18 Dirichlet L-functions, primes in arithmetic progressions

Having proved the prime number theorem, we would like to prove an analogous result for primes in arithmetic progressions. We begin with Dirichlet's theorem on primes in arithmetic progressions, a result that predates the prime number theorem by sixty years.

Theorem 18.1 (Dirichlet 1837)**.** *For all coprime integers a and m there are infinitely many primes $p \equiv a \bmod m$.*

In fact Dirichlet proved more than this. In a sense that we will make precise , he proved that for every fixed modulus m the primes are equidistributed among the residue classes in $(\mathbb{Z}/m\mathbb{Z})^\times$. The equidistribution statement that Dirichlet was able to prove is a bit weaker than one might like, but it is more than enough to establish Theorem 18.1.

Remark 18.2. Many of the standard tools of complex analysis we take for granted were not available to Dirichlet in 1837. Riemann was the first to seriously study $\zeta(s)$ as a function of a complex variable, some twenty years after Dirichlet proved Theorem 18.1. We will work in a more modern setting, but our approach follows the spirit of Dirichlet's proof.

18.1 Infinitely many primes

To motivate Dirichlet's method of proof, let us consider the following (admittedly clumsy) proof that there are infinitely many primes. It is sufficient to show that the Euler product

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

diverges as $s \to 1^+$. Of course we know $\zeta(s)$ has a pole at $s = 1$ (by Theorem 16.3), but let us suppose for the moment that we did not already know this. Taking logarithms yields

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

as $s \to 1^+$, where we have used the asymptotic bounds

$$-\log(1-x) = x + O(x^2) \quad (\text{as } x \to 0) \qquad \text{and} \qquad \sum_p O(p^{-2s}) = O(1) \quad (\mathrm{Re}(s) > 1/2 + \epsilon).$$

We can estimate $\sum_{p \le x} \frac{1}{p}$ via Mertens' second theorem, one of three he proved in [4].

Theorem 18.3 (Mertens 1874)**.** *As $x \to \infty$ we have*

(1) $\displaystyle\sum_{p \le x} \frac{\log p}{p} = \log x + R(x)$, *where* $|R(x)| < 2$.[1]

(2) $\displaystyle\sum_{p \le x} \frac{1}{p} = \log\log x + B + O\Big(\frac{1}{\log x}\Big)$, *where* $B = 0.261497\ldots$ *is Mertens' constant;*

(3) $\displaystyle\sum_{p \le x} \log\Big(1 - \frac{1}{p}\Big) = -\log\log x - \gamma + O\Big(\frac{1}{\log x}\Big)$, *where* $\gamma = 0.577216\ldots$ *is Euler's constant.*

Proof. See Problem Set 9. □

[1] In fact, $R(x) = -B_3 + o(1)$ where $B_3 = 1.332582\ldots$ is an explicit constant.

Thus not only does $\sum p^{-s}$ diverge as $s \to 1^+$, we can say with a fair degree of precision how quickly this happens. We should note, however, that Mertens' estimate is not as strong as the prime number theorem. Indeed, as you will prove on Problem Set 9, the Prime Number Theorem is equivalent to the statement

$$\sum_{p \le x} \frac{1}{p} = \log\log x + B + o\left(\tfrac{1}{\log x}\right),$$

which is (ever so slightly) sharper than Mertens' estimate.[2]

18.2 Dirichlet characters

We now define the notion of a Dirichlet character. Historically, these preceded the notion of a group character; they were introduced by Dirichlet in 1831, well before the notion of an abstract group was in common use.[3] In order to simplify the exposition we will occasionally invoke some standard facts about characters of finite abelian groups that we recall in §18.6.

Definition 18.4. A function $f\colon \mathbb{Z} \to \mathbb{C}$ is called an *arithmetic function*.[4] The function f is *multiplicative* if $f(1) = 1$ and $f(mn) = f(m)f(n)$ for all coprime $m, n \in \mathbb{Z}$; it is *totally multiplicative* (or *completely multiplicative*) if $f(1) = 1$ and $f(mn) = f(m)f(n)$ for all $m, n \in \mathbb{Z}$. For $m \in \mathbb{Z}_{>0}$ we say that f is *m-periodic* if $f(n+m) = f(n)$ for all $n \in \mathbb{Z}$, and we call m the *period* of f it is the least $m > 0$ for which this holds.

Definition 18.5. A *Dirichlet character* is a periodic totally multiplicative arithmetic function $\chi\colon \mathbb{Z} \to \mathbb{C}$.

The image of a Dirichlet character is a finite multiplicatively closed subset of $\mathbb{C}$, hence the union of a finite subgroup of $\mathrm{U}(1)$ and a subset of $\{0\}$. The constant function $\mathbb{1}(n) := 1$ is the *trivial Dirichlet character*; it is the unique Dirichlet character of period 1. Each m-periodic Dirichlet character χ restricts to a group character χ on $(\mathbb{Z}/m\mathbb{Z})^\times$. Conversely, every group character χ of $(\mathbb{Z}/m\mathbb{Z})^\times$ can be extended to a Dirichlet character χ by defining $\chi(n) = 0$ for $n \not\perp m$; this is called *extension by zero*.

Definition 18.6. A *Dirichlet character of modulus m* is an m-periodic Dirichlet character χ that is the extension by zero of a group character on $(\mathbb{Z}/m\mathbb{Z})^\times$; equivalently, an m-periodic Dirichlet character for which $\chi(n) \neq 0 \iff m \perp n$.

Remark 18.7. Some authors only define Dirichlet characters of modulus m, thinking of them as extensions by zero of group characters on $(\mathbb{Z}/m\mathbb{Z})^\times$, in which case every χ has an attached modulus m. But note that the function $\mathbb{Z} \to \mathbb{C}$ given by the extension by zero does not uniquely determine m (see Lemma 18.8 below). Indeed, the unique Dirichlet character of modulus 2 is a Dirichlet character of modulus 2^k for all $k \ge 1$.

