AN INTRODUCTION TO LOG-FREE ZERO DENSITY ESTIMATES

ASIF ZAMAN

ABSTRACT. Linnik proved his breakthrough result on the least prime in an arithmetic progression in 1944, and one of his pivotal innovations was a log-free zero density estimate. In these notes for the 2023 IPENT summer school, I will introduce log-free zero density estimates, skim some of their history and applications, and explore two essential questions. First, why does "log-free" matter? Second, how can you prove this kind of result?

CONTENTS

1. INTRODUCTION

For $q \geq 1, T \geq 3$, and $0 < \sigma < 1$, define

$$N_q(\sigma, T) := \sum_{\chi \pmod q} \#\{\rho \in \mathbb{C} : L(\rho, \chi) = 0, \sigma < \mathrm{Re}\{\rho\} < 1, |\mathrm{Im}\{\rho\}| \leq T\}. \tag{1.1}$$

Zero density estimates often take the form

$$N_q(\sigma, T) \ll (qT)^{A(1-\sigma)}(\log qT)^B \tag{1.2}$$

for some constants A, B which may depend on σ. There is a vast literature on such results, which we will not attempt to survey here. Efforts are usually spent on minimizing the exponent A for a wide range of σ, while allowing B to be a large absolute constant. However, our focus in this article will be when $B = 0$, namely *log-free* zero density estimates. These take the shape

$$N_q(\sigma, T) \ll (qT)^{C(1-\sigma)} \tag{1.3}$$

These results have a rich history dating back to celebrated works of Linnik [Lin44b, Lin44a] with many contributors since then, including Fogels [Fog65], Turan [Tur61], Bombieri [Bom87], Jutila [Jut69, Jut70, Jut77], Gallagher [Gal70], Graham [Gra81], Heath-Brown [HB92], and many more. In these informal notes, we aim to introduce the method at a high level omitting many technical details while keeping the most essential ideas.

Date: July 6, 2023.

2. Why log-free matters

When first reading these intricate arguments, you might ask:

What is so great about log-free density estimates?

We will explore two problems that demonstrate their greatness.

2.1. **Counting zeros.** Log-free zero density estimates are most advantageous near the line $\mathrm{Re}(s) = 1$ and close to the zero-free region. It is worthwhile to consider an explicit choice.

Let us compare counting zeros in the region

$$1 - \frac{100}{\log qT} < \mathrm{Re}(s) < 1, \qquad |\mathrm{Im}(s)| \leq T.$$

with (1.2) versus (1.3). First, the zero density estimate (1.2) shows that

$$N_q(1 - \tfrac{100}{\log(qT)}, T) \ll e^{100A(\sigma)}(\log qT)^{B(\sigma)}$$

so the number of such zeros is bounded by powers of $\log(qT)$. Unfortunately, when $T = 1$ and $B(\sigma) > 1$, this count is worse than trivial. On the other hand, the *log-free* zero density estimate (1.3) yields

$$N_q(1 - \tfrac{100}{\log(qT)}, T) \ll e^{100C}$$

so the number of such zeros is *absolutely* bounded (in terms of the constant C). This sample region is the prototypical situation where (1.3) supersedes (1.2).

Exercise 1 (Superiority complex). Determine precisely when (1.3) is superior to (1.2). Your answer should be an inequality involving σ. Draw this region in the complex plane and include the critical line, and the classical zero free-region.

2.2. **Counting primes.** Now, after this discussion, you might also justifiably retort:

Okay, I agree that you can sometimes save some logs when counting zeros with a log-free zero density estimate. But who cares? This should be negligible compared to the main contribution. A few logs never hurt anybody.

Surprisingly, every single log mattered for Linnik in 1944 when counting small primes in arithmetic progressions. He famously proved that there exists an absolute constant $L > 0$ such that if $\gcd(a, q) = 1$ then there exists a prime $p \equiv a \pmod q$ and $p \leq q^L$. A log-free density estimate was a critical new ingredient to establishing this spectacular result.

