Spring 5-1-2013

Dirichlet's Theorem and Applications

Nicholas Stanford
University of Connecticut - Storrs, mistersoccer@yahoo.com

Dirichlet's Theorem and Applications

Nicholas Stanford

B.S. Mathematics

B.A. Music

An Undergraduate Honors Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Bachelor of Science

at the

University of Connecticut

May 2013

APPROVAL PAGE

Bachelor of Science Honors Thesis

Dirichlet's Theorem and Applications

Presented by

Nicholas Stanford, B.S. Math., B.A. Music

Honors Major Advisor _____
William Abikoff

Honors Thesis Advisor _____
Keith Conrad

University of Connecticut

May 2013

ACKNOWLEDGMENTS

This thesis could not have been completed without the help of my thesis advisor, Professor Keith Conrad. Prof. Conrad offered to be my advisor during finals week of the fall semester of my senior year. I had not yet done any research toward my thesis and had taken just one class with Prof. Conrad. Yet he believed that I would be able to accomplish this and was willing to work tirelessly with me to make it happen.

Prof. Conrad has been both relentless and patient while guiding me through this process. I can confidently say that very few of my peers had their theses scrutinized in their entirety as many times as mine was. His attention to detail, ensuring that every line of this thesis was as well written as possible, was astounding. Furthermore, he pushed me to be exhaustive in my thinking. Specific direction was given when needed, but I was also allowed to consider the issues on my own. Perhaps I mistook it as simply being sarcastic humor, but Prof. Conrad never seemed to grow upset with me or the progress of my research (or lack thereof). Instead, he was understanding. He set feasible milestones for me to accomplish each week that ultimately led to the final product presented here. I sincerely thank Prof. Conrad for his help in the completion of this thesis, the completion of my University of Connecticut Honors requirements, and my acceptance into graduate school.

I would like to acknowledge the entire Mathematics Department at UConn, including both the faculty and my peers. They have helped me grow as a student of math since my first day as an undergraduate. My professors provided me with

challenging material to learn, but also ensured that they were available to help if necessary. Indeed, I have found that office hours are quite the valuable resource. The time spent collaborating with my peers, however, was just as beneficial. We discover new ways of thinking and strive to make our own work better when we work with others. The last four years spent learning with these individuals has prepared me for the next stage of my education in math.

Individuals outside of the Mathematics Department have also influenced me heavily. Music may not be my career, but it is my passion. The music faculty embraced me as a student and was encouraging of my studies. In particular, I would like to thank my piano teacher, Irma Vallecillo. She has been my most supportive teacher during my time at UConn. Irma not only taught me how to better express myself when playing piano, but became a mentor and a close friend. To the rest of the University of Connecticut community, I thank you.

I would be remiss if I did not recognize my friends and family here. While they may not have played a direct role in the writing of this thesis, they have been invaluable nonetheless. My roommates ensured that I actually left my room occasionally for such trivial activities as dinner or a game of frisbee. They kept me happy and made my time at UConn worthwhile on a personal level. I will miss each of them deeply when we go our separate ways in the fall. My family has loved and supported me for as long as I can recall. Never has my mom questioned my desire to continue my education instead of finding a job. Instead, she has done all that is in her capacity in order to allow my sister and me to excel in our desired fields. She is an inspiration to me. To my sister, thank you for putting up with me over the years; I really do appreciate having you around. To my father, I still think of you often and am inspired by you. Thank you once again to all of those who have helped me along my journey.

Dirichlet's Theorem and Applications

Nicholas Stanford, B.S., B.A.

University of Connecticut, May 2013

ABSTRACT

Dirichlet's theorem states that there exist an infinite number of primes in an arithmetic progression $a + mk$ when a and m are relatively prime and k runs over the positive integers. While a few special cases of Dirichlet's theorem, such as the arithmetic progression $2 + 3k$, can be settled by elementary methods, the proof of the general case is much more involved. Analysis of the Riemann zeta-function and Dirichlet L-functions is used.

The proof of Dirichlet's theorem suggests a method for defining a notion of density of a set of primes, called its Dirichlet density, and the primes of the form $a+mk$ have a Dirichlet density $1/\varphi(m)$, which is independent of a. While the definition of Dirichlet density is not intuitive, it is easier to compute than a more natural concept of density, and the two notions of density turn out to be equal when they both exist.

Dirichlet's theorem is often used to show a prime number exists satisfying a particular congruence condition while avoiding a finite set of "bad" primes. For example, it allows us to find the density of the set of primes p such that a given nonzero integer a is or is not a square mod p. More generally, it lets us find the density of the set of primes at which a finite set of integers have prescribed Legendre symbol values.

Contents

Chapter 1

Introduction

Dirichlet's theorem states that for two relatively prime integers a and m there exist infinitely many primes $p \equiv a \bmod m$. Table 1.0.1 below gives supporting numerical data for the case of primes $p \equiv a \bmod 9$ when $(a, 9) = 1$. We tabulate such $p \leq N$ as N runs through powers of 10.

N	a :	1	2	4	5	7	8
10^2		3	5	3	4	5	4
10^3		27	30	27	28	26	29
10^4		203	207	206	209	202	201
10^5		1592	1604	1601	1604	1591	1599
10^6		13063	13099	13070	13068	13098	13099

TABLE 1.0.1: Number of primes $p \leq N$ satisfying $p \equiv a \bmod 9$.

Dirichlet's theorem says more: the proportion of primes $p \equiv a \bmod m$ is $1/\varphi(m)$ when $\gcd(a, m) = 1$. For example, in Table 1.0.2 the proportions are all getting close to $1/\varphi(9) = 1/6 = .166\ldots$.

The background that led to Dirichlet's theorem was Legendre's unsuccessful at-

N	a :	1	2	4	5	7	8
10^2		.1200	.2000	.1200	.1600	.2000	.1600
10^3		.1607	.1786	.1607	.1607	.1548	.1726
10^4		.1652	.1684	.1676	.1701	.1644	.1635
10^5		.1660	.1672	.1669	.1672	.1659	.1667
10^6		.1664	.1669	.1665	.1650	.1669	.1669

TABLE 1.0.2: Proportion of primes $p \leq N$ satisfying $p \equiv a \bmod 9$.

tempt to prove quadratic reciprocity. He broke up quadratic reciprocity into eight cases, and to prove some of them he was led to conjecture that there are infinitely many primes in any arithmetic progression $a + mk$ where a and m are relatively prime and k runs over the natural numbers [9, pp. 6–8]. The first proof of Legendre's conjecture was given by Dirichlet in 1837, and his method was inspired by Euler's analytic proof from 1737 that there are infinitely many primes: Euler showed the series $\sum_p 1/p$ diverges, where p runs over the primes, and Dirichlet combined analysis with group theory to show the series $\sum_{p \equiv a \bmod m} 1/p$ diverges when $\gcd(a, m) = 1$, so the set $\{p \equiv a \bmod m\}$ is infinite. Our treatment of Dirichlet's theorem will use complex analysis, but Dirichlet's own proof did not. It preceded Riemann's work on complex analysis in number theory (for the zeta-function) by about twenty years.

To prove Dirichlet's theorem, in Chapter 2 we will introduce characters and discuss some essential results from complex analysis. These results will then be used in Chapter 3 in order to prove theorems about L-functions. These L-functions play a critical role in proving Dirichlet's theorem in Chapter 4. Lastly, in Chapter 5 we will look at applications of Dirichlet's theorem to the behavior of Legendre symbols.

Chapter 2

Characters and Complex Analysis

2.1 Characters

Definition 2.1.1. For a finite abelian group G, a *character* χ on G is a homomorphism from G to the unit circle S^1.

Example 2.1.2. If $G = \mathbf{Z}/4$, the functions $\chi(a \bmod 4) = i^a$ and $\chi(a \bmod 4) = (-1)^a$ are both characters of G.

Multiplication of two characters χ and ψ on G is defined by $(\chi\psi)(g) = \chi(g)\psi(g)$ for $g \in G$.

Theorem 2.1.3. *The set of all characters $\chi : G \to S^1$ forms a multiplicative group.*

Proof. Since the product of two elements in S^1 is still in S^1, and S^1 is abelian, the product of two characters still maps elements of G to S^1 and is a character. The identity character is the character χ_1 for which $\chi_1(g) = 1$ for all $g \in G$. We

call this the trivial character. The inverse of a character χ is $\overline{\chi}$, where we define $\overline{\chi}(g) = \overline{\chi(g)} = \chi(g)^{-1}$ for all $g \in G$. $\square$

We will denote the group of characters on G by $\widehat{G}$. If G is nontrivial, the construction of a nontrivial character on G is based on extending characters from subgroups to the whole group. Let us start with cyclic groups.

Theorem 2.1.4. *Suppose G is a finite cyclic group with order n. There are n distinct characters of G, and each one is uniquely determined by its value on a generator g_0 by sending g_0 to each of the n^{th} roots of unity in $\mathbf{C}$. Moreover, $\widehat{G}$ is cyclic.*

Proof. For any character χ of G and $g \in G$, $\chi(g)^n = \chi(g^n) = \chi(1) = 1$, so $\chi(g)$ is a root of unity with order dividing n. For any $g \in G$, we can write $g = {g_0}^\ell$ for some $\ell \in \mathbf{Z}$. Therefore, $\chi(g) = \chi({g_0}^\ell) = \chi(g_0)^\ell$, so the values that χ takes on G are completely determined by the value that χ takes at g_0.

We may assign the value of $\chi(g_0)$ to be any of the n^{th} roots of 1 in $\mathbf{C}$: for $\zeta \in \mathbf{C}$ with $\zeta^n = 1$, set $\chi({g_0}^\ell) = \zeta^\ell$ for all $\ell \in \mathbf{Z}$. This is well-defined: if ${g_0}^\ell = {g_0}^m$ then $\ell \equiv m \bmod n$, so $\zeta^\ell = \zeta^m$. It is easy to check that χ is a homomorphism, so χ is a character of G and $\chi(g_0) = \zeta$. Since there are n choices for the value of $\chi(g_0)$ and χ is completely determined by this value, there are n distinct characters on G. Thus $|\widehat{G}| = n$.

To show that $\widehat{G}$ is cyclic, for $0 \leq r < n$, let ψ_r be the character of G where $\psi_r(g_0) = e^{2\pi i r/n}$. Then $\psi_r(g_0) = \psi_1(g_0)^r$, so for all $\ell \in \mathbf{Z}$

$$\psi_r(g_0^\ell) = \psi_r(g_0)^\ell = e^{2\pi i r\ell/n} = \psi_1(g_0^\ell)^r.$$

This says $\psi_r(g) = \psi_1(g)^r$ for all $g \in G$. Thus $\psi_r = \psi_1^r$, so $\widehat{G}$ is generated by ψ_1. $\square$

Theorem 2.1.5. *For $g_1, g_2 \in G$ where $g_1 \neq g_2$, there exists a character χ of G such that $\chi(g_1) \neq \chi(g_2)$.*

Proof. To prove this, it will suffice to show that for every $g \in G$ where $g \neq 1$ there exists a χ such that $\chi(g) \neq 1$. Taking $g = g_1 g_2^{-1}$ then yields the desired result.

Suppose G is cyclic of order n with generator g_0. Every $g \in G$ with $g \neq 1$ can be written as $g = {g_0}^\ell$ with $1 \leq \ell < n$. By the previous theorem, we can choose $\chi \in \widehat{G}$ such that $\chi(g_0) = e^{2\pi i/n} \neq 1$, ensuring that $\chi(g) = \chi({g_0}^\ell) = e^{2\pi i \ell/n} \neq 1$ since $\ell < n$.

For general G, we know by group theory that G has a direct product decomposition $G \cong C_1 \times \cdots \times C_r$ where C_i is cyclic [3, pp. 344-345]. Let $g \in G$ have the form $(c_1, \ldots, c_r)$ for $c_i \in G_i$. If $g \neq 1$, then some $c_i \neq 1$. By the cyclic case, there exists a character $\psi : C_i \to S^1$ such that $\psi(c_i) \neq 1$. Let $\chi = \psi \circ \pi_i$ where $\pi_i : G \to C_i$ is the projection of G onto its ith factor. Then χ is a character on G and $\chi(g) = \psi(\pi_i g) = \psi(c_i) \neq 1$. □

Theorem 2.1.6. *Let G be a finite abelian group and let $\widehat{G}$ be its dual group. Then $|G| = |\widehat{G}|$, and in fact $\widehat{G} \cong G$.*

Proof. If G is cyclic, then it is clear from Theorem 2.1.4 that $\widehat{G} \cong G$.

To handle non-cyclic G, we first show for any finite abelian groups A and B that $\widehat{A \times B} \cong \widehat{A} \times \widehat{B}$.

Let χ be a character on $A \times B$. Let χ_A and χ_B be the restriction of χ to A and B respectively, i.e., $\chi_A(a) = \chi(a, 1)$ and $\chi_B(b) = \chi(1, b)$. Then χ_A and χ_B are characters on A and B respectively, and $\chi(a, b) = \chi((a, 1)(b, 1)) = \chi_A(a)\chi_B(b)$. In this manner, we obtain a mapping ϕ: $\widehat{A \times B} \to \widehat{A} \times \widehat{B}$ by $\chi \mapsto (\chi_A, \chi_B)$.

We will now check that ϕ is an isomorphism. To show that ϕ is a homomorphism,

for χ and χ' in $\widehat{A \times B}$ we have

$$\phi(\chi\chi') = ((\chi\chi')_A, (\chi\chi')_B) \text{ and } \phi(\chi)\phi(\chi') = (\chi_A, \chi_B)(\chi'_A, \chi'_B) = (\chi_A\chi'_A, \chi_B\chi'_B),$$

so we need to show $(\chi\chi')_A = \chi_A\chi'_A$ and $(\chi\chi')_B = \chi_B\chi'_B$. For $a \in A$,

$$(\chi\chi')_A(a) = (\chi\chi')(a,1) = \chi(a,1)\chi'(a,1) = \chi_A(a)\chi'_A(a) = (\chi_A\chi'_A)(a),$$

so $(\chi\chi')_A = \chi_A\chi'_A$. That $(\chi\chi')_B = \chi_B\chi'_B$ follows in the same way. The homomorphism is injective since if χ_A and χ_B are trivial, then $\chi(a,b) = \chi_A(a)\chi_B(b) = 1$, so χ is also trivial. To show ϕ is surjective, for any choice of (ψ, ψ') in $\widehat{A} \times \widehat{B}$, define $\chi : A \times B \to S^1$ by $\chi(a,b) = \psi(a)\psi'(b)$. This is a character on $A \times B$ and $\chi_A = \psi$, $\chi_B = \psi'$.

Returning to the theorem, write $G = C_1 \times \cdots \times C_m$ where each C_i is cyclic. By induction on the number of terms, $\widehat{A \times B} \cong \widehat{A} \times \widehat{B}$ generalizes to $\widehat{(A_1 \times \cdots \times A_m)} \cong \widehat{A_1} \times \cdots \times \widehat{A_m}$ for any finite abelian groups $A_1, \ldots, A_m$. Therefore, $\widehat{G} \cong \widehat{C_1} \times \cdots \times \widehat{C_m}$. From the cyclic case, $\widehat{C_i} \cong C_i$, so $\widehat{G} \cong \widehat{C_1} \times \cdots \times \widehat{C_m} \cong C_1 \times \cdots \times C_m \cong G$. □

Theorem 2.1.7. *For any finite abelian group G and $g \in G$,*

$$\sum_{\chi \in \widehat{G}} \chi(g) = \begin{cases} |G| & \text{if } g = 1, \\ 0 & \text{if } g \neq 1. \end{cases} \tag{2.1.1}$$

where χ runs through all characters of G. More generally, for all g and h in G,

$$\sum_{\chi \in \widehat{G}} \chi(g)\overline{\chi(h)} = \begin{cases} |G| & \text{if } g = h, \\ 0 & \text{if } g \neq h. \end{cases} \tag{2.1.2}$$

Proof. Let $S = \sum_\chi \chi(g)$. If $g = 1$, then $\chi(g) = 1$ for all $\chi \in \widehat{G}$, so S is equal to the size of $\widehat{G}$, which is $|G|$ by Theorem 2.1.6. If $g \neq 1$, choose a character ψ of G such that $\psi(g) \neq 1$. Such a ψ is guaranteed to exist by Theorem 2.1.5. Multiplying S by $\psi(g)$ yields

$$\psi(g)S = \psi(g)\sum_\chi \chi(g) = \sum_\chi (\psi\chi)(g) = \sum_\chi \chi(g) = S.$$

This implies that either $\psi(g) = 1$ or $S = 0$. But $\psi(g) \neq 1$ by assumption, so $S = 0$. This proves (2.1.1).

To prove (2.1.2), write the left side as $\sum_\chi \chi(g)\chi(h^{-1}) = \sum_\chi \chi(gh^{-1})$. Using gh^{-1} for g in (2.1.1), we get (2.1.2). □

Theorem 2.1.8. *For any finite abelian group G and $\chi \in \widehat{G}$,*

$$\sum_{g \in G} \chi(g) = \begin{cases} |G| & \text{if } \chi = \chi_1, \\ 0 & \text{if } \chi \neq \chi_1, \end{cases} \tag{2.1.3}$$

where g runs through all elements of G and χ_1 is the trivial character of G.

Proof. Set $S' = \sum_g \chi(g)$.

Suppose $\chi = \chi_1$. Then it is clear that $S' = |G|$.

Suppose $\chi \neq \chi_1$, so there is $g_0 \in G$ such that $\chi(g_0) \neq 1$. Then

$$\chi(g_0)S' = \sum_g \chi(g_0)\chi(g) = \sum_g \chi(g_0 g) = \sum_g \chi(g) = S'.$$

This implies that $\chi(g_0) = 1$ or $S' = 0$. But $\chi(g_0) \neq 1$ by assumption, so $S' = 0$. $\square$

In our study of Dirichlet's theorem, we will take $G = (\mathbf{Z}/m)^\times$. A character of $(\mathbf{Z}/m)^\times$ is called a *Dirichlet character*. In this case, the order of G is equal to $\varphi(m)$. A Dirichlet character χ on $(\mathbf{Z}/m)^\times$ can be defined on all of $\mathbf{Z}$ by setting

$$\chi(n) = \begin{cases} \chi(n \bmod m), & \text{if } (n,m) = 1, \\ 0, & \text{if } (n,m) \neq 1. \end{cases} \tag{2.1.4}$$

Then $\chi(n_1 n_2) = \chi(n_1)\chi(n_2)$ for all $n_1, n_2 \in \mathbf{Z}$. We define the trivial Dirichlet character mod m, denoted 1_m, as follows:

$$1_m(n) = \begin{cases} 1, & \text{if } (n,m) = 1, \\ 0, & \text{if } (n,m) \neq 1. \end{cases}$$

Note that $\chi(n)^{-1} = \overline{\chi}(n)$ when n and m are relatively prime but not when n and m are not relatively prime because $\chi(n) = 0$. However, $\overline{\chi}(n)$ is still a valid expression for all n by defining $\overline{\chi}(n) = \overline{\chi(n)}$. In particular, if $(n,m) \neq 1$, then $\overline{\chi}(n) = \overline{\chi(n)} = \overline{0} = 0$.

