MA4263 Introduction to Analytic Number Theory

Notes by Chan Heng Huat

Contents

References

The references are “Introduction to Analytic Number Theory” by T.M. Apostol (First three chapters and several parts of the notes), 1991 Analytic Number Theory notes (Chapter 5 and Chapter 7 of these notes) by A. Hildebrand and “Analytic Number Theory for undergraduates” by H.H. Chan.

1 Fundamental Theorem of Arithmetic

1.1 Least Integer Axiom and Mathematical Induction

Let

$$\mathbf{Z} = \{0, \pm 1, \pm 2, \cdots\}$$

be the set of integers. Let $\mathbf{N}$ denote the set of non-negative integers. The *Least Integer Axiom*, also known as the *Well Ordering Principle*, states that there is a smallest integer in every *nonempty* subset of non-negative integers.

Remark 1.1 One can show that the least integer axiom implies the principle of mathematical induction. Conversely, the principle of mathematical induction implies the least integer axiom.

1.2 The Division Algorithm

THEOREM 1.1 (Division Algorithm) Let a and b be integers such that $b > 0$. Then there exist *unique* integers q and r with

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

Proof
Let

$$S = \{y \in \mathbf{Z} | y = a - bx, \quad x \in \mathbf{Z} \text{ and } y \geq 0\}.$$

Note that since

$$a - b(-|a|) = a + b|a| \geq 0,$$

we find that

$$a + b|a| \in S,$$

and we conclude that S is nonempty. By the Least Integer Axiom, S contains a least non-negative integer, which we denote by r. We note that since $r \in S$,

$$r = a - bq,$$

for some integer q. We therefore conclude that

$$a = bq + r \quad \text{and} \quad r \geq 0.$$

We now show that $r < b$. Suppose $r \geq b$. Then

$$r - b \geq 0 \quad \text{and} \quad r - b = a - b(q + 1).$$

This implies that

$$r - b \in S.$$

By assumption, $b > 0$ and hence $r - b < r$. Hence, we have found a non-negative integer $r - b$ contained in S and smaller than r. This contradicts the minimality of r and we conclude that $r < b$.

Finally, we show that the integers q and r are unique. We suppose the contrary. Then there is a different representation of the form $a = bq' + r'$. This implies that

$$b(q' - q) = r - r' \tag{1.1}$$

and we conclude that $|r - r'|$ is a multiple of b. On the other hand, both $r, r' \in [0, b)$ and $|r - r'|$ can be a multiple of b only when $|r - r'| = 0$. In other words, $r = r'$ and by (1.1), $q = q'$. This contradicts the fact that the representations $a = bq' + r'$ $a = bq + r$ are different and therefore, the integers q and r must be unique. □

When $r = 0$ in Theorem 1.1, we have $a = bq$ and we say that b divides a and we write $b|a$. When $r > 0$, we say that b does not divide a and we write $b \nmid a$.

DEFINITION 1.1 If an integer b divides a, we say that b is a *divisor* of a and that a is a *multiple* of b.

DEFINITION 1.2 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 and n be nonzero integers. The following statements are true:

(a) For all nonzero integers k, $k|k$.

(b) If $d|n$ and $n|m$, then $d|m$.
(c) If $d|n$ and $d|m$, then $d|(an+bm)$.
(d) If $d|n$, then $ad|an$.
(e) If $ad|an$ and $a \neq 0$, then $d|n$.
(f) If $d|n$, then $|d| \leq |n|$.
(g) If $d|n$ and $n|d$, then $|d| = |n|$.
(h) If $d|n$, then $\left(\frac{n}{d}\right)|n$.

Proof
We will prove (c) and leave the rest of the statements as exercises. Since $d|n$, we find that $n = ds$ for some integer s. Similarly, $d|m$ implies that $m = dt$ for some integer t. Now,

$$an + bm = ads + bdt = d(as + bt).$$

This shows that $d|(an+bm)$ for any integers a and b. □

In Theorem 1.2 (h), we see that if d is a divisor of n, then n/d is also a divisor of n. If d is a divisor of n, then we call n/d is called the *conjugate divisor* of d.

We say that a is congruent to b modulo n when $n|(a-b)$. The notation is

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

With this notation, we conclude from Theorem 1.1 that given integers a and $b \geq 1$, there exists a unique r with $0 \leq r < b$ such that

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

THEOREM 1.3 (Basic Properties of Congruences) Let a, b, c, d, n be integers with $n > 0$. Then

(a) For all integers k, $k \equiv k \pmod{n}$.
(b) If $a \equiv b \pmod{n}$ then $b \equiv a \pmod{n}$.
(c) If $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$ then $a \equiv c \pmod{n}$.
(d) If $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$ then $a + c \equiv b + d \pmod{n}$ and $ac \equiv bd \pmod{n}$.

1.3 Greatest common divisors

DEFINITION 1.3 Let a and b be integers for which at least one of them is non-zero. A *common divisor* of integers a and b is an integer c with $c|a$ and $c|b$.

DEFINITION 1.4 A *greatest common divisor* of integers a and b is a number d with the following properties :

(a) The integer d is non-negative.
(b) The integer d is a common divisor of a and b.
(c) If e is any common divisor of a and b, then $e|d$.

The greatest common divisor of two integers (one of which is non-zero) is unique. It is written as

$$(a, b).$$

We will show later that the greatest common divisors of two integers a and b exists.

Remark 1.2 Note that if $b \neq 0$ and $a = 0$ then $|b| = (b, 0)$.

We will next show that the greatest common divisor of two integers exists. By Remark 1.2, it suffices to consider the case when both a and b are nonzero.

THEOREM 1.4 Let a and b be nonzero integers.Then there exists integers m, n such that

$$(a, b) = am + bn.$$

Proof
We give a proof of the above using facts from cyclic groups. We first recall that if G is a cyclic group and H is a subgroup of G, then H is cyclic. To see this, let G be generated by g. Since H is a subgroup of G,

$$H = \{g^{\ell} | \ell \in T\}$$

where T is a subset of $\mathbf{Z}$. Let r be the smallest positive integer in T. The existence of r is guaranteed by the least integer axiom. We claim that H is generated by g^r. Suppose not. Then by Theorem 1.1, there exists $\ell \in T$ such that

$$\ell = rq + s, 0 < s < r.$$

Note that

$$g^{rq} \in H \text{ and } g^{\ell} \in H$$

implies that

$$g^s = g^{\ell - rq} \in H.$$

Hence, $s \in T$ and $0 < s < r$, contradicting the minimality of r. Therefore, H is cyclic. Now, $(\mathbf{Z}, +)$ is a cyclic group generated by 1. The set

$$Q = \{am + bn | m, n \in \mathbf{Z}\}$$

is a subgroup of $\mathbf{Z}$. This can be seen using the subgroup criterion as

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

By the property of cyclic group, we conclude that Q is generated by a positive integer d.

Now, $d = a\alpha + b\beta$ for some integers α and β since $d \in Q$. Note that $a = a + b \cdot 0 \in Q$ and so $a = du$ since Q is generated by d. Therefore $d|a$. Similarly, $d|b$. Therefore d is a common divisor of a and b.

Next, let c be a common divisor of a and b. Write $a = c\nu$ and $b = c\omega$. Then

$$d = a\alpha + b\beta = c(\nu\alpha + \omega\beta)$$

implies that $c|d$. Since $d > 0$, we conclude that $d = (a, b)$ as d satisfies the conditions defining the greatest common divisors of a and b.

□

DEFINITION 1.5 We say that two integers a and b are *relatively prime* if

$$(a, b) = 1.$$

THEOREM 1.5 Let a and b be nonzero integers. Then $(a, b) = 1$ if and only if $1 = ax + by$ for some integers x and y.

Proof

Since $(a, b) = 1$, by Theorem 1.4,

$$1 = ax + by$$

for some integers x and y.

Conversely, if

$$1 = ax + by,$$

then $(a, b)|a$ and $(a, b)|b$, and therefore $(a, b)|1$. This implies that $(a, b) = 1$. □

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

THEOREM 1.6 Let a, b and c be nonzero integers. Then

(a) $(a, b) = (b, a)$

(b) $(a,(b,c)) = ((a,b),c)$ and
(c) $(ac,bc) = |c|(a,b)$.

Proof
We will prove only (c) and leave the proofs of the other statements as exercises. Let $d = (ac,bc)$ and $d' = |c|(a,b)$. By Theorem 1.4,

$$d = acx + bcy$$

for some integers x and y. Hence,

$$d = \frac{c}{|c|}\left(a \cdot |c| \cdot x + b \cdot |c| \cdot y\right). \tag{1.2}$$

Now, $d' = |c|(a,b)$ and since $(a,b)|a$ and $(a,b)|b$, we find that d' is a common divisor of $a \cdot |c|$ and $b \cdot |c|$ and therefore, by (1.2), $d'|d$.

Next, since $d'/|c| = (a,b)$, by Theorem 1.4,

$$\frac{d'}{|c|} = au + bv$$

for some integers u and v. This implies that

$$d' = a \cdot |c| \cdot u + b \cdot |c| \cdot v = \frac{|c|}{c}\left(acu + bcv\right).$$

But d is a common divisor of ac and bc and hence $d|d'$. Since $d'|d$ and $d|d'$, we conclude by Theorem 1.2 (g) that $|d| = |d'|$. Since both d and d' are positive, we deduce that $d = d'$. □

We have seen the definition of the greatest common divisor of two integers a and b. The greatest common divisor of m integers is defined in a similar way. It is a positive integer d which is the divisor of $a_1, \cdots, a_m$ satisfying the property that any common divisor of $a_1, \cdots, a_m$ divides d. The notation for the greatest common divisor of m integers is $(a_1, a_2, \cdots, a_m)$. For example, one can show that

$$(a,b,c) = (a,(b,c)) = ((a,b),c) = ((a,c),b).$$

1.4 The least common multiple

DEFINITION 1.6 The least common multiple of two integers a and b with $b \neq 0$ is defined as an integer m satisfying

(a) m is a positive integer,
(b) $a|m$ and $b|m$,
(c) If $a|\ell$ and $b|\ell$ then $m|\ell$.

The notation for the least common multiple of a and b is $[a, b]$.

An important identity relating (a, b) and $[a, b]$ is

THEOREM 1.7 Let a and b be positive integers. Then

$$ab = [a, b](a, b).$$

Proof
Our first step is to prove that if $(u, v) = 1$, $u|m$ and $v|m$ then $uv|m$. Write $m = u\alpha$ and $n = v\beta$. Note that $1 = (u, v)$ implies that

$$1 = u\nu + v\omega.$$

Hence,

$$m = mu\nu + mv\omega = v\beta u\nu + u\alpha v\omega = uv(\beta\nu + \alpha\omega).$$

Therefore $uv|m$.

For our second step, we show that if $c \in \mathbf{Z}^+$, then

$$c[h, k] = [ch, ck].$$

Let $\ell = [h, k]$ and $\ell' = [ch, ck]$. Now, $h|\ell$ and $k|\ell$, then $ch|c\ell$ and $ck|c\ell$. This implies that $\ell'|c\ell$ or $\ell'/c|\ell$. Next $ch|\ell'$ and $ck|\ell'$. Then $h|(\ell'/c)$ and $k|(\ell'/c)$ and $\ell|(\ell'/c)$. Therefore, $\ell = \ell'/c$ or $c\ell = \ell'$.

Now, suppose that $d = (a, b)$. Then

$$1 = (a/d, b/d),$$

since $c(a, b) = (ca, cb)$ for $c > 0$ (see (c)). By the first step,

$$[a/d, b/d] = ab/d^2.$$

Therefore,

$$ab = d(d[a/d, b/d]) = d[a, b]$$

by the second step.

□

1.5 Euclid's Lemma

We know that if $c \neq 0$ then $ca = cb$ implies that $a = b$. This is known as the law of cancelation for equality. The law is not true in general if we replace "=" by $\equiv$. For example, $15 \equiv 3 \pmod{12}$ but $5 \not\equiv 1 \pmod{12}$. The next result shows that the law of cancelation holds if we impose a condition on the integer c.

THEOREM 1.8 Let a, b, c and n be integers. If $ca \equiv cb \pmod{n}$ and $(c, n) = 1$, then $a \equiv b \pmod{n}$.

Proof
Since $(c, n) = 1$ there exist integers x and y such that $cx + ny = 1$. Multiplying a and b yields

$$acx + any = a$$

and

$$bcx + bny = b,$$

respectively. Since $ac \equiv bc \pmod{n}$, we conclude that $a - b \equiv (ac - bc)x \equiv 0 \pmod{n}$ and hence,

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

□

Theorem 1.8 can be used to prove the following result of Euclid.

COROLLARY 1.9 Let a and b be integers and p be a prime. If $p|(ab)$, then $p|a$ or $p|b$.

Proof
For any integer n, $(n, p) = 1$ or p since p has only two divisors. Suppose $p \nmid a$. Then $(p, a) = 1$. By Theorem 1.8, the relation

$$ab \equiv 0 \pmod{p}$$

then implies that

$$b \equiv 0 \pmod{p}.$$

□

By induction, we have the following:

COROLLARY 1.10 Let $a_1, a_2, \cdots, a_m$ be integers and let p be a prime. If $p|(a_1 a_2 \cdots a_m)$ then $p|a_k$ for some k.

1.6 Fundamental Theorem of Arithmetic

THEOREM 1.11 (Fundamental Theorem of Arithmetic) Every positive integer $n > 1$ can be expressed as a product of primes; this representation is unique apart from the order in which the factors occur.

Proof

We first show that n can be expressed as a prime or a product of primes. We use induction on n. The statement is clearly true for $n = 2$ since 2 is a prime. Suppose m is a prime or a product of primes for $2 \leq m \leq n-1$. If n is a prime then we are done. Suppose n is composite then $n = ab$, where $1 < a, b < n$. By induction each of the a and b is either a prime or a product of primes. Hence, $n = ab$ is a product of primes. By mathematical induction, every positive integer $n > 1$ is a prime or a product of primes.

To prove uniqueness, we use induction on n again. If $n = 2$ then the representation of n as a product of primes is clearly unique. Assume, then that it is true for all integers greater than 1 and less than n. We shall prove that it is also true for n. If n is prime, then there is nothing to prove. Assume, then, that n is composite and that n has two factorizations, say,

$$n = p_1 p_2 \cdots p_s = q_1 q_2 \cdots q_t. \tag{1.3}$$

Since p_1 divides the product $q_1 q_2 \cdots q_t$, it must divide at least one factor by Corollary 1.10. Relabel $q_1, q_2, ..., q_t$ so that $p_1 | q_1$. Then $p_1 = q_1$ since both p_1 and q_1 are primes. In (1.3), we may cancel p_1 on both sides to obtain

$$n/p_1 = p_2 \cdots p_s = q_2 \cdots q_t.$$

Now the induction hypothesis implies that the two factorizations of n/p_1 must be the same, apart from the order of the factors. Therefore, $s = t$ and the factorizations in (1.3) are also identical, apart from order. This completes the proof. □

In subsequent chapters, whenever we write

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r},$$

we mean that $p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$ is the prime power decomposition of n that is unique up to rearrangement of factors. When we write

$$n = \prod_{k=1}^{r} p_k^{\alpha_k}$$

we mean that $\alpha_j \neq 0, 1 \leq j \leq r$. If we write

$$n = \prod_{p} p^{\alpha_p},$$

then we understand that only finitely many α_p's are nonzero.

We also observe that using the Fundamental Theorem of Arithmetic, we deduce that if

$$a = \prod_p p^{\alpha_p} \qquad b = \prod_p p^{\beta_p},$$

then

$$(a, b) = \prod_p p^{\min(\alpha_p, \beta_p)}$$

and

$$[a, b] = \prod_p p^{\max(\alpha_p, \beta_p)}.$$

Hence,

$$(a, b)[a, b] = ab,$$

since

$$\min(h, k) + \max(h, k) = h + k.$$

This gives another proof of Theorem 1.7.

2 Arithmetical Functions

2.1 Arithmetical functions

DEFINITION 2.1 A real or complex-valued function defined on the set of positive integers is called an *arithmetical function*.

EXAMPLE 2.1 Here are examples of arithmetical functions:

1. The function $u(n) = 1$ for all positive integers n.
2. The function $N(n) = n$ for all positive integers n.
3. The function $d(n)$, the number of divisors of n.
4. The function $\sigma(n)$, the sum of divisors of n.

Given an arithmetical function $f(n)$, we can construct a new arithmetical function $g(n)$ by letting

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

Here the notation $\sum_{d|n} f(d)$ means the sum of $f(d)$ over all divisors of n.

Note that with the above notation, we may write

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

and

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

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. This construction of a new function from

a known function reminds us of constructing new continuous function through the integration of continuous function in Calculus.

Remark 2.1 If $d|n$ then $n = d(n/d)$ and this implies that n/d divides n. If n/d divides n then $d|n$. Summing over d is the same as summing over n/d since there is a one-one correspondence between these divisors. Therefore,

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

EXAMPLE 2.2 Show that

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

Solution

Note that

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

2.2 Multiplicative functions

DEFINITION 2.2 An arithmetical function f is said to be *multiplicative* if

$$f(1) = 1$$

and

$$f(mn) = f(m)f(n) \quad \text{whenever} \quad (m, n) = 1.$$

DEFINITION 2.3 An arithmetical function f is said to be *completely multiplicative* if $f(1) = 1$ and for all positive integers m and n,

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

EXAMPLE 2.3 The functions $u(n), N(n)$ are completely multiplicative. The functions $\varphi(n), d(n)$ and $\sigma(n)$ are all multiplicative but not completely multiplicative. (We will show later that $\varphi(n)$ is multiplicative.) The functions $\omega(n), \Omega(n)$ and $\Lambda(n)$ are not multiplicative.

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

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

and if f is multiplicative, then

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

This shows that if f is multiplicative, then its value at any positive integer n is determined by its values at prime powers.

If $f(n)$ is completely multiplicative, then

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

and the values of $f(n)$ is completely determined by the values of $f(p)$ for prime p.

We now prove a simple but useful result for multiplicative functions.

THEOREM 2.1 Let f be a multiplicative function. Then the function

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

is also multiplicative.

Proof

It is immediate that $g(1) = 1$. Let $(m, n) = 1$. Then observe that if $\ell | mn$, then we may write $\ell = \ell_1 \ell_2$ with $\ell_1 | m$ and $\ell_2 | n$ since m and n are relatively prime. To see this, suppose $\ell | mn$ and let $\ell_1 = (\ell, m), \ell_2 = (\ell, n)$. Note that $\ell_1 | m$ and $\ell_2 | n$ and

$$\ell_1 \ell_2 = (\ell, m)(\ell, n) = (mn, \ell(m, n), \ell^2) = (mn, \ell, \ell^2) = (mn, \ell) = \ell.$$

Therefore

$$\begin{aligned} g(mn) &= \sum_{\ell \mid mn} f(\ell) \\ &= \sum_{\ell_1 \mid m} \sum_{\ell_2 \mid n} f(\ell_1) f(\ell_2) \\ &= g(m)g(n). \end{aligned}$$

□

Remark 2.2

1. Let
$$\sigma_\alpha(n) = \sum_{d \mid n} d^\alpha,$$
where $\alpha \in \mathbf{Z}$. Note that $\sigma_1(n) = \sigma(n)$ and $\sigma_0(n) = d(n)$. Since d^α is multiplicative, by Theorem 2.1, $\sigma_\alpha(n)$ is multiplicative. Therefore $d(n)$ and $\sigma(n)$ are both multiplicative.
2. We can also show that $\varphi(n)$ is multiplicative using similar argument but at the moment, we will defer the proof of this fact.
3. Note that if $f(n)$ is completely multiplicative, $\sum_{d \mid n} f(d)$ may not be completely multiplicative. For example, $u(n)$ is completely multiplicative but $d(n) = \sum_{d \mid n} 1$ is not completely multiplicative.

2.3 Perfect numbers and $\sigma(n)$

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

$$\sigma(n) = 2n$$

or if and only if

$$\sigma(n) - n = n.$$

The following theorem gives the characterization of even perfect numbers:

THEOREM 2.2 Let n be a positive integer. An even integer N is perfect if and only if $N = 2^{k-1}(2^k - 1)$ where $2^k - 1$ is prime.

Proof
Let $N = 2^{k-1}(2^k - 1)$ with $2^k - 1$ a prime. Since $\sigma(n)$ is multiplicative,

$$\sigma(2^{k-1}(2^k - 1)) = \sigma(2^{k-1})\sigma(2^k - 1) = (2^k - 1)2^k = 2N.$$

Hence N is perfect.

Conversely, if N is even and perfect. Write $N = 2^{k-1}m$, $k \geq 2$ and m odd. Since $\sigma(n)$ is multiplicative and $(2^{k-1}, m) = 1$, we conclude that

$$\sigma(N) = \sigma(2^{k-1})\sigma(m) = \left(1 + 2 + \cdots + 2^{k-1}\right)\sigma(m) = (2^k - 1)\sigma(m). \tag{2.2}$$

But N is perfect and this implies that

$$\sigma(N) = 2N = 2^k m. \tag{2.3}$$

From (2.2) and (2.3), we deduce that

$$(2^k - 1)\sigma(m) = 2^k m.$$

Since $(2^k - 1, 2^k) = 1$, by Euclid's Lemma, we deduce that

$$(2^k - 1)|m. \tag{2.4}$$

By (2.4), we may write

$$m = (2^k - 1)s, \tag{2.5}$$

with $s \geq 1$. With this expression for m, we find using (2.2) and (2.3) that

$$\sigma(m) = 2^k s. \tag{2.6}$$

If $s > 1$ then (2.5) shows that 1, s and $(2^k - 1)s$ are all divisors of m. Hence,

$$\sigma(m) \geq 1 + s + (2^k - 1)s > s + (2^k - 1)s = 2^k s.$$

This contradicts (2.6). Therefore $s = 1$, $m = 2^k - 1$ and $\sigma(m) = 2^k$. But this means that 1 and $2^k - 1$ are the only divisors of m and hence $m = 2^k - 1$ must be a prime.

□

2.4 The Möbius function

Let us now introduce one of the most important arithmetical functions, namely, the Möbius function $\mu(n)$.

DEFINITION 2.4 Let $\mu(1) = 1$. If $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$, then define

$$\mu(n) = \begin{cases} (-1)^k & \text{if } \alpha_1 = \alpha_2 = \cdots = \alpha_k = 1, \\ 0 & \text{otherwise.} \end{cases}$$

The function $\mu(n)$ is known as the Möbius function.

DEFINITION 2.5 An arithmetical function $f(n)$ is additive if for any positive integers $(m, n) = 1$,

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

DEFINITION 2.6 The function $\omega(n)$ is defined by $\omega(1) = 0$ and $\omega(n)$ is the number of distinct prime divisors of n.

EXAMPLE 2.4 The function $\omega(n)$ is additive. For, if $(m, n) = 1$,

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

then

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

DEFINITION 2.7 Let m and n be positive integers. An arithmetical function $f(n)$ is completely additive if

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

DEFINITION 2.8 The function $\Omega(n)$ is defined by $\Omega(1) = 0$ and $\Omega(n)$ is the number of prime divisors of n.

EXAMPLE 2.5 The function $\Omega(n)$ is completely additive. This is because if

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

then

$$\Omega(mn) = \alpha_1 + \cdots + \alpha_k + \beta_1 + \cdots + \beta_t = \Omega(m) + \Omega(n).$$

If $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ and $\alpha_j = 1, 1 \leq j \leq k$, we say that n is squarefree. Note that in this case $\omega(n) = k$. Therefore, we have for $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$,

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

We now show that $\mu(n)$ is multiplicative. If either m or n is not squarefree, then $\mu(m)\mu(n) = 0$. Also, mn is not squarefree in this case and therefore $\mu(mn) = 0$. In other words,

$$\mu(mn) = \mu(m)\mu(n).$$

If both m and n are squarefree and $(m, n) = 1$, then

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

Therefore, $\mu(n)$ is multiplicative.

THEOREM 2.3 Let n be any positive integer and $[x]$ denote the integer part of a real number x. We have

$$\sum_{\ell|n} \mu(\ell) = \left[\frac{1}{n}\right] = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1. \end{cases}$$

Proof
By Theorem 2.1, we know that $g(n) = \sum_{\ell|n} \mu(\ell)$ is multiplicative. In other words, $g(1) = 1$ and

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

But

$$g(p^{\alpha_p}) = \mu(1) + \mu(p) + 0 + \cdots + 0 = 1 - 1 = 0.$$

In other words, if $n \neq 1$, then $g(n) = 0$. □

DEFINITION 2.9 Let n be any positive integer. The arithmetical function I is defined by

$$I(n) = \left[\frac{1}{n}\right] = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n > 1 \end{cases} \tag{2.7}$$

Using the above notation, we have

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

2.5 The Euler totient function

DEFINITION 2.10 The Euler totient $\varphi(n)$ is defined to be the number of positive integers not exceeding n which are relatively prime (see Definition 1.5) to n.

It is sometimes convenient to write $\varphi(n)$ as

$$\varphi(n) = \sum_{\substack{k=1 \\ (k,n)=1}}^{n} 1. \tag{2.8}$$

THEOREM 2.4 Let n be any positive integer. Then

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

Proof

If $g(k)$ is an arithmetical function, then

$$\sum_{\substack{k=1 \\ (k,n)=1}}^{n} g(k) = \sum_{k=1}^{n} g(k) I((k,n)),$$

where I is given by (2.7). Setting $g(k) = 1$, we find that

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

Now,

$$\varphi(n) = \sum_{k=1}^{n} I((k,n)) = \sum_{k=1}^{n} \sum_{\ell|(k,n)} \mu(\ell) = \sum_{k=1}^{n} \sum_{\substack{\ell|k \\ \ell|n}} \mu(\ell)$$

$$= \sum_{\ell|n} \mu(\ell) \sum_{q=1}^{n/\ell} 1 = \sum_{\ell|n} \mu(\ell)\frac{n}{\ell}.$$

This completes the proof of the theorem. □

Since $\mu(n)/n$ is multiplicative, we conclude from Theorem 2.1 that $\sum_{d|n} \mu(d)/d$ is multiplicative. This implies that $\varphi(n)/n$ is multiplicative. Since $N(n) = n$ is multiplicative, we deduce that $\varphi(n)$ is multiplicative and we state the result as follow:

THEOREM 2.5 If m and n are positive integers such that $(m, n) = 1$, then $\varphi(mn) = \varphi(m)\varphi(n)$.

