UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Ciencias e Ingenierías

Analytic Number Theory
and
The Prime Number Theorem

Ricardo David Mayorga Bayas

Matemáticas

Trabajo de titulación presentado como requisito para la obtención del título de Matemático

Quito, 12 de Diciembre de 2023

UNIVERSIDAD SAN FRANCISCO DE QUITO

Colegio de Ciencias e Ingenierías

HOJA DE CALIFICACIÓN DEL

TRABAJO DE TITULACIÓN

Analytic Number Theory
and
The Prime Number Theorem

Ricardo David Mayorga Bayas

Calificación: ______________________________

Nombre del Profesor: John R. Skukalek,

Título Académico: Ph.D.

Firma: ______________________________

Quito, 12 de Diciembre de 2023

© Derechos de Autor

Firma del estudiante: ______________________________

Nombres y apellidos: Ricardo David Mayorga Bayas

Código: 00212550

Cédula de identidad: 1804230892

Lugar y fecha: Quito, 12 de diciembre de 2023

ACLARACIÓN PARA PUBLICACIÓN

UNPUBLISHED DOCUMENT

Dedicado a

Mis padres,

mi novia,

por siempre estar para mí.

Gracias.

Resumen

En esta tesis, exploramos algunos de los conceptos y resultados más elementales de la teoría analítica de números, dirigidos hacia una demostración elemental del teorema de los números primos. En primer lugar, definiremos nociones y conceptos básicos con el fin de derivar relaciones asintóticas útiles y examinar la conexión entre ciertas funciones aritméticas y la distribución de números primos entre los enteros positivos. El teorema de los números primos se presenta siguiendo la demostración que Selberg publicó en 1949, y también discutiremos una generalización del teorema de los números primos.

Abstract

In this thesis, we explore some of the most elementary concepts and results of analytic number theory that are directed towards an elementary proof of the prime number theorem. We will first define basic notions and concepts in order to derive useful asymptotic relations and examine the connection between certain arithmetical functions and the distribution of prime numbers among the positive integers. The prime number theorem is presented followed by the proof that Selberg published in 1949 and we will also discuss a generalization of the prime number theorem.

Agradecimiento

A mis padres, que me han dado todo lo que he necesitado y más, que han estado ahí para mí y me han apoyado tanto en mi vida universitaria como personal. Agradezco a mi familia por hacer todo esto posible.

A mi novia, por ser una mujer maravillosa, que me ha apoyado desde el día cero en todas mis aventuras y desventuras, que ha puesto tachuelas en mis zapatos para que me acuerde que voy caminando; estaré siempre en deuda con ella.

A John R. Skukalek, persona y profesor excelente que ha sabido despertar mi curiosidad y amor por la matemática, estaré eternamente agradecido por la confianza que ha depositado en mí y por todas las clases que compartimos. También agradezco al resto de profesores de la USFQ que me han ayudado de sobremanera durante toda la carrera, especialmente a Nicola di Teodoro.

A mi gata, que me ha acompañado maullando y roncando incontables noches de desvelo a mi lado, que me ha dado la fuerza, como el resto de las personas que aquí agradezco, de seguir a pesar de las circunstancias.

Agradezco a mis compañeros de carrera, Pablo Padilla, José Palacios, Adrián Vásconez y a todas las personas que me han apoyado para cultivarme intelectual y personalmente.

Agradezco también a todas las personas que no han sido mencionadas, pero que me han hecho la vida más fácil y feliz, todos tienen un lugar en mis mejores memorias y en mi espíritu.

Gracias a ustedes.

Contents

Introduction

Mathematics is the queen of the sciences, and

number theory is the queen of mathematics.

- Carl Friedrich Gauss

Number theory, at a first glance, seems to be one of the most natural, intuitive and simple branches of mathematics, yet is one of the most puzzling and interesting ones to the extent that most of the greatest mathematicians have devoted hundreds or thousand of pages to unveil its secrets. We will dive into the classical question of how are primer numbers distributed and discuss about the analytical and elementary proofs of the Prime Number Theorem (PNT), its history and its importance for mathematics.

This work will be divided into x chapters, the first one regarding the history that precedes this problem and well-known results regarding prime numbers. The second chapter will define the necessary notation and present modern theorems from elementary number theory that, in their majority, won't be proved here but will be outsourced since they are either simple or out of the scope of this thesis. From the third chapter to the sixth one, we will prove various statements equivalent to the PNT and we will have a deep overview the first elementary proof of it, related to what's called Selberg's Asymptotic formula. Finally, in the last three chapter, we will discuss a sketch of the proof of the PNT and review the classical analytical proof, as well as enumerate related results and resources to learn more about this an related topics.

Chapter 1

Preliminaries

Natural numbers were created by God,

everything else is the work of men.

- Leopold Kronecker

1.1 Historical note on the study of numbers

Natural numbers, also called counting numbers or positive integers have been studied by possibly all known human civilizations since the comprehension of arithmetic operations is fundamental for any considerable population growth to happen. The capacity of man to count is essential for predicting climate seasons, organizing crop production and living stock, and, perhaps the most critical, commerce. We will use the (not uncommon) definition that the set of natural numbers "starts" with 1, specially useful for this field of mathematics, and denote it by $\mathbb{N}$.

Even though numbers, counting, adding, multiplying, subtracting and dividing are a great tool for many necessary tasks, they themselves became, thanks to the curious nature of men, objects of thought and study, the foundation of math itself[1]. That interest

[1]Needless to note that geometry is said to be equally important to mathematics as number theory (even though they have played different roles), and I do not disagree with that claim.

in arithmetic with natural numbers and the study of such objects have led philosophers and mathematicians to uncover very interesting results about natural numbers[2] and ways to classify them in a plethora of categories. We are specifically interested in the study of prime numbers, which come from the historical classification of numbers into polygonal numbers, i.e. numbers that count when things can be arranged into polygonal shapes like triangles, squares, etc., since we can also use irregular shapes like rectangles. In this fashion, composite numbers can be seen as rectangular numbers and prime numbers as "linear" numbers, since a rectangular array of a by b dots can represent the number ab, and for prime numbers either a or b have to be 1.

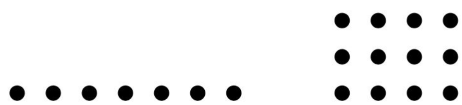

Figure 1.1: Rectangular representation for 7 and 12.
Note that 7 (a prime number) cannot have a rectangular representation of a different form than a rectangle with a side of 1.

The first relevant discovery about primes is the fact that there is no finite list that can contain all prime numbers, the proof was given by Euclid and is, perhaps, the most important result that from the greeks presented on number theory, even though they didn't study it deeply outside of the discoveries of the Pythagoreans, Euclid and Diophantus. After Diophantus, the subject of studying whole numbers and its properties remained slept in Europe at least up to the 17th century with Pierre de Fermat (who was inspired by Diophantus). After him, some of the greatest mathematicians such as Gauss, Euler, Legendre and Dirichlet made some of the major contributions.

Knowing that there are infinitely many primes leads to more questions about them, like how are they distributed between the rest of the whole numbers, or how do they grow, or if there is a formula for calculating prime numbers, etc. Some questions have been answered and some not, but the one that concerns us is the one about the distri-

[2] Deemed mystic and magical by some, thing that led, partially, to the general interest of studying properties of numbers.

bution of prime numbers. An answer was conjectured by Gauss an Legendre independently, by looking at tables of primes computed by hand, around 1800, but remained unproved until 1896 when, independently, J. Hadamard and C. J. de la Vallée Poussin proved it with a heavy use of complex analysis. This result is called the Prime Number Theorem or PNT and it was thought that no "elementary" proof was possible. However, in 1949 Paul Erdös and Atle Selberg came, also independently, with a proof that made no use of the famous Riemann zeta function ζ nor complex analysis, and from their discovery some more elementary proofs have been discovered. The PNT and Selberg's elementary proof are the main focus of this thesis.

In the next section we will define some notation and state some known theorems about natural and prime numbers.

1.2 Natural numbers

The set of natural numbers is

$$\mathbb{N} = \{1, 2, 3, 4, \ldots\}$$

which can be precisely constructed using the Peano Axioms for Arithmetic if we include 0, but which is not needed to prove the statements that we will present. We will also consider other sets of numbers through this work, such as the integers $\mathbb{Z}$, the real numbers $\mathbb{R}$ and the complex numbers $\mathbb{C}$. Addition and multiplication in $\mathbb{Z}$ are the usual binary operations and will always have a result in $\mathbb{N}$, in other words, the natural numbers are closed under addition and multiplication.

One of the main aspects of study in number theory is divisibility, an important relation on $\mathbb{N}$ that is the first step for defining many functions and also prime numbers. We will define it as

Definition 1 (Divisibility). Let $d, n \in \mathbb{N}$. We say that d divides n and write $d \mid n$ whenever $n = cd$ for some $c \in \mathbb{N}$. Equivalently, n is a multiple of d, d is a divisor of n, or d is a factor of n. If d does not divide n we write $d \nmid n$.

A divisor d of n such that $d < n$ is called a proper divisor of n.

It can be proven to be reflexive, transitive and linear, also that $1 \mid n$ and $n \mid 0$ for all naturals n, etc.

Examples.

- 6 has as divisors $1, 2, 3, 6$, notice $1 + 2 + 3 = 6$. A number which is equal to the sum of its proper divisors is called perfect.

- Numbers divisible by 2 are called even, the remaining are called odd.

With this definition, it comes naturally the question of how numbers are related by divisibility, and the answer comes in the form of another definition:

Definition 2 (Prime number). A number $p \in \mathbb{N}$ greater than 1 is said to be prime if $d \mid p$ implies $d = 1$ or $d = p$. A number that is not prime nor 1 is called composite[3].

So a number is prime when its only divisors are 1 and itself, i.e. it has no proper divisors greater than 1. This classifies the number into three categories: the unit 1, prime numbers and composite numbers.

Example. All the prime numbers less than 100 are $2, 3, 5, 7, 11, 13, 17, 19$, 23, 29, 31, $37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97$.

