Simultaneous prime values of binary forms

by

Cho Ho Lam

M.Sc., University of Hong Kong, 2014
B.Sc., Chinese University of Hong Kong, 2012

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department of Mathematics
Faculty of Science

SIMON FRASER UNIVERSITY
Summer 2019

Approval

Name:	**Cho Ho Lam**
Degree:	**Doctor of Philosophy (Mathematics)**
Title:	**Simultaneous prime values of binary forms**
Examining Committee:	**Chair:** Petr Lisonek Professor
	Stephen Choi Senior Supervisor Professor
	Peter Borwein Co-supervisor Professor Emeritus
	Imin Chen Supervisory Committee Associate Professor
	Nils Bruin Internal Examiner Professor
	Greg Martin External Examiner Professor Department of Mathematics University of British Columbia
Date Defended:	**August 16, 2019**

Abstract

The twin prime conjecture asserts that there are infinitely many positive integers x such that x and $x+2$ are simultaneously prime. In this thesis, we consider a two-variable analogue of this problem. Let $F(x,y)$ be a positive definite quadratic form and $G(x,y)$ a linear form, both with integer coefficients. Suppose for any prime p there exist ℓ, m such that $p \nmid F(\ell,m)G(\ell,m)$. Then we prove that there are infinitely many $\ell, m \in \mathbb{Z}$ such that both $F(\ell,m), G(\ell,m)$ are primes. In fact, our proof extends to primes in arithmetic progressions $F(\ell,m) \equiv a \pmod{q}$ and $G(\ell,m) \equiv b \pmod{q}$.

The main result (when $q = 1$) was first obtained independently by the author and another team of researchers D. Schindler and S. Y. Xiao. The extension is joint work with them.

Keywords: primes, binary quadratic forms, simultaneous prime values

Dedication

I am dedicating my dissertation to my parents, Wai-Keung Lam and Siu-Ling Mak, for their unyielding love and support. As I am the only child in the family, it was a difficult decision for them to let me pursue my study in Canada. I will never forget the sacrifice they have made for me.

I also dedicate my thesis to my girlfriend, Karen Ho, who has been the source of encouragement and support throughout my study. Pursuing a Ph.D. degree is a challenging experience for many people. I am grateful to have her walking alongside me on this journey.

Lastly, I dedicate this thesis to the Almighty God, who is the reason why I wanted to become a mathematician. The more mathematics I learn, the more I am amazed by His beautiful creation.

Acknowledgements

I would like to express my deepest gratitude to my senior supervisor, Professor Stephen Choi, for his encouragement and support throughout these five years. He has been very supportive of my research and has provided valuable advice on research, writing, networking and career development. There were some difficult times in my Ph.D. study and he was always there to help me out. I am eternally grateful for everything he has done for me.

I would like to thank my co-supervisor, Professor Peter Borwein, for supporting my study financially through his research grants.

I would also like to thank my supervisory committee member, Professor Imin Chen, for his support and guidance in my job applications.

I am deeply indebted to Professor Andrew Granville for his help on my dissertation. He pointed me to a result of Harald Helfgott, which is the last missing piece of the proof of my main result.

I am grateful to my collaborators on this project, Professor Damaris Schindler and Dr. Stanley Yao Xiao, for their trust and patience.

Finally, I would like to thank everyone in the Department of Mathematics at Simon Fraser University, including faculties, students and staffs. I had a lot of memorable moments here and I am honored to be a part of this community.

Table of Contents

Notations

$\Lambda(n)$ = von-Mangoldt function.
$\mu(n)$ = Möbius function.
$\tau(n) = \sum_{d|n} 1$ denotes the divisor functions.
$\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ denote the sets of all positive integers, all integers, all rational numbers, all real numbers and all complex numbers respectively.
$f(x) = O(g(x))$ means $|f(x)| \leqslant Cg(x)$ for $x \geq x_0$ and some absolute constant $C > 0$. Here $f(x)$ is a complex function of the real variable x and $g(x)$ is a positive function for $x \geq x_0$.
$\sum^{\flat}$ is a summation over positive squarefree integers.

Chapter 1

Introduction

1.1 Main Results

Many classical problems in number theory concern the infinitude of primes in some special subset of integers S. For example, the twin prime conjecture asserts that there are infinitely many primes in the set $S = \{p_n + 2 \mid n \in \mathbb{N}\}$, where p_n denotes the n-th prime. Landau's problem asks the same question for the set $S = \{n^2 + 1 \mid n \in \mathbb{N}\}$.

In this thesis, we are interested in studying pairs of primes. Another way to phrase the twin prime conjecture is that there are infinitely many n such that n and $n + 2$ are both primes. We are pretty far away from proving this and therefore we consider a two-variable analogue instead. For instance, are there infinitely many $x, y \in \mathbb{Z}$ such that both

$$x, \qquad x + 2y$$

are primes? The answer is obvious indeed: any pair of odd primes can be represented by them. For the higher degree cases, we wish to answer the following question.

Question 1. Given two binary forms $F(x, y), G(x, y) \in \mathbb{Z}[x, y]$, are there infinitely many $\ell, m \in \mathbb{Z}$ such that both $F(\ell, m)$ and $G(\ell, m)$ are primes?

In 1997, Fouvry and Iwaniec [5] studied the primes of the form $p = x^2 + y^2$ where y is also a prime, and they managed to prove that the number of these primes is infinite. In other words, the answer to Question 1 is affirmative when $F(x, y) = x^2+y^2$ and $G(x, y) = y$. In the Master's thesis [15] of the author, the case $F(x, y) = x^2 + 2y^2$ and $G(x, y) = y$ is settled. In [17] Pandey studied the case when $F(x, y) = x^2 - xy + y^2$ and $G(x, y) = 2x - y$. The author [16] also treated the case $F(x, y) = x^2 + Dy^2, G(x, y) = y$ when $D > 0$ and there is exactly one binary quadratic form of discriminant $-4D$ up to proper equivalence (see Chapter 4 for terminology on binary quadratic forms). These are the only known cases and we could make the following conjecture in general:

Conjecture 1.1.1. Let $F(x,y), G(x,y) \in \mathbb{Z}[x,y]$ be two irreducible binary forms. Assume that for every prime p there are $\ell, m \in \mathbb{Z}$ such that $p \nmid F(\ell,m)G(\ell,m)$. Then there exist infinitely many $\ell, m \in \mathbb{Z}$ such that both $F(\ell,m)$ and $G(\ell,m)$ are primes.

The assumption about p is very helpful. For example, for the pair of binary forms $F(x,y) = 2x^2 + xy + y^2$ and $G(x,y) = x$, we have $2|F(\ell,m)G(\ell,m)$ for all $\ell, m \in \mathbb{Z}$. Therefore, the prime pairs they could represent form a much thinner set (one of them has to be 2), and this makes counting these pairs much harder. If a prime p divides $F(\ell,m)G(\ell,m)$ for all $\ell, m \in \mathbb{Z}$, we say that we have a **local obstruction** at p.

By a simple change of variable, we can reduce to the case $G(x,y) = y$. If $G(x,y) = ax+by$ is a linear form with $\gcd(a,b) = 1$, then there exist integers c, d with $ad - bc = 1$. By setting $u = ax + by$ and $v = cx + dy$, we then have $x = du - bv, y = -cu + av$ and therefore

$$F(x,y) = F(du - bv, -cu + av)$$

is a binary quadratic form in u and v, which is still positive definite and irreducible.

To state our main result (Theorem 1.1.3), we need to introduce two more definitions. A binary quadratic form $F(x,y) = \alpha x^2 + \beta xy + \gamma y^2 \in \mathbb{Z}[x,y]$ is **primitive** if $\gcd(\alpha, \beta, \gamma) = 1$. For any positive integer d, we define $\rho(d)$ to be the number of solutions $\nu \pmod d$ to the congruence equation

$$F(1,\nu) \equiv 0 \pmod d.$$

Then we have the following proposition.

Proposition 1.1.2. Let $F(x,y) = \alpha x^2 + \beta xy + \gamma y^2 \in \mathbb{Z}[x,y]$ be a primitive positive definite quadratic form and X be a positive real number. Let $(\lambda(\ell))$ be a sequence of complex numbers supported on the natural numbers which satisfies the bound $|\lambda(\ell)| \leqslant \log^A X$ for all $\ell \in \mathbb{N}$ and some fixed $A > 0$. Suppose $q \in \mathbb{N}$, $q \leqslant (\log X)^Q$ for some $Q > 0$ and χ is a Dirichlet character modulo q. Then for any $B > 0$ we have

$$\sum_{F(\ell,m)\leqslant X} \sum \lambda(\ell)\chi(F(\ell,m))\Lambda(F(\ell,m)) = \sum_{\substack{F(\ell,m)\leqslant X \\ \gcd(\ell,\gamma m)=1 \\ \gcd(F(\ell,m),P_F)=1}} \sum \lambda(\ell)\chi(F(\ell,m))H_{F,q}(\ell) + O_{A,B,F,Q}(X(\log X)^{-B}),$$

where Λ is the von Mangoldt function,

$$H_{F,q}(\ell) = \prod_{p \nmid \ell q P_F} \left(1 - \frac{\rho(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-1} \prod_{p | \ell q P_F} \left(1 - \frac{1}{p}\right)^{-1},$$

and P_F is a positive integer that depends only on F (defined in (5.6)).

$H_{F,q}(\ell)$ is absolutely convergent by Lemma 2.1.1. Since $\rho(p) \leqslant 2$ for all prime p, $H_{F,q}(\ell)$ is positive if $\rho(2) \neq 2$. That means to check the local obstruction it suffices to consider $p = 2$. The flexibility provided by $\lambda(\ell)$ and $F(\ell, m) \equiv a \pmod q$ has applications in other prime-related problems too. Recently, Grimmelt proved Vinogradov's three primes theorem for Fouvry-Iwaniec primes and his proof needs such flexibility; see [9] for details.

The purpose of introducing P_F in our expression is to remove some small prime factors that forbids the use of a factorization proposition (see Proposition 4.2.4 in Section 4.2). By choosing $(\lambda(\ell))$, one can show that

Theorem 1.1.3 (Main Theorem)**.** Let $F(x, y) \in \mathbb{Z}[x, y]$ be a primitive positive definite quadratic form of discriminant $-\Delta$. Suppose $q \in \mathbb{N}$ and $q \leqslant (\log X)^Q$ for some $Q > 0$. Then for any $A > 0$ and $a, b \in \mathbb{Z}$ with $\gcd(ab, q) = 1$ we have

$$\sum_{\substack{F(\ell,m)\leqslant X\\ F(\ell,m)\equiv a \pmod q\\ \ell\equiv b \pmod q}}\sum \Lambda(\ell)\Lambda(F(\ell, m)) = \frac{H_q\rho(q; a, b)}{q\phi(q)}\frac{\pi X}{\sqrt{|\Delta|}} + O_{A,F,Q}(X(\log X)^{-A})$$

where

$$H_q = \prod_{p\nmid q}\left(1 - \frac{\rho(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-1}\prod_{p|q}\left(1 - \frac{1}{p}\right)^{-1}$$

and

$$\rho(d; a, b) = \sharp\{\nu \pmod d : F(b, \nu) \equiv a \pmod d\}.$$

When $q = 1$, we have

Corollary 1.1.4. Let $F(x, y) \in \mathbb{Z}[x, y]$ be a primitive positive definite quadratic form of discriminant $-\Delta$. Then for any $A > 0$ we have

$$\sum_{F(\ell,m)\leqslant X}\sum \Lambda(\ell)\Lambda(F(\ell, m)) = \frac{H\pi X}{\sqrt{|\Delta|}} + O_{A,F}(X(\log X)^{-A})$$

where

$$H = \prod_{p}\left(1 - \frac{\rho(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-1}.$$

As we promised before, we provide a new family that confirms Conjecture 1.1.1.

Corollary 1.1.5. Let F be a positive definite binary quadratic form and G be a binary linear form. Assume that for every prime p there are $\ell, m \in \mathbb{Z}$ such that $p \nmid F(\ell, m)G(\ell, m)$. Then there exist infinitely many $\ell, m \in \mathbb{Z}$ such that both $F(\ell, m)$ and $G(\ell, m)$ are primes.

To prove Proposition 1.1.2 (and hence Theorem 1.1.3 and the two corollaries), we will mostly follow the approach by Fouvry and Iwaniec in [5]. Their proof used the asymptotic sieve, and we provide an overview of sieve methods in Section 1.2 to 1.4. In Section 1.5, we present the outline of the proof of Proposition 1.1.2 and its consequences.

1.2 Sieve Methods

To count primes in integer sequences, one of the most powerful tools we have is the sieve method. Suppose (a_n) is a sequence of non-negative real numbers that is supported on S, and $\mathcal{P}(n)$ is defined by

$$\mathcal{P}(n) = \begin{cases} 1 & \text{if } n \text{ is prime,} \\ 0 & \text{otherwise.} \end{cases} \tag{1.1}$$

Then it suffices to deduce an asymptotic formula for

$$P(x) = \sum_{n \leqslant x} a_n \mathcal{P}(n). \tag{1.2}$$

If $P(x) \to \infty$ as $x \to \infty$, then this confirms that S contains infinitely many primes. Unfortunately, the weight $\mathcal{P}(n)$ lacks useful analytic or combinatorial properties and this hinders us from obtaining an asymptotic formula for $P(x)$. The magic of sieve methods comes from modifying $\mathcal{P}(n)$ to some better weight $\widetilde{\mathcal{P}}(n)$, but the cost is to partially sacrifice the prime-detecting property of it. Also, we might not be able to obtain an asymptotic formula but rather an upper bound or a lower bound for $P(x)$. Typically, the modified weight takes the form

$$\widetilde{\mathcal{P}}(n) = \sum_{\substack{d \mid n \\ d \in \mathcal{D}}} f(n) \tag{1.3}$$

where f is some nice arithmetical function and $\mathcal{D}$ is some relatively simple subset of positive integers. For example, $\mathcal{D}$ might contain only integers with a small number of prime factors, or integers less than a small power of x. For simplicity, we take $\mathcal{D} = \mathbb{N}$ for the rest of our discussion.

In this thesis, we will mostly work with the "asymptotic sieve", in which the sum

$$P(x) = \sum_{n \leqslant x} a_n \Lambda(n) \tag{1.4}$$

is considered. Here $\Lambda(n)$ is the von-Mangoldt function

$$\Lambda(n) = \begin{cases} \log p & \text{if } n \text{ is a perfect power of a prime } p, \\ 0 & \text{otherwise.} \end{cases}$$

Strictly speaking, Λ is supported on prime powers. However, the contribution from proper prime powers to $P(x)$ is usually negligible. Furthermore, Λ satisfies a fundamental identity

$$\Lambda(n) = -\sum_{d \mid n} \mu(d) \log d \tag{1.5}$$

where μ is the Möbius function

$$\mu(n) = \begin{cases} (-1)^j & \text{if } n \text{ is a product of } j \text{ distinct primes,} \\ 0 & \text{otherwise.} \end{cases}$$

A particularly important example of (1.4) is when $a_n = 1$ for all $n \in \mathbb{N}$, and the sum

$$P(x) = \sum_{n \leqslant x} \Lambda(n)$$

is known to be asymptotically equal to x by the prime number theorem. The complex-analytic proof of this statement begins with the following identity of Dirichlet series:

$$\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)}, \qquad \mathrm{Re}(s) > 1,$$

where $\zeta(s)$ is the well-known Riemann Zeta function. This identity is essentially (1.5), which provides us another reason to work with Λ.

Then by applying (1.5) in (1.4) and interchanging the order of summation,

$$\sum_{n \leqslant x} a_n \Lambda(n) = -\sum_{n \leqslant x} a_n \sum_{d \mid n} \mu(d) \log d = -\sum_{d \leqslant x} \mu(d)(\log d) \Bigg(\sum_{\substack{n \leqslant x \\ n \equiv 0 \pmod{d}}} a_n \Bigg).$$

Such interchange is always possible as long as our weight takes the form (1.3). For any $d \in \mathbb{N}$, we define

$$A_d(x) = \sum_{\substack{n \leqslant x \\ n \equiv 0 \pmod{d}}} a_n.$$

This quantity plays a crucial role in sieve theory. In most cases, we can provide an approximation

$$A_d(x) = g(d)A(x) + R_d(x) \tag{1.6}$$

where $A(x)$ depends only on x and g is a multiplicative function called the density function. One can regard $g(d)$ as the weighted probability that an element of S is divisible by d and $R_d(x)$ is the error that arises in this approximation. Assuming sufficient regularity conditions on $g(d)$ and some estimate of the remainder $R_d(x)$ on average, we expect that

$$\sum_{n \leqslant x} a_n \Lambda(n) \approx HA(x) \tag{1.7}$$

where

$$H = \prod_p (1 - g(p)) \Big(1 - \frac{1}{p}\Big)^{-1} = -\sum_d \mu(d) g(d) \log d.$$

This is because

$$\begin{aligned}\sum_{n\leqslant x} a_n\Lambda(n) &= -\sum_{d\leqslant x}\mu(d)(\log d)A_d(x)\\ &= \Big(-\sum_{d\leqslant x}\mu(d)g(d)\log d\Big)A(x) + O\Big(\sum_{d\leqslant x}|R_d(x)|\log d\Big)\\ &\approx \Big(-\sum_{d}\mu(d)g(d)\log d\Big)A(x).\end{aligned}$$

For most sequences (a_n) that arise in prime number theory, (1.7) actually agrees with the conjectural asymptotic formula. This is a very attractive feature of the asymptotic sieve (1.4). But compared to other sieve weights, much stronger arithmetical inputs are required here to make the above argument rigorous.

Define

$$R(x;D) = \sum_{d\leqslant D}|R_d(x)|. \tag{1.8}$$

One might attempt to show that $R(x;x)$ is small compared to the main term $HA(x)$. In practice, such an estimate is probably not available. The best we can hope for is to handle all d less than some threshold D, say $D = x^{1-\varepsilon}$, and apply something else to handle the remaining d. Precisely,

$$\sum_{n\leqslant x} a_n\Lambda(n) = -\sum_{\substack{d\leqslant x\\ d\leqslant D}}\mu(d)(\log d)A_d(x) - \sum_{\substack{d\leqslant x\\ d>D}}\mu(d)(\log d)A_d(x).$$

The first part can be tackled using a good estimate for $R(x;D)$; this is usually referred as a **Type I estimate**. For the second part, it leads us to

$$\sum_{\substack{d\leqslant x\\ d>D}}\mu(d)(\log d)A_d(x) = \sum_{\substack{mn\leqslant x\\ m>D}}\mu(m)a_{mn}\log m$$

and this is a sum over two variables m, n in which some of the variables can be quite large $(m > D)$. What we need here is the **Type II estimate**. Therefore to apply the asymptotic sieve, we need the following two major arithmetical inputs:

Question 2 (Type I Estimate)**.** Provide an adequate upper bound for

$$R(x;D) = \sum_{d\leqslant D}|R_d(x)|$$

for D as large as possible.

Question 3 (Type II Estimate)**.** Provide an adequate upper bound for

$$\sum_{m\sim M}\sum_{n\sim N}\alpha_m\beta_n a_{mn}$$

for M, N as wide as possible. Here $m \sim M$ means $M < m \leqslant 2M$ and similarly for $n \sim N$.

In the next two sections, we will explain more on how to obtain these estimates.

Remark 1.2.1. In fact, the weights $\mathcal{P}(n)$ also take the form (1.3) if we put $f(d) = \mu(d)$ and

$$\mathcal{D} = \{d \in \mathbb{N} \mid d \leqslant x \text{ and all prime factors of } d \text{ is less than } \sqrt{x}\}.$$

But the structure of $\mathcal{D}$ here is not simple enough for the application of sieve method.

1.3 Type I Estimate

To estimate $R(x; D)$, we first need to figure out what are $g(d), A(x)$ and $R_d(x)$ in the approximation (1.6). As an example, when $a_n = 1$, we have

$$A_d(x) = \sum_{\substack{n\leqslant x\\ n\equiv 0 \pmod d}} 1 = \left\lfloor \frac{x}{d} \right\rfloor$$

where $\lfloor\cdot\rfloor$ is the floor function. Therefore we can simply take $g(d) = 1/d, A(x) = x$ and $R_d(x) = -\{x/d\}$ ($\{\cdot\}$ is the fractional part function). This agrees with our intuition since we expect an integer is divisible by d with probability $1/d$. By the way, the main term in (1.7) is

$$HA(x) = A(x)\prod_p (1 - 1/p)\left(1 - \frac{1}{p}\right)^{-1} = x.$$

This is consistent with the prime number theorem.

