Class Number Statistics

LINUS ENVALL

Mathematics
Date: 18 July, 2025
Supervisor: Pär Kurlberg
Examiner: Pär Kurlberg
Department of Mathematics

Abstract

We expand on the Cohen-Lenstra heuristics by hypothesizing that there is a second order term to its predictions of divisibility of the class number and density of Sylow subgroups. This term is inspired by the second order term found to the Davenport-Heilbronn Theorem by M. Bhargava, A. Shankar, J. Tsimerman, T. Taniguchi and F. Thorne, as well as results by B. Hough. We then perform empirical tests, which seem to support our hypotheses. This paper also includes an introduction to class numbers and class groups.

Contents

Chapter 1

Introduction

Class numbers and class groups are two important concepts in number theory. Gauss introduced the concepts in *Disquisitiones Arithmeticae*, his groundbreaking monograph in number theory from 1801. Since then it has been an active area of research that intersect the development of both number theory and abstract algebra.

In this paper we use a probabilistic model to analyze the 3-, 5-, and 7-divisibility of the class number, as well as the 3-Sylow subgroups of the class group.

In chapter 2, we give some background and introduce the results that we will base our analysis on. This includes the Cohen-Lenstra heuristics, which gives an asymptotic prediction to the behavior of the class groups and class numbers. In chapter 3, we form our probabilistic model, as well as construct a test of its accuracy. In chapter 4 we use a computer to generate large amount of class numbers, to test our hypotheses using this data. Finally, in Chapter 5, we analyze the obtained results.

We discover that we can improve on the Cohen-Lenstra heuristics by introducing a second, higher order term inspired by a second order term to the Davenport-Heilbronn Theorem found by M. Bhargava, A. Shankar, J. Tsimerman, T. Taniguchi and F. Thorne. The errors to this new model seems to be normally distributed for both p-divisibility and 3-Sylow subgroups. We also discover that the size of this higher order term do not have a easy to state relationship with the Cohen-Lenstra prediction.

Chapter 2

Background

We will begin by giving an introduction to class numbers and class groups. There are two different, but equivalent, ways of constructing the class group. In section 2.1, we construct it using quadratic forms, which is how class groups first were discovered. In section 2.2, we construct them as the fractional ideal group of a quadratic number field, modulo principal ideals. This is a more modern way of looking at the class group. In section 2.3 and 2.4, we introduce some important results about class numbers, including the Cohen-Lenstra heuristics and the Davenport-Heilbronn Theorem.

2.1 Quadratic Forms

We will start by giving some definitions and terminology regarding quadratic forms.

Definition 1. An *integral quadratic form in two variables* (henceforth simply called *form*) is a function in two variables that can be written as

$$f(x, y) = ax^2 + bxy + cy^2, \quad a, b, c \in \mathbb{Z}.$$

A form is said to be *primitive* if $\gcd(a, b, c) = 1$. The *discriminant* of f is defined as $d := b^2 - 4ac$. Note that we must have $d \equiv 0 \mod 4$ (if b is even) or $d \equiv 1 \mod 4$ (if b is odd).

Definition 2. Two forms $f(x, y)$ and $g(x, y)$ is said to be *equivalent* if $f(x, y) = g(px + qy, rx + sx)$, where $\det \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) = 1$.

Definition 3. We say that a number m is *represented* by a form $f(x, y)$ if $m = f(x, y)$ for some $x, y \in \mathbb{Z}$.

We now state some basic results. For proofs the reader should consult a textbook on the topic, for example [1].

Theorem 1. Equivalent forms represent the same numbers, and the discriminant of equivalent forms are equal.

Theorem 2. If $d < 0$, $f(x, y)$ only represent numbers of the same sign. If $d < 0$ and $a > 0$, $f(x, y)$ only represents positive numbers and we call $f(x, y)$ *positive definite*. Similarly, if $d < 0$ and $a < 0$, then $f(x, y)$ only represents negative numbers and is called *negative definite*. If $d < 0$, then $f(x, y)$ is called *indefinite*.

We will mainly be concerned with primitive positive definite forms. Since equivalent forms represent the same numbers, it is natural to look at the equivalence classes.

Theorem 3. Equivalence of forms is an equivalence relation, and we denote the equivalence class of $f(x, y)$ as $[f(x, y)]$. Every primitive positive definite form is equivalent to a unique *reduced form*, which is a primitive positive definite form $ax^2 + bxy + cy^2$ with $|b| \leq a \leq c,$ and $b \geq 0$ if $|b| = a$ or $a = c$.

Definition 4. For any integer $d < 0$, let $h(d)$ denote the number of equivalence classes of primitive positive definite forms of discriminant d. By the theorem above, this is equal to the number of reduced forms of discriminant d. $h(d)$ is called the *class number* of d.

Since a reduced form $ax^2 + bxy + cy^2$ fulfills $|b| \leq a$ and $c \geq a$ we have

$$d = b^2 - 4ac \leq a^2 - 4a^2 = -3a^2 \implies a \leq \sqrt{\frac{-d}{3}}.$$

Through this inequality, it is possible to find all the reduced forms of a certain discriminant $d < 0$. For every $0 \leq a \leq \sqrt{-d/3}$, we can simply go through all $|b| \leq a$ and see if $c = \frac{b^2 - D}{4a}$ turns out to be an integer larger or equal to a. We also have to remember that for $ax^2 + bxy + cy^2$ to be a reduced form, we must have $\gcd(a, b, c) = 1$ and $b \geq 0$ if $|b| = a$ or $a = c$. This argument makes it clear that $h(d)$ is finite for all $d < 0$. Below we see the class numbers and the

d	$h(d)$	Reduced forms	d	$h(d)$	Reduced forms
-4	1	x^2+y^2	-3	1	x^2+xy+y^2
-8	1	x^2+2y^2	-7	1	$x^2+xy+2y^2$
-12	1	x^2+3y^2	-11	1	$x^2+xy+3y^2$
-16	1	x^2+4y^2	-15	2	$x^2+xy+4y^2$, $2x^2+xy+2y^2$
-20	2	x^2+5y^2, $2x^2+2xy+3y^2$	-19	1	$x^2+xy+5y^2$
-24	2	x^2+6y^2, $2x^2+3y^2$	-23	3	$x^2+xy+6y^2$, $2x^2 \pm xy+3y^2$
-28	1	x^2+7y^2	-27	1	$x^2+xy+7y^2$
-32	2	x^2+8y^2, $3x^2+2xy+3y^2$	-31	3	$x^2+xy+8y^2$, $2x^2 \pm xy+4y^2$
-36	2	x^2+9y^2, $2x^2+2xy+5y^2$	-35	2	$x^2+xy+9y^2$, $3x^2+xy+3y^2$
-40	2	x^2+10y^2, $2x^2+5y^2$	-39	4	$x^2+xy+10y^2$, $2x^2 \pm xy+5y^2$, $3x^2+3xy+4y^2$

Table 2.1: Class numbers for small $d < 0$

reduced forms for the first few negative discriminants. Note that $h(d) = 0$ if $d \equiv 2, 3 \mod 4$.

Let us look at the reduced forms of discriminant $d = -20$, $x^2 + 5y^2$ and $2x^2 + 2xy + 3y^2$. One might notice that two numbers represented by the form $x^2 + 5y^2$ (for example $6 = 1^2 + 5 \cdot 1^2$ and $9 = 2^2 + 5 \cdot 1^2$) will have their product represented by the same form. ($6 \cdot 9 = 54 = 7^2 + 5 \cdot 1^2$) This will always be the case, as

$$(x^2 + 5y^2)(z^2 + 5w^2) = (xz + 5yw)^2 + 5(xw - yz)^2.$$

One might also notice that the product of any two numbers represented by $2x^2 + 2xy + 3y^2$ is on the form $x^2 + 5y^2$. This can also be shown to always hold:

$$(2x^2+2xy+3y^2)(2z^2+2zw+3w^2) = (xz+xw+yz+3yw)^2+5(xw-yz)^2.$$

Furthermore, a number represented by $2x^2+2xy+3y^2$ multiplied by a number represented by x^2+5y^2 gives a number represented by $2x^2+2xy+3y^2$, namely

$$(2x^2 + 2xy + 3y^2)(z^2 + 5w^2) = 2X^2 + 2XY + 3Y^2$$
$$\text{where } X = xz - xw - 3yw \text{ and } Y = 2xw + yz + yw.$$

These observations hint at it being possible to define a Abelian group structure on the set $\{[x^2 + 5y^2], [2x^2 + 2xy + 3y^2]\}$ as follows:

$$[x^2 + 5y^2] * [x^2 + 5y^2] = [x^2 + 5y^2],$$

$$[x^2 + 5y^2] * [2x^2 + 2xy + 3y^2] = [2x^2 + 2xy + 3y^2],$$
$$[2x^2 + 2xy + 3y^2] * [2x^2 + 2xy + 3y^2] = [x^2 + 5y^2].$$

It also seems like this group is isomorphic to $\mathbb{Z}_2$. Although not obvious, it is always possible to define a group structure for any discriminant $d < 0$ in this manner, which is called the *class group*. It is however complicated by the fact that the product of two forms can sometimes be written as different, non-equivalent forms, which makes it necessary to somehow restrict the definition of composition of two forms presented above.

Gauss was the first to successfully construct the class group in *Disquisitiones Arithmeticae* from 1801. Writing Gauss's composition law explicitly is rather complicated, which is why we present a definition due to Dirichlet, who simplified and streamlined Gauss's method at the end of the 19th century.

Definition 5. To find the *composition* of two forms $[a_1x^2 + b_1xy + c_1y^2]$ and $[a_2x^2 + b_2xy + c_2y^2]$ of discriminant d, first make sure that $\gcd(a_1, a_2, (b_1 + b_2)/2) = 1$, finding fitting equivalences to $a_1x^2 + b_1xy + c_1y^2$ and $a_2x^2 + b_2xy + c_2y^2$ if needed. Then let $A = a_1a_2$, B be a solution to the system of congruences

$$\begin{cases} B \equiv b_1 & \mod 2a_1 \\ B \equiv b_2 & \mod 2a_2 \\ B^2 \equiv d & \mod 4a_1a_2 \end{cases}$$

(which can be shown to always have solutions), and $C = \frac{B^2-d}{4A}$. The *composition* of $[a_1x^2 + b_1xy + c_1y^2]$ and $[a_2x^2 + b_2xy + c_2y^2]$ is then defined as

$$[a_1x^2 + b_1xy + c_1y^2] * [a_2x^2 + b_2xy + c_2y^2] = [Ax^2 + Bxy + Cy^2].$$

It is then possible to show that this definition induces a group structure on the equivalence classes of primitive positive definite forms of discriminant d, summarized in theorem 4.

Theorem 4. Let $d < 0$ with $d \equiv 0, 1 \mod 4$, and let $H(d)$ be the set of equivalence classes of primitive positive definite forms of discriminant d. Then composition $*$ is a well-defined binary operation on $H(d)$, and it makes $H(d)$ a finite Abelian group of order $h(d)$, with identity element

$$\begin{cases} [x^2 + ny^2] & \text{if } d = -4n \\ [x^2 + xy + ny^2] & \text{if } d = -4n + 1 \end{cases}$$

and inverses $[ax^2 + bxy + cy^2]^{-1} = [ax^2 - bxy + cy^2]$. We call $H(d)$ the class group

Table 2.2 list $h(d)$ and $H(d)$ for some small values of d. It might look like all

d	$h(d)$	$H(d)$	d	$h(d)$	$H(d)$
-4	1	$\mathbb{Z}_1$	-3	1	$\mathbb{Z}_1$
-8	1	$\mathbb{Z}_1$	-7	1	$\mathbb{Z}_1$
-12	1	$\mathbb{Z}_1$	-11	1	$\mathbb{Z}_1$
-16	1	$\mathbb{Z}_1$	-15	2	$\mathbb{Z}_2$
-20	2	$\mathbb{Z}_2$	-19	1	$\mathbb{Z}_1$
-24	2	$\mathbb{Z}_2$	-23	3	$\mathbb{Z}_3$
-28	1	$\mathbb{Z}_1$	-27	1	$\mathbb{Z}_1$
-32	2	$\mathbb{Z}_2$	-31	3	$\mathbb{Z}_3$
-36	2	$\mathbb{Z}_2$	-35	2	$\mathbb{Z}_2$
-40	2	$\mathbb{Z}_2$	-39	4	$\mathbb{Z}_4$

Table 2.2: Class groups for small $d < 0$

class groups are cyclic. This is indeed most common case, but there are also examples like $H(-84) = \mathbb{Z}_2 \times \mathbb{Z}_2$.

2.2 The Ideal Class Group

It was through quadratic forms the theory of class groups originally took form, but a modern treatise of the subject usually constructs the class group using the tools of abstract algebra. This way of seeing the group might feel more abstract, but the proofs using this construction tend to be more straightforward than the corresponding proofs with quadratic forms. For completeness, we also include this way of defining class groups, with theorem 5 giving the isomorphism to the quadratic forms construction. The only part of this section that is referenced in the rest of the paper is the definition of a fundamental discriminant. We omit the proofs, and again refer interested readers to a textbook like [1].

Definition 6. A *number field* K is, for our purposes, a subset of $\mathbb{C}$ that is a finite-dimensional vector space over $\mathbb{Q}$. An *algebraic integer* is a complex number which is the root of some monic polynomial with integer coefficients. We denote the set of algebraic integers in K as $\mathcal{O}_\mathcal{K}$.

Definition 7. A *quadratic number field* is a field on the form $K = \mathbb{Q}(\sqrt{N})$ where $N \neq 0, 1$ is a squarefree integer. The *discriminant* of K is defined to

be

$$d_k = \begin{cases} N & \text{if } N \equiv 1 \mod 4 \\ 4N & \text{else} \end{cases}.$$

Definition 8. A *fundamental discriminant* is an integer d that is the discriminant of some quadratic field. That is, $d \neq 1$ and either $d \equiv 1 \mod 4$ and d is squarefree, or $d = 4N$ where $N \equiv 2$ or $3 \mod 4$ and N is squarefree.

