

Chapter 4

Elementary Results on the Distribution of Primes

4.1 Introduction

Definition 4.1. For real number $x > 0$, let $\pi(x)$ denote the number of primes not exceeding x .

The behavior of $\pi(x)$ as the function of x has been studied by many mathematicians ever since the eighteenth century. Inspection of tables of primes led C.F. Gauss (1792) and A.M. Legendre (1798) to conjecture that

$$\pi(x) \sim \frac{x}{\ln x}. \quad (4.1)$$

This conjecture was first proved independently by J. Hadamard and de la Vallée Poussin in 1896 and is now known as the *Prime Number Theorem*. We record the theorem as follows:

Theorem 4.1 (Prime Number Theorem). *Let x be a real positive number and $\pi(x)$ be the number of primes less than x . Then*

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

Proofs of the Prime Number Theorem are often classified as elementary or analytic. The proofs of J. Hadamard and de la Vallée Poussin are analytic, using complex function theory and properties of the Riemann zeta function $\zeta(s)$ (see Definition 3.3 for the definition of $\zeta(s)$ when $s \in \mathbf{R}$ and $s > 1$). Elementary proofs were discovered around 1949 by A. Selberg and P. Erdős. Their proofs do not involve $\zeta(s)$ and complex function theory, hence the name “elementary”.

There are other elementary proofs of the prime number theorem since the appearance of the work of Selberg and Erdős, one of which is [4]. The

proof given in [4] relies on proving an equivalent statement of the Prime Number Theorem and the mean value of $\mu(n)$. This equivalent statement of the Prime Number Theorem will be established in Section 4.5.

In this chapter, we derive some basic properties of $\pi(x)$ and establish several statements equivalent to the Prime Number Theorem. We will also use the results discussed in this chapter to study Bertrand's Postulate, which states that for $n \geq 2$, there exists a prime between n and $2n$.

4.2 The function $\psi(x)$

We recall the definition of Mangoldt's function (see Exercise 2.6, Problem 1).

Definition 4.2. Let n be a positive integer and let

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

Definition 4.3. For real number $x \geq 1$,

$$\psi(x) = \sum_{n \leq x} \Lambda(n) = \sum_{p^m \leq x} \ln p.$$

Theorem 4.2. *There exist positive constants c_1 and c_2 such that*

$$c_1 x \leq \psi(x) \leq c_2 x.$$

Proof. For $x \geq 4$, let

$$S = \sum_{n \leq x} \ln n - 2 \sum_{n \leq \frac{x}{2}} \ln n.$$

By Theorem 3.2 with $f(n) = \ln n$, we find that

$$\begin{aligned} \sum_{n \leq x} \ln n &= \int_1^x \ln t dt + \int_1^x \{t\} \frac{1}{t} dt - \{x\} \ln x + \{y\} \ln y \\ &= x \ln x - x + O(\ln x). \end{aligned} \tag{4.2}$$

This implies that

$$S = x \ln 2 + O(\ln x).$$

Therefore, there exists an $x_0 \geq 4$ such that

$$\frac{x}{2} \leq S \leq x \tag{4.3}$$

whenever $x \geq x_0 \geq 4$. Next, since (see Section 2.6, Problem 1)

$$\ln n = \sum_{d|n} \Lambda(d),$$

we find that

$$\begin{aligned} S &= \sum_{n \leq x} \sum_{d|n} \Lambda(d) - 2 \sum_{n \leq \frac{x}{2}} \sum_{d|n} \Lambda(d) \\ &= \sum_{d \leq x} \Lambda(d) \left[\frac{x}{d} \right] - 2 \sum_{d \leq \frac{x}{2}} \Lambda(d) \left[\frac{x}{2d} \right] \\ &= \sum_{d \leq \frac{x}{2}} \Lambda(d) \left\{ \left[\frac{x}{d} \right] - 2 \left[\frac{x}{2d} \right] \right\} + \sum_{\frac{x}{2} < d \leq x} \Lambda(d) \left[\frac{x}{d} \right]. \end{aligned}$$

Hence,

$$S = \sum_{d \leq \frac{x}{2}} \Lambda(d) \theta_d + \sum_{\frac{x}{2} < d \leq x} \Lambda(d) \left[\frac{x}{d} \right],$$

where

$$\theta_d = \left[\frac{x}{d} \right] - 2 \left[\frac{x}{2d} \right]. \quad (4.4)$$

Now, for

$$\frac{x}{2} < d \leq x,$$

we have

$$\left[\frac{x}{d} \right] = 1.$$

Therefore, we may simplify the second term on the right-hand side of (4.4) to obtain

$$S = \sum_{d \leq \frac{x}{2}} \Lambda(d) \theta_d + \sum_{\frac{x}{2} < d \leq x} \Lambda(d). \quad (4.5)$$

