191: Analytic Number Theory

Nir Elber

Spring 2023

Contents

How strange to actually have to see the path of your journey in order to make it.

—Neal Shusterman, [Shu16]

Theme 1
Arithmetic Progressions

1.1 January 18

Here we go.

1.1.1 House-Keeping

We're teaching analytic number theory. Here are some notes.

- We will be referencing [Dav80] mostly, but we will do some things that Davenport does not do. For example, we will discuss the circle method, for which we refer to [Dav05].
- We will assume complex analysis, at the level of Math 185. We will use some Fourier analysis, but we will discuss the relevant parts as we need them. Of course, because this is number theory, we will assume some algebra, such as characters on abelian groups.
- There is a website here, which includes a list of topics. Notably, there is a website for a previous version of the course.
- Grading is still up in there, as is the syllabus. Tentatively, grading will be as follows: by around the middle of the semester, there will be a list of recommended papers to read. Then we will write a 2–6-page report and present it to Professor Zhang. We will not have problem sets.
- Tentatively, office hours will be 90 minutes before lecture on Monday and Wednesday, in Evans 813.
- We should all write an email to Professor Zhang to introduce ourselves; for example, say what you're looking forward to in the course.

1.1.2 Facts about Dirichlet Series

In this first part of the course, we will be moving towards the following result.

Theorem 1.1 (Dirichlet)**.** Fix nonzero integers $a, q \in \mathbb{Z}$ such that $\gcd(a, q) = 1$. Then there exist infinitely many primes p such that $p \equiv a \pmod{q}$.

The statement of Theorem 1.1 is purely elementary, but the standard proof uses complex analysis.

The functions we will do analysis on are generalizations of the Riemann ζ function, defined as

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

which converges absolutely for $s \in \mathbb{C}$ such that $\operatorname{Re} s > 1$. Indeed, we can show this.

Lemma 1.2. Fix some open, connected subset $U \subseteq \mathbb{C}$ and some function $f : U \to \mathbb{C}$. Given holomorphic functions $f_n : U \to \mathbb{C}$ for each $n \in \mathbb{N}$, if $f_n \to f$ uniformly on all compact subsets $D \subseteq U$, then f is holomorphic.

Proof. The point is to use Morera's theorem. Each f_n is continuous, so we see f is continuous as well. Thus, fixing any closed piecewise C^1 path $\gamma : [0,1] \to U$, we would like to show

$$\oint_\gamma f(z)\,dz \stackrel{?}{=} 0.$$

Note $\operatorname{im}\gamma$ is compact, so $f_n \to f$ uniformly on $\operatorname{im}\gamma$. Thus, fixing any $\varepsilon > 0$, we can find some N such that

$$|f(z) - f_n(z)| < \varepsilon$$

for all $n > N$. Fixing any $n > N$, we find

$$\left|\oint_\gamma f(z)\,dz\right| = \left|\oint_\gamma f(z)\,dz - \oint_\gamma f_n(z)\,dz\right| \le \oint_\gamma |f(z) - f_n(z)|\,dz \le \varepsilon \ell(\gamma),$$

where $\ell(\gamma)$ is the length of γ. (Note $\ell(\gamma)$ is finite because γ is piecewise C^1.) Sending $\varepsilon \to 0^+$ finishes the proof. ∎

Proposition 1.3. Let $f : \mathbb{N} \to \mathbb{C}$ denote a sequence of complex numbers such that $|f(n)| = O(n^\sigma)$ for some $\sigma \in \mathbb{R}$. Then the series

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

converges absolutely for $s \in \mathbb{C}$ such that $\operatorname{Re} s > \sigma + 1$. Thus, $D(s)$ defines a holomorphic function in this region.

Proof. We are given $|f(n)| \le Cn^\sigma$ for some $C > 0$. Thus, showing the absolute convergence is direct: note

$$\sum_{n=1}^{\infty} \left|\frac{f(n)}{n^s}\right| \le C \sum_{n=1}^{\infty} \frac{1}{n^{\operatorname{Re}(s)-\sigma}},$$

which converges because $\operatorname{Re}(s) - \sigma > 1$.

We can now convert absolute convergence to uniform convergence of the partial sums $\{D_n\}_{n\in\mathbb{N}}$ of D, from which Lemma 1.2 will finish. Fix some compact subset $D \subseteq U$, and we want to show $D_n \to D$ uniformly on D. Because D is compact, there exists $s_0 \in D$ with minimal $\operatorname{Re} s_0$; define $\sigma_0 \coloneqq \operatorname{Re} s_0$. Now, the series

$$\sum_{n=1}^{\infty} \frac{|f(n)|}{n^{\sigma_0}}$$

converges by our absolute convergence.

As such, for any $\varepsilon > 0$, select N such that $n_0 > N$ implies

$$\sum_{n>n_0} \frac{|f(n)|}{n^{\sigma_0}} < \varepsilon.$$

Thus, for any $s \in \mathbb{C}$ and $n_0 > N$, we see

$$|D(s) - D_{n_0}(s)| = \left|\sum_{n>n_0} \frac{f(n)}{n^s}\right| \le \sum_{n>n_0} \frac{|f(n)|}{n^{\operatorname{Re} s}} \le \sum_{n>n_0} \frac{|f(n)|}{n^{\sigma_0}} < \varepsilon,$$

which is what we wanted. ∎

It follows from Proposition 1.3 that $\zeta(s)$ defines a holomorphic function on $\operatorname{Re} s > 1$.

1.1.3 The Euler Product

The following factorization is due to Euler.

Definition 1.4 (multiplicative)**.** Let $f\colon \mathbb{N} \to \mathbb{C}$ be a function. Then f is *multiplicative* if and only if $f(nm) = f(n)f(m)$ for any $n, m \in \mathbb{N}$ such that $\gcd(n, m) = 1$.

Proposition 1.5. Let $f\colon \mathbb{N} \to \mathbb{C}$ be a multiplicative function such that $|f(n)| = O(n^\sigma)$. For any $s \in \mathbb{C}$ such that $\operatorname{Re} s > \sigma + 1$, we have

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_{p \text{ prime}} \left(\sum_{k=0}^{\infty} \frac{f\left(p^k\right)}{p^{ks}} \right).$$

Proof. Fix $s \in \mathbb{C}$ with $\operatorname{Re} s > \sigma + 1$. Roughly speaking, this follows from unique prime factorization in $\mathbb{Z}$. For and N and M to be fixed later, define

$$P_{N,M} \coloneqq \prod_{p<N} \left(\sum_{k=0}^{M-1} \frac{f\left(p^k\right)}{p^{ks}} \right),$$

and define $P_{N,\infty}$ analogously. Define $A_{N,M}$ to be the set of integers n such that the prime factorization of n includes primes less than N each to a power less than M, and define $A_{N,\infty}$ analogously. Note $A_{N,M}$ is a finite set, so the distributive law implies

$$P_{N,M} = \sum_{n \in A_{N,M}} \frac{f(n)}{n^s}.$$

To begin, we fix N and claim

$$P_{N,\infty} \stackrel{?}{=} \sum_{n \in A_{N,\infty}} \frac{f(n)}{n^s}.$$

Note $P_{N,\infty} = \lim_{M \to \infty} P_{N,M}$, so we fix some $M > 0$ and compute

$$\left| P_{N,M} - \sum_{n \in A_{N,\infty}} \frac{f(n)}{n^s} \right| = \left| \sum_{n \in A_{N,\infty} \setminus A_{N,M}} \frac{f(n)}{n^s} \right| \le \sum_{n \notin A_{N,M}} \left| \frac{f(n)}{n^s} \right|.$$

Now, the smallest n such that $n \notin A_{N,M}$ is at least 2^M, so we see

$$\left| P_{N,M} - \sum_{n \in A_{N,\infty}} \frac{f(n)}{n^s} \right| \le \sum_{n \ge 2^M} \left| \frac{f(n)}{n^s} \right|,$$

which now vanishes as $M \to \infty$ because $\sum_{n=1}^{\infty} f(n)/n^s$ converges absolutely by Proposition 1.3. This completes the proof of the claim.

We now send $N \to \infty$ to finish the proof. For any $N > 0$, we use the claim to note

$$\left| P_{N,\infty} - \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \right| = \left| \sum_{n \notin A_{N,\infty}} \frac{f(n)}{n^s} \right| \le \sum_{n \notin A_{N,\infty}} \left| \frac{f(n)}{n^s} \right|.$$

Now, we note that the smallest $n \notin A_{N,\infty}$ is at least N because any $n < N$ has a prime factor less than N, so

$$\left| P_{N,\infty} - \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \right| \le \sum_{n \ge N} \left| \frac{f(n)}{n^s} \right|,$$

and now we see that the right-hand side goes to 0 as $N \to \infty$ because $\sum_{n=1}^{\infty} f(n)/n^s$ converges absolutely by Proposition 1.3. The proposition follows. ∎

Corollary 1.6. We have
$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ prime}} \frac{1}{1-p^{-s}}.$$

Proof. By Proposition 1.5, we see
$$\zeta(s) = \prod_{p \text{ prime}} \left(\sum_{k=0}^{\infty} \frac{1}{p^{ks}} \right) = \prod_{p \text{ prime}} \frac{1}{1-p^{-s}},$$
which is what we wanted. ∎

We can now use Corollary 1.6 to give a proof of the infinitude of primes.

Theorem 1.7. There are infinitely many primes. In fact,
$$\sum_{p \text{ prime}} \frac{1}{p} = +\infty.$$

Proof. Throughout the proof, s will be a real number greater than 1. The key estimate is to note
$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} \geq \int_1^{\infty} x^{-s}\,dx = -\frac{1}{1-s},$$
which goes to $+\infty$ as $s \to 1^+$. In particular, $\log \zeta(s) \to +\infty$ as $s \to 1^+$.

The last ingredient we need is to bound the Euler product of Corollary 1.6. In particular, we see
$$\log \zeta(s) = \log \left(\prod_{p \text{ prime}} \frac{1}{1-p^{-s}} \right) = \sum_{p \text{ prime}} -\log\left(1-p^{-s}\right).$$
(Formally, one should cap the number of factors and then send the number of factors to infinity.) Using the Taylor expansion of $-\log(1-x)$, we now see
$$\log \zeta(s) = \sum_{p \text{ prime}} \left(\sum_{k=1}^{\infty} \frac{1}{kp^{ks}} \right) = \left(\sum_{p \text{ prime}} \frac{1}{p^s} \right) + \sum_{p \text{ prime}} \left(\sum_{k=2}^{\infty} \frac{1}{kp^{ks}} \right).$$
We would like to focus on $\sum_p 1/p^s$, so we quickly show that the other sum converges. All terms are positive, so it suffices to show that it is bounded above, for which we see
$$\sum_{p \text{ prime}} \left(\sum_{k=2}^{\infty} \frac{1}{kp^{ks}} \right) \leq \sum_{p \text{ prime}} \left(\sum_{k=2}^{\infty} \frac{1}{p^k} \right) = \sum_{p \text{ prime}} \frac{1/p^2}{1-1/p} \leq \sum_{n=2}^{\infty} \frac{1}{n(n-1)} = \sum_{n=2}^{\infty} \left(\frac{1}{n-1} - \frac{1}{n} \right) = 1,$$
where we have telescoped in the last equality. Letting the value of this sum be $S(s)$, we see
$$\log \zeta(s) - S(s) = \sum_{p \text{ prime}} \frac{1}{p^s} < \sum_{p \text{ prime}} \frac{1}{p}.$$
Now, as $s \to 1^+$, we see $\log \zeta(s) - S(s) \to +\infty$, so the theorem follows. ∎

The proof of Theorem 1.1 more or less imitates the argument of Theorem 1.7. Roughly speaking, we will show that
$$\sum_{\substack{p \text{ prime} \\ p \equiv a \pmod q}} \frac{1}{p} = +\infty,$$
from which our infinitude follows. Finding a way to extract out the equivalence class $a \pmod q$ will use a little character theory.

1.1.4 Characters

Throughout, our groups will be finite and abelian, and actually we will be most interested in the abelian groups $\mathbb{Z}/n\mathbb{Z}$ and $(\mathbb{Z}/n\mathbb{Z})^\times$ for integers n. Formally, here is our definition.

Definition 1.8. Fix a positive integer n. Then we define $(\mathbb{Z}/n\mathbb{Z})^\times$ as the units in $\mathbb{Z}/n\mathbb{Z}$, which is $\{a \pmod n : \gcd(a,n) = 1\}$.

Remark 1.9. It is a fact that $(\mathbb{Z}/p\mathbb{Z})^\times$ is cyclic for any prime p. This is nontrivial to prove, but we will not show it here.

Notably, given a prime factorization $n = \prod_{p|n} p^{\nu_p(n)}$, there is an isomorphism of rings

$$\mathbb{Z}/n\mathbb{Z} \cong \prod_{p|n} \left(\mathbb{Z}/p^{\nu_p(n)}\right)$$

and hence also an isomorphism of multiplicative groups, upon taking units.

Having said all that, the theory is most cleanly build working with general finite abelian groups.

Definition 1.10 (dual group)**.** Let G be a group. Then the *dual group* is $\widehat{G} \coloneqq \operatorname{Hom}(G, \mathbb{C}^\times)$, where the operation is pointwise. Its elements are called *characters*.

Notation 1.11 (principal character)**.** There is a "trivial" character $1 \colon G \to \mathbb{C}^\times$ sending $g \mapsto 1$, which is the identity. We might call 1 the *principal character*; we might also denote 1 by χ_0.

Notation 1.12 (conjugate character)**.** If $\chi \colon G \to \mathbb{C}^\times$ is a character, then note that $\overline{\chi} \colon G \to \mathbb{C}^\times$ defined by $\overline{\chi}(g) \coloneqq \overline{\chi(g)}$ is also a character. Indeed, conjugation is a field homomorphism.

Remark 1.13. If G is a finite group, we note that any $\chi \in \widehat{G}$ and $g \in G$ has

$$\chi(g)^{\#G} = \chi\left(g^{\#G}\right) = 1,$$

so $\chi(g)$ is a $(\#G)$th root of unity. In particular, $|\chi(g)| = 1$, so $\overline{\chi(g)} = \chi(g)^{-1} = \chi\left(g^{-1}\right)$.

It will be helpful to have the following notation.

Notation 1.14. We might write $e \colon \mathbb{C} \to \mathbb{C}$ for the function $e(z) \coloneqq \exp(2\pi i z)$.

We now begin computing $\widehat{G}$ for finite abelian groups.

Lemma 1.15. Suppose G and H are groups. Then $\widehat{G} \times \widehat{H} \cong \widehat{G \times H}$ by sending (χ_G, χ_H) to $(g,h) \mapsto \chi_G(g)\chi_H(g)$.

Proof. We have the following checks. Let e_G and e_H be the identities of G and H, respectively.

- Well-defined: given $(\chi_G, \chi_H) \in \widehat{G} \times \widehat{H}$, define $\varphi(\chi_G, \chi_H) \colon G \times H \to \mathbb{C}^\times$ by $\varphi(\chi_G, \chi_H) \colon (g,h) \mapsto \chi_G(g)\chi_H(h)$. Note $\varphi(\chi_G, \chi_H)$ is a homomorphism: we have

$$\begin{aligned}
\varphi(\chi_G, \chi_H)((g,h)\cdot(g',h')) &= \varphi(\chi_G, \chi_H)(gg', hh') \\
&= \chi_G(gg')\chi_H(hh') \\
&= \chi_G(g)\chi_H(h)\chi_G(g')\chi_H(h') \\
&= \varphi(\chi_G, \chi_H)(g,h)\cdot\varphi(\chi_G, \chi_H)(g',h').
\end{aligned}$$

- Homomorphism: to show φ is a homomorphism, we have
$$\varphi((\chi_G,\chi_H)\cdot(\chi_G',\chi_H'))(g,h)=\chi_G(g)\chi_G'(g)\chi_H(h)\chi_H'(h)=\varphi(\chi_G,\chi_H)(g,h)\cdot\varphi(\chi_G',\chi_H')(g,h),$$
so $\varphi((\chi_G,\chi_H)\cdot(\chi_G',\chi_H'))=\varphi(\chi_G,\chi_H)\cdot\varphi(\chi_G',\chi_H')$.
- Injective: if $\varphi(\chi_G,\chi_H)=1$, then
$$\chi_G(g)\chi_H(h)=\varphi(\chi_G,\chi_H)(g,h)=1$$
for all $g\in G$ and $h\in H$. Setting $g=e_G$ shows that $\chi_H=1$, and similarly setting $h=e_H$ shows that $\chi_G=1$. Thus, $(\chi_G,\chi_H)=(1,1)$.
- Surjective: given a character $\chi\colon (G\times H)\to\mathbb{C}^\times$, define $\chi_G(g)\coloneqq\chi(g,e_H)$ and $\chi_H(h)\coloneqq\chi(e_G,h)$. Note χ_G is a character because
$$\chi_G(gg')=\chi(gg',e_H)=\chi(g,e_H)\chi(g',e_H)=\chi_G(g)\chi_G(g').$$
Switching the roles of G and H shows that χ_H is also a character. Lastly, we note $\varphi(\chi_G,\chi_H)=\chi$ because
$$\varphi(\chi_G,\chi_H)(g,h)=\chi(g,e_H)\chi(e_G,h)=\chi(g,h).$$
This completes the proof. ∎

Lemma 1.16. Suppose $G=\mathbb{Z}/n\mathbb{Z}$ for a positive integer n. Then $\chi_\bullet\colon\mathbb{Z}/n\mathbb{Z}\cong\widehat{G}$ by sending $[k]$ to the character $\chi_k\colon[\ell]\mapsto e(k\ell/n)$.

Proof. To begin, note $\chi_k\colon\mathbb{Z}\to\mathbb{C}^\times$ defines a homomorphism because
$$\chi_k(\ell+\ell')=e\left(\frac{k(\ell+\ell')}{n}\right)=e\left(\frac{k\ell}{n}\right)e\left(\frac{k\ell'}{n}\right)=\chi_k(\ell)\chi_k(\ell').$$
Further, note $\chi_k(n\ell)=e(k\ell)=1$ for any $n\ell\in\mathbb{Z}$, so $n\mathbb{Z}\subseteq\ker\chi_k$. It follows that χ_k produces a homomorphism $\chi_k\colon G\to\mathbb{C}^\times$.

We now note that $\chi_\bullet\colon\mathbb{Z}\to\widehat{G}$ defines a homomorphism: for any $[\ell]\in G$, we see
$$\chi_{k+k'}([\ell])=e\left(\frac{(k+k')\ell}{n}\right)=e\left(\frac{k\ell}{n}\right)e\left(\frac{k'\ell}{n}\right)=\chi_k([\ell])\chi_{k'}([\ell]).$$
Additionally, $\chi_{nk}([\ell])=e(k\ell)=1$, so $\chi_{nk}=1$, so $nk\in\ker\chi_\bullet$. It follows that $\chi_\bullet$ produces a homomorphism $\chi_\bullet\colon\mathbb{Z}/n\mathbb{Z}\to\widehat{G}$.

It remains to show that $\chi_\bullet$ is a bijection. We have two checks.

- Injective: suppose $\chi_k=1$ for $k\in\mathbb{Z}$. We must show $k\in n\mathbb{Z}$. Well, we must then have
$$1=\chi_k([1])=e(k/n),$$
which forces $n\mid k$.
- Surjective: given some character $\chi\colon G\to\mathbb{C}^\times$, we note $\chi([1])^n=\chi([0])=1$, so $\chi([1])$ is an nth root of unity. Thus, there exists k such that $\chi([1])=e(k/n)=\chi_k([1])$. Thus, for any $\ell\in\{0,1,\dots,n-1\}$, we see
$$\chi([\ell])=\chi(\underbrace{[1]+\cdots+[1]}_{\ell})=\underbrace{\chi([1])\cdot\ldots\cdot\chi([1])}_{\ell}=\underbrace{\chi_k([1])\cdot\ldots\cdot\chi_k([1])}_{\ell}=\chi_k([\ell]),$$
so $\chi=\chi_k$ follows. ∎

Proposition 1.17. Let G be a finite abelian group. Then $G \cong \widehat{G}$.

Proof. By the Fundamental theorem of finitely generated abelian groups, we may write

$$G \cong \prod_{i=1}^{n} \mathbb{Z}/n_i\mathbb{Z}$$

for some positive integers n_i. Thus, using Lemma 1.15 and Lemma 1.16, we compute

$$\widehat{G} \cong \widehat{\left(\prod_{i=1}^{n} \mathbb{Z}/n_i\mathbb{Z}\right)} = \prod_{i=1}^{n} \widehat{\mathbb{Z}/n_i\mathbb{Z}} \cong \prod_{i=1}^{n} \mathbb{Z}/n_i\mathbb{Z} \cong G,$$

which is what we wanted. ∎

Proposition 1.17 might look like we now understand dual groups perfectly, but the isomorphism given there is non-canonical because the isomorphism of Lemma 1.16 is non-canonical. In other words, given some $g \in G$, there is in general no good way to produce character $\chi \in \widehat{G}$.

However, there is a natural map $G \to \widehat{\widehat{G}}$ which is an isomorphism.