The Dirichlet characters of modulus m form a group under pointwise multiplication that is canonically isomorphic to the character group of $(\mathbb{Z}/m\mathbb{Z})^\times$. Not every m-periodic Dirichlet

[2] The error term in the PNT actually implies $\sum_{p \le x} \frac{1}{p} = \log\log x + B + O\left(\frac{1}{x}\right)$, but an $o(\frac{1}{\log x})$ bound is already enough to show $\pi(x) \sim x/\log x$. That the difference between a little-o and a big-O is the difference between proving the PNT and not proving it demonstrates how critical it is to understand error terms.

[3] Galois' seminal paper was rejected that same year; it wasn't published until 12 years after his death.

[4] Many authors restrict the domain of an arithmetic function to $\mathbb{Z}_{\ge 1}$; for the periodic arithmetic functions we are interested in here, this distinction is irrelevant, and it is slightly more natural to work with $\mathbb{Z}$.

character χ is a Dirichlet character of modulus m, since an m-periodic Dirichlet character need not vanish on $n \not\perp m$, but if χ has period m then this holds. More generally, we have the following lemma.

Lemma 18.8. *Let χ be a Dirichlet character of period m. Then χ is a Dirichlet character of modulus m' if and only if $m|m'|m^k$ for some k (which holds in particular for $m' = m$).*

Proof. To prove χ is a a Dirichlet character of modulus m we must show $\chi(n) \neq 0 \Leftrightarrow m \perp n$. Suppose $\chi(n) \neq 0$ with $m \not\perp n$, and let p be a common divisor of m and n. Then $\chi(p) \neq 0$, since $\chi(p)\chi(n/p) = \chi(n) \neq 0$, and for any $r \in \mathbb{Z}$ we have

$$\chi(r)\chi(p) = \chi(rp) = \chi(rp + m) = \chi(r + m/p)\chi(p),$$

which implies $\chi(r) = \chi(r + m/p)$, since $\chi(p) \neq 0$. Thus χ is (m/p)-periodic, but this contradicts the minimality of the period m. Therefore $\chi(n) \neq 0 \Rightarrow m \perp n$.

For any $n \perp m$ we can pick $a = n^e \equiv 1 \bmod m$ so that $\chi(1) = \chi(a) = \chi(n^e) = \chi(n)^e \neq 0$, in which case $\chi(n) \neq 0$. Thus $n \perp m \Rightarrow \chi(n) \neq 0$, so χ is a Dirichlet character of modulus m.

If $m|m'|m^k$, then the prime divisors of m' coincide with those of m. It follows that

$$n \perp m' \iff n \perp m \iff \chi(n) \neq 0,$$

and χ is clearly m'-periodic (since $m|m'$), so χ is a Dirichlet character of modulus m'.

Conversely, if χ is a Dirichlet character of modulus m', then χ is m'-periodic, and therefore $m|m'$, since m is the period of χ. And since χ is a Dirichlet character of modulus m and of modulus m', for each prime p we have

$$p|m \iff \chi(p) = 0 \iff p|m',$$

thus the prime divisors of m and m' coincide and m' must divide some power m^k of m. □

18.2.1 Primitive Dirichlet characters

Given a Dirichlet character χ_1 of modulus m_1 dividing m_2, we can always create a Dirichlet character χ_2 of modulus m_2 by taking the extension by zero of the restriction of χ_1 to $(\mathbb{Z}/m_2\mathbb{Z})^\times$; in other words, let $\chi_2(n) := \chi_1(n)$ for $n \in (\mathbb{Z}/m_2\mathbb{Z})^\times$ and $\chi_2(n) := 0$ otherwise. If m_2 is divisible by a prime p that does not divide m_1, the Dirichlet characters χ_1 and χ_2 will not be the same ($\chi_2(p) = 0 \neq \chi_1(p)$, for example), they will agree on $n \in (\mathbb{Z}/m_2\mathbb{Z})^\times$ but not on $n \in (\mathbb{Z}/m_1\mathbb{Z})^\times$.[5] We can create infinitely many new Dirichlet characters from χ_1 in this way, but they will differ from χ_1 only in a rather trivial sense. We would like to distinguish the Dirichlet characters that arise in this way from those that do not.

Definition 18.9. Let χ_1 and χ_2 be Dirichlet characters of modulus m_1 and m_2, respectively, with $m_1|m_2$. If $\chi_2(n) = \chi_1(n)$ for $n \in (\mathbb{Z}/m_2\mathbb{Z})^\times$ then χ_2 is *induced* by χ_1. A Dirichlet character that is not induced by any character other than itself is *primitive*.

Lemma 18.10. *A Dirichlet character χ_2 of modulus m_2 is induced by a Dirichlet character of modulus $m_1|m_2$ if and only if χ_2 is constant on residue classes in $(\mathbb{Z}/m_2\mathbb{Z})^\times$ that are congruent modulo m_1. When this holds, the Dirichlet character χ_1 of modulus m_1 that induces χ_2 is uniquely determined.*

[5] Note that while $\#(\mathbb{Z}/m_1\mathbb{Z})^\times \leq \#(\mathbb{Z}/m_2\mathbb{Z})^\times$, the set of integers $n \in (\mathbb{Z}/m_1\mathbb{Z})^\times$ (the n coprime to m_1) contains the set of integers $n \in (\mathbb{Z}/m_2\mathbb{Z})^\times$ (the n coprime to m_2) and is usually larger.