In general, $A(x)$ is simply $A_1(x) = \sum_{n\leqslant x} a_n$, or an approximation of it. The density function, $g(d)$, can be determined by some probabilistic heuristics. One can therefore say that $R_d(x)$ is uniquely determined by the expression $A_d(x) - g(d)A(x)$, but this does not provide us a feasible way to estimate it. To obtain an alternative representation for $R_d(x)$, one popular approach is to remove the condition $n \equiv 0 \pmod d$ by using some arithmetic identities.

For example, in the twin prime problem, we can take $a_n = \Lambda(n + 2)$. By the prime number theorem, we can set $A(x) = x$. For even d, it is necessary to set $g(d) = 0$ since the condition $n \equiv 0 \pmod d$ would imply $n + 2$ is not a prime. If $\gcd(2, d) = 1$, then $n + 2$ should belong to one of the $\phi(d)$ congruence classes that are coprime to d. Thus we should take $g(d) = 1/\phi(d)$ in this case. The remainder term $R_d(x)$ can be obtained by using

orthogonality of Dirichlet character:

$$\begin{aligned} A_d(x) &= \sum_{\substack{n\leqslant x\\ n\equiv 0 \pmod d}} \Lambda(n+2) \\ &= \sum_{\substack{n\leqslant x-2\\ n\equiv -2 \pmod d}} \Lambda(n) \\ &= \frac{1}{\phi(d)} \sum_{n\leqslant x-2} \Lambda(n) \Big(\sum_{\chi \pmod d} \chi(n)\overline{\chi}(-2) \Big) \\ &= \frac{1}{\phi(d)} \sum_{\chi \pmod d} \overline{\chi}(-2) \sum_{n\leqslant x-2} \chi(n)\Lambda(n). \end{aligned}$$

The principal character $\chi = \chi_0$ would provide the main term (up to a small error)

$$\frac{1}{\phi(d)} \sum_{\substack{n\leqslant x-2\\ \gcd(n,d)=1}} \Lambda(n) \approx \frac{x}{\phi(d)},$$

since $\chi_0(-2) = 1$. To complete the Type I estimate it suffices to estimate the contribution from non-principal characters. This is essentially how the classical Bombieri-Vinogradov theorem was proved.

In Landau's problem, we can take

$$a_n = \begin{cases} 1 & \text{if } n = m^2+1 \text{ for some } m \in \mathbb{Z}, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly we can set $A(x) = \sqrt{x}$. Let $\rho(d)$ be the number of solutions $\nu \pmod d$ to $\nu^2 + 1 \equiv 0 \pmod d$. Then in every d consecutive integers n exactly $\rho(d)$ of them satisfy $n^2 + 1 \equiv 0 \pmod d$. This suggests that $g(d) = \rho(d)/d$. By the orthogonality of additive characters,

$$\frac{1}{d} \sum_{m=0}^{d-1} e^{\frac{2\pi i m n}{d}} = \begin{cases} 1 & \text{if } n \equiv 0 \pmod d, \\ 0 & \text{otherwise,} \end{cases}$$

we deduce that

$$\begin{aligned} \sum_{\substack{n\leqslant x\\ n\equiv 0 \pmod d}} a_n &= \sum_{\substack{\nu \pmod d\\ \nu^2+1\equiv 0 \pmod d}} \sum_{\substack{m\leqslant \sqrt{x-1}\\ m\equiv \nu \pmod d}} 1 \\ &= \frac{1}{d} \sum_{\substack{\nu \pmod d\\ \nu^2+1\equiv 0 \pmod d}} \sum_{k=0}^{d-1} \sum_{m\leqslant \sqrt{x-1}} e^{\frac{2\pi i k(m-\nu)}{d}}. \end{aligned}$$

The contribution when $k = 0$ gives $\rho(d)\lfloor\sqrt{x-1}\rfloor/d$, which is roughly $g(d)A(x)$. To complete the Type I estimate, one needs to handle several exponential sums (Kloosterman sums in particular). Iwaniec [14] used this approach to show that there exist infinitely many $n \in \mathbb{N}$ such that $n^2 + 1$ has at most two prime factors.

After obtaining a nice expression for $R_d(x)$, we can proceed to estimate $R(x; D) = \sum_{d \leqslant D} |R_d(x)|$ for D as large as possible. Needless to say, obtaining a good Type I estimate is not an easy task; but over the years, mathematicians have developed many powerful analytic tools for this purpose. For instance, in the twin prime problem, the record is

$$R(x; x^{1/2-\varepsilon}) \ll_A x(\log x)^{-A}$$

for any $A > 0$ by the Bombieri-Vinogradov theorem. Although this is still far away from what we believe (the Elliott-Halberstam conjecture implies the same bound for $D = x^{1-\varepsilon}$), the above estimate is already essentially the Generalized Riemann Hypothesis on average. The topic in the next section, the Type II estimate, is a relatively new member of the arena of sieve theory and it presents a much harder challenge to us.

1.4 Type II Estimate

If our sequence is "multiplicative", say it can be split into $a_{mn} = c_m d_n$, then the Type II sum is simply the product of two simpler sums

$$\sum_{m\sim M}\sum_{n\sim N} \alpha_m\beta_n a_{mn} = \Big(\sum_{m\sim M} \alpha_m c_m\Big)\Big(\sum_{n\sim N} \beta_n d_n\Big). \tag{1.9}$$

We can then treat them separately using some standard techniques in analytic number theory. Otherwise, essentially the only known way to handle a Type II sum is via the Cauchy-Schwarz inequality:

$$\begin{aligned}\sum_{m\sim M}\sum_{n\sim N} \alpha_m\beta_n a_{mn} &= \sum_{m\sim M} \alpha_m \sum_{n\sim N} \beta_n a_{mn} \\ &\leqslant \Big(\sum_{m\sim M} |\alpha_m|^2\Big)^{\frac{1}{2}} \Big(\sum_{m\sim M}\Big|\sum_{n\sim N} \beta_n a_{mn}\Big|^2\Big)^{\frac{1}{2}}.\end{aligned} \tag{1.10}$$

The first part can be handled trivially but the second one is much messier. To handle it, we require some additional structure in a_{mn}. For example, if $a_n = r(n)$ is the number of ways to write n as a sum of two squares, then

$$r(mn) = \frac{1}{4} r(m)r(n) \qquad \text{when } \gcd(m, n) = 1.$$

This is basically Dirichlet composition of binary quadratic forms,

$$(a^2+b^2)(c^2+d^2) = (ac+bd)^2 + (ad-bc)^2,$$

or equivalently, the multiplicativity of the norm in $\mathbb{Z}[i]$,

$$N(a+bi)N(c+di) = N((a+bi)(c+di)).$$

This allows us to turn the last double sum in (1.10) into something close to (1.9), with some mild dependence between m and n. Unfortunately, such a factorization only exists for very few sequences. But this is almost a must for the successful estimation of the Type II sum.

For example, we know that there are infinitely many primes of the form $x^3+y^3+z^3$; but the proof is by showing that there are infinitely many primes of the form x^3+2y^3. The former does not have any such factorization structure we can exploit; on the other hand, x^3+2y^3 is simply the norm of $x+y\sqrt[3]{2}$ in $\mathbb{Q}[\sqrt[3]{2}]$ and it comes with a multiplicative structure. Surprisingly, adding an extra condition $y=z$ in $x^3+y^3+z^3$ actually makes the problem easier. See [11] for more details.

1.5 Outline of the Proof

In [5], Fouvry and Iwaniec considered the sum $P(x)$ in (1.2) with

$$a_n = \sum_{\substack{l\in\mathbb{N},m\in\mathbb{Z}\\ l^2+m^2=n}}\sum \lambda_l \tag{1.11}$$

where (λ_l) is a sequence of complex numbers that satisfies $|\lambda_l| \leqslant 1$. Eventually, they took

$$\lambda_l = \frac{\Lambda(l)}{\log l}$$

so that only those $n = l^2+m^2$ with l prime are counted (again, the contribution from higher prime powers is negligible). We will define our sequence (a_n) in a similar way with our own modifications.

Let $F(x,y) = \alpha x^2+\beta xy+\gamma y^2 \in \mathbb{Z}[x,y]$ be an irreducible binary quadratic form (not necessarily positive definite) and $q \leqslant (\log X)^Q$. We will work with the sequence

$$a_N = \sum_{\substack{\ell\in\mathbb{N},m\in\mathbb{Z}\\ F(\ell,m)=N\\ \gcd(\ell,\gamma m)=1}}\sum \lambda(\ell), \tag{1.12}$$

where $(\lambda(\ell))$ is a sequence of complex numbers. Here the sequence is allowed to depend on $F(x,y), X$ or even q. In particular, by choosing the support of λ, the above sum will be a finite sum. For any Dirichlet character χ of modulus q, define the sum

$$\begin{aligned} P(X;\chi) &= \sum_{\substack{N\leqslant X \\ \gcd(N,P_F)=1}} a_N\chi(N)\Lambda(N) \\ &= \sum_{\substack{\ell\in\mathbb{N},m\in\mathbb{Z} \\ F(\ell,m)\leqslant X \\ \gcd(\ell,\gamma m)=1}}\sum \lambda(\ell)\chi(F(\ell,m))\Lambda(F(\ell,m)). \end{aligned} \tag{1.13}$$

The character χ is present to detect the congruence condition modulo q. In Chapter 2, we will decompose the sum $P(X;\chi)$ using Proposition 9 from Section 7 of [5]. The main term ($HA(x)$ in (1.7)) will become apparent but two remainders $R(X;Y,Z;\chi)$ and $B(X;Y,Z;\chi)$ still need to be estimated. Their definitions are given in Section 2.3. $R(X;Y,Z;\chi)$ corresponds to our Type I estimate and will be treated in Chapter 3. In Chapter 4, we will cover background of binary quadratic forms, as well as the proof of a crucial factorization proposition. This factorization proposition, Proposition 4.2.4, is the most novel part of the thesis. We will use it to deal with $B(X;Y,Z;\chi)$ in Chapter 5. When all these ingredients are ready, we will prove Proposition 1.1.2, Theorem 1.1.3, Corollary 1.1.4 and Corollary 1.1.5 in Chapter 6.

Chapter 2

Setting up the sieve

2.1 Basic Properties of a_N

Let $X > 0$ and $F(x, y) = \alpha x^2 + \beta xy + \gamma y^2 \in \mathbb{Z}[x, y]$, where F is not necessarily positive definite at this moment. For any positive integer $N \in [0, X]$, we defined the sequence

$$a_N = \sum_{\substack{\ell \in \mathbb{N}, m \in \mathbb{Z} \\ F(\ell,m)=N \\ \gcd(\ell,\gamma m)=1}} \sum \lambda(\ell)$$

in (1.12). When F is indefinite, the sum is not well-defined since there could be infinitely many $\ell, m \in \mathbb{N}$ such that $F(\ell, m) = N$. Therefore we restrict λ so that its support lies in $[0, C_1\sqrt{X}]$, where C_1 is a positive constant that depends only on F. When F is positive definite, such restriction is in place automatically: the inequality $F(\ell, m) \leqslant X$ implies

$$(4\alpha\gamma - \beta^2)\ell^2 \leqslant (4\alpha\gamma - \beta^2)\ell^2 + (\beta\ell + 2\gamma m)^2 = 4\gamma F(\ell, m) \leqslant 4\gamma X.$$

Therefore λ is supported in $[0, C_1\sqrt{X}]$ where

$$C_1 = \sqrt{\frac{4\gamma}{4\alpha\gamma - \beta^2}}.$$

When F is indefinite, we will not define our C_1; for simplicity we can simply say $C_1 = 1$. The choice of C_1 does not affect the main argument at all. However, a good choice of C_1 will make the asymptotic formula more appealing since it comes from the integral

$$\iint_{\substack{F(\ell,m)\leqslant X \\ 0\leqslant \ell \leqslant C_1\sqrt{X}}} 1 \, d\ell \, dm.$$

Once we have the restriction $0 \leqslant \ell \leqslant C_1\sqrt{X}$, in the equality $F(\ell, m) = N$ we only have a finite number of choices for (ℓ, m) since $\ell \in [0, C_1\sqrt{X}], N \in [0, X]$ and m is a solution to

the quadratic equation $F(\ell, m) = N$.

The support of λ immediately implies that $|m| \leqslant C_2\sqrt{X}$ for some positive constant C_2. When F is positive definite, it follows from symmetry. When F is indefinite, note that

$$(4|\gamma| + C_1^2\Delta)X \geqslant 4|\gamma|X + \Delta\ell^2 \geqslant (2\gamma m + \beta\ell)^2.$$

Hence

$$\sqrt{4|\gamma| + C_1^2\Delta}\sqrt{X} \geqslant 2|\gamma|m - |\beta||\ell|$$

and we can take

$$C_2 = \frac{\sqrt{4|\gamma| + C_1^2\Delta} + |\beta|C_1}{2|\gamma|}.$$

Next, we wish to understand more about $\rho(d)$, the number of solutions $\nu \pmod d$ to the quadratic congruence $F(1, \nu) \equiv 0 \pmod d$. We have the following more general lemma from [4] and [13].

Lemma 2.1.1. Let $f(k)$ be an irreducible polynomial of degree n with integral coefficients. Let $\rho(a)$ denote the number of solutions of

$$f(k) \equiv 0 \pmod a, \qquad 0 \leqslant k < a.$$

Let D denote the discriminant of the polynomial $f(k)$. Then

(1) ρ is multiplicative;

(2) if $p \nmid D$, then $\rho(p) = \rho(p^\alpha) \leqslant n$;

(3) $\rho(p^\alpha) = O_f(1)$;

(4) $\rho(k) = O_f(n^{\omega(k)})$;

(5) there exists a constant $c > 0$ such that

$$\sum_{p \leqslant x} \frac{\rho(p)}{p} = \log\log x + c + O_f((\log x)^{-A}).$$

The error term we recorded in (5) is actually stronger than that in [4]. This follows from the prime ideal theorem (with a sufficiently good error bound) with partial summation. In our case, $F(1, x)$ is a quadratic polynomial and hence we have

Lemma 2.1.2. $\rho(N) \ll \tau(N)$ for all $N \in \mathbb{N}$.

Consequently, we have

Lemma 2.1.3. Suppose that λ satisfies the pointwise bound

$$|\lambda(\ell)| \leqslant (\log X)^A$$

for some fixed positive number A. Then $|a_N| \ll \tau(N)(\log X)^{A+1}$.

Proof. First of all, we have

$$|a_N| \leqslant (\log X)^A \sum_{\substack{F(\ell,m)\leqslant X \\ \gcd(\ell,\gamma m)=1 \\ \ell\in[0,C_1\sqrt{X}]}}\sum 1.$$

If $F(u,v) = N$ and $\gcd(u,\gamma v) = 1$, then $\gcd(u,N) = 1$ and

$$F(1, vu^{-1}) \equiv 0 \pmod N.$$

We wish to show that if we fix a pair $(u,v) \in \mathbb{Z}^2$ with $F(u,v) = N$, then there are very few other pairs $(w,z) \in \mathbb{Z}^2$ such that $F(u,v) = F(w,z) = N$ and $vu^{-1} \equiv zw^{-1} \pmod N$. Precisely, if F is positive definite, there are at most 5 other pairs; if F is indefinite, then there could be at most $O(\log X)$ other pairs.

Firstly, we suppose that F is positive definite with discriminant $\beta^2 - 4\alpha\gamma = -\Delta$. Then

$$\begin{aligned} N^2 &= (\alpha u^2 + \beta uv + \gamma v^2)(\alpha w^2 + \beta wz + \gamma z^2) \\ &= (\alpha uw + \frac{\beta}{2}uz + \frac{\beta}{2}vw + \gamma vz)^2 + \frac{\Delta}{4}(uz - vw)^2. \end{aligned}$$

If $\Delta \equiv 0 \pmod 4$, then $\beta \equiv 0 \pmod 2$. Then from $uz - vw \equiv 0 \pmod N$ we deduce that

$$\alpha uw + \frac{\beta}{2}uz + \frac{\beta}{2}vw + \gamma vz \equiv 0 \pmod N$$

as well. Therefore

$$1 = \left(\frac{\alpha uw + \frac{\beta}{2}uz + \frac{\beta}{2}vw + \gamma vz}{N}\right)^2 + \Delta\left(\frac{uz - vw}{N}\right)^2.$$

Since $\Delta > 0$, the equation $1 = U^2 + \Delta V^2$ has at most 4 solutions $(U,V) \in \mathbb{Z}^2$. For each such pair (U,V), (w,z) will be uniquely determined via the system of linear equations

$$\begin{aligned} (\alpha u + \frac{\beta}{2}v)w + (\frac{\beta}{2}u + \gamma v)z &= U, \\ uz - vw &= V. \end{aligned}$$

Thus when u, v are fixed, there are at most 3 other pairs (w, z) such that $F(u, v) = F(w, z)$ and $vu^{-1} \equiv zw^{-1} \pmod{N}$. Hence

$$|a_N| \leqslant 4(\log X)^A \rho(N) \leqslant 4\tau(N)(\log X)^A.$$

When $\Delta \equiv 3 \pmod 4$, then $\beta \equiv 1 \pmod 2$. Then the result follows from the similar identity

$$\begin{aligned} N^2 =& \left(\alpha uw + \frac{\beta-1}{2}uz + \frac{\beta+1}{2}vz + \gamma vz\right)^2 \\ &+ \left(\alpha uw + \frac{\beta-1}{2}uz + \frac{\beta+1}{2}vz + \gamma vz\right)(uz - wv) + \frac{1+\Delta}{4}(uz - wv)^2 \end{aligned}$$

except there could be at most 6 solutions to $1 = U^2 + UV + \frac{1+\Delta}{4}V^2$.

When F is indefinite, let $\beta^2 - 4\alpha\gamma = \Delta > 0$ be its discriminant. When $\Delta \equiv 0 \pmod 4$, we arrive at Pell's equation

$$1 = U^2 - \Delta V^2.$$

Without loss of generality, we assume that $U, V > 0$. If $x_1 + y_1\sqrt{\Delta}$ is the fundamental solution of the above Pell's equation, then we have

$$U + V\sqrt{\Delta} = (x_1 + y_1\sqrt{\Delta})^n$$

for some $n \in \mathbb{N}$. We are only interested in small solutions since

$$|V| = |uz - wv| \leqslant 2C_1C_2X.$$

But

$$n = \frac{\log(U + V\sqrt{\Delta})}{\log(x_1 + y_1\sqrt{\Delta})} = \frac{\log(\sqrt{1 + \Delta V^2} + V\sqrt{\Delta})}{\log(x_1 + y_1\sqrt{\Delta})} \ll \log X.$$

When $\Delta \equiv 1 \pmod 4$, we have the equation

$$1 = U^2 + UV + \frac{1-\Delta}{4}V^2.$$

or

$$4 = (2U + V)^2 - \Delta V^2.$$

A similar argument would show that we have at most $O(\log X)$ solutions here. Hence

$$|a_N| \ll (\log X)\rho(N)(\log X)^A \ll \tau(N)(\log X)^{A+1}.$$

□

Our estimate for a_N is certainly not optimal. However this is sufficient for our purpose.

2.2 Proposition by Fouvry and Iwaniec

To set up the sieve, we will rely on a sieving proposition developed by Fouvry and Iwaniec in [5]. The purpose of this section is to state their proposition, and explain how one of their assumptions can be modified to fit our sequence.