Proposition 1.18. Fix a finite abelian group G. Define the map $\mathrm{ev}_\bullet \colon G \to \widehat{\widehat{G}}$ by sending $g \in G$ to the map $\mathrm{ev}_g \in \widehat{\widehat{G}}$ defined by $\mathrm{ev}_g \colon \chi \mapsto \chi(g)$. Then $\mathrm{ev}_\bullet$ is an isomorphism.

Proof. We begin by checking that $\mathrm{ev}_\bullet$ is a well-defined homomorphism. For each $g \in G$, we see $\mathrm{ev}_g \colon \widehat{G} \to \mathbb{C}^\times$ is a homomorphism because

$$\mathrm{ev}_g(\chi\chi') = \chi(g)\chi(g') = \mathrm{ev}_g(\chi)\,\mathrm{ev}_g(\chi').$$

Further, $\mathrm{ev}_\bullet$ is a homomorphism because

$$\mathrm{ev}_{gg'}(\chi) = \chi(g)\chi(g') = \mathrm{ev}_g(\chi)\,\mathrm{ev}_{g'}(\chi).$$

It remains to show that $\mathrm{ev}_\bullet$ is an isomorphism. We claim that $\mathrm{ev}_\bullet$ is injective, which will be enough because $|G| = |\widehat{\widehat{G}}|$ by Proposition 1.17.

For this, we appeal to the following lemma.

Lemma 1.19. Fix a finite abelian group G with identity e. If $g \neq e$, then there exists $\chi \in \widehat{G}$ such that $\chi(g) \neq 1$.

Proof. Using the Fundamental theorem of finitely generated abelian groups, we may write

$$G \cong \prod_{i=1}^{n} \mathbb{Z}/n_i\mathbb{Z}$$

for positive integers $n_i \geq 2$. Moving our problem from G to the right-hand side, we are given some $(g_i)_{i=1}^n$ such that $[g_i] \neq [0]$ for at least one i, and we want a character χ such that $\chi\left((g_i)_{i=1}^n\right) \neq 1$. Without loss of generality, suppose that $g_1 \neq 0$ and define χ by

$$\chi\left((k_i)_{i=1}^n\right) \coloneqq e(k_1/n_1).$$

Certainly $\chi\left((g_i)_{i=1}^n\right) = e(g_1/n_1) \neq 1$, so it remains to show that χ is a character. This technically follows from Lemma 1.15, but we can see it directly by computing

$$\chi\left((k_i)_{i=1}^n + (k_i')_{i=1}^n\right) = e(k_1/n_1)e(k_1'/n_1) = \chi\left((k_i)_{i=1}^n\right)\chi\left((k_i')_{i=1}^n\right).$$

This completes the proof. ∎

The proof now follows quickly from Lemma 1.19. By contraposition, we see that any $g \in G$ such that $\chi(g) = 1$ for all $\chi \in \widehat{G}$ and must have $g = e$. But this is exactly the statement that $\mathrm{ev}_\bullet \colon G \to \widehat{\widehat{G}}$ is injective. ■

1.1.5 Finite Fourier Analysis

We now proceed to essentially do Fourier analysis for finite abelian groups. Here is the idea.

Idea 1.20. We can write general functions $G \to \mathbb{C}$ as linear combinations of characters.

Remark 1.21. When G is not abelian, one must work with function $G \to \mathbb{C}$ which are "locally constant" on conjugacy classes of G.

Here is our Fourier transform.

Notation 1.22. Let G be a finite abelian group. Given a function $f \colon G \to \mathbb{C}$, we define $\widehat{f} \colon \widehat{G} \to \mathbb{C}$ by

$$\widehat{f}(\chi) \coloneqq \sum_{g \in G} f(g)\overline{\chi(g)}.$$

Recall $\overline{\chi(g)} = \chi\left(g^{-1}\right)$ by Remark 1.13.

To manifest Idea 1.20 properly, we need the following orthogonality relations.

Proposition 1.23. Let G be a finite abelian group.

- For any fixed $\chi \in \widehat{G}$, we have
$$\sum_{g \in G} \chi(g) = \begin{cases} 0 & \text{if } \chi \neq 1, \\ \#G & \text{if } \chi = 1. \end{cases}$$
- For any $g \in G$, we have
$$\sum_{\chi \in \widehat{G}} \chi(g) = \begin{cases} 0 & \text{if } g \neq e, \\ \#G & \text{if } g = e. \end{cases}$$

Proof. We show these directly.

(a) If $\chi = 1$, then the sum is $\sum_{g \in G} 1 = \#G$.
Otherwise, $\chi \neq 1$, so there exists $g_0 \in G$ such that $\chi(g_0) \neq 1$. It follows

$$\chi(g_0) \sum_{g \in G} \chi(g) = \sum_{g \in G} \chi(g_0 g) \overset{*}{=} \sum_{g \in G} \chi(g),$$

so we must have $\sum_{g \in G} \chi(g) = 0$. Note that we have re-indexed the sum at $\overset{*}{=}$.

(b) If $g = e$, then the sum is $\sum_{\chi \in \widehat{G}} \chi(g) = \#(\widehat{G})$, which is $\#G$ by Proposition 1.17.
Otherwise, $g \neq e$, so by Lemma 1.19, there exists χ_0 such that $\chi_0(g) \neq 1$. Employing the same trick, it follows

$$\chi_0 \sum_{\chi \in \widehat{G}} \chi(g) = \sum_{\chi \in \widehat{G}} (\chi_0 \chi)(g) \overset{*}{=} \sum_{\chi \in \widehat{G}} \chi(g),$$

so we must have $\sum_{\chi \in \widehat{G}} \chi(g) = 0$. Again, we re-indexed at $\overset{*}{=}$. ■

Now here is our result.

Theorem 1.24 (Fourier inversion)**.** Let G be a finite abelian group. For any $f\colon G \to \mathbb{C}$, we have

$$f(g) = \frac{1}{\#G} \sum_{\chi \in \widehat{G}} \widehat{f}(\chi)\chi(g)$$

for any $g \in G$.

Proof. This is direct computation with Proposition 1.23. Indeed, for any $g_0 \in G$, we see

$$\sum_{\chi\in\widehat{G}} \widehat{f}(\chi)\chi(g_0) = \sum_{\chi\in\widehat{G}} \sum_{g\in G} f(g)\chi\left(g^{-1}\right)\chi(g_0) = \sum_{g\in G}\left(f(g)\sum_{\chi\in\widehat{G}}\chi\left(g^{-1}g_0\right)\right).$$

Now using Proposition 1.23, given $g \in G$, we see that the inner sum will vanish whenever $g \neq g_0$ and returns $\#G$ when $g = g_0$. In total, it follows

$$\frac{1}{\#G}\sum_{\chi\in\widehat{G}} \widehat{f}(\chi)\chi(g_0) = f(g_0),$$

which is exactly what we wanted. ∎

Here is our chief application.

Corollary 1.25. Let G be a finite abelian group. Fixing some $g_0 \in G$, we have

$$1_{g_0}(g) = \frac{1}{\#G}\sum_{\chi\in\widehat{G}} \overline{\chi(g_0)}\chi(g)$$

for any $g \in G$.

Proof. Note

$$\widehat{1}_{g_0}(\chi) = \sum_{g\in G} 1_{g_0}(g)\overline{\chi(g)} = \overline{\chi(g_0)}$$

because all terms except $g = g_0$ vanish. The result now follows from Theorem 1.24. ∎

1.1.6 Dirichlet Characters

We want to extend our characters on $(\mathbb{Z}/q\mathbb{Z})^\times$ to work on all $\mathbb{Z}$, but this requires some trickery because, for example, 0 is not in general represented in $(\mathbb{Z}/q\mathbb{Z})^\times$. Here is our definition.

Definition 1.26 (Dirichlet character)**.** Let q be a nonzero integer. A *Dirichlet character* $\pmod q$ is a function $\chi\colon \mathbb{Z} \to \mathbb{C}$ such that there exists a character $\widetilde{\chi}\colon (\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}^\times$ for which

$$\chi(a) = \begin{cases} 0 & \text{if } \gcd(a,q) > 1, \\ \widetilde{\chi}([a]) & \text{if } \gcd(a,q) = 1. \end{cases}$$

We might write this situation as $\chi \pmod q$. The Dirichlet character corresponding to 1 is denoted χ_0 and still called the *principal character*.

Remark 1.27. Note χ is periodic with period q.

We can finally define our generalization of ζ.

Definition 1.28 (Dirichlet L-function)**.** Fix a Dirichlet character $\chi \pmod q$. Then we define the *Dirichlet L-function* as

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

By Proposition 1.3, we have absolute convergence for $\operatorname{Re} s > 1$, and $L(s,\chi)$ defines a holomorphic function there.

Remark 1.29. Continuing in the context of the definition, we note Proposition 1.5 gives

$$L(s,\chi) = \prod_{p \text{ prime}} \left(\sum_{k=0}^{\infty} \frac{\chi(p)^k}{p^{ks}} \right) = \prod_{p \text{ prime}} \frac{1}{1-\chi(p)p^{-s}}$$

for $\operatorname{Re} s > 1$.

In fact, the summation for $L(s,\chi)$ defines a holomorphic function for $\operatorname{Re} s > 0$, but seeing this requires a little care.

1.2 January 20

A syllabus was posted. There are some extra references posted.

1.2.1 Continuing $L(s,\chi)$

We are going to need the following technical result. Roughly speaking, it allows us to estimate infinite sums with a discrete part and a continuous part by summing the discrete part and integrating the continuous part. Oftentimes, a sum is difficult because of the way it mixes discrete and continuous portions, so it is useful to be able to separate them.

Theorem 1.30 (Abel summation)**.** Let $\{a_n\}_{n\in\mathbb{N}}$ be a sequence of complex numbers, and define the partial sums be given by

$$A(t) \coloneqq \sum_{1 \le n \le t} a_n.$$

For any real numbers $x, y \in \mathbb{R}$ with $x < y$ and continuously differentiable function $f\colon (0,x] \to \mathbb{C}$, we have

$$\sum_{0<n\le x} a_n f(n) = A(x)f(x) - \int_0^x A(t)f'(t)\,dt.$$

Proof. The idea is to write $a_n = A(n) - A(n-1)$, so we write

$$\begin{aligned}
\sum_{n\le x} a_n f(n) &= \sum_{n\le x} A(n)f(n) - \sum_{n\le x} A(n-1)f(n) \\
&= \sum_{0<n\le x} A(n)f(n) - \sum_{-1<n\le x-1} A(n)f(n+1) \\
&= A(\lfloor x\rfloor)f(\lfloor x\rfloor) - A(-1)f(0) - \sum_{0<n\le x-1} A(n)\big(f(n+1)-f(n)\big).
\end{aligned}$$

Note $A(-1) = 0$. We now introduce an integral by noting $A(n)(f(n+1)-f(n)) = \int_n^{n+1} A(t)f'(t)\,dt$, which

upon summing over n yields

$$\sum_{0<n\le x} a_n f(n) = A(\lfloor x\rfloor)f(\lfloor x\rfloor) - \int_0^{\lfloor x\rfloor} A(t)f'(t)\,dt.$$

To finish, we see

$$A(\lfloor x\rfloor)f(\lfloor x\rfloor) = A(x)f(x) + A(\lfloor x\rfloor)\big(f(\lfloor x\rfloor) - f(x)\big) = A(x)f(x) - \int_{\lfloor x\rfloor}^{x} A(t)\,dt,$$

which when combined with the previous equality finishes. ∎

Remark 1.31. One can use the theory of Riemann–Stieltjes integration to turn Theorem 1.30 into just an application of integration by parts, but we will not need this.

As an example application, we may give $L(s,\chi)$ an analytic continuation to $\{s : \operatorname{Re} s > 0\}$ when χ is not the principal character.

Lemma 1.32 (Differentiation under the integral sign)**.** Let $U \subseteq \mathbb{C}$ be open, and let $f\colon \mathbb{R}\times U \to \mathbb{C}$ and $g\colon \mathbb{R}\to\mathbb{R}$ be functions satisfying the following properties.

- The function g is integrable and has $\int_{\mathbb{R}} g(t)\,dt < \infty$.
- For fixed t, the function $s\mapsto f(s,t)$ is holomorphic on U and has $|\partial f(s,t)/\partial s| \le g(t)$ for all s.
- For fixed s, the function $t\mapsto f(s,t)$ is measurable.

Then the function $F(s) \coloneqq \int_{\mathbb{R}} f(s,t)\,dt$ is holomorphic on U.

Proof. We use Morera's theorem. As such, there are two steps.

1. We show F is continuous. Fix some $z\in U$, and we show the result at z. Because U is open, we may find $r>0$ such that $z\in B(z,r)$. Now, for any $z'\in B(z,r)$, we let $\gamma\colon[0,1]\to B(z,r)$ denote a curve from z to z'. Thus, the Fundamental theorem of calculus grants

$$\begin{aligned}
|F(z') - F(z)| &= \left|\int_{\mathbb{R}} \big(f(z',t) - f(z,t)\big)\,dt\right| \\
&= \left|\int_{\mathbb{R}} \left(\int_\gamma \frac{\partial f(s,t)}{\partial s}\,ds\right) dt\right| \\
&\le \int_{\mathbb{R}} \left(\int_\gamma \left|\frac{\partial f(s,t)}{\partial s}\right| ds\right) dt \\
&\le \int_{\mathbb{R}} \ell(\gamma) g(t)\,dt \\
&= \ell(\gamma)\int_{\mathbb{R}} g(t)\,dt.
\end{aligned}$$

   Thus, letting γ denote the straight line from z' to z so that $\ell(\gamma) = |z'-z|$, sending $z'\to z$ forces $F(z')\to F(z)$, which is what we wanted.

2. Now, let $\gamma\colon[0,1]\to U$ denote a closed curve, and we show $\int_\gamma F(s)\,ds = 0$. Indeed, the absolute convergence of the previous step allows us to use Fubini's theorem to write

$$\int_\gamma F(s)\,ds = \int_\gamma \left(\int_{\mathbb{R}} \frac{\partial f(s,t)}{\partial s}\,dt\right) ds = \int_{\mathbb{R}} \left(\int_\gamma \frac{\partial f(s,t)}{\partial s}\,ds\right) dt.$$

   Now, $\int_\gamma (\partial f(s,t)/\partial s)\,ds$ is just 0 because $s\mapsto f(s,t)$ is holomorphic, so $s\mapsto \partial f(s,t)/\partial s$ is also holomorphic. This completes the proof. ∎

Remark 1.33. We do two small computations, using integration by parts. For $\sigma < -1$, we note that

$$\int_1^\infty (\log t)t^\sigma \, dt = \frac{(\log t)t^{\sigma+1}}{\sigma+1}\bigg|_1^\infty - \frac{1}{\sigma+1}\int_1^\infty t^\sigma \, dt < \infty.$$

Similarly, for $\sigma > -1$, we note that

$$\int_0^1 (\log t)t^\sigma \, dt = \frac{(\log t)t^{\sigma+1}}{\sigma+1}\bigg|_0^1 - \frac{1}{\sigma+1}\int_0^1 t^\sigma \, dt < \infty.$$

Proposition 1.34. Let $\chi \pmod q$ be a non-principal Dirichlet character. Then the function $L(s,\chi)$ admits an analytic continuation to $\{s : \operatorname{Re} s > 0\}$.

Proof. For given s with $\operatorname{Re} s > 1$, set $a_n \coloneqq \chi(n)$ and $f(x) \coloneqq 1/x^s$. Then the partial sums $A(t) \coloneqq \sum_{1\le n\le t} a_n$ have

$$\sum_{n=1}^{kq}\chi(n) = \sum_{a=0}^{k-1}\sum_{r=1}^{q}\chi(aq+r) = k\sum_{r=1}^{q}\chi(r) = k\sum_{\substack{1\le r\le q\\ \gcd(r,q)=1}}\chi(r) = k\cdot 0$$

for any $k \ge 0$, where in the last equality we have used Proposition 1.23. Thus, for any $t \ge 0$, find $k \in \mathbb{Z}$ such that $kq \le t < k(q+1)$, and we see

$$|A(t)| = \left|\sum_{1\le n\le t}\chi(n)\right| = \left|\sum_{1\le n\le kq}\chi(n) + \sum_{kq<n\le t}\chi(n)\right| \le \sum_{kq<n\le t}|\chi(n)| \le t - kq \le q.$$

Now, finally using Theorem 1.30, we see

$$L(s,\chi) = \sum_{n=1}^\infty \frac{\chi(n)}{n^s} = \left(\lim_{x\to\infty} A(x)x^{-s}\right) - \lim_{x\to\infty}\int_0^x \left(A(t)\cdot -st^{-s-1}\right)\, dt.$$

Because $\operatorname{Re} s > 1$, we see $|A(x)x^{-s}| \le qx^{-\operatorname{Re} s}$ goes to 0 as $x \to \infty$. Thus, we are left with

$$L(s,\chi) = s\int_0^\infty \frac{A(t)}{t^{s+1}}\, dt = s\underbrace{\int_1^\infty \frac{A(t)}{t^{s+1}}\, dt}_{I(s)}.$$

We claim that the right-hand side provides our analytic continuation to $\{s : \operatorname{Re} s > 0\}$. Indeed, it suffices to show that $I(s)$ is analytic on $\{s : \operatorname{Re} s > 0\}$. This is technical.

Roughly speaking, we want to write

$$\left|\int_1^\infty \frac{A(t)}{t^{s+1}}\, dt\right| \le q\int_1^\infty \frac{1}{t^{\operatorname{Re} s+1}}\, dt = q\cdot\frac{t^{\operatorname{Re} s}}{-\operatorname{Re} s}\bigg|_1^\infty = \frac{q}{\operatorname{Re} s}$$

for any $\operatorname{Re} s > 0$, meaning that the integral converges, so we ought to have a holomorphic function. To make this computation rigorous, we will show that $I(s)$ is holomorphic on $\{s : \operatorname{Re} s > \sigma\}$ for any $\sigma > 0$, which will be enough by taking the union over all σ. Indeed, for some fixed σ, we set $g(t) \coloneqq q(\log t)/t^{\sigma+1}$ for $t > 2$ and 0 elsewhere so that

$$\left|\frac{\partial}{\partial s}\frac{A(t)}{t^{s+1}}\right| = \left|\frac{A(t)\log t}{t^{s+1}}\right| \le g(t)$$

for $\operatorname{Re} s > \sigma$, and

$$\frac{1}{q}\int_{\mathbb{R}} g(t)\, dt = \frac{1}{q}\int_1^\infty \frac{\log t}{t^{\sigma+1}}\, dt < \infty$$

by Remark 1.33. Thus, Lemma 1.32 implies that $I(s)$ is holomorphic on $\{s : \operatorname{Re} s > \sigma\}$, finishing the proof. ∎

Remark 1.35. Using the notions and notations of the above proof, we see that

$$|L(s,\chi)| = \left| s\int_1^\infty \frac{A(t)}{t^{s+1}}\,dt \right| \le \frac{q|s|}{\operatorname{Re} s}$$

for $\operatorname{Re} s > 0$. This upper-bound is occasionally helpful.

One might wonder what happens to the principal character χ_0. It turns out its behavior is tied to ζ.

Lemma 1.36. Let $\chi_0 \pmod q$ be the principal Dirichlet character. Then for $\operatorname{Re} s > 1$, we have

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

Proof. By Remark 1.29, we see

$$L(s,\chi) = \prod_{p\text{ prime}} \frac{1}{1-\chi(p)p^{-s}} = \prod_{p\nmid q}\frac{1}{1-p^{-s}},$$

so

$$L(s,\chi)\prod_{p\mid q}\frac{1}{1-p^{-s}} = \prod_{p\text{ prime}}\frac{1}{1-p^{-s}} = \zeta(s)$$

by Corollary 1.6, which finishes. ∎

Thus, we are interested in continuing ζ. With a little more effort than Proposition 1.34, we may provide $\zeta(s)$ a meromorphic continuation to $\{s : \operatorname{Re} s > 0\}$. The main difficulty here is that we have a pole to deal with.

Proposition 1.37. The function $\zeta(s)$ has a meromorphic continuation to $\{s : \operatorname{Re} s > 0\}$. It is holomorphic everywhere except at $s = 1$, where it has a simple pole of residue 1.

Proof. For given s with $\operatorname{Re} s > 1$, set $a_n := 1$ and $f(x) := 1/x^s$. Then the partial sums $A(t) := \sum_{1\le n\le t} a_n$ have $A(t) = \lfloor t\rfloor$, so Theorem 1.30 grants

$$\zeta(s) = \sum_{n=1}^\infty \frac{1}{n^s} = \left(\lim_{x\to\infty}\lfloor x\rfloor\cdot x^{-s}\right) - \lim_{x\to\infty}\int_0^x \left(\lfloor t\rfloor\cdot -st^{-s-1}\right)\,dt.$$

Because $\operatorname{Re} s > 1$, we see $|\lfloor x\rfloor\cdot x^{-s}| \le x^{1-\operatorname{Re} s}$ goes to 0 as $x\to\infty$. Thus, we are left with

$$\zeta(s) = s\int_0^\infty \frac{\lfloor t\rfloor}{t^{s+1}}\,dt = s\int_1^\infty \frac{\lfloor t\rfloor}{t^{s+1}}\,dt.$$

To extract out a main term, we write $\lfloor t\rfloor = t + \{t\}$, giving

$$\zeta(s) = s\int_1^\infty t^{-s}\,dt + s\int_1^\infty \frac{\{t\}}{t^{s+1}}\,dt = \frac{s}{s-1} + s\underbrace{\int_1^\infty \frac{\{t\}}{t^{s+1}}\,dt}_{I(s)}.$$

We claim that the above expression defines our meromorphic continuation. Notably, the function $s/(s-1) = 1 + 1/(s-1)$ is holomorphic everywhere except at $s = 1$, where it has a simple pole of residue 1.

