ABSTRACT

Title of Dissertation:	DISTRIBUTION OF PRIME ORDER IDEAL CLASSES OF QUADRATIC CLASS GROUPS
	Melanka Saroad Wedige Doctor of Philosophy, 2024
Dissertation Directed by:	Professor Niranjan Ramachandran Department of Mathematics

The Cohen-Lenstra heuristics predicts that for any odd prime k, the k-part of quadratic class groups occur randomly on the space of k-groups with respect to a natural probability measure. However, apart from the first moments of the 3-torsion part of quadratic class groups, consequences of these heuristics still remain highly conjectural. The quadratic ideal classes have geometric representations on the modular curve as CM-points in the case of negative discriminants and as closed primitive geodesics in the case of positive discriminants. We mainly focus on the asymptotic distribution of these geometrical objects. As motivation, it is seen that in the case of imaginary quadratic fields, knowledge on the (joint) distribution of k-order CM-points leads in-general to the resolution of the Cohen-Lenstra conjectures on moments of the k-part of class groups. As a first step, inspired by the works of Duke, Hough conjectures that the k-order CM-points are equidistributed on the modular curve. Although the case with $k = 3$ was resolved by Hough himself, $k > 3$ remains unresolved.

In this dissertation, we revisit Hough's conjectures, with empirical evidence. We were able to reprove the conjecture for $k = 3$, and even stronger to show that the result holds along certain subfamilies of imaginary quadratic fields defined by local behaviors of their discriminants. In addition, we study the case for $k > 3$. We introduce a heuristics model, and show that this model agrees with Hough's conjectures. We also show that the difference between the actual asymptotics and the heuristic model reduces down to the distribution of solutions

to certain quadratic congruences. We, then again inspired by Duke's work, investigate an analog for the real quadratic fields. Backed by empirical evidence, we go on to conjecture the asymptotic behavior of the length of k-order geodesics on the modular curve. In addition, based on a theorem and its proof by Siegel, we prove certain results that may shed light on a probable proof direction of these conjectures.

DISTRIBUTION OF PRIME ORDER IDEAL CLASSES OF QUADRATIC CLASS GROUPS

by

Melanka Saroad Wedige

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2024

Advisory Committee:
Professor Niranjan Ramachandran, Chair/Advisor
Professor Lawrence Washington
Professor Patrick Brosnan
Professor Adam Kanigowski
Professor William Gasarch

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincerest gratitude to my graduate advisor, Professor Niranjan Ramachandran. His guidance, motivation and helpful suggestions, was significant in successfully completing this research. He has been a constant driving force throughout my graduate program.

I would also like to thank Professor Lawrence Washington for important insights, questions and feedback regarding the results of my thesis. Also, it was while reading with him that we discovered this research problem. I also extend my gratitude to the rest of my thesis committee; Professor Patrick Brosnan, Professor Adam Kanigowski and Professor William Gasarach for providing important feedback on my work.

I would like to thank Matthew Welsh for helping me with computing the lengths of closed geodesics and providing insights on the equidistribution of solutions to quadratic congruences, which were all central to my research work presented here.

Next I would like to thank the University of Maryland, in particular the Department of Mathematics. With immense gratitude I recall the support extended to me during the initial graduate application process by Professor Matei Machedon, the then graduate program director. I would also like to extend my gratitude towards the administrative staff, without whom I would have been overburdened with the numerous paper work.

I extend my heartfelt gratitude to my friends from both the mathematics graduate community and the Sri Lankan community at College Park. Their warmth and welcomingness made my graduate life pleasant and memorable.

Finally, I would extend my utmost gratitude to my parents and sister. Despite living far away from my family, they ensured that they were always with me through thick and thin. Without their constant love and care, graduate school would have been much more difficult.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Preliminaries

The work presented in this article is based on basic algebraic number theory. However, for the readers with less familiarity of this background and as introduction to the language used throughout, in this section we give a brief account of the definitions and notations used.

1.1.1 Number fields and class groups [22]

Given a known fixed field F, an object of interest is its polynomial ring $F[x]$. It is natural consider different quotient rings of $F[x]$. Of particular interest are rings of the form $E = F[x]/(f(x))$, where $f(x)$ is an irreducible polynomial of degree n over F. We call $[E : F] = n$ the degree of the extension. $F[x]$ being a PID, $(f(x))$ is a maximal ideal, and hence E is a degree n field extension of F. K and its subrings are objects of interest that may have properties distinct from its base field F.

$F = \mathbb{Q}$, we have the $\mathbb{Q}$ extensions, which in general are called number fields. Concretely, we will identify any number field K as a subfield of $\mathbb{C}$. An order $\mathcal{O}$ of K is a subring of K finitely generated as a $\mathbb{Z}$-module (of rank $[K : \mathbb{Q}]$) such that $\mathcal{O} \otimes_{\mathbb{Z}} \mathbb{Q} = K$. There is a unique maximal order for every number field K and is known as the ring of integers of K, denoted by $\mathcal{O}_K$; the equivalent of $\mathbb{Z}$ for $\mathbb{Q}$. Let $\alpha \in \mathcal{O}$, then the map $\alpha\cdot : \mathcal{O} \to \mathcal{O}$ given by multiplication with α is a finitely generated linear map, we define the trace of this map to be $tr(\alpha)$, trace of α. For $\alpha_1, \alpha_2, ..., \alpha_n \in \mathcal{O}$ we may compute $D(\alpha_1, \alpha_2, ..., \alpha_n) = det(tr(\alpha_i \alpha_j))$, and if α_i's form a $\mathbb{Z}$-basis of $\mathcal{O}$ we call $(\alpha_1, \alpha_2, ..., \alpha_n)$ the discriminant of $\mathcal{O}$ denoted by $disc(\mathcal{O})$, and it is independent of the basis chosen. The discriminant $disc_K$ of a number field is defined to be the discriminant of its ring of integers $\mathcal{O}_K$.

Ring of integers are Dedekind domains, meaning that each ideal has a unique prime ideal factorization. Most notably, being a PID every element of $\mathbb{Z}$ has a unique prime factorization. In general, the ring of integers of a number field may not be a PID. Hence arises the necessity of accessing whether the ring of integers of a particular number field. The fractional ideals $\mathcal{I}_{\mathcal{O}}$ of an order $\mathcal{O}$ forms a multiplicative abelian group. The principal fractional ideals $\mathcal{P}_{\mathcal{O}}$ and the totally positive principal fractional ideals (principal ideals generated by elements whose all possible real embeddings are positive) $\mathcal{P}_{\mathcal{O}}^+$ are subgroups of $\mathcal{I}_{\mathcal{O}}$. The class group $Cl(\mathcal{O})$ and the narrow class group $Cl^+(\mathcal{O})$ of $\mathcal{O}$ are defined as follows.

$$Cl(\mathcal{O}) = \mathcal{I}_{\mathcal{O}}/\mathcal{P}_{\mathcal{O}} \tag{1.1}$$

$$Cl^+(\mathcal{O}) = \mathcal{I}_{\mathcal{O}}/\mathcal{P}_{\mathcal{O}}^+ \tag{1.2}$$

It is proven each of these two groups are finite. The class number $h(\mathcal{O})$ and the narrow class number $h^+(\mathcal{O})$ are defined to be the order of $Cl(\mathcal{O})$ and $Cl^+(\mathcal{O})$ respectively. The class group $Cl(K)$ and the narrow class group $Cl^+(K)$ of K is defined to be the corresponding groups of its maximal order $\mathcal{O}_K$, $Cl(\mathcal{O}_K)$ and $Cl^+(\mathcal{O}_K)$ respectively. Similarly defined are the class number $h(K)$ and narrow class number $h^+(K)$ of K. $\mathcal{O}_K$ is a PID iff $h(K) = 1$ i.e. $Cl(K)$ is trivial. The structure of $Cl(K)$ qualitatively assess the nature of deviation of $\mathcal{O}_K$ from being a PID. The more complicated structure $Cl(K)$ has, the more $\mathcal{O}_K$ deviates from being a PID. In general, class groups are computationally expensive to calculate and difficult to predict.

We are interested in the simplest of the number fields, namely degree 2 (quadratic) extensions of $\mathbb{Q}$. More concretely, quadratic extensions of $\mathbb{Q}$ are constructed by adjoining the square-root of a non-square integer a, denoted as $K = \mathbb{Q}(\sqrt{a})$. We will call K an imaginary quadratic field or a real quadratic field depending on whether $a < 0$ or $a > 0$. The discriminant of a quadratic field K is called a fundamental discriminant and denoted D_K. Imaginary quadratic fields have negative discriminants and real quadratic fields have positive discriminants. The discriminant of an order $\mathcal{O}$ of K are of the form $f^2 D_K$ for some integer $f > 0$ called the conductor of $\mathcal{O}$. All quadratic discriminants are $0,\ 1 \mod 4$. Given an integer $D \equiv 0,\ 1 \mod 4$, there is a unique quadratic order $\mathcal{O}$ among all quadratic extensions of $\mathbb{Q}$ such that $disc(\mathcal{O}) = D$. So, with a slight abuse of notation we will denote the class group of $\mathcal{O}$ by $Cl(D)$, and similarly for the narrow class groups and class numbers.

For a quadratic field K let us define d_K as follows.

$$d_K = \begin{cases} \frac{D_K}{4} & \text{if } D_K \equiv 0 \mod 4 \\ D_K & \text{if } D \equiv 1 \mod 4 \end{cases} \tag{1.3}$$

d_K is always a square-free integer, hence a fundamental discriminant is always either a square-free integer or 4 times a square-free integer. The ring of integers $\mathcal{O}_K$ can be characterised as follows.

$$\mathcal{O}_K = \begin{cases} \mathbb{Z}[\sqrt{d_K}] & \text{if } d \equiv 2,\ 3 \mod 4 \\ \mathbb{Z}[\frac{1+\sqrt{d_K}}{2}] & \text{if } d \equiv 1 \mod 4 \end{cases} \tag{1.4}$$

Quadratic extensions $K/\mathbb{Q}$ are galois with $Gal(K/\mathbb{Q}) \cong \mathbb{Z}/2\mathbb{Z}$. We have the following two definitions for the norm Nm.

$$\begin{aligned} Nm &: K \to \mathbb{Q} \\ \alpha &\mapsto \prod_{\sigma \in Gal(K/\mathbb{Q})} \sigma(\alpha) \end{aligned} \tag{1.5}$$

When restricted to $\mathcal{O}_K$, takes values in $\mathbb{Z}$.

$$\begin{aligned} Nm &: \mathcal{I}_{\mathcal{O}} \to \mathbb{Q} \\ I &\mapsto \prod_{\sigma \in Gal(K/\mathbb{Q})} \sigma(I) \end{aligned} \tag{1.6}$$

When restricted to ideals, takes values in $\mathbb{Z}$.

It is useful to note that any ideal in any order is generated as a $\mathbb{Z}$-module by two elements. For an ideal $I \in \mathcal{I}_{\mathcal{O}}$, we typically choose $Nm(I)$ and $\frac{-b+\sqrt{D}}{2}$ where $D = disc(\mathcal{O})$ and $b^2 \equiv D \mod Nm(I)$. So we write $I = \left[Nm(I), \frac{-b+\sqrt{D}}{2}\right]$.

1.1.2 Binary quadratic forms [4, 17]

Given a ring R, binary quadratic forms over R are homogeneous polynomials of degree 2 in $R[x, y]$. A binary quadratic form $Q(x, y) = ax^2 + bxy + cy^2$ may be denoted as $[a, b, c]$ in short form. Of particular interest to us are the binary quadratic forms over $\mathbb{Z}$. For convenience, we will use the phrase quadratic form(s) for integral binary quadratic form(s) unless specified.

We call a quadratic form $[a, b, c]$ primitive if $gcd(a, b, c) = 1$. We say that an integer m is represented by a quadratic form $Q(x, y)$ if there exists $x_0, y_0 \in \mathbb{Z}$ such that $Q(x_0, y_0) = m$, and we say m is primitively represented if $gcd(x_0, y_0) = 1$. We also define a discriminant $D_Q = b^2 - 4ac$ for a quadratic form $Q = [a, b, c]$. Note that $D_Q \equiv 0,\ 1 \mod 4$. We are mostly interested in quadratic forms of non-square discriminants. For a fixed non-square integer D let $BQF(D)$ be the set of all primitive quadratic forms of discriminant D if $D > 0$ and the set of all primitive quadratic forms of discriminant D with positive x^2 coefficient if $D < 0$.

$SL_2(\mathbb{Z})$, the 2×2 integral matrices of determinant 1 has a natural action on the $BQF(D)$. Let $Q = [a, b, c] \in BQF(D)$ and $M = \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$, then the action of M on Q is defined as follows.

$$M \cdot Q(x, y) = Q((x, y)M) = Q(px + ry, qx + sy)$$

Interestingly, the discriminant is invariant under this action and preserves the sign of the x^2 coefficient in the case of $D < 0$, hence this qualifies as an action on $BQF(D)$. We simply say that two quadratic forms are equivalent if they lie in the same orbit under the $SL_2(\mathbb{Z})$ action. We denote the equivalence classes of $BQF(D)$ by $C(D)$. For fixed D, every quadratic form in $BQF(D)$ is equivalent to a reduced form $[a, b, c]$ such that $|b| \leq |a| \leq |c|$. It is not hard to check that the number of solutions (a, b, c) for $b^2 - 4ac = D$ under the condition for a reduced form is finite. Hence, $C(D)$ is finite.

Gauss in [9] presents a composition law $BQF(D) \times BQF(D) \to BQF(D)$ between two quadratic forms of the same discriminant, we now call as the Gauss composition law. This composition law is $SL_2(\mathbb{Z})$-invariant. Hence, under this composition law $C(D)$ is viewed as a finite abelian group. Following are some properties of the group $C(D)$.

1. The identity class is represented by $[1, 0, \frac{D}{4}]$ if $4 \mid D$ or $[1, 1, \frac{1-D}{4}]$ otherwise.

2. The inverse of a class represented by $[a, b, c]$ can be represented by $[a, -b, c]$.

1.1.3 Correspondence between $Cl^+(D)$ and $C(D)$ [4, 17]

Given an ideal $I = [\alpha, \beta]$, in a quadratic order $\mathcal{O}$ with discriminant D such that

$$\frac{\bar{\alpha}\beta - \alpha\bar{\beta}}{\sqrt{D}} > 0$$

we associate I to $Q_I \in BQF(D)$, where

$$Q_I(x, y) = \frac{(\alpha x - \beta y)(\bar{\alpha} x - \bar{\beta} y)}{Nm\,(I)}$$

This association results in a map

$$\begin{aligned} \psi : \mathcal{I}_{\mathcal{O}} &\to BQF(D) \\ I &\mapsto Q_I \end{aligned} \tag{1.7}$$

ψ projects down to a group isomorphism, $\tilde{\psi} : Cl^+(D) \to C(D)$.

For quadratic orders the class group is essentially the same as the narrow class group except for real quadratic orders whose fundamental unit ϵ is such that $Nm\,(\epsilon) = -1$. In that case the class group is isomorphic to the narrow class group modulo a subgroup of order 2.

$$\begin{cases} Cl(D) \cong Cl^+(D)/\langle -1 \rangle & \text{if } D > 0 \text{ and } \exists x, y \in \mathbb{Z}_{>0},\ x^2 - Dy^2 = -4 \\ Cl(D) \cong Cl^+(D) & \text{otherwise} \end{cases} \tag{1.8}$$

It is much easier to study the narrow class group since it can be identified with the equivalence classes of quadratic forms, which are concretely understood. Since we are concerned with the odd prime order ideal classes, there is no ambiguity in working with the equivalence classes of quadratic forms, and most of the time we continue to do so.

1.1.4 Notations and conventions

Throughout the text we will use the following notations in addition to those that will be defined within the text:

X	Large positive real number
μ	Mobius function
CRT	Chinese Remainder Theorem
$\mathcal{S}, \mathcal{R}$	Pair of disjoint finite set of rational primes, to specify local conditions on discriminants
k	Odd rational prime, that will denote the order of ideal classes of interest
Γ	$SL_2(\mathbb{Z})$, 2×2 integral matrices of determinant 1
$\mathcal{H}$	Complex upper half plane, $\Im z > 0$
$\Gamma \backslash \mathcal{H}$	Quotient space defined by the usual Γ action on $\mathcal{H}$; $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) \cdot z = \frac{pz+q}{rz+s}$
$\mathcal{F}$	$\left\{z \in \mathcal{H} \,;\, -\frac{1}{2} \leq \Re z < \frac{1}{2},\ \|\|z\|\| \geq 1\right\}$, the standard projection of $\Gamma \backslash \mathcal{H}$ on to $\mathcal{H}$
$\mathcal{F}_0$	$\left[-\frac{1}{2}, \frac{1}{3}\right) \times \left[\frac{\sqrt{3}}{2}, \infty\right) \subset \mathcal{H}$

Ω	Subset of $\Gamma\backslash\mathcal{H}$ with piecewise smooth boundary
ν	Hyperbolic probability on $\mathcal{H}$ invariant by Γ action ($d\nu = \frac{3}{\pi}\frac{1}{y^2}dydx$)
$\mathcal{D}$	$\{D \in \mathbb{Z}$ fundamental discriminant of quadratic number fields$\}$
$\mathcal{D}_X$	$\{D \in \mathcal{D}\ ;\ \lvert D\rvert \leq X\}$
$\mathcal{D}^{\mathcal{S},\mathcal{R}}$	$\{D \in \mathcal{D}\ ;\ p \mid D\ \forall p \in \mathcal{S},\ q \nmid D\ \forall q \in \mathcal{R}\}$
$\mathcal{D}_X^{\mathcal{S},\mathcal{R}}$	$\mathcal{D}^{\mathcal{S},\mathcal{R}} \cap \mathcal{D}_X$
${}^+\mathcal{D}$, ${}^-\mathcal{D}$	$\{D \in \mathcal{D}\ ;\ D > 0\}$, $\{D \in \mathcal{D}\ ;\ D < 0\}$
${}^+\mathcal{D}_X,\ {}^-\mathcal{D}_X$	${}^+\mathcal{D} \cap \mathcal{D}_X,\ {}^-\mathcal{D} \cap \mathcal{D}_X$
${}^+\mathcal{D}^{\mathcal{S},\mathcal{R}},\ {}^-\mathcal{D}^{\mathcal{S},\mathcal{R}}$	${}^+\mathcal{D} \cap \mathcal{D}^{\mathcal{S},\mathcal{R}},\ {}^-\mathcal{D} \cap \mathcal{D}^{\mathcal{S},\mathcal{R}}$
${}^+\mathcal{D}_X^{\mathcal{S},\mathcal{R}},\ {}^-\mathcal{D}_X^{\mathcal{S},\mathcal{R}}$	${}^+\mathcal{D} \cap \mathcal{D}_X^{\mathcal{S},\mathcal{R}},\ {}^-\mathcal{D} \cap \mathcal{D}_X^{\mathcal{S},\mathcal{R}}$
$\lvert A\rvert$	Cardinality if A is a set, absolute value if $A \in \mathbb{R}$
G_p	p-part of a group G, for prime p
$G[p]$	The set of all p-order elements in a group G

$Aut(G)$	Automorphism group of a group G
$M_{N_1 \times N_2}(R)$	Additive group of $N_1 \times N_2$ matrices over the ring R
$\mathbb{Z}_p$	The p-adic integers
$\left(\frac{a}{b}\right)$	Legendre/Jacobi symbol

1.2 Cohen - Lenstra conjectures and k-order ideal class asymptotics [3, 8, 18]

Fix an odd prime k. Let us denote by m_N the (probability) Haar measure on $M_{N\times(N+u)}(\mathbb{Z}_k)$ and Π_k the set of isomorphism classes of finite abelian k-groups. For a given non-negative integer u we define in a very natural way a probability measure $\mathbb{P}_k^u$ on Π_k as follows.

$$\mathbb{P}_k^u(H) = \lim_{N\to\infty} m_N(\{A \in M_{N\times(N+u)}(\mathbb{Z}_k)\ ;\ coker A \cong H\})\ ;\ H \in \Pi_k \tag{1.9}$$

In [8] Friedman and Washington showed that $\mathbb{P}_k^0(H)$ is well defined and that it is inversely proportional to $|Aut(G)|$ but anything else. Later, in [30] a similar result was in general proven for all $u \geq 0$. More precisely it is now known that,

$$\mathbb{P}_k^u(H) = \frac{\frac{1}{|H|^u|Aut(H)|}}{\sum_{G\in\Pi_k}\frac{1}{|G|^u|Aut(G)|}} \tag{1.10}$$

where

$$\sum_{G\in\Pi_k}\frac{1}{|G|^u|Aut(G)|} = \prod_{n\geq 1}\left(1 - \frac{1}{k^{n+u}}\right)$$

Backed by experimental evidence, Cohen and Lenstra [3] conjectured that for any odd prime k, the k-part of quadratic class groups are asymptotically equidistributed with respect to $\mathbb{P}_k^u$.

Conjecture 1.2.1 (Cohen - Lenstra [3]). *Let k be a fixed odd prime and $f : \Pi_k \to \mathbb{C}$ be any 'reasonably well behaving' function. Then we have the following,*

$$\lim_{X\to\infty} \frac{1}{|^{-}\mathcal{D}_X|} \sum_{D\in^{-}\mathcal{D}_X} f(Cl(D)_k) = \int_{\Pi_k} f \, d\mathbb{P}_k^0 \tag{1.11}$$

$$\lim_{X\to\infty} \frac{1}{|^{+}\mathcal{D}_X|} \sum_{D\in^{+}\mathcal{D}_X} f(Cl(D)_k) = \int_{\Pi_k} f \, d\mathbb{P}_k^1 \tag{1.12}$$

An insight into the Cohen-Lenstra conjectures is as follows. Fix a quadratic field K, and let S be a finite set of primes ideals in K. Let us denote by $\mathcal{O}_{K,S}^\times$ the S-units in $\mathcal{O}_K$ ($\alpha \in K$ is an S-unit if all the prime factors of the principal fractional ideal generated by α are in S) and $\mathcal{I}_{K,S}$ denote the fractional ideals generated by the the primes in S. Since the prime ideals generate the fractional ideals, and the class group is finite, for large enough S we have

$$1 \longrightarrow \mathcal{O}_K^\times \longrightarrow \mathcal{O}_{K,S}^\times \xrightarrow{\alpha\to(\alpha\cdot\mathcal{O}_K)} \mathcal{I}_{K,S} \longrightarrow Cl(K) \longrightarrow 1$$

now dropping the first two terms and tensoring with $\mathbb{Z}_k$ we have,

$$\mathcal{O}_{K,S}^\times \otimes_{\mathbb{Z}} \mathbb{Z}_k \longrightarrow \mathcal{I}_{K,S} \otimes_{\mathbb{Z}} \mathbb{Z}_k \longrightarrow Cl(K) \otimes_{\mathbb{Z}} \mathbb{Z}_k \longrightarrow 1$$

hence we have that

$$Cl(K)_k \cong coker(\mathcal{O}_{K,S}^\times \otimes_{\mathbb{Z}} \mathbb{Z}_k \to \mathcal{I}_{K,S} \otimes_{\mathbb{Z}} \mathbb{Z}_k)$$

Note that $\mathcal{I}_{K,S} \otimes_{\mathbb{Z}} \mathbb{Z}_k \cong \mathbb{Z}_k^{|S|}$, and by Dirichlet's S-unit theorem $\mathcal{O}_{K,S}^\times \otimes_{\mathbb{Z}} \mathbb{Z}_k \cong \mathbb{Z}_k^{|S|}$ if K is imaginary with discriminant < -4 and $\mathcal{O}_{K,S}^{\times} \otimes_{\mathbb{Z}} \mathbb{Z}_k \cong \mathbb{Z}_k^{|S|+1}$ if K is real. So as $|S| \to \infty$,

$$Cl(K)_k = coker(A_{K,S})$$

where

$$\begin{cases} A_{K,S} \in M_{|S| \times |S|}(\mathbb{Z}_k) & \text{if } K \text{ is imaginary quadratic} \\ A_{K,S} \in M_{|S| \times (|S|+1)}(\mathbb{Z}_k) & \text{if } K \text{ is real quadratic} \end{cases}$$

Hence it reduces to proving that the sequences of $A_{K,S}$ are distributed similar to the respective Haar measure. Some progress has been made in [30, Theorem 1.2], by relaxing the sufficient conditions on $A_{K,S}$ matrices, for them to distribute according to the Haar measure.

Despite apparently having such strong arguments in favour, the main Cohen-Lenstra conjecture as well as its consequential sub-conjectures (also by Cohen-Lenstra) remain largely untamed. We are interested on one of such sub-conjectures, regarding the moments of k-order ideal classes.

Conjecture 1.2.2 ([3, Section 9]). *Let k be a fixed odd prime and n be a positive integer then,*

$$\lim_{X \to \infty} \frac{1}{|^{-}\mathcal{D}_X|} \sum_{D \in {}^{-}\mathcal{D}_X} |Cl(D)[k]|^n = c_{k,n}^{-} \tag{1.13}$$

$$\lim_{X \to \infty} \frac{1}{|^{+}\mathcal{D}_X|} \sum_{D \in {}^{+}\mathcal{D}_X} |Cl(D)[k]|^n = c_{k,n}^{+} \tag{1.14}$$

where $c_{k,n}^{\pm}$ are constants ($c_{k,1}^{-} = 1$ and $c_{k,1}^{+} = \frac{1}{k}$)

The only case where this conjecture is known to be proven is when $k = 3, n = 1$ and was already established by Davenport-Heilbronn [5, Theorem 3] even before Cohen-Lenstra conjectures were published. Later in [1] the Davenport-Heilbronn result was reestablished with exactly determined secondary error terms using improvements on the original techniques of Davenport-Heilbronn. Separately, in [27] the result was reproven with the same secondary error terms using Shintani-Zeta functions.

Theorem 1.2.3 (Bhargava et al. [1, Theorem 4] and Taniguchi et al. [27, Theorem 1.2])**.**

$$\lim_{X\to\infty} \frac{1}{|^{-}\mathcal{D}_X|} \sum_{D\in {}^{-}\mathcal{D}_X} |Cl(D)[3]| = 1 + b_{-}X^{-1/6} + O\left(X^{-1/6-1/48+\epsilon}\right) \tag{1.15}$$

$$\lim_{X\to\infty} \frac{1}{|^{+}\mathcal{D}_X|} \sum_{D\in {}^{+}\mathcal{D}_X} |Cl(D)[3]| = \frac{1}{3} + b_{+}X^{-1/6} + O\left(X^{-1/6-1/48+\epsilon}\right) \tag{1.16}$$

where $b_-, b_+ < 0$ are explicit constants.

The ℓ-torsion conjecture states that ℓ-torsion part of a class group is arbitrarily small compared to the size of its discriminant. If proven it is known to have many ramifications in number field theory and elliptic curve theory. However, the ℓ-torsion conjecture is far from being resolved completely.

Conjecture 1.2.4 (ℓ-torsion conjecture [18, Conjecture 1.1])**.** *Let ℓ be any prime. $K/\mathbb{Q}$ be any number field with discriminant D and degree n. Then for $\epsilon > 0$,*

$$|Cl(K)_{\ell\text{-}torsion}| \ll_{\epsilon,n,\ell} |D|^{\epsilon}$$

Now let us in particular consider quadratic fields. By genus theory the conjecture is known to be true for $\ell = 2$. However, for $\ell \geq 3$ the best proven bound is $|D|^{1/2-1/(2\ell)+\epsilon}$ [14, Theorem 1.1]. In [18, Theorem 1.2] it is proven that conjecture 1.2.2 implies the l-torsion conjecture for quadratic number fields. Hence the Cohen-Lenstra conjectures, in particular conjecture 1.2.2 is of significant importance.

In the next section we will briefly discuss how distribution properties of prime order ideal classes connect with these important conjectures. Hence, the conjectures stated in this sub-section provides motivation to our work.

1.3 Prior work on class group equidistribution results [12, 7, 25, 19]

Linnik [19, Theorem V.2.1] proved that the lattice points $\left(\frac{a}{\sqrt{|D|}}, \frac{2b}{\sqrt{|D|}}, \frac{c}{\sqrt{|D|}}\right)$ on the Hyperboloid space $\{(x, y, z) \in \mathbb{R}^3 \ ; \ yz - x^2 = 1\}$ equidistribute (with respect to a natural hyperbolic measure), where $[a, 2b, c] \in BQF(D)$ as $D \to -\infty$. However, Linnik's result was conditioned on $\left(\frac{D}{p}\right) = 1$, for a fixed odd prime. Later Skubenko [25, Theorem 1] proved in a similar manner the same result for real quadratic discriminants.

The ideal classes of $Cl^+(D)$ have natural representations on $\Gamma\backslash\mathcal{H}$ as CM-points $CM(D)$ if $D < 0$ and closed primitive geodesics $\mathcal{G}(D)$ if $D > 0$. Duke [7, Theorem 1] interprets Linnik-Skubenko results in terms of these geometrical representations on $\Gamma\backslash\mathcal{H}$ and then goes on to prove them unconditionally without the need for any odd prime such that $\left(\frac{D}{p}\right) = 1$. Duke's proofs uses spectral theory on $\Gamma\backslash\mathcal{H}$ and proves good enough bounds on higher order Fourier coefficients of spectral Mass forms of half integer weights.

Theorem 1.3.1 (Duke [7, Theorem 1(i)]). *Suppose $D \in^- \mathcal{D}$ and $\Omega \in \Gamma\backslash\mathcal{H}$. Then for some $\delta = \delta(\Omega) > 0$.*

$$\frac{|CM(D) \cap \Omega|}{|CM(D)|} = \nu(\Omega) + O\left(|D|^{-\delta}\right) \ \text{as } D \to -\infty \tag{1.17}$$

$$\tag{1.18}$$

Theorem 1.3.2 (Duke [7, Theorem 1(ii)]). *Suppose $D \in^+ \mathcal{D}$ and $\Omega \in \Gamma\backslash\mathcal{H}$. Then for some $\delta = \delta(\Omega) > 0$.*

$$\frac{\sum\limits_{C \in \mathcal{G}(D)} ||C \cap \Omega||}{\sum\limits_{C \in \mathcal{G}(D)} ||C||} = \nu(\Omega) + O\left(|D|^{-\delta}\right) \ \text{as } D \to \infty \tag{1.19}$$

where $||*||$ denote the length of a geodesic (segment).

1.4 Prime order ideal class equidistribution and summary of our contributions

For imaginary quadratic fields let us introduce $CM(D)[k]$ as the set of CM-points on $\Gamma\backslash\mathcal{H}$ associated to $Cl^+(D)[k]$. Let us define $CM_X[k] = \bigcup_{D\in^-\mathcal{D}_X} CM(D)[k]$. Our work primarily focuses on the asymptotic distribution of $CM_X[k]$ on $\Gamma\backslash\mathcal{H}$, pioneered by Hough.

Hough [11, 12] inspired by theorem 1.3.1 and backed by empirical evidence, conjectured that $CM_X[k]$ is equidistributed on $\Gamma\backslash\mathcal{H}$ (with respect to the hyperbolic measure) as $X \to \infty$. He went on to prove the conjecture for $k = 3$, and in the process reproving theorem 1.2.3 for imaginary quadratics, as a special case. However, the conjecture is still open for $k > 3$.

It is shown in [11, p. 11], that assumptions of uniform marginal and joint densities of prime-order CM-points on $\Gamma\backslash\mathcal{H}$ provides a proof direction to conjecture 1.2.2. As shown in section 1.2, proving conjecture 1.2.2 will have a significant impact in arithmetic statistics. Thus, there is strong motivation to study the distribution of prime-order CM-points. In fact, Hough's conjecture asserts that the prime-order CM-points have uniform marginal distribution. Conjectures on joint distributions are yet to be realised.

Our work was motivated by the significance of Hough's conjecture as stated above. We started by conducting experiments on the distribution of CM-points to reaffirm Hough's conjecture. The empirical evidences we thus obtained overwhelmingly supports the conjecture. Then we pushed along the same path as Hough but with much emphasis on investigating the asymptotics of $CM(X)[k]$ for $k > 3$. For $k > 3$, we managed to reduce the problem of equidisdribution of $CM_X[k]$ to a problem of equiduistribution of solutions to quadratic congruences modulo certain families of integers. Specifically we showed for $k > 3$ and $\Omega \subseteq \Gamma\backslash\mathcal{H}$

that,

$$\left| \left(\sum_{z \in CM_X[k] \cap \Omega} 1 \right) - \nu(\Omega) \left| {}^{-}\mathcal{D}_X \right| \right| < E_k$$

where E_k is the discrepancy of the sequence $\left\{ \frac{x}{t^2} \right\}$ as x runs through the solutions to $x^2 \equiv A^k \mod t^2$ for $(A, t) \in \left(O\left(X^{1/2}\right), O\left(X^{k/4-1/2}\right)\right)$. We achieved this first by computing the asymptotics of a heuristic counting function of $CM_X[k]$ and then realising that the actual asymptotics differ from the heuristic asymptotics by E_k. However, we were unable to find a decent estimate of E_k, that could actually prove Hough's conjecture.

Additionally, we managed to extend Hough's results to show that his result for $k = 3$ is true along the sub families ${}^{-}\mathcal{D}^{\mathcal{S},\mathcal{R}}$ of imaginary quadratic fields. We were able to prove that,

$$\sum_{\substack{z \in CM_X[k] \cap \Omega \\ \text{discriminant } \in {}^{-}\mathcal{D}^{\mathcal{S},\mathcal{R}}}} 1 \sim \nu(\Omega) \left| {}^{-}\mathcal{D}_X^{\mathcal{S},\mathcal{R}} \right|$$

for any pair of mutually exclusive finite sets of primes $\mathcal{S}$ and $\mathcal{R}$.

Next we investigated the possibility of an analogue of Hough's conjecture for (odd) prime order ideal classes of real quadratic fields. Inspired by Duke's theorem (theorem 1.3.2), we conducted experiments to investigate any possibility of equidistribution of lengths of geodesics corresponding to k-order ideal classes. The experimental results provide strong evidence in favor. Let us denote by $\mathcal{G}_X[k]$ as the collection of all closed primitive geodesics in $\Gamma \backslash \mathcal{H}$ corresponding to the k-order ideal classes with fundamental positive discriminant less than X. Based on the empirical evidence we conjectured the following,

$$\frac{\sum_{C \in \mathcal{G}_X[k]} ||C \cap \Omega||}{\sum_{C \in \mathcal{G}_X[k]} ||C||} \sim \nu(\Omega)$$

for all odd primes $k \geq 3$.

Our effort to provide a proof for the above conjecture although not successful, lead us to generalize an asymptotic formula by Siegel on the sum of the lengths of geodesics associated to ideal classes. Siegel [24, p. 1] proved that,

$$\sum_{C} ||C|| \sim \frac{\pi^2}{9\zeta(3)} X^{3/2}$$

where the sum runs across all closed primitive geodesics on $\Gamma\backslash\mathcal{H}$ associated to the ideal classes of all quadratic orders with discriminant less than X. We proved a similar asymptotic for geodesics associated to ideal classes in maximal orders. We also showed that it is true for fundamental discriminants when restricted to ${}^{+}\mathcal{D}^{\mathcal{S},\mathcal{R}}$ for any two finite and mutually exclusive sets of primes $\mathcal{S}$, $\mathcal{R}$. Most importantly, we were able to show equidistribution of the sum of the lengths of these geodesics. More specifically for $\Omega \subseteq \Gamma\backslash\mathcal{H}$ we proved the following,

$$\sum_{C} ||C \cap \Omega|| \sim \frac{\pi^2}{9} \nu(\Omega) c_{\mathcal{S},\mathcal{R}} X^{3/2}$$

where the sum runs over the geodesics associated to ideal classes of fundamental discriminants in ${}^{+}\mathcal{D}_X^{\mathcal{S},\mathcal{R}}$. The constant $c_{\mathcal{S},\mathcal{R}}$ is explicitly determined.