THEOREM 2.6 Let n be any positive integer with prime factorization

$$n = \prod_{j=1}^{k} p_j^{\alpha_j}.$$

Then

$$\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right).$$

Proof
We are now required to compute $\varphi(p^k)$ for any prime p. Note that for $k \geq 1$,

$$\varphi(p) = p^k \sum_{d|p^k} \frac{\mu(d)}{d} = p^k \left(\frac{1}{1} + \frac{\mu(p)}{p}\right) = p^k - p^{k-1}.$$

□

Remark 2.3 The values of $\varphi(p^\alpha)$ can be computed directly. For $\alpha = 1$, since p is a prime, $\varphi(p) = p - 1$ since all integers less than p is relatively prime to p. For $\alpha > 1$, the integers less than p^α that is NOT relatively prime to p are multiples of p. There are $p^{\alpha-1}$ such integers. Therefore, there are $p^\alpha - p^{\alpha-1}$ integers less that p^α that are relatively prime to p^α, or

$$\varphi(p^\alpha) = p^\alpha - p^{\alpha-1}.$$

2.6 Dirichlet product of arithmetical functions and multiplicative functions

DEFINITION 2.11 Let f and g be two arithmetical functions. We define the *Dirichlet product* of f and g, denoted by $f * g$, as

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

We will often use $f * g$ to represent the function $(f * g)(n)$, suppressing the argument n.

Using the above notation, Theorem 2.4 can simply be written as

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

Our aim now is to show that the set of multiplicative functions, which we denote as $\mathcal{M}$, together with the operation $*$ forms an abelian group. We first note that $*$ is a binary operation on $\mathcal{M}$. The proof of this fact is similar to the proof of Theorem 2.1.

THEOREM 2.7 The function I is the identity function for $*$, that is, $I * f = f * I = f$ for every arithmetical function f.

Proof
By the definition of I, we find that

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

By the commutative law in Theorem 2.9, we conclude that

$$f * I = f.$$

□

Remark 2.4 Theorem 2.7 holds for any arithmetical function, not just multiplicative function.

THEOREM 2.8 Let f and g be multiplicative functions. Then $f * g$ is multiplicative.

Proof
Let $h = f * g$. Note that

$$h(1) = f(1)g(1) = 1.$$

Next, consider the expression

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

Given that $(m, n) = 1$, we can write $c = ab$, where $a|m$ and $b|n$. Therefore, we deduce that

$$\begin{aligned} h(mn) &= \sum_{a|m}\sum_{b|n} f(ab)g\left(\frac{m}{a}\frac{n}{b}\right) \\ &= \sum_{a|m}\sum_{b|n} f(a)f(b)g\left(\frac{m}{a}\right)g\left(\frac{n}{b}\right), \end{aligned}$$

since $(m/a, n/b) = 1$ and both f and g are multiplicative. This implies that

$$\begin{aligned} h(mn) &= \sum_{a|m} f(a)g\left(\frac{m}{a}\right)\sum_{b|n} f(b)g\left(\frac{n}{b}\right) \\ &= h(m)h(n). \end{aligned}$$

□

The following result shows that $*$ is both a commutative and associative operation on $\mathcal{M}$.

THEOREM 2.9 The Dirichlet product is commutative and associative, that is, for any arithmetical functions f, g, k, we have

$$f * g = g * f$$

and

$$(f * g) * k = f * (g * k).$$

Proof

The Dirichlet product of f and g is given by

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

Let $d_1 = n/d$ be the conjugate divisor of d. As d runs through all divisors of n, so does d_1. By (2.1),

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

To prove the associativity property, let $A = g * k$. Then

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

Similarly, if we set $B = (f * g)$, then

$$\begin{aligned}(B * k)(n) &= \sum_{d\cdot c=n} B(d)k(c) \\ &= \sum_{d\cdot c=n} \sum_{a\cdot b=d} f(a)g(b)k(c) \\ &= \sum_{a\cdot b\cdot c=n} f(a)g(b)k(c).\end{aligned}$$

Therefore,

$$(f * (g * k))(n) = ((f * g) * k)(n).$$

□

THEOREM 2.10 (Möbius inversion formula) If $f = g * u$, then $g = f * \mu$. Conversely, $g = f * \mu$ implies that $f = g * u$.

Proof
Suppose $f = g * u$. Then

$$f * \mu = (g * u) * \mu = g * (u * \mu) = g * I = g.$$

Conversely, if $g = f * \mu$ then

$$g * u = (f * \mu) * u = f * (\mu * u) = f * I = f.$$

□

We can now show the following identity that relates $N(n)$ to $\varphi(n)$.

THEOREM 2.11 Let n be any positive integer. Then

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

Proof

We have seen from Theorem 2.4 that

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

By Möbius inversion formula, we conclude that

$$N = u * \varphi.$$

□

We now show that for any arithmetical function $f(n)$ such that $f(1) \neq 0$ (not necessarily multiplicative), the inverse of f under $*$ exists.

THEOREM 2.12 Let f be an arithmetical function. If $f(1) \neq 0$, then there is a unique function g such that

$$f * g = I. \tag{2.9}$$

Proof
We show by induction on m that (2.9) has a unique solution $g(m)$. In order for (2.9) to hold, the function $g(n)$ must satisfy

$$f(1)g(1) = 1.$$

Since $f(1) \neq 0$, we find that

$$g(1) = \frac{1}{f(1)}$$

and $g(1)$ is uniquely determined. Suppose $m > 1$ and assume the values of $g(k)$ have been determined for $1 \leq k \leq m-1$. From (2.9), we find that

$$f(1)g(m) + \sum_{\substack{\ell | m \\ \ell > 1}} f(\ell) g\left(\frac{m}{\ell}\right) = 0.$$

Therefore,

$$g(m) = \frac{1}{f(1)} \left(- \sum_{\substack{\ell | m \\ \ell > 1}} f(\ell) g\left(\frac{m}{\ell}\right) \right)$$

and $g(m)$ is uniquely determined. By mathematical induction, there is a unique function $g(n)$ such that

$$f * g = I.$$

□

Remark 2.5 Theorem 2.12 holds for any arithmetical function f with $f(1) \neq 0$, not just multiplicative function.

DEFINITION 2.12 Given an arithmetical function f such that $f(1) \neq 0$. The unique function g such that $f * g = I$ is called the *Dirichlet inverse* of f. The notation for the Dirichlet inverse of f is f^{-1}.

EXAMPLE 2.6 From Theorem 2.3 which can be expressed as $I = u * \mu$, we conclude that the inverse of μ is u.

From the construction of f^{-1} in Theorem 2.12, it is not clear that the Dirichlet inverse of a multiplicative function f is multiplicative. To complete the proof that $(\mathcal{M}, *)$ forms an abelian group, it suffices to f^{-1} is multiplicative if f is multiplicative.

THEOREM 2.13 If both g and $f * g$ are multiplicative, then f is also multiplicative.

Proof
We prove the theorem by contradiction. Suppose f is not multiplicative. Let

$$h = f * g.$$

Since f is not multiplicative, there exist two relatively prime integers m and n such that

$$f(mn) \neq f(m)f(n).$$

We choose mn as small as possible. If $mn = 1$, then

$$f(1) \neq f(1)f(1),$$

which implies that $f(1) \neq 1$. Since $h(1) = f(1)g(1) = f(1) \neq 1$, we conclude that h is not multiplicative, which leads to a contradiction. Hence, $mn \neq 1$.

If $mn > 1$, then

$$f(ab) = f(a)f(b)$$

for all $1 \leq ab < mn$ and $(a, b) = 1$.

Now,

$$\begin{aligned} h(mn) &= \sum_{d|mn} f(d)g(mn/d) = \sum_{\substack{m=am' \\ n=bn'}} f(mn/(ab))g(ab) \\ &= f(mn) + \sum_{\substack{m=am' \\ n=bn' \\ ab\neq 1}} f(mn/(ab))g(ab) \\ &= f(mn) + \sum_{\substack{m=am' \\ n=bn' \\ ab\neq 1}} f(m/a)(n/b)g(a)g(b) \\ &= f(mn) + h(m)h(n) - f(m)f(n). \end{aligned}$$

Therefore $f(mn) = f(m)f(n)$, which contradicts our assumption that $f(mn) \neq f(m)f(n)$. □

THEOREM 2.14 If g is multiplicative, then the Dirichlet inverse g^{-1} is also multiplicative.

Proof
The functions g and $g * g^{-1} = I$ are multiplicative. By Theorem 2.13, g^{-1} is multiplicative. □

EXAMPLE 2.7 1. If f is completely multiplicative then $f^{-1} = \mu f$. This can be verified directly by show that $\mu f * f = I$.

2. The functions $\sigma^{-1} = \mu * \mu N$ and $\varphi^{-1} = u * \mu N$. The first identity follows from $\sigma = N * u$, which implies that $\sigma^{-1} = N^{-1} * u^{-1} = \mu N * \mu$ since $u^{-1} = \mu$ and $N^{-1} = \mu N$ (N being completely multiplicative).

2.7 Appendix

DEFINITION 2.13 Let f be an arithmetical function. The formal Dirichlet series associated with f is the formal series

$$D(f;s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}.$$

Note that if f and g are arithmetical functions, then

$$D(f+g;s) = D(f;s) + D(g;s)$$

and

$$D(f;s)D(g;s) = \sum_{\ell=1}^{\infty} \frac{f(\ell)}{\ell^s} \sum_{d=1}^{\infty} \frac{g(d)}{d^s} = \sum_{n=1}^{\infty} \frac{\sum_{d|n} f(n/d)g(d)}{n^s} = D(f*g;s).$$

Now, by Fundamental Theorem of Arithmetic, we may write $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$. If f is multiplicative, then the term $\dfrac{f(n)}{n^s}$ appears once in the expansion of the formal product

$$\prod_p \left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \cdots\right).$$

Conversely, if $D(f;s)$ can be expressed above product, then

$$f(n) = f(p_1^{\alpha_1}) \cdots f(p_k^{\alpha_k}).$$

This means that $f(mn) = f(m)f(n)$ whenever $(m,n) = 1$. This implies that f is multiplicative.

Therefore, f is multiplicative if and only if

$$D(f;s) = \prod_p \left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \cdots\right).$$

We next show that if f is multiplicative, then f^{-1} is multiplicative. Note that

$$D(f;s)D(f^{-1};s) = 1.$$

On the other hand,

$$\left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \cdots\right)\left(1 + \frac{f^{-1}(p)}{p^s} + \frac{f^{-1}(p^2)}{p^{2s}} + \cdots\right) = 1$$

since

$$\sum_{d|p^\alpha} f(d)f^{-1}(p^\alpha/d) = 0,$$

if $\alpha \geq 1$. This implies that

$$\begin{aligned} D(f;s)D(f^{-1};s) = 1 &= \prod_p \left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \cdots\right)\left(1 + \frac{f^{-1}(p)}{p^s} + \frac{f^{-1}(p^2)}{p^{2s}} + \cdots\right) \\ &= D(f;s)\left(1 + \frac{f^{-1}(p)}{p^s} + \frac{f^{-1}(p^2)}{p^{2s}} + \cdots\right). \end{aligned}$$

Hence,

$$D(f^{-1};s) = \prod_p \left(1 + \frac{f^{-1}(p)}{p^s} + \frac{f^{-1}(p^2)}{p^{2s}} + \cdots\right),$$

and therefore, f^{-1} is multiplicative.

3 Averages of Arithmetical Functions

3.1 Introduction

Let x be a positive real number. We use the notation

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

to denote the sum

$$f(1) + f(2) + \cdots + f([x]).$$

For positive real number x, the mean of the function f from 1 to x is defined by

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

The purpose of studying $\overline{f}(x)$ is because in general, $\overline{f}(x)$ behaves more regularly than $f([x])$, especially when x is large. For example, when f is the characteristic function for primes, namely,

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

The function

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

is usually written as $\pi(x)$ and the Prime Number Theorem states that $\overline{f}(x) = \pi(x)/x$ "behaves" like $1/\ln x$. On the other hand, we cannot predict the value of $f(n)$ for each $n = [x]$ since we do not know the location of primes in $\mathbf{Z}$.

We now introduce the "big-O" notation and the notion of asymptotic.

DEFINITION 3.1 Let a be any real number and let $g(x)$ be a real-valued function such that $g(x) > 0$ when $x \ge a$. We write

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

to mean that the quotient $f(x)/g(x)$ is bounded for $x \geq a$; that is, there exists a constant $M > 0$ such that

$$|f(x)| \leq Mg(x) \quad \text{for all } x \geq a.$$

Sometimes, we will also use the notation

$$f(x) \ll g(x)$$

to represent $f(x) = O(g(x))$.

EXAMPLE 3.1 The function $x^2 = O(x^3)$ when x is large. The function $x^n = O(e^x)$ for any positive integer n.

DEFINITION 3.2 If

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

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

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

EXAMPLE 3.2 The Prime Number Theorem can be written as

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

3.2 Partial summation and the Euler-Maclaurin summation formula

THEOREM 3.1 Let $a(n)$ be an arithmetic function and set

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

Let $0 \leq y < x$ be real numbers and f be a real-valued function with continuous derivative on $[y, x]$. Then

$$\sum_{y<n\leq x} a(n)f(n) = f(x)A(x) - f(y)A(y) - \int_y^x A(t)f'(t)\,dt. \tag{3.1}$$

Proof
We observe that

$$\begin{aligned}\int_y^x A(t)f'(t)\,dt &= \int_y^x \sum_{n\le t} a(n)f'(t)\,dt \\ &= \sum_{n\le x} a(n)\int_{\max(y,n)}^x f'(t)\,dt \\ &= \sum_{n\le x} a(n)[f(x)-f(\max(y,n))].\end{aligned} \tag{3.2}$$

Therefore,

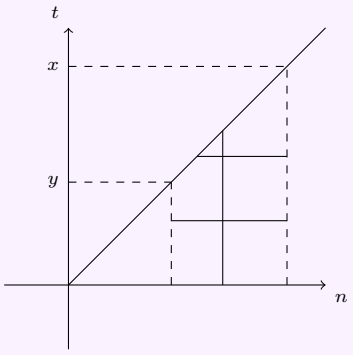
$$\begin{aligned}\int_y^x A(t)f'(t)\,dt &= f(x)A(x)-\sum_{n\le y} a(n)f(y)-\sum_{y<n\le x} a(n)f(n) \\ &= f(x)A(x)-f(y)A(y)-\sum_{y<n\le x} a(n)f(n).\end{aligned}$$

Simplifying, we find that

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

□

Remark 3.1 The second equality of (3.2) follows from interchanging the integral with the summation. We now explain the limits in the integral using Figure 3.1. Note that for a fixed t, the sum is over all $n \le t$ (consider the vertical line). For a fixed n, we integrate from y to x if $n < y$ and from n to x if $n \ge y$ (consider the two horizontal lines in the shaded region). Hence for a fixed n, we integrate from $\max(n, y)$ to x.



3.3 The Euler-Maclaurin summation formula (Special case)

In this section, we deduce the Euler-Maclaurin summation from Theorem 3.1.

THEOREM 3.2 (The Euler-Maclaurin summation formula) Let $0 < y < x$ and let $f(x)$ be a real-valued function with continuous derivative on $[y, x]$. Then

$$\sum_{y<n\leq x} f(n) = \int_y^x f(t)\,dt + \int_y^x \{t\}f'(t)\,dt - f(x)\{x\} + f(y)\{y\}. \tag{3.3}$$

Proof
By partial summation formula with $a(n) = 1$ and $A(x) = [x]$, we find that

$$\begin{aligned}
&\sum_{y<n\leq x} f(n) = f(x)[x] - f(y)[y] - \int_y^x [t]f'(t)\,dt \\
&= f(x)x - \{x\}f(x) + f(y)\{y\} - f(y)y - \int_y^x (t - \{t\})f'(t)\,dt \\
&= -\{x\}f(x) + f(y)\{y\} + \int_y^x \{t\}f'(t)\,dt + f(x)x - f(y)y - \int_y^x tf'(t)\,dt \\
&= -f(x)\{x\} + f(y)\{y\} + \int_y^x \{t\}f'(t)\,dt + \int_y^x f(t)\,dt.
\end{aligned}$$

□

3.4 Some elementary asymptotic formulas

DEFINITION 3.3 For each real number $s > 1$, we define the Riemann zeta function as

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

DEFINITION 3.4 The Euler constant γ is defined as

$$\gamma = \lim_{n\to\infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \ln n\right).$$

THEOREM 3.3 If $x \geq 1$, then

(a) $$\sum_{n \leq x} \frac{1}{n} = \ln x + \gamma + O\left(\frac{1}{x}\right),$$

(b) $$\sum_{n \leq x} \frac{1}{n^s} = \frac{x^{1-s}}{1-s} + C(s) + O(x^{-s}) \quad \text{if } s > 0 \text{ and } s \neq 1,$$

where

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

Proof

To prove (a), we first let $f(t) = 1/t$ in Theorem 3.2. Then by (3.3),

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

The improper integral

$$\int_1^\infty \{t\} t^{-2}\, dt$$

exists since it is dominated by

$$\int_1^\infty t^{-2}\, dt.$$

Furthermore,

$$0 \leq \int_x^\infty \frac{\{t\}}{t^2}\, dt \leq \int_x^\infty \frac{1}{t^2}\, dt = \frac{1}{x},$$

so the last equation becomes

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

This proves (a) with

$$\gamma = 1 - \int_1^\infty \frac{\{t\}}{t^2}\, dt = \lim_{x \to \infty} \left(\sum_{n \leq x} \frac{1}{n} - \ln x \right).$$

To prove (b), we use the same argument with

$$f(x) = x^{-s},$$

where $s > 0, s \neq 1$. The Euler-Maclaurin summation implies that

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

Therefore,

$$\sum_{n\leq x} \frac{1}{n^s} = \frac{x^{1-s}}{1-s} + C(s) + O(x^{-s}), \tag{3.4}$$

where

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

If $s > 1$ then the left-hand side of (3.4) 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

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

and (3.4) shows that

$$C(s) = \lim_{x\to\infty} \left(\sum_{n\leq x} \frac{1}{n^s} - \frac{x^{1-s}}{1-s} \right)$$

and this completes the proof of (b). □

EXAMPLE 3.3 We now note that

$$\begin{aligned}\sum_{n\leq x} d(n) &= \sum_{d\ell\leq x} 1 = \sum_{d\leq x} \sum_{\ell\leq x/d} 1 \\ &= \sum_{d\leq x} \left[\frac{x}{d}\right] = \sum_{d\leq x} \frac{x}{d} - \sum_{d\leq x} \left\{\frac{x}{d}\right\} \\ &= x(\ln x + C + O(1/x)) = x\ln x + O(x).\end{aligned}$$

In the above example, we find that

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

In general,

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

The sum on the left hand side can also be written as

$$\sum_{n\le x} u * f(n) = \sum_{\ell\le x} F(x/\ell)$$

where

$$F(x) = \sum_{n\le x} f(n).$$

EXAMPLE 3.4 Note that

$$\begin{aligned} 1 &= \sum_{n\le x}\sum_{d|n}\mu(d) = \sum_{d\le x}\mu(d)\left[\frac{x}{d}\right] \\ &= x\sum_{d\le x}\frac{\mu(d)}{d} - \sum_{d\le x}\mu(d)\left\{\frac{x}{d}\right\}. \end{aligned}$$

Therefore,

$$\begin{aligned} x\left|\sum_{d\le x}\frac{\mu(d)}{d}\right| &\le \left|1+\sum_{d\le x}\mu(d)\left\{\frac{x}{d}\right\}\right| \\ &\le 1+\sum_{d\le x}\left\{\frac{x}{d}\right\} \\ &= 1+\{x\}+[x]-1 = x. \end{aligned}$$

Hence,

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

3.5 The divisor function and Dirichlet's hyperbola method

In this section, we will first discuss the hyperbola method and then apply the method to study the mean value of the divisor function $d(n)$.

THEOREM 3.4 Let f and g be two arithmetic functions with

$$F(x) = \sum_{n\le x} f(n), \quad \text{and} \quad G(x) = \sum_{n\le x} g(n).$$

For $1 \le y \le x$, we have

$$\sum_{n\le x}(f*g)(n) = \sum_{n\le y} g(n)F\left(\frac{x}{n}\right) + \sum_{m\le \frac{x}{y}} f(m)G\left(\frac{x}{m}\right) - F\left(\frac{x}{y}\right)G(y).$$

Proof
First, we observe that

$$\sum_{n\le x}(f*g)(n)=\sum_{md\le x}f(m)g(d).$$

Next, for $y\le x$, we find that

$$\begin{aligned}\sum_{md\le x}f(m)g(d)&=\sum_{\substack{md\le x\\ d\le y}}f(m)g(d)+\sum_{\substack{md\le x\\ d>y}}f(m)g(d)\\&=\sum_{d\le y}g(d)F\left(\frac{x}{d}\right)+\sum_{m\le\frac{x}{y}}f(m)\left\{G\left(\frac{x}{m}\right)-G(y)\right\}.\end{aligned}$$

□

We now set $f=g=u$. Then

$$f*g=u*u=d.$$

Note that $F(x)=[x]=G(x)$. Let $y=\sqrt{x}$. Then by Theorem 3.4,

$$\begin{aligned}\sum_{n\le x}d(n)&=2\sum_{n\le\sqrt{x}}\left[\frac{x}{n}\right]-[\sqrt{x}]^2\\&=2x\sum_{n\le\sqrt{x}}\frac{1}{n}-x+O(\sqrt{x}).\end{aligned}$$

Using Theorem 3.3 (a), we conclude that

THEOREM 3.5 For all $x\ge 1$,

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

where γ is the Euler's constant.

As a corollary, we deduce that

$$\overline{d}(x)\sim\ln x. \tag{3.5}$$

In other words, the average order of $d(n)$ is $\ln n$.

Remark 3.2 The error term in Theorem 3.5 can be improved. In 1903 Voronoi proved that it is $O(x^{1/3}\log x)$. In 1928, J.G. van der Corput improved the error term to $O(x^{27/82})$ using the method of exponential sums. In 1988, H. Iwaniec and C.J. Mozzochi showed that the error term can be taken as $O(x^{7/22})$. The best possible error term is one given recently by M.N. Huxley in 2003, who showed that the error is $O(x^{131/416}(\ln x)^{26947/8320})$.

3.6 An application of the hyperbola method

An interesting question one can ask is:

"If two positive integers are randomly chosen, what is the probability that they are relatively prime?"

To answer this question, we first show the following result:

THEOREM 3.6 Let $\varphi(n)$ be the Euler φ function. For $x > 1$,

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

Remark 3.3 The above result shows that the average order of $\varphi(n)$ is $3n/\pi^2$.

Proof

We recall that $\varphi = \mu * N$. Applying Theorem 3.4 with $f = N$ and $g = \mu$, we find that

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

where

$$F(x) = \sum_{n\le x}N(n) = \frac{x^2}{2} + O(x)$$

and

$$G(x) = \sum_{n\le x}\mu(n) = O(x).$$

Therefore,

$$\begin{aligned}\sum_{n\le x}\varphi(n) = \frac{1}{2}\sum_{n\le y}\mu(n)\left(\frac{x}{n}\right)^2 + O\left(\left|\sum_{n\le y}\mu(n)\right|x\right)\\ + O\left(\sum_{m\le\frac{x}{y}}m\frac{x}{m}\right) + O\left(\left(\frac{x}{y}\right)^2 y\right).\end{aligned}$$

Let $y = \sqrt{x}$ and we conclude that

$$\sum_{n\le x}\varphi(n) = \sum_{n\le\sqrt{x}}\frac{\mu(n)}{n^2}\frac{x^2}{2} + O(x^{3/2}). \tag{3.6}$$

We will show in the chapter on Dirichlet series that

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

This implies that

$$\sum_{n=1}^{\infty}\frac{\mu(n)}{n^2}=\frac{6}{\pi^2},$$

since

$$\zeta(2)=\frac{\pi^2}{6}.$$

Given the above identity, we find that

$$\begin{aligned}\sum_{n\leq\sqrt{x}}\frac{\mu(n)}{n^2}&=\sum_{n=1}^{\infty}\frac{\mu(n)}{n^2}-\sum_{n>\sqrt{x}}\frac{\mu(n)}{n^2}\\&=\sum_{n=1}^{\infty}\frac{\mu(n)}{n^2}+O\left(\sum_{n>\sqrt{x}}\frac{1}{n^2}\right)\\&=\frac{6}{\pi^2}+O(x^{-1/2}).\end{aligned}$$

Substituting the above into (3.6), we conclude the proof of the theorem. □

Now, let T be a positive integer and

$$S=\{(m,n)|1\leq m\leq T, 1\leq n\leq T\}.$$

Then the total number of elements in S such that $(m,n)=1$ is given by

$$\begin{aligned}\sum_{\substack{n\leq T\\(m,n)=1}}\sum_{m\leq T}1&=1+2\sum_{\substack{m\leq T\\(m,n)=1}}\sum_{n<m}1\\&=1+2\sum_{m\leq T}\varphi(m)=\frac{6}{\pi^2}T^2+O(T^{3/2}).\end{aligned}$$

This shows that the probability that two randomly chosen positive integers are relatively prime is

$$\lim_{T\to\infty}\frac{|S|}{T^2}=6/\pi^2.$$

3.7 Some facts about Riemann-Stieltjes integrals

In this section, we give another proof of Theorem 3.2. Our main reference of this section is *Mathematical Analysis* (second edition) by T.M. Apostol (will be referred to as [MA-Apostol]).

DEFINITION 3.5 If $[a, b]$ is a compact interval, a set of points

$$P = \{a = x_0, x_1, x_2, \cdots, x_{n-1}, x_n = b\}$$

satisfying the inequalities

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b,$$

is called *a partition of* $[a, b]$. The collection of all possible partitions of $[a, b]$ will be denoted by $\mathcal{P}[a, b]$.

