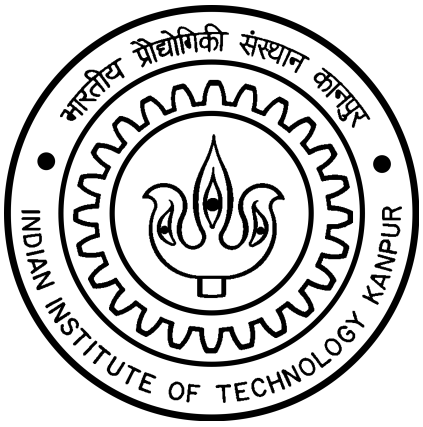
Uniqueness of Factorization in Quadratic Fields

Pritam Majumder

Supervisors: (i) Prof. G. Santhanam, (ii) Prof. Nitin Saxena



A project presented for the degree of
Master of Science

Department of Mathematics and Statistics
Indian Institute of Technology Kanpur
India
April, 2014

Contents

Chapter 0

Introduction

In this project, we will discuss the problem of finding all quadratic fields whose ring of integers are Unique Factorization Domains. Like most of the problems in Number Theory, this problem was also motivated by the problem of solving Diophantine equations. Let us see a few examples to see how the uniqueness of factorization in the ring of integers of quadratic fields can be helpful in solving Diophantine equations.

Example 0.1. *Solving the Diophantine equation* $y^2 = x^3 - 2$:

Suppose $x^3 = y^2 + 2$ for $x, y \in \mathbb{Z}$. If y is even, then x is even i.e. $4 \mid x^3$ and $4 \mid y^2$, which implies $4 \mid 2$ but this is absurd. Therefore y is odd. We will later show that the ring of integers of $\mathbb{Q}(\sqrt{-2})$ is $\mathbb{Z}[\sqrt{-2}]$ and this a UFD. Now in $\mathbb{Z}[\sqrt{-2}]$ we have,

$$(y + \sqrt{-2})(y - \sqrt{-2}) = x^3.$$

Suppose $c + d\sqrt{-2}$ is a common divisor of $y + \sqrt{-2}$ and $y - \sqrt{-2}$. Then,

$$c + d\sqrt{-2} \mid (y + \sqrt{-2}) + (y - \sqrt{-2}) = 2y$$

and

$$c + d\sqrt{-2} \mid (y + \sqrt{-2}) - (y - \sqrt{-2}) = 2\sqrt{-2}.$$

Then, taking norm, $c^2 + 2d^2 \mid 4y^2$ and $c^2 + 2d^2 \mid 8$ in $\mathbb{Z}$. If $c^2 + 2d^2 \nmid 4$, then $c^2 + 2d^2 \mid y^2$ (since $\gcd(x, y) = 1$), i.e. $c^2 + 2d^2$ is odd (since, y is odd) but this implies $c^2 + 2d^2 \nmid 8$. Therefore, $c^2 + 2d^2 \mid 4$, that is

$$c^2 + 2d^2 = 1, 2 \text{ or } 4.$$

If $c^2 + 2d^2 = 1$ then $d = 0$, $c = \pm 1$, i.e. $c + d\sqrt{-2}$ is a unit. If $c^2 + 2d^2 = 2$ then $d = \pm 1$, $d = 0$, again $c + d\sqrt{-2}$ is a unit. If $c^2 + 2d^2 = 4$, then $d = 0$,

$c = \pm 2$ or $d = \pm 2$, $c = 0$, but then $\pm 2 \mid y + \sqrt{-2}$ or $\pm 2\sqrt{-2} \mid y + \sqrt{-2}$ but none is possible since y is odd. Therefore, $y + \sqrt{-2}$ and $y - \sqrt{-2}$ are coprime. Then, since $\mathbb{Z}[\sqrt{-2}]$ is a UFD, we have

$$y + \sqrt{-2} = um^3 \quad \text{and} \quad y - \sqrt{-2} = u^{-1}n^3$$

for some unit u and $m, n \in \mathbb{Z}[\sqrt{-2}]$. But, $u = \pm 1$ which is a cube. This implies there exist $a, b \in \mathbb{Z}$, such that

$$y + \sqrt{-2} = (a + b\sqrt{-2})^3.$$

Now, equating coefficients of $\sqrt{-2}$ on both sides we get $1 = b(3a^2 - 2b^2)$. Therefore $b = \pm 1$. If $b = 1$, then $3a^2 - 2b^2 = 1$ implies $a = \pm 1$. If $b = -1$, then $3a^2 - 2b^2 = -1$ implies $3a^2 = 1$, but this can not happen. Therefore we have $a = \pm 1$, $b = 1$. Now, we also have $y = a^3 - 6ab^2$, which implies $y = -5, 5$ and then $x^3 = y^2 + 2 = 27$ implies $x = 3$. Hence, the solutions of $y^2 = x^3 - 2$ are given by $x = 3$, $y = \pm 5$.

Example 0.2. *Solving the Diophantine equation* $y^2 = x^3 - 1$:

Let $x, y \in \mathbb{Z}$ be such that $x^3 = y^2 + 1 = (y + i)(y - i)$, where $i := \sqrt{-1}$. Suppose $c + id$ is a common divisor of $y + i$ and $y - i$ in the ring $\mathbb{Z}[i]$. Then,

$$c + id \mid (y + i) + (y - i) = 2y$$

and

$$c + id \mid (y + i) - (y - i) = 2i.$$

Taking norms we have, $c^2 + d^2 \mid 4y^2$ and $c^2 + d^2 \mid 4$. We have three cases:

Case 1: If $c^2 + d^2 = 1$, then $c = 0$, $d = \pm 1$ or $c = \pm 1$, $d = 0$. That is, $c + id$ is a unit.

Case 2: If $c^2 + d^2 = 2$, then $c = \pm 1$, $d = \pm 1$. Now, if x is even then $y^2 \equiv -1 \equiv 3 \pmod 4$, which is not possible. Therefore x is odd and hence y is even. Suppose $1 + i \mid y + i$, then $y + i = (1 + i)(p + iq)$ for some $p, q \in \mathbb{Z}$. Then, $p - q = y$ and $p + q = 1$, which implies $y + 1 = 2p$ but this is a contradiction, since y is even. Hence we conclude that $(1 + i) \nmid (y + i)$. Similarly, we can also show that $(\pm 1 \pm i) \nmid (y + i)$.

Case 3: If $c^2 + d^2 = 4$, then $c = 0$, $d = \pm 2$ or $c = \pm 2$, $d = 0$. That is, $\pm 2 \mid y + i$ or $\pm 2i \mid y + i$. If $y + i = \pm 2(p + iq)$ for some $p, q \in \mathbb{Z}$, then $\pm 2q = 1$, which is impossible. Again, if $y + 1 = \pm 2i(p + iq)$ then $\pm 2p = 1$, which is also impossible.

Therefore $c + id$ is a unit. Hence, $y + i$ and $y - i$ are coprime. Now, using the fact that $\mathbb{Z}[i]$ is a UFD (which we will prove later), we have

$$y + i = um^3 \quad \text{and} \quad y - i = u^{-1}n^3$$

for some unit u and $m, n \in \mathbb{Z}[i]$. Now $u \in \{\pm 1, \pm i\}$ implies that u is a cube in $\mathbb{Z}[i]$. Therefore, there exist $a, b \in \mathbb{Z}$ such that

$$y + i = (a + ib)^3. \qquad (*)$$

Equating the coefficients of i we have, $1 = b(3a^2 - b^2)$, which implies $b = \pm 1$. If $b = 1$, then $3a^2 - b^2 = 1$ implies $3a^2 = 2$, which is not possible. If $b = -1$, then $3a^2 - b^2 = -1$ implies $a = 0$. From $(*)$ we also have, $y = a^3 - 3ab^2$. This implies $y = 0$ and $x^3 = y^2 + 1 = 1$ implies $x = 1$. Hence $x = 1$, $y = 0$ is the only solution of $y^2 = x^3 - 1$.

Note that, in the above examples we have used the crucial property of unique factorization in the rings $\mathbb{Z}[\sqrt{-2}]$ and $\mathbb{Z}[\sqrt{-1}]$. So it is a natural question to ask whether we can find all quadratic fields whose ring of integers has the property of unique factorization, so that we can have similar tools to solve a wide range of Diophantine equations. This problem is partially solved and partially unsolved. For imaginary quadratic fields, i.e. quadratic fields $\mathbb{Q}(\sqrt{d})$ with $d < 0$, Gauss showed that if $d = -1, -2, -3, -7, -11, -19, -43, -67, -167$, then the ring of integers of $\mathbb{Q}(\sqrt{d})$ is a UFD. Gauss also conjectured that these are the only imaginary quadratic UFDs, which was proved later by Heegner and Stark. For real quadratic fields (i.e. for $d > 0$) this question is still open. It is not even known whether there are infinitely many quadratic fields (or, even number fields!) whose ring of integers have unique factorization. Gauss conjectured that there are infinitely many real quadratic UFDs. Towards this direction of research in Number Theory, Cohen and Lenstra have given a very promising set of conjectures, known as Cohen-Lenstra Heuristics. According these conjectures a positive fraction of all real quadratic fields will have the property of unique factorization. There is also a connection between Cohen-Lenstra Heuristics and integer partition, due to J. Lengler and J. Fulman, which is also quite interesting. Fulman has given several probability measures on the set of all partitions which turn out to be equivalent to Cohen-Lenstra probability measure for p-groups. So, if some natural connection between real quadratic fields and partitions is found, then that might indicate some way to prove Cohen-Lenstra Heuristics for real quadratic fields.

Let me give a brief overview of the contents of this project here. In *Chapter 1*, we will develop some basic Algebraic Number Theory and prove that there are at least nine imaginary quadratic fields whose ring of integers is a UFD. In *Chapter 2*, we will discuss two main algorithms to compute class numbers (a quantity that measures uniqueness of factorization) of real quadratic fields. These are as follows: (i) *Reduced Form Algorithm*: Here one reduces every ideal in class group to an equivalent "reduced ideal". One then

shows that the set of all reduced ideals is finite and this set can be decomposed into finitely many cycles (the operator concerned is the 'reduction operator'). The number of such cycles gives the class number of the quadratic field. This number is counted by establishing a connection between "ideals" and "quadratic forms". (ii) Using *Analytic formula for Class Number*: Using this formula the problem boils down to computing the *regulator* of the quadratic field and the value of $L(1, \chi)$, where L is the Dirichlet L-function and χ is the quadratic character. We will study a method, due to Shanks, for computing the regulator. One defines a new notion of "distance" between ideals and with respect to this distance length of any cycle of reduced ideals has length almost equal to the regulator. So the problem of computing the regulator becomes almost like computing the order of a cyclic group. We will discuss a method to compute $L(1, \chi)$ by approximating a certain infinite product and we will also establish an identity which shows that $L(1, \chi)$ can be expressed as a finite sum which makes the direct computation of $L(1, \chi)$ possible (although this is not efficient). In *Chapter 3*, we will discuss Cohen-Lenstra Heuristics.We show that over the class of all finite abelian p-groups one can define a probability measure P such that $P(\{G\}) \propto 1/|\mathrm{Aut}(G)|$, where G is a p-group. We also establish that the probability measure P is a "natural" distribution by showing that this distribution can be obtained as a limiting distribution of *Haar measure* which is a generalization of the so called Lebesgue Measure to topological groups. We show that the above probability measure P can be extended to a reasonably large σ-algebra over the class of all finite abelian groups. Using this extension we will show that the probability that a real quadratic field has class number 1 is approximately 0.75. This will give us a hint that Gauss's conjecture (that there are infinitely many real quadratic field of class number one) might be true. Finally we will discuss the connection between Cohen-Lenstra heuristics and integer partitions. We will discuss several probability measures on partitions, due to J. Fulman, which are equivalent to the Cohen-Lenstra probability measure.

Chapter 1

Imaginary Quadratic Fields

Together with developing basic notions of Algebraic Number Theory, in this chapter our goal will be to prove that the ring of integers of a quadratic field $\mathbb{Q}(\sqrt{d})$ is a Unique Factorization Domain (UFD) for $d = -1, -2, -3, -7, -11, -19, -43, -67, -163$. In fact, this is the complete list of UFD quadratic fields with $d < 0$. We will prove this in two major steps. First we will show that ring of integers of $\mathbb{Q}(\sqrt{d})$ is a Euclidean Domain (and hence UFD) for $d = -1, -2, -3, -7, -11$. Next we will develop the notion of inverse of an ideal and class groups to prove the property of UFD for the rest of the values of d. We will borrow some results from [ST01].

Definition 1.1. *A field K is called a **Quadratic Field** if $\mathbb{Q} \subseteq K$ is an algebraic field extension of degree 2. (In general, a field K is called a **Number Field** if $\mathbb{Q} \subseteq K$ is a finite algebraic field extension.)*

Proposition 1.2. *If K is a quadratic field then $K = \mathbb{Q}(\sqrt{d})$, for some $d \in \mathbb{Z}$ such that d is square-free.*

Proof. See [ST01] Proposition 3.1. □

Next we define the norm of an element in a quadratic field.

Definition 1.3. *Let $K = \mathbb{Q}(\sqrt{d})$ be a quadratic field. Then, the **norm** of any $\alpha := r + s\sqrt{d} \in K$, where $r, s \in \mathbb{Q}$, is defined as*

$$N(\alpha) := r^2 - ds^2.$$

(In general, for any number field K and any $\alpha \in K$, we define

$$N(\alpha) := \prod_{i=1}^{n} \sigma_i(\alpha),$$

where σ_i's are the distinct embeddings of K in $\mathbb{C}$ and $n = [K : \mathbb{Q}]$.)

Definition 1.4. *A complex number α is said to be an **algebraic integer** if it satisfies some monic polynomial in $\mathbb{Z}[x]$.*

We will denote the set of all algebraic integers by $\mathbb{B}$. Then $\mathbb{B}$ is a subring of $\mathbb{C}$, see [ST01] Theorem 2.9.

Definition 1.5. *The **ring of integers** $\mathfrak{O}_K$ of any number field K is defined as*

$$\mathfrak{O}_K := K \cap \mathbb{B},$$

which is clearly a subring of K.

Lemma 1.6. *Let K be a number field and $\alpha \in K$. Then $\alpha \in \mathfrak{O}_K$ if and only if minimal polynomial of α over $\mathbb{Q}$ has integer coefficients.*

Proof. If minimal polynomial of α is in $\mathbb{Z}[x]$ then clearly $\alpha \in \mathfrak{O}_K$. Conversely if $\alpha \in \mathfrak{O}_K$ then α satisfies some $q \in \mathbb{Z}[x]$. Let $p \in \mathbb{Q}[x]$ be the minimal polynomial of α, then $p \mid q$ and by Gauss' lemma it follows that $p \in \mathbb{Z}[x]$. □

Theorem 1.7. *Let $K := \mathbb{Q}(\sqrt{d})$ be a quadratic field, where $d \in \mathbb{Z}$ is square-free as usual. Then,*

$$\mathfrak{O}_K = \begin{cases} \mathbb{Z}[\sqrt{d}] & \text{if } d \not\equiv 1 \pmod 4 \\ \mathbb{Z}\left[\frac{1+\sqrt{d}}{2}\right] & \text{if } d \equiv 1 \pmod 4 \end{cases}.$$

Proof. The proof follows from the Lemma 1.6 and elementary number theory, see [ST01] Theorem 3.2. □

It is easy to see that for a quadratic field K, norm is multiplicative i.e. $N(\alpha\beta) = N(\alpha)N(\beta)$ and $\alpha \in \mathfrak{O}_K$ implies $N(\alpha) \in \mathfrak{O}_K$ (follows by direct computation of norm). This is also true for any number field; see [ST01], section 2.5.

1.1 Imaginary Euclidean Quadratic Fields

Definition 1.8. *A domain D is called a **Euclidean Domain** (ED) if there exist a function $\phi\colon D \setminus \{0\} \to \mathbb{N}$ such that*

- *If $a, b \in D \setminus \{0\}$ and $a \mid b$ then $\phi(a) \leq \phi(b)$.*
- *If $a, b \in D \setminus \{0\}$ then there exists $q, r \in D$ such that $a = bq + r$, where either $r = 0$ or $\phi(r) < \phi(b)$.*

*and such a function ϕ is called **Euclidean function**.*

We will follow the following convention: we will say that a number field K is ED to mean that it's ring of integers $\mathfrak{O}_K$ is a Euclidean domain. In this section we will prove that for $d < 0$ the quadratic field $\mathbb{Q}(\sqrt{d})$ is ED if and only if $d = -1, -2, -3, -7, -11$.

Definition 1.9. *A number field K is called **norm-ED** if it's ring of integers $\mathfrak{O}_K$ is a ED with respect to the Euclidean function function $|N|$, where $N(\cdot)$ is the norm function.*

Lemma 1.10. *A number field K is norm-ED if and only if for all $x \in K$ there exists $m \in \mathfrak{O}_K$ such that $|N(x - m)| < 1$.*

Proof. Suppose K is a norm-ED. Let $x \in K$, then there exists $f \in \mathbb{Q}[x]$ such that $f(x) = 0$. By clearing out the denominators of f we have a polynomial $g \in \mathbb{Z}[x]$ such that $g(x) = 0$. Then cx satisfies a monic polynomial in $\mathbb{Z}[x]$, where c is the leading coefficient of g. Hence, $cx \in \mathfrak{O}_K$ and $c \neq 0$. Now by Euclideanity, there exist $q, r \in \mathfrak{O}_K$ such that $cx = cq + r$. If $r = 0$ then $cx = cq$. Hence $x = q \in \mathfrak{O}_K$ and we are done by taking $m = x$. Otherwise, $cx = cq + r$ and $|N(r)| < |N(c)|$, i.e. equivalently we have

$$x = q + \frac{r}{c} \quad \text{with} \quad \left|N\left(\frac{r}{c}\right)\right| < 1.$$

Then take $m = q$ and we are done.

Conversely, suppose $\forall x \in K$, there exist $m \in \mathfrak{O}_K$ such that $|N(x-m)| < 1$. Then we need to show that $N(\cdot)$ is a Euclidean function. Suppose $a \mid b$, where $a, b \in \mathfrak{O}_K \backslash \{0\}$. Then, $b = aa'$ for some $a' \in \mathfrak{O}_K$. This implies $|N(b)| = |N(a)| \cdot |N(a')|$ and hence $|N(a)| \leq |N(b)|$. Now let $a, b \in \mathfrak{O}_K \setminus \{0\}$. Then there exists $m \in \mathfrak{O}_K$ such that $\left|N\left(\frac{a}{b} - m\right)\right| < 1$ i.e. $|N(a - bm)| < |N(b)|$. Now, taking $q = m$ and $r = a - bm$, we have $a = bq + r$ and $|N(r)| < |N(b)|$. Hence we have proved that K is norm-ED. □

Theorem 1.11. *The quadratic field $\mathbb{Q}(\sqrt{d})$ is norm-ED if*

$$d = -1, -2, -3, -7, -11.$$

Proof. We will divide the proof into two cases:

Case 1: Let $d = -1, -2$, then $d \not\equiv 1 \pmod 4$. By the Lemma 1.10, given $r + s\sqrt{d} \in \mathbb{Q}(\sqrt{d})$, we need to find $x + y\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$ such that

$$|(r - x)^2 - d(s - y)^2| < 1 \quad \text{i.e.} \quad (r - x)^2 - d(s - y)^2 < 1$$

since $d < 0$. Now given $r, s \in \mathbb{Q}$, one can always choose $x, y \in \mathbb{Z}$ such that

$$|r - x| \leq \frac{1}{2} \quad \text{and} \quad |s - y| \leq \frac{1}{2}.$$

Hence we get

$$(r-x)^2 - d(s-y)^2 \leq \frac{1}{4} + 2 \cdot \frac{1}{4} = \frac{3}{4} < 1.$$

Therefore $\mathbb{Q}(\sqrt{d})$ is norm-ED for $d = -1, -2$.

Case 2: Let $d = -3, -7, -11$, then $d \equiv 1 \pmod 4$. Again by lemma, given $r + s\sqrt{d} \in \mathbb{Q}(\sqrt{d})$, we need to find

$$x + y\left(\frac{1+\sqrt{d}}{2}\right) = \left(x + \frac{y}{2}\right) + \frac{y}{2}\sqrt{d} \quad \in \mathbb{Z}\left[\frac{1+\sqrt{d}}{2}\right]$$

such that

$$\left(r - x - \frac{y}{2}\right)^2 - d\left(s - \frac{y}{2}\right)^2 < 1.$$

Now, first choose $y \in \mathbb{Z}$ such that $|2s - y| \leq 1/2$ and then choose $x \in \mathbb{Z}$ such that $|(r - y/2) - x| \leq 1/2$. Then we have,

$$\begin{aligned}\left(r - x - \frac{y}{2}\right)^2 - d\left(s - \frac{y}{2}\right)^2 &= \left(r - \frac{y}{2} - x\right)^2 - \frac{d}{4}(2s-y)^2 \\ &\leq \frac{1}{4} + \frac{11}{4} \cdot \frac{1}{4} \\ &= \frac{15}{16} < 1.\end{aligned}$$

Hence $\mathbb{Q}(\sqrt{d})$ is norm-ED for $d = -3, -7, -11$. □

Proposition 1.12. *Let K be a number field. Let $\alpha \in \mathfrak{O}_K$, then α is a unit if and only if $N(\alpha) = \pm 1$.*

Proof. Suppose α is a unit, i.e. $\alpha\beta = 1$ for some $\beta \in \mathfrak{O}_K$. This implies $N(\alpha)N(\beta) = N(1) = 1$ and hence $N(\alpha) = \pm 1$, since $N(\alpha), N(\beta)$ are integers.

Conversely, suppose $N(\alpha) = \pm 1$, then from the definition of norm

$$\prod_{i=1}^{n} \sigma_i(\alpha) = \alpha \cdot \prod_{i=2}^{n} \sigma_i(\alpha) = \pm 1$$

where σ_i's are the distinct embeddings of K in $\mathbb{C}$, $n = [K : \mathbb{Q}]$ and $\sigma_1 = \mathrm{id}$. Now since $\alpha \in \mathbb{B}$ and σ_i's are $\mathbb{Q}$-linear, we have $\sigma_i(\alpha) \in \mathbb{B}$ for each i. Hence $\beta := \prod_{i=1}^{2} \sigma_i(\alpha) \in \mathbb{B}$. Also $\beta = \alpha^{-1} \in K$. Hence $\beta \in \mathfrak{O}_K$ and $\alpha\beta = \pm 1$. Therefore α is a unit. □

Proposition 1.13. *Let $K = \mathbb{Q}(\sqrt{d})$ be a quadratic field, where $d < 0$ and square-free. Then the group of units (denoted by $\mathfrak{O}_K^\times$) of $\mathfrak{O}_K$ is given by*

- *For* $d = -1$, $\mathfrak{O}_K^\times = \{\pm 1, \pm i\}$.
- *For* $d = -3$, $\mathfrak{O}_K^\times = \{\pm 1, \pm\omega, \pm\omega^2\}$ *where* $\omega = e^{2\pi i/3}$.
- *For all other* $d < 0$, $\mathfrak{O}_K^\times = \{\pm 1\}$.

Proof. Since $d < 0$, the proof easily follows from Proposition 1.12 by solving easy diophantine equation. See [ST01] Proposition 4.2. □

We will need the following theorem on free $\mathbb{Z}$-module

Theorem 1.14. *Let G be a free $\mathbb{Z}$-module and let H be a submodule of G. Suppose $\{x_1, \dots x_n\}$ is a basis for G and $\{y_1, \dots y_n\}$ is a basis for H. If for each i, $y_i = \sum_{j=1}^n a_{ij}x_j$, where $a_{ij} \in \mathbb{Z}$. Then we have*

$$|G/H| = |\det([a_{ij}])|.$$

Proof. See [ST01] Theorem 1.17. □

Theorem 1.15. *If $d < -11$ and square-free then $\mathbb{Q}(\sqrt{d})$ is not a ED.*

Proof. Let $d < -11$ and $K = \mathbb{Q}(\sqrt{d})$. Suppose $\mathfrak{O}_K$ is a ED with some Euclidean function ϕ. Consider the set

$$S := \{\phi(x) : x \in \mathfrak{O}_K, x \neq 0 \text{ and } x \notin \mathfrak{O}_K^\times\} \subseteq \mathbb{N}.$$

Let $\phi(\alpha)$ be the least element of S (note that S is non-empty). Now, take any $\beta \in \mathfrak{O}_K$. By Euclideanity there exist $\gamma, \delta \in \mathfrak{O}_K$ such that $\beta = \alpha\gamma + \delta$. If $\delta \neq 0$ then $\phi(\delta) < \phi(\alpha)$ and from the minimality of $\phi(\alpha)$ we conclude that $\delta \in \mathfrak{O}_K^\times$, i.e. $\delta = \pm 1$ by Proposition 1.13. Hence δ can have at most three values 0,1 or -1 and this implies $|\mathfrak{O}_K/\langle\alpha\rangle| \leq 3$.

Now, $\mathfrak{O}_K$ is a free $\mathbb{Z}$-module of rank 2. Assume $d \not\equiv 1 \pmod 4$. Then $\langle\alpha\rangle$ is a submodule of $\mathfrak{O}_K$ with $\mathbb{Z}$-basis $\{\alpha, \alpha\sqrt{d}\}$. Let $\alpha = a + b\sqrt{d}$ where $a, b \in \mathbb{Z}$. Then $\langle\alpha\rangle$ has the $\mathbb{Z}$-basis $\{a + b\sqrt{d}, bd + a\sqrt{d}\}$. Now using Theorem 1.14 we have

$$|\mathfrak{O}_K/\langle\alpha\rangle| = \left|\begin{vmatrix} a & b \\ db & a \end{vmatrix}\right| = |a^2 - db^2| = a^2 - db^2.$$

Therefore from $|\mathfrak{O}_K/\langle\alpha\rangle| \leq 3$ we get $a^2 - db^2 \leq 3$. A similar calculation for $d \equiv 1 \pmod 4$ with $\alpha = a + b\frac{\sqrt{d}+1}{2}$ yields $a^2 - db^2 \leq 12$. Now one can easily check that the only solution to these diophantine equations are given by the trivial solution $a = \pm 1, b = 0$. Hence α is a unit and we have a contradiction. □

Now, we are left with three more square-free negative values of d, namely $d = -5, -6, -10$. We will show that for these values of d the quadratic field $\mathbb{Q}(\sqrt{d})$ is not a ED and this will imply that the values of d in Theorem 1.11 gives the complete list of Euclidean quadratic fields for $d < 0$.

