# ELEMENTARY AND ANALYTIC NUMBER THEORY 

W W L CHEN

## (C) W W L Chen, 1981 .

This work is available free, in the hope that it will be useful.
Any part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, with or without permission from the author.

## Chapter 1

## ARITHMETIC FUNCTIONS

### 1.1. Introduction

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

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

THEOREM 1.1. Suppose that the function $f: \mathbb{N} \rightarrow \mathbb{C}$ is multiplicative. Then the function $g: \mathbb{N} \rightarrow \mathbb{C}$, defined by

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

for every $n \in \mathbb{N}$, is multiplicative.
Here the summation $\sum_{m \mid n}$ denotes a sum over all positive divisors $m$ of $n$.
Proof of Theorem 1.1. 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 $u v$ is a positive divisor of $a b$. On the other hand, it is well-known that every positive divisor $m$ of $a b$ can be expressed uniquely in the form $m=u v$, where $u$ is a positive divisor of $a$ and $v$ is a positive divisor of $b$. It follows that

$$
\begin{aligned}
g(a b) & =\sum_{m \mid a b} f(m)=\sum_{u \mid a} \sum_{v \mid b} f(u v)=\sum_{u \mid a} \sum_{v \mid b} f(u) f(v) \\
& =\left(\sum_{u \mid a} f(u)\right)\left(\sum_{v \mid b} f(v)\right)=g(a) g(b)
\end{aligned}
$$

[^0]$\qquad$
$\qquad$

### 1.2. The Divisor Function

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

$$
\begin{equation*}
d(n)=\sum_{m \mid n} 1 \tag{1}
\end{equation*}
$$

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 $\sigma: \mathbb{N} \rightarrow \mathbb{C}$ by writing

$$
\begin{equation*}
\sigma(n)=\sum_{m \mid n} m \tag{2}
\end{equation*}
$$

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 1.2. Suppose that $n \in \mathbb{N}$ and that $n=p_{1}^{u_{1}} \ldots p_{r}^{u_{r}}$ is the canonical decomposition of $n$. Then

$$
d(n)=\left(1+u_{1}\right) \ldots\left(1+u_{r}\right) \quad \text { and } \quad \sigma(n)=\frac{p_{1}^{u_{1}+1}-1}{p_{1}-1} \ldots \frac{p_{r}^{u_{r}+1}-1}{p_{r}-1} .
$$

Proof. Every positive divisor $m$ of $n$ is of the form $m=p_{1}^{v_{1}} \ldots p_{r}^{v_{r}}$, where for every $j=1, \ldots, r$, the integer $v_{j}$ satisfies $0 \leq v_{j} \leq u_{j}$. It follows from (1) that $d(n)$ is the number of choices for the $r$-tuple $\left(v_{1}, \ldots, v_{r}\right)$. Hence

$$
d(n)=\sum_{v_{1}=0}^{u_{1}} \ldots \sum_{v_{r}=0}^{u_{r}} 1=\left(1+u_{1}\right) \ldots\left(1+u_{r}\right) .
$$

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

$$
\sigma(n)=\sum_{v_{1}=0}^{u_{1}} \ldots \sum_{v_{r}=0}^{u_{r}} p_{1}^{v_{1}} \ldots p_{r}^{v_{r}}=\left(\sum_{v_{1}=0}^{u_{1}} p_{1}^{v_{1}}\right) \ldots\left(\sum_{v_{r}=0}^{u_{r}} p_{r}^{v_{r}}\right) .
$$

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

$$
\sum_{v_{j}=0}^{u_{j}} p_{j}^{v_{j}}=1+p_{j}+p_{j}^{2}+\ldots+p_{j}^{u_{j}}=\frac{p_{j}^{u_{j}+1}-1}{p_{j}-1}
$$

The second result follows.

The result below is a simple deduction from Theorem 1.2.
THEOREM 1.3. The arithmetic functions $d: \mathbb{N} \rightarrow \mathbb{C}$ and $\sigma: \mathbb{N} \rightarrow \mathbb{C}$ are both multiplicative.
Natural numbers $n \in \mathbb{N}$ where $\sigma(n)=2 n$ 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.

Example. 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 1.4. (Euclid-Euler) Suppose that $m \in \mathbb{N}$. If $2^{m}-1$ is a prime, then the number $2^{m-1}\left(2^{m}-1\right)$ is an even perfect number. Furthermore, there are no other even perfect numbers.

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

$$
\left(2^{m-1}, 2^{m}-1\right)=1
$$

It follows from Theorems 1.2 and 1.3 that

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

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 1.2 , we have

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

so that

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

Note that $\sigma(u)$ and $u$ are integers and $\sigma(u)>u$. Hence $u /\left(2^{m}-1\right) \in \mathbb{N}$ and is a divisor of $u$. Since $m>1$, we have $2^{m}-1>1$, and so $u /\left(2^{m}-1\right) \neq u$. 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 $u /\left(2^{m}-1\right)=1$, so that $u=2^{m}-1$.

We are interested in the behaviour of $d(n)$ and $\sigma(n)$ as $n \rightarrow \infty$. 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 1.5. For any fixed real number $c>0$, the inequality $d(n) \ll(\log n)^{c}$ as $n \rightarrow \infty$ 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

$$
n=\left(p_{1} \ldots p_{\ell+1}\right)^{m}
$$

Then in view of Theorem 1.2, we have

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

where the positive constant

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

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 1.6. For any fixed real number $\epsilon>0$, we have $d(n) \ll_{\epsilon} n^{\epsilon}$ as $n \rightarrow \infty$.
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 1.2 that

$$
\frac{d(n)}{n^{\epsilon}}=\frac{\left(1+u_{1}\right)}{p_{1}^{\epsilon u_{1}}} \ldots \frac{\left(1+u_{r}\right)}{p_{r}^{\epsilon u_{r}}}
$$

$\qquad$

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

$$
p_{j}^{\epsilon u_{j}} \geq 2^{\epsilon u_{j}}=\mathrm{e}^{\epsilon u_{j} \log 2}>1+\epsilon u_{j} \log 2>\left(1+u_{j}\right) \epsilon \log 2
$$

so that

$$
\frac{\left(1+u_{j}\right)}{p_{j}^{\epsilon u_{j}}}<\frac{1}{\epsilon \log 2}
$$

On the other hand, if $p_{j} \geq 2^{1 / \epsilon}$, then $p_{j}^{\epsilon} \geq 2$, and so

$$
\frac{\left(1+u_{j}\right)}{p_{j}^{\epsilon u_{j}}} \leq \frac{1+u_{j}}{2^{u_{j}}} \leq 1
$$

It follows that

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

a positive constant depending only on $\epsilon$.
We see from Theorems 1.5 and 1.6 and the fact that $d(n)=2$ infinitely often that the magnitude of $d(n)$ fluctuates a great deal as $n \rightarrow \infty$. 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

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

THEOREM 1.7. (Dirichlet) As $X \rightarrow \infty$, we have

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

Here $\gamma$ is Euler's constant and is defined by

$$
\gamma=\lim _{Y \rightarrow \infty}\left(\sum_{n \leq Y} \frac{1}{n}-\log Y\right)=0.5772156649 \ldots
$$

Remark. It is an open problem in mathematics to determine whether Euler's constant $\gamma$ is rational or irrational.

The proof of Theorem 1.7 depends on the following intermediate result.
THEOREM 1.8. As $Y \rightarrow \infty$, we have

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

Proof. As $Y \rightarrow \infty$, we have

$$
\begin{aligned}
\sum_{n \leq Y} \frac{1}{n} & =\sum_{n \leq Y}\left(\frac{1}{Y}+\int_{n}^{Y} \frac{1}{u^{2}} \mathrm{~d} u\right)=\frac{[Y]}{Y}+\sum_{n \leq Y} \int_{n}^{Y} \frac{1}{u^{2}} \mathrm{~d} u \\
& =\frac{[Y]}{Y}+\int_{1}^{Y} \frac{1}{u^{2}}\left(\sum_{n \leq u} 1\right) \mathrm{d} u=\frac{[Y]}{Y}+\int_{1}^{Y} \frac{[u]}{u^{2}} \mathrm{~d} u \\
& =\frac{[Y]}{Y}+\int_{1}^{Y} \frac{1}{u} \mathrm{~d} u-\int_{1}^{Y} \frac{u-[u]}{u^{2}} \mathrm{~d} u \\
& =\log Y+1+O\left(\frac{1}{Y}\right)-\int_{1}^{\infty} \frac{u-[u]}{u^{2}} \mathrm{~d} u+\int_{Y}^{\infty} \frac{u-[u]}{u^{2}} \mathrm{~d} u \\
& =\log Y+\left(1-\int_{1}^{\infty} \frac{u-[u]}{u^{2}} \mathrm{~d} u\right)+O\left(\frac{1}{Y}\right)
\end{aligned}
$$

It is a simple exercise to show that

$$
1-\int_{1}^{\infty} \frac{u-[u]}{u^{2}} \mathrm{~d} u=\gamma
$$

Proof of Theorem 1.7. As $X \rightarrow \infty$, we have

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

We next turn our attention to the study of the behaviour of $\sigma(n)$ as $n \rightarrow \infty$. 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 1.6 that for any fixed real number $\epsilon>0$, we have

$$
\sigma(n) \leq n d(n) \ll_{\epsilon} n^{1+\epsilon}
$$

as $n \rightarrow \infty$.
In fact, it is rather easy to prove a slightly stronger result.
THEOREM 1.9. We have $\sigma(n) \ll n \log n$ as $n \rightarrow \infty$.
Proof. As $n \rightarrow \infty$, we have

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

As in the case of $d(n)$, the magnitude of $\sigma(n)$ fluctuates a great deal as $n \rightarrow \infty$. 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 1.7, we have the following result.

THEOREM 1.10. As $X \rightarrow \infty$, we have

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

Proof. As $X \rightarrow \infty$, we have

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

### 1.3. The Möbius Function

We define the Möbius function $\mu: \mathbb{N} \rightarrow \mathbb{C}$ 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. For a more detailed discussion, see Chapters 4,5 and 6 . At this point, it suffices to remark that the Möbius function is defined so that if we formally multiply the two series

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

then the product is identically equal to 1 . Heuristically, note that

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

It follows that the product is identically equal to 1 if

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

We shall establish this last fact and study some of its consequences over the next four theorems.
THEOREM 1.11. The Möbius function $\mu: \mathbb{N} \rightarrow \mathbb{C}$ is multiplicative.
Proof. Suppose that $a, b \in \mathbb{N}$ and $(a, b)=1$. If $a$ or $b$ is not squarefree, then neither is $a b$, and so $\mu(a b)=0=\mu(a) \mu(b)$. On the other hand, if both $a$ and $b$ are squarefree, then since $(a, b)=1, a b$ must also be squarefree. Furthermore, the number of prime factors of $a b$ must be the sum of the numbers of prime factors of $a$ and of $b$.

THEOREM 1.12. Suppose that $n \in \mathbb{N}$. Then

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

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

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

for every $n \in \mathbb{N}$. It follows from Theorems 1.1 and 1.11 that $f$ is multiplicative. For $n=1$, the result is trivial. To complete the proof, it therefore suffices to show that $f\left(p^{k}\right)=0$ for every prime $p$ and every $k \in \mathbb{N}$. Indeed,

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

Theorem 1.12 plays the central role in the proof of the following two results which are similar in nature.

THEOREM 1.13. (Möbius Inversion Formula) Given any function $f: \mathbb{N} \rightarrow \mathbb{C}$, suppose that the function $g: \mathbb{N} \rightarrow \mathbb{C}$ is defined by writing

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

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

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

Proof. The second equality is obvious. Also

$$
\begin{aligned}
\sum_{m \mid n} \mu(m) g\left(\frac{n}{m}\right) & =\sum_{m \mid n} \mu(m)\left(\sum_{k \left\lvert\, \frac{n}{m}\right.} f(k)\right)=\sum_{\substack{k, m \\
k m \mid n}} \mu(m) f(k) \\
& =\sum_{k \mid n} f(k)\left(\sum_{m \left\lvert\, \frac{n}{k}\right.} \mu(m)\right)=f(n)
\end{aligned}
$$

in view of Theorem 1.12.
THEOREM 1.14. For any function $g: \mathbb{N} \rightarrow \mathbb{C}$, if the function $f: \mathbb{N} \rightarrow \mathbb{C}$ 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

$$
\begin{aligned}
\sum_{m \mid n} f\left(\frac{n}{m}\right) & =\sum_{m \mid n}\left(\sum_{k \left\lvert\, \frac{n}{m}\right.} \mu\left(\frac{n}{m k}\right) g(k)\right)=\sum_{k \mid n} g(k)\left(\sum_{m \left\lvert\, \frac{n}{k}\right.} \mu\left(\frac{n / k}{m}\right)\right) \\
& =\sum_{k \mid n} g(k)\left(\sum_{m \left\lvert\, \frac{n}{k}\right.} \mu(m)\right)=g(n)
\end{aligned}
$$

in view of Theorem 1.12.
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 1.13 and 1.14. 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

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

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

$$
\sum_{k \left\lvert\, \frac{n}{m}\right.} A(k, m),
$$

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

$$
\sum_{m \left\lvert\, \frac{n}{k}\right.} A(k, m) .
$$

Summing over all positive divisors $k$ of $n$, we obtain

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

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

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

### 1.4. The Euler Function

We define the Euler function $\phi: \mathbb{N} \rightarrow \mathbb{C}$ 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 1.15. For every number $n \in \mathbb{N}$, we have

$$
\sum_{m \mid n} \phi(m)=n
$$

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 $n$,

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

If $x \in \mathcal{B}_{m}$, let $x=m x^{\prime}$. Then $\left(m x^{\prime}, n\right)=m$ if and only if $\left(x^{\prime}, n / m\right)=1$. Also $1 \leq x \leq n$ if and only if $1 \leq x^{\prime} \leq n / m$. Hence

$$
\mathcal{B}_{m}^{\prime}=\left\{x^{\prime}: 1 \leq x^{\prime} \leq n / m \text { and }\left(x^{\prime}, n / m\right)=1\right\}
$$

has the same number of elements as $\mathcal{B}_{m}$. Note now that the number of elements of $\mathcal{B}_{m}^{\prime}$ 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

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

Apply the Möbius inversion formula to the conclusion of Theorem 1.15, we obtain immediately the following result.

THEOREM 1.16. For every number $n \in \mathbb{N}$, we have

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

THEOREM 1.17. The Euler function $\phi: \mathbb{N} \rightarrow \mathbb{C}$ is multiplicative.
Proof. Since the Möbius function $\mu$ is multiplicative, it follows that the function $f: \mathbb{N} \rightarrow \mathbb{C}$, defined by $f(n)=\mu(n) / n$ for every $n \in \mathbb{N}$, is multiplicative. The result now follows from Theorem 1.1.

