Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

Robin Fissum

An introduction to primes in short intervals

Master's thesis in MSMNFMA
Supervisor: Kristian Seip
May 2023

$$\zeta(s)$$

NTNU
Norwegian University of
Science and Technology

Robin Fissum

An introduction to primes in short intervals

Master's thesis in MSMNFMA
Supervisor: Kristian Seip
May 2023

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

NTNU
Norwegian University of
Science and Technology

Abstract

In this thesis, we give an introduction to the study of primes in short intervals: the prime number theorem is equivalent to the statement that the number of primes in the interval $(x, x+h]$ is asymptotic to $h/\log x$ when $h = x$, and we investigate if not the same can be said for $h < x$.

We begin by discussing the Riemann zeta function, and how the Riemann hypothesis, Lindelöf hypothesis and the density hypothesis are related to its zeros. Afterwards, we derive the explicit formula for the second Chebyshev function, and relate this to the number of primes in an interval. Together with some moderate assumptions on the zeros of zeta, we prove Hoheisel's theorem: the assertion that $h = o(x)$ is permissible. Next, the Riemann hypothesis is assumed, and we show that an asymptotic formula then holds when $\sqrt{x}\log x = o(h)$; we also give an account of Cramér's theorem for prime gaps. Towards the end, we show how Selberg's mean-square method gives insight into the case $h < \sqrt{x}$, and some of the associated obstacles and heuristics for intervals of this length.

Sammendrag (abstract in Norwegian)

I denne oppgaven gir vi en introduksjon til studiet av primtall i korte intervaller: Primtallsatsen er ekvivalent med utsagnet om at antall primtall i intervallet $(x, x+h]$ er asymptotisk med $h/\log x$ når $h = x$, og vi undersøker om ikke det samme kan sies for $h < x$.

Vi begynner med å diskuter Riemanns zetafunksjon, og hvordan Riemannhypotesen, Lindelöf-hypotesen og tetthetshypotesen er relatert til dens nullpunkter. Deretter utleder vi den eksplisitte formelen for den andre Chebyshevfunksjonen, og relaterer denne til antall primtall i et intervall. Sammen med noen moderate antagelser om nullpunktene hos zeta, beviser vi Hoheisels teorem: utsagnet om at vi kan tillate oss $h = o(x)$. Videre antar vi Riemannhypotesen, og viser at en asymptotisk formel da holder når $\sqrt{x}\log x = o(h)$; vi kommer også inn på Cramérs teorem for primtallsgap. Mot slutten viser vi hvordan Selbergs kvadratgjennomsnittsmetode gir innsikt i tilfellet $h < \sqrt{x}$, samt noen av de assosierte utfordringene og formodningene rundt intervall av denne lengden.

Preface

This is my master's thesis in pure mathematics, which in practice means that it is filled symbols that are hard to typeset and conclusions that may seem distanced from reality. It was written as a compulsory part of the master's programme in pure mathematics at NTNU, The Norwegian University of Science and Technology, during the autumn 2022–spring 2023 period. That is, during this period a lot of material was read and attempted understood, many libraries were scouted for documents from the early 20th century, and at least seven books were utilised in parallel for composing this work. The project was supervised by Professor Kristian Seip, for which he deserves a special thanks. Indeed, we have had many fruitful discussions throughout the course of the project, and I would like to thank Seip for suggesting the topic of the project and pointing out the obvious mistakes in my calculation. This work would also not have been possible without useful feedback from Professor Harald Hance-Olsen, support from my family and friends and the expertise that I have borrowed from the MathOverflow community. Thank you.

In some sense, this work is one possible extension of Dalaker's bachelor's thesis [8], and it is also the natural next step after my own bachelor's thesis [12]. I therefore hope that the reader will enjoy (or, survive) it as much as I did.

Robin Fissum
Skaun, 2023

Contents

Abstract **i**

Preface **ii**

Contents **iii**

1 Introduction **1**
1.1 The Riemann zeta function 1
1.2 Some functions in analytic number theory 2
1.3 Asymptotic analysis 3

2 Three conjectures on the zeta function **4**
2.1 The number of nontrivial zeros 4
2.2 Growth of zeta on vertical lines 8
2.3 Zero-density estimates 12
2.4 Relationships between the hypotheses 14

3 The explicit formula for $\psi(x)$ **15**
3.1 Some lemmas 15
3.2 The explicit formula 21

4 Primes in short intervals: weak hypotheses **28**
4.1 Introduction 28
4.2 Hoheisel's theorem and generalisations 30

5 Primes in short intervals: the Riemann hypothesis **41**
5.1 Prime number theorem under the Riemann hypothesis 41
5.2 Cramér's theorem 50

6 Primes in short intervals: beyond the Riemann hypothesis **53**
6.1 The work of Selberg 53
6.2 Further investigations 58

Appendix **60**
A.1 Primes in long intervals 60
A.2 Miscellaneous results 61

References **63**

1 Introduction

In this thesis we investigate the ancient question:

How many primes are there in an interval?

Before we can get to the bottom of this, we give a short overview of the fundamental concepts and notation in this section. Afterwards, we discuss the Riemann, Lindelöf and density hypotheses in Section 2, and some relationships between these; this is necessary in order to understand the bigger picture. In Section 3 we prove the explicit formula for the second Chebyshev function, a fundamental tool for our investigation, which we use in Sections 4, 5 and 6, where we attempt to give a partial answer to the question above.

Our journey begins with the zeta function.

1.1 The Riemann zeta function

Throughout the text, $s = \sigma + ti$ denotes a complex number in which σ and t are real numbers; this has become standard practice in the field. The *Riemann zeta function*, $\zeta(s)$, is the function defined by

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

whenever $\sigma > 1$, and extended to a meromorphic function on $\mathbb{C}$ by means of analytic continuation.[1] This function is regular, that is, holomorphic and single-valued, except for a simple pole of residue 1 at $s = 1$, and satisfies everywhere the functional equation

$$\zeta(s) = 2^s \pi^{s-1} \sin\left(\frac{\pi s}{2}\right) \Gamma(1-s) \zeta(1-s), \tag{1}$$

or, in its symmetric form,

$$\zeta(s)\Gamma\left(\frac{s}{2}\right)\pi^{-s/2} = \zeta(1-s)\Gamma\left(\frac{1-s}{2}\right)\pi^{-(1-s)/2}, \tag{2}$$

with $\Gamma(s)$ denoting the usual gamma function.

The main reason that the zeta function is so important in number theory is its connection with prime numbers. This is because of the *Euler product*,

$$\zeta(s) = \prod_p (1 - p^{-s})^{-1}, \tag{3}$$

valid for $\sigma > 1$, where the product on the right runs over the prime numbers. Indeed, this identity is an analytic formulation of the fundamental theorem of arithmetic: 'analytic', since the left-hand side is a meromorphic function of a complex variable, and 'fundamental theorem of arithmetic', because the left- and right-hand sides are defined as a sum over the positive integers and as product over the primes, respectively.

Of particular importance is the question about the location of the zeros of $\zeta(s)$. From (3) one can deduce that $\zeta(s)$ is never zero for $\sigma > 1$, and from (1) it is straightforward to see that $\zeta(s)$ is zero whenever s is a negative even integer; these zeros are therefore known as the *trivial zeros* of $\zeta(s)$. Any other zero of $\zeta(s)$ is known as a *nontrivial zero*, which we denote in general by

$$\rho = \beta + \gamma i, \tag{4}$$

[1] n^{-s} is defined as $\mathrm{e}^{-s \log n}$, where log denotes the principal branch of the logarithm.

with β and γ real. A precise determination of the location of the nontrivial zeros would have far-reaching consequences in number theory. Indeed, the statement that $\zeta(s)$ has no zeros on the line $\sigma = 1$ is known to be equivalent to the prime number theorem, that is, to the assertion that

$$\pi(x) \sim \frac{x}{\log x}$$

as $x \to \infty$, where $\pi(x)$ denotes the prime counting function, and '$\sim$' an asymptotic equivalence (defined below).

The statements above, together with the relation $\zeta(\overline{s}) = \overline{\zeta(s)}$, show that the nontrivial zeros (4) must be confined to the vertical strip $0 < \sigma < 1$, lie symmetrically about (but not on) the real axis, and lie symmetrically about the line $\sigma = \frac{1}{2}$. This vertical strip is known as the *critical strip*, and the line $\sigma = \frac{1}{2}$, as the *critical line*. Riemann [29] conjectured that all of the nontrivial zeros lie on the critical line, a still-unproven statement now known as the *Riemann hypothesis:*

Conjecture 1.1 (Riemann hypothesis). *If $\rho = \beta + \gamma i$ is a nontrivial zero of the Riemann zeta function, where β and γ are real numbers, then $\beta = \frac{1}{2}$.*

1.2 Some functions in analytic number theory

Throughout the paper $\mathbb{N}$, $\mathbb{N}^+$, $\mathbb{R}$, $\mathbb{C}$, $\mathcal{P}$ and $\mathcal{P}^*$ denote the collections of nonnegative integers, positive integers, real numbers, complex numbers, prime numbers and prime powers, respectively. We sometimes write p_n for the nth prime, so that $p_1 = 2$, $p_2 = 3$, ... , and γ_n for the nth positive ordinate among the nontrivial zeros ρ, where $0 < \gamma_1 \leq \gamma_2 \leq \ldots$.

In addition to $\Gamma(s)$ denoting the gamma function and $\lfloor x \rfloor$ the greatest integer less than or equal to x, we define

$$\pi(x) = \sum_{p \leq x} 1 \qquad (\textit{the prime-counting function}),$$

$$\psi(x) = \sum_{p^m \leq x} \log p \qquad (\textit{the second Chebyshev function}),$$

$$\Pi(x) = \sum_{p^m \leq x} \frac{1}{m} \qquad (\textit{Riemann's auxiliary function}), \text{ and}$$

$$\Lambda(x) = \begin{cases} \log p & \text{if } x = p^m \text{ for } p \in \mathcal{P} \text{ and } m \in \mathbb{N}^+ \\ 0 & \text{otherwise} \end{cases} \qquad (\textit{the von Mangoldt function}),$$

for $x \in \mathbb{R}$, where $p \leq x$ and $p^m \leq x$ means that the summation is taken over all primes or prime powers less than or equal to x, respectively, and where an empty sum is defined equal to zero.

We also define (the *Landau xi function*)

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

Then $\xi(s)$ is an entire function, and the zeros of $\xi(s)$ are precisely the nontrivial zeros of $\zeta(s)$ (with the same multiplicities). Moreover, Equation (2) shows that $\xi(s)$ satisfies the particularly simple functional equation

$$\xi(s) = \xi(1-s).$$

Other functions will be specified when needed.

1.3 Asymptotic analysis

When we write $f(x) = O(g(x))$ or $f(x) \ll g(x)$, then we mean that there exists a constant $C > 0$ such that $|f(x)| \leq C|g(x)|$ for all values of x under consideration, typically for all $x \geq x_0$. The constant C (not unique) is sometimes referred to as the *implied constant* of the relation. If f and g depend on an additional parameter, such as ε, then so may C and x_0, in which case we write $f(x) = O_\varepsilon(g(x))$ or $f(x) \ll_\varepsilon g(x)$. The statement that $f(x, y) = O(g(x, y))$ holds uniformly over some range of x and y, say, $x \in X$ and $y \in Y$, means that the implied constant is *absolute*; that is, a fixed positive number, so long as x and y stay confined to these ranges.

Further, we write $f(x) \sim g(x)$ and $f(x) = o(g(x))$, as $x \to x_0$, if the limit

$$\lim_{x \to x_0} \frac{f(x)}{g(x)}$$

is equal to 1 or 0, respectively; the number x_0 may be infinite.

The notations above also apply to classes of functions. Thus, $O(1)$ denotes a bounded function, and $o(1)$ a vanishing function, with respect to the variables under consideration.

2 Three conjectures on the zeta function

In Section 2.4, we discuss the relationship between three unproven conjectures on the Riemann zeta function: the Riemann hypothesis, the Lindelöf hypothesis and the density hypothesis. Before we can state the two latter hypotheses, we need to introduce some notation and derive some results pertaining to the zeros, and to the growth of $\zeta(s)$ on vertical lines. We begin with a classical result that quantifies the density of nontrivial zeros inside the critical strip.

2.1 The number of nontrivial zeros

For a real number T, we denote by $N(T)$ the number of nontrivial zeros ρ with $0 < \gamma \leq T$, counted with multiplicity. It was proved nonrigorously by Riemann [29], and later rigorously by von Mangoldt [37], that the number $N(T)$ grows approximately like $T \log T$ as $T \to \infty$. We will follow a proof of this fact due to Backlund [3], which makes clever use of the functional equation for $\xi(s)$, but we first need two lemmas.

Lemma 2.1. *Suppose that $f(s)$ is a regular function on an open set containing the disc $|s - s_0| \leq R$, and that $f(s_0) \neq 0$. If $f(s)$ has m or more zeros when counted with multiplicity inside the disc $|s - s_0| \leq r$, where $0 < r < R$, then*

$$\left(\frac{R}{r}\right)^m \leq \frac{\max\limits_{|s-s_0|=R} |f(s)|}{|f(s_0)|}.$$

For proof, see [18] pp. 49–50.

Second, we need a lemma which bounds the vertical growth of $\zeta(s)$. In light of the next section, the following will suffice.

Lemma 2.2. *For any real σ_0, there exists a positive constant $A = A(\sigma_0)$, such that*

$$|\zeta(\sigma + ti)| = O(t^A)$$

uniformly for $\sigma \geq \sigma_0$ as $t \to \infty$.

Proof. We use Abel's summation formula (A2.1) to write

$$\sum_{n \leq X} n^{-s} = \frac{\lfloor X \rfloor}{X^s} + s \int_1^X \frac{\lfloor x \rfloor}{x^{s+1}} \mathrm{d}x = \frac{s}{s-1} - \frac{1}{2} - s \int_1^X \frac{P(x)}{x^{s+1}} \mathrm{d}x - \left(X^{-s} P(X) + \frac{1}{s-1} X^{-s+1}\right)$$

for $\sigma > 1$ and $X > 0$, where $P(x)$ is the 1-periodic function defined by[2]

$$P(x) = x - \lfloor x \rfloor - \frac{1}{2}. \tag{5}$$

Since $P(x)$ is bounded, the terms inside the brackets vanish as $X \to \infty$, and we are left with

$$\zeta(s) = \frac{s}{s-1} - \frac{1}{2} - s \int_1^\infty \frac{P(x)}{x^{s+1}} \mathrm{d}x \tag{6}$$

(in fact for $\sigma > 0$). Now apply integration by parts N times to put this in the form

$$\zeta(s) = \frac{s}{s-1} - \frac{1}{2} - s(s+1)\cdots(s+N) \int_1^\infty \frac{P_N(x)}{x^{s+N+1}} \mathrm{d}x, \tag{7}$$

[2] $P(x)$ is the 1-periodic extension of the restriction to $[0, 1)$ of the 1st Bernoulli-polynomial $B_1(x) = x - 1/2$.

where $P_0(x) = P(x)$ and $P_{n+1}(x) = \int_1^x P_n(y)\mathrm{d}y$ for $n \geq 0$. Then

$$P_N(x) = \frac{1}{(N+1)!}(x - \lfloor x \rfloor)^{N+1} - \frac{1}{2 \cdot N!}(x - \lfloor x \rfloor)^N$$

is bounded, which together with (7) shows that $|\zeta(\sigma+ti)| = O(|t|^{N+1})$ uniformly for $\sigma \geq -N+\delta > -N$ as $t \to \infty$. The claim follows upon choosing N so large that $-N < \sigma_0$. □

Theorem 2.3. *As $T \to \infty$, we have*

$$N(T) = \frac{T}{2\pi} \log\left(\frac{T}{2\pi}\right) - \frac{T}{2\pi} + O(\log T).$$

Proof. Let us work with $\xi(s)$ in place of $\zeta(s)$. This is simpler, because the functional equation $\xi(s) = \xi(1-s)$ allows us to break the problem into smaller pieces. Begin by assuming that $T > 0$, and that T does not coincide with any of the ordinates γ. Then $\xi(s)$ has exactly $2N(T)$ zeros in the interior of the rectangle $R = [-1, 2] \times [-T, T]$, and none on its boundary. Since $\xi(s)$ is an entire function, we may use the argument principle, without obstructions, to conclude that

$$4\pi N(T) = \big[\arg\, \xi(s)\big]_{\partial R},$$

where the right hand side denotes the increment of $\arg\, \xi(s)$ as the complex number s traverses once around the boundary of R in the positive sense. From the definition of $\xi(s)$, we have further

$$\big[\arg\, \xi(s)\big]_{\partial R} = \big[\arg\, \frac{1}{2}s(s-1)\big]_{\partial R} + \big[\arg\, \phi(s)\big]_{\partial R},$$

where $\phi(s) = \pi^{-s/2}\Gamma(s/2)\zeta(s)$. The first term on the right is 4π, again by the argument principle, since $s(s-1)/2$ has its only two zeros in the interior of R, and no poles. Since now $\phi(\overline{s}) = \overline{\phi(s)}$ and $\phi(s) = \phi(1-s)$ for all s, the second term must be equal to $4\big[\arg\, \phi(s)\big]_L$, where L consists of the segment from 2 to $2+Ti$ followed by the segment from $2+Ti$ to $\frac{1}{2}+Ti$. Putting this together, we obtain

$$\pi N(T) = \pi + \big[\arg\, \pi^{-s/2}\big]_L + \big[\arg\, \Gamma\Big(\frac{s}{2}\Big)\big]_L + \big[\arg\, \zeta(s)\big]_L. \tag{8}$$

The first term is

$$\big[\arg\, \pi^{-s/2}\big]_L = \big[-\frac{1}{2}t\log\pi\big]_L = -\frac{1}{2}T\log\pi.$$

The second term can be estimated using Stirling's formula (A2.2) (with $\alpha = 1/4$). This gives

$$\begin{aligned}\big[\arg\, \Gamma\Big(\frac{s}{2}\Big)\big]_L &= \big[\mathrm{Im}\log\Gamma\Big(\frac{s}{2}\Big)\big]_L = \mathrm{Im}\log\Gamma\Big(\frac{1}{4}+\frac{T}{2}i\Big) - \mathrm{Im}\log\Gamma(1)\\ &= \mathrm{Im}\Big(\Big(-\frac{1}{4}+\frac{T}{2}i\Big)\log\Big(\frac{T}{2}i\Big) - \frac{T}{2}i + \frac{1}{2}\log(2\pi)\Big) + O(T^{-1})\\ &= \frac{T}{2}\log\Big(\frac{T}{2}\Big) - \frac{T}{2} - \frac{\pi}{8} + O(T^{-1}).\end{aligned}$$

Substituted into (8), we find that[3]

$$N(T) = \frac{T}{2\pi}\log\left(\frac{T}{2\pi}\right) - \frac{T}{2\pi} + \frac{7}{8} + S(T) + O(T^{-1})$$

[3] A more precise version of the error term can be attained from a Stirling-series representation of the gamma function, see e.g., §§6.5–6.7 in Edward's book [11].

with

$$S(T) = \frac{1}{\pi}\big[\arg\, \zeta(s)\big]_L = \frac{1}{\pi}\arg\, \zeta\Big(\frac{1}{2} + Ti\Big),$$

where the last argument is defined by continuous variation along L.

So far the terms have been straightforward to estimate accurately, but this is not so easy for $S(T)$. The fact that

$$\mathrm{Re}\ \zeta(2+ti) = \sum_{n=1}^{\infty} \frac{\cos(t\log n)}{n^2} \geq 1 - \sum_{n=2}^{\infty}\frac{1}{n^2} = 2 - \frac{\pi^2}{6} = 0.35506\ldots$$

tells us that the variation in $\arg\, \zeta(s)$ along the vertical component of L is bounded in absolute value by π, so we can concentrate our attention on the horizontal segment from $2 + Ti$ to $\frac{1}{2} + Ti$. As a matter of fact, since $\zeta(\sigma)$ converges to 1 from above as $\sigma \to \infty$, the estimate

$$|\zeta(\sigma + ti) - 1| = \left|\sum_{n=2}^{\infty} n^{-(\sigma+ti)}\right| \leq \zeta(\sigma) - 1 \tag{9}$$

shows that $\zeta(\sigma + ti)$ converges to 1 uniformly in t as $\sigma \to \infty$. Thus, there must exist a real constant σ_*, such that $\mathrm{Re}\ \zeta(s) > 0$ whenever $\sigma > \sigma_*$. It would therefore suffice to consider the variation in $\arg\, \zeta(s)$ over the segment from $\sigma_* + Ti$ to $\frac{1}{2} + Ti$. The smallest constant σ_* with this property was computed to 100 digits in [10], the first of which are $\sigma_* = 1.19234\ldots$.

Here is one way of showing that $[\arg\, \zeta(s)]_L = O(\log T)$; Let m denote the number of distinct points s' in $\{2\} \times (0, T] \cup (\frac{1}{2}, 2] \times \{T\}$ for which $\mathrm{Re}\ \zeta(s') = 0$: then

$$\big|[\arg\, \zeta(s)]_L\big| \leq (m+1)\pi.$$

Indeed, when s traverses one of the $m + 1$ pieces of L subdivided by the s', then $\arg \zeta(s)$ does not change more that π, since $\mathrm{Re}\ \zeta(s)$ does not change sign there.

As argued above, there are no s' on the segment from 2 to $2 + Ti$, so m must be equal to the number of distinct points $\frac{1}{2} < \sigma < 2$ for which $\mathrm{Re}\ \zeta(\sigma + Ti) = 0$. This is in turn precisely equal to the number of distinct zeros of

$$g(s) = \frac{\zeta(s + Ti) + \zeta(s - Ti)}{2}$$

in the interval $\frac{1}{2} < s < 2$ on the real axis, because $g(s)$ is regular on $\mathbb{C}\backslash\{1 \pm Ti\}$ and equals $\mathrm{Re}\ \zeta(s+Ti)$ whenever s is a real number (this follows from $\zeta(\overline{s}) = \overline{\zeta(s)}$).

We can bound the number m by applying Lemma 2.1 to $g(s)$ and the discs $|s - 2| \leq 3/2$ and $|s - 2| \leq 7/4$. If we suppose that $T > 2$, then $g(s)$ is certainly regular in the largest disc, and the lemma yields

$$\Big(\frac{7}{6}\Big)^m \leq \frac{\max\limits_{|s-2|=7/4} |g(s)|}{|g(2)|} = \frac{\max\limits_{|s-2|=7/4} |\zeta(s+Ti) + \zeta(s-Ti)|}{2\mathrm{Re}\ \zeta(2+Ti)} \leq \frac{\max\limits_{|s-(2+Ti)|=7/4} |\zeta(s)|}{\mathrm{Re}\ \zeta(2+Ti)} \leq \frac{C(T+\frac{7}{4})^A}{2 - \frac{\pi^2}{6}},$$

where $C > 0$ and $A > 0$ are chosen using Lemma 2.2, such that $|\zeta(s)| \leq Ct^A$ for all $t \geq \frac{1}{4}$ and $\sigma \geq \frac{1}{4}$. Taking the logarithm on both sides, we get

$$m \leq \frac{\log\big(\frac{6C}{12-\pi^2}\big)}{\log(7/6)} + \frac{A}{\log(7/6)}\log\big(T + \tfrac{7}{4}\big) \leq D\log T \tag{10}$$

for some absolute constant $D > 0$. This proves the claim for $T > 2$ not coinciding with any γ. For T equal to some $\gamma > 0$, we may apply the bound (10) with $T + \varepsilon$ in place of T, and let $\varepsilon \to 0^+$. By continuity, the inequality remains valid for $\varepsilon = 0$, and our proof is complete. $\square$

Theorem 2.3 has the following useful corollary.

Corollary 2.4. *As $T \to \infty$, we have*

$$\sum_{0<\gamma\le T} \frac{1}{\gamma^k} = \begin{cases} O(1) & \text{if } k>1,\\ \dfrac{\log^2 T}{4\pi} - \dfrac{\log(2\pi)\log T}{2\pi} + O(1) & \text{if } k=1,\\ \dfrac{T^{1-k}\log T}{2\pi(1-k)} - \dfrac{1}{2\pi}\Big(\dfrac{\log(2\pi)}{1-k} + \dfrac{1}{(1-k)^2}\Big)T^{1-k} + O(1) & \text{if } 0<k<1,\\ N(T) = \dfrac{T\log T}{2\pi} - \dfrac{\log(2\pi\mathrm{e})T}{2\pi} + O(\log T) & \text{if } k=0,\\ \dfrac{T^{1-k}\log T}{2\pi(1-k)} - \dfrac{1}{2\pi}\Big(\dfrac{\log(2\pi)}{1-k} + \dfrac{1}{(1-k)^2}\Big)T^{1-k} + O(T^{-k}\log T) & \text{if } k<0, \end{cases}$$

and, if $k>1$,

$$\sum_{\gamma>T} \frac{1}{\gamma^k} = \frac{\log T}{2\pi(k-1)T^{k-1}} + \frac{1}{2\pi}\Big(\frac{1}{(k-1)^2} - \frac{\log(2\pi)}{k-1}\Big)\frac{1}{T^{k-1}} + O\Big(\frac{\log T}{T^k}\Big).$$

Proof. We begin by writing the asymptotic formula for $N(T)$ in the form

$$N(T) = \frac{T\log T}{2\pi} - \frac{\theta T}{2\pi} + O(\log T), \tag{11}$$

where $\theta = \log(2\pi\mathrm{e})$, and where the terms are arranged in decreasing order of magnitude. By Abel's summation formula (A2.1), we have

$$\sum_{0<\gamma\le T} \frac{1}{\gamma^k} = \frac{N(T)}{T^k} + k\int_1^T \frac{N(t)}{t^{k+1}}\mathrm{d}t \tag{12}$$

for $k \in \mathbb{R}$ and $T \le \infty$, where, in the case of $T=\infty$, both sides are finite only for $k>1$, with the first term on the right taking the value zero. If we substitute (11) into (12), then the result is

$$\sum_{0<\gamma\le T} \frac{1}{\gamma^k} = \Big(\frac{\log T}{2\pi T^{k-1}} - \frac{\theta}{2\pi T^{k-1}} + O\Big(\frac{\log T}{T^k}\Big)\Big) + \Big(\frac{k}{2\pi}\int_1^T \frac{\log t}{t^k}\mathrm{d}t - \frac{k\theta}{2\pi}\int_1^T \frac{\mathrm{d}t}{t^k} + O\Big(\int_1^T \frac{\log t}{t^{k+1}}\mathrm{d}t\Big)\Big).$$

Using the integration formulae (A2.6) through (A2.8) from the Appendix, this gives us

$$\begin{aligned}\sum_{0<\gamma\le T} \frac{1}{\gamma} &= \frac{\log T}{2\pi} - \frac{\theta}{2\pi} + O\Big(\frac{\log T}{T}\Big) + \frac{\log^2(T)}{4\pi} - \frac{\theta\log T}{2\pi} + O\Big(\int_1^T \frac{\log t}{t^2}\mathrm{d}t\Big)\\ &= \frac{\log^2 T}{4\pi} - \frac{\log(2\pi)\log T}{2\pi} + O(1).\end{aligned}$$

And, if $k \ne 1$,

$$\begin{aligned}\sum_{0<\gamma\le T} \frac{1}{\gamma^k} &= \frac{\log T}{2\pi T^{k-1}} - \frac{\theta}{2\pi T^{k-1}} + \frac{k}{2\pi}\Big(\frac{1}{(k-1)^2} - \frac{\log T}{(k-1)T^{k-1}} - \frac{1}{(k-1)^2T^{k-1}}\Big)\\ &\qquad - \frac{k\theta}{2\pi}\Big(\frac{1}{(1-k)T^{k-1}} - \frac{1}{1-k}\Big) + E(T,k)\\ &= \frac{\log T}{2\pi(1-k)T^{k-1}} - \frac{1}{2\pi}\Big(\frac{\log(2\pi)}{1-k} + \frac{1}{(k-1)^2}\Big)\frac{1}{T^{k-1}} + \frac{1}{2\pi}\Big(\frac{k(2-k)}{(k-1)^2} + \frac{k\log(2\pi)}{1-k}\Big) + E(T,k),\end{aligned}$$

with an error term

$$E(T,k) \ll \frac{\log T}{T^k} + \int_1^T \frac{\log t}{t^{k+1}}\mathrm{d}t \ll \begin{cases} 1 & \text{if } k > 0, \\ \log^2 T & \text{if } k = 0, \\ T^{-k}\log T & \text{if } k < 0. \end{cases}$$

When $k > 1$, we may similarly compute

$$\begin{aligned} \sum_{\gamma > T} \frac{1}{\gamma^k} &= \sum_{\gamma>0} \frac{1}{\gamma^k} - \sum_{0<\gamma\le T} \frac{1}{\gamma^k} = k\int_T^\infty \frac{N(t)}{t^{k+1}}\mathrm{d}t - \frac{N(T)}{T^k} \\ &= \Big(\frac{k}{2\pi}\int_T^\infty \frac{\log t}{t^k}\mathrm{d}t - \frac{k\theta}{2\pi}\int_T^\infty \frac{\mathrm{d}t}{t^k} + O\Big(\int_T^\infty \frac{\log t}{t^{k+1}}\mathrm{d}t\Big)\Big) - \Big(\frac{\log T}{2\pi T^{k-1}} - \frac{\theta}{2\pi T^{k-1}} + O\Big(\frac{\log T}{T^k}\Big)\Big) \\ &= \frac{k\log T}{2\pi(k-1)T^{k-1}} + \frac{k}{2\pi(k-1)^2 T^{k-1}} - \frac{k\theta}{2\pi(k-1)T^{k-1}} - \frac{\log T}{2\pi T^{k-1}} + \frac{\theta}{2\pi T^{k-1}} + O\Big(\frac{\log T}{T^k}\Big) \\ &= \frac{\log T}{2\pi(k-1)T^{k-1}} + \frac{1}{2\pi}\Big(\frac{1}{(k-1)^2} - \frac{\log(2\pi)}{k-1}\Big)\frac{1}{T^{k-1}} + O\Big(\frac{\log T}{T^k}\Big). \end{aligned}$$

□

2.2 Growth of zeta on vertical lines

As we saw above, there is an intricate connection between the distribution of nontrivial zeros of $\zeta(s)$ and its growth on vertical lines. For a real number σ, we define the *Lindelöf function* $\mu(\sigma)$ by

$$\mu(\sigma) = \inf\{a \ge 0 \ : \ \zeta(\sigma+ti) = O(t^a) \text{ as } t\to\infty\} = \limsup_{t\to\infty} \frac{\log|\zeta(\sigma+ti)|}{\log t}.$$

In light of Lemma 2.2, it is clear that $\mu(\sigma)$ is a well defined function. Moreover, it is easily seen that $\mu(\sigma) = 0$ for $\sigma > 1$, since the estimate (9) shows that $\zeta(s)$ is bounded in each half-plane $\sigma \ge 1+\delta$ with $\delta > 0$. This can be extended to $\mu(1) = 0$ without too much work. To see this, write

$$\zeta(s) = \sum_{n=1}^{N-1} n^{-s} + \sum_{n=N}^{\infty} n^{-s} = \sum_{n=1}^{N-1} n^{-s} + \frac{N^{1-s}}{s-1} + \frac{N^{-s}}{2} - s\int_N^\infty \frac{P(x)}{x^{s+1}}\mathrm{d}x$$

for $\sigma > 1$, where $P(x)$ is the function defined in Equation (5). Since $|P(x)| \le 1/2$ for all real x, we obtain

$$|\zeta(\sigma+ti)| \le 1 + \frac{1}{2} + \cdots + \frac{1}{N-1} + \frac{1}{\sqrt{(\sigma-1)^2+t^2}} + \frac{1}{2N} + \frac{\sqrt{\sigma^2+t^2}}{2N}.$$

Taking $N = \lfloor t \rfloor$ and letting $\sigma \to 1^+$, we conclude that $|\zeta(1+ti)| = O(\log t)$ as $t\to\infty$, which immediately implies $\mu(1) = 0$.[4]

Remark 2.5. Beware of the infimum in the definition of $\mu(\sigma)$: it says that, for any $\varepsilon > 0$, we have $\zeta(\sigma+ti) = O(t^{\mu(\sigma)+\varepsilon})$ as $t\to\infty$; the same need not be true if $\varepsilon = 0$. Also note that, since $\zeta(\overline{s}) = \overline{\zeta(s)}$, $\mu(\sigma)$ may equivalently be defined as the infimum of all $a \ge 0$ such that $\zeta(\sigma+ti) = O(|t|^a)$ as $|t|\to\infty$.