Proposition 1.16. *Let K be a number field and let $x \in \mathfrak{O}_K$ be such that $N(x)$ is prime. Then x is irreducible in $\mathfrak{O}_K$.*

Proof. Suppose $x = yz$ with $y, z \in \mathfrak{O}_K$, then we need to show that either y or z is a unit. Now $N(x) = N(y)N(z)$ is prime and hence either $N(y) = \pm 1$ or $N(z) = \pm 1$. Therefore $y \in \mathfrak{O}_K^\times$ or $z \in \mathfrak{O}_K^\times$. □

Theorem 1.17. *For $d = -5, -6, -10$ the ring of integers of $\mathbb{Q}(\sqrt{d})$ is not UFD and hence not ED.*

Proof. For each of these three values of d there exist an element in the ring of integers which has two distinct factorizations, namely in $\mathbb{Z}[\sqrt{-5}]$ we have

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

in $\mathbb{Z}[\sqrt{-6}]$ we have

$$-2 \cdot 3 = (\sqrt{-6}) \cdot (\sqrt{-6})$$

and in $\mathbb{Z}[\sqrt{-10}]$,

$$2 \cdot 7 = (2 + \sqrt{-10}) \cdot (2 - \sqrt{-10}).$$

We need to show that factors occurring in the above factorizations are irreducible in the respective ring of integers. This follows by considering the norm and solving easy diophantine equations. See [ST01] Theorem 4.10. □

Hence we have obtained our desired result of this section

Corollary 1.18. *Suppose d is square-free and $d < 0$. Then the ring of integers of $\mathbb{Q}(\sqrt{d})$ is a ED if and only if*

$$d = -1, -2, -3, -7, -11$$

and moreover the Euclidean function is given by the absolute value of the norm function.

Proof. Follows from Theorem 1.11, Theorem 1.15 and Theorem 1.17. □

1.2 Imaginary Quadratic UFDs

As mentioned at the beginning of this chapter that there are exactly nine imaginary quadratic fields whose ring of integers is a UFD. Since every ED is also a UFD we have already established that ring of integers of $\mathbb{Q}(\sqrt{d})$ is a UFD for $d = -1, -2, -3, -7, -11$. In this section our goal is to establish the existence of four more non-euclidean imaginary quadratic UFDs, namely for $d = -19, -43, -67$ and -163. To prove this we will first develop the theory of ideals and then the notion of class number of a number field.

1.2.1 Theory of Ideals

Definition 1.19. *Let R be a ring and K be it's field of fractions (for us K is a number field and $R = \mathfrak{O}_K$). Then an R-submodule M of K is called a* ***fractional ideal*** *of R if there exists $r \in R \setminus \{0\}$ such that $rM \subseteq R$.*

Note that $rM = \langle r \rangle M$ is an R-submodule of K contained in R. Hence rM is an ideal of R, say I. That is $M = r^{-1}I$. Hence, every fractional ideal of R is of the form $r^{-1}I$ for some ideal I of R and $r \in R$.

Now our goal is to show that the set of fractional ideals of $\mathfrak{O}_K$, for a number field K, has a group structure. For this we need to define the inverse of an ideal.

Definition 1.20. *Let R be a ring with field of fractions K and let $I \subseteq R$ be an ideal. Then we define the* ***inverse*** *of I as*

$$I^{-1} := \{x \in K : xI \subseteq R\}.$$

Proposition 1.21. *Let I and J be ideals of a ring R with field of fractions K, then the following holds:*

- $I \subseteq R \subseteq I^{-1} \subseteq K$.
- *If $I \subseteq J$ then $J^{-1} \subseteq I^{-1}$.*
- *I^{-1} is an R-submodule of K.*
- *If $I \neq 0$ then I^{-1} is a fractional ideal of R.*

Proof. All these directly follows from the definition of I^{-1}. For the last one, note that $cI^{-1} \subseteq R$ for any $c \in I \setminus \{0\}$. □

We want to know about the R-submodule II^{-1}. Note that, by the definition of I^{-1}, $II^{-1} \subseteq R$ and hence II^{-1} is an ideal of R. We will show that $II^{-1} = R$ for a special class of rings called Dedekind domains.

Definition 1.22. *Let R be a domain and K be it's field of fractions. Then R is called a* ***Dedekind domain*** *if it satisfies the following conditions:*

1. *R is Noetherian.*
2. *Every non-zero prime ideal of R is maximal.*
3. *R is integrally closed in K.*

Proposition 1.23. *Let K be a number field, then it's ring of integers $\mathfrak{O}_K$ is a Dedekind domain.*

Proof. Let I be an ideal of $\mathfrak{O}_K$. Now since $\mathfrak{O}_K$ is a finitely generated $\mathbb{Z}$-module and I is a submodule of $\mathfrak{O}_K$, we have that I is a finitely generated $\mathbb{Z}$-module. Hence, as an ideal of $\mathfrak{O}_K$, I is finitely generated.

Let P be a prime ideal of $\mathfrak{O}_K$ and let $a \in P$ be a non-zero element. Then

$$N := N(a) = a\prod_{i=2}^{n} \sigma_i(a) \in P.$$

where σ_i's are the distinct embeddings of K in $\mathbb{C}$ and σ_1=Id. Hence $\langle N\rangle \subseteq P$ and therefore $\mathfrak{O}_K/P$ is contained in $\mathfrak{O}_K/\langle N\rangle$ which is finite. Hence $\mathfrak{O}_K/P$ is a finite domain and hence a field. Therefore P is a maximal ideal.

The third condition is obvious. Hence $\mathfrak{O}_K$ is a Dedekind domain. □

Theorem 1.24. *Let R be a Dedekind domain and I be a non-zero ideal of R. Then $II^{-1} = R$.*

Proof. See [ST01] Theorem 5.6 (vi). □

Now one can show that the set of all fractional ideals of a Dedekind domain form a group. For any two fractional ideals $r^{-1}I$ and $s^{-1}J$, where $r, s \in R$ and I, J are ideals of R, define the group operation $\cdot$ as follows

$$(r^{-1}I) \cdot (s^{-1}J) := (rs)^{-1}IJ.$$

One can easily verify that this operation $\cdot$ is commutative, associative, R acts as identity and every $r^{-1}I$ has the inverse rI^{-1}.

Theorem 1.25. *Every non-zero proper ideal in a Dedekind domain can be written as a product of prime ideals.*

Proof. Let S be the set of all non-zero proper ideals of R which can not be written as a product of prime ideals. Suppose $S \neq \emptyset$. Then S has a maximal element (since R is Noetherian), say I_m. Then I_m is not prime. Now, since I_m is proper, it is contained in some maximal and hence prime ideal P of R. Then we have $R \subseteq P^{-1} \subseteq I_m^{-1}$ and multiplying by I_m we get

$$I_m \subseteq I_m P^{-1} \subseteq I_m I_m^{-1} \subseteq R.$$

Therefore $I_m P^{-1} \subseteq R$ and hence $I_m P^{-1}$ is an ideal of R. We claim that $I_m P^{-1} \neq I_m$, otherwise

$$I_m^{-1}(I_m P^{-1}) = I_m^{-1}(I_m) \quad \Rightarrow \quad P^{-1} = R \quad \Rightarrow \quad R = PP^{-1} = PR = P$$

but this can not happen since P is proper. Note that $I_m P^{-1} \neq 0$ (otherwise $I_m = 0$) and $I_m P^{-1} \neq R$ (otherwise I_m is prime) and also $I_m \subsetneq I_m P^{-1}$. Hence by maximality of I_m we have $I_m P^{-1} = P_1 \cdots P_r$ for some prime ideals $P_1, \ldots, P_r$. But this implies $I_m = P_1 \cdots P_r P$ which is a contradiction. Hence we conclude that $S = \emptyset$. □

In fact, this prime factorization is unique. To prove this we need the following propositions.

Definition 1.26. *Let R be a ring and I, J be ideals of R. Then we say I **divides** J (denoted by $I \mid J$) if there exist an ideal I' of R such that $J = II'$.*

Proposition 1.27. *Let I and J be ideals of a Dedekind domain R. Then $I \mid J$ if and only if $I \supseteq J$.*

Proof. If $I \mid J$ then clearly $I \supseteq J$. Now suppose $I \supseteq J$. Let $I' := JI^{-1}$ which is an R-submodule of $k(R)$. Then $J \subseteq I$ implies $I' = JI^{-1} \subseteq II^{-1} = R$ and hence I' is an ideal. But then $I \mid J$ because $II' = IJI^{-1} = J$. □

Proposition 1.28. *Let I, J and P be ideals of a Dedekind domain R and P is prime. Then, $P \mid IJ$ implies $P \mid I$ or $P \mid J$.*

Proof. Note that by the previous proposition it is enough to show the following

$$P \supseteq IJ \quad \Rightarrow \quad P \supseteq I \text{ or } P \supseteq J$$

Suppose there exist $a \in I \setminus P$ and $b \in J \setminus P$. Then $ab \in IJ \setminus P$ by the definition of prime ideal and we are done. □

Theorem 1.29. *The prime factorization of a non-zero proper ideal in a Dedekind domain is unique.*

Proof. Suppose a non-zero proper ideal I of a Dedekind domain has two prime factorizations namely

$$I = P_1 \cdots P_r = Q_1 \cdots Q_s.$$

W.l.o.g. assume $r \leq s$. Then $P_1 \mid Q_1 \cdots Q_s$. Now, since P_1 is prime, $P_1 \mid Q_i$ for some i; w.l.o.g. we can assume $P_1 \mid Q_1$. This implies $P_1 \supseteq Q_1$ and hence $P_1 = Q_1$ since Q_1 is maximal and P_1 is proper. Then multiplying byP_1^{-1} on the both sides we have

$$P_2 \cdots P_r = Q_2 \cdots Q_s.$$

Similarly we will have $P_2 = Q_2$. Continuing like this we get $P_i = Q_i$ for $i = 1, \ldots, r$. Now if $r = s$ then we are done. Otherwise if $r < s$ we have

$$R = Q_{r+1} \cdots Q_s \subseteq Q_{r+1}$$

but this can not happen since Q_{r+1} is prime and hence proper. □

1.2.2 Norm of an Ideal and Class Group

Definition 1.30. *Let K be a number field and I be a non-zero ideal of $\mathfrak{O}_K$. Then we define **norm** of I as*

$$N(I) := |\mathfrak{O}_K / I|.$$

Proposition 1.31. *Let K be a number field and I be a non-zero ideal of $\mathfrak{O}_K$. Then the following holds:*

- *N(I)=1 if and only if $I = \mathfrak{O}_K$.*
- *N(I) is finite.*
- *I divides the ideal generated by $N(I)$.*

Proof. The first one is trivial.

To show that $N(I) < \infty$, consider a non-zero element $\alpha \in I$. Then, if $\sigma_1 = \mathrm{Id}$, $\sigma_2, \ldots, \sigma_n$ are the distinct embeddings of K in $\mathbb{C}$, we have

$$N(\alpha) = \prod_{i=1}^{n} \sigma_i(\alpha) = \alpha \cdot \prod_{i=2}^{n} \sigma_i(\alpha) \in I$$

and hence $\langle N(\alpha) \rangle \subseteq I$. Then we have $\mathfrak{O}_K / I \subseteq \mathfrak{O}_K / \langle N(\alpha) \rangle$. Now $\mathfrak{O}_K / \langle N(\alpha) \rangle$ is finite because $\mathfrak{O}_K$ is a finitely generated $\mathbb{Z}$-module and $N(\alpha) \in \mathbb{Z}$. Hence we conclude that $N(I)$ is finite.

For the third, note that for any $x \in \mathfrak{O}_K$ we have $N(I) \cdot x = 0 \pmod{I}$ because the additive group of $\mathfrak{O}_K/I$ is of order $N(I)$. Putting $x = 1$ we get $N(I) \in I$. Hence $I \supseteq \langle N(I) \rangle$ and this implies $I \mid \langle N(I) \rangle$ since $\mathfrak{O}_K$ is a Dedekind domain. $\square$

Our next aim is to define the so called "class number" of a number field and relate it with the uniqueness of factorization in the ring of integers. One can think of this class number as a quantity which measures how far a number field is from being a UFD. We will show that the ring of integers of a number field is a UFD if and only if it's class number is 1.

Definition 1.32. *Let R be a Dedekind domain with field of fractions K. Then a fractional ideal $r^{-1}I$, where $r \in R$ and $I \subseteq R$ is an ideal, is called* ***principal*** *if I is a principal ideal.*

Definition 1.33. *Let K be a number with ring of integers $\mathfrak{O}_K$. Suppose $\mathcal{F}$ is the group of fractional ideals of $\mathfrak{O}_K$ and $\mathcal{P}$ be it's subgroup consisting of all principal fractional ideals. Then the quotient group $\mathcal{F}/\mathcal{P}$ is called the* ***class group*** *of K. The order of $\mathcal{F}/\mathcal{P}$ is called the* ***class number*** *of K and is denoted by $h(K)$.*

Let $\mathcal{F}$ and $\mathcal{P}$ are as in the above definition. Take a fractional ideal $M \in \mathcal{F}$ and let $[M] \in \mathcal{F}/\mathcal{P}$ be it's equivalence class. Then there exist an ideal of $\mathfrak{O}_K$ which is in the equivalence class of M. This follows from the fact that $M = r^{-1}I$ for some $r \in \mathfrak{O}_K$ and ideal $I \subseteq \mathfrak{O}_K$ and this implies $I = \langle r \rangle M$. In fact, for every fractional ideal M we can choose an ideal in it's equivalence class whose norm is bounded by a constant and this constant does not depend on M. We will use this fact to show that the class number of any number field is finite.

Definition 1.34. *Let K be a number field. Suppose $\{\alpha_1, \ldots, \alpha_n\}$ is a $\mathbb{Z}$-basis for $\mathfrak{O}_K$. Then the* ***discriminant*** *Δ of K is defined as*

$$\Delta := (\det[\sigma_i(\alpha_j)])^2$$

where $\sigma_1, \ldots, \sigma_n$ are the distinct embeddings of K in $\mathbb{C}$ and $n = [K : \mathbb{Q}]$.

Since we are particularly interested in quadratic fields we will compute their discriminant:

Proposition 1.35. *Let $K := \mathbb{Q}(\sqrt{d})$ be a quadratic field. Then the discriminant of K is given by*

$$\Delta = \begin{cases} 4d & \text{if } d \not\equiv 1 \pmod 4 \\ d & \text{if } d \equiv 1 \pmod 4 \end{cases}.$$

Proof. Note that the embeddings of K are given by the maps $\sqrt{d} \mapsto \sqrt{d}$ and $\sqrt{d} \mapsto -\sqrt{d}$. Also, Theorem 1.7 gives us the $\mathbb{Z}$-basis of $\mathfrak{O}_K$. Hence one can easily compute the discriminant directly from the definition to get the desired result $\square$

Theorem 1.36. *Let K be a number field of signature $\{s, t\}$ (i.e. s is the number of real embeddings and $2t$ is the number of non-real complex embeddings of K in $\mathbb{C}$). Let M be a fractional ideal of $\mathfrak{O}_K$. Then there exist an ideal I in the equivalence class (corresponding to the class group) of M such that*

$$N(I) \leq \left(\frac{4}{\pi}\right)^t \frac{n!}{n^n}\sqrt{\Delta}$$

where Δ is the discriminant of K and $n = s + 2t = [K : \mathbb{Q}]$.

Proof. The proof uses geometric techniques and Minkowski's convex body theorem. See [ST01] Corollary 10.3. $\square$

From now on we will denote the constant $\left(\frac{4}{\pi}\right)^t \frac{n!}{n^n}$ of the above theorem by M_{st} (called the Minkowski's constant).

Theorem 1.37. *Let K be a number field. Then the class number $h(K)$ of K is finite.*

Proof. First we show that the number of ideals of $\mathfrak{O}_K$ with a given norm is finite. Suppose $N \in \mathbb{N}$ and I be an ideal such that $N(I) = N$. Now, by prime factorization of ideals one can write $\langle N \rangle = P_1 \cdots P_r$, where P_i's are prime ideals. Then, since $I \mid \langle N(I) \rangle$, we have $I \mid P_1 \cdots P_r$. Now by uniqueness of prime factorization of I it follows that number of I can be finite.

Now, every equivalence class of fractional ideal contains an ideal of norm less than or equal to $M_{st}\sqrt{\Delta}$, where $\{s, t\}$ is the signature and Δ is the discriminant of K.Now since there are finitely many ideals with a given norm, we conclude that the number of such equivalence classes is finite. Hence the class number of K is finite. $\square$

Next, we will show that unique factorization in the ring of integers of a number field is equivalent to it's class group being trivial. This will directly follow from the following theorem.

Theorem 1.38. *Let K be a number field. Then the ring of integers $\mathfrak{O}_K$ is a UFD if and only if it is a PID.*

Proof. If $\mathfrak{O}_K$ is a PID then clearly it is a UFD. For the converse suppose $\mathfrak{O}_K$ is a UFD. Note that it is enough to show that every prime ideal of $\mathfrak{O}_K$ is principal because every ideal can be written as a product of prime ideals and product of principal ideals is again principal. Let P be a non-zero prime ideal. Then $P \mid \langle N(P) \rangle$. Now we can write $N(P) = \pi_1 \cdots \pi_s$, where π_i's are irreducibles in $\mathfrak{O}_K$ (this follows from the fact that during factorization norm of the factors keep decreasing and $N(a) = 1$ implies that a is a unit and hence the process of factorization into irreducibles ends in finitely many steps). Then $P \mid \langle \pi_1 \rangle \cdots \langle \pi_s \rangle$ and hence $P \mid \langle \pi_k \rangle$ for some k. Now π_k is irreducible implies π_k is prime (since $\mathfrak{O}_K$ is a UFD) and hence $\langle \pi_k \rangle$ is a prime ideal. But then by uniqueness of prime factorization we have $P = \langle \pi_k \rangle$ and hence P is principal □

Remark 1.39. *The above theorem is also true for any Dedekind domain but the same argument doesn't work since we don't have the notion of 'norm' in a general Dedekind domain. See [Mol10] Theorem 1.18 for the proof.*

Theorem 1.40. *Let K be a number field. Then $\mathfrak{O}_K$ is a UFD if and only if the class number $h(K) = 1$.*

Proof.

$$\begin{aligned} h(K) = 1 \quad &\text{iff} \quad \text{every fractional ideal of } \mathfrak{O}_K \text{ is principal} \\ &\text{iff} \quad \text{every ideal of } \mathfrak{O}_K \text{ is principal} \\ &\text{iff} \quad \mathfrak{O}_K \text{ is a UFD} \quad \text{(by previous theorem).} \end{aligned}$$

□

1.2.3 Imaginary Quadratic Fields of Class Number One

Now we proceed towards showing that for $d = -19, -43, -67, -163$ the quadratic field $\mathbb{Q}(\sqrt{d})$ has class number equal to one. This will give us the desired result mentioned at the beginning of this section.

Theorem 1.41. *Let K be a number field of signature $\{s, t\}$ and discriminant Δ. Suppose for every prime $p \in \mathbb{Z}$ such that $p \leq M_{st}\sqrt{\Delta}$ we have that every prime ideal of $\mathfrak{O}_K$ dividing $\langle p \rangle$ is principal. Then class number $h(K) = 1$.*

Proof. Let M be a fractional ideal of $\mathfrak{O}_K$, we want to show that M is principal. Now, by Theorem 1.36, there exist an ideal I of $\mathfrak{O}_K$ equivalent to M such that $N(I) \leq M_{st}\sqrt{\Delta}$. Note that, it is enough to show that I is principal, because if $I = \langle a \rangle$ and $M = r^{-1}\langle b \rangle I$ (since M is equivalent to I) for some $a, b, r \in \mathfrak{O}_K$ then $M = r^{-1}\langle ab \rangle$ which is principal. Suppose $N(I) = p_1 \cdots p_k$

where $p_i \in \mathbb{Z}$ are primes. Then, for each i, $p_i \leq N(I) \leq M_{st}\sqrt{\Delta}$ and hence by the hypothesis, for each i, every prime ideal dividing $\langle p_i \rangle$ is principal. Now, since $I \mid \langle N(I) \rangle$, we have $I \mid \langle p_1 \rangle \cdots \langle p_k \rangle$. Also we can write $I = P_1 \cdots P_s$ where P_j's are prime ideals. Then for each j,

$$P_j \mid \langle p_1 \rangle \cdots \langle p_k \rangle \quad \Rightarrow \quad P_j \mid \langle p_i \rangle \quad \text{for some } i$$

and so P_j is principal for all j. Hence I is principal. □

We will need the following theorem:

Theorem 1.42. *Let K be a number field of degree n. Suppose $\mathfrak{O}_K = \mathbb{Z}[\theta]$ for some $\theta \in \mathfrak{O}_K$. Given a prime $p \in \mathbb{Z}$, suppose the minimal polynomial $f \in \mathbb{Z}[x]$ of θ over $\mathbb{Q}$ gives rise to the factorization into irreducibles over $\mathbb{Z}_p$:*

$$\bar{f} = \bar{f}_1^{\,e_1} \cdots \bar{f}_r^{\,e_r}$$

where bar denotes the natural map $\mathbb{Z}[x] \to \mathbb{Z}_p[x]$. Then if $f_i \in \mathbb{Z}[x]$ is any polynomial mapping onto $\bar{f}_i$, then the ideal

$$P_i := \langle p \rangle + \langle f_i(\theta) \rangle$$

is prime and the prime factorization of $\langle p \rangle$ in $\mathfrak{O}_K$ is

$$\langle p \rangle = P_1^{e_1} \cdots P_r^{e_r}.$$

Proof. See [ST01] Theorem 10.1. □

Theorem 1.43. *Let $K := \mathbb{Q}(\sqrt{d})$ be a quadratic field of signature $\{s, t\}$ and discriminant Δ. Let $\mathfrak{O}_K = \mathbb{Z}[\theta]$ and $f(x) \in \mathbb{Z}[x]$ is the minimal polynomial of θ over $\mathbb{Q}$. Suppose, for all prime $p \in \mathbb{Z}$ such that $p \leq M_{st}\sqrt{\Delta}$, $\bar{f}(x) := f(x) \pmod{p} \in \mathbb{Z}_p[x]$ has no roots in $\mathbb{Z}_p$. Then K has class number equal to 1.*

Proof. Suppose for all prime $p \in \mathbb{Z}$ and $p \leq M_{st}\sqrt{\Delta}$, $\bar{f}(x)$ has no root in $\mathbb{Z}_p$ i.e. $\bar{f}(x)$ is irreducible in $\mathbb{Z}_p$ (since $\deg(\bar{f}) \leq 2$). Then by previous theorem $\langle p \rangle + \langle f(\theta) \rangle = \langle p \rangle$ is a prime ideal. Hence every prime ideal dividing $\langle p \rangle$ (which is $\langle p \rangle$ itself) is principal for all $p \leq M_{st}\sqrt{\Delta}$. Then by Theorem 1.41 we have $h(K) = 1$. □

Corollary 1.44. *The quadratic field $\mathbb{Q}(\sqrt{d})$ has class number 1 for*

$$d = -19, -43, -67, -163.$$

Proof. This follows by putting $d = -19, -43, -67, -163$ in Theorem 1.43 and verifying all the assumptions. Note that for quadratic fields all the parameters in the previous theorem, namely s, t, Δ, θ are known to us. □

Hence we have obtained our final result of this section;

Corollary 1.45. *Let $K := \mathbb{Q}(\sqrt{d})$. Then $\mathfrak{O}_K$ is a UFD for*

$$d = -1, -2, -3, -7, -11, -19, -43, -67, -163.$$

Proof. Follows from Corollary 1.18, Theorem 1.40 and Corollary 1.44. □

Remark 1.46. *For $d < 0$ and square-free, the converse of the above corollary is also true, that is this gives the complete list of imaginary quadratic fields whose ring of integers is a UFD. A proof was given by Heegner and Stark which uses modular functions, the proof has been discussed in [Kez12]. Another proof was given by Baker as an application of linear forms in logarithms.*

Chapter 2

Computing Class Number of Real Quadratic Fields

In this chapter we will address the problem of computing class number of a real quadratic field. Let $K = \mathbb{Q}(\sqrt{d})$, where $d > 0$, be a real quadratic field. All known unconditional algorithms for computing $h(\mathfrak{O}_K)$ have time complexity $O(d^{\alpha+\epsilon})$, where $\alpha < 1/2$ and under the assumption of Generalized Riemann Hypothesis the best known algorithm has time complexity $O(d^{1/5+\epsilon})$. In this chapter we will discuss two important methods for computing class number of a real quadratic fields: One is using the method of cycle counting, called the *Reduced Form Algorithm* and the other one is using the *Analytic Formula of Class Number*.

2.1 Reduced Form Algorithm

Let $d > 0$ be a square-free integer. Then, define r and ω as follows:

$$r := \begin{cases} 1 & \text{if } d \not\equiv 1 \pmod 4 \\ 2 & \text{if } d \equiv 1 \pmod 4 \end{cases} \quad \& \quad \omega := \frac{r - 1 + \sqrt{d}}{r}.$$

Now, if $K = \mathbb{Q}(\sqrt{d})$, then the free $\mathbb{Z}$-module $\mathfrak{O}_K$ has $\mathbb{Z}$-basis $\{1, \omega\}$ and the discriminant of K is given by $\Delta_K = 4d/r^2$.

Let I be an ideal of $\mathfrak{O}_K$, then I is a free $\mathbb{Z}$-submodule of $\mathfrak{O}_K$; in fact, we have the following

Theorem 2.1. *If I is an ideal of $\mathfrak{O}_K$ and $I \not\subseteq \mathbb{Z}$ then I has a $\mathbb{Z}$-basis $\{a, b + c\omega\}$ for some $a, b, c \in \mathbb{Z}$ such that $a \geq 0, c \geq 0, c \mid a$ and $c \mid b$.*

Proof. See [Jon61] Theorem 58, 59. □

Then, a is the least positive integer in I. We define $L(I) := a$.

Definition 2.2. *Let K be a real quadratic field. An ideal I of $\mathfrak{O}_K$ is said to be **primitive** if $L(I) = N(I)$.*

Proposition 2.3. *Let K be a real quadratic field. If an ideal $I \subseteq \mathfrak{O}_K$, with $\mathbb{Z}$-basis $\{a, b + c\omega\}$, is primitive then $c = 1$.*

Proof. Note that

$$N(I) = |\mathfrak{O}_K / I| = ac.$$

Then, $L(I) = N(I)$ implies $a = ac$ and therefore $c = 1$. □

Definition 2.4. *An ideal $I \subseteq \mathfrak{O}_K$ is said to be **reduced** if I is primitive and for all non-zero $\alpha \in I$, either $|\alpha| < L(I)$ or $|\overline{\alpha}| < L(I)$.*

Then, we have the following theorem which says that every ideal of $\mathfrak{O}_K$ is equivalent to a reduced ideal.

