi(x) = \infty$.
Observe that for small primes we see that $p_1 = 2 \leq 2^{2^0}$, $p_2 = 3 \leq 2^{2^1}$, $p_3 = 5 \leq 2^{2^2}$ and so on. Assume now that $p_n \leq 2^{2^{n-1}}$ holds and consider $N = p_1 p_2 \cdots p_n - 1$. We see that $\gcd(N, p_i) = 1$ for every $i = 1, ..., n$ and so the first prime p which can divide N is at least as big as p_{n+1}, that is

$$p_{n+1} \leq p \leq N < p_1 \cdots p_n \leq 2^{2^0} 2^{2^1} \cdots 2^{2^{n-1}} = 2^{2^0 + 2^1 + \cdots 2^{n-1}} = 2^{\frac{2^n - 1}{2-1}} < 2^{2^n}.$$

Thus we have proved by induction that $p_{n+1} < 2^{2^n}$ holds for every $n \in \mathbb{N}$. Let now $m \in \mathbb{N}$ be maximal such that $2^{2^m} \leq x$ holds, that is $2^{2^m} \leq x < 2^{2^{m+1}}$.
Then we have that $p_{m+1} \leq 2^{2^m} \leq x$ and so $\pi(x) \geq m + 1$. On the other hand we see that $\log(\log x) < (m+1)\log(2) + \log(\log(2))$ and so $\pi(x) \geq m + 1 > \frac{\log(\log x)}{\log(2)} - \frac{\log(\log(2))}{\log(2)} \geq c \log(\log x)$. □

Claim 3.0.2. *There exists some constant $c > 0$ such that $\pi(x) \geq c \log x$.*

Proof. The idea of this proof goes back to Dressler and Erdös.
Think of all square-free integers $n \leq x$ and their unique representation of the form $n = \prod_{i=1}^{\pi(x)} p_i^{\nu_i}$ where $\nu_i \in \{0, 1\}$ and $p_1, p_2, ..., p_{\pi(x)}$ are all primes up to x. Therefore the number of square-free integers up to x is at most $2^{\pi(x)}$.
On the other hand we have already showed that $\sum_{n \leq x} \mu(n) = \frac{x}{\zeta(2)} + \mathcal{O}(\sqrt{x})$ where $\mu(n)$ is one if n is square-free and zero else. Therefore there exists some $c' > 0$ such that $2^{\pi(x)} > c'x$ and so $\pi(x)\log(2) > \log(c') + \log x$. □

The true growth rate of $\pi(x)$, that is $\pi(x) \sim \frac{x}{\log x}$, was already predicted by Legendre and Gauss.

Conjecture of Gauss: $\pi(x) \sim Li(x) = \int_2^x \frac{1}{\log(t)} dt$ where Li is the so called offset logarithmic integral.

Remark 3.0.1. *For the logarithmic integral we get the asymptotics $Li(x) \sim \frac{x}{\log x}$.*

Proof. By applying L'Hôpital's rule we get

$$lim_{x \to \infty} \frac{Li(x)}{\frac{x}{\log x}} = lim_{x \to \infty} \frac{\frac{1}{\log x}}{\frac{1}{\log x} - \frac{1}{(\log x)^2}} = lim_{x \to \infty} \frac{1}{1 - \frac{1}{\log x}} = 1.$$

Note that $Li(x) \to \infty$ since $\log(t) < t^\alpha$. □

Remark 3.0.2. *We have that $\pi(x) \sim \frac{x}{\log x}$ holds if and only if for every $\epsilon > 0$ and some $x \geq N$ the inequalities*

$$(1-\epsilon)\frac{x}{\log x} \leq \pi(x) \leq (1+\epsilon)\frac{x}{\log x}$$

hold.

Let us note some historical efforts and results connected to the prime number theorem. Around 1850 Chebyshev showed $0.92\frac{x}{\log x} \leq \pi(x) \leq 1.11\frac{x}{\log x}$. In 1874 Mertens managed to show $\sum_{p\leq x}\frac{1}{p} \sim \log(\log x)$ and in 1896 de la Vallée Poussin - Hadamard got $\pi(x) \sim \frac{x}{\log x}$. Around 1980 Newman had a more elegant proof of the prime number theorem using a Tauberian Theorem, we will see Newman's proof later on.

Theorem 3.0.1 (Chebyshev)**.** *There exist two positive constants c_1 and c_2 such that for large enough x we get*

$$c_1\frac{x}{\log x} \leq \pi(x) \leq c_2\frac{x}{\log x}.$$

In order to prove this theorem we need the following lemma.

Lemma 3.0.1. *The following estimates hold for the von Mangoldt function:*

i) $\sum_{n\leq x}\Lambda(n)\left[\frac{x}{n}\right] = x\log x - x + \mathcal{O}(\log x)$

ii) $\sum_{n\leq x}\Lambda(n)\left(\left[\frac{x}{n}\right] - 2\left[\frac{x}{2n}\right]\right) = x\log(2) + \mathcal{O}(\log x)$.

Proof.

i) Recall that we have already seen $\sum_{d|n}\Lambda(d) = \log n$ and consider $\sum_{n\leq x}\log n$.

$$\sum_{n\leq x}\log n = \sum_{n\leq x}\sum_{d|n}\Lambda(d) = \sum_{d\leq x}\sum_{dk=n\leq x}\Lambda(d) = \sum_{d\leq x}\Lambda(d)\sum_{k\leq\frac{x}{d}}1 = \sum_{d\leq x}\Lambda(d)\left[\frac{x}{d}\right].$$

Now we estimate $\sum_{n\leq x}\log n$ via Euler summation formula, thus note that $f(x) = \log x$ is a monotone increasing function and so

$$\sum_{n\leq x}f(n) = \int_1^x f(t)dt + \mathcal{O}(f(x)).$$

$$\begin{aligned}\sum_{n\leq x}\log n &= \int_1^x \log(t)dt + \mathcal{O}(\log x) = \int_1^x [(t\log(t))' - 1]dt + \mathcal{O}(\log x) = \\ &= x\log x - 1\log(1) - (x-1) + \mathcal{O}(\log x) = x\log x - x + \mathcal{O}(\log x)\end{aligned}$$

Now we have two representations of the sum and the statement follows.

ii)

$$\begin{aligned}\sum_{n\leq x}\Lambda(n)\left(\left[\frac{x}{n}\right] - 2\left[\frac{x}{2n}\right]\right) &= \sum_{n\leq x}\Lambda(n)\left[\frac{x}{n}\right] - 2\sum_{n\leq x}\Lambda(n)\left[\frac{x}{2n}\right] = \\ &= \sum_{n\leq x}\Lambda(n)\left[\frac{x}{n}\right] - 2\sum_{n\leq\frac{x}{2}}\Lambda(n)\left[\frac{x/2}{n}\right] - 2\sum_{\frac{x}{2}<n\leq x}\Lambda(n)\left[\frac{x/2}{n}\right] = \\ &\overset{i)}{=} x\log x - x + \mathcal{O}(\log x) - 2\left(\frac{x}{2}\log\left(\frac{x}{2}\right) - \frac{x}{2} + \mathcal{O}(\log x)\right) = \\ &= x\log x - x + \mathcal{O}(\log x) - x(\log x - \log(2)) + x + \mathcal{O}(\log x) = \\ &= x\log(2) + \mathcal{O}(\log x).\end{aligned}$$

□

Now we can prove Theorem 3.0.1.

Proof. We will explore the function $\pi(x)$ through analysis of $\psi(x) = \sum_{n\leq x} \Lambda(x)$. Therefore observe that $\pi(x) = \sum_{p\leq x} 1$ and $\psi(x) = \sum_{n\leq x} \Lambda(n) = \sum_{p^\alpha \leq x} \log(p)$.

$$\sum_{n\leq x} \Lambda(n) = \sum_{p^\alpha\leq x} \log(p) = \sum_{p\leq x} \sum_{\alpha\leq \frac{\log x}{\log(p)}} \log(p) = \sum_{p\leq x} \log(p) \sum_{\alpha\leq\left[\frac{\log x}{\log(p)}\right]} 1 =$$
$$= \sum_{p\leq x} \log(p) \left[\frac{\log x}{\log(p)}\right] \leq \sum_{p\leq x} \log x = \log x\pi(x).$$

Note now that $[x] > x-1$ holds for every x and so $[a] - 2\left[\frac{a}{2}\right] < a - 2(\frac{a}{2}-1) = 2$ and $[a] - 2\left[\frac{a}{2}\right] \in \mathbb{Z}$ so we get $[a] - 2\left[\frac{a}{2}\right] \leq 1$. Therefore we get

$$\sum_{n\leq x} \Lambda(n)\cdot 1 \geq \sum_{n\leq x} \Lambda(n) \left(\left[\frac{x}{n}\right] - 2\left[\frac{x}{2n}\right]\right) \overset{Lemma\ 2.1\ ii)}{=} x\log(2) + \mathcal{O}(\log x).$$

This gives now $x\log(2) + \mathcal{O}(\log x) \leq \sum_{n\leq x} \Lambda(n) \leq \pi(x)\log x$ and so there exists some $c_1 > 0$ such that $c_1 \frac{x}{\log x} \leq \pi(x)$ holds for x large enough.
To get the upper bound of $\pi(x)$ we will use the dyadic partition of the interval $[1,x]$ in subintervals of the form $\left[\frac{x}{2^k}, \frac{x}{2^{k-1}}\right)$, that is $[1,x] = [1, \frac{x}{2^k}) \cup \left[\frac{x}{2^k}, \frac{x}{2^{k-1}}\right) \cup \cdots \cup \left[\frac{x}{2}, x\right]$. Where $k > 0$ is chosen such that $2^k \leq x < 2^{k+1}$. Then we get a telescopic sum of the form

$$\pi(x)\log x = \sum_{i=0}^{k} \left(\pi\left(\frac{x}{2^k}\right)\log\left(\frac{x}{2^k}\right) - \pi\left(\frac{x}{2^{k+1}}\right)\log\left(\frac{x}{2^{k+1}}\right)\right),$$

since $\pi(\frac{x}{2^{k+1}}) = 0$.
Consider now the difference

$$\pi(x)\log x - \pi\left(\frac{x}{2}\right)\log\left(\frac{x}{2}\right) = \log\left(\frac{x}{2}\right)\left(\pi(x) - \pi\left(\frac{x}{2}\right)\right) + \pi(x)\log(2) =$$
$$= \log\left(\frac{x}{2}\right)\left(\pi(x) - \pi\left(\frac{x}{2}\right)\right) + \mathcal{O}(x) = \log\left(\frac{x}{2}\right) \sum_{\frac{x}{2}<p\leq x} 1 + \mathcal{O}(x) = \mathcal{O}(\sum_{\frac{x}{2}<p\leq x} \log(p) + x) =$$
$$= \mathcal{O}(\sum_{\frac{x}{2}<n\leq x} \Lambda(n) + x) = \mathcal{O}\left(\sum_{\frac{x}{2}<n\leq x} \Lambda(n)\left(\left[\frac{x}{n}\right] - 2\left[\frac{x}{2n}\right]\right) + x\right) =$$
$$= \mathcal{O}\left(\sum_{n\leq x} \Lambda(n)\left(\left[\frac{x}{n}\right] - 2\left[\frac{x}{2n}\right]\right) + x\right) \overset{Lemma\ 2.1}{=} \mathcal{O}(x).$$

Here we use the fact that $[\frac{x}{n}] \geq 1$ and $[\frac{x}{2n}] = 0$ for $\frac{x}{2} < n \leq x$ and that $[a] - 2[\frac{a}{2}] > a - 1 - 2\frac{a}{2} = -1$ and so we can add the summands for $n < \frac{x}{2}$ to the error term.
Now since $\pi(x)\log x - \pi\left(\frac{x}{2}\right)\log\left(\frac{x}{2}\right) = \mathcal{O}(x)$ we get that $\pi\left(\frac{x}{2^k}\right)\log\left(\frac{x}{2^k}\right) - \pi\left(\frac{x}{2^{k+1}}\right)\log\left(\frac{x}{2^{k+1}}\right) = \mathcal{O}\left(\frac{x}{2^k}\right)$ and so

$$\pi(x)\log x = \sum_{i=0}^{k} \mathcal{O}\left(\frac{x}{2^k}\right) = \mathcal{O}\left(x\sum_{i=0}^{k} \frac{1}{2^k}\right) = \mathcal{O}(x)$$

because $\sum_{i=0}^{\infty} \frac{1}{2^k} < \infty$.
Therefore $\pi(x) = \mathcal{O}\left(\frac{x}{\log x}\right)$. □

Claim 3.0.3. *The statement* $\pi(x) \sim \frac{x}{\log x}$ *is equivalent to* $\psi(x) \sim x$.

Proof. We have to show that $\lim_{x\to\infty} \frac{\psi(x)}{x} = \lim_{x\to\infty} \pi(x)\frac{\log x}{x}$. In the proof of Chebyshev's theorem we have already seen $\psi(x) = \sum_{n\le x} \Lambda(n) \le \pi(x)\log x$, that is $\frac{\psi(x)}{x} \le \pi(x)\frac{\log x}{x}$. For any $1 < y < x$ we have

$$\pi(x) = \pi(y) + \sum_{y<p\le x} 1 \le \pi(y) + \sum_{y<p\le x} \frac{\log(p)}{\log(y)} \overset{Chebyshev}{\le} c_2\frac{y}{\log(y)} + \frac{\psi(x)}{\log(y)}.$$

Now multiplying both sides with $\frac{\log x}{x}$ and setting $y = \frac{x}{\log x}$ gives

$$\pi(x)\frac{\log x}{x} \le c_2\frac{1}{\log(y)} + \frac{\psi(x)}{x}\frac{\log x}{\log x - \log(\log x)} = c_2\frac{1}{\log x - \log(\log x)} + \frac{\psi(x)}{x}\frac{1}{1-\frac{\log(\log x)}{\log x}}.$$

All put together yields

$$\frac{\psi(x)}{x} \le \pi(x)\frac{\log x}{x} \le c_2\frac{1}{\log x - \log(\log x)} + \frac{\psi(x)}{x}\frac{1}{1-\frac{\log(\log x)}{\log x}}$$

and we see that $RHS \to \lim_{x\to\infty}\frac{\psi(x)}{x}$.
So we see $\lim_{x\to\infty}\frac{\psi(x)}{x} \le \lim_{x\to\infty}\pi(x)\frac{\log x}{x} \le \lim_{x\to\infty}\frac{\psi(x)}{x}$ and equality holds. Therefore we get the statement. □

Theorem 3.0.2 (Mertens, 1874)**.** *The following asymptotic approximations hold:*

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

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

Remark 3.0.3. *These claims are one step closer to the Prime Number Theorem after Chebyshev's Theorem, they are asymptotics* $\sum_{p^k\le x}\frac{\log(p)}{p} \sim \log x$ *whereas the PNT claims* $\sum_{p\le x}\log(p) \sim x$.

Proof.

i) From Lemma 3.0.1 we have an asymptotic formula for $\sum_{n\le x}\Lambda(n)\left[\frac{x}{n}\right]$ which is similar to $\sum_{n\le x}\log(p)\frac{x}{p}$ so we try to use it here. Consider

$$x\log x - x + \mathcal{O}(\log x) = \sum_{n\le x}\Lambda(n)\left[\frac{x}{n}\right] = \sum_{p\le x}\log(p)\left[\frac{x}{p}\right] + \sum_{\substack{p^\nu\le x\\ \nu\ge 2}}\log(p)\left[\frac{x}{p^\nu}\right] =$$

$$= \sum_{p\le x}\log(p)\frac{x}{p} - \sum_{p\le x}\left\{\frac{x}{p}\right\} + \mathcal{O}\left(\sum_{\substack{p^\nu\le x\\ \nu\ge 2}}\log(p)\frac{x}{p^\nu}\right) =$$

$$= x\sum_{p\le x}\frac{\log(p)}{p} + \mathcal{O}\left(\sum_{p\le x}\log(p)\right) + \mathcal{O}\left(x\sum_{n=1}^{\infty}\frac{\log n}{n^2}\right) =$$

$$= x\sum_{p\le x}\frac{\log(p)}{p} + \mathcal{O}\left(\log x\sum_{p\le x}1\right) + \mathcal{O}(x) = x\sum_{p\le x}\frac{\log(p)}{p} + \mathcal{O}\left(\log x c_2\frac{x}{\log x}\right) + \mathcal{O}(x) =$$

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

In the last step we used Theorem 3.0.1 and we get

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

ii) Here we are going to use $i)$ and Theorem 2.2.3.

$$\sum_{p\le x} \frac{1}{p} = \sum_{p\le x} \frac{\log(p)}{p}\frac{1}{\log(p)} = \frac{1}{\log x}\sum_{p\le x} \frac{\log(p)}{p} + \int_2^x \left(\sum_{p\le t} \frac{\log(p)}{p}\right)\frac{dt}{t(\log(t))^2} =$$

$$\overset{i)}{=} 1 + \mathcal{O}\left(\frac{1}{\log x}\right) + \int_2^x \left(\log(t) + \underbrace{\left(\sum_{p\le t}\frac{\log(p)}{p} - \log(t)\right)}_{\omega(t)}\right)\frac{dt}{t\log^2(t)} =$$

$$= 1 + \mathcal{O}\left(\frac{1}{\log x}\right) + \int_2^x \frac{dt}{t\log(t)} + \int_2^x \frac{\omega(t)t}{t\log^2(t)}.$$

Note that $(\log(f(t))' = \frac{1}{f(t)}f'(t)$ so we have $\frac{1}{\log(t)}\frac{1}{t} = \frac{1}{\log(t)}(\log(t))' = (\log(\log(t)))'$. Moreover by $i)$ we know that $\omega(t) = \mathcal{O}(1)$ for any t and so $\int_2^\infty \frac{\omega(t)dt}{t\log^2(t)} < \infty$. Therefore we get

$$\sum_{p\le x}\frac{1}{p} = \log(\log x) + \left(1 - \log(\log(2)) + \int_2^\infty \omega(t)\frac{dt}{t\log^2(t)}\right) + \mathcal{O}\left(\frac{1}{\log x} + \int_x^\infty |\omega(t)|\frac{dt}{t\log^2(t)}\right) =$$

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

□

3.1 Newman's proof of the Prime Number Theorem

Theorem 3.1.1 (Prime Number Theorem)**.** *The asymptotic equivalence $\psi(x) \sim x$ holds.*

The main steps in the proof of Newman are:

I. $\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}$ for $Re(s) > 1$. (Euler)

II. $\zeta(s) - \frac{1}{s-1}$ extends holomorphically to $Re(s) > 0$. (Riemann)

III. $\psi(x) = \mathcal{O}(x)$. (Chebyshev)

IV. $\zeta(s) \neq 0$ for $Re(s) \geq 1$ (Mertens) and $-\frac{\zeta'(s)}{\zeta(s)} - \frac{1}{s-1}$ is holomorphic for $Re(s) \geq 1$.

