Chapter 1

Euler's Product Formula

1.1 The Product Formula

The whole of analytic number theory rests on one marvellous formula due to Leonhard Euler (1707-1783):

$$\sum_{n \in \mathbb{N},\, n>0} n^{-s} = \prod_{\text{primes } p} \left(1 - p^{-s}\right)^{-1}.$$

Informally, we can understand the formula as follows. By the Fundamental Theorem of Arithmetic, each $n \geq 1$ is uniquely expressible in the form

$$n = 2^{e_2} 3^{e_3} 5^{e_5} \cdots ,$$

where $e_2, e_3, e_5, \ldots$ are non-negative integers (all but a finite number being 0).

Raising this to the power $-s$,

$$n^{-s} = 2^{-e_2 s} 3^{-e_3 s} 5^{-e_5 s} \cdots .$$

Adding for $n = 1, 2, 3, \ldots$,

$$\sum n^{-s} = \left(1 + 2^{-s} + 2^{-2s} + \cdots\right) \left(1 + 3^{-s} + 3^{-2s} + \cdots\right) \left(1 + 5^{-s} + 5^{-2s} + \cdots\right) \cdots ,$$

each term on the left arising from just one product on the right. But for each prime p,

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

and the result follows.

Euler's Product Formula equates the *Dirichlet series* $\sum n^{-s}$ on the left with the *infinite product* on the right.

To make the formula precise, we must develop the theory of infinite products, which we do in the next Section.

To understand the implications of the formula, we must develop the theory of Dirichlet series, which we do in the next Chapter.

1.2 Infinite products

1.2.1 Definition

Infinite products are less familiar than infinite series, but are no more complicated. Both are examples of limits of *sequences.*

Definition 1.1. *The infinite product*

$$\prod_{n\in\mathbb{N}} c_n$$

is said to converge to $\ell \neq 0$ *if the partial products*

$$P_n = \prod_{0\le m\le n} c_m \to \ell \text{ as } n \to \infty.$$

We say that the infinite product *diverges* if either the partial products do not converge, or else they converge to 0 (as would be the case for example if any factor were 0).

Proposition 1.1. *If* $\prod c_n$ *is convergent then*

$$c_n \to 1.$$

Proof ▶ We have

$$c_n = \frac{P_n}{P_{n-1}}.$$

Since $P_n \to \ell$ and $P_{n-1} \to \ell$, it follows that

$$c_n \to \frac{\ell}{\ell} = 1.$$

◀

It is usually more convenient to write the factors in the form

$$c_n = 1 + a_n.$$

In these terms, the last Proposition states that

$$\prod(1+a_n) \text{ convergent } \implies a_n \to 0.$$

1.2.2 The complex logarithm

The theory of infinite products requires some knowledge of the complex logarithmic function.

Suppose $z \neq 0$. Let
$$z = re^{i\theta},$$
where $r > 0$. We are interested in solutions $w \in \mathbb{C}$ of
$$e^w = z.$$
If $w = x + iy$ then
$$e^x = r,\ e^{-iy} = e^{-i\theta},$$
ie
$$x = \log r,\ y = \theta + 2n\pi$$
for some $n \in \mathbb{Z}$.

Just one of these solutions to $e^w = z$ satisfies
$$-\pi < y = \Im(w) \leq \pi.$$
We call this value of w the *principal logarithm* of z, and denote it by $\operatorname{Log} z$. Thus
$$e^{\operatorname{Log} z} = z,\ -\pi < \Im(z) \leq \pi.$$

The general solution of $e^w = z$ is
$$w = \operatorname{Log} z + 2n\pi i \qquad (n \in \mathbb{Z}).$$

Now suppose
$$w_1 = \operatorname{Log} z_1,\ w_2 = \operatorname{Log} z_2.$$
Then
$$e^{w_1+w_2} = z_1 z_2 = e^{\operatorname{Log}(z_1 z_2)}$$
It follows that
$$\operatorname{Log}(z_1 z_2) = \operatorname{Log} z_1 + \operatorname{Log} z_2 + 2n\pi i,$$
where it is easy to see that $n = 0, -1$ or 1.

If $\Re(z) > 0$ then $z = re^{i\theta}$ with $-\pi/2 < \theta < \pi/2$. It follows that
$$\Re(z_1), \Re(z_2) > 0 \implies -\pi/2 < \Im(\operatorname{Log} z_1), \Im(\operatorname{Log} z_2) < \pi/2;$$
and so
$$-\pi < \Im(\operatorname{Log} z_1 + \operatorname{Log} z_2) < \pi.$$
Thus
$$\Re(z_1), \Re(z_2) > 0 \implies \operatorname{Log}(z_1 z_2) = \operatorname{Log} z_1 + \operatorname{Log} z_2.$$
In particular, this holds if $|z_1|, |z_2| < 1$ (Fig 1.1).

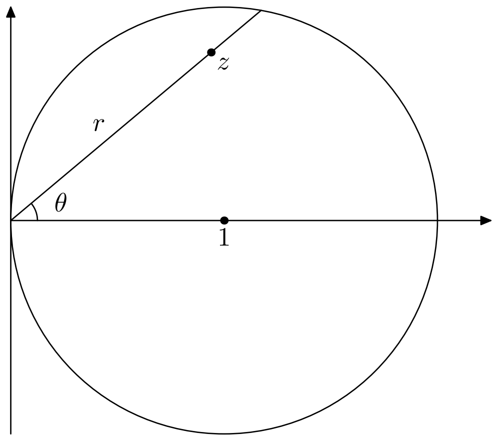

Figure 1.1: $|z-1| < 1,\ \operatorname{Log} z = \log r + i\theta$

1.2.3 Convergence

Proposition 1.2. *Suppose $a_n \neq -1$ for $n \in \mathbb{N}$. Then*

$$\prod (1+a_n) \ \textit{converges} \iff \sum \operatorname{Log}(1+a_n) \ \textit{converges.}$$

Proof ▶ Suppose $\sum \operatorname{Log}(1+a_n)$ converges to S. Let

$$S_n = \sum_{m \le n} \operatorname{Log}(1+a_m).$$

Then

$$e^{S_n} = \prod_{m \le n} (1+a_m).$$

But

$$S_n \to S \implies e^{S_n} \to e^S.$$

Thus $\prod(1+a_n)$ converges to e^s.

Conversely, suppose $\prod(1+a_n)$ converges. Let

$$P_n = \prod_{m \le N} (1+a_n).$$

Given $\epsilon > 0$ there exists N such that

$$|\frac{P_n}{P_m} - 1| < \epsilon$$

if $m, n \geq N$.

It follows that if $m, n \geq N$ then

$$\mathrm{Log}(P_n/P_N) = \mathrm{Log}(P_m/P_N) + \mathrm{Log}(P_n/P_m).$$

In particular (taking $m = n - 1$),

$$\mathrm{Log}(P_n/P_N) = \mathrm{Log}(P_{n-1}/P_N) + \mathrm{Log}(1 + a_n).$$

Hence

$$\mathrm{Log}(P_n/P_N) = \sum_{N<m\leq m} \mathrm{Log}(1 + a_m).$$

Since

$$P_n \to \ell \implies \mathrm{Log}(P_n/L_N) \to \mathrm{Log}(\ell/P_N),$$

we conclude that $\sum_{n>N}(1 + a_n)$ converges to $\mathrm{Log}(\ell/P_N)$; and in particular, $\sum_{n\geq 0} \mathrm{Log}(1 + a_n)$ is convergent. ◀

Proposition 1.3. *Suppose $a_n \neq -1$ for $n \in \mathbb{N}$. Then*

$$\sum |a_n| \textit{ convergent} \implies \prod (1 + a_n) \textit{ convergent.}$$

Proof ▶ The function $\mathrm{Log}(1 + z)$ is holomorphic in $|z| < 1$, with Taylor expansion

$$\mathrm{Log}(1 + z) = z - z^2/2 + z^3/3 - \cdots .$$

Thus if $|z| < 1/2$ then

$$\begin{aligned} |\mathrm{Log}(1 + z)| &\leq |z| + |z|^2 + |z|^3 + \cdots \\ &= \frac{|z|}{1 - |z|} \\ &\leq 2|z|. \end{aligned}$$

Now suppose $\sum |a_n|$ converges. Then $a_n \to 0$; and so

$$|a_n| \leq 1/2$$

for $n \geq N$. It follows that

$$|\mathrm{Log}(1 + a_n)| \leq 2|a_n|$$

for $n \geq N$. Hence

$$\sum \mathrm{Log}(1 + a_n) \text{ converges.}$$

◀

1.3 Proof of the product formula

Proposition 1.4. *For $\Re(s) > 1$,*

$$\sum_{n \in \mathbb{N},\, n>0} n^{-s} = \prod_{\textit{primes } p} \left(1 - p^{-s}\right)^{-1},$$

in the sense that each side converges to the same value.

Proof ▶ Let $\sigma = \Re(s)$. Then

$$|n^{-s}| = n^{-\sigma}.$$

Thus

$$|\sum_{M+1}^{N} n^{-s}| \le \sum_{M+1}^{N} n^{-\sigma}.$$

Now

$$n^{-\sigma} \le \int_{n-1}^{n} x^{-\sigma} dx;$$

and so

$$\begin{aligned}\sum_{M+1}^{N} n^{-\sigma} &\le \int_{M}^{N} x^{-\sigma} dx \\ &= \frac{1}{\sigma}\left(M^{-\sigma} - N^{-\sigma}\right) \\ &\to 0 \text{ as } M, N \to \infty.\end{aligned}$$

Hence $\sum n^{-s}$ is convergent, by Cauchy's criterion.

On the other hand,

$$\prod (1 - p^{-s})$$

is absolutely convergent, since

$$\sum |p^{-s}| = \sum p^{-\sigma} \le \sum n^{-\sigma},$$

which we just saw is convergent. Hence $\prod(1 - p^{-s})$ is convergent, by Proposition 1.3; and so therefore is

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

To see that the two sides are equal note that

$$\prod_{p \le N} \left(1 - \chi(p) p^{-s}\right)^{-1} = \sum_{n \le N} \chi(n) n^{-s} + {\sum}' \chi(n) n^{-s},$$

where the second sum on the right extends over those $n > N$ all of whose prime factors are $\le N$.

As $N \to \infty$, the right-hand side $\to \sum n^{-s}$, since this sum is absolutely convergent; while by definition, the left-hand side $\to \prod(1 - p^{-s})^{-1}$. We conclude that the two sides converge to the same value. ◀

1.4 Euler's Theorem

Proposition 1.5. *(Euler's Theorem)*

$$\sum_{primes\ p} \frac{1}{p} = \infty.$$

Proof ▶ Suppose $\sum 1/p$ is convergent. Then

$$\prod \left(1 - \frac{1}{p}\right)$$

is absolutely convergent, and so converges to ℓ say, by Proposition **??** It follows that

$$\prod_{p \le N} \left(1 - \frac{1}{p}\right) \to \ell^{-1}.$$

But

$$\sum_1^N \frac{1}{n} \le \prod_{p \le N} \left(1 - \frac{1}{p}\right),$$

since each n on the left is expressible in the form

$$n = p_1^{e_1} \cdots p_r^{e_r}$$

with $p_1, \ldots, p_r \le N$.

Hence $\sum 1/n$ is convergent. But

$$\frac{1}{n} > \int_n^{n+1} \frac{dx}{x}.$$

Thus

$$\sum_1^N n^{-1} \ge \int_1^{N+1} \frac{dx}{x} = \log(N+1).$$

Since $\log N \to \infty$ as $N \to \infty$ it follows that $\sum 1/n$ is divergent.

Our hypothesis is therefore untenable, and

$$\sum \frac{1}{p} \text{ diverges.}$$

◀

This is a remarkably accurate result; $\sum \frac{1}{p}$ only just diverges. For it follows from the Prime Number Theorem,

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

that if p_n denotes the n^{th} prime (so that $p_2 = 3,\ p_5 = 11$, etc) then

$$p_n \sim n \log n.$$

To see that, note that $\pi(p_n) = n$ (ie the number of primes $\le p_n$ is n). Thus setting $x = p_n$ in the Prime Number Theorem,

$$n \sim \frac{p_n}{\log p_n},$$

ie

$$\frac{p_n}{n \log p_n} \to 1.$$

Taking logarithms,

$$\log p_n - \log n - \log\log p_n \to 0;$$

hence

$$\frac{\log n}{\log p_n} \to 1,$$

ie

$$\log p_n \sim \log n.$$

We conclude that

$$p_n \sim n \log p_n \sim n \log n.$$

Returning to Euler's Theorem, we see that $\sum 1/p$ behaves like $\sum 1/n \log n$. The latter diverges, but only just, as we see by comparison with

$$\int \frac{dx}{x \log x} = \log\log x.$$

On the other hand,

$$\sum_p \frac{1}{p \log^\epsilon p}$$

converges for any $\epsilon > 0$, since

$$\sum_n \frac{1}{n \log^{1+\epsilon} n}$$

converges by comparison with

$$\int \frac{dx}{x \log^{1+\epsilon} x} = -\frac{1}{\epsilon} \log^{-\epsilon} x.$$

What is perhaps surprising is that it is so difficult to pass from Euler's Theorem to the Prime Number Theorem.

Chapter 2

Dirichlet series

2.1 Definition

Definition 2.1. *A* Dirichlet series *is a series of the form*

$$a_1 1^{-s} + a_2 2^{-s} + a_3 3^{-s} + \cdots ,$$

where $a_i \in \mathbb{C}$.

Remarks. 1. For $n \in \mathbb{N}$ we set

$$n^{-s} = e^{-s \log n},$$

taking the usual real-valued logarithm. Thus n^{-s} is uniquely defined for all $s \in \mathbb{C}$. Moreover,

$$m^{-s} n^{-s} = (mn)^{-s}, \quad n^{-s} n^{-s'} = n^{-(s+s')};$$

while

$$1^{-s} = 1$$

for all s.

2. The use of $-s$ rather than s is simply a matter of tradition. The series may of course equally well be written

$$a_1 + \frac{a_2}{2^s} + \frac{a_3}{3^s} + \cdots .$$

3. The term 'Dirichlet series' is often applied to the more general series

$$a_0 \lambda_0^{-s} + a_1 \lambda_1^{-s} + a_2 \lambda_2^{-s} + \cdots ,$$

where

$$0 < \lambda_0 < \lambda_1 < \lambda_2 < \cdots,$$

and

$$\lambda^{-s} = e^{-s \log \lambda}$$

Such series often appear in mathematical physics, where the λ_i might be, for example, the eigenvalues of an elliptic operator. However, we shall only make use of Dirichlet series of the more restricted type described in the definition above; and we shall always use the term in that sense, referring to the more general series (if at all) as *generalised Dirichlet series.*

4. It is perhaps worth noting that generalised Dirichlet series include power series

$$f(x) = \sum c_n x^n,$$

in the sense that if we make the substitution $x = e^{-s}$ then

$$f(e^{-s}) = \sum c_n e^{-ns} = \sum c_n (e^n)^{-s}.$$

2.2 Convergence

Proposition 2.1. *Suppose*

$$f(s) = a_1 1^{-s} + a_2 2^{-s} + \cdots$$

converges for $s = s_0$. Then it converges for all s with

$$\Re(s) > \Re(s_0).$$

Proof ▶ We use a technique that might be called 'summation by parts', by analogy with integration by parts.

Lemma 1. *Suppose $a_n, b_n (n \in \mathbb{N})$ are two sequences. Let*

$$A_n = \sum_{m \le n} a_m, \quad B_n = \sum_{m \le n} b_m.$$

Then

$$\sum_M^N a_n B_n = A_N B_{N+1} - A_{M-1} B_M - \sum_M^N A_n b_{n+1}.$$

Proof ▶ Substituting $a_n = A_n - A_{n-1}$,

$$\begin{aligned}\sum_M^N a_n B_n &= \sum_M^N (A_n - A_{n-1}) B_n \\ &= \sum_M^N A_n (B_n - B_{n+1}) + A_N B_{N+1} - A_{M-1} B_M \\ &= -\sum_M^N A_n b_{n+1} + A_N B_{N+1} - A_{M-1} B_M.\end{aligned}$$

◀

Lemma 2. *Suppose* $\sum a_n$ *converges and* $\sum b_n$ *converges absolutely. Then*

$$\sum a_n B_n$$

converges.

Proof ▶ By the previous Lemma,

$$\begin{aligned}\sum_M^N a_n B_n &= A_N B_{N+1} - A_{M-1} B_M - \sum_M^N A_n b_{n+1} \\ &= A_N (B_{N+1} - B_M) + (A_N - A_{M-1}) B_M - \sum_M^N A_n b_{n+1}.\end{aligned}$$

Let $\sum a_n = A$, $\sum b_n = B$. The partial sums of both series must be bounded; say

$$|A_n| \le C, \ |B_n| \le D.$$

Then

$$|\sum_M^N a_n B_n| \le C |B_{N+1} - B_M| + D |A_N - A_{M-1}| + C \sum_M^N |b_{n+1}|.$$

As $M, N \to \infty$,

$$B_{N+1} - B_M \to 0, \ A_N - A_{M-1} \to 0, \ \sum_M^N |b_{n+1}| \to 0.$$

Hence

$$\sum_M^N a_n B_n \to 0$$

as $M, N \to \infty$; and therefore $\sum a_n B_n$ converges, by Cauchy's criterion. ◀

Let $s' = s - s_0$. Then $\Re(s') > 0$. We apply the last Lemma with $a_n n^{-s_0}$ for a_n, and $n^{-s'}$ for B_n. Thus

$$\begin{aligned} b_n &= B_n - B_{n-1} \\ &= n^{-s'} - (n-1)^{-s'} \\ &= -s' \int_{n-1}^{n} x^{-s'} \frac{dx}{x}. \end{aligned}$$

Hence

$$\begin{aligned} |b_n| &\le |s'| \int_{n-1}^{n} |x^{-s'}| \frac{dx}{x} \\ &= |s'| \int_{n-1}^{n} x^{-\sigma'} \frac{dx}{x}, \end{aligned}$$

where $\sigma' = \Re(s')$.

Summing,

$$\begin{aligned} \sum_{M}^{N} |b_n| &\le |s'| \int_{M-1}^{N} x^{-\sigma'} \frac{dx}{x} \\ &= \frac{|s'|}{\sigma'} \left((M-1)^{-\sigma'} - N^{-\sigma'} \right). \end{aligned}$$

It follows that

$$\sum_{M}^{N} |b_n| \to 0 \text{ as } M, N \to \infty.$$

Thus $\sum |b_n|$ is convergent, and so the conditions of the last Lemma are fulfilled. We conclude that

$$\sum a_n n^{-s_0} n^{-s'} = \sum a_n n^{-(s_0+s')} = \sum a_n n^{-s}$$

is convergent. ◀

Corollary 2.1. *A Dirichlet series either*

1. *converges for all s,*
2. *diverges for all s, or*
3. *converges for all s to the right of a line*

$$\Re(s) = \sigma_0,$$

and diverges for all s to the left of this line.

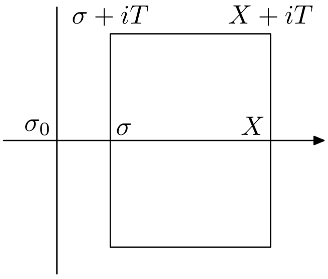

Figure 2.1: Uniform convergence

Definition 2.2. *We call σ_0 the* abscissa of convergence, *setting $\sigma_0 = -\infty$ if the series always converges, and $\sigma_0 = \infty$ if the series never converges.*

Proposition 2.2. *The function*

$$f(s) = a_1 1^{-s} + a_2 2^{-s} + \cdots$$

is holomorphic in the half-plane $\Re(s) > \sigma_0$.

Proof ▶ Suppose $\sigma > \sigma_0$. The argument in the proof of the last Proposition actually shows that $\sum a_n n^{-s}$ converges *uniformly* in any rectangle

$$\{S = x + it : \sigma \leq x \leq X; -T \leq t \leq T\}$$

strictly to the right of $\Re(s) = \sigma_0$ (Fig 2.1), since

$$\begin{aligned} \sum_M^N |b_n| &\leq \frac{|s'|}{\sigma'} M^{-\sigma'} \\ &\leq \frac{|s - s_0|}{\sigma - \sigma_0} M^{-(\sigma - \sigma_0)} \end{aligned}$$

in this region.

Thus $f(s)$ is holomorphic in this rectangle. We conclude that $f(s)$ is holomorphic in the half-plane $\Re(s) > \sigma_0$. ◀

2.3 Absolute convergence

Absolute convergence is simpler than convergence, since

$$\sum |a_n n^{-s}| = \sum |a_n| n^{-\sigma},$$

where $\sigma = \Re(s)$. Thus a Dirichlet series converges absolutely at all, or none, of the points on the line $\Re(s) = \sigma$.

Proposition 2.3. *If*

$$f(s) = a_1 1^{-s} + a_2 2^{-s} + \cdots$$

converges absolutely for $s = s_0$ *then it converges absolutely for all* s *with*

$$\Re(s) \geq \Re(s_0).$$

Proof ▶ This follows at once from the fact that each term

$$|a_n n^{-s}| = |a_n| n^{-\sigma}$$

is a decreasing function of σ. ◀

Corollary 2.2. *A Dirichlet series either*

1. *converges absolutely for all* s*,*
2. *does not converge absolutely for any* s*, or*
3. *converges absolutely for all* s *to the right of a line*

   $$\Re(s) = \sigma_1,$$

   and does not converge absolutely for any s *to the left of this line.*

Definition 2.3. *We call* σ_1 *the* abscissa of absolute convergence*, setting* $\sigma_1 = -\infty$ *if the series always converges absolutely, and* $\sigma_1 = \infty$ *if the series never converges absolutely.*

Proposition 2.4. *We have*

$$\sigma_0 \leq \sigma_1 \leq \sigma_0 + 1.$$

Proof ▶ Suppose

$$\Re(s) > \sigma_0.$$

Then

$$f(s) = a_1 1^{-s} + a_2 2^{-s} + \cdots$$

is convergent. Hence

$$a_n n^{-s} \to 0 \text{ as } n \to \infty.$$

In particular, $a_n n^{-s}$ is bounded, say

$$|a_n n^{-s}| \leq C.$$

But then

$$|a_n n^{-(s+1+\epsilon)}| \le C n^{-(1+\epsilon)}$$

for any $\epsilon < 0$. Since $\sum n^{-(1+\epsilon)}$ converges, it follows that

$$f(s+1+\epsilon)$$

converges absolutely. We have shown therefore that

$$\sigma > \sigma_0 \implies \sigma + 1 + \epsilon \ge \sigma_1$$

for any $\epsilon > 0$, from which it follows that

$$\sigma_0 + 1 \ge \sigma_1.$$

◀

Proposition 2.5. *If $a_n \ge 0$ then $\sigma_1 = \sigma_0$.*

Proof ▶ This is immediate, since in this case

$$\sum |a_n n^{-\sigma}| = \sum a_n n^{-\sigma}.$$

◀

2.4 The Riemann zeta function

Although we have already met the function $\zeta(s)$, it may be best to give a formal definition.

Definition 2.4. *The Riemann zeta function $\zeta(s)$ is defined by the Dirichlet series*

$$\zeta(s) = 1^{-s} + 2^{-s} + \cdots .$$

Remarks. 1. We shall often refer to the Riemann zeta function $\zeta(s)$ simply as *the* zeta function. This is slightly inaccurate, since the term 'zeta function' is applied to a wide range of related functions. However, the Riemann zeta function is the only such function we shall use; so it will cause no confusion if we use the unadorned term 'zeta function' to describe it.

2. For example, there is a zeta function $\zeta_k(s)$ corresponding to each number field k, defined by

$$\zeta_k(s) = \sum_{\mathfrak{a}} N(\mathfrak{a})^{-s},$$

where $\mathfrak{a}$ runs over the *ideals* in k (or rather, in the ring of integers $I(k) = k \cap \bar{\mathbb{Z}}$), and $N(\mathfrak{a})$ is the number of residue classes mod $\mathfrak{a}$.

Since the unique factorisation theorem holds for ideals, the analogue of Euler's product formula holds:

$$\zeta_k(s) = \prod_{\mathfrak{p}} \left(1 - N(\mathfrak{p})^{-s}\right)^{-1},$$

where the product runs over all prime ideals in $I(k)$.

This allows the Prime Number Theorem to be extended to give an approximate formula for the number of prime ideals $\mathfrak{p}$ in the number field k with $N(\mathfrak{p}) \le n$.

3. In another direction, the zeta function $\zeta_E(s)$ of an elliptic differential (or pseudo-differential) operator E is defined by

$$\zeta_E(s) = \sum \lambda_n^{-s},$$

where λ_n $(n = 0, 1, 2, \dots)$ are the eigenvalues of E (necessarily positive, if E is elliptic).

The Riemann zeta function $\zeta(s)$ can be interpreted in this sense as the zeta function of the Laplacian operator on the circle S^1.

Proposition 2.6. *The abscissa of convergence of the Riemann zeta function is*

$$\sigma_0 = 1.$$

Proof ▶ This follows at once from the fact that

$$\sum n^{-\sigma} < \infty \iff \sigma > 1.$$

Let us recall how this is established, by comparing the sum with the integral $\int x^{-\sigma} dx$. If $n - 1 \le x \le n$,

$$n^{-\sigma} \le x^{-\sigma} \le (n-1)^{-\sigma}.$$

Integrating,

$$n^{-\sigma} < \int_{n-1}^{n} x^{-\sigma} dx < (n-1)^{-\sigma},$$

Summing from $n = M + 1$ to N,

$$\sum_{M+1}^{N} n^{-\sigma} < \int_{M}^{N} x^{-\sigma} dx < \sum_{M}^{N+1} n^{-\sigma},$$

It follows that $\sum n^{-\sigma}$ and $\int^{\infty} x^{-\sigma} dx$ converge or diverge together.

But we can compute the integral directly: if $n = 1$ then

$$\int_X^Y x^{-1} dx = \log X - \log Y,$$

and so the integral diverges; while if $n \neq 1$ then

$$\int_X^Y x^{-\sigma} dx = \frac{1}{\sigma - 1}(M^{1-\sigma} - N^{1-\sigma}),$$

and so the integral converges if $\sigma > 1$ and diverges if $\sigma < 1$. ◀

Corollary 2.3. *The zeta function $\zeta(s)$ is holomorphic in the half-plane $\Re(s) > 1$.*

We can continue $\zeta(s)$ analytically to the half-plane $\Re(s) > 0$ in the following way.

Proposition 2.7. *The Dirichlet series*

$$f(s) = 1^{-s} - 2^{-s} + 3^{-s} - \cdots$$

has abscissa of convergence $\sigma_0 = 0$, and so defines a holomorphic function in the half-plane $\Re(s) > 0$.

Proof ▶ Suppose $\sigma > 0$. Then

$$f(\sigma) = 1^{-\sigma} - 2^{-\sigma} + 3^{-\sigma} - \cdots$$

converges, since the terms alternate in sign and decrease to 0 in absolute value. It follows, by Proposition 2.1, that $f(s)$ converges for $\Re(s) > 0$.

The series certainly does not converge for $\Re(s) < 0$, since the terms do not even $\to 0$. Thus $\sigma_0 = 0$. ◀

The abscissa of absolute convergence σ_1 of $f(s)$ is 1 since the terms have the same absolute value as those of $\zeta(s)$.

Proposition 2.8. *If $\Re(s) > 1$ then*

$$f(s) = (1 - 2^{1-s})\zeta(s).$$

Proof ▶ If $\Re(s) > 1$ then the Dirichlet series for $f(s)$ converves absolutely, so we may re-arrange its terms:

$$\begin{aligned} f(s) &= 1^{-s} - 2^{-s} + 3^{-s} - \cdots \\ &= (1^{-s} + 2^{-s} + 3^{-s} + \cdots) - 2(2^{-s} + 4^{-s} + \cdots) \\ &= \zeta(s) - 2 \cdot 2^{-s}(1^{-s} + 2^{-s} + \cdots) \\ &= \zeta(s) - 2 \cdot 2^{-s}\zeta(s) \\ &= (1 - 2^{1-s})\zeta(s). \end{aligned}$$

◀

Proposition 2.9. *The zeta function $\zeta(s)$ extends to a meromorphic function in $\Re(s) > 0$, with a single simple pole at $s = 1$ with residue 1.*

Proof ▶ We have

$$\zeta(s) = \frac{1}{1 - 2^{1-s}} f(s)$$

for $\Re(s) > 1$. But the right-hand side is meromorphic in $\Re(s) > 0$, and so defines an analytic continuation of $\zeta(s)$ (necessarily unique, by the theory of analytic continuation) to this half-plane.

Since $f(s)$ is holomorphic in this region, any pole of $\zeta(s)$ must be a pole of $1/(1 - 2^{1-s})$, ie a zero of $1 - 2^{1-s}$. But

$$2^{1-s} = e^{(1-s)\log 2}.$$

Hence

$$2^{1-s} = 1 \iff (1-s)\log 2 = 2n\pi i$$

for some $n \in \mathbb{Z}$. Thus $1/(1 - 2^{1-s})$ has poles at

$$s = 1 + \frac{2n\pi}{\log 2} i \qquad (n \in \mathbb{Z}).$$

At first sight this seems to give an infinity of poles of $\zeta(s)$ on the line $\Re(s) = 1$. However, the following argument shows that $f(s)$ must vanish at all these points except $s = 1$, thus 'cancelling out' all the poles of $1/(1-2^{1-s})$ except that at $s = 1$.