Proof. If χ_2 is induced by χ_1 then it must be constant on residue classes in $(\mathbb{Z}/m_2\mathbb{Z})^\times$ that are congruent modulo m_1, since χ_1 is. To prove the converse we first show that the surjective ring homomorphism $\mathbb{Z}/m_2\mathbb{Z} \to \mathbb{Z}/m_1\mathbb{Z}$ given by reduction modulo m_1 induces a surjective homomorphism $\pi\colon (\mathbb{Z}/m_2\mathbb{Z})^\times \to (\mathbb{Z}/m_1\mathbb{Z})^\times$ of unit groups.[6]

Suppose $u_1 \in \mathbb{Z}$ is a unit modulo m_1. Let a be the product of all primes dividing m_2/m_1 but not u_1. Then $u_2 = u_1 + m_1 a$ is not divisible by any prime $p|m_1$ (since u_1 isn't), nor is it divisible by any prime $p|(m_2/m_1)$: by construction, such a p divides exactly one of u_1 and $m_1 a$. Thus u_2 is a unit modulo m_2 that reduces to u_1 modulo m_1 and π is surjective.

If χ_2 is a Dirichlet character of modulus m_2 constant on fibers of π we can define a Dirichlet character χ_1 of modulus m_1 via $\chi_1(n_1) := \chi_2(n_2)$ for $n_1 \in (\mathbb{Z}/m_1\mathbb{Z})^\times$ with $n_2 \in \pi^{-1}(n_1)$ (any such n_2 will do). Thus χ_1 induces χ_2, and if χ_1' also induces χ_2 it must satisfy the same condition $\chi_1(n_1) = \chi_2(n_2)$ that uniquely determines χ_1. □

Definition 18.11. A Dirichlet character χ induced by $\mathbb{1}$ is called *principal* (and is primitive if and only if $\chi = \mathbb{1}$). For $m \in \mathbb{Z}_{>0}$ we use $\mathbb{1}_m$ to denote the principal Dirichlet character of modulus m; it corresponds to the trivial character of $(\mathbb{Z}/m\mathbb{Z})^\times$.

Lemma 18.12. *Let χ be a Dirichlet character of modulus m. Then*

$$\sum_{n \in \mathbb{Z}/m\mathbb{Z}} \chi(n) \neq 0 \quad \iff \quad \chi = \mathbb{1}_m.$$

Proof. We have $\chi(n) = 0$ for $n \notin (\mathbb{Z}/m\mathbb{Z})^\times$, and the sum over $(\mathbb{Z}/m\mathbb{Z})^\times$ is nonzero if and only if χ restricts to the trivial character on $(\mathbb{Z}/m\mathbb{Z})^\times$, by the orthogonality of characters; see Corollary 18.38. □

Note that the principal Dirichlet characters $\mathbb{1}_m$ and $\mathbb{1}_{m'}$ necessarily coincide when $m|m'|m^k$; for example the principal Dirichlet character of modulus 2 (the parity function) is the same as the principal Dirichlet character of modulus 4 (and every power of 2).

Theorem 18.13. *Every Dirichlet character χ is induced by a primitive Dirichlet character $\widetilde{\chi}$ that is uniquely determined by χ.*

Proof. Let us define a partial ordering $\preceq$ on the set of all Dirichlet characters by defining $\chi_1 \preceq \chi_2$ if χ_1 induces χ_2. The relation $\preceq$ is clearly reflexive, and it follows from Lemma 18.10 that it is transitive.

Let χ be a Dirichlet character of period m and consider the set $X = \{\chi' : \chi' \preceq \chi\}$. Each $\chi' \in X$ necessarily has period m' dividing m and there is at most one χ' of period m' for each divisor m' of m, by Lemma 18.10. Thus X is finite, and nonempty (since $\chi \in X$).

Suppose $\chi_1, \chi_2 \in X$ have periods m_1 and m_2, respectively. Then m_1 and m_2 both divide m, as does $m_3 = \gcd(m_1, m_2)$. We have a commutative square of surjective unit group homomorphisms induced by reduction maps:

$$\begin{array}{ccc} (\mathbb{Z}/m\mathbb{Z})^\times & \twoheadrightarrow & (\mathbb{Z}/m_1\mathbb{Z})^\times \\ \downarrow & & \downarrow \\ (\mathbb{Z}/m_2\mathbb{Z})^\times & \twoheadrightarrow & (\mathbb{Z}/m_3\mathbb{Z})^\times. \end{array}$$

[6] In fact, one can show that every surjective homomorphism of finite rings induces a surjective homomorphism of unit groups, but this does not hold in general (consider $\mathbb{Z} \to \mathbb{Z}/5\mathbb{Z}$, for example).

From Lemma 18.10 we know that χ is constant on residue classes in $(\mathbb{Z}/m\mathbb{Z})^\times$ that are congruent modulo either m_1 or m_2, and therefore χ is constant on residue classes in $(\mathbb{Z}/m\mathbb{Z})^\times$ that are congruent modulo m_3, as are χ_1 and χ_2 (which are determined by χ). It follows that there is a unique Dirichlet character χ_3 of modulus m_3 that induces χ, χ_1, and χ_2.

Thus every pair $\chi_1, \chi_2 \in X$ has a lower bound χ_3 under the partial ordering $\preceq$ that is compatible with the total ordering of X by period. This implies that X contains a unique element $\widetilde{\chi}$ that is minimal, both with respect to the partial ordering $\preceq$ and with respect to the total ordering by period; it must be primitive, by the transitivity of $\preceq$. □

Definition 18.14. The *conductor* of a Dirichlet character χ is the period of the unique primitive Dirichlet character $\widetilde{\chi}$ that induces χ.