V. $\int_1^\infty \frac{\psi(x)-x}{x^2}dx$ is a convergent integral. (Newman)

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

We will first discuss the ideas of Newman, that is we will formulate an Analytic Theorem of Newman, its Corollary, which will assure that V holds for more general functions and then the implication $V \Rightarrow VI$. Then we will specialize the argument for the function $\psi(x)$, which satisfies V because of I-IV. We only need to prove IV (and the statements of Newman) as I-III are already shown (more or less) due to previous lectures.

Theorem 3.1.2 (Analytic Theorem of Newman)**.** *Let $F(t)$ be a bounded complex-valued function $F : (0, \infty) \to \mathbb{C}$ which is integrable over every compact subset of $(0, \infty)$. Suppose that the Laplace transform of $F(t)$, given by $G(z) = \int_0^\infty F(t)e^{-zt}dt$ for $Re(z) > 0$, extends holomorphically to $Re(z) \geq 0$. Then the improper integral $\int_0^\infty F(t)dt$ converges and equals $G(0)$.*

Corollary 3.1.1. *Let $f(x)$ be a monotone non-decreasing function defined for $x \geq 1$ satisfying $f(x) = \mathcal{O}(x)$. Consider the Mellin transform of $f(x)$, given by $g(s) = s\int_1^\infty f(x)x^{-s-1}dx$ for $Re(s) > 1$, and assume that $g(s) - \frac{c}{s-1}$ is a holomorphic function in a region containing the closed half-plane $Re(s) \geq 1$. Then we have the asymptotic expression $f(x) \sim cx$.*

Proof. Define the function $F(t) := e^{-t}f(e^t) - c$. Then we see that $F(t)$ is bounded for $t \in (0, \infty)$ since $f(e^t) \leq ke^t$ and thus it is also integrable on every bounded subinterval of $(0, \infty)$, that is $\int_a^b F(t)dt$ exists for all $0 < a \leq b < \infty$. Consider now the Laplace transform of $F(t)$

$$G(z) = \int_0^\infty F(t)e^{-zt}dt = \int_0^\infty (e^{-t}f(e^t) - c)e^{-zt}dt \overset{e^t=x}{=} \int_1^\infty (x^{-1}f(x) - c)x^{-z}d\log x =$$
$$= \int_1^\infty (x^{-1}f(x) - c)x^{-z-1}dx = \int_1^\infty f(x)x^{-z-2}dx - c\int_1^\infty x^{-z-1}dx =$$
$$= \int_1^\infty f(x)x^{-z-2}dx - \frac{c}{z} = \frac{1}{z+1}g(z+1) - \frac{c}{z} = \frac{1}{z+1}\left(g(z+1) - \frac{c(z+1)}{z}\right) =$$
$$= \frac{1}{z+1}\left(g(z+1) - \frac{c}{z} - c\right).$$

Thus $G(z)$ is well defined for $Re(z) > 0$ and by assumption $g(z+1) - \frac{c}{z}$ is holomorphic for $Re(Z) \geq 0$. By the Analytic Theorem of Newman, for $t = \log x$, we have

$$G(0) = \int_0^\infty F(t)dt = \int_0^\infty (e^{-t}f(e^t) - c)t = \int_1^\infty \frac{f(x) - cx}{x^2}dx$$

that is that the integral $\int_1^\infty \frac{f(x)-cx}{x^2}$ converges. (Point V)
Assume now that $\limsup_{x\to\infty} \frac{f(x)}{x} > c$, that is for infinitely many arbitrary large x we have $\frac{f(x)}{x} > \lambda c$ for some $\lambda > 1$. Then take these infinitely many x and consider the integral

$$\int_x^{\lambda x} \frac{f(t) - ct}{t^2}dt \geq \int_x^{\lambda x} \frac{\lambda cx - ct}{t^2}dt = c\int_x^{\lambda x} \frac{\lambda x - t}{t^2}dt = c\int_x^{\lambda x} \frac{x^2(\lambda - \frac{t}{x})}{t^2}d\frac{t}{x} \overset{\frac{t}{\lambda}\to t}{=}$$
$$= c\int_1^\lambda \frac{\lambda - t}{t^2}dt > 0.$$

This works since $\lambda > 1$ and thus contradicts the integral $\int_1^\infty \frac{f(t)-ct}{t^2}$ being convergent, because for any $\epsilon > 0$ and μ_1, μ_2 large enough we should get $\left|\int_{\mu_1}^{\mu_2} \frac{f(t)-ct}{t^2}dt\right| < \epsilon$. Hence

$\limsup_{x\to\infty} \frac{f(x)}{x} \le c$.
Assume now that $\liminf_{x\to\infty} \frac{f(x)}{x} < c$ then there exists some $\mu < 1$ such that for infinitely many arbitrary big x we have $f(x) < \mu c x$. Then

$$\int_{\mu x}^{x} \frac{f(t)-ct}{t^2} dt < \int_{\mu x}^{x} \frac{f(x)-ct}{t^2} dt < \int_{\mu x}^{x} \frac{\mu c x - ct}{t^2} dt = c\int_{\mu x}^{x} \frac{x^2(\mu - \frac{t}{x})}{t^2} d\frac{t}{x} =$$
$$= c\int_{\mu}^{1} \frac{\mu - t}{t^2} dt < 0$$

Again this works since $\mu < 1$ and contradicts that the integral should be convergent, therefore $\liminf_{x\to\infty} \frac{f(x)}{x} \ge c$. So we get that $\lim_{x\to\infty} \frac{f(x)}{x} = c$ and so $f(x) \sim cx$. □

Theorem 3.1.3 (Prime Number Theorem)**.** *The asymptotic equivalence $\psi(x) \sim x$ holds.*

Proof. We have already shown that $\pi(x) \sim \frac{x}{\log x} \Leftrightarrow \psi(x) \sim x$. By Claim 2.6.1 we know that for $Re(s) > 1$ we have $-\frac{\zeta'(s)}{\zeta(s)} = s\int_1^\infty \frac{\psi(x)}{x^{s+1}}dx$, that is the Mellin transform of $\psi(x)$ is $-\frac{\zeta'(s)}{\zeta(s)} = D_\Lambda(s)$. Further we have that $\psi(x) = \sum_{n\le x}\Lambda(n) \le \sum_{p\le x}\log(p) \le \log x \sum_{p\le x} 1 = \log x \pi(x) \le \log x c_2 \frac{x}{\log x} \le c_2 x$ so $\psi(x)$ is non-decreasing and $\psi(x) = \mathcal{O}(x)$. So by the Corollary 3.1.1 it is enough to show that $-\frac{\zeta'(s)}{\zeta(s)}$ is holomorphically extendable past $Re(s) = 1$, that is $-\frac{\zeta'(s)}{\zeta(s)} - \frac{1}{s-1}$ is holomorphic in some region D containing the closed half plane $Re(s) \ge 1$. Recall that we have already seen (in Claim 2.7.1) that for $Re(s) > 0$ we have $\zeta(s) = \frac{1}{s-1} + 1 - s\int_1^\infty \frac{\{x\}}{x^{s+1}}dx$ and so $\zeta(s)$ is analytically continued from $Re(s) > 1$ to $Re(s) > 0$ with a simple pole at $s = 1$ with residue 1.
Therefore we have that $-\frac{\zeta'(s)}{\zeta(s)} - \frac{1}{s-1}$ is holomorphic, with a single exception at $s = 1$, as long as we can guarantee that $\zeta(s) \ne 0$.
Let $C = \{s \in \mathbb{C} : \ |s-1| < \delta\}$ and assume that in C we have $\zeta(s) = \frac{1}{s-1}(1 + h(s))$ for some analytic function h. Now actually $h(s) = (s-1)\left(1 - s\int_1^\infty \frac{\{x\}}{x^{s+1}}dx\right)$ is holomorphic for $Re(s) > 0$ and also bounded on compacts. Choose $\delta > 0$ small enough such that $|h(s)| < 1$ holds.We get

$$\zeta'(s) = -\frac{1}{(s-1)^2}(1+h(s)) + \frac{h'(s)}{s-1}$$

and so
$$-\frac{\zeta'(s)}{\zeta(s)} = \frac{s-1}{1+h(s)}\left(\frac{1+h(s)}{(s-1)^2} - \frac{h'(s)}{s-1}\right) = \frac{1}{s-1} - \frac{h'(s)}{1+h(s)}.$$

By the choice of the disk C with radius δ we have $1 + h(s) \ne 0$ in C and so $-\frac{\zeta'(s)}{\zeta(s)} - \frac{1}{s-1} = -\frac{h'(s)}{1+h(s)}$ is holomorphic in C aswell.
In order to apply Corollary 3.1.1 we need to assure that $-\frac{\zeta'(s)}{\zeta(s)} - \frac{1}{s-1}$ is holomorphic in some region D containing $Re(s) \ge 1$. For $Re(s) > 1$ we have the Euler product representation $\zeta(s) = \prod_p (1 - \frac{1}{p^s})^{-1}$ as in Claim 2.4.4 for $f(n) = 1$. Note that each of the factors $(1 - \frac{1}{p^s})^{-1} = \frac{p^s}{p^s - 1} > 1$, so we have that $\zeta(s) \ne 0$ for $Re(s) > 1$.
What is left to show is the same claim for $Re(s) = 1$ and $s \ne 1$. This will be the next Lemma.

Lemma 3.1.1 (Mertens)**.** *For $Re(s) = 1$ and $s \ne 1$ we get $\zeta(s) \ne 0$.*

Proof. First note that $3 + 4\cos(\varphi) + \cos(2\varphi) \geq 0$. Indeed using $\cos(2\alpha) = 2\cos^2(\alpha) - 1$ we get

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

If for $t \neq 0$ we have that $\zeta(1 + it) = 0$, then $\Theta(s) = \zeta(s)^3\zeta(s + it)^4\zeta(s + 2it)$ possesses zero at $s = 1$, because $\zeta(s)^3$ has a 3-gold pole which is prevailed by the 4-fold zero of $\zeta(s + it)$. Also $\Theta(s)$ is holomorphic around $s = 1$. Therefore $\lim_{s\to 1} \log(|\Theta(s)| = -\infty$. Recall that for some $x \in \mathbb{C}$ we have $|x| = \left|e^{Re(\log x)+iIm(\log x)}\right| = \left|e^{Re(\log x)}\right|$ and so $\log |x| = Re(\log x)$. Let us now approach $s = 1$ from the right along the real axis, therefore let $\sigma > 1$. Thus we know that we have an Euler product for the ζ function. Now we get

$$\log |\zeta(\sigma + it)| = Re(\log(\zeta(\sigma + it)) = Re(\log(\prod_p (1 - p^{-\sigma-it})^{-1})) = -Re(\sum_p \log(1 - p^{-\sigma-it})) =$$
$$= Re(\sum_p p^{-\sigma-it} + \frac{1}{2}(p^2)^{-\sigma-it} + \frac{1}{3}(p^3)^{-\sigma-it} + \cdots) = Re(\sum_n a_n n^{-\sigma-it})$$

for some non negative $a_n \geq 0$. Here we use the series expansion of $\log(1-x) = -\sum_{k=1}^{\infty} \frac{x^k}{k}$ for $|x| < 1$ and $|p^{-\sigma}| < 1$. Then we get

$$\log |\Theta(\sigma)| = 3\log |\zeta(\sigma)| + 4\log |\zeta(\sigma + it)| + \log |\zeta(\sigma + 2it)| =$$
$$= 3Re(\sum_n a_n n^{-\sigma}) + 4Re(\sum_n a_n n^{-\sigma-it}) + Re(\sum_n a_n n^{-\sigma-2it}) =$$
$$= Re(\sum_n a_n n^{-\sigma}(3 + 4n^[-it + n^{-2it})) = \sum_n a_n n^{-\sigma}(3 + 4\cos(t\log n) + \cos(2t\log n)).$$

The last step works because $n^{ix} = \cos(\log nx)+i\sin(\log nx)$. If we now use our observation we see that $\log |\Theta(\sigma)| \geq 0$ which contradicts $\lim_{s\to 1} |\Theta(s)| = -\infty$. Therefore we get $\zeta(1 + it) \neq 0$ for any $t \neq 0$. □

With this lemma we have everything we needed to show and the proof of the Prime Number Theorem is complete. □

What is left to show is that the Analytic Theorem of Newman holds. Therefore first recall the formulation of the Theorem:

Theorem 3.1.4 (Analytic Theorem of Newman)**.** *Let $F(t)$ be a bounded complex-valued function $F : (0, \infty) \to \mathbb{C}$ which is integrable over every compact subset of $(0, \infty)$. Suppose that the Laplace transform of $F(t)$, given by $G(z) = \int_0^\infty F(t)e^{-zt}dt$ for $Re(z) > 0$, extends holomorphically to $Re(z) \geq 0$. Then the improper integral $\int_0^\infty F(t)dt$ converges and equals $G(0)$.*

Proof. Without loss of generality we can assume that $|F(t)| \leq 1$ for all $t > 0$, since else we can just look at $F_1(t) = \frac{F(t)}{B}$.
For every $\lambda > 0$ define

$$G_\lambda(z) = \int_0^\lambda F(t)e^{-zt}dt,$$

which is holomorphic for every $z \in \mathbb{C}$ because $F(t)$ is compactly integrable and e^{-zt} is holomorphic itself. Now it is enough to show that

$$\lim_{\lambda\to\infty} G_\lambda(0) = \lim_{\lambda\to\infty} \int_0^\lambda F(t)dt = G(0),$$

that is $G(0) - G_\lambda(0)$ gets arbitrarily small for λ large enough.
Recall Cauchy's integral formula for simple closed, positively oriented curve γ and f some holomorphic function in and on γ (that is, f is holomorphic in an open subset U such that $\gamma \in U$). Then for any point a inside the area surrounded by γ we have

$$f(a) = \frac{1}{2\pi i} \oint_\gamma \frac{f(z)}{z-a} dz.$$

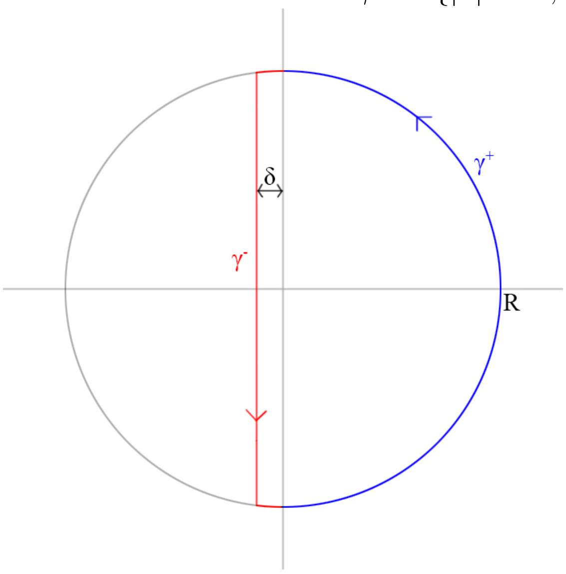
In particular if γ is a curve around $a = 0$ then we have

$$f(0) = \frac{1}{2\pi i} \int_\gamma \frac{f(z)}{z} dz.$$

Take R to be large enough and γ to be the boundary of the region $\{z \in \mathbb{C} : \ |z| \leq R, \ Re(z) \geq -\delta\}$ where δ is chosen small enough such that $G(z)$ is holomorphic on γ. Then by Cauchy's integral formula we get

$$\begin{aligned} G(0) - G_\lambda(0) &= \frac{1}{2\pi i} \int_\gamma (G(z) - G_\lambda(z)) e^{\lambda z} (1 + \frac{z^2}{R^2}) \frac{dz}{z} = \\ &= \underbrace{\int_{\gamma^+} \frac{1}{2\pi i} (G(z) - G_\lambda(z)) e^{\lambda z} (1 + \frac{z^2}{R^2}) \frac{dz}{z}}_{I_1} - \underbrace{\int_{\gamma^-} \frac{1}{2\pi i} G_\lambda(z) e^{\lambda z} (1 + \frac{z^2}{R^2}) \frac{dz}{z}}_{I_2} + \\ &+ \underbrace{\int_{\gamma^-} G(z) e^{\lambda z} (1 + \frac{z^2}{R^2}) \frac{dz}{z}}_{I_3}. \end{aligned}$$

Here we use the notation $\gamma^+ = \{|z| = R, \ Re(z) > 0\}$ and $\gamma^- = \gamma \backslash \gamma^+$.



