Dirichlet's theorem
a real variable approach

Robin Chapman

20 February 2008

We give a proof of Dirichlet's theorem on primes in arithmetic progressions:

> Let N be a positive integer, and a an integer coprime to N. Then there are infinitely many primes p with
>
> $$p \equiv a \pmod{N}.$$

Although our proof is standard (essentially the same as presented by Serre [2], whose was the first account that I understood) we eschew complex-variable theory. Instead we use simple estimates from real-variable theory. Only one new difficulty is introduced: proving that $L(1,\chi) \neq 0$ for nontrivial real-valued Dirichlet characters χ. We use an argument of Monsky [1] to prove this fact.

Our arguments use complex-valued functions of a real variable. A typical function will map an interval $I \subseteq \mathbf{R}$ to $\mathbf{C}$. Such a function may be considered as a pair, consisting its real part and its imaginary part. Standard notions of real analysis transfer to complex-valued functions by considering the real part and imaginary part. For instance a complex-valued function is continuous if and only if both its real and imaginary parts are continuous. We shall use such notions without further comment.

We use big-O notation from time to time. In all cases the reader may interpret this as follows:

$$f(s) = g(s) + O(h(s)) \qquad \text{as } s \to 1^+$$

means that there is C with $|f(s) - g(s)| \leq Ch(s)$ for all s with $1 < s < 2$. This is stronger than the standard definition, but in all cases considered here this stronger definition will hold.

I am indebted to Christophe Smet and Jeremy Scofield for pointing out errors in earlier versions.

1 The Riemann zeta-function

Let s be a real number. We define

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

whenever this sum converges. It is well-known that the sum converges for $s > 1$, but we shall derive this result for ourselves.

To investigate the convergence of the series for $\zeta(s)$ we use the integral test. For $s \leq 0$, $1/n^s \not\to 0$ so we shall assume that $s > 0$. In this case

$$\frac{1}{n^s} > \int_n^{n+1} \frac{dt}{t^s} > \frac{1}{(n+1)^s}.$$

If the infinite integral

$$\int_1^{\infty} \frac{dt}{t^s}$$

converges, then so does

$$\sum_{n=1}^{\infty} \frac{1}{(n+1)^s} = \sum_{n=2}^{\infty} \frac{1}{n^s}$$

and so also does the series for $\zeta(s)$. Conversely, if the integral diverges, then so does the series. We can now prove the following theorem.

Theorem 1 *The series for $\zeta(s)$ converges if and only if $s > 1$. In this case,*

$$\frac{s}{s-1} > \zeta(s) > \frac{1}{s-1}.$$

As a consequence

$$\lim_{s \to 1^+} (s-1)\zeta(s) = 1.$$

Proof We have

$$\int_1^N \frac{dt}{t^s} = \begin{cases} [1 - N^{1-s}]/(s-1) & \text{if } s \neq 1, \\ \log N & \text{if } s = 1. \end{cases}$$

The infinite integral

$$\int_1^{\infty} \frac{dt}{t^s}$$

converges, to $1/(s-1)$, if and only if $s > 1$. Summing the inequality

$$\frac{1}{n^s} > \int_n^{n+1} \frac{dt}{t^s} > \frac{1}{(n+1)^s}$$

over all positive integers n gives

$$\zeta(s) > \frac{1}{s-1} > \zeta(s) - 1$$

and so

$$\frac{s}{s-1} > \zeta(s) > \frac{1}{s-1}.$$

As a consequence, when $s > 1$, we have

$$s > (s-1)\zeta(s) > 1$$

and so

$$\lim_{s \to 1^+} (s-1)\zeta(s) = 1.$$

□

When we consider ζ as a function of a complex variable, $\zeta(s)$ has a simple pole with residue 1 at $s = 1$. This result is a real variable consequence of this fact.

We now need to prove the Euler product formula for ζ. This links the zeta function to the theory of prime numbers.

Theorem 2 *Let $s > 1$. Then*

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

where the product (here and in the sequel) is over all prime numbers p.

Proof Let M be a natural number. Then

$$\prod_{p \le M} \left(1 - \frac{1}{p^s}\right)^{-1} = \prod_{p \le M} \sum_{m=0}^{\infty} \frac{1}{p^{ms}} = \sum_{n \in A_M} \frac{1}{n^s}$$

where A_M is the set of natural numbers none of whose prime factors exceed M. Certainly A_M contains all natural numbers $n \le M$ and so

$$\sum_{n=1}^{M} \frac{1}{n^s} < \prod_{p \le M} \left(1 - \frac{1}{p^s}\right)^{-1} < \zeta(s).$$

Letting M tend to infinity we conclude that

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

□

2 Characters

In this section, let G denote a finite abelian group, having identity e, and the group operation denoted as multiplication. Let $\mathbf{C}^*$ denote the multiplicative group of nonzero complex numbers. A *character* of G is simply a homomorphism from G to $\mathbf{C}^*$. We denote the set of characters of G by $\widehat{G}$. Let χ and ψ be characters of G. Define $\chi\psi : G \to \mathbf{C}^*$ by $\chi\psi(a) = \chi(a)\psi(a)$ for $a \in G$. With this product, the set of characters of G becomes a group.