Chapter 2

Imaginary quadratic - Equidistribution of odd-prime order CM-points

2.1 Problem setup

In this chapter we consider imaginary quadratic extensions $K/\mathbb{Q}$. We will denote the fundamental discriminant by $D_K = -D$ and $d_K = -d$ (as defined in eq. (1.3)).

We have the following correspondence between the ideals $\mathcal{I}_K$ of ring of integers, the binary quadratic forms $BQF(-D)$ of discriminant $-D$ and CM-points on the complex upper half plane $\mathcal{H}$ [4, 17],

$$\begin{array}{ccccc} \mathcal{I}_K & \longrightarrow & BQF(-D) & \longrightarrow & \mathcal{H} \\ \left[a, \frac{b+\sqrt{-D}}{2}\right] & \longrightarrow & ax^2 - bxy + \frac{b^2+D}{4a}y^2 & \longrightarrow & \frac{b+\sqrt{-D}}{2a} \end{array} \tag{2.1}$$

The map $BQF(-D) \to \mathcal{H}$ projects down to a map between the Γ-equivalence classes on the two spaces. Let us denote,

$$CM(-D) = img(C(-D)) \subset \Gamma\backslash\mathcal{H}$$

The correspondence in eq. (2.1) projects down to a bijection between $Cl^+(-D)$, $C(-D)$ and

$CM(-D)$,

$$
\begin{array}{ccccc}
Cl^+(-D) & \longleftrightarrow & C(-D) & \longleftrightarrow & CM(-D) \\
\\
\left[a, \frac{b+\sqrt{-D}}{2}\right] & \longleftrightarrow & ax^2 - bxy + \frac{b^2+D}{4a}y^2 & \longleftrightarrow & \frac{b+\sqrt{-D}}{2a}
\end{array} \tag{2.2}
$$

Let us fix an odd prime k. Define the following,

$$CM(-D)[k] = img(Cl^+(-D)[k]) \subset CM(-D)$$

$$CM_X[k] = \bigcup_{-D \in {}^-\mathcal{D}} CM(-D)[k]$$

$$H_X[k] = |CM_X[k]|$$

$$H_X[k]\Omega = |CM_X[k] \cap \Omega|$$

$$H_X^\Omega[k] = |CM_X[k] \cap \Omega|$$

$$H_X(\mathcal{S}; \mathcal{R})[k] = \left| \bigcup_{-D \in {}^-\mathcal{D}^{\mathcal{S},\mathcal{R}}} CM(-D)[k] \right|$$

$$H_X^\Omega(\mathcal{S}; \mathcal{R})[k] = \left| \bigcup_{-D \in {}^-\mathcal{D}_X^{\mathcal{S},\mathcal{R}}} CM(-D)[k] \cap \Omega \right|$$

for $\Omega \subset \Gamma \backslash \mathcal{H}$.

Following Duke's theorem, a straight forward result is that CM_X is also equidistributed on $\Gamma\backslash\mathcal{H}$. But an interesting question along this line with a view towards Cohen-Lenstra is whether this equidistribution still holds when restricted to $CM_X[k]$.

2.2 Hough's work [12]

Hough [12, Conjecture 1] backed by several computational data conjectured the following,

Conjecture 2.2.1 (Hough)**.** *Fix any odd prime k. Let Φ be a compactly supported continuous function defined on $\Gamma\backslash\mathcal{H}$. Then $CM_X[k]$ is equidistributed on $\Gamma\backslash\mathcal{H}$ (with respect to the hyperbolic probability measure ν) in the following sense,*

$$\lim_{X\to\infty} \frac{\sum\limits_{-D\in{}^{-}\mathcal{D}} \left(\sum\limits_{z\in CM(-D)[k]} \Phi(z) \right)}{\sum\limits_{-D\in{}^{-}\mathcal{D}} 1} = \int_{\Gamma\backslash\mathcal{H}} \Phi(z)d\nu(z) \tag{2.3}$$

Since $\sum\limits_{-D\in{}^{-}\mathcal{D}} 1 \sim \frac{3}{\pi^2}X$ we restate the conjecture with same hypotheses as,

$$\sum_{z\in CM_X[k]} \Phi(z) = \left(\frac{3}{\pi^2} \int_{\Gamma\backslash\mathcal{H}} \Phi(z)d\nu(z) \right) X + o\left(X\right) \tag{2.4}$$

Hough [12, Theorem 2.1] proved the conjecture for $k = 3$. We restate it here in the form of eq. (2.4).

Theorem 2.2.2 (Hough, $k = 3$)**.** *Let Φ be a compactly supported continuous function defined on $\Gamma\backslash\mathcal{H}$. Then $CM_X[3]$ is equidistributed on $\Gamma\backslash\mathcal{H}$ (with respect to the hyperbolic probability measure ν) in the following sense,*

$$\sum_{z\in CM_X[3]} \Phi(z) = \left(\frac{3}{\pi^2} \int_{\Gamma\backslash\mathcal{H}} \Phi(z)d\nu(z) \right) X + cX^{5/6} + o\left((X^{5/6}\right)$$

where $c < 0$ is an explicitly known constant.

Note that the secondary term agrees with previous work [1, Theorem 4] and [27, Theorem 1.2].

For $k > 3$, Hough [12, Theorem 2.2] proved a milder version termed "equidistributed in the cusp". We restate it in the form of eq. (2.4).

Theorem 2.2.3. *Fix any odd prime $k > 3$. For each X fix a parameter $T = T(X) \in \left((X^{1/2-1/(k-2)+\epsilon}, X^{1/2-1/k-\epsilon}\right))$. Let Φ be a compactly supported continuous function defined on $\Gamma\backslash\mathcal{H}$. We use the notation $\Phi(z) = \Phi(\Re z, \Im z)$. Then $CM_X[k]$ is equidistributed in the cusp (with respect to the hyperbolic probability measure ν) in the following sense,*

$$\sum_{z\in CM_X[k]} \Phi\left(\Re z, \frac{\Im z}{T}\right) = \left(\frac{3}{\pi^2}\int_{\Gamma\backslash\mathcal{H}} \Phi\left(\Re z, \frac{\Im z}{T}\right) d\nu(z)\right) X + O\left(\frac{X^{k/4+\epsilon}}{T^{k/2}}\right)$$

It is noteworthy that this result shows equidistribution only on the "region" $\Gamma\backslash\mathcal{H} \cap \{\Im z = O(T)\}$ which tends towards the cusp as $X \to \infty$, hence we call it equidistribution at the cusp. This is a significantly small region with respect to the hyperbolic measure. As shown in [12, p. 5], the imaginary part of points in $CM_X[k]$ is bounded (upto a constant) by $X^{1/2-1/k}$. To see why, assume $\mathfrak{a}$ is an ideal in $\mathcal{O}_K$ that is of order k in the class group. Then $\mathfrak{a}^k = (x + y\sqrt{-D})$, where $x, y \in \mathbb{Z}\left[\frac{1}{2}\right]$ and $y \neq 0$, so

$$Nm\left(\mathfrak{a}^k\right) = x^2 + Dy^2 \geq \frac{D}{4} \Rightarrow Nm\left(\mathfrak{a}\right) > \frac{D^{1/k}}{4}$$

Suppose z represents $\mathfrak{a}$ on the upper half plane by the correspondence in eq. (2.2), then

$$\Im z = \frac{\sqrt{D}}{2Nm\left(\mathfrak{a}\right)} < \frac{1}{8}D^{1/2-1/k} \ll X^{1/2-1/k}$$

The region with imaginary part greater than $X^{1/2-1/k}$ has a hyperbolic volume of $X^{-1/2+1/k}$, and when scaled by X, an excess of $X^{1/2-1/k}$ for the asymptotics of counting $CM_X[k]$. This explains the negative error term we see in Hough's result for $k = 3$ as well as in [1, 27].

2.3 Experimental data and analysis

We start with concrete computations of class groups to study there asymptotic patterns experimentally. The computations were done on computer nodes on the University of Maryland High Performance Cluster [6]. Class group data readily available on LMFDB [20] as well as data generated using algorithms on Pari/Gp were used in the experiments.

We first investigate some statistical results on the limiting behaviour of $H_X[k]$. Cohen-Lenstra [3, Section 9] (C6, $\alpha = 1$) conjectures that $H_X[k] \sim |^{-}\mathcal{D}| \sim \frac{3}{\pi^2}X$. Table 2.1 shows computed data for $H_X[k]$ for small odd primes k and $X = 2^n \cdot 10^6$, $n = 0...14$. Figure 2.1 shows the asymptotic behaviour of $H_X[k]$ normalized by $\frac{3}{\pi^2}X$. The data shows almost smooth convergence from below. For $k = 3$ it is already proven that the secondary term is negative which explains this behaviour. For $k > 3$ based on Hough's work and this data it is safe to conjecture that the secondary term is negative. In this chapter we will show further support for this claim in our computations.

Table 2.1 and fig. 2.1 show strong support for Cohen-Lenstra. Another important observation is that rate of convergence improves as k increases. In fig. 2.2 we plot log error $=$ $\log|\frac{3}{\pi^2}X - H_X[k]|$ vs $\log(X)$, to determine the secondary term for the asymptotics of $H_X[k]$. The data points exhibit almost perfect linearity, implying that the secondary term should be a power term of the form X^a, where $a = a(k)$. As shown in fig. 2.2 regression analysis gives us approximations for $a(k)$. Based on Hough's work [12], it is reasonable to conjecture that $a(k) = \frac{1}{2} + \frac{1}{k}$. Table 2.2 gives a comparison between the conjectural values of $a(k)$ against the empirically determined values for small odd primes. It has been shown [1, 27, 12] that in fact $a(3) = \frac{1}{2} + \frac{1}{3}$, which is also reaffirmed by our experiments. However, $a(k)$ for $k > 3$ is yet to be established and the experiments show that the conjectural value lies within 5%

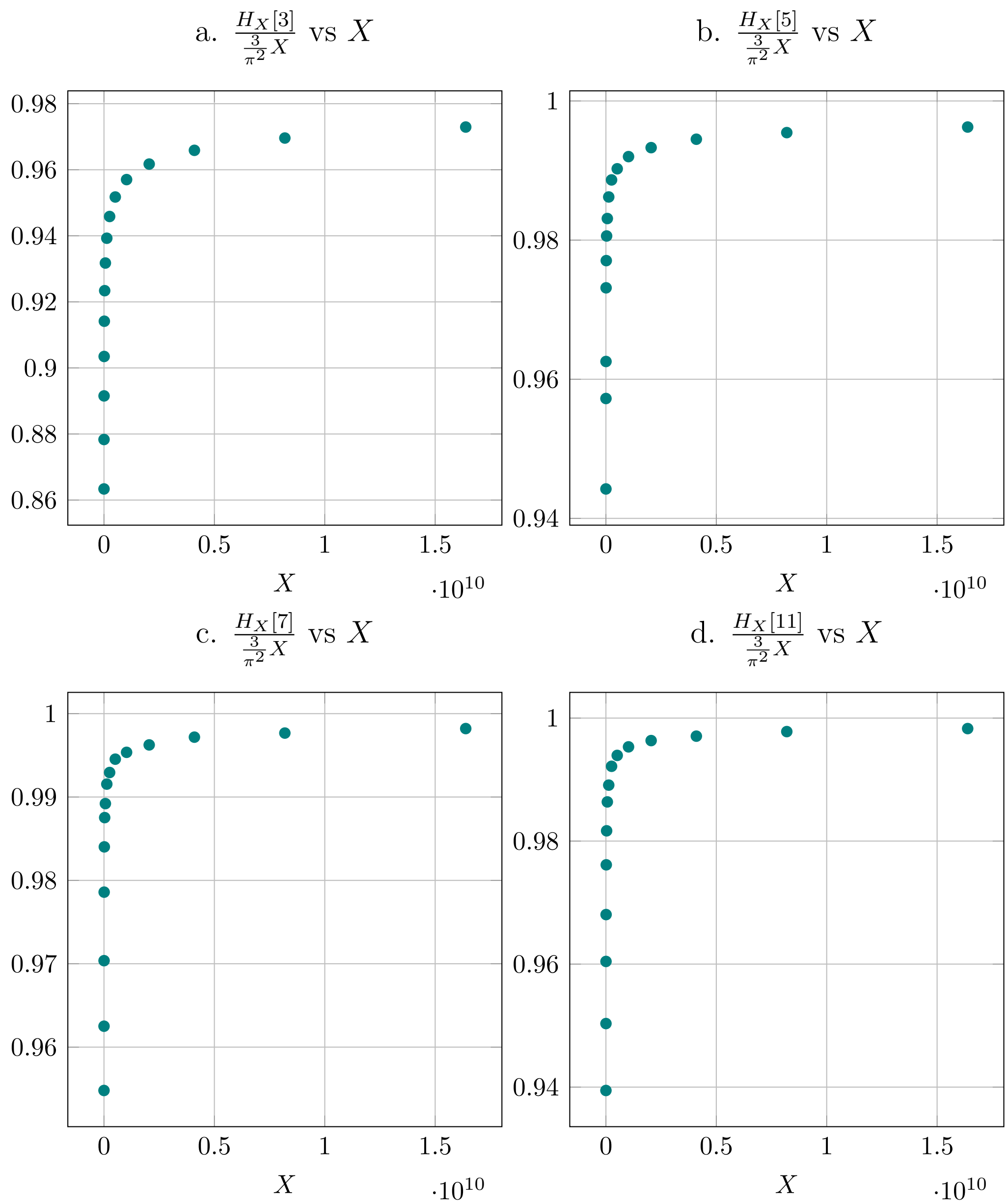

Figure 2.1: Asymptotic behaviour of $H_X[k]$

X	$\frac{3}{\pi^2}X$	$H_X[3]$	$H_X[5]$	$H_X[7]$	$H_X[11]$
$1 \cdot 10^6$	$3.0396 \cdot 10^5$	$2.6243 \cdot 10^5$	$2.8701 \cdot 10^5$	$2.9023 \cdot 10^5$	$2.8556 \cdot 10^5$
$2 \cdot 10^6$	$6.0793 \cdot 10^5$	$5.3396 \cdot 10^5$	$5.8192 \cdot 10^5$	$5.8513 \cdot 10^5$	$5.7773 \cdot 10^5$
$4 \cdot 10^6$	$1.2159 \cdot 10^6$	$1.084 \cdot 10^6$	$1.1703 \cdot 10^6$	$1.1798 \cdot 10^6$	$1.1678 \cdot 10^6$
$8 \cdot 10^6$	$2.4317 \cdot 10^6$	$2.197 \cdot 10^6$	$2.3664 \cdot 10^6$	$2.3796 \cdot 10^6$	$2.354 \cdot 10^6$
$1.6 \cdot 10^7$	$4.8634 \cdot 10^6$	$4.4458 \cdot 10^6$	$4.7518 \cdot 10^6$	$4.7857 \cdot 10^6$	$4.7474 \cdot 10^6$
$3.2 \cdot 10^7$	$9.7268 \cdot 10^6$	$8.9816 \cdot 10^6$	$9.538 \cdot 10^6$	$9.6055 \cdot 10^6$	$9.5486 \cdot 10^6$
$6.4 \cdot 10^7$	$1.9454 \cdot 10^7$	$1.8126 \cdot 10^7$	$1.9125 \cdot 10^7$	$1.9244 \cdot 10^7$	$1.9189 \cdot 10^7$
$1.28 \cdot 10^8$	$3.8907 \cdot 10^7$	$3.6544 \cdot 10^7$	$3.837 \cdot 10^7$	$3.8578 \cdot 10^7$	$3.8484 \cdot 10^7$
$2.56 \cdot 10^8$	$7.7815 \cdot 10^7$	$7.3601 \cdot 10^7$	$7.6931 \cdot 10^7$	$7.7265 \cdot 10^7$	$7.7205 \cdot 10^7$
$5.12 \cdot 10^8$	$1.5563 \cdot 10^8$	$1.4812 \cdot 10^8$	$1.5411 \cdot 10^8$	$1.5478 \cdot 10^8$	$1.5468 \cdot 10^8$
$1.02 \cdot 10^9$	$3.1126 \cdot 10^8$	$2.9788 \cdot 10^8$	$3.0877 \cdot 10^8$	$3.0982 \cdot 10^8$	$3.098 \cdot 10^8$
$2.05 \cdot 10^9$	$6.2252 \cdot 10^8$	$5.9869 \cdot 10^8$	$6.1834 \cdot 10^8$	$6.2018 \cdot 10^8$	$6.2024 \cdot 10^8$
$4.1 \cdot 10^9$	$1.245 \cdot 10^9$	$1.2026 \cdot 10^9$	$1.2382 \cdot 10^9$	$1.2415 \cdot 10^9$	$1.2414 \cdot 10^9$
$8.19 \cdot 10^9$	$2.4901 \cdot 10^9$	$2.4144 \cdot 10^9$	$2.4788 \cdot 10^9$	$2.4842 \cdot 10^9$	$2.4846 \cdot 10^9$
$1.64 \cdot 10^{10}$	$4.9801 \cdot 10^9$	$4.8453 \cdot 10^9$	$4.9614 \cdot 10^9$	$4.9713 \cdot 10^9$	$4.9716 \cdot 10^9$

Table 2.1: $H_X[k]$ for $k = 3, 5, 7, 11$

k	Conjectural; $\frac{1}{2} + \frac{1}{k}$	Empirically determined $a(k)$
3	0.83	0.83
5	0.70	0.73
7	0.64	0.67
11	0.59	0.62

Table 2.2: Comparison of $a(k)$ values - conjectural and empirical

error of the empirical value at least for $k = 3, 5, 7$. Data generated for much higher values of X may show improved support for the conjectural value. Later in this chapter we will show further support for the conjectural second term $X^{1/2+1/k}$.

Next we will focus on some empirical data to support Hough's conjecture. Let us first introduce some notations. In this section Ω will be a measurable set in $\Gamma\backslash\mathcal{H}$. We'll consider the following reformulation of conjecture 2.2.1,

$$\frac{H_X^{\Omega}[k]}{H_X[k]} \sim \nu(\Omega) \tag{2.5}$$

For the purpose of presentation in this section we define the sets $\Omega_{i,j} \subset \Gamma\backslash\mathcal{H}$ for $i \geq 1$ and

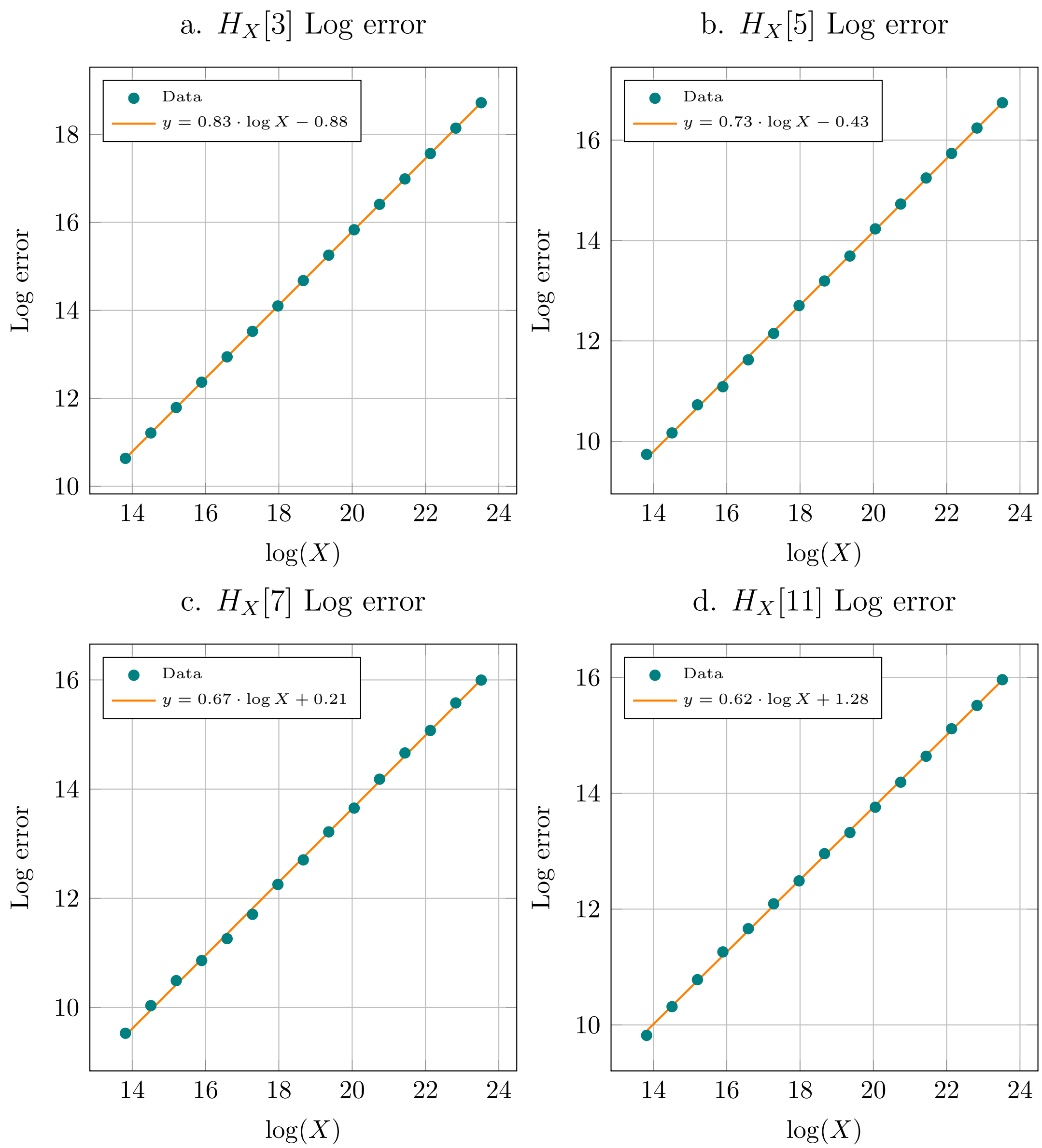

Figure 2.2: Secondary term estimation for $H_X[k]$

$j = 1, 2, 3$ as follows;

$$\Omega_{i,1} = \left\{ z \in \Gamma \backslash \mathcal{H} : \Re z \in [-\frac{1}{2}, -\frac{1}{6}), \Im z \geq i \right\}$$

$$\Omega_{i,2} = \left\{ z \in \Gamma \backslash \mathcal{H} : \Re z \in [-\frac{1}{6}, \frac{1}{6}), \Im z \geq i \right\}$$

$$\Omega_{i,3} = \left\{ z \in \Gamma \backslash \mathcal{H} : \Re z \in [\frac{1}{6}, \frac{1}{2}), \Im z \geq i \right\}$$

. Also, set $\Omega_i = \cup_{j=1}^{3} \Omega_{i,j}$. In following tables and diagrams we use the following definitions,

$$P_{i,j}(k, X) = \begin{cases} \frac{H_X^{\Omega_{i,j} \backslash \Omega_{i+1,j}}[k]}{H_X[k]} & \text{if } 1 \leq i < 10 \\ \frac{H_X^{\Omega_{i,j}}[k]}{H_X[k]} & \text{if } i \geq 10 \end{cases}$$

$$\nu_i = \nu_{i,j} = \begin{cases} \nu(\Omega_{i,j} \setminus \Omega_{i+1,j}) & \text{if } 1 \leq i < 10 \\ \nu(\Omega_{i,j}) & \text{if } i \geq 10 \end{cases}$$

Note that ν_i is well defined since $\nu(\Omega_{i,1}) = \nu(\Omega_{i,2}) = \nu(\Omega_{i,3})$ for all $i \geq 1$. For the sake of convenience we use the notation $P_{i,j}(k)$ instead of $P_{i,j}(k, X)$, dropping X. Also set $P_i(k) = \cup_{j=1}^{3} P_{i,j}(k)$.

Tables 2.3 to 2.5 show supporting evidence to eq. (2.5) for some chosen Ω. As we have seen in table 2.1, convergence rates seem to increase as k increases. We quantify the secondary term in fig. 2.3 for $\Omega_1 \setminus \Omega_2$. Since the data have been already normalised by $H_X[k] \asymp X$, we would expect the secondary term in this case to be of the order of $a(k) - 1$ and fig. 2.3 show strong support as expected.

X	ν_1	$P_{1,1}(3)$	$P_{1,2}(3)$	$P_{1,3}(3)$	$P_{1,1}(5)$	$P_{1,2}(5)$	$P_{1,3}(5)$	$P_{1,1}(7)$	$P_{1,2}(7)$	$P_{1,3}(7)$
$1 \cdot 10^6$	0.159	0.183	0.183	0.183	0.163	0.165	0.163	0.162	0.160	0.163
$2 \cdot 10^6$	0.159	0.180	0.180	0.180	0.163	0.165	0.163	0.162	0.161	0.162
$4 \cdot 10^6$	0.159	0.178	0.178	0.178	0.162	0.164	0.162	0.161	0.160	0.161
$8 \cdot 10^6$	0.159	0.176	0.176	0.176	0.161	0.163	0.161	0.160	0.160	0.160
$1.6 \cdot 10^7$	0.159	0.174	0.174	0.174	0.161	0.161	0.161	0.160	0.160	0.160
$3.2 \cdot 10^7$	0.159	0.172	0.172	0.172	0.161	0.161	0.161	0.160	0.160	0.160

Table 2.3: $P_{1,i}(k)$ for $i = 1, 2, 3$ and $k = 3, 5, 7$

X	ν_2	$P_{2,1}(3)$	$P_{2,2}(3)$	$P_{2,3}(3)$	$P_{2,1}(5)$	$P_{2,2}(5)$	$P_{2,3}(5)$	$P_{2,1}(7)$	$P_{2,2}(7)$	$P_{2,3}(7)$
$1 \cdot 10^6$	0.053	0.060	0.061	0.060	0.055	0.057	0.055	0.054	0.054	0.055
$2 \cdot 10^6$	0.053	0.060	0.060	0.060	0.054	0.055	0.054	0.054	0.054	0.054
$4 \cdot 10^6$	0.053	0.059	0.059	0.059	0.054	0.055	0.054	0.054	0.054	0.054
$8 \cdot 10^6$	0.053	0.058	0.058	0.058	0.054	0.055	0.054	0.054	0.054	0.054
$1.6 \cdot 10^7$	0.053	0.057	0.057	0.057	0.054	0.055	0.054	0.054	0.053	0.054
$3.2 \cdot 10^7$	0.053	0.057	0.057	0.057	0.054	0.054	0.054	0.053	0.053	0.053

Table 2.4: $P_{2,i}(k)$ for $i = 1, 2, 3$ and $k = 3, 5, 7$

X	ν_{10}	$P_{10,1}(3)$	$P_{10,2}(3)$	$P_{10,3}(3)$	$P_{10,1}(5)$	$P_{10,2}(5)$	$P_{10,3}(5)$	$P_{10,1}(7)$	$P_{10,2}(7)$	$P_{10,3}(7)$
$1 \cdot 10^6$	0.032	0.000	0.000	0.000	0.021	0.021	0.021	0.026	0.027	0.026
$2 \cdot 10^6$	0.032	0.000	0.000	0.000	0.023	0.023	0.023	0.026	0.027	0.026
$4 \cdot 10^6$	0.032	0.000	0.000	0.000	0.024	0.024	0.024	0.028	0.028	0.028
$8 \cdot 10^6$	0.032	0.003	0.003	0.003	0.026	0.026	0.026	0.029	0.029	0.029
$1.6 \cdot 10^7$	0.032	0.007	0.007	0.007	0.027	0.027	0.027	0.030	0.030	0.030
$3.2 \cdot 10^7$	0.032	0.009	0.010	0.009	0.028	0.028	0.028	0.030	0.030	0.030

Table 2.5: $P_{10,i}(k)$ for $i = 1, 2, 3$ and $k = 3, 5, 7$

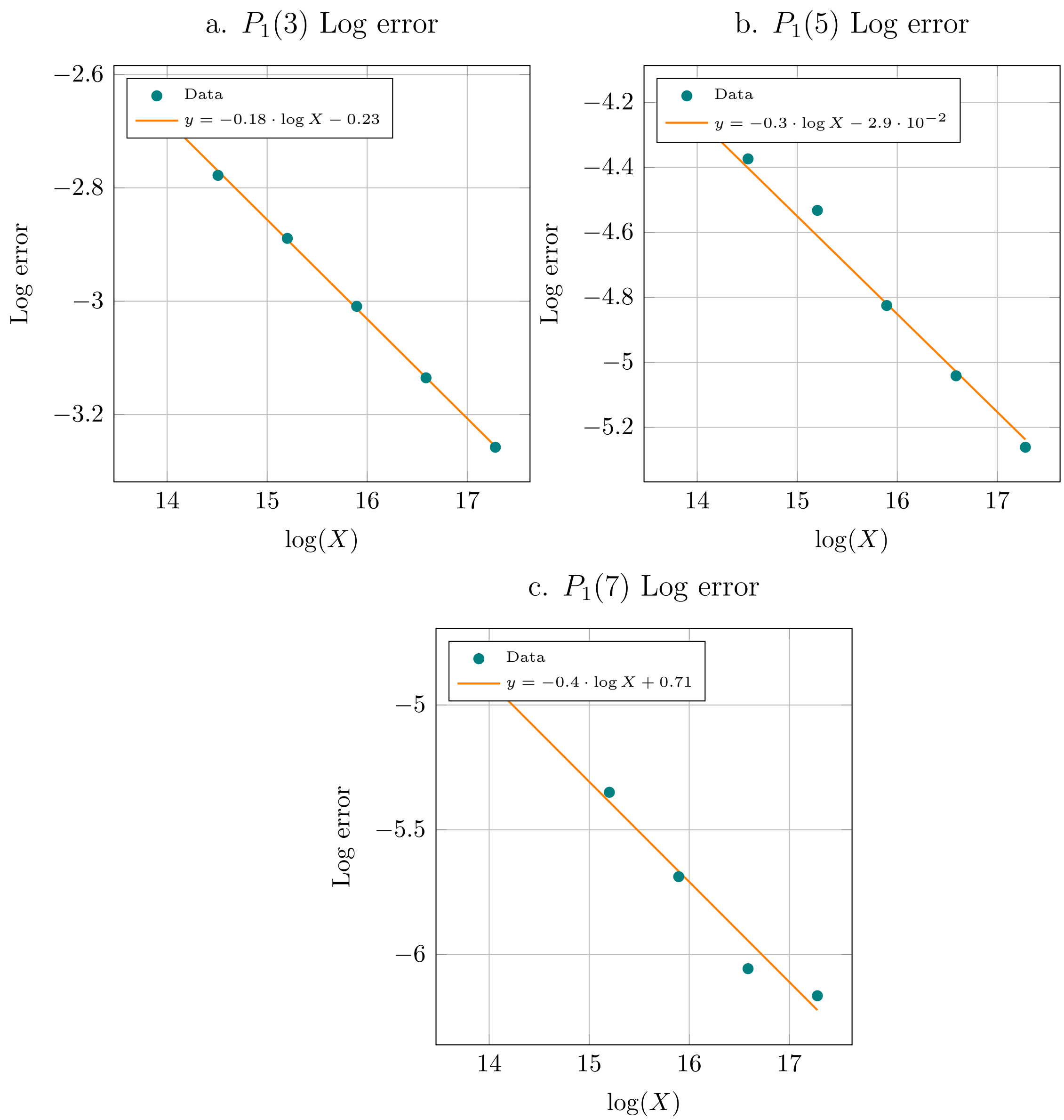

Figure 2.3: Secondary term estimation for $P_1(k)$

2.4 Theorems and proofs

2.4.1 Preliminaries on (equi)distribution (with respect to Lebegue mesaure) of real sequences on $[0,1)$ interval [16]

Let $\{x_n\}_{n\geq 1} \subset [0,1)$ be a finite sequence. We have the following definition for equidistribution (with respect to Lebegue mesaure).

Definition 2.4.1. *We say that the sequence* $\{x_n\}_{n\geq 1}$ *is equidistributed on* $[0,1)$ *if for all* $0 \leq \alpha < \beta < 1$ *we have that,*

$$\lim_{N\to\infty} \frac{|\{x_n \in [\alpha,\beta); n \leq N\}|}{N} = \beta - \alpha$$

However the definition is rarely used in practice to investigate the distribution of any sequence. Instead the Weyl's criterion allows a better way of evaluating the distribution of a sequence both qualitatively (whether or not equidistributed) and quantitatively (how well it is equidistributed). Let us define the following exponential sum associated with the sequence $\{x_n\}_{n\geq 1}$ for any integer h and positive integer N,

$$\mathbb{S}_h(\{x_n\}, N) = \sum_{n=1}^{N} e_h(x_n) \tag{2.6}$$

where $e_h(x) = e^{2\pi i h x}$.

Theorem (Weyl's criterion)**.** *The sequence* $\{x_n\}_{n\geq 1}$ *is equidistributed if and only if for all integers* $h \neq 0$

$$\lim_{n\to\infty} \frac{1}{N}\mathbb{S}_h(\{x_n\}, N) = 0 \tag{2.7}$$

Remark. *While Weyl's theorem provides a qualitative insight into the distribution of a sequence,* $\frac{1}{N}\mathbb{S}_h(\{x_n\}, N)$ *provides quantitative information of its distribution.*

We define the discrepancy of a sequence as follows.

Definition 2.4.2 (Discrepancy). *For a positive integer N, the discrepancy $\mathscr{D}(\{x_n\}, N)$ of a sequence $\{x_n\}_{n\geq 1}$ is defined as,*

$$\mathscr{D}(\{x_n\}, N) = \sup_{0\leq\alpha<\beta<1} \left| \frac{|\{x_n \in [\alpha, \beta)\}|}{N} - (\beta - \alpha) \right| \tag{2.8}$$

It is noteworthy that the discrepancy provides,

1. a straightforward evaluation of equidistributedness from the definition itself.

2. error estimate for counting sequence terms (assuming uniformity) in a subinterval of $[0, 1)$.

The Erdos-Turan inequality provides a computable upper bound for the discrepancy in terms of the exponential sums eq. (2.6).

Theorem (Erdos-Turan inequality). *For any positive integer H,*

$$\mathscr{D}(\{x_n\}, N) \leq \frac{6}{H+1} + \frac{4}{\pi} \sum_{h=1}^{H} \left| \frac{1}{N} \mathbb{S}_h(\{x_n\}, N) \right| \tag{2.9}$$

The inequality is obtained by estimating indicator functions of sub-intervals of $[0, 1)$ by finite exponential sums due to Fourier theory. The $\frac{6}{H+1}$ term results from such estimation errors by finite exponential sums.

It is evident that we may adapt the definition 2.4.2 to finite sets of real values on the interval $[0, 1)$, other than viewing them as finite subsequences of infinite sequences. Thus we do so hereafter. Notation wise given a finite set $\Sigma \in [0, 1)$ we write its discrepancy as $\mathscr{D}(\Sigma)$.

Definition 2.4.3 (Set of solutions to a quadratic congruences). *For a fixed integer A, a positive integer f and some large numbers T, R define $x_{T,R}(f, A) \subset [0, 1)$ as follows,*

$$x_{T,R}(f, A) = \left\{ \frac{x}{g} \in [0, 1); A \equiv x^2 \ mod\ g, (x, g) = 1, g = fr^2t^2, t < T, r < R \right\} \tag{2.10}$$

2.4.2 Main results

Following are the main theorems we state and prove in this chapter. In the following statements, suppose $\Omega \subset \Gamma\backslash\mathcal{H}$ have piecewise smooth boundary.

Theorem 2.4.4.

$$\left|H_X^\Omega[3] - \left(\nu(\Omega)\left|{}^{-}\mathcal{D}_X\right| + c_3 X^{1/2+1/3}\right)\right| = O\left(X^{9/10+\epsilon}\right) \tag{2.11}$$

where $c_3 = c_3(\Omega)$ is a negative constant.