DEFINITION 3.6 Let f be defined on $[a, b]$ and let $P = \{a = x_0, x_1, x_2, \cdots, x_{n-1}, x_n = b\}$ be a partition of $[a, b]$. If there exists a positive number M such that

$$\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| \leq M$$

for all partitions $P \in \mathcal{P}[a, b]$, then f is said to be of *bounded variation* on $[a, b]$.

DEFINITION 3.7 A partition P' is said to be *finer* than P if $P \subset P'$.

DEFINITION 3.8 Let $P = \{a = x_0, x_1, x_2, \cdots, x_{n-1}, x_n = b\}$ be a partition of $[a, b]$ and let t_k be a point in the subinterval $[x_{k-1}, x_k]$. A sum of the form

$$S(P, f, \alpha) = \sum_{k=1}^{n} f(t_k)(\alpha(x_k) - \alpha(x_{k-1}))$$

is called a *Riemann-Stieltjes sum of* f with respect to α. We say that f is Riemann-integrable with respect to α on $[a, b]$, and we write "$f \in R(\alpha)$ on $[a, b]$" if there exists a number A having the property : For every $\epsilon > 0$, there exists a partition P_ϵ of $[a, b]$ such that for every partition P finer than P_ϵ and for every choice of the points $t_k \in [x_{k-1}, x_k]$, we have

$$|S(P, f, \alpha) - A| < \epsilon.$$

The number A, if exists, is called the *Riemann-Stieltjes integral of* f *with respect to* α *on* $[a, b]$. The number A is denoted by

$$\int_a^b f d\alpha \quad \text{or} \quad \int_a^b f(t) d\alpha(t).$$

We have the following facts:

THEOREM 3.7 If f is continuous on $[a,b]$ and if α is of bounded variation on $[a,b]$, then $f \in R(\alpha)$ on $[a,b]$.

THEOREM 3.8 If $f \in R(\alpha)$ on $[a,b]$, then $\alpha \in R(f)$ on $[a,b]$ and we have

$$\int_a^b f(x)d\alpha(x) + \int_a^b \alpha(x)df(x) = f(b)\alpha(b) - f(a)\alpha(a).$$

For the proofs of Theorems 3.7 and 3.8, see [MA-Apostol, p.159, Theorem 7.27] and [MA-Apostol, p.144, Theorem 7.6] respectively.

We now give another proof of Theorem 3.1.

Second proof of Theorem 3.1

Note that if P is chosen so that $[x_{k-1}, x_k]$ contains just one integral point, then

$$\begin{aligned} S(P,f,\alpha) &= \sum_{k=1}^{n} f(t_k)(A(x_k) - A(x_{k-1})) \\ &= f(s_{[y+1]})a([y+1]) + \cdots + f(s_{[x]})a([x]), \end{aligned}$$

where $s_j = t_k$ for some k for which $[x_{k-1}, x_k]$ contains the integer j. Note that as the length of each interval of the partition P tends to 0, $t_k \to j$. Hence, we conclude that

$$\int_y^x f(t)dA(t) = \sum_{y<n\leq x} f(n)a(n). \tag{3.7}$$

By Theorem 3.8, we find that

$$\int_y^x f(t)dA(t) + \int_y^x A(t)df(t) = f(x)A(x) - f(y)A(y)$$

and this is precisely Theorem 3.1 by (3.7).

□

Remark 3.4 Let $y = 1$ in Theorem 3.1. We find that

$$\sum_{1<n\leq x} a(n)f(n) = A(x)f(x) - A(1)f(1) - \int_1^x A(t)f'(t)\,dt.$$

But

$$\begin{aligned} \sum_{1<n\leq x} a(n)f(n) + A(1)f(1) &= \sum_{1<n\leq x} a(n)f(n) + a(1)f(1) \\ &= \sum_{n\leq x} a(n)f(n). \end{aligned}$$

Consequently, we have

$$\sum_{n\le x} a(n)f(n) = A(x)f(x) - \int_1^x A(t)f'(t)\,dt. \tag{3.8}$$

4 Elementary Results on the Distribution of Primes

4.1 Introduction

We first recall the definition of $\pi(x)$:

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

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

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

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

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

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

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

There are other elementary proofs of the prime number theorem since the appearance of the work of Selberg and Erdös, one of which is due to A. Hildebrand. The proof given by Hildebrand relies on proving an equivalent statement of the Prime Number Theorem and the mean value of $\mu(n)$. In this chapter, we derive

some basic properties of $\pi(x)$ and establish several statements equivalent to the Prime Number Theorem.

4.2 The function $\psi(x)$

We recall the definition of Mangoldt's function

DEFINITION 4.2 Let n be a positive integer and let

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

EXAMPLE 4.1 We observe that if $n = \prod_{j=1}^{k} p_j^{\alpha_j}$ then

$$\sum_{d|n} \Lambda(n) = \sum_{j=1}^{k} \alpha_j \ln p_j = \ln n.$$

Also,

$$\begin{aligned} \sum_{n \le x} \Lambda\left[\frac{x}{n}\right] &= \sum_{n \le x} \Lambda * u(n) \\ &= \sum_{n \le x} \ln n = x \ln x - x + O(\ln x). \end{aligned} \tag{4.2}$$

DEFINITION 4.3 For real number $x \ge 1$,

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

THEOREM 4.2 There exist positive constants c_1 and c_2 such that

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

Proof
For $x \ge 4$, let

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

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

$$\begin{aligned}\sum_{n\le x} \ln n &= \int_1^x \ln t dt + \int_1^x \{t\}\frac{1}{t}dt - \{x\}\ln x + \{y\}\ln y \\ &= x\ln x - x + O(\ln x). \end{aligned} \tag{4.3}$$

This implies that

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

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

$$\frac{x}{2} \le S \le x \tag{4.4}$$

whenever $x \ge x_0 \ge 4$. Next, since

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

we find that

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

Hence,

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

where

$$\theta_d = \left[\frac{x}{d}\right] - 2\left[\frac{x}{2d}\right]. \tag{4.5}$$

Now, for

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

we have

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

Therefore, we may simplify the second term on the right-hand side of (4.5) to obtain

$$S = \sum_{d\le\frac{x}{2}}\Lambda(d)\theta_d + \sum_{\frac{x}{2}<d\le x}\Lambda(d). \tag{4.6}$$

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

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

Using (4.6), we deduce that

$$S \le \sum_{d \le \frac{x}{2}} \Lambda(d) + \sum_{\frac{x}{2} < d \le x} \Lambda(d) = \sum_{d \le x} \Lambda(d) = \psi(x) \tag{4.7}$$

and

$$S \ge \sum_{\frac{x}{2} < d \le x} \Lambda(d) = \psi(x) - \psi\left(\frac{x}{2}\right). \tag{4.8}$$

From (4.4) and (4.7),

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

Therefore,

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

To obtain a lower bound for $\psi(x)$, we first deduce from (4.4), (4.8) that

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

Therefore,

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

This implies that

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

for some positive real number c_2. □

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

DEFINITION 4.4 For real number $x \ge 1$, let

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

THEOREM 4.3 For real number $x \ge 1$, we have

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

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

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

Hence,

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

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

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

COROLLARY 4.4 For $x \ge 4$, there exist real positive constants c_1 and c_2 such that

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

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

THEOREM 4.5 For each positive real $x \ge 4$,

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

Proof
It suffices to prove that

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

by Theorem 4.3. We observe that

$$\begin{aligned}\pi(x) - \frac{\theta(x)}{\ln x} &= \sum_{p \le x} \left(1 - \frac{\ln p}{\ln x}\right) \\ &= \sum_{p \le x} \ln p \left(\frac{1}{\ln p} - \frac{1}{\ln x}\right). \qquad (4.9)\end{aligned}$$

If

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

then by Corollary 4.4,

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

The last expression in (4.9) is

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

□

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

COROLLARY 4.6 The Prime Number Theorem

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

is equivalent to each of the following relations:

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

4.4 Second proof of Chebyshev's estimate

In this section, we give another proof of Theorem 4.5 that is due to M. Nair (see Amer. Math. Monthly, 89, no.2, 126-129). Our presentation is a modification of Tenenbaum's Section 1.2.

The upper bound in Theorem 4.5 can already be found in Erdös' proof of Bertrand's postulate, a result that states that for any positive integer $n \ge 2$, there exists at least a prime p between n and $2n$. It is given as follows:

LEMMA 4.7 For $n \geq 2$,

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

Proof

Let $P(n)$ denote the statement. It is clear that $P(2)$ and $P(3)$ are true. If $m > 1$, then

$$\prod_{p \leq 2m+2} p = \prod_{p \leq 2m+1} p \leq 4^{2m+1} < 4^{2m+2}.$$

Therefore,

$$P(2m+1) \text{ implies } P(2m+2).$$

Suppose $n = 2m+1$. Then each prime in the interval $[m+2, 2m+1]$ is a factor of $\binom{2m+1}{m}$. This is because primes in the interval do not occur in the denominator of $\binom{2m+1}{m}$ (which is $m!(m+1)!$).

Since $P(m+1)$ holds, we find that

$$\prod_{p \leq 2m+1} p = \prod_{m+2 \leq p \leq 2m+1} p \prod_{p \leq m+1} p \leq \binom{2m+1}{m} 4^{m+1}.$$

But,

$$(1+1)^{2m+1} = \binom{2m+1}{0} + \binom{2m+1}{1} + \cdots + \binom{2m+1}{m} + \binom{2m+1}{m+1} + \cdots + \binom{2m+1}{2m+1} \geq 2\binom{2m+1}{m}.$$

Therefore,

$$\binom{2m+1}{m} < 4^m.$$

Hence,

$$\prod_{p \leq 2m+1} p \leq 4^m \cdot 4^{m+1} = 4^{2m+1}$$

and $P(2m+1)$ is true. □

Now, let t be an expression in terms of n which will be chosen later. from Lemma 4.7,

$$t^{\pi(n)-\pi(t)} < \prod_{t < p \leq n} p \leq 4^n.$$

This yields

$$\pi(n) \leq n\frac{\ln 4}{\ln t} + \pi(t) \leq n\frac{\ln 4}{\ln t} + t.$$

Choosing $t = \sqrt{n}$, we conclude that

$$\pi(n) \leq D\frac{n}{\ln n}$$

for some positive constant D. This gives the upper bound of Theorem 4.5.

The lower bound is harder to prove and the brilliant proof is due to Nair.

Let $d_n = [1, 2, 3, \cdots, n]$, i.e., d_n is the lowest common multiple of the integers from 1 to n. Note that if $j = \prod_p p^{\alpha_{j,p}}$, then

$$d_n = \prod_p p^{\max(\alpha_{1,p}, \alpha_{2,p}, \cdots, \alpha_{n,p})}.$$

This means that if $p^{\nu_p} \| d_n$ then $p^{\nu_p} = p^{\alpha_{k_p,p}}$ for some k_p and therefore,

$$p^{\nu_p} \leq k_p$$

for some k_p between 1 and n. In other words,

$$d_n = \prod_p p^{\max(\alpha_{1,p}, \alpha_{2,p}, \cdots, \alpha_{n,p})} \leq \prod_{p \leq n} k_p \leq \prod_{p \leq n} n = n^{\pi(n)}. \tag{4.10}$$

We claim that

LEMMA 4.8 For $n \geq 7$,

$$d_n \geq 2^n.$$

Assuming that Lemma 4.8 is true, then from (4.10), we deduce that

$$\pi(n) \geq \frac{\ln d_n}{\ln n} \geq \ln 2\frac{n}{\ln n},$$

which gives the lower bound for Theorem 4.5.

It remains to prove Lemma 4.8.

Proof

We first recall that the beta integral implies that

$$I(m, n) = \int_0^1 x^{m-1}(1-x)^{n-m} = \frac{(m-1)!(n-m)!}{n!} = \frac{1}{m\binom{n}{m}}. \tag{4.11}$$

Next, note that

$$I(m,n) = \sum_{j=0}^{n-m} (-1)^j \binom{n-m}{j} \int_0^1 x^{m+j-1} dx$$
$$= \sum_{j=0}^{n-m} (-1)^j \binom{n-m}{j} \frac{1}{m+j} \in \frac{1}{d_n}\mathbf{Z}.$$

Therefore, from (4.11), we deduce that

$$m\binom{n}{m} \mid d_n. \tag{4.12}$$

If we replace n by $2n$ and m by n in (4.12), we deduce that

$$n\binom{2n}{n} \mid d_{2n}. \tag{4.13}$$

Similarly, if we replace n by $2n+1$ and m by $n+1$ in (4.12), we deduce that

$$(n+1)\binom{2n+1}{n+1} \mid d_{2n+1}. \tag{4.14}$$

Using the identity

$$(n+1)\binom{2n+1}{n+1} = (2n+1)\binom{2n}{n},$$

we deduce from (4.14) that

$$(2n+1)\binom{2n}{n} \mid d_{2n+1}. \tag{4.15}$$

Now,

$$d_{2n} | d_{2n+1}$$

and therefore, (4.13) implies that

$$n\binom{2n}{n} \mid d_{2n+1}. \tag{4.16}$$

Since $(n, 2n+1) = 1$, we conclude from (4.15) and (4.16) that

$$n(2n+1)\binom{2n}{n} | d_{2n+1}.$$

Now,

$$d_{2n+1} \geq n(2n+1)\binom{2n}{n} \geq n(1+1)^{2n} = n2^{2n} \geq 2^{2n+1}$$

if $n \geq 2$. Also,

$$d_{2n+2} \geq d_{2n+1} \geq 2^{2n+2}$$

for $n \geq 4$. Therefore,

$$d_n \geq 2^n$$

if $n \geq 10$. The inequality for d_n for the remaining cases for $7 \leq n \leq 9$ can be checked directly.

□

4.5 Merten's estimates

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

THEOREM 4.9 (Merten's estimates) Let x be a positive real number greater than 1. We have

(a) $\displaystyle\sum_{n \leq x} \frac{\Lambda(n)}{n} = \ln x + O(1)$,

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

(c) $\displaystyle\sum_{p \leq x} \frac{1}{p} = \ln \ln x + A + O\left(\frac{1}{\ln x}\right)$, and

(d) (Merten's Theorem) $\displaystyle\prod_{p \leq x} \left(1 - \frac{1}{p}\right) = \frac{e^{-A}}{\ln x}\left(1 + O\left(\frac{1}{\ln x}\right)\right)$,

where A is a constant.

Proof

(a) First, using (4.2), we write

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

Now,

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

Hence, we deduce that

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

(b) We observe that

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

Hence,

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

(c) Let

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

where

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

Then, we find that

$$\begin{aligned}\sum_{p\le x}\frac{1}{p} &= \sum_{p\le x}\left(\frac{\ln p}{p}\right)\left(\frac{1}{\ln p}\right)\\ &= A(x)\frac{1}{\ln x} - \int_2^x A(t)\left(\frac{1}{\ln t}\right)' dt\\ &= \frac{A(x)}{\ln x} + \int_2^x \frac{A(t)}{t\ln^2 t}\, dt. \qquad (4.17)\end{aligned}$$

By Theorem 4.9 (b), we find that

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

with

$$R(t) \ll 1, \quad t \ge 2. \qquad (4.18)$$

Using (4.18) in the last term of (4.17), we deduce that

$$\begin{aligned}\int_2^x \frac{\ln t + R(t)}{t \ln^2 t} dt &= \int_2^x \frac{dt}{t \ln t} + \int_2^x \frac{R(t)}{t \ln^2 t} dt \\ &= \ln \ln x - \ln \ln 2 + \int_2^\infty \frac{R(t)}{t \ln^2 t} dt - \int_x^\infty \frac{R(t)}{t \ln^2 t}\, dt \\ &= \ln \ln x - \ln \ln 2 + A'' + O\left(\frac{1}{\ln x}\right). \end{aligned} \tag{4.19}$$

Substituting (4.19) into (4.17), we conclude our proof of (c).

(d) We observe that

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

where

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

Hence,

$$\begin{aligned}\ln \prod_{p \le x} \left(1 - \frac{1}{p}\right) &= \sum_{p \le x} r_p - \sum_{p \le x} \frac{1}{p} \\ &= -\ln \ln x + A + O\left(\frac{1}{\ln x}\right) + \sum_p r_p - \sum_{p > x} r_p. \end{aligned} \tag{4.20}$$

Now,

$$r_p = -\sum_{m=2}^{\infty} \frac{1}{m p^m} = O\left(\frac{1}{p^2}\right), \tag{4.21}$$

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

$$m p^m \ge 2^m.$$

Using (4.21) in (4.20), we deduce that

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

Hence,

$$\ln \prod_{p \le x} \left(1 - \frac{1}{p}\right) = -\ln \ln x + A' + O\left(\frac{1}{\ln x}\right). \tag{4.22}$$

Exponentiating both sides of (4.22), we arrive at

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

since $e^t = 1 + O(t)$. □

EXAMPLE 4.2 Find an asymptotic formula for

$$\sum_{n\le x}\omega(n),$$

where $\omega(n)$ is the number of distinct prime divisors of n.

Solution

Rewrite the sum as

$$\begin{aligned}\sum_{n\le x}\omega(n) &= \sum_{n\le x}\sum_{p\mid n}1\\ &= \sum_{p\le x}\sum_{\ell\le x/p}1 = \sum_{p\le x}\left[\frac{x}{p}\right]\\ &= \sum_{p\le x}\frac{x}{p} + O(\sum_{p\le x}1)\\ &= x\ln\ln x + Ax + O(x/\ln x),\end{aligned}$$

where we have used (c) and Chebyshev's estimate $\pi(x) = O(x/\ln x)$.

4.6 Bertrand's postulate (Erdös' proof)

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

THEOREM 4.10 (Bertrand's Postulate) Let n be an integer. Then for $n \ge 2$, there exists a prime p between n and $2n$.

We will need several elementary lemmas.

LEMMA 4.11 Let $r(p)$ satisfies

$$p^{r(p)} \leq 2n < p^{r(p)+1}. \tag{4.23}$$

Then

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

Proof

The number of integers less than n and divisible by m is $\left[\frac{n}{m}\right]$. Therefore, the number of integers from 1 to n that is exactly a multiple of p^j is

$$\left[\frac{n}{p^j}\right] - \left[\frac{n}{p^{j+1}}\right].$$

Hence, the exponent of p in $n!$ is

$$\begin{aligned}\left(\left[\frac{n}{p}\right] - \left[\frac{n}{p^2}\right]\right) + 2\left(\left[\frac{n}{p^2}\right] - \left[\frac{n}{p^3}\right]\right) + \cdots + (k-1)\left(\left[\frac{n}{p^{k-1}}\right] - \left[\frac{n}{p^k}\right]\right) + \left[\frac{n}{p^k}\right] \\ = \left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \cdots + \left[\frac{n}{p^k}\right],\end{aligned}$$

where k is such that

$$p^k \leq n < p^{k+1}.$$

Therefore the exponent of p is $\binom{2n}{n}$ is

$$\sum_{j=1}^{r(p)} \left\{\left[\frac{2n}{p^j}\right] - 2\left[\frac{n}{p^j}\right]\right\} \leq \sum_{j=1}^{r(p)} 1 = r(p).$$

Hence,

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

□

LEMMA 4.12 If $p > 2$ and

$$\frac{2n}{3} < p \leq n,$$

then

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

Proof
If p satisfies

$$\frac{2n}{3} < p \leq n,$$

then p occurs once in the factorization of $n!$. This is because if $2p \leq n$, then

$$p \leq \frac{n}{2} < \frac{2n}{3} < p,$$

which is a contradiction to our assumption. Now p occurs twice in $(2n)!$ because $3p > 2n$. Therefore,

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

□

Erdös' proof of Bertrand's postulate
We are now ready to prove Bertrand's postulate.

Suppose that Bertrand's postulate is false. Then there exists a positive integer $n > 1$ such that there is no prime p in the interval $[n, 2n)$. By Lemma 4.12, all prime factors of

$$\binom{2n}{n}$$

must satisfy $p \leq 2n/3$. Let $s(p)$ be the largest prime power of p that divides $\binom{2n}{n}$. By Lemma 4.11,

$$\prod_{p \leq 2n/3} p^{s(p)} = \binom{2n}{n} \mid \prod_{p \leq 2n} p^{r(p)}.$$

Therefore, $s(p) \leq r(p)$ and

$$p^{s(p)} \leq p^{r(p)} \leq 2n \tag{4.24}$$

by (4.23). If $s(p) > 1$, then $p^{s(p)} \geq p^2$ and thus,

$$p < \sqrt{2n}$$

since $p^{s(p)} < 2n$. In other words, no more than $[\sqrt{2n}]$ primes occur in $\binom{2n}{n}$ with exponent larger than 1. Now,

$$\begin{aligned}
\binom{2n}{n} &= \prod_{p \leq 2n/3} p^{s(p)} = \prod_{\substack{p \leq 2n/3 \\ s(p) > 1}} p^{s(p)} \prod_{\substack{p \leq 2n/3 \\ s(p) = 1}} p^{s(p)} \\
&\leq \prod_{p < \sqrt{2n}} p^{s(p)} \prod_{p \leq 2n/3} p \\
&< (2n)^{[\sqrt{2n}]} 4^{[2n/3]},
\end{aligned}$$

by (4.24) and Lemma 4.7.

Next, since

$$(1+1)^{2n} = \binom{2n}{0} + \cdots + \binom{2n}{n} + \cdots + \binom{2n}{2n} < (2n+1)\binom{2n}{n},$$

we conclude that

$$\frac{4^n}{2n+1} \leq \binom{2n}{n} \leq ((2n)^{\sqrt{2n}} 4^{2n/3},$$

which implies that

$$4^{n/3} \leq (2n+1)^{\sqrt{2n}+1}.$$

Therefore,

$$n\frac{\ln 4}{3} < (\sqrt{2n}+1)\ln(2n+1).$$

The above inequality is true for only small values of n, for example, $n < 469$. This implies that for $n \geq 750$, Bertrand's postulate is true. For $n < 750$, we verify directly that Bertrand's postulate is true by observing that 3 is a prime between 2 and 4, 5 is a prime between 3 and 6, 7 is a prime between 5 and 10, 13 is a prime between 7 and 14, 23 is a prime between 13 and 26, 43 is a prime between 23 and 46, 83 is a prime between 43 and 85, 163 is a prime between 83 and 166, 317 is a prime between 163 and 326, 631 is a prime between 317 and 634.

□

4.7 The Bertrand Postulate (Ramanujan's proof)

In this section, we present Ramanujan's proof of Bertrand's Postulate.

Ramanujan's proof was mentioned in an interesting article by P. Erdös titled "Ramanujan and I". Erdös' proof of Theorem 4.10 was published around 1932 and it was Kalmar who asked Erdös to look up on Ramanujan's proof and that was the first time Erdös heard about Ramanujan.

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

LEMMA 4.13 For each positive real number x,

$$\psi(x) = \theta(x) + \theta(\sqrt{x}) + \theta(\sqrt[3]{x}) + \cdots. \tag{4.25}$$

Next, we will show that

LEMMA 4.14

$$\ln([x]!) = \psi(x) + \psi\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right) + \cdots . \tag{4.26}$$

Proof
The function

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

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

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

where we have used properties of $\Lambda(n)$ for the last equality. □

We will now establish a few equalities and inequalities.

LEMMA 4.15 For positive real number x, we have

$$\psi(x) - 2\psi(\sqrt{x}) = \theta(x) - \theta(\sqrt{x}) + \theta(\sqrt[3]{x}) - \cdots , \tag{4.27}$$

$$\ln[x]! - 2\ln[x/2]! = \psi(x) - \psi\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right) - \cdots , \tag{4.28}$$

$$\psi(x) - 2\psi\left(\sqrt{x}\right) \le \theta(x) \le \psi(x) \tag{4.29}$$

and

$$\begin{aligned}\psi(x) - \psi\left(\frac{x}{2}\right) &\le \ln[x]! - 2\ln[x/2]! \\ &\le \psi(x) - \psi\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right).\end{aligned} \tag{4.30}$$

Proof of (4.27)

This follows directly from (4.25). More precisely,

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

□

Proof of (4.28)

This follows from (4.26), namely,

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

□

Proof of (4.29)

Note that $\theta(x)$ is increasing. Hence, from (4.27),

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

Also, from (4.25),

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

□

Proof of (4.30)
This follows immediately from (4.28). □

LEMMA 4.16 Let x be a real number. Then

$$\ln[x]! - 2\ln[x/2]! > \frac{2}{3}x \quad \text{if } x > 750, \tag{4.31}$$

$$\ln[x]! - 2\ln[x/2]! < \frac{3}{4}x \quad \text{if } x > 3, \tag{4.32}$$

$$\psi(x) - \psi(\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right) > \frac{2}{3}x \quad \text{if } x > 750, \tag{4.33}$$

and

$$\psi(x) - \psi\left(\frac{x}{2}\right) < \frac{3}{4}x \quad \text{if } x > 3. \tag{4.34}$$

Proof of (4.31)

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

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

The function $\Gamma(x)$ satisfies the well-known Stirling's formula [1]

$$\ln\Gamma(x) = \ln\sqrt{2\pi} + \left(x - \frac{1}{2}\right)\ln x - x + \frac{\vartheta_x}{12x}, 0 < \vartheta_x < 1. \tag{4.35}$$

[1] See p. 24 of E. Artin's "The Gamma function".

