1 Dirichlet’s theorem — set up

Dirichlet’s famous theorem from 1837 asserts that if \( k \) is a positive integer and \( a \) is an integer coprime to \( k \), then there are infinitely many primes \( p \) with \( p \equiv a \pmod{k} \). In this unit we will prove the stronger assertion that the sum of the reciprocals of the primes \( p \equiv a \pmod{k} \) is divergent.

Important in the proof is the use of the orthogonality relations for the Dirichlet characters mod \( k \) to pick out the residue class \( a \pmod{k} \). First note that if \( \chi \) is a character mod \( k \) and \( s > 1 \), then the Euler product

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

shows that \( L(s, \chi) \neq 0 \). One could say that of course it isn’t 0 since a product of non-zero numbers is not 0. But this is an infinite product and it might converge to 0. To see that it doesn’t, note that

\[
\left| (1 - \chi(p)p^{-s})^{-1} \right| \geq (1 + p^{-s})^{-1}.
\]

The logarithm of the product of these latter factors is

\[
-\sum_p \log (1 + p^{-s}) = \sum_p \sum_{j \geq 1} (-1)^{j-1} \frac{1}{j} p^{-js} = \sum_p p^{-s} + E(s),
\]

where

\[
|E(s)| \leq \sum_p \sum_{j \geq 1} \frac{1}{j} p^{-js} < \frac{1}{2} \sum_p \sum_{j \geq 2} p^{-js} = O \left( \sum_p p^{-2} \right) = O(1),
\]

for \( s \geq 1 \). Since \( \sum_p p^{-s} \) converges for \( s > 1 \), it follows that \( \log |L(s, \chi)| \) is defined for \( s > 1 \), so that \( L(s, \chi) \neq 0 \) for \( s > 1 \). Note too that if \( \chi \) is a real character, meaning that it takes values in \( \{0, 1, -1\} \), then \( L(s, \chi) > 0 \) for \( s > 1 \), since the factors \( (1 - \chi(p)p^{-s})^{-1} \) are all positive.

**Proposition 1.** For \( k \) a positive integer and \( s > 1 \),

\[
\prod_{\chi \pmod{k}} L(s, \chi) \geq 1.
\]

**Proof.** We take the log of the product of the complex numbers \( L(s, \chi) \). In analysis one learns how to take the log of a complex number (it is actually multi-valued) and how to take the
Taylor series of $\log(1 - z)$ for $|z| < 1$, which converges to one branch of the log. We shall assume that $\log$ has the usual properties (such as the log of a product is the sum of the logs) and

$$\log(1 - z) = \sum_{j=1}^{\infty} \frac{1}{j} z^j, \text{ for } |z| < 1.$$ 

The product in the proposition is of nonzero complex numbers, so using the Euler products for the $L$-functions $L(s, \chi)$, we have for $s > 1$,

$$\log \left( \prod_{\chi \pmod{k}} L(s, \chi) \right) = \sum_{\chi \pmod{k}} \log(L(s, \chi)) = \sum_{\chi \pmod{k}} \sum_p -\log \left( 1 - \chi(p)p^{-s} \right)$$

$$= \sum_{\chi \pmod{k}} \sum_p \sum_{j \geq 1} \frac{\chi(p)^j}{jp^{js}}.$$

Rewriting $\chi(p)^j$ as $\chi(p^j)$ and bringing the sum on $\chi$ inside, we have

$$\log \left( \prod_{\chi \pmod{k}} L(s, \chi) \right) = \sum_p \sum_{j \geq 1} \frac{1}{jp^{js}} \sum_{\chi \pmod{k}} \chi(p^j).$$

By the orthogonality relations, the inner sum is $\varphi(k)$ when $p^j \equiv 1 \pmod{k}$ and is 0 otherwise, so

$$\log \left( \prod_{\chi \pmod{k}} L(s, \chi) \right) = \varphi(k) \sum_{p, j: p^j \equiv 1 \pmod{k}} \frac{1}{jp^{js}} \geq 0.$$

Since this sum is real and non-negative, it follows that the product in the proposition is real and at least 1. This completes the proof. \qed

Note that it is easy to see that the product in the Proposition is real and positive for $s > 1$. We had already noted that for a real character $\chi$ that $L(s, \chi) > 0$ and for any character $\chi \pmod{k}$, $L(s, \chi) \neq 0$ when $s > 1$. In the product, each non-real character $\chi$ can be paired with its complex conjugate character $\overline{\chi}$, and then $L(s, \chi)L(s, \overline{\chi}) = |L(s, \chi)|^2 > 0$. The more substantive accomplishment of Proposition 1 is that the product is bounded away from 0.