For a generic sequence (a_n) of non-negative numbers, usually the first step to apply the sieve is to understand the approximation

$$A_d(x) = \sum_{\substack{n\leqslant x\\ n\equiv 0 \pmod d}} a_n = g(d)A(x) + R_d(x)$$

as in (1.6); however, they assumed each $A_d(x)$ is well-approximated by

$$M_d(x) = \frac{1}{d}\sum_{n\leqslant x} a_n(d) \tag{2.1}$$

and define $R_d(x) = A_d(x) - M_d(x)$. This approximation allows the main term $A(x)$ to be dependent on d. Furthermore, they assumed that every $a_n(d)$ is a linear combination of multiplicative functions $\varrho_l(d)$ in d, say

$$a_n(d) = \sum_l \lambda_l(n)\varrho_l(d) \tag{2.2}$$

with $\lambda_l(n) = 0$ for almost all l (the range of admissible l depends on n). Define

$$\begin{aligned}
R(x;y,z) &= \sum_{b\leqslant y}\mu(b)\Big\{R_b(x)\log\frac{x}{b} - \int_1^x R_b(t)\,dt - \sum_{c\leqslant z}\Lambda(c)R_{bc}(x)\Big\},\\
B(x;y,z) &= \sum_{\substack{bd\leqslant x\\ b>y}}\mu(b)\Big(\sum_{\substack{c\mid d\\ c>z}}\Lambda(c)\Big)a_{bd},\\
\delta_l(n;y,z) &= \sum_{b>y}\frac{\mu(b)}{b}\Big\{\varrho_l(b)\log\frac{n}{b} - \sum_{c\leqslant z}\frac{\Lambda(c)}{c}\varrho_l(bc)\Big\}.
\end{aligned} \tag{2.3}$$

They will serve as error terms. Finally let

$$\psi(l) = -\sum_b \frac{\mu(b)}{b}\varrho_l(b)\log b = \prod_p\Big(1-\frac{\varrho_l(p)}{p}\Big)\Big(1-\frac{1}{p}\Big)^{-1}. \tag{2.4}$$

Then Fouvry and Iwaniec proved that

Proposition 2.2.1. Let $c\in\mathbb{N}$. Suppose $\varrho_l(d)$ satisfies

$$\Big|\sum_{b\leqslant t}\frac{\mu(b)}{b}\varrho_l(bc)\Big| \leqslant \gcd(c,\ell)\tau(c)\Delta_l(t) \tag{2.5}$$

for all $t > 1$ with some $\Delta_l(t)$ such that $\Delta_l(t)(\log t)^2$ is decreasing. Then for $y, z \geqslant 1$ and $x > yz$ we have the identity

$$P(x) = \sum_{n\leqslant x}\sum_{l}\lambda_l(n)\{\psi(l) + \delta_l(n; y, z)\} + B(x; y, z) + R(x; y, z) + P(z).$$

The assumption of (2.5) can be weakened in certain aspects. To see how this assumption was used, we will briefly explain the proof of Proposition 2.2.1. The core of the proof of Proposition 2.2.1 is an identity developed by Vaughan [19]:

Proposition 2.2.2 (Vaughan's Identity)**.** For $y, z \geqslant 1$ and $n > z$, we have

$$\Lambda(n) = \sum_{\substack{b|n\\ b\leqslant y}}\mu(b)\log\frac{n}{b} - \sum_{\substack{bc|n\\ b\leqslant y, c\leqslant z}}\sum\mu(b)\Lambda(c) + \sum_{\substack{bc|n\\ b>y, c>z}}\sum\mu(b)\Lambda(c).$$

Then for a general sequence (a_n) and $P(x) = \sum_{n\leqslant x} a_n\Lambda(n)$, we have

$$\begin{aligned} P(x) - P(z) = &\sum_{b\leqslant y}\mu(b)\sum_{\substack{n\leqslant x\\ n\equiv 0 \pmod b}} a_n\log\frac{n}{b} - \sum_{b\leqslant y, c\leqslant z}\sum\mu(b)\Lambda(c)\sum_{\substack{n\leqslant x\\ n\equiv 0 \pmod{bc}}} a_n\\ &+ \sum_{\substack{bd\leqslant x\\ b>y}}\sum\mu(b)\Big(\sum_{\substack{c|d\\ c>z}}\Lambda(c)\Big)a_{bd}. \end{aligned}$$

The first two terms on the right will lead us to the Type I sum $R(x; y, z)$ after we isolated the main term $M_d(x)$ in (2.1); and the last term on the right is precisely the Type II sum $B(x; y, z)$. The main terms we isolated can be combined to form

$$M(x; y, z) = \sum_{n\leqslant x}\sum_{l}\lambda_l(n)\sum_{b\leqslant y}\frac{\mu(b)}{b}\Big\{\varrho_l(b)\log\frac{n}{b} - \sum_{c\leqslant z}\frac{\Lambda(c)}{c}\varrho_l(bc)\Big\}. \tag{2.6}$$

Assumption (2.5) allows us to extend the summation over b into an infinite series; the error made is $\delta_l(n; y, z)$ and the infinite sum can be factored into $\psi(l)$.

From the discussion above, we can see that we only need (2.5) in the range $c \leqslant z$ and $t > y$. Furthermore, the upper bound does not need to be this tight. For example, the estimate

$$\Big|\sum_{b\leqslant t}\frac{\mu(b)}{b}\varrho_l(bc)\Big| \ll_A \gcd(c, l)^{100}\tau(c)^{100}\tau(l)^{100}(\log t)^{-A} \qquad \text{for all } A > 0$$

is also acceptable. Comparing to (2.5), it is stronger in the $t-$aspect but weaker in the $c-$aspect. The high powers of gcd and τ would only cause some extra powers of $\log x$ in the end result, which can be balanced by choosing a sufficiently large A. The modified version of (2.5) we are going to use is given by (2.13).

2.3 Decomposition of $P(X;\chi)$

We will apply Proposition 2.2.1 to the sequence $(a_N\chi(N))$ with $\gcd(N,P_F)=1$. Recall that

$$P(X;\chi)=\sum_{\substack{N\leqslant X\\ \gcd(N,P_F)=1}} a_N\chi(N)\Lambda(N),$$

where P_F is a positive integer that depends only on F and χ is a Dirichlet character modulo q. Define

$$A_d(X;\chi)=\sum_{\substack{N\leqslant X\\ N\equiv 0 \pmod d\\ \gcd(N,P_F)=1}} a_N\chi(N) \tag{2.7}$$

for d a positive integer. Note that $A_d(X;\chi)=0$ if $\gcd(d,qP_F)>1$. We expect that $A_d(X;\chi)$ is approximated by

$$\begin{aligned} M_d(X;\chi)&=\frac{\rho(d)}{d}\sum_{\substack{N\leqslant X\\ \gcd(N,P_F)=1}}\sum_{\gcd(\ell,d)=1}\lambda(\ell;N)\\ &=\frac{\rho(d)}{d}\sum_{\substack{F(\ell,m)\leqslant X\\ \gcd(\ell,\gamma md)=1\\ \gcd(F(\ell,m),P_F)=1}}\sum\lambda(\ell)\chi(F(\ell,m)) \end{aligned}$$

when $\gcd(d,qP_F)=1$, where

$$\lambda(\ell;N)=\lambda(\ell)\chi(N)\sum_{\substack{m\in\mathbb{Z}\\ F(\ell,m)=N\\ \gcd(\ell,\gamma m)=1}}1 \qquad \text{if } \gcd(N,P_F)=1;$$

and $M_d(X;\chi)=0$ otherwise. With this we set

$$R_d(X;\chi)=A_d(X;\chi)-M_d(X;\chi). \tag{2.8}$$

Comparing to (2.1), we can therefore write $M_d(X;\chi)$ as

$$M_d(X;\chi)=\frac{1}{d}\sum_{N\leqslant X}a_N(d)$$

where

$$a_N(d)=\sum_{\ell\in\mathbb{Z}}\lambda(\ell;N)\rho_\ell(d), \qquad \rho_\ell(d)=\begin{cases}\rho(d) & \text{when } \gcd(d,\ell qP_F)=1,\\ 0 & \text{otherwise.}\end{cases}$$

By the way, if we follow Fouvry and Iwaniec's work closely, we should have defined $\rho_\ell(d)$ to be the number of solutions $\nu \pmod d$ to the quadratic congruence

$$F(\ell, \nu) \equiv 0 \pmod d.$$

Thanks to the extra assumption $\gcd(\ell, \gamma m) = 1$ we added in the definition of a_N, the condition $F(\ell, m) \equiv 0 \pmod d$ in $A_d(X; \chi)$ immediately implies that $\gcd(\ell, d) = 1$. Thus we can simply take $\rho_\ell(d) = \rho(d)$ when $\gcd(d, \ell q P_F) = 1$ and this saves us from some tedious calculations.

For a parameter D we define the complete remainder term $R(X, D; \chi)$ as

$$R(X, D; \chi) = \sum_{d \leqslant D} |R_d(X; \chi)|. \tag{2.9}$$

Let $Y, Z > 1$ be such that $X > YZ$. Put

$$R(X; Y, Z; \chi) = \sum_{b \leqslant Y} \mu(b) \left\{ R_b(X; \chi) \log \frac{X}{b} - \int_1^X R_b(t; \chi) \frac{\mathrm{d}t}{t} - \sum_{c \leqslant Z} \Lambda(c) R_{bc}(X; \chi) \right\} \tag{2.10}$$

and

$$B(X; Y, Z; \chi) = \sum_{\substack{bd \leqslant X \\ b > Y \\ \gcd(bd, P_F) = 1}} \sum \mu(b) \left(\sum_{\substack{c \mid d \\ c > Z}} \Lambda(c) \right) \chi(bd) a_{bd}. \tag{2.11}$$

We also define

$$\delta_\ell(N; Y, Z) = \sum_{\substack{b > Y \\ \gcd(b, \ell q P_F) = 1}} \frac{\mu(b)}{b} \left\{ \rho(b) \log \frac{N}{b} - \sum_{\substack{c \leqslant Z \\ \gcd(c, \ell q P_F) = 1}} \frac{\Lambda(c)}{c} \rho(bc) \right\}. \tag{2.12}$$

Clearly, they are $R(x; y, z)$, $B(x; y, z)$ and $\delta_l(n; y, z)$ in the last section for our sequence. Then we have the following proposition.

Proposition 2.3.1. Let $Y, Z \geq 1$, $X > YZ$ and $Z < Y^{1-\delta}$ for some $\delta > 0$. Assume that the discriminant of F divides P_F. Then we have the identity

$$\begin{aligned} P(X; \chi) = & \sum_{\substack{N \leqslant X \\ \gcd(N, P_F) = 1}} \sum_{\ell} \lambda(\ell; N) (H_{F,q}(\ell) + \delta_\ell(N; Y, Z)) \\ & + B(X; Y, Z; \chi) + R(X; Y, Z; \chi) + P(Z; \chi) \end{aligned}$$

where

$$H_{F,q}(\ell) = \prod_{p \nmid \ell q P_F} \left(1 - \frac{\rho(p)}{p}\right) \left(1 - \frac{1}{p}\right)^{-1} \prod_{p \mid \ell q P_F} \left(1 - \frac{1}{p}\right)^{-1}.$$

Proposition 1.1.2 would follow from Proposition 2.3.1 if we could provide acceptable estimates for $\delta_\ell(N;Y,Z), B(X;Y,Z;\chi), R(X;Y,Z;\chi)$ and $P(Z;\chi)$.

Proof. We will derive Proposition 2.3.1 from Proposition 2.2.1. From the discussion in Section 2.2, it suffices to prove a modified version of (2.5) with $\rho(bc)$ for all $t > Y, c \leqslant Z$. We claim that

$$\Big| \sum_{\substack{b\leqslant t\\ \gcd(bc,\ell qP_F)=1}} \frac{\mu(b)}{b}\rho(bc)\Big| \ll_{A,Q,\delta} \tau(c)^3\tau(\ell)(\log t)^{-A} \qquad \text{for any } A>0. \tag{2.13}$$

(Again, following (2.6), the sole purpose of such estimate is to extend the following finite sums over b

$$\sum_{\substack{b\leqslant Y\\ \gcd(b,\ell qP_F)=1}} \frac{\mu(b)\rho(b)}{b}, \sum_{\substack{b\leqslant Y\\ \gcd(b,\ell qP_F)=1}} \frac{\mu(b)\rho(b)\log b}{b}, \sum_{\substack{b\leqslant Y\\ \gcd(b,\ell qP_F)=1}} \frac{\mu(b)\rho(bc)\log b}{b},$$

to infinite sums. The first and the last one would vanish by (2.13) and the second one gives $-H_{F,q}(\ell)$. The error made in this process is the tail $\delta_\ell(N;Y,Z)$, which is finite by (2.13) and partial summation)

To prove (2.13), we can definitely assume that $\gcd(c,\ell qP_F)=1$. Since b is squarefree, we have $\gcd(b,c)|c$. By putting $e=\gcd(b,c)$, we deduce that

$$\Big| \sum_{\substack{b\leqslant t\\ \gcd(b,\ell qP_F)=1}} \frac{\mu(b)}{b}\rho(bc)\Big| \leqslant \sum_{\substack{e\leqslant t\\ e|c}} \frac{\rho(ec)}{e}\Big| \sum_{\substack{b\leqslant t/e\\ \gcd(b,\ell cqP_F)=1}} \frac{\mu(b)}{b}\rho(b)\Big|.$$

To this end, we require the following lemma, which is a modification of (2.4) in [7]:

Lemma 2.3.2. Let g be a multiplicative function such that $0\leqslant g(p)<1$ and $g(p)\ll p^{-1}$. Suppose for any $y\geqslant 2$, we have

$$\sum_{p\leqslant y} g(p) = \log\log y + c + O_A((\log y)^{-A}) \tag{2.14}$$

for any $A>0$. Then

$$\sum_{\substack{d\leqslant y\\ \gcd(d,\nu)=1}} \mu(d)g(d) \ll_{A'} \Big(\prod_{p|\nu}\Big(1+\frac{1}{\sqrt{p}}\Big)\Big)(\log y)^{-A'} \tag{2.15}$$

for any $\nu\geqslant 1, y\geqslant 2$ and $A'>0$.

In [7], Friedlander and Iwaniec assumed that (2.14) holds with $A=10$, and they used it to prove (2.15) with $A'=6$. Here we simply strengthen the assumption (2.14) and the

corresponding result (2.15) will be improved accordingly.

By Lemma 2.1.1, we can apply this lemma to $g(d) = \rho(d)/d$. Therefore

$$\sum_{\substack{b\leqslant t/e\\ \gcd(b,\ell cqP_F)=1}} \frac{\mu(b)}{b}\rho(b) \ll_A \Big(\prod_{p|\ell cqP_F}\Big(1+\frac{1}{\sqrt{p}}\Big)\Big)\Big(\log\frac{t}{e}\Big)^{-A}$$

Since $e \leqslant c \leqslant Z$ and $t^{1-\delta} > Y^{1-\delta} > Z$, we obtain

$$\log\frac{t}{e} \geqslant \log\frac{t}{Z} \geqslant \delta\log t.$$

This gives

$$\Big(\log\frac{t}{e}\Big)^{-A} \ll_\delta (\log t)^{-A}.$$

Since $\gcd(c, P_F) = 1$ and the discriminant of F divides P_F, we have $\rho(ec) = \rho(c)$ by Lemma 2.1.1. Therefore we obtain

$$\begin{aligned}\Big|\sum_{\substack{b\leqslant t\\ \gcd(b,\ell qP_F)=1}} \frac{\mu(b)}{b}\rho(bc)\Big| &\leqslant \sum_{\substack{e\leqslant t\\ e|c}}\frac{\rho(ec)}{e}\Big|\sum_{\substack{b\leqslant t/e\\ \gcd(b,\ell cqP_F)=1}}\frac{\mu(b)}{b}\rho(b)\Big|\\ &\ll_{A,\delta} \rho(c)(\log t)^{-A}\Big(\sum_{\substack{e\leqslant t\\ e|c}}\frac{1}{e}\Big)\Big(\prod_{p|\ell cqP_F}\Big(1+\frac{1}{\sqrt{p}}\Big)\Big)\\ &\ll_{A,\delta} \rho(c)\tau(c)\tau(\ell cqP_F)(\log t)^{-A}.\end{aligned}$$

By Lemma 2.1.2 and the fact that $\tau(mn) \leqslant \tau(m)\tau(n)$ for all $m, n \in \mathbb{N}$, we have

$$\Big|\sum_{\substack{b\leqslant t\\ \gcd(b,\ell qP_F)=1}} \frac{\mu(b)}{b}\rho(bc)\Big| \ll_{A,Q,\delta} \tau(c)^3\tau(\ell)(\log t)^{-A}$$

and this completes the proof of Proposition 2.3.1. □

Chapter 3

Type I Estimate

3.1 Main Goal

Our main goal in this chapter is to handle the Type I sum

$$R(X;Y,Z;\chi)=\sum_{b\leqslant Y}\mu(b)\left\{R_b(X;\chi)\log\frac{X}{b}-\int_1^X R_b(t;\chi)\frac{\mathrm{d}t}{t}-\sum_{c\leqslant Z}\Lambda(c)R_{bc}(X;\chi)\right\}$$

that appears in Proposition 2.3.1. Precisely, we will show that

Proposition 3.1.1. Suppose $q\in\mathbb{N}$ and $q\leqslant(\log X)^Q$ for some $Q>0$. Let $\varepsilon>0$. Assume that $Y,Z>1$ and $YZ<X^{1-\varepsilon}$. Then we have

$$R(X;Y,Z;\chi)\ll_{\varepsilon,Q}X^{1-\varepsilon/5}.$$

The central components in the definition of $R(X;Y,Z;\chi)$ are

$$R(X;D,\chi)=\sum_{d\leqslant D}|R_d(X;\chi)|=\sum_{d\leqslant D}|A_d(X;\chi)-M_d(X;\chi)|$$

for various choices of D. Below is our target estimate for $R(X;D,\chi)$:

Proposition 3.1.2. For $1\leqslant D\leqslant X$ we have the bound

$$R(X,D;\chi)\ll_\varepsilon q^3D^{1/4}X^{3/4+\varepsilon}.$$

To prove Proposition 3.1.2, it is convenient to remove the restrictive condition $\gcd(N,P_F)=1$ in $A_d(X;\chi)$ (see (2.7)). Furthermore, as in [5], we will first consider a smoothed version of Proposition 3.1.2. Since we need to accommodate the extra assumption $\gcd(\ell,\gamma m)=1$ (which is not present in [5]), we adopt the approach from [8] instead.

Let $\sqrt{X} \leqslant W \leqslant X$ be an additional parameter to be chosen later, and let $w : \mathbb{R}^+ \to \mathbb{R}$ be a smooth function with the following properties:

$$\begin{cases} w(u) = 0 & \text{if } u \notin [1, X], \\ 0 \leqslant w(u) \leqslant 1 & \text{if } u \in [1, X], \\ w(u) = 1 & \text{if } W \leqslant u \leqslant X - W, \\ w^{(j)}(u) \ll W^{-j} & \text{for } j = 1, 2. \end{cases} \tag{3.1}$$

For $a, \ell \geq 1$ we define the function

$$F_{a,\ell}(z) = \int_{-\infty}^{\infty} w(F(a\ell, at))e(-zt)\,\mathrm{d}t. \tag{3.2}$$

Let

$$\mathcal{A}_d(X; w, \chi) = \sum_{N \equiv 0 \pmod d} a_N w(N)\chi(N).$$

When $\gcd(d, q) = 1$, define

$$\mathcal{M}_d(X; w, \chi) = \frac{\rho(d)}{d} \sum_{\gcd(\ell, \gamma d)=1} \lambda(\ell)\Big(\frac{\sum_{k \pmod q} \chi(F(\ell, k))}{q}\Big)\Big(\prod_{p|\ell, p\nmid q}\Big(1 - \frac{1}{p}\Big)\Big)F_{1,\ell}(0) \tag{3.3}$$

as well as the smoothed remainder term

$$\mathcal{R}_d(X; w, \chi) = \mathcal{A}_d(X; w, \chi) - \mathcal{M}_d(X; w, \chi).$$

When $\gcd(d, q) > 1$, clearly $\mathcal{A}_d(X; w, \chi) = 0$ and accordingly we define

$$\mathcal{M}_d(X; w, \chi) = \mathcal{R}_d(X; w, \chi) = 0.$$

We obtain the following lemma.

Proposition 3.1.3. Let w and λ be as above and $1 \leqslant D \leqslant X$. Then one has

$$\sum_{d \leqslant D} |\mathcal{R}_d(X; w, \chi)| \ll_\varepsilon \frac{q^3 D^{1/2} X^{3/2+\epsilon}}{W}.$$

In Section 3.2, we will first prove a list of auxiliary lemmas that we use in the proof of Proposition 3.1.3. The complete proof of Proposition 3.1.3 is given in Section 3.3. We will consider the unsmoothed version of Proposition 3.1.3 in Section 3.4. We decide to leave it as a separate proposition (rather than proving Proposition 3.1.2 directly) since this unsmoothed version will be needed in the proof of Corollary 1.1.3. Finally, in Section 3.5, we will finish the proof of Proposition 3.1.2 and 3.1.1.

3.2 Auxiliary Lemmas

Lemma 3.2.1. For any $a, c, q, \ell \in \mathbb{N}$ with $\gcd(c, q) = 1$ and congruence classes $\nu \pmod c$ and $k \pmod q$, we have

$$\sum_{\substack{m \equiv \ell\nu \pmod c \\ m \equiv k \pmod q}} w(F(a\ell, am)) = \frac{1}{cq} \sum_{h \in \mathbb{Z}} e\Big(\frac{hk\bar{c}}{q}\Big) e\Big(\frac{h\ell\nu\bar{q}}{c}\Big) F_{a,\ell}\Big(\frac{h}{cq}\Big) \tag{3.4}$$

where $F_{a,\ell}(z)$ is defined in (3.2).

