University of Montreal

Linnik's theorem: a comparison of the classical and the pretentious approach

by

Joëlle Matte

Department of mathematics and statistics
Faculty of arts and sciences

Master's thesis presented to the Faculty of graduate studies
in order to obtain the title of
Master of Science (M.Sc.)
en Mathematics

janvier 2019

University of Montreal

Faculty of graduate studies

This master's thesis intitled

Linnik's theorem: a comparison of the classical and the pretentious approach

presented by

Joëlle Matte

was evaluated by the following jury:

Dimitris Koukoulopoulos

(Supervisor)

Matilde Lalin

(member of the jury)

Andrew Granville

(member of the jury)

Master's thesis accepted on:

December 20th 2018

Summary

The goal of this master's thesis is to understand Linnik's theorem, which gives us an upper bound for the first prime number in an arithmetic progression. We will analyze and compare two distinct methods: the classical approach and the pretentious approach. The first one relies on zeros of Dirichlet L-functions. The second one is based on Halász's theorem and distance functions. It was developped by Granville annd Soundarajan.

Keywords: Linnik, zeros of Dirichlet L-functions, pretentiousness, Halász's theorem.

Sommaire

Le but de ce mémoire est de comprendre le théorème de Linnik. Il nous donne une borne supérieure pour le premier nombre premier dans une progression arithmétique. Nous allons analyser et comparer deux méthodes distinctes: la classique et la prétentieuse. La première est basée sur les zéros des fonctions L de Dirichlet. La seconde méthode repose sur le théorème de Halász ainsi que sur la distance entre deux fonctions. Cette approche a été développée par Granville et Soundarajan.

Mots clés: Linnik, zéros des fonctions L de Dircihlet, théorie prétentieuse des nombres, théorème de Halász.

Table of content

List of tables

Acknowledgement

I would like to thank Dimitris without whom this project would not have been possible. I am forever grateful to my parents for the love and support they have given me throughout the years. I would also like to thank my brother for his unconditional kindness and patience. Last but not least, thank you Joël for showing me what true love is.

Introduction

0.1. Background and motivation

A prime number is an integer greater than one which cannot be obtained by multiplying two strictly smaller integers. Euclid was the first one to demonstrate the infinitude of primes. He was a Greek mathematician who lived around 300 BC. His proof was published in *Elements* (Book IX, Proposition 20). He supposed, on the contrary, that there were only $k < \infty$ primes. He demonstrated there would always be at least a $(k+1)$-th prime number. Hence, there are infinitely many of them.

Although Euclid was the first one to publish a proof, many mathematicians demonstrated the infinitude of primes. Euler's proof relies on the Fundamental Theorem of Arithmetic, which states that every integer has a unique prime factorization. He used the fact that the harmonic series diverges to prove that $\sum_{\text{p prime}} 1/p$ diverges as well. Paul Erdös [**21**] gave a third proof that also relies on the Fundamental Theorem of Arithmetic. However, he also used square-free integers. In 1950, Hillel Furstenberg gave a new proof using point-set topology. Other recent proofs were given by Pinasco, Whang, etc.

In 1644, Mengoli tried to understand the sum of $\sum_{n=1}^{\infty} \frac{1}{n^2}$. In 1733, Euler was able to approximate the value of the sum up to 20 decimals. Unfortunately, the convergence of this series is very slow. In 1735, in order to find the exact value, Euler began to study infinite sums. He defined, for $s > 1$:

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

He also demonstrated that

$$\zeta(s) = \prod_{i=1}^{\infty} \frac{1}{1 - p_i^{-s}}.$$

In 1755, Euler stated an interesting hypothesis: for a positive integer d, there are infinitely many primes of the form $1 + nd$ for a non-negative integer n. In 1765, Legendre conjectured that for any two positive coprime integers a and d there are infinitely many primes of the form $a + nd$, where n is a non-negative integer. Notice that Legendre's conjecture is a generalization of Euler's hypothesis. In 1808, Legendre believed he had proven it. Unfortunately, his proof relied on a false lemma. It said that for any two integers m and n relatively coprime and k odd prime numbers not dividing n, there exists at least one integer j between one and the k-th prime, denoted p_k, such that $-m + jn$ is not divisible by any of those k numbers. Thus, Legendre's conjecture remained unproven for many years.

In 1837, Dirichlet proved Legendre's conjecture assuming n was a prime number. A year later, he was able to prove it for every n. Legendre's conjecture became Dirichlet's theorem. Furthermore, in 1841, Dirichlet was able to generalize his proof to complex numbers for which the real part and the imaginary part are both integers. These numbers are called *Gaussian integers.* Dirichlet's demonstration linked Gauss' theory to Euler's ideas, because it used modular arithmetic and analytic number theory. It was a fairly difficult proof which required Complex Analysis and Cauchy's Residue Theorem. Selberg was the first one to find give an elementary[1] demonstration in 1950.

Once Dirichlet's theorem was proved, the next natural question was to count the number of primes p up to x. We denote this by $\pi(x)$. In 1791, Legendre conjectured that $\pi(x)$ can be approximated by

$$\frac{x}{A \log x + B}.$$

However, he did not specify the values of the two constants A and B. In 1808, he stated that $A = 1$ and $B = -1.08366$. Gauss also considered this problem at the age of fifteen. Eventually, Gauss conjectured the sequence of primes up to x has density $x/\log x$.

[1] In number theory, an elementary proof suggests we are not using Complex Analysis or Cauchy's Residue Theorem.

This yields the following function:

$$\mathrm{li}(x) = \int_0^x \frac{dt}{\log t},$$

which is asymptotic to $x/\log x$. Indeed, using integration by parts, we have

$$\mathrm{li}(x) = \frac{x}{\log x}\left(1 + \frac{1}{\log x} + \frac{2!}{\log^2 x} + \frac{3!}{\log^3 x} + ... + O\left(\frac{1}{\log^{k+1}(x)}\right)\right)$$

which means

$$\lim_{x\to\infty} \frac{\mathrm{li}(x)}{x/\log(x)} = \lim_{x\to\infty} \left(1 + \frac{1}{\log x} + \frac{2!}{\log^2 x} + \frac{3!}{\log^3 x} + ... + O\left(\frac{1}{\log^{k+1}(x)}\right)\right) = 1.$$

In two papers, published in 1848 and 1850, Chebychev tried to prove what will eventually become the asymptotic law of the distribution of prime numbers:

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\ln(x)} = 1.$$

He was able to demonstrate a weaker version. He proved that if the limit above exists, then it is necessarily equal to one. Moreover, he showed the ratio is bounded above and below by two explicitly given constants near 1, for all sufficiently large x.

Another crucial development in the distribution of primes was given by Riemann in 1859. In his memoir *On the Number of Primes Less Than a Given Magnitude*, Riemann explained the link between the distribution of prime numbers and the zeros of the analytically extended Riemann zeta function. He also suggested that it would be possible to use complex analysis to approximate $\pi(x)$.

Using the ideas outlined by Riemann, Hadamard and de la Vallée Poussin independently proved the law of the distribution of prime numbers in 1896. The two proofs used complex analysis. The crucial step was that $\zeta(s) \neq 0$ for $s = 1 + it$, when $t \neq 0$.

Thus, the approximation

$$\pi(x) \sim \frac{x}{\ln(x)} \quad (x \to \infty)^2$$

became the Prime Number Theorem. The initial demonstration used Complex Analysis. Elementary proofs were given by Selberg and Erdös in 1949 and 1950 respectively.

One of the next natural questions is to count the number of primes $p \leq x$ in an arithmetic progression. Thus, we define

$$\pi(x;q,a) = \#\{p \leq x | p \equiv a \mod q\}.$$

Eventually, de la Vallée Poussin showed that

$$\pi(x;q,a) \sim \frac{x}{\phi(q)\log x} \quad \text{for } q \text{ fixed}, \quad (x \to \infty). \tag{0.1.1}$$

A quantitative version of de la Vallée Poussin's proof implies that

$$\pi(x;q,a) \sim \frac{x}{\phi(q)\log x} \quad \text{for } q \leq (\log x)^{1-\epsilon}, \quad (x \to \infty). \tag{0.1.2}$$

Next, Walfisz and Siegel were able to demonstrate

$$\pi(x;q,a) \sim \frac{x}{\phi(q)\log x} \quad \text{for } x \geq \exp(q^{\epsilon}), \quad (x \to \infty). \tag{0.1.3}$$

Notice that in (0.1.3) and (0.1.2), both x and q are allowed to go to infinity.

One of the next natural questions was to ask how big is the first prime in an arithmetic progression [**1**]. Here is a well known approximation:

$$\pi(x;q,a) = \frac{\text{li}(x)}{\phi(q)} + O(\sqrt{x}\log x).^{3} \tag{0.1.4}$$

Many experts were quick to notice that for any positive integers m,q such that $(m,q) = 1$, the least prime p congruent to $m \mod q$, denoted $P(m,q)$, is not very large.

Let

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

---

[2] The reader not familiar with this notation may want to look at section 2.1.

[3] The reader not comfortable with this notation may want to look at the second section of Chapter 2.

$L(s,\chi)$ is called a Dirichlet L-function and $\chi(n)$ is a complex-valued completely multiplicative function. It is called a Dirichlet character. Its proprieties will be further explained in *Chapter 1: Prerequisites.* Nevertheless, the *Generalized Riemann Hypothesis* states that any non-trivial zero of $L(s,\chi)$ is on the line $\Re(s) = \frac{1}{2}$.

If we assume the Generalized Riemann Hypothesis holds, we can conclude from (0.1.4) that

$$P(a,q) = O(\phi^2(q)\log^4(q)).$$

Indeed, we need

$$\frac{x}{\phi(q)\log x} \geq \sqrt{x}\log x.$$

However, this is true if and only if

$$x \geq \phi(q)^2 \log^4 x.$$

However, it is conjectured that $P(a,q) \ll q^{1+\epsilon}$ [**19**]. By comparison, (0.1.3) gives the weak bound $P(q,a) \ll e^{q^{\epsilon}}$.

Before Linnik, the best unconditional lower bounds for the first prime p in an arithmetic progression were extremely distant from the conditional ones. Linnik was able to prove the following [**23**]:

Theorem 0.1.1 (Linnik's theorem)**.** *There are effective and computable constants $c,L \geq 1$ such that whenever $(a,q) = 1$, there exists a prime $p \leq cq^L$ congruent to $a \mod q$.*

As soon as Linnik published his paper, the constant L, now known as Linnik's constant, was numerically estimated by several other mathematicians. Here is a table containing some of the progress made through the years.The core of this memoir will be understanding Linnik's theorem. We will not try to find the best value possible for L.

Tab. 0.1. Approximation of L over the years

L $\leq$	Author	Year of publication
10,000	Pan	1957
10,000	Pan	1957
5448	Pan	1958
630	Jutila	1971
550	Jutila	1970
168	Chen	1977
36	Graham	1977
20	Graham	1981
16	Wang	1986
13.5	Chen and Liu	1989
8	Heath-Brown	1990
5.5	Heath-Brown	1992
5.18	Xylouris	2009
5	Xylouris	2011

Moreover, suppose $c_0 > 0$ is a computable constant for which $L(\sigma + it, \chi) \neq 0$ when $\sigma \geq 1 - \frac{c_0}{\log q}, |t| \leq 1$. In an article written by Heath-Brown [**2**], it is stated that for a constant $c_0 > 0$ and an integer q, both large enough, we can deduce $P(a,q) \ll q^{12/5+\epsilon}$, for any $\epsilon > 0$, by excluding a certain type of zeros.[4] The demonstration is based on zero density estimates developped by Huxley and Jutila.

[4]These zeros are called *Siegel zeros (exceptional zeros).* They will be explained in detail later.

0.2. Overview of Linnik's life

Yuri Vladimirovich Linnik was born in 1915 in Ukraine. His mother was a school teacher. His father, Vladimir Pavlovitch, got a job at the State Institute of Optics in 1926 before eventually being elected to the USSR Academy of Sciences. It is safe to say science was always a part of Linnik's life. After getting a high school diploma, he worked as a lab assistant for over a year before deciding to pursue his education. He studied physics for about three years in Leningrad before transferring to the state university to learn more about mathematics. His supervisor for his doctorate was Vladimir Tartakovski.

Unfortunately, World War II changed everything. In the winter of 1939-1940, Linnik was forced to serve in the Soviet Army. It was only in the spring of 1940, after being discharged, that he was able to submit his thesis. He worked on the *Representation of Big Numbers by Positive Ternary Quadratic Form*. It was very well received by the academic community.

However, his career took the back seat because of the German troops approaching the city. He joined the People's Guard to help defend it. Unfortunately, in September 1941, the enemies started a siege that lasted more than 872 days. Millions of people starved to death. Linnik would probably have been one of them had he not been sent to Kazan where the USSR Academy of Sciences had been moved to because of the war. In 1944, after the siege, he returned home and was appointed as professor of mathematics at Leningrad State University. From that point on, he dedicated his career to organizing the chair of probability theory and founding the famous *Leningrad School of Probability and Mathematical Statistics*. He also worked on number theory and statistics.

He introduced ergodic methods in his first work on the analytic theory of quadratic forms. In a 1941 paper, he introduced the large sieve method in number theory. He wanted to sieve out many residue classes $\bmod p$ from a set of integers, possibly increasing with p. The goal was to verify Vinogradov's hypothesis about the size of the smallest quadratic non-residue modulo p.

Many mathematicians helped expand this new theory: Selberg (1950), Alfréd Rényi (1950), Klaus Roth (1965), Enrico Bombieri (1965), Harold Davenport and Heini Halberstam (1966). The large sieve method led Linnik to study Dirichlet L-functions. Density theorems had been used from the 1930s to study primes and Linnik generalized these theorems to L-functions. Using this, Linnik constructed a series of papers in which he showed exceptional arithmetical consequences, including a variant of the Goldbach Conjecture. [**4**]

Furthermore, in 1950, he merged probability and number theory. He is the first notable mathematician to use number theoretic tools to solve probability problems. It is how the Behren-Fisher problem was demonstrated. In 1973, the authors of the Russian Mathematical Survey wrote:
"*In 1948-49, Linnik obtained results which contained, in principle, a complete solution to two central problems in the theory of the summation of variables forming a Markov chain. One of these, raised by Markov, the creator of the theory of chains, was: to find the conditions for the application of the integral limit theorem to the case of a singular chain. The first papers on this were written by Markov and Bernstein. Linnik substantially improved and developed the methods of his predecessors and gave an almost definitive solution of the problem for an inhomogeneous chain with an arbitrary finite number of events. The second problem concerned the conditions under which the local limit theorem for lattice type variables forming a chain holds. An important feature of the method used in this paper, which was largely responsible for its success, is the use of arguments from the study of trigonometric sums in the theory of numbers.*" [**4**]

Linnik is the co-author of *Characterisation Problems in Mathematical Statistics* which appeared in Russian in 1972. During his very productive career, he received many prestigious awards: State Prize (1947), Lenin Prize (1970) and an honorary doctorate by the University of Paris. He also held important positions. First of all, he was elected as the first president of the Leningrad Mathematical Society in 1959 before being elected to the Leningrad City Council six years later. He also wrote two volumes on number theory in the 80's: *The ergodic method and L-functions* and *L-functions and the dispersion method.* A volume has also been published of his work on probability theory (1981) and on mathematical statistics (1982).

0.3. Structure of this memoir

The main goal of this memoir is to understand the value of the least prime in an arithmetic progression. In 1944, Linnik showed there exists constants $c,L > 0$ such that the least prime in the arithmetic progression $a + nd$ is strictly bounded by cd^L.

To do so, we will understand two different proofs of the theorem: the classical proof and a new, innovative, proof which is said to be pretentious. After giving a short introduction, a few basic analytic number theory results will be explained.

The second section will give an introduction to pretentious number theory: how and why it was developped, the distance function, etc.

The third section will explain a few results in sieve theory: the Fundamental Theorems of Sieves and Selberg's sieve.

In the fourth section, a summary of the two different methods will be given: the classical approach and the pretentious approach.

The fifth section will give a detailed proof of the three principles: zero-free region, log-free zero-density estimate and the exceptional zero repulsion.

The sixth section and seventh section will give complete demonstrations of Linnik's theorem using the classical and the pretentious approach. We will also explain how to deduce Linnik's theorem from the two respective methods.

The last section is an appendix which states results taken as black boxes.

Chapter 1

Prerequisites

1.1. Notation

Throughout this memoir, we will use the following asymptotic notation:

- $f(x) \sim g(x) \quad (x \to x_0)$ is equivalent to saying $\lim_{x\to x_0} \frac{f(x)}{g(x)} = 1$.

- $f(x) = O(g(x))$ if there exists $c > 0$ and $X > 0$ such that $|f(x)| \leq c|g(x)|$ for every $x \in X$. This symbol is usually used for error terms. We can also write $f(x) \ll g(x)$ to bound quantities.

- $f(x) \asymp g(x)$ if $f(x) \ll g(x)$ and $g(x) \gg f(x)$. It means f is proportionate to g. Thus, the two functions have the same growth.

1.2. Convolution of two functions

The uniqueness of the prime factorization of any natural number yields many interesting questions. We might want to understand the arithmetic function $f : \mathbb{N} \to \mathbb{C}$ associating a positive integer to its number of distinct prime factors.

A function is multiplicative if $f(mn) = f(m)f(n)$ when $(m,n) = 1$. If the equality holds for any $m,n \in \mathbb{N}$, then it is completely multiplicative. There are many examples of multiplicative functions. Here are a few ones:

1.The Möbius function,

$$\mu(n) = \begin{cases} (-1)^r & \text{if } n \text{ is square free and has } r \text{ prime divisors,} \\ 0 & \text{otherwise.} \end{cases}$$

2. The Euler totient function,

$$\phi(n) = \#\{1 \leq a \leq n : (a,q) = 1\}.$$

3. The k-th divisor function,

$$\tau_k(n) = \#\{(d_1,d_2,...,d_k) \in \mathbb{N}^k : d_1 d_2 \cdots d_k = n\}.$$

Now, suppose we would like to create new arithmetic functions using existing ones. Assume we have two arithmetic functions $f,g : \mathbb{N} \to \mathbb{C}$. We define their Dirichlet convolution $f * g : \mathbb{N} \to \mathbb{C}$ by

$$(f * g)(n) = \sum_{ab=n} f(a)g(b) = \sum_{d|n} f(d)g(n/d).$$

For instance, with $k = 2$, we can see

$$\tau_2(n) = \tau(n) = \sum_{d|n} 1 = (1 * 1)(n).$$

Convolutions are commutative, associative and if f and g are multiplicative, then so is $f * g$. The unit of the convolution is the function:

$$\mathbf{1}(\mathrm{n}) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1. \end{cases}$$

Hence, for any multiplicative function f, there exists a unique multiplicative function g such that $(f * g)(n) = \mathbf{1}(\mathrm{n})$. We say that g is its Dirichlet inverse. For instance, the constant function 1 has Dirichlet inverse $\mu(n)$.

Here is another way of writing this:

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

As a direct consequence, if $f = g * 1$, then $g = \mu * f$. This is called the Möbius Inversion Formula.

Now, if we wish to study the distribution of primes, the Von Mangolt function Λ will be very useful:

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^k \text{ for some prime } p \text{ and some integer } k \geq 1, \\ 0 & \text{otherwise.} \end{cases}$$

We can easily see $\log n = (\Lambda * 1)(n)$, which means $\Lambda = \mu * \log$ by the Möbius Inversion Formula.

1.2.1. Dirichlet L-series

To every arithmetic function, we associate its Dirichlet series

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

defined for every $s \in \mathbb{C}$ for which the series converges. The most famous example is the Riemann zeta function

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

It can sometimes be useful to consider the shifted version of a Dirichlet series[1]:

$$L_y(s,f) = \sum_{\min\{p|n\}>y} \frac{f(n)}{n^s}.$$

When $f(n) = \chi(n)$, we get

$$L_y(s,\chi) = \sum_{\min\{p|n\}>y} \frac{\chi(n)}{n^s}.$$

[1]For instance, it is used in the proof of the Prime Number Theorem.

The first thing we need to study are the values of s for which the series converges. Suppose we let $f(n) = a_n$ and $s = \sigma + it$. Dirichlet series have domain of convergence that are half planes:

- If $L(\sigma_0 + it_0, f)$ converges absolutely, then $L(s, f)$ converges absolutely for each $s \in \mathbb{C}$ with $\Re(s) \geq \sigma_0$.

- If $L(\sigma_0 + it_0, f)$ converges, then $L(s, f)$ converges for each $s \in \mathbb{C}$ with $\Re(s) > \sigma_0$.

Now, we can define the abscissa of convergence $\sigma_c(L)$ of a Dirichlet series:

$$\sigma_c(L) = \inf\{\sigma \in \mathbb{R} : \text{exists } t \in \mathbb{R} \text{ such that } L(\sigma + it, f) \text{ converges.}\}$$

The abscissa of absolute converge $\sigma_a(L)$ can be defined in a similar way:

$$\sigma_a(L) = \inf\{\sigma \in \mathbb{R} : L(\sigma, f) \text{ converges absolutely.}\}$$

Another interesting concept is the possibility to bound $\sigma_a(L)$ using $\sigma_c(L)$:

$$\sigma_c(L) \leq \sigma_a(L) \leq \sigma_c(L) + 1.$$

However, it can sometimes be difficult to determine whether

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

converges absolutely or not. Suppose $f(n)$ is a multiplicative function. Assuming this, $L(s, f)$ converges absolutely if and only if the series

$$\sum_{\substack{p \text{ prime} \\ k \geq 1.}} \frac{f(p^k)}{p^{ks}}$$

converges absolutely. In that case, we can also write

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_{p \text{ prime}} \left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \cdots\right).$$

If $f(mn) = f(m)f(n)$ for every m, n, then

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_{p \text{ prime}} \left(1 - \frac{f(p)}{p^s}\right)^{-1}.$$

An interesting property of Dirichlet series is that they can be analytically or meromorphically continued to the left of their half-plane of convergence. This can be done in many ways, but usually we use the information given by the partial sums of their coefficients.

Suppose $f : \mathbb{N} \to \mathbb{C}$ is such that

$$\sum_{n\leq x} f(n) = cx + E(x)$$

with $c \in \mathbb{C}$ and

$$|E(x)| \leq Mx^{\theta} \quad (x \geq 1)$$

where $\theta \in [0,1), c \in \mathbb{C}$ and $M \geq 1$. Then,

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

has a meromorphic continuation to the half plane $\Re(s) > \theta$ using the formula

$$L(s,f) = \frac{cs}{s-1} + s\int_1^{\infty} \frac{E(x)}{x^{s+1}}dx.$$

Using this, we can see that $\zeta(s)$ can be meromorphically continued to the half-plane $\Re(s) > 0$ with a simple pole at $s = 1$ of residue 1 by the formula

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

Hence, the asymptotic behavior of f determines the analytic behavior of $L(s, f)$. We can easily see that the converse is also true. It suffices to apply *Perron's inversion formula*:

Let $f : \mathbb{N} \to \mathbb{C}$ be an arithmetic function with

$$|f(n)| \ll n^{\alpha} \log^{A}(n)$$

for some constants $\alpha, A \geq 0$. For $x,T \geq 2$ and $c \geq \alpha + 1 + 1/\log x$, we have

$$\sum_{n\leq x} f(n) = \frac{1}{2\pi i}\int_{\substack{\Re(s)=c\\|\Im(s)|\leq T}} L(s,f)\frac{x^s}{s}ds + O\left(\frac{x^c(\log x)^{A+1}}{T} + x^{\alpha}(\log x)^{A}\right).$$

Now, one of the key ideas of this master's thesis will be to study the zeros of

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

Thus, the next step is to understand in details the proprieties of $\chi(n)$.

1.3. Dirichlet characters

It is possible to rewrite the counting function $\pi(x;q,a)$ using $\mathbf{1}_{n\equiv a \mod q}$. Hence,

$$\begin{aligned}\pi(x;q,a) &= \#\{p \le x | p \equiv a \mod q\} \\ &= \sum_{p\le x} \mathbf{1}_{p\equiv a \mod q}.\end{aligned}$$

We will decompose the former in terms of Dirichlet characters. Indeed, we have

$$\mathbf{1}_{n\equiv a \mod q} = \frac{1}{\phi(q)} \sum_{\chi \mod q} \bar{\chi}(a)\chi(n). \tag{1.3.1}$$

In order to study primes in arithmetic progressions, we must explain the proprieties of Dirichlet characters.

Let q be an integer. A Dirichlet character $\mod q$ is a completely multiplicative function $\chi : \mathbb{Z} \to \mathbb{C}$ such that

- χ is q-periodic, which means $\chi(n+q) = \chi(n) \quad \forall n \in \mathbb{N}$;
- $\chi(n) \neq 0$ iff $(n,q) = 1$.

Here is an equivalent way of writing this. There is a homomorphism $\tilde{\chi}(n) : (\mathbb{Z}/q\mathbb{Z})^* \to \mathbb{C}$ such that

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

We will now discuss character theory on abelian groups $(G,\cdot)$. Let $\hat{G}$ denote its $|G|$ characters. We know that $\hat{G}$ forms a group with respect to the usual multiplication of complex-valued functions. Thus, we can define the group homomorphism $\chi : G \to \mathbb{C} \setminus \{0\}$.

Furthermore, $\hat{G}$ is a group with identity element $\mathbf{1_G}$ which equals the trivial character of value one. This element is the *principal character.* All the other elements of the group are called *non-principal characters.*

Now, we may notice that $|G| = \hat{G}$. If G is cyclic, say, $G = \mathbb{Z}/d\mathbb{Z}$, then every character is determined by its value at one so $\chi(1)$ has to be a d-th root of unity. However, if G is not cyclic, we can write it as the direct product of cyclic groups:

$$G \cong \mathbb{Z}/d_1\mathbb{Z} \times \mathbb{Z}/d_2\mathbb{Z} \times ... \times \mathbb{Z}/d_k\mathbb{Z}.$$

The next lemma allows us to conclude $|G| = \hat{G}$ when G is not cyclic:

Lemma 1.3.1. *If* $(G, \cdot)$, $(G_1, \cdot)$ *and* $(G_2, \cdot)$ *are abelian groups such that* $G = G_1 \times G_2$, *then the function*

$$\rho : \hat{G_1} \times \hat{G_2} \to \hat{G}$$

$$(\chi_1, \chi_2) \to \rho_{\chi_1,\chi_2}$$

where

$$\rho_{\chi_1,\chi_2}(g_1,g_2) = \chi_1(g_1)\chi_2(g_2),$$

is a group isomorphism.

