Dirichlet's Class Number Formula

Luke Giberson

4/26/13

These lecture notes are a condensed version of Tom Weston's exposition on this topic. The goal is to develop an analytic formula for the class number of an imaginary quadratic field.

Algebraic Motivation

Definition. Fix a squarefree positive integer n. The *ring of algebraic integers*, $\mathcal{O}_{-n} \leq \mathbb{Q}(\sqrt{-n})$ is defined as

$$\mathcal{O}_{-n} = \begin{cases} a + b\sqrt{-n} & a, b \in \mathbb{Z} & \text{if } n \equiv 1, 2 \mod 4 \\ a + b\sqrt{-n} & 2a, 2b \in \mathbb{Z}, 2a \equiv 2b \mod 2 & \text{if } n \equiv 3 \mod 4 \end{cases}$$

or equivalently

$$\mathcal{O}_{-n} = \{a + b\omega \colon a, b \in \mathbb{Z}\},$$

where

$$\omega = \begin{cases} \sqrt{-n} & \text{if } n \equiv 1, 2 \mod 4 \\ \dfrac{1 + \sqrt{-n}}{2} & \text{if } n \equiv 3 \mod 4. \end{cases}$$

This will be our primary object of study. Though the conditions in this definition appear arbitrary, one can check that elements in $\mathcal{O}_{-n}$ are precisely those elements in $\mathbb{Q}(\sqrt{-n})$ whose characteristic polynomial has integer coefficients.

Definition. We define the *norm*, a function from the elements of any quadratic field to the integers as $\mathrm{N}(\alpha) = \alpha\bar{\alpha}$. Working in an imaginary quadratic field, the norm is never negative.

The norm is particularly useful in transferring multiplicative questions from $\mathcal{O}_{-n}$ to the integers. Here are a few properties that are immediate from the definitions.

- If α divides β in $\mathcal{O}_{-n}$, then $\mathrm{N}(\alpha)$ divides $\mathrm{N}(\beta)$ in $\mathbb{Z}$.
- An element $\alpha \in \mathcal{O}_{-n}$ is a unit if and only if $\mathrm{N}(\alpha) = 1$.
- If $\mathrm{N}(\alpha)$ is prime, then α is irreducible in $\mathcal{O}_{-n}$.

Using the second bullet point above, we can calculate the units of $\mathcal{O}_{-n}$. In our case, we're just interested in the number of units, which we'll denote w_{-n}. It's not hard to show that

$$w_{-n} = \begin{cases} 4 & \text{if } n = 1 \\ 6 & \text{if } n = 3 \\ 2 & \text{otherwise.} \end{cases}$$

With a notion of irreducibility, we can now discuss factorization into irreducibles in this integer ring. The problem, however, is that unique factorization in $\mathcal{O}_{-n}$ is not guaranteed and is in fact dependent on the underlying imaginary quadratic field (the choice of n). For example, the classic counterexample to unique factorization in $\mathbb{Z}[\sqrt{-5}]$ is

$$6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5}).$$

To determine why unique factorization fails in some rings and not in others, we need to examine the ideal structure of $\mathcal{O}_{-n}$. It turns out that the ideals of this ring always factor uniquely, and the class number is a measure of the extent to which unique factorization fails in $\mathcal{O}_{-n}$.

Definition. An *ideal* of $\mathcal{O}_{-n}$ is a set of $\mathcal{O}_{-n}$-linear combinations. As such, we can talk about generators of an ideal, say $\alpha_1, \ldots, \alpha_r$ and write the ideal as

$$(\alpha_1, \ldots, \alpha_r) = \{\alpha_1 x_1 + \ldots + \alpha_r x_r : x_i \in \mathcal{O}_{-n}\}.$$

A *principal ideal* can be generated by a single element, say

$$(\alpha) = \{\alpha x : x \in \mathcal{O}_{-n}\}.$$

We'd like to develop a theory of factorization of these ideals. Multiplication of ideals can be defined in terms of sets of generators:

$$(\alpha_1, \ldots, \alpha_r)(\beta_1, \ldots, \beta_s) = (\{\alpha_i \beta_j\})$$

for $i = 1 \ldots r$ and $j = 1 \ldots s$. For ideals I, J, notice that $I \cdot J \subset I$, so the set of ideals do not form a multiplicative group.