We now make the following claim

Proposition 2.6. *The Lindelöf function satisfies*

$$\mu(\sigma) = \frac{1}{2} - \sigma + \mu(1-\sigma)$$

for all real σ.

[4] Here we implicitly used that the nth harmonic number is $O(\log n)$. See e.g., (A2.3) in the Appendix.

Proof. We begin with the asymmetric functional equation

$$\zeta(s) = 2^s \pi^{s-1} \sin\left(\frac{\pi s}{2}\right)\Gamma(1-s)\zeta(1-s),$$

where $s = \sigma + ti$. We assume throughout that σ is fixed and $t > 0$; we want to investigate what happens when $t \to \infty$. Taking absolute values and then the logarithm of both sides, we get

$$\begin{aligned}\log|\zeta(\sigma+ti)| &= \sigma\log(2) + (\sigma-1)\log(\pi) + \log\left|\sin\left(\frac{\pi\sigma}{2} + \frac{\pi t}{2} i\right)\right| \\ &\quad + \log|\Gamma(1-\sigma-ti)| + \log|\zeta(1-\sigma-ti)|. \end{aligned} \tag{13}$$

Since

$$\begin{aligned}|\sin(x+yi)|^2 &= |\sin(x)\cosh(y) + \cos(x)\sinh(y)i|^2 \\ &= \sin^2(x)\Big(\frac{e^y + e^{-y}}{2}\Big)^2 + \cos^2(x)\Big(\frac{e^y - e^{-y}}{2}\Big)^2 \\ &= \frac{e^{2y}}{4} + \frac{e^{-2y}}{4} + \frac{1}{2}\big(\sin^2(x) - \cos^2(x)\big) \\ &= \frac{e^{2y}}{4} + \frac{e^{-2y}}{4} - \frac{\cos(2x)}{2} \qquad \Big(= \frac{\cosh(2y) - \cos(2x)}{2}\Big)\end{aligned}$$

for real x and y, we have

$$\begin{aligned}\log\left|\sin\left(\frac{\pi\sigma}{2} + \frac{\pi t}{2}i\right)\right| &= \frac{1}{2}\log\left(\frac{1}{4}e^{\pi t} + \frac{1}{4}e^{-\pi t} - \frac{1}{2}\cos(\pi\sigma)\right) \\ &= \frac{1}{2}\log\left(\frac{1}{4}e^{\pi t}\right) + \frac{1}{2}\log\left(1 + e^{-2\pi t} - 2\cos(\pi\sigma)e^{-\pi t}\right) \\ &= \frac{\pi t}{2} - \log(2) + O(e^{-\pi t}).\end{aligned} \tag{14}$$

Now, if $t > 0$, then the principal logarithm

$$\log(-ti) = \log(t) - \frac{\pi}{2}i.$$

Using this in Stirling's formula (A2.2) (with $\alpha = 1 - \sigma$) then yields

$$\begin{aligned}\log\Gamma(1-\sigma-ti) &= \Big(\frac{1}{2} - \sigma - ti\Big)\Big(\log(t) - \frac{\pi}{2}i\Big) + ti + \frac{1}{2}\log(2\pi) + O(t^{-1}) \\ &= \Big(\frac{1}{2} - \sigma\Big)\log(t) - \frac{\pi t}{2} + \frac{1}{2}\log(2\pi) + \Big(t - \Big(\frac{1}{2} - \sigma\Big)\frac{\pi}{2} - t\log t\Big)i + O(t^{-1}).\end{aligned}$$

Therefore,

$$\begin{aligned}\log|\Gamma(1-\sigma-ti)| &= \mathrm{Re}\ \log\Gamma(1-\sigma-ti) \\ &= \Big(\frac{1}{2} - \sigma\Big)\log(t) - \frac{\pi t}{2} + \frac{1}{2}\log(2\pi) + O(t^{-1}).\end{aligned} \tag{15}$$

If we now plug (14) and (15) into (13), then we get as a result

$$\begin{aligned}\log|\zeta(\sigma+ti)| &= \sigma\log(2) + (\sigma-1)\log(\pi) + \frac{\pi t}{2} - \log(2) + O(e^{-\pi t}) \\ &\quad + \Big(\frac{1}{2} - \sigma\Big)\log(t) - \frac{\pi t}{2} + \frac{1}{2}\log(2\pi) + O(t^{-1}) + \log|\zeta(1-\sigma-ti)| \\ &= \Big(\sigma - \frac{1}{2}\Big)\log\Big(\frac{2\pi}{t}\Big) + O(t^{-1}) + \log|\zeta(1-\sigma-ti)|.\end{aligned}$$

Finally, if we exponentiate both sides and rearrange, then we see that

$$\left|\frac{\zeta(\sigma+ti)}{\zeta(1-\sigma-ti)}\right| = \left(\frac{2\pi}{t}\right)^{\sigma-1/2} \mathrm{e}^{O(1/t)} = \left(\frac{2\pi}{t}\right)^{\sigma-1/2} \left(1+O(t^{-1})\right) \tag{16}$$

holds.[5] This implies the stated result. To see this, note that for any $\varepsilon > 0$, we have $|\zeta(\sigma+ti)| = O(t^{\mu(\sigma)+\varepsilon})$ and $|\zeta(1-\sigma-ti)| = O(t^{\mu(1-\sigma)+\varepsilon})$. Substituting these estimates one by one into the equation above gives

$$|\zeta(1-\sigma-ti)| = O(t^{\mu(\sigma)+\sigma-1/2+\varepsilon}) \quad \text{and} \quad |\zeta(\sigma+ti)| = O(t^{\mu(1-\sigma)-\sigma+1/2+\varepsilon})$$

as $t \to \infty$. By definition of the Lindelöf function, this implies

$$\mu(1-\sigma) \leq \mu(\sigma) + \sigma - \frac{1}{2} + \varepsilon \quad \text{and} \quad \mu(\sigma) \leq \mu(1-\sigma) - \sigma + \frac{1}{2} + \varepsilon,$$

and, since $\varepsilon > 0$ was arbitrary, that $\mu(\sigma) = \frac{1}{2} - \sigma + \mu(1-\sigma)$. □

Since $\mu(\sigma)$ is zero for $\sigma \geq 1$, we immediately get

Corollary 2.7. *If $\sigma \leq 0$, then*

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

We can in fact say something more general about the behaviour of the Lindelöf function.

Proposition 2.8. *The Lindelöf function $\mu(\sigma)$ is convex.*

Proof. Let σ_1 and σ_2 be real numbers with $\sigma_1 < \sigma_2$. Suppose that for each of $j = 1, 2$, we have

$$|\zeta(\sigma_j + ti)| \leq A_j t^{p_j}$$

for $t \geq t_j$, where the A_j, p_j and t_j are positive constants. Let $t_0 = \max(t_1, t_2)$. We claim that there is a constant $C > 0$, such that

$$|\zeta(\sigma+ti)| \leq C t^{k(\sigma)}$$

for all $\sigma + ti$ in the vertical strip $V_{t_0} = \{\sigma_1 \leq \sigma \leq \sigma_2,\ t \geq t_0 > 0\}$, where

$$k(\sigma) = p_1 + \frac{p_2 - p_1}{\sigma_2 - \sigma_1}(\sigma - \sigma_1)$$

is the linear function that interpolates the points (σ_1, p_1) and (σ_2, p_2).

To this end, let $\varepsilon > 0$ and consider

$$F_\varepsilon(s) = \log|\zeta(s)| - k(\sigma)\log(t) - \varepsilon t.$$

Since the Laplacian of F_ε, $k(\sigma)t^{-2}$, is nonnegative wherever it is defined, the function F_ε is subharmonic on V_{t_0}, except at any point where $\zeta(s) = 0$, in which case F_ε has a singularity and assumes the value $-\infty$. Suppose for the moment that $\zeta(s)$ has no zeros in V_{t_0}. Then, on the vertical part of the boundary of V_{t_0}, we have

$$\begin{aligned} F_\varepsilon(\sigma_j + ti) &= \log|\zeta(\sigma_j+ti)| - k(\sigma_j)\log(t) - \varepsilon t \leq \log(A_j) + p_j\log(t) - k(\sigma_j)\log(t) - \varepsilon t \\ &= \log(A_j) - \varepsilon t \leq \log(A_j). \end{aligned}$$

[5]The estimate (16) for the *chi function* $\chi(s) = \zeta(s)/\zeta(1-s)$ can be found in Titchmarsh [35] p. 95 and incorrectly stated in Edwards [11] p. 185; see also Footnote †, p. 19 in Edward's book.

In other words, $F_\varepsilon(s)$ must be bounded above on ∂V_{t_0}, with

$$F_\varepsilon(s) \le \max\Big(\log A_1,\ \log A_2,\ \max_{\sigma_1\le\sigma\le\sigma_2}\Big(\log|\zeta(\sigma+t_0 i)| - k(\sigma)\log(t_0)\Big)\Big) \le \log C$$

for some appropriate choice of $C = C(t_0) > 0$. Further, by Lemma 2.2, we may find positive constants A and D, such that

$$|\zeta(s)| \le At^D$$

in V_{t_0}. Thus,

$$\begin{aligned} F_\varepsilon(\sigma+Ti) = &\ \log|\zeta(\sigma+Ti)| - k(\sigma)\log(T) - \varepsilon T \le \log(A) + D\log(T) - k(\sigma)\log(T) - \varepsilon T \\ &= O(\log T) - \varepsilon T. \end{aligned}$$

Since $\varepsilon > 0$, this implies that

$$F_\varepsilon(\sigma+Ti) \le \log C$$

for all $\sigma_1 \le \sigma \le \sigma_2$, provided T is chosen sufficiently large, say, $T \ge T_0$.

This shows that the subharmonic function F_ε is bounded above by $\log C$ on the boundary of $V_{t_0}^{(T)} = \{\sigma_1 \le \sigma \le \sigma_1,\ t_0 \le t \le T\}$. According to the weak maximum principle,[6] this is possible only if $F_\varepsilon(s) \le \log C$ on all of $V_{t_0}^{(T)}$. Since there is no loss in making T larger, we find that the same must be true in $V_{t_0}^{\infty} = V_{t_0}$. In other words,

$$|\zeta(s)| \le Ct^{k(\sigma)}\mathrm{e}^{\varepsilon t}$$

on all of V. Now, the choice of $\varepsilon > 0$ was arbitrary, so by continuity the same must be true if $\varepsilon = 0$. That is, $|\zeta(s)| \le Ct^{k(\sigma)}$ on V.

This proves our claim from the beginning at the proof, and it now only remains to relate this to the Lindelöf function. This is straightforward: for any $\delta > 0$, we have (by definition of the Lindelöf function) that

$$p_j = \mu(\sigma_j) + \delta$$

is an admissible choice for the p_j, in the sense that

$$|\zeta(\sigma_j + ti)| \le A_j(\delta)t^{\mu(\sigma_j)+\delta} \quad \text{for } t \ge t_j(\delta).$$

With this choice for the p_j, we have

$$k(\sigma) = k^*(\sigma) + \delta,$$

and hence

$$|\zeta(\sigma+ti)| \le C(\delta)t^{k^*(\sigma)+\delta} \quad \forall s \in V_{t_0(\delta)},$$

where $k^*(\sigma)$ is the linear function that interpolates the points $(\sigma_1, \mu(\sigma_1))$ and $(\sigma_2, \mu(\sigma_2))$. This implies that $\mu(\sigma) \le k^*(\sigma) + \delta$ and hence, since $\delta > 0$ was arbitrary, that $\mu(\sigma) \le k^*(\sigma)$ for $\sigma_1 \le \sigma \le \sigma_2$.

The same is true even if V_{t_0} contains points where $\zeta(s) = 0$. Indeed, we may exclude from V_{t_0} a collection of open discs around these zeros, whose radii are chosen independently and so small that that the bounds we used remain true on the boundary of the resulting 'perforated' strip; we omit the details.

This shows that $\mu(\sigma)$ is convex, given that the choice of σ_1 and σ_2 was arbitrary. □

[6] Suppose $f, g \in C(\overline{U}) \cap C^2(U)$, where U is a bounded domain. If f and g are subharmonic and harmonic on U, respectively, and if $f \le g$ on ∂U, then $f \le g$ holds throughout U.

Since $\mu(0) = \frac{1}{2}$ and $\mu(1) = 0$, we obtain at once

Corollary 2.9. *If* $0 \le \sigma \le 1$*, then*

$$\mu(\sigma) \le \frac{1}{2}(1-\sigma).$$

Besides the statement of Corollary 2.9, we have not commented on the value of $\mu(\sigma)$ for $0 < \sigma < 1$. This is because it is an open problem. However, since $\mu(\sigma)$ is a nonnegative function that is convex, zero for $\sigma \ge 0$, and equal to $\frac{1}{2} - \sigma$ for $\sigma \le 0$, we always have

$$\mu(\sigma) \ge \mu_{\mathrm{L}}(\sigma) = \begin{cases} \dfrac{1}{2} - \sigma & \text{if } \sigma \le \dfrac{1}{2}, \\ 0 & \text{if } \sigma \ge \dfrac{1}{2}. \end{cases}$$

By convexity of $\mu(\sigma)$, these functions are the same function if and only if it is true that $\mu(\frac{1}{2}) = 0$. The question about the truth of this equality is so important that it has its own name.

Conjecture 2.10 (Lindelöf hypothesis). *For any* $\varepsilon > 0$*, we have*

$$\zeta\Big(\frac{1}{2} + ti\Big) = O(t^{\varepsilon})$$

as $t \to \infty$*. Equivalently,* $\mu(\frac{1}{2}) = 0$*.*

2.3 Zero-density estimates

If σ and T are real numbers, then we let $N(\sigma, T)$ denote the number of nontrivial zeros ρ for which $\sigma \le \beta < 1$ and $0 < \gamma \le T$, counted with multiplicity. The function $N(\sigma, T)$ is nonincreasing as a function of σ (for T fixed), and nondecreasing as a function of T (for σ fixed). Also,

$$N(\sigma, T) + N(1-\sigma, T) = N(T),$$

provided that σ does not coincide with the abscissa β of any zero. The true size of $N(\sigma, T)$ as a function of σ and T hinges on the Riemann Hypothesis, and

$$N(\sigma, T) = \begin{cases} N(T) = N_0(T) & \text{if } \sigma \le 1/2, \\ 0 & \text{if } \sigma > 1/2, \end{cases}$$

provided that the Riemann hypothesis is true, where $N_0(T)$ denotes the number of zeros $\frac{1}{2} + \gamma i$ with $0 < \gamma \le T$, counted with multiplicity.

Upper bounds for $N(\sigma, T)$ in terms of σ and T are known as *zero-density estimates.* Since

$$\frac{T \log T}{4\pi} \sim \frac{1}{2} N(T) \le N(\tfrac{1}{2}, T) \le N(T) \sim \frac{T \log T}{2\pi},$$

the relation

$$N(\tfrac{1}{2}, T) = O(T \log T) = O(T^{1+\varepsilon}) \tag{17}$$

holds for any fixed $\varepsilon > 0$ as $T \to \infty$, while $N(\frac{1}{2}, T) = O(T)$ fails to be true for the same reason. We also have

$$N(1, T) = 0 = O(\log^{\varepsilon} T) = O(T^{\varepsilon}) \tag{18}$$

for any $\varepsilon > 0$ as $T \to \infty$.

For these reasons, it is common to write zero-density estimates on the form

$$N(\sigma, T) = O\big(T^{\lambda(\sigma)(1-\sigma)} f(T)\big),$$

where the bound holds uniformly in a range of σ as $T \to \infty$, with $f(T)$ typically being either T^{ε} for some fixed $\varepsilon > 0$, or some power of $\log T$. The factor $1-\sigma$ in the exponent is natural, for the following reason: if $\lambda(\sigma)$ is constant and equal to 2, then the resulting exponent, $2(1-\sigma)$, is the linear function which is 1 at $\sigma = \frac{1}{2}$ and 0 at $\sigma = 0$. In other words, it is the linear interpolation of the exponents of T in (17) and (18) in the limit as $\varepsilon \to 0$.

As we explain in the next section, the question of whether we can take $\lambda(\sigma) = 2$ is so interesting in and of itself that it also has its own name.

Conjecture 2.11 (Density hypothesis). *For any $\varepsilon > 0$, we have*

$$N(\sigma, T) = O(T^{2(1-\sigma)+\varepsilon})$$

uniformly in $\frac{1}{2} \le \sigma \le 1$ as $T \to \infty$.

Closely related to zero-density estimates are the so-called *zero-free regions* for $\zeta(s)$. These are regions that, as the name suggests, do not contain any zeros of $\zeta(s)$. As mentioned in the Introduction, all of the nontrivial zeros of the zeta function lie inside the critical strip and symmetrically about both the real and critical lines. Because of this symmetry, the zero-free regions are usually expressed by saying that $\zeta(s)$ has no zeros with

$$\sigma \ge 1 - A(t) \quad \text{and} \quad t \ge t_0,$$

where $A(t)$ is a function of t with $0 \le A(t) < \frac{1}{2}$ and t_0 is some positive constant.

Theorem 2.12. *Each of the following defines a region where $\zeta(s)$ has no zeros:*

$$\sigma \ge 1 - \frac{A}{\log t}, \quad t \ge t_0; \tag{19}$$

$$\sigma \ge 1 - \frac{A \log\log t}{\log t}, \quad t \ge t_0; \tag{20}$$

$$\sigma \ge 1 - \frac{A}{(\log t)^{10/11+\varepsilon}}, \quad t \ge t_0; \tag{21}$$

$$\sigma \ge 1 - \frac{A}{(\log t)^{2/3}(\log\log t)^{1/3}}, \quad t \ge t_0. \tag{22}$$

Here $A > 0$ and $t_0 > 3$ are constants that need not be the same in each case. Any $\varepsilon > 0$ may be taken in (21), possibly subject to increased values of A and t_0.

It is outside the scope of this project to prove these here, but we can give a summary of how they came to be. The zero-free region (19) was obtain independently by Hadamard and de la Vallée Poussin, and represents the first significant zero-free region to the left of the line $\sigma = 1$. To prove it, the 'only' thing that is needed is an equation that relates the logarithmic derivative of $\zeta(s)$ to its zeros, and the fact that $2(1+\cos\theta)^2 \ge 0$.[7]

The enlarged regions (20), (21) and (22) are due to Littlewood, Tchudakoff [34], and Koborov [24] and Vinogradov [36], respectively. In one way or another, they all rely on establishing cancellation among the terms in an exponential sum (also called trigonometric sum); that is, a sum of the form

$$\sum_{n \in S_X} \mathrm{e}^{2\pi i f(n)}, \tag{23}$$

[7] We are simplifying heavily here. A detailed proof may be found in [18].

where f is a function and S_X a set that may evolve with the parameter X. More specifically, (20) may be proved by estimating (23) using an appropriate integral, while (21) and (22) can be proved using Vinogradov's esimates for exponential sums (to quote Titchmarsh, 'This is in some ways very complicated').[8]

The region determined by (22) is effectively the largest known zero-free region for $\zeta(s)$; The regions (19)–(21) are included for historical reasons, but we will find that much can be proved using only Tchudakoff's region (21).

2.4 Relationships between the hypotheses

A relationship between the Lindelöf hypothesis and the number $N(\sigma, T)$ was established by Backlund [4], where we proved

Theorem 2.13 (Backlund, 1918). *The Lindelöf hypothesis is true if and only if*

$$N(\sigma, T+1) - N(\sigma, T) = o(\log T)$$

for every fixed $\sigma > \frac{1}{2}$ *as* $T \to \infty$.

If the Riemann hypothesis is true, then $N(\sigma, T)$ is zero whenever $\sigma > \frac{1}{2}$, and so

Corollary 2.14. *If the Riemann hypothesis is true, then so is the Lindelöf hypothesis.*

The density hypothesis has its origin in Ingham's article [19], where he showed the following.

Theorem 2.15 (Ingham, 1937). *If* $\zeta(\frac{1}{2} + ti) = O(t^c)$ *for some fixed* $c > 0$ *as* $t \to \infty$, *then*

$$N(\sigma, T) = O\big(T^{2(1+2c)(1-\sigma)} \log^5 T\big)$$

uniformly for $\frac{1}{2} \leq \sigma \leq 1$ *as* $T \to \infty$.

If the Lindelöf hypothesis is true, then c may be taken arbitrarily small and positive, which implies the following.

Corollary 2.16. *If the Lindelöf hypothesis is true, then so is the density hypothesis.*

In fact, we shall require a strengthened version of the density hypothesis in which the logarithmic powers are taken seriously.

Conjecture 2.17 (Strong density hypothesis). *There exists a constant* $\eta \geq 1$, *such that*

$$N(\sigma, T) = O\big(T^{2(1-\sigma)} \log^{\eta} T\big)$$

uniformly for $\frac{1}{2} \leq \sigma \leq 1$ *as* $T \to \infty$.

Proof of Theorems 2.13 and 2.15 can be found in textbooks on the zeta function.[9] It should be noted that Ingham's proof of Theorem 2.15 does not, at least without substantial modification, generalise in such a way that Corollary 2.16 remains true with the strong density hypothesis in place of the density hypothesis. In other words, this is the situation:

$$\begin{tikzcd}
& \text{Riemann hypothesis} \arrow[dl, "\text{Backlund}"'] \arrow[dr, "\text{obvious}"] & \\
\text{Lindelöf hypothesis} \arrow[dr, "\text{Ingham}"'] & & \text{Strong density hypothesis} \arrow[dl, "\text{obvious}"] \\
& \text{Density hypothesis} &
\end{tikzcd}$$

Question: *Is the Lindelöf hypothesis stronger than the strong density hypothesis, vice versa, or neither?*

[8] [35], p. 98.

[9] For example in Titchmarsh [35], Theorems 9.18 and 13.5.

3 The explicit formula for $\psi(x)$

The goal of this section is to prove the explicit formula for the second Chebyshev function $\psi(x)$. This formula expresses a fundamental relation between the distribution of prime powers and the distribution of zeros of the Riemann zeta function, which we put to use in the following sections. In order to derive the formula, we first consider some lemmas in Section 3.1 before we prove the explicit formula in Section 3.2.

3.1 Some lemmas

Our first lemma is a special case of *Perron's formula.* Some inequalities in the proof are marked with '†', indicating that an improvement of these inequalities may give a better error term in the lemma.

Lemma 3.1. *If $y > 0$ and $c > 0$ are fixed real numbers, then*

$$\frac{1}{2\pi i}\int_{c-\infty i}^{c+\infty i} \frac{y^s}{s}\mathrm{d}s = \begin{cases} 0 & \text{if } 0 < y < 1, \\ 1/2 & \text{if } y = 1, \\ 1 & \text{if } y > 1, \end{cases} \tag{24}$$

where y^s is defined as $\mathrm{e}^{s\log y}$, with $\log y$ denoting the real-valued logarithm of y. If the integral to the left (with the factor $\frac{1}{2\pi i}$) is denoted by $I(y)$, then we define it as the Cauchy principal value $\lim_{T\to\infty} I(y,T)$ in the case when $y = 1$, where

$$I(y,T) = \frac{1}{2\pi i}\int_{c-Ti}^{c+Ti} \frac{y^s}{s}\mathrm{d}s.$$

Moreover, if we write $I(y) = I(y,T) + \Delta(y,T)$ for $T \geq 0$, then

$$|\Delta(y,T)| < \begin{cases} \dfrac{y^c}{\pi T|\log y|} & \text{if } y \neq 1, \\ \dfrac{c}{\pi T} & \text{if } y = 1, \end{cases} \tag{25}$$

if $T > 0$, and

$$|\Delta(y,T)| < y^c \quad \text{for all } y > 0 \text{ and } T \geq 0. \tag{26}$$

Proof. Suppose in the first instance that $y > 1$, and let R be the rectangle whose vertices are

$$c + Vi, \quad -X + Vi, \quad -X - Ui, \quad c - Ui$$

for $X, U, V > 0$, see Figure 1.

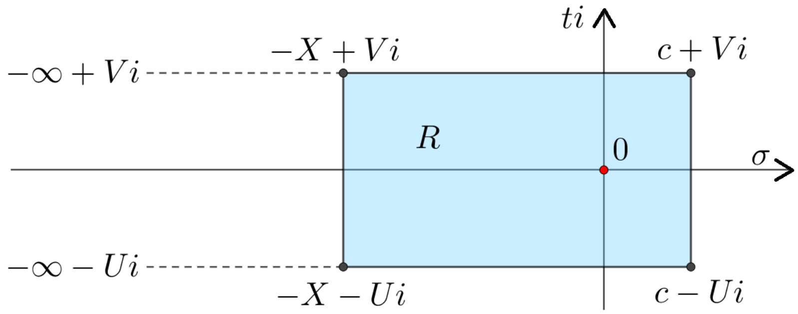

Figure 1

Since y^s/s is regular on $\mathbb{C}\setminus\{0\}$ with a simple pole of residue 1 at $s=0$, we have by the residue theorem that

$$\frac{1}{2\pi i}\oint_{\partial R}\frac{y^s}{s}\mathrm{d}s = 1. \tag{27}$$

On the segment of ∂R from $-X+Vi$ to $-X-Ui$, we have $|y^s/s| = y^{-X}/\sqrt{X^2+t^2} \le y^{-X}/X$, and hence

$$\left|\int_{-X+Vi}^{-X-Ui}\frac{y^s}{s}\mathrm{d}s\right| \le (U+V)\frac{y^{-X}}{X}.$$

Taking the limit as $X\to\infty$ in (27) while keeping U and V fixed therefore gives

$$\frac{1}{2\pi i}\int_{c-Ui}^{c+Vi}\frac{y^s}{s}\mathrm{d}s = 1 - \underbrace{\frac{1}{2\pi i}\int_{-\infty-Ui}^{c-Ui}\frac{y^s}{s}\mathrm{d}s}_{J(-U)} + \underbrace{\frac{1}{2\pi i}\int_{-\infty+Vi}^{c+Vi}\frac{y^s}{s}\mathrm{d}s}_{J(V)}. \tag{28}$$

The integrals $J(-U)$ and $J(V)$ are absolutely convergent. Indeed,

$$\begin{aligned}|J(-U)| &= \frac{1}{2\pi}\left|\int_{-\infty-Ui}^{c-Ui}\frac{y^s}{s}\mathrm{d}s\right| = \frac{1}{2\pi}\left|\int_{-\infty}^{c}\frac{y^{\sigma-Ui}}{\sigma-Ui}\mathrm{d}\sigma\right| \overset{\dagger}{\le} \frac{1}{2\pi}\int_{-\infty}^{c}\frac{y^\sigma}{\sqrt{\sigma^2+U^2}}\mathrm{d}\sigma \\ &< \frac{1}{2\pi}\int_{-\infty}^{c}\frac{y^\sigma}{U}\mathrm{d}\sigma = \frac{y^c}{2\pi U\log y} = \frac{y^c}{2\pi U|\log y|},\end{aligned} \tag{29}$$

and similarly

$$|J(V)| < \frac{y^c}{2\pi V|\log y|}. \tag{30}$$

If we now take the limit as $U,V\to\infty$ in (28), using the bounds (29) and (30), then we get (24) in the case when $y>1$. Also, upon taking $U=V=T>0$ in (28), we have

$$|\Delta(y,T)| = |I(y)-I(y,T)| = |1-I(y,T)| = |J(-T)-J(T)| \overset{\dagger}{\le} |J(-T)|+|J(T)| < \frac{y^c}{\pi T|\log y|},$$

and hence (25) holds for $y>1$.

Now suppose that $0<y<1$ is fixed, and let R^* denote the rectangle whose vertices are

$$X+Vi,\quad c+Vi,\quad c-Ui,\quad X-Ui$$

for $U,V>0$ and $X>c$, see Figure 2.