We will sketch a heuristic argument of Linnik's proof to illustrate why a log-free estimate was critical. For simplicity, we will ignore the existence of any Landau-Siegel zeros[1]. For $(a, q) = 1$, fixed $0 < \theta < 1$, and arbitrary $x \geq q$, an explicit formula roughly states that

$$\sum_{\substack{p \leq x \\ p \equiv a \pmod q}} \log p = \frac{x}{\varphi(q)} - \sum_{\chi \pmod q} \frac{\overline{\chi(a)}}{\varphi(q)} \sum_{\substack{\rho = \beta + i\gamma \\ |\mathrm{Im}(\rho)| \leq x^\delta}} \frac{x^\rho}{\rho} + O_\delta(x^{1-\delta}(\log x)^2)$$

Now, the goal is to show that the lefthand side is strictly positive when $x = q^L$ and L is a sufficiently large absolute constant. Dividing both sides by $\frac{x}{\varphi(q)}$, rearranging the main term,

[1]Dealing with Landau-Siegel zeros to count such primes was also a remarkable innovation of Linnik. However, that aspect is not the focus of this article so we have opted to ignore it.

and taking absolute values, it follows that

$$\Big|\frac{\varphi(q)}{x}\sum_{\substack{p\le x\\ p\equiv a \pmod q}}\log p-1\Big|\le\sum_{\chi \pmod q}\sum_{\substack{\rho=\beta+i\gamma\\ |\mathrm{Im}(\rho)|\le x^{\delta}}}\frac{x^{-(1-\beta)}}{|\rho|}+O_{\varepsilon}\Big(\frac{q(\log x)^2}{x^{\delta}}\Big). \tag{2.1}$$

To show that the sum over primes is non-zero, it suffices to prove that the RHS is $<\frac{1}{2}$ once $L=\frac{\log x}{\log q}$ is sufficiently large. Our attention will be entirely on the sum of zeros.

Roughly speaking, for $0<\sigma<1$ and $T\ge 1$, the subsum of zeros $\rho=\beta+i\gamma$ with $\beta\approx\sigma$ and $|\rho|\approx T$ contributes

$$\sum_{\chi \pmod q}\sum_{\substack{\beta\approx\sigma\\ |\gamma|\approx T}}\frac{x^{-(1-\beta)}}{|\rho|}\approx N_q(\sigma,T)\frac{x^{-(1-\sigma)}}{T}\ll\frac{1}{T}\Big(\frac{(qT)^C}{x}\Big)^{1-\sigma} \tag{2.2}$$

by the log-free zero density estimate (1.3). Since $x=q^L$ and $T\le x^{\delta}\le q^{L\delta}$, the above is

$$\le\frac{1}{T}(q^{C+CL\delta-L})^{1-\sigma}=\frac{1}{T}q^{(C+1-L)(1-\sigma)}$$

upon choosing $\delta=\frac{1}{CL}$. These contributions can be heuristically totaled by summing over $\sigma=\sigma_j:=1-\frac{j}{200\log(qT)}$ and $T=T_k:=e^k$ with with $1\le j\le J:=\lceil 100\log(qT)\rceil$ and $1\le k\le K:=\lceil\log(x^{\delta})\rceil$. Notice these choices roughly correspond to

$$\frac{1}{2}\le\sigma\le 1-\frac{1}{200\log(qT)},\qquad T\le x^{\delta}$$

Due to the classical zero-free region and the symmetry of the zeros, this covers all the relevant zeros in the critical strip contributing to the sum over zeros in (2.1).

Overall, this process roughly gives

$$\begin{aligned}\sum_{\chi \pmod q}\sum_{\substack{\rho=\beta+i\gamma\\ |\mathrm{Im}(\rho)|\le x^{\delta}}}\frac{x^{-(1-\beta)}}{|\rho|}&\approx\sum_{j=1}^{J}\sum_{k=1}^{K}\left(\sum_{\chi \pmod q}\sum_{\substack{\beta\approx\sigma_j\\ |\gamma|\approx T_k}}\frac{x^{-(1-\beta)}}{|\rho|}\right)d\sigma\\ &\ll\sum_{j=1}^{J}\sum_{k=1}^{K}\Big(\frac{e^{(C+1-L)j}}{e^k}\Big)\\ &\ll e^{C+1-L}\end{aligned}$$

provided $L>C+1$. Thus, taking L as a sufficiently large absolute constant ensures that the RHS of (2.1) is $<1/2$ when $x=q^L$.

