DISTRIBUTION OF PRIME NUMBERS

W W L CHEN

© W W L Chen, 1981, 2003.
This chapter originates from material used by the author at Imperial College, University of London, between 1981 and 1990.
It is available free to all individuals, on the understanding that it is not to be used for financial gains,
and may be downloaded and/or photocopied, with or without permission from the author.
However, this document may not be kept on any information storage and retrieval system without permission
from the author, unless such system is not accessible to any individuals other than its owners.

Chapter 1

ARITHMETIC FUNCTIONS

1.1. Introduction

By an arithmetic function, we mean a function of the form $f : \mathbb{N} \to \mathbb{C}$. We say that an arithmetic function $f : \mathbb{N} \to \mathbb{C}$ is multiplicative if $f(mn) = f(m)f(n)$ whenever $m, n \in \mathbb{N}$ and $(m, n) = 1$.

Example. The function $U : \mathbb{N} \to \mathbb{C}$, defined by $U(n) = 1$ for every $n \in \mathbb{N}$, is an arithmetic function. Furthermore, it is multiplicative.

THEOREM 1A. *Suppose that the function $f : \mathbb{N} \to \mathbb{C}$ is multiplicative. Then the function $g : \mathbb{N} \to \mathbb{C}$, defined by*

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

for every $n \in \mathbb{N}$, is multiplicative.

Here the summation $\sum_{m|n}$ denotes a sum over all positive divisors m of n.

Proof of Theorem 1A. Suppose that $a, b \in \mathbb{N}$ and $(a, b) = 1$. If u is a positive divisor of a and v is a positive divisor of b, then clearly uv is a positive divisor of ab. On the other hand, every positive divisor m of ab can be expressed uniquely in the form $m = uv$, where u is a positive divisor of a and v is a positive divisor of b. It follows that

$$g(ab) = \sum_{m|ab} f(m) = \sum_{u|a} \sum_{v|b} f(uv) = \sum_{u|a} \sum_{v|b} f(u)f(v) = \left(\sum_{u|a} f(u) \right) \left(\sum_{v|b} f(v) \right) = g(a)g(b).$$

This completes the proof. ○

1.2. The Divisor Function

We define the divisor function $d:\mathbb{N}\to\mathbb{C}$ by writing

$$d(n)=\sum_{m\mid n}1 \tag{1}$$

for every $n\in\mathbb{N}$. Here the sum is taken over all positive divisors m of n. In other words, the value $d(n)$ denotes the number of positive divisors of the natural number n. On the other hand, we define the function $\sigma:\mathbb{N}\to\mathbb{C}$ by writing

$$\sigma(n)=\sum_{m\mid n}m \tag{2}$$

for every $n\in\mathbb{N}$. Clearly, the value $\sigma(n)$ denotes the sum of all the positive divisors of the natural number n.

THEOREM 1B. *Suppose that $n\in\mathbb{N}$ and that $n=p_1^{u_1}\dots p_r^{u_r}$ is the canonical decomposition of n. Then*

$$d(n)=(1+u_1)\dots(1+u_r)\qquad \textit{and}\qquad \sigma(n)=\frac{p_1^{u_1+1}-1}{p_1-1}\dots\frac{p_r^{u_r+1}-1}{p_r-1}.$$

PROOF. Every positive divisor m of n is of the form $m=p_1^{v_1}\dots p_r^{v_r}$, where for every $j=1,\dots,r$, the integer v_j satisfies $0\le v_j\le u_j$. It follows from (1) that $d(n)$ is the number of choices for the r-tuple $(v_1,\dots,v_r)$. Hence

$$d(n)=\sum_{v_1=0}^{u_1}\dots\sum_{v_r=0}^{u_r}1=(1+u_1)\dots(1+u_r).$$

On the other hand, it follows from (2) that

$$\sigma(n)=\sum_{v_1=0}^{u_1}\dots\sum_{v_r=0}^{u_r}p_1^{v_1}\dots p_r^{v_r}=\left(\sum_{v_1=0}^{u_1}p_1^{v_1}\right)\dots\left(\sum_{v_r=0}^{u_r}p_r^{v_r}\right).$$

Note now that for every $j=1,\dots,r$, we have

$$\sum_{v_j=0}^{u_j}p_j^{v_j}=1+p_j+p_j^2+\dots+p_j^{u_j}=\frac{p_j^{u_j+1}-1}{p_j-1}.$$

The second assertion follows. ◯

The result below is a simple deduction from Theorem 1B.

THEOREM 1C. *The arithmetic functions $d:\mathbb{N}\to\mathbb{C}$ and $\sigma:\mathbb{N}\to\mathbb{C}$ are both multiplicative.*

Natural numbers $n\in\mathbb{N}$ where $\sigma(n)=2n$ are of particular interest, and are known as perfect numbers. A perfect number is therefore a natural number which is equal to the sum of its own proper divisors; in other words, the sum of all its positive divisors other than itself.

EXAMPLES. It is easy to see that $6=1+2+3$ and $28=1+2+4+7+14$ are perfect numbers.

It is not known whether any odd perfect number exists. However, we can classify the even perfect numbers.

THEOREM 1D. (EUCLID-EULER) *Suppose that $m \in \mathbb{N}$. If $2^m - 1$ is a prime, then the number $2^{m-1}(2^m - 1)$ is an even perfect number. Furthermore, there are no other even perfect numbers.*

PROOF. Suppose that $n = 2^{m-1}(2^m - 1)$, and $2^m - 1$ is prime. Clearly

$$(2^{m-1}, 2^m - 1) = 1.$$

It follows from Theorems 1B and 1C that

$$\sigma(n) = \sigma(2^{m-1})\sigma(2^m - 1) = \frac{2^m - 1}{2 - 1} 2^m = 2n,$$

so that n is a perfect number, clearly even since $m \geq 2$.

Suppose now that $n \in \mathbb{N}$ is an even perfect number. Then we can write $n = 2^{m-1}u$, where $m \in \mathbb{N}$ and $m > 1$, and where $u \in \mathbb{N}$ is odd. By Theorem 1B, we have

$$2^m u = \sigma(n) = \sigma(2^{m-1})\sigma(u) = (2^m - 1)\sigma(u),$$

so that

$$\sigma(u) = \frac{2^m u}{2^m - 1} = u + \frac{u}{2^m - 1}. \tag{3}$$

Note that $\sigma(u)$ and u are integers and $\sigma(u) > u$. Hence $u/(2^m - 1) \in \mathbb{N}$ and is a divisor of u. Since $m > 1$, we have $2^m - 1 > 1$, and so $u/(2^m - 1) \neq u$. It now follows from (3) that $\sigma(u)$ is equal to the sum of two of its positive divisors. But $\sigma(u)$ is equal to the sum of all its positive divisors. Hence u must have exactly two positive divisors, so that u is prime. Furthermore, we must have $u/(2^m - 1) = 1$, so that $u = 2^m - 1$. ◯

We are interested in the behaviour of $d(n)$ and $\sigma(n)$ as $n \to \infty$. If $n \in \mathbb{N}$ is a prime, then clearly $d(n) = 2$. Also, the magnitude of $d(n)$ is sometimes greater than that of any power of $\log n$. More precisely, we have the following result.

THEOREM 1E. *For any fixed real number $c > 0$, the inequality $d(n) \ll (\log n)^c$ as $n \to \infty$ does not hold.*

PROOF. The idea of the proof is to consider integers which are divisible by many different primes. Suppose that $c > 0$ is given and fixed. Let $\ell \in \mathbb{N} \cup \{0\}$ satisfy $\ell \leq c < \ell + 1$. For every $j = 1, 2, 3, \ldots$, let p_j denote the j-th positive prime in increasing order of magnitude, and consider the integer

$$n = (p_1 \ldots p_{\ell+1})^m.$$

Then in view of Theorem 1B, we have

$$d(n) = (m+1)^{\ell+1} > \left(\frac{\log n}{\log(p_1 \ldots p_{\ell+1})} \right)^{\ell+1} > K(c)(\log n)^{\ell+1} > K(c)(\log n)^c, \tag{4}$$

where the positive constant

$$K(c) = \left(\frac{1}{\log(p_1 \ldots p_{\ell+1})} \right)^{\ell+1}$$

depends only on c. The result follows on noting that the inequality (4) holds for every $m \in \mathbb{N}$. ◯

On the other hand, the order of magnitude of $d(n)$ cannot be too large either.

THEOREM 1F. *For any fixed real number $\epsilon > 0$, we have $d(n) \ll_\epsilon n^\epsilon$ as $n \to \infty$.*

PROOF. For every natural number $n > 1$, let $n = p_1^{u_1} \ldots p_r^{u_r}$ be its canonical decomposition. It follows from Theorem 1B that

$$\frac{d(n)}{n^\epsilon} = \frac{(1+u_1)}{p_1^{\epsilon u_1}} \ldots \frac{(1+u_r)}{p_r^{\epsilon u_r}}.$$

We may assume without loss of generality that $\epsilon < 1$. If $2 \le p_j < 2^{1/\epsilon}$, then

$$p_j^{\epsilon u_j} \ge 2^{\epsilon u_j} = \mathrm{e}^{\epsilon u_j \log 2} > 1 + \epsilon u_j \log 2 > (1+u_j)\epsilon \log 2,$$

so that

$$\frac{(1+u_j)}{p_j^{\epsilon u_j}} < \frac{1}{\epsilon \log 2}.$$

On the other hand, if $p_j \ge 2^{1/\epsilon}$, then $p_j^\epsilon \ge 2$, and so

$$\frac{(1+u_j)}{p_j^{\epsilon u_j}} \le \frac{1+u_j}{2^{u_j}} \le 1.$$

It follows that

$$\frac{d(n)}{n^\epsilon} < \prod_{p < 2^{1/\epsilon}} \frac{1}{\epsilon \log 2},$$

a positive constant depending only on ϵ. ○

We see from Theorems 1E and 1F and the fact that $d(n) = 2$ infinitely often that the magnitude of $d(n)$ fluctuates a great deal as $n \to \infty$. It may then be more fruitful to average the function $d(n)$ over a range of values n, and consider, for positive real numbers $X \in \mathbb{R}$, the value of the average

$$\frac{1}{X} \sum_{n \le X} d(n).$$

THEOREM 1G. (DIRICHLET) *As $X \to \infty$, we have*

$$\sum_{n \le X} d(n) = X \log X + (2\gamma - 1)X + O(X^{1/2}).$$

Here γ is Euler's constant and is defined by

$$\gamma = \lim_{Y \to \infty} \left(\sum_{n \le Y} \frac{1}{n} - \log Y \right) = 0.5772156649\ldots.$$

REMARK. It is an open problem in mathematics to determine whether Euler's constant γ is rational or irrational.

The proof of Theorem 1G depends on the following intermediate result.

THEOREM 1H. *As $Y \to \infty$, we have*

$$\sum_{n \le Y} \frac{1}{n} = \log Y + \gamma + O\left(\frac{1}{Y}\right).$$

PROOF. As $Y \to \infty$, we have

$$\begin{aligned}
\sum_{n \le Y} \frac{1}{n} &= \sum_{n \le Y} \left(\frac{1}{Y} + \int_n^Y \frac{1}{u^2} \, \mathrm{d}u \right) = \frac{[Y]}{Y} + \sum_{n \le Y} \int_n^Y \frac{1}{u^2} \, \mathrm{d}u = \frac{[Y]}{Y} + \int_1^Y \frac{1}{u^2} \left(\sum_{n \le u} 1 \right) \mathrm{d}u \\
&= \frac{[Y]}{Y} + \int_1^Y \frac{[u]}{u^2} \, \mathrm{d}u = \frac{[Y]}{Y} + \int_1^Y \frac{1}{u} \, \mathrm{d}u - \int_1^Y \frac{u-[u]}{u^2} \, \mathrm{d}u \\
&= \log Y + 1 + O\left(\frac{1}{Y}\right) - \int_1^\infty \frac{u-[u]}{u^2} \, \mathrm{d}u + \int_Y^\infty \frac{u-[u]}{u^2} \, \mathrm{d}u \\
&= \log Y + \left(1 - \int_1^\infty \frac{u-[u]}{u^2} \, \mathrm{d}u \right) + O\left(\frac{1}{Y}\right).
\end{aligned}$$

It is a simple exercise to show that

$$1 - \int_1^\infty \frac{u-[u]}{u^2} \, \mathrm{d}u = \gamma.$$

and this completes the proof. ○

PROOF OF THEOREM 1G. As $X \to \infty$, we have

$$\begin{aligned}
\sum_{n \le X} d(n) = \sum_{\substack{x,y \\ xy \le X}} 1 &= \sum_{x \le X^{1/2}} \sum_{y \le \frac{X}{x}} 1 + \sum_{y \le X^{1/2}} \sum_{x \le \frac{X}{y}} 1 - \sum_{x \le X^{1/2}} \sum_{y \le X^{1/2}} 1 \\
&= 2 \sum_{x \le X^{1/2}} \left[\frac{X}{x} \right] - [X^{1/2}]^2 = 2 \sum_{x \le X^{1/2}} \frac{X}{x} + O(X^{1/2}) - (X^{1/2} + O(1))^2 \\
&= 2X \left(\log X^{1/2} + \gamma + O\left(\frac{1}{X^{1/2}} \right) \right) + O(X^{1/2}) - X \\
&= X \log X + (2\gamma - 1)X + O(X^{1/2}).
\end{aligned}$$

This completes the proof. ○

We next turn our attention to the study of the behaviour of $\sigma(n)$ as $n \to \infty$. Every number $n \in \mathbb{N}$ has divisors 1 and n, so we must have $\sigma(1) = 1$ and $\sigma(n) > n$ if $n > 1$. On the other hand, it follows from Theorem 1F that for any fixed real number $\epsilon > 0$, we have

$$\sigma(n) \le n d(n) \ll_\epsilon n^{1+\epsilon} \quad \text{as } n \to \infty.$$

In fact, it is rather easy to prove a slightly stronger result.

THEOREM 1J. *We have $\sigma(n) \ll n \log n$ as $n \to \infty$.*

PROOF. As $n \to \infty$, we have

$$\sigma(n) = \sum_{m \mid n} \frac{n}{m} \le n \sum_{m \le n} \frac{1}{m} \ll n \log n.$$

This completes the proof. ○

As in the case of $d(n)$, the magnitude of $\sigma(n)$ fluctuates a great deal as $n \to \infty$. As before, we shall average the function $\sigma(n)$ over a range of values n, and consider some average version of the function. Corresponding to Theorem 1G, we have the following result.

THEOREM 1K. *As $X \to \infty$, we have*

$$\sum_{n \le X} \sigma(n) = \frac{\pi^2}{12} X^2 + O(X \log X).$$

PROOF. As $X \to \infty$, we have

$$\begin{aligned}
\sum_{n \le X} \sigma(n) &= \sum_{n \le X} \sum_{m \mid n} \frac{n}{m} = \sum_{m \le X} \sum_{\substack{n \le X \\ m \mid n}} \frac{n}{m} = \sum_{m \le X} \sum_{r \le X/m} r = \sum_{m \le X} \frac{1}{2} \left[\frac{X}{m}\right] \left(1 + \left[\frac{X}{m}\right]\right) \\
&= \frac{1}{2} \sum_{m \le X} \left(\frac{X}{m} + O(1)\right)^2 = \frac{X^2}{2} \sum_{m \le X} \frac{1}{m^2} + O\left(X \sum_{m \le X} \frac{1}{m}\right) + O\left(\sum_{m \le X} 1\right) \\
&= \frac{X^2}{2} \sum_{m=1}^{\infty} \frac{1}{m^2} + O\left(X^2 \sum_{m > X} \frac{1}{m^2}\right) + O(X \log X) = \frac{\pi^2}{12} X^2 + O(X \log X).
\end{aligned}$$

This completes the proof. ◯

1.3. The Möbius Function

We define the Möbius function $\mu : \mathbb{N} \to \mathbb{C}$ by writing

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1, \\ (-1)^r & \text{if } n = p_1 \ldots p_r, \text{ a product of distinct primes,} \\ 0 & \text{otherwise.} \end{cases}$$

REMARKS. (i) A natural number which is not divisible by the square of any prime is called a squarefree number. Note that 1 is both a square and a squarefree number. Furthermore, a number $n \in \mathbb{N}$ is squarefree if and only if $\mu(n) = \pm 1$.

(ii) The motivation for the definition of the Möbius function lies rather deep. To understand the definition, one needs to study the Riemann zeta function, an important function in the study of the distribution of primes. At this point, it suffices to remark that the Möbius function is defined so that if we formally multiply the two series

$$\sum_{n=1}^{\infty} \frac{1}{n^s} \qquad \text{and} \qquad \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s},$$

where $s \in \mathbb{C}$ denotes a complex variable, then the product is identically equal to 1. Heuristically, note that

$$\left(\sum_{k=1}^{\infty} \frac{1}{k^s}\right) \left(\sum_{m=1}^{\infty} \frac{\mu(m)}{m^s}\right) = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \sum_{\substack{m=1 \\ km=n}}^{\infty} \frac{\mu(m)}{n^s} = \sum_{n=1}^{\infty} \left(\sum_{m \mid n} \mu(m)\right) \frac{1}{n^s}.$$

It follows that the product is identically equal to 1 if

$$\sum_{m|n} \mu(m) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1. \end{cases}$$

We shall establish this last fact and study some of its consequences over the next four theorems.

THEOREM 1L. *The Möbius function $\mu : \mathbb{N} \to \mathbb{C}$ is multiplicative.*

PROOF. Suppose that $a, b \in \mathbb{N}$ and $(a, b) = 1$. If a or b is not squarefree, then neither is ab, and so $\mu(ab) = 0 = \mu(a)\mu(b)$. On the other hand, if both a and b are squarefree, then since $(a, b) = 1$, ab must also be squarefree. Furthermore, the number of prime factors of ab must be the sum of the numbers of prime factors of a and of b. ◯

THEOREM 1M. *Suppose that $n \in \mathbb{N}$. Then*

$$\sum_{m|n} \mu(m) = \begin{cases} 1 & \textit{if } n = 1, \\ 0 & \textit{if } n > 1. \end{cases}$$

PROOF. Consider the function $f : \mathbb{N} \to \mathbb{C}$ defined by writing

$$f(n) = \sum_{m|n} \mu(m)$$

for every $n \in \mathbb{N}$. It follows from Theorems 1A and 1L that f is multiplicative. For $n = 1$, the result is trivial. To complete the proof, it therefore suffices to show that $f(p^k) = 0$ for every prime p and every $k \in \mathbb{N}$. Indeed,

$$f(p^k) = \sum_{m|p^k} \mu(m) = \mu(1) + \mu(p) + \mu(p^2) + \ldots + \mu(p^k) = 1 - 1 + 0 + \ldots + 0 = 0.$$

This completes the proof. ◯

Theorem 1M plays the central role in the proof of the following two results which are similar in nature.

THEOREM 1N. (MÖBIUS INVERSION FORMULA) *For any function $f : \mathbb{N} \to \mathbb{C}$, if the function $g : \mathbb{N} \to \mathbb{C}$ is defined by writing*

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

for every $n \in \mathbb{N}$, then for every $n \in \mathbb{N}$, we have

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

PROOF. The second equality is obvious. Also

$$\sum_{m|n} \mu(m)\, g\left(\frac{n}{m}\right) = \sum_{m|n} \mu(m) \left(\sum_{k|\frac{n}{m}} f(k) \right) = \sum_{\substack{k,m \\ km|n}} \mu(m) f(k) = \sum_{k|n} f(k) \left(\sum_{m|\frac{n}{k}} \mu(m) \right) = f(n),$$

in view of Theorem 1M. ◯

THEOREM 1P. *For any function $g : \mathbb{N} \to \mathbb{C}$, if the function $f : \mathbb{N} \to \mathbb{C}$ is defined by writing*

$$f(n) = \sum_{m\mid n} \mu\left(\frac{n}{m}\right) g(m)$$

for every $n \in \mathbb{N}$, then for every $n \in \mathbb{N}$, we have

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

Proof. The second equality is obvious. Also

$$\sum_{m\mid n} f\left(\frac{n}{m}\right) = \sum_{m\mid n} \left(\sum_{k\mid \frac{n}{m}} \mu\left(\frac{n}{mk}\right) g(k) \right) = \sum_{k\mid n} g(k) \left(\sum_{m\mid \frac{n}{k}} \mu\left(\frac{n/k}{m}\right) \right) = \sum_{k\mid n} g(k) \left(\sum_{m\mid \frac{n}{k}} \mu(m) \right) = g(n),$$

in view of Theorem 1M. ○

Remark. In number theory, it occurs quite often that in the proof of a theorem, a change of order of summation of the variables is required, as illustrated in the proofs of Theorems 1N and 1P. This process of changing the order of summation does not depend on the summand in question. In both instances, we are concerned with a sum of the form

$$\sum_{m\mid n} \sum_{k\mid \frac{n}{m}} A(k,m).$$

This means that for every positive divisor m of n, we first sum the function A over all positive divisors k of n/m to obtain the sum

$$\sum_{k\mid \frac{n}{m}} A(k,m),$$

which is a function of m. We then sum this sum over all divisors m of n. Now observe that for every natural number k satisfying $k \mid n/m$ for some positive divisor m of n, we must have $k \mid n$. Consider therefore a particular natural number k satisfying $k \mid n$. We must find all natural numbers m satisfying the original summation conditions, namely $m \mid n$ and $k \mid n/m$. These are precisely those natural numbers m satisfying $m \mid n/k$. We therefore obtain, for every positive divisor k of n, the sum

$$\sum_{m\mid \frac{n}{k}} A(k,m).$$

Summing over all positive divisors k of n, we obtain

$$\sum_{k\mid n} \sum_{m\mid \frac{n}{k}} A(k,m).$$

Since we are summing the function A over the same collection of pairs (k,m), and have merely changed the order of summation, we must have

$$\sum_{m\mid n} \sum_{k\mid \frac{n}{m}} A(k,m) = \sum_{k\mid n} \sum_{m\mid \frac{n}{k}} A(k,m).$$

1.4. The Euler Function

We define the Euler function $\phi : \mathbb{N} \to \mathbb{C}$ as follows. For every $n \in \mathbb{N}$, we let $\phi(n)$ denote the number of elements in the set $\{1, 2, \ldots, n\}$ which are coprime to n.

THEOREM 1Q. *For every number $n \in \mathbb{N}$, we have*

$$\sum_{m|n} \phi(m) = n.$$

PROOF. We shall partition the set $\{1, 2, \ldots, n\}$ into $d(n)$ disjoint subsets $\mathcal{B}_m$, where for every positive divisor m of n,

$$\mathcal{B}_m = \{x : 1 \le x \le n \text{ and } (x, n) = m\}.$$

If $x \in \mathcal{B}_m$, let $x = mx'$. Then $(mx', n) = m$ if and only if $(x', n/m) = 1$. Also $1 \le x \le n$ if and only if $1 \le x' \le n/m$. Hence

$$\mathcal{B}'_m = \{x' : 1 \le x' \le n/m \text{ and } (x', n/m) = 1\}$$

has the same number of elements as $\mathcal{B}_m$. Note now that the number of elements of $\mathcal{B}'_m$ is exactly $\phi(n/m)$. Since every element of the set $\{1, 2, \ldots, n\}$ falls into exactly one of the subsets $\mathcal{B}_m$, we must have

$$n = \sum_{m|n} \phi\left(\frac{n}{m}\right) = \sum_{m|n} \phi(m).$$

This completes the proof. ○

Applying the Möbius inversion formula to the conclusion of Theorem 1Q, we obtain immediately the following result.

THEOREM 1R. *For every number $n \in \mathbb{N}$, we have*

$$\phi(n) = \sum_{m|n} \mu(m)\frac{n}{m} = n\sum_{m|n} \frac{\mu(m)}{m}.$$

THEOREM 1S. *The Euler function $\phi : \mathbb{N} \to \mathbb{C}$ is multiplicative.*

PROOF. Since the Möbius function μ is multiplicative, it follows that the function $f : \mathbb{N} \to \mathbb{C}$, defined by $f(n) = \mu(n)/n$ for every $n \in \mathbb{N}$, is multiplicative. The result now follows from Theorem 1A. ○

THEOREM 1T. *Suppose that $n \in \mathbb{N}$ and $n > 1$, with canonical decomposition $n = p_1^{u_1} \ldots p_r^{u_r}$. Then*

$$\phi(n) = n\prod_{j=1}^{r}\left(1 - \frac{1}{p_j}\right) = \prod_{j=1}^{r} p_j^{u_j-1}(p_j - 1).$$

PROOF. The second equality is trivial. On the other hand, for every prime p and every $u \in \mathbb{N}$, we have by Theorem 1R that

$$\frac{\phi(p^u)}{p^u} = \sum_{m|p^u} \frac{\mu(m)}{m} = 1 + \frac{\mu(p)}{p} = 1 - \frac{1}{p}.$$

The result now follows since ϕ is multiplicative. ○

We now study the magnitude of $\phi(n)$ as $n \to \infty$. Clearly $\phi(1) = 1$ and $\phi(n) < n$ if $n > 1$.

Suppose first of all that n has many different prime factors. Then n must have many different divisors, and so $\sigma(n)$ must be large relative to n. But then many of the numbers $1, \ldots, n$ cannot be coprime to n, and so $\phi(n)$ must be small relative to n. On the other hand, suppose that n has very few prime factors. Then n must have very few divisors, and so $\sigma(n)$ must be small relative to n. But then many of the numbers $1, \ldots, n$ are coprime to n, and so $\phi(n)$ must be large relative to n. It therefore appears that if one of the two values $\sigma(n)$ and $\phi(n)$ is large relative to n, then the other must be small relative to n. Indeed, our heuristics are upheld by the following result.

THEOREM 1U. *For every $n \in \mathbb{N}$, we have*

$$\frac{1}{2} < \frac{\sigma(n)\phi(n)}{n^2} \le 1.$$

PROOF. The result is obvious if $n = 1$, so suppose that $n > 1$. Let $n = p_1^{u_1} \ldots p_r^{u_r}$ be the canonical decomposition of n. Recall Theorems 1B and 1T. We have

$$\sigma(n) = \prod_{j=1}^{r} \frac{p_j^{u_j+1} - 1}{p_j - 1} = n \prod_{j=1}^{r} \frac{1 - p_j^{-u_j-1}}{1 - p_j^{-1}}$$

and

$$\phi(n) = n \prod_{j=1}^{r} (1 - p_j^{-1}).$$

Hence

$$\frac{\sigma(n)\phi(n)}{n^2} = \prod_{j=1}^{r} (1 - p_j^{-u_j-1}).$$

The upper bound follows at once. On the other hand,

$$\prod_{j=1}^{r} (1 - p_j^{-u_j-1}) \ge \prod_{p|n} (1 - p^{-2}) \ge \prod_{m=2}^{n} \left(1 - \frac{1}{m^2}\right) = \frac{n+1}{2n} > \frac{1}{2}$$

as required. ○

Combining Theorems 1J and 1U, we have the following result.

THEOREM 1V. *We have $\phi(n) \gg n/\log n$ as $n \to \infty$.*

We now consider some average version of the Euler function.

THEOREM 1W. (MERTENS) *As $X \to \infty$, we have*

$$\sum_{n \le X} \phi(n) = \frac{3}{\pi^2} X^2 + O(X \log X).$$

Proof. As $X \to \infty$, we have, by Theorem 1R, that

$$
\begin{aligned}
\sum_{n\le X}\phi(n) &= \sum_{n\le X}\sum_{m\mid n}\mu(m)\frac{n}{m} = \sum_{m\le X}\mu(m)\sum_{\substack{n\le X\\ m\mid n}}\frac{n}{m} = \sum_{m\le X}\mu(m)\sum_{r\le X/m} r \\
&= \sum_{m\le X}\mu(m)\frac{1}{2}\left[\frac{X}{m}\right]\left(1+\left[\frac{X}{m}\right]\right) = \frac{1}{2}\sum_{m\le X}\mu(m)\left(\frac{X}{m}+O(1)\right)^2 \\
&= \frac{X^2}{2}\sum_{m\le X}\frac{\mu(m)}{m^2} + O\left(X\sum_{m\le X}\frac{1}{m}\right) + O\left(\sum_{m\le X}1\right) \\
&= \frac{X^2}{2}\sum_{m=1}^{\infty}\frac{\mu(m)}{m^2} + O\left(X^2\sum_{m>X}\frac{1}{m^2}\right) + O(X\log X) \\
&= \frac{X^2}{2}\sum_{m=1}^{\infty}\frac{\mu(m)}{m^2} + O(X\log X).
\end{aligned}
$$

It remains to show that

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

But

$$\left(\sum_{n=1}^{\infty}\frac{1}{n^2}\right)\left(\sum_{m=1}^{\infty}\frac{\mu(m)}{m^2}\right) = \sum_{k=1}^{\infty}\frac{1}{k^2}\left(\sum_{\substack{n,m\\ nm=k}}\mu(m)\right) = \sum_{k=1}^{\infty}\frac{1}{k^2}\left(\sum_{m\mid k}\mu(m)\right) = 1,$$

in view of Theorem 1M. ○

1.5. Dirichlet Convolution

We shall denote the class of all arithmetic functions by $\mathcal{A}$, and the class of all multiplicative functions by $\mathcal{M}$.