THEOREM 1.18. Suppose that $n \in \mathbb{N}$ and $n>1$, and that $n=p_{1}^{u_{1}} \ldots p_{r}^{u_{r}}$ is the canonical decomposition of $n$. Then

$$
\phi(n)=n \prod_{j=1}^{r}\left(1-\frac{1}{p_{j}}\right)=\prod_{j=1}^{r} p_{j}^{u_{j}-1}\left(p_{j}-1\right)
$$

Proof. The second equality is trivial. On the other hand, for every prime $p$ and every $u \in \mathbb{N}$, we have by Theorem 1.16 that

$$
\frac{\phi\left(p^{u}\right)}{p^{u}}=\sum_{m \mid p^{u}} \frac{\mu(m)}{m}=1+\frac{\mu(p)}{p}=1-\frac{1}{p}
$$

The result now follows since $\phi$ is multiplicative.
We noe study the magnitude of $\phi(n)$ as $n \rightarrow \infty$. 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 1.19. For every $n \in \mathbb{N}$, we have

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

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 1.2 and 1.18. 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

$$
\phi(n)=n \prod_{j=1}^{r}\left(1-p_{j}^{-1}\right)
$$

Hence

$$
\frac{\sigma(n) \phi(n)}{n^{2}}=\prod_{j=1}^{r}\left(1-p_{j}^{-u_{j}-1}\right)
$$

The upper bound follows at once. On the other hand, the lower bound follows on observing that

$$
\prod_{j=1}^{r}\left(1-p_{j}^{-u_{j}-1}\right) \geq \prod_{p \mid n}\left(1-p^{-2}\right) \geq \prod_{m=2}^{n}\left(1-\frac{1}{m^{2}}\right)=\frac{n+1}{2 n}>\frac{1}{2}
$$

Combining Theorems 1.9 and 1.19, we have the following result.
THEOREM 1.20. We have $\phi(n) \gg n / \log n$ as $n \rightarrow \infty$.
We now consider some average version of the Euler function.
THEOREM 1.21. (Mertens) As $X \rightarrow \infty$, we have

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

Proof. As $X \rightarrow \infty$, we have, by Theorem 1.16, that

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

It remains to show that

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

But

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

in view of Theorem 1.12.

### 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 $f * g: \mathbb{N} \rightarrow \mathbb{C}$ by writing

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

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

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

Furthermore, the arithmetic function $I: \mathbb{N} \rightarrow \mathbb{C}$, 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 1.22. 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

$$
\begin{equation*}
g(1)=\frac{1}{f(1)} \tag{5}
\end{equation*}
$$

and

$$
\begin{equation*}
g(n)=-\frac{1}{f(1)} \sum_{\substack{d \mid n \\ d>1}} f(d) g\left(\frac{n}{d}\right) \tag{6}
\end{equation*}
$$

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 must be unique.

We now describe Theorem 1.12 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 1.23.

(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 1.12. To prove (ii), note that

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

To prove (iii), note that

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

We conclude this chapter by exhibiting some group structure within $\mathcal{A}$ and $\mathcal{M}$.
THEOREM 1.24. The sets $\mathcal{A}^{\prime}=\{f \in \mathcal{A}: f(1) \neq 0\}$ and $\mathcal{M}^{\prime}=\{f \in \mathcal{M}: f(1)=1\}$ 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 1.24. For $\mathcal{A}^{\prime}$, this is now trivial. We now consider $\mathcal{M}^{\prime}$. Clearly $I \in \mathcal{M}^{\prime}$. If $f, g \in \mathcal{M}^{\prime}$ and $(m, n)=1$, then

$$
\begin{aligned}
(f * g)(m n) & =\sum_{d \mid m n} f(d) g\left(\frac{m n}{d}\right)=\sum_{d_{1} \mid m} \sum_{d_{2} \mid n} f\left(d_{1} d_{2}\right) g\left(\frac{m n}{d_{1} d_{2}}\right) \\
& =\left(\sum_{d_{1} \mid m} f\left(d_{1}\right) g\left(\frac{m}{d_{1}}\right)\right)\left(\sum_{d_{2} \mid n} f\left(d_{2}\right) g\left(\frac{n}{d_{2}}\right)\right)=(f * g)(m)(f * g)(n),
\end{aligned}
$$

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

$$
g\left(p^{k}\right)=h\left(p^{k}\right)
$$

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

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

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

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

so that $g$ is an inverse of $f$.

# ELEMENTARY AND ANALYTIC NUMBER THEORY 

W W L CHEN

## (C) W W L Chen, 1981 .

This work is available free, in the hope that it will be useful.
Any part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, with or without permission from the author.

## Chapter 2

## DISTRIBUTION OF PRIMES <br> I : INTRODUCTION

### 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 2.1. 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

$$
S_{1}=\sum_{p \leq X} \frac{1}{p} \quad \text { and } \quad S_{2}=\sum_{p \leq X} \sum_{h=2}^{\infty} \frac{1}{h p^{h}}
$$

[^1]$\qquad$

Since

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

we have

$$
0 \leq S_{2} \leq \sum_{p} \frac{1}{p(p-1)} \leq \sum_{n=2}^{\infty} \frac{1}{n(n-1)}=1
$$

so that $0 \leq S_{2} \leq 1$. On the other hand, as $X \rightarrow \infty$, we have

$$
P_{X}=\prod_{p \leq X}\left(\sum_{h=0}^{\infty} \frac{1}{p^{h}}\right) \geq \sum_{n \leq X} \frac{1}{n} \rightarrow \infty
$$

The result follows.

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

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

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

$$
\lim _{X \rightarrow \infty} \frac{\pi(X) \log X}{X}=1
$$

We shall give an analytic proof of this in Chapter 5 . 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 $X \geq 2$, 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 $\Lambda: \mathbb{N} \rightarrow \mathbb{C}$, 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 2.2. 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

$$
\sum_{m \mid n} \Lambda(m)=\sum_{j=1}^{r} \sum_{v_{j}=1}^{u_{j}} \log p_{j}=\sum_{j=1}^{r} \log p_{j}^{u_{j}}=\log n
$$

THEOREM 2.3. As $X \rightarrow \infty$, 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 2.2 that

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

It therefore suffices to prove that as $X \rightarrow \infty$, we have

$$
\begin{equation*}
\sum_{n \leq X} \log n=X \log X-X+O(\log X) \tag{1}
\end{equation*}
$$

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

$$
\begin{aligned}
\sum_{n \leq X} \log n & =\sum_{2 \leq n \leq X} \log n \geq \int_{1}^{[X]} \log u \mathrm{~d} u \\
& =\int_{1}^{X} \log u \mathrm{~d} u-\int_{[X]}^{X} \log u \mathrm{~d} u \geq \int_{1}^{X} \log u \mathrm{~d} u-\log X .
\end{aligned}
$$

The inequality (1) now follows on noting that

$$
\int_{1}^{X} \log u \mathrm{~d} u=X \log X-X+1
$$

### 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 2.4. There exist positive absolute constants $c_{3}$ and $c_{4}$ such that

$$
\begin{equation*}
\sum_{m \leq X} \Lambda(m) \geq \frac{1}{2} X \log 2 \quad\left(X \geq c_{3}\right) \tag{2}
\end{equation*}
$$

and

$$
\begin{equation*}
\sum_{\frac{X}{2}<m \leq X} \Lambda(m) \leq c_{4} X \quad(X \geq 0) . \tag{3}
\end{equation*}
$$

Proof. If $m \in \mathbb{N}$ satisfies $X / 2<m \leq X$, then clearly $[X / 2 m]=0$. It follows from this and Theorem 2.3 that as $X \rightarrow \infty$, we have

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

Hence there exists a positive absolute constant $c_{5}$ such that for all sufficiently large $X$, we have

$$
\frac{1}{2} X \log 2<\sum_{m \leq X} \Lambda(m)\left(\left[\frac{X}{m}\right]-2\left[\frac{X}{2 m}\right]\right)<c_{5} X
$$

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 $[\alpha]-2[\alpha / 2] \leq 1$. It follows that for all sufficiently large $X$, we have

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

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

$$
\sum_{\frac{x}{2}<m \leq X} \Lambda(m) \leq c_{5} X
$$

The inequality (3) follows easily.
We now state and prove Tchebycheff's theorem.
THEOREM 2.5. (Tchebycheff) There exist positive absolute constants $c_{1}$ and $c_{2}$ such that for every real number $X \geq 2$, we have

$$
c_{1} \frac{X}{\log X}<\pi(X)<c_{2} \frac{X}{\log X} .
$$

Proof. To prove the lower bound, note that

$$
\begin{aligned}
\sum_{m \leq X} \Lambda(m) & =\sum_{\substack{p, n \\
p^{n} \leq X}} \log p=\sum_{p \leq X}(\log p) \sum_{1 \leq n \leq\left[\frac{\log X}{\log p}\right]} 1 \\
& =\sum_{p \leq X}(\log p)\left[\frac{\log X}{\log p}\right] \leq \pi(X) \log X
\end{aligned}
$$

It follows from (2) that

$$
\pi(X) \geq \frac{X \log 2}{2 \log X} \quad\left(X \geq c_{3}\right)
$$

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

$$
\sum_{X^{1 / 2}<p \leq X} \log p \leq \sum_{j=0}^{k} \sum_{\frac{X}{2 j+1}<p \leq \frac{X}{2 j}} \log p \leq c_{4} X \sum_{j=0}^{k} 2^{-j}<2 c_{4} X,
$$

so that

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

whence

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

for a suitable $c_{2}$.

### 2.4. Some Results of Mertens

We conclude this chapter by obtaining an improvement of Theorem 2.1.
THEOREM 2.6. (Mertens) As $X \rightarrow \infty$, we have

$$
\begin{gather*}
\sum_{m \leq X} \frac{\Lambda(m)}{m}=\log X+O(1)  \tag{4}\\
\sum_{p \leq X} \frac{\log p}{p}=\log X+O(1), \tag{5}
\end{gather*}
$$

and

$$
\begin{equation*}
\sum_{p \leq X} \frac{1}{p}=\log \log X+O(1) \tag{6}
\end{equation*}
$$

Proof. Recall Theorem 2.3. As $X \rightarrow \infty$, 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 $X \rightarrow \infty$, we have

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

It follows from (3) that

$$
\sum_{m \leq X} \Lambda(m) \leq \sum_{j=0}^{\infty} \sum_{\frac{X}{2 j+1}<m \leq \frac{X}{2 j}} \Lambda(m) \leq 2 c_{4} X
$$

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

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

The inequality (4) follows. Next, note that

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

As $X \rightarrow \infty$, we have

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

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

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

Then it follows from (5) that there exists a positive absolute constant $c_{6}$ such that $|T(X)-\log X|<c_{6}$ whenever $X \geq 2$. On the other hand,

$$
\begin{aligned}
\sum_{p \leq X} \frac{1}{p} & =\sum_{p \leq X} \frac{\log p}{p}\left(\frac{1}{\log X}+\int_{p}^{X} \frac{\mathrm{~d} y}{y \log ^{2} y}\right)=\frac{T(X)}{\log X}+\int_{2}^{X} \frac{T(y) \mathrm{d} y}{y \log ^{2} y} \\
& =\frac{T(X)-\log X}{\log X}+\int_{2}^{X} \frac{(T(y)-\log y) \mathrm{d} y}{y \log ^{2} y}+1+\int_{2}^{X} \frac{\mathrm{~d} y}{y \log y}
\end{aligned}
$$

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

$$
\left|\sum_{p \leq X} \frac{1}{p}-\log \log X\right|<\frac{c_{6}}{\log X}+\int_{2}^{X} \frac{c_{6} \mathrm{~d} y}{y \log ^{2} y}+1-\log \log 2=O(1)
$$

The inequality (6) follows.

# ELEMENTARY AND ANALYTIC NUMBER THEORY 

W W L CHEN

## (C) W W L Chen, 1990 .

This work is available free, in the hope that it will be useful.
Any part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, with or without permission from the author.

## Chapter 3

## DIRICHLET SERIES

### 3.1. Convergence Properties

A Dirichlet series is a series of the type

$$
\begin{equation*}
F(s)=\sum_{n=1}^{\infty} f(n) n^{-s} \tag{1}
\end{equation*}
$$

where $f: \mathbb{N} \rightarrow \mathbb{C}$ 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 convergence of Dirichlet series.
THEOREM 3.1. 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 $-\infty \leq \sigma_{0} \leq \sigma_{1} \leq \sigma_{2}<\infty$ 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 $\sigma_{0}-\epsilon<\sigma \leq \sigma_{0}$.
(ii) For every $\eta>0$, the series (1) converges uniformly on the set $\left\{s \in \mathbb{C}: \sigma>\sigma_{1}+\eta\right\}$ and does not converge uniformly on the set $\left\{s \in \mathbb{C}: \sigma>\sigma_{1}-\eta\right\}$.
(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 $\sigma_{2}-\epsilon<\sigma \leq \sigma_{2}$.

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.

[^2]$\qquad$

Proof of Theorem 3.1. Suppose that the series (1) converges for $s=s^{*}=\sigma^{*}+\mathrm{i} t^{*}$. Then $f(n) n^{-s^{*}} \rightarrow 0$ as $n \rightarrow \infty$, so that $\left|f(n) n^{-s^{*}}\right|=O(1)$, and so $|f(n)|=O\left(n^{\sigma^{*}}\right)$. It follows that for every $s \in \mathbb{C}$ with $\sigma>\sigma^{*}+1$, we have

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

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 $\sigma_{0} \leq \sigma_{2}$. 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

$$
\sup \left\{\left|\sum_{n=1}^{N} f(n) n^{-s}-\sum_{n=1}^{\infty} f(n) n^{-s}\right|: \sigma \geq \sigma_{2}+\delta\right\} \leq \sum_{n=N+1}^{\infty}|f(n)| n^{-\sigma_{2}-\delta}<\epsilon
$$

It follows that the series (1) converges uniformly on the set $\left\{s \in \mathbb{C}: \sigma \geq \sigma_{2}+\delta\right\}$. Now let

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

Clearly $\sigma_{0} \leq \sigma_{1} \leq \sigma_{2}+\delta$. Since $\delta>0$ is arbitrary, we must have $\sigma_{0} \leq \sigma_{1} \leq \sigma_{2}$.
A simple consequence of uniform convergence is the following result concerning differentiation term by term.

