DISTRIBUTION OF PRIME NUMBERS

W W L CHEN

Chapter 1

ARITHMETIC FUNCTIONS

1.1. Introduction

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

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

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

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

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

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

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

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

This completes the proof. ○

1.2. The Divisor Function

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

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

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

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

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

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

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

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 $0 \le v_j \le u_j$. It follows from (1) that $d(n)$ is the number of choices for the r-tuple $(v_1, \ldots, v_r)$. Hence

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

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 assertion follows. ○

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

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

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

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

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

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

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

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

It follows from Theorems 1B and 1C that

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

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

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

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

so that

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

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

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

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

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

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

Then in view of Theorem 1B, we have

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

where the positive constant

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

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

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

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

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

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

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

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

so that

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

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

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

It follows that

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

a positive constant depending only on ϵ. ○

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

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

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

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

Here γ is Euler's constant and is defined by

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

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

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

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

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

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

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

It is a simple exercise to show that

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

and this completes the proof. ○

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

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

This completes the proof. ○

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

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

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

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

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

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

This completes the proof. ○

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

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

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

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

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

This completes the proof. ○

1.3. The Möbius Function

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

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

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

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

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

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

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

It follows that the product is identically equal to 1 if

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

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

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

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

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

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

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

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

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

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

This completes the proof. ○

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

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

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

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

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

PROOF. The second equality is obvious. Also

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

in view of Theorem 1M. ○

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

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

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

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

Proof. The second equality is obvious. Also

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

in view of Theorem 1M. ○

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

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

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

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

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

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

Summing over all positive divisors k of n, we obtain

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

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

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

1.4. The Euler Function

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

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

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

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

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

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

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

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

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

This completes the proof. ○

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

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

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

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

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

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

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

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

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

The result now follows since ϕ is multiplicative. ○

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

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

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

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

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

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

and

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

Hence

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

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

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

as required. ○

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

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

We now consider some average version of the Euler function.

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

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

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

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

It remains to show that

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

But

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

in view of Theorem 1M. ○

1.5. Dirichlet Convolution

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

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

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

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

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

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

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

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

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

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

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

and

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

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

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

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

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

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

To prove (iii), note that

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

This completes the proof of Theorem 1Y. ○

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

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

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

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

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

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

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

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

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

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

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

so that g is an inverse of f. ○

Problems for Chapter 1

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

12. By first showing that the function $f(n) = (-1)^{n-1}$ is multiplicative, evaluate the sum
$$h(n) = \sum_{m \mid n} (-1)^{m-1} \mu\Big(\frac{n}{m}\Big) \qquad \text{for every } n \in \mathbb{N}.$$

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

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

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

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

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

18. Suppose that $n = p_1^{u_1} \ldots p_r^{u_r}$, where $p_1 < \ldots < p_r$ are primes and $u_1, \ldots, u_r \in \mathbb{N}$.
    (i) Write
$$s(n) = \sum_{\substack{m=1 \\ (m,n)=1}}^{n} m^2.$$
    Prove that
$$n^2 \sum_{d \mid n} \frac{s(d)}{d^2} = \frac{n(n+1)(2n+1)}{6}.$$
    (ii) Apply the Möbius inversion formula to deduce that
$$\sum_{\substack{m=1 \\ (m,n)=1}}^{n} m^2 = \frac{1}{3}\phi(n)n^2 + \frac{1}{6}(-1)^r \phi(n) p_1 \ldots p_r.$$

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

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

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

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

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

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

Prove that

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

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

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

Prove that

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

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

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

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

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

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

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

then

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

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

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

This is a generalization of Problems 21 and 22.]

25. For every $n\in\mathbb{N}$, let $f(n)=\displaystyle\sum_{m\mid n}\frac{\mu^2(m)}{\phi(m)}$.

(i) Prove that $f(n)=n/\phi(n)$ for every $n\in\mathbb{N}$.

(ii) Deduce that for every real number $X\ge 1$, we have $\displaystyle\sum_{n\le X}\frac{1}{\phi(n)}=\sum_{m\le X}\frac{\mu^2(m)}{m\phi(m)}\sum_{t\le X/m}\frac{1}{t}$.

(iii) Show that the series $\displaystyle\sum_{m=1}^{\infty}\frac{\mu^2(m)}{m\phi(m)}$ and $\displaystyle\sum_{m=1}^{\infty}\frac{\mu^2(m)\log m}{m\phi(m)}$ both converge.

(iv) Deduce that as $X\to\infty$, we have $\displaystyle\sum_{n\le X}\frac{1}{\phi(n)}\sim C\log X$, where $C=\displaystyle\sum_{m=1}^{\infty}\frac{\mu^2(m)}{m\phi(m)}$.

26. Consider a square lattice consisting of all points (a,b), where $a,b\in\mathbb{Z}$. Two lattice points P and Q are said to be mutually visible if the line segment which joins them contains no lattice points other than the endpoints P and Q.

(i) Prove that (a,b) and $(0,0)$ are mutually visible if and only if a and b are relatively prime.

(ii) We shall prove that the set of lattice points visible from the origin has density $6/\pi^2$. Consider a large square region on the xy-plane defined by the inequalities $|x|\le r$ and $|y|\le r$. Let $N(r)$ denote the number of lattice points in this square, and let $N'(r)$ denote the number of these which are visible from the origin. The eight lattice points nearest the origin are all visible from the origin. By symmetry, $N'(r)$ is equal to 8 plus 8 times the number of visible points in the region $\{(x,y):2\le x\le r \text{ and } 1\le y\le x\}$. Prove that

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

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

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