The classification of whole numbers in primes and composites is “global”, we can define a more “relative” classification with the following definition

Definition 3 (Greatest common divisor). Let $a, b \in \mathbb{N}$. The greatest common divisor of a and b is the unique[4] integer d that divides a and b such that $c \mid a$ and

---

[3]Since it will be “composed” of primes via multiplication.

[4]It can be proven to exist and be unique.

$c \mid b \implies c \mid d$. It is denoted by $\gcd(a,b)$ or (a,b).

If $\gcd(a,b) = 1$, we say a and b are coprime o relatively prime.

Thus we can say that prime numbers are numbers that are relatively prime to all the numbers they exceed.

Examples.

- $(a,b) = (b,a)\ \forall\ a,b \in \mathbb{N}$.
- If $d = (a,b)$ then $\exists\ x,y \in \mathbb{Z}$ s.t. $d = ax + by$. (*Bézout's identity*)

As mentioned before, Euclid proved the following theorem

Theorem 1 (Infinitude of Primes). *There are infinitely many prime numbers.*

Proof. Suppose $p_1, p_2, \ldots, p_n$ is a complete list of prime numbers. Notice that if $N = p_1 p_2 \cdots p_n + 1$, then $p_i \nmid N$ for all $i = 1, \ldots, n$ since N/p_i will always have a remainder of 1. Since any number > 1 is either prime or composite and no prime number divides N (assuming the initial list is complete), then N must be prime and our initial list would not be complete, contradicting our initial assumption. Thus there is no finite list that contains all prime numbers. ∎

Primes are extremely important as they are like the "building blocks" or "atoms" of numbers. Characterizing primes and learning about them has been found useful in fields outside number theory, such as Algebraic Geometry, Theory of Modular Forms and, in the more applied fashion, cryptography and security. We will state the most remarkable theorem for arithmetic with natural numbers[5].

1.2.1 A very remarkable result

Theorem 2 (The fundamental theorem of arithmetic). *Except for the arrangement of the factors, every positive integer $n > 1$ can be expressed uniquely as a product of primes.*

[5] Note that everything up to this point can be generalized to $\mathbb{Z}$ with minor alterations.

A small yet nice proof of theorem 2 can be found in the third chapter of [6]. If needed we will write the (unique) prime factorization of a number n as

$$n = \prod_{i=1}^{k} p_i^{\alpha_i} = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_n^{\alpha_n}$$

where p_i are prime numbers dividing n and $\alpha_i > 0$ being such that $p_i^{\alpha_i} \mid n$ but $p_i^{\alpha_i+1} \nmid n$ for $i = 1, \ldots, k$. Given a number n with the factorization above, any divisor of n is a product with the same primes but with the exponent of p_i being $\leq \alpha_i$ and possibly $\alpha_i = 0$ for some i (even for all i, then the (empty) product will be 1, a divisor of any number).

An example of application for the unique prime factorization of the Fundamental Theorem of Arithmetic or FTA is that given prime factorization of natural number a and b, we can calculate their gcd via the following proposition

Proposition 1. *If* $a, b \in \mathbb{N}$ *and*

$$a = \prod_{i=1}^{k} p_i^{\alpha_i}, \quad b = \prod_{i=1}^{k} p_i^{\beta_i}$$

are (not unique) prime factorizations of a *and* b *s.t.* α_i *and* β_i *are* ≥ 0*, then*

$$\gcd(a, b) = \prod_{i=1}^{k} p_i^{\min(\alpha_i, \beta_i)}.$$

to which a proof can be found in the first chapter of [1].

Having established a little bit if the importance of prime numbers and divisibility, we turn our attention to the functions that number-theorists work with, the so called arithmetical functions.

1.3 Arithmetical functions

We will define

Definition 4 (Arithmetical function). A function $f : \mathbb{N} \to \mathbb{R}$ (or $\mathbb{C}$) is called an arithmetical function or a number-theoretic function.

and state some examples of number theoretic functions and some of its properties.

Examples.

- The function $f_\alpha(n) = n^\alpha$ for all $n \in \mathbb{N}$, $\alpha \in \mathbb{R}$. The function f_1, usually called the identity function, is an arithmetical function that we will denote by N inspired by the notation of [1], we will denote f_α by N^α.

- We will call the function

$$f(n) = \left\lfloor \frac{1}{n} \right\rfloor = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{otherwise} \end{cases}$$

  for $n \in \mathbb{N}$ the identity function of arithmetical functions[6], we will denote it by I. Also $\lfloor x \rfloor$ is the greatest integer $\leq x$.

- The divisor function d, defined as the number of divisors of a number, equivalently written as

$$d(n) = \sum_{d|n} 1$$

  notice that the sum defining d runs over all divisors[7] of n.

- The Euler's totient function φ, defined as the number of numbers $< n$ relatively prime to n

$$\varphi(n) = \sum_{\substack{k=1 \\ (k,n)=1}}^{n} 1.$$

[6] We will make sense of this name when Dirichlet multiplication is defined.
[7] These kind sums are frequent and very important in number theory.

Apart from these examples, one of the main such functions is the so-called Möbius function μ, defined as

Definition (Möbius function)**.**

$$\mu(n) = \begin{cases} (-1)^k & \text{if } n \text{ is the product of } k \text{ distinct primes} \\ 0 & \text{otherwise} \end{cases}$$

This definition allows k to be 0 so $\mu(1) = 1$. Note that $\mu(n) \neq 0$ if and only if n is a square free positive integer.

An important property that these functions may have is multiplicativity

Definition 5 (Multiplicative arithmetical functions)**.** An arithmetical function f is said to be multiplicative if $f(mn) = f(m)f(n)$ whenever $(m, n) = 1$. f is said to be completely multiplicative if it happens even when $(m, n) > 1$.

Examples.

- N^α and I are completely multiplicative.
- μ and φ are multiplicative but not completely multiplicative.
- If f is multiplicative [completely multiplicative] then $f(1) = 1$ and in general $f(\prod p_i^{\alpha_i}) = \prod f(p_i^{\alpha_i})$ [$= \prod f(p_i)^{\alpha_i}$] for primes p_i.

The last example is one of the properties that make multiplicative functions behave so nicely with prime numbers, since they are almost entirely defined by their value at prime powers or at primes if completely multiplicative. We will make use of this property to prove the next result, a simple formula that relates μ to I:

Proposition 2. *For $n \in \mathbb{N}$ it follows that*

$$\sum_{d|n} \mu(d) = I(n).$$

Proof. Since the only divisor of 1 is 1, it is clearly true for $n = 1$.

For $n > 1$, we can make use of the prime factorization of n. However, since μ is multiplicative, we just need to show that the proposition holds for powers of primes. Let $n = p^\alpha$ for p prime, then $\sum_{d|p^\alpha} \mu(d) = \mu(1) + \mu(p) + \mu(p^2) + \cdots + \mu(p^\alpha) = 1 + (-1) + 0 + \cdots + 0 = 0$, thus it holds for any $n > 1$. ■

Notice that μ relates to the prime factorization of a number, but it's not the only arithmetical function that relates to primes. The next function, even though it is not multiplicative, is of a great importance for the elementary proof of the prime number theorem: the Mangoldt function.

Definition (Von Mangoldt function)**.** For $n \in \mathbb{N}$ we define the Mangoldt function Λ to be

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

One can easily see and prove the next relation that relates log closely to Λ, apart from the definition and as a direct consequence of the fundamental theorem of arithmetic

Proposition 3. *For $n \geq 1$ we have*

$$\log n = \sum_{d|n} \Lambda(n).$$

We will now be turning our attention to another key concept necessary to understand arithmetical functions and their relations.

1.4 Dirichlet product

There is a very important binary operation on the set of all arithmetic functions that turns a special subset of them into an abelian group with respect to this operation, this is the Dirichlet product.

> **Definition 6** (Dirichlet product). If f and g are two arithmetical functions, we define their Dirichlet product (or Dirichlet convolution) to be the arithmetical function h or $f * g$ defined by the equation

$$h(n) = (f * g)(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right) = \sum_{ab=n} f(a)g(b).$$

The Dirichlet product is associative and commutative, which can easily be seen by the second sum in the definition, and notice that

$$(f * I)(n) = \sum_{d|n} f(d) I\left(\frac{n}{d}\right) = f(1)I(n) + \cdots + f(n)I(1) = f(n)$$

and we can effectively write $f * I = I * f = f$ for all arithmetical functions f. This behaviour of I explains its name of identity function, since it is the identity with respect to the Dirichlet product. Now we just need some kind of inverses to define an abelian group of arithmetical functions, that inverse is given by the next theorem.

Theorem 3 (Dirichlet inverse). *If f is an arithmetical function with $f(1) \neq 0$, there is a unique arithmetical function f^{-1} called the Dirichlet inverse of f, such that*

$$f * f^{-1} = f * f^{-1} = I.$$

Proof. We will first prove uniqueness of such a function. Let's consider $n = 1$, then $(f^{-1} * f)(1) = f^{-1}(1)f(1) = I(1) = 1$ has one only possibility $f^{-1}(1) = (f(1))^{-1}$. Then let $n > 1$ suppose for all values $1 \leq k < n$, $f^{-1}(k)$ has been uniquely determined, we now just need to consider the equation $(f * f^{-1})(n) = I(n) = 0$ which leads to

$$\sum_{d|n} f\left(\frac{n}{d}\right) f^{-1}(d) = f(1)f^{-1}(n) + \sum_{\substack{d|n \\ d<n}} f\left(\frac{n}{d}\right) f^{-1}(d) = 0$$

and since for all values $< n$, f^{-1} have been uniquely determined, this will give only one possibility for $f^{-1}(n)$, namely

$$f^{-1}(n) = -\frac{1}{f(1)} \sum_{\substack{d|n \\ d<n}} f\left(\frac{n}{d}\right) f^{-1}(d).$$

and the result follows by induction. ■

So we can have inverses as long as $f(1) \neq 0$, this along with the algebraic properties of the Dirichlet product, and the easily verifiable fact that $(f*g)^{-1} = f^{-1}*g^{-1}$ whenever $f(1), g(1)$ are nonzero, makes the set of all arithmetical functions f s.t. $f(1) \neq 0$ into an abelian group with respect to $*$ with I being the identity element of this operation.