Theorem 3 *Let G be a finite abelian group. Under the multiplication introduced above, the set $\widehat{G}$ of characters of G forms an abelian group.*

Proof We first show that products of characters are characters. Let χ, $\psi \in \widehat{G}$ and a, $b \in G$. Then

$$\chi\psi(ab) = \chi(ab)\psi(ab) = \chi(a)\chi(b)\psi(a)\psi(b) = \chi\psi(a)\chi\psi(b).$$

Hence $\chi\psi \in \widehat{G}$.

Commutativity and associativity of multiplication of characters follows from the commutativity and associativity of multiplication of complex numbers. We need to show that identity and inverses exist in $\widehat{G}$. Define $\varepsilon : G \to \mathbf{C}^*$ by $\varepsilon(a) = 1$ for all $a \in G$. It is apparent that ε is a character, and that $\varepsilon\chi = \chi$ for all $\chi \in \widehat{G}$. For $\chi \in \widehat{G}$ define χ^{-1} by $\chi^{-1}(a) = \chi(a)^{-1}$. Then, for a, $b \in G$,

$$\chi^{-1}(ab) = \chi(ab)^{-1} = [\chi(a)\chi(b)]^{-1} = \chi(a)^{-1}\chi(b)^{-1} = \chi^{-1}(a)\chi^{-1}(b)$$

so that $\chi^{-1} \in \widehat{G}$ and also

$$\chi\chi^{-1}(a) = \chi(a)\chi^{-1}(a) = \chi(a)\chi(a)^{-1} = 1 = \varepsilon(a)$$

so that $\chi\chi^{-1} = \varepsilon$. Thus $\widehat{G}$ is an abelian group. □

This theorem is a special case of the result that the set of homomorphisms from one abelian group to a second itself is an abelian group.

We shall always denote the identity character, as in the proof of Theorem 3, by ε. If $|G| = n$ then $a^n = e$ for all $a \in G$ and so

$$\chi(a)^n = \chi(a^n) = 1.$$

Hence $\chi(a) \in \mu_n$, the group of n-th roots of unity in $\mathbf{C}^*$, and in particular $|\chi(a)| = 1$. This means that $\chi(a)^{-1} = \overline{\chi(a)}$. We define $\overline{\chi}$ by $\overline{\chi}(a) = \overline{\chi(a)}$. Then $\overline{\chi}$ is the inverse of χ in $\widehat{G}$.

Let H be a subgroup of G. If $\chi \in \widehat{G}$ then $\chi|_H$, the restriction of χ to H, is a character on H. We shall show that each character of H is the restriction of $|G : H|$ characters of G. We first prove a special case.

Theorem 4 *Let H be a subgroup of the finite abelian group G, and suppose that the quotient group G/H is cyclic. Then each character ψ of H is the restriction of $|G:H|$ characters of G.*

Proof Let $m = |G:H|$. Let aH be a generator of the cyclic group G/H. Then $a^m \in H$ and each element of G can be uniquely written as $a^j h$ for $0 \leq j < m$ and $h \in H$.

Let $\psi \in \widehat{H}$ and suppose that $\chi \in \widehat{G}$ with $\chi|_H = \psi$. Define $\eta = \chi(a)$. Then

$$\eta^m = \chi(a)^m = \chi(a^m) = \psi(a^m) \qquad (*)$$

and for $0 \leq j < m$ and $h \in G$,

$$\chi(a^j h) = \chi(a)^j \chi(h) = \eta^j \psi(h). \qquad (\dagger)$$

It follows that χ is determined by ψ and η. By $(*)$ there at most m possible values for η, and so at most m such characters χ. We need to show that for each of the m values of η satisfying $(*)$, defining χ by $(\dagger)$ really gives a character of G.

Take η to be one of the m m-th roots of $\psi(a^m)$ and define χ by $(\dagger)$. Let $0 \leq j, k < m$ and h, $h' \in H$. Then

$$\chi(a^j h)\chi(a^k h') = \eta^{j+k}\psi(h)\psi(h') = \eta^{j+k}\psi(hh').$$

If $j + k < m$ then

$$\chi(a^j h)\chi(a^k h') = \chi(a^{j+k} hh') = \chi((a^j h)(a^k h')).$$

Otherwise $0 \leq j + k - m < m$ and so

$$\begin{aligned} \chi(a^j h)\chi(a^k h') &= \eta^{j+k-m}\eta^m \psi(hh') = \eta^{j+k-m}\psi(a^m)\psi(hh') \\ &= \eta^{j+k-m}\psi(a^m hh') = \chi(a^{j+k-m}a^m hh') = \chi((a^j h)(a^k h')). \end{aligned}$$

Hence $\chi \in \widehat{G}$ and the proof is complete. □

We can now prove the general theorem.

