MA4263 Analytic Number Theory

Thang Pang Ern

Reference books:

(1). T. M. Apostol. *Introduction to Analytic Number Theory*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1976. ISBN: 9780387901633.

(2). H. Davenport. *Multiplicative Number Theory*. 3rd Edition. Revised by H. L. Montgomery. Graduate Texts in Mathematics, Vol. 74. Springer-Verlag, New York, 2000. ISBN: 9780387950976.

(3). G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers*. 6th Edition. Revised by D. R. Heath-Brown and J. H. Silverman. Oxford University Press, Oxford, 2008. ISBN: 9780199219865.

(4). H. H. Chan. *Analytic Number Theory for Undergraduates*. Monographs in Number Theory, No. 3. World Scientific Publishing Company, 2009. ISBN: 9789814271363.

(5). J.-M. De Koninck and F. Luca. *Analytic Number Theory: Exploring the Anatomy of Integers*. American Mathematical Society, 2012. ISBN: 9780821869470.

Contents

1 Fundamental Theorem of Arithmetic **3**
1.1 The Division Algorithm . . . 3
1.2 Greatest Common Divisors and Least Common Multiples . . . 4

2 Arithmetic Functions **8**
2.1 Arithmetic Functions and Multiplicative Functions . . . 8
2.2 Perfect Numbers and the Sum of Divisors Function $\sigma(n)$ . . . 10
2.3 The Möbius Function $\mu(n)$ . . . 11
2.4 The Euler Totient Function $\varphi(n)$ . . . 14
2.5 Dirichlet Products . . . 14
2.6 The Averages of Arithmetic Functions . . . 19
2.7 The Euler-Maclaurin Formula . . . 20
2.8 Dirichlet's Hyperbola Method . . . 24

3 The Prime Number Theorem **30**
3.1 Chebyshev's Functions $\theta(x)$ and $\psi(x)$ . . . 30
3.2 Merten's Estimates . . . 41
3.3 The Riemann Zeta Function . . . 42
3.4 Completing the Proof of the Prime Number Theorem . . . 44

4 Dirichlet Series **59**
4.1 Introduction to Dirichlet Series . . . 59
4.2 Multiplication of Dirichlet Series . . . 59
4.3 Conditional Convergence of Dirichlet Series . . . 59
4.4 Landau's Theorem . . . 59

5 Dirichlet's Theorem **60**
5.1 Dirichlet Characters . . . 60
5.2 Proof of Dirichlet's Theorem in Arithmetic Progressions . . . 60

6 Sieve Methods **69**
6.1 The Sieve of Eratosthenes . . . 69
6.2 The Large Sieve . . . 69
6.3 Brun's Sieve and Twin Primes . . . 69

7 Partitions **76**
7.1 The Rogers-Ramanujan Identities . . . 76

Chapter 1

Fundamental Theorem of Arithmetic

1.1 The Division Algorithm

Let $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$ denote the set of integers, and $\mathbb{N} = \{1, 2, \ldots\}$ denote the set of natural numbers or positive integers. The least integer axiom, or the well-ordering principle, states that there is a smallest integer in every non-empty subset of $\mathbb{N}$.

We recall the division algorithm (Theorem 1.1). One should have seen a proof of this result numerous times so we shall omit it.

Theorem 1.1 (division algorithm)**.** Let $a, b \in \mathbb{Z}$ such that $b > 0$. Then, there exist unique integers q and r such that

$$a = bq + r \quad \text{where } 0 \leq r < b.$$

Here, q is the quotient and r is the remainder.

When $r = 0$ in Theorem 1.1, we have $a = bq$ and we say that b divides a and we write $b \mid a$. When $r > 0$, we say that b does not divide a and we write $b \nmid a$. If $b \mid a$, we say that b is a *divisor* of a and that a is a multiple of b. From here, we say that a positive integer is a prime if it has exactly two divisors, namely 1 and itself.

We now state some elementary properties of divisibility.

Theorem 1.2. Let a, b, d, m, n be non-zero integers. Then, the following statements hold:

- **(i)** For every non-zero integer k, $k \mid k$
- **(ii)** **Transitivity of divisibility:** if $d \mid n$ and $n \mid m$, then $d \mid m$
- **(iii)** If $d \mid n$ and $d \mid m$, then $d \mid (an + bm)$
- **(iv)** If $d \mid n$, then $ad \mid an$
- **(v)** If $ad \mid an$ and $a \neq 0$, then $d \mid n$
- **(vi)** If $d \mid n$, then $|d| \leq |n|$
- **(vii)** **Antisymmetry of divisibility:** If $d \mid n$ and $n \mid d$, then $|d| = |n|$
- **(viii)** If $d \mid n$, then $\frac{n}{d} \mid n$

In Theorem 1.2, we see that if d is a divisor of n, then $\frac{n}{d}$ is also a divisor of n. If d is a divisor of n, we say that $\frac{n}{d}$ if the conjugate divisor of d.

Definition 1.1 (congruence modulo n)**.** We say that a is congruent to b modulo n if $n \mid (a-b)$. The notation is

$$a \equiv b \pmod{n}.$$

With the notation of congruence modulo n as in Definition 1.1, we conclude using the division algorithm (Theorem 1.1) that given positive integers a and b, there exists a unique r with $0 \leq r < b$ such that $a \equiv r \pmod{b}$. We then state some properties of congruences (Theorem 1.3).

Theorem 1.3. Let a, b, c, d, n be integers with $n > 0$. Then, the following hold:

(i) For all integers k, $k \equiv k \pmod{n}$

(ii) If $a \equiv b \pmod{n}$, then $b \equiv a \pmod{n}$

(iii) If $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$, then $a \equiv c \pmod{n}$

(iv) If $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$, then

$$a + c \equiv b + d \pmod{n} \quad \text{and} \quad ac \equiv bd \pmod{n}$$

1.2 Greatest Common Divisors and Least Common Multiples

Let $a, b \in \mathbb{Z}$ for which at leat one of them is non-zero. A common divisor of a and b is an integer c satisfying $c \mid a$ and $c \mid b$. From here, we define the greatest common divisor of two integers (Definition 1.2).

Definition 1.2 (greatest common divisor)**.** Let $a, b \in \mathbb{Z}$. A greatest common divisor of a and b is an integer d satisfying the following properties:

(i) d is a non-negative integer

(ii) d is a common divisor of a and b

(iii) If e is any common divisor of a and b, then $e \mid d$

The greatest common divisor of two integers a and b, of which one is non-zero, is unique. We write it as $\gcd(a, b)$ or simply (a, b). I prefer the former representation though. Note that if $b \neq 0$ and $a = 0$, then $\gcd(0, b) = |b|$. In Bézout's lemma (Lemma 1.1), we will show that the greatest common divisor of two integers exists. By our previous discussion, it suffices to only consider the case when both a and b are non-zero.

Lemma 1.1 (Bézout's lemma). Let a and b be non-zero integers. Then,

$$\text{there exist } m, n \in \mathbb{Z} \quad \text{such that} \quad am + bn = \gcd(a, b).$$

The conventional proof of Bézout's lemma *purely* involves the Euclidean algorithm as one would have seen in MA1100 Basic Discrete Mathematics. Here, we give a proof using Group Theory.

Proof. Recall from MA2202 Algebra 1 that if G is a cyclic group and $H \leq G$[†], then H is also cyclic. To see why, recall that every cyclic group can be generated by a single element. Suppose G is generated by g. Since $H \leq G$, then

$$H = \left\{ g^\ell : \ell \in T \right\} \quad \text{for some } T \subseteq \mathbb{Z}.$$

Let r be the smallest positive integer in T, where the existence of r is guaranteed by the well-ordering principle. We claim that H is generated by g^r. Suppose otherwise. Then, by the division algorithm,

$$\text{there exists } \ell \in T \quad \text{such that} \quad \ell = rq + s \text{ where } 0 < s < r.$$

Note that

$$g^{rq} \in H \text{ and } g^\ell \in H \quad \text{so} \quad g^s = g^{\ell - rq} \in H.$$

So, $s \in T$ and $0 < s < r$, contradicting the minimality of r. Hence, H is a cyclic group generated by g^r.

Now, let

$$G = (\mathbb{Z}, +) \quad \text{and} \quad H = \{am + bn : m, n \in \mathbb{Z}\}.$$

G is a cyclic group generated by 1, and $H \subseteq G$ since H consists of all linear combinations of $m, n \in \mathbb{Z}$. In fact, $H \leq G$ by using the subgroup criterion. That is to say, let $am + bn, am' + bn' \in H$. Then,

$$(am + bn) - (am' + bn') = a(m - m') + b(n - n') \in H.$$

Since $H \leq G$, then H must be generated by some positive integer, say d. Since $d \in H$, then

$$\text{there exist } \alpha, \beta \in \mathbb{Z} \quad \text{such that} \quad d = \alpha a + \beta b.$$

One can show that d is a common divisor of a and b. Next, let d' be another common divisor of a and b. Then,

$$\text{there exist } k_1, k_2 \in \mathbb{Z} \quad \text{such that} \quad a = k_1 d' \text{ and } b = k_2 d'.$$

[†]Recall that this means that H is a subgroup of G.

So, $d = d'(k_1\alpha + k_2\beta)$ which implies that $d' \mid d$. Since $d > 0$, we conclude that $d = \gcd(a, b)$ since d satisfies the conditions in Definition 1.2. □

Definition 1.3 (coprime). Let $a, b \in \mathbb{Z}$. We say that a and b are relatively prime or coprime if and only if $\gcd(a, b) = 1$.

Theorem 1.4. Let a and b be non-zero integers. Then, $\gcd(a, b) = 1$ if and only if there exist $x, y \in \mathbb{Z}$ such that $ax + by = 1$.

We now state some basic properties of the greatest common divisor of two integers.

Theorem 1.5. Let a, b, c be non-zero integers. Then, the following hold:

(i) $\gcd(a, b) = \gcd(b, a)$

(ii) $\gcd(a, \gcd(b, c)) = \gcd(\gcd(a, b), c)$

(iii) $(ac, bc) = |c| \gcd(a, b)$

At this juncture, we see that we have discussed the definition of the greatest common divisor of two integers a and b. We can extend this definition to the greatest common divisor of m integers. Suppose $d \in \mathbb{N}$ is the divisor of $a_1, \ldots, a_m$ satisfying the property that any common divisor of $a_1, \ldots, a_m$ divides d. The greatest common divisor of m integers is $\gcd(a_1, \ldots, a_m)$. For example, one can show what is known as the associativity and commutativity of the greatest common divisor as follows:

$$\gcd(a, b, c) = \gcd(a, \gcd(b, c)) = \gcd(\gcd(a, b), c) = \gcd(\gcd(a, c), b)$$

We now define the least common multiple of two integers a and b (Definition 1.4).

Definition 1.4 (least common multiple). The least common multiple of two integers a and b with $b \neq 0$ is defined as an integer m satisfying the following properties:

(i) m is a positive integer

(ii) $a \mid m$ and $b \mid m$

(iii) $a \mid \ell$ and $b \mid \ell$ implies $m \mid \ell$

The notation for the least common multiple of a and b is $\operatorname{lcm}(a, b)$ or $[a, b]$. Similarly, as mentioned before, I personally prefer the former representation. We finally introduce one important identity relating $\gcd(a, b)$ and $\operatorname{lcm}(a, b)$ (Theorem 1.6).

Theorem 1.6. Let $a, b \in \mathbb{N}$. Then,

$$\gcd(a, b) \operatorname{lcm}(a, b) = ab.$$

Recall the cancellation property for equality, which states that if $c \neq 0$, then $ca = cb$ implies $a = b$. This is not true for congruences. For example, we know that $15 \equiv 3 \pmod{12}$ since 15 and 3 differ by an integer multiple of 12. However, $5 \not\equiv 1 \pmod{12}$. Theorem 1.7 shows that the law of cancellation for congruences holds if we impose a condition on c.

Theorem 1.7. Let $a, b, c, n \in \mathbb{Z}$. If

$$ca \equiv cb \pmod{n} \text{ and } \gcd(c, n) = 1 \quad \text{then} \quad a \equiv b \pmod{n}.$$

Theorem 1.7 can be used to prove Euclid's lemma (Lemma 1.2).

Lemma 1.2 (Euclid's lemma)**.** Let $a, b \in \mathbb{Z}$ and p be a prime. If

$$p \mid ab \quad \text{then} \quad p \mid a \text{ or } p \mid b.$$

Corollary 1.1. Let $a_1, \ldots, a_m \in \mathbb{Z}$ and p be a prime. If

$$p \mid a_1 \ldots a_m \quad \text{then} \quad p \mid a_k \text{ for some } 1 \leq k \leq m.$$

Theorem 1.8 (Fundamental Theorem of Arithmetic)**.** Every positive integer $n > 1$ can be expressed as a product of primes, and this representation is unique apart from the order in which the factors occur.

One can use the fundamental theorem of arithmetic to prove Theorem 1.6.

Chapter 2

Arithmetic Functions

2.1 Arithmetic Functions and Multiplicative Functions

We now give an introduction to arithmetic functions. Simply said, any function $f : \mathbb{N} \to \mathbb{C}$ is an arithmetic function. We give some examples of arithmetic functions (Example 2.1).

Example 2.1 (examples of arithmetic functions)**.** For example, the constant unit function $u(n) = 1$ for all $n \in \mathbb{N}$ is an arithmetic function, and the identity function $N(n) = n$ for all $n \in \mathbb{N}$ is also an arithmetic function.

Let $d(n)$ denote the sum of divisors of n and $\sigma(n)$ denote the sum of divisors of n (see graph in Figures 1 and 2). Then, d and σ are both arithmetic functions. In fact, we will explore these functions in greater detail in Chapter 2.2.

Let $f(n)$ be an arithmetic function. It turns out that we can construct a new arithmetic function $g(n)$ by letting

$$g(n) = \sum_{d|n} f(d). \tag{2.1}$$

Clearly, $g : \mathbb{N} \to \mathbb{C}$ as well. In (2.1), g is constructed by summing $f(d)$ over all divisors of n. In fact, with this notation, we can write

$$d(n) = \sum_{\ell|n} 1 = \sum_{\ell|n} u(\ell) \quad \text{and} \quad \sigma(n) = \sum_{\ell|n} \ell.$$

In other words, $d(n)$ is constructed from $u(n)$ and $\sigma(n)$ is constructed from $N(n)$ via the summation over divisors of n. Well, such an approach is analogous to constructing a *new* continuous function via the integration of continuous functions in MA2002 Calculus.

Also, one should observe that there is a one-to-one correspondence between the divisors d of n. If $d \mid n$, then $n = d\left(\frac{n}{d}\right)$, and this implies that $\frac{n}{d} \mid n$. Conversely, $\frac{n}{d}$ divides n, then $d \mid n$. Summing over d is the same as summing over $\frac{n}{d}$ since we have seen that there is a one-to-one correspondence between these divisors. As such,

$$\sum_{d|n} f(d) = \sum_{\frac{n}{d}|n} f(d) = \sum_{d'|n} f\left(\frac{n}{d'}\right) = \sum_{d|n} f\left(\frac{n}{d}\right)$$

where we have let d' play the role of $\frac{n}{d}$. Based on our discussion, we have for example the identity

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

Proof. By definition of the sum of divisors function $\sigma(n)$,

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

Dividing both sides by n yields the desired result. □

Definition 2.1 (multiplicative function)**.** An arithmetic function $f : \mathbb{N} \to \mathbb{C}$ is said to be multiplicative if $f(1) = 1$ and for every $m, n \in \mathbb{N}$ such that $\gcd(m, n) = 1$, we have

$$f(mn) = f(m) f(n).$$

Definition 2.2 (completely multiplicative function)**.** An arithmetic function $f : \mathbb{N} \to \mathbb{C}$ is said to be completely multiplicative if $f(1) = 1$ and for every $m, n \in \mathbb{N}$, we have

$$f(mn) = f(m) f(n).$$

We see that every completely multiplicative function is also multiplicative, but we can *drop* the coprime condition.

Suppose $n > 1$ is an integer written in the form

$$n = \prod_{i=1}^{k} p_i^{\alpha_i}.$$

Say f is multiplicative. Then, $f(n)$ can be written as

$$f\left(\prod_{i=1}^{k} p_i^{\alpha_i}\right) = \prod_{i=1}^{k} f(p_i^{\alpha_i})$$

where we used the fact that prime numbers are coprime. This shows that if f is multiplicative, then its value at any positive integer n is determined by its values at prime powers. On the other hand, if f is completely multiplicative, then $f(n)$ can be written as

$$f\left(\prod_{i=1}^{k} p_i^{\alpha_i}\right) = \prod_{i=1}^{k} f(p_i^{\alpha_i}) = \prod_{i=1}^{k} [f(p_i)]^{\alpha_i}.$$

Now, the values of $f(n)$ are completely determined by the values of $f(p)$ for primes p.

Recall our construction of g from f in (2.1). We now give a simple yet useful result for multiplicative functions (Theorem 2.1).

Theorem 2.1. Let f be a multiplicative function. Then,

$$g(n) = \sum_{\ell \mid n} f(\ell) \quad \text{is also multiplicative.}$$

Proof. We have $g(1) = 1$. Suppose $m, n \in \mathbb{N}$ such that they are coprime. Suppose $\ell \mid mn$. Then, there exist $\ell_1, \ell_2 \in \mathbb{N}$ such that

$$\ell = \ell_1 \ell_2 \quad \text{with} \quad \ell_1 \mid m \text{ and } \ell_2 \mid n.$$

Here, we used the fact that m and n are coprime. In particular, choosing $\ell_1 = \gcd(\ell, m)$ and $\ell_2 = \gcd(\ell, n)$ works. As such,

$$g(mn) = \sum_{\ell \mid mn} f(\ell) = \sum_{\ell_1 \mid m} \sum_{\ell_2 \mid n} f(\ell_1) f(\ell_2) = g(m) g(n)$$

which completes the proof. □

2.2 Perfect Numbers and the Sum of Divisors Function $\sigma(n)$

An integer n is said to be perfect if the sum of its divisors less than n is equal to n. The first two perfect numbers are 6 and 28. Note that by the definition of $\sigma(n)$, we observe that a positive integer n is perfect if and only if $\sigma(n) = 2n$.

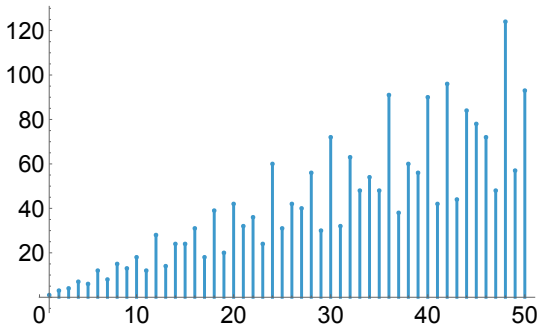

Figure 1: Graph of $\sigma(n)$ up to $n = 50$

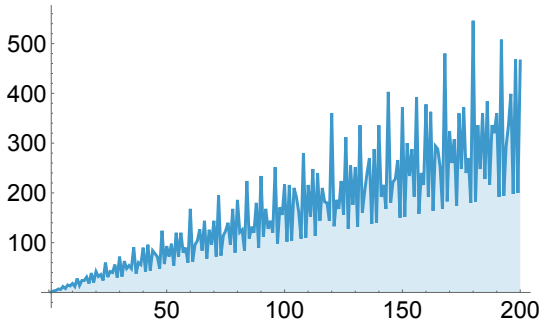

Figure 2: Graph of $\sigma(n)$ up to $n = 200$

We give a characterisation of even perfect numbers.

Theorem 2.2. Let $n \in \mathbb{N}$. An even integer N is perfect if and only if

$$\text{there exists a prime of the form } 2^k - 1 \quad \text{such that} \quad N = 2^{k-1}\left(2^k - 1\right).$$

Proof. □

2.3 The Möbius Function $\mu(n)$

We now introduce one of the most important arithmetic functions, the Möbius function $\mu(n)$ (Definition 2.3). The graph of $\mu(n)$ is shown in Figures 3 and 4.

Definition 2.3 (Möbius function)**.** Suppose the prime factorisation of n is

$$n = p_1^{\alpha_1} \dots p_k^{\alpha_k}.$$

Then, the Möbius function $\mu(n)$ is defined to be

$$\mu(n) = \begin{cases} (-1)^k & \text{if } \alpha_1 = \dots = \alpha_k = 1; \\ 0 & \text{otherwise.} \end{cases}$$

We also fix $\mu(1) = 1$.

Note that $\mu(n) = 0$ if and only if n has a square factor greater than 1. In other words, $\mu(n) = 0$ if and only if n is not squarefree.

