Prime Numbers

Gaurish Korpal[1]
gaurish.korpal@niser.ac.in

Summer Internship Project Report

[1] 3^{rd} year Int. MSc. Student, National Institute of Science Education and Research, Jatni (Bhubaneswar, Odisha)

Certificate

Certified that the summer internship project report "Prime Numbers" is the bona fide work of "Gaurish Korpal", 3^{rd} Year Int. MSc. student at National Institute of Science Education and Research, Jatni (Bhubaneswar, Odisha), carried out under my supervision during June 05, 2017 to July 15, 2017.

Place: Chennai
Date: July 15, 2017

Prof. Kotyada Srinivas
Supervisor
Professor H,
The Institute of Mathematical Sciences,
Tharamani, Chennai 600113

Abstract

The fact of being prime (or composite) is just a property of number itself, regardless of the way we write it. In general, the divisibility properties are independent of base system (decimal or binary) or writing system we choose (roman or hindu-arabic). In this report I will scratch the surface of the vast topic known as *distribution of primes*.

Acknowledgements

Foremost, I would like to express my sincere gratitude to my advisor *Prof. Kotyada Srinivas* for his motivation and immense knowledge. I am also thankful to my fellow interns *Sayan Kundu*[1] and *Aditya Kumar Shukla*[2] for the enlightening discussions.
Last but not the least, I would like to thank

- Donald Knuth for TeX
- Michael Spivak for AMS-TeX
- Sebastian Rahtz for TeX Live
- Leslie Lamport for LaTeX
- American Mathematical Society for AMS-LaTeX
- Hàn Thế Thành for pdfTeX
    - Christian Feuersänger & Till Tantau for PGF/Ti*k*Z interpreter
    - Heiko Oberdiek for `hyperref` package
    - Steven B. Segletes for `stackengine` package
    - Alan Jeffrey & Frank Mittelbach for `inputenc` package
    - Axel Sommerfeldt for `subcaption` package
    - David Carlisle for `graphicx` package
    - Javier Bezos for `enumitem` package
    - Hideo Umeki for `geometry` package
    - Philipp Lehman & Joseph Wright for `csquotes` package
    - Peter R. Wilson & Will Robertson for `epigraph` package
    - Sebastian Rahtz for `textcomp` package
    - Walter Schmidt for `gensymb` package
- Philipp Kühl & Daniel Kirsch for Detexify (a tool for searching LaTeX symbols)
- TeX.StackExchange community for helping me out with LaTeX related problems

GeoGebra was used to generate all the PGF/Ti*k*Z codes for this document.

[1]2nd year B.Sc.(Hons) Mathematics and Computing, IMA-Bhubaneswar
[2]3rd year B.Sc.(Hons) Mathematics, BHU-Varanasi

Contents

Introduction

> *"Prime numbers are what is left when you have taken all the patterns away. I think prime numbers are like life. They are very logical but you could never work out the rules, even if you spent all your time thinking about them."*
>
> — Christopher Boone*, narrator in Mark Haddon's The Curious Incident of the Dog in the Night Time*

Prime numbers are the integers bigger than 1 which are only divisible by 1 and themselves. We encounter these numbers as soon as we start analysing the integers, in form of *fundamental theorem of arithmetic*, which says that prime numbers are the "building blocks" of integers [21]. Moreover, in the past reports we saw that these numbers play an important role in solving Diophantine equations [24, §1.5, 1.6], generalising the concept of factorisation [27] and extending the concept of number system [28]. An exposition on further extension of these three ideas can be seen in Richard Taylor's lecture [21].

In this report I intend to introduce the basic tools involed in study of distribution of prime numbers. These tools will be developed with the help of complex function theory. Effectively we will be scratching the surface of the multiplicative aspects of analytic number theory. In analytic number theory one looks for good approximations. For the sort of quantity that one estimates in analytic number theory, one does not expect an exact formula to exist, except perhaps one of a rather artificial and unilluminating kind [9].

The first good guess about the asymptotic estimate for the primes emerged at the beginning of the nineteenth century, by Carl Friedrich Gauss at 16 years of age, that "the density of primes around x is about $1/\log(x)$". This statement was proved by the end of nineteenth century, by showing that Riemann zeta function $\zeta(s)$, has no zeros near the line $\mathrm{Re}(s) = 1$. I will discuss this proof in the first chapter of this report. Later in 1930s, Harald Cramér gave a probabilistic way of interpreting Gauss' prediction. The Gauss-Cramér model provides a nice way to think about distribution questions concerning the prime numbers, but it does not give proofs. For a discussion on distribution of primes based on this model, refer to Andrew Granville's article [9].

We can obtain two types of results about the distribution of primes – results of a *qualitative* and *quantitative* kind [2]. The second chapter is devoted to a result of qualitative nature. Long time ago it had been asserted that every arithmetic progression $a, a+m, a+2m, \ldots$, in which a and m has no common factor, includes infinitely many primes. The first proof was that of Peter Gustav Lejeune Dirichlet published in 1837 and to do this he introduced the L-functions which bear his name. But I will discuss that proof with a slight modification introduced by Charles Jean de la Vallée Poussin in 1896. One must note that the use of Dirichlet L-functions extends beyond the proof of this theorem [5, §16.6].

Unlike the lemma-theorem-proof style followed in my past reports, in this report I have tried to follow Harold Davenport's prose style [3]. My attempt is to let the reader appreciate how random appearing tricks or ideas come together to build this beautiful theory of prime numbers. It's similar to the idea that while reading a nicely written story, the reader is able to appreciate the ingenuity of ideas. Some steps in proofs may appear to be presented without any proper logical motivation behind them, and that's why it took great imagination to even conjecture the statements whose proof we try to understand in this report.

Chapter 1

Prime Number Theorem

> *"It should be pointed out that, among the most celebrated and cultivated of the Greeks, Eratosthenes was considered an extraordinary man who hurled the javelin, wrote poems, defeated the great runners, and solved astronomical problems. Various works of his have come down to posterity. He presented King Ptolemy III of Egypt with a table on which the prime numbers were etched on a metal plate, with those numbers with multiples marked with a small hole. And so they give the name of Eratosthenes' sieve to the process the wise astronomer used to draw up his table."*
>
> — Malba Tahan, *The Man who counted*

It's easy to prove that there are infintely many prime numbers [7, §1.4] and it's apparent from a glance at a list of primes that the sequence is rather irregular, but the list seems to thin out at a regular rate, even if it never ends [13]. Despite not having a good *exact formula* for the sequence of primes, we do have a fairly good *inexact formula* [20], in the form of *prime number theorem.*

1.1 Entire Functions

An *entire function* (or *integral function*) is a complex-valued function that is holomorphic at all finite points over the whole complex plane. We call an entire function $f(z)$ to be of finite order if there exists a number α such that

$$f(z) = O\left(e^{|z|^\alpha}\right) \quad \text{as } |z| \to \infty$$

We must have $\alpha > 0$, excluding the case when $f(z)$ is just a constant. The lower bound of the numbers α with this property is called the *order of* $f(z)$ [3, §11].

1.1.1 Finite Order without Zeros

Consider an entire function $f(z)$ of finite order with no zeros. Then the function $g(z) = \log(f(z))$ can be defined as to be single valued, and is itself an entire function. It satisfies

$$\mathrm{Re}(g(z)) = \log|f(z)| < 2r^\alpha \tag{1.1}$$

on any large circle $|z| = r$. Since $g(z)$ is an analytic function, we can write

$$g(z) = \sum_{n=0}^{\infty}(a_n + ib_n)z^n$$

and then for $z = re^{i\theta}$

$$\mathrm{Re}(g(z)) = \sum_{n=0}^{\infty}\left(a_n\cos(n\theta) - b_n\sin(n\theta)\right)r^n$$

Now we wish to obtain an inequality similar to Cauchy's inequality but involving upper bound of $\mathrm{Re}(g(z))$ [1, §2.53]. Since the series for $\mathrm{Re}(g(z))$ converges uniformly with respect to θ, we may multiply by $\cos(n\theta)$ or $\sin(n\theta)$ and integrate term by term to obtain

$$\int_0^{2\pi} \mathrm{Re}(g(z))\cos(n\theta)d\theta = \pi a_n r^n, \qquad \int_0^{2\pi} \mathrm{Re}(g(z))\sin(n\theta)d\theta = -\pi b_n r^n$$

for $n > 0$, while

$$\int_0^{2\pi} \mathrm{Re}(g(z))d\theta = 2\pi a_0 \tag{1.2}$$

Hence for $n > 0$

$$(a_n + ib_n)r^n = \frac{1}{\pi}\int_0^{2\pi} Re(g(z))e^{-in\theta}d\theta$$

But we know that [1, §2.31]

$$|(a_n + ib_n)r^n| = |a_n + ib_n|\, r^n \leq \frac{1}{\pi}\int_0^{2\pi} |Re(g(z))|\, d\theta$$

because $\left|e^{ik}\right| = 1$ for all $k \in \mathbb{R}$. Also, $|a_n| \leq |a_n + ib_n| = \sqrt{a_n^2 + b_n^2}$, therefore

$$|a_n|r^n \leq \frac{1}{\pi}\int_0^{2\pi} |Re(g(z))|\, d\theta$$

Adding (1.2) to this, we get

$$|a_n|r^n + 2a_0 \leq \frac{1}{\pi}\int_0^{2\pi} \Big(|Re(g(z))| + \mathrm{Re}(g(z))\Big)d\theta$$

Without loss of generality, we can take $g(0) = 0$ i.e. $a_0 = 0$, and then using (1.1) we get

$$|a_n|r^n \leq \frac{1}{\pi}\int_0^{2\pi} (2r^\alpha + 2r^\alpha)\, d\theta = 8r^\alpha$$

It follows, on making $r \to \infty$, that $a_n = 0$ if $n > \alpha$, and similarly for b_n. This proves that $g(z)$ is a polynomial, where $f(z) = e^{g(z)}$.

> An entire function of finite order with no zeros is necessarily of the form $e^{g(z)}$, where $g(z)$ is a polynomial, and its order is simply the order of $g(z)$ and so is an integer.

1.1.2 Finite Order with Zeros

Consider an entire function $f(z)$ of order ρ with zeros $z_1, z_2, \ldots, z_n$ in $|z| < R$ (multiple zeros being repeated as appropriate) and no zeros on $|z| = R$. For $\alpha > \rho$, we have

$$\log\left|f\left(Re^{i\theta}\right)\right| < R^\alpha$$

for all sufficiently large R. Then we can write

$$\frac{1}{2\pi}\int_0^{2\pi} \log\left|f\left(Re^{i\theta}\right)\right| d\theta < \frac{1}{2\pi}\int_0^{2\pi} R^\alpha d\theta = R^\alpha \tag{1.3}$$

Let $n(r)$ be the number of zeros of $f(z)$ in $|z| \leq r$ and $f(0) \neq 0$, then using Jensen's formula for analytic functions [1, §3.61]

$$\frac{1}{2\pi}\int_0^{2\pi} \log\left|f\left(Re^{i\theta}\right)\right| d\theta - \log|f(0)| = \log\left(\frac{R^n}{|z_1|\ldots|z_n|}\right) = \int_0^R \frac{n(r)}{r}dr$$

we can rewrite (1.3) as

$$\int_0^R \frac{n(r)}{r} dr + \log|f(0)| < R^\alpha$$

$$\Rightarrow \int_0^R \frac{n(r)}{r} dr < R^\alpha - \log|f(0)| < 2R^\alpha \tag{1.4}$$

Moreover, since $n(r)$ is non-decreasing

$$\int_R^{2R} \frac{n(r)}{r} dr \geq n(R) \int_R^{2R} \frac{1}{r} dr = n(R)\log(2)$$

Hence using (1.4), we conclude that

$$\boxed{n(R) = O(R^\alpha)}$$

Now let $r_k = |z_k|$ for $k = 1, 2, \ldots$, where z_k is a zero[1] of $f(z)$, then for $R = r_k$ and $\beta > \alpha > \rho$ we get

$$k < Ar_k^\alpha \quad \Longrightarrow \frac{1}{r_k^\beta} < \frac{A'}{n^{\frac{\beta}{\alpha}}}$$

Therefore,

$$\sum_{k=1}^{\infty} \frac{1}{r_k^\beta} < \sum_{k=1}^{\infty} \frac{A'}{n^{\frac{\beta}{\alpha}}} < \infty$$

since $\sum n^{-\ell}$ converges for $\ell > 1$.

> If $r_1, r_2, \ldots$ are the moduli of the zeros of an entire function $f(z)$ of order ρ, with $f(0) \neq 0$, then the series $\sum r_n^{-\beta}$ converges if $\beta > \rho$.

The lower bound of the positive numbers β for which $\sum r_n^{-\beta}$ is convergent is called the *exponent of cenvergence of the zeros*, and is denoted by ρ_1. What we proved above is that $\rho_1 \leq \rho$, But we may have $\rho_1 < \rho$; for example, if $f(z) = e^z$ then $\rho = 1$ but $\rho_1 = 0$ since there are no zeros. Moreover, any function with finite number of zeros has $\rho_1 = 0$. Therefore, $\rho_1 > 0$ implies that there are infinitely many zeros [1, §8.22].

1.1.3 Order 1 with Zeros

Consider an entire function $f(z)$ of order 1 with with zeros $z_1, z_2, \ldots$ (multiple zeros being repeated as appropriate) and $f(0) \neq 0$. We can then assert that $\sum r_n^{-1-\varepsilon}$ converges for any $\varepsilon > 0$, and in particular $\sum r_n^{-2}$ converges. Hence the product

$$P(z) = \prod_{n=1}^{\infty} \left(1 - \frac{z}{z_n}\right) e^{\frac{z}{z_n}}$$

converges absolutely for all z, and converges uniformly in any bounded domain not containing the points z_n [1, §1.44]. So, $P(z)$ is an entire function with zeros (of the appropriate multiplicities) at $z_1, z_2, \ldots$. If we put

$$f(z) = P(z)F(z) \tag{1.5}$$

then $F(z)$ is an integral function without zeros.

[1] An entire function which is not a polynomial may have infinitely many zeros z_n.

Since $\sum r_n^{-2}$ converges, the total length of all the intervals $\left(r_n - r_n^{-2}, r_n + r_n^{-2}\right)$ on the real line is finite, and consequently there exist arbitrarily large values of R with the property that

$$|R - r_n| > \frac{1}{r_n^2} \quad \text{for all } n \tag{1.6}$$

Put $P(z) = P_1(z)P_2(z)P_3(z)$, where these are the sub-products extended over the following sets of n:

$$\begin{aligned} P_1 &: \quad |z_n| < \frac{R}{2} \\ P_2 &: \quad \frac{R}{2} \leq |z_n| \leq 2R \\ P_3 &: \quad |z_n| > 2R \end{aligned}$$

For the factors of P_1 we have, on $|z| = R$

$$\left|\left(1 - \frac{z}{z_n}\right) e^{\frac{z}{z_n}}\right| \geq \left(\left|\frac{z}{z_n}\right| - 1\right) e^{-\frac{|z|}{|z_n|}} > \left(\frac{R}{R/2} - 1\right) e^{-\frac{R}{r_n}} = e^{\frac{-R}{r_n}}$$

and since

$$\sum_{r_n < \frac{R}{2}} \frac{1}{r_n} < \left(\frac{R}{2}\right)^{\varepsilon} \sum_{n=1}^{\infty} \frac{1}{r_n^{1+\varepsilon}}$$

it follows that

$$|P_1(z)| = \left|\prod_{|z_n| < \frac{R}{2}} \left(1 - \frac{z}{z_n}\right) e^{\frac{z}{z_n}}\right| > \prod_{|z_n| < \frac{R}{2}} e^{\frac{-R}{r_n}} = \exp\left(-R \sum_{r_n < \frac{R}{2}} \frac{1}{r_n}\right) > \exp\left(-\frac{R^{1+\varepsilon}}{2^{\varepsilon}} \sum_{n=1}^{\infty} \frac{1}{r_n^{1+\varepsilon}}\right)$$

$$\Rightarrow |P_1(z)| > \exp\left(-R^{1+2\varepsilon}\right) \tag{1.7}$$

For the factors of P_2 we have, on $|z| = R$

$$\left|\left(1 - \frac{z}{z_n}\right) e^{\frac{z}{z_n}}\right| > \left|\frac{z - z_n}{z_n}\right| e^{-\frac{|z|}{|z_n|}} > \frac{|z - z_n|}{2R} e^{-2} > \frac{\big||z| - |z_n|\big|}{2R} e^{-2} > \frac{e^{-2}}{2r_n^2 R} > \frac{C}{R^3}$$

where C is a positive constant as per (1.6). And by the result proved in subsection 1.1.2, $n(R) = O(R^{1+\varepsilon})$, we know that the number of factors of P_2 is less than $R^{1+\varepsilon}$.

$$\Rightarrow |P_2(z)| > \left(\frac{C}{R^3}\right)^{R^{1+\varepsilon}} > \exp(-R^{1+2\varepsilon}) \tag{1.8}$$

For the factors of P_3 we have, on $|z| = R$

$$\left|\left(1 - \frac{z}{z_n}\right) e^{\frac{z}{z_n}}\right| \geq \left(1 - \left|\frac{z}{z_n}\right|\right) e^{-\frac{|z|}{|z_n|}} > \left(1 - \frac{1}{2}\right) e^{-\frac{R}{r_n}} > e^{-c\left(\frac{R}{r_n}\right)^2}$$

for some positive constant c, and since

$$\sum_{r_n > 2R} \frac{1}{r_n^2} < \frac{1}{(2R)^{1-\varepsilon}} \sum_{n=1}^{\infty} \frac{1}{r_n^{1+\varepsilon}}$$

it follows that

$$|P_3(z)| = \left|\prod_{|z_n| > 2R} \left(1 - \frac{z}{z_n}\right) e^{\frac{z}{z_n}}\right| > \prod_{|z_n| > 2R} e^{-\left(\frac{R}{r_n}\right)^2} = \exp\left(\sum_{r_n > 2R} \frac{-R^2}{r_n^2}\right) > \exp\left(\frac{-R^{1+\varepsilon}}{2^{1-\varepsilon}} \sum_{n=1}^{\infty} \frac{1}{r_n^{1+\varepsilon}}\right)$$

$$\Rightarrow |P_3(z)| > \exp\left(-R^{1+2\varepsilon}\right) \tag{1.9}$$

From (1.7), (1.8) and (1.9) it follows that

$$|P(z)| > \exp\left(-R^{1+3\varepsilon}\right)$$

Hence from (1.5) and $f(z) = O\left(\exp(|z|^{1+\varepsilon})\right)$, it follows that

$$|F(z)| < \exp\left(R^{1+4\varepsilon}\right)$$

As seen in subsection 1.1.1, this implies that $F(z) = e^{g(z)}$ where $g(z)$ is a polynomial of degree at most 1. Finally we have

$$\boxed{f(z) = e^{A+Bz} \prod_{n=1}^{\infty} \left(1 - \frac{z}{z_n}\right) e^{\frac{z}{z_n}}}$$

where A and B are constants.

We know that the series $\sum r_n^{-1-\varepsilon}$ converges for any $\varepsilon > 0$. Then series $\sum r_n^{-1}$ may or may not converge, but if it does then $f(z)$ satisfies the inequality

$$|f(z)| < e^{C|z|}$$

for some constant C, because for all $w \in \mathbb{C}$ we have

$$|1 - w| \leq 1 + |w| < e^{|w|} \quad \text{and} \quad |e^w| = \left|e^{\mathrm{Re}(w)}\right| < e^{\sqrt{\mathrm{Re}(w)^2 + \mathrm{Im}(w)^2}} = e^{|w|}$$

$$\Rightarrow |(1 - w)e^w| \leq e^{2|w|}$$

An entire function $f(z)$ of order 1 necessarily has the form

$$f(z) = e^{A+Bz} \prod_{n=1}^{\infty} \left(1 - \frac{z}{z_n}\right) e^{\frac{z}{z_n}}$$

If $r_n = |z_n|$, where z_n are the zeros of $f(z)$, then $\sum r_n^{-1-\varepsilon}$ converges for any $\varepsilon > 0$. If $\sum r_n^{-1}$ converges, then $f(z)$ satisfies $|f(z)| < e^{C|z|}$.

1.2 Gamma Function

The gamma-function

$$\boxed{\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt}$$

is a uniformly convergent integral over any finite region throughout which $\mathrm{Re}(z) > 0$ [1, §1.51] and so a continuous function for $\mathrm{Re}(z) > 0$ [1, §1.52]. It is an analytic function, regular for $\mathrm{Re}(z) > 0$ [1, §2.85].