Another important aspect of character theory is the existence of two orthogonality relations:

For every $g \in G$,

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

For $\chi \in \hat{G}$,

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

Let's see how to get (1.3.1) using the two relations above:
We can assume $(n,q) = (a,q) = 1$, since otherwise the sum is 0 and then (1.3.1) holds trivially. Now, we suppose a has a multiplicative inverse $\bar{a} \mod q$. We now apply the second orthogonality relation with $g = n\bar{a}$. Notice that

$$\chi(a)\chi(g) = \chi(ag) = \chi(n).$$

So

$$\chi(g) = \frac{\chi(n)}{\chi(a)} = \chi(n)\overline{\chi(a)}.$$

We know $g = n\bar{a} \equiv 1 \mod q$ if and only if $n \equiv a \mod q$, so we obtain the desired result.

The next two concepts we will work on are primitive characters and the conductor of a character. Let $q_1|q$. It is possible that a character $\mod q$ can actually be a character $\mod q_1$ in disguise. The smallest such integer is called the *conductor* of χ. If g is the smallest integer, then χ is said to be *primitive* of conductor q.
Here is a more formal definition:

The character $\chi_1 \mod q_1$ induces the character $\chi_2 \mod q_2$ if

$$\chi_2(n) = \begin{cases} \chi_1(n) & \text{if } (n,q_2) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

This means a Dirichlet character always induces itself. The conductor is the smallest positive integer $q_1|q$ such that there exists a Dirichlet character $\chi_1 \mod q_1$ inducing χ. Also, if the conductor of χ is q, then χ is called primitive.

Here are two ways of determining whether a Dirichlet character $\chi \mod q$ is primitive or not:

- The character χ is imprimitive if and only if there is some $q_1|q$ with $1 \leq q_1 < q$ and $\chi(m) = \chi(n)$ when $m \equiv n \mod q_1$ and $(mn,q) = 1$.

- The character χ is imprimitive if and only if there is some $q_1|q$ with $1 \leq q_1 < q$ and $\chi(n) = 1$ when $n \equiv 1 \mod q_1$ and $(n,q) = 1$.

1.4. Dirichlet L-functions

Now, we have enough background to study the analytic proprieties of Dirichlet L-functions. We can suppose χ is a primitive character of conductor q. If $\chi_1 \mod q_1$ is imprimitive and is induced by $\chi \mod q$, we may write

$$L(s,\chi_1) = L(s,\chi) \prod_{p|q} \left(1 - \frac{\chi(p)}{p^s}\right).$$

This means the theory of $L-$functions with primitive characters has a direct translation to the theory of L-functions with a general character.

The first analytic propriety of Dirichlet L-functions is the functional equation. Let

$$\xi(s,\chi) = \left(\frac{s(s-1)}{2}\right)^{\mathbf{1}_{\chi=\mathbf{1}}} \left(\frac{q}{\pi}\right)^{\frac{s+a}{2}} \Gamma\left(\frac{s+a}{2}\right) L(s,\chi)$$

with

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

where Γ is the gamma function. We define the following functional equation:

$$\xi(1-s,\bar{\chi}) = \frac{i^a\sqrt{q}}{\mathcal{G}(\chi)}\xi(s,\chi),$$

where

$$\mathcal{G}(x) = \sum_{a=1}^{q} \chi(a) e^{\frac{2\pi i a}{q}}$$

and $\bar{\chi}$ is the multiplicative inverse of χ. $\xi(s,\chi)$ is an entire function. The functional equation shows $L(s,\chi)$ has a symmetry around the line $\Re(s) = 1/2$.

The Euler representation of $L(s,\chi)$ means $L(s,\chi) \neq 0$ for $\Re(s) > 1$. Hence, $\xi(s,\chi)$ does not equal zero for $\Re(s) > 1$. Furthermore, by the functional equation, $\xi(s,\chi)$ does not vanish for $\Re(s) < 0$.

The function $\Gamma(s)$ is analytic on the complex plane except at the points $0, -1, -2, \ldots$ where it has simple poles. This means $L(s,\chi) = 0$ when $-2n, n \in \mathbb{N}_{>0}$ for $a = 0$ and $-2n+1, n \in \mathbb{N}_{>0}$ for $a = 1$. All these zeros are simple.

They are called *trivial zeros*. All the other zeros are situated on the critical strip $\{s \in \mathbb{C} : 0 \leq \Re(s) \leq 1\}$ and are *non-trivial*.

Experts have divided the non-trivial zeros into stripes. The most famous one was named by Linnik himself, which he called the *Siegel stripe* [**23**]: $1 - \frac{c_0}{lnD} \leq \sigma \leq 1$. Here, c_0 is a small constant such that there are no zeros, except possibly one in the rectangle $1 - \frac{c_0}{\ln D} \leq \sigma \leq 1, |t| \leq D$, where D is the modulus. This zero, if it exists, is called the Siegel zero. Some mathematicians use the term *exceptional zero* instead. This will be further explained at the beginning of *Chapter 4: Summary of the different methods.*

Throughout this memoir, the non-trivial zeros of $L(s,\chi)$ will be denoted by $\rho_\chi = \beta_\chi + i\gamma_\chi$. The functional equation and the fact that $\overline{L(s,\chi)} = L(\bar{s},\bar{\chi})$ imply that if ρ is a trivial zero, then so is $1 - \bar{\rho}$.

Suppose we want to count the number of non-trivial zeros up to a given height. We will then define

$$N(T,\chi) = \#\{\rho \in \mathbb{C} : 0 \leq \Re(s) \leq 1, |\Im(\rho)| \leq T, L(\rho,\chi) = 0\},$$

where each zero ρ is counted with multiplicity. We may notice that when $\chi = 1$, this quantity counts the number of non-trivial zeros of $\zeta(s)$ in the rectangle above.

Usually, mathematicians want to find a good bound for $N(T,\chi)$. For example, it can be shown that for $T \geq 0$,

$$\frac{N(T,\chi)}{2} = \frac{T}{2\pi} \log\left(\frac{qT}{2\pi e}\right) + O\big(\log q(T+2)\big).$$

We now illustrate the importance of counting zeros of $L(s,\chi)$. Let

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

By formula (5.65) in [**19**],

$$\psi(x,\chi) = \delta_x x - \sum_{\substack{L(\rho,\chi)=0 \\ |\gamma| \leq R}} \frac{x^\rho - 1}{\rho} + O\left(\frac{x}{T}\log^2(xq)\right)$$

where

$$\delta_x = \begin{cases} 1 & \text{if } \chi = \chi_0, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, we will see the above yields a formula for $\psi(x,\chi)$ in terms of the zeros of $L(s,\chi)$. This formula will be key in the proof of Linnik's theorem.

Chapter 2

Pretentious number theory

2.1. How it all began

In 1859, Riemann published a nine page memoir in which he was able to prove that questions on the distribution of zeros of $\zeta(s)$ are more or less equivalent to questions on the distribution of primes. This man is recognized as the father of Analytic Number Theory. The methods he developed were pivotal in answering many arithmetic questions.

The theory of L-functions deals with multiplicative functions that are very special and whose Dirichlet series have very rigid properties. A parallel theory of multiplicative functions was pioneered in the second half of the twentieth century by Wirsing and Halász, and was further developed by various other mathematicians, such as Delange, Daboussi, Elliott, Erdős, Hall, Hildebrand, Montgomery-Vaughan and Tenenbaum, to name a few. Their methods allowed them to handle very general multiplicative functions whose Dirichlet series do not have as nice properties as Dirichlet L-functions.

Selberg's and Erdös' perspectives were also very inspiring. In 1948, Selberg gave an elementary proof of the following formula:

$$\sum_{\substack{\text{p prime} \\ p\leq x}} \log^2 p + \sum_{\substack{\text{p,q prime} \\ pq\leq x}} \log p \log q = 2x\log x + O(x). \tag{2.1.1}$$

Due to its closeness to the Prime Number Theorem, it seemed impossible to demonstrate it without the use of the zeros of $\zeta(s)$. However, Selberg was able to mute the influence of any zero near the 1-line. Hence, it could be proved in an elementary way.

Using (2.1.1), Erdös was able to give a proof of the Prime Number Theorem. Not long after that, Selberg also found a new demonstration.

In recent years, Granville and Soundararajan realized that a lot of this alternative theory can be cast in a unified and conceptual way, using the concept of a multiplicative function f mimicking the behaviour of another function g (we then say that f 'pretends' to be g). They called their approach *Pretentious Multiplicative Number Theory*. Even though the idea of multiplicative functions mimicking each other was implicit in the literature, the attempt to systematically build a coherent theory out of it is novel and has opened up new avenues, leading to exciting developments in number theory, such as in the study of character sums [**7**] and of multiplicative functions in short intervals [**6**].

In 2009, Granville and Soundararajan attended a conference at Princeton University given by Iwaniec. They were surprised to hear that Iwaniec, along with Friedlander, had found a new proof of Linnik's theorem without using zeros of L-functions. In light of this new evidence, Granville and Soundararjan were convinced the new techniques they had been working on could be used to prove every classical result. It was an enormous task. First of all, they gave a new proof of Linnik's theorem. Eventually, the two mathematicians were able to retrieve all the results in Davenport's book and in Bombieri's large sieve book.

Unfortunately, their theory had two major flaws. The first one was that they were unable to get a good error term for the Prime Number Theorem. This had a devastating effect on everything else. For instance, it is impossible to prove the Bombieri-Vinogradov theorem without having a good error term in the Prime Number Theorem. The second flaw was the some steps in the original proof seemed to be "magical". They did not have a global understanding as to why their proofs worked.

These two problems were fixed by two other mathematicians. Koukoulopoulos was able to obtain a strong version of the Prime Number Theorem. His result was just as precise as the classical one.

The second issue was solved by Adam Harper when he gave a new proof of Halász's theorem. It made it much easier to understand the method Granville and Soundararjan had developed.

At this point, Granville and Soundararjan believed they had built a strong theory that could rival the classical theory. *The pretentious approach to number theory* does not use zeros of $\zeta(s)$ or of any other L-function. It has great flexibility for a broad class of functions. For many problems, it allows us to go further than with classical number theory.

2.2. Halász's theorem

Halász's theorem is at the core of pretentious number theory. To understand it, we must define the notion of distance between two functions.

Intuitively, the distance between two functions f and g would be

$$\mathbb{D}^*(f,g;x) = \sum_{p\leq x} \frac{|f(p)-g(p)|^2}{p}.$$

However, if f and g both have values on the unit circle, then $|f(p)|^2 = |g(p)|^2 = 1$. Using this, notice that

$$\begin{aligned} |f(p)-g(p)|^2 &= \Big(f(p)-g(p)\Big)\Big(\overline{f(p)}-\overline{g(p)}\Big) \\ &= |f(p)|^2 + |g(p)|^2 - f(p)\overline{g(p)} - \overline{f(p)}g(p) \\ &= 2 - f(p)\overline{g(p)} - \overline{f(p)}g(p) \\ &= 2\left(1 - \Re(fg)(p)\right) \end{aligned}$$

if f and g are real functions. We will use the later definition for reasons we explain below.

Definition 2.2.1. *Suppose $f(n), g(n)$ are two multiplicative functions vith values on the unit circle. The distance between f and g is defined in the following way:*

$$\mathbb{D}^2(f,g;x) = \sum_{p\leq x} \frac{1-\Re(f\bar{g})(p)}{p}.$$

Similarly, let

$$\mathbb{D}^2(f,g;[y,x]) = \sum_{y<p\le x} \frac{1-\Re(f\bar{g})(p)}{p}.$$

One of the key applications of the distance function is the triangle inequality:

$$\mathbb{D}(f,g;x) + \mathbb{D}(g,h;x) \ge \mathbb{D}(f,h;x). \tag{2.2.1}$$

The next theorem, taken from [**17**], will give an alternative definition for $\mathbb{D}^2(f,g;[y,z])$:

Theorem 2.2.2. *Let f be a completely multiplicative function such that $|f| \le 1$. Suppose $F(s)$ is its Dirichlet series. If $s = \sigma + it$, $y \ge 2$ with $\sigma > 1$, then $|L_y(s,f)| \asymp 1$ when $\sigma \ge 1 + 1/\log y$. However, when $\sigma = 1 + 1/\log x \le 1 + 1/\log y$, then*

$$\log F_y(s) = \sum_{y<p\le x} \frac{f(p)}{p^{1+it}} + O(1).$$

This theorem yields the following 3 +0 result:

$$\mathbb{D}^2(f,g;[y,x]) = \log \left| L_y\left(1+\frac{1}{\log x}, f\bar{g}\right)\right| + O(1) \tag{2.2.2}$$

We can now state a crucial theorem in pretentious number theory [**22**].

Theorem 2.2.3 (Halász)**.** *Suppose $f(n)$ is a multiplicative function such that $|f(n)| \le 1$. If*

$$\left|\frac{1}{x}\sum_{n\le x} f(n)\right| \not\to 0 \text{ as } x \to \infty,$$

then there exists $t \in \mathbb{R}$ such that

$$\mathbb{D}(f(n), n^{it}, \infty) < \infty.$$

We thus see that if f has large partial sums, then it must be structured, in the sense that it is n^{it}-pretentious. The fact that multiplicative functions are structured is at the core of pretentious number theory.

We can easily see how Theorem 2.2.3 fits into the proof of the Prime Number Theorem:

$$\begin{aligned}\text{Prime Number Theorem is false} &\iff \frac{1}{x}\sum_{n\le x}\mu(n) \not\to 0 \quad \text{ as } x\to\infty \\ &\iff \mathbb{D}(\mu(n), n^{it}; \infty) < \infty \quad \text{(Halász.)}\end{aligned}$$

However,

$$\mathbb{D}^2(\mu, n^{it}; \infty) = \sum_{p \text{ prime}} \frac{1 - \Re(\mu(p)\overline{p^{it}})}{p} = \sum_{p \text{ prime}} \frac{1 - \Re(-p^{-it})}{p} = \mathbb{D}^2(1, -n^{it}; \infty).$$

But since

$$\mathbb{D}(\mu, n^{it}; \infty) < \infty,$$

then

$$\mathbb{D}(1, n^{2it}; \infty) < \infty.$$

Indeed, Using (2.2.1), we have

$$\begin{aligned}
\mathbb{D}(1, n^{2it}; \infty) &\leq \mathbb{D}(1, -n^{it}; \infty) + \mathbb{D}(-n^{it}, n^{2it}; \infty) \\
&= \mathbb{D}(1, -n^{it}; \infty) + \mathbb{D}(-1, n^{it}; \infty) \\
&= \mathbb{D}(1, -n^{it}; \infty) + \mathbb{D}(1, -n^{it}; \infty) \\
&= 2\mathbb{D}(1, -n^{it}; \infty) \\
&= 2\mathbb{D}(\mu, n^{it}, \infty).
\end{aligned}$$

This will imply

$$\zeta(1 + 2it) = \infty,$$

which is a contradiction unless $t = 0$. Note that $\mathbb{D}(\mu,1;\infty) = \infty$ by Mertens' theorem (Theorem A.1.4) so this case cannot occur as well.

Another interesting idea is the link between exceptional zeros and pretentiousness. Suppose β is an exceptional zero close to one. Thus,

$$L(\beta,\chi) = 0 \iff L(1,\chi) = \prod_p \left(1 + \frac{\chi(p)}{p} + \cdots\right) \ll \frac{1}{\log q}.$$

Roughly speaking, it is equivalent to saying $\chi(p) = -1 = \mu(p)$. We can say, approximately, that μ pretends to be χ. Hence, By Halász's theorem,

$$\mathbb{D}\left(\mu, \chi; [q, e^{\frac{1}{1-\beta}}]\right) < \infty.$$

However, not only the distance function is finite, but it is a small number.

Now, suppose we have a function f such that $f(p^k) = 0$ if $p > x$. Let us define $s = c + it$ with $c > 1$ and c close to one. Suppose

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

The link between the size of $F(s)$ and the distance function will now be studied. Thus,

$$\begin{aligned} F(s) &= \sum_{n \geq 1} \frac{f(n)}{n^s} \\ &= \prod_{p \leq x} \left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + ...\right) \\ &\approx \exp\left(\sum_{p \leq x} \frac{f(p)}{p^c} p^{-it}\right) \\ &\approx \exp\left(\sum_{p \leq x} \frac{f(p)}{p} p^{-it}\right) \\ &= \exp\left(-\sum_{p \leq x} \frac{-f(p)}{p} p^{-it}\right) \\ &\approx \log x \exp\left(\sum_{p \leq x} \frac{1 - f(p)p^{-it}}{p}\right). \end{aligned}$$

So

$$|F(s)| = |F(c + it)| \approx (\log x) \exp\Big(- \mathbb{D}^2(f(n), n^{it}; x)\Big). \tag{2.2.3}$$

Hence, the distance function helps understand the integrals in the classical theory. (i.e. Perron's formula.) It is crucial to understand the size of $F(s)$. Its maximum, up to height T, will occur when the distance between f and n^{it} is minimal.

We will use (2.2.3) to define $M_f(x,T)$:

$$\begin{aligned} M_f(x,T) &= \frac{1}{\log x} \max_{|t| \leq T} |F(c + it)| \\ &\approx \exp\left(- \min_{|t| \leq T} \mathbb{D}^2(f(n), n^{it}; x)\right) \\ &\ll 1. \end{aligned}$$

With this in mind, it is possible to define a second version of Halász's theorem [**22**].

Theorem 2.2.4 (Halász's theorem, encore)**.** *Let f be a multiplicative function on the unit circle. Then,*

$$\frac{1}{x}\sum_{n\le x} f(n) \ll \Big(1+|\log M_f(x,T)|\Big)M_f(x,T) \quad \textit{for } T \ge \log x.$$

Moreover, we can use pretentious number theory to demonstrate exceptional zeros repel each other. This will be a key element of this memoir.

Proof. Suppose

$$\sum_{n\le x} \mu(n)\chi(n) = o(x). \tag{2.2.4}$$

By Halász's first theorem, it means that

$$\mathbb{D}(\mu,\chi;\infty) = \infty.$$

We want to get a lower bound for $\mathbb{D}(\mu,\chi;\infty)$. However, we know

$$\mathbb{D}(\mu,\chi;\infty) = \infty \iff \sum_{p\le x} \frac{1+\chi(p)}{p} = \infty,$$

which implies

$$|L_y(1+1/\log x)| \asymp 1.$$

But

$$\mathbb{D}^2(\mu,\chi;x) = \sum_{p\le x}\frac{1+\chi(p)}{p} \to \infty \iff \sum_{p\le x}\frac{\Re(\chi(p))}{p} \ge -\log\log x + \xi(x), \quad (\xi(x)\to\infty)$$

$$\iff |L_y(1+1/\log x,\chi)| \gg \frac{e^{\xi(x)}}{\log x}, \quad (\xi(x)\to\infty).$$

Unfortunately, this is a very difficult bound to demonstrate directly. It explains why we will prove (2.2.4) using (2.2.1). We now suppose $\mathbb{D}(\mu,\chi;\infty) < \infty$. By (2.2.1),

$$\mathbb{D}(\mu,\chi;\infty) + \mathbb{D}(\mu,\psi;\infty) \ge \mathbb{D}(\chi,\psi;\infty).$$

If we let $\psi = \bar{\chi}$, we get

$$\mathbb{D}(\chi,\bar{\chi};\infty) < \infty$$

which implies

$$\mathbb{D}(1,\chi^2;\infty) < \infty.$$

In general, $\mathbb{D}(\chi, \psi; x)$ will be large if $\chi \neq \psi$. Thus, if χ is a complex number, then $\chi \neq \bar{\chi}$. So

$$\begin{aligned} 2\mathbb{D}(1, \chi; x) &= \mathbb{D}(\mu,\chi; x) + \mathbb{D}(\mu, \bar{\chi}; x) \\ &\geq \mathbb{D}(\chi, \bar{\chi}; x) \\ &= \mathbb{D}(1,\chi^2; x) \text{ is big.} \end{aligned}$$

It yields a contradiction if $L(1,\chi^2)$ is small. Indeed,

$$\begin{aligned} \mathbb{D}^2(1, \chi; x) = \sum_{p \leq x} \frac{1-\chi(p)}{p} \to \infty \iff & \sum_{p} \frac{\Re(\chi^2(p))}{p} \leq \log\log x - \varrho(x), \quad (\varrho(x) \to \infty) \\ \iff & |L_y(1 + 1/\log x, \chi)| \ll \frac{\log x}{e^{\varrho(x)}}, \quad (\varrho(x) \to \infty) \end{aligned}$$

which is equivalent to showing $|L(1,\chi^2)|$ has an upper bound. χ^2 is an even character. Letting $\chi^* = \chi^2$, we can bound $L(1,\chi^2) = L(1,\chi^*)$ by using the usual lower bounds on $L(1,\chi^*)$. □

In practice, we do not know what happens on $\chi(p)$ for small primes so it makes sense to consider instead sums of the form

$$\sum_{y<p\leq x} \frac{\chi(p)}{p}.$$

2.3. Auxiliary results

Here are a few theorems related to the previous section. They will be used when we will prove Linnik's theorem using the classical and the pretentious approach. All the results of this section are taken directly from [**17**].

Theorem 2.3.1. *Suppose $\chi \mod q$ is a non principal character. If $y \geq 2$, $s = \sigma + it$ and $y \geq q(|t| + 100)$, then*

$$L_y^{(j)}(s,\chi) \ll (\log y)^j$$

when $\sigma > 1 - 1/\log y$.

Theorem 2.3.2. *Suppose χ is a Dirichlet character $\mod q$ and $y \geq q^2(|t|+1)^4 + 3^8$. Then, there exists $c > 0$ such that if $\sigma \geq 1 - c/\log y$, then $L_y(s,\chi) \asymp 1$ when χ is not real or when $t \geq 1/\log y$.*

Here is another important lemma which proves $\mathbb{D}(1, \mu(n)n^{it}; [y,Y]) \ll 1$.

Lemma 2.3.1. *Let f be a multiplicative function of modulus $|f| \leq 1$ and F be its Dirichlet series. Fix $t \in \mathbb{R}$, $y \geq 2$ and assume $\sigma \to F(\sigma + it)$ is continuously differentiable for $\sigma \geq 1$ with $|F_y^{(j)}(\sigma + it, f)| \ll (\log y)^j$ uniformly for $j \in \{0,1\}$ and $\sigma \geq 1$. If $Y = y^{1/|F_y(1+it)|}$, then*

$$\sum_{y<p\leq Y} \frac{1 + \Re(f(p)p^{-it})}{p} = O(1)$$

and

$$\sum_{u<p\leq v} \frac{\Re(f(p)p^{-it})}{p} = O(1)$$

for $v \geq u \geq Y$.

Theorem 2.3.3. *[Zero repulsion]*
Let $\chi_1 \mod q_1$ and $\chi_2 \mod q_2$ be two real, non principal characters that are not induced by the same character. If $L_y(1,\chi_2) \geq L_y(1,\chi_1)$ for a $y \geq \max\{q_1,q_2\}$, then $L_y(1,\chi_2) \asymp 1$.

Proof. It is clear that $L_y(1,\chi_2) \ll 1$. Hence, to conclude $L_y(1,\chi_2) \asymp 1$, it suffices to prove $L_y(1,\chi_2) \gg 1$. But since $L_y(1,\chi_2) \geq L_y(1,\chi_1)$, then $\frac{1}{L_y(1,\chi_2)} \leq \frac{1}{L_y(1,\chi_1)}$. So we can write

$$Y_2 = y^{1/L_y(1,\chi_2)} \leq y^{1/L_y(1,\chi_1)} = Y_1.$$

By the definition of the distance between two multiplicative functions and Lemma 2.3.1, we have

$$\mathbb{D}(\chi_j,\mu; [y, Y_j]) \ll 1.$$

Indeed, the numerator is real and thus we can apply Lemma 2.3.1 for $j = 1,2$. The next step is to apply (2.2.1). It yields

$$\mathbb{D}(\chi_1,\chi_2; [y, Y_2]) \leq \mathbb{D}(\chi_1, \mu; [y, Y_2]) + \mathbb{D}(\mu, \chi_2; [y, Y_2]) \ll 1.$$

We bound $\mathbb{D}^2(\chi_1,\chi_2;[y,Y_2])$ using (2.2.2) :

$$\begin{aligned}\mathbb{D}^2(\chi_1,\chi_2;[y,Y_2]) &= 2\sum_{y<p\leq Y_2}\frac{1-\chi_1(p)\chi_2(p)}{p}\\ &= 2\log\frac{\log Y_2}{\log y} - 2\log L_y\left(1+\frac{1}{\log Y_2},\chi\right) + O(1)\\ &\geq 2\log\frac{\log Y_2}{\log y} + O(1)\end{aligned}$$

and so

$$O(1) \geq 2\log\frac{\log Y_2}{\log y} + O(1).$$

The only way this can be true is if

$$\log\frac{\log Y_2}{\log y} \ll 1,$$

i.e. $\log Y_2 \ll \log y$. Putting the definition of Y_2 in the inequality above, we get

$$\begin{aligned}&\log\left(y^{1/|L_y(1,\chi_2)|}\right) \ll \log y\\ &\Longrightarrow \frac{1}{|L_y(1,\chi_2)|}\log y \ll \log y\\ &\Longrightarrow 1 \ll |L_y(1,\chi_2)|.\end{aligned}$$

The fact that χ_2 is real concludes the proof. □

Here is one last important theorem.