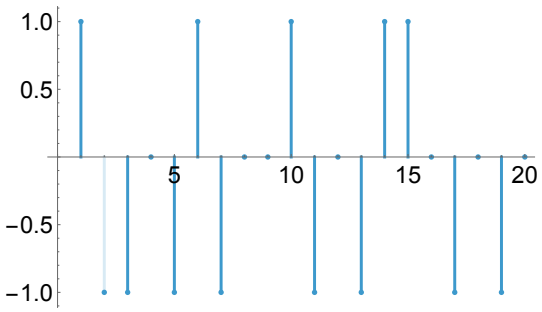

Figure 3: Graph of $\mu(n)$ up to $n = 20$

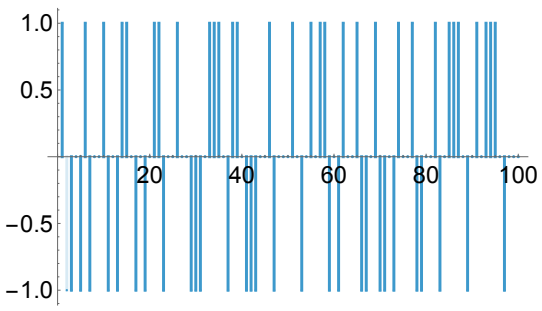

Figure 4: Graph of $\mu(n)$ up to $n = 100$

Recall we mentioned multiplicative functions in Definition 2.1. We now give a definition for additive functions.

Definition 2.4. An arithmetic function $f : \mathbb{N} \to \mathbb{C}$ is additive if for any coprime $m, n \in \mathbb{N}$, we have

$$f(mn) = f(m) + f(n).$$

Definition 2.5. An arithmetic function $f : \mathbb{N} \to \mathbb{C}$ is completely additive if for any $m, n \in \mathbb{N}$, we have

$$f(mn) = f(m) + f(n).$$

Definition 2.6 (prime omega functions)**.** Let $n \in \mathbb{N}$. The prime omega functions $\omega(n)$ and $\Omega(n)$ count the number of prime factors of n. The number of distinct prime factors is assigned to $\omega(n)$, while $\Omega(n)$ counts the total number of prime factors with multiplicity. We set $\omega(1) = 0$ and $\Omega(1) = 0$.

The graphs of $\omega(n)$ and $\Omega(n)$ are shown in Figures 5 and 6 respectively. From Definition 2.6, we see that if the prime factorisation of n is of the form $n = p_1^{\alpha_1} \dots p_k^{\alpha_k}$ for distinct primes p_i where $1 \leq i \leq k$, then

$$\omega(n) = k \quad \text{and} \quad \Omega(n) = \alpha_1 + \ldots + \alpha_k.$$

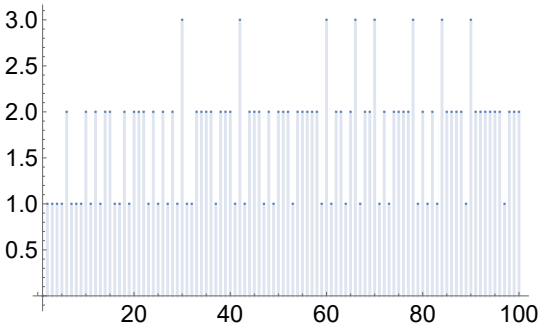

Figure 5: Graph of $\omega(n)$ up to $n = 100$

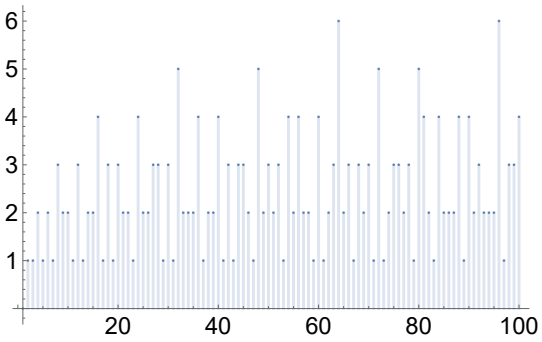

Figure 6: Graph of $\Omega(n)$ up to $n = 100$

Proposition 2.1. The arithmetic function $\omega(n)$ is additive.

Proof. Let $m, n \in \mathbb{N}$ be coprime. Suppose they have prime factorisations

$$m = \prod_{i=1}^{k} p_i^{\alpha_i} \quad \text{and} \quad n = \prod_{j=1}^{t} q_j^{\beta_j}.$$

Then, $\omega(mn) = k + t = \omega(m) + \omega(n)$. □

Proposition 2.2. The arithmetic function $\Omega(n)$ is completely additive.

Proof. Let $m, n \in \mathbb{N}$. Suppose they have prime factorisations

$$m = \prod_{i=1}^{k} p_i^{\alpha_i} \quad \text{and} \quad n = \prod_{j=1}^{t} q_j^{\beta_j}.$$

Then, $\Omega(mn) = \alpha_1 + \ldots + \alpha_k = \beta_1 + \ldots + \beta_t = \Omega(m) + \Omega(n)$. □

Let $n = p_1^{\alpha_1} \ldots p_k^{\alpha_k}$, where $\alpha_j = 1$ for all $1 \leq j \leq k$. Then, n is said to be squarefree. In this case, n has k distinct prime divisors so $\omega(n) = k$. Hence, for any $n = p_1^{\alpha_1} \ldots p_k^{\alpha_k}$,

$$\mu(n) = \begin{cases} (-1)^{\omega(n)} & \text{if } n \text{ is squarefree;} \\ 0 & \text{otherwise.} \end{cases} \tag{2.2}$$

Proposition 2.3. The arithmetic function $\mu(n)$ is multiplicative.

Proof. We proceed with casework. If neither m or n is squarefree, then $\mu(m)\mu(n) = 0$. As mn is not squarefree as well, then $\mu(mn) = 0$. Hence, $\mu(mn) = \mu(m)\mu(n)$. Next, if both m and n are coprime and squarefree, by (2.2) and the fact that ω is additive (Proposition 2.1),

$$\mu(mn) = (-1)^{\omega(mn)} = (-1)^{\omega(m)+\omega(n)} = \mu(m)\mu(n),$$

which shows that μ is multiplicative. □

Theorem 2.3. Let $n \in \mathbb{N}$ and $[x]$ denote the integer part of x (this can also be regarded as the floor of x). Then,

$$\sum_{\ell | n} \mu(\ell) = I(n),$$

where

$$I(n) = \left[\frac{1}{n}\right] = \begin{cases} 1 & \text{if } n = 1; \\ 0 & \text{if } n > 1 \end{cases} \quad \text{is the Kronecker delta symbol.}$$

Proof. From Theorem 2.1, we know that

$$g(n) = \sum_{\ell | n} \mu(\ell) \quad \text{is a multiplicative function.}$$

So, $g(1) = 1$ and

$$g\left(\prod_p p^{\alpha_p}\right) = \prod_p g\left(p^{\alpha_p}\right).$$

Since

$$g\left(p^{\alpha_p}\right) = \mu(1) + \mu(p) + 0 + \ldots + 0 = 1 - 1 = 0,$$

if $n \neq 1$, then $g(n) = 0$. The result follows. □

2.4 The Euler Totient Function $\varphi(n)$

We now introduce the Euler totient function (Definition 2.7), denoted by either $\phi(n)$ or $\varphi(n)$. See Figures 7 and 8 for a graph of the Euler totient function.

Definition 2.7 (Euler totient function). The Euler totient function $\varphi(n)$ is defined to be the number of positive integers not exceeding n which are relatively prime to n.

We can express $\varphi(n)$ as the following sum:

$$\varphi(n) = \sum_{\substack{1 \leq k \leq n \\ \gcd(k,n)=1}} 1. \tag{2.3}$$

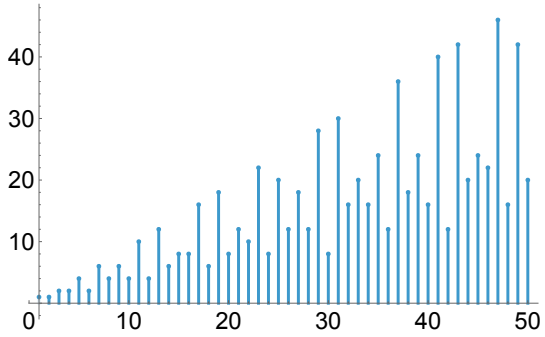

Figure 7: Graph of $\varphi(n)$ up to $n = 50$

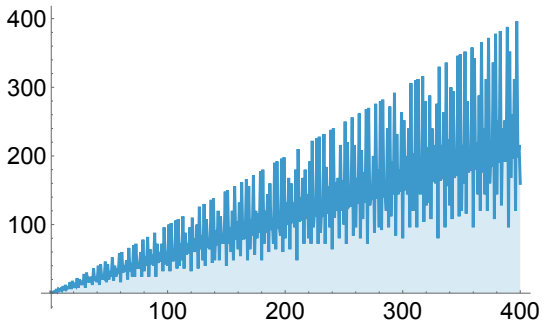

Figure 8: Graph of $\varphi(n)$ up to $n = 400$

2.5 Dirichlet Products

Definition 2.8. Let f and g be two arithmetic functions. We define the Dirichlet product of f and g, denoted by $f * g$, to be

$$(f * g)(n) = \sum_{\ell | n} f(\ell)\, g\left(\frac{n}{\ell}\right).$$

We will often use $f * g$ to represent the Dirichlet product. Using this notation, we can express identities such as Theorem 2.8 (which we will mention in due course) as

$$\varphi = \mu * N.$$

We now show that the set of multiplicative functions, which we shall denote by $\mathcal{M}$, equipped with the operation $*$, forms an Abelian group. There are several things we need to prove, which are namely $*$ being a binary operation on $\mathcal{M}$ (Theorem 2.4), $*$ is associative (Theorem 2.5), exhibit the identity function for the Dirichlet convolution, which is the Kronecker delta symbol I (Theorem 2.6), the existence of an inverse under $*$ (Theorem 2.9), and Corollary 2.1.

Theorem 2.4 ($*$ is a binary operation)**.** Let f and g be multiplicative functions. Then, $f * g$ is also multiplicative.

Proof. Let $h = f * g$. Note that $h(1) = f(1)\, g(1) = 1$. Next, consider the expression

$$h(mn) = \sum_{c \mid mn} f(c)\, g\left(\frac{mn}{c}\right).$$

Since f and g are multiplicative, given that $\gcd(m, n) = 1$, there exists $c = ab$ such that $a \mid m$ and $b \mid n$. Hence,

$$\begin{aligned} h(mn) &= \sum_{a \mid m} \sum_{b \mid n} f(ab)\, g\left(\frac{m}{a} \cdot \frac{n}{b}\right) \\ &= \sum_{a \mid m} \sum_{b \mid n} f(a)\, f(b)\, g\left(\frac{m}{a}\right) g\left(\frac{n}{b}\right) \quad \text{since } f, g \text{ are multiplicative} \end{aligned}$$

Here, we used the fact that $\gcd\left(\frac{m}{a}, \frac{n}{b}\right) = 1$. This shows that $h(mn) = h(m)\, h(n)$, implying that h is a multiplicative function. □

Theorem 2.5 (commutativity and associativity of $*$)**.** The Dirichlet product is commutative and associative. That is to say, let f, g, h be multiplicative functions. Then,

$$f * g = g * f \quad \text{and} \quad (f * g) * h = f * (g * h).$$

Proof. Showing that $*$ is commutative is easy — the Dirichlet product of f and g is given by

$$(f * g)(n) = \sum_{\ell \mid n} f(\ell)\, g\left(\frac{n}{\ell}\right).$$

Let $d_1 = \frac{n}{d}$ be the conjugate divisor of d. Since d runs through all divisors of n, then so does d_1. Hence,

$$(f * g)(n) = \sum_{d_1 | n} f\left(\frac{n}{d_1}\right) g(d_1) = (g * f)(n).$$

We then prove that $*$ is associative. Let $A = g * h$. Then,

$$\begin{aligned}(f * (g * h))(n) &= (f * A)(n) \\ &= \sum_{a|n} f(a) A\left(\frac{n}{a}\right) \\ &= \sum_{a \cdot d = n} f(a) \sum_{b \cdot c = d} g(b) h(c) \\ &= \sum_{a \cdot b \cdot c = n} f(a) g(b) h(c)\end{aligned}$$

Similarly, one can show that

$$((f * g) * h)(n) = \sum_{a \cdot b \cdot c = n} f(a) g(b) h(c)$$

so associativity of $*$ holds. □

Theorem 2.6 (identity element for $*$)**.** The function I is the identity function for $*$. That is, for any multiplicative function f, we have

$$I * f = f * I = f.$$

Note that Theorem 2.6 holds for any arithmetic function f, so f does not need to be multiplicative.

Proof. By the definition of I, we have

$$(I * f)(n) = \sum_{\ell | n} I(\ell) f\left(\frac{n}{\ell}\right) = f(n).$$

By the commutativity of $*$ (Theorem 2.5), we conclude that $f * I = f$. □

Theorem 2.7 (Möbius inversion formula)**.** If $f = g * u$, where $u(n) = 1$ is the constant unit function, then $g = f * \mu$, where μ denotes the Möbius function (Definition 2.3). The converse also holds.

Proof. Suppose $f = g * u$. Then,

$$\begin{aligned} f * \mu &= (g * u) * \mu \\ &= g * (u * \mu) \quad \text{by associativity of } * \\ &= g * I \end{aligned}$$

which is equal to g by Theorem 2.6. Note that the converse holds using a similar proof. Here, we used the fact that $u * \mu = I$. This can be rewritten as

$$(u * \mu)(n) = \sum_{\ell | n} \mu(\ell)\, u\left(\frac{n}{\ell}\right) = \sum_{\ell | n} \mu(\ell) = I(n)$$

where we used Theorem 2.3 in the last equality. □

Theorem 2.8. Let $n \in \mathbb{N}$. Then,

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

Proof. One should recall the classic identity

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

from MA2202 Algebra 1 or MA3265 Number Theory. To see why this holds, for each $1 \leq k \leq n$, let $d = \gcd(k, n)$. Then, $k = d\ell$ with $\gcd\left(\ell, \frac{n}{d}\right) = 1$. For each fixed $d \mid n$, there are exactly $\varphi\left(\frac{n}{d}\right)$ such ℓ. Summing over all $d \mid n$ yields

$$n = \sum_{d | n} \varphi\left(\frac{n}{d}\right) = \sum_{d | n} \varphi(d).$$

We can interpret the desired result (Theorem 2.8) as a Dirichlet convolution. Note that $u * \varphi = N$. By the Möbius inversion formula (Theorem 2.7), $\varphi = \mu * N$, so for each $n \in \mathbb{N}$, we have

$$\varphi(n) = \sum_{d | n} \mu(d)\, N\left(\frac{n}{d}\right) = \sum_{d | n} \mu(d) \cdot \frac{n}{d} = n \sum_{d | n} \frac{\mu(d)}{d}.$$

Dividing both sides by n yields the desired result. □

Theorem 2.9 (existence of inverse under $*$)**.** Let f be a multiplicative function. Then, there exists a unique function g such that $f * g = I$.

One can use induction to prove Theorem 2.9 but we omit the details. Also, Theorem 2.9 holds for any arithmetic function f with $f(1) \neq 0$, and not just for multiplicative functions. In fact, the function g in Theorem 2.9 has a name, and it is known as the Dirichlet inverse of f (Definition 2.9).

Definition 2.9 (Dirichlet inverse). Let f be an arithmetic function such that $f(1) \neq 0$. The unique function g satisfying the convolution $f * g = I$ is called the Dirichlet inverse of f, and it is denoted by f^{-1}.

From Definition 2.9, we see that

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

Example 2.2. Recall Theorem 2.3, which states that

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

As $I = u * \mu$, then the Dirichlet inverse of μ is u.

As mentioned, the proof of Theorem 2.9 uses a constructive proof of f^{-1} by induction. However, it is not clear that the Dirichlet inverse of a multiplicative function f is also multiplicative. To complete the proof that $(M, *)$ forms an Abelian group, it suffices to prove Theorem 2.10 then apply it to Corollary 2.1. The latter states that if f is multiplicative, then f^{-1} is multiplicative. One can prove it easily using contradiction so we omit the details.

Theorem 2.10. If both g and $f * g$ are multiplicative, then so is f.

Corollary 2.1. If g is multiplicative, then the Dirichlet inverse is also multiplicative.

Proof. Since g and $g * g^{-1} = I$ are multiplicative, the result follows by Theorem 2.10. □

Example 2.3. Suppose f is a completely multiplicative function. We first claim that

$$\mu f * f = I.$$

To see why,

$$(\mu f * f)(n) = \sum_{\ell | n} \mu(\ell) f(\ell) f\left(\frac{n}{\ell}\right) = \sum_{\ell | n} \mu(\ell) f(n)$$

where the second equality used the fact that f is completely multiplicative. So, the expression becomes

$$f(n)\sum_{\ell|n}\mu(\ell) = f(n)\,I(n) \quad \text{by Theorem 2.3}$$

Since $f(1) = 1$, it follows that $\mu f * f = I$. From here, we infer that $f^{-1} = \mu f$.

Example 2.4. Note that

$$\sigma = N * u.$$

In other words,

$$\sum_{\ell|n} N(\ell)\, u\left(\frac{n}{\ell}\right) = \sum_{\ell|n} \ell = \sigma(n)$$

so

$$\sigma^{-1} = N^{-1} * u^{-1} = \mu N * \mu = \mu * \mu N.$$

Example 2.5. Recall from Theorem 2.8 that

$$\varphi(n) = \sum_{d|n} \mu(d) \cdot \frac{n}{d} \quad \text{so} \quad \mu * N = \varphi.$$

Hence, $\varphi^{-1} = u * \mu N$.

2.6 The Averages of Arithmetic Functions

Let x be a positive real number. Then, of interest from this section onwards is the sum

$$\sum_{n\le x} f(n) = f(1) + f(2) + \ldots + f([x]).$$

The mean of the function f from 1 to x is defined by

$$\overline{f} = \frac{1}{x}\sum_{n\le x} f(n).$$

The purpose of studying this average is because in general, $\overline{f}$ behaves more regularly than $f([x])$ especially when x is large. For example, when f is the characteristic function for the prime numbers, namely $f(n) = 1$ if n is prime and 0 otherwise, the function

$$\sum_{n\le x} f(n)$$

is usually denoted by the prime counting function $\pi(x)$. The prime number theorem states that

$$\overline{f}(x) = \frac{\pi(x)}{x} \text{ behaves like } \frac{1}{\log x}.$$

On the other hand, we cannot predict the value of $f(n)$ for each $n = [x]$ since we do not precisely know the location of primes in $\mathbb{Z}$. We will study this beautiful topic in Chapter 3.

We first recall the big-O notation and the notion of asymptotic. Let $a \in \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ such that $g(x) > 0$ whenever $x \geq a$. We write

$$f(x) = \mathcal{O}(g(x))$$

to mean that the quotient $\frac{f(x)}{g(x)}$ is bounded for $x \geq a$. That is,

$$\text{there exists a constant } M > 0 \quad \text{such that} \quad |f(x)| \leq M\,|g(x)| \text{ for all } x \geq a.$$

Sometimes, we will also use the notation $f(x) \ll g(x)$ to represent $f(x) = \mathcal{O}(g(x))$. For example, for large x, we have $x^2 = \mathcal{O}(x^3)$ since $\frac{1}{x} = \frac{x^2}{x^3}$ is bounded by some positive constant for large x. Similarly, we have $x^n = \mathcal{O}(e^x)$ since exponential growth always dominates power growth. This is easy to see by the power series expansion of e^x, namely

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \geq \frac{x^n}{n!} \quad \text{where } x \geq 0.$$

As for asymptotic equivalence, if

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

we say that $f(x)$ is asymptotic to $g(x)$ as $x \to \infty$, and we write $f(x) \sim g(x)$. For example, based on our discussion of the prime number theorem, we have the asymptotic equivalence

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

2.7 The Euler-Maclaurin Formula

We will present the Euler-Maclaurin summation formula (Theorem 2.12) in this section. This formula provides a useful bridge between discrete sums and continuous integrals. Before that, we first introduce the Abel summation formula (Theorem 2.11). This is also known as the partial summation formula.

Theorem 2.11 (Abel summation formula). Let $a(n)$ be an arithmetic function and define

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

Let $0 \leq y < x$ be real numbers and $f : \mathbb{R} \to \mathbb{R}$ have continuous derivative on $[y, x]$. Then,

$$\sum_{y<n\leq x} a(n) f(n) = f(x) A(x) - f(y) A(y) - \int_y^x A(t) f'(t) \ dt. \tag{2.4}$$

Proof. Starting with the integral on the right side of (2.4), we have[†]

$$\int_y^x A(t) f'(t) \ dt = \int_y^x \sum_{n\leq t} a(n) f'(t) \ dt = \sum_{n\leq x} a(n) \int_{\max\{y,n\}}^x f'(t) \ dt$$

This is equal to

$$\sum_{n\leq x} a(n) \left[f(x) - f(\max\{y, n\}) \right].$$

Expanding the sum yields

$$f(x) A(x) - \sum_{n\leq y} a(n) f(y) - \sum_{y<n\leq x} a(n) f(n).$$

The result follows with some simple rearrangement. □

We then deduce the wonderful Euler-Maclaurin summation formula (Theorem 2.12) from the Abel summation formula (Theorem 2.11).