Theorem 5 *Let H be a finite subgroup of the finite abelian group G. Each character ψ of H is the restriction of $|G:H|$ characters of G. In particular $|\widehat{G}| = |G|$.*

Proof Let $G = \langle a_1, \ldots, a_r \rangle$. Define $H_j = H + \langle a_1, \ldots, a_j \rangle$ for $0 \leq j \leq r$. Then, $H_0 = H$, $H_r = G$, $H_j \subseteq H_{j+1}$ for $0 \leq j < r$ and H_{j+1}/H_j is cyclic. Applying Theorem 4 inductively we see that ψ is the restriction of

$$\prod_{j=0}^{r-1} |H_{j+1} : H_j| = |G:H|$$

characters of G.

The trivial group only has one character, so applying this to $H = \{e\}$ gives $|\widehat{G}| = |G : \{e\}| = |G|$. □

We now look at sums of character values and sums over the character group. We now look at sums over the character group.

Theorem 6 *Let G be a finite abelian group, and let $\chi \in \widehat{G}$. Then*

$$\sum_{a \in G} \chi(a) = \begin{cases} |G| & \text{if } \chi = \varepsilon, \\ 0 & \text{if } \chi \neq \varepsilon. \end{cases}$$

Proof If $\chi = \varepsilon$ then $\chi(a) = 1$ for all $a \in G$. hence

$$\sum_{a \in G} \chi(a) = \sum_{a \in G} 1 = |G|.$$

Now suppose that $\chi \neq \varepsilon$. Then there is $b \in G$ with $\chi(b) \neq 1$. As a ranges through the elements of G so does ba. Thus

$$\sum_{a \in G} \chi(a) = \sum_{a \in G} \chi(ba) = \sum_{a \in G} \chi(b)\chi(a) = \chi(b) \sum_{a \in G} \chi(a).$$

As $\chi(b) \neq 1$ then

$$\sum_{a \in G} \chi(a) = 0.$$

□

Theorem 7 *Let G be a finite abelian group, and let $a \in G$. Then*

$$\sum_{\chi \in \widehat{G}} \chi(a) = \begin{cases} |G| & \text{if } a = e, \\ 0 & \text{if } a \neq e. \end{cases}$$

Proof If $\chi \in \widehat{G}$ then $\chi(e) = 1$. Hence

$$\sum_{\chi \in \widehat{G}} \chi(e) = \sum_{\chi \in \widehat{G}} 1 = |\widehat{G}| = |G|.$$

Now suppose that $a \neq e$. Let $H = \langle a \rangle$. As each $\psi \in \widehat{H}$ is the restriction of $|G : H|$ characters of G, we have

$$\sum_{\chi \in \widehat{G}} \chi(a) = |G : H| \sum_{\psi \in \widehat{H}} \psi(a).$$

But $H = \langle a \rangle$ is cyclic, of order $m > 1$. The characters of H have the form $\psi_j : a^k \mapsto \exp(2\pi ijk/m)$ for $0 \leq j < m$. Thus

$$\sum_{\psi \in \widehat{H}} \psi(a) = \sum_{j=0}^{m-1} \psi_j(a) = \sum_{j=0}^{m-1} \exp(2\pi ij/m) = 0$$

by the formula for the sum of a geometric progression. Consequently,

$$\sum_{\chi \in \widehat{G}} \chi(a) = 0.$$

□

3 Dirichlet L-functions

Let N be a natural number. Let $\mathbf{Z}_N$ denote the ring of integers modulo N. Let $\mathbf{Z}_N^*$ denote the group of units of $\mathbf{Z}_N$. The group $\mathbf{Z}_N^*$ consists of the residues of the numbers coprime to N, and has $\phi(N)$ elements, where ϕ denotes Euler's totient function.

A *Dirichlet character* modulo N is a character of the group $\mathbf{Z}_N^*$. Define X_N to be the group of Dirichlet characters modulo N. We use ε to denote the trivial Dirichlet character. We consider a Dirichlet character $\chi \in X_N$ as a function on $\mathbf{Z}$ as follows: we set

$$\chi(a) = \begin{cases} \chi(a + N\mathbf{Z}) & \text{if } a \text{ is coprime to } N, \\ 0 & \text{if } a \text{ is not coprime to } N. \end{cases}$$

Let $\chi \in X_N$, and let s be real. We define the *Dirichlet L-function* $L(s, \chi)$ as

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

whenever this sum converges.

Before we can determine when the series for the L-function converges, we need a preliminary result

Theorem 8 *Let $\chi \in X_N$, and suppose that $\chi \neq \varepsilon$. Define $a_n = \sum_{j=1}^{n} \chi(j)$. Then $|a_n| \leq N$ for all nonnegative integers n.*

Proof First of all note that by Theorem 6

$$\sum_{j=1}^{N} \chi(j) = \sum_{a \in \mathbf{Z}_N^*} \chi(a) = 0.$$

Suppose that $n > N$. Then

$$a_n = \sum_{j=1}^{n} \chi(j) = \sum_{j=1}^{N} \chi(j) + \sum_{j=N+1}^{n} \chi(j) = 0 + \sum_{j=1}^{n-N} \chi(j) = a_{n-N}.$$

By iteration, $a_n = a_r$ where r is the least nonnegative residue of n modulo N. Then

$$|a_n| = |a_r| \leq \sum_{j=1}^{r} |\chi_j| = r \leq N.$$

□

We can now determine where the L-functions converge.