Proof. Define the function

$$G(x) = \sum_{\substack{m \equiv \ell\nu \pmod c \\ m \equiv k \pmod q}} w(F(a\ell, am + acqx)).$$

Then $G(x) = G(x+1)$. Computing its h-th Fourier coefficients,

$$\begin{aligned}
\int_0^1 G(x) e^{-2\pi i h x}\, dx &= \sum_{\substack{m \equiv \ell\nu \pmod c \\ m \equiv k \pmod q}} \int_0^1 w(F(a\ell, am + acqx)) e^{-2\pi i h x}\, dx \\
&= \frac{1}{cq} \sum_{\substack{m \equiv \ell\nu \pmod c \\ m \equiv k \pmod q}} \int_m^{m+cq} w(F(a\ell, ax)) e^{\frac{-2\pi i h(x-m)}{cq}}\, dx \\
&= \frac{1}{cq} \sum_{\substack{m \equiv \ell\nu \pmod c \\ m \equiv k \pmod q}} e^{\frac{2\pi i h m}{cq}} \int_m^{m+cq} w(F(a\ell, ax)) e^{\frac{-2\pi i h x}{cq}}\, dx.
\end{aligned}$$

By the Chinese remainder theorem, it follows

$$m \equiv \ell\nu q\bar{q} + kc\bar{c} \pmod{cq}.$$

Interchanging the integral and sum again,

$$\int_0^1 G(x) e^{-2\pi i h x}\, dx = \frac{1}{cq} e\Big(\frac{hk\bar{c}}{q}\Big) e\Big(\frac{h\ell\nu\bar{q}}{c}\Big) \int_{-\infty}^{\infty} w(F(a\ell, ax)) e^{\frac{-2\pi i h x}{cq}}\, dx.$$

Thus

$$G(x) = \frac{1}{cq} \sum_{h \in \mathbb{Z}} e\Big(\frac{hk\bar{c}}{q}\Big) e\Big(\frac{h\ell\nu\bar{q}}{c}\Big) \Big(\int_{-\infty}^{\infty} w(F(a\ell, at)) e^{\frac{-2\pi i h t}{cq}}\, dt \Big) e^{2\pi i h x}.$$

Therefore by putting $x = 0$,

$$\sum_{\substack{m \equiv \ell\nu \pmod c \\ m \equiv k \pmod q}} w(F(a\ell, am)) = \frac{1}{cq} \sum_{h \in \mathbb{Z}} e\Big(\frac{hk\bar{c}}{q}\Big) e\Big(\frac{h\ell\nu\bar{q}}{c}\Big) F_{a,\ell}\Big(\frac{h}{cq}\Big).$$

□

Lemma 3.2.2. When $\gcd(d,q)=1$,

$$\mathcal{M}_d(X;w,\chi)=\frac{\rho(d)}{q}\sum_{\gcd(a,\gamma)=1}\frac{\mu(a)}{ac}\sum_{\gcd(\ell,\gamma)=1}\lambda(a\ell)\sum_{k\ (\mathrm{mod}\ q)}\chi(F(a\ell,ak))\int_{-\infty}^{\infty}w(F(a\ell,t))\,dt. \tag{3.5}$$

Proof. Recall from (3.3) that

$$\mathcal{M}_d(X;w,\chi)=\frac{\rho(d)}{d}\sum_{\gcd(\ell,\gamma d)=1}\lambda(\ell)\Big(\frac{\sum_{k\ (\mathrm{mod}\ q)}\chi(F(\ell,k))}{q}\Big)\Big(\prod_{p\mid\ell,p\nmid q}\Big(1-\frac{1}{p}\Big)\Big)F_{1,\ell}(0)$$

when $\gcd(d,q)=1$. By interchanging the summations over a and ℓ, the long expression in (3.5) gives

$$\frac{\rho(d)}{dq}\sum_{\gcd(\ell,\gamma)=1}\lambda(\ell)F_{1,\ell}(0)\sum_{\substack{a\mid\ell\\ \gcd(a,\gamma)=1}}\frac{\mu(a)\gcd(a,d)}{a}\sum_{k\ (\mathrm{mod}\ q)}\chi(F(\ell,ak)).$$

We can assume that $\gcd(a,q)=1$; otherwise $\chi(F(\ell,ak))=0$. Therefore the above expression reduces to

$$\frac{\rho(d)}{d}\sum_{\gcd(\ell,\gamma)=1}\lambda(\ell)\Big(\frac{\sum_{k\ (\mathrm{mod}\ q)}\chi(F(\ell,k))}{q}\Big)F_{1,\ell}(0)\sum_{\substack{a\mid\ell\\ \gcd(a,q)=1}}\frac{\mu(a)\gcd(a,d)}{a}.$$

Note that

$$\sum_{\substack{a\mid\ell\\ \gcd(a,q)=1}}\frac{\mu(a)\gcd(a,d)}{a}=\prod_{p\mid\ell,p\nmid q}\Big(1-\frac{\gcd(p,d)}{p}\Big).$$

This equals to 0 if $\gcd(\ell,d)>1$; otherwise $\gcd(p,d)=1$ for all $p\mid\ell$ and the result follows. □

Lemma 3.2.3. Let $a,h,\ell\in\mathbb{N}$ be fixed. Put

$$T(u)=a^2\Big(\alpha\ell^2+\frac{\beta\ell u\sqrt{X}}{h}+\frac{\gamma Xu^2}{h^2}\Big).$$

Then

$$(w(T(u)))''\ll\frac{a^2X^2}{h^2W^2}.$$

Proof. Note that

$$(w(T(u)))''\ =\ w''(T(u))T'(u)^2+w'(T(u))T''(u).$$

From the support of w, we have

$$\begin{aligned} w''(T(u))T'(u)^2 &= w''(T(u))\Big(\frac{a^2\beta\ell\sqrt{X}}{h}+\frac{2a^2\gamma Xu}{h^2}\Big)^2 \\ &\ll \frac{a^4}{W^2}\Big(\frac{(\sqrt{X}/a)\sqrt{X}}{h}+\frac{X(h/a)}{h^2}\Big)^2 \\ &\ll \frac{a^2X^2}{h^2W^2} \end{aligned}$$

and

$$w'(T(u))T''(u) \;=\; \frac{2\gamma Xa^2}{h^2}w'(T(u)) \;\ll\; \frac{Xa^2}{h^2W} \;\ll\; \frac{a^2X^2}{h^2W^2}.$$

□

The last tool we need to prove Proposition 3.1.3 is a large sieve type inequality for the points $\nu/d \pmod 1$, where ν is a solution to the congruence equation $F(1,\nu)\equiv 0 \pmod d$. In [5], Fouvry and Iwaniec handled the case $\nu^2+1\equiv 0 \pmod d$. Some extensions were considered in [15] and [17]. The author and Choi [2] proved this for a large class of positive definite binary quadratic form, and Schindler and Xiao also developed an extension on their own too. These versions are limited to the positive definite case only. However, a completely general version of this was actually proved before all the aforementioned work in a paper by Balog, Blomer, Dartyge and Tenenbaum. This is Proposition 3 from [1].

Proposition 3.2.4. Let $F(x,y)=\alpha x^2+\beta xy+\gamma y^2\in\mathbb{Z}[x,y]$ be an arbitrary quadratic form whose discriminant is not a perfect square. For any sequence α_n of complex numbers, and for positive real numbers D,N, we have

$$\sum_{D\leqslant d\leqslant 2D}\sum_{F(\nu,1)\equiv 0 \pmod d}\Big|\sum_{n\leqslant N}\alpha_n e\Big(\frac{\nu n}{d}\Big)\Big|^2 \ll_F (D+N)\sum_n|\alpha_n|^2.$$

3.3 Proof of Proposition 3.1.3

Proof. To estimate $\mathcal{R}_d(X;w,\chi)$, it suffices to deal with the terms with $\gcd(d,q)=1$. Under this assumption, note that

$$\mathcal{A}_d(X;w,\chi)=\sum_{N\equiv 0 \pmod d}\chi(N)w(N)\sum_{\substack{F(\ell,m)=N\\ \gcd(\ell,\gamma m)=1}}\sum\lambda(\ell).$$

The conditions $\gcd(\ell,\gamma m)=1$ and $F(\ell,m)\equiv 0 \pmod d$ imply $\gcd(\ell,d)=1$; hence $\mathcal{A}_d(X;w,\chi)$ can be rewritten as

$$\mathcal{A}_d(X;w,\chi)=\sum_{\gcd(\ell,\gamma)=1}\lambda(\ell)\sum_{\substack{\nu \pmod d\\ F(1,\nu)\equiv 0 \pmod d}}\sum_{\substack{m\equiv \ell\nu \pmod d\\ \gcd(\ell,m)=1}}\chi(F(\ell,m))w(F(\ell,m)).$$

By Möbius inversion, we can trade the condition $\gcd(\ell, m) = 1$ with

$$\sum_{\gcd(a,\gamma)=1} \mu(a) \sum_{\gcd(\ell,\gamma)=1} \lambda(a\ell) \sum_{\substack{\nu \pmod d \\ F(1,\nu)\equiv 0 \pmod d}} \sum_{am \equiv a\ell\nu \pmod d} \chi(F(a\ell, am))w(F(a\ell, am)).$$

From the support of λ, a is bounded by $O(\sqrt{X})$. Now the innermost sum can be rewritten as

$$\sum_{k \pmod q} \chi(F(a\ell, ak)) \sum_{\substack{am \equiv a\ell\nu \pmod d \\ m \equiv k \pmod q}} w(F(a\ell, am)).$$

To simplify our notation, let $c = d/\gcd(a, d)$. Then the condition $am \equiv a\ell\nu \pmod d$ is the same as $m \equiv \ell\nu \pmod c$. By Lemma 3.2.1, we have

$$\sum_{\substack{m \equiv \ell\nu \pmod c \\ m \equiv k \pmod q}} w(F(a\ell, am)) = \frac{1}{cq} \sum_{h\in\mathbb{Z}} e\Big(\frac{hk\bar{c}}{q}\Big) e\Big(\frac{h\ell\nu\bar{q}}{c}\Big) F_{a,\ell}\Big(\frac{h}{cq}\Big)$$

where $F_{a,\ell}(z)$ is defined in (3.2). Now $\mathcal{A}_d(X; w, \chi)$ can be expressed as

$$\begin{aligned} \mathcal{A}_d(X; w, \chi) \;=\; & \frac{1}{q} \sum_{\gcd(a,\gamma)=1} \frac{\mu(a)}{c} \sum_{\gcd(\ell,\gamma)=1} \lambda(a\ell) \sum_{\substack{\nu \pmod d \\ F(1,\nu)\equiv 0 \pmod d}} \sum_{k \pmod q} \chi(F(a\ell, ak)) \\ & \sum_{h\in\mathbb{Z}} e\Big(\frac{hk\bar{c}}{q}\Big) e\Big(\frac{h\ell\nu\bar{q}}{c}\Big) F_{a,\ell}\Big(\frac{h}{cq}\Big). \end{aligned} \tag{3.6}$$

By Lemma 3.2.2, $\mathcal{M}_d(X; w, \chi)$ equals to the summand when $h = 0$. Accordingly, $\mathcal{R}_d(X; w, \chi)$ represents the terms with $h \neq 0$. We wish to sum $\mathcal{R}_d(X; w, \chi)$ dyadically and hence we define

$$\mathcal{R}(X, D; w, \chi) = \sum_{D<d\leqslant 2D} |\mathcal{R}_d(X; w, \chi)|.$$

Substituting $b = d/c = \gcd(a, d)$, each term in the above sum can be bounded by

$$|\mathcal{R}_d(X; w, \chi)| \;\leqslant\; \frac{1}{dq} \sum_a^\flat \sum_{\substack{bc=d \\ b|a}} \rho(b)b \sum_{\substack{\nu \pmod c \\ F(1,\nu)\equiv 0 \pmod c}} \sum_{k \pmod q} |W_a(c, \nu)|$$

where

$$W_a(c, \nu) = \sum_{h\neq 0} \sum_{(\ell,\gamma d)=1} \lambda(a\ell)\chi(F(a\ell, ak)) e\Big(\frac{hk\bar{c}}{q}\Big) e\Big(\frac{h\ell\nu\bar{q}}{c}\Big) F_{a,\ell}\Big(\frac{h}{cq}\Big).$$

Here $\sum^\flat$ means we are summing over positive squarefree integers. Hence

$$\mathcal{R}(X, D; w, \chi) \;\leqslant\; \frac{1}{Dq} \sum_{k \pmod q} \sum_a^\flat \sum_{b|a} \rho(b)bV_a(D/b) \tag{3.7}$$

where

$$V_a(C) = \sum_{C<c\leqslant 2C} \sum_{\substack{\nu \pmod c \\ F(1,\nu)\equiv 0 \pmod c}} |W_a(c,\nu)|.$$

By dyadic division,

$$V_a(C) \leqslant \sum_H \Big(V_a^+(C,H) + V_a^-(C,H) \Big) \tag{3.8}$$

where H is a power of 2,

$$V_a^+(C,H) = \sum_{C<c\leqslant 2C} \sum_{\substack{\nu \pmod c \\ F(1,\nu)\equiv 0 \pmod c}} \Big| \sum_{H\leqslant h<2H} \sum_{\gcd(\ell,\gamma)=1} \chi(F(a\ell, ak)) \lambda(a\ell) e\Big(\frac{hk\overline{c}}{q}\Big) e\Big(\frac{h\ell\nu\overline{q}}{c}\Big) F_{a,\ell}\Big(\frac{h}{cq}\Big)\Big|$$

and $V_a^-(C,H)$ is defined similarly for those $h<0$. We claim that

$$V_a^+(C,H) \ll q^3 X^{3/4} C^{1/2} \Big(C + \frac{H\sqrt{X}}{a}\Big)^{1/2} \log^{A+2}(HX) \min\Big\{ \frac{H^{1/2}}{a^{3/2}}, \frac{XC^2 a^{1/2}}{H^{3/2} W^2}\Big\}. \tag{3.9}$$

Assuming the above estimates hold, we deduce that

$$\begin{aligned} V_a^+(C,H) &\ll \frac{qX^{5/4} C H^{1/2}}{W^{1/2} a^{3/2}} \log^{A+2}(HX) \qquad \text{when } H \leqslant aC\sqrt{X} W^{-1}, \\ V_a^+(C,H) &\ll \frac{q^3 X^2 C^{5/2}}{W^2 H} \log^{A+2}(HX) \qquad \text{when } H > aC\sqrt{X} W^{-1}. \end{aligned}$$

Similarly, the same estimates hold for $V_a^-(C,H)$ as well. Applying the above estimates on (3.8), we obtain

$$V_a(C) \leqslant \sum_H \Big(V_a^+(C,H) + V_a^-(C,H)\Big) \ll \frac{q^3 X^{3/2} C^{3/2}}{aW} \log^{A+2} X. \tag{3.10}$$

Then (3.7) reduces to

$$\mathcal{R}(X,D;w,\chi) \leqslant \frac{q^3 X^{3/2}\sqrt{D}\log^{A+2} X}{W} \sum_{a\ll\sqrt{X}} \frac{\tau(a)}{a} \ll \frac{q^3 X^{3/2}\sqrt{D}\log^{A+4} X}{W}.$$

Here we have used the very crude estimate $\sum_{b|a} \rho(b) b^{-1/2} \leqslant \tau(a)$. Finally by summing dyadically,

$$\sum_{d\leqslant D} |\mathcal{R}_d(X;w,\chi)| \ll \sum_D \mathcal{R}(X,D;w,\chi) \ll \frac{q^3 D^{1/2} X^{3/2+\epsilon}}{W}.$$

Our proof of Proposition 3.1.3 will therefore be completed once we justify the estimate (3.9). The rest of this section is dedicated to that.

Recall that

$$V_a^+(C,H)=\sum_{C<c\leqslant 2C}\sum_{\substack{\nu \pmod c\\ F(1,\nu)\equiv 0 \pmod c}}\Big|\sum_{H\leqslant h<2H}\sum_{\gcd(\ell,\gamma)=1}\chi(F(a\ell,ak))\\ \lambda(a\ell)e\Big(\frac{hk\bar{c}}{q}\Big)e\Big(\frac{h\ell\nu\bar{q}}{c}\Big)F_{a,\ell}\Big(\frac{h}{cq}\Big)\Big|$$

and

$$F_{a,\ell}(z)=\int_{-\infty}^{\infty}w(F(a\ell,at))e(-zt)\,\mathrm{d}t.$$

Note that for any integer c, we have

$$F_{a,\ell}\Big(\frac{h}{cq}\Big)=\frac{\sqrt{X}}{|h|}\int_{-\infty}^{\infty}w\Big(a^2\Big(\alpha\ell^2+\frac{\beta\ell u\sqrt{X}}{h}+\frac{\gamma Xu^2}{h^2}\Big)\Big)e\Big(-\frac{\sqrt{X}u}{cq}\Big)\,du. \tag{3.11}$$

For a reduced residue class $t \pmod q$, we define

$$\alpha_{h,\ell,t}(u)=\chi(F(a\ell,ak))\lambda(a\ell)\frac{H}{h}w\Big(a^2\Big(\alpha\ell^2+\frac{\beta\ell u\sqrt{X}}{h}+\frac{\gamma Xu^2}{h^2}\Big)\Big)e\Big(\frac{hk\bar{t}}{q}\Big).$$

Then

$$V_a^+(C,H)\ll\frac{\sqrt{X}}{H}\int_{-C_2H/a}^{C_2H/a}\sum_{t \pmod q}^{*}\sum_{C<c\leqslant 2C}\sum_{\substack{\nu \pmod c\\ F(1,\nu)\equiv 0 \pmod c}}\Big|\sum_{H\leqslant h<2H}\\ \sum_{\gcd(\ell,\gamma)=1}\alpha_{h,\ell,t}(u)e\Big(\frac{h\ell\nu\bar{q}}{c}\Big)\Big|\,du. \tag{3.12}$$

The symbol $\sum^*$ means we are summing over reduced residue classes only. Notice that

$$\sum_h\sum_\ell\alpha_{h,\ell,t}(u)e\Big(\frac{h\ell\nu\bar{q}}{c}\Big)=\sum_{0\leqslant h_0,\ell_0<q}\sum_{\substack{h\equiv h_0 \pmod q\\ \ell\equiv\ell_0 \pmod q}}\alpha_{h,\ell,t}(u)e\Big(\frac{\nu n}{c}\Big)e\Big(\frac{\nu h_0\ell_0\bar{q}}{c}\Big)$$

where $n=(h\ell-h_0\ell_0)/q$. Hence for each fixed pair $0\leqslant h_0,\ell_0<q$, we only need to estimate

$$\sum_{C<c\leqslant 2C}\sum_{F(1,\nu)\equiv 0 \pmod c}\Big|\sum_{n\leqslant N}\alpha_n e\Big(\frac{\nu n}{c}\Big)\Big|$$

where

$$\alpha_n=\sum_{\substack{h\equiv h_0 \pmod q\\ \ell\equiv\ell_0 \pmod q\\ \gcd(\ell,\gamma)=1\\ h\ell=nq+h_0\ell_0}}\sum\alpha_{h,\ell,t}(u)$$

and

$$N = q + \frac{1}{q}(2H)\Big(\frac{C_1\sqrt{X}}{a}\Big) \ll \frac{H\sqrt{x}}{a}.$$

Applying Proposition 3.2.4 and Cauchy-Schwarz inequality to (3.12), we deduce that

$$V_a^+(C,H) \ll \frac{\sqrt{X}}{H}\cdot\frac{H}{a}\cdot q\cdot\bigg(\Big(C+\frac{H\sqrt{X}}{a}\Big)E\bigg)^{1/2}\cdot(C\log C)^{1/2}$$

where

$$E = \sum_n\bigg(\sum_{\substack{h\equiv h_0 \ (\mathrm{mod}\ q)\\ \ell\equiv \ell_0 \ (\mathrm{mod}\ q)\\ \gcd(\ell,\gamma)=1\\ h\ell = nq+h_0\ell_0}}\sum |\lambda(a\ell)|\bigg)^2 \ll \frac{H\sqrt{X}}{a}\log^{2A+3}(HX).$$