Theorem 2.12 (Euler-Maclaurin formula)**.** Let $0 < y < x$ and $f : \mathbb{R} \to \mathbb{R}$ have continuous derivative on $[y, x]$. Then,

$$\sum_{y<n\leq x} f(n) = \int_y^x f(t) + \int_y^x \{t\} f'(t) \ dt - f(x) \{x\} + f(y) \{y\}.$$

Proof. We apply the Abel summation formula (Theorem 2.11) with $a(n) = 1$, so $A(x) = [x]$. As such,

$$\sum_{y<n\leq x} f(n) = f(x) [x] - f(y) [y] - \int_y^x [t] f'(t) \ dt.$$

Recognising that $x = [x] + \{x\}$, the result follows. □

We give some elementary asymptotic formulae. First, we introduce the Riemann zeta function, denoted by $\zeta(s)$ (Definition 2.10).

[†]Try to understand why we can interchange the order of summation and integral and why the new limits of integration are as such.

Definition 2.10 (Riemann zeta function)**.** Let $s > 1$ be a real number. We define the Riemann zeta function as

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

We will give a definition for the complex analogue in Definition 3.5. Next, we define the Euler-Mascheroni constant γ.

Definition 2.11 (Euler-Mascheroni constant)**.** The Euler-Mascheroni constant is defined as

$$\gamma = \lim_{n\to\infty} (H_n - \ln n)$$

where H_n is the n^{th} harmonic number. γ carries a value of approximately 0.577.

Theorem 2.13. If $x \geq 1$, then

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

Theorem 2.13 is interesting — it states that the harmonic sum is given by an approximation consisting of the natural logarithm $\log x$, γ, and an error term of $\mathcal{O}\left(\frac{1}{x}\right)$, which means that the error is at most the order of $\frac{1}{x}$ for large x. This fundamental result in Analytic Number Theory gives a different proof of the divergence of the harmonic series!

Proof. Let $f(t) = \frac{1}{t}$ and $y = 1$. By the Euler-Maclaurin formula (Theorem 2.12),

$$\sum_{n\leq x} \frac{1}{n} = \int_1^x \frac{1}{t}\, dt - \int_1^x \frac{\{t\}}{t^2}\, dt + 1 - \frac{\{x\}}{x} = \log x + 1 + \mathcal{O}\left(\frac{1}{x}\right) - \int_1^x \frac{\{t\}}{t^2}\, dt.$$

For the integral in red, the idea is to break the interval into unit sub-intervals, i.e. $[n, n+1]$ for $n = 1, 2, \ldots$. This is a routine exercise from MA2002 Calculus so we omit the details. □

Example 2.6. Let $d(n)$ denote the number of divisors of n. Suppose there exist $d, \ell \in \mathbb{N}$ such that $n = dl$. Then,

$$\sum_{n\leq x} d(n) = \sum_{d\leq x} \sum_{\ell\leq \frac{x}{d}} 1 = \sum_{d\leq x} \left[\frac{x}{d}\right] = \sum_{d\leq x} \frac{x}{d} - \sum_{d\leq x} \left\{\frac{x}{d}\right\}.$$

Applying Theorem 2.13 to the sum in red yields

$$\sum_{d\leq x}\frac{x}{d}=x\log x+\gamma x+\mathcal{O}\left(1\right).$$

It follows that

$$\sum_{n\leq x}d\left(n\right)=x\log x+\mathcal{O}\left(x\right).$$

Example 2.7. The main result we wish to gear towards is

$$\left|\sum_{n\leq x}\frac{\mu\left(n\right)}{n}\right|\leq 1. \tag{2.5}$$

Note that

$$\sum_{n\leq x}\sum_{d|n}\mu\left(d\right)=1. \tag{2.6}$$

To see why, recall from Theorem 2.3 that

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

Summing over all $n\leq x$ indeed yields (2.6). Hence,

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

since the expression $\left[\frac{x}{d}\right]$ counts the number of $n\leq x$ such that $d\mid n$. Hence,

$$1=x\sum_{d\leq x}\frac{\mu\left(d\right)}{d}-\sum_{d\leq x}\mu\left(d\right)\left\{\frac{x}{d}\right\}.$$

So,

$$\begin{aligned}x\left|\sum_{d\leq x}\frac{\mu\left(d\right)}{d}\right|&\leq\left|1+\sum_{d\leq x}\mu\left(d\right)\left\{\frac{x}{d}\right\}\right|\\&\leq 1+\sum_{d\leq x}\left\{\frac{x}{d}\right\}\quad\text{by the triangle inequality and Definition 2.3}\end{aligned}$$

which is bounded above by $1+(x-1)=x$. This is easy to see because if $d=x$, then $\left\{\frac{x}{x}\right\}=0$ so it would not contribute anything to the sum. Dividing both sides by x yields the desired result.

Theorem 2.14. If $x \geq 1$, then

$$\sum_{n \leq x} \frac{1}{n^s} = \frac{x^{1-s}}{1-s} + C(s) + \mathcal{O}\left(\frac{1}{x^s}\right) \tag{2.7}$$

if $s > 0$ and $s \neq 1$, where

$$C(s) = \begin{cases} \zeta(s) & \text{if } s > 1; \\ \lim\limits_{x \to \infty} \left(\sum\limits_{n \leq x} \frac{1}{n^s} - \frac{x^{1-s}}{1-s} \right) & \text{if } 0 < s < 1. \end{cases}$$

Theorem 2.14 is the partial sum of what is known as the *Dirichlet series* (we will give a formal treatment in Chapter 4) of the Riemann zeta function. It states that if $s > 1$, the sum converges absolutely, and for $0 < s < 1$, the infinite series diverges but partial sums still make sense and can be approximated.

Proof. The proof is very similar to Theorem 2.13, but here we set $f(x) = x^{-s}$ where $s > 0$ and $s \neq 1$. By the Euler-Maclaurin formula (Theorem 2.12), we have

$$\sum_{n \leq x} \frac{1}{n^s} = \frac{x^{1-s}}{1-s} - \frac{1}{1-s} + 1 - s \int_1^\infty \frac{\{t\}}{t^{s+1}}\, dt + \mathcal{O}\left(x^{-s}\right).$$

Define $C(s)$ to be the expression in red. Then,

$$C(s) = 1 - \frac{1}{1-s} - s \int_1^\infty \frac{\{t\}}{t^{s+1}}\, dt.$$

If $s > 1$, then the left side of (2.7) approaches $\zeta(s)$ as x approaches ∞, and both x^{1-s} and x^{-s} approach 0. Hence, $C(s) = \zeta(s)$ if $s > 1$. If $0 < s < 1$, then

$$\lim_{x \to \infty} \frac{1}{x^s} = 0.$$

□

2.8 Dirichlet's Hyperbola Method

Here, we introduce the Dirichlet hyperbola method (Theorem 2.15) sand then apply it to study the mean value of the divisor function $d(n)$ as shown in Dirichlet's divisor problem (Theorem 2.16).

Theorem 2.15 (Dirichlet hyperbola method)**.** Let f be a multiplicative function. Define

$$F(n) = \sum_{k=1}^{n} f(k).$$

Suppose $f = g * h$, where g and h are multiplicative functions, so the sum becomes

$$F(n) = \sum_{k=1}^{n} \sum_{xy=k} g(x)\, h(y). \tag{2.8}$$

Then,

$$F(n) = \sum_{x=1}^{a} \sum_{y=1}^{n/x} g(x)h(y) + \sum_{y=1}^{b} \sum_{x=1}^{n/y} g(x)h(y) - \sum_{x=1}^{a} \sum_{y=1}^{b} g(x)h(y) \tag{2.9}$$

Although technically the inner sum in (2.8) runs over all ordered pairs $(x, y) \in \mathbb{N}^2$, we can leave our notation as mentioned. On the xy-plane, these pairs (x, y) lie on a hyperbola, and when the double sum is fully expanded, there is a bijection between the terms of the sum and the lattice points in the first quadrant on the hyperbolas of the form $xy = k$, where $1 \leq k \leq n$ is an integer.

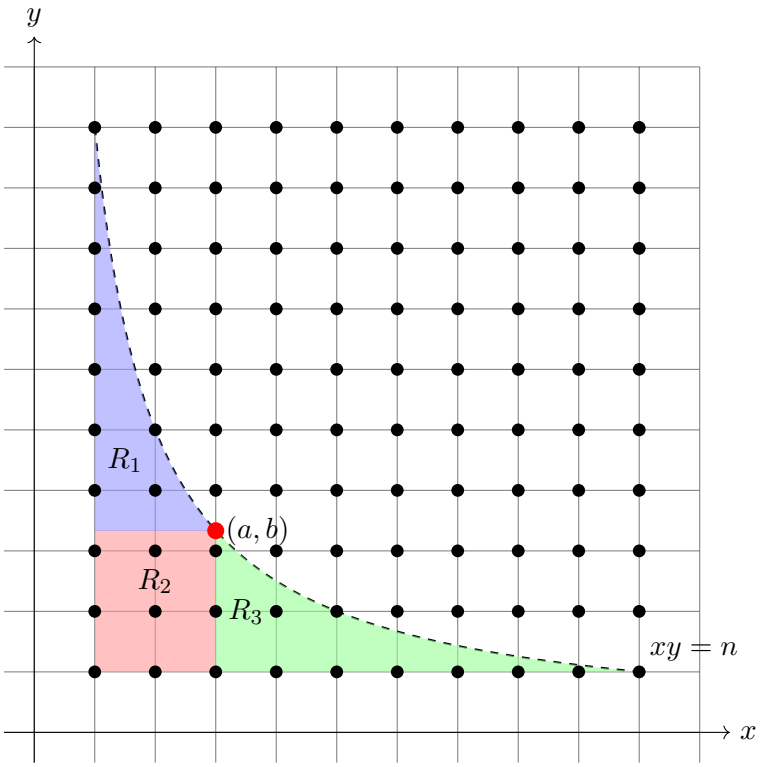

Figure 9: Geometric interpretation of the Dirichlet hyperbola method

Also, the geometric interpretation of the Dirichlet hyperbola method generally involves the principle of inclusion and inclusion. Let $a \in \mathbb{R}$ such that $1 < a < n$, and let $b = \frac{n}{a}$. Then the lattice points (x, y) can be split into three overlapping regions: one region is bounded by $1 \leq x \leq a$ and $1 \leq y \leq \frac{n}{x}$, another region is bounded by $1 \leq y \leq b$ and $1 \leq x \leq \frac{n}{y}$, and the third is bounded by $1 \leq x \leq a$ and $1 \leq y \leq b$. In Figure 9, the first region is the $R_1 \cup R_2$, the second region is $R_2 \cup R_3$, and the third region is R_2. Note that this $R_2 = R_1 \cap R_3$. By the principle of inclusion and exclusion, the sum in (2.8) is the sum over the first region, plus the sum over the second region, minus the sum over the third region. This yields (2.9).

Theorem 2.16 (Dirichlet's divisor problem)**.** For $x \geq 1$,

$$\sum_{n \leq x} d(n) = x \log x + (2\gamma - 1)\, x + \mathcal{O}\left(\sqrt{x}\right),$$

where γ denotes the Euler-Mascheroni constant (Definition 2.11).

From Theorem 2.16, we deduce that

$$\overline{d}(x) \sim \log x,$$

or equivalently, the average order of $d(n)$ is $\log n$. Note that the error term $\mathcal{O}(\sqrt{x})$ in Theorem 2.16 can be improved. It was Dirichlet in 1849 who first gave this estimate, and the Dirichlet divisor problem, precisely stated, is to improve this error bound by finding the smallest value of θ for which the error term $\mathcal{O}\left(x^{\theta+\varepsilon}\right)$ holds for all $\varepsilon > 0$. As of today, this problem remains unsolved and progress has been slow. Having said that, in 1903, Voronoi proved that the error term is $\mathcal{O}\left(x^{1/3}\log x\right)$. In 1928, van der Corput improved the error term to $\mathcal{O}\left(x^{27/82}\right)$ using exponential sums. The best possible error term is the one given by Huxley in 2003, who showed that it is $\mathcal{O}\left(x^{131/416}(\log x)^{26947/8320}\right)$. It is conjectured that $\inf\theta$ lies somewhere between $\frac{1}{4}$ and Huxley's estimate of 0.3149, and it is widely conjectured to be $\frac{1}{4}$.

See Figure 10 for the graphs of $\sum_{n\leq x} d(n)$ and the main term.

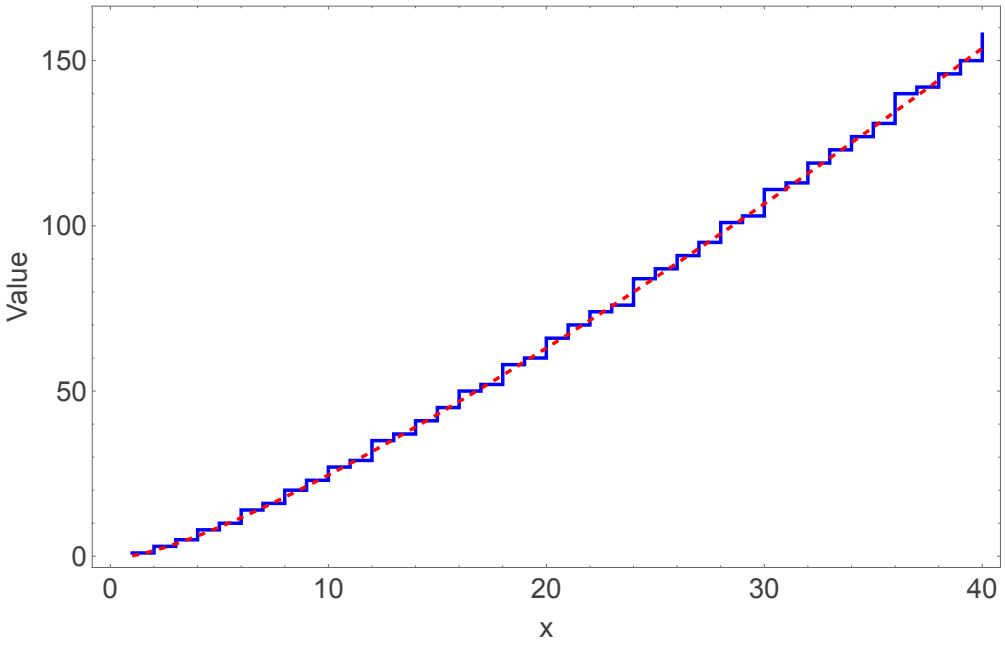

Figure 10: Graphs of $\displaystyle\sum_{n\leq x} d(n)$ and $x\log x + (2\gamma - 1)x$

We now deduce the original bound given by Dirichlet (Theorem 2.16).

Proof. Recall the Dirichlet hyperbola method (Theorem 2.15). Let $g = h = u$ be the constant unit function, which is multiplicative. Then,

$$g * h = u * u = d \quad \text{or equivalently} \quad d(n) = \sum_{d|n} 1.$$

Then, $G(x) = [x] = H(x)$. Let $y = \sqrt{x}$. By the hyperbola method, we have

$$\sum_{n \le x} d(n) = 2 \sum_{n \le \sqrt{x}} \left[\frac{x}{n}\right] - \left[\sqrt{x}\right]^2$$

and the result follows by Theorem 2.13. □

We now apply the Dirichlet hyperbola method to an interesting question: if two positive integers are randomly chosen, what is the probability that they are coprime? To answer this question, we need the result in Theorem 2.17.

Theorem 2.17. Let $\varphi(n)$ denote the Euler totient function. For $x > 1$, we have

$$\sum_{n \le x} \varphi(n) = \frac{3x^2}{\pi^2} + \mathcal{O}(x^{3/2}). \tag{2.10}$$

Proof. Note that $\varphi = \mu * N$, where $N(n) = n$. By applying the Dirichlet hyperbola method with $f = N$ and $g = \mu$ (Theorem 2.15), we have

$$\sum_{n \le x} \varphi(n) = \sum_{n \le y} \mu(n) F\left(\frac{x}{n}\right) + \sum_{m \le x/y} N(m) G\left(\frac{x}{m}\right) - F\left(\frac{x}{y}\right) G(y), \tag{2.11}$$

where

$$F(x) = \sum_{n \le x} N(n) = \frac{x^2}{2} + \mathcal{O}(x) \quad \text{and} \quad G(x) = \sum_{n \le x} \mu(n) = \mathcal{O}(x) \tag{2.12}$$

Letting $y = \sqrt{x}$, one can deduce that

$$\sum_{n \le x} \varphi(n) = \sum_{n \le \sqrt{x}} \frac{\mu(n)}{n^2} \frac{x^2}{2} + \mathcal{O}(x^{3/2}). \tag{2.13}$$

By the Möbius inversion formula (Theorem 2.7), we have

$$\zeta(2) \sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} = 1 \tag{2.14}$$

so

$$\sum_{n \le \sqrt{x}} \frac{\mu(n)}{n^2} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} + \mathcal{O}\left(\sum_{n > \sqrt{x}} \frac{1}{n^2}\right) = \frac{6}{\pi^2} + \mathcal{O}(x^{-1/2}). \tag{2.15}$$

The result follows. □

Now, let $N \in \mathbb{N}$ and S denote the set of $(a, b) \in \mathbb{N}^2$ such that $1 \leq a, b \leq N$. Then, the total number of elements in S such that $\gcd(a, b) = 1$ is given by

$$1 + 2 \sum_{b \leq N} \varphi(b) = \frac{6}{\pi^2} N^2 + \mathcal{O}(N^{3/2}), \tag{2.16}$$

where we used Theorem 2.17. This shows that the probability that two randomly chosen positive integers are relatively prime is $6/\pi^2$.

Chapter 3

The Prime Number Theorem

3.1 Chebyshev's Functions $\theta(x)$ and $\psi(x)$

Define the number of primes less than or equal to x to be $\pi(x)$. Then, we have the prime number theorem (Theorem 3.1) which is a beautiful result on the limiting behaviour of $\pi(x)$.

Theorem 3.1 (prime number theorem)**.** We have

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

Alternatively, we can think of this as an asymptotic equivalence, i.e.

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

We will eventually prove this remarkable result. Historically, the behaviour of $\pi(x)$ as a function of x has been studied by many Mathematicians since the 18th century. Inspection of tables of primes led Gauss and Legendre to independently conjecture in 1792 and 1798 respectively (the graph of $\pi(x)$ is shown in Figures 12, 13, and 11) that

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

It was not until about 100 years later where the conjecture was first proved independently by Hadamard and de la Vallée Poussin in 1896, and the conjecture, as expected, is now known as the prime number theorem. Proofs of the prime number theorem are often classified as either elementary or analytic. The proofs of Hadamard and de la Vallée Poussin are analytic. As mentioned, we will give a proof of this, where the idea leans more towards the ideas of Hadamard and de la Vallée Poussin. As such, the reader would expect some results in Complex Analysis and a fruitful discussion on the Riemann zeta function $\zeta(s)$. Elementary proofs were discovered around 1949 by Selberg and Erdős. Their proofs do not involve $\zeta(s)$ or Complex Analysis in general and hence the name 'elementary'.