Theorem 2.5. *Let $I \subseteq \mathfrak{O}_K$ be an ideal. Then there exists a reduced ideal J and an element $\lambda \in I$ such that $\langle \lambda \rangle J = \langle L(J) \rangle I$.*

Proof. See [Wil85] Corollary 3.3. □

We also have the following theorem

Theorem 2.6. *If an ideal $I \subseteq \mathfrak{O}_K$ is reduced then $L(I) < \sqrt{\Delta_K}$. On the other hand, if $I \subseteq \mathfrak{O}_K$ is a primitive ideal and $L(I) < \sqrt{\Delta_K}/2$ then I is reduced.*

Proof. See [Wil85] Theorem 5.2 and Theorem 5.3. □

2.1.1 Algorithm for computing reduced ideals

Let $I = a\mathbb{Z} + (b + \omega)\mathbb{Z}$ be a primitive ideal of $\mathfrak{O}_K$, where $a, b \in \mathbb{Z}$. Let us denote $\phi = (b + \omega)/a$. Then ϕ can be written as

$$\phi = \frac{P + \sqrt{d}}{Q}, \quad \text{where } P, Q \in \mathbb{Z} \text{ and } rQ \mid d - P^2.$$

Also, ϕ has a continued fraction expansion, say $\phi = \langle q_0, q_1, q_2, \ldots \rangle$. Then, for each $k \geq 0$, we define $\phi_k = \langle q_k, q_{k+1}, \cdots \rangle$. Then we can write

$$\phi_k = \frac{P_k + \sqrt{d}}{Q_k}$$

where the integers P_k and Q_k are given by the following recurrence relations:

$$P_0 = P, \quad Q_0 = Q, \quad q_0 = \left[\frac{P + \sqrt{d}}{Q}\right] = [\phi_0]$$

and for $i \geq 0$,

$$\begin{aligned} P_{i+1} &= q_i Q_i - P_i \\ Q_{i+1} &= \frac{d - P_{i+1}^2}{Q_i} \\ q_{i+1} &= [\phi_{i+1}] = \frac{P_{i+1} + \sqrt{d}}{Q_{i+1}}. \end{aligned}$$

Now for each $k \geq 0$, define the ideal I_k as follows

$$I_k := \frac{Q_k}{r}\mathbb{Z} + \frac{P_k + \sqrt{d}}{r}\mathbb{Z}.$$

Then we have the following

Theorem 2.7. *With the above notations we have*

- *I_k is equivalent to I for all $k \geq 0$.*
- *I_k is reduced if*

$$k > \max\left\{2, 4 + \log\left(\frac{Q_0}{2\sqrt{d}}\right)\frac{1}{2\log\tau}\right\}$$

  where $\tau = (1 + \sqrt{5})/2$.

- *If I_m is reduced for some m then I_k is reduced for all $k \geq m$.*
- *Suppose m is the least integer for which I_m is reduced. Now, if an ideal J is equivalent to I and J is reduced, then $J = I_k$ for some $k \geq m$.*

Proof. See [MW92] Theorem 2.7, Theorem 2.8. □

Note that, since the continued fraction expansion of ϕ is periodic, we have that each ideal is equivalent to only finitely many reduced ideals. Also, finiteness of class number implies that there are finitely many equivalence classes. Hence we conclude that there are finitely many reduced ideals and this finiteness will allow us to count the reduced ideals. Moreover, we have the following theorem which allows us to find all the reduced ideals in a finite process.

Theorem 2.8. *If $I \subseteq \mathfrak{O}_K$ is a reduced ideal and $I = \frac{Q}{r}\mathbb{Z} + \frac{P+\sqrt{d}}{r}\mathbb{Z}$ then $0 < Q < 2\sqrt{d}$ and $0 < P < \sqrt{d}$.*

Proof. See [MW92] Section 2. □

Now, Theorem 2.7 implies that, the finite set of all reduced ideals decomposes into finitely many cycles (the operator concerned is the reduction operator) and each such cycle corresponds to a representative element of the class group. Hence, to compute the class number we need to count the number of such cycles. One way is to count the number of cycles by brute force; although it is possible to set up this cycle counting technique for computing class number such that it has time complexity $O(d^{0.5076+\epsilon})$. Under the assumption of GRH the time complexity can be reduced to $O(d^{1/2+\epsilon})$, this observation is due to H. W. Lenstra, Jr.

2.1.2 Connection with Quadratic Forms

For actual computational purposes to compute the class number by cycle counting one uses the the theory of quadratic forms which computationally is easy to handle (although, theoretically it seems unmotivated). Here we will establish a connection between ideals and quadratic forms which will enable us to apply the algorithm in the setting of quadratic forms.

Definition 2.9. *A bivariate function $f(x, y) = ax^2 + bxy + cy^2$ is called a **quadratic form** and is denoted by $f := (a, b, c)$. The quantity $\Delta(f) = b^2 - 4ac$ is called the **discriminant** of the quadratic form f.*

Let $K = \mathbb{Q}(\sqrt{d})$ be a real quadratic field and let $D = \Delta_K = 4d/r^2$, the discriminant of K. Now define the set $\mathcal{F}$ as

$$\mathcal{F} := \{f := (a, b, c) : \Delta(f) = D\}/\mathrm{PSL}_2(\mathbb{Z}).$$

Then, the following theorem gives a correspondence between $\mathcal{F}$, the class group Cl_K and the *narrow class group* Cl_K^+ which is defined as the set of all fractional ideals modulo the set of all principal fractional ideals whose generators have positive norm.

Theorem 2.10. *There is a bijection between the sets $\mathcal{F}$ and Cl_K^+ and there is a bijection between $\mathcal{F}$ and Cl_K when the quadratic forms (a, b, c) and $(-a, b, -c)$ are identified.*

Proof. See [Coh96] Proposition 5.6.1. □

Definition 2.11. *A quadratic form $f = (a, b, c)$ is called* ***reduced*** *if*

$$\left|\sqrt{D} - 2|a|\right| < b < \sqrt{D},$$

where $D = \Delta(f)$ is the discriminant of f.

With the above definition we will show that there is a one-one correspondence between the reduced quadratic forms and reduced ideals and this will give a motivation for the above definition. Define the set

$$\Gamma_\infty := \left\{ \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} : m \in \mathbb{Z} \right\} \subseteq \mathrm{SL}_2(\mathbb{Z}).$$

Then $\Gamma_\infty \cong (\mathbb{Z}, +)$. Note that, Γ_∞ acts on the set $\{f := (a, b, c) : \Delta(f) = D\}$ by the following operation

$$\begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \cdot (a, b, c) = (a, b + 2am, c + bm + am^2).$$

Let us define the set F as

$$F := \{f := (a, b, c) : \Delta(f) = D\}/\Gamma_\infty$$

and let S be the set of all fractional ideals modulo the multiplicative group of rationals $\mathbb{Q}^*$ (which, in fact, is same as the class group). Then, there is an isomorphism $\psi \colon S \times \mathbb{Z}/2\mathbb{Z} \to F$ (for the proof refer to [Coh96] Theorem 5.2.4). Then we have the following

Theorem 2.12. *An ideal $I \subseteq \mathfrak{O}_K$ is reduced if and only if the quadratic form (a, b, c) is reduced, where $\psi(I, s) = [(a, b, c)]$.*

Proposition 2.13. *If (a, b, c) is a reduced quadratic form then*

$$|a| < \sqrt{D}, \quad b < \sqrt{D}, \quad |c| < \sqrt{D} \quad \text{and} \quad |a| + |c| < \sqrt{D}.$$

Also, (a, b, c) is a reduced quadratic form if and only if

$$\left|\sqrt{D} - 2|c|\right| < b < \sqrt{D},$$

where D as usual is the discriminant of (a, b, c).

Proof. See [Coh96] Proposition 5.6.3. □

Next we define the *reduction operator* ρ analogous to the reduction operator for ideals. Given $a, b \in \mathbb{Z}$ define $r(a, b)$ to be the unique integer r such that

$$r \equiv b \pmod{2a} \quad \& \quad \begin{cases} -|a| < r \leq |a| & \text{if } |a| > \sqrt{D} \\ \sqrt{D} - 2|a| < r < \sqrt{D} & \text{if } |a| < \sqrt{D} \end{cases}.$$

Now define ρ as follows

$$\rho(a, b, c) := \left(c, r(-b, c), \frac{r(-b, c)^2 - D}{4c} \right).$$

Then, one can show that $(a, b, c) \sim \rho(a, b, c)$ modulo $\mathrm{PSL}_2(\mathbb{Z})$ and that ρ is a permutation.

Now, we will describe the cycle counting technique for computing the class number using the language of quadratic forms. The proof of the following algorithm can be found in [Coh96] Proposition 5.6.6. We start with a quadratic form and keep applying the reduction operator. Eventually we will reach a reduced form after at most $2 + \log(|c|/\sqrt{D})$ steps and once we reach a reduced form we will keep on getting reduced forms. Then the set of all reduced forms will decompose into finitely many cycles under the reduction operator. So we need to count the number of such cycles. Note that, Proposition 2.13 implies that the number of reduced forms is less than or equal to D; in fact, this number is of the order $O(\sqrt{D} \ln D)$. So we list all reduced forms and count the number of orbits under the permutation ρ and this will give us the narrow class number $h^+(K) = |\mathrm{Cl}_K^+|$. While counting, if we identify the quadratic forms (a, b, c) and $(-a, b, -c)$, then we will get the actual class number $h(K)$. The time complexity of this algorithm is $O(D)$.

2.2 Analytic formula of Class Number

In this section our goal is to compute the class number of a real quadratic field using the analytic formula of class number. For that we need to introduce the zeta function of a number field, known as the Dedekind zeta function.

Definition 2.14. *Let K be a number field. We define the **Dedekind zeta function** of K as*

$$\zeta_K(s) := \sum_{n=1}^{\infty} \frac{a_K(n)}{n^s},$$

where $s \in \mathbb{C}$ and $a_K(n)$ is the number of ideals of $\mathfrak{O}_K$ of norm n.

Proposition 2.15. *We have that*

$$\zeta_K(s) = \sum_I \frac{1}{N(I)^s},$$

where the sum ranges over all ideals of $\mathfrak{O}_K$.

Proof. Immediate from the definition. □

Theorem 2.16. $\zeta_K(s)$ *converges absolutely for* $\mathrm{Re}(s) > 1$.

Proof. Let $\sigma = \mathrm{Re}(s) > 1$. Then it is enough to show that the partial sums

$$S_x := \sum_{N(I)\leq x} \frac{1}{N(I)^\sigma}$$

are bounded. We can write

$$\begin{aligned} S_x &\leq \prod_{N(P)\leq x} \left(1 + \frac{1}{N(P)^\sigma} + \frac{1}{N(P)^{2\sigma}} + \cdots\right) \\ &= \prod_{N(P)\leq x} \left(1 - \frac{1}{N(P)^\sigma}\right)^{-1} \end{aligned}$$

where P ranges over all prime ideals.

Claim. *If* P *is a prime ideal then* $N(P) = p^f$ *for some* $f \in \mathbb{N}$ *and a unique prime number* p.

Suppose $N(P)$ has prime factorization $N(P) = p_1 \cdots p_r$. Then we have $P \mid \langle p_1\rangle \cdots \langle p_r\rangle$. Now, if $P \mid \langle p_i\rangle$ and $P \mid \langle p_j\rangle$ for some distinct primes p_i, p_j, then $P \mid \langle 1\rangle$ (since $\gcd(p_i, p_j) = 1$) and this implies $P \supseteq \mathfrak{O}_K$ which is clearly a contradiction. Therefore $P \mid \langle p\rangle$ for some unique prime p. Then we have $N(P) \mid N(\langle p\rangle) = p^n$, where $n = [K : \mathbb{Q}]$. Hence $N(P) = p^m$ for some $m \leq n$ and we have established our claim.

Claim. *Fix a prime* p. *Then the number of prime ideals* P *such that* $N(P)$ *is a power of* p *is less than or equal to* $[K : \mathbb{Q}]$.

Let P be such a prime ideal. Then $P \mid N(P)$ implies $P \mid p$. Suppose the ideal $\langle p\rangle$ factorizes into prime ideals as

$$\langle p\rangle = P_1^{e_1} \cdots P_r^{e_r}. \qquad (*)$$

Then P has to be one of the P_is. Now, for each i, $P_i \mid \langle p \rangle$ which implies $N(P_i) \mid N(\langle p \rangle)$. Therefore we have $N(P_i) = p^{f_i}$ for some $f_i \in \mathbb{N}$. Now, taking norm on the both side of the equation $(*)$ we get

$$p^{[K:\mathbb{Q}]} = p^{e_1 f_1} \cdots p^{e_r f_r}$$

which implies $\sum_{i=1}^{r} e_i f_i = [K : \mathbb{Q}]$. Hence, we conclude that $r \leq [K : \mathbb{Q}]$, since $e_i, f_i \geq 1$ for each i. So we have proved the claim.

Now coming back to S_x, using the above claims we can write

$$\begin{aligned} S_x &\leq \prod_{N(P) \leq x} \left(1 - \frac{1}{p^\sigma}\right)^{-1} \\ &\leq \prod_{p \leq x} \left(1 - \frac{1}{p^\sigma}\right)^{-[K:\mathbb{Q}]} \\ &= \left(\prod_{p \leq x} \left(1 - \frac{1}{p^\sigma}\right)\right)^{-[K:\mathbb{Q}]} \end{aligned}$$

and this implies S_x is bounded, since the Riemann zeta function

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

is absolutely convergent for all $s \in \mathbb{C}$ such that $\mathrm{Re}(s) > 1$. □

Analogous to the Euler's product formula for Riemann zeta function, we have the following for Dedekind zeta function

Proposition 2.17. *If* $\mathrm{Re}(s) > 1$ *then we have*

$$\zeta_K(s) = \prod_{P} \left(1 - \frac{1}{N(P)^s}\right)^{-1}$$

where P ranges over all prime ideals.

Proof. This follows directly from Proposition 2.15 by writing every ideal as a product of prime ideals. □

Let K be a real quadratic field. Now our goal is to show that $(s-1)\zeta_K(s)$ extends to an analytic function in the region $\mathrm{Re}(s) > 1/2$. We will need the following result

Theorem 2.18. *Let K be a real quadratic field. Then we have*

- $a_K(n) = \sum_{\delta|n} \left(\frac{\Delta_K}{\delta}\right)$
- $\left|\sum_{n\leq x} \left(\frac{\Delta_K}{n}\right)\right| \leq \Delta_K$

where $\left(\frac{\cdot}{\cdot}\right)$ is the Kronecker symbol.

Proof. See [ME04] Exercise 10.2.5, 10.2.7. □

Lemma 2.19. *Let $g, h\colon \mathbb{N} \to \mathbb{N}$ and suppose f is defined as*

$$f(n) := \sum_{\delta|n} g(\delta) h\left(\frac{n}{\delta}\right).$$

Now define the functions $G, H\colon \mathbb{R} \to \mathbb{N}$ as follows

$$G(x) := \sum_{n\leq x} g(n) \quad \text{and} \quad H(x) := \sum_{n\leq x} h(n).$$

Then for any real number $y > 0$,

$$\sum_{n\leq x} f(n) = \sum_{\delta\leq y} g(\delta) H\left(\frac{x}{\delta}\right) + \sum_{\delta<\frac{x}{y}} h(\delta) G\left(\frac{x}{\delta}\right) - G(y) H\left(\frac{x}{y}\right).$$

Proof. Note that

$$\begin{aligned}
\sum_{n\leq x} f(n) &= \sum_{\delta e\leq x} g(\delta) h(e) \\
&= \sum_{\substack{\delta e\leq x \\ \delta\leq y}} g(\delta) h(e) + \sum_{\substack{\delta e\leq x \\ \delta> y}} g(\delta) h(e) \\
&= \sum_{\delta\leq y} g(\delta) H\left(\frac{x}{\delta}\right) + \sum_{e\leq\frac{x}{y}} h(e) \left\{ G\left(\frac{x}{e}\right) - G(y) \right\} \\
&= \sum_{\delta\leq y} g(\delta) H\left(\frac{x}{\delta}\right) + \sum_{e\leq\frac{x}{y}} h(e) G\left(\frac{x}{e}\right) - G(y) H\left(\frac{x}{y}\right).
\end{aligned}$$

□

In the next theorem we will give a linear bound on the number of ideals of $\mathfrak{O}_K$, for a quadratic field K, whose norm is less than or equal to some fixes number. This bound will be used later to get an analytic continuation of $(s-1)\zeta_K(s)$ for $\mathrm{Re}(s) > 1/2$.

Theorem 2.20. *Let K be a real quadratic field and $a_K(n)$ denotes the number of ideals in $\mathfrak{O}_K$ of norm n. Then,*

$$\sum_{n\leq x} a_K(n) = cx + O(\sqrt{x}),$$

where $c = \sum_{\delta=1}^{\infty} \left(\frac{\Delta_K}{\delta}\right) \frac{1}{\delta}$.

Proof. In Lemma 2.19 we take

$$g(\delta) = \left(\frac{\Delta_K}{\delta}\right), \quad h(\delta) = 1, \quad \text{and} \quad y = \sqrt{x}.$$

Note that we have

$$|G(x)| = \left|\sum_{n\leq x} \left(\frac{\Delta_K}{n}\right)\right| \leq \Delta_K \qquad (*)$$

by Theorem 2.18 and $H(x) = [x]$. Then by Lemma 2.19 we have

$$\begin{aligned}
\sum_{n\leq x} a_K(n) &= \sum_{\delta\leq\sqrt{x}} \left(\frac{\Delta_K}{\delta}\right) \left[\frac{x}{\delta}\right] + \sum_{\delta<\sqrt{x}} 1 \cdot G\left(\frac{x}{\delta}\right) - G(\sqrt{x})[\sqrt{x}] \\
&= \sum_{\delta\leq\sqrt{x}} \left(\frac{\Delta_K}{\delta}\right) \left[\frac{x}{\delta}\right] + O(\sqrt{x}) \qquad (\text{ by } (*)) \\
&= \sum_{\delta\leq\sqrt{x}} \left(\frac{\Delta_K}{\delta}\right) \frac{x}{\delta} + O(\sqrt{x}).
\end{aligned}$$

Now,

$$\begin{aligned}
\sum_{\delta\leq\sqrt{x}} \left(\frac{\Delta_K}{\delta}\right) \frac{1}{\delta} &= \sum_{\delta=1}^{\infty} \left(\frac{\Delta_K}{\delta}\right) \frac{1}{\delta} - \sum_{\delta>\sqrt{x}} \left(\frac{\Delta_K}{\delta}\right) \frac{1}{\delta} \\
&= c + O\left(\frac{1}{\sqrt{x}}\right)
\end{aligned}$$

and this implies $\sum_{n\leq x} a_K(n) = cx + O(\sqrt{x})$. □

Note that in the above theorem we have assumed that c is a real constant and for that we need to show that the series $\sum_{\delta=1}^{\infty} \left(\frac{\Delta_K}{\delta}\right) \frac{1}{\delta}$ converges. This is an immediate corollary of the following theorem.

Theorem 2.21. *Let $\{a_n\}$ be a sequence of complex numbers and*

$$A(x) := \sum_{n \leq x} a_n = O(x^\delta)$$

for some $\delta \geq 0$. Then $\sum_{n=1}^{\infty} \frac{a_n}{n^s}$ converges for $\mathrm{Re}(s) > \delta$ and

$$\sum_{n=1}^{\infty} \frac{a_n}{n^s} = s \int_1^{\infty} \frac{A(x)}{x^{s+1}}\, dx$$

Proof. For any $N > 1$ we have

$$\begin{aligned}
\sum_{n=1}^{N} \frac{a_n}{n^s} &= \sum_{n=1}^{N} \frac{A(n) - A(n-1)}{n^s} \\
&= \frac{A(N)}{N^s} + \sum_{n=1}^{N-1} A(n) \left(\frac{1}{n^s} - \frac{1}{(n+1)^s} \right) \\
&= \frac{A(N)}{N^s} + \sum_{n=1}^{N-1} A(n) \left(s \int_n^{n+1} \frac{dx}{x^{s+1}} \right) \\
&= \frac{A(N)}{N^s} + s \sum_{n=1}^{N-1} \int_n^{n+1} \frac{A(n)}{x^{s+1}}\, dx \\
&= \frac{A(N)}{N^s} + s \sum_{n=1}^{N-1} \int_n^{n+1} \frac{A(x)}{x^{s+1}}\, dx \\
&= \frac{A(N)}{N^s} + s \int_1^{N} \frac{A(x)}{x^{s+1}}\, dx.
\end{aligned}$$

Now taking limit $N \to \infty$ we have, $A(N)/N^s \to 0$ since $\mathrm{Re}(s) > \delta$ and $A(x) = O(x^\delta)$ and also $\int_1^{\infty} A(x)/x^{s+1}\, dx$ converges for the same reason. Therefore the series converges for $\mathrm{Re}(s) > \delta$ and

$$\sum_{n=1}^{\infty} \frac{a_n}{n^s} = s \int_1^{\infty} \frac{A(x)}{x^{s+1}}\, dx.$$

□

Corollary 2.22. *Let K be a quadratic field. Then the series $\sum_{\delta=1}^{\infty} \left(\frac{\Delta_K}{\delta}\right) \frac{1}{\delta}$ converges.*

Proof. In the previous theorem put $a_n = \left(\frac{\Delta_K}{n}\right)$ and $s = 1$. Now, we have

$$A(x) = \sum_{n \le x} \left(\frac{\Delta_K}{n}\right) = O(1)$$

by Theorem 2.18. Hence, using previous theorem, we conclude that the series $\sum_{n=1}^{\infty} \left(\frac{\Delta_K}{n}\right) \frac{1}{n}$ converges. □

Now we will prove the analytic continuation of $(s-1)\zeta_K(s)$ which is one of the important theorem of this section.

Theorem 2.23. *Let K be a real quadratic field. Then $(s-1)\zeta_K(s)$ extends to an analytic function for* $\mathrm{Re}(s) > 1/2$.

Proof. We have $\zeta_K(s) = \sum_{n=1}^{\infty} \frac{a_K(n)}{n^s}$ and

$$A_K(x) := \sum_{n \le x} a_K(n) = cx + O(\sqrt{x})$$

where c is a constant defined as in Theorem 2.20. Then by Theorem 2.21,

$$\begin{aligned}
\zeta_K(s) &= s \int_1^{\infty} \frac{A_K(x)}{x^{s+1}}\, dx \\
&= s \int_1^{\infty} \frac{cx + O(\sqrt{x})}{x^{s+1}}\, dx \\
&= cs \int_1^{\infty} \frac{1}{x^s}\, dx + s \int_1^{\infty} \frac{O(\sqrt{x})}{x^{s+1}}\, dx \\
&= \frac{cs}{s-1} + s \int_1^{\infty} \frac{O(\sqrt{x})}{x^{s+1}}\, dx \qquad (*)
\end{aligned}$$

which holds for $\mathrm{Re}(s) > 1$. Now the integral $\int_1^{\infty} O(\sqrt{x})/x^{s+1}\, dx$ converges for $\mathrm{Re}(s) + 1 - 1/2 > 1$, i.e. for $\mathrm{Re}(s) > 1/2$. Hence the RHS of $(*)$ is analytic for $\mathrm{Re}(s) > 1/2$ and $s \neq 1$. Therefore the equation $(*)$ is valid for $\mathrm{Re}(s) > 1/2$ and $s \neq 1$ and this implies $(s-1)\zeta_K(s)$ is analytic for $\mathrm{Re}(s) > 1/2$. □

Remark 2.24. *In fact, for any number field K, $(s-1)\zeta_K(s)$ extends to an entire function for all $s \in \mathbb{C}$. This was conjectured by Dedekind in 1877 and was proved by Hecke in 1917.*

Note that the equation $(*)$ of the preceding theorem implies that $\zeta_K(s)$ has a simple pole at $s = 1$. One can compute the residue of this pole by computing the limit $\lim_{s \to 1}(s-1)\zeta_K(s)$, the value of this residue is given by the following theorem. Note that taking limit $s \to 1$ is possible because the function $(s-1)\zeta_K(s)$ extends to an analytic function in some neighborhood around $s = 1$ by the previous theorem.

Theorem 2.25. *Let K be a real quadratic field. Then*

$$\lim_{s\to 1}(s-1)\zeta_K(s) = \frac{2R_K}{\sqrt{\Delta_K}} \cdot h_K$$

where $R_K := \log \varepsilon_0$ is the regulator (where ε_0 is the fundamental unit of $\mathfrak{O}_K$, i.e. the smallest unit greater than 1) and h_K, Δ_K are respectively class number and discriminant of K.

Proof. see [Jan96] Chapter 4, Theorem 2.12. □

Our next goal is to compute the limit $\lim_{s\to 1}(s-1)\zeta_K(s)$. For this, we will introduce Dirichlet L-function.

Definition 2.26. *Let K be a quadratic field. A map $\chi\colon \mathbb{N} \to \{-1, 0, 1\}$ is called a **quadratic character** if $\chi(1) = 1$, for all prime p*

$$\chi(p) = \begin{cases} +1 & \text{if } \langle p\rangle = P_1P_2 \quad (p \text{ decomposes}) \\ -1 & \text{if } \langle p\rangle = P \quad (p \text{ remains prime}) \\ 0 & \text{if } \langle p\rangle = P^2 \quad (p \text{ ramifies}) \end{cases}$$

where P, P_is are prime ideals and for any $n = \prod_{i=1}^{r} p_i$, where p_is are prime,

$$\chi(n) = \prod_{i=1}^{r} \chi(p_i).$$

Definition 2.27. *The **Dirichlet L-function** $L(s, \chi)$, where χ is a quadratic character and $s \in \mathbb{C}$, is defined as*

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

Then, we will show that $\zeta_K(s) = \zeta(s)L(s, \chi)$ where ζ is the Riemann zeta function.

Lemma 2.28. *Let K be a quadratic field. Then the function $a_K(n)$, which is the number of of ideals of $\mathfrak{O}_K$ of norm n, is multiplicative. That is*

$$\gcd(m, n) = 1 \quad \Rightarrow \quad a_K(mn) = a_K(m)a_K(n).$$

Proof. Let $I \subseteq \mathfrak{O}_K$ be an ideal such that $N(I) = mn$. Then $I \mid \langle N(I)\rangle$ implies $I \mid \langle m\rangle\langle n\rangle$. Let us write $\langle m\rangle = \prod P_i$ and $\langle n\rangle = \prod P'_j$ as product of prime ideals. Now $P_i \neq P'_j$ for all i, j (otherwise, there exists a i such

that $P_i \mid \langle m \rangle$ and $P_i \mid \langle n \rangle$ which implies $P_i \mid \langle m \rangle + \langle n \rangle = \mathfrak{O}_K$ but this is a contradiction). Now, since $I \mid \prod P_i \cdot \prod P'_j$, we have

$$I = \left(\prod_k P_{i_k} \right) \left(\prod_l P'_{j_l} \right) := J_1 J_2$$

where J_is are uniquely determined. Now from $J_1 \mid \langle m \rangle$ and $J_2 \mid \langle n \rangle$ we get $N(J_1) \mid m^2$ and $N(J_2) \mid n^2$. Also, we have $N(J_1)N(J_2) = N(J_1 J_2) = mn$ and $\gcd(m, n) = 1$. Therefore we get $N(J_1) = m$ and $N(J_2) = n$. Hence, for all ideal I such that $N(I) = mn$, there exist unique J_1, J_2 such that $N(J_1) = m$ and $N(J_2) = n$. We conclude that $a_K(mn) = a_K(m)a_K(n)$. □

Theorem 2.29. *Let K be a quadratic field. Then,*

$$a_K(n) = \sum_{\delta \mid n} \chi(\delta).$$

Proof. We will first compute $a_K(p)$ and $a_K(p^r)$ for prime p. Then we will use the multiplicative property of $a_K(n)$ to compute it for general n.