A real-valued function $F(x)$ defined on an open interval $a < x < b$ is called convex if for every $y \in (a, b)$,

$$G(x, y) = \frac{F(x) - F(y)}{x - y}$$

is a monotonically increasing function of x. It is known that $\ln \Gamma(x)$ is convex. The convexity of $\ln \Gamma(x)$ implies that [2]

$$(\ln \Gamma(x))'' \geq 0,$$

which leads to

$$\Gamma''(x)\Gamma(x) \geq 0.$$

Since $\Gamma(x)$ is positive for $x > 0$ (see the definition of $\Gamma(x)$, we deduce that

$$\Gamma''(x) \geq 0.$$

Observe now that $\Gamma(3) > \Gamma(2)$ and thus, by mean value theorem, there is a $c \in [2, 3]$ such that $\Gamma'(c) > 0$. Since $\Gamma''(x) \geq 0$ for $x > 0$, we deduce that $\Gamma'(x) > 0$ for all $x \geq c$. In other words, $\Gamma(x)$ is increasing for $x > 3 > c$ and we conclude that

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

To prove (4.31), it suffices to show that for $x > 750$,

$$\ln \Gamma(x) - 2 \ln \Gamma\left(\frac{1}{2}x + 1\right) > \frac{2x}{3}. \tag{4.36}$$

By (4.35), we deduce that

$$\begin{aligned} &\ln \Gamma(x) - 2 \ln \Gamma\left(\frac{1}{2}x + 1\right) \\ &\quad = \ln \sqrt{2\pi} + \left(x - \frac{1}{2}\right) \ln x - x + \frac{\vartheta_1}{12x} - 2 \ln \sqrt{2\pi} \\ &\qquad - 2\left(\frac{x}{2} + \frac{1}{2}\right) \ln\left(\frac{x}{2} + 1\right) + 2\left(\frac{x}{2} + 1\right) - \frac{\vartheta_2}{3x + 6}, \end{aligned} \tag{4.37}$$

where both ϑ_1, ϑ_2 belong to the interval $(0, 1)$. Since

$$2 + \frac{\vartheta_1}{12x} - \frac{\vartheta_2}{3x + 6} \geq 1,$$

we find from (4.37) that

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

[2] See p. 4 of E. Artin's "The Gamma Function".

Using the fact that $-\ln\sqrt{2\pi}+1>0$, $-1/2>-1$ and that for $x>2$,

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

we find that

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

It suffices to show that for $x>750$,

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

But if we let

$$f(x)=\ln 2x-\ln(x+2)-2\frac{\ln x}{x},$$

then

$$f'(x)=\frac{1}{x}-\frac{1}{x+2}-\frac{2}{x^2}+2\frac{\ln x}{x^2}.$$

But

$$\frac{1}{x}-\frac{1}{x+2}>0$$

and

$$2\frac{\ln x}{x^2}-\frac{2}{x^2}>0$$

if $x>3$. Hence if $x>3$, then $f'(x)>0$. Therefore, $f(x)$ is increasing. In other words, if $x>750$, then

$$f(x)>f(750)=0.672\cdots>\frac{2}{3},$$

and the proof of (4.36) is complete. □

Proof of (4.32)

The proof is similar to that for (4.31). We use the inequality

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

and Stirling's formula to conclude that (why?)

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

for all $x>3$.

□

Proof of (4.33) *and* (4.34)

These two inequalities follow immediately from (4.30)-(4.32). □

LEMMA 4.17 For each positive real number x, we have

$$\psi(x) < \frac{3}{2}x \quad \text{if } x > 3 \tag{4.38}$$

$$\begin{aligned}\psi(x) - \psi\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right) &\leq \theta(x) + 2\psi(\sqrt{x}) - \theta\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right) \\ &< \theta(x) - \theta\left(\frac{x}{2}\right) + \frac{x}{2} + 3\sqrt{x}\,. \end{aligned} \tag{4.39}$$

Proof of (4.38)

To prove (4.38), we use (4.34) repeatedly, with x replaced by $x/2$, $x/4, \cdots$ and add up the results. We find that

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

□

***Proof of* (4.39).**

From (4.29), we find that

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

Hence

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

Next, from (4.29),

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

Using the above inequalities, we deduce that

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

For the second inequality, we use (4.38) to deduce that

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

We are now ready to prove Bertrand's Postulate. By (4.33),

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

for $x > 750$. Hence we deduce from (4.39) that

$$\theta(x) - \theta(x/2) \geq 2x/3 - x/2 - 3\sqrt{x} > 0$$

whenever $x > 750$. This implies that for $n > 375$, there is a prime between n and $2n$.

We are now left with verifying that Bertrand's Postulate is true for $2 \leq n \leq 375$. This is straightforward and we leave it as an exercise.

5 The Prime Number Theorem

5.1 The Prime Number Theorem

In Chapter 4, Corollary 4.6, we proved that the Prime Number Theorem is equivalent to the statement

$$\psi(x) \sim x. \tag{5.1}$$

In this chapter, we will prove the following theorem.

THEOREM 5.1 For positive real number x, we have

$$\psi(x) = x + O\left(x \exp(-c \sqrt[10]{\ln x})\right),$$

where $c > 0$ is some constant independent of x.

We note that (5.1) follows immediately from Theorem 5.1.

Theorem 5.1, which was mentioned on page 169 of Gérald Tenenbaum's book "Introduction to Analytic and Probabilistic Number Theory", is weaker than the result obtained independently by J. Hadamard and de la Valleé Poussin, which states that

$$\psi(x) = x + O\left(x \exp(-c\sqrt{\ln x})\right).$$

But the treatment here (adapted from A. Hildebrand's 1991 "Analytic Number Theory" notes) allows us to appreciate the analytic method used in the proofs of the Prime Number Theorem with less technicalities.

5.2 The Riemann zeta function

In Chapter 3, Definition 3.3, we have encountered the Riemann zeta function for real $s > 1$. We now give the definition of the function when s is a complex number.

DEFINITION 5.1 Let $s = \sigma + it \in \mathbf{C}$ and $\sigma > 1$. Define

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

THEOREM 5.2 The Riemann zeta function $\zeta(s)$ is an analytic function for $\sigma > 1$.

Proof

Note that if $\sigma \geq 1 + \delta$, then

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

Now, for every $\epsilon > 0$, there exists $N_\epsilon > 0$ such that

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

for $M > m > N_\epsilon$. Hence, we conclude that

$$\sum_{n=m}^{M} \left| \frac{1}{n^s} \right| < \epsilon$$

for $M > m > N_\epsilon$. Therefore, by the Weierstrass M-test, the series

$$\sum_{n=1}^{\infty} \frac{1}{n^s}$$

is absolutely and uniformly convergent in any region $\sigma \geq 1 + \delta$, with $\delta > 0$. The Riemann zeta function $\zeta(s)$ is therefore an analytic function in $\sigma > 1$. □

5.3 Euler's product and the product representation of $\zeta(s)$

THEOREM 5.3 For $\sigma > 1$,

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

The above follows immediately from the next theorem.

DEFINITION 5.2 An infinite product

$$\prod_{n=1}^{\infty}(1+a_n)$$

is said to be absolutely convergent if

$$\sum_{n=1}^{\infty}\ln(1+a_n)$$

is absolutely convergent.

THEOREM 5.4 Let f be a multiplicative arithmetical function such that the series

$$\sum_{n=1}^{\infty}f(n)$$

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

$$\sum_{n=1}^{\infty}f(n)=\prod_{p}(1+f(p)+f(p^2)+\cdots), \tag{5.2}$$

extended over all primes.

The product above is called the Euler product of the series.

Proof

Consider the finite product

$$P(x)=\prod_{p\le x}(1+f(p)+f(p^2)+\cdots)$$

extended over all primes $p\le x$. Since this is a product of a finite number of absolutely convergent series we can multiply the series and rearrange the terms without altering the sum. A typical term is of the form

$$\prod_{p}f(p^{\alpha})=f\left(\prod_{p}p^{\alpha}\right),$$

since f is multiplicative. By the fundamental theorem of arithmetic we can write

$$P(x)=\sum_{n\in A}f(n)$$

where A consists of those n having all their prime factors less than or equal to x. Therefore,

$$\sum_{n=1}^{\infty} f(n) - P(x) = \sum_{n \in B} f(n),$$

where B is the set of n having at least one prime factor greater than x. Therefore,

$$\left| \sum_{n=1}^{\infty} f(n) - P(x) \right| \leq \sum_{n \in B} |f(n)| \leq \sum_{n > x} |f(n)|.$$

Since

$$\sum_{n=1}^{\infty} |f(n)|$$

is convergent,

$$\lim_{x \to \infty} \sum_{n > x} |f(n)| = 0.$$

Hence,

$$\lim_{x \to \infty} P(x) = \sum_{n=1}^{\infty} f(n).$$

We have proved that the infinite product is convergent. We now establish the absolute convergence of the infinite product. A necessary and sufficient condition for the absolute convergence of the product

$$\prod_{n} (1 + a_n)$$

is the convergence of the series

$$\sum_{n} |a_n|.$$

In this case, we have

$$\sum_{p \leq x} |f(p) + f(p^2) + f(p^3) + \cdots| \leq \sum_{p \leq x} (|f(p)| + |f(p^2)| + \cdots) \leq \sum_{n=2}^{\infty} |f(n)|.$$

Since the partial sums are bounded, the series of positive terms

$$\sum_{p \leq x} |f(p) + f(p^2) + f(p^3) + \cdots|$$

converges, and this implies absolute convergence of the product (5.2). □

Applying Theorem 5.4 with

$$f(n) = \frac{1}{n^s},$$

we obtain Theorem 5.3.

5.4 Analytic continuation of $\zeta(s)$ to $\sigma > 0$

THEOREM 5.5 The Riemann zeta function $\zeta(s)$ can be extended to a function that is analytic in $\sigma > 0$, except at $s = 1$, where it has a simple pole with residue 1.

Proof

Recall from Theorem 3.2 that

$$\sum_{n \le x} f(n) = f(1) + \int_1^x f(t)dt + \int_1^x f'(t)\{t\}dt - \{x\}f(x).$$

With s real,

$$x = N \in \mathbf{N} \quad \text{and} \quad f(n) = \frac{1}{n^s},$$

we have

$$\sum_{n=1}^{N} \frac{1}{n^s} = 1 + \int_1^N \frac{d\eta}{\eta^s} - \int_1^N \frac{s\{\eta\}}{\eta^{s+1}} d\eta.$$

By analytic continuation, the above identity is also valid for complex numbers $s = \sigma + it$ with $\sigma > 1$.

Now, assume $\sigma > 1$. Then

$$\lim_{N\to\infty} \sum_{n=1}^{N} \frac{1}{n^s} = \zeta(s),$$

$$\lim_{N\to\infty} \int_1^N \frac{d\eta}{\eta^s} = \int_1^\infty \frac{d\eta}{\eta^s} = \frac{1}{s-1}$$

and

$$\lim_{N\to\infty} \int_1^N \frac{\{\eta\}}{\eta^{s+1}} d\eta = \int_1^\infty \frac{\{\eta\}}{\eta^{s+1}} d\eta =: \Phi(s).$$

Therefore,

$$\zeta(s) = 1 + \frac{1}{s-1} - s\Phi(s), \sigma > 1.$$

But, $\Phi(s)$ is analytic for $\sigma > 0$. Define, for $\sigma > 0$, the extension of $\zeta(s)$ by

$$1 + \frac{1}{s-1} - s\Phi(s).$$

Note that this function has a pole at $s = 1$.

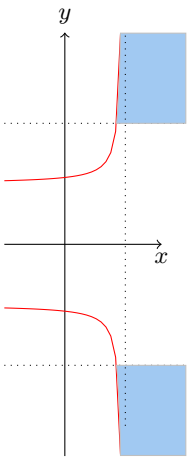


Figure 5.1 The shaded regions indicate the regions for which (5.4) and (5.5) hold.

To continue $\zeta(s)$ to $\sigma > -1$, we write

$$\Phi(s) = \int_1^\infty \frac{\{\eta\} - 1/2}{\eta^{s+1}} d\eta + \frac{1}{2s}.$$

This yields

$$\zeta(s) = \frac{s}{s-1} - s\int_1^\infty \frac{\{\eta\} - 1/2}{\eta^{s+1}} d\eta - \frac{1}{2}.$$

Incidently, this implies that $\zeta(0) = -1/2$.

□

5.5 Upper bounds for $|\zeta(s)|$ and $|\zeta'(s)|$ near $\sigma = 1$

THEOREM 5.6 Let A be a positive real number. If

$$|t| \geq 2 \quad \text{and} \quad \sigma \geq \max\left(\frac{1}{2}, 1 - \frac{A}{\ln|t|}\right), \tag{5.3}$$

then there are positive constants M and M' (depending on A) such that

$$|\zeta(s)| \leq M \ln|t| \tag{5.4}$$

$$|\zeta'(s)| \leq M' \ln^2|t|. \tag{5.5}$$

Proof

We recall that

$$\begin{aligned}\sum_{n=1}^{N}\frac{1}{n^s} &= 1+\int_1^N \frac{dx}{x^s} - s\int_1^N \frac{\{x\}}{x^{s+1}}dx\\ &= 1+\frac{N^{1-s}-1}{1-s} - s\Phi(s) + s\int_N^\infty \frac{\{x\}}{x^{s+1}}dx\\ &= \zeta(s)+\frac{N^{1-s}}{1-s}+s\int_N^\infty \frac{\{x\}}{x^{s+1}}dx.\end{aligned}$$

The above identity holds for $\sigma > 0$, where now,

$$\zeta(s) = 1+\frac{1}{s-1} - s\Phi(s), \sigma > 0,$$

with

$$\Phi(s) = \int_1^\infty \frac{\{\eta\}}{\eta^{s+1}}d\eta.$$

Now,

$$\begin{aligned}|\zeta(s)| &\le \sum_{n=1}^{N}\frac{1}{n^\sigma} + \frac{N^{1-\sigma}}{|1-s|} + |s|\int_N^\infty \frac{dx}{x^{\sigma+1}}\\ &\le \sum_{n=1}^{N}\frac{1}{n^\sigma} + \frac{N^{1-\sigma}}{|t|} + \frac{|s|}{\sigma N^\sigma}.\end{aligned}$$

Assume that s is in the region specified by (5.3) and let $N = [|t|]$. Then

$$N^{1-\sigma} \le \exp\left\{\frac{A\ln N}{\ln|t|}\right\} \le \exp(A).$$

This implies that

$$\begin{aligned}|\zeta(s)| &\le \sum_{n\le|t|}\frac{1}{n^\sigma} + \frac{e^A}{|t|} + \frac{|s|}{\sigma N}e^A\\ &\le \sum_{n\le|t|}\frac{1}{n^\sigma} + \frac{e^A}{2} + \frac{(\sigma+|t|)e^A}{\sigma(|t|/2)}\\ &\le \sum_{n\le|t|}\frac{1}{n^\sigma} + e^A\left(\frac{1}{2}+\frac{2}{|t|}+\frac{2}{\sigma}\right).\end{aligned} \tag{5.6}$$

Since $\sigma > 1/2$ and $|t| \ge 2$, we find that

$$\frac{1}{2}+\frac{2}{|t|}+\frac{2}{\sigma} < 6.$$

This shows that (5.6) may be written as

$$|\zeta(s)| \le \sum_{n\le|t|}\frac{1}{n^\sigma} + 6e^A. \tag{5.7}$$

For $\sigma \geq 1$,

$$\sum_{n \leq |t|} \frac{1}{n^\sigma} \leq \sum_{n \leq |t|} \frac{1}{n} = \ln |t| + O(1). \tag{5.8}$$

For

$$\max\left(\frac{1}{2}, 1 - \frac{A}{\ln |t|}\right) < \sigma < 1,$$

and $n \leq N$, we find that

$$\frac{1}{n^\sigma} \leq \frac{1}{n} n^{1-\sigma} \leq \frac{1}{n} N^{1-\sigma} \leq \frac{e^A}{n}.$$

Hence,

$$\sum_{n \leq |t|} \frac{1}{n^\sigma} \leq e^A \sum_{n \leq |t|} \frac{1}{n} = e^A \left(\ln |t| + u(t)\right), \tag{5.9}$$

where $u(t) = O(1)$.

Combining (5.8) and (5.9), we conclude that if s is in the region specified by (5.3), then

$$|\zeta(s)| \leq M \ln |t|,$$

where M is a positive constant depending on A. This proves (5.4).

We now prove (5.5). Differentiating the expression

$$\zeta(s) = \sum_{n \leq N} \frac{1}{n^s} - \frac{N^{1-s}}{1-s} - s \int_N^\infty \frac{\{t\}}{t^{s+1}}\, dt,$$

we deduce that

$$|\zeta'(s)| \leq \sum_{n \leq N} \frac{\ln n}{n^\sigma} + \frac{N^{1-\sigma} \ln N}{|1-s|} + \frac{N^{1-\sigma}}{|1-s|^2} + \int_N^\infty \frac{1}{t^{\sigma+1}} dt + |s| \int_N^\infty \frac{\ln t}{t^{\sigma+1}} dt$$

Let $N = [|t|]$, then from the above inequality, we find that

$$|\zeta'(s)| \leq \ln[|t|] \sum_{n \leq [|t|]} \frac{1}{n^\sigma} + \frac{e^A \ln |t|}{|1-s|} + \frac{e^A}{|1-s|^2} + \frac{1}{\sigma |t|^\sigma} + (\sigma + |t|) \int_N^\infty \frac{\ln t}{t^{\sigma+1}}$$

$$\leq \ln^2 |t| + C \ln |t| + e^A \left(\frac{\ln |t|}{|t|} + \frac{1}{|t|^2}\right) + \frac{|t|^{1-\sigma}}{\sigma |t|} + (\sigma + |t|) \left(\frac{\ln |t|}{\sigma [|t|]^\sigma} + \frac{1}{\sigma^2 [|t|]^\sigma}\right),$$

which leads to

$$|\zeta'(s)| \leq M' \ln^2 |t|,$$

after similar estimates as in the first case. □

5.6 The non-vanishing of $\zeta(1+it)$

THEOREM 5.7 For real number $t \neq 0$,

$$\zeta(1+it) \neq 0.$$

We first prove several simple lemmas.

LEMMA 5.8 For all $\theta \in \mathbf{R}$,

$$3 + 4\cos\theta + \cos 2\theta \geq 0.$$

Proof
The inequality follows immediately from the following computations:

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

□

LEMMA 5.9 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 representation of $\zeta(s)$, we find that

$$\begin{aligned} \zeta(s) &= \prod_p \left(1 - \frac{1}{p^s}\right)^{-1} \\ &= \exp\left(-\sum_p \ln\left(1 - \frac{1}{p^s}\right)\right) \\ &= \exp\left(\sum_p \sum_{m=1}^{\infty} \frac{1}{mp^{sm}}\right) = \exp(G(s)). \end{aligned}$$

□

LEMMA 5.10 For $\sigma > 1$, and all $t \in \mathbf{R}$,

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

Proof
By Lemma 5.9, we have for $\sigma > 1$,

$$\begin{aligned}\zeta(s) &= \exp\left(\sum_p \sum_{m=1}^{\infty} \frac{1}{mp^{ms}}\right) \\ &= \exp\left(\sum_p \sum_{m=1}^{\infty} \frac{1}{m} \exp\{-(\ln p)ms\}\right) \\ &= \exp\left(\sum_p \sum_{m=1}^{\infty} \frac{1}{m} \exp\{-m\sigma \ln p - itm \ln p\}\right),\end{aligned}$$

since $s = \sigma + it$. Hence,

$$\zeta(s) = \exp\left(\sum_p \sum_{m=1}^{\infty} \frac{1}{m} \frac{1}{p^{\sigma m}} \{\cos(tm \ln p) - i \sin(tm \ln p)\}\right).$$

Therefore,

$$|\zeta(s)| = \exp\left(\sum_p \sum_{m=1}^{\infty} \frac{1}{m} \frac{1}{p^{\sigma m}} \cos(tm \ln p)\right).$$

This implies that

$$\begin{aligned}&|\zeta(\sigma))|^3|\zeta(\sigma+it)|^4|\zeta(\sigma+2it)| \\ &\quad = \exp\left(\sum_p \sum_{m=1}^{\infty} \frac{1}{mp^{\sigma m}} (3 + 4\cos(tm \ln p) + \cos(2tm \ln p))\right) \\ &\quad \geq \exp(0) = 1.\end{aligned}$$

□

Proof of Theorem 5.7.

Suppose $\zeta(1 + it_0) = 0$ for some $t_0 \neq 0$. By Lemma 5.10, we deduce that for $\sigma > 1$,

$$|\zeta(\sigma)(\sigma-1)|^3 \left|\frac{\zeta(\sigma+it_0)}{\sigma-1}\right|^4 |\zeta(\sigma+2it_0)||\sigma-1| \geq 1 \tag{5.10}$$

Now, since $\zeta(\sigma)$ has a simple pole with residue 1 at $\sigma = 1$, we find that

$$\lim_{\sigma \to 1^+} \zeta(\sigma)(\sigma - 1) = 1. \tag{5.11}$$

Next,

$$\zeta(\sigma + it_0) = \zeta(1 + it_0) + (\sigma - 1)\zeta'(1 + it_0) + O((\sigma - 1)^2).$$

This implies that

$$\lim_{\sigma \to 1^+} \frac{\zeta(\sigma + it_0)}{\sigma - 1} = \zeta'(1 + it_0). \tag{5.12}$$

It is clear that

$$\zeta(\sigma + 2it_0) \to \zeta(1 + 2it_0). \tag{5.13}$$

Combining (5.11)–(5.13), we find that when σ approaches 1 from the right, $0 \geq 1$, which is clearly impossible. Hence, we conclude that

$$\zeta(1 + it) \neq 0$$

for all nonzero real t. □

5.7 A lower bound for $|\zeta(s)|$ near $\sigma = 1$

THEOREM 5.11 For $|t| \geq 2$, there exist positive constants c and d such that for

$$\sigma > 1 - \frac{c}{(\ln |t|)^9},$$

we have

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

Proof
For $\sigma \geq 2$,

$$\begin{aligned} |\zeta(s)| &= \left| \sum_{n=1}^{\infty} \frac{1}{n^s} \right| \geq 1 - \left| \sum_{n=2}^{\infty} \frac{1}{n^s} \right| \\ &\geq 1 - \sum_{n=2}^{\infty} \frac{1}{n^2} = 2 - \frac{\pi^2}{6} > \frac{1}{4}. \end{aligned}$$

Therefore, for $\sigma \geq 2$,

$$|\zeta(s)| \geq \frac{d}{(\ln |t|)^7},$$

provided that

$$d \leq \frac{\ln^7(2)}{4} \quad \text{and} \quad |t| \geq 2. \tag{5.14}$$

For $\delta > 0$, let

$$1 + \frac{\delta}{(\ln|t|)^9} \le \sigma \le 2, \quad |t| \ge 2.$$

By Lemma 5.10, we find that

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

Now, if $\sigma \le 2$,

$$\begin{aligned}\zeta(\sigma) = \sum_{n=1}^{\infty} \frac{1}{n^\sigma} &\le 1 + \int_1^\infty \frac{1}{x^\sigma} dx = 1 + \frac{1}{\sigma - 1} \\ &\le \frac{2}{\sigma - 1} \\ &\le \frac{2}{\delta}(\ln|t|)^9, \end{aligned} \tag{5.15}$$

since

$$\sigma \ge 1 + \frac{\delta}{(\ln|t|)^9}.$$

Suppose

$$\delta < \frac{\ln^9 2}{2}. \tag{5.16}$$

Then for $|t| > 2$, we have

$$\frac{1}{\ln|t|} < \frac{1}{\ln 2}$$

and therefore,

$$\frac{\delta}{\ln^9|t|} \le \frac{1}{2}\frac{\ln^9 2}{\ln^8|t| \ln|t|} < \frac{1}{2}\frac{\ln 2}{\ln|t|}.$$

In other words, if δ satisfies (5.16) and

$$\sigma > 1 - \frac{\delta}{\ln^9|t|},$$

we must have

$$\sigma > 1 - \frac{1}{2}\frac{\ln 2}{\ln|t|}. \tag{5.17}$$

By Theorem 5.6 with $A = \frac{1}{2}\ln 2$, we can find a constant $M > 0$ $\frac{1}{2}\ln 2$, such that

$$|\zeta(\sigma + 2it)| \le 2M \ln|t|. \tag{5.18}$$

Hence, by (5.18) and (5.15), we conclude that

$$|\zeta(\sigma+it)| \geq \left(\frac{\delta}{2\ln^9|t|}\right)^{3/4}\left(\frac{1}{2M\ln|t|}\right)^{1/4}$$
$$= \frac{\delta^{3/4}}{2M^{1/4}\ln^7|t|} \geq \frac{d}{\ln^7|t|},$$

for

$$d \leq \frac{\delta^{3/4}}{2M^{1/4}}. \tag{5.19}$$

Next, consider

$$1-\frac{\delta}{\ln^9|t|} \leq \sigma \leq 1+\frac{\delta}{\ln^9|t|}, \quad |t| \geq 2. \tag{5.20}$$

If

$$\sigma_0 = 1+\frac{\delta}{\ln^9|t|},$$

then we observe from (5.20) that

$$|\sigma-\sigma_0| \leq \frac{2\delta}{\ln^9|t|}. \tag{5.21}$$

We next show that $\zeta(\sigma+it)$ is close to $\zeta(\sigma_0+it)$.

$$|\zeta(\sigma+it)-\zeta(\sigma_0+it)| = \left|\int_\sigma^{\sigma_0}\zeta'(u+it)du\right|$$
$$\leq |\sigma-\sigma_0|\max_{\sigma\leq u\leq\sigma_0}|\zeta'(u+it)|.$$

Now, by Theorem 5.6, there exists an $M'>0$ such that

$$|\zeta'(u+it)| \leq M'\ln^2|t|,$$

for

$$|u| \geq \sigma \geq 1-\frac{1}{2}\frac{\ln 2}{\ln|t|} \quad \text{and} \quad |t| \geq 2.$$

Therefore,

$$|\zeta(\sigma+it)-\zeta(\sigma_0+it)| \leq \frac{2\delta}{\ln^7|t|}M',$$

where we have used (5.21) in the final inequality. M' is independent of δ provided

$$\frac{\delta}{\ln^9|t|} \leq \frac{1}{2}\frac{\ln 2}{\ln|t|}, |t| \geq 2.$$

This will be satisfied if

$$\delta \leq \frac{1}{2}\ln^9 2.$$

Hence,

$$\begin{aligned}|\zeta(\sigma+it)| &\geq |\zeta(\sigma_0+it)| - |\zeta(\sigma+it)-\zeta(\sigma_0+it)| \\ &\geq \frac{\delta^{3/4}}{2M^{1/4}\ln^7|t|} - \frac{2\delta}{\ln^7|t|}M' \\ &= \frac{\delta^{3/4}}{\ln^7|t|}\left(\frac{1}{2M^{1/4}} - 2\delta^{1/4}M'\right).\end{aligned}$$

We now choose a real positive number $\delta = \delta_1$ be such that

$$\left(\frac{1}{2M^{1/4}} - 2\delta_1^{1/4}M'\right) > 0.$$

Now, letting

$$0 < c < \min\left(\frac{1}{2}\ln^9(2), \delta_1\right)$$

and

$$0 < d < \min\left(\delta_1^{3/4}\left(\frac{1}{2M^{1/4}} - 2\delta_1^{1/4}M'\right), \frac{\ln^7(2)}{4}, \frac{\delta_1^{3/4}}{2M^{1/4}}\right),$$

we conclude that for $|t| \geq 2$ and $\sigma > 1 - \dfrac{c}{\ln^9|t|}$,

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

□

5.8 Perron's Formula

THEOREM 5.12 Let x be a half integer. Then for any $b \in [1,3]$ and any $T \geq 1$,

$$\psi(x) = \frac{1}{2\pi i}\int_{b-iT}^{b+iT}\left(-\frac{\zeta'}{\zeta}(s)\frac{x^s}{s}\right)ds + O\left(\frac{x^b}{T(b-1)} + x\frac{\ln^2 x}{T}\right).$$

We first begin with several lemmas.