Given arithmetic functions $f, g \in \mathcal{A}$, we define the function $f * g : \mathbb{N} \to \mathbb{C}$ by writing

$$(f*g)(n) = \sum_{m\mid n} f(m)\, g\Big(\frac{n}{m}\Big)$$

for every $n \in \mathbb{N}$. This function is called the Dirichlet convolution of f and g.

It is not difficult to show that Dirichlet convolution of arithmetic functions is commutative and associative. In other words, for every $f, g, h \in \mathcal{A}$, we have

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

Furthermore, the arithmetic function $I : \mathbb{N} \to \mathbb{C}$, defined by $I(1) = 1$ and $I(n) = 0$ for every $n \in \mathbb{N}$ satisfying $n > 1$, is an identity element for Dirichlet convolution. It is easy to check that $I*f = f*I = f$ for every $f \in \mathcal{A}$.

On the other hand, an inverse may not exist under Dirichlet convolution. Consider, for example, the function $f \in \mathcal{A}$ satisfying $f(n) = 0$ for every $n \in \mathbb{N}$.

THEOREM 1X. *For any $f \in \mathcal{A}$, the following two statements are equivalent:*
(i) We have $f(1) \neq 0$.
(ii) There exists a unique $g \in \mathcal{A}$ such that $f * g = g * f = I$.

PROOF. Suppose that (ii) holds. Then $f(1)g(1) = 1$, so that $f(1) \neq 0$. Conversely, suppose that $f(1) \neq 0$. We shall define $g \in \mathcal{A}$ iteratively by writing

$$g(1) = \frac{1}{f(1)} \tag{5}$$

and

$$g(n) = -\frac{1}{f(1)} \sum_{\substack{d \mid n \\ d>1}} f(d)\, g\Big(\frac{n}{d}\Big) \tag{6}$$

for every $n \in \mathbb{N}$ satisfying $n > 1$. It is easy to check that this gives an inverse. Moreover, every inverse must satisfy (5) and (6), and so the inverse must be unique. ○

We now describe Theorem 1M and Möbius inversion in terms of Dirichlet convolution. Recall that the function $U \in \mathcal{A}$ is defined by $U(n) = 1$ for all $n \in \mathbb{N}$.

THEOREM 1Y.
(i) We have $\mu * U = I$.
(ii) If $f \in \mathcal{A}$ and $g = f * U$, then $f = g * \mu$.
(iii) If $g \in \mathcal{A}$ and $f = g * \mu$, then $g = f * U$.

PROOF. (i) follows from Theorem 1M. To prove (ii), note that

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

To prove (iii), note that

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

This completes the proof of Theorem 1Y. ○

We conclude this chapter by exhibiting some group structure within $\mathcal{A}$ and $\mathcal{M}$.

THEOREM 1Z. *The sets $\mathcal{A}' = \{f \in \mathcal{A} : f(1) \neq 0\}$ and $\mathcal{M}' = \{f \in \mathcal{M} : f(1) = 1\}$ form abelian groups under Dirichlet convolution.*

REMARK. Note that if $f \in \mathcal{M}$ is not identically zero, then $f(n) \neq 0$ for some $n \in \mathbb{N}$. Since $f(n) = f(1)f(n)$, we must have $f(1) = 1$.

PROOF OF THEOREM 1Z. For $\mathcal{A}'$, this is now trivial. We now consider $\mathcal{M}'$. Clearly $I \in \mathcal{M}'$. If $f, g \in \mathcal{M}'$ and $(m, n) = 1$, then

$$\begin{aligned}(f * g)(mn) &= \sum_{d \mid mn} f(d)\, g\Big(\frac{mn}{d}\Big) = \sum_{d_1 \mid m} \sum_{d_2 \mid n} f(d_1 d_2)\, g\left(\frac{mn}{d_1 d_2}\right) \\ &= \left(\sum_{d_1 \mid m} f(d_1)\, g\left(\frac{m}{d_1}\right)\right) \left(\sum_{d_2 \mid n} f(d_2)\, g\left(\frac{n}{d_2}\right)\right) = (f * g)(m)(f * g)(n),\end{aligned}$$

so that $f * g \in \mathcal{M}$. Since $(f * g)(1) = f(1)g(1) \neq 0$, we have $f * g \in \mathcal{M}'$. It remains to show that if $f \in \mathcal{M}'$, then f has an inverse in $\mathcal{M}'$. Clearly f has an inverse in $\mathcal{A}'$ under Dirichlet convolution. Let this inverse be h. We now define $g \in \mathcal{A}$ by writing $g(1) = 1$,

$$g(p^k) = h(p^k)$$

for every prime p and $k \in \mathbb{N}$, and

$$g(n) = \prod_{p^k \| n} g(p^k)$$

for every $n > 1$. Then $g \in \mathcal{M}'$. Furthermore, for every integer $n > 1$, we have

$$(f * g)(n) = \prod_{p^k \| n} (f * g)(p^k) = \prod_{p^k \| n} (f * h)(p^k) = \prod_{p^k \| n} I(p^k) = I(n),$$

so that g is an inverse of f. ○

Problems for Chapter 1

1. Prove that $d(n) \leq d(2^n - 1)$ for every $n \in \mathbb{N}$.

2. Suppose that $n \in \mathbb{N}$ is composite. Prove that $\sigma(n) > n + \sqrt{n}$.

3. Prove that $d(n)$ is odd if and only if $n \in \mathbb{N}$ is a square.

4. Prove that $\displaystyle\prod_{m \mid n} m = n^{\frac{1}{2}d(n)}$ for every $n \in \mathbb{N}$.

5. Suppose that $n \in \mathbb{N}$. Show that the number N of solutions of the equation $x^2 - y^2 = n$ in natural numbers x and y satisfies

$$2N = \begin{cases} d(n) - e_n & \text{if } n \text{ is an odd number,} \\ 0 & \text{if } n \text{ is twice an odd number,} \\ d(n/4) - e_n & \text{if } 4 \mid n, \end{cases}$$

   where $e_n = 1$ if n is a perfect square, and $e_n = 0$ otherwise.

6. Prove that there are no squarefree perfect numbers apart from 6.

7. Prove that $\displaystyle\sum_{m \mid n} \frac{1}{m} = 2$ for every perfect number $n \in \mathbb{N}$.

8. Prove that every odd perfect number must have at least two distinct prime factors, exactly one of which has odd exponent.

9. Suppose that $a \in \mathbb{N}$ satisfy $a > 1$. Let d run over all the divisors of a that have no more than m prime divisors. Prove that

$$\sum \mu(d) \begin{cases} \geq 0 & \text{if } m \text{ is even,} \\ \leq 0 & \text{if } m \text{ is odd.} \end{cases}$$

   [Hint: Write down first the canonical decomposition of a.]

10. Suppose that $k \in \mathbb{N}$ is even, and the canonical decomposition of $a \in \mathbb{N}$ is of the form $a = p_1 p_2 \ldots p_k$, where $p_1, p_2, \ldots, p_k$ are distinct primes. Let d run over all the divisors of a such that $0 < d < \sqrt{a}$. Prove that $\sum \mu(d) = 0$.

11. Prove that $\displaystyle\sum_{d^2 \mid n} \mu(d) = \mu^2(n)$ for every $n \in \mathbb{N}$.

    [HINT: Distinguish between the cases when n is squarefree and when n is not squarefree.]

12. By first showing that the function $f(n) = (-1)^{n-1}$ is multiplicative, evaluate the sum

$$h(n) = \sum_{m \mid n} (-1)^{m-1} \mu\Big(\frac{n}{m}\Big) \qquad \text{for every } n \in \mathbb{N}.$$

13. Explain why $\displaystyle\sum_{m \mid n} \mu(m)\, \sigma\Big(\frac{n}{m}\Big) = n$ for every $n \in \mathbb{N}$.

14. Prove that $\displaystyle\sum_{\substack{m=1 \\ (m,n)=1}}^{n} m = \frac{n\phi(n)}{2}$ for every $n \in \mathbb{N}$.

15. Suppose that $n \in \mathbb{N}$ satisfies $\phi(n) \mid n$. Prove that $n = 2^a 3^b$ for some non-negative integers a and b.

16. Suppose that $p_1, p_2, \ldots, p_k \in \mathbb{N}$ are distinct primes, and that there are no other primes.
    (i) Let $a = p_1 p_2 \ldots p_k$. Explain why we must have $\phi(a) = 1$.
    (ii) Obtain a contradiction.
    [REMARK: This is yet another proof that there are infinitely many primes.]

17. Prove that $\sigma(n) + \phi(n) = nd(n)$ if and only if $n \in \mathbb{N}$ is prime.

18. Suppose that $n = p_1^{u_1} \ldots p_r^{u_r}$, where $p_1 < \ldots < p_r$ are primes and $u_1, \ldots, u_r \in \mathbb{N}$.
    (i) Write

$$s(n) = \sum_{\substack{m=1 \\ (m,n)=1}}^{n} m^2.$$

    Prove that

$$n^2 \sum_{d \mid n} \frac{s(d)}{d^2} = \frac{n(n+1)(2n+1)}{6}.$$

    (ii) Apply the Möbius inversion formula to deduce that

$$\sum_{\substack{m=1 \\ (m,n)=1}}^{n} m^2 = \frac{1}{3}\phi(n)n^2 + \frac{1}{6}(-1)^r \phi(n) p_1 \ldots p_r.$$

19. For every $n \in \mathbb{N}$, let $Q(n)$ denote the number of squarefree numbers not exceeding n.
    (i) Prove that $\displaystyle n - Q(n) \le \frac{n}{4} + \sum_{m=1}^{\infty} \frac{n}{(2m+1)^2}$, and deduce that $Q(n) > n/2$.
    (ii) Hence show that every natural number is a sum of two squarefree numbers.

20. An arithmetic function $f : \mathbb{N} \to \mathbb{C}$ is said to be completely multiplicative if f is not identically zero and $f(mn) = f(m)f(n)$ for all $m, n \in \mathbb{N}$.
   (i) Show that the Möbius function μ is not completely multiplicative.
   (ii) Show that the Euler function ϕ is not completely multiplicative.
   (iii) Suppose that $f : \mathbb{N} \to \mathbb{C}$ is multiplicative. Show that f is completely multiplicative if and only if its Dirichlet inverse f^{-1} satisfies $f^{-1}(n) = \mu(n)f(n)$ for all $n \in \mathbb{N}$.
   (iv) Prove that the Liouville function $\lambda : \mathbb{N} \to \mathbb{C}$, defined by $\lambda(1) = 1$ and $\lambda(n) = (-1)^{u_1+\ldots+u_r}$ if $n = p_1^{u_1} \ldots p_r^{u_r}$, is completely multiplicative. Prove also that for every $n \in \mathbb{N}$,

$$\sum_{m|n} \lambda(m) = \begin{cases} 1 & \text{if } n \text{ is a square,} \\ 0 & \text{otherwise,} \end{cases}$$

   and $\lambda^{-1}(n) = |\mu(n)|$.

21. Suppose that $F : \mathbb{R}^+ \to \mathbb{C}$, where $\mathbb{R}^+$ denotes the set of all positive real numbers. For any real number $X \geq 1$, let

$$G(X) = \sum_{n \leq X} F\left(\frac{X}{n}\right).$$

Prove that

$$F(X) = \sum_{n \leq X} \mu(n)\, G\left(\frac{X}{n}\right) \qquad \text{for every real number } X \geq 1.$$

22. Suppose that $G : \mathbb{R}^+ \to \mathbb{C}$. For any real number $X \geq 1$, let

$$F(X) = \sum_{n \leq X} \mu(n)\, G\left(\frac{X}{n}\right).$$

Prove that

$$G(X) = \sum_{n \leq X} F\left(\frac{X}{n}\right) \qquad \text{for every real number } X \geq 1.$$

23. Prove that each of the following identities is valid for every real number $X \geq 1$:
   (i) $\displaystyle\sum_{n \leq X} \mu(n) \left[\frac{X}{n}\right] = 1.$
   (ii) $\displaystyle\sum_{n \leq X} \phi(n) = \frac{1}{2} \sum_{n \leq X} \mu(n) \left[\frac{X}{n}\right]^2 + \frac{1}{2}.$
   (iii) $\displaystyle\sum_{n \leq X} \frac{\phi(n)}{n} = \sum_{n \leq X} \frac{\mu(n)}{n} \left[\frac{X}{n}\right].$

24. Suppose that the function $F : \mathbb{R}^+ \to \mathbb{C}$ satisfies $F(X) = 0$ whenever $0 < X < 1$. For any arithmetic function α, we define the function $\alpha \circ F : \mathbb{R}^+ \to \mathbb{C}$ by writing

$$(\alpha \circ F)(X) = \sum_{n \leq X} \alpha(n)\, F\left(\frac{X}{n}\right) \qquad \text{for every } X \in \mathbb{R}^+.$$

   (i) Prove that for any arithmetic functions α and β, we have $\alpha \circ (\beta \circ F) = (\alpha * \beta) \circ F$.

(ii) Suppose that the arithmetic function α has inverse α^{-1} under Dirichlet convolution. Prove that if

$$G(X) = \sum_{n \le X} \alpha(n)\, F\left(\frac{X}{n}\right) \qquad \text{for every real number } X \in \mathbb{R}^+,$$

then

$$F(X) = \sum_{n \le X} \alpha^{-1}(n)\, G\left(\frac{X}{n}\right) \qquad \text{for every real number } X \in \mathbb{R}^+.$$

[Hint: Note that the identity function I under Dirichlet convolution satisfies $I \circ F = F$.]
[Remark: If α is completely multiplicative, then $\alpha^{-1}(n) = \mu(n)\alpha(n)$ for every $n \in \mathbb{N}$ by Problem 20(iii). Hence

$$G(X) = \sum_{n \le X} \alpha(n)\, F\left(\frac{X}{n}\right) \qquad \text{if and only if} \qquad F(X) = \sum_{n \le X} \mu(n)\alpha(n)\, G\left(\frac{X}{n}\right).$$

This is a generalization of Problems 21 and 22.]

25. For every $n \in \mathbb{N}$, let $f(n) = \displaystyle\sum_{m \mid n} \frac{\mu^2(m)}{\phi(m)}$.
   (i) Prove that $f(n) = n/\phi(n)$ for every $n \in \mathbb{N}$.
   (ii) Deduce that for every real number $X \ge 1$, we have $\displaystyle\sum_{n \le X} \frac{1}{\phi(n)} = \sum_{m \le X} \frac{\mu^2(m)}{m\phi(m)} \sum_{t \le X/m} \frac{1}{t}$.
   (iii) Show that the series $\displaystyle\sum_{m=1}^{\infty} \frac{\mu^2(m)}{m\phi(m)}$ and $\displaystyle\sum_{m=1}^{\infty} \frac{\mu^2(m)\log m}{m\phi(m)}$ both converge.
   (iv) Deduce that as $X \to \infty$, we have $\displaystyle\sum_{n \le X} \frac{1}{\phi(n)} \sim C \log X$, where $C = \displaystyle\sum_{m=1}^{\infty} \frac{\mu^2(m)}{m\phi(m)}$.

26. Consider a square lattice consisting of all points (a, b), where $a, b \in \mathbb{Z}$. Two lattice points P and Q are said to be mutually visible if the line segment which joins them contains no lattice points other than the endpoints P and Q.
   (i) Prove that (a, b) and $(0, 0)$ are mutually visible if and only if a and b are relatively prime.
   (ii) We shall prove that the set of lattice points visible from the origin has density $6/\pi^2$. Consider a large square region on the xy-plane defined by the inequalities $|x| \le r$ and $|y| \le r$. Let $N(r)$ denote the number of lattice points in this square, and let $N'(r)$ denote the number of these which are visible from the origin. The eight lattice points nearest the origin are all visible from the origin. By symmetry, $N'(r)$ is equal to 8 plus 8 times the number of visible points in the region $\{(x, y) : 2 \le x \le r \text{ and } 1 \le y \le x\}$. Prove that

$$N'(r) = 8 \sum_{n=1}^{r} \phi(n).$$

Obtain an asymptotic formula for $N(r)$, and show that

$$\frac{N'(r)}{N(r)} \to \frac{6}{\pi^2} \qquad \text{as } r \to \infty.$$

DISTRIBUTION OF PRIME NUMBERS

W W L CHEN

© W W L Chen, 1981, 2003.
This chapter originates from material used by the author at Imperial College, University of London, between 1981 and 1990.
It is available free to all individuals, on the understanding that it is not to be used for financial gains,
and may be downloaded and/or photocopied, with or without permission from the author.
However, this document may not be kept on any information storage and retrieval system without permission
from the author, unless such system is not accessible to any individuals other than its owners.

Chapter 2

ELEMENTARY PRIME NUMBER THEORY

2.1. Euclid's Theorem Revisited

We have already seen the elegant and simple proof of Euclid's theorem, that there are infinitely many primes. Here we shall begin by proving a slightly stronger result.

THEOREM 2A. *The series*

$$\sum_p \frac{1}{p}$$

is divergent.

Proof. For every real number $X \geq 2$, write

$$P_X = \prod_{p \leq X} \left(1 - \frac{1}{p}\right)^{-1}.$$

Then

$$\log P_X = -\sum_{p \leq X} \log\left(1 - \frac{1}{p}\right) = S_1 + S_2,$$

where

$$S_1 = \sum_{p \leq X} \frac{1}{p} \qquad \text{and} \qquad S_2 = \sum_{p \leq X} \sum_{h=2}^{\infty} \frac{1}{hp^h}.$$

Since

$$0 \le \sum_{h=2}^{\infty} \frac{1}{hp^h} \le \sum_{h=2}^{\infty} \frac{1}{p^h} = \frac{1}{p(p-1)},$$

we have

$$0 \le S_2 \le \sum_p \frac{1}{p(p-1)} \le \sum_{n=2}^{\infty} \frac{1}{n(n-1)} = 1,$$

so that $0 \le S_2 \le 1$. On the other hand, we have

$$P_X = \prod_{p \le X} \left(\sum_{h=0}^{\infty} \frac{1}{p^h} \right) \ge \sum_{n \le X} \frac{1}{n} \to \infty \quad \text{as } X \to \infty.$$

The result follows. ◯

For every real number $X \ge 2$, we write

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

so that $\pi(X)$ denotes the number of primes in the interval $[2, X]$. This function has been studied extensively by number theorists, and attempts to study it in depth have led to major developments in other important branches of mathematics.

As can be expected, many conjectures concerning the distribution of primes were made based purely on numerical evidence, including the celebrated Prime number theorem, proved in 1896 by Hadamard and de la Vallée Poussin, that

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

We shall prove this in Chapter 5, and give another proof in Chapter 6. Here we shall be concerned with the weaker result of Tchebycheff, that there exist positive absolute constants c_1 and c_2 such that for every real number $X \ge 2$, we have

$$c_1 \frac{X}{\log X} < \pi(X) < c_2 \frac{X}{\log X}.$$

2.2. The Von Mangoldt Function

The study of the function $\pi(X)$ usually involves, instead of the characteristic function of the primes, a function which counts not only primes, but prime powers as well, and with weights. Accordingly, we introduce the von Mangoldt function $\Lambda : \mathbb{N} \to \mathbb{C}$, defined for every $n \in \mathbb{N}$ by writing

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^r, \text{ with } p \text{ prime and } r \in \mathbb{N}, \\ 0 & \text{otherwise.} \end{cases}$$

THEOREM 2B. *For every $n \in \mathbb{N}$, we have*

$$\sum_{m \mid n} \Lambda(m) = \log n.$$

PROOF. The result is clearly true for $n = 1$, so it remains to consider the case $n \geq 2$. Suppose that $n = p_1^{u_1} \ldots p_r^{u_r}$ is the canonical decomposition of n. Then the only non-zero contribution to the sum on the left hand side comes from those natural numbers m of the form $m = p_j^{v_j}$ with $j = 1, \ldots, r$ and $1 \leq v_j \leq u_j$. It follows that

$$\sum_{m|n} \Lambda(m) = \sum_{j=1}^{r} \sum_{v_j=1}^{u_j} \log p_j = \sum_{j=1}^{r} \log p_j^{u_j} = \log n.$$

This completes the proof. ○

THEOREM 2C. *As $X \to \infty$, we have*

$$\sum_{m \leq X} \Lambda(m) \left[\frac{X}{m}\right] = X \log X - X + O(\log X).$$

PROOF. It follows from Theorem 2B that

$$\sum_{n \leq X} \log n = \sum_{n \leq X} \sum_{m|n} \Lambda(m) = \sum_{m \leq X} \Lambda(m) \sum_{\substack{n \leq X \\ m|n}} 1 = \sum_{m \leq X} \Lambda(m) \left[\frac{X}{m}\right].$$

It therefore suffices to prove that

$$\sum_{n \leq X} \log n = X \log X - X + O(\log X) \quad \text{as } X \to \infty. \tag{1}$$

To prove (1), note that $\log X$ is an increasing function of X. In particular, for every $n \in \mathbb{N}$, we have

$$\log n \leq \int_n^{n+1} \log u \, \mathrm{d}u,$$

so that

$$\sum_{n \leq X} \log n - \log(X+1) \leq \int_1^X \log u \, \mathrm{d}u.$$

On the other hand, for every $n \in \mathbb{N}$, we have

$$\log n \geq \int_{n-1}^{n} \log u \, \mathrm{d}u,$$

so that

$$\sum_{n \leq X} \log n = \sum_{2 \leq n \leq X} \log n \geq \int_1^{[X]} \log u \, \mathrm{d}u = \int_1^X \log u \, \mathrm{d}u - \int_{[X]}^X \log u \, \mathrm{d}u \geq \int_1^X \log u \, \mathrm{d}u - \log X.$$

The inequality (1) now follows on noting that

$$\int_1^X \log u \, \mathrm{d}u = X \log X - X + 1.$$

This completes the proof. ○

2.3. Tchebycheff's Theorem

The crucial step in the proof of Tchebycheff's theorem concerns obtaining bounds on sums involving the von Mangoldt function. More precisely, we prove the following result.

THEOREM 2D. *There exist positive absolute constants c_3 and c_4 such that*

$$\sum_{m\le X}\Lambda(m)\ge\frac{1}{2}X\log 2\quad\text{if } X\ge c_3, \tag{2}$$

and

$$\sum_{\frac{X}{2}<m\le X}\Lambda(m)\le c_4X\quad\text{if } X\ge 0. \tag{3}$$

PROOF. If $m\in\mathbb{N}$ satisfies $X/2<m\le X$, then clearly $[X/2m]=0$. It follows from this and Theorem 2C that as $X\to\infty$, we have

$$\begin{aligned}\sum_{m\le X}\Lambda(m)\left(\left[\frac{X}{m}\right]-2\left[\frac{X}{2m}\right]\right)&=\sum_{m\le X}\Lambda(m)\left[\frac{X}{m}\right]-2\sum_{m\le\frac{X}{2}}\Lambda(m)\left[\frac{X}{2m}\right]\\&=(X\log X-X+O(\log X))-2\left(\frac{X}{2}\log\frac{X}{2}-\frac{X}{2}+O(\log X)\right)=X\log 2+O(\log X).\end{aligned}$$

Hence there exists a positive absolute constant c_5 such that for all sufficiently large X, we have

$$\frac{1}{2}X\log 2<\sum_{m\le X}\Lambda(m)\left(\left[\frac{X}{m}\right]-2\left[\frac{X}{2m}\right]\right)<c_5X.$$

We now consider the function $[\alpha]-2[\alpha/2]$. Clearly $[\alpha]-2[\alpha/2]<\alpha-2(\alpha/2-1)=2$. Note that the left hand side is an integer, so we must have $[\alpha]-2[\alpha/2]\le 1$. It follows that for all sufficiently large X, we have

$$\frac{1}{2}X\log 2<\sum_{m\le X}\Lambda(m).$$

The inequality (2) follows. On the other hand, if $X/2<m\le X$, then $[X/m]=1$ and $[X/2m]=0$, so that for all sufficiently large X, we have

$$\sum_{\frac{X}{2}<m\le X}\Lambda(m)\le c_5X.$$

The inequality (3) follows easily. ○

We now state and prove Tchebycheff's theorem.

THEOREM 2E. (TCHEBYCHEFF) *There exist positive absolute constants c_1 and c_2 such that for every real number $X\ge 2$, we have*

$$c_1\frac{X}{\log X}<\pi(X)<c_2\frac{X}{\log X}.$$

PROOF. To prove the lower bound, note that

$$\sum_{m\le X}\Lambda(m)=\sum_{\substack{p,n\\p^n\le X}}\log p=\sum_{p\le X}(\log p)\sum_{1\le n\le\left[\frac{\log X}{\log p}\right]}1=\sum_{p\le X}(\log p)\left[\frac{\log X}{\log p}\right]\le\pi(X)\log X.$$

It follows from (2) that

$$\pi(X) \geq \frac{X \log 2}{2 \log X} \quad \text{if } X \geq c_3.$$

Since $\pi(2) = 1$, we get the lower bound for a suitable choice of c_1.

To prove the upper bound, note that in view of (3) and the definition of the von Mangoldt function, the inequality

$$\sum_{\frac{X}{2^{j+1}} < p \leq \frac{X}{2^j}} \log p \leq c_4 \frac{X}{2^j}$$

holds for every integer $j \geq 0$ and every real number $X \geq 0$. Suppose that $X \geq 2$. Let the integer $k \geq 0$ be defined such that $2^k < X^{1/2} \leq 2^{k+1}$. Then

$$\sum_{X^{1/2} < p \leq X} \log p \leq \sum_{j=0}^{k} \sum_{\frac{X}{2^{j+1}} < p \leq \frac{X}{2^j}} \log p \leq c_4 X \sum_{j=0}^{k} 2^{-j} < 2c_4 X,$$

so that

$$\sum_{X^{1/2} < p \leq X} 1 \leq \sum_{X^{1/2} < p \leq X} \frac{\log p}{\log X^{1/2}} < \frac{4c_4 X}{\log X},$$

whence

$$\pi(X) \leq X^{1/2} + \frac{4c_4 X}{\log X} < \frac{c_2 X}{\log X}$$

for a suitable c_2. ○

2.4. Some Results of Mertens

We conclude this chapter by obtaining an improvement of Theorem 2A.