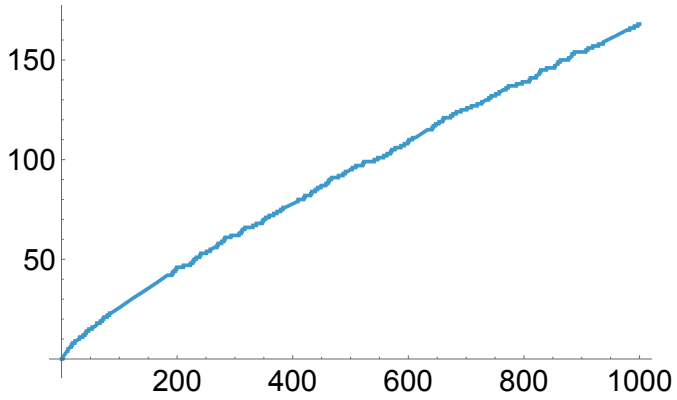

Figure 11: Graph of $\pi(n)$ up to $n = 1000$

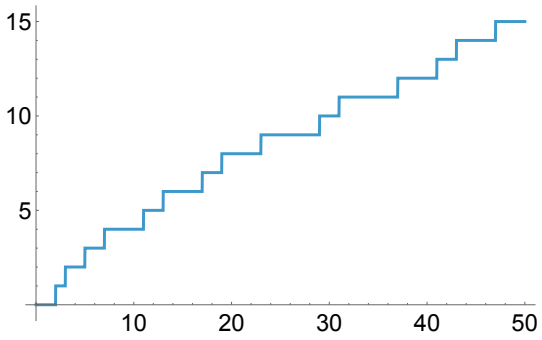

Figure 12: Graph of $\pi(n)$ up to $n = 50$

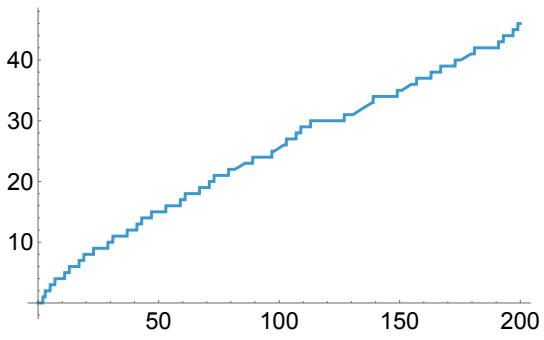

Figure 13: Graph of $\pi(n)$ up to $n = 200$

Definition 3.1 (Chebyshev's first function)**.** Also, known as Chebyshev's theta function, we represent it by $\vartheta(x)$ and it is defined by

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

We will also study Chebyshev's second function, but we shall first define the von Mangoldt function (Definition 3.2).

Definition 3.2 (von Mangoldt function)**.** Let $n, \alpha \in \mathbb{N}$. Define the von Mangoldt function $\Lambda(k)$ as follows:

$$\Lambda(k) = \begin{cases} \log k & \text{if } k = p^{\alpha}; \\ 0 & \text{if } k \neq p^{\alpha} \end{cases}$$

Here, p is a prime number.

Now, we are ready to define Chebyshev's second function (Definition 3.3).

Definition 3.3 (Chebyshev's second function)**.** This is known as the Chebyshev psi function, which is represented by $\psi(x)$. We define it to be

$$\psi(x) = \sum_{k \leq x} \Lambda(k) = \sum_{\alpha \in \mathbb{N}} \sum_{p^\alpha \leq x} \log p.$$

The latter representation of the Chebyshev psi function is more useful.

The graphs of $\vartheta(n)$ and $\psi(n)$ up to $n = 30$ are shown in Figures 14 and 15 respectively.

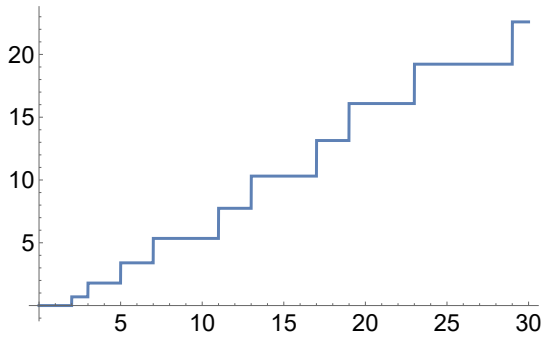

Figure 14: Graph of $\vartheta(n)$ up to $n = 30$

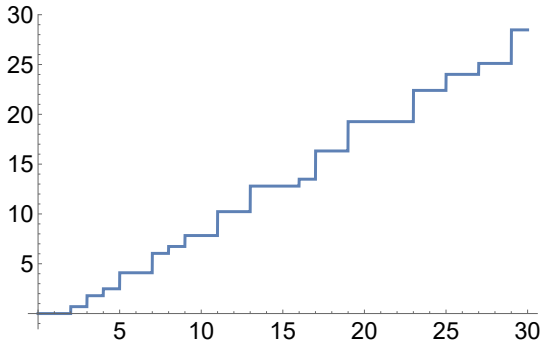

Figure 15: Graph of $\psi(n)$ up to $n = 30$

Theorem 3.2. There exist positive constants c_1 and c_2 such that

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

Theorem 3.3. For any real number $x \geq 1$, we have

$$\vartheta(x) = \psi(x) + \mathcal{O}\left(\sqrt{x}\right).$$

Proof. We note that

$$\psi(x) - \vartheta(x) = \sum_{\substack{p^m \leq x \\ m \geq 2}} \log p = \sum_{\substack{p \leq \sqrt{x} \\ m=2}} \log p + \sum_{p \leq x^{1/3}} \log p \sum_{3 \leq m \leq \frac{\log x}{\log p}} 1.$$

Hence,

$$\psi(x) - \vartheta(x) \leq \psi\left(\sqrt{x}\right) + \sum_{p \leq x^{1/3}} \log p \cdot \frac{\log x}{\log p} \leq \sqrt{x} + x^{1/3} \log x$$

and the result follows. □

In fact, the bound obtained in the proof of Theorem 3.3 can be strengthened. In 2010, Pierre Dusart obtained the following estimate in his paper 'Estimates of some Functions over Primes without the Riemann Hypothesis':

$$0.9999\sqrt{x} \leq \psi(x) - \vartheta(x) \leq 1.00007\sqrt{x} + 1.78\sqrt[3]{x} \quad \text{for } x \geq 121$$

From here, we obtain Corollary 3.1.

Corollary 3.1. For every $x \geq 4$, there exist positive real constants c_1 and c_2 such that

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

Theorem 3.4. For every $x \geq 4$, there exist positive constants c_1 and c_2 such that

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

We now discuss the steps required to prove the prime number theorem. Note that a continuation of the proof will be discussed in Chapter 3.4.

Definition 3.4 (smoothing function). Define $\psi_1(x)$ to be the following integral:

$$\psi_1(x) = \int_1^x \psi(t)\, dt$$

We would need to prove the following chain of implications:

$$\psi_1(x) \sim \frac{x^2}{2} \Rightarrow \psi(x) \sim x \Rightarrow \vartheta(x) \sim x \Rightarrow \pi(x) \sim \frac{x}{\log x} \tag{3.1}$$

In (3.1), we would first need to establish that $\psi_1(x) \sim x^2/2$ using methods in Complex Analysis. We break up (3.1) into several lemmas and prove each one of them individually. Before we state and prove Lemma 3.2, we need a preceding lemma (Lemma 3.1) on an upper bound for the Chebyshev theta function.

Lemma 3.1. $\vartheta(x) \leq x \log x$

Proof. We have

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

and we are done. □

Lemma 3.2. $\psi(x) \sim x \Rightarrow \vartheta(x) \sim x$

Proof. First, recall the latter representation of the Chebyshev psi function in Definition 3.3. Then,

$$\psi(x) = \sum_{p^\alpha \le x} \log p = \sum_{\alpha=1}^{\infty} \sum_{p^\alpha \le x} \log p = \sum_{\alpha=1}^{\infty} \sum_{p \le x^{1/\alpha}} \log p$$

We shall find an upper bound for α. We use the fact that $2^\alpha \le p^\alpha \le x$. This is obvious since the smallest prime number is 2 and $p^\alpha \le x$ arose in Definition 3.3. It is then clear that

$$\alpha \log 2 \le \alpha \log p \le \log x \quad \text{so} \quad \alpha \le \frac{\log x}{\log 2} \le \log_2 x.$$

As such, we have

$$\psi(x) = \sum_{1 \le \alpha \le \log_2 x} \sum_{p \le x^{1/\alpha}} \log p.$$

Recall that the summand

$$\sum_{p \le x^{1/\alpha}} \log p$$

reminds us of the Chebyshev theta function. In particular, it is $\vartheta\left(x^{1/\alpha}\right)$, so we have

$$\psi(x) = \sum_{1 \le \alpha \le \log_2 x} \vartheta\left(x^{1/\alpha}\right).$$

Now,

$$\lim_{x\to\infty} \frac{\psi(x)}{x} = \lim_{x\to\infty} \sum_{1 \le \alpha \le \log_2 x} \frac{\vartheta\left(x^{1/\alpha}\right)}{x} = \lim_{x\to\infty} \frac{\vartheta(x)}{x} + \lim_{x\to\infty} \sum_{2 \le \alpha \le \log_2 x} \frac{\vartheta\left(x^{1/\alpha}\right)}{x}$$

$$\lim_{x\to\infty} \frac{\vartheta(x)}{x} = 1 - \lim_{x\to\infty} \sum_{2 \le \alpha \le \log_2 x} \frac{\vartheta\left(x^{1/\alpha}\right)}{x}$$

$$\ge 1 - \lim_{x\to\infty} \sum_{2 \le \alpha \le \log_2 x} \frac{x^{1/\alpha} \log x}{\alpha x} \text{ by Lemma 3.1}$$

We now need to find an upper bound for

$$\sum_{2\leq\alpha\leq\log_2 x} \frac{x^{1/\alpha}\log x}{\alpha x},$$

which is obviously

$$\sum_{2\leq\alpha\leq\log_2 x} \frac{x^{1/2}\log x}{2x}.$$

As such,

$$\begin{aligned}\lim_{x\to\infty}\frac{\vartheta(x)}{x} &\geq 1-\lim_{x\to\infty}\frac{x^{1/2}\log x}{2x}(\log_2 x-1)\\ &= 1-\frac{1}{2}\lim_{x\to\infty}\frac{\log x}{\sqrt{x}}\left(\frac{\log x}{\log 2}-1\right)\\ &= 1\end{aligned}$$

so it shows that $\vartheta(x)\sim x$. □

Before we state and prove Lemma 3.3 on the asymptotic formula for $\pi(x)$ given that $\vartheta(x)\sim x$, one should recall the Abel summation formula (Theorem 2.11).

Lemma 3.3. $\vartheta(x)\sim x \Rightarrow \pi(x)\sim \dfrac{x}{\log x}$

Proof. Define $b_n=\log n$, where $n=p$ and p is prime. If p is not prime, $b_n=0$. Also, define $f(t)=1/\log t$. By the Abel summation formula (Theorem 2.11), we choose $y=1.5$ so that

$$\sum_{n=1}^{x} b_n f(n) = f(x)\sum_{n=1}^{x} b_n - f(1.5)\sum_{n=1}^{1.5} b_n - \int_{1.5}^{x} f'(t)\sum_{n=1}^{t} b_n\, dt.$$

However, note that

$$f(1.5)\sum_{n=1}^{1.5} b_n = 0,$$

so

$$\sum_{n=1}^{x} b_n f(n) = f(x)\sum_{n=1}^{x} b_n - \int_{1.5}^{x} f'(t)\sum_{n=1}^{t} b_n\, dt.$$

Substituting the known expressions for b_n and f at this stage, we have

$$\sum_{\substack{1\leq n\leq x\\ n \text{ prime}}}\frac{\log n}{\log n} = \frac{1}{\log x}\sum_{\substack{1\leq n\leq x\\ n \text{ prime}}}\log n + \int_{1.5}^{x}\frac{1}{t(\log t)^2}\sum_{\substack{1\leq n\leq t\\ n \text{ prime}}}\log n\, dt$$

$$\pi(x) = \frac{\vartheta(x)}{\log x} + \int_{1.5}^{x}\frac{\vartheta(t)}{t(\log t)^2}\, dt$$

Multiplying both sides by $\log x/x$, we have

$$\begin{aligned}\frac{\pi(x)\log x}{x} &= \frac{\vartheta(x)}{x} + \frac{\log x}{x}\int_{1.5}^{x}\frac{\vartheta(t)}{t(\log t)^2}\,dt\\ \lim_{x\to\infty}\frac{\pi(x)\log x}{x} &= \lim_{x\to\infty}\frac{\vartheta(x)}{x} - \lim_{x\to\infty}\frac{\log x}{x}\int_{1.5}^{x}\frac{\vartheta(t)}{t(\log t)^2}\,dt\\ &= 1 - \lim_{x\to\infty}\frac{\log x}{x}\int_{1.5}^{x}\frac{\vartheta(t)}{t(\log t)^2}\,dt\end{aligned}$$

For $x \geq 4$, it is a well-known fact that there exist positive c_1, c_2 such that $c_1 x \leq \vartheta(x) \leq c_2 x$ (Corollary 3.1). As such,

$$\begin{aligned}\lim_{x\to\infty}\frac{\log x}{x}\int_{1.5}^{x}\frac{\vartheta(t)}{t(\log t)^2}\,dt &\leq \lim_{x\to\infty}\frac{\log x}{x}\int_{1.5}^{x}\frac{c_2 t}{t(\log t)^2}\,dt\\ &= c_2\lim_{x\to\infty}\frac{\log x}{x}\int_{1.5}^{x}\frac{1}{(\log t)^2}\,dt\\ &= c_2\lim_{x\to\infty}\frac{\log x}{x}\left(\int_{1.5}^{\sqrt{x}}\frac{1}{(\log t)^2}\,dt + \int_{\sqrt{x}}^{x}\frac{1}{(\log t)^2}\,dt\right)\\ &\leq c_2\lim_{x\to\infty}\frac{\log x}{x}\left(\int_{1.5}^{\sqrt{x}}\frac{1}{(\log 2)^2}\,dt + \int_{\sqrt{x}}^{x}\frac{1}{(\log\sqrt{x})^2}\,dt\right)\\ &= c_2\lim_{x\to\infty}\frac{\log x}{x}\left(\frac{\sqrt{x}-1.5}{(\log 2)^2} + \frac{4\,(x-\sqrt{x})}{(\log x)^2}\right)\end{aligned}$$

We can split this limit as

$$\frac{1}{(\log 2)^2}\lim_{x\to\infty}\frac{\log x}{\sqrt{x}} - \frac{1.5}{(\log 2)^2}\lim_{x\to\infty}\frac{\log x}{x} + 4\lim_{x\to\infty}\frac{1}{\log x} - 4\lim_{x\to\infty}\frac{1}{\sqrt{x}\,(\log x)}$$

and we see that each of this limits tends to 0. Therefore, we conclude that

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

So, indeed, the implication holds. □

Lemma 3.4. $\psi_1(x) \sim \dfrac{x^2}{2} \Rightarrow \psi(x) \sim x$

Proof. Let

$$a(t) = \sum_{0\leq n\leq t} c_n, \text{ where } c_n \geq 0 \quad \text{and} \quad A(x) = \int_1^x a(t)\,dt.$$

Then, for all functions $a(t)$, if $A(x) \sim Cx^b$, then $a(t) \sim Cbx^{b-1}$. Observe that this is analogous to how the derivative works. That is to say, $a(t)$ is the formal derivative of

$A(t)$. For $\alpha > 1$, we have

$$A(\alpha x) - A(x) = \int_x^{\alpha x} a(t)\ dt.$$

As $a(t)$ is an increasing function, then

$$\begin{aligned} A(\alpha x) - A(x) &\geq \int_x^{\alpha x} a(x)\ dt \\ A(\alpha x) - A(x) &\geq (\alpha x - x)a(x) \\ \frac{A(\alpha x) - A(x)}{\alpha - 1} &\geq xa(x) \\ \lim_{x\to\infty} \frac{A(\alpha x) - A(x)}{x^b(\alpha - 1)} &\geq \lim_{x\to\infty} \frac{a(x)}{x^{b-1}} \end{aligned}$$

Since $A(x) \sim Cx^b$, then

$$\lim_{x\to\infty} \frac{a(x)}{x^{b-1}} \leq \frac{C\left(\alpha^b - 1\right)}{\alpha - 1}.$$

By first principles of differentiation, it is easy to see that the right side is just Cb, so

$$\lim_{x\to\infty} \frac{a(x)}{x^{b-1}} \leq Cb.$$

We shall examine the following too:

$$A(x) - A(\beta x) = \int_{\beta x}^{x} a(t)\ dt \quad \text{where } \beta < 1$$

In a similar fashion, we can prove that

$$Cb \leq \lim_{x\to\infty} \frac{a(x)}{x^{b-1}}.$$

By the squeeze theorem, we conclude that

$$\lim_{x\to\infty} \frac{a(x)}{x^{b-1}} = Cb,$$

so it implies that $a(x) \sim Cbx^{b-1}$. By definition of A, we apply this to ψ_1, i.e. if $\psi_1(x) \sim x^2/2$, then $\psi(x) \sim x$. □

The following integral in Theorem 3.5 looks bizarre and one might wonder what connection it might have with the prime number theorem. We will relate this to $\psi_1(x)$ in due course.

Theorem 3.5. Let $c, u > 0$. Then,

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{u^{-z}\Gamma(z)}{\Gamma(m+z+1)}\,dz = \begin{cases}\frac{1}{m!}(1-u)^m & \text{if } 0<u\leq 1;\\ 0 & \text{if } u>1.\end{cases},$$

where $\Gamma(z)$ denotes the gamma function and it is defined by

$$\Gamma(z) = \int_0^\infty e^{-t}t^{z-1}\,dt \quad \text{where } \Re(z) > 0.$$

We have the following well-known result involving the recurrence relation of the gamma function, which can be proven using integration by parts:

$$\Gamma(z+1) = z\Gamma(z) \quad \text{with initial condition } \Gamma(1) = 1$$

Also, note that the integral in Theorem 3.5 is absolutely convergent. We recall Cauchy's residue theorem from MA3211 Complex Analysis (Theorem 3.6).

Theorem 3.6 (Cauchy's residue theorem)**.** Let $U \subseteq \mathbb{C}$ be a simply connected open subset of the complex plane containing the points $z_1, z_2, \ldots, z_n$. Now, let $U_0 = U \setminus \{z_1, z_2, \ldots, z_n\}$ and C be a closed curve on U_0. Then,

$$\oint_C f(z)\,dz = 2\pi i\sum_{i=1}^{n}\operatorname{Res}(f(z), z = z_i).$$

We are now in a position to prove Theorem 3.5.

Proof. Note that the gamma function has some connection to the factorial. As such, by letting $f(z)$ denote the integrand in Theorem 3.5, we have

$$m!f(z) = \frac{u^{-z}m!}{z(z+1)(z+2)\ldots(z+m)}.$$

We wish to integrate $f(z)$ along a line in the complex plane. It would be ideal to integrate on a closed contour so that we can then apply Cauchy's residue theorem. We can add a circular arc to the line $z = R\cos\alpha$ to get a closed contour C. This contour will comprise an arc γ that is of radius R and centred at $z = 0$. We choose $R > 2m$ so that the integral on the circular arc γ can be evaluated easily for $0 < u \leq 1$.