THEOREM 3.2. For every $s \in \mathbb{C}$ with $\sigma>\sigma_{1}$, the series (1) may be differentiated term by term. In particular, $F^{\prime}(s)$ exists and

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

### 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 3.3. Suppose that

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

where $f: \mathbb{N} \rightarrow \mathbb{C}$ and $g: \mathbb{N} \rightarrow \mathbb{C}$ 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 $\sigma \geq \sigma_{3}$, 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 3.4. Suppose that

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

where $f: \mathbb{N} \rightarrow \mathbb{C}$ 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 $\sigma \geq \sigma_{3}$, 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 $|f(n)|=O\left(n^{\sigma_{3}}\right)$ for all $n \in \mathbb{N}$. Now let $\sigma \geq \sigma_{3}+2$. Then

$$
\begin{equation*}
\sum_{n=N}^{\infty} f(n) n^{-\sigma}=O\left(\sum_{n=N}^{\infty} n^{\sigma_{3}-\sigma}\right) \tag{2}
\end{equation*}
$$

Note next that $y^{\sigma_{3}-\sigma}$ is a decreasing function of $y$, so that

$$
\begin{equation*}
\sum_{n=N}^{\infty} n^{\sigma_{3}-\sigma}=N^{\sigma_{3}-\sigma}+\sum_{n=N+1}^{\infty} n^{\sigma_{3}-\sigma} \leq N^{\sigma_{3}-\sigma}+\int_{N}^{\infty} y^{\sigma_{3}-\sigma} \mathrm{d} y=O\left(N^{\sigma_{3}-\sigma+1}\right) \tag{3}
\end{equation*}
$$

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

$$
\begin{equation*}
\sum_{n=N}^{\infty} f(n) n^{-\sigma}=O\left(N^{\sigma_{3}-\sigma+1}\right) \tag{4}
\end{equation*}
$$

Using (4) with $N=2$, we obtain, for $\sigma \geq \sigma_{3}+2$,

$$
0=F(\sigma)=f(1)+\sum_{n=2}^{\infty} f(n) n^{-\sigma}=f(1)+O\left(2^{\sigma_{3}-\sigma+1}\right) \rightarrow f(1)
$$

as $\sigma \rightarrow+\infty$. 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 $\sigma \geq \sigma_{3}+2$,

$$
0=F(\sigma)=f(M) M^{-\sigma}+\sum_{n=M+1}^{\infty} f(n) n^{-\sigma}=f(M) M^{-\sigma}+O\left((M+1)^{\sigma_{3}-\sigma+1}\right)
$$

so that

$$
0=f(M)+O\left((M+1)^{\sigma_{3}+1}\left(\frac{M}{M+1}\right)^{\sigma}\right) \rightarrow f(M)
$$

as $\sigma \rightarrow+\infty$. 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 3.5. Suppose that for every $j=1,2,3$, we have

$$
F_{j}(s)=\sum_{n=1}^{\infty} f_{j}(n) n^{-s}
$$

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

$$
f_{3}(n)=\sum_{\substack{x, y \\ x y=n}} f_{1}(x) f_{2}(y)=\sum_{x \mid n} f_{1}(x) f_{2}\left(\frac{n}{x}\right)=\sum_{y \mid n} f_{1}\left(\frac{n}{y}\right) f_{2}(y) .
$$

$\qquad$

Then

$$
F_{1}(s) F_{2}(s)=F_{3}(s)
$$

provided that $\sigma>\max \left\{\sigma_{2}^{(1)}, \sigma_{2}^{(2)}\right\}$, 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

$$
\sum_{n=1}^{N} f_{3}(n) n^{-s}=\sum_{\substack{1 \leq x \leq N \\ 1 \leq y \leq N \\ x y \leq N}} f_{1}(x) x^{-s} f_{2}(y) y^{-s}
$$

so that

$$
\begin{aligned}
& \sum_{n=1}^{N} f_{3}(n) n^{-s}-\sum_{x \leq \sqrt{N}} f_{1}(x) x^{-s} \sum_{y \leq \sqrt{N}} f_{2}(y) y^{-s} \\
& \quad=\sum_{\sqrt{N}<x \leq N} f_{1}(x) x^{-s} \sum_{y \leq N / x} f_{2}(y) y^{-s}+\sum_{x \leq \sqrt{N}} f_{1}(x) x^{-s} \sum_{\sqrt{N}<y \leq N / x} f_{2}(y) y^{-s} .
\end{aligned}
$$

It follows that

$$
\begin{align*}
& \left|\sum_{n=1}^{N} f_{3}(n) n^{-s}-\sum_{x \leq \sqrt{N}} f_{1}(x) x^{-s} \sum_{y \leq \sqrt{N}} f_{2}(y) y^{-s}\right| \\
& <\left(\sum_{x>\sqrt{N}}\left|f_{1}(x)\right| x^{-\sigma}\right)\left(\sum_{y=1}^{\infty}\left|f_{2}(y)\right| y^{-\sigma}\right) \\
& \quad+\left(\sum_{x=1}^{\infty}\left|f_{1}(x)\right| x^{-\sigma}\right)\left(\sum_{y>\sqrt{N}}\left|f_{2}(y)\right| y^{-\sigma}\right) \tag{5}
\end{align*}
$$

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

$$
\sum_{x>\sqrt{N}}\left|f_{1}(x)\right| x^{-\sigma} \quad \text { and } \quad \sum_{y>\sqrt{N}}\left|f_{2}(y)\right| y^{-\sigma}
$$

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

$$
\sum_{x=1}^{\infty}\left|f_{1}(x)\right| x^{-\sigma} \quad \text { and } \quad \sum_{y=1}^{\infty}\left|f_{2}(y)\right| y^{-\sigma}
$$

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

$$
\sum_{x \leq \sqrt{N}} f_{1}(x) x^{-s} \quad \text { and } \quad \sum_{y \leq \sqrt{N}} f_{2}(y) y^{-s}
$$

converge to $F_{1}(s)$ and $F_{2}(s)$ respectively as $N \rightarrow \infty$. The result follows.
REmARK. Theorem 3.5 generalizes to a product of $k$ Dirichlet series $F_{1}(s), \ldots, F_{k}(s)$, where the general coefficient is

$$
\sum_{\substack{x_{1}, \ldots, x_{k} \\ x_{1} \ldots x_{k}=n}} f_{1}\left(x_{1}\right) \ldots f_{k}\left(x_{k}\right)
$$

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 $f: \mathbb{N} \rightarrow \mathbb{C}$ is multiplicative.

THEOREM 3.6. Suppose that the function $f: \mathbb{N} \rightarrow \mathbb{C}$ is multiplicative. Then for every $s \in \mathbb{C}$ satisfying $\sigma>\sigma_{2}$, the series (1) satisfies

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

Proof. By the Remark, if $p_{j}$ is the $j$-th prime in increasing order, then

$$
\prod_{j=1}^{k}\left(\sum_{h=0}^{\infty} f\left(p_{j}^{h}\right) p_{j}^{-h s}\right)=\sum_{n=1}^{\infty}\left(\sum_{\substack{h_{1}, \ldots, h_{k} \\ p_{1}^{h_{1}} \ldots p_{k}^{h_{k}}=n}} f\left(p_{1}^{h_{1}}\right) \ldots f\left(p_{k}^{h_{k}}\right)\right) n^{-s}
$$

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

$$
\prod_{j=1}^{k}\left(\sum_{h=0}^{\infty} f\left(p_{j}^{h}\right) p_{j}^{-h s}\right)=\sum_{n=1}^{\infty} \theta_{k}(n) f(n) n^{-s}
$$

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 as $k \rightarrow \infty$, we have

$$
\begin{gathered}
\prod_{j=1}^{k}\left(\sum_{h=0}^{\infty} f\left(p_{j}^{h}\right) p_{j}^{-h s}\right)-\sum_{n=1}^{\infty} f(n) n^{-s}=\sum_{n=1}^{\infty}\left(\theta_{k}(n)-1\right) f(n) n^{-s} \\
=O\left(\sum_{n=k+1}^{\infty}|f(n)| n^{-\sigma}\right) \rightarrow 0
\end{gathered}
$$

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

THEOREM 3.7. Suppose that the function $f: \mathbb{N} \rightarrow \mathbb{C}$ is totally multiplicative. Then for every $s \in \mathbb{C}$ satisfying $\sigma>\sigma_{2}$, the series (1) satisfies

$$
F(s)=\prod_{p}\left(1-f(p) p^{-s}\right)^{-1}
$$

Proof. The absolute convergence of the series

$$
\begin{equation*}
\sum_{h=0}^{\infty} f\left(p^{h}\right) p^{-h s} \tag{6}
\end{equation*}
$$

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 $\left(1-f(p) p^{-s}\right)^{-1}$.

3-6 $\qquad$ W W L Chen : Elementary and Analytic Number Theory

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

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

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

# ELEMENTARY AND ANALYTIC NUMBER THEORY 

W W L CHEN

(C) W W L Chen, 1990 .

This work is available free, in the hope that it will be useful.
Any part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, with or without permission from the author.

## Chapter 4

## DISTRIBUTION OF PRIMES <br> II : 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 4.1. Suppose that $q \in \mathbb{N}$ and $a \in \mathbb{Z}$ satisfy $(a, q)=1$. Then there are infinitely many primes $p \equiv a(\bmod q)$.

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

It is easy to prove Theorem 4.1 by elementary methods for some special values of $a$ and $q$.
Example. There are infinitely many primes $p \equiv-1(\bmod 4)$. Suppose on the contrary that $p_{1}, \ldots, p_{r}$ represent all such primes. Then $4 p_{1} \ldots p_{r}-1$ must have a prime factor $p \equiv-1(\bmod 4)$. But $p$ cannot be any of $p_{1}, \ldots, p_{r}$.

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

[^3]$\qquad$

### 4.2. A Special Case

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

$$
\sum_{p \equiv a(\bmod q)} \frac{1}{p}
$$

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

$$
\sum_{p \equiv a(\bmod q)} \frac{\log p}{p^{\sigma}} \rightarrow+\infty \quad \text { as } \sigma \rightarrow 1+
$$

Let us illustrate the idea of Dirichlet by studying the case $n \equiv 1(\bmod 4)$.
First of all, we need a function that distinguishes between integers $n \equiv 1(\bmod 4)$ and the others. Suppose that $n$ is odd. Then it is easy to check that

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

so that

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

Now the series

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

so it suffices to show that the series

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

converges as $\sigma \rightarrow 1+$.
The next idea is to show that if we consider the series

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

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 $\sigma \rightarrow 1+$.

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, $\chi(m n)=\chi(m) \chi(n)$ for every $m, n \in \mathbb{N}$. Write

$$
\begin{equation*}
L(\sigma)=\sum_{n=1}^{\infty} \frac{\chi(n)}{n^{\sigma}} \tag{2}
\end{equation*}
$$

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

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

It follows from Theorems 3.2 and 3.5 that for $\sigma>1$, we have

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

Hence

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

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

$$
L(\sigma) \rightarrow L(1)=1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\ldots>0
$$

and

$$
L^{\prime}(\sigma) \rightarrow \frac{\log 3}{3}-\frac{\log 5}{5}+\frac{\log 7}{7}-\ldots
$$

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 $\chi: \mathbb{N} \rightarrow \mathbb{C}$, known nowadays as the Dirichlet characters modulo $q$, which generalize the function $\chi$ in the special case and which satisfy

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

To understand Dirichlet's ideas, we shall first of all study group characters. This 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 $\chi$ on $G$ for which $\chi(u v)=\chi(u) \chi(v)$ 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 $\chi$ 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 $y_{j}\left(\bmod h_{j}\right)$ is uniquely determined for every $j=1, \ldots, s$. Since $u \neq e$, there exists $k=1, \ldots, s$ such that $y_{k} \not \equiv 0\left(\bmod h_{k}\right)$. Let $\chi\left(v_{k}\right)=e\left(1 / h_{k}\right)$, and let $\chi\left(v_{j}\right)=1$ for every $j=1, \ldots, s$ such that $j \neq k$. Clearly $\chi(u)=e\left(y_{k} / h_{k}\right) \neq 1$.

We shall denote by $\chi_{0}$ 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 4.2. Suppose that $G$ is a finite abelian group of order $h$ and with identity element $e$. Suppose further that $\chi_{0}$ is the principal character on $G$.
(i) For every character $\chi$ 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

$$
\sum_{\chi} \chi(u)=\left\{\begin{array}{cc}
c & \text { if } u=e, \\
0 & \text { if } u \neq e,
\end{array}\right.
$$

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

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

the last equality following from the fact that $u v$ 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 $\chi_{1}$ such that $\chi_{1}(u) \neq 1$, so that

$$
\chi_{1}(u) \sum_{\chi} \chi(u)=\sum_{\chi} \chi_{1}(u) \chi(u)=\sum_{\chi}\left(\chi_{1} \chi\right)(u)=\sum_{\chi} \chi(u),
$$

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 $\chi$ runs through all the characters on $G$. Hence

$$
\left(1-\chi_{1}(u)\right) \sum_{\chi} \chi(u)=0 .
$$

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

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

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 $n \equiv a(\bmod q)$ 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 $G=\left\{a_{1}, \ldots, a_{\phi(q)}\right\}$. We can now define a character $\chi$ 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

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

A function $\chi: \mathbb{N} \rightarrow \mathbb{C}$ of the form (3) is called a Dirichlet character modulo $q$. Note that $\chi$ is totally multiplicative. Also, clearly there are exactly $\phi(q)$ Dirichlet characters modulo $q$. Furthermore, the principal Dirichlet character $\chi_{0}$ 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 4.2.
THEOREM 4.3. Suppose that $q \in \mathbb{N}$. Suppose further that $\chi_{0}$ is the principal Dirichlet character modulo $q$.
(i) For every Dirichlet character $\chi$ modulo $q$, we have

$$
\sum_{n(\bmod q)} \chi(n)= \begin{cases}\phi(q) & \text { if } \chi=\chi_{0} \\ 0 & \text { if } \chi \neq \chi_{0}\end{cases}
$$

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

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

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

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

### 4.4. Some Dirichlet Series

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

$$
\begin{equation*}
\zeta(s)=\sum_{n=1}^{\infty} n^{-s} \tag{4}
\end{equation*}
$$

furthermore, for any Dirichlet character $\chi$ modulo $q$, let

$$
\begin{equation*}
L(s, \chi)=\sum_{n=1}^{\infty} \chi(n) n^{-s} \tag{5}
\end{equation*}
$$

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 3.7 that for $\sigma>1$, the series (4) and (5) have the Euler product representations

$$
\zeta(s)=\prod_{p}\left(1-p^{-s}\right)^{-1} \quad \text { and } \quad L(s, \chi)=\prod_{p}\left(1-\chi(p) p^{-s}\right)^{-1}
$$

respectively. The following are some simple properties of these functions.
$\qquad$

THEOREM 4.4. Suppose that $\sigma>1$. Then $\zeta(s) \neq 0$. Furthermore, $L(s, \chi) \neq 0$ for every Dirichlet character $\chi$ modulo $q$.