We now observe that $\theta_d = 0$ or 1 since

$$[y] - 2[y/2] = 0 \quad \text{or} \quad 1.$$

Using (4.5), we deduce that

$$S \leq \sum_{d \leq \frac{x}{2}} \Lambda(d) + \sum_{\frac{x}{2} < d \leq x} \Lambda(d) = \sum_{d \leq x} \Lambda(d) = \psi(x) \quad (4.6)$$

and

$$S \geq \sum_{\frac{x}{2} < d \leq x} \Lambda(d) = \psi(x) - \psi\left(\frac{x}{2}\right). \quad (4.7)$$

From (4.3) and (4.6),

$$\psi(x) \geq S \geq \frac{x}{2} \quad (x \geq x_0).$$

Therefore,

$$\psi(x) \geq c_1 x.$$

To obtain a lower bound for $\psi(x)$, we first deduce from (4.3), (4.7) that

$$\psi(x) - \psi\left(\frac{x}{2}\right) \leq S \leq x.$$

Therefore,

$$\begin{aligned} \psi(x) &\leq x + \psi\left(\frac{x}{2}\right), \quad x \geq x_0 \\ &\leq x + \frac{x}{2} + \psi\left(\frac{x}{4}\right), \quad x \geq 2x_0 \\ &\vdots \\ &\leq x + \frac{x}{2} + \cdots + \frac{x}{2^k} + \psi\left(\frac{x}{2^{k+1}}\right), \quad \frac{x}{2^{k+1}} < x_0 \leq \frac{x}{2^k}. \end{aligned}$$

This implies that

$$\psi(x) \leq 2x + \psi(x_0) \leq c_2 x$$

for some positive real number c_2 . □

4.3 The functions $\theta(x)$ and $\pi(x)$

Definition 4.4. For real number $x \geq 1$, let

$$\theta(x) = \sum_{p \leq x} \ln p.$$

Theorem 4.3. For real number $x \geq 1$, we have

$$\theta(x) = \psi(x) + O(\sqrt{x}).$$

Proof. We first note that the difference of $\psi(x)$ and $\theta(x)$ is

$$\begin{aligned} \psi(x) - \theta(x) &= \sum_{\substack{p^m \leq x \\ m \geq 2}} \ln p \\ &= \sum_{\substack{p \leq \sqrt{x} \\ m=2}} \ln p + \sum_{p \leq x^{1/3}} \ln p + \sum_{3 \leq m \leq \frac{\ln x}{\ln p}} 1. \end{aligned}$$

Hence,

$$\begin{aligned}\psi(x) - \theta(x) &\leq \psi(\sqrt{x}) + \sum_{p \leq x^{1/3}} \ln p \frac{\ln x}{\ln p} \\ &\ll \sqrt{x} + x^{1/3} \ln x \ll \sqrt{x},\end{aligned}$$

where $f(x) \ll g(x)$ is another notation for $f(x) = O(g(x))$ (see Definition 3.1). \square

Using Theorems 4.2 and 4.3, we deduce the following corollary.

Corollary 4.4. *For $x \geq 4$, there exist real positive constants c_1 and c_2 such that*

$$c_1 x \leq \theta(x) \leq c_2 x.$$

We give a relation between $\theta(x)$ and $\pi(x)$, where $\pi(x)$ is given by Definition 4.1.

Theorem 4.5. *For each positive real $x \geq 4$,*

$$\frac{c_1 x}{\ln x} \leq \pi(x) \leq \frac{c_2 x}{\ln x}.$$

Proof. It suffices to prove that

$$\pi(x) = \frac{1}{\ln x} \theta(x) + O\left(\frac{x}{\ln^2 x}\right)$$

by Theorem 4.3. We observe that

$$\begin{aligned}\pi(x) - \frac{\theta(x)}{\ln x} &= \sum_{p \leq x} \left(1 - \frac{\ln p}{\ln x}\right) \\ &= \sum_{p \leq x} \ln p \left(\frac{1}{\ln p} - \frac{1}{\ln x}\right).\end{aligned}\tag{4.8}$$

If

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

then by Corollary 4.4,

$$A(t) = \sum_{n \leq t} a(n) = \theta(t) \ll t.$$

The last expression in (4.8) is

$$\begin{aligned} \theta(x) \left(\frac{1}{\ln x} - \frac{1}{\ln x} \right) - \int_2^x \theta(t) \left(\frac{1}{\ln t} - \frac{1}{\ln x} \right)' dt \\ = \int_2^x \frac{\theta(t)}{t \ln^2 t} dt \ll \int_2^x \frac{dt}{\ln^2 t} \\ = \int_2^{\sqrt{x}} \frac{dt}{\ln^2 t} + \int_{\sqrt{x}}^x \frac{dt}{\ln^2 t} \\ \ll \sqrt{x} + \int_{\sqrt{x}}^x \frac{dt}{\ln^2 t} \ll \frac{x}{\ln^2 x}. \quad \square \end{aligned}$$

As corollaries of Theorems 4.3 and 4.5, we have the following results. We leave the details of the proofs of these corollaries to the readers.

