Diophantine equations in positive characteristic

Koymans, P.H.

Diophantine equations in positive characteristic

Proefschrift

ter verkrijging van
de graad Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C. J. J. M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 19 juni 2019
klokke 16.15 uur

door

Peter Hubrecht Koymans

geboren te Eindhoven

in 1992

Promotor	prof. dr. P. Stevenhagen	Leiden University
Copromotor	dr. J.-H. Evertse	Leiden University

Doctorate Committee

Chair	prof. dr. A. W. van der Vaart	Leiden University
Secretary	prof. dr. B. de Smit	Leiden University
Member	dr. A. Bartel	University of Glasgow
Member	prof. dr. E. Fouvry	University of Paris-Sud
Member	prof. dr. J. F. Voloch	University of Canterbury

Contents

Preface

In this preface we shall give a mathematical introduction to the various topics in the thesis. The thesis consists of three parts. The first part is devoted to exponential Diophantine equations in positive characteristic, while the second part revolves around class number statistics. These two parts form the main body of the thesis, whence the title of this thesis. The final and third part is a paper that solves the ternary Goldbach problem for Artin primes.

An exponential Diophantine equation is an equation where some of the variables occur as exponents. Famous examples of such equations are the Fermat equation

$$x^N + y^N = z^N \text{ in integers } N > 2,\ xyz \neq 0,$$

where N occurs as an exponent, and the Catalan equation

$$x^m - y^n = 1 \text{ in integers } x, y, m, n > 1,$$

where m and n occur as exponents. There is a well-known analogy between number fields and global function fields. Therefore, it is natural to solve these equations over global (or even more general) function fields instead of number fields. The advantage of global function fields is that one can use derivations, and this allows us to use elementary methods to establish our results.

Let K be a finitely generated field over $\mathbb{F}_p$ and fix $a, b \in K^*$. In the first chapter we shall study the generalized Catalan equation

$$ax^m + by^n = 1 \text{ in } x, y \in K \text{ and integers } m, n \text{ coprime with } p.$$

This equation was already studied by Silverman [66], but his main theorem is false as we shall demonstrate in the first chapter. We will prove that there are only finitely many solutions up to a natural equivalence relation provided that the pair (m, n) does not belong to an explicit finite list.

In the next chapter we shall study the so-called unit equation. Let K be a field of characteristic 0 and let G be a multiplicative subgroup of $K^* \times K^*$. Then the equation

$$x + y = 1 \text{ in } (x, y) \in G$$

is an exponential Diophantine equation. Siegel and Mahler showed finiteness of the solution set in important special cases, while Lang proved finiteness in general. Mahler and later Evertse [17] gave upper bounds for the solution sets in important special cases, while Beukers and Schlickewei [3] gave an upper bound in full generality. Namely, they showed that there are at most 2^{8r+8} solutions, where r is the rank of G. In characteristic $p > 0$ the situation turns out to be rather different. Indeed, if we have

$$x + y = 1 \text{ for some } (x, y) \in G,$$

we can apply Frobenius to find another solution

$$x^p + y^p = 1.$$

Voloch [78] gave an upper bound for the number of solutions up to a natural equivalence relation. His upper bound depends on both r and p, and he asked if the dependence on p could be removed. Together with Pagano I gave the upper bound $31 \cdot 19^r$, which answers Voloch's question. To do so, we adapt the method of Beukers and Schlickewei to positive characteristic.

The final chapter of the first part studies the Fermat surface

$$x^N + y^N + z^N = 1, \tag{1}$$

where $x, y, z \in \mathbb{F}_p(t)$ and N is a positive integer. The main result is that there are infinitely many primes N for which equation (1) has no solutions satisfying $x, y, z \notin \mathbb{F}_p(t^p)$ and $x/y, x/z, y/z \notin \mathbb{F}_p(t^p)$. We also show that the conditions on x, y and z can not be removed. This chapter is also joint work with Pagano.

The second part of the thesis revolves around the 2-part of the class groups of imaginary quadratic number fields. Cohen and Lenstra [10] put forward conjectures about the average behavior of such class groups. Let p be an odd prime. Their conjecture predicts that for all finite abelian p-groups A

$$\lim_{X \to \infty} \frac{|\{K \text{ imaginary quadratic} : |D_K| < X \text{ and } \mathrm{Cl}(K)[p^\infty] \cong A\}|}{|\{K \text{ imaginary quadratic} : |D_K| < X\}|} = \frac{\prod_{i=1}^{\infty} \left(1 - \frac{1}{p^i}\right)}{|\mathrm{Aut}(A)|},$$

where D_K and $\mathrm{Cl}(K)$ are respectively the discriminant and narrow class group of K. Although Cohen and Lenstra stated their conjecture already in 1984, there are very few proven instances despite significant effort. Davenport and Heilbronn [14] obtained partial results in the case $p = 3$, and the case $p > 3$ is still wide open. Although the conjecture was originally stated only for odd p, Gerth proposed the following modification; instead of $\mathrm{Cl}(K)[2^\infty]$, it is $(2\mathrm{Cl}(K))[2^\infty]$ that behaves randomly. This was recently proven by Smith [69] and can be considered a major breakthrough in the area.

One way to study $\mathrm{Cl}(K)[2^\infty]$ is by the use of governing fields. Let $k \geq 1$ and $d \not\equiv 2 \bmod 4$ be integers. Then Cohn and Lagarias [11] conjectured that there exists a finite normal field extension $M_{d,k}$ over $\mathbb{Q}$ such that

$$\dim_{\mathbb{F}_2} \frac{2^{k-1}\mathrm{Cl}(\mathbb{Q}(\sqrt{dp}))}{2^k\mathrm{Cl}(\mathbb{Q}(\sqrt{dp}))}$$

is determined by the splitting of p in $M_{d,k}$. Such a hypothetical field $M_{d,k}$ is called a governing field. Stevenhagen [70] showed in his thesis that governing field exists for $k \leq 3$ and all values of d. If one is able to give an explicit description of $M_{d,3}$, then one can get density results for $\mathrm{Cl}(\mathbb{Q}(\sqrt{dp}))[8]$ using the Chebotarev density theorem, where p varies over the primes.

It is a natural question to ask what happens for $\mathrm{Cl}(\mathbb{Q}(\sqrt{dp}))[16]$, and we analyze this problem for $d = -4$ and $d = -8$. This leads to the following density theorems, and we devote a chapter to each theorem.

Theorem (joint work with Milovic)**.** Let $h(-2p)$ be the class number of $\mathbb{Q}(\sqrt{-2p})$. Then we have

$$\lim_{X \to \infty} \frac{|\{p \leq X : p \text{ prime},\ p \equiv 1 \bmod 4 \text{ and } 16 \mid h(-2p)\}|}{|\{p \leq X : p \text{ prime}\}|} = \frac{1}{16}.$$

Theorem. Let $h(-p)$ be the class number of $\mathbb{Q}(\sqrt{-p})$. Then we have

$$\lim_{X \to \infty} \frac{|\{p \leq X : p \text{ prime and } 16 \mid h(-p)\}|}{|\{p \leq X : p \text{ prime}\}|} = \frac{1}{16}.$$

The proof of both theorems do not make any appeal to the theory of L-functions. Instead they rely on a method due to Vinogradov. This suggests that there is no governing field. The following theorem, which is proven in the final chapter of the second part, provides even more evidence towards the non-existence of governing fields.

Theorem (joint work with Milovic)**.** Assume a short character sum conjecture. Then the field $M_{-4,4}$ does not exist.

In the final part of this thesis we combine two classical problems in analytic number theory. The first problem is the well-known ternary Goldbach conjecture which states that every odd integer $n > 5$ can be written as the sum of three primes, i.e.

$$n = p_1 + p_2 + p_3$$

for primes p_1, p_2 and p_3. Vinogradov [74] showed that every sufficiently large odd integer admits such a representation, and Helfgott [34] settled the full ternary Goldbach conjecture. Another famous problem in analytic number theory is Artin's conjecture on primitive roots. Let g be an integer that is neither a square nor -1. Then Artin's conjecture states that there are infinitely many primes p such that g is a primitive root modulo p, or in other words g generates the group $(\mathbb{Z}/p\mathbb{Z})^*$. Hooley [35] showed the veracity of Artin's conjecture conditional on GRH.

We are interested in writing n as a sum of three primes, all of which have g as primitive root. The following is a simple corollary of our work that is particularly pleasing to state.

Corollary (joint with Frei and Sofos)**.** Assume GRH. Then there is a constant $C > 0$ such that for all odd integers $n > C$ we have the following equivalence: there are odd primes p_1, p_2, p_3 with 27 as primitive root and $n = p_1+p_2+p_3$ if and only if $n \equiv 3 \bmod 12$.

Chapter 1

The generalized Catalan equation in positive characteristic

Abstract

Let $K = \mathbb{F}_p(z_1, \ldots, z_r)$ be a finitely generated field over $\mathbb{F}_p$ and fix $a, b \in K^*$. We study the solutions of the generalized Catalan equation $ax^m + by^n = 1$ to be solved in $x, y \in K$ and integers $m, n > 1$ coprime with p.

1.1 Introduction

In this article we will bound m and n for the generalized Catalan equation in characteristic $p > 0$. Our main result is as follows.

Theorem 1.1.1. *Let $a, b \in K^*$ be given. Consider the equation*

$$ax^m + by^n = 1 \tag{1.1}$$

in $x, y \in K$ and integers $m, n > 1$ coprime with p satisfying

$$(m, n) \notin \{(2, 2), (2, 3), (3, 2), (2, 4), (4, 2), (3, 3)\}. \tag{1.2}$$

Then there is a finite set $\mathcal{T} \subseteq K^2$ such that for any solution (x, y, m, n) of (1.1), there is a $(\gamma, \delta) \in \mathcal{T}$ and $t \in \mathbb{Z}_{\geq 0}$ such that

$$ax^m = \gamma^{p^t}, by^n = \delta^{p^t}. \tag{1.3}$$

In the case $a = b = 1$, a stronger and effective result was proven in [42] based on the work of [6].

Let us now show that the conditions on m and n are necessary. If (1.2) fails, then (1.1) defines a curve of genus 0 or 1 over K. It is clear that (1.3) can fail in this case. It is also essential that m and n are coprime with p. Take for example $a = b = 1$. Then any solution of

$$x + y = 1$$

with $x, y \in K$ and $x, y \notin \overline{\mathbb{F}_p}$ gives infinitely many solutions of the form (1.3) after applying Frobenius.

The generalized Catalan equation over function fields was already analyzed in [66], where the main theorem claims that the generalized Catalan equation has no solutions for m and n sufficiently large. Unfortunately, it is not hard to produce counterexamples to the main theorem given there. Following the notation in [66], we choose $k = \mathbb{F}_p$, $K = k(u)$, $a = x = u$, $b = y = 1 - u$ and $m = n = p^t - 1$ for $t \in \mathbb{Z}_{\geq 0}$. Then we have

$$ax^m + by^n = u \cdot u^{p^t-1} + (1-u) \cdot (1-u)^{p^t-1} = 1$$

due to Frobenius, illustrating the need of (1.3).

1.2 Heights

Let K be a finitely generated extension of $\mathbb{F}_p$. The algebraic closure of $\mathbb{F}_p$ in K is a finite extension of $\mathbb{F}_p$, say $\mathbb{F}_q$ with $q = p^n$ for some $n \in \mathbb{Z}_{>0}$. There exists a projective variety V non-singular in codimension one defined over $\mathbb{F}_q$ with function field K.

Our goal will be to introduce a height function on K by using our variety V. For later purposes it will be useful to do this in a slightly more general setting. So let X be a projective variety, non-singular in codimension one, defined over a perfect field k. We write L for the function field of X and we assume that k is algebraically closed in L.

Fix a projective embedding of X such that $X \subseteq \mathbb{P}_k^M$ for some positive integer M. Then a prime divisor $\mathfrak{p}$ of X over k is by definition an irreducible subvariety of codimension one. Recall that for a prime divisor $\mathfrak{p}$ the local ring $\mathcal{O}_\mathfrak{p}$ is a discrete valuation ring, since X is non-singular in codimension one. Following [47] we will define heights on X. To do this, we start by defining a set of normalized discrete valuations

$$M_L := \{\mathrm{ord}_\mathfrak{p} : \mathfrak{p} \text{ prime divisor of } X\},$$

where $\mathrm{ord}_\mathfrak{p}$ is the normalized discrete valuation of L corresponding to $\mathcal{O}_\mathfrak{p}$. If $v = \mathrm{ord}_\mathfrak{p}$ is in M_L, we define for convenience $\deg v := \deg \mathfrak{p}$ with $\deg \mathfrak{p}$ being the projective degree in $\mathbb{P}_k^M$. Then the set M_L satisfies the sum formula for all $x \in L^*$

$$\sum_v v(x) \deg v = 0.$$

If P is a point in $\mathbb{P}^r(L)$ with coordinates $(y_0 : \ldots : y_r)$ in L, then its (logarithmic) height is

$$h_L(P) = -\sum_v \min_i\{v(y_i)\} \deg v.$$

Furthermore we define for an element $x \in L$

$$h_L(x) = h_L(1 : x). \tag{1.4}$$

We will need the following properties of the height.

Lemma 1.2.1. *Let $x, y \in L$ and $n \in \mathbb{Z}$. The height defined by (1.4) has the following properties:*

(a) $h_L(x) = 0 \Leftrightarrow x \in k$;

(b) $h_L(x + y) \leq h_L(x) + h_L(y)$;

(c) $h_L(xy) \leq h_L(x) + h_L(y)$;

(d) $h_L(x^n) = |n| h_L(x)$;

(e) Suppose that k is a finite field and let $C > 0$ be given. Then there are only finitely many $x \in L^*$ satisfying $h_L(x) \leq C$;

(f) $h_L(x) = h_{\overline{k} \cdot L}(x)$.

Proof. Property (a) is Proposition 4 of [46] (p. 157), while properties (b), (c) and (d) are easily verified. Property (e) is proven in [55]. Finally, property (f) can be found after Proposition 3.2 in [47] (p. 63). □

1.3 A generalization of Mason's ABC-theorem

For our proof we will need a generalization of Mason's ABC-theorem for function fields in one variable to an arbitrary number of variables. Such a result is given in [36]. For completeness we repeat it here.

Theorem 1.3.1. *Let X be a projective variety over an algebraically closed field k of characteristic $p > 0$, which is non-singular in codimension one. Let $L = k(X)$ be its function field and let M_L be as above. Let $L_1, \ldots, L_q$, $q \geq n+1$, be linear forms in $n+1$ variables over k which are in general position. Let $\mathbf{X} = (x_0 : \ldots : x_n) \in \mathbb{P}^n(L)$ be such that $x_0, \ldots, x_n$ are linearly independent over K^{p^m} for some $m \in \mathbb{N}$. Then, for any fixed* *finite subset S of M_L, the following inequality holds:*

$$
\begin{aligned}
&(q-n-1)h(x_0:\ldots:x_n)\\
&\leq \sum_{i=1}^{q}\sum_{v\notin S}\deg v\min\{np^{m-1}, v(L_i(\mathbf{X}))-\min_{0\leq j\leq n}\{v(x_j)\}\}\\
&+\frac{n(n+1)}{2}p^{m-1}\left(C_X+\sum_{v\in S}\deg v\right),
\end{aligned}
$$

where C_X is a constant depending only on X.

Proof. This is the main theorem in [36]. □

1.4 Proof of Theorem 1.1.1

In this section we proof our main theorem.

Proof of Theorem 1.1.1. Let (x,y,m,n) be an arbitrary solution. Let us first dispose with the case $ax^m \in \mathbb{F}_q$. Then also $by^n \in \mathbb{F}_q$, so we simply add $\mathbb{F}_q \times \mathbb{F}_q$ to $\mathcal{T}$. From now on we will assume $ax^m \notin \mathbb{F}_q$ and hence $by^n \notin \mathbb{F}_q$. It follows that

$$h_K(ax^m), h_K(by^n) \neq 0,$$

so we may write

$$ax^m = \gamma^{p^t}, by^n = \delta^{p^s}$$

for some $t, s \in \mathbb{Z}_{\geq 0}$ and $\gamma, \delta \notin K^p$. After substitution we get

$$\gamma^{p^t} + \delta^{p^s} = 1.$$

Extracting p-th roots gives $t = s$ and hence

$$\gamma + \delta = 1. \tag{1.5}$$

Our goal will be to apply the main theorem of [36] to (1.5). Note that Theorem 1.3.1 requires that the ground field k is algebraically closed. But a constant field extension does not change the height by Lemma 1.2.1(f). Hence we can keep working with our field K instead of $\overline{\mathbb{F}_p} \cdot K$. Define the following three linear forms in two variables X, Y

$$
\begin{aligned}
L_1 &= X\\
L_2 &= Y\\
L_3 &= X+Y.
\end{aligned}
$$

We apply Theorem 1.3.1 with our V, the above L_1, L_2, L_3 and $\mathbf{X} = (\gamma : \delta) \in \mathbb{P}^1(K)$. We claim that γ and δ are linearly independent over K^p. Suppose that there are $e, f \in K^p$ such that

$$e\gamma + f\delta = 0.$$

Together with $\gamma + \delta = 1$ we find that

$$0 = e\gamma + f\delta = e(1-\delta) + f\delta = e + (f-e)\delta.$$

If $e \neq f$, then this would imply that $\delta \in K^p$, contrary to our assumptions. Hence $e = f$, but then we find

$$0 = e\gamma + f\delta = e$$

and we conclude that $e = f = 0$ as desired.

We still have to choose the subset S of M_K to which we apply Theorem 1.3.1. First we need to make some preparations. From now on v will be used to denote an element of M_K. Define

$$\begin{aligned}
N_0 &:= \{v : v(a) \neq 0 \vee v(b) \neq 0\} \\
N_1 &:= \{v : v(a) = 0, v(b) = 0, v(\gamma) > 0\} \\
N_2 &:= \{v : v(a) = 0, v(b) = 0, v(\delta) > 0\} \\
N_3 &:= \{v : v(a) = v(b) = 0, v(\gamma) = v(\delta) < 0\}.
\end{aligned}$$

It is clear that N_0, N_1, N_2 and N_3 are finite disjoint sets. Before we proceed, we make a simple but important observation in the form of a lemma.

Lemma 1.4.1. *Let (γ, δ) be a solution of (1.5). If $v(\gamma) < 0$ or $v(\delta) < 0$, then*

$$v(\gamma) = v(\delta) < 0.$$

Proof. Obvious. □

Recall that

$$h_K(\gamma) = \sum_v \max(0, v(\gamma)) \deg v = \sum_v -\min(0, v(\gamma)) \deg v$$

and

$$h_K(\delta) = \sum_v \max(0, v(\delta)) \deg v = \sum_v -\min(0, v(\delta)) \deg v.$$

Lemma 1.4.1 tells us that

$$\sum_v -\min(0, v(\gamma)) \deg v = \sum_v -\min(0, v(\delta)) \deg v,$$

hence

$$\begin{aligned}
h_K(\gamma) = h_K(\delta) &= \sum_v \max(0, v(\gamma)) \deg v = \sum_v -\min(0, v(\gamma)) \deg v \qquad (1.6)\\
&= \sum_v \max(0, v(\delta)) \deg v = \sum_v -\min(0, v(\delta)) \deg v. \qquad (1.7)
\end{aligned}$$

We will use these different expressions for the height throughout. Let us now derive elegant upper bounds for N_1, N_2 and N_3. Again we will phrase it as a lemma.

Lemma 1.4.2. *Let (γ, δ) be a solution of (1.5). Then*

$$h_K(\gamma) = h_K(\delta) \geq m \sum_{v \in N_1} \deg v,$$
$$h_K(\gamma) = h_K(\delta) \geq n \sum_{v \in N_2} \deg v,$$
$$h_K(\gamma) = h_K(\delta) \geq \mathit{lcm}(m, n) \sum_{v \in N_3} \deg v.$$

Proof. We know that

$$h_K(\gamma) = h_K(\delta) = \sum_v \max(0, v(\gamma)) \deg v \geq \sum_{v \in N_1} \max(0, v(\gamma)) \deg v.$$

Now let $v \in N_1$. This means that $v(a) = v(b) = 0$ and $v(\gamma) > 0$. Then $ax^m = \gamma^{p^t}$ implies

$$v(a) + mv(x) = p^t v(\gamma)$$

and hence $mv(x) = p^t v(\gamma)$. But m and p are coprime by assumption, so we obtain $m \mid v(\gamma)$. Because $v(\gamma) > 0$, this gives $v(\gamma) \geq m$ and we conclude that

$$h_K(\gamma) = h_K(\delta) \geq m \sum_{v \in N_1} \deg v.$$

Using

$$h_K(\gamma) = h_K(\delta) = \sum_v \max(0, v(\delta)) \deg v \geq \sum_{v \in N_2} \max(0, v(\delta)) \deg v,$$

we find in a similar way that

$$h_K(\gamma) = h_K(\delta) \geq n \sum_{v \in N_2} \deg v.$$

It remains to be proven that

$$h_K(\gamma) = h_K(\delta) \geq \mathrm{lcm}(m, n) \sum_{v \in N_3} \deg v.$$

Now we use

$$\begin{aligned} h_K(\gamma) = h_K(\delta) &= \sum_v - \min(0, v(\gamma)) \deg v = \sum_v - \min(0, v(\delta)) \deg v \\ &\geq \sum_{v \in N_3} - \min(0, v(\gamma)) \deg v = \sum_{v \in N_3} - \min(0, v(\delta)) \deg v. \end{aligned}$$

Now take $v \in N_3$. Then $v(\gamma) = v(\delta) < 0$. In the same way as before, we can show that $m \mid v(\gamma)$ and $n \mid v(\delta)$. But $v(\gamma) = v(\delta) < 0$ by Lemma 1.4.1, so we find that

$$h_K(\gamma) = h_K(\delta) \geq \mathrm{lcm}(m, n) \sum_{v \in N_3} \deg v$$

as desired. □

Define

$$S := N_0 \cup N_1 \cup N_2 \cup N_3.$$

Suppose that $v \notin S$. We claim that

$$v(\gamma) = v(\delta) = 0.$$

But $v \notin S$ implies $v \notin N_0$, so certainly $v(a) = v(b) = 0$. Furthermore, we have that $v \notin N_1$ and $v \notin N_2$, which means that $v(\gamma) \leq 0$ and $v(\delta) \leq 0$. If $v(\gamma) < 0$ or $v(\delta) < 0$, then Lemma 1.4.1 gives $v \in N_3$, contradicting our assumption $v \notin S$. Hence $v(\gamma) = v(\delta) = 0$ as desired.

From our claim it follows that we have for $v \notin S$ and $i = 1, 2, 3$

$$v(L_i(\gamma, \delta)) = \min(v(\gamma), v(\delta)).$$

Theorem 1.3.1 tells us that

$$h_K(\gamma : \delta) \leq C_W + \sum_{v \in S} \deg v,$$

where C_W is a constant depending on W only. By Lemma 1.4.2 we find that

$$\begin{aligned}\sum_{v \in S} \deg v &= \sum_{v \in N_0} \deg v + \sum_{v \in N_1} \deg v + \sum_{v \in N_2} \deg v + \sum_{v \in N_3} \deg v \\ &\leq C_{a,b} + \left(\frac{1}{m} + \frac{1}{n} + \frac{1}{\operatorname{lcm}(m,n)}\right) h_K(\gamma),\end{aligned}$$

where $C_{a,b}$ is a constant depending on a and b only. Now (1.2) implies

$$\frac{1}{m} + \frac{1}{n} + \frac{1}{\operatorname{lcm}(m,n)} < 0.9,$$

hence

$$h_K(\gamma : \delta) \leq 10(C_W + C_{a,b}).$$

But $\gamma + \delta = 1$ gives

$$h_K(\gamma) = h_K(\delta) = h_K(\gamma : \delta).$$

The theorem now follows from Lemma 1.2.1(e). □

1.5 Discussion of Theorem 1.1.1

The conclusion of Theorem 1 tells us that there is a finite set $\mathcal{T} \subseteq K^2$ such that for any solution (x, y, m, n) of (1.1), there is a $(\gamma, \delta) \in \mathcal{T}$ and $t \in \mathbb{Z}_{\geq 0}$ such that

$$ax^m = \gamma^{p^t}, by^n = \delta^{p^t}.$$

Since $\mathcal{T}$ is finite, we may assume that γ and δ are fixed in the above two equations. It would be interesting to further study this equation.

Chapter 2

On the equation $x_1 + x_2 = 1$ in finitely generated multiplicative groups in positive characteristic[1]

Joint work with Carlo Pagano

Abstract

Let K be a field of characteristic $p > 0$ and let G be a subgroup of $K^* \times K^*$ with $\dim_{\mathbb{Q}}(G \otimes_{\mathbb{Z}} \mathbb{Q}) = r$ finite. Then Voloch proved that the equation $ax + by = 1$ in $(x, y) \in G$ for given $a, b \in K^*$ has at most $p^r(p^r + p - 2)/(p - 1)$ solutions $(x, y) \in G$, unless $(a, b)^n \in G$ for some $n \geq 1$. Voloch also conjectured that this upper bound can be replaced by one depending only on r. Our main theorem answers this conjecture positively. We prove that there are at most $31 \cdot 19^{r+1}$ solutions (x, y) unless $(a, b)^n \in G$ for some $n \geq 1$ with $(n, p) = 1$. During the proof of our main theorem we generalize the work of Beukers and Schlickewei to positive characteristic, which heavily relies on diophantine approximation methods. This is a surprising feat on its own, since usually these methods can not be transferred to positive characteristic.

2.1 Introduction

Let G be a subgroup of $\mathbb{C}^* \times \mathbb{C}^*$ with coordinatewise multiplication. Assume that the rank $\dim_{\mathbb{Q}} G \otimes_{\mathbb{Z}} \mathbb{Q} = r$ is finite. Beukers and Schlickewei [3] proved that the equation

$$x_1 + x_2 = 1$$

[1]A slightly modified version of this chapter appeared in the Quarterly Journal of Mathematics, volume 68, issue 3, pages 923-934.

in $(x_1, x_2) \in G$ has at most 2^{8r+8} solutions. A key feature of their upper bound is that it depends only on r.

In this chapter we will analyze the characteristic p case. To be more precise, let $p > 0$ be a prime number and let K be a field of characteristic p. Let G be a subgroup of $K^* \times K^*$ with $\dim_{\mathbb{Q}} G \otimes_{\mathbb{Z}} \mathbb{Q} = r$ finite. Then Voloch proved in [78] that an equation

$$ax_1 + bx_2 = 1 \text{ in } (x_1, x_2) \in G$$

for given $a, b \in K^*$ has at most $p^r(p^r + p - 2)/(p - 1)$ solutions $(x_1, x_2) \in G$, unless $(a, b)^n \in G$ for some $n \geq 1$.

Voloch also conjectured that this upper bound can be replaced by one depending only on r. Our main theorem answers this conjecture positively.

Theorem 2.1.1. *Let K, G, r, a and b be as above. Suppose that there is no positive integer n with $gcd(n, p) = 1$ such that $(a, b)^n \in G$. Then the equation*

$$ax_1 + bx_2 = 1 \textit{ in } (x_1, x_2) \in G \tag{2.1}$$

has at most $31 \cdot 19^{r+1}$ solutions.

Our main theorem will be a consequence of the following theorem.

Theorem 2.1.2. *Let K be a field of characteristic $p > 0$ and let G be a finitely generated subgroup of $K^* \times K^*$ of rank r. Then the equation*

$$x_1 + x_2 = 1 \textit{ in } (x_1, x_2) \in G \tag{2.2}$$

has at most $31 \cdot 19^r$ solutions (x_1, x_2) satisfying $(x_1, x_2) \notin G^p$.

Clearly, the last condition is necessary to guarantee finiteness. Indeed if we have any solution to $x_1 + x_2 = 1$, then we get infinitely many solutions $x_1^{p^k} + x_2^{p^k} = 1$ for $k \in \mathbb{Z}_{\geq 0}$ due to the Frobenius operator.

The set-up of the chapter is as follows. We start by introducing the basic theory about valuations that is needed for our proofs. Then we derive Theorem 2.1.2 by generalizing the proof of Beukers and Schlickewei [3] to positive characteristic. We remark that their proof heavily relies on techniques from diophantine approximation. Most of the methods from diophantine approximation can not be transferred to positive characteristic, so that this is possible with the method of Beukers and Schlickewei is a surprising feat on its own. It was more convenient for us to follow [18], which is directly based on the proof of Beukers and Schlickewei. In the final section we shall prove that Theorem 2.1.1 is a simple consequence of Theorem 2.1.2.

2.2 Valuations and heights

Our goal in this section is to recall the basic theory about valuations and heights without proofs. To prove Theorem 2.1.2 we may assume without loss of generality that

$K = \mathbb{F}_p(G)$. Thus, K is finitely generated over $\mathbb{F}_p$. Note that Theorem 2.1.2 is trivial if K is algebraic over $\mathbb{F}_p$, so from now on we further assume that K has positive transcendence degree over $\mathbb{F}_p$. The algebraic closure of $\mathbb{F}_p$ in K is a finite field, which we denote by $\mathbb{F}_q$. Then there is an absolutely irreducible, normal projective variety V defined over $\mathbb{F}_q$ such that its function field $\mathbb{F}_q(V)$ is isomorphic to K.

Fix a projective embedding of V such that $V \subseteq \mathbb{P}^M_{\mathbb{F}_q}$ for some positive integer M. A prime divisor $\mathfrak{p}$ of V over $\mathbb{F}_q$ is by definition an irreducible subvariety of V of codimension one. Recall that for a prime divisor $\mathfrak{p}$ the local ring $\mathcal{O}_\mathfrak{p}$ is a discrete valuation ring, since V is non-singular in codimension one. Following [47] we will define heights on V. To do this, we start by defining a set of normalized discrete valuations

$$M_K := \{\mathrm{ord}_\mathfrak{p} : \mathfrak{p} \text{ prime divisor of } V\},$$

where $\mathrm{ord}_\mathfrak{p}$ is the normalized discrete valuation of K corresponding to $\mathcal{O}_\mathfrak{p}$. If $v = \mathrm{ord}_\mathfrak{p}$ is in M_K, we set $\deg v := \deg \mathfrak{p}$ with $\deg \mathfrak{p}$ being the projective degree in $\mathbb{P}^M_{\mathbb{F}_q}$. Then the set M_K satisfies the sum formula

$$\sum_{v \in M_K} v(x) \deg v = 0$$

for $x \in K^*$. This is indeed a well-defined sum, since for $x \in K^*$ there are only finitely many valuations v satisfying $v(x) \neq 0$. Furthermore, we have $v(x) = 0$ for all $v \in M_K$ if and only if $x \in \mathbb{F}_q^*$. If P is a point in $\mathbb{A}^{n+1}(K) \setminus \{0\}$ with coordinates $(y_0, \ldots, y_n)$ in K, then its homogeneous height is

$$H_K^{\mathrm{hom}}(P) = -\sum_{v \in M_K} \min_i \{v(y_i)\} \deg v$$

and its height

$$H_K(P) = H_K^{\mathrm{hom}}(1, y_0, \ldots, y_n).$$

We will need the following properties of the height.

Lemma 2.2.1. *Let $P \in \mathbb{A}^{n+1}(K) \setminus \{0\}$. The height defined above has the following properties:*
1) $H_K^{hom}(\lambda P) = H_K^{hom}(P)$ for $\lambda \in K^*$.
2) $H_K^{hom}(P) \geq 0$ with equality if and only if $P \in \mathbb{P}^n(\mathbb{F}_q)$.

2.3 Proof of Theorem 2.1.2

This section is devoted to the proof of Theorem 2.1.2. We will follow the proof in [18], see Section 6.4, with some crucial modifications to take care of the presence of the Frobenius map. The general strategy of the proof in characteristic 0, and how we adapt it to characteristic p, will be explained after Lemma 2.3.9. Let us start with a simple lemma.

Lemma 2.3.1. *The equation*

$$x_1 + x_2 = 1 \text{ in } (x_1, x_2) \in G \tag{2.3}$$

has at most p^r solutions (x_1, x_2) satisfying $x_1 \notin K^p$ and $x_2 \notin K^p$.

Proof. Let $x = (x_1, x_2)$ and $y = (y_1, y_2)$ be two solutions of (2.3). We claim that $x \equiv y \mod G^p$ implies $x = y$. Indeed, if $x \equiv y \mod G^p$, we can write $y_1 = x_1\gamma^p$ and $y_2 = x_2\delta^p$ with $(\gamma, \delta) \in G$. In matrix form this means that

$$\begin{pmatrix} 1 & 1 \\ \gamma^p & \delta^p \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

For convenience we define

$$A := \begin{pmatrix} 1 & 1 \\ \gamma^p & \delta^p \end{pmatrix}.$$

If A is invertible, we find that $x_1, x_2 \in K^p$ contrary to our assumptions. So A is not invertible, which implies that $\gamma = \delta = 1$. This proves the claim.

The claim implies that the number of solutions is at most $|G/G^p|$. Let $\mathbb{F}_q$ be the algebraic closure of $\mathbb{F}_p$ in K. It is a finite extension of $\mathbb{F}_p$, since K is finitely generated over $\mathbb{F}_p$. It follows that $G^{\text{tors}} \subseteq \mathbb{F}_q^* \times \mathbb{F}_q^*$. Hence $|G^{\text{tors}}| \mid (q-1)^2$, which is co-prime to p. We conclude that $|G/G^p| = p^r$ as desired. □

Lemma 2.3.1 gives the following corollary.

Corollary 2.3.2. *The equation*

$$x_1 + x_2 = 1 \text{ in } (x_1, x_2) \in G \tag{2.4}$$

has at most p^r solutions (x_1, x_2) satisfying $(x_1, x_2) \notin G^p$.

Proof. Define

$$G' := \{(x_1, x_2) \in K \times K : (x_1^N, x_2^N) \in G \text{ for some } N \in \mathbb{Z}_{>0}\}.$$

It is a well known fact that G' is finitely generated if G and K are. It follows that G' is a finitely generated group of rank r. Our goal is to give an injective map from the solutions $(x_1, x_2) \in G$ of (2.4) satisfying $(x_1, x_2) \notin G^p$ to the solutions $(x_1', x_2') \in G'$ of (2.3) satisfying $(x_1', x_2') \notin K^p$ and then apply Lemma 2.3.1.

So let $(x_1, x_2) \in G$ be a solution of (2.4) satisfying $(x_1, x_2) \notin G^p$. We start by remarking that $x_1, x_2 \notin \mathbb{F}_q$. Hence we can repeatedly take p-th roots until we get $x_1', x_2' \notin K^p$. Using heights one can prove that this indeed stops after finitely many steps. Then it is easily verified that $(x_1', x_2') \in G'$ is a solution of (2.3) and that the map thus defined is injective. Now apply Lemma 2.3.1. □

By Corollary 2.3.2 we may assume that p is sufficiently large throughout, say $p > 7$. Both the proof in [18] and our proof rely on very special properties of the family of binary forms $\{W_N(X,Y)\}_{N\in\mathbb{Z}_{>0}}$ defined by the formula

$$W_N(X,Y) = \sum_{m=0}^{N} \binom{2N-m}{N-m}\binom{N+m}{m} X^{N-m}(-Y)^m.$$

We have for all positive integers N that $W_N(X,Y) \in \mathbb{Z}[X,Y]$. Furthermore, setting $Z = -X-Y$, the following statements hold in $\mathbb{Z}[X,Y]$.

Lemma 2.3.3. *1)* $W_N(Y,X) = (-1)^N W_N(X,Y)$.
2) $X^{2N+1}W_N(Y,Z) + Y^{2N+1}W_N(Z,X) + Z^{2N+1}W_N(X,Y) = 0$.
3) There exist a non-zero integer c_N such that

$$det\begin{pmatrix} Z^{2N+1}W_N(X,Y) & Y^{2N+1}W_N(Z,X) \\ Z^{2N+3}W_{N+1}(X,Y) & Y^{2N+3}W_{N+1}(Z,X)\end{pmatrix} = c_N(XYZ)^{2N+1}(X^2+XY+Y^2).$$

Proof. This is Lemma 6.4.2 in [18], which is a variant of Lemma 2.3 in [3]. □

Since the formulas in the previous lemma hold in $\mathbb{Z}[X,Y]$ they hold in every field K. But if $\mathrm{char}(K) = p > 0$ and $p \mid c_N$, then part 3) of Lemma 2.3.3 tells us that

$$det\begin{pmatrix} Z^{2N+1}W_N(X,Y) & Y^{2N+1}W_N(Z,X) \\ Z^{2N+3}W_{N+1}(X,Y) & Y^{2N+3}W_{N+1}(Z,X)\end{pmatrix} = 0$$

in $K[X,Y]$. The following remarkable identity will be handy later on, when we need that c_N does not vanish modulo p.

Lemma 2.3.4. *For every positive integer N, one has* $W_N(2,-1) = 4^N\binom{\frac{3}{2}N}{N}$.

Proof. It is enough to evaluate $\sum_{i=0}^{N}\binom{2N-i}{N}\binom{N+i}{N}2^{-i}$. We have

$$\sum_{i=0}^{N}\binom{2N-i}{N}\binom{N+i}{N}2^{-i} = \binom{2N}{N}F\left(-N, N+1, -2N, \frac{1}{2}\right),$$

where $F(a,b,c,z)$ is the hypergeometric function defined by the power series

$$F(a,b,c,z) := \sum_{i=0}^{\infty}\frac{(a)_i(b)_i}{i!(c)_i}z^n.$$

Here we define for a real t and a non-negative integer i $(t)_i = 1$ if $i = 0$ and for i positive $(t)_i = t(t+1)\cdot\ldots\cdot(t+i-1)$. Now the desired result follows from Bailey's formulas where special values of the function F are expressed in terms of values of the Γ-function, see [48] page 297. □

We obtain the following corollary.

Corollary 2.3.5. *Let p be an odd prime number and let N be a positive integer with $N < \frac{p}{3} - 2$. Then $c_N \not\equiv 0$ mod p.*

Proof. Indeed one has that

$$det\begin{pmatrix} Z^{2N+1}W_N(X,Y) & Y^{2N+1}W_N(Z,X) \\ Z^{2N+3}W_{N+1}(X,Y) & Y^{2N+3}W_{N+1}(Z,X) \end{pmatrix}$$

evaluated at $(X,Y,Z) = (2,-1,-1)$ gives up to sign $2W_N(2,-1)W_{N+1}(2,-1)$. By the previous proposition, this is a power of 2 times the product of two binomial coefficients whose top terms are less than p, hence it can not be divisible by p. $\square$

We now state and prove the analogues of Lemmata 6.4.3-6.4.5 from [18] for function fields of positive characteristic. These are variants of respectively Lemma 2.1, Corollary 2.2 and Lemma 2.3 from [3].

Lemma 2.3.6. *Let a, b, c be non-zero elements of K, and let $(\alpha_i, \beta_i, \gamma_i)$ for $i = 1, 2$ be two K-linearly independent vectors from K^3 such that $a\alpha_i + b\beta_i + c\gamma_i = 0$ for $i = 1, 2$. Then*

$$H_K^{hom}(a,b,c) \le H_K^{hom}(\alpha_1,\beta_1,\gamma_1) + H_K^{hom}(\alpha_2,\beta_2,\gamma_2).$$

Proof. The vector (a, b, c) is K-proportional to the vector with coordinates given by $(\beta_1\gamma_2 - \gamma_1\beta_2, \gamma_1\alpha_2 - \alpha_1\gamma_2, \alpha_1\beta_2 - \beta_1\alpha_2)$. So we have

$$\begin{aligned} H_K^{\text{hom}}(a,b,c) &= H_K^{\text{hom}}(\beta_1\gamma_2 - \gamma_1\beta_2, \gamma_1\alpha_2 - \alpha_1\gamma_2, \alpha_1\beta_2 - \beta_1\alpha_2) \\ &= \sum_{v \in M_K} -\min(v(\beta_1\gamma_2 - \gamma_1\beta_2), v(\gamma_1\alpha_2 - \alpha_1\gamma_2), v(\alpha_1\beta_2 - \beta_1\alpha_2)) \deg v \\ &\le \sum_{v \in M_K} (-\min(v(\beta_1), v(\gamma_1), v(\alpha_1)) - \min(v(\gamma_2), v(\alpha_2), v(\beta_2))) \deg v \\ &= H_K^{\text{hom}}(\alpha_1,\beta_1,\gamma_1) + H_K^{\text{hom}}(\alpha_2,\beta_2,\gamma_2), \end{aligned}$$

which was the claimed inequality. $\square$

We apply Lemma 2.3.6 to the equation $x_1 + x_2 = 1$.

Lemma 2.3.7. *Suppose $x = (x_1, x_2) \in G$ and $y = (y_1, y_2) \in G$ satisfy $x_1 + x_2 = 1$ and $y_1 + y_2 = 1$. Then we have $H_K(x) \le H_K(yx^{-1})$.*

Proof. Apply Lemma 2.3.6 with $(a, b, c) = (x_1, x_2, -1)$, $(\alpha_1, \beta_1, \gamma_1) = (1, 1, 1)$ and $(\alpha_2, \beta_2, \gamma_2) = (y_1x_1^{-1}, y_2x_2^{-1}, 1)$. Finally use the fact that $H_K^{\text{hom}}(1, 1, 1) = 0$. $\square$

The next Lemma takes advantage of the properties of $W_N(X,Y)$ listed in Lemma 2.3.3 and the non-vanishing of c_N modulo p obtained in Corollary 2.3.5.

Lemma 2.3.8. *Let x, y be as in Lemma 2.3.7. Let $N < \frac{p}{3} - 2$. Then there exists $M \in \{N, N+1\}$ such that $H_K(x) \le \frac{1}{M+1} H_K(yx^{-2M-1})$.*

Proof. The proof is almost the same as in Lemma 6.4.5 in [18], with only few necessary modifications. For completeness we give the full proof.

If x_1, and thus both x_1 and x_2 are roots of unity, we have that $H_K(x) = 0$ so the lemma is trivially true. By Lemma 2.3.3 part 2) we get that

$$x_1^{2M+1} W_M(x_2, -1) + x_2^{2M+1} W_M(-1, x_1) - W_M(x_1, x_2) = 0$$

for $M \in \{N, N+1\}$ as well as

$$x_1^{2M+1}(y_1 x_1^{-2M-1}) + x_2^{2M+1}(y_2 x_2^{-2M-1}) - 1 = 0.$$

Now we claim that there is $M \in \{N, N+1\}$ such that the vectors

$$(y_1, y_2, -1) \text{ and } ({x_1}^{2M+1} W_M(x_2, -1), {x_2}^{2M+1} W_M(-1, x_1), -W_M(x_1, x_2)) \qquad (2.5)$$

are linearly independent. Clearly, to prove the claim it is enough to prove that the two vectors

$$(x_1^{2M+1} W_M(x_2, -1), x_2^{2M+1} W_M(-1, x_1), -W_M(x_1, x_2)) \quad (M \in \{N, N+1\}) \qquad (2.6)$$

are linearly independent. But we know that for $M \in \{N, N+1\}$ we have $c_M \not\equiv 0 \bmod p$ by Corollary 2.3.5 and the assumption that $N < \frac{p}{3} - 2$. Furthermore, x_1 and x_2 are not algebraic over $\mathbb{F}_p$. Thus the identity Lemma 2.3.3 part 3) gives us the non-vanishing of the first 2×2 minor of the vectors in 2.6, which proves the claimed independence. So by applying to (2.5) the diagonal transformation that divides the first coordinate by x_1^{2M+1} and the second by x_2^{2M+1}, we deduce that the two vectors

$$(y_1 x_1^{-2M-1}, y_2 x_2^{-2M-1}, -1)$$

and

$$(W_M(x_2, -1), W_M(-1, x_1), -W_M(x_1, x_2)) =: (w_1, w_2, w_3)$$

are linearly independent. So by Lemma 2.3.6 we get that

$$(2M+1) H_K(x) \le H_K(y x^{-2M-1}) + H_K^{\text{hom}}(w_1, w_2, w_3)$$

But now the inequality

$$H_K^{\text{hom}}(w_1, w_2, w_3) \le M \cdot H_K(x)$$

follows immediately from the non-archimedean triangle inequality. So we indeed get

$$(M+1) H_K(x) \le H_K(y x^{-2M-1}),$$

completing the proof. □

Define

$$\text{Sol}(G) := \{(x_1, x_2) \in G \setminus G^{\text{tors}} : x_1 + x_2 = 1\}$$

and

$$\text{Prim-Sol}(G) := \{(x_1, x_2) \in G \setminus G^p : x_1 + x_2 = 1\}.$$

It is easily seen that Prim-Sol$(G) \subseteq \mathrm{Sol}(G)$. Finally define

$$S := \{v \in M_K : \text{ there is } (x_1, x_2) \in G \text{ with } v(x_1) \neq 0 \text{ or } v(x_2) \neq 0\}.$$

The set S is clearly finite. Write $s := |S|$, $S = \{v_1, \ldots, v_s\}$. Then we have a homomorphism $\varphi : G \to \mathbb{Z}^s \times \mathbb{Z}^s \subseteq \mathbb{R}^s \times \mathbb{R}^s$ defined by sending $(g_1, g_2) \in G$ to

$$(v_1(g_1) \deg v_1, \ldots, v_s(g_1) \deg v_s, v_1(g_2) \deg v_1, \ldots, v_s(g_2) \deg v_s).$$

Note that $\varphi(G)$ is a subgroup of $\mathbb{Z}^s \times \mathbb{Z}^s$ of rank r.

Let $u, v \in \mathrm{Sol}(G)$ be such that $\varphi(u) = \varphi(v)$. Suppose that $u \neq v$. Then Lemma 2.3.7 implies that $H_K(u) \leq 0$. Hence by Lemma 2.2.1 part 2) it follows that u and thus v are in G^{tors}. This implies that the restriction of φ to $\mathrm{Sol}(G)$ is injective. In particular the restriction of φ to Prim-Sol(G) is injective. We now call $\mathcal{S} := \varphi(\mathrm{Sol}(G))$ and $\mathcal{PS} := \varphi(\text{Prim-Sol}(G))$. To prove Theorem 2.1.2 it suffices to bound the cardinality of $\mathcal{PS}$.

Let $||\cdot||$ be the norm on $\mathbb{R}^s \times \mathbb{R}^s$ that is the average of the $||\cdot||_1$ norms on $\mathbb{R}^s$. More precisely, we define for $u = (u_1, u_2) \in \mathbb{R}^s \times \mathbb{R}^s$

$$||u|| = \frac{1}{2}(||u_1|| + ||u_2||).$$

We now state the most important properties of $\mathcal{S}$.

Lemma 2.3.9. *The set $\mathcal{S} \subseteq \mathbb{Z}^s \times \mathbb{Z}^s$ has the following properties:*
1) For any two distinct $u, v \in \mathcal{S}$, we have that $||u|| \leq 2||v - u||$.
2) For any two distinct $u, v \in \mathcal{S}$ and any positive integer N such that $N < \frac{p}{3} - 2$, there is $M \in \{N, N+1\}$ such that $||u|| \leq \frac{2}{M+1}||v - (2M+1)u||$.
3) $p\mathcal{S} \subseteq \mathcal{S}$.

Proof. Let $x = (x_1, x_2) \in G$. By construction we have

$$||\varphi(x)|| = H_K^{\mathrm{hom}}(1, x_1) + H_K^{\mathrm{hom}}(1, x_2).$$

Note the basic inequalities

$$H_K^{\mathrm{hom}}(x_1, x_2) \leq H_K^{\mathrm{hom}}(1, x_1) + H_K^{\mathrm{hom}}(1, x_2) \leq 2H_K^{\mathrm{hom}}(x_1, x_2).$$

It is now clear that Lemma 2.3.7 implies part 1) and Lemma 2.3.8 implies part 2). Finally, part 3) is due to the action of the Frobenius operator. □

Denote by V the real span of $\varphi(G)$. Then V is an r-dimensional vector space over $\mathbb{R}$. We will keep writing $||\cdot||$ for the restriction of $||\cdot||$ to V.

Recall that our goal is to bound $|\mathcal{PS}|$. We sketch the ideas behind our strategy here. Let us first describe the strategy in characteristic 0 as used in [3] and [18]. In their work the set $\mathcal{S}$ satisfies part 1) of Lemma 2.3.9 and part 2) of Lemma 2.3.9 without the condition $N < \frac{p}{3} - 2$.

To finish the proof, they subdivide the vector space V in B^r cones for some absolute constant B. In each cone one can use part 1) of Lemma 2.3.9 to show that two distinct points $u, v \in \mathcal{S}$ are not too close. But part 2) of Lemma 2.3.9 shows that inside the same cone two points $u, v \in \mathcal{S}$ can not be too far apart. Together with a lower bound for the height of $u, v \in \mathcal{S}$, this proves that there are at most finitely many points $u \in \mathcal{S}$, say A, in each cone. Hence we get an upper bound of the shape $A \cdot B^r$.

Now we describe how to modify this to characteristic p. Again we subdivide V in B^r cones for some absolute constant B. From now on we only consider points $u \in \mathcal{PS}$ inside a fixed cone C. Our goal is to show that there are at most A points $u \in \mathcal{PS} \cap C$, where A is an absolute constant. It follows that then all points $v \in \mathcal{S} \cap C$ are of the shape $v = p^k u$ for $u \in \mathcal{PS}$ and $k \in \mathbb{Z}_{\geq 0}$.

Part 1) of Lemma 2.3.9 tells us that two distinct points $u, v \in \mathcal{PS}$ are not too close. Using part 3) of Lemma 2.3.9 we can multiply two points $u, v \in \mathcal{PS}$ with a power of p in such a way that the then obtained $u', v' \in \mathcal{S}$ satisfy $1 \leq \frac{||u'||}{||v'||} \leq \sqrt{p}$. Then we are in the position to apply part 2) of Lemma 2.3.9, which shows that $||u'||$ and $||v'||$ are not too far apart. This allows us to deduce that $\mathcal{PS} \cap C$ contains at most A points.

The following lemma subdivides the vector space V in B^r cones for some absolute constant B.

Lemma 2.3.10. *Given a real number $\theta > 0$, one can find a set $\mathcal{E} \subseteq \{u \in V : ||u|| = 1\}$ satisfying*
1) $|\mathcal{E}| \leq (1 + \frac{2}{\theta})^r$,
2) for all $0 \neq u \in V$ there exists $e \in \mathcal{E}$ satisfying $||\frac{u}{||u||} - e|| \leq \theta$.

Proof. See Lemma 6.3.4 in [18], which is an improvement of Corollary 3.8 in [3]. □

Let $\theta \in (0, \frac{1}{9})$ be a parameter and fix a corresponding choice of a set $\mathcal{E}$ satisfying the above properties. Given $e \in \mathcal{E}$, we define the cone

$$\mathcal{S}_e := \left\{ u \in \mathcal{S} : \left|\left| \frac{u}{||u||} - e \right|\right| \leq \theta \right\}, \ \mathcal{PS}_e := \mathcal{S}_e \cap \mathcal{PS}.$$

Fix $e \in \mathcal{E}$. We proceed to bound $|\mathcal{PS}_e|$. We start by deducing a so-called gap principle from part 1) of Lemma 2.3.9.

Lemma 2.3.11. *Let u_1, u_2 be distinct elements of $\mathcal{S}_e$, with $||u_2|| \geq ||u_1||$. Then $||u_2|| \geq \frac{3-\theta}{2+\theta} ||u_1||$.*

Proof. Write $\lambda_i := ||u_i||$ for $i = 1, 2$. Then we have $u_i = \lambda_i e + u_i'$ where $||u_i'|| \leq \theta \lambda_i$, by definition of $\mathcal{S}_e$. Part 1) of Lemma 2.3.9 gives

$$\lambda_1 \leq 2||(\lambda_2 - \lambda_1)e + (u_2' - u_1')|| \leq 2(\lambda_2 - \lambda_1) + \theta(\lambda_2 + \lambda_1),$$

and after dividing by λ_1 we get that

$$1 \leq 2\left(\frac{\lambda_2}{\lambda_1} - 1\right) + \theta\left(\frac{\lambda_2}{\lambda_1} + 1\right).$$

This can be rewritten as $\frac{3-\theta}{2+\theta} \leq \frac{\lambda_2}{\lambda_1}$. $\square$

From part 2) of Lemma 2.3.9 we can deduce the following crucial Lemma.

Lemma 2.3.12. *Let u_1, u_2 be distinct elements of $\mathcal{S}_e$. Suppose that $\frac{||u_2||}{||u_1||} < \frac{2}{3}p - 3$. Then $\frac{||u_2||}{||u_1||} \leq \frac{10}{\theta}$.*

Proof. We follow the proof of Lemma 6.4.9 of [18] part (ii) with a few modifications. For completeness we write out the full proof.

Again define $\lambda_i = ||u_i||$ and $u_i' = u_i - \lambda_i e$, for $i = 1, 2$. Assume that $\lambda_2 \geq \frac{10}{\theta}\lambda_1$. Let N be the positive integer with $2N + 1 \leq \frac{\lambda_2}{\lambda_1} < 2N + 3$. Then $2N + 1 < \frac{2}{3}p - 3$ and hence $N < \frac{p}{3} - 2$. Applying part 2) of Lemma 2.3.9 gives an integer $M \in \{N, N+1\}$ satisfying

$$\lambda_1 \leq \frac{2}{M+1}||(\lambda_2 - (2M+1)\lambda_1)e + u_2' - (2M+1)u_1'||.$$

Furthermore, we have that

$$|\lambda_2 - (2M+1)\lambda_1| \leq 2\lambda_1$$

and $M > \frac{4}{\theta}$ from the assumption $\lambda_2 \geq \frac{10}{\theta}\lambda_1$. Hence

$$\begin{aligned}
\lambda_1 &\leq \frac{2}{M+1}||(\lambda_2 - (2M+1)\lambda_1)e + u_2' - (2M+1)u_1'|| \\
&\leq \frac{2}{M+1}(2\lambda_1 + \lambda_2\theta + (2M+1)\lambda_1\theta) \\
&\leq \frac{2}{M+1}(2 + (4M+4)\theta)\lambda_1 = \left(\frac{4}{M+1} + 8\theta\right)\lambda_1 < 9\theta\lambda_1.
\end{aligned}$$

It follows that $\lambda_1 < \frac{1}{1-9\theta}$. Now observe that for any non-negative integer h the elements $p^h u_1, p^h u_2$ of $\mathcal{S}_e$ satisfy all the assumptions made so far. We conclude that also $p^h \lambda_1 < \frac{1}{1-9\theta}$ for every non-negative integer h, which implies that $||u_1|| = 0$. This contradicts the fact that $u_1 \in \mathcal{S}_e$, completing the proof. $\square$

Remark 2.3.13. In characteristic 0, the analogue of Lemma 2.3.12 holds only when both u_1, u_2 have norms at least $\frac{1}{1-9\theta}$. Then one deals with the remaining points in $\mathcal{S}_e$ by using the analogue of part 1) of Lemma 2.3.9, together with a separate argument to deal with the "very small" solutions. In characteristic p, it is because of the additional tool given by the action of Frobenius that the condition that u_1, u_2 have norm at least $\frac{1}{1-9\theta}$ has disappeared.

Assume without loss of generality that $\mathcal{PS}_e$ is not empty, and fix a choice of $u_0 \in \mathcal{PS}_e$ with $||u_0||$ minimal. For any $u \in \mathcal{PS}_e$, denote by $k(u)$ the smallest non-negative integer such that $\frac{||u||}{p^{k(u)}||u_0||} < p$ and denote $\lambda(u) := \frac{||u||}{p^{k(u)}||u_0||}$.

We define $\mathcal{PS}_e(1) := \{u \in \mathcal{PS}_e : \lambda(u) \leq \sqrt{p}\}$ and $\mathcal{PS}_e(2) := \{u \in \mathcal{PS}_e : \lambda(u) > \sqrt{p}\}$. Since we may assume $p > 7$ by Corollary 2.3.2, we have $\frac{2p}{3} - 3 > \sqrt{p}$.

Lemma 2.3.14. *1) Let $i \in \{1, 2\}$ and let u_1, u_2 be distinct elements of $\mathcal{PS}_e(i)$ with $\lambda(u_2) \geq \lambda(u_1)$. Then $\lambda(u_2) \geq \frac{3-\theta}{2+\theta}\lambda(u_1)$ and $\lambda(u_2) \leq \frac{10}{\theta}\lambda(u_1)$.*
2) $\lambda(\mathcal{PS}_e(2)) \subseteq [\frac{\theta p}{10}, p)$.
3) λ is an injective map on $\mathcal{PS}_e$.

Proof. 1) If $k(u_2) \geq k(u_1)$, we put $u_1' := p^{k(u_2)-k(u_1)}u_1$, $u_2' := u_2$, and if instead $k(u_2) < k_(u_1)$, we put $u_1' := u_1$, $u_2' := p^{k(u_1)-k(u_2)}u_2$. Now apply Lemma 14 and Lemma 15 to u_1', u_2'. We stress that u_1', u_2' are distinct elements of $\mathcal{S}_e$, since u_1, u_2 are distinct elements of $\mathcal{PS}_e(i)$.
2) This follows from Lemma 2.3.12 applied to the pair $(u_1, p^{k(u_1)+1}u_0)$ for each u_1 in $\mathcal{PS}_e(2)$.
3) Use part 1) and the fact that $\frac{3-\theta}{2+\theta} > 1$ for $\theta \in (0, \frac{1}{9})$. □

Proof of Theorem 2.1.2. By part 3) of Lemma 2.3.14 it suffices to bound $|\lambda(\mathcal{PS}_e)|$. By part 1) and 2) of Lemma 2.3.14 it will follow that we can bound $|\lambda(\mathcal{PS}_e)|$ purely in terms of θ: thus collecting all the bounds for e varying in $\mathcal{E}$ we obtain a bound depending only on r. We now give all the details.

For any $\theta \in (0, \frac{1}{9})$ we have

$$\frac{3-\theta}{2+\theta} > \frac{26}{19}.$$

Then we find that $|\lambda(\mathcal{PS}_e(1))|$ is at most the biggest n such that

$$\left(\frac{26}{19}\right)^{n-1} \leq \frac{10}{\theta}$$

and similarly for $|\lambda(\mathcal{PS}_e(2))|$. We conclude that

$$|\mathcal{PS}_e| \leq 2 + 2\frac{\log(\frac{10}{\theta})}{\log(\frac{26}{19})}.$$

Multiplying by $|\mathcal{E}|$ gives that for every $\theta \in (0, \frac{1}{9})$

$$|\mathcal{PS}| \leq 2\left(1 + \frac{\log(\frac{10}{\theta})}{\log(\frac{26}{19})}\right)\left(1 + \frac{2}{\theta}\right)^r.$$

So letting θ increase to $\frac{1}{9}$ we obtain

$$|\mathcal{PS}| \leq 2\left(1 + \frac{\log(90)}{\log(\frac{26}{19})}\right)19^r < 31 \cdot 19^r.$$

This completes the proof of Theorem 2.1.2. □

2.4 Proof of Theorem 2.1.1

First suppose that G and K are finitely generated. Before we can start with the proof of Theorem 2.1.1, we will rephrase Theorem 2.1.2. Recall that we write $\mathbb{F}_q$ for the algebraic closure of $\mathbb{F}_p$ in K.

Then Theorem 2.1.2 implies that there is a finite subset T of G with $|T| \leq 31 \cdot 19^r$ such that any solution of

$$x_1 + x_2 = 1, (x_1, x_2) \in G$$

with $x_1 \notin \mathbb{F}_q$ and $x_2 \notin \mathbb{F}_q$ satisfies $(x_1, x_2) = (\gamma, \delta)^{p^t}$ for some $t \in \mathbb{Z}_{\geq 0}$ and $(\gamma, \delta) \in T$.

Now let $(x_1, x_2) \in G$ be a solution to

$$ax_1 + bx_2 = 1.$$

If $ax_1 \in \mathbb{F}_q$ or $bx_2 \in \mathbb{F}_q$, it follows that both $ax_1 \in \mathbb{F}_q$ and $bx_2 \in \mathbb{F}_q$, which implies that $(a, b)^{q-1} \in G$. This contradicts the condition on (a, b) in Theorem 2.1.1.

Hence $ax_1 \notin \mathbb{F}_q$ and $bx_2 \notin \mathbb{F}_q$. Define G' to be the group generated by G and the tuple (a, b). Then the rank of G' is at most $r+1$. Let $T \subseteq G'$ be as above, so $|T| \leq 31 \cdot 19^{r+1}$. We can write

$$(ax_1, bx_2) = (\gamma, \delta)^{p^t}$$

with $t \in \mathbb{Z}_{\geq 0}$ and $(\gamma, \delta) \in T$. Since $T \subseteq G'$, we can write

$$(\gamma, \delta) = (a^k y_1, b^k y_2)$$

with $k \in \mathbb{Z}$ and $(y_1, y_2) \in G$. This means that

$$(ax_1, bx_2) = (a^k y_1, b^k y_2)^{p^t},$$

which implies $(a, b)^{kp^t - 1} \in G$. If $kp^t - 1$ is co-prime to p, we have a contradiction with the condition on (a, b) in Theorem 2.1.1. But p can only divide $kp^t - 1$ if $t = 0$. Then we find immediately that there are at most $|T| \leq 31 \cdot 19^{r+1}$ solutions as desired.

We still need to deal with the case that K is an arbitrary field of characteristic p and G is a subgroup of $K^* \times K^*$ with $\dim_{\mathbb{Q}} G \otimes_{\mathbb{Z}} \mathbb{Q} = r$ finite. Suppose that $ax_1 + bx_2 = 1$ has more than $31 \cdot 19^{r+1}$ solutions $(x_1, x_2) \in G$. Then we can replace G by a finitely generated subgroup of G with the same property. We can also replace K by a subfield, finitely generated over its prime field, containing the coordinates of the new G and a, b. This gives the desired contradiction.

2.5 Acknowledgements

We are grateful to Julian Lyczak for explaining us how identities as in Lemma 2.3.4 follow from basic properties of hypergeometric functions. Many thanks go to Jan-Hendrik Evertse for providing us with this nice problem, his help throughout and the proofreading.

Addendum

Joint work with Carlo Pagano

On the 22nd October of 2018 Professor Felipe Voloch brought to our attention the unpublished master thesis of Yi-Chih Chiu, written under the supervision of Professor Ki-Seng Tan. In this work, Chiu establishes a special case of our main theorems [44, Theorem 1.1, Theorem 1.2]. We shall begin by explaining his result, and we will next compare it to our result.

Let p be a prime number. For a field extension K of $\mathbb{F}_p$ with transcendence degree equal to 1, we let k be the algebraic closure of $\mathbb{F}_p$ in K. Denote by Ω_K the set of valuations of K. Let S be a finite subset of Ω_K and fix $\alpha, \beta \in K^*$. The following theorem is proven in Chiu's master thesis.

Theorem 2.5.1. *The S-unit equation to be solved in $x, y \in \mathcal{O}_S^*$*

$$\alpha x + \beta y = 1,$$

has at most $3 \cdot 7^{2|S|-2}$ pairwise inequivalent non-trivial solutions if $\alpha, \beta \in \mathcal{O}_S^*$. If instead α, β are not both in $\mathcal{O}_S^*$, then it has at most $39 \cdot 7^{2|S|-2}$ non-trivial solutions.

Here a solution (x, y) is called trivial if $\frac{\alpha x}{\beta y} \in k$. Two solutions $(x_1, y_1), (x_2, y_2)$ are said to be equivalent if there exists $n \in \mathbb{Z}_{\geq 0}$ with

$$(\alpha x_1)^{p^n} = \alpha x_2, (\beta y_1)^{p^n} = \beta y_2 \quad \text{or} \quad (\alpha x_2)^{p^n} = \alpha x_1, (\beta y_2)^{p^n} = \beta y_1.$$

This result is a special case with slightly better constants of our theorems that we state now for the reader's convenience, see [44, Theorem 1.1, Theorem 1.2].

Theorem 2.5.2. *Let K be a field of characteristic $p > 0$. Take $\alpha, \beta \in K^*$ and let G be a finitely generated subgroup of $K^* \times K^*$ of rank $r := \dim_{\mathbb{Q}} G \otimes \mathbb{Q}$. Then the equation*

$$\alpha x + \beta y = 1,$$

to be solved in $(x, y) \in G$, has at most $31 \cdot 19^r$ pairwise inequivalent non-trivial solutions if $(\alpha, \beta)^n \in G$ for some $n > 0$. If instead $(\alpha, \beta)^n \notin G$ for all $n > 0$, then it has at most $31 \cdot 19^{r+1}$ non-trivial solutions.

Note that Theorem 2.5.2 applies to *any* finitely generated subgroup in *any* field of characteristic p. In contrast, Chiu's theorem applies only to the case of S-units of fields of transcendence degree 1 (with some care Chiu's theorem can be extended to S-units of function fields of projective varieties).

The reason for this difference in generality comes from the fact that Chiu's work is an adaptation of Evertse's work [17] to characteristic p. Our work is instead an adaptation of the work of Beukers and Schlickewei [3] to characteristic p. In both works [3, 17], there is a key use of a certain set of identities coming from hypergeometric functions, see [44, Lemma 3.3, Lemma 3.4]. In characteristic p these identities can be used only in a limited range, see [9, Proposition 2] and [44, Corollary 3.5] respectively.

Correspondingly, the solutions to the unit equations need to be counted only up to equivalence. One of the most important steps is to use this equivalence relation in such a way that one is inside this limited range. It is this step that allows one to obtain an upper bound that is independent of p. The reader can find this step in the two papers respectively at [9, Lemma 4] and at [44, Lemma 3.9].

Chapter 3

Unit equations and Fermat surfaces in positive characteristic

Joint work with Carlo Pagano

Abstract

In this article we study the three-variable unit equation $x + y + z = 1$ to be solved in $x, y, z \in \mathcal{O}_S^*$, where $\mathcal{O}_S^*$ is the S-unit group of some global function field. We give upper bounds for the height of solutions and the number of solutions. We also apply these techniques to study the Fermat surface $x^N + y^N + z^N = 1$.

3.1 Introduction

Let K be a finitely generated field over $\mathbb{F}_p$ of transcendence degree 1. Denote by $\mathbb{F}_q$ the algebraic closure of $\mathbb{F}_p$ inside K, which is a finite extension of $\mathbb{F}_p$. Let M_K be the set of places of K and let $S \subseteq M_K$ be a finite subset. To avoid degenerate cases, we will assume that $|S| \geq 2$ throughout the chapter. We define $\omega(S) = \sum_{v \in S} \deg(v)$ and we let H_K be the usual height. For a precise definition of $\deg(v)$ and H_K we refer the reader to Section 3.2. Mason [54] and Silverman [67] independently considered the equation

$$x + y = 1 \text{ in } x, y \in \mathcal{O}_S^*. \tag{3.1}$$

If $x, y \notin K^p$ is a solution to (3.1), they showed that

$$H_K(x) = H_K(y) \leq \omega(S) + 2g - 2, \tag{3.2}$$

where g is the genus of K. Previously, Stothers [72] proved (3.2) for polynomials $x, y \in \mathbb{C}[t]$.