Consider

$$g(s) = 1^{-s} + 2^{-s} - 2 \cdot 3^{-s} + 4^{-s} + 5^{-s} - 2 \cdot 6^{-s} + \cdots .$$

Like $f(s)$, this converges for all $\sigma > 0$. For if we group $g(\sigma)$ in sets of three terms

$$g(\sigma) = (1^{-\sigma} + 2^{-\sigma} - 2 \cdot 3^{-\sigma}) + (4^{-\sigma} + 5^{-\sigma} - 2 \cdot 6^{-\sigma}) + \cdots$$

we see that each set is > 0. Thus the series either converges (to a limit > 0), or else diverges to $+\infty$.

On the other hand, we can equally well group $g(\sigma)$ as

$$g(\sigma) = 1^{-\sigma} + 2^{-\sigma} - (2 \cdot 3^{-\sigma} - 4^{-\sigma} - 5^{-\sigma}) - (2 \cdot 6^{-\sigma} - 7^{-\sigma} - 8^{-\sigma}) + \cdots .$$

Now each group is < 0, if we omit the terms $1^{-\sigma} + 2^{-\sigma}$. Thus $g(\sigma)$ either converges (to a limit $< 1^{-\sigma} + 2^{-\sigma}$), or else diverges to $-\infty$.

We conclude the $g(\sigma)$ converges (to a limit between 0 and $1^{-\sigma} + 2^{-\sigma}$).

Hence $g(s)$ converges for $\Re(s) > 0$.

But if $\Re(s) > 1$ we can re-write $g(s)$ as

$$\begin{aligned} g(s) &= (1^{-s} + 2^{-s} + 3^{-s} + \cdots) - 3(3^{-s} + 6^{-s} + 9^{-s} + \cdots) \\ &= (1 - 3^{1-s})\zeta(s). \end{aligned}$$

Thus

$$\zeta(s) = \frac{1}{1 - 3^{1-s}} g(s).$$

The right hand side is meromorphic in the half-plane $\Re(s) > 0$, giving a second analytic continuation of $\zeta(s)$ to this region, which by the theory of analytic contination must coincide with the first.

But the poles of $1/(1 - 3^{1-s})$ occur where

$$(1 - s)\log 3 = 2m\pi i,$$

ie

$$s = 1 + \frac{2\pi m}{\log 3} i.$$

Thus $\zeta(s)$ can only have a pole where s is expressible in both forms

$$s = 1 + \frac{2\pi n}{\log 2} i = 1 + \frac{2\pi m}{\log 3} i \qquad (m, n \in \mathbb{Z}).$$

But this implies that

$$m \log 2 = n \log 3,$$

ie

$$2^m = 3^n,$$

which is of course impossible (by the Fundamental Theorem of Arithmetic) unless $m = n = 0$.

We have therefore eliminated all the poles except $s = 1$. At $s = 1$,

$$f(1) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \log 2.$$

(This follows on letting $x \to 1$ from below in $\log(1 + x) = x - x^2/2 + \cdots$.) On the other hand, if we set $s = 1 + s'$ then

$$\begin{aligned} 1 - 2^{1-s} &= 1 - e^{-s' \log 2} \\ &= s' \log 2 - s'^2/2! \log^2 2 + \cdots \\ &= s' \log 2(1 - s'/2 \log 2 + \cdots \end{aligned}$$

Thus

$$\begin{aligned}\frac{1}{1-2^{1-s}} &= \frac{1}{1-2^{-s'}}\\ &= \frac{1}{\log 2\, s'}(1+\frac{s'}{2}\log 2+\cdots)\\ &= \frac{1}{\log 2\, s'}+h(s),\end{aligned}$$

where $h(s)$ is holomorphic. Hence

$$\zeta(1+s') = \frac{1}{s'} + h(s)f(s).$$

We conclude that $\zeta(s)$ has a simple pole at $s = 1$ with residue 1. ◀

In Chapter 7 we shall see that $\zeta(s)$ can in fact be analytically continued to the whole of $\mathbb{C}$. It has no further poles; its only pole is at $s = 1$.

2.5 The Riemann-Stieltjes integral

It is helpful (although by no means essential) to introduce a technique which allows us to express sums as integrals, and brings 'summation by parts' into the more familiar guise of integration by parts.

Let us recall the definition of the Riemann integral $\int_a^b f(x)\,dx$ of a continuous function $f(x)$ on $[a, b]$. Note that $f(x)$ is in fact *uniformly* continuous on $[a, b]$, ie given $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|x-y| < \delta \implies |f(x)-f(y)| < \epsilon$$

for $x, y \in [a, b]$.

By a *dissection* Δ of $[a, b]$ we mean a sequence

$$\Delta : a = x_0 < x_1 < \cdots < x_n = b.$$

We set

$$\|\Delta\| = \max_{0\le i<n} |x_{i+1} - x_i|.$$

The dissection Δ' is said to be a *refinement* of Δ, and we write $\Delta' \subset \Delta$ if the set of dissection-points x_i of Δ is a subset of the set of dissection-points of Δ'. Evidently

$$\Delta' \subset \Delta \implies \|\Delta'\| \le \|\Delta\|.$$

Let

$$S(f,\Delta) = \sum_{0 \le i < n} f(x_i)(x_{i+1} - x_i).$$

Then $S(f,\Delta)$ is convergent as $\|\Delta\| \to 0$, ie given $\epsilon > 0$ there exists $\delta > 0$ such that

$$\|\Delta_1\|, \|\Delta_2\| < \delta \implies |S(f,\Delta_1) - S(f,\Delta_2)| < \epsilon.$$

This follows from 2 lemmas (each more or less immediate).

1. Suppose

   $$|x - y| < \delta \implies |f(x) - f(y)| < \epsilon.$$

   Then

   $$\|\Delta_1\| < \delta,\ \Delta_2 \subset \Delta \implies |S(f,\Delta_1) - S(f,\Delta_2)| < (b-a)\epsilon.$$

2. Given 2 dissections Δ_1, Δ_2 of $[a,b]$ we can always find a common refinement Δ_3, ie

   $$\Delta_3 \subset \Delta_1,\ \Delta_3 \subset \Delta_2.$$

These in turn imply

3. $$\|\Delta_1\|, \|\Delta_2\| < \delta \implies |S(f,\Delta_1) - S(f,\Delta_2)| < 2(b-a)\epsilon.$$

Thus, by Cauchy's criterion, $S(f,\Delta)$ converges as $\Delta \to 0$, ie there exists an $I \in \mathbb{C}$ such that ie given $\epsilon > 0$ there exists $\delta > 0$ such that

$$|S(f,\Delta) - I| < \epsilon \text{ if } \|\Delta\| < \delta.$$

Even if $f(x)$ is not continuous, we say that it is Riemann-integrable on $[a,b]$ with

$$\int_a^b f(x)\,dx = I$$

if

$$S(f,\Delta) \to I \text{ as } \|\Delta\| \to 0.$$

Now suppose $M(x)$ is an increasing (but not necessarily strictly increasing) function on $[a,b]$, ie

$$x \le y \implies f(x) \le f(y).$$

Then we set

$$S_M(f,\Delta) = \sum_{0 \le i < n} f(x_i)(M(x_{i+1}) - M(x_i)).$$

Proposition 2.10. *If $f(x)$ is continuous and $M(x)$ is increasing then*

$$S_M(f, \Delta) \text{ converges as } \|\Delta\| \to 0.$$

Proof ▶ The result follows in exactly the same way as for the Riemann integral above, with (1) replaced by

1'. Given $\epsilon > 0$, suppose $\delta > 0$ is such that

$$|x - y| < \delta \implies |f(x) - f(y)| < \epsilon.$$

Then if $\|\Delta_1\|, \|\Delta_2\| < \delta$,

$$|S(f, \Delta_1) - S(f, \Delta_2)| < (M(b) - M(a))\epsilon.$$

◀

Definition 2.5. *We call*

$$I = \lim_{\|\Delta\| \to 0} S_M(f, \Delta)$$

the Riemann-Stieltjes integral of $f(x)$ with respect to $M(x)$, and write

$$\int_a^b f(x)\, dM = I.$$

2.5.1 Functions of bounded variation

Definition 2.6. *A (real- or complex-valued) function $f(x)$ is said to be* of bounded variation *on the interval $[a, b]$ if there exists a constant C such that*

$$A(f, \Delta) = \sum |f(x_i) - f(x_{i-1})| \leq C$$

for all dissections Δ of $[a, b]$.

Proposition 2.11. *Any linear combination*

$$f(x) = \mu_1 f_1(x) + \cdots + \mu_r f_r(x) \qquad (\mu_1, \ldots, \mu_r \in \mathbb{C})$$

of functions $f_1(x), \ldots, f_r(x)$ of bounded variation is itself of bounded variation.

Proof ▶ This follows at once from the fact that

$$A(f, \Delta) \leq |\mu_1| A(f_1, \Delta) + \cdots + |\mu_r| A(f_r, \Delta).$$

◀

Proposition 2.12. *Any monotone increasing or decreasing function $f(x)$ is of bounded variation.*

Proof ▶ If $f(x)$ is increasing then

$$|f(x_i) - f(x_{i-1}| = f(x_i) - f(x_{i-1};$$

and so

$$A(f, \Delta) = f(b) - f(a).$$

If $f(x)$ is decreasing then $-f(x)$ is increasing, so the result follows from the last Proposition. ◀

Proposition 2.13. *A function $f(x)$ of class $C^1[a, b]$, ie with continuous derivative $f'(x)$ on $[a, b]$, is of bounded variation.*

Proof ▶ Since $f'(x)$ is continuous, it is bounded: say

$$|f'(x)| \le C.$$

Also, by the Mean Value Theorem,

$$f(x_i) - f(x_{i-1} = (x_i - x_{i-1})f'(\xi),$$

where $x_{i-1} < \xi < x_i$. Hence

$$|f(x_i) - f(x_{i-1}| \le C(x_i - x_{i-1});$$

and so

$$A(f, \Delta) \le C(b - a).$$

◀

Proposition 2.14. *A real-valued function $f(x)$ is of bounded variation on $[a, b]$ if and only if it can be expressed as the difference of two increasing functions:*

$$f(x) = M(x) - N(x),$$

where $M(x), N(x)$ are monotone increasing.

Proof ▶ If $f(x)$ is expressible in this form then it is of bounded variation, by Propositions 2.12 and 2.11.

For the converse, let

$$P(f, \Delta) = \sum_{i:f(x_i) \ge f(x_{i-1})} (f(x_i) - f(x_{i-1})),$$
$$N(f, \Delta) = -\sum_{i:f(x_i) < f(x_{i-1})} (f(x_i) - f(x_{i-1})).$$

for each dissection Δ of $[a, b]$. Then $P(f, \Delta), N(f, \Delta) \geq 0$; and

$$P(f, \Delta) - N(f, \Delta) = f(b) - f(a), \quad P(f, \Delta) + N(f, \Delta) = A(f, \Delta).$$

It follows that

$$0 \leq P(f, \Delta), N(f, \Delta) \leq A(f, \Delta).$$

Hence

$$P(f) = \sup_{\Delta} P(f, \Delta), \quad N(f) = \sup_{\Delta} N(f, \Delta)$$

are defined.

Lemma 3. *We have*

$$P(f) - N(f) = f(b) - f(a).$$

Proof ▶ Given $\epsilon > 0$ we can find dissections Δ_1, Δ_2 such that

$$\begin{aligned} P(f) &\geq P(f, \Delta_1) > P(f) - \epsilon, \\ N(f) &\geq N(f, \Delta_2) > N(f) - \epsilon. \end{aligned}$$

If now Δ is a common refinement of Δ_1, Δ_2 then

$$\begin{aligned} P(f) &\geq P(f, \Delta) \geq P(f, \Delta_1) > P(f) - \epsilon, \\ N(f) &\geq N(f, \Delta) \geq N(f, \Delta_2) > N(f) - \epsilon. \end{aligned}$$

But

$$P(f, \Delta) - N(f, \Delta) = f(b) - f(a).$$

It follows that

$$P(f) - N(f) - \epsilon \leq f(b) - f(a) \leq P(f) - N(f) + \epsilon.$$

Since this is true for all $\epsilon > 0$,

$$P(f) - N(f) = f(b) - f(a).$$

◀

Now suppose $a \leq x \leq b$. We apply the argument above to the interval $[a, x]$. Let $p(x), n(x)$ be the functions $P(f), N(f)$ for the interval $[a, x]$. By the last Lemma,

$$p(x) - n(x) = f(x) - f(a).$$

It is easy to see that $p(x), n(x)$ are increasing functions of x. For suppose

$$a \leq x < y \leq b.$$

To each dissection

$$\Delta : a = x_0 < x_1 < \cdots x_n = x$$

of $[a, x]$ we can associate the dissection

$$\Delta' : a = x_0 < x_1 < \cdots x_n < x_{n+1} = y$$

of $[a, y]$; and then

$$P(f, \Delta') \geq P(f, \Delta), \quad N(f, \Delta') \geq N(f, \Delta).$$

It follows that

$$p(y) \geq p(x),\ n(y) \geq n(x),$$

ie $p(x), n(x)$ are monotone increasing. Since

$$f(x) = (f(a) + p(x)) - n(x),$$

this establishes the result. ◀

Proposition 2.15. *The function $f(x)$ is of bounded variation on $[a, b]$ if and only if it can be expressed as a linear combination of increasing functions:*

$$f(x) = \mu_1 M_1(x) + \cdots + \mu_r M_r(x),$$

where $M_1(x), \ldots, M_r(x)$ are monotone increasing, and $\mu_1, \ldots, \mu_r \in \mathbb{C}$.

Proof ▶ It follows from Propositions 2.11 and 2.12 that a function of this form is of bounded variation.

For the converse, note that if $f(x)$ is complex-valued then it can be split into its real and imaginary parts:

$$f(x) = f_R(x) + if_I(x)$$

where $f_R(x), f_I(x)$ are real-valued functions. It is easy to see that if $f(x)$ is of bounded variation then so are $f_R(x)$ and $f_I(x)$. Hence each is expressible as a difference of increasing functions, say

$$f_R(x) = M_R(x) - N_R(x),\ f_I(x) = M_I(x) - N_I(x).$$

But then

$$f(x) = M_R(x) - N_R(x) + iM_I(x) - iN_I(x),$$

which is of the required form. ◀

This result allows us to extend the Riemann-Stieltjes integral to functions of bounded variation.

Suppose $U(x)$ is a function of bounded variation on $[a, b]$. We set

$$S_U(f, \Delta) = \sum f(x_i)(U(x_{i+1}) - U(x_i))$$

for any dissection Δ of $[a, b]$.

Proposition 2.16. *If $f(x)$ is continuous and $U(x)$ is of bounded variation then*

$$S_U(f, \Delta) \textit{ converges as } \|\Delta\| \to 0.$$

Proof ▶ By the last Proposition, we can express $U(x)$ as a linear combination of increasing functions. The result then follows from Proposition 2.10. ◀

Definition 2.7. *We call*

$$I = \lim_{\|\Delta\| \to 0} S_U(f, \Delta)$$

the Riemann-Stieltjes integral of $f(x)$ with respect to $U(x)$, and write

$$\int_a^b f(x)\, dU = I.$$

We extend the Riemann-Stieltjes integral to non-continuous functions $f(x)$ as we do the familiar Riemann integral. Thus if

$$S_U(f, \Delta) \to I \text{ as } \|\Delta\| \to 0$$

then we say that $f(x)$ is *Riemann-Stieltjes integrable* on $[a, b]$, with

$$\int_a^b f(x)\, dU = I.$$

Similarly, we extend the Riemann-Stieltjes integral to infinite ranges in the same was as the Riemann integral. Thus we set

$$\int_a^\infty f(x) dU = \lim_{X \to \infty} \int_a^X f(x) dU,$$

if the limit exists.

In one important case the Riemann-Stieltjes integral reduces to the familiar Riemann integral.

Proposition 2.17. *Suppose $U(x)$ is of class $C^1[a,b]$, ie $U(x)$ has continuous derivative $U'(x)$ on $[a,b]$; and suppose $f(x)U'(x)$ is Riemann integrable on $[a,b]$. Then $f(x)$ is Riemann-Stieltjes integrable, and*

$$\int_a^b f(x)dU = \int_a^b f(x)U'(x)dx.$$

Proof ▶ Suppose Δ is a dissection of $[a,b]$. We compare $S_U(f,\Delta)$ with $S(fU',\Delta)$.

By the Mean Value Theorem,

$$U(x_{i+1}) - U(x_i) = U'(\xi_i)$$

where $x_i < \xi_i < x_{i+1}$. Moreover, since $U'(x)$ is continuous on $[a,b]$, it is absolutely continuous; so given any $\epsilon > 0$ we can find $\delta > 0$ such that

$$|U'(x_i) - U'(\xi_i)| < \epsilon$$

if $x_{i+1} - x_i < \delta$.

Hence

$$|S_U(f,\Delta) - S(fU',\Delta)| \leq \max|f|(b-a)\epsilon$$

if $\|\Delta\| < \delta$, from which the result follows. ◀

2.5.2 Discontinuities

Proposition 2.18. *If $f(x)$ is a function of bounded variation on $[a,b]$ then the left limit*

$$f(x-0) = \lim_{t\to x-0} f(t)$$

exists for all $x \in [a,b)$; and the right limit

$$f(x+0) = \lim_{t\to x+0} f(t)$$

exists for all $x \in (a,b]$.

Proof ▶ The result is (almost) immediate if $f(x)$ is increasing. It follows for any function $f(x)$ of bounded variation by Proposition 2.15, since

$$f(x) = \mu_1 M_1(x) + \cdots + \mu_r M_r(x) \text{ for all } x \implies f(x-0) = \mu_1 M_1(x-0) + \cdots + \mu_r M_r(x-0),$$

and similarly for the right limit. ◀

The function $f(x)$ is continuous at $x = \xi$ if

$$f(\xi - 0) = f(\xi) = f(\xi + 0).$$

Otherwise $f(x)$ has a *discontinuity* at ξ.

Proposition 2.19. *The discontinuities of a function $f(x)$ of bounded variation are enumerable.*

Proof ▶ It is sufficient to prove the result for an increasing function; for if

$$f(x) = \mu_1 M_1(x) + \cdots + \mu_r M_r(x)$$

then the discontinuities of $f(x)$ lie in the union of the discontinuities of $M_1(x), \ldots, M_r(x)$; and a finite union of enumerable sets is enumerable.

Let us define the 'jump' at a discontinuity ξ to be

$$j(\xi) = f(\xi + 0) - f(\xi - 0).$$

Note that for an increasing function

$$f(\xi - 0) \le f(\xi) \le f(\xi + 0).$$

Thus $f(x)$ is discontinuous at ξ if and only if $j(\xi) > 0$.

Lemma 4. *Suppose $M(x)$ is increasing on $[a, b]$; and suppose*

$$a \le \xi_1 < \xi_2 < \cdots < \xi_n \le b.$$

Then

$$\sum_{1 \le i \le n} j(\xi_i) \le f(b) - f(a).$$

Proof ▶ Choose a dissection $x_0, x_1, \ldots, x_n$ of $[a, b]$ with

$$a = x_0 \le \xi_1 < x_1 < \xi_2 < x_2 < \cdots < x_{n-1} < \xi_n \le x_n = b.$$

Then it is easy to see that

$$f(x_i) - f(x_{i-1}) \ge j(\xi_i);$$

and so, on addition,

$$f(b) - f(a) \ge \sum j(\xi_i).$$

◀

Corollary 2.4. *Suppose $M(x)$ is increasing on $[a, b]$; Then the number of discontinuities with*

$$j(x) = f(x + 0) - f(x - 0) \geq 2^{-r}$$

is

$$\leq 2^r(f(b) - f(a)).$$

Using the Lemma, we can enumerate the discontinuities of $M(x)$ by first listing those with $j(x) \geq 1$, then those with $1 > j(x) \geq 2^{-1}$, then those with $2^{-1} > j(x) \geq 2^{-2}$, and so on. In this way we enumerate all the discontinuities:

$$\xi_0, \xi_1, \xi_2, \ldots .$$

◀

Remarks. 1. Note that we are not claiming that the discontinuities can be enumerated in *increasing order*, ie so that

$$\xi_0 < \xi_1 < \xi_2 < \cdots .$$

That is not so, in general; $f(x)$ could , for example, have a discontinuity at every rational point.

2. The discontinuity at ξ can be divided into two parts:

$$f(\xi) - f(\xi - 0) \text{ and } f(\xi + 0) - f(\xi).$$

However, if $f(x)$ is *right-continuous*, ie

$$f(x + 0) = f(x)$$

for all $x \in [a, b)$, then the second contribution vanishes, and the discontinuity is completely determined by

$$j(\xi) = f(\xi + 0) - f(\xi - 0) = f(\xi) - f(\xi - 0).$$

To simplify the discussion, the functions we use have all been chosen to be right-continuous; for example, we set

$$\pi(x) = \|\{p : p \leq x\}\|,$$

although we could equally well have taken the left-continuous function

$$\pi_1(x) = \|\{p : p < x\}\|.$$

(From a theoretical point of view, it might have been preferable to have imposed the symmetric condition

$$f(x) = \frac{1}{2}\left(f(x+0) + f(x-0)\right).$$

However, for our purposes the added complication would outweigh the theoretical advantage.)

Definition 2.8. *The* step function $H_\xi(x)$ *is defined by*

$$H_\xi(x) = \begin{cases} 0 & \textit{if } x < \xi, \\ 1 & \textit{if } x \geq \xi. \end{cases}$$

Proposition 2.20. *Suppose $U(x)$ is a right-continuous function of bounded variation on $[a, b]$. Then*

$$\sum_\xi j(\xi)$$

is absolutely convergent.

Proof ▶ It is sufficient to prove the result when $U(x)$ is increasing, by Proposition 2.15. But in that case $j(\xi) > 0$, and

$$\sum_\xi j(\xi) \leq f(b) - f(a),$$

by Lemma 4. ◀

Proposition 2.21. *Suppose $U(x)$ is a right-continuous function of bounded variation on $[a, b]$. Then $U(x)$ can be split into two parts,*

$$U(x) = J(x) + f(x),$$

where $f(x)$ is continuous, and

$$J(x) = \sum j(\xi) H_\xi(x),$$

the sum extending over all discontinuities ξ of $f(x)$ in $[a, b]$.

Proof ▶ It is sufficient to prove the result in the case where $U(x)$ is increasing, by Proposition 2.15.

Let

$$f(x) = U(x) - J(x).$$

We have to show that $f(x)$ is continuous.

The step function $H_\xi(x)$ is right-continuous. Hence $J(x)$ is right-continuous; and since $U(x)$ is right-continuous by hypothesis, it follows that $f(x)$ is right-continuous. We have to show that $f(x)$ is also left-continuous.

Suppose $x < y$. Then

$$\begin{aligned} J(y) - J(x) &= \sum_{x<\xi\leq y} j(\xi) \\ &\leq U(y) - U(x), \end{aligned}$$

by Proposition 4. Thus

$$f(x) = U(x) - J(x) \leq U(y) - J(y) = f(y),$$

ie $f(x)$ is increasing.

Moreover,

$$0 \leq f(y) - f(x) \leq U(y) - U(x).$$

Hence

$$0 \leq f(y) - f(y-0) \leq U(y) - U(y-0).$$

In particular, if $U(x)$ is left-continuous at y then so is $f(x)$.

Now suppose $U(x)$ has a discontinuity at y. If $x < y$ then

$$J(y) - J(x) \geq j(y) = U(y) - U(y-0).$$

Hence

$$J(y) - J(y-0) \geq U(y) - U(y-0),$$

ie

$$f(y-0) = U(y-0) - J(y-0) \geq f(y) = U(y) - J(y).$$

Since $f(x)$ is increasing, it follows that

$$f(y-0) = f(y),$$

ie $f(x)$ is left-continuous at y. ◀

Definition 2.9. *We call $f(x)$ the* continuous part *of $U(x)$, and $J(x)$ the* purely discontinous part.

Remarks. 1. This is our own terminology; there do not seem to be standard terms for these two parts of a function of bounded variation. That is probably because they are more generally studied through the *measure* or *distribution* dU, with the step function $H_\xi(x)$ replaced by the Dirac delta 'function' $\delta_\xi(x) = dH_\xi$.

2. Our definition of $J(x)$ entails that $J(a) = 0$. With that condition, the splitting of $U(x)$ is unique. If we drop the condition then $J(x)$ and $f(x)$ are defined up to a constant.

Proposition 2.22. *Suppose*

$$U(x) = \sum j(\xi) H_\xi(x)$$

is a purely discontinuous (but right-continuous) function of bounded variation on $[a, b]$*; and suppose* $f(x)$ *is continuous on* $[a, b]$*. Then*

$$\int_a^b f(x) dU = \sum j(\xi) f(\xi).$$

Proof ▶ Since $\sum j(\xi)$ is absolutely convergent, it is sufficient to prove the result for a single step function $H_\xi(x)$.

Suppose Δ is a dissection of $[a, b]$; and suppose

$$x_i < \xi \leq x_{i+1}.$$

Then

$$\begin{aligned} S_{H_\xi}(f, \Delta) &= f(x_i)(H_\xi(x_{i+1}) - H_\xi(x_i)) \\ &= f(x_i). \end{aligned}$$

Since

$$f(x_i) \to f(\xi) \text{ as } \|\Delta\| \to 0,$$

the result follows. ◀

In practice we shall encounter the Riemann-Stieltjes integral $\int_a^b f(x) dU$ in just two cases: the case above, where $f(x)$ is continuous and $U(x)$ is purely discontinuous; and the case where $U(x) \in C^1[a, b]$, when (as we saw in Proposition 2.17)

$$\int f(x) dU = \int f(x) U'(x) dx.$$

2.5.3 Integration by parts

Proposition 2.23. *Suppose* $U(x), V(x)$ *are of bounded variation on* $[a, b]$*; and suppose either* $U(x)$ *or* $V(x)$ *is continuous. Then*

$$\int_a^b U(x)\, dV + \int_a^b V(x)\, dU = U(b)V(b) - U(a)V(a).$$

Proof ▶ We may suppose that $U(x)$ is continuous. Then $\int U(x)dV$ is certainly defined; we must show that $\int V(x)dU$ is also defined.

Let

$$\Delta : a = x_0 < x_1 < \cdots < x_n = b$$

be a dissection of $[a, b]$.

Our proof is based on the formula for 'summation by parts' (Lemma 1), which we may re-write as

$$\sum_{i=1}^{n-1} A_i b_i + \sum_{i=1}^{n-1} a_{i+1} B_i = A_n B_{n-1} - A_1 B_0.$$

Adding $A_n b_n + a_1 B_0$ to each side, this becomes

$$\sum_{i=1}^{n} A_i b_i + \sum_{i=1}^{n} a_i B_{i-1} = A_n B_n - A_0 B_0.$$

Now let us substitute

$$A_i = U(x_i),\ B_i = V(x_i).$$

The first sum becomes

$$\sum_{i=1}^{n-1} A_i b_i = \sum_{i=1}^{n-1} U(x_i) \left(V(x_i) - V(x_{i-1}) \right).$$

This is almost $S_U(V, \Delta)$. There is a discrepancy because we are taking the value $U(x_i)$ at the top of the interval $[x_{i-1}, x_i]$ rather than at the bottom.

However, $U(x)$ is continuous, and so uniformly continuous, on $[a, b]$. Thus given $\epsilon > 0$ we can find $\delta > 0$ such that

$$|U(x_i) - U(x_{i-1}| < \epsilon$$

if $\|\Delta\| < \delta$. It follows that

$$\begin{aligned} |\sum A_i b_i - S_U(V, \Delta)| &\le \epsilon \sum |V(x_i) - V(x_{i-1})| \\ &= \epsilon A(V, \Delta). \end{aligned}$$

Since $V(x)$ is of bounded variation, there is a constant $C > 0$ such that

$$A(V, \Delta) \le C$$

for all dissections Δ of $[a, b]$. Thus

$$|\sum A_i b_i - S_U(V, \Delta)| \le C\epsilon.$$

Turning to the second term,

$$\begin{aligned}\sum_{i=1}^{n} a_i B_{i-1} &= \sum_{i=1}^{n} \left(U(x_i) - U(x_{i-1})\right) V(x_{i-1}) \\ &= S_U(V, \Delta).\end{aligned}$$

Now we know that

$$S_V(U, \Delta) \to \int_a^b U(x) dV \text{ as } \|\Delta\| \to 0.$$

It follows that $\int V\, dU$ is also defined, ie $V(x)$ is Riemann-Stieltjes integrable with respect to $U(x)$ over $[a, b]$, and

$$\begin{aligned}\int_a^b U(x) dV + \int_a^b V(x) dU &= U(b)V(b) - U(a)V(a) \\ &= \left[U(x)V(x)\right]_a^b.\end{aligned}$$

◀

2.5.4 The abscissa of convergence revisited

To see how the Riemann-Stieltjes integral can be used, we look again at the proof of Proposition 2.1. Let

$$f(s) = \sum a_n n^{-s}.$$

We have to show that

$$f(s) \text{ convergent } \implies f(s + s') \text{ convergent}$$

if $\Re(s') > 0$.