Remember the sum in proposition 2, we can define a function $u : \mathbb{N} \to \{1\}$ called the unit function and express said sum as a Dirichlet product

$$\mu * u = I$$

which reveals that the Dirichlet inverse of the Möbius function μ is the unit function u. Using this fact, the next result follows immediately.

Theorem 4 (Möbius inversion formula). *The equation*

$$f(n) = \sum_{d|n} g(d)$$

is true if and only if

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

This can be used to prove the following standard identities in number theory, presented as a proposition, whose proofs can be found, again, in [1].

Proposition 4.

- $\Lambda(n) = \sum_{d|n} \mu(d) \log \frac{n}{d} = -\sum_{d|n} \mu(d) \log d$ *for* $n \geq 1$.
- $n = \sum_{d|n} \varphi(d) \iff \varphi(n) = \sum_{d|n} d\mu\left(\frac{n}{d}\right)$ *for all*[8] $n \in \mathbb{N}$.

And even though there are many things that can be said about these and many other arithmetical functions, we shall continue with our journey to prove the prime number theorem. In the following chapter we will introduce functions of real variable that will be useful to define and study more properties of arithmetical functions using results from elementary calculus.

[8] Any of these can be proved to be true by other methods.

Chapter 2

Introducing real numbers

2.1 Partial sums

We know there are infinitely many primes, but a natural question that arises when thinking about series like the harmonic one, which is divergent, or the series $\sum_{n\geq 1}\frac{1}{n^2}$ which is convergent to $\frac{\pi^2}{6}$ is if the sum

$$\sum_{p \text{ prime}} \frac{1}{p} \tag{2.1}$$

of the reciprocals of primes converges, that would tells us how in "average" primes happen between natural numbers. Euler was the first one to show that (2.1) diverges without using Euclid's result about the existence of an infinite number of primes, implying in fact that result, being a stronger form of it. However, we can make more precise statements about how these kinds of sums grow as n grows. In order to do this we will introduce some definitions regarding the known big and little o notations that will be useful when studying the "asymptotic behavior" of partial sums of arithmetical functions f, namely $\sum_{1\leq n\leq k} f(n)$ and their averages $\frac{1}{n}\sum_{1\leq n\leq k} f(n)$.

2.1.1 Asymptotic notation

Definition 7 (Big O notation). For a real-valued[1] function f defined[2] on an unbounded subset of $\mathbb{R}^+$, if for a real valued function g, also defined in such a subset, $g(x) > 0$ for all $x \geq a$, $a \in \mathbb{R}$, we say $f(x)$ *is big o of* $g(x)$ and write

$$f(x) = O(g(x))$$

if the quotient $\frac{f(x)}{g(x)}$ is bounded for $x \geq a$.

We will often carelessly write equations like $f(x) = g(x) + O(h(x))$ meaning that $f(x) - g(x) = O(h(x))$ and manipulate them as needed.

Examples.

- If $f_1(x) = O(g(x))$, $f_2(x) = O(g(x))$ then $(f_1 + f_2)(x) = O(g(x))$ and $(cf_1)(x) = O(g(x))$ for $c \in \mathbb{R}$.
- If $f_1(x) = O(g_1(x))$, $f_2(x) = O(g_2(x))$ then $(f_1 f_2)(x) = O((g_1 g_2)(x))$.
- If $f(x) = O(g(x))$ then $\displaystyle\int_a^x f = O\left(\int_a^x g\right)$.

We will also introduce the following "stronger" notation

Definition 8 (Little o notation). For a real-valued function f defined on an unbounded subset of $\mathbb{R}^+$, if for a real valued function g, also defined in such a subset, $g(x) > 0$ for all $x \geq a$, $a \in \mathbb{R}$, we say $f(x)$ *is little o of* $g(x)$ and write

$$f(x) = o(g(x))$$

if

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = 0.$$

[1]This definition also works when f is complex-valued, however g will still need to be real-valued.

[2]But maybe also defined outside such subset, same with g. Also, needless to say that $\mathbb{R}^+$ is the interval $(0, \infty)$.

Notice that little o is stronger that big o in the sense that $f(x) = o(g(x)) \implies f(x) = O(g(x))$. General properties of o are similar to properties of O, here are some examples that illustrates this fact and some use of these notations

Examples.

- For $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ a polynomial of degree n, $f(x) = O(x^n)$ but $f(x) \neq o(x^n)$.
- For the same polynomial above, $f(x) = O(x^{n+1})$ and $f(x) = o(x^{n+1})$.
- $f = o(g) \implies cf = o(g)$ for $c \neq 0$.
- If $f = o(F)$, $g = o(G)$ then $fg = o(FG)$.
- If $f = o(g)$ and $g = o(h)$ then $f = o(h)$.

Finally we define when two functions are "asymptotically equivalent", meaning they eventually approach one another.

Definition 9 (Asymptotically equivalence). If for real valued functions f, g that are defined on unbounded subsets of $\mathbb{R}^+$ is true that

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$$

then we say that $f(x)$ is asymptotic to $g(x)$ as $x \to \infty$ and write

$$f(x) \sim g(x) \text{ as } x \to \infty.$$

Examples.

- $\sim$ is an equivalence relation in the subset of all function that satisfy the definition above.
- $f(x) \sim g(x)$ as $x \to \infty \iff f(x) = g(x) + o(g(x))$.

- $\sum_{k \le x} d(k) \sim x \log x$ as $x \to \infty$ where d is the divisor function defined on the examples of definition 4.

A proof of the last example and related identities involving partial sums of arithmetical functions can be found by applying some elementary calculus, we will use Euler's summation formula, which is found in the following subsection.

2.2 Euler's summation Formula

The idea behind the following theorem is to compare a partial sum or average by comparing it to an integral, from this formula we can obtain an exact expression for the error (which will be usually inside of a big or little o).

Theorem 5 (Euler's summation formula)**.** *If f has a continuous derivative f' on the interval $[y, x]$, where $0 < y < x$, then*

$$\begin{aligned} \sum_{y<n\le x} f(n) = \int_y^x f(t)\,\mathrm{d}t + \int_y^x (t - \lfloor t \rfloor) f'(t)\,\mathrm{d}t \\ + f(x)(\lfloor x \rfloor - x) - f(y)(\lfloor y \rfloor - y). \end{aligned} \tag{2.2}$$

Proof. Notice that whenever integers n, $n-1$ lie in $[y, x]$ we have that

$$\begin{aligned} \int_{n-1}^{n} \lfloor t \rfloor f'(t)\,\mathrm{d}t &= \int_{n-1}^{n} (n-1) f'(t)\,\mathrm{d}t \\ &= (n-1)(f(n) - f(n-1)) \\ &= (nf(n) - (n-1)f(n-1)) - f(n) \end{aligned}$$

so if we sum from $n = \lfloor y \rfloor + 2$ to $n = \lfloor x \rfloor$

$$\begin{aligned}\int_{\lfloor y\rfloor+1}^{\lfloor x\rfloor}\lfloor t\rfloor f'(t)\,\mathrm{d}t = &\sum_{n=\lfloor y\rfloor+2}^{\lfloor x\rfloor} nf(n)-(n-1)f(n-1)-f(n)\\ =&[(\lfloor y\rfloor+2)f(\lfloor y\rfloor+2)-(\lfloor y\rfloor+1)f(\lfloor y\rfloor+1)+(\lfloor y\rfloor+3)f(\lfloor y\rfloor+3)\\ &-(\lfloor y\rfloor+2)f(\lfloor y\rfloor+2)+\ldots+\lfloor x\rfloor f(\lfloor x\rfloor)-(\lfloor x\rfloor-1)f(\lfloor x\rfloor-1)]\\ &-\sum_{n=\lfloor y\rfloor+2}^{\lfloor x\rfloor} f(n)\end{aligned}$$

and notice that the sum inside square brackets telescopes leaving just the terms $\lfloor x\rfloor f(\lfloor x\rfloor)-(\lfloor y\rfloor+1)f(\lfloor y\rfloor+1)$, thus subtracting and adding $f(\lfloor y\rfloor+1)$ leaves us with

$$\begin{aligned}\sum_{y<n\le x} f(n) &= -\int_{\lfloor y\rfloor+1}^{\lfloor x\rfloor}\lfloor t\rfloor f'(t)\,\mathrm{d}t+\lfloor x\rfloor f(\lfloor x\rfloor)-\lfloor y\rfloor f(\lfloor y\rfloor+1)\\ &= -\int_y^x \lfloor t\rfloor f'(t)\,\mathrm{d}t+\lfloor x\rfloor f(x)-\lfloor y\rfloor f(y).\end{aligned}$$

Finally, notice that by parts

$$\int_y^x f(t)\,\mathrm{d}t = tf(t)\Big|_y^x-\int_y^x tf'(t)\,\mathrm{d}t$$

so, since the derivative of $\lfloor x\rfloor f(x)$ is $f'(x)$ on each interval within integers, we have that

$$\begin{aligned}\sum_{y<n\le x} f(n) = &\int_y^x f(t)\,\mathrm{d}t+\int_y^x (t-\lfloor t\rfloor)f'(t)\,\mathrm{d}t\\ &+f(x)(\lfloor x\rfloor-x)-f(y)(\lfloor y\rfloor-y).\end{aligned}$$

as needed. ■

2.3 Elementary examples of asymptotic formulas

We will proof the last statement in the examples of definition 9, usually called the average order of $d(n)$. However, we will first use Euler's summation formula by proving the following two formulas that are needed.

Proposition 5. *For* $x \geq 1$

(a) $\displaystyle\sum_{n\leq x} \frac{1}{n} = \log x + \gamma + O(1)$, *where* γ *is Euler's or Euler-Mascheroni's constant defined by* $\displaystyle\gamma = \lim_{n\to\infty}\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \log n\right)$.

(b) $\displaystyle\sum_{n\leq x} n^{\alpha} = \frac{x^{\alpha+1}}{\alpha+1} + O(x^{\alpha})$ *for* $\alpha \geq 0$.