It is important to note that the condition $x, y \notin K^p$ can not be removed. Indeed if we have a solution to (3.1), then we find that

$$x^{p^k} + y^{p^k} = 1$$

is also a solution to (3.1) for all integers $k \geq 0$ due to Frobenius, but the heights $H_K(x^{p^k})$ and $H_K(y^{p^k})$ become arbitrarily large. This new phenomenon is the main difficulty in dealing with two variable unit equations in positive characteristic.

The work of Mason and Silverman has been extended in various directions. Hsia and Wang [36] looked at the equation

$$x_1 + \cdots + x_n = 1 \text{ in } x_1, \ldots, x_n \in \mathcal{O}_S^*. \tag{3.3}$$

They were able to deduce a height bound similar to (3.2) under the condition that $x_1, \ldots, x_n$ are linearly independent over K^p. In particular it follows that under the same condition there are only finitely many solutions $x_1, \ldots, x_n$. Derksen and Masser [16] considered (3.3) without the restriction that $x_1, \ldots, x_n$ are linearly independent over K^p. In this case it is not a priori clear what the structure of the solution set should be, but Derksen and Masser give a completely explicit description that we repeat here in the special case that $n = 3$.

They define so-called one dimensional Frobenius families to be

$$\mathcal{F}(\mathbf{u}) := \{(u_1, u_2, u_3)^{p^e} : e \geq 0\}$$

for $\mathbf{u} = (u_1, u_2, u_3) \in (K^*)^3$ and two dimensional Frobenius families

$$\mathcal{F}_a(\mathbf{u}, \mathbf{v}) := \left\{ \left((u_1, u_2, u_3)(v_1, v_2, v_3)^{p^{af}} \right)^{p^e} : e, f \geq 0 \right\}$$

for $a \in \mathbb{Z}_{\geq 1}$, $\mathbf{u} = (u_1, u_2, u_3) \in (K^*)^3$, $\mathbf{v} = (v_1, v_2, v_3) \in (K^*)^3$, where all multiplications of tuples are taken coordinate-wise. Then Derksen and Masser prove that the solution set of

$$x + y + z = 1 \text{ in } x, y, z \in \mathcal{O}_S^* \tag{3.4}$$

is equal to a finite union of one dimensional and two dimensional Frobenius families. On top of that Derksen and Masser give effective height bounds for $\mathbf{u}$ and $\mathbf{v}$, which can be seen as another direct generalization of (3.2). In principle this also gives an upper bound on the total number of Frobenius families that one may need to describe the solution set of (3.4), but the resulting bounds are far from optimal. Leitner [49] computed the full solution set of (3.4) in the special case $S = \{0, 1, \infty\}$ and $K = \mathbb{F}_p(t)$.

In this chapter we give explicit upper bounds for the height of $\mathbf{u}$ and $\mathbf{v}$ in the case $n = 3$. Together with a "gap principle" we will use this to give an upper bound on the number of Frobenius families. For the two variable unit equation $x + y = 1$ such upper bounds have already been established by Voloch [78] and by Koymans and Pagano [44] using different methods than in this chapter. The upper bound in the latter paper has the particularly pleasant feature that it does not depend on p. This chapter is based on the paper of Beukers and Schlickewei [3], who had previously established a finiteness result for the two variable unit equation in characteristic 0.

Let g and γ be respectively the genus and the gonality of K. Put

$$c_{K,S} := 2\omega(S) + 4g - 4 + 4\gamma,\ c'_{K,S} := 2c_{K,S} \cdot (\omega(S) + 4c_{K,S} + 2g - 2) + 3c_{K,S}.$$

Define the following three sets

$$\begin{aligned}
A &:= \{\mathbf{x} = (x, y, z) \in (\mathcal{O}_S^*)^3 : x + y + z = 1, x, y, z \notin \mathbb{F}_q^*, H_K(x), H_K(y), H_K(z) \leq c'_{K,S}\}, \\
B_p &:= \{(\mathbf{u}, \mathbf{v}) \in (\mathcal{O}_S^*)^3 \times (\mathcal{O}_S^*)^3 : \mathbf{u}, \mathbf{v} \notin (\mathbb{F}_q^*)^3, (u_i, v_i) \notin \mathbb{F}_q^* \times \mathbb{F}_q^* \text{ for } i = 1, 2, 3, \\
&\qquad H_K(u_i) \leq c_{K,S} \text{ for } i = 1, 2, 3, \\
&\qquad H_K(v_i) \leq \omega(S) + 2g - 2 \text{ for } i = 1, 2, 3, \\
&\qquad u_1 v_1^{p^f} + u_2 v_2^{p^f} + u_3 v_3^{p^f} = 1 \text{ for all } f \in \mathbb{Z}_{\geq 0}\}, \\
B_q &:= \{(\mathbf{u}, \mathbf{v}) \in (\mathcal{O}_S^*)^3 \times (\mathcal{O}_S^*)^3 : \mathbf{u}, \mathbf{v} \notin (\mathbb{F}_q^*)^3, (u_i, v_i) \notin \mathbb{F}_q^* \times \mathbb{F}_q^* \text{ for } i = 1, 2, 3, \\
&\qquad H_K(u_i) \leq c_{K,S}, \text{ for } i = 1, 2, 3, \\
&\qquad H_K(v_i) \leq \frac{q}{p}(\omega(S) + 2g - 2), \text{ for } i = 1, 2, 3, \\
&\qquad u_1 v_1^{q^f} + u_2 v_2^{q^f} + u_3 v_3^{q^f} = 1 \text{ for all } f \in \mathbb{Z}_{\geq 0}\}.
\end{aligned}$$

Theorem 3.1.1. *For all $x, y, z \notin \mathbb{F}_q$ we have the following equivalence: x, y, z is a solution to (3.4) if and only if (x, y, z) is an element of one of the following three sets*

$$\bigcup_{\mathbf{x} \in A} \mathcal{F}(\mathbf{x}), \quad \bigcup_{(\mathbf{u}, \mathbf{v}) \in B_p} \mathcal{F}_1(\mathbf{u}, \mathbf{v}), \quad \bigcup_{(\mathbf{u}, \mathbf{v}) \in B_q} \mathcal{F}_{\log_p(q)}(\mathbf{u}, \mathbf{v}). \tag{3.5}$$

The novel feature of Theorem 3.1.1 is the excellent quality of the height bounds appearing in the definition of A, B_p and B_q. Because we are only dealing with the three variable unit equation, the descent step of Derksen and Masser becomes completely explicit. We make full use of this to improve on the height bounds obtained by Derksen and Masser.

We have the identity

$$\mathcal{F}_1(\mathbf{u}, \mathbf{v}) = \bigcup_{x=0}^{\log_p(q)-1} \mathcal{F}_{\log_p(q)}\left(\mathbf{u}, \mathbf{v}^{p^x}\right).$$

This allows us to remove the sets

$$\bigcup_{(\mathbf{u}, \mathbf{v}) \in B_p} \mathcal{F}_1(\mathbf{u}, \mathbf{v})$$

from Theorem 3.1.1 if desired. We have decided to state Theorem 3.1.1 in its current shape because the sets $\mathcal{F}_1(\mathbf{u}, \mathbf{v})$ naturally show up in the proof. Furthermore, it enables us to be more precise in our next theorem.

Theorem 3.1.2. *There are a subset C_1 of $(K^*)^3$ and subsets C_2 and C_3 of $(K^*)^3 \times (K^*)^3$ with the following properties*

- $|C_1| \leq 279 \cdot q^2 \cdot (\log_{\frac{5}{4}}(3c'_{K,S}) + 1)^2 \cdot (15 \cdot 10^6)^{|S|}$;
- $|C_2| \leq 2883 \cdot p^4 \cdot 19^{4|S|}$;
- $|C_3| \leq 2883 \cdot \log_p(q) \cdot q^4 \cdot 19^{4|S|}$;
- *for all $x, y, z \notin \mathbb{F}_q$ we have the following equivalence: x, y, z is a solution to (3.4) if and only if (x, y, z) is an element of one of the following three sets*

$$\bigcup_{\mathbf{x} \in C_1} \mathcal{F}(\mathbf{x}), \quad \bigcup_{(\mathbf{u},\mathbf{v}) \in C_2} \mathcal{F}_1(\mathbf{u}, \mathbf{v}), \quad \bigcup_{(\mathbf{u},\mathbf{v}) \in C_3} \mathcal{F}_{\log_p(q)}(\mathbf{u}, \mathbf{v}).$$

The work of Derksen and Masser quickly implies that there are finite subsets C_1, C_2 and C_3 satisfying the fourth condition in Theorem 3.1.2; indeed, Derksen and Masser show that C_1, C_2 and C_3 can be taken to be sets of bounded height. This gives effective upper bounds for $|C_1|$, $|C_2|$ and $|C_3|$, but the resulting bounds are rather poor. Our improvement comes from Theorem 3.1.1, the aforementioned "gap principle" and a reduction step to the two variable unit equation, which brings the results of [44] in play.

Let $N > 0$ be an integer. As is well known there is a strong relation between unit equations and the Fermat equation

$$x_1^N + \ldots + x_m^N = 1$$

to be solved in $x_1, \ldots, x_m \in k(t)$ for some field k. This relation has been used in characteristic 0 by for example Voloch [77] and Bombieri and Mueller [5]. However, it is not clear how these methods can be made to work in characteristic $p > 0$. For example it would be natural to try and use a height bound for (3.3), but this is only possible when $x_1^N, \ldots, x_m^N$ are linearly independent over K^p. In the special case $m = 2$ this problem has been considered by Silverman [66], but unfortunately his main theorem is false. A correct statement with proof can be found in [40]. Here we will analyze the case $m = 3$.

Definition 3.1.3. We say that an integer $N > 0$ is (x, p)-good if the congruence

$$ap^s + b \equiv 0 \mod N$$

has no solutions in integers $s \geq 0$, $0 < a, b \leq x$.

We remark that for a given tuple (x, p) a positive density of the primes is (x, p)-good. Indeed, if $N > 2$ is a prime satisfying

$$\left(\frac{-1}{N}\right) = -1, \quad \left(\frac{p}{N}\right) = 1, \quad \left(\frac{a}{N}\right) = 1 \text{ for } 0 < a \leq x,$$

then N is (x, p)-good.

Theorem 3.1.4. *Let $p > 480$ be a prime number and suppose that N is a $(480, p)$-good integer. If we further suppose that $\gcd(N, p) = 1$, then the Fermat surface*

$$x^N + y^N + z^N = 1 \tag{3.6}$$

has no solutions $x, y, z \in \mathbb{F}_p(t)$ satisfying $x, y, z \notin \mathbb{F}_p(t^p)$ and $x/y, x/z, y/z \notin \mathbb{F}_p(t^p)$.

Note that Theorem 3.1.4 is in stark contrast with the behavior of the Fermat surface in characteristic 0 [77]. Remarkably enough it turns out that Theorem 3.1.4 becomes false if we drop any of the last two conditions, see Section 3.6. We will also explain there why we need the condition that N is $(480, p)$-good. The rough reason is that if N is not $(1, p)$-good, then the Fermat surface is known to be unirational [63]. Our work shows that the unirationality of these surfaces is strongly related to the two-dimensional Frobenius families appearing in Theorem 3.1.1. For precise details, we refer the reader to Section 3.6.

3.2 Preliminaries

In this section we start by defining heights, which will play a key role throughout the chapter. Furthermore, we give two important lemmata about heights.

3.2.1 Definition of height

Recall that K is a finitely generated field over $\mathbb{F}_p$ of transcendence degree 1 and that $\mathbb{F}_q$ is the algebraic closure of $\mathbb{F}_p$ inside K. We further recall that M_K is the set of places of K. The valuation ring of a place $v \in M_K$ is given by

$$O_v := \{x \in K : v(x) \geq 0\}.$$

This is a discrete valuation ring with maximal ideal $m_v := \{x \in K : v(x) > 0\}$. The residue class field O_v/m_v naturally becomes a finite field extension of $\mathbb{F}_q$. Hence

$$\deg(v) := [O_v/m_v : \mathbb{F}_q]$$

is a well-defined integer. With these definitions it turns out that the sum formula holds for all $x \in K^*$, i.e.

$$\sum_v v(x) \deg(v) = 0,$$

where here and below $\sum_v$ denotes a summation over $v \in M_K$. This allows us to define the height for $x \notin \mathbb{F}_q$ as follows

$$H_K(x) := [K : \mathbb{F}_q(x)] = \sum_{v \in M_K} \max(v(x), 0) \deg(v) = \sum_{v \in M_K} -\min(v(x), 0) \deg(v).$$

For $x \in \mathbb{F}_q$ we set $H_K(x) := 0$. More generally, we define the projective height to be

$$H_K(x_0 : \ldots : x_n) := -\sum_{v \in M_K} \min(v(x_0), \ldots, v(x_n)) \deg(v)$$

for $(x_0 : \ldots : x_n) \in \mathbb{P}^n(K)$, which is well-defined due to the sum formula. One can recover the usual height by the identity $H_K(x) = H_K(1 : x)$.

3.2.2 Height lemmata

Pick $t \in K^*$ such that $K/\mathbb{F}_q(t)$ is of the minimal possible degree γ, the gonality of K. Then it follows that $K/\mathbb{F}_q(t)$ is a separable extension. Let D be the extension to K of the derivation $\frac{d}{dt}$ on $\mathbb{F}_q(t)$. We will fix such a derivation D for the remainder of the chapter. The following lemma will be important throughout.

Lemma 3.2.1. *The map $f : K^* \to K$ given by*

$$f(x) = \frac{Dx}{x}$$

is a homomorphism with kernel K^p.

Proof. The Leibniz rule implies that f is a homomorphism. Furthermore, the following is a standard fact regarding derivations

$$Dx = 0 \iff x \in K^p,$$

which immediately implies that the kernel of f is K^p. □

For every place $v \in M_K$, we choose an element z_v of K satisfying $v(z_v) = 1$. Since $K/\mathbb{F}_q(z_v)$ is a separable extension, we can uniquely extend the derivation $\frac{d}{dz_v}$ to K. For $x \in K^*$ we write $\omega(x) = \sum_{v : v(x) \neq 0} \deg(v)$.

Lemma 3.2.2. *Let $f \in K^*$. Then for $f \notin K^p$*

$$H_K\left(\frac{Df}{f}\right) \leq \omega(f) + 2g - 2 + 2\gamma,$$

where g is the genus of K.

Proof. We have

$$H_K\left(\frac{Df}{f}\right) = \frac{1}{2}\sum_v \left|v\left(\frac{Df}{f}\right)\right| \deg(v).$$

We may write

$$v\left(\frac{Df}{f}\right) = \left(v\left(\frac{df}{dz_v}\right) - v(f)\right) - v\left(\frac{dt}{dz_v}\right).$$

Therefore we get that

$$H_K\left(\frac{Df}{f}\right) = \frac{1}{2}\sum_v \left| v\left(\frac{Df}{f}\right)\right| \deg(v) \leq$$
$$\frac{1}{2}\cdot\left(\sum_v \left| v\left(\frac{df}{dz_v}\right) - v(f)\right| \deg(v) + \sum_v \left| v\left(\frac{dt}{dz_v}\right)\right| \deg(v)\right).$$

We call the two inner sums respectively T_1 and T_2.

Bound for T_1

By the Riemann-Roch Theorem, see e.g. equation (5) of page 96, chapter 6 in [54], we have for $f \notin K^p$ that

$$\sum_v v\left(\frac{df}{dz_v}\right) \deg(v) = 2g - 2 \tag{3.7}$$

and hence by the sum formula

$$\sum_v \left(v\left(\frac{df}{dz_v}\right) - v(f)\right) \deg(v) = 2g - 2.$$

Furthermore $v\left(\frac{df}{dz_v}\right) - v(f) < 0$ implies $v\left(\frac{df}{dz_v}\right) - v(f) = -1$. Therefore

$$\sum_{v : v\left(\frac{df}{dz_v}\right) < v(f)} \left| v\left(\frac{df}{dz_v}\right) - v(f)\right| \deg(v) \leq \omega(f)$$

and thus

$$\sum_{v : v\left(\frac{df}{dz_v}\right) \geq v(f)} \left(v\left(\frac{df}{dz_v}\right) - v(f)\right) \deg(v) \leq 2g - 2 + \omega(f).$$

In total we get that

$$T_1 \leq 2\omega(f) + 2g - 2.$$

Bound for T_2

We use equation (3.7) with $f = t$ to obtain

$$\sum_v v\left(\frac{dt}{dz_v}\right) \deg(v) = 2g - 2. \tag{3.8}$$

If $v(t) \geq 0$, then we clearly have $v\left(\frac{dt}{dz_v}\right) \geq 0$. On the other hand if $v(t) < 0$, we have

$$v\left(\frac{dt}{dz_v}\right) = v(t) - 1.$$

Hence

$$\sum_{v:v\left(\frac{dt}{dz_v}\right)<0} \left| v\left(\frac{dt}{dz_v}\right) \right| \deg(v) = \sum_{v:v(t)<0} (1 - v(t)) \deg(v) \leq -2 \sum_{v:v(t)<0} v(t) \deg(v) = 2\gamma, \tag{3.9}$$

which we can combine with equation (3.8) to deduce

$$\sum_{v:v\left(\frac{dt}{dz_v}\right)\geq 0} v\left(\frac{dt}{dz_v}\right) \deg(v) \leq 2g - 2 + 2\gamma \tag{3.10}$$

After adding equation (3.9) and equation (3.10), we conclude that

$$T_2 \leq 2g - 2 + 4\gamma.$$

Conclusion of proof
In total we get

$$H_K\left(\frac{Df}{f}\right) \leq \frac{1}{2}(T_1 + T_2) \leq \omega(f) + 2g - 2 + 2\gamma,$$

which is the desired inequality. □

We will repeatedly use the following two theorems.

Theorem 3.2.3. *Let $x, y \in \mathcal{O}_S^*$. If $x, y \notin K^p$ and*

$$x + y = 1,$$

then we have

$$H_K(x) = H_K(y) \leq \omega(S) + 2g - 2.$$

Proof. See [54] and [67]. □

Theorem 3.2.4. *Let K be a field of characteristic $p > 0$ and let G be a finitely generated subgroup of $K^* \times K^*$ of rank r. Then the equation*

$$x + y = 1 \text{ in } (x, y) \in G$$

has at most $31 \cdot 19^r$ solutions (x, y) satisfying $(x, y) \notin G^p$.

Proof. This is Theorem 1.2 of [44]. □

3.3 Proof of Theorem 3.1.1

Proof. By construction $\mathcal{F}(\mathbf{x})$ is a solution to (3.4) for $\mathbf{x} \in A$ and likewise all elements of $\mathcal{F}_a(\mathbf{u}, \mathbf{v})$ are solutions to (3.4). Hence it suffices to prove the only if part of Theorem 3.1.1. Let x, y, z be a solution of (3.4) with $x, y, z \notin \mathbb{F}_q$. Note that the sets as given in equation (3.5) are all invariant under taking p-th roots. Since $x, y, z \notin \mathbb{F}_q$, we can keep taking p-th roots of the tuple (x, y, z) until x, y or z is not in K^p. For ease of notation we will keep using the same letters for the new x, y and z. By symmetry we may assume that $z \notin K^p$. Then also $x \notin K^p$ or $y \notin K^p$. Again we may assume by symmetry that $y \notin K^p$. Now we distinguish two cases.

Case I: First suppose that $x \in K^p$. Then using

$$x + y + z = 1$$

we find after differentiating with respect to D

$$\frac{Dy}{y} y + \frac{Dz}{z} z = 0.$$

We can rewrite this as follows

$$x + y\left(1 - \frac{z}{Dz}\frac{Dy}{y}\right) = 1$$
$$x + z\left(1 - \frac{y}{Dy}\frac{Dz}{z}\right) = 1.$$

Define $a_2 := 1 - \frac{z}{Dz}\frac{Dy}{y}$ and $b_3 := 1 - \frac{y}{Dy}\frac{Dz}{z}$. Note that $a_2 = 0$ implies $x = 1$, contrary to our assumption $x \notin \mathbb{F}_q$. Similarly $b_3 \neq 0$. The above system of equations implies that either $b_3, a_2 \notin \mathcal{O}_S^*$ or $b_3, a_2 \in \mathcal{O}_S^*$. Consider first the case $b_3, a_2 \notin \mathcal{O}_S^*$. By Lemma 3.2.2 we have

$$H_K(b_3) \leq c_{K,S}. \tag{3.11}$$

We set $l := \lfloor \log_p c_{K,S} \rfloor + 1$ and claim that $b_3 z \notin K^{p^l}$. Take $v \notin S$ such that $v(b_3) \neq 0$; such a valuation exists by our assumption that $b_3 \notin \mathcal{O}_S^*$. From the height bound in equation (3.11) we deduce that

$$|v(b_3)| \leq H_K(b_3) \leq c_{K,S}.$$

Since $z \in \mathcal{O}_S^*$, we conclude that

$$0 \neq |v(b_3 z)| \leq c_{K,S}.$$

This immediately implies that $b_3 z \notin K^{p^l}$, which establishes our claim.

Write $x = \delta^{p^s}$ and $b_3 z = \epsilon^{p^s}$, with $\delta, \epsilon \notin K^p$. Note that $\delta + \epsilon = 1$, so an application of Theorem 3.2.3 gives

$$H_K(\delta) = H_K(\epsilon) \leq \omega(S) + 2c_{K,S} + 2g - 2,$$

where we used that $\omega(b_3) \leq 2H_K(b_3) \leq 2c_{K,S}$. We conclude that

$$H_K(x) = H_K(b_3 z) = p^s H_K(\delta) = p^s H_K(\epsilon) \leq c_{K,S} \cdot (\omega(S) + 2c_{K,S} + 2g - 2),$$

since $p^s \leq p^{l-1} \leq c_{K,S}$.

We now consider the case that $a_2, b_3 \in \mathcal{O}_S^*$. Since $x \notin \mathbb{F}_q$ there is $x' \notin K^p$ such that $x = x'^{p^s}$ for some $s > 0$. There are also $y', z' \in \mathcal{O}_S^*$ such that

$$x' + a_2 y' = 1 \tag{3.12}$$
$$x' + b_3 z' = 1. \tag{3.13}$$

Applying Theorem 3.2.3 again yields

$$H_K(x') = H_K(a_2 y') \leq \omega(S) + 2g - 2.$$

We conclude that

$$(x, y, z) \in \mathcal{F}_1((1, a_2^{-1}, b_3^{-1}), (x', a_2 y', b_3 z')),$$

with $a_2, b_3 \notin \mathbb{F}_q$, since otherwise $y, z \in K^p$, which would be a contradiction. Using the identity $a_2^{-1} + b_3^{-1} = 1$ and equations (3.12), (3.13), we quickly verify the identity

$$x'^{p^f} + a_2^{-1}(a_2 y')^{p^f} + b_3^{-1}(b_3 z')^{p^f} = 1$$

for all integers $f \geq 0$, so that indeed $((1, a_2^{-1}, b_3^{-1}), (x', a_2 y', b_3 z')) \in B_1$.

Case II: Now suppose $x \notin K^p$. We start by dealing with the case $\frac{x}{Dx} \neq \frac{y}{Dy}$, $\frac{x}{Dx} \neq \frac{z}{Dz}$, $\frac{y}{Dy} \neq \frac{z}{Dz}$. Then we find that

$$x + y + z = 1$$

and after differentiating with respect to D

$$\frac{Dx}{x}x + \frac{Dy}{y}y + \frac{Dz}{z}z = 0.$$

This is equivalent to

$$x\left(1 - \frac{z}{Dz}\frac{Dx}{x}\right) + y\left(1 - \frac{z}{Dz}\frac{Dy}{y}\right) = 1$$
$$x\left(1 - \frac{y}{Dy}\frac{Dx}{x}\right) + z\left(1 - \frac{y}{Dy}\frac{Dz}{z}\right) = 1.$$

For convenience we define

$$a_1 := 1 - \frac{z}{Dz}\frac{Dx}{x}, a_2 := 1 - \frac{z}{Dz}\frac{Dy}{y}, b_1 := 1 - \frac{y}{Dy}\frac{Dx}{x}, b_3 := 1 - \frac{y}{Dy}\frac{Dz}{z}.$$

By our assumption we know that the coefficients a_1, a_2, b_1 and b_3 are not zero. If one of the coefficients, say a_1, does not lie in $\mathcal{O}_S^*$, we can proceed exactly as before obtaining the bound

$$H_K(a_1 x) = H_K(a_2 y) \leq c_{K,S} \cdot (\omega(S) + 4c_{K,S} + 2g - 2).$$

So now suppose that $a_1, a_2, b_1, b_3 \in \mathcal{O}_S^*$, but also suppose that $d := \frac{a_1}{b_1} \notin \mathbb{F}_q^*$. In this case we have

$$H_K(d) \leq 2c_{K,S}$$

and therefore $a_1 x \notin K^{p^l}$ or $b_1 x \notin K^{p^l}$ with $l := \lfloor \log_p 2c_{K,S} \rfloor + 1$. Suppose that $a_1 x \notin K^{p^l}$. Then Theorem 3.2.3 gives

$$H_K(a_1 x) = H_K(a_2 y) \leq 2c_{K,S} \cdot (\omega(S) + 4c_{K,S} + 2g - 2)$$

and the other case can be dealt with in exactly the same way.

Finally suppose that $a_1, a_2, b_1, b_3 \in \mathcal{O}_S^*$ and $d \in \mathbb{F}_q^*$. If we additionally suppose that one of the coefficients is in $\mathbb{F}_q^*$, another application of Theorem 3.2.3 yields

$$H_K(a_1 x) = H_K(a_2 y) = H_K(b_1 x) = H_K(b_3 z) \leq \omega(S) + 2g - 2.$$

Hence we will assume that $a_1, a_2, b_1, b_3 \notin \mathbb{F}_q^*$ from now on. If $a_1 x \in \mathbb{F}_q^*$, we immediately get a height bound for x. So we may further assume that $a_1 x \notin \mathbb{F}_q^*$. Then let $l \geq 0$ be the largest integer such that $a_1 x \in K^{q^l}$. Define $x' \in \mathcal{O}_S^*$ as

$$(a_1 x')^{q^l} = a_1 x$$

and then define $y', z' \in \mathcal{O}_S^*$ such that

$$a_1 x' + a_2 y' = 1 \tag{3.14}$$
$$b_1 x' + b_3 z' = 1. \tag{3.15}$$

Furthermore,

$$H_K(a_1 x') = H_K(a_2 y') \leq \frac{q}{p}(\omega(S) + 2g - 2)$$

and

$$(x, y, z) \in \mathcal{F}_{\log_p(q)}((a_1^{-1}, a_2^{-1}, b_3^{-1}), (a_1 x', a_2 y', b_3 z')).$$

Once more, a direct verification using $a_2^{-1} + b_3^{-1} = 1$, $1 - \frac{a_1}{a_2} - \frac{b_1}{b_3} = 0$ and the equations (3.14), (3.15) shows that

$$a_1^{-1}(a_1 x')^{q^f} + a_2^{-1}(a_2 y')^{q^f} + b_3^{-1}(b_3 z')^{q^f} = 1$$

for all integers $f \geq 0$. We conclude that $((a_1^{-1}, a_2^{-1}, b_3^{-1}), (a_1 x', a_2 y', b_3 z')) \in B_q$. This deals with the case $x \notin K^p$ and $\frac{x}{Dx} \neq \frac{y}{Dy}$, $\frac{x}{Dx} \neq \frac{z}{Dz}$, $\frac{y}{Dy} \neq \frac{z}{Dz}$.

We still have to deal with the case $x \notin K^p$ and $\frac{x}{Dx} = \frac{y}{Dy}$ or $\frac{x}{Dx} = \frac{z}{Dz}$ or $\frac{y}{Dy} = \frac{z}{Dz}$. Recall that $y, z \notin K^p$ as well, hence the three cases are symmetrical. So we will only deal with the case $\frac{y}{Dy} = \frac{z}{Dz}$. Then we get the equations

$$x\left(1 - \frac{y}{Dy}\frac{Dx}{x}\right) = x\left(1 - \frac{z}{Dz}\frac{Dx}{x}\right) = 1$$

and hence

$$H_K(x) \leq c_{K,S}.$$

Our equation implies that $a_1 := b_1 := 1 - \frac{y}{Dy}\frac{Dx}{x} \in \mathcal{O}_S^*$. Substitution in the original equation yields

$$\frac{1}{a_1} + y + z = 1$$

or equivalently

$$y + z = 1 - \frac{1}{a_1} = \frac{a_1 - 1}{a_1}.$$

After putting $\alpha := \frac{a_1}{a_1 - 1}$ we get

$$\alpha y + \alpha z = 1.$$

Note that

$$H_K(\alpha) = H_K(a_1) = H_K(x) \leq c_{K,S}.$$

Suppose that $\alpha \notin \mathcal{O}_S^*$. Just as before we find that $\alpha y \notin K^{p^l}$, where $l := \lfloor \log_p c_{K,S} \rfloor + 1$. Then Theorem 3.2.3 gives

$$H_K(\alpha y) = H_K(\alpha z) \leq c_{K,S} \cdot (\omega(S) + c_{K,S} + 2g - 2).$$

The last case is $\alpha \in \mathcal{O}_S^*$. Suppose that $\alpha \in \mathbb{F}_q^*$. From Theorem 3.2.3 we deduce that

$$H_K(\alpha y) = H_K(\alpha z) \leq \omega(S) + 2g - 2.$$

So from now on we further assume that $\alpha \notin \mathbb{F}_q^*$. If $\alpha y \in \mathbb{F}_q^*$ or $\alpha z \in \mathbb{F}_q^*$, we immediately get a height bound for respectively y or z. So suppose that $\alpha y \notin \mathbb{F}_q^*$ and $\alpha z \notin \mathbb{F}_q$. Then there are $y', z' \notin K^p$ and $s \in \mathbb{Z}_{\geq 0}$ such that $y'^{p^s} = \alpha y$ and $z'^{p^s} = \alpha z$ and we get an equation

$$y' + z' = 1. \tag{3.16}$$

Applying Theorem 3.2.3 once more

$$H_K(y') = H_K(z') \leq \omega(S) + 2g - 2.$$

We conclude that

$$(x, y, z) \in \mathcal{F}_1((x, \alpha^{-1}, \alpha^{-1}), (1, y', z')).$$

A simple check using $x + \alpha^{-1} = 1$ and equation (3.16) shows that

$$x + \alpha^{-1}(y')^{p^f} + \alpha^{-1}(z')^{p^f} = 1$$

for all integers $f \geq 0$ and hence $((x, \alpha^{-1}, \alpha^{-1}), (1, y', z')) \in B_1$. This completes the proof. $\square$

3.4 Proof of Theorem 3.1.2

Define the set B_p' by

$$B_p' := \{(\mathbf{u}, \mathbf{v}) \in (\mathcal{O}_S^*)^3 \times (\mathcal{O}_S^*)^3 : \mathbf{u}, \mathbf{v} \notin (K^p)^3, (u_i, v_i) \notin \mathbb{F}_q^* \times \mathbb{F}_q^*, H_K(u_i) \leq c_{K,S},$$
$$H_K(v_i) \leq \omega(S) + 2g - 2, u_1 v_1^{p^f} + u_2 v_2^{p^f} + u_3 v_3^{p^f} = 1 \text{ for all } f \in \mathbb{Z}_{\geq 0}\}.$$

For the reader's convenience we recall that in the definition of B_p we only required that $\mathbf{u}, \mathbf{v} \notin (\mathbb{F}_q^*)^3$ instead of the stronger condition $\mathbf{u}, \mathbf{v} \notin (K^p)^3$. Nevertheless we have the equality

$$\bigcup_{(\mathbf{u},\mathbf{v}) \in B_p} \mathcal{F}_1(\mathbf{u}, \mathbf{v}) = \bigcup_{(\mathbf{u},\mathbf{v}) \in B'_p} \mathcal{F}_1(\mathbf{u}, \mathbf{v}). \tag{3.17}$$

To prove equality (3.17), we need two lemmata.

Lemma 3.4.1. *Let K be a field of characteristic $p > 0$ and let $n \geq 0$ be an integer. Suppose that $u_1, \ldots, u_n, v_1, \ldots, v_n \in K$ are such that*

$$u_1 v_1^{p^f} + \ldots + u_n v_n^{p^f} = 1$$

for all integers $f \geq 0$. Then $v_1, \ldots, v_n$ are linearly dependent over $\mathbb{F}_p$ or $u_1, \ldots, u_n \in \mathbb{F}_p$.

Proof. Define A to be the matrix

$$A := \begin{pmatrix} v_1 & v_2 & \ldots & v_n \\ v_1^p & v_2^p & \ldots & v_n^p \\ \vdots & & \vdots & \\ v_1^{p^{n-1}} & v_2^{p^{n-1}} & \ldots & v_n^{p^{n-1}} \end{pmatrix}.$$

It is a well-known fact that A is invertible if and only if $v_1, \ldots, v_n$ are linearly independent over $\mathbb{F}_p$; this can be proven by induction on n. We shall henceforth assume that A is invertible and prove that $u_1, \ldots, u_n \in \mathbb{F}_p$. But observe that

$$A \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad A \begin{pmatrix} u_1^{1/p} \\ \vdots \\ u_n^{1/p} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

This implies that $u_i = u_i^{1/p}$ for $i = 1, \ldots, n$, and the lemma follows. □

Lemma 3.4.2. *Let K be a field of characteristic $p > 0$ and let $n \geq 0$ be an integer. Suppose that $u_1, \ldots, u_n, v_1, \ldots, v_n \in K$ are such that*

$$u_1 v_1^{p^f} + \ldots + u_n v_n^{p^f} = 1$$

for all integers $f > 0$. Then we also have

$$u_1 v_1 + \ldots + u_n v_n = 1.$$

Proof. We proceed by induction on n with the base case $n = 0$ being trivial. We now apply Lemma 3.4.1. If $u_1, \ldots, u_n \in \mathbb{F}_p$, we clearly have

$$u_1 v_1 + \ldots + u_n v_n = 1$$

after taking p-th roots. So suppose that $v_1, \ldots, v_n$ are linearly dependent over $\mathbb{F}_p$. Without loss of generality we may write

$$v_n = \alpha_1 v_1 + \ldots + \alpha_{n-1} v_{n-1} \tag{3.18}$$

for some $\alpha_1, \ldots, \alpha_{n-1} \in \mathbb{F}_p$. Now substitution yields

$$(u_1 + \alpha_1 u_n) v_1^{p^f} + \ldots + (u_{n-1} + \alpha_{n-1} u_n) v_{n-1}^{p^f} = 1$$

for all integers $f > 0$. The induction hypothesis gives

$$(u_1 + \alpha_1 u_n) v_1 + \ldots + (u_{n-1} + \alpha_{n-1} u_n) v_{n-1} = 1. \tag{3.19}$$

Upon combining equation (3.18) with equation (3.19) we obtain the lemma. $\square$

Note that Lemma 3.4.2 with $n = 3$ readily implies the validity of equation (3.17). So our goal will be to give an upper bound for the cardinality of B'_p. Now let $(\mathbf{u}, \mathbf{v}) \in B'_p$. Then we know that

$$u_1 v_1^{p^f} + u_2 v_2^{p^f} + u_3 v_3^{p^f} = 1$$

for all $f \in \mathbb{Z}_{\geq 0}$. Since $\mathbf{u} \not\in (\mathbb{F}_p^*)^3$, an application of Lemma 3.4.1 shows that v_1, v_2, v_3 are indeed linearly dependent over $\mathbb{F}_p$. At the cost of multiplying our final bounds by 3, we may assume that

$$v_3 = \alpha_1 v_1 + \alpha_2 v_2$$

with $\alpha_1, \alpha_2 \in \mathbb{F}_p$. This yields

$$(u_1 + \alpha_1 u_3)\, v_1^{p^f} + (u_2 + \alpha_2 u_3)\, v_2^{p^f} = 1 \tag{3.20}$$

for all $f \in \mathbb{Z}_{\geq 0}$. Let us now apply Lemma 3.4.1 again. First suppose that v_1 and v_2 are linearly dependent over $\mathbb{F}_p$, we will show that this leads to a contradiction. Without loss of generality we may assume that $v_2 = \beta v_1$ for some $\beta \in \mathbb{F}_p$. Then we obtain

$$(u_1 + \alpha_1 u_3 + \beta u_2 + \beta \alpha_2 u_3) v_1^{p^f} = 1$$

for all $f \in \mathbb{Z}_{\geq 0}$, which implies $v_1 \in \mathbb{F}_p$. We deduce that $v_1, v_2, v_3 \in \mathbb{F}_p$, which is the desired contradiction. Hence Lemma 3.4.1 implies that

$$\lambda_1 := u_1 + \alpha_1 u_3 \in \mathbb{F}_p, \quad \lambda_2 := u_2 + \alpha_2 u_3 \in \mathbb{F}_p$$

and therefore $\lambda_1 v_1 + \lambda_2 v_2 = 1$. We claim that at most one of $\alpha_1, \alpha_2, \lambda_1, \lambda_2$ is equal to zero.

It is clear that α_1 and α_2 can not be simultaneously equal to zero, and the same holds for λ_1 and λ_2. If $\alpha_1 = \lambda_1 = 0$, we find that $u_1 = 0$, which contradicts $u_1 \in \mathcal{O}_S^*$. Now suppose that $\alpha_1 = \lambda_2 = 0$. In this case we deduce that $u_1, v_1 \in \mathbb{F}_p^*$, again contrary to our assumption $(\mathbf{u}, \mathbf{v}) \in B'_p$. The remaining two cases can be dealt with symmetrically, establishing our claim.

Let us first suppose that $\alpha_1, \alpha_2, \lambda_1, \lambda_2$ are all fixed and non-zero. Then we view the equations

$$\lambda_1 = u_1 + \alpha_1 u_3, \quad \lambda_2 = u_2 + \alpha_2 u_3, \quad \lambda_1 v_1 + \lambda_2 v_2 = 1$$

as unit equations to be solved in u_1, u_2, u_3, v_1, v_2. If one of the u_i is in K^p, then it turns out that all the u_i are in K^p, contradicting our assumption $\mathbf{u} \notin (K^p)^3$. Henceforth we may assume that $u_1, u_2, u_3 \notin K^p$ and similarly $v_1, v_2 \notin K^p$. Theorem 3.2.4 implies that there are at most $31 \cdot 19^{2|S|}$ solutions (u_1, u_3) to $\lambda_1 = u_1 + \alpha_1 u_3$ and at most $31 \cdot 19^{2|S|}$ solutions (v_1, v_2) to $\lambda_1 v_1 + \lambda_2 v_2 = 1$. Note that u_1 and u_3 determine u_2 and similarly v_1 and v_2 determine v_3. Hence there are at most $961 \cdot 19^{4|S|}$ possibilities for $(\mathbf{u}, \mathbf{v})$.

We will now treat the case $\lambda_2 = 0$ and $\alpha_1, \alpha_2, \lambda_1$ fixed and non-zero. In this case we can treat the unit equation

$$\lambda_1 = u_1 + \alpha_1 u_3$$

exactly as before; it has at most $31 \cdot 19^{2|S|}$ solutions (u_1, u_3). From $0 = \lambda_2 = u_2 + \alpha_2 u_3$ we see that u_2 is determined by u_1 and u_3. Note that $\lambda_2 = 0$ implies $\lambda_1 v_1 = 1$, i.e. $v_1 = \frac{1}{\lambda_1}$. We recall that

$$v_3 = \alpha_1 v_1 + \alpha_2 v_2$$

and therefore

$$v_3 = \frac{\alpha_1}{\lambda_1} + \alpha_2 v_2.$$

If $v_2 \in K^p$, then also $v_3 \in K^p$ and we conclude that $(v_1, v_2, v_3) \in (K^p)^3$. This is again a contradiction, so suppose that $v_2, v_3 \notin K^p$. We are now in the position to apply Theorem 3.2.4, which shows that there are at most $31 \cdot 19^{2|S|}$ solutions (v_2, v_3). Hence there are at most $961 \cdot 19^{4|S|}$ possibilities for $(\mathbf{u}, \mathbf{v})$.

Finally we will treat the case $\alpha_2 = 0$ and $\alpha_1, \lambda_1, \lambda_2$ still fixed and non-zero. We remark that the remaining two cases $\lambda_1 = 0$ and $\alpha_1 = 0$ can be dealt with using the same argument as the case $\lambda_2 = 0$ and $\alpha_2 = 0$ respectively. Note that $u_2 = \lambda_2 \in \mathbb{F}_p^*$. Using $\lambda_1 = u_1 + \alpha_1 u_3$ and $\mathbf{u} \notin (K^p)^3$, we deduce that $u_1, u_3 \notin K^p$. Hence the unit equation

$$\lambda_1 = u_1 + \alpha_1 u_3$$

has at most $31 \cdot 19^{2|S|}$ solutions (u_1, u_3). Similarly, the unit equation

$$\lambda_1 v_1 + \lambda_2 v_2 = 1$$

has at most $31 \cdot 19^{2|S|}$ solutions (v_1, v_2). Since v_1 determines v_3, we have proven that there are also at most $961 \cdot 19^{4|S|}$ possibilities for $(\mathbf{u}, \mathbf{v})$ in this case.

So far we have treated $\alpha_1, \alpha_2, \lambda_1, \lambda_2$ as fixed. To every element of B'_p we can attach a tuple $\mathbf{t} = (\alpha_1, \alpha_2, \lambda_1, \lambda_2)$. Clearly there are at most p^4 such tuples. Furthermore, we have shown that for each fixed tuple $\mathbf{t}$ there are at most $961 \cdot 19^{4|S|}$ $(\mathbf{u}, \mathbf{v}) \in B'_p$ that correspond to $\mathbf{t}$. Altogether we have proven that $|B'_p| \leq 3 \cdot 961 \cdot p^4 \cdot 19^{4|S|} = 2883 \cdot p^4 \cdot 19^{4|S|}$.

To deal with B_q one can use a very similar approach, so we will only sketch the proof. In this case we define

$$B'_q := \{(\mathbf{u}, \mathbf{v}) \in (\mathcal{O}_S^*)^3 \times (\mathcal{O}_S^*)^3 : \mathbf{u}, \mathbf{v} \notin (K^q)^3, (u_i, v_i) \notin \mathbb{F}_q^* \times \mathbb{F}_q^*, H_K(u_i) \leq c_{K,S},$$
$$H_K(v_i) \leq \frac{q}{p} (\omega(S) + 2g - 2), u_1 v_1^{q^f} + u_2 v_2^{q^f} + u_3 v_3^{q^f} = 1 \text{ for all } f \in \mathbb{Z}_{\geq 0}\}.$$

Note that we now only require that $\mathbf{u}, \mathbf{v} \notin (K^q)^3$ instead of $\mathbf{u}, \mathbf{v} \notin (K^p)^3$. In our new setting we find that $\alpha_1, \alpha_2, \lambda_1, \lambda_2 \in \mathbb{F}_q$ instead of $\alpha_1, \alpha_2, \lambda_1, \lambda_2 \in \mathbb{F}_p$. This means that we have q^4 tuples $(\alpha_1, \alpha_2, \lambda_1, \lambda_2)$. For each fixed tuple $\mathbf{t}$ there are at most $\log_p(q) \cdot 961 \cdot 19^{4|S|}$ $(\mathbf{u}, \mathbf{v}) \in B'_q$ that can map to $\mathbf{t}$. The extra factor $\log_p(q)$ comes from the fact that we merely know that $\mathbf{u}, \mathbf{v} \notin (K^q)^3$ when we apply Theorem 3.2.4. We conclude that $|B'_q| \leq 2883 \cdot \log_p(q) \cdot q^4 \cdot 19^{4|S|}$.

Our only remaining task is to bound $|A|$. We start by recalling a “gap principle”. Define

$$\mathcal{S} := \{(x_0 : x_1 : x_2 : x_3) \in \mathbb{P}^3(K) \setminus \mathbb{P}^3(\mathbb{F}_q) : x_0 + x_1 + x_2 = x_3,$$
$$v(x_0) = v(x_1) = v(x_2) = v(x_3) \text{ for every } v \in M_K \setminus S\}.$$

Then we have the following lemma.

Lemma 3.4.3 (Gap principle)**.** *Let B be a real number with $\frac{3}{4} < B < 1$, and let $P > 0$. Then the set of projective points $(x_0 : x_1 : x_2 : x_3)$ of $\mathcal{S}$ with*

$$P \leq H_K(x_0 : x_1 : x_2 : x_3) < \left(1 + \frac{4B - 3}{2}\right) P$$

is contained in the union of at most $4^{|S|}(e/(1 - B))^{3|S|-1}$ 1-dimensional projective subspaces of $x_0 + x_1 + x_2 = x_3$. Here e is the Euler constant.

Proof. This was proved in [18] for function fields in characteristic 0, but the proof works ad verbatim in characteristic p. □

Take any $P > 0$ and suppose that $(x, y, z) \in A$ is a solution to

$$x + y + z = 1$$

with $P \leq H_K(x : y : z : 1) < \left(1 + \frac{4B-3}{2}\right) P$. Then we can apply Lemma 3.4.3 to deduce that $(x : y : z : 1)$ is contained in some 1-dimensional projective subspace. This means that x, y, z satisfy an additional equation

$$ax + by + cz = d \tag{3.21}$$

for some $a, b, c, d \in K$, such that the equation is independent from the original equation $x + y + z = 1$. By adding the equation $x + y + z = 1$ to equation (3.21) if necessary, we can ensure that $a \neq 0$. We have

$$(a - b)y + (a - c)z = a - d. \tag{3.22}$$

If $a-b$, $a-c$ and $a-d$ are zero, we conclude that $a=b=c=d$. This is a contradiction, since we assumed that the equation $ax+by+cz=d$ was linearly independent from the equation $x+y+z=1$. If only one of $a-b$, $a-c$ and $a-d$ is not zero, we find that $y=0$, $z=0$ and $0=a-d\neq 0$ respectively, so we obtain a contradiction in every case. At the cost of multiplying our final bounds by 3, we may assume that $a-b\neq 0$. We will distinguish three cases.

Case I: $a-c\neq 0$, $a-d\neq 0$. In this case we view (3.22) as a unit equation. Since $(x,y,z)\in A$, it follows that $H_K(x), H_K(y), H_K(z)\leq c'_{K,S}$. We conclude that

$$H_K\left(\frac{a-b}{a-d}y\right)\in\left[H_K\left(\frac{a-b}{a-d}\right)-c'_{K,S}, H_K\left(\frac{a-b}{a-d}\right)+c'_{K,S}\right]. \tag{3.23}$$

If $(a-b)/(a-d)\notin\mathcal{O}_S^*$ or $(a-c)/(a-d)\notin\mathcal{O}_S^*$, we use Theorem 1.1 from [44]. In this case we see that there are at most $31\cdot 19^{2|S|-1}$ solutions. Now suppose that $(a-b)/(a-d)\in\mathcal{O}_S^*$ and $(a-c)/(a-d)\in\mathcal{O}_S^*$. Then we have

$$\frac{a-b}{a-d}y\in\mathbb{F}_q \quad \text{or} \quad y=\frac{a-d}{a-b}\cdot v^{p^z} \text{ for some } z\in\mathbb{Z}_{\geq 0}$$

with $v\notin(\mathcal{O}_S^*)^p$. The first case gives at most q^2 solutions. To treat the second case, we remark that there are at most $31\cdot 19^{2|S|-2}$ values of v due to Theorem 3.2.4. Furthermore, for fixed v, there are at most $\log_p(2c'_{K,S})+1$ choices of z by equation (3.23). Hence there are at most

$$q^2+(\log_p(2c'_{K,S})+1)\cdot 31\cdot 19^{2|S|}$$

solutions (y,z) to equation (3.22). From $x+y+z=1$ we see that y and z determine x.

We will now count the total contribution to the number of solutions from case I. Choose $B:=\frac{7}{8}$. Note that

$$H_K(x:y:z:1)\leq H_K(x)+H_K(y)+H_K(z)\leq 3c'_{K,S}.$$

Now define $l:=\log_{\frac{5}{4}}(3c'_{K,S})+1$. Then for every solution $(x,y,z)\in A$ there is i with $0\leq i<l$ such that

$$\left(\frac{5}{4}\right)^i\leq H_K(x:y:z:1)<\left(\frac{5}{4}\right)^{i+1}.$$

For fixed i every solution $(x:y:z:1)$ is contained in the union of at most $(2048e^3)^{|S|}$ 1-dimensional projective subspaces. Furthermore, we have just shown that each subspace contains at most $q^2+(\log_p(2c'_{K,S})+1)\cdot 31\cdot 19^{2|S|}$ solutions. This gives as total bound for A in case I

$$\begin{aligned}|A|&\leq(\log_{\frac{5}{4}}(3c'_{K,S})+1)\cdot(2048e^3)^{|S|}\cdot q^2\cdot(\log_p(2c'_{K,S})+1)\cdot 31\cdot 19^{2|S|}\\&\leq 31q^2\cdot(\log_{\frac{5}{4}}(3c'_{K,S})+1)^2\cdot(15\cdot 10^6)^{|S|}.\end{aligned} \tag{3.24}$$

Case II: $a-c\neq 0$, $a-d=0$. In this case (3.22) gives

$$z=-\frac{a-b}{a-c}y.$$

Substitution in $x + y + z = 1$ yields

$$x + \left(1 - \frac{a-b}{a-c}\right) y = 1. \tag{3.25}$$

If $a - b = a - c$, we see that $x = 1$, contrary to our assumption $x \notin \mathbb{F}_q$. So we will assume that $a - b \neq a - c$ and treat (3.25) as a unit equation. Then, following the proof of case I, we get the bound (3.24) for A in case II.

Case III: $a - c = 0$, $a - d \neq 0$. From (3.22) we deduce that

$$y = \frac{a-d}{a-b}.$$

If $a - b = a - d$, we conclude that $y = 1$, which is again a contradiction. Substitution in $x + y + z = 1$ gives

$$x + z = 1 - \frac{a-d}{a-b}. \tag{3.26}$$

Note that (3.26) is another unit equation and, just as before, we obtain the bound (3.24) for A in case III.

3.5 Application to Fermat surfaces

The goal of this section is to prove Theorem 3.1.4. We start off with a definition.

Definition 3.5.1. We say that a valuation v of K is D-generic if the following two conditions are satisfied

- first of all
$$v\left(\frac{Dx}{x}\right) = -1$$
for all $x \in K^*$ satisfying $p \nmid v(x)$;
- and secondly
$$v\left(\frac{Dx}{x}\right) \geq 0$$
for all $x \in K^*$ with $p \mid v(x)$.

In $\mathbb{F}_p(t)$ and D differentiation with respect to t, every valuation is D-generic except for the infinite valuation. In general only finitely many valuations are not generic.

In this section K and D will always be equal to respectively $\mathbb{F}_p(t)$ and differentiation with respect to t. Whenever we say that v is generic, we will mean generic with respect to this

D. Let N be a $(480, p)$-good integer coprime to p. In particular we have that $N > 480$, which we shall use at several points during the proof. Suppose that $x, y, z \in \mathbb{F}_p(t)$ is a solution to

$$x^N + y^N + z^N = 1 \tag{3.27}$$

satisfying the conditions of Theorem 3.1.4, i.e. $x, y, z, x/y, x/z, y/z \notin \mathbb{F}_p(t^p)$. By Lemma 3.2.1 this is equivalent to $\frac{Dx}{x} \neq 0$, $\frac{Dy}{y} \neq 0$, $\frac{Dz}{z} \neq 0$, $\frac{Dx}{x} \neq \frac{Dy}{y}$, $\frac{Dx}{x} \neq \frac{Dz}{z}$ and $\frac{Dy}{y} \neq \frac{Dz}{z}$. Then differentiation with respect to D yields

$$x^N \cdot \frac{NDx}{x} + y^N \cdot \frac{NDy}{y} + z^N \cdot \frac{NDz}{z} = 0,$$

and using that $(N, p) = 1$

$$x^N \cdot \frac{Dx}{x} + y^N \cdot \frac{Dy}{y} + z^N \cdot \frac{Dz}{z} = 0. \tag{3.28}$$

We multiply equation (3.28) with $\frac{z}{Dz}$ and subtract it from equation (3.27) to obtain

$$x^N \left(1 - \frac{z}{Dz}\frac{Dx}{x}\right) + y^N \left(1 - \frac{z}{Dz}\frac{Dy}{y}\right) = 1 \tag{3.29}$$

and similarly

$$x^N \left(1 - \frac{y}{Dy}\frac{Dx}{x}\right) + z^N \left(1 - \frac{y}{Dy}\frac{Dz}{z}\right) = 1. \tag{3.30}$$

Define

$$S := \{v \in M_K : v(x) \neq 0 \text{ or } v(y) \neq 0 \text{ or } v(z) \neq 0\}.$$

We may assume that x is such that

$$\omega(x) \geq \frac{\omega(S)}{3}. \tag{3.31}$$

If $N > 12$, thanks to Lemma 3.2.2 applied with $K = \mathbb{F}_p(t)$, we have

$$H_K(x^N) = NH_K(x) > 6\omega(x) \geq 2\omega(S) \geq H_K\left(1 - \frac{z}{Dz}\frac{Dx}{x}\right)$$

and similarly

$$H_K(x^N) > H_K\left(1 - \frac{y}{Dy}\frac{Dx}{x}\right).$$

Hence $x^N\left(1 - \frac{z}{Dz}\frac{Dx}{x}\right), x^N\left(1 - \frac{y}{Dy}\frac{Dx}{x}\right) \notin \mathbb{F}_p$ and therefore we can write

$$x^N\left(1 - \frac{z}{Dz}\frac{Dx}{x}\right) = \delta^{p^s} \tag{3.32}$$

$$x^N\left(1 - \frac{y}{Dy}\frac{Dx}{x}\right) = \epsilon^{p^r} \tag{3.33}$$

with $\delta, \epsilon \notin \mathbb{F}_p(t^p)$. Now we claim that for $N > 48$

$$\omega(\delta) \geq \frac{\omega(S)}{4}. \tag{3.34}$$

Indeed suppose for the sake of contradiction that $\omega(\delta) < \frac{\omega(S)}{4}$. Using equation (3.31) we find that there is a finite subset T of M_K with $\omega(T) \geq \frac{\omega(S)}{12}$ such that for all $v \in T$ we have $v(x) \neq 0$ and $v(\delta) = 0$. For such a valuation $v \in T$ we have due to equation (3.32)

$$v\left(1 - \frac{z}{Dz}\frac{Dx}{x}\right) = -Nv(x) \neq 0.$$

This implies that

$$4\omega(S) \geq 2H_K\left(1 - \frac{z}{Dz}\frac{Dx}{x}\right) \geq \sum_{v \in T}\left|v\left(1 - \frac{z}{Dz}\frac{Dx}{x}\right)\right| \deg(v) \geq N\frac{\omega(S)}{12}.$$

This is impossible for $N > 48$, so we have established (3.34). For convenience we define for a valuation v and $a, b \notin \mathbb{F}_p(t^p)$

$$f_v(a, b) := \left|v\left(1 - \frac{a}{Da}\frac{Db}{b}\right)\right|,$$

$$\begin{aligned} g_v(x, y, z) := |v(\delta)| + |v(\epsilon)| + f_v(x, y) + f_v(y, x) + \\ + f_v(x, z) + f_v(z, x) + f_v(y, z) + f_v(z, y). \end{aligned}$$

Our next claim is that there is a generic place $v \in M_K$ such that $v(\delta) \neq 0$ and

$$g_v(x, y, z) \leq 480. \tag{3.35}$$

Lemma 3.2.2 with $K = \mathbb{F}_p(t)$ shows that

$$\sum_{v \in M_K} f_v(x, y) \deg v = 2H_K\left(1 - \frac{x}{Dx}\frac{Dy}{y}\right) \leq 2\left(H_K\left(\frac{Dx}{x}\right) + H_K\left(\frac{Dy}{y}\right)\right) \leq 4\omega(S) \tag{3.36}$$

and similarly for the other f_v. Equation (3.29) and equation (3.32) combined with equation (3.36) show that

$$\sum_{\substack{v \in M_K \\ v(\delta) \neq 0 \text{ or } v(1-\delta) \neq 0}} \deg v \leq \omega(S) + \omega\left(1 - \frac{z}{Dz}\frac{Dx}{x}\right) + \omega\left(1 - \frac{z}{Dz}\frac{Dy}{y}\right) \leq 9\omega(S).$$

Indeed, if $v(\delta) \neq 0$, we have $v(x) \neq 0$, so $v \in S$, or $v\left(1 - \frac{z}{Dz}\frac{Dx}{x}\right) \neq 0$, while if $v(1 - \delta) \neq 0$, we have $v(y) \neq 0$, hence $v \in S$, or $v\left(1 - \frac{z}{Dz}\frac{Dy}{y}\right) \neq 0$. Similarly, equation (3.30) and equation (3.33) yield

$$\sum_{\substack{v \in M_K \\ v(\epsilon) \neq 0 \text{ or } v(1-\epsilon) \neq 0}} \deg v \leq 9\omega(S).$$

Then Theorem 3.2.3 gives

$$\sum_{v \in M_K} |v(\delta)| \deg v = 2H_K(\delta) \leq 18\omega(S) \tag{3.37}$$

and the same for $|v(\epsilon)|$. Hence we have

$$\sum_{\substack{v \in M_K \\ v(\delta) \neq 0}} g_v(x, y, z) \deg(v) \leq \sum_{v \in M_K} g_v(x, y, z) \deg(v) \leq 60\omega(S)$$

by equation (3.36) and equation (3.37). Note that there are at least two places such that $v(\delta) \neq 0$, so there is at least one generic place v such that $v(\delta) \neq 0$. Hence if $\omega(S) \leq 8$, (3.35) follows immediately. So suppose that $\omega(S) > 8$. Using (3.34) we conclude that

$$\begin{aligned}
\frac{\omega(S)}{8} \min_{\substack{v \in M_K \\ v(\delta) \neq 0 \\ v \text{ generic}}} g_v(x, y, z) &\leq \left(\frac{\omega(S)}{4} - 1\right) \min_{\substack{v \in M_K \\ v(\delta) \neq 0 \\ v \text{ generic}}} g_v(x, y, z) \\
&\leq (\omega(\delta) - 1) \min_{\substack{v \in M_K \\ v(\delta) \neq 0 \\ v \text{ generic}}} g_v(x, y, z) \\
&\leq 60\omega(S),
\end{aligned}$$

thus proving our claim, i.e. equation (3.35). From now on fix a generic $v \in M_K$ satisfying $v(\delta) \neq 0$ and equation (3.35). Note that equation (3.32) yields the following equality

$$v\left(1 - \frac{z}{Dz}\frac{Dx}{x}\right) + Nv(x) = p^s v(\delta). \tag{3.38}$$

We will next show that $s > 0$ and $r > 0$. Suppose not. Then we may assume that $s = 0$ by symmetry considerations. Equation (3.31) and (3.32) give

$$\frac{N\omega(S)}{6} \leq NH_K(x) \leq H_K(\delta) + H_K\left(1 - \frac{z}{Dz}\frac{Dx}{x}\right) \leq 11\omega(S),$$

where the last inequality follows from equation (3.36) and equation (3.37). If $N > 480$, this gives us the desired contradiction, so henceforth we may assume that $s, r > 0$.

If $p > 480$, we find that $v(x) \neq 0$ due to equation (3.38) and $s > 0$. We claim that

$$v\left(1 - \frac{z}{Dz}\frac{Dx}{x}\right) \neq 0. \tag{3.39}$$

Assume the contrary. Then equation (3.38) implies that N divides $v(\delta) \neq 0$, but this is impossible by construction of v and the fact that $N > 480$ thus establishing equation (3.39). Finally observe that

$$N \mid p^s v(\delta) - v\left(1 - \frac{z}{Dz}\frac{Dx}{x}\right).$$

We now distinguish two cases. First suppose that $v(\delta) > 0$. Then clearly also $v(x) > 0$. If furthermore $v\left(1 - \frac{z}{Dz}\frac{Dx}{x}\right) < 0$, we get that N divides $ap^s + b$ with $0 < a, b \leq 480$ contrary to our assumptions. Due to equation (3.39) we are left with the case

$$v\left(1 - \frac{z}{Dz}\frac{Dx}{x}\right) > 0. \tag{3.40}$$

Now comes the crucial observation that $p \nmid v(x)$. Indeed, otherwise we find by equation (3.38)

$$p \mid v\left(1 - \frac{z}{Dz}\frac{Dx}{x}\right),$$

which is not possible due to $p > 480$, equation (3.35) and equation (3.40). Hence we deduce for a generic valuation v that $v\left(\frac{Dx}{x}\right) = -1$. Combining this with equation (3.40) again we get that $v(z) \neq 0$. Equation (3.33) gives the equality

$$v\left(1 - \frac{y}{Dy}\frac{Dx}{x}\right) + Nv(x) = p^r v(\epsilon).$$

Recall that $v(x) > 0$, hence $v(\epsilon) > 0$. Using equation (3.30) and equation (3.33), we get

$$z^N\left(1 - \frac{y}{Dy}\frac{Dz}{z}\right) = (1 - \epsilon)^{p^r}.$$

Since $v(1 - \epsilon) = 0$, this shows

$$v\left(1 - \frac{y}{Dy}\frac{Dz}{z}\right) + Nv(z) = 0,$$

which is a contradiction for $N > 480$.

We still need to treat the case $v(\delta) < 0$. In that case we find that $v(x) < 0$ and $v\left(1 - \frac{z}{Dz}\frac{Dx}{x}\right) < 0$. Similarly as before we can show that this implies $p \mid v(z)$ for a generic valuation v. We will use equation (3.30) and equation (3.33) once more to deduce that

$$z^N\left(1 - \frac{y}{Dy}\frac{Dz}{z}\right) = (1 - \epsilon)^{p^r}.$$

Since $v(x) < 0$ implies that $v(\epsilon) < 0$, we find that

$$v\left(1 - \frac{y}{Dy}\frac{Dz}{z}\right) + Nv(z) = p^r v(1 - \epsilon) = p^r v(\epsilon). \tag{3.41}$$

Combining (3.41) with $p \mid v(z)$ we get that

$$p \mid v\left(1 - \frac{y}{Dy}\frac{Dz}{z}\right).$$

If $p > 480$, then (3.35) implies that $v\left(1 - \frac{y}{Dy}\frac{Dz}{z}\right) = 0$. Hence (3.41) gives $N \mid v(\epsilon)$. Using (3.35) and $N > 480$ once more we conclude that $v(\epsilon) = 0$, which is the desired contradiction.

3.6 Curves inside Fermat surfaces

The goal of this section is to show that Theorem 3.1.4 becomes false if we allow x, y, z, x/y, x/z or y/z to be in $\mathbb{F}_p(t^p)$. By symmetry it suffices to do this in the case x or y/z in $\mathbb{F}_p(t^p)$. We will do this by exhibiting explicit curves inside the Fermat surface.

Let us start by allowing $y/z \in \mathbb{F}_p(t^p)$. We can rewrite

$$x^N + y^N + z^N = 1$$

as

$$\frac{1}{1-x^N} y^N + \frac{1}{1-x^N} z^N = 1.$$

Then if N is odd, we have

$$\frac{1}{1-x^N} y^N + \frac{-x^N}{1-x^N} \frac{(-z)^N}{x^N} = 1.$$

The key point is that we can now put $\alpha := \frac{1}{1-x^N}$, $\tilde{z} = \frac{-z}{x}$, after which the last equation can be rewritten as

$$\alpha y^N + (1-\alpha)\tilde{z}^N = 1. \tag{3.42}$$

But it is rather straightforward to find solutions to this last equation. Indeed, we know that $N \mid p^k - 1$ for some $k > 0$. For such a k we put

$$y := \alpha^{\frac{p^k-1}{N}}, \tilde{z} := (1-\alpha)^{\frac{p^k-1}{N}},$$

and one easily verifies that y and $\tilde{z}$ satisfy (3.42). Going back to our original variables x, y and z we get that

$$y := \left(\frac{1}{1-x^N}\right)^{\frac{p^k-1}{N}}, z := -x\left(\frac{-x^N}{1-x^N}\right)^{\frac{p^k-1}{N}}.$$

There are two important remarks to make about the above construction. First of all, it is easily verified that $y/z \in \mathbb{F}_p(t^p)$ as we claimed. Secondly, we used that N is odd during our construction. However, we only need that -1 is an N-th power in $\mathbb{F}_p^*$.

Now suppose that $x \in \mathbb{F}_p(t^p)$. For simplicity we will again assume that N is odd. Then from the equation

$$x^N + y^N + z^N = 1$$

we find that

$$\left(\frac{1}{z}\right)^N + \left(\frac{-x}{z}\right)^N + \left(\frac{-y}{z}\right)^N = 1.$$

After putting $\tilde{x} = \frac{-y}{z}$, $\tilde{y} = \frac{-x}{z}$ and $\tilde{z} = \frac{1}{z}$ we get that

$$\tilde{x}^N + \tilde{y}^N + \tilde{z}^N = 1$$

with $\frac{\tilde{y}}{\tilde{z}} = -x \in \mathbb{F}_p(t^p)$. Hence we can apply the previous construction.

Finally we will explain why we need the condition that N is $(480, p)$-good. If $N = p^r + 1$ for some $r \geq 0$, it is possible to write down non-trivial lines on the Fermat surface, see Section 5.1-5.4 of [63]. It turns out that our method is unable to distinguish between the case $N = p^r + 1$ and $N = ap^r + b$ with $0 < a, b$ small. This may seem strange at first, but it is in fact quite natural.

Indeed, let us compare this with the situation in characteristic 0. In this case it follows from the work of Voloch [77] that for N sufficiently large the equation

$$x^N + y^N + z^N = 1$$

has no non-constant solutions $x, y, z \in \mathbb{C}(t)$. In fact, this is a rather easy consequence from his abc Theorem. However, it is a more difficult task to find the smallest N using abc Theorems, see for example [13]. Our Theorem 3.1.4 is also based on abc type arguments and for this reason it should not be surprising that we can not distinguish between the case $N = p^r + 1$, giving unirational surfaces [63], and $N = ap^r + b$ with $0 < a, b$ small.

Thus, morally, the notion of N being $(480, p)$-good in Theorem 3.6 can be interpreted as saying that N is "far enough" from an exponent that gives a unirational surface. In the proof we use this condition when we analyze the two dimensional Frobenius families. It is therefore instructive to notice here that there is a partial converse. Namely, we can use the description given at the beginning of Section 3.4 to produce non-trivial rational curves on Fermat surfaces. We will assume $p \equiv 1 \mod 4$ for simplicity: a similar computation can be carried out for the case $p \equiv 3 \mod 4$.