THEOREM 2F. (MERTENS) *As $X \to \infty$, we have*

(4) $$\sum_{m \leq X} \frac{\Lambda(m)}{m} = \log X + O(1),$$

(5) $$\sum_{p \leq X} \frac{\log p}{p} = \log X + O(1),$$

and

(6) $$\sum_{p \leq X} \frac{1}{p} = \log \log X + O(1).$$

Proof. Recall Theorem 2C. As $X \to \infty$, we have

$$\sum_{m \leq X} \Lambda(m) \left[\frac{X}{m} \right] = X \log X - X + O(\log X).$$

Clearly $[X/m] = X/m + O(1)$, so that as $X \to \infty$, we have

$$\sum_{m\le X} \Lambda(m)\left[\frac{X}{m}\right] = X\sum_{m\le X}\frac{\Lambda(m)}{m} + O\left(\sum_{m\le X}\Lambda(m)\right).$$

It follows from (3) that

$$\sum_{m\le X}\Lambda(m) \le \sum_{j=0}^{\infty}\sum_{\frac{X}{2^{j+1}}<m\le\frac{X}{2^j}}\Lambda(m) \le 2c_4X,$$

so that as $X \to \infty$, we have

$$X\sum_{m\le X}\frac{\Lambda(m)}{m} = X\log X + O(X).$$

The inequality (4) follows. Next, note that

$$\sum_{m\le X}\frac{\Lambda(m)}{m} = \sum_{\substack{p,k\\ p^k\le X}}\frac{\log p}{p^k} = \sum_{p\le X}\frac{\log p}{p} + \sum_{p\le X}(\log p)\sum_{2\le k\le\frac{\log X}{\log p}}\frac{1}{p^k}.$$

As $X \to \infty$, we have

$$\sum_{p\le X}(\log p)\sum_{2\le k\le\frac{\log X}{\log p}}\frac{1}{p^k} \le \sum_{p\le X}(\log p)\sum_{k=2}^{\infty}\frac{1}{p^k} = \sum_{p\le X}\frac{\log p}{p(p-1)} \le \sum_{n=2}^{\infty}\frac{\log n}{n(n-1)} = O(1).$$

The inequality (5) follows. Finally, for every real number $X \ge 2$, let

$$T(X) = \sum_{p\le X}\frac{\log p}{p}.$$

Then it follows from (5) that there exists a positive absolute constant c_6 such that $|T(X) - \log X| < c_6$ whenever $X \ge 2$. On the other hand,

$$\begin{aligned}\sum_{p\le X}\frac{1}{p} &= \sum_{p\le X}\frac{\log p}{p}\left(\frac{1}{\log X} + \int_p^X\frac{\mathrm{d}y}{y\log^2 y}\right) = \frac{T(X)}{\log X} + \int_2^X\frac{T(y)\,\mathrm{d}y}{y\log^2 y}\\ &= \frac{T(X)-\log X}{\log X} + \int_2^X\frac{(T(y)-\log y)\,\mathrm{d}y}{y\log^2 y} + 1 + \int_2^X\frac{\mathrm{d}y}{y\log y}.\end{aligned}$$

It follows that as $X \to \infty$, we have

$$\left|\sum_{p\le X}\frac{1}{p} - \log\log X\right| < \frac{c_6}{\log X} + \int_2^X\frac{c_6\,\mathrm{d}y}{y\log^2 y} + 1 - \log\log 2 = O(1).$$

The inequality (6) follows. ◯

Problems for Chapter 2

1. Prove that $\Lambda(n) + \sum_{m\mid n}\mu(m)\log m = 0$ for every $n \in \mathbb{N}$.

2. For any arithmetic function f, we define f' to be the arithmetic function given by $f'(n) = f(n)\log n$ for every $n \in \mathbb{N}$. Then for the arithmetic function U defined by $U(n) = 1$ for every $n \in \mathbb{N}$, we have $U'(n) = \log n$ and $U''(n) = \log^2 n$ for every $n \in \mathbb{N}$.
   (i) Suppose that f and g are arithmetic functions.
      (I) Prove that $(f+g)' = f' + g'$ and $(f * g)' = (f' * g) + (f * g')$.
      (II) Suppose that $f(1) \neq 0$. By noting that $(f * f^{-1})'(n) = 0$ for every $n \in \mathbb{N}$, prove that $(f^{-1})' = -f' * (f * f)^{-1}$.
   (ii) Explain why $\Lambda * U = U'$. Then establish Selberg's identity $\Lambda' + (\Lambda * \Lambda) = U'' * \mu$.

3. Prove that for every real number $X \geq 2$, we have $\displaystyle\prod_{p \leq X}\left(1 - \frac{1}{p}\right)^{-1} > \log X$.

4. Use the well-known inequality
$$\frac{t}{1+t} < \log(1+t) < t, \qquad \text{where } t > -1 \text{ and } t \neq 0,$$
to show that
$$\sum_{p \leq X} \frac{1}{p-1} > \log\log X \qquad \text{and} \qquad \sum_{p \leq X} \frac{1}{p} > \log\log X - 1.$$

5. Suppose that
   - λ_n is an increasing sequence of real numbers with limit infinity;
   - c_n is an arbitrary sequence of real or complex numbers; and
   - f has continuous derivative for $X \geq \lambda_1$.

   For every $X \geq \lambda_1$, let
$$C(X) = \sum_{\lambda_n \leq X} c_n.$$
   Establish the partial summation formula, that for every $X \geq \lambda_1$, we have
$$\sum_{\lambda_n \leq X} c_n f(\lambda_n) = C(X)f(X) - \int_{\lambda_1}^{X} C(y)f'(y)\,\mathrm{d}y.$$

6. Use Theorem 2F and partial summation to show that as $X \to \infty$, we have
$$\int_2^X \frac{\pi(y)}{y^2}\,\mathrm{d}y = \sum_{p \leq X} \frac{1}{p} + o(1) \sim \log\log X.$$

7. Derive the Prime number theorem, that
$$\pi(X) \sim \frac{X}{\log X} \quad \text{as } X \to \infty,$$
   from the hypothetical relation
$$\sum_{p \leq X} \log p \sim X \quad \text{as } X \to \infty,$$
   and the information
$$\int_2^X \frac{\mathrm{d}y}{\log y} = \frac{X}{\log X} + o\left(\frac{X}{\log X}\right) \quad \text{as } X \to \infty.$$

8. Show that the series $\displaystyle\sum_{p \leq X} \frac{1}{p \log p}$ converges as $X \to \infty$.

DISTRIBUTION OF PRIME NUMBERS

W W L CHEN

© W W L Chen, 1990, 2003.
This chapter originates from material used by the author at Imperial College, University of London, between 1981 and 1990.
It is available free to all individuals, on the understanding that it is not to be used for financial gains,
and may be downloaded and/or photocopied, with or without permission from the author.
However, this document may not be kept on any information storage and retrieval system without permission
from the author, unless such system is not accessible to any individuals other than its owners.

Chapter 3

DIRICHLET SERIES

3.1. Convergence Properties

A Dirichlet series is a series of the type

$$F(s)=\sum_{n=1}^{\infty} f(n)n^{-s}, \tag{1}$$

where $f:\mathbb{N}\to\mathbb{C}$ is an arithmetic function and $s\in\mathbb{C}$. We usually write $s=\sigma+\mathrm{i}t$, where $\sigma,t\in\mathbb{R}$.

Our first task is to investigate the convergence properties of Dirichlet series.

THEOREM 3A. *Suppose that the series* (1) *converges for some* $s\in\mathbb{C}$. *Then there exist unique real numbers* $\sigma_0,\sigma_1,\sigma_2$ *satisfying* $-\infty\le\sigma_0\le\sigma_1\le\sigma_2<\infty$ *and such that the following statements hold:*

(i) The series (1) *converges for every* $s\in\mathbb{C}$ *with* $\sigma>\sigma_0$. *Furthermore, for every* $\epsilon>0$, *the series* (1) *diverges for some* $s\in\mathbb{C}$ *with* $\sigma_0-\epsilon<\sigma\le\sigma_0$.

(ii) For every $\eta>0$, *the series* (1) *converges uniformly on the set* $\{s\in\mathbb{C}:\sigma>\sigma_1+\eta\}$ *and does not converge uniformly on the set* $\{s\in\mathbb{C}:\sigma>\sigma_1-\eta\}$.

(iii) The series (1) *converges absolutely for every* $s\in\mathbb{C}$ *with* $\sigma>\sigma_2$. *Furthermore, for every* $\epsilon>0$, *the series* (1) *does not converge absolutely for some* $s\in\mathbb{C}$ *with* $\sigma_2-\epsilon<\sigma\le\sigma_2$.

Example. The Dirichlet series

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

converges absolutely for every $s\in\mathbb{C}$ with $\sigma>1$ and diverges for every real $s<1$. It follows that $\sigma_0=\sigma_1=\sigma_2=1$ in this case.

Proof of Theorem 3A. Suppose that the series (1) converges for $s = s^* = \sigma^* + it^*$. Then $f(n)n^{-s^*} \to 0$ as $n \to \infty$, so that $|f(n)n^{-s^*}| = O(1)$, and so $|f(n)| = O(n^{\sigma^*})$. It follows that for every $s \in \mathbb{C}$ with $\sigma > \sigma^* + 1$, we have

$$|f(n)n^{-s}| = |f(n)n^{-\sigma}| = O(n^{\sigma^* - \sigma}),$$

so that the series (1) converges by the Comparison test. Now let

$$\sigma_0 = \inf\{u \in \mathbb{R} : \text{the series (1) converges for all } s \in \mathbb{C} \text{ with } \sigma > u\}$$

and

$$\sigma_2 = \inf\{u \in \mathbb{R} : \text{the series (1) converges absolutely for all } s \in \mathbb{C} \text{ with } \sigma > u\}.$$

Clearly (i) and (iii) follow, and $\sigma_0 \le \sigma_2$. To prove (ii), let $\delta > 0$ and $\epsilon > 0$ be chosen. Then there exists $N \in \mathbb{N}$ such that

$$\sum_{n=N+1}^{\infty} |f(n)| n^{-\sigma_2 - \delta} < \epsilon.$$

Hence

$$\sup\left\{ \left| \sum_{n=1}^{N} f(n)n^{-s} - \sum_{n=1}^{\infty} f(n)n^{-s} \right| : \sigma \ge \sigma_2 + \delta \right\} \le \sum_{n=N+1}^{\infty} |f(n)| n^{-\sigma_2 - \delta} < \epsilon.$$

It follows that the series (1) converges uniformly on the set $\{s \in \mathbb{C} : \sigma \ge \sigma_2 + \delta\}$. Now let

$$\sigma_1 = \inf\{u \in \mathbb{R} : \text{the series (1) converges uniformly on } \{z \in \mathbb{C} : \sigma \ge u\}\}.$$

Clearly $\sigma_0 \le \sigma_1 \le \sigma_2 + \delta$. Since $\delta > 0$ is arbitrary, we must have $\sigma_0 \le \sigma_1 \le \sigma_2$. ◯

A simple consequence of uniform convergence is the following result concerning differentiation term by term.

THEOREM 3B. *For every $s \in \mathbb{C}$ with $\sigma > \sigma_1$, the series (1) may be differentiated term by term. In particular, $F'(s)$ exists and*

$$F'(s) = -\sum_{n=1}^{\infty} f(n)(\log n)n^{-s}.$$

3.2. Uniqueness Properties

Our next task is to prove the uniqueness theorem of Dirichlet series, a result of great importance in view of the applications we have in mind.

THEOREM 3C. *Suppose that*

$$F(s) = \sum_{n=1}^{\infty} f(n)n^{-s} \qquad \text{and} \qquad G(s) = \sum_{n=1}^{\infty} g(n)n^{-s},$$

where $f : \mathbb{N} \to \mathbb{C}$ and $g : \mathbb{N} \to \mathbb{C}$ are arithmetic functions and $s \in \mathbb{C}$. Suppose further that there exists $\sigma_3 \in \mathbb{R}$ such that for every $s \in \mathbb{C}$ satisfying $\sigma \ge \sigma_3$, we have $F(s) = G(s)$. Then $f(n) = g(n)$ for every $n \in \mathbb{N}$.

It is clearly sufficient to prove the following special case.

THEOREM 3D. *Suppose that*

$$F(s)=\sum_{n=1}^\infty f(n)n^{-s},$$

where $f:\mathbb{N}\to\mathbb{C}$ is an arithmetic function and $s\in\mathbb{C}$. Suppose further that there exists $\sigma_3\in\mathbb{R}$ such that for every $s\in\mathbb{C}$ satisfying $\sigma\ge\sigma_3$, we have $F(s)=0$. Then $f(n)=0$ for every $n\in\mathbb{N}$.

PROOF. Since the series converges for $s=\sigma_3$, we must have $|f(n)|=O(n^{\sigma_3})$ for all $n\in\mathbb{N}$. Now let $\sigma\ge\sigma_3+2$. Then

$$\sum_{n=N}^\infty f(n)n^{-\sigma}=O\left(\sum_{n=N}^\infty n^{\sigma_3-\sigma}\right). \tag{2}$$

Note next that $y^{\sigma_3-\sigma}$ is a decreasing function of y, so that

$$\sum_{n=N}^\infty n^{\sigma_3-\sigma}=N^{\sigma_3-\sigma}+\sum_{n=N+1}^\infty n^{\sigma_3-\sigma}\le N^{\sigma_3-\sigma}+\int_N^\infty y^{\sigma_3-\sigma}\,\mathrm{d}y=O(N^{\sigma_3-\sigma+1}). \tag{3}$$

Combining (2) and (3), we see that for every $N\in\mathbb{N}$, we have

$$\sum_{n=N}^\infty f(n)n^{-\sigma}=O(N^{\sigma_3-\sigma+1}). \tag{4}$$

Using (4) with $N=2$, we obtain, for $\sigma\ge\sigma_3+2$,

$$0=F(\sigma)=f(1)+\sum_{n=2}^\infty f(n)n^{-\sigma}=f(1)+O(2^{\sigma_3-\sigma+1})\to f(1)\quad\text{as } \sigma\to+\infty.$$

Hence $f(1)=0$. Suppose now that $f(1)=f(2)=\ldots=f(M-1)=0$. Using (4) with $N=M+1$, we obtain, for $\sigma\ge\sigma_3+2$,

$$0=F(\sigma)=f(M)M^{-\sigma}+\sum_{n=M+1}^\infty f(n)n^{-\sigma}=f(M)M^{-\sigma}+O((M+1)^{\sigma_3-\sigma+1}),$$

so that

$$0=f(M)+O\left((M+1)^{\sigma_3+1}\left(\frac{M}{M+1}\right)^\sigma\right)\to f(M)\quad\text{as } \sigma\to+\infty.$$

Hence $f(M)=0$. The result now follows from induction. ○

3.3. Multiplicative Properties

Dirichlet series are extremely useful in tackling problems in number theory as well as in other branches of mathematics. The main properties that underpin most of these applications are the multiplicative aspects of these series.

THEOREM 3E. *Suppose that for every $j=1,2,3$, we have*

$$F_j(s)=\sum_{n=1}^\infty f_j(n)n^{-s},$$

where $f_j : \mathbb{N} \to \mathbb{C}$ is an arithmetic function and $s \in \mathbb{C}$. Suppose further that for every $n \in \mathbb{N}$, we have

$$f_3(n) = \sum_{\substack{x,y \\ xy=n}} f_1(x)f_2(y) = \sum_{x|n} f_1(x)f_2\left(\frac{n}{x}\right) = \sum_{y|n} f_1\left(\frac{n}{y}\right) f_2(y).$$

Then

$$F_1(s)F_2(s) = F_3(s),$$

provided that $\sigma > \max\{\sigma_2^{(1)}, \sigma_2^{(2)}\}$, where, for every $j = 1, 2$, the series $F_j(s)$ converges absolutely for every $s \in \mathbb{C}$ with $\sigma > \sigma_2^{(j)}$.

Proof. We have

$$\sum_{n=1}^{N} f_3(n)n^{-s} = \sum_{\substack{1 \le x \le N \\ 1 \le y \le N \\ xy \le N}} f_1(x)x^{-s} f_2(y)y^{-s},$$

so that

$$\begin{aligned} &\sum_{n=1}^{N} f_3(n)n^{-s} - \sum_{x \le \sqrt{N}} f_1(x)x^{-s} \sum_{y \le \sqrt{N}} f_2(y)y^{-s} \\ &\quad = \sum_{\sqrt{N} < x \le N} f_1(x)x^{-s} \sum_{y \le N/x} f_2(y)y^{-s} + \sum_{x \le \sqrt{N}} f_1(x)x^{-s} \sum_{\sqrt{N} < y \le N/x} f_2(y)y^{-s}. \end{aligned}$$

It follows that

$$\begin{aligned} (5) \qquad &\left| \sum_{n=1}^{N} f_3(n)n^{-s} - \sum_{x \le \sqrt{N}} f_1(x)x^{-s} \sum_{y \le \sqrt{N}} f_2(y)y^{-s} \right| \\ &\quad < \left(\sum_{x > \sqrt{N}} |f_1(x)|x^{-\sigma} \right) \left(\sum_{y=1}^{\infty} |f_2(y)|y^{-\sigma} \right) + \left(\sum_{x=1}^{\infty} |f_1(x)|x^{-\sigma} \right) \left(\sum_{y > \sqrt{N}} |f_2(y)|y^{-\sigma} \right). \end{aligned}$$

Suppose now that $\sigma > \max\{\sigma_2^{(1)}, \sigma_2^{(2)}\}$. Clearly

$$\sum_{x > \sqrt{N}} |f_1(x)|x^{-\sigma} \qquad \text{and} \qquad \sum_{y > \sqrt{N}} |f_2(y)|y^{-\sigma}$$

converge to 0 as $N \to \infty$. Furthermore, the series

$$\sum_{x=1}^{\infty} |f_1(x)|x^{-\sigma} \qquad \text{and} \qquad \sum_{y=1}^{\infty} |f_2(y)|y^{-\sigma}$$

are convergent. It follows that the right hand side of (5) converges to 0 as $N \to \infty$. On the other hand,

$$\sum_{x \le \sqrt{N}} f_1(x)x^{-s} \qquad \text{and} \qquad \sum_{y \le \sqrt{N}} f_2(y)y^{-s}$$

converge to $F_1(s)$ and $F_2(s)$ respectively as $N \to \infty$. The result follows. ○

REMARK. Theorem 3E generalizes to a product of k Dirichlet series $F_1(s), \ldots, F_k(s)$, where the general coefficient is

$$\sum_{\substack{x_1,\ldots,x_k\\ x_1\ldots x_k=n}} f_1(x_1)\ldots f_k(x_k).$$

In many applications, the coefficients $f(n)$ of the Dirichlet series will be given by various important arithmetic functions in number theory. We therefore study next some consequences when the function $f : \mathbb{N} \to \mathbb{C}$ is multiplicative.

THEOREM 3F. *Suppose that the function $f : \mathbb{N} \to \mathbb{C}$ is multiplicative. Then for every $s \in \mathbb{C}$ satisfying $\sigma > \sigma_2$, the series* (1) *satisfies*

$$F(s) = \prod_p \left(\sum_{h=0}^{\infty} f(p^h)p^{-hs}\right).$$

PROOF. By the Remark, if p_j is the j-th prime in increasing order, then

$$\prod_{j=1}^{k}\left(\sum_{h=0}^{\infty} f(p_j^h)p_j^{-hs}\right) = \sum_{n=1}^{\infty}\left(\sum_{\substack{h_1,\ldots,h_k\\ p_1^{h_1}\ldots p_k^{h_k}=n}} f(p_1^{h_1})\ldots f(p_k^{h_k})\right) n^{-s}.$$

By the uniqueness of factorization, the inner sum on the right hand side contains at most one term. Hence

$$\prod_{j=1}^{k}\left(\sum_{h=0}^{\infty} f(p_j^h)p_j^{-hs}\right) = \sum_{n=1}^{\infty} \theta_k(n)f(n)n^{-s},$$

where

$$\theta_k(n) = \begin{cases} 1 & \text{if all the prime factors of } n \text{ are among } p_1, \ldots, p_k,\\ 0 & \text{otherwise.}\end{cases}$$

It follows that

$$\begin{aligned}\prod_{j=1}^{k}\left(\sum_{h=0}^{\infty} f(p_j^h)p_j^{-hs}\right) - \sum_{n=1}^{\infty} f(n)n^{-s} &= \sum_{n=1}^{\infty}(\theta_k(n)-1)f(n)n^{-s}\\ &= O\left(\sum_{n=k+1}^{\infty}|f(n)|n^{-\sigma}\right) \to 0 \quad \text{as } k\to\infty.\end{aligned}$$

The result follows. ○

An arithmetic function $f : \mathbb{N} \to \mathbb{C}$ is said to be totally multiplicative or strongly multiplicative if $f(mn) = f(m)f(n)$ for every $m, n \in \mathbb{N}$.

THEOREM 3G. *Suppose that the function $f : \mathbb{N} \to \mathbb{C}$ is totally multiplicative. Then for every $s \in \mathbb{C}$ satisfying $\sigma > \sigma_2$, the series* (1) *satisfies*

$$F(s) = \prod_p (1 - f(p)p^{-s})^{-1}.$$

PROOF. The absolute convergence of the series

$$\sum_{h=0}^{\infty} f(p^h)p^{-hs} \tag{6}$$

is immediate for $\sigma > \sigma_2$ by comparison with the series

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

Furthermore, if f is not identically zero, then it is easy to see that $f(1) = 1$, so that the series (6) is now a convergent geometric series with sum $(1 - f(p)p^{-s})^{-1}$. ○

EXAMPLE. Consider again the Dirichlet series

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

For every $s \in \mathbb{C}$ satisfying $\sigma > 1$, we have

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

This is called the Euler product of the Riemann zeta function $\zeta(s)$.

PROBLEMS FOR CHAPTER 3

1. Prove that for $\sigma > 1$, we have
   (i) $\displaystyle\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}$; (ii) $\displaystyle\zeta^2(s) = \sum_{n=1}^{\infty} \frac{d(n)}{n^s}$; (iii) $\displaystyle\frac{\zeta(s)}{\zeta(2s)} = \sum_{n=1}^{\infty} \frac{|\mu(n)|}{n^s}$;
   (iv) $\displaystyle\frac{\zeta^3(s)}{\zeta(2s)} = \sum_{n=1}^{\infty} \frac{d(n^2)}{n^s}$; (v) $\displaystyle\frac{\zeta^4(s)}{\zeta(2s)} = \sum_{n=1}^{\infty} \frac{d^2(n)}{n^s}$.

2. Prove that for $\sigma > 2$, we have

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

DISTRIBUTION OF PRIME NUMBERS

W W L CHEN

© W W L Chen, 1990, 2003.
This chapter originates from material used by the author at Imperial College, University of London, between 1981 and 1990.
It is available free to all individuals, on the understanding that it is not to be used for financial gains,
and may be downloaded and/or photocopied, with or without permission from the author.
However, this document may not be kept on any information storage and retrieval system without permission
from the author, unless such system is not accessible to any individuals other than its owners.

Chapter 4

PRIMES IN ARITHMETIC PROGRESSIONS

4.1. Dirichlet's Theorem

The purpose of this chapter is to prove the following remarkable result of Dirichlet, widely regarded as one of the greatest achievements in mathematics.

THEOREM 4A. (DIRICHLET) *Suppose that $q \in \mathbb{N}$ and $a \in \mathbb{Z}$ satisfy $(a, q) = 1$. Then there are infinitely many primes $p \equiv a \pmod{q}$.*

Note that the requirement $(a, q) = 1$ is crucial. If $n \equiv a \pmod{q}$, then clearly $(a, q) \mid n$. It follows that if $(a, q) > 1$, then the residue class $n \equiv a \pmod{q}$ of natural numbers contains at most one prime. In other words, Dirichlet's theorem asserts that any residue class $n \equiv a \pmod{q}$ of natural numbers must contain infinitely many primes if there is no simple reason to support the contrary.

It is easy to prove Theorem 4A by elementary methods for some special values of a and q.

EXAMPLES. (i) There are infinitely many primes $p \equiv -1 \pmod 4$. Suppose on the contrary that $p_1, \ldots, p_r$ represent all such primes. Then $4p_1 \ldots p_r - 1$ must have a prime factor $p \equiv -1 \pmod 4$. But p cannot be any of $p_1, \ldots, p_r$.

(ii) There are infinitely many primes $p \equiv 1 \pmod 4$. Suppose on the contrary that $p_1, \ldots, p_r$ represent all such primes. Consider the number $4(p_1 \ldots p_r)^2 + 1$. Suppose that a prime p divides $4(p_1 \ldots p_r)^2 + 1$. Then $4(p_1 \ldots p_r)^2 + 1 \equiv 0 \pmod p$. It follows that -1 is a quadratic residue modulo p, so that we must have $p \equiv 1 \pmod 4$. Clearly p cannot be any of $p_1, \ldots, p_r$.

4.2. A Special Case

The idea of Dirichlet is to show that if $(a, q) = 1$, then the series

$$\sum_{p \equiv a \pmod{q}} \frac{1}{p}$$

is divergent. For technical reasons, it is easier to show that if $(a, q) = 1$, then

$$\sum_{p \equiv a \pmod{q}} \frac{\log p}{p^\sigma} \to +\infty \quad \text{as } \sigma \to 1+.$$

Let us illustrate the idea of Dirichlet by studying the case $n \equiv 1 \pmod 4$.

First of all, we need a function that distinguishes between integers $n \equiv 1 \pmod 4$ and the others. Suppose that n is odd. Then it is easy to check that

$$\frac{1 + (-1)^{\frac{n-1}{2}}}{2} = \begin{cases} 1 & \text{if } n \equiv 1 \pmod 4, \\ 0 & \text{if } n \equiv -1 \pmod 4; \end{cases}$$

so that

$$\sum_{p \equiv 1 \pmod 4} \frac{\log p}{p^\sigma} = \frac{1}{2} \sum_{p \text{ odd}} \frac{\log p}{p^\sigma} \left(1 + (-1)^{\frac{p-1}{2}} \right).$$

Now the series

$$\sum_{p \text{ odd}} \frac{\log p}{p^\sigma} \to +\infty \quad \text{as } \sigma \to 1+,$$

so it suffices to show that the series

$$\sum_{p \text{ odd}} \frac{(-1)^{\frac{p-1}{2}} \log p}{p^\sigma}$$

converges as $\sigma \to 1+$.

The next idea is to show that if we consider the series

$$\sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{(-1)^{\frac{n-1}{2}} \Lambda(n)}{n^\sigma} \tag{1}$$

instead, then the contribution from the terms corresponding to non-prime odd natural numbers n is convergent. It therefore suffices to show that the series (1) converges as $\sigma \to 1+$.

Note now that the function

$$\chi(n) = \begin{cases} (-1)^{\frac{n-1}{2}} & \text{if } n \text{ is odd}, \\ 0 & \text{if } n \text{ is even}, \end{cases}$$

is totally multiplicative; in other words, $\chi(mn) = \chi(m)\chi(n)$ for every $m, n \in \mathbb{N}$. Write

$$L(\sigma) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^\sigma}, \tag{2}$$

and note that for every $n \in \mathbb{N}$, we have

$$\chi(n)\log n = \chi(n)\sum_{m|n}\Lambda(m) = \sum_{m|n}\chi(m)\Lambda(m)\chi\left(\frac{n}{m}\right).$$

It follows from Theorems 3B and 3E that for $\sigma > 1$, we have

$$L'(\sigma) = -\sum_{n=1}^{\infty}\frac{\chi(n)\log n}{n^\sigma} = -\left(\sum_{n=1}^{\infty}\frac{\chi(n)\Lambda(n)}{n^\sigma}\right)\left(\sum_{n=1}^{\infty}\frac{\chi(n)}{n^\sigma}\right).$$

Hence

$$\sum_{\substack{n=1\\ n\text{ odd}}}^{\infty}\frac{(-1)^{\frac{n-1}{2}}\Lambda(n)}{n^\sigma} = \sum_{n=1}^{\infty}\frac{\chi(n)\Lambda(n)}{n^\sigma} = -\frac{L'(\sigma)}{L(\sigma)}.$$

Now as $\sigma \to 1+$, we expect

$$L(\sigma) \to L(1) = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots > 0 \qquad \text{and} \qquad L'(\sigma) \to \frac{\log 3}{3} - \frac{\log 5}{5} + \frac{\log 7}{7} - \ldots$$

which converges by the Alternating series test. We therefore expect the series (1) to converge to a finite limit.

4.3. Dirichlet Characters

Dirichlet's most crucial discovery is that for every $q \in \mathbb{N}$, there is a family of $\phi(q)$ functions $\chi : \mathbb{N} \to \mathbb{C}$, known nowadays as the Dirichlet characters modulo q, which generalize the function χ in the special case and satisfy

$$\frac{1}{\phi(q)}\sum_{\chi \bmod q}\frac{\chi(n)}{\chi(a)} = \begin{cases} 1 & \text{if } n \equiv a \pmod q,\\ 0 & \text{if } n \not\equiv a \pmod q,\end{cases}$$

where the summation is over the $\phi(q)$ distinct Dirichlet characters modulo q.

To understand Dirichlet's ideas, we shall first of all study group characters. Our treatment here is slightly more general than is necessary, but easier to understand.

Let G be a finite abelian group of order h and with identity element e. A character on G is a non-zero complex-valued function χ on G for which $\chi(uv) = \chi(u)\chi(v)$ for every $u, v \in G$. It is easy to check the following simple results.

Remark. We have

(i) $\chi(e) = 1$;

(ii) for every $u \in G$, $\chi(u)$ is an h-th root of unity;

(iii) the number c of characters is finite; and

(iv) the characters form an abelian group.

Slightly less trivial is the following.

Remark. If $u \in G$ and $u \neq e$, then there exists a character χ on G such that $\chi(u) \neq 1$. To see this, note that G can be expressed as a direct product of cyclic groups $G_1, \ldots, G_s$ of orders $h_1, \ldots, h_s$ respectively, where $h = h_1 \ldots h_s$. Suppose that for each $j = 1, \ldots, s$, the cyclic group G_j is generated by v_j. Then we can write $u = v_1^{y_1} \ldots v_s^{y_s}$, where $y_j \pmod{h_j}$ is uniquely determined for every $j = 1, \ldots, s$. Since $u \neq e$, there exists $k = 1, \ldots, s$ such that $y_k \not\equiv 0 \pmod{h_k}$. Let $\chi(v_k) = e(1/h_k)$, and let $\chi(v_j) = 1$ for every $j = 1, \ldots, s$ such that $j \neq k$. Clearly $\chi(u) = e(y_k/h_k) \neq 1$.

We shall denote by χ_0 the principal character on G. In other words, $\chi_0(u) = 1$ for every $u \in G$. Also, $\sum_\chi$ denotes a summation over all the distinct characters on G.

