A SURVEY OF RESULTS TOWARD THE CLASS NUMBER PROBLEM FOR REAL QUADRATIC FIELDS

by

JOSHUA BORWEIN NEVIN

B.Sc., California Institute of Technology, 2013

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Mathematics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April 2015

Abstract

The class number problem is one of the central open problems of algebraic number theory. It has long been known that there are only finitely many imaginary quadratic fields of class number one, and the full list of such fields is given by the Stark-Heegner theorem, but the corresponding problem for *real* quadratic fields is still open. It is conjectured that infinitely many real quadratic fields have class number one but at present it is still unknown even whether infinitely many algebraic number fields have class number one. This thesis reviews the relevant work that has been done on this problem in the last several decades. It is primarily concerned with a heuristic model of the sequence of real quadratic class groups called the *Cohen-Lenstra heuristics*, since this appears to offer the best hope of potentially solving the class number problem. The work of several other people who have put forward interpretations of the Cohen-Lenstra heuristics in other contexts, or who have generalized the heuristics, is also reviewed.

Preface

This thesis is the unpublished, independent work of the author, Joshua Borwein Nevin.

Table of Contents

Acknowledgements

I would like to thank my supervisor, Professor Sujatha Ramdorai, as well as Professor Greg Martin, who took over the task of helping me finish my thesis when Sujatha could not be here. I would also like to thank Professor Mike Bennett for acting as a reader for this thesis.

Introduction

Quadratic fields have been intensely studied, beginning with Gauss' development of the theory of binary quadratic forms. It is remarkable that Gauss developed the theory of class numbers without having any of the modern language of groups and fields. He conjectured that only finitely many negative fundamental discriminants have class number 1. This conjecture was proven by Heilbronn in 1934. The full list of imaginary quadratic fields with class number 1 was completed by Kurt Heegner and Harold Stark. It is conjectured that infinitely many real quadratic fields have class number one, but this problem is much more difficult than the corresponding problem for imaginary quadratic fields. Imaginary quadratic fields have unit group $\{\pm 1\}$ (except $\mathbb{Q}(\sqrt{-1})$ and $\mathbb{Q}(\sqrt{-3})$) but real quadratic fields have a unit group with a free part of rank 1 and so have a non-trivial regulator which is very difficult to control. Much of the work on this conjecture stems from a model created by Henri Cohen and Hendrik Lenstra. Cohen and Lenstra noticed that the (deterministic) sequence of quadratic class groups behaved as they were subject to a probability measure in which groups with many automorphisms were less likely to appear. If the heuristic assumptions of the Cohen-Lenstra model are true, the implications are far beyond what we can currently prove. For example, according to this model, the probability that a real quadratic class group has a trivial odd part is 75.446%.

The primary source for this model is from the original paper of Cohen and Lenstra, *Heuristics on Class Groups of Number Fields*, but the ideas contained therein have been expanded on by several other people, such as Lengler (see [3]) and Fulman (see [2]).

Chapter 1

An Overview of the Basic Theory

1.1 The Ideal Class Group

If K is a number field with number ring $\mathcal{O}_K$, then the group of fractional ideals of $\mathcal{O}_K$ form a free abelian group J_K generated by the primes. This contains a subgroup P_K generated by all fractional principal ideals. The *ideal class group* is the quotient group J_K/P_K. Crucially, this group is finite. There are many elegant proofs of the finiteness of this group. One such method starts by embedding the field K into the Euclidean space $\mathbb{R}^r \times \mathbb{C}^s$, where K has r embeddings into $\mathbb{R}$ and s conjugate pairs of embeddings into $\mathbb{C}$. The finitness of the ideal class group can then be deduced by regarding each nonzero ideal as a lattice in $\mathbb{R}^{r+2s}$ and applying Minkowski's theorem. The *class number* of the number field K is $|J_K/P_K|$.

For K a number field, the number ring $\mathcal{O}_K$ is a free $\mathbb{Z}$-module whose rank is equal to the number of embeddings of K into $\mathbb{C}$. Every field K comes equipped with a *discriminant*.

Definition 1.1.1 If K is a number field whose number ring $\mathcal{O}_K$ has basis $\{\alpha_1, \cdots, \alpha_n\}$ as a free $\mathbb{Z}$-module, and such that $\{\sigma_1, \sigma_2, \cdots, \sigma_n\}$ is the set of embeddings of K into $\mathbb{C}$ then the *discriminant* of K is the following quantity

$$\det \begin{pmatrix} \sigma_1(\alpha_1) & \sigma_2(\alpha_1) & \cdots & \sigma_n(\alpha_1) \\ \sigma_2(\alpha_1) & \sigma_2(\alpha_2) & \cdots & \sigma_2(\alpha_n) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_n(\alpha_1) & \sigma_n(\alpha_2) & \cdots & \sigma_n(\alpha_n) \end{pmatrix}.$$

We denote this quantity by $\text{disc}(K)$.

Our interest lies in the class groups of the quadratic number fields, that is, a number field of the form $K = \mathbb{Q}(\sqrt{k})$ for k squarefree.

Proposition 1.1.2 *The discriminant of the number field $\mathbb{Q}(\sqrt{k})$, for k squarefree, is*

$$\text{disc}\left(\mathbb{Q}(\sqrt{k})\right) = \begin{cases} k \text{ if } k \equiv 1 \mod 4 \\ 4k \text{ if } k \equiv 2, 3 \mod 4 \end{cases}.$$

Later on, we will develop the theory of binary quadratic forms. In the language of Gauss' theory of binary quadratic forms, the discriminant of a quadratic is a *fundamental discriminant*.

Definition 1.1.3 A *binary quadratic form* is a homogenous degree 2 polynomial $Q(x, y) = ax^2 + bxy + cy^2$ with integer coefficients. The *discriminant* of the binary quadratic form $ax^2 + bxy + cy^2$ is the invariant $b^2 - 4ac$. That is, a discriminant is an integer of the form $b^2 - 4ac$ where $a, b, c \in \mathbb{Z}$. A discriminant D is called *fundamental* if it cannot be written in the form $D = D_0 f^2$ for D_0, f integers with $f > 1$.

The discriminant of a quadratic number field (as in definition 1.1.1/proposition 1.1.2) is a fundamental discriminant (as in definition 1.1.3). Thus, the notation $h(d)$ will always be reserved for the class number of the unique quadratic field with fundamental discriminant d. For an algebraic number field K, we will also use the notation $h(K)$ to refer to the class number of K.

In the case where $d < 0$, the problem of studying $h(d)$ is made considerably simpler by the fact that the corresponding quadratic field has a finite unit group:

Proposition 1.1.4 *(Dirichlet's Unit Theorem)*

Let K be an algebraic number field with r real embeddings and s pairs of conjugate complex embeddings. Then the unit group $\mathcal{O}_K^*$ is isomorphic to $\mu(K) \times \mathbb{Z}^{r+s-1}$, where $\mu(K) \subseteq \mathbb{C}^*$ is the group of roots of

unity lying in K.

Thus, imaginary quadratic fields, for which $r = 0$ and $s = 1$, have no free part to their unit group. For $k > 0$ and $K = \mathbb{Q}(\sqrt{-k})$, the unit group of K is just $\{1, -1\}$, unless $k = 1$ or $k = 3$. But for any other algebraic number field K apart from the imaginary quadratic fields and $\mathbb{Q}$, K comes equipped with a *system of fundamental units*. A set $S \subseteq \mathcal{O}_K^*$ is called a system of fundamental units if it is a minimal generating set of the free part of $\mathcal{O}_K^*$.

In the particular case we are interested in, that of real quadratic fields, a system of fundamental units is just a singleton, since $r + s - 1 = 1$. In this case, we have an exact characterization of the fundamental units in terms of Diophantine equations. Namely, a fundamental unit of a real quadratic field is a minimal solution to *Pell's equation*:

Proposition 1.1.5 *Let $k > 1$ be a squarefree integer and d the discriminant of the real quadratic number field $K = \mathbb{Q}(\sqrt{k})$. Let x_1, y_1 be the uniquely determined minimal positive rational integer solution of the equation $x^2 - dy^2 = -4$, unless no such solutions exist. In that case, let x_1, y_1 be the uniquely determined minimal positive rational integer solutions of the equation $x^2 - dy^2 = 4$. Then $u_1 = \frac{x_1 + y_1\sqrt{d}}{2}$ is a fundamental unit of K.*

Proof: We consider two cases. In the first case $k \equiv 1 \mod 4$. In that case $k = d$ by proposition 1.1.2. Thus

$$N(u_1) = \frac{(x_1^2 - dy_1^2)}{4} = \pm 1.$$

Here the norm is with respect to the extension $K/\mathbb{Q}$. So in any case u_1 is a unit. Suppose it is not a fundamental unit. The number ring of $\mathbb{Q}(\sqrt{k})$ is $\mathbb{Z}\left[\frac{1+\sqrt{d}}{2}\right]$, as $k = d$. Therefore, there is a fundamental unit u expressible as

$$u = a + b\left(\frac{1 + \sqrt{d}}{2}\right) = \frac{(2a + b) + b\sqrt{d}}{2}$$

where $a, b \in \mathbb{Z}$.

Note we can assume that both $2a+b$ and b are nonnegative integers because if both are negative, then $-u$ is also a fundamental unit. If one of them is negative, then we can just pass to the Galois conjugate of u (or $-u$ if necessary), which is also a fundamental unit. So we may assume that u has $2a+b>0$ and $b>0$. Since u_1 is not a fundamental unit, there exists a $j \in \mathbb{Z}$ where $j \neq 0, 1, -1$, such that $u_1 = u^j$. By minimality, we have

$$0 < x_1 \leq 2a+b$$

$$0 < y_1 \leq b.$$

Since $2a+b, b, x_1, y_1$ are all positive integers, and $d \geq 5$ (since $d \equiv 1 \mod 4$), this means that $u_1 > 1$ and $u > 1$. Therefore, we are forced to have $j > 0$, otherwise $u_1 < 1$. But this contradicts the minimality of x_1, y_1.

Now we do the case where $k \equiv 2, 3 \mod 4$. Thus $d = 4k$ by 1.1.2. Furthermore, in this case the number ring of K is $\mathbb{Z}[\sqrt{k}]$. As before, we suppose $u_1 = \frac{x_1+\sqrt{d}y_1}{2}$ where x_1, y_1 are the minimal positive solutions to the Diophantine equation above. Suppose u_1 is not a fundamental unit of $\mathcal{O}_K^*$. In this case, the fundamental unit u of $\mathcal{O}_K$ can be expressed in the form $a+b\sqrt{k}$. Since $d = 4k$ we write this as $\frac{2a+b\sqrt{d}}{2}$. As before, we may suppose that $2a, b$ are nonnegative integers. Then for some $j \neq 0, 1, -1$ we have $u_1 = u^j$. Since $u_1 = \frac{x_1+2y_1\sqrt{d}}{2}$, we have $u_1 > 1$ so as above, $j > 1$, contradicting minimality.

□

1.2 The Dedekind ζ-function and the Class Number Formula

For any number field K we have a Dirichlet series

$$\sum_{I \neq 0} (NI)^{-s}$$

where the sum is taken over all nonzero ideals I of $\mathcal{O}_K$. The notation NI will always denote the ideal norm of the ideal I. That is, $NI = |\mathcal{O}_K : I|$.

This Dirichlet series converges on the half-plane $\mathrm{Re}(s) > 1$ and can be meromorphically continued to a function $\zeta_K(s)$ on the half-plane $\mathrm{Re}(s) > 1 - \frac{1}{|K:\mathbb{Q}|}$ with a simple pole at $s = 1$. In particular if $K = \mathbb{Q}$ this is just the ordinary Riemann-zeta function. This machinery is enough to provide an

analytic formula for the class number of K. Through Dirichlet's Unit Theorem, each algebraic number field K is associated an invariant $\mathrm{reg}(K)$, the *regulator* of K, which is a measure of the density of the unit group in the logarithmic space.

Definition 1.2.1 For an algebraic number field K with degree $n = r + 2s$, set $t = r + s - 1$. Then let $\{u_1, \cdots, u_t\}$ be a system of fundamental units of $\mathcal{O}_K^*$. Let $\tau_1, \cdots, \tau_{t+1}$ be the different embeddings of K into $\mathbb{R}$ or $\mathbb{C}$ (up to conjugation). Then we form a $t \times (t+1)$ matrix whose (i,j)th entry is $N_j \log |\tau_j(u_i)|$, where $N_j = 1$ if τ_j is a real embedding and 2 if it is a complex embedding up to conjugation. Then let R be the absolute value of the determinant of an arbitrary $t \times t$ minor of the matrix. Then R is called the *regulator* of K.

It can be shown that the determinant of the matrix above is independent of our choice of system of fundamental units. In the case of imaginary quadratic fields, we just set the regulator to be 1, because there are no fundamental units. In the case of a real quadratic field $\mathbb{Q}(\sqrt{k})$, suppose that we have a fundamental unit $a + b\sqrt{k}$. In that case, by the definition above, we form a 1×2 matrix

$$\begin{pmatrix} \log |a + b\sqrt{k}| \\ \log |a - b\sqrt{k}| \end{pmatrix}.$$

Since $a + b\sqrt{k}$ is a unit, it has norm 1. Thus $\log |a + b\sqrt{k}| = -\log |a - b\sqrt{k}|$, so we obtain the same absolute value regardless of which 1×1 minor we choose. The absolute value $|\log |a + b\sqrt{k}||$ is the regulator of $\mathbb{Q}(\sqrt{k})$.

The connection between the regulator and the ideal class group is given by the *class number formula*.

Proposition 1.2.2 *(The Class Number Formula)*

For a number field K with r real embeddings and s pairs of complex embeddings, the class number $h(K)$ is given by

$$\left(\lim_{s \to 1} \frac{\zeta_K(s)}{\zeta(s)}\right) \frac{\#\mu(K)\sqrt{|\mathrm{disc}(K)|}}{2^{r+s}\pi^s \mathrm{reg}(K)}.$$

This quantity $\left(\lim_{s\to 1} \frac{\zeta_K(s)}{\zeta(s)}\right)$ can now be studied by introducing a more general class of functions. Dirichlet L-functions, which were originally introduced to prove Dirichlet's theorem on arithmetic pro-

gression, can be used to give an expression for $\left(\lim_{s\to 1}\frac{\zeta_K(s)}{\zeta(s)}\right)$. For a Dirichlet character $\chi : (\mathbb{Z}/m\mathbb{Z})^* \to \mathbb{C}$, we associate the Dirichlet series

$$\sum_{n=0}^{\infty}\frac{\chi(n)}{n^s}.$$

This converges in the half-plane $\mathrm{Re}(s) > 1$ and can be analytically continued to a meromorphic function $L(s,\chi)$ on the complex plane. Now, if K is an abelian extension of $\mathbb{Q}$ with Galois group $G = \mathrm{Gal}(K/\mathbb{Q})$, and corresponding character group $\hat{G}$, then G is contained in some cyclotomic extension $\mathbb{Q}(e^{2\pi i/m})$ by the Kronecker-Weber theorem. Thus each element $\chi : G \to \mathbb{C}^*$ of the character group $\hat{G}$ may be regarded as a Dirichlet character mod m.