Theorem 2.3.4. *Let $\chi \mod q$ be a non-principal real character. If $Q = q^{1/L_q(1,\chi)}$, then*

$$\mathbb{D}^2(\mu,\chi;[y,z]) \ll \frac{\log z}{\log Q} + y^{-1/\log q} \quad (q \leq y \leq z \leq Q).$$

Chapter 3

Introduction to sieve theory

3.1. Fundamental theorem of sieve methods

First of all, around 300 BC, Euclid proved there are infinitely many primes. Around 50 years later, Eratosthenes of Cyren established a method to calculate primes quickly. He observed that if $n \geq 2$ is composite, then there is a prime number $p \leq \sqrt{n}$ that divides n. Using this fact, Eratosthenes developed an algorithm to count the number of primes up to x. All the integers who do no pass through it are exactly the primes up to x. Recall

$$\pi(x) = \#\{p \text{ prime} : p \leq x\}.$$

The main application of Eratosthenes' sieve is to give an approximation for $\pi(x)$. By the Prime Number Theorem, we know

$$\pi(x) \sim \frac{x}{\log x} \quad (x \to \infty).$$

We find it interesting to compare this approximation to the one obtained by Legendre. Notice that if an integer $n \leq x$ is not divisible by any prime $p \leq x$, then either $n = 1$ or n is a prime number in $[\sqrt{x}, x]$. Thus,

$$\begin{aligned} \pi(x) &= \#\{n \leq x : p|n \Rightarrow p > \sqrt{x}\} - 1 + \pi(\sqrt{x}) \\ &= \#\{n \leq x : p|n \Rightarrow p > \sqrt{x}\} + O(\sqrt{x}). \end{aligned}$$

We can now apply the Inclusion-Exclusion principle to rewrite $\#\{n \le x : p|n \Rightarrow p > \sqrt{x}\}$. So

$$\begin{aligned}\#\{n \le x : p|n \Rightarrow p > \sqrt{x}\} &= \#\Big(\bigcap_{i=1}^{r}\{n \le x : p_i \nmid n\}\Big)\\ &= \#\{n \le x\} - \#\Big(\bigcup_{i=1}^{r}\{n \le x : p_i|n\}\Big)\\ &= \#\{n \le x\} - \sum_{i=1}^{r}\#\{n \le x : p_i|n\} + \sum_{1 \le i < j \le r}^{r}\#\{n \le x : p_ip_j|n\} \pm ...\\ &= \lfloor x \rfloor - \sum_{i=1}^{r}\left\lfloor \frac{x}{p_i} \right\rfloor + \sum_{1 \le i < j \le r}\left\lfloor \frac{x}{p_ip_j} \right\rfloor \pm ...\end{aligned}$$

It is tempting to use the above and the fact that $\lfloor x \rfloor = x + O(1)$ to conclude that

$$\#\{n \le x : p|n \Rightarrow p > \sqrt{x}\} \approx x \prod_{p \le \sqrt{x}}\left(1 - \frac{1}{p}\right).$$

However, Mertens showed

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

where γ is the Euler-Mascheroni constant. Using this in the results above yields

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

But since $2e^{-\gamma} > 1$, then the Prime Number Theorem shows the Eratosthene-Legendre sieve overestimates $\pi(x)$. Legendre tried to improve the result. Indeed, for any $z \in [1,x]$, he observed that integers that have all their prime factors greater than z contain the primes in the interval $(z,x]$. Thus,

$$\begin{aligned}\pi(x) &\le \pi(x) + \#\{n \le x : p|n \Rightarrow p > z\}\\ &\le z + \#\{n \le x : p|n \Rightarrow p > z\}\end{aligned}$$

The Inclusion-Exclusion principle and Merten's theorem yield

$$\begin{aligned}
\#\{n \le x : p|n \Rightarrow p > z\} &= \sum_{p|d \Rightarrow p \le z} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor \\
&= \sum_{p|d \Rightarrow p \le z} \mu(d) \left(\frac{x}{d} + O(1) \right) \\
&= x \prod_{p \le x} \left(1 - \frac{1}{p} \right) + O\left(2^{\pi(z)}\right) \\
&\ll \frac{x}{\log z} + 2^z.
\end{aligned}$$

This means

$$\pi(x) \ll \frac{x}{\log z} + 2^z.$$

Taking $z = \frac{\log x}{2}$, he obtained the following bound:

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

Once again, it is not good enough. However, we can at least conclude that almost all integers are composite asymptotically.

The Prime Number Theorem and the proprieties of $\zeta(s)$ allow us to know quite a lot of the behavior of $\pi(x)$. Sieves are therefore not very useful for this problem. However, they are crucial when the theory of L-functions is unapplicable. Here are few examples [**17**]:

- Are there infinitely many pairs of integers $(n, n+2)$ which are both prime?
- Can any even integer greater than 2 be written as the sum of two primes?
- Is there a prime number between two consecutive squares?
- There are infinitely many many integers n such that $n^2 + 1$ has at most 2 prime factors (Iwaniec, 1980).

Altough the first three examples are still open, the last result was demonstrated approximately four decades ago using sieve theory.

We will now define some notation which can be applied to any sieving problem. Let $\mathcal{A}$ be a finite set of integers. We assume $z \geq 1$ is a real number. We now suppose

$$P(z) = \prod_{p<z} p$$

and

$$S(\mathcal{A}, z) = \#\{a \in \mathcal{A} : (a, P(z)) = 1\}.$$

Now, if we want to count the number of primes in the interval $(m^2, (m+1)^2)$, we need to take $\mathcal{A} = \{n \in (m^2, (m+1)^2)\}$ and $z = m + 1$ [**17**].

In general, the Möbius Inversion Formula allows us to write

$$S(\mathcal{A}, z) = \sum_{a \in \mathcal{A}} \sum_{d|(a, P(z))} \mu(d) = \sum_{d|P(z)} \mu(d)|\mathcal{A}_d|$$

with

$$\mathcal{A}_d = \{a \in \mathcal{A} : a \equiv 0 \mod d\}$$

and $d \in \mathbb{Z}$.

For many sets $\mathcal{A}$, we know $|\mathcal{A}_d|$ is asymptotically equal to $g(d)X$. Here, X is a positive number that approximates $|\mathcal{A}|$. Furthermore, $g : \mathbb{N} \to [0,1]$ is a multiplicative function such that $0 \leq g(p) < 1$, because $0 \leq |\mathcal{A}_d| \leq |\mathcal{A}_1|$. We suppose r_d is the remainder term for this approximation. So

$$r_d = |\mathcal{A}_d| - g(d)X.$$

This allows us to write

$$|\mathcal{A}_d| = g(d)X + r_d.$$

Here, g represents the probability that an element of $\mathcal{A}$ is divisible by d. In practice, we also assume $g(p) \leq \min\{c, k/p\}$ for some parameters $c > 1$ and $k > 0$. We can suppose g is multiplicative because $\mathcal{A}_{d_1}$ and $\mathcal{A}_{d_1}$ will be roughly independent when $(d_1, d_2) = 1$.

Finally, if you have true independence, the probability that an element of $\mathcal{A}$ has no prime factors strictly smaller than z is

$$V(z) = \prod_{p<z}(1-g(p)).$$

Inserting the formula for $|\mathcal{A}_d|$ in the definition of $S(\mathcal{A},z)$ yields

$$S(\mathcal{A},z) = X\sum_{d|P(z)}\mu(d)g(d) + \sum_{d|P(z)}\mu(d)r_d = X\ V(z) + \sum_{d|P(z)}\mu(d)r_d.$$

Taking into account only the first term gives

$$S(\mathcal{A},z) \approx XV(z).$$

Unfortunately, this is not true most of the time because $\sum_{d|P(z)}\mu(d)r_d$ has too many terms.

We will now describe a concept that will play a major role in this memoir. It is called the sifting dimension κ. Conceptually, it represents the average value of $pg(p)$ as p runs through all the primes.

Let I be an interval on the real line. If $\mathcal{A} = \{f(n) : n \in I\}$, then κ also corresponds to the average number of congruence classes that we need to 'remove' modulo each prime in order to extract primes (or products of primes) from the indexing set I. In this memoir, we will only use $\kappa = 0{,}1$. Sieving problems become harder as κ grows.

As seen earlier, one of the main issue of the Eratosthenes-Legendre sieve is that the error term has too many terms. Viggo Brun was the first mathematician to break new ground in this problem. To do so, he used combinatorial concepts. The Inclusion-Exclusion gives us the following formula:

$$\begin{aligned}\sum_{\substack{d|P(z)\\ \omega(d)\leq 2j+1}}\mu(d)|\mathcal{A}_d| \leq S(\mathcal{A},z) &= |\mathcal{A}| - \sum_{p_1<z}|\mathcal{A}_{p_1}| + \sum_{p_2<p_1<z}|\mathcal{A}_{p_1p_2}| - \sum_{p_3<p_2<p_1<z}|\mathcal{A}_{p_1p_2p_3}| \pm \cdots \\ &= |\mathcal{A}| + \sum_{j\geq 1}(-1)^j \sum_{p_j<\ldots<p_2<p_1<z}|\mathcal{A}_{p_1p_2\ldots p_j}| \\ &\leq \sum_{\substack{d|P(z)\\ \omega(d)\leq 2j}}\mu(d)|\mathcal{A}_d|.\end{aligned}$$

Notice that

$$V(z) = 1 + \sum_{j\geq 1}(-1)^j \sum_{p_j<\ldots<p_1} g(p_1p_2\cdots p_j).$$

We are eventually able to obtain the following approximation:

Theorem 3.1.1. *Let* $z \geq 1$ *and* $r = 3.6\log|V(z)|$. *Then,*

$$S(\mathcal{A}, z) = XV(z)\left(1 + O\left(\frac{1}{\sqrt{r}}\right)\right) + O\left(\sum_{\substack{d|P(z)\\ d\leq z^r,\, \omega(d)\leq r}} |r_d|\right),$$

where $\omega(n) = \sum_{p|n} 1$.

The interested reader might look at the details of this proof in chapter 20 of [**18**].

However, the second error term can sometimes be hard to control. We will now introduce a new method that will help diminish this issue. Buchstab noticed that

$$S(\mathcal{A},z) = |\mathcal{A}| - \sum_{p<z} S(\mathcal{A}_p,p).$$

This identity can be used to develop new combinatorial sieves. To do so, we define the following sets

$$\Pi_j \subset \{(p_1, p_2,..., p_j) : z > p_1 > p_2 > ... > p_j \text{ primes}\} \quad (j \geq 1)$$

such that

$$\Pi_{2j+1} \subset \Pi_{2j-1} \times \{p < z\}^2 \quad (j \geq 1)$$

and

$$\Pi_{2j+2} \subset \Pi_{2j} \times \{p < z\}^2 \quad (j \geq 1).$$

Also, suppose

$$\mathcal{D}^+ = \{1\} \cup \{d = p_1p_2...p_r > 1 : (p_1,p_2,...,p_j) \in \Pi_j, 1 \leq j \leq r, j \text{ odd}\}$$

and

$$\mathcal{D}^- = \{1\} \cup \{d = p_1p_2...p_r > 1 : (p_1,p_2,...,p_j) \in \Pi_j, 1 \leq j \leq r, j \text{ even}\}.$$

Hence, applying Buchstab's identity several times, we get

$$\begin{aligned} S(\mathcal{A},z) &= |\mathcal{A}| - \sum_{p_1<z} S(\mathcal{A}_{p_1},p_1) \\ &\leq |\mathcal{A}| - \sum_{p_1 \in \prod_1} S(\mathcal{A}_{p_1},z) \\ &= |\mathcal{A}| - \sum_{p_1 \in \prod_1} |\mathcal{A}_{p_1}| + \sum_{\substack{p_2<p_1 \\ p_1 \in \prod_1}} S(\mathcal{A}_{p_1 p_2},z). \end{aligned}$$

Repeating this step many times gives us

$$S(\mathcal{A},z) \leq \sum_{d|P(z)} \mu^+(d)|\mathcal{A}_d|,$$

where

$$\mu^+(d) = \begin{cases} \mu(d) & d \in \mathcal{D}^+, \\ 0 & \text{otherwise.} \end{cases}$$

We will now state the Fundamental Theorem of Sieve Methods, which will be very useful in this memoir. To do so, we take

$$\Pi_j = \left\{(p_1,p_2,...,p_j) : z > p_1 > \cdots > p_j, p_1 p_2 \cdots p_j < \frac{D}{p_i^\beta}, 1 \leq i \leq j, i \equiv j \mod 2\right\}$$

with an appropriate value of β, depending on the dimension of the sieve. This is called the β-sieve and it is due to Rosser and Iwaniec. Indeed, if

$$\beta_\kappa = \frac{2}{e^{\frac{1}{2\kappa}} - 1} + 1 < 1 + 4\kappa,$$

then we can approximate $S(\mathcal{A},z)$ asymptotically when $z = X^{o(1)}$.

Theorem 3.1.2 (Fundamental Theorem of Sieves)**.** *Let $\mathcal{A}$ be a finite set of integers which satisfies $|\mathcal{A}_d| = g(d)X + r_d$ and*

$$\frac{V(w)}{V(w')} = \prod_{w \leq p < w'} \frac{1}{1-g(p)} \leq K \left(\frac{\log w'}{\log w}\right)^\kappa \quad (3/2 \leq w \leq w')$$

for some $\kappa > 0$, $K \geq 1, z \geq 1$ and $u \geq \epsilon > 0$. Then,

$$S(\mathcal{A},z) = XV(z)\Big(1 + O(e^{-u \log u + O_{\kappa,K,\epsilon}(u)})\Big) + O\left(\sum_{\substack{d|P(z) \\ d \leq z^u}} |r_d|\right).$$

On the other hand, if

$$\frac{V(w)}{V(w')} \leq \left(1+\frac{C_1}{\log w}\right)\left(\frac{\log w'}{\log w}\right)^{\kappa} \quad (3/2 \leq w \leq w')$$

for $C_1 \geq 0$, then

$$S(\mathcal{A},z) \geq \frac{XV(z)}{8} + O\left(\sum_{\substack{d|P(z)\\ d\leq z^{\beta\kappa}}} |r_d|\right).$$

In 1930, Titchmarsh used Brun's sieve to prove that if $q < x^{1-\epsilon}$, then the following result holds [**24**]:

Theorem 3.1.3 (Brun–Titchmarsh)**.** *Let $\pi(x;q,a)$ denote the number of primes $p \equiv a \mod q$ such that $p \leq x$. Then,*

$$\pi(x;q,a) \ll \frac{x}{\phi(q)\log x}.$$

This bound represents the true order of magnitude of $\pi(x;q,a)$ in the whole range $q < x^{1-\epsilon}$. A stronger version of the Brun-Titchmarsh inequality, proven by Montgomery and Vaughan, if often used in modern litterature.

Theorem 3.1.4 (Brun–Titchmarsh, encore)**.** *Let $\pi(x;q,a)$ denote the number of primes $p \equiv a \mod q$ such that $p \leq x$. Then,*

$$\pi(x;q,a) \leq \frac{2x}{\phi(q)\log(x/q)}.$$

3.2. Selberg's sieve

We will now introduce Selberg's sieve. This method will be useful because it will allow us to build mollifiers which will be fundamental in the classical proof of Linnik's theorem.

Although the β- sieve uses optimization, it is implemented step by step. In Selberg's sieve, the weights are optimized globally. The Selberg sieve was developed in the 40's. Selberg replaced the values of the Möbius function which arise by a system of weights which are then optimized to fit the given problem. It enables us to get better upper bounds on $S(\mathcal{A},z)$.

Let $\{\lambda_n\}$ be a sequence of real numbers with $\lambda_1 = 1$. With that in mind, we can write

$$\begin{aligned} \mathbf{1}_{(n,P(z))=1} &\leq \left(\sum_{d|(n,P(z))} \lambda_d \right)^2 \\ &= \sum_{d|(n,P(z))} \sum_{\substack{d_1,d_2 \in \mathbb{N} \\ [d_1,d_2]=d}} \lambda_{d_1}\lambda_{d_2}. \end{aligned}$$

Thus, if $|\mathcal{A}_d| = Xg(d) + R_d$, then

$$\begin{aligned} S(\mathcal{A},z) &\leq \sum_{d_1,d_2|P(z)} \lambda_{d_1}\lambda_{d_2}|\mathcal{A}_{[d_1,d_2]}| \\ &\leq X \sum_{d_1,d_2|P(z)} \lambda_{d_1}\lambda_{d_2} g([d_1,d_2]) + \sum_{d|P(z)} |r_{[d_1,d_2]}\mu^+(d)| \\ &= XG + R, \end{aligned}$$

where $\mu^+(d) = \sum_{\substack{d_1,d_2 \in \mathbb{N} \\ [d_1,d_2]=d}} \lambda_{d_1}\lambda_{d_2}$.

We choose a value of λ_d that will minimize the first term of the sum in order to get the best bound possible. To make sure the error term is as small as possible, we compose the following condition: $\lambda_d = 0$ when $d > \sqrt{D}$ and $|\lambda_d| \leq 1$. The next step is to find a way to optimize

$$\begin{aligned} G &= \sum_{d_1,d_2|P(z)} \lambda_{d_1}\lambda_{d_2} g([d_1,d_2]) \\ &= \sum_{\substack{d_i|P(z), d_i \leq \sqrt{D} \\ i \in \{1,2\}}} \lambda_{d_1}\lambda_{d_2} g([d_1,d_2]) \end{aligned}$$

with $\lambda_1 = 1$. We know that

$$g((d_1,d_2))g([d_1,d_2]) = g(d_1)g(d_2)$$

because if $p^{\nu_1}\|d_1$ and $p^{\nu_2}\|d_2$, then we have $p^{\max\{\nu_1,\nu_2\}}\|[d_1,d_2]$.

We let $\mathbb{P} = \{p \text{ prime} : g(p) \neq 0\}$ and $\mathbb{P}_z = \mathbb{P} \cap (1,z)$. Also, we assume that all the $\lambda_d' s$ are supported on

$$\mathcal{D}_z = \Big\{ d \leq \sqrt{D} : d| \prod_{p \in \mathbb{P}_z} p \Big\}.$$

We may write

$$G = \sum_{d_1,d_2 \in \mathcal{D}_z} \lambda_{d_1}\lambda_{d_2} \frac{g(d_1)g(d_2)}{g((d_1,d_2))}.$$

We will now define

$$h(n) = \prod_{p|n} \left(\frac{g(p)}{1-g(p)} \right) > 0,$$

which allows us to write $1/g(n) = (1 * (1/h))(n)$ with n being a square-free integer for which $g(n) \neq 0$. Thus,

$$\begin{aligned} G &= \sum_{d_1,d_2 \in \mathcal{D}_z} \lambda_{d_1}\lambda_{d_2} g(d_1)g(d_2) \sum_{m|(d_1,d_2)} \frac{1}{h(m)} \\ &= \sum_{m \in \mathcal{D}_z} \frac{1}{h(m)} \sum_{\substack{d_i \in \mathcal{D}_z, m|d_i \\ i \in \{1,2\}}} \lambda_{d_1}\lambda_{d_2} g(d_1)g(d_2) \\ &= \sum_{m \in \mathcal{D}_z} \frac{1}{h(m)} \left(\sum_{\substack{d \in \mathcal{D}_z \\ d \equiv 0 \mod m}} \lambda_d g(d) \right)^2. \end{aligned}$$

We make the following change of variable

$$\xi_m = \sum_{\substack{d \in \mathcal{D}_z \\ d \equiv 0 \mod m}} \lambda_d g(d).$$

It yields

$$\begin{aligned} \sum_{\substack{m \in \mathcal{D}_z \\ m \equiv 0 \mod d}} \xi_m \mu(m/d) &= \sum_{\substack{m \in \mathcal{D}_z \\ m \equiv 0 \mod d}} \mu(m/d) \sum_{\substack{f \in \mathcal{D}_z \\ f \equiv 0 \mod m}} \lambda_f g(f) \\ &= \sum_{\substack{m \in \mathcal{D}_z \\ m \equiv 0 \mod d}} \lambda_f g(f) \sum_{m: d|m|f} \mu(m/d) \\ &= \lambda_d g(d). \end{aligned}$$

This highlights the one-to-one correspondence between λ_d and ξ_m. Furthermore, since $\lambda_d = 1$, we get

$$\sum_{m \in \mathcal{D}_z} \xi(m)\mu(m) = 1.$$

Assuming this, it will suffice to minimize

$$G = \sum_{m \in \mathcal{D}_z} \frac{\xi_m^2}{h(m)}.$$

Using Lagrange multipliers, the minimal value of G is obtained when $\xi_m = c\mu(m)h(m), m \in \mathcal{D}_z, c$ constant. So

$$c = \left(\sum_{m \in \mathcal{D}_z} h(m) \right)^{-1}.$$

This allows us to find the optimal value for λ_d:

$$\begin{aligned} \lambda_d &= \frac{1}{g(d)} \sum_{\substack{m \in \mathcal{D}_z \\ m \equiv 0 \mod d}} \xi_m \mu(m/d) \\ &= \frac{\mu(d)}{g(d)} \frac{\sum_{\substack{m \in \mathcal{D}_z \\ m \equiv 0 \mod d}} h(m)}{\sum_{m \in \mathcal{D}_z} h(m)}. \end{aligned}$$

Moreover,

$$G = \sum_{m \in \mathcal{D}_z} \frac{c^2 \mu^2(m) h^2(m)}{h(m)} = c^2 \sum_{m \in \mathcal{D}_z} h(m) = \left(\sum_{m \in \mathcal{D}_z} h(m) \right)^{-1}$$

by the definition of c above. Furthermore, G and λ_d will allow us to build mollifiers. They will help prove three principles on which the classical proof of Linnik's theorem is based on. Also, by the definition of λ_d above, we can easily see that $|\lambda_d| \leq 1$. Finally, notice that

$$\begin{aligned} |\mu^+(d)| &= \left| \sum_{\substack{d_1, d_2 \in \mathbb{N} \\ [d_1, d_2] = d}} \lambda_{d_1} \lambda_{d_2} \right| \\ &\leq \sum_{\substack{d_1, d_2 | P(z) \\ [d_1, d_2] = d}} 1 \\ &\leq \tau_3(d). \end{aligned}$$

Putting everything together yields Selberg's sieve.

Theorem 3.2.1 (Selberg's sieve)**.** *Suppose $\mathcal{A}$ is a set of integers such that $|\mathcal{A}_d| = Xg(d) + R_d$. Assume D and z are positive real numbers. Let g be a multiplicative function with $0 \leq g(p) < 1$. If $h(n) = \prod_{p|n} \left(\frac{g(p)}{1-g(p)} \right)$, then*

$$S(\mathcal{A}, z) \leq X \left(\sum_{s \leq \sqrt{D},\, s | P(z)} h(m) \right)^{-1} + \sum_{d \leq D,\, d | P(z)} \tau_3(d) |r_d|.$$

Here is an application of Selberg's sieve which will be useful in *Chapter 5: proof of the three principles.*

Theorem 3.2.2. *Let*

$$\lambda_d = \theta_b = \mu(d)\psi\left(\frac{\log d}{\log z}\right),$$

where $\psi \in C^\infty(\mathbb{R})$ *such that*

$$\psi(u) = \begin{cases} 1 & \text{if } u \leq \frac{\log w}{\log z}, \\ 0 & \text{if } u \geq 1, \end{cases}$$

and $0 \leq \psi(u) \leq 1$ *otherwise. Then,* λ_d *is the parameter* $\lambda(d)$ *of a Selberg sieve with* $\mathcal{D} = \{d \leq R : \mu^2(d) = 1\}$.

We will give a heuristic explaination instead of a formal proof.

As defined previously,

$$\begin{aligned} \lambda(d) &= \frac{\mu(d)}{G}\frac{1}{g(d)}\sum_{\substack{m\leq R \\ d|m}} h(m)\mu^2(d) \\ &= \frac{\mu(d)}{\sum_{m\leq R}\mu^2(m)h(m)}\frac{1}{g(d)}\sum_{m'\leq R/d} h(dm')\mu^2(dm') \end{aligned}$$

because we supposed $m = dm'$. We know that dm' being square-free is equivalent to d, m' both being square-free with $(d, m') = 1$. This means we can write

$$\lambda(d) = \frac{\mu(d)}{\sum_{m\leq R}\mu^2(m)h(m)}\frac{h(d)}{g(d)}\sum_{\substack{m'\leq R/d \\ (m',d)=1}} h(m')\mu^2(m').$$

Furthermore, we have

$$h(m) \approx \frac{\tau_k(m)}{m}$$

so

$$h(p) = \frac{g(p)}{1-g(p)} \approx \frac{\kappa/p}{1-\kappa/p} \approx \frac{\kappa}{p}.$$

Hence, there exists a constant C_h such that

$$\sum_{\substack{m\leq R/d \\ (m,d)=1}} \mu^2(m)h(m) \sim C_h\frac{g(d)}{h(d)}\left(\log\frac{R}{d}\right)^\kappa.$$

Putting everything together, we get

$$\lambda(d) = \frac{\mu(d)\frac{h(d)}{g(d)} \sum_{\substack{m \le R/d \\ (m,d)=1}} \mu^2(d) h(m)}{\sum_{m \le R} \mu^2(m) h(m)} \approx \mu(d) \frac{(\log(R/d))^\kappa}{(\log R)^\kappa}.$$

This is equivalent to saying

$$\lambda(d) = \mu(d) F\left(\frac{\log d}{\log R}\right)$$

where

$$F(x) = \begin{cases} (1-x)^\kappa & \text{if } 0 \le x < 1, \\ 0 & \text{if } x > 1. \end{cases}$$

This means

$$F\left(\frac{\log d}{\log R}\right) = \begin{cases} (1 - \frac{\log d}{\log R})^\kappa & \text{if } 0 \le \log d < \log R, \\ 0 & \text{if } \log d > \log R. \end{cases}$$

However, recall

$$\lambda_d = \mu(d) \psi\left(\frac{\log d}{\log z}\right),$$

where $\psi \in C^\infty(\mathbb{R})$ such that

$$\psi(x) = \begin{cases} 1 & \text{if } x \le \frac{\log w}{\log z}, \\ 0 & \text{if } x \ge 1, \end{cases}$$

and $0 \le \psi(x) \le 1$ otherwise. Thus, we can choose

$$\psi(x) = \begin{cases} 1 & \text{if } x \le \frac{\log w}{\log z}, \\ (1-x)^\kappa & \text{if } \frac{\log w}{\log z} < x < 1, \\ 0 & \text{if } x \ge 1, \end{cases}$$

So for every $x \ge 0$, we have $F(x + \log w / \log z) = \psi(x)$. This concludes our heuristic proof.

Chapter 4

Summary of the different methods

4.1. Classical proof

The classical demonstration will be based on the three following principles [**19**].

4.1.1. Three principles

In this subsection, we will state the three principles and give an *outline* of their proofs. Some steps will be skipped for now but will be further explained in *Chapter 5: Proof of the three principles*.