Here we have used the pointwise bound $|\lambda(a\ell)| \leqslant (\log X)^A$. The last term comes from the fact that

$$\sum_{C<c\leqslant 2C}\ \sum_{\substack{\nu \ (\mathrm{mod}\ c)\\ F(\nu,1)\equiv 0 \ (\mathrm{mod}\ c)}} 1 \leqslant \sum_{C<c\leqslant 2C}\tau(c) \ll C\log C.$$

Hence we obtain

$$V_a^+(C,H) \ll \frac{qX^{3/4}C^{1/2}H^{1/2}}{a^{3/2}}\Big(C+\frac{H\sqrt{X}}{a}\Big)^{1/2}\log^{A+2}(HX). \tag{3.13}$$

To develop a similar bound for large values of H, we apply integration by parts twice in (3.11) as in [8], followed with the large sieve type estimate. By Lemma 3.2.3, we have

$$\begin{aligned} F_{a,\ell}\Big(\frac{h}{cq}\Big) &= \frac{\sqrt{X}}{|h|}\int_{-\infty}^{\infty} w(T(u))e\Big(-\frac{\sqrt{X}u}{cq}\Big)\,du \\ &= -\frac{c^2q^2}{4\pi|h|\sqrt{X}}\int_{-\infty}^{\infty}(w(T(u)))''e\Big(-\frac{\sqrt{X}u}{cq}\Big)\,du \end{aligned}$$

and $(w(T(u)))'' \ll a^2X^2h^{-2}W^{-2}$. For a reduced residue class $t \ (\mathrm{mod}\ q)$, we define

$$\alpha_{h,\ell,t}(u) = \chi(F(a\ell,ak))\lambda(a\ell)\frac{H}{h}\frac{c^2}{C^2}\frac{(w(T(u))''}{(a^2X^2/h^2W^2)}e\Big(\frac{hk\bar{t}}{q}\Big).$$

Then we have again

$$\begin{aligned} V_a^+(C,H) \ll &\frac{q^2C^2}{H\sqrt{X}}\cdot\frac{a^2X^2}{H^2W^2}\cdot\int_{-C_1H/a}^{C_1H/a}\sideset{}{^*}\sum_{t \ (\mathrm{mod}\ q)}\sum_{C<c\leqslant 2C}\sum_{\substack{\nu \ (\mathrm{mod}\ c)\\ F(1,\nu)\equiv 0 \ (\mathrm{mod}\ c)}}\bigg|\sum_{H\leqslant h<2H}\\ &\sum_{\gcd(\ell,\gamma)=1}\alpha_{h,\ell,t}(u)e\Big(\frac{h\ell\nu\overline{q}}{c}\Big)\bigg|\,du. \end{aligned}$$

By another application of Proposition 3.2.4 and Cauchy-Schwarz inequality, we arrive at

$$V_a^+(C,H) \ll \frac{a^2q^2C^2X^{3/2}}{H^3W^2}\cdot\frac{H}{a}\cdot q\cdot\left(\left(C+\frac{H\sqrt{X}}{a}\right)\frac{H\sqrt{X}}{a}\log^{2A+3}(HX)\right)^{1/2}\cdot(C\log C)^{1/2}$$

After simplification, we have

$$V_a^+(C,H) \ll \frac{q^3X^{7/4}C^{5/2}a^{1/2}}{H^{3/2}W^2}\left(C+\frac{H\sqrt{X}}{a}\right)^{1/2}\log^{A+2}(HX). \tag{3.14}$$

Combining (3.13) and (3.14), we obtain our desired estimates for $V_a^+(C,H)$ in (3.9). □

3.4 The unsmoothed version

We let the scripted letters $\mathcal{A},\mathcal{M},\mathcal{R}$ denote the analogous quantities A,M,R which appeared in the Section 2.3, but without the condition $\gcd(N,P_F)=1$. Precisely,

$$\mathcal{A}_d(X;\chi) = \sum_{\substack{N\leqslant X\\ N\equiv 0 \pmod d}} a_N\chi(N) \tag{3.15}$$

which vanishes when $\gcd(d,q)>1$. We also define

$$\begin{aligned}\mathcal{M}_d(X;\chi) &= \frac{\rho(d)}{d}\sum_{N\leqslant X}\sum_{\substack{\ell\in\mathbb{N}\\ \gcd(\ell,d)=1}}\lambda(\ell;N)\\ &= \frac{\rho(d)}{d}\sum_{\substack{\ell\in\mathbb{N},m\in\mathbb{Z}\\ F(\ell,m)\leqslant X\\ \gcd(\ell,\gamma md)=1}}\sum\lambda(\ell)\chi(F(\ell,m))\end{aligned} \tag{3.16}$$

when $\gcd(d,q)=1$ and $M_d(X;\chi)=0$ otherwise. We then have the following analogue to Lemma 3.1.2:

Proposition 3.4.1. For $1\leqslant D\leqslant X$ we have the bound

$$\mathcal{R}(X,D;\chi) \ll_\varepsilon q^3D^{1/4}X^{3/4+\varepsilon}.$$

Proof. Our goal is to show that

$$\sum_{d\leqslant D}|\mathcal{R}_d(X;\chi)| \ll_\varepsilon q^3D^{1/4}X^{3/4+\varepsilon}.$$

It suffices to show that the error we made when we replace $\mathcal{A}_d(X;\chi)$ with $\mathcal{A}_d(X;w,\chi)$ is negligible as well, i.e. both $|\mathcal{A}_d(X;\chi)-\mathcal{A}_d(X;w,\chi)|$ and $|\mathcal{M}_d(X;\chi)-\mathcal{M}_d(X;w,\chi)|$ are

small. Note that by Lemma 2.1.3,

$$\sum_{\substack{d\leqslant D\\ \gcd(d,q)=1}} |\mathcal{A}_d(X;\chi)-\mathcal{A}_d(X;w,\chi)| \ll \sum_{\substack{N\leqslant W\\ \text{or } X>N\geqslant X-W}} |a_N|\tau(N)$$
$$\ll (\log X)^{A+1} \sum_{\substack{N\leqslant W\\ \text{or } X>N\geqslant X-W}} \tau^2(N)$$

Here we require a result in [18] (Lemma 2): for any $k\in\mathbb{N}$ and $\varepsilon>0$, we have

$$\sum_{x\leqslant n\leqslant x+y} \tau_k(n) \ll_{k,\varepsilon} y(\log x)^{k-1} \tag{3.17}$$

for any $x^\varepsilon\leqslant y\leqslant x$, where $\tau_k(n)$ is the number of representations of n as the product of k positive integers. By comparing the values of τ^2 and τ_5 at prime powers, we deduce that $\tau^2(n)\leqslant\tau_5(n)$ for all $n\in\mathbb{N}$. Hence

$$\sum_{X-W\leqslant N\leqslant X} \tau^2(N) \ll W(\log X)^4.$$

Therefore

$$\sum_{\substack{d\leqslant D\\ \gcd(d,q)=1}} |\mathcal{A}_d(X;\chi)-\mathcal{A}_d(X;w,\chi)| \ll W(\log X)^{A+5}.$$

For $|\mathcal{M}_d(X;\chi)-\mathcal{M}_d(X;w,\chi)|$, we first rewrite $\mathcal{M}_d(X;\chi)$ as

$$\begin{aligned}\mathcal{M}_d(X;\chi) &= \frac{\rho(d)}{d} \sum_{\substack{\ell\in\mathbb{N}\\ \gcd(\ell,\gamma d)=1}} \lambda(\ell) \sum_{\substack{m\in\mathbb{Z}\\ F(\ell,m)\leqslant X\\ \gcd(m,\ell)=1}} \chi(F(\ell,m))\\ &= \frac{\rho(d)}{d} \sum_{\substack{\ell\in\mathbb{N}\\ \gcd(\ell,\gamma d)=1}} \lambda(\ell) \sum_{k \pmod q} \chi(F(\ell,k)) \sum_{\substack{m\in\mathbb{Z}\\ F(\ell,m)\leqslant X\\ \gcd(m,\ell)=1\\ m\equiv k \pmod q}} 1.\end{aligned}$$

Therefore from (3.3), $|\mathcal{M}_d(X;\chi)-\mathcal{M}_d(X;w,\chi)|$ is bounded by

$$\frac{\rho(d)}{d} \sum_{\substack{\ell\in\mathbb{N}\\ \gcd(\ell,\gamma d)=1}} |\lambda(\ell)| \sum_{k \pmod q} |\chi(F(\ell,k))| \Bigg| \sum_{\substack{m\in\mathbb{Z}\\ F(\ell,m)\leqslant X\\ \gcd(m,\ell)=1\\ m\equiv k \pmod q}} 1 - \frac{1}{q}\prod_{p\mid\ell,p\nmid q}\left(1-\frac{1}{p}\right)\int_{-\infty}^{\infty} w(F(\ell,t))\,dt\Bigg|. \tag{3.18}$$

Due to the presence of $\chi(F(\ell,k))$, we can assume that $\gcd(\ell,q,k)=1$. By Möbius inversion, we have

$$\begin{aligned}\sum_{\substack{m\in\mathbb{Z}\\ F(\ell,m)\leqslant X\\ \gcd(m,\ell)=1\\ m\equiv k \pmod q}} 1 &= \sum_{\substack{a\mid\ell\\ \gcd(a,q)=1}}\mu(a)\sum_{\substack{m\in\mathbb{Z}\\ F(\ell,am)\leqslant X\\ am\equiv k \pmod q}} 1\\ &= \sum_{\substack{a\mid\ell\\ \gcd(a,q)=1}}\mu(a)\Big(\frac{1}{q}\int_{F(\ell,at)\leqslant X} 1\,dt+O(1)\Big)\\ &= \frac{1}{q}\prod_{p\mid\ell,p\nmid q}\Big(1-\frac{1}{p}\Big)\int_{F(\ell,t)\leqslant X}1\,dt+O(\tau(\ell)).\end{aligned}$$

Applying this to (3.18), we arrive at the bound

$$|\mathcal{M}_d(X;\chi)-\mathcal{M}_d(X;w,\chi)| \ll \frac{\rho(d)}{d}\Big(\sum_{\substack{\ell\in\mathbb{N},m\in\mathbb{Z}\\ F(\ell,m)\leqslant W\\ \text{or } X>F(\ell,m)\geqslant X-W}}\sum|\lambda(\ell)|\Big)+\frac{\rho(d)q}{d}\sum_{\ell\in\mathbb{N}}|\lambda(\ell)|\tau(\ell).$$

By Lemma 2.1.3 and (3.17) again, we deduce that

$$|\mathcal{M}_d(X;\chi)-\mathcal{M}_d(X;w,\chi)| \ll \frac{\rho(d)q}{d}W\log^{A+2}X.$$

After obtaining bounds for $|\mathcal{A}_d(X;\chi)-\mathcal{A}_d(X;w,\chi)|$ and $|\mathcal{M}_d(X;\chi)-\mathcal{M}_d(X;w,\chi)|$, we are ready to estimate $|\mathcal{R}_d(X;\chi)|$. Summing over d, we obtain

$$\begin{aligned}\sum_{d\leqslant D}|\mathcal{R}_d(X;\chi)| &\ll_\varepsilon \frac{q^3D^{1/2}X^{3/2+\epsilon}}{W}+qW(\log^{A+5}X)\Big(\sum_{d\leqslant D}\frac{\rho(d)}{d}\Big)\\ &\ll_\varepsilon \frac{q^3D^{1/2}X^{3/2+\epsilon}}{W}+qWX^{\varepsilon}.\end{aligned}$$

This is because by Lemma 2.1.1,

$$\sum_{d\leqslant D}\frac{\rho(d)}{d} \ll \sum_{d\leqslant D}\frac{\tau(d)}{d} \ll_\varepsilon D^{\varepsilon}\log D.$$

Finally, by choosing $W=D^{1/4}X^{3/4}$, we obtain

$$\mathcal{R}(X,D;\chi)=\sum_{d\leqslant D}|\mathcal{R}_d(X;\chi)| \ll_\varepsilon q^3D^{1/4}X^{3/4+\varepsilon}.$$

□

3.5 Proof of Proposition 3.1.2 and 3.1.1

Proof of Proposition 3.1.2. Let χ_0 be the principal character modulo P_F. Then $\chi\chi_0$ is a Dirichlet character modulo qP_F. For $1 \leqslant D \leqslant X$, by Proposition 3.4.1,

$$\mathcal{R}(X, D; \chi\chi_0) = \sum_{\substack{d \leqslant D \\ \gcd(d, qP_F)=1}} |\mathcal{A}_d(X; \chi\chi_0) - \mathcal{M}_d(X; \chi\chi_0)| \ll_\varepsilon q^3 D^{1/4} X^{3/4+\varepsilon}.$$

Note that when $\gcd(d, qP_F) = 1$,

$$\begin{aligned}
\mathcal{A}_d(X; \chi\chi_0) &= \sum_{\substack{N \leqslant X \\ N \equiv 0 \pmod d}} a_n \chi(N)\chi_0(N) \\
&= \sum_{\substack{N \leqslant X \\ N \equiv 0 \pmod d \\ \gcd(N, P_F)=1}} a_n \chi(N) \\
&= A_d(X; \chi)
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{M}_d(X; \chi\chi_0) &= \frac{\rho(d)}{d} \sum_{\substack{F(\ell,m) \leqslant X \\ \gcd(\ell, \gamma m d)=1}} \sum \lambda(\ell)\chi(F(\ell, m))\chi_0(F(\ell, m)) \\
&= \frac{\rho(d)}{d} \sum_{\substack{F(\ell,m) \leqslant X \\ \gcd(\ell, \gamma m d)=1 \\ \gcd(F(\ell,m), P_F)=1}} \sum \lambda(\ell)\chi(F(\ell, m)) \\
&= M_d(X; \chi).
\end{aligned}$$

Hence the result follows. □

Proof of Proposition 3.1.1. Recall that

$$R(X; Y, Z; \chi) = \sum_{b \leqslant Y} \mu(b) \left\{ R_b(X; \chi) \log \frac{X}{b} - \int_1^X R_b(t; \chi) \frac{\mathrm{d}t}{t} - \sum_{c \leqslant Z} \Lambda(c) R_{bc}(X; \chi) \right\}$$

The first and the last term together is bounded by

$$\sum_{d \leqslant YZ} |R_d(X; \chi)| \Big(\log \frac{X}{d} + \sum_{\substack{c \leqslant Z \\ c | d}} \Lambda(c) \Big) \leqslant R(X; Y, Z; \chi) \log X.$$

Thus we have the bound

$$|R(X; Y, Z; \chi)| \leqslant R(X, YZ; \chi) \log X + \int_1^X R(t, Y; \chi) \frac{\mathrm{d}t}{t}. \tag{3.19}$$

From Proposition 3.1.2, we already know that for $1 \leqslant t \leqslant Y$,

$$R(t;Y,\chi) = \sum_{d \leqslant Y} |R_d(t;\chi)| \ll_{\varepsilon'} q^3 Y^{1/4} t^{3/4+\varepsilon'}.$$

For $t > Y$, we have the trivial bound $R(t;Y,\chi) \ll t^{1+\varepsilon}(\log X)^{A+3}$ since by Lemma 2.1.3

$$\sum_{d \leqslant Y} |A_d(t;\chi)| \ \leqslant \ (\log X)^{A+1} \sum_{d \leqslant Y} \sum_{md \leqslant t} \tau(md) \ \ll_{\varepsilon} \ t^{1+\varepsilon}(\log X)^{A+3}$$

and

$$\sum_{d \leqslant Y} |M_d(t;\chi)| \ \leqslant \ (\log X)^{A+1} \Big(\sum_{d \leqslant Y} \frac{\rho(d)}{d} \Big) \Big(\sum_{F(\ell,m) \leqslant t} \sum 1 \Big) \ \ll_{\varepsilon} \ t^{1+\varepsilon}(\log X)^{A+2}.$$

Thus

$$\begin{aligned} \int_1^X \frac{R(t,Y;\chi)}{t}\, dt \ &= \ \int_Y^X \frac{R(t,Y;\chi)}{t}\, dt + \int_1^Y \frac{R(t,Y;\chi)}{t}\, dt \\ &\ll_{\varepsilon'} X^{3/4+\varepsilon'} Y^{1/4} + Y^{1+\varepsilon'}. \end{aligned}$$

Therefore if $YZ < X^{1-\varepsilon}$, we have

$$R(X;Y,Z;\chi) \ll_{\varepsilon,\varepsilon'} q^3 X^{1-\frac{\varepsilon}{4}+\varepsilon'}.$$

The result follows from choosing ε' accordingly. □

Chapter 4

Binary Quadratic Forms

4.1 Basic Terminology

In this section, we will review several basic notions for binary quadratic forms. We will mostly follow the terminology in [3]. For a binary quadratic form $f(x,y) = ax^2+bxy+cy^2 \in \mathbb{Z}[x,y]$, the discriminant is given by b^2-4ac. If the discriminant b^2-4ac is negative, we say that $f(x,y)$ is a *positive definite* binary quadratic form. **In the rest of the thesis, all forms are assumed to be positive definite.** $f(x,y)$ is called *primitive* if $\gcd(a,b,c) = 1$.

For example, x^2+y^2 is a positive definite binary quadratic form with discriminant -4. There are other binary quadratic forms of the same discriminant, but they are all essentially the same as x^2+y^2. To make this precise, we need the notion of equivalence between forms.

Definition 1 (Proper Equivalence)**.** Two forms $f(x,y)$ and $g(x,y)$ are *properly equivalent* if there are integers p,q,r and s such that

$$f(x,y) = g(px+qy, rx+sy) \qquad \text{and} \qquad ps-qr = 1.$$

If $ps-qr = \pm 1$, then $f(x,y)$ and $g(x,y)$ are said to be *equivalent*; we shall not use this.

Proper equivalence is indeed an equivalence relation. It is also clear that if $f(x,y)$ and $g(x,y)$ are properly equivalent, then they also have the same discriminant. Thus using proper equivalence, we can divide the set of all binary quadratic forms of a fixed discriminant into equivalence classes. For discriminant -4, it turns out that there is only one such class, and hence all forms are properly equivalent to x^2+y^2. For the positive definite case, there is a canonical way to choose a representative from each equivalence class. This can be achieved by using reduced forms:

Definition 2 (Reduced Forms)**.** A primitive positive definite form $ax^2+bxy+cy^2$ is *reduced* if

$$|b| \leqslant a \leqslant c, \text{ and } b \geqslant 0 \text{ if either } |b| = a \text{ or } a = c.$$

Then we have the following theorem, which is Theorem 1.2.8 of [3].

Theorem 4.1.1. Every primitive positive definite form is properly equivalent to a unique reduced form.

For any fixed (negative) discriminant, clearly there are only finitely many reduced forms. Hence the number of equivalence classes is finite. Furthermore, this finite set can be made into a finite Abelian group by the following operation:

Definition 3 (Dirichlet composition)**.** Let $f(x,y) = ax^2 + bxy + cy^2$ and $g(x,y) = \alpha x^2 + \beta xy + \gamma y^2$ be primitive positive definite forms of discriminant $-\Delta < 0$ which satisfy $\gcd(a, \alpha, (b+\beta)/2) = 1$. Then the *Dirichlet composition* of $f(x,y)$ and $g(x,y)$ is the form

$$h(x,y) = a\alpha x^2 + Bxy + \frac{B^2+\Delta}{4a\alpha}y^2$$

where B is any integer such that

$$\begin{aligned} B &\equiv b \pmod{2a} \\ B &\equiv \beta \pmod{2\alpha} \\ B^2 + \Delta &\equiv 0 \pmod{4a\alpha}. \end{aligned}$$

The term *composition* is justified by the following identity:

$$(au^2 + buv + cv^2)(\alpha X^2 + \beta XY + \gamma Y^2) = a\alpha W^2 + BWZ + \frac{B^2+\Delta}{4a\alpha}Z^2 \tag{4.1}$$

where

$$W = \left(u - \frac{B-b}{2a}v\right)X - \left(\frac{B-\beta}{2\alpha}u + \frac{(b+\beta)B - \Delta - b\beta}{4a\alpha}v\right)Y$$

and

$$Z = \alpha v X + \left(au + \frac{b+\beta}{2}v\right)Y.$$

The existence of a composition law was first provided by Gauss. However, the original formulation is a little bit messy and we will employ the version by Dirichlet. Finally, we recall the following theorem, which is a well-known consequence of Chebotarev's density theorem.

Theorem 4.1.2. Every primitive positive definite quadratic form represents infinitely many primes.