LEMMA 5.13 For $\sigma > 1$,

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

Proof
The proof is immediate using the formula

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

and the fact that

$$\sum_{n=1}^{\infty} \frac{f*g(n)}{n^s} = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \sum_{n=1}^{\infty} \frac{g(n)}{n^s}.$$

□

LEMMA 5.14 For $\sigma > 1$,

$$\left|\frac{\zeta'}{\zeta}(s)\right| \ll \frac{1}{\sigma - 1} + 1.$$

Proof
For $\sigma > 1$,

$$\begin{aligned}\left|\frac{\zeta'}{\zeta}(s)\right| &\leq \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^\sigma} = \sigma \int_1^{\infty} \sum_{n \leq t} \Lambda(n) \frac{d\eta}{\eta^{\sigma+1}} \\ &\leq \sigma \int_1^{\infty} \frac{c\eta}{\eta^{\sigma+1}} d\eta, \quad \text{by Theorem 4.2,} \\ &= c\frac{\sigma}{\sigma - 1} \ll 1 + \frac{1}{\sigma - 1}.\end{aligned}$$

□

LEMMA 5.15 For $b > 0$, $T \geq 1$, and $y > 0$, $y \neq 1$, we have

$$\frac{1}{2\pi i} \int_{b-iT}^{b+iT} \frac{y^s}{s} ds = \begin{cases} 1 + O\left(\dfrac{y^b}{T|\ln y|}\right) & \text{if } y > 1 \\ O\left(\dfrac{y^b}{T|\ln y|}\right) & \text{if } 0 < y < 1 \end{cases}$$

Proof
We will only prove the result when $y > 1$. By the Residue Theorem,

$$\frac{1}{2\pi i} \int_{b-iT}^{b+iT} \frac{y^s}{s} ds = \frac{1}{2\pi i} \int_{\Gamma_0} \frac{y^s}{s} ds = 1 + \sum_{j=1}^{3} \frac{1}{2\pi i} \int_{\Gamma_j} \frac{y^s}{s} ds.$$

Thus, it suffices to show that with $-a$ large enough,

$$\left|\int_{\Gamma_j} \frac{y^s}{s} ds\right| \ll \frac{y^b}{T|\ln y|},$$

with $j = 1, 2, 3$.

On Γ_2,

$$\left|\frac{y^s}{s}\right| = \frac{y^a}{|s|} \le y^a,$$

if $a \le -1$. This implies that

$$\left|\int_{\Gamma_2} \frac{y^s}{s} ds\right| \le y^a 2T.$$

Letting a approaches $-\infty$, we conclude that the above integral is 0.

On Γ_1 and Γ_3,

$$\left|\frac{y^s}{s}\right| = \frac{y^\sigma}{|s|} \le \frac{y^\sigma}{T},$$

since

$$|s| > |T|.$$

Hence, for $j = 1$ or 3,

$$\left|\int_{\Gamma_j} \frac{y^s}{s} ds\right| \le \int_a^b \frac{y^\sigma}{T} d\sigma \le \frac{1}{T}\int_{-\infty}^b e^{\sigma \ln y} d\sigma \ll \frac{y^b}{T|\ln y|}.$$

For the case $0 < y < 1$, we will leave it as exercise for the reader. □

Proof of Theorem 5.12.

Let

$$I = \frac{1}{2\pi i}\int_{b-iT}^{b+iT} \left(-\frac{\zeta'}{\zeta}(s)\right)\frac{x^s}{s} ds. \tag{5.22}$$

By Lemmas 5.13 and 5.15, we find that

$$\begin{aligned} I &= \frac{1}{2\pi i}\int_{b-iT}^{b+iT} \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}\frac{x^s}{s} ds \\ &= \sum_{n=1}^{\infty} \Lambda(n)\frac{1}{2\pi i}\int_{b-iT}^{b+iT} \frac{\left(\frac{x}{n}\right)^s}{s} ds \\ &= \sum_{n\le x} \Lambda(n) + \sum_{n=1}^{\infty} \Lambda(n) O\left(\frac{\left(\frac{x}{n}\right)^b}{T|\ln\frac{x}{n}|}\right), \\ &= \psi(x) + O\left(\frac{x^b}{T}\sum_{n=1}^{\infty}\frac{\Lambda(n)}{n^b|\ln\frac{x}{n}|}\right). \end{aligned}$$

Let

$$R = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^b |\ln \frac{x}{n}|}.$$

Then

$$R = \sum_{\frac{x}{2} \le n \le 2x} \frac{\Lambda(n)}{n^b |\ln \frac{x}{n}|} + \sum_{n \notin [\frac{x}{2}, 2x]} \frac{\Lambda(n)}{n^b |\ln \frac{x}{n}|} = R_1 + R_2.$$

Note that if $n > 2x$ or $n < x/2$ then $|\ln(x/n)| \ge \ln 2$. Furthermore, since $1 < b \le 3$, by Lemmas 5.13 and 5.14,

$$R_2 \le \frac{1}{\ln 2} \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^b} \ll 1 + \frac{1}{b-1} \le \frac{1}{b-1} + \frac{2}{3-1} \ll \frac{1}{b-1}.$$

Now, if

$$-\frac{1}{2} \le t < 1,$$

then

$$|\ln(1+t)| \ge \frac{|t|}{2}$$

and we deduce that

$$\left| \ln \frac{x}{n} \right| = \left| \ln \frac{n}{x} \right| = \left| \ln \left(1 + \frac{n-x}{x} \right) \right| \gg \left| \frac{n-x}{x} \right|. \tag{5.23}$$

Furthermore, since

$$\Lambda(n) \le \ln x \tag{5.24}$$

and

$$\frac{1}{n^b} \le \frac{2^b}{x^b} \tag{5.25}$$

for $x/2 < n$. Using (5.23)–(5.25), with the observations that

$$2^b \le 2^3 \quad \text{and} \quad |n| \le |x|,$$

we find that

$$R_1 = \sum_{\frac{x}{2} \le n \le 2x} \frac{\Lambda(n)}{n^b |\ln \frac{x}{n}|} \ll \frac{\ln x}{x^b} \sum_{\frac{x}{2} \le n \le 2x} \left| \frac{x}{x-n} \right|. \tag{5.26}$$

Since x is a half integer, the denominator in the sum

$$\sum_{\frac{x}{2} \le n \le 2x} \left| \frac{x}{x-n} \right|$$

is nonzero and we find that

$$\sum_{\frac{x}{2} \le n \le 2x} \left| \frac{x}{x-n} \right| \ll x \ln x. \tag{5.27}$$

Substituting (5.27) into (5.26), we conclude that

$$R_1 \ll \frac{\ln^2 x}{x^b} x.$$

Hence, the error term for I, given by (5.22), is

$$O\left(\frac{x^b}{T(b-1)} + x\frac{\ln^2 x}{T}\right).$$

□

5.9 Completion of the proof of the Prime Number Theorem

In this section, we can finally complete the proof of the prime number theorem.

Proof

Step 1.

Application of Perron's Formula:
Let

$$T \geq 1,\ x = N + \frac{1}{2} \geq 2 \quad \text{and} \quad b = 1 + \frac{1}{\ln x}.$$

Then

$$\psi(x) = \frac{1}{2\pi i}\int_{b-iT}^{b+iT}\left(-\frac{\zeta'}{\zeta}(s)\right)\frac{x^s}{s}ds + O\left(\frac{x\ln^2 x}{T}\right).$$

Step 2.

Shifting of path of integration:
Choose a sufficiently close to 1 so that

$$\zeta(s) \neq 0$$

for all $\sigma \geq a$, $|t| \leq T$. We note that the integrand

$$-\frac{\zeta'}{\zeta}(s)\frac{x^s}{s}$$

is analytic in the region enclosed by the old and new paths with an exception of a pole at $s = 1$, with residue x. By the Residue Theorem,

$$\frac{1}{2\pi i}\int_{b-iT}^{b+iT}\left(-\frac{\zeta'}{\zeta}(s)\right)\frac{x^s}{s}ds = x + \sum_{j=1}^{3}\frac{1}{2\pi i}\int_{\Gamma_j}\left(-\frac{\zeta'}{\zeta}(s)\right)\frac{x^s}{s}ds.$$

Step 3.

Estimation of $\displaystyle\int_{\Gamma_j}\left(-\frac{\zeta'}{\zeta}(s)\right)\frac{x^s}{s}\,ds$:
Let

$$B=\max_{s\in\Gamma_1,\Gamma_2,\Gamma_3}\left|\frac{\zeta'}{\zeta}(s)\right|.$$

The number B depends on T and will be estimated in **Step 4**.

Now, for $T\geq 2$,

$$\begin{aligned}\left|\int_{\Gamma_2}\left(-\frac{\zeta'}{\zeta}(s)\right)\frac{x^s}{s}\,ds\right| &\leq x^a B\int_{a-iT}^{a+iT}\frac{|ds|}{|s|}\\ &= 2x^aB\int_0^T\frac{dt}{|a+it|}\\ &\ll Bx^a\ln T. \end{aligned}\tag{5.28}$$

The last inequality follows from the fact that for $T\geq 2$,

$$\int_0^T\frac{dt}{|a+it|}\leq\int_1^T\frac{dt}{t}+\int_0^1\frac{dt}{a}\leq\ln T+2\ll\ln T.$$

We will now estimate the integral on Γ_3. The estimate of the integral on Γ_1 is similar. Since

$$b=1+\frac{1}{\ln x},$$

we find that

$$\begin{aligned}\left|\int_{\Gamma_3}\left(-\frac{\zeta'}{\zeta}(s)\right)\frac{x^s}{s}\,ds\right| &\leq \frac{B}{T}\int_a^b x^\sigma\,d\sigma\\ &\ll\frac{Bx^b}{T}\ll\frac{Bx}{T}.\end{aligned}\tag{5.29}$$

We therefore conclude from (5.28) and (5.29) that

$$\psi(x)=x+O\left(\frac{Bx}{T}\right)+O(Bx^a\ln T)+O\left(\frac{x\ln^2 x}{T}\right).$$

We note that the above holds for $T\geq 2$ and a suitable choice of a.

Step 4.

Choice of a and estimation of B:
For $|t|\leq 2$, we note that $\zeta(s)\neq 0$ for $s=1+it$. Therefore, there exists a $\delta>0$ such that for $|t|\leq 2$ and $\sigma>1-\delta$,

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

is analytic and bounded there. This implies that

$$\left|\frac{\zeta'}{\zeta}(s)\right|\ll\frac{1}{1-\sigma}\ll\ln^9 T\tag{5.30}$$

if $\sigma = 1 - \dfrac{A}{\ln^9 T}$.

Suppose $2 \le |t| \le T$. Then by Theorems 5.6 and 5.7, there exist c and d such that

$$|\zeta(s)| \ge \frac{d}{\ln^7 |t|} \quad \text{and} \quad |\zeta'(s)| \ll \ln^2 |t|$$

in the region

$$\sigma \ge 1 - \frac{c}{\ln^9 |t|}.$$

Note that we must choose c so that

$$c < \delta \ln^9 2. \tag{5.31}$$

The additional condition imposed on c is necessary for the validity of (5.30). Next, with $2 \le |t| \le T$, and

$$a = 1 - \frac{c}{\ln^9 |T|},$$

we conclude that

$$\left|\frac{\zeta'}{\zeta}(s)\right| \ll \ln^9 T.$$

Together with (5.30), we find that

$$B = \max_{s \in \Gamma_1, \Gamma_2, \Gamma_3} \left|\frac{\zeta'}{\zeta}(s)\right| \ll \ln^9 T.$$

Therefore,

$$\psi(x) = x + O\left(x \frac{\ln^9 T}{T}\right) + O\left(\frac{x \ln^2 x}{T}\right) + O\left(x \ln^{10} T \exp\left(-c \frac{\ln x}{\ln^9 T}\right)\right).$$

Now the first two error terms can be bounded by

$$O\left(x \frac{\ln^{10} x}{T}\right).$$

Hence

$$\psi(x) = x + O\left(x \frac{\ln^{10} x}{T}\right) + O\left(x \ln^{10} T \exp\left(-c \frac{\ln x}{\ln^9 T}\right)\right).$$

Step 5.

Choice of T:
Assume $2 \le T \le x$. The expression in the error term is minimal if

$$\frac{1}{T} = \exp\left\{-\frac{c \ln x}{\ln^9 T}\right\}.$$

Therefore,

$$T = \exp\{c_3^{1/10} \ln^{1/10} x\}.$$

With the choice of T, we have for sufficiently large $x \geq x_0$, $2 \leq T \leq x$,

$$\psi(x) = x + O\left(x \frac{\ln^{10} x}{\exp(c^{1/10} \ln^{1/10} x)}\right).$$

Since for any $\epsilon > 0$,

$$\ln^{10} x \ll \exp\left(\epsilon \ln^{1/10} x\right),$$

we conclude that

$$\psi(x) = x + O(x \exp(-c' \ln^{1/10} x))$$

with a suitable choice of $c' > 0$. For $2 \leq x \leq x_0$, we have

$$\psi(x) = x + O(x \exp(-c' \ln^{1/10} x)).$$

This completes the proof of the Prime Number Theorem. □

5.10 Prime Number Theorem without error term

In this section, we present J. Korevaar's proof of the Prime Number Theorem. His proof is a modification of D.J. Newman's proof. The section is adapted from Chapter 7 of "*Complex Analysis*" by R.B. Ash and W.P. Novinger.

THEOREM 5.16 (Auxiliary Tauberian Theorem) Let F be bounded and piecewise continuous on $[0, \infty)$ so that

$$G(z) = \int_0^\infty F(t) e^{-zt}\, dt$$

exists and is analytic on $\operatorname{Re} z > 0$.

Assume that G has an analytic continuation to a neighborhood of the imaginary axis $\operatorname{Re} z = 0$. Then $\displaystyle\int_0^\infty F(t)\, dt$ exists as an improper integral and is equal to $G(0)$.

A corollary of Theorem 5.16 is

COROLLARY 5.17 Let f be a non-negative, piecewise continuous and non-decreasing function on $[0, \infty)$ such that

$$f(x) = O(x).$$

Then its Mellin's transform

$$g(z) = z \int_1^\infty f(x) x^{-z-1}\, dx$$

exists for $\operatorname{Re} z > 1$ and defines an analytic function g. Assume that for some constant c, the function

$$g(z) - \frac{c}{z-1}$$

has an analytic continuation to a neighborhood of the line $\operatorname{Re} z = 1$. Then

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

Before we prove the main theorem and the corollary, we first deduce Prime Number Theorem from Corollary 5.17. Let $f(x) = x$ in the corollary. By Chebyschev's Theorem,

$$\psi(x) = O(x).$$

Note that

$$g(z) = z \int_1^\infty \psi(x) x^{-z-1}\, dx = -\frac{\zeta'(z)}{\zeta(z)}.$$

The function

$$g(z) - \frac{1}{z-1}$$

is analytic on $\operatorname{Re} z = 1$. Therefore, by Corollary 5.17,

$$\psi(x) \sim x$$

and Prime Number Theorem is true.

We now prove Corollary 5.17 using Theorem 5.16.

Proof of Corollary 5.17

Let $F(t) = e^{-t} f(e^t) - c$. Note that $f(x) = O(x)$ implies that $F(x)$ is bounded. Consider the integral

$$G(z) = \int_0^\infty (e^{-t} f(e^t) - c) e^{-zt}\, dt.$$

Set $x = e^t$. Then

$$G(z) = \int_1^\infty \left(\frac{f(x)}{x} - c\right) x^{-z-1}\, dx = \frac{g(z+1)}{z+1} - \frac{c}{z} = \frac{1}{z+1}\left(g(z+1) - \frac{c}{z} - c\right).$$

From the hypothesis, $G(z)$ has an analytic continuation to a neighborhood of the line $\operatorname{Re} z = 0$. Therefore, $\displaystyle\int_0^\infty F(t)\, dt$ exists and converges to $G(0)$. In terms of f, this implies that

$$\int_0^\infty (e^{-t} f(e^t) - c)\, dt$$

exists, or

$$\int_1^\infty \left(\frac{f(x)}{x} - c\right) \frac{dx}{x}$$

exists.

We need to show that

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

Suppose that

$$\overline{\lim} \frac{f(x)}{x} > c.$$

Then there exists a $\delta > 0$ such that $c+2\delta$ is not an upper bound of $f(x)/x$. This implies that

$$\frac{f(y)}{y} > (c + 2\delta)$$

for some positive real number y. Let $\rho = \dfrac{c+2\delta}{c+\delta} > 1$. Then for $y < x < \rho y$,

$$(c + \delta)x < (c + 2\delta)y < f(y) < f(x).$$

Here,

$$\int_y^{\rho y} \left(\frac{f(t)}{t} - c\right) \frac{dt}{t} \geq \int_y^{\rho y} \frac{\delta}{t}\, dt = \delta \ln\left(\frac{c+2\delta}{c+\delta}\right).$$

This implies that

$$\int_1^\infty \left(\frac{f(t)}{t} - c\right) \frac{dt}{t}$$

is not convergent. Therefore,

$$\overline{\lim} \frac{f(x)}{x} \leq c.$$

Similarly,

$$\underline{\lim} \frac{f(x)}{x} \geq c$$

and we must have

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

□

We now prove Theorem 5.16.

Proof of Theorem 5.16

Since F is bounded, $|F| \leq M$. By replacing F with F/M, we may assume that

$$|F(t)| \leq 1$$

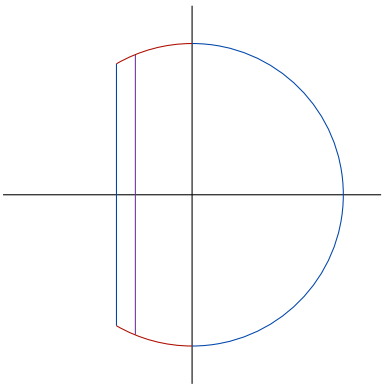
for $t \geq 0$. For $0 < \lambda < \infty$, define

$$G_\lambda(z) = \int_0^\lambda F(t)e^{-zt}\,dt.$$

We want to show that

$$\lim_{\lambda\to\infty} G_\lambda(0) = G(0).$$

Consider the following contour:



Denote the contour on the right of y-axis by γ_R^+ and the contour on the left by γ_R^-. Denote the union of the contour by γ_R. We suppose the straight line on the left is given by $x = -\delta(R)$ with $\delta(R) > 0$. Note that by Cauchy's integral formula

$$G(0) - G_\lambda(0) = \frac{1}{2\pi i}\int_{\gamma_R} (G(z) - G_\lambda(z))\, e^{\lambda z}\left(\frac{1}{z} + \frac{z}{R^2}\right) dz.$$

If $z \in \gamma_R^+$, then

$$\frac{1}{z} + \frac{z}{R^2} = 2\frac{\operatorname{Re} z}{R^2}.$$

Furthermore, since $|F| \leq 1$,

$$|G(z) - G_\lambda(z)| = |\int_\lambda^\infty F(t)e^{-zt}\,dt| \leq \int_\lambda^\infty e^{-xt}dt = \frac{e^{-\lambda x}}{x}.$$

This implies that

$$\left|(G(z) - G_\lambda(z))\, e^{\lambda z}\left(\frac{1}{z} + \frac{z}{R^2}\right)\right| \leq \frac{2}{R^2}.$$

Therefore,

$$\left|\frac{1}{2\pi i}\int_{\gamma_R^+} (G(z) - G_\lambda(z))\, e^{\lambda z}\left(\frac{1}{z} + \frac{z}{R^2}\right) dz\right| \leq \frac{1}{R}$$

and the integral tends to 0 as R tends to ∞.

Next, on γ_R^-, we have

$$\begin{aligned}\left|\frac{1}{2\pi i}\int_{\gamma_R^-}(G(z)-G_\lambda(z))\,e^{\lambda z}\left(\frac{1}{z}+\frac{z}{R^2}\right)\,dz\right| &\le \left|\frac{1}{2\pi i}\int_{\gamma_R^-}G(z)e^{\lambda z}\left(\frac{1}{z}+\frac{z}{R^2}\right)\,dz\right| \\ &\quad+\left|\frac{1}{2\pi i}\int_{\gamma_R^-}-G_\lambda(z)e^{\lambda z}\left(\frac{1}{z}+\frac{z}{R^2}\right)\,dz\right| \\ &\le I_1(R)+I_2(R).\end{aligned}$$

For $I_2(R)$, the function $G_\lambda(z)$ is entire and so we may replace γ_R^- by the semi-circle from iR to $-iR$ and deduce that

$$I_2(R)\le\frac{1}{R}$$

which tends to 0 as R tends to ∞.

For $I_1(R)$, we splits the contour γ_R^- into two parts. Suppose the purple vertical line in the diagram is given by $\operatorname{Re} z=-\delta_1, \delta_1>0$. Then if $|G(z)|\le M(R)$ for $z\in\gamma_R^-$, then

$$I_1(R)\le\frac{1}{2}M(R)e^{-\lambda\delta_1}\left(\frac{1}{\delta(R)}+\frac{1}{R}\right)R+\frac{1}{\pi}M(R)\left(\frac{1}{\delta(R)}+\frac{1}{R}\right)\sin^{-1}\frac{\delta_1}{R},$$

where the first estimate is from the contour to the left of $\operatorname{Re} z=-\delta_1$ and the second estimate is from the contour to the right of $\operatorname{Re} z=-\delta_1$. For the first estimate, we find that it tends to 0 as λ tends to ∞. For the second estimate, we see that it tends to 0 as δ_1 tends to 0. This completes the proof the theorem. □

6 Dirichlet Series

6.1 Absolute convergence of a Dirichlet series

DEFINITION 6.1 A Dirichlet series is a series of the form

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s}, s = \sigma + it,$$

where $f(n)$ is an arithmetical function.

Note that if $\sigma \geq a$ then $|n^s| \geq n^a$. Therefore,

$$\left| \frac{f(n)}{n^s} \right| \leq \frac{|f(n)|}{n^a}.$$

Therefore, if a Dirichlet series converges absolutely for $s = a + ib$, then by the comparison test, it also converges absolutely for all s with $\sigma \geq a$. This observation implies the following theorem.

THEOREM 6.1 Suppose the series

$$\sum_{n=1}^{\infty} \left| \frac{f(n)}{n^s} \right|$$

does not converge for all s or diverge for all s. Then there exists a real number σ_a called the abscissa of absolute convergence, such that the series

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

converges absolutely if $\sigma > \sigma_a$ but does not converge absolutely if $\sigma < \sigma_a$.

Proof
Let D be the set of all reals σ such that

$$\sum_{n=1}^{\infty} \left| \frac{f(n)}{n^s} \right|$$

diverges. Then D is not empty because the series does not converge for all s. The set D is bounded above since the series does not diverge for all s. Therefore, D has a least upper bound which we call σ_a. If $\sigma < \sigma_a$ then we claim that

$$\sum_{n=1}^{\infty} \frac{|f(n)|}{n^{\sigma}}$$

diverges. For otherwise,

$$\sum_{n=1}^{\infty} \frac{|f(n)|}{n^{\sigma}}$$

converges implies

$$\sum_{n=1}^{\infty} \left| \frac{f(n)}{n^{s}} \right|$$

converges for all Re $s > \sigma$. Hence, σ is an upper bound for D and since $\sigma < \sigma_a$, σ_a is not a least upper bound for D.

If $\sigma > \sigma_a$, then $\sigma \notin D$ since σ_a is an upper bound for D and the Dirichlet series converges absolutely. This proves the theorem. □

6.2 The Uniqueness Theorem

THEOREM 6.2 Let

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^{s}} \quad \text{and} \quad G(s) = \sum_{n=1}^{\infty} \frac{g(n)}{n^{s}}$$

be absolutely convergent for $\sigma > \sigma_0$. If $F(s) = G(s)$ for each s in an infinite sequence $\{s_k\}$ such that $\sigma_k \to \infty$ as $k \to \infty$, then $f(n) = g(n)$ for every n.

Proof

Let $h(n) = f(n) - g(n)$ and let $H(s) = F(s) - G(s)$. Then $H(s_k) = 0$ for each k. To prove that $h(n) = 0$ for all n we assume that $h(n) \neq 0$ for some n and obtain a contradiction.

Let N be the smallest integer n for which

$$h(n) \neq 0. \tag{6.1}$$

Then

$$H(s) = \sum_{n=N}^{\infty} \frac{h(n)}{n^{s}} = \frac{h(N)}{N^{s}} + \sum_{n=N+1}^{\infty} \frac{h(n)}{n^{s}}.$$

Hence,

$$h(N) = N^{s} H(s) - N^{s} \sum_{n=N+1}^{\infty} \frac{h(n)}{n^{s}}.$$

Putting $s = s_k$, we have $H(s_k) = 0$, and hence

$$h(N) = -N^{s_k} \sum_{n=N+1}^{\infty} \frac{h(n)}{n^{s_k}}.$$

Choose k so that $\sigma_k > c$ where $c > \sigma_a$. Now, note that

$$\frac{N^{\sigma_k}}{n^{\sigma_k}} = \frac{N^{\sigma_k - c}}{n^{\sigma_k - c}} \frac{N^c}{n^c} \leq \left(\frac{N}{N+1}\right)^{\sigma_k - c} \frac{N^c}{n^c}.$$

Then

$$|h(N)| \leq \left(\frac{N}{N+1}\right)^{(\sigma_k - c)} N^c \sum_{n=N+1}^{\infty} \frac{|h(n)|}{n^c} = \left(\frac{N}{N+1}\right)^{\sigma_k - c} A$$

where A is independent of k. Letting $k \to \infty$, we find that

$$\left(\frac{N}{N+1}\right)^{\sigma_k} \to 0.$$

Hence, $h(N) = 0$, a contradiction to (6.1). Consequently, $h(n) = 0$ for all positive integers n.