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

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

and

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

THEOREM 4.5. Suppose that $\chi_{0}$ is the principal Dirichlet character modulo $q$. Then for $\sigma>1$, we have

$$
L\left(s, \chi_{0}\right)=\zeta(s) \prod_{p \mid q}\left(1-p^{-s}\right)
$$

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

$$
\begin{aligned}
L\left(s, \chi_{0}\right) & =\prod_{p}\left(1-\chi_{0}(p) p^{-s}\right)^{-1}=\prod_{p \nmid q}\left(1-p^{-s}\right)^{-1} \\
& =\frac{\prod_{p}\left(1-p^{-s}\right)^{-1}}{\prod_{p \mid q}\left(1-p^{-s}\right)^{-1}}=\zeta(s) \prod_{p \mid q}\left(1-p^{-s}\right) .
\end{aligned}
$$

THEOREM 4.6. Suppose that $\sigma>1$. Then

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

Furthermore, for every Dirichlet character $\chi$ modulo $q$, we have

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

Proof. Since $\sigma>1$, it follows from Theorem 3.2 that

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

It now follows from Theorem 3.5 and

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

that

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

The first assertion follows. On the other hand, it also follows from Theorem 3.2 that

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

It now follows from Theorem 3.5 and

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

that

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

The second assertion follows.
THEOREM 4.7. If $\sigma>1$, then for every Dirichlet character $\chi$ modulo $q$, we have

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

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

$$
\begin{equation*}
\log L(s, \chi)=\log \prod_{p}\left(1-\chi(p) p^{-s}\right)^{-1}=\sum_{p} \log \left(1-\chi(p) p^{-s}\right)^{-1} \tag{6}
\end{equation*}
$$

so that

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

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

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

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

### 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 $\mathbb{C} \backslash\{1\}, g(s)=f(s)$ 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 $s=1$.

We shall prove the following results on analytic continuation of $\zeta(s)$ and $L(s, \chi)$.

THEOREM 4.8. The function $\zeta(s)$ admits an analytic continuation to the half plane $\sigma>0$. Furthermore, $\zeta(s)$ is analytic for $\sigma>0$ except for a simple pole at $s=1$ with residue 1 .

THEOREM 4.9. Suppose that $q \in \mathbb{N}$ and $\chi_{0}$ is the principal Dirichlet character modulo $q$. Then the function $L\left(s, \chi_{0}\right)$ admits an analytic continuation to the half plane $\sigma>0$. Furthermore, $L\left(s, \chi_{0}\right)$ is analytic for $\sigma>0$ except for a simple pole at $s=1$ with residue $\phi(q) / q$.

THEOREM 4.10. Suppose that $q \in \mathbb{N}$ and $\chi$ is a non-principal Dirichlet character modulo $q$. Then the function $L(s, \chi)$ admits an analytic continuation to the half plane $\sigma>0$. Furthermore, $L(s, \chi)$ is analytic for $\sigma>0$.

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

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

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

$$
\begin{equation*}
\sum_{n \leq X} a(n) n^{-s}=S(X) X^{-s}+s \int_{1}^{X} S(x) x^{-s-1} \mathrm{~d} x \tag{7}
\end{equation*}
$$

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

$$
\begin{equation*}
F(s)=s \int_{1}^{\infty} S(x) x^{-s-1} \mathrm{~d} x . \tag{8}
\end{equation*}
$$

Proof. To prove (7), simply note that

$$
\begin{aligned}
\sum_{n \leq X} a(n) n^{-s}-S(X) X^{-s} & =\sum_{n \leq X} a(n)\left(n^{-s}-X^{-s}\right)=\sum_{n \leq X} a(n) \int_{n}^{X} s x^{-s-1} \mathrm{~d} x \\
& =s \int_{1}^{X}\left(\sum_{n \leq x} a(n)\right) x^{-s-1} \mathrm{~d} x=s \int_{1}^{X} S(x) x^{-s-1} \mathrm{~d} x
\end{aligned}
$$

Also, (8) follows from (7) on letting $X \rightarrow \infty$.
The second technical result, standard in complex function theory, will be stated without proof.
THEOREM 4.12. Suppose that the path $\Gamma$ 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 $u^{\prime}(t)$ and $v^{\prime}(t)$ 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
(i) $f(s, w)$ is continuous for every $s \in D$ and every $w \in \Gamma$; and
(ii) for every $w \in \Gamma$, the function $f(s, w)$ is analytic in $D$.

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

Proof of Theorem 4.8. Let $F(s)=\zeta(s)$. In the notation of Theorem 4.11, 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

$$
\begin{aligned}
\zeta(s) & =s \int_{1}^{\infty}[x] x^{-s-1} \mathrm{~d} x=s \int_{1}^{\infty} x^{-s} \mathrm{~d} x-s \int_{1}^{\infty}\{x\} x^{-s-1} \mathrm{~d} x \\
& =1+\frac{1}{s-1}-s \int_{1}^{\infty}\{x\} x^{-s-1} \mathrm{~d} x
\end{aligned}
$$

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 4.12. To show (ii), note that for $\sigma>\delta$, we have

$$
\left|F_{n}(s)\right|=\left|\int_{n}^{n+1}\{x\} x^{-s-1} \mathrm{~d} x\right| \leq n^{-\sigma-1}<n^{-1-\delta}
$$

and (ii) follows from the Weierstrass $M$-test.
Proof of Theorem 4.9. Suppose that $\sigma>1$. Recall Theorem 4.5, that

$$
L\left(s, \chi_{0}\right)=\zeta(s) \prod_{p \mid q}\left(1-\frac{1}{p^{s}}\right)
$$

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 4.10 is left as an exercise.

### 4.6. Proof of Dirichlet's Theorem

We now attempt to prove Theorem 4.1. The following theorem will enable us to consider the analogue of (1).

THEOREM 4.13. Suppose that $\sigma>1$. Then

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

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

$$
\begin{aligned}
& \sum_{n \equiv a(\bmod q)} \frac{\Lambda(n)}{n^{\sigma}}-\sum_{p \equiv a(\bmod q)} \frac{\log p}{p^{\sigma}} \leq \sum_{p} \sum_{m=2}^{\infty} \frac{\log p}{p^{m \sigma}} \\
& \quad \leq \sum_{p} \sum_{m=2}^{\infty} \frac{\log p}{p^{m}}=\sum_{p} \frac{\log p}{p(p-1)} \leq \sum_{n=2}^{\infty} \frac{\log n}{n(n-1)}=O(1)
\end{aligned}
$$

The result follows.
Combining Theorems 4.13, 4.3 and 4.6, we have

$$
\begin{align*}
\sum_{p \equiv a(\bmod q)} \frac{\log p}{p^{\sigma}} & =\sum_{n \equiv a(\bmod q)} \Lambda(n) n^{-\sigma}+O(1) \\
& =\sum_{n=1}^{\infty}\left(\frac{1}{\phi(q)} \sum_{\chi(\bmod q)} \frac{\chi(n)}{\chi(a)}\right) \Lambda(n) n^{-\sigma}+O(1) \\
& =\frac{1}{\phi(q)} \sum_{\chi(\bmod q)} \frac{1}{\chi(a)} \sum_{n=1}^{\infty} \chi(n) \Lambda(n) n^{-\sigma}+O(1) \\
& =-\frac{1}{\phi(q)} \sum_{\chi(\bmod q)} \frac{1}{\chi(a)} \frac{L^{\prime}(\sigma, \chi)}{L(\sigma, \chi)}+O(1) \tag{9}
\end{align*}
$$

Suppose now that

$$
\begin{equation*}
\sum_{\substack{\chi(\bmod q) \\ \chi \neq \chi_{0}}} \frac{1}{\chi(a)} \frac{L^{\prime}(\sigma, \chi)}{L(\sigma, \chi)}=O(1) \quad \text { as } \sigma \rightarrow 1+ \tag{10}
\end{equation*}
$$

Then combining (9) and (10), as $\sigma \rightarrow 1+$, we have

$$
\sum_{p \equiv a} \frac{\log p}{p^{\sigma}}=-\frac{1}{\phi(q)} \frac{L^{\prime}\left(\sigma, \chi_{0}\right)}{L\left(\sigma, \chi_{0}\right)}+O(1)=\frac{1}{\phi(q)} \frac{1}{\sigma-1}+O(1) \rightarrow \infty
$$

since the function $L^{\prime}\left(s, \chi_{0}\right) / L\left(s, \chi_{0}\right)$ has a simple pole at $s=1$ with residue -1 by Theorem 4.9. 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 $\chi(\bmod q)$, we have $L(1, \chi) \neq 0$. Here we need to distinguish two cases, represented by the following two theorems.

THEOREM 4.14. Suppose that $q \in \mathbb{N}$ and $\chi$ is a non-real Dirichlet character modulo $q$. Then $L(1, \chi) \neq 0$.

Proof. For $\sigma>1$, we have, in view of Theorem 4.7, that

$$
\begin{aligned}
\sum_{\chi(\bmod q)} \log L(\sigma, \chi) & =\sum_{\chi(\bmod q)} \sum_{p} \sum_{m=1}^{\infty} \chi\left(p^{m}\right) m^{-1} p^{-m \sigma} \\
& =\sum_{p} \sum_{m=1}^{\infty}\left(\sum_{\chi(\bmod q)} \chi\left(p^{m}\right)\right) m^{-1} p^{-m \sigma} \\
& =\phi(q) \sum_{p} \sum_{\substack{m=1 \\
p^{m} \equiv 1(\bmod q)}}^{\infty} m^{-1} p^{-m \sigma}>0
\end{aligned}
$$

where the change of order of summation is justified since

$$
\sum_{\chi(\bmod q)} \sum_{p} \sum_{m=1}^{\infty}\left|\chi\left(p^{m}\right) m^{-1} p^{-m \sigma}\right|
$$

is finite. It follows that

$$
\begin{equation*}
\left|\prod_{\chi(\bmod q)} L(\sigma, \chi)\right|>1 \tag{11}
\end{equation*}
$$

Suppose that $\chi_{1}$ is a non-real Dirichlet character modulo $q$, and $L\left(1, \chi_{1}\right)=0$. Then $\chi_{1} \neq \overline{\chi_{1}}$, and $L\left(1, \overline{\chi_{1}}\right)=\overline{L\left(1, \chi_{1}\right)}=0$ also. It follows that these two zeros more than cancel the simple pole of $L\left(\sigma, \chi_{0}\right)$ 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 $\chi$ is real.
THEOREM 4.15. Suppose that $q \in \mathbb{N}$ and $\chi$ is a real, non-principal Dirichlet character modulo $q$. Then $L(1, \chi) \neq 0$.

Proof. Suppose that the result is false, so that there exists a real Dirichlet character $\chi$ 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 $n \in \mathbb{N}$,

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

Let the function $g: \mathbb{N} \rightarrow \mathbb{R}$ 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

$$
\begin{equation*}
f(n) \geq g(n) \tag{12}
\end{equation*}
$$

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

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

so that

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

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 3.2, 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),

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

by Theorem 3.2. It follows that for $0<r<3 / 2$, we have

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

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

# ELEMENTARY AND ANALYTIC NUMBER THEORY 

W W L CHEN

## Chapter 5

## DISTRIBUTION OF PRIMES

## III : 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 5.1. As $X \rightarrow \infty$, we have

$$
\pi(X) \sim \frac{X}{\log X}
$$

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

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

THEOREM 5.2. As $X \rightarrow \infty$, we have

$$
\psi(X) \sim X \quad \text { if and only if } \quad \pi(X) \sim \frac{X}{\log X}
$$

[^4]$\qquad$

Proof. Recall the proof of Theorem 2.5 due to Tchebycheff. We have

$$
\begin{align*}
\psi(X) & =\sum_{n \leq X} \Lambda(n)=\sum_{\substack{p, k \\
p^{k} \leq X}} \log p=\sum_{p \leq X}(\log p) \sum_{1 \leq k \leq \frac{\log X}{\log p}} 1 \\
& =\sum_{p \leq X}(\log p)\left[\frac{\log X}{\log p}\right] \leq \pi(X) \log X . \tag{1}
\end{align*}
$$

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

$$
\begin{align*}
\psi(X) & \geq \sum_{p \leq X} \log p \geq \sum_{X^{\alpha}<p \leq X} \log p \geq\left(\pi(X)-\pi\left(X^{\alpha}\right)\right) \log \left(X^{\alpha}\right) \\
& =\alpha\left(\pi(X)-\pi\left(X^{\alpha}\right)\right) \log X \tag{2}
\end{align*}
$$

Combining (1) and (2), we have

$$
\begin{equation*}
\alpha \frac{\pi(X)}{X / \log X}-\alpha \frac{\pi\left(X^{\alpha}\right)}{X / \log X} \leq \frac{\psi(X)}{X} \leq \frac{\pi(X)}{X / \log X} \tag{3}
\end{equation*}
$$

Since $\alpha<1$, it follows from Tchebycheff's theorem that as $X \rightarrow \infty$, we have

$$
\frac{\pi\left(X^{\alpha}\right)}{X / \log X} \rightarrow 0
$$

Suppose that as $X \rightarrow \infty$, we have $\pi(X) \sim X / \log X$. Then as $X \rightarrow \infty$, we have

$$
\alpha \frac{\pi(X)}{X / \log X}-\alpha \frac{\pi\left(X^{\alpha}\right)}{X / \log X} \rightarrow \alpha
$$

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

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

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

$$
\frac{\psi(X)}{X} \rightarrow 1
$$

as $X \rightarrow \infty$.
Note next that the inequalities (3) can be rewritten as

$$
\frac{\psi(X)}{X} \leq \frac{\pi(X)}{X / \log X} \leq \frac{1}{\alpha} \frac{\psi(X)}{X}+\frac{\pi\left(X^{\alpha}\right)}{X / \log X}
$$

Suppose that as $X \rightarrow \infty$, we have $\psi(X) \sim X$. Then a similar argument gives

$$
\frac{\pi(X)}{X / \log X} \rightarrow 1
$$

as $X \rightarrow \infty$.

### 5.2. A Smoothing Argument

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

$$
\begin{equation*}
\psi_{1}(X)=\int_{0}^{X} \psi(x) \mathrm{d} x \tag{4}
\end{equation*}
$$

THEOREM 5.3. Suppose that as $X \rightarrow \infty$, we have $\psi_{1}(X) \sim \frac{1}{2} X^{2}$. Then as $X \rightarrow \infty$, we have $\psi(X) \sim X$.