Claim. *For prime p, we have $a_K(p) = 1 + \chi(p)$.*

Suppose I is an ideal such that $N(I) = p$. Then I is a prime ideal. Now, we know $I \mid \langle N(I) \rangle = \langle p \rangle$. If $\langle p \rangle$ decomposes then there are two choices for I i.e. $a_K(p) = 2$. If $\langle p \rangle$ remains prime then $I = \langle p \rangle$, which is not possible since $N(\langle p \rangle) = p^2$; therefore $a_K(p) = 0$. Finally, if $\langle p \rangle$ ramifies then there is exactly one choice for I; hence in this case $a_K(p) = 1$. Hence we have shown that $a_K(p) = 1 + \chi(p)$.

Claim. *For prime power p^r, we have $a_K(p^r) = 1 + \chi(p) + \cdots + \chi(p)^r$.*

Suppose I is an ideal such that $N(I) = p^r$. We need to consider three cases separately, namely when $\chi(p) = 1, -1$ and 0.

Case 1: Suppose p decomposes into prime ideals, say $\langle p \rangle = P_1 P_2$. Then we have $I \mid \langle N(I) \rangle = P_1^r P_2^r$. Now, $N(P_1)N(P_2) = p^2$ and note that norm of any prime ideal dividing $\langle p \rangle$ is either p or p^2, hence we get $N(P_1) = N(P_2) = p$. Note that, I must be of the form $P_1^i P_2^j$ for some $i, j \geq 0$. Then $N(I) = p^i p^j$. Therefore $i + j = r$ and hence there are $r+1$ choices for I, i.e. $a_K(p^r) = r+1$. On the other hand, $\chi(p) = 1$ implies $1 + \chi(p) + \cdots + \chi(p)^r = r + 1$.

Case 2: Suppose p remains prime ideal, say $\langle p \rangle = P$. Then from the fact $I \mid \langle N(I) \rangle = P^r$, we get $I = P^i$ for some $i \geq 0$. Also, note that $N(P) = p^2$. Therefore $N(I) = p^{2i}$ and this implies $r = 2i$. Hence, $a_K(p^r) = 1$ if r is even and 0 otherwise. On the other hand, note that, since $\chi(p) = -1$, we have $1 + \chi(p) + \cdots + \chi(p)^r = 1$ if r is even and 0 if r is odd.

Case3: Suppose p ramifies, say $\langle p\rangle = P^2$. Then from $I \mid \langle N(I)\rangle = P^{2r}$ we get $I = P^i$ for some $i \geq 0$. Also, $p^2 = N(P)^2$ implies $N(P) = p$. Therefore $N(I) = p^i$ and this implies $i = r$. Hence there is exactly one choice for I, i.e. $a_K(p^r) = 1$. On the other hand, $\chi(p) = 0$ implies $1 + \chi(p) + \cdots + \chi(p)^r = 1$.

Hence we have established the claim for all the three cases.

Now suppose $n \in \mathbb{N}$ and has prime factorization $n = p_1^{r_1} \cdots p_k^{r_k}$. Then using the above claims we have:

$$\begin{aligned} a_K(n) &= \prod_{i=1}^{k} a_K(p_i^{r_i}) \\ &= \prod_{i=1}^{k} (1 + \chi(p_i) + \cdots + \chi(p_i)^{r_i}) \\ &= \sum_{\delta \mid n} \chi(\delta). \end{aligned}$$

The last equality follows by expanding the product and using multiplicative property of χ. □

Now we state the theorem which connects Dedekind zeta function, Riemann zeta function and Dirichlet L-function.

Theorem 2.30. *Let K be a quadratic field. Then, for* $\mathrm{Re}(s) > 1$,

$$\zeta_K(s) = \zeta(s) \cdot L(s, \chi).$$

Proof. The proof follows directly from definitions and the previous theorem. We have

$$\begin{aligned} \zeta(s) \cdot L(s, \chi) &= \left(\sum_{l=1}^{\infty} \frac{1}{l^s} \right) \left(\sum_{m=1}^{\infty} \frac{\chi(m)}{m^s} \right) \\ &= \sum_{n=1}^{\infty} \frac{1}{n^s} \left(\sum_{\delta \mid n} \chi(\delta) \right) \\ &= \sum_{n=1}^{\infty} \frac{a_K(n)}{n^s} \end{aligned}$$

and the last quantity is equal to $\zeta_K(s)$. □

Let K be a real quadratic field and χ be quadratic character of K. Then, from previous theorem we have

$$(s-1)\zeta_K(s) = (s-1)\zeta(s) \cdot L(s, \chi).$$

In the next theorem we will show that $(s-1)\zeta(s)$ extends to an analytic function for $\mathrm{Re}(s) > 0$ and $\lim_{s\to 1}(s-1)\zeta(s) = 1$. Then, taking limit $s \to 1^+$ on the both side of the above equation, we have

$$\frac{2R_K h_K}{\sqrt{\Delta_K}} = L(1,\chi), \qquad \text{(recall Theorem 2.25)}$$

which implies $h_K = L(1,\chi)\cdot\sqrt{\Delta_K}/2R_K$.

Theorem 2.31. *$(s-1)\zeta(s)$ extends to an analytic function for $\mathrm{Re}(s) > 0$ and $\lim_{s\to 1}(s-1)\zeta(s) = 1$.*

Proof. Note that $\zeta(s) = \sum_{n=1}^{\infty} a_n/n^s$, where $a_n := 1$ for all $n \geq 1$. Then $A(x) := \sum_{n\leq x} a_n = [x]$. Then, by Theorem 2.21, for $\mathrm{Re}(s) > 1$ we have

$$\begin{aligned}
\zeta(s) &= s\int_1^{\infty} \frac{[x]}{x^{s+1}}\,dx \\
&= s\int_1^{\infty} \frac{x-\{x\}}{x^{s+1}}\,dx \\
&= \frac{s}{s-1} - s\int_1^{\infty} \frac{\{x\}}{x^{s+1}}. \qquad (*)
\end{aligned}$$

Note that the integral $\int_1^{\infty}\{x\}/x^{s+1}\,dx$ converges for $\mathrm{Re}(s) > 0$ (since $\{x\} = O(1)$). Therefore RHS of $(*)$ is analytic for $\mathrm{Re}(s) > 0$ and $s \neq 1$. Hence, the equation $(*)$ is valid for $\mathrm{Re}(s) > 0$ and $s \neq 1$ and this implies $(s-1)\zeta(s)$ is analytic for $\mathrm{Re}(s) > 0$. Then, multiplying both side of $(*)$ by $(s-1)$ and taking $s \to 1$ we have

$$\begin{aligned}
\lim_{s\to 1}(s-1)\zeta(s) &= \lim_{s\to 1} s - \lim_{s\to 1} s(s-1)\int_1^{\infty} \frac{\{x\}}{x^{s+1}}\,dx \\
&= 1 - 0 \\
&= 1.
\end{aligned}$$

□

Corollary 2.32. *Let K be a real quadratic field. Then the class number h_K of K is given by*

$$h_K = \frac{\sqrt{\Delta_K}}{2R_K}\cdot L(1,\chi)$$

where R_K denotes the regulator of K.

Proof. Follows from Theorem 2.25, Theorem 2.30 and Theorem 2.31. □

Let K be a real quadratic field and χ is a character of K . Then, by the above corollary, to compute class number h_K we need to compute two things, namely, $L(1,\chi)$ and the regulator R_K. We will discuss methods for computing these one by one.

2.2.1 Computing $L(1, \chi)$

First we will discuss a method which computes $L(1, \chi)$ by approximating a certain infinite product. We need the following proposition.

Proposition 2.33. *Let K be a real quadratic field and χ be a quadratic character for K. Then, for all $n \in \mathbb{N}$,*

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

Proof. From Theorem 2.18 and Theorem 2.29 we have

$$\sum_{\delta|n} \chi(\delta) = \sum_{\delta|n} \left(\frac{\Delta_K}{\delta}\right).$$

If n is a prime, say $n = p$, then the above equation implies

$$1 + \chi(p) = 1 + \left(\frac{\Delta_K}{p}\right)$$

that is $\chi(p) = \left(\frac{\Delta_K}{p}\right)$. Hence the proposition is true for $n = p$. Now if n is a product of primes, say $n = p_1 \cdots p_r$, then using induction on r one can easily show that $\chi(n) = \left(\frac{\Delta_K}{n}\right)$. □

Hence by Proposition 2.33 and definition of $L(s, \chi)$ we have

$$L(1, \chi) = \sum_{n=1}^{\infty} \frac{1}{n} \left(\frac{\Delta_K}{n}\right).$$

Now, the RHS of the above equation converges by Corollary 2.22. Hence, using Euler's product formula we can write

$$L(1, \chi) = \prod_p \left(1 - \frac{1}{\left(\frac{\Delta_K}{p}\right)}\right)^{-1}$$

where p ranges over all primes. We compute $L(1, \chi)$ by approximating the above infinite product.

Next we we will discuss another way to compute $L(1, \chi)$. We will show that the infinite sum $\sum_{n=1}^{\infty} \chi(n)/n$ can be written as a finite sum and hence making the computation of $L(1, \chi)$ possible theoretically, although computationally evaluating this finite sum is not efficient. We will need the following:

Proposition 2.34. *Let Δ_K be the discriminant of a real quadratic field K. Then, for $m, n \in \mathbb{N}$,*

$$m \equiv n \pmod{\Delta_K} \quad \Rightarrow \quad \left(\frac{\Delta_K}{m}\right) = \left(\frac{\Delta_K}{n}\right).$$

Proof. See [ME04] Exercise 7.6.16. □

For the rest of this section K will denote a real quadratic field and χ will denote the quadratic character of K. By previous proposition we have

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \sum_{m=0}^{\Delta_K - 1} \chi(m) \sum_{\substack{n=1 \\ n \equiv m \pmod{\Delta_K}}}^{\infty} \frac{1}{n^s}$$

where $\chi(0) := 0$. We write the above equation as

$$L(s,\chi) = \sum_{m=0}^{\Delta_K - 1} \chi(m) \sum_{n=1}^{\infty} \frac{c_m(n)}{n^s}$$

where

$$c_m(n) := \begin{cases} 0 & \text{if } n \not\equiv m \pmod{\Delta_K} \\ 1 & \text{if } n \equiv m \pmod{\Delta_K} \end{cases}.$$

We can write $c_m(n)$ as follows

$$c_m(n) = \frac{1}{\Delta_K} \sum_{j=0}^{\Delta_K - 1} \gamma^{(m-n)j}$$

where γ is a primitive Δ_Kth root of unity. Then this implies

$$\begin{aligned} L(s,\chi) &= \sum_{m=0}^{\Delta_K - 1} \chi(m) \sum_{n=1}^{\infty} \frac{1}{\Delta_K} \sum_{j=0}^{\Delta_K - 1} \gamma^{(m-n)j} \\ &= \frac{1}{\Delta_K} \sum_{j=0}^{\Delta_K - 1} \left(\sum_{m=0}^{\Delta_K - 1} \chi(m) \gamma^{mj} \right) \sum_{n=1}^{\infty} \frac{\gamma^{-nj}}{n^s}. \end{aligned}$$

The inner sum $\mathfrak{g}(\chi, \gamma^j) := \sum_{m=0}^{\Delta_K - 1} \chi(m)\gamma^{mj}$ is called the *Gauss sum*. Let us denote the outer sum as $\ell(s) := \sum_{n=1}^{\infty} \frac{\gamma^{-nj}}{n^s}$, then we want to evaluate $\ell(1)$. To evaluate $\ell(1)$ we will use complex logarithm. For $z \in \mathbb{C}$ define

$$\log z := \ln |z| + i \operatorname{Arg} z, \qquad -\pi < \operatorname{Arg} z \leq \pi.$$

Also for $|z| < 1$ we have

$$\log(1-z) = -\left(z + \frac{z^2}{2} + \cdots + \frac{z^n}{n} + \cdots\right);$$

note that the RHS converges for $|z| < 1$.

Proposition 2.35. *The series $\sum_{n=1}^{\infty} \frac{z^n}{n}$ converges when z is a root of unity and $z \neq 1$.*

Proof. We have the Dirichlet series $\sum_{n=1}^{\infty} \frac{z^n}{n^s} = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$, where $a_n := z^n$ for each n. Now

$$\sum_{n \leq x} a_n = \sum_{n \leq x} z^n = O(1)$$

since $z \neq 1$ and z is a root of unity. Therefore by Theorem 2.21 the sum $\sum_{n=1}^{\infty} \frac{z^n}{n^s}$ converges for $\mathrm{Re}(s) > 0$. Hence the series $\sum_{n=1}^{\infty} \frac{z^n}{n}$ converges. □

Therefore if z is a root of unity and $z \neq 1$ then

$$-\sum_{n=1}^{\infty} \frac{z^n}{n} = \log(1-z) = \ln|1-z| + i\,\mathrm{Arg}(1-z).$$

Putting $z = \gamma^{-j}$ in the above equation we get $\ell(1) = -\log(1-\gamma^{-j})$ and this implies

$$L(1,\chi) = -\frac{1}{\Delta_K} \sum_{j=0}^{\Delta_K - 1} \mathfrak{g}(\chi, \gamma^j) \log(1-\gamma^{-j}).$$

Proposition 2.36. *Gauss sum has the following property*

$$\mathfrak{g}(\chi, \gamma^j) = \chi(j)\mathfrak{g}(\chi, \gamma)$$

Proof. First assume $\gcd(j, \Delta_K) = 1$. Then $\exists j'$ such that $jj' \equiv 1 \pmod{\Delta_K}$. Then,

$$\begin{aligned}
\mathfrak{g}(\chi, \gamma^j) &= \sum_{m=0}^{\Delta_K - 1} \chi(m)\gamma^{mj} \\
&= \sum_{m=0}^{\Delta_K - 1} \chi(mj')\gamma^{(mj')j} \\
&= \sum_{m=0}^{\Delta_K - 1} \chi(m)\chi(j')\gamma^{m} \\
&= \chi(j')\mathfrak{g}(\chi, \gamma).
\end{aligned}$$

Now $\chi(j)\chi(j') = 1$ implies $\chi(j) = \chi(j')$. Therefore $\mathfrak{g}(\chi, \gamma^j) = \chi(j)\mathfrak{g}(\chi, \gamma)$. Now, if $\gcd(j, \Delta_K) > 1$ then $\mathfrak{g}(\chi, \gamma^j) = 0$ (see [Jan96] Corollary C.6 for the proof) and hence $\mathfrak{g}(\chi, \gamma^j) = \chi(j)\mathfrak{g}(\chi, \gamma)$. □

Therefore using the above proposition, we can write $L(1, \chi)$ as

$$L(1, \chi) = -\frac{1}{\Delta_K}\mathfrak{g}(\chi, \gamma) \sum_{j=0}^{\Delta_K - 1} \chi(j) \log(1 - \gamma^{-j}).$$

Our next task is to evaluate the sum $S := \sum_{j=0}^{\Delta_K - 1} \chi(j) \log(1 - \gamma^{-j})$. Let us fix the value of γ as $\gamma = e^{2\pi i/\Delta_K}$. For $0 < j < \Delta_K$ we write

$$\begin{aligned} 1 - \gamma^{-j} &= \gamma^{-j/2}(\gamma^{j/2} - \gamma^{-j/2}) \\ &= 2i\gamma^{-j/2} \sin \frac{\pi j}{\Delta_K} \\ &= 2 \sin \frac{\pi j}{\Delta_K} \cdot e^{i(\pi/2 - \pi j/\Delta_K)} \end{aligned}$$

and hence $\log(1 - \gamma^{-j}) = 2 \sin \frac{\pi j}{\Delta_K} + i\left(\frac{\pi}{2} - \frac{\pi j}{\Delta_K}\right)$. We have,

$$\begin{aligned} S &= \sum_{j=0}^{\Delta_K - 1} \chi(j) \log(1 - \gamma^{-j}) \qquad (*) \\ &= \sum_{j=0}^{\Delta_K - 1} \chi(\Delta_K - j) \log\left(1 - \gamma^{(\Delta_K - j)}\right). \end{aligned}$$

Proposition 2.37. *We have* $\chi(\Delta_K - j) = \chi(j)$

Proof. Note that $\chi(\Delta_K - 1) = 1$ (see [Jan96] Chapter 6, Theorem 4.2 for proof). Then,

$$\begin{aligned} \chi(\Delta_K - j) &= \chi(j(\Delta_K - 1) - (j - 1)\Delta_K) \\ &= \chi(j(\Delta_K - 1)) \\ &= \chi(j)\chi(\Delta_K - 1) \\ &= \chi(j). \end{aligned}$$

□

Using the above proposition we have

$$S = \sum_{j=0}^{\Delta_K - 1} \chi(j) \log(1 - \gamma^{j}). \qquad (**)$$

Now, adding the equations $(*)$ and $(**)$ we get

$$2S = \sum_{j=0}^{\Delta_K - 1} \chi(j) \left(\log(1 - \gamma^{-j}) + \log(1 - \gamma^{j}) \right).$$

Now, since $\overline{1 - \gamma^j} = 1 - \gamma^{-j}$, we get

$$\log(1 - \gamma^{-j}) + \log(1 - \gamma^{j}) = 2 \ln |1 - \gamma^{-j}| = 2 \ln \left(2 \sin \frac{\pi j}{\Delta_K} \right).$$

Hence we have obtained

$$\begin{aligned} S &= \sum_{j=1}^{\Delta_K - 1} \chi(j) \ln \left(2 \sin \frac{\pi j}{\Delta_K} \right) \\ &= \left(\sum_{j=1}^{\Delta_K - 1} \chi(j) \right) \ln 2 + \sum_{j=1}^{\Delta_K - 1} \chi(j) \ln \left(\sin \frac{\pi j}{\Delta_K} \right) \\ &= \sum_{j=1}^{\Delta_K - 1} \chi(j) \ln \left(\sin \frac{\pi j}{\Delta_K} \right). \end{aligned}$$

Therefore we have

$$L(1, \chi) = -\frac{1}{\Delta_K} \mathfrak{g}(\chi, \gamma) \sum_{j=1}^{\Delta_K - 1} \chi(j) \ln \left(\sin \frac{\pi j}{\Delta_K} \right).$$

Now since $\chi(j) = \chi(\Delta_K - j)$ and $\sin(\pi - \alpha) = \sin \alpha$, we obtain the following expression

$$L(1, \chi) = -\frac{2}{\Delta_K} \mathfrak{g}(\chi, \gamma) \sum_{\substack{j=1 \\ \gcd(j, \Delta_K) = 1}}^{[\Delta_K / 2]} \chi(j) \ln \left(\sin \frac{\pi j}{\Delta_K} \right).$$

Note that, $L(1, \chi) \in \mathbb{R}$ and $L(1, \chi) > 0$ (since $L(1, \chi) = 2 R_K h_K / \sqrt{\Delta_K} > 0$); hence $L(1, \chi) = |L(1, \chi)|$. Now taking modulus on the both side of the previous equation and using the fact that $|\mathfrak{g}(\chi, \gamma)| = \sqrt{\Delta_K}$ (see [Jan96] Theorem C.7 for proof) we obtain the following finite sum for $L(1, \chi)$:

$$L(1, \chi) = \frac{2}{\sqrt{\Delta_K}} \left| \sum_{\substack{j=1 \\ \gcd(j, \Delta_K) = 1}}^{[\Delta_K / 2]} \chi(j) \ln \left(\sin \frac{\pi j}{\Delta_K} \right) \right|.$$

2.2.2 Computing the Regulator

Here we will describe a method for computing the regulator of quadratic field; the idea of this algorithm is due to Shanks. For this we need to define a new notion of "distance" between ideals.

Let K be a real quadratic field and $\alpha \in K^*$. Then we define

$$\operatorname{Log} \alpha := \frac{1}{2} \log \left| \frac{\sigma(\alpha)}{\alpha} \right| = \log |\sigma(\alpha)| - \frac{1}{2} \log |N(\alpha)|$$

where $\sigma(\alpha)$ is the conjugate of α. Then we have the following proposition:

Proposition 2.38. *We have the following facts:*

- *The map* $\operatorname{Log}\colon K^* \to (\mathbb{R}, +)$ *is a group homomorphism and the kernel of this map is* $\mathbb{Q}^* \cup \mathbb{Q}^* \sqrt{\Delta_K}$.
- *If* $a \in \mathfrak{O}_K^\times$ *then* $\operatorname{Log} a = -\log |a|$.
- $\operatorname{Log} \varepsilon_0 = -R_K$, *where* ε_0 *is the fundamental unit and* R_K *is the regulator of* K.
- *If* $\alpha \in K^*$ *then* $\operatorname{Log} \sigma(\alpha) = -\operatorname{Log} \alpha$.

Proof. Directly follows from definition of Log. □

Let $\mathcal{P}$ denote the set of all principal fractional ideals of $\mathfrak{O}_K$, where K is a real quadratic field. Then we define the map $d\colon \mathcal{P} \to \mathbb{R}/R_K\mathbb{Z}$ given by,

$$d(\alpha \mathfrak{O}_K) = \operatorname{Log} \alpha + R_K \mathbb{Z}, \qquad \text{where } \alpha \in K^*.$$

This d is interpreted as the distance between $\mathfrak{O}_K$ and $\alpha\mathfrak{O}_K$.

Proposition 2.39. *The map* d *is well defined.*

Proof. Suppose $\alpha\mathfrak{O}_K = \beta\mathfrak{O}_K$ for some $\alpha, \beta \in K^*$. Then we have $\alpha = \beta b$ and $\beta = \alpha a$ for some $a, b \in \mathfrak{O}_K$. Then this implies $\alpha/\beta \in \mathfrak{O}_K^\times$, that is $\alpha/\beta = \pm\varepsilon_0^n$ for some n. Then we have

$$\operatorname{Log} \alpha - \operatorname{Log} \beta = \operatorname{Log} \frac{\alpha}{\beta} = \operatorname{Log} \varepsilon_0^n = -nR_K$$

and hence we conclude that $d(\alpha\mathfrak{O}_K) = d(\beta\mathfrak{O}_K)$. □

Next we extend d to define "distance" between any two ideals in the same equivalence class. Suppose I and J are ideals of $\mathfrak{O}_K$ such that $J \sim I$, say $J = (\alpha\mathfrak{O}_K)I = \alpha I$ for some $\alpha \in K^*$. Then we define

$$d(I, J) := \operatorname{Log} \alpha + R_K\mathbb{Z}.$$

Suppose I and J are two ideals both equivalent to some ideal I_0. Then $I = \alpha I_0$ and $J = \beta I_0$ for some $\alpha, \beta \in K^*$. Then we have

$$d(I, J) = \operatorname{Log} \beta - \operatorname{Log} \alpha + R_K\mathbb{Z}.$$

Note that every ideal class contains a reduced ideal (as defined in Section 2.1). Let I_0 be a reduced ideal in some ideal class and suppose for $i \geq 0$, I_is are the reduced ideals equivalent to I_0. Then $I_0, I_1, \ldots, I_{l-1}, I_l = I_0$ forms a cycle for some l. Note that $I_{i+1} = \rho(I_i)$ for each i, where ρ is the reduction operator. Then, for each i, $I_{i+1} \sim I_i$, that is we have $I_{i+1} = \gamma_i I_i$ for some $\gamma_i \in K^*$. Then,

$$I_i = \left(\prod_{j=0}^{i-1} \gamma_j\right) I_0 = \alpha_i I_0,$$

where $\alpha_i := \prod_{j=0}^{i-1} \gamma_j$. Therefore, for each i, we can write

$$d(I_i, I_{i+1}) = \operatorname{Log} \gamma_i + R_K\mathbb{Z}$$

and

$$d(I_0, I_i) = \operatorname{Log} \alpha_i + R_K\mathbb{Z}.$$

Then with these notations we have the following theorem.

Theorem 2.40. *For each $i \geq 0$ we have:*

- *If (a_i, b_i, c_i) is the quadratic form equivalent to the ideal I_i, then*

$$\operatorname{Log} \gamma_i = \frac{1}{2} \log \frac{\sqrt{\Delta_K} + b_i}{\sqrt{\Delta_K} - b_i} > 0.$$

- $\operatorname{Log} \alpha_i < \operatorname{Log} \alpha_{i+1}$.
- $\operatorname{Log} \alpha_l = R_K$.

Proof. See [BV07] Lemma 10.1.5. □

Moreover, we have the following facts:

Theorem 2.41. *For each $i \geq 0$,*

- $\frac{1}{\sqrt{\Delta_K}} < \operatorname{Log} \gamma_i < \frac{1}{2} \log \Delta_K$
- $\operatorname{Log} \gamma_i + \operatorname{Log} \gamma_{i+1} > \log 2$
- $l \leq 1 + 2R_K / \log 2$.

Proof. Note that the third one is an immediate corollary of the second. For the proof of first two, see [BV07] Lemma 10.1.6. □

We started with a reduced ideal I_0 in some equivalence class and we got the following cycle of reduced ideals

$$\mathcal{C} = \{I_0, I_1, \ldots, I_{l-1}, I_l = I_0\}$$

such that $d(I_0, I_j) = \operatorname{Log} \alpha_j + R_K\mathbb{Z}$ for each j. Note that the length of this cycle $\mathcal{C}$ is equal to R_K by Theorem 2.40. So, to compute the regulator, we need to compute the length of the cycle $\mathcal{C}$. We will start with $I_0 = \mathfrak{O}_K$. We will give the set $\mathcal{C}$ a group like structure. For $I, I' \in \mathcal{C}$ define

$$I \star I' := \rho^k(II'),$$

where k is the least i such that $\rho^i(II')$ is reduced. Note that $(\mathcal{C}, \star)$ satisfies commutativity, associativity and has identity I_0. So, the problem of computing R_K is almost like computing the order of the "group" $(\mathcal{C}, \star)$. Note that, we also need to compute the distance from I_0 at each stage to compute the total length R_K of the cycle $\mathcal{C}$, i.e. one needs to keep track of the distance from I_0. This is done by the following formulas:

$$d(I_0, II') = d(I_0, I) + d(I_0, I')$$

(note that $I, I' \sim I_0$ implies $II' \sim I_0$) and

$$d(I_0, I \star I') = d(I_0, II') + d(II', I \star I').$$

One can show that $d(II', I \star I') < 2 \ln \sqrt{\Delta_K}$, then this says that $d(I_0, I \star I')$ is easy to compute since $d(II', I \star I')$ is very small. The first statement of Theorem 2.40 is used for computing $d(\cdot)$.