Definition 9. A *fractional ideal* is an ideal on the form $\alpha\mathcal{I}$ were $\alpha \in K$ and $\mathcal{I}$ is an ideal of $\mathcal{O}_K$. The set of fractional ideals I_K is a group under the multiplication $(\alpha\mathcal{I})(\beta\mathcal{J}) = (\alpha\beta)(\mathcal{I}\mathcal{J})$, where the product of ideals is defined as $\mathcal{I}\mathcal{J} := \{\sum_{k=1}^{n} i_k j_k : i_k \in \mathcal{I}, j_k \in \mathcal{J}, n \in \mathbb{N}\}$. A *principal fractional ideal* is defined to be a fractional ideal on the form $\alpha\mathcal{O}_K$ with α being a unit in K. The set of principal fractional ideals P_K forms a subgroup of I_K.

Definition 10. The *ideal class group* of K is the quotient I_K/P_K, denoted $H(\mathcal{O}_\mathcal{K})$.

Theorem 5. Let $d < 0$ be a fundamental discriminant. The class group $H(d)$ and the ideal class group of $K = \mathbb{Q}(\sqrt{d})$ are isomorphic, using the isomorphism

$$\phi : H(d) \longrightarrow H(\mathcal{O}_K),$$
$$\phi([ax^2 + bxy + cy^2]) = [\langle a, (-b + \sqrt{d})/2\rangle],$$

where $\langle a, (-b + \sqrt{d})/2\rangle = \{ma + n(-b + \sqrt{d})/2 : m, n \in \mathbb{Z}\} \in \mathcal{O}_K$.

2.3 The Cohen-Lenstra Heuristics

The connection between a discriminant $d < 0$, its class number $h(d)$ and the class group $H(d)$ turns out to be quite complicated. In this section we will list some conjectures and results that will be relevant for this paper.

The Cohen-Lenstra Heuristics, published 1984 by H. Cohen and H.W. Lensta, gives a probabilistic model for the behavior of $H(d)$.

Conjecture 1. (Cohen-Lenstra Heuristics)[2] For an odd prime p and $d < 0$,

$$\Pr(p|h(d)) = 1 - \prod_{k=1}^{\infty}\left(1 - \frac{1}{p^k}\right).$$

and

$$\Pr(\text{Syl}_p(H(d)) \cong G) = \frac{1}{\text{Aut}(G)} \prod_{k=1}^{\infty} \left(1 - \frac{1}{p^k}\right).$$

The *p-Sylow subgroup* of a group G is defined as follows. If G is a finite group and $|G| = kp^n$ where $p \nmid k$, a p-Sylow subgroup of G is a subgroup of order p^k. The Sylow theorems show that there always exists a p-Sylow subgroup, and that if there are more than one, they are isomorphic. We can thus talk about *the* p-Sylow subgroup of G, and we denote it as $\text{Syl}_p(G)$.

In Conjecture 1, $\Pr(p|h(d))$ should be understood as

$$\Pr(p|h(d)) = \lim_{X\to\infty} \frac{\#\{-X \le d < 0 : d \text{ fundamental, } p|h(d)\}}{\#\{-X \le d < 0 : d \text{ fundamental}\}},$$

and similarly for $\Pr(\text{Syl}_p(H(d)) \cong G)$. For example, the Cohen-Lenstra heuristics predicts that

$$\Pr(3|h(d)) = 1 - \prod_{k=1}^{\infty} \left(1 - \frac{1}{3^k}\right) \approx 0.44,$$

not $1/3$, as one might guess. Another interesting consequence of the Cohen-Lenstra heuristics is that the probability that the odd part of the class group is cyclic, is approximately equal to 98%.

Theorem 6. [2, p.57] Assuming the Cohen-Lenstra heuristics, we have the following asymptotic equivalence

$$\sum_{\substack{-X<d<0 \\ d \text{ fund.}}} |H(d)[p]| \sim 2 \sum_{\substack{-X<d<0 \\ d \text{ fund.}}} 1,$$

where $G[k]$ denotes the k-torsion subgroup of the group G, i.e. $G[k] = \{g \in G : g^k = 0\}$.

One of the few results that can bee seen as evidence for the Cohen-Lenstra heuristics is the Davenport-Heilbronn Theorem, proved by H. Davenport and H. Heilbronn in 1971, which says that theorem 6 holds for $p = 3$. We can also rewrite the right side using the standard result

$$\#\{-X \le d < 0 : d \text{ fundamental}\} \sim 3X/\pi^2.$$

Theorem 7. (Davenport-Heilbronn Theorem)[3]

$$\sum_{\substack{-X<d<0 \\ d \text{ fund.}}} |H(d)[3]| = 2 \sum_{\substack{-X<d<0 \\ d \text{ fund.}}} 1 + o(X) = C \cdot X + o(X)$$

where $C = 6/\pi^2 \approx 0.6079$.

A higher order term of size $X^{5/6}$ to the Davenport-Heilbronn Theorem was independently found and proved by M. Bhargava, A. Shankar and J.Tsimerman in 2012 and by T. Taniguchi and F. Thorne in 2013.

Theorem 8. [4][5] Let $H(d)[3]$ denote the p-torsion subgroup of G.

$$\sum_{\substack{-X<d<0 \\ d \text{ fund.}}} |H(d)[3]| = C \cdot X + D \cdot X^{5/6} + o(X^{5/6})$$

where $C = 6/\pi^2$, $D = \frac{8\sqrt{3}\zeta(1/3)}{5\Gamma(2/3)^3} \prod_p \left(1 - \frac{p^{1/3}+1}{p(p+1)}\right) \approx -0.4149$.

Finally, we note that

$$\sum_{\substack{-X<d<0 \\ d \text{ fund.}}} |H(d)[3]| \sim$$

$$\left(\sum_{\substack{-X<d<0 \\ d \text{ fund.}}} 1 \right) \cdot \left(\Pr(3 \nmid h(d)) + 3 \Pr(H(d)[3] = \mathbb{Z}_3) + 9 \Pr(H(d)[3] = \mathbb{Z}_3^2) + \cdots \right).$$

Since $\sum |H(d)[3]|$ has a second term, it is not unreasonable to suspect that there could be a secondary main term to $\Pr(3 \nmid h(d))$, and thus $\Pr(3|h(d))$, as well. We will investigate this in the next chapter.

2.4 Other results

The Cohen-Lenstra Heuristics only concern odd primes. The reason for this is that the prime 2 behaves in a quite special way in the class group. The following result is important in this regard.

Theorem 9. [1, Prop. 3.11] Let $d < 0$ be a fundamental discriminant and t be the number of prime divisors of d. Then $|H(d)[2]| = 2^{t-1}$

This shows that the 2-part of the class group is in a way "almost deterministic", which means that the probabilistic model of the Cohen-Lenstra heuristics is not suitable to analyze it.

Since -4 and -8 are the only negative fundamental determinants of the form -2^n, and thus the only two even negative fundamental determinants with only one prime divisor, we get the following corollary.

Corollary 1. For fundamental discriminants $d < -8$,

$$h(d) \text{ odd} \iff -d \text{ prime}.$$

Lastly, to compute $|\mathrm{Aut}(G)|$ in conjecture 1, we need the following theorem.

Theorem 10. [6] If $G = \mathbb{Z}/p^{e_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p^{e_n}\mathbb{Z}$ with $e_1 \leq e_2 \leq \cdots \leq e_n$, then

$$|\mathrm{Aut}(G)| = \prod_{k=1}^{n} \left(p^{d_k} - p^{k-1}\right) \prod_{j=1}^{n} (p^{e_j})^{n-d_j} \prod_{i=1}^{n} (p^{e_i-1})^{n-c_i+1}.$$

where $c_k := \min\{l : e_l = e_k\}$ and $d_k := \max\{l : e_l = e_k\}$.

In particular, we find that:

$$\begin{aligned} |\mathrm{Aut}(\mathbb{Z}_3)| &= 2, \\ |\mathrm{Aut}(\mathbb{Z}_9)| &= 6, \\ |\mathrm{Aut}(\mathbb{Z}_3 \times \mathbb{Z}_3)| &= 48, \\ |\mathrm{Aut}(\mathbb{Z}_{27})| &= 18, \\ |\mathrm{Aut}(\mathbb{Z}_9 \times \mathbb{Z}_3)| &= 108, \\ |\mathrm{Aut}(\mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3)| &= 11\,232. \end{aligned}$$

If $G = \mathbb{Z}_p^r$, then $|\mathrm{Aut}(G)|$ can also be found by the fact $\mathrm{Aut}(G) = \mathrm{GL}_r(\mathbb{F}_p)$.

Chapter 3

Method

To sidestep the points about the even part of the class group laid forward in section 2.4, we will focus on odd class numbers, which according to corollary 1 is the same thing as only looking at negative prime discriminants.

For the data used in this paper, we calculated $h(-d)$ for all primes $d \equiv 3 \mod 4$ satisfying $0 < d \leq 34\,576\,772\,507$. The number $3.4 \cdot 10^{10}$ is simply a consequence of how many class numbers it was possible to compute in 60 hours with the computing resources we had at hand. This resulted in $N_d = 744\,582\,042$ class numbers.

Furthermore, for all d for which $9||h(-d)$ or $27||h(-d)$, we also calculated the full group structure of $H(-d)$.[1]

The computations were made using the computer algebra system PARI/GP, with `qfbclassno` computing class numbers, and `quadclassunit` computing class groups. The data analysis and figures were made with Python.

[1] The notation $p^n||k$ means that $p^n|k$, but $p^{n-1} \nmid k$.

3.1 3-divisibility

Our goal is to find a probabilistic model that predict the 3-divisibility of $h(-d)$.

Let $\mathbb{X} = \{\mathbb{X}(d) : d \text{ prime}\}$ be a collection random variables defined as

$$\mathbb{X}(d) := \begin{cases} 1 & \text{with probability } f(d) \\ 0 & \text{with probability } 1 - f(d) \end{cases},$$

i.e. as Bernoulli distributions with parameters $f(d)$, for some function $f(d)$. For each d, the random variable $\mathbb{X}(d)$ models whether $h(-d)$ is divisible by three ($\mathbb{X}(d) = 1$) or not ($\mathbb{X}(d) = 0$). Whether $3|h(-d)$ or not is of course deterministic, but we model this behavior with a random variable $\mathbb{X}(d)$. Our goal is now to find a choice of the function $f(d)$ that fit our data.

If we assume that this probability does not depend on d, the Cohen-Lenstra heuristics suggests that $f(d) = 1 - \prod_{k=1}^{\infty}\left(1 - \frac{1}{3^k}\right) \approx 0.43987$ would be a good choice. Let us call this model the constant Cohen-Lenstra model.

However, as to be seen in chapter 4, this model do not fit the data for small d. The error does however shrink as we look at negative discriminants of larger magnitudes. This suggests that there might be a second, higher order correction to $f(d)$.

We will now form our hypothesis for $f(d)$, and then explain the theoretical reasoning behind it.

Hypothesis 1. A good choice for $f(d)$ is

$$f_3(d) = A_3 + \frac{B_3}{d^{1/6}}$$

where $A_3 = 1 - \prod_{k=1}^{\infty}\left(1 - \frac{1}{3^k}\right) \approx 0.43987$, for some B_3. We will compare this to the constant Cohen-Lenstra model $f_3^{CL}(d) = A_3$.

Hypothesis 1 is inspired by theorem 8, since under this hypothesis, [2]

$$\sum_{\substack{-X<-d<0 \\ 3|h(-d)}}{}' 1 \sim \frac{A_3}{2}\frac{X}{\log X} + \frac{3B_3}{5}\frac{X^{5/6}}{\log X}. \tag{3.1}$$

[2]The notation $\sum'$ is to be understood as taking the sum over all prime $d \equiv 3 \bmod 4$

The details can be found in appendix 6.1. This form is very similar to theorem 8, but with both terms divided by $\log X$, which is a remnant from us taking our sum over all prime discriminants, instead of all fundamental discriminants. The exponent $d^{-1/6}$ in hypothesis 1 was chosen to get the exponent $X^{5/6}$ in this calculation. As is shown in appendix 6.1, this was done by simply subtracting 1 from $5/6$.

To test a model $f(d)$, we group the N_d data points into sets of size N, denoted $(G_i)_{1\leq i\leq N_G}$, where $N_G = \lfloor N_d/N \rfloor$. This is done in such a way that G_1 contains the N smallest discriminants, G_2 contains the next N discriminants, and so on. For each G_i, we calculate

$$P_i := \Pr(3|h(-d) : d \in G_i) = \frac{\#\{d \in G_i : 3|h(-d)\}}{\#G_i} = \frac{\#\{d \in G_i : 3|h(-d)\}}{N}.$$

Since the magnitude of d does not change much within a specific G_i, the value P_i should be a good approximation of $f(d_i')$, where $d_i' := (\min G_i + \max G_i)/2$, assuming that $f(d)$ is continuous.

Furthermore, assuming that $f(d)$ is the right model, P_i can be seen as an observation of the variable

$$\frac{1}{N}\sum_{d\in G_i} \mathbb{X}(d) \approx \frac{1}{N}\sum_{n=1}^{N} \mathbb{X}(d_i') =: \overline{\mathbb{X}}_i.$$

The expectation and variance of $\overline{\mathbb{X}}_i$ are as follows:

$$\mathbb{E}[\overline{\mathbb{X}}_i] = f(d_i')$$
$$\text{Var}[\overline{\mathbb{X}}_i] = \frac{1}{N} f(d_i')(1 - f(d_i')).$$

In hypothesis 1, B_3 is left unspecified. Ideally, we would want to find this value using analytically, maybe by comparing equation (3.1) to theorem 8, but since the latter is summing over the 3-torsion subgroups this is quite hard to do. Instead, we will use the least square method to try to find an approximation of B_3.

From the data points (d_i', P_i), we can then use the least square method to find the coefficient B_3 that makes the model $P_i = A_3 + \frac{B_3}{(d_i')^{1/6}}$ best fit our data. Through a standard regression calculation, this coefficient is calculated to be

$$B_3 = \frac{\langle Y, Z\rangle}{\langle Z, Z\rangle}, \text{ where } Y = \begin{bmatrix} P_1 - A_3 \\ \vdots \\ P_{N_G} - A_3 \end{bmatrix} \text{ and } Z = \begin{bmatrix} (d_1')^{-1/6} \\ \vdots \\ (d_{N_G}')^{-1/6} \end{bmatrix}.$$

Finally, the normalized random variables

$$\mathbb{Y}_i := \frac{\overline{\mathbb{X}}_i - \mathbb{E}[\overline{\mathbb{X}}_i]}{\sqrt{\text{Var}[\overline{\mathbb{X}}_i]}}$$

should, according to the central limit theorem, approximately have the standard normal distribution $\mathcal{N}(0,1)$. We will thus also make a histogram of the values

$$Y_i := \frac{P_i - f(d_i')}{\sqrt{\frac{1}{N} f(d_i')(1 - f(d_i'))}},$$

to see if this holds. We will do this for both $f_3(d)$ and $f_3^{CL}(d)$ in place of $f(d)$. If the histogram is close to a standard bell curve, our model would seem to fit the data well.