This brings us to the crucial point. If there was a *single* extra $\log q$ from (2.2), then this excess would have propagated to this final estimate. This would force you to take L as a function of q, instead of taking L as a large absolute constant. The proof only succeeds because of a *log-free* zero density estimate. Hallelujah!

Remark 2.1. Notice the terms $x^{-(1-\beta)}/|\rho|$ in (2.1) are largest when β is close to 1 and $|\gamma|$ is small. This corresponds precisely to the ideal scenario described in Section 2.1 for counting zeros with a log-free zero density estimate. We want as few of these large terms as possible.

Exercise 2 (Background check). Prove that (2.1) follows from Theorem 12.10 of Montgomery and Vaughan.

Exercise 3 (Annoying details). The heuristic in (2.2) fails when $|\rho| < 1/4$. Estimate the contribution from these zeros separately. *Hint:* The real parts are very small, and $|\rho|$ cannot be too small.

Exercise 4 (Another crucial ingredient). The classical zero-free region is also essential to Linnik's proof. Explain how the heuristic argument would fail if you did not have any zero-free region. *Hint:* Look at the summation indices.

Exercise 5 (Release the rigour). Use the heuristic argument following (2.1) to formulate a rigorous and complete proof of Linnik's theorem assuming Landau-Siegel zeros do not exist. *Hint:* First divide the region into horizontal strips. On each piece, apply (1.3) using partial summation.

3. Basic strategy

The core ideas behind a log-free zero density estimate are mostly the same as standard zero density estimates. The distinctions are in the details so, before we dive into the novelty, let us do a high-level abstract recap of the two steps towards proving a density estimate. For simplicity, our exposition will assume henceforth

$$T = 100.$$

The letter T might still appear but this assumption of estimating bounded zeros will remain. While some important issues will be ignored, most of the core ideas will be illuminated with this choice alone.

3.1. **Detecting zeros.** The first step is to detect zeros. For each Dirichlet character $\chi \pmod q$, construct a Dirichlet series

$$F(s,\chi) = \sum_{n=1}^{\infty} f_n(s,\chi) \tag{3.1}$$

such that, for a suitably smooth non-negative weight $\phi : \mathbb{R}_{>0} \to [0,\infty)$, we have

$$\Big| \sum_n f_n(\rho,\chi)\phi(n) \Big| \geq \tfrac{1}{2} \tag{3.2}$$

whenever ρ is a non-trivial zero of χ. This initial choice is pivotal and possesses the key ideas, but it might be somewhat mysterious until proceeding through the calculations.

Now, we will take advantage of our simplifying assumption that $T = 100$. By a classical lemma of Linnik, each Dirichlet L-function has $\ll (1-\sigma)\log q$ zeros in the rectangle

$$\sigma < \mathrm{Re}(s) < 1, \qquad |\mathrm{Im}(s)| \leq 100. \tag{3.3}$$

Let $S = S_q(\sigma, T)$ be the set of characters $\chi \pmod q$ which have at least one zero ρ in (3.3). Thus, totaling our zero detectors for each character, it follows that

$$N_q(\sigma, 100) \ll (1-\sigma)\log q \sum_{\chi\in S} \Big| \sum_{n=1}^{\infty} f_n(\rho,\chi)\phi(n) \Big|^2. \tag{3.4}$$

All the zeros have therefore been detected.

3.2. **Applying a large sieve.** Expanding out the RHS of (3.4) and swapping sums, we get

$$\sum_{\chi}\Big|\sum_{n=1}^{\infty} f_n(\rho,\chi)\phi(n)\Big|^2 = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\phi(m)\phi(n)\sum_{\chi} f_m(\rho,\chi)\overline{f_n(\rho,\chi)}. \tag{3.5}$$

We must average over all of these characters and zeros. In an ideal world, the inner double sum over (ρ,χ) will act like an indicator function for whether $m=n$ or not. This wishful thinking comes from the belief that, for fixed $m\neq n$, the values $(f_m(\rho,\chi)\overline{f_n(\rho,\chi)})_\chi$ will be randomly distributed in a complex disk.