4.2 Factorization Proposition

The main purpose of this section is to prove a factorization proposition (Proposition 4.2.4) that will play an important role in Chapter 5. From Section 1.3, we understand that to produce a meaningful Type II estimate it is helpful to obtain some sort of factorization of

a_{mn} which takes approximately the form

$$\sum_{\substack{X,Y\in\mathbb{Z}\\ F(X,Y)=mn}}\sum \lambda(X).$$

In the Gaussian case studied by Fouvry and Iwanec, if $mn = X^2 + Y^2$ with $\gcd(m, n) = 1$, then we can write $m = a^2 + b^2, n = c^2 + d^2$ such that

$$ac + bd = X, \qquad ad - bc = Y.$$

This observation allows them to write

$$\sum_{m\sim M}\sum_{n\sim N}\Big(\sum_{\substack{X,Y\in\mathbb{Z}\\ X^2+Y^2=mn}}\sum \lambda(X)\Big) = \frac{1}{4}\sum_{\substack{a,b\in\mathbb{Z}\\ a^2+b^2\sim M}}\sum\sum_{\substack{c,d\in\mathbb{Z}\\ c^2+d^2\sim N}}\sum \lambda(ac+bd). \tag{4.2}$$

This identity is the basis of their Type II estimate. The composition law in the previous section gives us hope to generalize this to any primitive positive definite binary quadratic form.

In order to establish an analogue of (4.2), we study the equation

$$mn = F(X, Y)$$

when $\gcd(m, n) = 1$. From this equation it is possible to construct a binary quadratic form that represents m; and by composing its "inverse" with $F(x, y)$, we determine the form (up to proper equivalence) that represents n. However, the condition $\gcd(a, \alpha, (b + \beta)/2) = 1$ is needed when we apply Dirichlet composition. Therefore we have to be more selective about which representatives we used from the equivalence classes. The construction is given in the lemma below:

Lemma 4.2.1. Let t be a positive integer and $F(x, y) = \alpha x^2 + \beta xy + \gamma y^2$ be a primitive positive definite form with discriminant $-\Delta$. Then there exists a set, $\mathcal{S}_F(t)$, of binary quadratic forms of discriminant $-\Delta$ such that

1. every primitive binary quadratic form of discriminant $-\Delta$ is properly equivalent to exactly one element in $\mathcal{S}_F(t)$;
2. the principal form is contained in $\mathcal{S}_F(t)$; and
3. the set $\{f(1, 0) \ : \ f \in \mathcal{S}_F(t)\}$ consists of distinct primes that do not divide t.

Proof. If $\mathcal{S}_F(t)$ is the set of primitive reduced forms of discriminant $-\Delta$, then (1) and (2) are satisfied. Since each of these reduced forms represent infinitely primes by Theorem 4.1.2, we can transform the form so that the coefficient of x^2 is one of these primes and therefore

(3) is also satisfied. This can be done by the following lemma, which is a special case of Lemma 1.2.3 of [3].

Lemma 4.2.2. A form $f(x, y)$ represents a squarefree integer m if and only if $f(x, y)$ is properly equivalent to the form $mx^2 + Bxy + Cy^2$ for some $B, C \in \mathbb{Z}$.

□

Roughly speaking, the only information we used about F in defining $\mathcal{S}_F(t)$ is its discriminant; but in practice, the integer t would depend heavily on F.

We define $\mathcal{S}_F = \mathcal{S}_F(\alpha)$ and

$$Q_F = 2\alpha\gamma\Delta \prod_{f \in \mathcal{S}_F} f(1, 0). \tag{4.3}$$

We also pick an integer B with the following properties:

1. $B \equiv b \pmod{2a}$ for all $ax^2 + bxy + cy^2 \in \mathcal{S}_F$;
2. $B \equiv \beta \pmod{2\alpha}$; and
3. $B^2 + \Delta \equiv 0 \pmod{4a\alpha}$ for all $ax^2 + bxy + cy^2 \in \mathcal{S}_F$.

So B only depends on F and the choice of $\mathcal{S}_F$. Then we have the following crucial proposition, which will be used to treat our Type II sum.

Proposition 4.2.3. Let Δ be a positive integer and $F(x, y) = \alpha x^2 + \beta xy + \gamma y^2$ be a primitive binary quadratic form of discriminant $-\Delta$. Let m, n be positive integers such that $\gcd(mn, Q_F) = 1$. If $mn = F(X, Y)$ for some integers X, Y with $\gcd(X, Y) = 1$, then

(1) there exists a binary quadratic form $f(x, y) = ax^2 + bxy + cy^2 \in \mathcal{S}_F$ and integers u, v, w, z such that $\gcd(u, v) = \gcd(w, z) = 1$ and

$$\begin{aligned} au^2 + buv + cv^2 &= m, \\ a\alpha w^2 + Bwz + \frac{B^2 + \Delta}{4a\alpha} z^2 &= n, \\ \left(au + \frac{b + \beta}{2} v\right) w + \left(\frac{B - \beta}{2\alpha} u + \frac{(b + \beta)B + \Delta - b\beta}{4a\alpha} v\right) z &= X, \\ -\alpha vw + \left(u - \frac{B - b}{2a} v\right) z &= Y; \end{aligned}$$

(2) the choice of $f(x, y) \in \mathcal{S}_F$ is unique; and

(3) if $\Delta > 4$ then there is exactly one more tuple, namely $(-u, -v, -w, -z)$, that satisfies the properties. If $\Delta = 3$ or 4 we could have 6 or 4 solutions respectively.

Proof. **Proof of (1):**

Note that

$$4\alpha F(X,Y) = (2\alpha X + \beta Y)^2 + \Delta Y^2 \equiv 0 \pmod{m}.$$

If $\gcd(m, Y) > 1$, then $\gcd(m, \alpha X^2) > 1$ but this contradicts $\gcd(m, Q_F) = 1$. Therefore $\gcd(m, Y) = 1$ and

$$((2\alpha X + \beta Y)Y^{-1})^2 + \Delta \equiv 0 \pmod{m}.$$

Choose an integer ν such that $\nu \equiv (2\alpha X + \beta Y)Y^{-1} \pmod{m}$ and $\nu^2 + \Delta \equiv 0 \pmod{4}$. Since $\gcd(m, 2) = 1$, we have $4m | \nu^2 + \Delta$. Now we define the binary quadratic form

$$M(x,y) = mx^2 + \nu xy + \frac{\nu^2 + \Delta}{4m} y^2.$$

Since $\gcd(m, \Delta) = 1$, $M(x, y)$ is primitive. Thus $g(x, y)$ is properly equivalent to some $f(x, y) = ax^2 + bxy + cy^2 \in \mathcal{S}_F$. By definition, there exist integers u, v, r, s such that $us - rv = 1$ and

$$M(x,y) = f(ux + ry, vx + sy). \tag{4.4}$$

Therefore, $m = M(1, 0) = f(u, v) = au^2 + buv + cv^2$. By comparing the coefficients of xy in (4.4), we deduce that

$$\nu = 2aur + bus + brv + 2cvs$$

and it follows that

$$\begin{aligned} \nu v &= 2au(us - 1) + busv + brv^2 + 2cv^2 s \\ &\equiv -2au - buvs + brv^2 \pmod{m} \\ &\equiv -(2au + bv) \pmod{m}. \end{aligned}$$

Since $\nu \equiv (2\alpha X + \beta Y)Y^{-1} \pmod{m}$, we have

$$(2\alpha X + \beta Y)v \equiv -(2au + bv)Y \pmod{m}. \tag{4.5}$$

Consequently by $(\nu^2 + \Delta)vY \equiv 0 \pmod{m}$ we obtain

$$(2au + bv)(2\alpha X + \beta Y) - \Delta vY \equiv 0 \pmod{m}. \tag{4.6}$$

Now we have

$$m(mn) = (au^2 + buv + cv^2)(\alpha X^2 + \beta XY + \gamma Y^2) = a\alpha W^2 + BWZ + \frac{B^2 + \Delta}{4a\alpha} Z^2$$

where

$$W = \left(u - \frac{B - b}{2a} v\right) X - \left(\frac{B - \beta}{2\alpha} u + \frac{(b + \beta)B + \Delta - b\beta}{4a\alpha} v\right) Y$$

and

$$Z = \alpha v X + \Big(au + \frac{b+\beta}{2}v\Big)Y.$$

We claim that both W and Z are divisible by m. The second part follows from (4.5) and the fact that $\gcd(m, 2) = 1$. And

$$W = \frac{1}{4a\alpha}\Big((2au + bv)(2\alpha X + \beta Y) - \Delta v Y - 2BZ\Big).$$

hence the result follows from (4.6). Let

$$w = \frac{W}{m} \text{ and } z = \frac{Z}{m}. \tag{4.7}$$

Then w, z are both integers and $n = a\alpha w^2 + Bwz + \frac{B^2+\Delta}{4a\alpha}z^2$. Then by (4.7) we have

$$\Big(au + \frac{b+\beta}{2}v\Big)w + \Big(\frac{B-\beta}{2\alpha}u + \frac{(b+\beta)B + \Delta - b\beta}{4a\alpha}\Big)z = X,$$
$$-\alpha v w + \Big(u - \frac{B-b}{2a}v\Big)z = Y.$$

These equations also imply that $\gcd(u, v) | \gcd(X, Y)$ and $\gcd(z, w) | \gcd(X, Y)$, hence $\gcd(u, v) = \gcd(z, w) = 1$.

Proof of (2):

One can easily verify that

$$\Big(u - \frac{B-b}{2a}v\Big)X - \Big(\frac{B-\beta}{2\alpha}u + \frac{(b+\beta)B + \Delta - b\beta}{4a\alpha}v\Big)Y = mw \tag{4.8}$$

and

$$\alpha v X + \Big(au + \frac{b+\beta}{2}v\Big)Y = mz. \tag{4.9}$$

These imply

$$(2\alpha X + \beta Y)v \equiv -(2au + bv)Y \pmod{m}.$$

and

$$(2au + bv)(2\alpha X + \beta Y) - \Delta v Y \equiv 0 \pmod{m}.$$

Choose r, s so that $us - rv = 1$. Then the coefficient of xy in $H(x, y) = f(ux + ry, vx + sy)$ is congruent to $\nu \pmod{4m}$. Moreover, the coefficient of x^2 is $au^2 + buv + cv^2 = m$. By considering $H(x + ky, y)$ for a suitable k, $H(x, y)$, and hence $f(x, y)$ is properly equivalent to

$$M(x, y) = mx^2 + \nu xy + \frac{\nu^2 + \Delta}{4m}y^2.$$

Hence the choice of $f(x, y)$ is unique.

Proof of (3), for $\Delta > 4$:

Now suppose $\Delta > 4$ and there is another tuple (u_0, v_0, w_0, z_0) that satisfies the requirement. If $uv_0 - u_0v = 0$, then $u_0 = ku, v_0 = kv$ for some integer k and $n = k^2(au^2 + 2buv + cv^2) = k^2m$. Hence $k = \pm 1$ and we are done.

Suppose $uv_0 - u_0v \neq 0$. Then

$$4am = (2au + bv)^2 + \Delta v^2 = (2au_0 + bv_0)^2 + \Delta v_0^2. \tag{4.10}$$

Then since

$$zm = \alpha vX + \left(au + \frac{b+\beta}{2}v\right)Y,$$

we have

$$au + \frac{b+\beta}{2}v \equiv -(xy^{-1})\alpha v \pmod{m}$$

and similarly for u_0 and v_0. Therefore

$$\left(au + \frac{b+\beta}{2}v\right)v_0 - \left(au_0 + \frac{b+\beta}{2}v_0\right)v \equiv -(xy^{-1})\alpha vv_0 + (xy^{-1})\alpha v_0 v \equiv 0 \pmod{m}.$$

Hence $a(uv_0 - u_0v) \equiv 0 \pmod{m}$. By assumption we have $\gcd(m, a) = 1$, so it follows that $m|(uv_0 - u_0v)$. Therefore from (4.10) we deduce that

$$(4am)^2 = \left((2au + bv)(2au_0 + bv_0) + \Delta vv_0\right)^2 + \Delta\left(2a(uv_0 - u_0v)\right)^2$$

and $(2au + bv)(2au_0 + bv_0) + vv_0$ is a multiple of am. Thus,

$$16 = \left(\frac{(2au + bv)(2au_0 + bv_0) + \Delta vv_0}{am}\right)^2 + \Delta\left(\frac{2(uv_0 - u_0v)}{m}\right)^2. \tag{4.11}$$

But then

$$16 \geqslant 0 + 4\Delta > 16$$

and this gives us a contradiction.

Proof of (3), for $\Delta = 4$:

If $\Delta = 4$, we can take $\mathcal{S}_F = \{x^2 + y^2\}$ and $B = \beta$ (note that β must be even). Then we

have

$$\begin{aligned} u^2+v^2 &= m,\\ \alpha w^2+\beta wz+\gamma z^2 &= n,\\ \Big(u+\frac{\beta}{2}v\Big)w+\gamma vz &= X,\\ -\alpha vw+\Big(u-\frac{\beta}{2}v\Big)z &= Y. \end{aligned}$$

If (u,v,w,z) satisfies the above equations, it is straightforward to verify that all

$$(-u,-v,-w,-z),\Big(v,-u,\frac{\beta}{2}w+\gamma z,-\alpha w-\frac{\beta}{2}z\Big),\Big(-v,u,-\frac{\beta}{2}w-\gamma z,\alpha w+\frac{\beta}{2}z\Big)$$

satisfy them as well, and we claim these are all the possible solutions. Note that (4.11) becomes

$$1=\left(\frac{uu_0+vv_0}{m}\right)^2+\left(\frac{uv_0-u_0v}{m}\right)^2$$

and it forces $uu_0+vv_0=0$ and $uv_0-u_0v=\pm m$. Hence $(u_0,v_0)=(v,-u)$ or $(-v,u)$. To find (w_0,z_0) now we can simply solve the system of linear equations.

Proof of (3), for $\Delta=3$:
The case for $\Delta=3$ follows similarly. □

From Proposition 4.2.3, we are able to deduce the following factorization proposition for our sequence (a_n).

Proposition 4.2.4. If $\gcd(m,n)=\gcd(mn,Q_F)=1$, we have

$$a_{mn}=\frac{1}{2}\sum_{f\in\mathcal{S}_F}\sum_{\substack{(w,z)\in\mathbb{Z}^2\\ f^*(w,z)=m\\ \gcd(w,z)=1}}\sum_{\substack{(u,v)\in\mathbb{Z}^2\\ f(u,v)=n\\ \gcd(u,v)=1}}\lambda(\mathcal{Q}_F(u,v;w,z)) \tag{4.12}$$

where

$$f^*(w,z)=a\alpha w^2+Bwz+\frac{B^2+\Delta}{4a\alpha}z^2.$$

and

$$\mathcal{Q}_F(u,v;w,z)=\Big(au+\frac{b+\beta}{2}v\Big)w+\Big(\frac{B-\beta}{2\alpha}u+\frac{(b+\beta)B+\Delta-b\beta}{4a\alpha}v\Big)z.$$

Here we assume that $\lambda(\ell)=0$ if $\ell<0$. If $\Delta=3$ or 4, the constant $\frac{1}{2}$ before the summation should be $\frac{1}{6}$ and $\frac{1}{4}$ respectively.

Proof. Proposition 4.2.3 gives a 1 to 2 map (1 to 4 if $\Delta = 4$, 1 to 6 if $\Delta = 3$) from the set

$$\Big\{(X,Y) \in \mathbb{Z}^2 \mid F(X,Y) = mn\Big\} \tag{4.13}$$

to the set

$$\Big\{(f,u,v,w,z) \in \mathcal{S}_F \times \mathbb{Z}^4 \mid f(u,v) = m, f^*(w,z) = n, \gcd(u,v) = \gcd(w,z) = 1\Big\}. \tag{4.14}$$

For the backward direction, if we have $f(u,v) = m$ and $f^*(w,z) = n$, by Dirichlet composition they can produce $X, Y \in \mathbb{Z}$ such that $F(X,Y) = mn$ via

$$\begin{aligned}\left(au + \frac{b+\beta}{2}v\right)w + \left(\frac{B-\beta}{2\alpha}u + \frac{(b+\beta)B + \Delta - b\beta}{4a\alpha}v\right)z &= X,\\ -\alpha vw + \left(u - \frac{B-b}{2a}v\right)z &= Y.\end{aligned}$$

Assume $\gcd(m,n) = \gcd(u,v) = \gcd(w,z) = 1$. Then by (4.8) and (4.9), we have

$$\gcd(X,Y) | \gcd(mw, mz) = m$$

and similarly $\gcd(X,Y)|n$. Thus $\gcd(X,Y) = 1$. This establishes the backward map from (4.14) to (4.13) and hence (4.12) follows. □

Chapter 5

Type II Estimate

5.1 Outline

In this chapter we shall estimate $B(X;Y,Z;\chi)$ given in (2.11) by proving the following proposition:

Proposition 5.1.1. Let $F(x,y)=\alpha x^2+\beta xy+\gamma y^2\in\mathbb{Z}[x,y]$ be a positive definite binary quadratic form. Suppose $q\in\mathbb{N}$ and $q\leqslant(\log X)^Q$ for some $Q>0$. Let θ_1,θ_2 be two real numbers such that $1/2<\theta_1<1$ and $0<\theta_2<1-\theta_1$. Then for $Y=X^{\theta_1}$ and $Z=X^{\theta_2}$ and any $C>0$,

$$B(X;Y,Z;\chi)\ll_C X(\log X)^{-C}.$$

Recall from (2.11) and (1.12) that

$$B(X;Y,Z;\chi)=\sum_{\substack{bd\leqslant X\\ b>Y\\ \gcd(bd,P_F)=1}}\sum\mu(b)\left(\sum_{\substack{c\mid d\\ c>Z}}\Lambda(c)\right)\chi(bd)a_{bd}$$

and

$$a_N=\sum_{\substack{\ell\in\mathbb{N},m\in\mathbb{Z}\\ F(\ell,m)=N\\ \gcd(\ell,\gamma m)=1}}\sum\lambda(\ell).$$

We assumed that $|\lambda(\ell)|\leqslant\log^A X$ for some $A>0$; for simplicity, here we might simply assume that $|\lambda(\ell)|\leqslant 1$ for all $\ell\in\mathbb{N}$. Then the general case follows by considering the weights $\left(\frac{\lambda(\ell)}{\log^A X}\right)$.

At the end of the proof we will apply the following result by Helfgott, which is a special case of Lemma 3.3.6 of [10].

Lemma 5.1.2. Let $Q(x,y)$ be a primitive positive definite quadratic form. Let $H \leqslant (\log X)^N$. Then for any $h_1, h_2 \pmod H$, any $A > 0$ and sector $S \subset \mathbb{R}^2$,

$$\sum_{\substack{Q(x,y)\leqslant X \\ x\equiv h_1 \pmod H \\ y\equiv h_2 \pmod H \\ (x,y)\in S}}\sum \mu(Q(x,y)) \ll_{A,N} X(\log X)^{-A}.$$

Here sector refers to any region $S \subset \mathbb{R}^2$ given by

$$S = \{(x,y) \in \mathbb{R}^2 \mid \theta_1 \leqslant \arg(x+yi) \leqslant \theta_2\}$$

for some $\theta_1, \theta_2 \in \mathbb{R}$. Helfgott proved this for the Liouville function λ but the same proof also works for μ.

The proof of Proposition 5.1.1 is very long and we have divided it into three parts. The first part, given in Section 5.2, reduces $B(X;Y,Z;\chi)$ to a much simpler sum. In the second part, we apply Proposition 4.2.4 to unfold the multiplicative structure of a_{mn}; this will be done in Section 5.3. The last part of the proof involves a series of technical calculations and the application of Lemma 5.1.2. That will be covered in Section 5.4.