We will use the notation of Section 3.4. Rename $\tilde{\alpha_1} = \frac{\alpha_1}{\alpha_3}$ and $\tilde{\alpha_2} = \frac{\alpha_2}{\alpha_3}$. Choose $\tilde{\alpha_1}, \tilde{\alpha_2} \neq 0$ such that

$$\tilde{\alpha_1}^2 + \tilde{\alpha_2}^2 = -1$$

and put $\lambda_1 = i\tilde{\alpha_2}$ and $\lambda_2 = i\tilde{\alpha_1}$, where i is an element of $\mathbb{F}_p$ such that $i^2 = -1$. We further impose the conditions

$$u_1 = v_1, u_2 = v_2, u_3 = v_3.$$

With these choices, one can check that all the relevant equations in Section 3.4 are satisfied for $(v_1, v_2, v_3) = (\tilde{\alpha_1}t - i\tilde{\alpha_2}, \tilde{\alpha_2}t + i\tilde{\alpha_1}, t)$. Thus, since all the implications at the beginning of 3.4 are reversible, one deduces that the line $(\tilde{\alpha_1}t - i\tilde{\alpha_2}, \tilde{\alpha_2}t + i\tilde{\alpha_1}, t)$ is contained in *all* Fermat surfaces $x^{p^s+1} + y^{p^s+1} + z^{p^s+1} = 1$. Alternatively, one may directly verify that this yields lines on Fermat surfaces.

We conclude by remarking that the height bound in Theorem 3.1.1 *can not* be improved to a *linear* height bound in $\omega(S)$. Indeed, this follows easily by using the curves we constructed at the beginning of this section. A natural question is whether the quadratic dependency on $\omega(S)$ is sharp.

3.7 Acknowledgements

We thank Jan-Hendrik Evertse for giving us this problem, useful discussions and proof-reading. We would also like to thank Hendrik Lenstra and Ronald van Luijk for useful discussions. Finally we much appreciate the valuable remarks of the anonymous referee, which greatly improved the readability of the chapter.

Chapter 4

On the 16-rank of class groups of $\mathbb{Q}(\sqrt{-2p})$ for primes $p \equiv 1 \bmod 4$[1]

Joint work with Djordjo Milovic

Abstract

We use Vinogradov's method to prove equidistribution of a spin symbol governing the 16-rank of class groups of quadratic number fields $\mathbb{Q}(\sqrt{-2p})$, where $p \equiv 1 \bmod 4$ is a prime.

4.1 Introduction

Recently, the authors have used Vinogradov's method to prove density results about elements of order 16 in class groups in certain *thin* families of quadratic number fields parametrized by a single prime number, namely the families $\{\mathbb{Q}(\sqrt{-2p})\}_{p\equiv -1 \bmod 4}$ and $\{\mathbb{Q}(\sqrt{-p})\}_p$ [58, 41]. In this chapter, we establish a density result for the family $\{\mathbb{Q}(\sqrt{-2p})\}_{p\equiv 1 \bmod 4}$, thereby completing the picture for the 16-rank in families of imaginary quadratic fields with cyclic 2-class groups and even discriminant. Although our overarching methods are similar to those originally developed in the work of Friedlander et al. [24], the technical difficulties in the present case are different and require a more careful study of the spin symbols governing the 16-rank. The main distinguishing feature of the present work is that this careful study allows us to avoid relying on a conjecture about short character sums appearing in [24, 41], thus making our results unconditional.

[1] A slightly modified version of this chapter will appear in International Mathematics Research Notices.

More generally, given a sequence of complex numbers $\{a_n\}_n$ indexed by natural numbers, a problem of interest in analytic number theory is to prove an asymptotic formula for the sum over primes

$$S(X) := \sum_{\substack{p \text{ prime} \\ p \leq X}} a_p$$

as $X \to \infty$. Many sequences $\{a_n\}_n$ admit asymptotic formulas for $S(X)$ via various generalizations of the Prime Number Theorem, with essentially the best known error terms coming from ideas of de la Valée Poussin already in 1899 [15]. In 1947, Vinogradov [75, 76] invented another method to treat certain sequences which could not be handled with a variant of the Prime Number Theorem. His method has since been clarified and made easier to apply, most notably by Vaughan [73] and, for applications relating to more general number fields, by Friedlander et al. [24]. Nonetheless, there is a relative paucity of interesting sequences $\{a_n\}_n$ that admit an asymptotic formula for $S(X)$ via Vinogradov's method. The purpose of this chapter is to present yet another such sequence, of a similar nature as those appearing in [24, 41]; similarly as in [41], the asymptotics we obtain have implications in the arithmetic statistics of class groups of number fields.

Let $p \equiv 1 \bmod 4$ be a prime number, and let $\mathrm{Cl}(-8p)$ denote the class group of the quadratic number field $\mathbb{Q}(\sqrt{-2p})$ of discriminant $-8p$. The finite abelian group $\mathrm{Cl}(-8p)$ measures the failure of unique factorization in the ring $\mathbb{Z}[\sqrt{-2p}]$. By Gauss's genus theory [26], the 2-part of $\mathrm{Cl}(-8p)$ is cyclic and non-trivial, and hence determined by the largest power of 2 dividing the order of $\mathrm{Cl}(-8p)$. For each integer $k \geq 1$, we define a density $\delta(2^k)$, if it exists, as

$$\delta(2^k) := \lim_{X \to \infty} \frac{\#\{p \leq X : \ p \equiv 1 \bmod 4, \ 2^k | \#\mathrm{Cl}(-8p)\}}{\#\{p \leq X : \ p \equiv 1 \bmod 4\}}.$$

As stated above, the 2-part of $\mathrm{Cl}(-8p)$ is cyclic and non-trivial, so $\delta(2) = 1$. Rédei [62] proved that $4 | \#\mathrm{Cl}(-8p)$ if and only if p splits completely in $\mathbb{Q}(\zeta_8)$, and Stevenhagen [71] proved that $8 | \#\mathrm{Cl}(-8p)$ if and only if p splits completely in $\mathbb{Q}(\zeta_8, \sqrt[4]{2})$, where ζ_8 denotes a primitive 8th root of unity. It follows from these results and the Chebotarev Density Theorem (a generalization of the Prime Number Theorem) that $\delta(4) = \frac{1}{2}$ and $\delta(8) = \frac{1}{4}$. The qualitative behavior of divisibility by 16 departs from that of divisibility by lower 2-powers in that it can no longer be proved by a simple application of the Chebotarev Density Theorem. We instead use Vinogradov's method to prove

Theorem 4.1.1. *For a prime number $p \equiv 1 \bmod 4$, let $e_p = 0$ if $\mathrm{Cl}(-8p)$ does not have an element of order 8, let $e_p = 1$ if $\mathrm{Cl}(-8p)$ has an element of order 16, and let $e_p = -1$ otherwise. Then for all $X > 0$, we have*

$$\sum_{\substack{p \leq X \\ p \equiv 1 \bmod 4}} e_p \ll X^{1 - \frac{1}{3200}},$$

where the implied constant is absolute. In particular, $\delta(16) = \frac{1}{8}$.

In combination with [58], we get

Corollary 4.1.2. *For a prime number p, let $h_2(-2p)$ denote the cardinality of the 2-part of the class group* Cl($-8p$). *For an integer $k \geq 0$, let $\delta'(2^k)$ denote the natural density (in the set of all primes) of primes p such that $h_2(-2p) = 2^k$, if it exists. Then $\delta'(1) = 0$, $\delta'(2) = \frac{1}{2}$, $\delta'(4) = \frac{1}{4}$, and $\delta'(8) = \frac{1}{8}$.*

The power-saving bound in Theorem 4.1.1, similarly to the main results in [58] and [41], is another piece of evidence that *governing fields* for the 16-rank do *not* exist. For a sampling on previous work about governing fields, see [11], [12], [61], and [70].

The strategy to prove Theorem 4.1.1 is to construct a sequence $\{a_n\}_n$ which simultaneously carries arithmetic information about divisibility by 16 when n is a prime number congruent to 1 modulo 4 and is conducive to Vinogradov's method. On one hand, the criterion for divisibility by 16 cannot be stated naturally over the rational numbers $\mathbb{Q}$. For instance, even the criterion for divisibility by 8 is most naturally stated over a field of degree 8 over $\mathbb{Q}$. On the other hand, proving analytic estimates in a number field generally becomes more difficult as the degree of the number field increases, as exemplified by the reliance on a conjecture on short character sums in [24]. We manage to work over $\mathbb{Q}(\zeta_8)$, a field of degree 4. Although the methods of Friedlander et al. [24] narrowly miss the mark of being unconditional for number fields of degree 4, we manage to exploit the arithmetic structure of our sequence to ensure that Theorem 4.1.1 is unconditional.

Lastly, for work concerning the average behavior of the 2-parts of class groups of quadratic number fields in families that are *not* thin, i.e., for which the average number of primes dividing the discriminant grows as the discriminant grows, we point the reader to the extensive work of Fouvry and Klüners [19, 20, 21, 22] on the 4-rank and certain cases of the 8-rank and more recently to the work of Smith on the 8- and higher 2-power-ranks [68, 69]. While Smith's methods in [69] appear to be very powerful, the authors believe that they are unlikely to be applicable to thin families of the type appearing in this chapter.

Funding

This work was supported by the National Science Foundation [DMS-1128155 to D.Z.M.].

Acknowledgements

The authors thank Jan-Hendrik Evertse and Carlo Pagano for useful discussions.

4.2 Encoding the 16-rank of Cl($-8p$)

Given an integer $k \geq 1$, the 2^k-rank of a finite abelian group G, denoted by $\mathrm{rk}_{2^k} G$, is defined as the dimension of the $\mathbb{F}_2$-vector space $2^{k-1}G/2^kG$. If the 2-part of G is cyclic, then $\mathrm{rk}_{2^k} G \in \{0, 1\}$, and $\mathrm{rk}_{2^k} G = 1$ if and only if $2^k | \#G$. The order of a class group is called the class number, and we denote the class number of $\mathrm{Cl}(-8p)$ by $h(-8p)$.

The criterion for divisibility of $h(-8p)$ by 16 that we will use is due to Leonard and Williams [53, Theorem 2, p. 204]. Given a prime number $p \equiv 1 \bmod 8$ (so that $4|h(-8p)$), there exist integers u and v such that

$$p = u^2 - 2v^2, \qquad u > 0. \tag{4.1}$$

The integers u and v are *not* uniquely determined by p; nevertheless, if (u_0, v_0) is one such pair, then, every such pair (u, v) is of the form $u + v\sqrt{2} = \varepsilon^{2m}(u_0 \pm v_0\sqrt{2})$ for some $m \in \mathbb{Z}$, where $\varepsilon = 1 + \sqrt{2}$. The criterion for divisibility by 8 can be restated in terms of a quadratic residue symbol; one has

$$8|h(-8p) \Longleftrightarrow \left(\frac{u}{p}\right)_2 = 1.$$

Note that $1 = (u/p)_2 = (p/u)_2 = (-2/u)_2$, so that $8|h(-8p)$ if and only if $u \equiv 1, 3 \bmod 8$. Suppose that this condition is satisfied. As $\varepsilon^2(u + v\sqrt{2}) = (3u + 4v) + (2u + 3v)\sqrt{2}$ and v is even, we can always choose u and v in (4.1) so that $u \equiv 1 \bmod 8$. The criterion for divisibility of $h(-8p)$ by 16 states that if u and v are integers satisfying (4.1) and $u \equiv 1 \bmod 8$, then

$$16|h(-8p) \Longleftrightarrow \left(\frac{u}{p}\right)_4 = 1,$$

where $(u/p)_4$ is equal to 1 or -1 depending on whether or not u is a fourth power modulo p. To take advantage of the multiplicative properties of the fourth-power residue symbol, one has to work over a field containing $i = \sqrt{-1}$, a primitive fourth root of unity. Since u appears naturally via the splitting of p in $\mathbb{Q}(\sqrt{2})$, we see that the natural setting for the criterion above is the number field

$$M := \mathbb{Q}(\sqrt{2}, i) = \mathbb{Q}(\zeta_8),$$

of degree 4 over $\mathbb{Q}$. It is straightforward to check that the class number of M and each of its subfields is 1, that 2 is totally ramified in M, and that the unit group of its ring of integers $\mathcal{O}_M = \mathbb{Z}[\zeta_8]$ is generated by ζ_8 and $\varepsilon = 1 + \sqrt{2}$. Note that $M/\mathbb{Q}$ is a normal extension with Galois group isomorphic to the Klein four group, say $\{1, \sigma, \tau, \sigma\tau\}$, where σ fixes $\mathbb{Q}(i)$ and τ fixes $\mathbb{Q}(\sqrt{2})$.

$$\begin{tikzcd}
 & \mathbb{Q}(\zeta_8) \arrow[dl, dash, "\langle\sigma\rangle"'] \arrow[d, dash, "\langle\sigma\tau\rangle"] \arrow[dr, dash, "\langle\tau\rangle"] & \\
\mathbb{Q}(i) \arrow[dr, dash] & \mathbb{Q}(i\sqrt{2}) \arrow[d, dash] & \mathbb{Q}(\sqrt{2}) \arrow[dl, dash] \\
 & \mathbb{Q} &
\end{tikzcd}$$

Let $p \equiv 1 \bmod 8$ be a prime, so that p splits completely in M. Then there exists $w \in \mathcal{O}_M$ such that $\mathrm{N}(w) = p$, i.e., such that $p = w\sigma(w)\tau(w)\sigma\tau(w)$. Note that the inclusion $\mathbb{Z} \hookrightarrow \mathcal{O}_M$ induces an isomorphism $\mathbb{Z}/(p) \cong \mathcal{O}_M/(w)$, so that an integer n is a fourth power modulo p exactly when it is a fourth power modulo w. As $w\tau(w) \in \mathbb{Z}[\sqrt{2}]$, there exist integers u and v such that $w\tau(w) = u + v\sqrt{2}$. Then $u = (w\tau(w) + \sigma(w)\sigma\tau(w))/2$. With this in mind, we define, for any $\alpha \in \mathbb{Z}[\sqrt{2}]$,

$$\mathrm{r}(\alpha) = \frac{1}{2}\left(\alpha + \sigma(\alpha)\right)$$

and, for *any odd* (i.e., coprime to 2) $w \in \mathcal{O}_M$, not necessarily prime,

$$[w] := \left(\frac{\mathrm{r}(w\tau(w))}{w}\right)_4,$$

where $(\cdot/\cdot)_4$ is the quartic residue symbol in M; we recall the definition of $(\cdot/\cdot)_4$ in the next section. A simple computation shows that $\mathrm{r}(w\tau(w)) > 0$ for any non-zero $w \in \mathcal{O}_M$. Hence $16|h(-8p)$ if and only if $[w] = 1$, where w is any element of $\mathcal{O}_M$ such that $\mathrm{N}(w) = p$ and $\mathrm{r}(w) \equiv 1 \bmod 8$.

Given a Dirichlet character χ modulo 8, we define, for any odd $w \in \mathcal{O}_M$,

$$[w]_\chi := [w] \cdot \chi(\mathrm{r}(w\tau(w))).$$

Then

$$\frac{1}{4}\sum_{\chi \bmod 8}[w]_\chi = \begin{cases} [w] & \text{if } \mathrm{r}(w\tau(w)) \equiv 1 \bmod 8, \\ 0 & \text{otherwise,} \end{cases}$$

where the sum is over all Dirichlet characters modulo 8. Another simple computation shows that, for all odd $w \in \mathcal{O}_M$, we have $[\zeta_8 w] = [w]$. We note that $\mathrm{r}(\varepsilon^2\alpha) \equiv 3\cdot\mathrm{r}(\alpha) \bmod 8$ for any $\alpha \in \mathbb{Z}[\sqrt{2}]$, so that $\chi(\mathrm{r}(\varepsilon^2 w\tau(\varepsilon^2 w))) = \chi(\mathrm{r}(w\tau(w)))$ for every Dirichlet character χ modulo 8. Finally, we note that

$$[w] = \left(\frac{16\mathrm{r}(w\tau(w))}{w}\right)_4 = \left(\frac{8\sigma(w)\sigma\tau(w)}{w}\right)_4, \tag{4.2}$$

so that

$$[\varepsilon w] = \left(\frac{\sigma(\varepsilon)}{w}\right)_2 [w],$$

and hence $[\varepsilon^2 w] = [w]$. Having determined the action of the units $\mathcal{O}_M^\times$ on $[\cdot]_\chi$, we can define, for each Dirichlet character χ modulo 8, a sequence $\{a(\chi)_\mathfrak{n}\}_\mathfrak{n}$ indexed by *ideals* of $\mathcal{O}_M$ by setting $a(\chi)_\mathfrak{n} = 0$ if $\mathfrak{n}$ is even, and otherwise

$$a(\chi)_\mathfrak{n} := [w]_\chi + [\varepsilon w]_\chi, \tag{4.3}$$

where w is any generator of the odd ideal $\mathfrak{n}$. Again because $\mathrm{r}(\varepsilon^2\alpha) \equiv 3\cdot\mathrm{r}(\alpha) \bmod 8$ for any $\alpha \in \mathbb{Z}[\sqrt{2}]$, we see that if $8|h(-8p)$, then exactly one of $\mathrm{r}(w\tau(w))$ and $\mathrm{r}(\varepsilon w\tau(\varepsilon w))$ is 1 mod 8, and if $8 \nmid h(-8p)$, then neither is 1 mod 8. We have proved

Proposition 4.2.1. *Let $p \equiv 1 \bmod 8$ be a prime, and let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_M$ lying above p. Then*

$$\frac{1}{4}\sum_{\chi \bmod 8} a(\chi)_{\mathfrak{p}} = \begin{cases} 1 & \text{if } 16|h(-8p), \\ -1 & \text{if } 8|h(-8p) \text{ but } 16 \nmid h(-8p), \\ 0 & \text{otherwise,} \end{cases}$$

where the sum is over Dirichlet characters modulo 8.

4.3 Prerequisites

We now collect some definitions and facts that we will use in our proof of Theorem 4.1.1.

4.3.1 Quartic residue symbols and quartic reciprocity

Let L be a number field with ring of integers $\mathcal{O}_L$. Let $\mathfrak{p}$ be an odd prime ideal of $\mathcal{O}_L$ and let $\alpha \in \mathcal{O}_L$. One defines the *quadratic residue symbol* $(\alpha/\mathfrak{p})_{L,2}$ by setting

$$\left(\frac{\alpha}{\mathfrak{p}}\right)_{L,2} := \begin{cases} 0 & \text{if } \alpha \in \mathfrak{p} \\ 1 & \text{if } \alpha \notin \mathfrak{p} \text{ and } \alpha \equiv \beta^2 \bmod \mathfrak{p} \text{ for some } \beta \in \mathcal{O}_L \\ -1 & \text{otherwise.} \end{cases}$$

Then we have $(\alpha/\mathfrak{p})_{L,2} \equiv \alpha^{\frac{\mathrm{N}_{L/\mathbb{Q}}(\mathfrak{p})-1}{2}} \bmod \mathfrak{p}$. The quadratic residue symbol is then extended multiplicatively to all odd ideals $\mathfrak{n}$, and then also to all odd elements β in $\mathcal{O}_L$ by setting $(\alpha/\beta)_{L,2} = (\alpha/\beta\mathcal{O}_L)_{L,2}$. To define the quartic residue symbol, we assume that L contains $\mathbb{Q}(i)$. Then one can define the *quartic residue symbol* $(\alpha/\mathfrak{p})_{L,4}$ as the element of $\{\pm 1, \pm i, 0\}$ such that

$$\left(\frac{\alpha}{\mathfrak{p}}\right)_{L,4} \equiv \alpha^{\frac{\mathrm{N}_{L/\mathbb{Q}}(\mathfrak{p})-1}{4}} \bmod \mathfrak{p},$$

and extend this to all odd ideals $\mathfrak{n}$ and odd elements β in the same way as the quadratic residue symbol. A key property of the quartic residue symbol that we will use extensively is the following weak version of quartic reciprocity in $M := \mathbb{Q}(\zeta_8)$.

Lemma 4.3.1. *Let $\alpha, \beta \in \mathcal{O}_M$ with β odd. Then $(\alpha/\beta)_{M,4}$ depends only on the congruence class of β modulo $16\alpha\mathcal{O}_M$. Moreover, if α is also odd, then*

$$\left(\frac{\alpha}{\beta}\right)_{M,4} = \mu \cdot \left(\frac{\beta}{\alpha}\right)_{M,4},$$

where $\mu \in \{\pm 1, \pm i\}$ depends only on the congruence classes of α and β modulo $16\mathcal{O}_M$.

Proof. This follows from [50, Proposition 6.11, p. 199]. $\square$

4.3.2 Field lowering

A key feature of our proof is the reduction of quartic residue symbols in a quartic number field to quadratic residue symbols in a quadratic field. We do this by using the following three lemmas.

Lemma 4.3.2. *Let K be a number field and let $\mathfrak{p}$ be an odd prime ideal of K. Suppose that L is a quadratic extension of K such that L contains $\mathbb{Q}(i)$ and $\mathfrak{p}$ splits in L. Denote by ψ the non-trivial element in Gal(L/K). Then if ψ fixes $\mathbb{Q}(i)$ we have for all $\alpha \in \mathcal{O}_K$*

$$\left(\frac{\alpha}{\mathfrak{p}\mathcal{O}_L}\right)_{L,4} = \left(\frac{\alpha}{\mathfrak{p}\mathcal{O}_K}\right)_{K,2}$$

and if ψ does not fix $\mathbb{Q}(i)$ we have for all $\alpha \in \mathcal{O}_K$ with $\mathfrak{p} \nmid \alpha$

$$\left(\frac{\alpha}{\mathfrak{p}\mathcal{O}_L}\right)_{L,4} = 1$$

Proof. Since $\mathfrak{p}$ splits in L, we can write $\mathfrak{p} = \mathfrak{q}\psi(\mathfrak{q})$ for some prime ideal $\mathfrak{q}$ of L. Hence we have

$$\left(\frac{\alpha}{\mathfrak{p}\mathcal{O}_L}\right)_{L,4} = \left(\frac{\alpha}{\mathfrak{q}}\right)_{L,4} \left(\frac{\alpha}{\psi(\mathfrak{q})}\right)_{L,4}.$$

If ψ fixes i we find that

$$\left(\frac{\alpha}{\mathfrak{q}}\right)_{L,4} = \psi\left(\left(\frac{\alpha}{\mathfrak{q}}\right)_{L,4}\right) = \left(\frac{\psi(\alpha)}{\psi(\mathfrak{q})}\right)_{L,4} = \left(\frac{\alpha}{\psi(\mathfrak{q})}\right)_{L,4}.$$

Combining this with the previous identity gives

$$\left(\frac{\alpha}{\mathfrak{p}\mathcal{O}_L}\right)_{L,4} = \left(\frac{\alpha}{\mathfrak{q}}\right)_{L,4}^2 = \left(\frac{\alpha}{\mathfrak{q}}\right)_{L,2} = \left(\frac{\alpha}{\mathfrak{p}\mathcal{O}_K}\right)_{K,2},$$

establishing the first part of the lemma. If ψ does not fix i we find that

$$\left(\frac{\alpha}{\mathfrak{p}\mathcal{O}_L}\right)_{L,4} = \left(\frac{\alpha}{\mathfrak{q}}\right)_{L,4} \left(\frac{\alpha}{\psi(\mathfrak{q})}\right)_{L,4} = \left(\frac{\alpha}{\mathfrak{q}}\right)_{L,4} \psi\left(\left(\frac{\alpha}{\mathfrak{q}}\right)_{L,4}\right) = 1$$

by checking this for all values of $(\alpha/\mathfrak{q})_{L,4} \in \{\pm 1, \pm i\}$. This completes the proof. □

Lemma 4.3.3. *Let K be a number field and let $\mathfrak{p}$ be an odd prime ideal of K of degree 1 with residue field characteristic p. Suppose that L is a quadratic extension of K such that L contains $\mathbb{Q}(i)$ and $\mathfrak{p}$ stays inert in L. Then we have for all $\alpha \in \mathcal{O}_K$*

$$\left(\frac{\alpha}{\mathfrak{p}\mathcal{O}_L}\right)_{L,4} = \left(\frac{\alpha}{\mathfrak{p}\mathcal{O}_K}\right)_{K,2}^{\frac{p+1}{2}}.$$

Proof. We have

$$\left(\frac{\alpha}{\mathfrak{p}\mathcal{O}_L}\right)_{L,4} \equiv \alpha^{\frac{\mathrm{N}_L(\mathfrak{p})-1}{4}} \equiv \alpha^{\frac{p^2-1}{4}} \equiv \left(\alpha^{\frac{p-1}{2}}\right)^{\frac{p+1}{2}} \equiv \left(\alpha^{\frac{\mathrm{N}_K(\mathfrak{p})-1}{2}}\right)^{\frac{p+1}{2}} \equiv \left(\frac{\alpha}{\mathfrak{p}\mathcal{O}_K}\right)_{K,2}^{\frac{p+1}{2}} \bmod \mathfrak{p},$$

which immediately implies the lemma. □

Note that the previous lemmas only work if $\alpha \in \mathcal{O}_K$. Our last lemma gives a way to ensure that $\alpha \in \mathcal{O}_K$.

Lemma 4.3.4. *Let K be a number field and let L be a quadratic extension of K. Denote by ψ the non-trivial element in Gal(L/K). Suppose that $\mathfrak{p}$ is a prime ideal of K that does not ramify in L and further suppose that $\beta \in \mathcal{O}_L$ satisfies $\beta \equiv \psi(\beta) \bmod \mathfrak{p}\mathcal{O}_L$. Then there is $\beta' \in \mathcal{O}_K$ such that $\beta' \equiv \beta \bmod \mathfrak{p}\mathcal{O}_L$.*

Proof. Since by assumption $\mathfrak{p}$ does not ramify in L, we may assume that $\mathfrak{p}$ splits or stays inert in L. Let us first do the case that $\mathfrak{p}$ stays inert, which means precisely that $\psi(\mathfrak{p}) = \mathfrak{p}$. We conclude that ψ is in the decomposition group of $\mathfrak{p}$. Furthermore, the inertia group of $\mathfrak{p}$ is trivial by the assumption that $\mathfrak{p}$ does not ramify. Since ψ is not the identity, it follows that ψ must become the Frobenius map of the finite field extension $\mathcal{O}_K/\mathfrak{p} \hookrightarrow \mathcal{O}_L/\mathfrak{p}$. Then $\beta \equiv \psi(\beta) \bmod \mathfrak{p}\mathcal{O}_L$ means that β is fixed by the Frobenius map. We conclude that β comes from $\mathcal{O}_K/\mathfrak{p}$, which we had to prove.

We still have to prove the lemma if $\mathfrak{p}$ splits. In this case we can write $\mathfrak{p} = \mathfrak{q}\psi(\mathfrak{q})$ for some prime ideal $\mathfrak{q}$ of L. Note that

$$\mathcal{O}_K/\mathfrak{p} \hookrightarrow \mathcal{O}_L/\mathfrak{p}\mathcal{O}_L \cong \mathcal{O}_L/\mathfrak{q} \times \mathcal{O}_L/\psi(\mathfrak{q}). \tag{4.4}$$

One checks that ψ is the automorphism of $\mathcal{O}_L/\mathfrak{q} \times \mathcal{O}_L/\psi(\mathfrak{q})$ that maps the pair (x, y) to $(\psi(y), \psi(x))$. Hence $\beta \equiv \psi(\beta) \bmod \mathfrak{p}\mathcal{O}_L$ implies that there is some $x \in \mathcal{O}_L/\mathfrak{q}$ such that $\beta = (x, \psi(x))$ as an element of $\mathcal{O}_L/\mathfrak{q} \times \mathcal{O}_L/\psi(\mathfrak{q})$. Since $\mathcal{O}_K/\mathfrak{p} \cong \mathcal{O}_L/\mathfrak{q}$, we can pick $\beta' \in \mathcal{O}_K$ such that β' maps to x under the natural inclusion $\mathcal{O}_K/\mathfrak{p} \hookrightarrow \mathcal{O}_L/\mathfrak{q}$. Then it follows that β maps to $(\beta', \psi(\beta'))$ under the maps given as in (4.4). This implies that $\beta' \equiv \beta \bmod \mathfrak{p}\mathcal{O}_L$ as desired. □

4.3.3 A fundamental domain for the action of $\mathcal{O}_M^\times$

In defining $a(\chi)_\mathfrak{n}$ for odd ideals $\mathfrak{n}$ of $\mathcal{O}_M$, we had to choose a generator w for the ideal $\mathfrak{n}$. There are many such choices, since the group of units of $\mathcal{O}_M$ is quite large, i.e.,

$$\mathcal{O}_M^\times = \langle \zeta_8 \rangle \times \langle \varepsilon \rangle,$$

where $\varepsilon = 1 + \sqrt{2}$ as before. It will be important to us that we can choose generators that are in some sense as small as possible. We will do so by constructing a fundamental domain for the action (by multiplication) of $\mathcal{O}_M^\times$ on $\mathcal{O}_M$. The lemma that follows is usually implicitly proved in most number theory textbooks, but we have not been able to find a reference stating exactly the somewhat peculiar version that we will need. Below we deduce this version from [45, Lemma 1, p. 131].

More generally, let F be a number field of degree n over $\mathbb{Q}$ with ring of integers $\mathcal{O}_F$. Let $\sigma_1, \ldots, \sigma_r : F \hookrightarrow \mathbb{R}$ be the real embeddings of F and let $\tau_1, \bar{\tau_1}, \ldots, \tau_s, \bar{\tau_s} : F \hookrightarrow \mathbb{C}$ be the pairs of non-real complex conjugate embeddings of F (so that $r + 2s = n$). Let T be the subgroup of the unit group $\mathcal{O}_F^\times$ consisting of units of finite order. By Dirichet's Unit Theorem, there exists a free abelian subgroup $V \subset \mathcal{O}_F^\times$ of rank $r + s - 1$ such that $\mathcal{O}_F^\times = T \times V$; fix one such V.

Let $\eta = \{\eta_1, \ldots, \eta_n\}$ be an integral basis for $\mathcal{O}_F$; it defines an isomorphism $i_\eta : \mathbb{Q}^n \to F$ via the map $(a_1, \ldots, a_n) \mapsto a_1\eta_1 + \cdots a_n\eta_n$. For a subset $S \subset \mathbb{R}^n$ and an element $\alpha = a_1\eta_1 + \cdots + a_n\eta_n \in F$, we will say that α is in S (or $\alpha \in S$) to mean that $(a_1, \ldots, a_n) \in S$. Let $f_\eta \in \mathbb{Z}[x_1, \ldots, x_n]$ be the homogeneous polynomial of degree n in n variables defined by $f_\eta(x_1, \ldots, x_n) = \mathrm{N}(x_1\eta_1 + \cdots + x_n\eta_n)$. For a subset $S \subset \mathbb{R}^n$ and a real number $X > 0$, let $S(X)$ be the set of all $(s_1, \ldots, s_n) \in S$ such that $|f_\eta(s_1, \ldots, s_n)| \leq X$.

Lemma 4.3.5. *There exists a subset $\mathcal{D} \subset \mathbb{R}^n$ such that:*

(1) for all $\alpha \in \mathcal{O}_F \setminus \{0\}$, there exists a unique $v \in V$ such that $v\alpha \in \mathcal{D}$; moreover, the complete set of $u \in \mathcal{O}_F^\times$ such that $u\alpha \in \mathcal{D}$ is $\{\mu v : \mu \in T\}$;

(2) $\mathcal{D}(1)$ has an $(n-1)$-Lipschitz parametrizable boundary; and

(3) there exists a constant $C_\eta > 0$ such that for all $\alpha = a_1\eta_1 + \cdots + a_n\eta_n \in \mathcal{D}$ (with $a_i \in \mathbb{Z}$), we have $|a_i| \leq C_\eta \cdot \mathrm{N}(\alpha)^{\frac{1}{n}}$.

Proof. Let $J = \mathbb{R}^r \times \mathbb{C}^s$. Then $j = (\sigma_1, \ldots, \sigma_r, \tau_1, \ldots, \tau_s)$ defines an embedding $j : F \hookrightarrow J$. Moreover, $j \circ i_\eta : \mathbb{Q}^n \to J$ is a *linear* map of $\mathbb{Q}$-vector spaces. By extension of scalars, we extend this to a linear map

$$\bar{j} : \mathbb{R}^n \to J.$$

It follows from [45, Lemma 1, p. 131] and its proof that there is a subset $D \subset J^\times$ such that:

(1') for all $\alpha \in J^\times$, there exists a unique $v \in V$ such that $v\alpha \in D$; moreover, the complete set of $u \in \mathcal{O}_F^\times$ such that $u\alpha \in D$ is $\{\mu v : \mu \in T\}$; and

(2') $D(1) = \{(\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_s) \in D : \prod_{i=1}^r |\alpha_i| \prod_{j=1}^s |\beta_j|^2 \leq 1\}$ has an $(n-1)$-Lipschitz parametrizable boundary.

(3') for all non-zero $t \in \mathbb{R}$, we have $tD = D$.

Let $\mathcal{D} = \bar{j}^{-1}(D)$. Then (1) follows immediately from (1'). Since $\bar{j}$ is linear and hence Lipschitz continuous, (2') immediately implies (2) (after also taking into account the definitions of $D(1)$, f_η, and $\mathcal{D}(1)$). By (2), the set $\mathcal{D}(1) \subset \mathbb{R}^n$ is bounded, so we can set

$$C_\eta = \sup\{|a_i| : (a_1, \ldots, a_n) \in \mathcal{D}(1)\}.$$

Finally, again because $\bar{j}$ is linear, (3') implies that $t\mathcal{D} = \mathcal{D}$ for all non-zero $t \in \mathbb{R}$, so that $\mathcal{D}(t) = t^{1/n}\mathcal{D}(1)$. This proves (3). $\square$

4.3.4 General bilinear sum estimates

Let F, n, η, and V be as in Section 4.3.3. Fix a fundamental domain $\mathcal{D}$ for the action of V on $\mathcal{O}_F$ as in Lemma 4.3.5. Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be a pair of translates of $\mathcal{D}$, i.e., $\mathcal{D}_i = v_i\mathcal{D}$ for some $v_i \in V$. Let $\mathfrak{f}$ be a non-zero ideal in $\mathcal{O}_F$, and let $S_\mathfrak{f}$ be the set of elements in $\mathcal{O}_F$ coprime to $\mathfrak{f}$. Suppose γ is a map

$$\gamma : S_\mathfrak{f} \times \mathcal{O}_F \to \{-1, 0, 1\}$$

satisfying the following properties:

(P1) for every pair of invertible congruence classes ω and ζ modulo $\mathfrak{f}$, there exists $\mu(\omega, \zeta) \in \{\pm 1\}$ such that $\gamma(w, z) = \mu(\omega, \zeta)\gamma(z, w)$ whenever $w \equiv \omega \bmod \mathfrak{f}$ and $z \equiv \zeta \bmod \mathfrak{f}$;

(P2) for all $z_1, z_2 \in \mathcal{O}_F$ and all $w \in S_\mathfrak{f}$, we have $\gamma(w, z_1z_2) = \gamma(w, z_1)\gamma(w, z_2)$; similarly, for all $w_1, w_2 \in S_\mathfrak{f}$ and all $z \in \mathcal{O}_F$, we have $\gamma(w_1w_2, z) = \gamma(w_1, z)\gamma(w_2, z)$; and

(P3) for all non-zero $w \in S_\mathfrak{f}$, we have $\gamma(w, z_1) = \gamma(w, z_2)$ for all $z_1, z_2 \in \mathcal{O}_F$ with $z_1 \equiv z_2 \bmod \mathrm{N}w$; moreover, we have

$$\sum_{\xi \bmod w} \gamma(w, \xi) = 0$$

unless $\mathrm{N}w$ is squarefull.

We will consider bilinear sums of the type

$$B(M, N; \omega, \zeta) := \sum_{\substack{w \in \mathcal{D}_1(M) \\ w \equiv \omega \bmod \mathfrak{f}}} \sum_{\substack{z \in \mathcal{D}_2(N) \\ z \equiv \zeta \bmod \mathfrak{f}}} \alpha_w \beta_z \gamma(w, z), \tag{4.5}$$

where $\{\alpha_w\}_w$ and $\{\beta_z\}_z$ are bounded sequences of complex numbers, ω and ζ are invertible congruence classes modulo $\mathfrak{f}$, and M and N are positive real numbers. Recall that $w \in \mathcal{D}_1(M)$ if and only if $w \in \mathcal{D}_1$ and $\mathrm{N}(w) \leq M$, and similarly for $\mathcal{D}_2(N)$. Also recall that n is the degree of $F/\mathbb{Q}$. The following proposition is analogous to the bilinear sum estimates in [23, 24].

Proposition 4.3.6. *We have*

$$B(M, N; \omega, \zeta) \ll_\epsilon \left(M^{-\frac{1}{6n}} + N^{-\frac{1}{6n}}\right)(MN)^{1+\epsilon},$$

where the implied constant depends on ϵ, on the units v_1 and v_2, on the supremum norms of $\{\alpha_w\}_w$ and $\{\beta_z\}_z$, and the congruence classes ω and ζ modulo $\mathfrak{f}$.

Proof. We will prove that

$$B(M, N; \omega, \zeta) \ll_\epsilon M^{-\frac{1}{6n}}(MN)^{1+\epsilon} \tag{4.6}$$

whenever $N \geq M$; the proposition then immediately follows from the symmetry of the sum $B(M,N;\omega,\zeta)$ coming from property (P1). So suppose that $N \geq M$. We fix an integer $k \geq 2n$, and we apply Hölder's inequality (with $1 = \frac{k-1}{k} + \frac{1}{k}$) to the w variable to get

$$|B(M,N;\omega,\zeta)|^k \leq \left(\sum_w |\alpha_w|^{\frac{k}{k-1}}\right)^{k-1} \sum_w \left|\sum_z \beta_z \gamma(w,z)\right|^k,$$

where the summations over w and z are as above in (4.5). The first factor above is bounded trivially by $\ll M^{k-1}$, where the implied constant depends on the supremum norm of the sequence $\{\alpha_w\}_w$, on the fixed unit v_1, and on the constant C_η from part (3) of Lemma 4.3.5. We use property (P2), as well as the identity $|\alpha|^k = \alpha^k \cdot (|\alpha|/\alpha)^k$, to expand the inner sum in the second factor above, getting

$$|B(M,N;\omega,\zeta)|^k \ll M^{k-1} \sum_w \varepsilon(w) \sum_z \beta'_z \gamma(w,z),$$

where

$$\beta'_z = \sum_{\substack{z = z_1 \cdots z_k \\ z_1,\ldots,z_k \in \mathcal{D}_2(N) \\ z_1 \equiv \cdots \equiv z_k \equiv \zeta \bmod \mathfrak{f}}} \beta_{z_1} \cdots \beta_{z_k},$$

where $\varepsilon(w) = (|\sum_z \beta_z \gamma(w,z)|/\sum_z \beta_z \gamma(w,z))^k$, and where once again the summation conditions for w are as in (4.5). Since an ideal $\mathfrak{n}$ in $\mathcal{O}_F$ can be written as a product of k ideals in at most $\ll_\epsilon \mathrm{N}(\mathfrak{n})^\epsilon$ ways, and since $\mathcal{D}_2$ contains at most one generator of any principal ideal, we see that $\beta'_z \ll_\epsilon N^\epsilon$. Moreover, the coordinates of each $z_i \in \mathcal{D}_2$ $(1 \leq i \leq k)$ of norm at most N in the basis η are bounded by $N^{\frac{1}{n}}$ times a constant depending on the unit v_2 and on C_η from Lemma 4.3.5. Hence we may assume that the sum $\sum_z \beta'_z \gamma(w,z)$ above is over $z = a_1\eta_1 + \cdots + a_n\eta_n$ in a box $\mathcal{B}$ defined by $|a_j| \ll N^{\frac{k}{n}}$ $(1 \leq j \leq n)$, with the implied constant depending on v_2 and on the integral basis η. Next, we apply the Cauchy-Schwarz inequality to the z variable above and use property (P2) to get

$$\left|\sum_w \varepsilon(w) \sum_z \beta'_z \gamma(w,z)\right|^2 \ll_\epsilon N^{k+\epsilon} \sum_{w_1}\sum_{w_2} \varepsilon(w_1)\overline{\varepsilon(w_2)} \sum_z \gamma(w_1 w_2, z),$$

where the summation conditions for w_1 and w_2 are as those for w in (4.5), while the inner sum is over $z \in \mathcal{B}$. We break up the sum over z into congruence classes ξ modulo $\mathrm{N}(w_1 w_2)$ and note that, by property (P3),

$$\sum_{\xi \bmod w_1 w_2} \gamma(w_1 w_2, \xi) = 0$$

unless $\mathrm{N}(w_1 w_2)$ is squarefull. By counting points z in the box $\mathcal{B}$ and noting that $\mathrm{N}(w_1 w_2) \leq M^2$, this gives

$$\sum_z \gamma(w_1 w_2, z) \ll \begin{cases} N^k & \text{if } \mathrm{N}(w_1 w_2) \text{ is squarefull} \\ \sum_{i=1}^n M^{2i} N^{k(1-\frac{i}{n})} & \text{otherwise.} \end{cases}$$

Since we took $k \geq 2n$ and since $N \geq M$, we have $N^{\frac{k}{n}} \geq M^2$, so the last bound can be simplified to $M^2 N^{k(1-\frac{1}{n})}$. Hence, putting together all of the bounds above, we get

$$\begin{aligned} |B(M,N;\omega,\zeta)|^{2k} &\ll_\epsilon M^{2k-2}N^k\left(M \cdot N^k + M^2 \cdot M^2 N^{k(1-\frac{1}{n})}\right)(MN)^\epsilon \\ &\ll_\epsilon \left(M^{2k-1}N^{2k} + M^{2k+2}N^{2k(1-\frac{1}{2n})}\right)(MN)^\epsilon. \end{aligned}$$

Since $N \geq M$, if we take $k = 3n$, we get that $N^{2k\frac{1}{2n}} \geq M^3$, so that the first term above dominates the second term. With this choice of k, we get

$$|B(M,N;\omega,\zeta)| \ll_\epsilon M^{-\frac{1}{6n}}(MN)^{1+\epsilon},$$

and this finishes the proof of (4.6). □

4.3.5 The sieve

We will prove Theorem 4.1.1 by a sieve of Friedlander et al. [24] that generalizes the ideas of Vinogradov [75, 76] to the setting of number fields. Let χ be a Dirichlet character modulo 8, and let $a(\chi)_\mathfrak{n}$ be defined as in (4.3). We will prove the following two propositions.

Proposition 4.3.7. *For every $\epsilon > 0$, we have*

$$\sum_{\mathrm{N}(\mathfrak{n}) \leq X,\ \mathfrak{m}|\mathfrak{n}} a(\chi)_\mathfrak{n} \ll_\epsilon X^{1-\frac{1}{64}+\epsilon}$$

uniformly for all non-zero ideals $\mathfrak{m}$ of $\mathcal{O}_M$ and all $X \geq 2$.

Proposition 4.3.8. *For every $\epsilon > 0$, we have*

$$\sum_{\mathrm{N}(\mathfrak{m}) \leq M} \sum_{\mathrm{N}(\mathfrak{n}) \leq N} \alpha_\mathfrak{m}\beta_\mathfrak{n} a(\chi)_{\mathfrak{mn}} \ll_\epsilon (M+N)^{\frac{1}{24}}(MN)^{1-\frac{1}{24}+\epsilon}$$

uniformly for all $M, N \geq 2$ and sequences of complex numbers $\{\alpha_\mathfrak{m}\}$ and $\{\beta_\mathfrak{n}\}$ satisfying $|\alpha_\mathfrak{m}|, |\beta_\mathfrak{n}| \leq 1$.

From these two propositions we can apply [24, Proposition 5.2, p. 722] with $\theta_1 = \frac{1}{64}$ and $\theta_2 = \frac{1}{24}$ to prove

$$\sum_{\mathrm{N}(\mathfrak{n}) \leq X} a(\chi)_\mathfrak{n}\Lambda(\mathfrak{n}) \ll_\theta X^{1-\theta}$$

for all $\theta < 1/(49 \cdot 64) = 1/3136$. By partial summation, it follows that, say,

$$\sum_{\mathrm{N}(\mathfrak{p}) \leq X} a(\chi)_\mathfrak{p} \ll X^{1-\frac{1}{3200}}. \tag{4.7}$$

As

$$\sum_{\substack{\mathrm{N}(\mathfrak{p}) \leq X \\ \mathfrak{p} \text{ lies over } p \not\equiv 1 \bmod 8}} 1 \ll X^{\frac{1}{2}},$$

Theorem 4.1.1 follows from (4.7) and Proposition 4.2.1. It now remains to prove Propositions 4.3.7 and 4.3.8.

4.4 Proof of Proposition 4.3.7

Let χ be a Dirichlet character modulo 8. Let $\mathfrak{m}$ be an odd ideal of $\mathcal{O}_M$. In view of Proposition 4.2.1 we must bound the following sum

$$A(x) = A(x;\chi,\mathfrak{m}) := \sum_{\substack{\mathrm{N}(\mathfrak{a})\leq x \\ (\mathfrak{a},2)=1,\ \mathfrak{m}|\mathfrak{a}}} \left([\alpha]_\chi + [\varepsilon\alpha]_\chi\right),$$

where α is chosen to be any generator of $\mathfrak{a}$. Our proof is based on the argument in [41, Section 3, p. 12-19], which is in turn based on [24, Section 6, p. 722-733]. Let $\mathcal{D}$ be a fundamental domain for the action of $\mathcal{O}_M^\times$ on $\mathcal{O}_M \setminus \{0\}$ as in Lemma 4.3.5, with respect to the integral basis $\eta = \{1, \zeta_8, \zeta_8^2, \zeta_8^3\}$. Each non-zero ideal $\mathfrak{a}$ has exactly 8 generators $\alpha \in \mathcal{D}$. Set $u_1 = 1$ and $u_2 = \varepsilon$. Set $F = 16$. Note that $\chi(\mathrm{r}(\alpha\tau(\alpha)))$ depends only on the congruence class of α modulo 8. After splitting the above sum into congruence classes modulo F, and using (4.2) and Lemma 4.3.1, we find that

$$A(x) = \frac{1}{8}\sum_{i=1}^{2}\sum_{\substack{\rho \bmod F \\ (\rho,F)=1}} \mu(\rho,u_i)A(x;\rho,u_i),$$

where $\mu(\rho,u_i) \in \{\pm1,\pm i\}$ depends only on ρ and u_i and where

$$A(x;\rho,u_i) := \sum_{\substack{\alpha\in u_i\mathcal{D},\ \mathrm{N}(\alpha)\leq x \\ \alpha\equiv\rho \bmod F \\ \alpha\equiv 0 \bmod \mathfrak{m}}} \left(\frac{\sigma(\alpha)}{\alpha}\right)_{M,4}\left(\frac{\sigma\tau(\alpha)}{\alpha}\right)_{M,4}.$$

Our goal is to estimate $A(x;\rho,u_i)$ separately for each congruence class $\rho \bmod F$ such that $(\rho,F)=1$ and each unit u_i. We view $\mathcal{O}_M$ as a $\mathbb{Z}$-module of rank 4 and decompose it as $\mathcal{O}_M = \mathbb{Z}\oplus\mathbb{M}$, where $\mathbb{M} = \mathbb{Z}\zeta_8 \oplus \mathbb{Z}\zeta_8^2 \oplus \mathbb{Z}\zeta_8^3$ is a free $\mathbb{Z}$-module of rank 3. We can write α uniquely as

$$\alpha = a + \beta, \text{ with } a\in\mathbb{Z}, \beta\in\mathbb{M},$$

so that the summation conditions above are equivalent to

$$a+\beta\in u_i\mathcal{D}, \quad \mathrm{N}(a+\beta)\leq x, \quad a+\beta\equiv\rho \bmod F, \quad a+\beta\equiv 0 \bmod \mathfrak{m}. \qquad (*)$$

We may assume that $\sigma(\beta)\neq\beta$ and $\sigma\tau(\beta)\neq\beta$. Indeed, if $\sigma(\beta)=\beta$ or $\sigma\tau(\beta)=\beta$, the residue symbol in $A(x;\rho,u_i)$ is zero. We are now going to rewrite $(\sigma(\alpha)/\alpha)_{M,4}$ and $(\sigma\tau(\alpha)/\alpha)_{M,4}$ by using the same trick as in [24, p. 725]. Put

$$\sigma(\beta)-\beta = \eta^4 c_0 c \quad \text{ and } \quad \sigma\tau(\beta)-\beta = \eta'^4 c_0' c'$$

with $c_0, c_0', c, c', \eta, \eta' \in \mathcal{O}_M$, $c_0, c_0' \mid F$ not divisible by a non-trivial fourth power, $\eta,\eta' \mid F^\infty$ and $(c,F) = (c',F) = 1$. By multiplying with an appropriate unit we can even ensure that $c\in\mathbb{Z}[i]$ and $c'\in\mathbb{Z}[\sqrt{-2}]$. Indeed, observe that

$$\alpha' := \frac{\sigma(\alpha)-\alpha}{\zeta_8} = \frac{\sigma(\beta)-\beta}{\zeta_8} \in \mathbb{Z}[i], \tag{4.8}$$

and we have a similar identity for $\sigma\tau(\beta) - \beta$. Then we obtain, just as in [41, p. 14], by Lemma 4.3.1,

$$\left(\frac{\sigma(\alpha)}{\alpha}\right)_{M,4} = \mu_1 \cdot \left(\frac{a+\beta}{c\mathcal{O}_M}\right)_{M,4} \quad \text{and} \quad \left(\frac{\sigma\tau(\alpha)}{\alpha}\right)_{M,4} = \mu_2 \cdot \left(\frac{a+\beta}{c'\mathcal{O}_M}\right)_{M,4},$$

where $\mu_1, \mu_2 \in \{\pm 1, \pm i\}$ depend only on ρ and β. Hence

$$A(x; \rho, u_i) \leq \sum_{\beta \in \mathbb{M}} |T(x; \beta, \rho, u_i)|,$$

where

$$T(x; \beta, \rho, u_i) := \sum_{\substack{a \in \mathbb{Z} \\ a+\beta \text{ sat. } (*)}} \left(\frac{a+\beta}{c\mathcal{O}_M}\right)_{M,4} \left(\frac{a+\beta}{c'\mathcal{O}_M}\right)_{M,4}.$$

From now on we treat β as fixed and estimate $T(x; \beta, \rho, u_i)$. It is here that we deviate from [24] and [41]. Since we chose $c' \in \mathbb{Z}[\sqrt{-2}]$, we can factor the principal ideal $(c') \subset \mathbb{Z}[\sqrt{-2}]$ into prime ideals in $\mathbb{Z}[\sqrt{-2}]$ that do not ramify in M, say, $(c') = \prod_{i=1}^{k} \mathfrak{p}_i^{e_i}$, so that

$$\left(\frac{a+\beta}{c'\mathcal{O}_M}\right)_{M,4} = \prod_{i=1}^{k} \left(\frac{a+\beta}{\mathfrak{p}_i\mathcal{O}_M}\right)_{M,4}^{e_i}.$$

We claim that $((a+\beta)/\mathfrak{p}\mathcal{O}_M)_{M,4} = 1$ if $\mathfrak{p} \nmid a+\beta$. As a first step we can replace β by some $\beta' \in \mathbb{Z}[\sqrt{-2}]$ due to Lemma 4.3.4. Then Lemma 4.3.2 proves the claim if $\mathfrak{p}$ splits in M. Finally suppose that $\mathfrak{p}$ stays inert in M. If we define $p := \mathfrak{p} \cap \mathbb{Z}$, we find that $p \equiv 3 \bmod 8$. Hence Lemma 4.3.3 finishes the proof of the claim.

The factor $((a+\beta)/c\mathcal{O}_M)_{M,4}$ is handled more similarly to [24, (6.21), p. 727]. Since we chose $c \in \mathbb{Z}[i]$, we factor $(c) \subset \mathbb{Z}[i]$ in $\mathbb{Z}[i]$ as $(c) = \mathfrak{g}\mathfrak{q}$ in the unique way so that $q := N_{\mathbb{Q}(i)/\mathbb{Q}}(\mathfrak{q})$ is a squarefree odd integer and $g := N_{\mathbb{Q}(i)/\mathbb{Q}}(\mathfrak{g})$ is an odd squarefull integer coprime with q.

Lemma 4.3.4 and the Chinese remainder theorem imply that there exists $\beta' \in \mathbb{Z}[i]$ such that $\beta \equiv \beta' \bmod \mathfrak{q}\mathcal{O}_M$. Next, Lemma 4.3.2 and Lemma 4.3.3 imply that

$$((a+\beta')/\mathfrak{q}\mathcal{O}_M)_{M,4} = ((a+\beta')/\mathfrak{q})_{\mathbb{Q}(i),2}.$$

Finally, as q is squarefree, the Chinese remainder theorem guarantees the existence of a rational integer b such that $\beta' \equiv b \bmod \mathfrak{q}$. Combining all of this gives

$$\left(\frac{a+\beta}{c\mathcal{O}_M}\right)_{M,4} = \left(\frac{a+\beta}{\mathfrak{g}\mathcal{O}_M}\right)_{M,4} \left(\frac{a+b}{\mathfrak{q}}\right)_{\mathbb{Q}(i),2}.$$

Since c depends on β and not on a, we find that b depends on β and not on a. Now define g_0 as the radical of g, i.e., $g_0 := \prod_{p|g} p$. We observe that the quartic residue symbol $(\alpha/\mathfrak{g}\mathcal{O}_M)_{M,4}$ is periodic in α modulo $\mathfrak{g}^* := \prod_{\mathfrak{p}|\mathfrak{g}} \mathfrak{p}$. But clearly $\mathfrak{g}^*$ divides g_0, and hence we conclude that $((a+\beta)/\mathfrak{g}\mathcal{O}_M)_{M,4}$ is periodic of period g_0 when viewed as a function of $a \in \mathbb{Z}$. So we split $T(x;\beta,\rho,u_i)$ into congruence classes modulo g_0, giving

$$|T(x;\beta,\rho,u_i)| \leq \sum_{a_0 \bmod g_0} |T(x;\beta,\rho,u_i,a_0)|,$$

where

$$T(x;\beta,\rho,u_i,a_0) = \sum_{\substack{a\in\mathbb{Z} \\ a+\beta \text{ sat. } (*) \\ a\equiv a_0 \bmod g_0}} \left(\frac{a+b}{\mathfrak{q}}\right)_{\mathbb{Q}(i),2} \left(\frac{a+\beta}{c'\mathcal{O}_M}\right)_{M,4}.$$

We have already proven that $((a+\beta)/c'\mathcal{O}_M)_{M,4} = 1$ unless $\gcd(a+\beta,c') \neq (1)$ and in this case we have $((a+\beta)/c'\mathcal{O}_M)_{M,4} = 0$. An application of inclusion-exclusion gives

$$|T(x;\beta,\rho,u_i,a_0)| \leq \sum_{\substack{\mathfrak{d}\mid c'\mathcal{O}_M \\ \mathfrak{d} \text{ squarefree}}} |T(x;\beta,\rho,u_i,a_0,\mathfrak{d})|,$$

where

$$T(x;\beta,\rho,u_i,a_0,\mathfrak{d}) := \sum_{\substack{a\in\mathbb{Z} \\ a+\beta \text{ sat. } (*) \\ a\equiv a_0 \bmod g_0 \\ a+\beta\equiv 0 \bmod \mathfrak{d}}} \left(\frac{a+b}{\mathfrak{q}}\right)_{\mathbb{Q}(i),2}. \tag{4.9}$$

We unwrap the summation conditions above similarly as in [24, p. 728]. Certainly $a+\beta \in u_i\mathcal{D}$ implies that $|a| \leq Cx^{\frac{1}{4}}$, where $C>0$ depends only on one of the two fixed units u_i. The condition $\mathrm{N}_{M/\mathbb{Q}}(a+\beta) \leq x$ is for fixed β and x a polynomial inequality of degree 4 in a. Hence the summation variable $a \in \mathbb{Z}$ runs over at most 4 intervals of length $\leq Cx^{1/4}$ with endpoints depending on β and x.

Next, the congruence conditions $a+\beta \equiv \rho \bmod F$, $a+\beta \equiv 0 \bmod \mathfrak{m}$, $a \equiv a_0 \bmod g_0$ and $a+\beta \equiv 0 \bmod \mathfrak{d}$ imply that a runs over some arithmetic progression of modulus k dividing g_0mdF, where we define $m := \mathrm{N}_{M/\mathbb{Q}}(\mathfrak{m})$ and $d := \mathrm{N}_{M/\mathbb{Q}}(\mathfrak{d})$. Moreover, as $q = \mathrm{N}_{\mathbb{Q}(i)/\mathbb{Q}}(\mathfrak{q})$ is squarefree, $(\cdot/\mathfrak{q})_{\mathbb{Q}(i),2} : \mathbb{Z} \to \{\pm1, 0\}$ is the real primitive Dirichlet character of modulus q.

All in all, the sum in (4.9) can be rewritten as at most 4 incomplete real character sums of length $\ll x^{\frac{1}{4}}$ and modulus $q \ll x^{\frac{1}{2}}$, each of which runs over an arithmetic progression of modulus k. When the modulus q of the Dirichlet character divides the modulus k of the arithmetic progression, one does not get the desired cancellation. So for now we assume that $q \nmid k$, and we will handle the case $q \mid k$ later. As has been explained in [25, 7., p. 924-925], Burgess's bound for short character sums [8] implies that for each integer $r \geq 2$, we have

$$T(x;\beta,\rho,u_i,a_0,\mathfrak{d}) \ll_{\epsilon,r} x^{\frac{1}{4}\left(1-\frac{1}{r}\right)} \cdot x^{\frac{1}{2}\left(\frac{r+1}{4r^2}+\epsilon\right)},$$

so that on taking $r=2$, we obtain

$$T(x;\beta,\rho,u_i) \ll_\epsilon g_0 x^{\frac{1}{4}-\frac{1}{32}+\epsilon}. \tag{4.10}$$

It remains to do the case $q \mid k$. Certainly, this implies $q \mid md$. So (4.10) holds if $q \nmid md$. Recall that $(c) = \mathfrak{g}\mathfrak{q}$, hence we have (4.10) unless

$$p \mid \mathrm{N}_{\mathbb{Q}(i)/\mathbb{Q}}(\alpha') \implies p^2 \mid mdF\mathrm{N}_{\mathbb{Q}(i)/\mathbb{Q}}(\alpha') \tag{4.11}$$

for all primes p, where α' is defined as in (4.8). Define $A_\square(x; \rho, u_i)$ as the contribution to $A(x; \rho, u_i)$ from β satisfying (4.11). Then we get

$$A_\square(x; \rho, u_i) \leq |\{\alpha \in u_i\mathcal{D} : N_{M/\mathbb{Q}}(\alpha) \leq x,\ p \mid \mathrm{N}_{\mathbb{Q}(i)/\mathbb{Q}}(\alpha') \implies p^2 \mid mdF\mathrm{N}_{\mathbb{Q}(i)/\mathbb{Q}}(\alpha')\}|.$$

We decompose $\mathcal{O}_M$ as $\mathcal{O}_M = \mathbb{Z}[i] \oplus \mathbb{M}'$, where $\mathbb{M}' = \mathbb{Z}\zeta_8 \oplus \mathbb{Z}\zeta_8^3 = \mathbb{Z}[i] \cdot \zeta_8$ is a free $\mathbb{Z}$-module of rank 2. The linear map $\mathbb{M}' \to \mathbb{Z}[i]$ given by $\alpha \mapsto \alpha'$ is injective. Now suppose $\alpha \in u_i\mathcal{D}$ and $\mathrm{N}_{M/\mathbb{Q}}(\alpha) \leq x$. Then by Lemma 4.3.5, if we write $\alpha = a_1 + a_2 i + (a_3 + a_4 i)\zeta_8$, we have $a_j \ll x^{\frac{1}{4}}$ for $1 \leq j \leq 4$. Hence the norm $\mathrm{N}_{\mathbb{Q}(i)/\mathbb{Q}}(\cdot)$ of $\alpha' = -2(a_3 + a_4 i)$ is $\ll x^{\frac{1}{2}}$, and so

$$\begin{aligned} A_\square(x; \rho, u_i) \ll x^{\frac{1}{2}} |\{\alpha' \in \mathbb{Z}[i] : \mathrm{N}_{\mathbb{Q}(i)/\mathbb{Q}}(\alpha') \ll x^{\frac{1}{2}},& \\ p \mid \mathrm{N}_{\mathbb{Q}(i)/\mathbb{Q}}(\alpha') \implies p^2 \mid mdF\mathrm{N}_{\mathbb{Q}(i)/\mathbb{Q}}(\alpha')\}|.& \end{aligned} \tag{4.12}$$

Note that there are at most $\ll_\epsilon b^\epsilon$ elements $\alpha' \in \mathbb{Z}[i]$ such that $\mathrm{N}_{\mathbb{Q}(i)/\mathbb{Q}}(\alpha') = b$. This gives

$$A_\square(x; \rho, u_i) \ll_\epsilon x^{\frac{1}{2}+\epsilon} \sum_{\substack{b \ll x^{\frac{1}{2}};\\ p \mid b \implies p^2 \mid mdFb}} 1,$$

where b runs over the positive rational integers. We assume that $m \leq x$ because otherwise $A(x)$ is the empty sum. This shows that $md \ll x^2$ and we conclude that

$$A_\square(x; \rho, u_i) \ll_\epsilon x^{\frac{3}{4}+\epsilon}.$$

Let $A_0(x; \rho, u_i)$ be the contribution to $A(x; \rho, u_i)$ of the terms $\alpha = a + \beta$ not satisfying (4.11). Then we can split $A(x; \rho, u_i)$ as

$$A(x; \rho, u_i) = A_\square(x; \rho, u_i) + A_0(x; \rho, u_i).$$

To estimate $A_0(x; \rho, u_i)$ we can try to use our bound (4.10) for every relevant β, but for this we need g_0 to be small. Hence we make the further partition

$$A_0(x; \rho, u_i) = A_1(x; \rho, u_i) + A_2(x; \rho, u_i),$$

where β satisfies the additional constraint

$$\begin{aligned} g_0 &\leq Z \text{ in the sum } A_1(x; \rho, u_i), \\ g_0 &> Z \text{ in the sum } A_2(x; \rho, u_i). \end{aligned}$$

Here Z is at our disposal, and we choose it later. We estimate $A_1(x; \rho, u_i)$ as in [24] by using (4.10) and summing over $\beta = b_1\zeta_8 + b_2\zeta_8^2 + b_3\zeta_8^3 \in \mathbb{M}$ satisfying $b_i \ll x^{\frac{1}{4}}$ for $1 \leq i \leq 3$ to obtain

$$A_1(x; \rho, u_i) \ll_\epsilon Zx^{1-\frac{1}{32}+\epsilon}.$$

To finish the proof of Proposition 4.3.7 it remains to estimate $A_2(x;\rho,u_i)$. Note that $g_0 \le \sqrt{g}$ and $g \le \mathrm{N}_{\mathbb{Q}(i)/\mathbb{Q}}(c) \le \mathrm{N}_{\mathbb{Q}(i)/\mathbb{Q}}(\alpha') \ll x^{\frac{1}{2}}$. Hence, similarly as for $A_\Box(x;\rho,u_i)$, with $b = \mathrm{N}_{\mathbb{Q}(i)/\mathbb{Q}}(\alpha')$, we have

$$A_2(x;\rho,u_i) \ll_\epsilon x^{\frac{1}{2}+\epsilon} \sum_{Z<g_0\ll x^{\frac{1}{4}}} \sum_{\substack{b\ll x^{\frac{1}{2}} \\ g_0^2|b}} 1 \ll_\epsilon Z^{-1}x^{1+\epsilon}.$$

Picking $Z = x^{\frac{1}{64}}$ finishes the proof of Proposition 4.3.7.

4.5 Proof of Proposition 4.3.8

Let w and z be odd elements in $\mathcal{O}_M$. All quadratic and quartic residue symbols that follow are over M. By (4.2), we have

$$[wz] = \left(\frac{8\sigma(wz)\sigma\tau(wz)}{wz}\right)_4 = [w][z]\left(\frac{\sigma(w)}{z}\right)_4\left(\frac{\sigma\tau(w)}{z}\right)_4\left(\frac{\sigma(z)}{w}\right)_4\left(\frac{\sigma\tau(z)}{w}\right)_4.$$

By Lemma 4.3.1, we have, for some $\mu_1 \in \{\pm 1, \pm i\}$ that depends only on the congruence classes of w and z modulo 16,

$$\begin{aligned}\left(\frac{\sigma(w)}{z}\right)_4\left(\frac{\sigma(z)}{w}\right)_4 &= \mu_1\left(\frac{z}{\sigma(w)}\right)_4\left(\frac{\sigma(z)}{w}\right)_4 = \mu_1\left(\frac{z}{\sigma(w)}\right)_4\sigma\left(\frac{z}{\sigma(w)}\right)_4 \\ &= \mu_1\left(\frac{z}{\sigma(w)}\right)_2,\end{aligned}$$

because $\sigma(i) = i$. Similarly, for some $\mu_2 \in \{\pm 1, \pm i\}$ that depends only on the congruence classes of w and z modulo 16,

$$\left(\frac{\sigma\tau(w)}{z}\right)_4\left(\frac{\sigma\tau(z)}{w}\right)_4 = \mu_2\left(\frac{z}{\sigma\tau(w)}\right)_4\sigma\tau\left(\frac{z}{\sigma\tau(w)}\right)_4 = \mu_2\mathbb{1}_{\gcd(\sigma\tau(w),z)=1},$$

because $\sigma\tau(i) = -i$. Hence we get, for $\mu_3 = \mu_1\mu_2$,

$$[wz] = \mu_3[w][z]\left(\frac{z}{\sigma(w)}\right)_2\mathbb{1}_{\gcd(\sigma\tau(w),z)=1}. \tag{4.13}$$

This twisted multiplicativity formula for the symbol $[\cdot]$ is what makes the estimate in Proposition 4.3.8 possible; it is analogous to [23, Lemma 20.1, p. 1021], [24, (3.8), p. 708], [58, Proposition 8, p. 1010], and [41, (4.1), p. 19].