Corollary 18.15. *For a Dirichlet character χ of modulus m we have $\sum_{n\in\mathbb{Z}/m\mathbb{Z}} \chi(n) \neq 0$ if and only if χ has conductor 1.*

Proof. This follows immediately from Lemma 18.12. □

Corollary 18.16. *Let $M(m)$ denote the set of Dirichlet characters of modulus m, let $X(m)$ denote the set of primitive Dirichlet characters of conductor dividing m, and let $\widehat{G}(m)$ denote the character group of $(\mathbb{Z}/m\mathbb{Z})^\times$. We have canonical bijections*

$$\begin{array}{ccccc} M(m) & \xrightarrow{\sim} & X(m) & \xrightarrow{\sim} & \widehat{G}(m) \\ \chi & \longmapsto & \widetilde{\chi} & \longmapsto & (n \mapsto \widetilde{\chi}(n)). \end{array}$$

Proof. By Theorem 18.13, the map $\chi \to \widetilde{\chi}$ is injective, and it is also surjective: each $\widetilde{\chi} \in X(m)$ induces the character $\chi \in M(m)$ by setting $\chi(n) := \widetilde{\chi}(n)$ for $n \in (\mathbb{Z}/m\mathbb{Z})^\times$ and extending by zero. As previously noted, the map $\chi \to (m \mapsto \chi(m))$ defines a bijection $M \to \widehat{G}(m)$ (a group isomorphism, in fact), and this bijection factors through the map $\chi \mapsto \widetilde{\chi}$, since $\widetilde{\chi}(n) = \chi(n)$ for $n \in (\mathbb{Z}/m\mathbb{Z})^\times$. □

Remark 18.17. Corollary 18.16 implies that we can make $X(m)$ a group by defining $\widetilde{\chi_1}\widetilde{\chi_2} := \widetilde{\chi_1\chi_2}$. Note that $\widetilde{\chi_1\chi_2}$ is **not** the pointwise product of $\widetilde{\chi_1}$ and $\widetilde{\chi_2}$ (which is typically not primitive), it is the unique primitive character that induces the pointwise product.

Example 18.18. 12-periodic Dirichlet characters, ordered by period m and conductor c.

m	c	0	1	2	3	4	5	6	7	8	9	10	11	mod-12	principal	primitive
1	1	1	1	1	1	1	1	1	1	1	1	1	1	no	yes	yes
2	1	0	1	0	1	0	1	0	1	0	1	0	1	no	yes	no
3	1	0	1	1	0	1	1	0	1	1	0	1	1	no	yes	no
3	3	0	1	-1	0	1	-1	0	1	-1	0	1	-1	no	no	yes
4	4	0	1	0	-1	0	1	0	-1	0	1	0	-1	no	no	yes
6	1	0	1	0	0	0	1	0	1	0	0	0	1	yes	yes	no
6	3	0	1	0	0	0	-1	0	1	0	0	0	-1	yes	no	no
12	4	0	1	0	0	0	1	0	-1	0	0	0	-1	yes	no	no
12	12	0	1	0	0	0	-1	0	-1	0	0	0	1	yes	no	yes

The fact that $\chi(n) \in \{0, \pm 1\}$ for all 12-periodic Dirichlet characters χ follows from the fact that the exponent of $(\mathbb{Z}/m\mathbb{Z})^\times$ is 2; thus $(\operatorname{im}\chi) \cap \mathrm{U}(1) \subseteq \mu_2 = \{\pm 1\}$.

18.3 Dirichlet L-functions

Definition 18.19. The *Dirichlet L-function* associated to a Dirichlet character χ is

$$L(s,\chi) := \prod_p (1-\chi(p)p^{-s})^{-1} = \sum_{n\geq 1} \chi(n)n^{-s}.$$

The sum and product converge absolutely for $\operatorname{Re} s > 1$, since $|\chi(n)| \leq 1$, thus $L(s,\chi)$ is holomorphic on $\operatorname{Re}(s) > 1$.

For the trivial Dirichlet character $\mathbb{1}$ we have $L(s,\mathbb{1}) = \zeta(s)$. For the principal character $\mathbb{1}_m$ of modulus m induced by $\mathbb{1}$ we have

$$\zeta(s) = L(s,\mathbb{1}_m)\prod_{p|m}(1-p^{-s})^{-1}.$$

The product on the RHS is finite, hence bounded and nonzero as $s \to 1^+$, so the L-function $L(s,\mathbb{1}_m)$ has a simple pole at $s=1$ with residue

$$\operatorname{res}_{s=1} L(s,\mathbb{1}_m) = \lim_{s\to 1^+} (s-1)\zeta(s)\prod_{p|m}(1-p^{-s}) = \prod_{p|m}(1-p^{-1}) = \frac{\phi(m)}{m}.$$

The L-functions of non-principal Dirichlet characters do not have a pole at $s=1$.

Proposition 18.20. *Let χ be a non-principal Dirichlet character of modulus m. Then $L(s,\chi)$ extends to a holomorphic function on* $\operatorname{Re} s > 0$.

Proof. Define the function $T\colon \mathbb{R}_{\geq 0} \to \mathbb{C}$ by

$$T(x) := \sum_{0<n\leq x} \chi(n).$$

For any $x \in \mathbb{R}_{\geq 0}$ Lemma 18.15 implies

$$T(x+m) - T(x) = \sum_{x<n\leq x+m} \chi(n) = \sum_{n\in\mathbb{Z}/m\mathbb{Z}} \chi(n) = 0,$$

since χ is non-principal. Thus $T(x)$ is periodic modulo m and therefore bounded.

Writing $L(s,\chi)$ as a Stieltjes integral (see §18.5) and integrating by parts yields

$$\begin{aligned} L(s,\chi) &= \sum_{n\geq 1} \chi(n)n^{-s} \\ &= \int_0^\infty x^{-s}dT(x) \\ &= x^{-s}T(x)\Big|_0^\infty - \int_0^\infty T(x)d(x^{-s}) \\ &= 0 - \int_0^\infty T(x)(-sx^{-s-1})dx \\ &= s\int_0^\infty T(x)x^{-s-1}dx. \end{aligned}$$

The RHS is holomorphic on $\operatorname{Re} s > 0$, since it is the limit of the uniformly converging sequence of functions $\phi_n(s) := s\int_0^n T(x)x^{-s-1}dx$ (here we use the fact that $T(x)$ is bounded), and is thus the analytic continuation of $L(x,\chi)$ to $\operatorname{Re}(s) > 0$. □

Remark 18.21. In fact, $L(s,\chi)$ extends to a holomorphic function on $\mathbb{C}$ whenever χ is non-principal.