□

The above result is very useful. For example let $f(n)$ be a completely multiplicative function. Suppose

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}$$

and

$$G(s) = \sum_{n=1}^{\infty} \frac{f^{-1}(n)}{n^s}$$

are absolutely convergent for $\sigma \geq \sigma_0$. Then we know that

$$G(s) = 1/F(s) = \prod_p \left(1 - \frac{f(p)}{p^s}\right) = \sum_{n=1}^{\infty} \frac{\mu(n) f(n)}{n^s}.$$

By Theorem 6.2, this shows that

$$f^{-1}(n) = \mu(n) f(n).$$

6.3 Multiplication of Dirichlet series

The next theorem relates products of Dirichlet series with the Dirichlet convolution of their coefficients.

THEOREM 6.3 Given two functions $F(s)$ and $G(s)$ represented by Dirichlet series

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \quad \text{for } \sigma > a,$$

and

$$G(s) = \sum_{n=1}^{\infty} \frac{g(n)}{n^s} \quad \text{for } \sigma > b.$$

Then in the half plane where both series converge absolutely, we have

$$F(s)G(s) = \sum_{n=1}^{\infty} \frac{f * g(n)}{n^s}.$$

If

$$F(s)G(s) = \sum_{n=1}^{\infty} \frac{\alpha(n)}{n^s}$$

for all s in a sequence $\{s_k\}$ such that $\sigma_k \to \infty$ as $k \to \infty$ then $\alpha = f * g$.

Proof

For any s for which both series converge absolutely, we have

$$F(s)G(s) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{f(n)g(m)}{(nm)^s}.$$

Because of absolutely convergence, we can multiply these series together and arrange the terms in any way we please without altering the sum. Collect together those terms for which mn is constant, say $mn = k$. The possible values of k are $1, 2, \cdots$, hence,

$$F(s)G(s) = \sum_{k=1}^{\infty} \frac{\left(\sum_{mn=k} f(n)g(m) \right)}{k^s} = \sum_{k=1}^{\infty} \frac{h(k)}{k^s}$$

where

$$h(k) = \sum_{mn=k} f(n)g(m) = f * g(k).$$

This proves the first assertion. The second assertion follows from Theorem 6.2. □

6.4 Conditional convergence of Dirichlet series

THEOREM 6.4 For every Dirichlet series, there exists $\sigma_c \in [-\infty, \infty]$ such that the series converges (conditionally) for any s with $\sigma > \sigma_c$ and diverges for any s with $\sigma < \sigma_c$. Moreover,

$$\sigma_c \leq \sigma_a \leq \sigma_c + 1.$$

Proof

We will show that if

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

converges for $s = s_1$, then it also converges for every s with $\sigma > \sigma_1$.

Since

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

converges at $s = s_1$, we conclude that there exists a positive integer N_0 such that

$$\left| \sum_{y<n\leq x} \frac{f(n)}{n^{s_1}} \right| \leq 1$$

for all integers $x > y > N_0$. Now, let s with $\sigma > \sigma_1$ be given and let $x > y \geq N_0$. Let $\epsilon > 0$ be given. Then

$$\begin{aligned} \sum_{y<n\leq x} \frac{f(n)}{n^s} &= \sum_{y<n\leq x} \frac{f(n)}{n^{s_1}} n^{s_1-s} \\ &= \sum_{y<n\leq x} \frac{f(n)}{n^{s_1}} x^{s_1-s} - \frac{f(y)}{y^{s_1}} y^{s_1-s} - \int_y^x \sum_{y<n\leq t} \frac{f(n)}{n^{s_1}} t^{s_1-s-1}(s_1-s)dt. \end{aligned}$$

Therefore,

$$\begin{aligned} \left| \sum_{y<n\leq x} \frac{f(n)}{n^s} \right| &\leq 2y^{\sigma_1-\sigma} + \int_y^x |s_1 - s| t^{\sigma_1-\sigma-1} dt \\ &\leq 2y^{\sigma_1-\sigma} \left(1 + \frac{|s_1 - s|}{\sigma - \sigma_1} \right) \qquad (6.2) \\ &< \epsilon \end{aligned}$$

provided that

$$y \geq \left\{ \frac{2\left(1 + \dfrac{|s_1 - s|}{\sigma - \sigma_1}\right)}{\epsilon} \right\}^{1/(\sigma-\sigma_1)}.$$

We have therefore shown that for any $\epsilon > 0$ and a fixed s with $\sigma > \sigma_1$,

$$\left| \sum_{y<n\leq x} \frac{f(n)}{n^s} \right| < \epsilon$$

whenever

$$x \geq y \geq \max\left(N_0, \left\{ 2\frac{\left(1 + \frac{|s_1-s|}{\sigma-\sigma_1}\right)}{\epsilon} \right\}^{1/(\sigma-\sigma_1)} \right).$$

This shows the convergence of the Dirichlet series at s.

Now, let

$$\sigma_c := sup\left\{ \operatorname{Re} s | \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \quad \text{diverges} \right\}. \tag{6.3}$$

If $\sigma > \sigma_a$ then by previous argument, we conclude that $F(s)$ is convergent whenever $\sigma = \sigma_a + \delta$, $\delta > 0$. Therefore, we conclude that $\sigma_a \geq \sigma_c$.

It remains to show that $\sigma_a \leq \sigma_c + 1$. We first show that if

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

is convergent at $s = s_1$ then it is absolutely convergent at any s with $\sigma > \sigma_1 + 1$. The series

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^{s_1}}$$

is convergent implies that $f(n)n^{-s_1} \to 0$ as $n \to \infty$, or

$$\left| \frac{f(n)}{n^{s_1}} \right| \leq C$$

for all $n \in \mathbf{N}$ and some positive constant C. Given s with $\sigma > \sigma_1 + 1$,

$$\left| \frac{f(n)}{n^s} \right| = \left| \frac{f(n)}{n^{s_1}} \right| \frac{1}{n^{\sigma-\sigma_1}} \leq \frac{C}{n^{\sigma-\sigma_1}},$$

with $\sigma - \sigma_1 > 1$. Therefore, for any positive integer m,

$$\sum_{n=1}^{m} \left| \frac{f(n)}{n^s} \right| \leq \sum_{n=1}^{m} \frac{C}{n^{\sigma-\sigma_1}}.$$

Since $\sigma - \sigma_1 > 1$, the series

$$\sum_{n=1}^{\infty} \frac{1}{n^{\sigma-\sigma_1}}$$

converges. By comparison test, we conclude that

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

is absolutely convergent.

Now, we have shown that if

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

is convergent at $s_1 = \sigma_1 + it$, then $\sum_{n=1}^{\infty} \frac{f(n)}{n^s}$ is aboslutely convergent whenever $\sigma > \sigma_1 + 1$.

Therefore,

$$\sigma_1 + 1 \geq \sigma_a.$$

Now,

$$\sigma_1 = \sigma_c + \delta$$

for any positive δ and hence,

$$\sigma_c + 1 \geq \sigma_a.$$

□

6.5 Landau's Theorem for Dirichlet series

THEOREM 6.5 A Dirichlet series

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}$$

is analytic in $\sigma > \sigma_c$, where σ_c is given by (6.3).

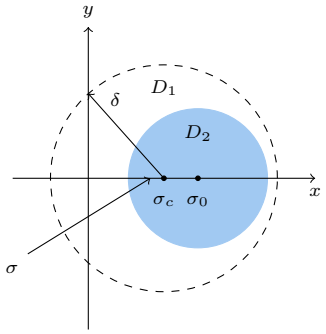
We now come to the main theorem of this chapter.

THEOREM 6.6 Let

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}$$

be a Dirichlet series with $f(n) \geq 0$ for all $n \in \mathbf{N}$ and $\sigma_c < \infty$. Then the function $F(s)$ has a singularity at $s = \sigma_c$.

Proof
Suppose $F(s)$ is analytic at σ_c. Then there exists $\delta > 0$ such that $F(s)$ is analytic in $D_1 := \{s : |s - \sigma_c| < \delta\}$. Fix a point on the real axis, say $\sigma_0 > \sigma_c$ contained in this disc and an $\epsilon > 0$ such that the whole disc $D_2 := \{s : |s - \sigma_0| < \epsilon\}$ is inside D_1 and $\sigma_c \in D_2$ (see the following diagram).



Since $F(s)$ is analytic in D_1, and hence analytic in D_2, we conclude that for $s \in D_2$,

$$F(s) = \sum_{n=0}^{\infty} \frac{F^{(n)}(\sigma_0)}{n!}(s - \sigma_0)^n.$$

Next, the Dirichlet series is convergent for $\sigma > \sigma_0$. So for σ close to σ_0 (for example, s is in an open ball centered at σ_0 that lies to the right of $x = \sigma_c$), $F(s)$ is given by

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

Therefore, we can differentiate the above term by term and substitute $s = \sigma_0$ to deduce that

$$F^{(\nu)}(\sigma_0) = \sum_{n=1}^{\infty} (-1)^{\nu} \frac{f(n) \ln^{\nu} n}{n^{\sigma_0}}.$$

Substituting this into the Taylor series expansion, we find that

$$F(s) = \sum_{\nu=0}^{\infty} \frac{(\sigma_0 - s)^{\nu}}{\nu!} \sum_{n=1}^{\infty} \left(\frac{f(n) \ln^{\nu} n}{n^{\sigma_0}} \right).$$

Now taking s real, say $\sigma_0 - \epsilon < s = \sigma < \sigma_c$, we have

$$\begin{aligned} F(\sigma) &= \sum_{\nu=0}^{\infty} \frac{(\sigma_0 - \sigma)^\nu}{\nu!} \sum_{n=1}^{\infty} \left(\frac{f(n) \ln^\nu n}{n^{\sigma_0}} \right) \\ &= \sum_{n=1}^{\infty} \frac{f(n)}{n^{\sigma_0}} \sum_{\nu=0}^{\infty} \frac{(\sigma_0 - \sigma)^\nu \ln^\nu n}{\nu!}, \end{aligned}$$

where the interchanging of the summations is valid since $f(n) \geq 0$ and $\sigma_0 - \sigma \geq 0$. Hence,

$$\begin{aligned} F(\sigma) &= \sum_{n=1}^{\infty} \frac{f(n)}{n^{\sigma_0}} \exp((\sigma_0 - \sigma) \ln n) \\ &= \sum_{n=1}^{\infty} \frac{f(n)}{n^{\sigma_0}} n^{\sigma_0 - \sigma} = \sum_{n=1}^{\infty} \frac{f(n)}{n^{\sigma}}. \end{aligned}$$

The last equality shows that Dirichlet series is convergent for some $\sigma < \sigma_c$ and this contradicts our assumption that σ_c is the abscissa of conditional convergence. □

7 Primes in Arithmetic Progression

7.1 Introduction

In Chapter 4, we proved that there are infinitely many primes by showing that (see Theorem 4.9 (c))

$$\sum_{p \le x} \frac{1}{p} = \ln \ln x + O(1). \tag{7.1}$$

The Dirichlet Theorem of primes in arithmetic progression states that for $(k, l) = 1$, there are infinitely many primes of the form $kn + l$. If we can prove a result similar to (7.1), with sum over primes p replaced by sum over primes of the form $kn + l$, then we would have Dirichlet's Theorem as a consequence. This strategy motivates the following theorem.

THEOREM 7.1 Let $k > 1$ and l be positive integers such that $(k, l) = 1$. Then

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

Theorem 7.1 immediately implies the Dirichlet Theorem of primes in arithmetic progression.

THEOREM 7.2 (Dirichlet's Theorem of primes in arithmetic progression) If k and l are positive integers such that $(k, l) = 1$, then there are infinitely many primes of the form $kn + l$.

7.2 Dirichlet's characters

DEFINITION 7.1 A Dirichlet character $\pmod k$ is an arithmetical function

$$\chi : \mathbf{N} \to \mathbf{C}$$

satisfying

(i) $\chi(mn) = \chi(m)\chi(n)$ for all $m, n \in \mathbf{N}$.
(ii) $|\chi(n)| = \begin{cases} 1 & \text{if } (n,k) = 1 \\ 0 & \text{otherwise.} \end{cases}$
(iii) $\chi(n + km) = \chi(n)$ for all $n, m \in \mathbf{N}$,
(iv) $\chi^{\varphi(k)}(n) = 1, (n,k) = 1$.

Remark 7.1

(a) The values of χ are 0 or $\varphi(k)$-th roots of unity. This follows from (iv).
(b) There are only finitely many characters $\pmod k$. This follows from the fact that χ is defined on $\varphi(k)$ values j with $1 \le j \le k$ and $(j,k) = 1$. Hence, from (iv), we see that for each j, there are $\varphi(k)$ values we can assign to $\chi(j)$. This shows that there can be at most $\varphi(k)^{\varphi(k)}$ characters.
(c) If χ_1 and χ_2 are characters $\pmod k$, then so is $\chi_1\chi_2$.
(d) A character $\chi \pmod k$ can be obtained from a homomorphism

$$\widetilde{\chi} : (\mathbf{Z}/k\mathbf{Z})^* \to \{z \in \mathbf{C} || z| = 1\}$$

where $(\mathbf{Z}/k\mathbf{Z})^*$ is the multiplicative group of residue classes

$$(\{[n]_k | (n,k) = 1\}, \cdot),$$

with multiplication $\cdot$ as group operation. Given a character $\widetilde{\chi}$, one defines

$$\chi(n) = \begin{cases} \widetilde{\chi}([n]_k), & (n,k) = 1 \\ 0, & \text{otherwise.} \end{cases}$$

Conversely, given χ, one obtains a homomorphism $\widetilde{\chi}$ given by

$$\widetilde{\chi}([n]_k) = \chi(n).$$

This shows that there is a one to one correspondence between Dirichlet's characters $\pmod k$ and homomorphisms from

$$(\mathbf{Z}/k\mathbf{Z})^* \quad \text{to} \quad \{z \in \mathbf{C} || z| = 1\}.$$

THEOREM 7.3 There are exactly $\varphi(k)$ characters $\pmod k$.

Proof

From Remark 7.1 (d) above, it suffices to show that there are exactly $\varphi(k)$ homomorphisms from

$$(\mathbf{Z}/k\mathbf{Z})^* \text{ to } \{z \in \mathbf{C} || z| = 1\}.$$

From the structure theorem of abelian group, we know that $(\mathbf{Z}/k\mathbf{Z})^*$ can be written as a direct sums of cyclic groups with prime power order, say,

$$(\mathbf{Z}/k\mathbf{Z})^* = C_{h_1} \times \cdots \times C_{h_r},$$

where h_i are prime powers and C_m denotes a cyclic group of order m.

Let $[a_i]_k$ be a generator for C_{h_i}, $1 \le i \le r$. Given $w_1, \cdots, w_r$ such that

$$w_i^{h_i} = 1,$$

set

$$\widetilde{\chi}([a_i]_k) = w_i, 1 \le i \le r.$$

If

$$[n]_k = \prod_i [a_i]_k^{\alpha_i},$$

then define

$$\widetilde{\chi}([n]_k) = \prod_i \widetilde{\chi}([a_i]_k)^{\alpha_i}.$$

Note that $\widetilde{\chi}$ is a homomorphism from

$$(\mathbf{Z}/k\mathbf{Z})^* \text{ to } \{z \in \mathbf{C} || z| = 1\}.$$

Therefore, we have at least

$$h_1 \cdots h_r = \varphi(k)$$

such homomorphisms.

Next, let $[a]_k \in (\mathbf{Z}/k\mathbf{Z})^*$. Then

$$[a]_k = [a_1]_k^{\alpha_1} \cdots [a_r]_k^{\alpha_r}$$

where

$$0 \le \alpha_i \le h_r - 1.$$

Now if $\widetilde{\chi}$ is a homomorphism from

$$(\mathbf{Z}/k\mathbf{Z})^* \text{ to } \{z \in \mathbf{C} || z| = 1\},$$

then

$$\widetilde{\chi}([a]_k) = \prod_i \widetilde{\chi}([a_i]_k)^{\alpha_i}.$$

The value $\widetilde{\chi}([a]_k)$ is dependent on the values $\widetilde{\chi}([a_i]_k), 1 \le i \le r$. The number of possible values for $\widetilde{\chi}([a_i]_k)$ is h_i, $1 \le i \le r$. Therefore, there can be at most $h_1 h_2 \cdots h_r = \varphi(k)$ characters. In conclusion, we deduce that there are exactly $\varphi(k)$ characters $\pmod k$. □

The character χ_0 will always denote the principal character $\pmod k$, that is,

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

The character $\overline{\chi}$ will denote the inverse of χ, or, $\chi \cdot \overline{\chi} = \chi_0$.

The set of Dirichlet characters forms an abelian group with binary operation $\chi_1\chi_2$ to be defined as $\chi_1\chi_2(n) = \chi_1(n)\chi_2(n)$. This group, which we will denote as $\hat{G}$ has identity χ_0.

Remark 7.2 The above Theorem is a special case from a more general Theorem in the theory of characters. In general a linear representation is a homomorphism $\rho : G \to GL_n(\mathbf{C})$. A character is defined to be $\chi(g) = \text{Trace}(\rho(g))$. In general, χ is not a homomorphism. However, when G is abelian, all irreducible representations are 1-dimensional over $\mathbf{C}$. In this case, $\chi(g) = \rho(g)$, and so, χ is a homomorphism. Furthermore, from character theory, we know that there are exactly C irreducible characters for a finite group with C conjugacy classes. When G is abelian, each element represents a single conjugacy class and so, there are exactly $|G|$ conjugacy classes, hence exactly $|G|$ characters. This explains why there are exactly $\varphi(k)$ characters $\pmod k$.

7.3 The orthogonal relations

In this section, we will often identify (see Remark 7.1 (d)) Dirichlet's characters χ with homomorphism $\widetilde{\chi}$ from

$$(\mathbf{Z}/k\mathbf{Z})^* \quad \text{to} \quad \{z \in \mathbf{C} || z| = 1\}.$$

THEOREM 7.4 (a) Let χ_1, χ_2 be two Dirichlet's characters modulo k. Then

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

(b) Let a_1, a_2 be integers with $(a_i, k) = 1$. Then

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

Proof
We will prove the following:

$$\sum_{a=1}^{k} \chi(a) = \begin{cases} \varphi(k) & \text{if } \chi = \chi_0, \\ 0 & \text{otherwise.} \end{cases} \tag{7.2}$$

We first observe that since $\chi(l) = 0$ whenever $(l, k) \neq 1$, we must have

$$\sum_{a=1}^{k} \chi(a) = \sum_{\substack{a=1 \\ (a,k)=1}}^{k} \chi(a).$$

If $\chi = \chi_0$ then for $(a, k) = 1$, $\chi(a) = 1$ and

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

If $\chi \neq \chi_0$, then there exists an a_0 relatively prime to k such that $\chi(a_0) \neq 1$. Now,

$$\begin{aligned} \chi(a_0) \sum_{a=1}^{k} \chi(a) &= \widetilde{\chi}([a_0]_k) \sum_{[a]_k \in (\mathbf{Z}/k\mathbf{Z})^*} \widetilde{\chi}([a]_k) \\ &= \sum_{[a]_k \in (\mathbf{Z}/k\mathbf{Z})^*} \widetilde{\chi}([a_0]_k [a]_k). \end{aligned}$$

Now, the multiplication of elements in $(\mathbf{Z}/k\mathbf{Z})^*$ by $[a_0]_k$ permutes the elements in $(\mathbf{Z}/k\mathbf{Z})^*$. Hence,

$$\sum_{[a]_k \in (\mathbf{Z}/k\mathbf{Z})^*} \widetilde{\chi}([a_0]_k [a]_k) = \sum_{[a]_k \in (\mathbf{Z}/k\mathbf{Z})^*} \widetilde{\chi}([a]_k) = \sum_{a=1}^{k} \chi(a).$$

Therefore, we conclude that

$$\sum_{a=1}^{k} \chi(a) = 0.$$

We now let $\chi = \chi_1 \overline{\chi_2}$ in (7.2) to complete the proof of (a). □

Proof of (b).

We will first show that

$$\sum_{\chi \pmod k} \chi(a) = \begin{cases} \varphi(k) & \text{if } a \equiv 1 \pmod k, \\ 0 & \text{otherwise.} \end{cases} \tag{7.3}$$

If $a \equiv 1 (\mathrm{mod}\ k)$ then $\chi(a) = 1$ for all characters χ. Since there are exactly $\varphi(k)$ such characters, we conclude that

$$\sum_{\chi} \chi(a) = \varphi(k).$$

Next, suppose $a \not\equiv 1 (\mathrm{mod}\ k)$. Then since characters are constructed by assigning roots of unity to the image of the characters on the generators of the cyclic components of $(\mathbf{Z}/k\mathbf{Z})^*$, there exists a character χ^* so that $\chi^*(a) \neq 1$. Therefore,

$$\chi^*(a) \sum_{\chi} \chi(a) = \sum_{\chi} \chi^*\chi(a) = \sum_{\chi} \chi(a),$$

where we have used the fact that multiplying the elements in the set of characters by χ^* permutes the elements in the set. This implies that

$$\sum_{\chi} \chi(a) = 0.$$

Now, in order to prove (b), we simply view χ as $\widetilde{\chi}$ and let $[a]_k = [a_1]_k \overline{[a_2]}_k$ where $\overline{[a]}_k$ denotes the inverse of $[a]_k$ in the group $(\mathbf{Z}/k\mathbf{Z})^*$, and observe that

$$\chi(a_1)\overline{\chi(a_2)} = \widetilde{\chi}([a_1]_k)\widetilde{\chi}(\overline{[a_2]}_k).$$

□

7.4 The Dirichlet L-series

DEFINITION 7.2 The Dirichlet L-series is defined as

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

THEOREM 7.5 (a) If $\chi = \chi_0$ then $L(s, \chi)$ can be analytically continued to the half-plane $\sigma > 0$, with the exception of the point $s = 1$ where it has a simple pole with residue $\varphi(k)/k$.

(b) If χ is not the principal character $(\mathrm{mod}\ k)$, then $L(s, \chi)$ can be analytically continued to $\sigma > 0$.

Proof of (a)

For $\sigma > 1$, we have by Theorem 5.4,

$$L(s, \chi) = \prod_{p} \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}.$$

Therefore,

$$L(s,\chi_0)=\prod_{p\nmid k}\left(1-\frac{1}{p^s}\right)^{-1}=\prod_{p}\left(1-\frac{1}{p^s}\right)^{-1}\prod_{p|k}\left(1-\frac{1}{p^s}\right).$$

The function $\zeta(s)$ has an analytic continuation with residue 1 at $s=1$. Therefore, the residue of $L(s,\chi_0)$ at $s=1$ is $\varphi(k)/k$ since

$$\lim_{s\to 1}\prod_{p|k}\left(1-\frac{1}{p^s}\right)=\frac{\varphi(k)}{k}.$$

□

Proof of (b)
If $\chi\neq\chi_0$, then

$$\sum_{n=1}^{k}\chi(n)=0.$$

Therefore,

$$\left|\sum_{n\le x}\chi(n)\right|\le k,$$

for $x\ge 1$. Hence, for any $\epsilon>0$,

$$\left|\sum_{y\le n\le x}\frac{\chi(n)}{n^s}\right|\le\frac{1}{|y^s|}\left|\sum_{y\le n\le x}\chi(n)\right|<\frac{k}{|y|^\sigma}<\epsilon,$$

whenever

$$|y|>\left(\frac{k}{\epsilon}\right)^{1/\sigma}.$$

This implies that L-series converges for $\sigma>0$. □

7.5 Proof of Dirichlet's Theorem

Step 1.

It suffices to show that if $x\ge 3$ and

$$\sigma=1+\frac{1}{\ln x},$$

then

$$\sum_{\substack{p\\ p\equiv l \pmod k}}\frac{1}{p^\sigma}=\frac{1}{\varphi(k)}\ln\left(\frac{1}{\sigma-1}\right)+O(1).$$

Let

$$\Sigma_1 = \sum_{p \equiv l \pmod{k}} \frac{1}{p^\sigma} \quad \text{and} \quad \Sigma_2 = \sum_{\substack{p \le x \\ p \equiv l \pmod{k}}} \frac{1}{p},$$

where

$$\sigma = 1 + \frac{1}{\ln x}.$$

Then

$$|\Sigma_1 - \Sigma_2| \le \underbrace{\sum_{p \le x} \left(\frac{1}{p} - \frac{1}{p^\sigma} \right)}_{\Sigma_3} + \underbrace{\sum_{p > x} \frac{1}{p^\sigma}}_{\Sigma_4}.$$

Now,

$$\begin{aligned}
\Sigma_3 &= \sum_{p \le x} \frac{1 - e^{-(\sigma-1)\ln p}}{p} = \sum_{p \le x} \sum_{m=1}^{\infty} \frac{-(-(\sigma-1)\ln p)^m}{m! p} \\
&= \sum_{p \le x} \frac{(\sigma-1)\ln p}{p} \sum_{m=1}^{\infty} \frac{-(-(\sigma-1)\ln p)^{m-1}}{m!} \\
&\le \frac{1}{\ln x} \sum_{p \le x} \frac{\ln p}{p} \sum_{m=1}^{\infty} \frac{((\sigma-1)\ln p)^{m-1}}{(m-1)!} \\
&\le \frac{1}{\ln x} \sum_{p \le x} \frac{\ln p}{p} e^{(\sigma-1)\ln p} \\
&\le \frac{1}{\ln x} \sum_{p \le x} \frac{\ln p}{p} x^{1/\ln x} = O(1).
\end{aligned}$$

Next,

$$\begin{aligned}
\Sigma_4 &= \lim_{y \to \infty} \sum_{x \le p \le y} \frac{1}{p^\sigma} \\
&= \lim_{y \to \infty} \left(\frac{1}{y^\sigma} \sum_{p \le y} 1 - \frac{1}{x^\sigma} \sum_{p \le x} 1 - \int_x^y \sum_{p \le t} 1 \left(-\frac{\sigma}{t^{\sigma+1}} \right) dt \right) \\
&= O(1) + \int_x^\infty O\left(\frac{t}{\ln t} \right) \frac{dt}{t^{\sigma+1}} \\
&= O(1) + O\left(\int_x^\infty \frac{dt}{t^\sigma \ln t} \right) \\
&= O(1) + O\left(\frac{1}{\ln x} \int_x^\infty \frac{dt}{t^\sigma} \right) = O(1).
\end{aligned}$$

Therefore, if

$$\sigma = 1 + \frac{1}{\ln x},$$

then

$$\sum_{\substack{p \\ p \equiv l \pmod k}} \frac{1}{p^\sigma} = \frac{1}{\varphi(k)} \ln \frac{1}{\sigma - 1} + O(1)$$

and Dirichlet's Theorem holds.