### 2 Dirichlet’s theorem — the main idea

Suppose that $k$ is a positive integer and the integer $a$ is coprime to $k$. We have from the orthogonality relations that for $s > 1$,

$$\sum_{p \equiv a \pmod{k}} \frac{1}{p^s} = \frac{1}{\varphi(k)} \sum_p \sum_{\chi \pmod{k}} \frac{\overline{\chi(a)}\chi(p)}{p^s} = \frac{1}{\varphi(k)} \sum_{\chi \pmod{k}} \overline{\chi(a)} \sum_p \frac{\chi(p)}{p^s}. \quad (2)$$
The magic here is that we have transformed a sum over the special primes in the residue class \(a\ mod\ k\) to a sum over all primes. In this last sum, one of the terms corresponds to the principal character \(\chi_1\ (mod\ k)\) (it is the characteristic function of the integers coprime to \(k\)). The contribution of this one character to the sum tends to \(+\infty\) as \(s \to 1^+\). Thus, the left side of the equation will also tend to \(+\infty\) as \(s \to 1^+\), so proving that \(\sum_{p=a(mod\ k)} 1/p\) is divergent, provided we show that

\[
\left| \sum_{p} \frac{\chi(p)}{p^s} \right| = O(1) \text{ for each } \chi \ mod\ k \text{ with } \chi \neq \chi_1 \text{ and for } s \geq 1. \tag{3}
\]

That is, (3) applied to (2) immediately implies Dirichlet’s theorem.

We have almost seen the sum \(\sum_{p} \chi(p)p^{-s}\) in the preceding section. In fact,

\[
\log(L(s, \chi)) - \sum_{p} \frac{\chi(p)}{p^s} = \sum_{p} \sum_{j \geq 2} \chi(p^j) / j p^s =: E_\chi(s), \tag{4}
\]

say. (This is the definition of \(E_\chi(s)\).) The series for \(E_\chi(s)\) converges absolutely when \(s \geq 1\), as we have seen several times, including in the last section. So, we have \(|E_\chi(s)| = O(1)\) uniformly for \(s \geq 1\). If \(L(1, \chi) \neq 0\), then \(L(s, \chi)\) is bounded away from 0 for \(s \geq 1\) and thus \(|\log(L(s, \chi))| = O(1)\) for these values of \(s\). We conclude that (4) and \(L(1, \chi) \neq 0\) imply that (3) holds for \(\chi\).

Since (3) has been seen to imply Dirichlet’s theorem, we have thus reduced Dirichlet’s theorem to the following result.

**Proposition 2.** For \(k\) a positive integer and \(\chi\) a non-principal character \(mod\ k\), we have \(L(1, \chi) \neq 0\).

### 3 Dirichlet’s theorem — the proof of Proposition 2

First note that for \(s > 1\), we have

\[
L(s, \chi_1) \leq \zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} < 1 + \int_{1}^{\infty} \frac{dt}{t^s} = \frac{s}{s - 1}. \tag{5}
\]

(The last inequality uses \(1/n^s < \int_{n-1}^{n} dt/t^s\) for \(n \geq 2\).) Next note that if \(L(1, \chi) = 0\), then

\[
|L(s, \chi)| = O(s - 1) \text{ for } s \geq 1. \tag{6}
\]

This follows from the Mean Value Theorem (MVT) in elementary calculus. Indeed, for \(\chi\) non-principal (which must be the case if \(L(1, \chi) = 0\)), we have

\[
L'(s, \chi) = -\sum_{n} \frac{\chi(n) \log n}{n^s}, \text{ so } |L'(s, \chi)| = O(1) \text{ for } s \geq 1.
\]
By the MVT, for each \( s > 1 \) there is some \( s_0 \in (1, s) \) with

\[
\frac{L(s, \chi) - L(1, \chi)}{s - 1} = L'(s_0, \chi).
\]

Using \( L(1, \chi) = 0 \), we multiply this equation by \( s - 1 \) to obtain (6).

Suppose \( \psi \) is a non-real character mod \( k \) and \( L(1, \psi) = 0 \). Then of course we also have
\( L(1, \bar{\psi}) = 0 \). In the product in Proposition 1 we have the three factors corresponding to \( \chi_1, \psi, \bar{\psi} \). Every other factor has absolute value that is \( O(1) \) for \( s \geq 1 \). Thus, using (5) and (6), the absolute value of the product is \( O(s(s - 1)) \), contradicting Proposition 1 for \( s \) close to 1. This proves Proposition 2 for \( \chi \) non-real.

It remains to consider the case of \( \chi \) real and non-principal. Let \( A = \chi \ast u \), so that

\[
A(n) = (\chi \ast u)(n) = \sum_{d \mid n} \chi(d).
\]

Then \( A \) is a multiplicative function. For a prime power \( p^a \), we have

\[
A(p^a) = 1 + \chi(p) + \cdots + \chi(p)^a = \begin{cases} 
1, & \chi(p) = 0, \\
1, & \chi(p) = -1, \ a \text{ even}, \\
0, & \chi(p) = -1, \ a \text{ odd.}
\end{cases}
\]

(Note that since \( \chi \) is real, we must have \( \chi(p) \in \{0, 1, -1\} \).) We conclude that \( A(n) \geq 0 \) for all \( n \) and \( A(m^2) \geq 1 \) for all \( m \).

