Contents

10 Analytic Number Theory

In this chapter, we discuss some fundamental results in analytic number theory. We begin by introducing the Riemann zeta function and establishing some of its properties, with the main goal being to establish Dirichlet's theorem on primes in arithmetic progressions, which along the way requires a fairly involved discussion of Dirichlet series, group characters, and L-series. We then construct the Dedekind zeta function of a quadratic integer ring and derive the analytic class number formula, which provides an analytic formula for the class number of a quadratic integer ring.

10.1 The Riemann Zeta Function and Dirichlet's Theorem on Primes in Arithmetic Progessions

- We begin our motivation for the use of analytic techniques in number theory by studying some basic properties of the Riemann zeta function.

10.1.1 The Riemann Zeta Function

- Here is the Riemann zeta function:
- Definition: If s is a complex number with $\text{Re}(s) > 1$, the Riemann zeta function is defined as $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$.
  - Remark: By tradition dating back to Dirichlet, the complex variable used in discussions of the zeta function is denoted $s = \sigma + it$ where $\sigma = \text{Re}(s)$ and $t = \text{Im}(s)$. (Supposedly, a printer's error changed it from the intended $s = \sigma + i\tau$.)
  - Explicitly, we have $\zeta(s) = 1 + \dfrac{1}{2^s} + \dfrac{1}{3^s} + \dfrac{1}{4^s} + \dfrac{1}{5^s} + \cdots$.
  - Observe in particular that $\zeta(1) = \sum_{n=1}^{\infty} \dfrac{1}{n}$ is the harmonic series, which famously diverges to ∞. This is the reason for our requirement that $\text{Re}(s) > 1$, to ensure the series actually converges.

- Indeed, for $\text{Re}(s) = \sigma > 1$ we have $|n^{-s}| = n^{-\sigma}$: thus $|\sum_{n=1}^{\infty} n^{-s}| \le \sum_{n=1}^{\infty} n^{-\sigma}$, and the latter series converges by comparison to the integral $\int_1^{\infty} x^{-\sigma}\,dx = \frac{x^{1-\sigma}}{1-\sigma}\Big|_{x=1}^{\infty} = \frac{1}{\sigma-1}$.
- Thus we see that the series for $\zeta(s)$ in fact converges absolutely whenever $\text{Re}(s) > 1$.

- One of the fundamental properties of the zeta function is that it can also be expressed as an infinite product.
  - The idea is quite simple and is simply an encapsulation of the unique prime factorization of positive integers.
  - As motivation, observe that $(1+\frac{1}{2^s}+\frac{1}{4^s})(1+\frac{1}{3^s}) = 1+\frac{1}{2^s}+\frac{1}{4^s}+\frac{1}{3^s}+\frac{1}{6^s}+\frac{1}{12^s} = 1+\frac{1}{2^s}+\frac{1}{3^s}+\frac{1}{4^s}+\frac{1}{6^s}+\frac{1}{12^s}$.
  - We can see that the product contains the terms n^{-s} for all n that are a product of one of $\{1,2,4\}$ with one of $\{1,3\}$.
  - For the same reason, if we multiply out the product $(1+\frac{1}{2^s}+\frac{1}{4^s}+\frac{1}{8^s})(1+\frac{1}{3^s}+\frac{1}{9^s})(1+\frac{1}{5^s})(1+\frac{1}{7^s})$ we will obtain terms n^{-s} for all n that are a product of one of $\{1,2,4,8\}$ with one of $\{1,3,9\}$ and one of $\{1,5\}$ and one of $\{1,7\}$, which in particular includes all terms n^{-s} with $1 \le n \le 10$ along with a number of others that are larger.
  - Euler's observation is that if we extend this product to include all necessary terms: namely, $1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots + \frac{1}{p^{ks}} + \cdots$ for each prime p, then the resulting distributed product over all primes p will be the sum over all terms n^{-s} for every positive integer n.
  - We can further condense the expansion by observing that $1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots + \frac{1}{p^{ks}} + \cdots$ is a geometric series with sum $\frac{1}{1-p^{-s}} = (1-p^{-s})^{-1}$.
  - Putting all of this together yields the following:

- <u>Proposition</u> (Euler Product for ζ): If s is a complex number with $\text{Re}(s) > 1$, then $\zeta(s) = \prod_{p \text{ prime}} (1-p^{-s})^{-1}$.
  - <u>Proof</u>: We only need to show that the partial products $P_d = \prod_{p \text{ prime} \le n} (1-p^{-s})^{-1}$ converge to the series for the zeta function as $d \to \infty$.
  - Each of the geometric series $(1-p^{-s})^{-1} = \sum_{k=0}^{\infty} p^{-ks}$ converges absolutely for $\text{Re}(s) > 1$, so we may arbitrarily rearrange the terms without changing the sum.
  - Therefore, if S_d represents the set of positive integers whose prime factors are all $\le d$, by expanding the product we have $\prod_{p \text{ prime} \le d} (1-p^{-s})^{-1} = \prod_{p \text{ prime} \le d} (1+p^{-s}+p^{-2s}+\cdots) = \sum_{n \in S_d} n^{-s}$.
  - Then $\left|\zeta(s) - \prod_{p \text{ prime} \le d}(1-p^{-s})^{-1}\right| = \left|\sum_{n \notin S_d} n^{-s}\right| \le \sum_{n=d+1}^{\infty} |n^{-s}|$ and the latter sum is a tail of the zeta function's expansion, hence tends to zero as $d \to \infty$.
  - We conclude that $\lim_{d\to\infty} \left|\zeta(s) - \prod_{p \text{ prime} \le d}(1-p^{-s})^{-1}\right| = 0$, meaning that $\zeta(s) = \lim_{d\to\infty} \prod_{p \text{ prime} \le d}(1-p^{-s})^{-1} = \prod_{p \text{ prime}}(1-p^{-s})^{-1}$, as claimed.

- By using suitably clever manipulations of the zeta function and its Euler product expansion, we can obtain some quite interesting results.
  - For example, as we have already noted above, $\zeta(1)$ diverges to infinity, in the sense that $\lim_{s\to 1^+} \zeta(s) = \infty$.
  - As an immediate and trivial consequence, we see that because $\prod_{p \text{ prime}}(1-p^{-1})^{-1}$ is infinite, the product must have infinitely many terms: thus, there are infinitely many primes.
  - By making slightly better estimates we can derive more interesting divergence results, such as the following:

- <u>Proposition</u> (Divergence of Prime Sum): The sum $\sum_{p \text{ prime}} 1/p$ diverges to ∞.
  - <u>Proof</u>: Taking the logarithm of product formula for the zeta function yields $\log(\zeta(s)) = \sum_{p \text{ prime}} -\log(1-p^{-s})$.

- Now, for $0 < x < 1/2$ we have the estimate $-\log(1-x) = \sum_{n=1}^{\infty} \frac{x^n}{n} < \sum_{n=1}^{\infty} x^n = \frac{x}{1-x} \leq 2x$.
- For $s > 1$ real we have $p^{-s} < 1/2$ for all primes p, so the estimate above yields $\log(\zeta(s)) = \sum_{p \text{ prime}} -\log(1-p^{-s}) \leq \sum_{p \text{ prime}} 2p^{-s}$.
- Now taking $s \to 1+$ shows that $\log(\zeta(1)) < \sum_{p \text{ prime}} 2p^{-1}$ and so the sum $\sum_{p \text{ prime}} 2p^{-1}$ must diverge to ∞. The desired result follows immediately.

- So far, we have mostly viewed $\zeta(s)$ as a function of a real variable, but considering $\zeta(s)$ as a function of a complex variable s is where the truly interesting results lie. Here are some basic complex-analytic properties of $\zeta(s)$:
- Theorem (Complex-Analytic Properties of ζ): The zeta function $\zeta(s)$ admits an analytic continuation to the region $\text{Re}(s) > 0$, and it is holomorphic on this region except for a simple pole with residue 1 at $s = 1$.
  - There are various ways to establish this fact, but most approaches rely on rearranging the series in some manner to increase the region of convergence. We will use a clever approach relying on a non-obvious rearrangement.
  - Proof: For $\text{Re}(s) > 2$, observe that all of the series in the calculation below converge absolutely, so the following rearrangements are permissible:

$$\begin{aligned}(s-1)\sum_{n=1}^{\infty}\frac{1}{n^s} &= -1+s+\sum_{n=2}^{\infty}\frac{s-1}{n^s} \\ &= \left[\sum_{n=1}^{\infty}\frac{1}{(n+1)^{s-1}} - \sum_{n=1}^{\infty}\frac{1}{n^{s-1}}\right] + \left[\sum_{n=1}^{\infty}\frac{s}{n^s} - \sum_{n=1}^{\infty}\frac{1}{(n+1)^s}\right] \\ &= \sum_{n=1}^{\infty}\frac{(n+1)}{(n+1)^s} - \sum_{n=1}^{\infty}\frac{1}{(n+1)^s} - \sum_{n=1}^{\infty}\frac{n}{n^s} + \sum_{n=1}^{\infty}\frac{s}{n^s} \\ &= \sum_{n=1}^{\infty}\left[\frac{n}{(n+1)^s} - \frac{n-s}{n^s}\right]\end{aligned}$$

  - Now notice that for $\text{Re}(s) > 0$ we have $\frac{n}{(n+1)^s} - \frac{n-s}{n^s} = n[\frac{1}{(n+1)^s} - \frac{1}{n^s}] + \frac{s}{n^s}$: then by a linearization we have $\frac{1}{(n+1)^s} - \frac{1}{n^s} \sim (-s)n^{-s-1} + O(n^{-s-1})$, so $\frac{n}{(n+1)^s} - \frac{n-s}{n^s} = O(n^{-s-1})$.
  - Therefore, by comparison to the corresponding integral, we see that $\sum_{n=1}^{\infty}\left[\frac{n}{(n+1)^s} - \frac{n-s}{n^s}\right]$ converges absolutely for all $\text{Re}(s) > 0$.
  - As a consequence, since the original definition and this new series agree for $\text{Re}(s) > 2$, we obtain an analytic continuation $\zeta(s) = \frac{1}{s-1}\sum_{n=1}^{\infty}\left[\frac{n}{(n+1)^s} - \frac{n-s}{n^s}\right]$ for all $\text{Re}(s) > 0$, except for $s = 1$.
  - Furthermore, since $\lim_{s\to 1}(s-1)\zeta(s)$ exists, we see immediately that $\zeta(s)$ has a simple pole at $s = 1$.
  - Finally, the residue at $s = 1$ is given by $\lim_{s\to 1}(s-1)\zeta(s) = \sum_{n=1}^{\infty}\left[\frac{n}{n+1} - \frac{n-1}{n}\right]$ which telescopes to 1.