Thus, it remains to show that $s \cdot I(s)$ is a holomorphic function for $\operatorname{Re} s > 0$, where it suffices to show that $I(s)$ is a holomorphic function for $\operatorname{Re} s > 0$. This is mildly technical. At a high level, we would like to just note that

$$\left| \int_1^\infty \frac{\{t\}}{t^{s+1}}\, dt \right| \le \int_1^\infty \frac{1}{t^{\operatorname{Re} s+1}}\, dt = \left. \frac{t^{-\operatorname{Re} s}}{-\operatorname{Re} s} \right|_1^\infty = \frac{1}{\operatorname{Re} s},$$

so the integral converges and ought to define a holomorphic function. To make this computation rigorous, we will show that $I(s)$ is holomorphic on $\{s : \operatorname{Re} s > \sigma\}$ for any $\sigma > 0$, which will be enough by taking the union over all σ. Indeed, for some fixed σ, we set $g(t) \coloneqq (\log t)/t^{\sigma+1}$ for $t > 1$ and 0 elsewhere so that

$$\left| \frac{\partial}{\partial s} \frac{\{t\}}{t^{s+1}} \right| = \left| \frac{\{t\} \log t}{t^{s+1}} \right| \le g(t)$$

for $\operatorname{Re} s > \sigma$, and

$$\int_{\mathbb{R}} g(t)\, dt = \int_1^\infty \frac{\log t}{t^{\sigma+1}}\, dt < \infty$$

by Remark 1.33. Thus, Lemma 1.32 implies that $I(s)$ is holomorphic on $\{s : \operatorname{Re} s > \sigma\}$, finishing the proof. ■

Remark 1.38. Using the notions and notation of the above proof, we see that

$$|\zeta(s)| \le \frac{|s|}{|s-1|} + \left| s \int_1^\infty \frac{\{t\}}{t^{s+1}}\, dt \right| \le \frac{|s|}{|s-1|} + \frac{|s|}{\operatorname{Re} s}.$$

For example, if $\operatorname{Re} s > 1$, then we get $|\zeta(s)| \le 1 + \frac{|s|}{\operatorname{Re} s} < |s| + 1$.

Remark 1.39. Doing repeated integration by parts, one can extend the continuations above further to the left, but we will not do this. Instead, we will use a functional equation to continue to all $\mathbb{C}$ in one fell swoop.

Corollary 1.40. Let $\chi_0 \pmod q$ denote the principal Dirichlet character. Then $L(s, \chi)$ has a meromorphic continuation to $\{s : \operatorname{Re} s > 0\}$. It is holomorphic everywhere except for a simple pole at $s = 1$.

Proof. Note that the function $\prod_{p \mid q} (1 - p^{-s})$ is entire and has its only zero at $s = 0$. Combining Lemma 1.36 and Proposition 1.37 completes the proof. ■

1.2.2 Reducing to $L(1, \chi)$

We now attack Theorem 1.1 directly. As in Theorem 1.7, we will want to understand $\log L(s, \chi)$.

Lemma 1.41. Let $\chi \pmod q$ be a Dirichlet character. For any s with $\operatorname{Re} s > 1$, we have

$$\log L(s, \chi) = \sum_{p \text{ prime}} \frac{\chi(p)}{p^s} + E(s, \chi),$$

where $|E(s, \chi)| \le 1$.

Proof. Fix s with $\operatorname{Re} s > 1$. Applying $\log$ to the Euler product of Remark 1.29, we see

$$\log L(s, \chi) = \sum_{p \text{ prime}} -\log\left(1 - \chi(p) p^{-s}\right) = \sum_{p \text{ prime}} \left(\sum_{k=1}^\infty \frac{\chi(p)^k}{k p^{ks}} \right).$$

The $k = 1$ term of the right-hand sum is the main term present in the statement, so we need to bound the terms with $k > 1$. Thus, for $\operatorname{Re} s > 1$, we compute

$$\left| \sum_{p \text{ prime}} \left(\sum_{k=2}^{\infty} \frac{\chi(p)^k}{kp^{ks}} \right) \right| \le \sum_{n=2}^{\infty} \left(\sum_{k=2}^{\infty} \frac{1}{n^k} \right) = \sum_{n=2}^{\infty} \frac{1/n^2}{1 - 1/n} = \sum_{n=2}^{\infty} \frac{1}{n(n-1)} = \sum_{n=2}^{\infty} \left(\frac{1}{n-1} - \frac{1}{n} \right) = 1,$$

where we have telescoped in the last equality. This completes the proof. ∎

As an aside, we note that Lemma 1.41 provides us with a relatively large zero-free region for $L(s, \chi)$.

Corollary 1.42. Let $\chi \pmod q$ be a Dirichlet character. For any s with $\operatorname{Re} s > 1$, we have $L(s, \chi) \neq 0$.

Proof. By Lemma 1.41, we see

$$|\log L(s, \chi)| \le \sum_{p \text{ prime}} \left| \frac{\chi(p)}{p^s} \right| + 1 \le \sum_{n=1}^{\infty} \frac{1}{n^{\operatorname{Re} s}} + 1,$$

which converges because $\operatorname{Re} s > 1$. Thus, $\log L(s, \chi)$ takes on a finite value for all s with $\operatorname{Re} s > 0$, which implies $L(s, \chi) \neq 0$. ∎

We now see that we can use Lemma 1.41 and Corollary 1.25 to extract a particular congruence class.

Lemma 1.43. Let q be an integer. For brevity, set $G \coloneqq (\mathbb{Z}/q\mathbb{Z})^\times$, and fix some $a \in G$. For any s with $\operatorname{Re} s > 1$, we have

$$\sum_{\substack{p \text{ prime} \\ p \equiv a \pmod q}} \frac{1}{p^s} = \frac{1}{\varphi(q)} \sum_{\chi \in \widehat{G}} \overline{\chi(a)} \log L(s, \chi) + E(s),$$

where $|E(s)| \le 1$.

Proof. Corollary 1.25 tells us

$$1_{[a]}(p) = \frac{1}{\varphi(q)} \sum_{\chi \in \widehat{G}} \overline{\chi}(a) \chi(p),$$

so

$$\sum_{\substack{p \text{ prime} \\ p \equiv a \pmod q}} \frac{1}{p^s} = \frac{1}{\varphi(q)} \sum_{\chi \in \widehat{G}} \left(\overline{\chi}(a) \sum_{p \text{ prime}} \frac{\chi(p)}{p^s} \right).$$

However, using the notation of Lemma 1.41, we see

$$\frac{1}{\varphi(q)} \sum_{\chi \in \widehat{G}} \left(\overline{\chi}(a) \sum_{p \text{ prime}} \frac{\chi(p)}{p^s} \right) = \frac{1}{\varphi(q)} \sum_{\chi \in \widehat{G}} \overline{\chi}(a) \log L(s, \chi) + \frac{1}{\varphi(q)} \sum_{\chi \in \widehat{G}} \overline{\chi}(a) E(s, \chi).$$

Because $\#\widehat{G} = \#G = \varphi(q)$ by Proposition 1.17, we conclude that the right-hand error term has magnitude bounded by 1, which completes the proof. ∎

We can now reduce Theorem 1.1 to analyzing $L(1, \chi)$.

Proposition 1.44. Let q be an integer. Suppose that $L(1,\chi) \neq 0$ for each non-principal Dirichlet character $\chi \pmod q$. Then, for all $a \in (\mathbb{Z}/q\mathbb{Z})^\times$, we have

$$\sum_{\substack{p \text{ prime} \\ p \equiv a \pmod q}} \frac{1}{p} = +\infty.$$

In particular, there are infinitely many primes $p \equiv a \pmod q$.

Proof. Note that $L(1,\chi)$ is at least a complex number for non-principal characters $\chi \pmod q$ by Proposition 1.34.

Let χ_0 denote the principal character. By Corollary 1.40, we see $L(s,\chi_0) \to +\infty$ as $s \to 1^+$: indeed, we know $L(s,\chi_0)$ must go to something in $\mathbb{R}_{\geq 0} \cup \{\infty\}$ because $L(s,\chi_0) \geq 1$ when $s > 1$ is real. But $L(s,\chi_0)$ cannot go to a finite value because then $L(s,\chi_0)$ would only have a removable singularity at $s = 1$.

Thus, we also have $\log L(s,\chi_0) \to +\infty$ as $s \to 1^+$. However, $\log L(s,\chi) \to \log L(1,\chi)$ as $s \to 1^+$ for non-principal characters χ, and by hypothesis, this is a finite limit. It follows that

$$\lim_{s \to 1^+} \frac{1}{\varphi(q)} \sum_{\chi \in \widehat{G}} \overline{\chi(a)} \log L(s,\chi) = +\infty,$$

so the result follows from Lemma 1.43. ∎

So we want to understand $L(1,\chi)$ when χ is a non-principal character. By paying closer attention to the above proof, we can control most of our characters χ.

Lemma 1.45. Let q be an integer, and set $G \coloneqq (\mathbb{Z}/q\mathbb{Z})^\times$ for brevity. For each Dirichlet character $\chi \pmod q$, let $v(\chi)$ denote the order of vanishing of $L(s,\chi)$ at $s = 1$. Then

$$\sum_{\chi \in \widehat{G}} v(\chi) \leq 0.$$

In other words, at most one non-principal character χ has $L(1,\chi) = 0$, in which case $L(s,\chi)$ has a simple zero at $s = 1$.

Proof. The idea here is that Lemma 1.43 has a certainly nonnegative sum on the left-hand side, so not too many of the $L(s,\chi)$s on the right-hand side may be 0, for otherwise the right-hand side would go to $-\infty$.

We make a few quick remarks on $v(\chi)$. Note Corollary 1.40 implies $v(\chi_0) = -1$, where χ_0 is the principal character. Additionally, $v(\chi) \geq 0$ for all non-principal characters χ by Proposition 1.34, and $v(\chi)$ is finite because $L(s,\chi)$ is not constantly zero by Corollary 1.42.

Thus, for each character χ, we may write $L(s,\chi) = (s-1)^{v(\chi)} L_0(s,\chi)$ for some function $L_0(s,\chi)$ holomorphic on $\{s : \operatorname{Re} s > 0\}$ with $L_0(1,\chi) \neq 0$. Setting up our application of Lemma 1.43, we see

$$\sum_{\chi \in \widehat{G}} \log L(s,\chi) = \left(\sum_{\chi \in \widehat{G}} v(\chi) \right) \log(s-1) + \left(\sum_{\chi \in \widehat{G}} \log L_0(s,\chi) \right)$$

for $\operatorname{Re} s > 1$. However, we now plug into Lemma 1.43 with $a \coloneqq 1$ so that $\overline{\chi(a)} = 1$ for all χ, giving

$$\sum_{\substack{p \text{ prime} \\ p \equiv 1 \pmod q}} \frac{1}{p^s} = \frac{1}{\varphi(q)} \left(\sum_{\chi \in \widehat{G}} v(\chi) \right) \log(s-1) + \frac{1}{\varphi(q)} \left(\sum_{\chi \in \widehat{G}} \log L_0(s,\chi) \right) + E(s)$$

for $\operatorname{Re} s > 0$. As $s \to 1^+$, the left-hand side remains nonnegative. On the right-hand side, the middle and right terms both remain finite, so the left term must also remain finite. However, $\log(s-1) \to -\infty$ as $s \to 1^+$, so we must have $\sum_\chi v(\chi) \leq 0$ to ensure this term is nonnegative.

We now show the last sentence. Indeed, we have

$$\sum_{\chi \in \widehat{G} \setminus \{\chi_0\}} v(\chi) \le -v(\chi_0) = 1,$$

so at most one $\chi \in \widehat{G} \setminus \{\chi_0\}$ may have $v(\chi) > 0$, in which case χ has $v(\chi) = 1$. ■

For example, the above lemma lets us control "complex" characters.

Lemma 1.46. Let q be an integer. If $\chi \pmod q$ is a non-principal Dirichlet character with $\chi \neq \overline{\chi}$, then $L(1,\chi) \neq 0$.

Proof. If $L(1,\chi) = 0$, then we see

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

But this grants two distinct characters χ and $\overline{\chi}$ with $L(1,\chi) = L(1,\overline{\chi}) = 0$, violating Lemma 1.45. ■

Thus, it remains to deal with the "real" non-principal characters χ with $\chi = \overline{\chi}$. This is genuinely difficult, so we will wait until next class for them.

1.3 January 23

Today we finish the proof of Theorem 1.1.

1.3.1 The Dirichlet Convolution

As motivation, we might be interested in the product of two Dirichlet series. Formally, we might write

$$\left(\sum_{k=1}^{\infty} \frac{a_k}{k^s}\right)\left(\sum_{\ell=1}^{\infty} \frac{b_\ell}{\ell^s}\right) = \sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty} \frac{a_k b_\ell}{(k\ell)^s} = \sum_{n=1}^{\infty} \left(\sum_{k\ell=n} a_k b_\ell\right) \frac{1}{n^s}.$$

Of course, we will want to formalize this intuitive argument to give the corresponding series the correct analytic properties, but we have at least arrived at the correct definition.

Definition 1.47 (Dirichlet convolution)**.** Fix functions $f, g \colon \mathbb{N} \to \mathbb{C}$. Then the *Dirichlet convolution* $(f * g) \colon \mathbb{N} \to \mathbb{C}$ is defined by

$$(f*g)(n) \coloneqq \sum_{k\ell=n} f(k)g(\ell) = \sum_{d \mid n} f(d)g(n/d).$$

And we may now take products of Dirichlet series.

Proposition 1.48. Fix functions $f, g \colon \mathbb{N} \to \mathbb{C}$ such that $|f(n)|, |g(n)| = O(n^\sigma)$ for some $\sigma \in \mathbb{R}$. Then define the series

$$F(s) \coloneqq \sum_{n=1}^{\infty} \frac{f(n)}{n^s}, \qquad G(s) \coloneqq \sum_{n=1}^{\infty} \frac{g(n)}{n^s}, \qquad D(s) \coloneqq \sum_{n=1}^{\infty} \frac{(f*g)(n)}{n^s}.$$

Then D converges absolutely for $\operatorname{Re} s > \sigma + 1$, where it defines a holomorphic function given by $D(s) = F(s)G(s)$.

Proof. Fix s with $\operatorname{Re} s > \sigma + 1$. We show that $D(s)$ converges absolutely and yields $D(s) = F(s)G(s)$, from which it follows that $D(s)$ is holomorphic over the region by using Proposition 1.3 on F and G. Let $F_n(s)$, $G_n(s)$, and $D_n(s)$ denote the nth partial sums. Then we see

$$F_N(s)G_N(s) = \left(\sum_{k=1}^{N} \frac{f(k)}{k^s}\right)\left(\sum_{\ell=1}^{N} \frac{g(\ell)}{\ell^s}\right) = \underbrace{\sum_{n=1}^{N}\left(\sum_{k\ell=n} f(k)g(\ell)\right)\frac{1}{n^s}}_{D_N(s)} + \underbrace{\sum_{\substack{1\le k,\ell\le N\\ k\ell>N}} \frac{f(k)g(\ell)}{(k\ell)^s}}_{R_N(s):=}.$$

Thus, the key claim is that $R_N(s) \to 0$ as $N \to \infty$. The main point is that $k\ell > N$ requires $k > \sqrt{N}$ or $\ell > \sqrt{N}$, so

$$|R_N(s)| \le \sum_{\substack{1\le k,\ell\le N\\ k\ell>N}} \frac{|f(k)|\cdot|g(\ell)|}{(k\ell)^{\operatorname{Re} s}} \le \left(\sum_{k>\sqrt{N}} \frac{|f(k)|}{k^{\operatorname{Re} s}}\right)\left(\sum_{\ell\ge 1} \frac{|g(\ell)|}{\ell^{\operatorname{Re} s}}\right) + \left(\sum_{k\ge 1} \frac{|f(k)|}{k^{\operatorname{Re} s}}\right)\left(\sum_{\ell>\sqrt{N}} \frac{|g(\ell)|}{\ell^{\operatorname{Re} s}}\right).$$

The absolute convergence of F and G at s now causes the right-hand side to be

$$\left(\sum_{k=1}^{\infty} \frac{|f(k)|}{k^{\operatorname{Re} s}}\right)\cdot 0 + 0 \cdot \left(\sum_{\ell=1}^{\infty} \frac{|g(\ell)|}{\ell^{\operatorname{Re} s}}\right) = 0$$

as $N \to \infty$, so we conclude $R_N(s) \to 0$ as $N \to \infty$. Thus, we conclude

$$F(s)G(s) = \lim_{N\to\infty}(F_N(s)G_N(s)) = \lim_{N\to\infty} D_N(s) + \lim_{N\to\infty} R_N(s) = D(s).$$

Lastly, we need to show that $D(s)$ actually converges absolutely. Well, we note that we can replace f with $|f|$ and g with $|g|$ and s with $\operatorname{Re} s$ everywhere in the above bounding to show that

$$\sum_{n=1}^{\infty}\left|\frac{(f*g)(n)}{n^s}\right| \le \sum_{n=1}^{\infty} \frac{(|f|*|g|)(n)}{n^{\operatorname{Re} s}} = \left(\sum_{k=1}^{\infty} \frac{|f(k)|}{k^{\operatorname{Re} s}}\right)\left(\sum_{\ell=1}^{\infty} \frac{|g(\ell)|}{\ell^{\operatorname{Re} s}}\right),$$

and the right-hand side converges because $F(s)$ and $G(s)$ converge absolutely. Thus, $D(s)$ converges absolutely. ∎

Example 1.49. Let $d(n)$ denote the number of divisors of n. Then we see

$$\zeta(s)^2 = \sum_{n=1}^{\infty} \frac{(1*1)(n)}{n^s} = \sum_{n=1}^{\infty} \frac{d(n)}{n^s}.$$

Here, $1\colon \mathbb{N} \to \mathbb{C}$ is the function which constantly returns 1.

We might be interested in an Euler product factorization for a product of two Dirichlet series (as in Proposition 1.5), but this notably requires the relevant functions to be multiplicative. Thus, we now show that the Dirichlet convolution sends multiplicative functions to multiplicative functions.

Lemma 1.50. Let $f, g\colon \mathbb{N} \to \mathbb{C}$ be multiplicative functions. Then $(f*g)\colon \mathbb{N} \to \mathbb{C}$ is still multiplicative.

Proof. Let n and m be coprime positive integers. We must show $(f*g)(nm) = (f*g)(n)\cdot(f*g)(m)$. The key point is that there is a bijection between divisors $d \mid nm$ and pairs of divisors $d_n \mid n$ and $d_m \mid m$ by sending (d_n, d_m) to d. We quickly show formally that this is a bijection.

- Well-defined: certainly $d_n \mid n$ and $d_m \mid m$ implies $d_n d_m \mid nm$.

- Injective: suppose $d_n d_m = d_n' d_m'$ for $d_n, d_n' \mid n$ and $d_m, d_m' \mid m$. We show $d_n = d_n'$, and $d_m = d_m'$ follows by symmetry. Well, for each $p \mid n$, we see $p \nmid m$ because $\gcd(n,m) = 1$, so $p \nmid d_m, d_m'$ as well, meaning
$$\nu_p(d_n) = \nu_p(d_n d_m) = \nu_p(d_n' d_m') = \nu_p(d_n')$$
for all $p \mid n$. However, $p \mid d_n, d_n'$ implies $p \mid n$, so we see that the prime factorizations of d_n and d_n' are the same, so $d_n = d_n'$.
- Surjective: for each $d \mid nm$, define $d_n \coloneqq \gcd(d,n)$ and $d_m \coloneqq \gcd(d,m)$. Certainly $d_n \mid n$ and $d_m \mid m$, so it remains to show $d = d_n d_m$. Well, for each $p \mid n$, we see $\nu_p(d_n) = \nu_p(d)$ because $d \mid n$; and similarly, each $p \mid m$ has $\nu_p(d_m) = \nu_p(m)$. Because each prime $p \mid nm$ divides exactly one of n or m, we see that
$$\nu_p(d_n d_m) = \nu_p(d_n) + \nu_p(d_m) = \nu_p(d)$$
by doing casework on $p \mid n$ or $p \mid m$.