For Dirichlet characters, viewed as functions on $\mathbf{Z}$, (2.1.2) becomes

$$\frac{1}{\varphi(m)} \sum_\chi \chi(n)\overline{\chi}(a) = \begin{cases} 1 & \text{if } n \equiv a \bmod m, \\ 0 & \text{if } n \not\equiv a \bmod m \end{cases} \tag{2.1.5}$$

for all $n \in \mathbf{Z}$, where the sum is over all Dirichlet characters mod m and $(a, m) = 1$. Formula (2.1.5) will be an important algebraic ingredient in the proof of Dirichlet's theorem.

Our last general result about characters in this section concerns the existence of a character taking prescribed values at "independent" elements of the group.

Definition 2.1.9. Let H be a finite abelian group. We call a set of elements $h_i \in H$ *multiplicatively independent* if $\prod_i h_i^{e_i} = 1$ implies that $h_i^{e_i} = 1$ for every i. Equivalently, no h_j in the set can have a nontrivial power written as a product of the powers of the h_i for $i \neq j$.

Remark 2.1.10. If the h_i's all have order 2, then multiplicative independence in H says no h_j is a product of the h_i for $i \neq j$.

Theorem 2.1.11. *Let H be a finite abelian group. For $h_1, \ldots, h_r \in H$, set $\varphi : \langle h_1 \rangle \times \cdots \times \langle h_r \rangle \to \langle h_1, \ldots, h_r \rangle$ by $\varphi(h_1^{e_1}, \ldots, h_r^{e_r}) = h_1^{e_1} \ldots h_r^{e_r}$. Then $\{h_1, \ldots, h_r\}$ is multiplicatively independent in H if and only if φ is an isomorphism.*

Proof. To check that φ is a homomorphism, note from H being abelian that

$$\begin{aligned} \varphi\left((h_1^{a_1}, \ldots, h_r^{a_r})(h_1^{b_1}, \ldots, h_r^{b_r})\right) &= \varphi\left(h_1^{a_1+b_1}, \ldots, h_r^{a_r+b_r}\right) \\ &= h_1^{a_1+b_1} \cdots h_r^{a_r+b_r} \\ &= h_1^{a_1} \cdots h_r^{a_r} h_1^{b_1} \cdots h_r^{b_r} \\ &= \varphi\left(h_1^{a_1}, \ldots, h_r^{a_r}\right) \varphi\left(h_1^{b_1}, \ldots, h_r^{b_r}\right). \end{aligned}$$

Suppose that $\varphi(h_1^{a_1}, \ldots, h_r^{a_r}) = 1$. Then $\prod_i h_i^{a_i} = 1$. If the h_i are multiplicatively independent in H, we must have that $h_i^{a_i} = 1$ for all i. Hence φ is injective.

The function φ is clearly surjective since all elements in the image are of the form $\prod_i h_i^{a_i}$. Therefore, multiplicative independence of the h_i in H implies that φ is an isomorphism.

Conversely, if φ is an isomorphism, then the h_i must be multiplicatively independent in H because the kernel of φ is trivial. □

Theorem 2.1.12. *In a finite abelian group H, let $\{h_1, \ldots, h_r\}$ be multiplicatively independent. Let m_i be the order of h_i in H, and let $\zeta_i \in S^1$ satisfy $\zeta_i^{m_i} = 1$. Then there exists a character $\chi \in \widehat{H}$ such that $\chi(h_i) = \zeta_i$ for all i.*

To have such a χ, it is necessary that $\zeta_i^{m_i} = 1$ because $\zeta_i^{m_i} = \chi(h_i)^{m_i} = \chi(h_i^{m_i}) = \chi(1) = 1$.

Proof. Define $\psi : \langle h_1, \ldots, h_r \rangle \to S^1$ by $\psi(h_1^{a_1} \cdots h_r^{a_r}) = \zeta_1^{a_1} \cdots \zeta_r^{a_r}$. To show this function is well-defined, suppose $h_1^{a_1} \cdots h_r^{a_r} = h_1^{b_1} \cdots h_r^{b_r}$. Then $1 = h_1^{a_1 - b_1} \cdots h_r^{a_r - b_r}$. Since the h_i are multiplicatively independent, this implies that $h_i^{a_i - b_i} = 1$ for all i, so $m_i | (a_i - b_i)$ and so $a_i \equiv b_i \bmod m_i$. Because $\zeta_i^{m_i} = 1$, we conclude that $\zeta_i^{a_i} = \zeta_i^{b_i}$, so $\zeta_1^{a_1} \cdots \zeta_r^{a_r} = \zeta_1^{b_1} \cdots \zeta_r^{b_r}$. One can also easily check that ψ is a homomorphism. By construction, $\psi(h_i) = \zeta_i$.

To complete the proof, we will extend ψ from a character on $\langle h_1, \ldots, h_r \rangle$ to a character on H. Let $K = \langle h_1, \ldots, h_r \rangle$. If $K = H$, we are done. If $K \neq H$, pick $h \in H - K$ and set $\langle K, h \rangle = \{kh^i : k \in K, i \in \mathbf{Z}\}$. If $\langle K, h \rangle \neq H$, then $\langle K, h \rangle \subset \langle K, h, g \rangle$ for some $g \in H$ such that $g \notin \langle K, h \rangle$. Since H is finite, we may repeat this process a finite number of times in order to make a tower of subgroups $K \subset \langle K, h \rangle \subset \langle K, h, g \rangle \subset \cdots \subset H$ where we successively adjoin one new element each time. If we can extend the character ψ on K to a character ψ' on $\langle K, h \rangle$, that process can be

repeated to extend ψ to a character on H. It therefore suffices to show that ψ can be extended from K to $\langle K, h\rangle$.

Let m be the minimal positive integer such that $h^m \in K$. The value of $\psi(h^m)$ is already known since $h^m \in K$. Choose $z \in \mathbf{C}^\times$ such that $z^m = \psi(h^m)$ so $z \in S^1$, and define $\psi' : \langle K, h\rangle \to S^1$ by $\psi'(kh^i) = \psi(k)z^i$.

To see that ψ' is well-defined, suppose $k_1 h^{i_1} = k_2 h^{i_2}$. Then $h^{i_1 - i_2} = k_2 k_1^{-1} \in K$, so $m|(i_1 - i_2)$ and $i_1 \equiv i_2 \bmod m$. Let $i_2 = i_1 + mc$ for some integer c. Then $k_1 = k_2 h^{i_2 - i_1} = k_2 h^{mc} = k_2 (h^m)^c$ and $h^m \in K$. This implies that

$$\psi(k_1) = \psi(k_2)\psi(h^m)^c = \psi(k_2)(z^m)^c = \psi(k_2)z^{mc} = \psi(k_2)z^{i_2 - i_1} \Rightarrow \psi(k_1)z^{i_1} = \psi(k_2)z^{i_2},$$

so ψ' is a well-defined function on $\langle K, h\rangle$. In particular, for $k \in K$, $\psi'(k) = \psi(k)z^0 = \psi(k)$, so $\psi' = \psi$ on K.

Lastly, we must check that ψ' is a homomorphism. Computation shows that $\psi'\left((k_1 h^{i_1})(k_2 h^{i_2})\right) = \psi(k_1 k_2)z^{i_1 + i_2} = \psi(k_1)z^{i_1}\psi(k_2)z^{i_2} = \psi'(k_1 h^{i_1})\psi'(k_2 h^{i_2})$. □

2.2 Theorems from Complex Analysis

We will next discuss some theorems from complex analysis that will be needed later.

For an open set $U \subset \mathbf{C}$, a function $f : U \to \mathbf{C}$ is called *holomorphic* (or *analytic*) if it is differentiable at each point in U. Cauchy's integral formula says for holomorphic f that

$$f(s) = \frac{1}{2\pi i}\int_\gamma \frac{f(\zeta)}{\zeta - s} d\zeta$$

for each $s \in U$, where γ is any continuous path in U that traces out a counterclockwise

loop once around s and can be shrunk to a point in U. Also, $\int_\gamma f(s)ds = 0$ for all closed paths γ in U that can be shrunk to a point in U. Conversely, Morera's theorem [11, p. 208] says that if $f : U \to \mathbf{C}$ is continuous and $\int_T f(s)ds = 0$ for all triangles T in U then f is holomorphic on U.

Theorem 2.2.1. *If D is an open disc in $\mathbf{C}$ and $f : D \to \mathbf{C}$ is holomorphic, then the power series expansion of f at the center of D converges and equals f on all of D.*

Proof. Our proof is based on [14, pp. 49-50]. Let s_0 be the center of D. Pick a circle C centered at s_0 with radius less than the radius of D. For any s inside the open disc at s_0 bounded by C, the Cauchy integral formula tells us that

$$f(s) = \frac{1}{2\pi i}\int_C \frac{f(\zeta)}{\zeta - s}d\zeta.$$

Write

$$\frac{1}{\zeta - s} = \frac{1}{\zeta - s_0 - (s - s_0)} = \frac{1}{\zeta - s_0}\frac{1}{1 - (\frac{s-s_0}{\zeta - s_0})}. \tag{2.2.1}$$

We wish to rewrite the last factor on the right of (2.2.1) using its geometric series expansion. Since $|s - s_0| < |\zeta - s_0|$ for all ζ on C, define a distance function $d : C \to [0, 1)$ by $d(\zeta) = |(s - s_0)/(\zeta - s_0)|$. This is a continuous mapping on a compact set, so there exists $r \in [0, 1)$ such that $d(\zeta) \leq r < 1$ for all $\zeta \in C$. (This is analogous to the extreme value theorem for continuous real-valued functions of a single real variable.) We may therefore write

$$\left|\frac{s - s_0}{\zeta - s_0}\right| \leq r < 1, \quad \text{so} \quad \frac{1}{1 - (\frac{s-s_0}{\zeta - s_0})} = \sum_{n=0}^{\infty}\left(\frac{s - s_0}{\zeta - s_0}\right)^n,$$

where the series converges uniformly for all $\zeta \in C$ since the series is bounded termwise

by $\sum_{n=0}^{\infty} r^n$, which is convergent. This allows us to interchange the series with the integral when we combine the above equations, thereby obtaining

$$f(s) = \sum_{n=0}^{\infty} \left(\frac{1}{2\pi i} \int_C \frac{f(\zeta)}{(\zeta - s_0)^{n+1}} d\zeta \right) \cdot (s - s_0)^n . \tag{2.2.2}$$

The coefficients of this series are the same for *any* s inside the disc bounded by C, and they are in fact independent of C, being $f^{(n)}(s_0)/n!$.

This proves the power series expansion of f converges and equals f on D. □

Theorem 2.2.2 (Cauchy's estimate)**.** *Let $f(s)$ be holomorphic on a closed disc D of radius $r > 0$ centered at a such that $|f(s)| \leq M$ for all $s \in D$. Then $|f'(a)| \leq M/r$.*

We had only defined holomorphic functions on open sets in $\mathbf{C}$. A function on a closed set in $\mathbf{C}$ is called holomorphic when it is the restriction to that set of a holomorphic function on a larger open set.

Proof. From Cauchy's integral formula or the coefficient in (2.2.2) at $n = 1$,

$$f'(a) = \frac{1}{2\pi i} \int_C \frac{f(\zeta)}{(\zeta - a)^2} d\zeta,$$

where C is the boundary of D, parameterized as $\zeta = a + re^{i\theta}$ for $\theta \in [0, 2\pi]$. Thus

$$\begin{aligned} |f'(a)| &\leq \frac{1}{2\pi} \int_0^{2\pi} \left| \frac{f(a + re^{i\theta})}{(re^{i\theta})^2} rie^{i\theta} \right| d\theta \\ &\leq \frac{M}{2\pi} \int_0^{2\pi} \frac{d\theta}{r} \\ &= \frac{M}{r}. \end{aligned}$$

□

Remark 2.2.3. By a similar argument for higher derivatives, Cauchy's estimate extends to $|f^{(n)}(a)| \leq n!M/r^n$ for all $n \geq 0$. We will only need the case $n = 1$.

Lemma 2.2.4. *For $s \in \mathbf{C}$ and $r \geq 0$, let $\overline{D}(s, r)$ denote the closed disc centered at s of radius r. For any nonempty open subset Ω of $\mathbf{C}$ and nonempty compact subset K of Ω, there is an $r > 0$ such that $\bigcup_{a \in K} \overline{D}(a, r) \subset \Omega$, and this union is compact.*

Proof. Since Ω is open, for each $s \in \Omega$ there exists a closed disc of positive radius centered at s that is a subset of Ω. The union of the interiors of such discs, as s runs over all points in K, is a covering of K. Since K is compact, the union of finitely many such open discs covers K. Therefore the union of the closures of these finitely many discs also covers K and is a subset of Ω.

Denote these open discs by $D_1, \ldots, D_n$ with centers s_i and radii R_i. For $1 \leq i \leq n$, define $g_i \colon K \to \mathbf{R}$ by $g_i(a) = \max\{0, R_i - |a - s_i|\}$. Then g_i is continuous and nonnegative on K. Define g pointwise by $g(a) = \max\{g_i(a)\}$ for $1 \leq i \leq n$. Then g is continuous on K. If $g(a) = 0$ for some $a \in K$, then $g_i(a) = 0$ for all i so $R_i \leq |a - s_i|$ for all i. This means $a \notin D_i$ for all i, a contradiction to the set of all D_i covering K. Hence $g(a) > 0$ for all $a \in K$. Since K is compact, g attains a minimum value r on K where $r > 0$. We will show $\bigcup_{a \in K} \overline{D}(a, r) \subset \Omega$ for this r.

By the definition of g, for each $a \in K$ we have $g(a) = g_i(a)$ for some i. Since $g(a) > 0$, we must have that $g_i(a) = R_i - |a - s_i|$ when $g(a) = g_i(a)$. Hence

$$r \leq g(a) = R_i - |a - s_i| \Longrightarrow |a - s_i| + r \leq R_i. \tag{2.2.3}$$

Then for all $y \in \overline{D}(a, r)$, we have from (2.2.3) and the triangle inequality

$$|y - s_i| \leq |y - a| + |a - s_i| \leq r + |a - s_i| \leq R_i,$$

so $\overline{D}(a,r) \subset \overline{D}(s_i, R_i) = \overline{D}_i \subset \Omega$. That proves $\bigcup_{a\in K} \overline{D}(a,r) \subset \Omega$.

To show $\widetilde{K} := \bigcup_{a\in K} \overline{D}(a,r)$ is compact, we will show it is closed and bounded. By definition

$$\widetilde{K} = \{s \in \mathbf{C} : |s-a| \leq r \text{ for some } a \in K\}, \tag{2.2.4}$$

and K is bounded because it is compact, so (2.2.4) shows that $\widetilde{K}$ is also bounded. To show $\widetilde{K}$ is closed, we will show its complement $\{s \in \mathbf{C} : |s-a| > r \text{ for all } a \in K\}$ is open. For each $s \in \widetilde{K}$, the inequality $|s-a| > r$ for all $a \in K$ can be improved to $|s-a| \geq r+\varepsilon$ for some $\varepsilon > 0$ because K is compact: the function $f_s \colon K \to \mathbf{R}$ given by $f_s(a) = |s-a| - r$ is continuous and positive on K, so f_s has a positive lower bound. Let ε be such a lower bound. From $|s-a| \geq r+\varepsilon$, the open ball $D(s,\varepsilon)$ is in the complement of $\widetilde{K}$ by the triangle inequality, since for $z \in D(s,\varepsilon)$ and $a \in K$ we have $|s-a| \leq |s-z| + |z-a| < \varepsilon + |z-a|$, so $|z-a| > |s-a| - \varepsilon \geq r$. □

Theorem 2.2.5. *Let $\{f_n\}$ be a sequence of holomorphic functions on an open subset Ω of $\mathbf{C}$. If $\{f_n\}$ converges uniformly on every compact subset of Ω to a limit function f, then f is holomorphic on Ω and the derivatives f_n' converge uniformly on every compact subset of Ω to f'.*

Proof. Our proof is based on [11, p. 214]. Since being holomorphic is a local property and $f_n \to f$ uniformly on any compact subset of any open disc in Ω, to prove that f is holomorphic we can assume that Ω is an open disc. Each f_n is continuous and the sequence $\{f_n\}$ converges uniformly to f on every compact disc in Ω, so f is continuous on Ω since continuity is a local property and a uniform limit of continuous functions is continuous [15, pp. 225-226]. Let T be a triangle contained completely within Ω. By "triangle" we mean the boundary, so it is a closed path. Since f is continuous, it

is integrable. Then

$$\left| \int_T f(s)ds - \int_T f_n(s)ds \right| \leq \int_T |f(s) - f_n(s)|\, ds \leq ||f - f_n||_T \cdot \text{length}(T),$$

where $||f - f_n||_T$ is the supremum norm on T. Since T is compact, $||f - f_n||_T \to 0$ as $n \to \infty$. We can shrink T to a point in Ω since we are taking Ω to be a disc, so by Cauchy's theorem $\int_T f_n(s)ds = 0$. Thus $|\int_T f(s)ds| \leq ||f - f_n||_T \cdot \text{length}(T) \to 0$ as $n \to \infty$, so $\int_T f(s)ds = 0$ for all triangles T in Ω. By Morera's theorem, we conclude that f is holomorphic on Ω.