10.1.2 Motivation for Dirichlet's Theorem

- Our main goal now is to establish Dirichlet's theorem on primes in arithmetic progressions:
- Theorem (Dirichlet's Theorem on Primes in Arithmetic Progressions): Suppose m is a positive integer and a is relatively prime to m. Then there exist infinitely many primes congruent to a modulo m: in other words, there are infinitely many primes in the arithmetic progression $\{a, a+m, a+2m, a+3m, \dots\}$.
  - Of course, if a is not relatively prime to m, then there cannot be very many primes congruent to a modulo m, since any integer congruent to a modulo m is divisible by $\gcd(a, m)$.

- Taking Dirichlet's theorem for granted at the moment, we see that there are $\varphi(m)$ residue classes modulo m that contain infinitely many primes. More precisely, we can still ask: how are the primes distributed among these residue classes?
  - In fact, the primes are asymptotically uniformly distributed among these residue classes: the proportion of primes congruent to a modulo m approaches $1/\varphi(m)$ upon taking an appropriate limit.
  - Explicitly, define the natural density of a set S of primes to be $\lim_{n\to\infty} \dfrac{S \cap \{1, 2, \dots, n\}}{\{\text{primes}\} \cap \{1, 2, \dots, n\}}$, provided the limit exists.
  - Then, as first proven by de la Vallée Poussin, the natural density of the primes congruent to a modulo m is $1/\varphi(m)$ when a is relatively prime to m.
- However, the natural density is somewhat difficult to handle with analytic methods. From the standpoint of zeta functions, a more natural choice is the Dirichlet density:
- Definition: If S is a set of primes, the Dirichlet density of S is the value $\delta_S = \lim_{s\to 1+} \dfrac{\sum_{p\in S \text{ prime}} p^{-s}}{\sum_{p \text{ prime}} p^{-s}}$, assuming the limit exists.
  - Note that the sum in the numerator is always finite for $\mathrm{Re}(s) > 1$ by comparison to the sum for the zeta function.
  - It is not hard to see that if S is finite, then its Dirichlet density is 0, since the numerator term is bounded while the denominator is unbounded, since as we proved earlier the sum $\sum_{p \text{ prime}} 1/p$ diverges.
  - More generally, one may prove that if a set has natural density δ, then its Dirichlet density is also δ.
  - The converse is not true, however. A (relatively) simple counterexample due to Serre is the set S of primes whose leading digit is 1 in base 10: its Dirichlet density is $\log_{10} 2$, but its natural density is undefined. Intuitively, the natural density of this set fluctuates too much when taking the limit, because there are so many n-digit primes with leading digit 1 relative to the number of primes having at most $n-1$ digits that the natural density limit does not converge.
- Our approach (and indeed, Dirichlet's original approach) is to show that the Dirichlet density of the set of primes congruent to a (mod m) has positive Dirichlet density: by our observation that finite sets have Dirichlet density zero, this would immediately imply that there are infinitely many primes congruent to a (mod m).

10.1.3 Dirichlet Series

- A classical area of study in elementary number theory over $\mathbb{Z}$ consists of the arithmetic functions related to divisors, such as the Euler φ-function, the divisor-counting function, and the sum-of-divisors function.
  - All of these are examples of multiplicative functions, which have the property that $f(ab) = f(a)f(b)$ whenever a, b are relatively prime.
  - Note the infelicitous terminology: if $f(ab) = f(a)f(b)$ for all a, b (rather than just relatively prime a, b) then f is instead called completely multiplicative.
  - If f is multiplicative then $f(1) = f(1)^2$ so either $f(1) = 0$ or $f(1) = 1$. The former case is quite trivial since it implies $f(a) = 0$ for all a, so we will also usually assume $f(1) = 1$ when f is multiplicative.
  - Then if n has prime factorization $n = \prod_i p_i^{a_i}$ and f is multiplicative, we immediately have $f(n) = \prod_i f(p_i^{a_i})$: thus, characterizing the values of a multiplicative function requires only knowing its values on prime powers.
- It is a standard principle in combinatorics that if we want to understand a sequence $\{a_n\}_{n\geq 1} = \{a_0, a_1, a_2, \dots\}$ we should study its generating function $f(x) = \sum_{n=0}^{\infty} a_n x^n$.
  - Since all of the terms of the sequence can be extracted from the function (they are, after all, merely its coefficients) the generating function, in principle, encapsulates all possible information about the sequence $\{a_n\}$.

- Typically, if the sequence has some kind of convenient description (e.g., a recurrence relation) then this information carries some kind of implication for the generating function: e.g., that it has some specific algebraic form, or satisfies a particular differential equation, or something similar of that nature.
- For example, if F_n is the nth Fibonacci number satisfying the recurrence $F_{n+1} = F_n + F_{n-1}$ for all $n \geq 1$, with associated generating function $f(x) = \sum_{n=0}^{\infty} F_n x^n$, then the recurrence implies that $(1 - x - x^2)f(x) = F_0 + (F_1 - F_0)x + (F_2 - F_1 - F_0)x^2 + \cdots = 1 + x$, and so we obtain an explicit formula $f(x) = \dfrac{1+x}{1-x-x^2}$.

- We would like to use a similar approach for studying multiplicative functions.
  - A natural first guess would be to use the same type of power series $F(x) = \sum_{n=0}^{\infty} f(n)x^n$.
  - However, this type of generating function is useful primarily for functions that behave additively. For number-theoretic functions, we instead want to use a different type of series, one that behaves multiplicatively.

- Definition: If $h : \mathbb{N} \to \mathbb{C}$ is a complex-valued function defined on positive integers, its associated Dirichlet series is $D_h(s) = \sum_{n=1}^{\infty} \dfrac{h(n)}{n^s}$.
  - Example: If $h(n) = 1$ for all n, then $D_h(s) = \sum_{n=1}^{\infty} \dfrac{1}{n^s} = \zeta(s)$, the Riemann zeta function. (So even for this extremely trivial function, we already obtain a very interesting Dirichlet series!)
  - In order for this series to converge, we need h not to grow too fast. One may check that if $|h(n)| = O(n^\alpha)$ then $D_h(s)$ is absolutely convergent for $\mathrm{Re}(s) > 1 + \alpha$ by comparison to the p-series $\sum_{n=1}^{\infty} n^{\alpha - \mathrm{Re}(s)}$.
  - We will mostly be able to ignore issues of convergence for the moment, since our functions will grow polynomially at worst.
  - In most cases, therefore, we may manipulate the series formally without worrying explicitly about convergence: in other words, by treating s as an indeterminate rather than some specific real or complex number.
  - If h is multiplicative, then it is a straightforward calculation to see that $D_h(s)$ has an Euler product expansion: $D_h(s) = \prod_{p \text{ prime}} (1 + \dfrac{h(p)}{p} + \dfrac{h(p^2)}{p^2} + \cdots)$, on the appropriate domain of convergence. (The argument is the same as for the Riemann zeta function: simply multiply out and collect terms.)

- The key property of Dirichlet series is that they reproduce desired number-theoretic behaviors under multiplication.
  - To motivate the idea, suppose f and g are any functions, and multiply out the product of the Dirichlet series

$$\begin{aligned} D_f(s)D_g(s) &= \left[\frac{f(1)}{1^s} + \frac{f(2)}{2^s} + \frac{f(3)}{3^s} + \frac{f(4)}{4^s} + \cdots\right] \cdot \left[\frac{g(1)}{1^s} + \frac{g(2)}{2^s} + \frac{g(3)}{3^s} + \frac{g(4)}{4^s} + \cdots\right] \\ &= \frac{f(1)g(1)}{1^s} + \frac{f(1)g(2) + f(2)g(1)}{2^s} + \frac{f(1)g(3) + f(3)g(1)}{3^s} + \frac{f(1)g(4) + f(2)g(2) + f(4)g(1)}{4^s} + \cdots. \end{aligned}$$

  - We can see in general that the coefficient of n^{-s} is the sum over divisors given by $\sum_{d|n} f(d)g(n/d)$.
  - This is a far more natural sum to consider for number-theoretic functions than the power series coefficient for x^n in the product $\sum_{n=0}^{\infty} f(n)x^n \cdot \sum_{n=0}^{\infty} g(n)x^n$, which is $\sum_{d=0}^{n} f(d)g(n-d)$.

- We record the definition of these coefficients:

- Definition: If $f, g : \mathbb{N} \to \mathbb{C}$ are functions, then their Dirichlet convolution $f * g$ is defined via $(f * g)(n) = \sum_{d|n} f(d)g(n/d)$ for each $n \geq 1$.
  - Equivalently, and more symmetrically, we have $(f * g)(n) = \sum_{ab=n} f(a)g(b)$.

- Here are various fundamental properties of Dirichlet convolution:

- Proposition (Properties of Dirichlet Convolution): Let $f, g, h : \mathbb{N} \to \mathbb{C}$ be functions with associated Dirichlet series $D_f(s)$, $D_g(s)$, and $D_h(s)$, and let $*$ represent Dirichlet convolution, where either we work with formal series or in a region where all series converge absolutely.