Theorem 2.4.5. *Suppose $k > 3$ is an odd prime. Then,*

$$\begin{aligned}&\left|H_X^\Omega[k] - \left(\nu(\Omega)\left|{}^{-}\mathcal{D}_X\right| + c_k X^{1/2+1/k}\right)\right| \\ &= O\left(\sqrt{X}\max_{f|4}\left(\sum_{\substack{lm<\sqrt{X}\\ (l,m)=1,\ \mu(l)\neq 0\\ T=\frac{\sqrt{lm^k}}{lm}}} T\ \mathscr{D}\left(x_{T,\sqrt{X}}\left(f, lm^k\right)\right)\right)\right) + O\left(X^{13/16+\epsilon}\right)\end{aligned} \tag{2.12}$$

where $c_k = c_k(\Omega)$ is a negative constant.

The equidistribution property of the CM-points seems to hold across subfamilies of $\mathcal{D}$. In particular, on families defined by prime ramification conditions. The following theorems generalizes theorems 2.4.4 and 2.4.5 for such families.

Theorem 2.4.6. *Suppose $\mathcal{S}$ and $\mathcal{R}$ are mutually exclusive finite sets of prime numbers of order $O(1)$. Then,*

$$\left|H_X^\Omega\left(\mathcal{S};\mathcal{R}\right)[3] - \left(\nu(\Omega)\left|{}^{-}\mathcal{D}_X^{\mathcal{S},\mathcal{R}}\right| + c_{3,\Omega,\mathcal{S},\mathcal{R}} X^{1/2+1/3}\right)\right| = O\left(X^{9/10+\epsilon}\right) \tag{2.13}$$

where $c_{3,\Omega,S,R}$ is a constant.

Theorem 2.4.7. *Suppose $\mathcal{S}$ and $\mathcal{R}$ are mutually exclusive finite sets of prime numbers of order $O(1)$. Suppose $k > 3$ is an odd prime. Then,*

$$\left| H_X^{\Omega}(\mathcal{S};\mathcal{R})[k] - \left(\nu(\Omega)\,|^{-}\mathcal{D}_X^{\mathcal{S},\mathcal{R}}| + c_{k,\Omega,\mathcal{S},\mathcal{R}} X^{1/2+1/k}\right)\right|$$
$$= O\left(\sqrt{X} \max_{f | \prod_{q\in\mathcal{S}\cup\mathcal{R}} q^2} \left(\sum_{\substack{lm<\sqrt{X} \\ (l,m)=1,\ \mu(l)\neq 0 \\ T=\frac{\sqrt{lm^k}}{lm}}} T\ \mathscr{D}\left(x_{T,\sqrt{X}}\left(f, lm^k\right)\right)\right)\right) + O\left(X^{13/16+\epsilon}\right) \tag{2.14}$$

where $c_{k,\Omega,\mathcal{S},\mathcal{R}}$ is a constant.

As a corollary of theorem 2.4.4 we immediately deduce the following previously proven result [5, 1], which is also a consequence of the Cohen-Lenstra heuristics [3, Section 9].

Corollary 2.4.8.

$$\lim_{X\to\infty} \frac{H_X[3]}{|^{-}\mathcal{D}_X|} = 1 \tag{2.15}$$

Remark. *As seen in theorems 2.4.4 to 2.4.7, we include the secondary term $X^{1/2+1/k}$, even though there is a much bigger error term of either $O\left(X^{9/10+\epsilon}\right)$ or $O\left(X^{13/16+\epsilon}\right)$ depending on whether $k = 3$ or $k > 3$. This is to show that our methods too yield this specific (secondary) term with a negative constant which is proven in [1, 12] and agrees with the experimental findings shown in section 2.3, and also hoping that further refinements of our methods will eventually reduce these error terms. The derivation of this secondary term is discussed in the proof and it has the same flavor of reasoning as discussed in [12], namely that this secondary term is a consequence of CM-points corresponding to ideals of fields with large discriminant but relatively small norms. However, our methods fail to prove that this is the actual secondary term and also its negative contribution to the asymptotic in theorems 2.4.6 and 2.4.7.*

2.4.3 Parametrization of k-order CM-points

The key to counting k-order CM-points is a parametrization of k-order ideal classes due to Soundararajan [26, Proposition 3]. This was also used by Hough [12] in his proofs. We state this parametrization as a lemma followed by its proof due to Hough.

Suppose I is an ideal in the ring of integers of a imaginary quadratic field of discriminant $-D$. Recall that $I = \left[a, \frac{b+i\sqrt{D}}{2}\right]$ for some integers a and b. We call I a primitive ideal if $a = Nm\,(I) \leq \frac{1}{\sqrt{3}}\sqrt{D}$ and if I is not divisible by any principal ideal generated by an integer prime i.e. $I = (p)J$ is not satisfied for any prime $p \in \mathbb{Z}$ and any ideal J in the ring of integers.

Lemma 2.4.9 (Parametrization of k-order ideal classes)**.** *Let $-D \leq -20$ be a fixed negative fundamental discriminant and k a fixed odd prime. K be the corresponding quadratic field. Let $-d = d_K$ as defined in eq.* (1.3)*, So we have the cases $d \equiv 1, 2, 3 \mod 4$. Then the following is true about the set $S_D(k)$ of k-order primitive ideals $I \in \mathcal{O}_K$.*

Case1: $d \equiv 2 \mod 4$

$S_D(k)$ is is in bijection with the set of tuples of integers (l, m, n, t) that satisfy,

$$\begin{aligned} lm^k &= l^2n^2 + t^2d \\ l &\mid d \\ (m, dn) &= 1 \\ lm &< \frac{2}{\sqrt{3}}\sqrt{d} \\ l, m, t &> 0 \end{aligned} \tag{2.16}$$

Furthermore, under this bijection the tuple (l, m, n, t) maps to the ideal $I = \left[a, \frac{b+i\sqrt{D}}{2}\right]$ where

$a = lm$, *b is even and* $\frac{b}{2} \equiv lnt^{-1} \mod lm$. *Here* t^{-1} *is the multiplicative inverse of* $t \mod m$.

Case2: $d \equiv 1 \mod 4$

$S_D(k)$ *is is in bijection with the set of tuples of integers* (c, l, m, n, t) *that satisfy,*

$$\begin{aligned} clm^k &= l^2n^2 + t^2d \\ l &\mid d \\ (m, 2dn) &= 1 \\ clm &< \frac{2}{\sqrt{3}}\sqrt{d} \\ l, m, t &> 0 \\ c &\in \{1, 2\} \end{aligned} \tag{2.17}$$

Furthermore, under this bijection the tuple (c, l, m, n, t) *maps to the ideal* $I = \left[a, \frac{b+i\sqrt{D}}{2}\right]$ *where* $a = clm$, *b is even and* $\frac{b}{2} \equiv lnt^{-1} \mod clm$. *Here* t^{-1} *is the multiplicative inverse of* $t \mod cm$.

Case3: $d \equiv 3 \mod 4$

$S_D(k)$ is is in bijection with the set of tuples of integers (l, m, n, t) that satisfy,

$$\begin{aligned} 4lm^k &= l^2n^2 + t^2d \\ l &\mid d \\ \begin{cases} (m, d\frac{n}{2}) = 1 & \textit{if } 2 \mid n \\ (m, dn) = 1 & \textit{otherwise} \end{cases} \\ lm &< \frac{1}{\sqrt{3}}\sqrt{d} \\ l, m, t &> 0 \end{aligned} \tag{2.18}$$

Furthermore, under this bijection the tuple (l, m, n, t) maps to the ideal $I = \left[a, \frac{b+i\sqrt{D}}{2}\right]$ where $a = lm$, b is odd and $b \equiv lnt^{-1} \mod 2lm$. Here t^{-1} is the multiplicative inverse of t $\mod 2m$.

Note that by the properties of fundamental discriminants, d and l are square-free integers

The simple idea behind the proof is that if the k-power of an ideal is principal, then its norm is a k-power of an integer. We only provide the proof for the case $d \equiv 2 \mod 4$ since the other two cases are almost identical proofs.

For Case1. First let us assume the existence of (l, m, n, t) satisfying eq. (2.16). By factorization in K we get, $lm^k = (ln + t\sqrt{-d})(ln - t\sqrt{-d})$. Since $l \mid d$, $(l \cdot \mathcal{O}_K) = \alpha^2$ for some ideal α that is a factor of the different ideal of K. We claim that α is the gcd of the ideals generated by $(ln + t\sqrt{-d})$ and $(ln - t\sqrt{-d})$. Since $\alpha = \bar{\alpha}$, it is easy to see that α is a common divisor of the two ideals. Suppose, that any other unramifiying integer prime p shares a divisor with the two ideals in question. Suppose p splits in K as $(p \cdot \mathcal{O}_K) = \mathfrak{p}\bar{\mathfrak{p}}$ as ideals and that w.l.o.g say $\mathfrak{p} \mid ((ln+t\sqrt{-d}) \cdot \mathcal{O}_K)$ by conjugation $\bar{\mathfrak{p}} \mid ((ln-t\sqrt{-d}) \cdot \mathcal{O}_K)$. But by assumption $\mathfrak{p} \mid ((ln-t\sqrt{-d}) \cdot \mathcal{O}_K)$ and hence again by conjugation $\bar{\mathfrak{p}} \mid ((ln+t\sqrt{-d}) \cdot \mathcal{O}_K)$. This shows that p is a common divisor of both $ln + t\sqrt{-d}$ and $ln - t\sqrt{-d}$, which violates the condition $(m, dn) = 1$. Similarly, for p unsplit prime we can get a contradiction working with conjugates. Claim proven.

So we write $((ln + t\sqrt{-d}) \cdot \mathcal{O}_K) = \alpha\beta$ and $((ln - t\sqrt{-d}) \cdot \mathcal{O}_K) = \alpha\bar{\beta}$ where β and $\bar{\beta}$ are coprime, and each of them only divisible by primes resulting from split integer primes, otherwise the condition $(m, dn) = 1$ will be violated. We have, $Nm\,(\beta) = \beta\bar{\beta} = (m \cdot \mathcal{O}_K)^k$, hence $\beta = \gamma^k$ for some ideal γ of norm m. Set $I = \alpha\gamma$. Note that I is not divisible by any unsplit primes by our choice of α and γ. Also, $Nm\,(I) = Nm\,(\alpha) = lm \leq \frac{2}{\sqrt{3}}\sqrt{d} = \frac{1}{\sqrt{3}}\sqrt{D}$. Hence, I is a primitive ideal.

Suppose I is principal then, $I = ((x + y\sqrt{-d}))$ for some integers x and y where $|y| \geq 1$ (since $y = 0$ violates primitivity of I). Then, $\frac{2}{\sqrt{3}}\sqrt{d} > Nm\,(I) = x^2 + b^2 d \geq b^2 d \geq d \Rightarrow d = 1$, which is a contradiction because $2 \mid d$. So I is non-principal. Now to show that I is k-torsion in $Cl(-D)$, consider $I^k = \alpha^k \gamma^k = \alpha^{k-1} \cdot \alpha\beta = (l \cdot \mathcal{O}_K)^{\frac{k-1}{2}} \cdot \alpha\beta$. Since k odd, $l^{\frac{k-1}{2}}$ is an integer, and $((ln + t\sqrt{-d}) \cdot \mathcal{O}_K) = \alpha\beta$ is a principal ideal. So I^k is principal.

To prove that no two solutions give the same ideal. Suppose there is two solutions to eq. (2.16), (l_1, m_1, n_1, t_1) and (l_2, m_2, n_2, t_2) that give the same ideal $I = I_1 = I_2$. Then $I_1\bar{I}_2 = (Nm\,(I) \cdot \mathcal{O}_K) = (x \cdot \mathcal{O}_K)$ for some integer x. We have $(x \cdot \mathcal{O}_K)^k = {I_1}^k {\bar{I}_2}^k = \left({l_1}^{\frac{k-1}{2}} \cdot \alpha_1\beta_1\right) \cdot \left({l_2}^{\frac{k-1}{2}} \cdot \alpha_2\bar{\beta}_2\right) = (l_1 l_2)^{\frac{k-1}{2}} \cdot \alpha_1\beta_1 \cdot \alpha_2\bar{\beta}_2 = (l_1 l_2)^{\frac{k-1}{2}}((l_1 n_1 + t_1\sqrt{-d}) \cdot \mathcal{O}_K)((l_2 n_2 - t_2\sqrt{-d}) \cdot \mathcal{O}_K)$, by equating the $\sqrt{-d}$ coefficient of the generators of these ideals we get that $l_1 n_1 t_2 = \pm l_2 n_2 t_1$, by the conditions in eq. (2.16) and since l_1, l_2 square-free, $(l_1 n_1, t_1) = (l_2 n_2, t_2) = 1 \Rightarrow t_1 = t_2$, $l_1 n_1 = \pm l_2 n_2 \Rightarrow l_1 {m_1}^k = l_2 {m_2}^k$, and then again by the conditions in eq. (2.16), $l_1 = l_2, m_1 = m_2$. But we still have that $n_1 = \pm n_2$. Suppose $n_1 = -n_2$, looking at the proof above we notice that this means $I = \bar{I}$ which means that I is 2-torsion in the classgroup, which cannot happen. So $n_1 = n_2$.

Now assume I is k-order primitive ideal. Let us factorise as $I = \alpha\gamma$, where α is product of all ramifying prime powers that divide I, and γ is the non-ramifying counterpart. By unique prime factorisation of ideals, this factorisation is unique. Set $Nm(\alpha) = l$ and $Nm(\gamma) = m$, then $Nm(I) = lm \leq \frac{2}{\sqrt{3}}\sqrt{d}$. Now, $I^k = ((x + y\sqrt{-d}) \cdot \mathcal{O}_K)$ for some integers x, y. To get rid of the ambiguity due to units, we choose to restrict y to be non-negative and since I is not principal, $y > 0$. Also, $I^k = (\alpha\gamma)^k$ and $\alpha^k = l^{\frac{k-1}{2}}\alpha$, so $l^{\frac{k-1}{2}} \mid (x, y)$. Set $x = l^{\frac{k-1}{2}} n_0$ and $y = l^{\frac{k-1}{2}} t$, we have that $\alpha\gamma^k = ((n_0 + t\sqrt{-d}) \cdot \mathcal{O}_K)$. Then taking norms $lm^k = {n_0}^2 + t^2 d$, since $l \mid d$, $l \mid n_0$ so set $n_0 = ln$. We finally have $lm^k = l^2 n^2 + t^2 d$, and it is easy to see that $l, m, t > 0$. By the primitive condition on I we get that $(m, dn) = 1$. Hence we have obtained a tuple (l, m, n, t) satisfying condition eq. (2.16). Also, it is easy to see that the mapping from such a tuple to I and I to the tuple are inverses of each other. Hence we have established the bijection.

Now let us determine a, b such that $I = \left[a, \frac{b+\sqrt{-D}}{2}\right]$ in terms of l, m, n, t. However, $\frac{b+\sqrt{-D}}{2} \in \mathcal{O}_K \Rightarrow \frac{b^2+D}{4} = Nm\left(\frac{b+\sqrt{-D}}{2}\right) \in \mathbb{Z}$ and since in this case $D = 4d$, b is even. So we replace b by $\frac{b}{2}$, still an integer, and recycling 'b' for convenience, we write $I = \left[a, b + \sqrt{-d}\right]$. From previous statements in this proof we have that $a = Nm(I) = lm$, so $I = \left[lm, b + \sqrt{-d}\right]$. Now $(l \cdot \mathcal{O}_K)\gamma^2 = I^2 = [l^2 m^2, lmb + lm\sqrt{-d}, b^2 - d + 2b\sqrt{-d}]$, considering the constant coefficients of the generators of the ideal I^2 we conclude that $l \mid (b^2 - d) \Rightarrow l \mid b^2 \Rightarrow l \mid b$, because l is square-free. So set $b = lb'$, hence we have $\gamma^2 = [lm^2, lmb' + m\sqrt{-d}, {b'}^2 - d/l + 2b'\sqrt{-d}]$. Note that $lm \in I$ and $lm^2, lmb' + m\sqrt{-d} \in \gamma^2$. Hence, $(lm)(lm^2)^{\frac{k-3}{2}}(lmb' + m\sqrt{-d}) \in I(\gamma^2)^{\frac{k-3}{2}}\gamma^2 = I\gamma^{k-1} = \alpha\gamma^k = \alpha\beta = ((ln + t\sqrt{-d}) \cdot \mathcal{O}_K)$. So for some $x, y \in \mathbb{Z}$, $l^{\frac{k+1}{2}} m^k b' + l^{\frac{k-1}{2}} m^{k-1}\sqrt{-d} = (ln + t\sqrt{-d})(x + y\sqrt{-d})$, multiplying both sides by m and factoring out lm^k on the left hand side we have $lm^k(l^{\frac{k-1}{2}} b' + l^{\frac{k-3}{2}}\sqrt{-d}) = (ln + t\sqrt{-d})(mx + my\sqrt{-d})$, factoring $lm^k = (ln + t\sqrt{-d})(ln - t\sqrt{-d})$ we can cancel out $(ln + t\sqrt{-d})$. We have $(ln - t\sqrt{-d})(l^{\frac{k-1}{2}} b' + l^{\frac{k-3}{2}}\sqrt{-d}) = mx + my\sqrt{-d} \Rightarrow (l^{\frac{k+1}{2}} nb' + l^{\frac{k-3}{2}} td) + (l^{\frac{k-1}{2}} n + l^{\frac{k-1}{2}} tb')\sqrt{-d} = mx + my\sqrt{-d}$, equating the $\sqrt{-d}$ term we have $l^{\frac{k-1}{2}} n \equiv l^{\frac{k-1}{2}} tb' \mod m \Rightarrow b' \equiv nt^{-1}$

mod $m \Rightarrow b \equiv lnt^{-1} \mod lm$.

Case2 and Case3, have identical proofs. $\square$

The extra c that comes in Case2 comes from the fact that 2 is ramified, but l is forbidden to be divisible by 2 since $l \mid d$ and d is odd unlike in Case1. So an extra $c = 2$ is possible in $Nm(\alpha) = cl$. The extra 4 in $4lm^k$ in Case3 comes from the fact that in this case $\mathcal{O}_K$ contains elements with half integers rather than integers. So, clearing out denominators give an extra 4. Also, note that for each (l, m, t) you can choose $\pm n$. The two solutions (l, m, n, t) and $(l, m, -n, t)$ gives conjugate ideals, both of which are k-order primitive.

2.4.4 Counting CM-points under a heuristic assumption

The CM-point z_I that correspond to the ideal $I = \left[a, \frac{b+i\sqrt{D}}{2}\right]$ is given by $z_I = \frac{b}{2a} + i\frac{\sqrt{D}}{2a}$. For primitive ideals, $z_I \in \{\Im z \geq \frac{\sqrt{3}}{2}\}$, and since in the case of the parametrization of k-order ideals b is defined $\mod 2a$ we may choose b such that the interested ideals fall into $\mathcal{F}_0$. Since $\mathcal{F} \subset \mathcal{F}_0$, it is sufficient to prove equidistribution in $\mathcal{F}_0$, which is a much more easier domain to work with.

Now notice that the following map transforms $\mathcal{H}$ from a hyperbolic measure space to a Lesbegue measure space.

$$\begin{aligned} (\mathcal{H}, \nu) &\to (\mathcal{H}, \lambda) \\ x + iy &\mapsto x + i\frac{1}{y} \\ \frac{1}{y^2} dydx &\mapsto dydx \end{aligned} \tag{2.19}$$

Hence with this we can simplify our problem to showing equidistribution with respect to Lesbegue measure λ on $\mathcal{F}_1 = \left[-\frac{1}{2}, \frac{1}{2}\right) \times \left(0, \frac{2}{\sqrt{3}}\right]$ with just a simple transformation of the CM-points. We prove theorems 2.4.4 and 2.4.5 for $\Omega = [0, u] \times (0, v] \in \mathcal{F}_1$ where $0 < u < \frac{1}{2}$ and $0 < v < \frac{2}{\sqrt{3}}$ and the case where $\Omega = [-u, 0] \times (0, v] \in \mathcal{F}_1$ follows from symmetry. It is easy to see that the results follows easily for general measurable sets from these types of rectangular domains.

We write down the full details of the proof only for Case1, since as the proofs are identical except for few minor technicalities. Now if $d \equiv 2 \mod 4$, a k-order ideal $I = \left[a, \frac{2b+\sqrt{-4d}}{2}\right] = \left[a, b+\sqrt{-d}\right]$ for integers $a = lm$ and $b \equiv lnt^{-1} \mod lm$ where l, m, n, t satisfy the conditions in eq. (2.16). Now in place of $z_I = \frac{nt^{-1}}{m} + i\frac{\sqrt{d}}{lm}$, and we consider the $\tilde{z_I} = \frac{nt^{-1}}{m} + i\frac{lm}{\sqrt{d}}$ given under eq. (2.19) . So now the problem simplifies to a problem counting solutions to a diophantine equation under certain conditions. More specifically, we now count the integer solutions to the following,

$$lm^k = l^2n^2 + t^2d \tag{2.20}$$

such that $l \mid d$, $(m, dn) = 1$, $l, m, t > 0$, $lm \leq v\sqrt{d}$ and nt^{-1} modulo m, when lifted to $\left[-\frac{1}{2}m, \frac{1}{2}m\right)$ lies in $[0, um]$. We also bound $d < X$ with a little abuse of notation (we were using X to bound $D = 4d$ prior to this point). Let us define the number of solutions to the above diophantine equation as $S(k, X, v, u)$.

We first find bounds for each of l, m, n, t. Using the condition that $lm \leq v\sqrt{d}$ we have,

$$\begin{aligned} lm \leq v\sqrt{d} &= v\sqrt{\frac{lm^k - l^2n^2}{t^2}} \\ \Rightarrow n^2 &\leq \frac{1}{l^2}(lm^k - \frac{l^2m^2t^2}{v^2}) \end{aligned}$$

Then bounding d,

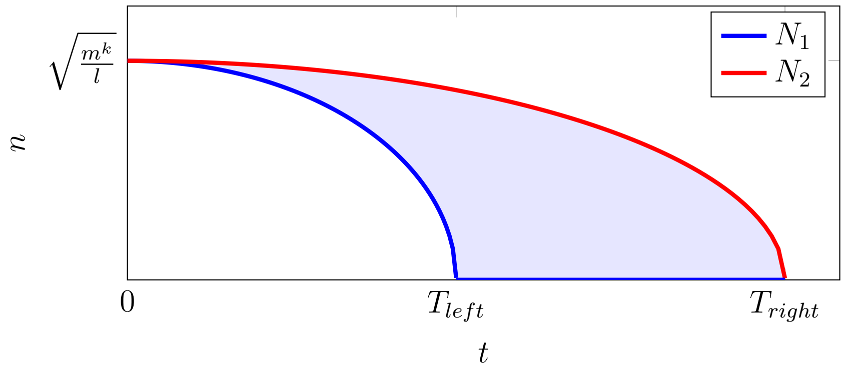

Figure 2.4: Domain of $|n|$ with respect to t for fixed l, m

$$\frac{lm^k - l^2n^2}{t^2} = d < X$$
$$\Rightarrow n^2 > \frac{1}{l^2}(lm^k - t^2X)$$

Now define,

$$N_1 = N_1(l, m, t) = \sqrt{\max\left(\frac{1}{l^2}(lm^k - t^2X),\ 0\right)} \tag{2.21}$$
$$N_2 = N_2(l, m, t) = \sqrt{\frac{1}{l^2}\left(lm^k - \frac{l^2m^2}{v^2}t^2\right)} \tag{2.22}$$

Then we have that $|n|$ is bounded between N_1 and N_2, $n \in (N_1, N_2] \cup [-N_2, -N_1)$. Define the following,

$$T_{left} = T_{left}(l, m) = \sqrt{\frac{lm^k}{X}} \tag{2.23}$$
$$T_{right} = T_{right}(l, m) = \sqrt{\frac{v^2m^{k-2}}{l}} \tag{2.24}$$

Hence, we have a domain for n in terms of $N_1, N_2, T_{left}, T_{right}$ as shown in fig. 2.4.

Also we have,

$$\begin{aligned}
lm \le v\sqrt{d} \le v\sqrt{X} \\
\Rightarrow m < \frac{1}{l} v\sqrt{X} \\
\Rightarrow l < v\sqrt{X}
\end{aligned}$$

Set,

$$M = M(l) = \frac{1}{l} v\sqrt{X} \tag{2.25}$$

$$L = v\sqrt{X} \tag{2.26}$$

Let us define,

$$Res(u, m) = \left\{ x \in \mathbb{Z} \ ; \ x \mod m \text{ when lifted to } \left[-\frac{1}{2}m, \frac{1}{2}m \right) \text{ lies in } [0, mu] \right\} \tag{2.27}$$

For co-prime integers c and $q > 1$ define $\rho_c(q)$ to be the number of unique solutions x

mod q to $c = x^2 \mod q$. Now for prime powers we have the following.

$$\rho_c(p^s) = \begin{cases} 0 & gcd(c,p) > 1 \\ 1 & s = 0 \\ 1 + \left(\frac{c}{p}\right) & p \nmid c,\ s \geq 1 \\ 1 & p = 2,\ s = 1,\ c \equiv 1 \mod 2 \\ 2 & p = 2,\ s = 2,\ c \equiv 1 \mod 4 \\ 4 & p = 2,\ s \geq 3,\ c \equiv 1 \mod 8 \end{cases} \tag{2.28}$$

It is easy to see by CRT that ρ_c is multiplicative, and hence we extend $\rho_c(q)$ for all positive integers q by multiplication. We will see later the exceptional behaviour at powers of 2 will add more technicalities to the proof.

We have the following rather complicated nested sum for $S(k, X, v, u)$,

$$S(k, X, v, u) = \sum_{\substack{l=1 \\ \text{square free}}}^{L} \sum_{\substack{m=1 \\ (2l,m)=1}}^{M} \sum_{\substack{t=1 \\ (lm,t)=1}}^{T_{right}} \sum_{\substack{|n| \in [N_1, N_2] \\ (m,n)=1 \\ nt^{-1} \in Res(u,m) \\ 2t^2 | lm^k - l^2 n^2}} \sum_{r^2 | \frac{lm^k - l^2 n^2}{t^2}} \mu(r) \tag{2.29}$$

Note that the map $c \mapsto \sum_{r^2|c} \mu(r)$, is actually the indicator function for square-free integers. Hence the last sum ensures that we are counting over fundamental discriminants. We treat the two cases $2 \mid l$ and $2 \nmid l$ separately. We will work the $2 \mid l$ case in detail first and later discuss the other case. Both cases are mostly identical. We set $S(k, X, v, u) = S_0(k, X, v, u) + S_1(k, X, v, u)$, where S_i corresponds to $l = i \mod 2$ case. Note that we can replace the condition $(2l, m) = 1$ by $(l, m) = 1$ in the m-sum, $2t^2 \mid lm^k - l^2n^2$ by $t^2 \mid lm^k - l^2n^2$ in the n-sum and most importantly can restrict r and t to be odd.

We resume by considering $S_0(k, X, v, u)$ as explained. Adapting eq. (2.29) to the new conditions and by changing order of summation of the last two nested sums we have,

$$S_0(k,X,v,u) = \sum_{\substack{l=1 \\ \text{square free} \\ 2|l}}^{L} \sum_{\substack{m=1 \\ (l,m)=1}}^{M} \sum_{\substack{t=1 \\ (lm,t)=1}}^{T_{right}} \sum_{\substack{r \\ r \text{ odd}}} \mu(r) \sum_{\substack{|n|\in[N_1,N_2] \\ (m,n)=1 \\ nt^{-1}\in Res(u,m) \\ lm^k=l^2n^2 \mod r^2t^2}} 1 \tag{2.30}$$

At this point we move away from computing $S(k,X,v,u)$. Alternatively, we compute $\tilde{S}(k,X,v,u)$, which we will describe shortly and show that $\tilde{S}$ is consistent with our expectations. We will call $\tilde{S}$ the heuristic version of S, equivalently the heuristic versions of S_0 and S_1 will be denoted by $\tilde{S}_0$ and $\tilde{S}_1$. $\tilde{S}_0(k,X,v,u)$ will be as follows,

$$\tilde{S}_0(k,X,v,u) = \sum_{\substack{l=1 \\ \text{square free} \\ 2|l}}^{L} \sum_{\substack{m=1 \\ (l,m)=1}}^{M} \sum_{\substack{t=1 \\ (lm,t)=1}}^{T_{right}} \sum_{\substack{r \\ r \text{ odd}}} 2\frac{N_2-N_1}{r^2t^2}\rho_{lm}(r^2t^2)u\frac{\phi(m)}{m} \tag{2.31}$$

Note that S_0, differs from $\tilde{S}_0$ only from the innermost sum. The rationale behind $\tilde{S}_0$ is as thus. Based on the coprimality conditions we deduce that $gcd(m, r^2t^2) = 1$, hence by CRT these local conditions modulo each of them be treated as 'independent events'. First there are $2\frac{N_2-N_1}{r^2t^2}\rho_{lm}(r^2t^2)$ number of n satisfying the condition $lm^k \equiv l^2n^2 \mod r^2t^2$. Relatively prime integers modulo an integer m is equidistributed [29, p. 23] with error $O(m^\epsilon)$, hence for a fixed t asymptotically there will be $u\phi(m)$ number of n relatively prime to m such that $nt^{-1} \in Res(u,m)$. As such we replace the inner most sum in eq. (2.30) with $2\frac{N_2-N_1}{r^2t^2}\rho_{lm}(r^2t^2)u\frac{\phi(m)}{m}$ and defined $\tilde{S}_0$. $_1$ shall be defined based on the same rationale and hence $\tilde{S}$ defined as $\tilde{S}_0 + \tilde{S}_1$.

Now we have the following for the r-sum in eq. (2.31),

$$\sum_{r \text{ odd}} \mu(r) 2 \frac{N_2 - N_1}{r^2 t^2} \rho_{lm}(r^2 t^2) u \frac{\phi(m)}{m} = 2u \frac{N_2 - N_1}{t^2} \rho_{lm}(t^2) \frac{\phi(m)}{m} \sum_{r \text{ odd}} \mu(r) \frac{\rho_{lm}(\frac{r^2}{(r^2,t^2)})}{r^2}$$
$$= 2u \frac{N_2 - N_1}{t^2} \rho_{lm}(t) \frac{\phi(m)}{m} \prod_{\substack{p \text{ prime} \\ p>2}} \left(1 - \frac{\rho_{lm}\left(\frac{p}{(p,t)}\right)}{p^2}\right)$$

To proceed further we state the following theorem in the spirit of [26, Lemma 2]. This lemma helps get rid of the ρ_{lm} terms, by averaging them over t and lm and invoking the Polya-Vinogradov inequality. We prove this lemma in a later section.

Lemma 2.4.10. *Let t_* be the odd part of t. Then,*

$$\sum_{\substack{l=1 \\ \text{square free} \\ 2|l}}^{L} \sum_{\substack{m=1 \\ (l,m)=1}}^{M} \frac{\phi(m)}{m} \sum_{\substack{t=1 \\ (lm,t)=1}}^{T_{right}} \left(\rho_{lm}(t_*) \prod_{\substack{p \text{ prime} \\ p>2}} \left(1 - \frac{\rho_{lm}\left(\frac{p}{(p,t_*)}\right)}{p^2}\right)\right) \frac{N_2 - N_1}{t^2}$$
$$= \sum_{\substack{l=1 \\ \text{square free} \\ 2|l}}^{L} \sum_{\substack{m=1 \\ (l,m)=1}}^{M} \frac{\phi(m)}{m} \sum_{\substack{t=1 \\ (lm,t)=1}}^{T_{right}} \prod_{\substack{p \text{ prime} \\ p\nmid lm}} \left(1 - \frac{1}{p^2}\right) \frac{N_2 - N_1}{t^2} + O\left(X^{13/16+\epsilon}\right) \tag{2.32}$$

Consider the t-sum in eq. (2.32), the Euler product is independent of t. So we actually need to consider the following sum,

$$F = \sum_{\substack{t=1 \\ (lm,t)=1}}^{T_{right}} \frac{N_2 - N_1}{t^2} \tag{2.33}$$

Set $\gamma = \gamma(l,m) = \frac{lm}{v\sqrt{X}} < 1$ and $t(a) = t = a\left(\frac{lm^k}{X}\right)^{1/2}$ for $a \in (0, \gamma^{-1}]$. Note that $t(1) = T_{left}$ and $t(\gamma^{-1}) = T_{right}$. Define the following,

$$f_A(a) = \frac{\sqrt{1-\gamma^2 a^2} - \sqrt{1-a^2}}{a^2}, a \in (0,1] \tag{2.34}$$
$$f_B(a) = \frac{\sqrt{1-\gamma^2 a^2}}{a^2}, a \in (1, \gamma^{-1}] \tag{2.35}$$

Then,

$$\frac{N_2 - N_1}{t^2} = \begin{cases} \frac{1}{l}\frac{X}{\sqrt{lm^k}} f_A(a) & a \in (0,1] \\ \frac{1}{l}\frac{X}{\sqrt{lm^k}} f_B(a) & a \in (1,\gamma^{-1}] \end{cases} \tag{2.36}$$

The t-sum can be expressed as,

$$\begin{aligned} F &= \frac{1}{l}\frac{X}{\sqrt{lm^k}} \left(\sum_{\substack{t=1 \\ (lm,t)=1}}^{T_{left}} f_A(a) + \sum_{\substack{t=T_{left} \\ (lm,t)=1}}^{T_{right}} f_B(a) \right) \\ &= \frac{1}{l}\frac{X}{\sqrt{lm^k}} \left(\sum_{t=1}^{T_{left}} \mathbb{1}_{lm}(t) f_A(a) + \sum_{t=T_{left}}^{T_{right}} \mathbb{1}_{lm}(t) f_B(a) \right) \\ &= \frac{1}{l}\frac{X}{\sqrt{lm^k}} (F_A + F_B) \end{aligned} \tag{2.37}$$

Where the F_A and F_B are the sub-sums corresponding to f_A and f_B respectively. The a in the sums should be understood to be that corresponding to each t through which we run the sum.

We state and prove the following lemma which we will immediately use for the t-sum and later for the m-sum. The lemma will be useful in case where we want to evaluate a sum over a large range of integers, the summand is the product of two arithmetic functions. The first function depends on the local properties of the summing variable the second depends on the archimedian properties of the summing variable. The idea is that on a 'short but sufficiently large' interval the second function is approximately constant, meanwhile the first function has 'reasonably random behaviour' with a constant asymptotic average. Hence the combined sum average over a 'short but sufficiently large' interval would approximate to the product of the asymptotic of the first function and the (constant) average of the second function.

Lemma 2.4.11. *Let $g : I \to \mathbb{R}_{>0}$ be a monotonic differentiable function on an interval I and $\beta : \mathbb{N} \to \mathbb{R}_{>0}$ an arithmetic function whose arithmetic average is asymptotic to a constant $\tilde{\beta}$,*

$$\frac{1}{\Delta n}\sum_{n=n_1}^{n_1+\Delta n}\beta(n)=\tilde{\beta}+O\left(\Delta n^{-c}\right)$$

for some constants $c>0$.