Assuming this ideal world occurs, the size of the LHS of (3.5) is roughly the same as only the diagonal terms of the RHS, i.e. when $m=n$. More formally, we hope to prove that

$$\sum_{\chi}\Big|\sum_{n=1}^{\infty} f_n(\rho,\chi)\phi(n)\Big|^2 \le B\sum_{n=1}^{\infty}|\phi(n)|^2 \tag{3.6}$$

for some constant $B=B_q(\sigma,T)>0$.

This averaging process is a kind of a "large sieve". As we shall see, our large sieve must primarily rely on the distribution of character values $(\chi(n))_\chi$ for any fixed $n\in\mathbb{Z}$. These are "well-spaced" by orthogonality of characters. We cannot rely much on the vertical distribution of zeros $\rho=\beta+i\gamma$ since our knowledge is not robust enough. That said, we will need to rely on distinguishing zeros ρ_1,ρ_2 when $|\gamma_1-\gamma_2|$ is somewhat big.

3.3. **Complete the proof.** Combining (3.4) and (3.6), we will get that

$$N_q(\sigma,T)\le 4B\sum_{n=1}^{\infty}|\phi(n)|^2 \tag{3.7}$$

where $B=B_q(\sigma,T)$ is some positive constant. This last sum is usually straightforward to calculate for the given choice of $\phi(n)$. This means we are finished!

Now, you may dispute this conclusion with a natural question.

How are we finished? We do not know anything about the RHS of (3.7).

Indeed, this concern is precisely where all the details come into play. The quantity $B\sum_n|\phi(n)|$ may depend on σ, q, or T but we have not yet indicated how. Based on (1.3), our goal will be to establish that

$$B\sum_{n=1}^{\infty}|\phi(n)|^2 \ll (qT)^{C(1-\sigma)} \tag{3.8}$$

for some positive absolute constant C. Reviewing our 2-step strategy from Sections 3.1 and 3.2, this observation leads to a couple of core questions.

A) *How do we choose the zero detector* (2.1) *to achieve* (3.8)*?*

Broadly speaking, there are two fundamental methods.

- **The mollifier[2] method.** The key tools are arithmetic in nature, inspired by sieve theory and cancellation of partial sums of the Mobius function. Most importantly, the coefficients $f_n(s,\chi)$ are supported on *sifted*[3] *integers* $n\in\mathbb{Z}$.

[2]This method is also known as the "large values approach".

[3]The phrase "sifted integers" means "integers without small prime factors"

- **The power sum method.** The key tools are spectral in nature, inspired by harmonic analysis and Diophantine questions. Most importantly, the coefficients $f_n(s,\chi)$ are supported *only on prime powers* $n \in \mathbb{Z}$.

B) *How can we apply a large sieve* (3.6) *to achieve* (3.8)*?*

There are many variations of the large sieve that succeed, but we will focus on a common thread for some of the best modern expositions: the duality principle. This will allow us to exchange a sum over χ with a dual sum over integers $n \in \mathbb{Z}$. This exchange is critical because sums over non-trivial zeros are mysterious whereas sums over sifted integers are much better understood.

For the sake of brevity, we will focus on the mollifier method. For examples of the power sum method, we refer the reader to [Tur61], [Bom87], or [TZ22]. In either case, the log-free requirement will force us to be ultra-efficient.

4. Mollifier method

Variants of the mollifier method[4] were applied to establish log-free zero density estimates by many authors, including Selberg, Motohashi, Jutila, Graham, and Heath-Brown. Our exposition is modeled after the elegant treatments of Graham and Heath-Brown. We do not aim to provide a fully rigorous proof, but instead outline the key steps and intuition.

4.1. **Motivating the mollifier.** We want to define the Dirichlet series in (3.1) as a product of Dirichlet series, namely

$$F(s,\chi) = L(s,\chi)M(s,\chi) \tag{4.1}$$

where the "mollifier" $M(s,\chi)$ is chosen to intuitively satisfy

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

Abstractly speaking, we therefore want $M(s,\chi)$ to be of the form

$$M(s,\chi) = \sum_{n=1}^{\infty} m_\chi(n)\chi(n)n^{-s},$$

where $m_\chi(n)$ should approximate $\mu(n)$ in some way. A natural choice would be to define $m_\chi(n)$ as a truncated version of the Mobius function but instead. However, the precise choice of $m_\chi(n)$ is actually much more complicated. To motivate its origin, we shift perspective.