3.2 5- and 7-divisibility

We will repeat the method for 3-divisibility for 5- and 7-divisibility as well, where we model the p-divisibility of $h(-d)$ as

$$\mathbb{X}(d) := \begin{cases} 1 & \text{with probability } f_p(d) \\ 0 & \text{with probability } 1 - f_p(d) \end{cases}.$$

Inspired by 3-divisibility, it would be natural to use the model

$$f_k(d) = A_p + \frac{B_p}{d^a}$$

for some a, where $A_p = 1 - \prod_{k=1}^{\infty} \left(1 - \frac{1}{p^k}\right)$.

Unfortunately, there is no corresponding result to theorem 8 for $p = 5$ or 7, so we don't know what a choose. However, in the paper "Equidistribution of bounded torsion CM points"[7] from 2019, B. Hough finds some theoretical indications of a higher order term of $X^{1/2+1/k}$ to the right side of conjecture 6. By a similar analysis as for the $k = 3$ case, a good choice for a might be $a = -(1/2 + 1/5 - 1) = 3/10$ for $k = 5$ and $a = -(1/2 + 1/7 - 1) = 5/14$ for $k = 7$. Our hypotheses will thus be:

Hypothesis 2. For some B_5,

$$f_5(d) = A_5 + \frac{B_5}{d^{3/10}}$$

where $A_5 = 1 - \prod_{k=1}^{\infty}\left(1 - \frac{1}{5^k}\right) \approx 0.23967$ is a good model of 5-divisibility of $h(-d)$.

Hypothesis 3. For some B_7,

$$f_7(d) = A_7 + \frac{B_7}{d^{5/14}}$$

where $A_7 = 1 - \prod_{k=1}^{\infty}\left(1 - \frac{1}{7^k}\right) \approx 0.16320$ is a good model of 7-divisibility of $h(-d)$.

We will also be comparing to the the constant Cohen-Lenstra models $f_5^{CL}(d) = A_5$ and $f_7^{CL}(d) = A_7$.

3.3 3-Sylow subgroups

Lastly, we will also look at the 3-Sylow subgroups of $H(-d)$, specifically on the cases $9||h(-d)$ and $27||h(-d)$.

In this case of $9||h(-d)$, the 3-Sylow subgroup $\text{Syl}_3(H(-d))$ can take on the form $\mathbb{Z}_9$ or $\mathbb{Z}_3 \times \mathbb{Z}_3$. According to conjecture 1, the probabilities for these are $\Pr(\text{Syl}_3(H(-d)) \cong \mathbb{Z}_9) \approx 0.56/6$, $\Pr(\text{Syl}_3(H(-d)) \cong \mathbb{Z}_3 \times \mathbb{Z}_3) \approx 0.56/48$, respectively.

Similarly, if $27||h(-d)$, then $\text{Syl}_3(H(-d))$ must be on the form $\mathbb{Z}_{27}$, $\mathbb{Z}_9 \times \mathbb{Z}_3$ or $\mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3$.

Also note that if $\text{Syl}_3(H(-d)) \cong \mathbb{Z}_9$ or $\mathbb{Z}_{27}$, then $H(-d)[3] \cong \mathbb{Z}_3$, if $\text{Syl}_3(H(-d)) \cong \mathbb{Z}_3^2$ or $\mathbb{Z}_9 \times \mathbb{Z}_3$, then $H(-d)[3] \cong \mathbb{Z}_3^2$, and if $\text{Syl}_3(H(-d)) \cong \mathbb{Z}_3^3$, then $H(-d)[3] \cong \mathbb{Z}_3^3$.

But as was seen at the end of section 2.3, the cases when $H(-d)[3] \cong \mathbb{Z}_3^n$, has different contributions to the sum $\sum_{-X<d<0} |H(d)[3]|$ for different n. So it could be worth looking into if $\Pr(\text{Syl}_3(H(-d)) \cong G)$ also has a $d^{-1/6}$-term. These will be our last hypotheses.

Hypotheses 4-8. Let $G = \mathbb{Z}_9, \mathbb{Z}_3^2, \mathbb{Z}_{27}, \mathbb{Z}_9 \times \mathbb{Z}_3$ or $\mathbb{Z}_3^3$. Then for some B_G,

$$f_G(d) = A_G + \frac{B_G}{d^{1/6}}$$

where $A_G = \frac{1}{|\mathrm{Aut}(G)|} \prod_{k=1}^{\infty} \left(1 - \frac{1}{3^k}\right)$, is a good model of weather the 3-Sylow subgroup of $H(-d)$ is congruent to G or not. Using theorem 10, we have

$$\begin{aligned} A_{\mathbb{Z}_9} &\approx 0.093354 \\ A_{\mathbb{Z}_3^2} &\approx 0.011669 \\ A_{\mathbb{Z}_{27}} &\approx 0.031118 \\ A_{\mathbb{Z}_9 \times \mathbb{Z}_3} &\approx 0.005186 \\ A_{\mathbb{Z}_3^3} &\approx 0.000050. \end{aligned}$$

We will also compare against the constant Cohen-Lenstra model $f_G^{CL}(d) = A_G$.

Chapter 4

Results

4.1 3-divisibility

As a first experiment we have chosen a partition of $N = 10^7$ points in each set G_i, which gives us $N_G = 74$ data points of the form (d_i', P_i). In figure 4.1, we see these data points plotted against the regression $A_3 + \frac{B_3}{d^{1/6}}$, where the least square approximation of the coefficient B_3 was found to be -0.2198. For reference, the Cohen-Lenstra prediction of $0.43987\ldots$ is also shown. Visually, it seems to be a good fit.

Figure 4.2 shows the density histogram for $(Y_i)_{1\leq i\leq 74}$. We have also calculated the mean and standard deviation of these values, μ_Y and σ_Y, and overlaid a normal distribution with these parameters. In this case, our calculated μ_Y and σ_Y are quite close to 0 and 1, but the data points are to few to see if they are normally distributed.

By choosing a smaller value of N, we get more data points (d_i', P_i). Figure 4.3 show the same plot with $N = 10^4$. This time, we find B_3 to be -0.2188, which is very close to the $N = 10^7$ case. Since the sets G_i have fewer data points, the variation in P_i is bigger. But as figure 4.4 shows, the data points Y_i are very close to being normally distributed. In appendix 6.2.1, histograms for other values of N can be found.

Table 4.1, shows the estimate of B_3 for various values of N. We can see that the estimate does not vary significantly around $B_3 \approx -0.219$.

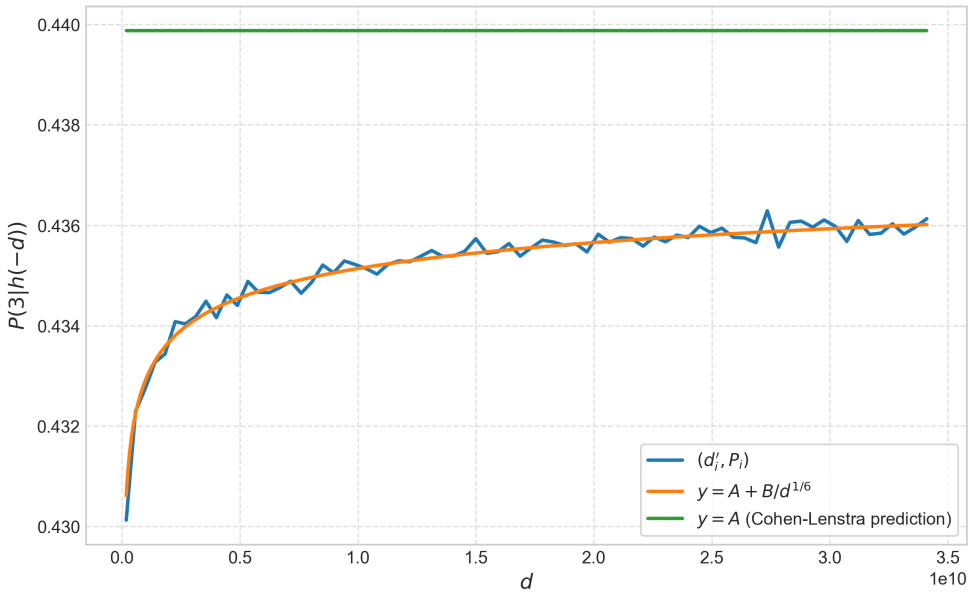

Figure 4.1: $N = 10^7, B \approx -0.2198$

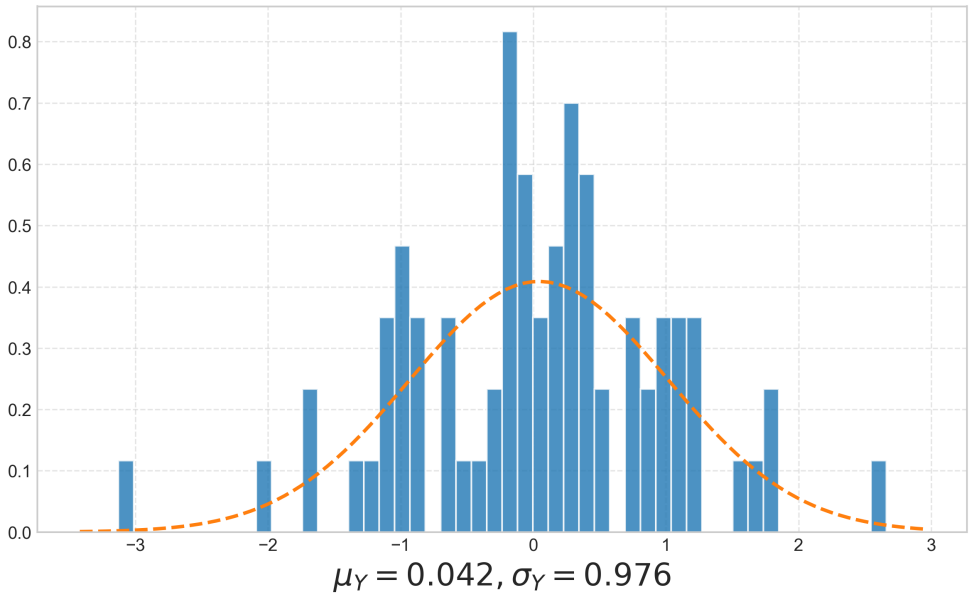

Figure 4.2: $N = 10^7, B \approx -0.2198$

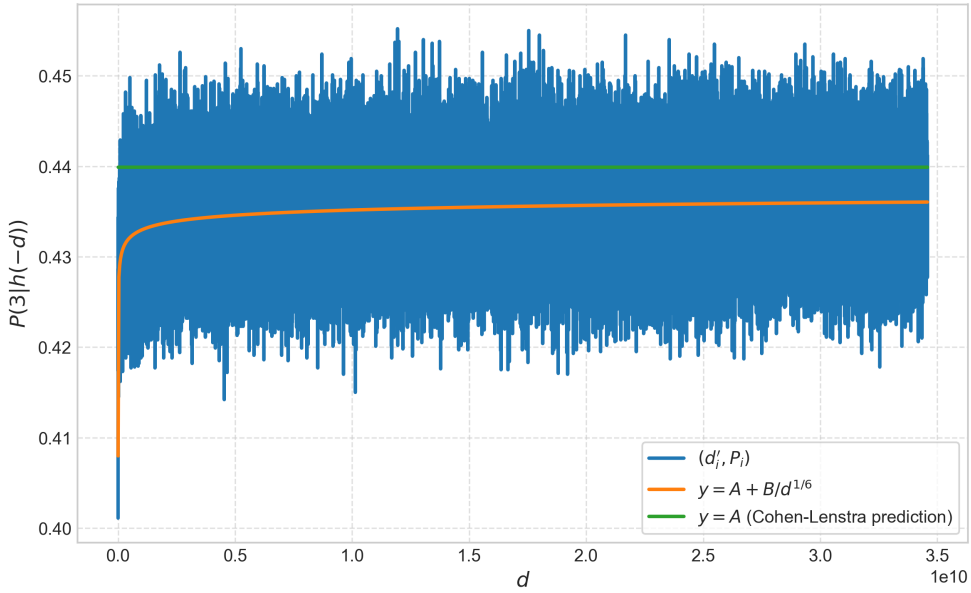

Figure 4.3: $N = 10^4, B \approx -0.2189$

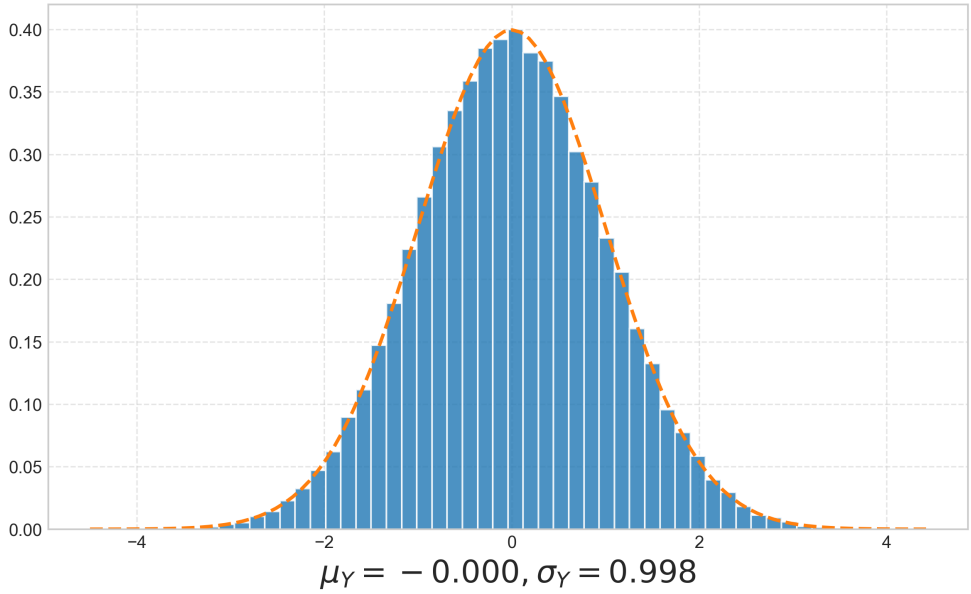

Figure 4.4: $N = 10^4, B \approx -0.2189$

N	N_G	B_3
10^7	74	-0.21980
10^6	744	-0.21908
10^5	7 445	-0.21891
10^4	74 458	-0.21888
10^3	744 582	-0.21888

Table 4.1: Estimate of B_3 for various values of N.