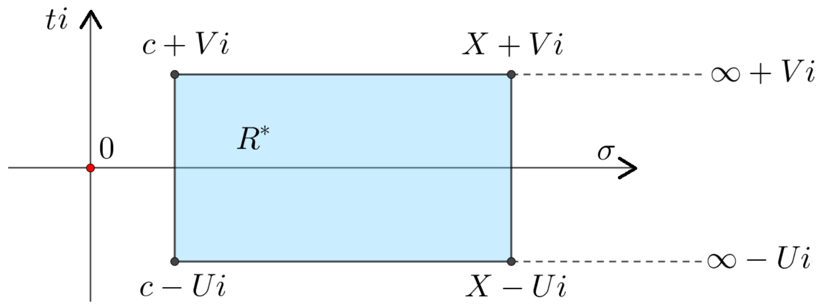

Figure 2

Since $X > c > 0$, the pole of y^s/s at $s = 0$ lies outside of R^*. Hence, by the residue theorem (now taking the integral with the negative orientation):

$$\frac{1}{2\pi i}\oint_{\partial R^*}\frac{y^s}{s}\mathrm{d}s = 0. \tag{31}$$

On the segment from $X + Vi$ to $X - Ui$, we have $|y^s/s| \leq y^X/\sqrt{X^2+t^2} \leq y^X/X$, so

$$\left|\int_{X+Vi}^{X-Ui}\frac{y^s}{s}\mathrm{d}s\right| \leq (U+V)\frac{y^X}{X}.$$

Taking the limit as $X \to \infty$ in (31) while keeping U and V fixed therefore gives

$$\frac{1}{2\pi i}\int_{c-Ui}^{c+Vi}\frac{y^s}{s}\mathrm{d}s = \underbrace{\frac{1}{2\pi i}\int_{c-Ui}^{\infty-Ui}\frac{y^s}{s}\mathrm{d}s}_{J^*(-U)} - \underbrace{\frac{1}{2\pi i}\int_{c+Vi}^{\infty+Vi}\frac{y^s}{s}\mathrm{d}s}_{J^*(V)}. \tag{32}$$

The integrals $J^*(-U)$ and $J^*(V)$ are also absolutely convergent, since

$$\begin{aligned}|J^*(-U)| &= \frac{1}{2\pi}\left|\int_{c-Ui}^{\infty-Ui}\frac{y^s}{s}\mathrm{d}s\right| = \frac{1}{2\pi}\left|\int_c^\infty\frac{y^{\sigma-Ui}}{\sigma-Ui}\mathrm{d}\sigma\right| \overset{\dagger}{\leq} \frac{1}{2\pi}\int_c^\infty\frac{y^\sigma}{\sqrt{\sigma^2+U^2}}\mathrm{d}\sigma \\ &< \frac{1}{2\pi}\int_c^\infty\frac{y^\sigma}{U}\mathrm{d}\sigma = \frac{-y^c}{2\pi U\log y} = \frac{y^c}{2\pi U|\log y|},\end{aligned} \tag{33}$$

and similarly

$$|J^*(V)| < \frac{y^c}{2\pi V|\log y|}. \tag{34}$$

Hence, letting $U, V \to \infty$ in (32) and using (33) and (34), we obtain (24) in the case when $0 < y < 1$. Moreover, if we take $U = V = T > 0$ in (32), then we get

$$|\Delta(y,T)| = |I(y) - I(y,T)| = |0 - I(y,T)| = |J^*(T) - J^*(-T)| \overset{\dagger}{\leq} |J^*(T)| + |J^*(-T)| \leq \frac{y^c}{\pi T|\log y|}.$$

Now suppose that $y = 1$. Then a straightforward computation gives

$$I(1,T) = \frac{1}{2\pi i}\int_{c-Ti}^{c+Ti}\frac{\mathrm{d}s}{s} = \frac{1}{2\pi}\int_{-T}^{T}\frac{\mathrm{d}t}{c+ti} = \frac{1}{2\pi}\int_{-T}^{T}\frac{c-ti}{c^2+t^2}\mathrm{d}t = \frac{1}{\pi}\int_0^T\frac{c}{c^2+t^2}\mathrm{d}t,$$

since the real and imaginary parts of $(c - ti)/(c^2 + t^2)$ are even and odd functions of t, respectively. Therefore,

$$I(1) \stackrel{\text{def}}{=} \lim_{T\to\infty} I(1,T) = \frac{1}{\pi}\int_0^\infty \frac{c}{c^2+t^2}\mathrm{d}t = \frac{1}{\pi}\arctan\Big(\frac{\infty}{c}\Big) = \frac{1}{2},$$

which proves (24) for $y = 1$. Furthermore,

$$0 < \Delta(1,T) = I(1) - I(1,T) = \tfrac{1}{2} - I(1,T) = \frac{1}{\pi}\int_T^\infty \frac{c}{c^2+t^2}\mathrm{d}t \stackrel{\dagger}{<} \begin{cases} \dfrac{c}{\pi}\displaystyle\int_T^\infty \frac{\mathrm{d}t}{t^2} = \frac{c}{\pi T} & \text{if } T > 0, \\ \dfrac{1}{\pi}\displaystyle\int_0^\infty \frac{c}{c^2+t^2}\mathrm{d}t = \frac{1}{2} < 1 = y^c, \end{cases}$$

so that (25) and (26) hold for $y = 1$.

To establish (26) in the case when $y \neq 1$ and $T > 0$, let Γ and Γ^* denote the arcs of the circle $|s| = \sqrt{c^2 + T^2}$ from $c + Ti$ to $c - Ti$ that lie to the left and right of the line $\sigma = c$, respectively, see Figure 3.

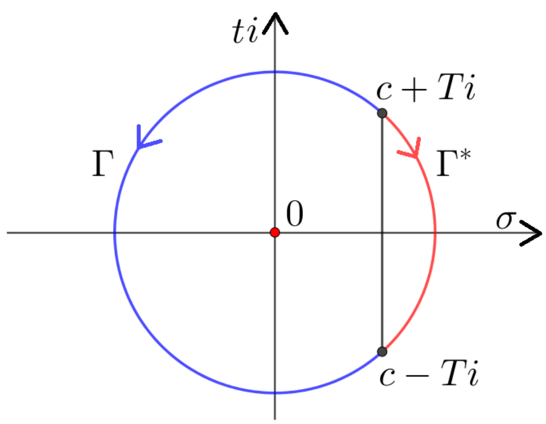

Figure 3

If we let $\overline{\Gamma} = \Gamma$ for $y > 1$, and $\overline{\Gamma} = \Gamma^*$ for $0 < y < 1$, then

$$\Delta(y,T) = I(y) - I(y,T) = I(y) - \Big(I(y,T) + \frac{1}{2\pi i}\int_{\overline{\Gamma}} \frac{y^s}{s}\mathrm{d}s\Big) + \frac{1}{2\pi i}\int_{\overline{\Gamma}} \frac{y^s}{s}\mathrm{d}s = \frac{1}{2\pi i}\int_{\overline{\Gamma}} \frac{y^s}{s}\mathrm{d}s,$$

since the bracketed term equals $I(y)$ by the residue theorem.[10] Accordingly,

$$|\Delta(y,T)| \stackrel{\dagger}{\leq} \frac{\text{length}(\overline{\Gamma})}{2\pi} \max_{s\in\overline{\Gamma}} \Big|\frac{y^s}{s}\Big| < \frac{2\pi\sqrt{c^2+T^2}}{2\pi} \cdot \frac{y^c}{\sqrt{c^2+T^2}} = y^c,$$

since $|y^s| = y^\sigma \leq y^c$ on $\overline{\Gamma}$.[11] By continuity of the integrand, this bound also holds if $T = 0$. This completes the proof of the lemma. $\square$

Remark 3.2. The inequalities (25) and (26) can be combined in saying that

$$|\Delta(y,T)| < \begin{cases} y^c \min\Big(1, \dfrac{1}{\pi T|\log y|}\Big) & \text{if } y \neq 1, \\ \min\big(1, \dfrac{c}{\pi T}\big) & \text{if } y = 1, \end{cases}$$

[10] This is true both for $y > 1$ and $0 < y < 1$.
[11] See Footnote 10.

for $T \geq 0$, where the finite value is understood if $T = 0$. The full proof was included for both of the cases $y > 1$ and $0 < y < 1$ because $J(-U)$ for $y > 1$ is not easily expressible in terms of, say, $J^*(U)$ for $0 < y < 1$.

We make a small digression on the exact error terms in Lemma 3.1, since the bounds attained are susceptible to improvement. A straightforward computation shows that

$$I(y,T) = \frac{1}{\pi}\int_0^T \mathrm{Re}\Big(\frac{y^{c+ti}}{c+ti}\Big)\mathrm{d}t = \frac{y^c}{\pi}\int_0^T \frac{c\cos(t\log y) + t\sin(t\log y)}{c^2+t^2}\mathrm{d}t,$$

so that $I(y)$, $I(y,T)$ and $\Delta(y,T)$ are all real for $y, T, c > 0$, with $I(y)$ and $\Delta(y,T)$ having a discontinuity at $y = 1$. Moreover, the proof of Lemma 3.1 shows that the exact error term is given by[12]

$$\Delta(y,T) = \begin{cases} J(-T) - J(T) & \text{if } y > 1, \\ \dfrac{1}{\pi}\displaystyle\int_T^\infty \frac{c}{c^2+t^2}\mathrm{d}t & \text{if } y = 1, \\ J^*(T) - J^*(-T) & \text{if } 0 < y < 1. \end{cases}$$

For $y > 1$, this can be written in various ways,

$$\begin{aligned} J(-T) - J(T) &= \frac{1}{2\pi i}\int_{-\infty-Ti}^{c-Ti} \frac{y^s}{s}\mathrm{d}s - \frac{1}{2\pi i}\int_{-\infty+Ti}^{c+Ti} \frac{y^s}{s}\mathrm{d}s = \frac{1}{2\pi i}\int_{-\infty}^{c} \frac{y^{\sigma-Ti}}{\sigma-Ti} - \frac{y^{\sigma+Ti}}{\sigma+Ti}\mathrm{d}\sigma \\ &= \frac{1}{\pi}\int_{-\infty}^{c} \mathrm{Im}\Big(\frac{y^{\sigma-Ti}}{\sigma-Ti}\Big)\mathrm{d}\sigma = \frac{1}{2\pi i}\int_{-\infty}^{c} \frac{y^\sigma}{\sigma^2+T^2}\big(y^{-Ti}(\sigma+Ti) - y^{Ti}(\sigma-Ti)\big)\mathrm{d}\sigma \\ &= \frac{1}{2\pi i}\int_{-\infty}^{c} \frac{y^c}{\sigma^2+T^2}\big(Ti(y^{Ti}+y^{-Ti}) - \sigma(y^{Ti}-y^{-Ti})\big)\mathrm{d}\sigma \\ &= \frac{1}{\pi}\int_{-\infty}^{c} \frac{y^\sigma}{\sigma^2+T^2}\big(T\cos(T\log y) - \sigma\sin(T\log y)\big)\mathrm{d}\sigma, \end{aligned}$$

and similarly for $0 < y < 1$:

$$J^*(T) - J^*(T) = \frac{1}{\pi}\int_c^\infty \mathrm{Im}\Big(\frac{y^{\sigma+Ti}}{\sigma+Ti}\Big)\mathrm{d}\sigma = \frac{1}{\pi}\int_c^\infty \frac{y^\sigma}{\sigma^2+T^2}\big(\sigma\sin(T\log y) - T\cos(T\log y)\big)\mathrm{d}\sigma.$$

Alternatively, one may use the circular arc $\overline{\Gamma}$ for $y \neq 1$, to write

$$\Delta(y,T) = \frac{1}{2\pi i}\int_{\overline{\Gamma}} \frac{y^s}{s}\mathrm{d}s = \frac{1}{2\pi}\int_{\theta_+}^{\theta_-} \mathrm{e}^{\alpha\cos(\theta)}\cos(\alpha\sin\theta)\mathrm{d}\theta + \frac{i}{2\pi}\int_{\theta_+}^{\theta_-} \mathrm{e}^{\alpha\cos(\theta)}\sin(\alpha\sin\theta)\mathrm{d}\theta,$$

with $\alpha = \log(y)\sqrt{c^2+T^2}$, $\theta_+ = \arctan(T/c)$ and $\theta_- = \begin{cases} 2\pi - \arctan(T/c) & \text{if } y > 1, \\ -\arctan(T/c) & \text{if } 0 < y < 1. \end{cases}$

To prove the explicit formula for $\psi(x)$ in the next section, we also need some lemmas bounding $\frac{\zeta'}{\zeta}(s)$ at various locations in the plane.[13]

Lemma 3.3. *There exists a constant $A_\Omega > 0$ such that*

$$\left|\frac{\zeta'}{\zeta}(s)\right| \leq A_\Omega \log(2|s|)$$

throughout the 'perforated' half-plane

$$\Omega = \{s \in \mathbb{C} \;:\; \sigma \leq -1 \text{ and } |s+2n| \geq \tfrac{1}{2} \text{ for all } n \in \mathbb{N}^+\}.$$

[12] Be sure to note the asymmetry between the formula for $y > 1$ and the formula for $0 < y < 1$.

[13] The notation $\frac{\zeta'}{\zeta}(s)$ is short for $\frac{\zeta'(s)}{\zeta(s)}$, the logarithmic derivative of the zeta function.

Remark 3.4. The set Ω contains every vertical line $\{-q+ti \ : \ t \in \mathbb{R}\}$, where q is an odd positive integer.

Lemma 3.5. *There exists a constant $A_{\equiv} > 0$ and real numbers $(T_m)_{m=1}^{\infty}$, such that*

$$m < T_m < m+1,$$

for all m, and such that

$$\left|\frac{\zeta'}{\zeta}(\sigma + T_m i)\right| \leq A_{\equiv} \log^2 T_m$$

for all m and all $-1 \leq \sigma \leq 2$.

For proof of Lemmas 3.3 and 3.5, the reader may consult Theorems 26 and 27 in Ingham's classical tract [18], and note that $\log(|s|+1) \leq \log(2|s|)$ throughout Ω.

Lemma 3.6. *We have*

$$-\frac{\zeta'}{\zeta}(1+\eta) < \frac{1}{\eta}$$

for all real $\eta > 0$.

Proof. Write Equation (6) on the form

$$\zeta(\sigma) = \frac{\sigma}{\sigma-1} - \sigma I(\sigma), \quad I(\sigma) = \int_1^{\infty} \frac{x - \lfloor x \rfloor}{x^{\sigma+1}} \mathrm{d}x,$$

for real $\sigma > 1$. If we differentiate this equation logarithmically, then we get

$$\frac{\zeta'}{\zeta}(\sigma) + \frac{1}{\sigma-1} = \frac{1-(2\sigma-1)I(\sigma) - \sigma(\sigma-1)I'(\sigma)}{\sigma - \sigma(\sigma-1)I(\sigma)}. \tag{35}$$

The denominator on the right is positive, being equal to $(\sigma-1)\zeta(\sigma)$. Also,

$$I'(\sigma) = -\int_1^{\infty} \frac{(x-\lfloor x \rfloor)\log x}{x^{\sigma+1}} \mathrm{d}x < 0,$$

so that $-\sigma(\sigma-1)I'(\sigma)$ in the numerator of (35) is positive as well. Thus, it suffices to show that the term $1-(2\sigma-1)I(\sigma)$ in the numerator is positive. But since $I(\sigma) = \frac{1}{\sigma-1} - \zeta(\sigma)/\sigma$, this is equivalent to showing that

$$\zeta(\sigma) > \frac{\sigma^2}{(\sigma-1)(2\sigma-1)}$$

when $\sigma > 1$. This is not too hard, since

$$\zeta(\sigma) = \sum_{n=1}^{\infty} n^{-\sigma} > 1 + \int_2^{\infty} x^{-\sigma} \mathrm{d}x = 1 + \frac{2^{1-\sigma}}{\sigma-1} = \frac{(\sigma-1+2^{1-\sigma})(2\sigma-1)}{(\sigma-1)(2\sigma-1)} > \frac{\sigma^2}{(\sigma-1)(2\sigma-1)}.$$

This shows that

$$\frac{\zeta'}{\zeta}(\sigma) + \frac{1}{\sigma-1} > 0$$

when $\sigma > 1$, and the claim now follows upon taking $\sigma = 1+\eta$. $\square$

3.2 The explicit formula

We are going to prove the explicit formula for $\psi(x)$ in this section, and our proof borrows inspiration and some notation from the classical proof.[14] This is going to require some amount of tedious calculation, and we therefore begin by giving an explanation of how the formula arises, and why the function $\psi(x)$ needs adjustment.

By applying logarithmic differentiation and thereafter Abel's summation formula to the Euler product (3), we obtain the identity

$$-\frac{\zeta'}{\zeta}(s) = s\int_1^\infty \frac{\psi(x)}{x^{s+1}}\mathrm{d}x \quad (\sigma > 1). \tag{36}$$

Thus, after the change of variable $x = \mathrm{e}^{2\pi y}$, we have

$$-\frac{1}{\sigma+ti}\cdot\frac{\zeta'}{\zeta}(\sigma+ti) = \int_{-\infty}^{\infty}\big(2\pi\psi(\mathrm{e}^{2\pi y})\mathrm{e}^{-2\pi\sigma y}\big)\mathrm{e}^{-2\pi t y i}\mathrm{d}y \tag{37}$$

(since $\psi(x)$ is zero when $x < 1$). The right hand side is the Fourier transform 'at t' of the function $f(y) = 2\pi\psi(\mathrm{e}^{2\pi y})\mathrm{e}^{-2\pi\sigma y}$, and it is conceivable that the Fourier inversion theorem would express this function as an integral involving $-(\zeta'/\zeta)(s)$. However, it is clear that Equation (36) will remain true even if $\psi(x)$ is modified on a set of (Lebesgue) measure zero, and so we cannot expect such an inverted formula to hold identically for all x. On the other hand, the function f is piecewise smooth and belongs to $L^1(\mathbb{R})$, and so the only thing missing for the inversion theorem to recover f is the property that

$$\lim_{\varepsilon\to 0}\frac{f(y-\varepsilon)+f(y+\varepsilon)}{2} = f(y)$$

at each point in its domain.[15] This is an easy fix: we define

$$\psi^\flat(x) = \lim_{\varepsilon\to 0}\frac{\psi(x-\varepsilon)+\psi(x+\varepsilon)}{2} = \frac{1}{2}\Big(\sum_{p^m\le x}\log p + \sum_{p^m<x}\log p\Big)$$

for $x\in\mathbb{R}$. Then $\psi(x)$ and $\psi^\flat(x)$ agree almost everywhere, equations (36) and (37) remain true with $\psi^\flat$ in place of ψ, and the inversion theorem yields

$$2\pi\psi^\flat(\mathrm{e}^{2\pi y})\mathrm{e}^{-2\pi\sigma y} = -\int_{-\infty}^{\infty}\frac{1}{\sigma+ti}\cdot\frac{\zeta'}{\zeta}(\sigma+ti)\mathrm{e}^{2\pi y t i}\mathrm{d}t.$$

Or, reverting back to x and noting that the right-hand side may be written as a path integral:

$$\psi^\flat(x) = -\frac{1}{2\pi i}\int_{\sigma-\infty i}^{\sigma+\infty i}\frac{x^s}{s}\cdot\frac{\zeta'}{\zeta}(s)\mathrm{d}s.$$

We would now get an explicit formula for $\psi^\flat(x)$, rather than $\psi(x)$, by writing the last integral as a sum of residues, using the residue theorem and an appropriate contour of integration.

The function $\psi^\flat(x)$, sometimes called the *normalised Chebyshev function*, is equal to $\psi(x)$ unless x coincides with a prime power: in which case it equals $\psi(x) - \frac{1}{2}\Lambda(x)$. The explicit formula is conveniently expressed in terms of $\psi^\flat(x)$, as shown in

[14] See e.g. [18] pp. 75–80, [9] §17, [22] §10, [27] §12 or [21] pp. 300–303. Observe also the difficulties in typesetting $-(\zeta'/\zeta)(s)$ in a satisfactory manner.

[15] See [38], Theorem 7.5, p. 171. Functions with this property may be called *JD-regular*, since they assume the arithmetic mean of their left and right limits at their Jump Discontinuities, and since they appear in the Jordan-Dirichlet test for the convergence of Fourier series.

Theorem 3.7. *If $x > 1$, then*

$$\psi^\flat(x) = x - \sum_\rho \frac{x^\rho}{\rho} - \frac{\zeta'}{\zeta}(0) - \frac{1}{2}\log\big(1 - x^{-2}\big). \tag{38}$$

The sum $\sum_\rho$ is taken over the nontrivial zeros of $\zeta(s)$, counted with multiplicity, and is defined as the limit of the symmetric sum

$$S(x, T) = \sum_{|\gamma|<T} \frac{x^\rho}{\rho}$$

as $T \to \infty$. Moreover, if we write $\sum_\rho x^\rho \rho^{-1} = S(x, T) + R(x, T)$, then there exists a constant $C > 0$ such that

$$|R(x, T)| \le C\Big(\frac{x\log^2(xT)}{T} + \log x\Big)$$

for all $x \ge 3$ and $T \ge 1$.[16]

Proof. Suppose $x > 1$, $c > 1$ and $T > 0$. Then

$$\begin{aligned}
\psi^\flat(x) &= \sum_{n=1}^\infty \Lambda(n) I(\tfrac{x}{n}) = \sum_{n=1}^\infty \Lambda(n) I(\tfrac{x}{n}, T) + \sum_{n=1}^\infty \Lambda(n)\Delta(\tfrac{x}{n}, T) \\
&= \sum_{n=1}^\infty \frac{\Lambda(n)}{2\pi i}\int_{c-Ti}^{c+Ti} \frac{(x/n)^s}{s}\mathrm{d}s + \sum_{n=1}^\infty \Lambda(n)\Delta(\tfrac{x}{n}, T) \\
&= \frac{1}{2\pi i}\int_{c-Ti}^{c+Ti} \frac{x^s}{s}\sum_{n=1}^\infty \frac{\Lambda(n)}{n^s}\mathrm{d}s + \sum_{n=1}^\infty \Lambda(n)\Delta(\tfrac{x}{n}, T) \\
&= \frac{1}{2\pi i}\int_{c-Ti}^{c+Ti} \frac{x^s}{s}\Big(-\frac{\zeta'}{\zeta}(s)\Big)\mathrm{d}s + \sum_{n=1}^\infty \Lambda(n)\Delta(\tfrac{x}{n}, T), \qquad (39)
\end{aligned}$$

where $I(\frac{x}{n})$, $I(\frac{x}{n}, T)$ and $\Delta(\frac{x}{n}, T)$ denote the quantities of Lemma 3.1, and where the interchange of the order of summation and integration is justified by uniform convergence of $\sum_{n=1}^\infty \Lambda(n)n^{-s}$ on a compact set containing the segment from $c - Ti$ to $c + Ti$. From our definition of $\Lambda(n)$, we have

$$\sum_{n=1}^\infty \Lambda(n)\Delta(\tfrac{x}{n}, T) = \Lambda(x)\Delta(1, T) + \sum_{\substack{n=1\\ n\ne x}}^\infty \Lambda(n)\Delta(\tfrac{x}{n}, T), \tag{40}$$

for all $x > 1$, wherein

$$|\Lambda(x)\Delta(1, T)| \begin{cases} = 0 & \text{if } x \notin \mathcal{P}^*, \\ < \log(x)\min(1, \frac{c}{\pi T}) & \text{for all } x > 1, \\ \le \Lambda(x)\min(1, \frac{c}{\pi T}) & \text{for all } x > 1, \end{cases} \tag{41}$$

and

$$\Big|\sum_{\substack{n=1\\ n\ne x}}^\infty \Lambda(n)\Delta(\tfrac{x}{n}, T)\Big| < \sum_{\substack{n=1\\ n\ne x}}^\infty \Lambda(n)(\tfrac{x}{n})^c \min(1, \frac{1}{\pi T|\log(x/n)|}). \tag{42}$$

[16]The lower bounds on x and T are imposed in order to deal with the complication that $\log x$ and $\log(xT)$ vanish for $x = 1$ and $xT = 1$, respectively.

Suppose $n \le \frac{3x}{4}$. Then $\frac{x}{n} \ge \frac{4}{3}$, so that $|\log\big(\frac{x}{n}\big)| = \log\big(\frac{x}{n}\big) \ge \log\big(\frac{4}{3}\big)$. This gives us

$$\sum_{n \le \frac{3x}{4}} \Lambda(n)(\tfrac{x}{n})^c \min(1, \frac{1}{\pi T|\log(x/n)|}) \le x^c \min(1, \frac{1}{\pi T \log(4/3)}) \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^c}$$

$$= x^c \min(1, \frac{1}{\pi T \log(4/3)})\Big(-\frac{\zeta'}{\zeta}(c)\Big). \tag{43}$$

Similarly, if $n \ge \frac{5x}{4}$, then $\frac{x}{n} \le \frac{4}{5}$, so that $|\log\big(\frac{x}{n}\big)| = \log\big(\frac{n}{x}\big) \ge \log\big(\frac{5}{4}\big)$. Therefore,

$$\sum_{n \ge \frac{5x}{4}} \Lambda(n)(\tfrac{x}{n})^c \min(1, \frac{1}{\pi T|\log(x/n)|}) \le x^c \min(1, \frac{1}{\pi T \log(5/4)})\Big(-\frac{\zeta'}{\zeta}(c)\Big). \tag{44}$$

It remains to consider the terms where $\frac{3x}{4} < n < \frac{5x}{4}$ and $n \ne x$. We split these into two different cases.

Case 1: $\boldsymbol{\frac{3x}{4} < n < x}$.
If there is no prime power n with $\frac{3x}{4} < n < x$, then there is nothing to consider, since all the terms of (40) corresponding to these values of n are zero in that case. Otherwise, let p_x^* denote the largest prime power in this interval. Then

$$\log\left(\frac{x}{p_x^*}\right) = -\log\left(\frac{p_x^*}{x}\right) = -\log\left(1 - \frac{x - p_x^*}{x}\right) > \frac{x - p_x^*}{x}.$$

Therefore, the term corresponding to $n = p_x^*$ to the right in (42) is bounded by

$$\Lambda(p_x^*)\left(\frac{x}{p_x^*}\right)^c \min(1, \frac{x}{\pi T(x - p_x^*)}) \le (\tfrac{4}{3})^c \log(x) \min(1, \frac{x}{\pi T(x - p_x^*)}). \tag{45}$$

For any other prime power n in this interval, we may write $n = p_x^* - \nu$ with $0 < \nu < x/4$. This yields

$$\log\left(\frac{x}{n}\right) \ge \log\left(\frac{p_x^*}{n}\right) = -\log\left(\frac{n}{p_x^*}\right) = -\log\left(1 - \frac{\nu}{p_x^*}\right) > \frac{\nu}{p_x^*}.$$

The contribution of these terms to the right in (42) is therefore bounded by

$$\sum_{\substack{0<\nu<x/4 \\ p_x^*-\nu \in \mathcal{P}^*}} \Lambda(p_x^* - \nu)(\tfrac{x}{p_x^*-\nu})^c \min(1, \frac{p_x^*}{\pi T \nu}) \le (\tfrac{4}{3})^c \log(x) \sum_{\substack{0<\nu<x/4 \\ \nu \in \mathbb{Z}}} \min(1, \frac{p_x^*}{\pi T \nu})$$

$$\le (\tfrac{4}{3})^c \log(x) \frac{p_x^*}{\pi T} \sum_{\substack{0<\nu<x/4 \\ \nu \in \mathbb{Z}}} \frac{1}{\nu}$$

$$\le \tfrac{1}{\pi}(\tfrac{4}{3})^c \log(x) \log(x+1) \frac{x}{T}, \tag{46}$$

since $\sum_{\substack{0<\nu<x/4 \\ \nu\in\mathbb{Z}}} \frac{1}{\nu} \le \log(x+1)$ for all $x \ge 0$ (cf. (A2.3) in the Appendix).

Case 2: $\boldsymbol{x < n < \frac{5x}{4}}$.
Again, if there is no prime power n with $x < n < \frac{5x}{4}$, then we are done. Otherwise, let P_x^* denote the smallest prime power in this interval. Then

$$\left|\log\left(\frac{x}{P_x^*}\right)\right| = -\log\left(\frac{x}{P_x^*}\right) = -\log\left(1 - \frac{P_x^* - x}{P_x^*}\right) > \frac{P_x^* - x}{P_x^*}.$$

Therefore, the term corresponding to $n = P_x^*$ to the right in (42) is bounded by

$$\Lambda(P_x^*)\left(\frac{x}{P_x^*}\right)^c \min(1, \frac{P_x^*}{\pi T(P_x^* - x)}) \le \log\left(\tfrac{5x}{4}\right) \min(1, \frac{5x}{4\pi T(P_x^* - x)}). \tag{47}$$

For any other prime power n in this interval, we may write $n = P_x^* + \nu$ with $0 < \nu < x/4$. This yields

$$\left|\log\left(\frac{x}{n}\right)\right| = \log\left(\frac{n}{x}\right) > \log\left(\frac{n}{P_x^*}\right) = -\log\left(\frac{P_x^*}{n}\right) = -\log\left(1 - \frac{\nu}{n}\right) > \frac{\nu}{n}.$$

The contribution of these terms in (42) is therefore bounded by

$$\begin{aligned}
\sum_{\substack{0<\nu<x/4\\ P_x^*-\nu\in\mathcal{P}^*}} \Lambda(P_x^*+\nu)(\tfrac{x}{P_x^*+\nu})^c \min(1, \frac{P_x^*+\nu}{\pi T\nu}) &\le \log\left(\tfrac{5x}{4}\right) \sum_{\substack{0<\nu<x/4\\ \nu\in\mathbb{Z}}} \min(1, \frac{5x}{4\pi T\nu}) \\
&\le \tfrac{5}{4\pi}\log\left(\tfrac{5x}{4}\right)\frac{x}{T} \sum_{\substack{0<\nu<x/4\\ \nu\in\mathbb{Z}}} \frac{1}{\nu} \\
&\le \tfrac{5}{4\pi}\log\left(\tfrac{5x}{4}\right)\log(x+1)\frac{x}{T}.
\end{aligned} \tag{48}$$

Collecting the terms.
For simplicity, let $\langle x\rangle = \mathrm{dist}(x, \mathcal{P}^* \setminus \{x\})$ denote the distance from x to the nearest prime power that is not equal to x. In particular, $x - p_x^*,\ P_x^* - x \ge \langle x\rangle$ provided that p_x^* and P_x^* exist. Equations (39) through (48) now give, for $x > 1, c > 1$ and $T > 0$:

$$\begin{aligned}
&\left|\psi^\flat(x) - \frac{1}{2\pi i}\int_{c-Ti}^{c+Ti} \frac{x^s}{s}\Big(-\frac{\zeta'}{\zeta}(s)\Big)\mathrm{d}s\right| \\
&\quad\le \Lambda(x)\min(1, \tfrac{c}{\pi T}) + (\tfrac{4}{3})^c\log(x)\min(1, \tfrac{x}{\pi T\langle x\rangle}) + \log\left(\tfrac{5x}{4}\right)\min(1, \tfrac{5x}{4\pi T\langle x\rangle}) \\
&\quad+ x^c\min(1, \frac{1}{\pi T\log(4/3)})\Big(-\frac{\zeta'}{\zeta}(c)\Big) + x^c\min(1, \frac{1}{\pi T\log(5/4)})\Big(-\frac{\zeta'}{\zeta}(c)\Big) \\
&\quad+ \tfrac{1}{\pi}(\tfrac{4}{3})^c\log(x)\log(x+1)\frac{x}{T} + \tfrac{5}{4\pi}\log\left(\tfrac{5x}{4}\right)\log(x+1)\frac{x}{T}.
\end{aligned}$$

Let us concretize these bounds. Take $c = c(x) = 1 + \frac{1}{\log(x+\mathrm{e}-1)}$ and suppose that $T \ge 3$. Then

(i) $1 < c < 2$ for $x > 1$, and $x^c \le \mathrm{e}x$.

(ii) In particular, $-\frac{\zeta'}{\zeta}(c) \le \log(x + \mathrm{e} - 1)$ by Lemma 3.6.

(iii) $\frac{c}{\pi T} \le \frac{2}{3\pi} < 1$, so that $\min(1, \frac{c}{\pi T}) = \frac{c}{\pi T} \le \frac{2}{\pi T}$.

(iv) $\frac{1}{\pi T\log(4/3)} \le \frac{1}{3\pi\log(4/3)} < 1$, so that $\min(1, \frac{1}{\pi T\log(4/3)}) = \frac{1}{\pi T\log(4/3)}$.

(v) $\frac{1}{\pi T\log(5/4)} \le \frac{1}{3\pi\log(5/4)} < 1$, so that $\min(1, \frac{1}{\pi T\log(5/4)}) = \frac{1}{\pi T\log(5/4)}$.