As such,

$$\oint_C f(z)\,dz = \lim_{R\to\infty}\int_{C-iR\sin\alpha}^{c+iR\sin\alpha} f(z)\,dz + \int_\gamma f(z)\,dz.$$

The trick is to let $R \to \infty$, so we have

$$\oint_C f(z)\ dz = \lim_{R\to\infty} \int_{C-i\infty}^{c+i\infty} f(z)\ dz + \int_\gamma f(z)\ dz.$$

To evaluate the contour integral

$$\oint_C f(z)\ dz,$$

note that $f(z)$ has simple poles (i.e. poles each of order 1) at $z = 0, z = -1, z = -2, \ldots, z = -m$. This can be easily seen by setting the denominator of $f(z)$ to be zero. Note that C encloses all these points. By Cauchy's residue theorem (Theorem 3.6),

$$\oint_C f(z)\ dz = 2\pi i \sum_{n=0}^{m} \operatorname{Res}\left(f(z), z = -n\right).$$

To compute each residue, we have for all $0 \le n \le m$,

$$\operatorname{Res}\left(f(z), z = -n\right) = \lim_{z\to -n} (z+n) f(z) = \lim_{z\to -n} \frac{u^{-z}(z+n)m!}{z(z+1)(z+2)\ldots(z+m)}.$$

Observe that the denominator contains $z + n$ as a factor because $n \le m$. Then, with some tedious but straightforward manipulations, we have

$$\lim_{z\to -n} \frac{u^{-z}m!}{z(z+1)\ldots(z+n-1)\,(z+n+1)\ldots(z+m)} = \binom{m}{n}(-u)^n$$

Substituting this into the residue theorem yields

$$\oint_C f(z)\ dz = 2\pi i \sum_{n=0}^{m} \binom{m}{n} (-u)^n = 2\pi i\,(1-u)^m,$$

where the last equality follows by the binomial theorem. Hence,

$$\begin{aligned} 2\pi i\,(1-u)^m &= \int_{c-i\infty}^{c+i\infty} f(z)\ dz + \lim_{R\to\infty} \int_\gamma f(z)\ dz \\ &= \int_{c-i\infty}^{c+i\infty} f(z)\ dz + \lim_{R\to\infty} \int_\gamma \frac{u^{-z}m!}{z(z+1)(z+2)\ldots(z+m)}\ dz \end{aligned}$$

From here, we can parametrise γ using $z = Re^{i\theta}$, where $\alpha \le \theta \le 2\pi - \alpha$, so

$$\begin{aligned} \int_\gamma \frac{u^{-z}m!}{z(z+1)(z+2)\ldots(z+m)}\ dz &= \int_\alpha^{2\pi-\alpha} \frac{iRe^{i\theta}u^{-R\exp(i\theta)}m!}{Re^{i\theta}\left(Re^{i\theta}+1\right)\left(Re^{i\theta}+2\right)\ldots\left(Re^{i\theta}+m\right)}\ d\theta \\ &= im! \int_\alpha^{2\pi-\alpha} \frac{u^{-R\exp(i\theta)}}{\left(Re^{i\theta}+1\right)\left(Re^{i\theta}+2\right)\ldots\left(Re^{i\theta}+m\right)}\ d\theta \end{aligned}$$

Hence,

$$\left| im! \int_{\alpha}^{2\pi-\alpha} \frac{u^{-R\exp(i\theta)}}{(Re^{i\theta}+1)\,(Re^{i\theta}+2)\ldots(Re^{i\theta}+m)}\, d\theta \right|$$

is equal to

$$\left| m! \int_{\alpha}^{2\pi-\alpha} \frac{u^{-R\exp(i\theta)}}{(Re^{i\theta}+1)\,(Re^{i\theta}+2)\ldots(Re^{i\theta}+m)}\, d\theta \right| \tag{3.2}$$

We use the triangle inequality to help us with bounding. For $0 < u \leq 1$, we have

$$\left| u^{-R\exp(i\theta)} \right| \leq u^{-c}$$

for some positive constant c. Then, we shall examine the denominator of the integral in (3.2), which comprises a product of m terms. Each term is of the form $Re^{i\theta} + n$, where $n \in \mathbb{N}$. By the triangle inequality,

$$\left| Re^{i\theta} + n \right| \geq \left| Re^{i\theta} \right| - n = R - n.$$

Earlier, we stated that $R > 2m$ and since $m \geq n$, then $R > 2n$. As such, $R - n > R - R/2 = R/2$, so

$$\left| Re^{i\theta} + n \right| > \frac{R}{2} \quad \text{which implies that} \quad \frac{1}{|Re^{i\theta}+n|} < \frac{2}{R}.$$

Now, we can properly bound our integral, so

$$\begin{aligned} \left| m! \int_{\alpha}^{2\pi-\alpha} \frac{u^{-R\exp(i\theta)}}{(Re^{i\theta}+1)\,(Re^{i\theta}+2)\ldots(Re^{i\theta}+m)}\, d\theta \right| &< m! \left(\frac{2}{R}\right)^m u^{-c} \int_{\alpha}^{2\pi-\alpha} d\theta \\ &= \frac{2^{m+1}\,(\pi-\alpha)\,m!u^{-c}}{R^m} \end{aligned}$$

As $R \to \infty$, the integral tends to zero!

Now, we can conclude that if $c > 0$ and $0 < u \leq 1$, then

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{u^{-z}\Gamma(z)}{\Gamma(m+z+1)}\, dz = \frac{(1-u)^m}{m!}.$$

If $u > 0$, one can show using the Cauchy-Goursat theorem (Theorem 3.7) that the integral is 0 because $f(z)$ has no poles inside the contour, say D, and so f is holomorphic on D. We are done with our proof. □

Theorem 3.7 (Cauchy-Goursat theorem). If $f(z)$ is holomorphic on a simply connected domain Ω, then for any simply closed contour C in Ω, that contour integral is zero, i.e.

$$\oint_C f(z)\, dz = 0.$$

In order to study the asymptotic behaviour of the smoothing function $\psi_1(x)$, we need to derive an expression for it! We will return to this in Chapter 3.4.

3.2 Merten's Estimates

Before returning to the prime number theorem, we pause to examine a set of classical results due to the German-Polish Mathematician Franz Merten, which provide explicit information about sums over the primes. He proved these results in 1874. Merten's estimates (Theorem 3.8) serve as some of the earliest quantitative refinements in the theory of prime numbers, predating the full development of the prime number theorem. They concern the asymptotic behaviour of three results on prime numbers.

Lemma 3.5. Let x be a positive real number greater than 1. Then,

$$\sum_{n\leq x} \frac{\Lambda(n)}{n} = \log x + \mathcal{O}(1).$$

Proof. The trick is to write

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

Observe that

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

We omit the remaining details (not much anyway). □

Theorem 3.8 (Merten's estimates). Let x be a positive real number greater than 1. Then, the following hold:

(i) We have

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

(ii) We have

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

(iii) We have

$$\prod_{p\leq x}\left(1-\frac{1}{p}\right)=\frac{e^{-A}}{\log x}\left(1+\mathcal{O}\left(\frac{1}{\log x}\right)\right)$$

Here, A is a constant which will be determined.

Example 3.1. Recall from Definition 2.6 that $\omega\left(n\right)$ is the number of distinct prime divisors of n. Find an asymptotic formula for

$$\sum_{n\leq x}\omega\left(n\right).$$

Solution. Rewrite the sum as

$$\sum_{n\leq x}\omega\left(n\right)=\sum_{n\leq x}\sum_{p\mid n}1=\sum_{p\leq x}\sum_{\ell\leq\frac{x}{p}}1=\sum_{p\leq x}\left[\frac{x}{p}\right]=x\sum_{p\leq x}\frac{1}{p}+\mathcal{O}\left(\sum_{p\leq x}1\right).$$

By using Merten's second estimate (Theorem 3.8), we bound the expression in red so an asymptotic formula is

$$x\log\log x+Ax+\mathcal{O}\left(\frac{x}{\log x}\right).$$

□

3.3 The Riemann Zeta Function

In Definition 2.10, we have encountered the Riemann zeta function for real $s>1$. We now give the definition for the *complex analogue* (Definition 3.5) — to be precise, we give the definition of the Riemann zeta function when s is a complex number. Note that a contour plot of $\zeta\left(s\right)$ is shown in Figure 16.

Definition 3.5 (complex analogue of Riemann zeta function)**.** Let $s=\sigma+it$ be a complex number such that $\sigma>1$ and its analytic continuation elsewhere. That is, $\Re(s)>1$. Then, $\zeta(s)$ is defined by

$$\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^{s}}. \tag{3.3}$$

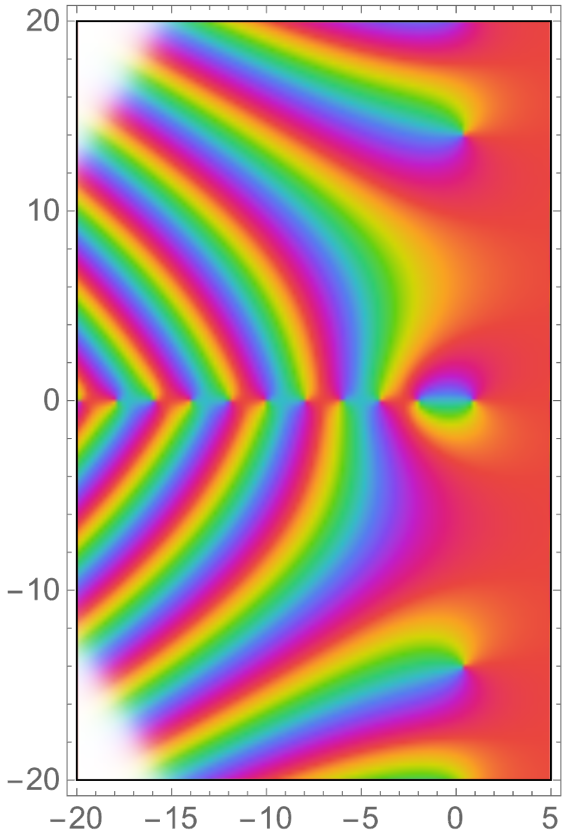

Figure 16: Contour plot of $\zeta(s)$

Theorem 3.9. The Riemann zeta function $\zeta(s)$ is an analytic function for $\sigma > 1$.

Proof. Suppose $\sigma \geq 1 + \delta$. Then,

$$\sum_{n=m}^{M} \left| \frac{1}{n^s} \right| \leq \sum_{n=m}^{M} \frac{1}{n^\sigma} \leq \sum_{n=m}^{M} \frac{1}{n^{1+\delta}}.$$

For every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $M > m > N$, we have

$$\sum_{n=m}^{M} \frac{1}{n^{1+\delta}} < \varepsilon.$$

Hence,

$$\sum_{n=m}^{M} \left| \frac{1}{n^s} \right| < \varepsilon \quad \text{for all } M > m > N.$$

By the Weierstrass M-test, the infinite series (3.3) is absolutely and uniformly convergent in any region $\sigma \geq 1 + \delta$ with $\delta > 0$. We conclude that $\zeta(s)$ is analytic for all $\sigma > 1$. □

Our next step is to study Euler products. Perhaps the most famous Euler product is for all $\sigma > 1$,

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

We omit the proof of this result as one can find many write-ups of it online.

Theorem 3.10. Let f be a multiplicative function such that the series

$$\sum_{n=1}^{\infty} f(n) \quad \text{is absolutely convergent.}$$

Then, the sum of the series can be expressed as an absolutely convergent infinite product, namely

$$\sum_{n=1}^{\infty} f(n) = \prod_p \left(1 + f(p) + f\left(p^2\right) + \ldots\right).$$

In relation to Theorem 3.10, recall that an infinite product

$$\prod_{n=1}^{\infty} (1 + a_n) \quad \text{is absolutely convergent}$$

if

$$\sum_{n=1}^{\infty} \log(1 + a_n) \quad \text{is absolutely convergent.}$$

3.4 Completing the Proof of the Prime Number Theorem

The main result that we need to prove here is as follows:

Theorem 3.11.

$$\frac{\psi_1(x)}{x^2} - \frac{1}{2}\left(1 - \frac{1}{x}\right)^2 = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} x^{z-1} h(z)\, dz,$$

where

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

We would need to study some important results first, especially some pertaining to the Riemann-Zeta Function, $\zeta(s)$. We consider the following theorem first:

Consider the following theorem:

Theorem 3.12.

$$\frac{\psi_1(x)}{x} = \sum_{k \le x} \left(1 - \frac{k}{x}\right) \Lambda(k)$$

Proof. Using the Abel Summation Formula,

$$\begin{aligned}\int_1^x \sum_{k\le t} \Lambda(k)\ dt &= x\sum_{k\le x}\Lambda(k) - \sum_{k\le 1}\Lambda(k) - \sum_{k\le x} k\Lambda(k)\\ &= \sum_{k\le x}(x-k)\,\Lambda(k)\end{aligned}$$

Recall that the integral of the sum of $\Lambda(k)$ at the start is $\psi_1(x)$, so dividing both sides by x, the result follows. □

Corollary 3.2.

$$\frac{\psi_1(x)}{x} = \sum_{k=1}^{\infty}\int_{c-i\infty}^{c+i\infty} f_k(z)\ dz,$$

where

$$2\pi i f_k(z) = \Lambda(k)\left(\frac{x}{k}\right)^z \frac{1}{z\,(z+1)}.$$

Proof. From Theorem 2.1, setting $m = 1$ and $u = k/x$, we have

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\left(\frac{k}{x}\right)^{-z}\frac{1}{z\,(z+1)}\ dz = 1 - \frac{k}{x}.$$

Multiplying by both sides by $\Lambda(k)$ and summing over all k which are $\le x$,

$$\frac{\psi_1(x)}{x} = \sum_{k\le x}\frac{\Lambda(k)}{2\pi i}\int_{c-i\infty}^{c+i\infty}\left(\frac{k}{x}\right)^{-z}\frac{1}{z\,(z+1)}\ dz$$

Observe that the integral vanishes if $k > x$. The result follows. □

Theorem 3.13. For all $\Re(s) > 1$,

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

Proof. The expression on the right prompts us to consider the derivative of $\log(\zeta(s))$. Using the Euler Product, it is easy to see that

$$\prod_{p\text{ prime}}\frac{1}{1-p^{-s}} = \prod_{p\text{ prime}}\left(\sum_{\alpha=0}^{\infty}\frac{1}{p^{\alpha s}}\right) = \sum_{n=1}^{\infty}\frac{1}{n^s} = \zeta(s).$$

Taking logarithms on both sides and differentiating, we have

$$\begin{aligned}\frac{d}{ds}\log\left(\zeta(s)\right) &= \frac{d}{ds}\left(\log\prod_{p\text{ prime}}\frac{1}{1-p^{-s}}\right)\\ \frac{\zeta'(s)}{\zeta(s)} &= -\frac{d}{ds}\left[\sum_{p\text{ prime}}\log\left(1-p^{-s}\right)\right]\\ &= -\sum_{p\text{ prime}}\frac{\log p}{p^s-1}\\ &= -\sum_{p\text{ prime}}\left(\log p\sum_{k=1}^{\infty}\frac{1}{p^{ks}}\right)\end{aligned}$$

and the result follows. □

We are now ready to prove Theorem 3.1.

Proof. From Corollary 3.1, we would need to change the order of summation and integration, so we shall make use of the monotone convergence theorem to justify that we can perform this operation. It suffices to prove that the series

$$\sum_{k=1}^{\infty}\int_{c-i\infty}^{c+i\infty}\left|f_k(z)\right|\,dz$$

is convergent, where

$$f_k(z)=\frac{1}{2\pi i}\left(\frac{x}{k}\right)^z\frac{\Lambda(k)}{z\left(z+1\right)}.$$

Consider the following N^{th} partial sum

$$\begin{aligned}\sum_{k=1}^{N}\int_{c-i\infty}^{c+i\infty}\left|\frac{1}{2\pi i}\left(\frac{x}{k}\right)^z\frac{\Lambda(k)}{z\left(z+1\right)}\right|dz &= \frac{1}{2\pi}\sum_{k=1}^{N}\Lambda(k)\int_{c-i\infty}^{c+i\infty}\left|\left(\frac{x}{k}\right)^z\cdot\frac{1}{z\left(z+1\right)}\right|dz\\ &= \frac{1}{2\pi}\sum_{k=1}^{N}\Lambda(k)\int_{c-i\infty}^{c+i\infty}\left|\left(\frac{x}{k}\right)^c\cdot\frac{1}{z\left(z+1\right)}\right|dz\\ &\le \frac{1}{2\pi}\sum_{k=1}^{N}\frac{\Lambda(k)}{k^c}\int_{c-i\infty}^{c+i\infty}\frac{x^c}{|z|^2}\,dz \text{ by the triangle inequality}\end{aligned}$$

We now split this integral into its real and imaginary part. Obviously, the integral evaluated over the real numbers is zero. We can consider setting $z=c+iy$, where $c,y\in\mathbb{R}$, so it is then clear that the integral becomes

$$\int_{-\infty}^{\infty}\frac{x^c}{c^2+y^2}\,dy=\frac{\pi x^c}{c}.$$

Thus,

$$\sum_{k=1}^{N}\int_{c-i\infty}^{c+i\infty}\left|\frac{1}{2\pi i}\left(\frac{x}{k}\right)^{z}\frac{\Lambda(k)}{z\,(z+1)}\right|\,dz\leq\frac{x^{c}}{2c}\sum_{k=1}^{N}\frac{\Lambda(k)}{k^{c}}$$

and since $x^c/2c$ is a constant, then the series

$$\sum_{k=1}^{\infty}\int_{c-i\infty}^{c+i\infty}|f_k(z)|\;dz$$

is convergent. Now,

$$\begin{aligned}\frac{\psi_1(x)}{x}&=\sum_{k=1}^{\infty}\int_{c-i\infty}^{c+i\infty}\frac{\Lambda(k)}{2\pi i}\left(\frac{x}{k}\right)^{z}\frac{1}{z\,(z+1)}\,dz\\ \frac{\psi_1(x)}{x^2}&=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{x^{z-1}}{z\,(z+1)}\sum_{k=1}^{\infty}\frac{\Lambda(k)}{k^{z}}\,dz\\ &=-\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{x^{z-1}}{z\,(z+1)}\left[\frac{\zeta'(z)}{\zeta(z)}\right]\,dz\text{ by Theorem 3.3}\end{aligned}$$

All that is left to prove is

$$\left(1-\frac{1}{x}\right)^{2}=\frac{1}{\pi i}\int_{c-i\infty}^{c+i\infty}\frac{x^{z-1}}{z\,(z+1)\,(z-1)}\,dz.$$

This is clear by Theorem 2.1. We can set $u=1/x$ and $m=2$ and the result follows. □

We now investigate the analytic continuation of the Riemann Zeta Function. It is of interest to study $\zeta(s)$ near the line $\sigma=1$. From here, we can obtain upper bounds for $|\zeta(s)|$ and $|\zeta'(s)|$. First, recall that the conventional way to write s is of the form $\sigma+it$.

Theorem 3.14. For $\Re(s)>0$ and $s\neq 1$,

$$\zeta(s)=\sum_{n=1}^{N}\frac{1}{n^{s}}+\frac{N^{1-s}}{s-1}-s\int_{N}^{\infty}\frac{\{x\}}{x^{s+1}}\,dx,$$

where $\{x\}=x-\lfloor x\rfloor$ denotes the fractional part of x.

Proof. We recall Abel's Summation Formula, so letting $a(n)=1$ and $f(n)=1/n^s$, as well as $y=0.9$ and $x=N$, we have

$$\begin{aligned}\sum_{0.9\leq n\leq N}\frac{1}{n^{s}}&=\frac{1}{N^{s}}\sum_{n\leq N}-\frac{1}{0.9^{s}}\sum_{n\leq 0.9}+\int_{0.9}^{N}\frac{s}{n^{s+1}}\sum_{n\leq t}\,dt\\ \sum_{n=1}^{N}\frac{1}{n^{s}}&=N^{1-s}+s\int_{1}^{N}\frac{\lfloor t\rfloor}{t^{s+1}}\,dt\end{aligned}$$

For the integral containing the floor function, for $s \neq 1$,

$$\begin{aligned} s\int_1^N \frac{\lfloor t \rfloor}{t^{s+1}}\,dt &= s\int_1^N \frac{t-\{t\}}{t^{s+1}}\,dt \\ &= s\int_1^N \frac{1}{t^s}\,dt - s\int_1^N \frac{\{t\}}{t^{s+1}}\,dt \\ &= \frac{s}{s-1} - \frac{sN^{1-s}}{s-1} - s\int_1^N \frac{\{t\}}{t^{s+1}}\,dt \end{aligned}$$

With some simple rearrangement and as t is a dummy variable, we can let it be x and so

$$\sum_{n=1}^N \frac{1}{n^s} = -\frac{N^{1-s}}{s-1} + \frac{s}{s-1} - s\int_1^N \frac{\{x\}}{x^{s+1}}\,dx$$

$$\sum_{n=1}^N \frac{1}{n^s} + \frac{N^{1-s}}{s-1} - s\int_N^\infty \frac{\{x\}}{x^{s+1}}\,dx = \frac{s}{s-1} - s\int_1^\infty \frac{\{x\}}{x^{s+1}}\,dx$$

We now need to prove that

$$\zeta(s) = \frac{s}{s-1} - s\int_1^\infty \frac{\{x\}}{x^{s+1}}\,dx.$$

Observe that

$$\int_1^\infty \frac{\{x\}}{x^{s+1}}\,dx = \int_1^2 \frac{x-1}{x^{s+1}}\,dx + \int_2^3 \frac{x-2}{x^{s+1}}\,dx + \int_3^4 \frac{x-3}{x^{s+1}}\,dx + \dots$$

so

$$\begin{aligned} \int_1^\infty \frac{\{x\}}{x^{s+1}}\,dx &= \sum_{j=1}^\infty \int_j^{j+1} \frac{x-j}{x^{s+1}}\,dx \\ &= \sum_{j=1}^\infty \left[\frac{j^{1-s}}{s\,(s-1)} - \frac{1}{(j+1)^s\,(s-1)} - \frac{j}{(j+1)^s s\,(s-1)}\right] \end{aligned}$$

Hence,

$$\begin{aligned} \frac{s}{s-1} - s\int_1^\infty \frac{\{x\}}{x^{s+1}}\,dx &= \frac{s}{s-1} - s\sum_{j=1}^\infty \left[\frac{j^{1-s}}{s\,(s-1)} - \frac{1}{(j+1)^s\,(s-1)} - \frac{j}{(j+1)^s s\,(s-1)}\right] \\ &= \frac{s}{s-1} - \frac{1}{s-1}\sum_{j=1}^\infty \frac{1}{j^{s-1}} + \frac{s}{s-1}\sum_{j=1}^\infty \frac{1}{(j+1)^s} + \frac{1}{s-1}\sum_{j=1}^\infty \frac{j}{(j+1)^s} \\ &= \frac{s}{s-1} - \frac{\zeta(s-1)}{s-1} + \frac{s\,(\zeta(s)-1)}{s-1} + \frac{1}{s-1}\left(\sum_{j=0}^\infty \frac{1}{j^{s-1}} + \frac{1}{j^s}\right) \\ &= \frac{s}{s-1} - \frac{\zeta(s-1)}{s-1} + \frac{s\,(\zeta(s)-1)}{s-1} + \frac{\zeta(s-1)+\zeta(s)}{s-1} \\ &= \zeta(s) \end{aligned}$$

Remark 3.1. Theorem 3.4 is very similar to Chapter 6 Problem 16 in Rudin's book on Principles of Mathematical Analysis. It asks the reader to prove that

$$\zeta(s) = \frac{s}{s-1} - s\int_1^\infty \frac{x - \lfloor x \rfloor}{x^{s+1}}\, dx.$$

□

As a corollary, we find an expression for $\zeta'(s)$.