Definition. The norm of an ideal, denoted $\mathrm{N}(I)$, is the positive integer d such that $I \cdot \bar{I} = (d)$.

The existence of such an object is not at all trivial, but in the interest of time, we'll take this as a definition. For a rigorous proof, one can argue that there is a subset of positive integers in $I \cdot \bar{I}$, whereupon choosing the smallest of these to be $\mathrm{N}(I)$ and showing that it divides any generic element $\alpha\bar{\beta}$ in $I \cdot \bar{I}$ proves the claim.

We have a few immediate properties of the ideal norm.

- $\mathrm{N}(I \cdot J) = \mathrm{N}(I)\mathrm{N}(J)$.

- For a principal ideal $I = (\alpha)$, $\mathrm{N}(I) = \mathrm{N}(\alpha)$.

The following property is the characterization of divisibility of ideals. In words, it says that "to divide an ideal is to contain that ideal as a subset". One direction is immediate from the definition of multiplication. The other direction can be proven using our new notion of the ideal norm.

Fact. *For ideals I and J, I divides J if and only if $J \subset I$.*

Definition. We will say that an ideal is *prime* if it satisfies Euclid's lemma for ideals. That is, I is prime if it satisfies the property that whenever I divides a product of ideals JK, then I divides J or I divides K.

Since every ideal I divides a principal ideal, namely, the ideal generated by its norm, if I is prime with $\mathrm{N}(I) = p_1^{\alpha_1} \dots p_m^{\alpha_m}$, then I divides $(\mathrm{N}(I)) = (p_1) \dots (p_r)$ and so I divides some principal ideal (p_k). Thus, we should look at the structure of principal ideals generated by primes.

Fact. *For primes $p \neq 2$,*

$$(p) = \begin{cases} (p) & \text{if -n is not a quadratic residue mod p} \\ (p, a + \sqrt{-n})(p, a - \sqrt{-n}) & \text{if -n is a non-zero quadratic residue mod p} \\ (p, \sqrt{-n})^2 & \text{if p divides -n,} \end{cases}$$

where each of the ideals on the right-hand side are prime.

Colloquially, we say that in the first case (p) is inert, in the second case (p) splits, and in the third case (p) is ramified. The conditions on these possibilities will later be encoded in a Dirichlet character called the Legendre symbol.

Corollary. *Ideals in $\mathcal{O}_{-n}$ factor uniquely.*

Proof. Induction on the ideal norm. □

Recall that ideals do not form a multiplicative group. To fix this strucure, consider the following equivalence relation. We'll say that two ideals I and J are similar, $I \sim J$, if there exist $\alpha, \beta \in \mathcal{O}_{-n}$ such that $(\alpha)I = (\beta)J$. Note that all principal ideals are similar because

$$(1)(\alpha) = (\alpha)\mathcal{O}_{-n}.$$

Definition. The *class group*, denoted $\mathbf{Cl}(-n)$, is the set of equivalence classes. The *class number*, h_{-n}, is the order of the class group.

We'll briefly justify that this is indeed a group. Let I be an ideal from a class $\mathcal{C}$ and let J be an ideal from a class $\mathcal{C}'$. The group operation is multiplication in that $\mathcal{C} \cdot \mathcal{C}' = \mathcal{C}_{IJ}$, where $\mathcal{C}_{IJ}$ is the ideal class that contains the product IJ. It can be shown that this operation is indepedent of the choice of I and J (and hence well-defined). The identity element, $\mathcal{C}_1$, is the class of all principal ideals.