Corollary 4.6. *The Prime Number Theorem*

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

is equivalent to each of the following relations:

- (a) $\theta(x) \sim x$, and
- (b) $\psi(x) \sim x$.

4.4 Merten’s estimates

In this section, we show that there are infinitely many primes by showing that $\sum_{p \leq x} \frac{1}{p}$ diverges.

Theorem 4.7 (Merten’s estimates). *Let x be a positive real number greater than 1. We have*

- (a) $\sum_{n \leq x} \frac{\Lambda(n)}{n} = \ln x + O(1)$,
- (b) $\sum_{p \leq x} \frac{\ln p}{p} = \ln x + O(1)$,
- (c) $\sum_{p \leq x} \frac{1}{p} = \ln \ln x + A + O\left(\frac{1}{\ln x}\right)$, and
- (d) (Merten’s Theorem) $\prod_{p \leq x} \left(1 - \frac{1}{p}\right) = \frac{e^{-A}}{\ln x} \left(1 + O\left(\frac{1}{\ln x}\right)\right)$,

where A is a constant.

Proof.

(a) First, we write

$$\begin{aligned} \sum_{n \leq x} \frac{\Lambda(n)}{n} &= \sum_{n \leq x} \left\{ \Lambda(n) \frac{1}{x} \left(\left[\frac{x}{n} \right] + O(1) \right) \right\} \\ &= \frac{1}{x} \sum_{n \leq x} \Lambda(n) \left[\frac{x}{n} \right] + O \left(\frac{1}{x} \sum_{n \leq x} \Lambda(n) \right). \end{aligned}$$

By (3.8), we find that

$$\sum_{n \leq x} \Lambda(n) \left[\frac{x}{n} \right] = \sum_{n \leq x} (\Lambda * u)(n).$$

Hence, we deduce using (2.12) and (4.2) that

$$\begin{aligned} \sum_{n \leq x} \frac{\Lambda(n)}{n} &= \frac{1}{x} \sum_{n \leq x} (\Lambda * u)(n) + O(1) \\ &= \frac{1}{x} \sum_{n \leq x} \ln n + O(1), \\ &= \ln x + O(1). \end{aligned}$$

(b) We observe that

$$\begin{aligned} 0 &\leq \sum_{n \leq x} \frac{\Lambda(n)}{n} - \sum_{p \leq x} \frac{\ln p}{p} = \sum_{p \leq \sqrt{x}} \ln p \sum_{2 \leq m \leq \frac{\ln x}{\ln p}} \frac{1}{p^m} \\ &\ll \sum_{p \leq \sqrt{x}} \frac{\ln p}{p^2} \ll 1. \end{aligned}$$

Hence,

$$\sum_{p \leq x} \frac{\ln p}{p} = \ln x + O(1).$$

(c) Let

$$A(x) = \sum_{n \leq x} a(n)$$

where

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

Then, we find that

$$\begin{aligned} \sum_{p \leq x} \frac{1}{p} &= \sum_{p \leq x} \left(\frac{\ln p}{p} \right) \left(\frac{1}{\ln p} \right) \\ &= A(x) \frac{1}{\ln x} - \int_2^x A(t) \left(\frac{1}{\ln t} \right)' dt \\ &= \frac{A(x)}{\ln x} + \int_2^x \frac{A(t)}{t \ln^2 t} dt. \end{aligned} \quad (4.9)$$

By Theorem 4.7 (b), we find that

$$A(t) = \ln t + R(t),$$

with

$$R(t) \ll 1, \quad t \geq 2. \quad (4.10)$$

Using (4.10) in the last term of (4.9), we deduce that

$$\begin{aligned} \int_2^x \frac{\ln t + R(t)}{t \ln^2 t} dt &= \int_2^x \frac{dt}{t \ln t} + \int_2^x \frac{R(t)}{t \ln^2 t} dt \\ &= \ln \ln x - \ln \ln 2 + \int_2^\infty \frac{R(t)}{t \ln^2 t} dt - \int_x^\infty \frac{R(t)}{t \ln^2 t} dt \\ &= \ln \ln x - \ln \ln 2 + A'' + O\left(\frac{1}{\ln x}\right). \end{aligned} \quad (4.11)$$

Substituting (4.11) into (4.9), we conclude our proof of (c).