THEOREM 4B. *Suppose that G is a finite abelian group of order h and with identity element e. Suppose further that χ_0 is the principal character on G.*

(i) For every character χ on G, we have

$$\sum_{u \in G} \chi(u) = \begin{cases} h & \text{if } \chi = \chi_0, \\ 0 & \text{if } \chi \neq \chi_0. \end{cases}$$

(ii) For every $u \in G$, we have

$$\sum_{\chi} \chi(u) = \begin{cases} c & \text{if } u = e, \\ 0 & \text{if } u \neq e, \end{cases}$$

where c denotes the number of distinct characters on G.

(iii) We have $c = h$.

(iv) For every $u, v \in G$, we have

$$\frac{1}{h} \sum_{\chi} \frac{\chi(u)}{\chi(v)} = \begin{cases} 1 & \text{if } u = v, \\ 0 & \text{if } u \neq v. \end{cases}$$

Proof. (i) If $\chi = \chi_0$, then the result is obvious. If $\chi \neq \chi_0$, then there exists $v \in G$ such that $\chi(v) \neq 1$, and so

$$\chi(v) \sum_{u \in G} \chi(u) = \sum_{u \in G} \chi(u)\chi(v) = \sum_{u \in G} \chi(uv) = \sum_{u \in G} \chi(u),$$

the last equality following from the fact that uv runs over all the elements of G as u runs over all the elements of G. Hence

$$(1 - \chi(v)) \sum_{u \in G} \chi(u) = 0.$$

The result follows since $\chi(v) \neq 1$.

(ii) If $u = e$, then the result is obvious. If $u \neq e$, then we have already shown that there exists a character χ_1 such that $\chi_1(u) \neq 1$, and so

$$\chi_1(u) \sum_{\chi} \chi(u) = \sum_{\chi} \chi_1(u)\chi(u) = \sum_{\chi} (\chi_1\chi)(u) = \sum_{\chi} \chi(u),$$

the last equality following from noting that the characters on G form an abelian group so that $\chi_1\chi$ runs through all the characters on G as χ runs through all the characters on G. Hence

$$(1 - \chi_1(u)) \sum_{\chi} \chi(u) = 0.$$

The result follows since $\chi_1(u) \neq 1$.

(iii) Note that

$$h = \sum_{\chi} \sum_{u \in G} \chi(u) = \sum_{u \in G} \sum_{\chi} \chi(u) = c.$$

(iv) Note that

$$\frac{1}{h} \sum_{\chi} \frac{\chi(u)}{\chi(v)} = \frac{1}{h} \sum_{\chi} \chi(u)\chi(v^{-1}) = \frac{1}{h} \sum_{\chi} \chi(uv^{-1}) = \begin{cases} c/h & \text{if } uv^{-1} = e, \\ 0 & \text{if } uv^{-1} \neq e. \end{cases}$$

The result follows since $h = c$. ○

We are now in a position to introduce Dirichlet characters. Let $q \in \mathbb{N}$ be given. Then there are exactly $\phi(q)$ residue classes $n \equiv a \pmod{q}$ satisfying $(a, q) = 1$. Under multiplication of residue classes, they form an abelian group of order $\phi(q)$. Suppose that these residue classes are represented by $a_1, \ldots, a_{\phi(q)}$ modulo q. Let $G = \{a_1, \ldots, a_{\phi(q)}\}$. We can now define a character χ on the group G as described earlier, interpreting the group elements as residue classes. Furthermore, we can extend the definition to cover the remaining residue classes. Precisely, for every $n \in \mathbb{N}$, let

$$\chi(n) = \begin{cases} \chi(a_j) & \text{if } n \equiv a_j \pmod{q} \text{ for some } j = 1, \ldots, \phi(q), \\ 0 & \text{if } (n, q) > 1. \end{cases} \tag{3}$$

A function $\chi : \mathbb{N} \to \mathbb{C}$ of the form (3) is called a Dirichlet character modulo q. Note that χ is totally multiplicative. Also, clearly there are exactly $\phi(q)$ Dirichlet characters modulo q. Furthermore, the principal Dirichlet character χ_0 modulo q is defined by

$$\chi_0(n) = \begin{cases} 1 & \text{if } (n, q) = 1, \\ 0 & \text{if } (n, q) > 1. \end{cases}$$

The following theorem follows immediately from these observations and Theorem 4B.

THEOREM 4C. *Suppose that $q \in \mathbb{N}$. Suppose further that χ_0 is the principal Dirichlet character modulo q.*

(i) For every Dirichlet character χ modulo q, we have

$$\sum_{n \pmod{q}} \chi(n) = \begin{cases} \phi(q) & \text{if } \chi = \chi_0, \\ 0 & \text{if } \chi \neq \chi_0. \end{cases}$$

(ii) For every $n \in \mathbb{N}$, we have

$$\sum_{\chi \pmod{q}} \chi(n) = \begin{cases} \phi(q) & \text{if } n \equiv 1 \pmod{q}, \\ 0 & \text{if } n \not\equiv 1 \pmod{q}. \end{cases}$$

(iii) For every $a \in \mathbb{Z}$ satisfying $(a, q) = 1$ and for every $n \in \mathbb{N}$, we have

$$\frac{1}{\phi(q)} \sum_{\chi \pmod{q}} \frac{\chi(n)}{\chi(a)} = \begin{cases} 1 & \text{if } n \equiv a \pmod{q}, \\ 0 & \text{if } n \not\equiv a \pmod{q}. \end{cases}$$

4.4. Some Dirichlet Series

Our next task is to introduce the functions analogous to the function (2) earlier. Let $s = \sigma + it \in \mathbb{C}$, where $\sigma, t \in \mathbb{R}$. For $\sigma > 1$, let

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s}; \tag{4}$$

furthermore, for any Dirichlet character χ modulo q, let

$$L(s,\chi) = \sum_{n=1}^{\infty} \chi(n) n^{-s}. \tag{5}$$

The functions (4) and (5) are called the Riemann zeta function and Dirichlet L-functions respectively. Note that the series are Dirichlet series and converge absolutely for $\sigma > 1$ and uniformly for $\sigma > 1 + \delta$ for any $\delta > 0$. Furthermore, the coefficients are totally multiplicative. It follows from Theorem 3G that for $\sigma > 1$, the series (4) and (5) have the Euler product representations

$$\zeta(s) = \prod_p (1 - p^{-s})^{-1} \qquad \text{and} \qquad L(s,\chi) = \prod_p (1 - \chi(p) p^{-s})^{-1}$$

respectively. The following are some simple properties of these functions.

THEOREM 4D. *Suppose that $\sigma > 1$. Then $\zeta(s) \neq 0$. Furthermore, $L(s,\chi) \neq 0$ for any Dirichlet character χ modulo q.*

PROOF. Since $\sigma > 1$, we have

$$|\zeta(s)| = \left| \prod_p (1 - p^{-s})^{-1} \right| \geq \prod_p (1 + p^{-\sigma})^{-1} = \prod_p \frac{1 - p^{-\sigma}}{1 - p^{-2\sigma}} = \frac{\zeta(2\sigma)}{\zeta(\sigma)} > 0$$

and

$$|L(s,\chi)| = \left| \prod_p (1 - \chi(p) p^{-s})^{-1} \right| \geq \prod_{p \nmid q} (1 + p^{-\sigma})^{-1} \geq \prod_p (1 + p^{-\sigma})^{-1} > 0.$$

This completes the proof. ○

THEOREM 4E. *Suppose that χ_0 is the principal Dirichlet character modulo q. Then for $\sigma > 1$, we have*

$$L(s,\chi_0) = \zeta(s) \prod_{p \mid q} (1 - p^{-s}).$$

PROOF. Since $\sigma > 1$, we have

$$L(s,\chi_0) = \prod_p (1 - \chi_0(p) p^{-s})^{-1} = \prod_{p \nmid q} (1 - p^{-s})^{-1} = \frac{\displaystyle\prod_p (1 - p^{-s})^{-1}}{\displaystyle\prod_{p \mid q} (1 - p^{-s})^{-1}} = \zeta(s) \prod_{p \mid q} (1 - p^{-s}).$$

This completes the proof. ○

THEOREM 4F. *Suppose that* $\sigma > 1$. *Then*

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

Furthermore, for every Dirichlet character χ *modulo* q, *we have*

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

PROOF. Since $\sigma > 1$, it follows from Theorem 3B that

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

It now follows from Theorem 3E and

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

that

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

The first assertion follows. On the other hand, it also follows from Theorem 3B that

$$-L'(s,\chi) = \sum_{n=1}^{\infty} \chi(n)(\log n) n^{-s}.$$

It now follows from Theorem 3E and

$$\chi(n)\log n = \sum_{m|n} \chi(m)\Lambda(m)\chi\left(\frac{n}{m}\right)$$

that

$$-L'(s,\chi) = \left(\sum_{n=1}^{\infty} \chi(n)\Lambda(n) n^{-s}\right)\left(\sum_{n=1}^{\infty} \chi(n) n^{-s}\right).$$

The second assertion follows. ○

THEOREM 4G. *If* $\sigma > 1$, *then for every Dirichlet character* χ *modulo* q, *we have*

$$\log L(s,\chi) = \sum_{p} \sum_{m=1}^{\infty} m^{-1} \chi(p^m) p^{-ms}.$$

PROOF. Taking logarithms on the Euler product representation, we have

(6) $$\log L(s,\chi) = \log \prod_{p} (1 - \chi(p)p^{-s})^{-1} = \sum_{p} \log(1 - \chi(p)p^{-s})^{-1},$$

so that

$$-\log L(s,\chi) = \sum_p \log(1-\chi(p)p^{-s}).$$

The justification for (6) is that the series on the right hand side converges uniformly for $\sigma > 1+\delta$, as can be deduced from the Weierstrass M-test on noting that

$$|\log(1-\chi(p)p^{-s})| \le 2|\chi(p)p^{-s}| \le 2p^{-1-\delta}.$$

The proof is now completed by expanding $\log(1-\chi(p)p^{-s})$. ◯

4.5. Analytic Continuation

Our next task is to extend the definition of $\zeta(s)$ and $L(s,\chi)$ to the half plane $\sigma > 0$. This is achieved by analytic continuation.

An example of analytic continuation is the following: Consider the geometric series

$$f(s) = \sum_{n=0}^{\infty} s^n.$$

This series converges absolutely in the set $\{s \in \mathbb{C} : |s| < 1\}$ and uniformly in the set $\{s \in \mathbb{C} : |s| < 1-\delta\}$ for any $\delta > 0$ to the sum $1/(1-s)$. Now let

$$g(s) = \frac{1}{1-s}$$

in $\mathbb{C}$. Then g is analytic in the set $\mathbb{C} \setminus \{1\}$, $g(s) = f(s)$ in the set $\{s \in \mathbb{C} : |s| < 1\}$, and g has a pole at $s = 1$. So g can be viewed as an analytic continuation of f to $\mathbb{C}$ with a pole at $s = 1$.

Returning to the functions $\zeta(s)$ and $L(s,\chi)$, we shall establish the following results on analytic continuation.

THEOREM 4H. *The function $\zeta(s)$ admits an analytic continuation to the half plane $\sigma > 0$, and is analytic in this half plane except for a simple pole at $s = 1$ with residue* 1.

THEOREM 4J. *Suppose that $q \in \mathbb{N}$ and χ_0 is the principal Dirichlet character modulo q. Then the function $L(s,\chi_0)$ admits an analytic continuation to the half plane $\sigma > 0$, and is analytic in this half plane except for a simple pole at $s = 1$ with residue $\phi(q)/q$.*

THEOREM 4K. *Suppose that $q \in \mathbb{N}$ and χ is a non-principal Dirichlet character modulo q. Then the function $L(s,\chi)$ admits an analytic continuation to the half plane $\sigma > 0$, and is analytic in this half plane.*

The proofs of these three theorems depend on the following two simple technical results. The first of these is basically a result on partial summation.

THEOREM 4L. *Suppose that* $a(n) = O(1)$ *for every* $n \in \mathbb{N}$. *For every* $x > 0$, *write*

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

Suppose further that for $\sigma > 1$, *we have*

$$F(s) = \sum_{n=1}^{\infty} a(n) n^{-s}.$$

Then for every $X > 0$ *and* $\sigma > 1$, *we have*

$$\sum_{n \le X} a(n) n^{-s} = S(X) X^{-s} + s \int_1^X S(x) x^{-s-1} \, \mathrm{d}x. \tag{7}$$

Furthermore, for $\sigma > 1$, *we have*

$$F(s) = s \int_1^{\infty} S(x) x^{-s-1} \, \mathrm{d}x. \tag{8}$$

Proof. To prove (7), simply note that

$$\begin{aligned}
\sum_{n \le X} a(n) n^{-s} - S(X) X^{-s} &= \sum_{n \le X} a(n)(n^{-s} - X^{-s}) = \sum_{n \le X} a(n) \int_n^X s x^{-s-1} \, \mathrm{d}x \\
&= s \int_1^X \left(\sum_{n \le x} a(n) \right) x^{-s-1} \, \mathrm{d}x = s \int_1^X S(x) x^{-s-1} \, \mathrm{d}x.
\end{aligned}$$

Also, (8) follows from (7) on letting $X \to \infty$. ◯

The second technical result, standard in complex function theory, will be stated without proof.

THEOREM 4M. *Suppose that the path* Γ *is defined by* $w(t) = u(t) + \mathrm{i}v(t)$, *where* $u(t), v(t) \in \mathbb{R}$ *for every* $t \in [0,1]$. *Suppose further that* $u'(t)$ *and* $v'(t)$ *are continuous on* $[0,1]$. *Let* D *be a domain in* $\mathbb{C}$. *For every* $s \in D$, *let*

$$F(s) = \int_{\Gamma} f(s, w) \, \mathrm{d}w,$$

where

- $f(s, w)$ *is continuous for every* $s \in D$ *and every* $w \in \Gamma$; *and*
- *for every* $w \in \Gamma$, *the function* $f(s, w)$ *is analytic in* D.

Then $F(s)$ *is analytic in* D.

Proof of Theorem 4H. Let $F(s) = \zeta(s)$. In the notation of Theorem 4L, we have $a(n) = 1$ for every $n \in \mathbb{N}$, so that $S(x) = [x]$ for every $x > 0$. It follows from (8) that

$$\zeta(s) = s \int_1^{\infty} [x] x^{-s-1} \, \mathrm{d}x = s \int_1^{\infty} x^{-s} \, \mathrm{d}x - s \int_1^{\infty} \{x\} x^{-s-1} \, \mathrm{d}x = 1 + \frac{1}{s-1} - s \int_1^{\infty} \{x\} x^{-s-1} \, \mathrm{d}x.$$

We shall show that the last term on the right hand side represents an analytic function for $\sigma > 0$. We can write

$$\int_1^{\infty} \{x\} x^{-s-1} \, \mathrm{d}x = \sum_{n=1}^{\infty} F_n(s),$$

where for every $n \in \mathbb{N}$,

$$F_n(s) = \int_n^{n+1} \{x\} x^{-s-1} \, \mathrm{d}x.$$

It remains to show that (i) for every $n \in \mathbb{N}$, the function $F_n(s)$ is analytic in $\mathbb{C}$; and (ii) for every $\delta > 0$, the series $\sum_{n=1}^{\infty} F_n(s)$ converges uniformly for $\sigma > \delta$. To show (i), note that by a change of variable,

$$F_n(s) = \int_0^1 t(n+t)^{-s-1} \, \mathrm{d}t = \int_0^1 t \mathrm{e}^{-(s+1)\log(n+t)} \, \mathrm{d}t,$$

and (i) follows from Theorem 4M. To show (ii), note that for $\sigma > \delta$, we have

$$|F_n(s)| = \left| \int_n^{n+1} \{x\} x^{-s-1} \, \mathrm{d}x \right| \le n^{-\sigma-1} < n^{-1-\delta},$$

and (ii) follows from the Weierstrass M-test. ○

Proof of Theorem 4J. Suppose that $\sigma > 1$. Recall Theorem 4E, that

$$L(s, \chi_0) = \zeta(s) \prod_{p \mid q} \left(1 - \frac{1}{p^s}\right).$$

Clearly the right hand side is analytic for $\sigma > 0$ except for a simple pole at $s = 1$. Furthermore, at $s = 1$, the function $\zeta(s)$ has a simple pole with residue 1, while

$$\prod_{p \mid q} \left(1 - \frac{1}{p}\right) = \frac{\phi(q)}{q}.$$

The result follows. ○

The proof of Theorem 4K is left as an exercise.

4.6. Proof of Dirichlet's Theorem

We now attempt to prove Theorem 4A. The result below will enable us to consider the analogue of (1).

THEOREM 4N. *Suppose that $\sigma > 1$. Then*

$$\sum_{p \equiv a \pmod q} \frac{\log p}{p^\sigma} = \sum_{n \equiv a \pmod q} \frac{\Lambda(n)}{n^\sigma} + O(1).$$

Proof. Note first of all that the sum on the left hand side does not exceed the first term on the right hand side. On the other hand, we have

$$\sum_{n \equiv a \pmod q} \frac{\Lambda(n)}{n^\sigma} - \sum_{p \equiv a \pmod q} \frac{\log p}{p^\sigma} \le \sum_p \sum_{m=2}^{\infty} \frac{\log p}{p^{m\sigma}}$$
$$\le \sum_p \sum_{m=2}^{\infty} \frac{\log p}{p^m} = \sum_p \frac{\log p}{p(p-1)} \le \sum_{n=2}^{\infty} \frac{\log n}{n(n-1)} = O(1).$$

The result follows. ○

Combining Theorems 4N, 4C and 4F, we have

$$(9)\qquad \sum_{p\equiv a \pmod q} \frac{\log p}{p^\sigma} = \sum_{n\equiv a \pmod q} \frac{\Lambda(n)}{n^\sigma} + O(1) = \sum_{n=1}^\infty \left(\frac{1}{\phi(q)} \sum_{\chi \pmod q} \frac{\chi(n)}{\chi(a)} \right) \frac{\Lambda(n)}{n^\sigma} + O(1)$$
$$= \frac{1}{\phi(q)} \sum_{\chi \pmod q} \frac{1}{\chi(a)} \sum_{n=1}^\infty \frac{\chi(n)\Lambda(n)}{n^\sigma} + O(1) = -\frac{1}{\phi(q)} \sum_{\chi \pmod q} \frac{1}{\chi(a)} \frac{L'(\sigma,\chi)}{L(\sigma,\chi)} + O(1).$$

Suppose now that

$$(10)\qquad \sum_{\substack{\chi \pmod q \\ \chi\neq\chi_0}} \frac{1}{\chi(a)} \frac{L'(\sigma,\chi)}{L(\sigma,\chi)} = O(1) \quad \text{as } \sigma \to 1+.$$

Then combining (9) and (10), we have

$$\sum_{p\equiv a \pmod q} \frac{\log p}{p^\sigma} = -\frac{1}{\phi(q)} \frac{L'(\sigma,\chi_0)}{L(\sigma,\chi_0)} + O(1) = \frac{1}{\phi(q)} \frac{1}{\sigma-1} + O(1) \to \infty \quad \text{as } \sigma \to 1+.$$

since the function $L'(s,\chi_0)/L(s,\chi_0)$ has a simple pole at $s=1$ with residue -1 by Theorem 4J. To complete the proof of Dirichlet's theorem, it remains to prove (10). Clearly (10) will follow if we can show that for every non-principal Dirichlet character $\chi \pmod q$, we have $L(1,\chi)\neq 0$. Here we need to distinguish two cases, represented by the next two theorems.

THEOREM 4P. *Suppose that $q\in\mathbb{N}$ and χ is a non-real Dirichlet character modulo q. Then we have $L(1,\chi)\neq 0$.*

PROOF. For $\sigma>1$, we have, in view of Theorem 4G, that

$$\sum_{\chi \pmod q} \log L(\sigma,\chi) = \sum_{\chi \pmod q} \sum_p \sum_{m=1}^\infty \chi(p^m) m^{-1} p^{-m\sigma}$$
$$= \sum_p \sum_{m=1}^\infty \left(\sum_{\chi \pmod q} \chi(p^m) \right) m^{-1}p^{-m\sigma} = \phi(q) \sum_p \sum_{\substack{m=1 \\ p^m\equiv 1 \pmod q}}^\infty m^{-1}p^{-m\sigma} > 0,$$

where the change of order of summation is justified since

$$\sum_{\chi \pmod q} \sum_p \sum_{m=1}^\infty |\chi(p^m) m^{-1} p^{-m\sigma}|$$

is finite. It follows that

$$(11)\qquad \left| \prod_{\chi \pmod q} L(\sigma,\chi) \right| > 1.$$

Suppose that χ_1 is a non-real Dirichlet character modulo q, and $L(1,\chi_1)=0$. Then $\chi_1\neq\overline{\chi_1}$, and $L(1,\overline{\chi_1}) = \overline{L(1,\chi_1)} = 0$ also. It follows that these two zeros more than cancel the simple pole of $L(\sigma,\chi_0)$ at $\sigma=1$, so that the product on the left hand side of (11) has a zero at $\sigma=1$. This gives a contradiction. ○

Clearly this approach does not work when χ is real.

THEOREM 4Q. *Suppose that $q \in \mathbb{N}$ and χ is a real, non-principal Dirichlet character modulo q. Then we have $L(1,\chi) \neq 0$.*

PROOF. Suppose that the result is false, so that there exists a real Dirichlet character χ modulo q such that $L(1,\chi) = 0$. Then the function $F(s) = \zeta(s)L(s,\chi)$ is analytic for $\sigma > 0$. Note that for $\sigma > 1$, we have

$$F(s) = \sum_{n=1}^{\infty} f(n)n^{-s},$$

where for every $n \in \mathbb{N}$,

$$f(n) = \sum_{m|n} \chi(m).$$

Let the function $g : \mathbb{N} \to \mathbb{R}$ be defined by

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

We shall first of all show that for every $n \in \mathbb{N}$, we have

$$f(n) \geq g(n). \tag{12}$$

Since χ is totally multiplicative, it suffices to prove (12) when $n = p^k$, where p is a prime and $k \in \mathbb{N}$. Indeed, since χ assumes only the values ± 1 and 0, we have

$$f(p^k) = 1 + \chi(p) + (\chi(p))^2 + \ldots + (\chi(p))^k = \begin{cases} 1 & \text{if } \chi(p) = 0, \\ k+1 & \text{if } \chi(p) = 1, \\ 1 & \text{if } \chi(p) = -1 \text{ and } k \text{ is even,} \\ 0 & \text{if } \chi(p) = -1 \text{ and } k \text{ is odd,} \end{cases}$$

so that

$$f(p^k) \geq g(p^k) = \begin{cases} 1 & \text{if } k \text{ is even,} \\ 0 & \text{if } k \text{ is odd.} \end{cases}$$

Suppose now that $0 < r < 3/2$. Since $F(s)$ is analytic for $\sigma > 0$, we must have the Taylor expansion

$$F(2-r) = \sum_{\nu=0}^{\infty} \frac{F^{(\nu)}(2)}{\nu!}(-r)^{\nu}.$$

Now by Theorem 3B, we have

$$F^{(\nu)}(2) = \sum_{n=1}^{\infty} f(n)(-\log n)^{\nu} n^{-2}.$$

It follows that for every $\nu \in \mathbb{N} \cup \{0\}$, we have, in view of (12),

$$\begin{aligned} \frac{F^{(\nu)}(2)}{\nu!}(-r)^{\nu} &= \frac{r^{\nu}}{\nu!} \sum_{n=1}^{\infty} f(n)(\log n)^{\nu} n^{-2} \geq \frac{r^{\nu}}{\nu!} \sum_{n=1}^{\infty} g(n)(\log n)^{\nu} n^{-2} = \frac{r^{\nu}}{\nu!} \sum_{k=1}^{\infty} (\log k^2)^{\nu}(k^2)^{-2} \\ &= \frac{(2r)^{\nu}}{\nu!} \sum_{k=1}^{\infty} (\log k)^{\nu} k^{-4} = \frac{(-2r)^{\nu}}{\nu!} \sum_{k=1}^{\infty} (-\log k)^{\nu} k^{-4} = \frac{(-2r)^{\nu}}{\nu!} \zeta^{(\nu)}(4) \end{aligned}$$

by Theorem 3B. It follows that for $0 < r < 3/2$, we have

$$F(2-r) \geq \sum_{\nu=0}^{\infty} \frac{(-2r)^{\nu}}{\nu!} \zeta^{(\nu)}(4) = \zeta(4-2r).$$

Now as $r \to 3/2-$, we must therefore have $F(2-r) \to +\infty$. This contradicts our assertion that $F(s)$ is analytic for $\sigma > 0$ and hence continuous at $s = 1/2$. ○

Problems for Chapter 4

1. Suppose that $q \in \mathbb{N}$, and that χ is a non-principal character modulo q.
   (i) Show that for every $X > 0$, we have

$$\left| \sum_{n \le X} \chi(n) \right| \le q.$$

   (ii) Noting that the function

$$S(X) = \sum_{n \le X} \chi(n)$$

   is constant between consecutive natural numbers, prove Theorem 4K.

DISTRIBUTION OF PRIME NUMBERS

W W L CHEN

© W W L Chen, 1990, 2003.
This chapter originates from material used by the author at Imperial College, University of London, between 1981 and 1990.
It is available free to all individuals, on the understanding that it is not to be used for financial gains,
and may be downloaded and/or photocopied, with or without permission from the author.
However, this document may not be kept on any information storage and retrieval system without permission
from the author, unless such system is not accessible to any individuals other than its owners.

Chapter 5

THE PRIME NUMBER THEOREM

5.1. Some Preliminary Remarks

In this chapter, we give an analytic proof of the famous Prime number theorem, a result first obtained in 1896 independently by Hadamard and de la Vallée Poussin.

THEOREM 5A. (PRIME NUMBER THEOREM) *We have*

$$\pi(X) \sim \frac{X}{\log X} \quad \textit{as } X \to \infty.$$

As in our earlier study of the distribution of primes, we use the von Mangoldt function Λ. For every $X > 0$, let

$$\psi(X) = \sum_{n \le X} \Lambda(n).$$

THEOREM 5B. *As $X \to \infty$, we have*

$$\psi(X) \sim X \qquad \textit{if and only if} \qquad \pi(X) \sim \frac{X}{\log X}.$$

Proof. Recall the proof of Theorem 2E due to Tchebycheff. We have

$$(1) \qquad \psi(X) = \sum_{n \le X} \Lambda(n) = \sum_{\substack{p,k \\ p^k \le X}} \log p = \sum_{p \le X} (\log p) \sum_{1 \le k \le \frac{\log X}{\log p}} 1 = \sum_{p \le X} (\log p) \left[\frac{\log X}{\log p} \right] \le \pi(X) \log X.$$

On the other hand, for any $\alpha \in (0,1)$, we have

$$\psi(X) \ge \sum_{p\le X} \log p \ge \sum_{X^\alpha < p \le X} \log p \ge (\pi(X) - \pi(X^\alpha)) \log(X^\alpha) = \alpha(\pi(X) - \pi(X^\alpha)) \log X. \tag{2}$$

Combining (1) and (2), we have

$$\alpha \frac{\pi(X)}{X/\log X} - \alpha \frac{\pi(X^\alpha)}{X/\log X} \le \frac{\psi(X)}{X} \le \frac{\pi(X)}{X/\log X}. \tag{3}$$

Since $\alpha < 1$, it follows from Tchebycheff's theorem that

$$\frac{\pi(X^\alpha)}{X/\log X} \to 0 \quad \text{as } X \to \infty.$$

Suppose that $\pi(X) \sim X/\log X$ as $X \to \infty$. Then

$$\alpha \frac{\pi(X)}{X/\log X} - \alpha \frac{\pi(X^\alpha)}{X/\log X} \to \alpha \quad \text{as } X \to \infty.$$

It follows that for any $\epsilon > 0$, the inequality

$$\alpha - \epsilon \le \frac{\psi(X)}{X} \le 1 + \epsilon$$

holds for all large X. Since $\alpha < 1$ is arbitrary, we must have

$$\frac{\psi(X)}{X} \to 1 \quad \text{as } X \to \infty.$$

Note next that the inequalities (3) can be rewritten as

$$\frac{\psi(X)}{X} \le \frac{\pi(X)}{X/\log X} \le \frac{1}{\alpha} \frac{\psi(X)}{X} + \frac{\pi(X^\alpha)}{X/\log X}.$$

Suppose that $\psi(X) \sim X$ as $X \to \infty$. Then

$$\frac{1}{\alpha} \frac{\psi(X)}{X} + \frac{\pi(X^\alpha)}{X/\log X} \to \frac{1}{\alpha} \quad \text{as } X \to \infty.$$

It follows that for every $\epsilon > 0$, the inequality

$$1 - \epsilon \le \frac{\pi(X)}{X/\log X} \le \frac{1+\epsilon}{\alpha}$$

holds for all large X. Since $\alpha < 1$ is arbitrary, we must have

$$\frac{\pi(X)}{X/\log X} \to 1 \quad \text{as } X \to \infty.$$

This completes the proof. ○

5.2. A Smoothing Argument

To prove the Prime number theorem, it suffices to show that $\psi(X) \sim X$ as $X \to \infty$. However, a direct discussion of $\psi(X)$ introduces various tricky convergence problems. We therefore consider a smooth average of the function ψ. For $X > 0$, let

$$\psi_1(X) = \int_0^X \psi(x)\,\mathrm{d}x. \tag{4}$$

THEOREM 5C. *Suppose that $\psi_1(X) \sim \frac{1}{2}X^2$ as $X \to \infty$. Then $\psi(X) \sim X$ as $X \to \infty$.*

PROOF. Suppose that $0 < \alpha < 1 < \beta$. Since $\Lambda(n) \geq 0$ for every $n \in \mathbb{N}$, the function ψ is an increasing function. Hence for every $X > 0$, we have

$$\psi(X) \leq \frac{1}{\beta X - X}\int_X^{\beta X} \psi(x)\,\mathrm{d}x = \frac{\psi_1(\beta X) - \psi_1(X)}{(\beta - 1)X},$$

so that

$$\frac{\psi(X)}{X} \leq \frac{\psi_1(\beta X) - \psi_1(X)}{(\beta - 1)X^2}. \tag{5}$$

On the other hand, for every $X > 0$, we have

$$\psi(X) \geq \frac{1}{X - \alpha X}\int_{\alpha X}^{X} \psi(x)\,\mathrm{d}x = \frac{\psi_1(X) - \psi_1(\alpha X)}{(1 - \alpha)X},$$

so that

$$\frac{\psi(X)}{X} \geq \frac{\psi_1(X) - \psi_1(\alpha X)}{(1 - \alpha)X^2}. \tag{6}$$

As $X \to \infty$, we have

$$\frac{\psi_1(\beta X) - \psi_1(X)}{(\beta - 1)X^2} \sim \frac{1}{\beta - 1}\left(\frac{1}{2}\beta^2 - \frac{1}{2}\right) = \frac{1}{2}(\beta + 1) \tag{7}$$

and

$$\frac{\psi_1(X) - \psi_1(\alpha X)}{(1 - \alpha)X^2} \sim \frac{1}{1 - \alpha}\left(\frac{1}{2} - \frac{1}{2}\alpha^2\right) = \frac{1}{2}(\alpha + 1). \tag{8}$$

Since α and β are arbitrary, we conclude, on combining (5)–(8), that $\psi(X)/X \sim 1$ as $X \to \infty$. ○

The rest of this chapter is concerned with establishing the following crucial result.