1.2.1 Analytic Continuation

Consider the function

$$f(z) = \int_C e^{-w} (-w)^{z-1} dw$$

where C consists of the real axis from ∞ to δ, the circle $|w| = \delta$ described in the positive direction, and the real axis from δ to ∞ again. Hence $f(z)$ is regular[2] for all finite values of z and the function

$$g(z) = \frac{1}{2}if(z)\text{cosec}(\pi z) = \frac{i}{2\sin(\pi z)}\int_C e^{-w}(-w)^{z-1}dw$$

is regular for all values of z except at the poles $z = 0, -1, -2, \ldots$ [1, §4.41]. We can therefore take $g(z)$ as the analytic continuation of $\Gamma(z)$ over $\mathbb{C}\setminus\{0,-1,-2,\ldots\}$. All the gamma-function formulae for real z [1, §1.86] can now be extended to general complex values of z. For example, we get the functional equations:

Integration by Parts:	$\Gamma(z+1) = z\Gamma(z)$	$\forall z \in \mathbb{C}\setminus -\mathbb{N}$
Euler's Reflection Formula:	$\Gamma(z)\Gamma(1-z) = \pi\text{cosec}(\pi z)$	$\forall z \in \mathbb{C}\setminus \mathbb{Z}$
Legendre's Duplication Formula:	$\Gamma(z)\Gamma\left(z+\frac{1}{2}\right) = 2^{1-2z}\sqrt{\pi}\Gamma(2z)$	$\forall z \in \mathbb{C}\setminus -\mathbb{N}/2$
Weierstrass' Formula:	$\Gamma(z) = \frac{e^{-\gamma z}}{z}\prod_{n=1}^{\infty}\left(1+\frac{z}{n}\right)^{-1}e^{\frac{z}{n}}$	$\forall z \in \mathbb{C}\setminus -\mathbb{N}$

where $\mathbb{N} = \{0,1,2,3,\ldots\}$ and γ is the *Euler-Mascheroni constant* given by

$$\gamma = -\log\prod_{n=1}^{\infty}\left(1+\frac{1}{n}\right)e^{\frac{-1}{n}} = \lim_{N\to\infty}\left(1+\frac{1}{2}+\ldots+\frac{1}{N}-\log N\right) \approx 0.57721$$

1.2.2 Stirling's Formula

When we try to observe the asymptotic behaviour of $\Gamma(x)$ as $x \to \infty$ we get the Stirlinig's theorem [1, §1.87]

$$\Gamma(x) = x^{x-\frac{1}{2}}e^{-x}\sqrt{2\pi}\,(1+o(1))$$

where $f(x) = o(g(x))$ means that $\lim\limits_{x\to\infty}\dfrac{f(x)}{g(x)} = 0$. By taking logarithm of the Weierstrass' formula

$$\log(\Gamma(z)) = \sum_{n=1}^{\infty}\left(\frac{z}{n} - \log\left(1+\frac{z}{n}\right)\right) - \gamma z - \log(z)$$

with each logarithm having its principal value. Then after appropriate manipulations we get [1, §4.42]

$$\boxed{\log(\Gamma(z)) = \left(z-\frac{1}{2}\right)\log(z) - z + \frac{1}{2}\log(2\pi) + O\left(\frac{1}{|z|}\right)}$$

for $-\pi+\delta \leq \arg(z) \leq \pi-\delta$, which is the extension of Stirling's formula to complex values of z. Therefore, for any constant a, we can write

$$\log(\Gamma(z+a)) = \left(z+a-\frac{1}{2}\right)\log(z) - z + \frac{1}{2}\log(2\pi) + O\left(\frac{1}{|z|}\right)$$

as $|z| \to \infty$, uniformly for $-\pi+\delta \leq \arg(z) \leq \pi-\delta$. And for any fixed value of x, as $y \to \pm\infty$

$$|\Gamma(x+iy)| \sim e^{-\frac{\pi|y|}{2}}|y|^{x-\frac{1}{2}}\sqrt{2\pi}$$

[2] The many-valued function $(-w)^{z-1} = e^{(z-1)\log(-w)}$ is made definite by taking $\log(-w)$ to be real at $w = -\delta$.

1.3 Zeta Function

1.3.1 Euler Zeta Function

For a given real number $k > 1$, consider the series

$$\sum_{n=1}^{\infty} \frac{1}{n^k} = 1 + \frac{1}{2^k} + \frac{1}{3^k} + \dots$$

Note that n^{-k} decreases steadily as n increases but is always positive and

$$\sum_{n=1}^{\infty} \frac{1}{n^k} < 1 + \lim_{\xi \to \infty} \int_1^{\xi} \frac{1}{x^k} dx$$

So the series $\sum_{n=1}^{\infty} n^{-k}$ converges because

$$\lim_{\xi \to \infty} \int_1^{\xi} \frac{1}{x^k} dx = \frac{1}{k-1}$$

is finite. Hence we conclude that the series $\sum_{n=1}^{\infty} n^{-k}$ is uniformly convergent for $a \le k \le b$, if $1 < a < b$.

Now let $\boxed{s = \sigma + it}$ be a complex number, then $\boxed{\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}}$ converges for $\sigma > 1$ since

$$\sum_{n=1}^{\infty} \frac{1}{n^s} < \sum_{n=1}^{\infty} \left|n^{-s}\right| = \sum_{n=1}^{\infty} \left|e^{-s \log(n)}\right| = \sum_{n=1}^{\infty} \left|\frac{1}{n^{\sigma}}\right|$$

and is called the *zeta function*. Also, the series $\sum_{n=1}^{\infty} n^{-s}$ is uniformly convergent throughout any finite region in which $\sigma \ge a > 1$. Therefore, the function $\zeta(s)$ is continuous at all points of the region $\sigma > 1$.

Since $\zeta(s)$ is absolutely convergent for $\sigma > 1$, by Dirichlet multiplication [1, §1.66(xi)] we conclude that

$$(\zeta(s))^2 = \sum_{a=1}^{\infty} \frac{1}{a^s} \sum_{b=1}^{\infty} \frac{1}{b^s} = \sum_{c=1}^{\infty} \frac{1}{c^s} \sum_{ab=c} 1 = \sum_{n=1}^{\infty} \frac{\mathrm{d}(n)}{n^s}, \quad \sigma > 1$$

where $\mathrm{d}(n)$ denotes the number of divisors of n. More generally

$$\boxed{(\zeta(s))^k = \sum_{n=1}^{\infty} \frac{\mathrm{d}_k(n)}{n^s}, \quad \sigma > 1}$$

where $k = 1, 2, \dots$ and $\mathrm{d}_k(n)$ is the number of ways of expressing n as a product of k factors.

1.3.2 Euler Product

The product

$$\prod_p \left(1 - \frac{1}{p^s}\right)$$

where p runs through the primes $2, 3, 5, \dots$ is uniformly convergent in any finite region throughout which $\sigma > 1$ because same thing is true of the series $\sum_p |p^{-s}|$, which consists of some of the terms of the series $\sum_{n=1}^{\infty} |n^{-s}|$ studied in subsection 1.3.1 [1, §1.44].

The value of the product is $\frac{1}{\zeta(s)}$. Since for

$$\left(1-\frac{1}{2^s}\right)\zeta(s) = 1+\frac{1}{3^s}+\frac{1}{5^s}+\dots$$

all terms containing the factor 2 being omitted on the right[3]. Next

$$\left(1-\frac{1}{2^s}\right)\left(1-\frac{1}{3^s}\right)\zeta(s) = 1+\frac{1}{5^s}+\frac{1}{7^s}+\dots$$

all terms containing the factors 2 or 3 being omitted. So generally, if p_n is the n^{th} prime,

$$\left(1-\frac{1}{2^s}\right)\left(1-\frac{1}{3^s}\right)\cdots\left(1-\frac{1}{p_n^s}\right)\zeta(s) = 1+\frac{1}{\ell^s}+\dots$$

where all numbers containing the factors $2, 3, \dots, p_n$ are omitted. Since all numbers upto p_n are of this form,

$$\left|\left(1-\frac{1}{2^s}\right)\left(1-\frac{1}{3^s}\right)\cdots\left(1-\frac{1}{p_n^s}\right)\zeta(s)-1\right| \le \left|\frac{1}{p_n+1}\right|^s + \left|\frac{1}{p_n+2}\right|^s + \dots$$

which tends to 0 as $p_n \to \infty$. Hence

$$\lim_{n\to\infty}\left(1-\frac{1}{2^s}\right)\left(1-\frac{1}{3^s}\right)\cdots\left(1-\frac{1}{p_n^s}\right)\zeta(s) = 1$$

hence proving the result[4] that

$$\boxed{\zeta(s) = \prod_p \left(1-\frac{1}{p^s}\right)^{-1}, \quad \sigma > 1}$$

This was proved by Leonhard Euler[5] in 1737. Since a convergent infinite product of non-zero factors is non-zero [4, §1.1], we conclude that:

$$\boxed{\zeta(s) \text{ has no zeros for } \sigma > 1.}$$

1.3.3 Analytic Properties of Euler Zeta Function

Since $u_n(s) = n^{-s}$ is analytic inside any region $D \subset \mathbb{C}$ with $\sigma > 1$, and the series $\sum_{n=1}^{\infty} u_n(s) = \sum_{n=1}^{\infty} n^{-s}$ is uniformly convergent throughout every region $D' \subset D$. The function $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ is an analytic function, regular[6] for $\sigma > 1$, and all its derivatives can be calculated by term-by-term differentiation [1, §2.8]. Therefore, in general

$$\boxed{\zeta^{(k)}(s) = (-1)^k \sum_{n=2}^{\infty} \frac{(\log(n))^k}{n^s}, \quad \sigma > 1}$$

[3] Notice the resemblance to the sieve of Eratosthenes.

[4] This identity is an analytic equivalent of the fundamental theorem of arithmetic [3, pp. 2].

[5] "Variae observationes circa series infinitas." St. Petersburg Acad., 1737. http://eulerarchive.maa.org/pages/E072.html

[6] A one-valued analytic function is regular at any point which is interior to one of the circles used in continuation from the original element [1, §4.2].

1.3.4 Riemann Zeta Function

Consider the following term-by-term integration over an infinite range

$$\begin{aligned}\int_0^\infty \frac{x^{s-1}}{e^x-1}dx &= \int_0^\infty \left(\sum_{n=1}^\infty x^{s-1}e^{nx}\right)dx \\ &= \sum_{n=1}^\infty \int_0^\infty x^{s-1}e^{-nx}dx \\ &= \sum_{n=1}^\infty n^{-s}\int_0^\infty y^{s-1}e^{-y}dy \\ &= \sum_{n=1}^\infty n^{-s}\Gamma(s)\end{aligned}$$

Then since $\sum_{n=1}^\infty n^{-s}$ is convergent for $\sigma > 1$, we have [1, §1.78]

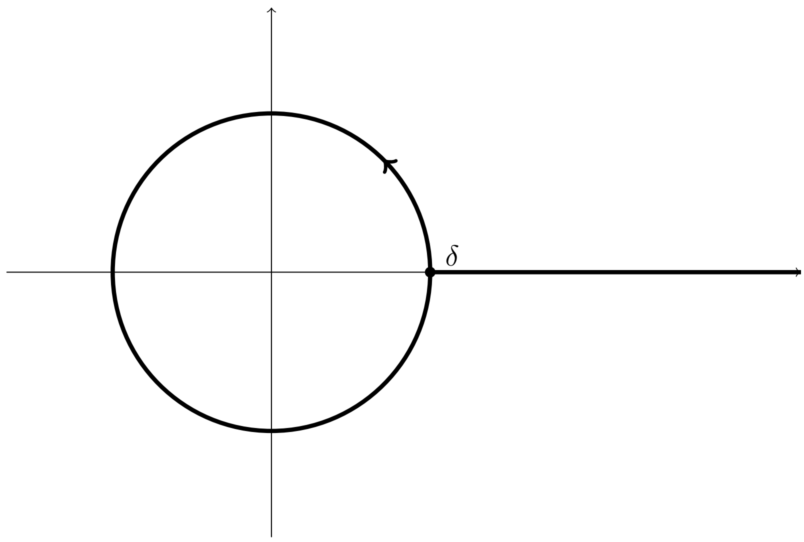
$$\boxed{\zeta(s) = \frac{1}{\Gamma(s)}\int_0^\infty \frac{x^{s-1}}{e^x-1}dx, \quad \sigma > 1}$$

We can use this formula to continue $\zeta(s)$ across $\sigma = 1$, in the same way that we continue $\Gamma(s)$ across $\sigma = 0$. In fact we can prove it precisely the same way as for gamma-function [1, §4.43]. Consider the complex integral

$$I(s) = \int_C \frac{w^{s-1}}{e^w-1}dw$$

where the contour C starts at infinity on the positive real axis, encircles the origin once in the positive direction, excluding the points $\pm 2i\pi, \pm 4i\pi, \ldots$ and returns to positive infinity[7]. Here the many-valued function $(w)^{s-1} = e^{(s-1)\log(w)}$ is made definite by taking $\log(w)$ to be real at the beginning of the contour; thus $\mathrm{Im}(\log(w))$ varies from 0 to 2π round the contour [4, §2.4]. If s is any integer, the integrand in $I(s)$ is one-valued, and $I(s)$ can be evaluated by the theorem of residues [1, §3.11].

We can take C to consist of the real axis from ∞ to δ $(0 < \delta < 2\pi)$, the circle $|w| = \delta$, and the real axis from δ to ∞.



[7] The only difference between this contour and the one used for gamma-function is that C must now exclude all the poles of $1/(e^w - 1)$ other than $w = 0$, i.e. the points $w = \pm 2i\pi, \pm 4i\pi, \ldots$

We note that

$$|w^{s-1}| = e^{(\sigma-1)\log|w| - t\arg(w)} \le |w|^{\sigma-1}e^{2\pi|t|}$$

$$|e^w - 1| = \left|w + \frac{w^2}{2} + \frac{w^3}{6} + \ldots\right| > A|w|, \quad \text{for some } A > 0$$

Hence on the circle $|w| = \delta$, for fixed s, we have

$$\left|\int_{|z|=\delta} \frac{w^{s-1}}{e^w - 1}dw\right| \le \frac{\delta^{\sigma-1}e^{2\pi|t|}}{A\delta} \cdot 2\pi\delta \to 0$$

as $\delta \to 0$ and $\sigma > 1$. Thus on letting $\delta \to 0$, if $\sigma > 1$, we get

$$\begin{aligned}
I(s) &= -\int_0^\infty \frac{x^{s-1}}{e^x - 1}dx + \int_0^\infty \frac{\left(xe^{2\pi i}\right)^{s-1}}{e^x - 1}dx \\
&= -\zeta(s)\Gamma(s) + e^{2\pi i(s-1)}\zeta(s)\Gamma(s) && \text{(using the boxed formula on previous page)} \\
&= \left(e^{2\pi i(s-1)} - 1\right)\zeta(s)\Gamma(s) \\
&= \left(e^{2\pi i(s-1)} - 1\right)\zeta(s)\frac{\pi}{\sin(\pi s)\Gamma(1-s)} && \text{(using Euler's Reflection Formula)} \\
&= \frac{e^{2\pi i s} - e^{2\pi i}}{e^{2\pi i}} \frac{\zeta(s)}{\Gamma(1-s)} \frac{2\pi i}{e^{i\pi s} - e^{-i\pi s}} \\
&= \frac{2\pi i e^{i\pi s}}{\Gamma(1-s)}\zeta(s) && \text{(using } e^{2\pi i} = 1\text{)}
\end{aligned}$$

Hence we get

$$\boxed{\zeta(s) = \frac{e^{-i\pi s}\Gamma(1-s)}{2\pi i}\int_C \frac{w^{s-1}}{e^w - 1}dw}$$

This formula has been proved for $\sigma > 1$. The integral $I(s)$, however, is uniformly convergent in any finite region of the complex plane, and so defines an integral function of s. Hence the formula provides the analytic continuation of $\zeta(s)$ over the whole complex plane. The only possible singularities are the poles of $\Gamma(1-s)$, i.e. $s = 1, 2, 3, 4, \ldots$. From the subsection 1.3.3 we know that $\zeta(s)$ is regular at $s = 2, 3, \ldots$, and in fact it follows at once from Cauchy's theorem [1, §2.33] that $I(s)$ vanishes at these points. Hence the only possible singularity is a simple pole at $s = 1$. For $s = 1$ the contour integral is equal to

$$I(1) = \int_C \frac{dw}{e^w - 1} = 2\pi i$$

by the residue theorem [1, §3.11], and $\Gamma(1-s)$ has a pole (as seen in section 1.2). Hence $s = 1$ is the only pole. Hence the above formula is analytic continuation of $\zeta(s)$ over $\mathbb{C} \setminus \{1\}$.

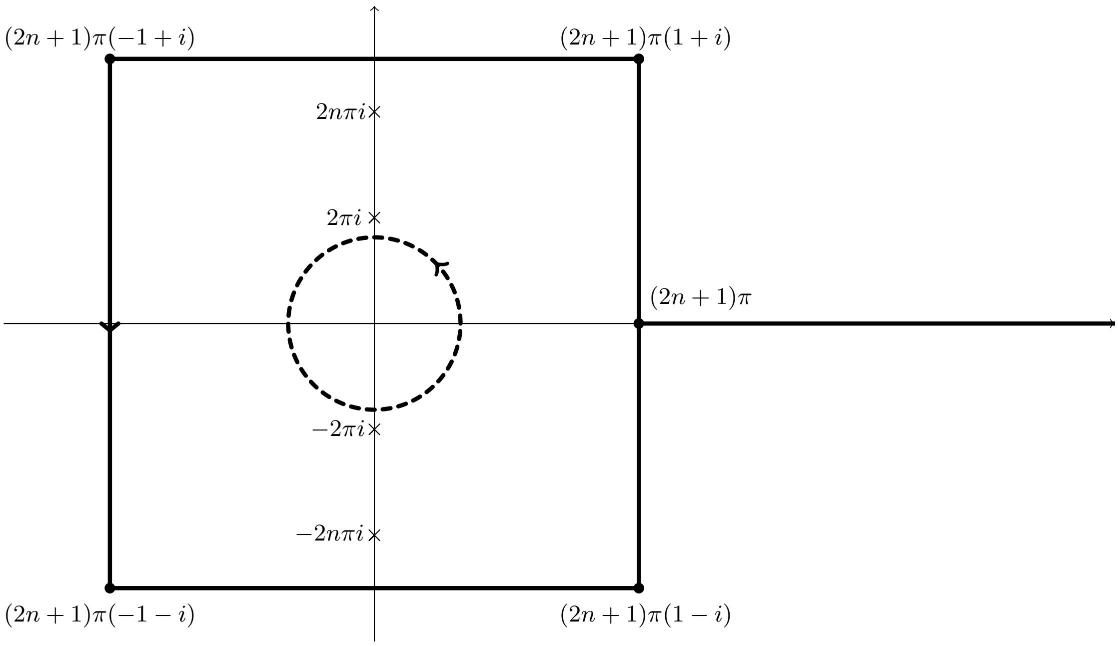
To deduce the *functional equation* from the analytic continuation formula, take the integral of $\dfrac{w^{s-1}}{e^w - 1}$ along the contour C_n, $n = 1, 2, 3, \ldots$, consisting of the positive real axis from infinity to $(2n+1)\pi$, then round the square with the corners $(2n+1)\pi(\pm 1 \pm i)$, and then back to infinity along the positive real axis. Between the contours C and C_n the integrand has poles at the points $\pm 2i\pi, \ldots, \pm 2in\pi$.

The residues at $2m\pi i$ and $-2m\pi i$ are taken together[8]

$$\lim_{z\to 2m\pi i}(z - 2m\pi i)\,\tfrac{(2m\pi i)^{s-1}}{e^{2m\pi i}-1} + \lim_{z\to -2m\pi i}(z + 2m\pi i)\,\tfrac{(-2m\pi i)^{s-1}}{e^{-2m\pi i}-1} = \left(2m\pi e^{\frac{i\pi}{2}}\right)^{s-1} + \left(2m\pi e^{\frac{3i\pi}{2}}\right)^{s-1}$$

[8] Can use L'Hopital's rule to evaluate this limit. This rule is a local statement i.e. it is concerned about behaviour of a function (single valued/multi valued) near a particular point and not the global issues (like branch cuts).

$$= (2m\pi)^{s-1} e^{i\pi(s-1)} 2\cos\left(\tfrac{\pi(s-1)}{2}\right)$$
$$= -2\,(2m\pi)^{s-1} e^{i\pi s} \sin\left(\tfrac{\pi s}{2}\right)$$



Hence by the theorem of residues [1, §3.11]

$$\int_{C_n} \frac{w^{s-1}}{e^w - 1} dw = \int_C \frac{w^{s-1}}{e^w - 1} dw + 2\pi i \left(-2e^{i\pi s} \sin\left(\frac{\pi s}{2}\right) \sum_{m=1}^{n} (2m\pi)^{s-1} \right)$$

Now let $\sigma < 0$ and make $n \to \infty$. The function $1/(e^z - 1)$ is bounded on the contours C_n, and $z^{s-1} = O\left(|z|^{\sigma-1}\right)$. Hence the integral round C_n tends to zero, and we obtain

$$\begin{aligned} I(s) &= 4\pi i e^{i\pi s} \sin\left(\frac{\pi s}{2}\right) \sum_{m=1}^{\infty} (2m\pi)^{s-1} \\ &= 4\pi i e^{i\pi s} \sin\left(\frac{\pi s}{2}\right) (2\pi)^{s-1} \zeta(1-s) \end{aligned}$$

Now using the analytic continuation formula, we deduce the desired functional equation:

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

This was proved by Bernhard Riemann[9] in 1859 [4, §2.4].