To prove that $f_n' \to f'$ uniformly on each compact subset of Ω, let K be a compact subset of Ω. By Lemma 2.2.4 there is an $r > 0$ such that $\widetilde{K} := \bigcup_{a \in K} \overline{D}(a, r)$ is a subset of Ω, and $\widetilde{K}$ is compact. Let M_n denote the maximum of $|f - f_n|$ on $\widetilde{K}$. For each $a \in K$ we can apply Theorem 2.2.2 to $f - f_n$ on $\overline{D}(a, r)$ in order to obtain $|f'(a) - f_n'(a)| \leq M_n/r$. Since f_n converges to f uniformly on each compact subset of Ω, such as $\widetilde{K}$, $M_n \to 0$ as $n \to \infty$. Thus f_n' converges uniformly to f' on $\widetilde{K}$, and thus on its subset K, so $f_n' \to f'$ uniformly on each compact subset of Ω. □

Definition 2.2.6. Let Ω be open in $\mathbf{C}$. A *logarithm* of a holomorphic function $f : \Omega \to \mathbf{C}$ is a holomorphic function $L_f : \Omega \to \mathbf{C}$ such that $e^{L_f(s)} = f(s)$ for all $s \in \Omega$.

Theorem 2.2.7. *If Ω is a connected and simply connected open set in $\mathbf{C}$ then any nonvanishing holomorphic function $f : \Omega \to \mathbf{C}$ admits a logarithm, and any two logarithms of f on Ω differ by an integral multiple of $2\pi i$.*

Proof. Our proof is based on [14, pp. 100-101]. Since Ω is connected and open, it is

path connected. Fix $s_0 \in \Omega$, and define $L_f : \Omega \to \mathbf{C}$ by

$$L_f(s) = \int_\gamma \frac{f'(w)}{f(w)} dw + c_0,$$

where γ is any path in Ω connecting s_0 to s, and $c_0 \in \mathbf{C}$ satisfies $e^{c_0} = f(s_0)$. To show $L_f(s)$ is independent of the choice of path, let γ_1 be another path connecting s_0 to s in Ω. The path φ obtained by traveling along γ from s_0 to s and then along γ_1 backwards from s to s_0 is a loop that can be shrunk to a point in Ω by simple connectedness. Since f is holomorphic on Ω, Cauchy's integral formula tells us

$$0 = \int_\varphi \frac{f'(w)}{f(w)} dw = \int_\gamma \frac{f'(w)}{f(w)} dw - \int_{\gamma_1} \frac{f'(w)}{f(w)} dw \Rightarrow \int_\gamma \frac{f'(w)}{f(w)} dw = \int_{\gamma_1} \frac{f'(w)}{f(w)} dw,$$

so $L_f(s)$ is independent of choice of path from s_0 to s.

To show that $L_f \colon \Omega \to \mathbf{C}$ is holomorphic, we will use the limit definition of the derivative:

$$L_f'(s) = \lim_{h \to 0} \frac{L_f(s+h) - L_f(s)}{h}.$$

For small $h \neq 0$, in $L_f(s+h)$ we can choose as the path from s_0 to $s+h$ the concatenation of the path from s_0 to s used for $L_f(s)$ followed by the straight line path from s to $s+h$. Then

$$\frac{L_f(s+h) - L_f(s)}{H} = \frac{1}{h} \int_{[s,s+h]} \frac{f'(w)}{f(w)} dw.$$

Since the path from s to $s+h$ is a straight line we can parameterize w along it as

$w = s + th$ for $t \in [0,1]$, so

$$\frac{1}{h}\int_{[s,s+h]} \frac{f'(w)}{f(w)}dw = \frac{1}{h}\int_0^1 \frac{f'(s+th)}{f(s+th)}h\ dt = \int_0^1 \frac{f'(s+th)}{f(s+th)}dt.$$

From continuity of f'/f at s,

$$\lim_{h\to 0}\int_0^1 \frac{f'(s+th)}{f(s+th)}dt = \int_0^1 \frac{f'(s)}{f(s)}dt = \frac{f'(s)}{f(s)},$$

so $L_f'(s) = f'(s)/f(s)$. Thus L_f is holomorphic.

To show that $e^{L_f(s)} = f(s)$, we calculate

$$\begin{aligned}\frac{d}{ds}\left(f(s)e^{-L_f(s)}\right) &= f'(s)e^{-L_f(s)} - f(s)L_f'(s)e^{-L_f(s)} \\ &= e^{-L_f(s)}f(s)\left(f'(s)/f(s) - L_f'(s)\right) \\ &= 0,\end{aligned}$$

so $f(s)e^{-L_f(s)}$ is constant (since Ω is connected). Evaluating this expression at s_0 we get the value $f(s_0)e^{-c_0} = 1$, so $f(s) = e^{L_f(s)}$ for all $s \in \Omega$. Thus L_f is a logarithm of f on Ω.

Suppose that f admits another logarithm $\tilde{L}_f$ on Ω, so $f(s) = e^{\tilde{L}_f(s)}$ for all $s \in \Omega$. Then for all $s \in \Omega$,

$$e^{L_f(s)} = e^{\tilde{L}_f(s)} \Rightarrow e^{L_f(s)-\tilde{L}_f(s)} = 1 \Rightarrow L_f(s) - \tilde{L}_f(s) \in 2\pi i\mathbf{Z}.$$

Since $L_f - \tilde{L}_f$ is continuous on Ω, its image is connected. The set $2\pi i\mathbf{Z}$ is discrete, so the image of $L_f - \tilde{L}_f$ on Ω is a point. Therefore there is an integer k such that $L_f(s) - \tilde{L}_f(s) = 2\pi ik$ for all $s \in \Omega$. □

Example 2.2.8. On $\{s : |s| < 1\}$, a logarithm of $1/(1-s)$ is $\sum_{k\geq 1} s^k/k$. That is, $\exp(\sum_{k\geq 1} s^k/k) = 1/(1-s)$. To prove this, both sides are holomorphic on $|s| < 1$ and equal for real s when $0 < s < 1$ by calculus, so they are equal for all s with $|s| < 1$.

Chapter 3

The zeta-function and L-functions

Proving Dirichlet's theorem requires the use of complex-valued functions called the Riemann zeta-function and Dirichlet L-functions. In this chapter, we will introduce these functions and prove various properties of them that will be essential in the next chapter.

3.1 The Riemann zeta-function

The *Riemann zeta-function* is defined for $\mathrm{Re}(s) > 1$ by $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$.

Following a common tradition in number theory, we will write the real and imaginary parts of s as $s = \sigma + it$ with $\sigma, t \in \mathbf{R}$.

Theorem 3.1.1. *The zeta-function is absolutely convergent for* $\mathrm{Re}(s) > 1$.

Proof. Note that

$$\sum_{n=1}^{\infty}\left|\frac{1}{n^s}\right| \leq \sum_{n=1}^{\infty}\frac{1}{|n^s|} = \sum_{n=1}^{\infty}\frac{1}{n^{\mathrm{Re}(s)}} = \sum_{n=1}^{\infty}\frac{1}{n^{\sigma}}.$$

Fix $\sigma > 0$. The function $f_\sigma(x) = 1/x^\sigma$ is a monotonic decreasing positive function for $x > 0$. By the integral comparison test, $\sum_{n=1}^{\infty} 1/n^\sigma$ is convergent if and only if $\int_1^\infty f_\sigma(x)dx$ is convergent. Evaluating,

$$\int_1^{\infty} f_\sigma(x)dx = \frac{x^{1-\sigma}}{1-\sigma}\Big|_1^{\infty} = \lim_{b\to\infty}\frac{b^{1-\sigma}}{1-\sigma} - \frac{1}{1-\sigma} = \frac{1}{\sigma-1} - \lim_{b\to\infty}\frac{1}{(\sigma-1)b^{\sigma-1}}.$$

This is convergent if (and only if) $\sigma > 1$. □

Intuitively, taking the limit of $\zeta(s)$ as $s \to 1^+$ yields the (divergent) harmonic series, so we expect $\zeta(s) \to \infty$ as $s \to 1^+$. Let us prove this.

Theorem 3.1.2. *For $s \in \mathbf{R}$, the limit of $\zeta(s)$ as $s \to 1^+$ is infinity.*

Proof. For each $s > 1$, since x^{-s} is a monotonic decreasing function for $x > 0$, we have the inequality

$$\frac{1}{(n+1)^s} < \int_n^{n+1} x^{-s}dx < \frac{1}{n^s}. \tag{3.1.1}$$

Summing (3.1.1) over $n \geq 1$ yields $\zeta(s) - 1 < \int_1^\infty x^{-s}dx < \zeta(s)$, and $\int_1^\infty x^{-s}dx = 1/(s-1)$. Rearranging terms, we get

$$1 < (s-1)\zeta(s) < s, \tag{3.1.2}$$

so $\lim_{s\to 1^+}(s-1)\zeta(s) = 1$. For $s > 1$, $\zeta(s) > 1/(s-1)$ by (3.1.2), so $\lim_{s\to 1^+}\zeta(s) = \infty$. □

By the next theorem, we will see that the zeta-function can be written as a product over the primes.

Theorem 3.1.3. *Let $h : \mathbf{Z}^+ \to \mathbf{C}$ be totally multiplicative such that $|h(n)| < 1$ for all n and $\sum_{n=1}^{\infty} h(n)$ is absolutely convergent. Then $\sum_{n=1}^{\infty} h(n) = \prod_p 1/(1-h(p))$, where the product runs over the primes and is also absolutely convergent.*

Proof. For $x \geq 2$, consider the finite product $P(x) = \prod_{p \leq x}\{1+h(p)+h(p^2)+\cdots\}$ over all primes $p \leq x$. Each series $1+h(p)+h(p^2)+\cdots$ is $1+h(p)+h(p)^2+\cdots$, which is an absolutely convergent geometric series. Thus $P(x)$ is a finite product of absolutely convergent series, so it may be expanded and rearranged into an absolutely convergent series whose general term is $h(p_1^{a_1})h(p_2^{a_2})\cdots h(p_r^{a_r})$, which is $h(p_1^{a_1}p_2^{a_2}\cdots p_r^{a_r})$ since h is totally multiplicative.

By unique factorization in $\mathbf{Z}^+$, $P(x) = \sum_{n \in A} h(n)$ where A is the set of positive integers having all prime factors less than or equal to x. Therefore, $\sum_{n=1}^{\infty} h(n) - P(x) = \sum_{n \in B} h(n)$, where B is the set of positive integers having at least one prime factor greater than x, and $|\sum_{n=1}^{\infty} h(n) - P(x)| \leq \sum_{n \in B} |h(n)| \leq \sum_{n>x} |h(n)|$. As $x \to \infty$, we have $\sum_{n>x} |h(n)| \to 0$ since the series $\sum_{n \geq 1} |h(n)|$ converges. This implies that $\lim_{x\to\infty} P(x) = \sum_{n=1}^{\infty} h(n)$.

Since h is totally multiplicative, $P(x) = \prod_{p \leq x}(\sum_{i=0}^{\infty} h(p)^i) = \prod_{p \leq x} 1/(1-h(p))$. Taking the limit of this as $x \to \infty$ yields the desired result. □

The above theorem, using $h(n) = 1/n^s$, shows that

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \frac{1}{1-1/p^s} \tag{3.1.3}$$

for $\mathrm{Re}(s) > 1$. The infinite product on the right is called the *Euler product* of $\zeta(s)$.

Euclid showed there are an infinitely many primes by contradiction. Here is another proof of this, due to Euler in 1737. It motivates the proof of Dirichlet's theorem.

Theorem 3.1.4. *As $s \to 1^+$, $\sum_p 1/p^s \to \infty$, where the sum is over all primes p.*

Proof. Our argument is based on [7, p. 250].

Let $\lambda_N(s) = \prod_{p>N} 1/(1-1/p^s)$ for $\mathrm{Re}(s) > 1$. By (3.1.3), we have for $\mathrm{Re}(s) > 1$ that

$$\zeta(s) = \prod_p \frac{1}{1-1/p^s} = \lambda_N(s) \prod_{p \le N} \frac{1}{1-1/p^s} \text{ where } \lambda_N(s) \to 1 \text{ as } N \to \infty. \tag{3.1.4}$$

Taking the logarithm of both sides of (3.1.4) when $s > 1$ and using the series expansion $-\ln(1-x) = \sum_{k=1}^{\infty} x^k/k$ for $|x| < 1$, we get for $s > 1$ that

$$\ln \zeta(s) = \ln \lambda_N(s) + \sum_{p \le N} \sum_{k=1}^{\infty} \frac{1}{kp^{ks}}. \tag{3.1.5}$$

Taking the limit as $N \to \infty$ of (3.1.5) yields

$$\ln \zeta(s) = \sum_p \sum_{k=1}^{\infty} \frac{1}{kp^{ks}} = \sum_p \frac{1}{p^s} + \sum_p \sum_{k=2}^{\infty} \frac{1}{kp^{ks}}. \tag{3.1.6}$$

The second double series on the right in (3.1.6) converges for each $s > 1/2$ since

$$\sum_p \sum_{k=2}^{\infty} \frac{1}{kp^{ks}} < \sum_p \sum_{k=2}^{\infty} \frac{1}{p^{ks}} = \sum_p \frac{1}{p^{2s}(1-1/p^s)} \le \frac{1}{1-1/2^s} \sum_p \frac{1}{p^{2s}} < \frac{1}{1-1/\sqrt{2}} \zeta(2s).$$

By Theorem 3.1.2 and that fact that $\lim_{x\to\infty} \ln(x) = \infty$, we have $\ln \zeta(s) \to \infty$ as $s \to 1^+$. For $s > 1$,

$$\sum_p \sum_{k=2}^{\infty} \frac{1}{kp^{ks}} < \frac{\zeta(2)}{1-1/\sqrt{2}},$$

so by (3.1.6) we get $\sum_p 1/p^s \to \infty$ as $s \to 1^+$.

□

Remark 3.1.5. The proof used $\sum_p \sum_{k\geq 2} 1/kp^{ks}$ only for $s \geq 1$, not $s > 1/2$. We will need to use $s > 1/2$ (really, $s > 1 - \varepsilon$) more seriously in Corollary 4.2.2.

Corollary 3.1.6. *The series $\sum_p 1/p$ diverges.*

If there were only finitely many primes, then $\sum_p 1/p$ would be convergent. Therefore, Corollary 3.1.6 shows that there are infinitely many prime numbers.

Proof. For $s > 1$, $1/p^s < 1/p$, so $\sum_p 1/p^s \leq \sum_p 1/p$. Letting $s \to 1^+$, Theorem 3.1.4 and the comparison test show that $\sum_p 1/p = \infty$. □

Here is an alternate heuristic proof of Corollary 3.1.6, more in the spirit of Euler.

Proof. Consider $\log(\sum_n 1/n)$. Using the Euler product at $s = 1$ (which, strictly speaking, is not valid),

$$\log\left(\sum_n \frac{1}{n}\right) = \log\left(\prod_p \frac{1}{1-1/p}\right) = \sum_p \log\left(\frac{1}{1-1/p}\right) = \sum_p \log\left(1 + \frac{1}{p-1}\right).$$

Since $e^x = 1 + x + x^2/2! + \cdots$, we have $x > \log(x+1)$ for $x > 0$. Therefore

$$\sum_p \frac{1}{p-1} \geq \sum_p \log\left(1 + \frac{1}{p-1}\right) = \log\left(\sum_n \frac{1}{n}\right),$$

so the divergence of $\sum_n 1/n$ implies that $\sum_p 1/(p-1)$ diverges. From $\frac{1}{p} \geq \frac{1}{2(p-1)}$, we get $\sum_p 1/p \geq \frac{1}{2}\sum_p 1/(p-1)$. Thus $\sum_p 1/p$ diverges. □

3.2 Dirichlet L-functions

For any Dirichlet character χ, viewed as a function on $\mathbf{Z}$, set

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} \tag{3.2.1}$$

for $\mathrm{Re}(s) > 1$. This is called the *Dirichlet L-series*, or Dirichlet L-function, of χ. Since $|\chi(n)| = 1$ (or 0), we may follow the proof of Theorem 3.1.1 to see that a Dirichlet L-series is absolutely convergent for $\mathrm{Re}(s) > 1$. While Dirichlet only used real s in his work, here we allow s to be a complex number.

Example 3.2.1. When χ_4 is the nontrivial character mod 4, $L(s,\chi_4) = 1 - 1/3^s + 1/5^s - 1/7^s + \ldots$, which converges for all real $s > 0$, and $L(1,\chi_4) = \pi/4$ by Leibniz's formula for $1 - 1/3 + 1/5 - 1/7 + \ldots$.

By Theorem 3.1.3, there is an Euler product representation

$$L(s,\chi) = \prod_p \frac{1}{1 - \chi(p)/p^s} \tag{3.2.2}$$

for $\mathrm{Re}(s) > 1$.

Example 3.2.2. If $\chi = 1_m$, the trivial character mod m, then

$$L(s,\chi) = \prod_{p \text{ not dividing } m} \frac{1}{1 - 1/p^s} = \prod_{p|m} \left(1 - \frac{1}{p^s}\right) \zeta(s).$$

We will see later (Corollary 3.3.6) that the L-series of every nontrivial Dirichlet character converges for $\mathrm{Re}(s) > 0$, although the Euler product in (3.2.2) has no obvious convergence when $0 < \mathrm{Re}(s) \leq 1$. The behavior of $L(s,\chi)$ near $s = 1$ is

central to the proof of Dirichlet's theorem.

3.3 Theorems on Dirichlet Series

We collect here several basic analytic facts about the types of series like $\zeta(s)$ and $L(s, \chi)$.

Definition 3.3.1. A *Dirichlet series* is a function of the form $f(s) = \sum_{n \geq 1} a_n/n^s$ where the a_n are complex constants and s is a complex variable.