The term $\left(\lim_{s\to 1}\frac{\zeta_K(s)}{\zeta(s)}\right)$ can be evaluated in terms of L-functions, at least in the case where K is abelian:

Proposition 1.2.3 *Let K be an abelian extension of $\mathbb{Q}$ contained in the cyclotomic extension $\mathbb{Q}(e^{2\pi i/m})$. Let $G = \mathrm{Gal}(K/\mathbb{Q})$ and let $\hat{G}$ be the corresponding character group. For each prime $p|m$, let f_p be the inertia degree of p in K and let r_p be the number of distinct primes in K above p. Then*

$$\left(\lim_{s\to 1}\frac{\zeta_K(s)}{\zeta(s)}\right) = \prod_{p|m}\left(1-p^{-1}\right)\left(1-p^{-f_p}\right)^{-r_p}\prod_{\substack{\chi\in\hat{G}\\ \chi\neq 1}}L(1,\chi).$$

Proof: See [6].

□

Before continuing, we introduce several important properties of characters.

Definition 1.2.4 Let $\chi : (\mathbb{Z}/n\mathbb{Z})^* \to \mathbb{C}$ be a character mod n. Suppose that m is a multiple of n, and so there is a projection $\pi : (\mathbb{Z}/m\mathbb{Z})^* \to (\mathbb{Z}/n\mathbb{Z})^*$. Then χ is said to *induce* the character mod m $\chi' = \chi \circ \pi$. A character χ' mod m is called *primitive* if it is not induced by a character mod n for any divisor n of m such that $n \neq m$. The character χ' is called *even* if $\chi'(-1) = 1$ and *odd* if $\chi'(-1) = -1$.

In the case of a quadratic field $K = \mathbb{Q}(\sqrt{k})$ with Galois group G, we are interested in describing the lone nontrivial character $\chi : G \to \mathbb{C}^*$ by considering an induced cyclotomic character. Let $m = |\mathrm{disc}(K)|$. Then K is contained in the cyclotomic field $\mathbb{Q}(e^{2\pi i/m})$. Thus $\mathrm{Gal}(\mathbb{Q}(e^{2\pi i/m})/\mathbb{Q})$ projects to $\mathrm{Gal}(K/\mathbb{Q})$. The first Galois group is $(\mathbb{Z}/m\mathbb{Z})^*$. Let $H \leq \mathrm{Gal}(\mathbb{Q}(\omega)/\mathbb{Q})$ be subgroup fixing K, so that $\mathrm{Gal}(K/\mathbb{Q}) \cong$

G/H. If p is an odd prime, with $p \nmid m$, (and so p is unramified in $\mathbb{Q}(\omega)$), then $p \mod m \in H$ if and only if k is a square mod p. Furthermore, if $p = 2$ then $p \mod m \in H$ if and only if $k \equiv 1 \mod 8$.

So, for the character χ mod m induced by a nontrivial character on $(\mathbb{Z}/m\mathbb{Z})^*/H$, χ takes the value of 1 on $a \in (\mathbb{Z}/m\mathbb{Z})^*$ if and only if $a \in H$. Otherwise, it takes the value -1. Thus, for odd primes $p \nmid m$ we obtain

$$\chi(p) = \begin{cases} 1 \text{ if } k \text{ is a square mod } p \\ -1 \text{ otherwise} \end{cases} = \left(\frac{k}{p}\right)$$

and

$$\chi(2) = \begin{cases} 1 \text{ if } k \equiv 1 \mod 8 \\ -1 \text{ if } k \equiv 5 \mod 8 \\ 0 \text{ otherwise} \end{cases}$$

So the character χ in the class number formula is the multiplicative extension of the character above. We introduce the Jacobi symbol. For $a \in \mathbb{Z}$ and odd $b > 0$ with $(a, b) = 1$ with $b = p_1^{r_1} \cdots p_k^{r_k}$ we define

$$\left(\frac{a}{b}\right) = \prod_{i=1}^{k} \left(\frac{a}{p_i}\right)^{r_i}$$

where $\left(\frac{a}{p_i}\right)$ is the usual Legendre symbol. Thus for odd n we obtain

$$\chi(n) = \left(\frac{k}{n}\right).$$

And with $\chi(2)$ defined above, χ is now defined over the positive integers.

It is straightforward to show that χ is a primitive (mod m) even character when $d > 0$ and a primitive (mod m) odd character when $d < 0$. So with this description in hand, for each fundamental discriminant d we set χ_d to be the unique nontrivial character above associated to the quadratic field with discriminant d, which is a character mod $|d|$. Furthermore, we note that this is a primitive character mod $|d|$. Then

we have the following formula for $h(d)$:

Proposition 1.2.5 *If $d < 0$, let μ_d be the group of roots of unity contained in the unique quadratic field with fundamental discriminant d (if $d > 0$ then the only roots of unity in the corresponding quadratic field are 1 and -1). If $d > 0$ then let u_d be a fundamental unit of the unique quadratic field with fundamental discriminant d. Then*

$$h(d) = \begin{cases} \frac{\#\mu_d \sqrt{|d|}}{2\pi} L(1, \chi_d) \text{ if } d < 0 \\ \\ \frac{\sqrt{d}}{2|\log(|u_d|)|} L(1, \chi_d) \text{ if } d > 0 \end{cases}.$$

Proof: See [6]

□

Thus, in the case of $d < 0$, studying $h(d)$ amounts to studying $L(1, \chi_d)$. The growth of $\frac{L(1,\chi_d)}{\sqrt{|d|}}$ as $|d| \to \infty$ depends on the size of the zero-free region of $L(s, \chi_d)$. For $d < 0$, χ_d is an odd, real primitive character mod $|d|$. Hecke proved the following theorem:

Proposition 1.2.6 *For $d < 0$ and χ an odd, real primitive character mod $|d|$, if $L(s, \chi) \neq 0$ in the region $s > 1 - \frac{c}{\log |d|}$ for s real and $c > 0$ some fixed absolute constant, then*

$$L(1, \chi_d) \gg \frac{1}{\log |d|}.$$

This theorem was the first step to solving a long-standing problem of Gauss. Gauss had conjectured that $\lim_{d \to -\infty} h(d) = \infty$. This is referred to as *Gauss' conjecture for imaginary quadratic fields.* Gauss also conjectured that infinitely many *real* quadratic fields have class number one. That is, he conjectured that $\liminf_{d \to \infty} h(d) = 1$. This is referred to as *Gauss' conjecture for real quadratic fields.*

The proof of Gauss' conjecture was then completed by Heilbronn in 1934, who showed that $h(d) \to \infty$ as $d \to -\infty$ unconditionally.

The behavior of the real quadratic class fields is much harder to study, owing to the difficulty of controlling the term $\frac{1}{|\log(|u_d|)|}$. So far, Gauss' conjecture for real quadratic field has resisted any attempt at an analytic proof (or any other kind of proof). The approach outlined in the following section is

perhaps the current best hope of achieving a full solution. In this model, the sequence of odd parts of the real quadratic class groups are treated as if they are subject to a probability measure. Under the assumption that the odd parts of the real quadratic class groups behave as if they were generated by a weighted random process with this probability measure, we can predict the probability that a particular odd part of a real quadratic class group will be trivial (we can predict much more, as we shall see). The proof of this heuristic assumption, if it exists, must lie very deep.

Chapter 2

The Cohen-Lenstra Model

Throughout this chapter we fix the following notation. The letter A will denote a number ring. The letter G will always denote a finite A-module. The notation $\mathcal{G}$ will denote the category of finite A-modules. For G a finite A-module, $\mathrm{Aut}(G)$ will denote the A-automorphisms of G. Furthermore, if B, C are A-modules, then unless otherwise indicated $\mathrm{Hom}(B, C)$ will always denote the A-module homomorphisms from B to C, i.e $\mathrm{Hom}_A(B, C)$. The letter I will always denote a nonzero ideal of A. The script letter $\mathfrak{p}$ will always denote a prime ideal of A. We present some of the proofs of the propositions which form the model, in order to give an overview of the methods being used, but we leave some proofs as cited.

2.1 Modules over Dedekind Domains

We review the theory of finite modules over Dedekind domains. A finite module G over A decomposes into a direct sum of the form

$$G = \bigoplus_{i=1}^{n} A/\mathfrak{p}_i^{m_i}$$

where $\mathfrak{p}_1, \cdots, \mathfrak{p}_n$ are prime ideals of A (not necessarily distinct). The ideal $\mathfrak{p}_1^{m_1} \cdots \mathfrak{p}_\mathfrak{n}{}^{m_n}$ of A is an invariant associated to G. We denote this ideal as $\chi_A(G)$. We refer to $\chi_A(G)$ as the *cardinal* of G. Note that in the case where $A = \mathbb{Z}$, then $\chi_\mathbb{Z}(G) = (\#G)\mathbb{Z}$. In the general case, $N\chi_A(G) = \#G$, where N is the ideal norm on A. We recall that $NI = |A : I|$ for $I \neq 0$ an ideal of A. Throughout this chapter,

NI will always denote the ideal norm of I. One advantage of χ_A is its multiplicativity through exact sequences of finite A-modules. That is, for

$$0 \to G_1 \to G \to G/G_1 \to 0$$

an exact sequence of finite A-modules we have

$$\chi_A(G) = \chi_A(G_1)\chi_A(G/G_1).$$

For a given prime ideal $\mathfrak{p}$ of A, the $\mathfrak{p}$-*rank* of G is the dimension of $G/\mathfrak{p}G$ as an $A/\mathfrak{p}$-vector space. Note that this is just the number of direct summands of G of the form $A/\mathfrak{p}^m$ in the decomposition above, with $m > 0$. We denote the $\mathfrak{p}$-rank of G by $r_{\mathfrak{p}}(G)$. We note that G has a nonzero $\mathfrak{p}$-rank if and only if $\mathfrak{p}|\chi_A(G)$. Furthermore, we define the *rank* of G to be the minimal number of generators of G as an A-module. We denote this by $r(G)$. Finally, for any prime ideal $\mathfrak{p}$ and $G \in \mathcal{G}_A$, the notation $G_{\mathfrak{p}}$ will denote the $\mathfrak{p}$-part of G. More generally, for any $\mathcal{P} \subseteq \mathrm{Spec}(A)$ and finite A-module G, the notation $G_{\mathcal{P}}$ will denote the $\mathcal{P}$-part of G. If there is only a single prime ideal $\mathfrak{p}|\chi_A(G)$ then we refer to G as a $\mathfrak{p}$A-module. More generally, if $\mathcal{P} \subseteq \mathrm{Spec}(A)$ and the primes dividing $\chi_A(G)$ all lie in $\mathcal{P}$ then we refer to G as a $\mathcal{P}A$-module.

For each integer $k \geq 1$ we introduce a k-*weight* to each finite A-module G

Definition 2.1.1

$$w_k(G) := \frac{\#\{\varphi \in \mathrm{Hom}_A(A^k, G) : \varphi \text{ surjective}\}}{(\#G)^k(\#\mathrm{Aut}(G))}.$$

$$w(G) := \frac{1}{\#(\mathrm{Aut}(G))}.$$

The purpose of introducing $w_k(G)$ is for ease of calculation. We will want to calculate the averages of certain functions $f : \mathcal{G}_A \to \mathbb{C}$ in which each $G \in \mathcal{G}_A$ is weighted by $w(G)$. It is usually easier to weight each $G \in \mathcal{G}_A$ by $w_k(G)$ and then let $k \to \infty$, since, as we shall see, $\lim_{k\to\infty} w_k(G) = w(G)$ for all $G \in \mathcal{G}_A$.

We note that $\frac{\#\{\varphi \in \mathrm{Hom}_A(A^k,G):\varphi \text{ surjective}\}}{(\#\mathrm{Aut}(G))}$ is simply $\#\{J \leq A^k : A^k/J \cong G\}$, because two surjective homomorphisms $\varphi_1, \varphi_2 : A^k \to G$ have the same kernel if and only if $\varphi_2 = \phi \circ \varphi_1$ for some $\phi \in \mathrm{Aut}(G)$.

We also note that $\#\{\varphi \in \mathrm{Hom}_A(A^k, G) : \varphi \text{ surjective}\} = 0$ if $k < r_{\mathfrak{p}}(G)$ for any prime ideal $\mathfrak{p}$ of A.

Finally, we note that $w(G)$ induces a local probability measure on $\mathcal{G}_A$ in the following sense. For each $\mathfrak{p}$ set $\mathcal{G}_{A,\mathfrak{p}}$ to be the set of finite $\mathfrak{p}A$-modules. Then for any $M \subseteq \mathcal{G}_{A,\mathfrak{p}}$, the probability that M occurs (with respect to $w(G)$) is just

$$\frac{\sum_{G \in M}(\#(\mathrm{Aut}(G)))^{-1}}{\sum_{G \in \mathcal{G}_{A,\mathfrak{p}}}(\#(\mathrm{Aut}(G)))^{-1}}.$$

This makes sense because $\sum_{G \in \mathcal{G}_{A,\mathfrak{p}}} \frac{1}{\#(\mathrm{Aut}(G))} < \infty$, which we shall see in the next section when we regard the weights $w(G)$ as Dirichlet coefficients.

Because the ideals of a number ring A are ordered by their norms, we can talk about the density of a particular subset of the set of all finite A-modules by ordering these modules by their cardinals. In view of this we introduce the following notations:

Definition 2.1.2 *For any function $f : \mathcal{G}_A \to \mathbb{C}$ and nonzero ideal I of A we introduce the notation*

$$\sum_{G(I)} f(G) := \sum_{\substack{G \text{ an isomorphism class} \\ \chi_A(G)=I}} f(G).$$

Definition 2.1.3 *For I a nonzero ideal of A and $k \geq 1$ an integer we set*

$$w_k(I) := \sum_{G(I)} w_k(G)$$

and

$$w(I) := \sum_{G(I)} w(G).$$

Definition 2.1.4 For any function $f : \mathcal{G}_A \to \mathbb{C}$ and $k \geq 1$ an integer we define the *CL-k-average of f* to be

$$M_{k,A}(f) = \lim_{x \to \infty} \frac{\sum_{NI \leq x}(NI)^{-1} \sum_{\substack{G(I) \\ \varphi \in \mathrm{Hom}(A,G)}} w_k(G) f(G/\mathrm{im}(\varphi))}{\sum_{NI \leq x}(NI)^{-1} \sum_{\substack{G(I) \\ \varphi \in \mathrm{Hom}(A,G)}} w_k(G)}$$

and we define the *CL-average of f* to be

$$M_A(f) = \lim_{x\to\infty} \frac{\sum_{NI\leq x}(NI)^{-1}\sum_{\substack{G(I)\\ \varphi\in \mathrm{Hom}(A,G)}} w(G)f(G/\mathrm{im}(\varphi))}{\sum_{NI\leq x}(NI)^{-1}\sum_{\substack{G(I)\\ \varphi\in \mathrm{Hom}(A,G)}} w(G)}.$$

If the ring A is clear from from the context, we just write $M_k(f)$ and $M(f)$ respectively. Of course, neither $M_k(f)$ nor $M(f)$ is guaranteed to exist for a particular function f, so the above definitions apply only if $M_k(f)$ and $M(f)$ exist respectively. The original paper of Cohen and Lenstra (see [5]) introduces a more general notion of a (k,u)-average with applications in calculating the ranks of groups of units, but the full generality of the model is not necessary for our purposes. If P is a property of a finite A-module and $f_P : \mathcal{G}_A \to \mathbb{C}$ is the corresponding indicator function, which takes the value of 1 if G satisfies P and 0 otherwise, then we refer to $M_k(f_P)$ as the *CL-k-probability* of P and we refer to $M(f_P)$ as the *CL-probability of P*.