Define,

$$H=\sum_{n=n_1}^{n_2}\beta(n)g(n)$$

Then for n_2-n_1 sufficiently large,

$$H=(n_2-n_1)\left(\tilde{\beta}+O\left((n_2-n_2)^{-c}\right)\right)g(n_2)+O\left((n_2-n_1)^2\sup|g'(x)|\right)$$

Proof. W.l.o.g we assume g is decreasing.

$$\begin{aligned}H&=\sum_{n=n_1}^{n_2}\beta(n)g(n_2)+\sum_{n=n_1}^{n_2}\beta(n)(g(n)-g(n_2))\\&=(n_2-n_1)\left(\tilde{\beta}+O\left((n_2-n_1)^{-c}\right)\right)g(n_2)+\sum_{n=n_1}^{n_2}\beta(n)(g(n)-g(n_2))\end{aligned}$$

$$\begin{aligned}\sum_{n=n_1}^{n_2}\beta(n)(g(n)-g(n_2))&\leq\sum_{n=n_1}^{n_2}\beta(n)(g(n_1)-g(n_2))\\&=(n_2-n_1)\left(\tilde{\beta}+O\left((n_2-n_2)^{-c}\right)\right)(g(n_1)-g(n_2))\\&=(n_2-n_1)^2\left(\tilde{\beta}+O\left((n_2-n_2)^{-c}\right)\right)\frac{g(n_1)-g(n_2)}{n_2-n_1}\\&=O\left((n_2-n_1)^2\sup|g'(x)|\right)\end{aligned}$$

$\square$

Let Δt be such sufficiently large but still $\Delta t \ll T_{left}, T_{right}$. Then set $\Delta t=\Delta a\left(\frac{lm^k}{X}\right)^{1/2}$. From now it is required l,m be such that $\left(\frac{lm^k}{X}\right)^{1/2}>X^\delta$ for some $\delta\geq 0$. Let $\mathcal{E}$ denote the error term due to $\left(\frac{lm^k}{X}\right)^{1/2} \leq X^\delta$, which we will compute later.

For T large enough consider the average,

$$\begin{aligned}
\alpha(l,m,s,T) &= \frac{1}{T} \sum_{\substack{t=s \\ (lm,t)=1}}^{s+T} 1 \\
&= \frac{1}{T} \sum_{t=s}^{s+T} \mathbb{1}_{lm}(t) \\
&= \frac{\phi(lm)}{lm} + O\left(T^{-1}(lm)^\epsilon\right) \qquad (2.38)
\end{aligned}$$

The last line follows from [29, p. 23]. We consider F_A as follows,

$$\begin{aligned}
F_A &= \sum_{t=1}^{T_{left}} \mathbb{1}_{lm}(t) f_A(a) \\
&= \sum_{t=1}^{T_1} \mathbb{1}_{lm}(t) f_A(a) + \sum_{t=T_1}^{T_2} \mathbb{1}_{lm}(t) f_A(a) + \sum_{t=T_2}^{T_{left}} \mathbb{1}_{lm}(t) f_A(a) \qquad (2.39) \\
&= F_{A,1} + F_{A,2} + F_{A,3}
\end{aligned}$$

Here we choose $a_1 \sim 1$ and $a_2 \sim 1$ (to be specified later) such that $T_1 = t(a_1)$ and $T_2 = t(a_2)$.

We first consider $F_{A,2}$ and proceed by applying lemma 2.4.11,

$$\begin{aligned}
F_{A,2} &= \sum_{t=T_1}^{T_2} \mathbb{1}_{lm}(t) f_A(a) \\
&= \sum_{h=0}^{\frac{T_2-T_1}{\Delta t}} \sum_{t=T_1+h\Delta t}^{T_1+(h+1)\Delta t} \mathbb{1}_{lm}(t) f_A(a) \\
&= \sum_{h=0}^{\frac{T_2-T_1}{\Delta t}} \left(\left(\frac{\phi(lm)}{lm} + \frac{1}{\Delta t} O\left((lm)^\epsilon\right) \right) \Delta t f_A(a_1 + (h+1)\Delta a) + O\left(\Delta t^2 \sup|\partial_t f_A(x)|\right) \right) \\
&= \left(\frac{\phi(lm)}{lm} + \left(\frac{X}{lm^k}\right)^{1/2} \frac{1}{\Delta a} O\left((lm)^\epsilon\right) \right) \sum_{h=0}^{\frac{T_2-T_1}{\Delta t}} \Delta t f_A(a_1 + (h+1)\Delta a) \\
&+ O\left((T_2 - T_1)\Delta t \sup|\partial_t f_A(x)|\right)
\end{aligned} \tag{2.40}$$

Computing the derivative of f with respect to t we get the following,

$$\partial_t f_A(a) = \left(\frac{X}{lm^k}\right)^{1/2} \left(\frac{2-a^2}{a^3\sqrt{1-a^2}} - \frac{2-\gamma^2 a^2}{a^3\sqrt{1-\gamma^2 a^2}} \right) > 0 \tag{2.41}$$

When $a \sim 1$,

$$\partial_t f_A(a) \leq \left(\frac{X}{lm^k}\right)^{1/2} 2 \frac{1}{\sqrt{1-a^2}} \leq 2 \left(\frac{X}{lm^k}\right)^{1/2} \frac{1}{\sqrt{1-a^2}}$$

When $a \sim 0$, using the expansion $\frac{1}{\sqrt{1-x^2}} = 1 + \frac{1}{2}x^2 + O\left(x^4\right)$,

$$\begin{aligned}
\partial_t f_A(a) &= O\left(\left(\frac{X}{lm^k}\right)^{1/2} \frac{1}{a^3} \left((2-a^2)\left(1 + \frac{1}{2}a^2\right) - (2-\gamma^2 a^2)\left(1 + \frac{1}{2}\gamma^2 a^2\right) \right) \right) \\
&= O\left(\left(\frac{X}{lm^k}\right)^{1/2} (1-\gamma^4) a \right)
\end{aligned}$$

Putting these together we have,

$$
\begin{aligned}
&O\left((T_2 - T_1)\Delta t \sup|\partial_t f_A(x)|\right) \\
&= O\left(\left(\frac{lm^k}{X}\right)^{1/2}\left(\Delta a\left(\frac{lm^k}{X}\right)^{1/2}\right)\left(\frac{X}{lm^k}\right)^{1/2}\frac{1}{\sqrt{1-a_2^2}}\right) \\
&= O\left(\left(\frac{lm^k}{X}\right)^{1/2}\frac{\Delta a}{\sqrt{1-a_2^2}}\right)
\end{aligned}
$$

We now approximate the main term in $F_{A,2}$ using an integral,

$$
\begin{aligned}
&\sum_{h=0}^{\frac{T_2-T_1}{\Delta t}} \Delta t f_A(a_1 + (h+1)\Delta a) \\
&= \sum_{h=0}^{\frac{a_2-a_1}{\Delta a}} \left(\left(\frac{lm^k}{X}\right)^{1/2}\Delta a\right) f_A(a_1 + (h+1)\Delta a) \\
&= \left(\frac{lm^k}{X}\right)^{1/2} \sum_{h=0}^{\frac{a_2-a_1}{\Delta a}} \Delta a f_A(a_1 + (h+1)\Delta a) \\
&= \left(\frac{lm^k}{X}\right)^{1/2}\left(\int_{a_1}^{a_2} f_A(a)da + O\left(\Delta a \sup|f_A'(a)|\right)\right) \\
&= \left(\frac{lm^k}{X}\right)^{1/2}\left(\int_0^1 f_A(a)da - \int_0^{a_1} f_A(a)da - \int_{a_2}^1 f_A(a)da + O\left(\Delta a \sup|f_A'(a)|\right)\right)
\end{aligned}
$$

As earlier we can show that,

$$
\sup|f_A'(a)| = O\left(\frac{\Delta a}{\sqrt{1-a_2^2}}\right)
$$

Putting these in eq. (2.40) and then all that in eq. (2.39) we get,

$$F_A = \frac{\phi(lm)}{lm}\left(\frac{lm^k}{X}\right)^{1/2}\int_0^1 f_A(a)da + \left(\sum_{t=1}^{T_1}\mathbb{1}_{lm}(t)f_A(a) - \frac{\phi(lm)}{lm}\left(\frac{lm^k}{X}\right)^{1/2}\int_0^{a_1} f_A(a)da\right)$$
$$+ \left(\sum_{t=T_2}^{T}\mathbb{1}_{lm}(t)f_A(a) - \frac{\phi(lm)}{lm}\left(\frac{lm^k}{X}\right)^{1/2}\int_{a_2}^{1} f_A(a)da\right) + O\left(\frac{1}{\Delta a}(lm)^{\epsilon}\right)$$
$$+ O\left(\left(\frac{lm^k}{X}\right)^{1/2}\frac{\Delta a}{\sqrt{1-a_2^2}}\right) \tag{2.42}$$

Since $a \sim 0$ in both the sum and integral in the second term of eq. (2.42), we use the expansion $\sqrt{1-x^2} = 1 - \frac{1}{2}x^2 + O(x^4)$ in $f_A(a)$ and it shows that this second term is $O\left(\left(\frac{lm^k}{X}\right)^{1/2}(1-\gamma)a_1\right)$. In the third term $a \sim 1$ so $f_A(a)$ is bounded, we have that the third term is $O\left(\left(\frac{lm^k}{X}\right)^{1/2}(1-a_2)\right)$. Hence, the second, third and last terms are,

$$O\left(\left(\frac{lm^k}{X}\right)^{1/2}(1-\gamma)a_1\right) + O\left(\left(\frac{lm^k}{X}\right)^{1/2}(1-a_2)\right)$$
$$+ O\left(\left(\frac{lm^k}{X}\right)^{1/2}\frac{\Delta a}{\sqrt{1-a_2^2}}\right) \tag{2.43}$$

We can have a_1 very small, but setting $(1-a_2) = \frac{\Delta a}{\sqrt{1-a_2^2}}$ we have $1 - a_2 \sim \Delta a^{2/3}$. Hence, eq. (2.43) is $O\left(\left(\frac{lm^k}{X}\right)^{1/2}\Delta a^{2/3}\right)$. The error terms in eq. (2.42) reduces to,

$$O\left(\frac{1}{\Delta a}(lm)^{\epsilon}\right) + O\left(\left(\frac{lm^k}{X}\right)^{1/2}\Delta a^{2/3}\right)$$

Setting Δa to,

$$\Delta a = (lm)^{3\epsilon/5}\left(\frac{X}{lm^k}\right)^{3/10} \tag{2.44}$$

we get,

$$F_A = \frac{\phi(lm)}{lm}\left(\frac{lm^k}{X}\right)^{1/2}\int_0^1 f_A(a)da + O\left((lm)^{2\epsilon/5}\left(\frac{lm^k}{X}\right)^{3/10}\right) \tag{2.45}$$

With similar treatment to F_B using lemma 2.4.11 and optimizing errors we get,

$$F_B = \frac{\phi(lm)}{lm}\left(\frac{lm^k}{X}\right)^{1/2}\int_1^{\gamma^{-1}} f_B(a)da + O\left((lm)^{2\epsilon/5}\left(\frac{lm^k}{X}\right)^{3/10}\right) \tag{2.46}$$

Then we have,

$$\begin{aligned} F_A + F_B = &\frac{\phi(lm)}{lm}\left(\frac{lm^k}{X}\right)^{1/2}\left(\int_0^1 f_A(a)da + \int_1^{\gamma^{-1}} f_B(a)da\right) \\ &+ O\left((lm)^{2\epsilon/5}\left(\frac{lm^k}{X}\right)^{3/10}\right) \end{aligned} \tag{2.47}$$

It is easy but a little tedious to prove that,

$$\int_0^1 f_A(a)da + \int_1^{\gamma^{-1}} f_B(a)da = \frac{\pi}{2}(1-\gamma)$$

Little explanation on computing the integrals; break the first integral as $\int_0^h + \int_h^1$, where $h > 0$ is arbitrarily small. Compute the first part using the expansion $\sqrt{1-x^2} = 1 - \frac{1}{2}x^2 + O(x^4)$ in $f_A(a)$ and then $h \to 0$ for $\int_0^h + \int_h^1$.

Putting this in eq. (2.47) and then putting all that in eq. (2.37) we have,

$$F = \frac{\pi}{2}\frac{X^{1/2}}{l}(1-\gamma)\frac{\phi(lm)}{lm} + O\left(\frac{1}{l}(lm)^{2\epsilon/5}\frac{X^{7/10}}{(lm^k)^{1/5}}\right) \tag{2.48}$$

We now define the following,

$$\alpha_l(m) = \left(\frac{\phi(m)}{m}\right)^2 \prod_{\substack{p \text{ prime} \\ p \nmid lm}} \left(1 - \frac{1}{p^2}\right) \mathbb{1}_l(m) \tag{2.49}$$

Now putting eq. (2.48) in eq. (2.31) using lemma 2.4.10 we have,

$$\tilde{S}_0(k, X, v, u) = u\pi X^{1/2} \sum_{\substack{l=1 \\ \text{square free} \\ 2|l}}^{L} \frac{1}{l}\frac{\phi(l)}{l} \sum_{m=1}^{M} \alpha_l(m)(1 - \gamma(l, m))$$
$$+ \sum_{\substack{l=1 \\ \text{square free} \\ 2|l}}^{L} \frac{1}{l^2}\frac{\phi(l)}{l} \sum_{m=1}^{M} O\left((lm)^{2\epsilon/5}\frac{X^{7/10}}{(lm^k)^{1/5}}\right)$$
$$+\mathcal{E} + O\left(X^{13/16+\epsilon}\right) \tag{2.50}$$

Note that,

$$\sum_{\substack{l=1 \\ \text{square free} \\ 2|l}}^{L} \frac{1}{l^2}\frac{\phi(l)}{l} \sum_{m=1}^{M} O\left((lm)^{2\epsilon/5}\frac{X^{7/10}}{(lm^k)^{1/5}}\right) = O\left(X^{7/10}\int_1^M m^{-(k/5-2\epsilon/5)}dm\right)$$
$$+ \begin{cases} O\left(X^{9/10+\epsilon}\right) & k = 3 \\ O\left(X^{7/10+\epsilon}\right) & k = 5 \\ O\left(X^{7/10}\right) & k > 5 \end{cases} \tag{2.51}$$

Define the following,

$$\alpha(l, s, Y) = \frac{1}{Y}\sum_{m=s}^{s+Y} \alpha_l(m) \tag{2.52}$$

Lemma 2.4.12.

$$\alpha(l, s, Y) = \frac{\phi(l)}{l}\prod_{p\nmid l}\left(1 - \frac{1}{p}\right)^2\left(1 + \frac{2}{p}\right) + O\left(Y^{-1}l^{\epsilon}(s + Y)^{\epsilon}\right) \tag{2.53}$$

For convenience we set $\alpha(l) = \frac{\phi(l)}{l}\prod_{p\nmid l}\left(1 - \frac{1}{p}\right)^2\left(1 + \frac{2}{p}\right)$. Note that $\left(1 - \frac{1}{p}\right)^2\left(1 + \frac{2}{p}\right) = 1 - \frac{3}{p^2} + \frac{2}{p^3} = 1 + O\left(p^{-2}\right)$, hence $\alpha(l)$ is well defined. We'll see the proof of this lemma in a later section.

To find the m-sum in eq. (2.50), we use lemma 2.4.11 as in the t-sums. Choose large enough Δm, set $a(m) = a = \frac{l}{vX^{1/2}}m$, hence $m = aM$ and $\Delta m = \Delta a M$, for appropriate Δa. Note that $\gamma(l,m) = c(m)$ and recall from eq. (2.25) that $M = \frac{vX^{1/2}}{l}$. So using lemma 2.4.11 and replacing Riemann sums by appropriate integral we have for the m-sum,

$$\begin{aligned}\sum_{m=1}^{M} \alpha_l(m)(1-\gamma(l,m)) &= M\left(\alpha(l) + O\left(M^{\epsilon}\Delta m^{-1}\right)\right)\int_0^1 (1-c)dc + O(M\Delta m \sup|\partial_m \gamma|) \\ &= \frac{1}{2}M\alpha(l) + O\left(M^{1+\epsilon}\Delta m^{-1}\right) + O\left(\Delta m\right) \\ &= \frac{1}{2}M\alpha(l) + O\left(M^{(1+\epsilon)/2}\right) \qquad (2.54)\end{aligned}$$

where the last line follows by setting by setting $\Delta m = M^{(1+\epsilon)/2}$. Now setting $M = \frac{vX^{1/2}}{l}$ in eq. (2.54) and then putting it in eq. (2.50) we have,

$$\tilde{S}_0(k,X,v,u) = \frac{\pi}{2}uvX\left(\sum_{\substack{l=1 \\ \text{square free} \\ 2|l}}^{L} \frac{1}{l^2}\frac{\phi(l)}{l}\alpha(l)\right) + \mathcal{E} + \begin{cases} O\left(X^{9/10+\epsilon}\right) & k=3 \\ O\left(X^{13/16+\epsilon}\right) & k>3 \end{cases} \qquad (2.55)$$

Now we have the following,

$$
\begin{aligned}
\alpha &= \sum_{\substack{l=1\\ \text{square free}\\ 2|l}}^{L} \frac{1}{l^2}\frac{\phi(l)}{l}\alpha(l)\\
&= \sum_{\substack{l=1\\ \text{square free}\\ 2|l}}^{L} \frac{1}{l^2}\left(\frac{\phi(l)}{l}\right)^2 \prod_{p\nmid l}\left(1-\frac{1}{p}\right)^2\left(1+\frac{2}{p}\right)\\
&= \prod_{p}\left(1-\frac{1}{p}\right)^2\left(1+\frac{2}{p}\right) \sum_{\substack{l=1\\ \text{square free}\\ 2|l}}^{L} \frac{1}{l^2}\left(\frac{\phi(l)}{l}\right)^2 \frac{1}{\prod_{p|l}\left(1-\frac{1}{p}\right)^2\left(1+\frac{2}{p}\right)}\\
&= \prod_{p}\left(1-\frac{1}{p}\right)^2\left(1+\frac{2}{p}\right) \sum_{\substack{l=1\\ \text{square free}\\ 2|l}}^{L} \frac{1}{l^2}\frac{1}{\prod_{p|l}\left(1+\frac{2}{p}\right)}\\
&= \prod_{p}\left(1-\frac{1}{p}\right)^2\left(1+\frac{2}{p}\right) \frac{1}{8} \sum_{\substack{l=1\\ \text{square free}\\ 2\nmid l}}^{L/2} \frac{1}{l^2}\frac{1}{\prod_{p|l}\left(1+\frac{2}{p}\right)}\\
&= \frac{1}{2}\prod_{p>2}\left(1-\frac{1}{p}\right)^2\left(1+\frac{2}{p}\right)\frac{1}{8}\left(\prod_{p>2}\left(1+\frac{1}{p^2}\frac{1}{1+\frac{2}{p}}\right)+O\left(L^{-1+\epsilon}\right)\right)\\
&= \frac{1}{16}\prod_{p>2}\left(1-\frac{1}{p}\right)^2\left(1+\frac{2}{p}\right)\prod_{p>2}\frac{\left(1+\frac{1}{p}\right)^2}{\left(1+\frac{2}{p}\right)}+O\left(L^{-1+\epsilon}\right)\\
&= \frac{1}{16}\prod_{p>2}\left(1-\frac{1}{p^2}\right)^2+O\left(L^{-1+\epsilon}\right)\\
&= \frac{1}{9}\left(\prod_{p}\left(1-\frac{1}{p^2}\right)\right)^2+O\left(L^{-1+\epsilon}\right)\\
&= \frac{1}{9}\frac{1}{\zeta(2)^2}+O\left(L^{-1}\right)=\frac{4}{\pi^4}+O\left(X^{-1/2+\epsilon}\right)
\end{aligned}
$$

We now estimate $\mathcal{E}$, the error term due to (l,m) such that $\left(\frac{lm^k}{X}\right)^{1/2}\leq X^{\delta}$ for some $\delta\geq 0$.

To compute this error we replace M by $\tilde{M}=\left(\frac{X^{2\delta+1}}{l}\right)^{1/k}$ and L by $\tilde{L}=\left(X^{2\delta+1}\right)^{1/k}$. Using methods similar as before, with the substitution $t=a\left(\frac{lm^k}{X}\right)^{1/2}$ we have,

$$\mathcal{E} = c' X^{1/2} \sum_{l}^{\tilde{L}} \frac{1}{l} \sum_{m=1}^{\tilde{M}} \left(\sum_{a=0}^{1} \frac{\sqrt{1-r^2a^2} - \sqrt{1-a^2}}{a^2} \Delta a + \sum_{a=1}^{1/r} \frac{\sqrt{1-r^2a^2}}{a^2} \Delta a \right)$$

for some constant c'. The a sums are of the order 1, hence we have,

$$\begin{aligned} \mathcal{E} &= c'' X^{1/2} \sum_{l} \frac{1}{l} \sum_{m=1}^{\tilde{M}} 1 \\ &= c_0 X^{1/2+1/k+2\delta/k} \end{aligned}$$

for some constant c. Since, we subtract $\mathcal{E}$ from the initial sum, $c < 0$. Recall that the purpose of such a δ is to maintain a sufficiently large $\Delta t = \Delta a \left(\frac{lm^k}{X}\right)^{1/2}$ in the calculations. By our choice of Δa in eq. (2.44), we have $\Delta t = (lm)^{3\epsilon/5} \left(\frac{lm^k}{X}\right)^{1/5} > (lm)^{3\epsilon/5} X^{2\delta/5}$, and it is sufficient to have $\Delta t = o\left(lm^{3\epsilon/5}\right)$, so we set $\delta = 0$. Hence,

$$\mathcal{E} = c_0 X^{1/2+1/k} \tag{2.56}$$

So putting this in eq. (2.55) we have,

$$\tilde{S}_0(k, X, v, u) = \frac{2}{\pi^3} uvX + c_0 X^{1/2+1/k} + \begin{cases} O\left(X^{9/10+\epsilon}\right) & k = 3 \\ O\left(X^{13/16+\epsilon}\right) & k > 3 \end{cases} \tag{2.57}$$

Now that we have discussed the case $2 \mid l$, let us see what changes we need to do in the case of $2 \nmid l$. We rewrite eq. (2.29) for $S_1(k, X, v, u)$ as,

$$S_1(k, X, v, u) = \sum_{\substack{l=1 \\ \text{square free} \\ 2 \nmid l}}^{L} \sum_{\substack{m=1 \\ (2l,m)=1}}^{M} \sum_{\substack{t=1 \\ (lm,t)=1}}^{T_{right}} \sum_{\substack{|n| \in [N_1, N_2] \\ (m,n)=1 \\ nt^{-1} \in Res(u,m) \\ 2t^2 | lm^k - l^2n^2}} \sum_{r^2 | \frac{lm^k - l^2n^2}{t^2}} \mu(r) \tag{2.58}$$

In order to enforce the fact that $2 \mid\mid \frac{lm^k - l^2n^2}{t^2}$ we write the inner most sums as follows,

$$\left(\sum_{\substack{r \\ r \text{ odd}}} \mu(r) \sum_{\substack{|n|\in[N_1,N_2] \\ (m,n)=1 \\ nt^{-1}\in Res(u,m) \\ lm^k = l^2n^2 \mod 2r^2t^2}} 1 \right) - \left(\sum_{\substack{r \\ r \text{ odd}}} \mu(r) \sum_{\substack{|n|\in[N_1,N_2] \\ (m,n)=1 \\ nt^{-1}\in Res(u,m) \\ lm^k = l^2n^2 \mod 4r^2t^2}} 1 \right)$$

Proceeding as earlier, adopting the heuristic model $\tilde{S}_1$ for S_1 we have,

$$\begin{aligned}
\tilde{S}_1(k,X,v,u) &= 2u \sum_{\substack{l=1 \\ \text{square free} \\ 2\nmid l}}^{L} \sum_{\substack{m=1 \\ (2l,m)=1}}^{M} \frac{\phi(m)}{m} \left(\frac{1}{2} \sum_{\substack{t=1 \\ (lm,t)=1}}^{T_{right}} \rho_{lm}(2t^2) \sum_{\substack{r \\ r \text{ odd}}} \frac{N_2-N_1}{r^2t^2} \rho_{lm}\left(\frac{r}{(r,t)}\right) \right. \\
&\qquad \left. - \frac{1}{4} \sum_{\substack{t=1 \\ (lm,t)=1}}^{T_{right}} \rho_{lm}(2t^2) \sum_{\substack{r \\ r \text{ odd}}} \frac{N_2-N_1}{r^2t^2} \rho_{lm}\left(\frac{r}{(r,t)}\right) \right) \\
&= 2u \sum_{\substack{l=1 \\ \text{square free} \\ 2\nmid l}}^{L} \sum_{\substack{m=1 \\ (2l,m)=1}}^{M} \frac{\phi(m)}{m} \left(\frac{1}{2} \sum_{\substack{t=1 \\ (lm,t)=1}}^{T_{right}} \rho_{lm}(2t^2) B_t - \frac{1}{4} \sum_{\substack{t=1 \\ (lm,t)=1}}^{T_{right}} \rho_{lm}(4t^2) B_t \right)
\end{aligned} \tag{2.59}$$

where $B_t = \sum_{r \text{ odd}} \frac{N_2-N_1}{r^2t^2} \rho_{lm}\left(\frac{r}{(r,t)}\right)$. For convenience we write the t-sum in eq. (2.59) as follows,

$$H(l,m) = \frac{1}{2} \sum_{t} \rho_{lm}(2t^2) B_t - \frac{1}{4} \sum_{t} \rho_{lm}(4t^2) B_t \tag{2.60}$$

Splitting the t-sum as odd and even, we compute $H(l,m)$ in the following cases separately,

1. $lm \equiv -1 \mod 4$

$$\begin{aligned}
H(l,m) &= \frac{1}{2} \left(\sum_{t \text{ odd}} \rho_{lm}(2t^2) B_t + \sum_{t \text{ even}} \rho_{lm}(2t^2) B_t \right) - \frac{1}{4} \left(\sum_{t \text{ odd}} \rho_{lm}(4t^2) B_t + \sum_{t \text{ even}} \rho_{lm}(4t^2) B_t \right) \\
&= \frac{1}{2} \left(\sum_{t \text{ odd}} \rho_{lm}(t_*) B_t + 0 \right) - (0+0) \\
&= \frac{1}{2} \sum_{t \text{ odd}} \rho_{lm}(t_*) B_t
\end{aligned}$$

2. $lm \equiv 1 \mod 4$ but $lm \not\equiv 1 \mod 8$

$$
\begin{aligned}
H(l,m) &= \frac{1}{2}\left(\sum_{t \text{ odd}} \rho_{lm}(t_*)B_t + 0\right) - \frac{1}{4}\left(\sum_{t \text{ odd}} 2\rho_{lm}(t_*)B_t + 0\right) \\
&= 0
\end{aligned}
$$

3. $lm \equiv 1 \mod 8$

$$
\begin{aligned}
H(l,m) &= \frac{1}{2}\left(\sum_{t \text{ odd}} \rho_{lm}(t_*)B_t + \sum_{t \text{ even}} 4\rho_{lm}(t_*)B_t\right) - \frac{1}{4}\left(\sum_{t \text{ odd}} 2\rho_{lm}(t_*)B_t + \sum_{t \text{ even}} 4\rho_{lm}(t_*)B_t\right) \\
&= \sum_{t \text{ even}} \rho_{lm}(t_*)B_t
\end{aligned}
$$

Using this we can rewrite eq. (2.59) as,

$$
\begin{aligned}
\tilde{S}_1(k,X,v,u) = 2u &\sum_{\substack{l=1 \\ \text{square free} \\ 2\nmid l}}^{L} \sum_{\substack{m=1 \\ (2l,m)=1 \\ lm\equiv -1 \mod 4}}^{M} \frac{\phi(m)}{m}\frac{1}{2} \sum_{\substack{t=1 \\ (lm,t)=1 \\ t \text{ odd}}}^{T_{right}} \rho_{lm}(t_*)B_t \\
&+ 2u \sum_{\substack{l=1 \\ \text{square free} \\ 2\nmid l}}^{L} \sum_{\substack{m=1 \\ (2l,m)=1 \\ lm\equiv 1 \mod 8}}^{M} \frac{\phi(m)}{m} \sum_{\substack{t=1 \\ (lm,t)=1 \\ t \text{ even}}}^{T_{right}} \rho_{lm}(t_*)B_t
\end{aligned} \tag{2.61}
$$

The computation of $\tilde{S}_1$ is almost identical to $\tilde{S}_0$ as we have seen, so by similar computations we obtain,

$$
\tilde{S}_1(k,X,v,u) = \frac{4}{\pi^3}uvX + c_1X^{1/2+1/k} + \begin{cases} O\left(X^{9/10+\epsilon}\right) & k=3 \\ O\left(X^{13/16+\epsilon}\right) & k>3 \end{cases} \tag{2.62}
$$

$$
\tilde{S}(k,X,v,u) = \frac{6}{\pi^3}uvX + cX^{1/2+1/k} + \begin{cases} O\left(X^{9/10+\epsilon}\right) & k=3 \\ O\left(X^{13/16+\epsilon}\right) & k>3 \end{cases} \tag{2.63}
$$

We replace X by $\frac{X}{4}$ as adjustment for using $d < X$, instead of $D = 4d < X$ in the proof. We have $\frac{1}{2\pi^2}X + O\left(X^{1/2}\right) = \frac{1}{6}|^{-}\mathcal{D}_X|$ number of case 1 discriminants in $^{-}\mathcal{D}_X$ and for our choice of $\Omega \in \mathcal{F}_0$, $\nu(\omega) = \frac{uv}{\frac{\pi}{3}}$. Hence, we rewrite the main term in eq. (2.63) as $\nu(\omega)\frac{1}{6}|^{-}\mathcal{D}_X|$. Both case 2 and case 3 discriminants can be dealt in similar methods as shown in this section (see remark below), so we have $\nu(\Omega)\frac{1}{6}|^{-}\mathcal{D}_X|$ and $\nu(\Omega)\frac{2}{3}|^{-}\mathcal{D}_X|$ respectively, with the same error terms. Let $\tilde{H}_X^{\Omega}[k]$ denote the heuristic version of $H_X^{\Omega}[k]$. So we have,

$$\tilde{H}_X^{\Omega}[k] = \nu(\Omega)|^{-}\mathcal{D}_X| + c_k X^{1/2+1/k} + \begin{cases} O\left(X^{9/10+\epsilon}\right) & k = 3 \\ O\left(X^{13/16+\epsilon}\right) & k > 3 \end{cases} \tag{2.64}$$

for some constant $c_k < 0$.

Remark. *The tricky part in cases 2 and 3 are enforcing the appropriate congruence class for $d \mod 4$, we achieve this by restricting l, m, t, n to certain congruence classes $\mod 4$. For example in case 2 ($d \equiv 1 \mod 4$), when $c = 1$ t, l, m, n are odd, then we enforce that $lm \equiv -1 \mod 4 \Rightarrow m \equiv -l \mod 4$. When $c = 1$ there are two cases t odd so l, m odd, n even with $lm \equiv 1 \mod 4 \Rightarrow m \equiv l \mod 4$, and the other case t even which is the laborious case. In this case l, m, n are even, then depending on the maximum power of 2 that divides t, we choose either $lm \equiv 1 \mod 4$ or $lm \equiv 1 \mod 8$ and count the solutions of n that satisfy the congruence which enforces $d \equiv 1 \mod 4$.*

2.4.5 Proofs of lemmas 2.4.10 and 2.4.11

Proof of lemma 2.4.10

Notation wise we use t instead of t_* for convenience. By the multiplicativity of the ρ-function and application of Hensel's lemma we have,

$$\rho_{lm}(t) = 1 + \sum_{\substack{q|t \\ 1<q<Q}} \mu^2(q)\chi_q(lm) + \sum_{\substack{q|t \\ q\geq Q}} \mu^2(q)\chi_q(lm)$$

where χ_q is the Kronecker symbol with modulus q, and Q is a parameter we will determine later. Also,

$$\prod_{\substack{p \text{ prime} \\ p>2}} \left(1 - \frac{\rho_{lm}\left(\frac{p}{(p,t)}\right)}{p^2}\right) = \prod_{\substack{p \text{ prime} \\ p|t \\ p\nmid lm}} \left(1 - \frac{1}{p^2}\right) \prod_{\substack{p \text{ prime} \\ p\nmid lmt}} \left(1 - \frac{1}{p^2} - \frac{\chi_p(lm)}{p^2}\right)$$

$$= \prod_{\substack{p \text{ prime} \\ p\nmid lm}} \left(1 - \frac{1}{p^2}\right) + \sum_{\substack{w>1 \\ (lmt,w)=1}} \left(\prod_{\substack{p \text{ prime} \\ p\nmid lmw}} \left(1 - \frac{1}{p^2}\right)\right) \mu(w) \frac{\chi_w(lm)}{w^2}$$

Together we get,

$$\rho_{lm}(t) \prod_{\substack{p \text{ prime} \\ p>2}} \left(1 - \frac{\rho_{lm}\left(\frac{p}{(p,t)}\right)}{p^2}\right) = \prod_{\substack{p \text{ prime} \\ p\nmid lm}} \left(1 - \frac{1}{p^2}\right)$$

$$+ \sum_{\substack{q|t \\ 1<q<Q}} \mu^2(q)\chi_q(lm) \prod_{\substack{p \text{ prime} \\ p\nmid lm}} \left(1 - \frac{1}{p^2}\right)$$

$$+ \sum_{\substack{q|t \\ 1<q<Q}} \mu^2(q)\chi_q(lm) \sum_{\substack{w>1 \\ (lmt,w)=1}} \left(\prod_{\substack{p \text{ prime} \\ p\nmid lmw}} \left(1 - \frac{1}{p^2}\right)\right) \mu(w) \frac{\chi_w(lm)}{w^2}$$

$$+ \sum_{\substack{q|t \\ q\geq Q}} \mu^2(q)\chi_q(lm) \prod_{\substack{p \text{ prime} \\ p>2}} \left(1 - \frac{\rho_{lm}\left(\frac{p}{(p,t)}\right)}{p^2}\right)$$

Now we can write the left hand sum Σ of eq. (2.32) as, $\Sigma = \Sigma_1 + \Sigma_2 + \Sigma_3 + \Sigma_4$. Here Σ_1 is the main term on the right hand side, so we need to show that the rest of the sum is small. From now on we drop the outer most sum in the variable l for convenience.

We can rewrite Σ_2 as follows after changing the order of summation,

$$\begin{aligned}
\Sigma_2 &= \sum_{\substack{m=1\\(l,m)=1}}^{M} \frac{\phi(m)}{m} \sum_{\substack{t=1\\(lm,t)=1}}^{T_{right}} \sum_{\substack{q|t\\1<q<Q}} \mu^2(q)\chi_q(lm) \prod_{\substack{p \text{ prime}\\p\nmid lm}} \left(1-\frac{1}{p^2}\right) \frac{N_2-N_1}{t^2} \\
&= \sum_{\substack{t=1\\(l,t)=1}}^{T'} \frac{1}{t^2} \sum_{\substack{q|t\\1<q<Q}} \mu^2(q)\chi_q(l) \prod_{\substack{p \text{ prime}\\p\nmid l}} \left(1-\frac{1}{p^2}\right) \sum_{\substack{m=M_1\\(lt,m)=1}}^{M_2} \chi_q(m) \frac{1}{\prod_{\substack{p \text{ prime}\\p|m}} \left(1+\frac{1}{p}\right)} (N_2-N_1)
\end{aligned}$$

where $T' = \gamma^{-1}\sqrt{\frac{v^k X^{(k-2)/2}}{l^{k-1}}}$, $M_1 = \left(\frac{\gamma^2 t^2 X}{l}\right)^{1/k}$ and $M_2 = M = \frac{vX^{1/2}}{l}$.