(d) We observe that

$$\begin{aligned} \ln \prod_{p \leq x} \left(1 - \frac{1}{p} \right) &= \sum_{p \leq x} \ln \left(1 - \frac{1}{p} \right) \\ &= \sum_{p \leq x} \left(-\frac{1}{p} + r_p \right), \end{aligned}$$

where

$$r_p = \ln \left(1 - \frac{1}{p} \right) + \frac{1}{p}.$$

Hence,

$$\begin{aligned} \ln \prod_{p \leq x} \left(1 - \frac{1}{p} \right) &= \sum_{p \leq x} r_p - \sum_{p \leq x} \frac{1}{p} \\ &= -\ln \ln x + A + O\left(\frac{1}{\ln x}\right) + \sum_p r_p - \sum_{p > x} r_p. \end{aligned} \quad (4.12)$$

Now,

$$r_p = - \sum_{m=2}^{\infty} \frac{1}{mp^m} = O\left(\frac{1}{p^2}\right), \quad (4.13)$$

since for $m \geq 1$ and $p \geq 2$,

$$mp^m \geq 2^m.$$

Using (4.13) in (4.12), we deduce that

$$\begin{aligned} \ln \prod_{p \leq x} \left(1 - \frac{1}{p}\right) &= -\ln \ln x + A' + O\left(\frac{1}{\ln x}\right) + O\left(\sum_{p > x} \frac{1}{p^2}\right) \\ &= -\ln \ln x + A' + O\left(\frac{1}{\ln x}\right) + O\left(\frac{1}{x-1}\right) \\ &= -\ln \ln x + A' + O\left(\frac{1}{\ln x}\right). \end{aligned}$$

Hence,

$$\ln \prod_{p \leq x} \left(1 - \frac{1}{p}\right) = -\ln \ln x + A' + O\left(\frac{1}{\ln x}\right). \quad (4.14)$$

Exponentiating both sides of (4.14), we arrive at

$$\begin{aligned} \prod_{p \leq x} \left(1 - \frac{1}{p}\right) &= \exp\left(-\ln \ln x + A' + O\left(\frac{1}{\ln x}\right)\right) \\ &= \frac{e^{A'}}{\ln x} \exp\left(O\left(\frac{1}{\ln x}\right)\right) \\ &= \frac{e^{A'}}{\ln x} \left(1 + O\left(\frac{1}{\ln x}\right)\right), \end{aligned}$$

since $e^t = 1 + O(t)$. □

4.5 Prime Number Theorem and $M(\mu)$

Definition 4.5. Let f be an arithmetical function. We define

$$M(f) = \lim_{x \rightarrow \infty} \frac{1}{x} \sum_{n \leq x} f(n)$$

if the limit on the right-hand side exists.

As mentioned in the introduction, there are several elementary proofs of the Prime Number Theorem. One of the proofs relies on showing that $M(\mu) = 0$ (see [4]). In this section, we will show that if $M(\mu) = 0$ then the Prime Number Theorem is true. Conversely, the Prime Number Theorem implies that $M(\mu) = 0$.

Theorem 4.8. *The Prime Number Theorem is equivalent to the relation*

$$M(\mu) = 0.$$

Proof. We will first show that the Prime Number Theorem implies that $M(\mu) = 0$.

Define

$$M_1(\mu) = \lim_{x \rightarrow \infty} \left(\frac{1}{x \ln x} \sum_{n \leq x} \mu(n) \ln n \right).$$

Note that

$$M(\mu) = 0 \text{ if and only if } M_1(\mu) = 0, \tag{4.15}$$

since

$$\frac{1}{x} \sum_{n \leq x} \left| \frac{\ln n}{\ln x} - 1 \right| \ll \frac{1}{\ln x}.$$

Assume the Prime Number Theorem in the form $\theta(x) \sim x$.

We observe that

$$\begin{aligned} \sum_{n \leq x} \mu(n) \ln n &= \sum_{n \leq x} \mu(n) \sum_{d|n} \Lambda(d) \\ &= \sum_{n \leq x} \mu(n) \sum_{p|n} \ln p, \end{aligned}$$

where we have used the fact that the terms with n non-squarefree are 0. Furthermore, the value $\Lambda(d)$ is nonzero only when d is a prime power. But since n is squarefree, the divisors $d|n$ that are prime powers are simply primes. Hence,

$$\begin{aligned} \sum_{n \leq x} \mu(n) \ln n &= \sum_{p \leq x} \ln p \sum_{\substack{n \leq x \\ p|n}} \mu(n) \\ &= \sum_{p \leq x} \ln p \sum_{\substack{n' \leq x/p \\ p \nmid n'}} (\mu(n')), \end{aligned}$$

where we have written $n = pn'$. Next, observe that

$$\begin{aligned} \sum_{n \leq x} \mu(n) \ln n &= - \sum_{p \leq x} \ln p \sum_{n \leq x/p} \mu(n) + O \left(\sum_{p \leq x} \ln p \sum_{\substack{n \leq x/p \\ p|n}} 1 \right) \\ &= - \sum_{p \leq x} \ln p \sum_{n \leq x/p} \mu(n) + O \left(x \sum_{p \leq x} \frac{\ln p}{p^2} \right) \\ &= - \sum_{n \leq x} \mu(n) \sum_{p \leq x/n} \ln p + O(x) \\ &= - \sum_{n \leq x} \mu(n) \theta \left(\frac{x}{n} \right) + O(x). \end{aligned}$$