  1. We have $D_f(s) \cdot D_g(s) = D_{f*g}(s)$.
     - Proof: We have $D_f(s)D_g(s) = \sum_{a=1}^{\infty} \frac{f(a)}{a^s} \sum_{b=1}^{\infty} \frac{g(b)}{b^s} = \sum_{a=1}^{\infty} \sum_{b=1}^{\infty} \frac{f(a)g(b)}{(ab)^s} = \sum_{n=1}^{\infty} \frac{1}{n^s} \sum_{ab=n} f(a)g(b) = \sum_{n=1}^{\infty} \frac{(f*g)(n)}{n^s} = D_{f*g}(s)$.
  2. Dirichlet convolution is commutative and associative, and has an identity element given by $I(n) = \begin{cases} 1 & \text{for } n = 1 \\ 0 & \text{for } n > 1 \end{cases}$.
     - Proof: For commutativity we have $(f * g)(n) = \sum_{ab=n} f(a)g(b) = \sum_{ab=n} g(a)f(b) = (g * f)(n)$ for each n.
     - For associativity we have $[(f*g)*h](n) = \sum_{bc=n}(f*g)(b)h(c) = \sum_{abc=n} f(a)g(b)h(c) = \sum_{ab=n} f(a)(g*h)(b) = [f * (g * h)](n)$ for each n.
     - For the identity we have $(f * I)(n) = \sum_{ab=n} f(a)I(b) = f(n)$ since all terms have $I(b) = 0$ except for the one with $b = 1$.
  3. The function f has an inverse under Dirichlet convolution if and only if $f(1) \neq 0$. In such a case, $D_{f^{-1}}(s) = \frac{1}{D_f(s)}$.
     - Proof: If $(f * g)(n) = I(n)$ then setting $n = 1$ yields $f(1)g(1) = 1$, so we must have $f(1) \neq 0$.
     - Conversely, if $f(1) \neq 0$ we can solve recursively for the values $g(n)$ using the recurrence $f(1)g(1) = 1$ and $\sum_{d|n} f(d)g(n/d) = 0$ for all $n \geq 2$.
     - Explicitly, we obtain $g(1) = f(1)^{-1}$ and $g(n) = -f(1)^{-1} \sum_{d|n,d>1} f(d)g(n/d)$ for each $n \geq 2$. The function g defined this way satisfies $f * g = I$, so it is an inverse of f.
     - The second statement follows immediately from (1) and the fact that $D_I(s) = 1$.
  4. If f is multiplicative with $f(1) = 1$, then its Dirichlet inverse f^{-1} is also multiplicative.
     - Proof: By using the explicit recurrence $g(1) = f(1)^{-1}$ and $g(n) = -f(1)^{-1} \sum_{d|n,d>1} f(d)g(n/d)$ for each $n \geq 2$ derived above in (3) for the Dirichlet inverse $g = f^{-1}$, we see that $g(1) = 1$ and $g(n) = -\sum_{d|n,d>1} f(d)g(n/d)$.
     - Then by a trivial induction on n, since each term in the sum $\sum_{d|n,d>1} f(d)g(n/d)$ is multiplicative (for f by definition, and for g by the induction hypothesis), the sum for g itself is also multiplicative.
  5. If two of f, g, and $f * g$ are multiplicative, then the third is also.
     - Proof: If f and g are multiplicative, then for $h = f * g$ and a, b relatively prime, then since any divisor $d|ab$ decomposes uniquely as $d = rs$ for $r|a$ and $s|b$, we obtain $h(ab) = \sum_{d|ab} f(d)g(ab/d) = \sum_{r|a,s|b} f(rs)g(ab/(rs)) = \sum_{r|a,s|b} f(r)f(s)g(a/r)g(b/s) = [\sum_{r|a} f(r)g(a/r)][\sum_{s|b} f(s)g(b/s)] = h(a)h(b)$ as desired.
     - If f and $f*g$ are multiplicative, then so is f^{-1} by (4), and thus so is $f^{-1}*[f*g] = [f^{-1}*f]*g = I*g = g$ by the calculation above and (2). The case where g and $f * g$ are multiplicative follows in the same way.

- By exploiting Dirichlet convolution, we can find the Dirichlet series for many basic multiplicative functions in terms of the Riemann zeta function:

- Proposition (Dirichlet Series Evaluations): Let n be a positive integer. Then the following hold:

  1. For $I(n) = \begin{cases} 1 & \text{for } n = 1 \\ 0 & \text{for } n > 1 \end{cases}$ we have $D_I(s) = 1$.
  2. For $1(n) = 1$ (for all n) we have $D_1(s) = \zeta(s)$.
  3. For $N(n) = n$ (for all n) we have $D_N(s) = \zeta(s - 1)$.

◦ Proofs: (1) and (2) are trivial. For (3), $D_N(s) = \sum_{n=1}^{\infty} \frac{n}{n^s} = \sum_{n=1}^{\infty} \frac{1}{n^{s-1}} = \zeta(s-1)$.

4. For the Möbius function $\mu(n) = \begin{cases} 0 & \text{if } n \text{ is not squarefree} \\ (-1)^r & \text{if } n \text{ is the product of } r \text{distinct primes} \end{cases}$ we have $\mu * 1 = I$ and $D_\mu(s) = \frac{1}{\zeta(s)}$.
   - Proof: First we observe that $\mu * 1 = I$, since $(\mu * 1)(n) = \sum_{d|n} \mu(d) 1(n/d) = \sum_{d|n} \mu(d) = \begin{cases} 1 & \text{for } n = 1 \\ 0 & \text{for } n > 1 \end{cases} = I(n)$ as follows by a straightforward induction.
   - Therefore, by multiplicativity of the Dirichlet series, we see $D_\mu(s) D_1(s) = D_I(s)$ so $D_\mu(s) = \frac{1}{\zeta(s)}$.
5. (Möbius Inversion) If $g(n) = \sum_{d|n} f(n)$ for each $n \geq 1$, then $f(n) = \sum_{d|n} \mu(d) g(n/d)$ for each $n \geq 1$.
   - Proof: The hypothesis on g states that $g = f * 1$, so since $\mu * 1 = I$, convolving with μ and using associativity yields $f = \mu * g$, which directly states that $f(n) = \sum_{d|n} \mu(d) g(n/d)$.
6. For the Euler φ-function φ, we have $D_\varphi(s) = \frac{\zeta(s-1)}{\zeta(s)}$.
   - Proof: First we observe that $\sum_{d|n} \varphi(d) = n$, since $\varphi(d)$ counts the number of integers A less than or equal to n with $\gcd(A, n) = n/d$. Summing over all possible gcds yields $\sum_{d|n} \varphi(d) = n$.
   - Equivalently, this result says that $\varphi * 1 = N$, so by composing with μ and using associativity, we see that $\varphi = \mu * N$.
   - Then $D_\varphi(s) = D_\mu(s) D_N(s) = \frac{\zeta(s-1)}{\zeta(s)}$.
   - Remark: In principle, we could have established this fornula for $D_\varphi(s)$ by manipulating the zeta function directly, but this method is both more difficult and requires knowing the actual (non-obvious) formula for the answer ahead of time.
7. For the divisor-counting function $d(n) = \sum_{d|n} 1$, we have $D_\varphi(s) = \zeta(s)^2$.
   - Proof: The definition equivalently states that $d(n) = \sum_{d|n} 1(d) 1(d/n)$ so $d = 1 * 1$. Then $D_d(s) = D_1(s)^2 = \zeta(s)^2$.
8. For the sum-of-divisors function $\sigma(n) = \sum_{d|n} d$, we have $D_\sigma(s) = \zeta(s-1)\zeta(s)$.
   - Proof: The definition equivalently states that $\sigma(n) = \sum_{d|n} N(d) 1(n/d)$ so $\sigma = N * 1$. Then $D_\sigma(s) = D_N(s) D_1(s) = \zeta(s-1)\zeta(s)$.

- There are many other useful arithmetic function whose Dirichlet series can be similarly computed explicitly, but we will content ourselves with these.
  - One of the main applications of computing the Dirichlet series for these various arithmetic functions is that we can extract information about growth rates from them.

10.1.4 Group Characters and Dirichlet Characters

- In order to progress towards Dirichlet's theorem, we require some basic facts about group characters, as they will will allow us to study the Dirichlet series that picks out primes congruent to a (mod m).
- Definition: Let G be a finite abelian group. A group character χ of G is a homomorphism $\chi : G \to \mathbb{C}^\times$.
  - Note that $\chi(1) = 1$ for every character, and also if $g \in G$ has order d, then $1 = \chi(1) = \chi(g^d) = \chi(g)^d$, so $\chi(g)$ is a dth root of unity. Thus in general, χ is a map from G to the group of complex $|G|$th roots of unity.
  - Example: For any G, the trivial character χ_{triv} has $\chi_{\text{triv}}(g) = 1$ for all $g \in G$.
  - Example: If $G = (\mathbb{Z}/p\mathbb{Z})^\times$, the quadratic residue symbol $\chi(a) = \left(\frac{a}{p}\right)$ is a group character.

- Example: If $G = (\mathbb{Z}/p\mathbb{Z})^\times$ has generator g (of order $p-1$), the map $\chi(g^d) = e^{2\pi i d/(p-1)}$ is a group character.
- Example: If $P = (\pi)$ is a prime ideal of $R = \mathbb{Z}[i]$, and $G = (R/P)^\times$, the quartic residue symbol $\chi(a) = \left[\frac{a}{\pi}\right]_4$ is a group character.

- We will be interested in the case where G is the group of units $(\mathbb{Z}/m\mathbb{Z})^\times$ in which case we call χ a Dirichlet character.
  - In some situations it is slightly more convenient to work with extended Dirichlet characters, which we extend to have domain $\mathbb{Z}/m\mathbb{Z}$ by setting $\chi(a) = 0$ whenever a is not relatively prime to m.
  - Equivalently, extended Dirichlet characters modulo m are functions $\chi : \mathbb{Z} \to \mathbb{C}$ such that (i) $\chi(a + bm) = \chi(a)$ for all a, b, (ii) $\chi(ab) = \chi(a)\chi(b)$ for all a, b, and (iii) $\chi(a) \neq 0$ if and only if a is relatively prime to m.
- We can multiply two group characters on G pointwise, and this operation makes them into a group:
- Proposition (Dual Group of G): The set of group characters on G forms a group under pointwise multiplication. The identity is the trivial character and the inverse of χ is its complex conjugate $\overline{\chi}$. This group is called the dual group of G and is denoted $\hat{G}$.
  - Proof: These properties can be checked directly (they are quite straightforward), or one may simply note that $\hat{G} = \text{Hom}(G, \mathbb{C}^\times)$, the set of group homomorphisms from G to $\mathbb{C}^\times$, which naturally carries a group structure.
- The dual group $\hat{G}$ is also an abelian group, so it is natural to wonder how its structure relates to G. In fact, it is isomorphic to G:
- Proposition (Dual Group, II): If G is a finite abelian group, its dual group $\hat{G}$ is isomorphic to G.
  - Proof: First consider the special case where G is a cyclic group of order n generated by g. Then $\chi(g^d) = \chi(g)^d$ for all d, so any group character χ is uniquely determined by the value of $\chi(g)$, which must be some nth root of unity.
  - Conversely, any such selection $e^{2\pi i a/n}$ for $\chi(g)$ yields a valid group character χ_a, namely with $\chi_a(g^d) = e^{2\pi i a d/n}$. Since $\chi_a \chi_b = \chi_{a+b}$ and χ_1^n is the trivial character, we see that the dual group $\hat{G}$ is cyclic of order n (the map $a \mapsto \chi_a$ is an isomorphism of $\mathbb{Z}/n\mathbb{Z}$ with $\hat{G}$).
  - Now suppose $G = H \times K$ is a direct product. If $\chi : H \times K \to \mathbb{C}^\times$ is a homomorphism, let $\chi_H : H \to \mathbb{C}^\times$ and $\chi_K : K \to \mathbb{C}^\times$ be the projections $\chi_H(h) = \chi(h, 1)$ and $\chi_K(k) = \chi(1, k)$. Then χ_H is a group character of H, χ_K is a group character of K, and $\chi(h, k) = \chi_H(h)\chi_K(k)$.
  - Conversely, any pair $(\chi_H, \chi_K) \in (\hat{H}, \hat{K})$ yields a character $\chi(h, k) = \chi_H(h)\chi_K(k) \in \hat{G}$, so we see $\hat{G} \cong \hat{H} \times \hat{K}$.
  - Since every finite abelian group is a direct product of cyclic groups, and the result holds for cyclic groups and direct products, we are done.
- The isomorphism between $\hat{G}$ and G above is non-canonical (i.e., it is not "coordinate-free", in the sense that we must pick specific generators for G and $\hat{G}$ to obtain the isomorphism).
  - However, there is a canonical isomorphism between $\hat{\hat{G}}$ (the double dual) and G given by the "evaluation map" φ, which maps an element $g \in G$ to the "evaluation-at-g" map e_g on characters $\chi \in \hat{G}$, defined by $e_g(\chi) = \chi(g)$.
  - This result is a special case of Pontryagin duality, and has an analogous statement for duals of finite-dimensional vector spaces.
  - In fact, it is really the algebraic analogue of Fourier inversion (the reason being that Fourier analysis on finite abelian groups involves sums over group characters in lieu of integrals). For a brief taste of the analogy, the main idea is to note that the map $e^{inx} : \mathbb{R} \to \mathbb{C}^\times$ is a group homomorphism, and thus is an "$\mathbb{R}$"-character.