5.2 Part I: Reduction

We define $\theta = (\log X)^{-2C}$ and write

$$\mathcal{B}(M,N) = \sum_{\substack{M<m\leqslant M' \\ \gcd(m,P_F)=1}} \left| \sum_{\substack{N<n<N' \\ \gcd(n,P_F)=1}} \mu(n)\chi(mn)a_{mn} \right|, \tag{5.1}$$

where $M' = e^\theta M$ and $N' = e^\theta N$. Using these sums for $M = e^{j\theta}Z$ and $N = e^{k\theta}Y$ we get

$$|B(X;Y,Z;\chi)| \leqslant (\log X) \sum_{\substack{\theta x<MN<X \\ M\geq Z, N\geq Y}}\sum \mathcal{B}(M,N) + O\left(\theta X(\log X)^4\right) \tag{5.2}$$

where the error term $O(\theta X(\log X)^4)$ represents a trivial bound for the contribution of $\mu(b)\chi(bd)a_{bd}$ with $\theta X \leqslant bd \leqslant e^{2\theta}\theta X$ or $X \leqslant bd \leqslant e^{2\theta}X$, which terms are not covered exactly. Precisely, we have

$$\sum_{\theta X\leqslant bd\leqslant e^{2\theta}\theta X}\sum |a_{bd}| \leqslant \sum_{\theta X\leqslant N\leqslant e^{2\theta}\theta X} \tau^2(N) \ll \theta^2 X(\log X)^4$$

and

$$\sum_{X\leqslant bd\leqslant e^{2\theta}X}\sum |a_{bd}| \leqslant \sum_{X\leqslant N\leqslant e^{2\theta}X} \tau^2(N) \ll \theta X(\log X)^4$$

by Lemma 2.1.3 and (3.17). Next, we need to show that each short sum $\mathcal{B}(M,N)$ satisfies

$$\mathcal{B}(M,N) \ll \theta^2 X. \tag{5.3}$$

In order to apply the factorization proposition, we need to somehow impose the extra condition $\gcd(m,n)=1$ in $\mathcal{B}(M,N)$. Let $\mathcal{B}_d(M,N)$ denote the sum (5.1) restricted to $\gcd(m,n)=d$. We have

$$\mathcal{B}(M,N) \leqslant \sum_{d<\theta^{-1}} \mathcal{B}_d(M,N) + O\left(\theta^2 X\right)$$

where the error term $O(\theta^2 X)$ represents a trivial bound for the contribution of $\mu(n)\chi(mn)a_{mn}$ with $\gcd(m,n)\geq\theta^{-1}$. Note that

$$\mathcal{B}_d(M,N) \leqslant \mathcal{B}_1(dM, N/d). \tag{5.4}$$

Therefore, the proof of Lemma 5.1.1 is reduced to showing the estimate

$$\mathcal{B}_1(M,N) \ll \theta^3 X \tag{5.5}$$

holds for any M,N with $M\geq Z, N\geq\theta Y$ and $\theta X < MN < X$.

5.3 Part II: Applying the Factorization Proposition

Define $\kappa(n)=\mu(n)\chi(n)$. The purpose of this section is to apply Proposition 4.2.4 to the sum

$$\mathcal{B}_1(M,N) = \sum_{\substack{M<m\leqslant 2M\\ \gcd(m,P_F)=1}} \left| \sum_{\substack{N<n<N'\\ \gcd(m,n)=1\\ \gcd(n,P_F)=1}} \kappa(n)a_{mn}\right|.$$

Proposition 4.2.4 allows us to decompose a_{mn} into solutions of

$$f(u,v)=m, \qquad f^*(w,z)=n$$

where f^* is the form constructed in Proposition 4.2.3,

$$f^*(x,y) = a\alpha x^2 + Bxy + \frac{B^2+\Delta}{4a\alpha}y^2.$$

Unfortunately, later in (5.8) we need to decompose the solutions of $f^*(w,z)=n$ again in a similar fashion. Therefore, we construct the set of binary quadratic forms $\mathcal{S}_{f^*}$ in the same way we construct $\mathcal{S}_F$ by taking

$$\mathcal{S}_{f^*} = \mathcal{S}_{f^*}\Big(\alpha \prod_{f\in\mathcal{S}_F} f(1,0)\Big).$$

We also pick an integer B such that

1. $B \equiv b \pmod{2a}$ for all $ax^2+bxy+cy^2 \in \mathcal{S}_F$;
2. $B \equiv e \pmod{2d}$ for all $dx^2+exy+fy^2 \in \mathcal{S}_{f^*}$;
3. $B \equiv \beta \pmod{2\alpha}$; and
4. $B^2+\Delta \equiv 0 \pmod{4ad\alpha}$ for all $ax^2+bxy+cy^2 \in \mathcal{S}_F$ and $dx^2+exy+fy^2 \in \mathcal{S}_{f^*}$.

Such B always exists since the coefficients of x^2 of elements in $\mathcal{S}_F$ or $\mathcal{S}_{f^*}$ are distinct primes. So B depends only on F and the choices of $\mathcal{S}_F$ and $\mathcal{S}_{f^*}$. We can make our choices of $\mathcal{S}_F$ and $\mathcal{S}_{f^*}$ canonical by ordering the forms by the sizes of their coefficients and using the smallest prime available for coefficient of x^2. In this way, the choice of B depends only on F. Finally, we define P_F by

$$P_F = Q_F \prod_{f\in\mathcal{S}_F} Q_{f^*} \tag{5.6}$$

where Q_{f^*} is defined as in (4.3). Precisely,

$$Q_{f^*} = 2f^*(1,0)f^*(0,1)\Delta \prod_{f\in\mathcal{S}_{f^*}} f(1,0).$$

Then the condition $\gcd(mn,P_F)=1$ would immediately imply $\gcd(mn,Q_F)=1$ and $\gcd(mn,Q_{f^*})=1$ all $f\in S_F$. By Proposition 4.2.4, we can bound $\mathcal{B}_1(M,N)$ by

$$\mathcal{B}_1(M,N) \leqslant \sum_{f\in\mathcal{S}_F} \sum_{\substack{M<f(u,v)\leqslant 2M\\ \gcd(f(u,v),P_F)=1\\ \gcd(u,v)=1}} \Bigg| \sum_{\substack{N<f^*(w,z)\leqslant N'\\ \gcd(f(u,v)P_F,f^*(w,z))=1\\ \gcd(w,z)=1}} \kappa(f^*(w,z))\lambda(\mathcal{Q}_F(u,v;w,z))\Bigg| \tag{5.7}$$

with

$$\mathcal{Q}_F(u,v;w,z) = \Big(au+\frac{b+\beta}{2}v\Big)w + \Big(\frac{B-\beta}{2\alpha}u + \frac{(b+\beta)B+\Delta-b\beta}{4a\alpha}v\Big)z.$$

We relax the condition $\gcd(f(u,v),f^*(w,z))=1$ by the classical identity

$$\sum_{r\mid \gcd(m,n)} \mu(r) = \begin{cases} 1 & \text{if } \gcd(m,n)=1,\\ 0 & \text{otherwise.}\end{cases}$$

Since n is squarefree, by Proposition 4.2.3 we can decompose $f^*(w,z)=n$ as

$$g(u_0,v_0)=r, g^*(w_0,z_0)=\frac{n}{r} \tag{5.8}$$

for some $g\in\mathcal{S}_{f^*}$ and we have the relations

$$\begin{aligned}\left(du_0+\frac{e+B}{2}v_0\right)w_0+\left(\frac{(B+e)B+\Delta-Be}{4da\alpha}v_0\right)z_0&=w,\\ -a\alpha v_0w_0+\left(u_0-\frac{B-e}{2d}v_0\right)z_0&=z.\end{aligned} \tag{5.9}$$

The inner sum of (5.7) then becomes

$$\sum_{g\in\mathcal{S}_{f^*}}\sum_{g(u_0,v_0)=r}\mu(r)\sum_{\substack{N<rg^*(w_0,z_0)<N'\\ \gcd(g^*(w_0,z_0),P_F)=1\\ \gcd(w_0,z_0)=1}}\kappa(rg^*(w_0,z_0))\lambda(\mathcal{Q}_F(u,v;w,z)).$$

Now it suffices to deal with

$$\sum_r{}^{\flat}\sum_{g(u_0,v_0)=r}\sum_{\substack{M<f(u,v)\leqslant 2M\\ \gcd(f(u,v),P_F)=1\\ \gcd(u,v)=1\\ r|f(u,v)}}\left|\sum_{\substack{N<rg^*(w_0,z_0)<N'\\ \gcd(g^*(w_0,z_0),P_F)=1\\ \gcd(w_0,z_0)=1}}\kappa(rg^*(w_0,z_0))\lambda(\mathcal{Q}_F(u,v;w,z))\right| \tag{5.10}$$

where w and z are determined by (5.9) and $\sum^{\flat}$ is a summation over squarefree integers. Next, we combine $f(u,v)$ and $g(u_0,v_0)$ via

$$f(u,v)g(u_0,v_0)=adP^2+BPQ+\frac{B^2+\Delta}{4ad}Q^2=H(P,Q),$$

where

$$\begin{aligned}P&=\left(u-\frac{B-b}{2a}v\right)u_0-\left(\frac{B-e}{2d}u+\frac{(b+e)B+\Delta-be}{4ad}v\right)v_0,\\ Q&=dvu_0+\left(au+\frac{b+e}{2}v\right)v_0.\end{aligned} \tag{5.11}$$

When (P,Q) and (u_0,v_0) are fixed, there is at most one pair (u,v) such that (5.11) holds. Therefore the sum (5.10) is bounded by

$$\sum_r{}^{\flat}\rho(r)\sum_{\substack{rM<H(P,Q)\leqslant 2rM\\ \gcd(H(P,Q),P_F)=1\\ r^2|H(P,Q)}}\left|\sum_{\substack{N<rg^*(w_0,z_0)<N'\\ \gcd(g^*(w_0,z_0),P_F)=1\\ \gcd(w_0,z_0)=1}}\kappa(rg^*(w_0,z_0))\lambda(\mathcal{Q}_F(u,v;w,z))\right|.$$

The above process is very similar to what we have done in (5.4), except now the calculation involves binary quadratic forms. Before we move on, we need to express $\mathcal{Q}_F(u,v;w,z)$ in terms of P,Q,w_0,z_0. By (5.9),

$$\begin{aligned}\mathcal{Q}_F(u,v;w,z) &= \Big(au+\frac{b+\beta}{2}v\Big)w+\Big(\frac{B-\beta}{2\alpha}u+\frac{(b+\beta)B+\Delta-b\beta}{4a\alpha}v\Big)z\\ &= \Big(au+\frac{b+\beta}{2}v\Big)\Big[\Big(du_0+\frac{e+B}{2}v_0\Big)w_0+\Big(\frac{(B+e)B+\Delta-Be}{4da\alpha}v_0\Big)z_0\Big]\\ &\quad+\Big(\frac{B-\beta}{2\alpha}u+\frac{(b+\beta)B+\Delta-b\beta}{4a\alpha}v\Big)\Big[-a\alpha v_0w_0+\Big(u_0-\frac{B-e}{2d}v_0\Big)z_0\Big].\end{aligned}$$

With (5.11), this gives

$$\mathcal{Q}_F(u,v;w,z)=\Big(adP+\frac{B+\beta}{2}Q\Big)w_0+\Big(\frac{B-\beta}{2\alpha}P+\frac{\Delta+B^2}{4da\alpha}Q\Big)z_0.$$

To simplify our notations, we define

$$U=\frac{B-\beta}{2\alpha}P+\frac{\Delta+B^2}{4da\alpha}Q,\qquad V=-adP-\frac{B+\beta}{2}Q.$$

Then it is not difficult to check that

$$h(U,V):=adU^2+BUV+\alpha\frac{\Delta+B^2}{4da\alpha}V^2=\frac{\Delta+\beta^2}{4\alpha}H(P,Q).$$

Therefore it suffices to prove that

$$\sum_r{}^{\flat}\rho(r)\sum_{\substack{rM<h(P,Q)\leqslant 2rM\\ \gcd(h(P,Q),P_F)=1\\ r^2\mid h(P,Q)}}\Big|\sum_{\substack{N<rg^*(w_0,z_0)<N'\\ \gcd(g^*(w_0,z_0),P_F)=1\\ \gcd(w_0,z_0)=1}}\kappa(rg^*(w_0,z_0))\lambda(z_0P-w_0Q)\Big|\ll\theta^3X$$

for any M,N with $M\geq Z, N\geq\theta Y$ and $\theta X<MN<X$.

Before we end this section, we wish to clean up the $\rho(r)$ in the first summation and insert the condition $\gcd(P,Q)=1$ in our sum. The latter condition will be useful in (5.15). First of all, estimating trivially, we find that the terms with $r\geq\theta^{-3}$ contribute

$$O\left(\theta MN\sum_{r>\theta^{-2}}\tau(r)^2r^{-2}\right)=O\left(\theta^4X(\log X)^4\right).$$

In the remaining terms, we ignore the conditions $r^2|h(P,Q)$, $\gcd(h(P,Q),P_F)=1$ and obtain the bound

$$\mathcal{B}_1(M,N) \leqslant \sum_{r<\theta^{-3}} \rho(r) \sum_{rM<h(P,Q)\leqslant 2rM} \Big| \sum_{\substack{N<rg^*(w_0,z_0)<N' \\ \gcd(g^*(w_0,z_0),P_F)=1 \\ \gcd(w_0,z_0)=1}} \kappa(rg^*(w_0,z_0))\lambda(z_0P-w_0Q)) \Big| + O(\theta^3 X).$$

Put

$$\mathcal{C}_r(M,N) = \sum_{M<h(P,Q)\leqslant 2M} \Big| \sum_{\substack{N<g^*(w_0,z_0)<N' \\ \gcd(g^*(w_0,z_0),P_F)=1 \\ \gcd(w_0,z_0)=1}} \kappa(rg^*(w_0,z_0))\lambda(z_0P-w_0Q) \Big|.$$

Our goal is to show that $\mathcal{C}_r(M,N)=O(\theta^6 X)$. We then write

$$\mathcal{C}_{c,r}(M,N) = \sum_{\substack{M<h(P,Q)\leqslant 2M \\ \gcd(P,Q)=1}} \Big| \sum_{\substack{N<g^*(w_0,z_0)<N' \\ \gcd(g^*(w_0,z_0),P_F)=1 \\ \gcd(w_0,z_0)=1}} \kappa(rg^*(w_0,z_0))\lambda(c(z_0P-w_0Q)) \Big|.$$

Then using the trivial bound $\mathcal{C}_{c,r}(M,N) \ll \theta MN$, we have

$$\mathcal{C}_r(M,N) = \sum_{c\geqslant 1} \mathcal{C}_{c,r}(M/c^2,N) = \sum_{1\leqslant c\leqslant \theta^4} \mathcal{C}_{c,r}(M/c^2,N) + O(\theta^5 X).$$

In conclusion, it then suffices to give a bound of the shape

$$\mathcal{C}_{c,r}(M,N) \ll \theta^6 MN$$

for every c,r,M,N with $c<\theta^{-4}, r<\theta^{-3}, M\geqslant \theta^8 Z, N>\theta^4 Y$, and $\theta^9 X<MN<X$. The precise powers of θ only have minor effects on our argument, since we will be able to save an arbitrary power of $\log X$ at the end. Our assumptions in Proposition 5.1.1 guarantee that M,N satisfy $N^\varepsilon < M < N^{1-\varepsilon}$ for some small $\varepsilon>0$. This assumption will be used in (5.21) and (5.25) and we will give a bound of the form

$$\mathcal{C}_{c,r}(M,N) \ll_j MN(\log N)^{-j} \tag{5.12}$$

for arbitrary $j>0$. We leave the final calculation to the next section.

5.4 Part III: Final Calculation

Let $A = \sqrt{N/r}, B = \sqrt{M}$ and $\kappa(u,v) = \kappa(rg^*(u,v))$. Then $\kappa(u,v)$ is supported in the annulus $A^2 < g^*(u,v) \leqslant 4A^2$. A simple application of Cauchy-Schwarz inequality gives

$$\begin{aligned}|\mathcal{C}_{cr}(M,N)| &\leqslant \sum_{\ell} |\lambda(\ell)| \sum_{\substack{M<h(w,z)<2M \\ \gcd(w,z)=1}} \Big| \sum_{vw-uz=\ell} \kappa(u,v)\Big| \\ &\leqslant A^{1/2}B^{3/2}\mathcal{D}(\kappa)^{1/2},\end{aligned}$$

where

$$\mathcal{D}(\kappa) = \sum_{\substack{(w,z)\in\mathbb{Z}^2 \\ \gcd(w,z)=1}} \psi(w,z) \sum_{\ell} \Big| \sum_{Q(u,v;w,z)=\ell} \kappa(u,v)\Big|^2 \tag{5.13}$$

and

$$Q(u,v;w,z) = vw - uz.$$

Here $\psi(w,z)$ can be any non-negative function with $\psi(w,z) \geq 1$ if $B^2 \leqslant h(w,z) \leqslant 4B^2$. We assume that $\psi(w,z)$ takes the form $\Psi(h(w,z))$, where

$$0 \leqslant \Psi(t) \leqslant 1, \Psi(t) = 1 \text{ if } B^2 \leqslant t \leqslant 4B^2,$$

$$\operatorname{supp}\Psi \subset [B^2/4, 9B^2], \Psi^{(j)} \ll B^{-2j}.$$

Our desired estimate for $\mathcal{D}(\kappa)$ is A^3B with a saving of an arbitrary power of $\log N$. Since ℓ runs over all integers (without any restriction), after squaring we obtain

$$\mathcal{D}(\kappa) = \sum_{\substack{(w,z)\in\mathbb{Z}^2 \\ \gcd(w,z)=1}} \psi(w,z) \sum_{Q(u,v;w,z)=0} (\kappa * \kappa)(u,v), \tag{5.14}$$

where

$$(\kappa * \kappa)(u,v) = \sum_{(s_1,t_1)-(s_2,t_2)=(u,v)} \kappa(s_1,t_1)\overline{\kappa}(s_2,t_2).$$

This equality follows because $Q(u,v;w,z)$ is a bilinear form. Note that

$$(\kappa * \kappa)(0,0) \ll A^2.$$

The orthogonality relation $Q(u,v;w,z) = 0$ in (5.14) is equivalent to

$$(u,v) = (cw, cz) \tag{5.15}$$

for some rational integer $c \in \mathbb{Z}$. It thus follows that

$$\begin{aligned}\mathcal{D}(\kappa) =& \sum_{c\in\mathbb{Z}} \sum_{\substack{(w,z)\in\mathbb{Z}^2\\ \gcd(w,z)=1}} \psi(w,z)(\kappa*\kappa)(cw,cz)\\ =& \mathcal{D}_0(\alpha) + 2\mathcal{D}^*(\kappa),\end{aligned} \tag{5.16}$$

say, where $\mathcal{D}_0(\kappa)$ denotes the contribution of $c = 0$ and $\mathcal{D}^*(\kappa)$ that of all $|c| > 0$. Thus

$$\mathcal{D}_0(\kappa) = \|\kappa\|^2 \sum_{\substack{(w,z)\in\mathbb{Z}^2\\ \gcd(w,z)=1}} \psi(w,z) \ll A^2B^2 \tag{5.17}$$

and

$$\mathcal{D}^*(\kappa) = \sum_{(s,t)\neq(0,0)} \psi\left(\frac{s}{\gcd(s,t)}, \frac{t}{\gcd(s,t)}\right)(\kappa*\kappa)(s,t). \tag{5.18}$$

We trade the primitivity condition for congruence conditions by means of Möbius inversion, getting

$$\mathcal{D}^*(\kappa) = \sum_{b,c>0} \mu(b)\mathcal{D}(\kappa;b,c) \tag{5.19}$$

where

$$\mathcal{D}(\kappa;b,c) = \sum_{(s,t)\equiv(0,0) \pmod{bc}} \psi\left(\frac{s}{c}, \frac{t}{c}\right)(\kappa*\kappa)(s,t). \tag{5.20}$$

Note that $g^*(s,t) \leqslant 2A$ (from the support of α) and $cB/2 < g^*(s,t) < 3cB$ (from the support of ψ). Observe that these imply that $c < 4AB^{-1}$, otherwise $\mathcal{D}(\kappa;b,c)$ is zero. Let Ξ be a parameter such that

$$1 \leqslant \Xi \leqslant 4AB^{-1} = C, \tag{5.21}$$

say. We will take Ξ to be a power of $\log N$ at the end and this explains why N needs to be larger than M, say $N^{1-\epsilon} > M$. By the trivial bound

$$\mathcal{D}(\kappa;b,c) \ll A^2B^2b^{-2}$$

we see that the terms with $b \geq \Xi$ or $c \leqslant C\Xi^{-1}$ contribute at most $O(A^3B\Xi^{-1})$ to $\mathcal{D}^*(\kappa)$ so

$$\mathcal{D}^*(\kappa) = \sum_{b\leqslant\Xi} \mu(b) \sum_{C\Xi^{-1}<c<C} \mathcal{D}(\kappa;b,c) + O\left(A^3B\Xi^{-1}\right). \tag{5.22}$$

Our next step is to rewrite $\psi(s/c, t/c)$ in $\mathcal{D}(\kappa; b, c)$ so that we can completely forget about the support of ψ. If $h(w,z) = Dw^2 + Ewz + Fz^2$ with $D > 0$, then

$$\begin{aligned}\psi(w,z) &= \Psi\Big(\frac{(2Dw+Ez)^2 + (4DF - E^2)z^2}{4D}\Big)\\ &= \Psi\Big(\Big(\frac{2Dw+Ez}{2\sqrt{D}}\Big)^2 + \Big(\frac{\sqrt{|\Delta|}z}{2\sqrt{D}}\Big)^2\Big).\end{aligned}$$