For a class $\mathcal{C}$ with an ideal I, its inverse is the class that contains the ideal $\bar{I}$, as shown in the following computation:

$$\mathcal{C}_I \cdot \mathcal{C}_{\bar{I}} = \mathcal{C}_{I\bar{I}} = \mathcal{C}_{(N(I))} = \mathcal{C}_1.$$

For some choice of n, if $h_{-n} = 1$, this means that $\mathcal{O}_{-n}$ is a principal ideal domain and thus a unique factorization domain. If the class number is not trivial, then we might interpret this as meaning we have "more" ideals than we have elements. We should stress that we currently don't know anything about the class number at this point - we don't even know that it's finite! To understand more, we move to lattices.

Geometric Computation

Definition. A *complex lattice*, $\Lambda \subset \mathbb{C}$ is a set of integer linear combinations

$$\Lambda = \{m\alpha + n\beta : m, n \in \mathbb{Z}\}$$

with α, β required to be $\mathbb{R}$-linearly independent. We call $< \alpha, \beta >$ a *basis* of Λ.

Definition. Two lattices, Λ and Λ' are said to be *homothetic* if there is some complex number γ such that $\gamma\Lambda = \Lambda'$. As this is an equivalence relation, we denote this $\Lambda \sim \Lambda'$.

So two lattices are homothetic if we can scale and rotate the lattices onto one another. After a little work, we can classify all lattices up to homothety by establishing that each lattice comes with a j-invariant. The next result says that two lattices are homothetic if and only if they have the same j-invariant.

Fact. *Set*

$$F = \{z \in \mathbb{C} : \Im(z) > 0, -\frac{1}{2} < \Re(z) \leq \frac{1}{2}, |z| \geq 1, \Re(z) \geq 0 \text{ whenever } |z| = 1\}.$$

Every lattice is homothetic to precisely one lattice of the form $< 1, j >$ with $j \in F$. That is, $\Lambda \sim \Lambda'$ if and only if $j_\Lambda = j_{\Lambda'}$.

Definition. We say that a lattice Λ has *complex multiplication (CM) by complex* γ if it has the property that $\gamma\Lambda \subset \Lambda$. When γ is an integer, this is a trivial property.

Note that complex multiplication is preserved by homothety. Indeed, if $\Lambda' = \delta\Lambda$ and Λ has complex multiplication by γ, then

$$\gamma\Lambda' = \gamma(\delta\Lambda) = \delta(\gamma\Lambda) \subset \delta\Lambda = \Lambda'.$$

That is, Λ' also has complex multiplication by γ. In this way, it makes sense to discuss homothety classes of lattices that have complex multiplication by some γ.

Fact. *For a lattice with complex multiplication by γ, one can show that γ is of the form*

$$\gamma = \begin{cases} \sqrt{-n} & \text{if n } \equiv 1, 2 \mod 4 \\ \dfrac{1+\sqrt{-n}}{2} & \text{if } n \equiv 3 \mod 4. \end{cases}$$

With this in mind, we will now talk exclusively about complex multiplication by ω, the characterization of which was given on the first page. It is now just a matter of playing around with the boundary conditions in the fundamental domain F to obtain the following result.

Fact. *Every lattice with CM by ω is homothetic to a unique lattice of the form $\left\langle 1, \frac{a+\sqrt{-n}}{b} \right\rangle$ with*

- $a, b \in \mathbb{Z}$ *(additionally, a odd and b even if $n \equiv 3 \mod 4$)*
- $0 < b \leq 2\sqrt{n/3}$
- $-b < 2a \leq b$
- $a^2 + n \geq b^2$ *and* $a \geq 0$ *if* $a^2 + n = b^2$
- b *divides* $a^2 + n$ *(additionally, $2b$ divides $a^2 + n$ if $n \equiv 3 \mod 4$).*

The relationship between $\mathcal{O}_{-n}$ and this classification of lattices is illuminated in the following three statements.

Fact. *An ideal I of $\mathcal{O}_{-n}$ can be thought of as a lattice with CM by ω. In particular,*

$$I =< m, a + b\sqrt{-n} >$$

where $m = N(I)$ and $a + b\sqrt{-n}$ is chosen with smallest possible b.