- We can also put the structure of an inner product on group characters. To establish this we first show some simple orthogonality relations:
- Proposition (Orthogonality Relations): If G is a finite abelian group and χ is a group character, the following hold:

  1. The sum $\sum_{g \in G} \chi(g) = \begin{cases} |G| & \text{if } \chi \text{is trivial} \\ 0 & \text{otherwise} \end{cases}$.
     - Proof: If χ is trivial the sum is clearly $|G|$. If χ is not trivial, say with $\chi(h) \neq 1$, then $\sum_{g \in G} \chi(g) = \sum_{g \in G} \chi(gh) = \chi(h) \sum_{g \in G} \chi(g)$ by reindexing (since $G = Gh$), and so $\sum_{g \in G} \chi(g) = 0$.
  2. The sum $\sum_{\chi \in \hat{G}} \chi(g) = \begin{cases} |G| & \text{if } g = 1 \\ 0 & \text{otherwise} \end{cases}$.
     - Proof: Apply duality to (1).
  3. (Orthogonality 1) For any characters χ_1 and χ_2, $\sum_{g \in G} \chi_1(g) \overline{\chi_2(g)} = \begin{cases} |G| & \text{if } \chi_1 = \chi_2 \\ 0 & \text{otherwise} \end{cases}$.
     - Proof: Apply (1) to $\chi = \chi_1 \overline{\chi_2}$.
  4. (Orthogonality 2) For any elements g_1 and g_2, $\sum_{\chi \in \hat{G}} \chi(g_1) \overline{\chi(g_2)} = \begin{cases} |G| & \text{if } g_1 = g_2 \\ 0 & \text{otherwise} \end{cases}$.
     - Proof: Apply (2) to $g = g_1 g_2^{-1}$, or apply duality to (3).
  5. The pairing $\langle f_1, f_2 \rangle_G = \frac{1}{|G|} \sum_{g \in G} f_1(g) \overline{f_2(g)}$ is a complex inner product on functions $f : G \to \mathbb{C}$, and the elements of the dual group $\hat{G}$ are an orthonormal basis with respect to this inner product.
     - Proof: The inner product axioms are straightforward, and the fact that $\hat{G}$ yields an orthonormal basis follows from (3).
  6. The pairing $\left\langle \hat{f}_1, \hat{f}_2 \right\rangle_{\hat{G}} = \frac{1}{|G|} \sum_{\chi \in \hat{G}} \hat{f}_1(\chi) \overline{\hat{f}_2(\chi)}$ is a complex inner product on functions $\hat{f} : \hat{G} \to \mathbb{C}$, and the elements of G are an orthonormal basis with respect to this inner product.
     - Proof: The inner product axioms are straightforward, and the fact that $G \cong \hat{\hat{G}}$ yields an orthonormal basis follows from (4), or apply duality to (5).
  7. (Fourier Inversion) For any function $f : G \to \mathbb{C}$, with the Fourier transform $\hat{f} : \hat{G} \to \mathbb{C}$ defined by $\hat{f}(\chi) = \langle f, \chi \rangle_G = \frac{1}{|G|} \sum_{g \in G} f(g) \overline{\chi(g)}$, we have $f(g) = \sum_{\chi \in \hat{G}} \hat{f}(\chi) \chi(g)$ for all $g \in G$.
     - Proof: This follows immediately from (5), since the elements of $\hat{G}$ are an orthonormal basis.

10.1.5 Dirichlet L-Series and Dirichlet's Theorem

- With the fundamentals taken care of, we can now focus on Dirichlet characters.
  - Studying primes congruent to a modulo m naturally leads to a question about Dirichlet characters via Fourier inversion, since we may decompose the characteristic function of [primes congruent to a modulo m] as a sum over Dirichlet characters for the group $G = (\mathbb{Z}/m\mathbb{Z})^\times$.
  - Explicitly, if $\delta_a(p)$ is 1 when $p \equiv a \pmod{m}$ and 0 otherwise, then $\hat{\delta_a}(\chi) = \frac{1}{\Phi(m)} \sum_{g \in G} \delta_a(g) \overline{\chi(g)} = \frac{1}{\Phi(m)} \overline{\chi(a)}$, since the only nonzero value of $\delta_a(g)$ occurs when $g \equiv a \pmod{m}$.
  - Then by Fourier inversion we have $\delta_a(p) = \sum_{\chi \in \hat{G}} \hat{\delta_a}(\chi) \chi(p) = \sum_{\chi \in \hat{G}} \frac{1}{\Phi(m)} \overline{\chi(a)} \chi(p)$. So the numerator for the Dirichlet density is $\sum_{p \equiv a \ (\text{mod } m)} |p|^{-s} = \sum_p \delta_a(p) |p|^{-s} = \frac{1}{\Phi(m)} \sum_{\chi \in \hat{G}} \left[\overline{\chi(a)} \sum_p \chi(p) |p|^{-s} \right]$.

- This is a bit complicated, but the point is that we have a sum over the Dirichlet characters of constants (namely $\overline{\chi(a)}$) times $\sum_p \dfrac{\chi(p)}{p^s}$, which is quite close to the Dirichlet series for the character χ: the only difference is that we are only summing over primes, rather than all integers.
- As we will see, we will be able to extract this sum over primes from the full Dirichlet series, which we now examine more closely.
- The main reason we go to this effort to use Fourier inversion is that the Dirichlet series for Dirichlet characters behave very nicely (far more nicely than the original series over primes congruent to a modulo m) because Dirichlet characters are completely multiplicative.

- Definition: If χ is a Dirichlet character modulo m, we define its associated Dirichlet L-series $L(s,\chi) = \sum_{n=1}^{\infty} \dfrac{\chi(n)}{n^s}$.
  - Note that this is just the Dirichlet series for χ, as we defined it previously. It is traditional to denote these series with the letter L (which was the letter Dirichlet used for such functions).
  - Since $|\chi(n)| \leq 1$ for all n, the series converges absolutely for $\mathrm{Re}(s) > 1$ by comparison to the series for the zeta function.
  - Furthermore, because Dirichlet characters are completely multiplicative, the L-series has a very simple Euler product: explicitly, $L(s,\chi) = \prod_{p \text{ prime}} [1-\chi(p)p^{-s}]^{-1}$, for $\mathrm{Re}(s) > 1$.
  - The Euler product is the key to calculating the Dirichlet density we wanted earlier: taking the logarithm of the Euler product gives $\log L(s,\chi) = -\sum_{p \text{ prime}} \log(1-\chi(p)/p^s) \approx \sum_{p \text{ prime}} \dfrac{\chi(p)}{p^s}$ using the Taylor approximation $-\log(1-x) \approx x$ which is accurate for small $|x|$.

- So our main task is to determine what happens to $\log L(s,\chi)$ as $s \to 1$, since this is the required input for calculating the Dirichlet density of the primes congruent to a modulo m.

- Example: For the trivial character χ_{triv} modulo m, we have $L(s,\chi_{\text{triv}}) = \prod_{p|m \text{ prime}}(1-p^{-s}) \cdot \zeta(s)$, since the terms with $p|m$ are missing from the Euler product for $L(s,\chi)$.
  - In particular, we see that $L(s,\chi_{\text{triv}})$ has an analytic continuation to the region with $\mathrm{Re}(s) > 0$, since $\zeta(s)$ does, and it has a single simple pole at $s = 1$.

- For other characters, the L-series is better behaved. In order to establish this fact, we require the discrete analogue of integration by parts:

- Proposition (Abel Summation): Suppose $\{a_n\}_{n\geq 1}$ and $\{b_n\}_{n \geq 1}$ are two complex sequences, and define $S_n = \sum_{k=1}^{n} a_k b_k$ and $B_n = \sum_{k=1}^{n} b_k$. Then $S_n = a_n B_n + \sum_{k=1}^{n-1} B_k(a_k - a_{k+1})$, and furthermore if $a_n B_n \to 0$ as $n \to \infty$, then $\sum_{k=1}^{\infty} a_k b_k = \sum_{k=1}^{\infty} B_k(a_k - a_{k+1})$.
  - Proof: First we show that $S_n = a_n B_n + \sum_{k=1}^{n-1} B_k(a_k - a_{k+1})$ by induction. The base case is trivial, and for the inductive step suppose that $S_n = a_n B_n + \sum_{k=1}^{n-1} B_k(a_k - a_{k+1})$.
  - Then $S_{n+1} = a_{n+1}b_{n+1} + S_n = a_{n+1}b_{n+1} + [a_n B_n + \sum_{k=1}^{n-1} B_k(a_k - a_{k+1})] = a_{n+1}B_{n+1} + (a_n - a_{n+1})B_n + \sum_{k=1}^{n-1} B_k(a_k - a_{k+1}) = a_{n+1}B_{n+1} + \sum_{k=1}^{n} B_k(a_k - a_{k+1})$ as required.
  - The second statement follows immediately by taking $n \to \infty$.