So, to compute the regulator, we need to compute the "order" of $(\mathcal{C}, \star)$ together with computing the distance from I_0 at each stage. One can compute this by brute force. Although, there is probabilistic algorithm by Shanks, called *Baby step, Giant step Algorithm*, which computes the regulator in time $O(\Delta_K^{1/4+\epsilon})$; this algorithm assumes some bounds on $|\mathcal{C}|$ and the algorithm depends on the correctness of those bounds, see [Coh96] Section 5.4.1 for more details.

There are also other algorithms to compute R_K. For example, one can try to compute the fundamental unit ε_0 directly, which is equivalent to finding the solutions of *Pell's Equation*. There are algorithms to find this solutions using the method of continued fractions, look at [Bur80] Section 13.5 for more details.

Chapter 3

Cohen-Lenstra Heuristics

In this chapter we will discuss Cohen-Lenstra Heuristics and it's connection with quadratic and number fields. There is a very promising set of conjectures given by Henri Cohen and Hendrik Lenstra Jr., according to which, a positive fraction of real quadratic fields has class number one. Note that, it is still an open question whether there are infinitely many real quadratic fields of class number one, which was conjectured by Gauss. So the conjectures of Cohen and Lenstra is definitely a big step towards Gauss' conjecture. We will discuss the Cohen-Lenstra probability model for imaginary quadratic fields, real quadratic fields and arbitrary number fields one by one.

3.1 Probabilistic model for Imaginary Quadratic Fields

We will denote the set of all finite abelian groups of odd order (i.e. groups with trivial 2-part) by $\mathcal{G}$ and for any fixed prime $p \neq 2$, $\mathcal{G}_p$ will denote the set of all finite abelian p-groups. Now, suppose X is a random variable taking values in $\mathcal{G}_p$. We will show that there is a natural probability distribution P which can be attached to X such that

$$P(X = G_p) \propto \frac{1}{|\mathrm{Aut}(G_p)|},$$

where $G_p \in \mathcal{G}_p$. We will explain why the above distribution is a natural distribution on $\mathcal{G}_p$ by showing that this distribution can be obtained from another natural measure called *Haar measure* which is a generalization of the Lebesgue measure to the topological groups. Once we have the above probability measure on $\mathcal{G}_p$, our next task will be to extend this probability measure to the class $\mathcal{G}$. Since any group can $G \in \mathcal{G}$ can be written as a

product of p-groups and since $\mathrm{Aut}(G_1 \times G_2) = \mathrm{Aut}(G_1) \times \mathrm{Aut}(G_2)$ for any two different p-groups G_1 and G_2, we might want to define P on $\mathcal{G}$ such that

$$P(X = G) \propto \frac{1}{|\mathrm{Aut}(G)|},$$

for any $G \in \mathcal{G}$. But, note that, the above definition does not give us a probability measure because,

$$\sum_{G \in \mathcal{G}} \frac{1}{|\mathrm{Aut}(G)|} \geq \sum_{p} \frac{1}{|\mathrm{Aut}(\mathbb{Z}/p\mathbb{Z})|} = \sum_{p} \frac{1}{p-1} = \infty$$

(where p ranges over all odd primes). Another way is to try to define P as follows: For $G \in \mathcal{G}$, such that $G = \prod_p G_p$ (where $G_p \in \mathcal{G}_p$), define

$$P(X = G) := \prod_{p} P(X = G_p),$$

but one can check that this again does not satisfy the properties of a measure. We will show how to extend this probability measure P from $\mathcal{G}_p$ to $\mathcal{G}$.

3.1.1 Cohen-Lenstra measure for p-groups

First we need the following lemma.

Lemma 3.1. *Fix a prime p. Suppose G_p is a finite abelian p-group and*

$$G_p = \prod_{i=1}^{k} (\mathbb{Z}/p^{e_i}\mathbb{Z})^{r_i}$$

for some $k \geq 0$, $e_1 > e_2 > \cdots > e_k > 0$ and $r_i \geq 0$. Then

$$|\mathrm{Aut}(G_p)| = \left(\prod_{i=1}^{k} \left(\prod_{s=1}^{r_i} (1 - p^{-s}) \right) \right) \left(\prod_{1 \leq i,j \leq k} p^{\min(e_i,e_j) r_i r_j} \right).$$

Proof. See [Len09] Theorem 1.2.10. □

Next we will prove the following theorem. The ideas used in the proof of this theorem will be will be necessary in many places of this chapter. So the reader is advised to go through this proof throughly.

Theorem 3.2. *For any fixed prime p,*

$$\sum_{G_p \in \mathcal{G}_p} \frac{1}{|\mathrm{Aut}(G_p)|} = \prod_{j=1}^{\infty} \left(1 - \frac{1}{p^j}\right)^{-1}.$$

Proof. The original proof, due to Cohen and Lenstra, of this theorem is long and uses zeta functions. We will present a more elegant and shorter combinatorial proof due to Hall.

Note that, the RHS can be written as a power series in $q := p^{-1}$:

$$\prod_{j=1}^{\infty}(1-p^{-j})^{-1} = \sum_{n=0}^{\infty} a_n q^n$$

where a_n is the number of partitions of size n. Now, note that there is an associated partition corresponding to every p-group and corresponding to every partition there is an associated p-group; this comes from writing p-groups uniquely as a product of cyclic groups. For example, if we write a p-group G_p as,

$$G_p = \prod_{i=1}^{k} \mathbb{Z}/p^{e_i}\mathbb{Z}$$

where $e_1 \geq e_2 \geq \cdots \geq e_k > 0$, then the associated partition λ is given by $\lambda = (e_1, e_2, \ldots, e_k)$. And, corresponding to every partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ the associated p-group G_λ is given by $G_\lambda = \prod_{i=1}^{k} \mathbb{Z}/p^{\lambda_i}\mathbb{Z}$. Note that, if $|\lambda|$ denotes the size of the partition λ, then the order of the p-group G_λ is given by $|G_\lambda| = p^{|\lambda|}$. Now, with the above notations, we need to show

$$\begin{aligned}\sum_{n=0}^{\infty} a_n q^n &= \sum_{G_p \in \mathcal{G}_p} \frac{1}{|\mathrm{Aut}(G_p)|} \\ &= \sum_{n=0}^{\infty} \sum_{\substack{G_\lambda \in \mathcal{G}_p \\ |\lambda|=n}} \frac{1}{|\mathrm{Aut}(G_\lambda)|}.\end{aligned}$$

Let $\lambda := (\lambda_1, \ldots, \lambda_l)$ be a partition of size n and suppose $\lambda' := (\lambda'_1, \ldots, \lambda'_m)$ is it's conjugate partition. Then, note that, in G_λ (as a product of cyclic groups), the factor $\mathbb{Z}/p^i\mathbb{Z}$ occurs exactly $\lambda'_i - \lambda'_{i+1}$ times (where $\lambda'_{m+1} := 0$). Then using Lemma 3.1 we can write

$$\begin{aligned}|\mathrm{Aut}(G_\lambda)| &= \left(\prod_{i=1}^{m}\left(\prod_{s=1}^{\lambda'_i-\lambda'_{i+1}}(1-p^{-s})\right)\right)\left(\prod_{1\leq i,j\leq m} p^{\min(i,j)(\lambda'_i-\lambda'_{i+1})(\lambda'_j-\lambda'_{j+1})}\right) \\ &= \left(\prod_{i=1}^{m}\left(\prod_{s=1}^{\lambda'_i-\lambda'_{i+1}}(1-p^{-s})\right)\right) p^{\sum_{1\leq i,j\leq m}\min(i,j)(\lambda'_i-\lambda'_{i+1})(\lambda'_j-\lambda'_{j+1})} \\ &= \left(\prod_{i=1}^{m}\left(\prod_{s=1}^{\lambda'_i-\lambda'_{i+1}}(1-p^{-s})\right)\right) p^{\sum_{i=1}^{m}(\lambda'_i)^2}.\end{aligned}$$

So, we need to show

$$\begin{aligned}
\sum_{n=0}^{\infty} a_n q^n &= \sum_{n=0}^{\infty} \sum_{\substack{G_\lambda \in \mathcal{G}_p \\ |\lambda|=n}} \frac{1}{|\mathrm{Aut}(G_\lambda)|} \\
&= \sum_{n=0}^{\infty} \sum_{\substack{G_\lambda \in \mathcal{G}_p \\ |\lambda|=n}} \left(\prod_{i=1}^{m} \left(\prod_{s=1}^{\lambda'_i - \lambda'_{i+1}} (1-p^{-s})^{-1} \right) \right) \left(\prod_{i=1}^{m} p^{-(\lambda'_i)^2} \right) \\
&= \sum_{n=0}^{\infty} \sum_{\substack{G_\lambda \in \mathcal{G}_p \\ |\lambda|=n}} \left(\prod_{i=1}^{m} \left(\prod_{s=1}^{\lambda'_i - \lambda'_{i+1}} (1-q^{s})^{-1} \right) \right) \left(\prod_{i=1}^{m} q^{(\lambda'_i)^2} \right) \\
&= \sum_{n=0}^{\infty} \sum_{\substack{G_\mu \in \mathcal{G}_p \\ |\mu|=n}} \left(\prod_{i=1}^{m} \left(\prod_{s=1}^{\mu_i - \mu_{i+1}} (1-q^{s})^{-1} \right) \right) \left(\prod_{i=1}^{m} q^{\mu_i^2} \right),
\end{aligned}$$

since $\mu := \lambda'$ varies over all partitions as λ varies over all partitions. Now, we have the following identity due to Euler,

$$\sum_{n=0}^{\infty} a_n q^n = \sum_{n=0}^{\infty} q^n \prod_{i=1}^{n} (1-q^i)^{-1}.$$

Hence, it is enough to show that

$$\begin{aligned}
q^n \prod_{i=1}^{n} (1-q^i)^{-1} &= \sum_{\substack{G_\mu \in \mathcal{G}_p \\ |\mu|=n}} \left(\prod_{i=1}^{m} \left(\prod_{s=1}^{\mu_i - \mu_{i+1}} (1-q^{s})^{-1} \right) \right) \left(\prod_{i=1}^{m} q^{\mu_i^2} \right) \\
&= \sum_{\substack{G_\mu \in \mathcal{G}_p \\ |\mu|=n}} \left(\prod_{i=1}^{m} \psi_{\mu_i, \mu_{i-1}-\mu_i}(q) \right) \left(\prod_{i=1}^{m} q^{\mu_i^2} \right),
\end{aligned}$$

where

$$\psi_{a,b}(q) := \frac{\prod_{i=1}^{a+b}(1-q^i)}{\prod_{i=1}^{a}(1-q^i)\prod_{i=1}^{b}(1-q^i)}, \quad \psi_{a,\infty}(q) := \frac{1}{\prod_{i=1}^{a}(1-q^i)}$$

and $\mu_0 := \infty$; note that, the coefficient of q^n in $\psi_{a,b}(q)$ is the number of partitions of n with hight at most a and width at most b. Now, to show the above equality, for each N, we will equate the coefficients of q^N on the both sides.

Note that, the coefficient of q^N on LHS is equal to the number of partitions of $N-n$ with greatest part (width) at most n, which is equal to the number of partitions of N with greatest part exactly equal to n. Let ν be a partition of N with greatest part equal to n; then, to each such ν we will associate a partition μ of size n on the RHS. Consider the conjugate ν' of ν and let D be the standard *Young diagram* of ν'. Note that ν' has hight equal to n. Now, define $\mu := (\mu_1, \ldots, \mu_m)$ as follows:

- Define μ_1 to be the largest integer such that $(\mu_1, \mu_1) \in D$.
- For $i \geq 2$, define μ_i to be the largest integer such that

$$(\mu_1 + \cdots + \mu_i, \mu_i) \in D.$$

(where $(i,j) \in D$ is defined as the block of D situated at the ith row from top and jth column from left). Then $|\mu| = n$. If M is the number of blocks outside the squares of size μ_i then $M = N - \mu_1^2 - \mu_2^2 - \cdots - \mu_m^2$. Let M_i be the number of blocks at the right of the block of size μ_i, i.e.

$$M_i := \left|\{(x,y) \in D : \mu_1 + \cdots + \mu_{i-1} < x < \mu_1 + \cdots + \mu_i,\ \mu_i < y\}\right|.$$

Then the blocks corresponding to M_i gives a partition of M_i of hight at most μ_i and width at most $\mu_{i-1} - \mu_i$ and hence this contributes to the coefficient of q^{M_i} in $\psi_{\mu_i, \mu_{i-1}-\mu_i}(q)$ on RHS. Note that $M = M_1 + \cdots + M_m$ which implies $M_1 + \cdots + M_m + \mu_1^2 + \cdots + \mu_m^2 = N$ and hence μ contributes to the coefficient of q^N on RHS.

Note that, the above construction can be reversed. Suppose μ is a partition which corresponds to the coefficient of q^N on RHS such that μ is specified by the numbers M_i, where $M_1 + \cdots + M_m + \mu_1^2 + \cdots + \mu_m^2 = N$, and partitions of M_i of hight at most μ_i and width at most $\mu_{i-1} - \mu_i$. Then we can construct the Young diagram D and construct the corresponding partition ν on LHS. Hence, we conclude that the coefficients of q^N on both sides are equal and this proves the theorem. □

Let p be a fixed prime . Then the above theorem will enable us to define the following probability measure on $\mathcal{G}_p$, the class of all p-groups.

Definition 3.3. *The **Cohen-Lenstra probability measure on** p**-groups** (or, the **local Cohen-Lenstra probability measure**), denoted by P_p, is defined as*

$$P_p(G_p) := \frac{1}{|\mathrm{Aut}(G_p)|} \prod_{j=1}^{\infty} \left(1 - \frac{1}{p^j}\right)$$

for each $G_p \in \mathcal{G}_p$.

Note that by Theorem 3.2,

$$\sum_{G_p \in \mathcal{G}_p} P_p(G_p) = 1$$

and hence P_p indeed defines a probability measure on $\mathcal{G}_p$. Next we will explain why this probability measure arises "naturally", by showing that this probability measure can be obtained from Haar measure.

First we need to ask what is the most natural way to generate random finite abelian p-groups. The first thought that comes to mind is to use the structure theorem for finite abelian p-groups. That is, choose randomly a decreasing sequence of natural numbers, say $e_1 \geq e_2 \geq \cdots \geq e_r$ and output the group $\prod_{i=1}^{r} \mathbb{Z}/p^{e_i}\mathbb{Z}$. But unfortunately, there is no way to choose even one single element randomly from $\mathbb{N}$. So this approach to generate random finite abelian p-groups does not work.

Next we will describe another natural way to generate random finite p-groups which at the end of the day will give us the local Cohen-Lenstra probability measure. The idea is to view finite abelian groups as free groups with finite number of generators and then impose relations on these generators. Let G be a finite abelian group; then we can write $G = \langle g_1, \ldots, g_r \rangle$, where $r = \operatorname{rank}(G)$, together with relations of the form $e_1 g_1 + \cdots + e_r g_r = 0$, where $e_i \in \mathbb{Z}$. Since $|G| < \infty$, we must have r relations. That is, we can write G as $G = \mathbb{Z}^r / \operatorname{im}(A)$, where $A \colon \mathbb{Z}^r \to \mathbb{Z}^r$ is an $r \times r$ matrix. Now, since we want to generate p- groups, we replace $\mathbb{Z}^r$ by $\mathbf{Z}_p^r$, where $\mathbf{Z}_p$ is the ring of p-adic integers, and take A to be an $r \times r$ matrix over $\mathbf{Z}_p$. Then, any finite abelian p-group can be generated this way and this is given by the following theorem.

Theorem 3.4. *Any finite abelian p-group G_p can be written as $G_p = \mathbf{Z}_p^r / \operatorname{im}(A)$, where $A \colon \mathbf{Z}_p \to \mathbf{Z}_p$ is an $r \times r$ matrix.*

Proof. Note that G_p is a $\mathbf{Z}_p$ module with respect to the following multiplication: For $g \in G_p$, where $o(g) = p^m$, define

$$(a_0 + a_1 p + a_2 p^2 + \cdots) g := a_0 g + a_1 (pg) + \cdots + a_{m-1}(p^{m-1} g).$$

Now, $|G_p| < \infty$ implies G_p is finitely generated as a $\mathbf{Z}_p$ module; suppose G_p is generated by $g_1, \ldots, g_r$. Then, we have a surjection $\pi \colon \mathbf{Z}_p^r \to G_p$ of $\mathbf{Z}_p$ modules such that $\pi(e_i) = g_i$ for $i = 1, \ldots, r$, where $e_1, \ldots, e_r$ is the canonical basis for $\mathbf{Z}_p^r$. Then $\mathbf{Z}_p^r / \ker(\pi) \cong G_p$. Now $|G_p| < \infty$ implies that the $\mathbf{Z}_p$ submodule $\ker(\pi)$ has rank r. Let $x_1, \ldots, x_r$ be a basis for $\ker(\pi)$ and let $A \in \mathbf{Z}_p^{r \times r}$ be such that $Ae_i = x_i$ for $i = 1, \ldots, r$. Then $\operatorname{im}(A) = \ker(\pi)$ and therefore we conclude that $\mathbf{Z}_p^r / \operatorname{im}(A) \cong G_p$. □

Then by above theorem, any $G_p \in \mathcal{G}_p$ is completely specified by a matrix in $\mathbf{Z}_p^{r\times r}$. To generate G_p randomly, we need to choose $A \in \mathbf{Z}_p^r$ randomly, i.e. we need to choose each element of A randomly from $\mathbf{Z}_p$ (note that, the distribution of A is the product of the distributions of each element in the matrix, since they are independent). Now, choosing an element randomly from $\mathbf{Z}_p$ is possible because, $\mathbf{Z}_p$ with the p-adic topology is a locally compact Hausdorff topological space and hence $\mathbf{Z}_p$ is equipped with a natural probability distribution coming from the Haar measure. We state the following theorem regarding Haar measure:

Theorem 3.5. *Let G be a locally compact Hausdorff topological space. Then there exists, up to a positive multiplicative constant, a unique countably additive non-trivial measure μ on the Borel subsets of G satisfying:*

1. *μ is left translation invariant, i.e., $\mu(gE) = \mu(E)$ for all $g \in G$ and for all Borel sets E.*
2. *$0 < \mu(K) < \infty$ for any compact subspace K of G.*
3. *μ is outer regular on Borel sets E, i.e.*
$$\mu(E) = \inf\{\mu(U) : E \subseteq U \text{ and } U \text{ is open }\}.$$
4. *μ is inner regular on open sets E, i.e.*
$$\mu(E) = \sup\{\mu(K) : K \subseteq E \text{ and } K \text{ is compact }\}.$$

Proof. Look at any standard text. □

We will use the first and second property in the proof of the next theorem. For us, we will specify $\mu(G) = 1$, since we want μ to be a probability measure; in this case μ is uniquely determined.

Coming back to p-groups, we have that, for any $G_p \in \mathcal{G}_p$, $G_p \cong \mathbf{Z}_p^r/\operatorname{im}(A)$ where $A \in \mathbf{Z}_p^{r\times r}$. We choose A randomly. Note that, we need A to be of full rank since $|G_p| < \infty$. So we should have that

$$\Pr(A \text{ does not have full rank}) = 0$$

where $\Pr(\cdot)$ denotes the probability measure coming from Haar measure. This is ensured by the following theorem.

Theorem 3.6. *Suppose A is a randomly chosen (w.r.t. Haar measure) matrix from $\mathbf{Z}_p^{n\times n}$. Then,*

$$\Pr(A \textit{ has full rank}) = 1$$

for all $n > 0$.

Proof. We will prove the theorem through many claims.

Claim. *$A \in \mathbf{Z}_p^{n\times n}$ has full rank if and only if there exists $e \geq 0$ such that $(p^e\mathbf{Z}_p)^n \subseteq \mathrm{im}(A)$.*

The 'only if' part is obvious. For the other direction, suppose for each i there exists $v_i \in \mathbf{Z}_p^n$ such that $Av_i = p^e e_i$, where e_i is the canonical basis with 1 at the ith position. Then $Av_i' = e_i$, where $v_i' := p^{-e}v_i \in \mathbf{Q}_p^n$. This implies that $\det(A) \in \mathbf{Q}_p \setminus \{0\}$. Now, since $A \in \mathbf{Z}_p^{n\times n}$, we have $\det(A) \in \mathbf{Z}_p \setminus \{0\}$. Therefore, A has full rank.

Claim. *For any $e' > e$, $(p^e\mathbf{Z}_p)^n \subseteq \mathrm{im}(A)$ iff $(p^e\mathbb{Z}/p^{e'}\mathbb{Z})^n \subseteq \mathrm{im}(A \mod p^{e'})$.*

The 'only if' part is obvious. For the other direction, we prove for $n = 1$ and this simply generalizes for any n. Take any

$$x := p^e(a_0 + a_1p + \cdots + a_rp^r + \cdots) \in (p^e\mathbf{Z}_p)^n.$$

Suppose $e + r = e'$. Then

$$\begin{aligned} x &= (p^ea_0 + \cdots + p^ea_{r-1}p^{r-1}) + p^r(p^ea_r + p^ea_{r+1}p + \cdots + p^ea_{2r-1}p^{r-1}) + \cdots \\ &= Ax_1 + p^r(Ax_2) + \cdots \qquad \text{(say)} \\ &= A(x_1 + p^rx_2 + p^{2r}x_3 + \cdots) \in \mathrm{im}(A). \end{aligned}$$

From the above claims we have

$$\begin{aligned} \Pr(A \text{ has full rank}) &= \sum_{e\geq 0} \Pr((p^e\mathbf{Z}_p)^n \subseteq \mathrm{im}(A)) \\ &\geq \Pr((p^e\mathbf{Z}_p)^n \subseteq \mathrm{im}(A)) \qquad \forall e \geq 0 \\ &= \Pr((p^e\mathbb{Z}/p^{e'}\mathbb{Z})^n \subseteq \mathrm{im}(A \mod p^{e'})) \qquad \forall e \geq 0, e' > e. \end{aligned}$$

Now, $p^e\mathbb{Z}/p^{e'}\mathbb{Z}$ is a discrete subgroup of $\mathbf{Z}_p$, since for any $x, y \in p^e\mathbb{Z}/p^{e'}\mathbb{Z}$ we have $|x - y|_p \geq p^{-e'}$.

Claim. *Haar measure on any discrete topological group (locally compact, Hausdorff) is equivalent to the counting measure.*

Let G be a discrete group and take any $g \in G$. Then $\{g\}$ is compact in G. Therefore $0 < \mu(\{g\}) < \infty$. Now for any other $h \in G$, we can write $g = (gh^{-1})h$. Then g is a left translation of h and this therefore implies that $\mu(\{g\}) = \mu(\{h\})$. Hence, for all $g \in G$, $\mu(\{g\}) = c$ for some constant $c > 0$, that is, μ is a counting measure.

So we can compute $\Pr(A \text{ has full rank})$ by counting the matrices $A \mod p^{e'}$ such that $(p^e\mathbb{Z}/p^{e'}\mathbb{Z})^n \subseteq \mathrm{im}(A)$. Let $N(e,e')$ denote the number of such matrices. Then one can directly compute $N(e,e')$ and verify that $N(e,e')/p^{n^2e'} \to 1$ as $e \to \infty$. See the paper [FW89] for reference. Now, we have

$$\begin{aligned}\Pr(A \text{ has full rank}) &\geq \Pr((p^e\mathbb{Z}/p^{e'}\mathbb{Z})^n \subseteq \mathrm{im}(A \mod p^{e'})) \\ &\geq N(e,e')/p^{n^2e'},\end{aligned}$$

since the total number of matrices over $\mathbb{Z}/p^{e'}\mathbb{Z}$ is $p^{n^2e'}$. Then, taking limit $e \to \infty$, we have that $\Pr(A \text{ has full rank}) = 1$. □

In the next theorem we show that the local Cohen-Lenstra measure can be obtained as a limiting distribution of the Haar measure and therefore it explains the naturality of this measure.

Theorem 3.7. *Suppose $G_p \in \mathcal{G}_p$ and $\mathrm{rank}(G_p) = r$. Let $A \in \mathbf{Z}_p^{n\times n}$ be a randomly chosen matrix w.r.t. Haar measure. Then*

$$\Pr(\mathbf{Z}_p^n/\mathrm{im}(A) \cong G_p) \to P_p(G_p) \qquad \text{as } n \to \infty,$$

where P_p is the local Cohen-Lenstra probability measure.

Proof. We will explicitly compute the probability $\Pr(\mathbf{Z}_p^n/\mathrm{im}(A) \cong G_p)$ and then take limit as $n \to \infty$. Let $\Gamma \subseteq \mathbf{Z}_p^n$ be a $\mathbf{Z}_p$-submodule such that $\mathbf{Z}_p^n/\Gamma \cong G_p$; fix such a Γ. We will compute $\Pr(\mathrm{im}(A) = \Gamma)$. Now, since Γ is free and $|\mathbf{Z}_p^n/\Gamma| < \infty$, we have $\mathrm{rank}(\Gamma) = n$. Let $\gamma_1, \ldots, \gamma_n$ be a $\mathbf{Z}_p$-basis for Γ. Let $A_0 \in \mathbf{Z}_p^{n\times n}$ be such that $A_0 e_i = \gamma_i$ for $i = 1, \ldots, n$, where e_is are the canonical basis for $\mathbf{Z}_p^n$; i.e. γ_is are the columns of A_0. Then, $\mathrm{im}(A_0) = \Gamma$.

Claim. $\{A \in \mathbf{Z}_p^{n\times n} : \mathrm{im}(A) = \Gamma\} = A_0 \cdot \mathrm{GL}_n(\mathbf{Z}_p)$.

Note that, for all $B \in \mathrm{GL}_n(\mathbf{Z}_p)$, $\mathrm{im}(A_0B) = \mathrm{im}(A_0) = \Gamma$ and this implies that $A_0 \cdot \mathrm{GL}_n(\mathbf{Z}_p) \subseteq \{A \in \mathbf{Z}_p^{n\times n} : \mathrm{im}(A) = \Gamma\}$. Conversely, let $A \in \mathbf{Z}_p^{n\times n}$ be such that $\mathrm{im}(A) = \Gamma$. Then there exists $v_i \in \mathbf{Z}_p^n$ such that $Av_i = \gamma_i$ for each $i = 1, \ldots, n$. This implies that $\{v_1, \ldots, v_n\}$ is a basis for $\mathbf{Z}_p^n$. Then, there exists $B \in \mathrm{GL}_n(\mathbf{Z}_p)$ such that $Bv_i = e_i$ for $i = 1, \ldots, n$. Then for each $i = 1, \ldots, n$,

$$A_0Bv_i = A_0e_i = \gamma_i = Av_i$$

and hence $A = A_0B$. Therefore, $\{A \in \mathbf{Z}_p^{n\times n} : \mathrm{im}(A) = \Gamma\} \subseteq A_0 \cdot \mathrm{GL}_n(\mathbf{Z}_p)$.