On the semicircle γ^+ we have

$$\begin{aligned} |G(z) - g_\lambda(z)| &= \left| \int_0^\infty F(t) e^{-zt} dt - \int_0^\lambda F(t) e^{-zt} dt \right| = \left| \int_\lambda^\infty F(t) e^{-zt} dt \right| \leq \\ &\leq \int_\lambda^\infty \left| e^{-zt} \right| dt = \int_\lambda^\infty e^{-xt} dt = \frac{1}{x} e^{-x\lambda} \end{aligned}$$

for $x = Re(z)$. Note that on the curve γ^+ we have $z\bar{z} = R^2$ and so

$$(a + \frac{z^2}{R^2})\frac{1}{z} = \frac{1}{z} + \frac{z}{R^2} = \frac{\bar{z}}{z\bar{z}} + \frac{z}{z\bar{z}} = \frac{\bar{z}+z}{R^2} = \frac{2x}{R^2}.$$

Therefore we get

$$|I_1| = \left|\frac{1}{2\pi i}\int_{\gamma^+}(G(z) - G_\lambda(z))e^{\lambda z}(1 + \frac{z^2}{R^2})\frac{1}{z}dz\right| \leq \frac{1}{2\pi}\int_{\gamma^+}\frac{1}{x}e^{-\lambda x}e^{\lambda x}\frac{2x}{R^2}dz =$$
$$= \frac{1}{\pi R^2}\int_{\gamma^+}dz = \frac{1}{\pi R^2}\pi R = \frac{1}{R}.$$

For estimating I_2 notice that $G_\lambda(z)$ is entire, that is holomorphic on all of $\mathbb{C}$, so by Cauchy's Theorem we can change the path of integration by looking at $\overline{\gamma_1} = \{z \in \mathbb{C} : |z| = R,\ Re(z) = 0\}$. Using again $|F(t)| \leq 1$ we get

$$|G_\lambda(z)| = \left|\int_0^\lambda F(t)e^{-zt}dt\right| \leq \int_0^\lambda |e^{-zt}|dt = \int_0^\lambda e^{-xt}dt <$$
$$< -\frac{1}{x}(e^{-x\lambda} - 1) < \frac{e^{-x\lambda}}{|x|}$$

since on $\overline{\gamma_1}$ we have $x < 0$. This gives

$$|I_2| = \left|\int_{\bar{\gamma_1}}\frac{1}{2\pi i}G_\lambda(z)e^{\lambda z}(1 + \frac{z^2}{R^2})\frac{dz}{z}\right| \leq \frac{1}{2\pi}\int_{\bar{\gamma_1}}|G_\lambda(z)|e^{\lambda x}\frac{2|x|}{R^2}dz <$$
$$< \frac{1}{2\pi}\int_{\bar{\gamma_1}}\frac{e^{-x\lambda}}{|x|}e^{\lambda x}\frac{2|x|}{R^2}dz = \frac{1}{\pi R^2}\int_{\bar{\gamma_1}}dz = \frac{1}{\pi R^2}\pi R = \frac{1}{R}.$$

Finally notice that in the estimate for I_3 we have the function $G(z)(1 + \frac{z^2}{R^2})\frac{1}{z}$ which does not depend on λ, it is holomorphic on γ^-, so it is bounded on the curve. Thus on γ^- we have

$$\left|G(z)(1 + \frac{z^2}{R^2})\frac{1}{z}\right| \leq K$$

for some $K = K(R, \delta) > 0$. Then we get

$$I_3 \leq \frac{1}{2\pi}\left|K\int_{\gamma^-}e^{\lambda z}dz\right| \leq \frac{K}{2\pi}\int_{\gamma^-}e^{\lambda x}dz.$$

Now for $\lambda \to \infty$ we have $e^{\lambda x} \to 0$ rapidly when $x < 0$ and uniformly on x, therefore we can assume that on γ^- we have $|e^{\lambda x}| < \frac{2\epsilon}{KR}$. Then this gives

$$|I_3| \leq \frac{K}{2\pi}\frac{2\epsilon}{KR}\pi R = \epsilon.$$

Thus for any $\epsilon > 0$ we get $|G(0) - G_\lambda(0)| < \epsilon + \frac{2}{R}$. Choose now $R > \frac{2}{\epsilon}$ then this yields $|G(0) - G_\lambda(0)| < 2\epsilon$ for large enough λ and so $\lim_{\lambda\to\infty} G_\lambda(0) = G(0)$. □

The following two Claims are corollaries of Corollary 3.1.1 of Newman.

Claim 3.1.1. *Let $f(x)$ be an arithmetic function $f : \mathbb{N} \to \mathbb{R}^+$ and consider the partial sum $P_f(x) = \sum_{n\leq x} f(n) = \mathcal{O}(x)$. Let the Dirichlet sereis $D_f(s) = \sum_{n=1}^\infty \frac{f(n)}{n^s}$ be holomorphic for $Re(s) > 1$ and let $D_f(s) - \frac{c}{s-1}$, for fixed c, be holomorphic in some region containing the closed half-plane $Re(s) \geq 1$. Then we have the asymptotic equivalence $P_f(x) \sim cx$.*

Proof. Recall that the Abel summation formula gives

$$\sum_{n=1}^{x} \frac{f(n)}{n^s} = \frac{P_f(x)}{x^s} - \int_1^x P_f(t)(t^{-s})' dt = \frac{P_f(x)}{x^s} + s\int_1^x \frac{P_f(t)}{t^{s+1}} dt.$$

If $P_f(x) = \mathcal{O}(x)$ and $Re(s) > 1$ then $\frac{P_f(x)}{x^s} \to_{x\to\infty} 0$ and the integral $\int_1^\infty \frac{P_f(t)}{t^{s+1}} dt$ converges. Thus we have an integral representation of the Dirichlet series for the arithmetic function $f(x)$, $D_f(s) = s\int_1^\infty \frac{P_f(x)}{x^{s+1}} dx$ which is also the Mellin transform of $P_f(x)$.
Then $D_f(s)$ plays the role of $g(s)$ from Corollary 3.1.1, $P_f(x) = \mathcal{O}(x)$ is monotone non-decreasing, and its Mellin transform can be analytically extended past $Re(s) \geq 1$, with the exception at $s = 1$, where it has a simple pole with residue c. Now with Corollary 3.1.1 it follows that $P_f(x) \sim cx$. □

Claim 3.1.2. *Let $f(n), g(n)$ be arithmetic functions such that $f : \mathbb{N} \to \mathbb{R}^+$ and $g(n) = \mathcal{O}(f(n))$, $P_f(x) = \sum_{n\leq x} f(n) = \mathcal{O}(x)$. If the Dirichlet series $D_f(s), D_g(s)$ are holomorphic for $Re(s) > 1$ and there exist constants c, γ such that $D_f - \frac{c}{s-1}$ and $D_g - \frac{\gamma}{s-1}$ are holomorphic for $Re(s) \geq 1$ then $P_g(x) = \sum_{n\leq x} g(n) \sim \gamma x$.*

Proof. For the proof we have to consider the cases whether g is real valued or g is complex valued. Let $K > 0$ such that $|g(n)| \leq K|f(n)| = Kf(n)$ for all $n \in \mathbb{N}$.
Case 1: g is real valued
Then we know that $Kf(n) + g(n) \geq 0$ holds for all $n \in \mathbb{N}$ and we consider the function $h(n) = Kf(n) + g(n)$. Obviously $h(n) \geq 0$ and $P_h(x) = KP_f(x) + P_g(x)$. Since $|P_g(x)| \leq \sum_{n\leq x} |g(n)| \leq K\sum_{n\leq x} f(n) = K\mathcal{O}(x)$ we have $P_g(x) = \mathcal{O}(x)$. By assumption we have $P_f(x) = \mathcal{O}(x)$ as well so we get $P_h(x) = \mathcal{O}(x)$.
Also $D_h(s) = KD_f(s) + D_g(s)$ so $D_h(S) - K\frac{c}{s-1} - \frac{\gamma}{s-1}$ is holomorphic for $Re(s) \geq 1$.
Now from Claim 3.1.1 we get $P_h(x) \sim (Kc + \gamma)x$. Since $P_f(x) \sim cx$, also from the above claim, we get that $P_g(x) = P_h(x) - KP_f(x) \sim \gamma x$.
Case 2: g is complex valued
Let $x = a + ib$ then we have $|x| = \sqrt{x\bar{x}} = \sqrt{a^2 + b^2}$ and so $max(|a|, |b|) \leq \sqrt{a^2 + b^2} = |x|$. Recall also that $x + \bar{x} = 2a$ and $x - \bar{x} = 2ib$.
Let us write $g(n) = g_1(n) + ig_2(n)$ where $g_1(n) = Re(g(n))$ and $g_2(n) = Im(g(n))$ with $g_1, g_2 : \mathbb{N} \to \mathbb{R}$. Note that

$$D_{g_1}(s) = \frac{1}{2}(D_g(s) + G_{\bar{g}}(s)) \quad D_{g_2}(s) = \frac{1}{2i}(D_g(s) - D_{\bar{g}}(s)).$$

Write $G(s) = D_g(s) - \frac{\gamma}{s-1}$ it is holomorphic for $Re(s) \geq 1$, so is $\overline{G(s)} = \overline{D_g(s)} - \frac{\bar{\gamma}}{\bar{s}-1} = D_{\bar{g}}(\bar{s}) - \frac{\bar{\gamma}}{\bar{s}-1}$ and also $\overline{G(\bar{s})} = \overline{D_g(\bar{s})} - \frac{\bar{\gamma}}{s-1} = D_{\bar{g}}(s) - \frac{\bar{\gamma}}{s-1}$. Therefore $\frac{1}{2}(D_g(s) + D_{\bar{g}}(s) - \frac{\gamma+\bar{\gamma}}{s-1})$ is holomorphic in $Re(s) \geq 1$.
We have $max(|g_1(n)|, |g_2(n)|) \leq |g(n)| \leq Kf(n)$ and after Case 1 we get that $P_{g_1}(x) \sim \frac{1}{2}(\gamma + \bar{\gamma})x$.
By analogy $\frac{1}{2i}(D_g(s) - D_{\bar{g}}(s) - \frac{\gamma-\bar{\gamma}}{s-1})$ is holomorphic in $Re(s) \geq 1$ and again by case 1 $P_{g_2}(x) \sim \frac{1}{2i}(\gamma - \bar{\gamma})x$. Hence $P_g(x) = P_{g_1}(x) + iP_{g_2}(x) \sim \gamma x$. □

Corollary 3.1.2. *For the Möbius function $\mu(n)$ and the Liouville function $\lambda(n)$ we have the following $\sum_{n\leq x} \mu(n) = o(x)$ and $\sum_{n\leq x} \lambda(n) = o(x)$.*

Proof. Recall Claim 2.6.1 where we showed that the associated Dirichlet series satisfy $D_\mu(s) = \frac{1}{\zeta(s)}$ and $D_\lambda(s) = \frac{\zeta(2s)}{\zeta(s)}$ for $Re(s) > 1$. Both can be continued analytically past

$Re(s) = 1$, because $\zeta(s) \neq 0$ at $Re(s) = 1$.
Thus in Claim 3.1.2 $g(n) = \mu(n)$ or $g(n) = \lambda(n)$ can be -1 but $|g(n)| \leq 1$ and for $f(n) = 1$ for all $n \in \mathbb{N}$ we have $P_f(x) = x = \mathcal{O}(x)$, $D_f(s) = D_1(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \zeta(s)$ is holomorphic for $Re(s) > 1$ and $\zeta(s) - \frac{1}{s-1}$, $D_g(s) - \frac{0}{s-1}$ are holomorphic for $Re(s) \geq 1$. This means that $\gamma = 0$ and so $\frac{P_\mu(x)}{x} \to_{x \to \infty} 0$, $\frac{P_\lambda(x)}{x} \to_{x \to \infty} 0$. $\square$

Remark 3.1.1. *Actually the two statements above are equivalent to the Prime Number Theorem.*

Chapter 4

Dirichlet's Theorem on primes in arithmetic progressions

Definition 4.0.1. *An* ***Arithmetic progression*** *is a sequence* $\{a + qn| \ \gcd(a, q) = 1, \ n \in \mathbb{N}\}$ *with an initial term* a *and a common difference* q.
For example consider the AP $m \equiv -1 \mod 3$ *that is* $m = -1 + 3k$.

Claim 4.0.1. *There exist infinitely many primes* $p \equiv -1 \mod 3$.

Proof. Assume that there are only finitely many primes $p \equiv -1 \mod 3$ $p_1, p_2, ..., p_N$. Then the number $P = 3p_1p_2 \cdots p_N - 1$ belongs to the same arithmetic Progression but is not divisible by any of $p_1, ..., p_N$ since $\gcd(p_i, P) = 1$ holds for all $i = 1, ...N$. Let $q_1, ..., q_k$ be all prime divisors of P, then at least one of them satisfies $p_i \equiv -1 \mod 3$ since otherwise all $q_j \equiv 1 \mod 3$ and so $P \equiv 1 \mod 3$. But then we see that $p_1, ..., p_N$ can not be all primes which are congruent to $-1 \mod 3$. □

Let us write now

$$\pi(x; a, q) = \sum_{\substack{p \leq x \\ p \equiv a \mod q}} 1 \quad \text{where } \gcd(a, q) = 1.$$

The probability that a prime p is in exactly one congruence class $a + \mathbb{Z}_q$, that is $p \equiv a \mod q$, is $\frac{1}{\varphi(q)}$ because all congruence classes, coprime with q, are $\varphi(q)$. As p is a prime this means that $\gcd(p, q) = 1$, except for $p|q$ which happens only for finitely many primes with density 0. Therefore heuristically we should get

$$\pi(x; a, q) \sim \frac{1}{\varphi(q)} \frac{x}{\log x}.$$

Later we will see this result as the Prime Number Theorem for arithmetic progressions.

4.1 Characters

When considering arithmetic progressions $m \equiv a \mod q$ we are dealing with residue classes, that is m is congruent to an element of the reduced residue group $(\mathbb{Z}/q\mathbb{Z})^*$.

Definition 4.1.1. *Let* G *be a finite abelian group. Then a homomorphism* $\chi : G \to \mathbb{C}^*$ *such that* $\chi(g_1g_2) = \chi(g_1)\chi(g_2)$ *for all* $g_1, g_2 \in G$ *is called a* ***character of the group G***.

Example 4.1.1. *When $G = \langle g \rangle$ is a cyclic group and $\chi : G \to \mathbb{C}^*$ then $\chi(g*1) = \chi(g)\chi(1)$ and therefore $\chi(1) = 1$ or 0. Take the case $\chi(1) = 1$.*
Then if $n = ord_G(g)$ is the order of the element $g \in G$, we have $\chi(g^n) = \chi(1) = \chi(g)^n$ and so $\chi(g) = e^{\frac{2\pi ik}{n}}$ for some $k \in \mathbb{Z}$. If $\chi_k(g) = e^{\frac{2\pi ik}{n}}$ then $k = 0, 1, ..., n-1$ generate exactly n different characters $\chi_0, \chi_1, ..., \chi_{n-1}$.
By the fundamental theorem of finite abelian groups we have $G = \langle g_1 \rangle \times \langle g_2 \rangle \times \cdots \times \langle g_s \rangle$ and each $g \in G$ has a representation of the form $g = g_1^{h_1} \cdots g_s 1^{h_s}$ for $0 \le h_i \le n_i = ord_G(g_i)$. Thus each character can be defined by some s-tuple $(\alpha_1, ..., \alpha_s)$ such that $\chi(g) = \chi(g_1^{h_1} \cdots g_s^{h_s}) = \prod_{i=1}^{s} e^{\frac{2\pi i}{n_i} h_i \alpha_i}$ for $0 \le \alpha_i \le n_i - 1$ and $i = 1, 2, ..., s$. This way we see that there are exactly as many characters as the number of elements of the group.

Definition 4.1.2. *Let $\hat{G}$ be the group of characters via pointwise multiplication, that is $\chi_1, \chi_2 \in \hat{G}$ then $\chi_1\chi_2(g) = \chi_1(g)\chi_2(g)$. Then*

i) χ_0 such that $\chi_0(g) = 1$ for all $g \in G$ is the identity element

ii) $\chi\overline{\chi}(g) = \chi\chi^{-1}(g) = \chi(g)\chi^{-1}(g) = \chi(g)\chi(\bar{g})$ so the complex conjugate of χ is the inverse element.

Claim 4.1.1. *We have $|\hat{G}| = |G|$.*

Claim 4.1.2. *The following orthogonality relations hold:*

i) For every $\chi \in \hat{G}$

$$\sum_{g \in G} \chi(g) = \begin{cases} |G| & , \text{ if } \chi = \chi_0 \\ 0 & , \text{ otherwise.} \end{cases}$$

ii)

$$\sum_{\chi \in \hat{G}} \chi(g) = \begin{cases} |\hat{G}| & , \text{ if } g = 1, \\ 0 & , \text{ otherwise.} \end{cases}$$

Proof. i) The case $\chi = \chi_0$ is obvious, since then we just sum over 1.
Let $\chi \ne \chi_0$. Then there exist $1 \ne g_1 \in G$ such that $\chi(g_1) \ne 1$ and so

$$\chi(g_1) \sum_{g \in G} \chi(g) = \sum_{g \in G} \chi(gg_1) = \sum_{gg_1 \in G} \chi(gg_1) = \sum_{g \in G} \chi(g).$$

Thus $(\chi(g_1) - 1) \sum_{g \in G} \chi(g) = 0$ and since $\chi(g_1) \ne 1$ we get $\sum_{g \in G} \chi(g) = 0$.
ii) The case $g = 1$ is obvious, since the sum is then again over 1.
Let $g \ne 1$, then there exist $\chi_1 \in \hat{G}$ such that $\chi_1(g) \ne 1$ and so

$$\chi_1(g) \sum_{\chi \in \hat{G}} \chi(g) = \sum_{\chi \in \hat{G}} \chi\chi_1(g) = \sum_{\chi \in \hat{G}} \chi(g)$$

and as above we get $\sum_{\chi \in \hat{G}} \chi(g) = 0$.

□

Now take $G = (\mathbb{Z}/q\mathbb{Z})^*$. We will extend the characters of G to arithmetic functions.