This yields

$$\begin{aligned}
&\left|\psi^\flat(x) - \frac{1}{2\pi i}\int_{c-Ti}^{c+Ti} \frac{x^s}{s}\Big(-\frac{\zeta'}{\zeta}(s)\Big)\mathrm{d}s\right| \\
&\quad\le \tfrac{2}{\pi T}\Lambda(x) + (\tfrac{4}{3})^2\log(x)\min(1, \tfrac{x}{\pi T\langle x\rangle}) + \log\left(\tfrac{5x}{4}\right)\min(1, \tfrac{5x}{4\pi T\langle x\rangle}) \\
&\quad+ \tfrac{\mathrm{e}}{\pi}\left(\tfrac{1}{\log(4/3)} + \tfrac{1}{\log(5/4)}\right)\tfrac{x\log(x+\mathrm{e}-1)}{T} \\
&\quad+ \tfrac{1}{\pi}(\tfrac{4}{3})^2\log(x)\log(x+1)\tfrac{x}{T} + \tfrac{5}{4\pi}\log\left(\tfrac{5x}{4}\right)\log(x+1)\tfrac{x}{T}.
\end{aligned} \tag{49}$$

Modifying the contour.
Let q be an odd positive integer, and R the rectangle whose vertices are

$$c + Ti, \quad -q + Ti, \quad -q - Ti, \quad c - Ti,$$

see Figure 4.

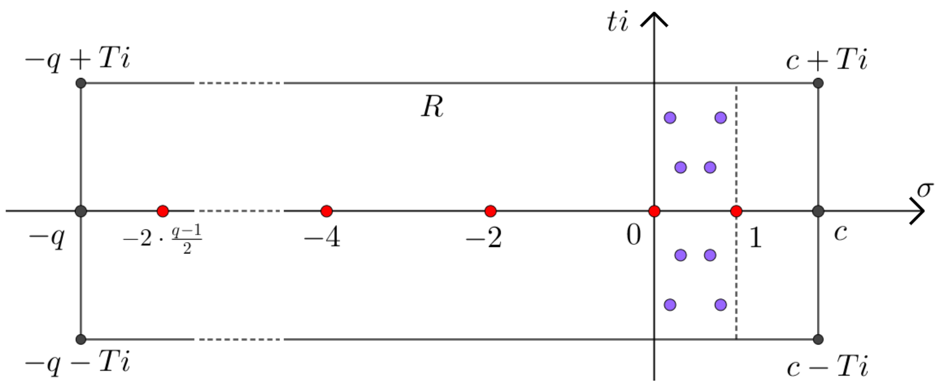

Figure 4: The red and purple dots are the poles of ζ'/ζ in R that lie outside and inside of the critical strip, respectively.

In order to avoid that $\zeta(s)$ has a zero on the horizontal segments of ∂R, replace T by $\widetilde{T}$, the smallest T_m from Lemma 3.5 that is greater than or equal to T. From the residue theorem applied to $g(s) = -\frac{x^s}{s}\frac{\zeta'}{\zeta}(s)$ over R, we get

$$\frac{1}{2\pi i}\int_{c-\widetilde{T}i}^{c+\widetilde{T}i} g(s)\mathrm{d}s = \sum_{\substack{\omega\in \mathrm{int}(R)\\ g(\omega)=\infty}} \operatorname*{Res}_{s=\omega} g(s) - \frac{1}{2\pi i}\Big(\int_{c+\widetilde{T}i}^{-q+\widetilde{T}i} + \int_{-q+\widetilde{T}i}^{-q-\widetilde{T}i} + \int_{-q-\widetilde{T}i}^{c-\widetilde{T}i}\Big) g(s)\mathrm{d}s. \tag{50}$$

The (necessarily simple) poles of g in $\mathrm{int}(R)$ are 1, 0, and $-2k$ for $1 \leq k \leq \frac{q-1}{2}$, as well as any nontrivial zero $\rho = \beta + \gamma i$ of $\zeta(s)$ with $|\gamma| < \widetilde{T}$. The residues at these points are[17]

$$\begin{aligned}
\operatorname*{Res}_{s=1} g(s) &= \lim_{s\to 1} -\frac{x^s}{s}(s-1)\frac{\zeta'}{\zeta}(s) = -x \operatorname*{Res}_{s=1}\frac{\zeta'}{\zeta}(s) = x,\\
\operatorname*{Res}_{s=0} g(s) &= \lim_{s\to 0} -x^s\frac{\zeta'}{\zeta}(s) = -\frac{\zeta'}{\zeta}(0),\\
\operatorname*{Res}_{s=-2k} g(s) &= \lim_{s\to -2k} -\frac{x^s}{s}(s+2k)\frac{\zeta'}{\zeta}(s) = -\frac{x^{-2k}}{-2k}\operatorname*{Res}_{s=-2k}\frac{\zeta'}{\zeta}(s) = -\frac{x^{-2k}}{-2k},\\
\operatorname*{Res}_{s=\rho} g(s) &= \lim_{s\to \rho} -\frac{x^s}{s}(s-\rho)\frac{\zeta'}{\zeta}(s) = -\frac{x^\rho}{\rho}\operatorname*{Res}_{s=\rho}\frac{\zeta'}{\zeta}(s) = -\frac{x^\rho}{\rho}N_\rho,
\end{aligned}$$

where N_ρ denotes the order of the nontrivial zero ρ of $\zeta(s)$. Accordingly,

$$\sum_{\substack{\omega\in \mathrm{int}(R)\\ g(\omega)=\infty}} \operatorname*{Res}_{s=\omega} g(s) = x - \sum_{|\gamma|<\widetilde{T}} \frac{x^\rho}{\rho} - \frac{\zeta'}{\zeta}(0) - \sum_{k=1}^{(q-1)/2}\frac{x^{-2k}}{-2k}. \tag{51}$$

[17] A zero of f of order n is a simple pole of f'/f with residue n, and a pole of f of order n is a simple pole of f'/f with residue $-n$.

The vertical integral to the right in (50) can be bounded using Lemma 3.3. We have

$$\left|\frac{1}{2\pi i}\int_{-q+\widetilde{T}i}^{-q-\widetilde{T}i} -\frac{x^s}{s}\frac{\zeta'}{\zeta}(s)\right| \leq \frac{1}{2\pi}\int_{-\widetilde{T}}^{\widetilde{T}} \frac{x^{-q}}{q} A_\Omega \log\left(2\sqrt{q^2+t^2}\right)\mathrm{d}t \leq \frac{A_\Omega \widetilde{T}}{\pi q x^q}\log\left(2\sqrt{q^2+\widetilde{T}^2}\right). \tag{52}$$

Since $x > 1$, this integral vanishes as $q \to \infty$ when $\widetilde{T}$ stays fixed.

For the horizontal integrals to the right in (50), we consider the ranges $-1 \leq \sigma \leq 2$ and $-q \leq \sigma \leq -1$ separately. We first have, by Lemma 3.5:

$$\begin{aligned}\left|\frac{1}{2\pi i}\int_{-1\pm\widetilde{T}i}^{c\pm\widetilde{T}i} -\frac{x^s}{s}\frac{\zeta'}{\zeta}(s)\mathrm{d}s\right| &\leq \frac{A_\equiv \log^2 \widetilde{T}}{2\pi\widetilde{T}}\int_{-1}^{c} x^\sigma \mathrm{d}\sigma \leq \frac{A_\equiv \log^2 \widetilde{T}}{2\pi\widetilde{T}}\int_{-\infty}^{c} x^\sigma \mathrm{d}\sigma \\ &= \frac{A_\equiv \log^2 \widetilde{T}}{2\pi\widetilde{T}}\frac{x^c}{\log x} \leq \frac{\mathrm{e}A_\equiv}{2\pi}\cdot\frac{x\log^2\widetilde{T}}{\log(x)\widetilde{T}}.\end{aligned} \tag{53}$$

For the leftmost part of these integrals, we have again by Lemma 3.3

$$\begin{aligned}\left|\frac{1}{2\pi i}\int_{-q\pm\widetilde{T}i}^{-1\pm\widetilde{T}i} -\frac{x^s}{s}\frac{\zeta'}{\zeta}(s)\mathrm{d}s\right| &\leq \frac{A_\Omega}{2\pi}\int_{-q}^{-1} x^\sigma \frac{\log\left(2\sqrt{\sigma^2+\widetilde{T}^2}\right)}{\sqrt{\sigma^2+\widetilde{T}^2}}\mathrm{d}\sigma \leq \frac{A_\Omega \log\left(2\widetilde{T}\right)}{2\pi\widetilde{T}}\int_{-\infty}^{-1} x^\sigma\mathrm{d}\sigma \\ &= \frac{A_\Omega\log\left(2\widetilde{T}\right)}{2\pi\widetilde{T}x\log(x)},\end{aligned} \tag{54}$$

since $v(\sigma) = \frac{\log\left(2\sqrt{\sigma^2+\widetilde{T}^2}\right)}{\sqrt{\sigma^2+\widetilde{T}^2}}$ is decreasing for[18] $\sigma \geq \sigma(\widetilde{T}) = \begin{cases}\sqrt{(\mathrm{e}/2)^2-\widetilde{T}^2} & \text{if } |\widetilde{T}| \leq \mathrm{e}/2,\\ 0 & \text{if } |\widetilde{T}| > \mathrm{e}/2,\end{cases}$ and since $\widetilde{T} \geq 3 > \mathrm{e}/2 \geq \sigma(\widetilde{T})$.

Conclusion

In the limit as $q \to \infty$, we get from equations (39), (49), (50), (51), (52), (53) and (54):

$$\psi^\flat(x) = x - \sum_{|\gamma|<\widetilde{T}}\frac{x^\rho}{\rho} - \frac{\zeta'}{\zeta}(0) - \frac{1}{2}\log(1-x^{-2}) + R(x,\widetilde{T}), \tag{55}$$

with

$$|R(x,\widetilde{T})| \leq \frac{\mathrm{e}A_\equiv}{\pi}\cdot\frac{x\log^2\widetilde{T}}{\log(x)\widetilde{T}} + \frac{A_\Omega\log\left(2\widetilde{T}\right)}{\pi\widetilde{T}x\log(x)} + Q(x,\widetilde{T}), \tag{56}$$

where $Q(x,\widetilde{T})$ is three-line upper bound specified by Equation (49). Since $Q(x,\widetilde{T}) \to 0$ when $\widetilde{T} \to \infty$ and $x > 1$ is kept fixed, we obtain the explicit formula (38) from (55) and (56) for $x > 1$. For $x \geq 3$ and $\widetilde{T} \geq 3$ specifically, we have $\log(\frac{5x}{4})$, $\log(x+\mathrm{e}-1)$, $\log(x+1) \leq 2\log(x)$ and $\log(x) > 1$, and so the bound takes the explicit form

$$\begin{aligned}|R(x,\widetilde{T})| \leq{}& \frac{\mathrm{e}A_\equiv}{\pi}\cdot\frac{x\log^2\widetilde{T}}{\log(x)\widetilde{T}} + \frac{A_\Omega}{\pi}\cdot\frac{\log\left(2\widetilde{T}\right)}{\widetilde{T}x\log(x)} + \tfrac{2}{\pi\widetilde{T}}\Lambda(x) + (\tfrac{4}{3})^2\log(x)\min(1,\tfrac{x}{\pi\widetilde{T}\langle x\rangle}) \\ &+ 2\log(x)\min(1,\tfrac{5x}{4\pi\widetilde{T}\langle x\rangle}) + \tfrac{2\mathrm{e}}{\pi}\left(\tfrac{1}{\log(4/3)}+\tfrac{1}{\log(5/4)}\right)\tfrac{x\log(x)}{\widetilde{T}} \\ &+ (\tfrac{32}{9\pi}+\tfrac{5}{\pi})\log^2(x)\tfrac{x}{\widetilde{T}} \\ \leq{}& M\frac{x\log^2(x\widetilde{T})}{\widetilde{T}} + K\log x \leq \max(M,K)\left(\frac{x\log^2(x\widetilde{T})}{\widetilde{T}} + \log x\right)\end{aligned}$$

[18] $v'(\sigma) = \sigma(1-\log\left(2\sqrt{\sigma^2+\widetilde{T}^2}\right))(\sigma^2+\widetilde{T}^2)^{-3/2}$.

with $M = \frac{\mathrm{e}A_{\equiv}}{\pi} + \frac{2\mathrm{e}}{\pi}(\frac{1}{\log(4/3)} + \frac{1}{\log(5/4)}) + \frac{32}{9\pi} + \frac{5}{\pi}$ and $K = \frac{A_\Omega}{\pi} + \frac{2}{3\pi} + (\frac{4}{3})^2 + 2$.

Now this applies only with $\widetilde{T}$ in place of T, but according to Lemma 3.5, we always have $\widetilde{T} - T < 2$ for $T \geq 3$. From the definition of $R(x,T)$ and Equation (38), we have additionally

$$R(x,\widetilde{T}) - R(x,T) = \sum_{T \leq |\gamma| < \widetilde{T}} \frac{x^\rho}{\rho},$$

with

$$\Bigg| \sum_{T \leq |\gamma| < \widetilde{T}} \frac{x^\rho}{\rho} \Bigg| \leq \frac{2x}{T} \sum_{T \leq \gamma < \widetilde{T}} 1 \leq \frac{2x}{T} \sum_{T \leq \gamma < T+2} 1 \leq 2A \frac{x \log T}{T}.$$

for some constant $A > 0$. Thus, we obtain

$$\begin{aligned} |R(x,T)| &\leq 2A\frac{x\log T}{T} + M\frac{x\log^2(x\widetilde{T})}{\widetilde{T}} + K\log x \\ &\leq 2A\frac{x\log T}{T} + M\frac{x\log^2(x(T+2))}{T} + K\log x \\ &\leq \widetilde{M}\frac{x\log^2(xT)}{T} + K\log x \leq \max(\widetilde{M},K)\left(\frac{x\log^2(xT)}{T} + \log x\right) \end{aligned} \tag{57}$$

for all $x, T \geq 3$ with $\widetilde{M} = 2A + M(\log(18)/\log(9))^2$, since $x(T+2) < 2xT$, and since $\log(2\lambda) \leq (\log(18)/\log(9))\log(\lambda)$ for $\lambda = xT \geq 9$.

The bound (57) can now be seen to hold in fact for $x \geq 3$ and $T \geq 1$. Indeed, the smallest ordinate of a nontrivial zero is $T_0 = 14.1347\ldots$ (see [39]), which implies that $R(x,T) = R(x,T_0)$ for all $0 < T \leq T_0$. It therefore suffices to show that $x\log^2(xT)/T \geq x\log^2(xT_0)/T_0$ for these values of T. The function $f(T) = \log^2(xT)/T$ is increasing on $(1/x, \mathrm{e}^2/x)$ and decreasing on $(\mathrm{e}^2/x, \infty)$, and so the claim is obvious if $x \geq \mathrm{e}^2$. If $x < \mathrm{e}^2$, then it suffices to show that $f(1) \geq f(T_0)$. The resulting equation $T_0\log^2(x) \geq \log^2(xT_0)$ can then be seen to hold for all $x \geq 3$. □

There are many formulas similar to (38) that can be proved in essentially the same way as Theorem 3.7 was proved above. We only mention one of these here, which we are going to need later.

Proposition 3.8. *If $\psi_1(x) = \int_1^x \psi(t)\mathrm{d}t = \sum_{n\leq x}\Lambda(n)(x-n)$ for $x \in \mathbb{R}$, then*

$$\psi_1(x) = \frac{x^2}{2} - \sum_\rho \frac{x^{\rho+1}}{\rho(\rho+1)} - x\frac{\zeta'}{\zeta}(0) + \frac{\zeta'}{\zeta}(-1) - \sum_{r=1}^{\infty}\frac{x^{1-2r}}{2r(2r-1)} \tag{58}$$

for $x \geq 1$.

To prove this, a generalisation of Lemma 3.1 is needed. A derivation can be found in [19].

4 Primes in short intervals: weak hypotheses

4.1 Introduction

The prime number theorem tells us that the number of primes in $[0, x]$ and $[x, 2x]$ are both asymptotic to $x(\log x)^{-1}$, and in fact that[19]

$$\psi(x + x^\vartheta) - \psi(x) \begin{cases} \sim x^\vartheta & \text{if } \vartheta \geq 1, \\ = o(x) & \text{if } \vartheta < 1, \end{cases} \quad \text{and} \quad \pi(x + x^\vartheta) - \pi(x) \begin{cases} \sim \dfrac{x^\vartheta}{\vartheta \log x} & \text{if } \vartheta \geq 1, \\ = o\Big(\dfrac{x}{\log x}\Big) & \text{if } \vartheta < 1. \end{cases} \tag{59}$$

Moreover, Gauss observed that the density of primes is well approximated by $(\log t)^{-1}$, in the sense that

$$\pi(x) \sim \int_2^x \frac{\mathrm{d}t}{\log t}.$$

We can therefore guess that

$$\pi(x + h) - \pi(x) \approx \int_x^{x+h} \frac{\mathrm{d}t}{\log t} \approx \frac{h}{\log x},$$

provided x is large and h is small compared to x. It would therefore be interesting to know if an asymptotic prime number theorem such as (59) holds for the interval $(x, x + h]$, where $h = h(x)$ is a nonnegative function that is smaller than x in the sense that $h = o(x)$ as $x \to \infty$. An interval of this form is usually called a *short interval*.

Such asymptotic formulae can of course not hold for arbitrarily small h. For example, if h is bounded, then the infinitude of x for which $\pi(x+h) - \pi(x) > 0$, has been established unconditionally only in the case when $h \geq 246$ (see [28]). Moreover, Maier [25] showed that the asymptotic formula $\pi(x + h) - \pi(x) \sim h(\log x)^{-1}$ is false for $h = \log^A x$, for any fixed number A. Our attention will therefore be restricted to those functions h that grow faster than this.

We shall mainly be concerned with estimates of the difference $\psi(x + h) - \psi(x)$, the reason being

Proposition 4.1. *Let h be a function of x such that $h = o(x \log x)$ and $\log x \log\log x = o(h)$. Then the statements*

$$\psi(x + h) - \psi(x) \sim h \quad \textit{and} \quad \pi(x + h) - \pi(x) \sim \frac{h}{\log x}$$

($as\ x \to \infty$), are either both true or both false.

Proof. For simplicity, let $\mathcal{F}_\bullet[\cdot]$ be the one-step forward difference operator, defined for functions f and g, and $x \in \mathbb{R}$, by

$$\mathcal{F}_g[f](x) = f\big(x + g(x)\big) - f(x).$$

Recall the definition of the function $\Pi(x) = \Sigma_{p^m \leq x}\, m^{-1}$, and let us write $\Pi(x) = \pi(x) + E(x)$. By

[19] See the Appendix for elaboration.

Abel's summation formula (A2.1), we then have, for $y \geq x > 1$,

$$
\begin{aligned}
\psi(y) - \psi(x) = \sum_{x<n\leq y} \Lambda(n) &= \log(y)\Pi(y) - \log(x)\Pi(x) - \int_x^y \frac{\Pi(t)}{t}\mathrm{d}t \\
&= \log(y)\pi(y) - \log(x)\pi(x) + \log(y)E(y) - \log(x)E(x) - \int_x^y \frac{\Pi(t)}{t}\mathrm{d}t \\
&= \mathcal{F}_{y-x}\big[\log(\cdot)\pi(\cdot)\big](x) + \mathcal{F}_{y-x}\big[\log(\cdot)E(\cdot)\big](x) - \int_x^y \frac{\Pi(t)}{t}\mathrm{d}t. \qquad (60)
\end{aligned}
$$

Take $y = x + h$, where $h = h(x) \geq 0$. Then

$$\mathcal{F}_h\big[\log(\cdot)f(\cdot)\big](x) = \log(x)\mathcal{F}_h[f](x) + \log\big(1 + \tfrac{h}{x}\big)f(x+h)$$

for 'any' function f. If we use this in Equation (60) for $f = \pi$ and $f = E$, and divide through by $\log x$, then we get

$$
\begin{aligned}
\frac{\mathcal{F}_h[\psi](x)}{\log x} &= \mathcal{F}_h[\pi](x) + \mathcal{F}_h[E](x) + \frac{\log\big(1+\frac{h}{x}\big)\pi(x+h)}{\log x} + \frac{\log\big(1+\frac{h}{x}\big)E(x+h)}{\log x} - \frac{1}{\log x}\int_x^{x+h} \frac{\Pi(t)}{t}\mathrm{d}t \\
&= \mathcal{F}_h[\pi](x) + \mathcal{F}_h[E](x) + \frac{\log\big(1+\frac{h}{x}\big)\Pi(x+h)}{\log x} - \frac{1}{\log x}\int_x^{x+h} \frac{\Pi(t)}{t}\mathrm{d}t. \qquad (61)
\end{aligned}
$$

Observe that

$$\frac{\log\big(1+\frac{h}{x}\big)\Pi(x)}{\log x} = \frac{\Pi(x)}{\log x}\int_x^{x+h}\frac{\mathrm{d}t}{t} \leq \frac{1}{\log x}\int_x^{x+h}\frac{\Pi(t)}{t}\mathrm{d}t \leq \frac{\Pi(x+h)}{\log x}\int_x^{x+h}\frac{\mathrm{d}t}{t} = \frac{\log\big(1+\frac{h}{x}\big)\Pi(x+h)}{\log x}.$$

This implies that the two last terms to the right in (61) together are nonnegative, with

$$
\begin{aligned}
0 &\leq \frac{\log\big(1+\frac{h}{x}\big)\Pi(x+h)}{\log x} - \frac{1}{\log x}\int_x^{x+h}\frac{\Pi(t)}{t}\mathrm{d}t \leq \frac{\log\big(1+\frac{h}{x}\big)}{\log x}\big(\Pi(x+h) - \Pi(x)\big) \\
&\leq \frac{h}{x\log x}\sum_{x<p^m\leq x+h}\frac{1}{m} = \frac{h}{x\log x}\sum_{x<p^m\leq x+h}\frac{\log p}{m\log p} \leq \frac{h}{x\log^2 x}\sum_{x<p^m\leq x+h}\log p \\
&= \frac{h}{x\log^2 x}\mathcal{F}_h[\psi](x) = o\Big(\frac{\mathcal{F}_h[\psi](x)}{\log x}\Big), \qquad (62)
\end{aligned}
$$

since $h = o(x\log x)$.

It remains only to consider $\mathcal{F}_h[E](x)$. To this end we take $N = \lfloor \log_2(x+h)\rfloor$, and use the definition of E to estimate

$$\begin{aligned}
\mathcal{F}_h[E](x) &= \sum_{\substack{x<p^m\leq x+h\\ m\geq 2}} \frac{1}{m} = \sum_{m=2}^{N} \sum_{x<p^m\leq x+h} \frac{1}{m} = \sum_{m=2}^{N} \frac{1}{m} \sum_{x^{1/m}<p\leq (x+h)^{1/m}} 1 \\
&= \sum_{m=2}^{N} \frac{\pi\big((x+h)^{1/m}\big)-\pi(x^{1/m})}{m} = \sum_{m=2}^{N} \frac{\pi\big(x^{1/m}+(x+h)^{1/m}-x^{1/m}\big)-\pi(x^{1/m})}{m} \\
&\leq \sum_{m=2}^{N} \frac{(x+h)^{1/m}-x^{1/m}+1}{m} = \sum_{m=2}^{N} \frac{x^{1/m}\big((1+\frac{h}{x})^{1/m}-1\big)+1}{m} \\
&\leq \sum_{m=2}^{N} \frac{hx^{1/m}}{xm^2} + \sum_{m=2}^{N} \frac{1}{m} \leq \frac{h}{\sqrt{x}}\Big(\sum_{m=2}^{\infty} \frac{1}{m^2}\Big) + \log N \\
&= O\Big(\frac{h}{\sqrt{x}} + \log\log x\Big) = o\Big(\frac{h}{\log x}\Big). \qquad (63)
\end{aligned}$$

In the third line we used the bound (A2.4) for the prime-counting function, and in the fourth line we used Bernoulli's inequality (A2.5) together with the bound (A2.3) for the harmonic series. The last line follows from the assumptions on h.

Thus, (61), (62) and (63) give

$$\mathcal{F}_h[\pi](x) = \frac{\mathcal{F}_h[\psi](x)}{\log x}\big(1+o(1)\big) + o\Big(\frac{h}{\log x}\Big), \qquad (64)$$

from which the stated claim follows easily. □

An asymptotic formula for $\pi(x+h)-\pi(x)$ also gives us for free an upper bound on prime gaps, which in our case can be stated as

Proposition 4.2. *Suppose that $h = h(x)$ is such that $\log x = o(h)$ and*

$$\pi(x+h)-\pi(x) \sim \frac{h}{\log x}$$

as $x \to \infty$. Then

$$p_{n+1} - p_n \leq h(p_n)$$

for all sufficiently large n.

Proof. By assumption, we have that $h(\log x)^{-1}$ goes to infinity with x, and the asymptotic equivalence therefore implies that there exists a positive x_0 such that $\pi(x+h)-\pi(x) \geq 1$ for all $x \geq x_0$; In particular, h is positive when $x \geq x_0$. Let p_n be any prime with $p_n \geq x_0$. Then $(p_n, p_n + h(p_n)]$ contains at least one prime, and one of these must be p_{n+1}. As such, we have

$$p_{n+1} - p_n \leq h(p_n).$$

□

4.2 Hoheisel's theorem and generalisations

We are now ready to begin our investigation proper. Our starting point is the explicit formula for $\psi(x)$, on the form

$$\psi(x) = x - \sum_{|\gamma|\leq T} \frac{x^\rho}{\rho} + O\Big(\frac{x\log^2(xT)}{T} + \log x\Big),$$

uniformly for $x \geq 3$ and $T \geq 1$. If h goes to infinity with x in such a way that $h = o(x)$ (or, in general, $h \ll x$), then

$$\psi(x+h) = x + h - \sum_{|\gamma| \leq T} \frac{(x+h)^\rho}{\rho} + O\Big(\frac{x \log^2(xT)}{T} + \log x\Big),$$

and hence

$$\psi(x+h) - \psi(x) - h = -\sum_{|\gamma| \leq T} C(\rho) + O\Big(\frac{x \log^2(xT)}{T} + \log x\Big), \tag{65}$$

uniformly for $x \geq x_0(h) \geq 3$ and $T \geq 1$, where

$$C(\rho) = \frac{(x+h)^\rho - x^\rho}{\rho}$$

(the dependence of $C(\rho)$ on x and h is implicitly understood).

The point is: if the right hand side of (65) is $o(h)$ as $x \to \infty$ for suitable choices of $h = h(x)$ and $T = T(x)$, then the conclusion is that

$$\psi(x+h) - \psi(x) \sim h$$

as $x \to \infty$. In our search for such a pair (h, T), we make the simplifying assumption that h and T are chosen in such a way that

$$\log x = o(h), \tag{66}$$

$$x \log^2(xT) = o(hT), \tag{67}$$

because this reduces the O-term in (65) to $o(h)$. The restriction (66) is not important in light of Maier's theorem, while (67) will be seen to put a restriction on h and T which is difficult to overcome using the strategies employed here; we discuss this briefly in Section 6.

We can derive two general bounds on the number $C(\rho)$. If $h \ll x$, then we have we have first

$$|C(\rho)| \leq \frac{(x+h)^\beta + x^\beta}{|\gamma|} \ll \frac{x^\beta}{|\gamma|} \tag{68}$$

uniformly over all $\rho = \beta + \gamma i$ as $x \to \infty$, by the triangle inequality. On the other hand, we have by integration

$$|C(\rho)| = \Big| \int_x^{x+h} t^{\rho-1} \mathrm{d}t \Big| \leq \int_x^{x+h} t^{\beta-1} \mathrm{d}t \ll h x^{\beta-1} \tag{69}$$

as $x \to \infty$.

Remark 4.3. In terms of real and imaginary parts, it can be computed that

$$C(\rho) = \frac{(x+h)^\beta P(x+h) - x^\beta P(x)}{\beta^2 + \gamma^2} + \frac{(x+h)^\beta Q(x+h) - x^\beta Q(x)}{\beta^2 + \gamma^2} i \quad \text{and}$$
$$|C(\rho)|^2 = \frac{(x+h)^{2\beta} - 2(x+h)^\beta x^\beta \cos\big(\gamma \log\big(1 + hx^{-1}\big)\big) + x^{2\beta}}{\beta^2 + \gamma^2},$$

where

$$P(t) = \beta \cos(\gamma \log t) + \gamma \sin(\gamma \log t),$$
$$Q(t) = \beta \sin(\gamma \log t) - \gamma \cos(\gamma \log t),$$
$$P^2(t) + Q^2(t) \equiv 1 \quad \text{and}$$
$$P(x+h)P(x) + Q(x+h)Q(x) = (\beta^2 + \gamma^2) \cos\big(\gamma \log\big(1 + hx^{-1}\big)\big).$$

We are going to use the bounds (68) and (69) to estimate the sum $\sum_{|\gamma|\le T} C(\rho)$. In general, it is useful to consider the nontrivial zeros of $\zeta(s)$ which lie to the left and right inside the critical strip separately.[20] We therefore introduce a constant $\frac{1}{2} \le \kappa < 1$ and write

$$\sum_{|\gamma|\le T} C(\rho) = \sum_{\substack{|\gamma|\le T \\ \beta<\kappa}} C(\rho) + \sum_{\substack{|\gamma|\le T \\ \beta\ge\kappa}} C(\rho). \tag{70}$$

We make a trivial estimate of the first sum using (69), to obtain

$$\sum_{\substack{|\gamma|\le T \\ \beta<\kappa}} C(\rho) \ll h \sum_{\substack{|\gamma|\le T \\ \beta<\kappa}} x^{\beta-1} \le hx^{\kappa-1}N(T) \ll hx^{\kappa-1}T\log T,$$

which is $o(h)$ if

$$T\log T = o(x^{1-\kappa}) \tag{71}$$

as $x \to \infty$.

For the first theorem we would like to prove, it suffices to use (69) when bounding the second sum to the right in (70) as well. Let $\beta_1 \le \beta_2 \le \ldots \le \beta_m$ denote the abscissas of the zeros $\rho = \beta + \gamma i$ of $\zeta(s)$ for which $\beta \ge \kappa$ and $0 < \gamma \le T$. If $\varphi(\sigma) = x^{\sigma-1}$, then

$$\begin{aligned}
\sum_{\substack{|\gamma|\le T \\ \beta\ge\kappa}} C(\rho) &\ll h \sum_{\substack{0<\gamma\le T \\ \beta\ge\kappa}} x^{\beta-1} = h\sum_{n=1}^{m} \varphi(\beta_n) = h\Big(\sum_{n=1}^{m} 1 - \sum_{n=1}^{m} \big(1-\varphi(\beta_n)\big)\Big) \\
&= h\Big(N(\kappa,T) - \sum_{n=1}^{m} \int_{\beta_n}^{1} \varphi'(\sigma)\mathrm{d}\sigma\Big) = h\Big(N(\kappa,T) - \int_{\kappa}^{1} \varphi'(\sigma) \sum_{\substack{n \\ \beta_n\le\sigma}} 1 \,\mathrm{d}\sigma\Big) \\
&= h\Big(N(\kappa,T) - \int_{\kappa}^{1} \varphi'(\sigma)\big(N(\kappa,T) - N(\sigma,T)\big)\mathrm{d}\sigma\Big) \\
&= h\Big(x^{\kappa-1}N(\kappa,T) + \log(x)\int_{\kappa}^{1} x^{\sigma-1}N(\sigma,T)\mathrm{d}\sigma\Big). \qquad (72)
\end{aligned}$$

The term

$$x^{\kappa-1}N(\kappa,T) \ll x^{\kappa-1}T\log T$$

vanishes assuming (71), so it remains only to investigate the integral in (72).