Now, by the properties of Haar measure, we have

$$\begin{aligned}\Pr(\mathrm{im}(A) = \Gamma) &= \frac{1}{|\det(A_0)|^n} \cdot \Pr(A \in \mathrm{GL}_n(\mathbf{Z}_p)) \\ &= \frac{1}{|G_p|^n} \cdot \Pr(A \text{ is invertible}).\end{aligned}$$

Claim. *$A \in \mathrm{GL}_n(\mathbf{Z}_p)$ if and only if $A \mod p \in \mathrm{GL}_n(\mathbb{Z}/p\mathbb{Z})$.*

If $A \in \mathrm{GL}_n(\mathbf{Z}_p)$ then there exists $B \in \mathrm{GL}_n(\mathbf{Z}_p)$ such that $AB = I$. Then $(A \mod p)(B \mod p) = I$ and hence $A \mod p \in \mathrm{GL}_n(\mathbb{Z}/p\mathbb{Z})$. For the converse, let $A \in \mathrm{M}_n(\mathbf{Z}_p)$ be such that $A \mod p \in \mathrm{GL}_n(\mathbb{Z}/p\mathbb{Z})$. Then $\det(A \mod p) = \det(A) \mod p \neq 0$ in $\mathbb{Z}/p\mathbb{Z}$. Therefore $p \nmid \det(A)$, i.e. $\det(A)$ is a unit in $\mathbf{Z}_p$; hence $\det(A)^{-1} \in \mathbf{Z}_p$. Now, $A \cdot \operatorname{adj} A = \det(A) \cdot I$ implies $A(\det(A)^{-1} \operatorname{adj} A) = I$. Therefore $A \in \mathrm{GL}_n(\mathbf{Z}_p)$.

Using the above claim we can write

$$\begin{aligned}\Pr(A \in \mathrm{GL}_n(\mathbf{Z}_p)) &= \Pr(A \mod p \in \mathrm{GL}_n(\mathbb{Z}/p\mathbb{Z})) \\ &= \frac{|\mathrm{GL}_n(\mathbb{Z}/p\mathbb{Z})|}{|(\mathbb{Z}/p\mathbb{Z})^{n\times n}|} \\ &= \frac{|\mathrm{GL}_n(\mathbb{Z}/p\mathbb{Z})|}{p^{n^2}}.\end{aligned}$$

Claim. $|\mathrm{GL}_n(\mathbb{Z}/p\mathbb{Z})| = p^{n^2}\prod_{i=1}^{n}(1-p^{-i})$.

We need to count the number of invertible matrices in $\mathbb{Z}/p\mathbb{Z}$. For any invertible matrix, the first column can be any non-zero vector, i.e. there are $p^n - 1$ choices for the first column; the second column should not be in the span of the first column, this gives $p^n - p$ choices for the second column; in general, the ith column should not be in the span of 1st, 2nd,...,$(i-1)$th columns, which gives $p^n - p^{i-1}$ choices for the i th column. Hence we get

$$\begin{aligned}|\mathrm{GL}_n(\mathbb{Z}/p\mathbb{Z})| &= \prod_{i=1}^{n}(p^n - p^{i-1}) \\ &= p^{n^2}\prod_{i=1}^{n}(1 - p^{i-1-n}) \\ &= p^{n^2}\prod_{i=1}^{n}(1 - p^{-i}).\end{aligned}$$

Then we get $\Pr(A \in \mathrm{GL}_n(\mathbf{Z}_p)) = \prod_{i=1}^{n}(1-p^{-i}))$. This implies

$$\Pr(\operatorname{im}(A) = \Gamma) = \frac{1}{|G_p|^n}\prod_{i=1}^{n}(1-p^{-i});$$

note that this is independent of Γ. Now,

$$\begin{aligned}\Pr(\mathbf{Z}_p^n/\operatorname{im}(A) \cong G_p) &= |\{\Gamma \subseteq \mathbf{Z}_p^n : \mathbf{Z}_p^n/\Gamma \cong G_p\}| \cdot \Pr(\operatorname{im}(A) = \Gamma) \\ &= \frac{1}{|G_p|^n}\prod_{i=1}^{n}(1-p^{-i}) \cdot |\{\Gamma \subseteq \mathbf{Z}_p^n : \mathbf{Z}_p^n/\Gamma \cong G_p\}|.\end{aligned}$$

Claim. *We have the following:*

$$|\{\Gamma \subseteq \mathbf{Z}_p^n : \mathbf{Z}_p^n/\Gamma \cong G_p\}| = \frac{|G_p|^n}{|\mathrm{Aut}(G_p)|} \prod_{i=n-r+1}^{n} (1-p^{-i}).$$

We will count the number of surjective homomorphisms, say s_n, of $\mathbf{Z}_p$-modules from $\mathbf{Z}_p^n$ to G_p. Note that,

$$|\{\Gamma \subseteq \mathbf{Z}_p^n : \mathbf{Z}_p^n/\Gamma \cong G_p\}| = \frac{s_n}{|\mathrm{Aut}(G_p)|},$$

because, each surjective homomorphism defines a submodule $\Gamma = \ker(\phi)$ and two such surjecive homomorphisms ϕ_1 and ϕ_2 defines the same Γ if and only if there exists an automorphism σ of G_p such that $\phi_1 = \sigma \circ \phi_2$. Let $\overline{s}_n$ be the number of surjective homomorphisms (of $\mathbb{Z}/p\mathbb{Z}$-modules) from $(\mathbb{Z}/p\mathbb{Z})^n$ to G_p/pG_p. Now, by Nakayama's lemma, $\phi \in \mathrm{Hom}_{\mathbf{Z}_p}(\mathbf{Z}_p^n, G_p)$ is surjective if and only if

$$\overline{\phi} := \phi \mod p \in \mathrm{Hom}_{\mathbb{Z}/p\mathbb{Z}}((\mathbb{Z}/p\mathbb{Z})^n, G_p/pG_p)$$

is surjective. Therefore,

$$s_n = \overline{s}_n \cdot |\{\phi \in \mathrm{Hom}_{\mathbf{Z}_p}(\mathbf{Z}_p^n, G_p) : \overline{\phi} = 0\}|.$$

Now, $\overline{\phi} = 0$ if and only if $\mathrm{im}(\phi) \in pG_p$, i.e. if and only if $\phi(e_i) \in pG_p$, where e_is are the canonical basis for $\mathbf{Z}_p^n$. Then

$$|\{\phi \in \mathrm{Hom}_{\mathbf{Z}_p}(\mathbf{Z}_p^n, G_p) : \overline{\phi} = 0\}| = |pG_p|^n.$$

Claim. $|pG_p| = \frac{1}{p^r}|G_p|$

We will show that $|G_p| = p^r|pG_p|$. Let $G_p = \langle g_1, \ldots, g_r \rangle$, note that $\mathrm{rank}(G_p) = r$. Then,

$$\begin{aligned} G_p &= \mathbb{Z}g_1 + \cdots + \mathbb{Z}g_r \\ &= (p\mathbb{Z}g_1 + \cdots + p\mathbb{Z}g_r) + (\mathbb{Z}/p\mathbb{Z}g_1 + \cdots + \mathbb{Z}/p\mathbb{Z}g_r) \\ &= pG_p + (\mathbb{Z}/p\mathbb{Z}g_1 + \cdots + \mathbb{Z}/p\mathbb{Z}g_r) \end{aligned}$$

and this implies that $|G_p| = |pG_p| \cdot p^r$.

Hence, we get that

$$s_n = \overline{s}_n \cdot \left(\frac{|G_p|}{p^r}\right)^n.$$

Next our task is count $\overline{s}_n$. Note that, $\overline{s}_n$ is equal to the number of $n \times r$ matrices of rank r over $\mathbb{Z}/p\mathbb{Z}$. For any such matrix, there are p^n choices for it's first column and in general $p^n - p^{i-1}$ choices for the ith column, where $i = 1, \ldots, r$. Hence $\overline{s}_n = \prod_{i=1}^{r}(p^n - p^{i-1})$. Then we get,

$$\begin{aligned} s_n &= \frac{|G_p|^n}{p^{rn}} \cdot \prod_{i=1}^{r}(p^n - p^{i-1}) \\ &= |G_p|^n \cdot \prod_{i=n-r+1}^{n}(1 - p^{-i}). \end{aligned}$$

Therefore we have,

$$\begin{aligned} \Pr(\mathbf{Z}_p^r/\operatorname{im}(A) \cong G_p) &= \frac{s_n}{|\operatorname{Aut}(G_p)|} \cdot \Pr(\operatorname{im}(A) = \Gamma)) \\ &= |G_p|^n \prod_{i=n-r+1}^{n}(1 - p^{-i}) \frac{1}{|\operatorname{Aut}(G_p)|} \cdot \frac{1}{|G_p|^n} \prod_{i=1}^{n}(1 - p^{-i}) \\ &= \frac{1}{|\operatorname{Aut}(G_p)|} \cdot \prod_{i=1}^{n}(1 - p^{-i}) \prod_{i=n-r+1}^{n}(1 - p^{-i}). \end{aligned}$$

Now taking limit as $n \to \infty$, we get

$$\Pr(\mathbf{Z}_p^r/\operatorname{im}(A) \cong G_p) \longrightarrow \frac{1}{|\operatorname{Aut}(G_p)|} \prod_{i=1}^{\infty}(1 - p^{-i}) = P_p(G_p).$$

□

3.1.2 Global Cohen-Lenstra measure

Next, we will see how to extend the local Cohen-Lenstra probability measure P_p to a bigger class of groups. We will define a σ-algebra Σ over the class of all groups with trivial 2-part (denoted by $\mathcal{G}$) and a probability measure P on Σ, such that, P is *compatible* with the local Cohen-Lenstra measure P_p, i.e. it satisfies

1. $P(\pi_p^{-1}(M)) = P_p(M)$ for all $M \subseteq \mathcal{G}_p$, where $\pi_p \colon \mathcal{G} \to \mathcal{G}_p$ is the natural projection to pth part.

2. For finitely many distinct primes $p_1, \ldots, p_k$,

$$P\left(\bigcap_{i=1}^{k} \pi_{p_i}^{-1}(M_i)\right) = \prod_{i=1}^{k} P(\pi_{p_i}^{-1}(M_i)) = \prod_{i=1}^{k} P_{p_i}(M_i).$$

Let $\mathbb{P}$ denote the set of all odd primes. Then condition 1 implies that $\pi_p^{-1}(M) \in \Sigma$ for all $M \subseteq \mathcal{G}_p$ and $p \in \mathbb{P}$. One can think of P as the joint distribution and P_ps as the marginal distributions. Although P looks like a product measure, but since $\mathcal{G} = \bigoplus_p \mathcal{G}_p$, we can not define a product measure on $\mathcal{G}$. Condition 2 says that, the p-parts of a group $G \in \mathcal{G}_p$ should be independent w.r.t. the probability measure P. The motivation for this condition comes from the following fact: If $G = G_1 \times G_2$, where G_is are different p-groups, then $\mathrm{Aut}(G) = \mathrm{Aut}(G_1) \times \mathrm{Aut}(G_2)$, i.e. $1/\left|\mathrm{Aut}(G)\right| = 1/\left|\mathrm{Aut}(G_1)\right| \cdot 1/\left|\mathrm{Aut}(G_2)\right|$.

The first thing that comes to mind is to define Σ as the smallest σ-algebra that contains the set $\{\pi_p^{-1}(M) : p \in \mathbb{P},\, M \subseteq \mathcal{G}_p\}$ and define P on Σ by the product formula. But unfortunately, this does not define a measure, as given by the following proposition.

Proposition 3.8. *Under the above definition, one can show that $P(\mathcal{G}) = 0$ and hence this does not define a probability measure.*

Proof. Let $G \in \mathcal{G}$ and suppose $G = G_{p_1} \times \cdots \times G_{p_r}$, where G_{p_i} is a p_i-group and p_is are distinct. Then we can write

$$\{G\} = \prod_{i=1}^{r} \pi_{p_i}^{-1}(\{G_{p_i}\}) \times \prod_{p \neq p_i} \pi_p^{-1}(\{0\}),$$

which is countable intersection of measurable sets and hence $\{G\} \in \Sigma$. Now, by condition 2, we have

$$\begin{aligned}
P(\{G\}) &= \prod_{i=1}^{r} P_{p_i}(\{G_{p_i}\}) \cdot \prod_{p \neq p_i} P_p(\{0\}) \\
&= \prod_{i=1}^{r} P_{p_i}(\{G_{p_i}\}) \cdot \prod_{p \neq p_i} \prod_{i=1}^{\infty} \left(1 - \frac{1}{p^i}\right) \\
&\leq \prod_{i=1}^{r} P_{p_i}(\{G_{p_i}\}) \cdot \prod_{p \neq p_i} \left(1 - \frac{1}{p}\right) \\
&\leq \prod_{i=1}^{r} P_{p_i}(\{G_{p_i}\}) \cdot \prod_{p \neq p_i} e^{-1/p} \\
&= \prod_{i=1}^{r} P_{p_i}(\{G_{p_i}\}) \cdot e^{-\sum_{p \neq p_i} 1/p} \\
&= 0,
\end{aligned}$$

since $\sum_{p \neq p_i} 1/p = \infty$. Hence, $P(\{G\}) = 0$ for all $G \in \mathcal{G}$. Now, since $\mathcal{G}$ is countable, we get

$$\begin{aligned} P(\mathcal{G}) &= P\left(\bigcup_{G \in \mathcal{G}} \{G\}\right) \\ &= \sum_{G \in \mathcal{G}} P(\{G\}) \qquad \text{(by countable additivity)} \\ &= 0. \end{aligned}$$

Therefore P does not define a probability measure, because for any probability measure P, we must have $P(\mathcal{G}) = 1$. □

Next we show that, the order of a group can not be measurable w.r.t. to any probability measure which extends the local Cohen-Lenstra measure.

Proposition 3.9. *For any $n \in \mathbb{N}$, define $S_n := \{G \in \mathcal{G} : |G| = n\}$. Then, for any σ-algebra Σ over $\mathcal{G}$ and probability measure P on Σ which is compatible with the local Cohen-Lenstra probability measure, we can not have that S_n is measurable for all $n \in \mathbb{N}$.*

Proof. Suppose S_n is measurable for all $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, define the set $I_n := \{p \in \mathbb{P} : p \nmid n\}$ and for each $p \in I_n$, let

$$T_p := \{G \in \mathcal{G} : \pi_p(G) = 0\}.$$

Then $T_p = \pi_p^{-1}(\{0\})$, which is measurable. Note that,

$$P(T_p) = P_p(\{0\}) = \prod_{i=1}^{\infty} \left(1 - \frac{1}{p^i}\right) \leq \left(1 - \frac{1}{p}\right).$$

Now, $S_n \subseteq T_p$ for all $p \in I_n$ (for, if $\pi_p(G) \neq 0$ for some $G \in S_n$, then $|\pi_p(G)| \mid n$, since $\pi_p(G)$ is a subgroup of G and this implies $p \mid n$). Therefore we have $S_n \subseteq \bigcap_{p \in I_n} T_p$ and also note that $\bigcap_{p \in I_n} T_p$ being countable of intersection of measurable sets is measurable. Then,

$$P(S_n) \leq P\left(\bigcap_{p \in I_n} T_p\right) = \prod_{p \in I_n} P(T_p) \leq \prod_{p \in I_n} \left(1 - \frac{1}{p}\right).$$

Now, since $e^x \geq 1 + x$ for all $x \in \mathbb{R}$, we get

$$P(S_n) \leq e^{-\sum_{p \in I_n} 1/p} = 0,$$

since $\sum_{p \in I_n} 1/p$ diverges. Therefore $P(S_n) = 0$ for all $n \in \mathbb{N}$. Now, note that $\mathcal{G} = \bigcup_{n \in \mathbb{N}} S_n$, which implies

$$P(\mathcal{G}) = \sum_{n \in \mathbb{N}} P(S_n) = 0$$

and we have a contradiction. □

Remark 3.10. *The above proposition also holds if Σ is an algebra and P is a finitely additive measure compatible with the local Cohen-Lenstra measure. We work with finite $F \subset I_n$ and take infimum over all such sets and the same proof works.*

Next, we define certain properties of groups, called uniform properties, which will become measurable w.r.t. our desired probability measure.

Definition 3.11. *A property $E \colon \mathcal{G} \to \{0, 1\}$ of $\mathcal{G}$ is said to be **uniform** if for all $G \in \mathcal{G}$, we have*

$$E(G) = 1 \quad \textit{iff} \quad E(G_p) = 1 \quad \forall p \in \mathbb{P},$$

where $G_p := \pi_p(G)$ ($\pi_p \colon \mathcal{G} \to \mathcal{G}_p$ is the natural projection).

The following are some examples of uniform properties:

1. Property of having order 1 (i.e. $E(G) = 1$ iff $|G| = 1$).
2. Property of having rank $\leq$ r (i.e. $E(G) = 1$ iff $\operatorname{rank}(G) \leq r$).

   This follows from the fact that

   $$\operatorname{rank}(G) = \max_{p \in \mathbb{P}} \operatorname{rank}(G_p)$$

   for all $G \in \mathcal{G}$, where $G_p = \pi_p(G)$.
3. Property of having *uniform order* $\leq n$, where uniform order $|G|_{\text{uni}}$ of $G \in \mathcal{G}$ is defined as

   $$|G|_{\text{uni}} := \max_{p \in \mathbb{P}} \log_p |G_p|.$$

   (Similarly, one can also define *uniform exponent* and *uniform rank*. Note that, uniform rank is same as the rank).

Then, we have the following theorem due to J. Lengler.

Theorem 3.12. *Let Σ be the smallest σ-algebra (over $\mathcal{G}$) containing the set*

$$\{E^{-1}(\{1\}) : E\colon \mathcal{G} \to \{0,1\} \text{ is a uniform property}\}.$$

Now, define a measure P on Σ, given by

$$P(E^{-1}(\{1\})) := \prod_{p\in\mathbb{P}} P_p(E_p^{-1}(\{1\})),$$

where $E_p := E|_{\mathcal{G}_p}$ and P_p is the local Cohen-Lenstra measure. Then, P defines a probability measure measure on Σ which makes all uniform properties measurable.

Proof. See [Len10b] Theorem 2.11. □

The probability measure P defined in the above theorem is called the *global Cohen-Lenstra probability measure* which is an extension of the local Cohen-Lenstra probability measure P_p. Note that, for all $r \in \mathbb{N}$, the set $\{G \in \mathcal{G} : \text{rank}(G) = r\}$ is measurable (w.r.t P), because

$$\{G \in \mathcal{G} : \text{rank}(G) = r\} = \{G \in \mathcal{G} : \text{rank}(G) \leq r\} \cap \{G \in \mathcal{G} : \text{rank}(G) \leq r-1\}^{\complement}$$

and both these sets are measurable. Similarly, the set $\{G \in \mathcal{G} : |G|_{\text{uni}} = n\}$ is also measurable and so is $\{G \in \mathcal{G} : |G| = 1\}$. For $G \in \mathcal{G}$, we have

$$P(G = \{0\}) = \prod_{p\in\mathbb{P}} P_p(G_p = \{0\}) = \prod_{p\in\mathbb{P}} \left(\prod_{j=1}^{\infty} \left(1 - \frac{1}{p^j}\right)\right).$$

Now, since $\prod_{j=1}^{\infty}(1-p^{-j}) \leq (1-1/p) \leq e^{-1/p}$, we get

$$P(G = \{0\}) \leq \prod_{p\in\mathbb{P}} e^{-1/p} = e^{-\sum_{p\in\mathbb{P}} 1/p} = 0.$$

Hence, the global Cohen-Lenstra probability of a group being trivial is 0.

Next we will state two interesting properties of the global Cohen-Lenstra probability measure P.

Proposition 3.13. *Let E be a uniform property . If $P(E^{-1}(\{1\})) > 0$ then $E(0) = 1$ and $E(\mathbb{Z}/p\mathbb{Z}) = 1$ for some $p \in \mathbb{P}$.*

Proof. Suppose $E(0) = 0$, then $E_p(0) = 0$ for each p (note that, $E_p = E|_{\mathcal{G}_p}$ and $0 \in \mathcal{G}_p$). This implies $E_p^{-1}(\{1\}) \subseteq \mathcal{G}_p \setminus \{0\}$ for each p. Hence we have

$$\begin{aligned} P_p(E_p^{-1}(\{1\})) &\leq P_p(\mathcal{G}_p \setminus \{0\}) \\ &= 1 - P_p(\{0\}) \\ &= 1 - \prod_{i=1}^{\infty} \left(1 - \frac{1}{p^i}\right). \end{aligned}$$

Now, if we wirte

$$\prod_{i=1}^{\infty}(1 - x^i) = 1 + \sum_{n=1}^{\infty} a_n x^n$$

then, a_ns are given by

$$a_n = \begin{cases} 0 & \text{if } n \neq k(3k \pm 1)/2,\ k \in \mathbb{N} \\ (-1)^k & \text{if } n = k(3k \pm 1)/2,\ k \in \mathbb{N} \end{cases}$$

and therefore $a_n \geq 1$ for each $n \in \mathbb{N}$. Hence, we can write

$$P_p(E_p^{-1}(\{1\})) \leq 1 - \left(1 - \sum_{i=1}^{\infty} \frac{1}{p^i}\right) = \frac{1/p}{1 - 1/p} = \frac{1}{p-1}.$$

Therefore

$$P(E^{-1}(\{1\})) = \prod_{p \in \mathbb{P}} P_p(E_p^{-1}(\{1\})) \leq \prod_{p \in \mathbb{P}} \frac{1}{p-1} = \frac{1}{\prod_p (p-1)} = 0.$$

but this is a contradiction. Hence, we conclude that $E(0) = 0$.

For the other part, suppose $E(\mathbb{Z}/p\mathbb{Z}) = 0$ for all $p \in \mathbb{P}$. This implies $E_p(\mathbb{Z}/p\mathbb{Z}) = 0$ for all $p \in \mathbb{P}$ (note that, $E_p = E|_{\mathcal{G}_p}$ and $\mathbb{Z}/p\mathbb{Z} \in \mathcal{G}_p$). Therefore, for all p, $E_p^{-1}(\{1\}) \subseteq \mathcal{G}_p \setminus \{\mathbb{Z}/p\mathbb{Z}\}$. Then,

$$\begin{aligned} P_p(E_p^{-1}(\{1\})) &\leq 1 - P_p(\{\mathbb{Z}/p\mathbb{Z}\}) \\ &= 1 - \frac{1}{p-1} \prod_{i=1}^{\infty} \left(1 - \frac{1}{p^i}\right) \\ &= 1 - p^{-1} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots\right) \prod_{i=1}^{\infty} \left(1 - \frac{1}{p^i}\right). \end{aligned}$$

Now, writing $\prod_{i=1}^{\infty}(1-1/p^i) = 1 + \sum_{n=1}^{\infty} a_n/p^n$, we get

$$\begin{aligned} P_p(E_p^{-1}(\{1\})) &\leq 1 - p^{-1}\left(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots\right)\left(1 + \sum_{n=1}^{\infty}\frac{a_n}{p^n}\right) \\ &= 1 - p^{-1}\left(\sum_{N=0}^{\infty}\left(\sum_{n=0}^{N} a_n\right)\frac{1}{p^n}\right) \qquad (\text{where } a_0 := 1) \\ &\leq 1 - p^{-1}\left(1 - 2\sum_{i=2}^{\infty}\frac{1}{p^i}\right) \end{aligned}$$

since, $a_0 = 1$, $a_0 + a_1 = 0$ and $\sum_{i=0}^{n} a_i \geq -2$ for all $n \geq 2$. Therefore, from above, we get

$$P_p(E_p^{-1}(\{1\})) \leq 1 - p^{-1} + \frac{2p^{-2}}{p-1} \leq 1 - \frac{1}{2p}$$

since $p > 2$. Therefore, writing $P(E^{-1}(\{1\})) = \prod_{p\in\mathbb{P}} P_p(E_p^{-1}(\{1\}))$ and using the above inequality we get

$$P(E^{-1}(\{1\})) \leq \prod_{p\in\mathbb{P}}\left(1 - \frac{1}{2p}\right) \leq \prod_{p\in\mathbb{P}} e^{-1/2p} = e^{-(1/2)\sum_{p\in\mathbb{P}} 1/p} = 0,$$

which is a contradiction. Therefore, $E(\mathbb{Z}/p\mathbb{Z}) = 1$ for some $p \in \mathbb{P}$. □

Proposition 3.14. *If $E(0) = 1$ and $E(\mathbb{Z}/p\mathbb{Z}) = 1$ for all $p \in \mathbb{P}$, then we have that $P(E^{-1}(\{1\})) > 0.75$; where E is a uniform property.*

Proof. Note that,

$$\{0, \mathbb{Z}/p\mathbb{Z}\} \subseteq E^{-1}(\{1\}) = E_p^{-1}(\{1\})$$

for each p. Then,

$$\begin{aligned} P_p(E_p^{-1}(\{1\})) &\geq P_p(\{0\}) + P_p(\{\mathbb{Z}/p\mathbb{Z}\}) \\ &= \prod_{i=1}^{\infty}\left(1 - \frac{1}{p^i}\right) + \frac{1}{p-1}\prod_{i=1}^{\infty}\left(1 - \frac{1}{p^i}\right) \\ &= \frac{p}{p-1}\prod_{i=1}^{\infty}\left(1 - \frac{1}{p^i}\right) \\ &= \prod_{i=2}^{\infty}\left(1 - \frac{1}{p^i}\right). \end{aligned}$$

Now, from $P(E^{-1}(\{1\})) = \prod_{p\in\mathbb{P}} P_p(E_p^{-1}(\{1\}))$ we get,

$$P(E^{-1}(\{1\})) \geq \prod_{p\in\mathbb{P}}\prod_{j=2}^{\infty}\left(1-\frac{1}{p^j}\right) \approx 0.75\ldots > 0.75$$

and this proves the proposition. □

3.1.3 Heuristic for Quadratic Fields

First, let us state what it means for a sequence of groups to be random with respect to the Cohen-Lenstra heuristic:

Definition 3.15. *Let $\{G_i\}_{i=1}^{\infty}$ be a sequence of finite abelian groups. Let Σ be the sigma algebra on $\mathcal{G}$ generated by uniform properties and let P be the global Cohen-Lenstra probability measure defined on Σ. We say that G_i* ***behaves as a random sequence*** *with respect to Cohen-Lenstra measure P if, for all measurable functions $f\colon \mathcal{G}\to\mathbb{C}$ we have,*

$$\lim_{n\to\infty}\frac{\sum_{i=1}^{n} f(G_i)}{n} = \int_{\mathcal{G}} f\, dP.$$

Now, coming to imaginary quadratic fields, let S be the the following sequence in $\mathcal{G}$:

$$S := \{\mathrm{Cl}_K^* : K\in\mathcal{K}\}$$

where, $\mathcal{K} := \{\mathbb{Q}(\sqrt{d}) : d \text{ is square-free and } d<0\}$ is the sequence of imaginary quadratic fields and Cl_K^* denotes the odd part of Cl_K (i.e. subgroup of elements of odd order). Then, the famous conjecture of Cohen and Lenstra states the following:

Conjecture 3.16. *The sequence S behaves as a random sequence with respect to the global Cohen-Lenstra probability measure P, that is, for any measurable function $f\colon \mathcal{G}\to\mathbb{C}$,*

$$\lim_{x\to\infty}\frac{\sum_{K\in\mathcal{K},\,|\Delta_K|\leq x} f(\mathrm{Cl}_K^*)}{\sum_{K\in\mathcal{K},\,|\Delta_K|\leq x} 1} = \mathrm{E}_P(f)$$

where, E_P denotes the expectation w.r.t. the probability measure P.