18.4 Primes in arithmetic progressions

We now return to our goal of proving Dirichlet's theorem on primes in arithmetic progressions. It suffices to show that for any $a \perp m$ the sum

$$\sum_{p \equiv a \bmod m} p^{-s}$$

is unbounded as $s \to 1^+$. To convert this to a sum over all primes we use Proposition 18.37 to construct the indicator function

$$\frac{1}{\phi(m)} \sum_{\chi \in X(m)} \chi(p/a) = \begin{cases} 1 & \text{if } p \equiv a \bmod m, \\ 0 & \text{otherwise} \end{cases}$$

where p/a is computed modulo m and χ ranges over primitive Dirichlet characters of conductor dividing m, which we identify with the character group of $(\mathbb{Z}/m\mathbb{Z})^\times$ via Corollary 18.16.

As $s \to 1^+$ we have

$$\begin{aligned} \sum_{p \equiv a \bmod m} p^{-s} &= \sum_p p^{-s} \frac{1}{\phi(m)} \sum_{\chi \in X(m)} \chi(p/a) \\ &= \sum_{\chi \in X(m)} \frac{\chi(1/a)}{\phi(m)} \sum_p \chi(p) p^{-s} \\ &= \sum_{\chi \in X(m)} \frac{\chi(1/a)}{\phi(m)} \Big(\log L(s,\chi) + O(1) \Big) \\ &= \frac{\log \zeta(s)}{\phi(m)} + \sum_{\substack{\chi \in X(m) \\ \chi \neq \mathbb{1}}} \frac{\chi(1/a)}{\phi(m)} \log L(s,\chi) + O(1). \end{aligned}$$

We now make the key claim that so long as χ is not principal, we have

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

This implies that $\log L(s,\chi) = O(1)$ as $s \to 1^+$ and therefore

$$\sum_{p \equiv a \bmod m} p^{-s} = \frac{\log \zeta(s)}{\phi(m)} + O(1)$$

is unbounded as $s \to 1^+$, since $\zeta(s)$ is. Moreover, Mertens' second theorem implies

$$\sum_{\substack{p \le x \\ p \equiv a \bmod m}} \frac{1}{p} \sim \frac{\log \log x}{\phi(m)}.$$

which proves that there are infinitely many primes $p \equiv a \bmod m$. We can makes this statement more precise by using the notion of *Dirichlet density*.

Definition 18.22. The *Dirichlet density* of a set of primes S is given by

$$d(S) := \lim_{s \to 1^+} \frac{\sum_{p \in S} p^{-s}}{\sum_p p^{-s}},$$

defined whenever this limit exists (one can also define notions of lower and upper Dirichlet density using lim inf and lim sup which both agree with $d(S)$ whenever it is defined). This definition differs from the more common notion of *natural density*

$$\delta(S) := \lim_{x\to\infty} \frac{\#\{p \le x : p \in S\}}{\#\{p \le x\}}.$$

The Dirichlet density of the set $S = \{p \equiv a \bmod m\}$ is

$$d(S) = \lim_{s\to 1^+} \frac{\sum_{p\in S} p^{-s}}{\sum_p p^{-s}} = \lim_{s\to 1^+} \frac{\log \zeta(s)/\phi(m)}{\log \zeta(s)} = \frac{1}{\phi(m)}.$$

The primes are thus equidistributed modulo m in the sense that for all $a \perp m$ we have

$$\sum_{\substack{p\le x\\ p\equiv a \bmod m}} \frac{1}{p} \sim \frac{1}{\phi(m)} \sum_{p\le x} \frac{1}{p} \sim \frac{1}{\phi(m)} \log\log x.$$

We should note that this statement is weaker than the prime number theorem for arithmetic progressions, which states that

$$\pi(x; m, a) \sim \frac{1}{\phi(m)} \pi(x),$$

where $\pi(x; m, a)$ counts the primes $p \le x$ for which $p \equiv a \bmod m$ (see Problem Set 9).

18.5 Stieltjes integrals

For the benefit of those who have not seen them before, we recall a few facts about Stieltjes integrals (also called Riemann-Stieltjes integrals), taken from [1, Ch. 7]. These generalize the Riemann integral but are less general than the Lebesgue integral; they provide a handy way for converting sums to integrals that is often used in analytic number theory.

Definition 18.23. Let f and g be (real or complex valued) functions defined on a nonempty real interval $[a, b]$. For any partition $P = (x_0, \dots, x_n)$ of $[a, b]$ and sequence $T = (t_1, \dots, t_n)$ with $t_k \in [x_{k-1}, x_k]$, we define the *Riemann-Stieltjes sum*

$$S(P, T, f, g) := \sum_{k=1}^{n} f(t_k)\Big(g(x_k) - g(x_{k-1})\Big)$$

We say that f is *Riemann-Stieltjes integrable with respect to* g and write $f \in S(g)$ if there is a (real or complex) number S such that for every $\epsilon > 0$ there is a partition P_ϵ of $[a, b]$ such that for every refinement $P = (x_0, \dots, x_n)$ of P_ϵ and every sequence $T = (t_1, \dots, t_n)$ with $t_k \in [x_{k-1}, x_k]$ we have $|S(P, T, f, g) - S| < \epsilon$.[7]

When such an S exists it is necessarily unique and we denote it by $\int_a^b f\, dg$, the *Riemann-Stieltjes integral of* f *with respect to* g. Improper Riemann-Stieltjes integrals are then defined as limits

$$\int_a^\infty f\, dg := \lim_{b\to\infty} \int_a^b f\, dg$$

(and similarly for the lower limit), and we define $\int_b^a f\, dg = -\int_a^b f\, dg$ and $\int_a^a f\, dg = 0$.