THEOREM 5D. *We have*

$$\psi_1(X) \sim \tfrac{1}{2}X^2 \quad \text{as } X \to \infty.$$

5.3. A Contour Integral

The following result brings the Riemann zeta function $\zeta(s)$ into the argument.

THEOREM 5E. *Suppose that $X > 0$ and $c > 1$. Then*

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

where the path of integration is the straight line $\sigma = c$.

A crucial step in the proof of Theorem 5E is provided by the following auxiliary result.

THEOREM 5F. *Suppose that $Y > 0$ and $c > 1$. Then*

$$\frac{1}{2\pi \mathrm{i}} \int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} \frac{Y^s}{s(s+1)}\,\mathrm{d}s = \begin{cases} 0 & \text{if } Y \le 1, \\ 1 - \dfrac{1}{Y} & \text{if } Y \ge 1. \end{cases}$$

PROOF. Note first of all that the integral is absolutely convergent, since

$$\left| \frac{Y^s}{s(s+1)} \right| \le \frac{Y^c}{|t|^2}$$

whenever $\sigma = c$. Let $T > 1$, and write

$$I_T = \frac{1}{2\pi \mathrm{i}} \int_{c-\mathrm{i}T}^{c+\mathrm{i}T} \frac{Y^s}{s(s+1)}\,\mathrm{d}s.$$

Suppose first of all that $Y \ge 1$. Consider the circular arc $A^-(c,T)$ centred at $s = 0$ and passing from $c - \mathrm{i}T$ to $c + \mathrm{i}T$ on the left of the line $\sigma = c$, and let

$$J_T^- = \frac{1}{2\pi \mathrm{i}} \int_{A^-(c,T)} \frac{Y^s}{s(s+1)}\,\mathrm{d}s.$$

Note that on $A^-(c,T)$, we have $|Y^s| = Y^\sigma \le Y^c$ since $Y \ge 1$; also we have $|s| = R$ and $|s+1| \ge R - 1$, where $R = (c^2 + T^2)^{1/2}$ is the radius of $A^-(c,T)$. It follows that

$$|J_T^-| \le \frac{1}{2\pi} \frac{Y^c}{R(R-1)} 2\pi R \le \frac{Y^c}{T-1} \to 0 \quad \text{as } T \to \infty.$$

By Cauchy's residue theorem, we have

$$I_T = J_T^- + \operatorname{res}\left(\frac{Y^s}{s(s+1)}, 0\right) + \operatorname{res}\left(\frac{Y^s}{s(s+1)}, -1\right) = J_T^- + 1 - \frac{1}{Y}.$$

The result for $Y \ge 1$ follows on letting $T \to \infty$.

Suppose now that $Y \le 1$. Consider the circular arc $A^+(c,T)$ centred at $s = 0$ and passing from $c - \mathrm{i}T$ to $c + \mathrm{i}T$ on the right of the line $\sigma = c$, and let

$$J_T^+ = \frac{1}{2\pi \mathrm{i}} \int_{A^+(c,T)} \frac{Y^s}{s(s+1)}\,\mathrm{d}s.$$

Note that on $A^+(c,T)$, we have $|Y^s| = Y^\sigma \le Y^c$ since $Y \le 1$; also we have $|s| = R$ and $|s+1| \ge R$, where $R = (c^2+T^2)^{1/2}$ is the radius of $A^+(c,T)$. It follows that

$$|J_T^+| \le \frac{1}{2\pi}\frac{Y^c}{R^2}2\pi R \le \frac{Y^c}{T} \to 0 \quad \text{as } T \to \infty.$$

By Cauchy's integral theorem, we have

$$I_T = J_T^+.$$

The result for $Y \le 1$ follows on letting $T \to \infty$. ○

Proof of Theorem 5E. Note that for $X \ge 1$, we have

$$\psi_1(X) = \int_0^X \psi(x)\,\mathrm{d}x = \int_1^X \psi(x)\,\mathrm{d}x = \int_1^X \left(\sum_{n\le x}\Lambda(n)\right)\mathrm{d}x = \sum_{n\le X}(X-n)\Lambda(n),$$

the last equality following from interchanging the order of integration and summation. Note also that the above conclusion holds trivially if $0 < X < 1$. It therefore follows from Theorem 5F that for every $X > 0$, we have

$$\frac{\psi_1(X)}{X} = \sum_{n\le X}\left(1-\frac{n}{X}\right)\Lambda(n) = \sum_{n\le X}\left(1-\frac{1}{X/n}\right)\Lambda(n) = \sum_{n=1}^\infty \frac{\Lambda(n)}{2\pi\mathrm{i}}\int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty}\frac{(X/n)^s}{s(s+1)}\,\mathrm{d}s,$$

where $c > 1$. Since $c > 1$, the order of summation and integration can be interchanged, as

$$\sum_{n=1}^\infty \int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty}\left|\frac{\Lambda(n)(X/n)^s}{s(s+1)}\right||\mathrm{d}s| \le X^c\sum_{n=1}^\infty\frac{\Lambda(n)}{n^c}\int_{-\infty}^{\infty}\frac{\mathrm{d}t}{c^2+t^2}$$

is finite. It follows that

$$\frac{\psi_1(X)}{X} = \frac{1}{2\pi\mathrm{i}}\int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty}\frac{X^s}{s(s+1)}\sum_{n=1}^\infty\frac{\Lambda(n)}{n^s}\,\mathrm{d}s = -\frac{1}{2\pi\mathrm{i}}\int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty}\frac{X^s}{s(s+1)}\frac{\zeta'(s)}{\zeta(s)}\,\mathrm{d}s$$

as required. ○

5.4. The Riemann Zeta Function

Recall first of all Theorem 4L. In the case of the Riemann zeta function, equation (7) of Chapter 4 becomes

(9)
$$\begin{aligned}\sum_{n\le X} n^{-s} &= s\int_1^X [x]x^{-s-1}\,\mathrm{d}x + [X]X^{-s}\\ &= s\int_1^X x^{-s}\,\mathrm{d}x - s\int_1^X \{x\}x^{-s-1}\,\mathrm{d}x + X^{-s+1} - \{X\}X^{-s}\\ &= \frac{s}{s-1} - \frac{s}{(s-1)X^{s-1}} - s\int_1^X\frac{\{x\}}{x^{s+1}}\,\mathrm{d}x + \frac{1}{X^{s-1}} - \frac{\{X\}}{X^s}.\end{aligned}$$

Letting $X \to \infty$, we deduce that

(10)
$$\zeta(s) = \frac{s}{s-1} - s\int_1^\infty \frac{\{x\}}{x^{s+1}}\,\mathrm{d}x$$

if $\sigma > 1$. Recall also that (10) gives an analytic continuation of $\zeta(s)$ to $\sigma > 0$, with s simple pole at $s = 1$. We shall use these formulae to deduce important information about the order of magnitude of $|\zeta(s)|$ in the neighbourhood of the line $\sigma = 1$ and to the left of it. Note that $\zeta(\sigma + it)$ and $\zeta(\sigma - it)$ are complex conjugates, so it suffices to study $\zeta(s)$ on the upper half plane.

THEOREM 5G. *For every $\sigma \geq 1$ and $t \geq 2$, we have*
(i) $|\zeta(s)| = O(\log t)$*; and*
(ii) $|\zeta'(s)| = O(\log^2 t)$.
Suppose further that $0 < \delta < 1$. *Then for every $\sigma \geq \delta$ and $t \geq 1$, we have*
(iii) $|\zeta(s)| = O_\delta(t^{1-\delta})$.

PROOF. For $\sigma > 0$, $t \geq 1$ and $X \geq 1$, we have, by (9) and (10), that

$$\zeta(s) - \sum_{n \leq X} \frac{1}{n^s} = \frac{s}{(s-1)X^{s-1}} - \frac{1}{X^{s-1}} + \frac{\{X\}}{X^s} - s\int_X^\infty \frac{\{x\}}{x^{s+1}}\,\mathrm{d}x = \frac{1}{(s-1)X^{s-1}} + \frac{\{X\}}{X^s} - s\int_X^\infty \frac{\{x\}}{x^{s+1}}\,\mathrm{d}x. \tag{11}$$

It follows that

$$|\zeta(s)| \leq \sum_{n \leq X} \frac{1}{n^\sigma} + \frac{1}{tX^{\sigma-1}} + \frac{1}{X^\sigma} + |s|\int_X^\infty \frac{\mathrm{d}x}{x^{\sigma+1}} \leq \sum_{n \leq X} \frac{1}{n^\sigma} + \frac{1}{tX^{\sigma-1}} + \frac{1}{X^\sigma} + \left(1 + \frac{t}{\sigma}\right)\frac{1}{X^\sigma}. \tag{12}$$

If $\sigma \geq 1$, $t \geq 1$ and $X \geq 1$, then

$$|\zeta(s)| \leq \sum_{n \leq X} \frac{1}{n} + \frac{1}{t} + \frac{1}{X} + \frac{1+t}{X} \leq (\log X + 1) + 3 + \frac{t}{X}.$$

Choosing $X = t$, we obtain

$$|\zeta(s)| \leq (\log t + 1) + 4 = O(\log t),$$

proving (i). On the other hand, if $\sigma \geq \delta$, $t \geq 1$ and $X \geq 1$, then it follows from (12) that

$$|\zeta(s)| \leq \sum_{n \leq X} \frac{1}{n^\delta} + \frac{1}{tX^{\delta-1}} + \left(2 + \frac{t}{\delta}\right)\frac{1}{X^\delta} \leq \int_0^{[X]} \frac{\mathrm{d}x}{x^\delta} + \frac{X^{1-\delta}}{t} + \frac{3t}{\delta X^\delta} \leq \frac{X^{1-\delta}}{1-\delta} + X^{1-\delta} + \frac{3t}{\delta X^\delta}.$$

Again choosing $X = t$, we obtain

$$|\zeta(s)| \leq t^{1-\delta}\left(\frac{1}{1-\delta} + 1 + \frac{3}{\delta}\right), \tag{13}$$

proving (iii). To deduce (ii), we may differentiate (11) with respect to s and proceed in a similar way. Alternatively, suppose that $s_0 = \sigma_0 + \mathrm{i}t_0$ satisfies $\sigma_0 \geq 1$ and $t_0 \geq 2$. Let C be the circle with centre s_0 and radius $\rho < 1/2$. Then Cauchy's integral formula gives

$$|\zeta'(s_0)| = \left|\frac{1}{2\pi\mathrm{i}}\int_C \frac{\zeta(s)}{(s-s_0)^2}\,\mathrm{d}s\right| \leq \frac{M}{\rho},$$

where $M = \sup_{s \in C}|\zeta(s)|$. Note next that for every $s \in C$, we clearly have $\sigma \geq \sigma_0 - \rho \geq 1 - \rho$ and $2t_0 > t \geq t_0 - \rho > 1$. It follows from (13), with $\delta = 1 - \rho$, that for every $s \in C$, we must have

$$|\zeta(s)| \leq (2t_0)^\rho\left(\frac{1}{\rho} + 1 + \frac{3}{1-\rho}\right) \leq \frac{10t_0^\rho}{\rho},$$

since $\rho < 1/2 < 1 - \rho < 1$. It follows that

$$|\zeta'(s_0)| \leq \frac{10t_0^\rho}{\rho^2}.$$

We now take $\rho = (\log t_0 + 2)^{-1}$. Then $t_0^\rho = \mathrm{e}^{\rho \log t_0} < \mathrm{e}$, and so

$$|\zeta'(s_0)| \leq 10\mathrm{e}(\log t_0 + 2)^2.$$

(ii) now follows. ○

THEOREM 5H. *The function $\zeta(s)$ has no zeros on the line $\sigma = 1$. Furthermore, there is a positive constant A such that as $t \to \infty$, we have, for $\sigma \geq 1$, that*

$$\frac{1}{\zeta(s)} = O\left((\log t)^A\right).$$

PROOF. For every $\theta \in \mathbb{R}$, we clearly have

$$3 + 4\cos\theta + \cos 2\theta = 2(1 + \cos\theta)^2 \geq 0. \tag{14}$$

On the other hand, it is easy to check that for $\sigma > 1$, we have

$$\log \zeta(s) = \sum_p \sum_{m=1}^\infty \frac{1}{mp^{ms}},$$

so that

$$\log|\zeta(\sigma + \mathrm{i}t)| = \mathfrak{Re}\left(\sum_{n=2}^\infty c_n n^{-\sigma - \mathrm{i}t}\right) = \sum_{n=2}^\infty c_n n^{-\sigma}\cos(t\log n), \tag{15}$$

where

$$c_n = \begin{cases} 1/m & \text{if } n = p^m, \text{ where } p \text{ is prime and } m \in \mathbb{N}, \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

Combining (14)–(16), we have

$$\log|\zeta^3(\sigma)\zeta^4(\sigma + \mathrm{i}t)\zeta(\sigma + 2\mathrm{i}t)| = \sum_{n=2}^\infty c_n n^{-\sigma}(3 + 4\cos(t\log n) + \cos(2t\log n)) \geq 0.$$

It follows that for $\sigma > 1$, we have

$$|(\sigma - 1)\zeta(\sigma)|^3 \left|\frac{\zeta(\sigma + \mathrm{i}t)}{\sigma - 1}\right|^4 |\zeta(\sigma + 2\mathrm{i}t)| \geq \frac{1}{\sigma - 1}. \tag{17}$$

Suppose that the point $s = 1 + it$ is a zero of $\zeta(s)$. Then since $\zeta(s)$ is analytic at the points $s = 1 + it$ and $s = 1 + 2it$ and has a simple pole with residue 1 at $s = 1$, the left hand side of (17) must converge to a finite limit as $\sigma \to 1+$, contradicting the fact that the right hand side diverges to infinity as $\sigma \to 1+$. Hence $s = 1 + it$ cannot be a zero of $\zeta(s)$. To prove the second assertion, we may assume without loss of generality that $1 \leq \sigma \leq 2$, since for $\sigma \geq 2$, we have

$$\left|\frac{1}{\zeta(s)}\right| = \left|\prod_p (1 - p^{-s})\right| \leq \prod_p (1 + p^{-\sigma}) < \zeta(\sigma) \leq \zeta(2).$$

Suppose now that $1 < \sigma \le 2$ and $t \ge 2$. Then by (17), we have

$$(\sigma-1)^3 \le |(\sigma-1)\zeta(\sigma)|^3|\zeta(\sigma+\mathrm{i}t)|^4|\zeta(\sigma+2\mathrm{i}t)| \le A_1|\zeta(\sigma+\mathrm{i}t)|^4\log(2t)$$

by Theorem 5G(i), where A_1 is a positive absolute constant. Since $\log(2t) \le 2\log t$, it follows that

$$|\zeta(\sigma+\mathrm{i}t)| \ge \frac{(\sigma-1)^{3/4}}{A_2(\log t)^{1/4}}, \tag{18}$$

where A_2 is a positive absolute constant. Note that (18) holds also when $\sigma = 1$. Suppose now that $1 < \eta < 2$. If $1 \le \sigma \le \eta$ and $t \ge 2$, then it follows from Theorem 5G(ii) that

$$|\zeta(\sigma+\mathrm{i}t)-\zeta(\eta+\mathrm{i}t)| = \left|\int_\sigma^\eta \zeta'(x+\mathrm{i}t)\,\mathrm{d}x\right| \le A_3(\eta-1)\log^2 t,$$

where A_3 is a positive absolute constant. Combining this with (18), we have

$$|\zeta(\sigma+\mathrm{i}t)| \ge |\zeta(\eta+\mathrm{i}t)| - A_3(\eta-1)\log^2 t \ge \frac{(\eta-1)^{3/4}}{A_2(\log t)^{1/4}} - A_3(\eta-1)\log^2 t. \tag{19}$$

On the other hand, if $\eta \le \sigma \le 2$ and $t \ge 2$, then in view of (18), the inequality (19) must also hold. It follows that inequality (19) holds if $1 \le \sigma \le 2$, $t \ge 2$ and $1 < \eta < 2$. We now choose η so that

$$\frac{(\eta-1)^{3/4}}{A_2(\log t)^{1/4}} = 2A_3(\eta-1)\log^2 t;$$

in other words, we choose

$$\eta = 1 + (2A_2A_3)^{-4}(\log t)^{-9},$$

where $t > t_0$, so that $\eta < 2$. Then

$$|\zeta(\sigma+\mathrm{i}t)| \ge A_3(\eta-1)\log^2 t = A_4(\log t)^{-7}$$

for $1 \le \sigma \le 2$ and $t > t_0$. ○

5.5. Completion of the Proof

We are now ready to complete the proof of Theorem 5D. By Theorem 5E, we have

$$\frac{\psi_1(X)}{X^2} = \frac{1}{2\pi\mathrm{i}}\int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} G(s)X^{s-1}\,\mathrm{d}s, \tag{20}$$

where $c > 1$ and $X > 0$, and where

$$G(s) = -\frac{1}{s(s+1)}\frac{\zeta'(s)}{\zeta(s)} = -\frac{1}{s(s+1)}\zeta'(s)\frac{1}{\zeta(s)}.$$

By Theorems 4H, 5G and 5H, we know that $G(s)$ is analytic for $\sigma \ge 1$, except at $s = 1$, and that for some positive absolute constant A, we have

$$G(s) = O\left(|t|^{-2}(\log|t|)^2(\log|t|)^A\right) < |t|^{-3/2} \tag{21}$$

for all $|t| > t_0$. Let $\epsilon > 0$ be given. We now consider a contour made up of the straight line segments

$$\begin{cases} L_1 = [1 - \mathrm{i}U, 1 - \mathrm{i}T], \\ L_2 = [1 - \mathrm{i}T, \alpha - \mathrm{i}T], \\ L_3 = [\alpha - \mathrm{i}T, \alpha + \mathrm{i}T], \\ L_4 = [\alpha + \mathrm{i}T, 1 + \mathrm{i}T], \\ L_5 = [1 + \mathrm{i}T, 1 + \mathrm{i}U], \end{cases}$$

where $T = T(\epsilon) > \max\{t_0, 2\}$, $\alpha = \alpha(T) = \alpha(\epsilon) \in (0, 1)$ and U are chosen to satisfy the following three conditions:

(i) We have

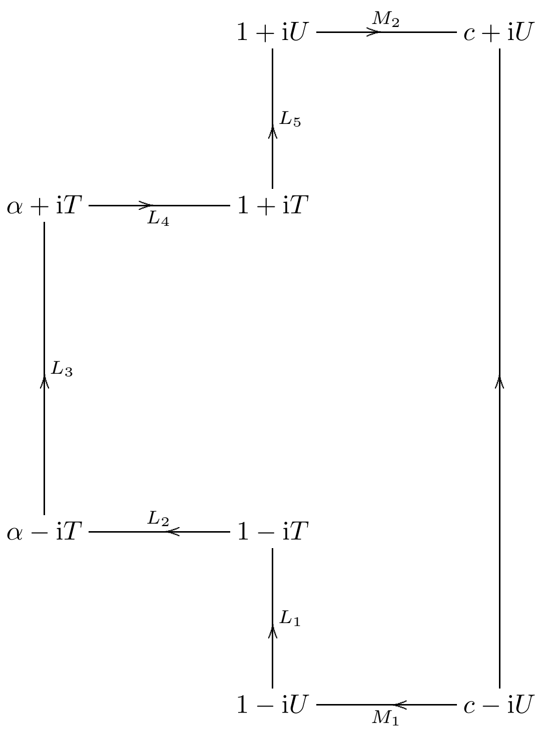
$$\int_T^\infty |G(1 + \mathrm{i}t)|\,\mathrm{d}t < \epsilon.$$

(ii) The rectangle $[\alpha, 1] \times [-T, T]$ contains no zeros of $\zeta(s)$. Note that this is possible since $\zeta(s)$ has no zeros on the line $\sigma = 1$ and, as an analytic function, has at most a finite number of zeros in the region $[1/2, 1) \times [-T, T]$.

(iii) We have $U > T$.

Furthermore, define the straight line segments

$$\begin{cases} M_1 = [c - \mathrm{i}U, 1 - \mathrm{i}U], \\ M_2 = [1 + \mathrm{i}U, c + \mathrm{i}U]. \end{cases}$$



By Cauchy's residue theorem, we have

$$\frac{1}{2\pi\mathrm{i}} \int_{c-\mathrm{i}U}^{c+\mathrm{i}U} G(s) X^{s-1}\,\mathrm{d}s = \frac{1}{2\pi\mathrm{i}} \sum_{j=1}^{2} \int_{M_j} G(s) X^{s-1}\,\mathrm{d}s + \frac{1}{2\pi\mathrm{i}} \sum_{j=1}^{5} \int_{L_j} G(s) X^{s-1}\,\mathrm{d}s + \mathrm{res}(G(s)X^{s-1}, 1), \tag{22}$$

where, for every $X > 1$, we have

$$\mathrm{res}(G(s)X^{s-1}, 1) = \tfrac{1}{2}. \tag{23}$$

Now

$$\left|\int_{L_1} G(s)X^{s-1}\,\mathrm{d}s\right| = \left|\int_{L_5} G(s)X^{s-1}\,\mathrm{d}s\right| \le \int_T^\infty |G(1+\mathrm{i}t)|\,\mathrm{d}t < \epsilon. \tag{24}$$

On the other hand,

$$\left|\int_{L_2} G(s)X^{s-1}\,\mathrm{d}s\right| = \left|\int_{L_4} G(s)X^{s-1}\,\mathrm{d}s\right| \le M\int_\alpha^1 X^{\sigma-1}\,\mathrm{d}\sigma \le \frac{M}{\log X} \tag{25}$$

and

$$\left|\int_{L_3} G(s)X^{s-1}\,\mathrm{d}s\right| \le 2TMX^{\alpha-1}, \tag{26}$$

where

$$M = M(\alpha, T) = M(\epsilon) = \sup_{L_2\cup L_3\cup L_4} |G(s)|. \tag{27}$$

Furthermore, by (21), we have, for $j = 1, 2$,

$$\left|\int_{M_j} G(s)X^{s-1}\,\mathrm{d}s\right| \le \int_1^c |U|^{-3/2}X^{\sigma-1}\,\mathrm{d}\sigma \le \frac{X^{c-1}}{\log X}|U|^{-3/2}. \tag{28}$$

Combining (22)–(28), we have

$$\left|\frac{1}{2\pi\mathrm{i}}\int_{c-\mathrm{i}U}^{c+\mathrm{i}U} G(s)X^{s-1}\,\mathrm{d}s - \frac{1}{2}\right| \le \frac{\epsilon}{\pi} + \frac{M}{\pi\log X} + \frac{TM}{\pi X^{1-\alpha}} + \frac{X^{c-1}|U|^{-3/2}}{\pi\log X}.$$

On letting $U \to \infty$, we have

$$\left|\frac{\psi_1(X)}{X^2} - \frac{1}{2}\right| \le \frac{\epsilon}{\pi} + \frac{M}{\pi\log X} + \frac{TM}{\pi X^{1-\alpha}}.$$

It then follows that

$$\lim_{X\to\infty}\left|\frac{\psi_1(X)}{X^2} - \frac{1}{2}\right| \le \frac{\epsilon}{\pi}.$$

Note finally that $\epsilon > 0$ is arbitrary, and the left hand side is independent of ϵ. It follows that

$$\lim_{X\to\infty}\frac{\psi_1(X)}{X^2} = \frac{1}{2}.$$

This completes the proof of Theorem 5D.

DISTRIBUTION OF PRIME NUMBERS

W W L CHEN

© W W L Chen, 1990, 2003.

This chapter originates from material used by the author at Imperial College, University of London, between 1981 and 1990.
It is available free to all individuals, on the understanding that it is not to be used for financial gains,
and may be downloaded and/or photocopied, with or without permission from the author.
However, this document may not be kept on any information storage and retrieval system without permission
from the author, unless such system is not accessible to any individuals other than its owners.

Chapter 6

THE RIEMANN ZETA FUNCTION

6.1. Riemann's Memoir

In Riemann's only paper on number theory, published in 1860, he proved the following result.

THEOREM 6A. (RIEMANN) *The function $\zeta(s)$ can be continued analytically over the whole complex plane $\mathbb{C}$, and satisfies the functional equation*

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

where Γ denotes the gamma function. In particular, the function $\zeta(s)$ is analytic everywhere, except for a simple pole at $s=1$ with residue 1.

Note that the functional equation (1) enables properties of $\zeta(s)$ for $\sigma<0$ to be inferred from properties of $\zeta(s)$ for $\sigma>1$.

Remarks. (i) As can be observed from the functional equation (1), the study of the Riemann zeta function depends intimately on properties of the gamma function. The latter is usually defined by the Euler integral

$$\Gamma(s)=\int_0^\infty \mathrm{e}^{-t}t^{s-1}\,\mathrm{d}t,$$

valid whenever $\mathfrak{Re}\,s>0$, and satisfies $\Gamma(s+1)=s\Gamma(s)$. The Weierstrass formula

$$\frac{1}{s\Gamma(s)}=\mathrm{e}^{\gamma s}\prod_{n=1}^\infty\left(1+\frac{s}{n}\right)\mathrm{e}^{-s/n},$$

where γ is Euler's constant, extends the gamma function to the whole complex plane $\mathbb{C}$. It is then easy to see that $\Gamma(s)$ has no zeros, but has simple poles at $s = 0, -1, -2, \ldots$.

(ii) The formulas

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

and

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

are particularly useful in the study of $\zeta(s)$, as we shall see later in the proof of Theorem 6V.

(iii) Stirling's asymptotic formula

$$\log \Gamma(s) = \left(s - \frac{1}{2}\right) \log s - s + \frac{\log 2\pi}{2} + O(|s|^{-1}) \quad \text{as } |s| \to \infty$$

is valid in any angle $-\pi + \delta < \arg s < \pi - \delta$ for any fixed $\delta > 0$. The same condition gives the estimate

$$\frac{\Gamma'(s)}{\Gamma(s)} = \log s + O(|s|^{-1}) \quad \text{as } |s| \to \infty.$$

(iv) The interested reader may refer to Chapters 12 and 13 in the volume *Modern Analysis* by Whittaker and Watson for detailed proofs of the above.