Recall that the Cohen-Lenstra probability of a group being trivial is 0. Hence, the above conjecture implies that,

$$P(\mathrm{Cl}_K^* = \{0\}) = 0,$$

which supports the fact that there are only finitely many imaginary quadratic fields of class number one.

3.2 Probabilistic model for Real Quadratic Fields

In this section we will modify our definition for local Cohen-Lenstra probability measure P_p. This modified definition will give us a probabilistic model for real quadratic fields. We will redefine P_p such that

$$P_p(G_p) \propto \frac{1}{|G_p| \cdot |\mathrm{Aut}(G_p)|}.$$

Note that, we have

$$\sum_{G_p \in \mathcal{G}_p} \frac{1}{|G_p| \cdot |\mathrm{Aut}(G_p)|} \leq \sum_{G_p \in \mathcal{G}_p} \frac{1}{|\mathrm{Aut}(G_p)|} < \infty$$

and this implies that P_p is indeed a probability measure. In fact, we have the following theorem:

Theorem 3.17. *Let p be a prime. Then,*

$$\sum_{G_p \in \mathcal{G}_p} \frac{1}{|G_p| \cdot |\mathrm{Aut}(G_p)|} = \prod_{j=2}^{\infty} \left(1 - \frac{1}{p^j}\right)^{-1}.$$

Proof. The original proof of this theorem is due to Cohen and Lenstra, which uses zeta functions. I will give a shorter and more elegant combinatorial proof. I will use similar ideas used in the proof of Theorem 3.2.

For $m \geq 0$, let a_m be the number of partitions of m with each part at least 2 and for $i, j \geq 0$, let $b_{i,j}$ be the number of partitions of i with greatest part exactly equal to j. Then the following claims are true.

Claim. *For all $m \geq 0$, the following holds*

$$a_m = \sum_{i+j=m} b_{i,j}.$$

We will give a bijection argument. Note that, the number of partitions of i with greatest part j, where $i + j = m$, is equal to the number of partitions of $i + j = m$ with greatest part j occurring at least twice. Hence $\sum_{i+j=m} b_{i,j}$ is equal to the number of partitions of m with greatest part occurring at least twice. Now, a partition of m has greatest part occurring at least twice if and only if it's conjugate partition has each part at least 2. This gives a bijection between the partitions of m with greatest part occurring at least twice and the partitions of m with each part at least 2. Therefore, $\sum_{i+j=m} b_{i,j} = a_m$.

Claim. *For each $n \geq 0$, let us define*

$$f_n(q) := \sum_{N=0}^{\infty} b_{N,n} q^N,$$

which is a formal power series in q. Then,

$$\prod_{j=2}^{\infty} (1-q^j)^{-1} = \sum_{n=0}^{\infty} f_n(q) q^n.$$

Note that,

$$\prod_{j=2}^{\infty} (1-q^j)^{-1} = \sum_{m=0}^{\infty} a_m q^m$$

since, for each m, the coefficient of q^m on LHS is equal to the number of partitions of m with each part at least 2. Then,

$$\begin{aligned}
\sum_{n=0}^{\infty} f_n(q) q^n &= \sum_{n=0}^{\infty} \left(\sum_{N=0}^{\infty} b_{N,n} q^N \right) q^n \\
&= \sum_{n=0}^{\infty} \sum_{N=0}^{\infty} b_{N,n} q^{N+n} \\
&= \sum_{m=0}^{\infty} \left(\sum_{i+j=m} b_{i,j} \right) q^m \\
&= \sum_{m=0}^{\infty} a_m q^m \\
&= \prod_{j=2}^{\infty} (1-q^j)^{-1}.
\end{aligned}$$

Now let us return to the proof of the theorem. Following a similar argument as given in the proof of Theorem 3.2, we can write

$$\sum_{G_p \in \mathcal{G}_p} \frac{1}{|G_p| \cdot |\mathrm{Aut}(G_p)|} = \sum_{n=0}^{\infty} \sum_{\substack{G_\mu \in \mathcal{G}_p \\ |\mu|=n}} q^n \left(\prod_{i=1}^{m} \psi_{\mu_i,\, \mu_{i-1}-\mu_i}(q) \right) \left(\prod_{i=1}^{m} q^{\mu_i^2} \right),$$

where the notations are same as in Theorem 3.2 and $q = p^{-1}$. Then, by second claim it is enough to show that, for each $n \geq 0$,

$$f_n(q) = \sum_{\substack{G_\mu \in \mathcal{G}_p \\ |\mu|=n}} \left(\prod_{i=1}^{m} \psi_{\mu_i,\, \mu_{i-1}-\mu_i}(q) \right) \left(\prod_{i=1}^{m} q^{\mu_i^2} \right).$$

That is, we need to equate coefficients of q^N on both sides, for each $N \geq 0$. Note that, coefficient of q^N on LHS is equal to $b_{N,n}$ which is the number of partitions of N with greatest part n and this is equal to the coefficient of q^N on the RHS (the proof of this is same as given the proof of Theorem 3.2). Therefore we have proved the theorem. □

Let p be a prime. Then the above theorem will enable us to define P_p as

$$P_p(G_p) := \frac{1}{|G_p| \cdot |\mathrm{Aut}(G_p)|} \prod_{j=2}^{\infty} \left(1 - \frac{1}{p^j}\right).$$

Note that, by Theorem 3.17, $\sum_{G_p \in \mathcal{G}_p} P_p(G_p) = 1$ and hence P_p is indeed a probability measure on $\mathcal{G}_p$. Note that the following random process gives us the above probability measure

1. Choose $H \in \mathcal{G}_p$ w.r.t. the local Cohen-Lenstra probability measure as in Definition 3.3.
2. Choose an element $g \in H$ uniformly at random.
3. Output the group $H/\langle g \rangle$.

Next, we will extend P_p to $\mathcal{G}$ so that it is compatible (as defined in imaginary case) with local P_p as follows: Define a probability measure P on $\mathcal{G}$ as,

$$P(G) = \prod_{p \in \mathbb{P}} P_p(G_p),$$

where, $G \in \mathcal{G}$ and $G_p = \pi_p(G)$ (where, $\pi_p \colon \mathcal{G} \to \mathcal{G}_p$ is the natural projection).

Theorem 3.18. *P indeed defines a probability measure on the whole power set of $\mathcal{G}$.*

Proof. Let $G \in \mathcal{G}$, then

$$\begin{aligned} P(G) &= \prod_{p \in \mathbb{P}} P_p(G_p) \\ &= \prod_{p \in \mathbb{P}} \frac{1}{|G_p| \cdot |\mathrm{Aut}(G_p)|} \prod_{j=2}^{\infty} \left(1 - \frac{1}{p^j}\right) \\ &= \frac{1}{|G| \cdot |\mathrm{Aut}(G)|} \prod_{p \in \mathbb{P}} \prod_{j=2}^{\infty} \left(1 - \frac{1}{p^j}\right) \\ &= \frac{0.75 \cdots}{|G| \cdot |\mathrm{Aut}(G)|} > 0. \end{aligned}$$

The second last equality follows from the fact that $|G| = \prod_{p\in\mathbb{P}} |G_p|$ and $|\mathrm{Aut}(G)| = \prod_{p\in\mathbb{P}} |\mathrm{Aut}(G_p)|$. Clearly, for any $S \subseteq \mathcal{G}$,

$$P(S) = \sum_{G\in S} P(G)$$

and this implies countable additivity. Also, we have

$$\sum_{G\in\mathcal{G}} P(G) = \sum_{\substack{G\in\mathcal{G} \\ G=\prod_p G_p}} \prod_{p\in\mathbb{P}} P_p(G_p) = \prod_{p\in\mathbb{P}} \sum_{G_p\in\mathcal{G}_p} P_p(G_p) = \prod_{p\in\mathbb{P}} 1 = 1.$$

Note that, the interchange of limits is possible because $P(G) > 0$ for all $G \in \mathcal{G}$. Hence, P is a probability measure. □

We also note the following observation

$$P(0) = \prod_{p\in\mathbb{P}} \prod_{j=2}^{\infty} \left(1 - \frac{1}{p^j}\right) \approx 0.75.$$

Now, let us come to the case of quadratic fields. Let S be the following sequence in $\mathcal{G}$:

$$S := \{\mathrm{Cl}_K^* : K \in \mathcal{K}\}$$

where, $\mathcal{K} := \{\mathbb{Q}(\sqrt{d}) : d \text{ is square free and } d > 0\}$ is the sequence of real quadratic fields. Then, Cohen-Lenstra heuristic states the following conjecture.

Conjecture 3.19. *The above sequence S behaves like a random sequence w.r.t. the probability measure P, that is, for any measurable function $f\colon \mathcal{G} \to \mathbb{C}$ we have,*

$$\lim_{x\to\infty} \frac{\sum_{K\in\mathcal{K},\, |\Delta_K|\le x} f(\mathrm{Cl}_K^*)}{\sum_{K\in\mathcal{K},\, |\Delta_K|\le x} 1} = \mathrm{E}_P(f),$$

where $\mathrm{E}_P(f)$ is the expectation of f w.r.t. the probability measure P.

Next, we will explain the motivation behind applying this probabilistic model to real quadratic fields. Recall from Chapter 2 that the class of all reduced ideals of a real quadratic field decomposes into finitely many cycles under the reduction operator ρ and the number of such cycles gives the class number of the real quadratic field. Also, the set of reduced ideals behaves as a 'group like' structure and the cycles behaves like cyclic groups with respect to some 'group like' operation defined before. Now, the group H of the random process is to be interpreted as the 'group' of reduced ideals,

which is chosen w.r.t. Cohen-Lenstra probability. Let H decomposes into cycles (or, cyclic groups) as $\{\langle u_1\rangle, \langle u_2\rangle, \ldots, \langle u_h\rangle\}$, where $u_i \in H$ and h is the number of reduced ideals i.e. the class number. Then Cl_K can be viewed as $H/\langle u_i\rangle$, for all i. Note that, the random process also chooses a u_i uniformly at random and outputs the group $H/\langle u_i\rangle$. So, this gives some justification why this probabilistic model should work for the sequence of real quadratic fields.

Note that, the Cohen-Lenstra conjecture implies that

$$P(\mathrm{Cl}_K^* = \{0\}) \approx 0.75.$$

We know that, if Δ_K is prime then $\mathrm{Cl}_K \cong \mathrm{Cl}_K^*$. Now, if d is prime and $d \equiv 1 \pmod 4$ then Δ_K is prime, and there are infinitely many primes of the form $4n+1$. Therefore in such cases $P(\mathrm{Cl}_K = \{0\}) \approx 0.75$. This gives a hint that there might be infinitely many real quadratic fields with class number 1, i.e. Gauss' conjecture is true.

3.3 Probabilistic model for Number Fields

Let us fix a set of primes B, which we will call the set of *bad primes*. For each $p \notin B$, we will denote the set of all p-groups by $\mathcal{G}_p$ and $\mathcal{G}$ will denote the set of all finite abelian groups with p-part trivial for all $p \in B$, i.e., $\mathcal{G} = \bigoplus_{p\notin B} \mathcal{G}_p$. In this section we will again modify the definition of local Cohen-Lenstra probability P_p to give us a probabilistic model for general number fields. We will redefine P_p such that, for all $G_p \in \mathcal{G}_p$ (where $p \notin B$), we have

$$P_p(G_p) \propto \frac{1}{|G_p|^u \cdot |\mathrm{Aut}(G_p)|},$$

where $u := \mathrm{rank}(\mathfrak{O}_K^\times) = r+s-1$ (where $\{r, s\}$ is the signature of the number field K) by Dirichlet's Unit Theorem . Again, note that

$$\sum_{G_p\in\mathcal{G}_p} \frac{1}{|G_p|^u \cdot |\mathrm{Aut}(G_p)|} \leq \sum_{G_p\in\mathcal{G}_p} \frac{1}{|\mathrm{Aut}(G_p)|} < \infty$$

and therefore P_p is indeed a probability measure. In fact, we have the following theorem:

Theorem 3.20. *Let p be a prime. Then,*

$$\sum_{G_p\in\mathcal{G}_p} \frac{1}{|G_p|^u \cdot |\mathrm{Aut}(G_p)|} = \prod_{j=u+1}^{\infty} \left(1 - \frac{1}{p^j}\right)^{-1}.$$

Proof. The original proof of this is due to Cohen and Lenstra using zeta functions. I will give a shorter combinatorial proof. This proof is a generalization of the proof of the theorem 3.17.

For $m \geq 0$, let a_m be the number of partitions of m with each part at least $u+1$ and for $i, j \geq 0$, let $b_{i,j}$ be the number of partitions of i with greatest part exactly equal to j. Then the following claims are true.

Claim. *For all $m \geq 0$, the following holds*

$$a_m = \sum_{i+uj=m} b_{i,j}.$$

We will give a bijection argument. Note that, the number of partitions of i with greatest part j, where $i + uj = m$, is equal to the number of partitions of $i + uj = m$ with greatest part j occurring at least $u + 1$ times. Hence $\sum_{i+uj=m} b_{i,j}$ is equal to the number of partitions of m with greatest part occurring at least $u + 1$ times. Now, a partition of m has greatest part occurring at least $u + 1$ times if and only if it's conjugate partition has each part at least $u + 1$. This gives a bijection between the partitions of m with greatest part occurring at least $u+1$ times and the partitions of m with each part at least $u + 1$. Therefore, $\sum_{i+uj=m} b_{i,j} = a_m$.

Claim. *For each $n \geq 0$, let us define*

$$f_n(q) := \sum_{N=0}^{\infty} b_{N,n} q^N,$$

which is a formal power series in q. Then,

$$\prod_{j=u+1}^{\infty} (1 - q^j)^{-1} = \sum_{n=0}^{\infty} f_n(q) q^{nu}.$$

Note that,

$$\prod_{j=u+1}^{\infty} (1 - q^j)^{-1} = \sum_{m=0}^{\infty} a_m q^m$$

since, for each m, the coefficient of q^m on LHS is equal to the number of

partitions of m with each part at least $u+1$. Then,

$$\begin{aligned}\sum_{n=0}^{\infty} f_n(q)q^{nu} &= \sum_{n=0}^{\infty}\left(\sum_{N=0}^{\infty} b_{N,n}q^N\right)q^{nu}\\ &= \sum_{n=0}^{\infty}\sum_{N=0}^{\infty} b_{N,n}q^{N+nu}\\ &= \sum_{m=0}^{\infty}\left(\sum_{i+uj=m} b_{i,j}\right)q^m\\ &= \sum_{m=0}^{\infty} a_m q^m\\ &= \prod_{j=u+1}^{\infty}(1-q^j)^{-1}.\end{aligned}$$

Now let us return to the proof of the theorem. Following a similar argument as given in the proof of Theorem 3.2, we can write

$$\sum_{G_p\in\mathcal{G}_p}\frac{1}{|G_p|^u\cdot|\mathrm{Aut}(G_p)|} = \sum_{n=0}^{\infty}\sum_{\substack{G_\mu\in\mathcal{G}_p\\ |\mu|=n}} q^{nu}\left(\prod_{i=1}^{m}\psi_{\mu_i,\,\mu_{i-1}-\mu_i}(q)\right)\left(\prod_{i=1}^{m}q^{\mu_i^2}\right),$$

where the notations are same as in Theorem 3.2 and $q=p^{-1}$. Then, by second claim it is enough to show that, for each $n\geq 0$,

$$f_n(q) = \sum_{\substack{G_\mu\in\mathcal{G}_p\\ |\mu|=n}}\left(\prod_{i=1}^{m}\psi_{\mu_i,\,\mu_{i-1}-\mu_i}(q)\right)\left(\prod_{i=1}^{m}q^{\mu_i^2}\right).$$

That is, we need to equate coefficients of q^N on both sides, for each $N\geq 0$. Note that, coefficient of q^N on LHS is equal to $b_{N,n}$ which is the number of partitions of N with greatest part n and this is equal to the coefficient of q^N on the RHS (the proof of this is same as given the proof of Theorem 3.2). Therefore we have proved the theorem. □

Let p be a prime. Then the above theorem will enable us to define P_p as

$$P_p(G_p) := \frac{1}{|G_p|^u\cdot|\mathrm{Aut}(G_p)|}\prod_{j=u+1}^{\infty}\left(1-\frac{1}{p^j}\right).$$

Note that, by Theorem 3.20, $\sum_{G_p\in\mathcal{G}_p}P_p(G_p)=1$ and hence P_p is indeed a probability measure on $\mathcal{G}_p$. Note that the following random process gives us the above probability measure:

1. Choose $H \in \mathcal{G}_p$ w.r.t. the local Cohen-Lenstra probability measure as in Definition 3.3.
2. Choose u elements $g_1, \ldots, g_u \in H$ uniformly at random.
3. Output the group $H/\langle g_1, \ldots, g_u \rangle$.

Next, we will extend P_p to $\mathcal{G}$ so that it is compatible (as defined in imaginary case) with local P_p as follows: Define a probability measure P on $\mathcal{G}$ as,

$$P(G) = \prod_{p \notin B} P_p(G_p),$$

where, $G \in \mathcal{G}$ and $G_p = \pi_p(G)$ (where, $\pi_p \colon \mathcal{G} \to \mathcal{G}_p$ is the natural projection).

Theorem 3.21. *P indeed defines a probability measure on the whole power set of $\mathcal{G}$.*

Proof. Let $G \in \mathcal{G}$, then

$$\begin{aligned} P(G) &= \prod_{p \notin B} P_p(G_p) \\ &= \prod_{p \notin B} \frac{1}{|G_p|^u \cdot |\mathrm{Aut}(G_p)|} \prod_{j=u+1}^{\infty} \left(1 - \frac{1}{p^j}\right) \\ &= \frac{1}{|G|^u \cdot |\mathrm{Aut}(G)|} \prod_{p \notin B} \prod_{j=u+1}^{\infty} \left(1 - \frac{1}{p^j}\right), \end{aligned}$$

since, $|G| = \prod_{p \notin B} |G_p|$ and $|\mathrm{Aut}(G)| = \prod_{p \notin B} |\mathrm{Aut}(G_p)|$. This implies that $P(G) > 0$, because the product

$$\prod_{p \notin B} \prod_{j=u+1}^{\infty} \left(1 - \frac{1}{p^j}\right) > 0$$

for $u \geq 1$ and for any set B. Clearly, for any $S \subseteq \mathcal{G}$,

$$P(S) = \sum_{G \in S} P(G)$$

and this implies countable additivity. Also, we have

$$\sum_{G \in \mathcal{G}} P(G) = \sum_{\substack{G \in \mathcal{G} \\ G = \prod G_p}} \prod_{p \notin B} P_p(G_p) = \prod_{p \notin B} \sum_{G_p \in \mathcal{G}_p} P_p(G_p) = \prod_{p \notin B} 1 = 1.$$

Note that, the interchange of limits is possible because $P(G) > 0$ for all $G \in \mathcal{G}$. Hence, P is a probability measure. □

Let us also note that, the probability that a group is trivial is given by

$$P(0) = \prod_{p \notin B} \prod_{j=u+1}^{\infty} \left(1 - \frac{1}{p^j}\right).$$

Now let us consider the case of number fields. Let K be a number field and Cl_K be the class group of K. Suppose Cl_K decomposes as $\mathrm{Cl}_K = \prod_{p \text{ prime}} G_p$ into p-groups. Let us define

$$\mathrm{Cl}_K^{\notin B} := \mathrm{Cl}_K \,/ \prod_{p \in B} G_p,$$

then $\mathrm{Cl}_K^{\notin B} \in \mathcal{G}$. Now, fix a signature $\{r, s\}$ and consider the following sequence S in $\mathcal{G}$

$$S := \{\mathrm{Cl}_K^{\notin B} : K \in \mathcal{K}\},$$

where $\mathcal{K} := \{K : K \text{ is a number field of signature } \{r, s\}\}$. Then, we have the following conjecture, due to Cohen and Lenstra:

Conjecture 3.22. *For a suitably chosen set of 'bad primes' B, the above sequence S behaves like a random sequence of groups w.r.t. the above probability measure P, that is, for any measurable function $f \colon \mathcal{G} \to \mathbb{C}$ we have,*

$$\lim_{x \to \infty} \frac{\sum_{K \in \mathcal{K}, |\Delta_K| \leq x} f(\mathrm{Cl}_K^{\notin B})}{\sum_{K \in \mathcal{K}, |\Delta_K| \leq x} 1} = \mathrm{E}_P(f),$$

where $\mathrm{E}_P(f)$ is the expectation of f w.r.t. the probability measure P.

Let us make a few remarks about 'bad primes'. Note that, for quadratic case, $B = \{2\}$ gives the probabilistic model for the class groups of real quadratic fields. One can show that, if $p \mid \deg(K)$ then $p \in B$. But this does not give the complete list of 'bad primes'. There are some conjectures on the set of bad primes given by Gunter Malle, we refer the reader to [Len09] Chapter 6 for details. The debate about the set B is still vivid and not finished.

3.4 Cohen-Lenstra measure on Partitions

Fix a prime p. Then there exist a bijection between the set of all p-groups and the set of all partitions. Therefore, the Cohen-Lenstra probability measure for p-groups will induce a similar probability measure on the set of all

partitions. Recall that the local Cohen-Lenstra probability measure P_p on $\mathcal{G}_p$ is given by

$$P_p(G_p) = \frac{1}{|G_p|^u \cdot |\mathrm{Aut}(G_p)|} \prod_{r=u+1}^{\infty} \left(1 - \frac{1}{p^r}\right)$$

for $G_p \in \mathcal{G}_p$. Note that, $u = 0$ corresponds to the case of imaginary quadratic fields, $u = 1$ corresponds to the case of real quadratic fields and $u = u$ corresponds to the case of any other number fields. Suppose the p-group G_p corresponds to the partition $\lambda = (\lambda_1, \ldots, \lambda_l)$. Let $\lambda' = (\lambda'_1, \ldots, \lambda'_t)$ be the conjugate of λ. Then, we know

$$\begin{aligned} |\mathrm{Aut}(G_p)| &= \prod_{i=1}^{t} \left(\prod_{s=1}^{\lambda'_i - \lambda'_{i+1}} (1 - p^{-s}) \right) p^{\sum_{i=1}^{t} (\lambda'_i)^2} \\ &= \prod_{i=1}^{t} \left(\prod_{s=1}^{m_i(\lambda)} (1 - p^{-s}) \right) p^{\sum_{i=1}^{t} (\lambda'_i)^2}, \end{aligned}$$

where $m_i(\lambda)$ denotes the number of parts of λ of size i. Note that $\lambda'_i = m_i(\lambda) + m_{i+1}(\lambda) + \cdots$ and this implies $m_i(\lambda) = \lambda'_i - \lambda'_{i+1}$. Therefore we can write the Cohen-Lenstra probability as

$$\begin{aligned} P_p(G_p) &= \frac{1}{p^{|\lambda| u} \prod_{i=1}^{t} \left(\prod_{s=1}^{m_i(\lambda)} (1 - p^{-s}) \right) p^{\sum_{i=1}^{t} (\lambda'_i)^2}} \prod_{r=u+1}^{\infty} \left(1 - \frac{1}{p^r}\right) \\ &= \left[\prod_{r=1}^{\infty} \left(1 - \frac{v}{p^r}\right) \right] \frac{v^{|\lambda|}}{p^{\sum_{i=1}^{t} (\lambda'_i)^2} \prod_{i=1}^{t} \left(\prod_{s=1}^{m_i(\lambda)} (1 - p^{-s}) \right)} \end{aligned}$$

where $v := p^{-u}$. Then we have the following Cohen-Lenstra probability measure on partitions: For a partition $\lambda = (\lambda_1, \ldots, \lambda_l)$ with conjugate $\lambda' = (\lambda'_1, \ldots, \lambda'_t)$ we define

$$P_p(\lambda) := \left[\prod_{r=1}^{\infty} \left(1 - \frac{v}{p^r}\right) \right] \frac{v^{|\lambda|}}{p^{\sum_{i=1}^{t} (\lambda'_i)^2} \prod_{i=1}^{t} \left(\prod_{s=1}^{m_i(\lambda)} (1 - p^{-s}) \right)}$$

where $v = p^{-u}$ and u is fixed. Next, we will give some equivalent versions of this probability measure on partitions.

3.4.1 The Column Algorithm

We have the following theorem, due to J. Fulman:

Theorem 3.23. *Let $\lambda'_0 = \infty$ and define in succession $\lambda'_1, \lambda'_2, \ldots$ according to the rule that, if $\lambda'_i = a$ then $\lambda'_{i+1} = b \leq a$ with probability*

$$K(a,b) = \frac{v^b \prod_{i=1}^{a}(1-p^{-i}) \prod_{i=1}^{a}(1-vp^{-i})}{p^{b^2} \prod_{i=1}^{a-b}(1-p^{-i}) \prod_{i=1}^{b}(1-p^{-i}) \prod_{i=1}^{b}(1-vp^{-i})}$$

and $\lambda'_{i+1} = b > a$ with probability 0. Then, the algorithm outputs a partition with probability 1 and the probability that the conjugate of the partition λ' is $\lambda = (\lambda_1, \ldots, \lambda_l)$ is equal to the Cohen-Lenstra probability measure $P_p(\lambda)$.