Principle 1: [Zero free region]
There is a positive constant c_1 such that $\prod_{\chi \mod q} L(s,\chi)$ has at most one real simple zero that corresponds to a real Dirichlet character in the region

$$\sigma \geq 1 - \frac{c_1}{\log(qT)}, |t| \leq T.$$

Principle 2: [Log-free zero-density estimate]
There are positive constants c_1, c_2 such that if $\frac{1}{2} \leq \alpha \leq 1$ and $T \geq 1$, then

$$N_q(\alpha, T) := \sum_{\chi \mod q} N(\alpha, T,\chi) \leq c_1(qT)^{c_2(1-\alpha)}$$

where $N(\alpha, T,\chi)$ is the number of zeros $\rho_\chi = \beta_\chi + i\gamma_\chi$ counted with multiplicity in the rectangle $\alpha < \sigma \leq 1, |t| \leq T$ and $\frac{1}{2} \leq \alpha \leq 1$.

Principle 3: [Exceptional zero repulsion]

There is a positive constant c_3 such that, if the exceptional zero β_1 exists, say $L(\beta_1,\chi_1) = 0$ with

$$1 - \frac{c_1}{\log(qT)} \leq \beta_1 < 1,$$

then the function $\prod_{\chi \mod q} L(s,\chi)$ has no other zeros in the region

$$\sigma \geq 1 - c_3 \frac{|\log(1-\beta_1)\log(qT)|}{\log(qT)}, |t| \leq T.$$

The proof of Principle 1 is the standart zero-free region for Dirichlet L-functions. The reader might want to look at Theorem 5.25 in Kowalski's book [**19**].

Sketch of the proof of principle 2: Principle 2 will be proven using the constant $c = 47$ but this method is capable of giving a much smaller constant by direct modifications.

- The first step will be to define a mollifier: λ_d. It is inspired by Selberg's sieve.
- Defining

  $$K(s,\chi) = \sum_{n=1}^{\infty} \left(\sum_{d|n} \lambda_d \right)^2 \frac{\chi(n)}{n^s}$$

  and

  $$K_x(s,\chi) = \sum_{n=1}^{x} \left(\sum_{d|n} \lambda_d \right)^2 \frac{\chi(n)}{n^s},$$

  we will prove $|K(s,\chi) - K_x(s,\chi)| \leq 1/2$. For every ρ with $L(\rho,\chi) = 0$, we will show this implies

  $$\left| \sum_{w<n\leq x} \left(\sum_{d|n} \lambda_d \right)^2 \frac{\chi(n)}{n^\rho} \right| \geq 1/2. \tag{4.1.1}$$

  Here, (4.1.1) will be our zero detector and w will be such that $\sum_{d|n} \lambda_d = 0$ when $1 < n \leq w$. Also, the value of x in the sum comes from the definition of $K_x(s,\chi)$.
- When $\chi = \chi_0$, the Vinogradov zero-free region and the Huxley density estimate will yield

  $$N(\alpha,T,\chi_0) \ll T^{3(1-\alpha)}. \tag{4.1.2}$$

- Using (4.1.2), it will suffice to prove

$$R = \sum_{\chi \neq \chi_0} N(\alpha, T, \chi) \ll (qT)^{47(1-\alpha)} \tag{4.1.3}$$

in order to conclude principle 2.

- The first step to prove (4.1.3) is to use (4.1.1) to get an upper bound for R. Thus,

$$R \leq 2 \sum_{w<n\leq x} \left| \left(\sum_{d|n} \lambda_d \right)^2 \right| \left| \sum_{\chi} \sum_{s \in S(\chi)} \frac{\chi(n) c_\chi(s)}{n^s} \right| \tag{4.1.4}$$

where $c_\chi(s)$ are numbers with norm one and $S(\chi)$ is the set of zeros of $L(s,\chi)$ in our rectangle.

- We will show this implies

$$R^2 \leq 4U_\alpha V, \tag{4.1.5}$$

where

$$U_\alpha = \sum_{w<n\leq x} \left(\sum_{d|n} \lambda_d \right)^2 n^{1-2\alpha}$$

and

$$V = \sum_n f(n) \left(\sum_{d|n} \lambda_d \right)^2 n^{2\alpha-1} \left| \sum_{\chi} \sum_{s \in S(\chi)} c_\chi(s) \frac{\chi(n)}{n^s} \right|^2,$$

where f is a non-negative function such that $f(n) \geq 1$ for $w < n \leq x$.

- A result by Granville, Koukoulopoulos and Maynard and partial summations will allow us to prove

$$U_\alpha \ll x^{2(1-\alpha)} \tag{4.1.6}$$

for every $1/2 \leq \alpha \leq 1$. We will also demonstrate

$$V \ll R(qT)^{1-\alpha}. \tag{4.1.7}$$

- Taking $x = (qT)^{23}$, we will combine (4.1.5), (4.1.6) and (4.1.7) to get

$$R \ll (qT)^{47(1-\alpha)}.$$

□

Sketch of the proof of principle 3: We will use the method of zero detectors from the previous proof on the function $\zeta(s)L(s+\delta_1,\chi_1)$ rather than on $\zeta(s)$. Here, $\chi_1 \mod q$ is the exceptional character associated to the exceptional zero of $L(s,\chi_1)$, β_1, which satisfies $\delta_1 = 1-\beta_1 \leq c_1(\log qT)^{-1}$.

However, the third principle will not be proven directly. The fact that $\prod_{\chi \mod q} L(s,\chi)$ has no other zeros in the region

$$\sigma \geq 1 - \frac{\log(c_0/\delta_1 \log qT)}{92 \log qT}, |t| \leq T \tag{4.1.8}$$

for an absolute constant $c_0 \geq 2c_1$ will be enough to conclude the proof of principle 3. Here is the strategy of the proof:

- Proving

$$1 \ll x^{4(1-\beta_1)}\delta_1 \log x \tag{4.1.9}$$

  yields (4.1.8) by isolating β_1.
- We show (4.1.9) holds by proving

$$1 \ll x^{4(1-\beta_1)}W \tag{4.1.10}$$

  and

$$W \ll \delta_1 \log x, \tag{4.1.11}$$

  where

$$W = \sum_{w<n\leq x} \frac{\nu(n)^2\rho(n)}{n}$$

  with $\nu(n) = \sum_{b|n} \lambda_b$ and $\rho(n) = \sum_{a|n} \frac{\chi_1(a)}{a^{\delta_1}}$.
- In order to prove (4.1.10), we use zero-detecting polynomials to demonstrate

$$\sum_{w<n\leq x} |\nu^2(n)| \frac{\rho(n)}{n^\beta} \geq \frac{1}{2},$$

  where ρ is a zero of $L(s,\chi)$. Applying Hölder's inequality gives us

$$16 \left(\sum_{w<n\leq x} \nu^2(n) n^{1-2\beta} \right)^2 \left(\sum_{w<n\leq x} \nu^2(n) \frac{\rho^3(n)}{n} \right) W \geq 1.$$

  Also, by the proof of the second principle, we know

$$\sum_{w<n\leq x} \nu^2(n) n^{1-2\beta} \ll x^{2(1-\beta)}.$$

Combining everything gives us (4.1.10).

- In order to demonstrate (4.1.11), we will define $W(s) = \sum_{n=1}^{\infty} \frac{\nu^2(n)\rho(n)}{n^s}$. Using Perron's formula, we are going to write $W = \text{res}_{s=1} W(s) \log \frac{x}{w} + O\left(\frac{1}{q}\right)$. Finding the value of the residue and bounding W using sieves will give the desired result.

□

4.1.2. Results leading to the classical approach

We will now use these three principle to prove Linnik's theorem. We will give an *outline* of the proof of Linnik's theorem using the classical method. It will be further explained *Chapter 6: Proof of the classical approach.*

Once the three principles have been proved, Linnik's theorem can be deduced from the following theorem:

Theorem 4.1.1. *Suppose* $\eta_1 = \frac{c_2}{2\log q}$ *and* $\eta_2 = c_3 \frac{|\log(2\delta_1 \log q)|}{2\log q}$. *For* $x \geq q^{4c_2}$, *then*

$$\psi(x;q,a) = \frac{x}{\phi(q)} \left(1 - \chi_1(a) \frac{x^{\beta_1 - 1}}{\beta_1} + O(cx^{-\eta_i/2})\right).$$

Here we suppose the β_1 *term does not exist if there is no exceptional zero. Furthermore,* $i = 2$ *if* β_1 *exists and* $i = 1$ *otherwise.*

Sketch of the proof of Theorem 4.1.1 .

- First of all, we need to find an approximation for $\psi(x;q,a)$. To do so, we use the fact that $\psi(x;q,a) = \frac{1}{\phi(q)} \sum_{\chi \mod q} \bar{\chi}(a)\psi(x,\chi)$. Formula (5.65) in [**19**] gives us

$$\psi(x,\chi) = \delta_x x - \sum_{\substack{L(\rho,\chi)=0 \\ |\gamma| \leq T}} \frac{x^\rho - 1}{\rho} + O\left(\frac{x}{T}\log^2(xq)\right),$$

where

$$\delta_x = \begin{cases} 1 & \text{if } \chi = \chi_0, \\ 0 & \text{otherwise.} \end{cases}$$

Putting everything together will yield

$$\psi(x;q,a) = \frac{x}{\phi(q)} - \frac{1}{\phi(q)} \sum_{\chi \mod q} \bar{\chi}(a) \sum_{T=2^m \leq R} \sum_{\substack{L(\rho,\chi)=0 \\ T/2 \leq |\gamma| \leq T}} \frac{x^\rho}{\rho} + O\left(\frac{x \log x}{R}\right).$$

- From the above, we can easily see

$$\psi(x;q,a) = \frac{x}{\phi(q)} - \frac{\chi_1(a)x^{\beta_1}}{\beta_1} - E_{R_c}$$

where

$$E_{R_c} = \frac{1}{\phi(q)} \sum_{\chi \mod q} \bar{\chi}(a) \sum_{T=2^m \leq R} \left(\sum_{\substack{L(\rho,\chi)=0, \rho \neq \beta_1 \\ T/2 \leq |\gamma| \leq T}} \frac{x^\rho}{\rho} \right) + O\left(\frac{x \log x}{R}\right).$$

- It suffices to prove $E_{R_c} \ll \frac{cx^{1-\eta_i/2}}{\phi(q)}$ in order to conclude Theorem 4.1.1.
- To do so, we bound

$$\left| \sum_{\chi \mod q} \bar{\chi}(a) \sum_{\substack{T/2 \leq |\gamma| \leq T \\ \rho_\chi \neq \beta_1}} \frac{x^{\rho_\chi}}{\rho_\chi} \right| \leq \frac{2}{T}\sqrt{x}N_q^*\left(\frac{1}{2},T\right) + \frac{2\log x}{T}\int_{1/2}^{1-\eta_i} N_q^*(\alpha,T)d\alpha$$

using mainly the triangle inequality and summation by parts. Here, the $*$ symbol means we are excluding the exceptional zero if it exists.
- The second principle will enable us to prove

$$\left| \sum_{\chi \mod q} \bar{\chi}(a) \sum_{\substack{T/2 \leq |\gamma| \leq T \\ \rho_\chi \neq \beta_1}} \frac{x^{\rho_\chi}}{\rho_\chi} \right| \leq \frac{4cx^{1-\eta_i/2}}{T}.$$

- Putting everything together gives us $E_{R_c} \ll \frac{cx^{1-\eta_i/2}}{\phi(q)}$ which is enough to conclude Theorem 4.1.1.

□

We will explain how to deduce Linnik's theorem from this in *Chapter 6: Proof of Linnik's theorem using the classical approach.*

4.2. Results leading to the pretentious approach

The classical approach requires in depth knowledge of the proprieties of Dirichlet L-functions. The pretentious approach uses multiplicative theory. Although the details of the proofs are very different, there are a lot of similarities at a conceptual and structural level.

The classical proof of Linnik's theorem is based on the principles we explained previously. We will see how to obtain what we will call the *Three pretentious principles.* We will compare the classical approach (c) with the pretentious approach (p).

(1) Principle 1

(c) $\prod_\chi L(s,\chi)$ has at most one zero in

$$\sigma \geq 1 - \frac{c_1}{\log qT}.$$

(p) There exists a character χ_1 such that if $\chi \neq \chi_0, \chi_1$, then $L_y(s,\chi) \asymp 1, \sigma > 1, y \geq q$.

$\chi_1 \mod q$ is a real, non-principal Dirichlet character with $L_q(1,\chi) = \min\{L_q(1,\chi) : \chi$ real and non-principal character $\mod q\}$. Here, clearly, χ_1 will be the character associated to a Siegel zero. It is an exceptional character. Let us define $C_q = \{\chi : \chi \neq \chi_0, \chi_1 \mod q\}$. Furthermore, we will see in the next section that one of the main ideas of the proof is demonstrating

$$\phi(q) \sum_{\substack{y<p\leq z \\ p\equiv a \mod q}} \frac{1}{p} = \sum_{y<p\leq z} \frac{1+\chi_1(ap)}{p} - E_{R_p} + O(1)$$

where

$$E_{R_p} = \int_y^z \sum_{\chi \in C_q} \bar{\chi}(a) \frac{L_y^{'}}{L_y} \left(1 + \frac{1}{\log u}, \chi\right) \frac{du}{u \log^2 u}.$$

This means the value of $\sum_{\substack{y<p\leq z \\ p\equiv a \mod q}} \frac{1}{p}$ depends on the value of our exceptional character. Notice that the sum only runs on $\chi \neq \chi_0, \chi_1$ so bounding E_{R_p} will be fairly easy because it excludes the Siegel zero so there are no potential problems. For every $\chi \neq \chi_0, \chi_1$, then $L(s,\chi) \asymp 1$ for $\sigma > 1, y \geq q$. This means that one has constructed a "pretentious zero-free region", which is equivalent to the first principle.

(2) Principle 2

(c)

$$N_q(\alpha,T)=\sum_{\chi \mod q}=N(\alpha,T,\chi)\leq c(qT)^{c_2(1-\alpha)}$$

for $\frac{1}{2}\leq\alpha\leq 1$ and $T\geq 1$.

(p)

Lemma 4.2.1. *If $y\geq q^2>1$ and $1<\sigma\leq 1+\frac{1}{\log y}$, then*

$$\sum_{\chi \mod q}\left|\sum_{\substack{n>y\\ P^-(n)>y}}\frac{\mu(n)\chi(n)}{n^\sigma}\right|^2\ll\frac{1}{(\sigma-1)^2\log^2 y}$$

and

$$\sum_{\substack{\chi \mod q\\ \chi\neq\chi_0}}|L_y'(\sigma,\chi)|^2\ll\log^2 y.$$

Proof. In order to obtain the first estimate, orthogonality is employed. So

$$\sum_{\chi \mod q}\left|\sum_{\substack{n>y\\ P^-(n)>y}}\frac{\mu(n)\chi(n)}{n^\sigma}\right|^2=\phi(q)\sum_{\substack{n_1>y\\ P^-(n_1)>y}}\frac{\mu(n_1)}{n_1^\sigma}\sum_{\substack{n_2>y\\ P^-(n_1)>y,n_1\equiv n_2 \mod q}}\frac{\mu(n_2)}{n_2^\sigma}.$$

By the Fundamental Theorem of Sieves (Theorem 3.1.2), the following set can be approximated:

$$\#\{n\leq x: n\equiv a \mod q, P^-(n)>y\}\ll\frac{x}{\phi(q)\log y}$$

for $x\geq y\geq q^2$. Now, by partial summation, we can see

$$\sum_{\substack{n_2>y,\ P^-(n_2)>y\\ n_1\equiv n_2 \mod q}}\frac{\mu(n_2)}{n_2^\sigma}\ll\frac{1}{\phi(q)(\sigma-1)\log y}. \tag{4.2.1}$$

Thus,

$$\sum_{\chi \mod q} \left| \sum_{\substack{n>y \\ P^-(n)>y}} \frac{\mu(n)\chi(n)}{n^\sigma} \right|^2 = \phi(q) \sum_{\substack{n_1>y \\ P^-(n_1)>y}} \frac{\mu(n_1)}{n_1^\sigma} \sum_{\substack{n_2>y \\ P^-(n_1)>y, n_1 \equiv n_2 \mod q}} \frac{\mu(n_2)}{n_2^\sigma}$$

$$\ll \phi(q) \sum_{\substack{n_1>y \\ P^-(n_1)>y}} \frac{\mu(n_1)}{n_1^\sigma} \left(\frac{1}{\phi(q)(\sigma-1)\log y} \right)$$

$$= \phi(q) \left(\frac{1}{(\sigma-1)\log y} \right) \left(\frac{1}{\phi(q)(\sigma-1)\log y} \right)$$

$$= \left(\frac{1}{(\sigma-1)\log y} \right)^2.$$

The last sum is approximated by taking $q = 1$ in (4.2.1).

The proof of the second estimate can be found in [**17**]. □

These last steps give us the second principle. One of the key ideas of the proof of the classical second principle is to construct Dirichlet polynomials which serve as zero detectors. Eventually, we get a bound for the number of zeros counted with multiplicity in the rectangle $\alpha < \sigma \leq 1, |t| \leq T$. In order words, it is known that for $\frac{1}{2} \leq \sigma \leq 1 - \frac{c}{\log qT}$ there is at most one exceptional zero. However, if we let $\sigma = 1 - \frac{C}{\log qT}$, for $c < C$, the log-free zero density will show that $\#\{\chi$ that have a zero in that region $\} = O(1)$.

To understand why Lemma 4.2.1 is analogous to a log-free zero-density estimate, consider the trivial bound

$$\left| \sum_{\substack{n>y \\ P^-(n)>y}} \frac{\mu(n)\chi(n)}{n^\sigma} \right| \leq \sum_{\substack{n>y \\ P^-(n)>y}} \frac{1}{n^\sigma}$$

$$\approx \frac{1}{(\sigma-1)\log y}.$$

The first sum above will be this big only if $L_y(1,\chi)$ is very small. To this end, let

$$N(\delta) = \#\left\{ \chi : \left| \sum_{\substack{n>y \\ P^-(n)>y}} \frac{\mu(n)\chi(n)}{n^\sigma} \right| \geq \frac{\delta}{(\sigma-1)\log y} \right\}.$$

Thus,

$$N(\delta)\left(\frac{\delta}{(\sigma-1)\log y}\right)^2 \le \sum_{\chi \mod q}\left|\sum_{\substack{n>y\\P^-(n)>y}}\frac{\mu(n)\chi(n)}{n^\sigma}\right|^2 = O\left(\frac{1}{(\sigma-1)\log y}\right)^2.$$

So

$$N(\delta) \ll \frac{1}{\delta^2}.$$

There are a finite number of "bad characters". This is meta equivalent to saying there is a bound for the number of zeros, which is the second principle.

(3) Principle 3

(c) There exists $c_3 > 0$ such that if $L(\beta_1,\chi_1) = 0$ with $1 - \frac{c_1}{\log(qT)} \le \beta_1 < 1$, then the function $\prod_{\chi \mod q} L(s,\chi)$ has no other zeros in the region

$$\sigma \ge 1 - c_3\frac{|\log(1-\beta_1)\log(qT)|}{\log(qT)}, |t| \le T.$$

(p) In the case of an exceptional character, we can sieve for primes using a sieve of dimension $o(1)$. The analogy is seen more clearly by examining the proof of the third principle:

$$\left|\sum_{w<n\le x}\left(\sum_{d|n}\lambda_d\right)^2\rho(n)\frac{\chi(n)}{n^s}\right| \ge \frac{1}{2}.$$

We have $\rho(n) \approx \sum_{a|n}\frac{\chi_1(a)}{a^{\delta_1}} = (1*\chi)(n)$. Also, $\sum_{d|n}\lambda_d$ is our mollifier, which annihilates the first terms of the series. Since $\chi(p) = -1$ for most primes, $\rho(n)$ sieves out most primes, thus rendering the sifting dimension $o(1)$. As we will see in *Chapter 7: Proof of Linnik's theorem using the pretentious approach,* it is very similar to the third case of the pretentious proof, which will occur when $\chi_1(a) = 1$ and $L_q(1,\chi_1) \le L^{-0.99}$. Indeed, $\chi(p) = -1$ for most primes in that case too.

Linnik's theorem can be deduced from the third case by proving

$$S(X,\sqrt{X};a,q) = \sum_{\substack{n \le x, n \equiv a mod q \\ p|n \to p > y}} (1 * \chi_1)(n) > 0$$

with $X = q^{L^{0.49}}$. Here notice that the summation starts at y, so the first terms have been sieved out. We also used $(1 * \chi)(n)$. The same idea is used for the two proofs.

4.3. The pretentious approach

In this subsection, we will give an *outline* of the proof of Linnik's theorem using the pretentious method. Some steps will be skipped for now but will be further explained in *Chapter 7: Proof of Linnik's theorem using the pretentious approach.*

First of all, we define $C_q = \{\chi \neq \chi_0, \chi_1 \mod q\}$.

Linnik's theorem, using this modern approach, can be deduced from a single theorem [**17**].

Theorem 4.3.1. *Let* $q \ge 4$*. If* $\chi_1 \mod q$ *is a real, non principal character with* $L_q(1,\chi_1) = \min\{L_q(1,\chi) : \chi$ *real and non principal character* $\mod q\}$*, then*

$$\sum_{\substack{y < p \le z \\ p \equiv a \mod q}} \frac{1}{p} = \frac{1}{\phi(q)} \sum_{y < p \le z} \frac{1 + \chi_1(ap)}{p} + O\left(\frac{1}{\phi(q)}\right).$$

Once it has been demonstrated, Linnik's theorem can be deduced from this. Three different cases arise which will be explained later in *Chapter 7: Proof of Linnik's theorem using the prententious approach.*

Here are the main steps to prove Theorem 4.3.1.

- Suppose we have the following equality:

$$\phi(q) \sum_{\substack{y<p\le z\\ p\equiv a \mod q}} \frac{1}{p} = \sum_{y<p\le z} \frac{1+\chi_1(ap)}{p} + O(1) - E_{R_p} \tag{4.3.1}$$

with

$$E_{R_p} = \int_y^z \sum_{\chi\in C_q} \bar{\chi}(a) \frac{L_y'}{L_y}\left(1+\frac{1}{\log u},\chi\right)\frac{du}{u\log^2 u}.$$

If this equality holds, Theorem 4.3.1 can be deduced by proving $E_{R_p} \ll 1$. It will be done using the Fundamental Theorem of Calculus, Theorem 4.2.1, the Fundamental Theorem of Sieves (Theorem 3.1.2), partial summation, pretentious zero free regions on $\chi \neq \chi_0, \chi_1$, Theorem 2.3.1, Cauchy-Schwartz, etc.

We will now try to demonstrate (4.3.1).

- To demonstrate (4.3.1), we start by using the definition of $\log L_y\left(1+\frac{1}{\log z},\chi\right) = \sum_{p>y} \frac{\chi(p)}{p^{1+1/\log z}} + O(1)$.
- By Brun-Titchmarsh's inequality (Theorem 3.1.4) and partial summation,

$$\sum_{\substack{p>y\\ p\equiv a \mod q}} \frac{1}{p^{1+1/\log y}} = O\left(\frac{1}{\phi(q)}\right),$$

so

$$\sum_{\chi \mod q} \bar{\chi}(a)\left(\log L_y\left(1+\frac{1}{\log z},\chi\right) - \log L_y\left(1+\frac{1}{\log y},\chi\right)\right) = \phi(q) \sum_{\substack{y<p\le z\\ p\equiv a \mod q}} \frac{1}{p} + O(1). \tag{4.3.2}$$

- Let $A(y,z,\chi) = \log L_y\left(1+\frac{1}{\log z},\chi\right) - \log L_y\left(1+\frac{1}{\log y},\chi\right)$. We can split the right hand side of (4.3.2) into

$$\sum_{\chi\in C_q} \bar{\chi}(a)A(y,z,\chi) + \sum_{\chi\notin C_q} \bar{\chi}(a)A(y,z,\chi) \tag{4.3.3}$$

- The first sum of right hand side of (4.3.3) can be evaluated using the Fundamental Theorem of Calculus.
- The second sum of the right hand side of (4.3.3) can be obtained using Theorem 2.2.2.
- Combining everything yields Theorem (4.3.1). The details will be given later.

Chapter 5

Proof of the three principles

In this section, we will give *complete* proofs of the three principles.

5.1. The first principle

Principle 1: (Zero free region)

There is a positive constant c_1 such that $\prod_{\chi \mod q} L(s,\chi)$ has at most one zero in the region

$$\sigma \geq 1 - \frac{c_1}{\log(qT)}, |t| \leq T.$$

This is a classical result due to Landau. The interested reader can look in Iwaniec-Kowalski's book on Analytic Number Theory [**19**].

5.1.1. The second principle

Recall $N(\alpha,T,\chi)$ is the number of zeros $\rho_\chi = \beta_\chi + i\gamma_\chi$ counted with multiplicity in the rectangle $\alpha < \beta \leq 1, |\gamma| \leq T$ and $1/2 \leq \alpha \leq 1$.

Theorem 5.1.1 (Log-free zero-density estimate)**.** *There are positive constants c_1,c_2 such that for every $\frac{1}{2} \leq \alpha \leq 1, T \geq 1$ and $qT \geq 64$, then*

$$N_q(\alpha,T) = \sum_{\chi \mod q} N(\alpha,T,\chi) \leq c(qT)^{c_2(1-\alpha)}.$$

We will demonstrate the second principle with $c_2 = 47$, but this method can yield better results by direct modifications. The main idea of this theorem is to build Dirichlet polynomials which serve as zeros detectors because they assume large values at the zeros of $L(s,\chi)$. To do so, a mollifier is constructed using sieving theory.

First of all, let

$$\lambda_d = \theta_b = \mu(d)\psi\left(\frac{\log d}{\log z}\right), \tag{5.1.1}$$

where $\psi \in C^\infty(\mathbb{R})$ such that

$$\psi(u) = \begin{cases} 1 & \text{if } u \leq \frac{\log w}{\log z}, \\ 0 & \text{if } u \geq 1, \end{cases}$$

and $0 \leq \psi(u) \leq 1$ otherwise. The definition above holds when $1 \leq d \leq z$ where $1 < w < z$ and we set $\lambda_d = 0$ if $d > z$. We will use a result inspired by [**25**] which will be a key in the classical proof of the second and third principle.