We note now that the denominator of $M_k(f)$ is simply

$$\sum_{NI\leq x} w_k(I)$$

and likewise for $k=\infty$, with $w(I)$ replacing $w_k(I)$. This is because for $\chi_A(G) = I$, $\#\mathrm{Hom}_A(A,G) = \#G = NI$.

The goal of the next section will be to express $M_k(f)$ and $M(f)$ in terms of Dirichlet series. To do this we must first provide formulae for $w_k(G)$ and $w_k(I)$.

Proposition 2.1.5 *For any $k \geq 1$ and $G \in \mathcal{G}_A$ with $\chi_A(G) = I$*

i)

$$w_k(G) = \begin{cases} \frac{1}{\#Aut(G)} \prod_{\mathfrak{p}|I} \prod_{i=k-r_{\mathfrak{p}}(G)+1}^{k} (1 - N\mathfrak{p}^{-i}) \text{ if } k \geq r_{\mathfrak{p}}(G) \\ 0 \text{ otherwise} \end{cases}$$

ii)

$$w(G) = \lim_{k\to\infty} w_k(G).$$

Proof: We first note that if $\mathfrak{p}_1, \cdots, \mathfrak{p}_j$ is the list of primes dividing I then

$$w_k(G) = \prod_{i=1}^{j} w_k(G_{\mathfrak{p}_i})$$

since choosing a surjective homomorphism $\varphi : A^k \to G$ just amounts to choosing a surjective homomorphism $\varphi : A^k \to G_{\mathfrak{p}_i}$ for each $1 \leq i \leq j$, likewise for A-automorphisms of G. So it suffices to prove the formula in i) for a $\mathfrak{p}$A-module G. So let $\mathfrak{p}$ be the lone prime dividing I and set $r = r_{\mathfrak{p}}(G)$. It is clear that if $k < r$ then $w_k(G) = 0$ as there are no surjective homomorphisms from A^k to G in that case. So we suppose $k \geq r$.

Now we note that $\varphi : A^k \to G$ induces a $A/\mathfrak{p}$-vector space homomorphism $\overline{\varphi} : (A/\mathfrak{p})^k \to G/\mathfrak{p}G$. To see this, we let $\pi \circ \varphi : A^k \to G/\mathfrak{p}G$ be the homomorphism from A^k to $G/\mathfrak{p}G$ induced by projection. Now, note that for each $1 \leq i \leq k$, $\pi \circ \varphi$ induces an A-module homomorphism $(\pi \circ \varphi)_i : A \to G/\mathfrak{p}G$ via $(\pi \circ \varphi)_i(x) = (\pi \circ \varphi)(0, \cdots, x, \cdots, 0)$, where $(0, \cdots, x, \cdots, 0) \in A^k$ is the k-tuple with x in the ith coordinate. Of course, the annihlator $\mathrm{Ann}_A(G/\mathfrak{p}G) \supseteq \mathfrak{p}$. Thus $(\pi \circ \varphi)_i$ factors through $A/\mathfrak{p}$, and so we get a well defined homomorphism $\overline{(\pi \circ \varphi)_i} : A/\mathfrak{p} \to G/\mathfrak{p}G$ for each $1 \leq i \leq k$. Thus we obtain a well-defined homomorphism $\overline{\varphi} : (A/\mathfrak{p})^k \to G/\mathfrak{p}G$.

It is clear that if φ is surjective then so is $\overline{\varphi}$. The converse is also true. If $\overline{\varphi} : (A/\mathfrak{p})^k \to G/\mathfrak{p}$ is a surjective $A/\mathfrak{p}$-vector space homomorphism and $\varphi : A^k \to G$ is an A-module homomorphism which induces $\overline{\varphi}$, then φ is also surjective.

To see this, we note that if $\overline{\varphi}$ is surjective, then there exist some $\{h_1, \cdots, h_s\} \in \varphi(A^k)$, where $s = \frac{\#G}{\#\mathfrak{p}G}$, which form a complete system of residue classes mod $\mathfrak{p}G$. Now suppose toward a contradiction that there is some $g \in G$ with $g \notin \varphi(A^k)$. Then we may write $g = h_{i_1} + p_1 g_1$ for some $p_1 \in \mathfrak{p}$ and $g_1 \in G$. Thus $p_1 g_1 \notin \varphi(A^k)$. We repeat this process. We write $g_1 = h_{i_2} + p_2 g_2$ for some $p_2 \in \mathfrak{p}$ and $g_2 \in G$. Thus $p_1 p_2 g_2 \notin \varphi(A^k)$. Continuing in this way, we obtain a sequence of elements $p_1 g_1, p_1 p_2 g_2, p_1 p_2 p_3 g_3, \cdots$, each lying outside $\varphi(A^k)$, where the nth element lies in $\mathfrak{p}^n G$. But this is clearly false. Since G is a finite $\mathfrak{p}A$-module, there exists some n such that $\mathfrak{p}^n G = 0$.

Thus we conclude that φ is surjective if and only if $\overline{\varphi}$ is surjective. Furthermore, for each $\overline{\varphi} : (A/\mathfrak{p})^k \to G/\mathfrak{p}G$, the number of lifts of $\overline{\varphi}$ to a surjective homomorphism from A^k to G is just $\#\{\phi \in \mathrm{Hom}(A^k, G) : \overline{\phi} = 0\}$, since two surjective homomorphisms $\varphi_1, \varphi_2 : A^k \to G$ pass to the same quotient homomorphism if and only if they lie in the same coset of $\mathrm{Hom}(A^k, G)$ modulo the submodule $\{\phi \in \mathrm{Hom}(A^k, G) : \overline{\phi} = 0\}$.

Thus we get an equality between $\#\{\varphi \in \mathrm{Hom}(A^k, G) : \varphi \text{ surjective}\}$ and $\#\{\overline{\varphi} \in \mathrm{Hom}_{A/\mathfrak{p}}((A/\mathfrak{p})^k, G/\mathfrak{p}G) : \overline{\varphi} \text{ surjective}\}\#\{\phi \in \mathrm{Hom}(A^k, G) : \overline{\phi} = 0\}$, as required.

Now we count $\#\{\overline{\varphi} \in \mathrm{Hom}_{A/\mathfrak{p}}((A/\mathfrak{p})^k, G/\mathfrak{p}G) : \overline{\varphi} \text{ surjective}\}$. $G/\mathfrak{p}G$ is just a $A/\mathfrak{p}$-vector space of dimension r. So we want to count the $k \times r$ matrices X which have entries in $A/\mathfrak{p}$ and which have rank r. So there are $N\mathfrak{p}^k$ choices for the first column vector v_1 of X, where $v_1 \in (A/\mathfrak{p})^k$, since $N\mathfrak{p}^k = \#(A/\mathfrak{p})^k$. Now we choose a second column vector $v_2 \in (A/\mathfrak{p})^k$ linearly independent to v_1. There are $N\mathfrak{p}^k - N\mathfrak{p}$ such choices, as we just avoid all the scalar multiples of v_1. Now we choose a $v_3 \in (A/\mathfrak{p})^k$ which is independent to v_1, v_2. There are $N\mathfrak{p}^k - N\mathfrak{p}^2$ such choices, as we just avoid all $A/\mathfrak{p}$-linear combinations of v_1, v_2, of which there are $N\mathfrak{p}^2$. Thus continuing in this way we obtain

$$\#\{\overline{\varphi} \in \mathrm{Hom}_{A/\mathfrak{p}}((A/\mathfrak{p})^k, G/\mathfrak{p}G) : \overline{\varphi} \text{ surjective}\} = \prod_{1 \leq i \leq r} (N\mathfrak{p}^k - N\mathfrak{p}^i).$$

Now we consider the other term. $\#\{\phi \in \mathrm{Hom}(A^k, G) : \overline{\phi} = 0\}$ is just the set of $\phi \in \mathrm{Hom}(A^k, G)$ such that $\phi(v) \in \mathfrak{p}G$ for all $v \in A^k$. The number of A-module homomorphisms $\phi : A \to \mathfrak{p}G$ is just $\#\mathfrak{p}G$, so this quantity is just

$$(\#\mathfrak{p}G)^k = \frac{\#G^k}{\#(G/\mathfrak{p}G)^k} = \frac{(NI)^k}{(N\mathfrak{p})^{kr}}.$$

Thus

$$\#\{\varphi \in \mathrm{Hom}(A^k, G) : \varphi \text{ surjective}\} = \frac{(NI)^k}{(N\mathfrak{p})^{kr}} \prod_{1 \leq i \leq r} (N\mathfrak{p}^k - N\mathfrak{p}^i).$$

Furthermore:

$$\prod_{1 \leq i \leq r} (N\mathfrak{p}^k - N\mathfrak{p}^i) = N\mathfrak{p}^{kr} \prod_{i=k-r+1}^{k} (1 - N\mathfrak{p}^{-i}).$$

Thus

$$w_k(G) = \frac{\#\{\varphi \in \mathrm{Hom}(A^k, G) : \ \varphi \text{ surjective}\}}{w(G)\#G^k} = \frac{1}{w(G)} \prod_{i=k-r+1}^{k} (1 - N\mathfrak{p}^{-i})$$

since $\#G = NI$. This completes the proof of i). Letting $k \to \infty$, the product $\prod_{i=k-r+1}^{k}(1-N\mathfrak{p}^{-i}) \to 1$, so ii) follows immediately.

□

Now that we have provided a formula for $w_k(G)$, we would like to provide a formula for $w_k(I)$ for I a nonzero ideal of A. We immediately note that w_k is multiplicative in coprime ideals of A since for I, J coprime and G, H finite A-modules with $\chi_A(G) = I$ and $\chi_A(H) = J$, $w_k(G)w_k(H) = w_k(G \oplus H)$ and so

$$\left(\sum_{\chi(G)=I} w_k(G)\right)\left(\sum_{\chi(H)=J} w_k(H)\right) = \sum_{\chi_A(K)=IJ} w_k(K).$$

Thus it suffices to provide a formula for $w_k(\mathfrak{p}^m)$.

Proposition 2.1.6 *For any $m \geq 0$ and $k \geq 1$*

i)

$$w_k(\mathfrak{p}^m) = N\mathfrak{p}^{-m}\left(\prod_{i=k}^{k+m-1}(1-N\mathfrak{p}^{-i})\right)\left(\prod_{i=1}^{m}(1-N\mathfrak{p}^{-i})^{-1}\right)$$

ii)

$$w(\mathfrak{p}^m) = N\mathfrak{p}^{-m}\left(\prod_{i=1}^{m}(1-N\mathfrak{p}^{-i})^{-1}\right).$$

Proof: i) See [5].

ii) now follows directly from i) by letting $k \to \infty$ and applying proposition 2.1.5.

□

2.2 The Weighted ζ-Function

Since $w_k(I)$ is multiplicative in coprime ideals of A, it is natural for us to consider the corresponding Dirichlet series associated to each $k \geq 1$:

$$\zeta_{k,w}(s) := \sum_{I \neq 0} w_k(I)(NI)^{-s}$$

and likewise

$$\zeta_{\infty,w}(s) := \sum_{I \neq 0} w(I)(NI)^{-s}.$$

We will see below that both of these Dirichlet series converge in the half plane $\mathrm{Re}(s) > 0$. The notation $\zeta_{k,w}(s)$ indicates that this is a weighted ζ-function with respect to the weight $w_k(I)$ given to each ideal. By multiplicativity, we have the following Euler product formula for these ζ-functions in the half plane:

$$\zeta_{k,w}(s) = \prod_{\mathfrak{p} \text{ prime}} \left(\sum_{a=0}^{\infty} w_k(\mathfrak{p}^a) N\mathfrak{p}^{-a} \right).$$

$$\zeta_{\infty,w}(s) = \prod_{\mathfrak{p} \text{ prime}} \left(\sum_{a=0}^{\infty} w(\mathfrak{p}^a) N\mathfrak{p}^{-a} \right).$$

The task now is to find a product formula for each $\mathfrak{p}$-term in the product above. Consider the following Dirichlet series:

$$\sum_{a=0}^{\infty} w_k(\mathfrak{p}^j)(N\mathfrak{p})^{-as}.$$

To find a formula for this, we first must provide a recursive formula for $w_{k+1}(\mathfrak{p}^a)$ in terms of $w_k(\mathfrak{p}^a)$:

Proposition 2.2.1

$$w_{k+1}(\mathfrak{p}^a) = \sum_{j=0}^{a} (N\mathfrak{p})^{-jk} w_1(\mathfrak{p}^j) w_k(\mathfrak{p}^{a-j}).$$

Proof: Let G be an A-module with $\chi_A(G) = \mathfrak{p}^a$. We begin by writing

$$\#\{\varphi \in \mathrm{Hom}(A^{k+1}, G) : \varphi \text{ surjective}\} = \sum_{H \leq G} \#\{\varphi \in \mathrm{Hom}(A^{k+1}, G) : \varphi \text{ surjective and } \varphi(A^k) = H\}.$$

We can rewrite this as

$$\sum_{H\leq G}\sum_{\substack{\phi\in\mathrm{Hom}(A^k,H)\\ \phi \text{ surjective}}} \#\{\varphi\in\mathrm{Hom}(A^{k+1},G):\varphi \text{ surjective and } \varphi|_{A^k}=\phi\}.$$

For a fixed H and fixed surjective $\phi\in\mathrm{Hom}(A^k,H)$, we have the following correspondence between the two sets $\{\varphi\in\mathrm{Hom}(A^{k+1},G):\varphi \text{ surjective and } \varphi|_{A^k}=\phi\}$ and $\{\psi\in\mathrm{Hom}(A,G/H):\psi \text{ surjective}\}$. For each $\varphi: A^{k+1}\to G$ with $\varphi|_{A^k}=\phi$, we restrict φ to the last coordinate of A^{k+1} and then project to G/H. Each surjective $\psi: A\to G/H$ then arises from exactly $\#H$ choices of $\varphi: A^{k+1}\to G$ in this way. Thus

$$\#\{\varphi\in\mathrm{Hom}(A^{k+1},G):\varphi \text{ surjective and } \varphi|_{A^k}=\phi\}=(\#H)\#\{\psi\in\mathrm{Hom}(A,G/H):\psi \text{ surjective}\}.$$

Thus

$$\#\{\varphi\in\mathrm{Hom}(A^{k+1},G):\varphi \text{ surjective}\}=\sum_{H\leq G}(\#H)\sum_{\substack{\phi\in\mathrm{Hom}(A^k,H)\\ \phi \text{ surjective}}}\#\{\psi\in\mathrm{Hom}(A,G/H):\ \psi \text{ surjective}\}.$$