1.3.5 Zeros of Riemann Zeta Function

Using the *Euler's reflection formula* and *Legendre's duplication formula* we can rewrite the functional equation as:

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

We can rewrite this as

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

[9] "Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse." Monatsberichte der Berliner Akademie, 1859. http://www.maths.tcd.ie/pub/HistMath/People/Riemann/Zeta/

which can be expressed by saying that the function on the left is an even function of $s - \frac{1}{2}$ [3, §8]. The functional equation allows the properties of $\zeta(s)$ for $\sigma < 0$ to be inferred from its properties for $\sigma > 1$. In particular, the only zeros of $\zeta(s)$ for $\sigma < 0$ are at the poles of $\Gamma\left(\frac{s}{2}\right)$, except 0 since it's a simple pole of $\zeta(1-s)$. So, the points $s = -2, -4, -6, \ldots$ are called the *trivial zeros*.

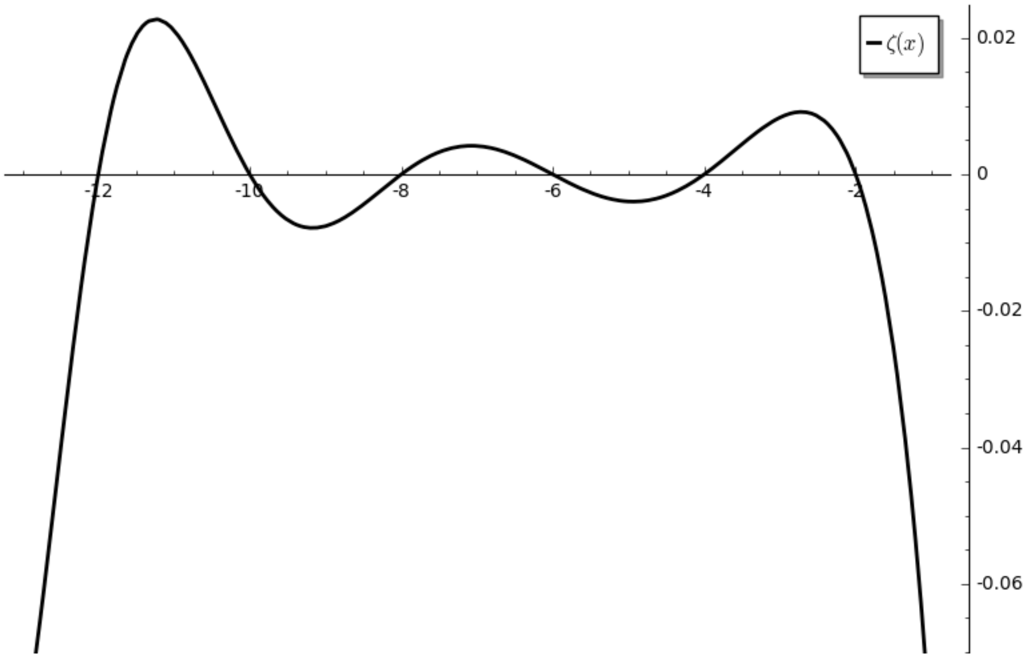


Figure 1.1: Plot of Riemann zeta function for real values of s, i.e. when $t = 0$, illustrating the *trivial zeros*. Plotted using `plot(zeta(x), (x,-13,-1), rgbcolor=(0,0,0), legend_label= '$\ zeta(x)$', thickness=2)` in SageMath 7.5.1.

The remainder of the plane, where $0 \le \sigma \le 1$, is called the *critical strip*. To analyse zeta function in this strip, we will consider the function $\xi(s)$ defined by

$$\xi(s) = \frac{1}{2}s(s-1)\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s) \tag{1.10}$$

This is an entire function because it has no pole for $\sigma \ge \frac{1}{2}$ and is an even function of $s - 1/2$. Note that $\xi(s) = \xi(1-s)$, hence it will suffice to prove the statements about $\xi(s)$ just for $\sigma \ge 1/2$. Observe that

$$\left|\frac{1}{2}s(s-1)\pi^{-\frac{s}{2}}\right| < \exp(C_1|s|) \tag{1.11}$$

for some constant C_1, and by Stirling's formula (subsection 1.2.2)

$$\left|\Gamma\left(\frac{s}{2}\right)\right| < \exp(C_2|s|\log|s|) \tag{1.12}$$

for some constant C_2, where $-\frac{\pi}{2} < \arg(s) < \frac{\pi}{2}$. And using Abel's partial summation (Appendix A) we can write for $\sigma > 1$

$$\sum_{n \le x}\frac{1}{n^s} = \frac{\lfloor x \rfloor}{x^s} + s\int_1^x \frac{\lfloor t \rfloor}{t^{s+1}}dt$$

and by taking $x \to \infty$ we get

$$\sum_{n=1}^{\infty}\frac{1}{x^s} = s\int_1^{\infty}\frac{1}{t^{s+1}}dt$$

Therefore we can rewrite $\zeta(s)$ for $\sigma > 1$ as

$$\begin{aligned} \zeta(s) &= s\int_1^\infty \frac{\lfloor x \rfloor}{x^{s+1}}dx \\ &= s\int_1^\infty \frac{x-\{x\}}{x^{s+1}}dx \\ &= \frac{s}{s-1} - s\int_1^\infty \frac{\{x\}}{x^{s+1}}dx \end{aligned} \tag{1.13}$$

Since the integral on the right is absolutely convergent for $\sigma > 0$, we get the analytic continuation[10] of $\zeta(s)$ for $\sigma > 0$. Hence for $\sigma \geq 1/2$ we have

$$|\zeta(s)| < C_3|s| \tag{1.14}$$

when $|s|$ is large. Combining (1.11), (1.12) and (1.14) we get

$$\boxed{|\xi(s)| < \exp(C|s|\log|s|)}$$

as $|s| \to \infty$ for some constant C. Hence $\xi(s)$ is of order at most 1. Moreover, as $s \to +\infty$ through real values, the above inequality is the best possible (apart from the value of C), since $\log(\Gamma(s)) \sim s\log(s)$ and $\zeta(s) \to 1$. Hence, if fact, $\xi(s)$ is of order 1.

Consequently, $\xi(s)$ does not satisfy the more precise inequality indicated in subsection 1.1.3, for entire functions of order 1. Hence $\sum r_n^{-1}$ diverges, i.e. $\xi(s)$ has infinitely many zeros.

The zeros of $\xi(s)$ are the non-trivial zeros of $\zeta(s)$, for in (1.10) the trivial zeros of $\zeta(s)$ are cancelled by the poles of $\Gamma(\frac{s}{2})$ (subsection 1.2.1), and $\frac{s}{2}\Gamma(\frac{s}{2})$ has no zeros, and the zero of $s-1$ is cancelled by the pole of $\zeta(s)$ (subsection 1.3.4). Then, recalling that $\zeta(s)$ has no zero for $\sigma > 1$ (subsection 1.3.2), we get

$$\boxed{\zeta(s) \text{ has infinitely many non-trivial zeros in the critical strip } 0 \leq \sigma \leq 1.}$$

This was proved by Jacques Hadamard[11] in 1893 [3, §12]. We can analyse the function $\xi(s)$ to write its product formula explicitly as shown in subsection 1.1.3:

$$\xi(s) = e^{A+Bs}\prod_\rho \left(1-\frac{s}{\rho}\right)e^{s/\rho} \tag{1.15}$$

where ρ is a zero of $\xi(s)$. Eliminating $\xi'(s)/\xi(s)$ from the logarithmic differentiation of (1.10) and (1.15) we get:

$$\boxed{\frac{\zeta'(s)}{\zeta(s)} = B - \frac{1}{s-1} + \frac{\log(\pi)}{2} - \frac{1}{2}\frac{\Gamma'\left(\frac{s}{2}+1\right)}{\Gamma\left(\frac{s}{2}+1\right)} + \sum_\rho \left(\frac{1}{s-\rho}+\frac{1}{\rho}\right)}$$

This exhibits the pole of $\zeta(s)$ at $s = 1$ and the non-trivial zeros at $s = \rho$.

Moreover, Riemann conjectured that any non-trivial zero of $\zeta(s)$ has $\sigma = 1/2$ (known as *Riemann Hypothesis*). This can be illustrated by the following figure [11][6, Chapter 8]

[10] So unlike previous proof of analytic continuation for all $s \in \mathbb{C}$, it's really easy to define $\zeta(s)$ as a meromorphic function for $\sigma > 0$ with only pole at $s = 1$.

[11] "Etude sur les propriétés des fonctions entières et en particulier d'une fonction considérée par Riemann." Journal de Mathématiques Pures et Appliquées (1893): 171-216. http://eudml.org/doc/234668

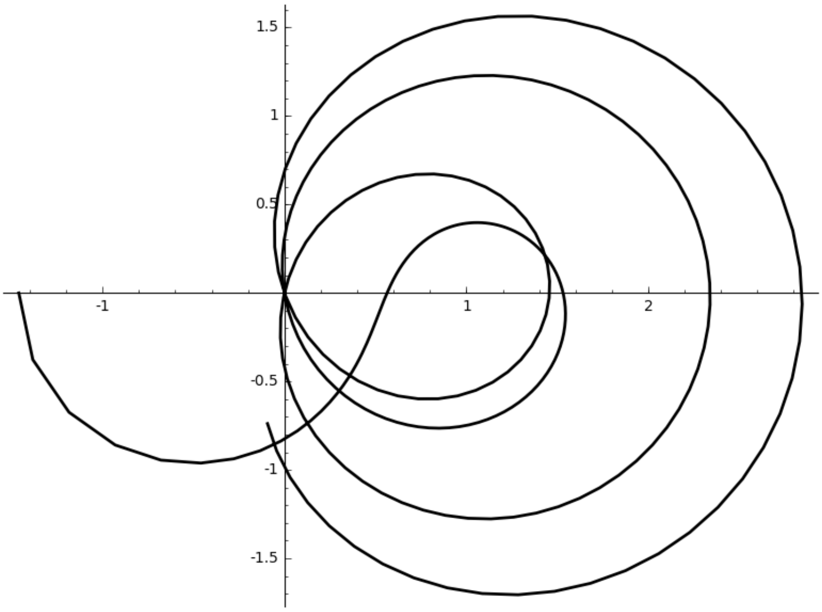


Figure 1.2: A plot of the real and imaginary parts of the Riemann zeta function $\zeta(1/2+it)$ for $0<t<30$. Plotted using `i = CDF.gen(); v = [zeta(0.5 + n/10 * i) for n in range(300)]; L = [(z.real(), z.imag()) for z in v]; line(L, rgbcolor=(0,0,0), thickness=2)` in SageMath 7.5.1.

1.4 von Mangoldt Function

We define the von Mangoldt function, $\Lambda(n)$, as:

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

Hence we have the identity: $\sum_{m|n} \Lambda(m) = \log(n)$.

Furthermore, from the *Euler product* we obtain that

$$\log(\zeta(s)) = -\sum_p \log\left(1 - \frac{1}{p^s}\right), \quad \sigma > 1 \tag{1.16}$$

And for the principal branch of the complex logarithm we have

$$-\log(1-z) = \sum_{n=1}^{\infty} \frac{z^n}{n}$$

where $|z| \leq 1$, $z \neq 1$. Hence we can re-write (1.16) as

$$\log(\zeta(s)) = \sum_p \sum_m \frac{1}{mp^{ms}} = \sum_p \sum_m \frac{1}{m} e^{-ms\log(p)}, \quad \sigma > 1$$

where p runs through all primes, and m through all positive integers. On differentiating both sides (since analytic), we get

$$\frac{\zeta'(s)}{\zeta(s)} = -\sum_p \sum_m \frac{\log(p)}{p^{ms}}, \quad \sigma > 1$$

Which can be re-written as

$$\boxed{-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}, \quad \sigma > 1}$$

and gives us the generating function for $\Lambda(n)$ [7, §17.7].

1.5 Prime Counting Function

We define $\pi(x)$ to be the number of primes which do note exceed x, i.e.

$$\pi(x) = \sum_{p \le x} 1$$

so that $\pi(1) = 0$, $\pi(2) = 1$, and so on. If p_n is the n^{th} prime, then $\pi(p_n) = n$. Therefore, $\pi(x)$ as a function of x and p_n as a function of n, are inverse functions [7, §1.5]. Therefore, to ask for an exact formula for $\pi(x)$ is same as to ask for an exact formula for p_n. No such formula is known.

The *prime number theorem*, states that

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

which is equivalent to

$$\boxed{p_n \sim n \log(n)}$$

because $\log(\log(x)) = o(\log(x))$ [7, §1.8].

Since there are infinitely many primes, $\pi(x) \to \infty$ as $x \to \infty$. Moreover, we can use the arguments used for proving existence of infinitely many primes to get a lower bound for $\pi(x)$ and upper bound for p_n. Following are the two bounds [7, §2.2, 2.6]:

Bound 1. $p_n < 2^{2^n}$ for any positive integer n and $\pi(x) \ge \log(\log(x))$ for all $x \ge 2$

Proof. Using induction, we can prove that $p_n < 2^{2^n}$ for any positive integer n. Also, we observe that for $n \ge 4$, $e^{n-1} > 2^n$ i.e. $e^{e^{n-1}} > e^{2^n} > 2^{2^n}$ (since e^x is an increasing function).

Now we divide the interval $x \ge 2$ into smaller disjoint intervals $\left[2, e^{e^3}\right]$ and $\left(e^{e^{n-1}}, e^{e^n}\right]$ for all $n \ge 4$. It's enough to prove for $x \in \left(e^{e^{n-1}}, e^{e^n}\right]$ for all $n \ge 4$, since for $x \in \left[2, e^{e^3}\right]$ the inequality follows from the fact that $\log(\log(x)) \in (-0.4, 3]$ and $\log(x)$ is an increasing function.

Let $x \in \left(e^{e^{n-1}}, e^{e^n}\right]$ where $n \ge 4$. Since $\pi(x)$ is a non-decreasing function, $p_n < 2^{2^n} < e^{e^{n-1}} < x$ implies that $\pi(p_n) \le \pi(x)$ i.e. $n \le \pi(x)$ for all $n \ge 4$. Since $\log(x)$ is an increasing function, $x < e^{e^n}$ implies that $\log(\log(x)) \le n$. Hence we conclude that $\log(\log(x)) \le n \le \pi(x)$ for all $n \ge 4$. Thus completing the proof. $\square$

Bound 2. $\pi(x) \ge \dfrac{\log(x)}{2\log(2)}$ for all $x \ge 1$ and $p_n \le 4^n$ for any positive integer n.

Proof. Suppose that $2, 3, 5, \ldots, p_j$ are the first j primes and let $N(x)$ be the number of positive integers, n, not exceeding x which are not divisible by any prime $p > p_j$. If we express such an n in the form $n = n_1^2 m$, where m is a square free integer i.e. not divisible by the square of any prime. Hence $m = 2^{b_1} 3^{b_2} \cdots p_j^{b_j}$, with b_i is either 0 or 1. There are just 2^j possible choices of the exponents and so not more than 2^j different values of m. Again, $n_1 \le \sqrt{n} \le \sqrt{x}$ and so there are not more than $\sqrt{x}$ different values of n_1. Hence

$$N(x) \le 2^j \sqrt{x}$$

Now we take, $j = \pi(x)$, so that $p_{j+1} > x$ and $N(x) = x$. Hence we have

$$N(x) = x \le 2^{\pi(x)} \sqrt{x} \quad \Longrightarrow \frac{\log(x)}{2\log(2)} \le \pi(x)$$

If we put $x = p_n$, so that $\pi(x) = n$, we get $p_n \le 4^n$. $\square$

The importance of $\zeta(s)$ in the theory of prime numbers lies in the fact that it combines two expressions, one of which contains the primes explicitly, while the other does not. The theory of primes is largely concerned with the function $\pi(x)$, the number of primes not exceeding x. We can transform the *Euler product* into a relation between $\zeta(s)$ and $\pi(x)$ [4, §1.1]. For $\sigma > 1$,

$$\begin{aligned}
\log(\zeta(s)) &= -\sum_p \log\left(1 - \frac{1}{p^s}\right) \\
&= -\sum_{n=2}^{\infty} (\pi(n) - \pi(n-1)) \log\left(1 - \frac{1}{n^s}\right) \\
&= -\sum_{n=2}^{\infty} \pi(n) \left(\log\left(1 - \frac{1}{n^s}\right) - \log\left(1 - \frac{1}{(n+1)^s}\right)\right) \\
&= \sum_{n=2}^{\infty} \pi(n) \int_n^{n+1} \left(\frac{d}{dx} \log\left(1 - \frac{1}{x^s}\right)\right) dx \\
&= \sum_{n=2}^{\infty} \pi(n) \int_n^{n+1} \frac{s}{x(x^s - 1)} dx \\
&= s \int_2^{\infty} \frac{\pi(x)}{x(x^s - 1)} dx
\end{aligned}$$

The rearrangement of the series is justified since $\log\left(1 - \frac{1}{n^s}\right) = O(n^{-\sigma})$ and $\pi(n) \leq n$. Therefore we have

$$\boxed{\log(\zeta(s)) = s \int_2^{\infty} \frac{\pi(x)}{x(x^s - 1)} dx}$$

1.6 Chebyshev Functions

The first Chebyshev function $\vartheta(x)$ is given by

$$\vartheta(x) = \sum_{p \leq x} \log(p) = \log\left(\prod_{p \leq x} p\right)$$

and the second Chebyshev function $\psi(x)$ is given by

$$\psi(x) = \sum_{p^k \leq x} \log(p) = \sum_{n \leq x} \Lambda(n)$$

Therefore, if p^m is the highest power of p not exceeding x $\log(p)$ occurs m times in $\psi(x)$. Also p^m is the highest power of p which divides any number upto x, so that

$$\psi(x) = \log(U(x))$$

where $U(x)$ is the least common multiple of all numbers upto x. We can also express $\phi(x)$ in the form

$$\psi(x) = \sum_{p \leq x} \left\lfloor \frac{\log(x)}{\log(p)} \right\rfloor \log(p)$$

The definition of $\vartheta(x)$ and $\psi(x)$ are more complicated than that of $\pi(x)$, but they are in reality more 'natural' functions [7, §22.1]. The function $\psi(x)$ can be thought of as a novel way of counting primes. Instead of adding 1 every time a prime occurs, it gives a greater weight to the primes and their powers, namely the weight $\log(p)$. With such weights the count of primes becomes almost equal to the upper bound of the interval in which they are counted. Hence the graph of $\psi(x)$ is almost a straight line (with jump discontinuities) that makes a 45° angle with horizontal axis, expecially for large numbers x [13].

Since $p^2 \leq x, p^3 \leq x, \ldots$ are equivalent to $p \leq x^{\frac{1}{2}}, p \leq x^{\frac{1}{3}}, \ldots$, we have

$$\psi(x) = \vartheta(x) + \vartheta\left(x^{\frac{1}{2}}\right) + \vartheta\left(x^{\frac{1}{3}}\right) + \ldots$$

The series breaks off when $x^{\frac{1}{m}} < 2$, i.e. when

$$m > \frac{\log(x)}{\log(2)}$$

And from the definition of $\vartheta(x)$, we can conclude that $\vartheta(x) < x\log(x)$ for all $x \geq 2$, and in fact

$$\vartheta\left(x^{\frac{1}{m}}\right) < x^{\frac{1}{m}}\log(x) \leq x^{\frac{1}{2}}\log(x)$$

for $m \geq 2$. Therefore

$$\sum_{m\geq 2}\vartheta\left(x^{\frac{1}{m}}\right) = O\left(x^{\frac{1}{2}}\left(\log(x)\right)^2\right)$$

since there are only $O(\log(x))$ terms in the series. Hence proving that

$$\psi(x) = \vartheta(x) + O\left(x^{\frac{1}{2}}\left(\log(x)\right)^2\right) \tag{1.17}$$

Next we note, without proof, that $\vartheta(x)$ is of the order x [7, §22.2],

$$\vartheta(x) \asymp x \quad \text{and} \quad \psi(x) \asymp x \quad \text{for } x \geq 2 \tag{1.18}$$

where $f(x) \asymp \phi(x)$ means $A\phi(x) < f(x) < B\phi(x)$ for some positive constants A and B independent of x, i.e. f is of the same order of magnitude as ϕ [7, §1.6]. Hence using (1.17) we can conclude that $\psi(x)$ is about the same as $\vartheta(x)$ when x is large.