Definition 4.1.3. *$\chi : G \to \mathbb{C}^*$ is a* ***Dirichlet character modulo q*** *if $\chi : G \to \mathbb{C}^*$ is a character of $G = (\mathbb{Z}/q\mathbb{Z})^*$, and for all $n \in \mathbb{N}$ we have*

$$\chi(n) = \begin{cases} 0 & \text{if } \gcd(n, q) > 1 \\ \chi(n \mod q) & \text{if } \gcd(n, q) = 1 \end{cases}.$$

4.2 Analytic properties of the Dirichlet L-series

Let $q \geq 2 \in \mathbb{Z}$ and $G = (\mathbb{Z}/q\mathbb{Z})^* = \{\bar{a} : \ \gcd(a, q) = 1\}$. G is a group of order $\varphi(q)$ under multiplication of residue classes. Let $\tilde{\chi} : G \to \mathbb{C}^*$ be any character of the group of the reduced residue system. We can lift it to a map $\chi : \mathbb{Z} \to \mathbb{C}$ called a **Dirichlet character modulo q** by setting

$$\chi(a) = \begin{cases} \tilde{\chi}(\bar{a}) & \text{if } \gcd(a, q) = 1; \\ 0 & \text{otherwise.} \end{cases}$$

The character χ has the following properties:

i) $\chi(1) = 1$

ii) $\chi(ab) = \chi(a)\chi(b)$ for $a, b \in \mathbb{Z}$

iii) $\chi(a) = \chi(b)$ if $a \equiv b \mod q$

iv) $\chi(a) = 0$ if $\gcd(a, q) > 1$.

Let $G(q)$ be the set of characters modulo q. It can be viewed as isomorphic to the group of characters $\hat{G}$ with the following operations: for $\chi_1, \chi_2 \in G(q)$ we define $\chi_1\chi_2(a) = \chi_1(a)\chi_2(a)$ for $a \in \mathbb{Z}$. Unit element is the **principle character modulo q**

$$\chi_0(a) = \begin{cases} 1 & \text{if } \gcd(a, q) = 1 \\ 0 & \text{otherwise.} \end{cases}$$

The inverse of $\chi \in G(q)$ is its complex conjugate $\bar{\chi} : a \mapsto \overline{\chi(a)}$. Note that $\bar{\chi}(a)\chi(a) = \chi^{-1}\chi(a) = \chi_0(a) = 1 = \chi(\bar{a}a) = \chi(\bar{a})\chi(a)$ where $\bar{a}$ is the inverse element to a modulo q, that is $\bar{a}a \equiv q \mod q$.
Then we have orthogonality relations also for the Dirichlet characters modulo q.

Claim 4.2.1. *Let $q \in \mathbb{Z}_{\geq 2}$ and let a run through a complete residue system modulo q. Then*

$$i) \sum_{a \mod q} \chi(a) = \begin{cases} \varphi(q) & \text{if } \chi = \chi_0 \\ 0 & \text{otherwise} \end{cases}$$

$$ii) \sum_{\chi \in G(q) = \hat{G}} \chi(a) = \begin{cases} \varphi(q) & \text{if } a \equiv 1 \mod q \\ 0 & \text{otherwise.} \end{cases}$$

This is Claim 4.1.2 applied for $G = (\mathbb{Z}/q\mathbb{Z})^*$.
Recall that the **Dirichlet L-series** is a Dirichlet series $D_\chi(s)$ where χ is a Dirichlet character modulo q. We denote it by

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

Claim 4.2.2. *If χ is not the principal character modulo q, then the Dirichlet L-series $L(\chi, s)$ is holomorphic on the half-plane $Re(s) > 0$.*
If $\chi = \chi_0$ then $L(\chi_0, s)$ is holomorphic in $Re(s) > 0$ except for the simple pole at $s = 1$ with residue $\frac{\varphi(q)}{q}$, that is $L(\chi_0, s)$ is analytic in $Re(s) > 1$ and $L(\chi_0, s) - \frac{\varphi(q)}{q}\frac{1}{s-1}$ can be analytically continued for $Re(s) > 0$.

Proof. Let $\chi \neq \chi_0$. Then

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

Let us look at the sum $\sum_{n=P}^{Q} \frac{\chi(n)}{n^s}$ and note that $\sum_{n=P+1}^{P+q} = 0$ by the orthogonaltiy relations in Claim 4.1.2 and so $\left|\sum_{n=P}^{Q} \chi(n)\right| \leq q$ for all $P < Q$. Then let $s = Re(s) > 0$ then by 2.2.2 Abel Transformation 2 we have

$$\left|\sum_{n=P}^{Q} \frac{\chi(n)}{n^s}\right| = \left|\frac{1}{(Q+1)^s} \sum_{n=P}^{Q} \chi(n) + \sum_{n=P}^{Q} \left(\frac{1}{n^s} - \frac{1}{(n+1)^s}\right) \sum_{k=P}^{n} \chi(k)\right| \leq$$
$$\frac{q}{(Q+1)^s} + q \sum_{n=P}^{Q} \left(\frac{1}{n^s} - \frac{1}{(n+1)^s}\right) = \frac{q}{(Q+1)^s} + q\left(\frac{1}{p^s} - \frac{1}{(Q+1)^s}\right) = \frac{q}{p^s}$$

When $s > 0$ the last expression tends to zero with $P \to \infty$, that is $L(\chi, s)$ converges. From the properties of the Dirichlet series (see Claim 2.5.1) we get that $L(\chi, s)$ is holomorphic for $Re(s) > 0$.
Let $\chi = \chi_0$. In general χ is strongly multiplicative, so

$$L(\chi, s) = \prod_{p} \frac{1}{1 - \frac{\chi(p)}{p^s}}.$$

Then

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

The Riemann zeta function $\zeta(s)$ has a pole at $s = 1$ but $\zeta(s) - \frac{1}{s-1}$ is analytically continuable for $Re(s) > 0$. Let $h(s) = \prod_{p|q}(1 - \frac{1}{p^s})$, then $h(s)$ is holomorphic and $h(1) = \frac{\varphi(q)}{q}$ by Corollary 2.1. Then

$$L(\chi_0, s) - \frac{\varphi(q)}{q} \frac{1}{s-1} = \zeta(s) h(s) - \frac{h(1)}{s-1} =$$
$$= \left(\zeta(s) - \frac{1}{s-1}\right) h(S) + \frac{h(s)}{s-1} - \frac{h(1)}{s-1} \to_{s \to 1} \left(\zeta(s) - \frac{1}{s-1}\right) h(s)|_{s=1}.$$

Therefore $L(\chi_0, s) - \frac{\varphi(q)}{q} \frac{1}{s-1}$ is holomorphic for $Re(s) > 0$. □

Claim 4.2.3. *At $Re(s) = 1$ we have $L(\chi, s) \neq 0$ for any Dirichlet character $\chi \in \hat{G}$.*

Proof. Let $\chi = \chi_0$, then we have $L(\chi_0, s) = \zeta(s) \prod_{p|q}(1 - \frac{1}{p^s})$ and we know that $\zeta(s) \neq 0$ for $Re(s) = 1$.
Let $\chi \neq \chi_0$ and assume that $L(\chi, 1 + it) = 0$ for some $t \in \mathbb{R}$. We will use the method of Mertens for $\zeta(s) \neq 0$ on $Re(s) = 1$. Therefore define

$$\theta(s) = L(\chi_0, s)^3 L(\chi, s + it)^4 L(\chi^2, s + 2it).$$

Case 1: $\chi^2 \neq \chi_0$.

Thus we have that χ is not a real character or $t \neq 0$. Then we get $L(\chi^2, s+2it) \neq L(\chi_0, s) = \zeta(s)h(s)$ and by the Claim above $L(\chi^2, s+2it)$ is holomorphic for $s = Re(s+2it) > 0$.
Then $\zeta(s)$ has a simple pole at $s=1$, thus $L(\chi_0, s) = \zeta(s)h(s)$ has a simple pole at $s=1$ and since $L(\chi^2, s+2it)$ is holomorphic in an area around $s=1$ we see that $\theta(s)$ has a zero at $s=1$. Therefore $lim_{s\to 1} \log|\theta(s)| = -\infty$.
Recall that for $x \in \mathbb{C}$ we have $\log|x| = Re(\log x)$, so for $\sigma > 1$

$$\log|L(\chi, \sigma+it)| = \log\left|\prod \left(1 - \frac{\chi(p)}{p^{\sigma+it}}\right)^{-1}\right| = -\sum_p \log\left|1 - \frac{\chi(p)}{p^{\sigma+it}}\right| =$$

$$= -\sum_p Re\left(\log\left(1 - \frac{\chi(p)}{p^{\sigma+it}}\right)\right) = \sum_p Re\left[\sum_{n=1}^{\infty} \frac{1}{n}\left(\frac{\chi(p)}{p^{\sigma+it}}\right)^n\right].$$

Here we use the series expansion of $\log(1-x) = \sum_{n=1}^{\infty} -\frac{x^n}{n}$. Note that $\chi(p) = 0$ if $p|q$ and $|\chi(p)| = 1$ otherwise. Write $\chi(p) = e^{i\arg\chi(p)}$ and $p^{ix} = e^{ix\log(p)}$ so

$$\frac{\chi(p)^n}{p^{itn}} = e^{i(n\arg\chi(p) - tn\log(p))}.$$

By de Moivre's formula we get

$$\log|L(\chi, \sigma+it)| = \sum_{p\,\nmid\,q}\sum_{n=1}^{\infty} \frac{\cos(n(\arg\chi(p) - t\log(p)))}{np^{\sigma n}}.$$

Then we get

$$\log|\theta(\sigma)| = \log|L(\chi_0, \sigma)^3 L(\chi, \sigma+it)^4 L(\chi^2, \sigma+2it)| =$$
$$= \sum_{p\,\nmid\,q}\sum_{n=1}^{\infty} \frac{1}{np^{\sigma n}}\left[3\cos(n\arg\chi_0(p)) + 4\cos(n(\arg\chi(p) - t\log p)) + \cos(n(\arg\chi^2(p) - 2t\log p)\right].$$

Note that $\chi_0(p) = 1 = e^0$ and so $\arg\chi_0(p) = 0$, therefore $\cos(n\arg\chi_0(p)) = 1$. Also $\arg\chi^2(p) = 2\arg\chi(p)$. Let $n(\arg\chi(p) - t\log p) = \alpha$ then $n(\arg\chi^2(p) - 2t\log p) = 2n(\arg\chi(p) - t\log p) = 2\alpha$ and we have the factor

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

and so $\log|\theta(\sigma)| \geq 0$ for any $\sigma > 1$ which gives a contradiction.
Case 2: $\chi^2 = \chi_0$
Thus we have $\chi : G \to \mathbb{R}$ such that $\chi(a) = \pm 1$ and $t = 0$. Then $L(\chi^2, s+2it) = L(\chi_0, s)$ has a pole at $s=1$. Consider now the product

$$L(\chi, s)\zeta(s) = D_\chi(s)D_I(s) = D_{\chi * I}(s) = \sum_{n=1}^{\infty} \frac{\sum_{k|n}\chi(k)}{n^s}.$$

If we assume that $L(\chi, 1) = 0$ then as $\zeta(s)$ has only a simple pole at $s=1$ it follows that $L(\chi, s)\zeta(s)$ is holomorphic and therefore convergent for all s in the half-plane $Re(s) > 0$. Recall the notation $S_\chi(n) = \sum_{k|n}\chi(k)$ for the sum function. Since χ is multiplicative, so is $S_\chi = \chi * I$. Also we have $S_\chi(p^\nu) = \sum_{j=0}^{\nu}\chi(p^j) = 1 + \sum_{j=1}^{\nu}\chi(p^j)$. Then since $\chi : G \to \mathbb{R}$ we have three possibilities:

i) $\chi(p) = 0$ and then $S_\chi(p^\nu) = 1$

ii) $\chi(p) = 1$ and then $S_\chi(p^\nu) = 1 + \nu$

iii) $\chi(p) = -1$ and then $S_\chi(p^\nu) = \begin{cases} 0 & \text{if } \nu \text{ is odd} \\ 1 & \text{if } \nu \text{ is even.} \end{cases}$

Then surely $S_\chi(n) \geq 0$ and for $n = k^2$ $S_\chi(n) = S_\chi(k^2) \geq 1$. Consider now $D_{\chi * I}(\frac{1}{2}) = \sum_{n=1}^{\infty} \frac{S_\chi(n)}{\sqrt{n}}$, we have

$$\sum_{n \leq x} \frac{S_\chi(n)}{\sqrt{n}} \geq \sum_{\substack{n \leq x \\ n = k^2}} \frac{1}{\sqrt{n}} = \sum_{k \leq \sqrt{x}} \frac{1}{k} \geq c \log \sqrt{x} \rightarrow_{x \rightarrow \infty} \infty.$$

This means that the Dirichlet series $D_{\chi * I}(s)$ is divergent as $s = \frac{1}{2}$ and therefore the abscissa of convergence is $\sigma_0 \geq \frac{1}{2}$. Recall that $\chi * I(n) = S_\chi(n) \geq 0$ for every $n \in \mathbb{N}$. Then by Landau's Theorem we have $D_{\chi * I}$ is not holomorphic past $s = \sigma_0$ for $Re(s) < \sigma_0$. Therefore $D_{\chi * I}$ should have a singularity in the region $Re(s) \geq \frac{1}{2}$. This contradicts the assumption $L(\chi, 1) = 0$ and so $D_{\chi * I}$ is holomorphic for $Re(s) > 0$. Therefore $L(\chi, s) \neq 0$. □

Claim 4.2.4. *For $Re(s) > 1$ and any Dirichlet character $\chi \in \hat{G}$ we have*

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

Proof. Take

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

and

$$\begin{aligned} -\frac{d}{ds} \log(L(\chi, s)) = \frac{L'(\chi, s)}{L(\chi, s)} &= \frac{d}{ds}\left(-\sum_p \log\left(1 - \frac{\chi(p)}{p^s}\right)\right) = \\ &= \sum_p -\chi(p)(-\log(p)p^{-s}) \frac{1}{1 - \frac{\chi(p)}{p^s}} = \sum_p \frac{\log(p)}{p^s} \frac{\chi(p)}{1 - \frac{\chi(p)}{p^s}} = \\ &= \sum_p \frac{\log(p)}{p^s} \sum_{\nu=0}^{\infty} \frac{\chi(p)^{\nu+1}}{p^{s\nu}} = \sum_p \sum_{k=1}^{\infty} \frac{\chi(p)^k \log(p)}{p^{sk}} = \\ &= \sum_p \sum_{k=1}^{\infty} \frac{\chi(p^k) \log(p)}{p^{sk}} = \sum_{n=1}^{\infty} \frac{\chi(n)\Lambda(n)}{n^s}. \end{aligned}$$

All these operations are allowed in $Re(s) > 1$ because $\log(1 - \frac{\chi(p)}{p^s} = -\frac{\chi(p)}{p^s} + \mathcal{O}(p^{-2\sigma})$ for $\sigma = Re(s)$ and the series $H(s) = -\sum_p \log(1 - \frac{\chi(p)}{p^s})$ is convergent absolutely and uniformly in any compact subset of $\{Re(s) > 1\}$. Thus $H(s)$ is a holomorphic function for $Re(s) > 1$ and $e^{H(s)} = L(\chi, s)$ is a well defined holomorphic function as well. □

We consider the function

$$\pi(x;a,q) = \sum_{\substack{p\le x\\ p\equiv a \mod 1}} 1 \quad \text{where } \gcd(a,q)=1$$

and by a heuristic argument we expect that $\pi(x;a,q) \sim \frac{1}{\varphi(q)}\frac{x}{\log x}$. Recall that for the proof of the Prime Number Theorem we introduced the Chebyshev's function $\psi(x) = \sum_{n\le x}\Lambda(n)$. We showed that the Prime Number Theorem is equivalent to $\psi(x)\sim x$ and then proved that $\psi(x)\sim x$ holds. Consider now the analogous function

$$\psi(x;a,q) = \sum_{\substack{n\le x\\ n\equiv a \mod q}} \Lambda(n) \quad \gcd(a,q)=1.$$

Lemma 4.2.1. *We have the asymptotic equivalence $\psi(x;a,q)\sim \frac{1}{\varphi(q)}x$.*

Proof. First note that the condition $n\equiv a \mod q$ in the definition of the function can be rewritten as $n\bar{a}\equiv 1 \mod q$. Then for $G=(\mathbb{Z}/q\mathbb{Z})^*$ by Claim 4.2.1 part ii) we have $\sum_{\chi\in\hat{G}}\chi(n\bar{a}) = \varphi(q)$ and so $\frac{1}{\varphi(q)}\sum_{\chi\in\hat{G}}\chi(n\bar{a}) = 1$. On the other hand if $a \mod q$ with $\gcd(a,q)=1$ is fixed we have

$$\frac{1}{\varphi(q)}\sum_{\chi\in\hat{G}}\chi(n\bar{a}) = \begin{cases} 1 & \text{if } n\equiv a \mod q\\ 0 & \text{if } n\not\equiv a \mod q.\end{cases}$$

Hence we can write

$$\psi(x;a,q) = \sum_{\substack{n\le x\\ n\equiv a \mod q}}\Lambda(n) = \sum_{n\le x}\Lambda(n)\frac{1}{\varphi(q)}\sum_{\chi\in\hat{G}}\chi(n\bar{a}) = \frac{1}{\varphi(q)}\sum_{\chi\in\hat{G}}\chi(\bar{a})\sum_{n\le x}\Lambda(n)\chi(n).$$

Recall Claim 3.1.2. In our case we have $\Lambda:\mathbb{N}\to\mathbb{R}^+$ and $\psi(s)=\sum_{n\le x}\Lambda(n) = \mathcal{O}(x)$. Since we want to evaluate $\sum_{n\le x}\chi(n)\Lambda(n)$ we consider the function $\chi\Lambda:\mathbb{N}\to\mathbb{C}$ where $\chi(n)\Lambda(n) = \mathcal{O}(\Lambda(n))$ since $|\chi(n)|=1$.
We know that $D_\Lambda(s) = \sum_{n=1}^{\infty}\frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)}$ is holomorphic for $Re(s)>1$ and $-\frac{\zeta'(s)}{\zeta(s)} - \frac{1}{s-1}$ is holomorphic in $Re(s)\ge 1$. In order to apply Claim 3.1.2 we need an analogous analytic continuation of the Dirichlet series $D_{\chi\Lambda}(s)$. In the last Claim we have seen that