Theorem 3.3.2. *Suppose that* $\sum |a_n|/n^\sigma$ *converges for at least one real* σ *and diverges for at least one other real* σ*. Then there exists a real number* σ_0 *such that* $\sum a_n/n^s$ *converges absolutely for* $\mathrm{Re}(s) > \sigma_0$ *but does not converge absolutely for* $\mathrm{Re}(s) < \sigma_0$*.*

Proof. When $s \in \mathbf{C}$ with real part σ, $\sum a_n/n^s$ is absolutely convergent if and only if $\sum a_n/n^\sigma$ is absolutely convergent since $|a_n/n^s| = |a_n|/n^\sigma$. Define D to be the set of real numbers σ such that $\sum |a_n|/n^\sigma$ diverges. By assumption, D and the complement of D are both nonempty. By the comparison test, if $\sigma \in D$ then $\sigma' \in D$ for all $\sigma' < \sigma$, and if $\sigma \notin D$ then $\sigma' \notin D$ for all $\sigma' > \sigma$. Therefore, since the complement of D is nonempty, D is bounded above. Denote the supremum of D by σ_0. The series must diverge for all $\sigma < \sigma_0$ by the comparison test: if $\sigma < \sigma_0$ then there is a $\sigma_1 \in D$ such that $\sigma < \sigma_1 < \sigma_0$, so $\sigma \in D$. The series converges for all $\sigma > \sigma_0$ because σ_0 is the supremum of D and hence an upper bound on all real numbers such that the series diverges. □

Theorem 3.3.3. *If the sequence* $\{a_n\}$ *is bounded, then* $\sum_{n \geq 1} a_n/n^s$ *converges absolutely for* $\mathrm{Re}(s) > 1$*.*

Proof. Since $\{a_n\}$ is bounded, there exists a number $B \in \mathbf{R}$ such that $|a_n| \leq B$ for all n. Hence $\sum_{n\geq 1} |a_n/n^s| \leq B \sum_{n\geq 1} 1/n^\sigma$, which converges for $\sigma > 1$ by Theorem 3.1.1. □

Theorem 3.3.4. *If $\{a_n\}$ is a sequence such that the partial sums $a_1 + \cdots + a_n$ are bounded for all n, then $\sum_{n\geq 1} a_n/n^s$ converges for* $\mathrm{Re}(s) > 0$.

Proof. We will use summation by parts:

$$\sum_{n=1}^{N} u_n(v_n - v_{n-1}) = u_N v_N - u_1 v_0 - \sum_{n=1}^{N-1} v_n(u_{n+1} - u_n)$$

for any sequences $u_1, \ldots, u_N$ and $v_0, v_1, \ldots, v_N$.

Set $b_n = a_1 + \cdots + a_n$ and let $|b_n| \leq b$ for all n. Then $a_n = b_n - b_{n-1}$ where $b_0 = 0$. For $N \geq 2$, we have

$$\begin{aligned}
\sum_{n=1}^{N} \frac{a_n}{n^s} &= \sum_{n=1}^{N} \frac{b_n - b_{n-1}}{n^s} \\
&= \sum_{n=1}^{N} \frac{1}{n^s}(b_n - b_{n-1}) \\
&= \frac{b_N}{N^s} - \sum_{n=1}^{N-1} b_n \left(\frac{1}{(n+1)^s} - \frac{1}{n^s} \right) \text{ by summation by parts} \\
&= \frac{b_N}{N^s} + \sum_{n=1}^{N-1} b_n \int_n^{n+1} \frac{s}{x^{s+1}} dx \\
&= \frac{b_N}{N^s} + s \sum_{n=1}^{N-1} b_n \int_n^{n+1} \frac{dx}{x^{s+1}}. \qquad (3.3.1)
\end{aligned}$$

If $\sigma = \mathrm{Re}(s) > 0$ then $|b_N/N^s| \leq b/N^\sigma \to 0$ as $N \to \infty$. Also,

$$\left| b_n \int_n^{n+1} \frac{dx}{x^{s+1}} \right| \leq b \int_n^{n+1} \frac{dx}{x^{\sigma+1}} < \frac{b}{n^{\sigma+1}},$$

so the series $\sum_{n\geq 1} b_n \int_n^{n+1} dx/x^{s+1}$ is absolutely convergent when $\sigma > 0$ by comparison with $\sum_{n\geq 1} b/n^{\sigma+1} = b\zeta(\sigma+1)$. Therefore the series in (3.3.1) converges as $N \to \infty$, so $\sum_{n\geq 1} a_n/n^s$ converges when $\mathrm{Re}(s) > 0$. □

Remark 3.3.5. The convergence of $\sum_{n\geq 1} a_n/n^s$ in Theorem 3.3.4 need not be absolute, e.g., the series $1 - 1/2^s + 1/3^s - 1/4^s + \dots$ with alternating coefficients ± 1 converges for $\mathrm{Re}(s) > 0$ but not absolutely for $0 < \mathrm{Re}(s) \leq 1$.

Corollary 3.3.6. *For any nontrivial Dirichlet character χ, the Dirichlet L-series $L(s,\chi) = \sum_{n\geq 1} \chi(n)/n^s$ converges for $\mathrm{Re}(s) > 0$.*

Proof. Let χ be a nontrivial Dirichlet character mod m. Since χ is nontrivial, by Theorem 2.1.8 we know that $\sum_{n=1}^{m} \chi(n) = \sum_{n\in(\mathbf{Z}/m)^\times} \chi(n) = 0$. Furthermore, $\sum_{n=i}^{m+i} \chi(n) = 0$ for all $i \geq 1$. The partial sums of the coefficients in $L(s,\chi)$ are bounded due to their cyclic vanishing, so $L(s,\chi)$ is convergent for nontrivial χ and $\mathrm{Re}(s) > 0$ by Theorem 3.3.4. □

We will need not just convergence of $L(s,\chi)$ in a half-plane, but analyticity too. This is justified by the next theorem.

Theorem 3.3.7. *If the Dirichlet series $\sum_{n\geq 1} a_n/n^s$ converges at $s = s_0$, then it converges in the open half-plane $\{s : \mathrm{Re}(s) > \mathrm{Re}(s_0)\}$, and in fact its partial sums converge uniformly on compact subsets of this half-plane.*

Proof. See [1, p. 235]. □

Corollary 3.3.8. *If the Dirichlet series $\sum_{n\geq 1} a_n/n^s$ converges in the open half-plane $\{s : \mathrm{Re}(s) > \sigma_0\}$, then it is holomorphic here and its derivative can be computed termwise as $\sum_{n\geq 1}(-a_n \log n)/n^s$.*

Proof. For $\varepsilon > 0$, Theorem 3.3.7 with $s_0 = \sigma_0 + \varepsilon$ implies the partial sums of $\sum_n a_n/n^s$ converge uniformly on compact subsets of $\{s : \mathrm{Re}(s) > \sigma_0 + \varepsilon\}$, so Theorem 2.2.5 implies $\sum_n a_n/n^s$ is holomorphic on $\{s : \mathrm{Re}(s) > \sigma_0 + \varepsilon\}$, and its derivative is computable termwise since the partial sums $\sum_{n=1}^N a_n/n^s$ are holomorphic with derivative $\sum_{n=1}^N -a_n(\log n)/n^s$. Since $\{s : \mathrm{Re}(s) > \sigma_0\}$ is the union of $\{s : \mathrm{Re}(s) > \sigma_0 + \varepsilon\}$, we are done. □

Corollary 3.3.9. *The function $\zeta(s) = \sum_{n\geq 1} 1/n^s$ is holomorphic on $\mathrm{Re}(s) > 1$, and the function $L(s,\chi) = \sum_{n\geq 1} \chi(n)/n^s$ is holomorphic on $\mathrm{Re}(s) > 0$ for nontrivial χ.*

Proof. Combine Theorem 3.1.1 and Corollary 3.3.8 to see that $\zeta(s)$ is holomorphic on $\mathrm{Re}(s) > 1$. Combine Corollary 3.3.6 and Corollary 3.3.8 to see that $L(s,\chi)$ is holomorphic on $\mathrm{Re}(s) > 0$ for nontrivial χ. □

Theorem 3.3.10. *For nontrivial χ, a logarithm of $L(s,\chi)$ for $\mathrm{Re}(s) > 1$ is*

$$\sum_{p^k} \frac{\chi(p^k)}{kp^{ks}} = \sum_p \sum_k \frac{(\chi(p)/p^s)^k}{k}. \tag{3.3.2}$$

Proof. Since $|\chi(p^k)/kp^{ks}| \leq 1/p^{k\sigma}$ and $\sum_{p^k} 1/p^{k\sigma}$ is absolutely convergent for $\sigma > 1$, we conclude by the comparison test that the series in (3.3.2) is absolutely convergent for $\sigma > 1$. Exponentiating (3.3.2), we find that

$$\exp\left(\sum_p \sum_k \frac{(\chi(p)/p^s)^k}{k}\right) = \prod_p \exp\left(\sum_k \frac{(\chi(p)/p^s)^k}{k}\right) = \prod_p \frac{1}{1-\chi(p)/p^s} = L(s,\chi)$$

by using Example 2.2.8, the fact that $|\chi(p)/p^s| < 1$ since $\mathrm{Re}(s) > 1$, and the Euler product (3.2.2). □

We can extend $\zeta(s)$ analytically to $\mathrm{Re}(s) > 0$, like $L(s,\chi)$, except there is a pole

at $s = 1$.

Theorem 3.3.11. *The zeta-function has a meromorphic continuation from* $\mathrm{Re}(s) > 1$ *to* $\mathrm{Re}(s) > 0$ *that is holomorphic everywhere except for a simple pole at* $s = 1$ *with residue* 1.

Proof. Consider the two series

$$\zeta_2(s) = 1 - \frac{1}{2^s} + \frac{1}{3^s} - \frac{1}{4^s} + \cdots = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n^s} \text{ and } \zeta_3(s) = 1 + \frac{1}{2^s} - \frac{2}{3^s} + \frac{1}{4^s} + \frac{1}{5^s} - \frac{2}{6^s} + \cdots,$$

where the coefficients in $\zeta_3(s)$ are periodically $1, 1, -2$. By Theorem 3.3.4, these series converge for $\mathrm{Re}(s) > 0$, so they are holomorphic there by Corollary 3.3.8.

Note that

$$\zeta(s) - \zeta_2(s) = \frac{2}{2^s} + \frac{2}{4^s} + \cdots = \frac{2}{2^s}\left(1 + \frac{1}{2^s} + \frac{1}{3^s} + \cdots\right) = \frac{2}{2^s}\zeta(s),$$

so

$$\zeta(s) = \frac{\zeta_2(s)}{1 - 1/2^{s-1}}. \tag{3.3.3}$$

The right side of (3.3.3) is meromorphic for $\mathrm{Re}(s) > 0$ except perhaps for simple poles at $s = 1 + 2k\pi i/\log 2, k \in \mathbf{Z}$, which is where $1/2^{s-1} = 1$.

Similarly,

$$\zeta(s) - \zeta_3(s) = \frac{3}{3^s}\zeta(s), \text{ so } \zeta(s) = \frac{\zeta_3(s)}{1 - 1/3^{s-1}}. \tag{3.3.4}$$

The right side of (3.3.4) is meromorphic for $\mathrm{Re}(s) > 0$ except possibly for simple poles at $s = 1 + 2\ell\pi i/\log 3$ where $\ell \in \mathbf{Z}$.

Since $\zeta(s)$ has two meromorphic continuations to $\mathrm{Re}(s) > 0$, the uniqueness of meromorphic continuation implies that $\zeta(s)$ can only have a pole for $\mathrm{Re}(s) > 0$ when

$s = 1 + 2k\pi i/\log 2 = 1 + 2\ell\pi i/\log 3$ for some integers k and ℓ. These formulas imply $k\log 3 = \ell \log 2$, so $2^\ell = 3^k$. This implies $k = \ell = 0$, so $\zeta(s)$ is holomorphic on $\mathrm{Re}(s) > 0$ away from $s = 1$. At $s = 1$, there is a simple pole since $\zeta(s)$ is meromorphic at 1 and $\lim_{s\to 1^+}(s-1)\zeta(s) = 1$ by (3.1.2), which tells us that the residue of $\zeta(s)$ at $s = 1$ is 1. □

Lemma 3.3.12 (Landau)**.** *Suppose that a Dirichlet series $f(s) = \sum_{n=1}^{\infty} a_n/n^s$ with nonnegative coefficients converges for $\mathrm{Re}(s) > \sigma_0$. If $f(s)$ extends analytically to some disc centered at σ_0, then the Dirichlet series $\sum_{n=1}^{\infty} a_n/n^s$ converges on the half-plane $\mathrm{Re}(s) > \sigma_0 - \varepsilon$ for some $\varepsilon > 0$.*

Proof. Our proof is based on [1, p. 237]. Using $f(s+\sigma_0) = \sum_{n=1}^{\infty} a_n n^{-\sigma_0}/n^s$ in place of $f(s)$, we can assume $\sigma_0 = 0$ and $f(s)$ has an analytic continuation to a neighborhood of 0, so to some disc $B(0,\delta)$. There is a small $\varepsilon < \delta$ such that the disc $\{s : |s-1| < 1+\varepsilon\}$ is entirely inside $B(0,\delta)\bigcup\{s : \mathrm{Re}(s) > 0\}$. By Theorem 2.2.1, the power series of $f(s)$ at $s = 1$ converges to $f(s)$ for any s such that $|s-1| < 1+\varepsilon$.

For $|\sigma - 1| < 1 + \varepsilon$ with $\sigma \in \mathbf{R}$, we have the power series

$$f(\sigma) = \sum_{k=0}^{\infty} \frac{f^{(k)}(1)}{k!}(\sigma - 1)^k.$$

By repeated use of Corollary 3.3.8, we can compute $f^{(k)}(1)$ by differentiating $f(s) = \sum_{n=1}^{\infty} a_n/n^s$ termwise to get

$$f^{(k)}(1) = \sum_{n=1}^{\infty} \frac{a_n(-\log n)^k}{n^s}\bigg|_{s=1} = (-1)^k \sum_{n=1}^{\infty} \frac{a_n(\log n)^k}{n}.$$

Therefore

$$f(\sigma) = (-1)^k \sum_{k=0}^{\infty} \sum_{n=1}^{\infty} \frac{a_n (\log n)^k}{k! n} (\sigma - 1)^k = \sum_{k=0}^{\infty} \sum_{n=1}^{\infty} \frac{a_n (\log n)^k}{k! n} (1 - \sigma)^k. \qquad (3.3.5)$$

When $-\varepsilon < \sigma < 1$, all terms on the right side of (3.3.5) are nonnegative, so this series can be rearranged to

$$f(\sigma) = \sum_{n=1}^{\infty} \frac{a_n}{n} \left(\sum_{k=0}^{\infty} \frac{(\log n)^k}{k!} (1 - \sigma)^k \right) = \sum_{n=1}^{\infty} \frac{a_n}{n} e^{(\log n)(1-\sigma)} = \sum_{n=1}^{\infty} \frac{a_n}{n} n^{1-\sigma} = \sum_{n=0}^{\infty} \frac{a_n}{n^{\sigma}}.$$

Therefore $f(\sigma) = \displaystyle\sum_{n=1}^{\infty} \frac{a_n}{n^{\sigma}}$ when $-\varepsilon < \sigma < 1$, so the Dirichlet series $\displaystyle\sum_{n=1}^{\infty} \frac{a_n}{n^s}$ converges for all $s \in \mathbf{C}$ such that $\mathrm{Re}(s) > -\varepsilon$ by Theorem 3.3.7. □

Chapter 4

Dirichlet's Theorem

In this chapter we will prove Dirichlet's theorem: for two relatively prime integers $a, m \geq 1$, there exist infinitely many primes $p \equiv a \bmod m$. The proof uses the Riemann zeta-function, Dirichlet characters, and Dirichlet L-series. These topics were covered in Chapters 2 and 3. In this chapter, we will first give proofs of elementary cases of Dirichlet's theorem and then discuss a hard theorem about L-series in detail, which is needed to prove Dirichlet's theorem in general.

4.1 Some Elementary Cases

To better appreciate the proof of Dirichlet's theorem, we will work out here some special cases by elementary algebraic methods that do not extend to the general case. Our proofs are based on [5, pp. 123-124].

Theorem 4.1.1. *There are infinitely many primes* $p \equiv 1 \bmod 4$.

Proof. One such prime is 5. If we have finitely many $p \equiv 1 \bmod 4$, say $p_1, \ldots, p_r$, we will construct another one. Let $N = (2p_1 \cdots p_r)^2 + 1$. Clearly neither 2 nor any p_i divides N. Since $N > 1$, it has a prime factor. Let p be some prime dividing N, so p is odd. By Fermat's little theorem, using $a = 2p_1 \cdots p_r$,

$$-1 \equiv a^2 \bmod p \Rightarrow (-1)^{(p-1)/2} \equiv a^{p-1} \equiv 1 \bmod p \Rightarrow (-1)^{(p-1)/2} = 1 \Rightarrow p \equiv 1 \bmod 4.$$

Hence N has a prime factor congruent to 1 mod 4 that is not any p_i. □

Theorem 4.1.2. *There are infinitely many primes $p \equiv 3 \bmod 4$.*

Proof. One such prime is 3. If we have finitely many $p \equiv 3 \bmod 4$, say $p_1, \ldots, p_r$, we will construct another one. Let $N = 4p_1 \cdots p_r - 1$, so $N \equiv -1 \equiv 3 \bmod 4$. Therefore N is odd and, since $N > 1$, at least one of the prime factors of N is congruent to 3 mod 4. (If this were not the case, N would be congruent to 1 mod 4.) Since $N \equiv -1 \bmod p_i$, no p_i divides N. Therefore N has a prime factor congruent to 3 mod 4 that is not any p_i. Iteration of this process produces infinitely many primes $p \equiv 3 \bmod 4$. □

The set of primes congruent to 3 mod 4 created using the above algorithm is probably not all of them. For instance, starting with the prime $p_1 = 3$, the next few primes generated (using the least prime factor of N that is congruent to 3 mod 4 at each step) are $p_2 = 11$, $p_3 = 131$, and $p_4 = 17,291$. Each of these numbers is equal to N in the above process, but that need not be the case (in fact, at the 6^{th} iteration N is composite); the process guarantees that we will find a prime $p \equiv 3 \bmod 4$ dividing N, not that N itself is prime.

Theorem 4.1.3. *There are infinitely many primes $p \equiv 1 \bmod 3$.*

Proof. One such prime is 7. If we have finitely many such primes, say $p_1, \ldots, p_r$, set $N = (2p_1 \cdots p_r)^2 + 3$. Then $N \equiv 1 \bmod 2$ and $N \equiv 1 \bmod 3$, so 2, 3, and p_i do not divide N. Obviously $N > 1$. Let p be a prime factor of N, implying that -3 is a square mod p. We will show that $-3 \equiv \square \bmod p$ implies $p \equiv 1 \bmod 3$.