Corollary 3.3.

$$\zeta'(s) = -\sum_{n=1}^{N} \frac{\log n}{n^s} + s\int_N^\infty \frac{\{x\}\log x}{x^{s+1}}\, dx - \int_N^\infty \frac{\{x\}}{x^{s+1}}\, dx - \frac{N^{1-s}\log N}{s-1} - \frac{N^{1-s}}{(s-1)^2}$$

Theorem 3.15. For $\sigma \geq 1 - \dfrac{a}{\log t}$, where $\sigma \geq \dfrac{1}{2}$, $a > 0$ and $t \geq 2$,

$$|\zeta(s)| \leq e^a (5 + \log t)$$

Proof. By the triangle inequality,

$$|\zeta(s)| \leq \underbrace{\left|\sum_{n=1}^{N} \frac{1}{n^s}\right|}_{A_1} + \underbrace{\left|\frac{N^{1-s}}{s-1}\right|}_{A_2} + \underbrace{\left|s\int_N^\infty \frac{\{x\}}{x^{s+1}}\, dx\right|}_{A_3}.$$

We first examine A_2. For $\sigma > 0$ and $t \neq 0$,

$$\left|\frac{N^{1-s}}{s-1}\right| = \left|\frac{N^{1-\sigma-it}}{\sigma + it - 1}\right| = \frac{N^{1-\sigma}\,|N^{-it}|}{\sqrt{(\sigma-1)^2 + t^2}} = \frac{N^{1-\sigma}}{\sqrt{(\sigma-1)^2 + t^2}} \leq \frac{N^{1-\sigma}}{|t|}.$$

We choose a positive constant μ such that $\mu \geq 1 - \sigma$. Thus,

$$\left|\frac{N^{1-s}}{s-1}\right| \leq \frac{N^\mu}{|t|} \leq \frac{3}{2}N^\mu$$

which is our bound for A_2. The last inequality follows from $t \geq 2$. Now, we work on A_1. This appears to be the N^{th} partial sum of $\zeta(s)$. We have

$$\left|\sum_{n=1}^{N} \frac{1}{n^s}\right| = \left|\sum_{n=1}^{N} \frac{1}{n^{\sigma+it}}\right| \leq \sum_{n=1}^{N} \frac{1}{n^\sigma} \leq \sum_{n=1}^{N} \frac{1}{n^{1-\mu}} = \sum_{n=1}^{N} \frac{n^\mu}{n} \leq \sum_{n=1}^{N} \frac{N^\mu}{n} = N^\mu\left(1 + \sum_{n=2}^{N} \frac{N^\mu}{n}\right),$$

but we are not done yet. There is something that can still be done to the modified harmonic series

$$\sum_{n=2}^{N} \frac{1}{n}.$$

In a trick similar to proving the divergence of the harmonic series using the integral test, it is easy to show that

$$\sum_{n=2}^{N} \frac{1}{n} \leq \int_{1}^{N} \frac{1}{x}\, dx = \log N.$$

Therefore, we have constructed a bound for A_1, which is

$$\left| \sum_{n=1}^{N} \frac{1}{n^s} \right| \leq N^{\mu} \left(1 + \log N\right).$$

The bound for A_3 is easy to establish too. For $\sigma \geq 1 - \mu$, $\sigma \geq 1/2 > 0$ and $t \geq 2$, we have

$$\begin{aligned} \left| s \int_{N}^{\infty} \frac{\{x\}}{x^{s+1}}\, dx \right| &\leq |\sigma + it| \left| \int_{N}^{\infty} \frac{1}{x^{\sigma+1+it}}\, dx \right| \\ &= \sqrt{\sigma^2 + t^2} \int_{N}^{\infty} \frac{1}{x^{\sigma+1}}\, dx \\ &= \frac{\sqrt{\sigma^2 + t^2}}{\sigma N^{\sigma}} \\ &\leq \frac{\sigma + t}{\sigma N^{\sigma}} \\ &= \frac{1}{N^{\sigma}} \left(1 + \frac{t}{\sigma}\right) \end{aligned}$$

Using the fact that $N \geq 1$ and $\sigma \geq 1/2$, it is easy to see that

$$\frac{1}{N^{\sigma}} \left(1 + \frac{t}{\sigma}\right) \leq \frac{1 + 2t}{N^{\sigma}} \leq \frac{1 + 2t}{N^{1-\mu}}.$$

Therefore,

$$|\zeta(s)| \leq N^{\mu} \left(\frac{3}{2} + \frac{1 + 2t}{N} + \log N\right).$$

Since N is an arbitrary positive integer, we can let $N \leq t \leq N + 1$ and so $N \geq 2$ (because $t \geq 2$). Consequently,

$$|\zeta(s)| \leq t^{\mu} \left(\frac{3}{2} + \frac{1 + 2\,(N+1)}{N} + \log t\right) \leq t^{\mu} \left(5 + \log t\right).$$

We shall set $a = \mu \log t$ and the result follows. □

Theorem 3.16. For $\sigma \geq 1 - \dfrac{a}{\log t}$, where $\sigma \geq \dfrac{1}{2}$, $a > 0$ and $t \geq 2$,

$$|\zeta'(s)| < e^{a} \left(\log t + 3\right)^2$$

Proof. We had an expression for $\zeta'(s)$ in Corollary 3.2. Using the triangle inequality, we have

$$|\zeta'(s)| \leq \underbrace{\left|\sum_{n=1}^{N} \frac{\log n}{n^s}\right|}_{B_1} + \underbrace{\left|s\int_N^\infty \frac{\{x\}\log x}{x^{s+1}}\,dx\right|}_{B_2} + \underbrace{\left|\int_N^\infty \frac{\{x\}}{x^{s+1}}\,dx\right|}_{B_3} + \underbrace{\left|\frac{N^{1-s}\log N}{s-1}\right|}_{B_4} + \underbrace{\left|\frac{N^{1-s}}{(s-1)^2}\right|}_{B_5}.$$

We first deal with B_1. Note that $N \leq t$, so

$$\left|\sum_{n=1}^{N} \frac{\log n}{n^s}\right| \leq \log N \left|\sum_{n=1}^{N} \frac{1}{n^s}\right| \leq \log t \left|\sum_{n=1}^{N} \frac{1}{n^s}\right| \leq N^\mu \log t\,(1+\log N)\,.$$

The last inequality was established in our proof of Theorem 3.5.

For B_4,

$$\left|\frac{N^{1-s}\log N}{s-1}\right| \leq \log t \left|\frac{N^{1-s}}{s-1}\right| \leq \log t \cdot \frac{|N^{1-s}|}{|t|} \leq \frac{e^a \log t}{2}.$$

As for B_5, it is easy to see that

$$\left|\frac{N^{1-s}}{(s-1)^2}\right| \leq \frac{e^a}{4}.$$

We adopted a similar method to find an upper bound for B_3 in the proof of Theorem 3.5 (see the proof of the upper bound for A_3). Thus,

$$\left|\int_N^\infty \frac{\{x\}}{x^{s+1}}\,dx\right| \leq \frac{1}{\sigma N^\sigma} = \frac{N^{1-\sigma}}{\sigma N} \leq \frac{N^{a/\log t}}{\sigma N} \leq \frac{t^{a/\log t}}{\sigma N} = \frac{2e^a}{N} \leq e^a.$$

Now, we are left to find an upper bound for B_2.

$$\begin{aligned}
\left| s\int_N^\infty \frac{\{x\}\log x}{x^{s+1}}\,dx \right| &\le |s| \left| \int_N^\infty \frac{\log x}{x^{s+1}}\,dx \right| \\
&\le |s| \int_N^\infty \frac{\log x}{x^{\sigma+1}}\,dx \\
&= -\frac{|s|}{\sigma}\left[\frac{\log x}{x^\sigma}\right]_N^\infty + \frac{|s|}{\sigma}\int_N^\infty \frac{1}{x^{\sigma+1}}\,dx \text{ using integration by parts} \\
&= \frac{|s|}{\sigma}\cdot\frac{\log N}{N^\sigma} + \frac{|s|}{\sigma}\cdot\frac{1}{\sigma N^\sigma} \\
&= \frac{|s|}{\sigma N^\sigma}\left(\log N + \frac{1}{\sigma}\right) \\
&\le \frac{1}{N^\sigma}\left(1+\frac{t}{\sigma}\right)(\log N + 2) \\
&\le \frac{1+2t}{N^{1-\mu}}(\log N + 2) \\
&\le N^\mu\left(2+\frac{3}{N}\right)(\log N + 2) \\
&\le \frac{7}{2}N^\mu(\log N + 2) \\
&\le \frac{7}{2}e^a(\log N + 2) \\
&\le \frac{7}{2}e^a(\log t + 2)
\end{aligned}$$

Therefore,

$$\begin{aligned}
|\zeta'(s)| &\le N^\mu \log t\,(1+\log N) + \frac{7}{2}N^\mu(\log t + 2) + e^a + \frac{e^a\log t}{2} + \frac{e^a}{4} \\
&= N^\mu\left(\frac{9}{2}\log t + \log t\log N + 7\right) + e^a\left(\frac{5}{4}+\frac{1}{2}\log t\right) \\
&\le e^a\left[\frac{9}{2}\log t + (\log t)^2 + 7 + \frac{5}{4} + \frac{1}{2}\log t\right] \\
&= e^a\left[(\log t)^2 + 5\log t + 8 + \frac{1}{4}\right] \\
&< e^a\left[(\log t)^2 + 6\log t + 9\right] \\
&= e^a(\log t + 3)^2
\end{aligned}$$

□

We are almost towards the end of the proof. We now talk about the Riemann Zeta Function being non-vanishing on the line $\Re(z) = 1$.

Theorem 3.17. For $t \neq 0$, $\zeta(1+it) \neq 0$

We would need some lemmas to prove this result.

Lemma 3.6. For all $\theta \in \mathbb{R}$, $3 + 4\cos\theta + \cos 2\theta \geq 0$

Lemma 3.7. For $\sigma > 1$,

$$\zeta(s) = e^{G(s)},$$

where

$$G(s) = \sum_{p} \sum_{m=1}^{\infty} \frac{1}{mp^{ms}}$$

Proof. Using the Euler Product, we have

$$\zeta(s) = \prod_{p \text{ prime}} \left(1 - \frac{1}{p^s}\right)^{-1}.$$

Then,

$$-\sum_{p \text{ prime}} \log\left(1 - \frac{1}{p^s}\right) = \sum_{p \text{ prime}} \sum_{m=1}^{\infty} \frac{1}{mp^{ms}} = G(s).$$

The result follows. □

Lemma 3.8. For $\sigma > 1$ and all $t \in \mathbb{R}$,

$$|\zeta(\sigma)|^3 \, |\zeta(\sigma + it)|^4 \, |\zeta(\sigma + 2it)| \geq 1$$

Proof. We mainly use Lemma 3.2 here.

$$\begin{aligned}
\zeta(s) &= \exp\left(\sum_{p \text{ prime}} \sum_{m=1}^{\infty} \frac{1}{mp^{ms}}\right) \\
&= \exp\left\{\sum_{p \text{ prime}} \sum_{m=1}^{\infty} \frac{1}{m} \exp\left[-(\log p)\, ms\right]\right\} \\
&= \exp\left[\sum_{p \text{ prime}} \sum_{m=1}^{\infty} \frac{1}{m} \exp\left(-m\sigma \log p - itm \log p\right)\right] \text{ since } s = \sigma + it
\end{aligned}$$

By applying Euler's Formula, it is then easy to establish that

$$|\zeta(s)| = \exp\left[\sum_{p \text{ prime}} \sum_{m=1}^{\infty} \frac{1}{mp^{\sigma m}} \cos\left(tm \log p\right)\right].$$

Hence,

$$|\zeta(\sigma)|^3 |\zeta(\sigma + it)|^4 \, |\zeta(\sigma + 2it)| = \exp\left\{\sum_{p \text{ prime}} \sum_{m=1}^{\infty} \frac{1}{mp^{\sigma m}} \left[3 + 4\cos\left(tm \log p\right) + \cos\left(2tm \log p\right)\right]\right\}.$$

We set $\theta = tm \log p$, so it is easy to establish that $|\zeta(\sigma)|^3 \, |\zeta(\sigma + it)|^4 \, |\zeta(\sigma + 2it)| \geq 1$ using Lemma 3.1. □

We will now prove Theorem 3.7.

Proof. We shall prove this theorem by contradiction. Suppose on the contrary that there exists some $t_0 \neq 0$ such that $\zeta(1+it_0)=0$. Using Lemma 3.3, by dividing both sides by $\sigma-1$,

$$|(\sigma-1)\,\zeta(\sigma)|^3\left|\frac{\zeta(\sigma+it)}{\sigma-1}\right|^4|\zeta(\sigma+2it)|\geq\frac{1}{\sigma-1}.$$

Because $(\sigma-1)\zeta(\sigma)$ is an analytic function, we can infer that $\zeta(\sigma)$ has a simple pole at $\sigma=1$. Its residue is 1. We see that

$$\lim_{\sigma\to1^+}(\sigma-1)\zeta(\sigma)=1,$$

so

$$\lim_{\sigma\to1^+}|(\sigma-1)\zeta(\sigma)|^3=1.$$

By using the first principles of the derivative, it is easy to see that

$$\lim_{\sigma\to1^+}\left|\frac{\zeta(\sigma+it)}{\sigma-1}\right|^4=\left[\zeta'(1+it_0)\right]^4.$$

Lastly,

$$\lim_{\sigma\to1^+}\zeta(\sigma+2it_0)=\zeta(1+2it_0).$$

We have been investigating the behaviour of σ as

$$f(\sigma)=|(\sigma-1)\,\zeta(\sigma)|^3\left|\frac{\zeta(\sigma+it)}{\sigma-1}\right|^4|\zeta(\sigma+2it)|$$

approaches to 1 from the right. We see that as $\sigma\to1^+$, $1/(\sigma-1)$ tends to positive infinity. However, by the analytic continuation of the Riemann Zeta Function, for $\Re(s)>0$, $\zeta(s)$ and $\zeta'(s)$ have no poles except at $s=1$. As such, $f(\sigma)$ is tending towards a finite value. This yields a contradiction and we conclude that there does not exist $t\neq0$ such that $\zeta(1+it)=0$. □

Theorem 3.18. If s is a complex number such that $\Re(s)>1$, then

$$\sum_{n=1}^{\infty}\frac{\mu(n)}{n^s}=\frac{1}{\zeta(s)}.$$

In Theorem 3.8, we say that the Dirichlet Series that generates the Möbius Function is the multiplicative inverse of the Riemann Zeta Function.

Theorem 3.19. For $|t|\geq2$, there exists a positive constant M such that

$$\left|\frac{1}{\zeta(s)}\right|<M\,(\log t)^7\quad\text{and}\quad\left|\frac{\zeta'(s)}{\zeta(s)}\right|<M\,(\log t)^9$$

whenever $\sigma\geq1$ and $t\geq e$.

Proof. For $\sigma \geq 2$, we have

$$\left|\frac{1}{\zeta(s)}\right| = \left|\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}\right| \leq \sum_{n=1}^{\infty} \frac{1}{n^2} = \zeta(2).$$

Also, as a Corollary of Theorem 3.3, we have

$$\left|\frac{\zeta'(s)}{\zeta(s)}\right| \leq \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^2}.$$

Theorem 3.9 holds for $\sigma \geq 2$. Now, we prove that the theorem also holds for $1 \leq \sigma \leq 2$ and $t \geq e$. From Lemma 3.3, taking the fourth root,

$$|\zeta(\sigma)|^{3/4}|\zeta(\sigma + 2it)|^{1/4} \geq \frac{1}{|\zeta(\sigma + it)|}.$$

Since $(\sigma - 1)\zeta(\sigma)$ is bounded in the interval $1 \leq \sigma \leq 2$, then we can set $(\sigma - 1)\zeta(\sigma) \leq M$ for some positive constant M. However, note that we need to find a bound for $\zeta(\sigma)$, so the restriction on the values of σ now becomes $1 < \sigma \leq 2$. Also, it is clear that $\zeta(\sigma + 2it) = \mathcal{O}(\log t)$ for $1 \leq \sigma \leq 2$. Therefore, one can easily establish that

$$\frac{1}{|\zeta(\sigma + it)|} \leq \frac{A(\log t)^{1/4}}{(\sigma - 1)^{3/4}},$$

where A is an absolute constant. Therefore, for some positive constant B, we have

$$|\zeta(\sigma + it)| > \frac{B(\sigma - 1)^{3/4}}{(\log t)^{1/4}} \text{ if } 1 < \sigma \leq 2 \text{ and } t \geq e.$$

The above holds trivially for $\sigma = 1$. Let α be a number such that $1 < \alpha < 2$. Then, if $1 \leq \sigma \leq \alpha$ and $t \geq e$, by Theorem 3.5, we have

$$\begin{aligned} |\zeta(\sigma + it) - \zeta(\alpha + it)| &\leq \int_{\sigma}^{\alpha} |\zeta'(u + it)| \ du \\ &\leq (\alpha - \sigma)\, M(\log t)^2 \\ &\leq (\alpha - 1)\, M(\log t)^2 \end{aligned}$$

By the triangle inequality,

$$\begin{aligned}
|\zeta(\sigma+it)| &= |\zeta(\alpha+it)+\zeta(\sigma+it)-\zeta(\alpha+it)| \\
&\geq |\zeta(\alpha+it)| - |\zeta(\sigma+it)-\zeta(\alpha+it)| \\
&\geq |\zeta(\alpha+it)| - (\alpha-1)\,M(\log t)^2 \\
&\geq \frac{B(\alpha-1)^{3/4}}{(\log t)^{1/4}} - (\alpha-1)\,M(\log t)^2 \\
&= \frac{(\alpha-1)^{3/4}}{(\log t)^{1/4}}\left[B-(\alpha-1)^{1/4}M(\log t)^{7/4}\right] \\
&= \frac{B(\alpha-1)^{3/4}}{(\log t)^{1/4}} - (\alpha-1)M(\log t)^2
\end{aligned}$$

We require

$$\frac{B(\alpha-1)^{3/4}}{(\log t)^{1/4}} = 2(\alpha-1)M(\log t)^2,$$

so

$$\alpha = 1+\left(\frac{B}{2M}\right)^4\frac{1}{(\log t)^9}.$$

Clearly, $\alpha > 1$ and also, $\alpha < 2$ if $t \geq t_0$ for some t_0. Thus, if $t \geq t_0$ and $1 \leq \sigma \leq 2$,

$$|\zeta(\sigma+it)| \geq \frac{C}{(\log t)^7}.$$

The inequality holds with a different C if $e \leq t \leq t_0$.

We have gotten an upper bound for $|1/\zeta(s)|$. The upper bound for $|\zeta'(s)/\zeta(s)|$ immediately follows from Theorem 3.6. □

Theorem 3.20. The function

$$F(s) = -\frac{\zeta'(s)}{\zeta(s)} - \frac{1}{s-1}$$

is analytic at $s = 1$.

Proof. This is obvious since $-\zeta'(s)/\zeta(s)$ and $1/(s-1)$ each has a first order pole of residue 1 at $s = 1$. Hence, their difference is analytic at $s = 1$. □

Lemma 3.9 (Riemann-Lebesgue Lemma). Let $f \in L^1(\mathbb{R}^n)$ be an integrable function, i.e. $f : \mathbb{R}^n \to \mathbb{C}$ is a measurable function such that

$$\|f\|_{L^1} = \int_{\mathbb{R}^n} |f(x)|\ dx < \infty.$$

Let $\widehat{f}$ denote the Fourier Transform of f. That is to say,

$$\widehat{f} : \mathbb{R}^n \to \mathbb{C},\ \xi \mapsto \int_{\mathbb{R}^n} f(x) e^{-i\xi x}\ dx.$$

Then, $\widehat{f}$ vanishes at infinity.

Theorem 3.21. For $x \geq 1$, we have

$$\frac{\psi_1(x)}{x^2} - \frac{1}{2}\left(1 - \frac{1}{x}\right)^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} h(1+iz) e^{iz\log x}\ dz,$$

where the integral

$$\int_{-\infty}^{\infty} |h(1+iz)|\ dz$$

converges.