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

and further we know that $L(\chi,s)$ is holomorphic for $Re(s)>0$ if $\chi\ne\chi_0$ ad $L(\chi,s)\ne 0$ for $Re(s)=1$ for any $\chi\in\hat{G}$. Therefore for $\chi\ne\chi_0$ we have that $-\frac{L'(\chi,s)}{L(\chi,s)} - \frac{0}{s-1}$ is holomorphic in $Re(s)\ge 1$.
When $\chi=\chi_0$ then $L(\chi_0,s)$ has a simple pole at $s=1$ with residue $\frac{\varphi(q)}{q}$. Then in a small disc $C=\{s\in\mathbb{C}:\ |s-1|<\delta\}$ we have $L(\chi_0,s) = g(s)+\frac{\varphi(q)}{q}\frac{1}{s-1}$ where $g(s)$ is holomorphic in C. Then $L(\chi_0,s) = \frac{1}{s-1}\left(\frac{\varphi(q)}{q}+h(s)\right)$ where $h(s)=(s-1)g(s)$ which is holomorphic in C. Therefore we have $L'(\chi_0,s) = -\frac{1}{(s-1)^2}\left(\frac{\varphi(q)}{q}+h(s)\right)+\frac{h'(s)}{s-1}$ and so

$$-\frac{L(\chi_0,s)}{L(\chi_0,s)} = \frac{s-1}{\frac{\varphi(q)}{q}+h(s)}\left(\frac{1}{(s-1)^2}\left(\frac{\varphi(q)}{q}+h(s)\right) - \frac{h'(s)}{s-1}\right) = \frac{1}{s-1} - \frac{h'(s)}{\frac{\varphi(q)}{q}+h(s)}.$$

We can choose $\delta > 0$ small enough such that $|h(s)| < \frac{\varphi(q)}{q}$ holds on C and hence the second fraction is a holomorphic function in C. Therefore $-\frac{L'(\chi_0,s)}{L(\chi_0,s)} - \frac{1}{s-1}$ is holomorphic for $Re(s) \geq 1$.
From Claim 3.1.2 we get that $\sum_{n\leq x} \chi_0(n)\Lambda(n) \sim x$ and $\sum_{n\leq x} \chi(n)\Lambda(n) = o(x)$ for $\chi \neq \chi_0$. Then

$$\begin{aligned}\psi(x;a,q) &= \frac{1}{\varphi(q)} \left(\sum_{\substack{\chi\in\hat{G} \\ \chi\neq\chi_0}} \chi(\bar{a}) \sum_{n\leq x} \chi(n)\Lambda(n) + \chi_0(\bar{a}) \sum_{n\leq x} \chi_0(n)\Lambda(n) \right) \\ &\sim \frac{1}{\varphi(q)} \left(\sum_{\substack{\chi\in\hat{G} \\ \chi\neq\chi_0}} o(x) + x \right) = o(x) + \frac{x}{\varphi(q)}.\end{aligned}$$

So we have $\psi(x;a,q) \sim \frac{x}{\varphi(q)}$. □

Theorem 4.2.1 (Prime Number Theorem for arithmetic progressions)**.** *For $q \in \mathbb{Z}_{\geq 2}$ and* $\gcd(a,q) = 1$ *we have the asymptotic formula*

$$\pi(x;a,q) \sim \frac{1}{\varphi(q)} \frac{x}{\log x}$$

.

Proof. We apply Newman's method for the Prime Number Theorem. Consider the function

$$\mathcal{L}(x;a,q) = \sum_{\substack{p\leq x \\ p\equiv a \mod q}} \log(p).$$

Then we get

$$\begin{aligned}\pi(x;a,q) &= \sum_{\substack{p\leq x \\ p\equiv a \mod q}} 1 = \sum_{\substack{p\leq x \\ p\equiv a \mod q}} \frac{\log(p)}{\log(p)} \overset{AT3}{=} \frac{1}{\log x}\mathcal{L}(x;a,q) - \int_2^x \mathcal{L}(t;a,q)\frac{-dt}{t(\log(t))^2} = \\ &= \frac{1}{\log x}\mathcal{L}(x;a,q) + \int_2^x \mathcal{L}(t;a,q)\frac{dt}{t(\log(t))^2}.\end{aligned}$$

Now

$$\begin{aligned}\mathcal{L}(x;a,q) &= \sum_{\substack{p\leq x \\ p\equiv a \mod q}} \log(p) = \sum_{\substack{p^n\leq x \\ p^n\equiv a \mod q}} \log(p) - \sum_{n\geq 2} \sum_{\substack{p^n\leq x \\ p^n\equiv a \mod q}} \log(p) = \\ &= \sum_{\substack{n\leq x \\ n\equiv a \mod q}} \Lambda(n) - \sum_{n\geq 2} \sum_{\substack{p^n\leq x \\ p^n\equiv a \mod q}} \log(p) = \psi(x;a,q) - \sum_{n\geq 2} \sum_{\substack{p^n\leq x \\ p^n\equiv a \mod q}} \log(p)\end{aligned}$$

If we now look at the second term we get

$$\begin{aligned}&\sum_{n\geq 2} \sum_{\substack{p^n\leq x \\ p^n\equiv a \mod q}} \log(p) \leq \sum_{n\geq 2} \sum_{p^n\leq x} \log(p) = \mathcal{O}\left(\log x \sum_{\substack{n\geq 2 \\ p^n\leq x}} 1 \right) = \mathcal{O}\left(\log x \sum_{2\leq n\leq \frac{\log x}{\log(2)}} \sum_{p\leq\sqrt{x}} 1 \right) = \\ &= \mathcal{O}(\log x\pi(\sqrt{x})\log x) = \mathcal{O}((\log x)^2 \frac{\sqrt{x}}{\log(\sqrt{x})}) = \mathcal{O}(\sqrt{x}\log x).\end{aligned}$$

Therefore we have $\mathcal{L}(x;a,q) = \psi(x;a,q) + \mathcal{O}(\sqrt{x}\log x)$ and so with the Lemma above $\mathcal{L}(x;a,q) \sim \frac{x}{\varphi(q)}$.
But then we get

$$\pi(x;a,q) \sim \frac{1}{\varphi(q)}\frac{x}{\log x} + \mathcal{O}\left(\int_2^x t\frac{dt}{t(\log(t))^2}\right) = \frac{1}{\varphi(q)}\frac{x}{\log x} + \mathcal{O}\left(\frac{x}{(\log x)^2}\right).$$

The latter follows from $\int_2^x \frac{dt}{(\log(t))^2} = Li(x) - \frac{x}{\log x} = \mathcal{O}\left(\frac{x}{(\log x)^2}\right)$. □

4.3 The error terms in the PNT and the PNT for APs

Note that the proof of Newman, using Tauberian Theorems, does not provide information on the error terms in the asymptotic formula for $\pi(x)$ and $\pi(x;a,q)$.

Theorem 4.3.1. *We have the asymptotics* $\pi(x) = Li(x) + \mathcal{O}\left(\frac{x}{e^{c\sqrt{\log x}}}\right)$ *for some constant* $x > 0$ *uniformly for* $x \geq 2$.

Note that the error terms is better than $\mathcal{O}\left(\frac{x}{(\log x)^2}\right)$ as $e^{c\sqrt{\log x}} > e^{2\log\log x} = (\log x)^2$. for any $c > 0$ and large enough x. On the other side $Li(x) = \int_2^x \frac{dt}{\log(t)} = \frac{x}{\log x} + \mathcal{O}\left(\frac{x}{(\log x)^2}\right)$, so a less precise formulation of the PNT is

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

Riemann Hypothesis (RH) All non trivial zeros of the Riemann ζ function lie on the line $Re(s) = \frac{1}{2}$.

Theorem 4.3.2. *Assume the RH, then for* $x \geq 2$ *we have* $\pi(x) = Li(x) + \mathcal{O}(x^{\frac{1}{2}}\log x)$.

From $\log\left(\frac{\sqrt{x}}{\log x}\right) > x\sqrt{\log x}$ for any $c > 0$ and x large enough we see that the error term under the RH is indeed stronger.

Theorem 4.3.3. $\pi(x;a,q) = \frac{Li(x)}{\varphi(q)} + \mathcal{O}_A\left(\frac{x}{e^{c_1\sqrt{\log x}}}\right)$ *for a given constant* $A > 0$, *where* $q \leq (\log x)^A$, $\gcd(a,q) = 1$ *and a certain constant* $c_1 > 0$.

Generalized Riemann Hypothesis (GRH) For any Dirichlet character χ modulo q and $s \in \mathbb{C}$ we have: if $L(\chi, s) = 0$ and $Re(s) > 0$ then $Re(s) = \frac{1}{2}$.
Remark that for $\chi = \chi_0$ we get the normal Riemann Hypothesis.

Theorem 4.3.4. *Let* q *be given and assume the GRH for all* L *functions modulo* q. *Then if* $\gcd(a,q) = 1$ *and* $x \geq 2$ *we get* $\pi(x;a,q) = \frac{Li(x)}{\varphi(q)} + \mathcal{O}(x^{\frac{1}{2}}\log x)$.

Chapter 5

The circle method and the ternary Goldbach's problem

Prelude to the circle method

For a general complex valued sequence $\{a_n\}_{n=0}^m$ we would like to demonstrate some asymptotic relation $a_n \sim F(n)$ for some function $F(n)$. We can take the power series generating function $f(z) = \sum_{n=0}^{\infty} a_n z^n$ with the assumption that the radius of convergence is $0 < r < 1$. By Cauchy's residue theorem we can express the members of the sequence $\{a_n\}$ by

$$I_n = 2\pi i a_n = \oint \frac{f(z)}{z^{n+1}} dz, \text{ for all } n \geq 0,$$

where the contour integral is taken over the circle with center 0 traversed once in anti clockwise direction.

The goal is to push the circle to $r = 1$ by having some insight on the singularities of $f(z)$ on $|z| = 1$. It turns out that in certain situations the roots of unity $\xi_s^r = e^{\frac{2\pi i r}{s}}$ with"small denominators", that is $s < N_0$ for some $N_0 \in \mathbb{N}$, give the "major contribution".

Remark 5.0.1. *Recall that the residues of the Dirichlet series $D_\Lambda(s)$ and $D_{\chi\Lambda}(s)$ at $s = 1$ played a major role for finding the main terms in the asymptotic formulas for $\sum_{n\leq x} \Lambda(n)$ and $\sum_{n\leq x} \chi(n)\Lambda(n)$ and thus in solving the PNT and the PNT for AP.*

One constructs the set of the **major arcs** $\mathfrak{M}$, which are arcs with centers ξ_s^r with small s and lengths chosen in such a way that two different arcs do not intersect. The complement of $\mathfrak{M}$ on the unit circle in $\mathbb{C}$ is then called the set of the **minor arcs** $\mathfrak{m}$, that is $\mathfrak{m} = S^1 \backslash \mathfrak{M}$.

Then, since we have chosen the major arcs so that they do not intersect each other, we aim at

$$I_n = I_{n,\mathfrak{M}} + I_{n,\mathfrak{m}} = \text{Main term} + o(MT).$$

Hardy-Littlewood (1920-1930): developed the circle method in connection to Waring's problem and the binary and ternary Goldbach's problem.

Vinogradov, 1937: modified the circle method by introducing finite trigonometric sums instead of power series generating functions.

The modification of Vinogradov, where instead of the unit circle, one takes an interval with length one, is the most established version by now. Let $e(z) = e^{2\pi iz}$ for $z \in \mathbb{R}$. We have that $e^{2\pi i} = \cos(2\pi) + i\sin(2\pi) = 1$ and so $e(z+1) = e(z)$, thus $e(z)$ is periodic with period 1. Furthermore, for $h \in \mathbb{Z}$ we have

$$\int_0^1 e(\alpha h)d\alpha = \begin{cases} 1, & \text{if } h = 0; \\ 0, & \text{otherwise.} \end{cases}$$

Now consider a partitioning of the unit interval in the following manner. Let us have the appropriately chosen parameters $Q = Q(n)$ and $\tau = \tau(n)$ and set a typical major arc

$$\mathfrak{M}(a,q) = \left\{ \alpha : \left| \alpha - \frac{a}{q} \right| \leq \tau \right\}.$$

The set of the major arcs is then given as the union

$$\mathfrak{M} = \bigcup_{q \leq Q} \bigcup_{\substack{1 \leq a \leq q-1 \\ \gcd(a,q)=1}} \mathfrak{M}(a,q).$$

The set of minor arcs is defined as the complement

$$\mathfrak{m} = (\tau, 1+\tau] \backslash \mathfrak{M}.$$

Then if we want to integrate any periodic function $f(z)$ with period 1 we can write

$$\int_0^1 f(z)dz = \int_\tau^{1+\tau} f(z)dz = \int_{\mathfrak{M}} f(z)dz + \int_{\mathfrak{m}} f(z)dz$$

as long as $\mathfrak{M}(a,q) \cap \mathfrak{M}(a',q') = \emptyset$ for $\frac{a}{q} \neq \frac{a'}{q'}$. The circle method will be illustrated in the treatment of the famous ternary Goldbach's problem.

5.1 The Goldbach's problems

These conjectures were formulated in a letter to Euler in 1742.

Binary Goldbach's problem: Every even number greater than 2 is a sum of two primes.

Ternary Goldbach's problem: Every odd number greater than 5 is a sum of three primes.

The binary version is still open, while the ternary Goldbach's problem was completely solved recently ($\approx$ 2013) by Helfgott. Vinogradov, using the circle method of Hardy-Littlewood, could show not only existence but an asymptotic formula for the number of presentations as a sum of three primes for any odd number $n > N_0$, for N_0 large enough. Let us consider the weighted quantity

$$R(n) = \sum_{\substack{p_1,p_2,p_3\\ p_1+p_2+p_3=n}} \log p_1 \log p_2 \log p_3.$$

Note that $R(n) = 0$ if there is not a triple p_1, p_2, p_3 which satisfies $p_1 + p_2 + p_3 = n$. We have the following asymptotic formula.

Theorem 5.1.1. *Suppose that $A > 0$ is a real constant. Then we get*

$$R(n) = \frac{1}{2}C(n)n^2 + \mathcal{O}\left(\frac{n^2}{\log^A n}\right)$$

where

$$C(n) = \prod_{p \nmid n}\left(1 + \frac{1}{(p-1)^3)}\right)\prod_{p|n}\left(1 - \frac{1}{(p-1)^2}\right)$$

is a positive real constant for which there are absolute constants $0 < c_1 < C(n) < c_2$ for any odd n.

Remark 5.1.1. *The method of Vinogradov ($\approx$ 1939) can produce the constant $N_0 = e^{e^{e^{41,96}}}$ such that any odd $n \geq N_0$ is presentable as sum of three primes. The constant N_0 was reduced many times until Helfgott reduced it to $N_0 = 10^{27}$. For the odd numbers $n < 10^{27}$ a computer verification of GRH yielded the ternary Goldbach's problem for odd $7 \leq n < 10^{27}$.*

Obviously, from Theorem 5.1.1 it follows that there exists a constant N_0 such that any odd $n > N_0$ is a sum of three primes.

5.2 Setting up the circle method

Let us consider the sum

$$S(\alpha) = S(\alpha, n) = \sum_{p\leq n} \log p \cdot e(\alpha p).$$

then

$$\begin{aligned} S(\alpha)^3 &= \sum_{p_1\leq n} \log p_1 e(\alpha p_1) \sum_{p_2\leq n} \log p_2 e(\alpha p_2) \sum_{p_3\leq n} \log p_3 e(\alpha p_3) = \\ &= \sum_{p_1,p_2,p_3\leq n} \log p_1 \log p_2 \log p_3 e\left(\alpha(p_1 + p_2 + p_3)\right). \end{aligned}$$

Recall the identity

$$\int_0^1 e(\alpha h)d\alpha = \begin{cases} 1, & \text{if } h = 0 \\ 0, & \text{otherwise.} \end{cases}$$

Then we have

$$R(n) = \sum_{p_1,p_2,p_3} \log p_1 \log p_2 \log p_3 \int_0^1 e\left(\alpha(p_1 + p_2 + p_3 - n)\right)d\alpha = \int_0^1 S(\alpha)^3 e(-n\alpha)d\alpha.$$

Note that $S(\alpha)$ is periodic with period 1. Now, let for a constant $B > 0$ we choose the parameters

$$Q = (\log n)^B,$$
$$\tau = \frac{a}{n} = \frac{(\log n)^B}{n}.$$

Then

$$\mathfrak{M}(a,q) = \left\{\alpha : \left|\alpha - \frac{a}{q}\right| \leq \frac{Q}{n}\right\}$$

denotes a typical major arc and the set of the major arcs is given by the union

$$\mathfrak{M} = \bigcup_{q \leq Q} \bigcup_{\substack{1 \leq a \leq q \\ \gcd(a,q)=1}} \mathfrak{M}(a,q).$$

Note that any two major arcs are disjoint. Indeed, let their centers be $\frac{a}{q}$ and $\frac{a'}{q'}$; then the distance between them is greater than twice the half-lengths of the arcs $\mathfrak{M}(a,q)$ and $\mathfrak{M}(a',q')$, so

$$\left|\frac{a}{a} - \frac{a'}{q'}\right| = \frac{|aq' - a'q|}{qq'} \geq \frac{1}{qq'}$$

since $\gcd(a,q) = \gcd(a',q') = 1$ and $aq' = a'q$ would imply $\frac{a}{q} = \frac{a'}{q'}$. Then $\frac{1}{qq'} \geq \frac{1}{Q^2} \geq \frac{2Q}{n} = 2\tau$ is a consequence of $n \geq 2Q^3 = 2(\log n)^B$ and thus is true for large enough n. Let us then consider the set of minor arcs

$$\mathfrak{m} = \left(\frac{Q}{n}, 1 + \frac{Q}{n}\right] \setminus \mathfrak{M}.$$

(As an exercise it is left to show that $\mathfrak{M} \subset (\tau, 1+\tau]$.)

We can then write

$$R(n) = \int_0^1 S(\alpha)^3 e(-n\alpha) d\alpha = \int_\tau^{1+\tau} S(\alpha)^3 e(-n\alpha) d\alpha =$$
$$= \int_{\mathfrak{M}} S(\alpha)^3 e(-n\alpha) d\alpha + \int_{\mathfrak{m}} S(\alpha)^3 e(-n\alpha) d\alpha = I_{\mathfrak{M}} + I_{\mathfrak{m}}.$$

We remark that the choice of $Q = (\log n)^B$ is dictated by an application of the Prime Number Theorem for arithmetic progressions, where we have a good error term only if $q \leq (\log n)^B$ holds. In this sense the major arcs are "sparse", hence the minor arcs constitute a larger part of the unit interval and it is harder to give a good upper bound for $I_{\mathfrak{m}}$ compared to other applications of the circle method with a larger parameter Q.