Proof. Suppose that $0<\alpha<1<\beta$. Since $\Lambda(n) \geq 0$ for every $n \in \mathbb{N}$, the function $\psi$ 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

$$
\begin{equation*}
\frac{\psi(X)}{X} \leq \frac{\psi_{1}(\beta X)-\psi_{1}(X)}{(\beta-1) X^{2}} \tag{5}
\end{equation*}
$$

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

$$
\begin{equation*}
\frac{\psi(X)}{X} \geq \frac{\psi_{1}(X)-\psi_{1}(\alpha X)}{(1-\alpha) X^{2}} \tag{6}
\end{equation*}
$$

As $X \rightarrow \infty$, we have

$$
\begin{equation*}
\frac{\psi_{1}(\beta X)-\psi_{1}(X)}{(\beta-1) X^{2}} \sim \frac{1}{\beta-1}\left(\frac{1}{2} \beta^{2}-\frac{1}{2}\right)=\frac{1}{2}(\beta+1) \tag{7}
\end{equation*}
$$

and

$$
\begin{equation*}
\frac{\psi_{1}(X)-\psi_{1}(\alpha X)}{(1-\alpha) X^{2}} \sim \frac{1}{1-\alpha}\left(\frac{1}{2}-\frac{1}{2} \alpha^{2}\right)=\frac{1}{2}(\alpha+1) \tag{8}
\end{equation*}
$$

Since $\alpha$ and $\beta$ are arbitrary, we conclude, on combining (5)-(8), that as $X \rightarrow \infty$, we have $\psi(X) / X \sim 1$.

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

THEOREM 5.4. As $X \rightarrow \infty$, we have

$$
\psi_{1}(X) \sim \frac{1}{2} X^{2} .
$$

### 5.3. A Contour Integral

The following theorem provides a link between $\psi_{1}(X)$ and the Riemann zeta function $\zeta(s)$.
THEOREM 5.5. Suppose that $X>0$ and $c>1$. Then

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

where the path of integration is the straight line $\sigma=c$.
A crucial step in the proof of Theorem 5.5 is provided by the following result concerning a particular contour integral.
$\qquad$

THEOREM 5.6. Suppose that $X>0$ and $c>1$. Then

$$
\frac{1}{2 \pi \mathrm{i}} \int_{c-\mathrm{i} \infty}^{c+\mathrm{i} \infty} \frac{X^{s}}{s(s+1)} \mathrm{d} s= \begin{cases}0 & (X \leq 1) \\ 1-\frac{1}{X} & (X \geq 1)\end{cases}
$$

Proof. Note first of all that the integral is absolutely convergent, as

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

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

$$
I_{T}=\frac{1}{2 \pi \mathrm{i}} \int_{c-\mathrm{i} T}^{c+\mathrm{i} T} \frac{X^{s}}{s(s+1)} \mathrm{d} s
$$

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

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

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

$$
\left|J_{T}^{-}\right| \leq \frac{1}{2 \pi} \frac{X^{c}}{R(R-1)} 2 \pi R \leq \frac{X^{c}}{T-1} \rightarrow 0 \quad \text { as } T \rightarrow \infty
$$

By Cauchy's residue theorem, we have

$$
I_{T}=J_{T}^{-}+\operatorname{res}\left(\frac{X^{s}}{s(s+1)}, 0\right)+\operatorname{res}\left(\frac{X^{s}}{s(s+1)},-1\right)=J_{T}^{-}+1-\frac{1}{X}
$$

The result for $X \geq 1$ follows on letting $T \rightarrow \infty$.
Suppose now that $X \leq 1$. Let $A^{+}(c, T)$ denote the circular arc centred at $s=0$ and passing from $c-\mathrm{i} T$ to $c+\mathrm{i} T$ on the right of the line $\sigma=c$. Furthermore, let

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

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

$$
\left|J_{T}^{+}\right| \leq \frac{1}{2 \pi} \frac{X^{c}}{R^{2}} 2 \pi R \leq \frac{X^{c}}{T} \rightarrow 0 \quad \text { as } T \rightarrow \infty
$$

By Cauchy's residue theorem, we have

$$
I_{T}=J_{T}^{+}
$$

The result for $X \leq 1$ follows on letting $T \rightarrow \infty$.
Proof of Theorem 5.5. Note that for $X \geq 1$, we have

$$
\psi_{1}(X)=\int_{0}^{X} \psi(x) \mathrm{d} x=\int_{1}^{X} \psi(x) \mathrm{d} x=\int_{1}^{X}\left(\sum_{n \leq x} \Lambda(n)\right) \mathrm{d} x=\sum_{n \leq X}(X-n) \Lambda(n)
$$

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 5.6 that for every $X>0$, we have

$$
\begin{aligned}
\frac{\psi_{1}(X)}{X} & =\sum_{n \leq X}\left(1-\frac{n}{X}\right) \Lambda(n)=\sum_{n \leq X}\left(1-\frac{1}{X / n}\right) \Lambda(n) \\
& =\sum_{n=1}^{\infty} \frac{\Lambda(n)}{2 \pi \mathrm{i}} \int_{c-\mathrm{i} \infty}^{c+\mathrm{i} \infty} \frac{(X / n)^{s}}{s(s+1)} \mathrm{d} s
\end{aligned}
$$

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

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

is finite. It follows that

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

as required.

### 5.4. The Riemann Zeta Function

Recall first of all Theorem 4.11. In the case of the Riemann zeta function, equation (7) of Chapter 4 becomes

$$
\begin{align*}
\sum_{n \leq X} n^{-s} & =s \int_{1}^{X}[x] x^{-s-1} \mathrm{~d} x+[X] X^{-s} \\
& =s \int_{1}^{X} x^{-s} \mathrm{~d} x-s \int_{1}^{X}\{x\} x^{-s-1} \mathrm{~d} x+X^{-s+1}-\{X\} X^{-s} \\
& =\frac{s}{s-1}-\frac{s}{(s-1) X^{s-1}}-s \int_{1}^{X} \frac{\{x\}}{x^{s+1}} \mathrm{~d} x+\frac{1}{X^{s-1}}-\frac{\{X\}}{X^{s}} \tag{9}
\end{align*}
$$

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

$$
\begin{equation*}
\zeta(s)=\frac{s}{s-1}-s \int_{1}^{\infty} \frac{\{x\}}{x^{s+1}} \mathrm{~d} x \tag{10}
\end{equation*}
$$

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 $\zeta(\sigma+\mathrm{i} t)$ and $\zeta(\sigma-\mathrm{i} t)$ are complex conjugates, so it suffices to study $\zeta(s)$ on the upper half plane.

THEOREM 5.7. For every $\sigma \geq 1$ and $t \geq 2$, we have
(i) $|\zeta(s)|=O(\log t)$; and
(ii) $\left|\zeta^{\prime}(s)\right|=O\left(\log ^{2} t\right)$.

Suppose further that $0<\delta<1$. Then for every $\sigma \geq \delta$ and $t \geq 1$, we have
(iii) $|\zeta(s)|=O_{\delta}\left(t^{1-\delta}\right)$.

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

$$
\begin{align*}
\zeta(s)-\sum_{n \leq X} \frac{1}{n^{s}} & =\frac{s}{(s-1) X^{s-1}}-\frac{1}{X^{s-1}}+\frac{\{X\}}{X^{s}}-s \int_{X}^{\infty} \frac{\{x\}}{x^{s+1}} \mathrm{~d} x \\
& =\frac{1}{(s-1) X^{s-1}}+\frac{\{X\}}{X^{s}}-s \int_{X}^{\infty} \frac{\{x\}}{x^{s+1}} \mathrm{~d} x \tag{11}
\end{align*}
$$

$\qquad$

It follows that

$$
\begin{align*}
|\zeta(s)| & \leq \sum_{n \leq X} \frac{1}{n^{\sigma}}+\frac{1}{t X^{\sigma-1}}+\frac{1}{X^{\sigma}}+|s| \int_{X}^{\infty} \frac{\mathrm{d} x}{x^{\sigma+1}} \\
& \leq \sum_{n \leq X} \frac{1}{n^{\sigma}}+\frac{1}{t X^{\sigma-1}}+\frac{1}{X^{\sigma}}+\left(1+\frac{t}{\sigma}\right) \frac{1}{X^{\sigma}} \tag{12}
\end{align*}
$$

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

$$
\begin{aligned}
|\zeta(s)| & \leq \sum_{n \leq X} \frac{1}{n^{\delta}}+\frac{1}{t X^{\delta-1}}+\left(2+\frac{t}{\delta}\right) \frac{1}{X^{\delta}} \\
& \leq \int_{0}^{[X]} \frac{\mathrm{d} x}{x^{\delta}}+\frac{X^{1-\delta}}{t}+\frac{3 t}{\delta X^{\delta}} \leq \frac{X^{1-\delta}}{1-\delta}+X^{1-\delta}+\frac{3 t}{\delta X^{\delta}}
\end{aligned}
$$

Again choosing $X=t$, we obtain

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

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

$$
\left|\zeta^{\prime}\left(s_{0}\right)\right|=\left|\frac{1}{2 \pi \mathrm{i}} \int_{C} \frac{\zeta(s)}{\left(s-s_{0}\right)^{2}} \mathrm{~d} s\right| \leq \frac{M}{\rho}
$$

where $M=\sup _{s \in C}|\zeta(s)|$. Now for every $s \in C$, we clearly have $\sigma \geq \sigma_{0}-\rho \geq 1-\rho$ and $2 t_{0}>t \geq$ $t_{0}-\rho>1$. It follows from (13) that for every $s \in C$, we must have ( $\delta=1-\rho$ )

$$
|\zeta(s)| \leq\left(2 t_{0}\right)^{\rho}\left(\frac{1}{\rho}+1+\frac{3}{1-\rho}\right) \leq \frac{10 t_{0}^{\rho}}{\rho}
$$

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

$$
\left|\zeta^{\prime}\left(s_{0}\right)\right| \leq \frac{10 t_{0}^{\rho}}{\rho^{2}}
$$

We now take $\rho=\left(\log t_{0}+2\right)^{-1}$. Then $t_{0}^{\rho}=\mathrm{e}^{\rho \log t_{0}}<\mathrm{e}$, and so

$$
\left|\zeta^{\prime}\left(s_{0}\right)\right| \leq 10 \mathrm{e}\left(\log t_{0}+2\right)^{2}
$$

(ii) now follows.

THEOREM 5.8. The function $\zeta(s)$ has no zeros on the line $\sigma=1$. Furthermore, there is a positive constant $A$ such that as $t \rightarrow \infty$, 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

$$
\begin{equation*}
3+4 \cos \theta+\cos 2 \theta=2(1+\cos \theta)^{2} \geq 0 \tag{14}
\end{equation*}
$$

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

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

so that

$$
\begin{equation*}
\log |\zeta(\sigma+\mathrm{i} t)|=\mathfrak{R}\left(\sum_{n=2}^{\infty} c_{n} n^{-\sigma-\mathrm{i} t}\right)=\sum_{n=2}^{\infty} c_{n} n^{-\sigma} \cos (t \log n), \tag{15}
\end{equation*}
$$

where

$$
c_{n}= \begin{cases}1 / m & \left(n=p^{m}, \text { where } p \text { is prime and } m \in \mathbb{N}\right),  \tag{16}\\ 0 & \text { (otherwise) }\end{cases}
$$

Combining (14)-(16), we have

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

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

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

Suppose that the point $s=1+\mathrm{it}$ is a zero of $\zeta(s)$. Then since $\zeta(s)$ is analytic at the points $s=1+\mathrm{i} t$ and $s=1+2 \mathrm{it}$ and has a simple pole with residue 1 at $s=1$, the left hand side of (17) must converge to a finite limit as $\sigma \rightarrow 1+$, contradicting the fact that the right hand side diverges to infinity as $\sigma \rightarrow 1+$. Hence $s=1+\mathrm{it}$ 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

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

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

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

by Theorem $5.7(\mathrm{i})$, where $A_{1}$ is a positive absolute constant. Since $\log (2 t) \leq 2 \log t$, it follows that

$$
\begin{equation*}
|\zeta(\sigma+\mathrm{i} t)| \geq \frac{(\sigma-1)^{3 / 4}}{A_{2}(\log t)^{1 / 4}} \tag{18}
\end{equation*}
$$

where $A_{2}$ is a positive absolute constant. Note that (18) holds also when $\sigma=1$. Suppose now that $1<\eta<2$. If $1 \leq \sigma \leq \eta$ and $t \geq 2$, then it follows from Theorem 5.7(ii) that

$$
|\zeta(\sigma+\mathrm{i} t)-\zeta(\eta+\mathrm{i} t)|=\left|\int_{\sigma}^{\eta} \zeta^{\prime}(x+\mathrm{i} t) \mathrm{d} x\right| \leq A_{3}(\eta-1) \log ^{2} t
$$

where $A_{3}$ is a positive absolute constant. Combining this with (18), we have

$$
\begin{equation*}
|\zeta(\sigma+\mathrm{i} t)| \geq|\zeta(\eta+\mathrm{i} t)|-A_{3}(\eta-1) \log ^{2} t \geq \frac{(\eta-1)^{3 / 4}}{A_{2}(\log t)^{1 / 4}}-A_{3}(\eta-1) \log ^{2} t \tag{19}
\end{equation*}
$$

$\qquad$

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

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

in other words,

$$
\eta=1+\left(2 A_{2} A_{3}\right)^{-4}(\log t)^{-9}
$$

where $t>t_{0}$ so that $\eta<2$. Then

$$
|\zeta(\sigma+\mathrm{i} t)| \geq A_{3}(\eta-1) \log ^{2} t=A_{4}(\log t)^{-7}
$$

for $1 \leq \sigma \leq 2$ and $t>t_{0}$.