Proof. For **(a)** we can apply Euler's summation for $f(t) = \frac{1}{t}$ then

$$\begin{aligned}\sum_{n\leq x} \frac{1}{n} &= 1 + \sum_{1<n\leq x} \frac{1}{n} \\ &= 1 + \int_1^x \frac{1}{t}\,\mathrm{d}t - \int_1^x \frac{(t-\lfloor t\rfloor)}{t^2}\,\mathrm{d}t + \frac{\lfloor x\rfloor - x}{x} - \frac{\lfloor 1\rfloor - 1}{1} \\ &= 1 + \log x - \int_1^{\infty} \frac{(t-\lfloor t\rfloor)}{t^2}\,\mathrm{d}t + \int_x^{\infty} \frac{(t-\lfloor t\rfloor)}{t^2}\,\mathrm{d}t + O\left(\frac{1}{x}\right)\end{aligned}$$

since $0 \leq t - \lfloor t\rfloor < 1$ for all $t \in \mathbb{R}$. The integral $\int_1^{\infty} \frac{(t-\lfloor t\rfloor)}{t^2}\,\mathrm{d}t \leq \int_1^{\infty} \frac{1}{t^2}\,\mathrm{d}t = 1$ so it converges. Similarly $0 \leq \int_x^{\infty} \frac{(t-\lfloor t\rfloor)}{t^2}\,\mathrm{d}t \leq \int_x^{\infty} \frac{1}{t^2}\,\mathrm{d}t = \frac{1}{x}$ so

$$\sum_{n\leq x} \frac{1}{n} = \log x + 1 - \int_1^{\infty} \frac{(t-\lfloor t\rfloor)}{t^2}\,\mathrm{d}t + O\left(\frac{1}{x}\right)$$

and notice that if we let $x \to \infty$ in **(a)** then

$$\lim_{x\to\infty}\left(\sum_{n\leq x} \frac{1}{n} - \log x\right) = \gamma = 1 - \int_1^{\infty} \frac{t-\lfloor t\rfloor}{t^2}\,\mathrm{d}t$$

thus proving **(a)**.

Now for **(b)** we use Euler's summation once again with $f(t) = t^\alpha$, then

$$\begin{aligned}
\sum_{n \le x} n^\alpha &= 1 + \sum_{1 < n \le x} n^\alpha \\
&= 1 + \int_1^x t^\alpha \, \mathrm{d}t + \alpha \int_1^x (t - \lfloor t \rfloor) t^{\alpha - 1} \, \mathrm{d}t + x^\alpha (\lfloor x \rfloor - x) \\
&= \frac{x^{\alpha+1}}{\alpha + 1} - \frac{1}{\alpha + 1} + O\left(\alpha \int_1^x t^{\alpha+1} \right) + O(x^\alpha) \\
&= \frac{x^{\alpha+1}}{\alpha + 1} + O(x^\alpha)
\end{aligned}$$

again, since $t - \lfloor t \rfloor$ is always between 0 and 1. ■

The following result involving partial sums of the divisor function was discovered by Dirichlet, we will prove it using the previous proposition and some clever counting under a curve.

Theorem 6. *For all* $x \ge 1$

$$\sum_{n \le x} d(n) = x \log x + (2\gamma - 1)x + O(\sqrt{x}).$$

Proof. We will be using both formulas that we just proved. ■

Chapter 3

Theorems on prime numbers

Mathematicians have tried in vain to this day to discover some order in the sequence of prime numbers, and we have reason to believe that it is a mystery into which the human mind would never penetrate.

- Leonard Euler

3.1 The Prime Number Theorem

In this section we will state the prime number theorem in its standard form and in some equivalent manners. We will prove these equivalences and try to understand precisely what needs to be known and proven. We will also look at two consequences of the PNT regarding prime numbers.

3.1.1 Statement of the PNT

We will introduce a new arithmetical function that is, not surprisingly, the main one when studying primes and the main focus of this work.

Definition 10 (Prime-counting function). For $x \in \mathbb{R}$ we define the prime counting function π as

$$\pi(x) = \#\text{ of primes } \leq x = \sum_{\substack{p \leq x \\ p \text{ prime}}} 1$$

For example, $\pi(10^2) = 25$, $\pi(10^5) = 9592$ and $\pi(10^{10}) = 455052511$. Euclid proved that $\pi(x) \to \infty$ as $x \to \infty$, the PNT deals with the asymptotic behaviour of π.

Theorem (The Prime Number Theorem). *For* $x \geq 1$

$$\lim_{x \to \infty} \frac{\pi(x)}{x/\log x} = 1,$$

equivalently

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

This simple yet curious statement is what took some of the best mathematicians almost a century to prove, and almost two centuries to prove with the use of elementary number theory and some basic calculus. It happens to be that there are equivalent formulations of the PNT that are more easy to work with, such as the ones we will prove with new functions we will define in the next subsection.

The history of the PNT was summarized in the last paragraph of the historical notes in the first chapter, however, we may now look more deeply at the history of its discovery and its place in mathematics. We begin by noting that this theorem was not the first that attempted to provide a description of $\pi(x)$. Legendre made a conjecture in 1798 that

$$\pi(x) = \frac{x}{\log x + A(x)}$$

where $A(x)$ approaches 1.08366 as $x \to \infty$, and even though this was not true, it was a first try on describing the behavior and distribution of prime numbers[1]. Legendre also conjectured what is now called the Dirichlet[2] theorem on primes in arithmetic sequences, that asserts that in any arithmetic sequence $\{kn+h\}_{n\in\mathbb{N}}$ there are infinitely many primes if and only if $(k,h)=1$, whose proof developed more tools and understanding that helped later prove the PNT.

Gauss also made, independently from Legendre, significant advances on the study of prime numbers, in fact, by 1793 he had made calculation and tabulation of primes his "hobby" and had studied distribution of primes in various lengthy intervals that grew up to the millions, some tabulations of Gauss are on the appendix of [3]. Gauss conjectured, in modern language, that primes occurred around n with "probability" $1/n$, or more precisely that for an interval $[a,b)$

$$\pi(b)-\pi(a)\approx\int_a^b\frac{1}{\log}$$

however, Gauss never published his work on the matter and did not found a proof for the PNT even though he stated it in a letter to Johann Encke in 1849.

As mentioned above, another big step forward in proving the theorem was Dirichlet's proof of primes in arithmetic sequences or progressions, for which he used "characters" of the group $\mathbb{Z}_n^*$ of residue classes modulo n. A character f is simply a homomorphism (a map that preserves the group operation) $\mathbb{Z}_n^*\to\mathbb{C}^*$ where $\mathbb{C}^*=\mathbb{C}\setminus\{0\}$ is the multiplicative group of (nonzero) complex numbers. It can be proven that $\mathbb{Z}_n^*$ has exactly $\varphi(n)$ characters f and then we can define for each character a related arithmetical function, called a Dirichlet character, $\chi_f:\mathbb{N}\to\mathbb{C}^*$ by $\chi_f(a)=f([a])$ if $(a,n)=1$ and $\chi_f(a)=0$ if $(a,n)>1$, where $[a]\in\mathbb{Z}_n^*$ is the residue class modulo n of $a\in\mathbb{N}$. Dirichlet proved some useful relations, usually called orthogonality relations, that helped him prove strong results about partial sums of these characters which can be reviewed seen

[1]The work of L.J. Goldstein in [3] deals in greater detail with $A(x)$.
[2]Since it was proved by him in 1837.

on chapters 6, 7, 8 and 10 of [1]. The second important thing he also developed is what is now called a Dirichlet L-series or Dirichlet L-function, namely

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

which are a generalization of the Riemann zeta function ζ, for which you just have to consider the unique Dirichlet character modulo 1. The proof that Dirichlet made, using these tools, of the infinitude of primes of the form $hn+k$ when $(h,k)=1$, made use of analysis to answer an arithmetic question. Even though the results of Dirichlet were of great importance, there was some resistance to the new ideas that came with the proof, mathematicians were always looking for proofs without the involvement of analysis to study arithmetic, but nowadays it is broadly accepted, used and recognized as a rightful and natural tool to study integers and not only as a shortcut.

The next step in the direction of the proof of the PNT was discovered by Pafnuty Chebyshev around 1850 and it is discussed in the next subsection.

3.1.2 Chebyshev's functions

We will introduce two new functions: ψ and ϑ, called Chebyshev's functions since they were introduced by the Russian mathematician Pafnuty Chebyshev in 1851, who proved some results that we will present here.

Definition 11 (Chebyshev ψ). For $x>0$ we define the Chebyshev ψ function by

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

since Λ is nonzero only for prime powers, we can rewrite it as

$$\sum_{m=1}^{\infty}\sum_{p\leq x^{1/m}}\log p=\sum_{m=1}^{\lfloor\log_2 x\rfloor}\sum_{p\leq x^{1/m}}\log p.$$

We also define a second function

Definition 12 (Chebyshev ϑ). For $x > 0$ we define the Chebyshev ϑ function by

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

where p runs over all primes $\leq x$.

this can be used to rewrite ψ as

$$\psi(x) = \sum_{m=1}^{\lfloor \log_2 x \rfloor} \vartheta\left(x^{1/m}\right).$$

and this equation is not the only relation they have. It so happens that these two functions play a major role in the proof of the PNT, in fact they grow alike in the sense of the following theorem.