Proof. This is actually a continuation of Theorem 3.1. In that particular theorem, we proved that if $c > 1$ and $x \geq 1$, then

$$\frac{\psi_1(x)}{x^2} - \frac{1}{2}\left(1 - \frac{1}{x}\right)^2 = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} x^{z-1} h(z)\ dz,$$

where

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

To prove Theorem 4.2, we first need to show that we can move the path of integration to the line $\sigma = 1$. Let R be the rectangle bounded by the vertices $(1 - iT), (c - iT), (c + iT), (1 + iT)$ such that the path is traversing anticlockwise. By the Cauchy-Goursat Theorem,

$$\oint_R x^{s-1} h(s)\ ds = 0$$

since the integrand is analytic inside and on R. Now, we prove that the integrals along the horizontal segments tend to zero as $T \to \infty$. Since each integral has the same absolute value at the conjugate points, it suffices to consider only the upper segment $t = T$. On this segment, we have the estimates

$$\left|\frac{1}{s(s+1)}\right| \leq \frac{1}{T^2} \text{ and } \left|\frac{1}{s(s+1)(s-1)}\right| \leq \frac{1}{T^3} \leq \frac{1}{T^2}.$$

We now apply Theorem 3.9. There exists a positive constant M such that $|\zeta'(s)/\zeta(s)| \leq M(\log t)^9$ if $\sigma \geq 1$ and $t \geq e$. Hence, if $T \geq e$, we have

$$|h(s)| \leq \frac{M(\log T)^9}{T^2}.$$

Consequently, it is easy to show that

$$\left|\int_1^c x^{s-1}h(s)\ ds\right| \leq \frac{Mx^{c-1}(c-1)(\log T)^9}{T^2}.$$

Hence,

$$\int_{c-i\infty}^{c+i\infty} x^{s-1}h(s)\ ds = \int_{1-i\infty}^{1+i\infty} x^{s-1}h(s)\ ds.$$

On the line $\sigma = 1$, we write $s = 1 + it$ to obtain

$$\frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} x^{s-1}h(s)\ ds = \frac{1}{2\pi}\int_{-\infty}^{\infty} h(1+it)e^{it\log x}\ dt.$$

Since t is a dummy variable, we can replace it with z. It is easy to see that

$$\int_e^{\infty} |h(1+iz)|\ dz$$

converges. The integral from $-\infty$ to $-e$ converges too, so the integral over $\mathbb{R}$ converges. We can then use the Riemann-Lebesgue Lemma (Lemma 4.1) to then establish that $\psi_1(x) \sim x^2/2$. □

Lemma 3.10. $\psi_1(x) \sim x^2/2$

We managed to prove Lemma 4.2 at the end of the proof of Theorem 4.2. As such, we have proven the Prime Number Theorem!

Chapter 4

Dirichlet Series

4.1 Introduction to Dirichlet Series

Definition 4.1 (Dirichlet series). Let f be an arithmetic function. Then, a Dirichlet series is a series of the form

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} \quad \text{where } s = \sigma + it.$$

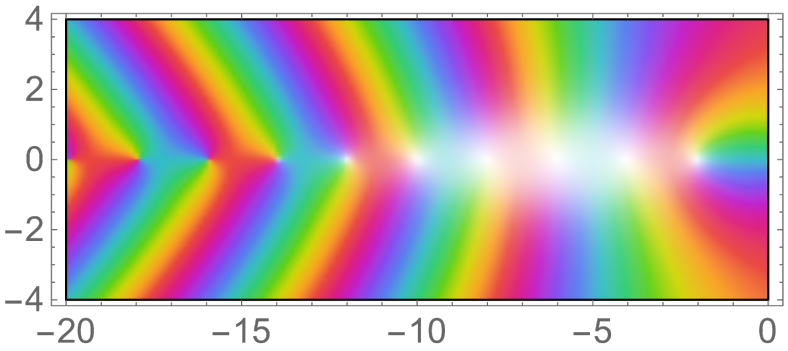

Figure 17: Contour plot of $\dfrac{1}{\zeta(s)} = \displaystyle\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}$

4.2 Multiplication of Dirichlet Series

4.3 Conditional Convergence of Dirichlet Series

4.4 Landau's Theorem

Chapter 5

Dirichlet's Theorem

5.1 Dirichlet Characters

5.2 Proof of Dirichlet's Theorem in Arithmetic Progressions

Theorem 5.1 (Dirichlet's theorem). Suppose $k, l \in \mathbb{N}$ and $k > 1$ such that $\gcd(k, l) = 1$. Then,

$$\sum_{\substack{p \le x \\ p \equiv l \pmod{k}}} \frac{1}{p} = \frac{\log(\log x)}{\varphi(k)} + \mathcal{O}(1). \tag{5.1}$$

Note that Dirichlet's theorem can be rephrased as follows: for any positive integers k and l such that $k > 1$ and k and l are coprime, then there are infinitely many primes of the form $kn + l$.

Definition 5.1. A Dirichlet character modulo k, denoted by χ, is an arithmetical function $\chi : \mathbb{N} \to \mathbb{C}$ satisfying the following:

(i) $\chi(mn) = \chi(m)\chi(n)$ for all $m, n \in \mathbb{N}$ (i.e. χ is completely multiplicative)

(ii) $|\chi(n)| = \begin{cases} 1 & \text{if } \gcd(n, k) = 1; \\ 0 & \text{otherwise} \end{cases}$

(iii) $\chi(n + km) = \chi(n)$ for all $m, n \in \mathbb{N}$

Definition 5.2. Let $\tilde{\chi} : (\mathbb{Z}/k\mathbb{Z})^* \to \{z \in \mathbb{C} : |z| = 1\}$.

In fact, $\tilde{\chi}$ is a group homomorphism that allows us to obtain a character χ modulo k. There are also precisely $\varphi(k)$ homomorphisms from $(\mathbb{Z}/k\mathbb{Z})^*$ to $\{z \in \mathbb{C} : |z| = 1\}$, which shows that there are precisely $\varphi(k)$ characters modulo k.

Definition 5.3. The principal character modulo k is denoted by χ_0. We have

$$\chi_0(n) = \begin{cases} 1 & \text{if } \gcd(n, k) = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{5.2}$$

Definition 5.4. Let $\overline{\chi}$ denote the inverse of χ.

The next part deals with the orthogonal relations.

Theorem 5.2 (Schur orthogonality relations). We have the following:

(i) Let χ_1, χ_2 be two Dirichlet characters modulo k. Then,

$$\sum_{a=1}^{k} \chi_1(a)\overline{\chi_2(a)} = \begin{cases} \varphi(k) & \text{if } \chi_1 = \chi_2 \\ 0 & \text{otherwise.} \end{cases} \tag{5.3}$$

(ii) Let $a_1, a_2 \in \mathbb{Z}$ such that $\gcd(a_j, k) = 1$ for $j = 1, 2$. Then,

$$\sum_{\chi \pmod k} \chi(a_1)\overline{\chi(a_2)} = \begin{cases} \varphi(k) & \text{if } a_1 \equiv a_2 \pmod k \\ 0 & \text{otherwise.} \end{cases} \tag{5.4}$$

Definition 5.5. The L-series is an example of a Dirichlet series, and it is defined to be

$$L(s, \chi) = \sum_{n \geq 1} \frac{\chi(n)}{n^s}, \quad \sigma > 1. \tag{5.5}$$

Theorem 5.3.

(i) If $\chi = \chi_0$, then $L(s, \chi)$ can be analytically continued to the half-plane $\chi > 0$, with the exception of the point $s = 1$ where it has a simple pole such that $\operatorname{Res}(1, \chi_0) = \varphi(k)/k$.

(ii) If $\chi \neq \chi_0$, then $L(s, \chi)$ can be analytically continued to $\sigma > 0$.

We need some lemmas before proving Dirichlet's theorem.

Lemma 5.1 (Merten's estimate). For any $x \geq 1$,

$$\sum_{p \leq x} \frac{\log p}{p} = \log x + \mathcal{O}(1). \tag{5.6}$$

We now prove Dirichlet's theorem.

Proof. It suffices to show that if $x \geq 3$ and $\sigma = 1 + \dfrac{1}{\log x}$, then

$$\sum_{p \equiv l \pmod k} \frac{1}{p^\sigma} = \frac{1}{\varphi(k)} \log\left(\frac{1}{\sigma - 1}\right) + \mathcal{O}(1), \tag{5.7}$$

where the sum on the LHS ranges over all p that satisfy the congruence.

Let

$$S_1 = \sum_{p \equiv l \pmod k} \frac{1}{p^\sigma} \quad \text{and} \quad S_2 = \sum_{\substack{p \leq x \\ p \equiv l \pmod k}} \frac{1}{p}. \tag{5.8}$$

Then,

$$|S_1 - S_2| = \left| \sum_{p \equiv l \pmod{k}} \frac{1}{p^\sigma} - \sum_{\substack{p \leq x \\ p \equiv l \pmod{k}}} \frac{1}{p} \right| \tag{5.9}$$

$$\leq \left| \sum_{p} \frac{1}{p^\sigma} - \sum_{p \leq x} \frac{1}{p} \right| \tag{5.10}$$

$$= \left| \sum_{p \leq x} \left(\frac{1}{p^\sigma} - \frac{1}{p} \right) + \sum_{p > x} \frac{1}{p^\sigma} \right| \tag{5.11}$$

$$\leq \sum_{p \leq x} \left(\frac{1}{p} - \frac{1}{p^\sigma} \right) + \sum_{p > x} \frac{1}{p^\sigma} \quad \text{by the triangle inequality} \tag{5.12}$$

Now, let

$$S_3 = \sum_{p \leq x} \left(\frac{1}{p} - \frac{1}{p^\sigma} \right) \quad \text{and} \quad S_4 = \sum_{p > x} \frac{1}{p^\sigma}. \tag{5.13}$$

We claim that

$$S_3 \leq \mathcal{O}(1) \quad \text{and} \quad S_4 = \mathcal{O}(1). \tag{5.14}$$

First, note that

$$S_3 = \sum_{p \leq x} \frac{1 - \exp\left[-(\sigma - 1)\log p\right]}{p} \tag{5.15}$$

$$\leq \sum_{p \leq x} \frac{(\sigma - 1)\log p}{p} \quad \text{since } e^x \geq 1 + x \text{ by series expansion} \tag{5.16}$$

$$= \frac{1}{\log x} \sum_{p \leq x} \frac{\log p}{p} \tag{5.17}$$

$$= \mathcal{O}(1) \quad \text{by Lemma 4.1} \tag{5.18}$$

Next,

$$S_4 = \sum_{p>x} \frac{1}{p^\sigma} \tag{5.19}$$

$$= \lim_{y\to\infty} \sum_{x\le p\le y} \frac{1}{p^\sigma} \tag{5.20}$$

$$= \lim_{y\to\infty} \left(\frac{1}{y^\sigma} \sum_{p\le y} 1 - \frac{1}{x^\sigma} \sum_{p\le x} 1 + \sigma \int_x^y \frac{1}{t^{\sigma+1}} \sum_{p\le t} 1 \; dt \right) \quad \text{by the Abel summation formula} \tag{5.21}$$

$$= \lim_{y\to\infty} \frac{1}{y^{\sigma-1}\log y} - \frac{1}{x^{\sigma-1}\log x} + \sigma \int_x^\infty \frac{1}{t^\sigma \log t} \; dt \quad \text{by the prime number theorem} \tag{5.22}$$

$$= \mathcal{O}(1) + \mathcal{O}\left(\frac{1}{\log x} \int_x^\infty \frac{1}{t^\sigma} \; dt \right) \tag{5.23}$$

$$= \mathcal{O}(1) \tag{5.24}$$

Since $\gcd(p, l) = 1$, for $\sigma > 1$, we have

$$S_1 = \sum_{p\equiv l \pmod k} \frac{1}{p^\sigma} = \sum_p \frac{1}{p^\sigma} \left[\frac{1}{\varphi(k)} \sum_{\chi \pmod k} \overline{\chi(l)}\chi(p) \right] \text{ by (4)} \tag{5.25}$$

$$= \frac{1}{\varphi(k)} \sum_p \frac{\chi(p)}{p^\sigma} \sum_{\chi \pmod k} \overline{\chi(l)} \tag{5.26}$$

Observe that $\sum_p \frac{\chi(p)}{p^\sigma}$ is an L-series so we shall denote it by $S(\sigma, \chi)$. Thus, we can write (26) as

$$\frac{1}{\varphi(k)} \sum_{\chi \pmod k} \overline{\chi(l)} S(\sigma, \chi). \tag{5.27}$$

Next, we show that

$$S(\sigma, \chi_0) = \log\left(\frac{1}{\sigma - 1} \right) + \mathcal{O}(1). \tag{5.28}$$

Observe that

$$S(\sigma,\chi_0) = \sum_p \sum_{m\geq 1} \frac{1}{mp^{m\sigma}} - \sum_{p|k} \sum_{m\geq 1} \frac{1}{mp^{m\sigma}} \tag{5.29}$$

$$= -\sum_p \log\left(1-\frac{1}{p^\sigma}\right) + \mathcal{O}(1) \quad \text{by series expansion} \tag{5.30}$$

$$= \log\left[\prod_p \left(1-\frac{1}{p^\sigma}\right)^{-1}\right] + \mathcal{O}(1) \tag{5.31}$$

$$= \log\zeta(\sigma) + \mathcal{O}(1) \quad \text{by the Euler product} \tag{5.32}$$

Next, to show that

$$\lim_{\sigma\to 1^+} (\sigma-1)\,\zeta(\sigma) = 1, \tag{5.33}$$

one can consider a corollary of Chapter 6 Exercise 16 in Walter Rudin's 'Principles of Mathematical Analysis' **?**, for which we have

$$(s-1)\zeta(s) = s - s(s-1)\int_1^\infty \frac{x-\lfloor x\rfloor}{x^{s+1}}\,dx. \tag{5.34}$$

To deal with the floor function in the integrand, simply consider a disjoint union of the intervals $[n, n+1)$, where $n \in \mathbb{N}$.

So for all σ in a sufficiently small neighbourhood of 1, we can write

$$\zeta(\sigma) = \frac{1}{\sigma-1} + g(\sigma), \tag{5.35}$$

where g is a function analytic at $\sigma = 1$. We thus see that the principal character contributes to (7), so it remains to show that $S(\sigma,\chi) = \mathcal{O}(1)$ for all $\sigma > 1$ and all non-principal characters $\chi \pmod k$.

It is clear by the steps used from (29) to (32) that

$$S(\sigma,\chi) = \ln(L(\sigma,\chi)) + \mathcal{O}(1). \tag{5.36}$$

For all non-principal characters χ, $L(s,\chi)$ is analytic for all $\sigma > 0$. So, $L(\sigma,\chi)$ is continuous for all $\sigma > 1$. As such,

$$\lim_{\sigma\to 1} L(\sigma,\chi) = L(1,\chi). \tag{5.37}$$

It remains to show $L(1,\chi) \neq 0$. We consider two cases, namely when χ (taken to be non-principal) is complex, and χ is real.

- **Case 1:** Suppose $\chi \neq \chi_0$ modulo k is complex. Let

$$P(\sigma) = \prod_{\chi \pmod k} L(\sigma, \chi). \tag{5.38}$$

For $\sigma > 1$, we have

$$\begin{aligned}
\log(P(\sigma)) &= \sum_{\chi \pmod k} \log(L(\sigma, \chi)) && (5.39)\\
&= \sum_{\chi \pmod k} \sum_{p} \sum_{m \geq 1} \frac{\chi(p^m)}{m p^{m\sigma}} \quad \text{by series expansion} && (5.40)\\
&= \sum_{p} \sum_{m \geq 1} \frac{1}{m p^{m\sigma}} \sum_{\chi \pmod k} \chi(p^m)\overline{\chi(1)} \quad \text{since } \chi(1) = \overline{\chi(1)} = 1 && (5.41)\\
&= \sum_{p} \sum_{\substack{m \geq 1 \\ p^m \equiv 1 \pmod k}} \frac{1}{m p^{m\sigma}} \quad \text{by (4)} && (5.42)\\
&\geq 0 && (5.43)
\end{aligned}$$

So we infer that for $\sigma > 1$, we have $P(\sigma) \geq 1$. Suppose on the contrary that $L(1, \chi) = 0$ for some character χ. Then, $L(1, \overline{\chi}) = 0$, which implies that P has two zeros at $\sigma = 1$. But we know that $L(\sigma, \chi_0)$ has a simple pole at $\sigma = 1$ as derived in (35). This implies that $P(1) = 0$, which contradicts $P(\sigma) \geq 1$. As such, $L(1, \chi) \neq 0$ when $\chi \neq \chi_0$ modulo k is complex.

- **Case 2:** Suppose $\chi \neq \chi_0$ modulo k is real. Consider $f = \chi * \mathbf{1}$, for which it is clear that $\mathbf{1}$ is an arithmetical function and f is the Dirichlet product of χ and $\mathbf{1}$.

Note that f is multiplicative. To see why, write

$$f(n) = \sum_{d|n} \chi(d). \tag{5.44}$$

From **(i)** of Definition 2.1, it is clear that $\chi(1) = [\chi(1)]^2$. By **(ii)**, since $\chi(1) \neq 0$, we have $\chi(1) = 1$. As such, $f(1) = 1$. Now, consider $m, n \in \mathbb{N}$ such that $\gcd(m, n) = 1$. Then, if $d|mn$, there exists $d_1, d_2 \in \mathbb{N}$ such that $d = d_1 d_2$ satisfying $d_1|m$ and $d_2|n$.

Hence,

$$f(mn) = \sum_{d|mn} \chi(d) \tag{5.45}$$

$$= \sum_{d_1|m} \sum_{d_2|n} \chi(d_1)\chi(d_2) \quad \text{by } \textbf{(i)} \text{ of Definition 2.1} \tag{5.46}$$

$$= \left(\sum_{d_1|m} \chi(d_1) \right) \left(\sum_{d_2|m} \chi(d_2) \right) \tag{5.47}$$

$$= f(m)f(n) \tag{5.48}$$

which shows that f is multiplicative.

Note that

$$\sum_{l=0}^{m} \chi\left(p^l\right) = \sum_{l=0}^{m} (\chi(p))^l \quad \text{since } \chi \text{ is multiplicative} \tag{5.49}$$

$$\begin{cases} = 1 & \text{if } p|k \\ \geq 1 & \text{if } p \text{ does not divide } k \text{ and } m \text{ is even} \\ 0 & \text{if } p \text{ does not divide } k \text{ and } m \text{ is odd} \end{cases} \tag{5.50}$$

So, $f(n) \geq 0$ for all $n \in \mathbb{N}$ and in particular, $f(n) \geq 1$ if n is square.

Let

$$F(\sigma) = \sum_{n \geq 1} \frac{f(n)}{n^\sigma}. \tag{5.51}$$

Using the fact that $f(n) \geq 1$ if n is square, it is easy to see that

$$F(\sigma) \geq \sum_{n \geq 1} \frac{1}{n^{2\sigma}} = \zeta(2\sigma). \tag{5.52}$$

F diverges when $\sigma = 1/2$, and so its abscissa of convergence, σ_c, is $\geq 1/2$. So, $F(s)$ must have a point of singularity at $s = \sigma_c \geq 1/2$.

On the other hand, as

$$F(s) = \sum_{n \geq 1} \frac{\chi(n)}{n^s} \sum_{n \geq 1} \frac{1}{n^s} = L(s, \chi)\zeta(s), \tag{5.53}$$

then if $L(1, \chi) = 0$, F would be analytic for $\sigma > 0$ and because $\sigma_c \geq 1/2$, it implies that F is analytic at $\sigma = \sigma_c$. This contradicts the earlier claim that F has a point of singularity at σ_c, so it follows that $L(1, \chi) \neq 0$ for all real characters $\chi \neq \chi_0$ modulo k.

We have thus proven Dirichlet's theorem. □

The proof of Dirichlet's theorem hinges on the fact that $L(1,\chi) \neq 0$ for all non-principal characters χ modulo k. The principal character χ_0 contributes to the $\log(\log k)$ term in the numerator of the main term in (1). We also used the second orthogonality relation, namely (4), to invoke the $1/\varphi(k)$.

Going back to (8), (18) and (24), we infer that $|S_1 - S_2|$ differ by a constant. The absolute value sign can be removed, so we are essentially saying that the difference of two series, S_1 and S_2, converges.