Theorem 9 *Let $\chi \in X_N$. If $\chi = \varepsilon$ then the series for $L(s, \chi)$ converges for $s > 1$. If $\chi \neq \varepsilon$ then the series for $L(s, \chi)$ converges for $s > 0$.*

Proof When $\chi = \varepsilon$, then $\chi(n) \in \{0, 1\}$ for all n. Then

$$0 \leq \sum_{n=1}^{M} \frac{\chi(n)}{n^s} \leq \sum_{n=1}^{M} \frac{1}{n^s}$$

and by comparison with the series for $\zeta(s)$ this series converges for $s > 1$.

Suppose that $\chi \neq \varepsilon$. We use the technique of partial summation. Let

$$a_n = \sum_{j=1}^{n} \chi(j).$$

Then $a_0 = 0$ and $\chi(n) = a_n - a_{n-1}$ for $n \geq 1$. Thus

$$\begin{aligned}
\sum_{n=1}^{M} \frac{\chi(n)}{n^s} &= \sum_{n=1}^{M} \frac{a_n - a_{n-1}}{n^s} \\
&= \sum_{n=1}^{M} \frac{a_n}{n^s} - \sum_{n=1}^{M} \frac{a_{n-1}}{n^s} \\
&= \sum_{n=1}^{M} \frac{a_n}{n^s} - \sum_{n=1}^{M-1} \frac{a_n}{(n+1)^s} \\
&= \frac{a_M}{M^s} + \sum_{n=1}^{M-1} a_n \left[\frac{1}{n^s} - \frac{1}{(n+1)^s} \right].
\end{aligned}$$

We apply the mean value theorem to the function $x \mapsto x^{-s}$ to get

$$\frac{1}{n^s} - \frac{1}{(n+1)^s} = \frac{s}{b_n^{s+1}}$$

where $n < b_n < n+1$. In particular

$$\frac{1}{n^s} - \frac{1}{(n+1)^s} < \frac{s}{n^{s+1}}$$

If $s > 0$ then $a_M/M^s \to 0$ as $M \to \infty$ as, by Theorem 8, the sequence (a_n) is bounded. Also

$$\sum_{n=1}^{M-1} a_n \left[\frac{1}{n^s} - \frac{1}{(n+1)^s}\right] \leq As \sum_{n=1}^{\infty} \frac{1}{n^{s+1}} = As\zeta(s+1)$$

where A is an upper bound for the sequence (a_n). Hence the series for $L(s,\chi)$ converges for $s > 0$, and indeed

$$L(s,\chi) = \sum_{n=1}^{\infty} a_n \left[\frac{1}{n^s} - \frac{1}{(n+1)^s}\right].$$

□

Dirichlet L-functions also have Euler products when $s > 1$.

Theorem 10 *Let $\chi \in X_N$ and let $s > 1$. Then*

$$L(s,\chi) = \prod_p \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}$$

where the product is over all prime numbers p.

Proof Let M be a natural number. Then

$$\prod_{p \leq M} \left(1 - \frac{\chi(p)}{p^s}\right)^{-1} = \prod_{p \leq M} \sum_{k=0}^{\infty} \frac{\chi(p)^k}{p^{ks}} = \sum_{n \in A_M} \frac{\chi(n)}{n^s}$$

where A_M is the set of natural numbers none of whose prime factors exceed M. This follows from the fact that χ is completely multiplicative: $\chi(ab) = \chi(a)\chi(b)$ for all a and b. Certainly A_M contains all natural numbers $n \leq M$ and so

$$\left| L(s,\chi) - \prod_{p \leq M} \left(1 - \frac{\chi(p)}{p^s}\right)^{-1} \right| \leq \sum_{n=M+1}^{\infty} \frac{1}{n^s}$$

which is the tail in the convergent series for $\zeta(s)$. Letting M tend to infinity we conclude that

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

□

When $\chi = \varepsilon$, the L-function is essentially the same as the ζ-function.

Theorem 11 *Let $\chi = \varepsilon \in X_N$. Then*

$$\lim_{s \to 1^+} (s-1)L(s, \chi) = \frac{\phi(N)}{N}$$

where ϕ denotes Euler's ϕ-function.

Proof Let $s > 1$. We have

$$L(s, \chi) = \prod_p \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}$$

and

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

But $\chi(p) = 0$ if $p \mid N$ and $\chi(p) = 1$ if $p \nmid N$. Hence

$$L(s, \chi) = \zeta(s) \prod_{p|N} \left(1 - \frac{1}{p^s}\right).$$

But $\lim_{s \to 1^+}(s-1)\zeta(s) = 1$. Hence

$$\lim_{s \to 1^+} (s-1)L(s, \chi) = \lim_{s \to 1^+} \prod_{p|N} \left(1 - \frac{1}{p^s}\right) = \prod_{p|N} \frac{p-1}{p} = \frac{\phi(N)}{N}.$$

□

For nontrivial characters χ we shall need to consider the value $L(1, \chi)$ and the behaviour of $L(s, \chi)$ for s just above $s = 1$.