Let

$$V(x) = \sum_{n \le x} a_n n^{-s}.$$

Then $V(x)$ has discontinuities at each integer point $x = n$, with $j(n) = n^{-s}$. Thus by Proposition 2.22

$$\begin{aligned}\sum_{M+1}^{N} a_n^{-(s+s')} &= \int_M^N x^{-s'} dV \\ &= \int_M^N U(x) dV,\end{aligned}$$

where

$$U(x) = x^{-s'}.$$

Integrating by parts (by Proposition 2.23),

$$\begin{aligned}\sum_{M+1}^{N} a_n^{-(s+s')} &= [U(x)V(x)]_M^N - \int_M^N V(x)dU \\ &= [U(x)V(x)]_M^N - \int_M^N V(x)U'(x)dx \\ &= [U(x)V(x)]_M^N - s'\int_M^N x^{-s'}V(x)\frac{dx}{x},\end{aligned}$$

since $U(x)$ has continuous derivative $s'x^{-(s'+1)}$.

Since $f(s)$ is convergent, $V(x)$ is bounded, say

$$V(x) \le V.$$

Thus if $\sigma' = \Re(s')$,

$$\begin{aligned}|\sum_{M+1}^{N} a_n^{-(s+s')}| &\le V\left(M^{-\sigma'} + N^{-\sigma'}\right) + |s'|\int_M^N x^{-\sigma'}\frac{dx}{x} \\ &= V\left(M^{-\sigma'} + N^{-\sigma'}\right) + \frac{V|s'|}{\sigma'}\left(M^{-\sigma'} - N^{-\sigma'}\right) \\ &\le V\left(M^{-\sigma'} + N^{-\sigma'} + \frac{|s'|}{\sigma'}M^{-\sigma'}\right) \\ &\to 0 \text{ as } M, N \to \infty.\end{aligned}$$

We conclude that $f(s + s')$ is convergent if $\sigma' = \Re(s') > 0$.

2.5.5 Analytically continuing $\zeta(s)$: an alternative approach

As another application of the Riemann-Stieltjes integral, we give an alternative method of extending $\zeta(s)$.

Let

$$G(x) = [x].$$

(This function is sometimes called the *Gauss function*.)

Suppose $\Re(s) > 1$. Then

$$\begin{aligned}
\zeta(s) &= \sum_1^\infty n^{-s} \\
&= \int_0^\infty x^{-s} dG \\
&= \left[x^{-s} G(x)\right]_0^\infty + s \int_0^\infty x^{-s} G(x) \frac{dx}{x} \\
&= s \int_1^\infty x^{-s} G(x) \frac{dx}{x}.
\end{aligned}$$

(Note that $G(x) = 0$ for $x \in [0,1)$; so there is no convergence problem at $x = 0$.)

We can write

$$x = G(x) + F(x),$$

where $F(x)$ is the 'fractional part' of x. Thus

$$0 \le F(x) \le 1.$$

Now

$$\begin{aligned}
\zeta(s) &= s \int_1^\infty x^{-s} dx - s \int_1^\infty x^{-s} F(x) \frac{dx}{x} \\
&= s \left[\frac{x^{1-s}}{1-s}\right]_1^\infty - s \int_1^\infty x^{-s} F(x) \frac{dx}{x} \\
&= \frac{s}{s-1} - s \int_1^\infty x^{-s} F(x) \frac{dx}{x}.
\end{aligned}$$

But the integral on the right converges if $\Re(s) > 0$, since

$$\begin{aligned}
|\int_X^Y x^{-s} F(x) \frac{dx}{x}| &\le \int_X^Y x^{-\sigma} \frac{dx}{x} \\
&= \frac{1}{\sigma} \left(X^{-\sigma} - Y^{-\sigma}\right) \\
&\to 0 \text{ as } X, Y \to \infty.
\end{aligned}$$

Thus

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

gives an analytic continuation of $\zeta(s)$ to $\Re(s) > 0$.

Moreover, since the integral is holomorphic in this region, we see that $\zeta(s)$ has a single simple pole at $s = 1$ with residue 1 in the half-plane $\Re(s) > 0$.

We can even extend $\zeta(s)$ further, to the half-plane $\Re(s) > -1$, if we take a little care. (In Chapter 7 we shall show by an entirely different method that $\zeta(s)$ can be continued analytically to the whole complex place $\mathbb{C}$; so the present exercise is just that — an exercise.)

Let

$$h(x) = F(x) - \frac{1}{2}.$$

Then

$$\int_n^{n+1} h(x)dx = 0.$$

Hence

$$H(x) = \int_0^x h(t)dt$$

is bounded; in fact

$$|H(x)| \leq \frac{1}{4}.$$

Suppose $\Re(s) > 0$. Integrating by parts (in the usual sense),

$$\begin{aligned}\int_1^\infty F(x)\frac{dx}{x} &= \frac{1}{2}\int_1^\infty x^{-s}\frac{dx}{x} + \int_1^\infty x^{-s}h(x)\frac{dx}{x}\\ &= \frac{1}{2}\left[\frac{x^{-s}}{-s}\right]_1^\infty + \left[x^{-s-1}H(x)\right]_1^\infty + \int_1^\infty x^{-s-2}H(x)dx\\ &= \frac{1}{2s} + \int_1^\infty x^{-s-2}H(x)dx.\end{aligned}$$

Thus

$$\zeta(s) = \frac{s}{s-1} - \frac{1}{2} - s\int_1^\infty x^{-s-2}H(x)dx.$$

But the integral on the right converges if $\Re(s) > -1$, since

$$\begin{aligned}|\int_X^Y x^{-s-2}H(x)dx| &\leq \frac{1}{4}\int_X^Y x^{-\sigma-2}dx\\ &= \frac{1}{4(\sigma+1)}\left(X^{-(\sigma+1)} - Y^{-(\sigma+1)}\right)\\ &\to 0 \text{ as } X, Y \to \infty.\end{aligned}$$

Thus

$$\zeta(s) = \frac{s}{s-1} - \frac{1}{2} - s\int_1^\infty x^{-(s+2)}H(x)dx$$

gives an analytic continuation of $\zeta(s)$ to $\Re(s) > -1$.

2.6 The relation between A_n and σ_0

Power series are simpler than Dirichlet series, in that the radius of convergence of a power series

$$\sum c_n x^n$$

is equal to the radius of absolute convergence, both being given by

$$r = \limsup |c_n|^{1/n}.$$

We must expect the corresponding result for a Dirichlet series

$$\sum a_n n^{-s}$$

to involve the partial sums

$$A_n = \sum_{m \le n} a_n$$

rather than the coefficients a_n themselves.

Proposition 2.24. *Suppose*

$$f(s) = \sum a_n n^{-s}$$

has abscissa of convergence σ_0. Then

$$A_n = o(n^\sigma)$$

for any $\sigma > \sigma_0$.

Conversely, if

$$A_n = O(n^\sigma)$$

then $\sigma \ge \sigma_0$.

Proof ▶ Suppose $\sigma > \sigma_0$. Choose σ' with

$$\sigma > \sigma' > \sigma_0.$$

Then

$$f(\sigma') = \sum a_n n^{-\sigma'}$$

is convergent. Hence $a_n n^{-\sigma'}$ is bounded, say

$$|a_n n^{-\sigma'}| \le C.$$

Then

$$|a_n| \le C n^{\sigma'}$$

ie

$$a_n = O(n^{\sigma'}) = o(n^{\sigma}).$$

Conversely, suppose

$$A_n = O(n^{\sigma}),$$

say

$$|A_n| \le Cn^{\sigma};$$

and suppose

$$\sigma' > \sigma.$$

Let

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

Then

$$|A(x)| = |A([x])| \le C[x]^{\sigma} \le Cx^{\sigma}.$$

Integrating by parts,

$$\begin{aligned}\sum_{M+1}^{N} a_n n^{-\sigma'} &= \int_M^N x^{-\sigma'} dA \\ &= \left[x^{-\sigma'} A(x)\right]_M^N + \sigma' \int_M^N x^{-\sigma'} A(x) \frac{dx}{x}.\end{aligned}$$

Hence

$$\begin{aligned}|\sum_{M+1}^{N} a_n n^{-\sigma'}| &\le C(M^{\sigma-\sigma'} + N^{\sigma-\sigma'}) + C\sigma \int_M^N x^{\sigma-\sigma'} \frac{dx}{x} \\ &\le C\left(2 + \frac{\sigma}{\sigma' - \sigma}\right) N^{\sigma-\sigma'}.\end{aligned}$$

Thus

$$|\sum_{M+1}^{N} a_n n^{-\sigma'}| \to 0 \text{ as } M, N \to \infty.$$

Hence, by Cauchy's criterion,

$$\sum a_n n^{-\sigma'}$$

is convergent; and so

$$\sigma' \le \sigma_0.$$

Since this holds for all $\sigma' > \sigma$,

$$\sigma \le \sigma_0.$$

◀

2.7 Dirichlet series with positive terms

If the Dirichlet series

$$f(s) = \sum a_n n^{-s}$$

has abscissa of convergence σ_0 then $f(s)$ is holomorphic in the half-plane $\Re(s) > \sigma_0$. But the converse is not in general true, ie we may be able to continue $f(s)$ analytically to a function $f(s)$ holomorphic in the half-plane $\Re(s) > \sigma'$, where $\sigma' < \sigma_0$.

For example, the abscissa of convergence of

$$\left(1 - 2^{1-s}\right)\zeta(s) = 1^{-s} - 2^{-s} + 3^{-s} - 4^{-s} + \cdots$$

is $\sigma_0 = 0$. (The terms in the series do not even $\to 0$ for $\Re(s) < 0$.) But as we shall see, this series extends analytically to an entire function, ie a function holomorphic in the whole of $\mathbb{C}$.

However, the following Proposition shows that if the coefficients a_n of the Dirichlet series are *positive* then the converse does hold — $f(s)$ cannot be extended holomorphically across the line $\Re(s) = \sigma_0$.

Proposition 2.25. *Suppose the Dirichlet series*

$$f(s) = \sum a_n n^{-s}$$

has abscissa of convergence σ_0; and suppose $a_n \geq 0$ for all n. If $f(s)$ can be extended to a function meromorphic in an open set containing $s = \sigma_0$ then $f(s)$ must have a pole at $s = \sigma_0$.

Proof ▶ Suppose $f(s)$ is holomorphic in

$$D_0 = D(\sigma_0, \delta) = \{z \in \mathbb{C} : |z - \sigma_0| < \delta\}$$

Let

$$\sigma = \sigma_0 + \frac{\delta}{4}.$$

Then $f(s)$ is holomorphic in

$$D_1 = D(\sigma, \frac{3\delta}{4}) = \{z \in \mathbb{C} : |z - \sigma| < \frac{3\delta}{4}\} \subset D_0.$$

It follows by Taylor's theorem that

$$f(s) = f(\sigma) + f'(\sigma)(s - \sigma) + \frac{1}{2!}f''(\sigma)(s - \sigma)^2$$

for $s \in D_1$.

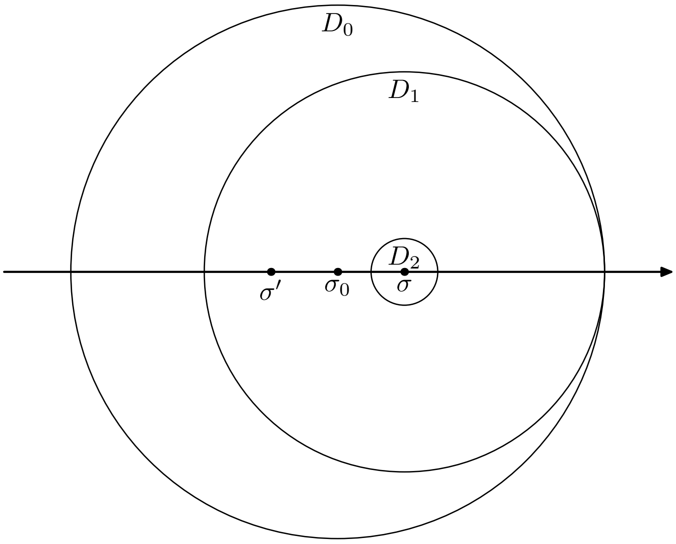

Figure 2.2: Convergence of Dirichlet series with positive terms

Now

$$f(s) = \sum a_n n^{-s}$$

near $s = \sigma$ (since this point is in the half-plane of convergence). Moreover this series converges uniformly and absolutely for s sufficiently close to σ, say inside

$$D_2 = D(\sigma, \frac{\delta}{8}.$$

It follows that we can differentiate term-by-term, as often as we like:

$$\begin{aligned} f'(s) &= -\sum a_n \log n\, n^{-s}, \\ f''(s) &= \sum a_n \log^2 n\, n^{-s}, \\ f'''(s) &= -\sum a_n \log^3 n\, n^{-s}, \end{aligned}$$

etc. In particular

$$f^{(k)}(\sigma) = (-1)^k \sum a_n \log^k nn^{-\sigma},$$

where $f^{(k)}$ denotes the kth derivative of $f(s)$.

Now let us apply Taylor's expansion to compute $f(\sigma')$, where

$$\sigma' = \sigma - \frac{\delta}{4}.$$

We have

$$\begin{aligned} f(\sigma') &= \sum_k \frac{1}{k!} f^{(k)}(\sigma)(\sigma' - \sigma)^k \\ &= \sum_k (-1)^k \frac{1}{k!} f^{(k)}(\sigma) \left(\frac{\delta}{4}\right)^k . \end{aligned}$$

Substituting from above for the $f^{(k)}$,

$$f(\sigma') = \sum_k \frac{1}{k!} \left(\frac{\delta}{4}\right)^k \sum_n a_n \log^k n n^{-\sigma}.$$

Since all the terms on the right are positive (the two factors $(-1)^k$ cancelling out), the double series is absolutely convergent, and we can invert the order of the summations:

$$f(\sigma') = \sum_n a_n n^{-\sigma} \sum_k \frac{1}{k!} \log^k n \left(\frac{\delta}{4}\right)^k$$

The series on the right may seem complicated, but common-sense tells us what the sum must be. We could have carried out the whole operation entirely within the half-plane of convergence, in which case we know that

$$f(\sigma') = \sum a_n n^{-\sigma'}.$$

Clearly this must still be true.

In fact,

$$\begin{aligned} \sum_k \frac{1}{k!} \log^k n \left(\frac{\delta}{4}\right)^k &= \sum_k \frac{1}{k!} \left(\frac{\delta \log n}{4}\right)^k \\ &= e^{\delta \log n/4} \\ &= n^{\delta/4} \\ &= n^{\sigma - \sigma'}, \end{aligned}$$

and so

$$\begin{aligned} f(\sigma') &= \sum_n a_n n^{-\sigma} n^{\sigma - \sigma'} \\ &= \sum_n a_n n^{-\sigma'}. \end{aligned}$$

Thus $f(\sigma')$ converges, which is impossible since $\sigma' < \sigma_0$. We conclude that our original assumption is untenable: $f(s)$ cannot be holomorphic in a neighbourhood of $s = \sigma_0$. ◀

Chapter 3

The Prime Number Theorem

3.1 Statement of the theorem

The Prime Number Theorem asserts that

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

It is more convenient — and preferable — to express this in a slightly different form.

Definition 3.1. *For $x \geq e$ we set*

$$\mathrm{Li}(x) = \int_e^x \frac{dt}{\log t}.$$

Proposition 3.1. *As $x \to \infty$,*

$$\mathrm{Li}(x) \sim \frac{x}{\log x}.$$

Proof ▶ Integrating by parts,

$$\begin{aligned}\mathrm{Li}(x) &= \int_e^x \frac{dt}{\log t}\\ &= \left[t\frac{1}{\log t}\right]_e^x + \int_e^x t\frac{1}{t\log^2 t}dt\\ &= \frac{x}{\log x} - e + \int_e^x \frac{dt}{\log^2 t}.\end{aligned}$$

It is clear from this that

$$\mathrm{Li}(x) \to \infty \text{ as } x \to \infty.$$

Thus the result will follow if we show that

$$\int_e^x \frac{dt}{\log^2 t} = o(\mathrm{Li}(x)).$$

But

$$\begin{aligned}\int_e^x \frac{dt}{\log^2 t} &= \int_e^{x^{1/2}} \frac{dt}{\log^2 t} + \int_{x^{1/2}}^x \frac{dt}{\log^2 t}\\ &\le x^{1/2} + \frac{1}{\log(x^{1/2})}\int_{x^{1/2}}^x \frac{dt}{\log t}\\ &\le x^{1/2} + \frac{2\,\mathrm{Li}(x)}{\log x}.\end{aligned}$$

From above,

$$\mathrm{Li}(x) \ge \frac{x}{\log x} - e.$$

Thus

$$x^{1/2} = o(\mathrm{Li}(x));$$

and so

$$\int_e^x \frac{dt}{\log^2 t} = o(\mathrm{Li}(x)),$$

as required. ◀

Remark. We can extend this result to give an *asymptotic expansion* of Li(x). Integrating by parts,

$$\begin{aligned}\int_e^x \frac{dt}{\log^n t} &= \left[t\frac{1}{\log^n t}\right]_e^x + \int_e^x t\frac{1}{nt\log^{n+1} t}dt\\ &= \frac{x}{\log^n x} - e + \frac{1}{n}\int_e^x \frac{dt}{\log^{n+1} t}.\end{aligned}$$

It follows that

$$\mathrm{Li}(x) = \frac{x}{\log x} + \frac{x}{\log^2 x} + \frac{1}{2!}\frac{x}{\log^3 x} + \cdots + \frac{1}{(n-1)!}\frac{x}{\log^n x} + O(\frac{x}{\log^{n+1} x}).$$

Corollary 3.1. *The Prime Number Theorem can be stated in the form:*

$$\pi(x) \sim \mathrm{Li}(x).$$

Remark. This is actually a more accurate form of the Prime Number Theorem, in the following sense. It has been shown that

$$\pi(x) - \mathrm{Li}(x)$$

changes sign infinitely often, ie however large x gets we find that sometimes $\pi(x) \geq \mathrm{Li}(x)$, and sometimes $\pi(x) < \mathrm{Li}(x)$.

On the other hand, it follows from the Remark above that $\mathrm{Li}(x)$ is substantially larger than $x/\log x$; and it has also been shown that $\pi(x) > x/\log x$ for all sufficiently large x.

3.2 Preview of the proof

The proof of the Prime Number Theorem is long and intricate, and divided into several more or less independent parts. A preview may therefore be helpful.

1. We start from Euler's Product Formula
$$\zeta(s) = \prod_{\text{primes } p} \left(1 - p^{-s}\right)^{-1}.$$

2. *Logarithmic differentiation* converts this to
$$\begin{aligned} -\frac{\zeta'(s)}{\zeta(s)} &= \sum_p \frac{\log p\, p^{-s}}{1 - p^{-s}} \\ &= \sum_n a_n n^{-s}, \end{aligned}$$
where
$$a_n = \begin{cases} \log p & \text{if } n = p^r \\ 0 & \text{otherwise.} \end{cases}$$

3. The function $\zeta'(s)/\zeta(s)$ has poles wherever $\zeta(s)$ has a pole or zero. It follows from Euler's Product Formula that $\zeta(s)$ has no zeros in $\Re(s) > 1$. Accordingly $\zeta'(s)/\zeta(s)$ has a pole at $s = 1$ (with residue 1) and no poles in $\Re(s) > 1$.

4. Although this is not essential, our argument is somewhat simplified if we 'hive off' the part of the Dirichlet series corresponding to higher prime-powers:
$$-\frac{\zeta'(s)}{\zeta(s)} = \sum \log p\, p^{-s} + h(s),$$

where

$$h(s) = \sum \log p \sum_{r>1} p^{-rs}.$$

The function $h(s)$ converges for $\Re(s) > 1/2$, as can be seen by comparison with $\zeta(2s)$. Its partial sums are therefore of order $o(n^{1/2+\epsilon})$, by Proposition 2.24. Consequently the contribution of $h(s)$ can be ignored in our argument.

5. We are left with the function

$$\begin{aligned}\Theta(s) &= \sum_p \log p\, p^{-s} \\ &= \int_0^\infty x^{-s} d\theta,\end{aligned}$$

where

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

6. A (fairly) simple exercise in summation by parts shows that

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

Accordingly, the proof of the Prime Number Theorem is reduced to showing that

$$\theta(x) \sim x,$$

ie

$$\theta(x) = x + o(x).$$

7. The dominant term x in $\theta(x)$ arises from the pole of $\Theta(s)$ at $s = 1$, in the following sense.

Consider the function $\zeta(s)$. This has a pole at $s = 1$ with residue 1, and it has partial sums

$$A(x) = \sum_{n \le x} 1 = [x] = x + O(1).$$

If now we subtract $\zeta(s)$ from $\Theta(s)$ then we 'remove' the pole at $s = 1$; and at the same time we subtract x from $\theta(x)$. More precisely, let

$$\begin{aligned}\Psi(s) &= \Theta(s) - \zeta(s) \\ &= \sum a_n n^{-s},\end{aligned}$$

where

$$a_n = \begin{cases} \log p - 1 & \text{if } n = p, \\ -1 & \text{otherwise.} \end{cases}$$

Then

$$\Psi(s) = \int_0^\infty x^{-s} d\psi,$$

where

$$\psi(x) = \theta(x) - [x] = \theta(x) - x + O(1).$$

The Prime Number Theorem, as we have seen, is equivalent to the statement that

$$\psi(x) = o(x).$$

8. Riemann hypothesised — we shall see why in Chapter 7 — that all the zeros of $\zeta(s)$ in the 'critical strip' $0 \le \Re(s) \le 1$ lie on the line $\Re(s) = 1/2$.

   If that were so then $\Psi(s)$ would be holomorphic in $\Re(s) > 1/2$, and it would follow from Proposition 2.24 that

   $$\psi(x) = o(x^{1/2+\epsilon})$$

   for any $\epsilon > 0$, which is more than enough to prove the Theorem.

   In fact, Riemann showed that with a little more care one can deduce from his hypothesis that

   $$\psi(x) = O(x^{1/2} \log x),$$

   ie

   $$\theta(x) = x + O(x^{1/2} \log x),$$

   from which it follows that

   $$\pi(x) = \mathrm{Li}(x) + O(x^{1/2} \log x).$$

   This — if it could be established — would constitute a remarkably strong version of the Prime Number Theorem.

9. The Riemann Hypothesis would allow us to push back the abscissa of convergence of $\Psi(s)$ all the way to $\sigma = 1/2$.

It would be sufficient for our purposes if we could push it back to any $\sigma < 1$, since this would imply that

$$\psi(x) = o(x^{\sigma+\epsilon})$$

for any $\epsilon > 0$.

Unfortunately, this has never been established. The best that we can do is to show that $\zeta(s)$ has no zeros actually *on* the line $\Re(s) = 1$:

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

for $t \in \mathbb{R} \setminus \{0\}$.

This proof of this result is, in a sense, the heart of the proof of the Prime Number Theorem. The argument we use is rather strange; we show that if $\zeta(s)$ had a zero at $s = 1 + it$ then it would have a pole at $s = 1 + 2it$, which we know is not the case.

10. This takes us a tiny step forward; it shows that $\Psi(s)$ is holomorphic in $\Re(s) \geq 1$.

    Proposition 2.24 only tells us that in this case

    $$\psi(x) = o(x^{1+\epsilon})$$

    for any $\epsilon > 0$, which is useless.

    We need a much stronger result which tells us that if the Dirichlet series $\sum a_n n^{-s}$ is holomorphic everywhere on its critical line $\Re(s) = \sigma_0$ (and satisfies some natural auxiliary conditions) then its partial sums satisfy

    $$A(x) = o(x^{\sigma_0}).$$

    Results of this kind — relating partial sums of Dirichlet series to the behaviour on the critical line — are known as *Tauberian* theorems, after Alfred Tauber, author of the first such result.

    Tauber's original result used real function theory, and was very difficult. Fortunately, complex function theory yields a Tauberian theorem sufficient for our purpose with relative ease.

    This allows us to conclude that

    $$\psi(x) = o(x),$$

    which as we have seen is tantamount to the Prime Number Theorem.

3.3 Logarithmic differentiation

Recall the notion of *logarithmic differentiation*. Suppose

$$f(x) = u_1(x) \cdots u_n(x),$$

where $u_i(x)$ is differentiable and $u_i(x) > 0$ for $1 \le i \le n$. Taking logarithms,

$$\log f(x) = \sum \log u_i(x).$$

Differentiating,

$$\frac{f'(x)}{f(x)} = \sum \frac{u_i'(x)}{u_i(x)}.$$

it is easy to establish this result without using logarithms: on differentiating the product,

$$f'(x) = \sum u_1(x) \cdots u_i'(x) \cdots u_n(x);$$

and the result follows on dividing by $f(x)$. This shows that the result holds without assuming that $u_i(x) > 0$. Indeed, by this argument the result holds for complex-valued functions: if

$$f(z) = \prod_{1 \le i \le n} u(z),$$

where $u_1(z), \dots, u_n(z)$ are holomorphic in U, then

$$\frac{f'(z)}{f(z)} = \sum \frac{u_i'(z)}{u_i(z)},$$

except where $z = 0$.

We want to extend this to infinite products.

Proposition 3.2. *Suppose $a_n(z)$ $(n \in \mathbb{N})$ is a sequence of holomorphic functions on the open set $U \subset \mathbb{C}$; and suppose the series*

$$\sum |a_n(z)|$$

is uniformly convergent on U. Then

$$f(z) = \prod_n (1 + a_n(z))$$

is holomorphic on U; and

$$\frac{f'(z)}{f(z)} = \sum_n \frac{a_n'(z)}{1 + a_n(z)}$$

on U.

Proof ▶ The partial products

$$P_n(z) = \prod_{m \le n} (1 + a_m(z))$$

converge uniformly to $f(z)$ in U:

$$P_n(z) \to f(z).$$

It follows that

$$P_n'(z) \to f'(z).$$

Hence

$$\frac{P_n'(z)}{P_n(z)} \to \frac{f'(z)}{f(z)}.$$

But

$$\frac{P_n'(z)}{P_n(z)} = \sum_{m \le n} \frac{a_m'(z)}{1 + a_m(z)}.$$

We conclude that

$$\sum_{n \in \mathbb{N}} \frac{a_m'(z)}{1 + a_m(z)} = \frac{f'(z)}{f(z)}.$$

◀

3.4 From $\pi(x)$ to $\theta(x)$

Definition 3.2. *We set*

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

Thus

$$\theta(x) = \begin{cases} 0 & \text{for } x < 2 \\ \log 2 & \text{for } 2 \le x < 3 \\ \log 6 & \text{ for } 3 \le x < 5 \\ & \dots \end{cases}$$

Proposition 3.3. $\pi(x) \sim \mathrm{Li}(x) \iff \theta(x) \sim x.$

Proof ▶ Suppose

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

Then

$$\begin{aligned}\theta(X) &= \sum_{p\le x} \log p\\ &= \log 2 + \int_e^X \log x\, d\pi\\ &= \log 2 + [\log x\, \pi(x)]_e^X - \int_e^X \frac{1}{x}\pi(x)dx\\ &= \log 2 + \pi(X)\log X - 1 - \int_e^X \frac{\pi(x)}{x}dx.\end{aligned}$$

Since $\pi(x) \sim x/\log x$,

$$\pi(x) \le C\frac{x}{\log x}$$

for some constant C; and so

$$0 \le \int_e^X \frac{\pi(x)}{x}dx \le C\int_e^X \frac{dx}{\log x} = C\,\mathrm{Li}(x) = o(x),$$

by Proposition 3.1. Thus

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

Conversely, suppose

$$\theta(x) \sim x.$$

Then

$$\begin{aligned}\pi(X) &= 1 + \int_e^X \frac{1}{\log x}d\theta\\ &= 1 + \left[\frac{\theta(x)}{\log x}\right]_e^X + \int_e^X \frac{\theta(x)}{x\log^2 x}dx\\ &= \frac{\theta(X)}{\log X} + (1-\log 2) + \int_e^X \frac{\theta(x)}{x\log x}dx.\end{aligned}$$

Now

$$\theta(x) \sim x \implies \theta(x) \le Cx$$

for some C; and so

$$\begin{aligned}0 &\le \int_e^X \frac{\theta(x)}{x\log^2 x}dx\\ &\le C\int_e^X \frac{dx}{\log^2 x}.\end{aligned}$$

Hence

$$\frac{\theta(x)}{\log x} \geq \pi(x) \geq \frac{\theta(x)}{\log x} + C\int_e^X \frac{dx}{\log^2 x}.$$

But

$$\theta(x) \sim x \implies \frac{\theta(x)}{\log x} \sim \frac{x}{\log x} \sim \mathrm{Li}(x),$$

while

$$\int_e^X \frac{dx}{\log^2 x} = o(\mathrm{Li}(x)),$$

as we saw in the proof of Proposition 3.1.

We conclude that

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

◀

Corollary 3.2. *The Prime Number Theorem is equivalent to:*

$$\theta(x) \sim x.$$

3.5 The zeros of $\zeta(s)$

Proposition 3.4. *The Riemann zeta function $\zeta(s)$ has no zeros in the half-plane $\Re(s) > 1$.*

Proof ▶ This follows at once from Euler's Product Formula:

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

For the right-hand side converges absolutely for $\Re(s) > 1$; and by the definition of convergence its value is $\neq 0$. ◀

We want to show that $\zeta(s)$ has no zeros on the line $\Re(s) = 1$. This is equivalent to showing that $\zeta'(s)/\zeta(s)$ has no poles on this line except at $s = 1$.