The product of Dirichlet series $L(s,\chi)M(s,\chi)$ can be expressed as

$$L(s,\chi)M(s,\chi) = \sum_{n=1}^{\infty}\Big(\sum_{n=de} m_\chi(d)\chi(d)\chi(e)\Big)n^{-s} = \sum_{n=1}^{\infty}\Big(\sum_{d|n} m_\chi(d)\Big)\chi(n)n^{-s}.$$

This reveals the source of the key idea. Indeed, if $m_\chi(d) \approx \mu(d)$ then

$$\sum_{d|n} m_\chi(d) \approx \sum_{d|n}\mu(d) = \begin{cases} 1 & n = 1 \\ 0 & n \neq 1 \end{cases},$$

[4] The approach of "pseudocharacters" pioneered by Selberg can be viewed as a type of mollifier method.

so, since $F(s,\chi) = \sum_n f_n(s,\chi)$ must satisfy (4.1), we should aim for

$$f_n(s,\chi) \approx \Big(\sum_{d|n} \mu(d)\Big)\chi(n). \tag{4.2}$$

Based on (4.2), we might guess that $f_n(s,\chi)$ could be of the form

$$\Big(\sum_{d|n} \psi_d\Big)\chi(n) \tag{4.3}$$

where $(\psi_d)_d$ are some coefficients approximating $\mu(d)$. This will barely fail by a logarithm. Instead, we will ultimately make a more clever choice:

$$f_n(s,\chi) := \Big(\sum_{d|n} \psi_d\Big)\Big(\sum_{d|n} \theta_d\Big)\chi(n) \tag{4.4}$$

where $(\psi_d)_d$ and $(\theta_d)_d$ are coefficients approximating $\mu(d)$ in different ways. This delicate line between success and failure will only be apparent when applying the large sieve. Before proceeding with the proof, we digress on how to approximate the Mobius function.

4.2. **Choosing sieve weights.** Arithmetic approximations for $\sum_{d|n} \mu(d)$ are at the heart of sieve theory, so there are many plausible choices. Our choice of coefficients $(\lambda_d)_d$ approximating $(\mu(d))_d$ will be motivated by some desirable analytic and arithmetic properties.

(i) *The quantity $\sum_{d|n} \lambda_d$ sifts out small prime factors of n.*
In other words, the quantity vanishes if n has a small prime factor. This can be achieved by assuming $\lambda_d = \mu(d)$ for $d \le z$ for a suitably chosen $z > 0$. The value of λ_d for $d > z$ can still be specified in many different ways.

(ii) *The coefficients λ_d smoothly decay to zero as d grows.*
This at least requires $\lambda_d = 0$ for $d > y$ for a suitably chosen $y > z$. The value of λ_d for $z < d < y$ can still be specified in many different ways.

(iii) *The density of $\big(\sum_{d|n} \lambda_d\big)^2$ is at most the density of primes for a long range of $n \in \mathbb{N}$.*
Somewhat more formally, we want a suitably chosen $x \ge y$ to satisfy that

$$\sum_{n\le N} \big(\sum_{d|n} \lambda_d\big)^2 \ll \frac{N}{\log N}$$

for $2 \le N \le x$.

These three properties are essentially encapsulated by a beautiful asymptotic result of Graham [Gra78] on the sieve weights of Barban and Vehov [BV68].

Theorem 4.1 (Graham)**.** *Fix $0 < z \le y$. Define $(\lambda_d)_{d\ge1}$ by*

$$\lambda_d = \begin{cases} \mu(d) & 1 \le d \le z, \\ \mu(d)\frac{\log(y/d)}{\log(y/z)} & z < d \le y, \\ 0 & d > y, \end{cases} \tag{4.5}$$

then for $N \ge 1$,

$$\sum_{n\le N} \Big(\sum_{d|n} \lambda_d\Big)^2 = \frac{N}{\log(y/z)} + O\Big(\frac{N}{(\log(y/z))^2}\Big).$$

Exercise 6 (Choices). Explain why properties (i) to (iii) are fulfilled with the choice $y = z^2$.