Theorem 12 *Let $\chi \in X_N$ with $\chi \neq \varepsilon$. There exists $C > 0$ such that*

$$L(s, \chi) = L(1, \chi) + O(s - 1)$$

as $s \to 1^+$. In particular

$$\lim_{s \to 1^+} L(s, \chi) = L(1, \chi).$$

Proof Let $1 < s < 2$. From the proof of Theorem 9 we have

$$L(s,\chi) - L(1,\chi) = \sum_{n=1}^{\infty} a_n \left[\left(\frac{1}{n^s} - \frac{1}{(n+1)^s} \right) - \left(\frac{1}{n} - \frac{1}{n+1} \right) \right]$$

where the sequence (a_n) is bounded. Applying the mean value theorem to the function $s \mapsto n^{-s} - (n+1)^{-s}$ gives a sequence (s_n) with $1 < s_n < s$ and

$$L(s,\chi) - L(1,\chi) = (s-1) \sum_{n=1}^{\infty} a_n \left[\frac{\log(n+1)}{(n+1)^{s_n}} - \frac{\log n}{n^{s_n}} \right].$$

A further application of the mean value theorem, to $x \mapsto \log x / x^{s_n}$, gives a sequence b_n with $n < b_n < n+1$ and

$$L(s,\chi) - L(1,\chi) = (s-1) \sum_{n=1}^{\infty} a_n \left[\frac{1 - s_n \log b_n}{b_n^{s_n+1}} \right].$$

Let A be an upper bound for the sequence $(|a_n|)$. Define

$$C = A \sum_{n=1}^{\infty} \frac{1 + 2\log(n+1)}{n^2}.$$

To see that this series is convergent first note that $\log(n+1) \le \log(2n) = \log 2 + \log n$ for $n \ge 1$. Now define $f(x) = \log x / x^{1/2}$ for $x > 0$. Then $f(x) \ge 0$ for $x \ge 1$ and

$$f'(x) = \frac{2 - \log x}{2x^{3/2}}.$$

Then $f'(x) > 0$ for $1 \le x < e^2$ and $f'(x) < 0$ for $x > e^2$. Hence $f(x) \le f(e^2)$ for all $x \ge 1$, that is

$$\frac{\log x}{x^{1/2}} \le \frac{2}{e}$$

for all $x \ge 1$. Hence

$$0 < \frac{1 + 2\log(n+1)}{n^2} \le \frac{1 + 2\log 2 + 2\log n}{n^2} \le \frac{1 + 2\log 2}{n^2} + \frac{4}{en^{3/2}}$$

for all positive integers n. It follows that

$$\sum_{n=1}^{N} \frac{1 + 2\log(n+1)}{n^2} \le (1 + \log 2)\zeta(2) + (4/e)\zeta(3/2)$$

and so the sum defining C is convergent. For each n,

$$\left| \frac{1 - s_n \log b_n}{b_n^{s_n+1}} \right| \le \frac{1 + s_n \log(n+1)}{n^{s_n+1}} \le \frac{1 + 2\log(n+1)}{n^2}$$

and so indeed

$$|L(s,\chi) - L(1,\chi)| \leq C(s-1)$$

for $1 < s < 2$. It is immediate that

$$\lim_{s \to 1^+} L(s,\chi) = L(1,\chi).$$

□

4 Proof of Dirichlet's theorem

We need the notion of the *Dirichlet density* of sets of prime numbers.

Theorem 13 *Let*

$$\psi(s) = \sum_p \frac{1}{p^s}$$

where the sum is over all primes p. Then this series is convergent for $s > 1$ and

$$\psi(s) = -\log(s-1) + O(1)$$

as $s \to 1^+$.

Proof Let $s > 1$. The series for $\psi(s)$ converges by comparison with that for $\zeta(s)$. From the Euler product for $\zeta(s)$, we have

$$\log \zeta(s) = -\sum_p \log\left(1 - \frac{1}{p^s}\right) = \sum_p \sum_{m=1}^{\infty} \frac{1}{mp^{ms}} = \psi(s) + \omega(s)$$

where

$$\omega(s) = \sum_p \sum_{m=2}^{\infty} \frac{1}{mp^{ms}}.$$

But

$$\begin{aligned} 0 < \omega(s) &< \sum_p \sum_{m=2}^{\infty} \frac{1}{2p^{ms}} = \frac{1}{2} \sum_p \frac{1}{p^{2s}} \left(1 - \frac{1}{p^s}\right)^{-1} = \frac{1}{2} \sum_p \frac{1}{(p^s - 1)p^s} \\ &< \frac{1}{2} \sum_p \frac{1}{(p-1)p} < \frac{1}{2} \sum_{m=1}^{\infty} \frac{1}{m(m+1)} = \frac{1}{2}. \end{aligned}$$

It follows that

$$\psi(s) = \log \zeta(s) + O(1)$$

as $s \to 1^+$. But

$$\frac{1}{s-1} < \zeta(s) < \frac{s}{s-1}$$

for $s > 1$, and so

$$-\log(s-1) < \log \zeta(s) < \log s - \log(s-1)$$

and hence

$$\log \zeta(s) = -\log(s-1) + O(1)$$

as $s \to 1^+$. Consequently

$$\psi(s) = -\log(s-1) + O(1)$$

as $s \to 1^+$. □

Let A denote a subset of the set of primes. The *Dirichlet density* of A is

$$\lim_{s \to 1^+} \frac{-1}{\log(s-1)} \sum_{p \in A} \frac{1}{p^s}$$

provided this limit exists. By Theorem 13, the set of all primes has Dirichlet density 1. Evidently any finite set of primes has Dirichlet density 0. Any set of primes with nonzero Dirichlet density is infinite.