We have written down all of this so that we may compute
$$\begin{aligned}(f*g)(nm) &= \sum_{d \mid nm} f(d) g(nm/d) \\ &= \sum_{d_n \mid n} \sum_{d_m \mid m} f(d_n d_m) g\left(\frac{n}{d_n} \cdot \frac{m}{d_m}\right) \\ &\overset{*}{=} \left(\sum_{d_n \mid n} f(d_n) g(n/d_n)\right)\left(\sum_{d_m \mid m} f(d_m) g(m/d_m)\right) \\ &= (f*g)(n) \cdot (f*g)(m).\end{aligned}$$
Here, we have used the multiplicativity at $\overset{*}{=}$, noting that $d_n \mid n$ and $d_m \mid m$ implies $\gcd(d_n, d_m) = 1$ because $\gcd(n,m) = 1$. ∎

1.3.2 The Mellin Transform

In this subsection, we pick up a few facts about the Mellin transform. Roughly speaking, we are doing Fourier analysis on the group $\mathbb{R}^+$ whose operation is multiplication. As such, the Haar measure is dx/x: for any Borel set $S \subseteq \mathbb{R}^+$ and $a \in \mathbb{R}^+$, we see
$$\int_{aS} \frac{dx}{x} = \int_S \frac{d(ax)}{ax} = \int_S \frac{a}{a} \cdot \frac{dx}{x} = \int_S \frac{dx}{x},$$
so dx/x is in fact a translation-invariant measure on $\mathbb{R}^+$. Anyway, here is our definition of the Mellin transform.

Definition 1.51 (decaying)**.** A function $\varphi\colon (0,\infty) \to \mathbb{C}$ is *decaying at a rate of* (α, β) for real numbers $\alpha < \beta$ if and only if the functions $x^\alpha \varphi(x)$ and $x^\beta \varphi(x)$ are bounded.

Example 1.52. If $\varphi\colon (0,\infty) \to \mathbb{C}$ has compact support, then φ decays at a rate of (α, β) for all $\alpha < \beta$. Indeed, for any γ, the function $x^\gamma \varphi(x)$ is a continuous function supported on a compact set and is thus bounded.

Definition 1.53 (Mellin transform)**.** Let $\varphi\colon (0,\infty) \to \mathbb{C}$ be a continuous function decaying at a rate of (α, β). Then the *Mellin transform* is the function $\mathcal{M}\varphi$ given by
$$(\mathcal{M}\varphi)(s) \coloneqq \int_0^\infty \varphi(x) x^s \frac{dx}{x}$$
for $\alpha < \operatorname{Re} s < \beta$.

Remark 1.54. We quickly check that the integral $\mathcal{M}\varphi$ (absolutely) converges for $\alpha < \operatorname{Re} s < \beta$. For each $\gamma \in \{\alpha, \beta\}$, find a constant $C_\gamma \in \mathbb{R}$ such that $|x^\gamma \varphi(x)| \le C_\gamma$ for all $x \in (0,\infty)$. For our absolute convergence, we set $\sigma \coloneqq \operatorname{Re} s \in (\alpha,\beta)$ and compute

$$\int_0^\infty |\varphi(x)x^s| \, \frac{dx}{x} \le \int_0^1 C_\alpha x^{-\alpha+\sigma-1} \, dx + \int_1^\infty C_\beta x^{-\beta+\sigma-1} \, dx,$$

so both of the right-hand integrals converge because $-\alpha+\sigma-1 > -1$ and $-\beta+\sigma-1 < -1$. Notably, this shows that $(\mathcal{M}\varphi)$ is uniformly bounded by

$$\int_0^1 C_\alpha x^{-\alpha+\alpha_0-1} \, dx + \int_1^\infty C_\beta x^{-\beta+\beta_0-1} \, dx$$

whenever $\sigma \in [\alpha_0, \beta_0]$.

Remark 1.55. Fixing some $\sigma \in (\alpha,\beta)$, let $\psi(u) \coloneqq e^{-\sigma u}\varphi(e^{-u})$. Provided that ψ is Schwarz, changing variables by $x = e^{-u}$ gives

$$(\mathcal{M}\varphi)(\sigma + 2\pi i t) = \int_0^\infty \varphi(x) x^{\sigma+2\pi i t} \, \frac{dx}{x} = \int_{\mathbb{R}} \varphi\left(e^{-u}\right) e^{-\sigma u - 2\pi i t u} \, du = (\mathcal{F}\psi)(t).$$

Here is a basic result on the Mellin transform.

Lemma 1.56. Fix a differentiable function $\varphi\colon (0,\infty) \to \mathbb{C}$ such that φ decays at a rate of (α,β). Defining $\psi(x) \coloneqq x\varphi'(x)$, for any $\alpha < \operatorname{Re} s < \beta$, the integral defining $(\mathcal{M}\psi)(s)$ converges, and

$$(\mathcal{M}\psi)(s) = -s(\mathcal{M}\varphi)(s).$$

Proof. This is by integration by parts. Indeed, we compute

$$\begin{aligned}(\mathcal{M}\psi)(s) &= \int_0^\infty x\varphi'(x)x^s \, \frac{dx}{x} \\ &= x^s\varphi(x)\Big|_0^\infty - s\int_0^\infty \varphi(x)x^s \, \frac{dx}{x} \\ &= -s(\mathcal{M}\varphi)(s),\end{aligned}$$

which is what we wanted. Note, $x^s\varphi(x) \to 0$ as $x \to 0^+$ and $x \to \infty$ because φ decays at a rate of (α,β) and $\operatorname{Re} s \in (\alpha,\beta)$. ∎

We will need two key properties of the Mellin transform.

Proposition 1.57. Let $\varphi\colon (0,\infty) \to \mathbb{C}$ be a continuous function decaying at a rate of (α,β).

(a) The function $\mathcal{M}\varphi$ is holomorphic on $\{s : \alpha < \operatorname{Re} s < \beta\}$.

(b) Suppose that φ is infinitely differentiable, and the nth derivatives decays at a rate of $(\alpha-n, \beta-n)$. Then for any integer $A \ge 0$ and $[\alpha_0,\beta_0] \subseteq (\alpha,\beta)$, the set

$$\left\{|s|^A(\mathcal{M}\varphi)(s) : \alpha_0 \le \operatorname{Re} s \le \beta_0\right\}$$

is bounded.

Proof. These are essentially bounding results.

(a) We use Lemma 1.32. Here, $f(s,t) \coloneqq \varphi(x)x^{s-1}$. We will show that $\mathcal{M}\varphi$ is holomorphic on the vertical strip $U \coloneqq \{s : -\alpha_0 < \operatorname{Re} s < \beta_0\}$ for any $\alpha < \alpha_0 < \beta_0 < \beta$, and the result will follow by taking the union over all α_0 and β_0.

By hypothesis on φ, we can find a constant C such that $|x^\alpha \varphi(x)| \le C$ and $\left|x^\beta \varphi(x)\right| \le C$ for each x. As such, we set

$$g(t) \coloneqq \begin{cases} 0 & x \le 0 \\ C(\log x)x^{-\alpha+\alpha_0-1} & \text{if } x \in (0,1], \\ C(\log x)x^{-\beta+\beta_0-1} & \text{if } x > 1. \end{cases}$$

Note $\int_{\mathbb{R}} g(t)\,dt < \infty$ by Remark 1.33 because $-\alpha+\alpha_0-1 > -1$ and $-\beta+\beta_0-1 < -1$. Thus, we see that $x \in (0,1]$ gives

$$\left|\frac{\partial}{\partial s}\varphi(x)x^{s-1}\right| = \left|\varphi(x)(\log x)x^{s-1}\right| \le C(\log x)x^{-\alpha+\operatorname{Re} s-1} \le C(\log x)x^{-\alpha+\alpha_0+1},$$

and similar for $x \in (1,\infty)$ comparing with β_0. The result now follows from Lemma 1.32.

(b) This follows from Lemma 1.56. Define $\varphi_0 \coloneqq \varphi$ and $\varphi_{n+1}(x) \coloneqq x\varphi_n'(x)$ for each n. By induction, φ_n decays at a rate of (α,β) for each n, and for each n, we see

$$|s^n(\mathcal{M}\varphi)(s)| = |(\mathcal{M}\varphi_n)(n)|$$

by Lemma 1.56. However, for each n, we see that $(\mathcal{M}\varphi_n)$ is uniformly bounded on $[\alpha_0,\beta_0]$ by Remark 1.54, which is what we wanted. ∎

Remark 1.58. The condition that $\varphi^{(n)}$ decay at a rate of $(\alpha-n,\beta-n)$ is essentially requiring that φ behave like a polynomial somewhat. These sorts of conditions more or less vanish for sufficiently good functions; for example, if φ is infinitely differentiable and has compact support, then all the derivatives have compact support, so $\varphi^{(n)}$ always decays at a rate of (α,β) for all $\alpha < \beta$ by Example 1.52.

Theorem 1.59. Let $\varphi\colon (0,\infty) \to \mathbb{C}$ be a function such that $\psi(u) \coloneqq e^{-\sigma u}\varphi\left(e^{-u}\right)$ is Schwarz. For any $\sigma \in \mathbb{R}$ and $x \in (0,\infty)$, we have

$$\varphi(x) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} (\mathcal{M}\varphi)(s)x^{-s}\,ds.$$

Proof. We translate everything to the Fourier setting with Remark 1.55, where Theorem A.10 finishes. Following this outline, we compute

$$\begin{aligned} \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} (\mathcal{M}\varphi)(s)x^{-s}\,ds &= \int_{\mathbb{R}} (\mathcal{M}\varphi)(\sigma+2\pi i t)x^{-\sigma-2\pi i t}\,dt \\ &= x^{-\sigma}\int_{\mathbb{R}} (\mathcal{F}\psi)(t)e^{2\pi i(-\log x)t}\,dt \\ &= x^{-\sigma}\cdot\psi(-\log x) \\ &= \varphi(x), \end{aligned}$$

which is what we wanted. ∎

1.3.3 Finishing Dirichlet's Theorem

We finish the proof of Theorem 1.1. By Proposition 1.44 and Lemma 1.46, we have left to show $L(1,\chi) \neq 0$ for real characters χ. We provide a slick proof of this result.

Lemma 1.60. Let $\chi \pmod q$ be a "real" non-principal Dirichlet character, meaning $\chi = \overline{\chi}$. We show

Proof. We combine two techniques called "positivity" and "smoothing." The main point is that $L(1,\chi) = 0$ implies that the zero of $L(s,\chi)$ at $s=1$ is able to cancel the pole of $\zeta(s)$ as $s=1$, implying that the function $\zeta(s)L(s,\chi)$ is holomorphic on $\{s : \operatorname{Re} s > 0\}$ by combining Propositions 1.34 and 1.37.

Anyway, we divide the proof in three steps.

1. Let's begin with our positivity result. Because we are interested in $\zeta(s)L(s,\chi)$, we will want to study the coefficients of this Dirichlet series, which are given by $(1*\chi)$ by Proposition 1.48. Note $(1*\chi)$ is multiplicative by Lemma 1.50.

   To set up our bounding, we claim that $(1*\chi)(n) \ge 0$ for all $n \in \mathbb{N}$, and $(1*\chi)\left(n^2\right) \ge 1$. Because $(1*\chi)$ is multiplicative, we may write

   $$(1*\chi)(n) = (1*\chi)\left(\prod_{p\mid n} p^{\nu_p(n)}\right) = \prod_{p\mid n} (1*\chi)\left(p^{\nu_p(n)}\right).$$

   Thus, it suffices to show $(1*\chi)\left(p^k\right) \ge 0$ for each prime-power p^k, and $(1*\chi)\left(p^k\right) \ge 1$ when k is even. Well, we can compute this directly as

   $$(1*\chi)\left(p^k\right) = \sum_{d\mid p^k} \chi(d) = \sum_{\nu=0}^{k} \chi\left(p^\nu\right) = \sum_{\nu=0}^{k} \chi(p)^\nu.$$

   Now, $\chi(p) = \overline{\chi(p)}$ by hypothesis on χ, so because $|\chi(p)| = 1$ by Remark 1.13, we conclude $\chi(p) \in \{\pm 1\}$. Thus, on one hand, if $\chi(p) = 1$, then $(1*\chi)\left(p^\nu\right) = \nu + 1 \ge 1$ always. On the other hand, if $\chi(p) = -1$, then $(1*\chi)\left(p^\nu\right)$ is 1 when ν is even and 0 if ν is odd. The claim follows.

   To finish, our positivity claim is that

   $$\sum_{x<n\le 2x} (1*\chi)(n) \ge \sum_{x<n^2\le 2x} (1*\chi)\left(n^2\right) \ge \sum_{\sqrt{x}<n\le\sqrt{2x}} 1 = \left\lfloor \sqrt{2x} \right\rfloor - \left\lfloor \sqrt{x} \right\rfloor \ge (\sqrt{2}-1)\sqrt{x} - 2.$$

   Thus, for x large enough, we see

   $$\sum_{x<n\le 2x} (1*\chi)(n) \ge \frac{1}{3}\sqrt{x}.$$

2. We now apply smoothing Let $\varphi\colon (0,\infty) \to (0,\infty)$ be an infinitely differentiable function with support contained in $[0.9, 2.1]$ such that $\varphi(x) = 1$ for $x \in [1,2]$. Then one sees

   $$\sum_{n=1}^{\infty} \varphi(n/x)(1*\chi)(m) \ge \sum_{x<n\le 2x} (1*\chi)(n) \ge \frac{1}{3}\sqrt{x}.$$

   Note that this sum is finite because only finitely many n have $n/x \le 2.1$.

   We now use the Mellin transform $\mathcal{M}\varphi$. Indeed, note that φ is decaying at a rate of (α,β) for any $\alpha < \beta$ by Remark 1.58. Further, for any $\sigma > 0$, the function $\psi(u) := e^{-\sigma u}\varphi\left(e^{-u}\right)$ has compact support and is infinitely differentiable, so $x^k\psi^{(\ell)}(x)$ is continuous of compact support for all k and ℓ and hence bounded. Thus, ψ is Schwarz, so we can use Theorem 1.59 to compute

   $$\sum_{n=1}^{\infty} \psi(n/x)(1*\chi)(n) = \frac{1}{2\pi i}\sum_{n=1}^{\infty} \int_{2-i\infty}^{2+i\infty} \left((\mathcal{M}\varphi)(s)x^s \cdot \frac{(1*\chi)(n)}{n^s}\right) ds.$$

   Thus, we see that we would like to exchange the integral and the sum so that we can sum over $(1*\chi)$ to finally make $\zeta(s)L(s,\chi)$ appear. It suffices to show that this iterated "integral" absolutely converges, so for any $\sigma > 0$, we may compute

   $$I_\sigma(x) := \int_{\sigma-i\infty}^{\sigma+i\infty} \sum_{n=1}^{\infty} \left|(\mathcal{M}\varphi)(s)x^s \cdot \frac{(1*\chi)(n)}{n^s}\right| ds = \int_{\sigma-i\infty}^{\sigma+i\infty} \left|(\mathcal{M}\varphi)(s)x^s \cdot \zeta(s)L(s,\chi)\right| ds$$

by Proposition 1.48. To bound this, we see $|x^s| \le x^{\operatorname{Re} s} = x^\sigma$ and

$$|\zeta(s)L(s,\chi)| \le q \cdot \frac{|s|}{\sigma} \cdot |s| \left(\frac{1}{|1-\sigma|} + \frac{1}{\sigma} \right) = C_0(q,\sigma)|s|^2$$

by Remarks 1.35 and 1.38, where $C_0(q,\sigma)$ is some constant. Thus,

$$I_\sigma(x) \le C_0(q,\sigma)x^c \int_{\sigma-i\infty}^{\sigma+i\infty} \left(|\mathcal{M}\varphi)(s)| \left(\sigma^2 + (\operatorname{Im} s)^2\right)\right) ds.$$

However, by Proposition 1.57 (and Remark 1.58), there is C such that $|(\mathcal{M}\varphi)(s)| \le C|s|^{-4} \le C(\operatorname{Im} s)^{-4}$ on the vertical strip of interest, so we bound

$$\begin{aligned}\frac{I(x)}{C_0(q,\sigma)x^c} \le C\left(\int_{\sigma-i\infty}^{\sigma-i} \frac{\left(\sigma^2 + (\operatorname{Im} s)^2\right)}{(\operatorname{Im} s)^4}\, ds\right) + C\left(\int_{\sigma+i}^{\sigma+i\infty} \frac{\left(\sigma^2 + (\operatorname{Im} s)^2\right)}{(\operatorname{Im} s)^4}\, ds\right) \\ + \left(\int_{\sigma-i}^{2+i} \left(|\mathcal{M}\varphi)(s)| \left(\sigma^2 + (\operatorname{Im} s)^2\right)\right) ds\right).\end{aligned}$$

The integrals on the top row are finite by direct computation (they are improper integrals avoiding 0 of decaying on the order of x^{-2} or faster), and the bottom integral is finite because it is a finite integral of a continuous function. We conclude that $I(x)$ converges, so we have absolute convergence.

In fact, the entire right-hand side of the above bound is merely some function of σ, so we have actually shown that

$$\int_{\sigma-i\infty}^{\sigma+i\infty} |(\mathcal{M}\varphi)(s)x^s \cdot \zeta(s)L(s,\chi)|\, ds \le C(q,\sigma)x^c \tag{1.1}$$

for some constant $C(q,\sigma)$.

3. Anyway, we now know we can write

$$\frac{1}{3}\sqrt{x} \le \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \underbrace{(\mathcal{M}\varphi)(s)x^s\zeta(s)L(s,\chi)}_{D(s)}\, ds$$

by exchanging the sum and the integral and using Proposition 1.48. In order to use (1.1), we would like to push the vertical line left from $\operatorname{Re} s = 2$ to $\operatorname{Re} s = 1/3$ (for example).

We will be allowed to do this by Cauchy's theorem because the function $D(s) = (\mathcal{M}\varphi)(s)x^s\zeta(s)L(s,\chi)$ is holomorphic on $\{s : \operatorname{Re} s > 0\}$. Indeed, the only possible pole among these functions is the pole of order 1 at $s = 1$ for $\zeta(s)$, but $L(s,\chi)$ has a zero there by assumption and thus cancels this out!

We now apply Cauchy's theorem. For any $T > 0$, we see

$$\left| \int_{1/3-iT}^{1/3+iT} D(s)\, ds - \int_{2-iT}^{2+iT} D(s)\, ds \right| \le \int_{1/3+iT}^{2+iT} |D(s)|\, ds + \int_{1/3-iT}^{2-iT} |D(s)|\, ds.$$

We would like to show that this right-hand side vanishes as $T \to \infty$. Because the length of each of these paths is finite, it suffices to show that $|D(s)|$ vanishes as $\operatorname{Im} s \to \infty$ on these paths. Well, utilizing our bounds from before, we see

$$|D(s)| \le |(\mathcal{M}\varphi)(s)| \cdot x^2 \cdot C_0(q,\sigma)\left(4 + (\operatorname{Im} s)^2\right).$$

Because $(\mathcal{M}\varphi)(s)$ is rapidly decaying as $\operatorname{Im} s \to \infty$ (recall Proposition 1.57), we see that this indeed goes to 0 as $\operatorname{Im} s \to \infty$.

In total, we see

$$\frac{1}{3}\sqrt{x} \le \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} D(s)\, ds = \frac{1}{2\pi i} \int_{1/3-i\infty}^{1/3+i\infty} D(s)\, ds \le C(q,1/3)x^{1/3},$$

where we have used (1.1) at the end. However, for x large enough, this is impossible: $x^{1/2-1/3} \to \infty$ as $x \to \infty$. So we have hit our contradiction. ∎

Remark 1.61. The product $\zeta(s)L(s,\chi)$ is the Dedekind ζ-function associated to a real quadratic field.

1.3.4 A Little on Quadratic Forms

To say something in the direction of Dirichlet's class number formula, we discuss quadratic forms. In particular, we will discuss the reduction theory, which shows that there are finitely many classes of binary quadratic forms of given discriminant.

Definition 1.62 (binary quadratic form)**.** A *binary quadratic form* is a function $f\colon \mathbb{Z}^2 \to \mathbb{Z}$ where $f(x,y) \coloneqq ax^2+bxy+cy^2$ where $a,b,c \in \mathbb{Z}$. If $\gcd(a,b,c)=1$, then we call the quadratic form *primitive*.

It is a problem of classical interest to determine when a quadratic form achieves a particular integer.

It is another problem of classical interest to count the number of binary quadratic forms. However, some binary quadratic forms are "the same," in the sense that they are just a variable change away.

Example 1.63. The quadratic forms $x_1^2+x_2^2$ and $y_1^2+2y_1y_2+2y_2^2$ are roughly the same by the change of variables given by

$$(y_1,y_2)=(x_1-x_2,x_2).$$

To define this correctly, we define a group action on the set of quadratic forms.

Lemma 1.64. Let $\mathcal{Q}$ be the set of binary quadratic forms. Then $\mathrm{SL}_2(\mathbb{Z})$ acts on the set of binary quadratic forms by

$$(\gamma \cdot f) \coloneqq f \circ \gamma^{-1},$$

where $f \in \mathcal{Q}$ and $\gamma \in \mathrm{SL}_2(\mathbb{Z})$.