[7] This definition (due to Pollard) is more general than that originally given by Stieltjes but is now standard.

Taking $g(x) = x$ yields the Riemann integral. The Riemann-Stieltjes integral satisfies the usual properties of linearity, summability, and integration by parts.

Proposition 18.24. *Let f, g, and h be functions on $[a, b]$ and let c_1 and c_2 be constants. The following hold:*

- *If $f, g \in S(h)$ then $\int_a^b (c_1 f + c_2 g)\, dh = c_1 \int_a^b f\, dh + c_2 \int_a^b g\, dh$.*
- *If $f \in S(g), S(h)$ then $\int_a^b f\, d(c_1 g + c_2 h) = c_i \int_a^b f\, dg + c_2 \int_a^b f\, dh$.*
- *If $f \in S(g)$ then for any $c \in [a, b]$ we have $\int_a^b f\, dg = \int_a^c f\, dg + \int_c^b f\, dg$.*
- *If $f \in S(g)$ then $g \in S(f)$ and $\int_a^b f\, dg + \int_a^b g\, df = f(b)g(b) - f(a)g(a)$.*
- *If $f = f_1 + if_2$ and $g = g_1 + ig_2$ with $f_1, f_2 \in S(g_1), S(g_2)$ then*

$$\int_a^b f\, dg = \left(\int_a^b f_1\, dg_1 - \int_a^b f_2\, dg_2 \right) + i \left(\int_a^b f_2\, dg_1 + \int_a^b f_1\, dg_2 \right).$$

Proof. See [1, Thm. 7.2-7,7.50]. □

The last identity allows us to reduce complex-valued integrals to real-valued integrals. The following proposition allows us to reduce Stieltjes integrals to Riemann integrals.

Proposition 18.25. *Let f and g be real-valued functions on $[a, b]$ and suppose g has a continuous derivative g' on $[a, b]$. Then*

$$\int_a^b f\, dg = \int_a^b f(x) g'(x) dx.$$

Proof. See [1, Thm. 7.8]. □

A key advantage of the Stieltjes integral $\int_a^b f\, dg$ is that neither the integrand f nor the integrator g is required to be continuous. It suffices for f and g to be of bounded variation and not share any discontinuities (and they can even share certain discontinuities, see Theorem 18.27).

Definition 18.26. Let f be a (real or complex valued) function defined on a nonempty real interval $[a, b]$. Then f is of *bounded variation* if there exists a real number M such that

$$\sum_{i=0}^{n-1} |f(x_{i+1}) - f(x_i)| < M$$

for every partition $P = (x_0, \ldots, x_n)$ of $[a, b]$. If f has a continuous derivative f' on $[a, b]$ this is equivalent to requiring $\int_a^b |f'(x)| dx < \infty$. Every piecewise monotone function is of bounded variation. In particular, any step function with finitely many discontinuities on $[a, b]$ is of bounded variation.

Theorem 18.27. *Let f and g be functions on $[a, b]$ of bounded variation such that for every $c \in [a, b]$ the function f is continuous from the left at c and the function g is continuous from the right at c. Then $\int_a^b f\, dg$ and $\int_a^b g\, df$ both exist.*

Proof. See [2, Thm. 3.7]. □

Corollary 18.28. *Let f and g be functions on $[a,b]$ such that f and g are not both discontinuous from the left or from the right at integers $n \in [a,b]$, and let $G(x) = \sum_{a<n\le x} g(n)$. Then*

$$\sum_{a<n\le b} f(n)g(n) = \int_a^b f(x)\,dG(x).$$

In particular, the integral on the RHS always exists.

Proof. See [1, Thm. 7.11]. □

As an example of using Stieltjes integrals, let us derive an asymptotic estimate for the harmonic sum

$$H(x) := \sum_{1\le n\le x} \frac{1}{n}.$$

Theorem 18.29. *For $x \in \mathbb{R}_{\ge 1}$, as $x \to \infty$ we have*

$$H(x) = \log x + \gamma + O\left(\tfrac{1}{x}\right)$$

where $\gamma = \lim_{x\to\infty}(H(x) - \log x)$ =0.577216... is Euler's constant.

Proof. Let $[t]$ denote the greatest integer function. Applying Corollary 18.28 with $g(t) = 1$ and $G(t) = \sum_{1\le n\le t} 1 = [t]$, we have

$$\begin{aligned}
H(x) = \sum_{1\le n\le x} \frac{1}{n} &= \int_{1^-}^x \frac{1}{t}\,d[t] \\
&= \frac{[t]}{t}\Big|_{1^-}^x - \int_{1^-}^x [t]\,d\frac{1}{t} \\
&= \frac{[x]}{x} + \int_{1^-}^x \frac{[t]}{t^2}\,dt \\
&= \frac{[x]}{x} + \int_{1^-}^x \frac{1}{t}\,dt - \int_{1^-}^x \frac{t-[t]}{t^2}\,dt \\
&= \frac{[x]}{x} + \log x - \int_{1^-}^x \frac{t-[t]}{t^2}\,dt,
\end{aligned}$$

where we used integration by parts in the second line and applied Proposition 18.25 to get the third line. Now let $\gamma = 1 - \int_{1^-}^\infty (t-[t])/t^2\,dt$. Then

$$\begin{aligned}
H(x) &= \frac{[x]}{x} + \log x - 1 + \gamma + \int_x^\infty \frac{t-[t]}{t^2}\,dt \\
&= \log x + \gamma + \left(\frac{[x]-x}{x} + \int_x^\infty \frac{t-[t]}{t^2}\,dt\right). \qquad (2)
\end{aligned}$$

Both summands in the parenthesized quantity in (2) are clearly $O\left(\frac{1}{x}\right)$; thus

$$\gamma = \lim_{x\to\infty} \left(H(x) - \log x\right),$$

and the theorem follows. □

18.6 A quick recap of the character theory of finite abelian groups

In this section we recall some standard results on characters of finite abelian groups.