- Now we can establish our first main result on the L-series for nontrivial Dirichlet characters.

- Proposition (L-Series for Nontrivial Characters): Let m be a modulus and χ be a nontrivial Dirichlet character modulo m. Then $L(s,\chi)$ has an analytic continuation to the region with $\mathrm{Re}(s) > 0$, and has no poles in this region.
  - Proof: First we show that the partial sums $B_N = \sum_{n=1}^{N} \chi(n)$ are bounded uniformly in N. Divide N by m to write $N = qm + r$ with $0 \leq r \leq m-1$.
  - Then since χ is periodic modulo m we have $\sum_{n=1}^{N} \chi(n) = q\sum_{n=0}^{m-1} \chi(n) + \sum_{n=1}^{r} \chi(n)$, and since $\sum_{n=0}^{m-1} \chi(n) = \sum_{n=0}^{m-1} \chi(n)\chi_{\text{triv}}(n) = 0$ by the orthogonality relations, we see $\sum_{n=1}^{N} \chi(n) = \sum_{n=1}^{r} \chi(n)$.

- Therefore, B_N is periodic modulo m and it is uniformly bounded above in absolute value by the maximum value of $|\sum_{n=1}^{r}\chi(n)|$ for $0 \le r \le m-1$. (Certainly this sum is at most $r \le \varphi(m)$ by the triangle inequality.)
- Now apply Abel summation with $a_n = n^{-s}$ and $b_n = \chi(n)$: since B_N is bounded and $a_n \to 0$ we have $a_n B_n \to 0$.
- Hence by the proposition, if we define $B_x = B_{\lfloor x \rfloor}$, then for $\mathrm{Re}(s) > 1$ we have $L(s,\chi) = \sum_{n=1}^{\infty} B_n[n^{-s} - (n+1)^{-s}] = \sum_{n=1}^{\infty} B_n s \int_n^{n+1} x^{-s-1}\,dx = s\int_1^{\infty} B_x x^{-s-1}\,dx$.
- By our bounding estimate we have $|B_x| \le \varphi(m)$, so the integral converges absolutely hence yields an analytic function for all $\mathrm{Re}(s) > 0$.

- As a consequence, we see that $L(s,\chi)$ has no pole at $s=1$ when $\chi \ne \chi_{\mathrm{triv}}$. Our next major goal is to prove that $L(1,\chi) \ne 0$ for $\chi \ne \chi_{\mathrm{triv}}$.
- Lemma (Character Product): Let χ be any Dirichlet character modulo m. Then for each prime p not dividing m, there exist a, b with $ab = \varphi(m)$ such that $\prod_{\chi \in \hat{G}} L(s,\chi) = \prod_{p \nmid m}(1 - |p|^{-as})^{-b}$.
  - Proof: For a fixed prime $p \nmid m$, as we have previously noted, the evaluation-at-p map $\chi \mapsto \chi(p)$ is a homomorphism from $\hat{G}$ to $\mathbb{C}^{\times}$.
  - Let the image be a cyclic group of order a and the kernel have size b: then $ab = \#\hat{G} = \#G = \varphi(m)$ by the first isomorphism theorem.
  - For this p, by grouping the fibers of the evaluation-at-p map together, for $\zeta = e^{2\pi i/p}$ we have $\prod_{\chi \in \hat{G}}(1 - \chi(p)p^{-s})^{-1} = \prod_{j=0}^{p-1}(1 - \zeta^j p^{-s})^{-b}$, and this last product equals $(1 - p^{-as})^{-b}$ since it is the evaluation of the polynomial $(1-t)(1-\zeta t)\cdots(1-\zeta^{p-1}t) = 1 - t^p$ at $t = p^{-s}$.
  - Thus, taking the product over all primes $p \nmid m$ yields the claimed $\prod_{\chi \in \hat{G}} L(s,\chi) = \prod_{\chi \in \hat{G}}\prod_{p \nmid m}(1 - \chi(p)p^{-s})^{-1} = \prod_{p \nmid m}(1 - p^{-as})^{-b}$ after reversing the order of the products.
- We next show that $L(1,\chi) \ne 0$ for nonreal Dirichlet characters χ:
- Lemma (Nonvanishing, I): Let χ be any Dirichlet character modulo m such that $\chi \ne \overline{\chi}$. Then $L(1,\chi) \ne 0$.
  - Proof: If we expand the product $\prod_{\chi \in \hat{G}} L(s,\chi) = \prod_{p \nmid m}(1 - p^{-as})^{-b}$ from the Lemma above, it yields a Dirichlet series with nonnegative coefficients and constant term 1.
  - Thus, if s is real and greater than 1 (so that the product converges), the value of the product is real and greater than 1.
  - If $\chi \ne \overline{\chi}$, then $\prod_{\chi \in \hat{G}} L(s,\chi) = L(s,\chi_{\mathrm{triv}})L(s,\chi)L(s,\overline{\chi}) \cdot [\text{other terms}]$.
  - Now suppose $L(1,\chi) = 0$: then we would have $L(1,\overline{\chi}) = 0$ also. But this would mean the product $\prod_{\chi \in \hat{G}} L(s,\chi)$ vanishes at $s=1$, because the only term that has a pole at $s=1$ is $L(s,\chi_{\mathrm{triv}})$ and that pole has order 1, but we have two zeroes at $s=1$ arising from $L(s,\chi)$ and $L(s,\overline{\chi})$.
  - But this is impossible because the value of the product is real and greater than 1 for $s > 1$. Thus, $L(1,\chi) \ne 0$.
- The case where $\chi = \overline{\chi}$ and $\chi \ne \chi_{\mathrm{triv}}$ (i.e., when χ has order 2 in $\hat{G}$) is quite a bit trickier, since we cannot get away with such a simple order-of-vanishing argument.
- Lemma (Nonvanishing, II): Let χ be any Dirichlet character of order 2 modulo m (i.e., such that $\chi = \overline{\chi}$ but $\chi \ne \chi_{\mathrm{triv}}$). Then $L(1,\chi) \ne 0$.
  - Proof: Suppose $\chi = \overline{\chi}$ but $\chi \ne \chi_{\mathrm{triv}}$, so that $\chi(p) \in \{\pm 1\}$ for $p \nmid m$, and also suppose by way of contradiction that $L(1,\chi) = 0$.
  - Define the function $G(s) = \dfrac{L(s,\chi_{\mathrm{triv}})L(s,\chi)}{L(2s,\chi_{\mathrm{triv}})} = \prod_{p \nmid m}\dfrac{(1-p^{-s})^{-1}(1-\chi(p)p^{-s})^{-1}}{(1-p^{-2s})^{-1}} = \prod_{p \nmid m}\dfrac{1+p^{-s}}{1-\chi(p)p^{-s}} = \prod_{p \nmid m, \chi(p)=1}\dfrac{1+p^{-s}}{1-p^{-s}} = \prod_{p \nmid m, \chi(p)=1}[1 + 2\sum_{k=1}^{\infty} p^{-ks}]$, where these manipulations are valid for $\mathrm{Re}(s) > 1$.

- By expanding this last expression for G, we can see that its Dirichlet series has all coefficients nonnegative: say $G(s) = \sum_{n=1}^{\infty} a_n n^{-s}$.
- The two numerator functions $L(s, \chi_{\text{triv}})$ and $L(s, \chi)$ are both analytic for $\text{Re}(s) > 0$ except for the pole of $L(s, \chi_{\text{triv}})$ at $s = 1$, but this is cancelled by the hypothesized zero of $L(s, \chi)$ at $s = 1$. Hence the numerator is analytic for $\text{Re}(s) > 0$.
- The denominator function $L(2s, \chi_{\text{triv}})$ is also analytic for $\text{Re}(s) > 0$ except for having a pole at $s = 1/2$, which since it is not offset by a pole of the numerator implies that $G(s) \to 0$ as $s \to 1/2$.
- Now, we would simply like to take $s \to 1/2$ in the product expansion $G(s) = \prod_{p \nmid m, \chi(p)=1} [1 + 2\sum_{k=1}^{\infty} p^{-ks}]$, which would give an immediate contradiction since clearly $G(s) \geq 1$ for positive s in this product (since all the coefficients are nonnegative), but we only know this expansion is valid when $\text{Re}(s) > 1$.
- So instead, we expand $G(s)$ as a power series centered at $s = 2$, as $G(s) = \sum_{j=0}^{\infty} b_j (s-2)^j$, and observe that the coefficients b_n are given by $\dfrac{1}{j!} G^{(j)}(2) = \dfrac{1}{j!} \sum_{n=1}^{\infty} \dfrac{d^j}{ds^j} [a_n n^{-s}] = \dfrac{1}{j!} \sum_{n=1}^{\infty} a_n (-\ln n)^j n^{-s} = (-1)^j c_j$ for some $c_j \geq 0$.
- Hence we obtain $G(s) = \sum_{j=0}^{\infty} (-1)^j c_j (s-2)^j = \sum_{j=0}^{\infty} c_j (2-s)^j$, which for real s with $1/2 < s < 2$ is a sum of nonnegative terms, so in particular $G(s) \geq c_0 = b_0 = 1$.
- This means $G(s) \geq 1$ for real $1/2 < s < 2$, which contradicts the fact that $G(s) \to 0$ as $s \to 1/2$. This is impossible so in fact $L(1, \chi) \neq 0$ as claimed.