Now, write

$$\sum_{n \leq x} \mu(n) \ln n = -x \sum_{n \leq x} \frac{\mu(n)}{n} - \sum_{n \leq x} \mu(n) R \left(\frac{x}{n} \right) + O(x), \quad (4.16)$$

where

$$R(y) = \theta(y) - y.$$

Since

$$1 = \sum_{n \leq x} \sum_{d|n} \mu(d) = \sum_{n \leq x} \mu(n) \left[\frac{x}{n} \right] = x \sum_{n \leq x} \frac{\mu(n)}{n} + O(x),$$

we find that

$$x \sum_{n \leq x} \frac{\mu(n)}{n} = O(x).$$

Therefore, (4.16) may be written as

$$\sum_{n \leq x} \mu(n) \ln n = - \sum_{n \leq x} \mu(n) R \left(\frac{x}{n} \right) + O(x).$$

This implies that

$$\frac{1}{x \ln x} \left| \sum_{n \leq x} \mu(n) \ln n \right| \leq \frac{1}{x \ln x} \sum_{n \leq x} \left| R \left(\frac{x}{n} \right) \right| + O \left(\frac{1}{\ln x} \right).$$

Let $\epsilon > 0$ be given. By the Prime Number Theorem,

$$\lim_{y \rightarrow \infty} \frac{R(y)}{y} = 0.$$

Therefore, there exists a y_0 such that

$$|R(y)| \leq \epsilon y \quad (y \geq y_0).$$

For $x \geq y_0$,

$$\begin{aligned} \sum_{n \leq x} \left| R\left(\frac{x}{n}\right) \right| &\leq \sum_{n \leq x/y_0} \epsilon \frac{x}{n} + \sum_{x/y_0 < n \leq x} \max_{y \leq y_0} |R(y)| \\ &\leq \epsilon x \ln x + O_\epsilon(x). \end{aligned}$$

Thus,

$$\overline{\lim} \frac{1}{x \ln x} \left(\sum_{n \leq x} \left| R\left(\frac{x}{n}\right) \right| \right) \leq \epsilon.$$

This shows that the Prime Number Theorem implies that $M_1(\mu) = 0$. By (4.15), we deduce that $M(\mu) = 0$.

To prove the converse, let

$$\ln n = d(n) - 2C + r(n)$$

where $r(n)$ is some arithmetical function. Let $y \geq 1$. By (4.2),

$$\begin{aligned} \sum_{n \leq y} r(n) &= \sum_{n \leq y} \ln n - \sum_{n \leq y} d(n) + 2Cy + O(1) \\ &= y(\ln y - 1) + O(\ln y) \\ &\quad - (y \ln y + (2C - 1)y + O(\sqrt{y})) + 2Cy + O(1) = O(\sqrt{y}). \end{aligned} \tag{4.17}$$

Next, since

$$\Lambda = \ln * \mu,$$

we conclude that

$$\begin{aligned} \sum_{n \leq x} \Lambda(n) &= \sum_{n \leq x} (\mu * \ln)(n) \\ &= \sum_{n \leq x} (\mu * d) - 2C \left(\sum_{n \leq x} \mu * u \right) + \sum_{n \leq x} (\mu * r) \\ &= [x] - 2C + \sum_{n \leq x} (\mu * r)(n). \end{aligned}$$

The last equality holds because

$$\sum_{n \leq x} \mu * u * u = [x].$$

Thus, the Prime Number Theorem follows from $\psi(x) \sim x$ (see Corollary 4.6) if

$$\lim_{x \rightarrow \infty} \frac{1}{x} \sum_{n \leq x} (\mu * r)(n) = 0.$$

Now, by Theorem 3.4,

$$\begin{aligned} \sum_{n \leq x} \mu * r(n) &= \sum_{n \leq x} \sum_{d_1 d_2 = n} \mu(d_1) r(d_2) \\ &= \sum_{d_1 \leq x} \sum_{\substack{d_2 \leq x \\ d_1 d_2 \leq x}} \mu(d_1) r(d_2) \\ &= S_1 + S_2 - S_3, \end{aligned}$$

where

$$\begin{aligned} S_1 &= \sum_{d_2 \leq y} \sum_{d_1 \leq x/d_2} \mu(d_1) r(d_2) \\ S_2 &= \sum_{d_1 \leq x/y} \sum_{d_2 \leq x/d_1} \mu(d_1) r(d_2) \\ S_3 &= \sum_{d_1 \leq x/y} \sum_{d_2 \leq y} \mu(d_1) r(d_2), \end{aligned}$$

and y is a parameter in $[1, x]$ to be chosen.