Let χ be a Dirichlet character modulo 8, and let $\{a(\chi)_\mathfrak{n}\}_\mathfrak{n}$ be the sequence defined in (4.3). Let $\{\alpha_\mathfrak{m}\}_\mathfrak{m}$ and $\{\beta_\mathfrak{n}\}_\mathfrak{n}$ be any two bounded sequences of complex numbers. Since each ideal of $\mathcal{O}_M$ has 8 different generators in $\mathcal{D}$, we have

$$\sum_{\mathrm{N}(\mathfrak{m})\le M}\sum_{\mathrm{N}(\mathfrak{n})\le N}\alpha_\mathfrak{m}\beta_\mathfrak{n}a(\chi)_{\mathfrak{mn}} = \frac{1}{8^2}\sum_{w\in\mathcal{D}(M)}\sum_{z\in\mathcal{D}(N)}\alpha_w\beta_z([wz]_\chi + [\varepsilon wz]_\chi).$$

Here $\varepsilon = 1+\sqrt{2}$, $\alpha_w := \alpha_{(w)}$ and $\beta_z := \beta_{(z)}$. Note that for any odd element $\alpha \in \mathcal{O}_M$, we have $[\alpha]_\chi = \mu_4 \cdot [\alpha]$ for some $\mu_4 \in \{\pm 1, \pm i\}$ that depends only on the congruence class of α modulo 8 (and so also modulo 16). Also note that (4.13) implies that $[\varepsilon wz] = \mu_5[wz]$ for some $\mu_5 \in \{\pm 1, \pm i\}$ that depends only on the congruence class of wz modulo 16. Hence, by restricting w and z to congruence classes modulo 16, we may break up the sum above into $2 \cdot 16^2$ sums of the shape

$$\mu_6 \sum_{\substack{w \in \mathcal{D}(M) \\ w \equiv \omega \bmod 16}} \sum_{\substack{z \in \mathcal{D}(N) \\ z \equiv \zeta \bmod 16}} \alpha_w \beta_z [wz],$$

where $\mu_6 \in \{\pm 1, \pm i\}$ depends only on the congruence classes ω and ζ modulo 16. Again by (4.13), we can replace α_w and β_z by $\alpha_w[w]$ and $\beta_z[z]$ to arrive at the sum

$$\mu_7 \sum_{\substack{w \in \mathcal{D}(M) \\ w \equiv \omega \bmod 16}} \sum_{\substack{z \in \mathcal{D}(N) \\ z \equiv \zeta \bmod 16}} \alpha_w \beta_z \left(\frac{z}{\sigma(w)}\right)_2 \mathbb{1}_{\gcd(\sigma\tau(w), z) = 1},$$

where $\mu_7 \in \{\pm 1, \pm i\}$ depends only on ω and ζ. One can now apply Proposition 4.3.6 with $\gamma(w, z) = \left(\frac{z}{\sigma(w)}\right)_2 \mathbb{1}_{\gcd(\sigma\tau(w), z) = 1}$ (and $F = \mathbb{Q}(\zeta_8)$, $n = 4$, $\mathfrak{f} = (16)$). Indeed, property (P1) follows from Lemma 4.3.1, while properties (P2) and (P3) follow from basic properties of the quadratic residue symbol in $\mathbb{Q}(\zeta_8)$. This finishes the proof of Proposition 4.3.8.

Chapter 5

The 16-rank of $\mathbb{Q}(\sqrt{-p})$

Abstract

Recently, a density result for the 16-rank of $\mathrm{Cl}(\mathbb{Q}(\sqrt{-p}))$ was established when p varies among the prime numbers, assuming a short character sum conjecture. In this chapter we prove the same density result unconditionally.

5.1 Introduction

If K is a quadratic number field with narrow class group $\mathrm{Cl}(K)$, there is an explicit description of $\mathrm{Cl}(K)[2]$ due to Gauss. Since then the class group of quadratic number fields has been extensively studied. If one is interested in the 2-part of the class group, i.e. $\mathrm{Cl}(K)[2^\infty]$, the explicit description of $\mathrm{Cl}(K)[2]$ is often very useful. It is for this reason that our current understanding of the 2-part of the class group is much better than the p-part for odd p.

In 1984, Cohen and Lenstra put forward conjectures regarding the average behavior of the class group $\mathrm{Cl}(K)$ of imaginary and real quadratic fields K. Despite significant effort, there has been relatively little progress in proving these conjectures. Almost all major results are about the 2-part with the most notable exception being the classical result of Davenport and Heilbronn [14] regarding the distribution of $\mathrm{Cl}(K)[3]$. Very little is known about $\mathrm{Cl}(K)[p]$ for $p > 3$. The non-abelian version of Cohen-Lenstra has recently also attracted great interest, see [1], [2], [39] and [80].

Gerth [27] studied the distribution of $2\mathrm{Cl}(K)[4]$, when the number of prime factors of the discriminant of K is fixed. Fouvry and Klüners [20] computed all the moments of $2\mathrm{Cl}(K)[4]$, when K varies among imaginary or real quadratic fields. In the paper [19], they deduced the probability that the 4-rank of a quadratic field has a given value. Their work was based on earlier ideas of Heath-Brown [33].

The study of $\mathrm{Cl}(K)[2^\infty]$ has often been conducted through the lens of *governing fields*. Let $k \geq 1$ be an integer and let d be an integer with $d \not\equiv 2 \bmod 4$. For a finite abelian group A we define the 2^k-rank of A to be $\mathrm{rk}_{2^k} A := \dim_{\mathbb{F}_2} 2^{k-1}A/2^kA$. Then a governing field $M_{d,k}$ is a normal field extension of $\mathbb{Q}$ such that

$$\mathrm{rk}_{2^k}\mathrm{Cl}\left(\mathbb{Q}\left(\sqrt{dp}\right)\right)$$

is determined by the splitting of p in $M_{d,k}$. Cohn and Lagarias [11] were the first to define the concept of a governing field, and conjectured that they always exist.

If $k \leq 3$, then governing fields are known to exist for all values of d. In case $k = 2$ this follows from work of Rédei [62] and Stevenhagen dealt with the case $k = 3$ [70]. The topic was recently revisited by Smith [68], who found a very explicit description for $M_{d,3}$ for most values of d. He then used this description to prove density results for $4\mathrm{Cl}(K)[8]$ assuming GRH. Not much later Smith [69] introduced *relative governing fields*, which allowed him to prove the most impressive result that $2\mathrm{Cl}(K)[2^\infty]$ has the expected distribution when K varies among all imaginary quadratic fields.

If we let $P(d, k)$ be the statement that a governing field $M_{d,k}$ exists, then there is currently not a single value of d for which the truth or falsehood of $P(d, 4)$ is known. This has been the most significant obstruction in proving density results for the 16-rank in thin families of the shape $\left\{\mathbb{Q}\left(\sqrt{dp}\right)\right\}_{p \text{ prime}}$.

This barrier was first broken by Milovic [58], who dealt with the 16-rank in the family $\left\{\mathbb{Q}\left(\sqrt{-2p}\right)\right\}_{p\equiv -1 \bmod 4}$. Milovic proves his density result with Vinogradov's method, and does not rely on the existence of a governing field. His use of Vinogradov's method was inspired by work of Friedlander et al. [24], which is based on earlier work of Friedlander and Iwaniec [23].

Milovic and the author established density results for the families $\left\{\mathbb{Q}\left(\sqrt{-2p}\right)\right\}_{p\equiv 1 \bmod 4}$ and $\{\mathbb{Q}(\sqrt{-p})\}_p$, see respectively [43] and [41] with the latter work being conditional on a short character sum conjecture. Both [43] and [41] follow the ideas of [24] closely in their treatment of the sums of type I, see Section 5.3 for a definition. However, if one applies the method of [24] to a number field of degree n, one is naturally lead to consider character sums of modulus q and length $q^{\frac{1}{n}}$.

In [43] we apply the method from [24] to a number field of degree 4. This leads to character sums just outside the range of Burgess' bound. Fortunately, the lemmas in Section 3.2 of [43] allow us to reduce the size of the modulus from q to $q^{\frac{1}{2}}$, and this enables us to deal with the sums of type I unconditionally. In [41] we use a criterion for the 16-rank of $\mathbb{Q}(\sqrt{-p})$ due to Bruin and Hemenway [7], and this criterion is stated most naturally over $\mathbb{Q}\left(\zeta_8, \sqrt{1+i}\right)$, which has degree 8. The resulting character sums are far outside the reach of Burgess' bound and we resort to assuming a short character sum conjecture, see [41, p. 8].

In this chapter we manage to deal with the 16-rank of $\mathbb{Q}(\sqrt{-p})$ unconditionally by using a criterion of Leonard and Williams [53], which one can naturally state over $\mathbb{Q}(\zeta_8)$. However, the Leonard and Williams criterion has the significant downside that it is the product of two residue symbols instead of one residue symbol, namely a quadratic and a quartic residue symbol. The resulting sums of type I can still not be treated unconditionally with the method from [24]. Instead, we use a rather ad hoc argument to deal with the resulting character sum.

Theorem 5.1.1. *Let $h(-p)$ be the class number of $\mathbb{Q}(\sqrt{-p})$. Then*

$$\lim_{X\to\infty} \frac{|\{p \text{ prime} : p \leq X \text{ and } 16 \mid h(-p)\}|}{|\{p \text{ prime} : p \leq X\}|} = \frac{1}{16}.$$

Milovic [57] has previously shown that there are infinitely many primes p with 16 dividing $h(-p)$. Theorem 5.1.1 gives an affirmative answer to conjectures in both [12] and [71]. For p a prime number, we define e_p by

$$e_p := \begin{cases} 1 & \text{if } 16 \mid h(-p) \\ -1 & \text{if } 8 \mid h(-p), 16 \nmid h(-p) \\ 0 & \text{otherwise.} \end{cases} \tag{5.1}$$

Theorem 5.1.1 is an immediate consequence of the following theorem.

Theorem 5.1.2. *We have*

$$\sum_{p\leq X} e_p \ll \frac{X}{\exp\left((\log X)^{0.1}\right)}.$$

It is natural to wonder if the other conditional results in [41] can be proven unconditionally using the methods from this chapter. This is likely to be the case, but it would require some effort to obtain suitable algebraic results similar to the Leonard and Williams [53] criterion used in this chapter.

Acknowledgements

I am very grateful to Djordjo Milovic for his support during this project. I would also like to thank Jan-Hendrik Evertse for proofreading.

5.2 Preliminaries

5.2.1 Quadratic and quartic reciprocity

Let K be a number field with ring of integers O_K. We say that an ideal $\mathfrak{n}$ of O_K is odd if $(\mathfrak{n}, 2) = (1)$. Similarly, we say that an element w of O_K is odd if the ideal generated by w is odd. If $\mathfrak{p}$ is an odd prime ideal of O_K and $\alpha \in O_K$, we define the quadratic residue symbol

$$\left(\frac{\alpha}{\mathfrak{p}}\right)_{2,K} := \begin{cases} 1 & \text{if } \alpha \notin \mathfrak{p} \text{ and } \alpha \equiv \beta^2 \bmod \mathfrak{p} \text{ for some } \beta \in O_K \\ -1 & \text{if } \alpha \notin \mathfrak{p} \text{ and } \alpha \not\equiv \beta^2 \bmod \mathfrak{p} \text{ for all } \beta \in O_K \\ 0 & \text{if } \alpha \in \mathfrak{p}. \end{cases}$$

Then Euler's criterion states

$$\left(\frac{\alpha}{\mathfrak{p}}\right)_{2,K} \equiv \alpha^{\frac{\mathrm{N}(\mathfrak{p})-1}{2}} \bmod \mathfrak{p}.$$

For a general odd ideal $\mathfrak{n}$ of O_K, we define

$$\left(\frac{\alpha}{\mathfrak{n}}\right)_{2,K} := \prod_{\mathfrak{p}^e \| \mathfrak{n}} \left(\left(\frac{\alpha}{\mathfrak{p}}\right)_{2,K}\right)^e.$$

Furthermore, for odd $\beta \in O_K$ we set

$$\left(\frac{\alpha}{\beta}\right)_{2,K} := \left(\frac{\alpha}{(\beta)}\right)_{2,K}.$$

We say that an element $\alpha \in K$ is totally positive if for all embeddings σ of K into $\mathbb{R}$ we have $\sigma(\alpha) > 0$. In particular, all elements of a totally complex number field are totally positive. We will make extensive use of the law of quadratic reciprocity.

Theorem 5.2.1. *Let $\alpha, \beta \in O_K$ be odd. If α or β is totally positive, we have*

$$\left(\frac{\alpha}{\beta}\right)_{2,K} = \mu(\alpha, \beta) \left(\frac{\beta}{\alpha}\right)_{2,K},$$

where $\mu(\alpha, \beta) \in \{\pm 1\}$ depends only on the congruence classes of α and β modulo 8.

Proof. This follows from Lemma 2.1 of [24]. □

If $K = \mathbb{Q}$, we shall drop the subscript. In this case the symbol $\left(\frac{\cdot}{\cdot}\right)$ is to be interpreted as the Kronecker symbol, which is an extension of the quadratic residue symbol to allow for even arguments in the bottom. We presume that the reader is familiar with the quadratic reciprocity law for the Kronecker symbol. Now let K be a number field containing $\mathbb{Q}(i)$ still with ring of integers O_K. For $\alpha \in O_K$ and $\mathfrak{p}$ an odd prime ideal of O_K, we define the quartic residue symbol $(\alpha/\mathfrak{p})_{4,K}$ to be the unique element in $\{\pm 1, \pm i, 0\}$ such that

$$\left(\frac{\alpha}{\mathfrak{p}}\right)_{4,K} \equiv \alpha^{\frac{\mathrm{N}(\mathfrak{p})-1}{4}} \bmod \mathfrak{p}.$$

We extend the quartic residue symbol to all odd ideals $\mathfrak{n}$ and then to all odd elements β in the same way as the quadratic residue symbol. Then we have the following theorem.

Theorem 5.2.2. *Let $\alpha, \beta \in \mathbb{Z}[\zeta_8]$ with β odd. Then for fixed α, the symbol $(\alpha/\beta)_{4,\mathbb{Q}(\zeta_8)}$ depends only on β modulo $16\alpha\mathbb{Z}[\zeta_8]$. Furthermore, if α is also odd, we have*

$$\left(\frac{\alpha}{\beta}\right)_{4,\mathbb{Q}(\zeta_8)} = \mu(\alpha, \beta) \left(\frac{\beta}{\alpha}\right)_{4,\mathbb{Q}(\zeta_8)},$$

where $\mu(\alpha, \beta) \in \{\pm 1, \pm i\}$ depends only on the congruence classes of α and β modulo 16.

Proof. Use Proposition 6.11 of Lemmermeyer [50, p. 199]. □

5.2.2 A fundamental domain

Let F be a number field of degree n over $\mathbb{Q}$ and let O_F be its ring of integers. Suppose that F has r real embeddings and s pairs of conjugate complex embeddings so that $r+2s=n$. Define T to be the torsion subgroup of O_F^*. Then, by Dirichlet's Unit Theorem, there exists a free abelian group $V \subseteq O_F^*$ of rank $r+s-1$ with $O_F^* = T \times V$. Fix one choice of such a V.

There is a natural action of V on O_F. The goal of this subsection is to construct a fundamental domain $\mathcal{D}$ for this action. Such a fundamental domain allows us to transform a sum over ideals into a sum over elements. It will be important that the resulting fundamental domain has nice geometrical properties, so that we have good control over the elements we are summing.

Fix an integral basis $\omega = \{\omega_1, \ldots, \omega_n\}$ for O_F. Then we get an isomorphism of $\mathbb{Q}$-vector spaces $i_\omega : \mathbb{Q}^n \to F$, where i_ω is given by $(a_1, \ldots, a_n) \mapsto a_1\omega_1 + \ldots + a_n\omega_n$. For a subset $S \subseteq \mathbb{R}^n$ and an element $\alpha \in F$, we will say that $\alpha \in S$ if $i_\omega^{-1}(\alpha) \in S$. Define for our integral basis ω and a real number $X > 0$

$$B(X,\omega) := \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n : \left| \prod_{i=1}^{n} \left(x_1\sigma_i(\omega_1) + \ldots + x_n\sigma_i(\omega_n)\right) \right| \leq X \right\},$$

where $\sigma_1, \ldots, \sigma_n$ are the embeddings of F into $\mathbb{C}$.

Lemma 5.2.3. *Let F be a number field with ring of integers O_F and integral basis $\omega = \{\omega_1, \ldots, \omega_n\}$. Choose a splitting $O_F^* = T \times V$, where T is the torsion subgroup of O_F^*. There exists a subset $\mathcal{D} \subseteq \mathbb{R}^n$ such that*

(i) for all $\alpha \in O_F \setminus \{0\}$, there exists a unique $v \in V$ such that $v\alpha \in \mathcal{D}$. Furthermore, we have the equality

$$\{u \in O_F^* : u\alpha \in \mathcal{D}\} = \{tv : t \in T\};$$

(ii) $\mathcal{D} \cap B(1,\omega)$ has an $(n-1)$-Lipschitz parametrizable boundary;

(iii) there is a constant $C(\omega)$ depending only on ω such that for all $\alpha \in \mathcal{D}$ we have $|a_i| \leq C(\omega) \cdot |\mathrm{N}(\alpha)|^{\frac{1}{n}}$, where $a_i \in \mathbb{Z}$ are such that $\alpha = a_1\omega_1 + \ldots + a_n\omega_n$.

Proof. This is Lemma 3.5 of [43]. □

We will use Lemma 5.2.3 for $F := \mathbb{Q}(\zeta_8)$; in order to do so we must make some choices. We choose $V := \langle 1 + \sqrt{2} \rangle$ and integral basis $\omega := \{1, \zeta_8, \zeta_8^2, \zeta_8^3\}$. The resulting fundamental domain will be called $\mathcal{D}$, and we define $\mathcal{D}(X) := \mathcal{D} \cap B(X, \omega)$.

5.3 The sieve

Let $\{a_p\}$ be a sequence of complex numbers indexed by the primes and define

$$S(X) := \sum_{p \leq X} a_p.$$

To prove our main theorem, we must prove oscillation of $S(X)$ for the specific sequence $\{e_p\}$ defined in equation (5.1). There are relatively few methods that can deal with such sums. The most common approach is to attach an L-function and then use the zero-free region. This approach requires that our sequence $\{e_p\}$ has good multiplicative properties. It turns out that $\{e_p\}$ is instead twisted multiplicative (see Lemma 5.6.1 and Lemma 5.6.3), and this suggests we use Vinogradov's method instead.

Recall that $h(-p)$ denotes the class number of $\mathrm{Cl}(\mathbb{Q}(\sqrt{-p}))$. By definition of e_p we have $e_p = 0$ if and only if $8 \nmid h(-p)$. It is well-known that $\mathbb{Q}(\zeta_8, \sqrt{1+i})$ is a *governing field* for the 8-rank of $\mathrm{Cl}(\mathbb{Q}(\sqrt{-p}))$, in fact a prime p splits completely in $\mathbb{Q}(\zeta_8, \sqrt{1+i})$ if and only if $8 \mid h(-p)$. This is extremely convenient. Indeed, if we apply Vinogradov's method to our governing field, primes of degree 1 will give the dominant contribution and these primes automatically have $e_p \neq 0$.

Unfortunately, $\mathbb{Q}(\zeta_8, \sqrt{1+i})$ is a field of degree 8, which is simply too large to make our analytic methods work unconditionally. Indeed, using the same approach for the sums of type I as [24], one ends up with short character sums of modulus q and length roughly $q^{\frac{1}{8}}$, which is far outside the reach of Burgess' famous bound. However, assuming a short character sum conjecture, one still obtains the desired oscillation and this is the approach taken in [41]. Instead we work over $\mathbb{Q}(\zeta_8)$; fortunately, $\mathbb{Q}(\zeta_8, \sqrt{1+i})$ is an abelian extension of $\mathbb{Q}(\zeta_8)$, which implies that the splitting of a prime $\mathfrak{p}$ of $\mathbb{Q}(\zeta_8)$ in the extension $\mathbb{Q}(\zeta_8, \sqrt{1+i})/\mathbb{Q}(\zeta_8)$ is determined by a congruence condition. Such a congruence condition can easily be incorporated in Vinogradov's method.

We will follow Section 5 of Friedlander et al. [24], who adapted Vinogradov's method to number fields. Define

$$\Lambda(\mathfrak{n}) := \begin{cases} \log \mathrm{N}\mathfrak{p} & \text{if } \mathfrak{n} = \mathfrak{p}^l \\ 0 & \text{otherwise} \end{cases}$$

and suppose that we want to prove oscillation of

$$S(X) := \sum_{\mathrm{N}\mathfrak{n} \leq X} a_{\mathfrak{n}} \Lambda(\mathfrak{n}),$$

where $a_{\mathfrak{n}}$ is of absolute value at most 1. The power of Vinogradov's method lies in the fact that one does not have to deal with $S(X)$ directly. Instead one has to prove cancellation of

$$A(X, \mathfrak{d}) := \sum_{\substack{\mathrm{N}\mathfrak{n} \leq X \\ \mathfrak{d} \mid \mathfrak{n}}} a_{\mathfrak{n}},$$

which are traditionally called sums of type I or linear sums, and

$$B(M,N) := \sum_{\mathrm{N}\mathfrak{m}\leq M} \sum_{\mathrm{N}\mathfrak{n}\leq N} \alpha_{\mathfrak{m}}\beta_{\mathfrak{n}}a_{\mathfrak{m}\mathfrak{n}},$$

which are traditionally called sums of type II or bilinear sums. It is important to remark that $S(X)$ depends only on $a_{\mathfrak{n}}$ with $\mathfrak{n}$ a prime power, while $A(X,\mathfrak{d})$ and $B(M,N)$ certainly do not. This gives a substantial amount of flexibility, since we may define $a_{\mathfrak{n}}$ on composite ideals $\mathfrak{n}$ in any way we like provided that we can prove oscillation of $A(X,\mathfrak{d})$ and $B(M,N)$. Constructing a suitable sequence $a_{\mathfrak{n}}$ will be the goal of Section 5.4. We are now ready to state the precise version of Vinogradov's method we are going to use.

Proposition 5.3.1. *Let F be a number field and let $a_{\mathfrak{n}}$ be a sequence of complex numbers, indexed by the ideals of O_F, with $|a_{\mathfrak{n}}| \leq 1$. If $0 < \theta_1, \theta_2 < 1$ and $\theta_3 > 0$ are such that*

- *we have for all ideals $\mathfrak{d}$ of O_F*

$$A(X,\mathfrak{d}) \ll_{F,a_{\mathfrak{n}},\theta_1} \frac{X}{\exp\left((\log X)^{\theta_1}\right)}; \tag{5.2}$$

- *we have for all sequences of complex numbers $\{\alpha_{\mathfrak{m}}\}$ and $\{\beta_{\mathfrak{n}}\}$ of absolute value at most 1*

$$B(M,N) \ll_{F,a_{\mathfrak{n}},\theta_2} (M+N)^{\theta_2}(MN)^{1-\theta_2}(\log MN)^{\theta_3}. \tag{5.3}$$

Then we have for all $c < \theta_1$

$$S(X) \ll_{c,F,a_{\mathfrak{n}},\theta_1,\theta_2,\theta_3} \frac{X}{\exp\left((\log X)^{c}\right)}.$$

Proof. This quickly follows from Proposition 5.1 of [24]. □

The remainder of this chapter is devoted to the three major tasks that are left. We start by constructing a suitable sequence $a_{\mathfrak{n}}$ in Section 5.4 to which we will apply Proposition 5.3.1 with $F = \mathbb{Q}(\zeta_8)$. The main result of Section 5.5 is Proposition 5.5.1, which proves equation (5.2) for $\theta_1 = 0.2$. Finally, we prove in Section 5.6 that (5.3) holds with $\theta_2 = \frac{1}{24}$; this is the content of Proposition 5.6.6. Once we have proven Proposition 5.5.1 and Proposition 5.6.6, the proof of Theorem 5.1.2 is complete.

5.4 Definition of the sequence

By Gauss genus theory we know that the 2-part of $\mathrm{Cl}(\mathbb{Q}(\sqrt{-p}))$ is cyclic, and the 2-part of $\mathrm{Cl}(\mathbb{Q}(\sqrt{-p}))$ is trivial if and only if $p \equiv 3 \bmod 4$. Let us recall a criterion for $16 \mid h(-p)$ due to Leonard and Williams [53]. We have

$$4 \mid h(-p) \Longleftrightarrow p \equiv 1 \bmod 8.$$

Now suppose that $4 \mid h(-p)$. There exist positive integers g and h satisfying

$$p = 2g^2 - h^2.$$

Then a classical result of Hasse [31] is

$$8 \mid h(-p) \Longleftrightarrow \left(\frac{g}{p}\right) = 1 \text{ and } p \equiv 1 \bmod 8$$

or equivalently

$$8 \mid h(-p) \Longleftrightarrow \left(\frac{-1}{g}\right) = 1 \text{ and } p \equiv 1 \bmod 8.$$

We are now ready to state the result of Leonard and Williams [53]. If p is a prime number with $8 \mid h(-p)$, we have

$$16 \mid h(-p) \Longleftrightarrow \left(\frac{g}{p}\right)_4 \left(\frac{2h}{g}\right) = 1.$$

With this in mind, we are going to define a sequence $\{a_{\mathfrak{n}}\}$, indexed by the integral ideals of $\mathbb{Z}[\zeta_8]$, such that for all unramified prime ideals $\mathfrak{p}$ in $\mathbb{Z}[\zeta_8]$ of norm p

$$a_{\mathfrak{p}} = \begin{cases} 1 & \text{if } 16 \mid h(-p) \\ -1 & \text{if } 8 \mid h(-p), 16 \nmid h(-p) \\ 0 & \text{otherwise.} \end{cases} \tag{5.4}$$

The sequence $\{a_{\mathfrak{n}}\}$ will be constructed in such a way that we can prove the two estimates in Proposition 5.5.1 and Proposition 5.6.6. Before we move on, it will be useful to recall some standard facts about $\mathbb{Z}[\zeta_8]$. The ring $\mathbb{Z}[\zeta_8]$ is a PID with unit group generated by ζ_8 and $\epsilon := 1 + \sqrt{2}$. Odd primes are unramified in $\mathbb{Z}[\zeta_8]$, while 2 is totally ramified. Furthermore, an odd prime p splits completely in $\mathbb{Z}[\zeta_8]$ if and only if $p \equiv 1 \bmod 8$ if and only if $4 \mid h(-p)$. We will make extensive use of the following field diagram.

$$\begin{tikzcd}
& M := \mathbb{Q}(\zeta_8) \arrow[dl, dash, "\langle\sigma\rangle"'] \arrow[d, dash, "\langle\sigma\tau\rangle"] \arrow[dr, dash, "\langle\tau\rangle"] & \\
\mathbb{Q}(i) \arrow[dr, dash] & \mathbb{Q}(i\sqrt{2}) \arrow[d, dash] & \mathbb{Q}(\sqrt{2}) \arrow[dl, dash] \\
& \mathbb{Q} &
\end{tikzcd}$$

If $\mathfrak{n}$ is not odd, we set $a_{\mathfrak{n}} := 0$. From now on $\mathfrak{n}$ is an odd, integral, non-zero ideal of $\mathbb{Z}[\zeta_8]$ and w is a generator of $\mathfrak{n}$. We can write w as

$$w = a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3$$

for certain $a, b, c, d \in \mathbb{Z}$. Define $u, v \in \mathbb{Z}$ by

$$w\tau(w) = u + v\sqrt{2}.$$

We can explicitly compute u and v using the following formulas

$$u = \frac{w\tau(w) + \sigma(w)\sigma\tau(w)}{2} = a^2 + b^2 + c^2 + d^2 \tag{5.5}$$

and

$$v = \frac{w\tau(w) - \sigma(w)\sigma\tau(w)}{2\sqrt{2}} = ab - ad + bc + cd. \tag{5.6}$$

Since w is odd, it follows that $\mathrm{N}w \equiv 1 \bmod 8$. Then it follows from

$$\mathrm{N}w = u^2 - 2v^2$$

that u is an odd integer and v is an even integer. Set

$$g := u + v, \quad h := u + 2v,$$

so that g is an odd integer and h is an odd integer, not necessarily positive. We claim that g is positive. Indeed

$$\begin{aligned} g &= a^2 + b^2 + c^2 + d^2 + ab - ad + bc + cd \\ &= \frac{1}{2}(a+b)^2 + \frac{1}{2}(a-d)^2 + \frac{1}{2}(b+c)^2 + \frac{1}{2}(c+d)^2 > 0. \end{aligned}$$

By construction g and h satisfy

$$\mathrm{N}w = 2g^2 - h^2.$$

We start by showing that the value of

$$\left(\frac{-1}{g}\right) \tag{5.7}$$

does not depend on the choice of generator w of our ideal $\mathfrak{n}$.

Lemma 5.4.1. *Let $\mathfrak{n}$ be an odd, integral ideal of $\mathbb{Z}[\zeta_8]$. Then the value of equation (5.7) is the same for all generators w of $\mathfrak{n}$.*

Proof. Suppose that we replace w by $\zeta_8 w$. Because $\zeta_8\tau(\zeta_8) = 1$, it follows that u and v, hence also g, do not change. Suppose instead that we replace w by ϵw. In this case u becomes $3u + 4v$ and v becomes $2u + 3v$, so g becomes $5u + 7v$. Hence our lemma boils down to

$$\left(\frac{-1}{u+v}\right) = \left(\frac{-1}{5u+7v}\right),$$

which holds if and only if

$$u + v \equiv 5u + 7v \bmod 4.$$

But recall that v is even by our assumption that w is odd. $\square$

We define for odd $w \in \mathbb{Z}[\zeta_8]$ the following symbol

$$[w] := \left(\frac{g}{w}\right)_{4,M} \left(\frac{2h}{g}\right),$$

where we remind the reader that M is defined to be $\mathbb{Q}(\zeta_8)$. We express this as

$$[w] = [w]_1[w]_2, \quad [w]_1 := \left(\frac{g}{w}\right)_{4,M}, \quad [w]_2 := \left(\frac{2h}{g}\right).$$

It is easily checked that $[\zeta_8 w] = [w]$. Unfortunately, it is not always true that $[\epsilon w] = [w]$. To get around this, we need the following lemma.

Lemma 5.4.2. *We have for all odd w*

$$[\epsilon^4 w] = [w].$$

Proof. We have for any odd w

$$[w]_1 = \left(\frac{g}{w}\right)_{4,M} = \left(\frac{u+v}{w}\right)_{4,M} = \left(\frac{\left(\frac{1}{2} - \frac{1}{2\sqrt{2}}\right)\sigma(w)\sigma\tau(w)}{w}\right)_{4,M}, \tag{5.8}$$

where we use the explicit formulas for u and v, see equation (5.5) and equation (5.6), in terms of w. From this expression it quickly follows that $[\epsilon^2 w]_1 = [w]_1$. We also have

$$\begin{aligned} [w]_2 &= \left(\frac{2h}{g}\right) = \left(\frac{2u+4v}{u+v}\right) = \left(\frac{2}{u+v}\right)\left(\frac{v}{u+v}\right) \\ &= \left(\frac{2}{u+v}\right)\left(\frac{-u}{u+v}\right) = \left(\frac{-2}{u+v}\right)\left(\frac{v}{u}\right)(-1)^{\frac{u-1}{2}\cdot\frac{u+v-1}{2}}. \end{aligned} \tag{5.9}$$

A straightforward computation shows that the u and v associated to $\epsilon^4 w$ are respectively $u_1 := 577u + 816v$ and $v_1 := 408u + 577v$. Then we have

$$\left(\frac{v}{u}\right) = \left(\frac{408u+577v}{577u+816v}\right) = \left(\frac{v_1}{u_1}\right) \tag{5.10}$$

due to Proposition 2 in Milovic [58]. It will be useful to observe that the following congruences hold true

$$u \equiv u_1 \bmod 8, \quad v \equiv v_1 \bmod 8.$$

This immediately implies

$$\left(\frac{-2}{u+v}\right) = \left(\frac{-2}{u_1+v_1}\right), \tag{5.11}$$

and therefore the lemma. $\square$

With this out of the way, we have all the tools necessary to define $a_{\mathfrak{n}}$. Suppose that $\mathfrak{n}$ is an odd, integral ideal of $\mathbb{Z}[\zeta_8]$ with generator w. Then we define

$$a_{\mathfrak{n}} := \begin{cases} \frac{1}{4}\left([w] + [\epsilon w] + [\epsilon^2 w] + [\epsilon^3 w]\right) & \text{if } w \text{ satisfies (5.7)} \\ 0 & \text{otherwise.} \end{cases} \tag{5.12}$$

for any generator w of $\mathfrak{n}$. Here we say that w satisfies equation (5.7) if $(-1/g) = 1$, where g is defined in terms of w as above. Then an application of Lemma 5.4.1 and Lemma 5.4.2 shows that (5.12) is indeed well-defined.

Lemma 5.4.3. *The sequence $a_{\mathfrak{n}}$ satisfies equation (5.4) for all unramified prime ideals $\mathfrak{p}$ of degree 1 in $\mathbb{Z}[\zeta_8]$.*

Proof. Let $\mathfrak{p}$ be an unramified prime ideal of degree 1 in $\mathbb{Z}[\zeta_8]$ and let w be a generator of $\mathfrak{p}$. Put $p := \mathrm{N}w$. Lemma 5.4.1 and the aforementioned result of Hasse imply

$$w \text{ does not satisfy (5.7)} \Longleftrightarrow 8 \nmid h(-p),$$

and $a_{\mathfrak{p}}$ is indeed 0 in this case. Now suppose that w does satisfy (5.7). Recall that

$$[w] = \left(\frac{g}{w}\right)_{4,M} \left(\frac{2h}{g}\right),$$

where g and h are explicit functions of w. We stress that these g and h are not necessarily the same g and h from Leonard and Williams. Indeed, Leonard and Williams require g and h to be positive, while our h is not necessarily positive. However, since w satisfies (5.7), their criterion remains valid irrespective of the sign of h. Then, the criterion implies

$$[w] = [\epsilon w] = [\epsilon^2 w] = [\epsilon^3 w].$$

Furthermore, the criterion also shows that

$$[w] = 1 \Longleftrightarrow 16 \mid h(-p).$$

This completes the proof of our lemma. $\square$

5.5 Sums of type I

The goal of this section is to bound the following sum

$$A(X, \mathfrak{d}) = \sum_{\substack{\mathrm{N}\mathfrak{n} \leq X \\ \mathfrak{d} | \mathfrak{n}}} a_{\mathfrak{n}} = \sum_{\substack{\mathrm{N}\mathfrak{n} \leq X \\ \mathfrak{d} | \mathfrak{n}, \mathfrak{n} \text{ odd}}} a_{\mathfrak{n}}.$$

By picking a generator for $\mathfrak{n}$ we obtain

$$A(X, \mathfrak{d}) = \frac{1}{8} \sum_{\substack{w \in \mathcal{D}(X) \\ w \equiv 0 \bmod \mathfrak{d} \\ w \text{ odd}}} a_{(w)} = \frac{1}{32} \sum_{\substack{w \in \mathcal{D}(X) \\ w \equiv 0 \bmod \mathfrak{d} \\ w \text{ odd}}} \mathbf{1}_{w \text{ sat. } (5.7)} \left([w] + [\epsilon w] + [\epsilon^2 w] + [\epsilon^3 w]\right).$$

We define for $i = 0, \dots, 3$ and ρ an invertible congruence class modulo 2^{10}

$$A(X, \mathfrak{d}, u_i, \rho) := \sum_{\substack{w \in u_i \mathcal{D}(X) \\ w \equiv 0 \bmod \mathfrak{d} \\ w \equiv \rho \bmod 2^{10}}} [w] = \sum_{\substack{w \in u_i \mathcal{D}(X) \\ w \equiv 0 \bmod \mathfrak{d} \\ w \equiv \rho \bmod 2^{10}}} \left(\frac{g}{w}\right)_{4,M} \left(\frac{2h}{g}\right),$$

where $u_i := \epsilon^i$. With this definition in place, we may split $A(X, \mathfrak{d})$ as follows

$$A(X, \mathfrak{d}) = \frac{1}{32} \sum_{i=0}^{3} \sum_{\rho \in (O_M/2^{10}O_M)^*} \mathbf{1}_{\rho \text{ sat. } (5.7)} A(X, \mathfrak{d}, u_i, \rho),$$

since the truth of equation (5.7) depends only on w modulo 4. Then it is enough to bound each individual sum $A(X, \mathfrak{d}, u_i, \rho)$. In order to bound this sum, our first step is to carefully rewrite the symbol $[w]$ in a more tractable form. While doing so, we will find some hidden cancellation between $[w]_1$ and $[w]_2$ that is vital for making our results unconditional.

Throughout this section we use the convention that $\mu(\cdot) \in \{\pm 1, \pm i\}$ is a function depending only on the variables between the parentheses; at each occurence $\mu(\cdot)$ may be a different function. Since our cancellation will come from fixing b, c and d while varying a, factors of the shape $\mu(\rho, b, c, d)$ will present no issues for us. Let us start by rewriting $[w]_2$. It follows from equation (5.9) that

$$\left(\frac{2h}{g}\right) = \left(\frac{v}{u}\right) \mu(\rho). \tag{5.13}$$

Using the formulas for u and v we get

$$\left(\frac{v}{u}\right) = \left(\frac{ab - ad + bc + cd}{a^2 + b^2 + c^2 + d^2}\right). \tag{5.14}$$

If v is not zero, we can uniquely factor v as

$$v := v_1 v_2 t,$$

where v_1 is an odd, positive integer satisfying $\gcd(v_1, b - d) = 1$, v_2 is an odd integer consisting only of primes dividing $b - d$ and t is positive and only divisible by powers of 2. Then we have

$$\left(\frac{ab - ad + bc + cd}{a^2 + b^2 + c^2 + d^2}\right) = \left(\frac{v_1}{a^2 + b^2 + c^2 + d^2}\right) \left(\frac{tv_2}{a^2 + b^2 + c^2 + d^2}\right). \tag{5.15}$$

Let ρ' be the congruence class of v_1 modulo 8. Using the following identity modulo v

$$a^2(b - d)^2 \equiv c^2(b + d)^2 \bmod v$$

and the fact that this identity continues to hold for any divisor of v, so in particular for v_1, we rewrite the first factor of equation (5.15) as follows

$$\begin{aligned}\left(\frac{v_1}{a^2+b^2+c^2+d^2}\right) &= \mu(\rho,\rho')\left(\frac{a^2+b^2+c^2+d^2}{v_1}\right)\\ &= \mu(\rho,\rho')\left(\frac{(a^2+b^2+c^2+d^2)(b-d)^2}{v_1}\right)\\ &= \mu(\rho,\rho')\left(\frac{a^2(b-d)^2+(b^2+c^2+d^2)(b-d)^2}{v_1}\right)\\ &= \mu(\rho,\rho')\left(\frac{c^2(b+d)^2+(b^2+c^2+d^2)(b-d)^2}{v_1}\right)\\ &= \mu(\rho,\rho')\left(\frac{(b^2+d^2)(2c^2+(b-d)^2)}{v_1}\right). \end{aligned}\tag{5.16}$$

Stringing together (5.13), (5.14), (5.15) and (5.16), we conclude that

$$\left(\frac{2h}{g}\right) = \mu(\rho,\rho')\left(\frac{(b^2+d^2)(2c^2+(b-d)^2)}{v_1}\right)\left(\frac{tv_2}{a^2+b^2+c^2+d^2}\right). \tag{5.17}$$

Our next goal is to simplify $[w]_1$. We have by equation (5.8) and Theorem 5.2.2

$$\left(\frac{g}{w}\right)_{4,M} = \left(\frac{\left(\frac{1}{2}-\frac{1}{2\sqrt{2}}\right)\sigma(w)\sigma\tau(w)}{w}\right)_{4,M} = \mu(\rho)\left(\frac{\sigma(w)\sigma\tau(w)}{w}\right)_{4,M}. \tag{5.18}$$

The quartic residue symbol in equation (5.18) is the product of two quartic residue symbols. One of them is equal to

$$\begin{aligned}\left(\frac{\sigma\tau(w)}{w}\right)_{4,M} &= \left(\frac{a+d\zeta_8-c\zeta_8^2+b\zeta_8^3}{a+b\zeta_8+c\zeta_8^2+d\zeta_8^3}\right)_{4,M} = \left(\frac{-2c\zeta_8^2+(d-b)(\zeta_8-\zeta_8^3)}{a+b\zeta_8+c\zeta_8^2+d\zeta_8^3}\right)_{4,M}\\ &= \left(\frac{\zeta_8^2}{a+b\zeta_8+c\zeta_8^2+d\zeta_8^3}\right)_{4,M}\left(\frac{-2c+(b-d)(\zeta_8+\zeta_8^3)}{a+b\zeta_8+c\zeta_8^2+d\zeta_8^3}\right)_{4,M}\\ &= \mu(\rho)\left(\frac{-2c+(b-d)(\zeta_8+\zeta_8^3)}{a+b\zeta_8+c\zeta_8^2+d\zeta_8^3}\right)_{4,M}, \end{aligned}\tag{5.19}$$

where the last equality is due to Theorem 5.2.2. For the remainder of this section we assume that $b-d$ is not zero. We factor $-2c+(b-d)(\zeta_8+\zeta_8^3)$ in the ring $\mathbb{Z}[\sqrt{-2}]$ as

$$-2c+(b-d)(\zeta_8+\zeta_8^3) = \eta^4 e_0 e$$

with η and e_0 consisting only of even prime factors, e_0 not divisible by a non-trivial fourth power and e odd. This factorization is unique up to multiplication by units. Then we have by Theorem 5.2.2

$$\left(\frac{-2c+(b-d)(\zeta_8+\zeta_8^3)}{a+b\zeta_8+c\zeta_8^2+d\zeta_8^3}\right)_{4,M} = \mu(\rho,b,c,d)\left(\frac{a+b\zeta_8+c\zeta_8^2+d\zeta_8^3}{e}\right)_{4,M}. \tag{5.20}$$

But a simple computation shows

$$a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3 \equiv \sigma\tau(a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3) \bmod e.$$

Let $\mathfrak{p}$ be a prime in $\mathbb{Z}[\sqrt{-2}]$ that divides e. Then we may replace $a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3$ by some element in $\mathbb{Z}[\sqrt{-2}]$ by Lemma 3.4 of [43]. In case $\mathfrak{p}$ splits in M, we apply Lemma 3.2 of [43]. While if $\mathfrak{p}$ remains inert, we see that $\mathfrak{p}$ is of degree 1 and $\mathrm{N}\mathfrak{p} \equiv 3 \bmod 8$. In this case we apply Lemma 3.3 of [43]. Hence in all cases

$$\left(\frac{a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3}{\mathfrak{p}}\right)_{4,M} = \mathbb{1}_{\gcd(a+b\zeta_8+c\zeta_8^2+d\zeta_8^3,\mathfrak{p})=(1)}.$$

This yields

$$\left(\frac{a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3}{e}\right)_{4,M} = \mathbb{1}_{\gcd(w,\sigma\tau(w))=(1)}. \tag{5.21}$$

We deduce from equation (5.19), (5.20) and (5.21) that

$$\left(\frac{\sigma\tau(w)}{w}\right)_{4,M} = \mu(\rho, b, c, d)\mathbb{1}_{\gcd(w,\sigma\tau(w))=(1)}. \tag{5.22}$$

We will now study the other quartic residue symbol in equation (5.18) using very similar methods. We start with the identity

$$\begin{aligned}\left(\frac{\sigma(w)}{w}\right)_{4,M} &= \left(\frac{a - b\zeta_8 + c\zeta_8^2 - d\zeta_8^3}{a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3}\right)_{4,M} = \left(\frac{-2\zeta_8(b + d\zeta_8^2)}{a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3}\right)_{4,M} \\ &= \left(\frac{-2\zeta_8}{a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3}\right)_{4,M}\left(\frac{b + d\zeta_8^2}{a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3}\right)_{4,M} \\ &= \mu(\rho)\left(\frac{b + d\zeta_8^2}{a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3}\right)_{4,M},\end{aligned} \tag{5.23}$$

where we use Theorem 5.2.2 once more. We choose $i := \zeta_8^2$ and factor $b + di$ in the ring $\mathbb{Z}[i]$ as

$$b + di = \eta'^4 e_0' e'$$

with η' and e_0' consisting only of even prime factors, e_0' not divisible by a non-trivial fourth power and e' odd. Such a factorization is unique up to multiplication by units. With this factorization we have due to Theorem 5.2.2

$$\left(\frac{b + di}{a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3}\right)_{4,M} = \mu(\rho, b, c, d)\left(\frac{a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3}{e'}\right)_{4,M}. \tag{5.24}$$

We claim that

$$\left(\frac{a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3}{e'}\right)_{4,M} = \left(\frac{a + c\zeta_8^2}{e'}\right)_{4,M} = \left(\frac{a + ci}{e'}\right)_{2,\mathbb{Q}(i)}. \tag{5.25}$$

Indeed, let $\mathfrak{p}$ be a prime in $\mathbb{Z}[i]$ that divides e'. If $\mathfrak{p}$ splits in M, Lemma 3.2 of [43] shows that

$$\left(\frac{a + c\zeta_8^2}{\mathfrak{p}}\right)_{4,M} = \left(\frac{a + ci}{\mathfrak{p}}\right)_{2,\mathbb{Q}(i)}.$$

Suppose instead that $\mathfrak{p}$ remains inert. Then $\mathfrak{p}$ is of degree 1 and $\mathrm{N}\mathfrak{p} \equiv 5 \bmod 8$. Now we apply Lemma 3.3 of [43] to obtain

$$\left(\frac{a + c\zeta_8^2}{\mathfrak{p}}\right)_{4,M} = \left(\frac{a + ci}{\mathfrak{p}}\right)_{2,\mathbb{Q}(i)}.$$

This establishes our claim and hence equation (5.24). Combining (5.23), (5.24) and (5.25) acquires the validity of

$$\left(\frac{\sigma(w)}{w}\right)_{4,M} = \mu(\rho, b, c, d) \left(\frac{a + ci}{e'}\right)_{2,\mathbb{Q}(i)}. \tag{5.26}$$

Put

$$f(w, \rho) := \mu(\rho, \rho', b, c, d)\mathbb{1}_{\gcd(w,\sigma\tau(w))=(1)} \left(\frac{tv_2}{a^2 + b^2 + c^2 + d^2}\right).$$

Using (5.17), (5.22) and (5.26), we conclude that

$$\left(\frac{g}{w}\right)_{4,M} \left(\frac{2h}{g}\right) = f(w, \rho) \left(\frac{(b^2 + d^2)(2c^2 + (b - d)^2)}{v_1}\right) \left(\frac{a + ci}{e'}\right)_{2,\mathbb{Q}(i)}. \tag{5.27}$$

Our hidden cancellation will come from comparing the Jacobi symbols

$$\left(\frac{b^2 + d^2}{v_1}\right) \text{ and } \left(\frac{a + ci}{e'}\right)_{2,\mathbb{Q}(i)}.$$

Our goal is to show that these two Jacobi symbols are equal up to some easily controlled factors. We can uniquely factor

$$b^2 + d^2 = z_1 z_2,$$

where z_1 and z_2 are positive integers satisfying

- $(z_1, z_2) = 1$;
- z_1 odd and squarefree;
- if p is odd and divides z_2, then also p^2 divides z_2.

With this factorization we have

$$\left(\frac{b^2 + d^2}{v_1}\right) = \left(\frac{z_1}{v_1}\right) \left(\frac{z_2}{v_1}\right) = \mu(\rho', b, c, d) \left(\frac{v_1}{z_1}\right) \left(\frac{z_2}{v_1}\right).$$

In a similar vein we uniquely factor, up to multiplication by units, e' in $\mathbb{Z}[i]$ as

$$e' = \gamma_1 \gamma_2$$

with $(\mathrm{N}\gamma_1, \mathrm{N}\gamma_2) = (1)$, $\mathrm{N}\gamma_1$ squarefree and $\mathrm{N}\gamma_2$ squarefull. The point of this factorization is that $\mathrm{N}\gamma_1 = z_1$. This gives

$$\left(\frac{v_1}{z_1}\right) = \left(\frac{v_1}{\gamma_1}\right)_{2,\mathbb{Q}(i)}.$$

We claim that

$$(tv_2, \gamma_1) = (d, \gamma_1) = (1). \tag{5.28}$$

We clearly have $(t, \gamma_1) = (1)$, so we first show that $(v_2, \gamma_1) = (1)$. Let $\mathfrak{p}$ be an odd prime of $\mathbb{Z}[i]$ above p such that $\mathfrak{p} \mid v_2$ and $\mathfrak{p} \mid \gamma_1$. Then we have $p \mid v_2$ and $\mathrm{N}\mathfrak{p} \mid \mathrm{N}\gamma_1$. However, v_2 is composed entirely of primes dividing $b - d$, while $\mathrm{N}\gamma_1$ divides $b^2 + d^2$. We conclude that p divides both b and d. But then p can not divide γ_1 by construction. We can prove in a similar way that $(d, \gamma_1) = (1)$, thus proving the claim.

From equation (5.28) we acquire the validity of

$$\begin{aligned}
\left(\frac{v_1}{z_1}\right) = \left(\frac{v_1}{\gamma_1}\right)_{2,\mathbb{Q}(i)} &= \mu(b, c, d, t) \left(\frac{v_2}{\gamma_1}\right)_{2,\mathbb{Q}(i)} \left(\frac{v}{\gamma_1}\right)_{2,\mathbb{Q}(i)} \\
&= \mu(b, c, d, t) \left(\frac{v_2}{\gamma_1}\right)_{2,\mathbb{Q}(i)} \left(\frac{a + ci}{\gamma_1}\right)_{2,\mathbb{Q}(i)} \left(\frac{-d(1+i)}{\gamma_1}\right)_{2,\mathbb{Q}(i)} \\
&= \mu(b, c, d, t) \left(\frac{v_2}{\gamma_1}\right)_{2,\mathbb{Q}(i)} \left(\frac{a + ci}{\gamma_1}\right)_{2,\mathbb{Q}(i)},
\end{aligned}$$

where we use the identity

$$v = ab - ad + bc + cd \equiv -ad(1 + i) + cd(1 - i) = -d(1 + i)(a + ci) \bmod \gamma_1.$$

We conclude that

$$\left(\frac{b^2 + d^2}{v_1}\right) \left(\frac{a + ci}{e'}\right)_{2,\mathbb{Q}(i)} = \mu(\rho, \rho', b, c, d, t) \left(\frac{z_2}{v_1}\right) \left(\frac{v_2}{\gamma_1}\right)_{2,\mathbb{Q}(i)} \left(\frac{a + ci}{\gamma_2}\right)_{2,\mathbb{Q}(i)} \mathbb{1}_{\gcd(a+ci,\gamma_1)=(1)}. \tag{5.29}$$

Put

$$g(w, \rho) := \mu(\rho, \rho', b, c, d, t) \left(\frac{tv_2}{a^2 + b^2 + c^2 + d^2}\right) \left(\frac{z_2}{v_1}\right) \left(\frac{v_2}{\gamma_1}\right)_{2,\mathbb{Q}(i)} \left(\frac{a + ci}{\gamma_2}\right)_{2,\mathbb{Q}(i)} \mathbb{1}_{\gcd(a+ci,\gamma_1)=\gcd(w,\sigma\tau(w))=(1)}.$$

After combining equations (5.27) and (5.29), we get

$$\begin{aligned}
\left(\frac{g}{w}\right)_{4,M} \left(\frac{2h}{g}\right) &= g(w, \rho) \left(\frac{2c^2 + (b - d)^2}{v_1}\right) \\
&= \mu(\rho, \rho', b, c, d, t) g(w, \rho) \left(\frac{v_1}{2c^2 + (b - d)^2}\right).
\end{aligned}$$

With this formula we have finally rewritten our symbol in a satisfactory manner; we now return to estimating the sum $A(X, \mathfrak{d}, u_i, \rho)$. We recall the factorization $v = v_1 v_2 t$, where v_1 is an odd, positive integer satisfying $\gcd(v_1, b-d) = 1$, v_2 is an odd integer consisting only of primes dividing $b - d$ and t is positive and only divisible by powers of 2. We further recall that ρ' is the congruence class of v_1 modulo 8.

Let 2^α be the closest integer power of 2 to $X^{\frac{1}{100}}$. We fix b, c, d such that $b-d$ has 2-adic valuation at most $\frac{\alpha}{2}$. If a modulo 2^α is given, we claim that v_{odd} is determined modulo 8, where v_{odd} is the odd part of

$$v = a(b-d) + c(b+d), \tag{5.30}$$

with the exception of $\ll X^{\frac{1}{200}}$ congruence classes ρ'' for a modulo 2^α. Note that, for fixed b, c and d, ρ'' determines v modulo 2^α. If $\alpha \geq 3$, v modulo 2^α determines v_{odd} modulo 8 unless v is divisible by $2^{\alpha-3}$. There are only 8 congruence classes modulo 2^α divisible by $2^{\alpha-3}$. Now take such a congruence class, say ρ'''. But there are $\ll X^{\frac{1}{200}}$ congruence classes ρ'' modulo 2^α with

$$\rho''(b-d) + c(b+d) \equiv \rho''' \bmod 2^\alpha$$

by our assumption that the 2-adic valuation of $b-d$ is at most $\frac{\alpha}{2}$, and our claim follows.

Similarly, we know the value of t with the exception of $\ll X^{\frac{1}{200}}$ congruence classes for a modulo 2^α. We remove all such congruence classes from the sum, which gives an error of size at most $X^{\frac{199}{200}}$. From now on we assume that a does not lie in such a congruence class. For the remaining congruence classes modulo 2^α, we observe that ρ' is determined by v_{odd} modulo 8 together with b, c and d. Hence both ρ' and t are determined by a modulo 2^α.

We would also like to treat v_2 as fixed, and we use a similar technique to achieve this. Once more we fix b, c and d. We assume that

$$\gcd(b-d, bc+cd) \leq \exp\left((\log X)^{0.25}\right).$$

We can uniquely factor a positive integer n as $x_1 x_2$, where $\gcd(x_1, x_2) = 1$, $x_1 > 0$ is squarefree and $x_2 > 0$ is squarefull. We call x_1 the squarefree part, and x_2 the squarefull part. We further assume that the squarefull part of $b-d$ is of size at most $\exp\left((\log X)^{0.25}\right)$. We now factor

$$\gcd(b-d, bc+cd) = \prod_{i=1}^{k} p_i^{f_i}.$$

Define $f_i'(p_i)$ to be the smallest integer such that

$$p_i^{f_i'(p_i)} \geq \exp\left(2\,(\log X)^{0.25}\right)$$

and define

$$G := \prod_{i=1}^{k} p_i^{f_i'(p_i)}.$$

Clearly, we have that $\gcd(b-d, bc+cd)$ divides G, since the squarefull part of $b-d$ is of size at most $\exp\left((\log X)^{0.25}\right)$. If a modulo G is given, we claim that v_2 is determined modulo G with the exception of at most

$$\ll \log X \min_{1 \leq i \leq k} \frac{G}{p_i^{f_i'(p_i)}}$$

congruence classes ρ'' for a modulo G. Take a prime divisor p_i of $b-d$. If p_i does not divide $bc+cd$, then clearly

$$p_i \nmid a(b-d) + bc + cd,$$

so we have found the p_i valuation of $a(b-d)+bc+cd$. Now suppose that p_i also divides $bc+cd$. Then we know the p_i valuation unless

$$a(b-d) + bc + cd \equiv 0 \bmod p_i^{f_i'(p_i)}.$$

However, we know that the p_i valuation of $b-d$ is at most $f_i'(p_i)/2$. Hence there are at most $p_i^{f_i'(p_i)/2}$ congruence classes for a modulo $p_i^{f_i'(p_i)/2}$ for which

$$a(b-d) + bc + cd \equiv 0 \bmod p_i^{f_i'(p_i)},$$

and we call such a congruence class forbidden. We let G_i be the set of forbidden congruence classes modulo $p_i^{f_i(p_i)}$. Now we discard all congruence classes ρ'' modulo G for which there exists a prime p_i dividing $\gcd(b-d, bc+cd)$ such that the reduction of ρ'' modulo $p_i^{f_i(p_i)}$ lies in G_i. This proves the claim.

Set

$$m := \operatorname{lcm}\left(G, z_2, \mathrm{N}\gamma_2, 2^\alpha, 2^{10}\right). \tag{5.31}$$

Then

$$\left(\frac{tv_2}{a^2+b^2+c^2+d^2}\right)\left(\frac{z_2}{v_1}\right)\left(\frac{v_2}{\gamma_1}\right)_{2,\mathbb{Q}(i)}\left(\frac{a+ci}{\gamma_2}\right)_{2,\mathbb{Q}(i)}$$

depends only on a modulo m, b, c and d. If we write $\beta := b\zeta_8 + c\zeta_8^2 + d\zeta_8^3$, we have the following estimate

$$A(X, \mathfrak{d}, u_i, \rho) \ll \sum_{\beta} \sum_{f \in \mathbb{Z}/m\mathbb{Z}} \left| \sum_{\substack{a \in \mathbb{Z} \\ a \text{ sat. } (*)}} \left(\frac{v_1}{2c^2 + (b-d)^2}\right) \mathbb{1}_{\gcd(a+ci,\gamma_1) = \gcd(a+\beta, \sigma\tau(a+\beta)) = (1)} \right|,$$

where $(*)$ are the simultaneous conditions

$$a + \beta \in u_i\mathcal{D}(X), \quad a + \beta \equiv 0 \bmod \mathfrak{d}, \quad a + \beta \equiv \rho \bmod 2^{10}, \quad a \equiv f \bmod m.$$

Recall that the condition $a + \beta \in u_i\mathcal{D}(X)$ implies $a, b, c, d \ll X^{\frac{1}{4}}$, see Lemma 5.2.3. We will only consider β satisfying the following five properties

- $z_2, \mathrm{N}\gamma_2 \leq X^{\frac{1}{200}}$;
- $\gcd(b-d, bc+cd) \leq \exp\left((\log X)^{0.25}\right)$;
- the 2-adic valuation of $b-d$ is at most $\frac{\alpha}{2}$;
- the squarefull part of $b-d$ is of size at most $\exp\left((\log X)^{0.25}\right)$;
- the odd, squarefree part of $2c^2+(b-d)^2$ is at least $X^{\frac{99}{200}}$.

We claim that there are at most

$$\ll \frac{X^{\frac{3}{4}}}{\exp\left((\log X)^{0.2}\right)}$$

elements β that do not satisfy all five conditions. To do so, we shall bound the number of β that fail a given bullet point in the above list. For the third and fourth bullet point this is easily verified. For the fifth bullet point, we use that $2c^2+(b-d)^2$ represents a given integer at most $\ll_\epsilon X^{\frac{1}{4}+\epsilon}$ times, and this reduces the problem to an easy counting problem. A similar argument disposes with the first bullet point. Finally, for the second bullet point, we count the number of β such that

$$\gcd(b-d, b+d) > \exp\left(\frac{1}{2}(\log X)^{0.25}\right) \text{ or } \gcd(b-d, c) > \exp\left(\frac{1}{2}(\log X)^{0.25}\right).$$

For those β, we bound the inner sum trivially by $\ll X^{\frac{1}{4}}/m$ inducing an error of size

$$\ll \frac{X}{\exp\left((\log X)^{0.2}\right)}.$$

For the remaining β, we have $G \ll_\epsilon X^\epsilon$ and hence $m \ll_\epsilon X^{\frac{1}{50}+\epsilon}$ by the first bullet point and the definition of m, see equation (5.31). Note that

$$\mathbb{1}_{\gcd(a+\beta, \sigma\tau(a+\beta))=(1)} = \mathbb{1}_{\gcd(a+\beta, \sigma\tau(\beta)-\beta)=(1)}.$$

We use the Möbius function to detect the coprimality conditions, which yields the following upper bound

$$A(X, \mathfrak{d}, u_i, \rho) \ll \sum_{\beta} \sum_{f \in \mathbb{Z}/m\mathbb{Z}} \sum_{\mathfrak{d}_1 | \gamma_1} \sum_{\mathfrak{d}_2 | \sigma\tau(\beta)-\beta} \left| \sum_{\substack{a \in \mathbb{Z} \\ a \text{ sat. } (**)}} \left(\frac{v_1}{2c^2+(b-d)^2}\right) \right|,$$

where $(**)$ are the simultaneous conditions

$$a+\beta \in u_i\mathcal{D}(X), \quad a+\beta \equiv 0 \bmod \mathfrak{d}, \quad a+\beta \equiv \rho \bmod 2^{10}, \quad a \equiv f \bmod m$$
$$a+ci \equiv 0 \bmod \mathfrak{d}_1, \quad a+\beta \equiv 0 \bmod \mathfrak{d}_2.$$

Define m' to be the smallest positive integer that is divisible by $\mathrm{lcm}(\mathfrak{d}, \mathfrak{d}_1, \mathfrak{d}_2)$. Put

$$M := \mathrm{lcm}\,(m, m')\,.$$

The congruence conditions for a in $(**)$ are equivalent to at most one congruence condition modulo M. We assume that it is equivalent to exactly one congruence condition modulo M, say F, otherwise the inner sum is empty. Then we have

$$A(X, \mathfrak{d}, u_i, \rho) \ll \sum_{\beta} \sum_{f \in \mathbb{Z}/m\mathbb{Z}} \sum_{\mathfrak{d}_1 | \gamma_1} \sum_{\mathfrak{d}_2 | \sigma\tau(\beta) - \beta} \left| \sum_{\substack{a \in \mathbb{Z} \\ a+\beta \in u_i \mathcal{D}(X) \\ a \equiv F \bmod M}} \left(\frac{v_1}{2c^2 + (b-d)^2} \right) \right|. \tag{5.32}$$

We assume that $M \leq X^{\frac{1}{8}}$, since otherwise the trivial bound suffices. Furthermore, for fixed β, the condition $a + \beta \in u_i\mathcal{D}(X)$ means that a runs over $\ll 1$ intervals with endpoints depending on β and u_i. Since $a \ll X^{\frac{1}{4}}$, we know that each interval has length $\ll X^{\frac{1}{4}}$. We have the factorization

$$2c^2 + (b-d)^2 = q_1 q_2,$$

where q_1 is the odd, squarefree part. We know that $q_2 \ll X^{\frac{1}{200}}$, and we split the sum over a in congruence classes modulo q_2. For fixed b, c and d, the condition $a \equiv F \bmod M$ implies that v_1 is a linear function of a with linear term not divisible by q_1 by our assumptions $q_1 \geq X^{\frac{99}{200}}$ and $M \leq X^{\frac{1}{8}}$. Hence we may employ Burgess' bound [8] to equation (5.32) with $r = 2$ and $q = q_1 \ll X^{\frac{1}{2}}$ to prove

$$A(X, \mathfrak{d}, u_i, \rho) \ll_\epsilon X^{\frac{31}{32} + \frac{1}{50} + \frac{1}{200} + \epsilon} + X^{\frac{199}{200}} + X^{\frac{15}{16}} + \frac{X}{\exp\left((\log X)^{0.2}\right)},$$

where the second term accounts for the discarded congruence classes for a, the third term accounts for those M with $M > X^{\frac{1}{8}}$ and the fourth term accounts for the discarded β. This establishes the following proposition.

Proposition 5.5.1. *We have for all ideals $\mathfrak{d}$ of $\mathbb{Z}[\zeta_8]$*

$$A(X, \mathfrak{d}) \ll \frac{X}{\exp\left((\log X)^{0.2}\right)}.$$

5.6 Sums of type II

During the proof of Lemma 5.4.2 we defined $[w]_1$ and $[w]_2$. We have the useful decomposition

$$[w] = [w]_1 [w]_2.$$

In this section we need to carefully study the multiplicative properties of $[w]$, and we do so by studying the multiplicative properties of $[w]_1$ and $[w]_2$. These properties will then be used to prove cancellation in sums of type II. We start by studying $[w]_1$; our treatment is almost identical to [43]. If w is an odd element of $\mathbb{Z}[\zeta_8]$, we have

$$[w]_1 = \left(\frac{\left(\frac{1}{2} - \frac{1}{2\sqrt{2}}\right)\sigma(w)\sigma\tau(w)}{w}\right)_{4,M} = \left(\frac{\left(2 - \sqrt{2}\right)\sigma(w)\sigma\tau(w)}{w}\right)_{4,M}.$$

Define

$$\gamma_1(w, z) := \left(\frac{\sigma(z)}{w}\right)_{2,M}. \tag{5.33}$$

For the remainder of this section, we use the convention that $\delta(w, z)$ is a function depending only on the congruence classes of w and z modulo 2^{10}; at each occurence $\delta(w, z)$ may be a different function.

Lemma 5.6.1. *We have for all odd $w, z \in \mathbb{Z}[\zeta_8]$*

$$[wz]_1 = \delta(w, z)[w]_1[z]_1\gamma_1(w, z)\mathbb{1}_{\gcd(w,\sigma\tau(z))=(1)}.$$

Proof. By definition of $[w]_1$ we have

$$\begin{aligned}[wz]_1 &= \left(\frac{\left(2 - \sqrt{2}\right)\sigma(wz)\sigma\tau(wz)}{wz}\right)_{4,M} \\ &= [w]_1[z]_1\left(\frac{\sigma(z)}{w}\right)_{4,M}\left(\frac{\sigma\tau(z)}{w}\right)_{4,M}\left(\frac{\sigma(w)}{z}\right)_{4,M}\left(\frac{\sigma\tau(w)}{z}\right)_{4,M}.\end{aligned}$$

Since σ fixes i and therefore any quartic residue symbol, Theorem 5.2.2 yields

$$\begin{aligned}\left(\frac{\sigma(z)}{w}\right)_{4,M}\left(\frac{\sigma(w)}{z}\right)_{4,M} &= \delta(w, z)\left(\frac{\sigma(z)}{w}\right)_{4,M}\left(\frac{z}{\sigma(w)}\right)_{4,M} \\ &= \delta(w, z)\left(\frac{\sigma(z)}{w}\right)_{4,M}\sigma\left(\left(\frac{\sigma(z)}{w}\right)_{4,M}\right) \\ &= \delta(w, z)\left(\frac{\sigma(z)}{w}\right)_{2,M}.\end{aligned}$$

If we do the same computation for $\sigma\tau$, we obtain

$$\left(\frac{\sigma\tau(z)}{w}\right)_{4,M}\left(\frac{\sigma\tau(w)}{z}\right)_{4,M} = \delta(w, z)\mathbb{1}_{\gcd(w,\sigma\tau(z))=(1)},$$

since $\sigma\tau$ does not fix i. This proves the lemma. $\square$

In the next lemma we collect the most important properties of $\gamma_1(w, z)$.

Lemma 5.6.2. *Let $w, z \in \mathbb{Z}[\zeta_8]$ be odd and define $\gamma_1(w, z)$ as in equation (5.33).*

(i) $\gamma_1(w,z)$ *is essentially symmetric*

$$\gamma_1(w,z) = \delta(w,z)\gamma_1(z,w).$$

(ii) $\gamma_1(w,z)$ *is multiplicative in both arguments*

$$\gamma_1(w,z_1z_2) = \gamma_1(w,z_1)\gamma_1(w,z_2), \quad \gamma_1(w_1w_2,z) = \gamma_1(w_1,z)\gamma_1(w_2,z).$$

Proof. This is straightforward. □

With this lemma we have completed our study of $[w]_1$ and $\gamma_1(w,z)$. We will now focus on $[w]_2$. Recall that

$$[w]_2 = \left(\frac{2h}{g}\right) = \delta(w)\left(\frac{v}{u}\right).$$

The second representation of $[w]_2$ is very convenient, since it allows us to use earlier work of Milovic [58]. Define

$$\gamma_2(w,z) := \left(\frac{\sigma(wz)\sigma\tau(wz)}{w\tau(w)}\right)_{2,K}, \tag{5.34}$$

where $K := \mathbb{Q}(\sqrt{2})$.