5.3 Treatment of the minor arcs

Our aim is to show that the contribution of $I_{\mathfrak{m}}$ is of smaller magnitude than n^2. Let us introduce the *symbol of Vinogradov* "$\ll$": we write $f(n) \ll g(n)$ if $f(n) = \mathcal{O}(g(n))$. Then the main result of this section is the following.

Theorem 5.3.1. *Let $A > 0$ be a positive constant. Then*

$$\int_{\mathfrak{m}} |S(\alpha)|^3 d\alpha \ll \frac{n^2}{(\log n)^A}.$$

Proof. We will need to adjust the choice of the parameter Q according to A, more precisely we will show soon that we need to have $B \geq 2A + 10$. Clearly we have

$$|I_{\mathfrak{m}}| = \left| \int_{\mathfrak{m}} S(\alpha)^3 e(-n\alpha) d\alpha \right| \leq \int_{\mathfrak{m}} |S(\alpha)|^3 d\alpha.$$

First observe that

$$\begin{aligned}
\int_0^1 |S(\alpha)|^2 d\alpha &= \int_0^1 \left(\sum_{p_1, p_2 \leq n} \log p_1 e(\alpha p_1) \overline{\log p_2 e(\alpha p_2)} \right) d\alpha = \\
&= \int_0^1 \sum_{p_1, p_2 \leq n} \log(p_1 + p_2) e(\alpha(p_1 - p_2)) d\alpha = \\
&= \sum_{p_1, p_2 \leq n} \log(p_1 + p_2) \int_0^1 e(\alpha(p_1 - p_2)) d\alpha = \\
&= \sum_{p \leq n} (\log p)^2 \ll \log n \sum_{p \leq n} \log p \leq \log n \psi(n) \ll n \log n.
\end{aligned}$$

Here we used Chebyshev's Theorem for the fuvntion $\psi(n)$.

The next crucial step is to give an upper bound of $\sup_{\alpha \in \mathfrak{m}} |S(\alpha)|$ of the order $n(\log n)^{4-B/2}$. In order to achieve this we rely on the following claim due to Vinogradov:

Claim 5.3.1 (Vinogradov)**.** *Let $\alpha \in \mathbb{R}$, $\gcd(a, q) = 1$, $q \leq n$, be such that*

$$\left| \alpha - \frac{a}{q} \right| \leq \frac{1}{q^2}.$$

Then we have

$$S(\alpha) \ll (\log n)^4 \left(nq^{-\frac{1}{2}} + n^{\frac{4}{5}} + (qn)^{\frac{1}{2}} \right).$$

Lemma 5.3.1 (Dirichlet's approximation Theorem)**.** *Let $\alpha \in \mathbb{R}$. Then for each real number $x \geq 1$ there exists a rational number $\frac{a}{q}$ with $\gcd(a, q) = 1$, $1 \leq q \leq x$ such that*

$$\left| \alpha - \frac{a}{q} \right| \leq \frac{1}{qx}.$$

Proof. It suffices to prove the inequality for $\gcd(a, q) \geq 1$ without assuming strictly $\gcd(a, q) = 1$. Indeed, if $q' = qs$ for $s > 1$, then $\left| \alpha - \frac{a'}{q'} \right| \leq \frac{1}{q'x} < \frac{1}{qx}$ where $a' = as$.

Let $m = [x]$, then the m numbers $\beta_q = \alpha q - [\alpha q]$ for $q = 1, 2, ..., m$ all lie in $[0, 1)$. Consider the $m + 1$ intervals

$$B_r = \left[\frac{r - 1}{m + 1}, \frac{r}{m + 1} \right) \quad r = 1, 2, ..., m + 1.$$

If $\beta_q \in B_1$ then from $\frac{\beta_q}{q} = \alpha - \frac{[\alpha q]}{q}$ we get

$$\left| \alpha - \frac{[\alpha q]}{q} \right| = \left| \frac{\beta_q}{q} \right| \leq \frac{1}{q(m+1)} < \frac{1}{qx}.$$

If $\beta_q \in B_{m+1}$ then $1 - \beta_q \leq 1/(m+1)$ and so

$$\left| \alpha - \frac{[\alpha q] + 1}{q} \right| = \frac{|1 - \beta_q|}{q} \leq \frac{1}{q(m+1)} < \frac{1}{qx}.$$

In the first case $a = [\alpha q]$, and in the second case $a = [\alpha q] + 1$.

If $\beta_q \notin B_1 \cup B_{m+1}$ then one of the $m-1$ intervals B_r with $2 \leq r \leq m$ contains at least two elements of the β_q (by the pigeonhole principle), say β_u and β_v with $u < v \leq m$. Then

$$\frac{1}{m+1} \geq |\beta_u - \beta_v| = \alpha(v-u) - ([\alpha v] - [\alpha u]).$$

Now we take $q = v - u$ and $a = [\alpha v] - [\alpha u]$. Then we have

$$\frac{1}{m+1} \geq \alpha q - a \quad \text{ and } \frac{1}{xq} \geq \frac{1}{(m+1)q} \geq \alpha - \frac{a}{q}.$$

□

Assuming now that Claim 5.3.1 and Lemma 5.3.1 hold we are able to bound $\sup_{\alpha \in \mathfrak{m}} |S(\alpha)|$. Indeed, we chose

$$x = \frac{n}{(\log n)^B} = \frac{n}{Q} = \frac{1}{\tau}.$$

If $\alpha \in \mathfrak{m}$ there are a and q with $\gcd(a,q) = 1$, $Q < q \leq x$, such that

$$\left| \alpha - \frac{a}{q} \right| \leq \frac{1}{qx} \leq \frac{1}{q^2}.$$

After Claim 5.3.1 we will then have

$$\begin{aligned} S(\alpha) &\ll (\log n)^4 \left(nq^{-\frac{1}{2}} + n^{\frac{4}{5}} + (nq)^{\frac{1}{2}} \right) \leq \\ &\leq (\log n)^4 \left(n(\log n)^{-\frac{B}{2}} + n^{\frac{4}{5}} + n^{\frac{1}{2}} n^{\frac{1}{2}} (\log n)^{-\frac{B}{2}} \right) \ll n(\log n)^{4 - \frac{B}{2}}. \end{aligned}$$

Now from $\int_0^1 |S(\alpha)|^2 d\alpha \ll n \log n$ and the latter estimate we get

$$\int_{\mathfrak{m}} |S(\alpha)|^3 d\alpha \ll \int_0^1 |S(\alpha)|^2 \sup_{\alpha \in \mathfrak{m}} |S(\alpha)| d\alpha \ll n^2 (\log n)^{5 - \frac{B}{2}}.$$

Now take $B \geq 10 + 2A$. This proves Theorem 5.3.1. □

Remember the Claim 5.3.1 in the past proof. In order to prove Vinogradov's Claim 5.3.1 we need some auxiliary Lemmas.

Lemma 5.3.2. *Let $\Theta \in \mathbb{R}$ and $[a,b] \subset [1,N]$. Then we have the estimate*

$$\left|\sum_{m\in[a,b]} e(\Theta m)\right| \ll \min\left\{N, \frac{1}{\|\Theta\|}\right\}$$

where $\|\Theta\| = \min_{z\in\mathbb{Z}} |z-\Theta|$ is the distance to the nearest integer.

Proof. Without loss of generality we can assume that $\Theta \in \left(-\frac{1}{2}, \frac{1}{2}\right]$, since $e(z)$ is periodic with period 1. Then $\|\Theta\| = |\Theta|$.

Case 1: $\Theta = 0$
Then

$$\left|\sum_{m\in[a,b]} e(\Theta m)\right| \le \sum_{m\in[a,b]} 1 \le |[a,b]| \le N.$$

Case 2: $\Theta \neq 0$
Then we have summing of a geometric progression

$$\left|\sum_{m\in[a,b]} e(\Theta m)|\right| = \left|e(a\Theta)\sum_{m\in[0,b-a]} e(\Theta m)\right| = \left|\frac{e(\Theta(b-a+1))-1}{e(\Theta)-1}\right| \le \frac{2}{|e(\Theta)-1|}.$$

Consider now

$$\begin{aligned}|e(\Theta)-1| &= \left|e\left(\frac{\Theta}{2}\right)\right|\left|e\left(\frac{\Theta}{2}\right) - e\left(-\frac{\Theta}{2}\right)\right| = \\ &= |\cos(\pi\Theta) + i\sin(\pi\Theta) - \cos(-\pi\Theta) - i\sin(-\pi\Theta)| = 2|\sin(\pi\Theta)| \gg |\Theta|.\end{aligned}$$

This proves the Lemma. □

Lemma 5.3.3. *Let $L \ge 1$, $n > 1$, $q \ge 1$ be given and $\alpha \in \mathbb{R}$ be such that*

$$\left|\alpha - \frac{a}{q}\right| \le \frac{1}{q^2} \quad \textit{for } \gcd(a,q) = 1.$$

Then we have

$$\sum_{l\le L} \min\left\{\frac{n}{l}, \frac{1}{\|\alpha l\|}\right\} \ll (nq^{-1} + L + q)\log(2Lq).$$

Proof. Let us write $\alpha = \frac{a}{q} + \beta$. Then $\left|\alpha - \frac{a}{q}\right| = |\beta| \le \frac{1}{q^2}$. We also partition the interval $[1,L]$ into parts of length q, that is we write $l = hq + r$ with $1 \le r \le q$ and $0 \le h \le \frac{L}{q}$. Then

$$\alpha l = \left(\frac{a}{q} + \beta\right)(hq + r) = ah + \frac{ra}{q} + hq\beta + r\beta \quad \text{and} \quad \|\alpha l\| = \left\|\frac{ra}{q} + hq\beta + r\beta\right\|.$$

We get

$$U = \sum_{l\le L} \min\left\{\frac{n}{l}, \frac{1}{\|\alpha l\|}\right\} = \sum_{0\le h\le \frac{L}{q}} \sum_{r=1}^{q} \min\left\{\frac{n}{hq+r}, \left\|\frac{ra}{q} + hq\beta + r\beta\right\|^{-1}\right\}.$$

Let us mention that we have $\|x\| = |x|$ when $|x| \le \frac{1}{2}$ and $\Big| \|x\| - \|y\| \Big| \le \|x+y\| \le \|x\| + \|y\|$.

Denote the contribution to the sum U for $h = 0, r \le \frac{q}{2}$ by U_0 so that $U = U_0 + U_1$. Then for $h = 0, 1 \le r \le \frac{q}{2}$ we have $|r\beta| \le \frac{q}{2}\frac{1}{q^2} = \frac{1}{2q}$ and $q \nmid ra$ implies that $\left\|\frac{ra}{q}\right\| \ge \frac{1}{q}$. In this case

$$\left\|\frac{ra}{q} + hq\beta + r\beta\right\| = \left\|\frac{ra}{q} + r\beta\right\| \ge \left| \left\|\frac{ra}{q}\right\| - \|r\beta\| \right| \ge \left\|\frac{ra}{q}\right\| - \frac{1}{2q},$$

and also $\frac{1}{2}\left\|\frac{ra}{q}\right\| \ge \frac{1}{2q}$, so that $\left\|\frac{ra}{q}\right\| - \frac{1}{2q} \ge \frac{1}{2}\left\|\frac{ra}{q}\right\|$.
Then we have

$$U_0 \ll \sum_{r \le \frac{q}{2}} \left(\left\|\frac{ra}{q}\right\| - \frac{1}{2q}\right)^{-1} \ll \sum_{r \le \frac{q}{2}} \left\|\frac{ra}{q}\right\|^{-1} \ll \sum_{\substack{-\frac{q}{2} < m \le \frac{q}{2} \\ m \ne 0}} \sum_{\substack{r \le \frac{q}{2} \\ ra \equiv m \pmod q}} \left\|\frac{ra}{q}\right\|^{-1} =$$

$$= \sum_{\substack{-\frac{q}{2} < m \le \frac{q}{2} \\ m \ne 0}} \left|\frac{m}{q}\right|^{-1} \sum_{\substack{1 \le r \le \frac{q}{2} \\ r \equiv m\bar{a} \pmod q}} 1 \ll q \sum_{1 \le m \le \frac{q}{2}} \frac{1}{m} \ll q \log q.$$

For the remaining summands in U_1 we have $hq + r \gg (h+1)q$. Indeed, if $h = 0$ and $r \ge \frac{q}{2}$ this holds with constant $\frac{1}{2}$. When $h \ge 1$ we have $hq + r \ge \frac{h+1}{2}q$. Then

$$U_1 \ll \sum_{0 \le h \le \frac{L}{q}} \sum_{r=1}^{q} \min\left\{\frac{n}{(h+1)q}, \left\|\frac{ra}{q} + hq\beta + r\beta\right\|^{-1}\right\}.$$

Let us consider any interval I with length $\frac{1}{q}$. Then for a fixed h the relation

$$\frac{ra}{q} + r\beta + hq\beta \pmod 1 \in I \tag{I}$$

has at most $\mathcal{O}(1)$ solutions $r \in [1, q]$. Indeed, if r, r' are two such solutions, then

$$\left\|\frac{ra}{q} + r\beta - \frac{r'a}{q} - r'\beta\right\| \le \frac{1}{q}.$$

If $s = r - r'$ then $0 \le |s| < q$ and $|s\beta| \le \frac{1}{q}$, so

$$\left\|\frac{sa}{q}\right\| - \frac{1}{q} \le \left\|\frac{sa}{q} - s\beta\right\| \le \frac{1}{q}.$$

As this is possible only for $0 \le \left\|\dfrac{sa}{q}\right\| \le \dfrac{2}{q}$ we have finite choices of $s \pmod q$, and thus of r', once r has been fixed.

Now we choose $I = I_s = \left[\dfrac{s}{q}, \dfrac{s+1}{q}\right]$ with $0 \le s \le q-1$. We group the $\mathcal{O}(1)$ r's such that r is a solution of (I_s) together. We note that when $s + 1 \le \frac{q}{2}$ this means that

$\left\|\frac{ra}{q} + hq\beta + r\beta\right\| \in I_s$, i.e. $\left\|\frac{ra}{q} + hq\beta + r\beta\right\| \geq \frac{s}{q}$.

If $s + 1 > \frac{q}{2}$ then for the solution of (I) we get

$$\left\|\frac{ra}{q} + hq\beta + r\beta\right\| \geq 1 - \frac{s+1}{q} = \frac{q-s-1}{q}.$$

Thus we can consider the summation only for $1 \leq s \leq \frac{q}{2}$ and get

$$U_1 \ll \sum_{0 \leq h \leq \frac{L}{q}} \sum_{s=1}^{\frac{q}{2}} \frac{q}{s} + \sum_{0 \leq h \leq \frac{L}{q}} \frac{n}{(h+1)q}$$

where in the last sum, for $s = 0$, we rather bounded by $\frac{n}{(h+1)q}$, also there only $\mathcal{O}(1)$ r's satisfy (I_0).
Then

$$U_1 \ll q \sum_{h \leq \frac{L}{q}} \log q + \frac{n}{q} \sum_{1 \leq h \leq \frac{L}{q}} \frac{1}{h} \ll L\log(q) + \frac{n}{q}\log(L).$$

Putting things together we get

$$U \ll |U_0| + |U_1| \ll q \log q + L \log Q + \frac{n}{q} \log L \ll (nq^{-1} + L + q) \log(2Lq).$$

□

We also need the following Lemma, which is a corollary of Vaughan's identity.

Lemma 5.3.4. *Let $U \geq 1$, $V \geq 1$, such that $UV \leq x$. Then for any arithmetic function f we have the estimate*

$$\sum_{U < n \leq x} f(n)\Lambda(n) \ll (\log x)T_1 + T_2,$$

where

$$T_1 = \sum_{l \leq UV} \max_{w} \left| \sum_{w < k \leq \frac{x}{l}} f(kl) \right|$$

$$T_2 = \left| \sum_{U < m \leq \frac{x}{V}} \sum_{V < k \leq \frac{x}{m}} \Lambda(m) b(k) f(mk) \right|$$

and $b(k)$ denotes an arithmetic function, depending only on V and satisfying $|b(k)| \leq \tau(k)$.

Proof. This is a Corollary of Vaughan's identity, see for example p. 194–196 in Brüdern's "Einführung in die analytische Zahlentheorie" [1] (Satz 6.1.2). □

Now we can finally prove Claim 5.3.1.