For a fixed $y \in [1, x]$,

$$|S_1| \leq \sum_{d_2 \leq y} |r(d_2)| \left| \sum_{d_1 \leq x/d_2} \mu(d_1) \right|.$$

Now, using the assumption that

$$M(\mu) = 0,$$

we find that

$$\lim_{x \rightarrow \infty} \left| \frac{1}{x} \sum_{d_1 \leq x/d_2} \mu(d_1) \right| = 0.$$

Next, using (4.17), we deduce that

$$\begin{aligned} |S_2| &\leq \sum_{d_1 \leq x/y} \left| \sum_{d_2 \leq x/d_1} r(d_2) \right| \leq c \sum_{d_1 \leq x/y} \sqrt{\frac{x}{d_1}} \\ &\leq c\sqrt{x} \sum_{d_1 \leq x/y} \frac{1}{\sqrt{d_1}} \leq c\sqrt{x} \left(1 + \int_1^{x/y} \frac{dt}{\sqrt{t}} \right) \leq c_1 \frac{x}{\sqrt{y}}. \end{aligned}$$

Finally, by using (4.17), we deduce that

$$|S_3| \leq \frac{x}{y} \left| \sum_{d_2 \leq y} r(d_2) \right| \leq \frac{x}{y} c_2 \sqrt{y} = c_2 \frac{x}{\sqrt{y}}.$$

Hence,

$$\overline{\lim}_{x \rightarrow \infty} \frac{1}{x} \left| \sum_{n \leq x} (\mu * r)(n) \right| \leq 0 + c_1 \frac{1}{\sqrt{y}} + \frac{c_2}{\sqrt{y}}.$$

Since y is arbitrary,

$$\lim_{x \rightarrow \infty} \frac{1}{x} \sum_{n \leq x} (\mu * r)(n) = 0.$$

□

4.6 The Bertrand Postulate

In this section, we will use the properties of the functions $\theta(x)$ and $\psi(x)$ to give a proof of the well-known Bertrand’s Postulate.

Theorem 4.9 (Bertrand’s Postulate). *Let n be an integer. Then for $n \geq 2$, there exists a prime p between n and $2n$.*

Most books that discuss Theorem 4.9 prove the result following Erdős’ approach (see Exercise 4.7, Problem 8). In this book, we present the proof due to S. Ramanujan. [9] This proof was mentioned in an interesting article by P. Erdős titled “Ramanujan and I” [3]. Erdős’ proof of Theorem 4.9 was published around 1932 and it was Kalmar who asked Erdős to look up on Ramanujan’s proof and that was the first time Erdős heard about Ramanujan [3].

By definitions of $\psi(x)$ and $\theta(x)$, we observe that

Lemma 4.10. *For each positive real number x ,*

$$\psi(x) = \theta(x) + \theta(\sqrt{x}) + \theta(\sqrt[3]{x}) + \dots \tag{4.18}$$

Next, we will show that

Lemma 4.11.

$$\ln([x]!) = \psi(x) + \psi\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right) + \dots \tag{4.19}$$

Proof. The function

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

where $\Lambda(n)$ is the von Mangoldt function. Hence

$$\begin{aligned} \sum_{k=1}^{\infty} \psi\left(\frac{x}{k}\right) &= \sum_{k=1}^{\infty} \sum_{\substack{n \leq \frac{x}{k} \\ k \geq 1}} \Lambda(n) = \sum_{\substack{kn \leq x \\ k \geq 1}} \Lambda(n) \\ &= \sum_{n \leq x} \sum_{k \leq \frac{x}{n}} \Lambda(n) = \sum_{n \leq x} \left[\frac{x}{n}\right] \Lambda(n) \\ &= \sum_{n \leq x} \sum_{d|n} \Lambda(d) = \ln[x]!, \end{aligned}$$

where the last two equalities follow from (3.8) and Exercise 2.6, Problem 1(b). \square

We will now establish a few equalities and inequalities.

Lemma 4.12. For positive real number x , we have

$$\psi(x) - 2\psi(\sqrt{x}) = \theta(x) - \theta(\sqrt{x}) + \theta(\sqrt[3]{x}) - \cdots, \quad (4.20)$$

$$\ln[x]! - 2\ln[x/2]! = \psi(x) - \psi\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right) - \cdots, \quad (4.21)$$

$$\psi(x) - 2\psi(\sqrt{x}) \leq \theta(x) \leq \psi(x) \quad (4.22)$$

and

$$\begin{aligned} \psi(x) - \psi\left(\frac{x}{2}\right) &\leq \ln[x]! - 2\ln[x/2]! \\ &\leq \psi(x) - \psi\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right). \end{aligned} \quad (4.23)$$

Proof of (4.20).

This follows directly from (4.18). More precisely,

$$\psi(x) - 2\psi(\sqrt{x}) = \sum_{k=1}^{\infty} \theta\left(\sqrt[k]{x}\right) - 2 \sum_{k=1}^{\infty} \theta\left(\sqrt[2k]{x}\right). \quad \square$$

Proof of (4.21).