Now since $\{\psi\in\mathrm{Hom}(A,G/H):\psi \text{ surjective}\}$ is just $\#\mathrm{Aut}(G/H)\#(G/H)w_1(G/H)$ (and $\#(G/H)=\frac{\#G}{\#H}$) and likewise $\#\{\phi\in\mathrm{Hom}(A^k,H):\phi \text{ surjective}\}$ is just $\#\mathrm{Aut}(H)(\#H)^k w_k(H)$, we obtain

$$\#\{\varphi\in\mathrm{Hom}(A^{k+1},G):\varphi \text{ surjective}\}=\#G\sum_{H\leq G}\#\mathrm{Aut}(G/H)w_1(G/H)\#\mathrm{Aut}(H)w_k(H)(\#H)^k.$$

Thus, since $w_{k+1}(G)=(\#G)^{-(k+1)}\#\mathrm{Aut}(G)^{-1}\#\{\varphi\in\mathrm{Hom}(A^{k+1},G):\ \varphi \text{ surjective}\}$

$$w_{k+1}(G)=\frac{1}{\#\mathrm{Aut}(G)}\sum_{H\leq G}\#(G/H)^{-k}\#\mathrm{Aut}(G/H)w_1(G/H)\#\mathrm{Aut}(H)w_k(H).$$

So now we sum over isomorphism classes of G with $\chi_A(G)=\mathfrak{p}^a$:

$$w_{k+1}(\mathfrak{p}^a) = \sum_{G(\mathfrak{p}^a)} \frac{1}{\#\mathrm{Aut}(G)} \sum_{H \leq G} \#(G/H)^{-k} \#\mathrm{Aut}(G/H) w_1(G/H) \#\mathrm{Aut}(H) w_k(H).$$

We can interchange the order of summation and then rewrite this as a sum over A-ideals dividing $\mathfrak{p}^a$, writing $C = G/H$ and $K = H$ to obtain

$$\sum_{j=0}^{a} (N\mathfrak{p})^{-jk} \left(\sum_{C(\mathfrak{p}^j)} \#\mathrm{Aut}(C) w_1(C) \right) \left(\sum_{K(\mathfrak{p}^{a-j})} \mathrm{Aut}(K) w_k(K) \right)$$

$$\times \left(\sum_{G(\mathfrak{p}^a)} \frac{1}{\#\mathrm{Aut}(G)} \#\{H \leq G : H \cong K \text{ and } G/H \cong C\} \right).$$

We emphasize that the rightmost term in this sum depends upon the two independent sums over $C(\mathfrak{p}^j)$ and $K(\mathfrak{p}^{a-j})$. From a lemma in [5] we have the following

$$\left(\sum_{G(\mathfrak{p}^a)} \frac{1}{\#\mathrm{Aut}(G)} \#\{H \leq G : H \cong K \text{ and } G/H \cong C\} \right) = (\#\mathrm{Aut}(K))^{-1} (\#\mathrm{Aut}(C))^{-1}.$$

Thus

$$w_{k+1}(\mathfrak{p}^a) = \sum_{j=0}^{a} (N\mathfrak{p})^{-jk} \left(\sum_{C(\mathfrak{p}^j)} w_1(C) \right) \left(\sum_{K(\mathfrak{p}^{a-j})} w_k(K) \right)$$

$$= \sum_{j=0}^{a} (N\mathfrak{p})^{-jk} w_1(\mathfrak{p}^j) w_k(\mathfrak{p}^{a-j}).$$

□

With this recursive formula in hand, the product form of the Dirichlet series above is a corollary:

Corollary 2.2.2 *For any nonzero prime ideal $\mathfrak{p}$ of A and $Re(s) > -1$*

$$\sum_{a=0}^{\infty} w_k(\mathfrak{p}^a)(N\mathfrak{p})^{-as} = \prod_{j=1}^{k} (1 - N\mathfrak{p}^{-j-s})^{-1}$$

and

$$\sum_{a=0}^{\infty} w(\mathfrak{p}^a)(N\mathfrak{p})^{-as} = \prod_{j=1}^{\infty}(1 - N\mathfrak{p}^{-j-s})^{-1}.$$

Proof: We prove this by induction on k. The base case is $k = 1$.

$$w_1(\mathfrak{p}^a) = N\mathfrak{p}^{-a}.$$

$$\sum_{a=0}^{\infty} w_1(\mathfrak{p}^a)N\mathfrak{p}^{-as} = \sum_{a=0}^{\infty} N\mathfrak{p}^{-a(s+1)} = (1 - N\mathfrak{p}^{-(s+1)})^{-1}.$$

This is justified since $-(s + 1) < 0$.

Now suppose the induction hypothesis is true in the kth case.

$$\sum_{a=0}^{\infty} w_k(\mathfrak{p}^a)(N\mathfrak{p})^{-as} = \prod_{j=1}^{k}(1 - N\mathfrak{p}^{-j-s})^{-1}.$$

Now we apply proposition 2.2.1:

$$\sum_{a=0}^{\infty} w_{k+1}(\mathfrak{p}^a)(N\mathfrak{p})^{-as} = \sum_{a=0}^{\infty} \left(\sum_{j=0}^{a} (N\mathfrak{p})^{-jk} w_1(\mathfrak{p}^j) w_k(\mathfrak{p}^{a-j}) \right) N\mathfrak{p}^{-as}.$$

Applying proposition 2.1.6 we obtain:

$$w_1(\mathfrak{p}^j) = N\mathfrak{p}^{-j}.$$

Thus

$$\sum_{a=0}^{\infty} w_{k+1}(\mathfrak{p}^a)(N\mathfrak{p})^{-as} = \sum_{a=0}^{\infty} \left(\sum_{j=0}^{a} w_k(\mathfrak{p}^{a-j}) N\mathfrak{p}^{-j(k+1)} \right) N\mathfrak{p}^{-as}.$$

Now interchange the order of summation to obtain

$$\sum_{j=0}^{\infty} N\mathfrak{p}^{-j(k+1)} \sum_{a=j}^{\infty} w_k(\mathfrak{p}^{a-j}) N\mathfrak{p}^{-as} = \sum_{j=0}^{\infty} N\mathfrak{p}^{-j(k+1)} \sum_{r=0}^{\infty} w_k(\mathfrak{p}^r) N\mathfrak{p}^{-(r+j)s}$$

$$= \sum_{j=0}^{\infty} N\mathfrak{p}^{-j(k+1+s)} \sum_{r=0}^{\infty} w_k(\mathfrak{p}^r) N\mathfrak{p}^{-rs} = (1 - N\mathfrak{p}^{-(k+1+s)})^{-1} \sum_{r=0}^{\infty} w_k(\mathfrak{p}^r) N\mathfrak{p}^{-rs}$$

which, by our induction hypothesis, equals

$$(1 - N\mathfrak{p}^{-(k+1+s)})^{-1} \prod_{j=1}^{k} (1 - N\mathfrak{p}^{-j-s})^{-1} = \prod_{j=1}^{k+1} (1 - N\mathfrak{p}^{-(j+s)})^{-1}.$$

This completes the induction proof. Thus for all $k \geq 1$ we obtain

$$\sum_{a=0}^{\infty} w_k(\mathfrak{p}^a)(N\mathfrak{p})^{-as} = \prod_{j=1}^{k} (1 - N\mathfrak{p}^{-j-s})^{-1}.$$

Letting $k \to \infty$, absolute converge of the product justifies us taking the limit inside the sum to obtain

$$\sum_{a=0}^{\infty} w(\mathfrak{p}^a)(N\mathfrak{p})^{-as} = \prod_{j=1}^{\infty} (1 - N\mathfrak{p}^{-j-s})^{-1}.$$

□

From this we conclude the following:

Corollary 2.2.3 *For $Re(s) > 0$, the Dirichlet series $\zeta_{k,w}(s)$ and $\zeta_{\infty,w}(s)$ converge, and:*

i)

$$\zeta_{k,w}(s) = \prod_{j=1}^{k} \zeta_A(s+j)$$

ii)

$$\zeta_{\infty,w}(s) = \prod_{j=1}^{\infty} \zeta_A(s+j)$$

where $\zeta_A(s)$ is the usual Dedekind ζ-function of A.

Proof: Applying the product formula for $\zeta_{k,w}$ and $\zeta_{\infty,w}$ and the result above we obtain

$$\zeta_{k,w}(s) = \prod_{\mathfrak{p} \text{ prime}} \prod_{j=1}^{k} (1 - N\mathfrak{p}^{-j-s})^{-1} = \prod_{j=1}^{k} \zeta_A(s+j).$$

This clearly converges for $\mathrm{Re}(s) > 0$, since each $\zeta_A(s+j)$ converges for $j \geq 1$. The result for $\zeta_{\infty,w}(s)$ follows by letting $k \to \infty$.

□

Thus we have formulae for the weighted ζ-functions $\zeta_{w,k}$ and $\zeta_{w,\infty}$ purely in terms of the values of the usual Dedekind ζ-function ζ_A. Note that from 2.2.2 we obtain the result we claimed in the previous section, namely that

$$\sum_{G \in \mathcal{G}_{A,\mathfrak{p}}} \frac{1}{\#(\mathrm{Aut}(G))} < \infty$$

since this is just

$$\sum_{a=0}^{\infty} w(\mathfrak{p}^a) = \prod_{j=1}^{\infty} (1 - N\mathfrak{p}^{-j})^{-1}.$$

Now we continue with the goal of finding ζ-function formulae for $M_k(f)$ and $M_\infty(f)$.

Definition 2.2.4 *For $f : \mathcal{G}_A \to \mathbb{C}$ and $k \geq 1$ an integer and I a nonzero ideal of A we set*

$$w_k(f, I) = \sum_{G(I)} w_k(G) f(G).$$

Now we introduce the following ζ-functions weighted with respect to the function f:

Definition 2.2.5 *For $k \geq 1$ and $f : \mathcal{G}_A \to \mathbb{C}$ we set:*

$$\zeta_{f,k}(s) := \sum_{I \neq 0} w_k(f, I)(NI)^{-s}$$

and

$$\zeta_{f,\infty}(s)L = \sum_{I \neq 0} w(f,I)(NI)^{-s}.$$

This is a purely formal notion, since we cannot guarantee the convergence of $\zeta_{k,f}$ or $\zeta_{\infty,f}$ anywhere in $\mathbb{C}$ for an arbitrary $f : \mathcal{G}_A \to \mathbb{C}$, but we will be interested in sufficiently nice functions f that their corresponding ζ-functions $\zeta_{f,k}$ and $\zeta_{f,\infty}$ converge in the half-plane.

Affer some work we arrive at the following ζ-function formula for the k-average of a function $f : \mathcal{G}_A \to \mathbb{C}$

Proposition 2.2.6 *For any $k \geq 1$ and $f : \mathcal{G}_A \to \mathbb{C}$, if $M_k(f)$ exists then the following holds:*

$$M_k(f) = \frac{\zeta_{f,k}(1)}{\zeta_{w,k}(1)}$$

and likewise if $M(f)$ exists then

$$M(f) = \frac{\zeta_{f,\infty}(1)}{\zeta_{w,\infty}(1)}.$$

Proof: See [5].

□

We can now use our formula for $M(f)$ to calculate the averages of functions $f : \mathcal{G}_A \to \mathbb{C}$ with respect to the weighting $w(G)$. Some examples are done below:

Proposition 2.2.7 *Let $\mathcal{P} \subseteq \mathrm{Spec}(A)$, Let L be a fixed finite $\mathcal{P}A$-module, and let Q denote the property that $\#G_{\mathcal{P}} \cong L$ as an A-module. Let f_Q denote the corresponding indicator function. Then the CL-probability that G satisfies Q, which is $M(f_Q)$, is given by*

$$\frac{1}{\#L\#(\mathrm{Aut}(L))} \prod_{\mathfrak{p} \in \mathcal{P}} \prod_{j=2}^{\infty} (1 - N\mathfrak{p}^{-j}).$$

Proof: We calculate the coefficients of $\zeta_{f_Q,\infty}(s)$. We have

$$w(f_Q, I) = \sum_{\substack{G(I) \\ G(Q)}} w(G)$$

where the sum $\sum_{\substack{G(I) \\ G(Q)}}$ is just taken over all G with $\chi_A(G) = I$ which satisfy property Q. Let J_1 denote the $\mathcal{P}$-part of the ideal I and J_2 the coprime-to-$\mathcal{P}$ part of I. Thus by multiplicativity:

$$w(f_Q, I) = \left(\sum_{H(J_2)} w(H) \right) \left(\sum_{\substack{K(J_1) \\ K(f_Q)}} w(K) \right).$$

Now note

$$\sum_{\substack{K(J_1) \\ K(f_Q)}} w(K) = \begin{cases} w(L) \text{ if } \chi_A(L) = J_1 \\ 0 \text{ otherwise} \end{cases}.$$

Thus

$$\sum_{I \neq 0} w(f_Q, I)(NI)^{-s} = \left(\sum_{\substack{J \neq 0 \\ J \text{ prime to } \mathcal{P}}} w(J)(NJ)^{-s} \right) w(L)(N\chi_A(L))^{-s}.$$

Thus

$$\zeta_{f_Q, \infty}(1) = \left(\sum_{\substack{J \neq 0 \\ J \text{ prime to } \mathcal{P}}} w(J)(NJ)^{-1} \right) w(L)(N\chi_A(L))^{-1}.$$

Whereas, again apply multiplicativity:

$$\zeta_{w,\infty}(1) = \sum_{I \neq 0} w(I)(NI)^{-1} = \left(\sum_{\substack{I \neq 0 \\ I \text{ prime to } \mathcal{P}}} w(I)(NI)^{-1} \right) \left(\sum_{\substack{I \neq 0 \\ \mathfrak{p} \mid I \to \mathfrak{p} \in \mathcal{P}}} w(I)(NI)^{-1} \right).$$

The rightmost term is simply the sum taken over all ideals I whose prime factors lie among $\mathcal{P}$.