Proof. We have the following checks.

- Identity: note $(\mathrm{id}\cdot f) = f\circ \mathrm{id}^{-1} = f\circ \mathrm{id} = f$.
- Composition: note $((\gamma\gamma')\cdot f) = f\circ(\gamma\gamma')^{-1} = f\circ(\gamma')^{-1}\circ\gamma^{-1} = \gamma\cdot(\gamma'\cdot f)$. ∎

Definition 1.65 (equivalent)**.** Two binary quadratic forms $f_1, f_2\colon \mathbb{Z}^2\to\mathbb{Z}$ are *equivalent* if and only if f_1 and f_2 live in the same orbit under the $\mathrm{SL}_2(\mathbb{Z})$-action. In other words, f_1 and f_2 are equivalent if and only if there exists $\gamma\in\mathrm{SL}_2(\mathbb{Z})$ such that

$$f_1 = f_2\circ\gamma.$$

Note that this is in fact an equivalence relation because the orbits of a group action form a partition.

Remark 1.66. For a binary quadratic form $f(x,y)\coloneqq ax^2+bxy+cy^2$, note that

$$f(v) = ax^2+bxy+cy^2 = \begin{bmatrix} x & y\end{bmatrix} \underbrace{\begin{bmatrix} a & b/2 \\ b/2 & c\end{bmatrix}}_{M\coloneqq} \begin{bmatrix} x \\ y\end{bmatrix} = v^{\intercal}Mv$$

for any $v=(x,y)\in\mathbb{Z}^2$. In fact, this symmetric matrix M is unique to f: if $v^{\intercal}Mv = v^{\intercal}M'v$ for all $v=(x,y)\in\mathbb{Z}^2$, then writing $M=(a_{ij})$ and $M'=(a'_{ij})$, we see

$$a_{11}x^2+2a_{12}xy+a_{22}^2 = v^{\intercal}Mv = v^{\intercal}M'v = a'_{11}x^2+2a'_{12}xy+a'_{22}y^2.$$

Plugging in $(x,y)\in\{(1,0),(0,1),(1,1)\}$ shows $M=M'$.

Remark 1.67. Associate a binary quadratic form f the matrix M as in Remark 1.66. Thus, for any $\gamma \in \mathrm{SL}_2(\mathbb{Z})$,

$$(\gamma \cdot f)(v) = f\left(\gamma^{-1}v\right) = \left(\gamma^{-1}v\right)^{\intercal} M\gamma^{-1}v = v^{\intercal}\left(\gamma^{-\intercal}M\gamma^{-1}\right)v,$$

so we can associate $\gamma \cdot f$ to the matrix $\gamma^{-\intercal}M\gamma^{-1}$. (Notably, this is still a symmetric matrix!) This allows for relatively easy computation of $\gamma \cdot f$.

So we would like to count the number of quadratic forms, up to equivalence. However, we will soon see that there are still infinitely many of equivalence classes, so we will want some stronger invariant to distinguish between them.

Definition 1.68 (discriminant)**.** The *discriminant* of the binary quadratic form $f(x,y) \coloneqq ax^2 + bxy + cy^2$ is given by $\operatorname{disc} f \coloneqq b^2 - 4ac$. The number of equivalence classes of quadratic forms of discriminant d is notated by $h(-d)$.

Remark 1.69. By definition, note that the discriminant of the binary quadratic form f is $4\det M$, where M is the matrix associated to f as in Remark 1.66. Using Remark 1.67, we see that the discriminant of $\gamma \cdot f$ is thus

$$4\det\left(\gamma^{-\intercal}\right)\det(M)\det\left(\gamma^{-1}\right) = 4\det M$$

for any $\gamma \in \mathrm{SL}_2(\mathbb{Z})$.

Remark 1.69 shows that the discriminant is invariant to equivalence class. Thus, for example, for each $d \in \mathbb{Z}$, we set

$$f_d(x,y) \coloneqq dxy$$

so that $\operatorname{disc} f = d^2$. Now letting d vary of $\mathbb{Z}$, we see that there are infinitely many equivalence classes of quadratic forms.

But once we bound our discriminant, there will be finitely many quadratic forms. Here is our goal.

Theorem 1.70. Let $d < 0$ be an integer. Then $h(d)$ is finite.

Remark 1.71. It is also true that $h(d)$ is finite when $d \geq 0$, but we will not show it here.

1.3.5 The Upper-Half Plane

To show Theorem 1.70, we will want to relate the action of $\mathrm{SL}_2(\mathbb{Z})$ on quadratic forms with the action of $\mathrm{SL}_2(\mathbb{R})$ on $\mathbb{H} \coloneqq \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$ given by

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} z \coloneqq \frac{az+b}{cz+d}.$$

Here are some checks on this action.

Lemma 1.72. Let $\mathbb{H} \coloneqq \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$ denote the upper-half plane.

(a) The group $\mathrm{SL}_2(\mathbb{R})$ acts on $\mathbb{H}$ by

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} z \coloneqq \frac{az+b}{cz+d}.$$

(b) The orbit of $i \in \mathbb{H}$ under $\mathrm{SL}_2(\mathbb{R})$ is all of $\mathbb{H}$.

(c) The stabilizer of $i \in \mathbb{H}$ is $\mathrm{SO}_2(\mathbb{R})$, the group of rotations.

Proof. We show the parts one at a time.

(a) To begin, we show that the action is well-defined: given z with $z \in \mathbb{H}$, we need to show that $\gamma \cdot z \in \mathbb{H}$ for any $\gamma \in \mathrm{SL}_2(\mathbb{R})$. Well, giving coefficients to γ, we compute

$$\gamma \cdot z = \begin{bmatrix} a & b \\ c & d \end{bmatrix} z = \frac{az+b}{cz+d} = \frac{(az+b)(c\overline{z}+d)}{|cz+d|^2} = \frac{\big(ac|z|^2+bd\big)+(adz+bc\overline{z})}{|cz+d|^2}.$$

To check $\gamma \cdot z \in \mathbb{H}$, we must check that the imaginary part here is positive. Well, we see

$$\operatorname{Im}(\gamma \cdot z) = \frac{(ad-bc)\operatorname{Im}(z)}{|cz+d|^2} = \frac{\operatorname{Im}(z)}{|cz+d|^2},$$

where the last equality is because $\det \gamma = 1$.

We now run our checks to have a group action.

- Identity: we compute
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} z = \frac{z+0}{0+1} = z.$$
- Composition: we compute
$$\begin{aligned} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \left(\begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} z \right) &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \frac{a'z+b'}{c'z+d'} \\ &= \frac{a \cdot \frac{a'z+b'}{c'z+d'} + b}{c \cdot \frac{a'z+b'}{c'z+d'} + d} \\ &= \frac{a(a'z+b')+b(c'z+d')}{c(a'z+b')+d(c'z+d')} \\ &= \frac{(aa'+bc')z+(ab'+bd')}{(ca'+dc')z+(cb'+dd')} \\ &= \begin{bmatrix} aa'+bc' & ab'+bd' \\ ca'+dc' & cb'+dd' \end{bmatrix} z \\ &= \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} \right) z. \end{aligned}$$

(b) Giving coefficients to some $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, we use the computation in (a) to see

$$\gamma \cdot i = \begin{bmatrix} a & b \\ c & d \end{bmatrix} i = \frac{\big(ac|i|^2+bd\big)+(adi+bc\overline{i})}{|ci+d|^2} = \frac{(ac+bd)+(ad-bc)i}{c^2+d^2} = \frac{(ac+bd)+i}{c^2+d^2}.$$

Thus, for any $a+bi \in \mathbb{H}$, we see

$$\begin{bmatrix} \sqrt{b} & a/\sqrt{b} \\ 0 & 1/\sqrt{b} \end{bmatrix} i = \frac{a/b+i}{1/b} = a+bi,$$

so the orbit of i is indeed all of $\mathbb{H}$.

(c) Using the computation of (b), we see that $\gamma \cdot i = i$ if and only if the usual coefficients of γ have $ac+bd=0$ and $c^2+d^2=1$. Thus, we see that any $\theta \in [0, 2\pi)$ will give

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} i = i$$

because $(\cos\theta)(\sin\theta)+(\cos\theta)(-\sin\theta)=0$ and $(\cos\theta)^2+(\sin\theta)^2=1$. It follows that $\mathrm{SO}_2(\mathbb{R})$ is certainly contained in the stabilizer of i.

Conversely, suppose γ stabilizes i and has the usual coefficients. Note that the pair (c,d) with $c^2+d^2=1$ has a unique $\theta \in [0, 2\pi)$ such that $c=\sin\theta$ and $d=\cos\theta$. To solve for a and b, we divide our work in two cases.

- If $c \neq 0$, then we see $a = -bd/c$. Further, $ad - bc = 1$, so we see $-bd^2/c - bc = 1$, which gives
$$b = -\frac{1}{d^2/c + c} = -\frac{c}{c^2 + d^2} = -c = -\sin\theta.$$
Thus, we see $a = -bd/c = d = \cos\theta$. Plugging everything in, we see $\gamma \in \mathrm{SO}_2(\mathbb{R})$.
- If $c = 0$, then $d \neq 0$, so we see $b = -ac/d$. Thus, $ad - bc = 1$, so we see $ad + ac^2/d = 1$, which gives
$$a = \frac{1}{d + c/d} = \frac{d}{c^2 + d^2} = d = \cos\theta.$$
Thus, we see $b = -ac/d = -c = -\sin\theta$. Plugging everything in, we again see $\gamma \in \mathrm{SO}_2(\mathbb{R})$.

The above cases complete the proof. ■

Remark 1.73. Parts (b) and (c) of Lemma 1.72 roughly show
$$\frac{\mathrm{SL}_2(\mathbb{R})}{\mathrm{SO}_2(\mathbb{R})} \cong \mathbb{H}.$$

Next class we will discuss how to build a fundamental domain for the induced action of $\mathrm{SL}_2(\mathbb{Z}) \subseteq \mathrm{SL}_2(\mathbb{R})$ on $\mathbb{H}$.

1.4 January 25

Today we continue discussing quadratic forms.

1.4.1 A Fundamental Domain

Recall from Remark 1.73 that
$$\frac{\mathrm{SL}_2(\mathbb{Z})}{\mathrm{SO}_2(\mathbb{R})} \cong \mathbb{H}.$$
Now, $\mathrm{SL}_2(\mathbb{Z}) \subseteq \mathrm{SL}_2(\mathbb{R})$ has a natural action on $\mathbb{H}$; this is a "discrete subgroup," so one might say that the action is discrete. (Concretely, we can see that the orbit of any $z \in \mathbb{H}$ under the action of $\mathrm{SL}_2(\mathbb{Z})$ is a discrete set.) We will be interested in a fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{H}$. Here is an example.

Proposition 1.74. Define the subset
$$D \coloneqq \{z \in \mathbb{H} : |z| > 1, -1/2 \leq \operatorname{Re} z < 1/2\} \cup \{z \in \mathbb{H} : |z| = 1, -1/2 \leq \operatorname{Re} z \leq 0\}.$$
Then D is a fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{H}$. In other words, for each $z \in \mathbb{H}$, there exists a unique $z_0 \in \mathbb{H}$ such that there exists $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $z = \gamma \cdot z_0$.

Proof. Omitted. Roughly speaking, one has to show that $\mathrm{SL}_2(\mathbb{Z})$ is generated by the elements
$$S \coloneqq \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \qquad \text{and} \qquad T \coloneqq \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$
Then one can use T to push all elements of $\mathbb{H}$ to $\{z \in \mathbb{H} : -1 \leq \operatorname{Re} z < 1\}$ and use S to push what's left over to S. We refer to [Ser12] for details. ■

1.4.2 Gauss Reduced Forms

We now use Proposition 1.74 for fun and profit.

Theorem 1.70. Let $d < 0$ be an integer. Then $h(d)$ is finite.

Proof. Roughly speaking, a quadratic form $f(x,y) \coloneqq ax^2+bxy+cz^2$ where $a, c > 0$ without loss of generality, we can study $f(x,1)$ to have a root

$$z_f \coloneqq \frac{-b+\sqrt{b^2-4ac}}{2a} = \frac{-b+\sqrt{d}}{2a}.$$

Now, in our case of interest, we have $d < 0$, so this describes an element of $\mathbb{H}$. (There is also a negative root, but we focus on z_f.) In fact, one can check that $z_{\gamma f} = \gamma z_f$, which is how we relate quadratic forms to $\mathbb{H}$.

In fact, by Proposition 1.74, we know there is some γf such that $z_{\gamma f} \in D$. The point here is that the number of quadratic forms up to equivalence is bounded above by the number of points in D with imaginary part $\sqrt{|d|}$. For example, the condition $|z_f| \geq 1$ implies that

$$\frac{b^2-d}{4a^2} = \frac{c}{a},$$

so $a \leq c$. Further, the condition $-1/2 \leq \operatorname{Re} z \leq 1/2$ implies $|b| \leq 2a$. Thus, we are counting the number of triples (a,b,c) with $a, c > 0$ such that $b^2 - 4ac = d$ and $|b| \leq a \leq c$, which we can see immediately is finite. Indeed, $b^2 \leq d$, so there are only finitely many possible b, but then for each b, we see $4ac = b^2 - d$, so there are only finitely many possible a and c. ∎

Remark 1.75. A quadratic form satisfying the above conditions on a, b, c is called "Gauss reduced."

1.4.3 Dirichlet's Class Number Formula

We take a moment to record Dirichlet's class number formula for completeness, though we will not prove it.

Theorem 1.76 (class number formula)**.** Let d be a "fundamental discriminant," meaning that $d \equiv 1 \pmod 4$ and is squarefree or $d = 4q$ where $q \equiv 2, 3 \pmod 4$ and is squarefree. Let $\chi_d = \left(\frac{d}{\bullet}\right)$ be the Kronecker symbol.

(a) If $d < 0$,

$$h(d) = \frac{w_d |d|^{1/2}}{2\pi} \cdot L(1, \chi_d),$$

where $w_d = 2$ if $d < -4$ and $w_d = 4$ if $d = -r$ and $w_d = 6$ if $d = -3$. (Namely, w_d is the number of roots of unity in $\mathbb{Q}(\sqrt{d})$.)

(b) If $d > 0$, then

$$h(d) \log \varepsilon_d = |d|^{1/2} L(1, \chi_d),$$

where ε_d is a fundamental unit for $\mathcal{O}_{\mathbb{Q}(\sqrt{d})}$. (Namely, $\varepsilon_d = (t_0 + u_0\sqrt{d})/2$ yields the least positive solution to $t_0^2 - du_0^2 = 4$.)

The point behind the fundamental discriminant is that $\operatorname{disc} \mathcal{O}_{\mathbb{Q}(\sqrt{d})} = d$.

Remark 1.77. The interested should now be able to do the first part of the first problem set.

THEME 2

THE PRIME NUMBER THEOREM

2.1 January 25

We now shift gears and move towards the Prime number theorem. Today, we begin by discussing Riemann's original paper on the topic.

Remark 2.1. For the rest of this course, any sum or product over an unnamed p will be a sum over primes.

2.1.1 The Statement

So far we have established the following facts about ζ.

- By Corollary 1.6, for $\operatorname{Re} s > 1$, there is an Euler product factorization

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p} \frac{1}{1-p^{-s}}.$$

- By Proposition 1.37, there is a meromorphic continuation of $\zeta(s)$ to $\operatorname{Re} s > 1$, where $\zeta(s)$ is analytic everywhere except for a pole of order 1 at $s = 1$.

Roughly speaking, we will show the Prime number theorem by being able to study $\zeta'(s)/\zeta(s) = \frac{d}{ds} \log \zeta(s)$. Let's establish some notation.

Definition 2.2. For $x \in \mathbb{R}$, we define the following functions.

$$\begin{aligned}
\pi(x) &:= \sum_{p \le x} 1, \\
\Lambda(n) &:= \begin{cases} \log p & \text{if } n = p^\nu \text{ for } \nu \in \mathbb{Z}^+, \\ 0 & \text{else,} \end{cases} \\
\psi(x) &:= \sum_{n \le x} \Lambda(n).
\end{aligned}$$

Quickly, we note that the prime-powers we have included in $\Lambda(n)$ and $\psi(x)$ don't actually matter.

Lemma 2.3. For any $x \geq 2$, we have

$$\psi(x) = \sum_{p \leq x} \log p + O\left(\sqrt{x}(\log x)^2\right).$$

Proof. Note

$$\psi(x) = \sum_{n \leq x} \Lambda(n) = \sum_{k=1}^{\infty} \left(\sum_{p^k \leq x} \log p \right).$$

Now, note $k \geq \log_2 x$ implies that $p^k \geq 2^k \geq x$ for all primes p, so we only need to sum up to $\log_2 x$. As such, we upper-bound the $k > 1$ sum as

$$\left| \sum_{k=2}^{\log_2 x} \left(\sum_{p^k \leq x} \log p \right) \right| \leq |\log_2 x - 1| \cdot \left| \sum_{n \leq \sqrt{x}} \log(\sqrt{x}) \right| \leq \frac{\sqrt{x}(\log x)^2}{2 \log 2}.$$

Adding the $k = 1$ sum back in, we see that

$$\psi(x) = \sum_{p \leq x} \log p + O\left(\sqrt{x}(\log x)^2\right),$$

which is what we wanted. ∎

Now, here is our statement.

Theorem 2.4 (Prime number)**.** We have $\pi(x) \sim x/\log x$ as $x \to \infty$.

Here is why we mentioned ψ.

Proposition 2.5. The following are equivalent.

(a) $\pi(x) \sim x/\log x$ as $x \to \infty$.

(b) $\psi(x) \sim x$ as $x \to \infty$.

Proof. This is by summation by parts. We have two implications to show.

(a) Suppose (b), and we'll show (a). Let $a_n = \log n$ be 1 if n is prime and 0 otherwise, and let $A(x) := \sum_{n \leq x} a_n$ denote the partial sums. By Lemma 2.3, we see

$$A(x) = \sum_{p \leq x} \log p = \psi(x) + O\left(\sqrt{x}(\log x)^2\right).$$

In particular, $A(x)/x = \psi(x) + O\left(x^{-1/2}(\log x)^2\right)$ goes to 1 as $x \to \infty$ by hypothesis.

Further, let $f(n) := 1/\log n$ if $n > 1$ and 0 at $n = 1$. Then Theorem 1.30 tells us

$$\pi(x) = \sum_{n \leq x} a_n f(n) = A(x)f(x) - \int_0^x A(t)f'(t)\,dt = \frac{A(x)}{\log x} + \int_2^x \frac{A(t)}{(\log t)^2}\,dt,$$

so

$$\frac{\pi(x)}{x/\log x} = \frac{A(x)}{x} + \frac{\log x}{x} \int_2^x \frac{A(t)}{(\log t)^2}\,dt.$$

As $x \to \infty$, we know $A(x)/x \to 1$, so we have left to show that the right-hand term goes to 0. Because $A(x)/x \to 1$, we can find N such that $x > N$ implies $A(x)/x \leq 2$. Thus, we can upper-bound

$$\frac{\log x}{x} \int_2^x \frac{A(t)}{(\log t)^2}\,dt \leq 2(\log x) \int_2^x \frac{1}{(\log t)^2}\,dt$$

for $x > N$.

■

2.1.2 Poisson Summation

Starting with the easier parts of Riemann's paper, we will show the functional equation for $\zeta(s)$. For this, we use the Poisson summation formula.

Theorem 2.6 (Poisson summation)**.** Let $f : \mathbb{R} \to \mathbb{C}$ be a Schwarz function. Then

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} (\mathcal{F}f)(n).$$

Proof. Consider the function

$$F(x) \coloneqq \sum_{n\in\mathbb{Z}} f(x+n).$$

The point is to compute the Fourier series of $F : \mathbb{R} \to \mathbb{C}$. Thus, we divide the proof into steps.

1. Note that F is continuous. Indeed, we will essentially see that the series (absolutely) converges uniformly on compact sets: let F_N denote the Nth partial sum, where $N \geq 1$. Thus, to show that F is continuous on some closed interval $[a, b]$, it suffices to show that $F_N \to F$ uniformly on $[a, b]$ because each F_N is continuous. This will be enough because each $x \in \mathbb{R}$ is contained in some closed interval $[x-1, x+1]$, so F is continuous at each $x \in \mathbb{R}$.

   Before doing anything, note $x \in [a, b]$ implies $|x| \leq m$ where $m \coloneqq \max\{|a|, |b|\}$, so we will take $N > 2m$ throughout. Now, the Schwartz condition on f lets us find a constant $C \in \mathbb{R}$ such that $\left|x^2 f(x)\right| \leq C$, so

   $$|F(x) - F_N(x)| \leq \sum_{|n|>N} |f(x+n)| \leq 2C \sum_{|n|>N} \frac{1}{(x+n)^2}.$$

   The sum now splits into

   $$|F(x) - F_N(x)| \leq \sum_{n<-N} \frac{1}{(x+n)^2} + \sum_{n>N} \frac{1}{(x+n)^2} = \sum_{n>N} \left(\frac{1}{(n-x)^2} + \frac{1}{(n+x)^2} \right).$$

   The summand is now decreasing in n, so we may upper-bound this by the integral test, writing

   $$|F(x) - F_N(x)| \leq \int_{N-1}^{\infty} \left(\frac{1}{(t-x)^2} + \frac{1}{(t+x)^2} \right) dt = \frac{1}{2(N-1-x)} + \frac{1}{2(N-1+x)},$$

   which does vanish as $N \to \infty$.