Proposition 3.5. *For $\Re(s) > 1$,*

$$\frac{\zeta'(z)}{\zeta(s)} = -\sum a_n n^{-s},$$

where

$$a_n = \begin{cases} \log p & \text{if } n = p^r \\ 0 & \text{otherwise.} \end{cases}$$

Proof ▶ The result follows at once on applying Proposition 3.2 to Euler's Product Formula, ◀

It is convenient to divide the Dirichlet series for $\zeta'(s)/\zeta(s)$ into two parts, the first corresponding to primes p, and the second to prime-powers $p^r (r \geq 2)$.

Definition 3.3. *We set*

$$\Theta(s) = -\sum_p \log p \, p^{-s}.$$

Proposition 3.6. *The function $\Theta(s)$ is holomorphic in $\Re(s) > 1$.*

Proof ▶ We know that

$$\zeta(s) = \sum n^{-s}$$

is uniformly convergent in $\Re(s) \geq \sigma$ for any $\sigma > 1$. It follows that we can differentiate term-by-term:

$$\zeta'(s) = \sum \log n \, n^{-s}$$

in $\Re(s) > 1$. Since the coefficients are all positive, the convergence is absolute. But the series for $\Theta(s)$ consists of some of the terms of $\zeta'(s)$, and so also converges absolutely in $\Re(s) > 1$. ◀

Proposition 3.7. *For $\Re(s) > 1$,*

$$\frac{\zeta'(z)}{\zeta(s)} = -\Theta(s) + h(s),$$

where $h(s)$ is holomorphic in $\Re(s) > 1/2$.

Proof ▶ We have

$$h(s) = -\sum_p \log p \sum_{r \geq 2} p^{-rs}.$$

If $\sigma = \Re(s)$ then $|p^{-s}| = p^{-\sigma}$. Thus

$$\begin{aligned}\sum_p \log p \sum_{r \geq 2} |p^{-rs}| &= \sum_p \log p \sum_{r \geq 2} p^{-r\sigma} \\ &= \sum_p \log p \frac{p^{-2\sigma}}{1 - p^{-\sigma}} \\ &\leq \frac{1}{1 - 2^{-\sigma}} \sum \log p \, p^{-2\sigma} \\ &= \frac{1}{1 - 2^{-\sigma}} \Theta(2\sigma),\end{aligned}$$

which converges for $2\sigma > 1$, ie $\sigma > 1/2$, by Proposition 3.6. ◀

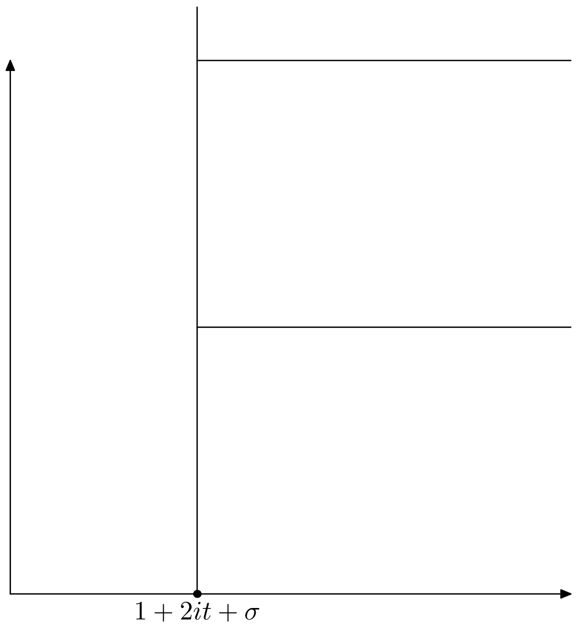

Figure 3.1: Comparing $\Theta(s)$ at three points

Proposition 3.8. *The Riemann zeta function $\zeta(s)$ has no zeros on the line $\Re(s) = 1$:*

$$\zeta(1+it) \neq 0 \qquad (t \in \mathbb{R} \setminus \{0\}).$$

Proof ▶ We shall show (in effect) that if $\zeta(s)$ has a zero at $s = 1+it$ then it must have a pole at $s = 1+2it$; but we know that is impossible, since the only pole of $\zeta(s)$ in $\Re(s) > 0$ is at $s = 1$.

We work with $\Theta(s)$ rather than $\zeta(s)$. If $\zeta(s)$ has a zero of multiplicity m at $1+it$ then $\zeta'(s)/\zeta(s)$ has a simple pole with residue m, and so $\Theta(s)$ has a simple pole with residue $-m$. Similarly, where $\zeta(s)$ has a pole of order M, $\Theta(s)$ has a simple pole with residue M.

We are going to compare

$$\Theta(1+\sigma),\ \Theta(1+it+\sigma),\ \Theta(1+2it+\sigma)$$

for small $\sigma > 0$ (Fig 3.1).

We have

$$\begin{aligned}\Theta(1+\sigma) &= \sum \log p\, p^{-(1+\sigma)},\\ \Theta(1+it+\sigma) &= \sum \log p\, p^{-(1+\sigma)} p^{-it},\\ \Theta(1+2it+\sigma) &= \sum \log p\, p^{-(1+\sigma)} p^{-2it}.\end{aligned}$$

Note that

$$\begin{aligned} p^{-it} &= \cos(t\log p) - i\sin(t\log p),\\ p^{-2it} &= \cos(2t\log p) - i\sin(2t\log p). \end{aligned}$$

Lemma 5. *For all $\theta \in \mathbb{R}$,*

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

Proof ▶ For $\tau \in \mathbb{R}$,

$$e^{i\tau} + e^{-i\tau} = 2\cos(\tau) \in \mathbb{R}.$$

Raising this to the fourth power,

$$\left(e^{i\tau} + e^{-i\tau}\right)^4 = e^{4i\tau} + e^{-4i\tau} + 4(e^{2i\tau}) + e^{-2i\tau}) + 6 \geq 0,$$

ie

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

The result follows on setting $\theta = 2\tau$. ◀

Lemma 6. *For $\sigma > 0$,*

$$\Re\left(\Theta(1 + 2i + \sigma) + 4\Theta(1 + it + \sigma) + 3\Theta(1 + \sigma)\right) \geq 0.$$

Proof ▶ We have

$$\Re\left(p^{-2it} + 4p^{-it} + 3\right) = \cos(2t\log p) + 4\cos(t\log p) + 3 \geq 0,$$

by the last Lemma.

The result follows on multiplying by $\log p\, p^{-(1+\sigma)}$ and summing. ◀

Remark. If we had taken squares instead of fourth powers, we would have found

$$\Re\left(\Theta(1 + it + \sigma) + \Theta(1 + \sigma)\right) \geq 0,$$

which is not quite sufficient for our purposes.

However, higher even powers would have done as well, eg sixth powers yield

$$\Re\left(\Theta(1 + 3it + \sigma) + 6\Re(1 + 2it + \sigma) + 15\Theta(1 + it + \sigma) + 10\Theta(1 + \sigma)\right) \geq 0,$$

which would have done.

Now suppose $\zeta(s)$ has a zero of multiplicity m at $s = 1 + it$, and suppose it also has a zero of multiplicity M at $s = 1 + 2it$, where we allow $M = 0$ if there is no zero. Then

$$\begin{aligned}\Theta(1+\sigma) &= \frac{1}{\sigma} + f_1(\sigma),\\ \Theta(1+it+\sigma) &= -\frac{m}{\sigma} + f_2(\sigma),\\ \Theta(1+2it+\sigma) &= -\frac{M}{\sigma} + f_3(\sigma),\end{aligned}$$

where $f_1(\sigma), f_2(\sigma), f_3(\sigma)$ are all continuous (and so bounded) for small σ.

Adding, and taking the real part,

$$\Re\left(\Theta(1+2i+\sigma) + 4\Theta(1+it+\sigma) + 3\Theta(1+\sigma)\right) = \frac{1-4m-3M}{\sigma} + f(\sigma),$$

where $f(\sigma)$ is continuous. By the last Lemma, this is ≥ 0 for all $\sigma > 0$. It follows that

$$1 - 4m - 3M \geq 0.$$

But that is impossible, since $m, n \in \mathbb{N}$ with $m > 0$. ◄

Remark. This proof is just a neat way of dressing up the following intuitive argument.

We know that $\Theta(s)$ has a pole at $s = 1$, with residue 1:

$$\Theta(1+\sigma) = \sum \log p\, p^{-(1+\sigma)} = \frac{1}{\sigma} + O(\sigma).$$

Note that the terms are all positive.

Now suppose $\zeta(s)$ has a zero of multiplicity m at $s = 1 + it$. Then $\Theta(s)$ has a pole at $s = 1 + it$ with residue $-m$:

$$\Theta(1+it+\sigma) = \sum \log p\, p^{-(1+\sigma)}\left(\cos(t\log p) + i\sin(t\log p)\right) = -\frac{m}{\sigma} + O(\sigma).$$

Comparing this with the formula for $\Theta(1+\sigma)$, and noting that

$$-1 \leq \cos(t\log p) \leq 1,$$

we see that in order to reach $-1/\sigma$ (let alone $-m/\sigma$), $\cos(t\log p)$ must be close to -1 for almost all p.

But

$$\cos\tau = -1 \implies \cos 2\tau = +1.$$

Thus follows that $\cos(2t\log p)$ is close to 1 for almost all p; and that in turn implies that

$$\Theta(1+2it+\sigma) = \sum \log p\, p^{-(1+\sigma}\left(\cos(2t\log p) + i\sin(2t\log p)\right)$$

is close to $1/\sigma$, which means that $\Theta(s)$ must have a pole with residue 1 (ie $\zeta(s)$ must have a simple pole) at $s = 1+2it$, which we know is not the case.

3.6 The Tauberian theorem

Proposition 3.9. *Suppose the function $f : [0,\infty) \to \mathbb{C}$ is*

1. *bounded; and*
2. *integrable over $[0, X]$ for all X.*

Then

$$F(s) = \int_0^\infty e^{-xs} f(x)dx$$

is defined and holomorphic in $\Re(s) > 0$.

Suppose $F(s)$ can be extended analytically to a holomorphic function in $\Re(s) \geq 0$. Then $f(x)$ is integrable on $[0,\infty)$, and

$$\int_0^\infty f(x)dx = F(0).$$

Proof ▶ Suppose

$$|f(x)| \leq C.$$

For each $X > 0$,

$$F_X(s) = \int_0^X e^{-xs} f(x)dx.$$

is an entire function, ie holomorphic in the whole of the complex plane $\mathbb{C}$.

Suppose $\sigma = \Re(s) > 0$. If $X < Y$ then

$$F_Y(s) - F_X(s) = \int_X^Y e^{-xs} f(x)dx.$$

Thus

$$\begin{aligned}|F_Y(s) - F_X(s)| &\leq C\int_X^Y e^{-x\sigma}dx \\ &= \frac{C}{\sigma}(e^{-X\sigma} - e{-}Y\sigma).\end{aligned}$$

Thus

$$F_Y(s) - F_X(s) \to 0 \text{ as } X, Y \to \infty.$$

Hence

$$F(s) = \int_0^\infty e^{-xs} f(x)dx$$

converges for $\Re(s) > 0$.

Moreover, our argument shows that this convergence is uniform in $\Re(s) \geq \sigma$ for each $\sigma > 0$. Hence $F(s)$ is holomorphic in each such half-plane, and so in $\Re(s) > 0$.

We have to show that

$$F_X(0) = \int_0^X f(x)dx \to F(0)$$

as $X \to \infty$. (This will prove both that $\int_0^\infty f(x)dx$ converges, and that its value is $F(0)$.)

By Cauchy's Theorem,

$$F_X(0) - F(0) = \frac{1}{2\pi i} \int_\gamma \left(F_X(s) - F(s)\right) \frac{ds}{s}$$

around any contour γ surrounding 0 within which $F(s)$ is holomorphic. We can even introduce a holomorphic factor $\lambda(s)$ satisfying $\lambda(0) = 1$:

$$F_X(0) - F(0) = \frac{1}{2\pi i} \int_\gamma \left(F_X(s) - F(s)\right) \lambda(s) \frac{ds}{s}.$$

We choose the contour γ in the following way. Suppose $R > 0$. (We shall later let $R \to \infty$.) By hypothesis, $F(s)$ is holomorphic at each point $s = it$ of the imaginary axis, ie it is holomorphic in some circle centred on $s = it$. It follows by a standard compactness argument that we can find a $\delta = \delta(R) > 0$ such that $F(s)$ is holomorphic in the rectangle

$$\{s = x + iy : -\delta \leq x \leq 0; -R \leq y \leq R\}.$$

To simplify the later computations we assume — as we evidently may — that $\delta \leq R$.

We take γ to be the contour formed by a large semicircle γ_1 of radius R in the positive half-plane, completed by 3 sides $\gamma_2 = \gamma_{2a} + \gamma_{2b} + \gamma_{2c}$ of the above rectangle in the negative half-plane (Fig 3.2).

We also choose our factor $\lambda(s)$ (for reasons that will become apparent) to be

$$\lambda(s) = e^{Xs}\left(1 + \frac{s^2}{R^2}\right).$$

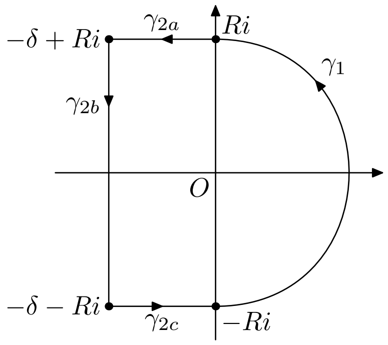

Figure 3.2: The contour γ

Note that we are playing with two constants, X and R, both tending to ∞. The interaction between them is subtle. First we fix R, and let $X \to \infty$. We shall show that there is a constant c such that

$$|F_X(0) - F(0)| \le c/R$$

for sufficiently large X. Since this holds for all R, it will show that

$$F_X(0) \to F(0) \text{ as } X \to \infty,$$

as required.

First we consider

$$I_1(X, R) = \int_{\gamma_1} (F_X(s) - F(s))\, \lambda(s) \frac{ds}{s}.$$

For $\sigma = \Re(s) > 0$,

$$F_X(s) - F(s) = \int_X^\infty e^{-xs} dx.$$

Thus

$$\begin{aligned} |F_X(s) - F(s)| &\le C \int_X^\infty e^{-x\sigma} dx \\ &= \frac{C}{\sigma} e^{-X\sigma}. \end{aligned}$$

As to the factor $\lambda(s)$,

$$|e^{Xs}| = e^{X\sigma};$$

while if $s = Re^{i\theta}$ then

$$1 + \frac{s^2}{R^2} = 1 + e^{2i\theta},$$

and so

$$|1 + \frac{s^2}{R^2}| = e^{i\theta} + e^{-i\theta} = 2\cos\theta = \frac{2\sigma}{R}.$$

Hence

$$\begin{aligned} |(F_X(s) - F(s))\,\lambda(s)| &\le \frac{C}{\sigma} e^{-X\sigma} \cdot e^{X\sigma} \frac{2\sigma}{R} \\ &= \frac{2C}{R}. \end{aligned}$$

(We see now how the two parts of $\lambda(s)$ were chosen to cancel out the factors $e^{-X\sigma}$ and $1/\sigma$.)

Since $s = Re^{i\theta}$,

$$\frac{ds}{s} = ie^{i\theta} d\theta;$$

and so

$$|I_1(X, R)| \le \frac{2C\pi}{R}.$$

Turning to the part γ_2 of the integral in the negative half-plane, we consider $F_X(s)$ and $F(s)$ separately:

$$I_2(X, R) = I_2'(X, R) + I_2''(X, R),$$

where

$$\begin{aligned} I_2'(X, R) &= \int_{\gamma_2} F_X(s)\lambda(s)\frac{ds}{s} \\ I_2''(X, R) &= \int_{\gamma_2} F(s)\lambda(s)\frac{ds}{s}. \end{aligned}$$

Since $F_X(s)$ is an entire function, we can replace the contour γ_2 in the integral $I_2'(X, R)$ by the half-circle γ_2' of radius R in the negative half-plane (Fig 3.3), ie the complementary half-circle to γ_1.

We have

$$F_X(s) = \int_0^X e^{-xs} f(x) dx.$$

Thus if $\sigma = \Re(s) \le 0$ then

$$\begin{aligned} |F_X(s)| &\le C \int_0^X e^{-x\sigma} dx \\ &\le \frac{C}{-\sigma} e^{-X\sigma}. \end{aligned}$$

As before,

$$|e^{Xs}| = e^{X\sigma};$$

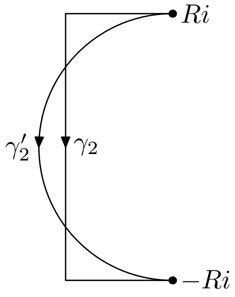

Figure 3.3: From γ_2 to $gamma_2'$

while

$$\begin{aligned}|1+\frac{s^2}{R^2}| &= |e^{i\theta}+e^{-i\theta}|\\ &= 2|\cos\theta|\\ &= \frac{-2\sigma}{R}.\end{aligned}$$

Thus

$$I_2'(X,R) \le \frac{2C\pi}{R}.$$

It remains to consider

$$I_2''(X,R) = \int_{\gamma_2} F(s)\lambda(s)\frac{ds}{s}.$$

We divide the integrand into two parts: the factor

$$e^{Xs} \to 0 \text{ as } X \to \infty$$

for all $s \in \gamma_2$ except for the two end-points $\pm Ri$; while the remaining factor

$$F(s)\left(1+\frac{s^2}{R^2}\right)\frac{1}{s}$$

is holomorphic in and on γ_2, and is therefore bounded there, say

$$|F(s)| \le D.$$

That is sufficient to show (for a given R) that

$$I_2''(X,R) \to 0 \text{ as } X \to \infty.$$

More precisely,

$$\begin{aligned}|I_{2a}(X,R)|, |I_{2c}(X,R)| &\le D\int_0^\delta e^{-X\sigma}d\sigma\\ &\le \frac{D}{X},\end{aligned}$$

while

$$|I_{2b}(X,R)| \le 2Re^{-X\delta}.$$

Thus all three parts of $I_2(X,R)$ tend to 0, and so

$$I_2(X,R) \to 0 \text{ as } X \to \infty$$

for each $R > 0$.

Putting all this together, we deduce that

$$|F_X(0) - F(0)| \le \frac{5C\pi}{R}$$

for sufficiently large X. It follows that

$$F_X(0) - F(0) \to 0 \text{ as } X \to \infty,$$

as required. ◀

3.7 Proof

We now have all the ingredients to complete the proof of the Prime Number Theorem.

Proof ▶ By Proposition 3.3, it is sufficient to prove that

$$\theta(x) \sim x.$$

We need to 'bootstrap' this result, by showing first that

$$\theta(x) = O(x).$$

Lemma 7. *There exists a constant C such that*

$$\theta(x) \le Cx$$

for all $x \ge 0$.

Proof ▶ Consider the binary coefficient

$$\binom{2n}{n} = \frac{(2n)(2n-1)\cdots(n+1)}{1\cdot 2\cdots n}.$$

This is of course an integer; and all the primes between n and $2n$ are factors, since each divides the top but not the bottom. Thus

$$\prod_{n<p\le 2n} p \le \binom{2n}{n}.$$

But

$$\binom{2n}{n} \le 2^{2n},$$

since the binomial coefficient is one term in the expansion of $(1+1)^{2n}$. Thus

$$\prod_{n<p\le 2n} p \le 2^{2n}.$$

Taking logarithms of both sides,

$$\theta(2n) - \theta(n) \le 2n \log 2.$$

Setting $n = 2^{m-1}, 2^{m-2}, \ldots$, successively,

$$\begin{aligned} \theta(2^m) - \theta(2^{m-1}) &\le 2^m \log 2, \\ \theta(2^{m-1}) - \theta(2^{m-2}) &\le 2^{m-1} \log 2, \\ &\cdots \\ \theta(2) - \theta(1) &\le 2 \log 2. \end{aligned}$$

Adding,

$$\begin{aligned} \theta(2^m) = \theta(2^m) - \theta(1) &\le (2^m + 2^{m-1} + \cdots + 2) \log 2 \\ &\le 2^{m+1} \log 2. \end{aligned}$$

Now suppose

$$2^{m-1} < x \le 2^m.$$

Then

$$\begin{aligned} \theta(x) &\le \theta(2^m) \\ &\le 2^{m+1} \log 2 \\ &= (4 \log 2) 2^{m-1} \\ &\le (4 \log 2) x. \end{aligned}$$

◀

Now let

$$\psi(x) = \theta(x) - x.$$

We have to show that

$$\psi(x) = o(x).$$

For $\Re(s) > 1$, let

$$\Psi(s) = \int_1^\infty x^{-s} d\psi.$$

Integrating by parts,

$$\begin{aligned}\int_1^X x^{-s} d\psi &= \left[x^{-s}\psi(x)\right]_1^X + s\int_1^X x^{-s}\psi(x)\frac{dx}{x} \\ &= X^{-s}\psi(X) - 1 + s\int_1^X x^{-s}\psi(x)\frac{dx}{x}.\end{aligned}$$

But

$$X^{-s}\psi(X) \to 0 \text{ as } X \to \infty$$

since

$$|\psi(X)| \le \max(\theta(X), X) \le C'X.$$

Thus

$$\begin{aligned}\Psi(x) &= 1 + s\int_1^\infty x^{-s}\psi(x)\frac{dx}{x} \\ &= 1 + s\int_1^\infty x^{-s}(\theta(x) - x)\frac{dx}{x} \\ &= 1 + s\left(\Theta(s) - \frac{1}{s-1}\right).\end{aligned}$$

Now $\Theta(s)$ has a pole at $s = 1$ with residue 1 (arising from the pole of $\zeta(s)$). It follows that $\Psi(s)$ is holomorphic at $s = 1$; and it has no poles elsewhere on $\Re(s) = 1$, since $\Theta(s)$ does not. Thus

$$\frac{1}{s}\left(\Psi(s) - 1\right) = \int_1^\infty x^{-s}\psi(x)\frac{dx}{x}$$

is holomorphic in $\Re(s) \ge 1$,

On making the change of variable $x = e^t$ (we can think of this as passing from the multiplicative group $\mathbb{R}^+$ to the additive group $\mathbb{R}$),

$$\begin{aligned}\frac{1}{s}\left(\Psi(s) - 1\right) &= \int_1^\infty x^{-s}\psi(x)\frac{dx}{x} \\ &= \int_0^\infty e^{-ts}\psi(e^t)dt.\end{aligned}$$

We are almost in a position to apply our Tauberian theorem. There is one last change; the theorem, as we expressed it, assumed that the critical line was the imaginary axis $\Re(s) = 0$. But the critical line of $\Psi(s)$ is $\Re(s) = 1$. We therefore set

$$s = 1 + s'.$$

We have

$$\begin{aligned}\frac{1}{s}\Psi(s) &= \int_0^\infty e^{-t(1+s')}\psi(e^t)dt\\ &= \int_0^\infty e^{-ts'}e^{-t}\psi(e^t)dt.\end{aligned}$$

Now we can apply the theorem, since

$$\begin{aligned}|e^{-t}\psi(e^t)| &\le |e^{-t}\theta(e^t|\\ &\le e^{-t}Ce^t\\ &\le C,\end{aligned}$$

ie $e^{-t}\psi(e^t)$ is bounded; while

$$\frac{1}{1+s'}\Psi(1+s')$$

is holomorphic on $\Re(s') = 0$.

We conclude that

$$\int_0^\infty e^{-t}\psi(e^t)dt$$

converges to $\Psi(1)$. (We only need the convergence, not the value.)

Changing variables back to $x = e^t$, we deduce that

$$\int_1^\infty \frac{\psi(x)}{x^2}dx = \int_1^\infty \frac{\theta(x) - x}{x^2}dx$$

converges.

It remains to show that this implies that

$$\theta(x) \sim x.$$

Suppose that were not so. Then either

$$\frac{\limsup \theta(x)}{x} > 1$$

or else

$$\frac{\liminf \theta(x)}{x} < 1$$

(or both). In other words, there exists a $\delta > 0$ such that either

$$\theta(X) \geq (1+\delta)X$$

for arbitrarily large X, or else

$$\theta(X) \leq (1-\delta)X$$

for arbitrarily large X.

Suppose

$$\theta(X) \geq (1+\delta)X.$$

Since $\theta(x)$ is increasing, it follows that

$$X \leq x \leq (1+\delta)X \implies \theta(x) \geq \theta(X) \geq (1+\delta)X \geq x,$$

ie

$$\theta(x) - x \geq 0$$

on the interval $[X, (1+\delta)X]$.

More precisely,

$$\begin{aligned}
\int_X^{(1+\delta)X} \frac{\theta(x)-x}{x^2} dx &\geq \int_X^{(1+\delta)X} \frac{(1+\delta)X - x}{x^2} dx \\
&\geq \int_1^{1+\delta} \frac{(1+\delta)-y}{y^2} dy, \text{ on setting } x = Xy, \\
&\geq \frac{1}{(1+\delta)^2} \int_1^{1+\delta} (1+\delta-y) dy \\
&\geq \frac{1}{(1+\delta)^2} \int_0^{\delta} u \, du \\
&= \frac{\delta^2}{2(1+\delta)^2}.
\end{aligned}$$

But the fact that there exist such intervals $[X, (1+\delta)X]$ with arbitrarily large X contradicts the convergence of

$$\int^{\infty} \frac{\theta(x)-x}{x^2} dx,$$

which we have already established. We conclude that

$$\limsup \frac{\theta(x)}{x} \leq 1.$$

Similarly, suppose

$$\theta(X) \le (1-\delta)X.$$

Since $\theta(x)$ is increasing, it follows that

$$(1-\delta)X \le x \le X \implies \theta(x) \le \theta(X) \le (1-\delta)X \le x,$$

ie

$$\theta(x) - x \le 0$$

on the interval $[(1-\delta)X, X]$.

More precisely,

$$\begin{aligned}
-\int_{(1-\delta)X}^{X} \frac{\theta(x)-x}{x^2} dx &= \int_{(1-\delta)X}^{X} \frac{x-\theta(x)}{x^2} dx \\
&\ge \int_{(1-\delta)X}^{X} \frac{x-(1-\delta)X}{x^2} dx \\
&\ge \int_{1-\delta}^{1} \frac{y-(1-\delta)}{y^2} dy \\
&\ge \frac{1}{(1-\delta)^2} \int_{1-\delta}^{1} (y-1+\delta) dy \\
&\ge \frac{1}{(1-\delta)^2} \int_0^{\delta} u\, du \\
&= \frac{\delta^2}{2(1-\delta)^2}.
\end{aligned}$$

Again, this contradicts the convergence of

$$\int^{\infty} \frac{\theta(x)-x}{x^2} dx.$$

Hence

$$\liminf \frac{\theta(x)}{x} \ge 1.$$

We have shown therefore that

$$\frac{\theta(x)}{x} \to 1,$$

ie

$$\theta(x) \sim x.$$

The proof of the Prime Number Theorem is complete. ◀

Chapter 4

The Dirichlet L-functions

4.1 Characters of a finite abelian group

4.1.1 Definition of a character

Definition 4.1. *A* character *of a finite abelian group A is a homomorphism*

$$\chi : A \to \mathbb{C}^\times.$$

The character defined by the trivial homomorphism is called the principal character *and is denoted by χ_1:*

$$\chi_1(a) = 1$$

for all $a \in A$.

Remarks. 1. We generally denote abelian groups multiplicatively — contrary perhaps to the usual practice — because the groups $(\mathbb{Z}/m)^\times$ to which we shall apply the theory are multiplicative.

2. For a map $\chi : A \to \mathbb{C}^\times$ to be a character it is sufficient that

$$\chi(ab) = \chi(a)\chi(b)$$

for all $a, b \in A$. For if that is so then

$$e^2 = e \implies \chi(e)^2 = \chi(e) \implies \chi(e) = 1;$$

and furthermore, if $a \in A$ then $a^n = e$ for some n by Lagrange's Theorem, so that

$$a^{-1} = a^{n-1},$$

and therefore

$$\chi(a^{-1}) = \chi(a^{n-1}) = \chi(a)^{n-1} = \chi(a)^{-1},$$

since

$$\chi(a)^n = \chi(a^n) = \chi(e) = 1.$$

Example. Suppose

$$A = C_n = \{e, g, g^2, \ldots, g^{n-1} : g^n = e\}.$$

Let $\omega = e^{2\pi i/n}$.

The cyclic group C_n has just n characters, namely

$$\chi^{(j)} : g^i \to \omega^{ij} \qquad (0 \le j < n).$$

For these are certainly characters of C_n; while conversely, if χ is such a character then

$$\begin{aligned} g^{n+1} = g &\implies \chi(g)^{n+1} = \chi(g^{n+1}) = \chi(g) \\ &\implies \chi(g) = \omega^j \text{ for some } j \in [0, n-1] \\ &\implies \chi = \chi^{(j)}. \end{aligned}$$

Proposition 4.1. *If χ is a character of the finite abelian group A then*

$$|\chi(a)| = 1$$

for all $a \in A$.