Now we can write

$$\vartheta(x) = \sum_{p\leq x}\log(p) \leq \log(x)\sum_{p\leq x}1 = \pi(x)\log(x)$$

and so by (1.18)

$$\pi(x) \geq \frac{\vartheta(x)}{\log(x)} > \frac{Ax}{\log(x)} \tag{1.19}$$

for some positive constant A independent of x. On the other hand, if $0 < \delta < 1$,

$$\begin{aligned}
\vartheta(x) &\geq \sum_{x^{1-\delta}<p\leq x}\log(p) \\
&\geq (1-\delta)\log(x)\sum_{x^{1-\delta}<p\leq x}1 \\
&= (1-\delta)\log(x)\left(\pi(x) - \pi\left(x^{1-\delta}\right)\right) \\
&\geq (1-\delta)\log(x)\left(\pi(x) - x^{1-\delta}\right)
\end{aligned}$$

and so

$$\pi(x) \leq x^{1-\delta} + \frac{\vartheta(x)}{(1-\delta)\log(x)} < \frac{Bx}{\log(x)} \tag{1.20}$$

for some positive constant B independent of x. By (1.19) and (1.20) we conclude that [7, §1.8]

$$\boxed{\pi(x) \asymp \frac{x}{\log(x)}}$$

Also, from (1.19) and (1.20) it follows that

$$1 \le \frac{\pi(x)\log(x)}{\vartheta(x)} \le \frac{x^{1-\delta}\log(x)}{\vartheta(x)} + \frac{1}{1-\delta}$$

For any $\varepsilon > 0$, we can choose $\delta = \delta(\varepsilon)$ so that

$$\frac{1}{1-\delta} < 1 + \frac{\varepsilon}{2}$$

and then choose $x_0 = x_0(\delta, \varepsilon) = x_0(\varepsilon)$ so that

$$\frac{x^{1-\delta}\log(x)}{\vartheta(x)} < \frac{C\log(x)}{x^{\delta}} < \frac{\varepsilon}{2}$$

for all $x > x_0$ and some positive constant C independent of x. Hence

$$1 \le \frac{\pi(x)\log(x)}{\vartheta(x)} < 1 + \varepsilon$$

for all $x > x_0$. Since ε is arbitrary, it follows that

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

Then by (1.17) and (1.18) it follows that

$$\boxed{\pi(x) \sim \frac{\vartheta(x)}{\log(x)} \sim \frac{\psi(x)}{\log(x)}}$$

1.7 The Proof

We will prove the following version of prime number theorem[12]

$$\pi(x) = \mathrm{Li}(x) + O\left(xe^{-c\sqrt{\log(x)}}\right)$$

where $\mathrm{Li}(x) = \int_2^x \frac{dt}{\log(t)}$ and c is some positive constant, following the method used by Charles Jean de la Vallée Poussin (1896, 1899) [3, §18].

Following is Terence Tao's very informal sketch of the proof [20]:

1. Create a "sound wave" (von Mangoldt function) which is noisy at prime number times, and quiet at other times.

2. "Listen" (Mellin transform) to this wave and record the notes that you hear (the zeros of the Riemann zeta function, or the *music of the primes*). Each such note corresponds to a hidden pattern in the distribution of the primes.

3. Show that certain types of notes do not appear in this music.

4. From this (and tools such as Fourier analysis) one can prove the prime number theorem.

We can illustrate the statement of *prime number theorem* by the following graph:

[12] A weaker result, where the error term in $O\left(xe^{-c(\log(x))^{1/10}}\right)$, can be proved without much machinery from complex analysis (like the theory of entire functions of finite order, Hadamard products, etc.). That proof was given by Edmund Landau (1903, 1912) and proceeds directly to the function $\pi(x)$ but involves a nonabsolutely convergent integral. For that proof one may refer to the book by Ayoub [2, pp. 47–72]

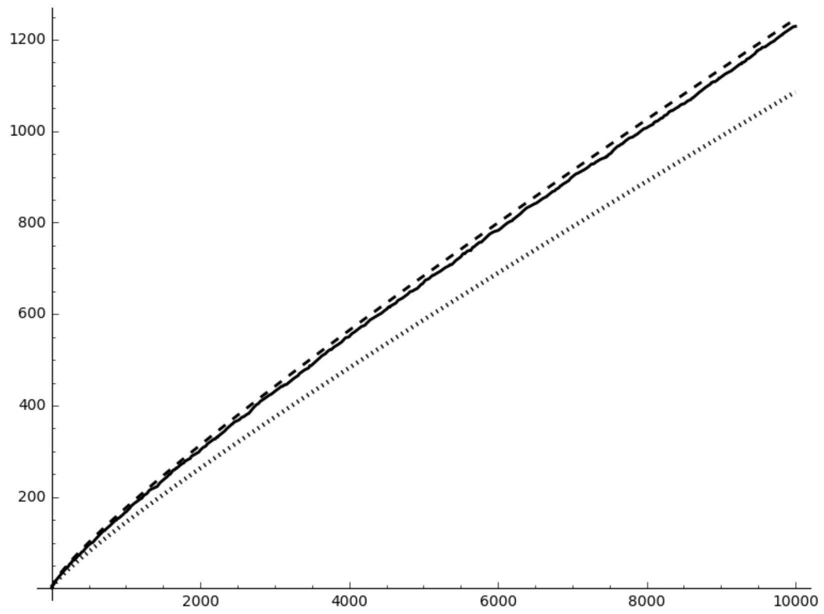


Figure 1.3: The dashed, solid and dotted curves represent $\mathrm{Li}(x)$, $\pi(x)$ and $\frac{x}{\log(x)}$ respectively for $2 \leq x \leq 10000$. Plotted using `P = plot(Li(x), (x,2,10000), linestyle="--", thickness=2, rgbcolor=(0,0,0)); Q = plot(prime_pi(x), (x,2,10000), thickness=2, rgbcolor=(0,0,0)); R =plot(x/log(x), (x,2,10000), linestyle=":", thickness=3, rgbcolor=(0,0,0)); (P+Q+R)` in SageMath 7.5.1.

1.7.1 Zero-Free Region for $\zeta(s)$

The vital step in the proof of prime number theorem is to show the existence of a thin zero-free region for $\zeta(s)$ to the left of $\sigma = 1$. We will work with the function $\zeta'(s)/\zeta(s)$, since its analytic continuation is easy and has poles only at the zeros of $\zeta(s)$ for $\sigma > 0$. From section 1.4 we know that

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}, \quad \sigma > 1$$

Therefore, we have

$$-\operatorname{Re}\left(\frac{\zeta'(s)}{\zeta(s)}\right) = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{\sigma}} \cos(t \log(n)), \quad \sigma > 1 \tag{1.21}$$

since $\operatorname{Re}(n^{-s}) = \operatorname{Re}\left(e^{-(\sigma+it)\log(n)}\right)$.

Now to proceed, we need to use following observation by Franz Mertens[13] for all $\theta \in \mathbb{R}$,

$$\begin{aligned} &2(1+\cos(\theta))^2 \geq 0 \\ \Rightarrow &2 + 2\cos^2(\theta) + 4\cos(\theta) \geq 0 \\ \Rightarrow &2 + \cos(2\theta) + 1 + 4\cos(\theta) \geq 0 \\ \Rightarrow &3 + 4\cos(\theta) + \cos(2\theta) \geq 0 \end{aligned}$$

Applied to (1.21) with t replaced[14] by $0, t, 2t$ in succession, it gives

$$3\left(-\frac{\zeta'(\sigma)}{\zeta(\sigma)}\right) + 4\left(-\operatorname{Re}\left(\frac{\zeta'(\sigma+it)}{\zeta(\sigma+it)}\right)\right) + \left(-\operatorname{Re}\left(\frac{\zeta'(\sigma+2it)}{\zeta(\sigma+2it)}\right)\right) \geq 0 \tag{1.22}$$

[13] According to Davenport [3, pp. 84], the usage of this trick was discovered in 1898, but I couldn't find that article. A less known fact is that the symbol μ for Möbius function was introduce by Mertens in 1874 [6, pp. 382].

[14] Note that the real parts are functions of $\cos(t)$ etc.

The behaviour of $-\zeta'(\sigma)/\zeta(\sigma)$ as $\sigma \to 1$, in view of the simple pole of $\zeta(s)$ at $s = 1$, we have

$$-\frac{\zeta'(\sigma)}{\zeta(\sigma)} < \frac{1}{\sigma - 1} + A_1 \tag{1.23}$$

for $1 < \sigma \leq 2$, where A_1 denotes a positive absolute constant.

The behaviour of the other two functions near $\sigma = 1$ is influenced by any zero that $\zeta(s)$ has to the left of $\sigma = 1$, at a height near to t or $2t$. Considered the following partial fraction formula shown at the end of subsection 1.3.5

$$-\frac{\zeta'(s)}{\zeta(s)} = \frac{1}{s-1} - B - \frac{\log(\pi)}{2} + \frac{1}{2}\frac{\Gamma'\left(\frac{s}{2}+1\right)}{\Gamma\left(\frac{s}{2}+1\right)} - \sum_{\rho}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right) \tag{1.24}$$

Using Stirling's formula (subsection 1.2.2) we know that

$$\log\left(\Gamma\left(\frac{s}{2}+1\right)\right) = \left(\frac{s+1}{2}\right)\log\left(\frac{s}{2}+1\right) - \left(\frac{s}{2}+1\right) + \frac{1}{2}\log(2\pi) + O\left(\frac{1}{\left|\frac{s}{2}+1\right|}\right)$$

Therefore, on differentiating it we get

$$\frac{\Gamma'\left(\frac{s}{2}+1\right)}{\Gamma\left(\frac{s}{2}+1\right)} = \frac{1}{2}\log\left(\frac{s}{2}+1\right) + \frac{s+1}{2(s+2)} - \frac{1}{2} + O\left(\frac{1}{\left|\frac{s}{2}+1\right|}\right) < C\log(s)$$

for some positive constant C. Now if we bound σ such that $1 \leq \sigma \leq 2$ but keep t unbounded such that $2 \leq t$, then we can replace $C\log(s)$ by some $C'\log(t)$. Hence for some positive constant A_2 we have

$$\mathrm{Re}\left(\frac{1}{s-1} - B - \frac{\log(\pi)}{2} + \frac{1}{2}\frac{\Gamma'\left(\frac{s}{2}+1\right)}{\Gamma\left(\frac{s}{2}+1\right)}\right) < A_2\log(t)$$

if $1 \leq \sigma \leq 2$ and $2 \leq t$. Hence, in this region using (1.24) we get

$$-\mathrm{Re}\left(\frac{\zeta'(s)}{\zeta(s)}\right) < A_2\log(t) - \sum_{\rho}\mathrm{Re}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right) \tag{1.25}$$

The sum over ρ is positive since[15]

$$\mathrm{Re}\left(\frac{1}{s-\rho}\right) = \frac{\sigma-\beta}{|s-\rho|^2} \quad \text{and} \quad \mathrm{Re}\left(\frac{1}{\rho}\right) = \frac{\beta}{|\rho|^2} \tag{1.26}$$

where $s = \sigma + it$ and $\rho = \beta + i\ell$.

We obtain a valid inequality when $s = \sigma + 2it$ by just omitting the sum:

$$-\mathrm{Re}\left(\frac{\zeta'(\sigma+2it)}{\zeta(\sigma+2it)}\right) < A_2\log(t) \tag{1.27}$$

For the case when $s = \sigma + it$, we choose t to coincide with the ordinate ℓ of the zero $\rho = \beta + i\ell$, with $\ell \geq 2$ and just the one term $1/(s-\rho)$ in the sum which corresponds to this zero:

$$-\mathrm{Re}\left(\frac{\zeta'(\sigma+it)}{\zeta(\sigma+it)}\right) < A_2\log(t) - \frac{1}{\sigma-\beta} \tag{1.28}$$

Substituting (1.23), (1.27) and (1.28) in (1.22) we obtain:

$$3\left(\frac{1}{\sigma-1} + A_1\right) + 4\left(A_2\log(t) - \frac{1}{\sigma-\beta}\right) + A_2\log(t) > 0$$

[15] Note that $2\,\mathrm{Re}(z) = z + \bar{z}$

$$\Rightarrow \frac{4}{\sigma - \beta} < \frac{3}{\sigma - 1} + A_3 \log(t)$$

Take $\sigma = 1 + \delta/\log(t)$, where δ is a positive constant. Then

$$\beta < 1 + \frac{\delta}{\log(t)} - \frac{4\delta}{(3 + A_3\delta)\log(t)}$$

and if δ is suitably chosen in relation to A_3, this gives

$$\beta < 1 - \frac{c}{\log(t)}$$

where c is a positive constant to which a numerical value could be assigned. Thus we have:

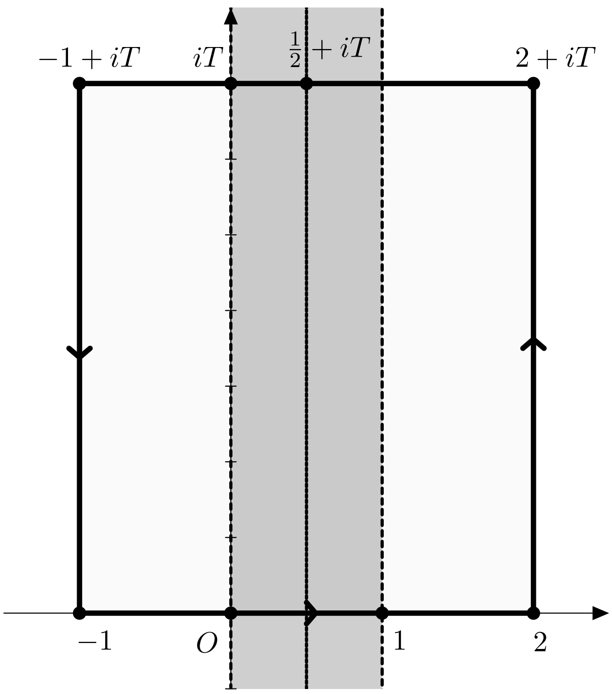
There exists a positive constant numerical constant c such that $\zeta(s)$ has no zero in the region

$$\sigma \geq 1 - \frac{c}{\log(t)}, \quad t \geq 2$$

Hence, $\zeta(s) \neq 0$ on $\sigma = 1$. This was proved independently by Hadamard[16] and de la Vallée Poussin[17] in 1896 [3, §13].

1.7.2 Counting the Zeros of $\zeta(s)$

Consider a rectangle $0 < \sigma < 1$, $0 < t < T$, and we want to find an approximate formula for $N(T)$, the number of zeros of $\zeta(s)$ in this rectangle. But, it is convenient to work initially with $\xi(s)$ rather than with $\zeta(s)$ because it has a simple functional equation and is an entire function (subsection 1.3.5), hence it's an analytic function.



[16] "Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques." Bulletin de la Société Mathématique de France 24 (1896): 199–220. http://eudml.org/doc/85858

[17] "Recherches analytiques la théorie des nombres premiers." Ann. Soc. scient. Bruxelles 20 (1896): 183–256. https://archive.org/details/recherchesanaly00pousgoog

If we assume (for simplicity) that T (which we suppose to be large) does not coincide with the ordinate of a zero of $\xi(s)$ and $-1 < \sigma < 2$, then using Cauchy's argument principle [1, §3.41] we can say that

$$N(T) = \frac{1}{2\pi}\Delta_R \arg\left(\xi(s)\right) \tag{1.29}$$

where Δ_R denotes the variation of $\arg\left(\xi(s)\right)$ round the contour R, which is a rectangle with vertices 2, $2+iT$, $-1+iT$, -1, described in the positive sense.

There is no change in $\arg(\xi(s))$ as s describes the base of the base of the rectangle, since $\xi(s)$ is then real and nowhere 0. Further, the change in $\arg(\xi(s))$ as s moves from $\frac{1}{2}+iT$ to $-1+iT$ and then to -1 is equal to the change as s moves from 2 to $2+iT$ and then to $\frac{1}{2}+iT$, since

$$\xi(\sigma+it) = \xi(1-\sigma-it) = \overline{\xi(1-\sigma+it)}$$

Hence, we can rewrite (1.29) as

$$N(T) = 2\left(\frac{1}{2\pi}\Delta_L \arg\left(\xi(s)\right)\right) = \frac{1}{\pi}\Delta_L \arg\left(\xi(s)\right) \tag{1.30}$$

where L consists of the edges of rectangle from 2 to $2+iT$ and then to $\frac{1}{2}+iT$.

The definition of $\xi(x)$, from (1.10) can be written as

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

Hence we have

$$\Delta_L \arg(\xi(s)) = \Delta_L \arg(s-1) + \Delta_L \arg\left(\pi^{-\frac{s}{2}}\right) + \Delta_L \arg\left(\Gamma\left(\frac{s}{2}+1\right)\right) + \Delta_L \arg\left(\zeta(s)\right) \tag{1.31}$$

We note that by using the facts that

$$\arctan(x) + \arctan\left(\frac{1}{x}\right) = \frac{\pi}{2}, \quad x > 0$$

and

$$\arctan(x) = -\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}x^{2n+1}, \quad |x| \leq 1$$

we get

$$\arctan(x) = \frac{\pi}{2} - \arctan\left(\frac{1}{x}\right) = \frac{\pi}{2} - \frac{1}{x} + \frac{1}{3x^3} - \frac{1}{5x^5} + \ldots$$

for large values of $x > 0$. Hence we have

$$\begin{aligned}
\Delta_L \arg(s-1) &= \arg\left(\frac{1}{2}+iT-1\right) - \arg(2-1) \\
&= \arg\left(-\frac{1}{2}+iT\right) \\
&= \pi - \arctan(2T) \\
&= \frac{\pi}{2} + \arctan\left(\frac{1}{2T}\right) \\
&= \frac{\pi}{2} + \frac{1}{2T} - \frac{1}{3(2T)^3} + \frac{1}{5(2T)^5} - \ldots \\
&= \frac{\pi}{2} + O\left(\frac{1}{T}\right)
\end{aligned} \tag{1.32}$$

We know that $a^b = e^{b\log(a)}$ hence using the definition of argument we get

$$\begin{aligned}
\Delta_L \arg\left(\pi^{-\frac{s}{2}}\right) &= \Delta_L \arg\left(\exp\left(-\frac{s}{2}\log(\pi)\right)\right) \\
&= \arg\left(\exp\left(-\frac{1+i2T}{4}\log(\pi)\right)\right) - \arg\left(\exp\left(-\log(\pi)\right)\right) \\
&= -\frac{T}{2}\log(\pi)
\end{aligned} \tag{1.33}$$

Using the fact that $\log(f(z)) = \log|f(z)| + i\arg(f(z))$ and Stirling's formula (subsection 1.2.2), we get

$$\begin{aligned}
\Delta_L \arg\left(\Gamma\left(\frac{s}{2}+1\right)\right) &= \Delta_L \operatorname{Im}\left(\log\left(\Gamma\left(\frac{s}{2}+1\right)\right)\right) \\
&= \operatorname{Im}\left(\log\left(\Gamma\left(\frac{5}{4}+i\frac{T}{2}\right)\right)\right) - \operatorname{Im}\left(\log\left(\Gamma\left(2\right)\right)\right) \\
&= \operatorname{Im}\left(\left(\frac{3}{4}+i\frac{T}{2}\right)\log\left(\frac{5}{4}+i\frac{T}{2}\right) - \frac{5}{4} - i\frac{T}{2} + \frac{1}{2}\log(2\pi) + O\left(\frac{1}{T}\right)\right) \\
&= \frac{3}{4}\arg\left(\frac{5}{4}+i\frac{T}{2}\right) + \frac{T}{2}\log\left|\frac{5}{4}+i\frac{T}{2}\right| - \frac{T}{2} + O\left(\frac{1}{T}\right) \\
&= \frac{3}{4}\arctan\left(\frac{2T}{5}\right) + \frac{T}{2}\log\left|\frac{\sqrt{25+4T^2}}{4}\right| - \frac{T}{2} + O\left(\frac{1}{T}\right) \\
&= \frac{3}{4}\left(\frac{\pi}{2} - \arctan\left(\frac{5}{2T}\right)\right) + \frac{T}{2}\log\left(\frac{T}{2}\right) - \frac{T}{2} + O\left(\frac{1}{T}\right) \\
&= \frac{3\pi}{8} + \frac{T}{2}\log\left(\frac{T}{2}\right) - \frac{T}{2} + O\left(\frac{1}{T}\right)
\end{aligned} \tag{1.34}$$

Now to find $\Delta_L \arg(\zeta(s))$, we will make use of (1.25) from the previous section, for $1 \le \sigma \le 2$ and $2 \le t$:

$$-\operatorname{Re}\left(\frac{\zeta'(s)}{\zeta(s)}\right) < A\log(t) - \sum_{\rho}\operatorname{Re}\left(\frac{1}{s-\rho}+\frac{1}{\rho}\right)$$

where A is some positive constant. In this formula we take $s = 2+iT$. Since $|\zeta'(s)/\zeta(s)|$ is bounded for such s, we obtain

$$\sum_{\rho}\operatorname{Re}\left(\frac{1}{s-\rho}+\frac{1}{\rho}\right) < A\log(T)$$

As seen earlier in (1.26), all the terms in both series are positive, and since

$$\operatorname{Re}\left(\frac{1}{s-\rho}\right) = \frac{2-\beta}{(2-\beta)^2+(T-\ell)^2} \ge \frac{1}{4+(T-\ell)^2}$$

we conclude that *if $\rho = \beta + i\ell$ runs through the non-trivial zeros of $\zeta(s)$, then for large T*

$$\sum_{\rho}\frac{1}{1+(T-\ell)^2} = O(\log(T))$$

From this it follows that

(a) the number of zeros with $T-1 < \ell < T+1$ is $O(\log(T))$

(b) the sum $\sum(T-\ell)^{-2}$ extended over the zeros with ℓ outside the interval $T-1 < \ell < T+1$ is also $O(\log(T))$

Also by (1.24), applied at s and $2+it$, and subtracted, we get

$$-\frac{\zeta'(s)}{\zeta(s)} = O(\log(t)) + \sum_{\rho}\left(\frac{1}{s-\rho} - \frac{1}{2+it-\rho}\right)$$

We can split the summation over ρ in two parts, namely when $|t-\ell| < 1$ and when $|t-\ell| \geq 1$. For the terms with $|t-\ell| \geq 1$, we have

$$\left|\frac{1}{s-\rho} - \frac{1}{2+it-\rho}\right| = \frac{2-\sigma}{|(s-\rho)(2+it-rho)|} \leq \frac{3}{|\ell-t|^2}$$

and sum of these is of $O(\log(t))$ by the conclusion (b) above. As for the terms with $|\ell-t| < 1$, we have $|2+it-\rho| \geq 1$, and the number of terms is $O(\log(t))$ by the conclusion (a) above. Hence we deduce that *for large t (not coinciding with the ordinate of a zero) and* $-1 \leq \sigma \leq 2$

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_{\rho'}\frac{1}{s-\rho'} + O(\log(t)) \tag{1.35}$$

where $\rho' = \beta + i\ell$ are those non-trivial zeros of $\zeta(s)$ for which $|t-\ell| < 1$.