For the rightmost term we apply multiplicativity one more time and write the sum as a product:

$$\sum_{\substack{I\neq 0\\ \mathfrak{p}\,|\,I\to\mathfrak{p}\in\mathcal{P}}} w(I)(NI)^{-1} = \prod_{\mathfrak{p}\in\mathcal{P}}\left(\sum_{a=0}^{\infty} w(\mathfrak{p}^a)N\mathfrak{p}^{-a}\right).$$

Now we just apply corollary 2.2.2:

$$\sum_{\substack{I\neq 0\\ \mathfrak{p}\,|\,I\to\mathfrak{p}\in\mathcal{P}}} w(I)(NI)^{-1} = \prod_{\mathfrak{p}\in\mathcal{P}}\prod_{j=1}^{\infty}(1-N\mathfrak{p}^{-j-1})^{-1}.$$

Thus

$$\frac{\zeta_{f_Q,\infty}(1)}{\zeta_{w,\infty}(1)} = \frac{w(L)}{N(\chi_A(L))}\prod_{\mathfrak{p}\in\mathcal{P}}\prod_{j=1}^{\infty}(1-N\mathfrak{p}^{-j-1})$$

$$= \frac{1}{\#L\#(\mathrm{Aut}(L))}\prod_{\mathfrak{p}\in\mathcal{P}}\prod_{j=1}^{\infty}(1-N\mathfrak{p}^{-j-1}).$$

□

This calculation demonstrates the usefulness of having a ζ-function expression for the average of a function f with respect to this probability measure. We do one more example:

Proposition 2.2.8 *Fix a nonzero prime $\mathfrak{p}$ of A. The CL-probability that $G_{\mathfrak{p}} \neq 0$ is*

$$1-\prod_{j=2}^{\infty}(1-N\mathfrak{p}^{-j}).$$

Proof: Let Q denote the property that $G_{\mathfrak{p}} = 0$ and let f_Q denote the corresponding indicator function. As before, we calculate the coefficients of $\zeta_{f_Q,\infty}(s)$.

$$w(f_Q, I) = \sum_{G(I)} f(G)w(G).$$

Let J denote the prime-to-$\mathfrak{p}$ part of I and $\mathfrak{p}^m$ the $\mathfrak{p}$ part of I (where k might be zero), so by multiplicativity we obtain

$$w(f_Q, I) = \left(\sum_{H(J)} w(H) \right) \left(\sum_{K(\mathfrak{p}^m)} f(K)w(K) \right).$$

Of course, the rightmost term is just

$$\sum_{K(\mathfrak{p}^m)} f(K)w(K) = \begin{cases} 1 \text{ if } m = 0 \\ \\ 0 \text{ otherwise} \end{cases}.$$

Thus

$$\zeta_{f_Q,\infty}(1) = \sum_{I \neq 0} w(f_Q, I)(NI)^{-1} = \sum_{\substack{I \neq 0 \\ \mathfrak{p} \nmid I}} w(I)(NI)^{-1}.$$

Whereas, applying multiplicativity of $w(I)$ and corollary 2.2.2

$$\zeta_{w,\infty}(1) = \left(\sum_{\substack{I \neq 0 \\ \mathfrak{p} \nmid I}} w(I)(NI)^{-1} \right) \left(\sum_{j=0}^{\infty} w(\mathfrak{p}^j) N\mathfrak{p}^{-j} \right)$$

$$= \left(\sum_{\substack{I \neq 0 \\ \mathfrak{p} \nmid I}} w(I)(NI)^{-1} \right) \left(\prod_{j=1}^{\infty} (1 - N\mathfrak{p}^{-j-1})^{-1} \right).$$

Thus

$$\frac{\zeta_{f_Q,\infty}(1)}{\zeta_{w,\infty}(1)} = \prod_{j=1}^{\infty} (1 - N\mathfrak{p}^{-j-1}).$$

□

With the tools of the previous section in hand, we will see how we can make calculations about the behavior of the odd parts of the real quadratic class groups. We introduce the heuristic assumption in the next section.

2.3 Applying The Heuristic Assumption of the Model

In order to introduce the heuristic assumption of the Cohen-Lenstra model, we will need to introduce some notions from commutative algebra. To each abelian group Γ of order n, we associate a ring A_Γ which is a finite product of rings of integers of number fields. Consider the $\mathbb{Q}$-algebra $\mathbb{Q}[\Gamma]/(\sum_{g\in\Gamma} g)$. The group ring $\mathbb{Q}[\Gamma]$ is a semisimple $\mathbb{Q}$-algebra by Maschke's theorem. Thus by Artin-Wedderburn, $\mathbb{Q}[\Gamma]$ is (as a $\mathbb{Q}$-algebra) isomorphic to a product of the form $R_1 \times \cdots \times R_k$, where each R_i is a $\mathbb{Q}$-algebra of the form $M_{n_i}(\Delta_i)$, where Δ_i is a division ring over $\mathbb{Q}$ and $M_{n_i}(\Delta_i)$ is the $\mathbb{Q}$-algebra of all $n_i \times n_i$ matrices over Δ_i. Of course, since $\mathbb{Q}[\Gamma]$ is commutative, this means Δ_i is a number field and each $n_i = 1$, i.e $\mathbb{Q}[\Gamma]$ is a finite product of number fields.

$\sum_{g\in\Gamma} g$ is a primitive central idempotent of $\mathbb{Q}[\Gamma]$, so under the Wedderburn decomposition of $\mathbb{Q}[\Gamma]$ it corresponds to a tuple having 1 in one position, say the ith, and 0 in all other positions, and so $\mathbb{Q}[\Gamma]/\left(\sum_{g\in\Gamma} g\right)$ is again a finite product of number fields. If Γ is the Galois group of an abelian extension $K/\mathbb{Q}$, then $\mathrm{Cl}(K)$ is a Γ-module in the usual way. The idempotent $\sum_{g\in\Gamma} g$ annihlates the prime-to-$|\Gamma|$ part of $\mathrm{Cl}(K)$, so $\mathrm{Cl}(K)$ is a $\mathbb{Z}[\Gamma]/\left(\sum_{g\in\Gamma} g\right)$-module.

Now we introduce the following definitions:

Definition 2.3.1 Let V be a finite-dimensional $\mathbb{Q}$-vector space. Then we define a *full $\mathbb{Z}$-lattice* in V to be a finitely generated $\mathbb{Z}$-submodule M of V such that $\mathbb{Q}M = V$, where $\mathbb{Q}M$ is the set of all finite $\mathbb{Q}$-linear combinations of M.

Definition 2.3.2 Let A be a $\mathbb{Q}$-algebra. A *$\mathbb{Z}$-order* in A is a unital subring Λ of A such that Λ is a full $\mathbb{Z}$-lattice in A. A *maximal* $\mathbb{Z}$-order in A is a $\mathbb{Z}$-order not properly contained in any other $\mathbb{Z}$-order of A.

We will now use the theory of $\mathbb{Z}$-orders to develop the Cohen-Lenstra model. We want to show that $\mathbb{Q}[\Gamma]/\left(\sum_{g\in\Gamma} g\right)$ has a unique maximal $\mathbb{Z}$-order, and that this unique maximal $\mathbb{Z}$-order is a finite product of number rings. We first note that if Λ is a $\mathbb{Z}$-order in a $\mathbb{Q}$-algebra A, then every element of Λ is integral over $\mathbb{Z}$. In particular, if K is a number field, then the number ring $\mathcal{O}_K$ is the unique maximal $\mathbb{Z}$-order in K. For our case, as $\mathbb{Q}[\Gamma]/\left(\sum_{g\in\Gamma} g\right)$ is a finite product of number fields, the unique maximal $\mathbb{Z}$-order in $\mathbb{Q}[\Gamma]/\left(\sum_{g\in\Gamma} g\right)$ is just the corresponding product of number rings.

Thus we conclude that for Γ a finite abelian group, there is a unique maximal $\mathbb{Z}$-order in $\mathbb{Q}[\Gamma]/\left(\sum_{g\in\Gamma} g\right)$ and this unique maximal $\mathbb{Z}$-order is a finite product of number rings. Let A_Γ denote this unique maximal $\mathbb{Z}$-order. Let Γ be an abelian group of N, and $r_1, r_2 \geq 0$ integers with $r_1 + 2r_2 = N$, let $\mathcal{F}_{N,r_1,r_2}$ denote

the family of abelian extensions of $\mathbb{Q}$ with Galois group Γ, r_1 real embeddings and r_2 pairs of complex embeddings up to conjugation.

We can order the fields in $\mathcal{F}_{N,r_1,r_2}$ by the absolute value of their discriminants, where the order among the finite list of fields in $\mathcal{F}_{N,r_1,r_2}$ having the same discriminant d does not matter. For $K \in \mathcal{F}_{N,r_1,r_2}$ let $H(K)$ denote the prime-to-N part of the class group $\mathrm{Cl}(K)$. Then, as indicated above, $H(K)$ is an A_Γ-module. For f a complex-valued function defined on the set of isomorphism classes of finite A_Γ-modules, we set

$$E(f) = \lim_{X\to\infty} \frac{\sum_{\substack{K\in\mathcal{F}_{\Gamma,r_1,r_2} \\ |\mathrm{disc}(K)|\leq X}} f(H(K)))}{\sum_{\substack{K\in\mathcal{F}_{\Gamma,r_1,r_2} \\ |\mathrm{disc}(K)|\leq X}} 1}.$$

That is, $E(f)$ is just the usual notion of the average value of $f(H(K))$ taken over all the fields in $K \in \mathcal{F}_{\Gamma,r_1,r_2}$.

Before introducing the heuristic assumption of the Cohen-Lenstra model, we need one more definition. Previously, we introduced the definition of the CL-average of a function $f : \mathcal{G}_A \to \mathbb{C}$, where $\mathcal{G}_A$ is the category of finite modules over a fixed number A. To state the heuristic assumption, we must extend this definition to functions defined on finite modules over finite products of number rings.

Definition 2.3.3 Let $A_1, \cdots, A_n$ be number rings and let $\mathbf{A} = \bigoplus_{i=1}^n A_i$ and $\mathcal{G}_\mathbf{A}$ denote the category of finite $\mathbf{A}$-modules. Let $f : \mathcal{G}_\mathbf{A} \to \mathbb{C}$ and for each $1 \leq i \leq n$ let $f_i : A_i \to \mathbb{C}$ be the corresponding coordinate function. Then the *CL-average* of f, denoted $M_\mathbf{A}(f)$ is defined to be

$$M_\mathbf{A}(f) := \prod_{i=1}^n M_{A_i}(f_i)$$

provided that each $M_{A_i}(f_i)$ exists.

We now have enough definitions in hand to introduce the *Cohen-Lenstra heuristic*.

Definition 2.3.4 Let $\mathcal{F}_{\Gamma,N,0}$ be the family of all totally real extensions $K/\mathbb{Q}$ of degree N (that is, $r_2 = 0$ and $r_1 = N$) having Galois group Γ, where Γ is an abelian group of order N. Let $\mathcal{A}_\Gamma$ be the category of all finite A_Γ-modules. Then the *Cohen-Lenstra heuristic* states that for all reasonably-behaved functions $f : \mathcal{A}_\Gamma \to \mathbb{C}$, $E(f)$ is equal to the CL-average $M_{A_\Gamma}(f)$ of f.

Intuitively, this heuristic assumption says that we should expect abelian groups with more automor-

phisms to be less likely to appear as class groups of abelian extensions of $\mathbb{Q}$. At the end of this chapter, we will give some justification for this heuristic, and make more precise the notion of "reasonably-behaved functions".

In the quadratic case, where $\Gamma = \mathbb{Z}/2\mathbb{Z}$, the ring $\mathbb{Q}[\Gamma]$ is just $\mathbb{Q}[x]/(x^2-1) \cong \mathbb{Q} \oplus \mathbb{Q}$, and so $\mathbb{Q}[\Gamma]/(\sum_{g\in\Gamma} g) \cong \mathbb{Q}$ so the corresponding maximal order A_Γ is just $\mathbb{Z}$.

The heuristic assumption implies that if f is the characteristic function of some property P of finite A_Γ-modules, then $M(f)$ is equal to the CL-probability of P. We note from the definition above that in this case, $M(f)$ is the probability that a field $K \in \mathcal{F}_{\Gamma,r_1,r_2}$ selected at random is such that the A_Γ-module $H(K)$ has property P.

Thus, for example, the heuristic assumption of the model, together with proposition 2.2.7 of the previous section, imply that the probability that a real quadratic field selected at random has a trivial odd part is

$$2\left(\prod_{j=1}^{\infty}(1-2^{-j})^{-1}\right)\left(\prod_{p \text{ prime}}\prod_{j=2}^{\infty}(1-p^{-j})\right) = 2\left(\prod_{j=1}^{\infty}(1-2^{-j})^{-1}\right)\prod_{j=2}^{\infty}\zeta(j)^{-1} \approx 0.75446.$$

We can generalize this to the following:

Proposition 2.3.5 *Let ℓ be an odd positive integer and L be an abelian group of order ℓ. Then the probability that a real quadratic field selected at random has odd part isomorphic to L is*

$$\frac{1}{2\ell(\#\mathrm{Aut}(L))}\prod_{j=2}^{\infty}\zeta(j)^{-1}\prod_{j=2}^{\infty}(1-2^{-j})^{-1}.$$

Proof: Apply proposition 2.2.7 to the ring $A = \mathbb{Z}$ and the set of primes $\mathrm{Spec}(\mathbb{Z}) \setminus \{2\}$.

□

In particular, if p is an odd prime, then the probability that a a real quadratic field selected at random has odd part of cardinality p is

$$\frac{1}{2p(p-1)}\prod_{j=2}^{\infty}\zeta(j)^{-1}\prod_{j=2}^{\infty}(1-2^{-j})^{-1} \approx \frac{0.75446}{2p(p-1)}.$$

We can also compute the probability that a real quadratic field selected at random has a class group with a cyclic odd part.

Proposition 2.3.6 *Let A be a number ring and let $\mathcal{P} \subseteq \mathrm{Spec}(A)$. Let Q be the property of a finite A-module G that the $\mathcal{P}$-part of G is cyclic. Then the CL-probability of Q is*

$$\prod_{\mathfrak{p}\in\mathcal{P}} \frac{1+(N\mathfrak{p})^{-1}-(N\mathfrak{p})^{-3}}{(1-(N\mathfrak{p})^{-2})(1-(N\mathfrak{p})^{-1})} \prod_{j=2}^{\infty}(1-N\mathfrak{p}^{-j}).$$

Proof: Let f_Q be the indicator function corresponding to Q. As before, we calculute the coefficients of $\zeta_{f_Q,\infty}(s)$.

$$w(f_Q, I) = \sum_{\substack{G(I)\\ G(Q}} w(G).$$

Here, the sum $\sum_{\substack{G(I)\\ G(Q)}}$ is taken over all finite A-modules G with $\chi_A(G) = I$ which satisfy property Q. Let J_1 denote the $\mathcal{P}$-part of the ideal I and J_2 the coprime-to-$\mathcal{P}$ part of I. Thus by multiplicativity

$$w(f_Q, I) = \left(\sum_{H(J_2)} w(H)\right)\left(\sum_{\substack{K(J_1)\\ K(Q)}} w(K)\right).$$

Now note

$$\sum_{\substack{K(J_1)\\ K(Q)}} w(K) = w(A/J_1) = \frac{1}{\#(\mathrm{Aut}(A/J_1))} = \frac{1}{\#((A/J_1)^*)}.$$

Here, $(A/J_1)^*$ is the multiplicative group of A/J_1.