Theorem 5.1.2. *Let λ_d and z be defined as in* (5.1.1)*. Let g be a multiplicative function such that $0 \leq g(p) < 1$. Then,*

$$\sum_{d_1,d_2|P(z)} \lambda_{d_1}\lambda_{d_2} g([d_1,d_2]) \ll \prod_{p\leq z}(1-g(p)).$$

Proof. The Monotonicity Principle (page 49 in [**5**]) allows us to assume that $g(p) = \min\{k/p, p-1\}$ for all primes p. We must show

$$\sum_{d_1,d_2|P(z)} \lambda_{d_1}\lambda_{d_2} g([d_1,d_2]) \ll \frac{1}{\log^k z}.$$

We set $\phi(u) = e^u\psi(u)$ and write ϕ using its Fourier Transform. We have

$$\begin{aligned}
\lambda_d &= \mu(d)\psi\left(\frac{\log d}{\log z}\right) \\
&= \frac{\mu(d)}{2\pi d^{1/\log z}} \int e^{i\xi \log d/\log z}\hat{\phi}(\xi)d\xi \\
&= \frac{\mu(d)}{2\pi} \int \frac{e^{\log(d^{i\xi/\log z})}}{d^{1/\log z}}\hat{\phi}(\xi)d\xi \\
&= \frac{\mu(d)}{2\pi} \int \frac{d^{i\xi/\log z}}{d^{1/\log z}}\hat{\phi}(\xi)d\xi \\
&= \frac{1}{2\pi} \int \mu(d)d^{(-1+i\xi)/\log z}\hat{\phi}(\xi)d\xi.
\end{aligned}$$

Hence,

$$S = \sum_{d_1,d_2|P(z)} \left(\frac{1}{2\pi}\int \mu(d_1){d_1}^{(-1+i\xi_1)/\log z}\hat{\phi}(\xi_1)d\xi_1\right)\left(\frac{1}{2\pi}\int \mu(d_2){d_2}^{(-1+i\xi_2)/\log z}\hat{\phi}(\xi_2)d\xi_2\right)g([d_1,d_2])$$

$$= \frac{1}{4\pi^2}\int\int \sum_{d_1,d_2} \mu(d_1)\mu(d_2)g([d_1,d_2])d_1^{(-1+\xi_1)/\log z}d_2^{(-1+\xi_2)/\log z}\hat{\phi}(\xi_1)\hat{\phi}(\xi_2)d\xi_1 d\xi_2.$$

$$= \frac{1}{4\pi^2}\int\int \prod_p \left(\sum_{\substack{d_1=\nu_1\\ d_2=\nu_2}} \mu(d_1)\mu(d_2)g(p^{\max\{\nu_1,\nu_2\}})p^{-\nu_1(-1+\xi_1)/\log z}p^{-\nu_2(-1+\xi_2)/\log z}\right)$$

where $\nu_1,\nu_2 \in \{0,1\}$ because of the definition of μ.

It is well-known [**17**] that if a function $f(n)$ is multiplicative and converges absolutely, then

$$\sum_{n=1}^{\infty} f(n) = \prod_p \left(1 + f(p) + f(p^2) + \cdots\right).$$

We will use the fact that $\mu, g, {d_1}^{(-1+i\xi_1)/\log z}$ and ${d_2}^{(-1+i\xi_2)/\log z}$ are multiplicative. Furthermore, $\mu(p^k) = 0$ for $k \geq 2$. This yields

$$S = \frac{1}{4\pi^2}\int\int \prod_p \left(1 - g(p)p^{(-1+i\xi_1)/\log z} - g(p)p^{(-1+i\xi_2)/\log z} + g(p)p^{(-2+i\xi_1+i\xi_2)/\log z}\right)\hat{\phi}(\xi_1)\hat{\phi}(\xi_2)d\xi_1 d\xi_2.$$

We know $\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}$, which means $\zeta^k(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-k}$. Since $g(p) = k/p$ for all p large enough, then

$$\left|\prod_p \left(1 - g(p)p^{(-1+i\xi_1)/\log z} - g(p)p^{(-1+i\xi_2)/\log z} + g(p)p^{(-2+i\xi_1+i\xi_2)/\log z}\right)\right|$$

$$\asymp \left|\left(\frac{\left(1 - p^{-1+\frac{i\xi_1-1}{\log z}}\right)\left(1 - p^{-1+\frac{i\xi_2-1}{\log z}}\right)}{\left(1 - p^{-1\frac{-2-i\xi_1-i\xi_2}{\log z}}\right)}\right)^k\right|$$

$$= \left|\prod_p \left(1 - \frac{1}{p^{1+\frac{2+i\xi_1+i\xi_2}{\log z}}}\right)^{-k}\prod_p \left(1 - \frac{1}{p^{1+\frac{1-i\xi_1}{\log z}}}\right)^{k}\prod_p \left(1 - \frac{1}{p^{1+\frac{1-i\xi_2}{\log z}}}\right)^{k}\right|$$

$$= \left|\zeta^k\left(1 + \frac{2+i\xi_1+i\xi_2}{\log z}\right)\zeta^{-k}\left(1 + \frac{1-i\xi_1}{\log z}\right)\zeta^{-k}\left(1 + \frac{1-i\xi_2}{\log z}\right)\right|$$

The crucial observation is is that the Fourier transform of ϕ decays very rapidly. Indeed, integration by parts yields

$$\hat{\phi}(\xi) \ll \frac{1}{(1+|\xi|)^A},$$

for any fixed $A > 0$. (Section 6 in [**25**].) This means that for any $A > 0$

$$
\begin{aligned}
S(g) &\ll \int\int \left|\zeta^k\left(1+\frac{2+i\xi_1+i\xi_2}{\log z}\right)\zeta^{-k}\left(1+\frac{1-i\xi_1}{\log z}\right)\zeta^{-k}\left(1+\frac{1-i\xi_2}{\log z}\right)\right| \frac{1}{(1+|\xi_1|)^A}\frac{1}{(1+|\xi_2|)^A}d\xi_1 d\xi_2 \\
&\ll \int\int (\log z)^{3k}\frac{1}{(1+|\xi_1|)^A}\frac{1}{(1+|\xi_2|)^A}d\xi_1 d\xi_2.
\end{aligned}
$$

When $\max\{|\xi_1|,|\xi_2|\} \geq \sqrt{\log y}$, we use that the product of the zetas above is $\ll (\log z)^{2k}$. Hence, this part of the integral is easily seen to be $\ll \frac{1}{(\log z)^k}$.

Thus, the mass of the integral is concentrated when ξ_1 and ξ_2 are both very small. The integral over

$$\max\{|\xi_1|,|\xi_2|\} \geq \sqrt{\log y}$$

is $\ll 1$. Finally, when $|\xi_1|$,$|\xi_2| \leq \sqrt{\log y}$, we use that $\zeta(s) \sim 1/(s-1)$ for s close to 1. It allows us to conclude

$$\left|\zeta^k\left(1+\frac{2+i\xi_1+i\xi_2}{\log z}\right)\zeta^{-k}\left(1+\frac{1-i\xi_1}{\log z}\right)\zeta^{-k}\left(1+\frac{1-i\xi_2}{\log z}\right)\right| \ll \frac{|1+i\xi_1|^k|1+i\xi_2|^k}{(\log z)^k|2+i\xi_1+i\xi_2|^k}.$$

The main consequence is that this part of the integral is

$$\ll \frac{1}{(\log z)^k}\int\int\frac{|1+i\xi_1|^{-2}|1+i\xi_2|^{-2}}{|2+i\xi_1+i\xi_2|^k}$$

by taking $A = k+2$. □

Now, suppose $y = (qT)^2, w = (qT)^7$ and $z = (qT)^8$. With this in mind, we define the following twisted series :

$$K(s,\chi) = \sum_{n=1}^{\infty}\left(\sum_{d|n}\lambda_d\right)^2\frac{\chi(n)}{n^s}.$$

Lemma 5.1.1. *Suppose* $x = (qT)^{23}$ *and* $qT \geq 64$. *Let the partial sum of* $K(s,\chi)$ *up to* $n = x$ *be denoted* $K_x(s,\chi)$. *Then,*

$$|K(s,\chi) - K_x(s,\chi)| \leq \frac{1}{2}.$$

Proof.

$$
\begin{aligned}
|K(s,\chi) - K_x(s,\chi)| &= \left|\sum_{n=1}^{\infty}\left(\sum_{d|n}\lambda_d\right)^2\frac{\chi(n)}{n^s} - \sum_{1\leq n\leq x}\left(\sum_{d|n}\lambda_d\right)^2\frac{\chi(n)}{n^s}\right| \\
&= \left|\sum_{n>x}\left(\sum_{d|n}\lambda_d\right)^2\frac{\chi(n)}{n^s}\right|
\end{aligned}
$$

By partial summation, if $\chi \neq \chi_0$, then

$$\begin{aligned}\left|\sum_{n=N}^{\infty} \frac{\chi(n)}{n^s}\right| &= \left|\sum_{n=N}^{\infty}\left(\sum_{m=N}^{n} \chi(m)\right)\left(\frac{1}{n^s} - \frac{1}{(n+1)^s}\right)\right| \\ &\leq \sum_{n=N}^{\infty}\left|\sum_{m=N}^{n} \chi(m)\right|\left|\frac{1}{n^s} - \frac{1}{(n+1)^s}\right| \\ &\leq q \sum_{n=N}^{\infty}\left|\frac{1}{n^s} - \frac{1}{(n+1)^s}\right| \\ &\leq \frac{2q|s|}{N^\sigma}\end{aligned}$$

because

$$\left|s \int_n^{n+1} \frac{dy}{y^{s+1}}\right| \leq |s| \int_n^{n+1} \frac{dy}{y^{\sigma+1}}.$$

This means

$$\left|\sum_{n>x} \frac{\chi(n)}{n^s}\right| \leq \frac{2q|s|}{x^\sigma}.$$

Hence, since $m = [b,d] \leq bd \leq yz$, we have

$$\begin{aligned}|K(s,\chi) - K_x(s,\chi)| &\leq \left|\sum_{n>x}\left(\sum_{d,b} \lambda_d \lambda_b\right)\left(\sum_{b,d|n} \frac{\chi(n)}{n^s}\right)\right| \\ &= \left|\left(\sum_{d,b} \frac{\lambda_d \lambda_b}{[b,d]}\right) \sum_{m>x/[b,d]} \frac{\chi(m)}{m^s}\right|\end{aligned}$$

Notice that the last sum can be bounded by the result given on the previous page. Thus,

$$|K(s,\chi) - K_x(s,\chi)| \leq \frac{2q|s|yz}{x^\sigma}.$$

But since $\lambda_d \neq 0$ when $d \leq z$, then $d \leq z$. Also, $|s| = \sqrt{\sigma^2 + T^2}$ and $\alpha < \sigma \leq 1$. Hence, for $T \geq 1$,

$$\begin{aligned}|s| &\leq \sqrt{1 + T^2} \\ &\leq \sqrt{3T^2 + T^2} \\ &= 2T.\end{aligned}$$

Furthermore, $x^{-\sigma} < x^{-\alpha}$ since $\alpha < \sigma \leq 1$. So

$$\begin{aligned}|K(s,\chi) - K_x(s,\chi)| &\leq \frac{2yzq|s|}{x^\sigma} \\ &\leq \frac{2yzq(2T)}{x^\alpha} \\ &= \frac{4qTyz}{x^\alpha}.\end{aligned}$$

In order to conclude this lemma, it is necessary to prove

$$\frac{4qTyz}{x^\alpha} \leq \frac{1}{2}$$

$$\begin{aligned}&\iff 4qT(qT)^2(qT)^8 x^{-\alpha} \leq \frac{1}{2} \\ &\iff (qT)^{11} x^{-\alpha} \leq \frac{1}{8} \\ &\iff (qT)^{11}((qT)^{23})^{-\alpha} \leq \frac{1}{8} \\ &\iff (qT)^{11-23\alpha} \leq \frac{1}{8}.\end{aligned}$$

But $\alpha \geq 1/2$ implies $-\alpha \leq \frac{-1}{2}$ so $11 - 23\alpha \leq 11 - \frac{23}{2} = -\frac{1}{2}$ This means

$$(qT)^{11-23\alpha} \leq \frac{1}{\sqrt{qT}} \leq \frac{1}{8}$$

i.e.

$$qT \geq 64.$$

□

Corollary 5.1.2. *Suppose ρ is a zero of $L(s,\chi)$. Then, for $qT \geq 64$,*

$$\left| \sum_{w<n\leq x} \left(\sum_{d|n} \lambda_d \right)^2 \frac{\chi(n)}{n^\rho} \right| \geq \frac{1}{2}.$$

Proof. Recall

$$K(s,\chi) = \sum_{n=1}^{\infty} \left(\sum_{d|n} \lambda_d \right)^2 \frac{\chi(n)}{n^s}$$

and

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

Let $m = [b,d]$. We suppose that every n in the sum above can be written as $n = tm$. So

$$\begin{aligned} K(s,\chi) &= \sum_{tm=1}^{\infty} \left(\sum_{d|tm} \lambda_d \right)^2 \frac{\chi(tm)}{(tm)^s} \\ &= \sum_{tm=1}^{\infty} \left(\sum_{d|tm} \lambda_d \right)^2 \frac{\chi(t)\chi(m)}{t^s m^s} \\ &= \sum_{t=1}^{\infty} \left(\sum_{m=1}^{\infty} \left(\sum_{[b,d]} \sum_{=m} \lambda_d \lambda_b \right) \frac{\chi(m)}{m^s} \right) \frac{\chi(t)}{t^s}. \end{aligned}$$

This means it is possible to factor out $L(s,\chi)$ and write $K(s,\chi) = L(s,\chi)M(s,\chi)$ with

$$M(s,\chi) = \sum_{m} \left(\sum_{[b,d]} \sum_{=m} \lambda_d \lambda_b \right) \frac{\chi(m)}{m^s}.$$

This factorization is useful because when $L(\rho,\chi) = 0$, then $K(\rho,\chi) = 0$. Hence, by the previous lemma,

$$|K(\rho,\chi) - K_x(\rho,\chi)| = |0 - K_x(\rho,\chi)| = |K_x(\rho,\chi)| \leq \frac{1}{2}.$$

Notice that

$$\begin{aligned} \left(\sum_{d|n} \lambda_d \right)^2 &= \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } 1 < n \leq w \end{cases} \\ &= \sum_{d|n} \mu(d) = 0 \text{ if } 1 < n \leq w. \end{aligned}$$

This allows us to write

$$\begin{aligned} |K_x(\rho,\chi)| &= \left| 1 + \sum_{w<n\leq x} \left(\sum_{d|n} \lambda_d \right)^2 \frac{\chi(n)}{\rho^s} \right| \\ &\leq \frac{1}{2}, \end{aligned}$$

which implies

$$\left| \sum_{w<n\leq x} \left(\sum_{d|n} \lambda_d \right)^2 \frac{\chi(n)}{\rho^s} \right| \geq \frac{1}{2}$$

because if $|1 + A| \leq \frac{1}{2}$, then $|A| > \frac{1}{2}$.

□

This is the zero detector because for every ρ such that $L(\rho,\chi) = 0$, then the value above will be greater or equal to one half. However, it will be supposed that $\chi \neq \chi_0$, because in that case the zero detector will be different.

Let $S(\chi)$ be the set of zeros of $L(s,\chi)$, with multiplicity, in the designated rectangle. Then,

$$R = \sum_{\chi \neq \chi_0} |S(\chi)| = \sum_{\chi \neq \chi_0} N(\alpha,T,\chi).$$

In order to prove the second principle, a bound is needed for $N_q(\alpha,T)$. Indeed,

$$\begin{aligned} N_q(\alpha,T) &= \sum_{\chi \mod q} N(\alpha,T,\chi) \\ &= N(\alpha,T,\chi_0) + \sum_{\chi \neq \chi_0 \mod q} N(\alpha,T,\chi) \\ &= N(\alpha,T,\chi_0) + R. \end{aligned}$$

i.e. We need to show

$$N_q(\alpha,T) = N(\alpha,T,\chi_0) + R \leq c(qT)^{47(1-\alpha)}.$$

However, the first term of this sum, $N(\alpha,T,\chi_0)$, is equal to the number of zeros of the Riemann zeta function in the rectangle. Equation (18.13) in [**19**] states

$$N_q(\alpha,T) \ll (qT)^{12/5(1-\alpha)}(\log qT)^A$$

where A is an absolute constant. Therefore, the inequality given in the second principle is new only for zeros near the line $\Re(s) = 1$, namely for α with

$$1 - \alpha \ll \frac{A \log \mathcal{L}}{\mathcal{L}},$$

where $\mathcal{L} = \log qT$.

Hence, as stated in equation (18.16) in [**19**], we can use the Vinogradov zero-free region (Theorem A.1.7 in appendix) and the Huxley Density Estimate (Theorem A.1.8 in appendix) to conclude

$$N(\alpha,T,\chi_0) \ll T^{3(1-\alpha)}.$$

So, if we show

$$R \ll (qT)^{47(1-\alpha)},$$

then we will be able to conclude

$$\begin{aligned} N_q(\alpha,T) &\ll T^{3(1-\alpha)} + (qT)^{47(1-\alpha)} \\ &\ll (qT)^{47(1-\alpha)}. \end{aligned}$$

This is the desired result. Hence, to conclude the second principle, it is enough to prove $R \ll (qT)^{47(1-\alpha)}$.

Lemma 5.1.3. *Let $S(\chi)$ be the set of zeros of $L(s,\chi)$ in the rectangle defined at the beginning of this section. Then,*

$$R^2 \leq 4U_\alpha V,$$

where

$$U_\alpha = \sum_{w<n\leq x} \left(\sum_{d|n} \lambda_d \right)^2 n^{1-2\alpha}$$

and

$$V = \sum_n f(n) \left(\sum_{d|n} \lambda_d \right)^2 n^{2\alpha-1} \left| \sum_\chi \sum_{s\in S(\chi)} c_\chi(s) \frac{\chi(n)}{n^s} \right|^2$$

for numbers $C_\chi(s)$ with norm one and f a non-negative function such that $f(n) \geq 1$ for $w < n \leq x$.

Proof. Recall that R represents the number of zeros in our rectangle with $\chi \neq \chi_0$. Another way of counting R is to go through the set of zeros counted with multiplicity in our rectangle for which we can detect a large polynomial. We then repeat this step for every χ. Thus, we get

$$\begin{aligned} R &= \sum_{\chi\neq\chi_0} N(\alpha,T,\chi) \\ &\leq 2\sum_\chi \sum_{s\in S(\chi)} \left| \sum_{w<n\leq x} \left(\sum_{d|n} \lambda_d \right)^2 \frac{\chi(n)}{n^s} \right| \\ &\leq 2 \sum_{w<n\leq x} \left| \left(\sum_{d|n} \lambda_d \right)^2 \right| \left| \sum_\chi \sum_{s\in S(\chi)} c_\chi(s) \frac{\chi(n)}{n^s} \right| \end{aligned}$$

for some numbers $c_\chi(s)$ with norm one. The last inequality is obtained by rearranging the sums. Finally, Cauchy-Schwartz's inequality will be applied. Indeed,

$$\begin{aligned} R^2 &\leq \left(2 \sum_{w<n\leq x} \left| \left(\sum_{d|n} \lambda_d \right)^2 \right| \left| \sum_{\chi} \sum_{s \in S(\chi)} c_\chi(s) \frac{\chi(n)}{n^s} \right| \right)^2 \\ &= \left(2 \sum_{w<n\leq x} \left| \left(\sum_{d|n} \lambda_d \right) n^{1/2-\alpha} \right| \left| \left(\sum_{d|n} \lambda_d \right) n^{\alpha-1/2} \right| \left| \sum_{\chi} \sum_{s \in S(\chi)} c_\chi(s) \frac{\chi(n)}{n^s} \right| \right)^2 . \end{aligned}$$

Thus, applying Cauchy-Schwartz yields

$$R^2 \leq 4 \left(\sum_{w<n\leq x} \left(\sum_{d|n} \lambda_d \right)^2 n^{1-2\alpha} \right) \left(\sum_{w<n\leq x} n^{2\alpha-1} \left(\sum_{d|n} \lambda_d \right)^2 \left| \sum_{\chi} \sum_{s \in S(\chi)} \frac{c_\chi(s)\chi(s)}{n^2} \right|^2 \right).$$

So

$$\begin{aligned} R^2 &\leq 4U_\alpha \left(\sum_{w<n\leq x} n^{2\alpha-1} \left(\sum_{d|n} \lambda_d \right)^2 \left| \sum_{\chi} \sum_{s \in S(\chi)} \frac{c_\chi(s)\chi(s)}{n^2} \right|^2 \right) \\ &\leq 4U_\alpha \sum_{n} f(n) \left(\sum_{d|n} \lambda_d \right)^2 n^{2\alpha-1} \left| \sum_{\chi} \sum_{s \in S(\chi)} c_\chi(s) \frac{\chi(n)}{n^s} \right|^2 \end{aligned}$$

for numbers $C_\chi(s)$ with norm one and f a non-negative function such that $f(n) \geq 1$ for $w < n \leq x$. ☐

Lemma 5.1.4. *Let* $x = (qT)^{23}, z = (qT)^8$, $w = (qT)^8$ *and* $1/2 \leq \alpha \leq 1$. *Then,*

$$U_\alpha \ll x^{2(1-\alpha)}.$$

PROOF. By Theorem 1.4 in an article written by Granville, Koukoulopoulos and Maynard [**25**], we have

$$\sum_{w<n\leq x} \left(\sum_{d|n} \lambda_d \right)^2 \ll \frac{x}{\log z}.$$

Hence,

$$\begin{aligned}
U_\alpha &= \sum_{w<n\le x}\left(\sum_{d|n}\lambda_d\right)^2 n^{1-2\alpha}\\
&\le \sum_{w<n\le x}\left(\sum_{d|n}\lambda_d\right)^2 x^{1-2\alpha}\\
&\ll x^{1-2\alpha}\frac{x}{\log z}\\
&\ll x^{2(1-\alpha)}.
\end{aligned}$$

□

The next step will be to find a good bound for V. To do so, we start by developing the square and changing the order of summation. This yields

$$V \le \sum_{\chi_1}\sum_{\chi_2}\sum_{s_1}\sum_{s_2}|B(\chi_1\bar{\chi_2}, s_1+\bar{s_2}+1-2\alpha)|$$

where

$$B(\chi,s) = \sum_n f(n)\left(\sum_{d|n}\lambda_d\right)^2\frac{\chi(n)}{n^s}.$$

We suppose the function f is supported on $[w/v, xv]$ and is continuous, bounded and piecewise monotonic. Thus, using summation by parts, we get

$$\sum_{n\equiv\alpha \mod q}\frac{f(dn)}{(dn)^s} = \frac{F(s)}{dq} + O\left(\frac{|s|v}{w}\right)$$

where

$$F(s) = \int\frac{f(\xi)}{\xi^s}d\xi.$$

The result above helps getting a good bound for V:

Theorem 5.1.3. *Let $S(\chi)$ be the set of zeros of $L(s,\chi)$ in the rectangle defined at the beginning of this section. Then,*

$$V \le \sum_{\chi}\sum_{s_1\in S(\chi)}\sum_{s_2\in S(\chi)}\frac{|F(s_1+\bar{s_2}+1-2\alpha)|}{\log y} + O\left(\frac{R^2Tqvy^2}{w}\right).$$

Here,

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

is a Selberg sieve. For this difficult proof, we refer to [**25**].

Corollary 5.1.5. *Let $S(\chi)$ be the set of zeros of $L(s,\chi)$ in the rectangle defined at the beginning of this section. Then,*

$$V \ll \frac{\log x}{\log y}\sum_{\chi}\sum_{s_1\in S(\chi)}\sum_{s_2\in S(\chi)}\frac{1}{(1+|\gamma_1-\gamma_2|\log v)^2}+\frac{R^2qvy^2}{w}.$$

Proof. We want to find a function f such that, for $\Re(s)\geq 1$, we have

$$F(s)\ll(1+|s-1|\log v)^{-2}\log x. \tag{5.1.2}$$

Indeed, we can take

$$f(\xi)=\begin{cases}\min\left(1-\frac{\log w/\xi}{\log v},1,1-\frac{\log \xi/x}{\log v}\right) & \text{if } w/v\leq\xi\leq xv,\\ 0 & \text{otherwise.}\end{cases}$$

So

$$f(\xi)\log v=\log^+(xv/\xi)-\log^+(w/\xi)+\log^+(w/v\xi),$$

where $f^+=\max(0,f)$.

Since

$$\int_0^\infty(\log^+\xi)\xi^{s-1}d\xi=\frac{1}{s^2},$$

we have

$$\begin{aligned}F(1-s)&=\hat{f}(s)\\&=\int_0^\infty f(\xi)\xi^{s-1}d\xi\\&=\frac{(xv)^s-x^s-w^s+(w/v)^s}{s^2\log v}\\&=\frac{(v^s-1)(x^s-w^2v^{-s})}{s^2\log v}\\&\ll\min\left(\log x,\frac{1}{|s|^2\log v}\right).\end{aligned}$$

This allows us to conclude (5.1.2) holds. Also, for $s=s_1+\bar{s}_2+1-2\alpha$, then $|s-1|=|\beta_1+\beta_2-2\alpha+i(\gamma_1-\gamma_2)|\geq|\gamma_1-\gamma_2|$. Using Theorem 5.1.3 and (5.1.2) concludes the proof. □

Before giving our bound for R, one last lemma is required.

Lemma 5.1.6. *Suppose* $\chi \mod q$ *is a non-trivial character. Let* $1/2 \leq \alpha \leq 1, v \geq 2$ *and* $t \in \mathbb{R}$. *If* A *is an absolute constant, then*

$$\sum_{\substack{L(\rho,\chi) \\ \beta \geq \alpha}} \frac{1}{(1+|\gamma - t|\log v)^2} \ll \frac{1}{2}\left(1-\alpha+\frac{1}{\log v}\right)\log\big(Avq(|t|+1)\big).$$

The proof is given in Lemma 18.3 of [**19**].