From this fact we can find $\Delta_L \arg(\zeta(s))$, as

$$\begin{aligned}
\Delta_L \arg(\zeta(s)) &= \arg\left(\zeta\left(\frac{1}{2}+iT\right)\right) - \arg(\zeta(2)) \\
&= \operatorname{Im}\left(\log\left(\zeta\left(\frac{1}{2}+iT\right)\right)\right) \\
&= O(1) - \int_{\frac{1}{2}+iT}^{2+iT} \operatorname{Im}\left(\frac{\zeta'(s)}{\zeta(s)}\right) ds \\
&= O(1) - \int_{\frac{1}{2}+iT}^{2+iT} \operatorname{Im}\left(\sum_{\rho'}\frac{1}{s-\rho'} + O(\log(t))\right) ds \\
&= O(1) - \sum_{\rho'}\int_{\frac{1}{2}+iT}^{2+iT} \operatorname{Im}\left(\frac{1}{s-\rho'}\right) ds + O(\log(T)) \\
&= O(1) - \sum_{\rho'}\Delta_{L'} \arg(s-\rho) + O(\log(T))
\end{aligned}$$

where $O(1)$ term is from the variation along $\sigma = 2$ and L' is the line joining $\frac{1}{2}+iT$ to $2+iT$. Since $|\Delta_{L'} \arg(s-\rho)|$ is at most π and the number of terms in the sum $\sum_{\rho'}\Delta_{L'} \arg(s-\rho)$ is $O(\log(T))$ by the conclusion (a) above, we get

$$\begin{aligned}
\Delta_L \arg(\zeta(s)) &= O(1) + O(\log(T)) + O(\log(T)) \\
&= O(\log(T))
\end{aligned} \tag{1.36}$$

Now using (1.32), (1.33), (1.34) and (1.36) in (1.31) we get

$$\Delta_L \arg(\xi(s)) = \frac{T}{2}\log\left(\frac{T}{2\pi}\right) - \frac{T}{2} + \frac{7\pi}{8} + O(\log(T))$$

And using this in (1.30) we finally get

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

This was proved by von Mangoldt, first in 1895 with a slightly less good error term and then fully in 1905. It would seem at first sight that we might as well omit the term 7/8; but as we shall see later, it has a certain significance [3, §15].

From the following plot for $T = 50$ we can see the number of zeros of the Riemann zeta function along the critical line[18] is 10, and $N(50) \approx 9.4$.

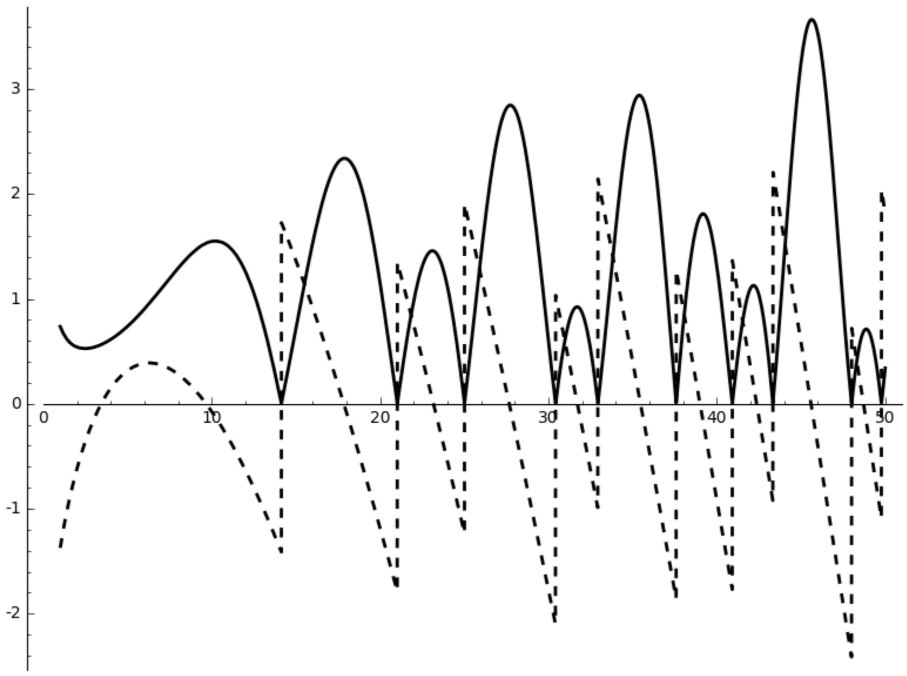


Figure 1.4: The dashed and solid curves represent the plot of $\arg(\zeta(s))$ and $|\zeta(s)|$ respectively, for $\sigma = 1/2$ and $1 \leq t \leq 50$. Plotted using `i = CDF.0; p1 = plot(lambda t: arg(zeta(0.5+t*i)), 1, 50,linestyle="--", thickness=2, rgbcolor=(0,0,0)); p2 = plot(lambda t: abs(zeta(0.5+t*i)), 1, 50,thickness=2, rgbcolor=(0,0,0)); p1+p2` in SageMath 7.5.1.

1.7.3 The Explicit Formula for $\psi(x)$

Riemann's great accomplishment was transforming the problem of describing prime numbers into the problem of describing the zeros of the Riemann zeta function - which can be attacked directly [13]. There is an *explicit formula* for $\psi(x)$, valid for $x > 1$, which consists of a sum over the complex (non-trivial) zeros ρ of $\zeta(s)$. It is astonishing that there can be such a formula, an exact expression for the number of primes upto x in terms of the zeros of a complicated function [9, §3]. To avoid some minor complications we shall suppose that $\boxed{x \geq 2}$, though the formula will be valid for $x > 1$.

Next we will use the inverse Mellin transform called Perron's Formula [1, §9.42], which states that if *x is not an integer, c is any positive number, and $\sigma > \sigma_0 - c$, then*

$$\sum_{n<x} \frac{a_n}{n^w} = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} f(s+w) \frac{x^s}{s} ds$$

where $f(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$ is a Dirichlet series. The particular case $w = 0$ is

$$\sum_{n<x} a_n = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} f(s) \frac{x^s}{s} ds \tag{1.37}$$

[18] Assuming Riemann hypothesis, all zeros in region $0 < \sigma < 1$ lie on $\sigma = 1/2$.

and applying this to the result obtained in section 1.4 we get

$$\sum_{n<x} \Lambda(n) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \left(-\frac{\zeta'(s)}{\zeta(s)} \right) \frac{x^s}{s} ds, \qquad c > 1$$

and since x is not an integer, we can write

$$\psi(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \left(-\frac{\zeta'(s)}{\zeta(s)} \right) \frac{x^s}{s} ds, \qquad x \notin \mathbb{Z}, c > 1 \tag{1.38}$$

If we can move the vertical line of integration away to negative infinity on the left we shall express $\psi(x)$ as the sum of the residues [1, §3.11] of the function $(-\zeta'(s)/\zeta(s))\, x^s/s$ at its poles. The pole of $\zeta(s)$ at $s = 1$ contributes x, the pole of $1/s$ at $s = 0$ contributes $-\zeta'(0)/\zeta(0)$; and each zero ρ of $\zeta(s)$, whether trivial or not, contributes $-x^\rho/\rho$.

The basic idea is to make use of following discontinuous integral [2, Theorem 3.2]:

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{y^s}{s} ds = \begin{cases} 0 & \text{if } 0 < y < 1 \\ \frac{1}{2} & \text{if } y = 1 \\ 1 & \text{if } y > 1 \end{cases} \tag{1.39}$$

where $c > 0$. We can obtain (1.37) from (1.39) by taking $y = x/n$ where $n < x$ such that $x \notin \mathbb{Z}$, as

$$\begin{aligned} \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} f(s) \frac{x^s}{s} ds &= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \sum_{n=1}^{\infty} \frac{a_n}{n^s} \frac{x^s}{s} ds \\ &= \frac{1}{2\pi i} \sum_{n=1}^{\infty} a_n \int_{c-i\infty}^{c+i\infty} \left(\frac{x}{n} \right)^s \frac{1}{s} ds \\ &= \sum_{n<x} a_n \end{aligned}$$

Moreover, the method used to prove[19] (1.39) can be extended to prove [3, §17] that if

$$\delta(y) = \begin{cases} 0 & \text{if } 0 < y < 1 \\ \frac{1}{2} & \text{if } y = 1 \\ 1 & \text{if } y > 1 \end{cases} \qquad \text{and} \qquad I(y, T) = \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{y^s}{s} ds$$

for $y > 0, c > 0, T > 0$, then

$$|\delta(y) - I(y, T)| < \begin{cases} y^c \min\left(1, \frac{1}{T|\log(y)|}\right) & \text{if } y \neq 1 \\ \frac{c}{T} & \text{if } y = 1 \end{cases} \tag{1.40}$$

To be able to use this result, we will have to rewrite $\psi(x)$. Note that, just like $\pi(x)$, $\psi(x)$ is also a discontinuous function with jump discontinuities at the points where x is a prime power. So we modify the definition by taking the mean of the values as

$$\begin{aligned} \psi_0(x) &= \frac{1}{2} \left(\sum_{n \le x} \Lambda(n) + \sum_{n<x} \Lambda(n) \right) \\ &= \begin{cases} \psi(x) - \frac{\Lambda(x)}{2} & \text{if } x = p^k \text{ for some prime } p \text{ and } k > 0 \\ \psi(x) & \text{otherwise} \end{cases} \end{aligned} \tag{1.41}$$

[19]We will have to use ML inequality [1, §2.31] after selecting the appropriate contour.

By replacing y by x/n, $\delta(y)$ by $\psi_0(x)$ and $I(y,T)$ by

$$J(x,T) = \frac{1}{2\pi i}\int_{c-iT}^{c+iT}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{x^s}{s}ds \tag{1.42}$$

we get from (1.40) that

$$|\psi_0(x) - J(x,T)| < \sum_{\substack{n=1\\ n\neq x}}^{\infty}\Lambda(n)\left(\frac{x}{n}\right)^c \min\left(1, \frac{1}{T\left|\log\left(\frac{x}{n}\right)\right|}\right) + \frac{c}{T}\Lambda(x) \tag{1.43}$$

where $c > 1$ and the term containing $\Lambda(x)$ is present only if x is a prime power.

Now we make a choice of c as per our convenience as

$$c = 1 + \frac{1}{\log(x)}$$

which is equivalent to $\boxed{x^c = ex}$.

We want to estimate the series on the right of (1.43), and we will achieve that by considering various cases[20] (so as to avoid $n = x$):

Case 1. If $n \leq \frac{3}{4}x$ or $n \geq \frac{5}{4}x$.

Then $|\log(x/n)|$ has a positive lower bound, so we get

$$\begin{aligned}\sum_n \Lambda(n)\left(\frac{x}{n}\right)^c \min\left(1, \frac{1}{T\left|\log\left(\frac{x}{n}\right)\right|}\right) &\ll \frac{x}{T}\sum_{n=1}^{\infty}\frac{\Lambda(n)}{n^c}\\ &= \frac{x}{T}\left(-\frac{\zeta'(c)}{\zeta(c)}\right)\\ &\ll \frac{x}{T}\log(x)\end{aligned}$$

Case 2. If $\frac{3}{4}x < n < x$.

Let x_1 be the largest prime power less than x, we can suppose that $\frac{3}{4}x < x_1 < x$, since otherwise the terms under consideration vanish.

(a) For the term $n = x_1$, using the series expansion of log function (section 1.4) we have

$$\log\left(\frac{x}{n}\right) = -\log\left(1 - \frac{x - x_1}{x}\right) \geq \frac{x - x_1}{x}$$

and therefore the contribution of this term is

$$\begin{aligned}\sum_n \Lambda(n)\left(\frac{x}{n}\right)^c \min\left(1, \frac{1}{T\left|\log\left(\frac{x}{n}\right)\right|}\right) &\ll \Lambda(x_1)\min\left(1, \frac{x}{T(x - x_1)}\right)\\ &\ll \log(x)\min\left(1, \frac{x}{T(x - x_1)}\right)\end{aligned}$$

(b) For other terms, we can put $n = x_1 - \nu$, where $0 < \nu < \frac{1}{4}x$, and then again using the series expansion of log function (section 1.4) we have

$$\log\left(\frac{x}{n}\right) \geq \log\left(\frac{x_1}{n}\right) = -\log\left(1 - \frac{\nu}{x_1}\right) \geq \frac{\nu}{x_1}$$

[20] We will be using Vinogradov's symbolism [3, pp. 107] where $f(x) \ll g(x)$ is equivalent to $f(x) = O(g(x))$.

Hence the contribution of these terms is

$$\sum_n \Lambda(n)\left(\frac{x}{n}\right)^c \min\left(1, \frac{1}{T\left|\log\left(\frac{x}{n}\right)\right|}\right) \ll \sum_{0<\nu<\frac{1}{4}x} \frac{\Lambda(x_1-\nu)}{T}\cdot\frac{x_1}{\nu}$$
$$\ll \frac{x}{T}\left(\log(x)\right)^2$$

Case 3. If $x < n < \frac{5}{4}x$.

Let x_2 be the least prime power greater than x, we can suppose that $x < x_2 < \frac{5}{4}x$, since otherwise the terms under consideration vanish.

(a) For the term $n = x_2$, using the series expansion of log function (section 1.4) we have

$$\left|\log\left(\frac{x}{n}\right)\right| = -\log\left(\frac{x}{n}\right) = -\log\left(1-\frac{x_2-x}{x_2}\right) \geq \frac{x_2-x}{x_2}$$

and therefore the contribution of this term is

$$\sum_n \Lambda(n)\left(\frac{x}{n}\right)^c \min\left(1, \frac{1}{T\left|\log\left(\frac{x}{n}\right)\right|}\right) \ll \Lambda(x_2)\min\left(1, \frac{x_2}{T(x_2-x)}\right)$$
$$\ll \log(x)\min\left(1, \frac{x}{T(x_2-x)}\right)$$

(b) For other terms, we can put $n = x_2+\nu$, where $0 < \nu < \frac{1}{4}x$, and then again using the series expansion of log function (section 1.4) we have

$$\left|\log\left(\frac{x}{n}\right)\right| = -\log\left(\frac{x}{n}\right) \geq \log\left(\frac{n}{x_2}\right) = -\log\left(1-\frac{\nu}{n}\right) \geq \frac{\nu}{n}$$

Hence the contribution of these terms is

$$\sum_n \Lambda(n)\left(\frac{x}{n}\right)^c \min\left(1, \frac{1}{T\left|\log\left(\frac{x}{n}\right)\right|}\right) \ll \sum_{0<\nu<\frac{1}{4}x} \frac{\Lambda(x_2+\nu)}{T}\cdot\frac{x_2+\nu}{\nu}$$
$$\ll \frac{x}{T}\left(\log(x)\right)^2$$

If we define $\langle x\rangle$ to be the distance from x to the nearest prime power, other than x itself in case x is a prime power and replace[21] $\psi_0(x)$ by $\psi(x)$, we can rewrite the above estimates for (1.43) as

$$|\psi(x) - J(x,T)| \ll \frac{x(\log(x))^2}{T} + \log(x)\min\left(1, \frac{x}{T\langle x\rangle}\right) \tag{1.44}$$

The next step is to replace the vertical line of integration in (1.42) by the other three sides of the rectangle with vertices at $c-iT, c+iT, -U+iT, -U-iT$ where U is a large odd integer. (in order to achieve what we earlier commented for (1.38)). Thus the left vertical side passes halfway between two of the trivial zeros of $\zeta(s)$.

Hence we can write

$$J(x,T) = \frac{1}{2\pi i}\int_C \left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{x^s}{s}ds + \frac{1}{2\pi i}\int_{-U+iT}^{c+iT}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{x^s}{s}ds$$

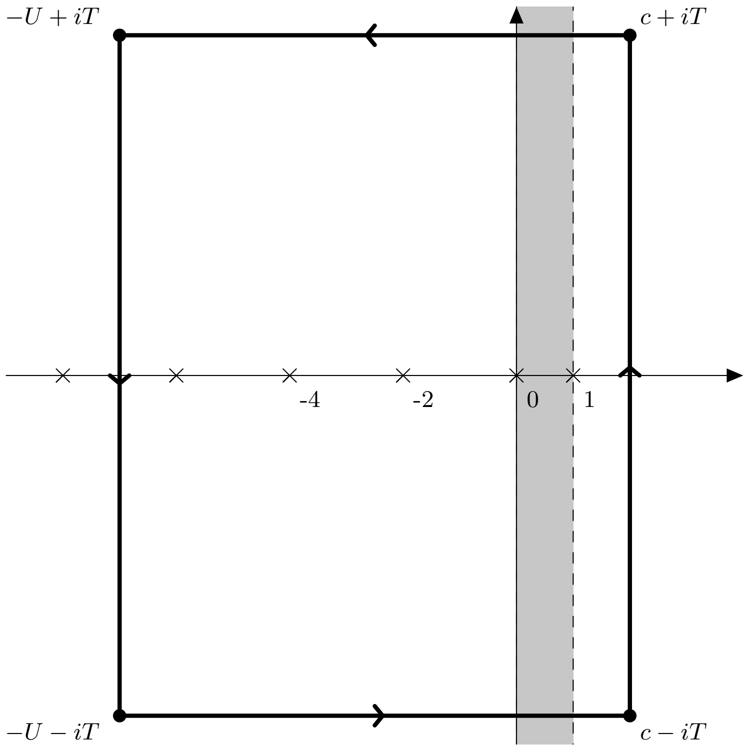
[21]Note that $\psi_0(x)$ is the same as $\psi(x)$ except that at its jump discontinuities (the prime powers) it takes the value halfway between the values to the left and the right. So while analysing the asymptotic behaviour, we can safely make this replacement.