### 5.5. Completion of the Proof

We are now ready to complete the proof of Theorem 5.4. By Theorem 5.5, we have

$$
\begin{equation*}
\frac{\psi_{1}(X)}{X^{2}}=\frac{1}{2 \pi \mathrm{i}} \int_{c-\mathrm{i} \infty}^{c+\mathrm{i} \infty} G(s) X^{s-1} \mathrm{~d} s \tag{20}
\end{equation*}
$$

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

$$
G(s)=-\frac{1}{s(s+1)} \frac{\zeta^{\prime}(s)}{\zeta(s)}=-\frac{1}{s(s+1)} \zeta^{\prime}(s) \frac{1}{\zeta(s)}
$$

By Theorems 4.8, 5.7 and 5.8, we know that $G(s)$ is analytic for $\sigma \geq 1$, except at $s=1$, and that for some positive absolute constant $A$, we have

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

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

$$
\left\{\begin{array}{l}
L_{1}=[1-\mathrm{i} U, 1-\mathrm{i} T], \\
L_{2}=[1-\mathrm{i} T, \alpha-\mathrm{i} T], \\
L_{3}=[\alpha-\mathrm{i} T, \alpha+\mathrm{i} T], \\
L_{4}=[\alpha+\mathrm{i} T, 1+\mathrm{i} T], \\
L_{5}=[1+\mathrm{i} T, 1+\mathrm{i} U],
\end{array}\right.
$$

where $T=T(\epsilon)>\max \left\{t_{0}, 2\right\}, \alpha=\alpha(T)=\alpha(\epsilon) \in(0,1)$ and $U$ are chosen to satisfy the following 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

$$
\left\{\begin{array}{l}
M_{1}=[c-\mathrm{i} U, 1-\mathrm{i} U] \\
M_{2}=[1+\mathrm{i} U, c+\mathrm{i} U] .
\end{array}\right.
$$

The above contours are indicated in the diagram below:


By Cauchy's residue theorem, we have

$$
\begin{align*}
& \frac{1}{2 \pi \mathrm{i}} \int_{c-\mathrm{i} U}^{c+\mathrm{i} U} G(s) X^{s-1} \mathrm{~d} s=\frac{1}{2 \pi \mathrm{i}} \sum_{j=1}^{2} \int_{M_{j}} G(s) X^{s-1} \mathrm{~d} s \\
& \quad+\frac{1}{2 \pi \mathrm{i}} \sum_{j=1}^{5} \int_{L_{j}} G(s) X^{s-1} \mathrm{~d} s+\operatorname{res}\left(G(s) X^{s-1}, 1\right) \tag{22}
\end{align*}
$$

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

$$
\begin{equation*}
\operatorname{res}\left(G(s) X^{s-1}, 1\right)=\frac{1}{2} \tag{23}
\end{equation*}
$$

Now

$$
\begin{equation*}
\left|\int_{L_{1}} G(s) X^{s-1} \mathrm{~d} s\right|=\left|\int_{L_{5}} G(s) X^{s-1} \mathrm{~d} s\right| \leq \int_{T}^{\infty}|G(1+\mathrm{i} t)| \mathrm{d} t<\epsilon \tag{24}
\end{equation*}
$$

On the other hand,

$$
\begin{equation*}
\left|\int_{L_{2}} G(s) X^{s-1} \mathrm{~d} s\right|=\left|\int_{L_{4}} G(s) X^{s-1} \mathrm{~d} s\right| \leq M \int_{\alpha}^{1} X^{\sigma-1} \mathrm{~d} \sigma \leq \frac{M}{\log X} \tag{25}
\end{equation*}
$$

and

$$
\begin{equation*}
\left|\int_{L_{3}} G(s) X^{s-1} \mathrm{~d} s\right| \leq 2 T M X^{\alpha-1} \tag{26}
\end{equation*}
$$

where

$$
\begin{equation*}
M=M(\alpha, T)=M(\epsilon)=\sup _{L_{2} \cup L_{3} \cup L_{4}}|G(s)| . \tag{27}
\end{equation*}
$$

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

$$
\begin{equation*}
\left|\int_{M_{j}} G(s) X^{s-1} \mathrm{~d} s\right| \leq \int_{1}^{c}|U|^{-3 / 2} X^{\sigma-1} \mathrm{~d} \sigma \leq \frac{X^{c-1}}{\log X}|U|^{-3 / 2} . \tag{28}
\end{equation*}
$$

Combining (22)-(28), we have

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

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

$$
\left|\frac{\psi_{1}(X)}{X^{2}}-\frac{1}{2}\right| \leq \frac{\epsilon}{\pi}+\frac{M}{\pi \log X}+\frac{T M}{\pi X^{1-\alpha}} .
$$

It clearly follows that as $X \rightarrow \infty$, we have

$$
\frac{\psi_{1}(X)}{X^{2}} \rightarrow \frac{1}{2}
$$

# ELEMENTARY AND ANALYTIC NUMBER THEORY 

W W L CHEN

(C) W W L Chen, 1990 .<br>This work is available free, in the hope that it will be useful.

Any part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, with or without permission from the author.

## 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 6.1. The function $\zeta(s)$ can be continued analytically over the whole plane, and satisfies the functional equation

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

where $\Gamma$ denotes the gamma function. In particular, $\zeta(s)$ 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 $\sigma>1$. In particular, 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 $0 \leq \sigma \leq 1$ is called the critical strip.
Riemann's paper is particularly remarkable in the conjectures it contains. While most of the conjectures have been proved, the famous Riemann hypothesis has so far resisted all attempts to prove or disprove it.

THEOREM 6.2. (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).

[^5]$\qquad$

THEOREM 6.3. (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 $0<\gamma \leq T$. Then

$$
\begin{equation*}
N(T)=\frac{T}{2 \pi} \log \frac{T}{2 \pi}-\frac{T}{2 \pi}+O(\log T) . \tag{2}
\end{equation*}
$$

THEOREM 6.4. (Hadamard 1893) The entire function

$$
\begin{equation*}
\xi(s)=\frac{1}{2} s(s-1) \pi^{-s / 2} \Gamma\left(\frac{s}{2}\right) \zeta(s) \tag{3}
\end{equation*}
$$

has the product representation

$$
\begin{equation*}
\xi(s)=\mathrm{e}^{A+B s} \prod_{\rho}\left(1-\frac{s}{\rho}\right) \mathrm{e}^{s / \rho} \tag{4}
\end{equation*}
$$

where $A$ and $B$ are constants and where $\rho$ runs over all the zeros of $\zeta(s)$ in the critical strip.
We comment here that the product representation (4) plays an important role in the first proof of the Prime number theorem.

The most remarkable of Riemann's conjectures is an explicit formula for the difference $\pi(X)-\operatorname{li}(X)$, 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 6.5. (von Mangoldt 1895) Let

$$
\psi(X)=\sum_{n \leq X} \Lambda(n) \quad \text { and } \quad \psi_{0}(X)=\frac{\psi(X-0)+\psi(X+0)}{2}
$$

Then

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

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 $\sigma=1 / 2$.

### 6.2. Riemann's Proof of Theorem 6.1

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

$$
\Gamma\left(\frac{s}{2}\right)=\int_{0}^{\infty} t^{s / 2-1} \mathrm{e}^{-t} \mathrm{~d} t=\left(n^{2} \pi\right)^{s / 2} \int_{0}^{\infty} x^{s / 2-1} \mathrm{e}^{-n^{2} \pi x} \mathrm{~d} x
$$

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

$$
\begin{align*}
\pi^{-s / 2} \Gamma\left(\frac{s}{2}\right) \zeta(s) & =\int_{1}^{\infty} x^{s / 2-1} \omega(x) \mathrm{d} x+\int_{0}^{1} y^{s / 2-1} \omega(y) \mathrm{d} y \\
& =\int_{1}^{\infty} x^{s / 2-1} \omega(x) \mathrm{d} x+\int_{1}^{\infty} x^{-s / 2-1} \omega\left(x^{-1}\right) \mathrm{d} x \tag{5}
\end{align*}
$$

We shall show that the function

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

satisfies the functional equation

$$
\begin{equation*}
\theta\left(x^{-1}\right)=x^{1 / 2} \theta(x) \tag{6}
\end{equation*}
$$

for every $x>0$. It then follows that

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

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

$$
\begin{align*}
\int_{1}^{\infty} x^{-s / 2-1} \omega\left(x^{-1}\right) \mathrm{d} x & =\int_{1}^{\infty} x^{-s / 2-1}\left(-\frac{1}{2}+\frac{1}{2} x^{1 / 2}+x^{1 / 2} \omega(x)\right) \mathrm{d} x \\
& =\frac{1}{s(s-1)}+\int_{1}^{\infty} x^{-s / 2-1 / 2} \omega(x) \mathrm{d} x \tag{7}
\end{align*}
$$

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

$$
\begin{equation*}
\pi^{-s / 2} \Gamma\left(\frac{s}{2}\right) \zeta(s)=\frac{1}{s(s-1)}+\int_{1}^{\infty}\left(x^{s / 2-1}+x^{-s / 2-1 / 2}\right) \omega(x) \mathrm{d} x \tag{8}
\end{equation*}
$$

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 $\omega(x)=O\left(\mathrm{e}^{-\pi x}\right)$ as $x \rightarrow+\infty$. 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 (6) for every $x>0$. In other words, we need to prove that for every $x>0$, we have

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

The starting point is the Poisson summation formula, that under certain conditions on a function $f(t)$, we have

$$
\begin{equation*}
\sum_{A \leq n \leq B}^{\prime} f(n)=\sum_{\nu=-\infty}^{\infty} \int_{A}^{B} f(t) \mathrm{e}^{2 \pi \mathrm{i} \nu t} \mathrm{~d} t \tag{9}
\end{equation*}
$$

where $\sum^{\prime}$ 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

$$
\sum_{n=-N}^{N} \mathrm{e}^{-n^{2} \pi / x}=\sum_{\nu=-\infty}^{\infty} \int_{-N}^{N} \mathrm{e}^{-t^{2} \pi / x} \mathrm{e}^{2 \pi \mathrm{i} \nu t} \mathrm{~d} t
$$

Letting $N \rightarrow \infty$, we obtain

$$
\begin{equation*}
\sum_{n=-\infty}^{\infty} \mathrm{e}^{-n^{2} \pi / x}=\sum_{\nu=-\infty}^{\infty} \int_{-\infty}^{\infty} \mathrm{e}^{-t^{2} \pi / x} \mathrm{e}^{2 \pi \mathrm{i} \nu t} \mathrm{~d} t \tag{10}
\end{equation*}
$$

This is justified by noting that

$$
\left(\int_{-\infty}^{-N}+\int_{N}^{\infty}\right) \mathrm{e}^{-t^{2} \pi / x} \mathrm{e}^{2 \pi \mathrm{i} \nu t} \mathrm{~d} t=2 \int_{N}^{\infty} \mathrm{e}^{-t^{2} \pi / x} \cos (2 \pi \nu t) \mathrm{d} t
$$

and that

$$
\left|\sum_{\nu \neq 0} \int_{N}^{\infty} \mathrm{e}^{-t^{2} \pi / x} \cos (2 \pi \nu t) \mathrm{d} t\right| \rightarrow 0
$$

as $N \rightarrow \infty$. Writing $t=x u$ and using (10), we have

$$
\begin{align*}
\sum_{n=-\infty}^{\infty} \mathrm{e}^{-n^{2} \pi / x} & =x \sum_{\nu=-\infty}^{\infty} \int_{-\infty}^{\infty} \mathrm{e}^{-u^{2} \pi x} \mathrm{e}^{2 \pi \mathrm{i} \nu x u} \mathrm{~d} u \\
& =x \sum_{\nu=-\infty}^{\infty} \int_{-\infty}^{\infty} \mathrm{e}^{-(u-\mathrm{i} \nu)^{2} \pi x-\nu^{2} \pi x} \mathrm{~d} u \\
& =x \sum_{\nu=-\infty}^{\infty} \mathrm{e}^{-\nu^{2} \pi x} \int_{-\infty}^{\infty} \mathrm{e}^{-(u-\mathrm{i} \nu)^{2} \pi x} \mathrm{~d} u \tag{11}
\end{align*}
$$

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 residue theorem that

$$
\begin{equation*}
\int_{-\infty}^{\infty} \mathrm{e}^{-(u-\mathrm{i} \nu)^{2} \pi x} \mathrm{~d} u=\int_{-\infty}^{\infty} \mathrm{e}^{-u^{2} \pi x} \mathrm{~d} u=A x^{-1 / 2} \tag{12}
\end{equation*}
$$

where

$$
\begin{equation*}
A \int_{-\infty}^{\infty} \mathrm{e}^{-y^{2} \pi} \mathrm{~d} y=1 \tag{13}
\end{equation*}
$$

Combining (11)-(13), we conclude that

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

This gives (6), and the proof of Theorem 6.1 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 6.2 and 6.4.

An entire function $f(s)$ is said to be of order 1 if

$$
\begin{equation*}
f(s)=O_{\alpha}\left(\mathrm{e}^{|s|^{\alpha}}\right) \quad \text { as }|s| \rightarrow \infty \tag{14}
\end{equation*}
$$

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

$$
\begin{equation*}
h(s)=O_{\alpha}\left(\mathrm{e}^{|s|^{\alpha}}\right) \quad \text { as }|s| \rightarrow \infty \tag{15}
\end{equation*}
$$

holds for every $\alpha>1$. Then

$$
\mathfrak{R} g\left(R \mathrm{e}^{\mathrm{i} \theta}\right)=\log \left|h\left(R \mathrm{e}^{\mathrm{i} \theta}\right)\right|=O_{\alpha}\left(R^{\alpha}\right) \quad \text { as } R \rightarrow \infty
$$

holds for every $\alpha>1$. Without loss of generality, we may suppose that $g(0)=0$. Then we can write

$$
g\left(R \mathrm{e}^{\mathrm{i} \theta}\right)=\sum_{k=1}^{\infty}\left(a_{k}+\mathrm{i} b_{k}\right) R^{k} \mathrm{e}^{\mathrm{i} k \theta} \quad\left(a_{k}, b_{k} \in \mathbb{R}\right)
$$

so that

$$
\mathfrak{R} g\left(R \mathrm{e}^{\mathrm{i} \theta}\right)=\sum_{k=1}^{\infty} a_{k} R^{k} \cos k \theta-\sum_{k=1}^{\infty} b_{k} R^{k} \sin k \theta
$$

Note now that for every $k, n \in \mathbb{N}$, we have

$$
\int_{0}^{2 \pi} \cos k \theta \cos n \theta \mathrm{~d} \theta=\left\{\begin{array}{cl}
\pi & \text { if } k=n \\
0 & \text { if } k \neq n
\end{array}\right.
$$

and

$$
\int_{0}^{2 \pi} \sin k \theta \cos n \theta \mathrm{~d} \theta=0
$$

It follows that

$$
\int_{0}^{2 \pi}\left(\Re g\left(R \mathrm{e}^{\mathrm{i} \theta}\right)\right) \cos n \theta \mathrm{~d} \theta=\pi a_{n} R^{n}
$$

so that

$$
\pi\left|a_{n}\right| R^{n} \leq \int_{0}^{2 \pi}\left|\Re g\left(R \mathrm{e}^{\mathrm{i} \theta}\right)\right| \mathrm{d} \theta=O_{\alpha}\left(R^{\alpha}\right) \quad \text { as } R \rightarrow \infty
$$

holds for every $\alpha>1$. On letting $R \rightarrow \infty$, 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 6.6. Suppose that the entire function $h(s)$ has no zeros on the plane, and that (15) holds for every $\alpha>1$. Then $h(s)=\mathrm{e}^{A+B s}$, 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 6.7. (Jensen's Formula) Suppose that $f(s)$ is an entire function satisfying $f(0) \neq 0$. 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

$$
\begin{equation*}
\frac{1}{2 \pi} \int_{0}^{2 \pi} \log \left|f\left(R \mathrm{e}^{\mathrm{i} \theta}\right)\right| \mathrm{d} \theta-\log |f(0)|=\log \frac{R^{n}}{\left|s_{1} \ldots s_{n}\right|} \tag{16}
\end{equation*}
$$