Now let

\[
B(x) = \sum_{n \leq x} \frac{A(n)}{\sqrt{n}}.
\]

Since \( A(m^2) \geq 1 \) and \( A(n) \geq 0 \) for \( n \) not a square, we have \( B(x) \geq \sum_{m \leq \sqrt{x}} 1/m \), so that \( B(x) \to \infty \) as \( x \to \infty \). We will now show that

\[
B(x) = 2\sqrt{x}L(1, \chi) + O(1). \tag{7}
\]

Note that if \( L(1, \chi) = 0 \), then \( B(x) = O(1) \), a contradiction. Thus, (7) will imply that \( L(1, \chi) \neq 0 \).

From the definitions of \( A(n) \) and of \( B(x) \), we have

\[
B(x) = \sum_{n \leq x} \sum_{d \mid n} \frac{\chi(d)}{\sqrt{n}} = \sum_{d \leq x} \chi(d) \sum_{d \mid n \leq x} \frac{1}{\sqrt{n}} = \sum_{d \leq x} \chi(d) \sum_{m \leq x/d} \frac{1}{\sqrt{m}} = \sum_{d \leq x} \frac{\chi(d)}{\sqrt{d}} \sum_{m \leq x/d} \frac{1}{\sqrt{m}}.
\]
Let’s work out an approximate formula for the inner sum here. Let
\[
\beta_m = \frac{1}{\sqrt{m}} - \int_m^{m+1} \frac{dt}{\sqrt{t}} = \frac{1}{\sqrt{m}} - 2 \left( \sqrt{m+1} - \sqrt{m} \right)
\]
\[
= \frac{1}{\sqrt{m}} - \frac{2}{\sqrt{m+1} + \sqrt{m}} = \frac{1}{(\sqrt{m+1} + \sqrt{m})^2 \sqrt{m}} = O \left( \frac{1}{m^{3/2}} \right).
\]
Thus, \( \beta := \sum_m \beta_m \) converges and \( \sum_{m>y} \beta_m = O(1/\sqrt{y}) \). We have
\[
\sum_{m \leq y} \frac{1}{\sqrt{m}} = \sum_{m \leq y} \beta_m + \int_1^{[y]+1} \frac{dt}{\sqrt{t}} = \beta + 2\sqrt{[y]} + 1 - 2 \sum_{m \geq y} \beta_m = 2\sqrt{y} + \beta - 2 + O \left( \frac{1}{\sqrt{y}} \right).
\]
In our last formula for \( B(x) \) we consider separately the cases \( d \leq \sqrt{x} \) and \( d > \sqrt{x} \). We have
\[
\sum_{d \leq \sqrt{x}} \sum_{m \leq x/d} \frac{\chi(d)}{\sqrt{d m}} = \sum_{d \leq \sqrt{x}} \frac{\chi(d)}{\sqrt{d}} \left( 2\sqrt{x} \frac{x}{d} + \beta - 2 + O \left( \frac{\sqrt{d}}{x} \right) \right)
\]
\[
= 2\sqrt{x} \sum_{d \leq \sqrt{x}} \frac{\chi(d)}{d} + (\beta - 2) \sum_{d \leq x} \frac{\chi(d)}{\sqrt{d}} + O \left( \frac{1}{\sqrt{x}} \sum_{d \leq \sqrt{x}} 1 \right).
\]
The \( O \)-term is \( O(1) \). The sum in the middle term is \( L(1/2, \chi) + O(1/x^{1/4}) \) (refer to the partial summation argument that \( L(s, \chi) \) converges for \( s > 0 \)). In particular, the middle term is \( O(1) \). And similarly the sum in the first term is \( L(1, \chi) + O(1/\sqrt{x}) \). Thus,
\[
\sum_{d \leq \sqrt{x}} \sum_{m \leq x/d} \frac{\chi(d)}{\sqrt{d m}} = 2\sqrt{x} L(1, \chi) + O(1).
\]
It remains to show that the contribution to \( B(x) \) for \( d > \sqrt{x} \) is just \( O(1) \), for then we would have (7). We have
\[
\sum_{\sqrt{x} < d \leq x} \sum_{m \leq x/d} \frac{\chi(d)}{\sqrt{d m}} = \sum_{m < \sqrt{x}} \frac{1}{\sqrt{m}} \sum_{\sqrt{x} < d \leq x/m} \frac{\chi(d)}{\sqrt{d}}.
\]
The same partial summation argument for the convergence of \( L(1/2, \chi) \) shows that the inner sum here is \( O(1/x^{1/4}) \), so the entire sum is
\[
O \left( \frac{1}{x^{1/4}} \sum_{m < \sqrt{x}} \frac{1}{\sqrt{m}} \right) = O(1).
\]
We have thus proved (7), and this completes our proof of Proposition 2.