Equipped with our sieve weights, we return to the main argument.

4.3. **Constructing zero detectors.** We are ready to formally define our zero detector. Define our universal parameter for measurement:

$$\mathscr{L} := \log(qT).$$

Take parameters $U = (qT)^u, V = (qT)^v, W = (qT)^w, X = (qT)^x$ where $0 < u < v$ and $w, x > 0$ are suitably chosen positive constants. Define $(\psi_d)_d$ and $(\theta_d)_d$ by

$$\psi_d = \begin{cases} \mu(d) & 1 \le d \le U, \\ \mu(d)\frac{\log(V/d)}{\log(V/U)} & U < d \le V, \\ 0 & d > V, \end{cases} \qquad \theta_d = \begin{cases} \mu(d)\frac{\log(W/d)}{\log W} & 1 \le d \le W, \\ 0 & d > W, \end{cases} \tag{4.6}$$

so both of them are examples of Barban-Vehov weights from (4.5). Define

$$F(s, \chi) := \sum_{n=1}^{\infty} \Big(\sum_{d|n} \psi_d\Big)\Big(\sum_{d|n} \theta_d\Big)\chi(n)n^{-s}$$

so indeed the coefficients $f_n(s, \chi)$ are of the form given by (4.4). Due to the choice of weights in (4.6), the mollifier $M(s, \chi)$ can be expressed as a finite Dirichlet polynomial.

Exercise 7 (Dirichlet series game)*.* Show that $F(s, \chi) = L(s, \chi)M(s, \chi)$ where

$$M(s, \chi) = \sum_{v \le V} \sum_{w \le W} \psi_v \theta_w \chi([v, w])[v, w]^{-s}.$$

Define the smooth weight $\phi : \mathbb{R}_{>0} \to [0, \infty)$ by

$$\phi(t) = e^{-t/X} - e^{-t\mathscr{L}^2/U}$$

so that, crudely speaking, $\phi(t) \approx 0$ for $0 < t < U$ and $t > X$, and $\phi(t) \approx 1$ for $U < t < X$. It remains to verify that these choices detect zeros as in (3.2).

Lemma 4.2 (Zero detector)**.** *If ρ is a non-trivial zero of χ with* $\mathrm{Re}(\rho) > 1/2$*, then*

$$\sum_{n=1}^{\infty} \Big(\sum_{d|n} \psi_d\Big)\Big(\sum_{d|n} \theta_d\Big)\chi(n)n^{-\rho}(e^{-n/X} - e^{-n\mathscr{L}^2/U}) = 1 + O(\mathscr{L}^{-1}).$$

provided $x > v + w + \frac{1}{4}$.

Proof. (Sketch) First, notice that

$$\sum_{n=1}^{\infty} \Big(\sum_{d|n} \psi_d\Big)\Big(\sum_{d|n} \theta_d\Big)\chi(n)n^{-\rho}e^{-n/X} = \int_{1-i\infty}^{1+i\infty} L(s + \rho, \chi)M(s + \rho, \chi)\Gamma(s)X^s ds$$

Shift this contour to the line $\mathrm{Re}(s) = \frac{1}{2} - \beta$. No poles are picked up as the integrand is holomorphic for $\mathrm{Re}(s) > -1$. Estimate the new integral using standard divisor bounds $d(n) \ll_\varepsilon n^\varepsilon$, the convexity bound for Dirichlet L-functions $|L(\frac{1}{2} + it, \chi)| \ll (q(1 + |t|))^{\frac{1}{4}}$, and Stirling's estimate $|\Gamma(\frac{1}{2} + it)| \ll e^{-|t|}$. The final error term will be much better than $O(\mathscr{L}^{-1})$.

Second, show that

$$\sum_{n=1}^{\infty}\Big(\sum_{d|n}\psi_d\Big)\Big(\sum_{d|n}\theta_d\Big)\chi(n)n^{-\rho}e^{-n\mathscr{L}^2/U} = e^{-\mathscr{L}^2/U}+\sum_{n>U}\Big(\cdots\Big)$$
$$= 1+O(\mathscr{L}^{-1})$$

Again, the error term can be much better than $O(\mathscr{L}^{-1})$. Finally, combine these two items to deduce the result. □

Exercise 8 (Zero to a hundred). Prove Lemma 4.2 by following the provided outline.