   As an aside, note that the above bounding has also shown that the series $F(x)$ absolutely converges because we showed that $\sum_{|n|>N} |f(x+n)|$ converges for some N depending on x (though this dependency is irrelevant here).

2. Note F is 1-periodic because rearranging the sum gives

   $$F(x+1) = \sum_{n\in\mathbb{Z}} f(x+n+1) = \sum_{n\in\mathbb{Z}} f(x+n) = F(x).$$

3. The next step is compute the Fourier coefficients of F, which for some $n \in \mathbb{Z}$ is

   $$a_n(F) = \int_0^1 \left(\sum_{k\in\mathbb{Z}} f(x+k) e^{-2\pi i n x} \right) dx.$$

We would like to exchange the integral and the sum, so we check the absolute convergence as

$$\int_0^1 \left(\sum_{k\in\mathbb{Z}} \left| f(x+k)e^{-2\pi inx} \right| \right) dx = \int_0^1 \left(\sum_{k\in\mathbb{Z}} |f(x+k)| \right) dx.$$

Now, we showed that the series $x \mapsto \sum_{k\in\mathbb{Z}} |f(x+k)|$ converges uniformly on compact closed intervals $[a,b]$, so it defines a continuous function on the closed interval $[0,1]$, so this integral converges. As such, we may now apply Fubini's theorem to get

$$a_n(F) = \sum_{k\in\mathbb{Z}} \int_0^1 f(x+k)e^{-2\pi inx}\,dx = \sum_{k\in\mathbb{Z}} \int_k^{k+1} f(x)e^{-2\pi inx}\,dx = \int_{-\infty}^{\infty} f(x)e^{-2\pi inx}\,dx = (\mathcal{F}f)(n).$$

4. We would like to build the Fourier series using Theorem A.20, but for this we must show that S_F converges absolutely and uniformly. Well, by Lemma A.6, we see that $n \neq 0$ have

$$(\mathcal{F}f)(n) = \frac{1}{2\pi in} \cdot (\mathcal{F}f')(n) = \frac{1}{-4\pi^2 n^2} \cdot (\mathcal{F}f'')(n).$$

Now, $(\mathcal{F}f'')$ is bounded by Remark A.5, so find M such that $|(\mathcal{F}f'')(s)| \le M$ for all s. Checking the absolute and uniform convergence, we see $N > 0$ lets us upper-bound

$$\sum_{|n|>N} \left| a_n(F)e^{2\pi inx} \right| \le \frac{M}{4\pi^2} \sum_{|n|>N} \frac{1}{n^2} = \frac{2M}{4\pi^2} \sum_{n>N} \frac{1}{n^2} = \frac{2M}{4\pi^2} \int_N^{\infty} \frac{1}{x^2}\,dx = \frac{2M}{4\pi^2 N},$$

which does vanish as $N \to \infty$.

5. The previous step gives our absolute and uniform convergence, so Theorem A.20 tells us

$$\sum_{n\in\mathbb{Z}} f(x+n) = F(x) = \sum_{n\in\mathbb{Z}} a_n(F)e^{2\pi inx} = \sum_{n\in\mathbb{Z}} (\mathcal{F}f)(n)e^{-2\pi inx}$$

for all $x \in \mathbb{R}$. Setting $x = 0$ completes the proof. ∎

Example 2.7. Let f be a Schwarz function, and define $f_x(t) \coloneqq f(tx)$ for any $x > 0$. Then $(\mathcal{F}f_x)(s) = \frac{1}{x}(\mathcal{F}f)\left(\frac{s}{x}\right)$, so Theorem 2.6 yields

$$\sum_{n\in\mathbb{Z}} f(nx) = \sum_{n\in\mathbb{Z}} f_x(n) = \sum_{n\in\mathbb{Z}} (\mathcal{F}f_x)(n) = \frac{1}{x} \sum_{n\in\mathbb{Z}} (\mathcal{F}f)(n/x).$$

2.2 January 27

We began class finishing the proof of Theorem 2.6. I have edited directly into that proof for continuity.

2.2.1 An Abstract Functional Equation

We now use Theorem 2.6 in order to show the functional equation for ζ, which provides us with its meromorphic continuation.

To work somewhat abstractly, suppose $f\colon \mathbb{R} \to \mathbb{R}$ is a Schwartz function which has both f and $\widehat{f}$ even and satisfies $f(0) = \widehat{f}(0) = 0$. Now define

$$I(f,s) \coloneqq \int_0^{\infty} \left(\sum_{n=1}^{\infty} f(nx) \right) x^s \frac{dx}{x}.$$

One can show by hand that $I(f,s)$ is absolutely convergent and then analytic for $\operatorname{Re} s > 0$. However, if we take the Fourier transform and use Theorem 2.6, we are able to continue this to all of $\mathbb{C}$. Now, for $\operatorname{Re} s > 1$, we have absolute convergence, so Fubini's theorem lets us write

$$I(f,s) = \sum_{n=1}^{\infty} \left(\int_0^{\infty} f(nx) x^s \, \frac{dx}{x} \right) = \sum_{n=1}^{\infty} \left(\frac{1}{n^s} \int_0^{\infty} f(x) x^s \, \frac{dx}{x} \right) = \zeta(s)(\mathcal{M}f)(s).$$

Notably, we used $\operatorname{Re} s > 1$ in order to write $\zeta(s)$ as the series. Now, everything has a continuation to $\operatorname{Re} s > 0$, so uniqueness of extension lets us extend to $\operatorname{Re} s > 0$.

However, we were able to continue $I(f,s)$ to all of $\mathbb{C}$ by applying Poisson summation. Indeed, Theorem 2.6 yields

$$I(f,s) = \int_0^{\infty} \left(\sum_{n=1}^{\infty} \widehat{f}(n/x) \right) x^{s-1} \, \frac{dx}{x}.$$

Now, $\widehat{f}$ is also Schwarz, so we have absolute convergence for $\operatorname{Re} s < 1$, so we have indeed produced our analytic continuation to all of $\mathbb{C}$. Manipulating, we see

$$I(f,s) = \int_0^{\infty} \left(\sum_{n=1}^{\infty} \widehat{f}(nx) \right) x^{1-s} \, \frac{dx}{x} = I(\widehat{f}, s-1).$$

Comparing, we see

$$(\mathcal{M}\widehat{f})(1-s)\zeta(1-s) = I(\widehat{f}, s-1) = I(f,s) = \zeta(s)(\mathcal{M}f)(s)$$

on $0 < \operatorname{Re} s < 1$. Thus, we see that we will have a good functional equation for ζ by choosing a sufficiently good f. Indeed, choosing f which is nonzero everywhere except for some finite set of points will grant us a meromorphic continuation of ζ to all of $\mathbb{C}$.

2.2.2 The Functional Equation

We now go back and use a specific value of f to give a functional equation we can really write down. In particular, we will know that ζ only has simple poles at $s = 0$ and $s = 1$, each of residue 1. Indeed, set $f \colon \mathbb{R} \to \mathbb{R}$ by $f(x) \coloneqq e^{-\pi x^2}$ so that $\widehat{f}(x) = f(x)$. Motivated by the above work, we set

$$\xi(s) \coloneqq \pi^{-s/2} \Gamma(s/2) \zeta(s),$$

where

$$\Gamma(s) = \int_0^{\infty} e^{-t} t^s \, \frac{dt}{t}$$

for $\operatorname{Re} s > 0$, and we can continue Γ to the left by the functional equation $\Gamma(s) = s\Gamma(s-1)$. (This functional equation is proven by integration by parts.)

Remark 2.8. In some sense, Γ is a continuous version of a Gauss sum: it's an integral of an additive character multiplied by a multiplicative character, over a suitable Haar measure.

Remark 2.9. Notably, $\Gamma(s)$ has simple poles for $s \in \{0, -1, -2, \ldots\}$. In fact, $1/\Gamma(s)$ is entire.

Remark 2.10. In some sense, we want to write

$$\xi(s) = \pi^{-s/2} \Gamma(s/2) \prod_{p \text{ prime}} \frac{1}{1 - p^{-s}}.$$

Here, $\pi^{-s/2}\Gamma(s/2)$ is an "archimedean local factor" corresponding to the infinite place ∞ of $\mathbb{Q}$, and each of the $(1-p^{-s})^{-1}$ are "nonarchimedean local factors." Roughly speaking, the rigorization of this intuition is Tate's thesis.

Now working through the arguments of the previous subsection, we see that $\xi(s)$ is entire except for simple poles at $s \in \{0, 1\}$, and

$$\xi(1-s) = \xi(s).$$

Let's give a few consequences.

Remark 2.11. Doing logarithmic differentiation, one finds

$$\frac{d}{ds}(-\log \zeta(s)) = -\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\Gamma(n)}{n^s}.$$

This explains why ψ is a "better" prime-counting function than π.

Remark 2.12. Ignoring convergence issues, we may compute

$$\psi(x) = \sum_{n \le x} \Gamma(n) = \sum_{n=1}^{\infty} 1_{[0,1]}(n/x)\Gamma(n) = \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \left(-\frac{\zeta'(s)}{\zeta(s)}\right) x^s \frac{ds}{s}.$$

Now, if we imagine that we could push this integral all the way to the left of $\mathbb{C}$, we will eventually vanish and only pick up on the poles of ζ'/ζ. As such, we expect to achieve a formula of the form

$$\psi(x) = x - \sum_{\rho} \frac{x^\rho}{\rho},$$

where the sum is over the roots ρ of ζ. Thus, we see that having more control over the zeroes of ζ will be able to get good bounds on $\psi(x) - x$. In particular, the Riemann hypothesis is equivalent to $\psi(x) = x + O(\sqrt{x})$. As another application, the discontinuity of ψ will imply that ζ must have infinitely many roots.

Remark 2.13. The previous remark, after some summation by parts, tells us that $\pi(x) - \operatorname{Li}(x)$ has a better error term than $\pi(x) - x/\log x$, where

$$\operatorname{Li}(x) = \int_2^x \frac{dt}{\log t}.$$

Next class we will show the functional equation.

2.3 January 30

Today we actually prove the functional equation.

2.3.1 The Functional Equation

The functional equation will be derived from the functional equation for the $\Theta\colon \mathbb{C} \to \mathbb{C}$ function, defined by

$$\Theta(x) \coloneqq \sum_{n \in \mathbb{Z}} e^{-\pi n^2 x}.$$

In particular, Theorem 2.6 implies

$$\Theta(x) = \sum_{n \in \mathbb{Z}} e^{-\pi n^2 x} = \sum_{n \in \mathbb{Z}} e^{-\pi \left(n\sqrt{x}\right)^2} = \sum_{n \in \mathbb{Z}} \frac{1}{\sqrt{x}} e^{-\pi \left(n/\sqrt{x}\right)^2} = \frac{1}{\sqrt{x}} \Theta\left(\frac{1}{x}\right)$$

for $x > 0$. The uniqueness of analytic continuation now implies $\Theta(z) = \frac{1}{\sqrt{z}}\Theta(z)$ for all $z \in \mathbb{C}$.

Remark 2.14. For $z \in \mathbb{H}$, we note that $\Theta(z+2) = \Theta(z)$ and $\Theta(-1/z) = \sqrt{z/i} \cdot \Theta(z)$, which shows that Θ is a modular form of weight $1/2$ and level

$$\left\langle \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\rangle.$$

We now use Θ for fun and profit. Consider the integral

$$\int_0^\infty \left(\frac{\Theta(x) - 1}{2} \right) x^{s/2} \, \frac{dx}{x}.$$

This will converge absolutely for $\operatorname{Re} s > 1$. However, for $\operatorname{Re} s > 1$, we can unwind this expression as

$$\begin{aligned} \int_0^\infty \left(\frac{\Theta(x) - 1}{2} \right) x^{s/2} \, \frac{dx}{x} &= \int_0^\infty \left(\sum_{n=1}^\infty e^{-\pi n^2 x} x^{s/2} \right) \frac{dx}{x} \\ &= \zeta(s) \int_0^\infty e^{-\pi x} x^{s/2} \, \frac{dx}{x} \\ &= \zeta(s) \pi^{-s/2} \Gamma(s/2) \end{aligned}$$

after enough rearranging. (Note we interchanged the sum and integral validly because everything absolutely converges.) Now, we define

$$\xi(s) \coloneqq \zeta(s) \pi^{-s/2} \Gamma(s/2).$$

The key result is as follows.

Theorem 2.15. The function $\Xi(s) \coloneqq \xi(s) s(1-s)$ has an analytic continuation to all of $\mathbb{C}$ and satisfies the functional equation

$$\xi(s) = \xi(1-s).$$

Proof. The trick is to break the integral

$$\int_0^\infty \left(\frac{\Theta(x) - 1}{2} \right) x^{s/2} \, \frac{dx}{x}$$

into pieces $(0,1)$ and $(1,\infty)$.

- On $(1,\infty)$, we see
$$I(s) \coloneqq \int_1^\infty \left(\frac{\Theta(x) - 1}{2} \right) dx$$
converges absolutely to a holomorphic function.
- We deal with $(0,1)$. Here,
$$\int_0^1 \frac{x^{s/2}}{2} \, \frac{dx}{x} = \frac{1}{s},$$
and for the Θ term, we use modularity to write
$$\begin{aligned} \int_0^1 \left(\frac{\Theta(x)}{2} \right) x^{s/2} \, \frac{dx}{x} &= \int_0^1 \left(\frac{\Theta(1/x)}{2\sqrt{x}} \right) x^{s/2} \, \frac{dx}{x} \\ &= \int_1^\infty \left(\frac{\Theta(x)}{2} \right) x^{(1-s)/2} \, \frac{dx}{x} \\ &= I(1-s) - \frac{1}{1-s}. \end{aligned}$$
Summing our two pieces, we get $I(1-s) - \frac{1}{s} - \frac{1}{1-s}$.

In total, we see

$$\xi(s) = I(s) + I(1-s) - \frac{1}{s} - \frac{1}{1-s},$$

where $I(s)$ is holomorphic. This completes the proof. ∎

2.3.2 Zeroes of ζ

We would like to understand the zeroes of $\zeta(s)$, for which we use Cauchy's formula. Roughly speaking, we will study integrals

$$\frac{1}{2\pi i} \oint_\gamma \frac{\zeta'(s)}{\zeta(s)}\, ds,$$

where γ is a contour over a very tall vertical strip in $\mathbb{C}$.

We will want the following bounds.

Lemma 2.16 (Stirling)**.** We have

$$\log \Gamma(s) = \left(s - \frac{1}{2}\right)\log s + \frac{1}{2}\log 2\pi + O_\delta(1/|s|),$$

and

$$\frac{\Gamma'(s)}{\Gamma(s)} = \log s + O_\delta(1/|s|)$$

as $|s| \to \infty$ for $-\pi + \delta < \arg s < \pi - \delta$.

Proof. Omitted. ∎

Remark 2.17. For $\sigma \in [a, b]$ for some fixed $a < b$, we have $\Gamma(\sigma + it) \sim e^{-\pi|t|/2}|t|^{\sigma - 1/2}$ as $\to \infty$. Again, we omit this proof.

We also want access to the Hadamard factorization theorem.

Theorem 2.18. Fix an entire function $\varphi\colon \mathbb{C} \to \mathbb{C}$. Let $n(r)$ denote the number of zeroes z such that $|z| < r$ and α the order of an entire function φ, one has $n(R) = O_\varepsilon\left(R^{\alpha+\varepsilon}\right)$ for any $\varepsilon > 0$.

Notably, for $\operatorname{Re} s > 1/2$, one has $|s(s-1)\zeta(s)| \ll |s|^3$. Further, one can check that Γ has order 1 as an entire function, so $s(1-s)\xi(s)$ has order at most 1. Thus, Hadamard's factorization theorem enforces

$$s(1-s)\xi(s) = e^{A+Bs} \prod_{\zeta(\rho)=0} \left(\left(1 - \frac{s}{\rho}\right) e^{s/\rho}\right).$$

Notably, this product will converge absolutely. For example, absolute convergence tells us

$$\sum_{\zeta(\rho)=0} \frac{1}{|\rho|^{1+\varepsilon}} < \infty$$

for any $\varepsilon > 0$. One also has the following result on the distribution of our ρ.

Theorem 2.19. Define

$$N(T) \coloneqq \#\{\rho : 0 \le \operatorname{Re}\rho \le 1, \operatorname{Im}\rho \ge 0 \zeta(\rho) = 0\}.$$

Then

$$N(T) = \frac{T}{2\pi}\log\left(\frac{T}{2\pi}\right) - \frac{T}{2\pi} + O(\log T)$$

as $T \to \infty$.

We will first show the following lemma.

Lemma 2.20. We have

$$\sum_{\rho} \frac{1}{1+|\operatorname{Im}\rho - T|^2} \ll \log(T+3).$$

Proof. This is by smoothing. By taking logarithmic differentiation

$$\frac{\Xi'(s)}{\Xi(s)} = B + \sum_{\rho} \left(\frac{1}{s-\rho} + \frac{1}{\rho} \right).$$

However, by using the above estimates, we see

$$\frac{\Xi'(s)}{\Xi(s)} = \frac{1}{s} - \frac{1}{1-s} - \frac{1}{2}\log\pi + \frac{1}{2}\frac{\Gamma'(s/2)}{\Gamma(s/2)} + \frac{\zeta'(s)}{\zeta(s)}$$

by definition of ξ. Now, the term ζ'/ζ is well-behaved for $\operatorname{Re} s$ large: we set $s \coloneqq 2+it$, and one can see that $|\zeta'(s)/\zeta(s)|$ is bounded by an absolute constant. Thus, we understand on the right-hand side here.

Continuing, we see

$$\operatorname{Re}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right) = \frac{2-\operatorname{Im} s}{(2-\operatorname{IRe} s)^2 + (T-\operatorname{Im} s)^2} + \frac{\beta}{(\operatorname{Re} s + \operatorname{Im} s)^2} \gg \frac{1}{1+|T-\operatorname{Im} s|^2}.$$

However, $\Gamma'(s)/\Gamma(s) \ll \log(T+3)$ by Stirling, so the result follows. ∎

2.4 February 1

Today we move towards a proof of the explicit formula.

Notation 2.21. A sum/product over ρ is over the zeroes of $\zeta(s)$.

2.4.1 Zeroes of ζ, Again

Let's provide a few applications of Lemma 2.20.

Corollary 2.22. We have

$$\#\{\rho : \zeta(\rho) = 0, \operatorname{Im}\rho \in [T, T+1], \operatorname{Re}\rho \in (0,1)\} = O(\log T).$$

Proof. This follows from Lemma 2.20 by separating out our zeroes into intervals. ∎

We will be interested in contours γ_T which look like large vertical rectangles; namely, they are the boundary of the rectangle $[-\varepsilon, 1+\varepsilon] \times [-T, T]$. Notably, the top and bottom of the rectangle's contours will cancel out by the functional equation, so we only need to pay attention to the vertical parts of this contour.

Lemma 2.23. For $t > 3$, we have

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_{|\operatorname{Im}\rho - t| \le 1} \frac{1}{s-\rho} + O(\log t)$$

for $\operatorname{Re} s \in [-1, 2]$.

Proof. We consider

$$\frac{\zeta'(s)}{\zeta(s)} - \frac{\zeta'(2+it)}{\zeta(2+it)}.$$

Thus, we recall

$$\frac{\Xi'(s)}{\Xi(s)} = \frac{1}{s} - \frac{1}{s-1} - \frac{1}{2}\log\pi + \frac{1}{2}\frac{\Gamma'(s/2)}{\Gamma(s/2)} + \frac{\zeta'(s)}{\zeta(s)},$$

so we can just bound everything. Notably, we can use the infinite product to bound Ξ'/Ξ and then compare everything. For example,

$$\frac{\zeta'(s)}{\zeta(s)} - \frac{\zeta'(2+it)}{\zeta(2+it)} = \sum_\rho \left(\frac{1}{s-\rho} - \frac{1}{2+it-\rho}\right) + O(\log t),$$

where the $O(\log t)$ includes the trivial zeroes of ζ. Now, we notice

$$\left|\frac{1}{s-\rho} - \frac{1}{2+it-\rho}\right| \ll \frac{1}{|\operatorname{Im}\rho - t|^2}$$

for $|\operatorname{Im}\rho - t| \geq 1$, so we can use Lemma 2.20 to absorb most terms into $O(\log t)$. In total, we see

$$\frac{\zeta'(s)}{\zeta(s)} - \frac{\zeta'(2+it)}{\zeta(2+it)} = \sum_{|\operatorname{Im}\rho - t|\leq 1} \frac{1}{s-\rho} + O(\log t),$$

and now the $2+it$ term can also be absorbed to $O(\log t)$. ∎

Remark 2.24. One can give more accurate bounding than the above, but we will not need it.