Theorem 7. *For $x > 0$ it follows that*

$$0 \leq \frac{\psi(x)}{x} - \frac{\vartheta(x)}{x} \leq \frac{\log^2 x}{2\sqrt{x}\log 2}.$$

Proof. Notice that the definition of $\vartheta(x)$ as the sum over primes $\leq x$ of their logarithms, we can say that $\vartheta(x) \leq \sum_{p \leq x} \log x \leq x \log x$. Also, given

$$\psi(x) = \sum_{m=1}^{\lfloor \log_2 x \rfloor} \vartheta\left(x^{1/m}\right)$$

it implies that

$$0 \leq \psi(x) - \vartheta(x) = \sum_{2 \leq m \leq \log_2 x} \vartheta\left(x^{1/m}\right)$$

so we can combine both inequalities and get that

$$0 \leq \psi(x) - \vartheta(x) \leq \sum_{2 \leq m \leq \log_2 x} x^{1/m} \log\left(x^{1/m}\right)$$

$$\begin{aligned} &\leq (\log_2 x)\sqrt{x}\log\sqrt{x} \\ &= \frac{\sqrt{x}\log^2 x}{2\log 2} \\ &= x\frac{\log^2 x}{2\sqrt{x}\log 2}. \end{aligned}$$ ■

This theorem not only gives a bound to the difference of $\psi(x)/x - \vartheta(x)/x$, it also implies that

Corollary 1. If one of the limits

$$\lim_{x\to\infty}\frac{\psi(x)}{x} \quad \textit{and} \quad \lim_{x\to\infty}\frac{\vartheta(x)}{x}$$

exists, then the other one exists, in which case both limits are equal.

This corollary means that both Chebyshev functions grow alike.

The relation of these functions to the prime number theorem will be reviewed in the next subsection, for that purpose we will need to find a formula that relates ϑ to π (the prime-counting function), for which we will use the following well known result:

Theorem 8 (Abel's summation formula). *For any arithmetical function $a(n)$ let the function $A(x) = \sum_{n\leq x} a(n)$ be the partial sum of a with $A(x) = 0$ when $x < 1$, if a function f has a continuous derivative on a nonempty interval $[y, x]$, $y > 0$, then we have*

$$\sum_{y<n\leq x} a(n)f(n) = A(x)f(x) - A(y)f(y) - \int_y^x A(t)f'(t)\,\mathrm{d}t\,.$$

Abel's formula will not be proved, however a short proof can be given noting that the left hand side can be expressed easily as a Riemann-Stieltjes integral as in chapter 4 of [1]. We'll use this identity, similarly as to when we used Euler's summation formula, to prove the next proposition.

Proposition 6. *For $x \geq 2$ we have*

$$\vartheta(x) = \pi(x)\log x - \int_2^x \frac{\pi(t)}{t}\,\mathrm{d}t \tag{3.1}$$

and

$$\pi(x) = \frac{\vartheta(x)}{\log x} + \int_2^x \frac{\vartheta(t)}{t\log^2 t}\,\mathrm{d}t\,. \tag{3.2}$$

Proof. If we let $a(n)$ be the characteristic function of the set $\{p \in \mathbb{N} \mid p \text{ is prime}\}$, then of course

$$\pi(x) = \sum_{1<n\leq x} a(n) \quad \text{and} \quad \vartheta(x) = \sum_{1<n\leq x} a(n)\log n.$$

To prove (3.1) we replace $f(x) = \log x$ and $y = 1$ in Abel's identity, thus given the expressions above,

$$\vartheta(x) = \pi(x)\log x - \pi(1)\log 1 - \int_1^x \frac{\pi(t)}{t}\,\mathrm{d}t$$

from which (3.1) is deduced since $\pi(t) = 0$ when $t < 2$. For (3.2) let $b(n) = a(n)\log n$, then

$$\pi(x) = \sum_{3/2<n\leq x} \frac{b(n)}{\log n} \quad \text{and} \quad \vartheta(x) = \sum_{1<n\leq x} b(n)$$

then if $f(x) = 1/\log x$ and $y = 3/2$ we get

$$\pi(x) = \frac{\vartheta(x)}{\log x} - \frac{\vartheta(3/2)}{\log 3/2} + \int_{3/2}^x \frac{\vartheta(t)}{t\log^2 t}\,\mathrm{d}t$$

and since $\vartheta(t) = 0$ if $t < 2$ the second equation is proved. ∎

We will also present a proposition that will be used, by not proved, but which are well-known results about Chebyshev's functions, together with related sums

Proposition 7. *For all $x \geq 1$ we have*

- $\displaystyle\sum_{n\leq x} \quad \left(\frac{x}{n}\right) = x\log x - x + O(\log x)$,
- $\displaystyle\sum_{n\leq x} \vartheta\left(\frac{x}{n}\right) = x\log x + O(x)$,
- $\displaystyle\sum_{n\leq x} \frac{\Lambda(n)}{n} = \log x + O(1)$,
- $\displaystyle\sum_{p\leq x} \frac{\log p}{p} = \log x + O(1)$,

also, $\vartheta(x) = O(x)$ *and* $\psi(x) = O(x)$.

Proves of all facts above can be given with Shapiro's Tauberian Theorem that involves relating partial sums of arithmetical functions $a(n)$ and $a(n)\lfloor x/n \rfloor$, an overview and proofs for these can be seen in [1].

3.1.3 Equivalent forms of the PNT

We already reviewed the PNT as an asymptotic relation that involves the prime-counting function π, the next result gives us alternative forms of the PNT.

Theorem 9. *The following identities are equivalent to one another*

$$\lim_{x\to\infty} \frac{\pi(x)\log x}{x} = 1, \tag{3.3}$$

$$\lim_{x\to\infty} \frac{\vartheta(x)}{x} = 1, \tag{3.4}$$

$$\lim_{x\to\infty} \frac{\psi(x)}{x} = 1. \tag{3.5}$$

Proof. Notice that the first equation in 6 implies that

$$\frac{\vartheta(x)}{x} = \frac{\pi(x)\log x}{x} - \frac{1}{x}\int_2^x \frac{\pi(t)}{t}\,\mathrm{d}t$$

so in order to show that (3.3) implies (3.4) we need to show it implies the integral above tends to 0 as $x \to \infty$. Notice that (3.3) implies that for $x \geq 2$

$$\frac{\pi(x)}{x} = O\left(\frac{1}{\log x}\right)$$

thus

$$\frac{1}{x}\int_2^x \frac{\pi(t)}{t}\,\mathrm{d}t = O\left(\frac{1}{x}\int_2^x \frac{1}{\log t}\,\mathrm{d}t\right)$$

and the integral inside the big O can be bounded by the rectangle with base of the area of integration and height of $1/\log s$ for s at the left end (since $1/\log$ is decreasing), thus

$$\int_2^x \frac{1}{\log t}\,\mathrm{d}t = \int_2^{\sqrt{x}} \frac{1}{\log t}\,\mathrm{d}t + \int_{\sqrt{x}}^x \frac{1}{\log t}\,\mathrm{d}t \leq \frac{\sqrt{x}-2}{\log 2} + \frac{x-\sqrt{x}}{\log\sqrt{x}}$$

which immediately implies $\lim_{x\to\infty} \frac{1}{x}\int_2^x \frac{1}{\log t}\,\mathrm{d}t = 0$, therefore (3.3) $\implies$ (3.4). For proving the reverse, just note that (3.2) implies

$$\frac{\pi(x)\log x}{x} = \frac{\vartheta(x)}{x} + \frac{\log x}{x}\int_2^x \frac{\vartheta(t)}{t\log^2 t}\,\mathrm{d}t$$

so we'll need to show that the above integral tends to 0 as $x \to \infty$. We'll proceed as before, noting that (3.4) implies that $\vartheta(t) = O(t)$ so

$$\frac{\log x}{x}\int_2^x \frac{\vartheta(t)}{t\log^2 t}\,\mathrm{d}t = O\left(\frac{\log x}{x}\int_2^x \frac{1}{\log^2 t}\,\mathrm{d}t\right)$$

we can do the same as we did for the integral we already bounded above to get

$$\int_2^x \frac{1}{\log^2 t}\,\mathrm{d}t \leq \frac{\sqrt{x}-2}{\log^2 2} + \frac{x-\sqrt{x}}{\log^2\sqrt{x}}$$

which proves $\lim_{x\to\infty} \frac{\log x}{x}\int_2^x \frac{1}{\log^2 t}\,\mathrm{d}t = 0$, hence showing (3.4) $\implies$ (3.3) and therefore proving both are equivalent. Since we already stated in corollary 1 that (3.4) and (3.5) are equivalent, we are done. ■

The proofs of the PNT usually relies in proving it in the forms of (3.4) or (3.5), written in an equivalent manner as $\vartheta(x) \sim x$ and $\psi(x) \sim x$ respectively, however, there are some other equivalences worth mentioning, but which will not be proved:

- $\pi(x) \sim \dfrac{x}{\log \pi(x)}$.
- $\displaystyle\lim_{n\to\infty} \frac{p_n}{n \log n} = 1$, where p_n is the nth prime.
- $\pi(x) \sim Li(x)$ where Li is the logarithmic integral function $Li(x) = \displaystyle\int_2^x \frac{1}{\log x}\,\mathrm{d}x$.
- $M(x) \sim x$, where $M(x)$ is the Mertens function $M(x) = \displaystyle\sum_{n\le x} \mu(n)$.

The prime number theorem gives us not only average information about the primes but concrete information in certain intervals. Some statements about this will be presented in the next subsection.