Step 2.

We observe that for $\sigma > 1$,

$$\begin{aligned} \sum_{\substack{p \\ p \equiv l \pmod k}} \frac{1}{p^\sigma} &= \sum_p \frac{1}{p^\sigma} \left(\frac{1}{\varphi(k)} \sum_{\chi \pmod k} \overline{\chi(l)} \chi(p) \right) \\ &= \frac{1}{\varphi(k)} \sum_{\chi \pmod k} \overline{\chi(l)} S(\sigma, \chi), \end{aligned}$$

where

$$S(\sigma, \chi) = \sum_p \frac{\chi(p)}{p^\sigma}.$$

Now,

$$\begin{aligned} S(\sigma, \chi_0) &= \sum_p \frac{\chi_0(p)}{p^\sigma} = \sum_p \frac{1}{p^\sigma} - \sum_{p \mid k} \frac{1}{p^\sigma} \\ &= \sum_p \frac{1}{p^\sigma} + O(1). \end{aligned} \tag{7.4}$$

But

$$-\sum_p \ln \left(1 - \frac{1}{p^\sigma}\right) - \sum_p \frac{1}{p^\sigma} = \sum_p \sum_{m \geq 1} \frac{1}{m p^{m\sigma}} - \sum_p \frac{1}{p^\sigma} = O(1), \tag{7.5}$$

since

$$\sum_p \sum_{m \geq 2} \frac{1}{m p^{m\sigma}} \leq \frac{1}{2} \sum_p \sum_{m \geq 2} \frac{1}{p^{m\sigma}} = \frac{1}{2} \sum_p \frac{1}{p^\sigma (p^\sigma - 1)} = O(1).$$

Hence, by (7.4) and (7.5),

$$\begin{aligned} S(\sigma, \chi) &= \sum_p \frac{1}{p^\sigma} + O(1) = -\sum_p \ln \left(1 - \frac{1}{p^\sigma}\right) + O(1) \\ &= \ln \prod_p \left(1 - \frac{1}{p^\sigma}\right)^{-1} + O(1) \\ &= \ln \zeta(\sigma) + O(1). \end{aligned}$$

Now,

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

where $g(\sigma)$ is a function analytic at 1. Hence,

$$\ln \zeta(\sigma) = \ln\left(\frac{1}{\sigma - 1}\right) + \ln(1 + (\sigma - 1)g(\sigma)).$$

Since $\ln(1 + (\sigma - 1)g(\sigma))) \to 0$ as $\sigma \to 1$, we conclude that

$$S(\sigma, \chi_0) = \ln\left(\frac{1}{\sigma - 1}\right) + O(1).$$

We conclude that the main term arises from the principal character χ_0. Hence, it remains to show that

$$S(\sigma, \chi) = O(1)$$

for $\sigma > 1$ and all non-principal characters $\chi \pmod k$.

Step 3.

Now, using computations similar to Step 2, we find that

$$\begin{aligned} S(\sigma, \chi) &= \sum_p \frac{\chi(p)}{p^\sigma} = \sum_p \sum_{m \geq 1} \frac{\chi(p)^m}{m p^{m\sigma}} + O(1) \\ &= -\sum_p \ln\left(1 - \frac{\chi(p)}{p^\sigma}\right)^{-1} + O(1) \\ &= \ln(L(\sigma, \chi)) + O(1). \end{aligned}$$

Now, for $\chi \neq \chi_0$ $L(s, \chi)$ is analytic in $\sigma > 0$. So, $L(\sigma, \chi)$ is continuous at $\sigma > 1$ and

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

If $L(1, \chi) \neq 0$ then we are done. It remains to show that $L(1, \chi) \neq 0$.

Step 4.

We first show that when $\chi \neq \chi_0$ is a complex character $\pmod k$, then

$$L(1, \chi) \neq 0.$$

Consider the expression

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

We find that for $\sigma > 1$,

$$\begin{aligned}\ln P(\sigma) &= \sum_{\chi \pmod k} \ln L(\sigma, \chi) \\ &= \sum_{\chi \pmod k} \sum_{p} \sum_{m \geq 1} \frac{\chi(p^m)}{mp^{m\sigma}} \\ &= \sum_{p} \sum_{m \geq 1} \frac{1}{mp^{m\sigma}} \sum_{\chi \pmod k} \chi(p^m)\overline{\chi(1)} \\ &= \sum_{p} \sum_{\substack{m \geq 1 \\ p^m \equiv 1 \pmod k}} \frac{1}{mp^{m\sigma}} \geq 0.\end{aligned}$$

Hence, for $\sigma > 1$,

$$P(\sigma) \geq 1. \tag{7.6}$$

Suppose that $L(1, \chi) = 0$ for some χ. Then $L(1, \bar{\chi}) = 0$. Hence, $P(s)$ has two zeros at $s = 1$. But $L(s, \chi_0)$ has a simple pole at $s = 1$, which means that $P(1) = 0$. This is a contradiction to (7.6).

Step 5.

In this final step, we show that for real character $\chi \neq \chi_0$, $L(1, \chi) \neq 0$. Consider the function $f = \chi * u$. Then f is multiplicative since it is the Dirichlet product of two multiplicative functions. Note that

$$\sum_{l=0}^{m} \chi(p^l) = \begin{cases} 1 & \text{if } p|k \\ \geq 1 & \text{if } p \nmid k, m \text{ even} \\ \geq 0 & \text{if } p \nmid k, m \text{ odd}. \end{cases}$$

Thus, $f(n) \geq 0$ for all n and $f(n) \geq 1$ when n is a square. Hence,

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

In particular, $F(\sigma)$ diverges at $\sigma = 1/2$ and so $\sigma_c \geq 1/2$. By Theorem 6.6, $F(s)$ must have a singularity at $s = \sigma_c \geq 1/2$.

On the other hand, for $\sigma > 1$,

$$F(s) = L(s, \chi)\zeta(s).$$

If $L(1, \chi) = 0$, then $F(s)$ would be analytic in $\sigma > 0$ and hence at $\sigma = \sigma_c$. This contradicts our previous observation that $F(s)$ has a singularity at σ_c and we must have $L(1, \chi) \neq 0$.

From Steps 3 and 4, we conclude that $L(1, \chi) \neq 0$ for all non-principal characters χ. This completes the proof of Dirichlet's Theorem.

Remark 7.3 The Dirichlet Theorem is a special case of the Chebotarev Density Theorem.

Remark 7.4 If p is a prime that satisfies the property that $p + 2$ is also a prime, then we call p a twin prime.The *twin primes conjecture* states that there are infinitely many twin primes. This statement remains an open problem.

Motivated by Merten's estimates and the proof of Dirichlet's Theorem of primes in arithmetical progression, it is natural to consider the sum

$$\sum_{p \leq x, p \in T} \frac{1}{p}$$

where T is the set of twin primes. If one can prove that the sum is divergent, then there would be infinitely many primes. Unfortunately, this sum turns out to be convergent (using sieve method). Consequently, this line of attack fails to provide a proof of the twin primes conjecture.

8 Introduction to Sieves

8.1 A weaker upper bound for $\pi(x)$

Sieve methods are important tools in analytic number theory. The earliest sieve is due to Eratothenes. The basic ideas of sieves are simple. Given a set of integers less than x, we want to sieve out set $\mathcal{A}$ which satisfies a certain property $\mathcal{P}$. For example, how many primes are there from 1 to x? We know that by elementary argument that an integer $n \leq x$ is a prime if $p \nmid n$ for all $p \leq \sqrt{x}$. In other words, if

$$P = \prod_{p \leq \sqrt{x}} p,$$

then to decide if $\sqrt{x} < n < x$ is a prime, it suffices to check if $(n, P) = 1$.

Let $x > 1$ be a real number and $\pi(x)$ be the number primes less than x. We find that

$$\begin{aligned}
\pi(x) - \pi(\sqrt{x}) &= \sum_{\substack{\sqrt{x} < n \leq x \\ (n,P)=1}} 1 \\
&= \sum_{\sqrt{x} < n \leq x} \sum_{\substack{d|P \\ d|n}} \mu(d) \\
&= \sum_{d|P} \mu(d) \sum_{\substack{\sqrt{x} < n \leq x \\ d|n}} 1 \\
&= \sum_{d|P} \mu(d) \left(\left[\frac{x}{d} \right] - \left[\frac{\sqrt{x}}{d} \right] \right) \\
&= x \sum_{d|P} \frac{\mu(d)}{d} + O\left(\sum_{d|P} 1 \right) - \sqrt{x} \sum_{d|P} \frac{\mu(d)}{d} \\
&= x \prod_{p \leq \sqrt{x}} \left(1 - \frac{1}{p} \right) + O\left(2^{\pi(\sqrt{x})} \right) - \sqrt{x} \prod_{p \leq \sqrt{x}} \left(1 - \frac{1}{p} \right).
\end{aligned}$$

But by Merten's estimate,

$$x \prod_{p \le \sqrt{x}} \left(1 - \frac{1}{p}\right) = O\left(\frac{x}{\ln x}\right)$$

and so, the error term $2^{\pi(\sqrt{x})}$ is larger than the main term, that is, (please check)

$$2^{\pi(\sqrt{x})} \gg \frac{x}{\ln x}.$$

We now modify the argument, introducing a new parameter $1 \le y \le x$.

Given an integer $n \le x$, a necessary condition for n to be composite is that $p|n$, $p \le y$. Then

$$\pi(x) - \pi(y) \le \alpha(x), \tag{8.1}$$

where

$$\alpha(x) = \sum_{\substack{n \le x \\ p|n \implies p > y}} 1 = \sum_{\substack{n \le x \\ (P,n)=1}} 1,$$

with

$$P = \prod_{p \le y} p.$$

To see why (8.1) is true, we note that the left hand side gives the number of primes between y and x. The function $\alpha(x)$ measures the number of integers $n \le x$ that are divisible by "large primes", that is, primes greater than y. This would include primes between y and x and hence the left hand side is less than the right hand side of (8.1). But $\alpha(x)$ also counts those integers $n \le x$ which are relatively prime to P.

Therefore,

$$\begin{aligned}\pi(x) - \pi(y) &= O\left(x \prod_{p \le y} \left(1 - \frac{1}{p}\right)\right) + O\left(2^{\pi(y)}\right) \\ &= O\left(\frac{x}{\ln y}\right) + O(2^y).\end{aligned}$$

Setting $y = \ln x$, we find that

$$x^{\ln 2} \ll \frac{x}{\ln \ln x},$$

and hence

$$\pi(x) - \pi(\ln x) + 1 \ll \frac{x}{\ln \ln x},$$

or, as a corollary,

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

This bound is of course weaker than Tchebychev's estimate but nonetheless, it is a non-trivial bound that is obtained from replacing $\sqrt{x}$ by y.

We will now study the situation closely. Let $\mathcal{A}$ be a subset of the set of integers less than x and $\mathcal{P}$ be the set of primes. Let

$$\mathcal{E} = \{n \in \mathcal{A} | n \not\equiv 0 \pmod{p} \quad \text{for all } p \in \mathcal{P}\}.$$

The set $\mathcal{E}$ is an example of a sifted set. Note that in our example above, our set $\mathcal{A}$ is the set of integers less than x and $\mathcal{P}$ is the set of primes less than y and $\alpha(x) = |\mathcal{E}|$.

8.2 The Large sieve and its applications

We now generalize the situation in the previous section. Let $\mathcal{A}$ be a subset of $\mathbf{N}$. Let $\mathcal{P}$ be the set of primes less than Q. Let Ω_p be a set of $\gamma(p)$ distinguished residue classes modulo p. Let

$$\mathcal{E} = \{n \in \mathcal{A} | n \pmod{p} \notin \Omega_p \quad \text{for all } p \in \mathcal{P}\}.$$

In the case when we are bounding $\pi(x)$, the distinguished residue class is 0 (mod p).

Let

$$S(\mathcal{A}, \mathcal{P}, Q) = |\mathcal{E}|.$$

Our aim is to bound $S(\mathcal{A}, \mathcal{P}, Q)$.

THEOREM 8.1 (The Large Sieve) Let N and Q be positive integers. Let $\mathcal{A}$ be the set of integers between 1 and N. Let $\mathcal{Q}$ be the set of $q \leq Q$ whose prime factors are in $\mathcal{P}$. Then

$$S(\mathcal{A}, \mathcal{P}, Q) \leq C\frac{N + Q^2}{L}$$

where C is some positive constant (which can be taken as 2π as seen in the proof) and

$$L = \sum_{q \in \mathcal{Q}} \mu^2(q) \prod_{p|q} \frac{\gamma(p)}{p - \gamma(p)}.$$

We will prove Theorem 8.1 in the next few sections. We begin with some applications of the Large Sieve.

EXAMPLE 8.1 Let $\mathcal{A} = \{n | n \leq x\}$ and $\mathcal{P} = \{p \text{ prime} | p \leq \sqrt{x}\}$. In other words, $Q = \sqrt{x}$. Let $\Omega_p = \{0\}$. This implies that $\gamma(p) = 1$. Let

$$\mathcal{E} = \{n \in \mathcal{A} | n \not\equiv 0 \pmod{p} \text{ for all } p \in \mathcal{P}.\}.$$

If n is a prime between $\sqrt{x}$ and x then $n \in \mathcal{E}$. Hence

$$\pi(x) - \pi(\sqrt{x}) \leq |\mathcal{E}|.$$

By Theorem 8.1, we conclude that

$$\pi(x) - \pi(\sqrt{x}) \leq C\frac{x + (\sqrt{x})^2}{L} \ll \frac{x}{L}$$

where

$$L = \sum_{q \in \mathcal{Q}} \mu^2(q) \prod_{p|q} \frac{1}{p-1}.$$

But

$$\begin{aligned} L &\gg \sum_{q \leq \sqrt{x}} \frac{\mu^2(q)}{q} \prod_{p|q} \left(\frac{1}{1 - \frac{1}{p}} \right) \\ &\gg \sum_{q \leq \sqrt{x}} \frac{\mu^2(q)}{q} \prod_{p|q} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots \right) \\ &\gg \sum_{n \leq \sqrt{x}} \frac{1}{n} = \ln x + O(1). \end{aligned}$$

The last estimate in the above follows by first writing

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k} q_1 \cdots q_j$$

where $\alpha_i \geq 2$. We then set

$$n' = p_1 p_2 \cdots p_k q_1 \cdots q_j$$

and

$$n'' = p_1^{\alpha_1 - 1} \cdots p_k^{\alpha_k - 1},$$

and observe that the term

$$\frac{1}{n'n''}$$

appears as a term in the form

$$\frac{\mu^2(n')}{n'} \frac{1}{n''}$$

on the left hand side.

Hence,

$$\pi(x) \ll \frac{x}{\ln x},$$

which is Chebyschev's estimate.

The next example illustrates the use of the Large Sieve in the study of twin primes.

EXAMPLE 8.2

Let $\pi_2(x)$ be the number of primes p less than x such that $p+2$ is also prime. We will show that

$$\pi_2(x) = O\left(\frac{x}{\ln^2 x}\right). \tag{8.2}$$

As a result, we have

$$\sum_{\substack{p\le x\\ p+2 \text{ is a prime}}} \frac{1}{p} = O(1).$$

The last conclusion follows from the expression that

$$\frac{\pi_2(x)}{x} + \int_2^x \frac{\pi_2(t)}{t^2}\,dt \ll 1 + \int_2^x \frac{1}{t\ln^2 t}\,dt \ll 1.$$

To prove (8.2), let $Q = \sqrt{x}$ and $\mathcal{P} = \{p \le \sqrt{x}\}$ and $\mathcal{A} = \{n \le x\}$. Let

$$\mathcal{E} = \{n \in \mathcal{A} | n \not\equiv 0 \pmod p \text{ and } n \not\equiv -2 \pmod p \text{ for all } p \in \mathcal{P}\}.$$

In other words, $\gamma(p) = 2$ if $p \neq 2$ and $\gamma(2) = 1$. Note that $\mathcal{E}$ contains twin primes $r \le x$. By Theorem 8.1, we find that

$$\pi_2(x) - \pi_2(\sqrt{x}) \le |\mathcal{E}| \ll \frac{x+1+x}{L},$$

where

$$L = \sum_{q\in\mathcal{Q}} \mu^2(q) \prod_{p|q} \frac{\gamma(p)}{p-\gamma(p)},$$

with $\mathcal{Q}$ containing integers with prime divisors $p \le \sqrt{x}$. Now,

$$\begin{aligned}
\sum_{q\in\mathcal{Q}} \mu^2(q) \prod_{p|q} \frac{\gamma(p)}{p-\gamma(p)} &= \sum_{q\in\mathcal{Q}} \mu^2(q) \prod_{\substack{p|q\\ p\neq 2}} \frac{2}{p-2} \frac{1}{2-1} \\
&= \frac{1}{2} \sum_{q\in\mathcal{Q}} \mu^2(q) \prod_{\substack{p|q\\ p\neq 2}} \frac{2}{p-2} \frac{2}{2-1} \\
&\gg \sum_{q\in\mathcal{Q}} \mu^2(q) \prod_{p|q} \frac{2}{p-1} \\
&= \sum_{q\in\mathcal{Q}} \frac{\mu^2(q) 2^{\omega(q)}}{q} \prod_{p|q} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots\right).
\end{aligned}$$

The sum

$$\sum_{q\in\mathcal{Q}} \frac{\mu^2(q) 2^{\omega(q)}}{q}$$

adds up term for which q is squarefree. But the sum

$$\sum_{q\in\mathcal{Q}} \frac{\mu^2(q)2^{\omega(q)}}{q} \prod_{p|q} \left(1+\frac{1}{p}+\frac{1}{p^2}+\cdots\right)$$

is a sum of the type

$$\sum_{q\in\mathcal{Q}} \frac{\mu^2(q)2^{\omega(q)}}{q} \left(\sum_{n\in\mathcal{B}} \frac{1}{n}\right),$$

with

$$\mathcal{B} = \{m| \quad p|m \implies p|q\}.$$

Therefore,

$$\sum_{q\in\mathcal{Q}} \frac{\mu^2(q)2^{\omega(q)}}{q} \sum_{n\in\mathcal{B}} \frac{1}{n} \geq \sum_{n\in\mathcal{Q}} \frac{2^{\omega(n)}}{n}. \tag{8.3}$$

The last inequality holds because an integer n appearing on the right hand side can be written as $q\cdot q'$ where q is the squarefree part of n and that the prime divisors of q' are prime divisors of q. Therefore, the term $\frac{2^{\omega(n)}}{n}$ corresponding to n can be written as

$$\frac{2^{\omega(q)}}{q}\cdot\frac{1}{q'}$$

and this term is present in the sum of the left hand side. Now,

$$\sum_{q\in\mathcal{Q}} \frac{2^{\omega(n)}}{n} = \sum_{q\leq\sqrt{x}} \frac{2^{\omega(n)}}{n}.$$

Since

$$\sum_{n\leq y} 2^{\omega(n)} = \frac{6}{\pi^2} y\ln y + O(y),$$

and

$$\sum_{n\leq y} \frac{2^{\omega(n)}}{n} = \frac{3}{\pi^2}\ln^2 y + O(\ln y),$$

we conclude from (8.3) that

$$L \gg \ln^2\sqrt{x}$$

and we complete the proof of (8.2).

8.3 The Large Sieve inequality

In this section, we prove the Large Sieve inequality.

THEOREM 8.2 Let $x_1, x_2, \cdots, x_r$ be δ-spaced. That is to say, if

$$\| x \| = \min_{n \in \mathbf{Z}} |x - n|,$$

then for $k \neq j$,

$$\| x_k - x_j \| \geq \delta.$$

Let

$$S(x) = \sum_{n=1}^{N} a_n e^{2\pi i n x}.$$

Then

$$\sum_{j=1}^{r} |S(x_j)|^2 \leq C \left(N + \frac{1}{\delta} \right) \sum_{n=1}^{N} |a_n|^2,$$

where C is some positive constant.

Remark 8.1 In the above theorem, we restrict our n from 1 to N. We may replace $[1, N]$ by $[M+1, M+N]$ by letting $a_n = b_{M+n} e^{2\pi i M x}$ and setting $k = n + M$. This gives

$$S^*(x) = \sum_{k=M+1}^{M+N} b_k e^{2\pi i M x} e^{2\pi i k x} = \sum_{n=1}^{N} a_n e^{2\pi i n x}.$$

The Large sieve inequality then yields

$$\sum_{j=1}^{R} |S^*(x_j)|^2 \leq C \left(N + \frac{1}{\delta} \right) \sum_{n=1}^{N} |a_n|^2 = C \left(N + \frac{1}{\delta} \right) \sum_{k=M+1}^{M+N} |b_k|^2,$$

since $|e^{2\pi i M x}| = 1$.

To prove Theorem 8.2, we need an intermediate lemma.

LEMMA 8.3 If f has a continuous derivative on $(x - \delta/2, x + \delta/2)$, then

$$|f(x)| \leq \frac{1}{\delta} \int_{x-\delta/2}^{x+\delta/2} |f(y)|\, dy + \frac{1}{2} \int_{x-\delta/2}^{x+\delta/2} |f'(y)|\, dy.$$

Proof
Assume $x = 0$. Note that

$$\int_0^{\delta/2} (\delta/2 - y)\, f'(y)\, dy = -\frac{\delta}{2} f(0) + \int_0^{\delta/2} f(y)\, dy$$

and

$$\int_{-\delta/2}^{0} (\delta/2+y)\, f'(y)\, dy = \frac{\delta}{2} f(0) - \int_{-\delta/2}^{0} f(y)\, dy.$$

Therefore,

$$-\int_{0}^{\delta/2} (\delta/2-y)\, f'(y)\, dy + \int_{-\delta/2}^{0} (\delta/2+y)\, f'(y)\, dy = \delta f(0) - \int_{-\delta/2}^{\delta/2} f(y)\, dy.$$

Now,

$$|\delta/2 - y| \le \delta/2 \text{ for } 0 \le y \le \delta/2$$

and

$$|\delta/2 + y| \le \delta/2 \text{ for } -\delta/2 \le y \le 0.$$

Hence,

$$|\delta f(0)| \le \delta/2 \int_{-\delta/2}^{\delta/2} |f'(y)|\, dy + \int_{-\delta/2}^{\delta/2} |f(y)|\, dy.$$

This proves the result with $x = 0$. Replacing $f(t)$ by $g(x+t)$, we complete the proof of the lemma.

□

We now prove Theorem 8.2.

Proof of Theorem 8.2
Applying Lemma 8.3 with $f(x) = S^2(x)$ and $x = x_i$, we deduce that

$$|S(x_i)|^2 \le \frac{1}{\delta} \int_{x_i-\delta/2}^{x_i+\delta/2} |S(y)|^2 dy + \frac{1}{2} \int_{x_i-\delta/2}^{x_i+\delta/2} |2S(y)S'(y)| dy.$$

Summing over i from 1 to r and observing that $(x_i - \delta/2, x_i + \delta/2)$ modulo 1 are non-overlapping for $i = 1, \cdots, r$, we deduce that

$$\sum_{i=1}^{r} |S(x_i)|^2 \le \frac{1}{\delta} \int_{\alpha}^{\alpha+1} |S(y)|^2 dy + \int_{\alpha}^{\alpha+1} |S(y)S'(y)| dy.$$

Note that

$$\int_{\alpha}^{1+\alpha} |S(y)|^2 dy = \sum_{n=1}^{N} |a_n|^2$$

and

$$\int_{\alpha}^{\alpha+1} |S(y)S'(y)| dy \le \left(\int_{\alpha}^{\alpha+1} |S(y)|^2 dy \right)^{1/2} \left(\int_{\alpha}^{\alpha+1} |S'(y)|^2 dy \right)^{1/2}.$$

Since

$$\int_{\alpha}^{\alpha+1} |S'(y)|^2 dy = (4\pi)^2 \sum_{n=1}^{N} |a_n|^2 n^2$$

and $n \leq N$, we deduce that

$$\sum_{i=1}^{r} |S(x_i)|^2 \leq \frac{1}{\delta} \sum_{n=1}^{N} |a_n|^2 + 2\pi N \left(\sum_{n=1}^{N} |a_n|^2 \right)^{1/2} \left(\sum_{n=1}^{N} |a_n|^2 \right)^{1/2}$$
$$\leq 2\pi \left(\frac{1}{\delta} + N \right) \sum_{n=1}^{N} |a_n|^2.$$

□

8.4 Farey sequence and Theorem 8.1

By a Farey sequence of order n, denoted $\mathcal{F}_n$, we mean a set of reduced fractions in the interval from 0 to 1, whose denominators are less than or equal to n, arranged in ascending order of magnitude.

One fact about elements in Farey sequence is that if $a/b, a'/b'$ are successive terms in $\mathcal{F}_n$, then

$$\frac{a'}{b'} - \frac{a}{b} = \frac{1}{bb'}.$$

Therefore,

$$\left| \frac{a}{b} - \frac{a'}{b'} \right| = \frac{1}{bb'} > \frac{1}{Q^2}$$

if $b \leq Q$. Therefore, the elements in the Farey sequence are $1/Q^2$ well spaced.