- Now that we know $L(1, \chi)$ vanishes for nontrivial characters χ, we can prove Dirichlet's theorem:

- <u>Theorem</u> (Dirichlet's Theorem on Primes in Arithmetic Progressions): Suppose m is a positive integer and a is relatively prime to m. Then the Dirichlet density of the set of primes congruent to $a \pmod m$ exists and is $1/\varphi(m)$. In particular, there are infinitely many such primes.
  - We have already obtained all of the necessary ingredients, so the proof is mostly a matter of putting them all together.
  - <u>Proof</u>: Recall the power series $-\log(1-x) = \sum_{k=1}^{\infty} x^k/k$, valid for $|x| < 1$.
  - Then for any Dirichlet character χ, we have $\log L(s, \chi) = \sum_p -\log(1 - \chi(p)p^{-s}) = \sum_p \left[\sum_{k=1}^{\infty} \dfrac{\chi(p)^k}{k} |p|^{-ks} \right] = \sum_p \dfrac{\chi(p)}{|p|^s} + \sum_p \sum_{k=2}^{\infty} \dfrac{\chi(p)^k}{k} p^{-ks}$. The absolute value of the second term is bounded by $\sum_p \sum_{k=2}^{\infty} \dfrac{1}{k} p^{-ks} \leq \sum_{n=2}^{\infty} n^{-2s}$, which is finite as $s \to 1+$.
  - Therefore, as $s \to 1+$, we have $\log L(s, \chi) = \sum_p \chi(p) p^{-s} + O(1)$. In particular, we see that $\sum_p p^{-s} = \log(s-1) + O(1)$ as $s \to 1+$, since $L(s, \chi_{\text{triv}})$ has a simple pole at $s = 1$.
  - Now, by Fourier inversion (as we previously worked out) we have $\sum_{p \equiv a \pmod m} p^{-s} = \sum_p \delta_a(p) p^{-s} = \dfrac{1}{\varphi(m)} \sum_{\chi \in \hat{G}} \left[\overline{\chi(a)} \sum_p \chi(p) p^{-s} \right]$.
  - So, the quotient for the Dirichlet density is $\dfrac{\sum_{p \equiv a \pmod m} p^{-s}}{\sum_p p^{-s}} = \dfrac{\frac{1}{\varphi(m)} \sum_{\chi \in \hat{G}} \left[\overline{\chi(a)} \sum_p \chi(p) p^{-s} \right]}{\sum_p p^{-s}} = \dfrac{1}{\varphi(m)} \left[\dfrac{\sum_{p \nmid m} p^{-s}}{\sum_p p^{-s}} + \dfrac{\sum_{\chi \neq \chi_{\text{triv}}} \overline{\chi(a)} \sum_p \chi(p) p^{-s}}{\sum_p p^{-s}} \right] = \dfrac{1}{\varphi(m)} \left[1 - \dfrac{\sum_{p \mid m} p^{-s}}{\log(s-1) + O(1)} + \dfrac{\sum_{\chi \neq \chi_{\text{triv}}} \log L(s, \chi) + O(1)}{\log(s-1) + O(1)} \right]$.
  - Now, taking the limit as $s \to 1+$ makes the second term go to zero (since the numerator is finite) and the third term go to zero (since $L(1, \chi) \neq 0$ for $\chi \neq \chi_{\text{triv}}$), and so the value of the limit is just $1/\varphi(m)$, as claimed.

- Now that we have established Dirichlet's theorem, we make some brief remarks about what improvements are required to show that the natural density of the primes congruent to a modulo m is equal to $1/\varphi(m)$, not just the Dirichlet density.
  - To do this requires showing that $L(s, \chi)$ is zero-free on a larger region: specifically, we need it to be zero-free for $\text{Re}(s) = 1$, rather than just at $s = 1$.
  - It is hypothesized in fact that the L-function $L(s, \chi)$ is actually zero-free on a much larger region: precisely, that all of its zeroes lie on the line $\text{Re}(s) = 1/2$: this is the generalized Riemann hypothesis.

10.2 The Dedekind Zeta Function and the Analytic Class Number Formula

- We now generalize the notion of the Riemann zeta function to give the analogous zeta function for a quadratic integer ring $\mathcal{O}_{\sqrt{D}}$.
  - We also use the resulting new kind of zeta function, called the Dedekind zeta function, to give an analytic formula for the class number.

10.2.1 The Dedekind Zeta Function

- Before beginning, we will highlight some of the salient features of the analogy between $\mathbb{Z}$ and $\mathcal{O}_{\sqrt{D}}$.
  - First, rather than having unique factorization of elements into products of prime elements, we have unique factorizations of ideals into products of prime ideals.
  - Therefore, the proper analogue of "positive integers n" is "nonzero ideals I", and the analogue of "prime numbers p" is "prime ideals P".
  - Furthermore, although we did not highlight it in $\mathbb{Z}$, we also have a natural way of measuring the size of an integer, namely via its absolute value. For ideals, we measure their size using the norm.
- Now we can define the Dedekind zeta function of a quadratic integer ring:
- Definition: If $K = \mathbb{Q}(\sqrt{D})$ and $R = \mathcal{O}_{\sqrt{D}}$ is the associated quadratic integer ring, the Dedekind zeta function for K is the Dirichlet series $\zeta_K(s) = \sum_{I \subseteq R} \dfrac{1}{N(I)^s}$, where the sum is over all nonzero ideals in $\mathcal{O}_{\sqrt{D}}$.
  - Remark: This definition works equally well for other finite-degree field extensions $K/\mathbb{Q}$, although we will focus only on the case of quadratic fields here.
  - By the uniqueness of prime factorizations of ideals in $\mathcal{O}_{\sqrt{D}}$ we obtain an analogous Euler product expansion: $\zeta_K(s) = \prod_{P \text{ prime}} (1 - N(P)^{-s})^{-1}$, where the product is over all prime ideals P of $\mathcal{O}_{\sqrt{D}}$.
  - Like with the Riemann zeta function, the series for $\zeta_K(s)$, and the Euler product, both converge absolutely for $\text{Re}(s) > 1$. However, this is a bit more difficult to prove than in the situation for the Riemann zeta function, because there can now be several ideals of the same norm.
  - Explicitly, as we have shown, there are at most 2 prime ideals of a given norm (if P is ramified or inert, then it is the only ideal of its norm, while if P is split then its conjugate is the only other ideal of the same norm).
  - By exploiting the Euler product and this fairly trivial estimate on the number of ideals of a given norm we can establish the desired convergence results:
- Proposition (Convergence of $\zeta_K(s)$): If $K = \mathbb{Q}(\sqrt{D})$ and $R = \mathcal{O}_{\sqrt{D}}$ then the Dedekind zeta function $\zeta_K(s) = \sum_{I \subseteq R} N(I)^{-s}$ converges absolutely for $\text{Re}(s) > 1$, and also has an Euler product expansion $\zeta_K(s) = \prod_{P \text{ prime}} (1 - N(P)^{-s})^{-1}$ which converges absolutely and is nonzero for $\text{Re}(s) > 1$.
  - Proof: Observe first that $\sum_{P \text{ prime}} N(P)^{-s}$ converges absolutely for $\text{Re}(s) > 1$, since $\sum_{P \text{ prime}} |N(P)^{-s}| = \sum_{P \text{ prime}} N(P)^{-\text{Re}(s)} \le \sum_{p \text{ prime}} 2p^{-\text{Re}(s)} \le \sum_{n=1}^{\infty} 2n^{-\text{Re}(s)} < \infty$, where the middle inequality follows since there are at most 2 prime ideals P lying above an integer prime p and their norms are at least p.
  - Now for a fixed X, observe that all terms in the partial sum $\sum_{I \subseteq R,\, N(I) \le X} N(I)^{-s}$ are included in the sum $\prod_{P \text{ prime},\, N(P) \le T} (1 + N(P)^{-s} + N(P)^{-2s} + \cdots)$ by unique factorization of ideals, since an ideal of norm $\le X$ is a product of prime powers of norm $\le X$.
  - Thus taking absolute values yields $\sum_{I \subseteq R,\, N(I) \le X} |N(I)^{-s}| \le \prod_{P \text{ prime},\, N(P) \le T} (1 + N(P)^{-\text{Re}(s)} + N(P)^{-2\text{Re}(s)} + \cdots) \le \prod_{P \text{ prime}} (1 + 3N(P)^{-\text{Re}(s)}) \le \exp[\sum_{P \text{ prime}} 3N(P)^{-\text{Re}(s)}] < \infty$, where the last step follows from the standard inequality $\prod(1 + x_i) \le e^{\sum x_i}$ for positive x_i.
  - Therefore, the partial sums of the Dedekind zeta function converge absolutely for $\text{Re}(s) > 1$. The Euler product also converges absolutely by the triangle inequality and the inequality $\prod(1 + x_i) \le e^{\sum x_i}$, and is nonzero since none of the terms can be zero. Finally, the sum and Euler product agree by comparing tails as with the Riemann zeta function.

- We will also remark that there is a simple relationship between the Dedekind zeta function and the Dirichlet L-series we have previously examined:

- Proposition (Dirichlet L-Series and $\zeta_K(s)$): For $K = \mathbb{Q}(\sqrt{D})$ of discriminant Δ, let χ be the Jacobi symbol modulo Δ (considered as a Dirichlet character). Then $\zeta_K(s) = \zeta(s) \cdot L(s, \chi)$ for $\text{Re}(s) > 1$.

  - Proof: We compare Euler products for both sides.
  - For $\zeta(s) = \prod_{p \text{ prime}} (1 - p^{-s})^{-1}$ the Euler factor at p is $(1 - p^{-s})^{-1}$.
  - For $L(s, \chi) = \prod_{p \text{ prime}} (1 - \chi(p)p^{-s})^{-1}$ the Euler factor at p is $(1 - \chi(p)p^{-s})^{-1}$.
  - Since $\chi(p) = \begin{cases} 1 & \text{when } p \text{ is a quadratic residue} \\ 0 & \text{when } p|\Delta \\ -1 & \text{when } p \text{ is a quadratic nonresidue} \end{cases} = \begin{cases} 1 & \text{when } p \text{ splits} \\ 0 & \text{when } p \text{ ramifies} \\ -1 & \text{when } p \text{ is inert} \end{cases}$ by Dedekind-Kummer, we see that the product of the Euler factors is $(1-p^{-s})^{-2}$ when p splits, it is $(1-p^{-s})^{-1}$ when p ramifies, and it is $(1 - p^{-s})^{-1}(1 + p^{-s})^{-1} = (1 - p^{-2s})^{-1}$ when p is inert.
  - For $\zeta_K(s) = \prod_{P \text{ prime}} (1 - N(P)^{-s})^{-1} = \prod_{p \text{ prime}} \prod_{P|p} (1 - N(P)^{-s})^{-1}$, the Euler factor at p is $\prod_{P|p}(1 - N(P)^{-s})^{-1}$.
  - This factor is $(1 - p^{-s})^{-2}$ when p splits (two factors with $N(P) = p$), it is $(1 - p^{-s})^{-1}$ when p ramifies (one factor with $N(P) = p$), and it is $(1 - p^{-2s})^{-1}$ when p is inert (one factor with $N(P) = p^2$).
  - In all three cases we see that the Euler factors are the same for the two expressions, so since the products all converge absolutely, we may rearrange them to conclude they are equal.

10.2.2 The Analytic Class Number Formula

- Our goal now is to show that $\zeta_K(s)$ has a pole at $s = 1$ and to calculate its residue there.