3.2 Distribution of prime numbers

Here we'll present two elementary theorems regarding the distribution of prime numbers between some bounds, namely where the nth prime is located, and strict bounds the function $\pi(n)$ for arbitrary n. But first, consider the sum $\sum_{p \text{ prime}} \frac{1}{p}$ of the reciprocals of primes, we know that Euler showed such sum diverges, but we haven't stated yet how it does so, thus let us consider the following proposition

Proposition 8. *There is a constant M such that*

$$\sum_{p\le x} \frac{1}{p} = \log\log x + M + O\left(\frac{1}{\log x}\right), \quad \textit{for all } x \ge 2.$$

Proof. We will make use of Abel's summation formula (theorem 8), for that purpose let $A(x) = \sum_{p\le x} (\log p)/p$ and let $a : \mathbb{N} \to \mathbb{R}$ be the characteristic function of primes, namely

$$a(n) = \begin{cases} 1 & \text{if } n \text{ is prime,} \\ 0 & \text{otherwise.} \end{cases}$$

so that our original sum to study will be the partial sums of $a(n)/n$, while A will be the partial sums of $a(n)(\log n)/n$. Thus for applying Abel's summation let $f(t) = 1/\log t$ and $1 < y < 2$ (then $A(y) = 0$)

$$\begin{aligned} \sum_{p \le x} \frac{1}{p} &= \sum_{y < n \le x} \left(\frac{a(n) \log n}{n} \right) \left(\frac{1}{\log n} \right) \\ &= \frac{A(x)}{\log x} + \int_2^x \frac{A(t)}{t \log^2 t} \, \mathrm{d}t \, . \end{aligned}$$

Notice that by proposition 7, $A(x) = \log x + O(1)$, so we can write $A(x) = \log x + R(x)$ for some function R such that $R(x) = O(1)$, then we get

$$\sum_{p \le x} \frac{1}{p} = 1 + O\left(\frac{1}{\log x} \right) + \int_2^x \frac{1}{t \log t} \, \mathrm{d}t + \int_2^x \frac{R(t)}{t \log^2 t} \, \mathrm{d}t \, .$$

The first integral equals $\log \log x - \log \log 2$, and the second one can be expressed as

$$\int_2^x \frac{R(t)}{t \log^2 t} \, \mathrm{d}t = \int_2^\infty \frac{R(t)}{t \log^2 t} \, \mathrm{d}t - \int_x^\infty \frac{R(t)}{t \log^2 t}$$

where the first improper (yet definite) integral converges since R is $O(1)$, and the second integral

$$\int_x^\infty \frac{R(t)}{t \log^2 t} \, \mathrm{d}t = O\left(\int_x^\infty \frac{1}{t \log^2 t} \, \mathrm{d}t \right) = O\left(\frac{1}{\log x} \right).$$

Finally, we can write

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

where M is defined by

$$M = 1 - \log\log 2 + \int_2^\infty \frac{R(t)}{t\log^2 t}\,\mathrm{d}t\,. \qquad \blacksquare$$

The constant M in the proposition above is called Meissel-Mertens constant, and its approximate value is $M \approx 0.2614972128$, more can be found in [2].

Now to the main results of this subsection. Much is talked about the prime counting function π, primes in an interval, etc., but according to Apostol in [1], even though better estimates can be given to approximate $\pi(n)$, the "correct" order is $n/\log n$, which is represented by the following theorem

Theorem 10. *For every integer $n \geq 2$ we have*

$$\frac{1}{6}\frac{n}{\log n} < \pi(n) < 6\frac{n}{\log n}$$

A proof is given in Chapter 4 of [1], is remarkable to say that Apostol's proof is almost entirely elementary and uses only combinatorics and inequalities for natural numbers until the end. This theorem opens room for another one

Theorem 11. *For $n \geq 1$, if p_n is the nth prime, we have*

$$\frac{1}{6}n\log n < p_n < 12\left(n\log n + n\log\frac{12}{e}\right).$$

Proof. If $n = \pi(k)$ (so that $k = p_n$), the last theorem gives us

$$n = \pi(k) < 6\frac{k}{\log k} = 6\frac{p_n}{\log p_n}$$

and rearranging the inequality we can get

$$p_n > \frac{1}{6}n\log p_n > \frac{1}{6}n\log n$$

proving the left hand side inequality of (11). For the upper bound we use the same theorem to get

$$n > \frac{1}{6}\frac{p_n}{\log p_n} \implies p_n < 6n \log p_n.$$

Now, since $\log x \leq \frac{2}{e}\sqrt{x}$ for $x \geq 1$ (notice $e \log x - 2\sqrt{x}$ has a maximum at $x = e^2$ and is < 0 otherwise), we also have $\log p_n \leq \frac{2}{e}\sqrt{p_n}$, thus the previous inequality yields

$$p_n < 6n \log p_n \leq \frac{12}{e} n \sqrt{p_n} \implies \sqrt{p_n} < \frac{12}{e} n$$

thus, applying log on both sides (since it is an increasing function) yields

$$\log p_n < 2 \log n + 2 \log \frac{12}{e}$$

and since $p_n < 6n \log p_n$, we have finally

$$p_n < 12\left(n \log n + n \log \frac{12}{e}\right),$$

proving the theorem. ■

3.3 Selberg's asymptotic formula

Almost any elementary proof (and some non-elementary proofs), specially the first ones by Selberg himself and Erdös independently, rely on the following identity proved by Selberg in 1948.

Theorem 12 (Selberg's asymptotic formula). *For $x > 0$ we have*

$$\vartheta(x) \log x + \sum_{p \leq x} \log p \, \vartheta\left(\frac{x}{p}\right) = 2x \log x + O(x).$$

Which we will prove via the following equivalent identity

$$\psi(x)\log x+\sum_{n\le x}\Lambda(n)\psi\Big(\frac{x}{n}\Big)=2x\log x+O(x)$$

using the following theorem

Theorem 13. *Let F be a real valued function defined on $(0,\infty)$, and let*

$$G(x)=\log x\sum_{n\le x}F\Big(\frac{x}{n}\Big).$$

Then

$$F(x)\log x+\sum_{n\le x}F\Big(\frac{x}{n}\Big)\Lambda(n)=\sum_{d\le x}\mu(d)G\Big(\frac{x}{d}\Big).$$

Proof. Notice that we can write $F(x)\log x$ as

$$\sum_{n\le x}F\Big(\frac{x}{n}\Big)\log\frac{x}{n}\left\lfloor\frac{1}{n}\right\rfloor=\sum_{n\le x}F\Big(\frac{x}{n}\Big)\log\frac{x}{n}\sum_{d|n}\mu(d), \tag{3.6}$$

since $\lfloor 1/n\rfloor$ is nonzero (and equal to 1) only at $n=1$, where the second expression comes from proposition 2. We can also write a similar equation using proposition 4

$$\sum_{n\le x}F\Big(\frac{x}{n}\Big)\Lambda(n)=\sum_{n\le x}F\Big(\frac{x}{n}\Big)\sum_{d|n}\mu(d)\log\frac{n}{d}. \tag{3.7}$$

We can add (3.6) and (3.7) to get that

$$\begin{aligned}F(x)\log x+\sum_{n\le x}F\Big(\frac{x}{n}\Big)\Lambda(n)&=\sum_{n\le x}F\Big(\frac{x}{n}\Big)\sum_{d|n}\mu(d)\Big(\log\frac{x}{n}+\log\frac{n}{d}\Big)\\&=\sum_{n\le x}\sum_{d|n}F\Big(\frac{x}{n}\Big)\mu(d)\log\frac{x}{d}\end{aligned}$$

and just let $n = qd$ so that

$$\sum_{n\le x}\sum_{d|n} F\Big(\frac{x}{n}\Big)\mu(d)\log\frac{x}{d} = \sum_{d\le x}\mu(d)\log\frac{x}{d}\sum_{q\le x/d} F\left(\frac{x}{qd}\right) = \sum_{d\le x}\mu(d)G\Big(\frac{x}{d}\Big). \qquad \blacksquare$$

And we can use this to prove the equivalent form of Theorem 12, whose proof goes as follows.

Proof of Selberg's asymptotic formula. Let $F_1(x) = \psi(x)$ and $F_2(x) = x - \gamma - 1$ where γ is Euler's constant which we already introduced in Chapter 2. Then we apply the previous theorem to these two functions so that for F_1, using theorem 7,

$$\begin{aligned} G_1(x) &= \log x \sum_{n\le x} \quad \Big(\frac{x}{n}\Big) \\ &= \log x(x\log x - x + O(\log x)) \\ &= x\log^2 x - x\log x + O(\log^2 x). \end{aligned}$$

For F_2 we have

$$\begin{aligned} G_2(x) &= \log x \sum_{n\le x}\Big(\frac{x}{n} - \gamma - 1\Big) \\ &= x\log x\sum_{n\le x}\frac{1}{n} - (\gamma+1)\log x\sum_{n\le x} 1 \\ &= x\log x\left(\log x + \gamma + O\left(\frac{1}{x}\right)\right) - (\gamma+1)\log x(x + O(1)) \\ &= x\log^2 x - x\log x + O(\log x) \end{aligned}$$

where the third line is justified by parts **(a)** and **(b)** of proposition 5, then we have that $G_1(x) - G_2(x) = O(\log^2 x) = O(\sqrt{x})$ (the second is less refined but it is enough to work). We can now study the right hand side of the previous theorem for $G_1 - G_2$ so that

$$\begin{aligned}\sum_{d\le x}\left(G_1\left(\frac{x}{d}\right)-G_2\left(\frac{x}{d}\right)\right) &= O\left(\sum_{d\le x}\sqrt{\frac{x}{d}}\right)\\ &= O\left(\sqrt{x}\sum_{d\le x}\frac{1}{\sqrt{d}}\right)\\ &= O\left(\sqrt{x}\frac{\sqrt{x}}{1/2}+O\left(\frac{1}{\sqrt{x}}\right)\right)\\ &= O(x)\end{aligned}$$

also using proposition 5 **(b)**, therefore

$$\begin{aligned}(F_1-F_2)(x)\log x+\sum_{n\le x}(F_1-F_2)\left(\frac{x}{n}\right)\Lambda(n) &= [\psi(x)-(x-\gamma-1)]\log x+\\ &\quad+\sum_{n\le x}\left[\quad\left(\frac{x}{n}\right)-\left(\frac{x}{n}-\gamma-1\right)\right]\Lambda(n)\\ &= O(x)\end{aligned}$$

from which we can get

$$\begin{aligned}\psi(x)\log x+\sum_{n\le x}\quad\left(\frac{x}{n}\right)\Lambda(n) &= (x-\gamma-1)\log x\\ &\quad+\sum_{n\le x}\left(\frac{x}{n}-\gamma-1\right)\Lambda(n)+O(x)\end{aligned}$$

and by proposition 7, remembering the definition of ψ as the partial sum of Λ, we can get

$$\begin{aligned}\psi(x)\log x+\sum_{n\le x}\quad\left(\frac{x}{n}\right)\Lambda(n) &= x\log x+x\sum_{n\le x}\frac{\Lambda(n)}{n}+O(x)\\ &= 2x\log x+O(x)\end{aligned}$$

as required, thus proving Selberg's asymptotic formula. ■

Chapter 4

Elementary proof of the PNT

4.1 Plan of prooof

The proof that I will present is in [5], the original article Selberg himself presented on the Annals of Mathematics in 1949, titled *An Elementary Proof of the Prime-Number Theorem*. It relies on his asymptotic formula (theorem 12) and estimates made by Chebyshev. We will follow Selberg's proof. The principal idea behind the proof is to find strict inequalities at sufficiently large scales where the PNT is approximately true, which will give us precise statements about when the error term in the approximation of the PNT is small enough for large enough quantities so we can prove it approximates 0 at ∞.