Theorem 5.1.4.

$$R \ll (qT)^{47(1-\alpha)}.$$

Proof. We can easily see that $1+\log(v^{1-\alpha}) = 1+(1-\alpha)\log v \leq v^{1-\alpha}$. Putting the result of Lemma 5.1.6 in Corollary 5.1.5 , we may write

$$V \ll \frac{\log x}{\log y}\sum_{\chi}\sum_{s_1 \in S(\chi)} \frac{1}{2}\left(1-\alpha+\frac{1}{\log v}\right)\log Avq(|\gamma_2|+1) + \frac{R^2qvy^2}{w},$$

where γ_2 is defined as in Corollary 5.1.5. Let $v = qT$ and recall the trivial bound $R \ll qT\log qT$. Thus, by Corollary 5.1.5 and Lemma 5.1.6,

$$\begin{aligned}
V &\ll Rv^{1-\alpha}\frac{\log x \log vqT}{\log y \log v} + \frac{R^2qvy^2}{w} \\
&\ll Rv^{1-\alpha}\frac{\log x}{\log y}\frac{\log(qTqT)}{\log qT} + R(qT\log qT)v\frac{((qT)^2)^2}{(qT)^7} \\
&= Rv^{1-\alpha}\frac{\log x}{\log y}\frac{\log((qT)^2)}{\log qT} + Rv\frac{\log qT}{(qT)^2} \\
&\ll Rv^{1-\alpha}\frac{\log x}{\log y} + Rv\frac{\log qT}{(qT)^2}.
\end{aligned}$$

So we have

$$\begin{aligned}
V &\ll Rv^{1-\alpha}\frac{\log((qT)^{23})}{\log((qT)^2)} + Rv^{1-\alpha}\frac{\log qT}{(qT)^2v^{-\alpha}} \\
&= Rv^{1-\alpha}\left(\frac{23}{3} + \frac{\log(qT)v^{\alpha}}{(qT)^2}\right) \\
&\leq Rv^{1-\alpha}\left(\frac{23}{2} + \frac{\log(qT)v}{(qT)^2}\right) \\
&= Rv^{1-\alpha}\left(\frac{23}{2} + \frac{\log(qT)}{(qT)}\right) \\
&\ll Rv^{1-\alpha}.
\end{aligned}$$

But since

$$R^2 \leq 4U_\alpha V,$$

then

$$R^2 \ll 4x^{2(1-\alpha)} R v^{1-\alpha}.$$

It means

$$R \ll ((qT)^{23})^{2(1-\alpha)} (qT)^{1-\alpha} = (qT)^{47(1-\alpha)}.$$

□

5.2. The Third Principle

Theorem 5.2.1 (Exceptional zero repulsion)**.** *There is a positive constant c_3 such that, if the exceptional zero β_1 exists, say $L(\beta_1,\chi_1) = 0$ with*

$$1 - \frac{c_1}{\log qT} \leq \beta_1 < 1,$$

then the function $\prod_{\chi \mod q} L(s,\chi)$ has no other zeros in the region

$$\sigma \geq 1 - c_3 \frac{|\log[(1-\beta_1)\log(qT)]|}{\log qT}, |t| \leq T.$$

Throughout this section, assume $\chi_1 \mod q$ is the exceptional character associated to the exceptional zero β_1 of $L(s,\chi_1)$. We also suppose $\delta_1 = 1 - \beta_1 \leq \frac{c_1}{\log qT}$. However, the third principle will not be proven directly. It suffices to show $\prod_{\chi \mod q} L(s,\chi)$ has no other zeros in the region

$$\sigma \geq 1 - \frac{\log(c_0/(\delta_1 \log(qT)))}{92 \log qT}, |t| \leq T \tag{5.2.1}$$

for an absolute constant $c_0 \geq 2c_1$. Assuming the last statement is true, it is easy to see there is no other zero in the region given by principle 3. Hence, it is enough to prove (5.2.1). The technique we used is very similar to the one employed in the proof of the second principle. A zero detector is created using the function

$$\begin{aligned} \zeta(s)L(s+\delta_1,\chi_1) &= \sum_{n=1}^{\infty} \frac{1}{n^s} \sum_{m=1}^{\infty} \left(\frac{\chi_1(m)}{m^{s+\delta_1}} \right) \\ &= \sum_{n=1}^{\infty} \left(\sum_{a|n} \frac{\chi_1(a)}{a^{\delta_1}} \right) \frac{1}{n^s} \\ &= \sum_{n=1}^{\infty} \frac{\rho(n)}{n^s}. \end{aligned}$$

Once again, $\sum_{d|n} \lambda_d$ is used as mollifier. Thus, for any non-principal character χ, we define the following twisted series:

$$K(s,\chi) = \sum_{n=1}^{\infty} \left(\sum_{d|n} \lambda_d \right)^2 \rho(n) \frac{\chi(n)}{n^s}.$$

Furthermore, recall the partial sum of $K(s,\chi)$ up to $n = x$, denoted by $K_x(s,\chi)$, is defined as

$$K_x(s,\chi) = \sum_{1 \le n \le x} \left(\sum_{d|n} \lambda_d \right)^2 \rho(n) \frac{\chi(n)}{n^s}.$$

Theorem 5.2.2. *Let $s = \sigma + it$ with $\sigma \ge 1/2$ and $|t| \le T$. Then, for $qT \ge 64$,*

$$|K(s,\chi) - K_x(s,\chi)| \le \frac{1}{2}.$$

Proof.

$$\begin{aligned}
|K(s,\chi) - K_x(s,\chi)| &= \left| \sum_{n=1}^{\infty} \left(\sum_{d|n} \lambda_d \right)^2 \rho(n) \frac{\chi(n)}{n^s} - \sum_{1 \le n \le x} \left(\sum_{d|n} \lambda_d \right)^2 \rho(n) \frac{\chi(n)}{n^s} \right| \\
&= \left| \sum_{n > x} \left(\sum_{d|n} \lambda_d \right)^2 \rho(n) \frac{\chi(n)}{n^s} \right| \\
&= \left| \sum_{n > x} \chi(n) \left(\sum_{d|n} \lambda_d \right)^2 \left(\sum_{a|n} \frac{\chi_1(a)}{a^{\delta_1}} \right) \frac{1}{n^s} \right| \\
&= \left| \sum_{n > x} \left(\sum_{d|n} \lambda_d \right)^2 \frac{\chi(n)}{n^s} (1 * t)(n) \right| \\
&= \left| \sum_{n > x} \left(\sum_{d|n} \lambda_d \right)^2 \frac{(\chi * \chi t(n))(n)}{n^s} \right|,
\end{aligned}$$

where $t(n) = \chi_1(n)/n^{\delta_1}$. We need to show this is bounded by $\frac{1}{2}$.

If $\chi = \chi_1$, then $\chi_1^2 = \chi_0$. The function $(\chi_1 * \frac{\chi_1^2}{\mathrm{Id}^{\delta_1}})$ is small:

$$\sum_{n \le N} \left(\chi_1 * \frac{\chi_1^2}{\mathrm{Id}^{\delta_1}} \right)(n) = \sum_{a \le A} \chi_1(a) \sum_{b \le x/a} \frac{1}{b^{\delta_1}} + \sum_{b \le x/A} \frac{1}{b^{\delta_1}} \sum_{A < a \le x/b} \chi_1(a).$$

Furthermore, we can use the Euler- MacLaurin forrmula to bound $\sum_{b\leq x/A}\frac{1}{b^{\delta_1}}$. Indeed, for $s\neq 1$,we have

$$\sum_{k=1}^{n}\frac{1}{k^s}=\zeta(s)+\frac{1}{1-s}n^{1-s}+O\left(\frac{1}{n^s}\right),$$

which gives us

$$\sum_{b\leq x/a}\frac{1}{b^{\delta_1}}\approx\frac{(N/a)^{\delta_1}}{1-\delta_1}.$$

Thus,

$$\begin{aligned}\sum_{n\leq N}\left(\chi_1*\frac{\chi_1^2}{\mathrm{Id}^{\delta_1}}\right)(n)&=\sum_{a\leq A}\chi_1(a)\frac{(N/a)^{\delta_1}}{1-\delta_1}+\sum_{b\leq x/A}\frac{1}{b^{\delta_1}}\sum_{A<a\leq x/b}\chi_1(a)\\&\approx\frac{N^{\delta_1}}{1-\delta_1}\sum_{a\leq A}\frac{\chi_1(a)}{a^{1-\delta_1}}+\sum_{b\leq x/A}\frac{1}{b^{\delta_1}}\sum_{A<a\leq x/b}\chi_1(a)\end{aligned}$$

We know $1-\delta_1=\beta_1$ is a zero of $L(s,\chi_1)$, so by the Pólya-Vinogradov inequality. (Theorem A.1.7 in appendix),

$$\sum_{a\leq A}\frac{\chi_1(a)}{a^{1-\delta_1}}=-\sum_{a>A}\frac{\chi_1(a)}{a^{1-\delta_1}}$$

is small. Also, we know $\sum_{A<a\leq x/b}\chi_1(a)$ is small using the Pólya-Vinogradov inequality. (Theorem A.1.7 in appendix). Since the two terms are little, we can bound the sum of the two terms by one half. □

Theorem 5.2.3. *Let $s=\rho$ be a zero of $L(s,\chi)$ which is different from β_1. Then,*

$$\left|\sum_{w<n\leq x}\left(\sum_{d|n}\lambda_d\right)^2\frac{\rho(n)\chi(n)}{n^\rho}\right|\geq\frac{1}{2}.$$

Proof. Since

$$K(s,\chi)=\sum_{n=1}^{\infty}\left(\sum_{d|n}\lambda_d\right)^2\rho(n)\frac{\chi(n)}{n^s}$$

and

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

we can clearly factorize $L(s,\chi)$ in $K(s,\chi)$.

Using the same techniques we used in the proof of Principle 2, we can write

$$K(s,\chi) = L(s,\chi)L(s+\delta_1, \chi\chi_1)M(s,\chi)$$

with

$$M(s,\chi) = \sum_m \left(\sum_{[b,d]} \sum_{=m} \lambda_d \lambda_b \right) \prod_{p|m} \left(\rho(p) - \frac{\chi(p)}{p^{s+2\delta_1}} \right) \frac{\chi(m)}{m^s}.$$

Indeed, we have

$$K(s,\chi) = \sum_m \left(\sum_{[b,d]} \sum_{=m} \lambda_d \lambda_b \right) \frac{\chi(m)}{m^s} \sum_n \frac{\rho(mn)\chi(n)}{n^s}.$$

Using the definition of $\rho(mn)$, we get

$$\begin{aligned}
\sum_a \frac{\chi_1(a)}{a^\delta} \sum_{n \equiv 0 \mod a/(a,m)} \frac{\chi(n)}{n^s} &= L(s,\chi) \sum_a \frac{\chi_1(a)}{a^{\delta_1}} \chi\left(\frac{a}{(a,m)}\right) \left(\frac{(a,m)}{a}\right)^s \\
&= L(s,\chi) \sum_{c|m} \frac{\chi_1(c)}{c^{\delta_1}} \sum_{(a,c)=1} \frac{\chi_1\chi(a)}{a^{s+\delta_1}} \text{ with } c = (a,m) \\
&= L(s,\chi)L(s+\delta_1, \chi\chi_1) \sum_{c|m} \frac{\chi_1(c)}{c^{\delta_1}} \prod_{p|c} \left(1 - \frac{\chi_1\chi(p)}{p^{s+\delta_1}}\right).
\end{aligned}$$

This gives us the desired factorization because m is square-free. It follows that $K(s,\chi)$ is holomorphic in the whole complex plane. Indeed, if $\chi = \chi_1$, then the pole of $L(s+\delta_1, \chi_0)$ at $s = 1-\delta_1 = \beta_1$ cancels with the zero of $L(s,\chi)$. For $s = \rho$, a zero of $L(s,\chi)$ which is different from β_1 if $\chi = \chi_1$, then by the factorization, we get $K(\rho,\chi) = 0$. By Lemma 5.1.6,

$$|K(\rho,\chi) - K_x(\rho,\chi)| = |K_x(\rho,\chi)| \leq \frac{1}{2}.$$

Recall

$$\begin{aligned}
\left(\sum_{d|n} \lambda_d\right)^2 &= \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } 1 < n \leq w \end{cases} \\
&= \sum_{d|n} \mu(d) \\
&= 0 \text{ if } 1 < n \leq w.
\end{aligned}$$

Since $\lambda_1 = \theta_1 = \rho(1) = 1$, then

$$|K_x(\rho,\chi)| = \left|1 + \sum_{w<n\leq x} \left(\sum_{d|n} \lambda_d\right)^2 \rho(n) \frac{\chi(n)}{n^s}\right| \leq \frac{1}{2}.$$

So

$$\left|\sum_{w<n\leq x} \left(\sum_{d|n} \lambda_d\right)^2 \rho(n) \frac{\chi(n)}{n^s}\right| \geq \frac{1}{2} \tag{5.2.2}$$

with $\rho = \beta + i\gamma, \beta \geq \frac{1}{2}$ and $|\gamma| \leq T$. □

This inequality is our zero detector. Here, $\rho(n)$ is small quite frequently. Furthermore, to simplify the notation, we suppose $\nu(n) = \sum_{d|n} \lambda_d$.

Lemma 5.2.1.

$$16 \left(\sum_{w<n\leq x} \nu^2(n) n^{1-2\beta}\right)^2 \left(\sum_{w<n\leq x} \nu^2(n)\rho^3(n)n^{-1}\right) \left(\sum_{w<n\leq x} \nu^2(n)\rho(n)n^{-1}\right) \geq 1.$$

Proof. It is enough to prove

$$4 \left(\sum_{w<n\leq x} \nu^2(n) n^{1-2\beta}\right) \sqrt{\sum_{w<n\leq x} \nu^2(n)\rho^3(n)n^{-1}} \sqrt{\sum_{w<n\leq x} \nu^2(n)\rho(n)n^{-1}} \geq 1.$$

However, we know that

$$\begin{aligned}
&4 \left(\sum_{w<n\leq x} \nu^2(n) n^{1-2\beta}\right) \sqrt{\sum_{w<n\leq x} \nu^2(n)\rho^3(n)n^{-1}} \sqrt{\sum_{w<n\leq x} \nu^2(n)\rho(n)n^{-1}} \\
&= 4 \left(\sum_{w<n\leq x} \nu^2(n) n^{1-2\beta}\right) \sqrt{\sum_{w<n\leq x} \left(\nu(n)\rho^{3/2}(n)n^{-1/2}\right)^2} \sqrt{\sum_{w<n\leq x} \left(\nu(n)\rho^{1/2}(n)n^{-1/2}\right)^2} \\
&\geq 4 \left(\sum_{w<n\leq x} \nu^2(n) n^{1-2\beta}\right) \sum_{w<n\leq x} \left(\nu(n)\rho^{3/2}(n)n^{-1/2}\right) \left(\nu(n)\rho^{1/2}(n)n^{-1/2}\right) \\
&= 4 \left(\sum_{w<n\leq x} \nu^2(n) n^{1-2\beta}\right) \sum_{w<n\leq x} \nu^2(n)\rho^2(n)n^{-1}
\end{aligned}$$

by applying Hölder's inequality with $p = q = 1/2$ (Theorem A.1.5 in appendix). However, we can write

$$\begin{aligned} 4\left(\sum_{w<n\le x}\nu^2(n)n^{1-2\beta}\right)\sum_{w<n\le x}\nu^2(n)\rho^2(n)n^{-1} &= 4\sum_{w<n\le x}\left(\nu(n)n^{1/2-\beta}\right)^2\sum_{w<n\le x}\left(\nu(n)\rho(n)n^{-1/2}\right)^2 \\ &\ge 4\left(\sum_{w<n\le x}\nu(n)n^{-\beta+1/2}\nu(n)\rho(n)n^{-1/2}\right)^2 \\ &= \left(2\sum_{w<n\le x}\nu^2(n)\rho(n)n^{-\beta}\right)^2 \end{aligned}$$

by applying Hölder's inequality with $p = q = 1/2$ (Theorem A.1.5 in appendix). Finally, applying (5.2.1) to the inequality above yields the desired result. □

Theorem 5.2.4. *Let* $W = \sum_{w<n\le x}\frac{\nu^2(n)\rho(n)}{n}$*. Then,*

$$x^{4(1-\beta)}W \gg 1.$$

Proof. By Lemma 5.2.1,

$$16\left(\sum_{w<n\le x}\nu^2(n)n^{1-2\beta}\right)^2\left(\sum_{w<n\le x}\nu^2(n)\frac{\rho^3(n)}{n}\right)W \ge 1.$$

However, from the proof of the second principle, for any $1/2 \le \alpha \le 1$,

$$\sum_{w<n\le x}\nu^2(n)n^{1-2\alpha} \ll x^{2(1-\alpha)},$$

so it implies

$$\left(\sum_{w<n\le x}\nu^2(n)n^{1-2\beta}\right)^2 \ll (x^{2(1-\beta)})^2 = x^{4(1-\beta)}.$$

For the second term, we can estimate $\rho^3(n)$ by $\tau^3(n)$. Then, if we let $P(x) = \prod_{p\le x}$,

$$\sum_{w<n\le x}\nu^2(n)\frac{\tau^3(n)}{n} \ll \sum_{d|P(x)}\frac{\nu^2(n)f(n)}{n},$$

where $f(p^k) = \tau(p^k)^3$ if $p \ge 11$, and $f(p^k) = 1$ for $p \le 7$. Opening the square and noting that λ is supported on integers $\le z$ yields that

$$\sum_{w<n\le x}\frac{\nu^2(n)f(n)}{n} \ll \sum_{d_1,d_2}\frac{\lambda_{d_1}\lambda_{d_2}}{[d_1,d_2]}\sum_{m|P(x)}\frac{f([d_1,d_2]m)}{m}.$$

The rightmost sum equals $h([d_1,d_2])S$, where $S \asymp (\log x)^8$ and h is multiplicative with $h(p) \le \min\{8, p-1\}$. Here, we use Theorem 5.1.2 to conclude

$$\sum_{w<n\le x} \nu^2(n)\frac{\rho^3(n)}{n} \ll 1.$$

Putting everything together yields

$$\begin{aligned} 1 &\le 16\left(\sum_{w<n\le x} \nu^2(n)n^{1-2\beta}\right)^2 \left(\sum_{w<n\le x} \nu^2(n)\frac{\rho^3(n)}{n}\right) W \\ &\ll x^{4(1-\beta)}W. \end{aligned}$$

□

Corollary 5.2.2.

$$W = \frac{\phi(q)}{q} M(1)L(1+\delta_1,\chi_1)\log\frac{x}{w} + O\left(\frac{1}{q}\right)$$

Proof. Recall

$$W = \sum_{w<n\le x} \frac{\nu(n)^2\rho(n)}{n}.$$

Suppose its sum goes to infinity instead of stopping at $n = x$. Let

$$\begin{aligned} W(s) :&= \sum_{n=1}^{\infty} \frac{\nu(n)^2\rho(n)}{n^s} \\ &= \sum_{n=1}^{\infty}\left(\sum_{d\mid n}\lambda_d\right)^2 \frac{\rho(n)}{n^s} \\ &= L(s,\chi_0)L(s+\delta_1,\chi_1)M(s,\chi_0) \end{aligned}$$

However, this is equal to $K(s,\chi_0)$ by the definition of $K(s,\chi)$ given above.

Contour integration yields a formula for W. Since $W(s)$ is holomorphic everywhere except at $s = 1$, then by Perron's formula

$$\begin{aligned} W &= \frac{1}{2\pi i}\int_{\Re(s)=1/\log x} \frac{W(s+1)(x^s - w^s)}{s} ds \\ &= \mathrm{res}_{s=1} W(s)\log\frac{x}{w} + O\left(\frac{1}{q}\right). \end{aligned}$$

However, since it is a simple pole,

$$\begin{aligned}\operatorname{Res}(W(s),1) &= \lim_{s\to 1}(s-1)W(s)\\ &= \lim_{s\to 1}(s-1)L(s,\chi_0)L(s+\delta_1,\chi_1)M(s)\\ &= \lim_{s\to 1}(s-1)\sum_{n=1}^{\infty}\frac{\mathbf{1_{(n,q)=1}}}{n^s}L(s+\delta_1,\chi_1)M(s)\\ &= \lim_{s\to 1}(s-1)\prod_{p\nmid q}\left(1-\frac{1}{p^s}\right)^{-1}L(s+\delta_1,\chi_1)M(s).\end{aligned}$$

Writing this product as a fraction, we obtain

$$\begin{aligned}\operatorname{Res}(W,1) &= \lim_{s\to 1}(s-1)\frac{\prod_{\text{p prime}}\left(1-\frac{1}{p^s}\right)^{-1}}{\prod_{p\mid q}\left(1-\frac{1}{p^s}\right)^{-1}}L(s+\delta_1,\chi_1)M(s)\\ &= \lim_{s\to 1}(s-1)\frac{\zeta(s)\phi(q)}{q}L(s+\delta_1,\chi_1)M(s)\\ &= \frac{\phi(q)}{q}L(1+\delta_1,\chi_1)M(1,\chi_0)\end{aligned}$$

Putting this last result in the definition of W above gives the desired result. □

Theorem 5.2.5.

$$W \ll \delta_1 \log x.$$

PROOF. By the definition of W given in the proof of Corollary 5.2.2, it suffices to get a good approximation for $M(1,\chi_0)$ and $L(1+\delta_1,\chi_1)$. First of all, using Corrollary 5.5 in [**5**] and Theorem 5.1.2, we get

$$\begin{aligned}M(1,\chi_0) &= \sum_{(m,q)=1}\left(\sum_{[b,d]}\sum_{=m}\lambda_d\lambda_b\right)\prod_{p\mid m}\left(\rho(p)-\frac{\chi(p)}{p^{1+2\delta_1}}\right)\frac{1}{m}\\ &\ll \prod_{\substack{p\le y\\ p\nmid q}}\left(1-\frac{\rho(p)}{p}\right)\\ &= \frac{q}{\phi(q)}\prod_{p\le y}\left(1-\frac{\rho(p)}{p}\right).\end{aligned}$$

To estimate $L(1+\delta_1,\chi_1)$, we use the fact that for $x > y \geq q^2$ and q large enough, then

$$\sum_{y<p\leq x} \frac{1+\chi_1(p)}{p} \leq 4\delta_1 \log x.$$

The inequality above is stated as Lemma 18.4 in [**19**]. By partial summation,

$$L(1+\delta_1,\chi_1) \ll \delta_1 \prod_{p\leq y}\left(1+\frac{\rho(p)}{p}\right).$$

Putting everything together gives

$$\begin{aligned} W &\ll \frac{\phi(q)}{q}\frac{q}{\phi(q)}\prod_{p\leq y}\left(1-\frac{\rho(p)}{p}\right)\delta_1\prod_{p\leq y}\left(1+\frac{\rho(p)}{p}\right)\log\frac{x}{w}+\frac{1}{q} \\ &= \delta_1\prod_{p\leq y}\left(1-\frac{\rho^2(p)}{p^2}\right)\log\frac{x}{w}+\frac{1}{q} \\ &\ll \delta_1 \log x. \end{aligned}$$

Notice that $\delta_1 \log x \geq 1/q$, because $x > y \leq q^2$ and q is big enough. □

Theorem 5.2.6. *Suppose c_1 is a small positive constant such that $c_0 \geq 2c_1$. If $\chi_1 \mod q$ is the exceptional character and β_1 is the exceptional zero of $L(s,\chi_1)$ which satisfies*

$$\delta_1 = 1-\beta_1 \leq \frac{c_1}{\log qT},$$

then there is no other zero in the region

$$\sigma \geq 1-\frac{\log(c_0/\delta_1\log qT)}{92\log qT}, |t|\leq T.$$

Proof. Recall $W = \sum_{w<n\leq x}\frac{\nu(n)\rho(n)}{n}$. Taking $\alpha=\beta_1$, we have

$$x^{4(1-\beta_1)}W \gg 1.$$

But, by Theorem 5.2.5,

$$W \ll \delta_1 \log x.$$

Putting these results together yields

$$1 \ll x^{4(1-\beta_1)}\delta_1 \log x. \tag{5.2.3}$$

Solving (5.2.3) for β_1 concludes the proof. □

Chapter 6

Proof of Linnik's theorem using the clasical approach

In this section, we will give a complete proof of the main theorem which will allow us to deduce Linnik's theorem using the classical method. We assume that c,c_1,c_2,c_3 are the absolute constants from the three principles. Also, suppose β_1 is the exceptional zero if it exists. Recall that $\rho_\chi = \beta_\chi + i\gamma_\chi$ is a zero of $L(s,\chi)$ with $\beta \geq 1/2$, $|T| \leq R$ and $R = x^{1/2c_2}$.

Linnik's theorem can be deduced from the following theorem:

Theorem 6.0.1. *Suppose* $\eta_1 = \frac{c_2}{2\log q}$ *and* $\eta_2 = c_3 \frac{|\log(2\delta_1 \log q)|}{2\log q}$. *For* $x \geq q^{4c_2}$, *then*

$$\psi(x;q,a) = \frac{x}{\phi(q)}\left(1 - \chi_1(a)\frac{x^{\beta_1 - 1}}{\beta_1} + O(cx^{-\eta_i/2})\right).$$

Here we suppose that the β_1 *term does not exist if there is no exceptional zero. Furthermore,* $i = 1$ *if* β_1 *does not exist and* $i = 2$ *otherwise.*

The purpose of this section is to prove the theorem above. First of all, we need to give an approximation for $\psi(x;q,a)$.

Lemma 6.0.1.

$$\psi(x;q,a) = \frac{x}{\phi(q)} - \frac{1}{\phi(q)} \sum_{\chi \mod q} \bar{\chi}(a) \sum_{\substack{L(\rho,\chi)=0 \\ \beta \geq \frac{1}{2}, |\gamma| \leq T}} \frac{x^\rho}{\rho} + O\left(\frac{x\log x}{R}\right).$$

Proof. First of all, recall

$$\psi(x;q,a) = \frac{1}{\phi(q)} \sum_{\chi \mod q} \bar{\chi}(a)\psi(x,\chi).$$

Furthermore, by formula (5.65) in [**19**],

$$\psi(x,\chi) = \delta_x x - \sum_{\substack{L(\rho,\chi)=0 \\ |\gamma|\le T}} \frac{x^\rho - 1}{\rho} + O\left(\frac{x}{T}\log^2(xq)\right)$$

where $1 \le T \le x$ and

$$\delta_x = \begin{cases} 1 & \text{if } \chi = \chi_0, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, by restricting the summation of $\psi(x;q,a)$ to the rectangle given at the beginning of this section, we may get the following truncated formula:

$$\psi(x;q,a) = \frac{x}{\phi(q)} - \frac{1}{\phi(q)} \sum_{\chi \mod q} \bar{\chi}(a) \sum_{\substack{L(\rho,\chi)=0 \\ \beta\ge 1/2, |\gamma|\le R}} \frac{x^{\rho_\chi}}{\rho_\chi} + O\left(\frac{x\log x}{R}\right).$$

□

Theorem 6.0.2. *Suppose the exceptional zero β_1, if it exists, is associated to the real character χ_1. Then,*

$$\psi(x;q,a) = \frac{x}{\phi(q)} - \frac{\chi_1(a)x^{\beta_1}}{\beta_1} - E_{R_c},$$

where

$$E_{R_c} = \frac{1}{\phi(q)} \sum_{\chi \mod q} \bar{\chi}(a) \sum_{T=2^m \le R} \left(\sum_{\substack{L(\rho,\chi)=0, \rho\ne\beta_1 \\ T/2\le|\gamma|\le T}} \frac{x^\rho}{\rho} \right) + O\left(\frac{x\log x}{R}\right).$$

If the exceptional zero does not exist, the second term is assumed to be zero.