Write $-3 \equiv b^2 \bmod p$. In $\mathbf{C}$, a cube root of unity is $(-1 + \sqrt{-3})/2$. Analogously, in $\mathbf{Z}/p$ the number $(-1 + b)/2$ cubes to 1:

$$\left(\frac{-1+b}{2}\right)^2 = \frac{1-2b+b^2}{4} = \frac{-1-b}{2} \text{ so } \left(\frac{-1+b}{2}\right)^3 = \frac{-1+b}{2} \cdot \frac{-1-b}{2} = \frac{1-b^2}{4} = 1.$$

And $(-1+b)/2 \not\equiv 1 \bmod p$ since if $(-1+b)/2 \equiv 1 \bmod p$ then $b \equiv 3 \bmod p$, which after squaring would imply $-3 \equiv 9 \bmod p$ so $p|12$, a contradiction. Therefore, $(-1 + b)/2$ has order 3 in $\mathbf{Z}/p$. Since $(\mathbf{Z}/p)^\times$ has size $p - 1$, this implies that $3|(p - 1)$ and hence $p \equiv 1 \bmod 3$.

Infinitely many primes that are 1 mod 3 can be constructed in this manner. □

Theorem 4.1.4. *There are infinitely many primes $p \equiv 2 \bmod 3$.*

Proof. One such prime is 2. If we have finitely many such primes, say $p_1, \ldots, p_r$, set $N = 3p_1 \cdots p_r - 1$. Clearly no p_i divides N. Since $N \equiv -1 \equiv 2 \bmod 3$ and $N > 1$, at least one prime factor p of N must be congruent to 2 mod 3. Since $p \notin \{p_1, \ldots, p_r\}$, infinitely many $p \equiv 2 \bmod 3$ can be added to the list by repeating this process. □

The proofs of these theorems involved creating specific formulas for N such that some prime dividing N satisfies certain conditions. These conditions varied in each case and do not appear to generalize in a uniform way. This suggests that the general method of proving Dirichlet's theorem must be of a very different form.

4.2 Proof of Dirichlet's Theorem

The following technical result is the hardest step in the proof of Dirichlet's theorem.

Theorem 4.2.1. *For any nontrivial Dirichlet character χ, the Dirichlet L-function $L(s,\chi)$ is nonzero at $s = 1$.*

Proof. We will show that $L(1,\chi) \neq 0$ by contradiction. Our argument is adapted from [2].

Set $H(s) = \zeta(s)^2 L(s,\chi)L(s,\overline{\chi})$ for $\mathrm{Re}(s) > 0$. Since $\zeta(s)$ is analytic for $\mathrm{Re}(s) > 0$ except for a simple pole at $s = 1$ by Theorem 3.3.11, and $L(s,\chi)$ and $L(s,\overline{\chi})$ are analytic for $\mathrm{Re}(s) > 0$ by Corollary 3.3.9, $H(s)$ is analytic for $\mathrm{Re}(s) > 0$ except possibly at $s = 1$, where it has at worst a double pole.

We have for $\mathrm{Re}(s) > 0$ that

$$\overline{L(s,\chi)} = \overline{\sum_{n\geq 1} \frac{\chi(n)}{n^s}} = \sum_{n\geq 1} \overline{\left(\frac{\chi(n)}{n^s}\right)} = \sum_{n\geq 1} \frac{\overline{\chi}(n)}{n^{\overline{s}}} = L(\overline{s},\overline{\chi}).$$

Setting $s = 1$, we get $\overline{L(1,\chi)} = L(1,\overline{\chi})$, so

$$L(1,\chi) = 0 \Longrightarrow L(1,\overline{\chi}) = 0.$$

The double pole at $s = 1$ in $\zeta(s)^2$ is therefore canceled by zeros at $s = 1$ of the two L-functions in $H(s)$. Thus $H(s)$ is analytic at $s = 1$, and hence for all s with $\mathrm{Re}(s) > 0$.

For $\mathrm{Re}(s) > 1$, the functions $\zeta(s)$, $L(s,\chi)$, and $L(s,\overline{\chi})$ can be represented by Euler

products:

$$\zeta(s) = \prod_p \frac{1}{1-1/p^s} \ , \ L(s,\chi) = \prod_p \frac{1}{1-\chi(p)/p^s} \ , \ L(s,\overline{\chi}) = \prod_p \frac{1}{1-\overline{\chi}(p)/p^s}. \quad (4.2.1)$$

Because $1/(1-z) = \exp(\sum_{k\geq 1} z^k/k)$ for $|z| < 1$ (Example 2.2.8), we have by (4.2.1) that for $\mathrm{Re}(s) > 1$

$$\begin{aligned} \zeta(s)^2 L(s,\chi)L(s,\overline{\chi}) &= \prod_p \left(\frac{1}{1-1/p^s}\right)^2 \prod_p \left(\frac{1}{1-\chi(p)/p^s}\right) \prod_p \left(\frac{1}{1-\overline{\chi}(p)/p^s}\right) \\ &= \prod_p e^{\sum_{k\geq 1} 2/kp^{ks}} \prod_p e^{\sum_{k\geq 1} \chi(p)^k/kp^{ks}} \prod_p e^{\sum_{k\geq 1} \overline{\chi}(p)^k/kp^{ks}} \\ &= \exp\left(\sum_p \sum_{k\geq 1} \frac{2+\chi(p)^k+\overline{\chi}(p)^k}{kp^{ks}}\right) \\ &= \exp\left(\sum_{p,k} \frac{2+\chi(p)^k+\overline{\chi}(p)^k}{kp^{ks}}\right), \end{aligned} \quad (4.2.2)$$

where k runs over the natural numbers and p over the primes. All series are absolutely convergent on $\mathrm{Re}(s) > 1$, which justifies rearrangements of series above and in what follows.

Look at the Dirichlet series in the exponential of (4.2.2). If p divides the modulus of χ, then the coefficient of $1/p^{ks}$ is $2/k$. If p does not divide the modulus of χ, we can write $\chi(p) = e^{i\theta_p}$. In that case, the coefficient of $1/p^{ks}$ is

$$\frac{2+\chi(p)^k+\overline{\chi}(p)^k}{k} = \frac{2+e^{ik\theta_p}+e^{-ik\theta_p}}{k} = \frac{2+2\cos(k\theta_p)}{k} = \frac{2(1+\cos(k\theta_p))}{k} \geq 0.$$

In both cases, the coefficient of $1/p^{ks}$ is nonnegative, so $H(s)$ is the exponential of a Dirichlet series with nonnegative coefficients that converges (absolutely) on $\mathrm{Re}(s) > 1$.

Since the power series of e^s and $\sum_{p,k}(2+\chi(p)^k+\overline{\chi}(p)^k)/kp^{ks}$ have nonnegative coefficients, we may expand and rearrange (4.2.2) to get a Dirichlet series representation of $H(s)$ with nonnegative coefficients for $\mathrm{Re}(s) > 1$. We will now show that this series in fact converges on the larger half-plane $\mathrm{Re}(s) > 0$.

Let C be the set of $\sigma > 0$ at which $\sum_n a_n/n^\sigma$ converges, so $(1,\infty) \subset C$. Let σ_0 be the infimum of C, so $0 \le \sigma_0 \le 1$. We will show $\sigma_0 = 0$. First we will show $\sum_n a_n/n^s$ converges when $\mathrm{Re}(s) > \sigma_0$: for each such s there is a $\sigma_1 \in C$ such that $\sigma_0 < \sigma_1 < \mathrm{Re}(s)$, so by the comparison test $\mathrm{Re}(s) \in C$, and by absolute convergence the series $\sum_n a_n/n^s$ converges. If $\sigma_0 > 0$, then Landau's lemma (Lemma 3.3.12) tells us that $\sum_n a_n/n^\sigma$ converges slightly to the left of σ_0 since $H(s)$ is analytic for $\mathrm{Re}(s) > 0$. This contradicts σ_0 being the infimum of C, so $\sigma_0 = 0$. Therefore $\sum_n a_n/n^s$ converges and equals $H(s)$ on the half-plane $\mathrm{Re}(s) > 0$.

We will next calculate the coefficient of $1/p^{2s}$ in the Dirichlet series representation of $H(s)$. Writing $H(s) = \sum_{n\ge1} a_n/n^s$, for a fixed prime p the subseries of $H(s)$ over powers of p is

$$\begin{aligned}
\sum_{r\ge0} \frac{a_{p^r}}{p^{rs}} &= \exp\left(\sum_{k\ge1} \frac{2+\chi(p)^k+\overline{\chi}(p)^k}{kp^{ks}}\right) \\
&= \prod_{k\ge1} \exp\left(\frac{2+\chi(p)^k+\overline{\chi}(p)^k}{kp^{ks}}\right) \\
&= \exp\left(\frac{2+\chi(p)+\overline{\chi}(p)}{p^s}\right) \exp\left(\frac{2+\chi(p)^2+\overline{\chi}(p)^2}{2p^{2s}}\right) \exp\left(\frac{2+\chi(p)^3+\overline{\chi}(p)^3}{3p^{3s}}\right)\cdots \\
&= \sum_{j\ge0} \frac{(2+\chi(p)+\overline{\chi}(p))^j}{j!p^{js}} \cdot \sum_{j\ge0} \frac{(2+\chi(p)^2+\overline{\chi}(p)^2)^j}{2^j(j!)p^{2js}} \sum_{j\ge0} \frac{(2+\chi(p)^3+\overline{\chi}(p)^3)^j}{3^j(j!)p^{3js}}\cdots \\
&= \left(1+\frac{2+\chi(p)+\overline{\chi}(p)}{p^s}+\frac{(2+\chi(p)+\overline{\chi}(p))^2}{2p^{2s}}+\cdots\right)\times \\
&\qquad \left(1+\frac{2+\chi(p)^2+\overline{\chi}(p)^2}{2p^{2s}}+\cdots\right)\left(1+\frac{2+\chi(p)^3+\overline{\chi}(p)^3}{3p^{3s}}+\cdots\right)\cdots .
\end{aligned}$$

Expanding products and rearranging terms, the coefficient of $1/p^{2s}$ is

$$\begin{aligned}\frac{(2+\chi(p)+\overline{\chi}(p))^2}{2}+\frac{2+\chi(p)^2+\overline{\chi}(p)^2}{2} &= 3+\chi(p)^2+\overline{\chi}(p)^2+2\chi(p)+2\overline{\chi}(p)+\chi(p)\overline{\chi}(p)\\ &= (\chi(p)+\overline{\chi}(p)+1)^2-\chi(p)\overline{\chi}(p)+2.\end{aligned}$$

Since $\chi(p)+\overline{\chi}(p)=2\mathrm{Re}(\chi(p))\in\mathbf{R}$ and $\chi(p)\overline{\chi}(p)=|\chi(p)|^2$ is 0 (if $\chi(p)=0$) or 1 (if $\chi(p)\neq 0$), we have $-\chi(p)\overline{\chi}(p)+2=2$ or 1, so the coefficient of $1/p^{2s}$ is real and ≥ 1.

Since $a_{p^2}\geq 1$ and each coefficient in the Dirichlet series for $H(s)$ is nonnegative, for real $s>0$ we have $H(s)\geq\sum_p a_{p^2}/p^{2s}\geq\sum_p 1/p^{2s}$. Plugging in $s=1/2$ gives us $H(1/2)\geq\sum_p 1/p$, but this series diverges by Corollary 3.1.6. This is a contradiction to the Dirichlet series for $H(s)$ being convergent for $\mathrm{Re}(s)>0$. We therefore conclude that $L(1,\chi)\neq 0$ for every nontrivial Dirichlet character χ.

□

Corollary 4.2.2. *For any nontrivial χ, $\sum_p \chi(p)/p^s$ converges as $s\to 1^+$.*

Proof. By Theorem 3.3.10, for $\mathrm{Re}(s)>1$ a logarithm of $L(s,\chi)$ is

$$\ell(s,\chi):=\sum_{p,k}\frac{\chi(p)^k}{kp^{ks}}=\sum_p\frac{\chi(p)}{p^s}+\sum_p\sum_{k\geq 2}\frac{\chi(p)^k}{kp^{ks}}. \tag{4.2.3}$$

The second term on the right in (4.2.3) converges absolutely for $\sigma>1/2$ by the proof of Theorem 3.1.4, and it is holomorphic on the half-plane $\sigma>1/2$ by Corollary 3.3.8. From (4.2.3), for $\mathrm{Re}(s)>1$ we may write

$$\sum_p\frac{\chi(p)}{p^s}=\ell(s,\chi)-\sum_p\sum_{k\geq 2}\frac{\chi(p)^k}{kp^{ks}}. \tag{4.2.4}$$

Since $L(s,\chi)\neq 0$ at $s=1$ by Theorem 4.2.1, $L(s,\chi)\neq 0$ near $s=1$ because

$L(s,\chi)$ is continuous there. By Theorem 2.2.7, $L(s,\chi)$ admits a logarithm $\tilde{\ell}(s,\chi)$ on a small open disc around 1. Because the holomorphic regions of $\ell(s,\chi)$ and $\tilde{\ell}(s,\chi)$ overlap in a small region to the right of $s=1$ and Theorem 2.2.7 says that $\tilde{\ell}(s,\chi) - \ell(s,\chi)$ is constant on this region, this implies that $\ell(s,\chi)$ has an analytic continuation to a neighborhood around $s=1$. (We are *not* saying that the series (4.2.3) converges on a neighborhood of 1.) Thus both terms on the the right side of (4.2.4) extend to continuous functions on a neighborhood of $s=1$, so they converge as $s \to 1^+$. Therefore $\sum_p \chi(p)/p^s$ converges as $s \to 1^+$. □

Although it will not be needed for our proof of Dirichlet's theorem, let us extend Theorem 4.2.1 from nonvanishing of $L(s,\chi)$ at $s=1$ to the whole line $\mathrm{Re}(s)=1$.

Theorem 4.2.3. *For nontrivial Dirichlet characters χ, $L(s,\chi) \neq 0$ when $\mathrm{Re}(s)=1$.*

Proof. For $y_0 \in \mathbf{R}$, we will show that $L(1+iy_0,\chi) \neq 0$ by contradiction.

Set $H^*(s) = \zeta(s)^2 L(s+iy_0,\chi)L(s-iy_0,\overline{\chi})$ for $\mathrm{Re}(s) > 0$. Since $L(s,\chi)$ and $L(s,\overline{\chi})$ are analytic for $\mathrm{Re}(s)>0$, $L(s+iy_0,\chi)$ and $L(s-iy_0,\overline{\chi})$ are also analytic for $\mathrm{Re}(s \pm iy_0) = \mathrm{Re}(s) > 0$. Thus $H^*(s)$ is analytic for all $\mathrm{Re}(s)>0$ except possibly at $s=1$.

Assume that $L(1+iy_0,\chi)=0$. For $\mathrm{Re}(s)>0$,

$$\overline{L(s+iy_0,\chi)} = \overline{\sum_{n\geq 1} \frac{\chi(n)}{n^{s+iy_0}}} = \sum_{n\geq 1} \overline{\left(\frac{\chi(n)}{n^{s+iy_0}}\right)} = \sum_{n\geq 1} \frac{\overline{\chi}(n)}{n^{\overline{s}-iy_0}} = L(\overline{s}-iy_0,\overline{\chi}).$$

Therefore, at $s=1$, $L(1-iy_0,\overline{\chi}) = \overline{L(1+iy_0,\chi)} = 0$.

As before, the double pole at $s=1$ in $\zeta(s)^2$ is canceled by the zeros in the two L-functions in $H^*(s)$ at $s=1$, so $H^*(s)$ is analytic at $s=1$ and thus for $\mathrm{Re}(s)>0$.

For $\mathrm{Re}(s) > 1$, $L(s + iy_0, \chi)$ and $L(s - iy_0, \overline{\chi})$ have Euler products

$$L(s + iy_0, \chi) = \prod_p \frac{1}{1 - \chi(p)p^{-iy_0}/p^s} \text{ and } L(s - iy_0, \overline{\chi}) = \prod_p \frac{1}{1 - \overline{\chi}(p)p^{iy_0}/p^s}.$$

This is the same as before except that $\chi(p)p^{-iy_0}$ replaces $\chi(p)$.

Following similar calculations as in the proof of Theorem 4.2.1,

$$H^*(s) = \zeta(s)^2 L(s + iy_0, \chi)L(s - iy_0, \overline{\chi}) = \exp\left(\sum_{p,k} \frac{2 + \left(\chi(p)p^{-iy_0}\right)^k + \left(\overline{\chi}(p)p^{iy_0}\right)^k}{kp^{ks}}\right).$$

In the series inside the exponential, the coefficient of $1/p^{ks}$ is again $2/k$ if $\chi(p) = 0$. If $\chi(p) \neq 0$ then write $\chi(p)p^{-iy_0} = e^{i\theta_{p,y_0}}$. The coefficient of $1/p^{ks}$ is

$$\frac{2 + \left(\chi(p)p^{-iy_0}\right)^k + \left(\overline{\chi}(p)p^{iy_0}\right)^k}{k} = \frac{2(1 + \cos(k\theta_{p,y_0}))}{k} \geq 0.$$

It follows from this nonnegativity that $H^*(s)$ has a Dirichlet series representation with nonnegative coefficients for $\mathrm{Re}(s) > 1$, so by Landau's lemma the Dirichlet series representation of $H^*(s)$ is valid for $\mathrm{Re}(s) > 0$. The coefficient of $1/p^{2s}$ in $H^*(s)$ is

$$\frac{\left(2 + \chi(p)p^{-iy_0} + \overline{\chi}(p)p^{iy_0}\right)^2}{2} + \frac{2 + \left(\chi(p)p^{-iy_0}\right)^2 + \left(\overline{\chi}(p)p^{iy_0}\right)^2}{2} = (z + \overline{z} + 1)^2 - z\overline{z} + 2,$$

where $z = \chi(p)p^{-iy_0}$. Since $|z| = 0$ or 1, this coefficient is real and ≥ 1 as before.

This leads to the conclusion that $H^*(1/2) \geq \sum_p 1/p$ as before, which is a contradiction to the Dirichlet series for $H^*(s)$ being convergent for $\mathrm{Re}(s) > 0$. Therefore $L(1 + iy_0, \chi) \neq 0$ for every nontrivial Dirichlet character χ and real number y_0. $\square$

We are now ready to prove Dirichlet's theorem.