We need to show that $L(1, \chi)$ is nonzero whenever $\chi \neq \varepsilon$. The next result is a start in this direction.

Theorem 14 *Let N be a positive integer. Then*

$$\prod_{\chi \in X_N} L(s, \chi) > 1$$

for all $s > 1$.

Proof Using Euler's product we have

$$\prod_{\chi \in X_N} L(s, \chi) = \prod_p \prod_{\chi \in X_N} \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}.$$

If $p \mid N$ then

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

Otherwise $\chi(p) \neq 0$ for all $\chi \in X_N$. Let H be the subgroup of $\mathbf{Z}_N^*$ generated by the residue of p, and let $r = |H|$. By Theorem 5

$$\prod_{\chi \in X_N} \left(1 - \frac{\chi(p)}{p^s}\right)^{-1} = \prod_{\psi \in \widehat{H}} \left(1 - \frac{\psi(p)}{p^s}\right)^{-\phi(N)/r}.$$

As H is cyclic of order r, generated by p, its characters are ψ_j where $\psi_j(p) = \exp(2\pi ij/r)$. Thus

$$\prod_{\psi \in \widehat{H}} \left(1 - \frac{\psi(p)}{p^s}\right) = \prod_{j=0}^{r-1} \left(1 - \frac{\exp(2\pi ij/r)}{p^s}\right) = \left(1 - \frac{1}{p^{rs}}\right)$$

Hence

$$\prod_{\chi \in X_N} \left(1 - \frac{\chi(p)}{p^s}\right)^{-1} = \left(1 - \frac{1}{p^{rs}}\right)^{-\phi(N)/r} > 1$$

for $s > 1$. It follows that

$$\prod_{\chi \in X_N} L(s, \chi) > 1$$

for $s > 1$. □

We can now prove that $L(1, \chi) \neq 0$ for all non-real valued Dirichlet characters χ.

Theorem 15 *Let $\chi \in X_N$, and suppose that $\chi \neq \overline{\chi}$. Then*

$$L(1, \chi) \neq 0.$$

Proof If $\chi \neq \overline{\chi}$ the three characters ε, χ and $\overline{\chi} \in X_N$ are all distinct. Suppose that $L(1, \chi) = 0$. Then

$$L(1, \overline{\chi}) = \sum_{n=1}^{\infty} \frac{\overline{\chi(n)}}{n} = \overline{\sum_{n=1}^{\infty} \frac{\chi(n)}{n}} = \overline{L(1, \chi)} = 0.$$

By Theorems 11 and 12, $L(s, \varepsilon) = O(1/(s-1))$, $L(s, \chi) = O(s-1)$ and $L(s, \overline{\chi}) = O(s-1)$ as $s \to 1^+$. Thus

$$\lim_{s \to 1^+} L(s, \varepsilon) L(s, \chi) L(s, \overline{\chi}) = 0.$$

Also

$$\lim_{s \to 1^+} \prod_{\psi \in X_N, \psi \neq \varepsilon, \chi, \overline{\chi}} L(s, \psi) = \prod_{\psi \in X_N, \psi \neq \varepsilon, \chi, \overline{\chi}} L(1, \psi)$$

It follows that

$$\lim_{s\to 1^+} \prod_{\psi\in X_N} L(s,\psi) = 0.$$

This is impossible, as

$$\prod_{\psi\in X_N} L(s,\psi) > 1$$

for all $s > 1$, by Theorem 14. This contradiction shows that it is impossible for $L(1,\chi)$ to equal zero. □

The most difficult part in the proof of Dirichlet's theorem is the proof that $L(1,\chi) \neq 0$ when χ is a nontrivial real-valued Dirichlet character. To achieve this some trick or other is needed. We shall employ an elegant device due to Monsky [1].

Theorem 16 *Let $\chi \in X_N$, and suppose that $\chi \neq \varepsilon$ and $\chi = \overline{\chi}$. Then*

$$L(1,\chi) \neq 0.$$

Proof We introduce a *Lambert series*

$$f(x) = \sum_{d=1}^{\infty} \frac{\chi(d)x^d}{1-x^d}.$$

This series converges for $0 < x < 1$ by comparison with

$$\sum_{d=1}^{\infty} \frac{x^d}{1-x^d}.$$

which converges by the ratio test. We calculate

$$f(x) = \sum_{d=1}^{\infty} \chi(d) \sum_{t=1}^{\infty} x^{dt} = \sum_{n=1}^{\infty} x^n \sum_{d|n} \chi(d) = \sum_{n=1}^{\infty} c_n x^n$$

where

$$c_n = \sum_{d|n} \chi(d).$$

We need to show that $c_n \geq 0$ for all n. As each $\chi(d)$ is real and has modulus 0 or 1 then $\chi(d) \in \{-1, 0, 1\}$ for all d. Clearly $c_1 = \chi(1) = 1$. We use induction on n. If $n > 1$ then $n = p^k m$ where p is prime, $k > 0$ and $p \nmid m$. Then

$$c_n = \sum_{d|p^k m} \chi(d) = \sum_{j=0}^{k} \sum_{r|m} \chi(p^j r) = \sum_{j=0}^{k} \sum_{r|m} \chi(p)^j \chi(r) = c_m \sum_{j=0}^{k} \chi(p)^j.$$