Lemma 5.6.3. *The following formula is valid for all odd* $w,z \in \mathbb{Z}[\zeta_8]$

$$[wz]_2 = \delta(w,z)[w]_2[z]_2\gamma_2(w,z).$$

Proof. Milovic [58, p. 1009] defines the following symbol

$$[u+v\sqrt{2}]_3 := \left(\frac{v}{u}\right).$$

Then it is easily seen that $[w]_2 = \delta(w)[w\tau(w)]_3$ and that $w\tau(w)$ is totally positive. Now apply Proposition 8 of Milovic [58]. □

To further our study of $\gamma_2(w,z)$, it will be convenient to define a second function $\mathrm{m}(w)$ by the following formula

$$\mathrm{m}(w) := \gamma_2(w,1) = \left(\frac{\sigma(w)\sigma\tau(w)}{w\tau(w)}\right)_{2,K}.$$

It turns out that $\gamma_2(w,z)$ is neither symmetric nor multiplicative. Instead, it is symmetric and multiplicative twisted by the factor m.

Lemma 5.6.4. *Let* $w,z \in \mathbb{Z}[\zeta_8]$ *be odd and define* $\gamma_2(w,z)$ *as in equation (5.34).*

(i) $\gamma_2(w,z)$ *is twisted symmetric*

$$\gamma_2(w,z)\gamma_2(z,w) = \mathrm{m}(wz).$$

(ii) $\gamma_2(w, z)$ *is twisted multiplicative in* z

$$\gamma_2(w, z_1 z_2) = \mathrm{m}(w)\gamma_2(w, z_1)\gamma_2(w, z_2).$$

Proof. Left to the reader. $\square$

With this out of the way we are ready to tackle the sums of type II. Let $\{\alpha_w\}$ and $\{\beta_z\}$ be sequences of complex numbers of absolute value at most 1 and let ρ and μ be invertible congruence classes modulo 2^{10}. We define

$$B_1(M, N, \rho, \mu) := \sum_{\substack{w \in \mathcal{D}(M) \\ w \equiv \rho \bmod 2^{10}}} \sum_{\substack{z \in \mathcal{D}(N) \\ z \equiv \mu \bmod 2^{10}}} \alpha_w \beta_z [wz],$$

where we suppress the dependence on $\{\alpha_w\}$ and $\{\beta_z\}$. Then we have the following proposition.

Proposition 5.6.5. *There is an absolute constant* $\theta_3 > 0$ *such that for all sequences of complex numbers* $\{\alpha_w\}$ *and* $\{\beta_z\}$ *of absolute value at most* 1*, all invertible congruence classes* ρ *and* μ *modulo* 2^{10}

$$B_1(M, N, \rho, \mu) \ll \left(M^{-\frac{1}{24}} + N^{-\frac{1}{24}}\right) MN(\log MN)^{\theta_3}.$$

Proof. We start by expanding $[wz]$ using Lemma 5.6.1 and Lemma 5.6.3. We may absorb $[w]_1$, $[w]_2$, $[z]_1$ and $[z]_2$ in the coefficients α_w and β_z. Then it suffices to prove for all sequences of complex numbers $\{\alpha_w\}$ and $\{\beta_z\}$ of absolute value at most 1 and all invertible congruence classes ρ and μ modulo 2^{10} the following estimate

$$\begin{aligned} B_2(M, N, \rho, \mu) &:= \sum_{\substack{w \in \mathcal{D}(M) \\ w \equiv \rho \bmod 2^{10}}} \sum_{\substack{z \in \mathcal{D}(N) \\ z \equiv \mu \bmod 2^{10}}} \alpha_w \beta_z \gamma_1(w, z)\gamma_2(w, z)\mathbb{1}_{\gcd(w, \sigma\tau(z)) = (1)} \\ &\ll \left(M^{-\frac{1}{24}} + N^{-\frac{1}{24}}\right) MN(\log MN)^{\theta_3}. \end{aligned}$$

Define

$$\gamma_3(w, z) := \left(\frac{\sigma(z)\sigma\tau(z)}{w\tau(w)}\right)_{2,K},$$

so that we have the factorization $\gamma_2(w, z) = \mathrm{m}(w)\gamma_3(w, z)$. Absorbing $\mathrm{m}(w)$ in α_w and using the identity

$$\gamma_3(w, z)\mathbb{1}_{\gcd(w, \sigma\tau(z)) = (1)} = \gamma_3(w, z),$$

we see that it is enough to establish

$$\begin{aligned} B_3(M, N, \rho, \mu) &:= \sum_{\substack{w \in \mathcal{D}(M) \\ w \equiv \rho \bmod 2^{10}}} \sum_{\substack{z \in \mathcal{D}(N) \\ z \equiv \mu \bmod 2^{10}}} \alpha_w \beta_z \gamma_1(w, z)\gamma_3(w, z) \\ &\ll \left(M^{-\frac{1}{24}} + N^{-\frac{1}{24}}\right) MN(\log MN)^{\theta_3}. \end{aligned}$$

Theorem 5.2.1 shows that $\gamma_3(w,z)$ is also essentially symmetric, i.e.

$$\gamma_3(w,z) = \delta(w,z)\gamma_3(z,w).$$

Due to the symmetry of $\gamma_1(w,z)$, see Lemma 5.6.2(i), and the symmetry of $\gamma_3(w,z)$, we may further reduce to the case $N \geq M$. We take $k := 12$ and apply Hölder's inequality with $1 = \frac{k-1}{k} + \frac{1}{k}$ to the w variable to obtain

$$|B_3(M,N,\rho,\mu)|^k \leq \left(\sum_{\substack{w \in \mathcal{D}(M) \\ w \equiv \rho \bmod 2^{10}}} |\alpha_w|^{\frac{k}{k-1}} \right)^{k-1} \sum_{\substack{w \in \mathcal{D}(M) \\ w \equiv \rho \bmod 2^{10}}} \left| \sum_{\substack{z \in \mathcal{D}(N) \\ z \equiv \mu \bmod 2^{10}}} \beta_z \gamma_1(w,z)\gamma_3(w,z) \right|^k .$$

The first factor is trivially bounded by $\ll M^{k-1}$ with absolute implied constant. Lemma 5.6.2(ii) implies that $\gamma_1(w,z)$ is multiplicative in z and Lemma 5.6.4(ii) implies that $\gamma_3(w,z)$ is multiplicative in z. Hence $\gamma_1(w,z)\gamma_3(w,z)$ is multiplicative in z. We conclude that

$$|B_3(M,N,\rho,\mu)|^k \ll M^{k-1} \sum_{\substack{w \in \mathcal{D}(M) \\ w \equiv \rho \bmod 2^{10}}} \epsilon(w) \sum_z \beta'_z \gamma_1(w,z)\gamma_3(w,z), \tag{5.35}$$

where

$$\epsilon(w) := \left(\frac{\left| \sum_{\substack{z \in \mathcal{D}(N) \\ z \equiv \mu \bmod 2^{10}}} \beta_z \gamma_1(w,z)\gamma_3(w,z) \right|}{\sum_{\substack{z \in \mathcal{D}(N) \\ z \equiv \mu \bmod 2^{10}}} \beta_z \gamma_1(w,z)\gamma_3(w,z)} \right)^k$$

and

$$\beta'_z := \sum_{\substack{z = z_1 \cdot \ldots \cdot z_k \\ z_1, \ldots, z_k \in \mathcal{D}(N) \\ z_1 \equiv \ldots \equiv z_k \equiv \mu \bmod 2^{10}}} \beta_{z_1} \cdot \ldots \cdot \beta_{z_k}.$$

We will now study the summation condition for z in the inner sum of equation (5.35) more carefully. By construction, $\mathcal{D}(N)$ contains exactly eight generators of any principal ideal. Furthermore, there are $\ll N^k$ values of z for which $\beta'_z \neq 0$. Hence we obtain the bound

$$\sum_z (\beta'_z)^2 \ll (\log N)^{\theta_3} N^k$$

for some absolute constant θ_3, since k is fixed. An application of the Cauchy-Schwarz inequality over the z variable yields

$$\left(\sum_{\substack{w\in\mathcal{D}(M)\\ w\equiv\rho \bmod 2^{10}}}\epsilon(w)\sum_{z}\beta'_z\gamma_1(w,z)\gamma_3(w,z)\right)^2 = \left(\sum_{z}\beta'_z\sum_{\substack{w\in\mathcal{D}(M)\\ w\equiv\rho \bmod 2^{10}}}\epsilon(w)\gamma_1(w,z)\gamma_3(w,z)\right)^2$$
$$\ll (\log N)^{\theta_3}N^k \sum_{\substack{w_1\in\mathcal{D}(M)\\ w_1\equiv\rho \bmod 2^{10}}}\ \sum_{\substack{w_2\in\mathcal{D}(M)\\ w_2\equiv\rho \bmod 2^{10}}}\epsilon(w_1)\overline{\epsilon(w_2)}\sum_{z}\gamma_1(w_1w_2,z)\gamma_3(w_1w_2,z), \quad (5.36)$$

because $\gamma_1(w,z)$ and $\gamma_3(w,z)$ are multiplicative in w. Conveniently, inequality (5.36) remains valid if we extend the sum over z to a larger domain. Let $z_1,\dots,z_k \in \mathcal{D}(N)$ and write

$$z_i = \sum_{j=1}^{4} a_{ij}\zeta_8^j.$$

Then we have $|a_{ij}| \ll N^{\frac{1}{4}}$. Now define

$$\mathcal{B}(C) := \left\{\sum_{j=1}^{4} a_j\zeta_8^j : a_j \in \mathbb{Z}, |a_j| \le CN^{\frac{k}{4}}\right\}.$$

Then, if C is sufficiently large, $\beta'_z \neq 0$ implies $z \in \mathcal{B}(C)$. For this choice of C, we extend the range of summation over z in equation (5.36) to the set $\mathcal{B}(C)$. We split the sum over z in congruence classes ζ modulo $\mathrm{N}(w_1w_2)$; we claim that for all odd w

$$\sum_{\zeta \bmod \mathrm{N}(w)} \gamma_1(w,\zeta)\gamma_3(w,\zeta) = 0$$

provided that $\mathrm{N}(w)$ is not squarefull. Substituting the definition of $\gamma_1(w,\zeta)$ and $\gamma_3(w,\zeta)$ gives

$$f(w) := \sum_{\zeta \bmod \mathrm{N}(w)} \gamma_1(w,\zeta)\gamma_3(w,\zeta) = \sum_{\zeta \bmod \mathrm{N}(w)} \left(\frac{\sigma(\zeta)\sigma\tau(\zeta)}{w\tau(w)}\right)_{2,K}\left(\frac{\sigma(\zeta)}{w}\right)_{2,M}.$$

Then a calculation shows that for all odd w and w' satisfying $(\mathrm{N}(w),\mathrm{N}(w')) = 1$

$$f(ww') = f(w)f(w').$$

Hence, to establish the claim, it is enough to prove that $f(w) = 0$ if w is an odd prime of degree 1. To do so, we start with the identity

$$\left(\frac{\sigma(\zeta)\sigma\tau(\zeta)}{w\tau(w)}\right)_{2,K} = \left(\frac{\sigma(\zeta)\sigma\tau(\zeta)}{w}\right)_{2,M}.$$

Here we rely in an essential way that w is an odd prime of degree 1, so we have an isomorphism of finite fields $O_M/w \cong O_K/w\tau(w)$. We use this to give a simple expression for $f(w)$

$$f(w) = \sum_{\zeta \bmod \mathrm{N}(w)} \left(\frac{\sigma\tau(\zeta)}{w}\right)_{2,M}\mathbb{1}_{(\sigma(\zeta),w)=(1)},$$

which apart from a non-zero factor is

$$\sum_{\zeta \bmod \sigma(w)\sigma\tau(w)} \left(\frac{\sigma\tau(\zeta)}{w}\right)_{2,M} \mathbb{1}_{(\sigma(\zeta),w)=(1)} = \sum_{\zeta \bmod \sigma\tau(w)} \left(\frac{\sigma\tau(\zeta)}{w}\right)_{2,M} \sum_{\zeta \bmod \sigma(w)} \mathbb{1}_{(\sigma(\zeta),w)=(1)} = 0.$$

Note that $\sigma(w)$ and $\sigma\tau(w)$ are coprime, so that we are allowed to expand the sum over $\sigma(w)\sigma\tau(w)$ as the product of the two sums over $\sigma(w)$ and $\sigma\tau(w)$. With the claim established, we can give an upper bound for the sum over $z \in \mathcal{B}(C)$

$$\sum_{z\in\mathcal{B}(C)} \gamma_1(w_1w_2, z)\gamma_3(w_1w_2, z) \ll \begin{cases} N^k & \text{if } \mathrm{N}(w_1w_2) \text{ is squarefull} \\ \sum_{i=1}^{4} M^{2i}N^{k\left(1-\frac{i}{4}\right)} & \text{otherwise,} \end{cases}$$

where the second bound uses the claim and $\mathrm{N}(w_1w_2) \leq M^2$. Because of our choice of k and $N \geq M$, we can simplify the second bound to $M^2N^{\frac{3}{4}k}$. Equation (5.35), equation (5.36) and the above bound acquire the validity of

$$\begin{aligned} |B_3(M, N, \rho, \mu)|^{2k} &\ll (\log N)^{\theta_3} M^{2k-2}N^k \left(M \cdot N^k + M^2 \cdot M^2N^{\frac{3}{4}k}\right) \\ &\ll (\log N)^{\theta_3} \left(M^{2k-1} \cdot N^k + M^{2k+2} \cdot N^{\frac{7}{4}k}\right). \end{aligned}$$

Since the first term above dominates the second term due to our choice of k and $N \geq M$, the proof of the proposition is complete. □

Having dealt with sums of type II for the symbol $[wz]$, we now turn to sums of type II with $a_{\mathfrak{mn}}$. For sequences of complex numbers $\{\alpha_{\mathfrak{m}}\}$ and $\{\beta_{\mathfrak{n}}\}$ of absolute value at most 1 we defined in Section 5.3 the following sum

$$B(M, N) = \sum_{\mathrm{N}\mathfrak{m}\leq M} \sum_{\mathrm{N}\mathfrak{n}\leq N} \alpha_{\mathfrak{m}}\beta_{\mathfrak{n}}a_{\mathfrak{mn}}.$$

Proposition 5.6.6. *There is an absolute constant $\theta_3 > 0$ such that for all sequences of complex numbers $\{\alpha_{\mathfrak{m}}\}$ and $\{\beta_{\mathfrak{n}}\}$ of absolute value at most* 1

$$B(M, N) \ll \left(M^{-\frac{1}{24}} + N^{-\frac{1}{24}}\right) MN(\log MN)^{\theta_3}.$$

Proof. By picking generators for $\mathfrak{m}$ and $\mathfrak{n}$ we obtain the following identity

$$B(M, N) = \sum_{\mathrm{N}\mathfrak{m}\leq M} \sum_{\mathrm{N}\mathfrak{n}\leq N} \alpha_{\mathfrak{m}}\beta_{\mathfrak{n}}a_{\mathfrak{mn}} = \frac{1}{64} \sum_{w\in\mathcal{D}(M)} \sum_{z\in\mathcal{D}(N)} \alpha_w\beta_z a_{(wz)}.$$

We split the sum $B(M, N)$ in congruence classes modulo 2^{10}. We need only consider invertible congruence classes, since otherwise $a_{wz} = 0$ by definition. Furthermore, condition (5.7) depends only on g modulo 4, which is in turn determined by w modulo 4.

Therefore, it suffices to bound the following sum

$$\sum_{\substack{w\in\mathcal{D}(M)\\ w\equiv\rho \bmod 2^{10}}}\ \sum_{\substack{z\in\mathcal{D}(N)\\ z\equiv\mu \bmod 2^{10}}} \alpha_w\beta_z\left([wz]+[\epsilon wz]+[\epsilon^2 wz]+[\epsilon^3 wz]\right),$$

where ρ and μ are invertible congruence classes modulo 2^{10} such that $g \equiv 1 \bmod 4$. From Lemma 5.6.1 and Lemma 5.6.3 we deduce that

$$[\epsilon wz] = \delta(w,z)[\epsilon][wz].$$

Now apply Proposition 5.6.5. □

Chapter 6

Joint distribution of spins

Joint work with Djordjo Milovic

Abstract

We answer a question of Iwaniec, Friedlander, Mazur and Rubin [24] on the joint distribution of spin symbols. As an application we give a negative answer to a conjecture of Cohn and Lagarias on the existence of governing fields for the 16-rank of class groups under the assumption of a short character sum conjecture.

6.1 Introduction

One of the most fundamental and most prevalent objects in number theory are extensions of number fields; they arise naturally as fields of definitions of solutions to polynomial equations. Many interesting phenomena are encoded in the splitting of prime ideals in extensions. For instance, if p and q are distinct prime numbers congruent to 1 modulo 4, the statement that p splits in $\mathbb{Q}(\sqrt{q})/\mathbb{Q}$ if and only if q splits in $\mathbb{Q}(\sqrt{p})/\mathbb{Q}$ is nothing other than the law of quadratic reciprocity, a common ancestor to much of modern number theory.

Let K be a number field, $\mathfrak{p}$ a prime ideal in its ring of integers $\mathcal{O}_K$, and α an element of the algebraic closure $\overline{K}$. Suppose we were to ask, as we vary $\mathfrak{p}$, how often $\mathfrak{p}$ splits completely in the extension $K(\alpha)/K$. If α is fixed as $\mathfrak{p}$ varies over all prime ideals in $\mathcal{O}_K$, a satisfactory answer is provided by the Chebotarev Density Theorem, which is grounded in the theory of L-functions and their zero-free regions. The Chebotarev Density Theorem, however, often cannot provide an answer if α varies along with $\mathfrak{p}$ in some prescribed manner. The purpose of this chapter is to fill this gap for quadratic extensions in a natural setting that arises in many applications. This setting, which we now describe, is inspired by the work of Friedlander, Iwaniec, Mazur, and Rubin [24] and is amenable to sieve theory involving sums of type I and type II, as opposed to the theory of L-functions.

Let $K/\mathbb{Q}$ be a Galois extension of degree n. Unlike in [24], we do *not* impose the very restrictive condition that $\mathrm{Gal}(K/\mathbb{Q})$ is cyclic. For the moment, let us restrict to the setting where K is totally real and where every totally positive unit in $\mathcal{O}_K$ is a square, as in [24]. To each non-trivial automorphism $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ and each odd principal prime ideal $\mathfrak{p} \subset \mathcal{O}_K$, we attach the quantity $\mathrm{spin}(\sigma, \mathfrak{p}) \in \{-1, 0, 1\}$, defined as

$$\mathrm{spin}(\sigma, \mathfrak{p}) = \left(\frac{\pi}{\sigma(\pi)}\right)_{K,2}, \tag{6.1}$$

where π is any totally positive generator of $\mathfrak{p}$ and $\left(\frac{\cdot}{\cdot}\right)_{K,2}$ denotes the quadratic residue symbol in K. If we let $\alpha^2 = \sigma^{-1}(\pi)$, then $\mathrm{spin}(\sigma, \mathfrak{p})$ governs the splitting of $\mathfrak{p}$ in $K(\alpha)$, i.e., $\mathrm{spin}(\sigma, \mathfrak{p}) = 1$ (resp., -1, 0) if $\mathfrak{p}$ is split (resp., inert, ramified) in $K(\alpha)/K$. In [24], under the assumptions that σ generates $\mathrm{Gal}(K/\mathbb{Q})$, that $n \geq 3$, and that the technical Conjecture C_n (see Section 6.2.5) holds true, Friedlander et al. prove that the natural density of $\mathfrak{p}$ that are split (resp., inert) in $K(\sqrt{\alpha})/K$ is $\frac{1}{2}$ (resp., $\frac{1}{2}$), just as would be the case were α not to vary with $\mathfrak{p}$.

More generally, suppose S is a subset of $\mathrm{Gal}(K/\mathbb{Q})$ and consider the *joint spin*

$$s_{\mathfrak{p}} = \prod_{\sigma \in S} \mathrm{spin}(\sigma, \mathfrak{p}),$$

defined for principal prime ideals $\mathfrak{p} = \pi\mathcal{O}_K$. If we let $\alpha^2 = \prod_{\sigma \in S} \sigma^{-1}(\pi)$, then $s_{\mathfrak{p}}$ is equal to 1 (resp., -1, 0) if $\mathfrak{p}$ is split (resp., inert, ramified) in $K(\alpha)/K$. If $\sigma^{-1} \in S$ for some $\sigma \in S$, then the factor $\mathrm{spin}(\sigma, \mathfrak{p})\mathrm{spin}(\sigma^{-1}, \mathfrak{p})$ falls under the purview of the usual Chebotarev Density Theorem as suggested in [24, p. 744] and studied precisely by McMeekin [56]. We therefore focus on the case that $\sigma \notin S$ whenever $\sigma^{-1} \in S$ and prove the following equidistribution theorem concerning the joint spin $s_{\mathfrak{p}}$, defined in full generality, also for totally complex fields, in Section 6.2.3.

Theorem 6.1.1. *Let $K/\mathbb{Q}$ be a Galois extension of degree n. If K is totally real, we further assume that every totally positive unit in $\mathcal{O}_K$ is a square. Suppose that S is a non-empty subset of* $\mathrm{Gal}(K/\mathbb{Q})$ *such that $\sigma \in S$ implies $\sigma^{-1} \notin S$. Foe each non-zero ideal $\mathfrak{a}$ in $\mathcal{O}_K$, define $s_{\mathfrak{a}}$ as in* (6.6)*. Assume Conjecture $C_{|S|n}$ holds true with $\delta = \delta(|S|n) > 0$ (see Section 6.2.5). Let $\epsilon > 0$ be a real number. Then for all $X \geq 2$, we have*

$$\sum_{\substack{\mathrm{N}(\mathfrak{p}) \leq X \\ \mathfrak{p} \ \mathit{prime}}} s_{\mathfrak{p}} \ll X^{1 - \frac{\delta}{54|S|^2 n(12n+1)} + \epsilon},$$

where the implied constant depends only on ϵ and K.

It may be possible to weaken our condition on S and instead require only that there exists $\sigma \in S$ with $\sigma^{-1} \notin S$.

The main theorem in [24] is the special case of Theorem 6.1.1 where $\mathrm{Gal}(K/\mathbb{Q}) = \langle\sigma\rangle$, $n \geq 3$, and $S = \{\sigma\}$. After establishing their equidistribution result, Friedlander et al. [24, p. 744] raise the question of the joint distribution of spins, and in particular the case of $\mathrm{spin}(\sigma, \mathfrak{p})$ and $\mathrm{spin}(\sigma^2, \mathfrak{p})$ where again $\mathrm{Gal}(K/\mathbb{Q}) = \langle \sigma \rangle$, but $S = \{\sigma, \sigma^2\}$ and $n \geq 5$. The following corollary of Theorem 6.1.1 applied to the set $S = \{\sigma, \sigma^2\}$ answers their question.

Theorem 6.1.2. *Let $K/\mathbb{Q}$ be a totally real Galois extension of degree n such that every totally positive unit in $\mathcal{O}_K$ is a square. Suppose that $S = \{\sigma_1, \ldots, \sigma_t\}$ is a non-empty subset of $\mathrm{Gal}(K/\mathbb{Q})$ such that $\sigma \in S$ implies $\sigma^{-1} \notin S$. Assume Conjecture C_{tn} holds true (see Section 6.2.5). Let $= (e_1, \ldots, e_t) \in \mathbb{F}_2^t$. Then, as $X \to \infty$, we have*

$$\frac{|\{\mathfrak{p} \text{ principal prime ideal in } \mathcal{O}_K :\ \mathrm{N}(\mathfrak{p}) \leq X,\ \mathrm{spin}(\sigma_i, \mathfrak{p}) = (-1)^{e_i} \text{ for } 1 \leq i \leq t\}|}{|\{\mathfrak{p} \text{ principal prime ideal in } \mathcal{O}_K :\ \mathrm{N}(\mathfrak{p}) \leq X\}|} \sim \frac{1}{2^t}.$$

We expect that Theorem 6.1.1 has several algebraic applications; see for example the original work of Friedlander et al. [24], but also [41], [43], and [58]. Here we give one such application by giving a negative answer to a conjecture of Cohn and Lagarias [11]. Given an integer $k \geq 1$ and a finite abelian group A, we define the 2^k-rank of A as

$$\mathrm{rk}_{2^k} A = \dim_{\mathbb{F}_2} 2^{k-1}A/2^k A.$$

Cohn and Lagarias [11] considered the one-prime-parameter families of quadratic number fields $\{\mathbb{Q}(\sqrt{dp})\}_p$, where d is a fixed integer $\not\equiv 2 \bmod 4$ and p varies over primes such that dp is a fundamental discriminant. Bolstered by ample numerical evidence as well as theoretical examples [11], they conjectured that for every $k \geq 1$ and $d \not\equiv 2 \bmod 4$, there exists a governing field $M_{d,k}$ for the 2^k-rank of the narrow class group $\mathcal{C}\ell(\mathbb{Q}(\sqrt{dp}))$ of $\mathbb{Q}(\sqrt{dp})$, i.e., there exists a finite normal extension $M_{d,k}/\mathbb{Q}$ and a class function

$$\phi_{d,k} :\ \mathrm{Gal}(M_{d,k}/\mathbb{Q}) \to \mathbb{Z}_{\geq 0}$$

such that

$$\phi_{d,k}(\mathrm{Art}_{M_{d,k}/\mathbb{Q}}(p)) = \mathrm{rk}_{2^k}\mathcal{C}\ell(\mathbb{Q}(\sqrt{dp})), \tag{6.2}$$

where $\mathrm{Art}_{M_{d,k}/\mathbb{Q}}(p)$ is the Artin conjugacy class of p in $\mathrm{Gal}(M_{d,k}/\mathbb{Q})$. This conjecture was proven for all $k \leq 3$ by Stevenhagen [70], but no governing field has been found for any value of d if $k \geq 4$. Interestingly enough, Smith [69] recently introduced the notion of relative governing fields and used them to deal with distributional questions for $\mathcal{C}\ell(K)[2^\infty]$ for imaginary quadratic fields K. Our next theorem, which we will prove in Section 6.5, is a relatively straightforward consequence of Theorem 6.1.1.

Theorem 6.1.3. *Assume conjecture C_n for all n. Then there is no governing field for the 16-rank of $\mathbb{Q}(\sqrt{-4p})$; in other words, there does not exist a field $M_{-4,4}$ and class function $\phi_{-4,4}$ satisfying* (6.2).

Acknowledgments

The authors are very grateful to Carlo Pagano for useful discussions. We would also like to thank Peter Sarnak for making us aware of the useful reference [4].

6.2 Prerequisites

Here we collect certain facts about quadratic residue symbols and unit groups in number fields that are necessary to give a rigorous definition of spins of ideals and that are useful in our subsequent arguments.

Throughout this section, let K be a number field which is Galois of degree n over $\mathbb{Q}$. Then either K is totally real, as in [24], or K is totally complex, in which case n is even. An element $\alpha \in K$ is called *totally positive* if $\iota(\alpha) > 0$ for all real embeddings $\iota : K \hookrightarrow \mathbb{R}$; if this is the case, we will write $\alpha \succ 0$. If K is totally complex, there are no real embeddings of K into $\mathbb{R}$, and so $\alpha \succ 0$ for every $\alpha \in K$ vacuously. Let $\mathcal{O}_K$ denote the ring of integers of K. If K is totally real, we assume that

$$(\mathcal{O}_K^\times)^2 = \left\{u^2 : u \in \mathcal{O}_K^\times\right\} = \left\{u \in \mathcal{O}_K^\times : u \succ 0\right\} = (\mathcal{O}_K^\times)_+, \tag{6.3}$$

where the first and last equalities are definitions and the middle equality is the assumption. This assumption, present in [24], implies that the narrow and the ordinary class groups of K coincide, and hence that every non-zero principal ideal $\mathfrak{a}$ in $\mathcal{O}_K$ can be written as $\mathfrak{a} = \alpha\mathcal{O}_K$ for some $\alpha \succ 0$. If K is totally complex, then the narrow and the ordinary class groups of K coincide vacuously. In either case, we will let $\mathcal{C}\ell = \mathcal{C}\ell(K)$ and $h = h(K)$ denote the (narrow) class group and the (narrow) class number of K.

6.2.1 Quadratic residue symbols and quadratic reciprocity

We define the quadratic residue symbol in K in the standard way. That is, given an odd prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$ (i.e., a prime ideal having odd absolute norm), and an element $\alpha \in \mathcal{O}_K$, define $\left(\frac{\alpha}{\mathfrak{p}}\right)_{K,2}$ as the unique element in $\{-1, 0, 1\}$ such that

$$\left(\frac{\alpha}{\mathfrak{p}}\right)_{K,2} \equiv \alpha^{\frac{\mathrm{N}_{K/\mathbb{Q}}(\mathfrak{p})-1}{2}} \bmod \mathfrak{p}.$$

Given an odd ideal $\mathfrak{b}$ of $\mathcal{O}_K$ with prime ideal factorization $\mathfrak{b} = \prod_{\mathfrak{p}} \mathfrak{p}^{e_{\mathfrak{p}}}$, define

$$\left(\frac{\alpha}{\mathfrak{b}}\right)_{K,2} = \prod_{\mathfrak{p}} \left(\frac{\alpha}{\mathfrak{p}}\right)_{K,2}^{e_{\mathfrak{p}}}.$$

Finally, given an element $\beta \in \mathcal{O}_K$, let (β) denote the principal ideal in $\mathcal{O}_K$ generated by β. We say that β is odd if (β) is odd and we define

$$\left(\frac{\alpha}{\beta}\right)_{K,2} = \left(\frac{\alpha}{(\beta)}\right)_{K,2}.$$

We will suppress the subscripts $K, 2$ when there is no risk of ambiguity. Although [24] focuses on a special type of totally real Galois number fields, the version of quadratic reciprocity stated in [24, Section 3] holds and was proved for a general number field. We recall it here. For a place v of K, finite or infinite, let K_v denote the completion of K with respect to v. Let $(\cdot,\cdot)_v$ denote the Hilbert symbol at v, i.e., given $\alpha, \beta \in K$, we let $(\alpha,\beta)_v \in \{-1,1\}$ with $(\alpha,\beta)_v = 1$ if and only if there exists $(x,y,z) \in K_v^3 \setminus \{(0,0,0)\}$ such that $x^2 - \alpha y^2 - \beta z^2 = 0$. As in [24, Section 3], define

$$\mu_2(\alpha,\beta) = \prod_{v\mid 2}(\alpha,\beta)_v \quad \text{and} \quad \mu_\infty(\alpha,\beta) = \prod_{v\mid\infty}(\alpha,\beta)_v.$$

The following lemma is a consequence of the Hilbert reciprocity law and local considerations at places above 2; see [24, Lemma 2.1, Proposition 2.2, and Lemma 2.3].

Lemma 6.2.1. *Let $\alpha, \beta \in \mathcal{O}_K$ with β odd. Then $\mu_\infty(\alpha,\beta)\left(\frac{\alpha}{\beta}\right)$ depends only on the congruence class of β modulo 8α. Moreover, if α is also odd, then*

$$\left(\frac{\alpha}{\beta}\right) = \mu_2(\alpha,\beta)\mu_\infty(\alpha,\beta)\left(\frac{\beta}{\alpha}\right).$$

The factor $\mu_2(\alpha,\beta)$ depends only on the congruence classes of α and β modulo 8.

We remark that if K is totally complex, then $(\alpha,\beta)_\infty = 1$ for all $\alpha,\beta \in K$. Also, if K is a totally real Galois number field and $\beta \in K$ is totally positive, then again $(\alpha,\beta)_\infty = 1$ for all $\alpha \in K$.

6.2.2 Class group representatives

As in [24, p. 707], we define a set of ideals $\mathcal{C}\ell$ and an ideal $\mathfrak{f}$ of $\mathcal{O}_K$ as follows. Let C_i, $1 \le i \le h$, denote the h ideal classes. For each $i \in \{1,\dots,h\}$, we choose two distinct odd ideals belonging to C_i, say $\mathfrak{A}_i$ and $\mathfrak{B}_i$, so as to ensure that, upon setting

$$\mathcal{C}\ell_a = \{\mathfrak{A}_1,\dots,\mathfrak{A}_h\}, \quad \mathcal{C}\ell_b = \{\mathfrak{B}_1,\dots,\mathfrak{B}_h\}, \quad \mathcal{C}\ell = \mathcal{C}\ell_a \cup \mathcal{C}\ell_b,$$

and

$$\mathfrak{f} = \prod_{\mathfrak{c}\in\mathcal{C}\ell}\mathfrak{c} = \prod_{i=1}^{h}\mathfrak{A}_i\mathfrak{B}_i,$$

the norm

$$f = \mathrm{N}(\mathfrak{f})$$

is squarefree. We define

$$F := 2^{2h+3} f D_K, \tag{6.4}$$

where D_K is the discriminant of K.

6.2.3 Definition of joint spin

We define a sequence $\{s_\mathfrak{a}\}_\mathfrak{a}$ of complex numbers indexed by non-zero ideals $\mathfrak{a} \subset \mathcal{O}_K$ as follows. Let S be a non-empty subset of $\mathrm{Gal}(K/\mathbb{Q})$ such that $\sigma \notin S$ whenever $\sigma^{-1} \in S$.

We define $r(\mathfrak{a})$ to be the indicator function of an ideal $\mathfrak{a}$ of $\mathcal{O}_K$ to be odd and principal, i.e.,

$$r(\mathfrak{a}) = \begin{cases} 1 & \text{if there exists an odd } \alpha \in \mathcal{O}_K \text{ such that } \mathfrak{a} = \alpha\mathcal{O}_K \\ 0 & \text{otherwise.} \end{cases}$$

Define $r_+(\alpha)$ to be the indicator function of an element $\alpha \in K$ to be totally positive, i.e.,

$$r_+(\alpha) = \begin{cases} 1 & \text{if } \alpha \succ 0 \\ 0 & \text{otherwise.} \end{cases}$$

Note that if K is a totally complex number field, then vacuously $r_+(\alpha) = 1$ for all α in K. If $\alpha \in K$ is odd and $r_+(\alpha) = 1$, then we define

$$\mathrm{spin}(\sigma, \alpha) = \left(\frac{\alpha}{\sigma(\alpha)}\right).$$

Fix a decomposition $\mathcal{O}_K^\times = T_K \times V_K$, where $T_K \subset \mathcal{O}_K^\times$ is the group of units of $\mathcal{O}_K$ of finite order and $V_K \subset \mathcal{O}_K^\times$ is a free abelian group of rank r_K (i.e., $r_K = n - 1$ if K is totally real and $r_K = \frac{n}{2} - 1$ if K is totally complex). With F as in (6.4), suppose that

$$\psi : (\mathcal{O}_K/F\mathcal{O}_K)^\times \to \mathbb{C} \tag{6.5}$$

is a map such that $\psi(\alpha \bmod F) = \psi(\alpha u^2 \bmod F)$ for all $\alpha \in \mathcal{O}_K$ coprime to F and all $u \in \mathcal{O}_K^\times$. We define

$$s_{\mathfrak{a}} = r(\mathfrak{a}) \sum_{t \in T_K} \sum_{v \in V_K/V_K^2} r_+(tv\alpha)\psi(tv\alpha \bmod F) \prod_{\sigma \in S} \mathrm{spin}(\sigma, tv\alpha), \tag{6.6}$$

where α is any generator of the ideal $\mathfrak{a}$ satisfying $r(\mathfrak{a}) = 1$. The averaging over V_K/V_K^2 makes the *spin* $s_{\mathfrak{a}}$ a well-defined function of $\mathfrak{a}$ since, for any unit $u \in \mathcal{O}_K^\times$, any totally positive $\alpha \in \mathcal{O}_K$ of odd absolute norm, and any $\sigma \in S$, we have

$$\mathrm{spin}(\sigma, u^2\alpha) = \left(\frac{u^2\alpha}{\sigma(u^2\alpha)}\right) = \left(\frac{u^2\alpha}{\sigma(\alpha)}\right) = \left(\frac{\alpha}{\sigma(\alpha)}\right) = \mathrm{spin}(\sigma, \alpha).$$

If K is a totally real (in which case we assume that K satisfies (6.3)), then, for an ideal $\mathfrak{a} = \alpha\mathcal{O}_K$, there is one and only one choice of $t \in T_K$ and $v \in V_K/V_K^2$ such that $r_+(tv\alpha) = 1$. Hence in this case

$$s_{\mathfrak{a}} = r(\mathfrak{a})\psi(\alpha \bmod F) \prod_{\sigma \in S} \mathrm{spin}(\sigma, \alpha),$$

where α is any totally positive generator of $\mathfrak{a}$. If in addition $n \geq 3$, $\mathrm{Gal}(K/\mathbb{Q}) = \langle\sigma\rangle$, and $S = \{\sigma\}$, then $s_{\mathfrak{a}}$ coincides with $\mathrm{spin}(\sigma, \mathfrak{a})$ in [24, (3.4), p. 706]. If we take instead $S = \{\sigma, \sigma^2\}$ and assume $n \geq 5$, then the distribution of $s_{\mathfrak{a}}$ has implications for [24, Problem, p. 744].

If K is totally complex, then vacuously $r_+(tv\alpha) = 1$ for all $t \in T_K$ and $v \in V_K/V_K^2$, so the definition of $s_{\mathfrak{a}}$ specializes to

$$s_{\mathfrak{a}} = r(\mathfrak{a}) \sum_{t \in T_K} \sum_{v \in V_K/V_K^2} \psi(tv\alpha \bmod F) \prod_{\sigma \in S} \mathrm{spin}(\sigma, tv\alpha).$$

6.2.4 Fundamental domains

We will need a suitable fundamental domain $\mathcal{D}$ for the action of the units on elements in $\mathcal{O}_K$.

In case that K is totally real and satisfies (6.3), we take $\mathcal{D} \subset \mathbb{R}_+^n$ to be the same as in [24, (4.2), p. 713]. We fix a numbering of the n real embeddings $\iota_1, \ldots, \iota_n : K \hookrightarrow \mathbb{R}$, and we say that $\alpha \in \mathcal{D}$ if and only if $(\iota_1(\alpha), \ldots, \iota_n(\alpha)) \in \mathcal{D}$. Hence every non-zero $\alpha \in \mathcal{D}$ is totally positive. Because of the assumption (6.3), every non-zero principal ideal in $\mathcal{O}_K$ has a totally positive generator, and $\mathcal{D}$ is a fundamental domain for the action of $(\mathcal{O}_K)_+^\times$ on the totally positive elements in $\mathcal{O}_K$, in the sense of [24, Lemma 4.3, p. 715].

In case that K is totally complex, we take $\mathcal{D} \subset \mathbb{R}^n$ to be the same as in [41, Lemma 3.5, p. 10]. In this case, we fix an integral basis $\{\eta_1, \ldots, \eta_n\}$ for $\mathcal{O}_K$. For an element $\alpha = a_1\eta_1 + \cdots + a_n\eta_n \in K$ with $a_1, \ldots, a_n \in \mathbb{Q}$ we say that $\alpha \in \mathcal{D}$ if and only if $(a_1, \ldots, a_n) \in \mathcal{D}$. Every non-zero principal ideal $\mathfrak{a}$ in $\mathcal{O}_K$ has exactly $|T_K|$ generators in $\mathcal{D}$; moreover, if one of the generators of $\mathfrak{a}$ in $\mathcal{D}$ is α, say, then the set of generators of $\mathfrak{a}$ in $\mathcal{D}$ is $\{t\alpha : t \in T_K\}$.

The main properties of $\mathcal{D}$ are listed in [24, Lemma 4.3, Lemma 4.4, Corollary 4.5] and [43, Lemma 3.5]. We will often use the property that if an element $\alpha \in \mathcal{D} \cap \mathcal{O}_K$ of norm $\mathrm{N}(\alpha) \leq X$ is written in an integral basis $\eta = \{\eta_1, \ldots, \eta_n\}$ as $\alpha = a_1\eta_1 + \cdots + a_n\eta_n \in \mathcal{O}_K$, $a_1, \ldots, a_n \in \mathbb{Z}$, then

$$|a_i| \ll X^{\frac{1}{n}}$$

for $1 \leq i \leq n$ where the implied constant depends only on η.

6.2.5 Short character sums

The following is a conjecture on short character sums appearing in [24]. It is essential for the estimates for sums of type I.

Conjecture 6.2.2. *For all integers $n \geq 3$ there exists $\delta(n) > 0$ such that for all $\epsilon > 0$ there exists a constant $C(n, \epsilon) > 0$ with the property that for all integers M, all integers $Q \geq 3$, all integers $N \leq Q^{\frac{1}{n}}$ and all real non-principal characters χ of modulus $q \leq Q$ we have*

$$\left| \sum_{M < m \leq M+N} \chi(m) \right| \leq C(n, \epsilon) Q^{\frac{1-\delta(n)}{n} + \epsilon}.$$

Instead of working directly with Conjecture C_n, we need a version of it for arithmetic progressions. If q is odd and squarefree, we let χ_q be the real Dirichlet character $\left(\frac{\cdot}{q}\right)$.

Corollary 6.2.3. *Assume Conjecture C_n. Then for all integers $n \geq 3$ there exists $\delta(n) > 0$ such that for all $\epsilon > 0$ there exists a constant $C(n, \epsilon) > 0$ with the property that for all odd squarefree integers $q > 1$, all integers $N \leq q^{\frac{1}{n}}$, all integers M, l and k*

with $q \nmid k$, we have

$$\left| \sum_{\substack{M<m\leq M+N \\ n\equiv l \bmod k}} \chi_q(m) \right| \leq C(n,\epsilon) q^{\frac{1-\delta(n)}{n}}.$$

Proof. This is an easy generalization of Corollary 7 in [41]. □

6.2.6 The sieve

We will prove the following oscillation results for the sequence $\{s_{\mathfrak{a}}\}_{\mathfrak{a}}$. First, for any non-zero ideal $\mathfrak{m} \subset \mathcal{O}_K$ and any $\epsilon > 0$, we have

$$\sum_{\substack{\mathrm{N}(\mathfrak{a})\leq X \\ \mathfrak{a}\equiv 0 \bmod \mathfrak{m}}} s_{\mathfrak{a}} \ll_\epsilon X^{1-\frac{\delta}{54n|S|^2}+\epsilon}, \tag{6.7}$$

where δ is as in Conjecture C_n. Second, for any $\epsilon > 0$, we have

$$\sum_{\mathrm{N}(\mathfrak{a})\leq x} \sum_{\mathrm{N}(\mathfrak{b})\leq y} v_{\mathfrak{a}} w_{\mathfrak{b}} s_{\mathfrak{a}\mathfrak{b}} \ll_\epsilon \left(x^{-\frac{1}{6n}} + y^{-\frac{1}{6n}}\right)(xy)^{1+\epsilon}, \tag{6.8}$$

for any pair of bounded sequences of complex numbers $\{v_{\mathfrak{m}}\}$ and $\{w_{\mathfrak{n}}\}$ indexed by non-zero ideals in $\mathcal{O}_K$. Then [24, Proposition 5.2, p. 722] implies that for any $\epsilon > 0$, we have

$$\sum_{\substack{\mathrm{N}(\mathfrak{p})\leq X \\ \mathfrak{p}\text{ prime ideal}}} s_{\mathfrak{p}} \ll_\epsilon X^{1-\theta+\epsilon},$$

where

$$\theta := \frac{\delta(|S|n)}{54|S|^2 n(12n+1)}.$$

Hence, in order to prove Theorem 6.1.1, it suffices to prove the estimates (6.7) and (6.8). We will deal with (6.7) in Section 6.3 and with (6.8) in Section 6.4.

6.3 Linear sums

We first treat the case that K is totally real. Let $\mathfrak{m}$ be an ideal coprime with F and $\sigma(\mathfrak{m})$ for all $\sigma \in S$. Following [24] we will bound

$$A(x) = \sum_{\substack{\mathrm{N}\mathfrak{a}\leq x \\ (\mathfrak{a},F)=1, \mathfrak{m}|\mathfrak{a}}} r(\mathfrak{a})\psi(\alpha \bmod F) \prod_{\sigma\in S} \mathrm{spin}(\sigma,\alpha), \tag{6.9}$$

where α is any totally positive generator of $\mathfrak{a}$. We pick for each ideal $\mathfrak{a}$ with $r(\mathfrak{a}) = 1$ its unique generator α satisfying $\mathfrak{a} = (\alpha)$ and $\alpha \in \mathcal{D}^*$, where $\mathcal{D}^*$ is the fundamental domain from Friedlander et al. [24]. After splitting (6.9) in residue classes modulo F we obtain

$$A(x) = \sum_{\substack{\rho \bmod F \\ (\rho,F)=1}} \psi(\rho) A(x;\rho) + \partial A(x),$$

where by definition

$$A(x;\rho) := \sum_{\substack{\alpha\in\mathcal{D}, \mathrm{N}\alpha \leq x \\ \alpha \equiv \rho \bmod F \\ \alpha \equiv 0 \bmod \mathfrak{m}}} \prod_{\sigma \in S} \mathrm{spin}(\sigma, \alpha). \tag{6.10}$$

The boundary term $\partial A(x)$ can be dealt with using the argument in [24, p. 724], which gives $\partial A(x) \ll x^{1-\frac{1}{n}}$. Here and in the rest of our arguments the implied constant depends only on K unless otherwise indicated. We will now estimate $A(x;\rho)$ for each $\rho \bmod F$, $(\rho, F) = 1$. Let $1, \omega_2, \ldots, \omega_n$ be an integral basis for $\mathcal{O}_K$ and define

$$\mathbb{M} := \omega_2\mathbb{Z} + \cdots + \omega_n\mathbb{Z}.$$

Then, just as in [24, p. 725], we can decompose α uniquely as

$$\alpha = a + \beta, \qquad \text{with } a \in \mathbb{Z}, \beta \in \mathbb{M}.$$

Hence the summation conditions in (6.10) can be rewritten as

$$a + \beta \in \mathcal{D}, \quad \mathrm{N}(a+\beta) \leq x, \quad a + \beta \equiv \rho \bmod F, \quad a + \beta \equiv 0 \bmod \mathfrak{m}. \tag{$*$}$$

From now on we think of a as a variable satisfying $(*)$ while β is inactive. We have the following formula

$$\mathrm{spin}(\sigma, \alpha) = \left(\frac{\alpha}{\sigma(\alpha)}\right) = \left(\frac{a+\beta}{a+\sigma(\beta)}\right) = \left(\frac{\beta - \sigma(\beta)}{a+\sigma(\beta)}\right).$$

If $\beta = \sigma(\beta)$ for some $\sigma \in S$ we get no contribution. So from now on we can assume $\beta \neq \sigma(\beta)$ for all $\sigma \in S$. Define $\mathfrak{c}(\sigma, \beta)$ to be the part of the ideal $(\beta - \sigma(\beta))$ coprime to F. Then, as explained on [24, p. 726], quadratic reciprocity gives

$$A(x;\rho) = \sum_{\beta \in \mathbb{M}} \pm T(x;\rho,\beta),$$

where $T(x;\rho,\beta)$ is given by

$$\begin{aligned} T(x;\rho,\beta) := \sum_{\substack{a\in\mathbb{Z} \\ a+\beta \text{ sat. } (*)}} \prod_{\sigma\in S} \left(\frac{a+\sigma(\beta)}{\mathfrak{c}(\sigma,\beta)}\right) &= \sum_{\substack{a\in\mathbb{Z} \\ a+\beta \text{ sat. } (*)}} \prod_{\sigma\in S} \left(\frac{a+\beta}{\mathfrak{c}(\sigma,\beta)}\right) \\ &= \sum_{\substack{a\in\mathbb{Z} \\ a+\beta \text{ sat. } (*)}} \left(\frac{a+\beta}{\prod_{\sigma\in S}\mathfrak{c}(\sigma,\beta)}\right). \end{aligned} \tag{6.11}$$

Define $\mathfrak{c} := \prod_{\sigma \in S} \mathfrak{c}(\sigma, \beta)$ and factor $\mathfrak{c}$ as

$$\mathfrak{c} = \mathfrak{g}\mathfrak{q}, \tag{6.12}$$

where by definition $\mathfrak{g}$ consists of those prime ideals $\mathfrak{p}$ dividing $\mathfrak{c}$ that satisfy one of the following three properties

- $\mathfrak{p}$ has degree greater than one;
- $\mathfrak{p}$ is unramified of degree one and some non-trivial conjugate of $\mathfrak{p}$ also divides $\mathfrak{c}$;
- $\mathfrak{p}$ is unramified of degree one and $\mathfrak{p}^2$ divides $\mathfrak{c}$.

Note that there are no ramified primes dividing $\mathfrak{c}$, since $\mathfrak{c}$ is coprime to the discriminant by construction of F. Putting all the remaining prime ideals in $\mathfrak{q}$, we note that $q := \mathrm{N}\mathfrak{q}$ is a squarefree number and $g := \mathrm{N}\mathfrak{g}$ is a squarefull number coprime with q. The Chinese Remainder Theorem implies that there exists a rational integer b with $b \equiv \beta \bmod \mathfrak{q}$. We stress that $\mathfrak{c}$, $\mathfrak{g}$, $\mathfrak{q}$, g, q and b depend only on β. Define g_0 to be the radical of g. Then the quadratic residue symbol $(\alpha/\mathfrak{g})$ is periodic in α modulo g_0. Hence the symbol $((a+\beta)/\mathfrak{g})$ as a function of a is periodic of period g_0. Splitting the sum (6.11) in residue classes modulo g_0 we obtain

$$|T(x; \rho, \beta)| \leq \sum_{a_0 \bmod g_0} \left| \sum_{\substack{a \equiv a_0 \bmod g_0 \\ a+\beta \text{ sat. } (*)}} \left(\frac{a+b}{\mathfrak{q}}\right) \right|. \tag{6.13}$$

Following the argument on [24, p. 728], we see that (6.13) can be written as n incomplete character sums of length $\ll x^{\frac{1}{n}}$ and modulus $q \ll x^{|S|}$. Furthermore, the conditions $(*)$ and $a \equiv a_0 \bmod g_0$ imply that a runs over a certain arithmetic progression of modulus k dividing $g_0 F m$, where $m := \mathrm{N}\mathfrak{m}$. So if $q \nmid k$, Corollary 6.2.3 yields

$$T(x; \rho, \beta) \ll_\epsilon g_0 x^{\frac{1-\delta}{n} + \epsilon} \tag{6.14}$$

with $\delta := \delta(|S|n) > 0$. Since $q \mid k$ implies $q \mid m$, we see that (6.14) holds if $q \nmid m$. Recalling (6.12) we conclude that (6.14) holds unless

$$p \mid \prod_{\sigma \in S} \mathrm{N}(\beta - \sigma(\beta)) \Rightarrow p^2 \mid mF \prod_{\sigma \in S} \mathrm{N}(\beta - \sigma(\beta)). \tag{6.15}$$

Our next goal is to count the number of $\beta \in \mathbb{M}$ satisfying both $(*)$ for some $a \in \mathbb{Z}$ and (6.15). For β an algebraic integer of degree n, we denote by $\beta^{(1)}, \ldots, \beta^{(n)}$ the conjugates of β. Now if β satisfies $(*)$ for some $a \in \mathbb{Z}$, we have $|\beta^{(i)}| \ll x^{\frac{1}{n}}$. So to achieve our goal, it suffices to estimate the number of $\beta \in \mathbb{M}$ satisfying $|\beta^{(i)}| \leq x^{\frac{1}{n}}$ and (6.15).

To do this, we will need two lemmas. So far we have followed [24] rather closely, but we will have to significantly improve their estimates for the various error terms given on [24, p. 729-733]. One of the most important tasks ahead is to count squarefull norms in a certain $\mathbb{Z}$-submodule of $\mathcal{O}_K$. This problem is solved in [24] by simply counting squarefull norms in the full ring of integers. For our application this loss is unacceptable. In our first lemma we directly count squarefull norms in this submodule, a problem described in [24, p. 729] as potentially "very difficult".

Lemma 6.3.1. *Factor* $\mathfrak{c}(\sigma,\beta)$ *as*

$$\mathfrak{c}(\sigma,\beta)=\mathfrak{g}(\sigma,\beta)\mathfrak{q}(\sigma,\beta)$$

just as in (6.12). Let K^σ *be the subfield of* K *fixed by* σ *and let* $\mathcal{O}_{K^\sigma}$ *be its ring of integers. Decompose* $\mathcal{O}_K$ *as*

$$\mathcal{O}_K=\mathcal{O}_{K^\sigma}\oplus\mathbb{M}'.$$

Let $ord(\sigma)$ *be the order of* σ *in* $\mathrm{Gal}(K/\mathbb{Q})$. *If* $g_0(\sigma,\beta)$ *is the radical of* $\mathrm{N}\mathfrak{g}(\sigma,\beta)$, *then we have for all* $\epsilon>0$

$$|\{\beta\in\mathbb{M}':|\beta^{(i)}|\le x^{\frac{1}{n}},g_0(\sigma,\beta)>Z\}|\ll_\epsilon x^{1-\frac{1}{ord(\sigma)}+\epsilon}Z^{-1+\frac{2}{ord(\sigma)}}.$$

Proof. The argument given here is a generalization of [41, p. 17-18]. We start with the simple estimate

$$|\{\beta\in\mathbb{M}':|\beta^{(i)}|\le x^{\frac{1}{n}},g_0(\sigma,\beta)>Z\}|\le\sum_{\substack{\mathfrak{g}\\ g_0>Z}}A_{\mathfrak{g}},\tag{6.16}$$

where

$$A_{\mathfrak{g}}:=|\{\beta\in\mathbb{M}':|\beta^{(i)}|\le x^{\frac{1}{n}},\beta-\sigma(\beta)\equiv 0 \bmod \mathfrak{g}\}|.$$

Let $\mathbb{M}''$ be the image of $\mathbb{M}'$ under the map $\beta\mapsto\beta-\sigma(\beta)$ and fix a $\mathbb{Z}$-basis $\eta_1,\dots,\eta_r$ of $\mathbb{M}''$. We remark that $r=n\left(1-\frac{1}{\mathrm{ord}(\sigma)}\right)$, which will be important later on. Because $|\beta^{(i)}|\le x^{\frac{1}{n}}$, we can write $\beta-\sigma(\beta)$ as $\beta-\sigma(\beta)=\sum_{i=1}^r a_i\eta_i$ with $|a_i|\le C_Kx^{\frac{1}{n}}$, where C_K is a constant depending only on K. Hence we have

$$A_{\mathfrak{g}}\le|\Lambda_{\mathfrak{g}}\cap S_x|,$$

where by definition

$$\Lambda_{\mathfrak{g}}:=\{\gamma\in\mathbb{M}'':\gamma\equiv 0 \bmod \mathfrak{g}\}$$
$$S_x:=\{\gamma\in\mathbb{M}'':\gamma=\sum_{i=1}^r a_i\eta_i,|a_i|\le C_Kx^{\frac{1}{n}}\}.$$

Using our fixed $\mathbb{Z}$-basis $\eta_1,\dots,\eta_r$ we can view $\mathbb{M}''$ as a subset of $\mathbb{R}^r$ via the map $\eta_i\mapsto e_i$, where e_i is the i-th standard basis vector. Under this identification $\mathbb{M}''$ becomes $\mathbb{Z}^r$ and $\Lambda_{\mathfrak{g}}$ becomes a sublattice of $\mathbb{Z}^r$. We have

$$A_{\mathfrak{g}}\le|\Lambda_{\mathfrak{g}}\cap T_x|,\tag{6.17}$$

where

$$T_x := \{(a_1, \ldots, a_r) \in \mathbb{R}^r : |a_i| \leq C_K x^{\frac{1}{n}}\}.$$

Let us now parametrize the boundary of T_x. We start off by observing that $T_x = x^{\frac{1}{n}} T_1$, which implies that $\mathrm{Vol}(T_x) = x^{\frac{r}{n}} \mathrm{Vol}(T_1)$. Because T_1 is an r-dimensional hypercube, we conclude that its boundary ∂T_1 can be parametrized by Lipschitz functions with Lipschitz constant L depending only on K. Therefore ∂T_x can also be parametrized by Lipschitz functions with Lipschitz constant $x^{\frac{1}{n}} L$. Theorem 5.4 of [79] gives

$$\left| |\Lambda_{\mathfrak{g}} \cap T_x| - \frac{\mathrm{Vol}(T_x)}{\det \Lambda_{\mathfrak{g}}} \right| \ll_L \max_{0 \leq i < r} \frac{x^{\frac{i}{n}}}{\lambda_{\mathfrak{g},1} \cdot \ldots \cdot \lambda_{\mathfrak{g},i}}, \tag{6.18}$$

where $\lambda_{\mathfrak{g},1}, \ldots, \lambda_{\mathfrak{g},r}$ are the successive minima of $\Lambda_{\mathfrak{g}}$. Since L depends only on K, it follows that the implied constant in (6.18) depends only on K, so we may simply write $\ll$ by our earlier conventions.

Our next goal is to give a lower bound for $\lambda_{\mathfrak{g},1}$. So let $\gamma \in \Lambda_{\mathfrak{g}}$ be non-zero. By definition of $\Lambda_{\mathfrak{g}}$ we have $\mathfrak{g} \mid \gamma$ and hence $g \mid \mathrm{N}\gamma$. Write

$$\gamma = \sum_{i=1}^{r} a_i \eta_i.$$

If $a_1, \ldots, a_r \leq C'_K g^{\frac{1}{n}}$ for a sufficiently small constant C'_K, we find that $\mathrm{N}\gamma < g$. But this is impossible, since $g \mid \mathrm{N}\gamma$ and $\mathrm{N}\gamma \neq 0$. So there is an i with $a_i > C'_K g^{\frac{1}{n}}$. If we equip $\mathbb{R}^r$ with the standard Euclidean norm, we conclude that the length of γ satisfies $||\gamma|| \gg g^{\frac{1}{n}}$ and hence

$$\lambda_{\mathfrak{g},1} \gg g^{\frac{1}{n}}. \tag{6.19}$$

Minkowski's second theorem and (6.19) imply that

$$\det \Lambda_{\mathfrak{g}} \gg g^{\frac{r}{n}}. \tag{6.20}$$

Combining (6.18), (6.19), (6.20) and $g \leq x$ gives

$$|\Lambda_{\mathfrak{g}} \cap T_x| \ll \frac{x^{\frac{r}{n}}}{g^{\frac{r}{n}}} + \frac{x^{\frac{r-1}{n}}}{g^{\frac{r-1}{n}}} \ll \frac{x^{\frac{r}{n}}}{g^{\frac{r}{n}}}. \tag{6.21}$$

Plugging (6.17) and (6.21) back in (6.16) yields

$$|\{\beta \in \mathbb{M}' : |\beta^{(i)}| \leq x^{\frac{1}{n}}, g_0(\sigma, \beta) > Z\}| \leq \sum_{\substack{\mathfrak{g} \\ g_0 > Z}} A_{\mathfrak{g}} \leq \sum_{\substack{\mathfrak{g} \\ g_0 > Z}} |\Lambda_{\mathfrak{g}} \cap T_x| \ll \sum_{\substack{\mathfrak{g} \\ g_0 > Z}} \frac{x^{\frac{r}{n}}}{g^{\frac{r}{n}}}.$$

If we define $\tau_K(g)$ to be the number of ideals of K of norm g, we can bound the last sum as follows

$$\sum_{\substack{\mathfrak{g} \\ g_0>Z}} \frac{x^{\frac{r}{n}}}{g^{\frac{r}{n}}} = x^{\frac{r}{n}} \sum_{\substack{g\leq x \\ g \text{ squarefull} \\ g_0>Z}} \frac{\tau_K(g)}{g^{\frac{r}{n}}} \ll_\epsilon x^{\frac{r}{n}+\epsilon} \sum_{\substack{g\leq x \\ g \text{ squarefull} \\ g_0>Z}} \frac{1}{g^{\frac{r}{n}}}$$

$$= x^{\frac{r}{n}+\epsilon} \sum_{\substack{g\leq x \\ g \text{ squarefull} \\ g_0>Z}} g^{\frac{1}{2}-\frac{r}{n}} \frac{1}{g^{\frac{1}{2}}} \leq x^{\frac{r}{n}+\epsilon} Z^{1-\frac{2r}{n}} \sum_{\substack{g\leq x \\ g \text{ squarefull} \\ g_0>Z}} \frac{1}{g^{\frac{1}{2}}}$$

$$\leq x^{\frac{r}{n}+\epsilon} Z^{1-\frac{2r}{n}} \sum_{\substack{g\leq x \\ g \text{ squarefull}}} \frac{1}{g^{\frac{1}{2}}} \ll_\epsilon x^{\frac{r}{n}+\epsilon} Z^{1-\frac{2r}{n}}.$$

Recalling that $r = n\left(1 - \frac{1}{\operatorname{ord}(\sigma)}\right)$ completes the proof of Lemma 6.3.1. □

Lemma 6.3.2. *Let $\sigma, \tau \in S$ be distinct. Recall that*

$$\mathcal{O}_K = \mathbb{Z} \oplus \mathbb{M}.$$

Fix an integral basis $\omega_2, \ldots, \omega_n$ of $\mathbb{M}$ and define the polynomials $f_1, f_2 \in \mathbb{Z}[x_2, \ldots, x_n]$ by

$$f_1(x_2, \ldots, x_n) = \mathrm{N}\left(\sum_{i=2}^n x_i(\sigma(\omega_i) - \omega_i)\right)$$

$$f_2(x_2, \ldots, x_n) = \mathrm{N}\left(\sum_{i=2}^n x_i(\tau(\omega_i) - \omega_i)\right).$$

For $\beta \in \mathbb{M}$ with $\beta = \sum_{i=2}^n a_i\omega_i$ we define $f_1(\beta) := f_1(a_2, \ldots, a_n) = \mathrm{N}(\sigma(\beta) - \beta)$ and similarly for $f_2(\beta)$. Then

$$|\{\beta \in \mathbb{M} : |\beta^{(i)}| \leq x^{\frac{1}{n}}, \gcd(f_1(\beta), f_2(\beta)) > Z\}| \ll_\epsilon x^{\frac{n-1}{n}+\epsilon} Z^{-\frac{1}{18}} + x^{\frac{n-2}{n}} + Z^{\frac{2n-4}{3}}.$$

Proof. Let Y be the closed subscheme of $\mathbb{A}^{n-1}_{\mathbb{Z}}$ defined by $f_1 = f_2 = 0$. We claim that Y has codimension 2, i.e. f_1 and f_2 are relatively prime polynomials. Suppose not. Note that f_1 and f_2 factor in $K[x_2, \ldots, x_n]$ as

$$f_1(x_2, \ldots, x_n) = \prod_{\sigma' \in \mathrm{Gal}(K/\mathbb{Q})} \left(\sum_{i=2}^n x_i(\sigma'\sigma(\omega_i) - \sigma'(\omega_i))\right)$$

$$f_2(x_2, \ldots, x_n) = \prod_{\tau' \in \mathrm{Gal}(K/\mathbb{Q})} \left(\sum_{i=2}^n x_i(\tau'\tau(\omega_i) - \tau'(\omega_i))\right).$$

Hence if f_1 and f_2 are not relatively prime, there are $\sigma', \tau' \in \mathrm{Gal}(K/\mathbb{Q})$ and $\kappa \in K^*$ such that

$$\sum_{i=2}^n x_i(\sigma'\sigma(\omega_i) - \sigma'(\omega_i)) = \kappa \sum_{i=2}^n x_i(\tau'\tau(\omega_i) - \tau'(\omega_i))$$

for all $x_2, \ldots, x_n \in \mathbb{Z}$. Put $\beta = \sum_{i=2}^n x_i\omega_i$. Then we can rewrite this as

$$\sigma'\sigma(\beta) - \sigma'(\beta) = \kappa(\tau'\tau(\beta) - \tau'(\beta)) \tag{6.22}$$

for all $\beta \in \mathbb{M}$. But this implies that (6.22) holds for all $\beta \in K$. Now we apply the Artin-Dedekind Lemma, which gives a contradiction in all cases due to our assumptions $\sigma, \tau \in S$ and $\sigma \neq \tau$.