At this point, we need to apply heavier machinery to get further. It suffices to use (a weakened version of) Tchudakoff's zero-free region (21): Tchudakoff's result implies that $\zeta(s)$ has no zeros in a region of the form $\sigma \ge L(t)$, $t \ge t_0$, where

$$L(t) = 1 - \frac{A(t)\log\log t}{\log t},$$

and where $A(t)$ is an increasing function that goes to infinity with t.

Now assume that

$$N(\sigma,T) = O\big(T^{\alpha(1-\sigma)}\log^{\eta} T\big)$$

[20]Throughout we assume, of course, that there may exist zeros which are nontrivial and do not lie on the critical line.

uniformly in $\frac{1}{2} \leq \sigma \leq 1$ as $T \to \infty$, for some constants $\alpha \geq 2$ and $\eta \geq 1$. Then

$$\begin{aligned}\log(x)\int_\kappa^1 x^{\sigma-1}N(\sigma,T)\mathrm{d}\sigma &= \log(x)\int_\kappa^{L(T)} x^{\sigma-1}N(\sigma,T)\mathrm{d}\sigma \\ &\ll \log(x)\log^\eta(T)\int_\kappa^{L(T)}\left(\frac{x}{T^\alpha}\right)^{\sigma-1}\mathrm{d}\sigma \\ &= \frac{\log(x)\log^\eta(T)}{\log(x/T^\alpha)}\left(\left(\frac{x}{T^\alpha}\right)^{L(T)-1} - \left(\frac{x}{T^\alpha}\right)^{\kappa-1}\right).\end{aligned}$$

Take $T = x^\theta$, where $0 < \theta < \alpha^{-1} \leq \frac{1}{2}$.[21] Then $x/T^\alpha = x^{1-\theta\alpha} > 1$, and so

$$\begin{aligned}\log(x)\int_\kappa^1 x^{\sigma-1}N(\sigma,T)\mathrm{d}\sigma &\ll \frac{\log(x)\log^\eta(T)}{\log(x/T^\alpha)}\left(\left(\frac{x}{T^\alpha}\right)^{L(T)-1} - \left(\frac{x}{T^\alpha}\right)^{\kappa-1}\right) \\ &\leq \frac{\log^\eta(x)}{1-\theta\alpha}x^{(1-\theta\alpha)(L(T)-1)} \\ &\ll \log^\eta(x)\exp\left(-(1-\theta\alpha)\log(x)A(T)\frac{\log\log T}{\log T}\right) \\ &= \log^\eta(x)\exp\left(-(1-\theta\alpha)\log(x)A(x^\theta)\frac{\log(\theta)+\log\log x}{\theta\log x}\right) \\ &\ll \log^\eta(x)\exp\big(-(1-\theta\alpha)\theta^{-1}A(x^\theta)\log\log x\big) \\ &= (\log x)^{\eta-(1-\theta\alpha)\theta^{-1}A(x^\theta)}. \qquad (73)\end{aligned}$$

This vanishes as $x \to \infty$ provided that

$$(1-\theta\alpha)\theta^{-1}A(x^\theta) - \eta \qquad (74)$$

is positive and bounded away from zero for all sufficiently large x. It is sufficient (but not necessary) to assume that θ is constant with $0 < \theta < \alpha^{-1}$ for this to hold. But then we are close to a conclusion, since with our choice of T, we have

$$T\log T = \theta x^\theta \log x = o(x^{1-\kappa})$$

for any fixed $\kappa \in [\frac{1}{2}, 1-\theta) \subset [\frac{1}{2}, 1)$, so that (71) is satisfied. To get to the goal, we need only choose h such that (66) and (67) are true. But with $T = x^\theta$, this is equivalent to

$$x^{1-\theta}\log^2 x = o(h),$$

and together with the assumption that $h = o(x)$, it is clear that we can take $h = x^\vartheta$ for any fixed $1-\theta < \vartheta < 1$. Since θ may by taken arbitrarily close (but not equal) to α^{-1}, we arrive at the following conclusion, first proved by Hoheisel [17].[22]

Theorem 4.4 (Hoheisel, 1930). *If $N(\sigma,T) = O(T^{\alpha(1-\sigma)}\log^\eta T)$ uniformly for $\frac{1}{2} \leq \sigma \leq 1$ as $T \to \infty$, then*

$$\psi(x+x^\vartheta) - \psi(x) \sim x^\vartheta$$

as $x \to \infty$, for any fixed $1-\alpha^{-1} < \vartheta < 1$.

[21] We use the notation $T = x^\theta$ and $h = x^\vartheta$ repeatedly throughout the chapter.

[22] This is not word-for-word what Hoheisel proved. For example, the zero-free region $\sigma \geq L(t)$ was established by Tchudakoff after the publication of Hoheisel's article. Hoheisel worked with the specific values $\alpha = 4$ and $\eta = 6$, and a slightly weaker zero-density estimate: it is apparently usual to cite authors in the field not by the results that they proved, but by the immediate generalizations of their results, which were discovered later.

The best zero-density estimate one can hope for that is of the form required in Theorem 4.4 is the truth of the density hypothesis ($\alpha = 2$). In light of Propositions 4.1 and 4.2, this yields

Corollary 4.5. *If the strong density hypothesis is true, i.e., if $N(\sigma, T) = O(T^{2(1-\sigma)} \log^\eta T)$ uniformly for $\frac{1}{2} \le \sigma \le 1$ as $T \to \infty$, where $\eta \ge 1$, then*

$$\psi(x + x^\vartheta) - \psi(x) \sim x^\vartheta \qquad \textit{as } x \to \infty,$$
$$\pi(x + x^\vartheta) - \pi(x) \sim \frac{x^\vartheta}{\log x} \qquad \textit{as } x \to \infty,$$
$$p_{n+1} - p_n = O(p_n^\vartheta) \qquad \textit{as } n \to \infty,$$

all hold for any fixed $\frac{1}{2} < \vartheta < 1$.

Remark 4.6. Equation (59) and Corollary 4.5 can be combined in saying that

$$\pi(x + x^\vartheta) - \pi(x) \sim \min(1, \vartheta^{-1}) \frac{x^\vartheta}{\log x}$$

for any fixed $\vartheta > \frac{1}{2}$ as $x \to \infty$, provided that the strong density hypothesis is true.

Before we consider an unconditional result that follows from Hoheisel's theorem, we want emphasise that (74) can be made to vanish also if one allows θ to depend on x and approach α^{-1} from below as $x \to \infty$, although the improvement of the conclusion is minimal. For example, in agreement with (21), the choice $A(t) = \log\log t$ is admissible. Thus, if we write $\theta(x) = \alpha^{-1} - \Delta(x)$ for $0 < \Delta(x) = o(1)$, then it is sufficient that there exists a constant $C > 0$, such that $x \ge x_0$ implies

$$\begin{aligned}
(1 - \theta\alpha)\theta^{-1} A(x^\theta) - \eta &\ge C \\
(1 - \theta\alpha) \log(\theta \log x) - (C + \eta)\theta &\ge 0 \\
\Delta(x) \log\log x &\ge (C + \eta)\alpha^{-1} - (C + \eta)\Delta(x) - \Delta(x) \log\big(\alpha^{-1} - \Delta(x)\big) \\
\Delta(x) &\ge \frac{(C + \eta)\alpha^{-1} + o(1)}{\log\log x}.
\end{aligned}$$

So, for any constant $D > (C + \eta)\alpha^{-1}$, we can in fact take

$$T = x^\theta = x^{\alpha^{-1} - D(\log\log x)^{-1}} = x^{\alpha^{-1}} \exp\Big(- D \frac{\log x}{\log\log x} \Big).$$

The resulting requirement on h obtained from (67) is

$$x^{1-\alpha^{-1}} \log^2(x) \exp\Big(D \frac{\log x}{\log\log x} \Big) = o(h),$$

and the factor $\log^2(x) \exp(\ldots)$ on the left grows slower than any power of x, but faster than any power of $\log x$.

It is known that

$$N(\sigma, T) = O(T^{\frac{12}{5}(1-\sigma)} \log^9 T) \tag{75}$$

uniformly for $\frac{1}{2} \le \sigma \le 1$ as $T \to \infty$, and this may be the best known zero-density estimate of this form over the range $\frac{1}{2} \le \sigma \le 1$.[23] Using this in Theorem 4.4 yields unconditionally

$$\psi(x + x^{\frac{7}{12}+\varepsilon}) - \psi(x) \sim x^{\frac{7}{12}+\varepsilon},$$

[23] See [21], Equation (11.29), p. 275.

for any fixed $\varepsilon > 0$ as $x \to \infty$.

Corollary 4.5 shows how sharp a result we get under the strong density hypothesis, but we have not yet commented on what is obtainable using merely the density hypothesis. For this we can employ almost exactly the same arguments used in deriving Hoheisel's theorem, except with one caveat. The density hypothesis on the form $N(\sigma, T) = O(T^{2(1-\sigma)+\varepsilon})$ has the defect that the exponent of T approaches ε (> 0) as $\sigma \to 1$ from the left, although there are no zeros to the right of $\sigma = 1$. As a consequence, we would end up with an extra factor $T^\varepsilon = x^{\varepsilon\theta}$ in Equation (73), resulting in an unacceptable conclusion.

To remedy this defect, we can use the fact that

$$N(\sigma, T) = O\big(T^{\frac{35}{36}(1-\sigma)} \log^{16} T\big) \tag{76}$$

uniformly for $\frac{152}{155} \le \sigma \le 1$ as $T \to \infty$.[24] Assuming $\kappa < \frac{152}{155}$, we reconsider the integral from (72) by writing

$$\log(x) \int_\kappa^1 x^{\sigma-1} N(\sigma, T) \mathrm{d}\sigma = \log(x) \int_\kappa^{\frac{152}{155}} x^{\sigma-1} N(\sigma, T) \mathrm{d}\sigma + \log(x) \int_{\frac{152}{155}}^1 x^{\sigma-1} N(\sigma, T) \mathrm{d}\sigma.$$

Using the density hypothesis, we have first

$$\begin{aligned} \log(x) \int_\kappa^{\frac{152}{155}} x^{\sigma-1} N(\sigma, T) \mathrm{d}\sigma &\ll \log(x) T^\varepsilon \int_\kappa^{\frac{152}{155}} \Big(\frac{x}{T^2}\Big)^{\sigma-1} \mathrm{d}\sigma = \frac{T^\varepsilon \log(x)}{\log(x/T^2)} \bigg(\Big(\frac{x}{T^2}\Big)^{-\frac{3}{155}} - \Big(\frac{x}{T^2}\Big)^{\kappa-1}\bigg) \\ &\ll \frac{T^\varepsilon \log(x)}{\log(x/T^2)} \Big(\frac{x}{T^2}\Big)^{-\frac{3}{155}} = \frac{1}{1-2\theta} x^{\varepsilon\theta - (1-2\theta)\frac{3}{155}}, \end{aligned}$$

which vanishes as $x \to \infty$ if $\theta < \frac{1}{2}(1 + \frac{155}{6}\varepsilon)^{-1} < \frac{1}{2}$. To the right, (76) yields

$$\begin{aligned} \log(x) \int_{\frac{152}{155}}^1 x^{\sigma-1} N(\sigma, T) \mathrm{d}\sigma &\ll \log^{17}(x) \int_{\frac{152}{155}}^{L(T)} \Big(\frac{x}{T^{\frac{35}{36}}}\Big)^{\sigma-1} \mathrm{d}\sigma = \frac{\log^{17}(x)}{\log\Big(\frac{x}{T^{\frac{35}{36}}}\Big)} \bigg(\Big(\frac{x}{T^{\frac{35}{36}}}\Big)^{L(T)-1} - \Big(\frac{x}{T^{\frac{35}{36}}}\Big)^{-\frac{3}{155}}\bigg) \\ &\ll \frac{\log^{17}(x)}{\log\Big(\frac{x}{T^{\frac{35}{36}}}\Big)} \Big(\frac{x}{T^{\frac{35}{36}}}\Big)^{L(T)-1} = \frac{\log^{16}(x)}{1 - \frac{35}{36}\theta} x^{-(1-\frac{35}{36}\theta) A(x^\theta) \log\log(x^\theta) (\log x^\theta)^{-1}} \\ &\ll (\log x)^{16 - (1-\frac{35}{36}\theta)\theta^{-1} A(x^\theta) \log\log x}, \end{aligned}$$

which vanishes automatically because $A(x^\theta) \to \infty$ and $\theta < \frac{1}{2} < \frac{36}{35}$. Since $\varepsilon > 0$ may be chosen arbitrarily small, and since the assumptions are otherwise the same as in the derivation of Hoheisel's theorem, we reach the following conclusion.

Proposition 4.7. *Corollary 4.5 remains true if 'strong density hypothesis' is replaced by 'density hypothesis'. In particular, Corollary 2.16 then implies that Corollary 4.5 holds if the Lindelöf hypothesis is true.*

There are two comments to be made here. First, the density hypothesis in its natural form is really the assertion that, for any $\varepsilon > 0$,

$$N(\sigma, T) = O\big(T^{(2+\varepsilon)(1-\sigma)}\big)$$

uniformly for $\frac{1}{2} \le \sigma \le 1$ as $T \to \infty$. Indeed, the zero-density estimate (76) shows that this is equivalent to density hypothesis 2.11, except that value of $N(\sigma, T)$ near $\sigma = 1$ has been corrected.

[24] See [21], Theorem 11.3, p. 277.

Second, it is tempting to try to use a zero-density estimate[25] of Halász and Turán [15] over the interval $\sigma \in [\frac{3}{4}+\delta, \frac{152}{155}]$ when estimating the integral in (72), in an attempt to improve the conclusion under the Lindelöf hypothesis. However, the best known bound for $N(\sigma, T)$ over this interval, under the Lindelöf hypothesis, is still only that which is given by the density hypothesis. But the resulting estimate

$$\log(x)\int_{\frac{1}{2}}^{\frac{3}{4}+\delta} x^{\sigma-1}N(\sigma,T)\mathrm{d}\sigma \ll x^{\theta\varepsilon-(\frac{1}{4}-\delta)(1-2\theta)}$$

requires that we take $\theta < \frac{1}{2}$, so no stronger result is attained under this assumption.

It is also possible to reach interesting conclusions if we merely assume that the strong density hypothesis has been established in the rightmost part of critical strip, as is the case for the estimate (76). An example of such a result is given by Ivić.[26]

Theorem 4.8 (Ivić, 1979). *Suppose $N(\sigma,T) = O(T^{\alpha(\sigma)(1-\sigma)}\log^{\eta} T)$ uniformly for $\frac{1}{2} \le \sigma \le 1$ as $T \to \infty$, where $\eta \ge 1$, and where $\alpha(\sigma)(1-\sigma) \le 1$ for all σ. Suppose further that there exist constants $\frac{1}{2} < \kappa < 1$ and $\alpha_0 > 2$, such that*

1. *$\alpha(\sigma) \le \alpha_0$ for $\frac{1}{2} \le \sigma \le \kappa$,*
2. *$\alpha(\sigma) \le 2$ for $\kappa \le \sigma \le 1$.*

Then

$$\psi(x+h) - \psi(x) \sim h$$

as $x \to \infty$, provided

$$h(x) \ge x^{1-\alpha_0^{-1}}\log^b x,$$

where b is any fixed number satisfying

$$b > \frac{\eta+2}{\alpha_0(1-\kappa)}.$$

Proof. To prove this, we need to reconsider the sum in (65). As before, we split the sum at $\sigma = \kappa$,

$$\sum_{|\gamma|\le T} C(\rho) = \sum_{\substack{|\gamma|\le T\\ \beta<\kappa}} C(\rho) + \sum_{\substack{|\gamma|\le T\\ \beta\ge\kappa}} C(\rho), \tag{77}$$

but we now assume that $T > x/h$, as we may in light of (67). Let us begin by considering the sum over the zeros ρ whose real parts are at least κ . Since we have two bounds for $C(\rho)$, namely $C(\rho) \ll hx^{\beta-1}$ and $C(\rho) \ll x^{\beta}/|\gamma|$, it is convenient to further split these zeros into two classes, according to if $|\gamma| \le x/h$ or $|\gamma| > x/h$. This allows us to use the best bound in each case, as follows.

$$\begin{aligned}\sum_{\substack{|\gamma|\le T\\ \beta\ge\kappa}} C(\rho) &\ll \frac{h}{x}\sum_{\substack{|\gamma|\le x/h\\ \beta\ge\kappa}} x^{\beta} + \sum_{\substack{x/h<|\gamma|\le T\\ \beta\ge\kappa}} \frac{x^{\beta}}{|\gamma|}\\ &= \frac{h}{x}\sum_{\substack{|\gamma|\le x/h\\ \beta\ge\kappa}} (x^{\beta}-x^{\kappa}) + \frac{hx^{\kappa}}{x}\sum_{\substack{|\gamma|\le x/h\\ \beta\ge\kappa}} 1 + \sum_{\substack{x/h<|\gamma|\le T\\ \beta\ge\kappa}} x^{\beta}\Big(\frac{1}{|\gamma|}-\frac{1}{T}\Big) + \frac{1}{T}\sum_{\substack{x/h<|\gamma|\le T\\ \beta\ge\kappa}} x^{\beta}.\end{aligned} \tag{78}$$

[25] For any $\varepsilon, \delta > 0$, we have $N(\sigma,T) = O(T^{\varepsilon})$ uniformly for $\frac{3}{4}+\delta \le \sigma \le 1$ as $T \to \infty$, assuming the Lindelöf hypothesis.

[26] See [21], Theorem 12.8, p. 316, or [20]. Beware that there is a typo in lines 12 and 13 on p. 317 of [21], where two instances of 'T' should be replaced by 't'.

In the second line we have 'recentred' the summation, since this allows for easy bounding by comparing with appropriate integrals. We now bound these four terms, starting with the second and the fourth.[27]

$$\frac{hx^\kappa}{x}\sum_{\substack{|\gamma|\le x/h\\ \beta\ge\kappa}} 1 = 2x^\kappa N(\kappa,\tfrac{x}{h})(\tfrac{x}{h})^{-1} \ll \max_{\kappa\le\sigma\le 1}\big(x^\sigma \max_{x/h\le t\le T}(N(\sigma,t)t^{-1})\big). \tag{79}$$

In a similar manner, the computation (72) gives us

$$\begin{aligned}\frac{1}{T}\sum_{\substack{x/h<|\gamma|\le T\\ \beta\ge\kappa}} x^\beta \le \frac{x}{T}\sum_{\substack{|\gamma|\le T\\ \beta\ge\kappa}} x^{\beta-1} &= 2x^\kappa N(\kappa,T)T^{-1} + 2\log(x)\int_\kappa^1 x^\sigma N(\sigma,T)T^{-1}\mathrm{d}\sigma\\ &\ll \log(x)\max_{\kappa\le\sigma\le 1}\big(x^\sigma \max_{x/h\le t\le T}(N(\sigma,t)t^{-1})\big).\end{aligned} \tag{80}$$

To handle the first and third term of (78), let 1_S denote the indicator function for the set S. Then we have, for $0\le a<b$,

$$\begin{aligned}\sum_{\substack{a<|\gamma|\le b\\ \beta\ge\kappa}}(x^\beta - x^\kappa) &= \log(x)\sum_{\substack{a<|\gamma|\le b\\ \beta\ge\kappa}}\int_\kappa^\beta x^\sigma\mathrm{d}\sigma = \log(x)\sum_{\substack{a<|\gamma|\le b\\ \beta\ge\kappa}}\int_\kappa^1 x^\sigma 1_{[\kappa,\beta]}(\sigma)\mathrm{d}\sigma\\ &= \log(x)\int_\kappa^1 x^\sigma \sum_{\substack{a<|\gamma|\le b\\ \beta\ge\kappa}} 1_{[\kappa,\beta]}(\sigma)\mathrm{d}\sigma = \log(x)\int_\kappa^1 x^\sigma\sum_{\substack{a<|\gamma|\le b\\ \beta\ge\sigma}} 1\,\mathrm{d}\sigma\\ &= 2\log(x)\int_\kappa^1 x^\sigma N(\sigma,a,b)\mathrm{d}\sigma,\end{aligned} \tag{81}$$

where $N(\sigma,a,b)$ denotes $N(\sigma,b)-N(\sigma,a)$. Using this with $a=0$ and $b=x/h$ gives us the first term of (78) as

$$\begin{aligned}\frac{h}{x}\sum_{\substack{|\gamma|\le x/h\\ \beta\ge\kappa}}(x^\beta - x^\kappa) = \frac{2\log(x)h}{x}\int_\kappa^1 x^\sigma N(\sigma,\tfrac{x}{h})\mathrm{d}\sigma &= 2\log(x)\int_\kappa^1 x^\sigma N(\sigma,\tfrac{x}{h})(\tfrac{x}{h})^{-1}\mathrm{d}\sigma\\ &\ll \log(x)\max_{\kappa\le\sigma\le 1}\big(x^\sigma\max_{x/h\le t\le T}(N(\sigma,t)t^{-1})\big).\end{aligned} \tag{82}$$

To bound the third term of (78), we again compare with an integral.

$$\begin{aligned}\sum_{\substack{x/h<|\gamma|\le T\\ \beta\ge\kappa}} x^\beta\Big(\frac{1}{|\gamma|}-\frac{1}{T}\Big) &= \sum_{\substack{x/h<|\gamma|\le T\\ \beta\ge\kappa}} x^\beta\int_{|\gamma|}^T\frac{\mathrm{d}t}{t^2} = \sum_{\substack{x/h<|\gamma|\le T\\ \beta\ge\kappa}} x^\beta\int_{x/h}^T\frac{1_{[|\gamma|,T]}(t)}{t^2}\mathrm{d}t\\ &= \int_{x/h}^T\frac{1}{t^2}\sum_{\substack{x/h<|\gamma|\le T\\ \beta\ge\kappa}} x^\beta 1_{[|\gamma|,T]}(t)\mathrm{d}t = \int_{x/h}^T\frac{1}{t^2}\sum_{\substack{x/h<|\gamma|\le t\\ \beta\ge\kappa}} x^\beta\mathrm{d}t\\ &= \int_{x/h}^T\frac{1}{t^2}\sum_{\substack{x/h<|\gamma|\le t\\ \beta\ge\kappa}}(x^\beta-x^\kappa)\mathrm{d}t + \int_{x/h}^T\frac{1}{t^2}\sum_{\substack{x/h<|\gamma|\le t\\ \beta\ge\kappa}} x^\kappa\mathrm{d}t\\ &= 2\log(x)\int_{x/h}^T\frac{1}{t^2}\int_\kappa^1 x^\sigma N(\sigma,\tfrac{x}{h},t)\mathrm{d}\sigma\mathrm{d}t + 2\int_{x/h}^T\frac{x^\kappa}{t^2}N(\kappa,\tfrac{x}{h},t)\mathrm{d}t,\end{aligned} \tag{83}$$

[27]The double maximum exists because $N(\sigma,t)t^{-1}$ for σ fixed is piecewise nonincreasing and upper semicontinuous function of t, and so is $x^\sigma \max_{x/h\le t\le T}(N(\sigma,t)t^{-1})$ as a function of $1-\sigma$.

where we used (81) with $a = x/h$ and $b = t$ in the last line. Here,

$$\begin{aligned}
\int_{x/h}^{T} \frac{1}{t^2} \int_{\kappa}^{1} x^{\sigma} N(\sigma, \tfrac{x}{h}, t) \mathrm{d}\sigma \mathrm{d}t &= \int_{x/h}^{T} \frac{1}{t^2} \int_{\kappa}^{1} x^{\sigma} \big(N(\sigma, t) - N(\sigma, \tfrac{x}{h})\big) \mathrm{d}\sigma \mathrm{d}t \\
&= \int_{x/h}^{T} \frac{1}{t} \int_{\kappa}^{1} x^{\sigma} \frac{N(\sigma, t)}{t} \mathrm{d}\sigma \mathrm{d}t - \int_{x/h}^{T} \frac{1}{t^2} \int_{\kappa}^{1} x^{\sigma} N(\sigma, \tfrac{x}{h}) \mathrm{d}\sigma \mathrm{d}t \\
&\leq \big(\log(T) - \log\big(\tfrac{x}{h}\big)\big) \max_{\kappa \leq \sigma \leq 1} \big(x^{\sigma} \max_{x/h \leq t \leq T} (N(\sigma, t) t^{-1})\big) - \binom{\text{nonnegative}}{\text{quantity}} \\
&\ll \log(x) \max_{\kappa \leq \sigma \leq 1} \big(x^{\sigma} \max_{x/h \leq t \leq T} (N(\sigma, t) t^{-1})\big), \qquad (84)
\end{aligned}$$

since $x/h < T \leq x$. Moreover,

$$\begin{aligned}
\int_{x/h}^{T} \frac{x^{\kappa}}{t^2} N(\kappa, \tfrac{x}{h}, t) \mathrm{d}t &= \int_{x/h}^{T} \frac{x^{\kappa}}{t^2} \big(N(\kappa, t) - N(\kappa, \tfrac{x}{h})\big) \mathrm{d}t \\
&= \int_{x/h}^{T} \frac{1}{t} x^{\kappa} N(\kappa, t) t^{-1} \mathrm{d}t - x^{\kappa} N(\kappa, \tfrac{x}{h}) \int_{x/h}^{T} \frac{\mathrm{d}t}{t^2} \\
&\leq \big(\log(T) - \log\big(\tfrac{x}{h}\big)\big) x^{\kappa} \max_{x/h \leq t \leq T} (N(\kappa, t) t^{-1}) - \binom{\text{nonnegative}}{\text{quantity}} \\
&\ll \log(x) \max_{\kappa \leq \sigma \leq 1} \big(x^{\sigma} \max_{x/h \leq t \leq T} (N(\sigma, t) t^{-1})\big). \qquad (85)
\end{aligned}$$

Equations (78) through (85) now give us the collective bound

$$\sum_{\substack{|\gamma| \leq T \\ \beta \geq \kappa}} C(\rho) \ll \log^2(x) \max_{\kappa \leq \sigma \leq 1} \big(x^{\sigma} \max_{x/h \leq t \leq T} (N(\sigma, t) t^{-1})\big), \qquad (86)$$

where the $\log^2(x)$ comes from equations (83) and (84).

For the term of (77) taken over the roots with $\beta < \kappa$, we can consider the roots to the left or right of the critical line separately, so that

$$\sum_{\substack{|\gamma| \leq T \\ \beta < \kappa}} C(\rho) = \sum_{\substack{|\gamma| \leq T \\ \beta \leq 1/2}} C(\rho) + \sum_{\substack{|\gamma| \leq T \\ 1/2 < \beta < \kappa}} C(\rho). \qquad (87)$$

Using (69), we trivially have

$$\sum_{\substack{|\gamma| \leq T \\ \beta \leq 1/2}} C(\rho) \ll \frac{h}{\sqrt{x}} N(T) \ll \frac{hT \log T}{\sqrt{x}} \ll \frac{hT \log x}{\sqrt{x}} = o(h) \qquad (88)$$

since $T \leq x$, and provided that

$$T = o(\sqrt{x} (\log x)^{-1}). \qquad (89)$$

And, in the same way that (86) was deduced, we get

$$\sum_{\substack{|\gamma| \leq T \\ 1/2 < \beta < \kappa}} C(\rho) \ll \log^2(x) \max_{1/2 \leq \sigma \leq \kappa} \big(x^{\sigma} \max_{x/h \leq t \leq T} (N(\sigma, t) t^{-1})\big). \qquad (90)$$

It is time to concretise the bounds obtained so far, and for this we use the Koborov-Vinogradov zero-free region. Let therefore

$$L(t) = 1 - A(\log t)^{-2/3} (\log \log t)^{-1/3}$$

denote the function to the right in (22), and put

$$\begin{aligned} h &= x^a \log^b(x) \quad \text{and} \quad T = x^c \log^d(x), \ \text{ so} \\ x/h &= x^{1-a} \log^{-b}(x) \quad \text{and} \quad hT = x^{a+c} \log^{b+d}(x), \end{aligned}$$

for constants a, b, c, d, that should be chosen such that h is minimised. The restrictions imposed on h and T as $x \to \infty$ give us the additional information that

(i) $h = o(x) \implies a < 1 \vee (a = 1 \wedge b < 0)$,

(ii) $\lim_{x\to\infty} h(x) = \infty \implies a > 0 \vee (a = 0 \wedge b > 0)$,

(iii) $T \leq x \implies c < 1 \vee (c = 1 \wedge d \leq 0)$,

(iv) $T \geq 1 \implies c > 0 \vee (c = 0 \wedge d \geq 0)$,

(v) $T = o(\sqrt{x}(\log x)^{-1}) \implies c < \frac{1}{2} \vee (c = \frac{1}{2} \wedge d < -1)$,

(vi) $x \log^2 x = o(hT) \implies a + c > 1 \vee (a + c = 1 \wedge b + d > 2)$.

The first two give $0 \leq a \leq 1$, but the case $a = 1$ can be ignored, since it gives a much larger h than we desire. We can also ignore $a = 0$, since by Maier's theorem this choice of h is too small. Thus, $0 < a < 1$. The fourth and fifth give $0 \leq c \leq \frac{1}{2}$, but we will see later that we need in fact $c < \frac{1}{2}$. The sixth gives $a + c \geq 1$, or in other words $a \geq 1 - c$. Since we want to minimize h, we should take a as small as possible. For a given c, this achieved by choosing $a = 1 - c$, and we better choose $c > 0$. But then $a + c = 1$, so the sixth implication forces $b + d > 2$. We can handle this by demanding $b + d = 3$, i.e., $d = 3 - b$, and assume that $b > 0$. In summary, we have

$$h = x^{1-c} \log^b(x), \ \ T = x^c \log^{3-b}(x), \ \ x/h = x^c \log^{-b}(x)$$

for $0 < c \leq \frac{1}{2}$ and $b > 0$.

Equation (86) together with the assumptions on $N(\sigma, T)$ for $\sigma \geq \kappa$, yield

$$\begin{aligned} \sum_{\substack{|\gamma| \leq T \\ \beta \geq \kappa}} C(\rho) &\ll \log^2(x) \max_{\kappa \leq \sigma \leq 1} \left(x^\sigma \max_{x/h \leq t \leq T} (N(\sigma, t) t^{-1}) \right) \\ &\ll \log^2(x) \max_{\kappa \leq \sigma \leq L(x)} \left(x^\sigma \max_{x/h \leq t \leq T} (t^{\alpha(\sigma)(1-\sigma)-1} \log^\eta t) \right) \\ &\leq (\log x)^{2+\eta} \max_{\kappa \leq \sigma \leq L(x)} \left(x^\sigma \max_{x/h \leq t \leq T} (t^{\alpha(\sigma)(1-\sigma)-1}) \right) \\ &= (\log x)^{2+\eta} \max_{\kappa \leq \sigma \leq L(x)} (x^\sigma (x/h)^{\alpha(\sigma)(1-\sigma)-1}) \\ &= h(\log x)^{2+\eta} \max_{\kappa \leq \sigma \leq L(x)} \left(x \cdot (x/h)^{-\alpha(\sigma)} \right)^{\sigma - 1} \\ &= h(\log x)^{2+\eta} \max_{\kappa \leq \sigma \leq L(x)} \left(x \cdot (x^c \log^{-b}(x))^{-\alpha(\sigma)} \right)^{\sigma-1} \\ &\leq h(\log x)^{2+\eta} \max_{\kappa \leq \sigma \leq L(x)} \left(x \cdot (x^c \log^{-b}(x))^{-2} \right)^{\sigma-1} \\ &= h(\log x)^{2+\eta} \max_{\kappa \leq \sigma \leq L(x)} \left(x^{(1-2c)(\sigma-1)} (\log x)^{2b(\sigma-1)} \right) \\ &\leq h(\log x)^{2+\eta} \max_{\kappa \leq \sigma \leq L(x)} x^{(1-2c)(\sigma-1)} \\ &= h(\log x)^{2+\eta} \exp\Big(-A(1-2c)(\log x)^{1/3} (\log\log x)^{-1/3} \Big) \\ &= o(h) \end{aligned}$$

as $x \to \infty$. In the third line we used $t \leq x$ to pull out the logarithm, and in the fourth line we used that $\alpha(\sigma)(1-\sigma) \leq 1$, so that the inner maximum is attained at $t = x/h$. The seventh line used that

$\alpha(\sigma) \le 2$ for $\kappa \le \sigma \le 1$, together with the fact that the outer exponent $\sigma - 1$ is negative. In the ninth line we used that the exponent $2b(\sigma - 1)$ of the logarithm is negative. Finally, in the tenth line we assumed that $c < \frac{1}{2}$, so that $(1-2c)(\sigma-1)$ is an increasing function of σ, and the maximum happens at $\sigma = L(x)$.