Proof ▶ By Lagrange's Theorem, $a^n = e$ for some n. Hence

$$\chi(a)^n = \chi(a^n) = \chi(e) = 1 \implies |\chi(a)| = 1.$$

◀

Proposition 4.2. *For any character χ of A,*

$$\chi(a^{-1}) = \overline{\chi(a)}.$$

Proof ▶ This follows at once from Proposition 4.1, since

$$|z| = 1 \implies z^{-1} = \bar{z}.$$

◀

4.1.2 The dual group A^*

Proposition 4.3. *The characters of a finite abelian group A form a group A^* under multiplication:*

$$(\chi\chi')(a) = \chi(a)\chi'(a).$$

The principal character χ_1 is the identity of A^*; and the inverse of χ is the character

$$\chi^{-1}(a) = \chi(a^{-1}) = \overline{\chi a}.$$

Proof ▶ The first part follows at once, since

$$\begin{aligned}(\chi\chi')(ab) &= \chi(ab)\chi'(ab)\\ &= (\chi(a)\chi(b))(\chi'(a)\chi'(b))\\ &= (\chi(a)\chi'(a))(\chi(b)\chi'(b))\\ &= (\chi\chi')(a)(\chi\chi')(b)\end{aligned}$$

The last two parts are trivial. ◀

Definition 4.2. *The group A^* of characters is called the* dual group *of A.*

Example. If $A = C_n$ then, as we have seen,

$$A^* = \{\chi^{(0)}, \chi^{(1)}, \dots, \chi^{(n-1)}\},$$

where $\chi^{(j)}(g^i) = \omega^{ij}$. It is easy to see that

$$\chi^{(i)}\chi^{(i')} = \chi^{(i+i' \bmod n)}.$$

It follows that the characters can be identified with the group $\mathbb{Z} \bmod m$; hence

$$C_n^* \cong \mathbb{Z}/(n) \cong C_n.$$

We may say that the cyclic group C_n is *self-dual.*

Proposition 4.4. *Every finite abelian group A is self-dual, ie*

$$A^* \cong A.$$

Proof ▶ We know that A is expressible as a product of cyclic groups:

$$A = C_{n_1} \times \cdots \times C_{n_r}.$$

Lemma 8. *If $A = B \times C$ then*

$$A^* = B^* \times C^*.$$

Proof ▶ We can identify B, C with the subgroups $B \times \{e\}$, $\{e\} \times C$ of A. Thus each character χ of A defines characters χ_B, χ_C of B, C by restriction. Moreover, since

$$(b, c) = (b, e) \cdot (e, c)$$

it follows that

$$\chi(b, c) = \chi_B(b)\chi_C(c).$$

This gives a one-one correspondence

$$\chi \longleftrightarrow (\chi_B, \chi_C)$$

between characters of A, and pairs of characters of B and C; and it is straightforward to verify that this is an isomorphism. ◀

It follows from the Lemma that

$$A^* = C_{n_1}^* \times \cdots \times C_{n_r}^*.$$

But we have seen that

$$C_n^* \cong C_n$$

for any cyclic group C_n. It follows that

$$A^* \cong A.$$

◀

Remark. This isomorphism is *non-canonical*, in the sense that there is no natural way of picking out one such isomorphism.

More precisely, the *functor*

$$A \rightsquigarrow A^*$$

is *contravariant*, ie each homomorphism

$$\alpha : A \to B$$

gives rise to a homomorphism

$$\alpha^* : B^* \to A^*$$

in the opposite direction; and there is no way in general of choosing an isomorphism $\theta : A \to A^*$ such that the diagram

$$\begin{tikzcd}
A \arrow[r, "\theta"] \arrow[d, "\alpha"] & A^* \\
A \arrow[r, "\theta"] & A^* \arrow[u, "\alpha^*"']
\end{tikzcd}$$

is commutative for all α.

If $a \in A$ then the map

$$\chi \mapsto \chi(a)$$

defines a character of the group A^*. This gives a (natural) homomorphism

$$A \to A^{**}.$$

Proposition 4.5. *For any finite abelian group*

$$A^{**} = A.$$

Proof ▶ Since

$$|A^{**}| = |A^*| = |A|,$$

it is sufficient (by the Pigeon-Hole Principle) to show that the map

$$A \to A^{**}$$

is injective.

Lemma 9. *If $a \neq e$ then there exists a character χ such that*

$$\chi(a) \neq 1.$$

Proof ▶ The elements

$$B = \{b : \chi(b) = 1 \text{ for all } \chi \in A^*\}$$

form a subgroup $B \subset A$; and every character of A is a character of the quotient-group A/B, ie

$$A^* = (A/B)^*.$$

But that is impossible unless $B = \{e\}$, since otherwise

$$|(A/B)^*| = |A/B| < |A| = |A^*|.$$

◀

We conclude that the homomorphism

$$A \to A^{**}$$

is injective, and is therefore an isomorphism. ◀

Remark. The character theory of finite abelian groups is a more-or-less trivial case of the character theory of locally compact abelian groups.

Each such group A has a dual A^*, consisting of the characters, ie continuous homomorphisms $\chi : A \to \mathbb{C}^\times$ such that $|\chi(a)| = 1$ for all a. (For compact abelian groups this last condition necessarily holds. But in the non-compact case we must impose it.)

For example, the additive group $\mathbb{R}$ is self-dual: $\mathbb{R}^* = \mathbb{R}$. This is the basis of the Fourier integral.

The dual of the torus $\mathbb{T}$ is the additive group of integers: $\mathbb{T}^* = \mathbb{Z}$. That is the basis of Fourier series.

The character theory of general locally compact abelian groups is sometimes called *generalised Fourier analysis.*

4.1.3 Sums over elements

Proposition 4.6. *Suppose χ is a character of the finite abelian group A. Then*

$$\sum_{a \in A} \chi(a) = \begin{cases} |A| & \text{if } \chi = \chi_1, \\ 0 & \text{otherwise.} \end{cases}$$

Proof ▶ If $\chi = \chi_1$, ie $\chi(a) = 1$ for all $a \in A$ then the sum is clearly $|A|$.

Suppose $\chi \neq \chi_1$. Then we can find a $b \in A$ such that

$$\chi(b) \neq 1.$$

As a runs over A so does ab. Hence

$$\begin{aligned} \sum_{a \in A} \chi(a) &= \sum_{a \in A} \chi(ab) \\ &= \chi(b) \sum_{a \in A} \chi(a). \end{aligned}$$

Thus

$$(\chi(b) - 1) \sum_a \chi(a) = 0,$$

and so

$$\sum_a \chi(a) = 0.$$

◀

Proposition 4.7. *Suppose χ, χ' are characters of the finite abelian group A. Then*

$$\sum_{a\in A} \overline{\chi(a)}\chi'(a) = \begin{cases} |A| & \text{if } \chi = \chi', \\ 0 & \text{otherwise.} \end{cases}$$

Proof ▶ By Proposition 4.3,

$$\begin{aligned} \overline{\chi(a)}\chi'(a) &= \chi^{-1}(a)\chi'(a) \\ &= (\chi^{-1}\chi')(a). \end{aligned}$$

Hence

$$\sum_{a\in A} \overline{\chi(a)}\chi'(a) = \sum_{a\in A} (\chi^{-1}\chi')(a),$$

and the result follows from Proposition 4.6, since

$$\chi^{-1}\chi' = \chi_1 \iff \chi = \chi'.$$

◀

4.1.4 Sums over characters

Proposition 4.8. *Suppose $a \in A$, where A is a finite abelian group. Then*

$$\sum_{\chi\in A^*} \chi(a) = \begin{cases} |A| & \text{if } a = e, \\ 0 & \text{otherwise.} \end{cases}$$

Proof ▶ If $a = e$ then $\chi(a) = 1$ for all $\chi \in A^*$ and the sum is evidently $|A|$.

Suppose $a \neq e$. By the Lemma to Proposition 4.5, we can find a $\chi' \in A^*$ such that

$$\chi'(a) \neq 1.$$

As χ runs over A^* so does $\chi'\chi$. Hence

$$\begin{aligned} \sum_{\chi\in A^*} \chi(a) &= \sum_{\chi\in A^*} (\chi'\chi)(a) \\ &= \chi'(a) \sum_{\chi\in A^a st} \chi(a). \end{aligned}$$

Thus

$$(\chi'(a) - 1) \sum_{\chi} \chi(a) = 0,$$

and so

$$\sum_{\chi} \chi(a) = 0.$$

◀

Proposition 4.9. *Suppose $ab \in A$, where A is a finite abelian group. Then*

$$\sum_{\chi \in A^*} \overline{\chi(a)}\chi(b) = \begin{cases} |A| & \textit{if } a = b, \\ 0 & \textit{otherwise.} \end{cases}$$

Proof ▶ Since

$$\begin{aligned} \overline{\chi(a)}\chi(b) &= \chi(a^{-1})\chi(a) \\ &= \chi(a^{-1}b). \end{aligned}$$

the result follows at once from Proposition 4.8. ◀

Remark. Alternatively, Propositions 4.8 and 4.9 follow at once from Propositions 4.6 and 4.7, on applying the latter to the dual group A^*, and using the fact that $A^{**} = A$.

4.1.5 Functions on a finite abelian group

Suppose A is a finite abelian group. The functions

$$f : A \to \mathbb{C}$$

form a vector space $C(A)$ (over $\mathbb{C}$) of dimension $|A|$, of which the $|A|$ characters of A are elements.

It is convenient to introduce an inner product in the space $C(A)$ of functions on A.

Definition 4.3. *If $f(a), g(a) \in C(A)$ we set*

$$\langle f \,|g\rangle = \frac{1}{|A|} \sum_{a \in A} \overline{f(a)} g(a).$$

It is a straightforward matter to verify that this is a positive-definite hermitian form, ie

1. $\langle g \,|f\rangle = \overline{\langle f \,|g\rangle}$;
2. $\langle f \,|f\rangle \geq 0$, and $\langle f \,|f\rangle = 0 \iff f = 0$;
3. $\langle f \,|\lambda_1 g_1 + \lambda_2 g_2\rangle = \lambda_1 \langle f \,|g_1\rangle + \lambda_2 \langle f \,|g_2\rangle$.

Now we can re-state Proposition 4.7 as follows.

Proposition 4.10. *The characters of A form an orthonormal set:*

$$\langle \chi' \,|\chi\rangle = \begin{cases} 1 & \text{if } \chi = \chi', \\ 0 & \text{otherwise.} \end{cases}$$

Corollary 4.1. *The characters are linearly independent.*

Proof ▶ Suppose

$$\sum_i \lambda_i \chi_i = 0.$$

Then

$$\begin{aligned} 0 &= \left\langle \chi_j \middle| \sum_i \lambda_i \chi_i \right\rangle \\ &= \sum_i \lambda_i \left\langle \chi_j | \chi_i \right\rangle \\ &= \lambda_j \end{aligned}$$

for all j. ◀

Corollary 4.2. *The characters form a basis for $C(A)$. Explicitly, each function $f : A \to \mathbb{C}$ is uniquely expressible as a linear combination of characters:*

$$f = \sum_\chi \lambda_\chi \chi,$$

with

$$\lambda_\chi = \langle \chi \,| f\rangle = \frac{1}{|A|} \sum_{a \in A} \overline{\chi(a)} f(a).$$

Proof ▶ The characters must form a basis for $C(A)$, since they are linearly independent and there are

$$|A| = \dim C(A)$$

of them.

So certainly

$$f = \sum_\chi \lambda_\chi \chi,$$

for some $\lambda_\chi \in \mathbb{C}$. To determine these coefficients, we take the inner-product with χ':

$$\begin{aligned} \langle \chi' \,| f\rangle &= \sum \langle \chi' \,| \lambda_\chi \chi\rangle \\ &= \lambda_{\chi'}. \end{aligned}$$

◀

We shall make use of one particular case of this.

Corollary 4.3. *Let $c_b(x)$ denote the* characterestic function *of the element b, ie*

$$c_b(a) = \begin{cases} 1 & \text{if } a = b, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\langle c_b | \chi \rangle = \overline{\chi(b)}/|A|$, and so

$$c_b = \frac{1}{|A|} \sum_{\chi \in A^*} \overline{\chi(b)} \chi.$$

4.2 Multiplicative characters mod$m$

Suppose $m \in \mathbb{N}$, $m \neq 0$. We denote the ring of residue classes mod$m$ by $\mathbb{Z}/(m)$. We can identify the classes in $\mathbb{Z}/(m)$ with their representatives r, $0 \leq r < m$.

Recall that we denote by $\phi(m)$ the number of residue classes coprime to m, ie

$$\phi(m) = \|\{r : 0 \leq r < m, \ \gcd(r, m) = 1\}\|.$$

Proposition 4.11. *The $\phi(m)$ residue classes coprime to m form a multiplicative group.*

Proof ▶ If r, s are coprime to m then so is rs. It remains to show that each such residue class has an inverse s mod m:

$$rs \equiv 1 \bmod m.$$

If $\gcd(r, m) = 1$ then the map

$$x \mapsto rx \bmod m : \mathbb{Z}/(m) \to \mathbb{Z}/(m)$$

is injective, since

$$rx \equiv ry \bmod m \implies m \mid r(x - y) \implies m \mid (x - y) \implies x \equiv y \bmod m.$$

It follows (by the Pigeon-Hole Principle) that this map is surjective. In particular there exists an s such that

$$rs \equiv 1 \bmod 1.$$

◀

Definition 4.4. *We denote this multiplicative group by* $\mathbb{Z}/m)^{\times}$.

Example.

$$\begin{aligned}
(\mathbb{Z}/1)^{\times} &= \{1\},\\
(\mathbb{Z}/2)^{\times} &= \{1\},\\
(\mathbb{Z}/3)^{\times} &= \{1,2\} = \{\pm 1\} \cong C_2,\\
(\mathbb{Z}/4)^{\times} &= \{1,3\} = \{\pm 1\} \cong C_2,\\
(\mathbb{Z}/5)^{\times} &= \{1,2,3,4\} = \{\pm 1, \pm 2\} \cong C_4,\\
(\mathbb{Z}/6)^{\times} &= \{1,5\} = \{\pm 1\} \cong C_2,\\
(\mathbb{Z}/7)^{\times} &= \{1,2,3,4,5,6\} = \{\pm 1, \pm 2, \pm 3\} \cong C_6,\\
(\mathbb{Z}/8)^{\times} &= \{1,3,5,7\} = \{\pm 1, \pm 3\} \cong C_2 \times C_2,\\
(\mathbb{Z}/9)^{\times} &= \{1,2,4,5,7,8\} = \{\pm 1, \pm 2, \pm 4\} \cong C_6,
\end{aligned}$$

Proposition 4.12. *Suppose* $m = m_1 m_2$*, where* $\gcd(m_1, m_2) = 1$*. Then*

$$(\mathbb{Z}/m)^{\times} = (\mathbb{Z}/m_1)^{\times} \times (\mathbb{Z}/m_2)^{\times}.$$

Proof ▶ By the Chinese Remainder Theorem, the ring-homomorphism

$$\Theta : \mathbb{Z}/(m) \to \mathbb{Z}/(m_1) \times \mathbb{Z}/(m_2) : x \mapsto (x \bmod m_1, x \bmod m_2)$$

is an isomorphism.

Suppose $r \in \mathbb{Z}/(m)$. Then

$$\gcd(r, m) = 1 \iff \gcd(r, m_1) = 1 = \gcd(r, m_2).$$

Hence Θ maps $(\mathbb{Z}/m)^{\times}$ onto $(\mathbb{Z}/m_1)^{\times} \times (\mathbb{Z}/m_2)^{\times}$, which proves the result. ◀

Example. Since $\gcd(4, 3) = 1$,

$$(\mathbb{Z}/12)^{\times} = (\mathbb{Z}/4)^{\times} \times (\mathbb{Z}/3)^{\times},$$

with the pairings

$$\begin{aligned}
1 &\mapsto (1,1),\\
5 &\mapsto (1,2),\\
7 &\mapsto (3,1),\\
11 &\mapsto (3,2).
\end{aligned}$$

Corollary 4.4. *If* $\gcd(m,n)=1$ *then*

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

Corollary 4.5. *Suppose*

$$m=p_1^{e_1}\cdots p_r^{e_r},$$

where $p_1,\dots,p_r$ *are distinct primes. Then*

$$(\mathbb{Z}/m)^\times=(\mathbb{Z}/p_1^{e_1})^\times\cdots\times(\mathbb{Z}/p_r^{e_r})^\times.$$

Thus the structure of the groups $(\mathbb{Z}/m)^\times$ is reduced to the structure of the groups $\mathbb{Z}/p^r)^\times$. Although we shall not make use of the following results, it may be helpful to know what these groups look like.

Proposition 4.13. *If* p *is prime then the group* $(\mathbb{Z}/p)^\times$ *is cyclic.*

Proof ▶ We have

$$(\mathbb{Z}/p)^\times=\{1,2,\dots,p-1\}.$$

Since each element of the ring $\mathbb{Z}/(p)$ except 0 is invertible, $Z/(p)$ is in fact a field.

By Lagrange's Theorem, if G is a group of order n then

$$g^n=e$$

for all $g\in G$. The smallest number $e>0$ such that

$$g^e=e$$

is called the *exponent* of G. By Lagrange's Theorem, $e\mid n$.

Lemma 10. *The exponent of* $(\mathbb{Z}/p)^\times$ *is* $p-1$.

Proof ▶ Each element $r\in(\mathbb{Z}/p)^*$ satisfies the equation

$$x^e-1=0$$

over the field $\mathbb{Z}/(p)$. But this equation has at most e roots. It follows that

$$p-1\le e.$$

Since $e\mid(p-1)$ it follows that

$$e=p-1.$$

◀

Lemma 11. *Suppose A is a finite abelian group of exponent e. Then A has an element of order e.*

Proof ▶ Let
$$e = p_1^{e_1} \cdots p_r^{e_r}.$$

For each i there must be an element a_i whose order is divisible by $p_i^{e_i}$; for otherwise p_i would occur to a lower power in the exponent e. Let
$$\operatorname{order}(a_i) = p_i^{e_i} q_i.$$

Then
$$b_i = a_i^{q_i}$$
has order e_i.

But if A is a finite abelian group, and $a, b \in A$ have orders r, s then
$$\gcd(r, s) = 1 \implies \operatorname{order}(ab) = rs.$$

For suppose $\operatorname{order}(ab) = n$. Then
$$(ab)^{rs} = 1 \implies n \mid rs.$$

On the other hand, since r, s are coprime we can find $x, y \in \mathbb{Z}$ such that
$$rx + sy = 1.$$

But then
$$(ab)^{sy} = a^{sy} = a^{1-rx} = a.$$

It follows that $r \mid n$. Similarly $s \mid n$. Since $\gcd(r, s) = 1$ this implies that
$$rs \mid n.$$

Hence
$$n = rs.$$

Applying this to
$$a = b_1 \cdots b_r$$
we conclude that a has order
$$p_1^{e_1} \cdots p_r^{e_r} = e.$$

◀

By these two Lemmas, we can find an element $a \in (\mathbb{Z}/p)^\times$ of order $p-1$. Hence $(\mathbb{Z}/p)^\times$ is cyclic. ◀

Generators of $(\mathbb{Z}/p)^\times$ are called *primitive roots* mod$p$.

If a is a primitive root mod$p$ then it is easy to see that a^r is a primitive root if and only if $\gcd(r, p-1) = 1$. It follows that there are $\phi(p-1)$ primitive roots mod$p$.

For example, there are just $\phi(6) = 2$ primitive roots $\bmod 7$, namely 3 and $5 = 3^{-1} \bmod 7$.

Proposition 4.14. *If p is an odd prime number then the multiplicative group*

$$(\mathbb{Z}/p^e)^\times$$

is cyclic for all $e \geq 1$.

Proof ▶ We have proved the result for $e = 1$. We derive the result for $e > 1$ in the following way.

The group $(\mathbb{Z}/p^e)^\times$ has order

$$\phi(p^e) = p^{e-1}(p-1).$$

By the last Proposition, there exists an element a with

$$\text{order}(a \bmod p) = p - 1.$$

Evidently

$$\text{order}(a \bmod p) \mid \text{order}(a \bmod p^e).$$

Thus the order of $a \bmod p$ is divisible by $p - 1$. It is therefore sufficient by Lemma 11 to show that there exists an element of order p^{e-1} in the group.

The elements of the form $x = 1 + py$ form a subgroup

$$S = \{x \in (\mathbb{Z}/p^e)^\times : x \equiv 1 \bmod p^e\}$$

of order p^{e-1}. It suffices therefore to show that this subgroup is cyclic.

That is relatively straightforward, since this group is 'almost additive'. Each element of the group has order p^j for some j. We have to show that some element $x = 1 + py$ has order p^{e-1}, ie

$$(1 + py)^{p^{e-2}} \not\equiv 1 \bmod p^e.$$

By the binomial theorem,

$$(1 + py)^{p^{e-2}} = 1 + p^{e-2}py + \binom{p^{e-2}}{2}p^2y^2 + \binom{p^{e-2}}{3}p^3y^3 + \cdots.$$

We claim that all the terms after the first two are divisible by p^e, ie

$$p^e \mid \binom{p^{e-2}}{r} p^r y^r$$

for $r \geq 2$. To see this, note that

$$\begin{aligned}\binom{p^{e-2}}{r} p^r &= \frac{p^{e-2}(p^{e-2}-1)\cdots(p^{e-2}-r+1)}{1\cdot 2\cdots r} p^r \\ &= \frac{(p^{e-2}-1)\cdots(p^{e-2}-r+1)}{1\cdot 2\cdots (r-1)} p^{e-2}\frac{p^r}{r} \\ &= \binom{p^{e-2}-1}{r-1} p^{e-2}\frac{p^r}{r}.\end{aligned}$$

Thus it is sufficient to show that

$$p^2 \mid \frac{p^r}{r}$$

for $r \geq 2$; and that follows at once from the fact that

$$p^{r-1} > r,$$

eg because

$$p^{r-1} > (1+1)^{r-1} \geq 1 + (r-1) = r.$$

Thus any element of the form $1 + py$ where y is not divisible by p (for example, $1 + p$) must have multiplicative order p^{e-1}, and so must generate S. In particular the subgroup S is cyclic, and so $(\mathbb{Z}/p^e)^\times$ is cyclic. ◀

Turning to $p = 2$, it is evident that $(\mathbb{Z}/2)^\times$ is trivial, while $(\mathbb{Z}/4)^\times = C_2$.

Proposition 4.15. *If $e \geq 3$ then*

$$(\mathbb{Z}/2^e)^\times \cong C_2 \times C_{2^{e-2}}.$$

Proof ▶ Since

$$\phi(2^e) = 2^{e-1},$$

$(\mathbb{Z}/2^e)^\times$ contains 2^{e-1} elements. By the Structure Theorem for finite abelian groups, it is sufficient to show that $(\mathbb{Z}/2^e)^\times$ has exponent 2^{e-2}. For then one of the factors in

$$(\mathbb{Z}/2^e)^\times = C_{2^{e_1}} \times \cdots \times C_{2^{e_r}}$$

must be $C_{2^{e-2}}$, and the remaining factor must be C_2.

This is certainly true for $(\mathbb{Z}/8)^\times = \{\pm 1, \pm 3\}$, since

$$(\pm 1)^2 = (\pm 3)^2 = 1.$$

It follows that $(\mathbb{Z}/2^e)^\times$ cannot be cyclic for $e > 3$; for if a generated $(\mathbb{Z}/2^e)^\times$ then it would generate $\mathbb{Z}/8)^\times$. (In effect, $(\mathbb{Z}/8)^\times$ is a quotient group of $(\mathbb{Z}/2^e)^\times$.) Thus the Proposition will be proved if we can find an element of order 2^{e-2} mod 2^e.

We argue as we did for odd p, except that now we take $x = 1 + 2^2 y$. By the binomial theorem,

$$(1 + 2^2 y)^{2^{e-3}} = 1 + 2^{e-3} 2^2 y + \binom{2^{e-3}}{2} 2^4 y^2 + \binom{2^{e-3}}{3} 2^6 y^3 + \cdots .$$

As before, all the terms after the first two are divisible by 2^e, ie

$$2^e \mid \binom{p^{e-3}}{r} 2^{2r} y^r$$

for $r \geq 2$. For

$$\begin{aligned}
\binom{2^{e-3}}{r} 2^{2r} &= \frac{2^{e-3}(2^{e-3}-1)\cdots(2^{e-3}-r+1)}{1 \cdot 2 \cdots r} 2^{2r} \\
&= \frac{(2^{e-3}-1)\cdots(2^{e-3}-r+1)}{1 \cdot 2 \cdots (r-1)} 2^{e-3} \frac{2^{2r}}{r} \\
&= \binom{2^{e-3}-1}{r-1} 2^{e-3} \frac{2^{2r}}{r}.
\end{aligned}$$

Thus it is sufficient to show that

$$2^3 \mid \frac{2^{2r}}{r}$$

for $r \geq 2$; and that follows at once from the fact that

$$2^{2(r-1)} > r,$$

eg because

$$2^{2(r-1)} = (1+1)^{2(r-1)} \geq 1 + 2(r-1) + 1 = 2r.$$

Thus any element of the form $1 + 2^2 y$ with y odd (for example, 5) must have multiplicative order 2^{e-1}, which as we have seen is sufficient to prove the result. ◀

4.2.1 Characters and multiplicative functions

Suppose χ is a character of $(\mathbb{Z}/m)^\times$. Thus in principle χ is a function

$$\chi : (\mathbb{Z}/m)^* \to \mathbb{C}^\times.$$

However, we extend χ to a function

$$\chi : \mathbb{Z}/(m) \to \mathbb{C},$$

by setting

$$\chi(r) = 0 \text{ if } \gcd(r, m) > 1.$$

Now we extend χ to a function

$$\chi : \mathbb{N} \to \mathbb{Z}/(m) \to \mathbb{C}$$

by composition. (It should cause no confusion that we use the same symbol χ for all three functions.)

For example, suppose $m = 6$. Since $\phi(6) = 2$, there are just 2 multiplicative characters mod6, the principal character χ_1 and the character

$$\chi(r) = \begin{cases} 1 & \text{if } r \equiv 1 \bmod 6, \\ -1 & \text{if } r \equiv 5 \bmod 6. \end{cases}$$

The corresponding function $\chi : \mathbb{N} \to \mathbb{C}$ is given by

$$\chi(n) = \begin{cases} 0 & \text{if } n \equiv 0, 2, 3, 4 \bmod 6, \\ 1 & \text{if } n \equiv 1 \bmod 6, \\ -1 & \text{if } n \equiv 5 \bmod 6. \end{cases}$$

Recall that a function

$$\chi : \mathbb{N} \to \mathbb{C}$$

is said to be *multiplicative* if

$$\gcd(m, n) = 1 \implies \chi(mn) = \chi(m)\chi(n),$$

and $\chi(0) = 0,\ \chi(1) = 1$. (We include the latter condition to exclude the case where $f(n) = 0$ for all n).

We say that $\chi(n)$ is *strictly multiplicative* if

$$\chi(mn) = \chi(m)\chi(n)$$

for all $m, n \in \mathbb{N}$, and $\chi(0) = 0\ \chi(1) = 1$.

Proposition 4.16. *If χ is a multiplicative character* mod$m$ *then the corresponding function*

$$\chi : \mathbb{N} \to \mathbb{C}$$

is strictly multiplicative.

Proof ▶ This is immediate; for if r or s is not coprime to m then neither is rs, and so

$$\chi(rs) = 0 = \chi(r)\chi(s).$$

◀

Let us say that a function $f : \mathbb{N} \to \mathbb{C}$ is *modular* with modulus m if

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

for all n.

If is clear that if $d \mid m$ then any multiplicative character mod$d$ defines a function which is modular with modulus m.

The following result shows that every function $f : \mathbb{N} \to \mathbb{C}$ which is both strictly multiplicative and modular arises in this way.

Proposition 4.17. *Suppose $f : \mathbb{N} \to \mathbb{C}$ is modular* mod$m$*. Then $f(n)$ is strictly multiplicative if and only if it is defined by a multiplicative character* mod$d$ *for some $d \mid m$.*

Proof ▶ We argue by induction on m.

Suppose

$$f(d) \neq 0$$

for some proper divisor $d \mid m$, $1 < d < m$. Then

$$\begin{aligned} r \equiv s \bmod m/d &\implies rd \equiv sd \bmod m \\ &\implies f(rd) = f(sd) \\ &\implies f(r)f(d) = f(s)f(d) \\ &\implies f(r) = f(s). \end{aligned}$$

Thus $f(n)$ is modular modm/d. It follows from our inductive hypothesis that $f(n)$ is defined by a multiplicative character mod$e$ for some $e \mid d \mid m$.