Fact. *Ideals are similar if and only if their corresponding lattice representations given above are homothetic.*

Fact. *A lattice with CM by ω is homothetic to some ideal of $\mathcal{O}_{-n}$.*

These results give a bijection

$$\text{ideal classes in } \mathcal{O}_{-n} \longleftrightarrow \text{homothety classes of lattices with CM by } \omega.$$

In particular, since there are only a finite number of choices for a and b in the classification above (and thus a finite number of homothety classes with CM by ω), the number of ideal classes in $\mathcal{O}_{-n}$ must also be finite. That is, the class number h_{-n} is finite.

Before moving on to the analytic portion of the notes, we'll need a lemma that can be stated entirely in the lattice framework.

Lemma. *Let $\Lambda =< \alpha, \beta >$ and let P_0 be the period parallelogram from the origin with area A. Fix $t > 0$ and set C_t as the origin-centered circle of radius t. Of interest is to count the number of lattice points that fall within this circle. One can show that*

$$\left| \#\Lambda \cap C_t - \frac{\pi t^2}{A} \right| = O(t).$$

Analytic Connection

Definition. We define the *Legendre symbol* as a function of primes $p \neq 2$ by

$$\left(\frac{-n}{p}\right) = \begin{cases} 1 & \text{if -n is a non-zero quadratic residue mod p} \\ -1 & \text{if -n is not a quadratic residue mod p} \\ 0 & \text{if p divides -n.} \end{cases}$$

Excluding the case when $p = 2$ for brevity, we can extend this to the *Jacobi symbol*, a function on all integers by using the fundamental theorem of arithmetic:

$$\left(\frac{-n}{m}\right) = \left(\frac{-n}{p_1}\right)^{\alpha_1} \cdots \left(\frac{-n}{p_k}\right)^{\alpha_k}.$$

By its definition, the Jacobi symbol is multiplicative and is nonzero precisely when $(-n, m) = 1$. It can be shown to have period $4n$. Therefore, the Jacobi symbol is a Dirichlet character, so all the results from class will apply.

Definition. We define the *quadratic L-function* to be the Dirichlet series using the Jacobi character. That is,

$$L_{-n}(s) = \sum_{m=1}^{\infty} \left(\frac{-n}{m}\right) m^{-s}.$$

Fact. *The quadratic L-function has analytic continuation to $Re(s) > 0$ and has Euler product expansion*

$$L_{-n}(s) = \prod_{p} \left(1 - \left(\frac{-n}{p}\right) p^{-s}\right)^{-1}$$

valid for $Re(s) > 1$.

Proof. We can bound the partial sums by

$$A_M = \sum_{m=1}^{M} \left(\frac{-n}{m}\right) \leq 4n = O(M^0).$$

Thus, by (2.1.5), we have that $L_{-n}(s)$ can be continued to $\mathrm{Re}(s) > 0$. The Euler product expansion comes from (2.2.2). □

Definition. The *Dedekind zeta function* is defined as

$$\zeta_{-n}(s) = \sum_{m=1}^{\infty} a_m m^{-s}$$

where a_m = number of ideals in $\mathcal{O}_{-n}$ of norm m.

Note that the sequence (a_m) is multiplicative because of the bijection

$$\{\text{ideals of norm } m\} \times \{\text{ideals of norm } n\} \longleftrightarrow \{ \text{ ideals of norm } mn\},$$

which follows directly from the unique factorization of ideals in $\mathcal{O}_{-n}$.

Lemma. *We have that*

$$\sum_{j=0}^{\infty} a_{p^j} p^{-js} = \begin{cases} (1-p^{-s})^{-2} & \text{if } \left(\frac{-n}{p}\right) = 1 \\ (1-p^{-s})^{-1} & \text{if } \left(\frac{-n}{p}\right) = 0 \\ (1-p^{-2s})^{-1} & \text{if } \left(\frac{-n}{p}\right) = -1. \end{cases}$$

Proof. In the first case, (p) splits into two ideals I and J, each with norm p. So $a_{p^j} = \#\{I^i J^{j-i} : i = 0 \dots j\} = j+1$. Therefore,

$$\sum_{j=0}^{\infty} (j+1)p^{-js} = \frac{d}{dp} \sum_{j=0}^{\infty} p^{-js+1} = (1-p^{-s})^{-2}.$$