Definition 18.30. A *character* of a group G is a homomorphism $\chi\colon G \to \mathrm{U}(1)$.[8] The *character group* (or *dual group*) of G is the abelian group

$$\widehat{G} := \mathrm{Hom}(G, \mathrm{U}(1))$$

under pointwise multiplication: $(\chi_1\chi_2)(g) := \chi_1(g)\chi_2(g)$. The inverse of χ is given by complex conjugation: $\chi^{-1}(g) = \overline{\chi}(g) := \overline{\chi(g)}$. The identity element of $\widehat{G}$ is the *trivial character* $g \mapsto 1$.

Remark 18.31. This definition generalizes to locally compact abelian groups G, in which case each character $\chi\colon G \to \mathrm{U}(1)$ is a homomorphism of topological groups and the dual group $\widehat{G}$ is locally compact under the *compact-open topology* which has a basis of neighborhoods of the identity the sets $U(C,V) := \{\chi \in \widehat{G} : \chi(C) \subseteq V\}$, where C ranges over compact subsets of G and V ranges over open neighborhoods of the identity in $\mathrm{U}(1)$. The locally compact group $\widehat{G}$ is called the *Pontryagin dual* of G.[9] When G is finite it necessarily has the discrete topology (since it must be Hausdorff), every homomorphism $G \to \mathrm{U}(1)$ is automatically continuous, and the compact-open topology on $\widehat{G}$ is also discrete.

Proposition 18.32. *Let G be a finite abelian group with character group $\widehat{G}$. Then $G \simeq \widehat{G}$.*

Proof. As a finite abelian group we can write G as a direct product of cyclic groups

$$G = \langle g_1\rangle \times \cdots \times \langle g_r\rangle \simeq \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_r\mathbb{Z},$$

with $n_i = |g_i|$, and each $g \in G$ can be uniquely written as $g = \prod_i g_i^{e_i}$ with $0 \le e_i < n_1$. Now fix (not necessarily distinct) primitive n_i-th roots of unity $\alpha_i \in \mathrm{U}(1)$ and for $1 \le i \le r$ define $\chi_i \in \widehat{G}$ via

$$\chi_i(g_j) := \begin{cases} \alpha_i & \text{if } i = j, \\ 1 & \text{if } i \ne j. \end{cases}$$

Then $|\chi_i| = |\alpha_i| = n_i$, and each $\chi \in \widehat{G}$ can be written uniquely as $\prod_i \chi_i^{e_i}$ with $0 \le e_i < n_i$, where $\chi(g_i) = \alpha_i^{e_i}$ (because a character is completely determined by its values on generators and the χ_i are clearly orthogonal). Therefore

$$\widehat{G} = \langle \chi_1\rangle \times \cdots \times \langle \chi_n\rangle \simeq \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_r\mathbb{Z} \simeq G. \qquad \square$$

Corollary 18.33. *Let G be a finite abelian group. Then $g \in G$ is the identity if and only if $\chi(g) = 1$ for all $\chi \in \widehat{G}$ and $\chi \in \widehat{G}$ is the identity if and only if $\chi(g) = 1$ for all $g \in G$.*

The isomorphism in Proposition 18.33 is not canonical. Indeed, there are $\#\mathrm{Aut}(G)$ distinct ways to choose the α_i used to construct the isomorphism $G \simeq \widehat{G}$. But there is a canonical isomorphism from G to the character group of $\widehat{G}$, the *double dual* of G.

[8] Some authors call these *unitary characters*, allowing characters to have image in $\mathbb{C}^\times$. When G is finite every character is a unitary character, so this distinction won't concern us.

[9] Some authors define the topology on the Pontryagin duality using uniform convergence on compact sets; for topological groups this is equivalent to the compact-open topology. The unitary group $\mathrm{U}(1) \simeq \mathbb{R}/\mathbb{Z}$ is also referred to as the 1-*torus* or *circle group* and may be denoted $\mathbb{T}$ or S^1 and viewed as an additive group.

Corollary 18.34. *Let G be a finite abelian group. The evaluation map*

$$g \mapsto (\chi \mapsto \chi(g))$$

is a canonical isomorphism from G to its double dual.

Proof. It is clear that the map above is a homomorphism, and Proposition 18.33 implies that G is isomorphic to its dual group $\widehat{G}$, which is in turn isomorphic to its dual group, the double dual of G). So it suffices to show the map is injective, which follows from Corollary 18.34: if g lies in the kernel then $\chi(g) = 1$ for all $\chi \in \widehat{G}$ and $g = 1_G$, by Corollary 18.34, □

Corollary 18.35 allows us to view G as the character group of $\widehat{G}$ by defining $g(\chi) := \chi(g)$.

Remark 18.35. Corollary 18.35 is a special case of *Pontryagin duality*, which applies to any locally compact abelian group G. For infinite groups, G and $\widehat{G}$ need not be isomorphic; for example, the character group of $\mathbb{Z}$ is isomorphic to $\mathrm{U}(1)$ (but in some cases they are, as when G is $\mathbb{R}$ or $\mathbb{Q}_p$, or any local field, see [3, XV, Lemma 2.2.1]). But the canonical isomorphism between G and its double dual always holds.