Suppose to the contrary that

$$d \mid m,\ d > 1 \implies f(d) = 0;$$

and suppose $d = \gcd(r, m) > 1$, say

$$r = dr',\ m = dm'.$$

Then

$$f(r) = f(d)f(r') = 0.$$

On the other hand, if $\gcd(r, m) = 1$ then r has a multiplicative inverse $\bmod m$, say

$$rs \equiv 1 \bmod m;$$

and so

$$f(r)f(s) = f(1) = 1 \implies f(r) \neq 0.$$

It follows that $f(n)$ is defined by a function

$$\chi : (\mathbb{Z}/m)^\times \to \mathbb{C}^\times,$$

which is readily seen to be a multiplicative character $\bmod m$. ◀

4.3 Dirichlet's L-functions

Dirichlet observed that Euler's Product Formula could be extended to include mutliplicative factors. Informally, if $\chi(n)$ is multiplicative then

$$\sum \chi(n)n^{-s} = \prod_{\text{primes } p} F_p(s),$$

where

$$F_p(s) = 1 + \chi(p)p^{-s} + \chi(p^2)p^{-2s} + \cdots .$$

This follows from the fact that if

$$n = p_1^{e_1} \cdots p_r^{e_r}$$

then

$$\chi(n) = \chi(p_1^{e_1}) \cdots \chi(p_r^{e_r}),$$

and so

$$\chi(n)n^{-s} = \left(\chi(p_1^{e_1})n^{-e_1 s}\right) \cdots \left(\chi(p_r^{e_r})n^{-e_r s}\right),$$

If $\chi(n)$ is strictly multiplicative then

$$\begin{aligned} F_p(s) &= 1 + \chi(p)p^{-s} + \chi(p)^2 p^{-2s} + \cdots \\ &= \left(1 - \chi(p)p^{-s}\right)^{-1}; \end{aligned}$$

and so

$$\sum \chi(n)n^{-s} = \prod_p \left(1 - \chi(p)p^{-s}\right)^{-1}.$$

Definition 4.5. *Suppose χ is a multiplicative character* $\bmod m$*, regarded as a function $\chi : \mathbb{N} \to \mathbb{C}$. Then the Dirichlet L-function corresponding to χ is defined by the Dirichlet series*

$$L_\chi(s) = \sum_{n \in \mathbb{N}} \chi(n) n^{-s}.$$

Proposition 4.18. *Suppose χ is a multiplicative character* $\bmod m$*.*

If $\chi \neq \chi_1$ then the Dirichlet series $L_\chi(s)$ converges in the half-plane $\Re(s) > 0$, and thus defines a holomorphic function there.

If $\chi = \chi_1$ then $L_\chi(s)$ converges in the half-plane $\Re(s) > 1$. However, this function can be continued analytically to the half-plane $\Re(s) > 0$, in which it has a single simple pole at $s = 1$, with residue $\phi(m)/m$.

Proof ▶ Let

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

Lemma 12. *If $\chi \neq \chi_1$ then $S(x)$ is bounded. More precisely,*

$$|S(x)| \leq \phi(m).$$

Proof ▶ By Proposition 4.1,

$$\sum_{r \in (\mathbb{Z}/m)^\times} \chi(r) = 0.$$

It follows that

$$\sum_{r \in \mathbb{Z}/(m)} \chi(r) = 0,$$

ie $\sum \chi(r)$ vanishes over any complete set of residues. Hence

$$S(mq - 1) = \sum_{n=0}^{mq-1} \chi(n) = 0.$$

for any q. Now suppose $mq \leq x < m(q + 1)$. Then

$$S(x) = \sum_{n=mq}^{[x]} \chi(n).$$

This sum contain $\leq m$ terms, of which at most $\phi(m)$ are non-zero. Since $|\chi(n)| = 1$ for each of these terms, we conclude that

$$|S(x)| \leq \phi(m).$$

◀

Remark. In fact it is easy to see that

$$|S(x)| \le \frac{\phi(m)}{2}.$$

For

$$S(x) = \sum_{n=mq}^{[x]} \chi(n) = -\sum_{[x]+1}^{m(q+1)-1} \chi(n);$$

and these two sums together contain $\phi(m)$ non-zero terms, so one of them contains $\le \phi(m)/2$ such terms.

Integrating by parts,

$$\begin{aligned}\sum_M^N \chi(n)n^{-s} &= \int_M^N x^{-s} dS \\ &= [x^{-s}S(x)]_M^N + s\int_M^N x^{-s}S(x)\frac{dx}{x}.\end{aligned}$$

Thus

$$\begin{aligned}|\sum_M^N \chi(n)n^{-s}| &\le \phi(m)(M^{-\sigma} + N^{-\sigma}) + |s|\phi(m)\int_M^N x^{-\sigma}\frac{dx}{x} \\ &= \phi(m)\left(M^{-\sigma} + N^{-\sigma} + \frac{|s|}{\sigma}(M^{-\sigma} - N^{-\sigma}\right).\end{aligned}$$

Since $M^{-\sigma}, N^{-\sigma} \to 0$ as $M, N \to \infty$, it follows that

$$|\sum_M^N \chi(n)n^{-s}| \to 0$$

as $M, N \to \infty$. Hence the series is convergent, by Cauchy's criterion.

Now suppose $\chi = \chi_1$. Let

$$h(s) = L_\chi(s) - \frac{\phi(m)}{m} = \sum a(n)n^{-s},$$

where

$$a(n) = \begin{cases} 1 - \phi(m)/m & \text{if } \gcd(n,m) = 1 \\ -\phi(m)/m & \text{if } \gcd(n,m) > 1 \end{cases}$$

Evidently,

$$\sum_{r\in\mathbb{Z}/(m)} a(r) = 0,$$

while $|a(n)| < 1$ for all $n \in \mathbb{N}$. It follows by the argument we used above that the Dirichlet series

$$h(s) = \sum_{n \geq 1} a(n) n^{-s}$$

converges in $\Re(s) > 0$, and so defines a holomorphic function there.

We conclude that

$$L_\chi(s) = \frac{\phi(m)}{m} \zeta(s) + h(s)$$

defines the analytic continuation of $L_\chi(s)$ to $\Re(s) > 0$, with the only pole arising from the pole of $\zeta(s)$ at $s = 1$. ◀

Proposition 4.19. *Suppose χ is a multiplicative character* mod$m$. *Then*

$$L_\chi(s) = \prod_{primes\ p} \left(1 - \chi(p) p^{-s}\right)^{-1}$$

for $\Re(s) > 1$.

Proof ▶ This follows in exactly the same way as for $\zeta(s)$. Thus if $\Re(s) > 1$ then

$$\prod_{p \leq N} \left(1 - \chi(p) p^{-s}\right)^{-1} = \sum_{n \leq N} \chi(n) n^{-s} + \sum{}' \chi(n) n^{-s},$$

where the second sum on the right extends over those $n > N$ all of whose prime factors are $\leq N$.

The sum

$$\sum_{n \in \mathbb{N},\, n \neq 0} \chi(n) n^{-s}$$

converges absolutely for $\Re(s) > 1$, by comparison with $\zeta(s)$, since

$$\sum_M^N |\chi(n) n^{-s}| \leq \sum_M^N |n^{-s}|.$$

It follows that

$$\prod_{p \leq N} \left(1 - \chi(p) p^{-s}\right)^{-1} \to L_\chi(s)$$

as $N \to \infty$.

◀

Chapter 5

Dirichlet's Theorem

Definition 5.1. *Suppose* $r, m \in \mathbb{N}$*. We denote by* $\pi_{r,m}$ *the number of primes* $p \le x$ *congruent to* $r \bmod m$*:*

$$\pi_{r,m}(x) = \|\{p \le x : p \equiv r \bmod m\}\|.$$

If we suppose — as we may — that $0 \le r < m$ then $\pi_{r,m}(x)$ measures the number of primes $\le x$ in the arithmetic sequence

$$r,\ r+m,\ r+2m,\ \dots.$$

If r and m have a factor in common then clearly there is at most one prime in this sequence, namely its first element r if r is prime:

$$\gcd(r, m) > 1 \implies \pi_{r,m}(x) \le 1.$$

We are not interested in this trivial case.

Proposition 5.1. *(Dirichlet's Theorem) If* $\gcd(r, m) = 1$ *then*

$$\pi_{r,m} \sim \frac{\mathrm{Li}(x)}{\phi(m)} \sim \frac{1}{\phi(m)} \frac{x}{\log x}.$$

Remarks. 1. It is not strictly accurate to speak of this as *Dirichlet's Theorem*, since Dirichlet only showed that if $\gcd(r, m) = 1$ then there are an infinity of primes in the arithmetic sequence $r, r+m, r+2m, \dots$.

However, his argument, when combined with the techniques used to prove the Prime Number Theorem in Chapter 3, immediately yields the stronger result above; so it is not unreasonable to give Dirichlet's name to the theorem.

2. Our proof of Dirichlet's Theorem closely mirrors our earlier proof of the Prime Number Theorem; and where the arguments are identical we refer to the earlier proof for details.

   As in the earlier case, we start with a preview, followed by some preliminary results, before giving the proof proper.

5.1 Preview of the proof

This preview should be read in conjunction with our earlier preview (Section 3.2) of the proof of the Prime Number Theorem.

1. We start from the analogue to Euler's Product Formula:

$$L_\chi(s) = \prod_{\text{primes } p} \left(1 - \chi(p)p^{-s}\right)^{-1}.$$

2. Logarithmic differentiation converts this to

$$\begin{aligned} \frac{L'_\chi(s)}{L_\chi(s)} &= -\sum_p \frac{\chi(p)\log p\, p^{-s}}{1-\chi(p)p^{-s}} \\ &= -\sum_n a_n \chi(n) n^{-s}, \end{aligned}$$

   where

$$a_n = \begin{cases} \log p & \text{if } n = p^e \\ 0 & \text{otherwise.} \end{cases}$$

3. Now we use the fact that we can pick out a particular residue class by taking an appropriate linear combination of characters:

$$\frac{1}{\phi(m)} \sum_\chi \overline{\chi(r)} \frac{L'_\chi(s)}{L_\chi(s)} = \sum_{n \equiv r \bmod m} a_n n^{-s},$$

   where the sum on the left runs over all the multiplicative characters $\bmod m$.

4. As before, it is convenient to 'hive off' the part of the Dirichlet series on the right corresponding to higher prime-powers:

$$\sum a_n n^{-s} = \Theta_{r,m}(s) + h(s),$$

   where

$$\Theta_{r,m}(s) = \sum_{p \equiv r \bmod} \log p\, p^{-s},$$

while $h(s)$ converges absolutely for $\Re(s) > 1/2$, by comparison with $\zeta(2s)$, and so may be ignored in our argument.

5. As before (again!),
$$\Theta_{r,m}(s) = \int_0^\infty x^{-s} d\theta_{r,m},$$
where
$$\theta_{r,m}(x) = \sum_{p \le x,\, p \equiv r \bmod m} \log p.$$

6. The argument by which we showed before that
$$\pi(x) \sim \frac{x}{\log x} \iff \theta(x) \sim x$$
now shows that
$$\pi_{r,m}(x) \sim \frac{1}{\phi(m)} \frac{x}{\log x} \iff \theta_{r,m}(x) \sim \frac{x}{\phi(m)}.$$
Accordingly, the proof of Dirichlet's Theorem is reduced to showing that
$$\theta_{r,m}(x) \sim \frac{x}{\phi(m)},$$
ie
$$\theta_{r,m}(x) = \frac{x}{\phi(m)} + o(x).$$

7. The function $L'_\chi(s)/L_\chi(s)$ has poles wherever $L_\chi(s)$ has a pole or zero. It follows from the Product Formula that $L_\chi(s)$ has no zeros in $\Re(s) > 1$. Accordingly
$$\Theta_{r,m}(s) = -\frac{1}{\phi(m)} \sum_\chi \overline{\chi(r)} \frac{L'_\chi(s)}{L_\chi(s)} + h(s)$$
is holomorphic in $\Re(s) > 1$.

8. As with the Prime Number Theorem, the fundamental problem is to determine what happens on the line $\Re(s) = 1$. The heart of Dirichlet's Theorem is the proof that none of the L-functions has a zero on this line:
$$\Re(s) = 1 \implies L_\chi(s) \ne 0.$$

The proof that $L_\chi(1+it) \neq 0$ for $t \neq 0$ is straightforward; in effect, the proof that $\zeta(1+it) \neq 0$ carries over unchanged. But now we have to prove also that

$$L_\chi(1) \neq 0$$

for $\chi \neq \chi_1$; and this turns out to be a much more formidable task.

9. Having got over this hurdle, it follows that $\Theta_{r,m}(s)$ has a simple pole at $s = 1$, arising from the pole of $L_{\chi_1}(s)$, with residue $1/\phi(m)$, and no other poles on the line $\Re(s) = 1$.

10. The rest of the proof is as before. We 'remove' the pole at $s = 1$ by subtracting an appropriate multiple of $\zeta(s)$. Thus

$$\Psi_{r,m}(s) = \Theta_{r,m}(s) - \frac{1}{\phi(m)}\zeta(s)$$

is holomorphic in $\Re(s) \geq 1$; and

$$\Psi_{r,m}(s) = \int_1^\infty x^{-s} d\psi_{r,m},$$

where

$$\begin{aligned}\psi_{r,m}(x) &= \theta_{r,m}(x) - \frac{1}{\phi(m)}[x] \\ &= \theta_{r,m}(x) - \frac{1}{\phi(m)}x + O(1).\end{aligned}$$

11. The Tauberian Theorem now shows that

$$\int_1^\infty \frac{\psi_{r,m}(x)}{x^2} dx$$

converges. (Note that the bootstrap lemma — Lemma 7 — carries over since

$$\theta_{r,m}(x) \leq \theta(x)$$

for all x.)

From this we deduce, as before, that

$$\theta_{r,m}(x) \sim \frac{x}{\phi(m)};$$

and that, as we have seen, establishes Dirichlet's Theorem.

5.2 From $\pi_{r,m}(x)$ to $\theta_{r,m}(x)$

Definition 5.2. *For $r, m \in \mathbb{N}$ we set*

$$\theta_{r,m}(x) = \sum_{p \le x,\ p \equiv r \bmod m} \log p.$$

Proposition 5.2. *If* $\gcd(r,m) = 1$ *then*

$$\pi_{r,m}(x) \sim \frac{\mathrm{Li}(x)}{\phi(m)} \iff \theta_{r,m}(x) \sim \frac{x}{\phi(m)}.$$

Proof ▶ This is in effect a re-wording of Proposition 3.3, taking $\phi(m)\pi_{r,m}(x)$ in place of $\pi(x)$, and $\phi(m)\theta_{r,m}(x)$ in place of $\theta(x)$, ◀

Corollary 5.1. *Dirichlet's Theorem is equivalent to:*

$$\theta_{r,m}(x) \sim \frac{x}{\phi(m)}$$

for $\gcd(r,m) = 1$*.*

5.3 Picking out the residue class

Definition 5.3. *For $r, m \in \mathbb{N}$ we set*

$$\Theta_{r,m}(s) = \sum_{p \equiv r \bmod m} \log p\, p^{-s}.$$

Proposition 5.3. *If* $\gcd(r,m) = 1$ *then*

$$\frac{1}{\phi(m)} \sum_{\chi} \bar{\chi}(r) \frac{L'_\chi(s)}{L_\chi(s)} = -\Theta_{r,m}(s) + h(s),$$

where $h(s)$ is holomorphic in $\Re(s) > 1/2$.

Proof ▶ If $\Re(s) > 1$ then by Proposition 4.19

$$L_\chi(s) = \prod \left(1 - \chi(p)p^{-s}\right)^{-1}.$$

Differentiating logarithmically,

$$\begin{aligned} \frac{L'_\chi(s)}{L_\chi(s)} &= -\sum_p \frac{\chi(p)\log p\, p^{-s}}{1 - \chi(p)p^{-s}} \\ &= -\Theta_{r,m}(s) + h_{r,m}(s), \end{aligned}$$

where

$$h_{r,m}(s) = -\sum_p \log p \sum_{p^e \equiv r \bmod m} p^{-es}.$$

Since the function $h_{r,m}(s)$ consists of certain terms taken from the corresponding series for $h(s)$ in Proposition 3.7, and since we showed that this series converges absolutely for $\Re(s) > 1/2$, it follows that $h_{r,m}(s)$ also converges absolutely, and so is holomorphic, in $\Re(s) > 1/2$. ◀

5.4 The zeros of $L_\chi(s)$

Proposition 5.4. *Suppose χ is a multiplicative character* mod$m$. *Then $L_\chi(s)$ has no zeros in $\Re(s) > 1$.*

Proof ▶ This follows at once from the product formula for $L_\chi(s)$, like the corresponding result for $\zeta(s)$. ◀

Proposition 5.5. *If $t \neq 0$ then*

$$L_\chi(1 + it) \neq 0.$$

Proof ▶ Consider

$$\begin{aligned}\Theta_{1,m}(s) &= \sum_{p \equiv 1 \bmod m} \log p \, p^{-s} \\ &= -\frac{1}{\phi(m)} \sum_\chi \frac{L'_\chi(s)}{L_\chi(s)} + h(s),\end{aligned}$$

where $h(s)$ is holomorphic in $\Re(s) > 1/2$ (and so may be ignored).

Each character χ for which $L_\chi(1 + it) = 0$ will contribute to the residue of $\Theta_{r,m}(s)$ at $s = 1 + it$. More precisely, if the multiplicity of this zero is m_χ then

$$\mathrm{res}_{1+it}(\Theta_{r,m}) = -\frac{1}{\phi(m)} \sum_\chi m_\chi.$$

(If $L_\chi(1 + it) \neq 0$ then we set $m_\chi = 0$.) Similarly, if each $L_\chi(s)$ has a zero with multiplicity M_χ at $s = 1 + 2it$ then

$$\mathrm{res}_{1+2it}(\Theta_{r,m}) = -\frac{1}{\phi(m)} \sum_\chi M_\chi.$$

We know that $L_{\chi_1}(s)$ has a simple pole at $s = 1$. Suppose that, for $\chi \neq \chi_1$, $L_\chi(s)$ has a zero with multiplicity μ_χ at $s = 1$. Then

$$\mathrm{res}_1(\Theta_{r,m}) = \frac{1}{\phi(m)} \left(1 - \sum_{\chi \neq \chi_1} \mu_\chi\right).$$

But now, applying Lemma 6 to $\Theta_{1,m}(s)$ in exactly the same way that we applied it to $\Theta(s)$,

$$\Re\left(\Theta_{1,m}(1+2i+\sigma)+4\Theta_{1,m}(1+it+\sigma)+3\Theta_{1,m}(1+\sigma)\right)\geq 0$$

for any $\sigma>0$; and from this it follows, as before, that

$$\operatorname{res}_{1+2i}(\Theta_{1,m})+4\operatorname{res}_{1+i}(\Theta_{1,m})+3\operatorname{res}_{1}(\Theta_{1,m})\geq 0,$$

ie

$$\sum M_\chi+4\sum m_\chi+3\sum \mu_\chi\leq 3.$$

Since M_χ, m_χ, μ_χ are all non-negative integers, this implies that

$$m_\chi=0 \text{ for all } \chi.$$

(For if $m_\chi\geq 1$ for any χ this will already 'out-vote' the right-hand side.) In other words,

$$L_\chi(1+it)\neq 0.$$

◀

Proof ▶ In the proof above, we lumped all the $L_\chi(s)$ together. We can equally well consider the $L_\chi(s)$ separately, by modifying Lemma 6 slightly, as follows.

Lemma 13. *Let*

$$\Theta_\chi(s)=\sum\chi(p)\log p\,p^{-s}.$$

Then

$$\Re\left(\Theta_\chi(1+2i+\sigma)+4\Theta_\chi(1+it+\sigma)+3\Theta_\chi(1+\sigma)\right)\geq 0$$

for any $\sigma>0$.

Proof ▶ If $\chi(p)\neq 0$ then $|\chi(p)|=1$, say

$$\chi(p)=e^{i\theta_p}.$$

Since $\chi(n)$ is strictly multiplicative,

$$\chi^2(p)=(\chi(p))^2=e^{2i\theta_p}.$$

It follows that

$$\Re\left(\chi^2(p)p^{-2it}+4\chi(p)p^{-it}+3\right)=\cos\left(2(t\log p+\theta_p)\right)+4\cos(t\log p+\theta_p)+3\geq 0,$$

by Lemma 5, with $\theta=t\log p+\theta_p$. ◀

We deduce, as before, that

$$\mathrm{res}_{1+2i}(\Theta_{\chi^2}) + 4\,\mathrm{res}_{1+i}(\Theta_\chi) + 3 \geq 0,$$

ie

$$-M_\chi - 4m_\chi + 3 \geq 0,$$

where M_χ, m_χ are the multiplicities of the zeros of $L_{\chi^2}(s)$ at $s = 1 + 2it$ and of $L_\chi(s)$ at $s = 1 + it$. Since M_χ and m_χ are both non-negative integers, it follows that

$$m_\chi = 0,$$

ie

$$L_\chi(1 + it) \neq 0.$$

◀

There is one important difference between the proofs of the Prime Number Theorem and Dirichlet's Theorem. In the earlier proof, we knew that $\zeta(s)$ had a simple pole at $s = 1$. But now, while we know that $L_{\chi_1}(s)$ has a simple pole at $s = 1$ we must also consider the behaviour of $L_\chi(s)$ at $s = 1$ for $\chi \neq \chi_1$.

Of course, $L_\chi(s)$ cannot have a pole at $s = 1$ if $\chi \neq \chi_1$, since we know by Proposition 4.18 that $L_\chi(s)$ is holomorphic in $\Re(s) > 0$. However, it could have a zero at $s = 1$, and this would affect the residue of $\Theta_{r,m}(s)$ at $s = 1$, and that in turn would affect the number of primes in the arithmetic sequence.

We must show that this does not in fact occur, ie

$$L_\chi(1) \neq 0$$

if $\chi \neq \chi_1$.

It turns out that there are two very different cases to consider, according as χ is real or not. For non-real characters, the result follows easily by the argument used above to show that $L_\chi(1 + it) \neq 0$. However, the real case is a much harder nut to crack.

Definition 5.4. *The multiplicative character* $\chi(n) \bmod m$ *is said to be* real *if*

$$\bar{\chi} = \chi,$$

ie

$$\chi(n) \in \mathbb{R} \text{ for all } n \in \mathbb{N}.$$

Proposition 5.6. *The character χ is real if and only if*

$$\chi(n) \in \{0, \pm 1\}$$

for all $n \in \mathbb{N}$.

Proof ▶ If $\chi(n) \in \{0, \pm 1\}$ then evidently χ is real.

Conversely, suppose χ is real. If $\chi(n) \neq 0$ then $|\chi(n)| = 1$. Hence $\chi(n) = \pm 1$, since these are the only reals on the unit circle in $\mathbb{C}$. ◀

Corollary 5.2. *Suppose χ is a multiplicative character* mod$m$. *Then*

$$\chi \text{ real} \iff \chi^2 = \chi_1.$$

Proposition 5.7. *If χ is non-real then*

$$L_\chi(1) \neq 0.$$

Proof ▶ We have in effect already proved this result, in both of the proofs of Proposition 5.5.

Thus in the first proof, taking *any* point $s = 1 + it$ (whether $L_\chi(s)$ has a zero there or not) it follows that

$$\sum_\chi \mu_\chi \leq 1.$$

In other words,

$$L_\chi(1) = 0$$

for at most one character χ.

But

$$\begin{aligned} L_\chi(1) = 0 &\implies L_\chi(\sigma) \to 0 \text{ as } \sigma \to 1 + 0 \\ &\implies \sum \chi(n) n^\sigma \to 0 \\ &\implies \sum \overline{\chi(n)} n^\sigma \to 0 \\ &\implies L_{\bar\chi}(\sigma) \to 0 \\ &\implies L_{\bar\chi}(1) = 0. \end{aligned}$$

Thus if $L_\chi(1) = 0$ and χ is non-real then

$$\sum_\chi m_\chi \geq 2,$$

which as we have seen is impossible.

As for the second proof, although we assumed that $t \neq 0$, our argument actually shows that

$$\operatorname{res}_{1+2i}(\Theta_{\chi^2}) + 4\operatorname{res}_{1+i}(\Theta_\chi) + 3 \geq 0$$

even if $t = 0$, ie

$$\operatorname{res}_1(\Theta_{\chi^2}) + 4\operatorname{res}_1(\Theta_\chi) + 3 \geq 0$$

But now if χ is not real then $\chi^2 \neq \chi_1$ and so Θ_{χ^2} does not have a pole at $s = 1$. Hence both residues are negative, and we deduce as before that $\Theta_\chi(s)$ cannot have a zero at $s = 1$. ◀

Remark. These proofs might be considered something of overkill. More simply,

$$\Theta_{1,m}(1+\sigma) = \sum_{p \equiv \bmod m} \log p \, p^{1+\sigma} \geq 0$$

for $\sigma > 0$. Hence

$$\operatorname{res}_1(\Theta_{1,m}) \geq 0,$$

ie

$$1 - \sum_\chi \mu_\chi \geq 0,$$

from which it follows that $m_\chi > 0$ for at most one χ.

Proposition 5.8. *If $\chi \neq \chi_1$ is real then*

$$L_\chi(1) \neq 0.$$

Proof ▶ Suppose χ is real; and suppose $L_\chi(1) = 0$. Consider the product

$$F(s) = \zeta(s)L_\chi(s).$$

The putative zero of $L_\chi(s)$ at $s = 1$ cancels out the pole of $\zeta(s)$, leaving a function $F(s)$ holomorphic in $\Re(s) > 0$.

The following result on the product of two Dirichlet series is readily established.

Lemma 14. *Suppose the two Dirichlet series*

$$f(s) = \sum a_n n^{-s}, \quad g(s) = \sum b_n n^{-s}$$

are absolutely convergent in $\Re(s) > \sigma$*. Then the product series*

$$f(s)g(s) = \sum c_n n^{-s},$$

where

$$c_n = \sum_{n=de} a_d b_e,$$

is also absolutely convergent in $\Re(s) > \sigma$*.*

Applying this result to $F(s) = \zeta(s)L_\chi(s)$, we see that for $\Re(s) > 1$

$$F(s) = \sum f(n)n^{-s},$$

where

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

Lemma 15. *1. $f(n)$ is multiplicative;*

2. $f(n) \geq 0$ for all n;

3. $f(n^2) > 0$.

Proof ▶ 1. In general, if $\chi(n)$ is multiplicative then so is

$$f(n) = \sum d \mid n\chi(d).$$

For suppose $n = n_1 n_2$, where $\gcd(n_1, n_2) = 1$. Then any factor $d \mid n$ splits into two coprime factors $d = d_1 d_2$, where $d_1 \mid n_1$ and $d_2 \mid n_2$. It follows that

$$\begin{aligned} f(n) &= \sum_{d|n} \chi(d) \\ &= \sum_{d_1|n_1,\, d_2|n_2} \chi(d_1 d_2) \\ &= \sum_{d_1|n_1} \chi(d_1) \sum_{d_2|n_2} \chi(d_2) \\ &= f(n_1)f(n_2). \end{aligned}$$

2. Suppose

$$n = p_1^{e_1} \cdots p_r^{e_r}.$$

Since $f(n)$ is multiplicative,

$$f(n) = f(p_1^{e_1}) \cdots f(p_r^{e_r}.$$

But

$$\begin{aligned} f(p^e) &= \chi(1) + \chi(p) + \cdots + \chi(p^e) \\ &= \chi(1) + \chi(p) + \cdots + \chi(p)^e, \end{aligned}$$

since χ is strictly multiplicative. Recall that $\chi(n) \in \{0, \pm 1\}$. It follows that

$$f(p^e) = \begin{cases} 1 & \text{if } \chi(p) = 0, \\ e + 1 & \text{if } \chi(p) = 1, \\ (-1)^e + 1 & \text{if } \chi(p) = -1s. \end{cases}$$

In particular

$$f(p^e) \ge 0$$

in all cases, and so

$$f(n) \ge 0.$$

3. Each prime factor in n^2 occurs to an even power p^{2e}. It follows from the expression for $f(p^e)$ above that

$$f(p^e) = 1, 2e + 1 \text{ or } 1$$

according as $\chi(p) = 0, 1$ or -1. In all cases,

$$f(p^{2e}) > 0,$$

and so

$$f(n^2) > 0.$$

◀

Now suppose

$$F(s) = \sum f(n)n^{-s}$$

has abscissa of convergence σ_0. Since the coefficients are non-negative this is also the abscissa of absolute convergence.

By Proposition 2.25, since $F(s)$ is holomorphic in $\Re(s) > 0$ it follows that

$$\sigma_0 \le 0.$$

This is amazing; it tells us that

$$\sum f(n)n^{-\sigma} < \infty$$

for all $\sigma > 0$.

But we know that

$$f(n^2) \geq 1.$$

These terms alone contribute

$$\begin{aligned}\sum (n^2)^{-\sigma} &= \sum n^{-2\sigma}\\ &= \zeta(2\sigma).\end{aligned}$$

But we know that $\zeta(\sigma)$ diverges if $\sigma \leq 1$. It follows that $F(s)$ diverges for $\sigma \leq 1/2$, contradicting our assertion that $\sigma_0 \leq 0$.

Thus our original assumption that $L_\chi(1) = 0$ is untenable:

$$L_\chi(1) \neq 0$$

for any real character $\chi \neq \chi_1$. ◀

5.5 Proof of Dirichlet's Theorem

We now have all the ingredients for our proof, which as we have said (many times) closely imitates that of the Prime Number Theorem.

Proof ▶ Since $L_\chi(s)$ has no zeros in $\Re(s) \geq 1$, by Propositions 5.4, 5.5, 5.7 and 5.8, it follows that if $\chi \neq \chi_1$ then

$$\frac{L'_\chi(s)}{L_\chi(s)}$$

is holomorphic in $\Re(s) \geq 1$; while on the other hand, $L_{\chi_1}(s)$ has a simple pole at $s = 1$, by Proposition 4.18, and so

$$\frac{L'_{\chi 1(s)}}{L_{\chi 1(s)}}$$

has a simple pole with residue 1 at $s = 1$, and no other poles in $\Re(s) \geq 1$.