Theorem 4.2.4 (Dirichlet, 1837)**.** *For any relatively prime integers $a, m \geq 1$, there exist infinitely many primes $p \equiv a \bmod m$.*

Proof. Let $s > 1$ (here s is real, not just complex). From (2.1.5), we can write

$$\begin{aligned}
\sum_{p \equiv a \bmod m} \frac{1}{p^s} &= \frac{1}{\varphi(m)} \sum_p \sum_\chi \frac{\chi(p)\overline{\chi}(a)}{p^s} \\
&= \frac{1}{\varphi(m)} \sum_\chi \overline{\chi}(a) \left(\sum_p \frac{\chi(p)}{p^s} \right) \\
&= \frac{1}{\varphi(m)} \sum_{(p,m)=1} \frac{1}{p^s} + \frac{1}{\varphi(m)} \sum_{\chi \neq 1_m} \overline{\chi}(a) \left(\sum_p \frac{\chi(p)}{p^s} \right) \\
&= \frac{1}{\varphi(m)} \sum_p \frac{1}{p^s} - \frac{1}{\varphi(m)} \sum_{p|m} \frac{1}{p^s} + \frac{1}{\varphi(m)} \sum_{\chi \neq 1_m} \overline{\chi}(a) \left(\sum_p \frac{\chi(p)}{p^s} \right).
\end{aligned}$$

Dividing through by $\sum_p 1/p^s$ yields

$$\frac{\sum_{p \equiv a \bmod m} 1/p^s}{\sum_p 1/p^s} = \frac{1}{\varphi(m)} - \frac{1}{\varphi(m)} \cdot \frac{\sum_{p|m} 1/p^s}{\sum_p 1/p^s} + \frac{1}{\varphi(m)} \cdot \frac{\sum_{\chi \neq 1_m} \overline{\chi}(a) \left(\sum_p \chi(p)/p^s \right)}{\sum_p 1/p^s}.$$

Letting $s \to 1^+$, we know that $\sum_p 1/p^s$ tends to infinity by Theorem 3.1.4. Only a finite number of primes divide m, so the second term tends to zero in the limit. The third term on the right also tends to zero by Corollary 4.2.2. Therefore

$$\lim_{s \to 1^+} \frac{\sum_{p \equiv a \bmod m} 1/p^s}{\sum_p 1/p^s} = \frac{1}{\varphi(m)} > 0, \tag{4.2.5}$$

so $\sum_{p \equiv a \bmod m} 1/p^s$ tends to infinity as $s \to 1^+$. Thus there are an infinite number of primes $p \equiv a \bmod m$ for relatively prime a and m. □

For $s > 1$, $1/p^s < 1/p$, so

$$\sum_{p \equiv a \bmod m} \frac{1}{p^s} < \sum_{p \equiv a \bmod m} \frac{1}{p}. \tag{4.2.6}$$

Letting $s \to 1+$, we conclude from (4.2.6) that

$$\sum_{p \equiv a \bmod m} \frac{1}{p} = \infty, \tag{4.2.7}$$

which is an analogue of Corollary 3.1.6. Dirichlet's original idea for proving that the set $\{p : p \equiv a \bmod m\}$ is infinite was to show (4.2.7).

4.3 Dirichlet Density

The proof of Theorem 4.2.4 suggests a method for defining the density of a set of primes within the set of all primes.

Definition 4.3.1. The *Dirichlet density* $d(P)$ of a set of primes P is

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

if the limit exists.

While the Dirichlet density of P is not as intuitive as the natural density of P, which is

$$\lim_{x \to \infty} \frac{\#\{p \leq x : p \in P\}}{\#\{p \leq x\}},$$

it is more computationally accessible. Furthermore, the two are equal when both values exist [6, pp. 256-257].

Proposition 4.3.2. *A finite set of primes has Dirichlet density* 0*. Therefore a set of primes with positive Dirichlet density is infinite.*

Proof. If P is a finite set, then the numerator of the limit converges as $s \to 1^+$. The denominator tends to infinity as $s \to 1^+$. Hence $d(P) = 0$. □

When $(a, m) = 1$, the set $P = \{p : p \equiv a \bmod m\}$ has Dirichlet density $1/\varphi(m)$ by (4.2.5). This density is independent of a, so the primes are "equally distributed" among the congruence classes mod m that are relatively prime to m. Using Theorem 4.2.3 in place of Theorem 4.2.1, it can be shown that $\{p : p \equiv a \bmod m\}$ has natural density $1/\varphi(m)$ when $(a, m) = 1$, which is reflected by the data in Table 1.0.2 for $m = 9$.

Theorem 4.3.3. *Let P_1 and P_2 be sets of prime numbers such that P_1, P_2, and $P_1 \bigcap P_2$ have Dirichlet densities. Then $d(P_1 \bigcup P_2) = d(P_1) + d(P_2) - d(P_1 \bigcap P_2)$. In particular, if P_1 and P_2 have Dirichlet densities and their intersection is finite, then $d(P_1 \bigcup P_2) = d(P_1) + d(P_2)$.*

Proof. By definition,

$$d(P_1 \bigcup P_2) = \lim_{s \to 1^+} \frac{\sum_{p \in (P_1 \bigcup P_2)} p^{-s}}{\sum_p p^{-s}},$$

if the limit exists. For $s > 1$,

$$\begin{aligned} \frac{\sum_{p \in (P_1 \bigcup P_2)} p^{-s}}{\sum_p p^{-s}} &= \frac{\sum_{p \in P_1} p^{-s} + \sum_{p \in P_2} p^{-s} - \sum_{p \in (P_1 \bigcap P_2)} p^{-s}}{\sum_p p^{-s}} \\ &= \frac{\sum_{p \in P_1} p^{-s}}{\sum_p p^{-s}} + \frac{\sum_{p \in P_2} p^{-s}}{\sum_p p^{-s}} - \frac{\sum_{p \in (P_1 \bigcap P_2)} p^{-s}}{\sum_p p^{-s}}. \end{aligned}$$

As $s \to 1^+$, the limit is $d(P_1) + d(P_2) - d(P_1 \bigcap P_2)$. □

Example 4.3.4. Let $a_1, \ldots, a_k$ be relatively prime to m and incongruent mod m. Taking $P = \bigcup_{i=1}^{k} P_i$ where $P_i = \{p : p \equiv a_i \bmod m\}$, Theorem 4.3.3 and induction on k shows that $d(P) = \sum_{i=1}^{k} d(P_i) = k/\varphi(m)$.

Chapter 5

Applications of Dirichlet's Theorem

Quadratic reciprocity, in essence, is the study of the congruence $x^2 \equiv a \bmod p$: we are asking for which $a \in \mathbf{Z}$ and primes p is a a perfect square mod p. For fixed p, one can easily describe all such a. For example, $x^2 \equiv a \bmod 7$ has a nonzero solution if and only if $a \equiv 1$, 2, or 4 mod 7. It is not nearly as simple, however, to describe p with fixed a. For example, how can we characterize all p such that $6 \bmod p$ is a square? This is where interesting results connected to quadratic reciprocity arise. Dirichet's theorem will help us calculate the density of primes modulo which a fixed integer, or finite set of integers, is congruent to a perfect square.

Dirichlet's theorem has many other applications beyond the setting of quadratic reciprocity, which we only mention in passing. It is used in the proof of the classification of quadratic forms over $\mathbf{Q}$ [12, Sect. 2.2 and 3.2], in the determination of torsion points on the elliptic curve $y^2 = x^3 - n^2x$ [8, p. 44], in the calculation of a bound on finite subgroups of $\mathrm{GL}_n(\mathbf{Q})$ [13, Sect. 1.3], and in the construction of the countable

random graph, called the Rado graph [4].

5.1 The Legendre Symbol

Definition 5.1.1. We define the *Legendre symbol* for $a \in \mathbf{Z}$ and odd primes p by

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \equiv x^2 \bmod p \text{ for some } x \in \mathbf{Z} \text{ and } a \not\equiv 0 \bmod p, \\ -1 & \text{if } a \not\equiv x^2 \bmod p \text{ for all } x \in \mathbf{Z}, \\ 0 & \text{if } p|a. \end{cases}$$

If $(\frac{a}{p}) = 1$, we call a a *quadratic residue* mod p. If $(\frac{a}{p}) = -1$, we call a a *quadratic nonresidue* mod p. For example, mod 7 the quadratic residues are $1 \equiv 1^2$, $2 \equiv 3^2$, and $4 \equiv 5^2$ while the quadratic nonresidues are 3, 5, and 6.

We will use the above assumptions on a and p throughout the chapter: $a \in \mathbf{Z}$ and p is an odd prime. Next we present some basic properties of the Legendre symbol.

Lemma 5.1.2. *If $a \equiv b \bmod p$, then $(\frac{a}{p}) = (\frac{b}{p})$.*

Proof. If $p|a$, then $a \equiv 0 \bmod p$ so $b \equiv 0 \bmod p$ and hence $p|b$, so $(\frac{a}{p}) = (\frac{b}{p}) = 0$. Now assume $a \not\equiv 0 \bmod p$. Then a is a quadratic residue if and only if $a \equiv x^2 \bmod p$ for some x. Since $a \equiv b \bmod p$, this implies that $b \equiv x^2 \bmod p$ if and only if $a \equiv x^2 \bmod p$. Hence if $a \equiv b \bmod p$, then a and b are either both quadratic residues or both quadratic nonresidues. □

Lemma 5.1.3. *For prime p and $a \in (\mathbf{Z}/p)^\times$, $a^{(p-1)/2} \equiv 1 \bmod p$ if and only if a is a quadratic residue* mod *p.*

The exponent $(p-1)/2$ is an integer since p is odd.

Proof. Suppose a is a quadratic residue, that is, $a \equiv x^2 \bmod p$ for some $x \in \mathbf{Z}$ and $x \not\equiv 0 \bmod p$. Then

$$a^{(p-1)/2} \equiv (x^2)^{(p-1)/2} \equiv x^{p-1} \equiv 1 \bmod p$$

by Fermat's little theorem.

Since $\mathbf{Z}/p$ is a field, the congruence $a^{(p-1)/2} \equiv 1 \bmod p$ has at most $(p-1)/2$ solutions in $\mathbf{Z}/p$. We saw that every quadratic residue is a solution, so we will next count how many quadratic residues there are mod p. The only solutions to $t^2 \equiv 1 \bmod p$ are $t \equiv \pm 1 \bmod p$, so the squaring homomorphism $(\mathbf{Z}/p)^\times \to (\mathbf{Z}/p)^\times$ has kernel of size 2 and thus image of size $(p-1)/2$ since $|(\mathbf{Z}/p)^\times| = p-1$. Therefore, $(p-1)/2$ elements in $(\mathbf{Z}/p)^\times$ are quadratic residues, so the solutions to $a^{(p-1)/2} \equiv 1 \bmod p$ are the quadratic residues mod p. □

Lemma 5.1.4 (Euler's Criterion)**.** *For all $a \in \mathbf{Z}$, we have $(\frac{a}{p}) \equiv a^{(p-1)/2} \bmod p$.*

Proof. The theorem is trivial if $p|a$: both sides are $0 \bmod p$. Therefore, assume that p does not divide a, so $a \in (\mathbf{Z}/p)^\times$. If a is a quadratic residue mod p, then $a^{(p-1)/2} \equiv 1 \bmod p$ by Lemma 5.1.3 and $(\frac{a}{p}) = 1$ by definition. If a is a quadratic nonresidue mod p, then $a^{(p-1)/2} \not\equiv 1 \bmod p$ by Lemma 5.1.3 and $(\frac{a}{p}) = -1$. Since $(a^{(p-1)/2})^2 \equiv a^{p-1} \equiv 1 \bmod p$ and $a^{(p-1)/2} \not\equiv 1 \bmod p$, we must have $a^{(p-1)/2} \equiv -1 \bmod p$. □

Corollary 5.1.5. *The Legendre symbol is multiplicative. That is, $(\frac{a}{p})(\frac{b}{p}) = (\frac{ab}{p})$ for all integers a and b.*

Proof. By Lemma 5.1.4, $(\frac{a}{p})(\frac{b}{p}) \equiv a^{(p-1)/2}b^{(p-1)/2} \bmod p$ and $(\frac{ab}{p}) \equiv (ab)^{(p-1)/2} \bmod p$. But $a^{(p-1)/2}b^{(p-1)/2} = (ab)^{(p-1)/2}$, so $(\frac{a}{p})(\frac{b}{p}) \equiv (\frac{ab}{p}) \bmod p$. Since 0, 1, and -1 are incongruent mod p, we get $(\frac{a}{p})(\frac{b}{p}) = (\frac{ab}{p})$. □

Lemma 5.1.6. *For all odd primes p, $(\frac{-1}{p}) = (-1)^{(p-1)/2}$.*

Proof. Substituting $a = -1$ into Lemma 5.1.4 yields $(\frac{-1}{p}) \equiv (-1)^{(p-1)/2} \bmod p$. Since $(\frac{-1}{p}) = \pm 1$, $(-1)^{(p-1)/2} = \pm 1$, and $1 \not\equiv -1 \bmod p$, we have $(\frac{-1}{p}) = (-1)^{(p-1)/2}$. □

Lemma 5.1.7. *For all odd primes p, $(\frac{2}{p}) = (-1)^{(p^2-1)/8}$.*

Proof. Our argument is from [10, p. 150]. First note that this is equivalent to 2 being a quadratic residue if $p \equiv 1, 7 \bmod 8$ and a quadratic nonresidue if $p \equiv 3, 5 \bmod 8$. That is what we will show.

Suppose $p \equiv 1 \bmod 4$. We then have

$$\begin{aligned}
2^{(p-1)/2}\left(\frac{p-1}{2}\right)! &= 2^{(p-1)/2} \cdot 1 \cdot 2 \cdot 3 \cdots \left(\frac{p-1}{2}\right) \\
&= 2 \cdot 4 \cdot 6 \cdots (p-1) \\
&= 2 \cdot 4 \cdot 6 \cdots \left(\frac{p-1}{2}\right)\left(\frac{p+3}{2}\right) \cdots (p-3)(p-1) \\
&\equiv 2 \cdot 4 \cdot 6 \cdots \left(\frac{p-1}{2}\right)\left(\frac{3-p}{2}\right) \cdots (-3)(-1) \bmod p \\
&\equiv 2 \cdot 4 \cdot 6 \cdots \left(\frac{p-1}{2}\right)\left(\frac{p-3}{2}\right) \cdots (3)(1)(-1)^{(p-1)/4} \bmod p \\
&\equiv 1 \cdot 2 \cdot 3 \cdots \left(\frac{p-1}{2}\right)(-1)^{(p-1)/4} \bmod p \\
&\equiv (-1)^{(p-1)/4}\left(\frac{p-1}{2}\right)! \bmod p.
\end{aligned}$$

Because $\left(\frac{p-1}{2}\right)!$ is a product of invertible numbers mod p, it is invertible mod p. Therefore $2^{(p-1)/2} \equiv (-1)^{(p-1)/4} \bmod p$. Plugging in $p \equiv 1 \bmod 8$ with $p = 8k + 1$ yields $2^{(p-1)/2} \equiv (-1)^{2k} \equiv 1 \bmod p$, whereas plugging in $p \equiv 5 \bmod 8$ with $p = 8k + 5$ yields $2^{(p-1)/2} \equiv (-1)^{2k+1} \equiv -1 \bmod p$. Thus 2 is a quadratic residue when $p \equiv 1 \bmod 8$, but is a quadratic nonresidue when $p \equiv 5 \bmod 8$.

We can work similarly for $p \equiv 3 \bmod 4$. Computation shows that

$$\begin{aligned} 2^{(p-1)/2}\left(\frac{p-1}{2}\right)! &= 2 \cdot 4 \cdot 6 \cdots \left(\frac{p-3}{2}\right)\left(\frac{p+1}{2}\right)\cdots(p-3)(p-1) \\ &\equiv 2 \cdot 4 \cdot 6 \cdots \left(\frac{p-3}{2}\right)\left(\frac{1-p}{2}\right)\cdots(-3)(-1) \bmod p \\ &\equiv 1 \cdot 2 \cdot 3 \cdots \left(\frac{p-1}{2}\right)(-1)^{(p+1)/4} \bmod p \\ &\equiv (-1)^{(p+1)/4}\left(\frac{p-1}{2}\right)! \bmod p. \end{aligned}$$

Here we find that $2^{(p-1)/2} \equiv (-1)^{(p+1)/4} \bmod p$. If we plug in $p \equiv 3 \bmod 8$, we find that $2^{(p-1)/2} \equiv -1 \bmod p$. On the other hand, if $p \equiv 7 \bmod 8$ then $2^{(p-1)/2} \equiv 1 \bmod p$. □

Lemmas 5.1.6 and 5.1.7 are called the supplementary laws of quadratic reciprocity. We now present without proof the main law of quadratic reciprocity.

Theorem 5.1.8. *For distinct odd primes p and q, $\left(\frac{q}{p}\right) = (-1)^{((p-1)/2)\cdot((q-1)/2)}\left(\frac{p}{q}\right)$. Equivalently,*

$$\left(\frac{q}{p}\right) = \begin{cases} \left(\frac{p}{q}\right) & \text{if } p \text{ or } q \equiv 1 \bmod 4, \\ -\left(\frac{p}{q}\right) & \text{if } p \text{ and } q \equiv 3 \bmod 4. \end{cases}$$

Proof. See [7, pp. 58-60]. □

To demonstrate the usefulness of quadratic reciprocity, we will now compute some examples.

Example 5.1.9. Is $87 = \square \bmod 61$? Since $87 \equiv 26 \bmod 61$, by Lemma 5.1.2 we can reduce this to $\left(\frac{87}{61}\right) = \left(\frac{26}{61}\right) = \left(\frac{2}{61}\right)\left(\frac{13}{61}\right)$, where the second equality holds by Corollary 5.1.5. Since $13 \equiv 1 \bmod 4$, $\left(\frac{13}{61}\right) = \left(\frac{61}{13}\right) = \left(\frac{9}{13}\right)$ by Theorem 5.1.8. Clearly $\left(\frac{9}{13}\right) = 1$

because $9 = 3^2$. By Lemma 5.1.7, $(\frac{2}{61}) = -1$ since $61 \equiv 5 \bmod 8$. Hence $(\frac{87}{61}) = (\frac{2}{61})(\frac{9}{13}) = -1 \cdot 1 = -1$, so 87 is not a square mod 61.