This follows from (4.19), namely,

$$\ln[x]! - 2\ln[x/2]! = \sum_{k=1}^{\infty} \psi\left(\frac{x}{k}\right) - 2 \sum_{k=1}^{\infty} \psi\left(\frac{x}{2k}\right). \quad \square$$

Proof of (4.22).

Note that $\theta(x)$ is increasing. Hence, from (4.20),

$$\psi(x) - 2\psi(\sqrt{x}) \leq \theta(x).$$

Also, from (4.18),

$$\psi(x) \geq \theta(x). \quad \square$$

Proof of (4.23).

This follows immediately from (4.21). □

Lemma 4.13. *Let x be a real number. Then*

$$\ln[x!] - 2 \ln[x/2]! > \frac{2}{3}x \quad \text{if } x > 750, \tag{4.24}$$

$$\ln[x!] - 2 \ln[x/2]! < \frac{3}{4}x \quad \text{if } x > 0, \tag{4.25}$$

$$\psi(x) - \psi\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right) > \frac{2}{3}x \quad \text{if } x > 750, \tag{4.26}$$

and

$$\psi(x) - \psi\left(\frac{x}{2}\right) < \frac{3}{4}x \quad \text{if } x > 0. \tag{4.27}$$

Proof of (4.24).

For real number z , the Gamma function $\Gamma(z)$ is given by

$$\Gamma(z) = \lim_{n \rightarrow \infty} \frac{(n-1)!n^z}{z(z+1)\cdots(z+n-1)}.$$

The function $\Gamma(x)$ satisfies the well-known Stirling's formula [14, p. 253]

$$\ln \Gamma(x) = \ln \sqrt{2\pi} + \left(x - \frac{1}{2}\right) \ln x - x + \frac{\vartheta_x}{12x}, \quad 0 < \vartheta_x < 1. \tag{4.28}$$

Since $\Gamma(x)$ is increasing when $x > 3$, we conclude that

$$\ln[x!] - 2 \ln[x/2]! \geq \ln \Gamma(x) - 2 \ln \Gamma\left(\frac{1}{2}x + 1\right).$$

To prove (4.24), it suffices to show that for $x > 750$,

$$\ln \Gamma(x) - 2 \ln \Gamma\left(\frac{1}{2}x + 1\right) > \frac{2x}{3}. \tag{4.29}$$

By (4.28), we deduce that

$$\begin{aligned} \ln \Gamma(x) - 2 \ln \Gamma\left(\frac{1}{2}x + 1\right) &= \ln \sqrt{2\pi} + \left(x - \frac{1}{2}\right) \ln x - x + \frac{\vartheta_1}{12x} - 2 \ln \sqrt{2\pi} \\ &\quad - 2 \left(\frac{x}{2} + \frac{1}{2}\right) \ln\left(\frac{x}{2} + 1\right) + 2 \left(\frac{x}{2} + 1\right) - \frac{\vartheta_2}{3x + 6}, \end{aligned}$$

where both ϑ_1, ϑ_2 belong to the interval $(0, 1)$. Simplifying the above inequality, we deduce that

$$\ln \Gamma(x) - 2 \ln \Gamma\left(\frac{1}{2}x + 1\right) > x \ln\left(\frac{2x}{x+2}\right) - 2 \ln x.$$

For $x > 750$,

$$x \ln\left(\frac{2x}{x+2}\right) - 2 \ln x > \frac{2x}{3}.$$

This completes the proof of (4.29). \square

Proof of (4.25).

The proof is similar to that for (4.24). We use the inequality

$$\ln[x]! - 2 \ln[x/2]! \leq \ln \Gamma(x + 1) - 2 \ln \Gamma\left(\frac{1}{2}x + \frac{1}{2}\right)$$

and Stirling's formula to conclude that (see Problem 8 of Exercise 4.7)

$$\ln[x]! - 2 \ln[x/2]! \leq \frac{3}{4}x$$

for all $x > 0$. \square

Proof of (4.26) and (4.27).

These two inequalities follow immediately from (4.23)–(4.25). \square

Lemma 4.14. *For each positive real number x , we have*

$$\psi(x) < \frac{3}{2}x \quad \text{if } x > 0 \tag{4.30}$$

$$\begin{aligned} \psi(x) - \psi\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right) &\leq \theta(x) + 2\psi(\sqrt{x}) - \theta\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right) \\ &< \theta(x) - \theta\left(\frac{x}{2}\right) + \frac{x}{2} + 3\sqrt{x}. \end{aligned} \tag{4.31}$$

Proof of (4.30).

To prove (4.30), we use (4.27) repeatedly, with x replaced by $x/2$, $x/4, \dots$ and add up the results. We find that

$$\psi(x) \leq \frac{3}{4}x \left(1 + \frac{1}{2} + \dots \right) < \frac{3}{2}x.$$

Proof of (4.31).