4.2 5- and 7-divisibility

We repeat the calculations for 5-divisibility and 7-divisibility.

Starting with 5-divisibility, figure 4.5 shows the plot for $N = 10^7$, which results in the estimate $B_5 = -0.3026$ in the equation $f_5(d) = A_5 + \frac{B_5}{d^{3/10}}$.

As seen in figure 4.6, with $N = 10^4$, (Y_i) seems to be normally distributed. In appendix 6.2.2, histograms for other values of N can be found.

With 7-divisibility, figure 4.7 shows the plot for $N = 10^7$, which results in the estimate $B_7 = -0.2684$ in the equation $f_7(d) = A_7 + \frac{B_7}{d^{5/14}}$.

In figure 4.8, with $N = 10^4$, we can see that (Y_i) seem to be normally distributed in this case as well. In appendix 6.2.3, histograms for other values of N can be found.

Table 4.2 shows the estimates of B_5 and B_7 for various values of N.

N	B_5	B_7
10^7	-0.30261	-0.26844
10^6	-0.29796	-0.25398
10^5	-0.29437	-0.25039
10^4	-0.29357	-0.24359

Table 4.2: Estimates of B_5 and B_7 for different N.

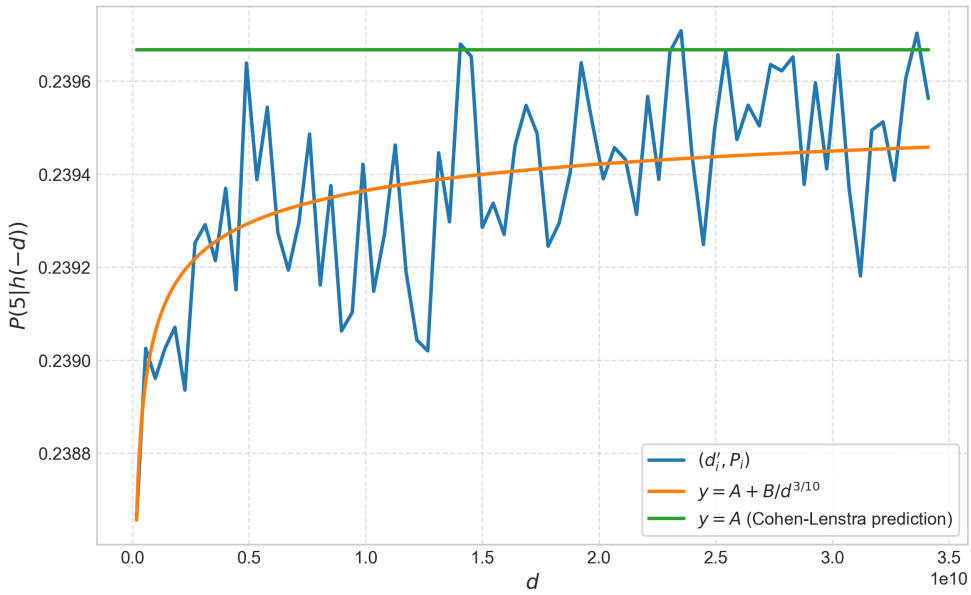

Figure 4.5: $N = 10^7$, $B_5 \approx -0.3026$

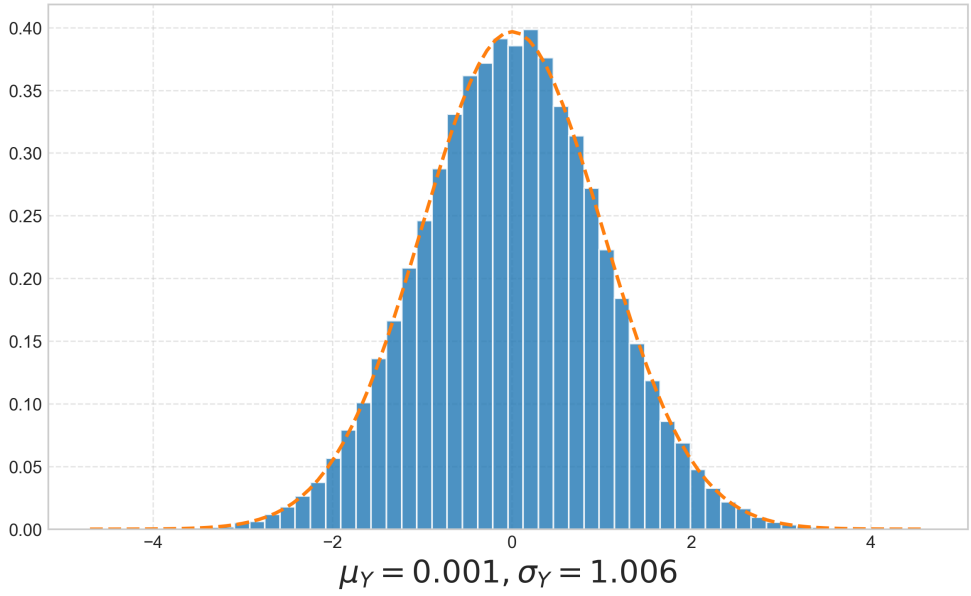

Figure 4.6: $N = 10^4$, $B_5 \approx -0.2935$

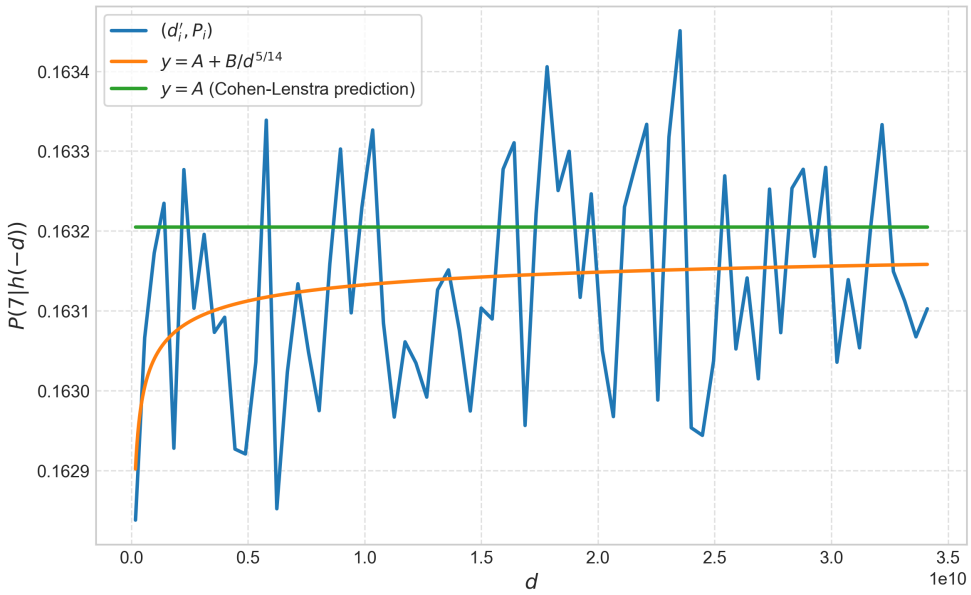

Figure 4.7: $N = 10^7, B_7 \approx -0.2684$

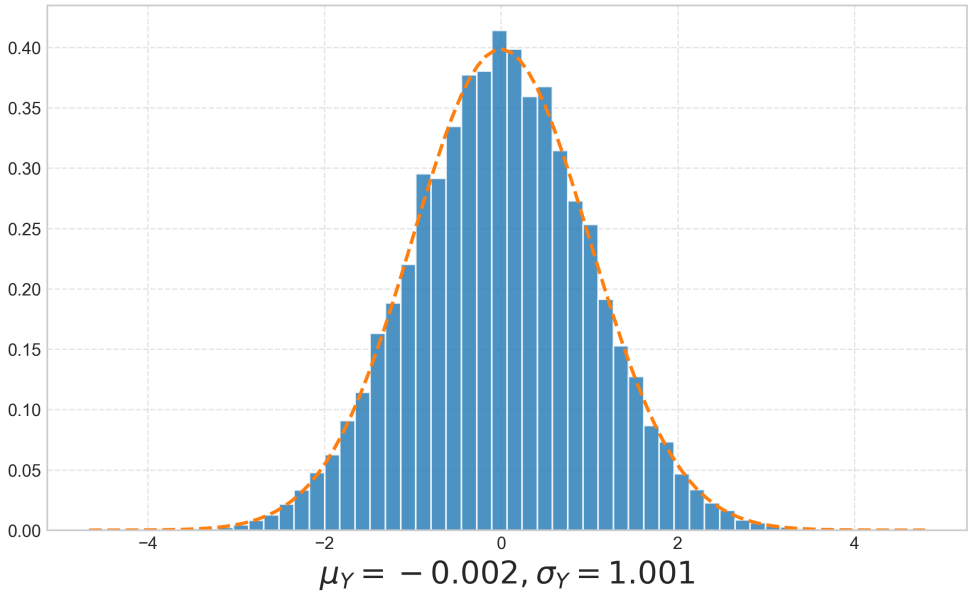

Figure 4.8: $N = 10^4, B_7 \approx -0.2436$

4.3 3-Sylow: $\mathbb{Z}_9$ and $\mathbb{Z}_3 \times \mathbb{Z}_3$

We now turn to the 3-Sylow subgroup. We first focus on the case where $9||h(-d)$, i.e. where the possible 3-Sylow subgroups are $\mathbb{Z}_9$ and $\mathbb{Z}_3 \times \mathbb{Z}_3$.

Starting with $\text{Syl}_3(H(-d)) \cong \mathbb{Z}_9$, figure 4.9 shows the plot for $N = 10^7$, which results in the estimate $B_{\mathbb{Z}_9} = -0.0353$ in the equation $f_{\mathbb{Z}_9}(d) = A_{\mathbb{Z}_9} + \frac{B_{\mathbb{Z}_9}}{d^{1/6}}$.

In figure 4.10, with $N = 10^4$, we can see that (Y_i) seem to be normally distributed. In appendix 6.2.4, histograms for other values of N can be found.

With $\text{Syl}_3(H(-d)) \cong \mathbb{Z}_3 \times \mathbb{Z}_3$, figure 4.11 shows the plot for $N = 10^7$, which results in the estimate $B_{\mathbb{Z}_3^2} = -0.0361$ in the equation $f_{\mathbb{Z}_3^2}(d) = A_{\mathbb{Z}_3^2} + \frac{B_{\mathbb{Z}_3^2}}{d^{1/6}}$.

In figure 4.12, with $N = 10^4$, we can see that (Y_i) looks to be normally distributed in this case too. In appendix 6.2.5, histograms for other values of N can be found.

Table 4.3, shows the estimates of $B_{\mathbb{Z}_9}$ and $B_{\mathbb{Z}_3^2}$ for various values of N.

N	$B_{\mathbb{Z}_9}$	$B_{\mathbb{Z}_3^2}$
10^7	-0.03532	-0.03608
10^6	-0.03525	-0.03598
10^5	-0.03524	-0.03595
10^4	-0.03523	-0.03594

Table 4.3: Estimate of $B_{\mathbb{Z}_9}$ and $B_{\mathbb{Z}_3^2}$ for different N.