Recall the following theorem from Real Analysis:

Theorem 5.4. Let $\{a_n\}_{n=1}^{\infty}, \{b_n\}_{n=1}^{\infty}$ be two sequences of real numbers. If $\sum_{n=1}^{\infty}(a_n - b_n)$ converges, then

$$\sum_{n=1}^{\infty} a_n \text{ and } \sum_{n=1}^{\infty} b_n \text{ converge} \quad \text{or} \quad \sum_{n=1}^{\infty} a_n \text{ and } \sum_{n=1}^{\infty} b_n \text{ diverge.} \tag{5.54}$$

Proof. We prove by contraposition. First, consider the sets

$$X = \left\{ \sum_{n=1}^{\infty} a_n \text{ diverges} \right\} \quad \text{and} \quad Y = \left\{ \sum_{n=1}^{\infty} b_n \text{ converges} \right\}. \tag{5.55}$$

and note that $((X' \cap Y) \cup (X \cap Y'))' = X \cap Y$.

By symmetry, it suffices to show that if

$$\sum_{n=1}^{\infty} a_n \text{ diverges} \quad \text{and} \quad \sum_{n=1}^{\infty} b_n \text{ converges,} \tag{5.56}$$

then $\sum_{n=1}^{\infty}(a_n - b_n)$ diverges. Suppose on the contrary that $\sum_{n=1}^{\infty}(a_n - b_n)$ converges. Then,

$$\sum_{n=1}^{\infty}(a_n - b_n + b_n) \text{ converges} \tag{5.57}$$

because the sum of two convergent series is also convergent. This implies that $\sum_{n=1}^{\infty} a_n$ diverges, which is a contradiction. The result follows. □

We showed that S_1 diverges, and we wanted to prove that S_2 diverges (recall that S_1 is the LHS of (7) and S_2 is the LHS of (1)). By Theorem 5.1, S_2 diverges so S_1

diverges too, asserting the infinitude of primes of the form $kn + l$, where $k, l \in \mathbb{N}$, $k > 1$ and $\gcd(k, l) = 1$.

In fact, we adopted the clever choice of $\sigma = 1 + 1/\log x$, where $x \geq 3$. So, for $\sigma \geq 1.92$ (note that $1 + 1/\log 3 \approx 1.9102$), Dirichlet's theorem follows too.

Some texts would refer to Dirichlet's theorem as the following statement. For example, one can turn to Theorem 7.3 of Tom Apostol's book 'Introduction to Analytic Number Theory' **?**. We modify the symbols too.

Theorem 5.5 (Dirichlet's theorem). Suppose $k > 0$ and $\gcd(k, l) = 1$. Then, for all $x \geq 1$, we have

$$\sum_{\substack{p \leq x \\ p \equiv l \pmod{k}}} \frac{\log p}{p} = \frac{\log x}{\varphi(k)} + \mathcal{O}(1). \tag{5.58}$$

In fact, Theorem 5.2 implies Theorem 1.1 by the Abel summation formula.

Chapter 6

Sieve Methods

6.1 The Sieve of Eratosthenes

6.2 The Large Sieve

6.3 Brun's Sieve and Twin Primes

The Brun sieve **?** is a method in Analytic Number Theory for estimating the density of sets of integers that satisfy certain congruences. It uses sieving techniques to remove integers not meeting these criteria, allowing for approximating the distribution of primes and almost-primes within specific intervals.

Definition 6.1. Let $\mu(n)$ denote the Möbius function, where $n \in \mathbb{N}$. We define it as follows:

$$\mu(n) = \begin{cases} 1 & \text{if } n \text{ is square-free with an even number of prime factors;} \\ -1 & \text{if } n \text{ is square-free with an odd number of prime factors;} \\ 0 & \text{if } n \text{ is divisible by some square.} \end{cases} \tag{6.1}$$

Theorem 6.1. For any $n \in \mathbb{N}$, we have

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1; \\ 0 & \text{otherwise.} \end{cases} \tag{6.2}$$

We will also denote $\sum_{d|n} \mu(d)$ by $E(n)$, so $E(n)$ behaves like an indicator function.

Proof. When $n = 1$, the result is clear because n is not divisible by any square (excluding 1 by definition) and 1 has no prime factors. Next, suppose $n > 1$. Write $n = \prod_{i=1}^{r} p_i^{\alpha_i}$, where $\alpha_i \geq 0$ but the α_i's cannot be all non-zero. Define $m = \prod_{i=1}^{r} p_i$. By considering m,

$$\sum_{d|n} \mu(d) = 1 - \binom{r}{1} + \binom{r}{2} + \ldots + (-1)^r \binom{r}{r} = 0.$$

□

In Brun's sieve, we intend to introduce two functions μ_1 and μ_2 such that

$$\mu_1 * \mathbf{1} \leq E \leq \mu_2 * \mathbf{1}. \tag{6.3}$$

In fact, the sieve of Eratosthenes is based on the identity

$$\mu * \mathbf{1} = E. \tag{6.4}$$

Recall that these are simply Dirichlet products, or rather, convolutions, defined as follows:

Definition 6.2. Let f and g be arithmetic functions. Then,

$$(f * g)(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right) = \sum_{d|n} f\left(\frac{n}{d}\right) g(d). \tag{6.5}$$

Obviously, the identity function $\mathbf{1}$ is an arithmetic function. One should also recall that μ is an arithmetic function.

The issue with the sieve of Eratosthenes is that it is not an efficient method to *sieve* primes. We have the following corollary as a result of it:

Corollary 6.1. By the sieve of Eratosthenes,

$$\pi(x) \leq \frac{x}{\log(\log x)}. \tag{6.6}$$

On the other hand, Brun's sieve is more efficient.

Corollary 6.2. By Brun's sieve,

$$\pi(x) \leq \frac{x \log(\log x)}{\log x}. \tag{6.7}$$

Remark 6.1. Let $f(x) = \dfrac{x}{\log(\log x)}$ and $g(x) = \dfrac{x \log(\log x)}{\log x}$. Then, for $x \geq e$, $g(x) \leq f(x)$.

Proof. Trivial by considering $u = \log x$. □

Theorem 6.2. Let $\chi_t = \{n \in \mathbb{Z} : \omega(n) \leq t\}$. Recall that $\omega(n)$ counts the number of distinct prime factors of n. Then, for all $h \in \mathbb{Z}_{\geq 0}$,

$$\mu_i(n) = \mu(n) \chi_{2h+2-i}(n) \tag{6.8}$$

satisfies (3) for $i = 1, 2$.

Corollary 6.3. Let $\mathcal{A}$ be a finite set of integers and $\mathcal{P}$ be a set of prime numbers. Define

$$\mathcal{A}_d = |\{a \in \mathcal{A} : a \equiv 0 \pmod d\}| \tag{6.9}$$

$$P(y) = \prod_{\substack{p \leq y \\ p \in \mathcal{P}}} p \tag{6.10}$$

$$\mathcal{S}(\mathcal{A}, \mathcal{P}, y) = |\{a \in \mathcal{A} : \gcd(a, P(y)) = 1\}| \tag{6.11}$$

Then, for all $h \in \mathbb{Z}_{\geq 0}$,

$$\sum_{\substack{d|P(y) \\ \omega(d)\leq 2h+1}} \mu(d)\mathcal{A}_d \leq \mathcal{S}(\mathcal{A},\mathcal{P},y) \leq \sum_{\substack{d|P(y) \\ \omega(d)\leq 2h}} \mu(d)\mathcal{A}_d. \tag{6.12}$$

We now provide a proof of Corollary 1.2.

Proof. From Corollary 1.3, set

$$\mathcal{A} = \{n \in \mathbb{Z} : n \leq x\} \tag{6.13}$$

$$\mathcal{P} = \text{all primes} \tag{6.14}$$

$$P = P(y) = \prod_{p\leq y} p \tag{6.15}$$

Thus, $\mathcal{S}(\mathcal{A},\mathcal{P},y)$ denotes the cardinality of the set of positive integers n which are $\leq x$ such that $\gcd(n,p) = 1$ for all primes $p \leq y$.

Consider the upper bound of Corollary 1.3. We have

$$\mathcal{S}(\mathcal{A},\mathcal{P},y) \leq \sum_{\substack{d|P(y) \\ \omega(d)\leq 2h}} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor. \tag{6.16}$$

Using

$$\left\lfloor \frac{x}{d} \right\rfloor = \frac{x}{d} + \mathcal{O}(1), \tag{6.17}$$

the RHS of (16) can be written as

$$x \sum_{\substack{d|P(y) \\ \omega(d)\leq 2h}} \frac{\mu(d)}{d} + \mathcal{O}\left(\sum_{\substack{d|P(y) \\ \omega(d)\leq 2h}} 1 \right). \tag{6.18}$$

Recall that

$$\sum_{d|n} \frac{\mu(d)}{d} = \prod_{p|n} \left(1 - \frac{1}{p}\right), \tag{6.19}$$

which is easily justified by the definition of $\zeta(s)$ for $s = 1$.

Since

$$\sum_{\substack{d|P(y) \\ \omega(d)\leq 2h}} \frac{\mu(d)}{d} = \sum_{d|P(y)} \frac{\mu(d)}{d} - \sum_{\substack{d|P(y) \\ \omega(d)>2h}} \frac{\mu(d)}{d} \tag{6.20}$$

and noting that $\mu(d)/d \leq 1/d$, it follows that (18) can be written as

$$x \sum_{p \leq y} \left(1 - \frac{1}{p}\right) + \mathcal{O}\left(\sum_{\substack{d|P(y) \\ \omega(d) \leq 2h}} 1 + x \sum_{\substack{d|P(y) \\ \omega(d) > 2h}} \frac{1}{d} \right). \tag{6.21}$$

In a similar fashion, the lower bound for $\mathcal{S}(\mathcal{A}, \mathcal{P}, y)$ can be written as

$$\mathcal{S}(\mathcal{A}, \mathcal{P}, y) \geq x \prod_{p \leq y} \left(1 - \frac{1}{p}\right) + \mathcal{O}\left(\sum_{\substack{d|P(y) \\ \omega(d) \leq 2h+1}} 1 + x \sum_{\substack{d|P(y) \\ \omega(d) > 2h+1}} \frac{1}{d} \right) \tag{6.22}$$

Since $P(y)$ is the product of all p which are $\leq y$, by considering the extreme case, we obtain the following upper bound for the sum in the error term in (21) and (22):

$$\sum_{\substack{d|P(y) \\ \omega(d) \leq 2h}} 1, \quad \sum_{\substack{d|P(y) \\ \omega(d) \leq 2h+1}} 1 \leq y^{2h+1} \tag{6.23}$$

For the second error term in (21), by the multinomial theorem,

$$\sum_{\substack{d|P(y) \\ \omega(d) > 2h}} \frac{1}{d} \leq \sum_{k > 2h} \frac{1}{k!} \left(\sum_{p \leq y} \frac{1}{p} \right)^k. \tag{6.24}$$

By Merten's estimate (Theorem 1.3), we can further bound (24) to obtain

$$\sum_{k > 2h} \frac{1}{k!} \left(\sum_{p \leq y} \frac{1}{p} \right)^k \leq \sum_{k > 2h} \frac{(\log(\log y) + c)^k}{k!}, \tag{6.25}$$

where $c = \beta + \mathcal{O}\left(\frac{1}{\log x}\right)$. We will define β in Theorem 1.3.

By Stirling's approximation,

$$\sum_{k > 2h} \frac{(\log(\log y) + c)^k}{k!} \leq \frac{1}{\sqrt{2\pi}} \sum_{k > 2h} \frac{(e \log(\log y) + ce)^k}{k^{k+1/2} \exp\left(\frac{1}{12k+1}\right)} \tag{6.26}$$

$$\leq \sum_{k > 2h} \left(\frac{e \log(\log y) + ce}{k} \right)^k \tag{6.27}$$

For all $h \geq e \log(\log y) + ce$, we conclude that

$$\sum_{\substack{d|P(y) \\ \omega(d) > 2h}} \frac{1}{d} \leq \left(\frac{1}{2}\right)^{2h} \sum_{r=0}^{\infty} \frac{1}{2^r} = 2^{1-2h} \leq 2^{1-2e \log(\log x) - ce} \leq \frac{1}{(\log x)^2}. \tag{6.28}$$

With this choice of h, we have $y^{2h+1} \leq \dfrac{x}{(\log x)^2}$. In fact, let $y = \exp\left(\dfrac{\log x}{10\log(\log x)}\right)$.

We use Corollary 1.4 (consequence of Merten's estimate) and the squeeze theorem to conclude that

$$\mathcal{S}(\mathcal{A}, \mathcal{P}, y) = \frac{xe^{-\gamma}}{\log y} + \frac{xe^{-\gamma}}{\log y}\mathcal{O}\left(\frac{1}{\log y}\right) + \mathcal{O}\left(\frac{x}{(\log x)^2}\right) \tag{6.29}$$

$$= \frac{10xe^{-\gamma}\log(\log x)}{\log x} + \mathcal{O}\left(\frac{x}{(\log x)^2}\right) \tag{6.30}$$

$$\sim \frac{x\log(\log x)}{\log x} \tag{6.31}$$

The result follows by considering $\mathcal{S}(\mathcal{A}, \mathcal{P}, y) \geq \pi(x) - \pi(y) \geq \pi(x) - y \geq \pi(x) + \mathcal{O}(\sqrt{x})$ and by our choice of y earlier. □

Theorem 6.3 (Merten's estimate). ? There exists $\beta \in \mathbb{R}$ such that

$$\sum_{p \leq x} \frac{1}{p} = \log(\log x) + \beta + \mathcal{O}\left(\frac{1}{\log x}\right), \tag{6.32}$$

where

$$\beta = 1 - \log(\log 2) + \int_2^\infty \frac{1}{t(\log t)^2}\, dt. \tag{6.33}$$

Proof. We use the Abel summation formula in our proof. Recall that if $a(n)$ is an arithmetic function and

$$A(x) = \sum_{n \leq x} a(n), \tag{6.34}$$

if $0 \leq y < x$ and f is a real-valued function with a continuous derivative on $[y, x]$, then

$$\sum_{y \leq n < x} a(n)f(n) = f(x)A(x) - f(y)A(y) - \int_y^x A(t)f'(t)\, dt. \tag{6.35}$$

Set

$$a(n) = \begin{cases} \log p/p & \text{if } n = p; \\ 0 & \text{otherwise} \end{cases} \quad \text{and } f(t) = \frac{1}{\log t}. \tag{6.36}$$

It is clear that

$$A(x) = \sum_{n \leq x} a(n) = \sum_{p \leq x} \frac{\log p}{p} = \log x + \mathcal{O}(1). \tag{6.37}$$

It follows that

$$\sum_{p\le x}\frac{1}{p} = 1+\frac{\mathcal{O}(1)}{\log x}+\int_2^x \frac{1}{t\log t}+\mathcal{O}(1)\int_2^x \frac{1}{t(\log t)^2}\,dt \tag{6.38}$$

$$= 1+\frac{\mathcal{O}(1)}{\log x}+\int_2^x \frac{1}{t\log t}+\mathcal{O}(1)\int_2^\infty \frac{1}{t(\log t)^2}\,dt-\mathcal{O}(1)\int_x^\infty \frac{1}{t(\log t)^2}\,dt \tag{6.39}$$

$$= 1+\frac{\mathcal{O}(1)}{\log x}+\log(\log x)-\log(\log 2)+\mathcal{O}(1)\left(\frac{1}{\log 2}-\frac{1}{\log x}\right) \tag{6.40}$$

and the result follows. □

Corollary 6.4. As a result of Theorem 1.3,

$$\prod_{p\le x}\left(1-\frac{1}{p}\right) = \frac{e^{-\gamma}}{\log x}\left(1+\mathcal{O}\left(\frac{1}{\log x}\right)\right), \tag{6.41}$$

where γ is the Euler-Mascheroni constant.

Proposition 6.1. ? Let $\mathcal{J} = \{p : p \text{ and } p+2 \text{ are primes}\}$ and $\mathcal{J}(x) = |\{p\in\mathcal{J} : p\le x\}|$. Then,

$$\mathcal{J}(x) \le \frac{x\,(\log(\log x))^2}{(\log x)^2}. \tag{6.42}$$

We omit the proof as it is not important. Anyway, the proof hinges on Corollary 1.3 by setting $\mathcal{A} = \{n(n+2) : n\le x\}$.

Theorem 6.4 (Brun's theorem). The sum of reciprocals of twin primes converges. That is, the sum

$$\sum_{p,p+2 \text{ primes}}\frac{1}{p} \tag{6.43}$$

converges.

It is unknown whether there are infinitely many twin primes but we know that the sum of reciprocals of twin primes is convergent by Brun's theorem.

We now prove Brun's theorem.

Proof. Observe that

$$\mathcal{J}(n)-\mathcal{J}(n-1) = \begin{cases} 1 & \text{if } p \text{ and } p+2 \text{ are primes;} \\ 0 & \text{otherwise.}\end{cases} \tag{6.44}$$

So,

$$\sum_{p,p+2 \text{ primes}} \frac{1}{p} = \sum_{n=2}^{\infty} \frac{\mathcal{J}(n) - \mathcal{J}(n-1)}{n} \tag{6.45}$$

$$= \sum_{n=1}^{\infty} \frac{\mathcal{J}(n)}{n(n+1)} \tag{6.46}$$

$$\leq \sum_{n=2}^{\infty} \frac{(\log(\log n))^2}{(n+1)(\log n)^2} \tag{6.47}$$

$$\leq \sum_{n=2}^{\infty} \frac{(\log(\log n))^2}{n(\log n)^2} \tag{6.48}$$

Let $g(n)$ denote the summand in (48). Consider $\int_2^{\infty} g(t)\ dt$ and substitute $u = \log t$. The integral becomes

$$\int_2^{\infty} \frac{(\log u)^2}{u^2}\ du = 1 + \log 2 + \frac{1}{2}(\log 2)^2 \tag{6.49}$$

using integration by parts. By the integral test, (43) converges. □

Chapter 7

Partitions

7.1 The Rogers-Ramanujan Identities

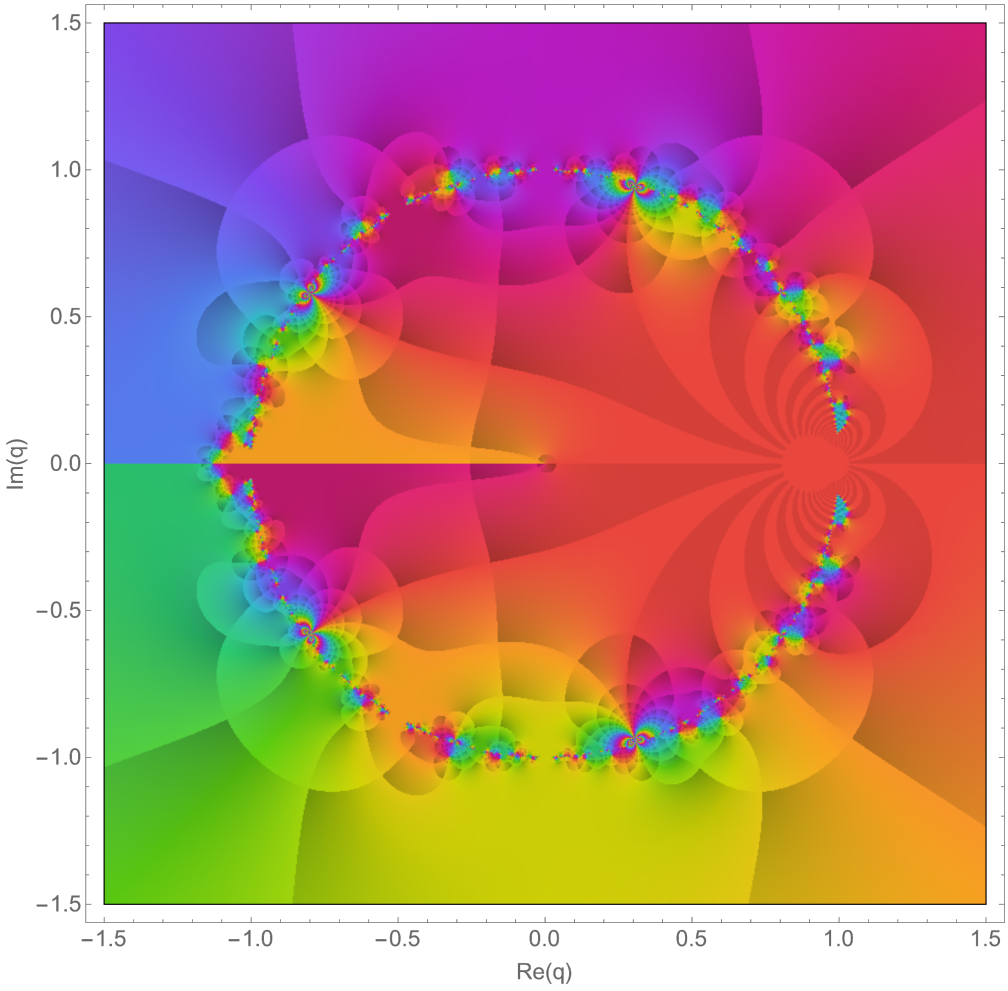

Figure 18: Domain colouring of $R(q)$