Proof. Recall that $S(\alpha) = \sum_{k \le n} \Lambda(k) e(\alpha k)$. Then

$$S(\alpha) = \sum_{k \le U} \Lambda(k) e(\alpha k) + \sum_{U < k \le n} \Lambda(k) e(\alpha k) \ll U + \sum_{U < k \le n} \Lambda(k) e(\alpha k)$$

where we used Chebyshev's Theorem for the first sum. For the second sum we will apply Lemma 5.3.4 with $x = n$, $f(k) = e(\alpha k)$ and $U = V$. We will choose later the parameter $U \le n$ in a suitable way.
Then

$$S(\alpha) \ll U + (\log n) T_1 + T_2,$$

where

$$T_1 = \sum_{l \le U^2} \max_{w} \left| \sum_{w \le k \le \frac{n}{l}} e(\alpha k l) \right|$$

$$T_2 = \left| \sum_{U < m \le \frac{n}{U}} \sum_{U < k \le \frac{n}{m}} \Lambda(m) b(k) e(\alpha k m) \right| .$$

For the inner sum of T_1 we immediately apply Lemma 5.3.2:

$$\left| \sum_{w \le k \le \frac{n}{l}} e(\alpha l k) \right| \ll \min \left\{ \frac{n}{l}, ||\alpha l||^{-1} \right\}$$

and therefore

$$T_1 \ll \sum_{l \le U^2} \min \left\{ \frac{n}{l}, ||\alpha l||^{-1} \right\} .$$

In order to estimate the sum T_2 we first exchange the order of summation. Then from $U \le k \le \frac{n}{m}$ we get $U \le k \le \frac{n}{U}$, since $U < m \le \frac{n}{U}$ and $\frac{n}{m} \le \frac{n}{U}$. Then we split the interval $\left[U, \frac{n}{U}\right]$ into dyadic intervals $K < k \le 2K$ where $K = 2^\nu U$ and $K \le \frac{n}{U}$. Clearly $\log K = \nu \log 2 + \log U$ and $\nu \ll \log n$. Then we have

$$T_2 = \left| \sum_{U < k \le \frac{n}{U}} b(k) \sum_{U < m \le \frac{n}{k}} \Lambda(m) e(\alpha m k) \right| \ll (\log n) \max_{U < K \le \frac{n}{U}} T(K)$$

with

$$T(K) = \left| \sum_{K < k \le 2K} b(k) \sum_{U < m \le \frac{n}{k}} \Lambda(m) e(\alpha m k) \right| .$$

Recall the Cauchy-Schwarz inequality $| \langle x, y \rangle |^2 \le \|x\|^2 \cdot \|y\|^2$. Then

$$T(K)^2 \le \sum_{K < k \le 2K} |b(k)|^2 \sum_{K < k \le 2K} \left| \sum_{U < m \le \frac{n}{k}} \Lambda(m) e(\alpha m k) \right|^2 .$$

By Lemma 5.3.4 we know that $|b(k)| \le \tau(k)$, and we also use without proof the estimate

$$\sum_{k \le z} \tau^2(k) \ll z (\log z)^3.$$

(Recall that in Claim 2.4.2 we showed that $\sum_{k\le z}\tau(k) \ll z\log z$.) Then we get

$$T(K)^2 \ll K(\log K)^3 \sum_{K<k\le 2K}\sum_{U<m_1,m_2\le \frac{n}{k}} \Lambda(m_1)\Lambda(m_2)e(\alpha k(m_1-m_2)).$$

The terms with $m_1=m_2$ give

$$\sum_{K<k\le 2K}\sum_{U<m\le\frac{n}{k}}\Lambda(m)^2 \ll \sum_{K<k\le 2K}\left(\log\frac{n}{k}\right)^2\cdot\frac{n}{k} \ll K(\log n)^2\cdot\frac{n}{K} \ll n(\log n)^2.$$

For the other terms we have

$$\begin{aligned}\sum_{K<k\le 2K}\sum_{\substack{U<m_1,m_2\le\frac{n}{k}\\ m_1\ne m_2}}\Lambda(m_1)\Lambda(m_2)e(\alpha k(m_1-m_2)) \ll\\ \ll \sum_{\substack{U<m_1,m_2\le\frac{n}{K}\\ m_1\ne m_2}}\Lambda(m_1)\Lambda(m_2)\left|\sum_{K<k\le 2K}e(\alpha k(m_1-m_2))\right| \ll\\ \ll (\log n)^2\sum_{U<m_2<m_1\le\frac{n}{K}}\min\{K,\|\alpha(m_1-m_2)\|^{-1}\}\end{aligned}$$

where we again used Lemma 5.3.2.

Further, put $l=m_1-m_2$. When $U<m_2<m_1\le\frac{n}{K}$ then $1\le l\le\frac{n}{K}$ and for any l the equation $l=m_1-m_2$ has at most $\frac{n}{K}$ solutions. Then, since $K\le\frac{n}{l}$ we can write the last sum as

$$\sum_{l\le\frac{n}{K}}\min\left\{\frac{n}{l},\|\alpha l\|^{-1}\right\}.$$

Putting together the pieces up to now we have

$$\begin{aligned}T(K)^2 &\ll K(\log K)^3\left(n(\log n)^2+(\log n)^2\sum_{l\le\frac{n}{K}}\min\left\{\frac{n}{l},\|\alpha l\|^{-1}\right\}\right) \ll\\ &\ll Kn(\log n)^5+n(\log n)^5+n(logn)^5\sum_{l\le\frac{n}{K}}\min\left\{\frac{n}{l},\|\alpha l\|^{-1}\right\}.\end{aligned}$$

Thus both T_1 and $T(K)$, i.e T_2, got reduced to estimating a sum of the type treated in Lemma 5.3.3. If we have the same condition for existence of a good rational approximation of α like in Claim 5.3.1, Lemma 5.3.3 then gives

$$\begin{aligned}T(K)^2 &\ll Kn(\log n)^5+n(\log n)^5\log\left(2q\frac{n}{K}\right)\left(nq^{-1}+\frac{n}{K}+q\right) \ll\\ &\ll Kn(\log n)^5+(n^2q^{-1}+n^2K^{-1}+nq)(\log n)^6.\end{aligned}$$

Then

$$T(K) \ll (Kn)^{\frac12}(\log n)^{\frac52}+\left(nq^{-\frac12}+nK^{-\frac12}+(nq)^{-\frac12}\right)(\log n)^3.$$

Using the condition $U\le K\le\frac{n}{U}$ we get $K^{-\frac12}\le U^{-\frac12}$ and $K^{\frac12}\le n^{\frac12}U^{-\frac12}$, so

$$\begin{aligned}T_2 &\ll (\log n)\max_{U<K\le\frac{n}{U}}T(K) \ll (\log n)^4(nU^{-\frac12}+nq^{-\frac12}+nU^{-\frac12}+(nq)^{\frac12}) \ll\\ &\ll (\log n)^4(nU^{-\frac12}+nq^{-\frac12}+(nq)^{\frac12}).\end{aligned}$$

Similarily, we apply Lemma 5.3.3 to the sum T_1 to get

$$T_1 \ll \sum_{l \leq U^2} \min\left\{\frac{n}{l}, \|\alpha l\|^{-1}\right\} \ll \log\left(2U^2 q\right)\left(nq^{-1} + U^2 + q\right).$$

Now we choose $U = n^{\frac{2}{5}}$. Then $U^2 = n^{\frac{4}{5}}$ and $\log(2U^2q) \ll \log n$ because $q \leq n$. Then

$$T_1 \ll (\log n)(nq^{-1} + n^{\frac{4}{5}} + q)$$
$$T_2 \ll (\log n)^4(n^{1-\frac{1}{5}} + nq^{-\frac{1}{2}} + (nq)^{\frac{1}{2}}).$$

Noticing that for $q \leq n$ we have $q \leq (nq)^{\frac{1}{2}}$, we finally get

$$\begin{aligned} S(\alpha) &\ll n^{\frac{2}{5}} + (\log n)^2(nq^{-1} + n^{\frac{4}{5}} + q) + (\log n)^4(n^{\frac{4}{5}} + nq^{-\frac{1}{2}} + (nq)^{\frac{1}{2}}) \\ &\ll (\log n)^4\left(n^{\frac{4}{5}} + nq^{-\frac{1}{2}} + (qn)^{\frac{1}{2}}\right). \end{aligned}$$

□

Let us note here that we defined $S(\alpha) = S(\alpha, n) = \sum_{p \leq n} \log p \cdot e(\alpha p)$ but we proved Claim 5.3.1 for $S^*(\alpha) = \sum_{k \leq n} \Lambda(k) e(\alpha k)$. However, since

$$S^*(\alpha) = S(\alpha) + \sum_{\substack{p^\beta \leq n \\ \beta \geq 2}} \log(p) e(\alpha p^\beta) = S(\alpha) + \Sigma$$

where

$$|\Sigma| \leq \sum_{2 \leq \beta \leq \frac{\log n}{\log 2}} \sum_{p \leq n^{\frac{1}{2}}} \log p \leq \sum_{\beta \leq \frac{\log n}{\log 2}} \sum_{p \leq n^{\frac{1}{2}}} 1 \ll \log n \sum_{\beta \leq \log n} \frac{n^{\frac{1}{2}}}{\log n} \ll n^{\frac{1}{2}} \log n.$$

One of the terms on the RHS of Claim 5.3.1 is of magnitude $(\log n)^4 n^{\frac{4}{5}}$, so having proven $S^*(\alpha) \ll RHS$ gives also $S(\alpha) \ll RHS$.

5.4 Treatment of the major arcs

Now recall that we have already seen some error terms for the Prime number theorem and the Prime number theorem for arithmetic progressions. We will need the following version which is a corollary of the famous Siegel-Walfisz Theorem (Corollary 11.21 in [3]).

Theorem 5.4.1 (Siegel-Walfisz). *Let $A > 0$ and $q \leq (\log x)^A$, $\gcd(a, q) = 1$. Then there exists a constant $c > 0$ such that*

$$\theta(x; a, q) := \sum_{\substack{p \leq x \\ p \equiv a \mod q}} \log p = \frac{x}{\varphi(q)} + \mathcal{O}_A(xe^{-c\sqrt{\log x}}).$$

Definition 5.4.1. *The **Ramanujan's sum** is defined as*

$$c_q(h) = \sum_{\substack{a=1 \\ \gcd(a,q)=1}}^{q} e\left(\frac{ah}{q}\right).$$

The following identity holds.

Lemma 5.4.1. *The Ramanujan's sum satisfies*

$$c_q(h) = \varphi(q)\frac{\mu\left(\frac{q}{\gcd(q,h)}\right)}{\varphi\left(\frac{q}{\gcd(q,h)}\right)}.$$

Proof. First note that for $h \in \mathbb{Z}$ we have

$$\sum_{a=1}^{q} e\left(\frac{ah}{q}\right) = \begin{cases} q, & \text{if } h \equiv 0 \pmod q; \\ 0, & \text{otherwise.} \end{cases}$$

The first case is trivial. In the second case we have a geometric progression

$$e\left(\frac{h}{q}\right)\left(e(0) + e\left(\frac{h}{q}\right) + \cdots + e\left(\frac{(q-1)h}{q}\right)\right) = e\left(\frac{h}{q}\right)\frac{e\left(\frac{qh}{q}\right) - 1}{e\left(\frac{h}{q}\right) - 1}$$

$$= e\left(\frac{h}{q}\right)\frac{e(h) - 1}{e\left(\frac{h}{q}\right) - 1} = 0.$$

We can write

$$\sum_{a=1}^{q} e\left(\frac{ah}{q}\right) = \sum_{d|q}\sum_{\substack{a=1 \\ \gcd(a,q)=d}}^{q} e\left(\frac{ah}{q}\right) = \sum_{d|q}\sum_{\substack{1\le \frac{a}{d} \le \frac{q}{d} \\ \gcd\left(\frac{a}{d},\frac{q}{d}\right)=1}} e\left(\frac{\frac{a}{d}h}{\frac{q}{d}}\right)$$

$$= \sum_{d|q} c_{\frac{q}{d}}(h) = \sum_{d|q} c_d(h) \left(= S_{c_q}(h)\right).$$

Now recall the Möbius inversion formula $f = \mu * S_f$ and also note taht for any $d|q$ we will have $S_{c_d}(h) = d$ if $h \equiv 0 \mod q$, and 0 otherwise. Then

$$c_q(h) = \sum_{\substack{d|q \\ h\equiv 0 \mod q}} \mu\left(\frac{q}{d}\right) S_{c_d}(h) = \sum_{d|\gcd(q,h)} \mu\left(\frac{q}{d}\right) d.$$

In particular, $c_q(h)$ is real-valued and it is multiplicative regarding q, when h is fixed. Also $c_q(h) = c_q(\gcd(a,q))$. Then it is enough to verify the statement of the Lemma only for prime powers, that is

$$c_q(h) = \mu\left(\frac{q}{\gcd(q,h)}\right)\frac{\varphi(q)}{\varphi\left(\frac{q}{\gcd(q,h)}\right)}$$

holds for $q = p^k$ and $k \geq 1$.

Let $p^\beta | h$ but $p^{\beta+1} \nmid h$ (denoted by $p^\beta || h$). Then $\gcd(h,q) = \gcd(p^\beta h_1, p^k) = p^{\min\{\beta,k\}}$.

Case 1: $k \leq \beta$ Then $c_{p^k}(h) = c_{p^k}(p^k)$ and

$$c_{p^k}(p^k) = \sum_{d|p^k} \mu\left(\frac{p^k}{d}\right) d = \varphi(p^k) = \varphi(q).$$

We used that $S_\varphi(n) = n$ and by Möbius inversion $\varphi(n) = \mu * S_\varphi(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) d$. In this case $\frac{q}{\gcd(q,h)} = \frac{p^k}{p^k} = 1$.

Case 2: $k = \beta + 1$ Then $\gcd(q, h) = p^\beta$ and

$$c_{p^{\beta+1}}(h) = \sum_{d|p^\beta} \mu\left(\frac{p^{\beta+1}}{d}\right) d = \mu(p)p^\beta = \mu(p)\frac{\varphi(p^{\beta+1})}{\varphi(p)} = \mu\left(\frac{q}{gcd(q,h)}\right)\frac{\varphi(q)}{\varphi\left(\frac{q}{gcd(q,h)}\right)}.$$

We used that $\frac{p^{\beta+1}}{d} \leq p \Leftrightarrow p^\beta \leq d$ that is $d = p^\beta$, all the remaining summands with $\mu\left(\frac{p^{\beta+1}}{d}\right)$ factors are zero.

Case 3: $k \geq \beta + 2$ Again $\gcd(q, h) = p^\beta$ and in the sum

$$c_{p^k}(h) = \sum_{d|p^\beta} \mu\left(\frac{p^k}{d}\right) d$$

all the factors with μ are 0, since $\dfrac{p^k}{d} \geq \dfrac{p^k}{p^\beta} \geq p^2$. Thus we again have

$$c_{p^k}(h) = 0 = \mu(p^{k-\beta})\frac{\varphi(q)}{\varphi\left(\frac{q}{\gcd(q,h)}\right)}.$$

□

After the estimate of $S(\alpha)$ at the minor arcs in Claim 5.3.1 it is now time to estimate it on the major arcs.

Claim 5.4.1. *Let $\beta = \alpha - \frac{a}{q}$ and $T(\beta) = \sum_{m=1}^{n} e(\beta m)$ with $1 \leq a \leq q \leq Q$ such that $\gcd(a, q) = 1$. Then for any $\alpha \in \mathfrak{M}(a, q)$ there exists a constant $c > 0$ such that*

$$S(\alpha) = \frac{\mu(q)}{\varphi(q)}T(\beta) + \mathcal{O}(ne^{-c\sqrt{\log n}}).$$

Proof. Consider first

$$S\left(\frac{a}{q}, x\right) = \sum_{p \leq x} (\log p) e\left(\frac{a}{q}p\right).$$

Then we have

$$\begin{aligned} S\left(\frac{a}{q}, x\right) &= \sum_{r=1}^{q} e\left(\frac{q}{q}r\right) \sum_{\substack{p \leq x \\ p \equiv r \mod q}} \log p \\ &= \sum_{\substack{r=1 \\ \gcd(r,q)=1}}^{q} e\left(\frac{a}{q}r\right) \theta(x; r, q) + \sum_{\substack{r=1 \\ \gcd(r,q) \geq 2}}^{q} e\left(\frac{a}{q}r\right) \sum_{\substack{p \leq x \\ p \equiv r \mod q}} \log p. \end{aligned}$$

Note that in the second sum if $\gcd(r, q) = d$ then $p \equiv r \mod q$ yields $d|p$ and since $d \geq 2$ we should have $d = p$. But then in the second summation there is only one term for each p dividing q in the inner sum and the whole contribution is $\mathcal{O}(\omega(q)\log x) \ll \log q \log x$. Here $\omega(q)$ denotes the number of prime divisors of q and we use the trivial bound $\omega(q) \ll \log q$.

Now we can apply the Theorem of Siegel-Walfisz for estimating $\theta(x; r, q)$.

$$S\left(\frac{a}{q}, x\right) = \sum_{\substack{r=1 \\ \gcd(r,q)=q}}^{q} e\left(\frac{a}{q}r\right)\left(\frac{x}{\varphi(q)} + \mathcal{O}(xe^{-c\sqrt{\log x}})\right) + \mathcal{O}(\log x \log q) =$$

$$= \frac{x}{\varphi(q)} \sum_{\substack{r=1 \\ \gcd(r,q)=q}}^{q} e\left(\frac{a}{q}r\right) + \mathcal{O}(qxe^{-c\sqrt{\log x}}) = \frac{x}{\varphi(q)}c_q(a) + \mathcal{O}(xe^{-c\sqrt{\log x}}),$$

since $q \leq (\log x)^B \leq (\log n)^B = Q$, and we consider $x \leq n$. Also, the constant c in the error terms might vary in the different instances. Clearly if $\gcd(a, q) = 1$ then $\frac{q}{\gcd(a,q)} = q$ and by the above Lemma 5.4.1 we have $c_q(a) = \mu(q)$. So we get

$$S\left(\frac{a}{q}, x\right) = \frac{\mu(q)}{\varphi(q)}x + \mathcal{O}(xe^{-c\sqrt{\log x}}). \tag{5.1}$$

Something more, for all $x \leq n$ we have

$$S\left(\frac{a}{q}, x\right) = \frac{\mu(q)}{\varphi(q)}x + \mathcal{O}(ne^{-c\sqrt{\log n}}). \tag{5.2}$$

Indeed, if $x \leq \sqrt{n}$ then this is trivial. If $\sqrt{n} \leq x \leq n$ we have

$$\frac{1}{e^{c\sqrt{\log n}}} \gg \frac{1}{e^{c\sqrt{\log x}}} \quad \text{thus} \quad \frac{x}{e^{c\sqrt{\log x}}} \ll \frac{n}{e^{c\sqrt{\log n}}}.$$

So (5.2) follows from (5.1).