We estimate Σ_2 by applying partial summation to the inner m-sum. First we need to estimate the sum of $\chi_q(m)$ over m relatively prime to lt. The idea is basically Polya-Vinogradov inequality with some technicalities to handle the coprimality condition. For $1 \le H_1 < H_2$ and square-free q,

$$\begin{aligned}
\left| \sum_{\substack{m=H_1\\(lt,m)=1}}^{H_2} \chi_q(m) \right| &= \left| \sum_{m=H_1}^{H_2} \sum_{c|(lt,m)} \mu(c)\chi_q(m) \right| \\
&= \left| \sum_{c|lt} \mu(c)\chi_q(c) \sum_{a=H_1/c}^{H_2/c} \chi_q(a) \right| \\
&\le \sum_{c|lt} \left| \sum_{a=H_1/c}^{H_2/c} \chi_q(a) \right| \\
&= O\left(\tau(lt)\sqrt{q}\log(q)\right)
\end{aligned}$$

Next we estimate the following sum. The trick is to use a sort of partial summation in multiplicative form,

$$
\begin{aligned}
A(H_1, H_2) = \sum_{\substack{m=H_1 \\ (lt,m)=1}}^{H_2} \frac{1}{\prod\limits_{\substack{p \text{ prime} \\ p|m}} \left(1+\frac{1}{p}\right)} \chi_q(m) &= \sum_{\substack{m=H_1 \\ (lt,m)=1}}^{H_2} h(m)\chi_q(m) \\
&= \sum_{\substack{m=H_1 \\ (lt,m)=1}}^{H_2} \left(\sum_{b|m} g(b)\right) \chi_q(m) \\
&= \sum_{\substack{m=H_1 \\ (lt,m)=1}}^{H_2} \left(\sum_{ab=m} g(b)\chi_q(b)\chi_q(a)\right) \\
&= \sum_{\substack{b=1 \\ (lt,b)=1}}^{H_2} g(b)\chi_q(b) \sum_{\substack{a=H_1/b \\ (lt,a)=1}}^{H_2/b} \chi_q(a)
\end{aligned}
$$

where g defined by Mobius inversion of h as follows,

$$
\begin{aligned}
h(m) &= \frac{1}{\prod\limits_{\substack{p \text{ prime} \\ p|m}} \left(1+\frac{1}{p}\right)} \\
g(m) &= \sum_{ab=m} h(a)\mu(b) \\
\Rightarrow h(m) &= \sum_{b|m} g(m)
\end{aligned}
$$

Note the following; for any prime p, $g(p^s) = -\frac{1}{p+1}$ if $s = 1$ else 0, and that g is multiplicative. Hence for g we have,

$$
\begin{aligned}
g(m) = \mu(m) \prod_{\substack{p \text{ prime} \\ p \nmid m}} \frac{1}{p+1} \\
|g(m)| \leq \frac{\mu(m)^2}{m}
\end{aligned}
$$

Hence we have the following bound for $A(H_1, H_2)$,

$$|A(H_1, H_2)| = O\left(\log(H_2)\tau(lt)\sqrt{q}\log(q)\right)$$

Now we are ready to consider the m-sum in Σ_2. Set $f(m) = N_2(m) - N_1(m) > 0$, note that it is a piece-wise monotonic function of m, initial increasing and decreasing after $M' = m = \left(\frac{t^2X}{l}\right)^{1/k}$. By partial summation we can write this innermost sum as follows,

$$\sum_{\substack{m=M_1\\(lt,m)=1}}^{M_2} f(m)\frac{1}{\prod\limits_{\substack{p\text{ prime}\\p|m}}\left(1+\frac{1}{p}\right)}\chi_q(m) = \sum_{m=M_1}^{M_2-1} A(1,m)(f(m)-f(m+1))$$

$$+ A(1,M_2)f(M_2) - A(1,M_1-1)f(M_1)$$

so we have,

$$\left|\sum_{\substack{m=M_1\\(lt,m)=1}}^{M_2} f(m)\frac{1}{\prod\limits_{\substack{p\text{ prime}\\p|m}}\left(1+\frac{1}{p}\right)}\chi_q(m)\right|$$
$$= O\left(\log(M_2)\tau(lt)\sqrt{q}\log(q)\right)\left(\sum_{m=M_1}^{M_2-1}|f(m)-f(m+1)|+f(M_2)+f(M_1)\right)$$
$$= O\left(\log(M_2)\tau(lt)\sqrt{q}\log(q)f(M')\right)$$
$$= O\left(\log(X)\tau(lt)\sqrt{q}\log(q)\sqrt{X}l^{-1}t\right)$$

Now $\tau(t) = o\left(t^{\epsilon}\right)$, $log(t) = O\left(\log(X)\right) = O\left(X^{\epsilon}\right)$ and choosing $Q < X$,

$$|\Sigma_2| = O\left(\tau(l)l^{-1}\sqrt{X}\log(X)\sqrt{Q}\log(Q)\right)\sum_{t=1}^{T'}\frac{\tau(t)^2}{t}$$
$$= O\left(\tau(l)l^{-1}X^{1/2+\epsilon}\sqrt{Q}\right)$$

Σ_3 can be dealt in a very similar fashion. The difference is that in that case we will be applying Polya-Vinogradov to a sum of χ_{qw} characters, and then each such sum will be weighed by a factor $\frac{1}{w^2}$. Since the series of $\frac{1}{w^2}$ converges we would have no problem in achieving the same asymptotic as for Σ_2. Hence, we have that,

$$\Sigma_2 + \Sigma_3 = O\left(\tau(l) l^{-1} X^{1/2+\epsilon} \sqrt{Q}\right)$$

Next we consider Σ_4,

$$\Sigma_4 = \sum_{\substack{m=1 \\ (l,m)=1}}^{M} \frac{\phi(m)}{m} \sum_{\substack{t=1 \\ (lm,t)=1}}^{T_{right}} \prod_{\substack{p \text{ prime} \\ p>2}} \left(1 - \frac{\rho_{lm}\left(\frac{p}{(p,t)}\right)}{p^2}\right) \frac{N_2 - N_1}{t^2} \sum_{\substack{q|t \\ q \geq Q}} \mu^2(q) \chi_q(lm)$$

To estimate this sum we need the following lemma,

Lemma 2.4.13. *Fix sufficiently large C. Then for any non-square integer $m > 0$,*

$$\sum_{c<C} \mu(c)^2 \chi_c(m) \ll C^{1/2} m^{1/4} \log(m)^{1/2}$$

Proof.

$$\begin{aligned}
\sum_{c \leq C} \mu(c)^2 \chi_c(m) &= \sum_{c<C} \sum_{e^2|c} \mu(e) \chi_c(m) \\
&= \sum_{e \leq \sqrt{C}} \mu(e) \sum_{f \leq C/e^2} \chi_s(m)
\end{aligned}$$

where the last step follows from the substitution $c = e^2 f$. Treating $\chi_s(m)$ as a Dirichlet character of modulus m (upto multiplication by 4) and applying Polya-Vinogradov we have,

$$\begin{aligned}
\sum_{c\leq C}\mu(c)^2\chi_c(m) &\ll \sum_{e\leq\sqrt{C}}\min\left(\frac{C}{e^2},\sqrt{m}\log(m)\right)\\
&\ll \sum_{e\leq E}\sqrt{m}\log(m)+\sum_{E<e\sqrt{C}}\frac{C}{e^2}\\
&\ll E\sqrt{m}\log(m)+\frac{C}{E}\\
&\ll C^{1/2}m^{1/4}\log(m)^{1/2}
\end{aligned}$$

where at the last step we set $E = C^{1/2}m^{-1/4}\log(m)^{-1/2}$ □

Now for non-square lm and fixed T,

$$\begin{aligned}
\sum_{\substack{t=1\\(lm,t)=1}}^{T}\sum_{\substack{q|t\\q\geq Q}}\mu^2(q)\chi_q(lm) &= \sum_{\substack{t=1\\(lm,t)=1}}^{T}\sum_{\substack{q|t\\q\leq t/Q}}\mu^2(t/q)\chi_{t/q}(lm)\\
&= \sum_{q\leq T/Q}\sum_{Q\leq c\leq T/q}\mu(c)^2\chi_c(lm)\\
&\ll \sum_{q\leq T/Q}\left(\frac{T}{q}\right)^{1/2}(lm)^{1/4}\log(lm)^{1/2}\\
&\ll \frac{T}{\sqrt{Q}}(lm)^{1/4}\log(lm)^{1/2}
\end{aligned}$$

Factoring out an infinite product of Euler factors independent of t from the t-sum in Σ_4 we have,

$$h(t)=\prod_{p|t}\frac{1-\frac{1}{p^2}}{1-\frac{\rho_{lm}(p)}{p^2}}$$

Similar to what we did in the case of Σ_2, using Mobius inversion on $h(t)$ we obtain a function

$g(t)$ bounded by $\frac{\mu(t)^2}{t^{2-\epsilon}}$. Now,

$$\begin{aligned}
&\sum_{\substack{t=1\\(lm,t)=1}}^{T} h(t) \sum_{\substack{q|t\\q\geq Q}} \mu^2(q)\chi_q(lm)\\
&= \sum_{\substack{t=1\\(lm,t)=1}}^{T} \left(\sum_{h|t} g(h)\right) \left(\sum_{\substack{q|t\\q\geq Q}} \mu^2(q)\chi_q(lm)\right)\\
&= \sum_{\substack{h=1\\(lm,h)=1}}^{T} g(h) \sum_{\substack{t=1\\(lm,t)=1}}^{T/h} \left(\sum_{\substack{q|t\\q\geq Q}} \mu^2(q)\chi_q(lm)\right)\\
&\ll \frac{T}{\sqrt{Q}}(lm)^{1/4}\log(lm)^{1/2}
\end{aligned}$$

Notice that $(N_2(t)-N_1(t))/t^2$ is a piece-wise monotonic function of t, increasing till $t = T_{left}$ and then decreasing. So again as in the case of Σ_2, we can bound the t-sum in Σ_4 using the trick of partial summations,

$$\begin{aligned}
&\sum_{\substack{t=1\\(lm,t)=1}}^{T_{right}} \prod_{\substack{p \text{ prime}\\p>2}} \left(1 - \frac{\rho_{lm}\left(\frac{p}{(p,t)}\right)}{p^2}\right) \frac{N_2 - N_1}{t^2} \sum_{\substack{q|t\\q\geq Q}} \mu^2(q)\chi_q(lm)\\
&\ll \frac{N_2(T_{left}) - N_1(T_{left})}{T_{left}^2} \frac{T_{right}}{\sqrt{Q}}(lm)^{1/4}\log(lm)^{1/2}\\
&= O\left(\frac{X^{1/2}}{l} \frac{(lm)^{1/4}\log(lm)^{1/2}}{\sqrt{Q}}\right)
\end{aligned}$$

So we have,

$$\begin{aligned}
\Sigma_4 &= O\left(l^{-3/4}\sqrt{X}\frac{M^{5/4}\log(lM)^{1/2}}{\sqrt{Q}}\right)\\
&= O\left(l^{-2}X^{1/2+\epsilon}\frac{X^{5/8}}{\sqrt{Q}}\right)
\end{aligned}$$

Finally,

$$\begin{aligned}\Sigma_2+\Sigma_3+\Sigma_4 &= O\left(\tau(l)l^{-1}X^{1/2+\epsilon}\sqrt{Q}\right)+O\left(l^{-2}X^{1/2+\epsilon}\frac{X^{5/8}}{\sqrt{Q}}\right)\\ &= O\left(\frac{\sqrt{\tau(l)}}{l^{3/2}}X^{13/16+\epsilon}\right)\end{aligned}$$

by setting $Q=\frac{1}{\tau(l)l}X^{5/8}$. Note that the perfect square case of lm in Σ_4 gives a contribution of $X^{3/4+\epsilon}\ll X^{13/16+\epsilon}$.

Proof of lemma 2.4.12

We use Mobius inversion. We first rearrange the euler products as follows,

$$\begin{aligned}Y\alpha(l,s,Y) &= \sum_{\substack{m=s\\(l,m)=1}}^{s+Y}\left(\frac{\phi(m)}{m}\right)^2\prod_{\substack{p\text{ prime}\\p\nmid lm}}\left(1-\frac{1}{p^2}\right)\\ &= \prod_{\substack{p\text{ prime}\\p\nmid l}}\left(1-\frac{1}{p^2}\right)\sum_{\substack{m=s\\(l,m)=1}}^{s+Y}\prod_{\substack{p\text{ prime}\\p\mid m}}\frac{1-\frac{1}{p}}{1+\frac{1}{p}}\end{aligned}\tag{2.65}$$

Set $h(m)=\prod_{\substack{p\text{ prime}\\p\mid m}}\frac{1-\frac{1}{p}}{1+\frac{1}{p}}$. Define $g(m)$ as,

$$g(m)=\sum_{m=ab}h(a)\mu(b)$$

Note that g is multiplicative and for any prime p,

$$g(p^r)=\begin{cases}1 & r=0\\ \frac{-\frac{2}{p}}{1+\frac{1}{p}} & r=1\\ 0 & \text{otherwise}\end{cases}$$

Hence,

$$|g(m)| \leq \mu(m)^2 \frac{2^{\omega(m)}}{m} = O\left(m^{-1+\epsilon}\right)$$

Now using Mobius inversion formula (in the second step) we have,

$$\begin{aligned}
&\sum_{\substack{m=s \\ (l,m)=1}}^{s+Y} h(m) \\
&= \sum_{\substack{m=s \\ (l,m)=1}}^{s+Y} \sum_{q|m} g(q) \\
&= \sum_{\substack{q=1 \\ (l,q)=1}}^{s+Y} \left(\frac{\phi(l)}{l} + O\left(Y^{-1} l^{\epsilon}\right)\right) \left(\left\lfloor \frac{s+Y}{q} \right\rfloor - \left\lfloor \frac{s}{q} \right\rfloor\right) g(q) \\
&= \sum_{\substack{q=1 \\ (l,q)=1}}^{s+Y} \left(\frac{\phi(l)}{l} + O\left(Y^{-1} l^{\epsilon}\right)\right) \left(\frac{Y}{q} + O(1)\right) g(q) \\
&= Y \frac{\phi(l)}{l} \sum_{\substack{q=1 \\ (l,q)=1}}^{s+Y} \frac{1}{q} g(q) + O\left(\sum_{q=1}^{s+Y} |g(q)|\right) + O\left(l^{\epsilon} \sum_{q=1}^{s+Y} \frac{1}{q}\right) \\
&= Y \frac{\phi(l)}{l} \left(\sum_{\substack{q \geq 1 \\ (l,q)=1}} \frac{1}{q} g(q) + O\left(\sum_{q \geq s+Y} \frac{1}{q} |g(q)|\right)\right) + O\left(\sum_{q=1}^{s+Y} m^{-1+\epsilon}\right) + O\left(l^{\epsilon} \log(s+Y)\right) \\
&= Y \frac{\phi(l)}{l} \left(\sum_{\substack{q \geq 1 \\ (l,q)=1 \\ \mu(q) \neq 0}} \frac{1}{q} g(q) + O\left(\sum_{q \geq s+Y} q^{-2+\epsilon}\right)\right) + O\left((s+Y)^{\epsilon}\right) + O\left(l^{\epsilon} \log(s+Y)\right) \\
&= Y \frac{\phi(l)}{l} \left(\prod_{\substack{p \text{ prime} \\ p \nmid l}} \left(1 + \frac{-\frac{2}{p}}{p\left(1 + \frac{1}{p}\right)}\right) + O\left((s+Y)^{-1+\epsilon}\right)\right) + O\left(l^{\epsilon} (s+Y)^{\epsilon}\right) \\
&= Y \frac{\phi(l)}{l} \prod_{\substack{p \text{ prime} \\ p \nmid l}} \frac{(1 - \frac{1}{p})(1 + \frac{2}{p})}{1 + \frac{1}{p}} + O\left(l^{\epsilon} (s+Y)^{\epsilon}\right)
\end{aligned}$$

Putting this in eq. (2.65) we get the needed result.

2.4.6 Reconciling the counting problem with its heuristic form and the case with the distribution of solutions to quadratic congruences

The difference between $H_X^{\Omega}[k]$ and its heuristic version $\tilde{H}_X^{\Omega}[k]$ is quantified by the disparity between the actual distribution of solutions to a certain set of quadratic congruences and its heuristically presumed equidistribution. In order to help quantify this disparity we define the following,

For positive integers A, g and integers $Y_2 > Y_1$ let us define,

$$e_A(g, Y_1, Y_2) = \left(\sum_{\substack{y \in [Y_1, Y_2] \\ y^2 \equiv A \mod g}} 1 \right) - \rho_A(g) \frac{Y_2 - Y_1}{g} \tag{2.66}$$

Suppose $G_2 > G_1$ and let $U \in [G_1, G_2]$. Define,

$$E_A(U, Y_1, Y_2) = \sum_{g \in U} e_A(g, Y_1, Y_2) \tag{2.67}$$

We may identify $e_A(g, Y_1, Y_2)$ as the localised error at some g and $E_A(U, Y_1, Y_2)$ as the total error over a given set U. It is this total error that we hope to have an asymptotic proportional to $|U|^{\delta}$ for some $\delta > 0$, the smaller the better. Recall that we wanted to study the distribution of the solutions to $lm^k \equiv l^2 n^2 \mod r^2 t^2$ in eq. (2.30). In particular we are interested in the case where U is more or less a bounded set of square numbers divisible by some fixed r^2.

Based on our work in section 2.4.4 we have the following disparity among $H_X^{\Omega}[k]$ and $\tilde{H}_X^{\Omega}[k]$,

$$|H_X^{\Omega}[k] - \tilde{H}_X^{\Omega}[k]| < \left| \max_{f=1,2,4} \left(\sum_{\substack{lm < \sqrt{X} \\ (l,m)=1 \\ l \text{ square-free}}} \sum_{\substack{r < \sqrt{X} \\ r \text{ odd}}} \mu(r) E_{lm^k} \left(\{ f r^2 t^2 \}_{t < T_{right}}, N_1, N_2 \right) \right) \right| \tag{2.68}$$

We can relate the right hand-side of eq. (2.68) to better concise forms. For that purpose from this point onward we consider the two cases $k = 3$ and $k > 3$ separately.

Case $k > 3$

Note that when t is sufficiently small such that $N_1, N_2 \gg r^2t^2$, the heuristic version turns out to be the actual situation, since we can treat it as follows. Divide the $N_2 - N_1$ into $\frac{N_2-N_1}{r^2t^2}$ number of r^2t^2 length intervals of which each would contain $\rho_{lm}(rt)$ number of solutions to the quadratic congruence of interest, hence we estimate the total count of solutions to be exactly the heuristic with a tolerable error of $O\left(\rho_{lm}(rt)\right)$. However, large t values do not give space for similar justification of the heuristic. As such the right hand-side of eq. (2.68) is dominated by large t, henceforth we assume that t is sufficiently large.

The r-sum on the right hand-side of eq. (2.68) can be written as,

$$\sum_{\substack{r<\sqrt{X} \\ r \text{ odd}}} \mu(r) E_{lm^k}\left(\{fr^2t^2\}_{t<T_{right}}, N_1, N_2\right) = \sum_{\substack{r<\sqrt{X} \\ r \text{ odd} \\ \mu(r)=1}} E_{lm^k}\left(\{fr^2t^2\}_{t<T_{right}}, N_1, N_2\right) \tag{2.69}$$

$$- \sum_{\substack{r<\sqrt{X} \\ r \text{ odd} \\ \mu(r)=-1}} E_{lm^k}\left(\{fr^2t^2\}_{t<T_{right}}, N_1, N_2\right) \tag{2.70}$$

Now we recall definition 2.4.3. Note that this r-sum corresponds to error terms associated to equidistribution of $x_{T_{right},\sqrt{X}}(f, lm^k)$ in very small intervals $\left[\frac{N_1}{f(rt)^2}, \frac{N_2}{f(rt)^2}\right) \subset [0, 1)$ of size $\sim \left(lm^k\right)^{-1/2}$, hence we are now able to introduce discrepancy of $x_{T_{right},\sqrt{X}}(f, lm^k)$ into the picture. The rate of change of both $\frac{N_1}{f(rt)^2}$ and $\frac{N_2}{f(rt)^2}$ with respect to t is very small, typically of the magnitude $X^{-(k-3)/2}$. Hence we may very well regard these intervals to remain constant at least for a considerable range of t-values. Set $T = \frac{\sqrt{lm^k}}{lm}$ and recall that $\gamma = \gamma(l, m) = \frac{lm}{v\sqrt{X}}$.

Now it follows that,

$$|H_X^{\Omega}[k] - \tilde{H}_X^{\Omega}[k]| = O\left(\sqrt{X}\max_{f|4}\left(\sum_{\substack{lm<\sqrt{X}\\(l,m)=1,\ \mu(l)\neq 0}} T\ \mathscr{D}\left(x_{T,\sqrt{X}}\left(f, lm^k\right)\right)\right)\right) \tag{2.71}$$

The problem of resolving the actual disparity between $H_X^{\Omega}[k]$ and $\tilde{H}_X^{\Omega}[k]$ reduces to estimating the discrepancy of $x_{T,\sqrt{X}}\left(f, lm^k\right)$. By the Erdos-Turan inequality, we are left to computing the exponential sums *eq.* (2.6) of $x_{T,\sqrt{X}}\left(f, lm^k\right)$ and we would wish that the estimates are small enough. Instead of the exact problem that we have at hand, let us consider a simplified version.

Definition 2.4.14. *Let w be a fixed positive integer, A be a fixed integer and G be a large positive integer then define,*

$$x_G^w(A) = \left\{\frac{x}{g} \in [0,1); A \equiv x^2\ mod\ g^w, (x,g) = 1, g < G\right\} \tag{2.72}$$

The simplified version of $x_{T,R}(f, A)$ is $x_G^2(A)$, which will be the object of interest in the discussion that follows.

Results on the asymptotics of $x_G^2(A)$ (as $G \to \infty$) is almost non-existent in literature. However, the distribution of $x_G^1(A)$ was first studied by Hooley in [10], he proved that $\mathcal{S}_h(x_G^1(A)) = O\left(A\max(|h|^{4/5+\epsilon} + G^{3/4+\epsilon})\right)$ for $h \neq 0$. His proof was based on parametrizing solutions to the quadratic coungruence $A \equiv x^2 mod\ g$, using binary quadratic forms of discriminant A that represent g. Unfortunately, Hooley's proof methods cannot be adapted to prove similar results for $x_G^2(A)$. Bykovskii [2] refines Hooeley's result proving for $A < 0$ and $h \neq 0$ that $\mathcal{S}_h(x_G^1(A)) = O\left(G^{2/3+\epsilon} + G^{1/2+\epsilon}|h|^{1/4} + |h|^{2/3+\epsilon}\right)$. The results of Hooley and Bykovskii gives $\mathscr{D}(x_G^1(A))$ bounds of the order $G^{-1/4+\epsilon}$ and $G^{-1/3+\epsilon}$ respectively. However, to have a meaningful result at least for $k = 5$ out of eq. (2.71), we would need

$\mathscr{D}\left(x_{T,R}\left(f,A\right)\right) = O\left(\left(RT\right)^{-3/5}\right)$. Hence, unfortunately known similar results are hardly adequate to proceed beyond eq. (2.71).

Soundararajan and Kowalski [15] quite recently published a very generalized result, that rigorously formalizes the intuition that would tempt one to believe that sequences such as $x_G^w(A)$ should exhibit equidistribution i.e. by Chinese Remainder Theorem, the solutions to quadratic coungruences *mod g* are compounded by solutions to similar quadratic congruences mod the primes that divide g, and such a process deems to produce randomness. Their result is generalizable to all sequences constructed through CRT, subjected to a very mild condition. However, the asymptotics from their results produce very large error terms of the order $G^{1-\epsilon}$ for $x_G^w(A)$, and hence practically of no use to our situation.

The t-sum in the heuristic setup we discussed in section 2.4.4 results in a term of order $X^{1/2}$ despite the sum running over a (very large) range in the order of $X^{k/4-1/2}$ integers. This is primarily because the summand results from counting the (normalized) solutions to quadratic congruences, that fall within a (very small) interval of the order $X^{-(k/4-1)}$, and we heuristically believe that this count (asymptotically) is proportional to the size of the interval. Under such circumstances, rigorous proof of our heuristic claims should come with considerably very small error terms to overcome the high rate of discounts received by the main term by the very small interval as we have shown above. But if we could show that the error resulting from a small interval is also proportional in a sense to the size of the interval, we could further relax the high expectations from the asymptotics of $x_{T,R}\left(f, lm^k\right)$. Attempts to adapt the proof of Erdos-Turan inequality failed miserably due to the high levels of variational errors we have to deal with when approximating indicator functions of very small intervals. However, very recently Marklof and Welsh [21] studies the distribution of $x_G^1(A)$ in very small intervals.

Alternatively, we could expect to see some useful results by fully exploiting the cancellations apparently available in eq. (2.71) due to the mobius function. This way of viewing this problem is related to the Sarnak conjecture [23], which says that in simpler terms that for a sequence of complex numbers $\{y_n\}$, $\frac{1}{N}\sum_{n<N}\mu(n)y_n \to 0$ as $N \to \infty$. The idea is that the mobius function behaves randomly with expectation 0. This idea is widely believed and sometimes referred to as the Mobius Randomness Law [13, p. 338]. However, such randomness is yet to be rigorously established.

Case $k = 3$

As we have seen in case $k > 3$, the heuristic is justified with a reasonable error term if $N_2 - N_1 > (rt)^2$. We will show that when $k = 3$, this is partly satisfied under certain restrictions on r and that the rest of it we settle for a tolerable error term.

From eqs. (2.21) to (2.25) we see that $N_1, N_2 \sim X^{3/4}$. For $r < X^{1/8}$ we have that $r^2t^2 \ll X^{3/4}$, so $N_1, N_2 \gg r^2t^2$.

We focus on eq. (2.68), consider its right hand side.

$$O\left(\sum_{lm<\sqrt{X}}\sum_{r<X^{1/8}} E_{lm^k}\left(\{fr^2t^2\}_{t<T_{right}}, N_1, N_2\right)\right) + O\left(\sum_{lm<\sqrt{X}}\sum_{X^{1/8}<r<X^{1/2}} E_{lm^k}\left(\{fr^2t^2\}_{t<T_{right}}, N_1, N_2\right)\right) \quad (2.73)$$

We have the following for $r < X^{1/8}$.

$$\left|E_{lm^3}\left(\{fr^2t^2\}_{t<T_{right}}, N_1, N_2\right)\right| = O\left(\rho_{lm}(frt)\right) = X^{\epsilon}$$

Since $\rho_c(h) \leq 2^{\omega(h)} \leq \sigma_0(h) = O\left(h^{\epsilon}\right)$. So we have,

$$O\left(\sum_{lm<\sqrt{X}} \sum_{r<X^{1/8}} E_{lm^k}\left(\{fr^2t^2\}_{t<T_{right}}, N_1, N_2\right)\right) = O\left(\sum_{lm<X^{1/2}} X^{1/8+\epsilon} T_{right}\right)$$
$$= O\left(X^{7/8+\epsilon}\right) \tag{2.74}$$

For $r > X^{1/8}$, we give a very crude upper-bound for the number of solutions to $m^3 \equiv ln^2 \mod r^2t^2$. Looking at it in a slightly different point of view and based on the bounds in eqs. (2.21), (2.22) and (2.25), for given (r,t) we say that we are looking to estimate the number of multiples of $(rt)^2$ in the form $w = x^3 - y^2 > 0$ for uniformly distributed $0 < x < m, 0 < y < M^{3/2}$. We first compute the density of $0 < w < M^3$. Let us approximate the distribution F of w. Then,

$$F(z) = P(0 < w < z) = \sum_{y=0}^{M^{3/2}} P\left(y^{2/3} < x < \left(y^2+z\right)^{1/3}\right) P(y)$$
$$= \sum_{y=0}^{M^{3/2}} \frac{\left(y^2+z\right)^{1/3} - y^{2/3}}{M} \frac{1}{M^{2/3}}$$

then for the density $f = F'$ of w we have

$$f(z) \sim \frac{1}{M^{5/2}} \sum_{y=0}^{M^{3/2}} \left(y^2+z\right)^{-2/3} \asymp \frac{1}{M^{5/2}} \int_0^{M^{3/2}} \left(y^2+z\right)^{-2/3} dy = \frac{1}{M^{5/2}} O\left(M^{1/2} z^{-2/3}\right)$$
$$= O\left(\frac{z^{-2/3}}{M^2}\right)$$

Now given an interval $[h_1, h_2] \subset [0, M^3]$ we have $(h_2-h_1)M^3 O\left(\frac{z^{-2/3}}{M^2}\right) = (h_2-h_1) O\left(M z^{-2/3}\right)$ number of w in that interval. Divide $[0, M^3]$ into equal intervals of length h. Now we bound the number of multiples of $(rt)^2$ of the form $w = x^3 - y^2 > 0$ as follows.

$$\sum_{s=1}^{\frac{M^3}{h}} \frac{h}{r^2t^2} O\left(M(sh)^{-2/3}\right) = \frac{Mh^{1/3}}{r^2t^2} \sum_{s=1}^{\frac{M^3}{h}} O\left(s^{-2/3}\right) = O\left(\frac{M^2}{r^2t^2}\right) = O\left(\frac{X}{r^2t^2}\right)$$

Summing r between $\left(X^{1/8}, X^{1/2}\right)$ and $t < T_{right}$ we have,

$$O\left(\sum_{lm<\sqrt{X}} \sum_{X^{1/8}<r<X^{1/2}} E_{lm^k}\left(\{fr^2t^2\}_{t<T_{right}}, N_1, N_2\right)\right) = O\left(X^{7/8}\right) \tag{2.75}$$

Hence by putting eqs. (2.74) and (2.75) in eq. (2.68) for $k = 3$ we have,

$$|H_X^{\Omega}[3] - \tilde{H}_X^{\Omega}[3]| = O\left(X^{7/8+\epsilon}\right)$$

Using eq. (2.64) finally we have,

$$H_X^{\Omega}[3] = \nu(\Omega)|^{-}\mathcal{D}_X| + c_3 X^{1/2+1/3} + O\left(X^{9/10+\epsilon}\right) \tag{2.76}$$

for some negative constant c_3.

2.4.7 Extending the heuristic results across local conditions

In this section we discuss the extension of results on ${}^{-}\mathcal{D}_X$ from section 2.4.4 to ${}^{-}\mathcal{D}_X^{\mathcal{S},\mathcal{R}}$. Denote the heuristic version of $H_X^{\Omega}(\mathcal{S};\mathcal{R})[k]$ as $\tilde{H}_X^{\mathcal{S};\mathcal{R}}(\Omega)[k]$. Our aim is to prove that,

$$\frac{\tilde{H}_X^{\mathcal{S};\mathcal{R}}(\Omega)[k]}{\tilde{H}_X^{\Omega}[k]} \sim \frac{|{}^{-}\mathcal{D}_X^{\mathcal{S},\mathcal{R}}|}{|{}^{-}\mathcal{D}_X|} \text{ as } X \to \infty \tag{2.77}$$

We will not present a full detail of the proof, but will explain the process of adapting the previous proofs in section 2.4.4. Let us also restrict the explanation only to case 1 in lemma 2.4.9 with $2 \mid l$, the other cases can be similarly handled.

Let Q be a finite set of odd primes each of order $O(1)$. Partition Q into three sets as Q_1, Q_2, Q_3. Define g_1, g_2 and g_3 as,

$$g_i = \prod_{q \in Q_i} q$$

Now define S_{Q_1,Q_2,Q_3} by modifying eq. (2.29),

$$\begin{aligned} S_{Q_1,Q_2,Q_3} &= S_{Q_1,Q_2,Q_3}(k,X,v,u) \\ &= \sum_{\substack{l=1 \\ \text{square free} \\ g_1|l \\ (g_2g_3,l)=1}}^{L} \sum_{\substack{m=1 \\ (2lg_2g_3,m)=1}}^{M} \sum_{\substack{t=1 \\ (lm,t)=1}}^{T_{right}} \sum_{\substack{|n|\in[N_1,N_2] \\ (m,n)=1 \\ nt^{-1}\in Res(u,m) \\ 2t^2|lm^k-l^2n^2 \\ g_2|\frac{lm^k-l^2n^2}{t^2} \\ g_3^2|\frac{lm^k-l^2n^2}{t^2}}} \sum_{\substack{(r,g_2g_3)=1 \\ r^2|\frac{lm^k-l^2n^2}{t^2}}} \mu(r) \end{aligned} \tag{2.78}$$

S_{Q_1,Q_2,Q_3} counts the number of solutions to case 1 in lemma 2.4.9, where d is divisible by all the primes in Q, l is divisible by all the primes in Q_1 but by none in $Q_2 \cup Q_3$, and $d = g_3^2 d'$ for some square-free odd positive integer d'.

For convenience we further restrict to $2 \mid l$, and with abuse of notation rewrite S_{Q_1,Q_2,Q_3} by modifying eq. (2.30) as follows,

$$S_{Q_1,Q_2,Q_3} = \sum_{\substack{l=1 \\ \text{square free} \\ 2|l \\ g_1|l \\ (g_2g_3,l)=1}}^{L} \sum_{\substack{m=1 \\ (g_2g_3l,m)=1}}^{M} \sum_{\substack{t=1 \\ (lm,t)=1}}^{T_{right}} \sum_{\substack{r \\ r \text{ odd} \\ (r,g_2g_3)=1}} \mu(r) \sum_{\substack{|n|\in[N_1,N_2] \\ (m,n)=1 \\ nt^{-1}\in Res(u,m) \\ lm^k=l^2n^2 \mod g_2g_3^2r^2t^2}} 1 \tag{2.79}$$

We now switch to $\tilde{S}_{Q_1,Q_2,Q_3}$ the heuristic version of S_{Q_1,Q_2,Q_3} and then we have the equivalent of eq. (2.31),

$$\tilde{S}_{Q_1,Q_2,Q_3} = \sum_{\substack{l=1 \\ \text{square free} \\ 2|l \\ g_1|l \\ (g_2g_3,l)=1}}^{L} \sum_{\substack{m=1 \\ (g_2g_3l,m)=1}}^{M} \sum_{\substack{t=1 \\ (lm,t)=1}}^{T_{right}} \sum_{\substack{r \\ r \text{ odd} \\ (r,g_2g_3)=1}} 2\frac{N_2-N_1}{g_2g_3^2r^2t^2}\rho_{lm}(g_2g_3^2r^2t^2)u\frac{\phi(m)}{m} \tag{2.80}$$

As we have dealt earlier, the r-sum can be written as follows,

$$\begin{aligned}
&\sum_{\substack{r \\ r \text{ odd} \\ (r,g_2g_3)=1}} \mu(r)2\frac{N_2-N_1}{g_2g_3^2r^2t^2}\rho_{lm}(g_2g_3^2r^2t^2)u\frac{\phi(m)}{m} \\
&= 2u\frac{1}{g_2g_3^2}\frac{N_2-N_1}{t^2}\rho_{lm}(g_2g_3^2t^2)\frac{\phi(m)}{m}\sum_{\substack{r \\ r \text{ odd} \\ (r,g_2g_3)=1}} \mu(r)\frac{\rho_{lm}(\frac{r^2}{(r^2,t^2)})}{r^2} \\
&= 2u\frac{1}{g_2g_3^2}\frac{N_2-N_1}{t^2}\rho_{lm}(g_2g_3^2t^2)\frac{\phi(m)}{m}\prod_{\substack{p \text{ prime} \\ p>2 \\ p\nmid g_2g_3}}\left(1-\frac{\rho_{lm}\left(\frac{p}{(p,t)}\right)}{p^2}\right)
\end{aligned}$$

Then by a similar result as lemma 2.4.10 we have the following,

$$\tilde{S}_{Q_1,Q_2,Q_3} = 2u\frac{1}{g_2g_3^2}\sum_{\substack{l=1 \\ \text{square free} \\ 2|l \\ g_1|l \\ (g_2g_3,l)=1}}^{L}\sum_{\substack{m=1 \\ (g_2g_3l,m)=1}}^{M}\frac{\phi(m)}{m}\sum_{\substack{t=1 \\ (lm,t)=1}}^{T_{right}}\prod_{\substack{p \text{ prime} \\ p\nmid g_2g_3lm}}\left(1-\frac{1}{p^2}\right)\frac{N_2-N_1}{t^2}+O(X^{13/16+\epsilon}) \tag{2.81}$$

We can compute the t-sum similar to that in section 2.4.4 and we then obtain the following equivalent of eq. (2.50),

$$\begin{aligned}
\tilde{S}_{Q_1,Q_2,Q_3} &= \frac{1}{g_2g_3^2}u\pi X^{1/2}\sum_{\substack{l=1 \\ \text{square free} \\ 2|l \\ g_1|l \\ (g_2g_3,l)=1}}^{L}\frac{1}{l}\frac{\phi(l)}{l}\sum_{m=1}^{M}\alpha_{g_2g_3l}(m)(1-\gamma(l,m))+\mathcal{E} \\
&+\begin{cases} O\left(X^{9/10+\epsilon}\right) & k=3 \\ O\left(X^{7/10+\epsilon}\right) & k=5 \\ O\left(X^{7/10}\right) & k>5 \end{cases}
\end{aligned} \tag{2.82}$$

Then by lemma 2.4.12 and computing the m-sum we get the equivalent of eq. (2.55),

$$\tilde{S}_{Q_1,Q_2,Q_3} = \frac{1}{g_2 g_3^2}\frac{\pi}{2}uvX\left(\sum_{\substack{l=1\\ \text{square free}\\ 2|l\\ g_1|l\\ (g_2g_3,l)=1}}^{L}\frac{1}{l^2}\frac{\phi(l)}{l}\alpha(g_2g_3l)\right) + \mathcal{E} + \begin{cases} O\left(X^{9/10+\epsilon}\right) & k=3 \\ O\left(X^{13/16+\epsilon}\right) & k>3 \end{cases} \tag{2.83}$$

Then finally computing the l-sum and $\mathcal{E}$, we have the final form of $\tilde{S}_{Q_1,Q_2,Q_3}$ which is equivalent to eq. (2.57),

$$\tilde{S}_{Q_1,Q_2,Q_3} = \frac{\frac{1}{g_1^2}}{\prod_{q|g_1}\left(1+\frac{1}{q}\right)^2}\frac{\frac{1}{g_2g_3^2}}{\prod_{q|g_2g_3}\left(1-\frac{1}{q}\right)\left(1+\frac{1}{q}\right)^2}\frac{2}{\pi^3}uvX + c_0X^{1/2+1/k} + \begin{cases} O\left(X^{9/10+\epsilon}\right) & k=3 \\ O\left(X^{13/16+\epsilon}\right) & k>3 \end{cases} \tag{2.84}$$

for some negative constant c_0.