  - The first portion of the argument is to exploit the class group structure to decompose the sum $\zeta_K(s) = \sum_{I \subseteq R} \frac{1}{N(I)^s}$ into separate pieces for each ideal class $\mathcal{C}$ in the class group G.

- Proposition (Decomposition of $\zeta_K(s)$ Over Ideal Classes): Let $K = \mathbb{Q}(\sqrt{D})$ and $R = \mathcal{O}_{\sqrt{D}}$, with Dedekind zeta function $\zeta_K(s) = \sum_{I \subseteq R} N(I)^{-s}$ and ideal class group G. Then we have the following:

  1. We have $\zeta_K(s) = \sum_{\mathcal{C} \in G} \sum_{I \in \mathcal{C}} N(I)^{-s}$ where the outer sum is over ideal classes $\mathcal{C}$ and the inner sum is over all ideals I in the class $\mathcal{C}$.
     - Proof: Since $\zeta_K(s)$ is the sum over all nonzero ideals I in R and each ideal I lies in a unique ideal class $\mathcal{C}$, grouping the ideals together by ideal class yields the claimed formula.
  2. For a fixed ideal class $\mathcal{C}$ choose any ideal $J_C \in \mathcal{C}^{-1}$ and let S be any set of nonzero elements $\alpha \in J_C$ containing exactly one element from each associate class (i.e., representatives for $J_C \backslash \{0\}$ up to associates). Then $\sum_{I \in \mathcal{C}} N(I)^{-s} = N(J_c)^s \sum_{\alpha \in S} |N(\alpha)|^{-s}$.
     - Proof: By definition, for any $I \in \mathcal{C}$ the ideal product $J_C I$ represents the trivial class, hence is some principal ideal (α). Furthermore, we see that the correspondence $I \mapsto (\alpha)$ is a bijection between ideals $J \in \mathcal{C}$ and nonzero principal ideals (α) is divisible by J_C: equivalently, such that $\alpha \in J_C$.
     - Taking norms yields $N(J_C I) = N((\alpha)) = |N(\alpha)|$ whence $N(I) = N(J_C)^{-1} |N(\alpha)|$.
     - Thus $\sum_{I \in \mathcal{C}} N(I)^{-s} = N(J_c)^s \sum_{J_C|(\alpha)} |N(\alpha)|^{-s}$, where the sum is over all principal ideals (α) containing J_C. By selecting a generator $\alpha \in J_C$ of (α), which is unique up to associates, we see that $\sum_{J_C|(\alpha)} |N(\alpha)|^{-s} = \sum_{\alpha \in S} |N(\alpha)|^{-s}$ as claimed.
  3. Suppose $D < 0$ and let $\pi : \mathcal{O}_{\sqrt{D}} \to \mathbb{R}^2$ be the Minkowski embedding $\pi(\alpha) = (\text{Re}\alpha, \text{Im}\alpha)$. If $\mathcal{O}_{\sqrt{D}}$ has $\omega(K)$ units, then each nonzero element $\pi(\alpha) \in \pi(\mathcal{O}_{\sqrt{D}})$ has a unique associate in the fundamental region X given by the set of points (r, θ) in polar coordinates with $r > 0$ and $0 \le \theta < 2\pi/\omega(K)$.
     - Note that the group of roots of unity in K is finite: they are the 6 sixth roots of unity in $O_{\sqrt{-3}}$, the 4 fourth roots of unity in $\mathcal{O}_{\sqrt{-1}}$, and the 2 square roots of unity ± 1 in all other cases.

- ◦ Proof: Since the only units in K are the $\omega(K)$ roots of unity, which act by rotation in increments of $2\pi/\omega(K)$ radians, the fundamental domain is simply the quotient of $\mathbb{C}^\times$ by this group, which is precisely the claimed region X.

4. Suppose $D > 0$ and let $\pi : \mathcal{O}_{\sqrt{D}} \to \mathbb{R}^2$ be the Minkowski embedding $\pi(\alpha) = (\alpha, \overline{\alpha})$. If $\mathcal{O}_{\sqrt{D}}$ has fundamental unit $\epsilon > 1$, then each nonzero element $\pi(\alpha) \in \pi(\mathcal{O}_{\sqrt{D}})$ has a unique associate in the fundamental region X given by the set of points (x, y) in rectangular coordinates with $x > 0$ and with $\epsilon^{-2} |x| < |y| \le |x|$.
   - ◦ Proof: Since $\epsilon > 1$ is the fundamental unit of $\mathcal{O}_{\sqrt{D}}$, the group of units is $\pm\epsilon^n$ for integers n, per our study of Pell's equation. Then $\pi(\epsilon) = (\epsilon, \overline{\epsilon}) = (\epsilon, \pm 1/\epsilon)$ and $\pi(-1) = (-1, -1)$.
   - ◦ Therefore, given any nonzero element $(x, y) \in \pi(J_C)$, we may rescale its first coordinate to be positive, which fully accounts for the unit factor ± 1. Then scaling α by ϵ^n rescales $|y| / |x|$ by ϵ^{-2n}, hence we may rescale by a unique power of ϵ to ensure that $1 \le |y| / |x| < \epsilon^{-2}$, which fully accounts for the unit factor ϵ^n. Thus each element has a unique associate in the region X described above, as claimed.

- We can see that both of the fundamental regions X over whose lattice points we seek to sum in (3) and (4) of the proposition above are cones: sets such that $x \in X$ implies $\lambda x \in X$ for any $\lambda > 0$.
  - ◦ Explicitly, suppose that f is a homogeneous function on X of degree 2: a function such that $f(\lambda x) = \lambda^2 f(x)$ for all $x \in X$ and all $\lambda > 0$. (Note that the norm function satisfies this condition in both the real and imaginary cases.)
  - ◦ Also suppose Λ is a lattice in $\mathbb{R}^2$ whose fundamental parallelogram has area Δ.
  - ◦ Then the zeta functions in the proposition above are both of the form $\zeta_{\Lambda,F}(s) = \sum_{x \in \Lambda \cap X} \dfrac{1}{f(x)^s}$. We now seek to give a general result for calculating the residue at $s = 1$ for zeta functions of this type.
  - ◦ Roughly speaking, the idea is to use the homogeneity of f to rescale the sum and then compare it to the Riemann zeta function.

- Theorem (Zeta Functions on Cones): Suppose X is a cone in $\mathbb{R}^2$, f is a homogeneous function on X of degree 2, $T = \{(x, y) \in X : f(x, y) \le 1\}$ is bounded and has area A, and Λ is a lattice in $\mathbb{R}^2$ whose fundamental parallelogram P has area Δ. Then the following hold:

1. We have $\lim_{s\to\infty} \dfrac{\#\{x \in \Lambda : f(x) \le s^2\}}{s^2} = \dfrac{A}{\Delta}$.
   - ◦ Proof: For any set S and any positive real scalar $s > 0$ let sS be the set obtained by scaling all elements of S by s.
   - ◦ Then $s\Lambda$ is a lattice with fundamental parallelogram sP of area $s^2\Delta$, sX is merely X itself, and $sT = \{sx : x \in X,\ f(x) \le 1\} = \{x \in X : f(x) \le s^2\}$ by the homogeneity of f and $sX = X$.
   - ◦ Then $\Lambda \cap sT = \{(x, y) \in \Lambda : f(x) \le s^2\}$, so we wish to calculate $\lim_{s\to\infty} \#(\Lambda \cap sT)/s^2$.
   - ◦ Consider $\Lambda \cap sR$ and imagine filling in a copy of P with center at each point of $\Lambda \cap sT$: the area of these tiled copies of P differs from the area of sT by at most the perimeter of sT times the area of P.
   - ◦ Since the area of all these copies of P is $\#(\Lambda \cap sT) \cdot \text{Area}(P)$, the area of sT is s^2 times the area of T while the perimeter of sT is s times the perimeter of T, we see that $\left|\#(\Lambda \cap sT) \cdot \text{Area}(P) - s^2\text{Area}(T)\right| \le s \cdot \text{Perimeter}(T) \cdot \text{Area}(P)$, so dividing yields $\left|\dfrac{\#(\Lambda \cap sT)}{s^2} - \dfrac{\text{Area}(T)}{\text{Area}(P)}\right| \le \dfrac{1}{s}\text{Perimeter}(T)$ which tends to 0 as $s \to \infty$.
   - ◦ Hence we see that $\lim_{s\to\infty} \dfrac{\#(\Lambda \cap sT)}{s^2} = \dfrac{\text{Area}(T)}{\text{Area}(P)} = \dfrac{A}{\Delta}$, as desired.
2. If we label the countably many points $\{x_1, x_2, x_3, \dots\}$ of $\Lambda \cap X$ in increasing order of their associated value of f, so that $0 < f(x_1) \le f(x_2) \le f(x_3) \le \cdots$, then we have $\lim_{n\to\infty} \dfrac{n}{f(x_n)} = \dfrac{A}{\Delta}$.
   - ◦ Proof: Note as in (1) that $\sqrt{f(x_n)}T = \{x \in X : f(x) \le f(x_n)\}$.

- Then each of $x_1, x_2, \dots, x_n$ lie in $\sqrt{f(x_n)}T$, but x_n does not lie in $(\sqrt{f(x_n)} - \epsilon)T$ for any $\epsilon > 0$. Therefore, we have $\#[(\sqrt{f(x_n)} - \epsilon)T] \le n \le \#[\sqrt{f(x_n)}T]$.
- Thus $\dfrac{\#[(\sqrt{f(x_n)} - \epsilon)T]}{(\sqrt{f(x_n)} - \epsilon)^2} \cdot \left[\dfrac{\sqrt{f(x_n)} - \epsilon}{\sqrt{f(x_n)}}\right]^2 \le \dfrac{n}{f(x_n)} \le \dfrac{\#[\sqrt{f(x_n)}T]}{f(x_n)}$. As $k \to \infty$ we also have $f(x_n) \to \infty$ since there are only finitely many x_n with $f(n_k) \le B$ for any B.
- Thus by (1) the lower bound has limit $\dfrac{A}{\Delta} \cdot 1^2$ while the upper bound also has limit $\dfrac{A}{\Delta}$, and so we conclude that $\lim_{n\to\infty} \dfrac{n}{f(x_n)} = \dfrac{A}{\Delta}$, as desired.