Using non-zeroes elements of the Farey sequence of order Q, we conclude from Theorem 8.2 that

$$\sum_{q \leq Q} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \left| S\left(\frac{a}{q}\right) \right|^2 \leq C\left(N + Q^2\right) \sum_{n=1}^{N} |a_n|^2. \tag{8.4}$$

Proof of Theorem 8.1
To prove Theorem 8.1, we will choose a_n such that

$$a_n = 0 \qquad \text{whenever } n \notin \mathcal{E}. \tag{8.5}$$

We will show that it suffices to prove the following inequality:

$$\left| \sum_{n=1}^{N} a_n \right|^2 \mu^2(q) \prod_{p|q} \frac{\gamma(p)}{p - \gamma(p)} \leq \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \left| S\left(\frac{a}{q}\right) \right|^2. \tag{8.6}$$

If (8.6) is true, then by letting

$$a_n = \begin{cases} 1 & \text{if } n \in \mathcal{E}, \\ 0 & \text{otherwise.} \end{cases}$$

From (8.6), we find that

$$\left|\sum_{n=1}^{N} a_n\right|^2 \sum_{q\le Q}\left(\mu^2(q)\prod_{p|q}\frac{\gamma(p)}{p-\gamma(p)}\right) \le \sum_{q\le Q}\sum_{\substack{a=1\\(a,q)=1}}^{q}\left|S\left(\frac{a}{q}\right)\right|^2.$$

Hence,

$$\begin{aligned}|\mathcal{E}|^2\left(\sum_{q\in\mathcal{Q}}\mu^2(q)\prod_{p|q}\frac{\gamma(p)}{p-\gamma(p)}\right) &\le \sum_{q\le Q}\sum_{\substack{a=1\\(a,q)=1}}^{q}\left|S\left(\frac{a}{q}\right)\right|^2\\ &\le C(N+Q^2)\sum_{n=1}^{N}|a_n|^2\\ &\le C(N+Q^2)|\mathcal{E}|.\end{aligned}$$

Therefore,

$$|\mathcal{E}| \le C\frac{N+Q^2}{L},$$

where

$$L=\sum_{q\in\mathcal{Q}}\mu^2(q)\prod_{p|q}\frac{\gamma(p)}{p-\gamma(p)}.$$

We now prove (8.6). First, we observe that if q is not squarefree, then $\mu(q)=0$ and (8.6) is true since its right hand side is non-negative. From now, we may assume q to be **squarefree**.

Let

$$J(q)=\mu^2(q)\prod_{p|q}\frac{\gamma(p)}{p-\gamma(p)}.$$

We now rewrite (8.6) as

$$\sum_{\substack{a=1\\(a,q)=1}}^{q}\left|S\left(\frac{a}{q}\right)\right|^2 \ge |S(0)|^2 J(q). \tag{8.7}$$

We will now show how to establish (8.7) by first showing that it is true prime p. Note that in (8.7), $q=p_1p_2\cdots p_k$ where $p_j, 1\le j\le k$ are distinct primes. Suppose

$$\sum_{\substack{a=1\\(a,p)=1}}^{p}\left|S\left(\frac{a}{p}\right)\right|^2 \ge |S(0)|^2\frac{\gamma(p)}{p-\gamma(p)} = |S(0)|^2 J(p). \tag{8.8}$$

By replacing a_n by $a_ne^{2\pi in\beta}$, we find that

$$\sum_{\substack{a=1\\(a,p)=1}}^{p}\left|S\left(\frac{a}{p}+\beta\right)\right|^2 \ge |S(\beta)|^2 J(p). \tag{8.9}$$

Next, suppose we have proved (8.8) for p, p' with $(p, p') = 1$. Then by the Chinese Remainder Theorem,(8.8) and (8.9), we find that

$$\begin{aligned}\sum_{\substack{c=1\\(a,pp')=1}}^{pp'} \left|S\left(\frac{c}{pp'}\right)\right|^2 &= \sum_{\substack{a=1\\(a,p)=1}}^{p} \sum_{\substack{b=1\\(b,p')=1}}^{p'} \left|S\left(\frac{a}{p}+\frac{b}{p'}\right)\right|^2 \\ &\geq \sum_{\substack{a=1\\(a,p)=1}}^{p} \left|S\left(\frac{a}{p}\right)\right|^2 J(p') \\ &\geq |S(0)|^2 J(p)J(p') = |S(0)|^2 J(pp').\end{aligned}$$

By induction, we conclude that (8.7) holds since q is a product of k distinct primes.

We have seen that it suffices to prove (8.8). Let

$$Z(p,a) = \sum_{\substack{n=1\\ n\equiv a \,(\mathrm{mod}\ p)}}^{N} a_n.$$

We note that

$$|Z(p,a)|^2 = \sum_{\substack{1\leq m,n\leq N\\ m\equiv n\equiv a \,(\mathrm{mod}\ p)}} a_n\overline{a_m}. \tag{8.10}$$

Furthermore, if $n \equiv a \pmod p$ and $a \in \Omega_p$, then $n \notin \mathcal{E}$. This is because an element n in $\mathcal{E}$ must satisfy $n \not\equiv a \pmod p$ for all $a \in \Omega_p$. Hence,

$$Z(p,a) = 0 \quad \text{if } a \in \Omega_p \tag{8.11}$$

since (8.5) implies that $a_n = 0$ whenever $n \notin \mathcal{E}$.

Now,

$$\begin{aligned}\sum_{a=0}^{p-1}\left|S\left(\frac{a}{p}\right)\right|^2 &= \sum_{a=0}^{p-1}\left|\sum_{n=1}^{N} a_n e^{2\pi i n\frac{a}{p}}\right|^2 \\ &= \sum_{a=0}^{p-1}\sum_{1\leq n,m\leq N} a_n\overline{a_m}e^{2\pi i(n-m)\frac{a}{p}} \\ &= p\sum_{\substack{1\leq n,m\leq N\\ n\equiv m \,(\mathrm{mod}\ p)}} a_n\overline{a_m} \\ &= p\sum_{a=0}^{p-1}\sum_{\substack{1\leq n,m\leq N\\ m\equiv n\equiv a \,(\mathrm{mod}\ p)}} a_n\overline{a_m} \\ &= p\sum_{a=0}^{p-1}|Z(p,a)|^2,\end{aligned}$$

where we have used (8.10) in the last equality. In other words, we have

$$\sum_{a=0}^{p-1}\left|S\left(\frac{a}{p}\right)\right|^2 = p\sum_{a=0}^{p-1}|Z(p,a)|^2. \tag{8.12}$$

Let

$$\chi_a = \begin{cases} 1 & \text{if } a \notin \Omega_p \\ 0 & \text{otherwise.} \end{cases}$$

Note that

$$\sum_{a=0}^{p-1} Z(p,a)\chi_a = S(0).$$

By Cauchy's inequality

$$\left|\sum a_n b_n\right|^2 \le \sum |a_n|^2 \sum |b_n|^2,$$

we find that

$$\left|\sum_{a=0}^{p-1} Z(p,a)\chi_a\right|^2 \le \left(\sum_{a=0}^{p-1}\chi_a^2\right)\sum_{a=0}^{p-1}|Z(p,a)|^2 \le (p-\gamma(p))\sum_{a=0}^{p-1}|Z(p,a)|^2. \tag{8.13}$$

But by (8.11) and the definition of $S(x)$, we find that

$$\left|\sum_{a=0}^{p-1} Z(p,a)\chi_a\right|^2 = |S(0)|^2. \tag{8.14}$$

Using (8.12), (8.14) and (8.13), we conclude that

$$|S(0)|^2 \le \frac{p-\gamma(p)}{p}\left(\sum_{a=1}^{p-1}\left|S\left(\frac{a}{p}\right)\right|^2 + |S(0)|^2\right).$$

Simplifying, we find that

$$\sum_{a=1}^{p-1}\left|S\left(\frac{a}{p}\right)\right|^2 \ge \frac{\gamma(p)}{p-\gamma(p)}|S(0)|^2,$$

and this completes the proof of (8.7) for prime number p and the proof of Theorem 8.1 is complete. □

9 Roth's Theorem on Arithmetic Progression

This Chapter is a modified version of H. Halberstam's lecture notes on Circle Method. See also Chapter 10 of R. C. Vaughn's book "Hardy-Littlewood Circle Method."

9.1 Sets without three terms in arithmetic progression

Let n be a positive integer and $M^{(3)}(n)$ denote the largest number of integers in $[1, n]$ having no three terms in arithmetic progression among them. Define

$$\mu^{(3)}(n) = \frac{M^{(3)}(n)}{n}.$$

In this chapter, we will show that

$$\mu^{(3)}(n) = O\left(\frac{1}{\ln \ln n}\right).$$

EXAMPLE 9.1 When $n = 27$, $M^{(3)}(27) = 9$ and s possible set is $\{1, 2, 4, 9, 13, 14, 20, 24, 26\}$.

9.2 Basic inequalities associated with $M^{(3)}(n)$

The first observation about $M^{(3)}(n)$ is the "triangle inequality"

LEMMA 9.1 Let m and n be positive integers. Then

$$M^{(3)}(m+n) \leq M^{(3)}(m) + M^{(3)}(n). \tag{9.1}$$

Let $\mathfrak{M}$ be a set of integers from $[1, m+n]$ with $M^{(3)}(m+n)$ elements. Then $\mathfrak{M} \cap [1, m]$ has at most $M^{(3)}(m)$ elements. The set $\mathfrak{M} \cap [m+1, m+n]$ has at most $M^{(3)}(n)$ elements since $\mathfrak{M} - m \cap [1, n]$ is a subset of $[1, n]$ with no three term in arithmetic progression. Therefore (9.1) holds.

Our next lemma contains several inequalities associated with $\mu^{(3)}(n)$ and will be used throughout the proof of Roth's Theorem.

LEMMA 9.2 Let m and n be positive integers. The following are true:

(i) If $m|n$, then

$$\mu^{(3)}(n) \leq \mu^{(3)}(m). \tag{9.2}$$

(ii) If $m \leq n$, then

$$\mu^{(3)}(n) \leq \left(1 + \frac{1}{[n/m]}\right) \mu^{(3)}(m). \tag{9.3}$$

(iii) The limit $\lim_{n\to\infty} \mu^{(3)}(n) = \mu^{(3)}$ exists.

Proof

If $m|n$ then from (9.1), we deduce that

$$M^{(3)}(n) \leq \frac{n}{m} M^{(3)}(m)$$

and (i) follows immediately.

Suppose $m \leq n$. Write

$$n = qm + r, 0 \leq r < m$$

with $q = [n/m]$. By (9.1),

$$M^{(3)}(n) = M^{(3)}(qm + r) \leq M^{(3)}(qm) + \mu^{(3)}(r) \leq qM^{(3)}(m) + M^{(3)}(r).$$

Since

$$M^{(3)}(r) \leq M^{(3)}(m),$$

we conclude that

$$\begin{aligned} \frac{M^{(3)}(n)}{n} &\leq \frac{q+1}{qm+r} M^{(3)}(m) \leq \frac{(q+1)m}{qm+r} \frac{M^{(3)}(m)}{m} \\ &\leq \mu^{(3)}(m) + \frac{m-r}{qm+r} \mu^{(3)}(m) \leq \left(1 + \frac{1}{q}\right) \mu^{(3)}(m). \end{aligned}$$

This completes the proof of (ii).

Now let $n \to \infty$ in (ii) and deduce that

$$\limsup \mu^{(3)}(n) \leq \mu^{(3)}(m).$$

Let $m \to \infty$ to deduce that

$$\limsup \mu^{(3)}(n) \leq \liminf \mu^{(3)}(m)$$

and this implies that the limit $\lim_{n\to\infty} \mu^{(3)}(n) = \mu^{(3)}$ exists and (iii) is true.

□

9.3 $M^{(3)}(n)$ as an integral

Let $\mathfrak{M}$ be a set no three term arithmetic progression and $|\mathfrak{M}| = M^{(3)}(n)$. If $m_1 + m_3 = 2m_2$ and m_1, m_2, m_3 are in $\mathfrak{M}$, then $m_1 = m_2 = m_3$. This is because if $m_1 \neq m_2$ then the equality will imply that $m_1, m_1+(m_2-m_1), m_1+2(m_2-m_1)$ will be three terms in $\mathfrak{M}$ which are in arithmetic progression. In other words, if $e(\alpha) = e^{2\pi i\alpha}$ and

$$f(\alpha) = \sum_{m\in\mathfrak{M}} e(m\alpha),$$

then

$$M^{(3)}(n) = \int_0^1 f^2(\alpha) f(-2\alpha)\, d\alpha.$$

The above follows from the fact that

$$\int_0^1 e(\alpha t)\, dt = \begin{cases} 1 & \text{if } \alpha = 0 \\ 0 & \text{otherwise.} \end{cases},$$

and the inner integral of the right hand side of

$$\int_0^1 f^2(\alpha) f(-2\alpha)\, d\alpha = \sum_{m_1,m_2,m_3\in\mathfrak{M}} \int_0^1 e((m_1 + m_3 - 2m_2)t)\, dt$$

is non-zero only when $m_1 = m_2 = m_3$.

Write

$$f(\alpha) = \sum_{r=1}^{n} \kappa(r) e(\alpha r) \tag{9.4}$$

where

$$\kappa(r) = \begin{cases} 1 & \text{if } r \in \mathfrak{M} \\ 0 & \text{otherwise.} \end{cases}$$

In the next section, we will use $f(\alpha)$ to obtain more information for $M^{(3)}(n)$.

9.4 Roth's Theorem in arithmetic progression

Let

$$\nu(\alpha) = \mu^{(3)}(m) \sum_{r=1}^{n} e(\alpha r) \tag{9.5}$$

where m will be chosen later.

We now introduce the function

$$E(\alpha) = \nu(\alpha) - f(\alpha) = \sum_{r=1}^{n} (\mu^{(3)}(m) - \kappa(r)) e(\alpha r), \tag{9.6}$$

where f is given by (9.4).

Let

$$c(r) = \mu^{(3)}(m) - \kappa(r). \tag{9.7}$$

Observe that

$$|c(r)| \leq 1 \tag{9.8}$$

since $0 \leq \mu^{(3)}(m) \leq 1$ and $\kappa(r)$ is either 0 or 1.

Let

$$F(\alpha) = \sum_{s=0}^{m-1} e(-\alpha s) \tag{9.9}$$

and supposed that $n \geq qm$. Note that

$$F(\alpha q)E(\alpha) = \sum_{s=0}^{m-1} \sum_{r=1}^{n} c(r)e(\alpha(r - qs)), \tag{9.10}$$

where $c(r)$ is given by (9.7).

Let $h = r - qs$. We rewrite (9.10) as

$$\begin{aligned} F(\alpha q)E(\alpha) &= \sum_{s=0}^{m-1} \sum_{h=1-qs}^{n-qs} c(h + qs)e(\alpha h) \\ &= \sum_{s=0}^{m-1} \left(\sum_{h=1}^{n-qm} c(h + qs)e(\alpha h) - \sum_{h=1-qs}^{0} c(h + qs)e(\alpha h) - \sum_{h=n-qm}^{n-qs-1} c(h + qs)e(\alpha h) \right) \\ &= \sum_{s=0}^{m-1} \sum_{h=1}^{n-qm} c(h + qs)e(\alpha h) - R(\alpha), \end{aligned} \tag{9.11}$$

where

$$|R(\alpha)| \leq \sum_{s=0}^{m-1} \sum_{h=1-qs}^{0} |c(h + qs)| + \sum_{h=n-qm}^{n-qs-1} |c(h + qs) \leq qm^2, \tag{9.12}$$

by (9.8).

Let $\sigma(h) = \sum_{s=0}^{m-1} c(h + qs)$. We may then rewrite (9.11) as

$$F(\alpha q)E(\alpha) = \sum_{h=1}^{n-qm} \sigma(h)e(\alpha h) - R(\alpha). \tag{9.13}$$

The next lemma is important.

LEMMA 9.3 If $n > qm$, then $\sigma(h) \geq 0$.

Proof
Note that

$$\sigma(h) = \sum_{s=0}^{m-1} \Big(\mu^{(3)}(m) - \kappa(h+qs)\Big) = M^{(3)}(m) - K,$$

where K is the number of elements of $\mathfrak{M}$ among $\{h+jq | 0 \le j \le m-1\}$. Let us write these K integers as $\{h+s_i q | 1 \le s \le K\}$. Then $\{1+s_i | 1 \le s \le K\}$ is a subset of $[1,m]$ without 3 term arithmetic progression. In other words, $K \le M^{(3)}(m)$ and therefore,

$$\sigma(h) = M^{(3)}(m) - K \ge 0.$$

□

We now recall a lemma known as Dirichlet approximation theorem.

LEMMA 9.4 Let $\alpha \in \mathbf{R}$ and N be any positive integer. Then there exists positive integers a and q with $q \le N$ such that

$$\left|\alpha - \frac{a}{q}\right| \le \frac{1}{Nq}.$$

Proof
Consider the numbers $0, \{\alpha\}, \{2\alpha\}, \cdots, \{N\alpha\}$. By pigeonhole principle, there exists ℓ and k $(\ell > k)$ such that $\{k\alpha\}$ and $\{\ell\alpha\}$ lie in an interval of the form $[s/N, (s+1)/N]$. This means that

$$|\{\ell\alpha\} - \{k\alpha\}| \le \frac{1}{N},$$

which implies that

$$|(\ell - k)\alpha - ([\ell\alpha] - [k\alpha])| \le \frac{1}{N}.$$

Choosing $a = [\ell\alpha] - [k\alpha]$ and $q = (\ell - k)$ completes the proof. □

We are now ready to estimate $E(\alpha)$ (see (9.6) for the definition of $E(\alpha)$).

LEMMA 9.5 Suppose $n > 2m^2$. Then for every $\alpha \in \mathbf{R}$,

$$|E(\alpha)| \le \frac{\pi}{2} n \Big(\mu^{(3)}(m) - \mu^{(3)}(n)\Big) + 7m^2.$$

Proof
From Lemma 9.4, there are integers a and q such that

$$|q\alpha - a| \le \frac{1}{2m}$$

with $q \leq 2m$. Let $\beta = q\alpha - a$. The function $F(\alpha q)$ (see (9.9) for the definition of $F(\alpha)$) can then be written as

$$F(\alpha q) = \sum_{s=0}^{m-1} e(-s\alpha q) = \sum_{s=0}^{m-1} e(-s(\beta - a)) = F(\beta).$$

Now,

$$|F(\beta)| = \left| \frac{1 - e(-\beta m)}{1 - e(-\beta)} \right| = \left| \frac{\sin \pi \beta m}{\sin \pi \beta} \right|.$$

Since $\sin \pi x \leq \pi x$ and $\sin \pi x \geq 2x$ for $0 \leq x \leq 1/2$, we conclude that

$$|F(\alpha q)| = |F(\beta)| = \left| \frac{\sin \pi \beta m}{\sin \pi \beta} \right| \geq \frac{2m}{\pi}.$$

Therefore, from (9.13), we conclude that

$$\begin{aligned} \frac{2m}{\pi} |E(\alpha)| \leq |F(\alpha q)||E(\alpha)| &\leq \sum_{h=0}^{n-mq} \sigma(h) + |R(\alpha)| \\ &= F(0)E(0) - R(0) + |R(\alpha)|, \end{aligned}$$

where we have used (9.3) to conclude that $|\sigma(h)| = \sigma(h)$ and (9.13) to deduce that

$$\sum_{h=0}^{n-mq} \sigma(h) = F(0)E(0) - R(0).$$

Hence,

$$|E(\alpha)| \leq \frac{\pi}{2m} \left(mE(0) + 2qm^2 \right)$$

where we have used (9.12) to deduce that

$$-R(0) + |R(\alpha)| \leq |R(0)| + |R(\alpha)| \leq 2qm^2.$$

Here we check that $n \geq 2m^2 > qm$. Recall from (9.6) that

$$E(0) = \nu(0) - f(0) = n\mu^{(3)}(m) - n\mu^{(3)}(()n).$$

Therefore,

$$|E(\alpha)| \leq \frac{\pi}{2} n \left(\mu^{(3)}(m) - \mu^{(3)}(n) \right) + 2m^2\pi$$

and the proof is complete after bounding 2π by 7.

□

Recall that

$$M^{(3)}(n) = \int_0^1 f^2(\alpha) f(-2\alpha)\, d\alpha = I - \int_0^1 f^2(\alpha) E(-2\alpha)\, d\alpha,$$

where we have used (9.6). Note that

$$I = \mu^{(3)}(m) \sum_{a\in\mathfrak{M}} \sum_{b\in\mathfrak{M}} \sum_{r=1}^{n} \int_0^1 e(\alpha(a+b-2r)\, d\alpha$$
$$= \begin{cases} 1 & \text{if } a+b \equiv 0 \pmod 2 \\ 0 & \text{otherwise.} \end{cases}.$$

Let M_E and M_O be the number of even integers and odd integers in $\mathfrak{M}$ respectively. Then

$$I = \mu^{(3)}(m)\left(M_E^2 + M_O^2\right) \geq \frac{\mu^{(3)}(m)}{2}\left(M_E + M_O\right)^2 = \frac{\mu^{(3)}(m)}{2} M^{(3)}(n)^2, \quad (9.14)$$

where we have used the inequality

$$2(s^2+t^2) \geq (s+t)^2.$$

Now,

$$|M^{(3)}(n) - I| \leq \int_0^1 |f(\alpha)|^2 |E(-2\alpha)|\, d\alpha \leq \int_0^1 |f(\alpha)|^2\, d\alpha \left(\frac{\pi}{2} n \left(\mu^{(3)}(m) - \mu^{(3)}(n)\right) + 7m^2\right),$$

where we have used Lemma 9.5. By Parseval's identity,

$$\int_0^1 |f(\alpha)|^2\, d\alpha = M^{(3)}(n),$$

and therefore, if $n \geq 2m^2$,

$$|M^{(3)}(n) - I| \leq M^{(3)}(n) \left(\frac{\pi}{2} n \left(\mu^{(3)}(m) - \mu^{(3)}(n)\right) + 7m^2\right). \quad (9.15)$$

Next,

$$|M^{(3)}(n) - I| \geq I - M^{(3)}(n) \geq \frac{\mu^{(3)}(m)}{2} M^{(3)}(n)^2 - M^{(3)}(n)$$

where we have used (9.14).Therefore, we may deduce from (9.15) that

$$\frac{M^{(3)}(n)}{2} \mu^{(3)}(m) - 1 \leq \frac{\pi}{2} n \left(\mu^{(3)}(m) - \mu^{(3)}(n)\right) + 7m^2.$$

Dividing both sides of the inequality by n, we conclude that

$$\mu^{(3)}(n)\mu^{(3)}(m) \leq \pi \left(\mu^{(3)}(m) - \mu^{(3)}(n)\right) + \frac{14m^2}{n} + \frac{2}{n}. \quad (9.16)$$

Letting $n \to \infty$ followed by $m \to \infty$ and using the fact that $\lim_{n\to\infty} \mu^{(3)}(n) = \mu^{(3)}$ exists (see Lemma 9.2), we conclude that

$$\left(\mu^{(3)}\right)^2 \leq 0$$

and therefore,

$$\lim_{n\to\infty} \mu^{(3)}(n) = 0$$

and the weak form of Roth's Theorem is true.

To obtain the strong form of Roth's theorem, let $n = 2^{3^k}$ and $m = 2^{3^{k-1}}$. Note that $m^3 = n$ and therefore $m^2/n = 1/m \leq 2$ which means that $n \geq 2m^2$. Let $\lambda(k) = \mu^{(3)}(2^{3^k})$. Then from (9.16), we deduce that

$$\lambda(k-1)\lambda(k) \leq \pi(\lambda(k-1) - \lambda(k)) + \frac{14m^2}{n} + \frac{2}{n},$$

or

$$1 \leq \pi\left(\frac{1}{\lambda(k)} - \frac{1}{\lambda(k-1)}\right) + \frac{30}{2^{3^k}\lambda^2(k)},$$

where we have used

$$\lambda(k) \leq \lambda(k-1).$$

From now on, we will not worry about the constant 30 and simply replace it by c. Now we sum k from ℓ to $2\ell + 1$ to deduce that

$$\ell \leq \frac{\pi}{\lambda(2\ell)} + \frac{c'\ell}{2^{3^\ell}\lambda^2(2\ell)}.$$

Here we have used the bound $\lambda(2\ell) \leq 2\lambda(k)$ and $2^{3^\ell} \leq 2^{3^k}$ for $\ell \leq k \leq 2\ell + 1$. We claim that

$$\lambda(2\ell) \leq \frac{C}{\ell}$$

for some constant C. If this were true, then we are done. Suppose

$$\lambda(2\ell) > \frac{C}{\ell}.$$

Then

$$\lambda(2\ell) \leq \frac{\pi}{\ell} + \frac{c'}{2^{3^\ell}\lambda(2\ell)} \leq \frac{\pi}{\ell} + \frac{c'\ell}{2^{3^\ell}C}.$$

Choose ℓ large enough so that

$$\frac{\ell}{2^{3^\ell}} \leq \frac{C^*}{\ell}$$

and we conclude that

$$\lambda(2\ell) \leq \frac{C^\dagger}{\ell}$$

for some constant $C^\dagger$. So we know that for sufficiently large ℓ,

$$\lambda(2\ell) \leq \frac{d}{\ell}$$

for some constant d. Using the same argument, we may deduce that

$$\lambda(2\ell + 1) \leq \frac{d'}{\ell}.$$

Therefore

$$\lambda(\ell) \leq \frac{d''}{\ell}.$$

Let n be sufficiently large and choose ℓ such that

$$2^{3^\ell} \leq n < 2^{3^{\ell+1}}.$$

Note that ℓ is of the same order as $\ln \ln n$. Therefore,

$$\mu^{(3)}(n) = \lambda(\ell) \ll \frac{1}{\ell}$$

implies that

$$\mu^{(3)}(n) = O\left(\frac{1}{\ln \ln n}\right)$$

and this completes the proof of Roth's Theorem on arithmetic progression.