Now let's fix Q_1 and vary the partitioning of $Q \setminus Q_1$ by Q_2, Q_3. Consider the following sum as it runs across the bi-partitions of $Q \setminus Q_1$,

$$\tilde{S}_{Q_1}(Q) = \sum_{\substack{Q_1,Q_2\\ Q_1\cap Q_2=\{\}\\ Q_1\cup Q_2=Q\setminus Q_1}} (-1)^{|Q_3|}\tilde{S}_{Q_1,Q_2,Q_3} \tag{2.85}$$

By the principle of inclusion-exclusion, $\tilde{S}_{Q_1}(Q)$ (as a heuristic) counts the number of solutions to case 1 (with $2 \mid l$) of lemma 2.4.9 where d is a square-free integer divisible by all primes in Q and l is divisible by all primes in Q_1 but none in $Q \setminus Q_1$. By easy computations it is easy to see that the q-products of the main term of $\tilde{S}_{Q_1}(Q)$ is

$$\frac{\prod_{q\in Q_1}\frac{1}{q^2}}{\prod_{q\in Q_1}\left(1+\frac{1}{q}\right)^2}\frac{\prod_{q\in Q\setminus Q_1}\frac{1}{q}}{\prod_{q\in Q\setminus Q_1}\left(1+\frac{1}{q}\right)^2}$$

Hence for some constant c' we have,

$$\tilde{S}_{Q_1}(Q)=\prod_{q\in Q_1}\frac{1}{q}\prod_{q\in Q}\frac{\frac{1}{q}}{\left(1+\frac{1}{q}\right)^2}\frac{2}{\pi^3}uvX+c'X^{1/2+1/k}+\begin{cases}O\left(X^{9/10+\epsilon}\right) & k=3\\ O\left(X^{13/16+\epsilon}\right) & k>3\end{cases} \tag{2.86}$$

Now we sum $\tilde{S}_{Q_1}(Q)$ as Q_1 runs over all subsets of Q, which (as a heuristic) gives the number of solutions to case 1 (with $2\mid l$) of lemma 2.4.9 where d is a square-free integer divisible by all primes in Q,

$$\begin{aligned}\tilde{S}(Q)&=\sum_{Q_1\subseteq Q}\tilde{S}_{Q_1}(Q) \qquad (2.87)\\ &=\left[\prod_{q\in Q}\frac{\frac{1}{q}}{\left(1+\frac{1}{q}\right)}\right]\frac{2}{\pi^3}uvX+c''X^{1/2+1/k}+\begin{cases}O\left(X^{9/10+\epsilon}\right) & k=3\\ O\left(X^{13/16+\epsilon}\right) & k>3\end{cases}\end{aligned}$$

for some constant c''.

Although tedious, it is easy to see how that we can extend this to the other cases cited in lemma 2.4.9 such that the discriminant is divisible by all the primes in a Q, and will give the same q-product. The cases 1, 2 and 3 of lemma 2.4.9 allow us to extended the result beyond the odd primes, so here after we will be considering local conditions for discriminants defined by all primes. Setting $\mathcal{S}=Q$ and $\mathcal{R}=\{\}$ we have for some constant c''',

$$\tilde{H}_X^{\mathcal{S};\{\}}(\Omega)[k]=\left[\prod_{q\in\mathcal{S}}\frac{\frac{1}{q}}{\left(1+\frac{1}{q}\right)}\right]\nu(\Omega)|^-\mathcal{D}_X|+c'''X^{1/2+1/k}+\begin{cases}O\left(X^{9/10+\epsilon}\right) & k=3\\ O\left(X^{13/16+\epsilon}\right) & k>3\end{cases} \tag{2.88}$$

Suppose now we want to consider $\mathcal{R} \neq \{\}$. Again by inclusion-exclusion the following sum that runs through all subsets of $\mathcal{R}$ gives the required result,

$$\begin{aligned}\tilde{H}_X^{\mathcal{S};\mathcal{R}}(\Omega)[k] &= \sum_{\mathcal{R}'\subseteq\mathcal{R}} (-1)^{|\mathcal{R}'|} \tilde{H}_X^{\mathcal{S}\cup\mathcal{R}';\{\}}(\Omega)[k] \\ &= \left[\prod_{q\in\mathcal{S}} \frac{\frac{1}{q}}{\left(1+\frac{1}{q}\right)}\right]\left[\prod_{q\in\mathcal{R}} \frac{1}{\left(1+\frac{1}{q}\right)}\right] \nu(\Omega)|^{-}\mathcal{D}_X| \quad + c_{k,\Omega,\mathcal{S},\mathcal{R}} X^{1/2+1/k} \\ &\quad + \begin{cases} O\left(X^{9/10+\epsilon}\right) & k=3 \\ O\left(X^{13/16+\epsilon}\right) & k>3 \end{cases}\end{aligned} \tag{2.89}$$

where $c_{k,\Omega,\mathcal{S},\mathcal{R}}$ is a constant that depends on k, Ω, $\mathcal{S}$ and $\mathcal{R}$. By noting that

$$|^{-}\mathcal{D}_X^{\mathcal{S},\mathcal{R}}| = \left[\prod_{q\in\mathcal{S}} \frac{\frac{1}{q}}{\left(1+\frac{1}{q}\right)}\right]\left[\prod_{q\in\mathcal{R}} \frac{1}{\left(1+\frac{1}{q}\right)}\right] |^{-}\mathcal{D}_X| + O\left(X^{1/2}\right)$$

we finally have that,

$$\tilde{H}_X^{\mathcal{S};\mathcal{R}}(\Omega)[k] = \nu(\Omega)|^{-}\mathcal{D}_X^{\mathcal{S},\mathcal{R}}| + c_{k,\Omega,\mathcal{S},\mathcal{R}} X^{1/2+1/k} + \begin{cases} O\left(X^{9/10+\epsilon}\right) & k=3 \\ O\left(X^{13/16+\epsilon}\right) & k>3 \end{cases} \tag{2.90}$$

Now by the same reasoning for the case $k=3$ in section 2.4.6 we have proven,

$$H_X^{\Omega}(\mathcal{S};\mathcal{R})[3] = \nu(\Omega)|^{-}\mathcal{D}_X^{\mathcal{S},\mathcal{R}}| + c_{3,\Omega,\mathcal{S},\mathcal{R}} X^{1/2+1/k} + O\left(X^{9/10+\epsilon}\right) \tag{2.91}$$

Chapter 3

Real quadratic - Equidistribution of odd-prime order Geodesic lengths

In this chapter we study ceratin asymptotics of closed primitive geodesics corresponding to quadratic ideal classes. We will state and prove a result extending a theorem by Siegel. Our main aim is to empirically investigate an equidistribution result for ideal classes of odd-prime order of real quadratic fields analogous to what we have seen in the preceding chapter. Then we state several conjectures based on our emperical analysis.

3.1 Setup

We consider real quadratic extensions $K/\mathbb{Q}$ with fundamental discriminant $D \in \mathcal{D}$. Let $\mathcal{O}$ be an order of K, with conductor f, $disc(\mathcal{O}) = f^2 D$. Suppose $Q(x, y) = ax^2 + bxy + \frac{b^2 - f^2 D}{4a} y^2$ is a primitive binary quadratic form of discriminant $f^2 D$. The two distinct roots r_1, r_2 of $Q(x, 1)$ lie on the boundary $\partial\mathcal{H}$ of the upper half plane $\mathcal{H}$. There is unique geodesic G_{r_1,r_2} on $\mathcal{H}$ whose limit points on $\partial\mathcal{H}$ are r_1 and r_2. We say G_{r_1,r_2} is positively oriented if $a > 0$ and negatively oriented if $a < 0$. Hence, we have the following association from the ideals

$\mathcal{I}_{\mathcal{O}}$, to the set of geodesics $\mathcal{G}$ on the complex upper half plane $\mathcal{H}$ [17, 21],

$$\begin{array}{ccccc} \mathcal{I}_{\mathcal{O}} & \longrightarrow & BQF(f^2D) & \longrightarrow & \mathcal{G} \\ \left[a, \frac{-b+f\sqrt{D}}{2}\right] & \longrightarrow & ax^2+bxy+\frac{b^2-f^2D}{4a}y^2 & \longrightarrow & G_{r_1,r_2} \end{array} \tag{3.1}$$

For any $G \in \mathcal{G}$, that has an automorphism in Γ, the action of Γ on $\mathcal{G}$ projects down to a closed primitive geodesic $\tilde{G}$ in $\Gamma\backslash\mathcal{H}$. In particular, for any quadratic form $[a, b, c] \in BQF(f^2D)$ its automorphism group is generated by $\begin{pmatrix} \frac{t+bu}{2} & au \\ -cu & \frac{t-bu}{2} \end{pmatrix} \in \Gamma$, where t, u is the smallest positive integral solution to $t^2 - f^2Du^2 = 4$, and it fixes r_1, r_2, G_{r_1,r_2}. Then we have a bijective map between the Γ-equivalence classes of $BQF(f^2D)$ and $img(BQF(f^2D)) \in \mathcal{G}$. Let us denote,

$$\mathcal{G}(f^2D) = img(C(f^2D)) \subset \mathcal{G}$$

Then we have the following bijection.

$$\begin{array}{ccccc} Cl^+(f^2D) & \longleftrightarrow & C(f^2D) & \longleftrightarrow & \mathcal{G}(f^2D) \\ \left[a, \frac{-b+f\sqrt{D}}{2}\right] & \longleftrightarrow & ax^2+bxy+\frac{b^2-f^2D}{4a}y^2 & \longleftrightarrow & \tilde{G}_{r_1,r_2} \end{array} \tag{3.2}$$

Let us fix an odd primes k. We define the following as the collection of all geodesics over all quadratic orders with discriminant less than X,

$$\mathcal{G}'_X = \bigcup_{\substack{0<D<X \\ D=disc(\mathcal{O}) \\ \mathcal{O} \text{ quadratic order}}} \mathcal{G}(D)$$

The following are definitions related to geodesics over fundamental discriminants.

$$\mathcal{G}_X = \bigcup_{D \in {}^+\mathcal{D}_X} \mathcal{G}(D)$$

$$\mathcal{G}_X^{\mathcal{S},\mathcal{R}} = \bigcup_{D \in {}^+\mathcal{D}_X^{\mathcal{S},\mathcal{R}}} \mathcal{G}(D)$$

$$\mathcal{G}(D)[k] = img(Cl^+(D)[k]) \subset \mathcal{G}(D)$$

$$\mathcal{G}_X[k] = \bigcup_{D \in {}^+\mathcal{D}_X} \mathcal{G}(D)[k]$$

$$L_X = \sum_{C \in \mathcal{G}_X} ||C||$$

$$L_X\left(\mathcal{S};\mathcal{R}\right) = \sum_{C \in \mathcal{G}_X^{\mathcal{S},\mathcal{R}}} ||C||$$

$$L_X^{\Omega} = \sum_{C \in \mathcal{G}_X} ||C \cap \Omega||$$

$$L_X[k] = \sum_{C \in \mathcal{G}_X[k]} ||C||$$

$$L_X^{\Omega}[k] = \sum_{C \in \mathcal{G}_X[k]} ||C \cap \Omega||$$

$$L_X\left(\mathcal{S};\mathcal{R}\right)[k] = \sum_{C \in \mathcal{G}_X[k] \cap \mathcal{G}_X^{\mathcal{S},\mathcal{R}}} ||C||$$

$$L_X^{\Omega}\left(\mathcal{S};\mathcal{R}\right)[k] = \sum_{C \in \mathcal{G}_X[k] \cap \mathcal{G}_X^{\mathcal{S},\mathcal{R}}} ||C \cap \Omega||$$

where $\Omega \in \Gamma \backslash \mathcal{H}$. In addition, notation-wise, if $\mathcal{R} = \{\}$ we write $L_X\left(\mathcal{S};\right)$ $\left(L_X\left(\mathcal{S};\right)[k]\right)$ and if $\mathcal{S} = \{\}$ we write $L_X\left(;\mathcal{R}\right)$ $\left(L_X\left(;\mathcal{R}\right)[k]\right)$.

3.2 Related prior work

In this section we state two theorems that will provide inspiration to our investigation and conjectures later in this chapter.

The first of which we recall theorem 1.3.2 from section 1.3.

Theorem 1.3.2 (Duke [7, Theorem 1(ii)])**.** *Suppose $D \in^{+} \mathcal{D}$ and $\Omega \in \Gamma\backslash\mathcal{H}$. Then for some $\delta = \delta(\Omega) > 0$.*

$$\frac{\sum\limits_{C\in\mathcal{G}(D)} ||C \cap \Omega||}{\sum\limits_{C\in\mathcal{G}(D)} ||C||} = \nu(\Omega) + O\left(|D|^{-\delta}\right) \text{ as } D \to \infty \tag{1.19}$$

where $||{*}||$ denote the length of a geodesic (segment).

The second is a theorem from Siegel [24] in which he proves an asymptotic formula on the total length of geodesics in $\mathcal{G}'_X$.

Theorem 3.2.1 (Siegel [24, p. 1])**.**

$$\sum_{C\in\mathcal{G}'_X} ||C|| = \frac{\pi^2}{9\zeta(3)} X^{3/2} + O\left(X \log X\right) \tag{3.3}$$

3.3 Experiments on prime order geodesic lengths

We start with concrete computations of closed geodesics in $\Gamma\backslash\mathcal{H}$ to study there asymptotic patterns experimentally. The computations were done on computer nodes of the University of Maryland High Performance Cluster [6]. Class group data readily available on LMFDB [20] as well as data generated using algorithms on Pari/Gp were used in the experiments. Generating geodesic related data is computationally heavy, and limitations on computational resources permits us only to compute data on up to around $X = 2^{14} \cdot 10^3 = 16384000$.

Inspired by Siegel's theorem 3.2.1 we are interested in studying the asymptotic behaviour of L_X and $L_X(S;R)$ for large enough X. We present experimental data on these asymptotics in table 3.1. Later we will state and prove these asymptotics in full agreeance with table 3.1.

X	$\frac{L_X}{X^{3/2}}$	$\frac{L_X(\{3\};)}{L_X}$	$\frac{L_X(\{5\};)}{L_X}$	$\frac{L_X(\{7\};)}{L_X}$	$\frac{L_X(;\{3\})}{L_X}$	$\frac{L_X(;\{5\})}{L_X}$	$\frac{L_X(;\{7\})}{L_X}$
1,000	0.5247	0.2553	0.1746	0.1324	0.7447	0.8254	0.8676
2,000	0.5457	0.2415	0.1709	0.1298	0.7585	0.8291	0.8702
4,000	0.5582	0.2391	0.1728	0.1293	0.7609	0.8272	0.8707
8,000	0.5657	0.2372	0.1665	0.1273	0.7628	0.8335	0.8727
16,000	0.5728	0.2339	0.1649	0.1255	0.7661	0.8351	0.8745
32,000	0.577	0.2324	0.1645	0.1251	0.7676	0.8355	0.8749
64,000	0.5799	0.2315	0.163	0.1244	0.7685	0.837	0.8756
$1.28 \cdot 10^5$	0.5822	0.2305	0.1627	0.1238	0.7695	0.8373	0.8762
$2.56 \cdot 10^5$	0.5837	0.2301	0.1622	0.1236	0.7699	0.8378	0.8764
$5.12 \cdot 10^5$	0.5849	0.2297	0.1618	0.1234	0.7703	0.8382	0.8766
$1.02 \cdot 10^6$	0.5857	0.2293	0.1616	0.1231	0.7707	0.8384	0.8769
$2.05 \cdot 10^6$	0.5863	0.2291	0.1614	0.123	0.7709	0.8386	0.877

Table 3.1: $\frac{L_X}{X^{3/2}}$, $\frac{L_X(\{p\};)}{L_X}$ and $\frac{L_X(;\{p\})}{L_X}$ for $p = 3, 5, 7$

Now we investigate the asymptotics of $L_X[k]$ and $L_X(S;R)[k]$. Figure 3.1 shows $\log(L_X[k])$ vs $\log(X)$ for $k = 3, 5, 7$. The regression on the plotted data shows that $L_X[k]$ is asymptotic to $X^{3/2}$ (up to a constant). Suppose $L_X[k] \sim c_k X^{3/2}$.

We now investigate the coefficient c_k of the main term in $L_X[k]$. Table 3.2 shows $L_X[k]$ for $k = 3, 5, 7$ normalized by $X^{3/2}$ and fig. 3.2 plots the results against $\log(X)$.

The data does not seem to show smooth convergence of normalized $L_X[k]$, but shows some sort of orderly behavior for large enough X. Unfortunately, the computational resources at hand limits us from experimenting with larger X. Set $\tilde{c}_k = \frac{1}{k^2}\left(1 - \frac{1}{k^2}\right)$. However, the ratios $\frac{L_X[5]}{L_X[3]}$ and $\frac{L_X[7]}{L_X[3]}$ tend to show much smoother behavior as shown in fig. 3.3, and suggests that

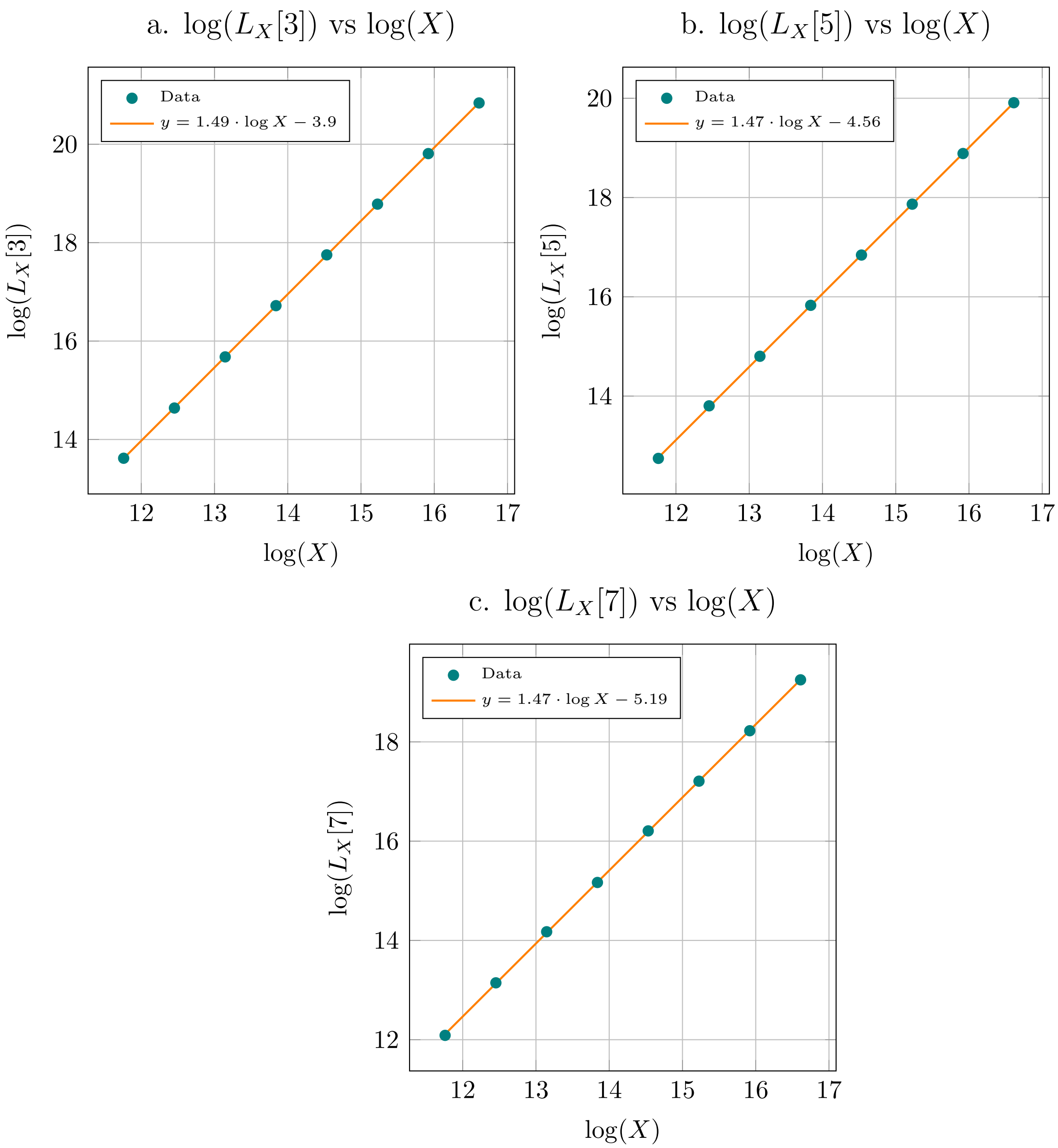

Figure 3.1: $\log(L_X[k])$ vs $\log(X)$ and main power term estimation for $\log(L_X[k])$

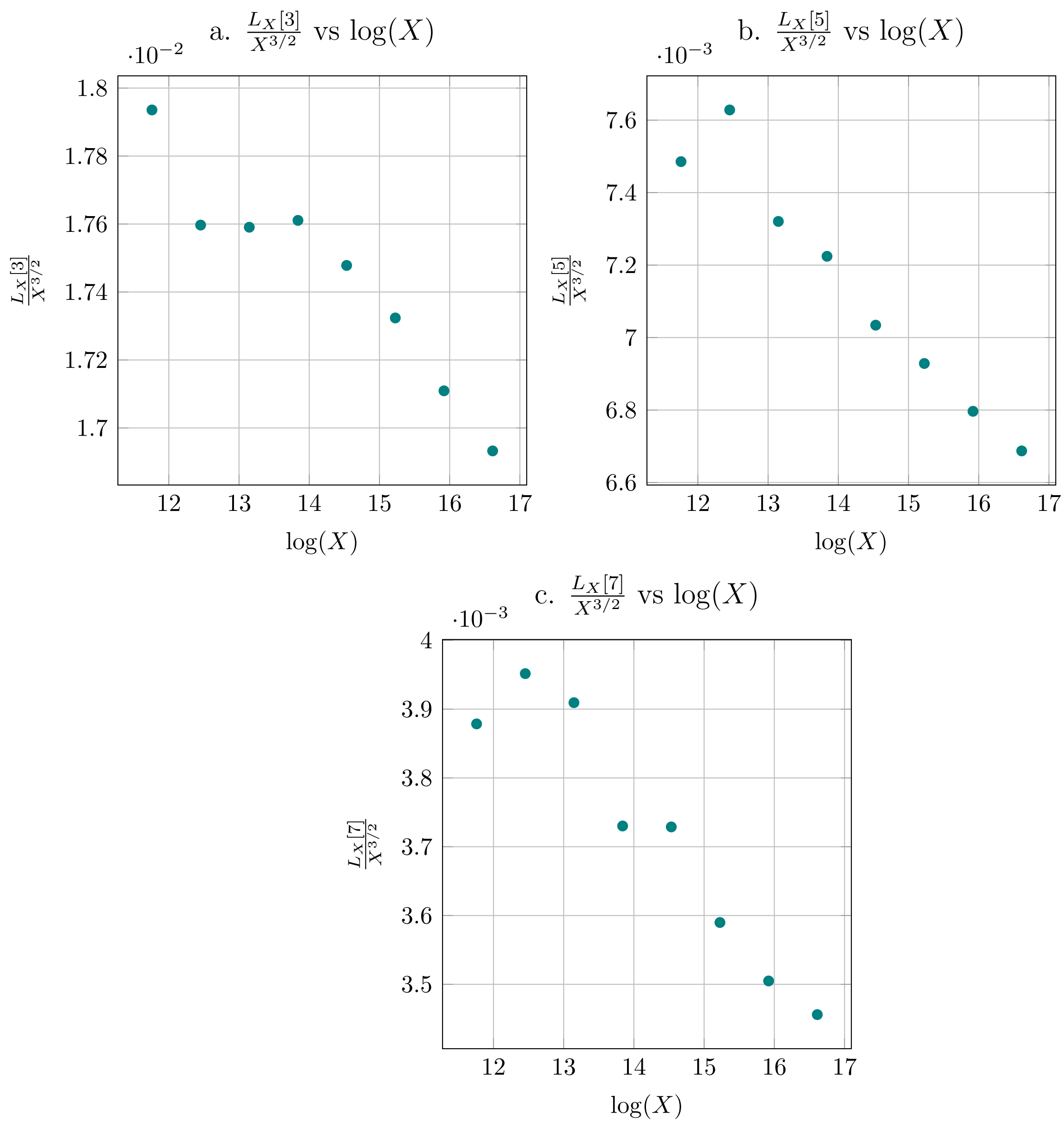

Figure 3.2: $\frac{L_X[k]}{X^{3/2}}$ vs $\log(X)$ for $k = 3, 5, 7$

X	$\frac{L_X[3]}{X^{3/2}}$	$\frac{L_X[5]}{X^{3/2}}$	$\frac{L_X[7]}{X^{3/2}}$	$\frac{L_X[5]}{L_X[3]}$	$\frac{L_X[7]}{L_X[3]}$
$1.28 \cdot 10^5$	0.0179	$7.4855 \cdot 10^{-3}$	$3.8786 \cdot 10^{-3}$	0.4174	0.2162
$2.56 \cdot 10^5$	0.0176	$7.6281 \cdot 10^{-3}$	$3.9516 \cdot 10^{-3}$	0.4335	0.2246
$5.12 \cdot 10^5$	0.0176	$7.3205 \cdot 10^{-3}$	$3.9094 \cdot 10^{-3}$	0.4162	0.2222
$1.02 \cdot 10^6$	0.0176	$7.2242 \cdot 10^{-3}$	$3.7302 \cdot 10^{-3}$	0.4102	0.2118
$2.05 \cdot 10^6$	0.0175	$7.0342 \cdot 10^{-3}$	$3.7288 \cdot 10^{-3}$	0.4025	0.2133
$4.1 \cdot 10^6$	0.0173	$6.9285 \cdot 10^{-3}$	$3.5898 \cdot 10^{-3}$	0.3999	0.2072
$8.19 \cdot 10^6$	0.0171	$6.7965 \cdot 10^{-3}$	$3.5049 \cdot 10^{-3}$	0.3972	0.2049
$1.64 \cdot 10^7$	0.0169	$6.6872 \cdot 10^{-3}$	$3.4559 \cdot 10^{-3}$	0.3949	0.2041

Table 3.2: $\frac{L_X[3]}{X^{3/2}}$ for $k = 3, 5, 7$ and $\frac{L_X[k]}{L_X[3]}$ for $k = 5, 7$

$c_k \propto \tilde{c}_k$.

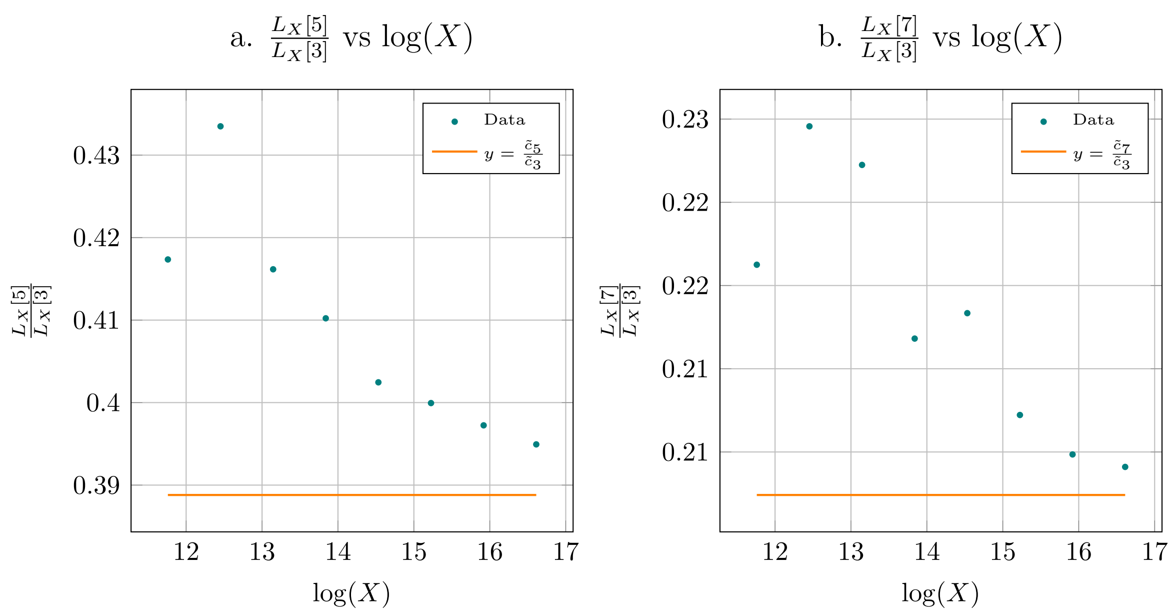

Figure 3.3: $\frac{L_X[k]}{L_X[3]}$ vs $\log(X)$ for $k = 5, 7$

We also estimate c_k by minimizing the asymptotics of the possible secondary term, $|L_X[k] - cX^{3/2}|$ as c runs through positive real values. We do this by trial and error. We thus obtain the following estimates $\hat{c}_k$,

$$\hat{c}_3 = 0.017, \hat{c}_5 = 0.0066, \hat{c}_k = 0.00341$$

the estimates agree with the results we have seen in table 3.2 of $\frac{L_X[k]}{X^{3/2}}$. Set $Y_k = |L_X[k] - \hat{c}_k X^{3/2}|$. Figure 3.4 shows $\log(Y_k)$ vs $\log(X)$, and regression on the data (on large enough X) shows a secondary term of $X^{1+\epsilon}$, which agrees with the secondary term in Siegel's theorem 3.2.1.

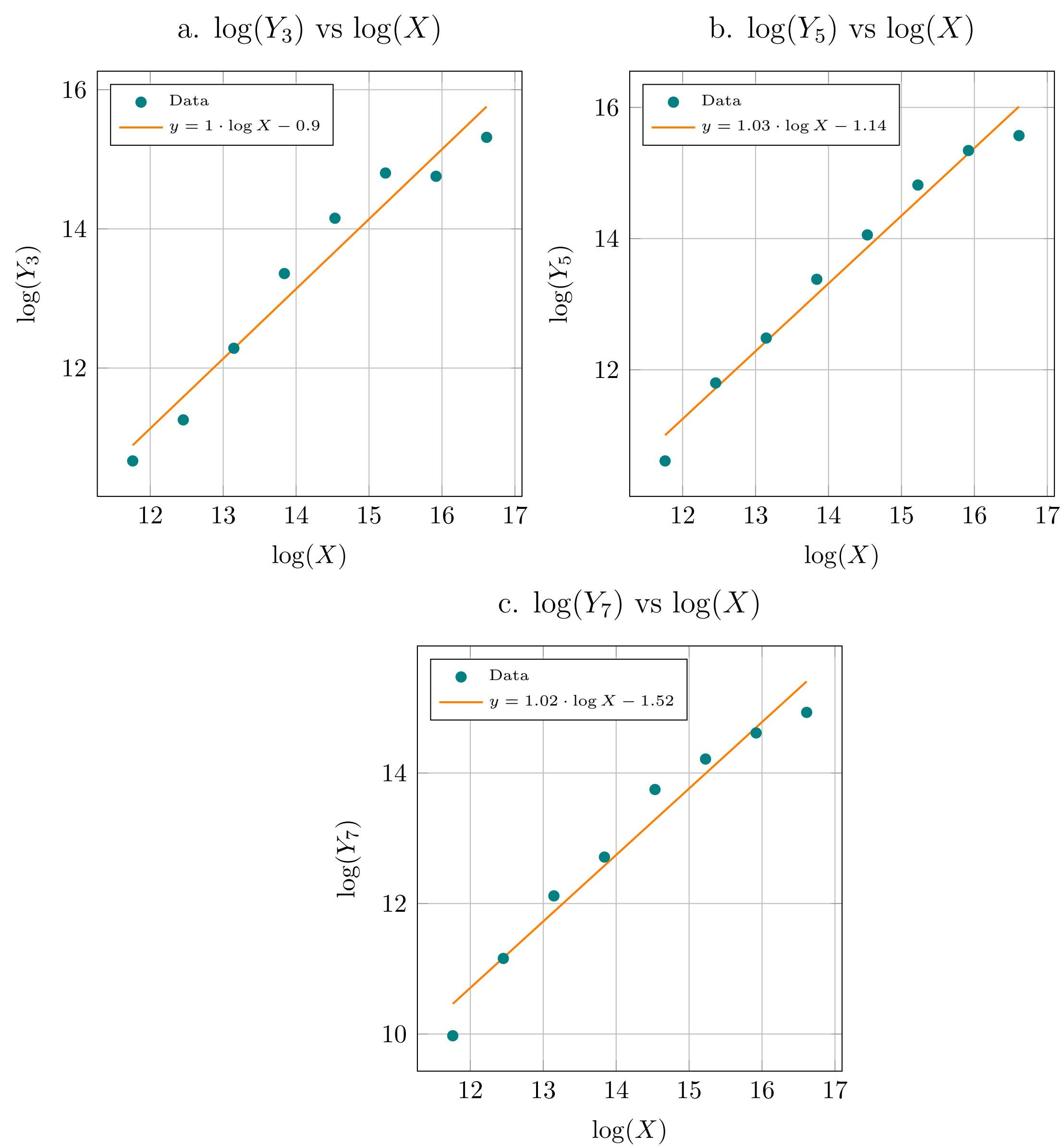

Figure 3.4: $\log(Y_k)$ vs $\log(X)$ and secondary term estimation for $\log(L_X[k])$

Next we investigate the asymptotic behavior of $L_X[k]$ along families of discriminants conditioned by local properties, in particular $L_X(S;R)[k]$. Based on tables 3.3 and 3.4 we note the following observations.