Thus

$$\zeta_{f_Q,\infty}(1) = \sum_{I\neq 0} w(f_Q, I)(NI)^{-1} = \left(\sum_{\substack{I\neq 0\\ I \text{ prime to } \mathcal{P}}} w(I)(NI)^{-1}\right)\left(\sum_{\substack{I\neq 0\\ \mathfrak{p}\mid I\to\mathfrak{p}\in\mathcal{P}}} (\#(A/I)^*)^{-1}(NI)^{-1}\right).$$

Whereas

$$\zeta_{w,\infty}(1) = \sum_{I \neq 0} w(I)(NI)^{-1} = \left(\sum_{\substack{I \neq 0 \\ I \text{ prime to } \mathcal{P}}} w(I)(NI)^{-1} \right) \left(\sum_{\substack{I \neq 0 \\ \mathfrak{p} \mid I \to \mathfrak{p} \in \mathcal{P}}} w(I)(NI)^{-1} \right).$$

Now, for I, J coprime ideals, we note that $\#(A/IJ)^* = \#(A/I)^* \#(A/J)^*$ by the Chinese Remainder theorem. Thus, we may write $\sum_{\substack{I \neq 0 \\ \mathfrak{p} \mid I \to \mathfrak{p} \in \mathcal{P}}} (\#(A/I)^*)^{-1}(NI)^{-1}$ as an Euler product:

$$\sum_{\substack{I \neq 0 \\ \mathfrak{p} \mid I \to \mathfrak{p} \in \mathcal{P}}} (\#(A/I)^*)^{-1}(NI)^{-1} = \prod_{\mathfrak{p} \in \mathcal{P}} \left(\sum_{m=0}^{\infty} (\#(A/\mathfrak{p}^m)^*)^{-1} N\mathfrak{p}^{-m} \right).$$

To determine $\sum_{m=0}^{\infty} (\#(A/\mathfrak{p}^m)^*)^{-1} N\mathfrak{p}^{-m}$, we note that we can calculuate $\#(A/\mathfrak{p}^m)$ with the generalized Euler totient function φ_A (where $\varphi_A(I)$ gives the number of invertible elements of A/I. We first note that $(A/\mathfrak{p})^*$ has a cardinality of $N\mathfrak{p}-1$, as $A/\mathfrak{p}$ is a field, and furthermore, the projection map $(A/\mathfrak{p}^m) \to A/\mathfrak{p}$ is such that $a + \mathfrak{p}^m \in (A/\mathfrak{p}^m)^*$ if and only if $a + \mathfrak{p} \in (A/\mathfrak{p})^*$. Furthermore, each $a + \mathfrak{p} \in (A/\mathfrak{p})^*$ has exactly $N\mathfrak{p}^{m-1}$ preimages in $(A/\mathfrak{p}^m)^*$. Thus we conclude that

$$\varphi_A(\mathfrak{p}^m) = N\mathfrak{p}^{m-1}(N\mathfrak{p} - 1)$$

which is what we expect. We need to be somewhat careful here. This formula is valid if $m \geq 1$. If $m = 0$ then of course the trivial A-module has an automorphism group of size 1. We take this into account below.

Thus

$$\sum_{m=0}^{\infty} (\#(A/\mathfrak{p}^m)^*)^{-1} N\mathfrak{p}^{-m} = 1 + (N\mathfrak{p} - 1)^{-1} \sum_{m=1}^{\infty} N\mathfrak{p}^{1-2m}$$

$$= 1 + \frac{N\mathfrak{p}}{(N\mathfrak{p} - 1)} \frac{N\mathfrak{p}^{-2}}{(1 - N\mathfrak{p}^{-2})} = 1 + \frac{N\mathfrak{p}^{-2}}{(1 - (N\mathfrak{p})^{-1})(1 - (N\mathfrak{p})^{-2})}$$

$$= \frac{(1 - (N\mathfrak{p})^{-1})(1 - (N\mathfrak{p})^{-2}) + N\mathfrak{p}^{-2}}{(1 - (N\mathfrak{p})^{-1})(1 - (N\mathfrak{p})^{-2})} = \frac{1 - (N\mathfrak{p})^{-1} + (N\mathfrak{p})^{-3}}{(1 - (N\mathfrak{p})^{-1})(1 - (N\mathfrak{p})^{-2})}.$$

$$\frac{\zeta_{f_Q,\infty}(1)}{\zeta_{w,\infty}(1)} = \prod_{\mathfrak{p}\in\mathcal{P}} \frac{1-(N\mathfrak{p})^{-1}+(N\mathfrak{p})^{-3}}{(1-(N\mathfrak{p})^{-1})(1-(N\mathfrak{p})^{-2})} \prod_{j=1}^{\infty}(1-N\mathfrak{p}^{-j-1}).$$

Here, we have again used the fact that

$$\sum_{\substack{I\neq 0 \\ I \text{ prime to } \mathcal{P}}} w(I)(NI)^{-1} = \prod_{\mathfrak{p}\notin\mathcal{P}} \prod_{j=1}^{\infty}(1-N\mathfrak{p}^{-j-1})$$

and likewise for the set of ideals I with all prime divisors in $\mathcal{P}$. Thus, the CL-probability that a finite A-module has cyclic $\mathcal{P}$-part is

$$\prod_{\mathfrak{p}\in\mathcal{P}} \frac{1-(N\mathfrak{p})^{-1}+(N\mathfrak{p})^{-3}}{(1-(N\mathfrak{p})^{-1})(1-(N\mathfrak{p})^{-2})} \prod_{j=2}^{\infty}(1-N\mathfrak{p}^{-j}).$$

□

Thus, in the real quadratic case, where $A_\Gamma = \mathbb{Z}$, we see that the probability of selecting a real quadratic field at random which has a class group with cyclic odd part is

$$\frac{\frac{1}{2}\cdot\frac{3}{4}}{1-\frac{1}{2}+\frac{1}{8}} \prod_{j=2}^{\infty}(1-2^{-j})^{-1} \prod_{p \text{ prime}} \frac{1-p^{-1}+p^{-3}}{(1-p^{-1})(1-p^{-2})} \prod_{j=2}^{\infty}(1-p^{-j})$$

$$= \frac{3}{5} \prod_{j=2}^{\infty}(1-2^{-j})^{-1} \prod_{j=2}^{\infty}\zeta(j)^{-1} \prod_{p \text{ prime}} \frac{1-p^{-1}+p^{-3}}{(1-p^{-1})(1-p^{-2})}.$$

$$\prod_{p \text{ prime}} \frac{1-p^{-1}+p^{-3}}{(1-p^{-1})(1-p^{-2})} = \zeta(2) \prod_{p \text{ prime}} \left(1+\frac{p^{-3}}{1-p^{-1}}\right).$$

So the probability that a real quadratic calss field selected at random will be cyclic (i.e have rank ≤ 1) is

$$\frac{3\cdot\zeta(2)}{5} \left(\prod_{j=2}^{\infty}\zeta(j)^{-1}\right) \left(\prod_{j=2}^{\infty}(1-2^{-j})^{-1}\right) \left(\prod_{p \text{ prime}} \left(1+\frac{1}{p^3(1-p^{-1})}\right)\right)$$

which is approximately 0.33267.

We can go further than this and ask what is the probability that a real quadratic class group selected at random will have rank r, but producing a formula like the one in 2.3.3 is not possible in the general case. The local version of this question is answerable, however. Fix a prime p. Then we can find the probability that a real quadratic class group selected at random has a p-rank of r.

Proposition 2.3.7 *For a number ring A and a fixed prime $\mathfrak{p}$ of A and an integer $r \geq 0$, the CL-probability that a finite A-module G has $r_{\mathfrak{p}}(G) = r$ is*

$$N\mathfrak{p}^{-r(r+1)} \left(\prod_{j=r+1}^{\infty} 1 - N\mathfrak{p}^{-j} \right) \left(\prod_{j=1}^{r+1} (1 - N\mathfrak{p}^{-j})^{-1} \right).$$

Proof: See [5].

□

2.4 Justification for the Cohen-Lenstra Heuristic

In order to give some heuristic justification for 2.3.3, it is convenient to pass between the language of ideal classes and the language of binary quadratic forms to describe the class group of a number field. Recall from the first section that a *binary quadratic form* is a homogeneous polynomial $Q(x, y) = ax^2+bxy+cy^2$. We will pass between the quadratic form $ax^2+bxy+cy^2$ and the tuple of coefficients (a, b, c). An *integral quadratic form* is a binary quadratic form with integer coefficients. An integral quadratic form (a, b, c) is called *primitive* if a, b, c are coprime. From now on in this section, all binary quadratic forms unless otherwise stated will be integral quadratic forms.

Definition 2.4.1 Let $P(X, Y)$ and $Q(A, B)$ be two binary quadratic forms. Then $P(X, Y)$ and $Q(A, B)$ are called *equivalent* if there exists a linear transformation $T \in \mathrm{SL}_2(\mathbb{Z})$ such that $P(X, Y) = Q(TX, TY)$. That is, there are m, n, p, q such that $P(X, Y) = Q(mX + nY, pX + qY)$ and $mq - np = \pm 1$.

It is clear that this defines an equivalence relation on the binary quadratic forms. It is also clear that this defines a group action of $\mathrm{SL}_2(\mathbb{Z})$ onto $\{Q(x, y) \in \mathbb{Z}[X, Y] : \deg(Q) = 2,\ Q, \text{homogeneous}\}$ where $(T, Q(X, Y)) \to Q(TX, TY)$.

Recall that a binary quadratic form $Q(x, y) = ax^2 + bxy + cy^2$ has *discriminant* $D(Q) = b^2 - 4ac$. Clearly, for any binary quadratic form Q, $D(Q) \equiv 0 \mod 4$ or $D(Q) \equiv 1 \mod 4$. Furthermore, it

is clear that any integer a which is 0 or 1 mod 4 is of the form $D(Q)$, since a differs from a perfect square by a multiple of 4. It is also clear that two equivalent binary quadratic forms represent the same integer.

Proposition 2.4.2 *Let $P(X,Y)$ and $Q(A,B)$ be two equivalent binary quadratic forms. Then $D(P) = D(Q)$.*

Proof: Suppose that $P(X,Y) = Q(mX+nY, pX+qY)$, where m,n,p,q are integers with $mq-np = \pm 1$. Suppose that $Q(X,Y) = aX^2+bXY+cY^2$. Thus

$$P(X,Y) = a(mX+nY)^2 + b(mX+nY)(pX+qY) + c(pX+qY)^2$$

$$= (am^2+bmp+cp^2)X^2 + (2amn+bmq+bnp+2cpq)XY + (an^2+bnq+cq^2)Y^2.$$

Then

$$D(Q) = (2amn+bmq+bnp+2cpq)^2 - 4(am^2+bmp+cp^2)(an^2+bnq+cq^2)$$

$$= (b^2-4ac)(mq-np)^2 = b^2-4ac$$

$$= D(P).$$

□

Thus, we may speak of the discriminant of an equivalence class of binary quadratic forms (i.e a $\mathrm{SL}_2(\mathbb{Z})$-orbit of $\{Q(x,y) \in \mathbb{Z}[X,Y] : \deg(Q) = 2,\ Q\text{homogeneous}\}$).

Gauss defined an operation called *composition* on the set of equivalence calsses of binary quadratic forms having a fixed discriminant D. The deepest insight of Gauss was that for any discriminant D, there are only finitely many equivalence classes of binary quadratic forms representing D, and this finite family of equivalence classes forms a group under composition. This group is isomorphic to the class group of $\mathbb{Q}(\sqrt{D})$.

Definition 2.4.3 A binary quadratic form (a, b, c) of discriminant $D < 0$ is called reduced if the following conditions hold:

$$\begin{cases} |\sqrt{D} - 2|a|| < b < \sqrt{D} \text{ if } D > 0 \\ \\ \begin{cases} |b| \leq a \leq c \\ b \geq 0 \text{ if } |b| = a \text{ or } a = c \end{cases} \text{ if } D < 0 \end{cases}.$$

For a fixed discriminant D, we define $\mathcal{R}_D$ to be the set of reduced forms of discriminant D. An element of $\mathcal{R}_D$ may be written as (a, b), since $c = \frac{b^2 - D}{4a}$.

Proposition 2.4.4 *$\mathcal{R}_D$ has only finitely many elements.*

Proof: Let (a, b, c) be a form of discriminant D.

We first do the case where $D > 0$. Suppose that $|\sqrt{D} - 2|a|| < b < \sqrt{D}$. Thus $(\sqrt{D} - 2|a|)^2 < D$ and so $D + 4|a|^2 - 4|a|\sqrt{D} < D$. Thus $4|a|^2 < 4|a|\sqrt{D}$, and so $|a| < \sqrt{D}$. From this it follows immediately that $\mathcal{R}_D$ is finite, since there are only finitely many choices of a, and for each a there are only finitely many choices of b, and for each choice of a, b there is at most one choice of c. However, more is true in this case. We have

$$D + 4|a|^2 - 4|a|\sqrt{D} < b^2 < D.$$

Subtracting D from all three terms we get

$$4|a|^2 - 4|a|\sqrt{D} < 4ac < 0.$$

So taking absolute values and reversing the inequalities

$$4|a|\sqrt{D} - 4|a|^2 > 4|a||c|.$$

Thus

$$\sqrt{D} > |c| + |a|.$$

Now we do the case where $D < 0$. In this case, $0 \leq a \leq c$. Since $D = b^2 - 4ac$, we have $D \leq b^2 - 4a^2$. Since $b^2 \leq a^2$, we have $D \leq -3a^2$, we have $|D|/3 \geq a^2$, so there are only finitely many choices of a and thus, only finitely many choices of b. So again there are only finitely many elements of $\mathcal{R}_D$. $\square$

Gauss devised a *reduction algorithm* to transfer any binary quadratic form into an equivalent reduced form. The reduction Algorithm for Binary Quadratic Forms is as follows.

Let $P(X, Y) = aX^2 + bXY + cY^2$ be a binary quadratic form of discriminant D. Appy to $P(X, Y)$ an element $T \in \mathrm{SL}_2(\mathbb{Z})$ of the following form:

$$T_m = \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}$$

where $m \in \mathbb{Z}$.

Note that applying T to $P(X, Y)$ produces the equivalent binary quadratic form

$$P(X + mY, Y) = a(X + mY)^2 + b(X + mY)Y + cY^2$$

$$= aX^2 + (b + 2am)XY + (c + bm + m^2)Y^2.$$

If we denote this new binary quadratic form by (a', b', c'), then we note that we can choose b' to be in any real interval of length $2|a|$ by an appropriate choice of $m \in \mathbb{Z}$, since the translates are all the points differing from b by an integer multiple of $2a$.

We can also apply to $P(X, Y)$ the element $S \in \mathrm{SL}_2(\mathbb{Z})$

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

When we apply S to $P(X,Y)$, we get

$$P(-Y,X) = aY^2 - bXY + cX^2.$$

Thus, applying S to (a,b,c) we obtain the equivalent binary quadratic form $(c,-b,a)$.

The algorithm consists of the following steps:

1. Start with an integer binary quadratic form (a,b,c) of discriminant D.

2. If $D<0$ then choose an $m \in \mathbb{Z}$ so that when T_m is applied to (a,b,c), the resulting binary quadratic form (a',b',c') has $b' \in \{x \in \mathbb{R} : -|a| < x \leq |a|\}$.

If $D>0$ and $0<|a|<\sqrt{D}$, choose an $m \in \mathbb{Z}$ so that when T_m is applied to (a,b,c), the resulting binary quadratic form (a',b',c') has $b' \in \{x \in \mathbb{R} : -|a| < x < |a|\}$.