Proof. We may clearly write

$$
f(s)=\left(s-s_{1}\right) \ldots\left(s-s_{n}\right) k(s)
$$

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

$$
\frac{1}{2 \pi} \int_{0}^{2 \pi} \log k\left(R \mathrm{e}^{\mathrm{i} \theta}\right) \mathrm{d} \theta=\log k(0)
$$

Taking real parts, we obtain

$$
\begin{equation*}
\frac{1}{2 \pi} \int_{0}^{2 \pi} \log \left|k\left(R \mathrm{e}^{\mathrm{i} \theta}\right)\right| \mathrm{d} \theta=\log |k(0)|=\log |f(0)|-\log \left|s_{1} \ldots s_{n}\right| \tag{17}
\end{equation*}
$$

Unfortunately, for every $j=1, \ldots, n$, we cannot apply a similar argument to $\log \left|s-s_{j}\right|$, 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

$$
\left|\frac{R^{2}-\overline{s_{j}} s}{R}\right|=\left|s-s_{j}\right|
$$

on the circle $|s|=R$, so that

$$
\begin{equation*}
\frac{1}{2 \pi} \int_{0}^{2 \pi} \log \left|R \mathrm{e}^{\mathrm{i} \theta}-s_{j}\right| \mathrm{d} \theta=\frac{1}{2 \pi} \int_{0}^{2 \pi} \log \left|\frac{R^{2}-\overline{s_{j}} R \mathrm{e}^{\mathrm{i} \theta}}{R}\right| \mathrm{d} \theta \tag{18}
\end{equation*}
$$

Clearly the function

$$
\log \frac{R^{2}-\overline{s_{j}} s}{R}
$$

is analytic in $|s| \leq R$. 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

$$
\log \left|f\left(R \mathrm{e}^{\mathrm{i} \theta}\right)\right|=\sum_{j=1}^{n} \log \left|R \mathrm{e}^{\mathrm{i} \theta}-s_{j}\right|+\log \left|k\left(R \mathrm{e}^{\mathrm{i} \theta}\right)\right|
$$

so that

$$
\frac{1}{2 \pi} \int_{0}^{2 \pi} \log \left|f\left(R \mathrm{e}^{\mathrm{i} \theta}\right)\right| \mathrm{d} \theta=n \log R+\log |f(0)|-\log \left|s_{1} \ldots s_{n}\right|,
$$

and the result follows.

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 6.2 and 6.4.
(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

$$
F(0)=\frac{1}{2 \pi} \int_{0}^{2 \pi} F\left(R \mathrm{e}^{\mathrm{i} \theta}\right) \mathrm{d} \theta
$$

A simple consequence of Jensen's formula is the following result on the zeros of entire functions.
THEOREM 6.8. Suppose 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 $\left|s_{1}\right| \leq\left|s_{2}\right| \leq\left|s_{3}\right| \leq \ldots$. Then for every $\alpha>1$, the series

$$
\sum_{n=1}^{\infty}\left|s_{n}\right|^{-\alpha}
$$

is convergent.
Proof. Note that the right hand side of (16) is equal to

$$
\int_{0}^{R} r^{-1} n(r) \mathrm{d} r
$$

where, for every non-negative $r \leq R, n(r)$ denotes the number of zeros of $f(s)$ in $|s| \leq r$. To see this, note that

$$
\begin{aligned}
\int_{0}^{R} r^{-1} n(r) \mathrm{d} r & =\sum_{j=1}^{n-1} \int_{\left|s_{j}\right|}^{\left|s_{j+1}\right|} r^{-1} j \mathrm{~d} r+\int_{\left|s_{n}\right|}^{R} r^{-1} n \mathrm{~d} r \\
& =\sum_{j=1}^{n-1} j\left(\log \left|s_{j+1}\right|-\log \left|s_{j}\right|\right)+n\left(\log R-\log \left|s_{n}\right|\right) \\
& =n \log R-\log \left|s_{1}\right|-\ldots-\log \left|s_{n}\right|
\end{aligned}
$$

For every $\alpha>1$, write $\alpha^{*}=(\alpha+1) / 2$, so that $1<\alpha^{*}<\alpha$. Then

$$
\log \left|f\left(R \mathrm{e}^{\mathrm{i} \theta}\right)\right|=O_{\alpha}\left(R^{\alpha^{*}}\right) \quad \text { as } R \rightarrow \infty
$$

so that by Jensen's formula, we have

$$
\int_{0}^{R} r^{-1} n(r) \mathrm{d} r=O_{\alpha}\left(R^{\alpha^{*}}\right)-\log |f(0)|=O_{\alpha}\left(R^{\alpha^{*}}\right) \quad \text { as } R \rightarrow \infty .
$$

On the other hand, note that

$$
\int_{R}^{2 R} r^{-1} n(r) \mathrm{d} r \geq n(R) \int_{R}^{2 R} r^{-1} \mathrm{~d} r=n(R) \log 2
$$

It follows that

$$
n(R)=O_{\alpha}\left(R^{\alpha^{*}}\right) \quad \text { as } R \rightarrow \infty
$$

Now

$$
\sum_{n=1}^{\infty}\left|s_{n}\right|^{-\alpha}=\int_{0}^{\infty} r^{-\alpha} \mathrm{d} n(r)=\alpha \int_{0}^{\infty} r^{-\alpha-1} n(r) \mathrm{d} r<\infty
$$

as required.
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 $\left|s_{1}\right| \leq\left|s_{2}\right| \leq\left|s_{3}\right| \leq \ldots$. Then for every $\epsilon>0$, the series

$$
\sum_{n=1}^{\infty}\left|s_{n}\right|^{-1-\epsilon}
$$

$\qquad$
converges, so that the series

$$
\sum_{n=1}^{\infty}\left|s_{n}\right|^{-2}
$$

converges, and so the product

$$
\begin{equation*}
P(s)=\prod_{n=1}^{\infty}\left(1-\frac{s}{s_{n}}\right) \mathrm{e}^{s / s_{n}} \tag{19}
\end{equation*}
$$

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

$$
\begin{equation*}
f(s)=P(s) h(s) \tag{20}
\end{equation*}
$$

where $h(s)$ is an entire function without zeros. If (15) holds for every $\alpha>1$, then $h(s)=\mathrm{e}^{A+B s}$, where $A$ and $B$ are constants, and so

$$
\begin{equation*}
f(s)=\mathrm{e}^{A+B s} \prod_{n=1}^{\infty}\left(1-\frac{s}{s_{n}}\right) \mathrm{e}^{s / s_{n}} . \tag{21}
\end{equation*}
$$

THEOREM 6.9. Under the hypotheses of Theorem 6.8, 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

$$
\sum_{n=1}^{\infty}\left|s_{n}\right|^{-2}
$$

is convergent, the set

$$
\mathcal{S}=\bigcup_{n=1}^{\infty}\left(\left|s_{n}\right|-\left|s_{n}\right|^{-2},\left|s_{n}\right|+\left|s_{n}\right|^{-2}\right)
$$

has finite total length. It follows that there exist arbitrarily large positive values $R$ such that $R \notin \mathcal{S}$, so that

$$
\begin{equation*}
\left|R-\left|s_{n}\right| \| \geq\left|s_{n}\right|^{-2} \quad \text { for every } n \in \mathbb{N} .\right. \tag{22}
\end{equation*}
$$

For any such $R$, write, for $j=1,2,3$,

$$
P_{j}(s)=\prod_{(23 . j)}\left(1-\frac{s}{s_{n}}\right) \mathrm{e}^{s / s_{n}},
$$

where the products are taken over all $n \in \mathbb{N}$ satisfying

$$
\begin{gather*}
\left|s_{n}\right|<\frac{R}{2}  \tag{23.1}\\
\frac{R}{2} \leq\left|s_{n}\right|<2 R,  \tag{23.2}\\
\left|s_{n}\right| \geq 2 R, \tag{23.3}
\end{gather*}
$$

respectively. Clearly

$$
\begin{equation*}
P(s)=P_{1}(s) P_{2}(s) P_{3}(s) . \tag{24}
\end{equation*}
$$

Let $\epsilon>0$ be chosen and fixed. Suppose first of all that (23.1) holds. Then on $|s|=R$, we have

$$
\left|\left(1-\frac{s}{s_{n}}\right) \mathrm{e}^{s / s_{n}}\right| \geq\left(\left|\frac{s}{s_{n}}\right|-1\right) \mathrm{e}^{-|s| /\left|s_{n}\right|}>\mathrm{e}^{-R /\left|s_{n}\right|}
$$

and so it follows from

$$
\sum_{(23.1)}\left|s_{n}\right|^{-1}<\left(\frac{R}{2}\right)^{\epsilon} \sum_{n=1}^{\infty}\left|s_{n}\right|^{-1-\epsilon}
$$

that

$$
\begin{equation*}
\left|P_{1}(s)\right| \gg_{\epsilon} \mathrm{e}^{-R^{1+2 \epsilon}} \quad \text { as } R \rightarrow \infty \tag{25}
\end{equation*}
$$

Suppose next that (23.2) holds. Then on $|s|=R$, we have

$$
\left|\left(1-\frac{s}{s_{n}}\right) \mathrm{e}^{s / s_{n}}\right| \geq\left|\frac{s_{n}-s}{s_{n}}\right| \mathrm{e}^{-|s| /\left|s_{n}\right|}>\frac{\left|\left|s_{n}\right|-R\right|}{2 R} \mathrm{e}^{-2} \gg R^{-3}
$$

in view of (22). Note that there are at most $O_{\epsilon}\left(R^{1+\epsilon}\right)$ values of $n$ for which (23.2) holds. Hence on $|s|=R$, we have

$$
\begin{equation*}
\left|P_{2}(s)\right| \gg_{\epsilon}\left(R^{-3}\right)^{R^{1+\epsilon}} \gg_{\epsilon} \mathrm{e}^{-R^{1+2 \epsilon}} \quad \text { as } R \rightarrow \infty \tag{26}
\end{equation*}
$$

Suppose finally that (23.3) holds. Then on $|s|=R$, we have

$$
\begin{equation*}
\left|\left(1-\frac{s}{s_{n}}\right) \mathrm{e}^{s / s_{n}}\right|>\mathrm{e}^{-c\left(R /\left|s_{n}\right|\right)^{2}} \tag{27}
\end{equation*}
$$

for some positive constant $c$ (see the Remark below), and so it follows from

$$
\sum_{(23.3)}\left|s_{n}\right|^{-2} \leq(2 R)^{-1+\epsilon} \sum_{n=1}^{\infty}\left|s_{n}\right|^{-1-\epsilon}
$$

that

$$
\begin{equation*}
\left|P_{3}(s)\right| \gg_{\epsilon} \mathrm{e}^{-R^{1+2 \epsilon}} \quad \text { as } R \rightarrow \infty \tag{28}
\end{equation*}
$$

It now follows from (24)-(26) and (28) that on $|s|=R$, we have

$$
\begin{equation*}
|P(s)| \gg_{\epsilon} \mathrm{e}^{-R^{1+3 \epsilon}} \quad \text { as } R \rightarrow \infty \tag{29}
\end{equation*}
$$

The result now 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

$$
\begin{equation*}
\left|(1-z) \mathrm{e}^{z}\right|>\mathrm{e}^{-c|z|^{2}} \tag{30}
\end{equation*}
$$

where $|z| \leq 1 / 2$. Write $z=x+\mathrm{i} y$, where $x, y \in \mathbb{R}$. Then (30) will follow if we show that

$$
(1-x)^{2} \mathrm{e}^{2 x}>\mathrm{e}^{-2 c x^{2}}
$$

whenever $|x| \leq 1 / 2$. 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 6.10. Under the hypotheses of Theorem 6.8, suppose further that the series

$$
\sum_{n=1}^{\infty}\left|s_{n}\right|^{-1}
$$

is convergent. Then there exists a positive constant $c$ such that

$$
f(s)=O\left(\mathrm{e}^{c|s|}\right) \quad \text { as }|s| \rightarrow \infty
$$

holds.
Proof. This follows from (21) and the inequality $\left|(1-z) \mathrm{e}^{z}\right| \leq \mathrm{e}^{2|z|}$ which holds for every $z \in \mathbb{C}$.

### 6.4. Proof of Theorems 6.2 and 6.4

Recall that $\xi(s)$, defined by (3), 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 6.4 , we shall use Theorem 6.9. We therefore first need to show that

$$
\xi(s)=O_{\alpha}\left(\mathrm{e}^{|s|^{\alpha}}\right) \quad \text { as }|s| \rightarrow \infty
$$

holds for every $\alpha>1$.
We shall in fact prove the following stronger result.
THEOREM 6.11. There exists a positive constant $c$ such that

$$
\begin{equation*}
|\xi(s)|<\mathrm{e}^{c|s| \log |s|} \quad \text { as }|s| \rightarrow \infty \tag{31}
\end{equation*}
$$

Furthermore, the inequality

$$
\begin{equation*}
|\xi(s)|<\mathrm{e}^{c|s|} \quad \text { as }|s| \rightarrow \infty \tag{32}
\end{equation*}
$$

does not hold for any positive constant $c$.
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

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

as $|s| \rightarrow \infty$ 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 $s \rightarrow+\infty$ through real values, we have

$$
\log \Gamma\left(\frac{s}{2}\right) \sim \frac{s}{2} \log \frac{s}{2} \quad \text { and } \quad \zeta(s) \rightarrow 1
$$

so that (32) does not hold.
To complete the proof of Theorems 6.2 and 6.4 , note that by Theorem 6.10, the series

$$
\sum_{\rho}|\rho|^{-1}
$$

is divergent, where $\rho$ denotes the zeros of $\xi(s)$ and so the zeros of $\zeta(s)$ in the critical strip. Theorem 6.2 follows immediately. Theorem 6.4 now follows from Theorems 6.9 and 6.11 .