Then we can define

$$\psi_0(w,z) = \psi\Big(\frac{\sqrt{|\Delta|}w - Ez}{\sqrt{D|\Delta|}}, \frac{2\sqrt{D}z}{\sqrt{\Delta}}\Big).$$

This gives

$$\psi(w,z) = \psi_0\Big(\frac{2Dw+Ez}{2\sqrt{D}}, \frac{\sqrt{|\Delta|}z}{2\sqrt{D}}\Big) \quad \text{and} \quad \psi_0(w,z) = \Psi(w^2+z^2).$$

Hence if we define

$$\phi(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi_0(w,z) e(-(xw+yz))\, dw\, dz,$$

by a standard change of variables we obtain

$$\phi(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi\Big(\frac{2Dw+Ez}{2\sqrt{D}}, \frac{\sqrt{|\Delta|}z}{2\sqrt{D}}\Big) e(-(xw+yz))\, dw\, dz.$$

As $\phi(x,y)$ depends only on x^2+y^2, we can set $\phi(x,y) = \Phi(x^2+y^2)$. By inversion and an another change of variables, we deduce that

$$\begin{aligned}\psi(w,z) &= \psi_0\Big(\frac{2Dw+Ez}{2\sqrt{D}}, \frac{\sqrt{|\Delta|}z}{2\sqrt{D}}\Big)\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi(x,y) e\Big(x\cdot\frac{2Dw+Ez}{2\sqrt{D}} + y\cdot\frac{\sqrt{|\Delta|}z}{2\sqrt{D}}\Big)\, dx\, dy\\ &= \frac{2}{\sqrt{|\Delta|}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi\Big(\frac{x}{\sqrt{|D|}}, -\frac{Ex}{\sqrt{D|\Delta|}} + \frac{2\sqrt{D}y}{\sqrt{|\Delta|}}\Big) e(xw+yz)\, dx\, dy\\ &= \frac{2}{\sqrt{|\Delta|}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Phi\Big(\frac{4h(-y,x)}{|\Delta|}\Big) e(xw+yz)\, dx\, dy.\end{aligned}$$

Thus

$$\psi\Big(\frac{w}{c}, \frac{z}{c}\Big) = \frac{2c^2}{\sqrt{|\Delta|}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Phi\Big(\frac{4c^2 h(-y,x)}{|\Delta|}\Big) e(xw+yz)\, dx\, dy \tag{5.23}$$

and consequently

$$\mathcal{D}(\kappa; b, c) = \frac{2c^2}{\sqrt{|\Delta|}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Phi\Big(\frac{4c^2 h(-y,x)}{|\Delta|}\Big) S_{bc}(x,y)\, dx\, dy$$

where

$$\begin{aligned} S_d(x,y) &= \sum_{(s,t)\equiv(0,0) \pmod d} (\kappa * \kappa)(s,t) e(xs+yt) \\ &= \sum_{\substack{s_1 \equiv s_2 \pmod d \\ t_1 \equiv t_2 \pmod d}} \sum \kappa(s_1,t_1)\overline{\kappa}(s_2,t_2) e(xs_1+yt_1)e(-xs_2-yt_2) \\ &= \sum_{d_1,d_2 \pmod d} \Big| \sum_{\substack{s \equiv d_1 \pmod d \\ t \equiv d_2 \pmod d}} \kappa(s,t) e(xs+yt) \Big|^2. \end{aligned}$$

By (9.14) of [5],

$$c^2 \Phi\Big(\frac{4c^2 h(-y,x)}{|\Delta|}\Big) \ll \frac{c^2 B^2}{(1+c^2B^2h(-y,x))^{3/2}} \ll \frac{A^2\Xi}{(1+h(-y,x)A^2)^{3/2}}.$$

Hence

$$\mathcal{D}(\kappa; b, c) \ll \Xi A^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(x,y) S_{bc}(x,y)\, dx\, dy$$

where

$$H(x,y) = \frac{1}{(1+h(-y,x)A^2)^{3/2}}.$$

By grouping $d = bc$ and setting $D = C\Xi$, we obtain from (5.22)

$$\mathcal{D}^*(\alpha) \ll M\Xi^3 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(x,y)\Big(\sum_{d \leqslant D} d^2 S_d(x,y)\Big)\, dx\, dy + A^3 B \Xi^{-1}. \tag{5.24}$$

To account for the large d appearing in the above sum, we need to invoke Proposition 15 of [5].

Proposition 5.4.1. Suppose $A \geqslant D \geqslant 1$. Let f be a complex-valued function on $\mathbb{Z}[i]$ supported on the disc $|z| \leqslant A$. Define

$$S_f(D) = \sum_{d \leqslant D} d^2 \sum_{\delta \pmod d} \Big| \sum_{z \equiv \delta \pmod d} f(z) \Big|^2.$$

Then for any $G \geqslant 1$ we have

$$S_f(D) \leqslant 2DS_f(G) + O_\varepsilon(AD(D^{1+\epsilon} + AG^{\epsilon-1})||f||^2) \tag{5.25}$$

where

$$||f|| = \sum_{|z| \leqslant A} |f(z)|^2.$$

For $m+ni \in \mathbb{Z}[i]$, we take $f(m+ni) = \kappa(m,n)e(xm+yn)$. Thus

$$\sum_{d\leqslant D} d^2 S_d(x,y) \leqslant 2\alpha D \sum_{d\leqslant G} d^2 S_d(x,y) + O(A^5 B^{-1}\Xi G^{\epsilon-1}) \tag{5.26}$$

where

$$\mathcal{D}_d(\kappa) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(x,y)S_d(x,y)\,dx\,dy.$$

Similar to Ξ, we expect G is a power of $\log N$. To apply (5.25) we need $DG < A^{1-\varepsilon}$, which is valid if $B > A^{\varepsilon}$. By taking $G = \Xi^6$ and substituting (5.17), (5.24) and (5.26) into (5.16), we arrive at

$$\mathcal{D}(\kappa) \ll AB\Xi^4 \sum_{d\leqslant \Xi^6} d^2 \mathcal{D}_d(\kappa) + A^2(B^2 + AB\Xi^{-1})$$

where

$$\begin{aligned}\mathcal{D}_d(\kappa) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(x,y)S_d(x,y)\,dx\,dy \\ &= \sum_{(s,t)\equiv(0,0)\ (\mathrm{mod}\ d)} (\kappa*\kappa)(s,t)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(x,y)e(xs+yt)\,dx\,dy.\end{aligned}$$

Our final obstacle is to develop an estimate of $\mathcal{D}_d(\kappa)$ for small values of d. Here the modulus d is less than a power of $\log N$, which is analogous to the classical Siegel-Walfisz theorem. As in (5.23), after some changes of variables the above integral can be expressed as

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(x,y)e(xw+yz)\,dx\,dy = \frac{2\pi}{A^2}\exp\Big(-\frac{4\pi\sqrt{h(w,z)}}{A\sqrt{|\Delta|}}\Big).$$

This gives

$$\mathcal{D}_d(\kappa) = 2\pi A^{-2} \sum_{\substack{s_1\equiv s_2\ (\mathrm{mod}\ d)\\ t_1\equiv t_2\ (\mathrm{mod}\ d)}}\sum \kappa(s_1,t_1)\overline{\kappa}(s_2,t_2)\exp\Big(-\frac{4\pi\sqrt{h(s_1,t_1;s_2,t_2)}}{A\sqrt{|\Delta|}}\Big)$$

where

$$h(s_1,t_1;s_2,t_2) = h(s_1-s_2,t_1-t_2).$$

Note that

$$\begin{aligned}\mathcal{D}_d(\kappa) \ll \max_{d_1,d_2\ (\mathrm{mod}\ d)} \max_{N<g^*(s_0,t_0)<N'} \Bigg| \sum_{\substack{(s,t)\equiv(d_1,d_2)\ (\mathrm{mod}\ d)\\ N<g^*(s,t)<N'}} \\ \mu(g^*(s,t))\chi(g^*(s,t))\exp\Big(-\frac{4\pi\sqrt{h(s,t;s_0,t_0)}}{A\sqrt{|\Delta|}}\Big)\Bigg|.\end{aligned}$$

Hence it suffices to show that

$$\sum_{\substack{(s,t)\equiv(d_1,d_2) \pmod d \\ N<g^*(s,t)<N'}} \mu(rg^*(s,t))\chi(rg^*(s,t)) \exp\Big(-\frac{4\pi\sqrt{h(s,t;s_0,t_0)}}{A\sqrt{|\Delta|}}\Big) \ll N\eta.$$

Define $\eta = (\log N)^{-j}$. We can divide the region $N < g^*(s,t) < N'$ into non-overlapping sectors of the form

$$R(Z,\xi) = \{(s,t) \in \mathbb{Z}^2 : Z - \sqrt{N}\eta < g^*(s,t) \leqslant Z, \xi < \arg(s+ti) \leqslant \xi + \eta\}$$

and there are at most η^{-2} regions. For a fixed $(S,T) \in R(Z,\xi)$ and any $(s,t) \in R(Z,\xi)$, we always have

$$\exp\Big(-\frac{4\pi\sqrt{h(s,t;s_0,t_0)}}{A\sqrt{|\Delta|}}\Big) = \exp\Big(-\frac{4\pi\sqrt{h(S,T;s_0,t_0)}}{A\sqrt{|\Delta|}}\Big) + O(\eta).$$

Hence it suffices to show that

$$\sum_{\substack{(s,t)\in R(Z,\xi) \\ (s,t)\equiv(d_1,d_2) \pmod d \\ \gcd(g^*(s,t),r)=1}} \mu(g^*(s,t))\chi(g^*(s,t)) \ll N\eta^3.$$

The proof is complete by applying Lemma 5.1.2.

Chapter 6

Proof of the Main Theorem

With all these ingredients we can prove our main theorems and corollaries.

Proof of Proposition 1.1.2. Define $Y = X^{5/7}, Z = X^{1/7}$. Combining Proposition 2.3.1 with Proposition 3.1.1 and Proposition 5.1.1, we have shown that

$$P(X;\chi) = \sum_{\substack{N\leqslant X\\ \gcd(N,P_F)=1}} \sum_{\ell} \lambda(\ell;N)(H_{F,q}(\ell)+\delta_\ell(N;Y,Z))+P(Z;\chi)+O_{A,B,F,Q}(X(\log X)^{-B}).$$

By estimating $\delta_\ell(N;Y,Z)$ using (2.13) and $P(Z;\chi)$ using Lemma 2.1.3, we deduce that

$$\begin{aligned} P(X;\chi) &= \sum_{\substack{\ell\in\mathbb{N},m\in\mathbb{Z}\\ F(\ell,m)\leqslant X\\ \gcd(\ell,\gamma m)=1\\ \gcd(F(\ell,m),P_F)=1}}\sum \lambda(\ell)\chi(F(\ell,m))\Lambda(F(\ell,m)) \\ &= \sum_{\substack{\ell\in\mathbb{N},m\in\mathbb{Z}\\ F(\ell,m)\leqslant X\\ \gcd(\ell,\gamma m)=1\\ \gcd(F(\ell,m),P_F)=1}}\sum \lambda(\ell)\chi(F(\ell,m))H_{F,q}(\ell) + O_{A,B,F,Q}(X(\log X)^{-B}). \end{aligned} \tag{6.1}$$

The condition $\gcd(\ell,\gamma m)=1$ and $\gcd(F(\ell,m),P_F)=1$ on the first line can be removed (with an acceptable error) due to the presence of $\Lambda(F(\ell,m))$. □

Proof of Theorem 1.1.3. Define $\lambda(\ell)=\Lambda(\ell)$ if $\ell\equiv b \pmod q$ and $\lambda(\ell)=0$ otherwise. Then

$$\begin{aligned} H_{F,q}(\ell) &= \prod_{p\nmid \ell q P_F}\left(1-\frac{\rho(p)}{p}\right)\left(1-\frac{1}{p}\right)^{-1}\prod_{p\mid \ell q P_F}\left(1-\frac{1}{p}\right)^{-1} \\ &= \prod_{p}\left(1-\frac{\rho(p)}{p}\right)\left(1-\frac{1}{p}\right)^{-1}\prod_{p\mid \ell q P_F}\left(1-\frac{\rho(p)}{p}\right)^{-1} \\ &= \prod_{p}\left(1-\frac{\rho(p)}{p}\right)\left(1-\frac{1}{p}\right)^{-1}\prod_{p\mid q P_F}\left(1-\frac{\rho(p)}{p}\right)^{-1}\prod_{p\mid \ell, p\nmid q P_F}\left(1-\frac{\rho(p)}{p}\right)^{-1}. \end{aligned}$$

Define $\theta(\ell)$ to be the third product in the last line. Then if ℓ is a power of p and $p \nmid qP_F$, we have

$$\theta(\ell) = \frac{p}{p - \rho(p)} = 1 + O\Big(\frac{1}{p}\Big).$$

Thus

$$\begin{aligned} &\sum_{\substack{\ell\in\mathbb{N},m\in\mathbb{Z}\\ F(\ell,m)\leqslant X}}\sum \lambda(\ell)\chi(F(\ell,m))\Lambda(F(\ell,m)) \\ &= H_{F,q} \sum_{\substack{\ell\in\mathbb{N},m\in\mathbb{Z}\\ F(\ell,m)\leqslant X\\ \gcd(\ell,\gamma m)=1\\ \gcd(F(\ell,m),P_F)=1}}\sum \lambda(\ell)\chi(F(\ell,m)) + O_{A,F,Q}(X(\log X)^{-A}) \end{aligned} \tag{6.2}$$

where

$$H_{F,q} = \prod_{p\nmid qP_F}\Big(1 - \frac{\rho(p)}{p}\Big)\Big(1 - \frac{1}{p}\Big)^{-1} \prod_{p\mid qP_F}\Big(1 - \frac{1}{p}\Big)^{-1}.$$

Next, by using Möbius inversion, the double summation in the right side of (6.2) can be rewritten as

$$\sum_{\substack{\ell\in\mathbb{N},m\in\mathbb{Z}\\ F(\ell,m)\leqslant X\\ \gcd(\ell,\gamma m)=1\\ \gcd(F(\ell,m),P_F)=1}}\sum \lambda(\ell)\chi(F(\ell,m)) \;=\; \sum_{d\mid P_F}\mu(d) \sum_{\substack{N\leqslant X\\ N\equiv 0 \pmod d}} a_N\chi(N).$$

If $\gcd(d,q) = 1$, the summation over N is in fact $\mathcal{A}_d(X;\chi)$ in (3.15). Hence by (3.15) and (3.16),

$$\begin{aligned} \sum_{\substack{\ell\in\mathbb{N},m\in\mathbb{Z}\\ F(\ell,m)\leqslant X\\ \gcd(\ell,\gamma m)=1\\ \gcd(F(\ell,m),P_F)=1}}\sum \lambda(\ell)\chi(F(\ell,m)) \;&=\; \sum_{\substack{d\mid P_F\\ \gcd(d,q)=1}} \mu(d)\mathcal{A}_d(X;\chi) \\ &= \sum_{\substack{d\mid P_F\\ \gcd(d,q)=1}} \mu(d)\mathcal{M}_d(X;\chi) + O(\mathcal{R}(X,P_F;\chi)). \end{aligned}$$

The condition $\gcd(\ell,\gamma md) = 1$ in $\mathcal{M}_d(X;\chi)$ is negligible. Therefore by Proposition 3.4.1,

$$\sum_{\substack{\ell\in\mathbb{N},m\in\mathbb{Z}\\ F(\ell,m)\leqslant X}}\sum \lambda(\ell)\chi(F(\ell,m))\Lambda(F(\ell,m)) = H_q \sum_{\substack{\ell\in\mathbb{N},m\in\mathbb{Z}\\ F(\ell,m)\leqslant X}}\sum \lambda(\ell)\chi(F(\ell,m)) + O_{A,F,Q}(X(\log X)^{-A}).$$

Using the orthogonality of χ and definition of $\lambda(\ell)$, it gives

$$\sum_{\substack{\ell\in\mathbb{N},m\in\mathbb{Z}\\ F(\ell,m)\leqslant X\\ F(\ell,m)\equiv a \pmod q\\ \ell\equiv b \pmod q}}\sum \Lambda(\ell)\Lambda(F(\ell,m)) = H_q \sum_{\substack{\ell\in\mathbb{N},m\in\mathbb{Z}\\ F(\ell,m)\leqslant X\\ F(\ell,m)\equiv a \pmod q\\ \ell\equiv b \pmod q}}\sum \Lambda(\ell) + O_{A,F,Q}(X(\log X)^{-A}).$$

Finally, the result follows from the prime number theorem in arithmetic progression

$$\sum_{\substack{\ell\in\mathbb{N}\\ F(\ell,m)\leqslant X\\ \ell\equiv b \pmod q}} \Lambda(\ell) = \frac{1}{\phi(q)} \sum_{\substack{\ell\in\mathbb{N}\\ F(\ell,m)\leqslant X}} \Lambda(\ell) + O_A(X(\log X)^{-A}).$$

The implied constant is ineffective since we applied Siegel's theorem. □

Proof of Corollary 1.1.4. Simply take $q = a = b = 1$. □

Proof of Corollary 1.1.5. Without loss of generality we assume that $G(\ell, m) = \ell$. Then for every prime p, there are $\ell, m \in \mathbb{Z}$ such that $p \nmid \ell$ and $p \nmid F(\ell, m)$. Therefore $\nu \equiv m\ell^{-1} \pmod p$ would be a solution to

$$F(1, \nu) \not\equiv 0 \pmod p.$$

This implies $\rho(p) < p$ for all p. By Corollary 1.1.4, we have $H > 0$ and the result follows. □

Bibliography

[1] A. Balog, V. Blomer, C. Dartyge and G. Tenenbaum, *Friable values of binary forms*, Comment. Math. Helv. **87** (2012), no. 3, 639-667.

[2] S. Choi and P. C. H. Lam, *Large sieve type inequality for the roots of quadratic congruence*, unpublished.

[3] D. Cox, *Primes of the form* $x^2 + ny^2$, Pure and Applied Mathematics (Hoboken), John Wiley & Sons, Inc., Hoboken, NJ, 2013.

[4] P. Erdos, *On the sum* $\sum_{k=1}^{x} d(f(k))$, J. London Math. Soc. **27**, (1952). 7-15.

[5] E. Fouvry and H. Iwaniec, *Gaussian primes*, Acta Arith. **79** (1997), no. 3, 249- 287.

[6] J. Friedlander and H. Iwaniec, *The polynomial* $X^2 + Y^4$ *captures its primes*, Ann. of Math. (2) **148** (1998), no. 3, 945-1040.

[7] J. Friedlander and H. Iwaniec, *Asymptotic sieve for primes*, Ann. of Math. (2) **148** (1998), no. 3, 1041-1065.

[8] J. Friedlander and H. Iwaniec, *Gaussian sequences in arithmetic progressionss*, Funct. Approx. Comment. Math., **37**, (2007), part 1, 149-157.

[9] L. Grimmelt, *Vinogradov's theorem with Fouvry-Iwaniec primes*, submitted.

[10] H. Helfgott, *Root numbers and the parity problem*, Ph. D. Thesis, 2003.

[11] D. R. Heath-Brown, *Primes represented by* $x^3 + 2y^3$, Acta Math. **186** (2001), 1-84.

[12] D. R. Heath-Brown, X. Li, *Prime values of* $a^2 + p^4$, Invent. Math (2) **208** (2017), 441-499.

[13] C. Hooley, *On the distribution of the roots of polynomial congruences*, Mathematika **11** (1964), 39–49.

[14] H. Iwaniec, *Almost-primes represented by quadratic polynomials*, Invent. Math. **47** (1978), no. 2, 171-188.

[15] P. C. H. Lam, *Primes of the form* $x^2 + Dy^2$, M. Phil. Thesis, The University of Hong Kong (2014).

[16] P. C. H. Lam, *Primes of the form* $\alpha x^2 + \beta xy + \gamma y^2$, submitted.

[17] M. Pandey, *On Eisenstein primes*, Integers **18** (2018), A59

[18] J. Galambos, K.-H. Indlekofer and I. Kátai, *A renewal theorem for random walks in multidimensional time*, Trans. Amer. Math. Soc. **300** (1987), no. 2, 759–769.

[19] R. C. Vaughan, *Mean value theorems in prime number theory*, J. London Math. Soc. **10** (1975), 153-162.