If $D>0$ and $a=0$, then first apply S to (a,b,c) to get $(c,-b,a)$ and then choose $m \in \mathbb{Z}$ so that when T_m is applied to $(c,-b,a)$ the resulting binary quadratic form (a',b',c') has $b' \in \{x \in \mathbb{R} : -|c| < x < |c|\}$, unless $c=0$ as well. If $c=0$ as well, then first apply T_1 to $(0,b,0)$ to get $(0,b,b+1)$ and then redo the above.

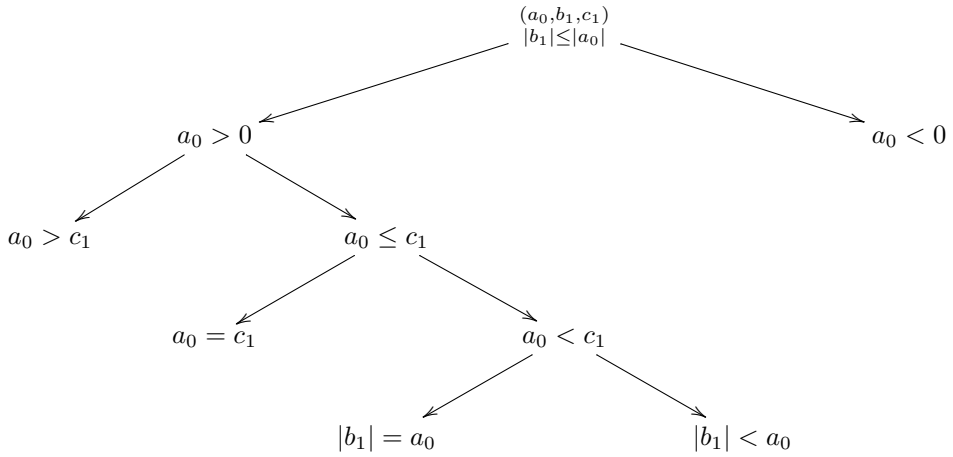
If $D>0$ and $|a| \geq \sqrt{D}$, choose an $m \in \mathbb{Z}$ so that when T_m is applied to (a,b,c), the resulting binary quadratic form (a',b',c') has $b' \in \{x \in \mathbb{R} : \sqrt{D}-2|a| < x < \sqrt{D}\}$.

3. After step 2, we have obtained a binary quadratic form (a',b',c'). If (a',b',c') is reduced, we are done. If not, replace (a',b',c') with $(c',b',-a')$ using S, and go back to step 2, now applying step 2 to $(c',b',-a')$.

Now we can show that the algorithm does terminate with the desired output.

Proposition 2.4.5 *For any binary quadratic form (a,b,c) as input, the algorithm above will terminate and its output will be a reduced form (a',b',c') equivalent to (a,b,c).*

Proof: Let (a_0,b_0,c_0) have discriminant D. We first do the case where $D<0$. We first note that in this case, both $a_0 \neq 0$ and $c_0 \neq 0$, otherwise $D = b_0^2 > 0$. Suppose we start with $|b_0| > a_0$. By the right choice of T_{m_0} we get a new form (a_0,b_1,c_1) with $b_1 \in \{x \in \mathbb{R} : -|a_0| < x \leq |a_0|\}$. Thus $|b_1| \leq |a_0|$. Now we break all the possible cases satisfied by the form (a_0,b_1,c_1) into the tree below



The idea is to show that in each of the subcases on the left main fork of the tree, the algorithm terminates with the desired output, and then show that if (a_0, b_1, c_1) lies in the right main fork of the tree, then we can run the algorithm to obtain an equivalent form which lies in the left main fork, and then we are done.

We start from the bottom of the left main fork. If $|b_1| < a_0$ and $0 < a_0 < c_1$ then we are done. We have a reduced form. On the other hand, suppose we have a form satisfying $c_1 > a_0 > 0$ and $|b_1| = a_0$. If $b_1 > 0$ we have a reduced form, so we are done. If not, then we have the form $(a_0, -a_0, c_1)$. This form is almost reduced, except that the XY-coefficient has the wrong sign. Applying the element T_1 we obtain the form $(a_0, a_0, c_1 + 1 - a_0)$. This is not reduced since $0 \leq c_1 + 1 - a_0 \leq a_0$, so we go to step 3 and obtain the form $(c_1 + 1 - a_0, -a_0, a_0)$. If $c_1 \leq 2a_0 - 1$ then we have the form $(c_2, -a_0, a_0)$ where $0 < c_2 \leq a_0$. Then we apply T_1 to this form to obtain $(c_2, 2c_2 - a_0, 1)$. Then we go to step 3 to obtain the form $(1, a_0 - 2c_2, c_2)$ and this form finally is reduced since if $c_2 = 1$ then $c_1 = a_0$ which is false, and $a_0 - 2c_2 \leq 0$. On the other hand, if $c_1 > 2a_0 - 1$ then $(c_1 + 1 - a_0, -a_0, a_0)$ is already reduced.

So now both branches of the $a_0 < c_1$ subcase are done. So now suppose $a_0 = c_1$. Then we have the form (a_0, b_1, a_0). If $b_1 > 0$ then we are done. If not, we just apply S to obtain $(a_0, -b_1, a_0)$ and we are done. So now, both branches of the $a_0 \leq c_1$ subcase are done.

So now suppose $a_0 > 0$ and $a_0 > c_1$ in the form (a_0, b_1, c_1). If this form is not reduced, then we go to step 3 and apply S to obtain $(c_1, -b_1, a_0)$. If $|a_0| \leq c_1$ we are done, since that just reduces to the right fork of $a_0 > 0$ in the diagram above, which we just completed. On the other hand, if $|b_0| > c_1$ then we apply some T_{m_1} to produce the form (c_1, b_2, c_2) with $|b_2| \leq c_1$. If $c_1 \leq c_2$ then we are done, since

that just reduces to the right form of $a_0 > 0$ in the diagram above. If $c_1 > c_2$, we apply S to obtain $(c_2, -b_2, c_1)$

We repeat until we obtain a form satisfying the right fork of $a_0 > 0$.

This completes the case where $a_0 > 0$. On the other hand, if $a_0 < 0$ then we can always pass to an equivalent reduced form (a, b, c) satisfying $a_0 > 0$ by some appropriate sequence of T_m and S.

This finally completes the proof in the case where $D < 0$. The case where $D > 0$ is more difficult. The proof for this case is in [4].

□

Note that this provides another way proof of the finiteness of the class group without using any Minkowski theory: Since every form of discriminant D is equivalent to a form in $\mathcal{R}_D$, and there are only finitely many such forms, the class group of D is finite. In the case where $D < 0$, each binary quadratic form of discriminant D is equivalent exactly *one* reduced form.

Proposition 2.4.6 *Let $D < 0$ and let $Q(x, y)$ be a form of discriminant D. Then there is exactly one form in $\mathcal{R}_D$ equivalent to $Q(x, y)$.*

Proof: We have already proven that for any form (a, b, c), there is an equivalent reduced form. Thus we may assume that (a, b, c) is a reduced form. We want to show that any reduced form equivalent to (a, b, c) is (a, b, c). Suppose that (d, e, f) is equivalent to (a, b, c). So let

$$T = \begin{pmatrix} m & n \\ p & q \end{pmatrix}$$

be the transformation that goes between (a, b, c) and (d, e, f). We recall from the proof of 2.4.2 that

$$d = am^2 + bmp + cp^2.$$

Since $mq - np = \pm 1$, mq is coprime to np. Since (a, b, c) is reduced, we have $|b| \leq a \leq c$. We want to show that $a \leq d$. We can assume that $m \neq 0$. If $m = 0$ then $d = cp^2 \geq a$, so we are done.

Now we write

$$am^2 + bmp + cp^2 = am^2 \left(1 + \frac{b}{a}\frac{p}{m}\right) + cp^2.$$

We also write

$$am^2 + bmp + cp^2 = am^2 + cp^2 \left(1 + \frac{b}{a}\frac{m}{p}\right).$$

We know that $\left|\frac{b}{a}\right| \leq 1$. There are two cases to consider Either $\left|\frac{p}{m}\right| < 1$ or $\left|\frac{p}{m}\right| > 1$. In the first case, we apply the first identity to obtain $0 \leq a \leq d$. In the second case, we apply the second identity to again obtain $0 \leq a \leq d$.

Thus, if (a, b, c) is a any reduced form, and (d, e, f) is any equivalent reduced form, then a is minimal among the first coordinates of all reduced forms equivalent to (a, b, c). Thus $a = d$. Therefore we get

$$a = am^2 \left(1 + \frac{b}{a}\frac{p}{m}\right) + cp^2.$$

$$a = am^2 + cp^2 \left(1 + \frac{b}{a}\frac{m}{p}\right).$$

Since $0 \leq a \leq c$, if $\left|\frac{p}{m}\right| < 1$ then the first identity implies that $p = 0$ and $m = 1$, and if $\left|\frac{m}{p}\right| < 1$, the second identity implies that $p = 0$ and $m = 1$. Thus our matrix T must be of the form

$$T = \begin{pmatrix} 1 & n \\ 0 & \pm 1 \end{pmatrix}.$$

We require $q = \pm 1$ since $\det(T) = \pm 1$.

We recall from 2.4.2 that

$$e = 2amn + bmq + bnp + 2cpq = 2an \pm b.$$

Since $a = d$, we must have $e \in (-a, a]$ since $|e| \leq a$ and (d, e, f) is reduced. Thus $n = 0$. Finally, we cannot have $q = -1$. If we let $q = -1$ and then apply T to (a, b, c) we obtain $(c, -b, a)$. If $(c, -b, a)$ is also reduced, it implies that $a = c$ since $a \leq c$ and $c \leq a$. But that requires that $b \geq 0$ and $-b \geq 0$, so $b = 0$. But then $e = 0$ as well by the identity above, so $(a, b, c) = (d, e, f)$, which implies that $q = 1$.

Thus $(a, b, c) = (d, e, f)$. □

Thus, in the case of $D < 0$, the set $\mathcal{R}_D$ can be identified with the ideal class group of $\mathbb{Q}(\sqrt{D})$. This is not true in the case where $D > 0$. In the case where $D > 0$, a binary quadratic form of discriminant D is equivalent to finitely many reduced forms.

In the case $D > 0$ we can describe $\mathcal{R}_D$ in the following way. Let $\rho : \mathcal{R}_D \to \mathcal{R}_D$ be a map which takes a reduced binary quadratic form (a, b, c) and assigns it to (a', b', c'), where (a', b', c') satisfies the following:

- $a' = c$
- b' is the unique integer in $(\sqrt{D} - 2|c|,\ \sqrt{D})$ which is congruent to $-b \bmod 2|c|$

We note that with a' and b' defined, c' is defined as well since it is determined by a', b' and D.

We can view the map ρ as performing a reduction step in accordance with the algorithm above on a binary quadratic form which is already reduced.

Proposition 2.4.7 *For each $D > 0$ the map $\rho : \mathcal{R}_D \to \mathcal{R}_D$ as described above is a permutation of $\mathcal{R}_D$.*

Proof: See [1].

□

Since ρ is a permutation of $\mathcal{R}_D$, we can examine the cycle decomposition of $\mathcal{R}_D$. We note that the leading coefficients of (a, b, c) and $\rho((a, b, c))$ have opposite signs. To see this, we note that $|a| < \sqrt{D}$, since (a, b, c) is reduced, and thus $a^2 < b^2 - 4ac$, so $4ac < b^2 - a^2 < 0$, the right inequality again following from the fact that (a, b, c) is reduced. Thus each ρ-cycle in $\mathcal{R}_D$ has even length.

Proposition 2.4.8 *Two binary quadratic forms (a, b, c) and (a', b', c') in $\mathcal{R}_D$ are equivalent iff they belong to the same ρ-cycle.*

Proof: One direction just follows from the definition of ρ. The other direction may be found in [1].

□

Thus, in the case $D > 0$, the ideal class group of $\mathbb{Q}(\sqrt{D})$ can be identified with the ρ-cycles of $\mathcal{R}_D$. Thus, $\mathcal{R}_D$ comes equipped with a *principal cycle*, the ρ-cycle corresponding to the identity element of $\mathrm{Cl}(\mathbb{Q}(\sqrt{D}))$.

The Cohen-Lenstra heuristic assumption then asserts that $\mathrm{Cl}(\mathbb{Q}(\sqrt{D})$ behaves like a random group of the form $G/\langle\sigma\rangle$, where G is a random group weighted by $1/|\mathrm{Aut}(G)|$, where $G/\langle\sigma\rangle$ can be thought of as $\mathcal{R}_D$ modulo the principal cycle in the space of ρ-cycles of $\mathcal{R}_D$. Although $\mathcal{R}_D$ is not a group, and each ρ-cycle in $\mathcal{R}_D$ does not have the same number of cycles, $\mathcal{R}_D$ has a "group-like" structure which is described in [3], which makes the notion of $\mathcal{R}_D$ modulo the principle cycle sensible.

We have seen that the Cohen-Lenstra heuristic assumption implies results much stronger than anything we can currently prove. Although the ideas of the previous section provide some intuitive grounds to believe the Cohen-Lenstra heuristic, they aren't a formal proof. In the next and final chapter, we will review some interpretations of the Cohen-Lenstra probability measure. That is, we will show that the Cohen-Lenstra probability measure on the family of finite abelian groups corresponds to several other probability measures.

Chapter 3

Interpretations of the Cohen-Lenstra Heuristics

In this section, we will be concerned with connecting the Cohen-Lenstra probability measure on finite abelian groups, to other probability measures that occur in other contexts. Throughout this chapter we will be focused only on the *local* Cohen-Lenstra probability measure.

Definition 3.0.9 Fix an integer prime p. Then let $\mathcal{G}_p$ denote the category of all finite Abelian p-groups. The *local Cohen-Lenstra probability measure* on $\mathcal{G}_p$ is a map $P : \mathcal{P}(\mathcal{G}_{Ap}) \to \mathbb{R}_{\geq 0}$ defined as

$$P(M) = \frac{\sum_{G \in M} |\mathrm{Aut}(G)|^{-1}}{\sum_{G \in \mathcal{G}_p} |\mathrm{Aut}(G)|^{-1}} \text{ for } M \subseteq \mathcal{G}_p.$$

If M has a single element G, we will just write $P(G)$ in place of $P(\{G\})$.

We will keep p as a fixed prime throughout this chapter.

3.1 Young Diagrams of Partitions

Each element of $\mathcal{G}_{Ap}$ clearly corresponds to an *integer partition*. An integer partition is just a multiset with elements from $\mathbb{N}$. For example, the integer partitions of 5 are just the mulisets $\{5\}$, $\{1, 4\}$, $\{2, 3\}$,

$\{1,1,3\}$, $\{1,2,2\}$, $\{1,1,1,2\}$, $\{1,1,1,1,1\}$. The multset $\{a_1, a_2, \cdots, a_n\}$ corresponds to the abelian group

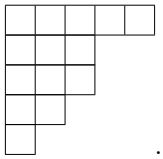
$$\mathbb{Z}/p^{a_1}\mathbb{Z} \oplus \mathbb{Z}/p^{a_2}\mathbb{Z} \cdots \mathbb{Z}/p^{a_n}\mathbb{Z}.$$

Clearly there is a 1-1 correspondence between the abelian groups in $\mathcal{G}_{Ap}$ and the integer partitions. Thus, for an integer partition λ, we will write $P(\lambda)$ to denote the probability of selecting the abelian group corresponding to λ, with respect to the probability measure P above.