Let $S = S_q(\sigma, T)$ be the set of characters $\chi \pmod q$ which have at least one zero ρ in (3.3). By (3.4) and Lemma 4.2, we have that

$$N_q(\sigma,T) \ll ((1-\sigma)\log q)\,|S| \quad \text{and} \quad |S| \le 4\sum_{\chi}\Big|\sum_{n=1}^{\infty}a_{\chi,n}b_n\Big|^2, \tag{4.7}$$

where

$$a_{\chi,n} = \Big(\sum_{d|n}\theta_d\Big)\chi(n)n^{\frac{1}{2}-\rho}(e^{-n/X}-e^{-n\mathscr{L}^2/U})^{1/2},$$
$$b_n = \Big(\sum_{d|n}\psi_d\Big)n^{-\frac{1}{2}}(e^{-n/X}-e^{-n\mathscr{L}^2/U})^{1/2}.$$

All zeros have been detected, so we are ready for the next step.

4.4. **Applying a large sieve via the duality principle.** In our case of (4.7), the duality principle implies: if there exists $B > 0$ such that

$$\sum_{n}\Big|\sum_{\chi}a_{\chi,n}c_{\chi}\Big|^2 \le B\sum_{\chi}|c_{\chi}|^2 \tag{4.8}$$

for all choices of coefficients c_{χ}, then

$$\sum_{\chi}\Big|\sum_{n}a_{\chi,n}b_n\Big|^2 \le B\sum_{n}|b_n|^2$$

for all choices of coefficients b_n.

Exercise 9 (Dual dualities). The duality principle is a statement about linear operators on Banach spaces. Namely, if $x \mapsto Ax$ is a linear operator between Banach spaces $X \to Y$, then the norm of A is equal to the norm of its adjoint A^*. Show that this formulation implies the above formulation.

Assuming (4.8) holds, we will obtain from Lemma 4.2 and partial summation that

$$\sum_{\chi}\Big|\sum_{n}a_{\chi,n}b_n\Big|^2 \le B\sum_{n=1}^{\infty}\Big(\sum_{d|n}\psi_d\Big)^2 n^{-1}(e^{-n/X}-e^{-n\mathscr{L}^2/U}) \ll B \tag{4.9}$$

since $x > v$ according to Lemma 4.2.

Exercise 10 (Warm up). Verify the calculations justifying (4.9).

Exercise 11 (Missing you). What would happen if the term $(\sum_{d|n}\psi_d)^2$ were missing from the expression in (4.9)? Estimate $\sum_{n=1}^{\infty} n^{-1}(e^{-n/X} - e^{-n\mathscr{L}^2/U})$ with partial summation.

It suffices to establish (4.8) with some constant B. Expanding the LHS of (4.8), all terms of the form

$$c_\chi \overline{c_{\chi'}} \sum_{n=1}^{\infty} \Big(\sum_{d|n} \theta_d\Big)^2 \chi(n)\overline{\chi'(n)} n^{1-\rho-\overline{\rho'}}(e^{-n/X} - e^{-n\mathscr{L}^2/U}).$$

For the non-diagonal terms where $\chi \neq \chi'$, the same ideas from Lemma 4.2 show that

$$\sum_{n=1}^{\infty} \Big(\sum_{d|n} \theta_d\Big)^2 \chi(n)\overline{\chi'(n)} n^{1-\rho-\overline{\rho'}}(e^{-n/X} - e^{-n\mathscr{L}^2/U}) \ll \mathscr{L}^{-1}$$

provided $u > 2w + \frac{1}{4}$. The total contribution of non-diagonal terms is therefore

$$\ll \sum_{\chi}\sum_{\chi'} |c_\chi||c_{\chi'}|\mathscr{L}^{-1} \leq \frac{1}{2}\mathscr{L} \sum_{\chi}\sum_{\chi'} \big(|c_\chi|^2 + |c_{\chi'}|^2\big) = |S|\mathscr{L}^{-1}\Big(\sum_{\chi}|c_\chi|^2\Big)$$