The version of the theorem that will be used is (3.4), namely that $\vartheta(x) \sim x$ as $x \to \infty$, equivalently

$$\lim_{x\to\infty} \frac{\vartheta(x)}{x} = 1.$$

But first we will, noting that $\vartheta(x) = O(x)$, write $\vartheta(x) = x + R(x)$ (where $R(x) = O(1)$). This will make so the PNT is equivalent to the statement

$$\lim_{x\to\infty} \frac{R(x)}{x} = 0,$$

that will be proved studying the properties of $R(x)$, where we will make use of clever bounds an properties of the logarithm to conclude that if for some $\alpha < 8$, $R(x) < \alpha x$, then for some constant $K > 0$

$$|R(x)| < \alpha\left(1 - \frac{\alpha^2}{300K}\right)x$$

and since the sequence defined recursively by $\alpha_{n+1} = \alpha_n\left(1 - \frac{\alpha_n^2}{300K}\right)$ converges to zero, this will prove the main result.

4.2 Preliminary results

4.2.1 Bounds for the error term

The main idea is that $\vartheta = O(x)$, therefore, as we planned, we can write $\vartheta(x) = x + R(x)$, so we can bound cleverly the error term $R(x) = \vartheta(x) - x$. To do this, first consider the following proposition, taking into account that from now on not only p but also q will denote only prime numbers.

Proposition 9. *For $x \geq 4$ we have*

$$\vartheta(x)\log x = \sum_{pq \leq x} \frac{\log p \log q}{\log pq}\vartheta\left(\frac{x}{pq}\right) + O(x \log\log x)$$

The proof can be found in [5], however the idea is to prove, with Abel's summation formula and Möbius inversion, the inequality

$$\sum_{p \leq x} \log^2 p + \sum_{pq \leq x} \log p \log q = x \sum_{d \leq x} \frac{\mu(d)}{d} \log^2 \frac{x}{d} + O(x) \tag{4.1}$$

together with

$$\sum_{p \leq x} \log^2 p = \vartheta(x) \log x + O(x)$$

and some asymptotic formulas for $\sum_{n\le x}\frac{1}{n}$ (which we already had in proposition 5) and $\sum_{n\le x} d(n)/n$ where d is the divisor function.

Proposition 9 will imply, replacing $\vartheta(x)$ with $x+R(x)$, that

$$(x+R(x))\log x = \sum_{pq\le x}\frac{\log p\log q}{\log pq}\left(\frac{x}{pq}+R\left(\frac{x}{pq}\right)\right)+O(x\log\log x) \tag{4.2}$$

$$= x\sum_{pq\le x}\frac{\log p\log q}{pq\log pq}+\sum_{pq\le x}\frac{\log p\log q}{\log pq}R\left(\frac{x}{pq}\right)+O(x\log\log x). \tag{4.3}$$

The first sum can be dealt with the help of the inequality

$$\sum_{p\le x}\frac{\log p}{p}=\log x+O(1) \tag{4.4}$$

of proposition 7, by applying Abel's summation formula with $a(n)=\sum_{pq=n}\log p\log q/(pq)$ and $f(n)=1$, we get

$$\sum_{pq\le x}\frac{\log p\log q}{pq}=\frac{1}{2}\log^2 x+O(\log x),$$

and applying Abel's summation again we get

$$\sum_{pq\le x}\frac{\log p\log q}{pq\log pq}=\log x+O(\log\log x)$$

so that (4.3) becomes

$$(x+R(x))\log x = x\log x+O(x\log\log x)+\sum_{pq\le x}\frac{\log p\log q}{\log pq}R\left(\frac{x}{pq}\right)+O(x\log\log x)$$

thus leaving us with

$$R(x)\log x=\sum_{pq\le x}\frac{\log p\log q}{\log pq}R\left(\frac{x}{pq}\right)+O(x\log\log x). \tag{4.5}$$

A similar inequality can be given by replacing $\vartheta(x)$ with $x+R(x)$ in Selberg's asymptotic formula (theorem 12), namely that

$$(x+R(x))\log x + \sum_{p\le x}\log p\left(\frac{x}{p}+R\left(\frac{x}{p}\right)\right) = 2x\log x + O(x)$$

so that with (4.4) we get

$$R(x)\log x = -\sum_{p\le x}\log pR\left(\frac{x}{p}\right) + O(x). \tag{4.6}$$

Notice that then, by (4.5) and (4.6), we can bound $R(x)$ with the use of the triangle inequality as follows

$$2|R(x)|\log x \le \sum_{p\le x}\log p\left|R\left(\frac{x}{p}\right)\right| + \sum_{pq\le x}\frac{\log p\log q}{\log pq}\left|R\left(\frac{x}{pq}\right)\right| + O(x\log\log x)$$

which with partial summation and help from (4.1) yields, finally, that

$$|R(x)| \le \frac{1}{\log x}\sum_{n\le x}\left|R\left(\frac{x}{n}\right)\right| + O\left(x\frac{\log\log x}{\log x}\right). \tag{4.7}$$

4.2.2 Approximating the PNT

The last inequality in proposition 7 gives us that

$$\sum_{n\le x}\frac{\vartheta(n)}{n^2} = \log x + O(1)$$

thus, remembering that $\vartheta(x) = x + R(x)$, we have

$$\sum_{n\le x}\frac{\vartheta(n)}{n^2} = \sum_{n\le x}\frac{n+R(n)}{n^2} = \log n + \sum_{n\le x}\frac{R(n)}{n^2}$$

implying

$$\sum_{n\le x}\frac{R(n)}{n^2}=O(1).$$

then, by definition, for some $a>0$, for all $x>a$ and $x'>x$, we have

$$\left|\sum_{x\le n\le x'}\frac{R(n)}{n^2}\right|<K_1 \tag{4.8}$$

for some positive constant K_1. Moreover, if it happens that $R(x)$ does not change sign in $[x,x']$ we have

$$\inf_{x\le y\le x'}\left|\frac{R(y)}{y}\right|\left(\log\frac{x'}{x}+O(1)\right)\le\left|\sum_{x\le n\le x'}\frac{R(n)}{n^2}\right|<K_1$$

then there exist $y\in[x,x']$ such that for some positive constant $K_2\ge 1$

$$\left|\frac{R(y)}{y}\right|<\frac{K_2}{\log\frac{x'}{x}} \tag{4.9}$$

however, since if it changes sign there exists y in that interval with $|R(y)|<\log y$, the result follows for all intervals $[x,x']$.

Notice that the interval just requires $x'>x>a$, the length of the interval is arbitrarily large or small. This implies, together with (4.8) and (4.9) that for any $0<\delta<1$ and $x>a$ there will exist $y\in\left[x,e^{K_2/\delta}x\right]$ such that

$$|R(y)|<\delta y. \tag{4.10}$$

then using the following inequality

$$\sum_{p\le x}\log p+\sum_{pq\le x}\frac{\log p\log q}{\log pq}=2x+O\left(\frac{x}{\log x}\right)$$

which can be easily derived from partial summation with (4.1), we can get that for $y'>y$,

$y > 4$,

$$0 \leq \sum_{y<p\leq y'} \log p \leq 2(y'-y) + O\left(\frac{y'}{\log y'}\right)$$

thus showing that

$$|R(y') - R(y)| \leq y' - y + O\left(\frac{y'}{\log y'}\right)$$

which imply a nicer statement

$$|R(y')| \leq |R(y)| + |y - y'| + O\left(\frac{y'}{\log y'}\right) \tag{4.11}$$

for $y/2 \leq y' \leq 2y$ and $y > 4$.

Finally, we have that for $x > 4$ and $y \in \left[x, e^{K_2/\delta}x\right]$ s.t. (4.10) follows, if $y' \in [y/2, 2y]$, then by (4.11)

$$|R(y')| \leq \delta y + |y' - y| + \frac{K_3 y'}{\log x}$$

so $|R(y')/y'| < 2\delta + |1 - y'/y| + K_3/\log x$, thus if $x > e^{K_3/\delta}$ and $y'/y \in [e^{-\delta/2}, e^{\delta/2}]$, the last equation becomes

$$\left|\frac{R(y')}{y'}\right| < 2\delta + (e^{\delta/2} - 1) + \delta < 4\delta.$$

Then we have proved that for large enough x, more precisely, $x > e^{K_3/\delta}$, the interval $[x, e^{K_2/\delta}x]$ contains another interval $[y_1, e^{\delta/2}y_1]$ such that $|R(z)| < 4\delta z$ for all z in such interval.

4.3 The proof

Finally, the proof of the prime number theorem goes as follows.

Proof of the Prime Number Theorem. We know that for $\vartheta(x) = x + R(x)$, the PNT is equivalent to the statement

$$\lim_{x\to\infty}\frac{\vartheta(x)}{x}=1 \qquad \Longleftrightarrow \qquad \lim_{x\to\infty}\frac{R(x)}{x}=0.$$

We have that if $x>1$ then since $R(x)=O(x)$, $|R(x)<K_4x|$ for some positive constant K_4. Assume that for some $0<\alpha<8$ s.t. $|R(x)|<\alpha x$ holds for all $x>x_0$ for some x_0. Let $\delta=\alpha/8$ and assume now $x_0>e^{K_3/\delta}$ then by our previous results, all intervals $[x, e^{K_2/\delta}x]$ with $x>x_0$ contains intervals $[y_1, e^{\delta/2}y_1]$ with

$$|R(z)|<4\frac{\alpha}{8}z=\frac{\alpha}{2}z$$

for all z in that interval.