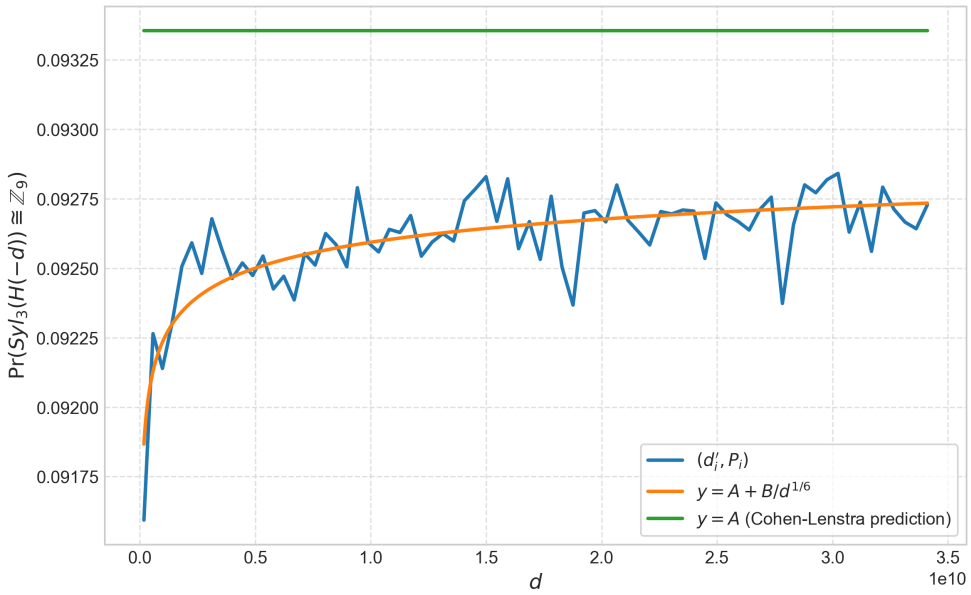

Figure 4.9: $N = 10^7$, $B_{\mathbb{Z}_9} \approx -0.0353$

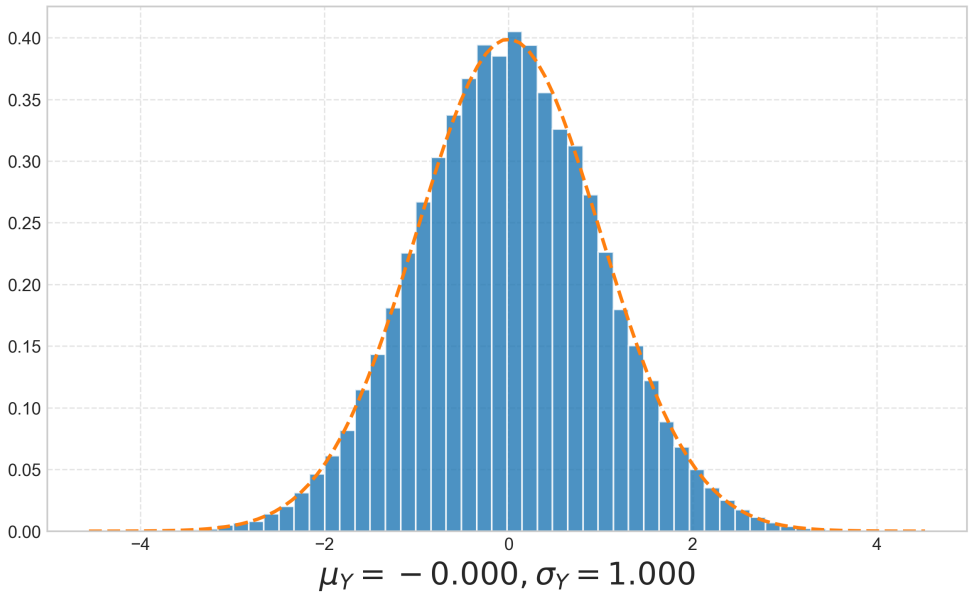

Figure 4.10: $N = 10^4$, $B_{\mathbb{Z}_9} \approx -0.0352$

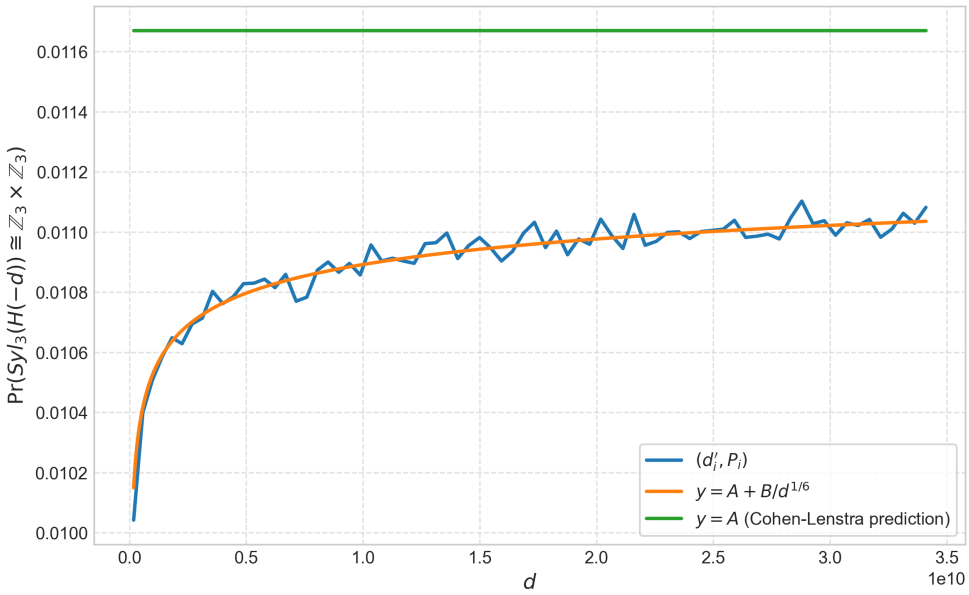

Figure 4.11: $N = 10^7$, $B_{\mathbb{Z}_3^2} \approx -0.0361$

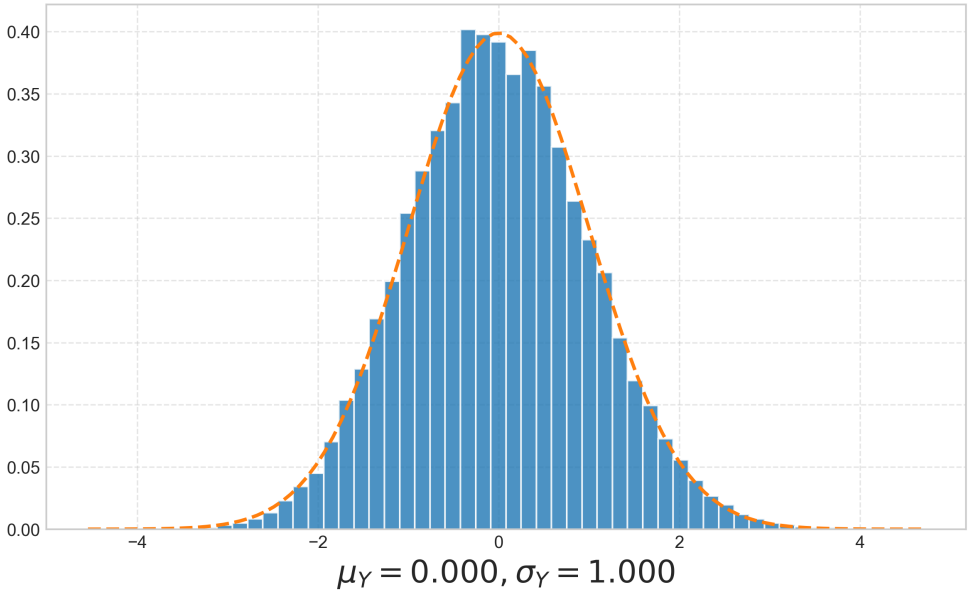

Figure 4.12: $N = 10^4$, $B_{\mathbb{Z}_3^2} \approx -0.0359$

4.4 3-Sylow: $\mathbb{Z}_{27}$, $\mathbb{Z}_9 \times \mathbb{Z}_3$ and $\mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3$

Finally, we turn to the case $27||h(-d)$, where the 3-Sylow subgroup can take the form $G = \mathbb{Z}_{27}$, $\mathbb{Z}_9 \times \mathbb{Z}_3$ or $\mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3$. The calculated value of B_G for various N is presented in table 4.4, the graphs using $N = 10^7$ is shown in figures 4.13-4.15, and the histograms of (Y_i) can be found in appendices 6.2.6-6.2.8.

N	$B_{\mathbb{Z}_{27}}$	$B_{\mathbb{Z}_9 \times \mathbb{Z}_3}$	$B_{\mathbb{Z}_3^3}$
10^7	-0.01201	-0.01612	-0.000628
10^6	-0.01197	-0.01607	-0.000625
10^5	-0.01195	-0.01606	-0.000624
10^4	-0.01196	-0.01606	-0.000624

Table 4.4: Estimate of $B_{\mathbb{Z}_9}$ and $B_{\mathbb{Z}_3^2}$ for different N.

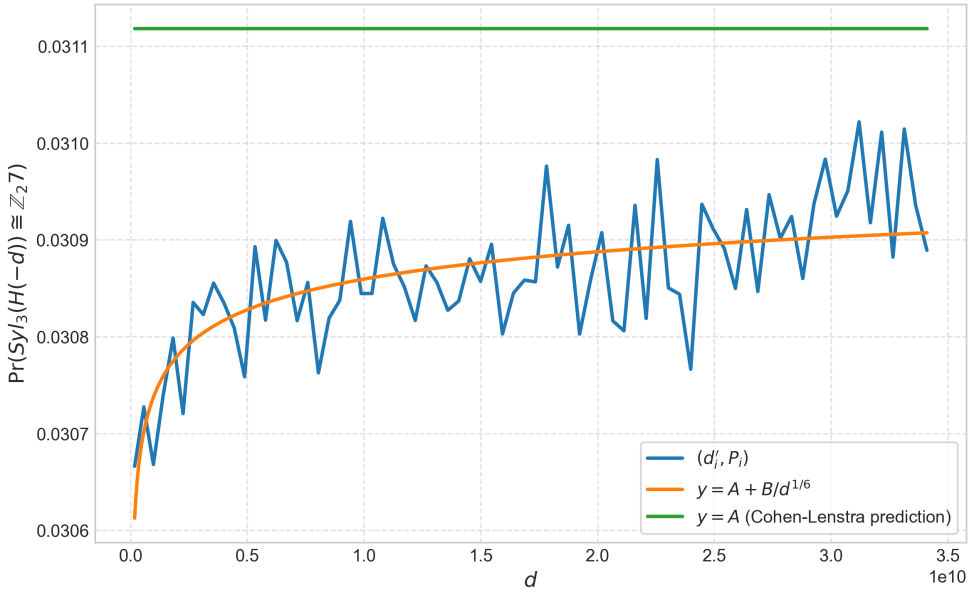

Figure 4.13: $N = 10^7$, $B_{\mathbb{Z}_{27}} \approx -0.0361$

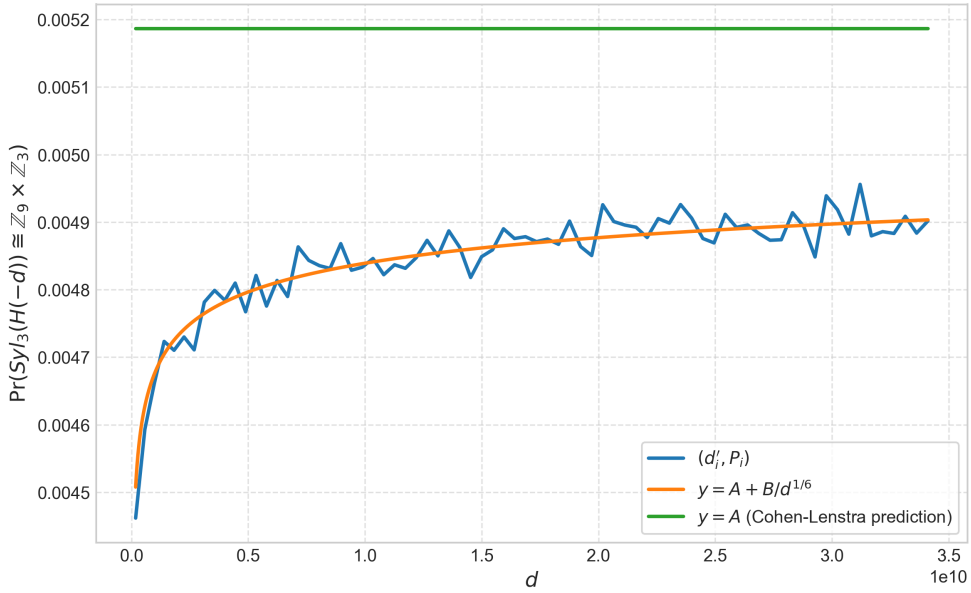

Figure 4.14: $N = 10^7, B_{\mathbb{Z}_9 \times \mathbb{Z}_3} \approx -0.0359$

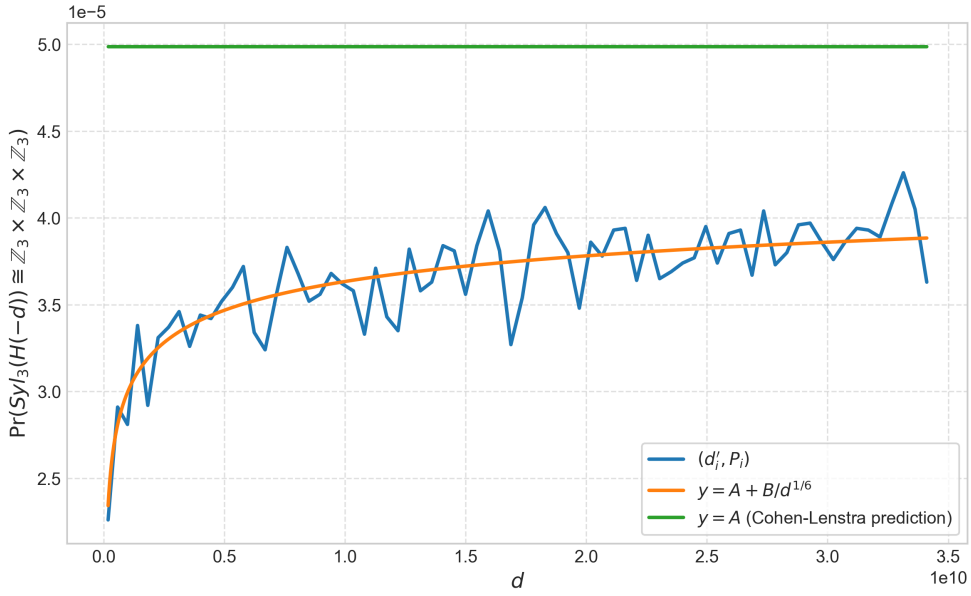

Figure 4.15: $N = 10^7, B_{\mathbb{Z}_3^3} \approx -0.0359$

Chapter 5

Discussion

In this section we discuss the results obtained in chapter 4. First we present a table of the values of $a.$, $A.$ and $B.$ for our models

$$f.(d) = A. + \frac{B.}{d^{a.}}.$$

For $B.$, we have written down the number of decimals that seem to be unchanging when varying N between different powers of 10, as seen in tables 4.1-4.4.

$\cdot$	$a.$	$A.$	$B.$
3	$1/6$	0.43987	-0.229
5	$3/10$	0.23967	-0.29
7	$5/14$	0.16320	-0.24
$\mathbb{Z}_9$	$1/6$	0.09335	-0.0352
$\mathbb{Z}_3^2$	$1/6$	0.01167	-0.0359
$\mathbb{Z}_{27}$	$1/6$	0.03112	-0.0120
$\mathbb{Z}_9 \times \mathbb{Z}_3$	$1/6$	0.00519	-0.0161
$\mathbb{Z}_3^3$	$1/6$	0.00005	-0.00062

Table 5.1: Summary of our results

5.1 3-divisibility

Looking at figure 4.2, the model

$$f_3(d) = A_3 + \frac{B_3}{d^{1/6}},$$

with $A_3 = 1 - \prod_{k=1}^{\infty}\left(1 - \frac{1}{3^k}\right) \approx 0.43987$ and $B_3 = -0.219$, seem to visually fit the data very well. In figure 4.4, the data fluctuate far wider, but this is expected, since N is smaller and $\text{Var}[\overline{\mathbb{X}}_i] \propto 1/N$.

Looking at histograms 4.2 and 4.4, as well as those in appendix 6.2.1, it seems like the data points (Y_i) follow a standard normal distribution, since the fit gets better with more data points.

This means the errors from this model are normally distributed around 0, and with the same standard deviation as we calculated theoretically in section 3.1.

On the other side, if we instead compute the errors to the constant Cohen-Lenstra prediction $f_3^{CL}(d) = A_3$, our errors are no longer centered around 0, and are not normally distributed, as seen in the plots to the right in appendix 6.2.1.