In a similar manner, Equations (87) through (90) yield

$$\begin{aligned}\sum_{\substack{|\gamma|\le T\\ \beta<\kappa}} C(\rho) &\ll hx^{-1/2}T\log(x) + \log^2(x) \max_{1/2\le\sigma\le\kappa}\left(x^{\sigma}\max_{x/h\le t\le T}(N(\sigma,t)t^{-1})\right)\\
&\ll o(h) + (\log x)^{2+\eta}\max_{1/2\le\sigma\le\kappa}\left(x^{\sigma}(x/h)^{\alpha(\sigma)(1-\sigma)-1}\right)\\
&= o(h) + (\log x)^{2+\eta}\max_{1/2\le\sigma\le\kappa}\left(x^{\sigma+c\alpha(\sigma)(1-\sigma)-c}(\log x)^{-b\alpha(\sigma)(1-\sigma)+b}\right)\\
&= o(h) + x^{-c}(\log x)^{2+\eta+b}\max_{1/2\le\sigma\le\kappa}\left(x^{\sigma+c\alpha(\sigma)(1-\sigma)}(\log x)^{-b\alpha(\sigma)(1-\sigma)}\right)\\
&= o(h) + x^{-c}(\log x)^{2+\eta+b}\max_{1/2\le\sigma\le\kappa}\left(x^{\sigma}(x^{c}\log^{-b}(x))^{\alpha(\sigma)(1-\sigma)}\right)\\
&\le o(h) + x^{-c}(\log x)^{2+\eta+b}\max_{1/2\le\sigma\le\kappa}\left(x^{\sigma}(x^{c}\log^{-b}(x))^{\alpha_0(1-\sigma)}\right),\end{aligned}$$

where the last line used that $\alpha(\sigma) \le \alpha_0$ and $1-\sigma \ge 0$ for $\frac{1}{2} \le \sigma \le \kappa$. If we suppose that $c \le \alpha_0^{-1}$, then this last maximum must occur at $\sigma = \kappa$, because then the combined exponent $\sigma + c\alpha_0(1-\sigma)$ of x is nondecreasing, and the combined exponent $-b\alpha_0(1-\sigma)$ of $\log x$ is increasing. Therefore,

$$\begin{aligned}\sum_{\substack{|\gamma|\le T\\ \beta<\kappa}} C(\rho) &\ll o(h) + x^{-c}(\log x)^{2+\eta+b}x^{\kappa}(x^{c}\log^{-b}(x))^{\alpha_0(1-\kappa)}\\
&= o(h) + hx^{\kappa+c\alpha_0(1-\kappa)-1}(\log x)^{2+\eta-b\alpha_0(1-\kappa)}.\end{aligned}$$

The exponent $\kappa + c\alpha_0(1-\kappa) - 1$ is nonpositive when $c \le \alpha_0^{-1}$, so that the power of x is bounded. To make the entire expression $o(h)$ it therefore suffices to choose b such that the exponentiated logarithm vanishes. That is, such that

$$b > \frac{2+\eta}{\alpha_0(1-\kappa)}.$$

This completes the proof of Ivić's theorem. □

As an application, it is known by the works of M. Julita [23] that, for any $\varepsilon > 0$, we have $N(\sigma, T) = O\big(T^{\alpha(\sigma)(1-\sigma)}\log^{\eta} T\big)$ uniformly for $\frac{11}{14} + \varepsilon \le \sigma \le 1$, where $\eta \ge 1$ and $\alpha(\sigma) \le 2$.[28] If we use this together with the estimate (75) in Theorem 4.8 (with $\alpha_0 = \frac{12}{5}$, $\eta = 9$, $\kappa = \frac{11}{14} + \varepsilon$, $0 < \varepsilon < 10^{-3}$), then we obtain

$$\psi(x+h) - \psi(x) \sim h$$

for $h \ge x^{\frac{7}{12}}\log^{22} x$ as $x \to \infty$.

At this point it is clear that the methods we have employed so far are close to their breaking point, with a fundamental barrier for h in the vicinity of $\sqrt{x}$. We may therefore ask how strong a result that one can possibly hope for. This is discussed in the next section.

[28] See also [21], §11.7.

5 Primes in short intervals: the Riemann hypothesis

In the present section, we are going to see what information we can squeeze out about primes in short intervals by assuming the Riemann hypothesis. This will ultimately require a reconsideration of the explicit formula for $\psi(x)$, in which the error term (introduced by replacing $\sum_\rho$ with $\sum_{|\gamma|\le T}$) is reduced by a logarithmic factor.

As a point of reference, the estimates further down may be compared to the classical results due to Littlewood and von Koch:

Theorem 5.1. *On the one hand, we have*

$$\psi(x) - x = O(x^\Theta \log^2 x)$$

as $x \to \infty$, *where* Θ *denotes the supremum of the real parts of the nontrivial zeros of* $\zeta(s)$. *On the other hand,*

$$\psi(x) - x = \Omega_\pm(\sqrt{x}\log\log\log x)$$

as $x \to \infty$. *Here,* $f(x) = \Omega_\pm(g(x))$ *means that there exists a constant* $C > 0$ *and sequences* $(x_n)_{n=1}^\infty$ *and* $(x'_n)_{n=1}^\infty$ *tending to infinity, such that*

$$f(x_n) > Cg(x_n) \quad \text{and} \quad f(x'_n) < -Cg(x'_n)$$

for all n.

5.1 Prime number theorem under the Riemann hypothesis

A convenience of Theorem 5.1, is that it gives at once

Corollary 5.2. *If the Riemann hypothesis is true and* $\sqrt{x}\log^2 x = o(h)$, *then* $\psi(x+h) - \psi(x) \sim h$ *as* $x \to \infty$.

Proof. If the Riemann hypothesis is true, then we have $\psi(x) - x = O(\sqrt{x}\log^2 x)$ by Theorem 5.1, so

$$\psi(x+h) - \psi(x) - h \ll \sqrt{x+h}\log^2(x+h) \ll \max(\sqrt{x}\log^2 x, \sqrt{h}\log^2 h) = o(h).$$

□

We are next going to show that the conclusion of Corollary 5.2 can be improved by a logarithmic factor of x. This was first made possible by the work of Cramér [5, 6], although he did not state this result explicitly. Cramér's approach was to investigate the function $V(z) = \sum_{\gamma>0} \mathrm{e}^{\rho z}$, for z in the upper half-plane, and, using estimates for this function, he further derived several estimates pertaining to $\psi(x) - x$, as well as Cramér's theorem (discussed in the next section). For example, he showed that

$$\psi(x) = x - \sum_\rho \frac{x^\rho}{\rho}\mathrm{e}^{-|\gamma|/x^3} + O(\log^2 x)$$

as $x \to \infty$, where the convergence of the sum on the right has been accelerated using so-called convegence factors.

This is as much as we will dwell on Cramér's results, and that is simply because shorter proofs of these results have been discovered later. Our investigation follows Goldston's article [13]. In short, Goldston's contribution was to enlarge the range of T's for which the estimate

$$\psi(x) = x - \sum_{|\gamma|\le T} \frac{x^\rho}{\rho} + O(\sqrt{x}\log x)$$

holds, from $T \ge \sqrt{x}$ to $T \ge \sqrt{x}(\log x)^{-1}$. This will be our next goal as well, and we begin with a lemma due to Littlewood.

Lemma 5.3 (Littlewood). *If z and w are complex numbers with $|z| \le \frac{1}{2}$ and $|zw| \le 2$, then*

$$\big|(1+z)^w - 1 - wz\big| \le \frac{13}{5}|w|(|w|+1)|z|^2.$$

Proof. For simplicity of notation, let $r = |z|$ and $\mu = |w|$. The claim is trivially true if $zw = 0$, so suppose that $r > 0$ and $\mu > 0$. By the generalized binomial theorem, we have

$$\begin{aligned}|(1+z)^w - 1 - wz| &= \Big|\sum_{n=2}^{\infty} \frac{w(w-1)\cdots(w-n+1)}{n!} z^n\Big| \le \sum_{n=2}^{\infty} \frac{\mu(\mu+1)\cdots(\mu+n-1)}{n!} r^n \\ &= (1-r)^{-\mu} - 1 - r\mu,\end{aligned}$$

so that

$$\frac{|(1+z)^w - 1 - wz|}{|w|(|w|+1)|z|^2} \le \sum_{n=2}^{\infty} \frac{\mu(\mu+1)\cdots(\mu+n-1)}{\mu(\mu+1)n!} r^{n-2} = \frac{(1-r)^{-\mu} - 1 - r\mu}{\mu(\mu+1)r^2}.$$

If r is fixed, then this upper bound is an increasing function of μ, since the coefficients in front of the r^{n-2} in the second sum are increasing functions of μ if $n \ge 3$ (and constant if $n = 2$). Its maximal value is therefore attained when $\mu = 2/r$ (where $|zw| = 2$), implying

$$\frac{|(1+z)^w - 1 - wz|}{|w|(|w|+1)|z|^2} \le \frac{(1-r)^{-2/r} - 3}{2(r+2)}.$$

We claim that the function on the right increases strictly with r for $0 < r < 1/2$ (this is in fact true for all $r < 1$, if we take care of the removable singularity at $r = -2$). Indeed, its derivative is

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}r}\left[\frac{(1-r)^{-2/r} - 3}{2(r+2)}\right] &= \frac{2(r+2)(1-r)^{-2/r}\big(2r^{-2}\log(1-r) + 2r^{-1}(1-r)^{-1}\big) - 2\big((1-r)^{-2/r} - 3\big)}{2^2(r+2)^2} \\ &= \frac{(r+2)\big(2r^{-2}\log(1-r) + 2r^{-1}(1-r)^{-1}\big) - 1}{2(r+2)^2(1-r)^{2/r}} + \frac{3}{2(r+2)^2} \\ &= \frac{(r+2)(1 + (2-\frac{2}{3})r + (2-\frac{2}{4})r^2 + (2-\frac{2}{5})r^3 + \cdots) - 1}{2(r+2)^2(1-r)^{2/r}} + \frac{3}{2(r+2)^2} \\ &> 0,\end{aligned}$$

where we used Taylor expansions $\log(1-r) = -\sum_{n=1}^{\infty} \frac{r^n}{n}$ and $(1-r)^{-1} = \sum_{n=0}^{\infty} r^n$ in the second-to-last line. Thus, the maximum in our case must be attained where $r = |z| = 1/2$, finally yielding

$$\frac{|(1+z)^w - 1 - wz|}{|w|(|w|+1)|z|^2} \le \frac{(1-\frac{1}{2})^{-\frac{2}{1/2}} - 3}{2(\frac{1}{2}+2)} = \frac{13}{5},$$

as desired. ∎

We are now ready to derive a stronger version of the explicit formula for $\psi(x)$.

Theorem 5.4 (Goldston). *If the Riemann hypothesis is true, then there exists a constant $T_0 > 1$, such that*

$$\Big|\psi(x) - x + \sum_{|\gamma| \le T} \frac{x^\rho}{\rho}\Big| < \frac{x}{2T} + 2\sqrt{x}\log T \tag{91}$$

whenever $x \ge 3$ and $T \ge T_0$.

Proof. Let $\psi_1(x)$ denote the function from Proposition 3.8. Since $\psi(x)$ is a nondecreasing function, we have

$$\frac{\psi_1(x-\ell)-\psi_1(x)}{-\ell}=\frac{1}{\ell}\int_{x-\ell}^{x}\psi(t)\mathrm{d}t\le\psi(x)\le\frac{1}{\ell}\int_{x}^{x+\ell}\psi(t)\mathrm{d}t=\frac{\psi_1(x+\ell)-\psi_1(x)}{\ell}\tag{92}$$

for $\ell>0$, and our first objective is to estimate the difference quotients on the left and right. Assume that $x\ge 3$, and that ℓ satisfies $1\le\ell\le x/2$. Then $x\pm\ell>1$, and the explicit formula (58) yields

$$\frac{\psi_1(x\pm\ell)-\psi_1(x)}{\pm\ell}=x\pm\frac{\ell}{2}-\sum_{\rho}\frac{(x\pm\ell)^{\rho+1}-x^{\rho+1}}{\pm\rho(\rho+1)\ell}-\frac{\zeta'}{\zeta}(0)-\sum_{r=1}^{\infty}\frac{(x\pm\ell)^{1-2r}-x^{1-2r}}{\pm2r(2r-1)\ell}.$$

If we denote the sum of the last two terms on the right by $K=K(x,\ell)$, then it is the case that $|K|<3$. Indeed,

$$\frac{\zeta'}{\zeta}(0)=\log(2\pi)=1.8378\ldots<2,$$

and, by the triangle inequality,

$$\begin{aligned}\Big|\sum_{r=1}^{\infty}\frac{(x\pm\ell)^{1-2r}-x^{1-2r}}{\pm2r(2r-1)\ell}\Big|&\le\sum_{r=1}^{\infty}\frac{(x-\ell)^{1-2r}+x^{1-2r}}{2r(2r-1)}\le\sum_{r=1}^{\infty}\frac{(\frac{x}{2})^{1-2r}+x^{1-2r}}{2r(2r-1)}\\&\le\sum_{r=1}^{\infty}\frac{(\frac{3}{2})^{1-2r}+3^{1-2r}}{2r(2r-1)}\le\sum_{r=1}^{\infty}\frac{1}{2r(2r-1)}\\&=\log 2=0.6931\ldots<1,\end{aligned}$$

since $\ell\le x/2$, and since each exponent $1-2r$ is negative.

We will handle the sum over the ρ's by splitting the sum at height T. The sum over the large ρ's can then be bounded as follows, using Corollary 2.4,

$$\begin{aligned}\Big|\sum_{|\gamma|>T}\frac{(x+\ell)^{\rho+1}-x^{\rho+1}}{\pm\rho(\rho+1)\ell}\Big|&\le 2\frac{(x+\ell)^{\frac32}+x^{\frac32}}{\ell}\sum_{\gamma>T}\frac{1}{\gamma^2}\le 2\frac{(x+\frac{x}{2})^{\frac32}+x^{\frac32}}{\ell}\sum_{\gamma>T}\frac{1}{\gamma^2}\\&=2\cdot\Big(\Big(\frac32\Big)^{\frac32}+1\Big)\frac{x^{\frac32}}{\ell}\Big(\frac{\log T}{2\pi T}+O\Big(\frac1T\Big)\Big)\le\frac{24}{25}\cdot\frac{x^{\frac32}\log T}{\ell T}\end{aligned}$$

whenever $T\ge T_1$, where $T_1>1$ is an absolute constant.

For the small ρ's, we write

$$\sum_{|\gamma|\le T}\frac{(x\pm\ell)^{\rho+1}-x^{\rho+1}}{\pm\rho(\rho+1)\ell}=\sum_{|\gamma|\le T}\frac{x^\rho}{\rho}+\sum_{|\gamma|\le T}v_\rho,$$

with

$$v_\rho=\frac{(x\pm\ell)^{\rho+1}-x^{\rho+1}\mp\ell(\rho+1)x^\rho}{\pm\rho(\rho+1)\ell}=x^{\rho+1}\cdot\frac{(1\pm\frac{\ell}{x})^{\rho+1}-1\mp(\rho+1)\frac{\ell}{x}}{\pm\rho(\rho+1)\ell}.$$

Suppose that $T\le\frac{x}{\ell}$. Then the numerator of the last factor on the right satisfies the assumptions of the lemma with $z=\pm\frac{\ell}{x}$ and $w=\rho+1$. Indeed $|z|\le\frac{x/2}{x}=1/2$, and

$$|zw|=\Big|\frac{\ell}{x}(\rho+1)\Big|\le\frac{\ell}{x}(|\rho|+1)<\frac{\ell}{x}\Big(|\gamma|+\frac32\Big)\le\frac{\ell}{x}\Big(T+\frac32\Big)\le\frac{\ell}{x}\cdot\frac{x}{\ell}+\frac{3\ell}{2x}\le1+\frac34<2.$$

As such, we get

$$|v_\rho| \le x^{\frac{3}{2}} \cdot \frac{\frac{13}{5}|\rho+1|(|\rho+1|+1)\frac{\ell^2}{x^2}}{|\rho||\rho+1|\ell} = \frac{13}{2}\frac{\ell}{\sqrt{x}}\frac{|\rho+1|+1}{|\rho|} \le \frac{13}{2}\frac{\ell}{\sqrt{x}}\frac{|\rho|+2}{|\rho|}$$
$$= \frac{13}{2}\frac{\ell}{\sqrt{x}}\Big(1+\frac{2}{|\rho|}\Big) < \frac{13}{2}\frac{\ell}{\sqrt{x}}\Big(1+\frac{2}{|\gamma|}\Big) < \frac{13}{2}\frac{\ell}{\sqrt{x}}\Big(1+\frac{2}{14}\Big) < 3\frac{\ell}{\sqrt{x}}.$$

This gives further

$$\Big|\sum_{|\gamma|\le T} v_\rho\Big| \le 6\frac{\ell}{\sqrt{x}}N(T) = 6\frac{\ell}{\sqrt{x}}\Big(\frac{T\log T}{2\pi}+O(T)\Big) < \frac{\ell T\log T}{\sqrt{x}}$$

whenever $T_2 \le T \le x/\ell$, with $T_2 > 1$ another absolute constant.

Let us collect the results obtained so far. If we let $g = O_{<1}(f)$ mean that $|g| < |f|$, then we have

$$\frac{\psi_1(x\pm\ell)-\psi_1(x)}{\pm\ell} = x \pm \frac{\ell}{2} - \sum_{|\gamma|\le T}\frac{x^\rho}{\rho} + O_{<1}\Big(\frac{x^{\frac{3}{2}}\log T}{\ell T}\Big) + O_{<1}\Big(\frac{\ell T\log T}{\sqrt{x}}\Big),$$

if $x \ge 3$, $1 \le \ell \le x/2$ and $T_3 \le T \le x/\ell$, where $T_3 > 1$ is yet another absolute constant. The term $K(x,\ell)$ was absorbed into the first error term, possibly subject to an increase in the smallest permitted value of T, because

$$\Big|\frac{24}{25}\cdot\frac{x^{\frac{3}{2}}\log T}{\ell T} + K\Big| < \frac{24}{25}\cdot\frac{x^{\frac{3}{2}}\log T}{\ell T} + 3 = \frac{x^{\frac{3}{2}}\log T}{\ell T}\Big(\frac{24}{25}+3\frac{\ell T}{x^{3/2}\log T}\Big)$$
$$\le \frac{x^{\frac{3}{2}}\log T}{\ell T}\Big(\frac{24}{25}+3\frac{1}{\sqrt{x}\log T}\Big) \le \frac{x^{\frac{3}{2}}\log T}{\ell T}\Big(\frac{24}{25}+\frac{1}{\sqrt{3}\log T}\Big)$$
$$< \frac{x^{\frac{3}{2}}\log T}{\ell T}$$

if, say, $T \ge \exp(\frac{25}{\sqrt{3}})+1$ (in particular, we may take $T_3 = \max(\exp(\frac{25}{\sqrt{3}})+1, T_1, T_2)$).

If we now use (92), then we obtain

$$\Big|\psi(x) - x + \sum_{|\gamma|\le T}\frac{x^\rho}{\rho}\Big| < \frac{\ell}{2} + \frac{x^{\frac{3}{2}}\log T}{\ell T} + \frac{\ell T\log T}{\sqrt{x}} = \frac{\ell}{2} + \sqrt{x}\log T\Big(\frac{x}{\ell T}+\frac{\ell T}{x}\Big).$$

under the same restrictions. We see that the contribution of the two last terms is minimal when $\ell T = x$, in which case

$$\Big|\psi(x) - x + \sum_{|\gamma|\le T}\frac{x^\rho}{\rho}\Big| < \frac{x}{2T} + 2\sqrt{x}\log T. \tag{93}$$

The dependence on the parameter ℓ may now be obliterated. Indeed, our investigation has led us to the assumptions $x \ge 3$, $1 \le \ell \le x/2$ and $T_3 \le T = x/\ell$, which implies that (93) holds independently of ℓ provided only that the inequality $T_3 \le T \le x$ is satisfied.

Finally, if $T \ge x \ge 3$, then the truncated explicit formula presented in Theorem 3.7 has an error term which is less than or equal to an absolute constant times

$$\frac{x\log^2(xT)}{T} + \log x = \frac{x\log^2 x}{T} + \frac{2x\log x\log T}{T} + \frac{x\log^2 T}{T} + \log x$$
$$= 2\sqrt{x}\log T\Big(\frac{\sqrt{x}\log^2 x}{2T\log T} + \frac{\sqrt{x}\log x}{T} + \frac{\sqrt{x}\log T}{2T} + \frac{\log x}{2\sqrt{x}\log T}\Big)$$
$$= (2\sqrt{x}\log T)O\big((\log T)^{-1}\big),$$

where the implied constant in the last line is absolute, as can be seen by substituting the estimates $T \geq x$ and $T \geq \sqrt{xT}$ across the denominators. Thus, the estimate (93) is seen to hold also if both $T \geq x \geq 3$ and $T \geq T_4$ are true, with $T_4 > 3$ an absolute constant. It therefore holds for all $x \geq 3$ and $T \geq \max(T_3, T_4)$. $\square$

Remark 5.5. The difference quotients used in the estimate (92) can be thought of as weighted sums of the form $\sum_{n=1}^{\infty} w(n)\Lambda(n)$. Indeed, using the identity $\psi_1(x) = \sum_{n\leq x}(x-n)\Lambda(n)$, we see that

$$\frac{\psi_1(x \pm \ell) - \psi_1(x)}{\pm \ell} = \sum_{n=1}^{\infty} w_{\pm}(n)\Lambda(n),$$

where

$$w_+(t) = \begin{cases} 1 & \text{if } 1 \leq t \leq x, \\ \frac{(x+\ell)-t}{\ell} & \text{if } x < t \leq x+\ell, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad w_-(t) = \begin{cases} 1 & \text{if } 1 \leq t \leq x-\ell, \\ \frac{x-t}{\ell} & \text{if } x-\ell < t \leq x, \\ 0 & \text{otherwise.} \end{cases}$$

These quotients therefore correspond to over- and under-estimates of $\psi(x)$ using piecewise linear weight functions.

Corollary 5.6. *Assume that the Riemann hypothesis is true. Then*

$$\psi(x) = x - \sum_{|\gamma|\leq T} \frac{x^\rho}{\rho} + O(\sqrt{x}\log x) \tag{94}$$

uniformly for $T \geq \sqrt{x}(\log x)^{-1}$ *as* $x \to \infty$. *Moreover,*

$$\Bigg|\psi(x) - x + \sum_{|\gamma|\leq \sqrt{x}(\log x)^{-1}} \frac{x^\rho}{\rho}\Bigg| < \frac{3}{2}\sqrt{x}\log x \tag{95}$$

for all sufficiently large x.

Proof. By Goldston's theorem, 5.4, if $\sqrt{x}(\log x)^{-1} \leq T \leq x$ and x is sufficiently large, then

$$\Bigg|\psi(x) - x + \sum_{|\gamma|\leq T} \frac{x^\rho}{\rho}\Bigg| < \frac{x}{2\sqrt{x}(\log x)^{-1}} + 2\sqrt{x}\log x = \frac{5}{2}\sqrt{x}\log x.$$

As we showed in the proof of Goldston's theorem, if $T \geq x$, then the explicit formula from Theorem 3.7 has error term

$$(2\sqrt{x}\log T)O((\log T)^{-1}) \ll \sqrt{x} = o(\sqrt{x}\log x),$$

and so (94) follows. To get (95), we simply substitute $T = \sqrt{x}(\log x)^{-1}$ into the estimate (91). This yields

$$\begin{aligned} \Bigg|\psi(x) - x + \sum_{|\gamma|\leq \sqrt{x}(\log x)^{-1}} \frac{x^\rho}{\rho}\Bigg| &< \frac{x}{2\sqrt{x}(\log x)^{-1}} + 2\sqrt{x}\log\big(\sqrt{x}(\log x)^{-1}\big) \\ &= \frac{3}{2}\sqrt{x}\log(x) - 2\sqrt{x}\log\log x \\ &< \frac{3}{2}\sqrt{x}\log x \end{aligned}$$

for all sufficiently large x. $\square$

We also get this corollary, whose first appearance in print was likely in Selberg's article [31].[29]

Corollary 5.7. *If the Riemann hypothesis is true, then*

$$\psi(x+h) - \psi(x) \sim h,$$
$$\pi(x+h) - \pi(x) \sim \frac{h}{\log x},$$

both hold as $x \to \infty$, provided that $h \leq x$ and $\sqrt{x} \log x = o(h)$. The statement is meaningful and true so long as $h = h(x)$ is defined on a sequence $(x_n)_{n=1}^{\infty}$ tending to infinity, and the bound $h \leq x$ needs only hold for all sufficiently large x.

We give two proofs of this fact. The first proof uses Goldston's theorem, while the second proof is more direct. They both ultimately rely on the explicit formula for $\psi_1(x)$. By virtue of Proposition 4.1, it suffices to prove the first formula.

Proof 1. We apply Goldston's theorem, assuming that $x \geq 3$, $x + h \geq 3$ and $T \geq T_0$. This gives us

$$\Big|\psi(x) - x + \sum_{|\gamma| \leq T} \frac{x^\rho}{\rho}\Big| < \frac{x}{2T} + 2\sqrt{x}\log T \quad \text{and}$$
$$\Big|\psi(x+h) - (x+h) + \sum_{|\gamma| \leq T} \frac{(x+h)^\rho}{\rho}\Big| < \frac{x+h}{2T} + 2\sqrt{x+h}\log T.$$

Now suppose that $h \ll x$. Then the bounds on the right are both

$$\ll \frac{x}{T} + \sqrt{x}\log T =: f(x,T).$$

On one hand, we want to choose $T = T(x)$ such that $f(x,T)$ is as small as possible. On the other hand, we want T to be as small as possible, such that the resulting sums $\Sigma_{|\gamma| \leq T}$ become as small as possible. For a given $x \geq 3$, the minimum of $f(x,T)$ is attained at $T = \sqrt{x}$, where $f(x, \sqrt{x}) = \frac{1}{2}\sqrt{x}\log x + \sqrt{x}$ (in particular, there is no way to avoid $\sqrt{x}\log x \ll f(x,T)$ using this strategy). But T can be taken smaller than this, since a simple inspection shows that $f(x,T) \ll \sqrt{x}\log x$ so long as $\sqrt{x}(\log x)^{-1} \ll T$ and $\log T \ll \log x$. We therefore take $T = \sqrt{x}(\log x)^{-1}$, and obtain as a result

$$\psi(x+h) - \psi(x) = h - \sum_{|\gamma| \leq \sqrt{x}(\log x)^{-1}} \frac{(x+h)^\rho - x^\rho}{\rho} + O(\sqrt{x}\log x)$$

as $x \to \infty$.

Under the Riemann hypothesis, we have from (68) and (69),

$$C(\rho) = \frac{(x+h)^\rho - x^\rho}{\rho} \ll \min\Big(\frac{h}{\sqrt{x}}, \frac{\sqrt{x}}{|\gamma|}\Big),$$

where the bound $h/\sqrt{x}$ is generally better for small values of $|\gamma|$, and $\sqrt{x}/|\gamma|$ the best for large $|\gamma|$. Let us try to find the 'sweet spot' which balances between these two bounds. We introduce a parameter $1 \leq \lambda \leq \sqrt{x}(\log x)^{-1}$, and write

$$\sum_{|\gamma| \leq \sqrt{x}(\log x)^{-1}} C(\rho) = \sum_{|\gamma| \leq \lambda} C(\rho) + \sum_{\lambda < |\gamma| \leq \sqrt{x}(\log x)^{-1}} C(\rho).$$

[29] Equation 4, p. 88. See also Footnote 3 on the same page.

The contribution of the small zeros is then

$$\sum_{|\gamma|\le\lambda} C(\rho) \ll \frac{hN(\lambda)}{\sqrt{x}} \ll \frac{h\lambda\log\lambda}{\sqrt{x}}.$$

This is not $o(h)$ unless $\lambda = o(\sqrt{x}(\log x)^{-1})$. Indeed, if we take $\lambda = \sqrt{x}(\varphi\log x)^{-1}$, where $1 \le \varphi(x) \le \sqrt{x}(\log x)^{-1}$, then

$$\frac{\lambda\log\lambda}{\sqrt{x}} = \frac{\log\big(\sqrt{x}(\varphi\log x)^{-1}\big)}{\varphi\log x} = \frac{1}{2\varphi} - \frac{\log\varphi}{\varphi\log x} - \frac{\log\log x}{\varphi\log x} = \frac{1}{2\varphi} + o(1),$$

which vanishes if and only if φ goes to infinity with x.

The contribution of the large zeros is then, according to Corollary 2.4,

$$\begin{aligned}
&\sum_{\lambda<|\gamma|\le\sqrt{x}(\log x)^{-1}} C(\rho) \ll \sqrt{x}\sum_{\sqrt{x}(\varphi\log x)^{-1}<\gamma\le\sqrt{x}(\log x)^{-1}} \frac{1}{\gamma}\\
&= \sqrt{x}\bigg(\frac{\log^2(\sqrt{x}(\log x)^{-1})}{4\pi} + O\Big(\log\frac{\sqrt{x}}{\log x}\Big) - \frac{\log^2(\sqrt{x}(\varphi\log x)^{-1})}{4\pi} - O\Big(\log\frac{\sqrt{x}}{\varphi\log x}\Big)\bigg)\\
&\ll \sqrt{x}\bigg(\Big(\frac{\log x}{2} - \log\log x\Big)^2 - \Big(\frac{\log x}{2} - \log\varphi(x) - \log\log x\Big)^2\bigg) + O(\sqrt{x}\log x)\\
&= \sqrt{x}\big(\log(x)\log(\varphi) - \log^2(\varphi) - 2\log(\varphi)\log(\log x)\big) + O(\sqrt{x}\log x)\\
&\ll \sqrt{x}\log(x)\log(\varphi) + O(\sqrt{x}\log x)
\end{aligned}$$

We therefore have

$$\psi(x+h) - \psi(x) = h + o(h) + O\big(\sqrt{x}\log(x)\log(\varphi)\big) + O(\sqrt{x}\log x),$$

and, if we take $h = \sqrt{x}\log(x)\varphi(x)$,

$$\psi(x+h) - \psi(x) \sim h.$$

□

The second proof follows an outline by Montgomery and Vaughan [27].[30]

Proof 2. The estimation of $\psi(x+h) - \psi(x) = \sum_{x<n\le x+h}\Lambda(n)$ may be more forgiving if the sum is not so brutally started at $n = x$ and ended at $n = x+h$ (compare this with Remark 5.5). That is, it may be easier to estimate if the sum is 'smoothed' somewhat.