Proof. The first part is easy to see. If $a, b \neq 0$, then the probability that $a = b$ is given by

$$K(a,a) = \frac{v^b}{p^{b^2}} = \frac{1}{p^{ub+b^2}} \longrightarrow 0 \qquad \text{as } b \to \infty,$$

which implies $b < a$ with probability 1. Therefore, with probability 1, the sequence $\lambda'_1, \lambda'_2, \ldots$ is strictly decreasing. Hence, with probability 1, the algorithm outputs a partition. For the second part, we can write

$$\begin{aligned}
&P_p(\lambda'_1 = r_1, \lambda'_2 = r_2, \ldots) \\
&= P_p(\lambda'_0 = \infty) \frac{P_p(\lambda'_0 = \infty, \lambda'_1 = r_1)}{P_p(\lambda'_0 = \infty)} \prod_{i=1}^{\infty} \frac{P_p(\lambda'_0 = \infty, \lambda'_1 = r_1, \ldots, \lambda'_{i+1} = r_{i+1})}{P_p(\lambda'_0 = \infty, \lambda'_1 = r_1, \ldots, \lambda'_i = r_i)}.
\end{aligned}$$

So, it is enough to show that

$$\begin{aligned}
&\frac{P_p(\lambda'_0 = \infty, \lambda'_1 = r_1, \ldots, \lambda'_{i-1} = r_{i-1}, \lambda'_i = a, \lambda'_{i+1} = b)}{P_p(\lambda'_0 = \infty, \lambda'_1 = r_1, \ldots, \lambda'_{i-1} = r_{i-1}, \lambda'_i = a)} \\
&= \frac{v^b \prod_{i=1}^{a}(1-p^{-i}) \prod_{i=1}^{a}(1-vp^{-i})}{p^{b^2} \prod_{i=1}^{a-b}(1-p^{-i}) \prod_{i=1}^{b}(1-p^{-i}) \prod_{i=1}^{b}(1-vp^{-i})} \qquad (1)
\end{aligned}$$

for all $i, a, b, r_1, \ldots, r_{i-1} \geq 0$. Let $i \geq 1$, then

$$\begin{aligned}
&P_p(\lambda'_0 = \infty, \lambda'_1 = r_1, \ldots, \lambda'_{i-1} = r_{i-1}, \lambda'_i = a) \\
&= \sum_{\substack{\lambda:\ \lambda'_1 = r_1, \ldots, \lambda'_{i-1} = r_{i-1} \\ \lambda'_i = a}} P_p(\lambda) \\
&= \sum_{\substack{\mathcal{G}_\mu \in \mathcal{G}_p \\ \mu'_1 = a}} \left[\prod_{r=1}^{\infty}\left(1 - \frac{v}{p^r}\right)\right] \frac{v^{r_1+\cdots+r_{i-1}+|\mu|}}{p^{r_1^2+\cdots+r_{i-1}^2+\sum_j(\mu'_j)^2} \prod_{j=1}^{i-1}\left(\prod_{s=1}^{r_j - r_{j+1}}(1-p^{-s})\right) \prod_{j=1}^{\infty}\left(\prod_{s=1}^{\mu'_j - \mu'_{j+1}}(1-p^{-s})\right)} \\
&= \frac{v^{r_1+\cdots+r_{i-1}}}{p^{r_1^2+\cdots+r_{i-1}^2} \prod_{s=1}^{r_1-r_2}(1-p^{-s}) \cdots \prod_{s=1}^{r_{i-2}-r_{i-1}}(1-p^{-s}) \prod_{s=1}^{r_{i-1}-a}(1-p^{-s})} \cdot P_p(\mu'_1 = a)
\end{aligned}$$

(where $r_i = \mu_1 = a$). Similarly, we have

$$P_p(\lambda'_0 = \infty, \lambda'_1 = r_1, \ldots, \lambda'_{i-1} = r_{i-1}, \lambda'_i = a, \lambda'_{i+1} = b)$$
$$= \frac{v^{r_1+\cdots+r_{i-1}+a}}{p^{r_1^2+\cdots+r_{i-1}^2+a^2} \prod_{s=1}^{r_1-r_2}(1-p^{-s}) \cdots \prod_{s=1}^{r_{i-1}-a}(1-p^{-s}) \prod_{s=1}^{a-b}(1-p^{-s})} \cdot P_p(\mu'_1 = b).$$

Therefore,

$$\frac{P_p(\lambda'_0 = \infty, \lambda'_1 = r_1, \ldots, \lambda'_{i-1} = r_{i-1}, \lambda'_i = a, \lambda'_{i+1} = b)}{P_p(\lambda'_0 = \infty, \lambda'_1 = r_1, \ldots, \lambda'_{i-1} = r_{i-1}, \lambda'_i = a)}$$
$$= \frac{v^a}{p^{a^2} \prod_{s=1}^{a-b}(1-p^{-s})} \cdot \frac{P_p(\mu'_1 = b)}{P_p(\mu'_1 = a)} \qquad (*)$$

where μ is any arbitrary partition.

Claim. *We have that*

$$P_p(\mu'_1 = a) = \frac{v^a \prod_{s=1}^{\infty}(1-vp^{-s})}{p^{a^2} \prod_{s=1}^{a}(1-p^{-s}) \prod_{s=1}^{a}(1-vp^{-s})}.$$

Note that a is fixed. Now, taking sum over all b such that $b \leq a$ on the both sides of $(*)$, we have

$$\sum_{b \leq a} \frac{v^a}{p^{a^2} \prod_{s=1}^{a-b}(1-p^{-s})} \frac{P_p(\mu'_1 = b)}{P_p(\mu'_1 = a)} = 1. \qquad (**)$$

We will prove the claim by induction on a, using the above equation for induction step. If $a = 0$, then

$$P_p(\mu'_1 = 0) = P_p(\lambda = 0) = \prod_{r=1}^{\infty}\left(1 - \frac{v}{p^r}\right)$$

and hence the claim is true for $a = 0$. Assume the claim for $\mu'_1 = a$. From $(**)$ we have,

$$\frac{v^{a+1}}{p^{(a+1)^2}} \cdot \frac{1}{P_p(\mu'_1 = a+1)} \left(\frac{P_p(\mu'_1 = 0)}{\prod_{s=1}^{a+1}(1-p^{-s})} + \frac{P_p(\mu'_1 = 1)}{\prod_{s=1}^{a}(1-p^{-s})} + \cdots + \frac{P(\mu'_1 = a+1)}{\prod_{s=1}^{1}(1-p^{-s})} \right) = 1.$$

Then putting the values of $\mu'_1 = k$ for $k \leq a$ one can verify that the claim is true for $\mu'_1 = a + 1$. We will give another proof of this claim later.

Note that, this claim also proves the equation (1) for $i = 0$. Now, putting the values of $P_p(\mu'_1 = a)$ and $P_p(\mu'_1 = b)$ from the claim into the equation $(*)$ we have, for $i \geq 1$,

$$\frac{P_p(\lambda'_0 = \infty, \lambda'_1 = r_1, \ldots, \lambda'_{i-1} = r_{i-1}, \lambda'_i = a, \lambda'_{i+1} = b)}{P_p(\lambda'_0 = \infty, \lambda'_1 = r_1, \ldots, \lambda'_{i-1} = r_{i-1}, \lambda'_i = a)}$$
$$= \frac{v^a}{p^{a^2} \prod_{s=1}^{a-b}(1-p^{-s})} \cdot \frac{v^b \prod_{s=1}^{\infty}(1-vp^{-s})}{p^{b^2} \prod_{s=1}^{b}(1-p^{-s}) \prod_{s=1}^{b}(1-vp^{-s})} \cdot \frac{p^{a^2} \prod_{s=1}^{a}(1-p^{-s}) \prod_{s=1}^{a}(1-vp^{-s})}{v^a \prod_{s=1}^{\infty}(1-vp^{-s})}$$
$$= \frac{v^b}{p^{b^2}} \cdot \frac{\prod_{s=1}^{a}(1-p^{-s}) \prod_{s=1}^{a}(1-vp^{-s})}{\prod_{s=1}^{a-b}(1-p^{-s}) \prod_{s=1}^{b}(1-p^{-s}) \prod_{s=1}^{b}(1-vp^{-s})}.$$

Therefore, (1) is true for $i \geq 0$ and for all $a, b, r_1, \ldots, r_{i-1} \geq 0$. Hence, we have proved the theorem □

3.4.2 Young Tableau Algorithm

First let us state the algorithm. The *Young Tableau Algorithm* is given by the following steps:

- **Step 0**: Start with λ the empty partition and $N = 1$. Also start with a collection of coins indexed by natural numbers such that coin i has probability v/p^i of heads and has probability $1 - v/p^i$ of tails.
- **Step 1**: Flip coin N.
- **Step 2a**: If coin N comes up tails, leave λ unchanged, set $N = N + 1$ and go to step 1.
- **Step 2b**: If coin N comes up heads, choose an integer $S > 0$ according to the following rule. Set $S = 1$ with probability $(p^{N-\lambda'_1} - 1)/(p^N - 1)$ and set $S = s > 1$ with probability $(p^{N-\lambda'_s} - p^{N-\lambda'_{s-1}})/(p^N - 1)$. Then increase the size of column s of λ by one and go to step 1.

Note that, we use the convention that all undefined entries of λ are 0 and λ' as usual denotes the conjugate of λ. We have the following theorem.

Theorem 3.24. *With probability 1, the algorithm outputs a finite partition. For any given partition λ, the probability that the algorithm outputs λ is equal to the Cohen-Lenstra probability $P_p(\lambda)$.*

Proof. The proof is due to Fulman given in [Ful99]. Note that, if $\lambda'_s = \lambda'_{s-1}$ for some s, then the probability that we increase the column s by size one is 0. Hence, with probability 1, we get λ'_is for $i = 1, 2, \ldots$ to be in decreasing

order. To show that the output λ of the algorithm is a partition (finite), we need to prove that, with probability 1, only finitely many λ_i's are non-zero. Let A_N be the event that the coin N comes up heads at least once. Then, $\text{Prob}(A_N) = v/p^N$ and this implies

$$\sum_N \text{Prob}(A_N) = \sum_{N=1}^{\infty} \frac{v}{p^N} = \frac{v(1/p)}{1-1/p} < \infty.$$

By Borel-Cantelli lemma, with probability 1, only finitely many A_Ns occur. For each N, let $B_{N,m}$ be the event that coin N comes up heads at least m times (where $m \geq 1$). Then, $\text{Prob}(B_{N,m}) = (v/p^N)^m$ and this implies

$$\sum_m \text{Prob}(B_{N,m}) = \sum_{m=1}^{\infty} \frac{v}{p^{Nm}} = \frac{v/p^N}{1-1/p^N} < \infty.$$

Hence, by Borel-Cantelli lemma, if a coin N comes up heads at least once, then with probability 1, it does so only finitely many times. Note that, the size of a column increases only when a coin comes up with heads. We conclude that, with probability 1, the algorithm outputs a finite partition. Now, let $\text{Prob}^N(\lambda)$ denote the probability that the algorithm outputs λ when the coin N comes up tails. We will show by induction that,

$$\text{Prob}^N(\lambda) = \frac{v^{|\lambda|}\prod_{s=1}^{N}(1-vp^{-s})\prod_{s=1}^{N}(1-p^{-s})}{\prod_{s=1}^{N-\lambda'_1}(1-p^{-s})} \prod_{i\geq 1} \frac{1}{p^{(\lambda'_i)^2}\prod_{s=1}^{m_i(\lambda)}(1-p^{-s})}$$

if $\lambda'_1 \leq N$ and 0 otherwise. Then note that, as $N \to \infty$,

$$\text{Prob}^N(\lambda) \longrightarrow P_p(\lambda).$$

This will imply that the algorithm generates partitions according to Cohen-Lenstra probability measure P_p. Now let us prove the above assertion. Clearly, $\text{Prob}^N(\lambda) = 0$ if $N < \lambda'_1$, since the size of any column increases only when a coin comes up with heads; which implies $N \geq \lambda'_1$ with probability 1. For $N \geq \lambda'_1$, we prove the assertion by induction on $|\lambda|$. Suppose $|\lambda| = 0$, i.e. λ is empty partition. Then all the coins $1, 2, \ldots, N$ comes up with tails. This implies

$$\text{Prob}^N(\lambda) = \prod_{i=1}^{N}\left(1 - \frac{v}{p^i}\right)$$

and hence the assertion is true for $|\lambda| = 0$. Now, for induction step, let $s_1 \leq s_2 \leq \cdots \leq s_k$ be the columns of λ such that changing λ by decreasing the column s_i by 1 gives the partition λ^{s_i}. Then λ is obtained from one

of these partitions λ^{s_i} by increasing the column s_i by 1. We will use the induction hypothesis for these λ^{s_i} to prove the assertion for λ. Note that, we have three possible cases:

Case 1: The partition λ was already obtained when coin $N-1$ came up with tails. The probability of this event is $\text{Prob}^{N-1}(\lambda) \cdot (1 - v/p^N)$.

Case 2: The partition λ was obtained by increasing the first column of λ^1 by 1. Note that, in this case the coin N came up with heads. The probability of this event is

$$\frac{v}{p^N} \cdot \frac{p^{N-\lambda'_1} - 1}{p^N - 1} \cdot \text{Prob}^N(\lambda^1).$$

Case 3: The partition λ was obtained by increasing the column s_i of λ^{s_i} by 1, where $s_i \neq 1$. Then the probability of this event is

$$\frac{v}{p^N} \cdot \frac{p^{N-\lambda'_{s_i}+1} - p^{N-\lambda'_{s_i-1}}}{p^N - 1} \cdot \text{Prob}^N(\lambda^{s_i}).$$

Hence we get

$$\begin{aligned}\text{Prob}^N(\lambda) = &\left(1 - \frac{v}{p^N}\right) \text{Prob}^{N-1}(\lambda) + \frac{v}{p^N} \frac{p^{N-\lambda'_1} - 1}{p^N - 1} \text{Prob}^N(\lambda^1) \\ &+ \sum_{s_i > 1} \frac{v}{p^N} \frac{p^{N-\lambda'_{s_i}+1} - p^{N-\lambda'_{s_i-1}}}{p^N - 1} \text{Prob}^N(\lambda^{s_i}).\end{aligned}$$

Now, using the induction hypothesis one can put the values of $\text{Prob}^{N-1}(\lambda)$, $\text{Prob}^N(\lambda^1)$ and $\text{Prob}^N(\lambda^{s_i})$ in the above expression to compute $\text{Prob}^N(\lambda)$. One can check that the assertion holds. This computation can be found in [Ful99]. □

3.4.3 Interpretation in Young Lattice

First we will define Young Lattice. *Young Lattice* is defined as a directed graph whose nodes are the partitions of all natural numbers and there is a directed edge from a partition λ to a partition μ if the Young diagram of λ is contained in the Young diagram of μ and $|\mu| = |\lambda| + 1$, i.e. μ is obtained from λ by adding one block. Then we have the following theorem.

Theorem 3.25. *Put weights $m_{\lambda,\mu}$ on the edges of the Young Lattice according to the rules:*

1. *$m_{\lambda,\mu} = v/p^{\lambda'_1}(p^{\lambda'_1+1} - 1)$ if the Young diagram of μ is obtained from that of λ by adding one block to column 1.*

2. $m_{\lambda,\mu} = v(p^{-\lambda'_s} - p^{-\lambda'_{s-1}})/(p^{\lambda'_1} - 1)$ *if the Young diagram of* μ *is obtained from that of* λ *by adding one block to column* $s > 1$.

Then, the following holds:

$$P_p(\lambda) = \left[\prod_{r=1}^{\infty}\left(1 - \frac{v}{p^r}\right)\right] \sum_{\lambda} \prod_{i=0}^{|\lambda|-1} m_{\gamma_i,\gamma_{i+1}}$$

where, the sum is over all directed paths γ *from the empty partition to* λ *and* γ_i *are the partitions along the path* γ.

Proof. For the proof we need to define Young Tableau. Let λ be a partition. Then the *Young Tableau* of λ is obtained by assigning the numbers $1, 2, \ldots, |\lambda|$ to each block of the Young diagram of λ, such that, each block is assigned a distinct number and the numbers increases as one moves from up to down in each row and from left to right in each column. Note that Young Tableau for a partition λ is not unique. Note that, the Young Tableau Algorithm constructs a Young Tableau (whenever we add ith block to some column of the Young diagram, we assign the number i to that block). Let T be a Young tableau and $\lambda(T)$ be the partition corresponding to T. Let $\mathrm{Prob}(T)$ be the probability that the Young tableau algorithm outputs T and let $\mathrm{Prob}^N(T)$ be the probability that the Young tableau algorithm outputs T when coin N comes up tails. Let $T_{i,j}$ denote the entry at the (i,j)th position in T. For $j \geq 2$, let $A_{i,j}$ be the number of entries in column $j-1$ which are less than $T_{i,j}$, i.e. $A_{i,j} = |\{(i', j-1) : T_{i',j-1} < T_{i,j}\}|$ and let $B_{i,j}$ be the number of entries in the first column which are less than $T_{i,j}$, i.e. $B_{i,j} = |\{(i', 1) : T_{i',1} < T_{i,j}\}|$. Then we have the following

Claim. *For a Young tableau* T,

$$\mathrm{Prob}^N(T) = \frac{v^{|\lambda(T)|}}{p^{\lambda'_1(T)^2}\prod_{s=1}^{\lambda'_1(T)}(1-p^{-s})} \frac{\prod_{r=1}^{N}(1-vp^{-r})\prod_{r=1}^{N}(1-p^{-r})}{\prod_{r=1}^{N-\lambda'_1(T)}(1-p^{-r})} \prod_{\substack{(i,j)\in\lambda(T)\\ j\geq 2}} \frac{p^{1-i} - p^{-A_{i,j}}}{b^{B_{i,j}} - 1}$$

if $\lambda'_1(T) \leq N$, *and 0 otherwise.*

First let us show that the above claim proves the theorem. Note that, for a partition λ,

$$P_p(\lambda) = \sum_{T:\ \lambda(T)=\lambda} \left(\lim_{N\to\infty} \mathrm{Prob}^N(T)\right).$$

As $N \to \infty$ we have,

$$\begin{aligned}
\mathrm{Prob}(T) &= \frac{v^{|\lambda(T)|}}{p^{\lambda_1'(T)^2}\prod_{s=1}^{\lambda_1'(T)}(1-p^{-s})}\prod_{r=1}^{\infty}\left(1-\frac{v}{p^r}\right)\prod_{\substack{(i,j)\in\lambda(T)\\ j\geq 2}}\frac{p^{1-i}-p^{-A_{i,j}}}{p^{B_{i,j}}-1} \\
&= \prod_{r=1}^{\infty}\left(1-\frac{v}{p^r}\right)\left(\prod_{i=1}^{\lambda_1'(T)-1}\frac{v}{p^{\lambda_1'(T)}(p^{\lambda_1'(T)+1}-1)}\prod_{\substack{(i,j)\in\lambda(T)\\ j\geq 2}}\frac{v(p^{1-i}-p^{-A_{i,j}})}{p^{B_{i,j}}-1}\right) \\
&= \prod_{r=1}^{\infty}\left(1-\frac{v}{p^r}\right)\cdot\prod_{i=0}^{|\lambda(T)|-1} m_{\gamma(T)_i,\gamma(T)_{i+1}}
\end{aligned}$$

where, $\gamma(T)$ is the unique path from empty partition to $\lambda(T)$, which is uniquely determined by T. Then,

$$\begin{aligned}
P_p(\lambda) &= \prod_{r=1}^{\infty}\left(1-\frac{v}{p^r}\right)\cdot\sum_{T:\,\lambda(T)=\lambda}\prod_{i=0}^{|\lambda(T)|-1} m_{\gamma(T)_i,\gamma(T)_{i+1}} \\
&= \prod_{r=1}^{\infty}\left(1-\frac{v}{p^r}\right)\cdot\sum_{\gamma}\prod_{i=0}^{|\lambda|-1} m_{\gamma_i,\gamma_{i+1}}
\end{aligned}$$

where γ ranges over all partitions from the empty partition to λ.

Now let us prove the claim. Note that, $\mathrm{Prob}^N(T) = 0$ if $N < \lambda_1'(T)$, since the size of any column increases only when a coin comes up with heads, which implies $N \geq \lambda_1'(T)$ with probability 1. For $n \geq \lambda_1'(T)$ we prove by induction on $|\lambda(T)|$. The proof is almost same as in the Theorem 3.24. Suppose $|\lambda(T)| = 0$, i.e. $\lambda(T)$ is the empty partition. Then all the coins $1, 2, \ldots, N$ comes up tails, i.e.,

$$\mathrm{Prob}^N(T) = \prod_{i=1}^{N}\left(1-\frac{v}{p^i}\right)$$

and hence the claim is true for $|\lambda(T)| = 0$. For the induction step, we have two cases:

Case 1: The largest entry in T occurs in column $s > 1$. Let T^s be the tableau obtained after removing the largest entry from T. Then,

$$\mathrm{Prob}^N(T) = \left(1-\frac{v}{p^N}\right)\mathrm{Prob}^{N-1}(T) + \frac{v}{p^N}\frac{p^{N-\lambda_s'+1}-p^{N-\lambda_{s-1}'}}{p^N-1}\,\mathrm{Prob}^N(T^s).$$

Now using induction hypothesis, putting the values for $\mathrm{Prob}^{N-1}(T)$ and $\mathrm{Prob}^{N}(T^s)$, one can check that the claim holds for $\mathrm{Prob}^{N}(T)$. See [Ful99] for this calculation.

Case 2: Suppose the largest entry of T occurs in column 1. Then,

$$\mathrm{Prob}^{N}(T) = \left(1 - \frac{v}{p^N}\right) \mathrm{Prob}^{N-1}(T) + \frac{v}{p^N} \frac{p^{N-\lambda'_1} - 1}{p^N - 1} \mathrm{Prob}^{N}(T^1)$$

and one can verify this also by putting the values of $\mathrm{Prob}^{N-1}(T)$ and $\mathrm{Prob}^{N}(T^1)$. This completes the proof of the theorem. □

Next, we will give a second proof of the Claim in the Theorem 3.23 using the theorem we just proved.

Proposition 3.26. *The following holds:*

$$\sum_{\lambda:\, \lambda'_1 = a} P_p(\lambda) = \frac{v^a \prod_{s=1}^{\infty}(1 - vp^{-s})}{p^{a^2} \prod_{s=1}^{a}(1 - p^{-s}) \prod_{s=1}^{a}(1 - vp^{-s})}.$$

Proof. We sum over all Young tableau with a parts. Let T be a Young tableau with a parts. Let h_m, for $1 \leq m \leq a-1$, be the number of blocks added to T after it becomes a tableau with m parts and before it becomes a tableau with $m+1$ parts and let h_a be the number of blocks added to T after it becomes a tableau with a parts. Suppose one takes a step along the Young lattice from a partition with m parts. Then Theorem 3.25 implies that the weight for adding a block to column 1 is $v/p^m(p^{m+1} - 1)$ and the sum of weights for adding to any other column is v/p^m. Then, the probability that the Young Tableau Algorithm yields the tableau T is equal to

$$\begin{aligned} &\prod_{s=1}^{\infty}\left(1 - \frac{v}{p^s}\right) \cdot \prod_{m=1}^{a} \frac{v}{p^m(p^{m+1} - 1)} \prod_{m=1}^{a} \left(\frac{v}{p^m}\right)^{h_m} \\ &= \frac{v^a \prod_{s=1}^{\infty}(1 - vp^{-s})}{p^{a^2} \prod_{s=1}^{a}(1 - p^{-s})} \prod_{m=1}^{a} \left(\frac{v}{p^m}\right)^{h_m}. \end{aligned}$$

Then, summing over all possible values of $h_m \geq 0$,

$$\begin{aligned} \sum_{\lambda:\, \lambda'_1 = a} P_p(\lambda) &= \frac{v^a \prod_{s=1}^{\infty}(1 - vp^{-s})}{p^{a^2} \prod_{s=1}^{a}(1 - p^{-s})} \prod_{m=1}^{a} \sum_{h_m=0}^{\infty} \left(\frac{v}{p^m}\right)^{h_m} \\ &= \frac{v^a \prod_{s=1}^{\infty}(1 - vp^{-s})}{p^{a^2} \prod_{s=1}^{a}(1 - p^{-s})} \prod_{m=1}^{a} \frac{1}{1 - v/p^m} \\ &= \frac{v^a \prod_{s=1}^{\infty}(1 - vp^{-s})}{p^{a^2} \prod_{s=1}^{a}(1 - p^{-s})} \frac{1}{\prod_{s=1}^{a}(1 - vp^{-s})}. \end{aligned}$$

□

Next, we will state a connection between Cohen-Lenstra probability measure P_p and conjugacy classes in $\mathrm{GL}_n(\mathbb{F}_p)$. One can show that, each conjugacy class of $\mathrm{GL}_n(\mathbb{F}_p)$ is uniquely specified by the following data: For each monic irreducible polynomial ϕ over $\mathbb{F}_p$, associate a partition λ_ϕ such that

1. $|\lambda_X| = 0$
2. $\sum_\phi |\lambda_\phi| \deg(\phi) = n$.

This follows by considering the rational canonical forms of conjugacy classes of $\mathrm{GL}_n(\mathbb{F}_p)$. See [Len10a] and [Ful97] for details. We have the following theorem.

Theorem 3.27. *Let ϕ be a monic polynomial over $\mathbb{F}_p$ of degree 1 and let λ be a partition. Now pick a random matrix in $\mathrm{GL}_n(\mathbb{F}_p)$ uniformly at random and consider it's conjugacy class and the associated set of partitions for that class, say, this is $\{\lambda_\phi : \phi$ is monic irreducible polynomial $\}$. Then, the probability that $\lambda_\phi = \lambda$ tends to the Cohen-Lenstra probability measure $P_p(\lambda)$ as $n \to \infty$.*

Proof. See [Ful97] for the proof. □

Bibliography

[Bur80] D. M. Burton. *Elementary Number Theory.* Allyn and Bacon, Inc., 1980.

[BV07] J. Buchmann and U. Vollmer. *Binary Quadratic Forms.* Springer, 2007.

[Coh96] H. Cohen. *A course in Computational Algebraic Number Theory.* Springer, 1996.

[Ful97] J. Fulman. *Probability in the Classical Groups over Finite Fields: Symmetric Functions, Stochastic Algorithms, and Cycle Indices.* PhD thesis, Harvard University, 1997.

[Ful99] J. Fulman. A probabilistic approach toward conjugacy classes in the finite general linear and unitary groups. *Journal of Algebra*, 1999.

[FW89] E. Friedman and L. Washington. On the distribution of divisor class groups of curves over a finite field. *Theorie des nombers*, 1989.

[Jan96] G. J. Janusz. *Algebraic Number Fields.* American Mathematical Society, second edition, 1996.

[Jon61] B. W. Jones. *The Arithmetic Theory of Quadratic Forms.* The Mathematical Association of America, 1961.

[Kez12] Y. Kezuka. The class number problem. Master's thesis, Imperial College London, 2012.

[Len09] J. Lengler. *The Cohen-Lenstra Heuristic for Finite Abelian Groups.* PhD thesis, Universitat des Saarlandes, 2009.

[Len10a] J. Lengler. The cohen-lenstra heuristic: Methodology and results. *Journal of Algebra*, 2010.

[Len10b] J. Lengler. The global cohen-lenstra heuristic. *arXiv:0912.4977v2*, 2010.

[ME04] M. R. Murty and J. Esmonde. *Problems in Algebraic Number Theory*. Springer, second edition, 2004.

[Mol10] R. A. Mollin. *Algebraic Number Theory*. CRC Press, second edition, 2010.

[MW92] R. A. Mollin and H. C. Williams. Computation of the class number of a real quadratic field. 1992.

[ST01] I. Stewart and D. Tall. *Algebraic Number Theory and Fermat's Last Theorem*. A K Peters, third edition, 2001.

[Wil85] H. C. Williams. Continued fractions and number theoritic computations. *Rocky Mountain Journal of Mathematics*, 15(2), 1985.