In the second case, (p) is ramified and $a_{p^j} = 1$ for all j. In the third case, (p) is inert and $a_{p^j} = 1$ if j is even and 0 otherwise. A similar type of argument using geometric series can be made for both of these situations to give the desired result. □

Corollary. *The Dedekind zeta function has Euler expansion*

$$\zeta_{-n}(s) = \prod_{p,\left(\frac{-n}{p}\right)=1} (1-p^{-s})^{-2} \cdot \prod_{p,\left(\frac{-n}{p}\right)=0} (1-p^{-s})^{-1} \cdot \prod_{p,\left(\frac{-n}{p}\right)=-1} (1-p^{-2s})^{-1}$$

valid for $Re(s) > 1$.

Proof. This follows from the previous lemma and the multiplicative nature of (a_m). □

To get an analytic continuation of the Dedekind zeta function, we'll need to bound the partial sums, $A_M = \sum_{m=1}^{M} a_m$. To make this approximation easier, we'll use the natural partition of the ideal classes, getting bounds on the number of ideals up to norm M in a general ideal class and then summing across all the classes. Fix any ideal class $\mathcal{C}$ and introduce

$$a_m(\mathcal{C}) = \text{number of ideals in } \mathcal{C} \text{ with norm } m$$
$$A_M(\mathcal{C}) = \sum_{m=1}^{M} a_m(\mathcal{C})$$

so that

$$A_M = \sum_{\mathcal{C}} A_M(\mathcal{C}) = h_{-n} A_M(\mathcal{C}).$$

The following lemma approximates the quantity $A_M(\mathcal{C})$.

Lemma. *For any ideal class $\mathcal{C}$ and for any $M \geq 1$, we have*

$$\left| A_M(\mathcal{C}) - \frac{\pi}{A_{-n} w_{-n}} M \right| = O(\sqrt{M}),$$

where

$$A_{-n} = \begin{cases} \sqrt{n} & \text{if } n \equiv 1, 2 \mod 4 \\ \sqrt{n}/2 & \text{if } n \equiv 3 \mod 4. \end{cases}$$

Proof. Let $\frac{a+\sqrt{-n}}{b}$ be the j-invariant of $\mathcal{C}^{-1}$ so that the ideal $J = (b, a+\sqrt{-n})$ is an element of $\mathcal{C}^{-1}$. The group structure of the ideal class group gives a bijection

$$\{I \in \mathcal{C} : \mathrm{N}(I) = m\} \longleftrightarrow \{I' \subset J : I' \text{ principal }, \mathrm{N}(I') = m\mathrm{N}(J)\}$$

by the mapping

$$I \mapsto I \cdot J = I'.$$

With this is mind, let $b_m(J)$ denote the number of elements in J with norm $m \cdot \mathrm{N}(J)$ and set $B_M(J) = \sum_{m=1}^{M} b_m(J)$. Then we have

$$\begin{aligned} B_M(J) &= \sum_{m=1}^{M} b_m(J) \\ &= \#\{\alpha \in J : \mathrm{N}(\alpha) \leq M \cdot \mathrm{N}(J)\} \\ &= \#\{\alpha \in J : |\alpha| \leq \sqrt{M \cdot \mathrm{N}(J)}\}. \end{aligned}$$

We might interpret this last equality to count the lattice points of $< b, a + \sqrt{-n} >$ that fall within a circle of radius $\sqrt{M \cdot \mathrm{N}(J)}$. The area of this period parallelogram is $b\sqrt{n}$, so we use lemma stated at the end of the geometry section to approximate

$$\left| B_M(J) - \frac{\pi}{b\sqrt{n}} M \cdot \mathrm{N}(J) \right| = O(\sqrt{M}).$$

From the bijection,

$$A_M(\mathcal{C}) = \frac{1}{w_{-n}} B_M(J),$$

where we divide by the number of units to avoid doublecounting associates. Applying this substitution and setting $A_{-n} = \frac{b\sqrt{n}}{\mathrm{N}(J)}$ completes the proof. □

Corollary. *Summing over all h_{-n} ideal classes gives the approximation*

$$\left| A_M - \frac{\pi h_{-n}}{A_{-n} w_{-n}} M \right| = O(\sqrt{M}).$$

Remark. The quantity A_{-n} above is precisely the area of the period parallelogram of the lattice $< 1, \omega_{-n} >$.