We can also associate to each integer partition a *Young diagram*.

Definition 3.1.1 Let $\{a_1, \cdots, a_n\}$ be an integer partition, such that $a_1 \geq a_2 \geq \cdots a_n$. Then the *Young diagram* corresponding to $\{a_1, \cdots, a_n\}$ is a grid of boxes with n rows, with a_i boxes in the ith row.

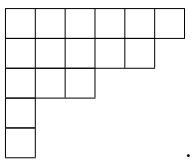
For example, the Young diagram corresponding to the integer partition $\{5,3,3,2,1\}$ is



Clearly there is a 1-1 correspondence between integer partitions and Young diagrams. Through its Young diagram, every integer partition is associated with a *conjugate partition*.

Definition 3.1.2 Let $\{a_1, \cdots, a_n\}$ be an integer partition, with $a_1 \geq \cdots \geq a_n$. The *conjugate partition* of $\{a_1, \cdots, a_n\}$ is the integer partition corresponding to the Young diagram whose rows are the columns of the Young diagram of $\{a_1, \cdots, a_n\}$.

Using the previous Young diagram as an example, if we make the columns of the Young diagram of $\{5,3,3,2,1\}$ into rows, we obtain the following Young diagram:



Thus, the conjugate partition to $\{5,3,3,2,1\}$ is $\{6,5,3,1,1\}$.

For a partition λ and an integer $s \geq 1$, the notation λ_s will denote the number of blocks in the sth-row of the Young diagram of λ (read from top to bottom). For example, if $\lambda = \{5, 3, 3, 2, 1\}$ as above, then $\lambda_1 = 5$, $\lambda_2 = 3$, $\lambda_3 = 3$, $\lambda_4 = 2$, $\lambda_5 = 1$ and $\lambda_s = 0$ for $s > 5$. If λ is the unique integer partition satisfying $\lambda_s = 0$ for all $s \geq 1$ then we refer to λ as the *empty partition*. For a partition λ, the notation λ' will always denote the conjugate partition to λ.

Definition 3.1.3 Let $\{a_1, \cdots, a_n\}$ be an integer partition with $a_1 \geq \cdots \geq a_n$, and $\sum_{i=1}^{n} a_i = M$. A *Young tableau* of $\{a_1, \cdots, a_n\}$ is the Young diagram of $\{a_1, \cdots, a_n\}$ with each box filled in with one element from the set $\{1, \cdots, M\}$, so that each element of $\{1, \cdots, M\}$ appears exactly once in the diagram.

For example, a Young tableau corresponding to the integer partition $\{5, 3, 3, 2, 1\}$ would be

2	7	9	10	11
1	6	14		
3	8	13		
4	5			
12				

.

Clearly there are $M!$ distinct Young tableaux corresponding to a Young diagram of an integer partition $\{a_1, \cdots, a_n\}$ with $\sum_{i=1}^{n} a_i = M$. A Young tableau is called *standard* if each row has entries of increasing value when read from left to right. The above Young tableau is a standard Young tableau.

Several probabilistic interpretations of the Cohen-Lenstra probability measure on $\mathcal{G}_p$ were given by Jason Fulman. We describe some of these below.

3.2 The Young Tableau Algorithm

In this section we examine the *Young Tableau Algorithm*. This is a probabilistic algorithm which outputs an integer partition $\{a_1, \cdots, a_n\}$ with probability $w(G)$, where G is the abelian p-group corresponding to $\{a_1, \cdots, a_n\}$ (recall that throughout this chapter, p is a fixed prime). The algorithm is not a deterministic algorithm, but at each step, makes a decision with a certain probability.

The steps of the algorithm are as follows:

0. Start with the empty partition λ and a collection of weighted coins indexed by $\mathbb{N} = \{1, 2, \cdots\}$, such

that coin i has probability p^{-i} of heads, and $1-p^{-i}$ of tails. Also start with an $N=1$.

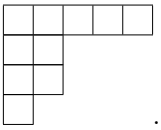
1. Flip coin N. If this coin comes up tails, then set $N := N+1$ and redo step 1. Otherwise go to step 2.

2. Choose an integer $S \geq 1$ according to the following probabilistic rule: Set $S := 1$ with probability $\frac{p^{N-\lambda'_1}-1}{p^N-1}$. For $s > 1$, set $S := s$ with probability $\frac{p^{N-\lambda'_s}-p^{N-\lambda'_{s-1}}}{p^N-1}$. In either case, after S is chosen in this way, add one block to the Sth column of the Young diagram of λ, and then return to step 1.

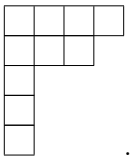
In other words, what we are doing is starting with a partition and an integer N. We flip the Nth coin. If it comes up tails, we flip the $N+1$th coin and repeat. If not, we add another block to the Young diagram of our partition and then start again. Note the step 3 is well-defined because if we are at step 2 and have partition λ and integer N, and λ has k columns, then the probability of selecting $s > k+1$ is zero.

Below we do an example of following several steps of the algorithm:

Suppose we are currently at step 1, with $N = 5$ and the partition $\lambda = \{5,2,2,1\}$, so we have the following Young diagram

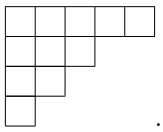


Thus the conjugate partition is



Now we flip coin 5. With probability p^{-5} we get heads, and with probability $1-p^{-5}$ we get tails. If we get tails, we redo the process with coin 6. Otherwise, we choose an integer S according to the following probabilities. We choose $S := 1$ with probability $\frac{p-1}{p^5-1}$. We choose $S := 2$ with probability $\frac{p^2-p}{p^5-1}$. We choose $S := 3$ with probability $\frac{p^4-p^2}{p^5-1}$. We choose $S := 4$ with probability 0. We choose $S := 5$ with probability 0. We choose $S := 6$ with probability $\frac{p^5-p^4}{p^5-1}$. We choose $S > 6$ with probability 0.

We note that this probability distribution is well defined since $\sum_{s\geq 1} P(S := s) = \frac{(p-1)+(p^2-p)+(p^4-p^2)+(p^5-p^4)}{p^5-1} = \frac{p^5-1}{p^5-1} = 1$. If we have selected, say, $S := 3$ then we replace the Young diagram of λ with the Young diagram



which is the Young diagram of $\{5, 3, 2, 1\}$. Now we go back to step 1.

This algorithm doesn't halt, but with probability 1, it ouputs a finite partition. This means that if, after some finite number of steps, the algorithm has arrived at the partition ω, there is a *positive* probability that in all the infinitely many steps of the algorithm after this point, the algorithm will add no more blocks to the Young diagram of ω (in other words, the algorithm outputs ω with positive probability).

Proposition 3.2.1 *With probability 1, the algorithm above outputs an integer partition. That is, if we treat the possible outputs of the algorithm as infinite sequences of Young diagrams, the probability that the algorithm will output a sequence which stops adding blocks after finitely many steps is one.*

Proof: For each $N \geq 1$ let A_N denote the event "the Nth coin comes up heads at least once". The conditional probability that the Nth coin always comes up tails, given that the algorithm reaches the integer N in finite time, is equal to the conditional probability that the Nth coin comes up tails the first time the algorithm reaches N. since a coin can only come up tails once. This is equal to $1 - p^{-N}$. Thus the conditional probability of A_N, given that the algorithm reaches N is finite time, is p^{-N}. Thus

$$\sum_{N\geq 1} P(A_N) \leq \sum_{N\geq 1} p^{-N} < \infty.$$

Thus, by the Borel-Cantenelli lemma, with probability 1 only finitely many of the events $\{A_N : N \geq 1\}$ occur. For N, k let $B_{N,k}$ denote the event that the Nth coin comes up heads k times.

Suppose then that the algorithm outputs a sequence in which the events $A_{N_1}, \cdots, A_{N_r}$ occur. Then $P(B_{N_i,k} \mid A_{N_i}) = p^{-N_i k}$. Thus

$$\sum_{k\geq 1} P(B_{N_i,k}) < \infty.$$

Thus a second application of the Borel-Cantenelli lemma implies that any coin coming up heads at least once, does so only finitely many times with probability 1.

Thus with probability 1, the algorithm outputs a sequence in which for some i sufficiently large, after the ith stage all coin tosses come up tails. That is, the algorithm outputs a sequence which stops adding blocks to the Young diagram after some finite step. So, with probability 1, the algorithm outputs a finite partition in finitely many steps.

□

The connection between the algorithm and the local Cohen-lenstra probability measure is as follows:

Proposition 3.2.2 *For a particular integer partition ω, the probability that the algorithm above outputs the conjugate partition to ω (that is, outputs an infinite sequence which reaches ω' after finitely many steps, and then stops adding blocks) is $P(\omega)$.*

Proof: See [2].

□

3.3 The Conjugacy Class Interpretation

Another interpretation of the CL probability measure on $\mathcal{G}_{Ap}$ provided by Fulman is in terms of conjugacy classes of elements of $\mathrm{GL}(n,p)$ (that is, the group of invertible $n \times n$ matrices over $\mathbb{F}_p$). To move from conjugacy classes to integer partitions, we define the *Jordan-Chevalley normal form.*

Definition 3.3.1 Let $\phi(X) \in \mathbb{F}_p[X]$ be a monic polynomial of degree m, where $\phi(X) = X^m + a_{m-1}X^{m-1} + \cdots + a_1X + a_0$. To this polynomial we associate a *companion matrix* over $\mathbb{F}_p$ with size $m \times m$. This matrix is denoted by $C(\phi)$ and is defined as

$$C(\phi) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{m-1} \end{pmatrix}.$$

Definition 3.3.2 *A* Jordan-Chevalley normal form *matrix in* $\mathrm{GL}(n,p)$ *is a block matrix of the following form*

$$\begin{pmatrix} J_1 & 0 & 0 & \cdots & 0 \\ 0 & J_2 & 0 & \cdots & 0 \\ 0 & 0 & J_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & J_r \end{pmatrix}$$

where the block J_i *is of the form* $C(\phi_i^{s_i})$*, where* $\phi_i(X) \in \mathbb{F}_p[X]$ *is a monic irreducible polynomial and* s_i *is a positive integer, such that*

$$n = \sum_{i=1}^{r} s_i \deg(\phi_i).$$

That is, a Jordan-Chevalley normal form is given by associating to each irreducible polynomial $\phi(X) \in \mathbb{F}_p[X]$ an integer partition, where all but finitely many irreducible monic polynomials in $\mathbb{F}_p[X]$ recieve the empty partition. As an example, if a particular monic irreducible $\psi(X) \in \mathbb{F}_p[X]$ recieves the partition $\lambda_\psi = \{4,4,3,3,3,1\}$, this means that the corresponding Jordan-Chevalley normal form contains two blocks of the form $C(\psi^4)$, three blocks of the form $C(\psi^3)$ and one block of the form $C(\psi)$.

Thus, a matrix $A \in \mathrm{GL}(n,p)$ in Jordan-Chevalley normal form corresponds to a family $(\lambda_\phi)_\phi$, where each λ_ϕ is an integer partition and ϕ runs over all irreducible monic polynomials in $\mathbb{F}_p[X]$, all but finitely many of the λ_ϕ are empty partitions. Clearly the monic irreducible polynomial X must recieve the empty partition, otherwise the matrix A is not invertible. Furthermore, if we let $\lambda_{\phi,s}$ denote the number of blocks in the s-th row of the Young diagram of $\lambda_{\phi,s}$, then the condition that the sum of the sizes of the Jordan blocks adds to n is expressed as

$$\sum_{\phi,s} \deg(\phi) \cdot \lambda_{\phi,s} = n.$$

Conversely, suppose we start with a family $(\lambda_\phi)_\phi$ of integer partitions indexed by all the monic irreducible polynomials in $\mathbb{F}_p[X]$, and this family satisfiees the following properties

- $\lambda_X = \varnothing$
- $\sum_{\phi,s} \deg(\phi) \cdot s \cdot \lambda_{\phi,s} = n.$

Then the family $(\lambda_\phi)_\phi$ determines a conjugacy class in $\mathrm{GL}(n,p)$.

Now that we have the appropriate dictionary between integer partitions and conjugacy classes in $\mathrm{GL}(n,p)$, and integer partitions can be regarded as elements of $\mathcal{G}_p$, we have the following

Proposition 3.3.3 *Fix a monic polynomial $\phi(X) \in \mathbb{F}_p[X]$ of degree 1. Let λ be any integer partition. For each conjugacy class $[A]$ of $\mathrm{GL}(n,p)$, let $(\lambda_{[A],\phi})_\phi$ denote the corresponding family of integer partitions. Then if we choose a conjugacy class $[A]$ in $\mathrm{GL}(n,p)$ uniformly at random we get*

$$\lim_{n\to\infty} \mathbb{P}(\lambda_{[A],\phi} = \lambda) = w(\lambda).$$

In other words, if we choose a conjugacy class $[A]$ of $\mathrm{GL}(n,p)$ uniformly at random, with corresponding partition family $(\lambda_{[A],\phi})_\phi$ the probability that $\lambda_{[A],\phi} = \lambda$ converges to $P(\lambda)$ as $n \to \infty$.

Proof: See [2].

□

Chapter 4

Conclusion

The approach of Cohen and Lenstra appears to be a promising one for solving the class number problem for real quadratic fields. Applying the methods and results of the previous chapter, results in group theory and representation theory about conjugacy classes of $\mathrm{GL}(n, p)$ can be transferred to the study of the Cohen-Lenstra probability measure. This offers a potential avenue for proving the Cohen-Lenstra heuristic and therefore solving the class number problem for real quadratic fields.

Bibliography

[1] Dirichlet, P.G. and Dedekind, R., *Vorlesungen über Zahlentheorie*. New York: Springer-Verlag, 1968 (reprint).

[2] Fulman, J., *A Probabilistic Approach Toward Conjugacy Classes in the Finite General Linear and Unitary Groups*. *Journal of Algebra* 212, 557-590, 1999.

[3] Lengler, J., *The Cohen-Lenstra heuristic: Methodology and results*. *Journal of Algebra* 323, 2960-2967, 2010.

[4] Lenstra, H.W., *On the Calculation of Regulators and Class Numbers of Quadratic Fields*. Cambridge: University Press, 1982.

[5] Lenstra, H. and Cohen, H.W., *Heuristics on Class Groups of Number Fields*. *Lecture Notes on Math vol. 1086*. Berlin: Springer-Verlag, 1984.

[6] Marcus, A., *Number Fields*. New York: Springer-Verlag, 1977.

[7] Neukirch, J., *Algebraic Number Theory*. Berlin: Springer-Verlag, 1999.

[8] Reiner, I., *Maximal Orders*. Oxford: Clarendon Press, 2003.

[9] Stark, H., *The Gauss class-number problems*. Clay Mathematics Proceedings, Volume 7, 2007.