We now return to the proof of Theorem 2.19.

Proof of Theorem 2.19. Use the argument principle on Ξ on the box $[-1,2]\times[-T,T]$. In particular, by the functional equation, it suffices to just look at the right and top edges of this box. The hope is that we can use Lemma 2.23 and the ideas in its proof to do the bounding for us. In particular, we will be working with the equation

$$\frac{\Xi'(s)}{\Xi(s)} = \frac{1}{s} - \frac{1}{1-s} - \frac{1}{2}\log\pi + \frac{1}{2}\frac{\Gamma'(s/2)}{\Gamma(s/2)} + \frac{\zeta'(s)}{\zeta(s)}.$$

Now, the main term in the argument will come from $\Gamma'(s)/\Gamma(s)$, which one can see using Stirling's asymptotics. Most of these terms are not going to matter on our contour. It turns out that the only difficulty is integrating ζ'/ζ over the line $\{a+Ti : a\in[-1,2]\}$. Well, using the above estimates, we recall

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_{|\operatorname{Im}\rho - T|\leq 1} \frac{1}{s-\rho} + O(\log T),$$

where now the integral of the $1/(s-\rho)$ term is bounded by a constant, and the number of terms is $O(\log T)$ by Corollary 2.22, so everything is absorbed into the error term. ∎

2.4.2 The Explicit Formula

Let's move towards the explicit formula. Here is our statement.

Theorem 2.25. When x is not a prime-power, we have

$$\psi(x) = x - \sum_\rho \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2}\log\left(1-x^{-2}\right).$$

Proof. We first describe a heuristic. The main idea is to use contour integration, noting that

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} y^s \frac{ds}{s} = \begin{cases} 0 & \text{if } y < 1, \\ 1 & \text{if } y > 1, \end{cases} \tag{2.1}$$

where $c > 1$ and $y \in (0,\infty) \setminus \{1\}$. The proof of this is essentially complex analysis where we "complete the contour" of this vertical line either off to $-\infty$ or off to $+\infty$ depending on $y < 1$ or $y > 1$.

Now, the point is that we can write

$$\psi(x) = \sum_{n \le x} \Lambda(x) \approx \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} -\frac{\zeta'(s)}{\zeta(s)} x^s \frac{ds}{s},$$

where we have ignored convergence issues to exchange the Dirichlet series for ζ'/ζ with this integral. The point now is that we can integrate ζ'/ζ appropriately to give the formula.

To make this more rigorous, we need to do only finite computations. Thus, we define

$$I(y,T) \coloneqq \frac{1}{2\pi i}\int_{c-iT}^{c+iT} y^s \frac{ds}{s}.$$

We now note that our extra variable c will be later set to $1 + 1/\log T$, so it is important to have this degree of freedom. Now, the proof of (2.1) grants

$$|I(y,T) - 1_{>1}(y)| \ll y^c \min\{1, 1/|T\log y|\},$$

where the implied constant is absolute. Integrating over this, we see

$$\left|\psi(x) - \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} -\frac{\zeta'(s)}{\zeta(s)} x^s\, ds\right| \le \left(\sum_{n=1}^{\infty} \Lambda(n)(x/n)^c \min\left\{1, \frac{1}{|T\log(x/n)|}\right\}\right).$$

Now, setting $c = 1 + 1/\log T$, we get an upper-bound of $O\left(x(\log x)^2/T\right)$. Roughly speaking, we note that x away from particular n are small; however, when n is close to x, we can explicitly evaluate the logarithm as about $1/(n-x)$, and there the sum is roughly harmonic and thus grants a logarithmic growth.

We now do a contour shift on the integral

$$\int_{c-i\infty}^{c+i\infty} -\frac{\zeta'(s)}{\zeta(s)} x^s\, ds.$$

Roughly speaking, we will expand our box to look like $[-U, 1+\varepsilon]\times[-T,T]$, sending $U \to \infty$ for fixed T. We will then send $T \to \infty$, always remembering to choose T avoiding zeroes of $\zeta(s)$. In particular, by Corollary 2.22, we can be at least $1/\log T$ away from any particular zeroes. We will finish the proof next lecture. ∎

APPENDIX A
FOURIER ANALYSIS

Ring around the rosie,
A pocket full of posies.
Ashes! Ashes!
We all fall down!

—"Ring Around the Rosie"

A.1 The Fourier Transform

It will pay off for us later to have established a little Fourier analysis right now. Our exposition follows [SS03, Chapter 5].

Definition A.1 (Schwarz)**.** A function $f : \mathbb{R} \to \mathbb{C}$ is *Schwarz* if and only if f is infinitely differentiable and the nth derivative $f^{(n)}$ satisfies that the function $x^A \cdot f^{(n)}(x)$ is bounded for all nonnegative integers A.

Remark A.2. Because the linear combination of bounded sets remains bounded, we see that Schwarz functions form a $\mathbb{C}$-vector space. Also, by definition, if f is Schwarz, then any derivative is also Schwarz.

Remark A.3. If $f : \mathbb{R} \to \mathbb{R}$ is Schwarz, we note that $\left|x^A f(x)\right|$ is integrable over $\mathbb{R}$ for any $A \geq 0$. Indeed, let k be an integer greater than $A + 2$, and we are granted a constant C such that $\left|x^k f(x)\right| \leq C$. Thus,

$$\int_{\mathbb{R}} \left|x^A f(x)\right| \, dx \leq \int_{[-1,1]} \left|x^A f(x)\right| + \int_{|x|\geq 1} \frac{C}{x^{A-k}} \, dx,$$

which is finite because $A - k < -2$.

Definition A.4 (Fourier transform)**.** Let $f : \mathbb{R} \to \mathbb{C}$ be a Schwarz function. Then we define the *Fourier transform* to be the function $\mathcal{F}f : \mathbb{R} \to \mathbb{C}$ given by

$$(\mathcal{F}f)(s) \coloneqq \int_{\mathbb{R}} f(x) e^{-2\pi i s x} \, dx.$$

Remark A.5. The integral converges because it absolutely converges: we note

$$\int_{\mathbb{R}} \left| f(x)e^{-2\pi isx} \right| \, dx = \int_{\mathbb{R}} |f(x)| \, dx$$

is finite by Remark A.3. In fact, this shows that $\mathcal{F}f$ is bounded.

We now run some quick checks on the Fourier transform.

Lemma A.6. Let $f : \mathbb{R} \to \mathbb{C}$ be a Schwartz function.

(a) For some $\lambda > 0$, define $f_\lambda(x) := f(\lambda x)$. Then $f_\lambda : \mathbb{R} \to \mathbb{C}$ is Schwartz, and $(\mathcal{F}f_\lambda)(s) = \frac{1}{\lambda}(\mathcal{F}f)\left(\frac{s}{\lambda}\right)$.

(b) We have $(\mathcal{F}f')(s) = 2\pi is(\mathcal{F}f)(s)$.

(c) The function $\mathcal{F}f$ is differentiable, and $(\mathcal{F}f)'(s)$ is the Fourier transform of the function $g(x) := -2\pi ixf(x)$.

(d) The function $\mathcal{F}f$ is Schwarz.

Proof. We show these one at a time.

(a) To show f_λ is Schwarz, we note $f_\lambda^{(n)}(x) = \lambda^n f^{(n)}(\lambda x)$ for all $n \geq 0$, so $x^n \cdot f_\lambda^{(n)}(x)$ is bounded because $(\lambda x)^n f^{(n)}(\lambda x)$ is. The last equality is a direct computation. We see

$$\begin{aligned}
(\mathcal{F}f_\lambda)(s) &= \int_{\mathbb{R}} f_\lambda(x)e^{-2\pi isx} \, dx \\
&= \int_{\mathbb{R}} f(\lambda x)e^{-2\pi i(s/\lambda)\lambda x} \, dx \\
&= \frac{1}{\lambda} \int_{\mathbb{R}} f(x)e^{-2\pi i(s/\lambda)x} \, dx \\
&= \frac{1}{\lambda}(\mathcal{F}f)\left(\frac{s}{\lambda}\right).
\end{aligned}$$

(b) Note f' is Schwarz by Remark A.2, so the statement at least makes sense. Now, by integration by parts, we see

$$\begin{aligned}
(\mathcal{F}f')(s) &= \int_{\mathbb{R}} f'(x)e^{-2\pi isx} \, dx \\
&= f(x)e^{-2\pi isx}\Big|_{-\infty}^{\infty} - \int_{\mathbb{R}} f(x)\left(2\pi ise^{-2\pi isx}\right) \, dx \\
&= 2\pi is(\mathcal{F}f)(s).
\end{aligned}$$

To justify the last equality, we see that $f(x)e^{-2\pi isx} \to 0$ as $x \to \pm\infty$ because f is Schwarz: note $\left|f(x)e^{-2\pi isx}\right| = |f(x)|$, and $|xf(x)|$ is bounded, so there is a constant C such that $f(x) \leq C/|x|$ for all $x \neq 0$, meaning that $|f(x)| \to 0$ as $x \to \pm\infty$.

(c) Note g is the product of infinitely differentiable functions and thus infinitely differentiable. Further, by induction, derivatives of g are the $\mathbb{C}$-linear of terms of the form $x^k f^{(\ell)}(x)$. Thus, for any integers $k, \ell \geq 0$, the function $|x|^k \left|g^{(k)}(x)\right|$ is a $\mathbb{C}$-linear combination of bounded functions because f is Schwarz, so it follows that $|x|^k \left|g^{(\ell)}(x)\right|$ is bounded, so g is Schwarz.

The rest of the proof is a direct computation. For $t, t' \in \mathbb{R}$, we see

$$\int_t^{t'} (\mathcal{F}g)(s) \, ds = \int_t^{t'} \left(\int_{\mathbb{R}} -2\pi ixf(x)e^{-2\pi isx} \, dx \right) \, ds.$$

We would like to exchange the two integrals. Well, by Fubini's theorem, we note that $\int_{\mathbb{R}} |xf(x)|\,dx < \infty$ is finite by Remark A.3, so

$$\int_t^{t'} \left(\int_{\mathbb{R}} \left| -2\pi i x f(x) e^{-2\pi i s x} \right| \, dx \right) ds \le 2\pi |t' - t| \int_{\mathbb{R}} |xf(x)|\,dx < \infty.$$

Thus, we may write

$$\begin{aligned} \int_t^{t'} (\mathcal{F}g)(s)\,ds &= \int_{\mathbb{R}} \left(\int_t^{t'} -2\pi i x f(x) e^{-2\pi i s x}\,ds \right) dx \\ &= \int_{\mathbb{R}} f(x) \left(e^{-2\pi i t' x} - e^{-2\pi i t x} \right) dx \\ &= (\mathcal{F}f)(t') - (\mathcal{F}f)(t). \end{aligned}$$

Thus, by the Fundamental theorem of calculus, we see

$$(\mathcal{F}f)'(t) = \lim_{t' \to t} \frac{(\mathcal{F}f)(t') - (\mathcal{F}f)(t)}{t' - t} = \lim_{t' \to t} \left(\frac{1}{t' - t} \int_t^{t'} (\mathcal{F}g)(s)\,ds \right) = (\mathcal{F}g)(t).$$

(d) By Remark A.5, the Fourier transform of a Schwarz function is bounded. Thus, it suffices to note that, for any nonnegative integers k and ℓ, the function $s \mapsto s^k (\mathcal{F}f)^{(\ell)}(s)$ is the Fourier transform of the function

$$x \mapsto \frac{1}{(2\pi i)^k} \left(\frac{d}{dx'} \right)^{(k)} \left((-2\pi i x')^\ell f(x') \right) \Big|_{x'=x}$$

by combining (b) and (c). This completes the proof. ∎

As an application, we can compute the Fourier transform of the Gaussian.

Exercise A.7 (Gaussian). Define $g \colon \mathbb{R} \to \mathbb{R}$ by $g(x) := e^{-\pi x^2}$. Then g is Schwarz, and $(\mathcal{F}g)(x) = g(x)$.

Proof. We build a differential equation that $\mathcal{F}g$ solves, and then we solve that differential equation. Namely, by using Lemma A.6 repeatedly, we see

$$\begin{aligned} (\mathcal{F}g)'(s) &= \int_{\mathbb{R}} -2\pi i x g(x) e^{-2\pi i s x}\,dx \\ &= i \int_{\mathbb{R}} g'(x) e^{-2\pi i s x}\,dx \\ &= i(\mathcal{F}g')(s) \\ &= -2\pi x (\mathcal{F}g)(s), \end{aligned}$$

so $(\mathcal{F}g)$ solves the differential equation $y' + 2\pi y = 0$. To solve this differential equation, we define $f(x) := (\mathcal{F}g)(x) e^{\pi x^2}$ and use the differential equation to write

$$f'(x) = (\mathcal{F}g)(x) \cdot 2\pi x e^{\pi x^2} - 2\pi x (\mathcal{F}g)(x) \cdot e^{\pi x^2} = 0.$$

Thus, f is a constant function, so there exists $a \in \mathbb{C}$ such that $(\mathcal{F}g)(x) = a e^{-\pi x^2}$ for all $x \in \mathbb{R}$.

To finish, we need to introduce an initial condition. Well, we compute $(\mathcal{F}g)(0) = 1$ in the usual way,

writing

$$\begin{aligned}(\mathcal{F}g)(0)^2 &= \left(\int_{\mathbb{R}} e^{-\pi x^2}\,dx\right)^2 \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} e^{-\pi\left(x^2+y^2\right)}\,dxdy \\ &= \int_0^{2\pi}\int_0^{\infty} e^{-\pi r^2}\,rdrd\theta \\ &= \int_0^{2\pi} \frac{1}{2\pi}\,d\theta \\ &= 1.\end{aligned}$$

However, surely $(\mathcal{F}g)(0) \geq 0$, so we conclude $(\mathcal{F}g)(0) = 1$. It follows that $a = 1$, so $(\mathcal{F}g)(x) = e^{-\pi x^2}$ for all $x \in \mathbb{R}$. ■

A.2 Fourier Inversion

The goal of this subsection is to prove the Fourier inversion theorem; we continue to roughly follow [SS03, Chapter 5]. Roughly speaking, this will follow from understanding the Gaussian. Here are the necessary tools.

Lemma A.8. Define the Gaussian $g\colon \mathbb{R} \to \mathbb{R}$ by $g(x) \coloneqq e^{-\pi x^2}$. For any $\delta > 0$, we have

$$\lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon}\int_{|t|\geq\delta} g(t/\varepsilon)\,dt = 0.$$

Proof. Changing variables, we see

$$\lim_{\varepsilon\to 0^+}\int_{|t|\geq\delta} g(t/\varepsilon) = \lim_{\varepsilon\to 0^+}\int_{|t|\geq\delta/\varepsilon} g(t)\,dt = \lim_{N\to\infty}\int_{|t|\geq N} g(t)\,dt,$$

where $N = \delta/\varepsilon$ in the last equality. However, g is Schwarz by Exercise A.7, so $\int_{\mathbb{R}} g(t)\,dt$ is finite by Remark A.3, so

$$\lim_{N\to\infty}\int_{|t|\leq N} g(t)\,dt = \int_{\mathbb{R}} g(t)\,dt.$$

Rearranging, we see

$$\lim_{N\to\infty}\int_{|t|\geq N} g(t)\,dt = 0,$$

which is what we wanted. ■

Lemma A.9. Define the Gaussian $g\colon \mathbb{R} \to \mathbb{R}$ by $g(x) \coloneqq e^{-\pi x^2}$. For any bounded and continuous function $f\colon \mathbb{R} \to \mathbb{R}$, we have

$$f(0) = \lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon}\int_{\mathbb{R}} f(t)g(t/\varepsilon)\,dt.$$

Proof. The point here is that, for any $\varepsilon > 0$, we have

$$\frac{1}{\varepsilon}\int_{\mathbb{R}} g(t/\varepsilon)\,dt = \int_{\mathbb{R}} g(t)\,dt = (\mathcal{F}g)(0) = g(0) = 1 \tag{A.1}$$

by Exercise A.7. However, the functions $t \mapsto g(t/\varepsilon)$ concentrate at $t = 0$ as $\varepsilon \to 0^+$, so we expect that adding in an $f(t)$ to our integral will force the output to be $f(0)$.

As an aside, we go ahead and check that these integrals converge for each $\varepsilon > 0$. Indeed, they absolutely converge: because f is bounded, we may find $M_f \in \mathbb{R}$ such that $|f(x)| \le M_f$ for each $x \in \mathbb{R}$, which gives

$$\int_{\mathbb{R}} |f(t)g(t/\varepsilon)|\,dt \le M_f \int_{\mathbb{R}} g(t/\varepsilon)\,dt = \varepsilon M_f,$$

where we have used (A.1).

We now proceed with the proof, which is somewhat technical. For psychological reasons, we set $h(x) \coloneqq f(x) - f(0)$ for all $x \in \mathbb{R}$. Note h is still bounded and continuous (it's a shift away from f). Further, for each $\varepsilon > 0$, we see

$$\frac{1}{\varepsilon}\int_{\mathbb{R}} h(t)g(t/\varepsilon)\,dt = \frac{1}{\varepsilon}\int_{\mathbb{R}} f(t)g(t/\varepsilon)\,dt - \frac{f(0)}{\varepsilon}\int_{\mathbb{R}} g(t/\varepsilon)\,dt = \frac{1}{\varepsilon}\int_{\mathbb{R}} f(t)g(t/\varepsilon)\,dt - f(0),$$

where we have used (A.1) in the last equality, so it suffices to show

$$\lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon}\int_{\mathbb{R}} h(t)g(t/\varepsilon)\,dt \stackrel{?}{=} 0.$$

Well, fix any $\delta > 0$. Note h is continuous at 0 and has $h(0) = 0$, so we may find $\delta_0 > 0$ such that $|h(t)| < \delta$ for $|t| < \delta_0$. For the other values of t, we note h is bounded, so we may find $M_h \ge 0$ such that $|h(t)| < M_h$ for all t. Thus, we upper-bound

$$\begin{aligned}
\left|\frac{1}{\varepsilon}\int_{\mathbb{R}} h(t)g(t/\varepsilon)\,dt\right| &\le \frac{1}{\varepsilon}\int_{|t|\le\delta_0} |h(t)g(t/\varepsilon)|\,dt + \frac{1}{\varepsilon}\int_{|t|\ge\delta_0} |h(t)g(t/\varepsilon)|\,dt \\
&\le \frac{\delta}{\varepsilon}\int_{|t|\le\delta_0} g(t/\varepsilon)\,dt + \frac{M_h}{\varepsilon}\int_{|t|\ge\delta_0} g(t/\varepsilon)\,dt \\
&\le \frac{\delta}{\varepsilon}\int_{\mathbb{R}} g(t/\varepsilon)\,dt + \frac{M_h}{\varepsilon}\int_{|t|\ge\delta_0} g(t/\varepsilon)\,dt \\
&= \delta + \frac{M_h}{\varepsilon}\int_{|t|\ge\delta_0} g(t/\varepsilon)\,dt.
\end{aligned}$$

(As usual, we have used (A.1) in the last equality.) Thus, using Lemma A.8, sending $\varepsilon \to 0^+$ shows that

$$\lim_{\varepsilon\to 0^+}\left|\frac{1}{\varepsilon}\int_{\mathbb{R}} h(t)g(t/\varepsilon)\,dt\right| \le \delta$$

for any $\delta > 0$, so sending $\delta \to 0^+$ completes the proof. ∎

And here is our main attraction.

Theorem A.10 (Fourier inversion)**.** Let $f : \mathbb{R} \to \mathbb{R}$ be a Schwarz function. For any $x \in \mathbb{R}$, we have

$$f(x) = \int_{\mathbb{R}} (\mathcal{F}f)(s)e^{2\pi ixs}\,ds.$$

Proof. Expanding out the definition of $\mathcal{F}f$, we are computing

$$\int_{\mathbb{R}}\left(\int_{\mathbb{R}} f(t)e^{-2\pi its}\,dt\right)e^{2\pi ixs}\,ds.$$

We would like to exchange the two integrals, but we do not have absolute convergence. As such, we employ a trick: fix some $\varepsilon > 0$, and define the integral

$$f_\varepsilon(x) \coloneqq \int_{\mathbb{R}}\int_{\mathbb{R}} f(t)e^{2\pi i(x-t)s}e^{-\pi\varepsilon^2 s^2}\,dtds.$$

Notably, we expect $f_\varepsilon(x) \to \int_{\mathbb{R}} (\mathcal{F}f)(s) e^{2\pi i x s}\,ds$ as $\varepsilon \to 0^+$. As such, we compute the behavior of $\varepsilon \to 0^+$ in two ways.

- We integrate over dt first. Namely, we would like to send $\varepsilon \to 0^+$, for which we use the Dominated convergence theorem. For each $\varepsilon > 0$, note that we have the bound
$$\left| \int_{\mathbb{R}} f(t) e^{2\pi i (x-t)s} e^{-\pi \varepsilon^2 s^2}\,dt \right| \le e^{-\pi\varepsilon^2 s^2} \int_{\mathbb{R}} |f(t)|\,dt.$$
Now, $s \mapsto e^{-\pi\varepsilon^2 s^2}$ is Schwarz by Exercise A.7 (combined with (a) of Lemma A.6), so we may integrate the right-hand function over all $s \in \mathbb{R}$ by Remark A.3.