Example 5.1.10. Is $33 = \square \bmod 97$? By Corollary 5.1.5, $(\frac{33}{97}) = (\frac{3}{97})(\frac{11}{97})$. We then find that because $97 \equiv 1 \bmod 4$,

$$\left(\frac{3}{97}\right) = \left(\frac{97}{3}\right) = \left(\frac{1}{3}\right) = 1 \text{ and } \left(\frac{11}{97}\right) = \left(\frac{97}{11}\right) = \left(\frac{9}{11}\right) = 1.$$

Hence 33 is a square mod 97. Quadratic reciprocity tells us that the congruence $33 \equiv x^2 \bmod 97$ has a solution. It does *not* tell us what a solution is. From an explicit search, $33 \equiv 324 \equiv 18^2 \bmod 97$.

Example 5.1.11. Find all primes $p \neq 2, 3$ such that $6 \equiv \square \bmod p$. Write $(\frac{6}{p}) = 1$ as $(\frac{2}{p})(\frac{3}{p}) = 1$. This is satisfied if either 2 and 3 are both quadratic residues mod p or if both are quadratic nonresidues mod p. The number $(\frac{2}{p})$ is completely determined by $p \bmod 8$. Since

$$\left(\frac{3}{p}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{3-1}{2}} \left(\frac{p}{3}\right) = (-1)^{\frac{p-1}{2}} \left(\frac{p}{3}\right),$$

where $(-1)^{(p-1)/2}$ is determined by $p \bmod 4$ and $(\frac{p}{3})$ is determined by $p \bmod 3$, $(\frac{3}{p})$ is determined by $p \bmod 12$. So $(\frac{6}{p})$ is determined by $p \bmod 24$.

Below we present a table of values of $(\frac{6}{p})$ for $p \bmod 24 \in (\mathbf{Z}/24)^\times$, and from the bottom row we see $(\frac{6}{p}) = 1 \iff p \equiv 1, 5, 19, 23 \bmod 24$.

p mod 24	1	5	7	11	13	17	19	23
Sample p	73	5	7	11	13	17	19	23
$(\frac{2}{p})$	1	-1	1	-1	-1	1	-1	1
$(-1)^{(p-1)/2}$	1	1	-1	-1	1	1	-1	-1
$(\frac{p}{3})$	1	-1	1	-1	1	-1	1	-1
$(\frac{6}{p})$	1	1	-1	-1	-1	-1	1	1

5.2 Statistics of One Legendre Symbol

Our first application of Dirichlet's theorem to the behavior of the Legendre symbol is a proof that when $a \in \mathbf{Z}$ is not a square, $(\frac{a}{p}) = -1$ infinitely often. It depends on the following property of the Legendre symbol.

Lemma 5.2.1. *For any nonzero integer a, if $p \equiv 1 \bmod 4a$ then $(\frac{a}{p}) = 1$.*

Proof. We induct on $|a|$. If $a = 1$ the result is trivial, and if $a = -1$ the result is clear from the formula $(\frac{-1}{p}) = (-1)^{(p-1)/2}$.

If $|a| > 1$, then a has a prime factor, say q. Write $a = qa'$, so $(\frac{a}{p}) = (\frac{q}{p})(\frac{a'}{p})$. Since $|a'| < |a|$ and $p \equiv 1 \bmod 4a'$, by induction $(\frac{a'}{p}) = 1$. We also have $p \equiv 1 \bmod 4q$, and we will derive $(\frac{q}{p}) = 1$ from this using quadratic reciprocity.

If $q = 2$ then $p \equiv 1 \bmod 8$, so $(\frac{q}{p}) = (\frac{2}{p}) = 1$ by the supplementary law of quadratic reciprocity.

If q is odd then

$$\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}} \left(\frac{p}{q}\right),$$

and we have $(-1)^{\frac{p-1}{2}} = 1$ from $p \equiv 1 \bmod 4$, while $(\frac{p}{q}) = 1$ from $p \equiv 1 \bmod q$. $\square$

Theorem 5.2.2. *For any nonzero integer a that is not a perfect square, there exist infinitely many odd primes p such that $(\frac{a}{p}) = -1$.*

Proof. First we will assume that a is squarefree. If $a = -1$, use any of the infinitely many primes $p \equiv 3 \bmod 4$, which exist either by Dirichlet's theorem or by the elementary method in Theorem 4.1.2. If $a \neq -1$ then a has a prime factor q. Write $a = qa'$, so $(q, a') = 1$.

If a is odd, so $q \neq 2$, pick an integer x such that $x \bmod q$ is not a square. By the Chinese remainder theorem, there exists an integer m satisfying

$$m \equiv \begin{cases} x \bmod q, \\ 1 \bmod 4a', \end{cases} \tag{5.2.1}$$

and $(m, 4qa') = 1$ since $(x, q) = 1$. By Dirichlet's theorem, there are infinitely many primes p such that $p \equiv m \bmod 4a$ (note that $4a = 4a'q$). For any such prime p, $(\frac{a}{p}) = (\frac{q}{p})(\frac{a'}{p}) = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}(\frac{p}{q})(\frac{a'}{p})$. Since $p \equiv 1 \bmod 4$, the first factor is 1. The Legendre symbol $(\frac{p}{q})$ equals $(\frac{x}{q}) = -1$, and the Legendre symbol $(\frac{a'}{p})$ is 1 by Lemma 5.2.1 since $p \equiv 1 \bmod 4a'$, so $(\frac{a}{p}) = (1)(-1)(1) = -1$.

If a is even, let $q = 2$, so a' is odd since a is assumed to be squarefree. By the Chinese remainder theorem, there exists an integer m satisfying

$$m \equiv \begin{cases} 5 \bmod 8, \\ 1 \bmod a', \end{cases} \tag{5.2.2}$$

and $(m, 8a') = 1$. By Dirichlet's theorem there are infinitely many primes p such that $p \equiv m \bmod 4a$ (note that $4a = 8a'$). For any such p, $(\frac{a}{p}) = (\frac{q}{p})(\frac{a'}{p}) = (\frac{2}{p})(\frac{a'}{p})$. Since

$p \equiv 5 \bmod 8$, $(\frac{2}{p}) = -1$ by Lemma 5.1.7. Since $p \equiv 1 \bmod 4$ and $p \equiv 1 \bmod a'$, we have $p \equiv 1 \bmod 4a'$, so $(\frac{a'}{p}) = 1$ by Lemma 5.2.1. Thus $(\frac{a}{p}) = (-1)(1) = -1$.

If a is not squarefree, we may write $a = bc^2$ where b is squarefree and $b \neq 1$. Then $(\frac{a}{p}) = (\frac{b}{p})(\frac{c}{p})^2$. If p does not divide c, then $(\frac{c}{p})^2 = 1$ and $(\frac{a}{p}) = (\frac{b}{p})$. From the squarefree case, for all odd p satisfying a particular congruence condition mod $4b$ we have $(\frac{b}{p}) = -1$. Infinitely many such p are guaranteed to exist by Dirichlet's theorem, and infinitely many such p do not divide c. For these p, $(\frac{a}{p}) = (\frac{b}{p}) = -1$. □

Theorem 5.2.2 can be proved without Dirichlet's theorem [7, pp. 57-58].

Corollary 5.2.3. *A nonzero integer a is a perfect square if and only if $(\frac{a}{p}) = 1$ for all but finitely many odd primes p.*

Proof. If a is not a perfect square, then by Theorem 5.2.2 there are infinitely many odd primes p such that $(\frac{a}{p}) = -1$, so contrapositively if $(\frac{a}{p}) = 1$ for all but finitely many odd primes p then a is a perfect square.

Conversely, suppose a is a perfect square. Then we can write $a = k^2$ for some integer k and it follows that $(\frac{a}{p}) = (\frac{k^2}{p}) = (\frac{k}{p})^2$. If p does not divide a, clearly p does not divide k, so $(\frac{a}{p}) = (\frac{k}{p})^2 = (\pm 1)^2 = 1$. □

Our next application of Dirichlet's theorem, at the end of this section, is a proof that when $a \in \mathbf{Z}$ is not a perfect square, $(\frac{a}{p}) = -1$ for half of the primes p in the sense of Dirichlet density. To show this, we will first need to show that the Legendre symbol $(\frac{a}{p})$ for fixed nonzero a and varying odd prime p can be interpreted as a Dirichlet character on the group $(\mathbf{Z}/4a)^\times$. In order to do this, we need to build up some of the mechanical properties of the Legendre symbol that follow from quadratic reciprocity.

Lemma 5.2.4. *For odd primes p, q, and r, if $pq \equiv r \bmod 4$, then $(\frac{-1}{p})(\frac{-1}{q}) = (\frac{-1}{r})$.*

Proof. The conclusion says $(-1)^{(p-1)/2}(-1)^{(q-1)/2} = (-1)^{(r-1)/2}$, or equivalently

$$\frac{p-1}{2} + \frac{q-1}{2} \equiv \frac{r-1}{2} \bmod 2.$$

Multiplying through by 2, that congruence is the same as $(p-1)+(q-1) \equiv r-1 \bmod 4$. Since $pq \equiv r \bmod 4$, we can rewrite $(p-1)+(q-1) \equiv r-1 \bmod 4$ as $(p-1)+(q-1) \equiv pq-1 \bmod 4$, and after rearranging terms and factoring, this becomes $(p-1)(q-1) \stackrel{?}{\equiv} 0 \bmod 4$. Since p and q are odd, $p-1$ and $q-1$ are even, so we are done. □

Lemma 5.2.5. *For odd primes p, q, and r, if $pq \equiv r \bmod 8$, then $(\frac{2}{p})(\frac{2}{q}) = (\frac{2}{r})$.*

Proof. First note by the supplementary law of quadratic reciprocity that

$$\left(\frac{2}{p}\right)\left(\frac{2}{q}\right) = (-1)^{(p^2-1)/8+(q^2-1)/8} \text{ and } \left(\frac{2}{r}\right) = (-1)^{(r^2-1)/8}.$$

Thus our claim is true if and only if $\frac{p^2-1}{8} + \frac{q^2-1}{8} \equiv \frac{r^2-1}{8} \bmod 2$, which is equivalent to

$$(p^2-1)+(q^2-1) \equiv r^2-1 \bmod 16. \tag{5.2.3}$$

For integers a and b, if $a \equiv b \bmod 2^m$ then $a^2 \equiv b^2 \bmod 2^{m+1}$: from $a-b \equiv 0 \bmod 2^m$ and $a+b \equiv 0 \bmod 2$, $a^2-b^2 = (a-b)(a+b) \equiv 0 \bmod 2^{m+1}$ so $a^2 \equiv b^2 \bmod 2^{m+1}$. Taking $a = pq$, $b = r$, and $m = 3$, we get $p^2q^2 \equiv r^2 \bmod 16$. That makes (5.2.3) the same as

$$(p^2-1)+(q^2-1) \equiv p^2q^2-1 \bmod 16,$$

which is equivalent to $p^2q^2 - p^2 - q^2 + 1 \equiv 0 \bmod 16$, and the left side can be put in factored form:

$$(p^2 - 1)(q^2 - 1) \stackrel{?}{\equiv} 0 \bmod 16.$$

For any odd number a, $a^2 \equiv 1 \bmod 8$ (since $a^2 - 1 = (a+1)(a-1)$, with one factor being divisible by 2 and the other by 4). Thus $(p^2 - 1)(q^2 - 1)$ is divisible by 64, and thus by 16. □

Theorem 5.2.6. *Assume that a is a nonzero integer and that p, q, and r are odd primes not dividing a. If $pq \equiv r \bmod 4a$, then $(\frac{a}{p})(\frac{a}{q}) = (\frac{a}{r})$.*

Remark 5.2.7. The preceeding lemmas are the special cases $a = -1$ and $a = 2$. The theorem would make no sense if $a = 0$ since the condition $pq \equiv r \bmod 4a$ would become $pq \equiv r \bmod 0$, *i.e.*, $pq = r$, which is false.

Proof. Suppose a has a square factor, say $a = bc^2$. Then $(\frac{a}{p}) = (\frac{b}{p})(\frac{c}{p})^2$. Since p does not divide a, and thus does not divide c, $(\frac{c}{p})^2 = (\pm 1)^2 = 1$, so $(\frac{a}{p}) = (\frac{b}{p})$. Likewise $(\frac{a}{q}) = (\frac{b}{q})$ and $(\frac{a}{r}) = (\frac{b}{r})$. Moreover, from $pq \equiv r \bmod 4a$ we get $pq \equiv r \bmod 4b$. Therefore we can replace a with b throughout and thus assume without loss of generality that a is squarefree.

Let a have factorization $\varepsilon\ell_1\ell_2\cdots\ell_n$, where $\varepsilon = \pm 1$ and the ℓ_i are distinct primes. Then

$$\left(\frac{a}{p}\right)\left(\frac{a}{q}\right) = \left(\frac{\varepsilon}{p}\right)\left(\frac{\ell_1}{p}\right)\cdots\left(\frac{\ell_n}{p}\right)\left(\frac{\varepsilon}{q}\right)\left(\frac{\ell_1}{q}\right)\cdots\left(\frac{\ell_n}{q}\right)$$

and

$$\left(\frac{a}{r}\right) = \left(\frac{\varepsilon}{r}\right)\left(\frac{\ell_1}{r}\right)\cdots\left(\frac{\ell_n}{r}\right),$$

so the equation $(\frac{a}{p})(\frac{a}{q}) = (\frac{a}{r})$ would follow from

$$\left(\frac{\varepsilon}{p}\right)\left(\frac{\varepsilon}{q}\right) = \left(\frac{\varepsilon}{r}\right) \quad \text{and} \quad \left(\frac{\ell_i}{p}\right)\left(\frac{\ell_i}{q}\right) = \left(\frac{\ell_i}{r}\right) \quad \text{for each } \ell_i. \tag{5.2.4}$$

The first equation in (5.2.4) is obvious when $\varepsilon = 1$ and it is Lemma 5.2.4 when $\varepsilon = -1$.

To prove the second equation in (5.2.4), we can use Lemma 5.2.5 if $\ell_i = 2$. If $\ell_i \neq 2$, the main law of quadratic reciprocity tells us that

$$\left(\frac{\ell_i}{p}\right)\left(\frac{\ell_i}{q}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{\ell_i-1}{2}}\left(\frac{p}{\ell_i}\right)(-1)^{\frac{q-1}{2}\cdot\frac{\ell_i-1}{2}}\left(\frac{q}{\ell_i}\right) = (-1)^{(\frac{p-1}{2}+\frac{q-1}{2})\frac{\ell_i-1}{2}}\left(\frac{pq}{\ell_i}\right) \tag{5.2.5}$$

and

$$\left(\frac{\ell_i}{r}\right) = (-1)^{\frac{r-1}{2}\cdot\frac{\ell_i-1}{2}}\left(\frac{r}{\ell_i}\right). \tag{5.2.6}$$

Since $pq \equiv r \bmod 4a$ implies $pq \equiv r \bmod 4$, $(-1)^{(\frac{p-1}{2}+\frac{q-1}{2})\frac{\ell_i-1}{2}} = (-1)^{\frac{r-1}{2}\cdot\frac{\ell_i-1}{2}}$ by Lemma 5.2.4 (raise both sides of $(\frac{-1}{p})(\frac{-1}{q}) = (\frac{-1}{r})$ to the power $\frac{\ell_i-1}{2}$). Since $pq \equiv r \bmod 4a$ implies $pq \equiv r \bmod \ell_i$ too, $(\frac{pq}{\ell_i}) = (\frac{r}{\ell_i})$. □

Now we are ready to interpret $(\frac{a}{p})$ as a Dirichlet character of the denominator.

Theorem 5.2.8. *For a fixed nonzero $a \in \mathbf{Z}$, define $\chi : (\mathbf{Z}/4a)^\times \to \{\pm 1\}$ by $\chi(m) = (\frac{a}{\ell})$ for any odd prime ℓ such that $\ell \equiv m \bmod 4a$. Then χ is a well-defined group homomorphism. Furthermore, χ is nontrivial when a is not a perfect square.*

Proof. By Dirichlet's theorem there exist odd prime representatives for each of the congruence classes in $(\mathbf{Z}/4a)^\times$, and the Legendre symbol $(\frac{a}{\ell})$ makes sense with any odd prime ℓ. When ℓ does not divide $2a$, $(\frac{a}{\ell})$ is ± 1, not 0.

We must ensure that $\chi(m)$ is independent of the prime ℓ such that $\ell \equiv m \bmod 4a$. That is, if for odd primes ℓ and p such that $\ell \equiv p \bmod 4a$, we need to check that $(\frac{a}{\ell}) = (\frac{a}{p})$. Let a have the factorization $a = \varepsilon q_1 q_2 \cdots q_n$ where $\varepsilon = \pm 1$ and the q_i are (not necessarily distinct) primes.

Case 1: $a > 0$ and a is odd. Thus $\varepsilon = 1$. Computation shows that

$$\left(\frac{a}{\ell}\right) = \prod_{i=1}^{n} \left(\frac{q_i}{\ell}\right) = (-1)^{\frac{\ell-1}{2}\sum_{i=1}^{n}\frac{q_i-1}{2}} \prod_{i=1}^{n} \left(\frac{\ell}{q_i}\right)$$

and

$$\left(\frac{a}{p}\right) = \prod_{i=1}^{n} \left(\frac{q_i}{p}\right) = (-1)^{\frac{p-1}{2}\sum_{i=1}^{n}\frac{q_i-1}{2}} \prod_{i=1}^{n} \left(\frac{p}{q_i}\right).$$

For each i, $(\frac{\ell}{q_i}) = (\frac{p}{q_i})$ because $\ell \equiv p \bmod a$, hence $\ell \equiv p \bmod q_i$. Also, $(-1)^{(\ell-1)/2} = (-1)^{(p-1)/2}$ because $\ell \equiv p \bmod 4$. Therefore, $(\frac{a}{\ell}) = (\frac{a}{p})$ when $\ell \equiv p \bmod 4a$.

Case 2: $a > 0$ and a is even. Write $a = 2^e a'$, where $e \geq 1$ and a' is odd. Set $q_1 = 2$ and $a' = q_2 \cdots q_n$. Then $(\frac{a}{\ell}) = (\frac{2}{\ell})^e (\frac{a'}{\ell})$ and $(\frac{a}{p}) = (\frac{2}{p})^e (\frac{a'}{p})$. By Case 1, $(\frac{a'}{\ell}) = (\frac{a'}{p})$ since a' is odd and positive and $\ell \equiv p \bmod 4a'$. It remains to check that $(\frac{2}{\ell}) = (\frac{2}{p})$. That means that we require $(-1)^{(\ell^2-1)/8} = (-1)^{(p^2-1)/8}$, or equivalently $\ell^2 \equiv p^2 \bmod 16$. Since $\ell \equiv p \bmod 4a$ and $2|a$, we have $\ell \equiv p \bmod 8$. By the proof of Lemma 5.2.5, we conclude that $\ell^2 \equiv p^2 \bmod 16$. (Alternatively, $\ell \equiv p \bmod 8 \Rightarrow (\frac{2}{\ell}) = (\frac{2}{p})$ by Lemma 5.1.7.)