Having established our claim, we are in position to apply Theorem 3.3 of [4]. We embed $\mathbb{M}$ in $\mathbb{R}^{n-1}$ by sending ω_i to e_i, the i-th standard basis vector. Note that the image under this embedding is $\mathbb{Z}^{n-1}$. Write $\beta = \sum_{i=2}^n a_i\omega_i$. Since $|\beta^{(i)}| \leq x^{\frac{1}{n}}$, it follows that $|a_i| \leq C_K x^{\frac{1}{n}}$ for some constant C_K depending only on K. Let B be the compact region in $\mathbb{R}^{n-1}$ given by $B := \{(a_2, \ldots, a_n) : |a_i| \leq C_K\}$. Theorem 3.3 of [4] with our B, Y and $r = x^{\frac{1}{n}}$ gives

$$|\{\beta \in \mathbb{M} : |\beta^{(i)}| \leq x^{\frac{1}{n}}, p \mid \gcd(f_1(\beta), f_2(\beta)), p > M\}| \ll \frac{x^{\frac{n-1}{n}}}{M \log M} + x^{\frac{n-2}{n}}, \tag{6.23}$$

where M is any positive real number. Factor

$$\begin{aligned} f_1(\beta) &:= g_1q_1, \quad (g_1, q_1) = 1, \quad g_1 \text{ squarefull}, \quad q_1 \text{ squarefree} \\ f_2(\beta) &:= g_2q_2, \quad (g_2, q_2) = 1, \quad g_2 \text{ squarefull}, \quad q_2 \text{ squarefree.} \end{aligned}$$

By Lemma 6.3.1 we conclude that for all $A > 0$ and $\epsilon > 0$

$$|\{\beta \in \mathbb{M} : |\beta^{(i)}| \leq x^{\frac{1}{n}}, g_1 > A\}| \ll_\epsilon x^{\frac{n-1}{n}+\epsilon} A^{-\frac{1}{2}+\frac{1}{\mathrm{ord}(\sigma)}}.$$

With the same argument applied to τ we obtain

$$|\{\beta \in \mathbb{M} : |\beta^{(i)}| \leq x^{\frac{1}{n}}, g_1 > A \text{ or } g_2 > A\}| \ll_\epsilon x^{\frac{n-1}{n}+\epsilon} A^{-\frac{1}{2}+\frac{1}{\mathrm{ord}(\sigma)}} + x^{\frac{n-1}{n}+\epsilon} A^{-\frac{1}{2}+\frac{1}{\mathrm{ord}(\tau)}}. \tag{6.24}$$

We discard those β that satisfy (6.23) or (6.24). From (6.24) we deduce that the remaining β certainly satisfy $\gcd(q_1, q_2) > \frac{Z}{A^2}$. Furthermore, by discarding those β satisfying (6.23), we see that $\gcd(q_1, q_2)$ has no prime divisors greater than M. This implies that $\gcd(q_1, q_2)$ is divisible by a squarefree number between $\frac{Z}{A^2}$ and $\frac{ZM}{A^2}$. So we must still give an upper bound for

$$\left|\left\{\beta \in \mathbb{M} : |\beta^{(i)}| \leq x^{\frac{1}{n}}, r \mid \gcd(q_1, q_2), \frac{Z}{A^2} < r \leq \frac{ZM}{A^2}\right\}\right|. \tag{6.25}$$

Let r be a squarefree integer and let $\mathfrak{r}_1, \mathfrak{r}_2$ be two ideals of K with norm r. Define

$$E_{\mathfrak{r}_1,\mathfrak{r}_2} := \left|\left\{\beta \in \mathbb{M} : |\beta^{(i)}| \leq x^{\frac{1}{n}}, \mathfrak{r}_1 \mid \sigma(\beta) - \beta, \mathfrak{r}_2 \mid \tau(\beta) - \beta\right\}\right|.$$

We will give an upper bound for $E_{\mathfrak{r}_1,\mathfrak{r}_2}$ following [24, p. 731-733]. Write $\beta = \sum_{i=2}^n a_i\omega_i$. Then $|\beta^{(i)}| \leq x^{\frac{1}{n}}$ implies $a_i \ll x^{\frac{1}{n}}$ and

$$\sum_{i=2}^n a_i(\sigma(\omega_i) - \omega_i) \equiv 0 \bmod \mathfrak{r}_1 \tag{6.26}$$

$$\sum_{i=2}^n a_i(\tau(\omega_i) - \omega_i) \equiv 0 \bmod \mathfrak{r}_2. \tag{6.27}$$

We split the coefficients $a_2, \dots, a_n$ according to their residue classes modulo r. Suppose that $p \mid r$ and let $\mathfrak{p}_1$, $\mathfrak{p}_2$ be the unique prime ideals of degree one dividing $\mathfrak{r}_1$ and $\mathfrak{r}_2$ respectively. Then we get

$$\sum_{i=2}^{n} a_i(\sigma(\omega_i) - \omega_i) \equiv 0 \bmod \mathfrak{p}_1 \tag{6.28}$$

$$\sum_{i=2}^{n} a_i(\tau'\tau(\omega_i) - \tau'(\omega_i)) \equiv 0 \bmod \mathfrak{p}_1, \tag{6.29}$$

where τ' satisfies $\tau'^{-1}(\mathfrak{p}_1) = \mathfrak{p}_2$. If we further assume that $\mathfrak{p}_1$ is unramified, we claim that the above two equations are linearly independent over $\mathbb{F}_p$. Indeed, consider the isomorphism

$$\mathcal{O}_K/p \cong \mathbb{F}_p \times \cdots \times \mathbb{F}_p.$$

Note that $\tau'\tau \notin \{\mathrm{id}, \sigma\}$ or $\tau' \notin \{\mathrm{id}, \sigma\}$ due to our assumption that σ and τ are distinct elements of S. Let us deal with the case $\tau'\tau \notin \{\mathrm{id}, \sigma\}$, the other case is dealt with similarly. Then there exists $\beta \in \mathcal{O}_K$ such that $\beta \equiv 1 \bmod \mathfrak{p}_1$, $\beta \equiv 1 \bmod \sigma^{-1}(\mathfrak{p}_1)$, $\beta \equiv 1 \bmod \tau'^{-1}(\mathfrak{p}_1)$ and β is divisible by all other conjugates of $\mathfrak{p}_1$. By our assumption on $\tau'\tau$ it follows that $\beta \equiv 0 \bmod \tau^{-1}\tau'^{-1}(\mathfrak{p}_1)$. Hence we obtain

$$\sigma(\beta) - \beta \equiv 0 \bmod \mathfrak{p}_1, \quad \tau'\tau(\beta) - \tau'(\beta) \equiv -1 \bmod \mathfrak{p}_1.$$

However, for $\mathfrak{p}_1$ an unramified prime, we know that $\sigma(\beta) - \beta \equiv 0 \bmod \mathfrak{p}_1$ can not happen for all $\beta \in \mathcal{O}_K$, unless σ is the identity. This proves our claim.

If we further split the coefficients $a_2, \dots, a_n$ according to their residue classes modulo p, our claim implies that there are p^{n-3} solutions $a_2, \dots, a_n$ modulo p satisfying (6.28) and (6.29), provided that p is unramified. For ramified primes we can use the trivial upper bound p^{n-1}. Then we deduce from the Chinese Remainder Theorem that there are $\ll r^{n-3}$ solutions $a_2, \dots, a_n$ modulo r satisfying (6.26) and (6.27). This yields

$$E_{\mathfrak{r}_1, \mathfrak{r}_2} \ll r^{n-3} \left(\frac{x^{\frac{1}{n}}}{r} + 1 \right)^{n-1} \ll x^{\frac{n-1}{n}} r^{-2} + r^{n-3}.$$

Therefore we have the following upper bound for (6.25)

$$\begin{aligned}
\sum_{\frac{Z}{A^2} < r \leq \frac{ZM}{A^2}} \sum_{\substack{\mathfrak{r}_1, \mathfrak{r}_2 \\ \mathrm{N}\mathfrak{r}_1 = \mathrm{N}\mathfrak{r}_2 = r}} E_{\mathfrak{r}_1, \mathfrak{r}_2} &\ll \sum_{\frac{Z}{A^2} < r \leq \frac{ZM}{A^2}} \sum_{\substack{\mathfrak{r}_1, \mathfrak{r}_2 \\ \mathrm{N}\mathfrak{r}_1 = \mathrm{N}\mathfrak{r}_2 = r}} x^{\frac{n-1}{n}} r^{-2} + r^{n-3} \\
&\ll_\epsilon x^\epsilon \sum_{\frac{Z}{A^2} < r \leq \frac{ZM}{A^2}} x^{\frac{n-1}{n}} r^{-2} + r^{n-3} \\
&\ll_\epsilon x^\epsilon \left(x^{\frac{n-1}{n}} \frac{A^2}{Z} + \left(\frac{ZM}{A^2} \right)^{n-2} \right).
\end{aligned}$$

Note that $\sigma \in S$ implies $\mathrm{ord}(\sigma) \geq 3$. Now choose $A = M = Z^{\frac{1}{3}}$ to complete the proof of Lemma 6.3.2. $\square$

With Lemma 6.3.1 and Lemma 6.3.2 in hand we return to estimating the number of $\beta \in \mathbb{M}$ satisfying $|\beta^{(i)}| \leq x^{\frac{1}{n}}$ and (6.15). We choose a $\sigma \in S$ and we will consider it as fixed for the remainder of the proof. Note that any integer $n > 0$ can be factored uniquely as

$$n = q'g'r',$$

where q' is a squarefree integer coprime to mF, g' is a squarefull integer coprime to mF and r' is composed entirely of primes from mF. This allows us to define $\mathrm{sqf}(n, mF) := q'$. We start by giving an upper bound for

$$\left|\left\{\beta \in \mathbb{M} : |\beta^{(i)}| \leq x^{\frac{1}{n}}, \mathrm{sqf}(\mathrm{N}(\beta - \sigma(\beta)), mF) \leq Z\right\}\right|.$$

To do this, we need a slight generalization of the argument on [24, p. 729]. Recall that K^σ is the subfield of K fixed by σ and $\mathcal{O}_{K^\sigma}$ its ring of integers. Decompose $\mathcal{O}_K$ as

$$\mathcal{O}_K = \mathcal{O}_{K^\sigma} \oplus \mathbb{M}'.$$

Then we have

$$\begin{aligned}\left|\left\{\beta \in \mathbb{M} : |\beta^{(i)}| \leq x^{\frac{1}{n}}, \mathrm{sqf}(\mathrm{N}(\beta - \sigma(\beta)), mF) \leq Z\right\}\right|& \\ \ll x^{\frac{1}{\mathrm{ord}(\sigma)} - \frac{1}{n}} \left|\left\{\beta \in \mathbb{M}' : |\beta^{(i)}| \leq x^{\frac{1}{n}}, \mathrm{sqf}(\mathrm{N}(\beta - \sigma(\beta)), mF) \leq Z\right\}\right|.& \quad (6.30)\end{aligned}$$

The map $\mathbb{M}' \to \mathcal{O}_K$ given by $\beta \mapsto \beta - \sigma(\beta)$ is injective. Set $\gamma := \beta - \sigma(\beta)$. Furthermore, the conjugates of γ satisfy $|\gamma^{(i)}| \leq 2x^{\frac{1}{n}}$, which gives

$$\begin{aligned}\left|\left\{\beta \in \mathbb{M}' : |\beta^{(i)}| \leq x^{\frac{1}{n}}, \mathrm{sqf}(\mathrm{N}(\beta - \sigma(\beta)), mF) \leq Z\right\}\right|& \\ \leq \left|\left\{\gamma \in \mathcal{O}_K : |\gamma^{(i)}| \leq 2x^{\frac{1}{n}}, \mathrm{sqf}(\mathrm{N}(\gamma), mF) \leq Z\right\}\right|.& \quad (6.31)\end{aligned}$$

Instead of counting algebraic integers γ, we will count the principal ideals they generate, where each given ideal occurs no more than $\ll (\log x)^n$ times. This yields the bound

$$\begin{aligned}\left|\left\{\gamma \in \mathcal{O}_K : |\gamma^{(i)}| \leq 2x^{\frac{1}{n}}, \mathrm{sqf}(\mathrm{N}(\gamma), mF) \leq Z\right\}\right|& \\ \ll (\log x)^n \left|\left\{\mathfrak{b} \subseteq \mathcal{O}_K : \mathrm{N}(\mathfrak{b}) \leq 2^n x, \mathrm{sqf}(\mathrm{N}(\mathfrak{b}), mF) \leq Z\right\}\right|.&\end{aligned}$$

We conclude that

$$\left|\left\{\gamma \in \mathcal{O}_K : |\gamma^{(i)}| \leq 2x^{\frac{1}{n}}, \mathrm{sqf}(\mathrm{N}(\gamma), mF) \leq Z\right\}\right| \ll (\log x)^n \sum_{\substack{b \leq 2^n x \\ \mathrm{sqf}(b, mF) \leq Z}} \tau_K(b), \quad (6.32)$$

where we remind the reader that $\tau_K(b)$ denotes the number of ideals in K of norm b.

Let us count the number of $b \leq 2^n x$ satisfying $\mathrm{sqf}(b, mF) \leq Z$. We do this by counting the number of possible $g', r' \leq 2^n x$ that can occur in the factorization $b = q'g'r'$. First of all, there are $\ll x^{\frac{1}{2}}$ squarefull integers g' satisfying $g' \leq 2^n x$. To bound the number of $r' \leq 2^n x$, we observe that we may assume $m \leq x$, because otherwise the sum in (6.9) is empty. This implies that the number of integers $r' \leq 2^n x$ that are composed entirely of primes from mF is $\ll_\epsilon x^\epsilon$. Obviously there are at most Z squarefree integers q' coprime to mF satisfying $q' \leq Z$. We conclude that the number of $b \leq 2^n x$ satisfying $\mathrm{sqf}(b, mF) \leq Z$ is $\ll_\epsilon Z x^{\frac{1}{2}+\epsilon}$. Combined with the upper bound $\tau_K(b) \ll_\epsilon x^\epsilon$ we obtain

$$(\log x)^n \sum_{\substack{b \leq 2^n x \\ \mathrm{sqf}(b, mF) \leq Z}} \tau_K(b) \ll_\epsilon Z x^{\frac{1}{2}+\epsilon}. \tag{6.33}$$

Stringing together the inequalities (6.30), (6.31), (6.32) and (6.33) we conclude that

$$\left|\left\{\beta \in \mathbb{M} : |\beta^{(i)}| \leq x^{\frac{1}{n}}, \mathrm{sqf}(\mathrm{N}(\beta - \sigma(\beta)), mF) \leq Z\right\}\right| \ll_\epsilon Z x^{\frac{1}{2}+\frac{1}{\mathrm{ord}(\sigma)}-\frac{1}{n}+\epsilon}. \tag{6.34}$$

Now in order to give an upper bound for the number of β satisfying $|\beta^{(i)}| \leq x^{\frac{1}{n}}$ and (6.15), that is

$$p \mid \prod_{\sigma \in S} \mathrm{N}(\beta - \sigma(\beta)) \Rightarrow p^2 \mid mF \prod_{\sigma \in S} \mathrm{N}(\beta - \sigma(\beta)),$$

we start by picking $Z = x^{\frac{1}{3n}}$ and discarding all β satisfying (6.34) for the $\sigma \in S$ we fixed earlier. For this $\sigma \in S$ and varying $\tau \in S$ with $\tau \neq \sigma$ we apply Lemma 6.3.2 to obtain

$$|\{\beta \in \mathbb{M} : |\beta^{(i)}| \leq x^{\frac{1}{n}}, \gcd(\mathrm{N}(\beta - \sigma(\beta)), \mathrm{N}(\beta - \tau(\beta))) > x^{\frac{1}{3n|S|}}\}| \ll_\epsilon x^{\frac{n-1}{n}-\frac{1}{54n|S|}+\epsilon}. \tag{6.35}$$

We further discard all β satisfying (6.35) for some $\tau \in S$ with $\tau \neq \sigma$. Now it is easily checked that the remaining β do not satisfy (6.15). Hence we have completed our task of estimating the number of β satisfying $|\beta^{(i)}| \leq x^{\frac{1}{n}}$ and (6.15).

Let $A_0(x; \rho)$ be the contribution to $A(x; \rho)$ of the terms $\alpha = a + \beta$ for which (6.15) does not hold and let $A_\square(x; \rho)$ be the contribution to $A(x; \rho)$ for which (6.15) holds. Then we have the obvious identity

$$A(x; \rho) = A_0(x; \rho) + A_\square(x; \rho).$$

Next we make a further partition

$$A_0(x; \rho) = A_1(x; \rho) + A_2(x; \rho),$$

where the components run over $\alpha = a + \beta$, $\beta \in \mathbb{M}$ with β such that

$$\begin{aligned} g_0 &\leq Y \text{ in } A_1(x; \rho) \\ g_0 &> Y \text{ in } A_2(x; \rho). \end{aligned}$$

Here Y is at our disposal and we choose it later. From (6.34) and (6.35) we deduce that

$$A_\square(x; \rho) \ll_\epsilon x^{1-\frac{1}{54n|S|}+\epsilon}.$$

To estimate $A_1(x;\rho)$ we apply 6.14 and sum over all $\beta \in \mathbb{M}$ satisfying $|\beta^{(i)}| \leq x^{\frac{1}{n}}$, ignoring all other restrictions on β, to obtain

$$A_1(x;\rho) \ll_\epsilon Y x^{1-\frac{\delta}{n}+\epsilon}.$$

We still have to bound $A_2(x;\rho)$. Recall that

$$\mathfrak{c} = \prod_{\sigma \in S} \mathfrak{c}(\sigma, \beta),$$

leading to the factorization $\mathfrak{c} = \mathfrak{g}\mathfrak{q}$ in (6.12). We further recall that g_0 is the radical of $N\mathfrak{g}$. Now factor each term $\mathfrak{c}(\sigma, \beta)$ as

$$\mathfrak{c}(\sigma, \beta) = \mathfrak{g}(\sigma, \beta)\mathfrak{q}(\sigma, \beta) \tag{6.36}$$

just as in (6.12). The point of (6.36) is that

$$\mathfrak{g} \mid \prod_{\sigma \in S} \mathfrak{g}(\sigma, \beta) \prod_{\substack{\sigma, \tau \in S \\ \sigma \neq \tau}} \gcd(\mathfrak{c}(\sigma, \beta), \mathfrak{c}(\tau, \beta))$$

and therefore

$$g_0 \mid \prod_{\sigma \in S} g_0(\sigma, \beta) \prod_{\substack{\sigma, \tau \in S \\ \sigma \neq \tau}} \gcd(\mathfrak{c}(\sigma, \beta), \mathfrak{c}(\tau, \beta)).$$

We use Lemma 6.3.1 to discard all β satisfying $g_0(\sigma, \beta) > Y^{\frac{1}{|S|^2}}$. Similarly, we use Lemma 6.3.2 to discard all β satisfying $\gcd(\mathfrak{c}(\sigma, \beta), \mathfrak{c}(\tau, \beta)) > Y^{\frac{1}{|S|^2}}$. Then the remaining β satisfy $g_0 \leq Y$. Furthermore, we have removed

$$\ll_\epsilon x^{\frac{n-1}{n}+\epsilon} Y^{-\frac{1}{18|S|^2}} + x^{\frac{n-2}{n}} + Y^{\frac{2n-4}{3|S|^2}} + x^{\frac{n-1}{n}+\epsilon} Y^{-\frac{1}{3|S|^2}}$$

β in total and hence

$$A_2(x;\rho) \ll_\epsilon x^{1+\epsilon} Y^{-\frac{1}{18|S|^2}} + x^{\frac{n-1}{n}} + x^{\frac{1}{n}} Y^{\frac{2n-4}{3|S|^2}} + x^{1+\epsilon} Y^{-\frac{1}{3|S|^2}}.$$

After picking $Y = x^{\frac{\delta}{2n}}$ we conclude that

$$A(x) \ll_\epsilon x^{1-\frac{\delta}{54n|S|^2}+\epsilon}.$$

We will now sketch how to modify this proof for totally complex K. We have to bound

$$A(x) = \sum_{\substack{N\mathfrak{a} \leq x \\ (\mathfrak{a}, F)=1, \mathfrak{m}|\mathfrak{a}}} r(\mathfrak{a}) \sum_{t \in T_K} \sum_{v \in V_K/V_K^2} \psi(tv\alpha \bmod F) \prod_{\sigma \in S} \text{spin}(\sigma, tv\alpha). \tag{6.37}$$

We use the fundamental domain constructed for totally complex fields form subsection 6.2.4 and we pick for each principal $\mathfrak{a}$ its generator in $\mathcal{D}$. Then equation (6.37) becomes

$$\begin{aligned} A(x) &= \sum_{t\in T_K} \sum_{v\in V_K/V_K^2} \sum_{\substack{\alpha\in\mathcal{D}, \mathrm{N}\alpha\leq x \\ \alpha\equiv\rho \bmod F \\ \alpha\equiv 0 \bmod \mathfrak{m}}} \psi(tv\alpha \bmod F) \prod_{\sigma\in S} \mathrm{spin}(\sigma, tv\alpha) \\ &= \sum_{t\in T_K} \sum_{v\in V_K/V_K^2} \sum_{\substack{\alpha\in tv\mathcal{D}, \mathrm{N}\alpha\leq x \\ \alpha\equiv\rho \bmod F \\ \alpha\equiv 0 \bmod \mathfrak{m}}} \psi(\alpha \bmod F) \prod_{\sigma\in S} \mathrm{spin}(\sigma, \alpha). \end{aligned}$$

We deal with each sum of the shape

$$\sum_{\substack{\alpha\in tv\mathcal{D}, \mathrm{N}\alpha\leq x \\ \alpha\equiv\rho \bmod F \\ \alpha\equiv 0 \bmod \mathfrak{m}}} \psi(\alpha \bmod F) \prod_{\sigma\in S} \mathrm{spin}(\sigma, \alpha) \tag{6.38}$$

exactly in the same way as for real quadratic fields K, where it is important to note that the shifted fundamental domain $tv\mathcal{D}$ still has the essential properties we need. Combining our estimate for each sum in equation (6.38), we obtain the desired upper bound for $A(x)$.

6.4 Bilinear sums

Let $x, y > 0$ and let $\{v_\mathfrak{a}\}_\mathfrak{a}$ and $\{w_\mathfrak{b}\}_\mathfrak{b}$ be two sequences of complex numbers bounded in modulus by 1. Define

$$B(x,y) = \sum_{\mathrm{N}(\mathfrak{a})\leq x} \sum_{\mathrm{N}(\mathfrak{b})\leq y} v_\mathfrak{a} w_\mathfrak{b} s_{\mathfrak{a}\mathfrak{b}}. \tag{6.39}$$

We wish to prove that for all $\epsilon > 0$, we have

$$B(x,y) \ll_\epsilon \left(x^{-\frac{1}{6n}} + y^{-\frac{1}{6n}}\right) (xy)^{1+\epsilon}, \tag{6.40}$$

where the implied constant is uniform in all choices of sequences $\{v_\mathfrak{a}\}_\mathfrak{a}$ and $\{w_\mathfrak{b}\}_\mathfrak{b}$ as above.

We split the sum $B(x,y)$ into h^2 sums according to which ideal classes $\mathfrak{a}$ and $\mathfrak{b}$ belong to. In fact, since $s_{\mathfrak{a}\mathfrak{b}}$ vanishes whenever $\mathfrak{a}\mathfrak{b}$ does not belong to the principal class, it suffices to split $B(x,y)$ into h sums

$$B(x,y) = \sum_{i=1}^{h} B_i(x,y), \quad B_i(x,y) = \sum_{\substack{\mathrm{N}(\mathfrak{a})\leq x \\ \mathfrak{a}\in C_i}} \sum_{\substack{\mathrm{N}(\mathfrak{b})\leq y \\ \mathfrak{b}\in C_i^{-1}}} v_\mathfrak{a} w_\mathfrak{b} s_{\mathfrak{a}\mathfrak{b}}.$$

We will prove the desired estimate for each of the sums $B_i(x,y)$. So fix an index $i \in \{1, \ldots, h\}$, let $\mathfrak{A} \in \mathcal{C}\ell_a$ be the ideal belonging to the ideal class C_i^{-1}, and let

$\mathfrak{B} \in \mathcal{C}\ell_b$ be the ideal belonging to the ideal class C_i. The conditions on $\mathfrak{a}$ and $\mathfrak{b}$ above mean that

$$\mathfrak{a}\mathfrak{A} = (\alpha), \quad \alpha \succ 0$$

and

$$\mathfrak{b}\mathfrak{B} = (\beta), \quad \beta \succ 0.$$

Since $\mathfrak{A} \in C_i^{-1}$ and $\mathfrak{B} \in C_i$, there exists an element $\gamma \in \mathcal{O}_K$ such that

$$\mathfrak{A}\mathfrak{B} = (\gamma), \quad \gamma \succ 0.$$

We are now in a position to use the factorization formula for $\mathrm{spin}(\mathfrak{a}\mathfrak{b})$ appearing in [24, (3.8), p. 708], which in turn leads to a factorization formula for $s_{\mathfrak{a}\mathfrak{b}}$. We note that the formula [24, (3.8), p. 708] also holds in case K is totally complex, with exactly the same proof. We have

$$\mathrm{spin}(\sigma, \alpha\beta/\gamma) = \mathrm{spin}(\sigma, \gamma)\delta(\sigma; \alpha, \beta)\left(\frac{\alpha\gamma}{\sigma(\mathfrak{a}\mathfrak{B})}\right)\left(\frac{\beta\gamma}{\sigma(\mathfrak{b}\mathfrak{A})}\right)\left(\frac{\alpha}{\sigma(\beta)\sigma^{-1}(\beta)}\right), \tag{6.41}$$

where $\delta(\sigma; \alpha, \beta) \in \{\pm 1\}$ is a factor which comes from an application of quadratic reciprocity and which depends only on σ and the congruence classes of α and β modulo 8.

If K is real quadratic, then we set

$$v'_{\mathfrak{a}} = v_{\mathfrak{a}} \prod_{\sigma \in S}\left(\frac{\alpha\gamma}{\sigma(\mathfrak{a}\mathfrak{B})}\right), \quad w'_{\mathfrak{b}} = w_{\mathfrak{b}} \prod_{\sigma \in S}\left(\frac{\beta\gamma}{\sigma(\mathfrak{b}\mathfrak{A})}\right),$$

and

$$\delta(\alpha, \beta) = \psi(\alpha\beta \bmod F) \prod_{\sigma \in S} \delta(\sigma; \alpha, \beta), \quad s(\gamma) = \prod_{\sigma \in S} \mathrm{spin}(\sigma, \gamma),$$

so that we can rewrite the sum $B_i(x, y)$ as

$$B_i(x, y) = s(\gamma) \sum_{\substack{\alpha \in \mathcal{D} \\ \mathrm{N}(\alpha) \leq x\mathrm{N}(\mathfrak{A}) \\ \alpha \equiv 0 \bmod \mathfrak{A}}} \sum_{\substack{\beta \in \mathcal{D} \\ \mathrm{N}(\beta) \leq y\mathrm{N}(\mathfrak{B}) \\ \beta \equiv 0 \bmod \mathfrak{B}}} \delta(\alpha, \beta) v'_{(\alpha)/\mathfrak{A}} w'_{(\beta)/\mathfrak{B}} \prod_{\sigma \in S}\left(\frac{\alpha}{\sigma(\beta)\sigma^{-1}(\beta)}\right). \tag{6.42}$$

Now set

$$v_{\alpha} = \mathbf{1}(\alpha \equiv 0 \bmod \mathfrak{A}) \cdot v'_{(\alpha)/\mathfrak{A}}$$

and

$$w_{\beta} = \mathbf{1}(\beta \equiv 0 \bmod \mathfrak{B}) \cdot w'_{(\beta)/\mathfrak{B}},$$

where $\mathbf{1}(P)$ is the indicator function of a property P. Also, for $\alpha, \beta \in \mathcal{O}_K$ with β odd, we define

$$\phi(\alpha, \beta) = \prod_{\sigma \in S}\left(\frac{\alpha}{\sigma(\beta)\sigma^{-1}(\beta)}\right).$$

Finally, we further split $B_i(x,y)$ according to the congruence classes of α and β modulo F, so as to control the factor $\delta(\alpha,\beta)$, which now depends on congruence classes of α and β modulo F due to the presence of $\psi(\alpha\beta \bmod F)$. We have

$$B_i(x,y) = s(\gamma) \sum_{\alpha_0\in(\mathcal{O}_K/(F))^\times} \sum_{\beta_0\in(\mathcal{O}_K/(F))^\times} \delta(\alpha_0,\beta_0)B_i(x,y;\alpha_0,\beta_0),$$

where

$$B_i(x,y;\alpha_0,\beta_0) = \sum_{\substack{\alpha\in\mathcal{D}(x\mathrm{N}(\mathfrak{A}))\\ \alpha\equiv\alpha_0 \bmod F}} \sum_{\substack{\beta\in\mathcal{D}(y\mathrm{N}(\mathfrak{B}))\\ \beta\equiv\beta_0 \bmod F}} v_\alpha w_\beta \phi(\alpha,\beta).$$

To prove the bound (6.40), at least in the case that K is totally real, it now suffices to prove, for each $\epsilon > 0$, the bound

$$B_i(x,y;\alpha_0,\beta_0) \ll_\epsilon \left(x^{-\frac{1}{6n}} + y^{-\frac{1}{6n}}\right)(xy)^{1+\epsilon}, \tag{6.43}$$

where the implied constant is uniform in all choices of uniformly bounded sequences of complex numbers $\{v_\alpha\}_\alpha$ and $\{w_\beta\}_\beta$ indexed by elements of $\mathcal{O}_K$. Each of the sums $B_i(x,y;\alpha_0,\beta_0)$ is of the same shape as $B(M,N;\omega,\zeta)$ in Chapter 4; in the notation of Chapter 4, $\mathfrak{f} = (F)$, α_w corresponds to v_α, β_z corresponds to w_β, and $\gamma(w,z)$ corresponds to $\phi(\alpha,\beta)$ (unfortunately with the arguments α and β flipped). Our desired estimate for $B_i(x,y;\alpha_0,\beta_0)$, and hence also $B(x,y)$, would now follow from Proposition 4.3.6, provided that we can verify properties (P1)-(P3) for the function $\phi(\alpha,\beta)$.

We now verify (P1)-(P3), thereby proving the bound (6.43) and hence also the bound (6.40). Property (P1) follows from the law of quadratic reciprocity, since for odd α and β we have

$$\begin{aligned}
\phi(\alpha,\beta) &= \prod_{\sigma\in S}\left(\frac{\alpha}{\sigma(\beta)}\right)\left(\frac{\alpha}{\sigma^{-1}(\beta)}\right)\\
&= \prod_{\sigma\in S}\mu(\sigma;\alpha,\beta)\left(\frac{\sigma(\beta)}{\alpha}\right)\left(\frac{\sigma^{-1}(\beta)}{\alpha}\right)\\
&= \left(\prod_{\sigma\in S}\mu(\sigma;\alpha,\beta)\right)\cdot\prod_{\sigma\in S}\left(\frac{\beta}{\sigma^{-1}(\alpha)}\right)\left(\frac{\beta}{\sigma(\alpha)}\right)\\
&= \left(\prod_{\sigma\in S}\mu(\sigma;\alpha,\beta)\right)\cdot\phi(\beta,\alpha),
\end{aligned}$$

where $\mu(\sigma;\alpha,\beta)$ depends only on σ and the congruence classes of α and β modulo 8. Property (P2) follows immediately from the multiplicativity of each argument of the quadratic residue symbol $(\cdot/\cdot)$. Finally, for property (P3), since $\sigma^{-1}\notin S$ whenever $\sigma\in S$, we see that

$$\varphi(\beta) = \prod_{\sigma\in S}\sigma(\beta)\sigma^{-1}(\beta)$$

divides $\mathrm{N}(\beta) = \prod_{\sigma \in \mathrm{Gal}(K/\mathbb{Q})} \sigma(\beta)$; thus, the first part of (P3) indeed holds true. It now suffices to prove that

$$\sum_{\xi \bmod \mathrm{N}(\beta)} \left(\frac{\xi}{\varphi(\beta)}\right)$$

vanishes if $|\mathrm{N}(\beta)|$ is not squarefull. The sum above is a multiple of the sum

$$\sum_{\xi \bmod \varphi(\beta)} \left(\frac{\xi}{\varphi(\beta)}\right),$$

which vanishes if the principal ideal generated by $\varphi(\beta)$ is not the square of an ideal. The proof now proceeds as in [24, Lemma 3.1]. Supposing $|\mathrm{N}(\beta)|$ is not squarefull, we take a rational prime p such that $p \mid \mathrm{N}(\beta)$ but $p^2 \nmid \mathrm{N}(\beta)$. This implies that there is a degree-one prime ideal divisor $\mathfrak{p}$ of β such that $(\beta) = \mathfrak{p}\mathfrak{c}$ with $\mathfrak{c}$ coprime to p, i.e., coprime to all the conjugates of $\mathfrak{p}$. Hence $\varphi(\beta)$ factors as

$$(\varphi(\beta)) = \prod_{\sigma \in S} \sigma(\mathfrak{p})\sigma^{-1}(\mathfrak{p}) \prod_{\sigma \in S} \sigma(\mathfrak{c})\sigma^{-1}(\mathfrak{c}),$$

where the evidently non-square $\prod_{\sigma \in S} \sigma(\mathfrak{p})\sigma^{-1}(\mathfrak{p})$ is coprime to $\prod_{\sigma \in S} \sigma(\mathfrak{c})\sigma^{-1}(\mathfrak{c})$, hence proving that $(\varphi(\beta))$ is not a square. This proves that property (P3) holds true, and then Proposition 4.3.6 implies the estimate (6.43) and hence also (6.40), at least in the case that K is totally real.

If K is totally complex, fix $t \in T_K$ and $v \in V_K/V_K^2$. Then replacing α by $tv\alpha$ in (6.41), we get

$$\begin{aligned}\mathrm{spin}(\sigma, tv\alpha\beta/\gamma) = \mathrm{spin}(\sigma, \gamma)\delta(\sigma; tv\alpha, \beta)& \\ \left(\frac{tv\alpha\gamma}{\sigma(\mathfrak{a}\mathfrak{B})}\right)\left(\frac{\beta\gamma}{\sigma(\mathfrak{b}\mathfrak{A})}\right)&\left(\frac{tv}{\sigma(\beta)\sigma^{-1}(\beta)}\right)\left(\frac{\alpha}{\sigma(\beta)\sigma^{-1}(\beta)}\right),\end{aligned}$$

where now $\delta(\sigma; \alpha, \beta; t, v) = \delta(\sigma; tv\alpha, \beta)\left(\frac{tv}{\sigma(\beta)\sigma^{-1}(\beta)}\right) \in \{\pm 1\}$ depends only on σ, t, v, and the congruence classes of α and β modulo 8. Then instead of (6.42), we have

$$\begin{aligned}B_i(x, y) = s(\gamma) \sum_{t \in T_K} \sum_{v \in V_K/V_K^2} \sum_{\substack{\alpha \in \mathcal{D} \\ \mathrm{N}(\alpha) \leq x\mathrm{N}(\mathfrak{A}) \\ \alpha \equiv 0 \bmod \mathfrak{A}}} \sum_{\substack{\beta \in \mathcal{D} \\ \mathrm{N}(\beta) \leq y\mathrm{N}(\mathfrak{B}) \\ \beta \equiv 0 \bmod \mathfrak{B}}} \delta(\alpha, \beta; t, v)& \\ v(t, v)'_{(\alpha)/\mathfrak{A}} w'_{(\beta)/\mathfrak{B}} \prod_{\sigma \in S} \left(\frac{\alpha}{\sigma(\beta)\sigma^{-1}(\beta)}\right),& \quad (6.44)\end{aligned}$$

where now

$$v(t, v)'_{\mathfrak{a}} = v_{\mathfrak{a}} \prod_{\sigma \in S} \left(\frac{tv\alpha\gamma}{\sigma(\mathfrak{a}\mathfrak{B})}\right), \quad w'_{\mathfrak{b}} = w_{\mathfrak{b}} \prod_{\sigma \in S} \left(\frac{\beta\gamma}{\sigma(\mathfrak{b}\mathfrak{A})}\right),$$

and

$$\delta(\alpha, \beta; t, v) = \psi(tv\alpha\beta \bmod F) \prod_{\sigma \in S} \delta(\sigma; \alpha, \beta; t, v), \quad s(\gamma) = \prod_{\sigma \in S} \mathrm{spin}(\sigma, \gamma).$$

The rest of the proof now proceeds identically to the case when K is totally real.

6.5 Governing fields

Let $E = \mathbb{Q}(\zeta_8, \sqrt{1+i})$ and let $h(-4p)$ be the class number of $\mathbb{Q}(\sqrt{-4p})$. It is well-known that E is a governing field for the 8-rank of $\mathbb{Q}(\sqrt{-4p})$; in fact 8 divides $h(-4p)$ if and only if p splits completely in E. We assume that K is a hypothetical governing field for the 16-rank of $\mathbb{Q}(\sqrt{-4p})$ and derive a contradiction. If K' is a normal field extension of $\mathbb{Q}$ containing K, then K' is also a governing field. Therefore we can reduce to the case that K contains E. In particular, K is totally complex.

We have $\mathrm{Gal}(E/\mathbb{Q}) \cong D_4$ and we fix an element of order 4 in $\mathrm{Gal}(E/\mathbb{Q})$ that we call r. Let p be a rational prime that splits completely in E. Since E is a PID, we can take π to be a prime in $\mathcal{O}_E$ above p. It follows from Proposition 6.2 of [41], which is based on earlier work of Bruin and Hemenway [7], that there exists an integer F and a function $\psi_0 : (\mathcal{O}_E/F\mathcal{O}_E)^\times \to \mathbb{C}$ such that for all p with $(p, F) = 1$ we have

$$16 \mid h(-4p) \Leftrightarrow \psi_0(\pi \bmod F) \left(\frac{r(\pi)}{\pi}\right)_{E,2} = 1, \tag{6.45}$$

where $\psi_0(\alpha \bmod F) = \psi_0(\alpha u^2 \bmod F)$ for all $\alpha \in \mathcal{O}_K$ coprime to F and all $u \in \mathcal{O}_K^\times$. We take S equal to the inverse image of our fixed automorphism r under the natural surjective map $\mathrm{Gal}(K/\mathbb{Q}) \to \mathrm{Gal}(E/\mathbb{Q})$. Then it is easily seen that $\sigma \in S$ implies $\sigma^{-1} \notin S$. If $\mathfrak{p}$ is a principal prime of K with generator w of norm p, we have

$$\begin{aligned}\prod_{\sigma\in S} \mathrm{spin}(\sigma, w) &= \prod_{\sigma\in S} \left(\frac{w}{\sigma(w)}\right)_{K,2} = \left(\frac{w}{r(\mathrm{N}_{K/E}(w))}\right)_{K,2} \\ &= \psi_1(w \bmod 8) \left(\frac{r(\mathrm{N}_{K/E}(w))}{w}\right)_{K,2} = \psi_1(w \bmod 8) \left(\frac{r(\mathrm{N}_{K/E}(w))}{\mathrm{N}_{K/E}(w)}\right)_{E,2}.\end{aligned}$$

We are now going to apply Theorem 6.1.1 to the number field K, the function

$$\psi(w \bmod F) := \psi_1(w \bmod 8)\psi_0\left(\mathrm{N}_{K/E}(w) \bmod F\right).$$

and S as defined above. Then for a principal prime $\mathfrak{p}$ of K with generator w and norm p

$$\begin{aligned}s_{\mathfrak{p}} &= \sum_{t\in T_K} \sum_{v\in V_K/V_K^2} \psi\left(tvw \bmod F\right) \prod_{\sigma\in S} \mathrm{spin}(\sigma, tvw) \\ &= 2|T_K||V_K/V_K^2| \left(\mathbf{1}_{16|h(-p)} - \frac{1}{2}\right),\end{aligned} \tag{6.46}$$

since the equivalence in (6.45) does not depend on the choice of π. Theorem 6.1.1 shows oscillation of the sum

$$\sum_{\substack{\mathrm{N}(\mathfrak{p})\leq X \\ \mathfrak{p}\text{ principal}}} s_{\mathfrak{p}}.$$

The dominant contribution of this sum comes from prime ideals of degree 1 and for these primes equation (6.46) is valid. But if K were to be a governing field, $s_{\mathfrak{p}}$ has to be constant on unramified prime ideals of degree 1, which is the desired contradiction.

Chapter 7

Vinogradov's three primes theorem with primes having given primitive roots

Joint work with Christopher Frei and Efthymios Sofos

Abstract

The first purpose of this chapter is to show how Hooley's celebrated method leading to his conditional proof of the Artin conjecture on primitive roots can be combined with the Hardy–Littlewood circle method. We do so by studying the number of representations of an odd integer as a sum of three primes, all of which have prescribed primitive roots. The second purpose is to analyse the singular series. In particular, using results of Lenstra, Stevenhagen and Moree, we provide a partial factorisation as an Euler product and prove that this does not extend to a complete factorisation.

7.1 Introduction

Can we represent an odd integer as a sum of 3 odd primes all of which have 27 as a primitive root? Lenstra [51] was the first to address the problem of primes with a fixed primitive root and lying in an arithmetic progression. One of his results [51, Th.(8.3)] states that if $b \neq 5 \pmod{12}$ then either there are no primes $p \equiv b \pmod{12}$ having 27 as a primitive root or there is exactly one such prime, namely $p = 2$. Hence, unless $n \equiv 3 \pmod{12}$, no such representation exists.

In this chapter, we are interested in the converse direction, at least for all sufficiently large values of n. The existence of infinitely many primes with a given primitive root a is currently not known unconditionally for any $a \in \mathbb{Z}$, so we need to be content with working under the assumption of a certain generalised Riemann Hypothesis, sometimes called *Hooley's Riemann Hypothesis.* For any non-zero integer a, we will write HRH(a) for the hypothesis that

> for all square-free $k \in \mathbb{N}$, the Dedekind zeta function of the number field $\mathbb{Q}(\zeta_k, \sqrt[k]{a})$, where $\zeta_k \in \mathbb{C}$ is a primitive k-th root of unity, satisfies the Riemann hypothesis.

Our main theorem can be seen as a combination of the classical conditional result of Hardy and Littlewood [30] towards ternary Goldbach with Hooley's [35] conditional proof of Artin's conjecture.

Theorem 7.1.1. *Let* $\mathbf{a} = (a_1, a_2, a_3) \in \mathbb{Z}^3$ *such that no* a_i *is* -1 *or a perfect square. Assuming* HRH(a_i) *for* $i = 1, 2, 3$, *we have*

$$\sum_{\substack{p_1+p_2+p_3=n \\ \forall i:\ \mathbb{F}_{p_i}^* = \langle a_i \rangle}} \prod_{i=1}^{3} \log p_i = \mathcal{A}_{\mathbf{a}}(n) n^2 + o(n^2), \qquad \textit{as} \qquad n \to +\infty, \tag{7.1}$$

with an explicit factor $\mathcal{A}_{\mathbf{a}}(n) \in \mathbb{R}_{\geq 0}$ *that satisfies* $\mathcal{A}_{\mathbf{a}}(n) \gg_{\mathbf{a}} 1$ *whenever* $\mathcal{A}_{\mathbf{a}}(n) > 0$.

The bulk of this chapter will be devoted to the description and investigation of the factor $\mathcal{A}_{\mathbf{a}}(n)$. In particular, a product decomposition of $\mathcal{A}_{\mathbf{a}}(n)$ will allow us to interpret Theorem 7.1.1 as a local-global principle and gives the following as a simple consequence.

Corollary 7.1.2. *Assume* HRH(27). *Let* n *be a sufficiently large odd integer. Then there are odd primes* p_1, p_2, p_3 *with* 27 *as a primitive root and* $n = p_1 + p_2 + p_3$ *if and only if* $n \equiv 3 \bmod 12$.

We can also get an explicit saving in the error term, for the price of working under a stronger generalised Riemann hypothesis. Let HRH'(a) be the hypothesis that

> for each square-free $k > 0$ all Hecke L-functions of the number field $\mathbb{Q}(\zeta_k, \sqrt[k]{a})$ satisfy the Riemann hypothesis.

Theorem 7.1.3. *Let* a_1, a_2, a_3 *be three integers none of which is* -1 *or a perfect square. Assuming* HRH'(a_i) *for* $i = 1, 2, 3$, *we have for* $\beta \in (0, 1)$,

$$\sum_{\substack{p_1+p_2+p_3=n \\ \forall i:\ \mathbb{F}_{p_i}^* = \langle a_i \rangle}} \prod_{i=1}^{3} \log p_i = \mathcal{A}_{\mathbf{a}}(n) n^2 + O_{\mathbf{a},\beta}(n^2 (\log n)^{-\beta}), \tag{7.2}$$

where the implied constant is effective and depends at most on a_1, a_2, a_3 *and* β.

Before returning to the explicit description of our factor $\mathcal{A}_{\mathbf{a}}(n)$, let us briefly review the relevant literature on Artin's conjecture and the ternary Goldbach problem, and introduce some necessary notation along the way.

7.1.1 Artin's conjecture

Fix an integer $a \neq -1$ which is not a perfect square. A question going back to Gauss regards the infinitude of primes having a as a primitive root. It was realised by Artin that the question admits an interpretation through algebraic number theory. Denote by ζ_k a primitive kth root of unity and define for any positive square-free integer k the number field

$$G_{a,k} := \mathbb{Q}(a^{1/k}, \zeta_k). \tag{7.3}$$

Artin's criterion states that the prime p has a as a primitive root if and only if for every prime q the rational prime p does not split completely in $G_{a,q}$. This led to the formulation of the following conjecture via a collective effort due to Artin, Lehmer and Heilbronn. Define

$$\Delta_a := \mathrm{Disc}(\mathbb{Q}(\sqrt{a})), \text{ the discriminant of } \mathbb{Q}(\sqrt{a}) \tag{7.4}$$

$$h_a := \max\{m \in \mathbb{N} : a \text{ is an } m\text{th power}\}, \tag{7.5}$$

$$\mathcal{A}_a := \prod_{p \mid h_a}\left(1 - \frac{1}{p-1}\right)\prod_{p \nmid h_a}\left(1 - \frac{1}{p(p-1)}\right) \tag{7.6}$$

and for positive integers q let

$$f_a^{\ddagger}(q) := \Big(\prod_{p \mid q, p \mid h_a}(p-2)^{-1}\Big)\Big(\prod_{p \mid q, p \nmid h_a}(p^2-p-1)^{-1}\Big). \tag{7.7}$$

Here, and throughout the chapter, the letter p is reserved for rational primes. We furthermore define

$$\mathcal{L}_a := \mathcal{A}_a \cdot \big(1 + \mu(2|\Delta_a|)f_a^{\ddagger}(|\Delta_a|)\big), \tag{7.8}$$

where μ is the Möbius function. Artin's conjecture then states that

$$\lim_{x \to +\infty} \frac{\#\{p \leq x : \mathbb{F}_p^* = \langle a \rangle\}}{\#\{p \leq x\}} = \mathcal{L}_a. \tag{7.9}$$

This conjecture is of substantial difficulty: there is no value of a for which the limit is known to be positive. In fact, it is not even known whether for every integer a that is not a square or -1 there exists a prime having primitive root a.

A significant step in the subject has been the, conditional under HRH(a), resolution of Artin's conjecture by Hooley [35]. His method is pivotal in the present work. Notable progress was later made by Heath-Brown [32], who building on work of Gupta and Murty [28], showed unconditionally that at least $\gg x/(\log x)^2$ primes $p \leq x$ have primitive root q, r or s, where $\{q, r, s\}$ is any set of non-zero integers which is multiplicative independent and such that none of $q, r, s, -3qr, -3qs, -3rs$ or qrs is a square. There is a rather extensive list of further results, as well as certain cryptographic applications; the reader is referred to the comprehensive survey of Moree [60]. Lenstra [51] used Hooley's method to show, conditionally on HRH(a), the existence of the Dirichlet density of primes in an arithmetic progression and with a as primitive root. An explicit formula for these densities was given later by Moree [59]. To describe Moree's result we need the following notation. Let

$$\beta_a(q) := \begin{cases} (-1)^{\frac{\frac{\Delta_a}{\gcd(q,\Delta_a)}-1}{2}} \gcd(q,\Delta_a), & \text{if } \frac{\Delta_a}{\gcd(q,\Delta_a)} \equiv 1 \,(\mathrm{mod}\ 2) \\ 1 & \text{otherwise,} \end{cases} \tag{7.10}$$

and observe that $\beta_a(q)$ is a fundamental discriminant in case $\Delta_a/\gcd(q,\Delta_a) \equiv 1 \bmod 2$. For positive integers q let

$$f_a^\dagger(q) := \prod_{p|h_a, p|q} \left(1-\frac{1}{p-1}\right)^{-1} \prod_{p\nmid h_a, p|q} \left(1-\frac{1}{p(p-1)}\right)^{-1}. \tag{7.11}$$

Definition 7.1.4. Assume that $a \neq -1$ is a non-square integer, let Δ_a, h_a be as in (7.4), (7.5) and assume that x, q are integers with $q > 0$. We define

$$\mathcal{A}_a(x \bmod q) := \mathcal{A}_a \cdot \begin{cases} \frac{f_a^\dagger(q)}{\phi(q)} \prod_{p|x-1, p|q} \left(1-\frac{1}{p}\right), & \text{if } \gcd(x-1,q,h_a) = \gcd(x,q) = 1, \\ 0, & \text{otherwise,} \end{cases} \tag{7.12}$$

and

$$\delta_a(x \bmod q) := \mathcal{A}_a(x \bmod q)\left(1+\mu\left(\frac{2|\Delta_a|}{\gcd(q,\Delta_a)}\right)\left(\frac{\beta_a(q)}{x}\right) f_a^\ddagger\left(\frac{|\Delta_a|}{\gcd(q,\Delta_a)}\right)\right),$$

where $\phi(\cdot)$ is the Euler totient function and $\left(\frac{\cdot}{\cdot}\right)$ is the Kronecker quadratic symbol.

Moree's result [59] states that, conditionally under HRH(a), the Dirichlet density of primes in an arithmetic progression and with a as primitive root equals $\delta_a(x \bmod q)$. His work will prove of central importance in our interpretation of the Artin factor for the ternary Diophantine problem under study.

7.1.2 Ternary Goldbach problem

The ternary Goldbach problem has been one of the most central subjects in analytic number theory; it asserts that every odd integer greater than 5 is the sum of 3 primes. Hardy and Littlewood [30] used the circle method to provide the first serious approach to the problem; they proved an asymptotic formula for the number of representations of n as a sum of k primes ($k \geq 3$) conditionally on the veracity of the generalised Riemann hypothesis. This hypothesis was removed later by Vinogradov [74]. His result states that for every $\beta > 0$ one has for all odd integers n that

$$\sum_{p_1+p_2+p_3=n} \prod_{i=1}^{3} \log p_i = \frac{1}{2}\left(\prod_p \varrho_p(n)\right) n^2 + O_\beta(n^2(\log n)^{-\beta}),$$

where the product is over all primes, the implied constant depends at most on β, and

$$\varrho_p(n) := p\Bigg(\sum_{\substack{b_1,b_2,b_3\in(\mathbb{Z}/p\mathbb{Z})^* \\ b_1+b_2+b_3\equiv n \pmod p}} \frac{1}{(p-1)^3}\Bigg). \tag{7.13}$$

This can be thought as the ratio of the probability that a random vector $\mathbf{b}\in((\mathbb{Z}/p\mathbb{Z})^*)^3$ satisfies $\sum_{1\le i\le 3} b_i \equiv n \pmod p$ to the probability that a random vector $\mathbf{b}\in(\mathbb{Z}/p\mathbb{Z})^3$ satisfies $\sum_{1\le i\le 3} b_i \equiv n \pmod p$, as made clear from

$$p = \Bigg(\sum_{\substack{b_1,b_2,b_3 \pmod p \\ b_1+b_2+b_3\equiv n \pmod p}} \frac{1}{p^3}\Bigg)^{-1}. \tag{7.14}$$

It should be mentioned that Helfgott [34] recently settled the ternary Goldbach problem. Using recent developments in additive combinatorics, Shao [65] provided general conditions for an infinite subset $\mathcal{P}$ of the primes that allow solving $n = p_1 + p_2 + p_3$ for large odd n with each p_i in $\mathcal{P}$. The result most related to our work is [65, Th.1.3]; it states that if there exists $\delta > 0$ such that the intersection of $\mathcal{P}$ with each invertible residue class modulo every integer q has density at least $\delta/\phi(q)$, then, under suitable additional assumptions, $n = p_1 + p_2 + p_3$ is soluble within $\mathcal{P}$. This does not cover our situation, since if $h_a > 1$ then the densities $\delta_a(1 \bmod h_a)$ vanish. Furthermore, if $h_a = 1$ then these densities could become arbitrarily close to zero. Indeed, if q is of the form $\prod_{p\le T} p$ for some $T > 2$ then it is easy to see that

$$\delta_a(1 \bmod q)\phi(q) \le \prod_{p\le T}\Big(1-\frac{1}{p}\Big) \ll \frac{1}{\log\log q}.$$

It would be interesting to modify his approach in order to recover some of our results, for example a lower bound of the correct order of magnitude as the one provided by Theorem 7.1.1. This approach would still require HRH(a_i) and besides the focal point of the chapter is the 'Artin factor' $\mathcal{A}_{\mathbf{a}}(n)$ in Theorem 7.1.1. A further result related to ours is that of Kane [38]. A very special case of his work provides an asymptotic for the number of solutions of $n = p_1 + p_2 + p_3$ when each p_i lies in a prefixed Chebotarev class of a Galois extension of $\mathbb{Q}$. Primes with a prescribed primitive root do admit a Chebotarev description, however the number of conditions involved is not fixed.

7.1.3 The factor $\mathcal{A}_{\mathbf{a}}(n)$

Let us now describe the representation of $\mathcal{A}_{\mathbf{a}}(n)$ that is obtained directly from the proof of Theorem 7.1.1. Define for $q > 0$ and square-free $k > 0$ the number field $F_{a,q,k} := \mathbb{Q}(\zeta_q, \zeta_k, a^{1/k})$, so that $G_{a,k} = F_{a,k,k}$. Moreover, for $b \in \mathbb{Z}$ with $\gcd(b, q) = 1$, we let $c_{a,q,k}(b) := 1$ if the restriction of the automorphism $\sigma_b : \zeta_q \mapsto \zeta_q^b$ of $\mathbb{Q}(\zeta_q)$ to

$\mathbb{Q}(\zeta_q) \cap G_{a,k}$ is the identity and we otherwise let $c_{a,q,k}(b) := 0$. We use the usual notation $\mathrm{e}_q(z) := \exp(2\pi i z/q)$, for $z \in \mathbb{C}, q \in \mathbb{N}$. The exponential sum

$$S_{a,q,k}(z) := \sum_{b \in (\mathbb{Z}/q\mathbb{Z})^*} c_{a,q,k}(b)\mathrm{e}_q(zb) \tag{7.15}$$

and the entities

$$L_{\mathbf{a},q,\mathbf{k}}(z) := \prod_{i=1}^{3} S_{a_i,q,k_i}(z), \tag{7.16}$$

$$d_{\mathbf{a},\mathbf{k}}(q) := \prod_{i=1}^{3} [F_{a_i,q,k_i} : \mathbb{Q}] \tag{7.17}$$

will play a central role throughout this chapter. For positive square-free k_1, k_2, k_3 we define

$$\mathfrak{S}_{\mathbf{a},\mathbf{k}}(n) := \sum_{q=1}^{\infty} \frac{1}{d_{\mathbf{a},\mathbf{k}}(q)} \sum_{\substack{z \in \mathbb{Z}/q\mathbb{Z} \\ \gcd(z,q)=1}} \mathrm{e}_q(-nz) L_{\mathbf{a},q,\mathbf{k}}(z). \tag{7.18}$$

It will be made clear in §7.2 that this is the *singular series* for the representation problem $n = p_1 + p_2 + p_3$ where for each i the prime p_i splits completely in G_{a_i,k_i}. The absolute convergence of the sum over q will be verified in Lemma 7.3.2. With this notation in place, the leading factor in Theorem 7.1.1 and Theorem 7.1.3 is given by

$$\mathcal{A}_{\mathbf{a}}(n) = \frac{1}{2}\Bigg(\sum_{\mathbf{k} \in \mathbb{N}^3} \mu(k_1)\mu(k_2)\mu(k_3) \mathfrak{S}_{\mathbf{a},\mathbf{k}}(n) \Bigg). \tag{7.19}$$

The sum over $\mathbf{k}$ will be shown to be absolutely convergent in Lemma 7.3.2. It is desirable to describe the integers n for which $\mathcal{A}_{\mathbf{a}}(n) \neq 0$. An important remark is that if the method of Hooley works in an Artin conjecture-related problem then it provides a leading constant which is an infinite alternating sum of Euler products that is not obviously equal to the conjectured Artin constant. Such a phenomenon is well documented and can be observed for instance in the work of Lenstra [51], who studied the density of primes in arithmetic progressions and with a prescribed primitive root, as well as the work of Serre [64], who studied the density of primes p for which the reduction of an elliptic curve over $\mathbb{F}_p$ is cyclic. Artin constants have not been studied in the context of Diophantine problems prior to the present work, however, we will show that $\mathcal{A}_{\mathbf{a}}(n)$ factorises partially and we shall provide an interpretation for $\mathcal{A}_{\mathbf{a}}(n)$. For every positive integer d we define the densities

$$\sigma_{\mathbf{a},n}(d) := d \Bigg(\sum_{\substack{b_1,b_2,b_3 (\bmod\ d) \\ b_1+b_2+b_3 \equiv n (\bmod\ d)}} \prod_{i=1}^{3} \frac{\delta_{a_i}(b_i \bmod d)}{\mathcal{L}_{a_i}} \Bigg). \tag{7.20}$$

The factor d has an explanation that is identical to the explanation of the factor p in (7.13)-(7.14). Let $[\cdot]$ denote the least common multiple, $\nu_p(\cdot)$ be the p-adic valuation and define

$$\mathfrak{D}_{\mathbf{a}} := 2^{\min\{\nu_2(\Delta_{a_i}):1 \le i \le 3\} - \max\{\nu_2(\Delta_{a_i}):1 \le i \le 3\}} [\Delta_{a_1}, \Delta_{a_2}, \Delta_{a_3}]. \tag{7.21}$$

Theorem 7.1.5. *The factor $\mathcal{A}_{\mathbf{a}}(n)$ in Theorems 7.1.1 and 7.1.3 factorises as follows,*

$$\mathcal{A}_{\mathbf{a}}(n) = \frac{1}{2}\Big(\prod_{i=1}^{3} \mathcal{L}_{a_i}\Big)\sigma_{\mathbf{a},n}(\mathfrak{D}_{\mathbf{a}}) \prod_{p\nmid \mathfrak{D}_{\mathbf{a}}} \sigma_{\mathbf{a},n}(p). \tag{7.22}$$

Furthermore, whenever $\mathcal{A}_{\mathbf{a}}(n) > 0$, we have

$$\mathcal{A}_{\mathbf{a}}(n) \gg \prod_{i=1}^{3} \frac{\phi(h_{a_i})}{|\Delta_{a_i}|^2 h_{a_i}}, \tag{7.23}$$

with an absolute implied constant.

For an interpretation of the right side of (7.22) see §7.1.4. The proof of (7.22) (that will be provided in §7.4.1) requires adroit manoeuvring. This is because the densities $\delta_a(b_i \bmod d)$ in (7.20) have a complicated dependence on b_i and also do not exhibit good factorisation properties with respect to d.

Let us furthermore comment that in contrast to the usual applications of the circle method, the constant in (7.22) does not factorise as an Euler product, see §7.4.6 for a precise statement of this phenomenon. The following consequence of Theorem 7.1.1 and Theorem 7.1.5 can be interpreted as a local-global principle.

Corollary 7.1.6. *Let a_1, a_2, a_3 be three integers none of which is -1 or a perfect square, and assume $\mathrm{HRH}(a_i)$ for $i = 1, 2, 3$. For every sufficiently large odd integer n, the following statements are equivalent:*

1. *There are primes p_1, p_2, p_3 not dividing $6\Delta_{a_1}\Delta_{a_2}\Delta_{a_3}$ such that each a_i is a primitive root modulo p_i and $p_1 + p_2 + p_3 = n$.*
2. *For $d \in \{3, \mathfrak{D}_{\mathbf{a}}\}$, there are primes p_1, p_2, p_3 with $\gcd(p_1p_2p_3, 2d) = 1$ such that a_i is a primitive root for p_i for every $i = 1, 2, 3$ and $p_1 + p_2 + p_3 \equiv n \bmod d$.*

Though part *(2)* of Corollary 7.1.6 may not look like a purely local statement, it is one. In fact, for any d in $\mathbb{N}$, solubility of the congruence modulo d in primes not dividing $2d$ with prescribed primitive roots is equivalent to the statement that $\sigma_{\mathbf{a},d}(n) > 0$. In Lemma 7.4.7, we shall see that $\sigma_{\mathbf{a},n}(p) > 0$ whenever $p \nmid 3\Delta_{a_1}\Delta_{a_2}\Delta_{a_3}$. Moreover, it is clear from the definition in (7.20), that whether $\sigma_{\mathbf{a},d}(n) = 0$ or not is a local condition modulo d.

7.1.4 Interpretation of the Artin factor for the ternary Goldbach problem

Studying the constants in any counting problem of flavour similar to that of Artin's conjecture is a non-trivial task and has been analysed rather extensively. The problems involve primes with a fixed primitive root, primes in progressions and with a fixed primitive root and primes such that the reduction of a fixed elliptic curve over the corresponding finite field is cyclic, see the work of Serre [64]. The reader that is interested in an overview of the work that has been done on these constants so far is directed at the work of and Lenstra–Stevenhagen–Moree [52], as well as the survey of Moree [60].

We now focus on the interpretation of the "Artin-factor" $\mathcal{A}_{\mathbf{a}}(n)$ with the help of (7.22). First, the factor $1/2$ is related to the density of solutions in $\mathbb{R}$ of $\sum_{1\leq i\leq 3} x_i = n$ and it has the exact same interpretation as in the classical situation of ternary Goldbach, and therefore, we do not further comment on this.

The term

$$\mathcal{L}_{a_1}\mathcal{L}_{a_2}\mathcal{L}_{a_3}$$

in (7.22) should be thought of as the "probability" that for all $i = 1, 2, 3$, a random prime p_i has primitive root a_i, see (7.9).

The factors $\sigma_{\mathbf{a},n}(d)$ for $d \in \{\mathfrak{D}_{\mathbf{a}}\} \cup \{p \text{ prime } : p \nmid \mathfrak{D}_{\mathbf{a}}\}$ admit an explanation that is comparable to the analogous densities in the classical case of the ternary Goldbach problem, see (7.13). There is only one difference, namely that one has to use the weight

$$\frac{\delta_{a_i}(b_i \bmod d)}{\mathcal{L}_{a_i}}$$

instead of $1/(p-1)$. This new weight equals the *conditional* probability that a random prime lies in the arithmetic progression $b_i \pmod d$ given that it has primitive root a_i.

It would be desirable to use algebraic considerations (for example, the approach of 'entanglement' of splitting fields as in the work of Lenstra–Stevenhagen–Moree [52]), to provide a prediction for $\mathcal{A}_{\mathbf{a}}(n)$ with a method that is different to the one in §7.4.1.

7.1.5 The case where all primitive roots are equal

In our next theorem, we provide an explicit description of the local conditions in Corollary 7.1.6, but for space considerations we do so only in the important case where

$$a_1 = a_2 = a_3 =: a.$$

The first row of the following table contains the discriminant of $\mathbb{Q}(\sqrt{a})$ and the second row contains the power properties of a. For example, if a is a cube but not a fifth power we shall write $a \in \mathbb{Z}^3 \setminus \mathbb{Z}^5$.

Theorem 7.1.7. *Let $a \neq -1$ be a non-square integer and $n \in \mathbb{N}$. Then the 'Artin factor'*

$$\mathcal{A}_{(a,a,a)}(n)$$

is strictly positive if and only if n satisfies one of the congruence conditions in the third row of the following table. The second to last row refers to all integers a not considered *in a row above it, as long as a is a third power. The last row refers to every integer a not considered in a row above it.*

$\mathrm{Disc}(\mathbb{Q}(\sqrt{a}))$	*Power properties of a*	*Congruence conditions for n*
-3	$\mathbb{Z}\setminus(\{-1\}\cup\mathbb{Z}^2)$	$3\,(\mathrm{mod}\ 6)$
-4	$\mathbb{Z}\setminus(\{-1\}\cup\mathbb{Z}^2)$	$1\,(\mathrm{mod}\ 4)$
5	$\mathbb{Z}\setminus(\{-1\}\cup\mathbb{Z}^2)$	$1\,(\mathrm{mod}\ 2)$ *and not* $0\,(\mathrm{mod}\ 5)$
12	$\mathbb{Z}\setminus(\{-1\}\cup\mathbb{Z}^2\cup\mathbb{Z}^3)$	$3,5,7,9\,(\mathrm{mod}\ 12)$
12	$\mathbb{Z}^3\setminus(\{-1\}\cup\mathbb{Z}^2)$	$3\,(\mathrm{mod}\ 12)$
-15	$\mathbb{Z}\setminus(\{-1\}\cup\mathbb{Z}^2\cup\mathbb{Z}^3\cup\mathbb{Z}^5)$	$1\,(\mathrm{mod}\ 2)$ *and not* $0\,(\mathrm{mod}\ 15)$
-15	$\mathbb{Z}^3\setminus(\{-1\}\cup\mathbb{Z}^2\cup\mathbb{Z}^5)$	$1\,(\mathrm{mod}\ 2)$ *and* $3,6,9,12\,(\mathrm{mod}\ 15)$
-15	$\mathbb{Z}^5\setminus(\{-1\}\cup\mathbb{Z}^2\cup\mathbb{Z}^3)$	$1\,(\mathrm{mod}\ 2)$ *and not* $0,1,2,7,8,14\,(\mathrm{mod}\ 15)$
-15	$\mathbb{Z}^{15}\setminus(\{-1\}\cup\mathbb{Z}^2)$	$12\,(\mathrm{mod}\ 15)$
-20	$\mathbb{Z}^5\setminus(\{-1\}\cup\mathbb{Z}^2)$	$1\,(\mathrm{mod}\ 2)$ *and not* $1\,(\mathrm{mod}\ 20)$
21	$\mathbb{Z}^7\setminus(\{-1\}\cup\mathbb{Z}^2\cup\mathbb{Z}^3)$	$1\,(\mathrm{mod}\ 2)$ *and not* $8\,(\mathrm{mod}\ 21)$
21	$\mathbb{Z}^3\setminus(\{-1\}\cup\mathbb{Z}^2\cup\mathbb{Z}^7)$	$3\,(\mathrm{mod}\ 6)$
21	$\mathbb{Z}^{21}\setminus(\{-1\}\cup\mathbb{Z}^2)$	$1\,(\mathrm{mod}\ 2)$ *and* $3,6,12,15\,(\mathrm{mod}\ 21)$
± 24	$\mathbb{Z}^3\setminus(\{-1\}\cup\mathbb{Z}^2)$	$3\,(\mathrm{mod}\ 6)$
60	$\mathbb{Z}^3\setminus(\{-1\}\cup\mathbb{Z}^2)$	$3\,(\mathrm{mod}\ 6)$
60	$\mathbb{Z}^5\setminus(\{-1\}\cup\mathbb{Z}^2\cup\mathbb{Z}^3)$	$1\,(\mathrm{mod}\ 2)$ *and not* $31,41\,(\mathrm{mod}\ 60)$
-84	$\mathbb{Z}^3\setminus(\{-1\}\cup\mathbb{Z}^2)$	$3\,(\mathrm{mod}\ 6)$
105	$\mathbb{Z}^3\setminus(\{-1\}\cup\mathbb{Z}^2)$	$3\,(\mathrm{mod}\ 6)$
± 120	$\mathbb{Z}^3\setminus(\{-1\}\cup\mathbb{Z}^2)$	$3\,(\mathrm{mod}\ 6)$
± 168	$\mathbb{Z}^3\setminus(\{-1\}\cup\mathbb{Z}^2)$	$3\,(\mathrm{mod}\ 6)$
-420	$\mathbb{Z}^3\setminus(\{-1\}\cup\mathbb{Z}^2)$	$3\,(\mathrm{mod}\ 6)$
± 840	$\mathbb{Z}^3\setminus(\{-1\}\cup\mathbb{Z}^2)$	$3\,(\mathrm{mod}\ 6)$
other values	$\mathbb{Z}^3\setminus(\{-1\}\cup\mathbb{Z}^2)$	$3\,(\mathrm{mod}\ 6)$
every other value	$\mathbb{Z}\setminus(\{-1\}\cup\mathbb{Z}^2)$	$1\,(\mathrm{mod}\ 2)$

Theorem 7.1.7 enables one to describe all large enough integers having a representation as a sum of 3 primes with a prescribed primitive root.One such example is Corollary 7.1.2, whose proof we give now.

Proof of Corollary 7.1.2. If n is a sum of 3 odd primes all of which have primitive root 27, we saw in the first paragraph of this chapter that n must be $3 \bmod 12$. For the opposite direction we observe that if $a = 27$ then we have $\mathrm{Disc}(\mathbb{Q}(\sqrt{a})) = 12$ and $a \in \mathbb{Z}^3 \setminus (\{-1\} \cup \mathbb{Z}^2)$, hence alluding to the fifth row in the table of Theorem 7.1.7 we see that, conditionally on HRH(27), every sufficiently large integer $n \equiv 3\,(\mathrm{mod}\ 12)$ is a sum of three odd primes with primitive root 27. $\square$

7.1.6 Structure of the chapter

We study a generalisation of the ternary Goldbach problem in §7.2, where each of the three primes involved satisfies certain splitting conditions in a different number field extension of $\mathbb{Q}$. The main result of §7.2 is Proposition 7.2.1, whose proof is given in §7.2.3.