Thus

$$c_n = \begin{cases} (k+1)c_m & \text{if } \chi(p) = 1, \\ c_m & \text{if } \chi(p) = 0, \\ c_m & \text{if } \chi(p) = -1 \text{ and } k \text{ is even}, \\ 0 & \text{if } \chi(p) = -1 \text{ and } k \text{ is odd}. \end{cases}$$

Inductively, as $c_m \geq 0$, then $c_n \geq 0$. Also $c_{n^k} = 1$ for all $k \geq 0$ and so $\sum_{r=1}^{\infty} c_r$ diverges. For each $M > 0$,

$$\limsup_{x \to 1^-} f(x) \geq \lim_{x \to 1^-} \sum_{n=1}^{M} c_n x^n = \sum_{n=1}^{M} c_n.$$

Hence $f(x) \to \infty$ as $x \to 1^-$.

Suppose that $L(1, \chi) = 0$. Then

$$-f(x) = \frac{L(1,\chi)}{1-x} - f(x) = \sum_{n=1}^{\infty} \chi(n) \left[\frac{1}{n(1-x)} - \frac{x^n}{1-x^n} \right].$$

Let

$$b_n(x) = \frac{1}{n(1-x)} - \frac{x^n}{1-x^n}.$$

Then for $0 < x < 1$,

$$\begin{aligned} (1-x)(b_n(x) - b_{n+1}(x)) &= \frac{1}{n} - \frac{1}{n+1} - \frac{x^n(1-x)}{(1-x^n)} - \frac{x^{n+1}(1-x)}{(1-x^{n+1})} \\ &= \frac{1}{n(n+1)} - \frac{x^n(1-x)^2}{(1-x^n)(1-x^{n+1})}. \end{aligned}$$

By the arithmetic-geometric mean inequality,

$$\frac{1-x^n}{1-x} = \sum_{j=0}^{n-1} x^j = \frac{1}{2} \sum_{j=0}^{n-1} (x^j + x^{n-1-j}) \geq n x^{(n-1)/2}.$$

Replacing n by $n+1$ gives

$$\frac{1-x^{n+1}}{1-x} \geq (n+1) x^{n/2}.$$

Hence

$$(1-x)(b_n(x) - b_{n+1}(x)) \geq \frac{1-x^{1/2}}{n(n+1)}.$$

Thus $(b_n(x))$ is a decreasing sequence whenever $0 < x < 1$.

Set $a_m = \sum_{j=1}^{m} \chi(j)$. The sequence (a_m) is bounded by Theorem 8; suppose that $|a_m| \leq A$ for all m. Then

$$\sum_{n=1}^{M} \chi(n) b_n(x) = \sum_{n=1}^{M} (a_n - a_{n-1}) b_n(x) = a_M b_M(x) + \sum_{n=1}^{M-1} a_n (b_n(x) - b_{n+1}(x)).$$

As $M \to \infty$, $b_M(x) \to 0$ and so

$$-f(x) = \sum_{n=1}^{\infty} a_n (b_n(x) - b_{n+1}(x)).$$

Hence

$$|f(x)| \leq \sum_{n=1}^{\infty} |a_n| (b_n(x) - b_{n+1}(x)) \leq A \sum_{n=1}^{\infty} (b_n(x) - b_{n+1}(x)) = A b_1(x).$$

But

$$b_1(x) = \frac{1}{1-x} - \frac{x}{1-x} = 1.$$

It follows that $f(x)$ is bounded for $0 < x < 1$. But this contradicts the fact that $f(x) \to \infty$ as $x \to 1^-$. Hence we cannot have $L(1, \chi) = 0$. □

We now look at a sum over primes weighted by a nontrivial Dirichlet character.

Theorem 17 *Let $\chi \in X_N$ with $\chi \neq \varepsilon$. Define*

$$\psi_\chi(s) = \sum_p \frac{\chi(p)}{p^s}$$

where the sum is over all primes p. Then this series is convergent for $s > 1$ and

$$\psi_\chi(s) = O(1)$$

as $s \to 1^+$.

Proof By Theorem 13 the series

$$\psi(s) = \sum_p \frac{1}{p^s}$$

converges for $s > 1$. By comparison the series for $\psi_\chi(s)$ also converges for $s > 1$.