3. The series $\sum_{x \in \Lambda \cap X} f(x)^{-s}$ converges absolutely for $\mathrm{Re}(s) > 1$ with a simple pole at $s = 1$ of residue A/Δ.
   - The idea here is to compare $\zeta_{\Lambda,f}$ to the Riemann zeta function and use our earlier calculation of its pole and residue at $s = 1$.
   - Proof: Let $\epsilon > 0$. By (2) we know that there exists N such that for all $n \ge N$ we have $\left|\dfrac{n}{f(x_n)} - \dfrac{A}{\Delta}\right| < \epsilon$, meaning that $\left[\dfrac{A}{\Delta} - \epsilon\right]\dfrac{1}{n} < \dfrac{1}{f(x_n)} < \left[\dfrac{A}{\Delta} + \epsilon\right]\dfrac{1}{n}$.
   - Raising to the sth power and taking absolute values shows $\sum_{n=N}^{\infty} \dfrac{1}{|f(x_n)|^s} \le \left[\dfrac{A}{\Delta} + \epsilon\right]^{\mathrm{Re}(s)} \zeta(\mathrm{Re}(s))$ which converges absolutely for $\mathrm{Re}(s) > 1$.
   - For the pole and residue, multiplying by $s - 1$ for real $s > 1$ then yields $\left[\dfrac{A}{\Delta} - \epsilon\right]^s (s-1)\sum_{n=N}^{\infty} \dfrac{1}{n^s} < (s-1)\sum_{n=N}^{\infty} \dfrac{1}{f(x_n)^s} < \left[\dfrac{A}{\Delta} + \epsilon\right]^s (s-1)\sum_{n=N}^{\infty} \dfrac{1}{n^s}$.
   - Now we make the trivial observations that $\lim_{s\to 1}(s-1)\sum_{n=1}^{N-1} \dfrac{1}{n^s} = 0 = \lim_{s\to 1+}(s-1)\sum_{n=1}^{N-1} \dfrac{1}{f(x_n)^s}$ since both sums are finite.
   - Thus $\left[\dfrac{A}{\Delta} - \epsilon\right]^s (s-1)\sum_{n=1}^{\infty} \dfrac{1}{n^s} - \delta_s < (s-1)\sum_{n=1}^{\infty} \dfrac{1}{f(x_n)^s} < \left[\dfrac{A}{\Delta} + \epsilon\right]^s (s-1)\sum_{n=1}^{\infty} \dfrac{1}{n^s} + \delta_s$ for some $\delta_s \to 0$ as $s \to 1+$.
   - Hence when we take $s \to 1+$, the lower bound approaches $\dfrac{A}{\Delta} - \epsilon$ while the upper bound approaches $\dfrac{A}{\Delta} + \epsilon$. So $\dfrac{A}{\Delta} - \epsilon \le \liminf_{s\to 1+}(s-1)\sum_{n=1}^{\infty} \dfrac{1}{f(x_n)^s} \le \limsup_{s\to 1+}(s-1)\sum_{n=1}^{\infty} \dfrac{1}{f(x_n)^s} \le \dfrac{A}{\Delta} + \epsilon$.
   - Finally, taking $\epsilon \to 0$ shows that in fact the liminf and limsup must equal $\dfrac{A}{\Delta}$, and hence the limit exists and we have $\lim_{s\to 1+}(s-1)\sum_{n=1}^{\infty} \dfrac{1}{f(x_n)^s} = \dfrac{A}{\Delta}$, as desired.

- Now we assemble these various pieces to compute the residue of the Dedekind zeta function at $s = 1$, which gives an analytic expression for the class number:
- Theorem (Analytic Class Number Formula): Let $K = \mathbb{Q}(\sqrt{D})$ and $R = \mathcal{O}_{\sqrt{D}}$, with Dedekind zeta function $\zeta_K(s) = \sum_{I \subseteq R} N(I)^{-s}$, discriminant Δ, and ideal class group G and class number $h(K)$.

  1. If $D < 0$ then $\zeta_K(s)$ has a simple pole at $s = 1$ with residue equal to $\dfrac{2\pi h(K)}{\omega(K)\sqrt{|\Delta|}}$ where $\omega(K)$ is the number of roots of unity in K.
     - Proof: By our proposition on decomposing the Dirichlet series, we have $\zeta_K(s) = \sum_{\mathcal{C} \in G} \sum_{I \in \mathcal{C}} N(I)^{-s}$ where $\sum_{I \in \mathcal{C}} N(I)^{-s} = N(J_c)^s \sum_{\alpha \in \pi(J_c) \cap X} |N(\alpha)|^{-s}$ and X is the fundamental region given by the set of points (r, θ) in polar coordinates such that $r > 0$ and $0 \le \theta < 2\pi/\omega(K)$.
     - Now, since X is a cone, and $\pi(J_c)$ is a lattice Λ whose fundamental parallelogram has area $\sqrt{|\Delta|}N(J_C)$, applying our theorem about zeta functions on cones yields that $\sum_{\alpha \in \pi(J_c) \cap X} |N(\alpha)|^{-s}$ has a simple pole at $s = 1$ with residue A/Δ, where A is the area of the region T with $N(x + iy) \le 1$ inside X.

- Since the Minkowski map has $\pi(\alpha) = (\text{Re}\alpha, \text{Im}\alpha)$, the norm map is $N(\alpha) = \alpha\overline{\alpha} = [\text{Re}\alpha]^2 + [\text{Im}\alpha]^2$ which on coordinates is $N(x, y) = x^2 + y^2$.
- Therefore, the set $T = \{(x, y) \in X : N(x, y) \le 1\}$ is just a circular sector of radius 1 and angle $\dfrac{2\pi}{\omega(K)}$, hence has area $A = \dfrac{2\pi}{\omega(K)}$. Therefore the pole residue at $s = 1$ of $\sum_{\alpha \in \pi(J_c) \cap X} |N(\alpha)|^{-s}$ is $\dfrac{A}{\Delta} = \dfrac{2\pi}{\omega(K)\sqrt{|\Delta|}N(J_C)}$, hence the pole residue of $N(J_c)^s \sum_{\alpha \in \pi(J_c) \cap X} |N(\alpha)|^{-s}$ is $\dfrac{2\pi}{\omega(K)\sqrt{|\Delta|}}$.
- Finally, summing over all terms shows that $\sum_{\mathcal{C} \in G} \sum_{I \in \mathcal{C}} N(I)^{-s}$ has a simple pole at $s = 1$ and its residue is the sum of the residues for each term. Since there are $h(K)$ terms each with residue $\dfrac{2\pi}{\omega(K)\sqrt{|\Delta|}}$, the sum is $\dfrac{2\pi h(K)}{\omega(K)\sqrt{|\Delta|}}$ as claimed.

2. If $D > 0$ then $\zeta_K(s)$ has a simple pole at $s = 1$ with residue equal to $\dfrac{2h(K)\ln\epsilon}{\sqrt{\Delta}}$ where $\epsilon > 1$ is the fundamental unit of K.
   - Proof: The argument is the same as in (1), except now X is the fundamental region given by the set of points (x, y) in rectangular coordinates with $x > 0$ and with $\epsilon^{-2}|x| < |y| \le |x|$.
   - Since the Minkowski map has $\pi(\alpha) = (\alpha, \overline{\alpha})$ and the norm map is $N(\alpha) = \alpha\overline{\alpha}$, the norm map on coordinates is $N(x, y) = xy$.
   - Therefore, the set $T = \{(x, y) \in X : N(x, y) \le 1\}$ is the region defined by $\epsilon^{-2}|x| < |y| \le |x|$ and $xy \le 1$.
   - This region consists of two pieces symmetric across the x-axis, with the piece above the axis bounded by the lines $y = x$, $y = \epsilon^{-2}x$ and the curve $N(xy) = xy = 1$.
   - The area of each piece is therefore $\int_0^1 (x - \epsilon^{-2}x)\,dx + \int_1^\epsilon (x^{-1} - \epsilon^{-2}x)\,dx = \ln\epsilon$, so the area of the fundamental region is $2\ln\epsilon$.
   - Just as in (1), since the area of the fundamental parallelogram is $\sqrt{\Delta}N(J_C)$, so the pole residue of the pole residue of $N(J_c)^s \sum_{\alpha \in \pi(J_c) \cap X} \dfrac{1}{|N(\alpha)|^s}$ is $\dfrac{2\ln\epsilon}{\sqrt{\Delta}}$. Each of the $h(K)$ terms has the same residue, so the residue for the sum is $\dfrac{2h(K)\ln\epsilon}{\sqrt{\Delta}}$ as claimed.

- As an immediate corollary, because we have the decomposition $\zeta_K(s) = \zeta(s)L(s, \chi)$ for the Jacobi symbol χ modulo D, equating the residues at $s = 1$ yields the following:
- Corollary (Class Numbers and L-Series): Let $K = \mathbb{Q}(\sqrt{D})$ have discriminant Δ and let χ be the Jacobi symbol modulo Δ. Then for $D < 0$ we have $L(1, \chi) = \dfrac{2\pi h(K)}{\omega(K)\sqrt{|\Delta|}}$ and for $D > 0$ we have $L(1, \chi) = \dfrac{2h(K)\ln\epsilon}{\sqrt{\Delta}}$.
  - We will remark here that this calculation provides an alternate proof that $L(1, \chi) \ne 0$ for the quadratic character χ, since all of the terms in the expression are obviously nonzero (in fact, this was essentially Dirichlet's original argument for the nonvanishing of these L-series).
  - Proof: Since $\zeta_K(s)$ and $\zeta(s)L(s, \chi)$ both have simple poles at $s = 1$, the limit $L(1, \chi) = \lim_{s \to 1^+} L(s, \chi) = \lim_{s \to 1^+} \dfrac{\zeta_K(s)}{\zeta(s)} = \lim_{s \to 1^+} \dfrac{(s-1)\zeta_K(s)}{(s-1)\zeta(s)}$ is just the ratio of their residues. Since the zeta function has residue 1, we obtain the claimed formulas immediately.
- Example: Verify the analytic class number formula for $D = -1$.
  - For the quadratic character modulo 4, we have $L(1, \chi) = \sum_{n=0}^{\infty} \dfrac{(-1)^n}{2n+1} = 1 - \dfrac{1}{3} + \dfrac{1}{5} - \dfrac{1}{7} - \cdots = \dfrac{\pi}{4}$ since this is Leibniz's formula.
  - Here we also have $\omega = 4$, $\Delta = -4$, and $h = 1$, so $\dfrac{2\pi h}{\omega \cdot \sqrt{|\Delta|}} = \dfrac{\pi}{4}$, which agrees.

Well, you're at the end of my handout. Hope it was helpful.