Let $2 \le \Delta \le h \le x$, and define the piecewise linear weight function w by

$$w(t) = w(t,x,h,\Delta) = \begin{cases} 0 & \text{if } t \le x-\Delta,\\ \frac{t-(x-\Delta)}{\Delta} & \text{if } x-\Delta \le t \le x,\\ 1 & \text{if } x \le t \le x+h,\\ \frac{(x+h+\Delta)-t}{\Delta} & \text{if } x+h \le t \le x+h+\Delta,\\ 0 & \text{if } t \ge x+h+\Delta. \end{cases}$$

[30] Exercise 2, Section 13.1.1.

Using the formula $\psi_1(\lambda) = \sum_{n\le\lambda}\Lambda(n)(\lambda-n)$, together with some algebra, we can write the von Mangoldt function weighted by w as follows:

$$\begin{aligned}
\sum_{n=1}^{\infty}\Lambda(n)w(n) &= \sum_{x-\Delta<n\le x}\Lambda(n)\frac{n-x+\Delta}{\Delta} + \sum_{x<n\le x+h}\Lambda(n) + \sum_{x+h<n\le x+h+\Delta}\Lambda(n)\frac{x+h+\Delta-n}{\Delta}\\
&= \frac{1}{\Delta}\Big(\sum_{x-\Delta<n\le x}\Lambda(n)(n-x+\Delta) + \sum_{x<n\le x+h}\Lambda(n)\Delta + \sum_{x+h<n\le x+h+\Delta}\Lambda(n)(x+h+\Delta-n)\Big)\\
&= \frac{1}{\Delta}\Big(\psi_1(x+h+\Delta) + \sum_{x-\Delta<n\le x}\Lambda(n)(n-x+\Delta) + \sum_{x<n\le x+h}\Lambda(n)\Delta - \sum_{n\le x+h}\Lambda(n)(x+h+\Delta-n)\Big)\\
&= \frac{1}{\Delta}\Big(\psi_1(x+h+\Delta) - \psi_1(x+h) + \sum_{x-\Delta<n\le x}\Lambda(n)(n-x+\Delta) - \sum_{n\le x}\Lambda(n)\Delta\Big)\\
&= \frac{1}{\Delta}\Big(\psi_1(x+h+\Delta) - \psi_1(x+h) - \sum_{x-\Delta<n\le x}\Lambda(n)(x-n) - \sum_{n\le x-\Delta}\Lambda(n)\Delta\Big)\\
&= \frac{1}{\Delta}\Big(\psi_1(x+h+\Delta) - \psi_1(x+h) - \psi_1(x) + \sum_{n\le x-\Delta}\Lambda(n)(x-\Delta-n)\Big)\\
&= \frac{\psi_1(x+h+\Delta) - \psi_1(x+h) - \psi_1(x) + \psi_1(x-\Delta)}{\Delta}.
\end{aligned}$$

But if we use the explicit formula for ψ_1 on this last expression, then we are left with

$$h+\Delta-\frac{1}{\Delta}\sum_{\rho}S(\rho)+O\Big(\frac{1}{x\Delta}\Big),$$

assuming $x-\Delta\ge 1$ and $\Delta = o(x)$, where

$$S(\rho) = \frac{(x+h+\Delta)^{\rho+1}-(x+h)^{\rho+1}-x^{\rho+1}+(x-\Delta)^{\rho+1}}{\rho(\rho+1)}.$$

We may bound $S(\rho)$ in three different ways as follows. First,

$$S(\rho) = \int_{x-\Delta}^{x+h+\Delta}\Delta w(t)t^{\rho-1}\mathrm{d}t \ll \Delta\sup_{t\in\mathbb{R}}w(t)\big((x+h+\Delta)-(x-\Delta)\big)(x-\Delta)^{-\frac{1}{2}} \ll \frac{h\Delta}{\sqrt{x}}.$$

Second,

$$\begin{aligned}
|S(\rho)| &\le \frac{1}{|\gamma|}\bigg(\Big|\frac{(x+h+\Delta)^{\rho+1}-(x+h)^{\rho+1}}{\rho+1}\Big| + \Big|\frac{x^{\rho+1}-(x-\Delta)^{\rho+1}}{\rho+1}\Big|\bigg)\\
&= \frac{1}{|\gamma|}\bigg(\Big|\int_{x+h}^{x+h+\Delta}t^{\rho}\mathrm{d}t\Big| + \Big|\int_{x-\Delta}^{x}t^{\rho}\mathrm{d}t\Big|\bigg) \le \frac{\Delta}{|\gamma|}(\sqrt{x+h+\Delta}+\sqrt{x})\\
&\ll \frac{\Delta\sqrt{x}}{|\gamma|}.
\end{aligned}$$

Third, we have, by estimating trivially,

$$S(\rho) \ll \frac{x^{\frac{3}{2}}}{\gamma^2}.$$

Comparing these bounds, we find that the best one for different sizes of $|\gamma|$ can be summarised as

$$S(\rho) \ll \begin{cases} \dfrac{h\Delta}{\sqrt{x}} & \text{if } |\gamma| \le x/h, \\[2ex] \dfrac{\Delta\sqrt{x}}{|\gamma|} & \text{if } x/h < |\gamma| \le x/\Delta, \\[2ex] \dfrac{x^{3/2}}{\gamma^2} & \text{if } |\gamma| > x/\Delta. \end{cases}$$

Applying these to bound $\frac{1}{\Delta}\sum_\rho S(\rho)$, we get

$$\frac{1}{\Delta}\sum_{|\gamma|\le x/h} S(\rho) \ll \frac{1}{\Delta}\cdot\frac{h\Delta}{\sqrt{x}} N\Big(\frac{x}{h}\Big) \ll \frac{h}{\sqrt{x}}\cdot\frac{x}{h}\log\frac{x}{h} = \sqrt{x}\log\frac{x}{h} \ll \sqrt{x}\log x,$$
$$\frac{1}{\Delta}\sum_{|\gamma|> x/\Delta} S(\rho) \ll \frac{x^{\frac{3}{2}}}{\Delta}\sum_{\gamma>x/\Delta}\frac{1}{\gamma^2} \ll \frac{x^{\frac{3}{2}}}{\Delta}\cdot\frac{\log\frac{x}{\Delta}}{\frac{x}{\Delta}} = \sqrt{x}\log\frac{x}{\Delta} \ll \sqrt{x}\log x,$$

and

$$\begin{aligned}\frac{1}{\Delta}\sum_{x/h<|\gamma|\le x/\Delta} S(\rho) &\ll \frac{1}{\Delta}\cdot\Delta\sqrt{x}\sum_{x/h<\gamma\le x/\Delta}\frac{1}{\gamma} \\ &= \sqrt{x}\Big(\frac{\log^2(\frac{x}{\Delta})}{4\pi} + O(\log\frac{x}{\Delta}) - \frac{\log^2(\frac{x}{h})}{4\pi} - O(\log\frac{x}{h})\Big) \\ &\ll \sqrt{x}\log\bigg(\frac{x^2}{\Delta h}\bigg)\log\bigg(\frac{h}{\Delta}\bigg) + O(\sqrt{x}\log x) \\ &\ll \sqrt{x}\log(x)\log\bigg(\frac{2h}{\Delta}\bigg),\end{aligned}$$

where the factor 2 has been included in the numerator to absorb the $O(\sqrt{x}\log x)$. Putting this together, we get

$$\sum_{n=1}^{\infty}\Lambda(n)w(n,x,h,\Delta) = h + \Delta + O\Big(\sqrt{x}(\log x)\log\Big(\frac{2h}{\Delta}\Big)\Big),$$

where the dependence of w on all the parameters has been made explicit again. In particular,

$$\begin{aligned}\sum_{n=1}^{\infty}\Lambda(n)w(n,x+\Delta,h-2\Delta,\Delta) &= (h-2\Delta)+\Delta+O\Big(\sqrt{x+\Delta}\log(x+\Delta)\log\Big(\frac{2(h-2\Delta)}{\Delta}\Big)\Big) \\ &= h-\Delta+O\Big(\sqrt{x}(\log x)\log\Big(\frac{2h}{\Delta}\Big)\Big)\end{aligned}$$

so long as $\Delta \le h/3$.

The point is that, since $\psi(x+h)-\psi(x) = \sum_{x<n\le x+h}\Lambda(n)$, we have

$$\sum_{n=1}^{\infty}\Lambda(n)w(n,x+\Delta,h-2\Delta,\Delta) \le \psi(x+h)-\psi(x) \le \sum_{n=1}^{\infty}\Lambda(n)w(n,x,h,\Delta)$$

by construction, and hence

$$\psi(x+h)-\psi(x) = h + O(\Delta) + O\Big(\sqrt{x}(\log x)\log\Big(\frac{2h}{\Delta}\Big)\Big).$$

Taking $\Delta = \sqrt{x}\log x$, this implies in particular

$$\psi(x+h) - \psi(x) = h + O\Big(\sqrt{x}(\log x)\log\Big(\frac{2h}{\sqrt{x}\log x}\Big)\Big)$$

uniformly for $3\sqrt{x}\log x \le h \le x$, and

$$\psi(x+h) - \psi(x) \sim h$$

for $h = \sqrt{x}\log(x)f(x)$; provided f goes to infinity with x in such a way that $3 \le f \le \sqrt{x}(\log x)^{-1}$. $\square$

5.2 Cramér's theorem

The results of the preceding section are interesting by themselves, but what do they have to say when it comes to prime gaps? Well, using Proposition 4.2, we get

$$p_{n+1} - p_n \le \sqrt{p_n}\log^2(p_n)f(p_n) \quad \text{and}$$
$$p_{n+1} - p_n \le \sqrt{p_n}\log(p_n)f(p_n)$$

for all sufficiently large n, from Corollaries 5.2 and 5.7, respectively, where f is any function that goes to infinity with x. It is in fact possible to remove the appearance of the annoying function f altogether, and this, Cramér was the first to put in print. We follow a simpler proof.[31]

Theorem 5.8 (Cramér, 1920). *Suppose that the Riemann hypothesis is true. Then*

$$p_{n+1} - p_n = O(\sqrt{p_n}\log p_n)$$

as $n \to \infty$.

Proof. Let $0 < \varepsilon < \frac{1}{2}$ be fixed, and h a parameter depending on x, such that $x^{\varepsilon} < h < x^{1-\varepsilon}$ whenever $x \ge 3$. We will consider the moving interval $I(x, 2h) = (x, x+2h]$ as $x \to \infty$, in two different ways.

Suppose first that $I(x, 2h)$ does not contain any primes. Then neither does the subinterval $I(t, h) = (t, t+h]$ for $x \le t \le x+h$, which by necessity implies (with $N = \lfloor \log_2(t+h)\rfloor$):

$$\begin{aligned}
\psi(t+h) - \psi(t) &= \sum_{\substack{t < p^m \le t+h \\ m \ge 2}} \log p = \sum_{m=2}^{N} \sum_{t^{1/m} < p \le (t+h)^{1/m}} \log p \\
&\le \sum_{m=2}^{N} \frac{\log(t+h)}{m}\big(\pi((t+h)^{1/m}) - \pi(t^{1/m})\big) \\
&= \log(t+h)\sum_{m=2}^{N} \frac{\pi(t^{1/m} + (t+h)^{1/m} - t^{1/m}) - \pi(t^{1/m})}{m} \\
&\le \log(t+h)\sum_{m=2}^{N} \frac{(t+h)^{1/m} - t^{1/m} + 1}{m} \le \log(t+h)\sum_{m=2}^{N} \frac{t^{1/m}((1+\frac{h}{t})^{1/m} - 1) + 1}{m} \\
&\le \log(t+h)\sum_{m=2}^{N} \frac{t^{1/m}\cdot\frac{h}{t}\cdot\frac{1}{m} + 1}{m} \\
&\le \frac{h\log(t+h)}{\sqrt{t}}\sum_{m=2}^{\infty}\frac{1}{m^2} + \log(t+h)\sum_{m=2}^{N}\frac{1}{m} = O_{\varepsilon}\Big(\frac{h\log x}{\sqrt{x}} + \log x\log\log x\Big),
\end{aligned}$$

[31] Ivić [21], Theorem 12.10. This proof was outlined by Ingham [19]: see Footnote §, p. 256.

where the implied constant depends only on ε (this fact is crucial for the proof). In particular, we conclude from this that

$$\int_x^{x+h} \psi(t+h)-\psi(t)\mathrm{d}t = O_\varepsilon\Big(\frac{h^2\log x}{\sqrt{x}}+h\log x\log\log x\Big)=o(h^2). \tag{96}$$

We now want to estimate this integral once again by using the explicit formula for $\psi(\,\cdot\,)$.

We recall that

$$\psi(t)=t-\sum_{|\gamma|\le T}\frac{x^\rho}{\rho}+O\Big(\frac{t\log^2(tT)}{T}+\log t\Big)$$

uniformly for $t\ge 3$ and $T\ge 1$, so that

$$\psi(t+h)=t+h-\sum_{|\gamma|\le T}\frac{(t+h)^\rho}{\rho}+O\Big(\frac{(t+h)\log^2\big((t+h)T\big)}{T}+\log(t+h)\Big)$$

uniformly for $t+h\ge 3$ and $T\ge 1$. Restricting t again to the interval $x\le t\le x+h$, then $t\ll_\varepsilon x$ and $t+h\ll_\varepsilon x$ as $x\to\infty$, so that

$$\psi(t+h)-\psi(t)=h-\sum_{|\gamma|\le T}C(\rho)+O_\varepsilon\Big(\frac{x\log^2(xT)}{T}+\log x\Big)$$

(now $C(\rho)=((t+h)^\rho-t^\rho)/\rho)$. Take $T=x$, with the conclusion that

$$\psi(t+h)-\psi(t)=h-\sum_{|\gamma|\le x}C(\rho)+O_\varepsilon(\log^2 x).$$

We can now integrate this relation over $t\in(x,x+2h]$, obtaining

$$\int_x^{x+h}\psi(t+h)-\psi(t)\,\mathrm{d}t=h^2-\sum_{|\gamma|\le x}\int_x^{x+h}C(\rho)\mathrm{d}t+O_\varepsilon(h\log^2 x).$$

We introduce an auxiliary parameter $3<\lambda<x$. Then we have first

$$\begin{aligned}\sum_{|\gamma|\le\lambda}\int_x^{x+h}C(\rho)\mathrm{d}t &= \sum_{|\gamma|\le\lambda}\int_x^{x+h}\frac{(t+h)^\rho-t^\rho}{\rho}\mathrm{d}t=\int_x^{x+h}\sum_{|\gamma|\le\lambda}\int_t^{t+h}z^{\rho-1}\mathrm{d}z\mathrm{d}t\\ &=\int_x^{x+h}\int_t^{t+h}z^{-\frac12}\sum_{|\gamma|\le\lambda}z^{\beta i}\mathrm{d}z\mathrm{d}t\le\int_x^{x+h}\int_t^{t+h}z^{-\frac12}\cdot 2N(\lambda)\mathrm{d}z\mathrm{d}t\\ &=O\Big(\frac{h^2\lambda\log\lambda}{\sqrt{x}}\Big),\end{aligned} \tag{97}$$

where the implied constant is absolute, since $z^{-\frac12}\le t^{-\frac12}\le x^{-\frac12}$. Also, by actually computing the integral, we get

$$\begin{aligned}\sum_{\lambda<|\gamma|\le x}\int_x^{x+h}\frac{(t+h)^\rho-t^\rho}{\rho}\mathrm{d}t &= \sum_{\lambda<|\gamma|\le x}\frac{(x+2h)^{\rho+1}-2(x+h)^{\rho+1}+x^{\rho+1}}{\rho(\rho+1)}\\ &\ll_\varepsilon x^{\frac32}\sum_{|\gamma|>\lambda}\gamma^{-2}=O_\varepsilon\Big(x^{\frac32}\frac{\log\lambda}{\lambda}\Big)\end{aligned} \tag{98}$$

by Corollary 2.4.

Thus, the estimates (96), (97) and (98) yield

$$o(h^2) = h^2 + O\Big(\frac{h^2\lambda\log\lambda}{\sqrt{x}}\Big) + O_\varepsilon\Big(x^{\frac{3}{2}}\frac{\log\lambda}{\lambda}\Big) + O_\varepsilon(h\log^2 x).$$

Now take $\lambda = x/h$. Then this reduces to the statement that

$$o(h^2) = h^2 + O_\varepsilon(h\sqrt{x}\log x).$$

But this is impossible if $h = C\sqrt{x}\log x$ with $C > 0$ sufficiently large. Thus, $(x, x + 2C\sqrt{x}\log x]$ contains a prime for all $x \geq x_0$, and so

$$p_{n+1} - p_n \leq 2C\sqrt{p_n}\log p_n$$

for all primes $p_n \geq x_0$. This completes the proof. □

The proof of Cramér's theorem given above, although instructive, can be rendered unnecessary: if the Riemann hypothesis is true, then 'Cramér's bound' follows from Corollary 5.7 and the following result.

Proposition 5.9. *Let $\mathcal{F}$ denote the collection of all real-valued partial functions on $\mathbb{R}$ such that*

1. *Each $h = h(x) \in \mathcal{F}$ is defined for arbitrarily large values of x, and*
2. *Each $h = h(x) \in \mathcal{F}$ satisfies $h(x) \leq x$ for all sufficiently large values of x in its domain.*

Suppose that

$$\forall h \in \mathcal{F} : \Bigg(\lim_{\substack{x\to\infty \\ x\in\mathrm{domain}(h)}} \frac{h(x)}{\sqrt{x}\log x} = \infty\Bigg) \implies \Big(\pi\big(x + h(x)\big) - \pi(x) \sim \frac{h(x)}{\log x} \text{ as } x \to \infty\Big).$$

Then Cramér's bound $p_{n+1} - p_n = O(\sqrt{p_n}\log p_n)$ holds as $n \to \infty$.

Proof. For the sake of readability, let $p(n) := p_n$ denote the nth prime. If Cramér's bound does not hold, then there exists a strictly increasing sequence of positive integers $(n_k)_{k=1}^\infty$ such that

$$\lim_{k\to\infty}\frac{p(n_k+1) - p(n_k)}{\sqrt{p(n_k)}\log p(n_k)} = \infty.$$

Let $\mathcal{D} = (p(n_k))_{k=1}^\infty$, and define $h : \mathcal{D} \to \mathbb{R}$ by

$$h\big(p(n_k)\big) = p(n_k+1) - p(n_k) \quad \text{for } k \geq 1.$$

On the one hand, it follows from Theorem 4.4, Equation (75), and Propositions 4.1 and 4.2, that $h(p(n_k)) = o(p(n_k))$ as $k \to \infty$ (thus, $h(x) \leq x$ for all sufficiently large $x \in \mathcal{D}$; also $h \in \mathcal{F}$). On the other hand,

$$\lim_{\substack{x\to\infty \\ x\in\mathrm{domain}(h)}} \frac{h(x)}{\sqrt{x}\log x} = \lim_{k\to\infty}\frac{h(p(n_k))}{\sqrt{p(n_k)}\log p(n_k)} = \lim_{k\to\infty}\frac{p(n_k+1) - p(n_k)}{\sqrt{p(n_k)}\log p(n_k)} = \infty.$$

Thus, we have by assumption

$$\pi\big(x + h(x)\big) - \pi(x) \sim \frac{h(x)}{\log x}$$

as $x \to \infty$. However, taking $x = p(n_k)$, we see that this implies

$$1 = \pi\big(p(n_k+1)\big) - \pi\big(p(n_k)\big) \sim \frac{p(n_k+1) - p(n_k)}{\log p(n_k)} \geq \frac{p(n_k+1) - p(n_k)}{\sqrt{p(n_k)}\log p(n_k)} \to \infty,$$

which is impossible.[32] Thus, Cramér's bound must hold. □

[32] The same contradiction can be attained also if one demands h to be a function defined on (a, ∞) for some $a > 0$: simply extend the function in the proof by defining $h(x) = x^\vartheta$ for $x \neq p(n_k)$, where ϑ is fixed with $1/2 < \vartheta < 1$.

6 Primes in short intervals: beyond the Riemann hypothesis

Although it is outside the scope of this project to make an extensive list of every fact concerning the difference $\psi(x+h)-\psi(x)$, it would nevertheless be interesting to consider what kind of results that could conceivably be proved in the future. Much of subsequent work and conjectures on primes in short intervals is based on Selberg's article [31] as well as probabilistic models for the prime numbers. The latter will not be discussed here, but we will discuss some of Selberg's results.

6.1 The work of Selberg

Selberg relaxed the requirement that an asymptotic prime number theorem has to hold for all $x>0$, to 'almost all $x>0$':

A statement P is said to be true for *almost all* $x>0$ if there exists a set $E\subset(0,\infty)$ such that P is true as $x\to\infty$ through any sequence of numbers in $(0,\infty)\setminus E$, and such that $\lambda((0,x)\cap E)=o(x)$ as $x\to\infty$, where λ denotes the Lebesgue measure.[33] In other words, the statement P is true except on a (asymptotically) small set E of exceptional points.

Note that this is completely unrelated to the notion of 'almost everywhere' from measure theory.

Such a relaxation allows for results that are valid for much shorter intervals than in the preceding sections. Selberg proved, among other things:

Theorem 6.1 (Selberg, 1943). *Suppose that*

$$N(\sigma,T)=O\big(T^{\alpha(1-\sigma)}\log^{\eta}T\big)$$

uniformly for $\frac{1}{2}\le\sigma\le 1$ as $T\to\infty$, where $\alpha\ge 2$ and $\eta\ge 1$. Let h be a positive and increasing function of x such that

1. *$h(x)/x$ is decreasing for $x>0$ with $\lim\limits_{x\to\infty}h(x)/x=0$, and*
2. *$\liminf\limits_{x\to\infty}\dfrac{\log h(x)}{\log x}>1-2\alpha^{-1}$.*

Then we have

$$\pi\big(x+h(x)\big)-\pi(x)\sim\frac{h(x)}{\log x}$$

for almost all $x>0$.

Proof. We follow Selberg's proof except for some small modifications, allowing for general values of α and η.[34]

If we assume throughout that $3\le T\le x$, then we may use the explicit formula (38) to write

$$\psi(x)=x-\sum_{|\gamma|<T}\frac{x^{\rho}}{\rho}+O\Big(\frac{x\log^2 x}{T}\Big).$$

Let also y be a real number confined to the range $x\le y\le 2x$, and $r=r(x)\ge 1$ a real parameter which we may specify later at our convenience. If we define δ implicitly by $\mathrm{e}^{\delta}=1+r^{-1}$, then

$$\psi\Big(y+\frac{y}{r}\Big)-\psi(y)-\frac{y}{r}=-\sum_{|\gamma|<T}\frac{\mathrm{e}^{\delta\rho}-1}{\rho}y^{\rho}+O\Big(\frac{x\log^2 x}{T}\Big).$$

[33] We write 'statement' here, and not 'proposition function $P(x)$', because the statement may not be local around x, but global, as is the case in Selberg's theorem.

[34] The generalised statement appears in Montgomery [26], p. 131. Selberg worked with $\alpha=\frac{77}{29}+\varepsilon$ and $\eta=5$, the best known values at the time. Cf. Footnote 22 below Hoheisel's theorem.

For a given $0 < \varepsilon < 1/\alpha$, take $T = x^{2/\alpha - \varepsilon}$. Moreover, write $\log x \le x^{\varepsilon/4}$, assuming $x \ge x_0(\varepsilon) \ge 3$, say, to obtain

$$\psi\Big(y + \frac{y}{r}\Big) - \psi(y) - \frac{y}{r} = -\sum_{|\gamma|<T} \frac{\mathrm{e}^{\delta\rho} - 1}{\rho} y^{\rho} + O\big(x^{1-2/\alpha+3\varepsilon/2}\big).$$

From this, it now follows that

$$\begin{aligned}\frac{1}{x}\int_x^{2x} \Big|\psi\Big(y + \frac{y}{r}\Big) - \psi(y) - \frac{y}{r}\Big|^2 \mathrm{d}y = O\bigg(\frac{1}{x}\int_x^{2x} \sum_{|\gamma|<T}\sum_{|\gamma'|<T} \frac{\mathrm{e}^{\delta\rho} - 1}{\rho} \cdot \frac{\mathrm{e}^{\delta\overline{\rho'}} - 1}{\overline{\rho'}} y^{\rho+\overline{\rho'}} \mathrm{d}y\bigg)&\\ + O\big(x^{2-4/\alpha+3\varepsilon}\big)&\end{aligned} \tag{99}$$

(we used the estimate $(A+B)^2 \ll A^2 + B^2$ here, together with the relation $|z|^2 = z\overline{z}$).

Note that

$$\Big|\frac{\mathrm{e}^{\delta\rho} - 1}{\rho}\Big| = \Big|\int_0^{\delta} \mathrm{e}^{\lambda\rho}\mathrm{d}\lambda\Big| \le \int_0^{\delta} \mathrm{e}^{\lambda\beta}\mathrm{d}\lambda = \frac{\mathrm{e}^{\delta\beta} - 1}{\beta} \overset{*}{\le} \mathrm{e}\delta = \mathrm{e}\log\Big(1 + \frac{1}{r}\Big) < \frac{\mathrm{e}}{r},$$

where the inequality $*$ holds by virtue of

$$0 < \delta\beta < \delta = \log\Big(1 + \frac{1}{r}\Big) < \frac{1}{r} \le 1 < 1.7507\ldots,$$

with $1.7507\ldots$ denoting the positive real solution to the equation $\mathrm{e}^x - 1 = \mathrm{e}x$. Therefore,

$$\begin{aligned}\frac{1}{x}\int_x^{2x} \sum_{|\gamma|<T}\sum_{|\gamma'|<T} \frac{\mathrm{e}^{\delta\rho} - 1}{\rho} \cdot \frac{\mathrm{e}^{\delta\overline{\rho'}} - 1}{\overline{\rho'}} y^{\rho+\overline{\rho'}} \mathrm{d}y &= O\bigg(\frac{1}{xr^2}\sum_{|\gamma|<T}\sum_{|\gamma'|<T}\int_x^{2x} y^{\rho+\overline{\rho'}}\mathrm{d}y\bigg)\\ &= O\bigg(\frac{1}{r^2}\sum_{|\gamma|<T}\sum_{|\gamma'|<T} \frac{2^{\rho+\overline{\rho'}+1} - 1}{\rho + \overline{\rho'} + 1} x^{\rho+\overline{\rho'}}\bigg)\\ &= O\bigg(\frac{1}{r^2}\sum_{|\gamma|<T}\sum_{|\gamma'|<T} \frac{1}{1 + |\gamma - \gamma'|} x^{\beta+\beta'}\bigg)\\ &= O\bigg(\frac{1}{r^2}\sum_{0<\gamma<T} x^{2\beta}\sum_{|\gamma'|<T} \frac{1}{1 + |\gamma - \gamma'|}\bigg).\end{aligned} \tag{100}$$

In the third line we used the simple estimates

$$\begin{aligned}&|2^{\rho+\overline{\rho'}+1} - 1| \le 2^{\beta+\beta'+1} + 1 < 9,\\ &\frac{\frac{1}{|\rho+\overline{\rho'}+1|}}{\frac{1}{1+|\gamma-\gamma'|}} = \frac{1 + |\gamma - \gamma'|}{\sqrt{(1 + \beta + \beta')^2 + |\gamma - \gamma'|^2}} < \sup_{\theta \ge 0} \frac{1 + \theta}{\sqrt{1 + \theta^2}} = \sqrt{2},\end{aligned}$$

and for the last line we used

$$\begin{aligned}\sum_{\substack{|\gamma|<T\\ |\gamma'|<T}} \frac{x^{\beta+\beta'}}{1 + |\gamma - \gamma'|} &= \sum_{\substack{|\gamma|<T\\ |\gamma'|<T\\ \beta\ge\beta'}} \frac{x^{\beta+\beta'}}{1 + |\gamma - \gamma'|} + \sum_{\substack{|\gamma|<T\\ |\gamma'|<T\\ \beta<\beta'}} \frac{x^{\beta+\beta'}}{1 + |\gamma - \gamma'|} \le 2\sum_{\substack{|\gamma|<T\\ |\gamma'|<T\\ \beta\ge\beta'}} \frac{x^{\beta+\beta'}}{1 + |\gamma - \gamma'|}\\ &= 4\sum_{\substack{0<\gamma<T\\ |\gamma'|<T\\ \beta\ge\beta'}} \frac{x^{\beta+\beta'}}{1 + |\gamma - \gamma'|} \le 4\sum_{0<\gamma<T} x^{2\beta}\sum_{\substack{|\gamma'|<T\\ \beta'\le\beta}} \frac{1}{1 + |\gamma - \gamma'|}\\ &\le 4\sum_{0<\gamma<T} x^{2\beta}\sum_{|\gamma'|<T} \frac{1}{1 + |\gamma - \gamma'|}.\end{aligned}$$

The inner sum in (100) can be estimated as follows:

$$\sum_{|\gamma'|<T} \frac{1}{1+|\gamma-\gamma'|} = \sum_{n=1}^{\infty} \sum_{\substack{|\gamma'|<T \\ n-1\le|\gamma-\gamma'|<n}} \frac{1}{1+|\gamma-\gamma'|} \le \sum_{n=1}^{\lfloor 2T\rfloor+1} \frac{1}{n} \sum_{\substack{|\gamma'|<T \\ n-1\le|\gamma-\gamma'|<n}} 1$$

$$\ll \sum_{n=1}^{\lfloor 2T\rfloor+1} \frac{\log T}{n} \ll \log^2 T \le \log^2 x, \tag{101}$$

since

$$\{\rho' : |\gamma'| < T \text{ and } n-1 \le |\gamma-\gamma'| < n\} = \varnothing$$

when $n \ge 2T+1 > |\gamma| + T + 1$, and since

$$N(n) - N(n-1) = O(\log n) = O(\log T).$$[35]

Now take $1 \le r \le x^{2/\alpha-2\varepsilon}$. Then

$$x^{2-4/\alpha+3\varepsilon} = \frac{x^2}{(x^{2/\alpha-2\varepsilon})^2 (x^{\varepsilon/4})^4} \le \frac{x^2}{r^2\log^4 x} \tag{102}$$

for $x \ge x_0(\varepsilon)$.