For the diagonal terms, partial summation and Theorem 4.1 imply that

$$\sum_{n=1}^{\infty} \Big(\sum_{d|n} \theta_d\Big)^2 \chi_0(n) n^{1-2\beta}(e^{-n/X} - e^{-n\mathscr{L}^2/U}) \ll |c_\chi|^2 \frac{X^{2-2\beta}}{(2-2\beta)\log W} \ll \frac{X^{2-2\beta}}{(2-2\beta)\log W}$$

The total contribution of all diagonal terms will therefore be

$$\ll \sum_{\chi} |c_\chi|^2 \frac{X^{2-2\beta}}{(2-2\beta)\log W} \ll \frac{X^{2-2\sigma}}{(2-2\sigma)\log W}\Big(\sum_{\chi}|c_\chi|^2\Big).$$

Thus, the choice

$$B = \frac{X^{2-2\sigma}}{(2-2\sigma)\log W} + \frac{|S|}{\mathscr{L}}$$

may be taken in (4.8).

Overall, by (4.7) and (4.9), this shows that

$$N_q(\sigma, T) \ll (1-\sigma)\log q |S| \ll (1-\sigma)\log q \frac{X^{2-2\sigma}}{(2-2\sigma)\log W} \ll q^{C(1-\sigma)}$$

for some absolute positive constant $C > 0$. That was a long and complicated story!

Remark 4.3. Many details were omitted at various steps, especially due to the assumption that $T = 100$. If you are interested in a more detailed, complete, and self-contained exposition, see Iwaniec and Kowalski [IK04] for example.

Exercise 12 (Nitpicking). Confirm the estimates in the final step.

Exercise 13 (A big challenge). Identify what goes wrong when you remove the assumption that $T = 100$, especially near the end of the proof. There is more than one thing. How might you try to resolve these issues?

References

[Bom87] Enrico Bombieri. Le grand crible dans la théorie analytique des nombres. *Astérisque*, (18):103, 1987.

[BV68] M. B. Barban and P. P. Vehov. An extremal problem. *Trudy Moskov. Mat. Obšč.*, pages 83–90, 1968.

[Fog65] E. Fogels. On the zeros of L-functions. *Acta Arith*, pages 67–96, 1965.

[Gal70] P. X. Gallagher. A large sieve density estimate near $\sigma = 1$. *Invent. Math.*, 11:329–339, 1970.

[Gra78] S. Graham. An asymptotic estimate related to Selberg's sieve. *J. Number Theory*, 10(1):83–94, 1978.

[Gra81] S. Graham. On Linnik's constant. *Acta Arith.*, 39(2):163–179, 1981.

[HB92] D. R. Heath-Brown. Zero-free regions for Dirichlet L-functions, and the least prime in an arithmetic progression. *Proc. London Math. Soc. (3)*, 64(2):265–338, 1992.

[IK04] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[Jut69] Matti Jutila. On two theorems of Linnik concerning the zeros of Dirichlet's L-functions. *Ann. Acad. Sci. Fenn. Ser. A I No.*, page 32, 1969.

[Jut70] Matti Jutila. A new estimate for Linnik's constant. *Ann. Acad. Sci. Fenn. Ser. A I No.*, page 8, 1970.

[Jut77] Matti Jutila. On Linnik's constant. *Math. Scand.*, 41(1):45–62, 1977.

[Lin44a] U. V. Linnik. On the least prime in an arithmetic progression. I. The basic theorem. *Rec. Math. [Mat. Sbornik] N.S.*, pages 139–178, 1944.

[Lin44b] U. V. Linnik. On the least prime in an arithmetic progression. II. The Deuring-Heilbronn phenomenon. *Rec. Math. [Mat. Sbornik] N.S.*, pages 347–368, 1944.

[Tur61] P. Turán. On a density theorem of Yu. V. Linnik. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 6:165–179, 1961.

[TZ22] Jesse Thorner and Asif Zaman. An explicit version of bombieri's log-free density estimate and sárközy's theorem for shifted primes, 2022.

Asif Zaman, Department of Mathematics, University of Toronto, 40 St. George Street, Room 6290, Toronto, ON M5S 2E4, CANADA

Email address: `asif.zaman@utoronto.ca`