It follows that

$$\frac{1}{\phi(m)} \sum_\chi \bar{\chi}(r) \frac{L'_\chi(s)}{L_\chi(s)}$$

has a simple pole with residue $1/\phi(m)$ at $s = 1$, and no other poles in $\Re(s) \geq 1$. The same is therefore true of $\Theta_{r,s}(s)$, by Proposition 5.3.

Thus

$$\Psi_{r,m}(s) = \Theta_{r,m}(s) - \frac{1}{\phi(m)}\zeta(s)$$

is holomorphic in $\Re(s) \geq 1$; and since

$$\begin{aligned}\Psi_{r,m}(s) &= \int_1^\infty x^{-s} d\psi_{r,m} \\ &= s\int_1^\infty x^{-s}\psi_{r,m}(x)\frac{dx}{x} \\ &= s\int_0^\infty e^{-st}\psi_{r,m}(e^t)dt,\end{aligned}$$

for $\Re(s) > 1$, we can apply our Tauberian Theorem, Proposition 3.9, with

$$F(s) = \frac{1}{s+1}\Psi_{r,m}(s+1)$$

and

$$f(x) = e^{-x}\psi_{r,m}(e^x).$$

(As we noted earlier, the condition that $f(x)$ is bounded follows at once from the fact that

$$\theta_{r,m}(x) \leq \theta(x) \leq Cx$$

for some constant C.)

We conclude that

$$\begin{aligned}\int_0^\infty e^{-t}\psi_{r,m}(e^t)dt &= \int_1^\infty \frac{\psi_{r,m}(x)}{x^2}dx \\ &= \int_1^\infty \frac{\theta_{r,m}(x) - x/\phi(m)}{x^2}dx\end{aligned}$$

converges; and from this we deduce, as before, that

$$\theta_{r,m}(x) \sim \frac{x}{\phi(m)},$$

from which Dirichlet's Theorem follows, by Corollary 5.1. ◀

Chapter 6

The gamma function

6.1 Definition

Definition 6.1. *For $\Re(s) > 0$ we set*

$$\Gamma(s) = \int_0^\infty x^s e^{-x} \frac{dx}{x}$$

The integral converges as $x \to \infty$ for all s, since $e^{-x} \to 0$ faster than any power $x^n \to \infty$. It converges at 0 for $\Re(s) > 0$ since

$$|x^{s-1} e^{-x}| \le x^{\sigma - 1}.$$

Proposition 6.1. *$\Gamma(s)$ is a holomorphic function for $\Re(s) > 0$.*

Proof ▶ The finite integral

$$\int_0^X x^s e^{-x} \frac{dx}{x}$$

is holomorphic for each $X > 0$, by one of the standard results of complex function theory.

Moreover, it is readily verified that if $\Re(s) \ge \sigma > 0$ then

$$\int_0^X x^s e^{-x} \frac{dx}{x} \to \Gamma(s)$$

uniformly as $X \to \infty$.

It follows that $\Gamma(s)$ is holomorphic. ◀

6.2 The first identity

Proposition 6.2. *For $\Re(s) > 0$,*

$$\Gamma(s+1) = s\Gamma(s).$$

Proof ▶ Integrating by parts,

$$\begin{aligned}\Gamma(s+1) &= \int_0^\infty x^s e^{-x} dx \\ &= \left[x^s \cdot -e^{-x}\right]_0^\infty + s\int_0^\infty x^{s-1}e^{-x} dx \\ &= s\Gamma(s).\end{aligned}$$

◀

Corollary 6.1. *For $n \in \mathbb{N}$,*

$$\Gamma(n+1) = n!$$

Proof ▶ For $n = 0$,

$$\begin{aligned}\Gamma(1) &= \int_0^\infty e^{-x} dx \\ &= \left[-e^{-x}\right]_0^\infty \\ &= 1.\end{aligned}$$

The result for general n follows on repeated application of the Proposition. ◀

6.3 Analytic continuation

Proposition 6.3. $\Gamma(s)$ *can be continued analytically to a meromorphic function in the whole plane, with simple poles at $s = 0, -1, -2, \dots,$ the pole at $s = -n$ having residue $(-1)^n/n!$.*

Proof ▶ By repeated application of the last Proposition,

$$\Gamma(s) = \frac{1}{s(s+1)\cdots(s+n-1)}\Gamma(s+n).$$

This holds for $\Re(s) > 0$. But the right-hand side is defined for $\Re(s) > -n$, and so extends $\Gamma(s)$ to this region.

By putting together these extensions for different n (which must coincide on their overlap by the theory of analytic continuation), we can extend $\Gamma(s)$ to the whole complex plane.

If $r < n$ then we see from the formula above that $\Gamma(s)$ has a simple pole at $s = -r$ with residue

$$\begin{aligned}\frac{1}{(-r)(-r+1)\cdots(-1)(1)(2)\cdots(-r+n-1)}\Gamma(n-r) &= (-1)^r\frac{\Gamma(n-r)}{r!(n-r-1)!}\\ &= \frac{(-1)^r}{r!}.\end{aligned}$$

◀

6.4 Analytic continuation: an alternative approach

There is an entirely different way of extending $\Gamma(s)$ to the whole plane, which has special significance for us, since we shall later apply the same method to extend $\zeta(s)$ and $L_\chi(s)$ to the whole plane.

Let us 'cut' the complex plane along the positive real axis from 0 to $+\infty$. Then we can define $\log z$ holomorphically in the cut plane by setting

$$\log(Re^{i\theta}) = \log R + i\theta \qquad (0 \le \theta \le 2\pi).$$

(The cut prevents us encircling 0 and thus passing from one branch of $\log z$ to another.) On the upper edge of the cut $\theta = 0$, and so

$$\log z = \log x$$

at $z = x > 0$. On the lower edge $\theta = 2\pi$, and so

$$\log z = \log x + 2\pi i$$

at $z = x > 0$.

Passing to

$$z^s = e^{s\log z},$$

we have

$$z^s = x^s$$

at $z = x$ on the upper edge of the cut, while

$$z^s = e^{2\pi i s}x^s$$

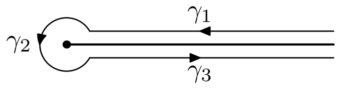

Figure 6.1: The contour $\gamma = \gamma_1 + \gamma_2 + \gamma_3$

at $z = x$ on the lower edge.

Now let us consider the integral

$$I(s) = \int_\gamma z^s e^{-z} \frac{dz}{z},$$

around the contour $\gamma = \gamma_1 + \gamma_2 + \gamma_3$ (Fig 6.1), which comes in from $+\infty$ to ϵ along the upper edge of the cut (γ_1), travels around the circle radius ϵ around 0 in the positive, or anti-clockwise, direction (γ_2) and then returns to $+\infty$ along the lower edge of the cut (γ_3).

Note that by Cauchy's Theorem $I(s)$ is *independent of* ϵ. For, writing $I_\epsilon(s)$ temporarily for $I(s)$, the difference

$$I_{\epsilon_1}(s) - I_{\epsilon_2}(s) = \int_C z^s e^{-z} \frac{dz}{z}$$

where C is the contour shown in Figure 6.2, within which the integrand is holomorphic. Hence

$$I_{\epsilon_1}(s) - I_{\epsilon_2}(s) = 0,$$

ie $I(s)$ is independent of ϵ.

(Cauchy's Theorem can be expressed in topological terms as follows. Suppose $f(z)$ is meromorphic in the open set U, with poles at $z_0, z_1, \ldots$. Let us 'puncture' U at these points, ie pass to $U' = U \setminus \{z_0, z_1, \ldots\}$. If now one contour γ in U' can be deformed into another contour γ', without passing through any poles, then

$$\int_\gamma f(z)\, dz = \int_{\gamma'} f(z)\, dz.$$

In other words,

$$\int_\gamma f(z)\, dz$$

depends only on the *homotopy class* of γ.)

Proposition 6.4. *If* $\Re(s) > 0$*,*

$$\Gamma(s) = \frac{1}{e^{2\pi i s} - 1} \int_\gamma z^s e^{-z} \frac{dz}{z}.$$

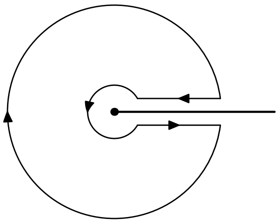

Figure 6.2: The difference $I_{\epsilon_1}(s) - I_{\epsilon_2}(s)$

Proof ▶ As $\epsilon \to 0$,

$$I_1(s) \to -\int_0^\infty x^s e^{-x} \frac{dx}{x} = -\Gamma(s).$$

Similarly,

$$I_3(s) \to e^{2\pi i s} \int_0^\infty x^s e^{-x} \frac{dx}{x} = e^{2\pi i s}\Gamma(s).$$

Also, if $\sigma = \Re(s)$,

$$\begin{aligned} |I_2(s)| &\le 2\pi\epsilon \cdot \epsilon^{\sigma-1} \\ &= 2\pi\epsilon^\sigma \\ &\to 0. \end{aligned}$$

We conclude that

$$I(s) \to (e^{2\pi i s} - 1)\Gamma(s)$$

as $\epsilon \to 0$. Since $I(s)$ is in fact independent of ϵ, it follows that

$$I(s) = (e^{2\pi i s} - 1)\Gamma(s),$$

ie

$$\Gamma(s) = \frac{1}{e^{2\pi i s} - 1} I(s)$$

for $\Re(s) > 0$. ◀

The integral $I(s)$ converges for all $s \in \mathbb{C}$, since the ‘diversion’ round 0 along γ_2 avoids the problem of convergence at $s = 0$; it therefore defines an entire function.

Proposition 6.5. *The formula*

$$\Gamma(s) = \frac{1}{e^{2\pi i s} - 1} \int_\gamma z^s e^{-z} \frac{dz}{z}.$$

extends $\Gamma(s)$*to a meromorphic function in the whole of* $\mathbb{C}$*, with simple poles at* $s = 0, -1, -2, \ldots$.

Proof ▶ Since $I(s)$ is an entire function, the only poles of $\Gamma(s)$ must arise from poles of

$$\frac{1}{e^{2\pi i s} - 1}.$$

But this function has simple poles with residue $1/2\pi i$ at each integer point $s = n \in \mathbb{Z}$. That is clear at $s = 0$, since

$$e^{2\pi i s} - 1 = 2\pi i s + O(s^2)$$

in the neighbourhood of $s = 0$; and the same result holds at $s = n$ since the function is periodic with period 1.

However, $I(s) = 0$ if $s = n > 0$, since the integrand is in fact holomorphic in the *uncut* plane. This cancels out the pole; and in any case we know that $\Gamma(n+1) = n!$.

For $n = -n \leq 0$, it is still true that the integrand is holomorphic in $\mathbb{C} \setminus \{0\}$, but now it has a pole of order $n+1$ at $s = 0$. The residue of the pole is given by the coefficient of z^n in e^{-z}. Thus

$$I(s) = \frac{2\pi i}{n!};$$

and so $\Gamma(s)$ has a simple pole at $s = -n$ with residue $1/n!$, as we saw before. ◀

6.5 Γ(s) as a limit

Euler originally defined the gamma function as a limit, in the following way.

Definition 6.2. *For* $n \in \mathbb{N}$*, we set*

$$\Gamma(s, n) = \frac{n! n^s}{s(s+1)\cdots(s+n)}.$$

Proposition 6.6. *As* $n \to \infty$*,*

$$\Gamma(s, n) \to \Gamma(s).$$

Proof ▶ Recall that

$$\left(1 - \frac{x}{n}\right)^n \to e^{-x}$$

as $n \to \infty$. This follows on taking logarithms, since

$$\begin{aligned} \log\left(1 - \frac{x}{n}\right)^n &= -n\left(\frac{x}{n} + \frac{x^2}{2n^2} + \cdots\right) \\ &= -t + O(\frac{1}{n}). \end{aligned}$$

In fact, since each term $-x, -\frac{x^2}{2n}, \ldots$ increases with n, this argument shows that $(1 - x/n)^n$ increases monotonically to e^{-x}, for each $x \geq 0$.

Let

$$f(x,n) = \begin{cases} (1 - x/n)^n & \text{if } 0 \leq x \leq n \\ 0 & \text{if } x > n \end{cases}$$

Then

$$f(x,n) \to e^{-x}$$

uniformly in any finite range $[0, X]$; and

$$0 \leq f(x,n) \leq e^{-x}$$

for all x.

It follows that if $\Re(s) > 0$ then

$$\int_0^n x^s \left(1 - \frac{1}{x}\right)^n \frac{dx}{x} = \int_0^\infty x^s f(x,n) \frac{dx}{x} \to \Gamma(s)$$

as $n \to \infty$.

But we can compute this integral by repeated integration by parts. Thus

$$\begin{aligned}
\int_0^n x^s \left(1-\frac{1}{x}\right)^n \frac{dx}{x} &= \Gamma(s,n) \\
&= \int_0^n x^{s-1}\left(1-\frac{x}{n}\right)^n dx \\
&= \left[\frac{x^s}{s}\left(1-\frac{x}{n}\right)^n\right]_0^n + \int_0^n \frac{x^s}{s}\left(1-\frac{x}{n}\right)^{n-1} dx \\
&= \frac{n}{ns}\int_0^n x^s\left(1-\frac{x}{n}\right)^{n-1} dx \\
&= \frac{n(n-1)}{n^2 s(s+1)}\int_0^n x^{s+1}\left(1-\frac{x}{n}\right)^{n-2} dx \\
&= \cdots \\
&= \frac{n(n-1)(n-2)\cdots 2}{n^{n-1}s(s+1)\cdots(s+n-2)}\int_0^n x^{s+n-2}\left(1-\frac{x}{n}\right) dx \\
&= \frac{n!}{n^n s(s+1)\cdots(s+n-1)}\int_0^n x^{s+n-1} dx \\
&= \frac{n!}{n^n s(s+1)\cdots(s+n)}\left[x^{s+n}\right]_0^n \\
&= \frac{n!}{n^n s(s+1)\cdots(s+n)} n^{s+n} \\
&= \frac{n! n^s}{s(s+1)\cdots(s+n)} \\
&= \Gamma(s,n).
\end{aligned}$$

We have therefore established that

$$\Gamma(s,n) \to \Gamma(s)$$

as $n \to \infty$, provided $\Re(s) > 0$. We can extend the result to all s (except $s = 0, -1, -2, \dots$) by noting that

$$\begin{aligned}
\Gamma(s+r,n) &= \frac{n^{r+s}n!}{(s+r)(s+r+1)\cdots(s+r+n)} \\
&= n^r \frac{s(s+1)\cdots(s+n)}{(s+r)(s+r+1)\cdots(s+r+n)}\Gamma(s,n).
\end{aligned}$$

Thus if $n \geq r$,

$$\Gamma(s,n) = \frac{\Gamma(s+r,n)}{s(s+1)\cdots(s+r-1)}\frac{n^r}{(s+n+1)\cdots(s+n+r)}.$$

Now suppose $\Re(s) > -r$. From above,

$$\Gamma(s+r, n) \to \Gamma(s+r).$$

Moreover,

$$\frac{n^r}{(s+n+1)\cdots(s+n+r)} = \frac{1}{(1+\frac{s+1}{n})\cdots(1+\frac{s+r}{n}} \to 1$$

as $n \to \infty$. It follows that

$$\Gamma(s,n) \to \frac{\Gamma(s+r)}{s(s+1)\cdots(s+r-1)} = \Gamma(s).$$

We have thus extended the result to $\Re(s) > -r$, and so to the whole plane (excluding the poles $s = 0, -1, -2, \dots$). ◀

We can re-write $\Gamma(s,n)$ as

$$\Gamma(s,n) = \frac{n^s}{s}\frac{1}{(1+s)(1+\frac{s}{2})\cdots(1+\frac{s}{n})}.$$

Thus

$$s\Gamma(s,n) = n^s \prod_{1 \le m \le n}\left(1+\frac{s}{m}\right)^{-1}.$$

We can also re-write n as

$$\begin{aligned} n &= \frac{2}{1}\frac{3}{2}\cdots\frac{n}{n-1} \\ &= \prod_{1 \le m \le (n-1)}\left(1+\frac{1}{m}\right). \end{aligned}$$

Thus

$$n^s = \prod_{1 \le m \le (n-1)}\left(1+\frac{1}{m}\right)^s.$$

Hence

$$s\Gamma(s,n) = \left(1+\frac{1}{n}\right)^{-s}\prod_{1 \le m \le n}\left\{\left(1+\frac{s}{m}\right)^{-1}\left(1+\frac{1}{m}\right)^s\right\}.$$

Since $(1+\frac{1}{n})^s \to 1$, it follows that

$$\prod_{1 \le m \le n}\left\{\left(1+\frac{s}{m}\right)^{-1}\left(1+\frac{1}{m}\right)^s\right\} \to s\Gamma(s).$$

In other words, $\Gamma(s)$ can be expressed as the infinite product

$$\Gamma(s) = \frac{1}{s} \prod_{m \geq 1} (1 + a_m),$$

where

$$1 + a_m = \left(1 + \frac{s}{m}\right)^{-1} \left(1 + \frac{1}{m}\right)^s.$$

This infinite product converges absolutely, since

$$\begin{aligned} 1 + a_m &= \left(1 + \frac{s}{m}\right)^{-1} \left(1 + \frac{1}{m}\right)^s \\ &= \left(1 - \frac{s}{m} + \frac{s^2}{m^2} + O(\frac{1}{m^3})\right) \left(1 + \frac{s}{m} + \frac{s(s-1)}{2m^2} + O(\frac{1}{m^3}\right) \\ &= 1 - \frac{s(s-1)}{2m^2} + O(\frac{1}{m^3}), \end{aligned}$$

and we know of course that $\sum m^{-2}$ converges.

Since the series $\sum |a_m|$ is uniformly convergent in any compact (ie closed and bounded) subset C not containing any of the poles, the function defined by the infinite product is holomorphic in C. This gives a third way of extending $\Gamma(s)$ holomorphically to the entire plane.

6.6 The second identity

Proposition 6.7. *For all $s \in \mathbb{C} \setminus \mathbb{Z}$,*

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}.$$

Proof ▶ We have

$$\begin{aligned} \Gamma(s,n)\Gamma(1-s,n) &= \frac{n^s}{s(1+s)(1+\frac{s}{2})\cdots(1+\frac{s}{n})} \frac{n^{1-s}}{(1-s)(1-\frac{s}{2})\cdots(1-\frac{s}{n})} \frac{1}{1-s+n} \\ &= \frac{1}{s} \prod_{1 \leq m \leq n} \left(1 - \frac{s^2}{m^2}\right)^{-1} \frac{n}{1-s+n} \end{aligned}$$

But we saw in Chapter 1 that

$$\sin \pi s = \pi s \prod \left(1 - \frac{s^2}{m^2}\right).$$

It follows that

$$\Gamma(s,n)\Gamma(1-s,n) \to \frac{\pi}{\sin \pi s},$$

from which the result follows. ◀

We shall give another proof of this result below.

Proposition 6.8. $\Gamma(1/2) = \sqrt{\pi}$.

Proof ▶ Setting $s = 1/2$ in the identity above,

$$\begin{aligned}\Gamma(1/2)^2 &= \frac{\pi}{\sin\frac{\pi}{2}}\\ &= \pi.\end{aligned}$$

Thus

$$\Gamma(1/2) = \pm\sqrt{\pi}.$$

Since

$$\Gamma(1/2) = \int_0^\infty x^{1/2} e^{-x} \frac{dx}{x} > 0,$$

it follows that

$$\Gamma(1/2) = \sqrt{\pi}.$$

◀

Corollary 6.2. *For each $n \in \mathbb{N}$,*

$$\Gamma(n + \frac{1}{2}) = \frac{1}{2}\frac{3}{2}\cdots(n - \frac{1}{2})\frac{1}{\sqrt{\pi}}$$

6.7 The third identity

We can write $(2n)!$ as

$$\begin{aligned}(2n)! &= (1\cdot 3\cdot 5\cdots(2n-1))\,(2\cdot 4\cdot 6\cdots(2n))\\ &= 2^{2n}\left(\frac{1}{2}\frac{3}{2}\cdots(n-\frac{1}{2})\right)n!\\ &= 2^{2n}\frac{\Gamma(n+\frac{1}{2})}{\Gamma(\frac{1}{2})}n!.\end{aligned}$$

Dividing each side by $2n$,

$$\Gamma(2n) = 2^{2n-1}\frac{\Gamma(n+\frac{1}{2})\Gamma(n)}{\Gamma(\frac{1}{2})}$$

ie

$$\Gamma(n)\Gamma(n+\tfrac{1}{2}) = 2^{1-2n}\sqrt{\pi}\Gamma(2n).$$

This strongly suggests — but does not establish— the following result.

Proposition 6.9. *For all s,*

$$\Gamma(s)\Gamma(s+\tfrac{1}{2}) = 2^{1-2s}\sqrt{\pi}\Gamma(2s).$$

Proof ▶ We have

$$\begin{aligned}\Gamma(s,n)\Gamma(s+\tfrac{1}{2}) &= \frac{n^{2s+\frac{1}{2}}(n!)^2}{s(s+\frac{1}{2})(s+1)(s+\frac{3}{2})\cdots(s+n)(s+n+\frac{1}{2})}\\ &= \frac{2^{2n+2}n^{2s+\frac{1}{2}}(n!)^2}{2s(2s+1)\cdots(2s+2n)(2s+2n+1)}\end{aligned}$$

while

$$\begin{aligned}\Gamma(2s,2n) &= \frac{(2n)^{2s}(2n)!}{2s(2s+1)\cdots(2s+2n)}\\ &= \frac{2^{2s}n^{2s}(2n)!}{2s(2s+1)\cdots(2s+2n)}.\end{aligned}$$

Thus

$$\begin{aligned}\frac{2^{2s}\Gamma(s)\Gamma(s+\frac{1}{2})}{\Gamma(2s)} &= \frac{2^{2n+2}n^{\frac{1}{2}}(n!)^2}{(2n)!}\frac{1}{2s+2n+1}\\ &= \frac{2^{2n}n^{\frac{1}{2}}(n-1)!^2}{(2n-1)!}\frac{2n}{2s+2n+1}\\ &= \frac{2^{2n}n^{\frac{1}{2}}\Gamma(n)^2}{\Gamma(2n)}\frac{2n}{2s+2n+1}.\end{aligned}$$

Note that the right-hand side is independent of s, except for the factor $2n/(2s+2n+1)$, which tends to 1 and can thus be ignored. We have to show that the right-hand side $\to \sqrt{\pi}$ as $n \to \infty$, ie

$$\frac{2^{2n}n^{\frac{1}{2}}\Gamma(n)^2}{\Gamma(2n)} \to \sqrt{\pi}.$$

It follows that if the result holds for one s then it will hold for all s. But we saw in the introduction to the Proposition that the result holds for positive integers $s = m > 0$. We conclude that it holds for all s. ◀

6.8 The Eulerian integral

Definition 6.3. *For* $\Re(u) > 0$, $\Re(v) > 0$, *we set*

$$B(u,v) = \int_0^1 t^{u-1}(1-t)^{v-1}dt.$$

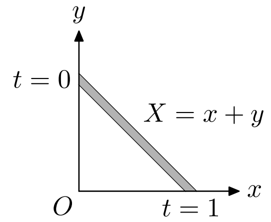

Figure 6.3: The double integral for $\Gamma(u)\Gamma(v)$

The integral converges at 0 if $\Re(u) > 0$; it converges at 1 if $\Re(v) > 0$. Setting $t = \sin^2\theta$, the integral can be written in the form

$$B(u, v) = 2\int_0^{\pi/2} \sin^{2u}\theta \cos^{2v}\theta \, d\theta.$$

$B(u, v)$ is often called the *Eulerian integral of the first kind*; the integral by which we defined $\Gamma(s)$ being the Eulerian integral of the second kind.

Proposition 6.10. *For* $\Re(u) > 0,\ \Re(v) > 0$,

$$B(u, v) = \frac{\Gamma(u)\Gamma(v)}{\Gamma(u + v)}.$$

Proof ▶ We compute $\Gamma(u)\Gamma(v)$ as a double integral:

$$\begin{aligned}\Gamma(u)\Gamma(v) &= \int_0^\infty x^{u-1}e^{-x}dx \int_0^\infty y^{v-1}e^{-y}dy \\ &= \int\int x^{u-1}y^{v-1}e^{-(x+y)}dx\,dy,\end{aligned}$$

where the double integral extends over the first quadrant.

Now let us change variables to

$$X = x + y,\ t = \frac{x}{x + y}.$$

Inversely,

$$x = Xt,\ y = X(1 - t).$$

The Jacobian is

$$\frac{\partial(x, y)}{\partial(X, t)} = \det\begin{pmatrix} t & 1 - t \\ X & -t \end{pmatrix} = -X.$$

Thus the integral becomes

$$\begin{aligned}\Gamma(u)\Gamma(v) &= \int_0^\infty \int_0^1 X^{u+v-2} t^{u-1}(1-t)^{v-1} e^{-X} X dX\, dt\\ &= \int_0^\infty X^{u+v-1} e^{-X} dX \int_0^1 t^{u-1}(1-t)^{v-1} dt\\ &= \Gamma(u+v)B(u,v),\end{aligned}$$

as required. ◀

This provides an alternative proof of our second identity

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}.$$

For suppose $0 < \Re(s) < 1$. Then

$$\begin{aligned}\Gamma(s)\Gamma(1-s) &= \Gamma(1)B(s, 1-s)\\ &= \int_0^1 t^s (1-t)^{-s} \frac{dt}{t}\\ &= \int_0^1 \left(\frac{t}{1-t}\right)^s \frac{dt}{x}.\end{aligned}$$

Let

$$u = \frac{t}{1-t}.$$

As t increases from 0 to 1, u increases from 0 to ∞. Also

$$t = \frac{u}{u+1} = 1 - \frac{1}{u+1},$$

and so

$$\begin{aligned}\frac{dt}{t} &= \frac{u+1}{u} \frac{du}{(u+1)^2}\\ &= \frac{du}{u(u+1)}.\end{aligned}$$

Thus

$$\Gamma(s)\Gamma(1-s) = \int_0^\infty \frac{u^s}{u(u+1)} du$$

Now we can play the same 'trick' that we used to continue $\Gamma(s)$ analytically:

$$\Gamma(s)\Gamma(1-s) = \frac{1}{e^{2\pi i s} - 1} \int_\gamma \frac{z^s}{z(z+1)} dz,$$

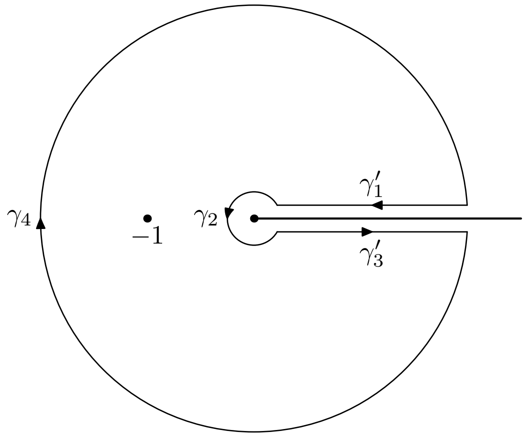

Figure 6.4: The contour $\gamma = \gamma_1 + \gamma_2 + \gamma_3$

where γ is the contour shown in Fig 6.1, with the proviso now that $\epsilon < 1$, to avoid the pole at $s = -1$.

But now let us 'complete' the contour with a large circle radius R (Fig 6.4). Let

$$I'(s) = \int_{\gamma'} \frac{z^s}{z(z+1)} dz,$$

where $\gamma' = \gamma_1' + \gamma_2 + \gamma_3' + \gamma_4$, with corresponding definitions of $I_1'(s), I_2(s), I_3'(s), I_4(s)$. As $R \to \infty$,

$$I_1'(s) \to I_1(s), \quad I_3'(s) \to I_3(s),$$

Also

$$|I_4(s)| \le 2\pi R \frac{R^\sigma}{R(R-1)};$$

and so

$$I_4(s) \to 0$$

as $R \to \infty$.

In fact $I'(s)$ is independent of R (provided $R > 1$) by the same argument that showed $I(s)$ was independent of ϵ. Hence

$$I'(s) = I(s).$$

But now we can compute $I'(s)$ by Cauchy's Theorem. Since we are going round γ' in the 'wrong way' (clockwise),

$$\begin{aligned} I'(s) &= -2\pi i \operatorname{res}_{-1}(\frac{z^s}{z(z+1)} \\ &= -2\pi i \frac{(-1)^s}{-1} \\ &= 2\pi i e^{\pi i s}. \end{aligned}$$

We conclude that

$$\begin{aligned} \Gamma(s)\Gamma(1-s) &= \frac{2\pi i e^{\pi i s}}{e^{2\pi i s} - 1} \\ &= \frac{2i}{e^{\pi i s} - e^{-\pi i s}} \pi \\ &= \frac{\pi}{\sin \pi s}, \end{aligned}$$

since $\sin z = (e^{iz} - e^{-iz})/2i$.

Chapter 7

The functional equation for $\zeta(s)$

7.1 Analytic continuation of $\zeta(s)$

Proposition 7.1. *For* $\Re(s) > 0$,

$$\Gamma(s)\zeta(s) = \int_0^\infty \frac{x^s}{e^x - 1}\frac{dx}{x}.$$

Proof ▶ The rôle of the gamma function in the theory of $\zeta(s)$ stems from the following result.