As N gets smaller (and thus N_G gets bigger), the constant Cohen-Lenstra histograms does however seem to get closer to a normal distribution. The reason for this is that when N gets smaller, the standard deviation of $\overline{\mathbb{X}}_i$ gets bigger, which means that after normalizing, the difference between A_3 and $A_3 + B_3/d^{1/6}$ gets smaller. But note that the plots on the left in appendix 6.2.1 all have $\mu_Y \approx 0$ and $\sigma_Y \approx 1$ for all N.

This is to say that the constant Cohen-Lenstra prediction $f_3^{CL}(d) = A_3$ is a bad fit for our data, but adding the term $B_3/d^{1/6}$ does seem to fit the data well.

5.2 5- and 7 divisibility

We now turn to 5-divisibility and figure 4.7. Visually, it is not as clear that we have found the right fit. Comparing it to the 3-divisibility plot in figure 4.1, it might look like the variations are larger in the 5-divisibility case. However, if

one takes a closer look at the y-axis, the variations are actually smaller, as we would expect from the form of $\mathrm{Var}[\overline{\mathbb{X}}_i]$. The difference is that for 5-divisibility our model is closer to the Cohen-Lenstra prediction, since $1/d^{3/10}$ diminishes faster than $1/d^{1/6}$.

Looking at the histograms in appendix 6.2.2, the normalized errors does seem to follow a standard normal distribution, as opposed to the errors to the constant Cohen-Lenstra model. The deviations are smaller that in the 3-divisibility case, for the reasons given above, but is nevertheless clear that the model with the higher-degree term follows the data better, compared to the constant Cohen-Lenstra model.

Another difference from the 3-divisibility histograms is that it is hard to notice deviations from a normal distribution for the constant Cohen-Lenstra prediction errors. The relatively large shifts in μ_Y do however make it clear that these do not follow a *standard* normal distribution.

Therefore, hypothesis 2 seem to be compatible with our data, and the constant Cohen-Lenstra model does not seem to be.

For 7-divisibility, the observations from 3- and 5-divisibility seem to continue to hold, but here the deviation between the two models are even smaller. This gives a dismissal of the constant Cohen-Lenstra prediction a lower confidence than for the 3- and 5-divisibility cases. However, comparing the histograms on the right and the left side in appendix 6.2.2, the improvement does seem significant enough to warrant skepticism around the constant Cohen-Lenstra model.

5.3 3-Sylow: $\mathbb{Z}_9$ and $\mathbb{Z}_3 \times \mathbb{Z}_3$

The results for the 3-Sylow subgroups are similar to the ones for 3-divisibility. The differences in the histograms on the right and left side in appendix 6.2.1 are larger than the ones for 5- and 7-divisibility, which could be explained by these terms being proportional to $d^{-1/6}$, which diminishes slower than $d^{-3/10}$ and $d^{-5/14}$.

It is interesting that the coefficients $B_{\mathbb{Z}_9}$ and $B_{\mathbb{Z}_3^2}$ are so close to each other. This means that the higher order term in $\Pr(\mathrm{Syl}_3(H(-d)) \cong \mathbb{Z}_9)$ and $\Pr(\mathrm{Syl}_3(H(-d)) \cong$

$\mathbb{Z}_3 \times \mathbb{Z}_3$) is almost identical, even though the former is around 8 times larger then the latter.

It is worth noting that since the probability $\Pr(\text{Syl}_3(H(-d)) \cong \mathbb{Z}_3 \times \mathbb{Z}_3)$ is much lower than the other ones we have looked at, around 1.1%. This means that if $N = 10^4$, there will only be around 110 positive results in each group, which will mean that our data becomes discretized in a more obvious way. This is the reason the constant Cohen-Lenstra prediction histogram looks a bit weird for $N = 10^4$; the height of the histogram bars also depends on how many of these discrete numbers lie in their bin. The same effect can be observed in some of the histograms of the 3-sylow groups of order 27.

5.4 3-Sylow: $\mathbb{Z}_{27}$, $\mathbb{Z}_9 \times \mathbb{Z}_3$ and $\mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3$

From figures 4.13-4.15 and the histograms in appendices 6.2.6-6.2.8, it is clear that the behavior observed in section 5.3 continue to hold for the 3-Sylow subgroups of order 27.

However, this time $B_{\mathbb{Z}_{27}}$, $B_{\mathbb{Z}_9 \times \mathbb{Z}_3}$ and $B_{\mathbb{Z}_3^3}$ are not close to each other. This might point at the observation made in section 6.2.4 that $B_{\mathbb{Z}_9}$ and $B_{\mathbb{Z}_3^2}$ are approximately equal, might have been a coincidence.

It is however interesting that $B_{\mathbb{Z}_9 \times \mathbb{Z}_3}$ is around 30% larger than $B_{\mathbb{Z}_{27}}$, even though the Cohen-Lensta prediction $A_{\mathbb{Z}_9 \times \mathbb{Z}_3}$ is just a sixth of $A_{\mathbb{Z}_{27}}$. It seems to be hard to predict the size of the higher degree term $B_G/d^{1/6}$ solely from the size of $|\text{Aut}(G)|$.

Chapter 6

Appendices

6.1 The sum $\sum'_{3|h(-d)} 1$

Here we calculate the sum

$$\sum_{\substack{-X<-d<0 \\ 3|h(-d)}}{}' 1 = A_3 \left(\sum_{-X<-d<0}{}' 1 \right) + B_3 \left(\sum_{-X<-d<0}{}' \frac{1}{d^{1/6}} \right)$$

under hypothesis 1.

The first sum on the right hand side can, according to the Dirichlet's theorem on arithmetic progressions, be approximated as:

$$\sum_{-X<-d<0}{}' 1 = \sum_{\substack{0<d<X \\ d \text{ prime} \\ d \equiv 3 \bmod 4}} 1 \sim \frac{1}{2} \sum_{\substack{0<d<X \\ d \text{ prime}}} 1 = \frac{\pi(X)}{2} \sim \frac{X}{2 \log X}$$

We now move to the second sum. First, according to Dirichlet's theorem on arithmetic progressions,

$$\sum_{-X \leq -d<0}{}' \frac{1}{d^{1/6}} = \sum_{\substack{0<d \leq X \\ d \text{ prime} \\ d \equiv 3 \bmod 4}} \frac{1}{d^{1/6}} \sim \frac{1}{2} \sum_{\substack{0<d \leq X \\ d \text{ prime}}} \frac{1}{d^{1/6}}.$$

Now we use Abel's summation formula, stated below.

Theorem 11. (Abel's summation formula) If $(a_n)_{n\in\mathbb{N}^+}$ is a sequence and $f(x)$ is a continuously differentiable function on $[1,\infty)$, then:

$$\sum_{n\leq x} a_n f(n) = A(x)f(x) - \int_1^x A(t)f'(t)\,dt$$

where $A(x) = \sum_{n\leq x} a_n$.

If we let $f(x) = 1/x^{1/6}$ and

$$a_n = \begin{cases} 1 & \text{if } n \text{ is prime} \\ 0 & \text{else} \end{cases},$$

we get $A(x) = \sum_{n\leq x} a_n = \pi(x)$. Applying Abel's summation formula, we get:

$$\sum_{\substack{0<d\leq X \\ d \text{ prime}}} \frac{1}{d^{1/6}} = \sum_{n\leq X} \frac{a_n}{n^{1/6}} = \frac{\pi(X)}{X^{1/6}} + \frac{1}{6}\int_2^X \frac{\pi(t)}{t^{5/6}}\,dt.$$

To calculate the integral, we use the prime number theorem,

$$\begin{aligned}\int_2^X \frac{\pi(t)}{t^{5/6}}\,dt &\sim \int_2^X \frac{1}{t^{1/6}\log t}\,dt \quad \left\{\begin{array}{l} t = u^{6/5} \\ dt = \frac{6}{5}u^{1/5}\,du \end{array}\right\} \\ &= \int_{2^{5/6}}^{X^{5/6}} \frac{1}{\log u}\,du \sim \mathrm{Li}(X^{5/6}) \sim \frac{X^{5/6}}{\log X^{5/6}} = \frac{6}{5}\frac{X^{5/6}}{\log X}.\end{aligned}$$

Thus we have:

$$\sum_{\substack{0<d\leq X \\ d \text{ prime}}} \frac{1}{d^{1/6}} \sim \frac{1}{X^{1/6}}\frac{X}{\log X} + \frac{1}{6}\left(\frac{6}{5}\frac{X^{5/6}}{\log X}\right) = \frac{6}{5}\frac{X^{5/6}}{\log X},$$

and subsequently

$$\sum_{-X\leq -d<0}{}' \frac{1}{d^{1/6}} \sim \frac{3}{5}\frac{X^{5/6}}{\log X}.$$

Thus, under hypothesis 1 we have

$$\sum_{\substack{-X<-d<0 \\ 3|h(-d)}}{}' 1 \sim \frac{A_3}{2}\frac{X}{\log X} + \frac{3B_3}{5}\frac{X^{5/6}}{\log X}.$$

Note that the exponent $5/6$ arose from adding one to the exponent $-1/6$ in the sum we started with.

6.2 Histograms

In this section, the complete collection of the histograms of (Y_i) are presented. The definition of Y_i is

$$Y_i := \frac{P_i - f(d_i')}{\sqrt{\frac{1}{N} f(d_i')(1 - f(d_i'))}},$$

and for a good model $f(d)$, Y_i should be normally distributed with a mean of 0 and standard deviation of 1.

For each conjecture, we present two histograms for $N = 10^7, 10^6, 10^5$ and 10^4.[1] For each N, the left histogram use the Cohen-Lenstra heuristc $f_\cdot^{CL}(d) = A_\cdot$, and for the right histogram we include our higher order term, $f_\cdot(d) = A_\cdot + \frac{B_\cdot}{d^a}$, where $B_\cdot$ is our least square estimate, documented in tables 4.1-4.4.

Below each figure, the mean and standard deviation of (Y_i) (μ_Y and σ_Y) is presented, and a normal distribution with these parameters, $\mathcal{N}(\mu_Y, \sigma_Y)$, is drawn in orange.

[1] Since class groups with $\mathrm{Syl}_3(H(-d)) \cong \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3$ are so uncommon, we only present histograms for $N = 10^7$ and $N = 10^6$ in section 6.2.8.