  Thus, our integrand in $f_\varepsilon(x)$ is dominated by an integrable function, so the Dominated convergence theorem implies
$$\lim_{\varepsilon\to 0^+} f_\varepsilon(x) = \int_{\mathbb{R}} \left(\lim_{\varepsilon \to 0^+} e^{-\pi\varepsilon^2 s^2} \int_{\mathbb{R}} f(x) e^{2\pi i (x-t)}\,dt \right) ds = \int_{\mathbb{R}} (\mathcal{F}f)(s) e^{2\pi i x s}\,ds.$$
- We integrate over ds first. As such, we begin by justifying our application of Fubini's theorem: checking for absolute convergence, we compute
$$\int_{\mathbb{R}} \int_{\mathbb{R}} \left| f(t) e^{2\pi i (x-t)s} e^{-\pi\varepsilon^2 s^2} \right| dt ds = \left(\int_{\mathbb{R}} |f(t)|\,dt \right) \left(\int_{\mathbb{R}} e^{-\pi\varepsilon^2 s^2}\,ds \right).$$
Now, f is Schwarz by hypothesis, as is $s \mapsto e^{-\pi\varepsilon^2 s^2}$ by Exercise A.7, so both of these integrals are finite by Remark A.3.

  Thus, we may switch the order of our integration. Setting up notation, we let $g(x) \coloneqq e^{-\pi x^2}$ denote the Gaussian (so that $(\mathcal{F}g)(s) = g(s)$ for all $s \in \mathbb{R}$) and $g_\varepsilon(x) \coloneqq g(\varepsilon x)$. Then we see
$$\begin{aligned} f_\varepsilon(x) &= \int_{\mathbb{R}} \int_{\mathbb{R}} f(t) e^{2\pi i (x-t)s} e^{-\pi\varepsilon^2 s^2}\,ds dt \\ &= \int_{\mathbb{R}} f(t) \left(\int_{\mathbb{R}} e^{-\pi(\varepsilon s)^2} e^{-2\pi i (t-x)}\,ds \right) dt \\ &= \int_{\mathbb{R}} f(t) (\mathcal{F}g_\varepsilon)(t-x)\,dt \\ &\overset{*}{=} \frac{1}{\varepsilon} \int_{\mathbb{R}} f(t) g\left(\frac{t-x}{\varepsilon} \right) dt \\ &= \frac{1}{\varepsilon} \int_{\mathbb{R}} f(t+x) g\left(t/\varepsilon\right) dt, \end{aligned}$$
where we have used part (a) of Lemma A.6 at $\overset{*}{=}$. Sending $\varepsilon \to 0^+$, Lemma A.9 tells us that
$$f(x) = \lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon} \int_{\mathbb{R}} f(t+x) g(t/\varepsilon)\,dt = \lim_{\varepsilon \to 0^+} f_\varepsilon(x).$$

Combining the above two computations completes the proof. ∎

A.3 Fourier Coefficients

In order to say that we've done some Fourier analysis, we will also say a few things about Fourier series. We follow [SS03, Chapter 2].

The idea here is that the functions $e_n \colon x \mapsto e^{2\pi i n x}$ for $n \in \mathbb{Z}$ form an orthonormal set of continuous functions $\mathbb{R} \to \mathbb{C}$, where our (Hermitian) inner product is given by
$$\langle f, g \rangle \coloneqq \frac{1}{2\pi i} \int_0^1 f(x) \overline{g(x)}\,dx.$$

Indeed, for any $n, m \in \mathbb{Z}$, we see

$$\langle e_n, e_m \rangle = \int_0^1 e^{2\pi i n x} \overline{e^{2\pi i m x}}\, dx = \int_0^1 e^{2\pi i (m-n)x}\, dx = \begin{cases} 1 & \text{if } m = n, \\ 0 & \text{if } m \neq n. \end{cases} \tag{A.2}$$

Now, the functions e_n are varied enough that we might hope that all sufficiently smooth 1-periodic functions $f\colon \mathbb{R} \to \mathbb{C}$ can be written in terms of our orthonormal functions as

$$f(x) = \sum_{n \in \mathbb{Z}} a_n e^{2\pi i n x}$$

for some coefficients $a_n \in \mathbb{Z}$. Thus, we might hope we can extract out the nth coefficient by

$$\langle f, e_n \rangle = \int_0^1 f(x) e^{-2\pi i n x}\, dx.$$

This motivates the following definition.

Definition A.11 (Fourier coefficient)**.** Fix a continuous 1-periodic function $f\colon \mathbb{R} \to \mathbb{C}$. Then we define the *$n$th Fourier coefficient* as

$$a_n(f) \coloneqq \int_0^1 f(x) e^{-2\pi i n x}\, dx.$$

Remark A.12. Note that the integral defining $a_n(f)$ converges absolutely. Indeed, f is continuous on $[0,1]$ and hence bounded because $[0,1]$ is compact. Thus, we may find M such that $|f(x)| \le M$ for $x \in [0,1]$, which implies

$$|a_n(f)| \le \int_0^1 \left| f(x) e^{-2\pi i n x} \right| dx \le M \int_0^1 dx = M.$$

Of course, one can weaken the requirement that f be continuous, but we will have no need for these levels of generality.

Remark A.13. In fact, we note

$$a_n(f) = \int_t^{t+1} f(x) e^{2\pi i n x}\, dx$$

for any $t \in \mathbb{R}$. Because $x \mapsto f(x) e^{2\pi i n x}$ is 1-periodic, it suffices to show this for $t \in [0,1)$. Then the integral over $[t, t+1) = [t,1) \sqcup [1, 1+t)$ is equal to the integral over $[0,t) \sqcup [t,1) = [0,1)$, where we have used the 1-periodicity.

Here is some basic arithmetic with these coefficients.

Lemma A.14. Fix continuous 1-periodic functions $f, g\colon \mathbb{R} \to \mathbb{C}$.

(a) For any $z, w \in \mathbb{C}$ and $n \in \mathbb{Z}$, we see $a_n(zf + wg) = z a_n(f) + w a_n(g)$.

(b) For any $n \in \mathbb{Z}$, we see $a_n(\overline{f}) = \overline{a_{-n}(f)}$.

(c) Given $x_0 \in \mathbb{R}$, define $g(x) \coloneqq f(x + x_0)$. Then $a_n(g) = e^{-2\pi i n x_0} a_n(f)$.

Proof. Here we go.

(a) This follows from the fact that $\langle \cdot, \cdot \rangle$ is an inner product. Indeed,

$$a_n(zf + wg) = z \int_0^1 f(x) e^{-2\pi i n x}\, dx + w \int_0^1 g(x) e^{-2\pi i n x}\, dx = z a_n(f) + w a_n(g).$$

(b) We compute

$$a_n(\overline{f}) = \int_0^1 \overline{f(x)} e^{-2\pi inx}\,dx = \overline{\int_0^1 f(x)e^{2\pi inx}\,dx} = \overline{a_{-n}(f)}.$$

(c) We compute

$$a_n(g) = \int_0^1 f(x+x_0)e^{2\pi inx}\,dx = e^{-2\pi inx_0}\int_0^1 f(x+x_0)e^{2\pi in(x+x_0)}\,dx = e^{-2\pi inx_0}a_n(f),$$

where the last inequality used Remark A.13. ∎

Here is a slightly harder computation, still akin to Lemma A.6.

Lemma A.15. Fix a continuously differentiable 1-periodic function $f\colon \mathbb{R}\to\mathbb{C}$. For $n\neq 0$, we have

$$a_n(f') = -2\pi i n a_n(f).$$

Proof. This is by integration by parts. Indeed, we compute

$$\begin{aligned} a_n(f') &= \int_0^1 f'(x)e^{-2\pi inx}\,dx \\ &= \frac{f(x)e^{-2\pi inx}}{-2\pi in}\bigg|_0^1 - \frac{1}{-2\pi in}\int_0^1 f(x)e^{-2\pi inx}\,dx \\ &= 0 + \frac{1}{2\pi in}\cdot a_n(f), \end{aligned}$$

which is what we wanted. ∎

The following is our key result.

Lemma A.16. Fix a continuous 1-periodic function $f\colon\mathbb{R}\to\mathbb{R}$ such that $f(0)\neq 0$. Then $a_n(f)\neq 0$ for some $n\in\mathbb{Z}$.

Proof. Define the function $p\colon\mathbb{R}\to\mathbb{C}$ by $p(x)\coloneqq e^{-2\pi inx}$. Roughly speaking, the idea is that $a_n(f)=0$ for all $n\in\mathbb{Z}$ implies that any "polynomial in p" named $q\in\mathbb{C}\left[p,p^{-1}\right]$ written as

$$q\coloneqq\sum_{n\in\mathbb{Z}} q_n p^n,$$

where all but finitely many of the q_n vanish, will have

$$\int_{-1/2}^{1/2} f(x)q(x)\,dx = \sum_{n\in\mathbb{Z}}\left(q_n\int_{-1/2}^{1/2} f(x)e^{2\pi inx}\,dx\right) = \sum_{n\in\mathbb{Z}} q_n a_n(f) = 0$$

by Remark A.13. Indeed, we will be able to build a function $q\in\mathbb{C}\left[p,p^{-1}\right]$ which is "concentrated at 0" so that $f(0)\neq 0$ is incompatible with all these integrals vanishing.

We now proceed with the proof. Quickly, we replace $f(x)$ with $f(x)/f(0)$, which is still continuous, 1-periodic, and has $a_n(f/f(0)) = a_n(f)/f(0)$ for all $n\in\mathbb{Z}$, so $a_n(f/f(0))\neq 0$ implies $a_n(f)\neq 0$. Thus, we may assume $f(0)=1$, and we still want to show $a_n(f)\neq 0$ for some n.

We now set up some bounding, in steps.

1. Note f is continuous on the compact set $[-1/2,1/2]$, so we may find some M_f such that $|f(x)|\leq M_f$ for all $x\in[-1/2,1/2]$.

2. Because f is continuous, we may find $\delta_f > 0$ such that $|f(x) - 1| \le 1/2$ for $|x| < \delta_f$. In particular, we see $f(x) \ge 1/2$ for $|x| < \delta_f$. By making δ_f smaller if necessary, we will enforce $\delta_f \le 1/4$.

3. Now, define $q_1(x) \coloneqq 2\varepsilon + p(x) + p(x)^{-1} = 2\varepsilon + \cos(2\pi x)$, for $\varepsilon \coloneqq \frac{2}{3}(1 - \cos(2\pi\delta_f))$. Note $\cos(2\pi x)$ is decreasing in the region in $[\delta_f, 1/2]$, so in fact

$$\varepsilon \le \frac{1}{3}(1 - \cos(2\pi x))$$

for $x \in [\delta_f, 1/2]$. Rearranging, we see

$$q_1(x) = 2\varepsilon + \cos(2\pi x) \le 1 - \varepsilon$$

for $x \in [\delta_f, 1/2]$. In fact, because $q_1(x) \ge -1 + 2\varepsilon$, we see that $|q_1(x)| \le 1 - \varepsilon$ for $x \in [\delta_f, 1/2]$. Lastly, because q_1 is even, these inequalities hold on $[-1/2, -\delta_f] \cup [\delta_f, 1/2]$.

4. Lastly, choose $\delta_q > 0$ such that $|q_1(x) - q_1(0)| \le \varepsilon$ for $|x| < \delta_q$. In particular, $q_1(x) \ge 1 - \varepsilon$ for $|x| < \delta_q$. By making δ_q smaller if necessary, we may assume $\delta_q < \delta_f$, though this is actually implied.

To finish, we define $q_N \coloneqq q_1^N$ for $N \in \mathbb{N}$. (Notably, $q_1 = q_1^1$.) The point is that $k \to \infty$ makes q_N blow up at 0 around points where f is bounded below by $1/2$, but q_N will vanish elsewhere. Indeed, using Remark A.13, we compute

$$\begin{aligned}
\int_{-1/2}^{1/2} f(x)q_N(x)\,dx &= \int_{|x|\le\delta_q} f(x)q_N(x)\,dx + \int_{\delta_q\le|x|\le\delta_f} f(x)q_N(x)\,dx + \int_{\delta_f\le|x|\le1/2} f(x)q_N(x)\,dx \\
&\ge 2\delta_q \cdot \frac{1}{2}(1+\varepsilon)^N + 2(\delta_f - \delta_q)\cdot\frac{1}{2}\cdot 0 - 2\left(\frac{1}{2} - \delta_f\right)B(1-\varepsilon)^N \\
&\ge \delta_q(1+\varepsilon)^N - \delta_f B(1-\varepsilon)^N.
\end{aligned}$$

Thus, as $N \to \infty$, the integral goes to $+\infty$. In particular, we can (in theory) find an (explicit) N such that $\int_{-1/2}^{1/2} f(x)q_N(x)\,dx > 0$. Now, we may write

$$q_N = \left(2\varepsilon + p + p^{-1}\right)^N = \sum_{n=-N}^{N} q_{N,n}p^n$$

for some coefficients $q_{N,n} \in \mathbb{R}$. Thus,

$$0 < \int_{-1/2}^{1/2} f(x)q_N(x)\,dx = \sum_{n=-N}^{N}\left(q_{N,n}\int_{-1/2}^{1/2} f(x)e^{-2\pi inx}\,dx\right) = \sum_{n=-N}^{N} q_{N,n}a_n(f),$$

where we have used Remark A.13. Thus, there exists n with $|n| \le N$ such that $a_n(f) \ne 0$. ∎

Proposition A.17. Fix a continuous 1-periodic function $f\colon \mathbb{R} \to \mathbb{C}$ such that $a_n(f) = 0$ for all $n \in \mathbb{N}$. Then $f(x) = 0$ for all $x \in \mathbb{R}$.

Proof. This follows from Lemma A.16 and the following reductions.

- It suffices to show the result for real-valued functions f. Indeed, we may write $f(x) \coloneqq u(x) + iv(x)$ for some real-valued, continuous, and 1-periodic functions $u, v\colon \mathbb{R} \to \mathbb{R}$. (Namely, $u = \operatorname{Re} f$ and $v = \operatorname{Im} f$, and each adjective is inherited from f.) However, for each $n \in \mathbb{N}$, we use Lemma A.14 to see

$$a_n(u) = a_n\left(\frac{f + \overline{f}}{2}\right) = \frac{1}{2}\left(a_n(f) + \overline{a_{-n}(f)}\right) = 0,$$

and

$$a_n(v) = a_n\left(\frac{f-\overline{f}}{2i}\right) = \frac{1}{2i}\left(a_n(f) - \overline{a_{-n}(f)}\right) = 0.$$

Thus, if we can prove the result for real-valued functions, we see $a_n(u) = a_n(v) = 0$ for all $n \in \mathbb{Z}$ forces $u = v = 0$, so $f = u + iv = 0$ also.

- It suffices to show that $f(0) = 0$, which is Lemma A.16. Indeed, for some fixed $x_0 \in \mathbb{R}$, we define $g(x) \coloneqq f(x + x_0)$. Note g is still continuous and 1-periodic. Further, Lemma A.14 tells us that $a_n(g) = e^{2\pi i n x_0} a_n(f) = 0$ for each $n \in \mathbb{Z}$. Thus, Lemma A.16 implies $g(0) = 0$, so $f(x_0) = g(0) = 0$ follows. ∎

The point is that we know the linear transformation sending a continuous 1-periodic function f to the tuple of its coefficients $\{a_n(f)\}_{n\in\mathbb{N}}$ is injective. We now expect that we can construct a partial inverse map by sending the tuple of coefficients to the corresponding Fourier series, which is what we show next.

A.4 Fourier Series

Now that we have our coefficients, we can define our Fourier series. We continue to follow [SS03, Chapter 2].

Definition A.18 (Fourier series)**.** Fix a continuous 1-periodic function $f\colon \mathbb{R} \to \mathbb{C}$. Then we define the Nth partial sum of the *Fourier series* of f as

$$S_{f,N}(x) \coloneqq \sum_{n=-N}^{N} a_n(f) e^{2\pi i n x}.$$

The *Fourier series* is defined as $S_f(x) \coloneqq \lim_{N\to\infty} S_{f,N}(x)$, when this limit converges.

The main goal of this subsection is to provide smoothness conditions on f which will imply $f(x) = S(x)$ for all $x \in \mathbb{R}$.

We will begin by figuring out when this series will converge.

Lemma A.19. Fix a twice continuously differentiable 1-periodic function $f\colon \mathbb{R} \to \mathbb{C}$. Then the series $S_f(x)$ converges absolutely and uniformly.

Proof. This follows from Lemma A.15. Indeed, for $n \neq 0$, we see that

$$a_n(f) = \frac{1}{-2\pi i n} \cdot a_n(f') = \frac{a_n(f'')}{4\pi^2 n^2}.$$

Because f'' is continuous, Remark A.12 grants $M \in \mathbb{R}$ such that $|a_n(f'')| \le M$, so it follows that $|a_n(f)| \le M/\left(4\pi^2 n^2\right)$ for $n \neq 0$. Thus, we see the series S_f converges absolutely because

$$\sum_{n\in\mathbb{Z}} \left|a_n(f) e^{2\pi i n x}\right| \le a_0(f) + \frac{2M}{4\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$$

for any $x \in \mathbb{R}$. To get the uniform convergence, for any $N \in \mathbb{N}$, we compute

$$|S_f(x) - S_{f,N}(x)| = \left|\sum_{|n|>N} a_n(f) e^{2\pi i n x}\right| \le \sum_{|n|>N} \frac{M}{4\pi^2 n^2} = \frac{2M}{4\pi^2} \sum_{n>N} \frac{1}{n^2} < \frac{2M}{4\pi^2} \int_N^\infty \frac{1}{t^2}\, dt = \frac{2M}{4\pi^2 N},$$

which does vanish as $N \to \infty$. ∎

And in this situation, we can show that our Fourier series is well-behaved.

Theorem A.20. Fix a continuous 1-periodic function $f\colon \mathbb{R} \to \mathbb{C}$. If the series S_f converges absolutely and uniformly, then $S_f(x) = f(x)$ for all $x \in \mathbb{R}$.

Proof. The point is to show that $a_n(S_f) = a_n(f)$ for all $n \in \mathbb{Z}$ so that the result will follow from Proposition A.17.

Quickly, note that the uniform convergence provided by hypothesis implies that S_f is a continuous function because the partial sums $S_{f,N}$ are continuous. Further, S_f is 1-periodic: for any $x \in \mathbb{R}$, we see

$$S_f(x+1) = \lim_{N\to\infty} \sum_{n=-N}^{N} a_n(f) e^{2\pi i n(x+1)} = \lim_{N\to\infty} \sum_{n=-N}^{N} a_n(f) e^{2\pi i n x} = S_f(x).$$

Thus, we are allowed to compute the Fourier coefficients

$$a_n(S_f) = \int_0^1 \left(\sum_{m\in\mathbb{Z}} a_n(f) e^{2\pi i (m-n)x} \right) dx$$

for $n \in \mathbb{Z}$. We would like to exchange the sum and the integral, for which we use Fubini's theorem. Indeed, we see

$$\int_0^1 \left(\sum_{m\in\mathbb{Z}} \left| a_n(f) e^{2\pi i (m-n)x} \right| \right) dx = \left(\int_0^1 dx \right) \sum_{m\in\mathbb{Z}} |a_n(f)| = \sum_{m\in\mathbb{Z}} |a_n(f)|,$$

which converges because $S_f(0)$ converges absolutely by hypothesis. Thus, Fubini's theorem lets us write

$$a_n(S_f) = \sum_{m\in\mathbb{Z}} \left(\int_0^1 a_n(f) e^{2\pi i (m-n)x}\, dx \right) = a_n(f),$$

where we have used (A.2) in the last equality. To finish the proof, we note $a_n(S_f - f) = 0$ by Lemma A.14. As such, $S_f - f = 0$ by Proposition A.17, which finishes the proof. ∎

BIBLIOGRAPHY

[Dav80] Harold Davenport. *Multiplicative number theory*. eng. Second Edition. Vol. 74. Graduate Texts in Mathematics. New York, NY: Springer, 1980. ISBN: 9781475759297.

[SS03] Elias M. Stein and Rami Shakarchi. *Fourier Analysis: An Introduction*. Vol. 1. Princeton University Press, 2003. URL: https://books.google.com/books?id=I6CJngEACAAJ.

[Dav05] Harold Davenport. *Analytic methods for Diophantine equations and Diophantine inequalities*. eng. 2nd ed. / this edition edited and prepared for publication by T, D. Browning. Cambridge mathematical library. Cambridge, UK ; Cambridge University Press, 2005. ISBN: 0521605830.

[Ser12] Jean-Pierre Serre. *A Course in Arithmetic*. Graduate Texts in Mathematics. Springer New York, 2012. URL: https://books.google.com/books?id=8fPTBwAAQBAJ.

[Shu16] Neal Shusterman. *Scythe*. Arc of a Scythe. Simon & Schuster, 2016.

List of Definitions