Next, §7.3.1 contains the first steps for the combination of Hooley's argument [35] and the Hardy–Littlewood circle method. Theorem 7.1.1 will be proved in §7.3.2, while Theorem 7.1.3 is verified in §7.3.3.

The rest of our chapter, namely §7.4, deals with the 'Artin factor' $\mathcal{A}_{\mathbf{a}}(n)$. The former part of Theorem 7.1.1, viz. (7.22), is verified in §7.4.1, while the latter part, viz. (7.23), is established in §7.4.2. Corollary 7.1.6 and Theorem 7.1.7 are proved in §7.4.4 and §7.4.5 respectively. Finally, we show that $\mathcal{A}_{\mathbf{a}}(n)$ does not factorise as an Euler product in §7.4.6.

Notation 7.1.8. The letters p and ℓ will always denote a rational prime. The entities $a_i, h_{a_i}, \Delta_{a_i}$ are considered constant throughout our work, thus the dependence of implied constants on them will not be recorded. On several occasions our implied constants are absolute, this will always be specified. Finally, we will use the notation

$$\mathrm{e}(z) := \exp(2\pi i z) \text{ and } \mathrm{e}_q(z) := \exp(2\pi i z/q), (z \in \mathbb{C}, q \in \mathbb{N}).$$

Acknowledgements. This work was completed while Christopher Frei and Peter Koymans were visiting the Max Planck Institute in Bonn, the hospitality of which is greatly acknowledged.

7.2 Uniform ternary Goldbach with certain splitting conditions

In this section the letters k, k_i shall refer exclusively to positive square-free integers. Recall (7.3) and define

$$\mathrm{Spl}\,(G_{a,k}) := \{p \text{ prime in } \mathbb{N} : p \text{ splits completely in } G_{a,k}\}. \tag{7.24}$$

We study the asymptotics of the representation function

$$V_{\mathbf{a},\mathbf{k}}(n) := \sum_{\substack{p_1+p_2+p_3=n \\ \forall i:\ p_i \in \mathrm{Spl}\left(G_{a_i,k_i}\right)}} \prod_{i=1}^{3} \log p_i. \tag{7.25}$$

We will see that the singular series related to the estimation of $V_{\mathbf{a},\mathbf{k}}(n)$ is the series $\mathfrak{S}_{\mathbf{a},\mathbf{k}}(n)$ introduced in (7.18). Kane [38] studied a very general set of problems, one case of which is that of evaluating $V_{\mathbf{a},\mathbf{k}}(n)$ asymptotically. His work provides a function $f_{\mathbf{a}}$ such that for each $B > 0$ and square-free k_1, k_2, k_3 we have

$$V_{\mathbf{a},\mathbf{k}}(n) = \frac{1}{2}\mathfrak{S}_{\mathbf{a},\mathbf{k}}(n) n^2 + O_B\left(|f_{\mathbf{a}}(\mathbf{k})| \frac{n^2}{(\log n)^B}\right), \tag{7.26}$$

where the implied constant depends at most on $\mathbf{a}$ and B. This can be deduced by taking

$$N := n,\ X := n,\ k := 3,\ a_i := 1,\ K_i := G_{a_i,k_i} \quad\text{and}\quad C_i := \mathrm{id}_{G_{a_i,k_i}}$$

in [38, Th.2]. With this choice the constant C_∞ in [38, Th.2] equals $n^2/2$ and a long but straightforward computation allows one to show that the 'singular series' $\mathfrak{S}_{\mathbf{a},\mathbf{k}}(n)$ can be factored into the remaining parts of the main term in the asymptotic formula [38, Eq.(1.2)].

Our aim in this section is to prove the following result, conditional on the hypothesis HRH'(a_i) introduced before Theorem 7.1.3. It constitutes a version of (7.26) that has a power saving in the error term and an explicit and polynomial dependence on the k_i. As is surely familiar to circle method experts, an error term of this quality is currently out of reach unconditionally even in the setting of the classical ternary Goldbach problem.

Proposition 7.2.1. *Assume* HRH'(a_i) *for $i = 1, 2, 3$. The following estimate holds for all square-free k_1, k_2, k_3 with $1 \le k_1, k_2, k_3 \le n$ and with an implied constant depending at most on* $\mathbf{a}$*,*

$$V_{\mathbf{a},\mathbf{k}}(n) = \frac{1}{2}\mathfrak{S}_{\mathbf{a},\mathbf{k}}(n)n^2 + O\Big(n^{11/6}(\log n)^6\big(\max_{1\le i\le 3} k_i\big)^6\Big).$$

7.2.1 Algebraic considerations

We shall need explicit bounds for certain algebraic quantities associated to $G_{a,k}$. This subsection is mostly devoted to providing the necessary estimates.

Recall the definitions of Δ_a and h_a, given in (7.4) and (7.5). We begin by determining the degree of the number field $F_{a,q,k}$ defined at the start of §7.1.3 (see [59, Lemma 2.3]).

Lemma 7.2.2. *For k square-free, set $k' := k/\gcd(k, h_a)$. Then we have*

$$[F_{a,q,k} : \mathbb{Q}] = k'\phi([q,k])/\epsilon(q,k),$$

where

$$\epsilon(q,k) = \begin{cases} 2, & \text{if } 2 \mid k \text{ and } \Delta_a \mid [q,k], \\ 1, & \text{otherwise.} \end{cases}$$

Lemma 7.2.3. *Let $k' = k/\gcd(k, h_a)$ and $a = g_1^{\gcd(k,h_a)} g_2^k$, with g_1 free of k'-th powers. Then*

$$\frac{\log|\mathrm{Disc}(F_{a,q,k})|}{[F_{a,q,k} : \mathbb{Q}]} \le \log k' + \log([q,k]) + 2\log|g_1|.$$

Proof. We have $|\mathrm{Disc}(F_{a,q,k})| = \mathfrak{N}(\Delta_{F_{a,q,k}/\mathbb{Q}(\zeta_{[q,k]})})|\mathrm{Disc}(\mathbb{Q}(\zeta_{[q,k]}))|^{[F_{a,q,k}:\mathbb{Q}(\zeta_{[q,k]})]}$, where $\mathfrak{N}$ is the absolute norm of an ideal and $\Delta_{F_{a,q,k}/\mathbb{Q}(\zeta_{[q,k]})}$ is the relative discriminant ideal. Any k'-th root $\alpha \in F_{a,q,k}$ of g_1 generates $F_{a,q,k}$ over $\mathbb{Q}(\zeta_{[q,k]})$, so it's different $d(\alpha) \ne 0$ is in the different ideal of $F_{a,q,k}/\mathbb{Q}(\zeta_{[q,k]})$. Since the minimal polynomial of α over $\mathbb{Q}(\zeta_{[q,k]})$

divides $x^{k'} - g_1$, we find that $k'\alpha^{k'-1}$ is a multiple of $d(\alpha)$ in $\mathcal{O}_{F_{a,q,k}}$, and thus in the different ideal as well. Hence,

$$\begin{aligned}\mathfrak{N}(\Delta_{F_{a,q,k}/\mathbb{Q}(\zeta_{[q,k]})}) &\leq |N_{F_{a,q,k}/\mathbb{Q}}(k'\alpha^{k'-1})| \\ &\leq (k')^{[F_{a,q,k}:\mathbb{Q}]}|g_1|^{(k'-1)\varphi([q,k])} \\ &\leq (k')^{[F_{a,q,k}:\mathbb{Q}]}|g_1|^{2[F_{a,q,k}:\mathbb{Q}]}.\end{aligned}$$

Now use

$$|\mathrm{Disc}(\mathbb{Q}(\zeta_{[q,k]}))| = [q,k]^{\varphi([q,k])} \prod_{p|qk} p^{-\varphi([q,k])/(p-1)} \leq [q,k]^{\varphi([q,k])}$$

to complete the proof. □

Clearly, the intersection $\mathbb{Q}(\zeta_q) \cap G_{a,k}$ contains $\mathbb{Q}(\zeta_{\gcd(q,k)})$. More precisely, it is determined as follows (see [59, Lemma 2.4]).

Lemma 7.2.4. *We have*

$$[\mathbb{Q}(\zeta_q) \cap G_{a,k} : \mathbb{Q}(\zeta_{\gcd(q,k)})] = \begin{cases} 2 & \text{if } 2 \mid k,\ \Delta_a \nmid k \text{ and } \Delta_a \mid [q,k] \\ 1 & \text{otherwise.} \end{cases}$$

In the first case, the integer $\beta_a(q)$ defined in (7.10) *is a fundamental discriminant and we have $\mathbb{Q}(\zeta_q) \cap G_{a,k} = \mathbb{Q}(\zeta_{\gcd(q,k)}, \sqrt{\beta_a(q)})$.*

Since both $\mathbb{Q}(\zeta_q)$ and $G_{a,k}$ are normal, the same holds for their compositum $F_{a,q,k}$. We investigate the existence of certain elements of the Galois group $\mathrm{Gal}(F_{a,q,k}/\mathbb{Q})$. Recall the definitions of σ_b and $c_{a,q,k}(b)$ from the start of §7.1.3.

Lemma 7.2.5. *Let $b \in \mathbb{Z}$ with* $\gcd(b,q) = 1$*. The following are equivalent:*

1. *there is an automorphism $\sigma \in \mathrm{Gal}(F_{a,q,k}/\mathbb{Q})$ with*

$$\sigma|_{\mathbb{Q}(\zeta_q)} = \sigma_b \quad \text{and} \quad \sigma|_{G_{a,k}} = \mathrm{id}_{G_{a,k}}, \tag{7.27}$$

2. $c_{a,q,k}(b) = 1$,

3. *with $\beta_a(q)$ defined in* (7.10)*, we have*

$$b \equiv 1 \pmod{\gcd(q,k)}, \qquad \text{and} \tag{7.28}$$

$$2 \mid k,\ \Delta_a \nmid k,\ \Delta_a \mid [q,k] \quad \text{implies that} \quad \left(\frac{\beta_a(q)}{b}\right) = 1. \tag{7.29}$$

Moreover, if σ as in (1) exists, it is unique and in the center of $\mathrm{Gal}(F_{a,q,k})/\mathbb{Q}$.

Proof. Write $I := \mathbb{Q}(\zeta_q) \cap G_{a,k}$. The map $\sigma \mapsto (\sigma|_{\mathbb{Q}(\zeta_q)}, \sigma|_{G_{a,k}})$ provides an isomorphism

$$\mathrm{Gal}(F_{a,q,k}/\mathbb{Q}) \cong \{(\sigma_1, \sigma_2) \in \mathrm{Gal}(\mathbb{Q}(\zeta_q)/\mathbb{Q}) \times \mathrm{Gal}(G_{a,k}/\mathbb{Q}) \ : \ \sigma_1|_I = \sigma_2|_I\}.$$

Thus, an automorphism σ with (7.27) exists if and only if $c_{a,q,k}(b) = 1$, proving the equivalence of *(1)* and *(2)*. In this case σ is necessarily unique and clearly in the center of $\mathrm{Gal}(F_{a,q,k}/\mathbb{Q})$, because the Galois group $\mathrm{Gal}(\mathbb{Q}(\zeta_q)/\mathbb{Q})$ is abelian and $\mathrm{id}_{G_{a,k}}$ is in the center of $\mathrm{Gal}(G_{a,k}/\mathbb{Q})$. Thus, let us study the conditions under which $c_{a,q,k}(b) = 1$.

Since $\mathbb{Q}(\zeta_{\gcd(q,k)}) \subset I$ and $\sigma_b|_{\mathbb{Q}(\zeta_{\gcd(q,k)})}$ coincides with the automorphism given by $\zeta \mapsto \zeta^{b \pmod{\gcd(q,k)}}$, the condition (7.28) is clearly necessary. Thus, we assume it to hold from now on, whence $\sigma_b|_{\mathbb{Q}(\zeta_{\gcd(q,k)})} = \mathrm{id}_{G_{a,k}}$. If the antecedent in (7.29) is false, then we have $I = \mathbb{Q}(\zeta_{\gcd(q,k)})$ by Lemma 7.2.4, and thus $c_{a,q,k}(b) = 1$. If the antecedent in (7.29) holds, then, invoking Lemma 7.2.4 once more, we find that $\sqrt{\beta_a(q)} \in \mathbb{Q}(\zeta_q)$ and $c_{a,q,k}(b) = 1$ is equivalent to

$$\sigma_b(\sqrt{\beta_a(q)}) = \sqrt{\beta_a(q)}. \tag{7.30}$$

Since $\beta_a(q)$ is a fundamental discriminant, we may invoke [59, Lemma 2.2] to see that (7.30) is equivalent to $\left(\frac{\beta_a(q)}{b}\right) = 1$. □

7.2.2 Consequences of HRH'(a)

In this section we use the hypothesis HRH'(a) to provide estimates for certain exponential sums related to the estimation of $V_{\mathbf{a},\mathbf{k}}(n)$.

Lemma 7.2.6. *Assume* HRH'(a)*. For any square-free k and coprime integers c, q we have*

$$\sum_{\substack{p \le x \\ p \in \mathrm{Spl}(G_{a,k})}} (\log p) \mathrm{e}_q(cp) = \frac{x}{\varphi(q)[G_{a,k} : \mathbb{Q}]} \sum_{\substack{\chi \pmod{q} \\ \chi \circ \mathfrak{N} = \chi_0}} \overline{\chi(c)} \tau(\chi) + O(k^2 \sqrt{qx} (\log qx)^2).$$

Here, χ runs through all Dirichlet characters modulo q for which $\chi \circ \mathfrak{N}$, considered as a ray class character modulo $q\mathcal{O}_{G_{a,k}}$, is the trivial ray class character χ_0. Moreover, $\tau(\chi)$ denotes the Gauss sum $\tau(\chi) = \sum_{y \pmod{q}} \chi(y) \mathrm{e}_q(y)$.

Proof. We have

$$\sum_{\substack{p \le x \\ p \in \mathrm{Spl}(G_{a,k})}} (\log p) \mathrm{e}_q(cp) = \sum_{\substack{p \le x, p \nmid q \\ p \in \mathrm{Spl}(G_{a,k})}} (\log p) \mathrm{e}_q(cp) + O((\log q)^2). \tag{7.31}$$

Bringing into play the Dirichlet characters modulo q allows us to inject, for $p \nmid q$,

$$\mathrm{e}_q(cp) = \frac{1}{\varphi(q)} \sum_{b \pmod{q}} \sum_{\chi \pmod{q}} \chi(b) \overline{\chi(cp)} \mathrm{e}_q(b) = \frac{1}{\varphi(q)} \sum_{\chi \pmod{q}} \overline{\chi(cp)} \tau(\chi)$$

into (7.31), thus acquiring the validity of

$$\sum_{\substack{p\leq x\\ p\in\mathrm{Spl}(G_{a,k})}}(\log p)\mathrm{e}_q(cp)=\frac{1}{\varphi(q)}\sum_{\chi(\mathrm{mod}\ q)}\overline{\chi(c)}\tau(\chi)\psi_{a,k}(x,\overline{\chi})+O((\log q)^2),\tag{7.32}$$

where

$$\begin{aligned}\psi_{a,k}(x,\chi):=&\sum_{\substack{p\leq x\\ p\in\mathrm{Spl}(G_{a,k})}}(\log p)\chi(p)=\frac{1}{[G_{a,k}:\mathbb{Q}]}\sum_{\substack{\mathfrak{N}\mathfrak{p}\leq x\\ \deg(\mathfrak{p})=1}}(\log\mathfrak{N}\mathfrak{p})\chi(\mathfrak{N}\mathfrak{p})\\ =&\frac{1}{[G_{a,k}:\mathbb{Q}]}\sum_{\mathfrak{N}\mathfrak{n}\leq x}\Lambda(\mathfrak{n})\chi(\mathfrak{N}\mathfrak{n})+O(\sqrt{x}\log x).\end{aligned}$$

Here and for the rest of this section $\mathfrak{p}$ denotes a prime ideal in $\mathcal{O}_{G_{a,k}}$, $\deg(\mathfrak{p})$ denotes its inertia degree over $\mathbb{Q}$, $\mathfrak{n}$ denotes an ideal in $\mathcal{O}_{G_{a,k}}$, and Λ is the von Mangoldt function on ideals of $\mathcal{O}_{G_{a,k}}$, defined by $\Lambda(\mathfrak{p}^e):=\log\mathfrak{N}\mathfrak{p}$ for $e\geq 1$ and $\Lambda(\mathfrak{n}):=0$ in all other cases. Observing that $\chi\circ\mathfrak{N}$ defines a character of the ray class group of $G_{a,k}$ modulo $q\mathcal{O}_{G_{a,k}}$, we consider its Hecke L-function,

$$L(s,\chi):=\sum_{\mathfrak{n}\neq 0}\chi(\mathfrak{N}\mathfrak{n})(\mathfrak{N}\mathfrak{n})^{-s}.$$

It is now easy to see that

$$-L'(s,\chi)/L(s,\chi)=\sum_{\mathfrak{n}\neq 0}\Lambda(\mathfrak{n})\chi(\mathfrak{N}\mathfrak{n})(\mathfrak{N}\mathfrak{n})^{-s}.$$

The Ramanujan–Petersson conjecture is obviously true for $L(s,\chi)$, since it is true for any Hecke L-function. Hence Theorem 5.15 from [37] implies that

$$\sum_{\mathfrak{N}\mathfrak{n}\leq x}\Lambda(\mathfrak{n})\chi(\mathfrak{N}\mathfrak{n})=r_\chi x+O(x^{\frac{1}{2}}(\log x)\log(x^{[G_{a,k}:\mathbb{Q}]}\mathfrak{q}(\chi))),$$

where r_χ is the order of the pole of $L(s,\chi)$ at $s=1$. For the definition of $\mathfrak{q}(\chi)$, see page 95 of [37]. Furthermore, on page 129 of [37] it is proven that

$$\mathfrak{q}(\chi)\leq 4^{[G_{a,k}:\mathbb{Q}]}|\mathrm{Disc}(G_{a,k})|q^{[G_{a,k}:\mathbb{Q}]}.$$

Our next task is to make explicit the value of r_χ. If $\chi\circ\mathfrak{N}$ is the trivial ray class character χ_0 modulo $\mathcal{O}_{G_{a,k}}$, then we have $r_\chi=1$; otherwise we have $r_\chi=0$. Using $|\tau(\chi)|\leq\sqrt{q}$ and Lemma 7.2.3 we can substitute in (7.32) to find that

$$\begin{aligned}\frac{1}{\varphi(q)}\sum_{\chi(\mathrm{mod}\ q)}\overline{\chi(c)}\tau(\chi)\psi_{a,k}(x,\overline{\chi})=\frac{x\varphi(q)^{-1}}{[G_{a,k}:\mathbb{Q}]}\sum_{\substack{\chi(\mathrm{mod}\ q)\\ \chi\circ\mathfrak{N}=\chi_0}}\overline{\chi(c)}\tau(\chi)+&\\ O([G_{a,k}:\mathbb{Q}]\sqrt{qx}(\log qx)^2),&\end{aligned}$$

thus concluding our proof upon observing that $[G_{a,k}:\mathbb{Q}]=[F_{a,k,k}:\mathbb{Q}]\leq k^2$. □

Although it is possible to directly evaluate the main term in Lemma 7.2.6, we will instead use the following trick.

Lemma 7.2.7. *Under the same conditions as in Lemma 7.2.6 we have*

$$\sum_{\substack{p\leq x\\ p\in\mathrm{Spl}(G_{a,k})}} (\log p)\mathrm{e}_q(cp) = \frac{S_{a,q,k}(c)}{[F_{a,q,k}:\mathbb{Q}]}x + o_{q,k}(x), \ \ \textit{as } x\to+\infty.$$

Proof. Partitioning in progressions modulo q we see that, owing to (7.31), the sum over p in our lemma is equal to the following quantity up to an error of size $o_{q,k}(x)$,

$$\sum_{b\in(\mathbb{Z}/q\mathbb{Z})^*} \mathrm{e}_q(bc) \sum_{\substack{p\leq x\\ p\equiv b (\mathrm{mod}\ q)\\ p\in\mathrm{Spl}(G_{a,k})}} \log p.$$

By Lemma 7.2.5 there exists an automorphism σ of $F_{a,q,k}$ satisfying

$$\sigma|_{\mathbb{Q}(\zeta_q)} = \sigma_b \ \text{ and } \ \sigma|_{G_{a,k}} = \mathrm{id}_{G_{a,k}}$$

if and only if $c_{a,q,k}(b) = 1$. Furthermore, if such an automorphism exists, it is unique. The lemma is now a consequence of Chebotarev's density theorem. □

Combining Lemma 7.2.6 and Lemma 7.2.7 proves the following lemma.

Lemma 7.2.8. *Under the same assumptions as in Lemma 7.2.6 we have*

$$\sum_{\substack{p\leq x\\ p\in\mathrm{Spl}(G_{a,k})}} (\log p)\mathrm{e}_q(cp) = \frac{S_{a,q,k}(c)x}{[F_{a,q,k}:\mathbb{Q}]} + O(k^2\sqrt{qx}\log^2 qx).$$

Define for a square-free integer $k>0$ the exponential sum

$$f_{a,k}(\alpha) = \sum_{\substack{p\leq n\\ p\in\mathrm{Spl}(G_{a,k})}} (\log p)\mathrm{e}(\alpha p), \ \ (\alpha\in\mathbb{R}). \tag{7.33}$$

The next lemma is easily proved via partial summation and Lemma 7.2.8.

Lemma 7.2.9. *Assume* HRH'(a)*. Let k be square-free integer and define $\alpha = c/q+\beta$, where $(c,q)=1$. Then*

$$f_{a,k}(\alpha) = \frac{S_{a,q,k}(c)}{[F_{a,q,k}:\mathbb{Q}]}\int_0^n \mathrm{e}(\beta x)\mathrm{d}x + O\left(k^2(1+|\beta|n)\sqrt{qn}(\log qn)^2\right).$$

It will be necessary to gain a better understanding of the exponential sums $S_{a,q,k}(c)$. We start by studying $c_{a,q,k}(\cdot)$ in the next lemma, whose proof flows directly from (7.28) and (7.29).

Lemma 7.2.10. *Let b, q be coprime integers and factor q as $q = d\prod_{i=1}^{l} p_i^{e_i}$ with d an integer composed of primes dividing Δ_a and p_i distinct prime numbers not dividing Δ_a. Then we have for any square-free integer k,*

$$c_{a,q,k}(b) = c_{a,d,k}(b)\prod_{i=1}^{l} c_{a,p_i^{e_i},k}(b).$$

Lemma 7.2.11. *Let k be square-free, assume that b, q are coprime integers and suppose that $q = q_1q_2$, $b = b_1q_2 + b_2q_1$, with q_1, q_2 coprime. If $\gcd(q_1, \Delta_a) = 1$ or $\gcd(q_2, \Delta_a) = 1$ then we have*

$$S_{a,q,k}(b) = S_{a,q_1,k}(b_1)S_{a,q_2,k}(b_2).$$

Proof. By the Chinese remainder theorem we can write each element $y \in \mathbb{Z}/q\mathbb{Z}$ as $y_1q_2 + y_2q_1$, where $y_i \in \mathbb{Z}/q_i\mathbb{Z}$, thus showing that $\mathrm{e}_q(by) = \mathrm{e}_{q_1}(b_1y_1q_2)\mathrm{e}_{q_2}(b_2y_2q_1)$. This leads to

$$\begin{aligned} S_{a,q,k}(b) &= \sum_{y\in(\mathbb{Z}/q\mathbb{Z})^*} c_{a,q,k}(y)\mathrm{e}_q(by) \\ &= \sum_{y_1\in(\mathbb{Z}/q_1\mathbb{Z})^*} \mathrm{e}_{q_1}(b_1y_1q_2) \sum_{y_2\in(\mathbb{Z}/q_2\mathbb{Z})^*} \mathrm{e}_{q_2}(b_2y_2q_1)c_{a,q,k}(y_1q_2 + y_2q_1). \end{aligned}$$

By Lemma 7.2.10 we have $c_{a,q,k}(y_1q_2 + y_2q_1) = c_{a,q_1,k}(y_1q_2 + y_2q_1)c_{a,q_2,k}(y_1q_2 + y_2q_1)$. The entity $c_{a,q,k}(y)$ is periodic $(\mathrm{mod}\ q)$ as a function of y, thus we can write $S_{a,q,k}(b)$ as

$$\sum_{y_1\in(\mathbb{Z}/q_1\mathbb{Z})^*} \mathrm{e}_{q_1}(b_1y_1q_2)c_{a,q_1,k}(y_1q_2) \sum_{y_2\in(\mathbb{Z}/q_2\mathbb{Z})^*} \mathrm{e}_{q_2}(b_2y_2q_1)c_{a,q_2,k}(y_2q_1)$$

and a simple linear change of variables in each sum completes the proof. □

Lemma 7.2.12. *For k square-free, b an integer and p a prime with $p \nmid b\Delta_a$ we have*

$$|S_{a,p^j,k}(b)| = \begin{cases} 1, & j = 1 \\ 0, & j > 1. \end{cases}$$

Proof. Let us observe that (7.29) always holds for $q = p^j$ as in the lemma, as the antecedent is never satisfied. We first handle the case $j = 1$. If $p \nmid k$ then by Lemma 7.2.5, $S_{a,p,k}(b)$ is the classical Ramanujan sum and the result follows, while in the remaining case, $p \mid k$, the result is also immediate from (7.28). Now suppose $j > 1$. Again, if $p \nmid k$, the sum in the lemma is a Ramanujan sum and the result follows. We are therefore free to assume that $p \mid k$. Writing $y = 1 + px$ we see that

$$S_{a,p^j,k}(b) = \sum_{\substack{y(\mathrm{mod}\ p^j) \\ y\equiv 1(\mathrm{mod}\ p)}} \mathrm{e}_{p^j}(by) = \mathrm{e}_{p^j}(b) \sum_{x(\mathrm{mod}\ p^{j-1})} \mathrm{e}_{p^{j-1}}(bx),$$

which is clearly sufficient since the inner sum vanishes. □

Lemma 7.2.13. *Let $r, Q, c \in \mathbb{Z}$ be such that $rQ \neq 0, \gcd(c, Q) = 1$, r divides Q and assume that a function $f : \mathbb{Z} \to \mathbb{C}$ has period $|r|$. If we have $|r| < |Q|$ then the following sum vanishes,*

$$\sum_{b(\bmod\ |Q|)} \mathrm{e}_{|Q|}(bc)f(b).$$

Proof. The claim becomes clear upon writing the sum in our lemma as

$$\sum_{b_0(\bmod\ |r|)} \mathrm{e}_{|Q|}(b_0c)f(b_0) \sum_{x(\bmod\ |Q/r|)} \mathrm{e}_{|Q/r|}(xc)$$

and observing that if $|Q/r| \neq 1$ then each exponential sum over x vanishes. □

Lemma 7.2.14. *Let k be a square-free integer, suppose that q is composed of primes dividing Δ_a and let b be an integer with $\gcd(b, q) = 1$. If $q \nmid \Delta_a$, then $S_{a,q,k}(b) = 0$.*

Proof. First suppose $2 \nmid k$ or $\Delta_a \mid k$ or $\Delta_a \nmid [q, k]$ and write $q = p_1^{e_1} \cdots p_l^{e_l}$. We have

$$c_{a,q,k}(b) = \prod_{i=1}^{l} c_{a,p_i^{e_i},k}(b),$$

therefore $S_{a,q,k}(b) = 0$ can now be easily proved as before, as our hypotheses imply that $e_j > 1$ for at least one j.

Now suppose that $2 \mid k$ and $\Delta_a \nmid k$ and $\Delta_a \mid [q, k]$. For $y \in \mathbb{Z}$, let $f(y) := 1$ if $y \equiv 1 \bmod \gcd(k, q)$ and $\left(\frac{\beta_a(q)}{y}\right) = 1$, and $f(y) := 0$ otherwise. By Lemma 7.2.5 we have

$$S_{a,q,k}(b) = \sum_{y(\bmod\ q)} f(y)\mathrm{e}_q(by).$$

Since $\gcd(k, q) \mid \gcd(\Delta_a, q) = |\beta_a(q)|$ and $\beta_a(q)$ is a fundamental discriminant, we see that f has period $\gcd(\Delta_a, q)$, strictly dividing q by our hypotheses. Apply Lemma 7.2.13. □

Combining Lemmas 7.2.11, 7.2.12 and 7.2.14 allows us to conclude that

$$S_{a,q,k}(b) \ll 1, \tag{7.34}$$

where the implied constant depends at most on a.

7.2.3 Proof of Proposition 7.2.1

Recall (7.33). Our starting point is the circle method identity,

$$\sum_{\substack{p_1+p_2+p_3=n \\ p_i \in \mathrm{Spl}(G_{a_i,k_i})}} \prod_{i=1}^{3} (\log p_i) = \int_0^1 f_{a_1,k_1}(\alpha) f_{a_2,k_2}(\alpha) f_{a_3,k_3}(\alpha) \mathrm{e}(-n\alpha) \mathrm{d}\alpha. \tag{7.35}$$

Corollary 7.2.15. *Assume* HRH'(a)*, and suppose* α, c, q *fulfil* $|\alpha - c/q| \le q^{-1}n^{-2/3}$*,* $\gcd(c,q) = 1$*,* $q \le n^{2/3}$ *and that* k *is square-free. Then we have*

$$f_{a,k}(\alpha) \ll (n/q + k^2 n^{5/6})(\log n)^2.$$

Proof. Observe that Lemma 7.2.2 gives

$$[F_{a,q,k} : \mathbb{Q}]^{-1} \ll \varphi([q,k])^{-1} \le \varphi(q)^{-1} \ll (\log q) q^{-1},$$

hence, by Lemma 7.2.9 and (7.34) one obtains

$$f_{a,k}(\alpha) \ll n(\log n)q^{-1} + k^2(1 + n^{1/3}q^{-1})\sqrt{qn}(\log n)^2.$$

Our proof can then be concluded by using $q \le n^{2/3}$. $\square$

Define $P := n^\nu$, for an absolute constant $\nu \in (0, 1/6]$ that will be chosen later. In our situation the major arc $\mathfrak{M}(c,q)$ is defined for coprime c, q via

$$\mathfrak{M}(q,c) := \{\alpha : |\alpha - c/q| \le q^{-1}n^{-2/3}\},$$

while we let $\mathfrak{M}$ be the union of all $\mathfrak{M}(q,c)$ with $1 \le q \le P$, $1 \le c \le q, \gcd(c,q) = 1$ and define the minor arcs through $\mathfrak{m} := [0,1] \setminus \mathfrak{M}$. We note here that the major arcs are disjoint owing to $(qq')^{-1} > (qn^{2/3})^{-1} + (q'n^{2/3})^{-1}$ that can be proved for all $n > 8$ due to $q, q' \le n^{1/3}$.

Corollary 7.2.16. *Assume* HRH'(a_i) *for* $1 \le i \le 3$*. Then*

$$\int_{\mathfrak{m}} |f_{a_1,k_1}(\alpha) f_{a_2,k_2}(\alpha) f_{a_3,k_3}(\alpha)| \mathrm{d}\alpha \ll n^{2-\nu}(\log n)^3 \min_i k_i^2.$$

Proof. By Dirichlet's approximation theorem, for each α there exist coprime integers c, q with $|\alpha - c/q| \le q^{-1}n^{-2/3}$ and $1 \le q \le n^{2/3}$. If $\alpha \in \mathfrak{m}$ then $q > n^\nu$, hence Corollary 7.2.15 yields the estimate $f_{a,k}(\alpha) \ll k^2 n^{1-\nu}(\log n)^2$. We may assume $k_1 \le k_2, k_3$ with no loss of generality, therefore the integral in our lemma is

$$\ll k_1^2 n^{1-\nu}(\log n)^2 \int_0^1 |f_{a_2,k_2}(\alpha) f_{a_3,k_3}(\alpha)| \mathrm{d}\alpha,$$

thus Cauchy's inequality yields the following bound for the last integral,

$$\ll \left(\int_0^1 |f_{a_2,k_2}(\alpha)|^2 \mathrm{d}\alpha\right)^{1/2} \left(\int_0^1 |f_{a_3,k_3}(\alpha)|^2 \mathrm{d}\alpha\right)^{1/2}.$$

Both integrals are at most $\sum_{p \le n} (\log p)^2 \ll n \log n$, which provides the desired result. $\square$

Note that if $\beta + c/q \in \mathfrak{M}(q,c)$ for some $q \le n^{1/3}$ then Lemma 7.2.9 shows that

$$f_{a_i,k_i}(\alpha) = \frac{S_{a_i,q,k_i}(c)}{[F_{a_i,q,k_i} : \mathbb{Q}]} \int_0^n \mathrm{e}(\beta x) \mathrm{d}x + O\left(\frac{n^{5/6}}{q^{1/2}} (\log n)^2 \max_i k_i^2\right).$$

Hence the estimates

$$\int_0^n \mathrm{e}(\beta x)\mathrm{d}x \ll \min\{n,|\beta|^{-1}\} \quad \text{and} \quad \frac{S_{a,q,k}(c)}{[F_{a,q,k}:\mathbb{Q}]} \ll \varphi(q)^{-1}$$

show that $f_{a_1,k_1}(c/q+\beta)f_{a_2,k_2}(c/q+\beta)f_{a_3,k_3}(c/q+\beta)-L_{\mathbf{a},q,\mathbf{k}}(c)d_{\mathbf{a},\mathbf{k}}(q)^{-1}\left(\int_0^n \mathrm{e}(\beta x)\mathrm{d}x\right)^3$ is

$$\ll \frac{\min\{n^2,|\beta|^{-2}\}}{\varphi(q)^2}\frac{n^{5/6}}{q^{1/2}}(\log n)^2 \max_i k_i^2 + \frac{n^{15/6}}{q^{3/2}}(\log n)^6 \max_i k_i^6. \tag{7.36}$$

The major arcs make the following contribution towards (7.35),

$$\sum_{1\leq q\leq n^\nu}\sum_{\substack{1\leq c\leq q\\ \gcd(c,q)=1}}\int_{-q^{-1}n^{-2/3}}^{q^{-1}n^{-2/3}} f_{a_1,k_1}(c/q+\beta)f_{a_2,k_2}(c/q+\beta)f_{a_3,k_3}(c/q+\beta)\mathrm{e}(-n(c/q+\beta))\mathrm{d}\beta,$$

and a straightforward analysis utilising (7.36) reveals that the last expression equals

$$\sum_{1\leq q\leq n^\nu}\sum_{\substack{1\leq c\leq q\\ \gcd(c,q)=1}}\frac{\mathrm{e}_q(-cn)L_{\mathbf{a},q,\mathbf{k}}(c)}{d_{\mathbf{a},\mathbf{k}}(q)}\int_{-q^{-1}n^{-2/3}}^{q^{-1}n^{-2/3}}\left(\int_0^n \mathrm{e}(\beta x)\mathrm{d}x\right)^3 \mathrm{e}(-n\beta)\mathrm{d}\beta +$$

$$O\left(\frac{n^{11/6}(\log n)^6}{\max_i k_i^{-6}}\right).$$

The integral over β can be estimated as $n^2/2+O(q^2n^{4/3})$, thus by (7.34) the sum over q is $\mathfrak{S}_{\mathbf{a},\mathbf{k}}(n)n^2/2+O((n^{4/3+\nu}+n^{2-\nu})(\log n)^3)$ and the choice $\nu=1/6$ concludes the proof of Proposition 7.2.1.

7.3 The circle method and Hooley's approach

7.3.1 Opening phase

The aim of §7.3 is to prove Theorem 7.1.1 and Theorem 7.1.3. We commence in this subsection by calling upon parts of Hooley's work [35] that will prove useful. We will make an effort to keep the notation in line with his as much as possible. In this section, the letters p,q will be reserved for primes. Two primes p,q are said to satisfy the property $R_a(q,p)$ if both of the following conditions hold,

$$q|(p-1); a \text{ is a } q\text{th power residue } \pmod p.$$

A standard index calculus argument shows that for a prime $p\nmid a$ the integer a is a primitive root $\pmod p$ if and only if $R_a(q,p)$ fails for all primes q. For any $\eta,\eta_1,\eta_2\in\mathbb{R}_{>0}$ we define

$$\mathrm{N}_a(n,\eta) := \#\{p\leq n : R_a(q,p) \text{ fails for all primes } q\leq\eta\}$$

and

$$\mathrm{M}_a(n,\eta_1,\eta_2) := \#\big\{p \le n : \text{ there exists } q \in (\eta_1,\eta_2] \text{ such that } R_a(q,p) \text{ holds}\big\}.$$

Letting

$$\mathrm{N}_a(n) := \#\{p \le n : a \text{ is a primitive root modulo } p\}$$

we see from the work of Hooley [35, Eq.(1)] that for each $\xi_1,\xi_2,\xi_3 \in \mathbb{R}$ with

$$1 \le \xi_1 < \xi_2 < \xi_3 < n-1$$

we have

$$\mathrm{N}_a(n) = \mathrm{N}_a(n,\xi_1) + O\big(\mathrm{M}_a(n,\xi_1,\xi_2) + \mathrm{M}_a(n,\xi_2,\xi_3) + \mathrm{M}_a(n,\xi_3,n-1)\big). \tag{7.37}$$

Hooley makes specific choices for the parameters ξ_i; we will keep the same choice for ξ_2 and ξ_3, namely $\xi_2 := n^{\frac{1}{2}}(\log n)^{-2}$, $\xi_3 := n^{\frac{1}{2}}\log n$, however, we shall later choose a different value for ξ_1. For the moment we shall only demand that $1 < \xi_1 \le (\log n)(\log\log n)^{-1}$. The estimates proved in [35, Eq.(2), Eq.(3)] provide us with

$$\mathrm{N}_a(n) = \mathrm{N}_a(n,\xi_1) + O\big(\mathrm{M}_a(n,\xi_1,\xi_2) + n(\log\log n)(\log n)^{-2}\big). \tag{7.38}$$

The argument in [35, Eq.(33)] shows that for each ξ_1 as above, one has under HRH(a) that

$$\mathrm{M}_a(n,\xi_1,\xi_2) \ll \frac{n}{\log n}\sum_{q>\xi_1}\frac{1}{q^2} + \frac{n}{\log^2 n},$$

which, once combined with the simple estimate $\sum_{q>\xi_1} q^{-2} \ll \xi_1^{-1}$ and (7.38) provides us with

$$\mathrm{N}_a(n) = \mathrm{N}_a(n,\xi_1) + O\left(\frac{n}{\log n}\frac{1}{\xi_1} + \frac{n\log\log n}{\log^2 n}\right), \tag{7.39}$$

with an implied constant depending at most on a.

Lemma 7.3.1. *For any $\beta \in (0,1)$ and any sets of primes $\mathcal{P}_i \subset [1,n]$ of cardinality $\epsilon(\mathcal{P}_i)n/\log n$ the following estimate holds with an implied constant that depends at most on β,*

$$\sum_{\substack{p_1+p_2+p_3=n\\ \exists i: p_i\in\mathcal{P}_i}}\prod_{i=1}^{3}\log p_i \ll_\beta n^2(\max_i \epsilon(\mathcal{P}_i))^\beta.$$

Proof. Define $r_2(m) := \#\{(p_1,p_2) : p_i \text{ prime}, p_1+p_2=m\}$. The sum in the lemma is at most

$$(\log n)^3\sum_{i=1}^{3}\sum_{\substack{p_1+p_2+p_3=n\\ p_i\in\mathcal{P}_i}} 1 = (\log n)^3\sum_{i=1}^{3}\sum_{p<n}\mathbf{1}_{\mathcal{P}_i}(p)r_2(n-p)$$

and using Hölder's inequality with exponents $(1/\beta, 1/(1-\beta))$ allows us to bound the inner sum on the right by

$$\epsilon(\mathcal{P}_i)^\beta n^\beta(\log n)^{-\beta}\Big(\sum_{p<n} r_2(n-p)^{1/(1-\beta)}\Big)^{1-\beta}.$$

Straightforwardly, there exists $c = c(\beta) > 0$ with $(1-z)/(1-2z) \le (1+cz)^{1-\beta}$ for all $0 < z \le 1/3$. Using this for $z = 1/p'$ and alluding to the following classical bound (that can be found in [29, Eq. (7.2)], for example),

$$r_2(m) \ll \frac{m}{(\log m)^2} \prod_{p'|m, p' \neq 2} \frac{p'-1}{p'-2}$$

yields

$$r_2(m) \ll_\beta \frac{m}{(\log m)^2} \prod_{p'|m} \left(1 + \frac{c}{p'}\right)^{1-\beta}.$$

Therefore the quantity in the lemma is

$$\ll (\log n)^3 \Big(\frac{n \max_i \epsilon(\mathcal{P}_i)}{\log n}\Big)^\beta \Big(\Big(\frac{n}{(\log n)^2}\Big)^{1/(1-\beta)} \sum_{p<n} \prod_{p'|n-p} (1 + c/p')\Big)^{1-\beta}$$

and to finish our proof it remains to show that

$$\sum_{p<n} \prod_{p'|n-p} (1 + c/p') \ll_c \frac{n}{\log n}.$$

Rewriting this sum as $\sum_{d \le n} \mu(d)^2 c^{\omega(d)} d^{-1} \#\{p < n : p \equiv n \,(\mathrm{mod}\ d)\}$ we see that the contribution from integers $d > n^{1/2}$ is $\ll \sum_{n^{1/2} < d \le n} c^{\omega(d)} d^{-1}(n/d + 1) \ll n^{1/2+1/100}$. By Brun–Titchmarsh, the contribution of terms with $d \le n^{1/2}$ is

$$\ll n(\log n)^{-1} \sum_{d \le n^{1/2}} c^{\omega(d)} (d\phi(d))^{-1} \ll n(\log n)^{-1},$$

thus concluding our proof. □

Let us define the set

$$\mathcal{P}_i := \big\{p : p|a_i\big\} \cup \big\{p \le n : R_{a_i}(q,p) \text{ holds for some prime } q > \xi_1\big\}.$$

The arguments bounding $\mathrm{M}_a(n, \xi_1, n-1)$ in the deduction of (7.39) show under HRH(a) that

$$\#\mathcal{P}_i \ll \frac{n}{\xi_1 \log n} + \frac{n \log\log n}{\log^2 n}. \tag{7.40}$$

We can now apply Lemma 7.3.1 and to do so let us observe that by (7.40) we have

$$\epsilon(\mathcal{P}_i) = \frac{\log n}{n} \#\mathcal{P}_i \ll \frac{1}{\xi_1} + \frac{\log\log n}{\log n} \ll \frac{1}{\xi_1}.$$

Therefore, under HRH(a_i) for $i = 1, 2, 3$, and for each fixed $\beta \in (0,1)$ we acquire the validity of

$$\sum_{\substack{p_1+p_2+p_3=n \\ \forall i:\ \mathbb{F}_{p_i}^* = \langle a_i \rangle}} \prod_{i=1}^3 \log p_i = \sum_{\substack{p_1+p_2+p_3=n, p_i \nmid a_i \\ \forall i, \forall q \le \xi_1:\ R_{a_i}(q,p_i) \text{ fails}}} \prod_{i=1}^3 \log p_i + O_\beta\Big(\frac{n^2}{\xi_1^\beta}\Big). \tag{7.41}$$

Bringing into play the following quantity for each square-free positive integer k_i,

$$\mathrm{P}_{\mathbf{a},\mathbf{k}}(n) := \sum_{\substack{p_1+p_2+p_3=n,\ p_i \nmid a_i \\ \forall i:\ q|k_i \Rightarrow R_{a_i}(q,p_i) \text{ holds}}} \prod_{i=1}^{3} \log p_i, \tag{7.42}$$

makes the following estimate available, once the inclusion-exclusion principle has been used,

$$\sum_{\substack{p_1+p_2+p_3=n \\ \forall i:\ \mathbb{F}_{p_i}^* = \langle a_i \rangle}} \prod_{i=1}^{3} \log p_i = \sum_{\substack{\mathbf{k}\in\mathbb{N}^3 \\ p|k_1k_2k_3 \Rightarrow p \leq \xi_1}} \mu(k_1)\mu(k_2)\mu(k_3)\mathrm{P}_{\mathbf{a},\mathbf{k}}(n) + O_\beta\Big(n^2\xi_1^{-\beta}\Big). \tag{7.43}$$

The entity $\mathrm{P}_{\mathbf{a},\mathbf{k}}(n)$ is analogous to $\mathrm{P}_a(k)$ that is present in the work of Hooley [35, §3]. Indeed, the inclusion-exclusion argument above is inspired by the argument leading to [35, Eq.(5)].

Using the arguments in [35, §4] we shall first translate the $R_{a_i}(q,p_i)$-condition present in (7.42) into a condition related to the factorisation properties of the prime p_i in certain number fields. Recall the definition of h_a given in (7.5). For any positive square-free integer k_i we define $k_i' := k_i/\gcd(k_i, h_{a_i})$. Then, as explained in [35, Eq.(8)], for a prime $p \nmid a_i$ and a square-free integer k_i, the conditions $R_{a_i}(q,p)$ hold for all $q \mid k_i$ if and only if

$$x^{k_i'} \equiv a_i \,(\mathrm{mod}\ p) \text{ is soluble} \quad \text{and} \quad p \equiv 1 \,(\mathrm{mod}\ k_i)\,.$$

It is then proved following [35, Eq.(8)] that, in light of the Kummer–Dedekind theorem, this is in turn equivalent to the property that p is completely split in the number field $\mathbb{Q}(a_i^{1/k_i'}, \zeta_{k_i})$. Recall (7.3) and let us see why

$$G_{a_i,k_i} = \mathbb{Q}(a_i^{1/k_i'}, \zeta_{k_i}).$$

It is clearly sufficient to show that $a_i^{1/k_i} \in \mathbb{Q}(a_i^{1/k_i'}, \zeta_{k_i})$. Writing $a_i = b^{h_{a_i}}$ and using $\mu(k_i)^2 = 1$, we see that $\gcd(h_{a_i}\gcd(k_i,h_{a_i}), k_i) | h_{a_i}$, hence there are integers x, y with

$$h_{a_i}\gcd(k_i,h_{a_i})x + k_i y = h_{a_i}.$$

This leads to the equality $a_i^{1/k_i} = (b^{1/k_i})^{h_{a_i}} = b^y(a_i^{1/k_i'})^x$, which completes the argument.

Recalling the definition of $\mathrm{Spl}\,(G_{a_i,k_i})$ in (7.24), we infer by (7.42) that for all $\mathbf{k} \in \mathbb{N}^3$ with each k_i square-free we have

$$\mathrm{P}_{\mathbf{a},\mathbf{k}}(n) = \sum_{\substack{p_1+p_2+p_3=n,\ p_i \nmid a_i \\ \forall i:\ p_i \in \mathrm{Spl}(G_{a_i,k_i})}} \prod_{i=1}^{3} \log p_i = V_{\mathbf{a},\mathbf{k}}(n) + O_\beta(n^2((\log n)/n)^\beta),$$

for any $\beta \in (0,1)$. For the second equality, recall (7.25) and use Lemma 7.3.1. Injecting this into (7.43) we have proved that whenever $1 < \xi_1 \leq (\log n)(\log\log n)^{-1}$ and $0 < \beta < 1$ then

$$\sum_{\substack{p_1+p_2+p_3=n \\ \forall i:\ \mathbb{F}_{p_i}^* = \langle a_i \rangle}} \prod_{i=1}^{3} \log p_i = \sum_{\substack{\mathbf{k}\in\mathbb{N}^3 \\ p|k_1k_2k_3 \Rightarrow p\leq\xi_1}} \mu(k_1)\mu(k_2)\mu(k_3)V_{\mathbf{a},\mathbf{k}}(n) + O_\beta\Big(n^2\xi_1^{-\beta}\Big), \tag{7.44}$$

where, for $2-\beta<\delta<2$, the estimate

$$\begin{aligned}\sum_{\substack{\mathbf{k}\in\mathbb{N}^3 \\ p|k_1k_2k_3 \Rightarrow p\leq\xi_1}} |\mu(k_1)\mu(k_2)\mu(k_3)|n^\delta &\leq n^\delta\Big(\sum_{\substack{k\in\mathbb{N} \\ p|k\Rightarrow p\leq\xi_1}} |\mu(k)|\Big)^3 = n^\delta 2^{3\#\{p\leq\xi_1\}} \\ &\leq n^\delta \mathrm{e}^{3\xi_1} \leq n^{\delta+\frac{3}{\log\log n}} \\ &\ll_{\beta,\delta} n^2(\log n)^{-\beta}(\log\log n)^\beta \leq n^2\xi_1^{-\beta}\end{aligned}$$

Before concluding the proofs of Theorem 7.1.1 and Theorem 7.1.3, we need a preparatory lemma.

Lemma 7.3.2. *The series defining $\mathfrak{S}_{\mathbf{a},\mathbf{k}}(n)$ in* (7.18) *and representing $\mathcal{A}_{\mathbf{a}}(n)$ in* (7.19) *are absolutely convergent. For each $\epsilon>0$ and $z\geq 1$ we have*

$$\begin{aligned}\sum_{\substack{\mathbf{k}\in\mathbb{N}^3 \\ \exists i,p:\ p|k_i \text{ and } p\geq z}} |\mathfrak{S}_{\mathbf{a},\mathbf{k}}(n)|\Big(\prod_{i=1}^{3}|\mu(k_i)|\Big) &\leq \sum_{\substack{\mathbf{k}\in\mathbb{N}^3 \\ \exists i:\ k_i\geq z}} \Big(\prod_{i=1}^{3}|\mu(k_i)|\Big)\sum_{q=1}^{\infty}\frac{1}{d_{\mathbf{a},\mathbf{k}}(q)}\sum_{x\in(\mathbb{Z}/q\mathbb{Z})^*}|L_{\mathbf{a},q,\mathbf{k}}(x)| \\ &\ll_\epsilon \frac{1}{z^{1-\epsilon}},\end{aligned}$$

with an implied constant depending at most on $\mathbf{a}$ and ϵ.

Proof. The first inequality is clear by (7.18). Observe that $k_i' \geq k_i/h_{a_i} \gg k_i$, hence by Lemma 7.2.2 we obtain

$$\frac{1}{d_{\mathbf{a},\mathbf{k}}(q)} \ll \prod_{i=1}^{3}\frac{1}{k_i\varphi([q,k_i])} = \frac{1}{\varphi(q)^3}\prod_{i=1}^{3}\frac{\varphi(\gcd(q,k_i))}{k_i\varphi(k_i)}.$$

Combining this with (7.34) we see by (7.18) that for $\epsilon>0$ and square-free k_i,

$$\begin{aligned}\sum_{q=1}^{\infty}\frac{1}{d_{\mathbf{a},\mathbf{k}}(q)}\sum_{x\in(\mathbb{Z}/q\mathbb{Z})^*}|L_{\mathbf{a},q,\mathbf{k}}(x)| &\ll \prod_{i=1}^{3}\frac{1}{k_i\varphi(k_i)}\sum_{q=1}^{\infty}\frac{\varphi(\gcd(q,k_1))\varphi(\gcd(q,k_2))\varphi(\gcd(q,k_3))}{\varphi(q)^2} \\ &\ll_\epsilon \frac{\gcd(k_1,k_2,k_3)}{(k_1k_2k_3)^{2-\epsilon}}.\end{aligned}$$

Therefore, the inner sum our lemma is

$$\ll \sum_{k_1\geq z}\frac{|\mu(k_1)|}{k_1^{2-\epsilon}}\sum_{k_2\in\mathbb{N}}\frac{|\mu(k_2)|}{k_2^{2-\epsilon}}\sum_{k_3\in\mathbb{N}}\frac{|\mu(k_3)|\gcd(k_1,k_2,k_3)}{k_3^{2-\epsilon}}.$$

Using the estimates

$$\sum_{k_3\in\mathbb{N}} |\mu(k_3)| \gcd(k_3,m) k_3^{-2+\epsilon} \ll_\epsilon m^\epsilon \quad \text{and} \quad \sum_{k_1\geq z} \frac{|\mu(k_1)|}{k_1^{2-\epsilon}} \ll z^{-1+\epsilon}$$

concludes our proof of the desired bound, which implies absolute convergence of the sum in (7.19). □

7.3.2 The proof of Theorem 7.1.1

Recall (7.26). Now note that, replacing $f_{\mathbf{a}}(\mathbf{x})$ by a larger function if necessary, we may assume in the statement of (7.26) that $f_{\mathbf{a}}([1,\infty)^3)$ is a subset of $(1,\infty)$. Fix any $B>0$. The function

$$x \mapsto \log(1+x) + \sum_{1\leq k_1,k_2,k_3\leq x} f_{\mathbf{a}}(\mathbf{k}),$$

is strictly increasing, hence it has an inverse, which we call $h_{\mathbf{a}}(x)$. Define the function $\xi_1:(1,\infty)\to\mathbb{R}$ through

$$\xi_1(x) := \frac{1}{2}\cdot\min\left\{\frac{\log x}{\log\log x},\ \log(h_{\mathbf{a}}((\log x)^{B/2}))\right\} \tag{7.45}$$

and observe that

$$\lim_{x\to+\infty} \xi_1(x) = +\infty, \tag{7.46}$$

however, owing to the non-explicit error term in [38, Th.2] we cannot have any further control on the rate of divergence in the last limit. For $n \gg 1$, the definition of ξ_1 implies

$$\sum_{1\leq k_1,k_2,k_3\leq \mathrm{e}^{2\xi_1(n)}} f_{\mathbf{a}}(\mathbf{k}) \leq (\log n)^{B/2}.$$

Noting that a square-free integer with all of its prime factors bounded by $\xi_1(n)$ must be at most $\prod_{p\leq\xi_1(n)} p \leq \exp(2\xi_1(n))$ and injecting (7.26) into (7.44) yields the following with an implied constant depending on β and B,

$$\begin{aligned}\sum_{\substack{p_1+p_2+p_3=n\\ \forall i:\ \mathbb{F}_{p_i}^*=\langle a_i\rangle}} \prod_{i=1}^{3}\log p_i &= \frac{n^2}{2}\sum_{\substack{\mathbf{k}\in\mathbb{N}^3\\ p\mid k_1k_2k_3\Rightarrow p\leq\xi_1(n)}} \Big(\prod_{i=1}^{3}\mu(k_i)\Big)\mathfrak{S}_{\mathbf{a},\mathbf{k}}(n) + \\ &\qquad O\left(\frac{n^2}{\xi_1^\beta} + \frac{n^2}{(\log n)^B}\Big(\sum_{\substack{\mathbf{k}\in\mathbb{N}^3\\ \forall i:\ k_i\leq \mathrm{e}^{2\xi_1(n)}}} f_{\mathbf{a}}(\mathbf{k})\Big)\right) \\ &= \frac{n^2}{2}\sum_{\substack{\mathbf{k}\in\mathbb{N}^3\\ p\mid k_1k_2k_3\Rightarrow p\leq\xi_1(n)}} \Big(\prod_{i=1}^{3}\mu(k_i)\Big)\mathfrak{S}_{\mathbf{a},\mathbf{k}}(n) + O\left(\frac{n^2}{\xi_1^\beta} + \frac{n^2}{(\log n)^{B/2}}\right).\end{aligned}$$

An application of Lemma 7.3.2 with $\epsilon = 1 - \beta$ shows that

$$\sum_{\substack{p_1+p_2+p_3=n \\ \forall i:\ \mathbb{F}_{p_i}^* = \langle a_i \rangle}} \prod_{i=1}^{3} \log p_i - \frac{1}{2}\Big(\sum_{\mathbf{k}\in\mathbb{N}^3} \mu(k_1)\mu(k_2)\mu(k_3)\mathfrak{S}_{\mathbf{a},\mathbf{k}}(n)\Big) n^2$$
$$\ll_{\beta,B} \frac{n^2}{\min\{(\log n)^{B/2}, \xi_1(n)^\beta\}},$$

and the proof of Theorem 7.1.1 is concluded upon invoking (7.46), up to the assertion that $\mathcal{A}_{\mathbf{a}}(n) \gg_a 1$ whenever $\mathcal{A}_{\mathbf{a}}(n) > 0$. This follows immediately from Theorem 7.1.5, proved in §7.4. Moreover, we have confirmed the shape of $\mathcal{A}_{\mathbf{a}}(n)$ given in (7.19). $\square$

Note that the reason for the non-explicit error term in Theorem 7.1.1 is that the function ξ_1 in (7.45) is not explicit.

7.3.3 The proof of Theorem 7.1.3

Let β be any real number in $(0,1)$ and define

$$\xi_1(n) := \frac{\log n}{\log\log n}.$$

Injecting Proposition 7.2.1 into (7.44) provides us with

$$\sum_{\substack{p_1+p_2+p_3=n \\ \forall i:\ \mathbb{F}_{p_i}^* = \langle a_i \rangle}} \prod_{i=1}^{3} \log p_i - \frac{n^2}{2} \sum_{p \mid k_1k_2k_3 \Rightarrow p \leq \xi_1} \mathfrak{S}_{\mathbf{a},\mathbf{k}}(n) \prod_{i=1}^{3} \mu(k_i)$$
$$\ll_\beta \frac{n^2}{\xi_1^\beta} + \frac{(\log n)^6}{n^{-11/6}} \Big(\sum_{\substack{k\in\mathbb{N} \\ p\mid k \Rightarrow p\leq \xi_1}} k^6 |\mu(k)| \Big)^3.$$

For $n \gg 1$, each k in the sum satisfies $k \leq \prod_{p\leq\xi_1} p \leq n^{\frac{2}{\log\log n}}$, hence the cube of the sum over k is at most $n^{\frac{\theta}{\log\log n}}$ for some absolute positive constant θ. This shows that the right side above is $\ll_\beta n^2 \xi_1^{-\beta}$. Appealing to Lemma 7.3.2 completes the proof of Theorem 7.1.3. $\square$

7.4 Artin's factor for ternary Goldbach

In this section, we study in detail the leading factor $\mathcal{A}_{\mathbf{a}}(n)$ in Theorems 7.1.1 and 7.1.3, and thus prove Theorem 7.1.5, Corollary 7.1.6 and Theorem 7.1.7. Recall that we have already confirmed the equality (7.19) in the proof of Theorem 7.1.1 in §7.3.2.

7.4.1 The proof of (7.22)

Recall the definitions of $F_{a,q,k}(b)$ and $c_{a,q,k}(b)$ from the start of §7.1.3. It was shown by Lenstra [51, Th.(3.1),Eq.(2.15)] conditionally under HRH(a), that for all integers b and $q>0$ the Dirichlet density of the primes p satisfying the following conditions exists,

$$\mathbb{F}_p^* = \langle a \rangle \text{ and } p \equiv b \,(\mathrm{mod}\ q)\,,$$

and, furthermore, that it equals $\sum_{k\in\mathbb{N}} \mu(k) c_{a,q,k}(b)[F_{a,q,k} : \mathbb{Q}]^{-1}$. This topic was later revisited by Moree [59], who showed that

$$\sum_{k\in\mathbb{N}} \frac{\mu(k)c_{a,q,k}(b)}{[F_{a,q,k}:\mathbb{Q}]} = \delta_a(b \,\mathrm{mod}\ q), \tag{7.47}$$

where $\delta_a(b \,\mathrm{mod}\ q)$ is the arithmetic function given in Definition 7.1.4. We will make consistent use of Moree's result in this section.

Lemma 7.4.1. *We have*

$$\sum_{\mathbf{k}\in\mathbb{N}^3} \mu(k_1)\mu(k_2)\mu(k_3)\mathfrak{S}_{\mathbf{a},\mathbf{k}}(n) = \sum_{q=1}^{\infty} \sum_{c\in(\mathbb{Z}/q\mathbb{Z})^*} \mathrm{e}_q(-nc) \prod_{i=1}^{3} \left(\sum_{b_i\in\mathbb{Z}/q\mathbb{Z}} \mathrm{e}_q(b_ic)\delta_{a_i}(b_i \,\mathrm{mod}\ q) \right).$$

Proof. Recall (7.15) and (7.18). Lemma 7.3.2 allows us to rearrange terms, thus we can rewrite the sum over $\mathbf{k}$ in our lemma as

$$\sum_{q=1}^{\infty} \sum_{\substack{c\in\mathbb{Z}/q\mathbb{Z}\\ \gcd(c,q)=1}} \mathrm{e}_q(-cn) \prod_{i=1}^{3} \left(\sum_{k_i\in\mathbb{N}} \frac{\mu(k_i)S_{a_i,q,k_i}(c)}{[F_{a_i,q,k_i}:\mathbb{Q}]} \right).$$

By (7.15) the sum over k_i equals

$$\sum_{\substack{b_i\in\mathbb{Z}/q\mathbb{Z}\\ \gcd(b_i,q)=1}} \mathrm{e}_q(b_ic) \sum_{k_i\in\mathbb{N}} \frac{\mu(k_i)c_{a_i,q,k_i}(b_i)}{[F_{a_i,q,k_i}:\mathbb{Q}]}$$

and using (7.47) concludes our proof. □

The difficulty of converting the sum over $\mathbf{k}$ in (7.19) into a product comes from the fact that the terms $\delta_{a_i}(b_i \,\mathrm{mod}\ q)$ in Lemma 7.4.1 are not a multiplicative function of q. These terms would be multiplicative in the classical Vinogradov setting, where one has $\mathbf{1}_{\gcd(b_i,q)=1}(b_i)/\phi(q)$ in place of $\delta_{a_i}(b_i \,\mathrm{mod}\ q)$.

For brevity, we will write from now on $\beta_i(q)$ and Δ_i for $\beta_{a_i}(q)$ and Δ_{a_i}.

Lemma 7.4.2. *If the odd part of a positive integer q is not square-free then the following expression vanishes,*

$$\prod_{i=1}^{3} \left(\sum_{b_i\in\mathbb{Z}/q\mathbb{Z}} \mathrm{e}_q(b_ic)\delta_{a_i}(b_i \,\mathrm{mod}\ q) \right).$$

Furthermore, the expression vanishes if $\nu_2(q) > \min\{\nu_2(\Delta_i) : i=1,2,3\}$.

Proof. In the present proof we write $[P] := 1$ if a proposition P holds, and $[P] := 0$ otherwise. For $1 \leq i \leq 3$, we factorise each positive integer q as $q = q_{i,0}q_{i,1}$, where the positive integers $q_{i,0}, q_{i,1}$ are uniquely defined through the conditions $p \mid q_{i,0} \Rightarrow p|\Delta_i$ and $\gcd(q_{i,1}, \Delta_i) = 1$. Now owing to Definition 7.1.4 the quantity $\delta_{a_i}(b_i \bmod q)/\mathcal{A}_{a_i}$ equals

$$\left(\left[\gcd(b_i, q_{i,1}) \gcd(b_i - 1, q_{i,1}, h_{a_i}) = 1\right] \frac{f^\dagger_{a_i}(q_{i,1})}{\phi(q_{i,1})} \prod_{p|b_i-1, p|q_{i,1}} \left(1 - \frac{1}{p}\right)\right) \times$$
$$\left(\frac{f^\dagger_i(q_{i,0})}{\phi(q_{i,0})} \prod_{p|b_i-1, p|q_{i,0}} \left(1 - \frac{1}{p}\right)\right) \times \left[\gcd(b_i, q_{i,0}) \gcd(b_i - 1, q_{i,0}, h_{a_i}) = 1\right] \times$$
$$\left(1 + \left(\frac{\beta_i(q_{i,0})}{b_i}\right) \mu\left(\frac{2|\Delta_i|}{\gcd(q_{i,0}, \Delta_i)}\right) f^\ddagger_{a_i}\left(\frac{|\Delta_i|}{\gcd(q_{i,0}, \Delta_i)}\right)\right).$$

The integers $q_{i,0}$ and $q_{i,1}$ are coprime, hence we may write $b_i = q_{i,0}b_{i,1} + q_{i,1}b_{i,0}$ and use the Chinese remainder theorem to write the sum over b_i in the lemma as the product of

$$\mathcal{A}_{a_i} \cdot \frac{f^\dagger_{a_i}(q_{i,0})}{\phi(q_{i,0})} \frac{f^\dagger_{a_i}(q_{i,1})}{\phi(q_{i,1})} \sum_{\substack{b_{i,1} (\mathrm{mod}\ q_{i,1}) \\ \gcd(b_{i,1}, q_{i,1})=1 \\ \gcd(b_{i,1}q_{i,0}-1, q_{i,1}, h_{a_i})=1}} \mathrm{e}(b_{i,1}c/q_{i,1}) \prod_{p|(b_{i,1}q_{i,0}-1, q_{i,1})} \left(1 - \frac{1}{p}\right)$$

and

$$\sum_{\substack{b_{i,0} (\mathrm{mod}\ q_{i,0}) \\ \gcd(b_{i,0}, q_{i,0})=1 \\ \gcd(b_{i,0}q_{i,1}-1, q_{i,0}, h_{a_i})=1}} \frac{\mathrm{e}(b_{i,0}c/q_{i,0})}{\prod_{p|(b_{i,0}q_{i,1}-1, q_{i,0})}(1 - \frac{1}{p})^{-1}} \times$$
$$\left(1 + \left(\frac{\beta_i(q_{i,0})}{b_{i,0}q_{i,1}}\right) \mu\left(\frac{2|\Delta_i|}{\gcd(q_{i,0}, \Delta_i)}\right) f^\ddagger_{a_i}\left(\frac{|\Delta_i|}{\gcd(q_{i,0}, \Delta_i)}\right)\right).$$

To study the sum over $b_{i,1}$ we use Lemma 7.2.13 with

$$Q := q_{i,1}, \quad r := \prod_{p|q_{i,1}} p, \quad f(b) := [\gcd(b, r) \gcd(b - 1, r, h_{a_i}) = 1] \prod_{p|b-1, p|r} \left(1 - \frac{1}{p}\right)$$

to deduce that if the expression in our lemma is non-vanishing then for each i the integer $q_{i,1}$ must be square-free. Now assume that the prime p satisfies $p \nmid \gcd(\Delta_1, \Delta_2, \Delta_3)$. Then there exists $i \in \{1, 2, 3\}$ such that $p \nmid \Delta_i$ and then the non-vanishing of the expression in the lemma implies that $q_{i,1}$ must be square-free, thus $\nu_p(q) = \nu_p(q_{i,1}) \leq 1$.