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

The sum of the residues of the integrand at its poles inside the rectangle is (as earlier commented for (1.38))

$$\frac{1}{2\pi i}\int_C\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{x^s}{s}ds=x-\frac{\zeta'(0)}{\zeta(0)}-\sum_{|\ell|<T}\frac{x^\rho}{\rho}-\sum_{0<2m<U}\frac{x^{-2m}}{-2m} \tag{1.45}$$

where $\rho=\beta+i\ell$ is the non-trivial zero of $\zeta(s)$.



Now we need to carefully choose the T. By the conclusion (a) of subsection 1.7.2, we know that for any large T, the number of zeros with $|\ell-T|<1$ is $O(\log(T))$, i.e. $N(T)\ll\log(T)$. Among the ordinates of these zeros there must be a gap of length[22] $\gg\frac{1}{\log(T)}$. Hence by varying T by a bounded amount, we can ensure that

$$\frac{1}{\log(T)}\ll|\ell-T|$$

for all zeros $\rho=\beta+i\ell$.

Now we will determine the contribution made by horizontal integrals depending upon the choice of σ

Case 1. If $-1\le\sigma\le 2$

Recall (1.35) from subsection 1.7.2 that

$$\frac{\zeta'(s)}{\zeta(s)}=\sum_{|\ell-T|<1}\frac{1}{s-\rho}+O\left(\log(T)\right)$$

[22]Recall that $f(x)=O(g(x))\iff\exists M,x_0\in\mathbb{R},M>0$ such that $|f(x)|<M|g(x)|\forall x>x_0$ and hence $\frac{1}{g(x)}=O\left(\frac{1}{f(x)}\right)$. In Vinogradov symbolism, $f(x)\ll g(x)\iff\frac{1}{g(x)}\ll\frac{1}{f(x)}$.

for $s = \sigma + iT$ and $-1 \le \sigma \le 2$. With the present choice of T, each term is $\ll \log(T)$, and the number of terms is also $\ll \log(T)$. Hence on the new horizontal lines of integration we have

$$\frac{\zeta'(s)}{\zeta(s)} = O\left((\log(T))^2\right)$$

The contribution made to the horizontal integrals by this range of σ, $-1 \le \sigma \le 2$ is therefore

$$\begin{aligned}\frac{1}{2\pi i}\int_{-U+iT}^{c+iT}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{x^s}{s}ds + \frac{1}{2\pi i}\int_{c-iT}^{-U-iT}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{x^s}{s}ds &\ll (\log(T))^2\int_{-1}^{c}\left|\frac{x^s}{s}\right|d\sigma \\ &\ll \frac{(\log(T))^2}{T}\int_{-\infty}^{c}x^\sigma d\sigma \\ &\ll \frac{x\,(\log(T))^2}{T\log(x)}\end{aligned} \tag{1.46}$$

Case 2. If $-U \le \sigma \le -1$.

We need to estimate for $|\zeta'(s)/\zeta(s)|$ for $\sigma \le -1$. From the functional equation stated at the end of subsection 1.3.4

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

we note that if $1-\sigma \le -1$ the functions on the right have to be considered only for $\sigma \ge 2$. The logarithmic derivative is

$$-\frac{\zeta'(1-s)}{\zeta(1-s)} = -\log(2) - \log(\pi) - \frac{\pi}{2}\tan\left(\frac{\pi s}{2}\right) + \frac{\Gamma'(s)}{\Gamma(s)} + \frac{\zeta'(s)}{\zeta(s)}$$

Observe that $\tan\left(\frac{\pi s}{2}\right)$ is bounded if $|s-(2m+1))| \ge \frac{1}{2}$, that is, if $|(1-s)+2m| \ge \frac{1}{2}$. Also, $\frac{\Gamma'(s)}{\Gamma(s)} \ll \log|s|$ and therefore $\ll \log\left(2|1-s|\right)$ for $\sigma \ge 2$. And $\frac{\zeta'(s)}{\zeta(s)}$ is bounded. Hence it follows that

$$\left|\frac{\zeta'(s)}{\zeta(s)}\right| \ll \log(2|s|) \tag{1.47}$$

in the half-plane $\sigma \le -1$, provided that the circles of radius $1/2$ (say) around all the trivial zeros at $s = -2, -4, \ldots$ are excluded. Hence the contribution to the remainder of the horizontal integrals for $-U \le \sigma \le -1$ is

$$\begin{aligned}\frac{1}{2\pi i}\int_{-U+iT}^{c+iT}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{x^s}{s}ds + \frac{1}{2\pi i}\int_{c-iT}^{-U-iT}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{x^s}{s}ds &\ll \frac{\log(2T)}{T}\int_{-U}^{-1}x^\sigma d\sigma \\ &\ll \frac{\log(T)}{Tx\log(x)}\end{aligned}$$

which is negligible compared with (1.46).

Also by using (1.47) we get the contribution made by the vertical integral at $\sigma = -U$ as

$$\begin{aligned}\frac{1}{2\pi i}\int_{-U-iT}^{-U+iT}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{x^s}{s}ds &\ll \frac{\log(2U)}{U}\int_{-T}^{T}x^{-U}dt \\ &\ll \frac{T\log(U)}{Ux^U}\end{aligned}$$

which vanishes as $U \to \infty$.

Making $U \to \infty$ and adding the estimate (1.46) to (1.45), then using (1.38) and (1.44) we obtain

$$\psi(x) = x - \sum_{|\ell|<T} \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2}\log\left(1 - \frac{1}{x^2}\right) + R(x,T) \tag{1.48}$$

where

$$|R(x,T)| \ll \frac{x(\log(xT))^2}{T} + \log(x)\min\left(1, \frac{x}{T\langle x\rangle}\right) \tag{1.49}$$

As $T \to \infty$ for any given $x \geq 2$, we have $R(x,T) \to 0$, and therefore we get the desired explicit formula

$$\boxed{\psi(x) = x - \sum_{\rho} \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2}\log\left(1 - \frac{1}{x^2}\right)}$$

where the sum over the non-trivial zeros ρ of $\zeta(s)$ is to be understood in the symmetric sense as $\lim_{T\to\infty} \sum_{|\ell|<T} \frac{x^\rho}{\rho}$. The convergence is uniform in any closed interval of x which doesn't contain a prime power, but not otherwise, since $\psi(x)$ is discontinuous at each prime power value of x. This was proved by von Mangoldt in 1895 [3, §17].

The result (1.48) and (1.49) constitute the more precise form of the explicit formula. We proved them subject to a restriction on T, but this can now be removed. The effect of varying T by a bounded amount is to change the sum over ρ by $O(\log(T))$ terms, and each term is $O(x/T)$. Hence the variation in the' sum is $O(x(\log(T))/T)$, and this is covered by the estimate on the right of (1.49).

We note for future reference that, *if x is an integer, then $\langle x\rangle \geq 1$, and (1.49) takes the simpler form*

$$|R(x,T)| \ll \frac{x}{T}(\log(xT))^2 \tag{1.50}$$

The results (1.48) and (1.49) continue to hold for $1 < x < 2$, with a slight modification in the form of the estimate for $R(x,T)$.

1.7.4 Completing the Proof

We have to estimate the sum $\sum_{|\ell|<T} \frac{x^\rho}{\rho}$ in (1.48) of subsection 1.7.3. Firstly we will estimate x^ρ using the fact that the real part β of ρ is not too near 1 (subsection 1.7.1). It follows from the conclusion of subsection 1.7.1 that if $\ell < T$, where T is large, then

$$\beta < 1 - \frac{c_1}{\log(T)}$$

where c_1 is a positive absolute constant. Hence

$$|x^\rho| = x^\beta < x\exp\left(-c_1\frac{\log(x)}{\log(T)}\right) \tag{1.51}$$

Secondly we will estimate $1/\rho$. Since $|\rho| \geq \ell$ for $\ell > 0$, so we just need to estimate $\sum_{0<\ell<T} \frac{1}{\ell}$. If $N(t)$ denotes, as in subsection 1.7.2, the number of zeros in the critical strip with ordinates less than t, this sum is

$$\sum_{0<\ell<T} \frac{1}{\ell} = \int_0^T \frac{1}{t}dN(t) = \frac{1}{T}N(T) + \int_0^T \frac{1}{t^2}N(t)dt$$

From the conclusion of subsection 1.7.2 we know that $N(t) \ll t\log(t)$ for large t, hence we have

$$\sum_{0<\ell<T} \frac{1}{|\rho|} < \sum_{0<\ell<T} \frac{1}{\ell} \ll (\log(T))^2 \tag{1.52}$$

Hence combining (1.51) and (1.52) we conclude that

$$\sum_{|\ell|<T}\left|\frac{x^{\rho}}{\rho}\right| \ll x\left(\log(T)\right)^{2}\exp\left(-c_1\frac{\log(x)}{\log(T)}\right) \tag{1.53}$$

Now, without loss of generality, we can take x to be an integer. Then using (1.53) and (1.50) in (1.48) of subsection 1.7.3, it follows that

$$|\psi(x)-x| \ll \frac{x(\log(xT))^2}{T} + x(\log(T))^2\exp\left(-c_1\frac{\log(x)}{\log(T)}\right) \tag{1.54}$$

for large x. Next we determine T by a function of x, by equating $(\log(T))^2$ to $\log(x)$ so that

$$\frac{1}{T} = \exp\left(-\sqrt{\log(x)}\right)$$

and using this in (1.54) we get

$$\begin{aligned}|\psi(x)-x| &\ll x(\log(x))^2\exp\left(-\sqrt{\log(x)}\right) + x\log(x)\exp\left(-c_1\sqrt{\log(x)}\right)\\ &\ll x\exp\left(-c_2\sqrt{\log(x)}\right)\end{aligned}$$

provided c_2 is a constant that is less than both 1 and c_1. This proves

$$\boxed{\psi(x) = x + O\left(x\exp\left(-c_2\sqrt{\log(x)}\right)\right)}$$

Now from this we will derive the desired relation for $\pi(x)$. First we consider the function

$$\pi_1(x) = \sum_{n\le x}\frac{\Lambda(n)}{\log(n)} \tag{1.55}$$

This can be expresses in terms of the function $\psi(x)$ by

$$\begin{aligned}\pi_1(x) &= \sum_{n\le x}\Lambda(n)\int_n^x\frac{dt}{t(\log(t))^2} + \frac{1}{\log(x)}\sum_{n\le x}\Lambda(n)\\ &= \int_2^x\frac{\psi(t)}{t(\log(t))^2}dt + \frac{\psi(x)}{\log(x)}\end{aligned}$$

We can replace $\psi(t)$ by t to get

$$\begin{aligned}\int_2^x\frac{1}{(\log(t))^2}dt + \frac{x}{\log(x)} &= \int_2^x t\frac{d}{dt}\left(-\frac{1}{\log(t)}\right)dt + \frac{x}{\log(x)}\\ &= \left[\frac{-t}{\log(t)}\right]_2^x + \int_2^x\frac{1}{\log(t)}dt + \frac{x}{\log(x)}\\ &= \mathrm{Li}(x) + \frac{2}{\log(2)}\end{aligned}$$

Thus we have

$$\pi_1(x) = \mathrm{Li}(x) + \frac{2}{\log(2)} + E(x,t)$$

where

$$|E(x,t)| \ll \int_2^x\exp\left(-c_2\sqrt{\log(t)}\right)dt + x\exp\left(-c_2\sqrt{\log(x)}\right)$$

The contribution of the range $t < x^{\frac{1}{4}}$ to the integral is trivially less than $x^{\frac{1}{4}}$, and in the rest of the range we have $\sqrt{\log(t)} > \frac{1}{2}\sqrt{\log(x)}$.

Hence
$$\pi_1(x) = \mathrm{Li}(x) + O\left(x \exp\left(-c_3\sqrt{\log(x)}\right)\right) \tag{1.56}$$
where $c_3 = \frac{c_2}{2}$.

Using the definition of $\Lambda(n)$ from section 1.4, we can rewrite (1.55) as
$$\begin{aligned}\pi_1(x) &= \sum_{p^m \le x} \frac{\log(p)}{m \log(p)} \\ &= \pi(x) + \frac{1}{2}\pi\left(x^{\frac{1}{2}}\right) + \frac{1}{3}\pi\left(x^{\frac{1}{3}}\right) + \ldots\end{aligned}$$
since $\pi(x) = \sum_{p \le x} 1$ from section 1.5. Also, since $\pi\left(x^{\frac{1}{2}}\right) \le x^{\frac{1}{2}}, \pi\left(x^{\frac{1}{3}}\right) \le x^{\frac{1}{3}}, \ldots$, the difference between $\pi_1(x)$ and $\pi(x)$ is $O(x^{\frac{1}{2}})$. Thus using this in (1.56), we get
$$\boxed{\pi(x) = \mathrm{Li}(x) + O\left(x \exp\left(-c_3\sqrt{\log(x)}\right)\right)}$$
This was proved by de la Vallée Poussin in 1899 [3, §18].

1.8 Some Remarks

1.8.1 Lindelöf Hypothesis

Besides the Riemann hypothesis, another important conjecture involving the Riemann zeta function is the Lindelöf hypothesis, given by Ernst Leonard Lindelöf in 1908. The conjecture states that
$$\zeta\left(\frac{1}{2} + it\right) = O(t^{\varepsilon})$$
for every positive ε. In other words, it's concerned with the growth of the Riemann zeta function on the line $\sigma = \frac{1}{2}$, and conjectures that the modulus of $\zeta(1/2+it)$ grows slower than any positive power t as t tends to infinity.

Riemann hypothesis implies Lindelöf hypothesis, but Lindelöf hypothesis does not imply Riemann hypothesis. It was the Lindelöf hypothesis that lead to the study of moments of Riemann zeta function. The $2k^{\text{th}}$ moment of the modulus of the Riemann zeta function is defined as
$$I_k(T) = \frac{1}{T}\int_0^T \left|\zeta\left(\frac{1}{2} + it\right)\right|^{2k} dt$$
The Lindelöf hypothesis is equivalent to the statement that for any k and any positive ε, $I_k(T) = O(T^{\varepsilon})$.

However, moments are now appriciated in their own right becuase they represent mean values of the Riemann zeta function on the critical line over a finite interval and estimate of these average values can provide information about the zeros of $\zeta(s)$. A nice exposition about the moments of Riemann zeta function is given in Jennifer Beineke and Chris Hughes's article [11].

1.8.2 Elementary Proof

A simple question like "How many primes are there up to x?" deserves a simple answer, one that uses elementary methods rather than all the methods of complex analysis, which may seem far from the question at hand [9, §3]. In 1948, using a fundamental inequality of Atle Selberg,

Selberg and Paul Erdős succeeded in giving an elementary proof of the prime number theorem. The proof is based on Selberg's formula

$$\sum_{p<x}(\log(p))^2 + \sum_{pq<x}(\log(p))(\log(q)) = 2x\log(x) + O(x)$$

which is completely combinatorial in nature. For that proof one may refer [7, Ch. XXII]. Much about this elementary proof can be summarised by quoting following lines by Carl Pomerance [17]

> Thus, far from being an isolated intellectual challenge, the elementary proof of the prime number theorem was a signal that good ideas and strong tools are close at hand. We already had an inkling of this in Riemann's era when Chebyshev used combinatorial methods to show that there is a prime in $[n, 2n]$ for every natural number n. And a century ago, the elementary proof of Brun, stating that most primes are not part of twin-prime pairs, opened the door for combinatorial sieve methods and their many glorious consequences.
>
> Since the elementary proof, some of the most profound and exciting results in the field have had strong elementary and combinatorial leanings. After Roth used the (analytic) circle method to show that dense sets of integers must have 3-term arithmetic progressions, Szemeredi used an elementary (and very complicated) proof to generalize this to k-term arithmetic progressions. This result became an intrinsic tool in the recent Green–Tao proof that the set of primes contains arbitrarily long arithmetic progressions.

The Green-Tao theorem was proved by Ben Green and Terence Tao[23] in 2004 [20][9].

1.8.3 Heuristic Approach

A heuristic technique is any approach to problem-solving that employs a practical method not guaranteed to be optimal or perfect, but sufficient for the immediate goals. Common heuristic techniques are drawing a picture, working backward by assuming a solution, examining a concrete example if problem is too abstract and examining a more general problem if the problem is too specific.

But for the prime number theorem, there is a problem with an attempt at a heuristic explanation because the sieve of Eratosthenes does not behave as one might guess it would from pure probabilistic considerations [9, §2]. Sieving out the composites under x using primes upto $\sqrt{x}$ would lead to (as seen in subsection 1.3.2)

$$x\prod_{p<\sqrt{x}}\left(1-\frac{1}{p}\right)$$

which turns out to be asymptotic to $\frac{2}{e^{\gamma}\log(x)}$, instead of $\frac{x}{\log(x)}$ (proved by Franz Mertens in 1874). Thus the sieve is about 11% more efficient at eliminating composites than one might expect. In 2006, Hugh L. Montgomery and Stan Wagon transformed an old heuristic approach into a proof of following result[16]

> If $x/\pi(x)$ is asymptotic to an increasing function, then $\pi(x) \sim x/\log(x)$.

The proof involves the use of elementary calculus and Tauberian result, instead of complex function theory. Hence, there is a possibility of finding another proof of prime number theorem by proving that $x/\pi(x)$ is asymptotic to an increasing function.

[23] "The primes containing arbitrarily long arithmetic progressions." *Annals of Mathematics* 167, no. 2 (2008), 481–547. http://arxiv.org/abs/math/0404188

Chapter 2

Primes in Arithmetic Progression

"The great fusion between arithmetic and analysis–between counting and measuring, between numbers staccato and numbers legato–came about as the result of an inquiry into prime numbers, conducted by Lejeune Dirichlet in 1830s"

— John Derbyshire, *Prime Obsession*

It is an interesting question whether it is possible to define an infinite number of prime numbers using only addition, subtraction and multiplication. Let's consider the polynomials in one variable whose coefficients are natural numbers. Peter Gustav Lejeune Dirichlet proved that the linear polynomial $ax + b$, where a, b, and x are positive integers with $\gcd(a, b) = 1$, defines an infinite number of prime numbers. It is easy to prove that no non-constant polynomial can generate only prime values, but it is still unknown whether there exists a polynomial of degree 2 or more that generates infinitely many primes [8][23].

2.1 Dirichlet Density

The Dirichlet density, $\mathrm{D}(\mathcal{P})$, of a set $\mathcal{P}$ of prime numbers, if exists, is given by

$$\mathrm{D}(\mathcal{P}) = \lim_{\sigma\to 1} \frac{-1}{\log(\sigma - 1)} \sum_{p\in\mathcal{P}} \frac{1}{p^\sigma}$$

We can draw the following conclusions from this definition [5, §16.1]:

(a) If $\mathcal{P}$ has finitely many elements, then $D(\mathcal{P}) = 0$

(b) If $\mathcal{P}$ consists all but finitely many positive primes, then $D(\mathcal{P}) = 1$

(c) If $\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2$ where $\mathcal{P}_1$ and $\mathcal{P}_2$ are disjoint and $\mathrm{D}(\mathcal{P}_1)$ and $\mathrm{D}(\mathcal{P}_2)$ both exist, then $\mathrm{D}(\mathcal{P}) = \mathrm{D}(\mathcal{P}_1) + \mathrm{D}(\mathcal{P}_2)$.

2.2 Dirichlet Characters

Let m be a fixed positive integer. Let $\chi' : (\mathbb{Z}/m\mathbb{Z})^\times \to \mathbb{C}^\times$ be a homomorphism. Given χ', define $\chi : \mathbb{Z} \to \mathbb{C}^\times$ as follows:

$$\chi(a) = \begin{cases} \chi'(a) & \text{if } \gcd(a, m) = 1 \\ 0 & \text{if } \gcd(a, m) \neq 1 \end{cases}$$

The functions χ defined in this manner are called *Dirichlet characters modulo m*.

2.2.1 Number of Dirichlet Characters

If G is a multiplicatively written finite abelian group, then the *character on* G is a homomorphism from G to $\mathbb{C}^\times$ [28, Chapter 2]. We will denote the set of all such characters by $\widehat{G}$. If $\chi, \chi' \in \widehat{G}$ define $\chi\chi'$ to be the function which takes $g \in G$ to $\chi(g)\chi'(g)$. Then $\chi\chi'$ is also a character. Let's define χ_0 to be the *trivial character*, i.e. $\chi_0(g) = 1$ for all $g \in G$. If $\chi \in \widehat{G}$ define χ^{-1} by $\chi^{-1}(g) = \frac{1}{\chi(g)}$ for all $g \in G$. With these definitions $\widehat{G}$ becomes an abelian group with χ_0 as the identity element.