This is analogous to the situation with vector spaces: a finite dimensional vector space (which may be an infinite abelian group) is non-canonically isomorphic to its dual space but canonically isomorphic to its double dual via the evaluation map. We should note that for a locally compact topological vector space V over a field k, the Pontryagin dual is not the same thing as the vector space dual: the Pontryagin dual corresponds to $\mathrm{Hom}(V, \mathrm{U}(1))$ (morphisms of locally compact groups) while the vector space dual corresponds to $\mathrm{Hom}_k(V, k)$ (morphisms of topological k-vector spaces). For example, the vector space dual of $\mathbb{Q}$ is isomorphic to $\mathbb{Q}$, as is its double dual, but the Pontryagin dual of $\mathbb{Q}$ is uncountable (thus not isomorphic to $\mathbb{Q}$), even though the Pontryagin double dual is isomorphic to $\mathbb{Q}$.

Proposition 18.36. *Let G be a finite abelian group. For all $g_1, g_2 \in G$ we have*

$$\langle g_1, g_2\rangle := \frac{1}{\#G}\sum_{\chi\in\widehat{G}} \chi(g_1)\overline{\chi(g_2)} = \begin{cases} 1 & \text{if } g_1 = g_2, \\ 0 & \text{if } g_1 \neq g_2, \end{cases}$$

and for all $\chi_1, \chi_2 \in \widehat{G}$ we have

$$\langle \chi_1, \chi_2\rangle := \frac{1}{\#G}\sum_{g\in G} \chi_1(g)\overline{\chi_2(g)} = \begin{cases} 1 & \text{if } \chi_1 = \chi_2, \\ 0 & \text{if } \chi_1 \neq \chi_2. \end{cases}$$

Proof. By duality it suffices to consider $\langle g_1, g_2\rangle$. If $g_1 = g_2$ then $\chi(g_1)\overline{\chi(g_2)} = 1$ for all $\chi \in \widehat{G}$ and $\langle g_1, g_2\rangle = \#\widehat{G}/\#G = 1$. If $g_1 \neq g_2$ then by Corollary 18.34 there exists $\lambda \in \widehat{G}$ for which $\alpha := \lambda(g_1)\overline{\lambda(g_2)} = \lambda(g_1 g_2^{-1}) \neq 1$. We then have

$$\alpha\langle g_1, g_2\rangle = \frac{1}{\#G}\sum_{\chi\in\widehat{G}} (\lambda\chi)(g_1)\overline{(\lambda\chi)(g_2)} = \frac{1}{\#G}\sum_{\chi\in\lambda\widehat{G}} \chi(g_1)\overline{\chi(g_2)} = \langle g_1, g_2\rangle,$$

which implies $\langle g_1, g_2\rangle = 0$, since $\alpha \neq 1$. □

Corollary 18.37. *For $\chi \in \widehat{G}$ we have $\sum_{g\in G}\chi(g) \neq 0$ if and only χ is the trivial character. For $g \in G$ we have $\sum_{\chi\in\widehat{G}}\chi(g) \neq 0$ if and only if g is the trivial element.*

Remark 18.38. The *orthogonality of characters* given by Proposition 18.37 is a special case of the orthogonality of characters one encounters in Fourier analysis on compact groups; since G is finite, the weighted sum over G amounts to integrating against its Haar measure (the counting measure μ normalized so that $\mu(G) = 1$).

We conclude our discussion of character groups with a theorem analogous to the fundamental theorem of Galois theory.

Proposition 18.39. *Let G be a finite abelian group. There is an inclusion reversing bijection φ between subgroups H of G and subgroups K of $\widehat{G}$ defined by*

$$\varphi(H) \coloneqq \{\chi \in \widehat{G} : \chi(h) = 1 \text{ for all } h \in H\}.$$

The inverse bijection ϕ is given by

$$\phi(K) \coloneqq \{g \in G : \chi(g) = 1 \text{ for all } \chi \in K\},$$

and $\widehat{H} \simeq \widehat{G}/\varphi(H)$ and $\widehat{K} \simeq G/\phi(K)$; in particular, $\#H = [\widehat{G}:\varphi(H)]$ and $\#K = [G:\phi(K)]$.

Proof. Its clear from the definitions that φ and ϕ are inclusion reversing. Let H be a subgroup of G. The group $K = \varphi(H)$ consists of the characters of G whose kernel contains H. It is clear that $H' \coloneqq \phi(K)$ contains H, since it is equal to the intersection of these kernels, and by duality it is similarly clear that $K' \coloneqq \varphi(H')$ contains K. We then have $H \subseteq H'$ and $\varphi(H) \subseteq \varphi(H')$, but φ is inclusion reversing so $H = H'$; thus $\phi \circ \varphi$ is the identity map, and by duality, so is $\varphi \circ \phi$.

The restriction map $\widehat{G} \to \widehat{H}$ defined by $\chi \mapsto \chi_{|H}$ is a group homomorphism with kernel $K = \varphi(H)$. It is surjective because if we let $\chi_1 \coloneqq 1_{\widehat{G}}$ then we have

$$\#H\#K = \sum_{h\in H}\sum_{\chi\in K}\chi(h) = \sum_{h\in H}\sum_{\chi\in K}\chi(h)\overline{\chi_1(h)} = \sum_{g\in G}\sum_{\chi\in K}\chi(g)\overline{\chi_1(g)} = \#G,$$

by Proposition 18.37, and therefore $\#\widehat{H}\#K = \#\widehat{G}$ (by Proposition 18.33). It follows that $\widehat{H} \simeq \widehat{G}/\varphi(H)$, and by duality, $K \simeq G/\phi(K)$. □

References

[1] Tom Apostol, *Mathematical analysis*, 2nd edition, Addison-Wesley, 1974.

[2] Paul Bateman and Harold Diamond, *Analytic number theory: An introductory course*, World Scientific, 2004.

[3] J.W.S. Cassels and A. Fröhlich, *Algebraic number theory*, 2nd edition, London Mathematical Society, 2010.

[4] Franz Mertens, *Ein Beitrag zur analytischen Zahlentheorie*, J. Reine Angew. Math., **78** (1874), 46–62.