Let $s > 1$. From the Euler product for $L(s, \chi)$, we have

$$\log L(s,\chi) = -\sum_p \log\left(1 - \frac{\chi(p)}{p^s}\right) = \sum_p \sum_{m=1}^{\infty} \frac{\chi(p^m)}{mp^{ms}} = \psi_\chi(s) + \omega_\chi(s)$$

where

$$\omega_\chi(s) = \sum_p \sum_{m=2}^{\infty} \frac{\chi(p^m)}{mp^{ms}}.$$

But

$$|\omega_\chi(s)| \leq \omega(s) = \sum_p \sum_{m=2}^{\infty} \frac{1}{mp^{ms}}$$

and from the proof of Theorem 13 the series $\omega(s) = O(1)$ as $s \to 1^+$. As $L(1, \chi) \neq 0$ and

$$\lim_{s \to 1^+} L(s,\chi) = L(1,\chi)$$

then

$$\log L(s,\chi) = O(1)$$

as $s \to 1^+$ and so

$$\psi_\chi(s) = O(1)$$

as $s \to 1^+$. □

The proof of Dirichlet's theorem is now almost immediate.

Theorem 18 (Dirichlet) *Let N be a positive integer, and let a be an integer coprime to N. Then the set*

$$P_{a,N} = \{p : p \text{ prime}, p \equiv a \pmod{N}\}$$

has Dirichlet density $1/\phi(N)$. In particular, $P_{a,N}$ is infinite, that is, there are infinitely many primes congruent to a modulo N.

Proof For $\chi \in X_N$ let

$$\psi_\chi(s) = \sum_p \frac{\chi(p)}{p^s}$$

for $s > 1$. If $\chi \neq \varepsilon$, we have seen that $\psi_\chi(s) = O(1)$ as $s \to 1^+$. For $\chi = \varepsilon$ we have

$$\psi_\varepsilon(s) = -\sum_{p|N} \frac{1}{p^s} + \sum_p \frac{1}{p^s}.$$

We have seen that

$$\sum_p \frac{1}{p^s} = -\log(s-1) + O(1)$$

as $s \to 1^+$, and so

$$\psi_\varepsilon(s) = -\log(s-1) + O(1)$$

as $s \to 1^+$.

Let a be coprime to N and consider

$$\theta_a(s) = \sum_{\chi \in X_N} \overline{\chi(a)}\psi_\chi(s) = \sum_{\chi \in X_N} \overline{\chi(a)} \sum_p \frac{\chi(p)}{p^s} = \sum_p \frac{1}{p^s} \sum_{\chi \in X_N} \overline{\chi(a)}\chi(p).$$

Let b be a positive integer with $ab \equiv 1 \pmod N$. Then $\chi(b) = \overline{\chi(a)}$ and so by Theorem 7

$$\sum_{\chi \in X_N} \overline{\chi(a)}\chi(p) = \sum_{\chi \in X_N} \chi(bp) = \begin{cases} |X_N| & \text{if } bp \equiv 1 \pmod N, \\ 0 & \text{if } bp \not\equiv 1 \pmod N. \end{cases}$$

As $bp \equiv 1 \pmod N$ if and only if $p \equiv a \pmod N$ then

$$\theta_a(s) = |X_N| \sum_{p \in P_{a,N}} \frac{1}{p^s}.$$

But

$$\theta_a(s) = -\log(s-1) + O(1)$$

and so $P_{a,N}$ has Dirichlet density $1/|X_N| = 1/|\mathbf{Z}_N^*| = 1/\phi(N)$. In particular $P_{a,N}$ is infinite. □

5 Remarks on complex variables

We briefly indicate on how this proof may be simplified when we use complex analysis.

In accordance with the conventions of analytic number theory, let $s = \sigma + it$ denote a complex variable in this section. Then the series for $\zeta(s)$ converges for $\sigma > 1$. The convergence is uniform on sets $\{s : \sigma > c\}$ where $c > 1$ so that $\zeta(s)$ is holomorphic for $\sigma > 1$. Similarly using partial summation, $L(s, \chi)$ is holomorphic for $\sigma > 0$ whenever χ is a nontrivial character. Then Theorem 12 become trivial (but also unnecessary). Also $\zeta(s) - 1/(s-1)$ has an analytic continuation on the set $\{s : \sigma > 0\}$. A corresponding result is true for $L(s, \varepsilon)$.

If $L(1, \chi)$ vanishes for some $\chi \in X_N$ then

$$f(s) = \prod_{\chi \in X_N} L(s, \chi)$$

is holomorphic on $\{s : \sigma > 0\}$. But for $s > 1$

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

where the a_n are nonnegative. A standard result on series of this form (*Dirichlet series*) shows that if all a_n are nonnegative and $f(s)$ analytically continues to $\{s : \sigma > 0\}$, then the series is convergent there. However examination of our series shows that there is a real number s with $0 < s < 1$ such that the series fails to converge. This gives the crucial non-vanishing result for the $L(1, \chi)$. This approach is conceptually simpler, and much less *ad hoc* than the proof we give here.

References

[1] Paul Monsky, 'Simplifying the proof of Dirichlet's theorem', *American Mathematical Monthly*, **100** (1993) 861–862.

[2] Jean-Pierre Serre, *A Course in Arithmetic*, Springer-Verlag, 1973.