Case 3: $a < 0$. Write $a = -a'$. Then $\ell \equiv p \bmod 4a'$, so $(\frac{a'}{\ell}) = (\frac{a'}{p})$ from Cases 1 and 2. Therefore showing that $(\frac{a}{\ell}) = (\frac{a}{p})$ reduces to showing that $(\frac{-1}{\ell}) = (\frac{-1}{p})$, or equivalently that $(-1)^{(\ell-1)/2} = (-1)^{(p-1)/2}$, which is true because $\ell \equiv p \bmod 4$.

To be a group homomorphism, we must have $\chi(mn) = \chi(m)\chi(n)$ for all $m, n \in (\mathbf{Z}/4a)^\times$. Choose odd primes ℓ, p, and r such that $\ell \equiv m \bmod 4a$, $p \equiv n \bmod 4a$, and

$r \equiv mn \bmod 4a$. Then $\ell p \equiv mn \equiv r \bmod 4a$. We have $\chi(mn) = (\frac{a}{r})$ and $\chi(m)\chi(n) = (\frac{a}{\ell})(\frac{a}{p})$, so $\chi(mn) = \chi(m)\chi(n)$ by Theorem 5.2.6.

The function χ is nontrivial when a is not a perfect square by Theorem 5.2.2. □

Theorem 5.2.9. *If a is not a perfect square, then the sets of primes $\{p : (\frac{a}{p}) = 1\}$ and $\{p : (\frac{a}{p}) = -1\}$ both have Dirichlet density $1/2$.*

Proof. Let χ be the Dirichlet character mod $4a$ from Theorem 5.2.8, so $\chi(p) = (\frac{a}{p})$ if p does not divide $4a$. By Theorem 5.2.2, $\chi : (\mathbf{Z}/4a)^\times \to \{\pm 1\}$ is surjective. Since χ is nontrivial, we know from group theory that the kernel of χ has index 2. Therefore, half of all elements of $(\mathbf{Z}/4a)^\times$ are mapped to 1 by χ and the other half to -1. That is, the condition $(\frac{a}{p}) = 1$ is equivalent to p lying in half of the congruence classes of $(\mathbf{Z}/4a)^\times$. By Example 4.3.4, each of these sets has Dirichlet density $1/2$. □

Remark 5.2.10. Using the natural density version of Dirichlet's theorem, these two sets of primes also have natural density $1/2$.

5.3 Statistics of Multiple Legendre Symbols

A set of nonzero integers $a_1, \ldots, a_r$ is called *independent mod squares* if no product of the a_i's is a perfect square. For example, $\{12, 21, 27\}$ is not independent mod squares since $12 \cdot 27 = 18^2$, but $\{6, 13, 35\}$ is: none of 6, 13, 35, $6 \cdot 13$, $6 \cdot 35$, $13 \cdot 35$, $6 \cdot 13 \cdot 35$ is a square. The goal of this section is to prove the following:

Theorem 5.3.1. *Given any integers $a_1, \ldots, a_r$ that are independent mod squares and any $\varepsilon_1, \ldots, \varepsilon_r \in \{\pm 1\}$, there exist infinitely many odd primes p such that $(\frac{a_i}{p}) = \varepsilon_i$ for $i = 1, \ldots, r$.*

This says that each $(\frac{a_i}{p})$ is "independent" of the other Legendre symbols $(\frac{a_j}{p})$ as p varies. Later we will show that the Dirichlet density of p such that $(\frac{a_i}{p}) = \varepsilon_i$ for $i = 1, \ldots, r$ is $1/2^r$, which is independent of the choice of prescribed ε_i values.

The case $r = 1$ of Theorem 5.3.1 follows from Lemma 5.2.1 (for $(\frac{a_1}{p}) = 1$) and Theorem 5.2.2 (for $(\frac{a_1}{p}) = -1$).

Lemma 5.3.2. *Let $a \in \mathbf{Z} - \{0, 1, -1\}$ be squarefree and $d|a$ with $|d| < |a|$. For each $\varepsilon \in \{\pm 1\}$, there are infinitely many primes p such that $p \equiv 1 \bmod 4d$ and $(\frac{a}{p}) = \varepsilon$.*

Proof. Since $|d| < |a|$, there is a prime ℓ dividing a/d. Write $a = \ell a'$, so $d|a'$ and $(\frac{a}{p}) = (\frac{\ell}{p})(\frac{a'}{p})$ for all odd primes p. Since a is squarefree, ℓ does not divide a'.

Case 1: $\ell \neq 2$. For any odd prime $p \neq \ell$, $(\frac{a}{p}) = (-1)^{(p-1)/2 \cdot (\ell-1)/2}(\frac{p}{\ell})(\frac{a'}{p})$.

Choose an integer c such that $(\frac{c}{\ell}) = \varepsilon$. By the Chinese remainder theorem, there is an integer k such that $k \equiv c \bmod \ell$ and $k \equiv 1 \bmod 4a'$ since ℓ and $4a'$ are relatively prime. Dirichlet's theorem then tells us that there are infinitely many primes p such that $p \equiv k \bmod 4a$. (Note $4a = 4a\ell'$.) For such primes, $p \equiv k \equiv 1 \bmod 4a'$, so $(\frac{a'}{p}) = 1$ by Lemma 5.2.1 and $(\frac{a}{p}) = (-1)^{(p-1)/2 \cdot (\ell-1)/2}(\frac{p}{\ell})(\frac{a'}{p}) = 1 \cdot (\frac{c}{\ell}) \cdot 1 = \varepsilon$. We have $p \equiv 1 \bmod 4d$ since $d|a'$.

Case 2: $\ell = 2$. We have $a = 2a'$ with a' odd. For any odd prime p, $(\frac{a}{p}) = (\frac{2}{p})(\frac{a'}{p})$.

If $\varepsilon = 1$, let $k \in \mathbf{Z}$ satisfy $k \equiv 1 \bmod 8$ and $k \equiv 1 \bmod a'$, while if $\varepsilon = -1$ let $k \in \mathbf{Z}$ satisfy $k \equiv 5 \bmod 8$ and $k \equiv 1 \bmod a'$. For every prime p satisfying $p \equiv k \bmod 8$, we have $(\frac{2}{p}) = \varepsilon$. Dirichlet's theorem tells us that there are infinitely many $p \equiv k \bmod 4a$. For these p we have $p \equiv k \bmod 8$ since $2|a$, and also $p \equiv 1 \bmod 4a'$, so $(\frac{a}{p}) = (\frac{2}{p})(\frac{a'}{p}) = \varepsilon \cdot 1 = \varepsilon$. We have $p \equiv 1 \bmod 4d$ since $k \equiv 1 \bmod 4$ and $d|a'$. $\square$

To prove the case $r \geq 2$ in Theorem 5.3.1, we will use the notion of multiplicatively independent elements in a group, which was defined in Definition 2.1.9.

Theorem 5.3.3. *Suppose a set of nonzero integers $a_1, \ldots, a_r$ is independent mod squares. Then the Dirichlet characters $\chi_i : (\mathbf{Z}/4a_1 \cdots a_r)^\times \to S^1$, where $\chi_i(p) = (\frac{a_i}{p})$ for primes p not dividing $4a_1 \cdots a_r$, are multiplicatively independent in the group of characters of $(\mathbf{Z}/4a_1 \cdots a_r)^\times$.*

Proof. Since each a_i is not a square, each χ_i has order 2 by Theorem 5.2.8. Therefore if $\{\chi_1, \ldots, \chi_r\}$ were not multiplicatively independent, we have $\chi_i = \prod_{j \in S} \chi_j$ for some i and some $S \subset \{1, \ldots, r\} - \{i\}$. Then $(\frac{a_i}{p}) = (\frac{\prod_{j \in S} a_j}{p})$ for all primes p not dividing $4a_1 \cdots a_r$, so $(\frac{a_i \prod a_j}{p}) = 1$ for all such p. Therefore, $a_i \prod_{j \in S} a_j$ is a perfect square by Corollary 5.2.3, a contradiction of the independence of the a_i's mod squares. □

Theorem 5.3.4. *Let $a_1, \ldots, a_r$ be nonzero integers that are independent mod squares. For $\varepsilon_1, \ldots, \varepsilon_r \in \{\pm 1\}$, there are infinitely many odd primes p such that $(\frac{a_i}{p}) = \varepsilon_i$ for each i.*

Proof. Let $\chi_i \colon (\mathbf{Z}/4a_1 \cdots a_r)^\times \to \{\pm 1\}$ by $\chi_i(n) = (\frac{a_i}{p})$ for prime $p \equiv n \bmod 4a_1 \cdots a_r$. By Theorem 5.3.3, since the a_i are independent mod squares the characters $\chi_1, \ldots, \chi_r$ are multiplicatively independent in the group of characters of $(\mathbf{Z}/4a_1 \cdots a_r)^\times$. We want to apply Theorem 2.1.12, using for H in that theorem the group of characters of $(\mathbf{Z}/4a_1 \cdots a_r)^\times$. Let us first see how $\widehat{H}$ can be viewed as $(\mathbf{Z}/4a_1 \cdots a_r)^\times$.

For any finite abelian group G, the group $\widehat{\widehat{G}}$ can be interpreted as G: to each $g \in G$ we have the evaluation mapping ev_g where $ev_g : \widehat{G} \to S^1$ by $ev_g(\chi) = \chi(g)$. Each evaluation is a homomorphism since $ev_g(\chi_1\chi_2) = (\chi_1\chi_2)(g) = \chi_1(g)\chi_2(g) = ev_g(\chi_1)ev_g(\chi_2)$. Thus $ev_g \in \widehat{\widehat{G}}$ for all $g \in G$. The mapping $\psi : G \to \widehat{\widehat{G}}$ by $\psi(g) = ev_g$ is a homomorphism: to show that $\psi(g_1g_2) = \psi(g_1)\psi(g_2)$, i.e. $ev_{g_1g_2} = ev_{g_1}ev_{g_2}$ on $\widehat{G}$, pick $\chi \in \widehat{G}$. Then $ev_{g_1g_2}(\chi) = \chi(g_1g_2) = \chi(g_1)\chi(g_2) = ev_{g_1}(\chi)ev_{g_2}(\chi) = (ev_{g_1}ev_{g_2})(\chi)$. Furthermore, by Theorem 2.1.5, for any two distinct $g, h \in G$, there exists some $\chi \in \widehat{G}$

such that $\chi(g) \neq \chi(h)$, so $ev_g(\chi) \neq ev_h(\chi)$ which means ψ is injective. By Theorem 2.1.6, $|G| = |\widehat{G}|$ and $|\widehat{G}| = |\widehat{\widehat{G}}|$, so $|G| = |\widehat{\widehat{G}}|$. Therefore, $G \cong \widehat{\widehat{G}}$ by $g \mapsto ev_g$.

In Theorem 2.1.12, let H be the group of characters of $(\mathbf{Z}/4a_1 \cdots a_r)^\times$, so $\widehat{H} = (\mathbf{Z}/4a_1 \cdots a_r)^\times$, and let $h_i = \chi_i$. Each χ_i has order 2. Since $\varepsilon_i^2 = 1$, by Theorem 2.1.12 there exists an $n \in (\mathbf{Z}/4a_1 \cdots a_r)^\times$ such that $\chi_i(n) = \varepsilon_i$ for every i. By Dirichlet's theorem there are infinitely many prime $p \equiv n \bmod 4a_1 \cdots a_r$, so $\chi_i(p) = \varepsilon_i$ for all i, which says $(\frac{a_i}{p}) = \varepsilon_i$ for all i. □

Theorem 5.3.5. *Let $a_1, \ldots, a_r$ be nonzero integers that are independent mod squares. For $\varepsilon_1, \ldots, \varepsilon_r \in \{\pm 1\}$, the Dirichlet density of the set of primes p such that $(\frac{a_1}{p}) = \varepsilon_1, \ldots, (\frac{a_r}{p}) = \varepsilon_r$ is $1/2^r$.*

Proof. Let $\chi_i : (\mathbf{Z}/4a_1 \cdots a_r)^\times \to \{\pm 1\}$ by $\chi_i(p) = (\frac{a_i}{p})$ for primes p not dividing $4a_1 \cdots a_r$. Direct computation shows that

$$\begin{aligned} \left\{p : \left(\frac{a_1}{p}\right) = \varepsilon_1, \ldots, \left(\frac{a_r}{p}\right) = \varepsilon_r\right\} &= \{p \bmod 4a_1 \cdots a_r : \chi_i(p) = \varepsilon_i \text{ for } 1 \leq i \leq r\} \\ &= \left\{p \bmod 4a_1 \cdots a_r \in f^{-1}(\varepsilon_1, \ldots, \varepsilon_r)\right\}, \qquad (5.3.1) \end{aligned}$$

where $f : (\mathbf{Z}/4a_1 \cdots a_r)^\times \to \{\pm 1\}^r$ by $f(n) = (\chi_1(n), \ldots, \chi_r(n))$. This is a homomorphism. The function f is surjective by Theorem 5.3.4, so the set of primes p such that $f(p) = (\varepsilon_1, \ldots, \varepsilon_r)$ is equal to a coset of the kernel $K = \ker(f)$ in $G = (\mathbf{Z}/4a_1 \cdots a_r)^\times$, say gK. In particular, since f is surjective, there are 2^r cosets of K in G. Using Example 4.3.4, the set of primes in (5.3.1) has Dirichlet density

$$\frac{|gK|}{|G|} = \frac{|K|}{|G|} = \frac{1}{|G|/|K|} = \frac{1}{[G : K]} = \frac{1}{2^r}.$$

□

Remark 5.3.6. Using the natural density version of Dirichlet's theorem, the set of primes in Theorem 5.3.5 has natural density $1/2^r$.

Example 5.3.7. Let $r = 3$ with $a_1 = 6$, $a_2 = 13$, and $a_3 = 35$. Below we provide a table counting the number of primes $p \leq N$ such that $(\frac{6}{p})$, $(\frac{13}{p})$, and $(\frac{35}{p})$ respectively take the values ε_1, ε_2, ε_3, labeled by the vector $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$. The corresponding proportions in Table 5.3.2 have limit $1/2^3 = 1/8 = .125$ as $N \to \infty$.

$(\varepsilon_1, \varepsilon_2, \varepsilon_3)$	$N:$	10^2	10^3	10^4	10^5	10^6
$(1,1,1)$		2	19	143	1178	9722
$(1,1,-1)$		1	17	160	1201	9850
$(1,-1,1)$		4	17	154	1214	9829
$(1,-1,-1)$		2	26	149	1192	9809
$(-1,1,1)$		1	19	149	1197	9921
$(-1,1,-1)$		2	24	157	1209	9746
$(-1,-1,1)$		2	21	169	1190	9777
$(-1,-1,-1)$		5	20	143	1206	9839

TABLE 5.3.1: Number of $p \leq N$ where $(\frac{6}{p}) = \varepsilon_1$, $(\frac{13}{p}) = \varepsilon_2$, $(\frac{35}{p}) = \varepsilon_3$.

$(\varepsilon_1, \varepsilon_2, \varepsilon_3)$	$N:$	10^2	10^3	10^4	10^5	10^6
$(1,1,1)$		.1200	.1131	.1164	.1228	.1239
$(1,1,-1)$		.0400	.1012	.1302	.1252	.1255
$(1,-1,1)$		.1600	.1012	.1253	.1266	.1252
$(1,-1,-1)$		.0800	.1548	.1212	.1243	.1250
$(-1,1,1)$		.0400	.1131	.1212	.1248	.1264
$(-1,1,-1)$		.0800	.1429	.1277	.1260	.1242
$(-1,-1,1)$		.0800	.1250	.1375	.1241	.1246
$(-1,-1,-1)$		.2000	.1190	.1164	.1257	.1253

TABLE 5.3.2: Proportion of $p \leq N$ where $(\frac{6}{p}) = \varepsilon_1$, $(\frac{13}{p}) = \varepsilon_2$, $(\frac{35}{p}) = \varepsilon_3$.

Bibliography

[1] T. Apostol, *Introduction to Analytic Number Theory*, Springer-Verlag, New York, 1976.

[2] P. T. Bateman, *Theorem of Ingham implying that Dirichlet's L-functions have no zeros with real part one*, L'Enseignement Mathématique, 43 (1997), pp. 281–284.

[3] J. Beachy and W. Blair, *Abstract Algebra*, Waveland Press, Inc., Illinois, 3rd. ed., 2006.

[4] P. Cameron, *The Random Graph*, The Mathematics of Paul Erdős, II, 14 (1997), pp. 333–351.

[5] K. Conrad, *Dirichlet L-functions, Generalizations, and Applications*. Undergraduate lecture notes at Park City Mathematics Institute, 2009.

[6] R. Descombes, *Éléments de Théorie des Nombres*, Presses Universitaires de France, 1986.

[7] K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory*, Springer-Verlag, New York, 1982.

[8] N. Koblitz, *Introduction to Elliptic Curves and Modular Forms*, Springer-Verlag, New York, 1984.

[9] F. Lemmermeyer, *Reciprocity Laws: From Euler to Eisenstein*, Springer-Verlag, New York, 1976.

[10] Á. Lozano-Robledo, *Number Theory.* Class Notes on Undergraduate Number Theory.

[11] W. Rudin, *Real and Complex Analysis*, McGraw-Hill Book Company, New York, 3rd. ed., 1987.

[12] J.-P. Serre, *A Course in Arithmetic*, Springer-Verlag, New York, 1973.

[13] ——, *Bounds for the Orders of the Finite Subgroups of* $G(k)$, in Group Representation Theory, EPFL Press, Lausanne, 2007, pp. 405–450.

[14] E. Stein and R. Shakarchi, *Complex Analysis*, Princeton University Press, Princeton, 2003.

[15] W. R. Wade, *An Introduction to Analysis*, Pearson Prentice Hall, New Jersey, 4th ed., 2010.