Proof. We put the result obtained in Lemma 6.0.1 in the definition of $\psi(x;q,a)$, we obtain

$$\begin{aligned}\psi(x;q,a) &= \frac{1}{\phi(q)} \sum_{\chi \mod q} \bar{\chi}(a) \left(\delta_x x - \sum_{\substack{L(\rho,\chi)=0 \\ |\gamma|\le R}} \frac{x^\rho - 1}{\rho} + O\left(\frac{x}{T}\log^2(xq)\right) \right) \\ &= \frac{1}{\phi(q)} \sum_{\chi \mod q} \bar{\chi}(a)\delta_x x - \frac{1}{\phi(q)} \sum_{\chi \mod q} |\bar{\chi}(a)| \sum_{\substack{L(\rho,\chi)=0 \\ |\gamma|\le R}} \left(\frac{x^\rho - 1}{\rho}\right) \\ &+ O\left(\frac{1}{\phi(q)} \sum_{\chi \mod q} |\bar{\chi}(a)| \frac{x}{T}\log^2(xq) \right).\end{aligned}$$

Hence, we can write

$$\begin{aligned}\psi(x;q,a) &= \frac{x}{\phi(q)} - \frac{1}{\phi(q)} \sum_{\chi \mod q} \bar{\chi}(a) \sum_{\substack{L(\rho,\chi)=0 \\ |\gamma|\le R}} \frac{x^\rho}{\rho} + O\left(\frac{1}{\phi(q)} \sum_{\chi \mod q} |\bar{\chi}(a)| \frac{x}{T} \log^2(xq)\right) \\ &= \frac{x}{\phi(q)} - \frac{1}{\phi(q)} \sum_{\chi \mod q} \bar{\chi}(a) \sum_{\substack{L(\rho,\chi)=0 \\ |\gamma|\le R, \beta\ge 1/2}} \frac{x^\rho}{\rho} + O\left(\frac{x \log x}{R}\right) \\ &= \frac{x}{\phi(q)} - \frac{1}{\phi(q)} \sum_{\chi \mod q} \bar{\chi}(a) \sum_{T=2^m\le R} \left(\sum_{\substack{L(\rho,\chi)=0 \\ T/2\le|\gamma|\le T}} \frac{x^\rho}{\rho}\right) + O\left(\frac{x \log x}{R}\right).\end{aligned}$$

□

In light of Theorem 6.0.2, it suffices to show there exist $c > 0$ such that

$$E_{R_c} \ll \frac{c x^{1-\eta_i/2}}{\phi(q)}$$

in order to conclude Theorem 6.0.1. However, the value of η_i will change depending on wheter the exceptional zero exists or not.

Lemma 6.0.2. *Let $1 - \eta_i$ be the biggest number such that there are no zeros of $L(s,\chi)$ in $1 - \eta_i \le \sigma \le 1$. Suppose β_1 is the an exceptional zero with Dirichlet character χ_1 if it exists. Then,*

$$\left| \sum_{\chi \mod q} \bar{\chi}(a) \sum_{\substack{T/2\le|\gamma_\chi|\le T \\ \rho_\chi\ne\beta_1}} \frac{x^{\rho_\chi}}{\rho_\chi} \right| \le \frac{2}{T}\sqrt{x} N_q\left(\frac{1}{2},T\right) + \frac{2\log x}{T} \int_{1/2}^{1-\eta_i} N_q(\alpha,T) x^\alpha d\alpha.$$

Proof.

$$\begin{aligned}\left| \sum_{\chi \mod q} \bar{\chi}(a) \sum_{\substack{T/2\le|\gamma_\chi|\le T \\ \rho_\chi\ne\beta_1}} \frac{x^{\rho_\chi}}{\rho_\chi} \right| &\le \sum_{\chi \mod q} |\bar{\chi}(a)| \left| \sum_{\substack{T/2\le|\gamma_\chi|\le T \\ \rho_\chi\ne\beta_1}} \frac{x^{\beta_\chi+i\gamma_\chi}}{\beta_\chi + i\gamma_\chi} \right| \\ &\le \sum_{\chi \mod q} \sum_{\substack{T/2\le|\gamma_\chi|\le T \\ \rho_\chi\ne\beta_1}} \left| \frac{x^{\beta_\chi+i\gamma_\chi}}{\beta_\chi + i\gamma_\chi} \right|.\end{aligned}$$

But since $|\rho_\chi| \geq |\gamma_\chi| \geq T/2$,

$$\left| \sum_{\chi \mod q} \bar{\chi}(a) \sum_{\substack{T/2 \leq |\gamma_\chi| \leq T \\ \rho_\chi \neq \beta_1}} \frac{x^{\rho_\chi}}{\rho_\chi} \right| \leq \frac{2}{T} \sum_{\chi \mod q} \sum_{\substack{|\gamma_\chi| \leq T \\ \rho_\chi \neq \beta_1}} x^{\beta_\chi}$$
$$= -\frac{2}{T} \int_{1/2}^{1-\eta_i} x^\alpha dN_q(\alpha,T).$$

The last line is obtained by Theorem A.1.1 and A.1.2 in the appendix. Also, the star above $N_q(\alpha,T)$ means we are excluding the potential exceptional zero. Now, using integration by parts with $dv = dN_q(\alpha,T)$ and $u = x^\alpha$, we get

$$-\frac{2}{T} \int_{1/2}^{1-\eta_i} x^\alpha dN_q^*(\alpha,T) = \frac{2}{T}\sqrt{x} N_q^*\left(\frac{1}{2},T\right) - \frac{2}{T} x^{1-\eta_i} N_q^*(1-\eta_i,T) + \frac{2\log x}{T} \int_{1/2}^{1-\eta_i} N_q^*(\alpha,T) x^\alpha d\alpha.$$

But $N_q^*(1-\eta_i,T) = 0$ which implies

$$-\frac{2}{T} \int_{1/2}^{1-\eta_i} x^\alpha dN_q^*(\alpha,T) = \frac{2}{T}\sqrt{x} N_q^*\left(\frac{1}{2},T\right) + \frac{2\log x}{T} \int_{1/2}^{1-\eta_i} N_q^*(\alpha,T) x^\alpha d\alpha.$$

The desired result is obtained by combining everything. □

Lemma 6.0.3. *Let* $1 - \eta_i$ *be the biggest number such that there are no zeros of* $L(s,\chi)$ *in* $1 - \eta_i \leq \sigma \leq 1$. *Then,*

$$\left| \sum_{\chi \mod q} \bar{\chi}(a) \sum_{\substack{T/2 \leq |\gamma_\chi| \leq T \\ \rho_\chi \neq \beta_1}} \frac{x^{\rho_\chi}}{\rho_\chi} \right| \leq \frac{4cx^{1-\eta_i/2}}{T}.$$

Proof. By Lemma 6.0.2,

$$K = \left| \sum_{\chi \mod q} \bar{\chi}(a) \sum_{\substack{T/2 \leq |\gamma_\chi| \leq T \\ \rho_\chi \neq \beta_1}} \frac{x^{\rho_\chi}}{\rho_\chi} \right| \leq \frac{2}{T}\sqrt{x} N_q^*(1/2,T) + \frac{2\log x}{T} \int_{1/2}^{1-\eta_i} N_q^*(\alpha,T) x^\alpha d\alpha.$$

By the second principle,

$$N_q(\alpha,T) \le cq^{2c_2(1-\alpha)}.$$

So

$$\begin{aligned}\left|\sum_{\chi\neq\chi_1 \mod q}\bar{\chi}(a)\sum_{T/2\le|\gamma_\chi|\le T}\frac{x^{\rho_\chi}}{\rho_\chi}\right| &\le \frac{2c}{T}\sqrt{x}q^{c_2}+\frac{2cx}{T}\log x\int_{1/2}^{1-\eta_i}\frac{q^{2c_2(1-\alpha)}}{x^{1-\alpha}}d\alpha\\ &= \frac{2c}{T}\sqrt{x}q^{c_2}+\frac{2cx\log x}{T}\frac{q^{2c_2}}{x}\int_{1/2}^{1-\eta_i}\left(\frac{x}{q^{2c_2}}\right)^{\alpha}d\alpha\\ &= \frac{2c}{T}\sqrt{x}q^{c_2}+\frac{2c}{T}\log x\frac{q^{2c_2}}{\log(xq^{-2c_2})}\left(\frac{x^{1-\eta_i}}{q^{2c_2(1-\eta_i)}}-\frac{\sqrt{x}}{q^{c_2}}\right)\\ &= \frac{2c}{T}\sqrt{x}q^{c_2}+\frac{2c\log x}{T\log(xq^{-2c_2})}\left(x^{1-\eta_i}q^{2c_2\eta_i}-\sqrt{x}q^{c_2}\right)\\ &\le \frac{2c}{T}\sqrt{x}q^{c_2}+\frac{2c\log x}{T\log(xq^{-2c_2})}x^{1-\eta_i}q^{2c_2\eta_i}-\frac{2c}{T}\sqrt{x}q^{c_2}\\ &= \frac{2c\log x}{T\log(xq^{-2c_2})}x^{1-\eta_i}q^{2c_2\eta_i}\\ &\le \frac{4}{T}cx^{1-\eta_i/2}.\end{aligned}$$

□

Theorem 6.0.3. *Let c_1, c_3 be the two constant disribed in the three principles. Assume $\eta_1 = \frac{c_1}{\log qT}$ and $\eta_2 = c_3\frac{\log 2\delta_1 \log q}{2\log q}$. Then,*

$$E_{R_c} \ll \frac{cx^{1-\eta_i/2}}{\phi(q)},$$

with $i = 1$ if there is an exptional zero β_1 and $i = 2$ otherwise.

Proof. Suppose there is no exceptional zero β_1. Then, according to principle 1, there is no zero in the region $1 - \frac{c_1}{\log qT} \le \sigma \le 1$, which explains why $\eta_1 = \frac{c_1}{\log qT}$. If β_1 exists, then the third principle suggest there is no zero in the region $c_3\frac{\log 2\delta_1 \log q}{2\log q} \le \sigma \le 1$. Hence, $\eta_2 = c_3\frac{\log 2\delta_1 \log q}{2\log q}$.

By Theorem 6.0.2, we have

$$\left|\sum_{\chi \mod q}\bar{\chi}(a)\sum_{\substack{T/2\le|\gamma_\chi|\le T\\ \rho_\chi\neq\beta_1}}\frac{x^{\rho_\chi}}{\rho_\chi}\right| \le \frac{4cx^{1-\eta_i/2}}{T}.$$

$$|E_{R_c}| = \left| \frac{1}{\phi(q)} \sum_{\chi \mod q} \bar{\chi}(a) \sum_{T=2^m \leq R} \left(\sum_{\substack{L(\rho,\chi)=0 \\ T/2 \leq |\gamma_\chi| \leq T}} \frac{x^{\rho_\chi}}{\rho_\chi} \right) \right| + O\left(\frac{x \log x}{R} \right)$$

$$\leq \frac{1}{\phi(q)} \frac{1}{T} \sum_{T=2^m \leq R} \frac{4cx^{1-\eta_i/2}}{\phi(q)} + O\left(\frac{x \log x}{R} \right)$$

$$\ll \frac{4cx^{1-\eta_i/2}}{\phi(q)} + \frac{x \log x}{R}.$$

From the above, we can see that if

$$\frac{x \log x}{R} = x^{1-1/2c_2} \log x \ll \frac{cx^{1-\eta_i/2}}{\phi(q)},$$

then

$$E_R \ll \frac{cx^{1-\eta_i/2}}{\phi(q)}.$$

We know this is true because either $\eta_1 = c_1/2 \log T$ or $\eta_2 = c_3 \frac{\log 2\delta_1 \log q}{2 \log q}$ Thus,

$$E_{R_c} \ll \frac{cx^{1-\eta_i/2}}{\phi(q)}.$$

□

Using the definition of $\psi(x;q,a)$ given in Theorem 6.0.2, we get

$$\left| \psi(x;q,a) - \frac{x}{\phi(q)} + \frac{\chi_1(a)x^{\beta_1}}{\beta_1} \right| = |E_{R_c}| = O\left(\frac{cx^{1-\eta_i/2}}{\phi(q)} \right)$$

which means[1]

$$\psi(x;q,a) = \frac{x}{\phi(q)} - \frac{\chi_1(a)x^{\beta_1}}{\beta_1} + O\left(\frac{cx^{1-\eta_i/2}}{\phi(q)} \right).$$

This yields Theorem 6.0.1.

[1]Here we assume the β_1 term equals 0 if there is no exeptional zero.

6.0.1. Deducing Linnik's theorem

We will deduce Linnik's theorem by proving $\psi(x;q,a) > 0$ for both cases.

Theorem 6.0.4 (Linnik's theorem). *Suppose Theorem 6.0.1 holds. Then,*

$$\psi(x;q,a) > 0.$$

PROOF.

Case 1: No exceptional zero β_1

First of all, by Theorem 6.0.1, when there is no exceptional zero, we can write

$$\psi(x;q,a) = \frac{x}{\phi(q)}\left(1 + O(cx^{-\eta_1/2})\right)$$

and $x^{-\eta_1/2} = e^{-c\frac{\log x}{\log q}}$ is small enough if $x \geq q^L$ for L large enough. Hence, $\psi(x;q,a) > 0$.

Case 2: an exceptional zero β_1

- If $\chi_1(a) = -1$

  By Theorem 6.0.1, we have

$$\begin{aligned}\psi(x;q,a) &= \frac{x}{\phi(q)}\left(1 - \frac{\chi_1(a)x^{\beta_1-1}}{\beta_1} + O(cx^{-\eta_2/2})\right)\\ &= \frac{x}{\phi(q)}\left(1 + \frac{x^{\beta_1-1}}{\beta_1} + O(cx^{-\eta_2/2})\right).\end{aligned}$$

  Here, as in the previous case, we can clearly see that the main term will be bigger than the error term so $\psi(x;q,a) > 0$.

- If $\chi_1(a) = 1$

  One again, by Theorem 6.0.1, when there is an exceptional zero β_1, we have

$$\begin{aligned}\psi(x;q,a) &= \frac{x}{\phi(q)}\left(1 - \frac{\chi_1(a)x^{\beta_1-1}}{\beta_1} + O(cx^{-\eta_2/2})\right)\\ &= \frac{x}{\phi(q)}\left(1 - \frac{x^{\beta_1-1}}{\beta_1} + O(cx^{-\eta_2/2})\right).\end{aligned}$$

We need to demonstrate

$$1 - \frac{x^{\beta_1 - 1}}{\beta_1} + O(cx^{-\eta_2/2}) > 0.$$

We let $\beta_1 = 1 - \delta_1$. It suffices to show

$$1 - \frac{x^{-\delta}}{1 - \delta_1} + O(cx^{-\eta_2/2}) > 0.$$

We know $\delta_1 \leq \frac{c_1}{2 \log q}$ so we define $\delta_1 = \frac{1}{2M_1 \log q}$ with $M_1 > 0$. Furthermore, recall that $\eta_2 = \frac{\log(2\delta_1 \log q)}{2 \log q}$. Plugging δ_1 in the definiton of η_2 yields

$$\eta_2 = \frac{\log M_1}{2 \log q}.$$

Since $x = q^L$, then

$$\begin{aligned}
1 - \frac{x^{\beta_1 - 1}}{\beta_1} + O(cx^{-\eta_2/2}) &= 1 - \frac{q^{-L\delta_1}}{1 - \delta_1} + O(cq^{-\eta_2 L/2}) \\
&= 1 - \frac{q^{\frac{-L}{2M_1 \log q}}}{1 - \delta_1} + O(cq^{\frac{-L \log M_1}{4 \log q}}) \\
&= 1 - \frac{e^{-L/2M_1}}{1 - \delta_1} + O(e^{-L \log M_1/4}) \\
&= \frac{q^{\frac{-L}{2M_1 \log q}}}{1 - \delta_1} + O(cq^{\frac{-L \log M_1}{4 \log q}}) \\
&= 1 - \frac{e^{-L/2M_1}}{1 - \delta_1} + O\left(\frac{1}{M_1^{L/4}}\right).
\end{aligned}$$

Thus,

$$\frac{1 - \delta_1 - e^{-L/2M_1}}{1 - \delta_1} \asymp 1 - e^{-L/2M_1} \asymp \begin{cases} \frac{L}{2M_1} & \text{if } \frac{L}{2M_1} \leq 1 \\ 1 & \text{if } \frac{L}{2M_1} \geq 1. \end{cases}$$

The main term clearly dominates the error term so $\psi(x; q,a) > 0$. Notice that $L = 5$ is enough if M is large. $\square$

Chapter 7

Proof of Linnik's theorem using the pretentious approach

7.1. Proof of the pretentious approach

In this section, we will give a *complete* proof of Linnik's theorem using the pretentious method. The proof of Linnik's theorem is divided into two parts. The first part deals with the contribution of all but one Dirichlet characters. First, recall the definition of the distance function:

Suppose $f(n), g(n)$ are two multiplicative functions on the unit circle. Hence,

$$\mathbb{D}^2(f,g;[q,x]) = \sum_{q<p\leq x} \frac{1-\Re(f\bar{g})(p)}{p}.$$

The following key result can now be obtained.

Theorem 7.1.1. *Let $q \geq 4$. Suppose $\chi_1 \pmod q$ is a real non principal Dirichlet character with $L_q(1,\chi_1) = \min\{L_q(1,\chi) : \chi$ real and non principal character (mod q)$\}$. Then,*

$$\sum_{\substack{y<p\leq z \\ p\equiv a(mod q)}} \frac{1}{p} = \frac{1}{\phi(q)} \sum_{y<\leq z} \frac{1+\chi_1(ap)}{p} + O\left(\frac{1}{\phi(q)}\right),$$

where a is a fixed constant.

Linnik's theorem is usually obtained by using log-free density estimates. As seen previously, new advances allow us to use the theory of multiplicative functions instead. Once the theorem above is established, we can deduce Linnik's theorem from three cases.

We will explain them at the end of this section.

Before proving Theorem 7.1.1, a few results and lemmas will be established.

7.1.1. The main theorem

Lemma 7.1.1. *Letting* $C_q = \{\chi \mod q : \chi \neq \chi_0, \chi_1\}$, *then*

$$\phi(q) \sum_{\substack{y<p\leq z \\ p\equiv a \mod q}} \frac{1}{p} = \sum_{y<\leq z} \frac{1+\chi_1(ap)}{p} + O(1) - E_{R_p},$$

where

$$E_{R_p} = \int_y^z \sum_{\chi\in C_q} \bar{\chi}(a) \frac{L_y'}{L_y}\left(1+\frac{1}{\log u}, \chi\right) \frac{du}{u\log^2 u}.$$

Proof. First of all,

$$\log F_y(s) = \log\left(\prod_{p>y}\left(1-\frac{f(p)}{p^2}\right)^{-1}\right) = \sum_{p>y}\frac{f(p)}{p^s} + O(1)$$

for a multiplicative function f with $|f|\leq 1$ and Dirichlet series F. This means

$$\log L_y\left(1+\frac{1}{\log z}, \chi\right) = \sum_{p>y}\frac{\chi(p)}{p^{1+1/\log z}} + O(1).$$

Using this,

$$\begin{aligned}\sum_{\chi \mod q} \bar{\chi}(a)\log L_y\left(1+\frac{1}{\log z},\chi\right) &= \sum_{\chi \mod q}\bar{\chi}(a)\left(\sum_{p>y}\frac{\chi(p)}{p^{1+1/\log z}} + O(1)\right) \\ &= \sum_{\chi \mod q}\sum_{p>y}\frac{\chi(p)\bar{\chi}(a)}{p^{1+1/\log z}} + O\left(\sum_{\chi \mod q}|\bar{\chi}(a)|\right) \\ &= \phi(q)\sum_{\substack{p>y \\ p\equiv a \mod q}}\frac{1}{p^{1+1/\log z}} + O\left(\sum_{\chi \mod q}|\bar{\chi}(a)|\right),\end{aligned}$$

because

$$\sum_{\chi \mod q}\chi(p)\bar{\chi}(a) = 0$$

unless $p\equiv a \mod q$ and in that case the sum is $\phi(q)$.

Hence,

$$\sum_{\chi \mod q} \bar{\chi}(a) \left(\log L_y\left(1+\frac{1}{\log z},\chi\right) - \log L_y\left(1+\frac{1}{\log y},\chi\right)\right)$$
$$= \phi(q)\left(\sum_{\substack{p>y \\ p\equiv a \mod q}} \frac{1}{p^{1+1/\log z}} - \sum_{\substack{p>y \\ p\equiv a \mod q}} \frac{1}{p^{1+1/\log y}}\right) + O\left(\sum_{\chi \mod q} |\bar{\chi}(a)|\right).$$

Recall Brun-Titchmarsh's inequality (Theorem 3.1.4). Combining this with partial summation gives us

$$\sum_{\substack{p>y \\ p\equiv a \mod q}} \frac{1}{p^{1+1/\log y}} = O\left(\frac{1}{\phi(q)}\right).$$

So

$$\sum_{\substack{p>y \\ p\equiv a \mod q}} \frac{\phi(q)}{p^{1+1/\log z}} - \sum_{\substack{p>y \\ p\equiv a \mod q}} \frac{\phi(q)}{p^{1+1/\log y}} = \phi(q) \sum_{\substack{y<p\leq z \\ p\equiv a \mod q}} \frac{1}{p} + O(1).$$

Putting this in the previous equation,

$$\sum_{\chi \mod q} \bar{\chi}(a) \left(\log L_y\left(1+\frac{1}{\log z},\chi\right) - \log L_y\left(1+\frac{1}{\log y},\chi\right)\right)$$
$$= \phi(q) \sum_{\substack{y<p\leq z \\ p\equiv a \mod q}} \frac{1}{p} + O(1).$$

We can now split the proof into two distinct cases: when $\chi \notin C_q$ and when $\chi \in C_q$.

Case 1: when $\chi \in C_q$

Here we have $L(s,\chi) \asymp 1$. By the Fundamental Theorem of Calculus,

$$\log L_y\left(1+\frac{1}{\log z},\chi\right) - \log L_y\left(1+\frac{1}{\log y},\chi\right) = -\int_y^z \frac{L_y'}{L_y}\left(1+\frac{1}{\log u},\chi\right)\frac{du}{u\log^2 u}.$$

Case 2: when $\chi \notin C_q$

When $\chi \notin \{\chi_0, \chi_1\}$, we do not have $L(s,\chi) \asymp 1$. Also, since $\sigma \leq 1 + 1/\log y$, Theorem 2.2.2 gives $\log F_y(\sigma + it) = \sum_{y<p\leq x} \frac{f(p)}{p^{1+it}} + O(1)$. This means

$$\log L_y\left(1 + \frac{1}{\log z}\right) = \sum_{y<p\leq z} \frac{\chi(p)}{p} + O(1),$$

so

$$\begin{aligned}\log L_y\left(1 + \frac{1}{\log z}\right) - \log L_y\left(1 + \frac{1}{\log y}\right) &= \sum_{y<p\leq z} \frac{\chi(p)}{p} - \sum_{y<p\leq y} \frac{\chi(p)}{p} + O(1) \\ &= \sum_{y<p\leq z} \frac{\chi(p)}{p} + O(1).\end{aligned}$$

Thus, combining Case 1 and Case 2, we get

$$\begin{aligned}\phi(q) \sum_{\substack{y<p\leq z \\ p\equiv a \mod q}} \frac{1}{p} &= \sum_{\chi \mod q} \bar{\chi}(a) \left(\log L_y\left(1 + \frac{1}{\log z},\chi\right) - \log L_y\left(1 + \frac{1}{\log y},\chi\right)\right) \\ &= \sum_{\chi\in C_q} \bar{\chi}(a) \left(\log L_y\left(1 + \frac{1}{\log z},\chi\right) - \log L_y\left(1 + \frac{1}{\log y},\chi\right)\right) \\ &+ \sum_{\chi\notin C_q} \bar{\chi}(a) \left(\log L_y\left(1 + \frac{1}{\log z},\chi\right) - \log L_y\left(1 + \frac{1}{\log y},\chi\right)\right).\end{aligned}$$

Hence,

$$\begin{aligned}\phi(q) \sum_{\substack{y<p\leq z \\ p\equiv a \mod q}} \frac{1}{p} &= -\sum_{\chi\in C_q} \bar{\chi}(a) \int_y^z \frac{L_y'}{L_y}\left(1 + \frac{1}{\log u},\chi\right) \frac{du}{u\log^2 u} + \sum_{\chi\notin C_q} \bar{\chi}(a) \left(\sum_{y<p\leq z} \frac{\chi(p)}{p} + O(1)\right) \\ &= O(1) - E_{R_p} + (1 + \chi_1(ap)) \left(\sum_{y<p\leq z} \frac{1}{p}\right) + O(1) \\ &= \sum_{y<p\leq z} \frac{1 + \chi_1(ap)}{p} + O(1) - E_{R_p}.\end{aligned}$$

□

In light of Lemma 7.1.1, it suffices to prove $E_{R_p} \ll 1$.