6.2.1 3-divisibility

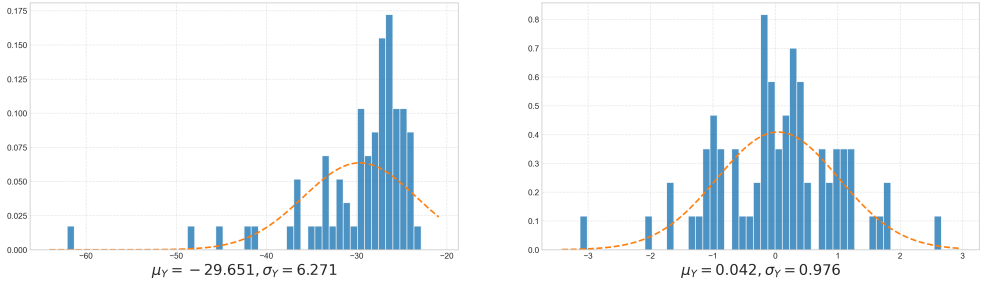

Figure 6.1: $N = 10^7$

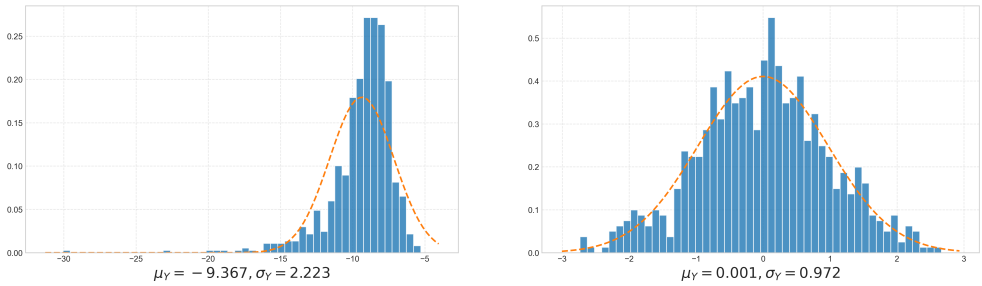

Figure 6.2: $N = 10^6$

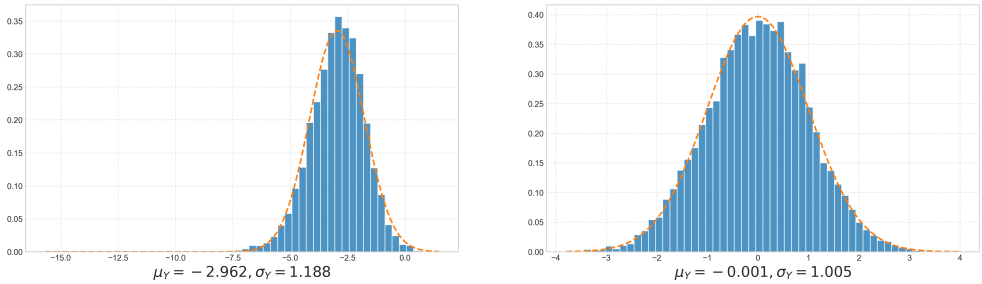

Figure 6.3: $N = 10^5$

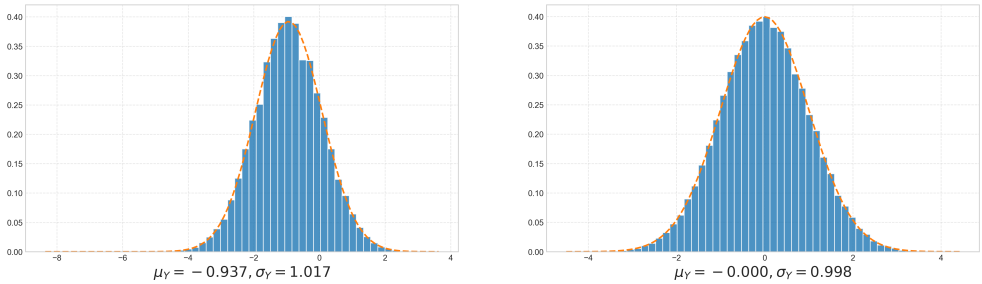

Figure 6.4: $N = 10^4$

6.2.2 5-divisibility

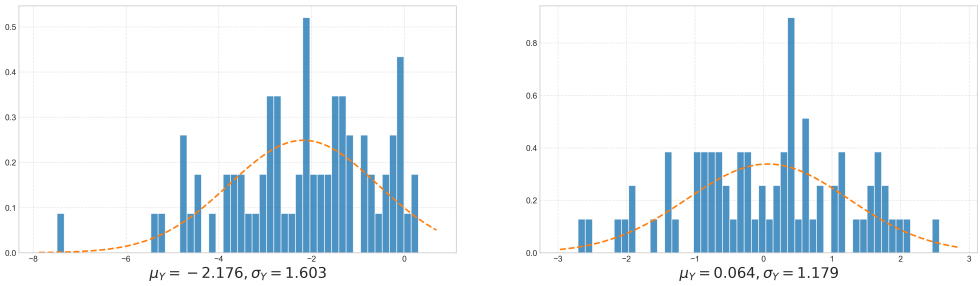

Figure 6.5: $N = 10^7$

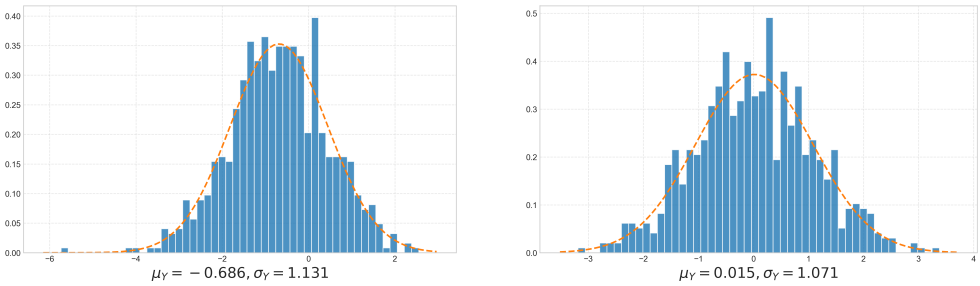

Figure 6.6: $N = 10^6$

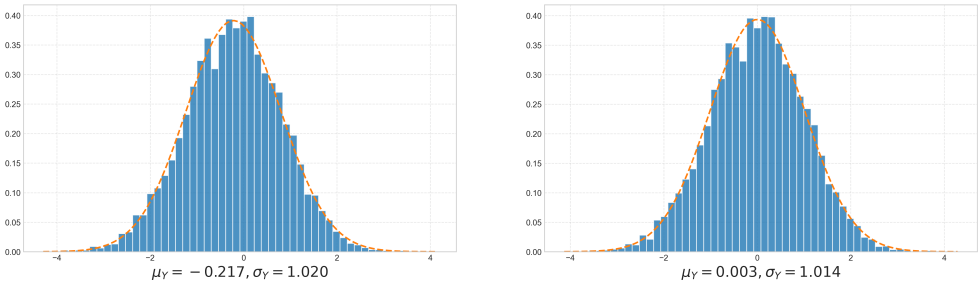

Figure 6.7: $N = 10^5$

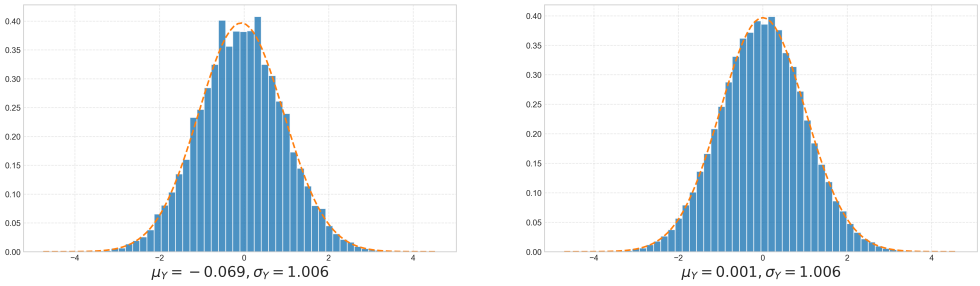

Figure 6.8: $N = 10^4$

6.2.3 7-divisibility

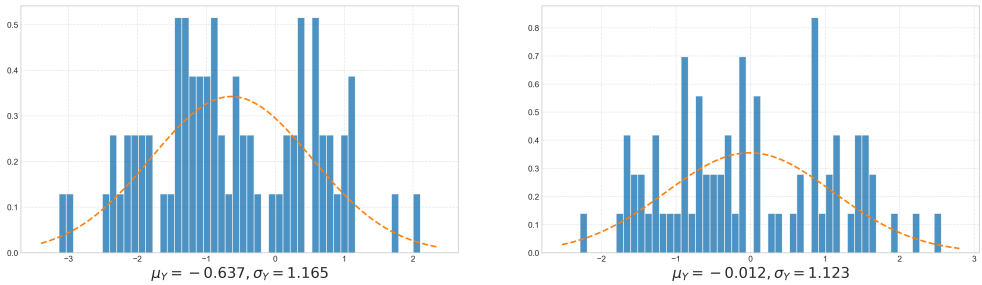

Figure 6.9: $N = 10^7$

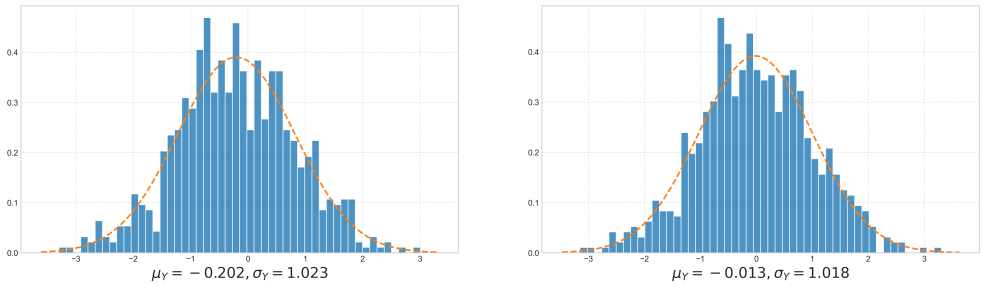

Figure 6.10: $N = 10^6$

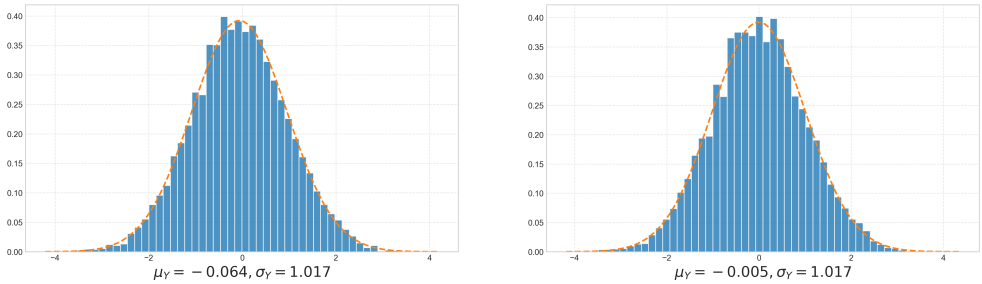

Figure 6.11: $N = 10^5$

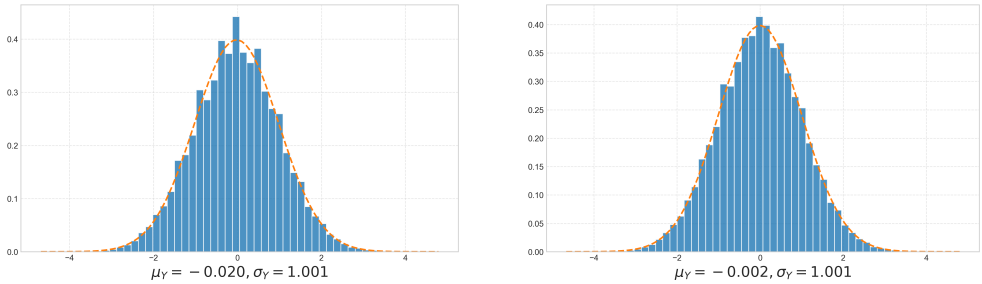

Figure 6.12: $N = 10^4$

6.2.4 3-Sylow subgroup $\mathbb{Z}_9$

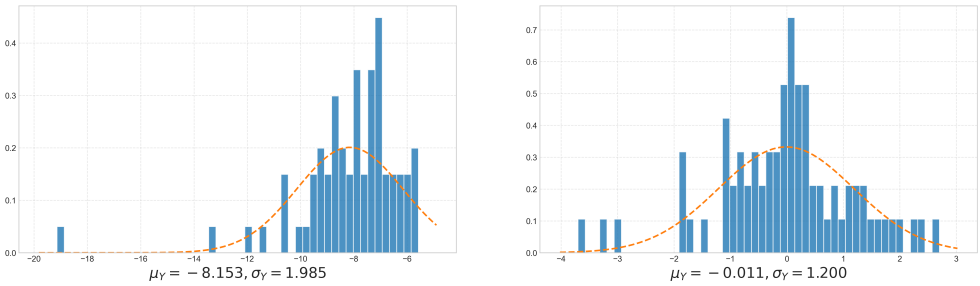

Figure 6.13: $N = 10^7$

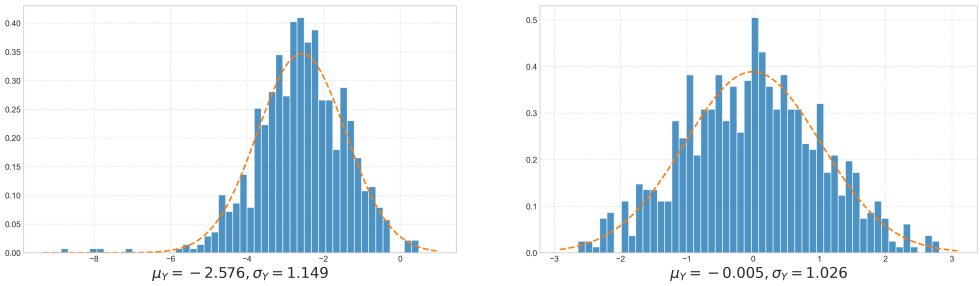

Figure 6.14: $N = 10^6$

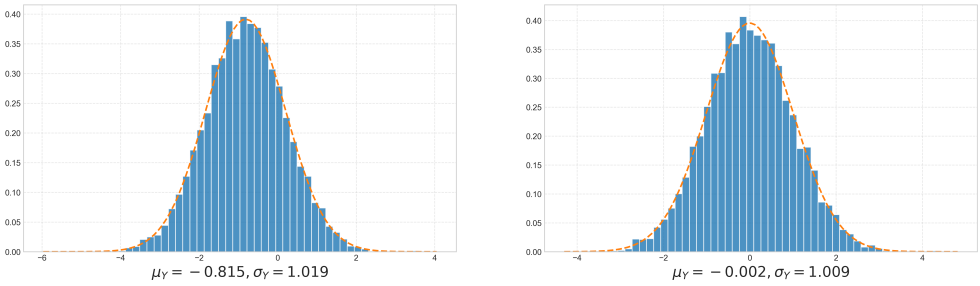

Figure 6.15: $N = 10^5$

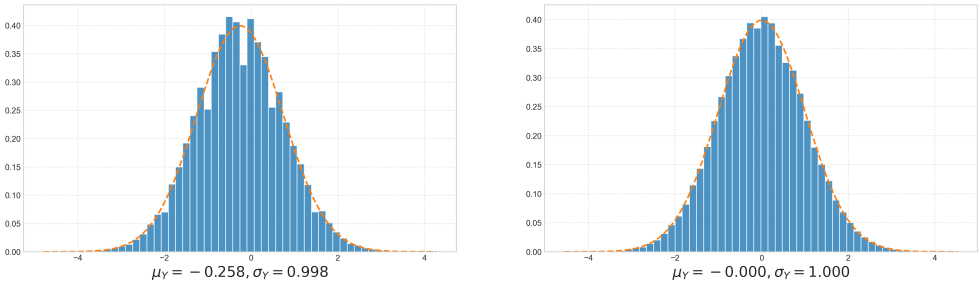

Figure 6.16: $N = 10^4$

6.2.5 3-Sylow subgroup $\mathbb{Z}_3 \times \mathbb{Z}_3$

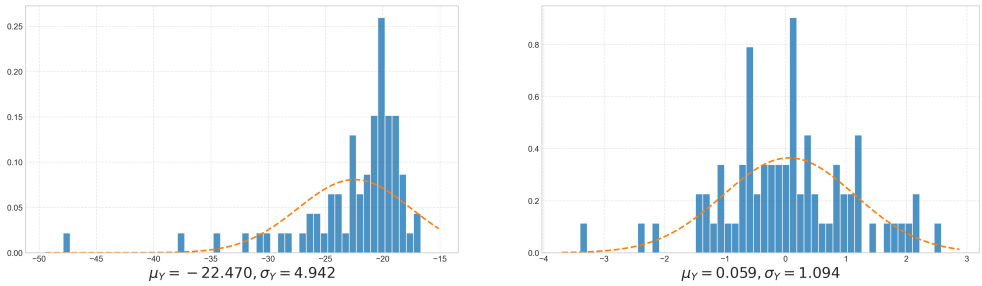

Figure 6.17: $N = 10^7$

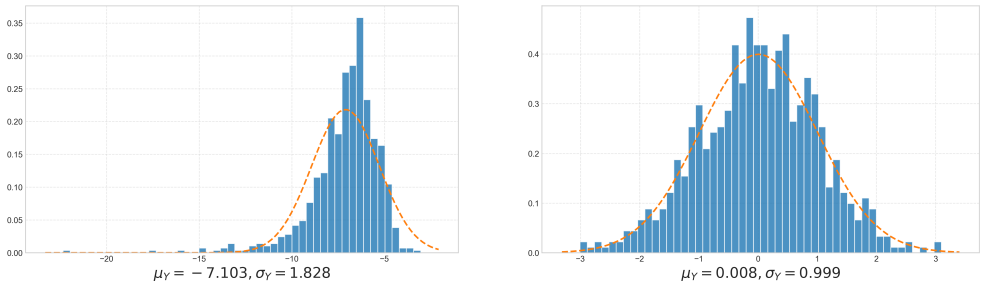

Figure 6.18: $N = 10^6$

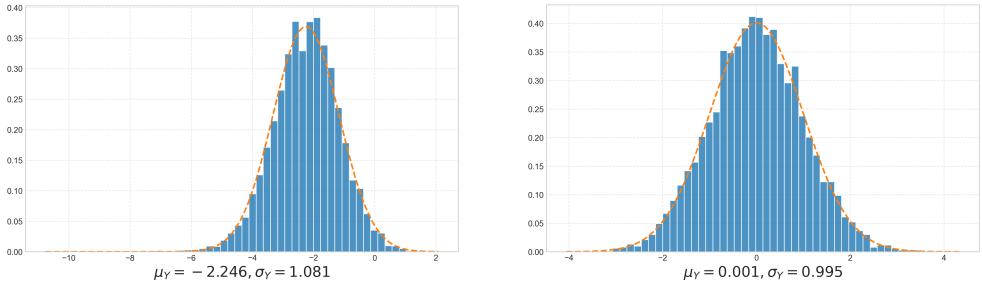

Figure 6.19: $N = 10^5$

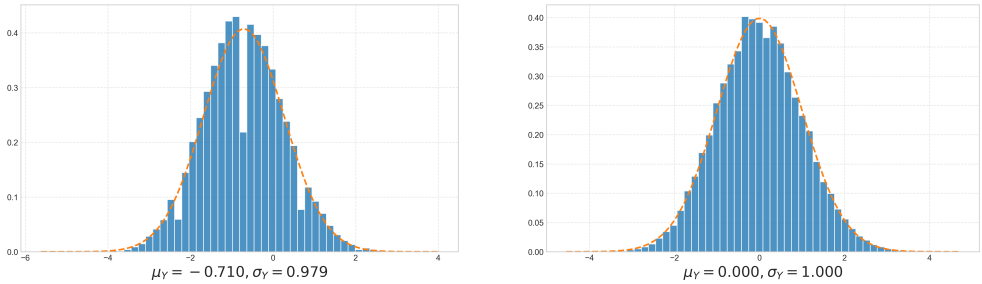

Figure 6.20: $N = 10^4$

6.2.6 3-Sylow subgroup $\mathbb{Z}_{27}$

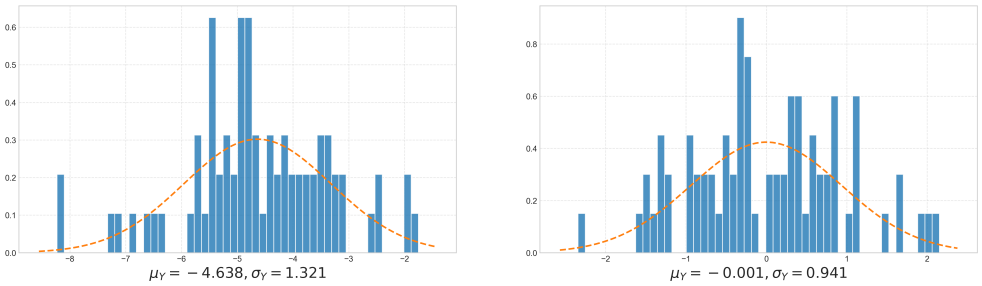

Figure 6.21: $N = 10^7$

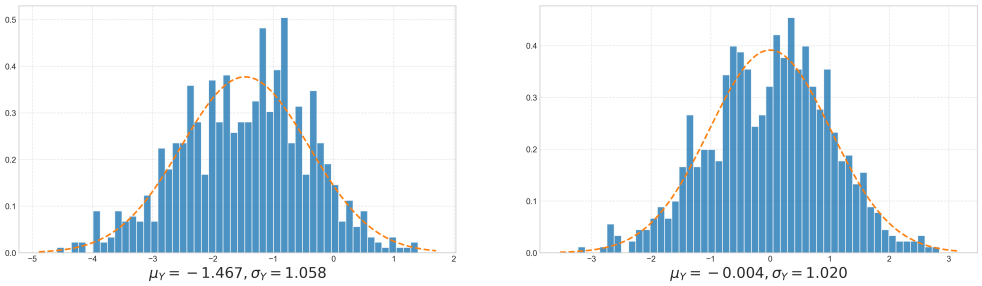

Figure 6.22: $N = 10^6$

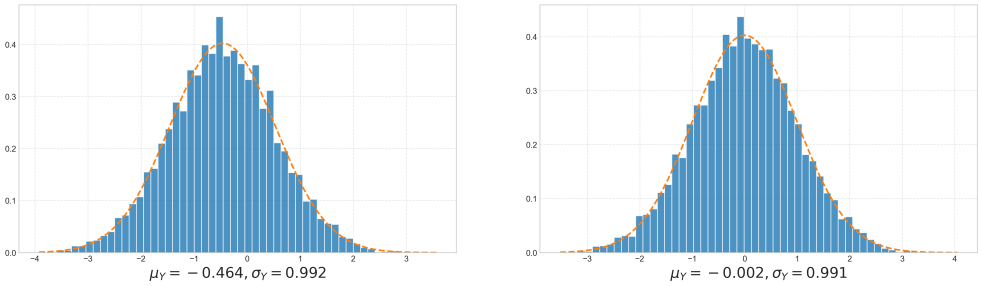

Figure 6.23: $N = 10^5$

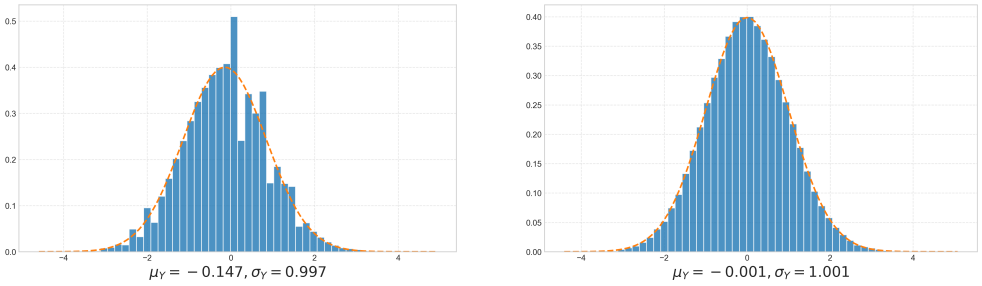

Figure 6.24: $N = 10^4$

6.2.7 3-Sylow subgroup $\mathbb{Z}_9 \times \mathbb{Z}_3$

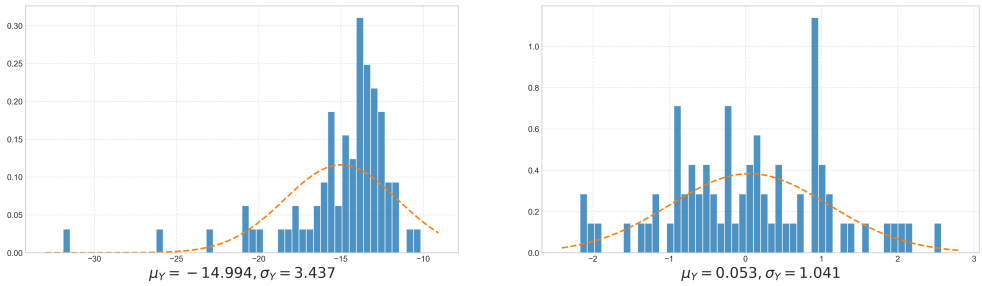

Figure 6.25: $N = 10^7$