1. For both $k=5$ and $k=7$, $\frac{L_X(\{3\};)[k]}{L_X[k]}$ tend to approach a common constant ≈ 0.133
2. The ratio $\frac{L_X(\{5\};)[5]}{L_X[5]} : \frac{L_X(\{3\};)[5]}{L_X[5]}$ tend to approach ≈ 0.6, also note that $\frac{\left|{}^{+}\mathcal{D}_X^{\{5\},\{\}}\right|}{\left|{}^{+}\mathcal{D}_X^{\{3\},\{\}}\right|} \sim 0.667$.
3. The ratio $\frac{L_X(\{7\};)[5]}{L_X[5]} : \frac{L_X(\{3\};)[5]}{L_X[5]}$ tend to approach ≈ 0.51, also note that $\frac{\left|{}^{+}\mathcal{D}_X^{\{7\},\{\}}\right|}{\left|{}^{+}\mathcal{D}_X^{\{3\},\{\}}\right|} \sim 0.5$.
4. The product $\frac{L_X(\{3\};)[5]}{L_X[5]} \cdot \frac{L_X(\{5\};)[5]}{L_X[5]}$ tend to approach ≈ 0.0106 and $\frac{L_X(\{3,5\};)[5]}{L_X[5]}$ tend to approach ≈ 0.012.
5. Set $e=\frac{8}{15}$ and for $p=3,5,7$ set $e_p = e\dfrac{\frac{1}{p}}{1+\frac{1}{p}} \sim e\frac{\left|{}^{+}\mathcal{D}_X^{\{p\},\{\}}\right|}{\left|{}^{+}\mathcal{D}_X\right|}$, then $\frac{L_X(\{p\};)[5]}{L_X[5]}$ tends to approach e_p and $\frac{L_X(\{3\};)[7]}{L_X[7]}$ tends to approach e_3. Also, $\frac{L_X(\{3,5\};)[5]}{L_X[5]}$ tend to approach e_3e_5.
6. As expected $\frac{L_X(;\{p\})[5]}{L_X[5]} \approx 1 - \frac{L_X(\{p\};)[5]}{L_X[5]}$ for $p=3,5,7$.
7. The product $\frac{L_X(;\{3\})[5]}{L_X[5]} \cdot \frac{L_X(;\{5\})[5]}{L_X[5]}$ tend to approach ≈ 0.795 and $\frac{L_X(;\{3,5\})[5]}{L_X[5]}$ tend to approach ≈ 0.795 too.
8. The product $\frac{L_X(\{3\};)[5]}{L_X[5]} \cdot \frac{L_X(;\{5\})[5]}{L_X[5]}$ tend to approach ≈ 0.123 and $\frac{L_X(;\{3,5\})[5]}{L_X[5]}$ tend to approach ≈ 0.121.

These observations show evidence to results similar to that promised in section 2.4.7 in the previous chapter.

Next in our experiments we investigate the possibility of a theorem 1.3.2 like result of prime-order geodesics. We compute $\dfrac{L_X^{\Omega}[k]}{L_X[k]}$ for a few chosen types of $\Omega \in \Gamma\backslash\mathcal{H}$.

Let us define $\Omega_{i,j} \in \Gamma\backslash\mathcal{H}$ for $i \geq 1$ and $j = 1,2,3$ in terms of its projection on $\mathcal{F}$. Define, $\Omega_{i,1} = \{z \in \Gamma\backslash\mathcal{H} : \Re z \in [-\frac{1}{2}, -\frac{1}{6}), \Im z \geq i\}$, $\Omega_{i,2} = \{z \in \Gamma\backslash\mathcal{H} : \Re z \in [-\frac{1}{6}, \frac{1}{6}), \Im z \geq i\}$ and

X	$\frac{L_X(\{3\};)[5]}{L_X[5]}$	$\frac{L_X(\{5\};)[5]}{L_X[5]}$	$\frac{L_X(\{7\};)[5]}{L_X[5]}$	$\frac{L_X(\{3,5\};)[5]}{L_X[5]}$	$\frac{L_X(\{3\};)[7]}{L_X[7]}$
$1.28 \cdot 10^5$	0.1327	0.0768	0.0629	0.0102	0.1017
$2.56 \cdot 10^5$	0.1356	0.0801	0.0600	0.0121	0.1126
$5.12 \cdot 10^5$	0.1330	0.0816	0.0609	0.0121	0.1290
$1.02 \cdot 10^6$	0.1291	0.0862	0.0662	0.0120	0.1301
$2.05 \cdot 10^6$	0.1330	0.0824	0.0685	0.0128	0.1300
$4.1 \cdot 10^6$	0.1348	0.0808	0.0683	0.0127	0.1358
$8.19 \cdot 10^6$	0.1345	0.0829	0.0687	0.0125	0.1335
$1.64 \cdot 10^7$	0.1334	0.0817	0.0689	0.0124	0.1356

Table 3.3: $\frac{L_X(S;)[5]}{L_X[5]}$ for $S = \{3\}, \{5\}, \{7\}, \{3,5\}$ and $\frac{L_X(\{3\};)[7]}{L_X[7]}$

X	$\frac{L_X(;\{3\})[5]}{L_X[5]}$	$\frac{L_X(;\{5\})[5]}{L_X[5]}$	$\frac{L_X(;\{7\})[5]}{L_X[5]}$	$\frac{L_X(;\{3,5\})[5]}{L_X[5]}$	$\frac{L_X(\{3\};\{5\})[5]}{L_X[5]}$
$1.28 \cdot 10^5$	0.8673	0.9232	0.9371	0.8007	0.1225
$2.56 \cdot 10^5$	0.8644	0.9199	0.9400	0.7964	0.1235
$5.12 \cdot 10^5$	0.8670	0.9184	0.9391	0.7975	0.1209
$1.02 \cdot 10^6$	0.8709	0.9138	0.9338	0.7968	0.1171
$2.05 \cdot 10^6$	0.8670	0.9176	0.9315	0.7974	0.1202
$4.1 \cdot 10^6$	0.8652	0.9192	0.9317	0.7971	0.1221
$8.19 \cdot 10^6$	0.8655	0.9171	0.9313	0.7951	0.1220
$1.64 \cdot 10^7$	0.8666	0.9183	0.9311	0.7973	0.1210

Table 3.4: $\frac{L_X(;R)[5]}{L_X[5]}$ for $R = \{3\}, \{5\}, \{7\}, \{3,5\}$ and $\frac{L_X(\{3\};\{5\})[5]}{L_X[5]}$

$\Omega_{i,3} = \{z \in \Gamma\backslash\mathcal{H} : \Re z \in [\frac{1}{6}, \frac{1}{2}), \Im z \geq i\}$. Note that for any $i \geq 1$, $\nu(\Omega_{i,1}) = \nu(\Omega_{i,2}) = \nu(\Omega_{i,3})$, so set $\nu_i = \nu(\Omega_{i,j})$ for $j = 1, 2, 3$.

X	ν_1	$\frac{L_X^{\Omega_{1,1}}[3]}{L_X[3]}$	$\frac{L_X^{\Omega_{1,2}}[3]}{L_X[3]}$	$\frac{L_X^{\Omega_{1,3}}[3]}{L_X[3]}$	$\frac{L_X^{\Omega_{1,1}}[5]}{L_X[5]}$	$\frac{L_X^{\Omega_{1,2}}[5]}{L_X[5]}$	$\frac{L_X^{\Omega_{1,3}}[5]}{L_X[5]}$	$\frac{L_X^{\Omega_{1,1}}[7]}{L_X[7]}$	$\frac{L_X^{\Omega_{1,2}}[7]}{L_X[7]}$	$\frac{L_X^{\Omega_{1,3}}[7]}{L_X[7]}$
$1.28 \cdot 10^5$	0.3183	0.3182	0.3177	0.3176	0.3181	0.3178	0.3180	0.3191	0.3176	0.3184
$2.56 \cdot 10^5$	0.3183	0.3180	0.3180	0.3177	0.3184	0.3179	0.3182	0.3185	0.3179	0.3183
$5.12 \cdot 10^5$	0.3183	0.3180	0.3181	0.3179	0.3183	0.3180	0.3182	0.3185	0.3179	0.3183
$1.02 \cdot 10^6$	0.3183	0.3181	0.3180	0.3180	0.3182	0.3181	0.3181	0.3184	0.3181	0.3182
$2.05 \cdot 10^6$	0.3183	0.3182	0.3181	0.3181	0.3182	0.3182	0.3181	0.3183	0.3181	0.3182
$4.1 \cdot 10^6$	0.3183	0.3182	0.3182	0.3181	0.3183	0.3182	0.3182	0.3183	0.3182	0.3182
$8.19 \cdot 10^6$	0.3183	0.3182	0.3182	0.3182	0.3183	0.3183	0.3182	0.3183	0.3182	0.3182

Table 3.5: $\frac{L_X^{\Omega_{1,j}}[k]}{L_X[k]}$ for $j = 1, 2, 3$ and $k = 3, 5, 7$

X	ν_2	$\frac{L_X^{\Omega_{2,1}}[3]}{L_X[3]}$	$\frac{L_X^{\Omega_{2,2}}[3]}{L_X[3]}$	$\frac{L_X^{\Omega_{2,3}}[3]}{L_X[3]}$	$\frac{L_X^{\Omega_{2,1}}[5]}{L_X[5]}$	$\frac{L_X^{\Omega_{2,2}}[5]}{L_X[5]}$	$\frac{L_X^{\Omega_{2,3}}[5]}{L_X[5]}$	$\frac{L_X^{\Omega_{2,1}}[7]}{L_X[7]}$	$\frac{L_X^{\Omega_{2,2}}[7]}{L_X[7]}$	$\frac{L_X^{\Omega_{2,3}}[7]}{L_X[7]}$
$1.28 \cdot 10^5$	0.1592	0.1519	0.1523	0.1519	0.1540	0.1539	0.1537	0.1552	0.1541	0.1548
$2.56 \cdot 10^5$	0.1592	0.1535	0.1538	0.1534	0.1554	0.1555	0.1552	0.1557	0.1555	0.1556
$5.12 \cdot 10^5$	0.1592	0.1548	0.1551	0.1548	0.1561	0.1562	0.1560	0.1565	0.1565	0.1565
$1.02 \cdot 10^6$	0.1592	0.1559	0.1559	0.1558	0.1568	0.1569	0.1567	0.1572	0.1572	0.1572
$2.05 \cdot 10^6$	0.1592	0.1566	0.1566	0.1566	0.1574	0.1575	0.1574	0.1575	0.1575	0.1575
$4.1 \cdot 10^6$	0.1592	0.1572	0.1572	0.1572	0.1578	0.1579	0.1578	0.1579	0.1579	0.1578
$8.19 \cdot 10^6$	0.1592	0.1576	0.1577	0.1576	0.1582	0.1582	0.1582	0.1582	0.1582	0.1582

Table 3.6: $\frac{L_X^{\Omega_{2,j}}[k]}{L_X[k]}$ for $j = 1, 2, 3$ and $k = 3, 5, 7$

X	ν_3	$\frac{L_X^{\Omega_{3,1}}[3]}{L_X[3]}$	$\frac{L_X^{\Omega_{3,2}}[3]}{L_X[3]}$	$\frac{L_X^{\Omega_{3,3}}[3]}{L_X[3]}$	$\frac{L_X^{\Omega_{3,1}}[5]}{L_X[5]}$	$\frac{L_X^{\Omega_{3,2}}[5]}{L_X[5]}$	$\frac{L_X^{\Omega_{3,3}}[5]}{L_X[5]}$	$\frac{L_X^{\Omega_{3,1}}[7]}{L_X[7]}$	$\frac{L_X^{\Omega_{3,2}}[7]}{L_X[7]}$	$\frac{L_X^{\Omega_{3,3}}[7]}{L_X[7]}$
$1.28 \cdot 10^5$	0.1061	0.0971	0.0971	0.0971	0.0994	0.0993	0.0992	0.0997	0.0994	0.0997
$2.56 \cdot 10^5$	0.1061	0.0991	0.0991	0.0990	0.1011	0.1011	0.1010	0.1009	0.1010	0.1009
$5.12 \cdot 10^5$	0.1061	0.1008	0.1008	0.1007	0.1020	0.1020	0.1019	0.1024	0.1024	0.1024
$1.02 \cdot 10^6$	0.1061	0.1020	0.1020	0.1020	0.1030	0.1030	0.1030	0.1034	0.1035	0.1034
$2.05 \cdot 10^6$	0.1061	0.1029	0.1029	0.1029	0.1039	0.1039	0.1039	0.1039	0.1039	0.1039
$4.1 \cdot 10^6$	0.1061	0.1036	0.1036	0.1036	0.1044	0.1044	0.1044	0.1044	0.1044	0.1044
$8.19 \cdot 10^6$	0.1061	0.1042	0.1042	0.1041	0.1048	0.1049	0.1048	0.1049	0.1049	0.1049

Table 3.7: $\frac{L_X^{\Omega_{3,j}}[k]}{L_X[k]}$ for $j = 1, 2, 3$ and $k = 3, 5, 7$

The tables 3.5 to 3.8 provide evidence of $\dfrac{L_X^{\Omega}[k]}{L_X[k]}$ converging to $\nu(\Omega)$. So we would expect $L_X^{\Omega}[k]$ to have a $X^{3/2}$ main term. Then we would expect a secondary term of $O\left(X\log(X)\right)$ similar to that in Siegel's theorem 3.2.1, so hence $\dfrac{L_X^{\Omega}[k]}{L_X[k]}$ to have a secondary term of $O\left(X^{-(1/2-\epsilon)}\right)$.

X	ν_4	$\frac{L_X^{\Omega_{4,1}}[3]}{L_X[3]}$	$\frac{L_X^{\Omega_{4,2}}[3]}{L_X[3]}$	$\frac{L_X^{\Omega_{4,3}}[3]}{L_X[3]}$	$\frac{L_X^{\Omega_{4,1}}[5]}{L_X[5]}$	$\frac{L_X^{\Omega_{4,2}}[5]}{L_X[5]}$	$\frac{L_X^{\Omega_{4,3}}[5]}{L_X[5]}$	$\frac{L_X^{\Omega_{4,1}}[7]}{L_X[7]}$	$\frac{L_X^{\Omega_{4,2}}[7]}{L_X[7]}$	$\frac{L_X^{\Omega_{4,3}}[7]}{L_X[7]}$
$1.28 \cdot 10^5$	0.0796	0.0699	0.0696	0.0697	0.0722	0.0721	0.0721	0.0723	0.0725	0.0722
$2.56 \cdot 10^5$	0.0796	0.0723	0.0720	0.0721	0.0740	0.0739	0.0739	0.0738	0.0739	0.0737
$5.12 \cdot 10^5$	0.0796	0.0740	0.0738	0.0738	0.0750	0.0750	0.0750	0.0753	0.0752	0.0752
$1.02 \cdot 10^6$	0.0796	0.0752	0.0752	0.0751	0.0760	0.0760	0.0760	0.0765	0.0765	0.0765
$2.05 \cdot 10^6$	0.0796	0.0761	0.0761	0.0761	0.0770	0.0770	0.0770	0.0772	0.0771	0.0771
$4.1 \cdot 10^6$	0.0796	0.0769	0.0769	0.0768	0.0776	0.0776	0.0776	0.0777	0.0777	0.0777
$8.19 \cdot 10^6$	0.0796	0.0775	0.0775	0.0775	0.0782	0.0782	0.0781	0.0783	0.0782	0.0782

Table 3.8: $\frac{L_X^{\Omega_{4,j}}[k]}{L_X[k]}$ for $j = 1, 2, 3$ and $k = 3, 5, 7$

Linear regression on the log-errors vs $\log(X)$ in figs. 3.5 and 3.6 provide further evidence to this speculation.

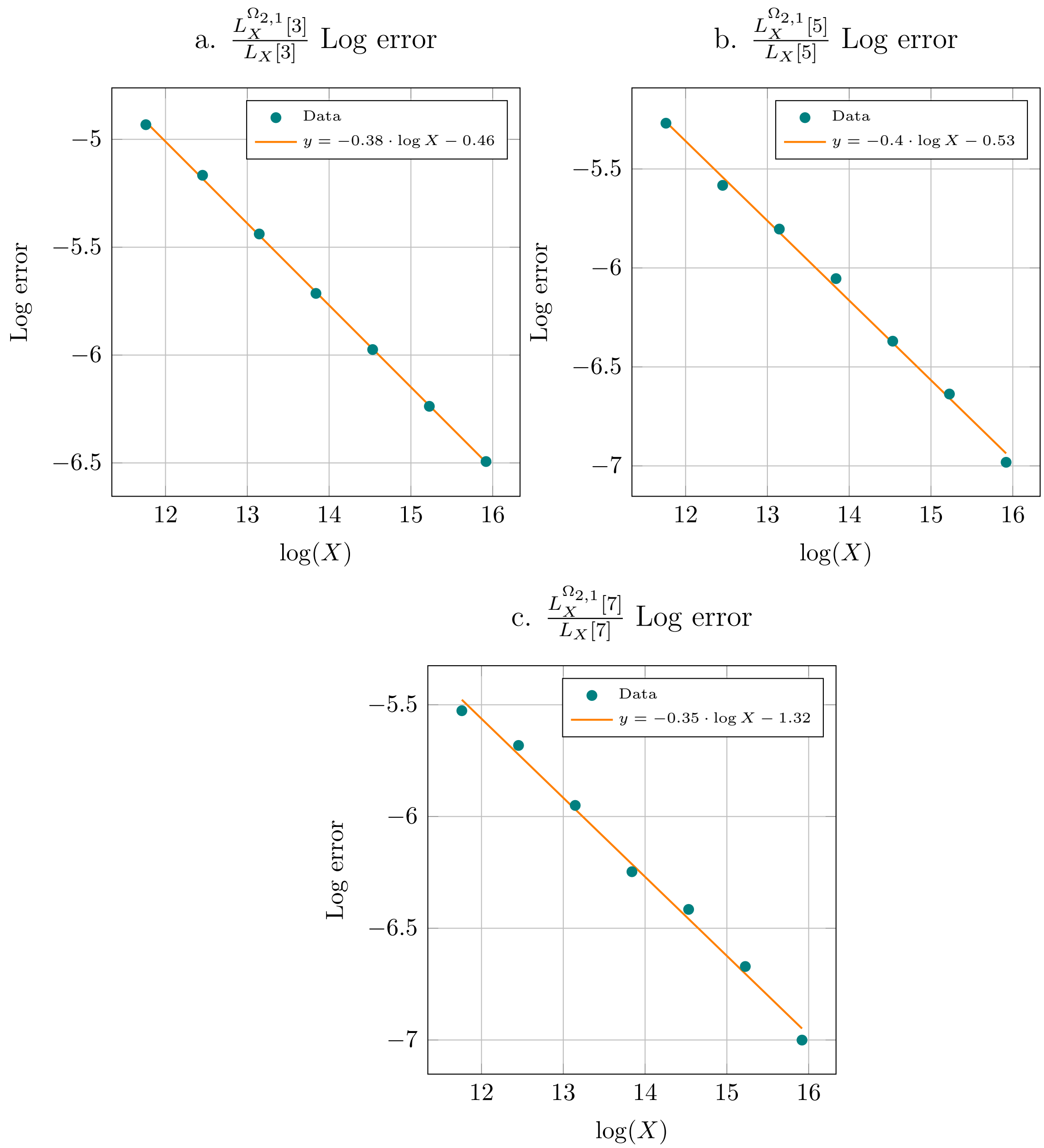

Figure 3.5: Secondary term estimation for $\frac{L_X^{\Omega_{2,1}}[k]}{L_X[k]}$

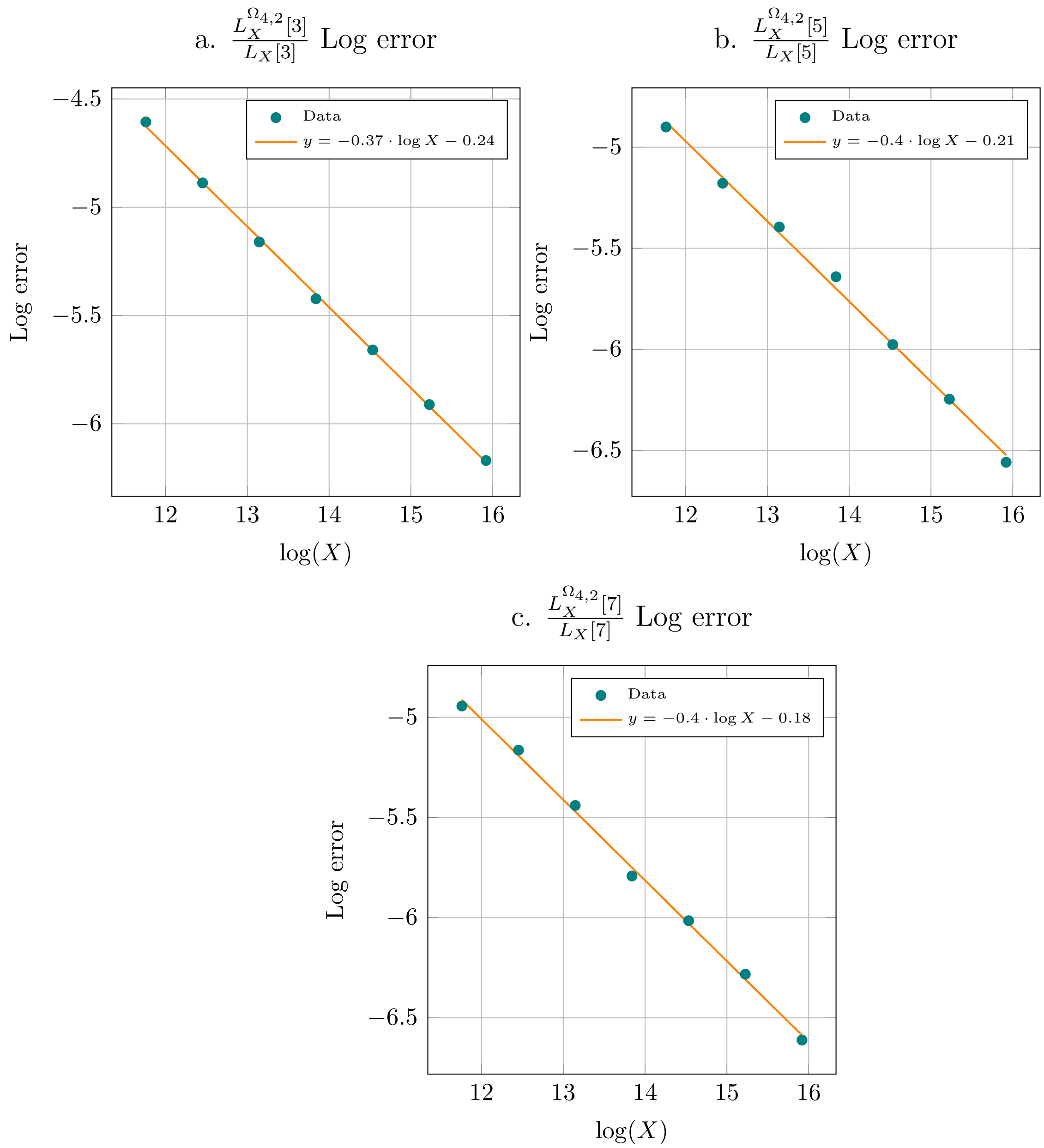

Figure 3.6: Secondary term estimation for $\frac{L_X^{\Omega_{4,2}}[k]}{L_X[k]}$

3.4 Main results

In this chapter we will use Siegel's proof for theorem 3.2.1, to derive the following asymptotic formula for $\mathcal{G}_X^{\mathcal{S},\mathcal{R}}$ which also incorporates a cumulative form of Duke's theorem 1.3.2.

Theorem 3.4.1.

$$\begin{aligned} L_X^{\Omega}\left(\mathcal{S};\mathcal{R}\right) &= \sum_{C\in\mathcal{G}_X^{\mathcal{S},\mathcal{R}}} \|C\cap\Omega\| \\ &= \nu(\Omega)\left(\prod_{p\in\mathcal{S}} s(p)\right)\left(\prod_{p\in\mathcal{R}} r(p)\right)\left(\prod_{p\notin\mathcal{S}\cup\mathcal{R}} h(p)\right)\frac{\pi^2}{9}X^{3/2} + O\left(X\log X\right) \end{aligned} \tag{3.4}$$

where,

$$h(p) = \left(1-\frac{1}{p}\right)\left(1+\frac{1}{p}-\frac{1}{p^3}\right)$$

$$s(p) = \frac{1}{p}\left(1-\frac{1}{p}\right)^2\left(1+\frac{1}{p}\right),\ r(p) = h(p)-s(p) = 1-\frac{1}{p}$$

Note that theorem 3.4.1 agrees very well with table 3.1. For example, $\prod_{p \text{ prime}} h(p) \approx 0.58775$ matches with the data shown for $\frac{L_X}{X^{3/2}}$.

In the process of proving theorem 3.4.1, we prove the following lemma which we will state here due to its seemingly general applicability.

Lemma 3.4.2. *Let $\Lambda\subset\mathbb{Z}^3$ be uniformly distributed in $\mathbb{Z}^3$ with density $\rho>0$.*

$$|W_N\cap\Lambda| = \rho N^3 + O\left(N^{\omega}\right),\ \omega<3$$

where $W_N\subset\mathbb{R}^3$ is a cube of size $N\times N\times N$. Then,

$$\sum_{\substack{C\in\mathcal{G}'_X \\ C \text{ from } [a,b,c],\ (a,b,c)\in\Lambda}} \|C\cap\Omega\| = \rho\frac{\pi^2}{9}\nu(\Omega)X^{3/2} + O\left(X^{\omega/2}\log(X)\right) \tag{3.5}$$

3.5 Asymptotic conjectures on prime order Geodesic lengths

Based on the experimental evidence in the precious section and inspired by the results of Duke and Siegel in theorems 1.3.2 and 3.2.1 respectively we state the following conjectures.

Conjecture 3.5.1. *For a fixed odd prime k,*

$$L_X[k] = c_k X^{3/2} + O\left(X \log(X)\right) \tag{3.6}$$

where $c_k = \alpha \tilde{c}_k$ for some constant α independent of k.

Based on the experimental results $\alpha \approx 0.17$, and $\tilde{c}_k \sim \frac{1}{k^2}\left(1 - \frac{1}{k^2}\right)$.

Conjecture 3.5.2. *Let $\mathcal{S}$ and $\mathcal{R}$ be mutually exclusive finite sets of prime numbers. Then for a fixed odd prime k,*

$$L_X\left(S; R\right)[k] = \left(\prod_{p\in S} e_p\right)\left(\prod_{p\in R}(1-e_p)\right) L_X[k] + O\left(X\log(X)\right) \tag{3.7}$$

where $e_p = \beta\tilde{e}_p$ for some constant β independent of both k and p.

Based on the experimental results $\beta \approx \frac{8}{15}$, and $\tilde{e}_p \sim \dfrac{\frac{1}{p}}{1+\frac{1}{p}} \sim \frac{\left|{}^{+}\mathcal{D}_X^{\{p\},\{\}}\right|}{\left|{}^{+}\mathcal{D}_X\right|}$.

Remark. *Surprisingly, $\frac{L_X(\mathcal{S};\mathcal{R})}{L_X}$ (established by theorem 3.4.1 and data in table 3.1), does not seem to agree with $\frac{L_X(\mathcal{S};\mathcal{R})[k]}{L_X[k]}$ (tables 3.3 and 3.4). Also, the extra β factor in e_p was rather unexpected and works against intuition. It could be the case that the limited data at hand may not be giving an accurate picture.*

Conjecture 3.5.3. *Let $\Omega \subseteq \Gamma\backslash\mathcal{H}$, then for a fixed odd prime k,*

$$\frac{L_X^{\Omega}[k]}{L_X[k]} = \nu(\Omega) + O\left(X^{-1/2}\log(X)\right) \tag{3.8}$$

Finally, combining all three conjectures above, we have the following generalized conjecture.

Conjecture 3.5.4. *Let $\mathcal{S}$ and $\mathcal{R}$ be mutually exclusive finite sets of prime numbers, and $\Omega \subseteq \Gamma \backslash \mathcal{H}$. Then for a fixed odd prime k,*

$$L_X^{\Omega}\left(\mathcal{S};\mathcal{R}\right)[k] = \nu(\Omega)\left(\prod_{p\in\mathcal{S}} e_p\right)\left(\prod_{p\in\mathcal{R}}(1-e_p)\right) c_k X^{3/2} + O\left(X\log(X)\right) \tag{3.9}$$

where e_p and c_k are defined as in the previous statements.

3.6 Proof of theorem 3.4.1

In [24] Siegel proves theorem 3.2.1 using two different proof methods; one based on the class number formula and the other appealing directly to the geometry of primitive closed geodesics parameterized by primitive binary quadratic forms. In this section, we will restate the latter proof method, with some changes to suit our proof of theorem 3.4.1.

3.6.1 Restating Siegel's proof, with generalisation to Ω

We will interchangeably view the upper complex plane as an euclidean space and hyperbolic space depending on context. For the proofs consider $\mathcal{F}$ the usual standard projection of $\Gamma\backslash\mathcal{H}$ onto $\mathcal{H}$ however we will work with $\mathcal{F}_0$ instead of $\mathcal{F}$ for convenience. Suppose $\Omega \in \mathcal{F}_0$ be a connected and convex domain. Specifically, in our proofs we will assume $\Omega = [u_1, u_2) \times [v, \infty)$ for some $-\frac{1}{2} \leq u_1 < u_2 < \frac{1}{2}$ and $v \geq \frac{\sqrt{3}}{2}$, for convenience. However, it is easily seen how the results can be generalised from those proven for these simplified Ω's.

Let

$$Q = Q(x,y) = ax^2 + bxy + cy^2 = [a,b,c]$$

be a binary quadratic form with discriminant (non-square) $D = b^2 - 4ac > 0$, $a \neq 0$ and $Q(x,1) = 0$ roots $r_1 > r_2$, so $Q = a(x - r_1y)(x - r_2y)$. For λ a positive real parameter define the positive quadratic form $P = [\alpha, \beta\gamma]$ as follows.

$$P = P(x, y) = |a|\Big(\lambda^{-1}(x - r_1 y)^2 + \lambda(x - r_2 y)^2\Big) = \alpha x^2 + 2\beta xy + \gamma y^2 \tag{3.10}$$

Note that $D = \alpha\gamma - \beta^2$.Now let $\tau = \xi + i\eta$ be the root of $P(x, 1) = 0$ in $\mathcal{H}$. Then,

$$a(\xi^2 + \eta^2) + b\xi + c = 0 \tag{3.11}$$

and

$$\frac{\tau - r_1}{\tau - r_2} = i\lambda \tag{3.12}$$

Note that when $\lambda \to 0$ then $\tau \to r_2$ and when $\lambda \to \infty$ then $\tau \to r_1$. Also, the vectors defined by $\tau - r_1$ and $\tau - r_2$ are orthogonal (in euclidean sense). Hence we see that as λ varies from 0 to ∞, τ traces the half circle on the upper half plane with ends r_1 and r_2, in other words it describes the geodesic G_{r_1,r_2}.

We will call Q acceptable, if for at least one λ, τ lies in $\mathcal{F}_0$. For Q to be acceptable note that at least one of $\pm\frac{1}{2} + i\frac{\sqrt{3}}{2}$ should lie inside the circle described by eq. (3.11). So we have $a + \frac{1}{2}b + c \leq 0$ or $a - \frac{1}{2}b + c \leq 0$, hence $a + c \leq \frac{1}{2}|b|$. Since

$$(|b| - a)^2 + 3a^2 = \frac{1}{4}(4a - |b|)^2 + \frac{3}{4}b^2 = D + 4a\left(a + c - \frac{1}{2}|b|\right)$$

we have that,

$$a^2 \leq \frac{1}{3}D,\ b^2 \leq \frac{4}{3}D,\ 4a|c| = |b^2 - D| \leq D \tag{3.13}$$

It is clear that the intersection of G_{r_1,r_2} with Ω is an arc $G_{r_1,r_2,\Omega}$ that corresponds to an interval $\lambda_1 \leq \lambda \leq \lambda_2$. Let us denote the hyperbolic length of $G_{r_1,r_2,\Omega}$ by $\ell_\Omega = \ell_\Omega(a, b, c)$, a continuous function with respect to a, b, c. Using eq. (3.12), we may replace the hyperbolic line element $ds = \frac{(d\xi^2+d\eta^2)^{1/2}}{\eta}$ by $\frac{d\lambda}{\lambda}$. Hence,

$$\ell_\Omega = \ell_\Omega(a,b,c) = \begin{cases} 0 & G_{r_1,r_2,\Omega} = \{\} \\ \int_{\lambda_1}^{\lambda_2} \frac{d\lambda}{\lambda} = \log\frac{\lambda_2}{\lambda_1} & \text{otherwise} \end{cases} \tag{3.14}$$

Now for Q, the corresponding closed primitive geodesic $\tilde{G}_Q$ in $\Gamma\backslash\mathcal{H}$ projects down on $\mathcal{F}$ as a collection of arcs of geodesics with end points that are roots of binary quadratic forms equivalent to Q. Hence we have the following,

$$||\tilde{G}_Q \cap \Omega|| = \sum_{Q\sim[a',b',c']} \ell_\Omega(a',b',c')$$

So now we have,

$$\sum_{C\in\mathcal{G}(D)} ||C\cap\Omega|| = 2\sum_{\substack{b^2-4ac=D\\ a>0}} \ell_\Omega(a,b,c)$$

$$\sum_{C\in\mathcal{G}'_X} ||C\cap\Omega|| = 2\sum_{\substack{b^2-4ac\le X\\ b^2-4ac \text{ non-square}\\ a>0}} \ell_\Omega(a,b,c) \tag{3.15}$$

the factor 2 on the right hand side accounts for the forms with $a<0$, since for each form $[a,-b,c]$, there will be a corresponding $[-a,b,-c]$, such that $\ell_\Omega(a,-b,c) = \ell_\Omega(-a,b,-c)$.