### 6.5. An Important Formula

It follows from (4) that

$$
\log \xi(s)=A+B s+\sum_{\rho}\left(\frac{s}{\rho}+\log \left(1-\frac{s}{\rho}\right)\right) .
$$

Differentiating with respect to $s$, we obtain

$$
\begin{equation*}
\frac{\xi^{\prime}(s)}{\xi(s)}=B+\sum_{\rho}\left(\frac{1}{\rho}+\frac{1}{s-\rho}\right) . \tag{33}
\end{equation*}
$$

On the other hand, it follows from (3) that

$$
\begin{aligned}
\log \xi(s) & =\log (s-1)-\frac{s}{2} \log \pi+\log \left(\frac{s}{2} \Gamma\left(\frac{s}{2}\right)\right)+\log \zeta(s) \\
& =\log (s-1)-\frac{s}{2} \log \pi+\log \Gamma\left(\frac{s}{2}+1\right)+\log \zeta(s) .
\end{aligned}
$$

Differentiating with respect to $s$, we obtain

$$
\begin{equation*}
\frac{\xi^{\prime}(s)}{\xi(s)}=\frac{1}{s-1}-\frac{1}{2} \log \pi+\frac{1}{2} \frac{\Gamma^{\prime}\left(\frac{s}{2}+1\right)}{\Gamma\left(\frac{s}{2}+1\right)}+\frac{\zeta^{\prime}(s)}{\zeta(s)} \tag{34}
\end{equation*}
$$

Combining (33) and (34), we obtain the following result.
THEOREM 6.12. We have

$$
\begin{equation*}
\frac{\zeta^{\prime}(s)}{\zeta(s)}=B-\frac{1}{s-1}+\frac{1}{2} \log \pi-\frac{1}{2} \frac{\Gamma^{\prime}\left(\frac{s}{2}+1\right)}{\Gamma\left(\frac{s}{2}+1\right)}+\sum_{\rho}\left(\frac{1}{\rho}+\frac{1}{s-\rho}\right) \tag{35}
\end{equation*}
$$

where $B$ is a constant and where $\rho$ denotes the zeros of $\zeta(s)$ in the critical strip.
The formula (35) exhibits the pole of $\zeta(s)$ at $s=1$ and the zeros $\rho$ in the critical strip. The trivial zeros are exhibited by the term

$$
-\frac{1}{2} \frac{\Gamma^{\prime}\left(\frac{s}{2}+1\right)}{\Gamma\left(\frac{s}{2}+1\right)}
$$

To see this last point, we start from 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 $\gamma$ is Euler's constant. This gives

$$
\frac{1}{\Gamma\left(\frac{s}{2}+1\right)}=\mathrm{e}^{\gamma s / 2} \prod_{n=1}^{\infty}\left(1+\frac{s}{2 n}\right) \mathrm{e}^{-s / 2 n}
$$

Taking logarithms, we obtain

$$
-\log \Gamma\left(\frac{s}{2}+1\right)=\frac{1}{2} \gamma s+\sum_{n=1}^{\infty}\left(\log \left(1+\frac{s}{2 n}\right)-\frac{s}{2 n}\right)
$$

Differentiating with respect to $s$, we obtain

$$
-\frac{1}{2} \frac{\Gamma^{\prime}\left(\frac{s}{2}+1\right)}{\Gamma\left(\frac{s}{2}+1\right)}=\frac{1}{2} \gamma+\sum_{n=1}^{\infty}\left(\frac{1}{s+2 n}-\frac{1}{2 n}\right)
$$

### 6.6. 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

$$
\frac{1}{2 \pi \mathrm{i}} \int_{C} \frac{F^{\prime}(s)}{F(s)} \mathrm{d} s=\frac{1}{2 \pi} \Delta_{C} \arg F(s)
$$

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

$$
N(T)=\frac{1}{2 \pi \mathrm{i}} \int_{C} \frac{\xi^{\prime}(s)}{\xi(s)} \mathrm{d} s=\frac{1}{2 \pi} \Delta_{C} \arg \xi(s)
$$

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 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

$$
\begin{equation*}
\pi N(T)=\Delta_{L} \arg \xi(s) \tag{36}
\end{equation*}
$$

Recall that

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

It follows that

$$
\begin{equation*}
\Delta_{L} \arg \xi(s)=\Delta_{L} \arg (s-1)+\Delta_{L} \arg \pi^{-s / 2}+\Delta_{L} \arg \Gamma\left(\frac{s}{2}+1\right)+\Delta_{L} \arg \zeta(s) \tag{37}
\end{equation*}
$$

Clearly

$$
\begin{equation*}
\Delta_{L} \arg (s-1)=\arg \left(-\frac{1}{2}+\mathrm{i} T\right)=\frac{1}{2} \pi+O\left(T^{-1}\right) \tag{38}
\end{equation*}
$$

and

$$
\begin{equation*}
\Delta_{L} \arg \pi^{-s / 2}=\Delta_{L}\left(-\frac{1}{2} t \log \pi\right)=-\frac{1}{2} T \log \pi \tag{39}
\end{equation*}
$$

On the other hand,

$$
\Delta_{L} \arg \Gamma\left(\frac{s}{2}+1\right)=\Im \log \Gamma\left(\frac{5}{4}+\frac{1}{2} \mathrm{i} T\right)
$$

By Stirling's formula,

$$
\log \Gamma\left(\frac{5}{4}+\frac{1}{2} \mathrm{i} T\right)=\left(\frac{3}{4}+\frac{1}{2} \mathrm{i} T\right) \log \left(\frac{5}{4}+\frac{1}{2} \mathrm{i} T\right)-\frac{5}{4}-\frac{1}{2} \mathrm{i} T+\frac{1}{2} \log \pi+O\left(T^{-1}\right)
$$

so that

$$
\begin{equation*}
\Delta_{L} \arg \Gamma\left(\frac{s}{2}+1\right)=\frac{T}{2} \log \frac{T}{2}+\frac{3}{8} \pi-\frac{T}{2}+O\left(T^{-1}\right) \tag{40}
\end{equation*}
$$

Combining (36)-(40), we have

$$
\begin{aligned}
N(T) & =\frac{1}{2}-\frac{T}{2 \pi} \log \pi+\frac{T}{2 \pi} \log \frac{T}{2}+\frac{3}{8}-\frac{T}{2 \pi}+S(T)+O\left(T^{-1}\right) \\
& =\frac{T}{2 \pi} \log \frac{T}{2 \pi}-\frac{T}{2 \pi}+\frac{7}{8}+S(T)+O\left(T^{-1}\right)
\end{aligned}
$$

where

$$
\pi S(T)=\Delta_{L} \arg \zeta(s)
$$

To prove Theorem 6.3, it suffices to prove the following result.
THEOREM 6.13. We have $S(T)=O(\log T)$ as $T \rightarrow \infty$.
Proof. Note first of all that $\arg \zeta(2)=0$. On the other hand,

$$
\arg \zeta(s)=\tan ^{-1}\left(\frac{\Im \zeta(s)}{\mathfrak{R} \zeta(s)}\right)
$$

and $\mathfrak{R} \zeta(s) \neq 0$ on the line $\sigma=2$. It follows that

$$
|\arg \zeta(2+\mathrm{i} T)|<\frac{\pi}{2}
$$

Suppose now that $\mathfrak{R} \zeta(s)$ 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, $\mathfrak{R} \zeta(s)$ 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 $\pi$. It follows that for $s=\sigma+\mathrm{i} T$, we have

$$
\begin{equation*}
|\arg \zeta(s)| \leq(q+1) \pi+\frac{1}{2} \pi \quad\left(\frac{1}{2} \leq \sigma \leq 2\right) \tag{41}
\end{equation*}
$$

To prove our result, it remains to find a suitable bound for $q$. For $s=\sigma+\mathrm{i} T$, we have

$$
\mathfrak{R} \zeta(s)=\frac{1}{2}(\zeta(\sigma+\mathrm{i} T)+\zeta(\sigma-\mathrm{i} T)) .
$$

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,

$$
\begin{equation*}
q \leq n\left(\frac{3}{2}\right) \tag{42}
\end{equation*}
$$

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 $\zeta\left(\frac{1}{2}+\mathrm{i} T\right) \neq 0$, we have

$$
\begin{equation*}
\int_{0}^{7 / 4} \frac{n(r)}{r} \mathrm{~d} r=\frac{1}{2 \pi} \int_{0}^{2 \pi} \log \left|f_{T}\left(2+\frac{7}{4} \mathrm{e}^{\mathrm{i} \theta}\right)\right| \mathrm{d} \theta-\log \left|f_{T}(2)\right| \tag{43}
\end{equation*}
$$

On the other hand,

$$
\begin{equation*}
\int_{0}^{7 / 4} \frac{n(r)}{r} \mathrm{~d} r \geq \int_{3 / 2}^{7 / 4} \frac{n(r)}{r} \mathrm{~d} r \geq n\left(\frac{3}{2}\right) \int_{3 / 2}^{7 / 4} \frac{1}{r} \mathrm{~d} r=n\left(\frac{3}{2}\right) \log \frac{7}{6} \tag{44}
\end{equation*}
$$

Observe that

$$
\begin{aligned}
\left|f_{T}(2)\right| & =\left|\frac{1}{2}(\zeta(2+\mathrm{i} T)+\zeta(2-\mathrm{i} T))\right|=|\Re \zeta(2+\mathrm{i} T)| \\
& =\left|\Re\left(\sum_{n=1}^{\infty} \frac{1}{n^{2+\mathrm{i} T}}\right)\right| \geq 1-\sum_{n=2}^{\infty} \frac{1}{n^{2}}=2-\frac{\pi^{2}}{6}>0
\end{aligned}
$$

so that

$$
\begin{equation*}
-\log \left|f_{T}(2)\right|=O(1) \tag{45}
\end{equation*}
$$

Finally, recall that $|\zeta(s)| \ll T^{3 / 4}$ for every $\sigma \geq 1 / 4$. It follows that for every $\theta \in[0,2 \pi]$, we have

$$
\left|f_{T}\left(2+\frac{7}{4} \mathrm{e}^{\mathrm{i} \theta}\right)\right| \leq \frac{1}{2}\left(\left|\zeta\left(2+\frac{7}{4} \mathrm{e}^{\mathrm{i} \theta}+\mathrm{i} T\right)\right|+\left|\zeta\left(2+\frac{7}{4} \mathrm{e}^{\mathrm{i} \theta}-\mathrm{i} T\right)\right|\right) \ll T^{3 / 4}
$$

so that

$$
\begin{equation*}
\log \left|f_{T}\left(2+\frac{7}{4} \mathrm{e}^{\mathrm{i} \theta}\right)\right| \ll \log T \tag{46}
\end{equation*}
$$

Combining (43)-(46), we conclude that

$$
\begin{equation*}
n\left(\frac{3}{2}\right) \ll \log T \tag{47}
\end{equation*}
$$

The result now follows on combining (41), (42) and (47).

### 6.7. Sketch of the Proof of Theorem 6.5

It can be shown that for $c>1$, we have

$$
\frac{1}{2 \pi \mathrm{i}} \int_{c-\mathrm{i} \infty}^{c+\mathrm{i} \infty} \frac{y^{s}}{s} \mathrm{~d} s= \begin{cases}0 & \text { if } 0<y<1 \\ 1 / 2 & \text { if } y=1 \\ 1 & \text { if } y>1\end{cases}
$$

It follows that

$$
\begin{aligned}
\psi_{0}(X) & =\sum_{n \leq X}^{\prime} \Lambda(n)=\sum_{n=1}^{\infty} \Lambda(n)\left(\frac{1}{2 \pi \mathrm{i}} \int_{c-\mathrm{i} \infty}^{c+\mathrm{i} \infty} \frac{(X / n)^{s}}{s} \mathrm{~d} s\right) \\
& =\frac{1}{2 \pi \mathrm{i}} \int_{c-\mathrm{i} \infty}^{c+\mathrm{i} \infty}\left(\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{s}}\right) \frac{X^{s}}{s} \mathrm{~d} s=-\frac{1}{2 \pi \mathrm{i}} \int_{c-\mathrm{i} \infty}^{c+\mathrm{i} \infty} \frac{\zeta^{\prime}(s)}{\zeta(s)} \frac{X^{s}}{s} \mathrm{~d} s
\end{aligned}
$$

where $\sum^{\prime}$ denotes that the term in the sum corresponding to $n=X$ is $\frac{1}{2} \Lambda(X)$.
The idea is now to move the line of integration away to infinity on the left, so that we can express $\psi_{0}(X)$ as a sum of the residues at the poles of the function

$$
-\frac{\zeta^{\prime}(s)}{\zeta(s)} \frac{X^{s}}{s}
$$

Here, the pole at $s=1$ gives rise to a residue $X$, while the pole at $s=0$ gives rise to a residue $-\zeta^{\prime}(0) / \zeta(0)$. On the other hand, each zero $\rho$ of $\zeta(s)$ in the critical strip gives rise to a residue $-X^{\rho} / \rho$, while each trivial zero $-2 n$ of $\zeta(s)$ gives rise to a residue $X^{-2 n} / 2 n$. Finally, note that

$$
-\frac{1}{2} \log \left(1-\frac{1}{X^{2}}\right)=\sum_{n=1}^{\infty} \frac{X^{-2 n}}{2 n} .
$$

The first step in this process is to consider the integral

$$
-\frac{1}{2 \pi \mathrm{i}} \int_{c-\mathrm{i} T}^{c+\mathrm{i} T} \frac{\zeta^{\prime}(s)}{\zeta(s)} \frac{X^{s}}{s} \mathrm{~d} s
$$

and regard the path of integration as one side of a rectangle with vertices at $c \pm \mathrm{i} T$ and $-U \pm \mathrm{i} T$, where $U>0$ is large. Here $T$ has to be chosen carefully so that the horizontal sides of the rectangle should avoid the zeros of $\zeta(s)$ in the critical strip. Also $U$ has to be chosen carefully so that the left vertical side of the rectangle should avoid the trivial zeros of $\zeta(s)$. On taking $U \rightarrow \infty$, this will result in a finite version of Theorem 6.5, of the form

$$
\psi_{0}(X)=X-\sum_{|\gamma|<T} \frac{X^{\rho}}{\rho}-\frac{\zeta^{\prime}(0)}{\zeta(0)}-\frac{1}{2} \log \left(1-\frac{1}{X^{2}}\right)+R(X, T)
$$

where the error term $R(X, T)$ can be estimated.


[^0]:    $\dagger$ This chapter was first used in lectures given by the author at Imperial College, University of London, in 1981.

[^1]:    $\dagger$ This chapter was first used in lectures given by the author at Imperial College, University of London, in 1981.

[^2]:    $\dagger$ This chapter was first used in lectures given by the author at Imperial College, University of London, in 1990.

[^3]:    $\dagger$ This chapter was first used in lectures given by the author at Imperial College, University of London, in 1990.

[^4]:    $\dagger$ This chapter was first used in lectures given by the author at Imperial College, University of London, in 1990.

[^5]:    $\dagger$ This chapter was first used in lectures given by the author at Imperial College, University of London, in 1990.