Lemma 7.1.2.

$$\begin{aligned}E_{R_p} &= \int_y^z \sum_{\chi\in C_q} \bar{\chi}(a) L_y'(1+1/\log u,\chi)\frac{du}{u\log^2 u} \\ &+ \int_y^z \sum_{\chi\in C_q} \bar{\chi}(a) L_y'\left(1+\frac{1}{\log u},\chi\right) \sum_{\substack{n>y\\P^-(n)>y}} \frac{\mu(n)\chi(n)}{n^{1+1/\log y} u\log^2 u} du \\ &+ \int_{y\le w\le u\le z} \sum_{\chi\in C_q} \bar{\chi}(a) L_y'(1+/\log w,\chi)\frac{L_y'(1+/\log u,\chi)}{L_y^2(1+1/\log w,\chi)}\frac{du}{u\log^2 u}\frac{dw}{w\log^2 w}\end{aligned}$$

Proof. The main idea is to use the Fundamental Theorem of Calculus.

$$\begin{aligned}\frac{1}{L_y(1+1/\log u,\chi)} &= \frac{1}{L_y(1+1/\log y,\chi)} + \int_y^u \frac{L_y'}{L_y^2}\left(1+\frac{1}{\log w},\chi\right)\frac{w}{w\log^2 w} \\ &= 1 + \sum_{\substack{n>y\\P^-(n)>y}} \frac{\mu(n)\chi(n)}{n^{1+1/\log y}} + \int_y^y \frac{L_y'}{L_y^2}\left(1+\frac{1}{\log w},\chi\right)\frac{dw}{w\log^2 w}.\end{aligned}$$

Here, we must remember that

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

Since

$$E_{R_p} = \int_y^z \sum_{\chi\in C_q} \bar{\chi}(a)\frac{L_y'}{L_y}\left(1+\frac{1}{\log u},\chi\right)\frac{du}{u\log^2 u},$$

it is possible to replace $\frac{1}{L_y(1+1/\log u,\chi)}$ in the definition above to get

$$\begin{aligned}E_{R_p} &= \int_y^z \sum_{\chi\in C_q} \bar{\chi}(a) L_y'(1+1/\log u,\chi)\frac{du}{u\log^2 u} \\ &+ \int_y^z \sum_{\chi\in C_q} \bar{\chi}(a) L_y'\left(1+\frac{1}{\log u},\chi\right) \sum_{\substack{n>y\\P^-(n)>y}} \frac{\mu(n)\chi(n)}{n^{1+1/\log y} u\log^2 u} du \\ &+ \int_{y\le w\le u\le z} \sum_{\chi\in C_q} \bar{\chi}(a) \frac{L_y'(1+/\log u,\chi)}{L_y(1+1/\log w,\chi)}\frac{L_y'(1+/\log w,\chi)}{L_y(1+1/\log w,\chi)}\frac{du}{u\log^2 u}\frac{dw}{w\log^2 w}\end{aligned}$$

which is the result we need. □

There are three sums in E_{R_p} and each one of them has to be $\ll 1$.

Corollary 7.1.3.

$$\int_y^z \sum_{\chi\in C_q} \bar{\chi}(a) L_y'(1+1/\log u,\chi)\frac{du}{u\log^2 u} \ll 1.$$

PROOF. Recall that $C_q = \{\chi \mod q : \chi \neq \chi_0, \chi_1\}$ and that χ_1 is an non-principal character. With that in mind, for $u > y$, we have $\sigma = 1 + 1/\log u > 1 - 1/\log y$. Thus, by Theorem 2.3.1, $L_y^{'}(\sigma, \chi) \ll \log y$.

Hence,

$$\sum_{\chi \in C_q} \bar{\chi}(a) L_y^{'}(1 + 1/\log u, \chi) = \sum_{\chi \in C_q, \chi_1} \bar{\chi}(a) L_y^{'}(1 + 1/\log u, \chi) - \bar{\chi}_1(a) L_y^{'}(1 + 1/\log u, \chi_1).$$

But since $L_y^j(f, s) \ll (\log y)^j$, we have

$$\begin{aligned}
\sum_{\chi \in C_q} \bar{\chi}(a) L_y^{'}(1 + 1/\log u, \chi) &= \sum_{\chi \in C_q, \chi_1} \bar{\chi}(a) L_y^{'}(1 + 1/\log u, \chi) + O(\log y) \\
&= \sum_{\chi \neq \chi_0} \bar{\chi}(a) L_y^{'}(1 + 1/\log u, \chi) + O(\log y) \\
&= \phi(q) \left(\sum_{\substack{n > y, P^-(n) > y \\ n \equiv a \mod q}} \frac{\log n}{n^\sigma} - \sum_{\substack{n > y \\ P^-(n) > y}} \frac{\log n}{n^\sigma} \right) + O(\log y).
\end{aligned}$$

Now, applying the Fundamental Lemma of Sieves (Theorem 3.1.2) with $y^u = x^{1/3}$, we can approximate the following set:

$$\#\{n \leq x : P^-(n) > y, n \equiv a \mod q\} = \frac{\#\{n \leq x : P^-(n) > y\}}{\phi(q)} + O\left(\frac{x^{1-1/\log y}}{\phi(q) \log y} + x^{1/3} \right).$$

By partial summation and Theorem 2.3.1,

$$\sum_{\chi \in C_q} \bar{\chi}(a) L_y^{'}(1 + 1/\log u, \chi) \ll \log y.$$

Finally, integration yields

$$\begin{aligned}
\int_y^z \sum_{\chi \in C_q} \bar{\chi}(a) L_y^{'}(1 + 1/\log u, \chi) \frac{du}{u \log^2 u} &\ll \log y \int_y^\infty \frac{du}{u \log^2 u} \\
&= 1.
\end{aligned}$$

□

To bound E_{R_p}'s last two terms, two important lemmas are needed.

Lemma 7.1.4.

$$E_{R_p} \ll 1 + \int_y^z S\left(1+\frac{1}{\log u}\right)^{1/2} W\left(1+\frac{1}{\log y}\right)^{1/2} \frac{du}{u\log^2 u}$$
$$+ \int_{y\le w\le z} S\left(1+\frac{1}{\log u}\right)^{1/2} S\left(1+\frac{1}{\log w}\right)^{1/2} \frac{du}{u\log^2 u}\frac{dw}{w\log^2 w},$$

where

$$S(\sigma) = \sum_{\chi\in C_q} |L_q'(\sigma,\chi)|^2$$

and

$$W(\beta) = \sum_{\chi\in C_q} \left| \sum_{\substack{n>y \\ P^-(n)>y}} \frac{\mu(n)\chi(n)}{n^\beta} \right|^2.$$

Proof. We start by combining Lemma 7.1.2 and Corollary 7.1.3.

$$E_{R_p} = O(1) + \int_y^z \sum_{\chi\in C_q} \bar\chi(a) L_y'\left(1+\frac{1}{\log u},\chi\right) \sum_{\substack{n>y \\ P^-(n)>y}} \frac{\mu(n)\chi(n)}{n^{1+1/\log y}} \frac{du}{u\log^2 u}$$
$$+ \int_{y\le w\le u\le z} \sum_{\chi\in C_q} \bar\chi(a) L_y'(1+/\log u,\chi) \frac{L_y'(1+/\log w,\chi)}{L_y^2(1+1/\log w,\chi)} \frac{du}{u\log^2 u}\frac{dw}{w\log^2 w}.$$

By taking the absolute values, we get

$$E_{R_p} \ll 1 + \int_y^z \sum_{\chi\in C_q} \left| L_y'\left(1+\frac{1}{\log u},\chi\right)\right| \left| \sum_{\substack{n>y \\ P^-(n)>y}} \frac{\mu(n)\chi(n)}{n^{1+1/\log y}} \right| \frac{du}{ulog^2 u}$$
$$+ \int_{y\le w\le u\le z} \sum_{\chi\in C_q} |L_y'(1+/\log u,\chi)| \left| \frac{L_y'(1+/\log w,\chi)}{L_y^2(1+1/\log w,\chi)} \right| \frac{du}{u\log^2 u}\frac{dw}{w\log^2 w}.$$

The implicit zero-free region on C_q is used. Since $|L_y(1+1/\log w,\chi)| \asymp 1$ for $\chi\in C_q$, then $\left|\frac{1}{L_y(1+1/\log w,\chi)}\right| \ll 1$. Thus,

$$E_{R_p} \ll 1 + \int_y^z \sum_{\chi\in C_q} \left| L_y'\left(1+\frac{1}{\log u},\chi\right)\right| \left| \sum_{\substack{n>y \\ P^-(n)>y}} \frac{\mu(n)\chi(n)}{n^{1+1/\log y}} \right| \frac{du}{u\log^2 u}$$
$$+ \int_{y\le w\le u\le z} \sum_{\chi\in C_q} \left|L_y'(1+/\log u,\chi)\right| \left|L_y'(1+/\log w,\chi)\right| \frac{du}{u\log^2 u}\frac{dw}{w\log^2 w}.$$

Applying the Cauchy-Schwartz inequality concludes the proof. ☐

Proof of theorem 7.1.1: By Lemma 7.1.1, it suffices to prove $E_{R_p} \ll 1$. Applying Lemma 4.2.1 gives $S(\sigma) \ll \log^2(y)$ and $W(\beta) \ll 1/(\beta-1)^2 \log^2 y$. Using these results in Lemma 7.1.4 gives

$$\begin{aligned}
E_{R_p} &\ll 1 + \int_y^z S\left(1+\frac{1}{\log u}\right)^{1/2} W\left(1+\frac{1}{\log y}\right)^{1/2} \frac{du}{u\log u^2} \\
&+ \int_{y\le w\le u\le z} S\left(1+\frac{1}{\log u}\right)^{1/2} S\left(1+\frac{1}{\log w}\right)^{1/2} \frac{du}{u\log^2 u}\frac{dw}{w\log^2 w} \\
&\ll 1 + \int_y^z (\log^2 y)^{1/2}\left(\frac{1}{(1+1/\log y - 1)^2\log^2 y}\right)^{1/2} \frac{du}{u\log^2 u} \\
&+ \int_{u=y}^z\int_{w=y}^u \log^2 y \frac{du}{u\log^2 u}\frac{dw}{w\log^2 w} \\
&\le 1 + \log y\int_y^\infty \frac{du}{u\log^2 u} + \log^2 y \int_{u=y}^\infty\int_{w=y}^\infty \frac{du}{u\log^2 u}\frac{dw}{w\log^2 w} \\
&= 3.
\end{aligned}$$

□

Combining the two cases yields Theorem 6.0.1.

7.1.2. Deducing Linnik's theorem using the pretentious approach

Case 1: $\chi_1(a) = -1$

Since χ_1 is a Siegel zero, we have $\sigma \le 1 + 1/\log y$ and thus it is possible to write

$$\sum_{y<p\le z} \frac{\chi_1(p)}{p} = \log L_y(\sigma,\chi_1) + O(1)$$

with $\sigma = 1 + \frac{1}{\log z}$. Furthermore, by Lemma 2.3.1, for $y \ge q(|t|+100)$ and $\sigma > 1 - 1/\log y$, we have $L_y^j(s,\chi) \ll (\log y)^j$ because $\chi \ne \chi_0$. Since $j = 0$, it is possible to bound $\log L_y(1+1/\log z,\chi) \le O(1)$.

It implies $\sum_{y<p\le z} \frac{\chi_1(p)}{p} \le O(1)$. Assuming Theorem 7.1.1, elementary computation shows

$$\begin{aligned}
\frac{1}{\phi(q)}\mathbb{D}^2(1,\chi_1(a)\mu\chi_1;[y,z]) &= \frac{1}{\phi(q)}\mathbb{D}^2(1,-\mu\chi_1;[y,z]) \\
&= \frac{1}{\phi(q)}\mathbb{D}^2(1,\chi_1;[y,z])
\end{aligned}$$

This means we can write

$$\frac{1}{\phi(q)}\mathbb{D}^2(1,\chi_1;[y,z]) = \frac{1}{\phi(q)}\sum_{y<p\le z}\frac{1-\chi_1(p)}{p}$$
$$= \frac{1}{\phi(q)}\left(\sum_{y<p\le z}\frac{1}{p} + O(1)\right)$$
$$= \frac{1}{\phi(q)}\sum_{y<p\le z}\frac{1}{p} + O\left(\frac{1}{\phi(q)}\right)$$
$$= \frac{1}{\phi(q)}\left(\log\left(\frac{\log z}{\log y}\right) + O(1)\right).$$

Merten's theorem will be applied twice (Theorem A.1.4). Now, letting $z = q^L$ and $y = q^2$,

$$\sum_{p\le q^L}\frac{1}{p} = \log(\log q^L) + \gamma + O\left(\frac{1}{\log q^L}\right)$$
$$= \log L + \log\log q + \gamma + O\left(\frac{1}{L\log q}\right).$$

Also,

$$\sum_{p\le q^2}\frac{1}{p} = \log(\log q^2) + \gamma + O\left(\frac{1}{\log q^2}\right)$$
$$= \log 2 + \log\log q + \gamma + O\left(\frac{1}{2\log q}\right).$$

Hence,

$$\sum_{q^2\le p\le q^L}\frac{1}{p} = \log L - \log 2 + O\left(\frac{1}{2\log q} + \frac{1}{L\log q}\right)$$
$$> \log L + O(1).$$

This means

$$\sum_{q^2<p\le q^L}\frac{1+\chi_1(ap)}{p} = 2\mathbb{D}^2(1,\chi_1;[q^2,q^L]) \ge \log L + O(1),$$

which implies, by Theorem 7.1.1,

$$\sum_{\substack{y<p\le z\\ p\equiv a \bmod q}}\frac{1}{p} \ge \frac{\log L}{\phi(q)} > 0.$$

Concretly, this means there exists a computable constant $L \geq 1$ such that for $q \geq 3, (a,q) = 1$, there exists a prime $p \leq q^L$ in a congruence class modulo q. This is exactly the statement of Linnik's theorem.

Case 2: $\chi_1(a) = 1, L_q(1,\chi_1) \geq L^{-0.99}$

Taking $y = q^{L^{0.99}}$ and $z = q^L$, it suffices to apply Theorem 2.3.1. Using this, we conclude

$$\sum_{y<p\leq z} \frac{\chi_1(p)}{p} = O(1).$$

Thus,

$$\begin{aligned}
\sum_{q^{L^{0.99}}<p<q^L} \frac{1+\chi_1(p)}{p} &= \sum_{q^{L^{0.99}}<p<q^L} \frac{1}{p} + O(1) \\
&= \sum_{p<q^L} \frac{1}{p} - \sum_{p<q^{L^{0.99}}} \frac{1}{p} + O(1) \\
&= \log(L \log q) - \log(L^{0.99} \log q) + O(1).
\end{aligned}$$

The last line is due to Merten's theorem. It yields

$$\begin{aligned}
\sum_{q^{L^{0.99}}<p<q^L} \frac{1+\chi_1(p)}{p} &= \log L + \log\log q - \log(L^{0.99}) - \log\log q + O(1) \\
&> 0.01 \log L + O(1) \\
&= \frac{\log L}{100} + O(1).
\end{aligned}$$

If L is big enough, the logarithmic term will dominate the $O(1)$ term. Hence, assuming Theorem 7.1.1,

$$\begin{aligned}
\sum_{\substack{q^{L^{0.99}}<p\leq q^L \\ p\equiv a (mod q)}} \frac{1}{p} &= \frac{1}{\phi(q)} \sum_{q^{L^{0.99}}<p\leq q^L} \frac{1+\chi_1(p)}{p} + O\left(\frac{1}{\phi(q)}\right) \\
&= \frac{1}{\phi(q)} \left(\frac{\log L}{100} + O(1)\right) + O\left(\frac{1}{\phi(q)}\right) \\
&= \frac{\log L}{100\phi(q)} + O\left(\frac{1}{\phi(q)}\right) \\
&> 0
\end{aligned}$$

since the main term dominates the error term.

Case 3: $\chi_1(a) = 1, L_q(1,\chi_1) \leq L^{-0.99}$

Recall that χ_1 is the exceptional character so it is real and non-principal. Thus, we can use Theorem 2.3.4. Let $Q = q^{\frac{1}{L_q(1,\chi_1)}}$. Then,

$$\sum_{q^{\log L} < p \leq q^{L^{0.49}}} \frac{1+\chi_1(p)}{p} \ll \frac{\log q^{L^{0.49}}}{\log q^{1/L_q(1,\chi_1)}} + \frac{1}{L}$$
$$= \frac{L^{0.49}\log q}{\frac{1}{L_q(1,\chi_1)}\log q} + \frac{1}{L}$$
$$= L_q(1,\chi_1)L^{0.49} + \frac{1}{L}$$
$$\leq L^{-0.99+0.49}$$
$$\leq \frac{1}{\sqrt{L}}.$$

Since this bound is very small, it means that for most primes in this interval, $1+\chi_1(p) = 0$, i.e. $\chi_1(p) = -1$. Now, suppose

$$S(X,y;a,q) = \sum_{\substack{n \leq x, n \equiv a mod q \\ p|n \to p > y}} (1 * \chi_1)(n).$$

Notice that

$$S(q^{L^{0.49}}, \sqrt{q^{L^{0.49}}}; a,q) = \sum_{\substack{n \leq q^{L^{0.49}}, n \equiv a mod q \\ p|n \Rightarrow p > \sqrt{q^{L^{0.49}}}}} (1 * \chi_1)(n) > 0 \Rightarrow \text{Linnik's theorem.}$$

The convolution above is used to pre-sift any prime p with $\chi_1(p) = -1$. As seen previously, this is the majority of primes. Thus, it becomes a zero-dimensional sieve problem. The following theorem will be very useful to prove the statement above. It is stated as Theorem in Chapter 20 of [**17**] .

Theorem 7.1.2. *Let* χ_1 *be a real, non principal character* $\mod q$ *and* $X = q^{L^{0.49}}$. *Suppose* $(b,q) = 1, \chi_1(b) = 1, x \in [\sqrt{X},X], 2 \leq y \leq \sqrt{X}$ *and* $u = \frac{\log x}{\log y}$. *Then,*

$$S(x,y;q,b) = (2+\epsilon)xL(1,\chi_1)\prod_{\substack{\ell<y \\ a\nmid q}}\left(1 - \frac{1+\chi_1(\ell)}{l} + \frac{\chi_1(\ell)}{\ell^2}\right)$$

where $\epsilon = O\left(e^{-u} + \frac{1}{\log x}\right)$

Corollary 7.1.5. *Let* $X = q^{L^{0.49}}$. *Then,*

$$S(X,\sqrt{X},q,a) > 0.$$

Proof. Let $y = q^{\log L}$ and $\bar{b}$ be the inverse of $b \mod q$. Recall $P^-(n)$ is the smallest prime factor of p. The first step is to use Buchstab's identity. Now, we can write

$$S(X,\sqrt{X};q,a) = S(X,y;q,a) - \sum_{y<p\leq\sqrt{X}}\sum_{\substack{n\equiv a \mod q \\ P^-(n)=p}}(1*\chi_1)(n)$$
$$= S(X,y;q,a) - \sum_{\substack{y<p\leq\sqrt{X} \\ j\geq 1}}(1*\chi_1)(p^j)S(x/p^j,p;q,\bar{p}^j a).$$

For $j \geq 2$, we can use the fact that $S(x/p^j,p;q,\bar{p}^j a) \leq x/p^j$. It means the contribution is $O(x/y)$. Any other term can be estimated using Theorem 7.1.2. Using the fact that $(1*\chi_1)(p) = 1+\chi_1(p)$, we have

$$S(X,\sqrt{X};q,a) = \frac{2xL(1,\chi_1)(1+O(L^{-1/2}+1/\log x))}{q}\prod_{\substack{\ell<q^{\log L} \\ \ell\nmid q}}\left(1 - \frac{1+\chi_1(\ell)}{\ell} + \frac{\chi_1(\ell)}{\ell^2}\right).$$

Notice that the error term $O(x/y)$ can be absorbed into the other error term. As stated in Theorem 11.5 in [**17**], if χ is a non-principal real character $\mod q$, then

$$L(1,\chi) \gg \frac{1}{\sqrt{q}\log^2 q}.$$

Hence, for L big enough, we have $S(X,\sqrt{X},q,a) > 0$.

□

Bibliography

[1] Ang Li. Dirichlet's theorem about primes in arithmetic progressions *Ann. of Math.* , 2001.

[2] D.R. Heath-Brown, Siegel zeros and the least primes in arithmetic progression, The Quarterly Journal of Mathematics, Volume 41, Issue 4, 1 December 1990, Pages 405–418.

[3] Eric Bach and Jonathan Sorenson, Explicit Bounds for Primes in Residue Classes, Mathematics of Computation,Vol. 65, No. 216 (Oct., 1996), pp. 1717-1735.

[4] University of Scotland, MacTutor History of Mathematics archive: Linnik's biography, http://www-groups.dcs.st-and.ac.uk/history/Biographies/Linnik.html, may 2018.

[5] John Friedlander, Henry Iwaniec, Opera de cribo, Colloquium Publications,Volume: 57; 2010; 527 pp.

[6] Kaisa Matomaki, Maksym Radziwill, Multiplicative functions in short intervals, arXiv:1501.04585 [math.NT], January 2015.

[7] Andrew Granville, Kannan Soundararajan, Large character sums: Pretentious characters and the Polya-Vinogradov Theorem, arXiv:math/0503113 [math.NT], March 2005.

[8] Yoichi Motohashi, An Extension of the Linnik Phenomenon, arXiv:1204.0149v1 [math.NT] 1 Apr 2012.

[9] U. V. Linnik, On the least prime in an arithmetic progression. I. The basic theorem, Rec. Math. [Mat. Sbornik] N.S., 1944, Volume 15(57), Number 2, 139–178.

[10] U. V. Linnik, On the least prime in an arithmetic progression. II. The Deuring–Heilbronn phenomenon, Rec. Math. [Mat. Sbornik] N.S., 1944, Volume 15(57), Number 3, 347–368.

[11] Pan Cheng Tung, On the least prime in an arithmetic progression, Pekin University, New Ser. Vol. I, no 5, 1957.

[12] Matti Jutila, A new estimate for Linnik's constant, Annales academiae scientiarum fenniecae, Series A, I Mathematica, 471.

[13] S.Graham, An asymptotic estimate related to Selberg's sieve, J Number Theory, (1978), 83-94.

[14] David Lowry-Duda, A friendly intro to sieves with a look towards recent progress on the twin primes conjecture, 2000 Mathematics Subject Classification. 11-02; 11N35.

[15] G.J.O. Jameson, Notes on the large sieve.

[16] Andrew Granville, Journal de Théorie des Nombres de Bordeaux 21.1 (2009): 159-173.

[17] Dimitris Koukoulopoulos, The distribution of prime numbers, University of Montreal, 2010 Mathematics Subject Classification. Primary.

[18] Dimitris Koukoulopoulos, Sieve method, University of Montreal, 2015.

[19] Henryk Iwaniec,Emmanuel Kowalski, American Mathematical Society, Colloquium Publications, vol. 53,2004.

[20] S.Graham, An asymptotic estimate related to Selberg's sieve, J. Number Theory 10, (1978),83-94.

[21] Aigner, Martin, Ziegler, Günter M, Proofs from THE BOOK, Springer, 2014.

[22] Andrew Granville, Adam J Harper, K. Soundararajan, A new proof of Halász's Theorem, and its consequences, Submitted on 12 Jun 2017, MSC classes 11N56, 11M20.

[23] U.V. Linnik, On the least prime in an arithmetic progression. I. The basic theorem, N.S., 1944, Volume 15(57),Number 2,139-178.

[24] E.C. Titchmarsh, A divisor problem, Rend. Circ. Mat. Palermo, 54 (1930), 414-429.

[25] Koukoulopoulos, Dimitris, Granville, Andrew and Maynard, James. Sieve weights and their smoothings 2010 Mathematics Subject Classification. Primary: 11N35, 11N37; Secondary: 11T06, 20B30, 05A05.

Appendix A

Useful results

A.1. Appendix

In this section, a few results are stated.They are taken as black boxes.

Theorem A.1.1. *Let* $\alpha : [a,b] \to \mathbb{C}$ *be a step function with a partition* $\mathbb{P} = \{x_0,x_1,...,x_n\}$ *of* $[a,b]$ *such that* α *is constant in each interval of the form* $(x_{j-1},x_j), 1 \leq j \leq n$*. Let*

$$\alpha_j = \begin{cases} \alpha(a^+) - \alpha(a) & \text{if } j = 0 \\ \alpha(x_j^+) - \alpha(x_j^-) & \text{if } 1 \leq j \leq n-1 \\ \alpha(b) - \alpha(b^-) & \text{if } j = n. \end{cases}$$

If $f : [a,b] \to \mathbb{R}$ *is a function such that there is no* $x \in [a,b]$ *for which both* f *and* α *are simultaneously discontinuous from the right or from the left, then*

$$\int_a^b f(x)d\alpha(x) = \sum_{j=0}^{n} f(x_j)\alpha_j.$$

Theorem A.1.2. *(Summation by parts)*
Suppose a_n *is a sequence of complex numbers and* $f \in C^1([a,b])$*. Also, let*

$$\sum_{n \leq x} a_n = M(x) + R(x)$$

where $M \in C^1([a,b])$ *and* R *is the remainder term in the approximation of* $\sum_{n\leq x} a_n$ *by* $M(x)$*. If* $a \leq y \leq z \leq b$*, then*

$$\sum_{y<n\leq z} a_n f(n) = \int_y^z f(t)M'(t)dt + R(z)f(z) - R(y)f(y) - \int_y^z R(t)f'(t)dt.$$

Theorem A.1.3. *(Pólya–Vinogradov inequality, 1918)*

Let

$$S(\chi) = \max_{m,n} \left| \sum_{M<n<N+M} \chi(n) \right|.$$

Then, for any non principal character χ*,*

$$S(\chi) \ll \sqrt{q} \log q.$$

Theorem A.1.4. *(Merten's theorem)*

$$\sum_{p \leq x} \frac{1}{p} = \log\log x + \gamma + O\left(\frac{1}{\log x}\right),$$

where γ *is the Euler-Mascheroni constant.*

Theorem A.1.5. *(Hölder)*

Let $p \geq 1, q \leq \infty$ *such that* $1/p + 1/q = 1$*. Then,*

$$\sum_{k=1}^{n} |x_k y_k| \leq \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p} \left(\sum_{k=1}^{n} |y_k|^q\right)^{1/q}.$$

Theorem A.1.6. *(Plancherel)*

Let

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx.$$

Then

$$\int_{-\pi}^{\pi} |f(x)|^2 dx = 2\pi \sum_{n \in \mathbb{Z}} |c_n|^2.$$

Theorem A.1.7. *(Vinogradov-Korobov, 1957)*

There exists absolute constants $\gamma, \delta > 0$ *such that*

$$\zeta(\sigma + it) \leq \gamma (\log t)^{2/3}.$$

Theorem A.1.8. *(Huxley Density Estimate)*

For any $\alpha > 5/6$ *and* $T \geq 2$*, we have*

$$N(\alpha, T) \ll T^{\frac{3(1-\alpha)}{3\alpha-1}} (\log A)^T,$$

where $A = \frac{300}{(\alpha - 5/6)^2}$.