Let $\vartheta > 0$ be a 'small' constant that we will determine later. For convenience, we set $Y^2 = X$ and define the following.

$$J = \frac{1}{Y^3}\sum_{\substack{b^2-4ac\le Y^2\\ a>0}} \ell_\Omega(a,b,c) \tag{3.16}$$

$$J_\vartheta = \frac{1}{Y^3}\sum_{\substack{b^2-4ac\le Y^2\\ a>\vartheta Y}} \ell_\Omega(a,b,c) \tag{3.17}$$

Our interest lies in computing J. The strategy is to view J as a Riemann sum and estimate it using an integral. However, as $\frac{a}{Y} \to 0$, $\ell_\Omega(a,b,c)$ is unbounded and also the domain of integration unbounded too. So, instead we estimate J_ϑ, using an integral closer enough to J by suitably choosing small enough ϑ. With that in mind we define the following integrals. Note that in these integrals, a, b, c are real numbers rather than integers.

$$I = \iiint\limits_{\substack{b^2-4ac\leq 1\\ a>0}} \ell_\Omega(a,b,c)da\,db\,dc \tag{3.18}$$

$$I_\vartheta = \iiint\limits_{\substack{b^2-4ac\leq 1\\ a>\vartheta}} \ell_\Omega(a,b,c)da\,db\,dc \tag{3.19}$$

We will use some estimations proven by Siegel in [24, p. 9], and have the following.

$$|J - J_\vartheta| = O\left(\vartheta\log(\vartheta^{-1})\right) + O\left(\vartheta\log(Y)\right) \tag{3.20}$$

Similarly we have,

$$|I - I_\vartheta| = O\left(\vartheta\log(\vartheta^{-1})\right) \tag{3.21}$$

Now we want to estimate J_ϑ using I_ϑ. Note that two binary quadratic forms $[a,b,c]$ and $[ta,tb,tc]$ $(t \neq 0)$ with real coefficients, corresponds to the same geodesic, and hence $\ell_\Omega(a,b,c) = \ell_\Omega(ta,tb,tc)$. Hence we may normalise a, b, c (by dividing out each of them by Y) in J_ϑ, such that $a \geq \vartheta$ and $D \leq 1$. Hence we have that,

$$J_\vartheta = \frac{1}{Y^3} \sum_{\substack{b^2-4ac\leq 1\\ a>\vartheta\\ (Ya,Yb,Yc)\in\mathbb{Z}^3}} \ell_\Omega(a,b,c) \tag{3.22}$$

To quantify $|J_\vartheta - I_\vartheta|$ the maximum variation of ℓ_Ω within a small region in $\mathbb{R}^3$ should be considered. We state and prove the following lemma for that purpose.

Lemma 3.6.1. $|\ell_\Omega(a,b,c) - \ell_\Omega(a+\Delta a, b+\Delta b, c+\Delta c)| = O\left(\frac{1}{a}\max \Delta a, \Delta b, \Delta c\right)$

Proof. By mean value theorem we know that $\ell_\Omega(a,b,c) - \ell_\Omega(a+\Delta a, b+\Delta b, c+\Delta c) = \Delta a \partial_a \ell_\Omega(a',b',c') + \Delta b \partial_b \ell_\Omega(a',b',c') + \Delta c \partial_c \ell_\Omega(a',b',c')$, for some $a' \in (a, a+\Delta a), b' \in (b, b+\Delta b), c' \in (c, c+\Delta c)$. To find the derivatives of ℓ_Ω, we consider eq. (3.12). We have that $\lambda^2 = \frac{|\tau - r_1|^2}{|\tau - r_2|^2}$, hence

$$\log \lambda = \log\left((\xi - r_1)^2 + \eta^2\right) - \log\left((\xi - r_2)^2 + \eta^2\right)$$

By eq. (3.14) to bound the derivatives of ℓ_Ω it is sufficient to bound the derivatives of $\log\left((\xi - r_i)^2 + \eta^2\right)$ where $\xi + i\eta$ represents any point of intersection of $\partial\Omega$ with the geodesic of interest. We will show by computation how $\partial_b \ell_\Omega$ can be bounded, the other derivatives can be dealt similarly.

$$\partial_b \log\left((\xi - r)^2 + \eta^2\right) = \frac{2(\xi - r)}{(\xi - r)^2 + \eta^2}\frac{\partial \xi}{\partial b} + \frac{-2(\xi - r)}{(\xi - r)^2 + \eta^2}\frac{\partial r}{\partial b} + \frac{2\eta}{(\xi - r)^2 + \eta^2}\frac{\partial \eta}{\partial b}$$

From eq. (3.11) we have

$$a\left(2\xi\frac{\partial \xi}{\partial b} + 2\eta\frac{\partial \eta}{\partial b}\right) + \xi + b\frac{\partial \xi}{\partial b} = 0$$

By our assumptions on the type of Ω, $\partial\Omega$ is composed of horizontal and vertical lines or line segments (in the Euclidean sense). So either, $\frac{\partial \xi}{\partial b} = 0$ (when geodesic intersects at a vertical boundary region) or $\frac{\partial \eta}{\partial b} = 0$ (when geodesic intersects at a horizontal boundary region).

1. If $\frac{\partial \eta}{\partial b} = 0$, $\frac{\partial \xi}{\partial b} = \frac{-\xi}{2a\left(\xi + \frac{b}{2a}\right)}$. $\left|\xi + \frac{b}{2a}\right|$ pertains to the horizontal distance between the center of the circle and the point of intersection of the circle with $\partial\Omega$ at the horizontal boundary region, since $\frac{\partial \eta}{\partial b} = 0$. Hence it is easy to geometrically understand that in the event of $\left|\xi + \frac{b}{2a}\right|$ being very small, the segment of the geodesic inside Ω is very small. So we may reasonably assume that $\left|\xi + \frac{b}{2a}\right|$ is bounded away from 0. So $\left|\frac{\partial \xi}{\partial b}\right| = O\left(a^{-1}\right)$.

2. If $\frac{\partial \xi}{\partial b} = 0$, $\frac{\partial \eta}{\partial b} = \frac{-\xi}{2a\eta}$, so $\left|\frac{\partial \eta}{\partial b}\right| = O\left(a^{-1}\right)$.

3. $\frac{\partial r_i}{\partial b} = \frac{1}{2a}\left(-1 \pm \frac{b}{\sqrt{D}}\right)$. By using eq. (3.13) we have $\left|\frac{\partial r_i}{\partial b}\right| = O\left(a^{-1}\right)$.

Now it follows that $|\partial_b \log(\lambda)| = O\left(a^{-1}\right)$. Similarly for $|\partial_a \log(\lambda)|$ and $|\partial_c \log(\lambda)|$. Hence we get what's needed. □

For large enough a, the variation of ℓ_Ω is bounded. Summing the variation of ℓ_Ω separately over small and large a we get that,

$$|J_\vartheta - I_\vartheta| = O\left(\frac{1}{Y^4\vartheta^3} + \frac{1}{Y}\right) \tag{3.23}$$

Combining eqs. (3.20), (3.21) and (3.23) we have.

$$\begin{aligned} |J - I| &= O\left(\vartheta \log(\vartheta^{-1})\right) + O\left(\vartheta \log(Y)\right) + O\left(\frac{1}{Y^4\vartheta^3} + \frac{1}{Y}\right) \\ &= O\left(Y^{-1}\log(Y)\right) \end{aligned} \tag{3.24}$$

where the last line follows from setting $\vartheta = Y^{-1}$. Now, it remains to compute I. To compute I we appeal to the setup of the binary quadratic form Q and the corresponding form P parameterised by λ that we defined at the beginning. However, instead of treating them as quadratic forms with integer coefficients now they are treated as quadratic forms with real coefficients and positive discriminants less than 1. Note that the geometric facts pertaining to them remain unchanged. We rewrite I as,

$$I = \iiint\limits_{\substack{b^2-4ac\leq 1\\ a>0}} \ell_\Omega(a,b,c)\,da\,db\,dc = \int_0^\infty \left(\iiint\limits_{\substack{b^2-4ac\leq 1\\ a>0\\ \tau\in\Omega}} da\,db\,dc\right)\frac{d\lambda}{\lambda} \tag{3.25}$$

For a fixed λ, the inner integral computes the volume of the domain R_λ that contains $(a,b,c) \in \mathbb{R}^3$, where $a > 0$, $b^2 - 4ac < 1$ and the corresponding τ lies in Ω. To compute this integral, we map the domain R_λ of $Q = [a, b, c]$ to the domain $\tilde{R}$ of $P = [\alpha, \beta, \gamma]$, whose root with positive imaginary part lies in Ω and such that $0 < \alpha\gamma - \beta^2 < 1$. The domain $\tilde{R}$ is independent of λ.

Equating coefficients of P in eq. (3.10) we have $a^{-1}\alpha = \lambda^{-1} + \lambda$, $a^{-1}\beta = -\lambda^{-1}r_1 - \lambda r_2$, $a^{-1}\gamma = \lambda^{-1}r_1^2 + \lambda r_2^2$. So we obtain,

$$d\alpha = (\lambda^{-1} + \lambda)da, \ \frac{d(a^{-1}\beta, a^{-1}\gamma)}{d(r_1, r_2)} = 2(r_1 - r_2)$$

From Q, we have $a^{-1}b = -r_1 - r_2$, $a^{-1}c = r_1 r_2$, so we have

$$\frac{d(a^{-1}b, a^{-1}c)}{d(r_1, r_2)} = r_2 - r_1$$

Using chain rule and other properties of Jacobians we have

$$\frac{d(\alpha, \beta, \gamma)}{d(a, b, c)} = -2(\lambda^{-1} + \lambda)$$

Using change of variable in eq. (3.25) we have,

$$I = \frac{1}{2}\int_0^\infty \frac{d\lambda}{\lambda^2 + 1}\int_{\tilde{R}} d\alpha\, d\beta\, d\gamma = \frac{\pi}{4}\int_{\tilde{R}} d\alpha\, d\beta\, d\gamma$$

To compute the volume of $\tilde{R}$ we change variables from α, β, γ to D, ξ, η. From $\alpha^{-1}\beta = -\xi$, $\alpha^{-1}\gamma = \xi^2 + \eta^2$, $D = \alpha\gamma - \beta^2 = (\alpha\eta)^2$ we obtain,

$$\frac{d(\alpha^{-1}\beta, \alpha^{-1}\gamma)}{d(\xi, \eta)} = -2\eta, \ 2\alpha\eta\, d(\alpha\eta) = dD$$

So we have,

$$\frac{d(\alpha, \beta, \gamma)}{d(D, \xi, \eta)} = -\alpha\eta^{-1} = -D^{1/2}\eta^{-2}$$

Hence,

$$\begin{aligned}
\int_{\tilde{R}} d\alpha\, d\beta\, d\gamma &= \int_0^1 D^{1/2} dD \int_\Omega \frac{d\xi d\eta}{\eta^2} \\
&= \frac{2}{3}\frac{\pi}{3}\int_\Omega \frac{3}{\pi}\frac{d\xi d\eta}{\eta^2} \\
&= \frac{2\pi}{9}\nu(\Omega) \\
\Rightarrow I &= \frac{\pi^2}{18}\nu(\Omega)
\end{aligned}$$

Finally we have by re-substituting $Y = X^{1/2}$,

$$\begin{aligned}
\sum_{C\in\mathcal{G}'_X} ||C\cap\Omega|| &= IY^3 + O\left(Y^2\log(Y)\right) \\
&= \frac{\pi^2}{9}\nu(\Omega)X^{3/2} + O\left(X\log(X)\right) \qquad (3.26)
\end{aligned}$$

Now that we have followed Siegel's proof to obtain an asymptotic formula of the sum of geodesic lengths across all discriminants, next we will restrict this sum to only fundamental discriminants ${}^+\mathcal{D}_X$ or even fundamental discriminants restricted by certain local conditions ${}^+\mathcal{D}_X^{\mathcal{S},\mathcal{R}}$.

3.6.2 Local conditions on the asymptotic

We first compute the asymptotic densities of the following sets.

$$H = \{(a,b,c)\in\mathbb{Z}^3\ ;\ b^2-4ac \in^+ \mathcal{D}\} \qquad (3.27)$$

$$H^{\mathcal{S},\mathcal{R}} = \{(a,b,c)\in\mathbb{Z}^3\ ;\ b^2-4ac \in^+ \mathcal{D}^{\mathcal{S},\mathcal{R}}\} \qquad (3.28)$$

Let $N > 0$ be large, now consider a cube $W_N \subset \mathbb{R}^3$ of size $N\times N\times N$. For integer $n > 1$, let $A_n = \{(a,b,c)\in\mathbb{Z}^3;\quad b^2-4ac\equiv 0 \mod n^2\}$, $B_n = \{(a,b,c)\in\mathbb{Z}^3;\quad b^2-4ac\equiv n \mod n^2\}$ and $C_n = \{(a,b,c)\in\mathbb{Z}^3;\ \ b^2-4ac\equiv 0 \mod n\}$. First we compute the cardinality of $W_N \cap A_p$ for prime p.

Let $p = o(N)$ be an odd prime. We consider three mutually exclusive and exhaustive cases where $p^2 \mid b^2 - 4ac$.

1. $p \nmid b$: There are $\frac{p-1}{p}N + O(1)$ possibilities for b. Since in this case both a and c should is indivisible by p, given b, there are $\frac{p-1}{p}N + O(1)$ possibilities for a, and both a, b, there are only $\frac{1}{p^2}N + O(1)$ possibilities for c. So $|W_N \cap A_p \cap \{p \nmid b\}| = \frac{1}{p^2}\left(1 - \frac{1}{p}\right)^2 N^3 + O(N^2)$in this case.

2. $p \mid b$ and exactly one of a, c is divisible by p : There are $\frac{1}{p}N + O(1)$ possibilities for b. Of a, c suppose $p \mid a$ then $p^2 \mid a$ and $p \nmid c$, so there are $\frac{1}{p^2}N + O(1)$ possibilities for a, and $\frac{p-1}{p}N + O(1)$ possibilities for c. Similarly considering the other sub-case where $p \mid c$ we have have that $|W_N \cap A_p \cap \{p \mid b \text{ and } p \nmid gcd(a, b, c)\}| = \frac{2}{p^3}\left(1 - \frac{1}{p}\right) N^3 + O\left(\frac{N^2}{p}\right)$ in this case.

3. $p \mid gcd(a, b, c)$: Each of a, b, c have $\frac{1}{p}N + O(1)$ possibilities. Hence in this case $|W_N \cap A_p \cap \{p \mid gcd(a, b, c)\}| = \frac{1}{p^3}N^3 + O\left(\frac{N^2}{p^3}\right)$.

Adding up all the cases we have,

$$|W_N \cap A_p| = \left(\frac{1}{p^2} + \frac{1}{p^3} - \frac{1}{p^4}\right) N^3 + O\left(N^2\right) \tag{3.29}$$

Going about with a similar method we have that,

$$|W_N \cap C_p| = \frac{1}{p}N^3 + O\left(N^2\right) \tag{3.30}$$

$$\begin{aligned} |W_N \cap B_p| = |W_N \cap C_p| - |W_N \cap A_p| &= \left(\frac{1}{p} - \frac{1}{p^2} - \frac{1}{p^3} + \frac{1}{p^4}\right) N^3 + O\left(N^2\right) \\ &= s(p)N^3 + O\left(N^2\right) \end{aligned} \tag{3.31}$$

Let $\mathcal{D}_2$ be the $(a, b, c) \in \mathbb{Z}^3$ solutions such that $b^2 - 4ac = D$ is a fundamental discriminant in

$\mathbb{Z}_2$. Now we compute $|W_N \cap \mathcal{D}_2|$. Again we consider three mutually exclusive and exhaustive cases.

1. $b^2 - 4ac \equiv 1 \mod 4$: So $b \equiv 1 \mod 2$ for any a, c. So we have $|W_N \cap \{b^2 - 4ac \equiv 1 \mod 4\}| = \frac{1}{2}N^3 + O\left(N^2\right)$.

2. $b^2 - 4ac \equiv 12 \mod 16$: Then $(a, b, c) \equiv (2, 2, 1), (2, 2, 3), (1, 0, 1) \mod 4$ exhaust all the possibilities. So we have $|W_N \cap \{b^2 - 4ac \equiv 12 \mod 16\}| = \frac{3}{32}N^3 + O\left(N^2\right)$.

3. $b^2 - 4ac \equiv 8 \mod 16$: Then $(a, b, c) \equiv (2, 0, 1), (2, 0, 3), (3, 2, 1) \mod 4$ exhaust all the possibilities. So we have $|W_N \cap \{b^2 - 4ac \equiv 8 \mod 16\}| = \frac{3}{32}N^3 + O\left(N^2\right)$.

Adding all these we have,

$$|W_N \cap \mathcal{D}_2| = \frac{11}{16}N^3 + O\left(N^2\right) \tag{3.32}$$

Now by chinese remainder theorem, for odd square-free n we have,

$$\begin{aligned}
\frac{1}{N^3}|W_N \cap A_n \cap \mathcal{D}_2| &= \left(\frac{1}{N^3}|W_N \cap \mathcal{D}_2|\right) \prod_{p|n} \frac{1}{N^3}|W_N \cap A_p| \\
&= \frac{11}{16} \prod_{p|n} \left(\frac{1}{p^2} + \frac{1}{p^3} - \frac{1}{p^4}\right) + O\left(\frac{1}{Nn^2}\right) \\
\Rightarrow |W_N \cap A_n \cap \mathcal{D}_2| &= \frac{11}{16} \left(\prod_{p|n} \left(\frac{1}{p^2} + \frac{1}{p^3} - \frac{1}{p^4}\right)\right) N^3 + O\left(\frac{N^2}{n^2}\right)
\end{aligned}$$

The set of fundamental discriminant solutions of $b^2 - 4ac$ is equal to $W_N \cap \left(\bigcap_{p \text{ odd prime}} A_p\right) \cap \mathcal{D}_2$. Hence by inclusion-exclusion principle we have the following.

$$\begin{aligned}
|W_N \cap H| &= \sum_{n \text{ odd}} \mu(n) |W_N \cap A_n \cap \mathcal{D}_2| \\
&= \frac{11}{16} \sum_{n \text{ odd}} \mu(n) \prod_{p|n} \left(\frac{1}{p^2} + \frac{1}{p^3} - \frac{1}{p^4}\right) N^3 + O\left(\frac{N^2}{n^2}\right) \\
&= \frac{11}{16} \left(\prod_{p \text{ odd prime}} \left(1 - \frac{1}{p^2} - \frac{1}{p^3} + \frac{1}{p^4}\right)\right) N^3 + O\left(N^2 \sum_{n<N} \frac{1}{n^2}\right) \\
&= \left(\prod_{p \text{ prime}} h(p)\right) N^3 + O\left(N^2\right) \qquad (3.33)
\end{aligned}$$

Similarly we can prove that,

$$\left|W_N \cap H^{\mathcal{S},\mathcal{R}}\right| = \left(\prod_{p \in S} s(p)\right) \left(\prod_{p \in R} r(p)\right) \left(\prod_{p \notin S \cup R} h(p)\right) N^3 + O\left(N^2\right) \qquad (3.34)$$

For $\Lambda = H^{\mathcal{S},\mathcal{R}}$, applying lemma 3.4.2 we prove theorem 3.4.1.

Now we prove lemma 3.4.2.

Proof of lemma 3.4.2. Let Y be chosen as in the previous section. Based on the proof we have seen in the previous section it is sufficient to prove that Λ behaves well with J_ϑ (refer eqs. (3.16) and (3.22). That is we prove that,

$$Y^3 J_\vartheta^\Lambda = \rho Y^3 J_\vartheta + O\left(Y^\omega \log(X)\right)$$

where

$$J_\vartheta^\Lambda = \frac{1}{Y^3} \sum_{\substack{b^2 - 4ac \leq 1 \\ a > \vartheta \\ (Ya, Yb, Yc) \in \Lambda}} \ell_\Omega(a, b, c)$$

Set $R=\{(a,b,c)\in\mathbb{R}^3\ ;\ 0<b^2-4ac\leq 1,\ a>\vartheta,\ (Ya,Yb,Yc)\in\mathbb{Z}^3\}$ and $\tilde{\Lambda}=\frac{1}{Y}\Lambda$. So,

$$Y^3 J_\vartheta^\Lambda=\sum_{(a,b,c)\in R\cap\tilde{\Lambda}}\ell_\Omega(a,b,c)$$

We divide R into (non overlapping) cubes with sides parallel to the standard axes of $\mathbb{R}^3$ and edge lengths parameterized by a subset Ξ of the a coordinates of R. In fact, for a given cube $\mathcal{C}$ the a-coordinate corresponding to its left face parameterize the side length of $\mathcal{C}$. Let N_a be the number of lattice points of R that lie on one edge of a cube $\mathcal{C}$ corresponding to an $a\in\Xi$. The parameterization and hence the choice of Ξ will be specified later. Then, there are N_a^3 number of R-lattice points in $\mathcal{C}$ and its edge length is $O\left(\frac{1}{Y}N_a\right)$. Define F_a as the set that contains all the cubes that correspond to $a\in\Gamma$. Note that $|F_a|=O\left(\left(\frac{Y}{N_a}\right)^2\right)$. Define $\ell_\Omega^{\mathcal{C}}$ as the average ℓ_Ω over $\mathcal{C}$. Since ℓ_Ω continuous we also know that $\ell_\Omega^{\mathcal{C}}=\ell_\Omega(a_{\mathcal{C}},b_{\mathcal{C}},c_{\mathcal{C}})$ for some $(a_{\mathcal{C}},b_{\mathcal{C}},c_{\mathcal{C}})\in\mathcal{C}$.

$$\begin{aligned} Y^3 J_\vartheta^\Lambda &=\sum_{\alpha\in\Xi}\sum_{\mathcal{C}\in F_\alpha}\sum_{(a,b,c)\in\mathcal{C}\cap\Lambda}\ell_\Omega(a,b,c)\\ &=\sum_{\alpha\in\Xi}\sum_{\mathcal{C}\in F_\alpha}\sum_{(a,b,c)\in\mathcal{C}\cap\Lambda}\ell_\Omega^{\mathcal{C}}+\sum_{\alpha\in\Xi}\sum_{\mathcal{C}\in F_\alpha}\sum_{(a,b,c)\in\mathcal{C}\cap\Lambda}\left(\ell_\Omega(a,b,c)-\ell_\Omega^{\mathcal{C}}\right) \qquad (3.35)\\ & \qquad (3.36)\end{aligned}$$

Note that,

$$
\begin{aligned}
&\sum_{\alpha\in\Xi}\sum_{\mathcal{C}\in F_\alpha}\sum_{(a,b,c)\in\mathcal{C}\cap\Lambda}\ell_\Omega^{\mathcal{C}}\\
&=\sum_{\alpha\in\Xi}\sum_{\mathcal{C}\in F_\alpha}\ell_\Omega^{\mathcal{C}}\sum_{(a,b,c)\in\mathcal{C}\cap\Lambda}1\\
&=\sum_{\alpha\in\Xi}\sum_{\mathcal{C}\in F_\alpha}\ell_\Omega^{\mathcal{C}}\left(\rho N_\alpha^3+O\left(N_\alpha^\omega\right)\right)\\
&=\rho\sum_{\alpha\in\Xi}\sum_{\mathcal{C}\in F_\alpha}\ell_\Omega^{\mathcal{C}}N_\alpha^3+\sum_{\alpha\in\Xi}\sum_{\mathcal{C}\in F_\alpha}\ell_\Omega^{\mathcal{C}}O\left(N_\alpha^\omega\right)\\
&=\rho\sum_{(a,b,c)\in R}\ell_\Omega(a,b,c)+\sum_{\alpha\in\Xi}\sum_{\mathcal{C}\in F_\alpha}\ell_\Omega^{\mathcal{C}}O\left(N_\alpha^\omega\right)\\
&=\rho Y^3 J_\vartheta+\sum_{\alpha\in\Xi}\ell_\Omega^{\mathcal{C}}\sum_{\mathcal{C}\in F_\alpha}O\left(N_\alpha^\omega\right)
\end{aligned}
$$

We use the following result proven by Siegel in [24]

$$
\ell_\Omega(a,b,c)<2\log(a^{-1})
$$

Now consider,

$$
\begin{aligned}
&\sum_{\alpha\in\Xi}\ell_\Omega^{\mathcal{C}}\sum_{\mathcal{C}\in F_\alpha}O\left(N_\alpha^\omega\right)\\
&=O\left(\sum_{\alpha\in\Xi}\log(\alpha^{-1})\left(\frac{Y}{N_a}\right)^2N_\alpha^\omega\right)\\
&=O\left(Y^2\sum_{\alpha\in\Xi}N_\alpha^{\omega-2}\log(\alpha^{-1})\right)
\end{aligned}
$$

Let us index (using non-negative integers) the elements in $\alpha\in\Xi$ such that $\alpha_{i+1}>\alpha_i$, where $\alpha_0=\vartheta=Y^{-1}$, and $\alpha_{|\Xi|}=O\left(1\right)$. Now it is clear that $\alpha_{i+1}=\alpha_i+\frac{1}{Y}N_a$, i.e. two α's differ by the edge length of a cube corresponding to the smaller α. Now we specify the set Ξ, by starting at $\alpha_0 Y^{-1}$ and then setting $\alpha_{i+1} - \alpha_i = \frac{2^i}{Y}$, so $\alpha_i = \frac{2^i}{Y}$ and $N_{\alpha_i} = 2^i$. Then,

$$\begin{aligned}
&\sum_{\alpha \in \Xi} \ell_\Omega^{\mathcal{C}} \sum_{C \in F_\alpha} O\left(N_\alpha^\omega\right) \\
&= O\left(Y^\Omega \sum_{n=0}^{\log(Y)} 2^{\omega n} \log\left(\frac{2^n}{Y}\right)\right) \\
&= O\left(Y^\Omega \log(Y)\right)
\end{aligned}$$

Now using lemma 3.6.1 and our choice of Ξ above we have,

$$\begin{aligned}
&\sum_{\alpha \in \Xi} \sum_{\mathcal{C} \in F_\alpha} \sum_{(a,b,c) \in \mathcal{C} \cap \Lambda} \left|\ell_\Omega(a,b,c) - \ell_\Omega^{\mathcal{C}}\right| \\
&= \sum_{\alpha \in \Xi} \sum_{\mathcal{C} \in F_\alpha} \sum_{(a,b,c) \in \mathcal{C}} O\left(\frac{1}{\alpha} \max\left(|a - a_{\mathcal{C}}|, |b - b_{\mathcal{C}}|, |c - c_{\mathcal{C}}|\right)\right) \\
&= \sum_{\alpha \in \Xi} \sum_{\mathcal{C} \in F_\alpha} O\left(\frac{1}{\alpha} \sum_{r=0}^{N_\alpha} \frac{r}{Y}\right) \\
&= \sum_{\alpha \in \Xi} \sum_{\mathcal{C} \in F_\alpha} O\left(\frac{1}{\alpha} \frac{N_\alpha^2}{Y}\right) \\
&= \sum_{\alpha \in \Xi} O\left(\frac{1}{\alpha} \frac{N_\alpha^2}{Y}\right) \sum_{\mathcal{C} \in F_\alpha} 1 \\
&= \sum_{\alpha \in \Xi} O\left(Y \frac{1}{\alpha}\right) \\
&= O\left(Y \sum_{n=0}^{\log(Y)} \frac{Y}{2^n}\right) \\
&= O\left(Y^2\right)
\end{aligned}$$

Combining all of these in eq. (3.35), we prove the lemma. □

Chapter 4

Future Explorations

In this final concluding chapter we discuss briefly on three related directions for future exploration.

4.1 Possible improvements in the distribution results of roots of quadratic congruences

As we have shown in section 2.4.6 going by our methods, proving asymptotics of odd prime order CM points reduces to estimating the discrepancy of $x_{T,R}(f,A)$ (definition 2.4.3), simplified to $x_G^2(A)$ (definition 2.4.14). However, to effectively prove the conjectural results for $k>3$, we would need discrepancy $\mathscr{D}\left(x_G^2(A)\right) = O\left(G^{-(1-2/k)}\right)$, (for $k=5$, $G^{-3/5}$). This seems to be a very far shot from the best results we have seen for $x_G^1(A)$ [2, 10], where $\mathscr{D}\left(x_G^1(A)\right) = O\left(G^{-1/3+\epsilon}\right)$. The only result generalised to suit $x_G^2(A)$ [15, Theorem 1.2], shows $\mathscr{D}\left(x_G^2(A)\right) = O\left(G^{-\epsilon}\right)$. Further investigating $x_G^2(A)$, may at least achieve the results for $k=5$. Although the results Soundararajan and Kowalski [15] give very small power savings, their idea of exploiting the randomness created by CRT seems to be a very natural approach. So, their idea may have some promising prospects.

In reference to eq. (2.69), exploiting the expected random cancellations among the Mobius function may further help. For a bounded function $f:\mathbb{N}\to\mathbb{C}$ Sarnak's conjecture says that,

$$\sum_{n<X} \mu(n) f(n) = o\left(X\right)$$

and is proven by Landau for f constant. It is also known that a much stronger result

$$\sum_{n<X} \mu(n) = O\left(X^{1/2+\epsilon}\right)$$

is equivalent to the Riemann Hypothesis [28]. So assuming RH, we can replace the $\sqrt{X}$ factor in the error term of theorems 2.4.5 and 2.4.7 by $X^{1/4+\epsilon}$. Hence it would be sufficient to prove $\mathscr{D}\left(x_G^2(A)\right) = O\left(G^{-(1-2/(k-1))}\right)$ for ($k = 5$, $G^{-1/2}$). Due such observations improvements in the asymptotics of $\sum_{n<X} \mu(n)$, are also of great interest.

4.2 Investigations into the joint distribution of prime order CM-points

Hough in [11, p. 11] gives an outline of a proof of conjecture 1.2.2, assuming his conjectures on the distribution of prime order CM-points along with some further assumptions on the joint distributions of the k-order CM-points. More generally suppose $f : \Gamma\backslash\mathcal{H} \to \mathbb{C}$ is a 'well-behaved' function. For a fixed positive integer n we are interested in computing the following,

$$\begin{aligned} M_{f,X}^n &= \frac{1}{\left|{}^{-}\mathcal{D}_X\right|} \sum_{D \in {}^{-}\mathcal{D}_X} \left(\sum_{z \in CM(D)[k]} f(z) \right)^n \\ &= \frac{1}{\left|{}^{-}\mathcal{D}_X\right|} \sum_{D \in {}^{-}\mathcal{D}_X} \sum_{z_i \in CM(D)[k]} f(z_1) f(z_2) \cdots f(z_n) \end{aligned}$$

For $f \equiv 1$, $M_{f,X}^n$ gives the n^{th} moments. Conjecture 2.2.1 predicts the case for $n = 1$. For general n we need further knowledge on the joint distribution of prime order CM-points. For example, if they are independent then

$$M_{f,X}^{n} = \left(M_{f,X}^{1}\right)^{n}$$

Based, on Hough's illustration of a possible proof of conjecture 1.2.2 independence should be expected given the CM-points do not come from the same k-cyclic group. Hence, the investigations into possible conjectures on the joint distribution of $CM_X[k]$ has great consequences in relation to the Cohen-Lenstra conjectures.

4.3 Resolving conjectures 3.5.1 to 3.5.4 on asymptotics of $L_X[k]$

As we have seen in section 2.4.3 for imaginary quadratic fields, k-order ideals (ideal classes) of real quadratic fields can be parameterized. The proofs are very similar to those in section 2.4.3. For example, suppose $D > 0$ is a fundamental discriminant such that $D \equiv 8 \mod 16$, set $d = \frac{D}{4}$. Then any k-order ideal (or integral binary quadratic form $Q(x,y) = ax^2 + bxy + cy^2$ such that $a > 0$) can be parameterized as

$$lm^k = l^2n^2 - t^2d,\ (l,m,n,t) \in \mathbb{Z}^4,\ l,m > 0,\ gcd(l,m,n) = 1$$

then

$$a = lm,\ b \equiv lnt^{-1} \mod a,\ \text{and } c = \frac{b^2 - D}{4a}$$

To prove asymptotics involving geodesics of k-order ideal classes, by virtue of lemma 3.4.2 it is sufficient to prove that (a,b,c) defined as above, to be uniformly distributed in $\mathbb{Z}^3$. For the moment, we leave this as possible future work.

Bibliography

[1] M. Bhargava, A. Shankar, and J. Tsimerman. “On the Davenport–Heilbronn theorems and second order terms”. In: *Invent. math.* 193 (2013), pp. 439–499.

[2] V. A. Bykovskii. “Spectral Decompostions of certain Automorphic Functions and their Number-Theoretic Applications”. In: *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov.* 134 (1984), pp. 15–33.

[3] H. Cohen and Jr H. W. Lenstra. “Heuristics on Class Groups of Number Fields”. In: *Number theory, Noordwijkerhout* (1983), pp. 33–62.

[4] David A. Cox. *Primes of the form* $x^2 + ny^2$. 2nd ed. Massachusetts: John Wiley & Sons, 1989.

[5] H. Davenport and H. Heilbronn. “On the density of discriminants of cubic fields. II”. In: *Proc.Roy. Soc. London Ser.* A 322.1551 (1971), pp. 405–420.

[6] UMD DIT. *University of Maryland High Performance Clusters.* `https://hpcc.umd.edu/`.

[7] W. Duke. “Hyperbolic distribution problems and half-integral weight Maass forms”. In: *Invent.Math.* 92 (1988), pp. 73–90.

[8] Eduardo Friedman and Lawrence C. Washington. “On the Distribution of Divisor Class groups of Curves over a Finite field”. In: (1989).

[9] Carl Gauss. *Disquisitiones Arithmeticae.* 1801.

[10] C. Hooley. “On the Number of Divisors of Quadratic Polynomials”. In: *Acta Mathematica* 110 (1963), pp. 97–114.

[11] Robert Hough. “Distribution Problems in Number Theory”. PhD thesis. Stanford University, 2012.

[12] Robert Hough. “Equidistribution of bounded torsion CM points”. In: *JAMA* 138 (2019), pp. 765–797.

[13] Henrryk Iwaniec and Emmanuel Kowalski. *Analytic Number Theory.* American Mathematical Society, 2004.

[14] Lillian B. Pierce Jordan Ellenberg and Melanie Matchett Wood. “On ℓ-torsion of Class Groups of Number Fields”. In: *Algebra & Number Theory* 11 (2017), pp. 1739–1778.

[15] E. Kowalski and K. Soundararajan. “Equidistribution from the Chinese Remainder Theorem”. In: *Advances in Mathematics* 385 (2021).

[16] L. Kuipers and H. Niederreiter. *Uniform Distribution of Sequences*. New York: John Wiley & Sons, 1974.

[17] Franz Lemmermeyer. *Binary Quadratic Forms*. 2010.

[18] Caroline L. Turnage-Butterbaugh Lillian B. Pierce and Melanie Matchett Wood. "On a Conjecture for ℓ-torsion in Class Groups of Number Fields from the perspective of Moments". In: *Mathematical Research Letters* 28 (2021), pp. 575–621.

[19] Yu. V. Linnik. *Ergodic Properties of Algebraic Fields*. Trans. by M. S. Keane. New York: Springer-Verlag, 1968.

[20] The LMFDB Collaboration. *The L-functions and modular forms database*. `https://www.lmfdb.org`.

[21] Jens Marklof and Matthew Welsh. "Fine-scale distribution of roots of quadratic congruences". In: (2022).

[22] James S. Milne. *Algebraic Number Theory (v3.08)*. Available at www.jmilne.org/math/. 2020.

[23] J. Kułaga-Przymus S. Ferenczi and M. Lemańczyk. "Sarnak's Conjecture – what's new". In: *Advances in Mathematics* (2017).

[24] C. L. Siegel. "The Average Measure of Quadratic Forms with given Determinant and Signature". In: *Annals of Mathematics* 45 (1944).

[25] B. F. Skubenko. "The Asymptotic Distribution of Integers on a Hyperboloid of One Sheet and Ergodic Theorems". In: *Izv. Akad. Nauk SSSR Ser. Mat.* 26 (1962), pp. 721–752.

[26] K. Soundararajan. "Divisibility of Class Numbers of Imaginary Quadratic Fields". In: *Journal of the London Mathematical Society* 61 (2000), pp. 681–690.

[27] Takashi Taniguchi and Frank Thorne. "Secondary terms in counting functions for cubic fields". In: *Duke Math. J.* 162.13 (2013), pp. 2451–2508.

[28] Terence Tao. *The Chowla conjecture and the Sarnak conjecture*. `https://terrytao.wordpress.com/2012/10/14/the-chowla-conjecture-and-the-sarnak-conjecture/`.

[29] I. M. Vinogradov. *An Introduction to the Theory of Numbers*. London & New York: Pergamon Press, 1955.

[30] Melanie Matchett Wood. "Random Integral Matrices and the Cohen Lenstra Heuristics". In: *American Journal of Mathematics* 141 (2019), pp. 383–398.