In view of Remark (i) above, the only zeros of $\zeta(s)$ for $\sigma < 0$ are at the poles of $\Gamma(s/2)$; in other words, at the points $s = -2, -4, -6, \ldots$. These are called the trivial zeros of $\zeta(s)$.

The part of the plane with $0 \le \sigma \le 1$ is called the critical strip.

Riemann's paper is particularly remarkable for the conjectures it contains. While most of these conjectures have been proved, the famous Riemann hypothesis has so far resisted all attempts to prove or disprove it.

THEOREM 6B. (HADAMARD 1893) *The function $\zeta(s)$ has infinitely many zeros in the critical strip.*

It is easy to see that the zeros of $\zeta(s)$ in the critical strip are placed symmetrically with respect to the line $t = 0$ as well as with respect to the line $\sigma = 1/2$, the latter observation being a consequence of the functional equation (1).

THEOREM 6C. (HADAMARD 1893) *The entire function*

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

has the product representation

$$\xi(s) = \mathrm{e}^{A+Bs} \prod_{\rho} \left(1 - \frac{s}{\rho}\right) \mathrm{e}^{s/\rho}, \tag{3}$$

where A and B are constants and where ρ runs over all the zeros of the function $\zeta(s)$ in the critical strip.

We comment here that the product representation (3) plays an important role in the first proof of the Prime number theorem.

THEOREM 6D. (VON MANGOLDT 1905) *Let $N(T)$ denote the number of zeros $\rho = \beta + \mathrm{i}\gamma$ of the function $\zeta(s)$ in the critical strip with $0 < \gamma \le T$. Then*

$$N(T) = \frac{T}{2\pi}\log\frac{T}{2\pi} - \frac{T}{2\pi} + O(\log T). \tag{4}$$

The most remarkable of Riemann's conjectures is an explicit formula for the difference $\pi(X) - \mathrm{li}(X)$, containing a term which is a sum over the zeros of $\zeta(s)$ in the critical strip. This shows that the zeros of $\zeta(s)$ plays a crucial role in the study of the distribution of primes. Here we state a result closely related to this formula.

THEOREM 6E. (VON MANGOLDT 1895) *Let*

$$\psi(X) = \sum_{n \le X} \Lambda(n) \qquad \textit{and} \qquad \psi_0(X) = \frac{\psi(X-0) + \psi(X+0)}{2}.$$

Then

$$\psi_0(X) - X = -\sum_{\rho} \frac{X^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2}\log\left(1 - \frac{1}{X^2}\right),$$

where the terms in the sum arising from complex conjugates are taken together.

However, there remains one of Riemann's conjectures which is still unsolved today. The open question below is arguably the most famous unsolved problem in the whole of mathematics.

CONJECTURE. (RIEMANN HYPOTHESIS) *The zeros of the function $\zeta(s)$ in the critical strip all lie on the line $\sigma = 1/2$.*

We shall nevertheless establish the following rather weak partial result which gives a zero-free region for $\zeta(s)$. This will be sufficient to give another proof of the Prime number theorem, via the explicit formula given in Theorem 6E.

THEOREM 6F. (DE LA VALLÉE-POUSSIN 1899) *There exists an absolute constant $c > 0$ such that the function $\zeta(s)$ has no zeros in the region*

$$\sigma \ge 1 - \frac{c}{\log t} \qquad \textit{and} \qquad t \ge 2.$$

6.2. Riemann's Proof of the Functional Equation

Suppose that $\sigma > 0$. Writing $t = n^2\pi x$, we have

$$\Gamma\left(\frac{s}{2}\right) = \int_0^\infty t^{s/2-1}\mathrm{e}^{-t}\,\mathrm{d}t = (n^2\pi)^{s/2}\int_0^\infty x^{s/2-1}\mathrm{e}^{-n^2\pi x}\,\mathrm{d}x,$$

so that

$$\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)n^{-s} = \int_0^\infty x^{s/2-1}\mathrm{e}^{-n^2\pi x}\,\mathrm{d}x.$$

It follows that for $\sigma > 1$, we have

$$\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \sum_{n=1}^{\infty}\int_0^{\infty} x^{s/2-1}\mathrm{e}^{-n^2\pi x}\,\mathrm{d}x = \int_0^{\infty} x^{s/2-1}\left(\sum_{n=1}^{\infty}\mathrm{e}^{-n^2\pi x}\right)\mathrm{d}x,$$

where the change of order of summation and integration is justified by the convergence of

$$\sum_{n=1}^{\infty}\int_0^{\infty} x^{\sigma/2-1}\mathrm{e}^{-n^2\pi x}\,\mathrm{d}x.$$

Now write

$$\omega(x) = \sum_{n=1}^{\infty}\mathrm{e}^{-n^2\pi x}.$$

Then for $\sigma > 1$, we have

$$\begin{aligned}\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s) &= \int_1^{\infty} x^{s/2-1}\omega(x)\,\mathrm{d}x + \int_0^1 y^{s/2-1}\omega(y)\,\mathrm{d}y \\ &= \int_1^{\infty} x^{s/2-1}\omega(x)\,\mathrm{d}x + \int_1^{\infty} x^{-s/2-1}\omega(x^{-1})\,\mathrm{d}x.\end{aligned} \tag{5}$$

We shall show that for every $x > 0$, the function

$$\theta(x) = \sum_{n=-\infty}^{\infty}\mathrm{e}^{-n^2\pi x} = 1 + 2\omega(x)$$

satisfies the functional equation $\theta(x^{-1}) = x^{1/2}\theta(x)$ which can be written in the form

$$\sum_{n=-\infty}^{\infty}\mathrm{e}^{-n^2\pi/x} = x^{1/2}\sum_{n=-\infty}^{\infty}\mathrm{e}^{-n^2\pi x}. \tag{6}$$

It then follows that

$$2\omega(x^{-1}) = \theta(x^{-1}) - 1 = x^{1/2}\theta(x) - 1 = -1 + x^{1/2} + 2x^{1/2}\omega(x),$$

so that for $\sigma > 1$, we have

$$\begin{aligned}\int_1^{\infty} x^{-s/2-1}\omega(x^{-1})\,\mathrm{d}x &= \int_1^{\infty} x^{-s/2-1}\left(-\frac{1}{2} + \frac{1}{2}x^{1/2} + x^{1/2}\omega(x)\right)\mathrm{d}x \\ &= \frac{1}{s(s-1)} + \int_1^{\infty} x^{-s/2-1/2}\omega(x)\,\mathrm{d}x.\end{aligned} \tag{7}$$

It follows on combining (5) and (7) that for $\sigma > 1$, we have

$$\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \frac{1}{s(s-1)} + \int_1^{\infty}\left(x^{s/2-1} + x^{-s/2-1/2}\right)\omega(x)\,\mathrm{d}x. \tag{8}$$

Note now that the integral on the right hand side of (8) converges absolutely for any s, and uniformly in any bounded part of the plane, since $\omega(x) = O(\mathrm{e}^{-\pi x})$ as $x \to +\infty$. Hence the integral represents an entire function of s, and the formula gives the analytic continuation of $\zeta(s)$ over the whole plane. Note also that the right hand side of (8) remains unchanged when s is replaced by $1-s$, so that the functional equation (1) follows immediately. Finally, note that the function

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

is analytic everywhere. Since $s\Gamma(s/2)$ has no zeros, the only possible pole of $\zeta(s)$ is at $s=1$, and we have already shown earlier that $\zeta(s)$ has a simple pole at $s=1$ with residue 1.

It remains to establish the functional equation (6) for every $x>0$. The starting point is the Poisson summation formula, that under certain conditions on a function $f(t)$, we have

$$\sum_{A\le n\le B}{}' f(n) = \sum_{\nu=-\infty}^{\infty}\int_A^B f(t)\mathrm{e}^{2\pi i\nu t}\,\mathrm{d}t, \tag{9}$$

where $\sum'$ denotes that the terms in the sum corresponding to $n=A$ and $n=B$ are $\frac{1}{2}f(A)$ and $\frac{1}{2}f(B)$ respectively. Using (9) with $A=-N$, $B=N$ and $f(t)=\mathrm{e}^{-t^2\pi/x}$, we have

$$\sum_{n=-N}^{N}{}' \mathrm{e}^{-n^2\pi/x} = \sum_{\nu=-\infty}^{\infty}\int_{-N}^{N}\mathrm{e}^{-t^2\pi/x}\mathrm{e}^{2\pi i\nu t}\,\mathrm{d}t.$$

Letting $N\to\infty$, we obtain

$$\sum_{n=-\infty}^{\infty}\mathrm{e}^{-n^2\pi/x} = \sum_{\nu=-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{e}^{-t^2\pi/x}\mathrm{e}^{2\pi i\nu t}\,\mathrm{d}t. \tag{10}$$

This is justified by noting that

$$\left(\int_{-\infty}^{-N}+\int_N^{\infty}\right)\mathrm{e}^{-t^2\pi/x}\mathrm{e}^{2\pi i\nu t}\,\mathrm{d}t = 2\int_N^{\infty}\mathrm{e}^{-t^2\pi/x}\cos(2\pi\nu t)\,\mathrm{d}t,$$

and that

$$\left|\sum_{\nu\ne0}\int_N^{\infty}\mathrm{e}^{-t^2\pi/x}\cos(2\pi\nu t)\,\mathrm{d}t\right| \to 0 \quad \text{as } N\to\infty.$$

Writing $t=xu$ and using (10), we have

$$\begin{aligned}\sum_{n=-\infty}^{\infty}\mathrm{e}^{-n^2\pi/x} &= x\sum_{\nu=-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{e}^{-u^2\pi x}\mathrm{e}^{2\pi i\nu xu}\,\mathrm{d}u = x\sum_{\nu=-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{e}^{-(u-i\nu)^2\pi x-\nu^2\pi x}\,\mathrm{d}u \\ &= x\sum_{\nu=-\infty}^{\infty}\mathrm{e}^{-\nu^2\pi x}\int_{-\infty}^{\infty}\mathrm{e}^{-(u-i\nu)^2\pi x}\,\mathrm{d}u.\end{aligned} \tag{11}$$

Note now that the function $\mathrm{e}^{-z^2\pi x}$ is an entire function of the complex variable z. It follows from Cauchy's integral theorem that

$$\int_{-\infty}^{\infty}\mathrm{e}^{-(u-i\nu)^2\pi x}\,\mathrm{d}u = \int_{-\infty}^{\infty}\mathrm{e}^{-u^2\pi x}\,\mathrm{d}u = Ax^{-1/2}, \tag{12}$$

where

$$A = \int_{-\infty}^{\infty}\mathrm{e}^{-y^2\pi}\,\mathrm{d}y = 1. \tag{13}$$

The functional equation (6) now follows on combining (11)–(13), and the proof of Theorem 6A is now complete.

6.3. Entire Functions

In this section, we shall prove some technical results on entire functions for use later in the proof of Theorems 6B and 6C.

An entire function $f(s)$ is said to be of order 1 if

$$f(s) = O_\alpha\left(\mathrm{e}^{|s|^\alpha}\right) \quad \text{as } |s| \to \infty \tag{14}$$

holds for every $\alpha > 1$ and fails for every $\alpha < 1$.

Suppose that the entire function $h(s)$ has no zeros on the plane. Then the function $g(s) = \log h(s)$ can be defined as a single valued function and is also entire. Suppose that

$$h(s) = O_\alpha\left(\mathrm{e}^{|s|^\alpha}\right) \quad \text{as } |s| \to \infty \tag{15}$$

holds for every $\alpha > 1$. Then

$$\mathfrak{Re}\, g(R\mathrm{e}^{\mathrm{i}\theta}) = \log |h(R\mathrm{e}^{\mathrm{i}\theta})| = O_\alpha(R^\alpha) \quad \text{as } R \to \infty$$

holds for every $\alpha > 1$. Without loss of generality, we may suppose that $g(0) = 0$. Then we can write

$$g(R\mathrm{e}^{\mathrm{i}\theta}) = \sum_{k=1}^{\infty} (a_k + \mathrm{i}b_k) R^k \mathrm{e}^{\mathrm{i}k\theta}, \quad \text{where } a_k, b_k \in \mathbb{R},$$

so that

$$\mathfrak{Re}\, g(R\mathrm{e}^{\mathrm{i}\theta}) = \sum_{k=1}^{\infty} a_k R^k \cos k\theta - \sum_{k=1}^{\infty} b_k R^k \sin k\theta.$$

Note now that for every $k, n \in \mathbb{N}$, we have

$$\int_0^{2\pi} \cos k\theta \cos n\theta \,\mathrm{d}\theta = \begin{cases} \pi & \text{if } k = n, \\ 0 & \text{if } k \neq n, \end{cases}$$

and

$$\int_0^{2\pi} \sin k\theta \cos n\theta \,\mathrm{d}\theta = 0.$$

It follows that

$$\int_0^{2\pi} \left(\mathfrak{Re}\, g(R\mathrm{e}^{\mathrm{i}\theta})\right) \cos n\theta \,\mathrm{d}\theta = \pi a_n R^n,$$

so that

$$\pi |a_n| R^n \leq \int_0^{2\pi} \left|\mathfrak{Re}\, g(R\mathrm{e}^{\mathrm{i}\theta})\right| \mathrm{d}\theta = O_\alpha(R^\alpha)$$

holds for all sufficiently large R and every $\alpha > 1$. On letting $R \to \infty$, we see that $a_n = 0$ for every $n > 1$. A similar argument using the function $\sin n\theta$ instead of the function $\cos n\theta$ gives $b_n = 0$ for every $n > 1$. We have therefore proved the following result.

THEOREM 6G. *Suppose that the entire function $h(s)$ has no zeros on the complex plane $\mathbb{C}$, and that (15) holds for every $\alpha > 1$. Then $h(s) = \mathrm{e}^{A+Bs}$, where A and B are constants.*

REMARK. In the preceding argument, note that it is enough to assume that the estimates for $h(s)$ hold for a sequence of values R with limit infinity.

Our next task is to study the distribution of the zeros of an entire function. The first step in this direction is summarized by the result below.

THEOREM 6H. (JENSEN'S FORMULA) *Suppose that an entire function $f(s)$ satisfies $f(0) \neq 0$. Suppose further that $s_1, \ldots, s_n$ are the zeros of $f(s)$ in $|s| < R$, counted with multiplicities, and that there are no zeros of $f(s)$ on $|s| = R$. Then*

$$\frac{1}{2\pi} \int_0^{2\pi} \log |f(R\mathrm{e}^{\mathrm{i}\theta})| \, \mathrm{d}\theta - \log |f(0)| = \log \frac{R^n}{|s_1 \ldots s_n|}. \tag{16}$$

PROOF. We may clearly write

$$f(s) = (s - s_1) \ldots (s - s_n) k(s),$$

where $k(s)$ is analytic and has no zeros in $|s| \leq R$, so that $\log k(s)$ is analytic in $|s| \leq R$. It follows from Gauss's mean value theorem that

$$\frac{1}{2\pi} \int_0^{2\pi} \log k(R\mathrm{e}^{\mathrm{i}\theta}) \, \mathrm{d}\theta = \log k(0).$$

Taking real parts, we obtain

$$\frac{1}{2\pi} \int_0^{2\pi} \log |k(R\mathrm{e}^{\mathrm{i}\theta})| \, \mathrm{d}\theta = \log |k(0)| = \log |f(0)| - \log |s_1 \ldots s_n|. \tag{17}$$

Unfortunately, for every $j = 1, \ldots, n$, we cannot apply a similar argument to $\log |s - s_j|$, since the function $s - s_j$ has a zero at s_j. Note, however, that the function

$$\frac{R^2 - \overline{s_j} s}{R}$$

has no zeros in $|s| \leq R$ and satisfies

$$\left| \frac{R^2 - \overline{s_j} s}{R} \right| = |s - s_j|$$

on the circle $|s| = R$, so that

$$\frac{1}{2\pi} \int_0^{2\pi} \log |R\mathrm{e}^{\mathrm{i}\theta} - s_j| \, \mathrm{d}\theta = \frac{1}{2\pi} \int_0^{2\pi} \log \left| \frac{R^2 - \overline{s_j} R\mathrm{e}^{\mathrm{i}\theta}}{R} \right| \, \mathrm{d}\theta. \tag{18}$$

Clearly the function

$$\log \frac{R^2 - \overline{s_j} s}{R}$$

is analytic in $|s| \le R$. Applying Gauss's mean value theorem over the circle $|s| = R$ on this function and taking real parts, we conclude that the right hand side of (18) is equal to $\log R$. Finally, note that

$$\log |f(Re^{i\theta})| = \sum_{j=1}^{n} \log |Re^{i\theta} - s_j| + \log |k(Re^{i\theta})|,$$

so that

$$\frac{1}{2\pi} \int_0^{2\pi} \log |f(Re^{i\theta})| \, d\theta = n \log R + \log |f(0)| - \log |s_1 \ldots s_n|.$$

This completes the proof. ○

Remarks. (i) It is important to point out that Jensen's formula was in fact only discovered after Hadamard's work in connection with Theorems 6B and 6C.

(ii) Gauss's mean value theorem states that the value of an analytic function at the centre of a circle is equal to the arithmetic mean of its values on the circle. In particular, if the function $F(s)$ is analytic for $|s| < R_0$, then for every $R < R_0$, we have

$$F(0) = \frac{1}{2\pi} \int_0^{2\pi} F(Re^{i\theta}) \, d\theta.$$

A simple consequence of Jensen's formula is the following result on the zeros of entire functions.

THEOREM 6J. *Suppose that $f(s)$ is an entire function satisfying $f(0) \ne 0$, and that (14) holds for every $\alpha > 1$. Suppose further that $s_1, s_2, s_3, \ldots$ are the zeros of $f(s)$, counted with multiplicities and where $|s_1| \le |s_2| \le |s_3| \le \ldots$. Then for every $\alpha > 1$, the series*

$$\sum_{n=1}^{\infty} |s_n|^{-\alpha}$$

is convergent.

Proof. Note that the right hand side of (16) is equal to

$$\int_0^R r^{-1} n(r) \, dr,$$

where, for every non-negative $r \le R$, $n(r)$ denotes the number of zeros of $f(s)$ in $|s| \le r$. To see this, note that

$$\begin{aligned} \int_0^R r^{-1} n(r) \, dr &= \sum_{j=1}^{n-1} \int_{|s_j|}^{|s_{j+1}|} r^{-1} j \, dr + \int_{|s_n|}^{R} r^{-1} n \, dr \\ &= \sum_{j=1}^{n-1} j(\log |s_{j+1}| - \log |s_j|) + n(\log R - \log |s_n|) \\ &= n \log R - \log |s_1| - \ldots - \log |s_n|. \end{aligned}$$

For every $\alpha > 1$, write $\alpha^* = (\alpha + 1)/2$, so that $1 < \alpha^* < \alpha$. Then

$$\log |f(Re^{i\theta})| = O_\alpha(R^{\alpha^*}) \quad \text{as } R \to \infty,$$

so that by Jensen's formula, we have

$$\int_0^R r^{-1}n(r)\,\mathrm{d}r = O_\alpha(R^{\alpha^*}) - \log|f(0)| = O_\alpha(R^{\alpha^*}) \quad \text{as } R \to \infty.$$

On the other hand, note that

$$\int_R^{2R} r^{-1}n(r)\,\mathrm{d}r \geq n(R)\int_R^{2R} r^{-1}\,\mathrm{d}r = n(R)\log 2.$$

It follows that

$$n(R) = O_\alpha(R^{\alpha^*}) \quad \text{as } R \to \infty.$$

Hence

$$\sum_{n=1}^{\infty} |s_n|^{-\alpha} = \int_0^\infty r^{-\alpha}\,\mathrm{d}n(r) = \alpha\int_0^\infty r^{-\alpha-1}n(r)\,\mathrm{d}r < \infty.$$

This completes the proof. ○

Suppose now that $f(s)$ is an entire function satisfying $f(0) \neq 0$, and that (14) holds for every $\alpha > 1$. Suppose further that $s_1, s_2, s_3, \ldots$ are the zeros of $f(s)$, counted with multiplicities and where $|s_1| \leq |s_2| \leq |s_3| \leq \ldots$. Then for every $\epsilon > 0$, the series

$$\sum_{n=1}^{\infty} |s_n|^{-1-\epsilon}$$

converges, so that the series

$$\sum_{n=1}^{\infty} |s_n|^{-2}$$

converges, and so the product

$$P(s) = \prod_{n=1}^{\infty}\left(1 - \frac{s}{s_n}\right)\mathrm{e}^{s/s_n} \tag{19}$$

converges absolutely for every $s \in \mathbb{C}$, and uniformly in any bounded domain not containing any zeros of $f(s)$. It follows that $P(s)$ is an entire function, with zeros at $s_1, s_2, s_3, \ldots$. Now write

$$f(s) = P(s)h(s), \tag{20}$$

where $h(s)$ is an entire function without zeros. If (15) holds for every $\alpha > 1$, then $h(s) = \mathrm{e}^{A+Bs}$, where A and B are constants, and so

$$f(s) = \mathrm{e}^{A+Bs}\prod_{n=1}^{\infty}\left(1 - \frac{s}{s_n}\right)\mathrm{e}^{s/s_n}. \tag{21}$$

THEOREM 6K. *Under the hypotheses of Theorem 6J, the inequality* (15) *holds for every $\alpha > 1$, where the function $h(s)$ is defined by* (19) *and* (20). *In particular, the function $f(s)$ can be expressed in the form* (21), *where A and B are constants.*

Proof. To show that the inequality (15) holds for every $\alpha > 1$, it clearly suffices, in view of (14) and (20), to establish a suitable lower bound for $|P(s)|$. Since the series

$$\sum_{n=1}^{\infty} |s_n|^{-2}$$

is convergent, the set

$$\mathcal{S} = \bigcup_{n=1}^{\infty} (|s_n| - |s_n|^{-2}, |s_n| + |s_n|^{-2})$$

has finite total length. It follows that there exist arbitrarily large positive real numbers R such that $R \notin \mathcal{S}$. It is easy to see that for any such real number $R \notin \mathcal{S}$, we have

$$|R - |s_n|| \ge |s_n|^{-2} \quad \text{for every } n \in \mathbb{N}. \tag{22}$$

The idea now is to split up the product $P(s)$ into three products according to the size of $n \in \mathbb{N}$ relative to R. More precisely, for any such $R \notin \mathcal{S}$, write

$$P(s) = P_1(s)P_2(s)P_3(s), \tag{23}$$

where for every $j = 1, 2, 3$, we have

$$P_j(s) = \prod_{(24.j)} \left(1 - \frac{s}{s_n}\right) \mathrm{e}^{s/s_n},$$

where the products are taken over all $n \in \mathbb{N}$ satisfying

$$|s_n| < \frac{R}{2}, \tag{24.1}$$

$$\frac{R}{2} \le |s_n| < 2R, \tag{24.2}$$

$$|s_n| \ge 2R, \tag{24.3}$$

respectively. Let $\epsilon > 0$ be chosen and fixed.

Suppose first of all that (24.1) holds. Then on $|s| = R$, we have

$$\left| \left(1 - \frac{s}{s_n}\right) \mathrm{e}^{s/s_n} \right| \ge \left(\left|\frac{s}{s_n}\right| - 1 \right) \mathrm{e}^{-|s|/|s_n|} > \mathrm{e}^{-R/|s_n|},$$

and so it follows from

$$\sum_{(24.1)} |s_n|^{-1} < \left(\frac{R}{2}\right)^{\epsilon} \sum_{n=1}^{\infty} |s_n|^{-1-\epsilon}$$

that

$$|P_1(s)| \gg_\epsilon \mathrm{e}^{-R^{1+2\epsilon}} \quad \text{as } R \to \infty. \tag{25}$$

Suppose next that (24.2) holds. Then on $|s| = R$, we have

$$\left| \left(1 - \frac{s}{s_n}\right) \mathrm{e}^{s/s_n} \right| \ge \left| \frac{s_n - s}{s_n} \right| \mathrm{e}^{-|s|/|s_n|} > \frac{||s_n| - R|}{2R} \mathrm{e}^{-2} \gg R^{-3},$$

in view of (22). Note that there are at most $O_\epsilon(R^{1+\epsilon})$ values of n for which (24.2) holds. Hence on $|s| = R$, we have

$$|P_2(s)| \gg_\epsilon (R^{-3})^{R^{1+\epsilon}} \gg_\epsilon \mathrm{e}^{-R^{1+2\epsilon}} \quad \text{as } R \to \infty. \tag{26}$$

Suppose finally that (24.3) holds. Then on $|s| = R$, we have

$$\left|\left(1 - \frac{s}{s_n}\right)\mathrm{e}^{s/s_n}\right| > \mathrm{e}^{-c(R/|s_n|)^2} \tag{27}$$

for some positive constant c (see the Remark below), and so it follows from

$$\sum_{(24.3)} |s_n|^{-2} \le (2R)^{-1+\epsilon} \sum_{n=1}^{\infty} |s_n|^{-1-\epsilon}$$

that

$$|P_3(s)| \gg_\epsilon \mathrm{e}^{-R^{1+2\epsilon}} \quad \text{as } R \to \infty. \tag{28}$$

It now follows from (23), (25), (26) and (28) that on $|s| = R$, we have

$$|P(s)| \gg_\epsilon \mathrm{e}^{-R^{1+3\epsilon}} \quad \text{as } R \to \infty. \tag{29}$$

The result then follows on combining (20) and (29), and noting that the inequality (14) holds for $\alpha = 1+\epsilon$. ○

REMARK. Note that the inequality (27) is of the form

$$|(1-z)\mathrm{e}^z| > \mathrm{e}^{-c|z|^2}, \tag{30}$$

where $|z| \le 1/2$. Write $z = x + \mathrm{i}y$, where $x, y \in \mathbb{R}$. Then (30) will follow if we show that

$$(1-x)^2\mathrm{e}^{2x} > \mathrm{e}^{-2cx^2}$$

whenever $|x| \le 1/2$. This last inequality can easily be established by using the theory of real valued functions of a real variable.

Finally, we make the following simple observation.

THEOREM 6L. *Under the hypotheses of Theorem 6J, suppose further that the series*

$$\sum_{n=1}^{\infty} |s_n|^{-1}$$

is convergent. Then there exists a positive constant c such that

$$f(s) = O(\mathrm{e}^{c|s|}) \quad \textit{as } |s| \to \infty.$$

PROOF. This follows from (21) and the inequality $|(1-z)\mathrm{e}^z| \le \mathrm{e}^{2|z|}$ which holds for every $z \in \mathbb{C}$. ○

6.4. Zeros of the Zeta Function

Recall that the function $\xi(s)$, defined by (2), is an entire function, and that $\xi(0) \neq 0$. Note also that the zeros of $\xi(s)$ are precisely the zeros of $\zeta(s)$ in the critical strip. In order to establish Theorem 6C, we shall use Theorem 6K. We therefore first need to show that for every $\alpha > 1$, we have

$$\xi(s) = O_\alpha\left(\mathrm{e}^{|s|^\alpha}\right) \quad \text{as } |s| \to \infty.$$

We shall in fact prove the following stronger result.

THEOREM 6M. *There exists a positive constant c such that*

$$|\xi(s)| < \mathrm{e}^{c|s|\log|s|} \quad \textit{as } |s| \to \infty. \tag{31}$$

Furthermore, for any positive constant c, the inequality

$$|\xi(s)| < \mathrm{e}^{c|s|} \quad \textit{as } |s| \to \infty \tag{32}$$

does not hold.

PROOF. Since $\xi(s) = \xi(1-s)$ for every $s \in \mathbb{C}$, it suffices to prove the inequality (31) for $\sigma \geq 1/2$. First of all, there exists a positive constant c_1 such that

$$\left|\frac{1}{2}s(s-1)\pi^{-s/2}\right| < \mathrm{e}^{c_1|s|}.$$

Next, Stirling's formula

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

as $|s| \to \infty$ is valid in the angle $-\pi/2 < \arg s < \pi/2$, and so there exists a positive constant c_2 such that

$$\left|\Gamma\left(\frac{s}{2}\right)\right| < \mathrm{e}^{c_2|s|\log|s|}.$$

Finally, note that the formula

$$\zeta(s) = \frac{s}{s-1} - s\int_1^\infty \{x\}x^{-s-1}\,\mathrm{d}x$$

is valid for $\sigma > 0$, and the integral is bounded for $\sigma \geq 1/2$, so that there exists a positive constant c_3 such that

$$|\zeta(s)| < c_3|s|.$$

This proves (31). On the other hand, note that as $s \to +\infty$ through real values, we have

$$\log\Gamma\left(\frac{s}{2}\right) \sim \frac{s}{2}\log\frac{s}{2} \qquad \text{and} \qquad \zeta(s) \to 1,$$

so that (32) does not hold. ○

To complete the proof of Theorems 6B and 6C, note that by Theorem 6L, the series

$$\sum_\rho |\rho|^{-1}$$

is divergent, where ρ denotes the zeros of $\xi(s)$ and so the zeros of $\zeta(s)$ in the critical strip. Theorem 6B follows immediately. Theorem 6C now follows from Theorems 6K and 6M.