Fact. *The Dedekind zeta function has analytic continuation to $Re(s) > 1$. It also has a simple pole at $s = 1$ with residue $\frac{\pi h_{-n}}{A_{-n} w_{-n}}$.*

Proof. The corollary above shows that $|A_M| = O(M)$, so by (2.1.5), we get the desired convergence. For the residue result, define

$$f(s) = \sum_{m=1}^{\infty} \left(a_m - \frac{\pi h_{-n}}{A_{-n} w_{-n}} \right) m^{-s}.$$

Bounding these partial sums using the corollary above, we see that

$$\left| \sum_{m=1}^{M} \left(a_m - \frac{\pi h_{-n}}{A_{-n} w_{-n}} \right) \right| = \left| A_M - \frac{\pi h_{-n}}{A_{-n} w_{-n}} \right| = O(\sqrt{M}),$$

so by (2.1.5), we get convergence of $f(s)$ on $\mathrm{Re}(s) > 1/2$. Note that $f(s)$ was defined so that

$$\zeta_{-n}(s) = f(s) + \frac{\pi h_{-n}}{A_{-n} w_{-n}} \zeta(s).$$

Multiplying by $(s-1)$ and taking the limit as $s \to 1^+$ gives the result. □

Fact. *For $Re(s) > 1$, we can use Euler products to factor the Dedekind zeta function,*

$$\zeta_{-n}(s) = \zeta(s) \cdot L_{-n}(s).$$

Proof. We have

$$\begin{aligned}
\zeta_{-n}(s) &= \prod_{\left(\frac{-n}{p}\right)=1} (1-p^{-s})^{-2} \cdot \prod_{\left(\frac{-n}{p}\right)=0} (1-p^{-s})^{-1} \cdot \prod_{\left(\frac{-n}{p}\right)=-1} (1-p^{-2s})^{-1} \\
&= \prod_{p} (1-p^{-s})^{-1} \cdot \prod_{p,\left(\frac{-n}{p}\right)=1} (1-p^{-s})^{-1} \cdot \prod_{p,\left(\frac{-n}{p}\right)=-1} (1+p^{-s})^{-1} \\
&= \prod_{p} (1-p^{-s})^{-1} \cdot \prod_{p} \left(1 - \left(\frac{-n}{p}\right) p^{-s}\right)^{-1} \\
&= \zeta(s) \cdot L_{-n}(s).
\end{aligned}$$

□

Theorem (Dirichlet)**.** *Having established the continuity of $L_{-n}(s)$ for $Re(s) > 0$, we have that*

$$L_{-n}(1) = \begin{cases} \dfrac{\pi h_{-n}}{\sqrt{n} w_{-n}} & \text{if } n \equiv 1, 2 \mod 4 \\ \dfrac{2\pi h_{-n}}{\sqrt{n} w_{-n}} & \text{if } n \equiv 3 \mod 4. \end{cases}$$

Proof. From the previous factorization,

$$\begin{aligned}
L_{-n}(1) &= \lim_{s \to 1^+} L_{-n}(s) \\
&= \lim_{s \to 1^+} \frac{(s-1)\zeta_{-n}(s)}{(s-1)\zeta(s)} \\
&= \frac{\pi h_{-n}}{A_{-n} w_{-n}},
\end{aligned}$$

with

$$A_{-n} = \begin{cases} \sqrt{n} & \text{if } n \equiv 1, 2 \mod 4 \\ \sqrt{n}/2 & \text{if } n \equiv 3 \mod 4 \end{cases}$$

as before. □