Lemma 16. *If* $\Re(s) > 0$,

$$\int_0^\infty x^s e^{-nx}\frac{dx}{x} = n^{-s}\Gamma(s).$$

Proof ▶ On making the change of variable $y = nx$,

$$\begin{aligned}\int_0^\infty x^s e^{-nx}\frac{dx}{x} &= \int_0^\infty n^{-s}y^s e^{-y}\frac{dy}{y}\\ &= n^{-s}\Gamma(s).\end{aligned}$$

◀

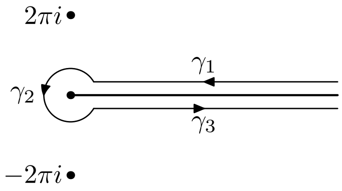

Figure 7.1: The contour γ

Summing this result for $n = 1, 2, 3, \ldots,$

$$\begin{aligned}
\zeta(s)\Gamma(s) &= \sum_{n=1}^{\infty} n^{-s}\Gamma(s) \\
&= \sum_{n=1}^{\infty} \int_0^{\infty} x^s e^{-nx} \frac{dx}{x} \\
&= \int_0^{\infty} x^s \sum_{n=1}^{\infty} e^{-nx} \frac{dx}{x} \\
&= \int_0^{\infty} x^s \frac{e^{-x}}{1-e^{-x}} \frac{dx}{x} \\
&= \int_0^{\infty} \frac{x^s}{e^x - 1} \frac{dx}{x},
\end{aligned}$$

the interchange of sum and integral being justified by the absolute convergence of the two together. ◀

Now we can play the same ‘trick’ that we used to analytically continue the gamma function, integrating around the contour $\gamma = \gamma_1 + \gamma_2 + \gamma_3$ in the cut plane introduced in Proposition 6.1, with the added proviso in this case that we must take the radius of the small circle $\epsilon < 2\pi$, to avoid the poles of $1/(e^z - 1)$ at $\pm 2\pi i$ (Fig 7.1).

Proposition 7.2. *The Riemann zeta function $\zeta(s)$ can be analytically continued to the whole complex plane $\mathbb{C}$ through the formula*

$$\Gamma(s)\zeta(s) = \frac{1}{e^{2\pi i s} - 1} \int_\gamma \frac{z^s}{e^z - 1} \frac{dz}{z}.$$

Proof ▶ Let

$$I(s) = \int_\gamma \frac{z^s}{e^z - 1} \frac{dz}{z},$$

so that

$$I(s) = I_1(s) + I_2(s) + I_3(s),$$

where $I_1(s), I_2(s), I_3(s)$ denote the corresponding integrals along $\gamma_1, \gamma_2, \gamma_3$. As in Section 6.3, $I(s)$ is independent of ϵ, by Cauchy's Theorem. And as there,

$$z^s = x^s = e^{s \log x}$$

at $z = x$ on the upper edge of the cut, while

$$z^s = e^{s(\log x + 2\pi i)} = e^{2\pi i s} x^s$$

at $z = x$ on the lower edge of the cut. Thus

$$\begin{aligned} I_1(s) + I_3(s) &= (e^{2\pi i s} - 1) \int_\epsilon^\infty \frac{x^s}{e^x - 1} \frac{dx}{x} \\ &\to (e^{2\pi i s} - 1)\zeta(s) \text{ as } \epsilon \to 0, \end{aligned}$$

by Proposition 7.1.

On the other hand, the function

$$f(z) = \frac{z}{e^z - 1}$$

is holomorphic, and so bounded, in $|z| \le \pi$, say

$$|f(z)| \le C,$$

ie

$$|\frac{1}{e^z - 1}| \le C|z|^{-1}.$$

Hence

$$|I_2(s)| \le 2\pi C \epsilon^{\sigma - 1}.$$

Thus if $\Re(s) > 1$ then

$$I_2(s) \to 0 \text{ as } \epsilon \to 0.$$

Since $I(s)$ is independent of ϵ, it follows that

$$I(s) = (e^{2\pi i s} - 1)\zeta(s),$$

ie

$$\zeta(s) = \frac{1}{e^{2\pi i s} - 1} I(s).$$

◀

The following alternative form of this result is often more convenient.

Corollary 7.1. *For all s,*

$$\zeta(s) = \frac{\Gamma(1-s)}{2\pi i} e^{-\pi i s} \int_\gamma \frac{z^s}{e^z - 1} \frac{dz}{z}.$$

Proof ▶ By Proposition 6.7,

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}.$$

But

$$\begin{aligned} e^{2\pi i s - 1} &= e^{\pi i s}\left(e^{\pi i s} - e^{-\pi i s}\right) \\ &= 2ie^{\pi i s} \sin \pi s. \end{aligned}$$

Thus

$$\begin{aligned} \zeta(s) &= \frac{1}{\Gamma(s)(e^{2\pi i s} - 1)} I(s) \\ &= \frac{\Gamma(1-s)}{2\pi i} e^{-\pi i s} I(s). \end{aligned}$$

◀

Proposition 7.3. *The only pole of $\zeta(s)$ in the entire complex plane $\mathbb{C}$ is the simple pole (with residue 1) at $s = 1$.*

Proof ▶ $The function \Gamma(1-s)$ has poles at $s = 1, 2, 3, \ldots$, since $\Gamma(s)$ has poles at $s = 0, -1, -2, \ldots$. On the other hand, the function $I(s)$ is entire, as is $e^{-\pi i s}$.

It follows from the Corollary to the last Proposition that $\zeta(s)$ can only have poles at $s = 1, 2, 3, \ldots$. But we know that $\zeta(s)$ is holomorphic for $\Re(s) > 1$. Thus the only possible pole is at $s = 1$, and we already know that there is a simple pole there with residue 1. ◀

7.2 The functional equation

Proposition 7.4. *The Riemann zeta function $\zeta(s)$ satisfies the functional equation*

$$\zeta(1-s) = 2\cos\tfrac{\pi s}{2}(2\pi)^{-s}\Gamma(s)\zeta(s).$$

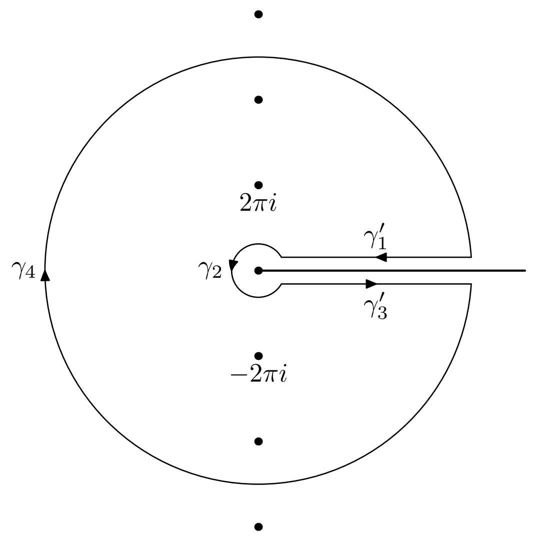

Figure 7.2: The contour γ'

Proof ▶ Suppose $\sigma = \Re(s) < 0$. Let

$$F(z) = \frac{z^s}{e^z - 1}\frac{1}{z};$$

and let

$$I'(s) = \int_{\gamma'} F(z)dz$$

around the *clockwise* contour

$$\gamma' = \gamma_1' + \gamma_2 + \gamma_3' + \gamma_4$$

(Fig 7.2), where γ_1' runs from R to ϵ along the upper edge of the cut, γ_2 is a small circle radius ϵ as before, γ_3' runs from ϵ to R along the lower edge of the cut, and γ_4 is the circle radius $R = (2n+1)\pi$ considered above. Let us denote the corresponding integrals along these contours by $I_1'(s), I_2(s), I_3'(s), I_4(s)$, so that

$$I'(s) = I_1'(s) + I_2(s) + I_3'(s) + I_4(s).$$

To avoid the poles of $1/(e^z - 1)$ at $z = 2n\pi i$ let us take

$$R = (2n+1)\pi,$$

so that the circle γ_4 passes mid-way between two poles at the top, and similarly at the bottom.

As $n \to \infty$ (and so $R \to \infty$),

$$I_1'(s) \to I_1(s),\ I_3'(s) \to I_3(s),$$

On the other hand, we shall show that, since $\Re(s) < 0$,

$$I_4(s) \to 0$$

as $n \to \infty$. It will follow that

$$I'(s) \to I(s).$$

The function

$$f(z) = \frac{1}{e^z - 1}$$

has poles at

$$z = 2n\pi i \quad (n \in \mathbb{Z}).$$

The following Lemma shows that provided we keep a reasonable distance away from the poles, the function $f(z)$ will remain reasonably small.

Lemma 17. *There is a constant C such that*

$$\frac{1}{|e^z - 1|} \le C.$$

provided

$$|z - 2n\pi i| \ge 1$$

for all $n \in \mathbb{Z}$.

Proof ▶ Since $f(z) = 1/(e^z - 1)$ is periodic with period $2\pi i$, it is sufficient to consider its value in the strip

$$S = \{z = x + iy : -\pi \le y \le \pi\}$$

outside the disk

$$D = \{z : |z| \le 1\}$$

(Fig 7.3).

The function

$$g(z) = zf(z) = \frac{z}{e^z - 1}$$

is holomorphic in S, and is therefore bounded in any finite part of this strip, say

$$|g(z)| \le c$$

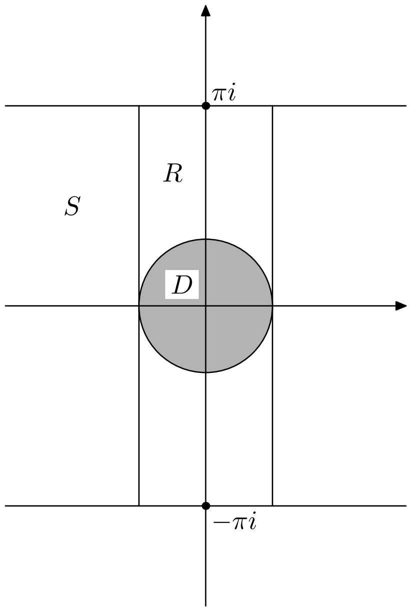

Figure 7.3: Determining sup$|1/(e^z - 1)|$

in

$$R = \{z : -1 \le \Re(z) \le 1\}$$

(Fig 7.3). Thus

$$|f(z)| \le c$$

in $R \setminus D$ (since $|z| \ge 1$ outside D).

On the other hand, if $\Re(z) \ge 1$ then

$$|e^z - 1| \ge |e^z| - 1 \ge e - 1;$$

while if $\Re(z) \le -1$ then

$$|e^z - 1| \ge 1 - |e^z| \ge 1 - e^{-1}.$$

It follows that

$$\frac{1}{|e^z - 1|} \le C = \max(c, 1/(1 - e^{-1}).$$

◀

By the Lemma,

$$\frac{1}{|e^z - 1|} \le C$$

on the large circle γ_4; while

$$|z^s| = R^\sigma$$

on this circle. Hence

$$\begin{aligned} |I_4(s)| &\le 2\pi C R^\sigma \\ &\to 0 \text{ as } n \to \infty, \end{aligned}$$

since $\sigma = \Re(s) < 0$.

It follows that

$$I'(s) \to I(s) = \left(e^{2\pi i s} - 1\right) \Gamma(s)\zeta(s)$$

as $n \to \infty$.

But now. since the contour γ' is closed, we can compute the integral $I'(s)$ by Cauchy's Theorem, from the residues of $F(z)$ at its poles within the contour. Since the contour runs in the 'wrong' direction (clockwise rather than anti-clockwise), we must negate the sum. Thus

$$I'(s) = -2\pi i \sum_{0<m\le n} \left(\mathrm{res}_{2m\pi i}(F) + \mathrm{res}_{-2m\pi i}(F)\right).$$

In the neighbourhood of $z = 0$,

$$f(z) = \frac{1}{e^z - 1} = \frac{1}{z + z^2/2 + \cdots} = \frac{1}{z} + h(z),$$

where $h(z)$ is holomorphic at $z = 0$. It follows that $f(z)$ has a simple pole with residue 1 at $z = 0$. Therefore, since $f(z)$ is periodic with period $2\pi i$, it has a simple pole with residue 1 at $z = 2n\pi i$ for each $n \in \mathbb{Z}$. Thus

$$\mathrm{res}_{2n\pi i}(F) = \frac{(2n\pi i)^s}{2n\pi i}, \quad \mathrm{res}_{-2n\pi i}(F) = \frac{(-2n\pi i)^s}{-2n\pi i}.$$

We must take care to compute the powers correctly. Recall that if

$$z = re^{i\theta} \quad (0 \le \theta \le 2\pi)$$

then we must take

$$z^s = r^s e^{i\theta s}.$$

Thus

$$z = 2n\pi i = 2n\pi e^{i\pi/2} \implies z^s = (2n\pi)^s e^{\pi i s/2},$$

while

$$z = -2n\pi i = 2n\pi e^{3\pi i/2} \implies z^s = (2n\pi)^s e^{3\pi i s/2}.$$

It follows that

$$\begin{aligned} I'(s) &= -2\pi i \sum_{0<m\le n} (2n\pi)^s \left(\frac{e^{\pi i/2}}{2n\pi i} + \frac{e^{3\pi i/2}}{-2n\pi i} \right) \\ &= (2\pi)^s \sum_{0<m\le n} n^{s-1} \left(e^{3\pi i/2} - e^{\pi i/2} \right). \end{aligned}$$

Since

$$\sum n^{s-1} = \zeta(1-s),$$

we conclude that

$$\begin{aligned} \Gamma(s)\zeta(s) &= \frac{1}{e^{2\pi is}-1} I(s) \\ &= \frac{1}{e^{2\pi is}-1} \lim_{n\to\infty} I'(s) \\ &= \frac{e^{3\pi is/2} - e^{\pi is/2}}{e^{2\pi is}-1} (2\pi)^s \zeta(1-s) \\ &= \frac{e^{\pi is/2} - e^{-\pi is/2}}{e^{\pi is} - e^{-\pi is}} (2\pi)^s \zeta(1-s) \\ &= \frac{1}{e^{\pi is/2} + e^{-\pi is/2}} (2\pi)^s \zeta(1-s) \\ &= \frac{1}{2\cos(\pi s/2)} (2\pi)^s \zeta(1-s), \end{aligned}$$

ie

$$2\cos(\pi s/2)\Gamma(s)\zeta(s) = (2\pi)^s \zeta(1-s).$$

All this was on the assumption that $\Re(s) < 0$. But now it follows by analytic continuation that the result holds for all $s \in \mathbb{C}$. ◀

The functional equation can be re-written in various ways, using the properties of $\Gamma(s)$ established in Chapter 6. In particular we can express it in a form invariant under the transformation $s \to 1-s$. (In geometric terms, this transformation describes reflection in the point $s = 1/2$.)

Proposition 7.5. *Let*

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

Then $\xi(s)$ is an entire function; and

$$\xi(1-s) = \xi(s).$$

Proof ▶ The function $\Gamma(s/2)$ has poles at $s = 0, -2, -4, \ldots$, while $\zeta(s)$ has zeros at $s = -2, -4, \ldots$. This leaves a pole at $s = 0$ which is cancelled by the zero of the factor s, In addition, the pole of $\zeta(s)$ at $s = 1$ is cancelled by the zero of the factor $s - 1$. Thus all possible poles of $\xi(s)$ are accounted for, and this function must be entire.

By the second gamma function identity (Proposition 6.7), with $(1-s)/2$ in place of s,

$$\begin{aligned}\Gamma\left(\tfrac{1-s}{2}\right)\Gamma\left(\tfrac{1+s}{2}\right) &= \frac{\pi}{\sin\frac{\pi(1-s)}{2}}\\ &= \frac{\pi}{\cos\frac{\pi s}{2}},\end{aligned}$$

since $\sin(\pi/2 - \tau) = \cos\tau$.

By the third gamma function identity (Proposition 6.9), with $s/2$ in place of s,

$$\Gamma\left(\tfrac{s}{2}\right)\Gamma\left(\tfrac{1+s}{2}\right) = 2^{1-s}\pi^{\frac{1}{2}}\Gamma(s).$$

Dividing one relation by the other,

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

But the functional equation can be written

$$\begin{aligned}\frac{\zeta(1-s)}{\zeta(s)} &= 2^{1-s}\cos\frac{\pi s}{2}\Gamma(s)\pi^{-s}\\ &= \frac{\Gamma(\frac{s}{2})}{\Gamma(\frac{1-s}{2})}\pi^{\frac{1}{2}-s}.\end{aligned}$$

Thus if we set

$$\eta(s) = \Gamma(\frac{s}{2})\zeta(s)$$

then

$$\frac{\eta(1-s)}{\eta(s)} = \pi^{\frac{1}{2}-s}.$$

But now if we set

$$\theta(s) = \pi^{\frac{s}{2}}$$

then

$$\frac{\theta(1-s)}{\theta(s)} = \pi^{\frac{1}{2}-s}.$$

Hence

$$\frac{\eta(1-s)}{\eta(s)} = \frac{\theta(1-s)}{\theta(s)}$$

ie

$$\beta(1-s) = \beta(s),$$

where

$$\begin{aligned}\beta(s) &= \eta(s)\theta(s)\\ &= \pi^{-\frac{s}{2}}\Gamma\left(\tfrac{s}{2}\right)\zeta(s).\end{aligned}$$

We conclude that, since the function $s(s-1)$ is invariant under $s \mapsto 1-s$ (we include it to remove the pole at $s=1$),

$$\begin{aligned}\xi(s) &= s(s-1)\beta(s)\\ &= s(s-1)\pi^{-\frac{s}{2}}\Gamma\left(\tfrac{s}{2}\right)\zeta(s)\end{aligned}$$

satisfies

$$\xi(1-s) = \xi(s).$$

◀

7.3 The behaviour of $\zeta(s)$ for $\Re(s) \le 0$

The functional equation allows us to determine how $\zeta(s)$ behaves 'on the far side' of the critical strip $0 \le \Re(s) \le 1$; for the map

$$s \mapsto 1-2$$

sends the left-hand half-plane $\Re(s) < 0$ into the half-plane $\Re(s) > 1$, where $\zeta(s)$ is well-behaved.

We already know that $\zeta(s)$ has no poles in $\Re(s) \le 0$, by Proposition 7.3. It does however have zeros, as we shall see.

Proposition 7.6. *The Riemann zeta function $\zeta(s)$ has simple zeros at $s = -2, -4, -6, \ldots$. It has no other zeros (or poles) in $\Re(s) \le 0$.*

Proof ▶ Since $\pi^{-s/2}$ and $\Gamma(s/2)$ have no zeros anywhere, it follows that any zero of

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

must be a zero of $\zeta(s)$, except possibly for $s = 0, 1$.

At $s = 1$, $\zeta(s)$ has a simple pole which is cancelled out by the zero of $s - 1$. Thus $\xi(1) \neq 0$; and since $\xi(0) = \xi(1)$ by the functional equation $\xi(1-s) = \xi(s)$, it follows that $\xi(0) \neq 0$. Thus

$$\xi(s) = 0 \implies \zeta(s) = 0.$$

Now we know that $\zeta(s)$ has no zeros in $\Re(s) \geq 1$ by Propositions 3.4 and 3.8. Hence $\xi(s)$ has no zeros in $\Re(s) \geq 1$. Thus $\xi(s)$ has no zeros in $\Re(s) \leq 0$, since $\xi(1-s) = \xi(s)$.

It follows that $\zeta(s)$ has zeros in $\Re(s) \leq 0$ just at those points where $s(s-1)\Gamma(s/2)$ has poles. Now $\Gamma(s/2)$ has simple poles at $s = 0, -2, -4, \ldots$; but the pole at $s = 0$ is cancelled by the zero of s at this point. We conclude that $\zeta(s)$ has simple zeros at $s = -2, -4, -6, \ldots$, and that these are the only zeros of $\zeta(s)$ in $\Re(s) \leq 0$. ◀

7.4 The values of $\zeta(2n)$

The functional equation allows us to express $\zeta(2n)$ in terms of $\zeta(1-2n)$. Although at first sight this might seem a step backwards, it turns out that the latter can be determined with relative ease, using Cauchy's Residue Theorem.

Interestingly, the argument only works for even values; there seem to be no simple expressions for

$$\zeta(3), \zeta(5), \zeta(7), \ldots.$$

7.4.1 The Bernouilli numbers

Our formulae for $\zeta(2n)$ involve the *Bernouilli numbers*, rational numbers which occur in many mathematical formulae.

Definition 7.1. *The Bernouilli numbers $B_n (n \in \mathbb{N})$ are defined by*

$$\frac{z}{e^z - 1} = \sum_{n \in \mathbb{N}} B_i \frac{z^i}{i!}.$$

Remarks. 1. Different authors use slightly different notations for the Bernouilli numbers. As we shall see, the odd Bernouilli numbers all vanish after the first. What we call B_{2n} is sometimes denoted by B_n.

Again, it will follow from our formulae for $\zeta(2n)$ that $B_2, B_4, B_6, \dots$ are alternatively positive and negative. Sometimes B_n is used to denote the absolute value, so that $B_n \geq 0$ for all n.

However, we shall stick with the definition above.

2. We can compute the Bernouilli numbers recursively from the identity

$$\left(1 + \tfrac{1}{2}z + \tfrac{1}{6}z^2 + \tfrac{1}{24}z^3 + \tfrac{1}{120}z^4 + \cdots\right)\left(B_0 + B_1 z + \tfrac{1}{2}B_2 z^2 + \tfrac{1}{6}B_3 z^3 + \tfrac{1}{24}B_4 z^4 + \cdots\right) = 1.$$

Comparing constant terms,

$$B_0 = 1.$$

Comparing coefficients of z, z^2, z^3, z^4,

$$\begin{aligned}
B_1 + \tfrac{1}{2}B_0 = 0 &\implies B_1 = -\tfrac{1}{2},\\
\tfrac{1}{2}B_2 + \tfrac{1}{2}B_1 + \tfrac{1}{6}B_0 = 0 &\implies B_2 = -\tfrac{1}{6},\\
\tfrac{1}{6}B_3 + \tfrac{1}{4}B_2 + \tfrac{1}{6}B_1 + \tfrac{1}{24}B_0 = 0 &\implies B_3 = 0,\\
\tfrac{1}{24}B_4 + \tfrac{1}{12}B_3 + \tfrac{1}{12}B_2 + \tfrac{1}{24}B_1 + \tfrac{1}{24}B_0 = 0 &\implies B_4 = -\tfrac{1}{30}.
\end{aligned}$$

Proposition 7.7. *The odd Bernouilli numbers after B_1 all vanish:*

$$B_{2n+1} = 0 \qquad (n = 1, 2, 3, \dots).$$

Proof ▶ Let

$$f(z) = \frac{z}{e^z - 1}.$$

Then

$$\begin{aligned}
f(-z) &= \frac{-z}{e^{-z} - 1}\\
&= \frac{ze^z}{e^z - 1}.
\end{aligned}$$

Thus

$$f(z) - f(-z) = -z.$$

On the other hand,

$$f(z) - f(-z) = 2\sum_{n \text{ odd}} B_n \frac{z^n}{n!}.$$

It follows that

$$B_1 = -\frac{1}{2},\ B_3 = B_5 = \cdots = 0.$$

◀

7.4.2 Determining $\zeta(1-2n)$

Proposition 7.8. *For $n = 1, 2, 3, \ldots,$*

$$\zeta(1-2n) = -\frac{B_{2n}}{2n}.$$

Proof ▶ By the Corollary to Proposition 7.2, setting $s = 1 - 2n$,

$$\zeta(1-2n) = \frac{\Gamma(2n)}{2\pi i} e^{-\pi i(1-2n)} \int_\gamma \frac{z^{-2n}}{e^z - 1} dz.$$

Now the function

$$F(z) = \frac{z^{-2n}}{e^z - 1}$$

is meromorphic in the complex plane. In particular, the values of $F(z)$ at $z = x$ on the upper and lower edges of the cut coincide. It follows that the integrals of $F(z)$ along γ_1 and γ_3 cancel out, leaving

$$\zeta(1-2n) = -\frac{\Gamma(2n)}{2\pi i} I_2,$$

where

$$\begin{aligned} I_2 &= \int_{\gamma_2} F(z)dz \\ &= 2\pi i \operatorname{res}_0(F), \end{aligned}$$

by Cauchy's Theorem.

But

$$\begin{aligned} F(z) &= \frac{z^{-2n}}{e^z - 1} \\ &= z^{-2n-1} \frac{z}{e^z - 1} \\ &= z^{-2n-1} \sum_{r \ge 0} B_r \frac{z^r}{r!} \\ &= \sum_{r \ge 0} B_r \frac{z^{r-2n-1}}{r!}. \end{aligned}$$

By definition, $\operatorname{res}_0(F)$ is the coefficient of z^{-1} in this expansion. Thus

$$\operatorname{res}_0(F) = \frac{B_{2n}}{(2n)!}.$$

Hence

$$\begin{aligned}\zeta(1-2n) &= \frac{\Gamma(2n)}{2\pi i}(-2\pi i)\frac{B_{2n}}{(2n)!}\\ &= -\frac{(2n-1)!}{(2n)!}B_{2n}\\ &= -\frac{B_{2n}}{2n}.\end{aligned}$$

◀

7.4.3 Determining $\zeta(2n)$

Proposition 7.9. *For $n = 1, 2, 3, \ldots,$*

$$\zeta(2n) = (-1)^{n-1}2^{2n-1}\pi^{2n}\frac{B_{2n}}{(2n)!}.$$

Proof ▶ By the functorial equation, Proposition 7.4,

$$\zeta(1-2n) = 2\cos\tfrac{2\pi n}{2}(2\pi)^{-2n}\Gamma(2n)\zeta(2n).$$

Thus

$$\zeta(2n) = (-1)^n 2^{2n-1}\pi^{2n}\frac{\zeta(1-2n)}{\Gamma(2n)}$$

$$= (-1)^{n-1}2^{2n-1}\pi^{2n}\frac{B_{2n}}{2n\Gamma(2n)}$$

$$= (-1)^{n-1}2^{2n-1}\pi^{2n}\frac{B_{2n}}{(2n)!}.$$

◀

For $n = 1$ this gives

$$\zeta(2) = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots = \pi^2 B_2 = \frac{\pi^2}{6},$$

a result which is probably familiar, and which can be proved in several ways.

For $n = 2$ it gives

$$\zeta(4) = 1 + \frac{1}{2^4} + \frac{1}{3^4} + \cdots = -\tfrac{1}{3}\pi^4 B_4 = \frac{\pi^4}{90}.$$

7.5 Postscript

In his seminal paper (Appendix B), Riemann gave a second proof of the functional equation. Although at first sight this seems more complicated than his first proof (given above) it has turned out to have far greater significance.

By Lemma 16, with πn^2 in place of n,

$$\Gamma(s)(\pi n^2)^{-s} = \int_0^\infty x^s e^{-\pi n^2 x} \frac{dx}{x}.$$

Summing over n, as before,

$$\Gamma(s)\pi^{-s}\zeta(2s) = \int_0^\infty z^s \frac{\psi(x) - 1}{2} \frac{dx}{x},$$

where

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

Some 20 years before Riemann's work, Jacobi had published a study of the function $\psi(x)$, in the course of which he showed that $\psi(x)$ satisfies the functional equation

$$\psi\left(\tfrac{1}{x}\right) = x^{-\frac{1}{2}}\psi(x).$$

It is a straightforward matter to derive the functional equation for $\zeta(s)$ from this.

It follows from Jacobi's identity that the *theta function*

$$\theta(x) = \psi(x/i) = \sum e^{piin^2 x}$$

satisfies the equation

$$\theta\left(\tfrac{1}{x}\right) = \sqrt{\tfrac{i}{x}}\theta(x).$$

It is clear that $\theta(x)$ is also periodic with period 1:

$$\theta(x+1) = \theta(x).$$

Now the transformations $x \mapsto 1/x,\ x \mapsto x+1$ generate the *modular group* consisting of the transformations

$$z \mapsto \frac{az+b}{cz+d} \qquad (a,b,c,d \in \mathbb{Z},\ ad-bc=1).$$

This group can be identified with the group of 2×2 matrices

$$SL_2(\mathbb{Z}) = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} : ad - bc = 1\}$$

The relation between zeta functions and modular functions — functions invariant, or nearly invariant, under the modular group — has proved remarkably fruitful. Andrew Wiles' proof of Fermat's Last Theorem, for example, was based on this correspondence.

Another advantage of this approach is that it leads to a functional equation for $L_\chi(s)$, although one relating $L_\chi(1-s)$ to $L_{\bar\chi}(s)$, where $\bar\chi$ is the conjugate character to χ, given by

$$\bar\chi(a) = \overline{\chi(a)} = \chi(a^{-1}).$$

This identity in turn suggests the *Generalised Riemann Hypothesis*, which asserts that *the zeros of* $L_\chi(s)$ *in the critical strip* $0 < \Re(s) < 1$ *all lie on the line* $\Re(s) = 1/2$.

Incidentally, the zeta functions $\zeta_k(s)$ of number fields k, which we briefly alluded to earlier, can all be expressed in terms of the Riemann zeta function $\zeta(s)$ and the L-functions $L_chi(s)$; and the Riemann Hypothesis for $\zeta_k(s)$ would follow from the Generalised Riemann Hypothesis. In that sense, the Generalised Riemann Hypothesis is as general as one would wish.

Contents