If n is the order of G then for $g \in G$ we have $g^n = e$, where e is the identity element of G. So if $\chi \in \widehat{G}$ then $(\chi(g))^n = 1$, i.e. $\boxed{\text{the values of } \chi \text{ are } n^{\text{th}} \text{ roots of unity}}$. Therefore $\overline{\chi(g)} = \frac{1}{\chi(g)} = \chi^{-1}(a)$, where bar denotes the complex conjugation. Therefore, χ^{-1} can be written as $\overline{\chi}$ and called the *conjugate character of* χ.

In general, G is a direct product of cyclic groups, i.e. there are elements $g_1, g_2, \ldots, g_t \in G$ such that order of g_k is n_k with $n = n_1 n_2 \cdots n_t$ and every element $g \in G$ can be uniquely written in the form $g = g_1^{m_1} g_2^{m_2} \cdots g_t^{m_t}$ where $0 \le m_k < n_k$ for all k. Hence we can conclude that $\widehat{G}$ is generated by χ_k for all $1 \le k \le t$ such that $\chi_k(g_k) = e^{\frac{2\pi i}{n_k}}$ and $\chi_k(g_\ell) = 1$ for $k \ne \ell$. And hence we have $G \cong \widehat{G}$ where $\widehat{G}$ is the direct product of the cyclic subgroups generated by the χ_k with order of χ_k being n_k [5, §16.3]. Therefore, $|\widehat{G}| = |G| = n$. If we take $G = (\mathbb{Z}/m\mathbb{Z})^\times$ and note that though Dirichlet characters are defined on $\mathbb{Z}$ but are induced by the elements in the character group of $(\mathbb{Z}/m\mathbb{Z})^\times$, we can conclude that:

$$\boxed{\text{There are exactly } \phi(m) \text{ Dirichlet characters modulo } m, \text{ where } \phi(m) = |\,(\mathbb{Z}/m\mathbb{Z})^\times\,|.}$$

2.2.2 Orthogonality Relations

Let $\chi \in \widehat{G}$, then we have $\sum_{g \in G} \chi(g) = n$ for $\chi = \chi_0$. But if $\chi \ne \chi_0$ then there is a $g' \in G$ such that $\chi(g') \ne 1$, and hence

$$\sum_{g \in G} \chi(g) = \sum_{g \in G} \chi(gg') = \chi(g') \sum_{g \in G} \chi(g)$$

So $\sum_{g \in G} \chi(g) = 0$ if $\chi \ne \chi_0$. Therefore for $\chi, \chi' \in \widehat{G}$, we have

$$\sum_{g \in G} \chi\chi'^{-1}(g) = \sum_{g \in G} \chi(g)\overline{\chi'(g)} = \begin{cases} n & \text{if } \chi = \chi' \\ 0 & \text{if } \chi \ne \chi' \end{cases} \tag{2.1}$$

Also for $g \in G$, we have $\sum_{\chi \in \widehat{G}} \chi(g) = n$ if $g = e$. But if $g \ne e$ then $g = g_1^{m_1} g_2^{m_2} \cdots g_t^{m_t}$ with $0 \le m_k < n_k$ for all k and at least one $m_k \ne 0$. Since $\chi(g) = \chi_1(g)^{e_1} \chi_2(g)^{e_2} \cdots \chi_t(g)^{e_t}$ with $0 \le e_k < n_k$, we have $\chi_k(g) = \chi_k(g_k^{m_k}) = \chi_k(g_k)^{m_k} = e^{\frac{2\pi i m_k}{n_k}} \ne 1$, and hence

$$\sum_{\chi \in \widehat{G}} \chi(g) = \sum_{\chi \in \widehat{G}} \chi_k \chi(g) = \chi_k(g) \sum_{\chi \in \widehat{G}} \chi(g)$$

So $\sum_{\chi \in \widehat{G}} \chi(g) = 0$ for $a \ne e$. Therefore for $g, g' \in G$, we have

$$\sum_{\chi \in \widehat{G}} \chi(gg'^{-1}) = \sum_{\chi \in \widehat{G}} \chi(g)\overline{\chi(g')} = \begin{cases} n & \text{if } g = g' \\ 0 & \text{if } g \ne g' \end{cases} \tag{2.2}$$

From (2.1) and (2.2) we conclude that

Let χ and χ' be Dirichlet characters modulo m, and $a, b \in \mathbb{Z}$, then

(a) $$\sum_{a=0}^{m-1} \chi(a)\overline{\chi'(a)} = \begin{cases} \phi(m) & \text{if } \chi = \chi' \\ 0 & \text{if } \chi \neq \chi' \end{cases}$$

(b) $$\sum_{\chi} \chi(a)\overline{\chi(b)} = \begin{cases} \phi(m) & \text{if } a \equiv b \pmod{m} \\ 0 & \text{if } a \not\equiv b \pmod{m} \end{cases}$$

2.3 L-function

2.3.1 Dirichlet L-function

Let χ be a Dirichlet character modulo m. We define the *Dirichlet L-function associated to* χ by the formula

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

where $\boxed{s = \sigma + it}$ is a complex number. Since $|\chi(n)| = 1$ (subsection 2.2.1) we have

$$\left| \frac{\chi(n)}{n^s} \right| \leq \frac{1}{n^s}$$

and hence we see that the terms of $L(s,\chi)$ are dominated in absolute value by the corresponding terms of Euler zeta function $\zeta(s)$ (subsection 1.3.1). Thus $L(s,\chi)$ converges and is analytic for $\sigma > 1$.

2.3.2 Product Formula

Since χ is completely multiplicative we have a product formula for $L(s,\chi)$ in exactly the same way as for $\zeta(s)$ (subsection 1.3.2):

$$\boxed{L(s,\chi) = \prod_{p} \left(1 - \frac{\chi(p)}{p}\right)^{-1}, \quad \sigma > 1}$$

Since $\chi(p) = 0$ for $p|m$ the above product is over positive primes not dividing m. There is a close connection between $L(s,\chi_0)$ and $\zeta(s)$. In fact,

$$\begin{aligned} L(s,\chi_0) &= \prod_{p\nmid m} \left(1 - \frac{1}{p^s}\right)^{-1} \\ &= \prod_{p|m} \left(1 - \frac{1}{p^s}\right) \prod_{p} \left(1 - \frac{1}{p^s}\right)^{-1} \\ &= \prod_{p|m} \left(1 - \frac{1}{p^s}\right) \zeta(s) \end{aligned} \tag{2.3}$$

From this we conclude that

$$\boxed{L(s,\chi_0) \text{ has a pole at } s = 1, \text{ hence } L(1,\chi_0) \neq 0.}$$

2.3.3 Logarithm Formula

The values of $L(s,\chi)$ are in general complex (even if we restrict s to be real). So we choose the principal branch of the logarithm of product formula of $L(s,\chi)$ and then do its series expansion (section 1.4) to get

$$\boxed{G(s,\chi) = \log(L(s,\chi)) = \sum_p \sum_k \frac{\chi\left(p^k\right)}{kp^{ks}}, \quad \sigma > 1}$$

where p runs through all primes, and k through all positive integers. Moreover, we have

$$G(s,\chi) = \sum_{p \nmid m} \frac{\chi(p)}{p^s} + \sum_p \sum_{k=2}^{\infty} \frac{\chi\left(p^k\right)}{kp^{ks}}$$

and using triangle inequality and $(1-x)^{-1} = 1 + x + x^2 + \dots$ for $|x| < 1$, we get

$$\begin{aligned}
\left|\sum_p \sum_{k=2}^{\infty} \frac{\chi\left(p^k\right)}{kp^{ks}}\right| &\leq \sum_p \sum_{k=2}^{\infty} \left|\frac{\chi\left(p^k\right)}{kp^{ks}}\right| \\
&= \sum_p \sum_{k=2}^{\infty} \frac{1}{kp^{ks}} \\
&\leq \sum_p \sum_{k=2}^{\infty} \frac{1}{p^{ks}} \\
&= \sum_p \frac{1}{p^{2s}} \left(1 - \frac{1}{p^s}\right)^{-1} \\
&\leq \left(1 - \frac{1}{2^s}\right)^{-1} \sum_p \frac{1}{p^{2s}} \\
&\leq 2\zeta(2)
\end{aligned}$$

Therefore, we have

$$G(s,\chi) = \sum_{p \nmid m} \frac{\chi(p)}{p^s} + R_\chi(s), \quad \sigma > 1 \tag{2.4}$$

where $R_\chi(s)$ remains bounded around $s = 1$. Now let's multiply both sides of (2.4) by $\overline{\chi(a)}$ where $a \in \mathbb{Z}$ with $\gcd(a,m) = 1$ and then sum over all Dirichlet characters modulo m, to get

$$\sum_\chi \overline{\chi(a)} G(s,\chi) = \sum_{p \nmid m} \frac{1}{p^s} \sum_\chi \overline{\chi(a)}\chi(p) + \sum_\chi \overline{\chi(a)} R_\chi(s)$$

Using the result (b) from subsection 2.2.2, we see

$$\sum_\chi \overline{\chi(a)} G(s,\chi) = \phi(m) \sum_{p \equiv a \pmod m} \frac{1}{p^s} + R_{\chi,a}(s), \quad \sigma > 1 \tag{2.5}$$

where $R_{\chi,a}(s)$ remains bounded around $s = 1$.

Now restricting s to be real, i.e. $t = 0$ and $s = \sigma$, in (2.4) and (2.5) we get

Let $\sigma > 1$ be a real number, then

(a) $G(\sigma, \chi) = \sum_{p \nmid m} \frac{\chi(p)}{p^\sigma} + R_\chi(\sigma)$

where $R_\chi(\sigma)$ remains bounded as $\sigma \to 1$.

(b) $\sum_\chi \overline{\chi(a)} G(\sigma, \chi) = \phi(m) \sum_{p \equiv a \pmod m} \frac{1}{p^\sigma} + R_{\chi,a}(\sigma)$

where $R_{\chi,a}(\sigma)$ remains bounded as $\sigma \to 1$.

2.3.4 Analytic Continuation

We want to analytically continue $L(s, \chi)$ to $\sigma > 0$ from $\sigma > 1$ when χ is a non-trivial Dirichlet character modulo m. We have already seen such analytic continuation of $\zeta(s)$ in subsection 1.3.5 as (1.13). We can use same technique to extend $L(s, \chi)$. Let $A(x) = \sum_{a \le x} \chi(a)$, then by taking $x \to \infty$ in Abel's partial summation (Appendix A), we get

$$\boxed{L(s, \chi) = s \int_1^\infty \frac{A(x)}{x^{s+1}} dx}$$

Since $\chi(a + m) = \chi(a)$ for all $a \in \mathbb{Z}$ ($\chi \neq \chi_0$), we note that if $N = \lfloor x \rfloor = qm + r$ for some $q, r \in \mathbb{Z}$ with $0 \le r < m$ then

$$\begin{aligned} |A(x)| = \left| \sum_{a=0}^{N} \chi(a) \right| &= \left| q \left(\sum_{a=0}^{m-1} \chi(a) \right) + \sum_{a=0}^{r} \chi(a) \right| \\ &= \left| \sum_{a=0}^{r} \chi(a) \right| \\ &\le \sum_{a=0}^{m-1} |\chi(a)| \\ &= \phi(m) \end{aligned}$$

using the result (a) of subsection 2.2.2 which says $\sum_{a=0}^{m-1} \chi(a) = 0$ and that $|\chi(a)| = 1$. From this we conclude that $|A(x)| \le \phi(m)$ for all x, and hence we get that the above integral converges. Therefore if χ is non-trivial Dirichlet character modulo m, then $L(s, \chi)$ can be continued to a analytic function in the region $\{s \in \mathbb{C} | \sigma > 0\}$.

2.3.5 Product of L-functions

Let $\boxed{F(s) = \prod_\chi L(s, \chi)}$, where the product is over all Dirichlet character modulo m. Assume s is real and $s > 1$, i.e. $s = \sigma > 1$. From subsection 2.3.3 we know that

$$G(\sigma, \chi) = \sum_p \sum_{k=1}^{\infty} \frac{\chi(p^k)}{k p^{k\sigma}}$$

Summing over χ and using the conclusion (b) of subsection 2.2.2, we get

$$\sum_\chi G(\sigma, \chi) = \phi(m) \sum_p \sum_k \frac{1}{k p^{k\sigma}} \tag{2.6}$$

where the sum is over all primes p and integers k such that $p^k \equiv 1 \pmod m$. The right-hand side of (2.6) is non-negative, hence taking the exponential of both sides we get

For s real and $s > 1$, i.e. $s = \sigma > 1$ we have $F(s) = F(\sigma) \ge 1$.

2.4 The Proof

Let $\mathcal{P}(a;m)$ be the set of prime numbers p such that $p \equiv a \pmod{m}$. We wish to prove that $\mathcal{P}(a;m)$ has infinite number of elements. We will divide the proof into two parts, while restricting s to real, i.e. $t = 0$ we will write $s = \sigma$ [5, §16.4].

2.4.1 Trivial Dirichlet Character

Let χ_0 denote the trivial character modulo m, then $L(\sigma, \chi_0)$ is a real function of positive real numbers. From (2.3) it follows that

$$G(\sigma, \chi_0) = \sum_{p|m} \log\left(1 - \frac{1}{p^\sigma}\right) + \log(\zeta(\sigma)) \tag{2.7}$$

As seen in subsection 1.3.1, for $s = \sigma$,

$$\lim_{\sigma \to 1} (\sigma - 1)\zeta(\sigma) = 1$$

which implies that [5, §16.1]

$$\lim_{\sigma \to 1} \frac{-\log(\zeta(\sigma))}{\log(\sigma - 1)} = 1 \tag{2.8}$$

Using (2.8) in (2.7) we get

$$\boxed{\lim_{\sigma \to 1} \frac{-G(\sigma, \chi_0)}{\log(\sigma - 1)} = 1}$$

2.4.2 Non-trivial Dirichlet Character

Let χ be a non-trivial Dirichlet character modulo m. We need to ensure that $L(1, \chi) \neq 0$, and to prove this we will consider following two cases.

Complex Character

Let χ be a complex character modulo m, i.e. a character which takes only non-real values. From the series defining $L(s, \chi)$ we see that for $s = \sigma$ real, $\sigma > 1$,

$$\overline{L(\sigma, \chi)} = L(\sigma, \overline{\chi})$$

Assume $L(1, \chi) = 0$, then $L(1, \overline{\chi}) = 0$. Hence the functions $L(\sigma, \chi)$ and $L(\sigma, \overline{\chi})$ are distinct and both have a zero at $\sigma = 1$. In the product $F(\sigma) = \prod_\chi L(\sigma, \chi)$ we know $L(\sigma, \chi_0)$ has a simple pole at $\sigma = 1$ (subsection 2.3.2) and all other factors are analytic about $\sigma = 1$ (subsection 2.3.4). It follows that $F(1) = 0$. But from subsection 2.3.5 we know that $F(\sigma) \geq 1$ for all $\sigma > 1$. This contradicts our assumption. Therefore $L(1, \chi) \neq 0$ when χ is a non-trivial complex character modulo m.

Real Character

Let χ be a non-trivial real character, i.e. $\chi(a) = 0, 1$ or -1 for all $a \in \mathbb{Z}$. We will make use of following result whose proof is analogous to the proof of Euler's product formula (subsection 1.3.2) [5, §16.5]

> Suppose f is a non-negative, multiplicative function of $\mathbb{Z}^+$, i.e. for all $a, b > 0$ with $\gcd(a, b) = 1$, $f(ab) = f(a)f(b)$. Assume there is a constant c such that $f\left(p^k\right) < c$ for all prime powers p^k. Then $\sum_{n=1}^{\infty} f(n)n^{-s}$ converges for all $\sigma > 1$. Moreover
>
> $$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_p \left(1 + \sum_{k=1}^{\infty} \frac{f\left(p^k\right)}{p^{ks}}\right) \tag{2.9}$$

Assume $L(1,\chi)=0$ and consider the function

$$\omega(s)=\frac{L(s,\chi)L(s,\chi_0)}{L(2s,\chi_0)} \tag{2.10}$$

The zero of $L(s,\chi)$ at $s=1$ cancels the simple pole of $L(s,\chi_0)$ so the numerator is analytic on $\sigma>0$ (subsection 2.3.4). The denominator is non-zero and analytic for $\sigma>\frac{1}{2}$. Thus $\omega(s)$ is analytic on $\sigma>\frac{1}{2}$ (compare with subsection 1.3.5). Moreover, since $L(2s,\chi_0)$ has a pole at $s=\frac{1}{2}$ (subsection 2.3.2) we have for real s, i.e. $s=\sigma$, $\omega(\sigma)\to 0$ as $\sigma\to\frac{1}{2}$.

We can re-write (2.10) to get the infinite product expansion (subsection 2.3.2)

$$\begin{aligned}\omega(s)&=\prod_p\left(1-\frac{\chi(p)}{p^s}\right)^{-1}\left(1-\frac{\chi_0(p)}{p^s}\right)^{-1}\left(1-\frac{\chi_0(p)}{p^{2s}}\right)\\&=\prod_{p\nmid m}\left(1-\frac{\chi(p)}{p^s}\right)^{-1}\left(1-\frac{1}{p^s}\right)^{-1}\left(1-\frac{1}{p^{2s}}\right)\\&=\prod_{p\nmid m}\frac{(1+p^{-s})}{(1-\chi(p)p^{-s})}\end{aligned}$$

If $\chi(p)=-1$ the $p-$factor is equal to 1. Thus

$$\omega(s)=\prod_{\chi(p)=1}\frac{1+p^{-s}}{1-p^{-s}} \tag{2.11}$$

where the product over all p such that $\chi(p)=1$. Moreover,

$$\begin{aligned}\frac{1+p^{-s}}{1-p^{-s}}&=\left(1+\frac{1}{p^s}\right)\left(\sum_{k=0}^{\infty}\frac{1}{p^{ks}}\right)\\&=1+\frac{2}{p^s}+\frac{2}{p^{2s}}+\dots\end{aligned}$$

Using this and applying (2.9) in (2.11) we find that

$$\omega(s)=\sum_{k=1}^{\infty}\frac{a_k}{k^s} \tag{2.12}$$

where $a_k\geq 0$ and the series converges for $\sigma>1$. Note that $a_1=1$.

Since $\omega(s)$ is analytic for $\sigma>\frac{1}{2}$, expanding $\omega(s)$ in a power series about $s=2$ using Taylor's theorem [1, §2.43], we get

$$\omega(s)=\sum_{j=0}^{\infty}\frac{\omega^{(j)}(2)}{j!}(s-2)^j \tag{2.13}$$

where $\omega^{(j)}(s)$ is the j^{th} derivative of $\omega(s)$ with the radius of convergence at least $\frac{3}{2}$ [1, §7.1].

But from (2.12) we get

$$\omega^{(j)}(2)=\sum_{k=1}^{\infty}\frac{a_k(-\log(k))^j}{k^2}=(-1)^jc_j \tag{2.14}$$

where $c_j\geq 0$. Now using (2.14) in (2.13) we get

$$\omega(s)=\sum_{j=0}^{\infty}c_j(2-s)^j$$

where c_j is non-negative and

$$c_0 = \omega(2) = \sum_{k=1}^{\infty} \frac{a_k}{k^2} \geq a_1 = 1$$

It follows that for real s, i.e. $s = \sigma$, in the interval $\frac{1}{2} < \sigma < 2$ we have $\omega(\sigma) \geq 1$. This contradicts our assumption that $\omega(\sigma) \to 0$ as $\sigma \to \frac{1}{2}$, and so $L(1, \chi) \neq 0$.

Combining the above two cases, we can say that

> If χ is a non-trivial Dirichlet character modulo m, then $G(\sigma, \chi)$ remains bounded as $\sigma \to 1$.

This was proved by de la Vallée Poussin[1] in 1896 [3, §4].

2.4.3 Completing the Proof

Now we simply divide all terms on both sides of the conclusion (b) of subsection 2.3.3 by $-\log(\sigma - 1)$ and take the limit $\sigma \to 1$

$$\lim_{\sigma \to 1} \sum_{\chi} \overline{\chi(a)} \frac{-G(\sigma, \chi)}{\log(\sigma - 1)} = \phi(m) \lim_{\sigma \to 1} \frac{-1}{\log(\sigma - 1)} \sum_{p \in \mathcal{P}(a;m)} \frac{1}{p^\sigma} + \lim_{\sigma \to 1} \frac{-R_{\chi,a}(\sigma)}{\log(\sigma - 1)}$$

From the results of subsection 2.4.1 and subsection 2.4.2, we get the limit on left-hand side is 1 (because all except trivial-character will vanish) whereas the limit on the right-hand side is $\phi(m)\mathrm{D}(\mathcal{P}(a;m))$ (section 2.1). Thus $\mathrm{D}(\mathcal{P}(a;m)) = 1/\phi(m)$ and we are done.