6.5. An Important Formula

It follows from (3) that

$$\log \xi(s) = A + Bs + \sum_{\rho} \left(\frac{s}{\rho} + \log \left(1 - \frac{s}{\rho} \right) \right).$$

Differentiating with respect to s, we obtain

$$\frac{\xi'(s)}{\xi(s)} = B + \sum_{\rho} \left(\frac{1}{\rho} + \frac{1}{s-\rho} \right). \tag{33}$$

On the other hand, it follows from (2) and $s\Gamma(s) = \Gamma(s+1)$ that

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

Differentiating with respect to s, we obtain

$$\frac{\xi'(s)}{\xi(s)} = \frac{1}{s-1} - \frac{1}{2} \log \pi + \frac{1}{2} \frac{\Gamma'(\frac{s}{2}+1)}{\Gamma(\frac{s}{2}+1)} + \frac{\zeta'(s)}{\zeta(s)}. \tag{34}$$

Combining (33) and (34), we obtain the following result.

THEOREM 6N. *We have*

$$\frac{\zeta'(s)}{\zeta(s)} = B - \frac{1}{s-1} + \frac{1}{2} \log \pi - \frac{1}{2} \frac{\Gamma'(\frac{s}{2}+1)}{\Gamma(\frac{s}{2}+1)} + \sum_{\rho} \left(\frac{1}{\rho} + \frac{1}{s-\rho} \right), \tag{35}$$

where B is a constant and where ρ denotes the zeros of the function $\zeta(s)$ in the critical strip.

The formula (35) clearly exhibits the pole of $\zeta(s)$ at $s = 1$ and the zeros ρ in the critical strip. On the other hand, the trivial zeros are exhibited by the term

$$-\frac{1}{2} \frac{\Gamma'(\frac{s}{2}+1)}{\Gamma(\frac{s}{2}+1)}.$$

To see this last point, we start from the Weierstrass formula

$$\frac{1}{\Gamma(s+1)} = \frac{1}{s\Gamma(s)} = \mathrm{e}^{\gamma s} \prod_{n=1}^{\infty} \left(1 + \frac{s}{n} \right) \mathrm{e}^{-s/n},$$

where γ is Euler's constant. This gives

$$\frac{1}{\Gamma(\frac{s}{2}+1)} = \mathrm{e}^{\gamma s/2} \prod_{n=1}^{\infty} \left(1 + \frac{s}{2n} \right) \mathrm{e}^{-s/2n}.$$

Taking logarithms, we obtain

$$-\log \Gamma \left(\frac{s}{2} + 1 \right) = \frac{1}{2} \gamma s + \sum_{n=1}^{\infty} \left(\log \left(1 + \frac{s}{2n} \right) - \frac{s}{2n} \right).$$

Differentiating with respect to s, we obtain

$$-\frac{1}{2}\frac{\Gamma'(\frac{s}{2}+1)}{\Gamma(\frac{s}{2}+1)}=\frac{1}{2}\gamma+\sum_{n=1}^{\infty}\left(\frac{1}{s+2n}-\frac{1}{2n}\right).$$

6.6. A Zero-Free Region

Recall Theorem 5H, where we show that the function $\zeta(s)$ has no zeros on the line $\sigma=1$ by using the function $\log\zeta(s)$ together with the observation that

$$3+4\cos\theta+\cos 2\theta=2(1+\cos\theta)^2\geq 0.$$

Here it is more convenient to work with the logarithmic derivative $\zeta'(s)/\zeta(s)$, since its only poles for $\sigma>0$ are at $s=1$ and the zeros of $\zeta(s)$ in the critical strip. Starting from the Dirichlet series

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

valid for $\sigma>1$, we immediately deduce that

$$\mathfrak{Re}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)=\sum_{n=1}^{\infty}\frac{\Lambda(n)}{n^\sigma}\cos(t\log n).$$

It follows that for every $\sigma>1$,

$$3\left(-\frac{\zeta'(\sigma)}{\zeta(\sigma)}\right)+4\mathfrak{Re}\left(-\frac{\zeta'(\sigma+\mathrm{i}t)}{\zeta(\sigma+\mathrm{i}t)}\right)+\mathfrak{Re}\left(-\frac{\zeta'(\sigma+2\mathrm{i}t)}{\zeta(\sigma+2\mathrm{i}t)}\right)\geq 0. \tag{36}$$

The simple pole of $\zeta(s)$ at $s=1$ leads to a simple pole of $-\zeta'(s)/\zeta(s)$ there with residue 1. Hence there exists a positive absolute constant A_1 such that

$$-\frac{\zeta'(\sigma)}{\zeta(\sigma)}<\frac{1}{\sigma-1}+A_1\quad\text{if } 1<\sigma\leq 2. \tag{37}$$

On the other hand, it is well known that there exists a positive absolute constant A_2 such that the gamma function $\Gamma(s)$ satisfies the inequality

$$\frac{1}{2}\frac{\Gamma'(\frac{s}{2}+1)}{\Gamma(\frac{s}{2}+1)}<A_2\log t\quad\text{if } 1<\sigma\leq 2 \text{ and } t\geq 2.$$

It follows from the identity (35) that there exists a positive absolute constant A_3 such that

$$\mathfrak{Re}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)<A_3\log t-\sum_{\rho}\mathfrak{Re}\left(\frac{1}{\rho}+\frac{1}{s-\rho}\right)\quad\text{if } 1<\sigma\leq 2 \text{ and } t\geq 2. \tag{38}$$

Suppose that $\rho=\beta+\mathrm{i}\gamma$, where $\beta,\gamma\in\mathbb{R}$, is a zero of the function $\zeta(s)$ in the critical strip. Then $0<\beta<1$, and since $\sigma>1$, we have

$$\mathfrak{Re}\frac{1}{\rho}=\frac{\beta}{|\rho|^2}>0\qquad\text{and}\qquad\mathfrak{Re}\frac{1}{s-\rho}=\frac{\sigma-\beta}{|s-\rho|^2}>0.$$

This means that the inequality (38) remains valid if we omit any term from the sum on the right hand side. In particular, when $s = \sigma + 2\mathrm{i}t$, we have the inequality

$$\mathfrak{Re}\left(-\frac{\zeta'(\sigma+2\mathrm{i}t)}{\zeta(\sigma+2\mathrm{i}t)}\right) < A_3\log(2t) < A_4\log t \quad \text{if } 1<\sigma\le 2 \text{ and } t\ge 2, \tag{39}$$

where A_4 is a positive absolute constant.

Suppose now that $t \ge 2$ is fixed and there exists a real number β such that $\rho = \beta + \mathrm{i}t$ is a zero of the function $\zeta(s)$ in the critical strip. Then removing all but one term from the sum on the right hand side of (38), we have

$$\mathfrak{Re}\left(-\frac{\zeta'(\sigma+\mathrm{i}t)}{\zeta(\sigma+\mathrm{i}t)}\right) < A_3\log t - \mathfrak{Re}\frac{1}{(\sigma+\mathrm{i}t)-(\beta+\mathrm{i}t)} = A_3\log t - \frac{1}{\sigma-\beta} \quad \text{if } 1<\sigma\le 2. \tag{40}$$

Combining (36), (37), (39) and (40), we obtain

$$0 < \frac{3}{\sigma-1} + 3A_1 + (4A_3+A_4)\log t - \frac{4}{\sigma-\beta} \quad \text{if } 1<\sigma\le 2.$$

In other words, there exists a positive absolute constant A_5 such that

$$\frac{4}{\sigma-\beta} < \frac{3}{\sigma-1} + A_5\log t \quad \text{if } 1<\sigma\le 2.$$

Let $\sigma = 1+\delta/\log t$, where $\delta>0$ will be chosen later, sufficiently small to guarantee that $1<\sigma\le 2$. Then elementary calculation gives the inequality

$$\beta < 1 - \frac{\delta(1-A_5\delta)}{(3+A_5\delta)\log t}.$$

We now choose δ in terms of A_5 to conclude that there exists a positive absolute constant c such that

$$\beta < 1 - \frac{c}{\log t}. \tag{41}$$

In conclusion, we have shown that if $t\ge 2$ and $\beta+\mathrm{i}t$ is a zero of the function $\zeta(s)$ in the critical strip, then the inequality (41) must hold. Theorem 6F follows immediately.

6.7. Counting Zeros in the Critical Strip

The starting point of our discussion is based on the Argument principle. Suppose that the function $F(s)$ is analytic, apart from a finite number of poles, in the closure of a domain D bounded by a simple closed positively oriented Jordan curve C. Suppose further that $F(s)$ has no zeros or poles on C. Then

$$\frac{1}{2\pi\mathrm{i}}\int_C \frac{F'(s)}{F(s)}\,\mathrm{d}s = \frac{1}{2\pi}\Delta_C \arg F(s)$$

represents the total number of zeros of $F(s)$ in D minus the total number of poles of $F(s)$ in D, counted with multiplicities. Here $\Delta_C \arg F(s)$ denotes the change of argument of the function $F(s)$ along C.

It is convenient to use the function $\xi(s)$, since it is entire and its zeros are precisely the zeros of $\zeta(s)$ in the critical strip. To calculate $N(T)$, it is convenient to take the domain $(-1,2)\times(0,T)$, so that C is the rectangular path passing through the vertices

$$2, \quad 2+\mathrm{i}T, \quad -1+\mathrm{i}T, \quad -1$$

in the anticlockwise direction. If no zeros of $\zeta(s)$ has imaginary part T, then

$$N(T) = \frac{1}{2\pi\mathrm{i}}\int_C \frac{\xi'(s)}{\xi(s)}\,\mathrm{d}s = \frac{1}{2\pi}\Delta_C \arg\xi(s).$$

Let us now divide C into the following parts. First, let L_1 denote the line segment from -1 to 2. Next, let L_2 denote the line segment from 2 to $2+\mathrm{i}T$, followed by the line segment from $2+\mathrm{i}T$ to $\frac{1}{2}+\mathrm{i}T$. Finally, let L_3 denote the line segment from $\frac{1}{2}+\mathrm{i}T$ to $-1+\mathrm{i}T$, followed by the line segment from $-1+\mathrm{i}T$ to -1.

Since $\xi(s)$ is real and non-zero on L_1, clearly $\Delta_{L_1} \arg \xi(s) = 0$. On the other hand,

$$\xi(\sigma + \mathrm{i}t) = \xi(1 - \sigma - \mathrm{i}t) = \overline{\xi(1 - \sigma + \mathrm{i}t)},$$

so that $\Delta_{L_2} \arg \xi(s) = \Delta_{L_3} \arg \xi(s)$. If we write $L = L_2$, so that L denotes the line segment from 2 to $2 + \mathrm{i}T$, followed by the line segment from $2 + \mathrm{i}T$ to $\frac{1}{2} + \mathrm{i}T$, then

$$\pi N(T) = \Delta_L \arg \xi(s). \tag{42}$$

Recall that

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

It follows that

$$\Delta_L \arg \xi(s) = \Delta_L \arg(s-1) + \Delta_L \arg \pi^{-s/2} + \Delta_L \arg \Gamma\left(\frac{s}{2} + 1\right) + \Delta_L \arg \zeta(s). \tag{43}$$

Clearly

$$\Delta_L \arg(s-1) = \arg\left(-\frac{1}{2} + \mathrm{i}T\right) = \frac{1}{2}\pi + O(T^{-1}) \tag{44}$$

and

$$\Delta_L \arg \pi^{-s/2} = \Delta_L \left(-\frac{1}{2} t \log \pi\right) = -\frac{1}{2} T \log \pi. \tag{45}$$

On the other hand,

$$\Delta_L \arg \Gamma\left(\frac{s}{2} + 1\right) = \mathfrak{Im} \log \Gamma\left(\frac{5}{4} + \frac{1}{2}\mathrm{i}T\right).$$

By Stirling's formula,

$$\log \Gamma\left(\frac{5}{4} + \frac{1}{2}\mathrm{i}T\right) = \left(\frac{3}{4} + \frac{1}{2}\mathrm{i}T\right) \log\left(\frac{5}{4} + \frac{1}{2}\mathrm{i}T\right) - \frac{5}{4} - \frac{1}{2}\mathrm{i}T + \frac{1}{2}\log \pi + O(T^{-1}),$$

so that

$$\Delta_L \arg \Gamma\left(\frac{s}{2} + 1\right) = \frac{T}{2} \log \frac{T}{2} + \frac{3}{8}\pi - \frac{T}{2} + O(T^{-1}). \tag{46}$$

Combining (42)–(46), we have

$$\begin{aligned} N(T) &= \frac{1}{2} - \frac{T}{2\pi}\log \pi + \frac{T}{2\pi}\log\frac{T}{2} + \frac{3}{8} - \frac{T}{2\pi} + S(T) + O(T^{-1}) \\ &= \frac{T}{2\pi}\log\frac{T}{2\pi} - \frac{T}{2\pi} + \frac{7}{8} + S(T) + O(T^{-1}), \end{aligned}$$

where

$$\pi S(T) = \Delta_L \arg \zeta(s).$$

To prove Theorem 6D, it suffices to prove the following result.

THEOREM 6P. *We have $S(T) = O(\log T)$ as $T \to \infty$.*

Note first of all that $\arg\zeta(2) = 0$. On the other hand,

$$\arg\zeta(s) = \tan^{-1}\left(\frac{\mathfrak{Im}\zeta(s)}{\mathfrak{Re}\zeta(s)}\right)$$

and $\mathfrak{Re}\zeta(s) \neq 0$ on the line $\sigma = 2$. It follows that

$$|\arg\zeta(2 + \mathrm{i}T)| < \frac{\pi}{2},$$

and so $\Delta_{[2,2+\mathrm{i}T]}\arg\zeta(s) = O(1)$. Hence we may assume, without loss of generality, that L is the line segment from $2 + \mathrm{i}T$ to $\frac{1}{2} + \mathrm{i}T$. We shall give two proofs, the first of which uses Jensen's formula.

Suppose that $\mathfrak{Re}\zeta(s)$ vanishes q times on the line segment from $2 + \mathrm{i}T$ to $\frac{1}{2} + \mathrm{i}T$. Then this line segment can be divided into $q + 1$ parts, where in each subinterval, $\mathfrak{Re}\zeta(s)$ may vanish only at one or both of the endpoints and has constant sign strictly in between, so that the variation of $\arg\zeta(s)$ in each such subinterval does not exceed π. It follows that

$$S(T) \ll (q+1)\pi + \frac{1}{2}\pi. \tag{47}$$

To prove our result, it remains to find a suitable bound for q. For $s = \sigma + \mathrm{i}T$, we have

$$\mathfrak{Re}\zeta(s) = \frac{1}{2}(\zeta(\sigma + \mathrm{i}T) + \zeta(\sigma - \mathrm{i}T)).$$

Let T be fixed, and consider the function

$$f_T(s) = \frac{1}{2}(\zeta(s + \mathrm{i}T) + \zeta(s - \mathrm{i}T))$$

(note that we no longer insist that $s = \sigma + \mathrm{i}T$). Then q is the number of zeros of $f_T(s)$ on the line segment from $1/2$ to 2, and so is bounded above by the number of zeros of $f_T(s)$ in the disc $|s - 2| \leq 3/2$. In other words,

$$q \leq n\left(\frac{3}{2}\right), \tag{48}$$

where, for every $r \geq 0$, $n(r)$ denotes the number of zeros of $f_T(s)$ in the disc $|s - 2| \leq r$. By Jensen's formula and noting that we may assume that $\zeta(\frac{1}{2} + \mathrm{i}T) \neq 0$, we have

$$\int_0^{7/4}\frac{n(r)}{r}\,\mathrm{d}r = \frac{1}{2\pi}\int_0^{2\pi}\log\left|f_T\left(2 + \frac{7}{4}\mathrm{e}^{\mathrm{i}\theta}\right)\right|\,\mathrm{d}\theta - \log|f_T(2)|. \tag{49}$$

On the other hand,

$$\int_0^{7/4}\frac{n(r)}{r}\,\mathrm{d}r \geq \int_{3/2}^{7/4}\frac{n(r)}{r}\,\mathrm{d}r \geq n\left(\frac{3}{2}\right)\int_{3/2}^{7/4}\frac{1}{r}\,\mathrm{d}r = n\left(\frac{3}{2}\right)\log\frac{7}{6}. \tag{50}$$

Observe that

$$\begin{aligned}|f_T(2)| &= \left|\frac{1}{2}(\zeta(2 + \mathrm{i}T) + \zeta(2 - \mathrm{i}T))\right| = |\mathfrak{Re}\zeta(2 + \mathrm{i}T)|\\ &= \left|\mathfrak{Re}\left(\sum_{n=1}^{\infty}\frac{1}{n^{2+\mathrm{i}T}}\right)\right| \geq 1 - \sum_{n=2}^{\infty}\frac{1}{n^2} = 2 - \frac{\pi^2}{6} > 0,\end{aligned}$$

so that

$$-\log|f_T(2)| = O(1). \tag{51}$$

Finally, recall that $|\zeta(s)| \ll T^{3/4}$ for every $\sigma \ge 1/4$. It follows that for every $\theta \in [0, 2\pi]$, we have

$$\left|f_T\left(2+\frac{7}{4}\mathrm{e}^{\mathrm{i}\theta}\right)\right| \le \frac{1}{2}\left(\left|\zeta\left(2+\frac{7}{4}\mathrm{e}^{\mathrm{i}\theta}+\mathrm{i}T\right)\right| + \left|\zeta\left(2+\frac{7}{4}\mathrm{e}^{\mathrm{i}\theta}-\mathrm{i}T\right)\right|\right) \ll T^{3/4},$$

so that

$$\log\left|f_T\left(2+\frac{7}{4}\mathrm{e}^{\mathrm{i}\theta}\right)\right| \ll \log T. \tag{52}$$

Combining (49)–(52), we conclude that

$$n\left(\frac{3}{2}\right) \ll \log T. \tag{53}$$

Theorem 6P now follows on combining (47), (48) and (53).

The starting point of our second proof of Theorem 6P is the observation that

$$\Delta_{[\frac{1}{2}+\mathrm{i}T,2+\mathrm{i}T]} \arg\zeta(s) = \int_{\frac{1}{2}+\mathrm{i}T}^{2+\mathrm{i}T} \mathfrak{Im}\frac{\zeta'(s)}{\zeta(s)}\,\mathrm{d}s,$$

since $\arg\zeta(s) = \mathfrak{Im}\log\zeta(s)$. We therefore need to study the logarithmic derivative of $\zeta(s)$ on the line segment between $\frac{1}{2}+\mathrm{i}T$ and $2+\mathrm{i}T$, and show that

$$\int_{\frac{1}{2}+\mathrm{i}T}^{2+\mathrm{i}T} \mathfrak{Im}\frac{\zeta'(s)}{\zeta(s)}\,\mathrm{d}s = O(\log T).$$

This approach has the added bonus of providing some intermediate results which are useful in the deduction of the asymptotic formula given in Theorem 6E. Recall the inequality (38). If we write $s = 2+\mathrm{i}T$, where $T \ge 2$, then

$$\mathfrak{Re}\left(-\frac{\zeta'(2+\mathrm{i}T)}{\zeta(2+\mathrm{i}T)}\right) < A_3\log T - \sum_\rho \mathfrak{Re}\left(\frac{1}{\rho}+\frac{1}{2+\mathrm{i}T-\rho}\right),$$

where A_3 is a positive absolute constant. Clearly

$$\mathfrak{Re}\left(-\frac{\zeta'(2+\mathrm{i}T)}{\zeta(2+\mathrm{i}T)}\right) = \sum_{n=1}^\infty \frac{\Lambda(n)}{n^2}\cos(T\log n) = O\left(\sum_{n=1}^\infty \frac{\log n}{n^2}\right) = O(1),$$

with the immediate consequence that

$$\sum_\rho \mathfrak{Re}\left(\frac{1}{\rho}+\frac{1}{2+\mathrm{i}T-\rho}\right) = O(\log T).$$

Writing $\rho = \beta+\mathrm{i}\gamma$, where $\beta,\gamma\in\mathbb{R}$, we see that

$$\mathfrak{Re}\frac{1}{\rho} > 0 \qquad \text{and} \qquad \mathfrak{Re}\frac{1}{2+\mathrm{i}T-\rho} = \frac{2-\beta}{(2-\beta)^2+(T-\gamma)^2} \ge \frac{1}{4+(T-\gamma)^2}.$$

We have proved the following result.

THEOREM 6Q. *For all sufficiently large positive real numbers T, we have*

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

where ρ denotes the zeros of the function $\zeta(s)$ in the critical strip.

This has two immediate consequences. Their proofs are left as exercises.

THEOREM 6R. *For all sufficiently large positive real numbers T, the number of zeros of the function $\zeta(s)$ in the critical strip with $|\gamma - T| < 1$ is $O(\log T)$.*

THEOREM 6S. *For all sufficiently large positive real numbers T, we have*

$$\sum_{\substack{\rho \\ |\gamma-T|\ge 1}} \frac{1}{(T-\gamma)^2} = O(\log T),$$

where ρ denotes the zeros of the function $\zeta(s)$ in the critical strip.

The crucial estimate is given by the following result. The range for σ is greater than for our present need, but will be necessary for later use.

THEOREM 6T. *For every $s = \sigma + \mathrm{i}T$, where $-1 \le \sigma \le 2$ and where T is sufficiently large and $T \ne \gamma$ for any zero $\rho = \beta + \mathrm{i}\gamma$ of the function $\zeta(s)$ in the critical strip, we have*

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_{\substack{\rho \\ |\gamma-T|<1}} \frac{1}{s-\rho} + O(\log T), \tag{54}$$

where ρ denotes the zeros of the function $\zeta(s)$ in the critical strip.

Proof. We start with the formula

$$\frac{\zeta'(s)}{\zeta(s)} = B - \frac{1}{s-1} + \frac{1}{2}\log\pi - \frac{1}{2}\frac{\Gamma'(\frac{s}{2}+1)}{\Gamma(\frac{s}{2}+1)} + \sum_{\rho}\left(\frac{1}{\rho} + \frac{1}{s-\rho}\right), \tag{55}$$

given in Theorem 6N. Writing $s = 2 + \mathrm{i}T$, we have

$$\frac{\zeta'(2+\mathrm{i}T)}{\zeta(2+\mathrm{i}T)} = B - \frac{1}{1+\mathrm{i}T} + \frac{1}{2}\log\pi - \frac{1}{2}\frac{\Gamma'(2+\frac{1}{2}\mathrm{i}T)}{\Gamma(2+\frac{1}{2}\mathrm{i}T)} + \sum_{\rho}\left(\frac{1}{\rho} + \frac{1}{2+\mathrm{i}T-\rho}\right). \tag{56}$$

Note now that $\zeta'(2+\mathrm{i}T)/\zeta(2+\mathrm{i}T) = O(1)$ and $\Gamma'(2+\frac{1}{2}\mathrm{i}T)/\Gamma(2+\frac{1}{2}\mathrm{i}T) = O(\log T)$ for every T under consideration, and $\Gamma'(\frac{s}{2}+1)/\Gamma(\frac{s}{2}+1) = O(\log T)$ for every s under consideration. It follows on subtracting (56) from (55) that

$$\begin{aligned}\frac{\zeta'(s)}{\zeta(s)} &= \sum_{\rho}\left(\frac{1}{s-\rho} - \frac{1}{2+\mathrm{i}T-\rho}\right) + O(\log T) \\ &= \sum_{\substack{\rho \\ |\gamma-T|<1}} \frac{1}{s-\rho} - \sum_{\substack{\rho \\ |\gamma-T|<1}} \frac{1}{2+\mathrm{i}T-\rho} + \sum_{\substack{\rho \\ |\gamma-T|\ge 1}}\left(\frac{1}{s-\rho} - \frac{1}{2+\mathrm{i}T-\rho}\right) + O(\log T).\end{aligned} \tag{57}$$

On the one hand, every zero ρ of the function $\zeta(s)$ in the critical strip satisfies $|2+\mathrm{i}T-\rho| > 1$. Hence

$$\sum_{\substack{\rho \\ |\gamma-T|<1}} \frac{1}{2+\mathrm{i}T-\rho} = O\left(\sum_{\substack{\rho \\ |\gamma-T|<1}} 1\right) = O(\log T), \tag{58}$$

in view of Theorem 6R. On the other hand, if $|\gamma - T| \ge 1$, then

$$\left|\frac{1}{s-\rho} - \frac{1}{2+\mathrm{i}T-\rho}\right| = \frac{2-\sigma}{|(s-\rho)(2+\mathrm{i}T-\rho)|} \le \frac{3}{(T-\gamma)^2},$$

and so

$$\sum_{\substack{\rho \\ |\gamma - T| \geq 1}} \left(\frac{1}{s-\rho} - \frac{1}{2 + \mathrm{i}T - \rho} \right) = O\left(\sum_{\substack{\rho \\ |\gamma - T| \geq 1}} \frac{1}{(T-\gamma)^2} \right) = O(\log T), \tag{59}$$

in view of Theorem 6S. The inequality (54) now follows on combining (57)–(59). ○

Taking imaginary parts in the inequality (54) on the line segment between $\frac{1}{2} + \mathrm{i}T$ and $2 + \mathrm{i}T$, where T is sufficiently large and $T \neq \gamma$ for any zero $\rho = \beta + \mathrm{i}\gamma$ of the function $\zeta(s)$ in the critical strip, gives

$$\mathfrak{Im} \frac{\zeta'(s)}{\zeta(s)} = \sum_{\substack{\rho \\ |\gamma - T| < 1}} \mathfrak{Im} \frac{1}{s-\rho} + O(\log T).$$

Combining this with the simple observation that

$$\left| \int_{\frac{1}{2}+\mathrm{i}T}^{2+\mathrm{i}T} \mathfrak{Im} \frac{1}{s-\rho} \, \mathrm{d}s \right| = \left| \Delta_{[\frac{1}{2}+\mathrm{i}T, 2+\mathrm{i}T]} \arg(s-\rho) \right| < \pi$$

and Theorem 6R, we obtain

$$\int_{\frac{1}{2}+\mathrm{i}T}^{2+\mathrm{i}T} \mathfrak{Im} \frac{\zeta'(s)}{\zeta(s)} \, \mathrm{d}s = O\left(\sum_{\substack{\rho \\ |\gamma - T| < 1}} 1 \right) + O(\log T) = O(\log T).$$

This completes the proof of Theorem 6P.

6.8. An Asymptotic Formula

In this section, we shall establish the asymptotic formula in Theorem 6E when $X > \mathrm{e}$. Here the starting point is the simple observation that

$$\psi_0(X) = \sideset{}{'}\sum_{n \leq X} \Lambda(n),$$

where $\sum'$ denotes that the term in the sum corresponding to $n = X$ is $\frac{1}{2}\Lambda(n)$. Analogous to the discussion for $\psi_1(X)$ in Chapter 5, we want to write

$$\psi_0(X) = \sideset{}{'}\sum_{n \leq X} \Lambda(n) = \sum_{n=1}^{\infty} \Lambda(n) I\left(\frac{X}{n}\right), \tag{60}$$

where the function

$$I(Y) = \begin{cases} 0 & \text{if } 0 < Y < 1, \\ 1/2 & \text{if } Y = 1, \\ 1 & \text{if } Y > 1, \end{cases} \tag{61}$$

brings the function $\zeta(s)$ into play.

The following is a suitable analogue of Theorem 5F. However, the proof is much more complicated since we do not have absolute convergence.

THEOREM 6U. *Suppose that $Y>0$ and $c>0$. Let*

$$I(Y)=\frac{1}{2\pi\mathrm{i}}\int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty}\frac{Y^s}{s}\,\mathrm{d}s,$$

where the integral in the case $Y=1$ is defined to be the limit of

$$I(Y,T)=\frac{1}{2\pi\mathrm{i}}\int_{c-\mathrm{i}T}^{c+\mathrm{i}T}\frac{Y^s}{s}\,\mathrm{d}s$$

as $T\to\infty$. Then (60) *and* (61) *hold. Furthermore, for every $T>0$, we have*

$$|I(Y)-I(Y,T)|\le\begin{cases}Y^c\min\{1,(\pi T|\log Y|)^{-1}\} & \text{if } Y\ne 1,\\ c(\pi T)^{-1} & \text{if } Y=1.\end{cases}\tag{62}$$

PROOF. For every $T_1,T_2>0$, write

$$I(Y,T_1,T_2)=\frac{1}{2\pi\mathrm{i}}\int_{c-\mathrm{i}T_1}^{c+\mathrm{i}T_2}\frac{Y^s}{s}\,\mathrm{d}s.$$

We shall consider three cases, corresponding to $Y>1$, $0<Y<1$ and $Y=1$.