Now let us apply (4.7) together with $|R(z)|<K_4x$ to get

$$|R(x)|<K_4\frac{x}{\log x}\sum_{x/x_0<n\le x}\frac{1}{n}+\frac{x}{\log x}\sum_{n\le x/x_0}\frac{1}{n}\left|\frac{1}{n}R\left(\frac{x}{n}\right)\right|+O\left(\frac{x}{\sqrt{\log x}}\right)$$

now, letting $\rho=e^{K_2/\delta}$, and using $|R(x)|<\alpha x$ for $x>x_0$ and $|R(z)|<\alpha z/2$ for the interval mentioned above, we can write

$$\begin{aligned}
R(x)&<\frac{\alpha x}{\log x}\sum_{n\le x/x_0}\frac{1}{n}-\frac{\alpha x}{2\log x}\sum_{1\le\nu\le(\log(x/x_0))/\log p}\ \sum_{\substack{y_\nu\le n\le y_\nu e^{\delta/2}\\ \rho^{\nu-1}<y_\nu\le\rho^\nu e^{-\delta/2}}}\frac{1}{n}+O\left(\frac{x}{\sqrt{\log x}}\right)\\
&=\alpha x-\frac{\alpha x}{2\log x}\sum_{1\le\nu\le(\log(x/x_0))/\log p}\frac{\delta}{2}+O\left(\frac{x}{\sqrt{\log x}}\right)\\
&=\alpha x-\frac{\alpha\delta}{4\log\rho}x+O\left(\frac{x}{\sqrt{\log x}}\right)
\end{aligned}$$

and remembering that $\delta=\alpha/8$ we have

$$R(x)<\alpha\left(1-\frac{\alpha\delta}{32K_2}\right)x=\alpha\left(1-\frac{\alpha^2}{256K_2}\right)x<\alpha\left(1-\frac{\alpha^2}{300K_2}\right)x.$$

Notice that from $|R(x)| < \alpha x$ for $x > 1$ we concluded $|R(x)| < \alpha\left(1 - \frac{\alpha^2}{300K_2}\right)x$ for any $x > x_1$. Since the sequence defined recursively by

$$\alpha_{n+1} = \alpha_n\left(1 - \frac{\alpha_n^2}{300K_2}\right)x$$

converges to 0 as $n \to \infty$, this proves that $R(x)/x$ approaches 0 as $x \to \infty$, since we can eventually bound it by above with a sequence converging to 0, and thus that

$$\lim_{x\to\infty} \frac{\vartheta(x)}{x} = 1. \qquad \blacksquare$$

Chapter 5

More on the PNT

5.1 About the analytical proof of the PNT

Since the first proofs of the PNT were complex-analytic in nature, these are called classical or analytic proofs (as opposed to elementary proofs). A complete classical proof can be found in [1], in which the main idea is to consider statement

$$\psi_1(x) = \int_1^x \psi(t)\,\mathrm{d}t \sim \frac{1}{2}x^2 \text{ as } x \to \infty$$

which would imply the PNT in the form $\psi(x) \sim x$.

For that purpose, it is necessary to dive into contour integration in the complex plane and the fact that

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

for $Re(s) > 1$, so that for $c > 1$ and $x \geq 1$

$$\frac{\psi_1(x)}{x^2} = \frac{1}{2\pi i}\int_{c-\infty i}^{c+\infty i} \frac{x^{s-1}}{s(s+1)}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\mathrm{d}s$$

which then becomes

$$\frac{\psi_1(x)}{x^2} - \frac{1}{2}\left(1 - \frac{1}{x}\right)^2 = \frac{x^{c-1}}{2\pi}\int_{-\infty}^{\infty} h(c+it)e^{it\log x}\,\mathrm{d}t \tag{5.1}$$

for

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

The PNT then ends if (5.1) could be used exactly at $c = 1$, which is what needs to be proved. No more of the proof will be presented, but the rest can be summarized in that last task, which nonetheless implies a lot of work studying contour integrals and the behaviour of ζ and ζ' near $Re(s) = 1$.

The deepness on the theorems that go into this proof is much more intricate with complex analysis compared to what we did in Chapter 4, which was mostly elementary, therefore making a "disconnection" with the world of number theory or even analytic number theory that relies mostly on basic calculus theorems and properties of natural numbers.

5.2 Dirichlet Theorem on primes in arithmetic progressions

5.2.1 Primes in arithmetic progressions

An arithmetic progression with initial term h and difference k is a sequence $\{h + kn\}_{n\in\mathbb{N}\cup\{0\}}$, i.e. $\{h, h+k, h+2k, \ldots\}$. One can ask about primes in a given sequence (equivalently, given h and k). For example, the sequence of positive integers ($h = 1$, $k = 1$) has all the primes and has been know to have infinitely many primes since, at least, Euclid, the sequence of positive odd numbers ($h = 1$, $k = 2$) has infinitely many primes and a similar story, while positive even numbers have only 1 prime, and any sequence with non prime h and k with $(h, k) > 1$ has exactly 0 prime numbers.

Playing with sequences might get you to believe that no matter h, k, if $(h, k) = 1$ then the sequence $\{h + k_n\}_{n \in \mathbb{N} \cup \{0\}}$ must contain a prime number if not infinitely many. The next theorem states exactly that

Theorem 14 (Dirichlet's Theorem on Primes in Arithmetical Progressions). *If* $h, k \in \mathbb{N}$ *such that* $(h, k) = 1$*, then there are infinitely many primes of the form* $kn + h$ *for* $n \in \mathbb{N} \cup \{0\}$.

A complete proof can be found in chapter 7 of [1]. It is said that the proof of this theorem by Dirichlet, conjectured first by Legendre, was a huge step in the right direction to the proof of the PNT, since it uses techniques and mathematical tools that Dirichlet himself developed.

However, there's a more precise statement about this fact

Theorem 15. *If* $k > 0$ *and* $(h, k) = 1$ *we have, for all* $x > 1$*,*

$$\sum_{\substack{p \le x \\ p \equiv h \pmod k}} \frac{\log p}{p} = \frac{1}{\varphi(k)} \log x + O(1).$$

Notice that this last theorem is stronger than the original statement of Dirichlet's theorem, since it implies not only that for coprime h and k there are infinitely many primes of the form $h + kn$, but that each of the $\varphi(k)$ reduced residue classes modulo k contributes the same to the already presented identity

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

5.2.2 A generalization of the PNT

We can define a prime counting function like π for primes of the form $kn+h$ (or congruent to h module k). For $k>0$ and $(a,k)=1$ we define

$$\pi_a(x) = \sum_{\substack{p\leq x \\ p\equiv a \pmod l}} 1.$$

Then $\pi_a(x)$ counts the number of primes of the form $a+kn$ up to $\lfloor x \rfloor$. Dirichlet's theorem just proves $\pi_a(x) \to \infty$ as $x \to \infty$, but there's a generalization of the PNT, called the prime number theorem for arithmetical progressions which states that

$$\pi_a(x) \sim \frac{\pi(x)}{\varphi(k)} \sim \frac{1}{\varphi(k)}\frac{x}{\log x}$$

in a similar fashion to theorem 15.

It is needless to say that the main focus of study has always been the original PNT and this is not the only generalization of the PNT that is done, another, perhaps more abstract, generalization involving defining a new notion of primality can be found in [4]

.

Chapter 6

Conclusions

The prime number theorem has been a great milestone in the history of mathematics, it has caused a lot of progress in many different fields, since even if it seems pointless or “useless” to study topics in number theory, mathematicians are always puzzled and amazed by what such simple and natural[1] objects can present so many challenges and questions that may seem elementary but may be deeper than noticed by a first sight. The greatests mathematicians at their times couldn’t find a proof of the PNT for almost a century, and not an elementary proof for almost two centuries.

This work presented the elementary proof in its entirety, using only results from number theory, elementary asymptotic relations and properties of elementary real functions. In chapter 5 it was also explained in a summary the principal ideas in the shorter analytic proof, noticing that it was very different to the elementary proof. In fact, the nature of the analytic proof is so different to anything alike in number theory that G. H. Hardy proposed in 1921 a “hierarchy” of results in mathematics based on what is necessary to prove them. Anything that used deep results, such as results from complex analysis, was deemed to be deep and thus unprovable by elementary methods, category in which the PNT fit until Selberg’s and Erdös’s proofs.

[1]Kronecker was the one who said the famous quote I included in the preliminaries: Natural numbers were created by God, everything else is the work of men.

After the first elementary proofs, many more elementary, analytic and hybrid proofs appeared, even a very short proof by D. J. Newman that is not elementary but also not too deep into the complex realm. The lesson from the history of the PNT is, in conclusion, that if a result can be provable by complex and advanced or intricate methods, it may still be the case that one can prove the same result using only the most basic stuff in mathematics. Famous results such as Fermat's last theorem have been proved only by using advanced mathematics, specifically for the mentioned it makes heavy use of harmonic analysis and modular forms, but it may still be possible that an elementary proof (for such an "elementary" problem) can be found, but it may be a long time, as happened with the PNT, until someone makes enough progress to tie the (many) knots that may join elementary results and theorems that today can be seen as deep.

Bibliography

[1] Apostol, T. M. (1976). *Introduction to analytic number theory* (1st ed.). Springer New York.

[2] Finch, S. R. (2003). *Mathematical constants* (1st ed.). Cambridge University Press.

[3] Goldstein, L. J. (1973). A history of the prime number theorem. *The American Mathematical Monthly*, *80*(6), 599–615. Retrieved November 12, 2023, from http://www.jstor.org/stable/2319162

[4] Hall, R. (1972). The prime number theorem for generalized primes. *Journal of Number Theory*, *4*(4), 313–320. https://doi.org/https://doi.org/10.1016/0022-314X(72)90066-2

[5] Selberg, A. (1949). An elementary proof of the prime-number theorem. *Annals of Mathematics*, *50*(2), 305–313. Retrieved September 14, 2023, from http://www.jstor.org/stable/1969455

[6] Tattersall, J. J. (2005). *Elementary number theory in nine chapters* (2nd ed.). Cambridge University Press.