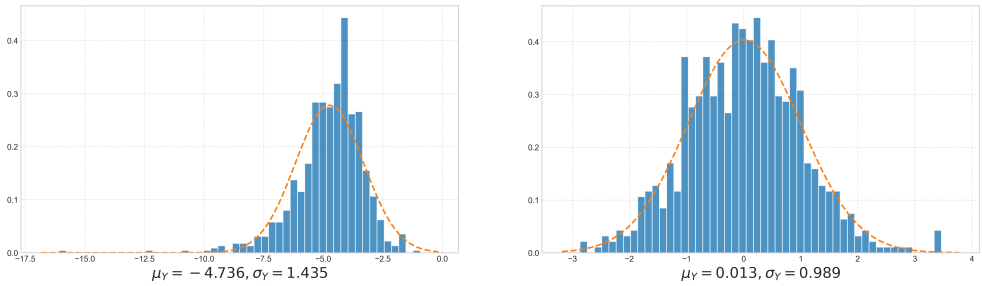

Figure 6.26: $N = 10^6$

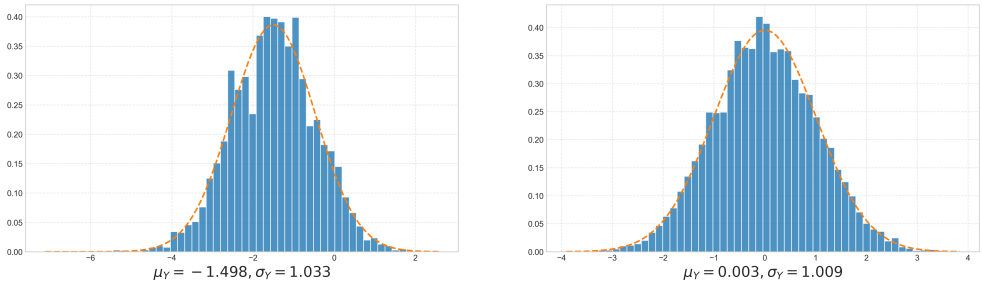

Figure 6.27: $N = 10^5$

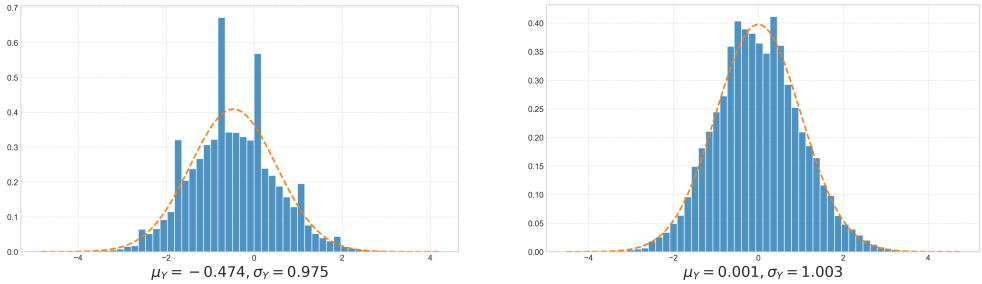

Figure 6.28: $N = 10^4$

6.2.8 3-Sylow subgroup $\mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3$

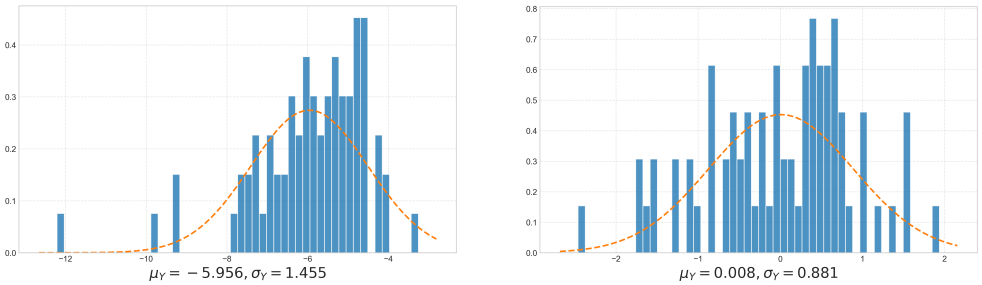

Figure 6.29: $N = 10^7$

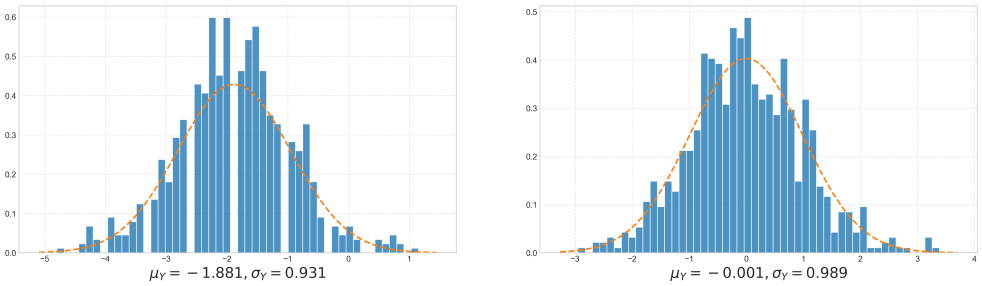

Figure 6.30: $N = 10^6$

Bibliography

[1] David A. Cox. *Primes of the Form $x^2 + ny^2$: Fermat, Class Field Theory, and Complex Multiplication*. Pure and Applied Mathematics. New York: John Wiley & Sons, 1989. ISBN: 978-0-471-50654-2.

[2] H. Cohen and H. W. Lenstra. "Heuristics on class groups of number fields". In: *Number Theory Noordwijkerhout 1983*. Ed. by Hendrik Jager. Berlin, Heidelberg: Springer Berlin Heidelberg, 1984, pp. 33–62. ISBN: 978-3-540-38906-4.

[3] H. Davenport and H. Heilbronn. "On the Density of Discriminants of Cubic Fields. II". In: *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* 322.1551 (1971), pp. 405–420. ISSN: 00804630. URL: http://www.jstor.org/stable/77760 (visited on 06/16/2025).

[4] Manjul Bhargava, Arul Shankar, and Jacob Tsimerman. *On the Davenport-Heilbronn theorems and second order terms*. 2012. arXiv: 1005.0672 [math.NT]. URL: https://arxiv.org/abs/1005.0672.

[5] Takashi Taniguchi and Frank Thorne. "Secondary terms in counting functions for cubic fields". In: *Duke Mathematical Journal* 162.13 (Oct. 2013). ISSN: 0012-7094. DOI: 10.1215/00127094-2371752. URL: http://dx.doi.org/10.1215/00127094-2371752.

[6] Christopher J. Hillar and Darren Rhea. *Automorphisms of finite Abelian groups*. 2006. arXiv: math/0605185 [math.GR]. URL: https://arxiv.org/abs/math/0605185.

[7] Bob Hough. "Equidistribution of bounded torsion CM points". In: *Journal d'Analyse Mathématique* 138.2 (July 2019), pp. 765–797. ISSN: 1565-8538. DOI: 10.1007/s11854-019-0044-4. URL: http://dx.doi.org/10.1007/s11854-019-0044-4.