We return to estimating $S(\alpha)$ at $\alpha = \frac{a}{q} + \beta$. By partial summation (2.2.3) we get

$$\begin{aligned}
S\left(\frac{a}{q} + \beta\right) &= \sum_{p \leq n} (\log p) e\left(\frac{a}{q}p\right) e(\beta p) = \\
&= e(\beta n) \sum_{p \leq n} e\left(\frac{a}{q}p\right) - \int_1^n \sum_{p \leq x} \log(p) e\left(\frac{a}{q}p\right) e(\beta x)' dx = \\
&= e(\beta n) S\left(\frac{a}{q}\right) - 2\pi i\beta \int_1^n S\left(\frac{a}{q}, x\right) e(\beta x) dx = \\
&= e(\beta n)\left(\frac{\mu(q)}{\varphi(q)}n + \mathcal{O}(ne^{-c\sqrt{\log n}})\right) - 2\pi i\beta \int_1^n \left(\frac{\mu(q)}{\varphi(q)}x + \mathcal{O}(ne^{-c\sqrt{\log n}})\right) e(\beta x) dx = \\
&= \frac{\mu(q)}{\varphi(q)}\left(ne(\beta n) - 2\pi i\beta \int_1^n xe(\beta x) dx\right) + \mathcal{O}\left(n(1 + n|\beta|)e^{-c\sqrt{\log n}}\right).
\end{aligned}$$

Recall that $T(\beta) = \sum_{k \leq n} e(\beta k)$ and again by summation by parts we get

$$T(\beta) = ne(\beta n) - 2\pi i\beta \int_1^n [x] e(\beta x) dx.$$

Then we use that $|\beta| \leq \tau = \dfrac{(\log n)^B}{n}$ and so $n|\beta| \ll (\log n)^B$. We get

$$\begin{aligned} S\left(\frac{a}{q}+\beta\right) &= \frac{\mu(q)}{\varphi(q)}T(\beta) + \mathcal{O}\left(\beta \int_1^n \{x\} e(\beta x) dx\right) + \mathcal{O}(ne^{-c\sqrt{\log n}}) \\ &= \frac{\mu(q)}{\varphi(q)}T(\beta) + \mathcal{O}\left(ne^{-c\sqrt{\log n}}\right), \end{aligned}$$

which proves the Claim. □

Now we are able to prove the estimate on the major arcs.

Theorem 5.4.2. *Suppose that A is a positive constant and $B \geq 2A$. Then*

$$\int_{\mathfrak{M}} S(\alpha)^3 e(-\alpha n) d\alpha = \frac{1}{2}C(n)n^2 + \mathcal{O}(n^2(\log n)^{-A})$$

where

$$C(n) = \prod_{p\nmid n}(1+(p-1)^{-3}) \prod_{p|n}(1-(p-1)^{-2}).$$

Proof. First, by Claim 5.4.1, we have

$$S\left(\frac{a}{q}+\beta\right)^3 = \left(\frac{\mu(q)}{\varphi(q)}T(\beta) + \mathcal{O}(ne^{-c\sqrt{\log n}})\right)^3 = \frac{\mu(q)}{\varphi(q)}T(\beta)^3 + \mathcal{O}(n^3e^{-c\sqrt{\log n}})$$

because $\left|\dfrac{\mu(q)}{\varphi(q)^3}T(\beta)\right| \leq n$ by a trivial estimate.

When $\alpha \in \mathfrak{M}(a,q)$ we have $\alpha = \frac{a}{q} + \beta$ and

$$\begin{aligned} &\int_{\mathfrak{M}(a,q)} S(\alpha)^3 e(-\alpha n) d\alpha = \\ &\quad = \frac{\mu(q)}{\varphi(q)^3} e\left(-\frac{a}{q}n\right) \int_{\mathfrak{M}(a,q)} T(\beta)^3 e(-\beta n) d\alpha + \mathcal{O}\left(\int_{\mathfrak{M}(a,q)} n^3 e^{-c\sqrt{\log n}} e(-\alpha n) d\alpha\right) = \\ &\quad = \frac{\mu(q)}{\varphi(q)^3} e\left(-\frac{a}{q}n\right) \int_{-\tau}^{\tau} T(\beta)^3 e(-\beta n) d\beta + \mathcal{O}(n^2 e^{-c\sqrt{\log n}}) \end{aligned}$$

where we used that $|\beta| \leq \tau = \dfrac{(\log n)^B}{n}$.

Now integrating over all major arcs $\mathfrak{M}$ we get

$$\begin{aligned} \int_{\mathfrak{M}} S(\alpha)^3 e(-\alpha n) d\alpha &= \sum_{q\leq Q} \sum_{\substack{a=1 \\ \gcd(a,q)=1}}^{q} \int_{\mathfrak{M}(a,q)} S(\alpha)^2 e(-\alpha n) d\alpha = \\ &= \sum_{q\leq Q} \sum_{\substack{a=1 \\ \gcd(a,q)=1}}^{q} \frac{\mu(q)}{\varphi(q)^3} e\left(-\frac{a}{q}n\right) \int_{-\tau}^{\tau} T(\beta)^3 e(-\beta n) d\beta + \mathcal{O}(Q^2 n^2 e^{-c\sqrt{\log n}}) \end{aligned}$$

since

$$\sum_{q\le Q}\sum_{\substack{a=1\\ \gcd(a,q)=1}}^{q} 1 \le \sum_{q\le Q} q = 1+2+\cdots+Q = \frac{(Q+1)Q}{2} \ll Q^2.$$

Then we can write, using the appropriate notation,

$$I_{\mathfrak{M}} = C^*(n,Q)\int_{-\tau}^{\tau} T(\beta)^3 e(-\beta n)d\beta + \mathcal{O}(n^2 e^{-c\sqrt{\log n}}).$$

Recall Lemma 5.3.2: Since for $\beta \in \left(-\frac{1}{2},\frac{1}{2}\right)$ we have $\|\beta\| = |\beta|$, we will get

$$|T(\beta)| \ll \frac{1}{\|\beta\|} = \frac{1}{\beta}.$$

Thus

$$\int_{-\tau}^{\tau} T(\beta)^3 e(-\beta n)d\beta = \int_{-\frac{1}{2}}^{\frac{1}{2}} T(\beta)^3 e(-\beta n)d\beta + \mathcal{O}\left(\int_{\tau}^{\frac{1}{2}} |T(\beta)|^3 d\beta\right) =$$

$$=: J(n) + \mathcal{O}\left(\int_{\tau}^{\frac{1}{2}} \frac{d\beta}{\beta^3}\right) = J(n) + \mathcal{O}(\tau^{-2}) = J(n) + \mathcal{O}(n^2 Q^{-2}).$$

Then we have

$$I_{\mathfrak{M}} = C^*(n,Q)J(n) + \mathcal{O}\left(\sum_{q\le Q}\varphi(q)^{-2} n^2 Q^{-2}\right) + \mathcal{O}\left(n^2(\log n)^{-2B}\right).$$

Recall that we have already seen that $n^{1-\epsilon} = o(\varphi(n))$ and so $\varphi(n)^{-1} \ll n^{\epsilon-1}$. Therefore we have

$$\sum_{q\le Q}\varphi(q)^{-2} \ll \sum_{q\le Q} q^{-(2-\epsilon)} \le \sum_{q=1}^{\infty} q^{-(2-\epsilon)} \ll 1.$$

Then the second error term is absorbed by the last one and we have

$$I_{\mathfrak{M}} = C^*(n,Q)J(n) + \mathcal{O}\left(n^2(\log n)^{-2B}\right).$$

We also used that $Q^2 n^2 e^{-c\sqrt{\log n}} \ll n^2(\log n)^{-2B}$. We note that

$$J(n) = \int_{-\frac{1}{2}}^{\frac{1}{2}} T(\beta)^3 e(-\beta n)d\beta = \sum_{\substack{1\le m_1,m_2,m_3\le n\\ m_1+m_2+m_3=n}} 1,$$

which follows from the orthogonality relation

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} e(\alpha h)d\alpha = \begin{cases} 1, & h=0\\ 0, & h\in\mathbb{Z}\setminus\{0\}.\end{cases}$$

But

$$\sum_{\substack{1\le m_i\le n\\ m_1+m_2+m_3=n}} 1 = \sum_{\substack{1\le m_i\le n\\ m_1+m_2=n-m_3}} 1 = \sum_{1\le m_3\le n-2} (n-m_3-1)$$

$$= (n-1)\sum_{1\le m_3\le n-2} 1 - \sum_{1\le m_3\le n-2} m_3 =$$

$$= (n-1)(n-2) - \frac{1}{2}(n-2)(n-1) = \frac{1}{2}(n-1)(n-2).$$

We already showed that $|C^*(n,Q)| \ll 1$, thus we can proceed by writing

$$I_{\mathfrak{M}} = \frac{1}{2}C^*(n,Q)n^2 + \mathcal{O}(n^2(\log n)^{-2B}).$$

Finally, let us complete the singular series by introducing

$$C^*(n) = \sum_{q=1}^{\infty} \frac{\mu(q)}{\varphi(q)^3} \sum_{\substack{a=1 \\ \gcd(a,q)=1}}^{q} e\left(-\frac{a}{q}n\right).$$

Clearly

$$C^*(n,Q) = C^*(n) + \mathcal{O}\left(\sum_{q>Q} \varphi(q)^{-2}\right).$$

As $\varphi(q)^{-2} \ll q^{-(2-\epsilon)}$ we have

$$\sum_{q>Q} \varphi(q)^{-2} \ll \sum_{q>Q} q^{-(2-\epsilon)} \ll Q^{-(1-\epsilon)} \ll Q^{-\frac{1}{2}}.$$

Then

$$C^*(n,Q) = C^*(n) + \mathcal{O}(Q^{-\frac{1}{2}})$$

and

$$\begin{aligned} I_{\mathfrak{M}} &= \frac{1}{2}C^*(n)n^2 + \mathcal{O}(n^2Q^{-\frac{1}{2}}) + \mathcal{O}(n^2Q^{-2}) = \frac{1}{2}n^2C^*(n) + \mathcal{O}(n^2Q^{-\frac{1}{2}}) \\ &= \frac{1}{2}n^2C^*(n) + \mathcal{O}\left(n^2(\log n)^{-\frac{B}{2}}\right). \end{aligned}$$

Now if we choose B such that $(\log n)^{-\frac{B}{2}} \leq (\log n)^{-A}$ the error term would be of the desired shape.

The only thing left to check is that $C^*(n) = C(n)$. Notice that

$$\sum_{\substack{a=1 \\ gcd(a,q)=1}}^{Q} e\left(-\frac{a}{q}n\right) = c_q(n)$$

is the Ramanujan's sum. So

$$C^*(n) = \sum_{q=1}^{\infty} \frac{\mu(q)}{\varphi(q)^3} c_q(n) = \prod_p \left(1 + \frac{\mu(p)}{\varphi(p)^3} c_p(n)\right).$$

Now by Lemma 5.4.1

$$c_p(n) = \varphi(p)\frac{\mu\left(\frac{p}{\gcd(p,n)}\right)}{\varphi\left(\frac{p}{\gcd(p,n)}\right)} = \begin{cases} \mu(p), & \text{if } \gcd(p,n)=1; \\ \varphi(p), & \text{if } \gcd(p,n)=p. \end{cases}$$

Then

$$\begin{aligned} C^*(n) &= \prod_{p \nmid n}\left(1 + \frac{1}{\varphi(p)^3}\right)\prod_{p \mid n}\left(1 - \frac{1}{\varphi(p)^2}\right) \\ &= \prod_{p \nmid n}\left(1 + \frac{1}{(p-1)^3}\right)\prod_{p \mid n}\left(1 - \frac{1}{(p-1)^2}\right) = C(n). \end{aligned}$$

One easily sees that $C(n) \gg 1$, and $C(n) = 0$ when n is even. This proves the Theorem. $\square$

Combining the estimates of the minor and major arcs from Theorem 5.3.1 and Theorem 5.4.2 we get the desired asymptotic formula for the quantity $R(n)$, thus we prove Theorem 5.1.1.

5.5 Other applications of the circle method

5.5.1 Exceptional set for the binary Goldbach's problem

Recall that for the ternary Goldbach's problem we showed that the sum

$$R(n) = \sum_{\substack{p_1,p_2,p_3 \\ p_1+p_2+p_3=n}} \log p_1 \log p_2 \log p_3$$

satisfies the asymptotic relation

$$R(n) = \frac{1}{2}C(n)n^2 + \mathcal{O}\left(\frac{n^2}{(\log n)^A}\right)$$

for any constant $A > 0$ and $1 \ll C(n) \ll \infty$ is the singular series

$$C(n) = \sum_{q=1}^{\infty} \frac{\mu(q)}{\varphi(q)^3} c_q(n).$$

It is conjectured that in the binary Goldbach's problem the corresponding weighted sum

$$R_2(n) = \sum_{\substack{p_1,p_2 \\ p_1+p_2=n}} \log p_1 \log p_2$$

satisfies the asymptotic relation

$$R_2(n) = nC_2(n) + \mathcal{O}\left(\frac{n}{(\log n)^A}\right)$$

for any constant $A > 0$ and the corresponding singular series

$$C_2(n) = \sum_{q=1}^{\infty} \frac{\mu(q)^2}{\varphi(q)^2} c_q(n).$$

Question: Why can't the circle method handle the binary Goldbach's problem?

In the ternary case we used that for the set of minor arcs $\mathfrak{m}$ and $S(\alpha) = \sum_{p \le n} \log p \cdot e(\alpha p)$ we have

$$\int_{\mathfrak{m}} |S(\alpha)|^2 d\alpha = \sum_{p \le n} (\log p)^2 \ll n \log n$$

and

$$\sup_{\alpha \in \mathfrak{m}} |S(\alpha)| \ll n(\log n)^{4-\frac{B}{2}}.$$

Then we saw in Theorem 5.3.1 that $|I_{\mathfrak{m}}| \ll \frac{n^2}{(\log n)^A}$ where $I_{\mathfrak{m}}$ denotes as usual the integral over the minor arcs. This suffices, since the main term of $R(n)$ is of magnitude n^2. However, if we follow the same idea for the binary Goldbach's problem, we would get

$$\int_{\mathfrak{m}} |S(\alpha)| d\alpha \ll n$$

and

$$\left| \int_{\mathfrak{m}} S(\alpha)^2 e(-n\alpha) d\alpha \right| \ll \frac{n^2}{(\log n)^A}$$

which is a problem in the binary case since there we expect a main term of magnitude n. Actually, we can show that over the major arcs

$$\int_{\mathfrak{M}} S(\alpha)^2 e(-n\alpha) d\alpha = nC_2(n) + o(n).$$

Combining the estimates over the major and minor arcs only provides an upper bound $\mathcal{O}\left(\frac{n^2}{(\log n)^A}\right)$ for $R_2(n)$ and does not guarantee $R_2(n) > 0$.

Still, the circle method can be used to prove the following second moment version of the binary problem.

Theorem 5.5.1. *Let $A > 0$ be any constant. Then*

$$\sum_{m=1}^{n} |R_2(m) - mC_2(m)|^2 \ll \frac{n^3}{(\log n)^A}.$$

Using this theorem one can give a non-trivial upper bound of the exceptional set of even numbers which are *not* presentable as the sum of two primes.

Corollary 5.5.1. *Let $E(n)$ be the number of even numbers m not exceeding n for which m is not the sum of two primes. Then*

$$E(n) \ll \frac{n}{(\log n)^A}.$$

Proof. Note that trivially $E(n) \leq \left[\frac{n}{2}\right]$ so the corollary gives a non-trivial improvement.

For each m, counted by $E(n)$, we have $R_2(m) = 0$ and so

$$m^{-2}|R_2(m) - mC_2(m)|^2 = C_2(m)^2 \gg 1.$$

Hence

$$E(n) \ll \sum_{m=1}^{n} m^{-2}|R_2(m) - mC_2(m)|^2 =$$
$$= n^{-2} \sum_{m=1}^{n} |R_2(m) - mC_2(m)|^2 + \mathcal{O}\left(\int_1^n \frac{dt}{(\log t)^A}\right) \ll \frac{n}{(\log n)^A}$$

by the Abel transformation. □

5.5.2 Waring's problem

Theorem 5.5.2. *For any $n \geq 2$ there exists $k = k(n)$ such that every $N \in \mathbb{N}$ can be presented as a sum of at most k n-th powers of positive integers, i.e. the Diophantine equation*

$$x_1^n + \cdots + x_k^n = N$$

has a solution for $x_i \in \mathbb{N}$.

Theorem 5.5.3 (Jacobi,1834)**.** *For any $n \in \mathbb{N}$ let $R(n)$ denote the number of solutions of*

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = n$$

with $x_i \in \mathbb{Z}$. Then

$$R(n) = 8 \sum_{\substack{d|n \\ d \not\equiv 0 \mod 4}} d.$$

Theorem 5.5.4. *Let $n \geq 2$ and $k \geq 2^n + 1$. Then there are $\delta = \delta(k,n) > 0$, $c_1 = c_1(k,n) > 0$, $c_2 = c_2(k,n) > 0$ independent of N, such that if $I_{k,n}$ denotes the number of k-tuples $(x_1, \ldots, x_n) \in \mathbb{N}^k$ satisfying*

$$x_1^n + \cdots + x_k^n = N,$$

we have the asymptotic formula

$$I_{k,n}(N) = \frac{\Gamma\left(1 + \frac{1}{n}\right)^k}{\Gamma\left(\frac{k}{n}\right)} C_{k,n}(N) N^{\frac{k}{n}-1} + \mathcal{O}\left(N^{\frac{k}{n}-1-\delta}\right),$$

where $c_1 \leq C_{k,n}(N) \leq c_2$, and $\Gamma(\alpha)$ is the gamma function.

Set-up of the circle method Take

$$Q = \left(N^{\frac{1}{n}}\right)^{\frac{1}{100}}, \tau = \frac{Q}{N} = \frac{1}{N^{1-\frac{1}{100n}}}.$$

Then

$$\mathfrak{M}(a,q) = \{\alpha : \left|\alpha - \frac{a}{q}\right| \leq \tau\}, \quad \mathfrak{M} = \bigcup_{q \leq Q} \bigcup_{\substack{1 \leq a \leq q-1 \\ \gcd(a,q)=1}} \mathfrak{M}(a,q)$$

and

$$\mathfrak{m} = [\tau, 1+\tau] \setminus \mathfrak{M}.$$

The condition $k \geq 2^n + 1$ plays a major role, it guarantees that the singular series $C_{k,n}(N)$ is absolutely convergent.

In general the circle method provides asymptotic formulae for additive Diophantine problems only if the number of variables is large enough, still, it provides heuristics what to expect for a smaller number of variables as well.