From (4.22), we find that

$$\psi(x) - 2\psi(\sqrt{x}) \leq \theta(x).$$

Hence

$$\psi(x) \leq \theta(x) + 2\psi(\sqrt{x}).$$

Next, from (4.22),

$$\theta(x/2) \leq \psi(x/2).$$

Using the above inequalities, we deduce that

$$\psi(x) - \psi(x/2) + \psi(x/3) \leq \theta(x) + 2\psi(\sqrt{x}) - \theta(x/2) + \psi(x/3).$$

For the second inequality, we use (4.30) to deduce that

$$2\psi(\sqrt{x}) + \psi(x/3) \leq 3\sqrt{x} + x/2.$$

We are now ready to prove Bertrand's Postulate. By (4.26),

$$\psi(x) - \psi(x/2) + \psi(x/3) \geq \frac{2}{3}x$$

for $x > 750$. Hence we deduce from (4.31) that

$$\theta(x) - \theta(x/2) \geq 2x/3 - x/2 - 3\sqrt{x},$$

and there is a prime between x and $2x$ for $x > 750$.

We are now left with verifying that Bertrand's Postulate is true for $x \leq 750$. This is straightforward and we leave it as an exercise.

Example 4.1. Suppose S is a set of consecutive integers which contains a prime. Show that there exists an integer in S that is relatively prime to all the other integers in S . Conversely, show that if for every finite set of consecutive integers with at least a prime contains a number that is relatively prime to the others, then Bertrand's Postulate is true.

Proof. Let $S := \{n, n+1, \dots, n+k\}$ be any set of consecutive integers with at least a prime. Suppose p is the largest prime in the set S . If $2p \leq n+k$, then by Bertrand's Postulate, there exists another prime q such that $p < q < 2p$ and this contradicts our assumption that p is the largest prime in the set. Hence $2p > n+k$ and p is relatively prime to all the other integers in S .

Conversely, let $n > 1$ be given and consider the set $T = \{2, 3, \dots, 2n\}$. Obviously there is at least a prime in T and hence by hypothesis, there exists an integer q such that q is relatively prime to the other numbers in T . This implies that $2q > 2n$ (for, if $2q < 2n$ then q will not be prime to $2q$.) Hence $n < q < 2n$. To prove that q is prime, we suppose $a|q$ and $a \neq 1$. Then $a \in T$ and so, a is not relatively prime to q . This contradicts the choice of q . Hence q must be prime. This means that there is a prime between n and $2n$ if $n > 1$, and this is Bertrand's Postulate. \square

4.7 Exercises

1. Prove that the following two relations are equivalent.

$$(a) \pi(x) = \frac{x}{\ln x} + O\left(\frac{x}{\ln^2 x}\right)$$

$$(b) \theta(x) = x + O\left(\frac{x}{\ln x}\right).$$

2. Show that if p, q are primes, then

$$\sum_{pq \leq x} \frac{1}{pq} = (\ln \ln x)^2 + O(\ln \ln x).$$

3. Let $\omega(n)$ be defined as in Definition 3.6. Show that

$$\sum_{n \leq x} \omega(n) = x \ln \ln x + O(x).$$

4. Show that for $x \geq 2$,

$$\sum_{n \leq x} \psi\left(\frac{x}{n}\right) = x \ln x + O(x).$$

5. Show that for $x \geq 1$,

$$\sum_{n \leq x} \frac{\theta(n)}{n^2} = \ln x + O(1).$$

6. Show that

$$\sum_{p \leq x} \frac{\ln p}{p-1} = \ln x + O(1).$$

7. Let x be a positive real number. Verify using Stirling's formula that

$$\ln[x]! - 2 \ln[x/2]! < \frac{3}{4}x \quad \text{if } x > 0.$$

8. Follow Erdős' proof of Bertrand's Postulate by completing the following exercises.

(a) Let $r(p)$ be such that

$$p^{r(p)} \leq 2n < p^{r(p)+1}.$$

Show that

$$\binom{2n}{n} \mid \prod_{p \leq 2n} p^{r(p)}.$$

(b) Show that if $p > 2$ and $\frac{2n}{3} < p \leq n$ then

$$p \nmid \binom{2n}{n}.$$

(c) Show that

$$\prod_{p \leq n} p < 4^n.$$

(d) Use the above results to deduce Bertrand's Postulate.

9. Show that if n is any positive integer greater than 1, then $n!$ is never a perfect square.
10. Let n be a positive integer. Assume that when $m > 20$, there exists a prime between $m/2$ and $m - 6$. Show that n can be expressed as a sum of distinct primes when $n > 6$. (Challenge: Try to prove the above result without making the first assumption.)

11. Use Bertrand's Postulate to show that for every positive integer n , the number

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$$

can never be an integer.

12. Use induction and Bertrand's Postulate to show that if p_n denote the n -th prime, then for $n > 3$,

$$p_n < p_1 + p_2 + \cdots + p_{n-1}.$$