Suppose first of all that $Y>1$. We consider the rectangular path with vertices

$$c-\mathrm{i}T_1,\quad c+\mathrm{i}T_2,\quad -u+\mathrm{i}T_2,\quad -u-\mathrm{i}T_1,$$

where $u>0$, followed in the anticlockwise direction. Applying Cauchy's residue theorem, we obtain

$$I(Y,T_1,T_2)-1=\frac{1}{2\pi\mathrm{i}}\int_{-u-\mathrm{i}T_1}^{-u+\mathrm{i}T_2}\frac{Y^s}{s}\,\mathrm{d}s+\frac{1}{2\pi\mathrm{i}}\int_{-u+\mathrm{i}T_2}^{c+\mathrm{i}T_2}\frac{Y^s}{s}\,\mathrm{d}s-\frac{1}{2\pi\mathrm{i}}\int_{-u-\mathrm{i}T_1}^{c-\mathrm{i}T_1}\frac{Y^s}{s}\,\mathrm{d}s.\tag{63}$$

On the vertical edge $\sigma=-u$, we have $|Y^s/s|\le Y^{-u}/u$, and so

$$\left|\frac{1}{2\pi\mathrm{i}}\int_{-u-\mathrm{i}T_1}^{-u+\mathrm{i}T_2}\frac{Y^s}{s}\,\mathrm{d}s\right|\le\frac{(T_1+T_2)Y^{-u}}{2\pi u}.\tag{64}$$

On the horizontal edge $t=T_2$, we have $|Y^s/s|\le Y^\sigma/T_2$, and so

$$\left|\frac{1}{2\pi\mathrm{i}}\int_{-u+\mathrm{i}T_2}^{c+\mathrm{i}T_2}\frac{Y^s}{s}\,\mathrm{d}s\right|\le\frac{1}{2\pi}\int_{-\infty}^{c}\frac{Y^\sigma}{T_2}\,\mathrm{d}\sigma=\frac{Y^c}{2\pi T_2\log Y}.\tag{65}$$

On the horizontal edge $t=-T_1$, we have $|Y^s/s|\le Y^\sigma/T_1$, and so

$$\left|\frac{1}{2\pi\mathrm{i}}\int_{-u-\mathrm{i}T_1}^{c-\mathrm{i}T_1}\frac{Y^s}{s}\,\mathrm{d}s\right|\le\frac{1}{2\pi}\int_{-\infty}^{c}\frac{Y^\sigma}{T_1}\,\mathrm{d}\sigma=\frac{Y^c}{2\pi T_1\log Y}.\tag{66}$$

Combining (63)–(66), we obtain

$$|I(Y,T_1,T_2)-1|\le\frac{(T_1+T_2)Y^{-u}}{2\pi u}+\frac{Y^c}{2\pi T_1\log Y}+\frac{Y^c}{2\pi T_2\log Y}.$$

Since the left hand side is independent of u, it follows on letting $u\to\infty$ that

$$|I(Y,T_1,T_2)-1|\le\frac{Y^c}{2\pi T_1\log Y}+\frac{Y^c}{2\pi T_2\log Y}.$$

Letting $T_1, T_2 \to \infty$ gives (61). Letting $T = T_1 = T_2$ gives one of the inequalities in (62). To deduce the other inequality, we use the circular arc $A^-(c,T)$ centred at $s = 0$ and passing from $c - \mathrm{i}T$ to $c + \mathrm{i}T$ on the left of the line $\sigma = c$, as in the proof of Theorem 5F. Then Cauchy's residue theorem gives

$$I(Y,T) - I(Y) = I(Y,T) - 1 = \frac{1}{2\pi\mathrm{i}} \int_{A^-(c,T)} \frac{Y^s}{s}\,\mathrm{d}s. \tag{67}$$

On the circular arc $A^-(c,T)$, we have $|Y^s/s| \le Y^c/R$, where $R = (c^2 + T^2)^{1/2}$ is the radius of $A^-(c,T)$. It follows that

$$\left| \frac{1}{2\pi\mathrm{i}} \int_{A^-(c,T)} \frac{Y^s}{s}\,\mathrm{d}s \right| \le \frac{1}{2\pi} \frac{Y^c}{R} 2\pi R = Y^c. \tag{68}$$

The inequality $|I(Y) - I(Y,T)| \le Y^c$ now follows on combining (67) and (68).

Suppose next that $0 < Y < 1$. We consider the rectangular path with vertices

$$c - \mathrm{i}T_1, \quad c + \mathrm{i}T_2, \quad u + \mathrm{i}T_2, \quad u - \mathrm{i}T_1,$$

where $u > 0$, followed in the clockwise direction. Applying Cauchy's integral theorem, we obtain

$$I(Y,T_1,T_2) = \frac{1}{2\pi\mathrm{i}} \int_{u-\mathrm{i}T_1}^{u+\mathrm{i}T_2} \frac{Y^s}{s}\,\mathrm{d}s + \frac{1}{2\pi\mathrm{i}} \int_{c-\mathrm{i}T_1}^{u-\mathrm{i}T_1} \frac{Y^s}{s}\,\mathrm{d}s - \frac{1}{2\pi\mathrm{i}} \int_{c+\mathrm{i}T_2}^{u+\mathrm{i}T_2} \frac{Y^s}{s}\,\mathrm{d}s. \tag{69}$$

On the vertical edge $\sigma = u$, we have $|Y^s/s| \le Y^u/u$, and so

$$\left| \frac{1}{2\pi\mathrm{i}} \int_{u-\mathrm{i}T_1}^{u+\mathrm{i}T_2} \frac{Y^s}{s}\,\mathrm{d}s \right| \le \frac{(T_1 + T_2)Y^u}{2\pi u}. \tag{70}$$

On the horizontal edge $t = -T_1$, we have $|Y^s/s| \le Y^\sigma/T_1$, and so

$$\left| \frac{1}{2\pi\mathrm{i}} \int_{c-\mathrm{i}T_1}^{u-\mathrm{i}T_1} \frac{Y^s}{s}\,\mathrm{d}s \right| \le \frac{1}{2\pi} \int_c^\infty \frac{Y^\sigma}{T_1}\,\mathrm{d}\sigma = -\frac{Y^c}{2\pi T_1 \log Y} = \frac{Y^c}{2\pi T_1 |\log Y|}. \tag{71}$$

On the horizontal edge $t = T_2$, we have $|Y^s/s| \le Y^\sigma/T_2$, and so

$$\left| \frac{1}{2\pi\mathrm{i}} \int_{c+\mathrm{i}T_2}^{u+\mathrm{i}T_2} \frac{Y^s}{s}\,\mathrm{d}s \right| \le \frac{1}{2\pi} \int_c^\infty \frac{Y^\sigma}{T_2}\,\mathrm{d}\sigma = -\frac{Y^c}{2\pi T_2 \log Y} = \frac{Y^c}{2\pi T_2 |\log Y|}. \tag{72}$$

Combining (69)–(72), we obtain

$$|I(Y,T_1,T_2)| \le \frac{(T_1 + T_2)Y^u}{2\pi u} + \frac{Y^c}{2\pi T_1 |\log Y|} + \frac{Y^c}{2\pi T_2 |\log Y|}.$$

Since the left hand side is independent of u, it follows on letting $u \to \infty$ that

$$|I(Y,T_1,T_2)| \le \frac{Y^c}{2\pi T_1 |\log Y|} + \frac{Y^c}{2\pi T_2 |\log Y|}.$$

Letting $T_1, T_2 \to \infty$ gives (61). Letting $T = T_1 = T_2$ gives one of the inequalities in (62). To deduce the other inequality, we use the circular arc $A^+(c,T)$ centred at $s = 0$ and passing from $c - \mathrm{i}T$ to $c + \mathrm{i}T$ on the right of the line $\sigma = c$, as in the proof of Theorem 5F. Then Cauchy's integral theorem gives

$$I(Y,T) - I(Y) = I(Y,T) = \frac{1}{2\pi\mathrm{i}} \int_{A^+(c,T)} \frac{Y^s}{s}\,\mathrm{d}s. \tag{73}$$

On the circular arc $A^+(c,T)$, we have $|Y^s/s| \le Y^c/R$, where $R=(c^2+T^2)^{1/2}$ is the radius of $A^+(c,T)$. It follows that

$$\left|\frac{1}{2\pi \mathrm{i}}\int_{A^+(c,T)}\frac{Y^s}{s}\,\mathrm{d}s\right| \le \frac{1}{2\pi}\frac{Y^c}{R}2\pi R = Y^c. \tag{74}$$

The inequality $|I(Y)-I(Y,T)|\le Y^c$ now follows on combining (73) and (74).

Suppose finally that $Y=1$. Then

$$I(1,T)=\frac{1}{2\pi \mathrm{i}}\int_{c-\mathrm{i}T}^{c+\mathrm{i}T}\frac{\mathrm{d}s}{s}=\frac{1}{2\pi}\int_{-T}^{T}\frac{\mathrm{d}t}{c+\mathrm{i}t}=\frac{1}{2\pi}\int_{-T}^{T}\frac{c-\mathrm{i}t}{c^2+t^2}\,\mathrm{d}t.$$

Note that the imaginary part of the integrand of the last integral is an odd function, while the real part is an even function. It follows that

$$I(1,T)=\frac{1}{\pi}\int_0^T\frac{c}{c^2+t^2}\,\mathrm{d}t \to \frac{1}{2} \quad \text{as } T\to\infty.$$

On the other hand, we have

$$|I(1)-I(1,T)|=\left|\frac{1}{2}-I(1,T)\right|=\frac{1}{\pi}\int_T^\infty\frac{c}{c^2+t^2}\,\mathrm{d}t \le \frac{1}{\pi}\int_T^\infty\frac{c}{t^2}\,\mathrm{d}t = c(\pi T)^{-1}.$$

This completes the proof. ○

In view of the identity (60), it is now reasonable to compare $\psi_0(X)$ with the sum

$$\begin{aligned}\psi(X,T)&=\sum_{n=1}^\infty \Lambda(n) I\left(\frac{X}{n},T\right)=\frac{1}{2\pi \mathrm{i}}\sum_{n=1}^\infty\Lambda(n)\int_{c-\mathrm{i}T}^{c+\mathrm{i}T}\frac{X^s}{sn^s}\,\mathrm{d}s\\ &=\frac{1}{2\pi \mathrm{i}}\int_{c-\mathrm{i}T}^{c+\mathrm{i}T}\left(\sum_{n=1}^\infty\frac{\Lambda(n)}{n^s}\right)\frac{X^s}{s}\,\mathrm{d}s=\frac{1}{2\pi \mathrm{i}}\int_{c-\mathrm{i}T}^{c+\mathrm{i}T}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{X^s}{s}\,\mathrm{d}s.\end{aligned}$$

Using Theorem 6U, we see that

$$\begin{aligned}|\psi_0(X)-\psi(X,T)| &\le \sum_{n=1}^\infty \Lambda(n)\left|I\left(\frac{X}{n}\right)-I\left(\frac{X}{n},T\right)\right|\\ &\le \sum_{\substack{n=1\\ n\ne X}}^\infty \Lambda(n)\left(\frac{X}{n}\right)^c \min\left\{1,\left(\pi T\left|\log\frac{X}{n}\right|\right)^{-1}\right\}+c(\pi T)^{-1}\Lambda(X),\end{aligned} \tag{75}$$

with the understanding that the last term is present only if X is a prime power.

Let $X>\mathrm{e}$ be given and fixed. We shall choose

$$c=1+(\log X)^{-1}, \qquad \text{so that} \qquad X^c=\mathrm{e}X. \tag{76}$$

Note that $c<2$. We can write

$$\sum_{\substack{n=1\\ n\ne X}}^\infty \Lambda(n)\left(\frac{X}{n}\right)^c \min\left\{1,\left(\pi T\left|\log\frac{X}{n}\right|\right)^{-1}\right\}=\sum\nolimits_1+\sum\nolimits_2+\sum\nolimits_3+\sum\nolimits_4, \tag{77}$$

where

$$\sum\nolimits_1 = \sum_{n\le 3X/4} \Lambda(n)\left(\frac{X}{n}\right)^c \min\left\{1, \left(\pi T\left|\log\frac{X}{n}\right|\right)^{-1}\right\},$$

$$\sum\nolimits_2 = \sum_{3X/4<n<X} \Lambda(n)\left(\frac{X}{n}\right)^c \min\left\{1, \left(\pi T\left|\log\frac{X}{n}\right|\right)^{-1}\right\},$$

$$\sum\nolimits_3 = \sum_{X<n<4X/3} \Lambda(n)\left(\frac{X}{n}\right)^c \min\left\{1, \left(\pi T\left|\log\frac{X}{n}\right|\right)^{-1}\right\},$$

$$\sum\nolimits_4 = \sum_{n\ge 4X/3} \Lambda(n)\left(\frac{X}{n}\right)^c \min\left\{1, \left(\pi T\left|\log\frac{X}{n}\right|\right)^{-1}\right\}.$$

Suppose first of all that $n \le 3X/4$ or $n \ge 4X/3$. Then it is easy to see that $|\log(X/n)| \ge \log(4/3)$. In view of (76) and (37), we have

(78) $$\sum\nolimits_1 + \sum\nolimits_4 \ll XT^{-1}\sum_{n=1}^{\infty}\frac{\Lambda(n)}{n^c} = XT^{-1}\left(-\frac{\zeta'(c)}{\zeta(c)}\right) \ll XT^{-1}\log X.$$

Suppose next that $3X/4 < n < X$. In this case, let X_1 denote the largest prime power less than X. We may assume, without loss of generality, that $3X/4 < X_1 < X$, for otherwise we have $\sum_2 = 0$. For the term $n = X_1$, we have

$$\log\frac{X}{n} = -\log\frac{X_1}{X} = -\log\left(1-\frac{X-X_1}{X}\right) \ge \frac{X-X_1}{X}.$$

It follows that the contribution of this term to the sum $\sum_2$ is

$$\ll \Lambda(X_1)\min\left\{1, \frac{X}{T(X-X_1)}\right\} \ll (\log X)\min\left\{1, \frac{X}{T(X-X_1)}\right\}.$$

The other terms form a subcollection of $n = X_1 - m$, where $0 < m < X/4$, and we have

$$\log\frac{X}{n} \ge \log\frac{X_1}{n} = -\log\frac{n}{X_1} = -\log\left(1-\frac{m}{X_1}\right) \ge \frac{m}{X_1}.$$

It follows that the contribution of these terms to the sum $\sum_2$ is

$$\ll \sum_{0<m<X/4} \Lambda(X_1-m)\frac{X_1}{Tm} \ll XT^{-1}\sum_{0<m<X/4}\frac{\Lambda(X_1-m)}{m} \ll XT^{-1}(\log X)^2.$$

Hence

(79) $$\sum\nolimits_2 \ll XT^{-1}(\log X)^2 + (\log X)\min\left\{1, \frac{X}{T\langle X\rangle}\right\},$$

where $\langle X\rangle$ denotes the distance of X to the nearest prime power.

Suppose finally that $X < n < 4X/3$. In this case, let X_2 denote the smallest prime power greater than X. We may assume, without loss of generality, that $X < X_2 < 4X/3$, for otherwise we have $\sum_3 = 0$. For the term $n = X_2$, we have

$$\left|\log\frac{X}{n}\right| = \log\frac{X_2}{X} = \log\left(1+\frac{X_2-X}{X}\right) \ge \frac{X_2-X}{X}.$$

It follows that the contribution of this term to the sum $\sum_3$ is

$$\ll \Lambda(X_2)\min\left\{1,\frac{X}{T(X_2-X)}\right\} \ll (\log X)\min\left\{1,\frac{X}{T(X_2-X)}\right\}.$$

The other terms form a subcollection of $n = X_2 + m$, where $0 < m < X/3$, and we have

$$\left|\log\frac{X}{n}\right| \geq \left|\log\frac{X_2}{n}\right| = \log\frac{n}{X_2} = \log\left(1+\frac{m}{X_2}\right) \geq \frac{m}{X_2}.$$

It follows that the contribution of these terms to the sum $\sum_3$ is

$$\ll \sum_{0<m<X/3} \Lambda(X_2+m)\frac{X_2}{Tm} \ll XT^{-1} \sum_{0<m<X/3} \frac{\Lambda(X_2+m)}{m} \ll XT^{-1}(\log X)^2.$$

Hence

(80) $$\sum\nolimits_3 \ll XT^{-1}(\log X)^2 + (\log X)\min\left\{1,\frac{X}{T\langle X\rangle}\right\}.$$

Combining (77)–(80), we conclude that

$$\sum_{\substack{n=1\\ n\neq X}}^{\infty} \Lambda(n)\left(\frac{X}{n}\right)^c \min\left\{1,\left(\pi T\left|\log\frac{X}{n}\right|\right)^{-1}\right\} \ll XT^{-1}(\log X)^2 + (\log X)\min\left\{1,\frac{X}{T\langle X\rangle}\right\},$$

and so it follows from (75) and (76) that

(81) $$|\psi_0(X) - \psi(X,T)| \ll XT^{-1}(\log X)^2 + (\log X)\min\left\{1,\frac{X}{T\langle X\rangle}\right\} \quad \text{if } c = 1 + (\log X)^{-1}.$$

We now need to study the term $\psi(X,T)$.

Consider a rectangular path with vertices

$$c - \mathrm{i}T, \quad c + \mathrm{i}T, \quad -U + \mathrm{i}T, \quad -U - \mathrm{i}T,$$

followed in the anticlockwise direction and where U and T are chosen carefully to satisfy the following two conditions:

(i) U is a large odd positive integer to ensure that the left edge of the rectangular path passes halfway between two consecutive trivial zeros of $\zeta(s)$.

(ii) T is chosen so that $|\gamma - T| \gg (\log T)^{-1}$ for any zero $\rho = \beta \pm \mathrm{i}\gamma$ of $\zeta(s)$ in the critical strip. This is clearly possible by varying T by a bounded amount, in view of Theorem 6R.

Applying Cauchy's residue theorem, we obtain

(82) $$\psi(X,T) = \frac{1}{2\pi\mathrm{i}}\int_{-U+\mathrm{i}T}^{c+\mathrm{i}T}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{X^s}{s}\,\mathrm{d}s + \frac{1}{2\pi\mathrm{i}}\int_{-U-\mathrm{i}T}^{-U+\mathrm{i}T}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{X^s}{s}\,\mathrm{d}s$$
$$- \frac{1}{2\pi\mathrm{i}}\int_{-U-\mathrm{i}T}^{c-\mathrm{i}T}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{X^s}{s}\,\mathrm{d}s + X - \frac{\zeta'(0)}{\zeta(0)} - \sum_{\substack{\rho\\ |\gamma|<T}} \frac{X^\rho}{\rho} + \sum_{m=1}^{[U/2]} \frac{X^{-2m}}{2m}.$$

We study first the integral

$$\frac{1}{2\pi \mathrm{i}}\int_{-U+\mathrm{i}T}^{c+\mathrm{i}T}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{X^s}{s}\,\mathrm{d}s=\frac{1}{2\pi \mathrm{i}}\int_{-U+\mathrm{i}T}^{-1+\mathrm{i}T}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{X^s}{s}\,\mathrm{d}s+\frac{1}{2\pi \mathrm{i}}\int_{-1+\mathrm{i}T}^{c+\mathrm{i}T}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{X^s}{s}\,\mathrm{d}s. \tag{83}$$

For every $s=\sigma+\mathrm{i}T$, where $-1\le\sigma\le 2$, we have

$$\left|\frac{1}{s-\rho}\right|\le\frac{1}{|T-\gamma|}\ll\log T,$$

in view of condition (ii). This, combined with Theorems 6R and 6T, gives the bound

$$\left|\frac{\zeta'(s)}{\zeta(s)}\right|\ll(\log T)^2.$$

It follows on recalling $X^c=\mathrm{e}X$ and $c<2$ that

$$\frac{1}{2\pi \mathrm{i}}\int_{-1+\mathrm{i}T}^{c+\mathrm{i}T}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{X^s}{s}\,\mathrm{d}s\ll\frac{(\log T)^2}{T}\int_{-\infty}^{c}X^\sigma\,\mathrm{d}\sigma\ll\frac{X(\log T)^2}{T\log X}. \tag{84}$$

To study the first integral on the right hand side of (83), we need the following estimate.

THEOREM 6V. *Suppose that $\mathfrak{Re}\, s\le -1$ and $|s+2m|\ge 1$ for every $m\in\mathbb{N}$. Then*

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

Proof. We start with the formula

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

Combining this with the functional equation (1), we obtain the functional equation in unsymmetric form

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

On taking logarithmic derivatives, we have

$$\frac{\zeta'(s)}{\zeta(s)}=C-\frac{\Gamma'(1-s)}{\Gamma(1-s)}-\frac{\zeta'(1-s)}{\zeta(1-s)}+\frac{\pi}{2}\tan\frac{\pi(1-s)}{2},$$

where C is a constant. Note next that if $\mathfrak{Re}\, s\le -1$, then $\mathfrak{Re}(1-s)\ge 2$. Furthermore, if $|s+2m|\ge 1$, then $|(1-s)-(2m+1)|\ge 1$, so that $\tan\frac{1}{2}\pi(1-s)$ is bounded. The result now follows on noting that

$$\left|\frac{\Gamma'(1-s)}{\Gamma(1-s)}\right|=O(\log|1-s|)=O(\log 2|s|)\qquad\text{and}\qquad\left|\frac{\zeta'(1-s)}{\zeta(1-s)}\right|=O(1)$$

for $\mathfrak{Re}(1-s)\ge 2$. ◯

It now follows from Theorem 6V that

$$\frac{1}{2\pi \mathrm{i}}\int_{-U+\mathrm{i}T}^{-1+\mathrm{i}T}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{X^s}{s}\,\mathrm{d}s\ll\frac{\log 2T}{T}\int_{-U}^{-1}X^\sigma\,\mathrm{d}\sigma\ll\frac{\log T}{TX\log X}. \tag{85}$$

Combining (83)–(85), we obtain

$$\frac{1}{2\pi \mathrm{i}}\int_{-U+\mathrm{i}T}^{c+\mathrm{i}T}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{X^s}{s}\,\mathrm{d}s \ll \frac{X(\log T)^2}{T\log X}. \tag{86}$$

A similar consideration gives the analogous estimate

$$\frac{1}{2\pi \mathrm{i}}\int_{-U-\mathrm{i}T}^{c-\mathrm{i}T}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{X^s}{s}\,\mathrm{d}s \ll \frac{X(\log T)^2}{T\log X}. \tag{87}$$

Note also that Theorem 6V also leads to the estimate

$$\frac{1}{2\pi \mathrm{i}}\int_{-U-\mathrm{i}T}^{-U+\mathrm{i}T}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{X^s}{s}\,\mathrm{d}s \ll \frac{\log 2U}{U}\int_{-T}^{T} X^{-U}\,\mathrm{d}t \ll \frac{T\log U}{UX^U}. \tag{88}$$

Combining (82) and (86)–(88), we obtain

$$\psi(X,T) = X - \frac{\zeta'(0)}{\zeta(0)} - \sum_{\substack{\rho\\ |\gamma|<T}} \frac{X^\rho}{\rho} + \sum_{m=1}^{[U/2]} \frac{X^{-2m}}{2m} + O\left(\frac{X(\log T)^2}{T\log X}\right) + O\left(\frac{T\log U}{UX^U}\right),$$

valid for arbitrarily large values of U. Keeping T fixed and letting $U\to\infty$, we deduce that

$$\psi(X,T) = X - \frac{\zeta'(0)}{\zeta(0)} - \sum_{\substack{\rho\\ |\gamma|<T}} \frac{X^\rho}{\rho} + \sum_{m=1}^{\infty} \frac{X^{-2m}}{2m} + O\left(\frac{X(\log T)^2}{T\log X}\right).$$

Combining this with (81), we obtain

$$\psi_0(X) = X - \frac{\zeta'(0)}{\zeta(0)} - \sum_{\substack{\rho\\ |\gamma|<T}} \frac{X^\rho}{\rho} + \sum_{m=1}^{\infty} \frac{X^{-2m}}{2m} + R(X,T), \tag{89}$$

where the error term $R(X,T)$ satisfies the bound

$$R(X,T) \ll \frac{X(\log XT)^2}{T} + (\log X)\min\left\{1, \frac{X}{T\langle X\rangle}\right\}. \tag{90}$$

However, we have to recognize that (89) and (90) have been established under a restriction on the value of T which has made it necessary for us to vary its value by a bounded amount. This has the effect of changing the number of terms in the sum over ρ in (89) by $O(\log T)$ terms, in view of Theorem 6R, and each such term clearly contributes at most $O(XT^{-1})$. It follows that the error incurred on relaxing the restriction on T is at most $O(XT^{-1}\log T)$ which is easily absorbed in the error estimate (90). Hence (89) and (90) remain valid for all large values of T.

Note now that for fixed X, we have $R(X,T)\to 0$ as $T\to\infty$. It follows that

$$\psi_0(X) = X - \frac{\zeta'(0)}{\zeta(0)} - \sum_{\rho} \frac{X^\rho}{\rho} + \sum_{m=1}^{\infty} \frac{X^{-2m}}{2m} = X - \frac{\zeta'(0)}{\zeta(0)} - \sum_{\rho} \frac{X^\rho}{\rho} - \log\left(1 - \frac{1}{X^2}\right).$$

6.9. The Prime Number Theorem

The estimates (89) and (90) give us another opportunity to establish the Prime number theorem, in the form $\psi_0(X) \sim X$ as $X \to \infty$. Clearly we need a good estimate for the sum

$$\sum_{\substack{\rho \\ |\gamma|<T}} \frac{X^\rho}{\rho}, \tag{91}$$

as well as an appropriate estimate for the error term $R(X,T)$, by choosing suitably large values for the parameter T.

To obtain a good estimate for the sum (91), the idea here is to make use of the zero-free region given by Theorem 6F. A consequence of this zero-free region is that for every zero $\rho = \beta + \mathrm{i}\gamma$ of $\zeta(s)$ in the critical strip with $|\gamma| < T$, where $T > 2$, we have

$$\beta \le 1 - \frac{c}{\log|\gamma|} \le 1 - \frac{c}{\log T},$$

where c is a positive absolute constant. It follows that for any such zero ρ, we have

$$|X^\rho| = X^\beta \le X\mathrm{e}^{-c(\log X)/(\log T)},$$

and so

$$\sum_{\substack{\rho \\ |\gamma|<T}} \frac{X^\rho}{\rho} \ll X\mathrm{e}^{-c(\log X)/(\log T)} \sum_{\substack{\rho \\ 0<\gamma<T}} \frac{1}{\gamma}.$$

Note next that

$$\sum_{\substack{\rho \\ 0<\gamma<T}} \frac{1}{\gamma} = \int_0^T t^{-1}\,\mathrm{d}N(t) = \frac{N(T)}{T} + \int_0^T t^{-2} N(t)\,\mathrm{d}t, \tag{92}$$

where $N(T)$ denotes the number of zeros $\rho = \beta + \mathrm{i}\gamma$ of $\zeta(s)$ in the critical strip with $0 < \gamma < T$. But $N(T) = O(T\log T)$ for large T by Theorem 6D, so the sum (92) is $\ll (\log T)^2$. Hence

$$\sum_{\substack{\rho \\ |\gamma|<T}} \frac{X^\rho}{\rho} \ll X(\log T)^2\mathrm{e}^{-c(\log X)/(\log T)}. \tag{93}$$

Combining (89), (90) and (93), we now have

$$\psi_0(X) - X \ll \frac{X(\log XT)^2}{T} + (\log X)\min\left\{1, \frac{X}{T\langle X\rangle}\right\} + X(\log T)^2\mathrm{e}^{-c(\log X)/(\log T)}.$$

We may assume, without loss of generality, that X is an integer, so that $\langle X\rangle \ge 1$. Then

$$\psi_0(X) - X \ll \frac{X(\log XT)^2}{T} + X(\log T)^2\mathrm{e}^{-c(\log X)/(\log T)}.$$

We now choose T to satisfy $(\log T)^2 = \log X$, so that $T^{-1} = \mathrm{e}^{-(\log X)^{1/2}}$. Then

$$\psi_0(X) - X \ll X(\log X)^2\mathrm{e}^{-(\log X)^{1/2}} + X(\log X)\mathrm{e}^{-c(\log X)^{-1/2}} \ll X\mathrm{e}^{-c'(\log X)^{-1/2}},$$

where $c' < \min\{1, c\}$ is a positive absolute constant.

Problems for Chapter 6

1. Prove Theorem 6R.
2. Prove Theorem 6S.