> If $a, m \in \mathbb{Z}$, with $\gcd(a, m) = 1$, then
>
> $$\mathrm{D}(\mathcal{P}(a;m)) = \frac{1}{\phi(m)}$$
>
> where $\phi(m)$ is the number of positive integers n, less than m, with $\gcd(n, m) = 1$.

Since $D(\mathcal{P}(a;m)) \neq 0$, this implies that $\mathcal{P}(a;m)$ has infinite number of elements.

2.5 Some Remarks

2.5.1 Chebotarëv Density Theorem

Chebotarëv's density theorem may be regarded as the least common generalization of Dirichlet's theorem on primes in arithmetic progressions (1837) and a theorem of Frobenius (1880; published 1896) [28, §1.1]. Generally, Frobenius's theorem for $f(X) = X^m - 1$ is implied by Dirichlet's theorem for the same m, but not conversely. One can formulate a sharper version of Frobenius's theorem that for $f(X) = X^m - 1$ does come down to Dirichlet's theorem. To do this we need do something called *Frobenius substitution of* p, σ_p. Once the Frobenius substitution has been defined, one can wonder about the density of the set of primes p for which σ_p is equal to a given element of the Galois group G of $f(X)$, corresponding to the field K generated by the zeros of $f(X)$. Note that the Frobenius map Frob_p is an automorphism of the field $\mathbb{F}_p$ of characteristic p, and the Frobenius substitution σ_p is going to be an automorphism of the field K of characteristic zero.

Theorem (Reformulation of Dirichlet's Theorem)**.** *If $f(X) = X^m - 1$ for some positive integer m, then the set of prime numbers p for which σ_p is equal to a given element of the Galois group G of $f(X)$ has a density, and this density equals $1/|G|$.*

[1] "Recherches analytiques sur la théorie des nombres premiers." Deuxieme partie. *Ann. Soc. Sci. Bruxelles.* 20, 281-362 (1896).

Thus the Frobenius substitution is equidistributed over the Galois group if p varies over all primes not dividing m. This leads to the desired generalization of the theorems of Dirichlet and Frobenius. It was formulated as a conjecture by Ferdinand Georg Frobenius, and was ultimately proved by Nikolaĭ Chebotarëv in 1922. Chebotarëv's theorem extends this to all $f(X)$.

Theorem (Chebotarëv's Density Theorem)**.** *Let $f(X)$ be a polynomial with integer coefficients and with leading coefficient 1. Assume that the discriminant $\Delta(f)$ of $f(X)$ does not vanish. Let C be a conjugacy class of the Galois group G of $f(X)$. Then the set of primes p not dividing $\Delta(f)$ for which σ_p belongs to C has a density, and this density equals $|C|/|G|$.*

For details refer to the article by P. Stevenhagen and H. W. Lenstra Jr.[14].

2.5.2 Prime Number Theorem for Arithmetic Progressions

We can combine the the results in the two chapters to talk about *quantitative* aspects of distribution of primes in arithmetic progression

$$\pi(x; m, a) \sim \frac{\pi(x)}{\phi(m)}$$

The analytic proof of this statement is analogous to that of prime number theorem, with $\zeta(s)$ replaced by $L(s, \chi)$. For that proof one can refer to Davenport's book [3, §20]. For discussion regarding the elementary proof on may refer to this MathOverflow discussion initiated by Qiaochu Yuan: *Is a "non-analytic" proof of Dirichlet's theorem on primes known or possible?*, URL (version: 2011-09-29): https://mathoverflow.net/q/16735.

One can make enormous improvements in the error term of this estimate, provided there are no *Siegel zeros*. And the Siegel zeros are rare as a consequence of *Deuring-Heilbronn phenomenon*, which states that, the zeros of L-functions repel each other just like in Roth's theorem [26] different algebraic numbers repel each other. There is a close relation between Siegel zeros and *class numbers* [27, §4.1]. For details refer to Andrew Granville's article [9].

2.5.3 Tchébychev Bias

This phenomenon was first observed in a letter written by Pafnuty Tchébychev to M. Fuss on 23 March 1853:

> There is a notable difference in the splitting of the prime numbers between the two forms $4k + 3$, $4k + 1$: the first form contains a lot more than the second.

This bias is perhaps unexpected because as per the the prime number theorem for arithmetic progressions discussed above, the primes tend to be equally split amongst the various forms $p \equiv a \pmod{m}$ with $\gcd(a, m) = 1$ for any given modulus m. Hence we have

$$\pi(x; 4, 1) \sim \pi(x; 4, 3) \sim \frac{x}{2\log(x)}$$

i.e. half the primes are of the form $4k + 1$, and half of the form $4k + 3$. This asymptotic result does not inform us about any of the fine details of these prime number counts, so neither verifies nor contradicts the observation that $\pi(x; 4, 3) > \pi(x; 4, 1)$.

But as we count the number of primes for various values of x, we observe that from time to time, more primes of the form $4k + 1$ than of the form $4k + 3$, but this lead is held only very briefly and then relinquished for a long stretch. Now one might guess that $4n + 1$ will occasionally take the lead as we continue to watch for bigger x. Indeed this is the case, as John Edensor Littlewood discovered in 1914:

Theorem (Littlewood, 1914)**.** *There are arbitrarily large values of x for which there are more primes of the form $4k + 1$ up to x than primes of the form $4k + 3$. In fact there are arbitrarily large values of x for which*

$$\pi(x; 4, 1) - \pi(x; 4, 3) \geq \frac{\sqrt{x}\log(\log(\log(x)))}{2\log(x)}$$

At first sight, this seems to be the end of the story. But in 1962, Stanislaw Knapowski and Pál Turán made a conjecture that is consistent with Littlewood's result but also bears out Tchébychev's observation:

> As $X \to \infty$, the percentage of integers $x \leq X$ for which there are more primes of form $4k + 3$ up to x than of the form $4k + 1$ goes to 100%

This conjecture may be paraphrased as "Tchébychev was correct *almost* all of the time." For details and recent develpoments refer to the article by Andrew Granville and Greg Martin [15].

Conclusion

> *"Riemann discovered that prime numbers too could be studied by harmonic analysis, albeit of a slightly different kind. He realized that the psi function can be thought of as a sum of elementary waveforms;...... To me, that the distribution of prime numbers can be so accurately represented in a harmonic analysis is absolutely amazing and incredibly beautiful. It tells of an arcane music and a secret harmony composed by the prime numbers."*
>
> — Enrico Bombieri, *Prime Territory*

One very important tool for our understanding of primes, that I couldn't discuss in this report, is the computational tool which enables us to establish faith in the conjectures that we are trying to prove. For example, Alan Turing[2] (1912–1954) was very much interested in disproving Riemann hypothesis by finding a counter-example using an automatic computer. For details about Turing's attempts in this direction refer to the lecture by Yuri Matiyasevich [22].

As pointed in the beginning of second chapter, there can't be any polynomial in one variable that gives only prime number values. But we have following remarkable result by Julia Robinson and Hilary Putnam:

> There exists an exponential polynomial $R(x_0, \ldots, x_k)$ such that all of its positive values for positive integer variables are prime numbers, and every prime number is so presented.

One can find a good discussion on this theorem by Yuri Matiyasevich [8], who in 1971 extended this result[23]. This result played an important role in solving Hilbert's Tenth Problem whose statement I have discussed earlier [24, Conclusion]. But this result is of no practical use, since it will be computationally difficult to calculate prime numbers by solving this system of equations.

Luckily we have better ways of determining whether a given number is a prime or not. In 1970s, Gary Miller[3] showed how a generalization of the Riemann Hypothesis allows a deterministic, polynomial-time procedure for recognizing primes, a few decades later in 2002, Manindra Agrawal, Neeraj Kayal, and Nitin Saxena[4] showed the same with completely elementary (and rigorous) methods [17]. But, so far, this new polynomial-time primality test is not good in practice. But there exist ways for primanlity testing, like using the arithmetic of elliptic curves, which are conjectured to run in polynomial time but we have not even proved that they always terminates [10].

Practical importance of prime numbers lies in the modern public key encryption [25] which is based on the belief that there can't exist a polynomial-time algorithm for factorising composite numbers (unlike the efficient Euclid's algorithm for finding the greatest common divisor). Most practical factoring algorithms are based on unproved but reasonable-seeming hypotheses about

[2] He is widely considered to be the father of theoretical computer science and artificial intelligence.

[3] "Riemann's Hypothesis and tests for primality." *Proceedings of the seventh annual ACM symposium of Theory of Computing* 1975, pp. 234–239. doi:10.1145/800116.803773

[4] "PRIMES is in P." *Annals of Mathematics* 160, no. 2 (2004), 781–793. doi: 10.4007/annals.2004.160.781.

the natural numbers. Although we may not know how to prove rigorously that these methods will always produce a factorisation, or do so quickly, in practice they do. Broadly speaking there are two types of factorisation methods, sieve-based methods [12] and elliptic curve method. A nice exposition about computational aspects of analytic number theory (the distribution of primes and the Riemann hypothesis), Diophantine equations (Fermat's last theorem and the abc conjecture) and elementary number theory (primality and factorisation) can be found in Carl Pomerance's article [10].

I will end my report by stating the two recent developments towards our understanding of prime numbers:

- In 2013, Yitang Zhang[5] proved that

$$\liminf_{n\to\infty}(p_{n+1}-p_n) < 7\times 10^6$$

  where p_n denotes the n^{th} prime. This was a result of Zhang's attempt to prove the *Twin Primes conjecture*. The proof was inspired by idea of Bombieri-Friedlander-Iwaniec[6] and Goldston-Pintz-Yíldírím[7] sieve (which makes the connection between prime gaps and primes in arithmetic progression). Few months later James Maynard and Terence Tao (independently) gave a different simpler approach than Zhang's which reduced the bounded gaps between primes to 246 [23]. One interesting aspect of the better bounds is that the best bounds were obtained by using important theorems in Algebraic Geometry, what are known as the 'Weil Conjectures', which at first glance appear to be far removed from Twin Primes [18].

- In 2016, Kannan Soundararajan and Robert Lemke Oliver[8] discovered that that primes seem to avoid being followed by another prime with the same final digit. They presented both numerical and theoretical evidence that prime numbers repel other would-be primes that end in the same digit, and have varied predilections for being followed by primes ending in the other possible final digits [19].

[5] "Bounded gaps between primes." *Annals of Mathematics* 179, no. 3 (2014), 1121–1174. doi: 10.4007/annals.2014.179.3.7

[6] Bombieri, E., Friedlander, J. B. and Iwaniec, H. "Primes in arithmetic progressions to large moduli." *Acta Mathematica* 156 (1986), 203–251. doi:10.1007/BF02399204. http://projecteuclid.org/euclid.acta/1485890416.

[7] Goldston, D. A., Pintz, J. and Yíldírím, C. Y. "Primes in Tuples I." *Annals of Mathematics* 170, no. 2 (2009): 819–862. doi:10.4007/annals.2009.170.819. https://arxiv.org/abs/math/0508185

[8] "Unexpected Biases in the Distribution of Consecutive Primes." *Proceedings of the National Academy of Sciences* 113, no. 31 (2016): E4446–E4454. doi:10.1073/pnas.1605366113. https://arxiv.org/abs/1603.03720

Appendix A

Abel's Summation Formula

This formula is based on our understanding of Riemann–Stieltjes integral. We will follow [7, §22.5].

Theorem. *Suppose that $a_1, a_2, \ldots$ is a sequence of numbers, such that $A(t) = \sum_{n \le t} a_n$ and $f(t)$ is any function of t. Then*

$$\sum_{n \le x} a_n f(n) = \sum_{n \le x-1} A(n)\Big(f(n) - f(n+1)\Big) + A(x)f(\lfloor x \rfloor) \tag{*}$$

If, in addition, $a_j = 0$ for all j less than some positive integer m and $f(t)$ has a continuous derivative for $t \ge m$, then

$$\sum_{n \le x} a_n f(n) = A(x)f(x) - \int_m^x A(t)f'(t)dt \tag{**}$$

Proof. We rewrite the left-hand-side of (*) as

$$\begin{aligned}
\sum_{n \le x} a_n f(n) &= A(1)f(1) + \Big(A(2) - A(1)\Big)f(2) + \ldots + \Big(A(\lfloor x \rfloor) - A(\lfloor x \rfloor - 1)\Big)f(\lfloor x \rfloor) \\
&= A(1)\Big(f(1) - f(2)\Big) + \ldots + A(\lfloor x \rfloor - 1)\Big(f(\lfloor x \rfloor - 1) - f(\lfloor x \rfloor)\Big) + A(\lfloor x \rfloor)f(\lfloor x \rfloor)
\end{aligned}$$

Since $A(\lfloor x \rfloor) = A(x)$, this proves (*). To deduce (**) we observe that $A(t) = A(n)$, a constant, when $n \le t < n+1$ and so

$$A(n)\Big(f(n) - f(n+1)\Big) = -\int_n^{n+1} A(t)f'(t)dt$$

Also $A(t) = 0$ when $t < m$. Hence proving (**). $\square$

If we put $a_n = 1$ and $f(t) = 1/t$, we have $A(x) = \lfloor x \rfloor$ and (**) becomes

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

Observe that

$$
\begin{aligned}
\left|\int_1^x \frac{\{t\}}{t^2} dt\right| &\leq \int_1^x |\{t\}| \frac{1}{t^2} dt \\
&\leq \int_1^x \frac{1}{t^2} dt \\
&= 1 - \frac{1}{x}
\end{aligned}
$$

Hence the integral $\displaystyle\int_1^x \frac{\{t\}}{t^2} dt$ converges to a limit as $x \to \infty$. Thus we can write

$$
\int_1^x \frac{\{t\}}{t^2} dt = \int_1^\infty \frac{\{t\}}{t^2} dt - \int_x^\infty \frac{\{t\}}{t^2} dt
$$

Therefore, we have

$$
\begin{aligned}
\sum_{n \leq x} \frac{1}{n} &= \log(x) + \left(1 - \int_1^\infty \frac{\{t\}}{t^2} dt\right) + \left(\int_x^\infty \frac{\{t\}}{t^2} dt - \frac{\{x\}}{x}\right) \\
&= \log(x) + \gamma + \left(\int_x^\infty \frac{O(1)}{t^2} dt + O\left(\frac{1}{x}\right)\right)
\end{aligned}
$$

Therefore

$$
\boxed{\sum_{n \leq x} \frac{1}{n} = \log(x) + \gamma + O\left(\frac{1}{x}\right)}
$$

where γ is the *Euler-Mascheroni constant* given by

$$
\gamma = \lim_{N \to \infty} \left(1 + \frac{1}{2} + \ldots + \frac{1}{N} - \log N\right) \approx 0.57721
$$

Bibliography

[1] Titchmarsh, E. C. *The Theory of Functions*. London: Oxford University Press, 1939. https://archive.org/details/TheTheoryOfFunctions

[2] Ayoub, R. *An Introduction to the Analytic Theory of Numbers* (Mathematical Surveys and Monographs 10), Providence: American Mathematical Society, 1963. http://doi.org/10.1090/surv/010

[3] Davenport, H. *Multiplicative Number Theory*. New York: Springer-Verlag, 1980. http://doi.org/10.1007/978-1-4757-5927-3

[4] Titchmarsh, E. C. *The Theory of the Riemann Zeta-Function*. Oxford: Oxford University Press, 1986.

[5] Ireland, K. and Rosen, M. *A Classical Introduction to Modern Number Theory*. New York: Springer-Verlag, 1990. http://doi.org/10.1007/978-1-4757-2103-4

[6] Derbyshire, J. *Prime Obsession: Bernhard Riemann and the Greatest Unsolved Problem in Mathematics*. Washington, D.C.: Joseph Henry Press, 2003. http://www.johnderbyshire.com/Books/Prime/page.html

[7] Hardy, G. H. and Wright E. M. *An Introduction to the Theory of Numbers*. Oxford: Oxford University Press, 2008.

[8] Matiyasevich, Y. "Formulas for Prime Numbers." in *Kvant Selecta: Algebra and Analysis - II*, edited by S. Tabachnikov, pp. 13–24. Providence: The American Mathematical Society, 1999 (The Russian original was published in *Kvant* 1975. no. 5. pp. 5-13. http://kvant.mccme.ru/1975/05/formuly_dlya_prostyh_chisel.htm; Translated by N. K. Kulman)

[9] Granville, A. "Analytic Number Theory." in *The Princeton Companion for Mathematics*, edited by T. Gowers, J. Barrow-Green and I. Leader, pp. 332–348. Princeton and Oxford: Princeton University Press, 2008. Preprint available at http://www.dms.umontreal.ca/~andrew/PDF/PrinceComp.pdf

[10] Pomerance, C. "Computational Number Theory." in *The Princeton Companion for Mathematics*, edited by T. Gowers, J. Barrow-Green and I. Leader, pp. 348–362. Princeton and Oxford: Princeton University Press, 2008. Preprint available at https://math.dartmouth.edu/~carlp/PDF/pcm0049.pdf

[11] Beineke, J. and Hughes, C. "Great Moments of the Riemann zeta Function." in *Biscuits of Number Theory*, edited by A. T. Benjamin and E. Brown, pp. 199–215. Providence: The Mathematical Association of America, 2009.

[12] Pomerance, C. "The Search for Prime Numbers." *Scientific American* 247, no. 6 (1982), 136–147. doi: 10.1038/scientificamerican1282-136. https://math.dartmouth.edu/~carlp/PDF/search.pdf

[13] Bombieri, E. "Prime Territory." *The Sciences* 32, no. 5 (1992), 30–36. http://doi.org/10.1002/j.2326-1951.1992.tb02416.x

[14] Stevenhagen, P. and Lenstra, H. W. "Chebotarëv and His Density Theorem." *The Mathematical Intelligencer* 18, no. 2 (1996), 26–37. http://dx.doi.org/10.1007/bf03027290

[15] Granville, A. and Martin, G. "Prime Number Races." *The American Mathematical Monthly* 113, no. 1 (2006), 1–33. http://doi.org/10.2307/27641834

[16] Montgomery, H. L. and Wagon, S. "A Heuristic for the Prime Number Theorem." *The Mathematical Intelligencer* 28, no. 3 (2006), 6–9. http://doi.org/10.1007/BF02986877

[17] Spencer, J. and Graham, R. "The Elementary proof of Prime Number Theorem." *The Mathematical Intelligencer* 31, no. 3 (2009), 18–23. http:/doi.org/10.1007/s00283-009-9063-9.

[18] Sreekantan, R. "Yitang Zhang and The Twin Primes Conjecture." *At Right Angles* 2, no. 3 (2013), 14–17. http://teachersofindia.org/en/article/yitang-zhang-and-twin-primes-conjecture

[19] Klarreich, E. "Mathematicians Discover Prime Conspiracy." *Quanta Magazine* March 13, 2016. https://www.quantamagazine.org/mathematicians-discover-prime-conspiracy-20160313/

[20] Tao, T. (22 Jan 2009) *Structure and Randomness in the Prime Numbers* [Video File]. University of California, Los Angeles. Retrieved from https://www.youtube.com/watch?v=PtsrAw1LR3E. Lecture slides available at https://www.math.ucla.edu/~tao/preprints/Slides/primes.pdf

[21] Taylor, R. (27 April 2012) *Primes and Equations* [Video file]. Institute for Advanced Study, Princeton. Retrieved from https://www.youtube.com/watch?v=5pTaZu3C--s

[22] Matiyasevich, Y. (10 Dec 2012) *Alan Turing and Number Theory* [Video File]. Alan Turing Centenary Conference Manchester, 2012. Retrieved from https://www.youtube.com/watch?v=5AD0T72v_9U. Lecture slides available at http://videolectures.net/turing100_matiyasevich_number_theory/

[23] Granville, A. (20 August 2015) *The Patterns in the Primes* [Video file]. Yale University, New Haven. Retrieved from https://www.youtube.com/watch?v=p07Egc5Dtqs. Lecture slides available at http://www.dms.umontreal.ca/~andrew/PrimeLect.html

[24] Korpal, G. "Diophantine Equations." *Summer Internship Project Report*, guided by Prof. S. A. Katre (18 May 2015 – 16 June 2015)

[25] Korpal, G. "Enigma Cryptanalysis." *Summer Internship Project Report*, guided by Prof. Geetha Venkataraman (6 July 2015 – 26 July 2015)

[26] Korpal, G. "Diophantine Approximations." *Winter Internship Project Report*, guided by Prof. R. Thangadurai (13 December 2015 – 8 January 2016)

[27] Korpal, G. "Number Fields." *Summer Internship Project Report*, guided by Prof. Ramesh Sreekantan (1 June 2016 – 31 July 2016)

[28] Korpal, G. "Reciprocity Laws." *Winter Internship Project Report*, guided by Prof. Chandan Dalawat (9 December 2016 – 7 January 2017)

Prepared in $\LaTeX\,2_\varepsilon$ by Gaurish Korpal