Thus, we get from (99), (100), (101) and (102);

$$\frac{1}{x}\int_x^{2x} \left|\psi\Big(y+\frac{y}{r}\Big) - \psi(y) - \frac{y}{r}\right|^2 \mathrm{d}y = O\Big(\frac{x^2\log^2 x}{r^2} \sum_{0<\gamma<T} x^{2\beta-2}\Big) + O\Big(\frac{x^2}{r^2\log^4 x}\Big). \tag{103}$$

From (21), it follows that $\zeta(s)$ has no zeros in a region of the form $\sigma \ge L(t)$, $t \ge t_0$, where

$$L(t) = 1 - (\log\log t)^2 (\log t)^{-1}.$$

(*continued on the next page.*)

[35] Ingham [18], Theorem 25 a, p. 70.

Using this together with the identity $\int_{-1}^{\beta} x^{2(\sigma-1)}\mathrm{d}\sigma = \frac{x^{2\beta-2}-x^{-4}}{2\log x}$, we get

$$\begin{aligned}
\sum_{0<\gamma<T} x^{2\beta-2} &= 2\log(x) \sum_{0<\gamma<T} \int_{-1}^{\beta} x^{2(\sigma-1)}\mathrm{d}\sigma + x^{-4} \sum_{0<\gamma<T} 1 \\
&\ll \log(x) \sum_{0<\gamma<T} \int_{-1}^{\beta} x^{2(\sigma-1)}\mathrm{d}\sigma \\
&= \log(x) \int_{-1}^{1} \sum_{\substack{0<\gamma<T \\ \beta\geq\sigma}} x^{2(\sigma-1)}\mathrm{d}\sigma \\
&= \log(x) \int_{-1}^{1} x^{2(\sigma-1)} N(\sigma,T)\mathrm{d}\sigma \\
&\ll (\log x)^{\eta+1} \int_{-1}^{L(x)} \left(\frac{T^\alpha}{x^2}\right)^{1-\sigma} \mathrm{d}\sigma \\
&\leq (\log x)^{\eta+1} \left(\frac{T^\alpha}{x^2}\right)^{1-L(x)} \int_{-1}^{L(x)} \mathrm{d}\sigma \\
&\leq 2(\log x)^{\eta+1} \left(\frac{x^{(2/\alpha-\varepsilon)\alpha}}{x^2}\right)^{\frac{(\log\log x)^2}{\log x}} \\
&\leq 2(\log x)^{\eta+1} x^{-\varepsilon(\log\log x)^2(\log x)^{-1}} \\
&= 2(\log x)^{\eta+1} \mathrm{e}^{-\varepsilon(\log\log x)^2} \\
&\ll \frac{1}{\log^6 x}.
\end{aligned}$$

In the fifth line we used that $L(T) \leq L(x)$ ultimately holds for x sufficiently large[36], and in the sixth line we used that the integrand increases with σ (since $T^\alpha/x^2 = x^{-\alpha\varepsilon} < 1$).

Plugging this into (103) yields the bound

$$\frac{1}{x}\int_x^{2x} \left|\psi\Big(y+\frac{y}{r}\Big) - \psi(y) - \frac{y}{r}\right|^2 \mathrm{d}y = O\Big(\frac{x^2}{r^2\log^4 x}\Big). \tag{104}$$

Now suppose that $h(x)$ satisfies conditions *1* and *2*. Then we may make ε smaller if necessary, and so assume that $h(x) \geq 2x^{1-2/\alpha+2\varepsilon}$ for $x \geq x_1(\varepsilon) \geq x_0(\varepsilon)$. Thus, if we take $r = \frac{x'}{h(x')}$ for $x_1(\varepsilon) \leq x \leq x' \leq 2x$, then

$$r < \frac{x'}{2(x')^{1-2/\alpha+2\varepsilon}} = \frac{1}{2}(x')^{2/\alpha-2\varepsilon} \leq 2^{2/\alpha-1-2\varepsilon} x^{2/\alpha-2\varepsilon} < x^{2/\alpha-2\varepsilon}.$$

That is, this choice of r satisfies the assumptions for (104). Using the fact that this $r = r(x')$ increases with x', we get

$$\frac{\log x}{x}\int_{x'}^{x'+\frac{x}{[\log x]}} \left|\psi\Big(y+\frac{h(x')}{x'}y\Big) - \psi(y) - \frac{h(x')}{x'}y\right|^2 \mathrm{d}y = O\Big(\frac{h^2(x)}{\log^3 x}\Big), \tag{105}$$

provided $x \leq x' < x' + \frac{x}{[\log x]} \leq 2x$. This implies that

$$\psi\Big(y+\frac{h(x')}{x'}y\Big) - \psi(y) - \frac{h(x')}{x'}y < \frac{h(y)}{\log y}$$

[36] Recall that $T \leq x^{2/\alpha-\varepsilon} \leq x$. The function $1-L(t)$ is decreasing for $\mathrm{e} < t < \exp(\exp(2))$ and increasing for $t > \exp(\exp(2)) = 1618.177\ldots$

for $x' \leq y \leq x' + \frac{x}{\lfloor \log x \rfloor}$, except in a subset E_1 of $\left[x', x' + \frac{x}{\lfloor \log x \rfloor}\right]$ of Lebesgue measure

$$\lambda(E_1) = O\Big(\frac{x}{\log^2 x}\Big). \tag{106}$$

Indeed,

$$\frac{\log x}{x} \int_{E_1} \Big|\psi\Big(y + \frac{h(x')}{x'} y\Big) - \psi(y) - \frac{h(x')}{x'} y\Big|^2 \mathrm{d}y \geq \frac{\log x}{x} \int_{E_1} \frac{h^2(y)}{\log^2 y} \mathrm{d}y \gg \frac{h^2(x)\lambda(E_1)}{x \log x},$$

which would violate (105) if (106) did not hold.

Moreover, if $y \in \left[x', x' + \frac{x}{\lfloor \log x \rfloor}\right]$ does not belong to this exceptional subset E_1, then

$$\begin{aligned} \psi\big(y + h(y)\big) - \psi(y) - h(y) &\leq \psi\Big(y + \frac{h(x')}{x'} y\Big) - \psi(y) - \frac{h(x')}{x'} y + \frac{h(x')}{x'} y - h(y) \\ &< \frac{h(y)}{\log y} + h(y)(\frac{y}{x'} - 1) \leq \frac{h(y)}{\log y} + \frac{h(y)}{\lfloor \log x \rfloor} \\ &\leq \frac{h(y)}{\log y} + \frac{h(y)}{\lfloor \log \frac{y}{2} \rfloor} < \frac{3h(y)}{\log y}, \end{aligned}$$

since $h(y)/y \leq h(x')/x'$ and $h(x') \leq h(y)$, and where the last inequality holds for $y > 15$, say.

If we apply the argument above for the specific choices

$$x, \quad x + \frac{x}{\lfloor \log x \rfloor}, \quad x + \frac{2x}{\lfloor \log x \rfloor}, \ \ldots, \ x + \frac{\lfloor \log x \rfloor - 1}{\lfloor \log x \rfloor} x$$

for x', then we find that

$$\psi\big(y + h(y)\big) - \psi(y) - h(y) < \frac{3h(y)}{\log y} \tag{107}$$

for $y \in [x, 2x]$, except in a subset E_2 of $[x, 2x]$ of measure

$$\lambda(E_2) = O\Big(\frac{x}{\log^2 x} \cdot \log x\Big) = O\Big(\frac{x}{\log x}\Big).$$

If we substitute $\frac{x}{2}$, $\frac{x}{4}$, $\frac{x}{6}$, $\ldots$ in place of x, then, since the implied constant in the bound for $\lambda(E_2)$ will be the same in all cases, we get that (107) holds for all $y \in (0, x)$, except in a subset E_3 of $(0, x)$ of measure

$$\lambda(E_3) = O\Big(\frac{x}{\log x}\Big).$$

An identical argument can be used to show that

$$\psi\big(y + h(y)\big) - \psi(y) - h(y) > -\frac{3h(y)}{\log y}$$

for $y \in (0, x)$, except in a subset E_4 of $(0, x)$ of measure

$$\lambda(E_4) = O\Big(\frac{x}{\log x}\Big).$$

Thus,

$$\big|\psi\big(y + h(y)\big) - \psi(y) - h(y)\big| < \frac{3h(y)}{\log y}$$

for $y \in (0, x)$, except in a subset E of $(0, x)$ of measure

$$\lambda(E) = \lambda(E_3 \cup E_4) = O\Big(\frac{x}{\log x}\Big) = o(x).$$

That is,

$$\psi\big(x + h(x)\big) - \psi(x) = h(x) + O\Big(\frac{h(x)}{\log x}\Big)$$

holds for almost all $x > 0$. By virtue of Equation (64), this implies the stated result. $\square$

By taking $h(x) = x^\varepsilon$ and assuming the strong density hypothesis, we get:

Corollary 6.2. *Suppose that the strong density hypothesis is true. That is, suppose that $N(\sigma, T) = O\big(T^{2(1-\sigma)} \log^\eta T\big)$ uniformly for $\frac{1}{2} \leq \sigma \leq 1$ as $T \to \infty$, where $\eta \geq 1$. Let $\varepsilon > 0$. Then we have*

$$\pi(x + x^\varepsilon) - \pi(x) \sim \frac{x^\varepsilon}{\log x}$$

for almost all $x > 0$. In particular, the measure of the collection of $X \in (0, x)$ such that $[X, X + X^\varepsilon]$ does not contain a prime is $o(x)$.

These results show that we can get an asymptotic prime number theorem for much shorter intervals if we allow for a few exceptions. Even more can be proved under the Riemann hypothesis: Selberg showed the Riemann hypothesis implies $\pi(x + h) - \pi(x) \sim h(\log x)^{-1}$ for almost all $x > 0$, so long as $h = h(x)$ is increasing, h/x is decreasing, and

$$\lim_{x\to\infty} \frac{x}{h} = \lim_{x\to\infty} \frac{h}{\log^2 x} = \infty.$$

That is, we would get an 'almost- prime number theorem' so long as the length of the interval grows faster than $\log^2 x$.

6.2 Further investigations

In addition to the results above, there are many conditional results and conjectures about primes in short intervals. For example, by assuming the Riemann hypothesis and various additional properties of the distribution of the ordinates γ of the nontrivial zeros of $\zeta(s)$, Heath-Brown [16] showed that some improvements are possible:

$$\psi(x) - x = o(\sqrt{x} \log^2 x),$$
$$p_{n+1} - p_n = O\big(\sqrt{p_n \log p_n}\,\big),$$

and even

$$\pi(x + h) - \pi(x) > 0$$

for almost all $x > 0$, so long as $\log x = o(h)$.

Sadly, Heath-Brown's estimate for the size of prime gaps is nowhere close to what we expect to be the truth. For example, assuming $\pi(x+h)-\pi(x) \sim h(\log x)^{-1}$, Proposition 4.2 gives $p_{n+1}-p_n \leq h(p_n)$ for n sufficiently large. This is clearly not optimal, since the assumption says that the interval $(x, x+h]$ contains about $h(\log x)^{-1}$ primes. Thus, if the primes were evenly distributed over this interval (which they are unlikely to be), then our best naive guess would be that the bound is closer to

$$p_{n+1} - p_n \leq \frac{\text{length of interval}}{\text{number of primes}} = \frac{h}{h(\log x)^{-1}} = \log x.$$

This is much smaller than h if, say, $h = x^{\vartheta}$ for some fixed $\vartheta > 0$. The problem here is, of course, that the primes may all be clumped together in one half of the interval.

A cause of this discrepancy can be illustrated as follows. If we take $x = p_n$ and $x + h = p_{n+1}$ in the explicit formula (38), then we see that

$$p_{n+1} - p_n = \sum_{\rho} \frac{p_{n+1}^{\rho} - p_n^{\rho}}{\rho} + O(\log p_n),$$

if there are no prime powers between p_n and p_{n+1}. This shows explicitly the relation between prime gaps and sums of the form

$$\sum_{\rho} \frac{(x+h)^{\rho} - x^{\rho}}{\rho} \quad \text{or} \quad \sum_{\rho} \frac{x^{\rho}}{\rho}.$$

Estimating these sums is in fact the crux of the matter: the proofs in Section 4 consisted of estimating the first sum using a rather naive strategy of 'taking the absolute value and using the triangle inequality.' Unfortunately, this eliminates all possibility that there may be some significant cancellation among the different terms of the sum.

To round off, we mention this reformulation of Selberg's result due to Saffari and Vaughan [30]:

Theorem 6.3. *If the Riemann hypothesis is true, then*

$$\frac{1}{x} \int_x^{2x} \left|\psi(t + \theta t) - \psi(t) - \theta t\right|^2 \mathrm{d}t = O\Big(\theta x \log^2\Big(\frac{2}{\theta}\Big)\Big)$$

uniformly for $x \geq 4$ and $0 < \theta \leq 1$.[37]

The left hand side of Saffari and Vaughan's estimate is the average value of the integrand for $x \leq t \leq 2x$, and an argument similar to the one in the proof of Proposition 6.1 reveals that

$$\psi(t + \theta t) - \psi(t) - \theta t = O(\sqrt{\theta x} \log x)$$

for $x^{-1} \leq \theta \leq x$ and $x \leq t \leq 2x$, expect for t in a subset of $[x, 2x]$ of Lebesgue measure $o(x)$. Taking $h = \theta t$, then h and θx have the same order of magnitude, which gives some justification of the following conjecture of Montgomery and Vaughan.[38]

Conjecture 6.4. *For every $\varepsilon > 0$, we have*

$$\psi(x + h) - \psi(x) = h + O_{\varepsilon}(\sqrt{h}\, x^{\varepsilon})$$

uniformly for $2 \leq h \leq x$ as $x \to \infty$.

In other words, it is conceivable that an asymptotic prime number theorem for the interval $(x, x+h]$ holds for small h close to Maier's lower bound, $\log^A x$ (all $A > 0$), but the exact threshold remains a mystery.

[37] For more estimates of similar type, the reader may want to have a look at [14]. We would like to thank Ofir Gorodetsky and 'user 2734364041' for making us aware of the result: this matter, as well as the density hypothesis were discussed on the threads [1] and [2].

[38] [27], Conjecture 13.4.

Appendix

A.1 Primes in long intervals

By the prime number theorem, we have

$$\lim_{x\to\infty}\frac{\psi(x)}{x^\vartheta}=\lim_{x\to\infty}\frac{\pi(x)}{x^\vartheta/\log x}=\begin{cases}0 & \text{if } \vartheta>1,\\ 1 & \text{if } \vartheta=1,\\ \infty & \text{if } \vartheta<1,\end{cases}\tag{A1.1}$$

and also

$$1=\lim_{x\to\infty}\frac{\psi(x+x^\vartheta)}{x+x^\vartheta}=\lim_{x\to\infty}\frac{\psi(x+x^\vartheta)}{x^\vartheta}\cdot\frac{1}{1+x^{1-\vartheta}}=\lim_{x\to\infty}\frac{\psi(x+x^\vartheta)}{x}\cdot\frac{1}{1+x^{\vartheta-1}}.$$

Thus,

$$\lim_{x\to\infty}\frac{\psi(x+x^\vartheta)}{x^\vartheta}=1+\lim_{x\to\infty}x^{1-\vartheta}=\begin{cases}1 & \text{if } \vartheta>1,\\ 2 & \text{if } \vartheta=1,\\ \infty & \text{if } \vartheta<1,\end{cases}\quad\text{and}\tag{A1.2}$$

$$\lim_{x\to\infty}\frac{\psi(x+x^\vartheta)}{x}=1+\lim_{x\to\infty}x^{\vartheta-1}=\begin{cases}\infty & \text{if } \vartheta>1,\\ 2 & \text{if } \vartheta=1,\\ 1 & \text{if } \vartheta<1.\end{cases}\tag{A1.3}$$

In a similar manner, the prime number theorem gives

$$\begin{aligned}1&=\lim_{x\to\infty}\frac{\pi(x+x^\vartheta)}{x+x^\vartheta}\log\bigl(x+x^\vartheta\bigr)\\
&=\vartheta\lim_{x\to\infty}\frac{\pi(x+x^\vartheta)}{x^\vartheta/\log x}\cdot\frac{1}{1+x^{1-\vartheta}}+\lim_{x\to\infty}\underbrace{\frac{\pi(x+x^\vartheta)}{\frac{x+x^\vartheta}{\log(x+x^\vartheta)}}}_{\sim 1}\cdot\underbrace{\frac{\log\bigl(1+x^{1-\vartheta}\bigr)}{\log(x+x^\vartheta)}}_{F_\vartheta(x)} &\text{(A1.4)}\\
&=\lim_{x\to\infty}\frac{\pi(x+x^\vartheta)}{x/\log x}\cdot\frac{1}{1+x^{\vartheta-1}}+\lim_{x\to\infty}\underbrace{\frac{\pi(x+x^\vartheta)}{\frac{x+x^\vartheta}{\log(x+x^\vartheta)}}}_{\sim 1}\cdot\underbrace{\frac{\log\bigl(1+x^{\vartheta-1}\bigr)}{\log(x+x^\vartheta)}}_{G_\vartheta(x)}. &\text{(A1.5)}\end{aligned}$$

An elementary calculation shows that

$$\lim_{x\to\infty}F_\vartheta(x)=\begin{cases}0 & \text{if } \vartheta\ge 1,\\ 1-\vartheta & \text{if } \vartheta<1,\end{cases}\quad\text{and}\quad\lim_{x\to\infty}G_\vartheta(x)=\begin{cases}1-\vartheta^{-1} & \text{if } \vartheta>1,\\ 0 & \text{if } \vartheta\le 1.\end{cases}$$

Hence, Equations (A1.4) and (A1.5) yield

$$\lim_{x\to\infty}\frac{\pi(x+x^\vartheta)}{x^\vartheta/\log x}=\begin{cases}\infty & \text{if } \vartheta=0,\\ \vartheta^{-1}(1-F_\vartheta(\infty))(1+\lim\limits_{x\to\infty}x^{1-\vartheta}) & \text{if } \vartheta\ne 0,\end{cases}=\begin{cases}\vartheta^{-1} & \text{if } \vartheta>1,\\ 2 & \text{if } \vartheta=1,\\ \infty & \text{if } \vartheta<1,\end{cases}\tag{A1.6}$$

$$\lim_{x\to\infty}\frac{\pi(x+x^\vartheta)}{x/\log x}=(1-G_\vartheta(\infty))(1+\lim_{x\to\infty}x^{\vartheta-1})=\begin{cases}\infty & \text{if } \vartheta>1,\\ 2 & \text{if } \vartheta=1,\\ 1 & \text{if } \vartheta<1.\end{cases}\tag{A1.7}$$

In particular, Equations (A1.1)–(A1.3) and (A1.6)–(A1.7) give

$$\lim_{x\to\infty}\frac{\psi(x+x^{\vartheta})-\psi(x)}{x^{\vartheta}}=\begin{cases}1-0 & \text{if } \vartheta>1,\\ 2-1 & \text{if } \vartheta=1,\\ \infty-\infty & \text{if } \vartheta<1,\end{cases}=\begin{cases}1 & \text{if } \vartheta\geq 1,\\ \text{indeterminate} & \text{if } \vartheta<1,\end{cases}$$

$$\lim_{x\to\infty}\frac{\psi(x+x^{\vartheta})-\psi(x)}{x}=\begin{cases}\infty-1 & \text{if } \vartheta>1,\\ 2-1 & \text{if } \vartheta=1,\\ 1-1 & \text{if } \vartheta<1,\end{cases}=\begin{cases}\infty & \text{if } \vartheta>1,\\ 1 & \text{if } \vartheta=1,\\ 0 & \text{if } \vartheta<1,\end{cases}$$

$$\lim_{x\to\infty}\frac{\pi(x+x^{\vartheta})-\pi(x)}{x^{\vartheta}/\log x}=\begin{cases}\vartheta^{-1}-0 & \text{if } \vartheta>1,\\ 2-1 & \text{if } \vartheta=1,\\ \infty-\infty & \text{if } \vartheta<1,\end{cases}=\begin{cases}\vartheta^{-1} & \text{if } \vartheta\geq 1,\\ \text{indeterminate} & \text{if } \vartheta<1,\end{cases}\quad\text{and}$$

$$\lim_{x\to\infty}\frac{\pi(x+x^{\vartheta})-\pi(x)}{x/\log x}=\begin{cases}\infty-1 & \text{if } \vartheta>1,\\ 2-1 & \text{if } \vartheta=1,\\ 1-1 & \text{if } \vartheta<1,\end{cases}=\begin{cases}\infty & \text{if } \vartheta>1,\\ 1 & \text{if } \vartheta=1,\\ 0 & \text{if } \vartheta<1,\end{cases}$$

which implies (59).

A.2 Miscellaneous results

Let $\lambda_1 \leq \lambda_2 \leq \ldots$ be a sequence of real numbers with $\lim\limits_{n\to\infty}\lambda_n = \infty$, and $(c_n)_{n=1}^{\infty}$ any sequence of complex numbers. For real t, define $S_t = \{n \in \mathbb{N}^+ : \lambda_n \leq t\}$; this is then a finite set.

If $X \geq \lambda_1$ and $\varphi \in C^1([\lambda_1, X], \mathbb{C})$, then Abel's summation formula holds true:

$$\sum_{n\in S_X} c_n\varphi(\lambda_n) = -\int_{\lambda_1}^{X}\varphi'(t)\sum_{n\in S_t} c_n\,\mathrm{d}t + \varphi(X)\sum_{n\in S_X} c_n. \tag{A2.1}$$

In particular, if $\lambda_n = \gamma_n$ exhaust the ordinates of the nontrivial zeros of $\zeta(s)$ in the upper half plane, and each $c_n = 1$, then $\sum_{n\in S_t} c_n = N(t)$. A proof of the formula may be found in [18] (Theorem A, p. 18).

Let $\alpha \in \mathbb{C}$ and $0 < \delta \leq \pi$. Then the following version of Stirling's formula is true, which states that

$$\log\Gamma(s+\alpha) = \left(s+\alpha-\frac{1}{2}\right)\log(s) - s + \frac{1}{2}\log(2\pi) + O(|s|^{-1}) \tag{A2.2}$$

uniformly in the angle $|\arg\, s| \leq \pi-\delta$ as $|s| \to \infty$, where the branches of the logarithms are chosen to be real on the positive real axis. This formula appears as stated in [18], but an instructive and modernised derivation may be found in [33] (Theorem 2.3 in the Appendix and its associated exercises).

We always have

$$\sum_{2\leq n\leq x}\frac{1}{n}\begin{cases}=0 & \text{if } x<2,\\ \leq\log x & \text{if } x\geq 1.\end{cases} \tag{A2.3}$$

Proof. It suffices to prove the logarithmic bound assuming $x \geq 2$. Since t^{-1} is a decreasing function of t, we have

$$\sum_{2\leq n\leq x}\frac{1}{n} \leq \int_1^x t^{-1}\mathrm{d}t = \log x.$$

□

For all $\alpha, \beta \geq 0$, we have

$$\pi(\alpha+\beta)-\pi(\alpha) < \beta+1 \tag{A2.4}$$

Proof. Exercise. □

Using the convention that $0^0 = 1$, we have the (Bernoulli) inequality

$$(1+a)^b \leq 1 + ab \tag{A2.5}$$

for all real $a \geq -1$ and $0 \leq b \leq 1$.

Proof. The only nontrivial case is when $a > -1$ and $0 < b < 1$. Take $t = 1 + a$, so that we must prove $t^b - 1 \leq b(t-1)$ for $t > 0$. Both sides are equal to zero if $t = 1$ (where $a = 0$). Now differentiate both sides with respect to t to find $\frac{\partial}{\partial t}[t^b - 1] = bt^{b-1}$ and $\frac{\partial}{\partial t}[b(t-1)] = b$. Since $0 < b < 1$, we have $bt^{b-1} < b \cdot 1^{b-1} = b$ if $t > 1$ and $bt^{b-1} > b$ if $0 < t < 1$. This implies the stated result. □

On the positive axis, we have

$$\int \frac{\log t}{t^m} \mathrm{d}t \equiv \begin{cases} -\dfrac{\log t}{(m-1)t^{m-1}} - \dfrac{1}{(m-1)^2 t^{m-1}} & \text{if } m \neq -1, \\ \dfrac{1}{2}\log^2 t & \text{if } m = 1. \end{cases} \tag{A2.6}$$

In particular, if $T > 0$, then

$$\int_1^T \frac{\log t}{t^m} \mathrm{d}t = \begin{cases} \dfrac{1}{(m-1)^2} - \dfrac{\log T}{(m-1)T^{m-1}} - \dfrac{1}{(m-1)^2 T^{m-1}} & \text{if } m \neq 1, \\ \dfrac{1}{2}\log^2 T & \text{if } m = 1, \end{cases} \tag{A2.7}$$

$$\int_T^\infty \frac{\log t}{t^m} \mathrm{d}t = \frac{\log T}{(m-1)T^{m-1}} + \frac{1}{(m-1)^2 T^{m-1}} \qquad \text{if } m > 1. \tag{A2.8}$$

The end.

References

[1] AfterMath. What consequences would follow from the density hypothesis? https://mathoverflow.net/questions/431567/what-consequences-would-follow-from-the-density-hypothesis, October 2022. Question on MathOverflow.

[2] AfterMath. What heuristic arguments support Montgomery's conjecture for primes in short intervals? https://mathoverflow.net/questions/444951/what-heuristic-arguments-support-montgomerys-conjecture-for-primes-in-short-int, April 2023. Question on MathOverflow.

[3] R. J. Backlund. Sur les zéros de la fonction $\zeta(s)$. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 158:1979–1981, 1914.

[4] R. J. Backlund. Über die Beziehung zwischen Anwachsen und Nullstellen der Zetafunktion. *Öfversigt af Finska Vetenskaps-Societetens Förhandlingar*, 61(9):1–8, 1918–1919.

[5] H. Cramér. Studien über die Nullstellen der Riemannschen Zetafunktion. *Mathematische Zeitschrift*, 4(1/2):104–130, 1919. Also found in [7], 97–123.

[6] H. Cramér. Some theorems concerning prime numbers. *Arkiv för Matematik, Astronomi och Fysik*, 15(5):1–33, 1920. Also found in [7], 138–170.

[7] H. Cramér. *Collected works I.* Springer-Verlag Berlin Heidelberg, 1994.

[8] L. Dalaker. The Riemann hypothesis, the Lindelöf hypothesis and the density hypothesis - consequences and relations. Bachelor's thesis, NTNU, 2020. https://hdl.handle.net/11250/2980237.

[9] H. Davenport. *Multiplicative Number Theory.* Markham Publishing Company, Chicago, IL, 1967. Revised third edition by H. L. Montgomery, Springer-Verlag Berlin Heidelberg 2000.

[10] J. A. de Reyna, R. P. Brent, and J. van de Lune. A note on the real part of the Riemann zeta-function, 2014.

[11] H. M. Edwards. *Riemann's Zeta Function.* Pure and applied mathematics. Academic Press, Inc., 1974. Republication by Dover Publications in Dover books on mathematics, 2001.

[12] R. Fissum. Digit sums and the number of prime factors of the factorial $n! = 1 \cdot 2 \cdots n$. Bachelor's thesis, NTNU, 2020. https://hdl.handle.net/11250/2980249.

[13] D. A. Goldston. On a result of Littlewood concerning prime numbers. *Acta Arithmetica*, 40:263–271, 1982.

[14] D. A. Goldston and H. L. Montgomery. *Pair Correlation of Zeros and Primes in Short Intervals*, pages 183–203. Birkhäuser Boston, Boston, MA, 1987.

[15] G. Halász and P. Turán. On the distribution of roots of Riemann zeta and allied functions, I. *Journal of Number Theory*, 1(1):121–137, 1969.

[16] D. R. Heat-Brown. Gaps between primes, and the pair correlation of zeros of the zeta-function. *Acta Arithmetica*, 41(1):85–99, 1982.

[17] G. Hoheisel. *Primzahlprobleme in der Analysis.* Physikalisch-mathematische Klasse. Sitzungsberichten der preussischen Akademie der Wissenschaften, 1930.

[18] A. E. Ingham. *The Distribution of Prime Numbers.* Number 30 in Cambridge Tracts in Mathematics and Mathematical Physics. Cambridge University Press, 1932. Paperback reissue 1990, from the Cambridge Mathematical Library series.

[19] A. E. Ingham. On the difference between consecutive primes. *The Quarterly Journal of Mathematics*, os-8(1):255–266, 01 1937.

[20] A. Ivić. On sums of large differences between consecutive primes. *Mathematische Annalen*, 241:1–9, 1979.

[21] A. Ivić. *The Riemann Zeta-Function: Theory and Applications*. Dover Books on Mathematics Series. Dover Publications, Incorporated, 2013. Reprint 2020.

[22] H. Iwaniec. *Lectures on the Riemann Zeta Function*, volume 62 of *University Lecture Series*. American Mathematical Society, 2014.

[23] M. Jutila. Zero-density estimates for L-functions. *Acta Arithmetica*, 32:55–62, 1977.

[24] N. M. Korobov. Estimates of trigonometric sums and their applications. *Uspekhi Matematicheskikh Nauk*, 13(4):185–192, 1958.

[25] H. Maier. Primes in short intervals. *Michigan Mathematical Journal*, 32(2):221 – 225, 1985.

[26] H. L. Montgomery. *Topics in Multiplicative Number Theory*, volume 227 of *Lecture Notes in Mathematics*. Springer Berlin, Heidelberg, November 2006.

[27] H. L. Montgomery and R. C. Vaughan. *Multiplicative Number Theory I: Classical Theory*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2006.

[28] D. H. J. Polymath. Variants of the Selberg sieve, and bounded intervals containing many primes, 2014.

[29] B. Riemann. Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse. *Monatsberichte der Königlichen Preussische Akademie des Wissenschaften zu Berlin*, pages 671–680, November 1959. An English translation may be found in [11].

[30] B. Saffari and R. C. Vaughan. On the fractional parts of x/n and related sequences. II. *Annales de l'Institut Fourier*, 27(2):1–30, 1977.

[31] A. Selberg. On the normal density of primes in small intervals, and the difference between consecutive primes. *Archiv for Mathematik og Naturvidenskab*, 47:87–105, 1943. Also found in [32], 160–178.

[32] A. Selberg. *Collected Papers I*. New York: Springer Verlag, 1989.

[33] E. M. Stein and R. Shakarchi. *Complex analysis*, volume II of *Princeton lectures in analysis*. Princeton University Press, 2003.

[34] N. G. Tchudakoff. On zeros of Dirichlet's L-functions. *Recueil mathématique (Moscou)*, 43:591–602, 1936.

[35] E. C. Titchmarsh. *The Theory of the Riemann Zeta-function*. Oxford science publications. Oxford University Press, 1986. Revised second edition by D.R. Heath-Brown, 2007.

[36] I. M. Vinogradov. A new estimate for $\zeta(\frac{1}{2} + it)$. *Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya*, 22(2):161–164, 1958.

[37] H. von Mangoldt. Zur Verteilung der Nullstellen der Riemannschen Funktion $\xi(t)$. *Mathematische Annalen*, 60:1–19, 1905.

[38] A. Vretblad. *Fourier Analysis and Its Applications*, volume 223 of *Graduate Texts in Mathematics*. Springer New York, NY, 2003. Corrected second printing, 2005.

[39] R. G. Wilson. A058303 - OEIS. Decimal expansion of the imaginary part of the first nontrivial zero of the Riemann zeta function. https://oeis.org/A058303, December 2000.

NTNU

Norwegian University of
Science and Technology