Now the sum over $b_{i,0}$ can be studied via Lemma 7.2.13 with $Q := q_{i,0}$, $r := \gcd(q_{i,0}, \Delta_i)$ and with $f(b)$ being the product of $[\gcd(b, r) \gcd(bq_{i,1} - 1, r, h_{a_i}) = 1]$ and

$$\left\{1 + \left(\frac{\beta(q_{i,0})}{b}\right) \mu\left(\frac{2|\Delta_i|}{\gcd(q_{i,0}, \Delta_i)}\right) f^\ddagger_i\left(\frac{|\Delta_i|}{\gcd(q_{i,0}, \Delta_i)}\right)\right\} \prod_{p|(bq_{i,1}-1, r)} \left(1 - \frac{1}{p}\right).$$

We have used the fact that $p \mid q_{i,0} \Leftrightarrow p \mid r$ and that the Kronecker symbol has period $|\beta(q_{i,0})| = r$. Lemma 7.2.13 shows that unless the expression in our lemma vanishes, we have $\gcd(q_{i,0}, \Delta_i) = q_{i,0}$, thus for every i we must have $q_{i,0} \mid \Delta_i$. Now if a prime p satisfies $p \mid \gcd(\Delta_1, \Delta_2, \Delta_3)$ we have that for every i, $\nu_p(q) = \nu_p(q_{i,0}) \leq \nu_p(\Delta_i)$, thus $\nu_p(q) \leq \min\{\nu_p(\Delta_i) : i = 1, 2, 3\}$. If $p \neq 2$ then this shows that $\nu_p(q) \leq 1$ since the odd part of a fundamental discriminant is square-free, while if $p = 2$ then we must have $\nu_2(q) \leq \min\{\nu_2(\Delta_i) : i = 1, 2, 3\}$. $\square$

Lemma 7.4.2 allows us to simplify the summation over q in Lemma 7.4.1 since the only integers q making a contribution towards the sum must satisfy

$$\forall p, i: \ p|\Delta_i, p|q \Rightarrow \nu_p(q) \leq \nu_p(\Delta_i) \quad \text{and} \quad p|q, p \nmid \Delta_1\Delta_2\Delta_3 \Rightarrow \nu_p(q) \leq 1.$$

To keep track of every factorisation we introduce for every $q \in \mathbb{N}$ and $\mathbf{w} \in \{0, 1\}^3$ the positive integer

$$q(\mathbf{w}) := \prod_{\substack{p: \\ \forall i: \ p|\Delta_i \Leftrightarrow \mathbf{w}(i)=0}} p^{\nu_p(q)}$$

so that $q = \prod_{\mathbf{w} \in \mathbb{F}_2^3} q(\mathbf{w})$. Furthermore, $\mathbf{w} \neq \mathbf{u}$ implies $\gcd(q(\mathbf{w}), q(\mathbf{u})) = 1$. Note that for a given q, $q(\mathbf{w})$ is uniquely characterised by the properties

$$\gcd(q(\mathbf{w}), \prod_{i:\mathbf{w}(i)=1} \Delta_i) = 1 \quad \text{and} \quad q(\mathbf{w}) \mid \gcd\{\Delta_i : \mathbf{w}(i) = 0\}. \tag{7.48}$$

In the case $\mathbf{w} = (1, 1, 1)$, the latter condition is interpreted as vacuous. It may be that for certain values of a_i and for all q some $q(\mathbf{w})$ are equal to 1; for example, this happens if $a_1 = a_2 = a_3$, in which case we have $\mathbf{w} \notin \{(0, 0, 0), (1, 1, 1)\} \Rightarrow q(\mathbf{w}) = 1$. We now use the definition of $q(\mathbf{w})$, Lemma 7.4.1 and Lemma 7.4.2 to infer

$$\sum_{\mathbf{k} \in \mathbb{N}^3} \mu(k_1)\mu(k_2)\mu(k_3)\mathfrak{S}_{\mathbf{a},\mathbf{k}}(n) = \sum_{\substack{(q(\mathbf{w})) \in \mathbb{N}^8, \\ (7.48) \text{ holds} \\ \mu(q((1,1,1)))^2=1}} \sum_{\substack{c \left(\mathrm{mod} \ \prod_{\mathbf{w}} q(\mathbf{w})\right) \\ \gcd(c, \prod_{\mathbf{w}} q(\mathbf{w}))=1}} \mathrm{e}(-nc \prod_{\mathbf{w}} q(\mathbf{w})^{-1}) \times$$

$$\prod_{i=1}^{3} \left(\sum_{b_i \left(\mathrm{mod} \ \prod_{\mathbf{w}} q(\mathbf{w})\right)} \mathrm{e}\Big(b_i c \prod_{\mathbf{w}} q(\mathbf{w})^{-1}\Big) \delta_{a_i}\Big(b_i \bmod \prod_{\mathbf{w}} q(\mathbf{w})\Big) \right). \tag{7.49}$$

Noting that the integers $\prod_{\mathbf{w}(i)=0} q(\mathbf{w})$ and $\prod_{\mathbf{w}(i)=1} q(\mathbf{w})$ are coprime, that

$$\gcd\Big(\Delta_i, \prod_{\mathbf{w}} q(\mathbf{w})\Big) = \prod_{\mathbf{w}(i)=0} q(\mathbf{w})$$

and recalling Definition 7.1.4 we see that

$$\delta_{a_i}\Big(b_i \bmod \prod_{\mathbf{w}} q(\mathbf{w})\Big) = \delta_{a_i}\Big(b_i \bmod \prod_{\mathbf{w}(i)=0} q(\mathbf{w})\Big) \mathcal{A}_{a_i}\Big(b_i \bmod \prod_{\mathbf{w}(i)=1} q(\mathbf{w})\Big) \mathcal{A}_{a_i}^{-1}.$$

Writing $b_i = b_i' \prod_{\mathbf{w}(i)=1} q(\mathbf{w}) + b_i'' \prod_{\mathbf{w}(i)=0} q(\mathbf{w})$ and using the Chinese remainder theorem we obtain

$$\begin{aligned}
&\sum_{b_i \left(\bmod\ \prod_{\mathbf{w}} q(\mathbf{w})\right)} \mathrm{e}\Big(b_i c \prod_{\mathbf{w}} q(\mathbf{w})^{-1}\Big) \delta_{a_i}\Big(b_i \bmod \prod_{\mathbf{w}} q(\mathbf{w})\Big) \\
&= \sum_{b_i' \left(\bmod\ \prod_{\mathbf{w}(i)=0} q(\mathbf{w})\right)} \mathrm{e}\Big(b_i' c \prod_{\mathbf{w}(i)=0} q(\mathbf{w})^{-1}\Big) \delta_{a_i}\Big(b_i' \prod_{\mathbf{w}(i)=1} q(\mathbf{w}) \bmod \prod_{\mathbf{w}(i)=0} q(\mathbf{w})\Big) \times \\
&\times \sum_{b_i'' \left(\bmod\ \prod_{\mathbf{w}(i)=1} q(\mathbf{w})\right)} \mathrm{e}\Big(b_i'' c \prod_{\mathbf{w}(i)=1} q(\mathbf{w})^{-1}\Big) \mathcal{A}_{a_i}^{-1} \mathcal{A}_{a_i}\Big(b_i'' \prod_{\mathbf{w}(i)=0} q(\mathbf{w}) \bmod \prod_{\mathbf{w}(i)=1} q(\mathbf{w})\Big).
\end{aligned}$$

For the further analysis of the expressions above, we introduce for $r \in \mathbb{N}$, $c \in \mathbb{Z}$ the quantity

$$\mathcal{M}_a(c, r) := \frac{1}{\mathcal{A}_a} \sum_{b (\bmod\ r)} \mathrm{e}_r(bc) \mathcal{A}_a(b \bmod r), \tag{7.50}$$

and for $\mathbf{r} \in \mathbb{N}^k$, $\mathbf{c} \in \mathbb{Z}^k$ define

$$\mathcal{D}_a(\mathbf{c}, \mathbf{r}) := \sum_{b (\bmod\ r_1 \cdots r_k)} \mathrm{e}\Big[b\Big(\sum_{i=1}^{r} \frac{c_i}{r_i}\Big)\Big] \delta_a(b \bmod r_1 \cdots r_k).$$

Hence, writing

$$c = \sum_{\mathbf{w} \in \{0,1\}^3} c^{[\mathbf{w}]} \prod_{\mathbf{v} \neq \mathbf{w}} q(\mathbf{v}),$$

we see that $\prod_{\mathbf{w}(i)=1} \mathcal{M}_{a_i}(c^{[\mathbf{w}]}, q(\mathbf{w}))$ equals

$$\mathcal{A}_{a_i}^{-1} \sum_{b_i'' \left(\bmod\ \prod_{\mathbf{w}(i)=1} q(\mathbf{w})\right)} \mathrm{e}\Big(b_i'' c \prod_{\mathbf{w}(i)=1} q(\mathbf{w})^{-1}\Big) \mathcal{A}_{a_i}\Big(b_i'' \prod_{\mathbf{w}(i)=0} q(\mathbf{w}) \bmod \prod_{\mathbf{w}(i)=1} q(\mathbf{w})\Big)$$

and that $\mathcal{D}_{a_i}((c^{[\mathbf{w}]})_{\mathbf{w}(i)=0}, (q(\mathbf{w}))_{\mathbf{w}(i)=0})$ is

$$\sum_{b_i' \left(\bmod\ \prod_{\mathbf{w}(i)=0} q(\mathbf{w})\right)} \mathrm{e}\Big(b_i' c \prod_{\mathbf{w}(i)=0} q(\mathbf{w})^{-1}\Big) \delta_{a_i}\Big(b_i' \prod_{\mathbf{w}(i)=1} q(\mathbf{w}) \bmod \prod_{\mathbf{w}(i)=0} q(\mathbf{w})\Big).$$

Let us bring into play the entities

$$\Delta_{\mathbf{w}} := \prod_{p \nmid \prod_{\mathbf{w}(i)=1} \Delta_i} p^{\min\{\nu_p(\Delta_i)\ :\ \mathbf{w}(i)=0\}},$$

which we interpret as 1 in case $\mathbf{w} = (1,1,1)$, and note that $\prod_{\mathbf{w}} \Delta_{\mathbf{w}}$ coincides with the entity $\mathfrak{D}_{\mathbf{a}}$ introduced in (7.21). We see that the sum in (7.49) becomes

$$\sum_{\substack{(q(\mathbf{w}))\in\mathbb{N}^8\\ \mathbf{w}\neq(1,1,1)\Rightarrow q(\mathbf{w})|\Delta_{\mathbf{w}}\\ \mu(q((1,1,1)))^2=1\\ \gcd(q((1,1,1)),\Delta_1\Delta_2\Delta_3)=1}} \sum_{(c^{[\mathbf{w}]})\in\prod_{\mathbf{w}}(\mathbb{Z}/q(\mathbf{w})\mathbb{Z})^*} \Big(\prod_{\mathbf{w}} \mathrm{e}_{q(\mathbf{w})}(-nc^{[\mathbf{w}]})\Big)\times$$

$$\times \prod_{i=1}^{3}\Big\{\mathcal{D}_{a_i}((c^{[\mathbf{w}]})_{\mathbf{w}(i)=0},(q(\mathbf{w}))_{\mathbf{w}(i)=0})\prod_{\mathbf{w}(i)=1}\mathcal{M}_{a_i}(c^{[\mathbf{w}]},q(\mathbf{w}))\Big\}.$$

Clearly, the terms corresponding to $q((1,1,1))$ can be separated, thus, in light of (7.49), we are led to

$$\sum_{\mathbf{k}\in\mathbb{N}^3}\mu(k_1)\mu(k_2)\mu(k_3)\mathfrak{S}_{\mathbf{a},\mathbf{k}}(n) = S_{\mathbf{a},0}(n)S_{\mathbf{a},1}(n), \tag{7.51}$$

where

$$S_{\mathbf{a},0}(n) := \sum_{\substack{(q(\mathbf{w}))_{\mathbf{w}\neq(1,1,1)}\in\mathbb{N}^7\\ q(\mathbf{w})|\Delta_{\mathbf{w}}}} \sum_{(c^{[\mathbf{w}]})\in\prod_{\mathbf{w}\neq(1,1,1)}(\mathbb{Z}/q(\mathbf{w})\mathbb{Z})^*} \Big(\prod_{\mathbf{w}\neq(1,1,1)} \mathrm{e}_{q(\mathbf{w})}(-nc^{[\mathbf{w}]})\Big)\times$$

$$\prod_{i=1}^{3}\Big\{\mathcal{D}_{a_i}((c^{[\mathbf{w}]})_{\mathbf{w}(i)=0},(q(\mathbf{w}))_{\mathbf{w}(i)=0})\prod_{\substack{\mathbf{w}(i)=1\\ \mathbf{w}\neq(1,1,1)}}\mathcal{M}_{a_i}(c^{[\mathbf{w}]},q(\mathbf{w}))\Big\}$$

and

$$S_{\mathbf{a},1}(n) := \sum_{\gcd(q((1,1,1)),\Delta_1\Delta_2\Delta_3)=1} \mu(q((1,1,1)))^2\times$$

$$\sum_{c^{[(1,1,1)]}\in(\mathbb{Z}/q((1,1,1))\mathbb{Z})^*} \mathrm{e}_{q((1,1,1))}(-nc^{[(1,1,1)]})\prod_{i=1}^{3}\mathcal{M}_{a_i}(c^{[(1,1,1)]},q((1,1,1))). \tag{7.52}$$

Lemma 7.4.3. *For any $q\in\mathbb{N}$ and $\mathbf{w}\in\{0,1\}^3$ define $d_{\mathbf{w}} := \Delta_{\mathbf{w}}/q(\mathbf{w})$.*

1. *Let $i\in\{1,2,3\}$ and for each $\mathbf{w}$ with $\mathbf{w}(i)=0$ let $c^{[\mathbf{w}]}\in(\mathbb{Z}/q(\mathbf{w})\mathbb{Z})^*$. Then*
$$\mathcal{D}_{a_i}((c^{[\mathbf{w}]})_{\mathbf{w}(i)=0},(q(\mathbf{w}))_{\mathbf{w}(i)=0}) = \mathcal{D}_{a_i}((c^{[\mathbf{w}]}d_{\mathbf{w}})_{\mathbf{w}(i)=0},(\Delta_{\mathbf{w}})_{\mathbf{w}(i)=0}).$$

2. *Let $i\in\{1,2,3\}$, $\mathbf{w}\in\{0,1\}^3\setminus\{(1,1,1)\}$ with $\mathbf{w}(i)=1$ and $c^{[\mathbf{w}]}\in(\mathbb{Z}/q(\mathbf{w})\mathbb{Z})^*$. Then*
$$\mathcal{M}_{a_i}(c^{[\mathbf{w}]},q(\mathbf{w})) = \mathcal{M}_{a_i}(c^{[\mathbf{w}]}d_{\mathbf{w}},\Delta_{\mathbf{w}}).$$

Proof. (1): Define

$$Q := \prod_{\mathbf{w}:\mathbf{w}(i)=0} q(\mathbf{w}) = \prod_{\mathbf{w}:\mathbf{w}(i)=0}\frac{\Delta_{\mathbf{w}}}{d_{\mathbf{w}}} \quad\text{and}\quad D := \prod_{\mathbf{w}:\mathbf{w}(i)=0}\Delta_{\mathbf{w}}.$$

If we assume HRH(a_i) then it is immediately clear from Moree's interpretation of δ_{a_i} as Dirichlet densities [59] that the following holds,

$$\delta_{a_i}\big(m \bmod Q\big) = \sum_{\substack{b(\bmod\ D)\\ b\equiv m(\bmod\ Q)}} \delta_{a_i}\big(b \bmod D\big).$$

One can also prove this unconditionally directly from Definition 7.1.4 via a tedious but straightforward calculation that we do not reproduce here. To conclude the proof we observe that

$$\begin{aligned}\mathcal{D}_{a_i}((c^{[\mathbf{w}]})_{\mathbf{w}(i)=0},(q(\mathbf{w}))_{\mathbf{w}(i)=0}) &= \sum_{m(\bmod\ Q)} \mathrm{e}\bigg(m \sum_{\mathbf{w}:\mathbf{w}(i)=0} \frac{c^{[\mathbf{w}]}}{q(\mathbf{w})}\bigg)\delta_{a_i}(m \bmod Q)\\ &= \sum_{b(\bmod\ D)} e\bigg(b \sum_{\mathbf{w}:\mathbf{w}(i)=0} \frac{c^{[\mathbf{w}]}d_{\mathbf{w}}}{\Delta_{\mathbf{w}}}\bigg)\delta_{a_i}(b \bmod D)\\ &= \mathcal{D}_{a_i}((c^{[\mathbf{w}]}d_{\mathbf{w}})_{\mathbf{w}(i)=0},(\Delta_{\mathbf{w}})_{\mathbf{w}(i)=0}).\end{aligned}$$

(2): Due to the assumption that $\mathbf{w}(i)=1$ we have $\gcd(\Delta_{\mathbf{w}},\Delta_i)=1$, and thus,

$$\frac{\mathcal{A}_{a_i}(m \bmod \Delta_{\mathbf{w}})}{\mathcal{A}_{a_i}} = \frac{\delta_{a_i}(m \bmod \Delta_{\mathbf{w}})}{\mathcal{L}_{a_i}}.$$

We similarly have

$$\frac{\mathcal{A}_{a_i}(m \bmod \Delta_{\mathbf{w}}/d_{\mathbf{w}})}{\mathcal{A}_{a_i}} = \frac{\delta_{a_i}(m \bmod \Delta_{\mathbf{w}}/d_{\mathbf{w}})}{\mathcal{L}_{a_i}}.$$

By HRH(a_i) it then follows that

$$\mathcal{A}_{a_i}(m \bmod \Delta_{\mathbf{w}}/d_{\mathbf{w}}) = \sum_{\substack{b(\bmod\ \Delta_{\mathbf{w}})\\ b\equiv m(\bmod\ \Delta_{\mathbf{w}}/d_{\mathbf{w}})}} \mathcal{A}_{a_i}(b \bmod \Delta_{\mathbf{w}}),$$

which can also be shown unconditionally as above. The rest of the proof is conducted as in the first part. □

For the analysis of $S_{\mathbf{a},1}(n)$, we recall the definition of $\sigma_{\mathbf{a},n}(d)$ in (7.20) and use the following lemma.

Lemma 7.4.4. *If $p \nmid \Delta_1\Delta_2\Delta_3$, then*

$$\sigma_{\mathbf{a},n}(p) = 1 + \sum_{c\in(\mathbb{Z}/p\mathbb{Z})^*} \mathrm{e}_p(-nc)\prod_{i=1}^{3}\mathcal{M}_{a_i}(c,p)$$

Proof. The easily verified equality $\sum_{b(\text{mod } p)} \mathcal{A}_{a_i}(b\text{mod } p) = \mathcal{A}_{a_i}$ shows that the expression on the right-hand side is equal to

$$\sum_{c\in\mathbb{Z}/p\mathbb{Z}} \mathrm{e}_p(-cn)\prod_{i=1}^{3}\mathcal{M}_{a_i}(c,p) = \sum_{\mathbf{b}\in(\mathbb{Z}/p\mathbb{Z})^3}\left(\prod_{i=1}^{3}\frac{\mathcal{A}_{a_i}(b_i \bmod p)}{\mathcal{A}_{a_i}}\right) \sum_{c\in\mathbb{Z}/p\mathbb{Z}} \mathrm{e}_p(c(b_1+b_2+b_3-n)),$$

which is in turn equal to

$$p\sum_{\substack{\mathbf{b}\in(\mathbb{Z}/p\mathbb{Z})^3\\ \sum_{i=1}^3 b_i\equiv n(\text{mod } p)}}\prod_{i=1}^{3}\frac{\mathcal{A}_{a_i}(b_i \bmod p)}{\mathcal{A}_{a_i}}.$$

Since $p\nmid\Delta_1\Delta_2\Delta_3$, we see that $\mathcal{A}_{a_i}(b_i \mod p)/\mathcal{A}_{a_i} = \delta_{a_i}(b_i \mod d)/\mathcal{L}_{a_i}$. □

Using (7.52), multiplicativity and Lemma 7.4.4, we infer that

$$S_{\mathbf{a},1}(n) = \prod_{p\nmid\Delta_1\Delta_2\Delta_3}\Big(1+\sum_{c\in(\mathbb{Z}/p\mathbb{Z})^*}\mathrm{e}_p(-nc)\prod_{i=1}^{3}\mathcal{M}_{a_i}(c,p)\Big) = \prod_{p\nmid\Delta_1\Delta_2\Delta_3}\sigma_{\mathbf{a},n}(p). \quad (7.53)$$

We now turn our attention to $S_{\mathbf{a},0}(n)$. Letting $d_{\mathbf{w}} := \Delta_{\mathbf{w}}/q(\mathbf{w})$ we use Lemma 7.4.3 to obtain

$$S_{\mathbf{a},0}(n) = \sum_{\substack{(d_{\mathbf{w}})_{\mathbf{w}\neq(1,1,1)}\in\mathbb{N}^7\\ d_{\mathbf{w}}|\Delta_{\mathbf{w}}}}\ \sum_{(c^{[\mathbf{w}]})\in\prod_{\mathbf{w}\neq(\mathbf{1},\mathbf{1},\mathbf{1})}\left(\frac{\mathbb{Z}}{(\Delta_{\mathbf{w}}/d_{\mathbf{w}})\mathbb{Z}}\right)^*}\Big(\prod_{\mathbf{w}\neq(1,1,1)}\mathrm{e}\Big(-nc^{[\mathbf{w}]}d_{\mathbf{w}}/\Delta_{\mathbf{w}}\Big)\Big)\times$$
$$\prod_{i=1}^{3}\Big\{\mathcal{D}_{a_i}((c^{[\mathbf{w}]}d_{\mathbf{w}})_{\mathbf{w}(i)=0},(\Delta_{\mathbf{w}})_{\mathbf{w}(i)=0})\prod_{\substack{\mathbf{w}(i)=1\\ \mathbf{w}\neq(1,1,1)}}\mathcal{M}_{a_i}(c^{[\mathbf{w}]}d_{\mathbf{w}},\Delta_{\mathbf{w}})\Big\}.$$

For any $d_{\mathbf{w}}$ with $d_{\mathbf{w}} \mid \Delta_{\mathbf{w}}$ the elements $y^{[\mathbf{w}]}\,(\text{mod } \Delta_{\mathbf{w}})$ that satisfy the condition $\gcd(y^{[\mathbf{w}]},\Delta_{\mathbf{w}}) = d_{\mathbf{w}}$ are exactly those of the form

$$y^{[\mathbf{w}]} = c^{[\mathbf{w}]}d_{\mathbf{w}}, \quad c^{[\mathbf{w}]}\in\Big(\frac{\mathbb{Z}}{(\Delta_{\mathbf{w}}/d_{\mathbf{w}})\mathbb{Z}}\Big)^*.$$

We thus obtain that the sum over $d_{\mathbf{w}}, c^{[\mathbf{w}]}$ equals

$$\sum_{(y^{[\mathbf{w}]})\in\prod_{\mathbf{w}\neq(\mathbf{1},\mathbf{1},\mathbf{1})}(\mathbb{Z}/\Delta_{\mathbf{w}}\mathbb{Z})}\Big(\prod_{\mathbf{w}\neq(1,1,1)}\mathrm{e}\Big(-ny^{[\mathbf{w}]}/\Delta_{\mathbf{w}}\Big)\Big)\times$$
$$\times\prod_{i=1}^{3}\Big\{\mathcal{D}_{a_i}((y^{[\mathbf{w}]})_{\mathbf{w}(i)=0},(\Delta_{\mathbf{w}})_{\mathbf{w}(i)=0})\prod_{\substack{\mathbf{w}(i)=1\\ \mathbf{w}\neq(1,1,1)}}\mathcal{M}_{a_i}(y^{[\mathbf{w}]},\Delta_{\mathbf{w}})\Big\}.$$

By definition, $\Delta_{(1,1,1)} = 1$, so $\mathfrak{D}_{\mathbf{a}} = \prod_{\mathbf{w}\neq(1,1,1)} \Delta_{\mathbf{w}}$. Note that $\gcd(\Delta_{\mathbf{w}}, \Delta_{\mathbf{v}}) = 1$ for $\mathbf{w} \neq \mathbf{v}$. Using the Chinese remainder theorem and writing every $y \left(\bmod \prod_{\mathbf{v}\neq(1,1,1)} \Delta_{\mathbf{w}}\right)$ as

$$y = \sum_{\mathbf{w}\neq(1,1,1)} y^{[\mathbf{w}]} \prod_{\mathbf{v}\notin\{\mathbf{w},(1,1,1)\}} \Delta_{\mathbf{v}},$$

we see that the sum over $y^{[\mathbf{w}]}$ equals

$$\sum_{y(\bmod \mathfrak{D}_{\mathbf{a}})} \mathrm{e}(-ny/\mathfrak{D}_{\mathbf{a}}) \prod_{i=1}^{3} \left(\sum_{b_i(\bmod \mathfrak{D}_{\mathbf{a}})} \mathrm{e}(b_i y/\mathfrak{D}_{\mathbf{a}}) \delta_{a_i}(b_i \bmod \mathfrak{D}_{\mathbf{a}}) \right).$$

This is clearly

$$\mathfrak{D}_{\mathbf{a}} \sum_{\substack{\mathbf{b}(\bmod \mathfrak{D}_{\mathbf{a}}) \\ \sum_{i=1}^{3} b_i \equiv n (\bmod \mathfrak{D}_{\mathbf{a}})}} \prod_{i=1}^{3} \delta_{a_i}(b_i \bmod \mathfrak{D}_{\mathbf{a}}),$$

thus, recalling (7.20), we have shown that

$$S_{\mathbf{a},0}(n) = \sigma_{\mathbf{a},n}(\mathfrak{D}_{\mathbf{a}}) \prod_{i=1}^{3} \mathcal{L}_{a_i}. \tag{7.54}$$

The proof of (7.22) is concluded upon combining (7.51), (7.53) and (7.54).

7.4.2 The proof of (7.23)

We begin by finding an explicit expression for $\sigma_{\mathbf{a},n}(p)$, for $p \nmid \Delta_1\Delta_2\Delta_3$, that is explicit in terms of n and the h_{a_i}. Define

$$\theta_a(p) := \begin{cases} 1, & \text{if } p \mid h_a, \\ \frac{1}{p}, & \text{if } p \nmid h_a. \end{cases}$$

Lemma 7.4.5. *For an integer c and a prime p with $p \nmid c$ we have*

$$\mathcal{M}_a(c,p) = -\frac{(1+\theta_a(p)\mathrm{e}_p(c))}{(p-1-\theta_a(p))}.$$

Proof. Combining (7.12) and (7.50) we immediately infer

$$\mathcal{M}_a(c,p) = \frac{1}{(p-1-\theta_a(p))} \sum_{\substack{b(\bmod p) \\ \gcd(b,p)=1 \\ \gcd(b-1,p,h_a)=1}} \mathrm{e}_p(bc) \prod_{\substack{\ell \text{ prime} \\ \ell \mid \gcd(b-1,p)}} \left(1 - \frac{1}{\ell}\right).$$

It is now easy to see that the sum over b equals $-1-\mathrm{e}_p(c)$ or $-1-\mathrm{e}_p(c)/p$ according to whether $p \mid h_a$ or $p \nmid h_a$. $\square$

Let us denote the elementary symmetric polynomials in $\theta_{a_i}(p)$ by

$$\begin{aligned}
\Xi_0(p) &:= 1,\\
\Xi_1(p) &:= \theta_{a_1}(p) + \theta_{a_2}(p) + \theta_{a_3}(p),\\
\Xi_2(p) &:= \theta_{a_1}(p)\theta_{a_2}(p) + \theta_{a_2}(p)\theta_{a_3}(p) + \theta_{a_1}(p)\theta_{a_3}(p),\\
\Xi_3(p) &:= \theta_{a_1}(p)\theta_{a_2}(p)\theta_{a_3}(p).
\end{aligned}$$

Lemma 7.4.6. *For every odd integer n and prime $p \nmid \prod_{i=1}^{3} \Delta_i$ we have*

$$\sigma_{\mathbf{a},n}(p) = 1 - \frac{p}{\prod_{1\le i\le 3}(p-1-\theta_{a_i}(p))}\Big(\sum_{\substack{0\le j\le 3\\ j\equiv n \,(\mathrm{mod}\ p)}} \Xi_j(p)\Big) + \prod_{1\le i\le 3}\Big(\frac{1+\theta_{a_i}(p)}{p-1-\theta_{a_i}(p)}\Big).$$

Proof. By Lemma 7.4.4 and Lemma 7.4.5 we see that

$$\sigma_{\mathbf{a},n}(p) = 1 - \frac{1}{\prod_{1\le i\le 3}(p-1-\theta_{a_i}(p))}\sum_{c\in(\mathbb{Z}/p\mathbb{Z})^*} \mathrm{e}_p(-cn)\prod_{1\le i\le 3}(1+\theta_{a_i}(p)\mathrm{e}_p(c)).$$

The sum over c equals

$$\sum_{0\le j\le 3}\Xi_j(p)\sum_{c\in(\mathbb{Z}/p\mathbb{Z})^*}\mathrm{e}_p(c(j-n)) = p\Big(\sum_{\substack{0\le j\le 3\\ j\equiv n \,(\mathrm{mod}\ p)}} \Xi_j(p)\Big) - \prod_{1\le i\le 3}(1+\theta_{a_i}(p))$$

and the proof is complete. □

Lemma 7.4.7. *Let n be an odd integer. If $3 \mid \Delta_1\Delta_2\Delta_3$, then $\prod_{p\nmid\Delta_1\Delta_2\Delta_3}\sigma_{\mathbf{a},n}(p) \neq 0$. If $3\nmid \Delta_1\Delta_2\Delta_3$, then the following are equivalent:*

1. $\prod_{p\nmid\Delta_1\Delta_2\Delta_3}\sigma_{\mathbf{a},n}(p) = 0$,
2. $\sigma_{\mathbf{a},n}(3) = 0$,
3. *One of the following two conditions holds,*

   3 *divides every element in the set* $\{h_{a_1}, h_{a_2}, h_{a_3}\}$ *and* $3\nmid n$, *or*
   3 *divides exactly two elements in the set* $\{h_{a_1}, h_{a_2}, h_{a_3}\}$, *and* $n\equiv 1 \,(\mathrm{mod}\ 3)$.

Furthermore, $\prod_{p\nmid\Delta_1\Delta_2\Delta_3}\sigma_{\mathbf{a},n}(p)\neq 0$ implies $\prod_{p\nmid\Delta_1\Delta_2\Delta_3}\sigma_{\mathbf{a},n}(p) \gg 1$, with an absolute implied constant.

Proof. For a prime $p\nmid\Delta_1\Delta_2\Delta_3$ with $p\ge 5$ there exists at most one $0\le j\le 3$ satisfying $j\equiv n\,(\mathrm{mod}\ p)$, therefore

$$\sum_{\substack{0\le j\le 3\\ j\equiv n \,(\mathrm{mod}\ p)}} \Xi_j(p) \le 3.$$

Invoking Lemma 7.4.6 we obtain

$$\sigma_{\mathbf{a},n}(p) > 1 - \frac{3p}{(p-2)^3} + \frac{1}{(p-1)^3}.$$

Recall that no a_i is a square, hence $2 \nmid h_{a_1}h_{a_2}h_{a_3}$. The fact that n is odd implies that

$$\sum_{\substack{0\leq j\leq 3\\ j\equiv n (\bmod\ 2)}} \Xi_j(2) = \Xi_1(2) + \Xi_3(2) = \frac{13}{8},$$

hence if $\Delta_1\Delta_2\Delta_3$ is odd we can use Lemma 7.4.6 to show that $\sigma_{\mathbf{a},n}(2) = 2$. We have shown that for odd n one has

$$\prod_{\substack{p\nmid \Delta_1\Delta_2\Delta_3\\ p\neq 3}} \sigma_{\mathbf{a},n}(p) \gg 1$$

with an absolute implied constant and it remains to study $\sigma_{\mathbf{a},n}(3)$. One can find an explicit formula for this density by fixing the congruence class of $n \,(\bmod\ 3)$. For example, in the case that $n \equiv 1 \,(\bmod\ 3)$ we have

$$\sigma_{\mathbf{a},n}(3) = 1 - \frac{3(\theta_{a_1}(3) + \theta_{a_2}(3) + \theta_{a_3}(3))}{\prod_{1\leq i\leq 3}(2-\theta_{a_i}(3))} + \prod_{1\leq i\leq 3}\Big(\frac{1+\theta_{a_i}(3)}{2-\theta_{a_i}(3)}\Big)$$

and we can check that $\sigma_{\mathbf{a},n}(3) = 0$ if and only if at most one of the θ_i is equal to $1/3$. A case by case analysis reveals that if $n \equiv 2 \,(\bmod\ 3)$ then $\sigma_{\mathbf{a},n}(3) = 0$ if and only if $(\theta_{a_i}(3))_i = (1,1,1)$ and that if $n \equiv 0 \,(\bmod\ 3)$ then $\sigma_{\mathbf{a},n}(3)$ never vanishes. Noting that $\sigma_{\mathbf{a},n}(3)$ attains only finitely many values as it only depends on $n \,(\bmod\ 3)$ and the choice of $(\theta_{a_i}(3))_i \in \{1, \frac{1}{3}\}^3$, we see that there exists an absolute constant c such that if $\sigma_{\mathbf{a},n}(3) > 0$ then $\sigma_{\mathbf{a},n}(3) > c$, thus concluding our proof. $\square$

We next provide a lower bound for $S_{\mathbf{a},0}(n)$, see (7.54). One could proceed by finding explicit expressions, however, this will lead to rather more complicated formulas than the one for $S_{\mathbf{a},1}(n)$ in Lemma 7.4.6. We shall instead opt to bound the densities $\delta_a(b_i \bmod \mathfrak{D}_{\mathbf{a}})$ from below in (7.54) and then count the number of solutions of the equation $n \equiv x_1 + x_2 + x_3 \,(\bmod\ \mathfrak{D}_{\mathbf{a}})$ such that for every i we have $\delta_a(x_i \bmod \mathfrak{D}_{\mathbf{a}}) \neq 0$.

Lemma 7.4.8. *For any integers q and x such that q is positive and $\delta_a(x \bmod q) > 0$ we have*

$$\delta_a(x \bmod q) \gg \frac{\phi(h_a)}{qh_a},$$

with an absolute implied constant.

Proof. Under the assumptions of our lemma we have the following due to Definition 7.1.4,

$$\delta_a(x \bmod q)\mathcal{A}_a^{-1}\frac{\phi(q)}{f_a^{\dagger}(q)}\prod_{p\mid x-1, p\mid q}\Big(1-\frac{1}{p}\Big)^{-1} =$$
$$1 + \mu\left(\frac{2|\Delta_a|}{\gcd(q,\Delta_a)}\right)\left(\frac{\beta_a(q)}{x}\right) f_a^{\ddagger}\left(\frac{|\Delta_a|}{\gcd(q,\Delta_a)}\right).$$

The right-hand side is either ≥ 1 or equal to $1-f_a^\ddagger(|\Delta_a|\gcd(q,\Delta_a)^{-1})$. In the latter case, since the right-hand side must be positive and $f_a^\ddagger(|\Delta_a|\gcd(q,\Delta_a)^{-1})^{-1}$ is an integer, we see that the right-hand side is $\geq 1/2$. Therefore, under the assumptions of our lemma we have

$$\delta_a(x \bmod q) \geq \frac{\mathcal{A}_a}{2}\frac{f_a^\dagger(q)}{\phi(q)}\prod_{p|x-1,p|q}\Big(1-\frac{1}{p}\Big).$$

It is obvious that $\mathcal{A}_a f_a^\dagger(q) \gg \phi(h_a)/h_a$, with an implied absolute constant. This is sufficient for our lemma owing to $\prod_{p|x-1,p|q}(1-\frac{1}{p}) \geq \phi(q)/q$. □

Recalling (7.20) we see that

$$\sigma_{\mathbf{a},n}(\mathfrak{D}_\mathbf{a})\prod_{i=1}^3\mathcal{L}_{a_i} = \mathfrak{D}_\mathbf{a}\sum_{\substack{b_1,b_2,b_3 (\text{mod } \mathfrak{D}_\mathbf{a})\\ b_1+b_2+b_3\equiv n (\text{mod } \mathfrak{D}_\mathbf{a})}}\prod_{i=1}^3\delta_{a_i}(b_i \bmod \mathfrak{D}_\mathbf{a}),$$

thus, if $\sigma_{\mathbf{a},n}(\mathfrak{D}_\mathbf{a})>0$ then there exist $x_1,x_2,x_3 \pmod{\mathfrak{D}_\mathbf{a}}$ such that

$$\prod_{i=1}^3\delta_{a_i}(x_i \mod \mathfrak{D}_\mathbf{a})>0$$

and $x_1+x_2+x_3\equiv n \pmod{\mathfrak{D}_\mathbf{a}}$. Invoking Lemma 7.4.8 we see that if $\sigma_{\mathbf{a},n}(\mathfrak{D}_\mathbf{a})>0$ then

$$\sigma_{\mathbf{a},n}(\mathfrak{D}_\mathbf{a})\prod_{i=1}^3\mathcal{L}_{a_i} \geq \mathfrak{D}_\mathbf{a}\prod_{i=1}^3\delta_{a_i}(x_i \mod \mathfrak{D}_\mathbf{a}) \gg \mathfrak{D}_\mathbf{a}^{-2}\prod_{i=1}^3\frac{\phi(h_{a_i})}{h_{a_i}}.$$

Recalling (7.21) we obtain $\mathfrak{D}_\mathbf{a}\leq[\Delta_1,\Delta_2,\Delta_3]\leq|\Delta_1\Delta_2\Delta_3|$, hence

$$\sigma_{\mathbf{a},n}(\mathfrak{D}_\mathbf{a})\prod_{i=1}^3\mathcal{L}_{a_i} \gg \prod_{i=1}^3\frac{\phi(h_{a_i})}{|\Delta_i|^2 h_{a_i}}, \tag{7.55}$$

with an absolute implied constant. Combined with Lemma 7.4.7, this concludes the proof of (7.23).

7.4.3 The proof of Theorem 7.1.5

The proof of the first part of Theorem 7.1.5, which is (7.22) is spread throughout §7.4.1. The proof of the second (and last) part of Theorem 7.1.5, which is (7.23), is spread throughout §7.4.2.

7.4.4 The proof of Corollary 7.1.6

Obviously, *(1)* implies *(2)*. For the reverse direction, let $d\in\{3,\mathfrak{D}_\mathbf{a}\}$ and let p_1,p_2,p_3 be primes not dividing $2d$, such that each a_i is a primitive root modulo p_i and

$$p_1+p_2+p_3\equiv n \bmod d.$$

Thus, for every $i = 1, 2, 3$ the progression $p_i \pmod d$ satisfies $\gcd(p_i, d) = 1$ and contains an odd prime having a_i as a primitive root. We can now use the following observation due to Lenstra [51, p.g.216]: if $\gcd(x, d) = 1$ and $\delta_a(x \bmod d) = 0$ then either there is no prime $p \equiv x \pmod d$ with $\mathbb{F}_p^* = \langle a \rangle$ or there is one such prime, which must be equal to 2. This shows that we must have $\delta_a(x_i \bmod d) > 0$ for every $i = 1, 2, 3$. Using the fact that $x_1 + x_2 + x_3 \equiv n \pmod d$, as well as Definition (7.20) shows that $\sigma_{\mathbf{a},n}(\mathfrak{D}_{\mathbf{a}})\sigma_{\mathbf{a},n}(3) > 0$. By Lemma 7.4.7, we get $\mathcal{A}_{\mathbf{a}}(n) > 0$, and thus $\mathcal{A}_{\mathbf{a}}(n) \gg 1$ by (7.23). Thus, *(1)* follows immediately from Theorem 7.1.1 and the trivial estimate

$$\sum_{\substack{p_1+p_2+p_3=n \\ \exists i:\ p_i | 6\Delta_1\Delta_2\Delta_3}} \left(\prod_{i=1}^{3} \log p_i \right) \ll n(\log n)^3.$$

7.4.5 The proof of Theorem 7.1.7

First note that $\mathfrak{D}_{(a,a,a)} = |\Delta_a|$. It is clear that for the proof of Theorem 7.1.7 we need to find equivalent conditions for n to satisfy

$$\sigma_{(a,a,a),n}(|\Delta_a|) \prod_{p \nmid \Delta_a} \sigma_{(a,a,a),n}(p) > 0.$$

By Lemma 7.4.7 the condition $\prod_{p \nmid \Delta_a} \sigma_{(a,a,a),n}(p) \neq 0$ is equivalent to

$$\begin{cases} n \equiv 3 \pmod 6, & \text{if } 3 \mid h_a \text{ and } 3 \nmid \Delta_a, \\ n \equiv 1 \pmod 2, & \text{otherwise.} \end{cases} \tag{7.56}$$

Hence it remains to find equivalent conditions for n to satisfy $\sigma_{(a,a,a),n}(|\Delta_a|) > 0$.

Proposition 7.4.9. *Assume that n is an odd positive integer.*

1. *If $3 \nmid \gcd(\Delta_a, h_a)$ or $3 \mid n$, and if Δ_a has a prime divisor that is greater than 7, then $\sigma_{(a,a,a),n}(|\Delta_a|) > 0$.*
2. *If $3 \mid \gcd(\Delta_a, h_a)$ and $3 \nmid n$, then $\sigma_{(a,a,a),n}(|\Delta_a|) = 0$.*

Proof. It can be seen directly from Definition 7.1.4 that the quantity $\delta_a(x_i \bmod |\Delta_a|)$ is non-zero if and only if

$$\gcd(x_i - 1, \Delta_a, h_a) = 1,\ \gcd(x_i, \Delta_a) = 1 \text{ and } \left(\frac{\Delta_a}{x_i}\right) = -1. \tag{7.57}$$

In view of Definition 7.20, we need to find conditions under which there are $x_1, x_2, x_3 \in \mathbb{Z}$ with $x_1 + x_2 + x_3 \equiv n \bmod \Delta_a$, such that each x_i satisfies (7.57).

To prove *(2)*, we observe that the first two conditions in (7.57) imply that $x_i \equiv 2 \bmod 3$, hence $3 \mid n$.

Let us now prove *(1)*. We can write $\Delta_a = \prod_{p \mid \Delta_a} D_p$, where $D_2 \in \{-8, -4, 8\}$ and $D_p = (-1)^{(p-1)/2} p$ for $p \geq 3$. Let $p' > 7$ be the largest prime divisor of Δ_a. For every $p < p'$, we find $x_1^{(p)}, x_2^{(p)}, x_3^{(p)}$ mod D_p that solve the congruence $x_1^{(p)} + x_2^{(p)} + x_3^{(p)} \equiv n$ mod D_p and satisfy $\gcd(x_i^{(p)} - 1, \Delta_a, h_a) = \gcd(x_i^{(p)}, \Delta_a) = 1$. If $p > 3$, this is possible for every n by a simple application of the Cauchy–Davenport Theorem. If $p = 3$, it is possible precisely by our assumption that then $3 \nmid h_a$ or $3 \mid n$. Finally, for $p = 2$, it is possible since $2 \nmid n h_a$.

Let us now define $x_i^{(p')}$. Consider the sets

$$R := \left\{ x \in \mathbb{Z}/p'\mathbb{Z} : \left(\frac{x}{p'} \right) = 1, x \not\equiv 1 \,(\mathrm{mod}\ p') \right\}, \ N := \left\{ x \in \mathbb{Z}/p'\mathbb{Z} : \left(\frac{x}{p'} \right) = -1 \right\}.$$

If $\prod_{\substack{p \mid \Delta_a \\ p < p'}} \left(\frac{D_p}{x_i^{(p)}} \right) = 1$, we pick $x_i^{(p')} \in N$, and if $\prod_{\substack{p \mid \Delta_a \\ p < p'}} \left(\frac{D_p}{x_i^{(p)}} \right) = -1$, we pick $x_i^{(p')} \in R$. We can always do so and achieve $x_1^{(p')} + x_2^{(p')} + x_3^{(p')} \equiv n$ mod p', as the sets

$$R + R + R, \ R + R + N, \ R + N + N, \ N + N + N$$

cover all of $\mathbb{Z}/p'\mathbb{Z}$. This follows from a direct computation if $p' = 11$ and from the Cauchy–Davenport Theorem if $p' \geq 13$.

To finish our proof of *(1)*, we pick integers x_i that satisfy $x_i \equiv x_i^{(p)}$ mod D_p for all $p \mid \Delta_a$. Then quadratic reciprocity ensures that

$$\left(\frac{\Delta_a}{x_i} \right) = \left(\frac{x_i^{(p')}}{p'} \right) \prod_{\substack{p \mid \Delta_a \\ p < p'}} \left(\frac{D_p}{x_i^{(p)}} \right) = -1$$

for all i. Hence, the x_i satisfy (7.57), and moreover $x_1 + x_2 + x_3 \equiv n$ mod Δ_a. □

Proof of Theorem 7.1.7. First let us note that the fundamental discriminants with every prime smaller than 11 are of the form

$$D_2^{i_1} (-3)^{i_2} 5^{i_3} (-7)^{i_4},$$

where D_2 is an integer in the set $\{-4, 8, -8\}$ and every exponent i_j is either 0 or 1. This gives a finite set of values for Δ_a and it is straightforward to use a computer program that finds all congruence classes $n \,(\mathrm{mod}\ \Delta_a)$ such that $n \equiv x_1 + x_2 + x_3 \,(\mathrm{mod}\ \Delta_a)$ for some $\mathbf{x} \in (\mathbb{Z}/\Delta_a\mathbb{Z})^3$ satisfying all of the conditions (7.57) for $1 \leq i \leq 3$.

By Definition 7.1.4 these conditions are equivalent to $\delta_a(x_i \,\mathrm{mod}\ |\Delta_a|) \neq 0$ and when combined with (7.56) they provide the congruence classes for n in every row of the table in Theorem 7.1.7 apart from the last two rows. For the last two rows, Δ_a has a prime factor greater than 7, so one sees by Proposition 7.4.9 that we only have to provide conditions on n that are equivalent to $\prod_{p \nmid \Delta_a} \sigma_{(a,a,a),n}(p) > 0$, which was already done in (7.56). □

7.4.6 Non-factorisation of $\mathcal{A}_{\mathbf{a}}(n)$

We finish by showing that the right side in (7.22) does not always factorise as an Euler product of a specific form. Namely, assume that for every non-square integer $a \neq -1$ we are given a sequence of real numbers $\lambda_a : \mathbb{Z}^2 \to [0, \infty)$ such that for every prime p and integers x, x' we have

$$\delta_a(x \bmod p) > 0 \Rightarrow \lambda_a(x, p) > 0 \tag{7.58}$$

and

$$x \equiv x' \pmod p \Rightarrow \lambda_a(x, p) = \lambda_a(x', p).$$

Now, in parallel with (7.20) let us define

$$\varpi_{p,a}(n) := \Bigg(\sum_{\substack{b_1,b_2,b_3 \pmod p \\ b_1+b_2+b_3 \equiv n \pmod p}} \prod_{i=1}^{3} \lambda_a(x, p) \Bigg) \Bigg(\sum_{\substack{b_1,b_2,b_3 \pmod p \\ b_1+b_2+b_3 \equiv n \pmod p}} \frac{1}{p^3} \Bigg)^{-1}.$$

The fact that the quantities $\varpi_{p,a}(n)$ are well-defined follows from the periodicity of λ_a.

We will see that one cannot have the following factorisation for all odd integers n,

$$\mathcal{L}_a^3 \sigma_{(a,a,a),n}(|\Delta_a|) = \prod_{p \mid \Delta_a} \varpi_{p,a}(n). \tag{7.59}$$

Indeed, if $a := (-15)^5 = -759375$ then by Definition 7.1.4 we easily see that

$$\delta_{-759375}(x \bmod 15) > 0 \Leftrightarrow x \pmod{15} \in \{7, 13, 14 \pmod{15}\},$$

hence for all integers $n \equiv 7 \pmod{15}$ we have $\sigma_{(a,a,a),n}(|\Delta_a|) = 0$ due to (7.20) and the fact that for all $\mathbf{x} \in \{7, 13, 14\}^3$ one has $\sum_{i=1}^{3} x_i \not\equiv 7 \pmod{15}$. Definition 7.1.4 furthermore implies that

$$\delta_{-759375}(x \bmod 3) > 0 \Leftrightarrow x \pmod 3 \in \{1, 2 \pmod 3\}$$

and

$$\delta_{-759375}(y \bmod 5) > 0 \Leftrightarrow y \pmod 5 \in \{2, 3, 4 \pmod 5\},$$

therefore whenever $n \equiv 7 \pmod{15}$ then the vectors $\mathbf{x} = (1, 1, 2)$ and $\mathbf{y} = (4, 4, 4)$ satisfy

$$\sum_{i=1}^{3} x_i \equiv n \pmod 3, \quad \sum_{i=1}^{3} y_i \equiv n \pmod 5$$

and

$$\prod_{i=1}^{3} \delta_{-759375}(x_i \bmod 3) \delta_{-759375}(y_i \bmod 5) > 0.$$

By (7.58) this implies that $\varpi_{3,-759375}(n) > 0$ and $\varpi_{5,-759375}(n) > 0$. This contradicts equation (7.59) due to $\sigma_{(a,a,a),n}(|\Delta_a|) = 0$.

Bibliography

[1] B. Alberts. Cohen-Lenstra Moments for Some Nonabelian Groups. August 2016.

[2] B. Alberts and J. Klys. The distribution of H_8-extensions of quadratic fields. June 2017.

[3] F. Beukers and H. P. Schlickewei. The equation $x + y = 1$ in finitely generated groups. *Acta Arith.*, 78(2):189–199, 1996.

[4] M. Bhargava. The geometric sieve and the density of squarefree values of invariant polynomials. January 2014.

[5] E. Bombieri and J. Mueller. The generalized Fermat equation in function fields. *J. Number Theory*, 39(3):339–350, 1991.

[6] B. Brindza. The Catalan equation over finitely generated integral domains. *Publ. Math. Debrecen*, 42(3-4):193–198, 1993.

[7] Nils Bruin and Brett Hemenway. On congruent primes and class numbers of imaginary quadratic fields. *Acta Arith.*, 159(1):63–87, 2013.

[8] D. A. Burgess. On character sums and L-series. II. *Proc. Lond. Math. Soc. (3)*, 13:524–536, 1963.

[9] Y.-C. Chiu. S-unit equation over algebraic function fields of characteristic $p > 0$. *Master Thesis, National Taiwan University*, 2002.

[10] H. Cohen and H. W. Lenstra, Jr. Heuristics on class groups of number fields. In *Number theory, Noordwijkerhout 1983 (Noordwijkerhout, 1983)*, volume 1068 of *Lecture Notes in Math.*, pages 33–62. Springer, Berlin, 1984.

[11] H. Cohn and J. C. Lagarias. On the existence of fields governing the 2-invariants of the classgroup of $\mathbf{Q}(\sqrt{dp})$ as p varies. *Math. Comp.*, 41(164):711–730, 1983.

[12] H. Cohn and J. C. Lagarias. Is there a density for the set of primes p such that the class number of $\mathbf{Q}(\sqrt{-p})$ is divisible by 16? In *Topics in classical number theory, Vol. I, II (Budapest, 1981)*, volume 34 of *Colloq. Math. Soc. János Bolyai*, pages 257–280. North-Holland, Amsterdam, 1984.

[13] Pietro Corvaja and Umberto Zannier. An *abcd* theorem over function fields and applications. *Bull. Soc. Math. France*, 139(4):437–454, 2011.

[14] H. Davenport and H. Heilbronn. On the density of discriminants of cubic fields. II. *Proc. Roy. Soc. London Ser. A*, 322(1551):405–420, 1971.

[15] C.-J. de la Vallée Poussin. *Mém. Couronnés Acad. Roy. Belgique*, 59:1–74, 1899.

[16] H. Derksen and D. Masser. Linear equations over multiplicative groups, recurrences, and mixing I. *Proc. Lond. Math. Soc. (3)*, 104(5):1045–1083, 2012.

[17] J.-H. Evertse. On equations in S-units and the Thue-Mahler equation. *Invent. Math.*, 75(3):561–584, 1984.

[18] Jan-Hendrik Evertse and Kálmán Győry. *Unit equations in Diophantine number theory*, volume 146 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2015.

[19] É. Fouvry and J. Klüners. Cohen-Lenstra heuristics of quadratic number fields. In *Algorithmic number theory*, volume 4076 of *Lecture Notes in Comput. Sci.*, pages 40–55. Springer, Berlin, 2006.

[20] É. Fouvry and J. Klüners. On the 4-rank of class groups of quadratic number fields. *Invent. Math.*, 167(3):455–513, 2007.

[21] É. Fouvry and J. Klüners. On the negative Pell equation. *Ann. of Math. (2)*, 172(3):2035–2104, 2010.

[22] É. Fouvry and J. Klüners. The parity of the period of the continued fraction of $\sqrt{d}$. *Proc. Lond. Math. Soc. (3)*, 101(2):337–391, 2010.

[23] J. B. Friedlander and H. Iwaniec. The polynomial $X^2 + Y^4$ captures its primes. *Ann. of Math. (2)*, 148(3):945–1040, 1998.

[24] J. B. Friedlander, H. Iwaniec, B. Mazur, and K. Rubin. The spin of prime ideals. *Invent. Math.*, 193(3):697–749, 2013.

[25] J. B. Friedlander, H. Iwaniec, B. Mazur, and K. Rubin. Erratum to: The spin of prime ideals. *Invent. Math.*, 202(2):923–925, 2015.

[26] C. F. Gauss. *Disquisitiones arithmeticae*. Springer-Verlag, New York, 1986. Translated and with a preface by Arthur A. Clarke, Revised by William C. Waterhouse, Cornelius Greither and A. W. Grootendorst and with a preface by Waterhouse.

[27] F. Gerth, III. The 4-class ranks of quadratic fields. *Invent. Math.*, 77(3):489–515, 1984.

[28] Rajiv Gupta and M. Ram Murty. A remark on Artin's conjecture. *Invent. Math.*, 78(1):127–130, 1984.

[29] H. Halberstam and H.-E. Richert. *Sieve methods.* Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], London-New York, 1974. London Mathematical Society Monographs, No. 4.

[30] G. H. Hardy and J. E. Littlewood. Some problems of 'Partitio numerorum'; III: On the expression of a number as a sum of primes. *Acta Math.*, 44(1):1–70, 1923.

[31] H. Hasse. Über die Klassenzahl des Körpers $P(\sqrt{\ } - 2p)$ mit einer Primzahl $p \neq 2$. *J. Number Theory*, 1:231–234, 1969.

[32] D. R. Heath-Brown. Artin's conjecture for primitive roots. *Quart. J. Math. Oxford Ser. (2)*, 37(145):27–38, 1986.

[33] D. R. Heath-Brown. The size of Selmer groups for the congruent number problem. II. *Invent. Math.*, 118(2):331–370, 1994. With an appendix by P. Monsky.

[34] A. H. Helfgott. The ternary Goldbach problem. January 2015.

[35] Christopher Hooley. On Artin's conjecture. *J. Reine Angew. Math.*, 225:209–220, 1967.

[36] Liang-Chung Hsia and Julie Tzu-Yueh Wang. The *ABC* theorem for higher-dimensional function fields. *Trans. Amer. Math. Soc.*, 356(7):2871–2887, 2004.

[37] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications.* American Mathematical Society, Providence, RI, 2004.

[38] Daniel M. Kane. An asymptotic for the number of solutions to linear equations in prime numbers from specified Chebotarev classes. *Int. J. Number Theory*, 9(4):1073–1111, 2013.

[39] J. Klys. Moments of unramified 2-group extensions of quadratic fields. October 2017.

[40] P. Koymans. The generalized Catalan equation in positive characteristic. October 2016.

[41] P. Koymans and D. Z. Milovic. Spins of prime ideals and the negative Pell equation. *Compos. Math.*, 155(1):100–125, 2019.

[42] P. H. Koymans. The Catalan equation. *Indag. Math. (N.S.)*, 28(2):321–352, 2017.

[43] Peter Koymans and Djordjo Milovic. On the 16-rank of class groups of $\mathbb{Q}(\sqrt{-2p})$ for primes p 1 mod 4. *International Mathematics Research Notices*, page rny010, 2018.

[44] Peter Koymans and Carlo Pagano. On the equation $X_1 + X_2 = 1$ in finitely generated multiplicative groups in positive characteristic. *Q. J. Math.*, 68(3):923–934, 2017.

[45] S. Lang. *Algebraic Number Theory.* Springer-Verlag, New York, second edition, 1986.

[46] Serge Lang. *Introduction to algebraic geometry.* Addison-Wesley Publishing Co., Inc., Reading, Mass., 1972. Third printing, with corrections.

[47] Serge Lang. *Fundamentals of Diophantine geometry.* Springer-Verlag, New York, 1983.

[48] J. L. Lavoie, F. Grondin, and A. K. Rathie. Generalizations of Whipple's theorem on the sum of a ${}_3F_2$. *J. Comput. Appl. Math.*, 72(2):293–300, 1996.

[49] Dominik Leitner. Linear equations over multiplicative groups in positive characteristic II. *J. Number Theory*, 180:169–194, 2017.

[50] F. Lemmermeyer. *Reciprocity laws.* Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2000. From Euler to Eisenstein.

[51] H. W. Lenstra, Jr. On Artin's conjecture and Euclid's algorithm in global fields. *Invent. Math.*, 42:201–224, 1977.

[52] H. W. Lenstra, Jr., P. Stevenhagen, and P. Moree. Character sums for primitive root densities. *Math. Proc. Cambridge Philos. Soc.*, 157(3):489–511, 2014.

[53] P. A. Leonard and K. S. Williams. On the divisibility of the class numbers of $Q(\sqrt{-p})$ and $Q(\sqrt{-2p})$ by 16. *Canad. Math. Bull.*, 25(2):200–206, 1982.

[54] R. C. Mason. *Diophantine equations over function fields*, volume 96 of *London Mathematical Society Lecture Note Series.* Cambridge University Press, Cambridge, 1984.

[55] D. W. Masser. Mixing and linear equations over groups in positive characteristic. *Israel J. Math.*, 142:189–204, 2004.

[56] C. McMeekin. On the Asymptotics of a Prime Spin Relation. November 2018.

[57] D. Milovic. The infinitude of $\mathbb{Q}(\sqrt{-p})$ with class number divisible by 16. *ArXiv e-prints*, February 2015.

[58] D. Milovic. On the 16-rank of class groups of $\mathbb{Q}(\sqrt{-8p})$ for $p \equiv -1 \mod 4$. *Geom. Funct. Anal.*, 27(4):973–1016, 2017.

[59] Pieter Moree. On primes in arithmetic progression having a prescribed primitive root. II. *Funct. Approx. Comment. Math.*, 39(part 1):133–144, 2008.

[60] Pieter Moree. Artin's primitive root conjecture—a survey. *Integers*, 12(6):1305–1416, 2012.

[61] P. Morton. Density result for the 2-classgroups of imaginary quadratic fields. *J. Reine Angew. Math.*, 332:156–187, 1982.

[62] L. Rédei. Arithmetischer Beweis des Satzes über die Anzahl der durch vier teilbaren Invarianten der absoluten Klassengruppe im quadratischen Zahlkörper. *J. Reine Angew. Math.*, 171:55–60, 1934.

[63] Matthias Schütt, Tetsuji Shioda, and Ronald van Luijk. Lines on Fermat surfaces. *J. Number Theory*, 130(9):1939–1963, 2010.

[64] J.-P. Serre. Résumé des cours de 1977-1978. *Annuaire du Collège de France*, pages 67–70, 1978.

[65] Xuancheng Shao. A density version of the Vinogradov three primes theorem. *Duke Math. J.*, 163(3):489–512, 2014.

[66] Joseph H. Silverman. The Catalan equation over function fields. *Trans. Amer. Math. Soc.*, 273(1):201–205, 1982.

[67] Joseph H. Silverman. The S-unit equation over function fields. *Math. Proc. Cambridge Philos. Soc.*, 95(1):3–4, 1984.

[68] A. Smith. Governing fields and statistics for 4-Selmer groups and 8-class groups. *ArXiv e-prints*, July 2016.

[69] A. Smith. 2^∞-Selmer groups, 2^∞-class groups, and Goldfeld's conjecture. *ArXiv e-prints*, February 2017.

[70] P. Stevenhagen. Ray class groups and governing fields. In *Théorie des nombres, Année 1988/89, Fasc. 1*, Publ. Math. Fac. Sci. Besançon, page 93. Univ. Franche-Comté, Besançon, 1989.

[71] P. Stevenhagen. Divisibility by 2-powers of certain quadratic class numbers. *J. Number Theory*, 43(1):1–19, 1993.

[72] W. W. Stothers. Polynomial identities and Hauptmoduln. *Quart. J. Math. Oxford Ser. (2)*, 32(127):349–370, 1981.

[73] R.-C. Vaughan. Sommes trigonométriques sur les nombres premiers. *C. R. Acad. Sci. Paris Sér. A-B*, 285(16):A981–A983, 1977.

[74] I. M. Vinogradov. Representation of an odd number as a sum of three primes. *C. R. (Dokl.) Acad. Sci. URSS*, 15:291–294, 1937.

[75] I. M. Vinogradov. The method of trigonometrical sums in the theory of numbers. *Trav. Inst. Math. Stekloff*, 23:109, 1947.

[76] I. M. Vinogradov. *The method of trigonometrical sums in the theory of numbers.* Dover Publications, Inc., Mineola, NY, 2004. Translated from the Russian, revised and annotated by K. F. Roth and Anne Davenport, Reprint of the 1954 translation.

[77] José Felipe Voloch. Diagonal equations over function fields. *Bol. Soc. Brasil. Mat.*, 16(2):29–39, 1985.

[78] José Felipe Voloch. The equation $ax+by=1$ in characteristic p. *J. Number Theory*, 73(2):195–200, 1998.

[79] M. Widmer. Counting primitive points of bounded height. *Trans. Amer. Math. Soc.*, 362(9):4793–4829, 2010.

[80] M. M. Wood. Non-Abelian Cohen-Lenstra moments. July 2018.

Samenvatting

Dit proefschrift “Diophantische vergelijkingen in positieve karakteristiek” bestaat uit drie losse delen, die elk afzonderlijk kunnen worden gelezen. In het eerste deel bestuderen we zogenaamde exponentiële Diophantische vergelijkingen in positieve karakteristiek. Een voorbeeld van een exponentiële Diophantische vergelijking is

$$2^a \cdot 3^b \cdot 5^c - 2^d \cdot 3^e \cdot 5^f = 1,$$

waar we op zoek zijn naar oplossingen met a, b, c, d, e, f gehele getallen. Het is een bekende stelling dat een dergelijke vergelijking slechts eindig veel oplossingen heeft. Er is ook uitgebreid onderzoek gedaan naar een bovengrens voor het aantal oplossingen.

In de bovenstaande vergelijking zijn 2, 3 en 5 allemaal gehele getallen. Als we de gehele getallen vervangen door vaste elementen uit een lichaam van positieve karakteristiek, kunnen we ons nog steeds afvragen of het mogelijk is om een bovengrens te berekenen. In dit geval is het echter niet langer waar dat er maar slechts eindig veel oplossingen zijn. Wel kunnen de oplossingen op een natuurlijke manier in equivalentieklassen worden verdeeld. Een van de belangrijkste resultaten in dit proefschrift, bewezen samen met Carlo Pagano, is een bovengrens voor het aantal equivalentieklassen.

Twee andere bekende exponentiële Diophantische vergelijkingen zijn de zogenaamde Fermat-vergelijking en de Catalan-vergelijking. De Catalan-vergelijking is

$$x^n - y^m = 1,$$

waar we zoeken naar gehele oplossingen $x, y, m, n > 1$. In 1844 sprak Eugène Catalan al het vermoeden uit dat de enige oplossing $x = 3$, $n = 2$, $y = 2$ en $m = 3$ is. Pas zeer recent (2002) heeft Preda Mihăilescu bewezen dat dit inderdaad klopt! In dit proefschrift bestuderen we het analogon van deze vergelijking over lichamen van positieve karakteristiek. Opnieuw blijken er oneindig veel oplossingen te zijn; we bewijzen dat er slechts eindig veel oplossingen zijn op een natuurlijke equivalentierelatie na.

We sluiten het eerste deel af met een uitgebreide studie van het Fermat-oppervlak gegeven door de vergelijking

$$x^N + y^N + z^N = 1.$$

We zijn deze keer geïnteresseerd voor welke waarden van N er oplossingen x, y en z in het lichaam $\mathbb{F}_p(t)$ bestaan. Als we dezelfde vraag zouden stellen over een lichaam van

karakteristiek 0, bijvoorbeeld de rationale getallen, dan is het nog steeds een compleet open probleem om de oplossingsverzameling van het Fermat-oppervlak te vinden. Samen met Carlo Pagano heb ik bewezen dat er oneindig veel priemgetallen N zijn waarvoor de vergelijking geen oplossingen heeft onder een extra technische voorwaarde. We laten ook zien dat de stelling niet langer waar is zonder deze technische voorwaarde.

Het tweede deel van dit proefschrift betreft statistische eigenschappen van klassegroepen. Al eeuwenlang zijn wiskundigen gefascineerd door klassegroepen en hun relatie met unieke factorisatie. Cohen en Lenstra hebben in 1984 een groot aantal vermoedens uitgesproken over de statistische eigenschappen van klassegroepen. Sindsdien is er uitgebreid onderzoek hiernaar gedaan en dit proefschrift gaat hier verder op in. We kunnen de belangrijkste stellingen uit dit proefschrift als volgt informeel samenvatten.

Stelling. Laat p een priemgetal zijn en $h(-p)$ het klassegetal van $\mathbb{Q}(\sqrt{-p})$. Dan is de dichtheid van de priemen met de eigenschap $16 \mid h(-p)$ gelijk aan $\frac{1}{16}$.

Stelling (samen met Djordjo Milovic)**.** Laat p een priemgetal zijn en $h(-2p)$ het klassegetal van $\mathbb{Q}(\sqrt{-2p})$. Dan is de dichtheid van de priemen met de eigenschap $p \equiv 1 \bmod 4$ en $16 \mid h(-2p)$ gelijk aan $\frac{1}{16}$.

In het derde deel bekijken we een van de meest klassieke problemen in de analytische getaltheorie, namelijke het vermoeden van Goldbach. Christian Goldbach sprak in 1742 het vermoeden uit dat elk oneven getal n groter dan 5 kan worden geschreven als de som van drie priemgetallen. Ivan Vinogradov bewees in de jaren 30 van de vorige eeuw dat dit waar is voor voldoende grote n, en Harald Helfgott heeft het in 2013 voor alle n groter dan 5 bewezen.

Laat $g > 1$ een geheel getal zijn dat geen kwadraat is. We bekijken de vergelijking

$$p_1 + p_2 + p_3 = n,$$

waar de priemen p_1, p_2 en p_3 allemaal g als primitieve wortel hebben, d.w.z. g brengt de groep $(\mathbb{Z}/p_i\mathbb{Z})^*$ voort voor $i = 1, 2, 3$. We laten zien dat de bovenstaande vergelijking voor voldoende grote n altijd een oplossing heeft onder aanname van de veralgemeende Riemann-hypothese, zolang n voldoet aan zekere congruentie condities. Dit artikel is samen met Christopher Frei en Efthymios Sofos geschreven.

Acknowledgements

First, I would like to thank my supervisors Jan-Hendrik and Peter for their support and help. I am especially grateful to Jan-Hendrik for helping me many times throughout my years in Leiden. I also received mathematical advice from Bas, Hendrik and Ronald on several occasions.

During my time in Leiden I had the pleasure of working together with Carlo, Djordjo and Efthymios on various projects. This experience was invaluable to me, and all three of them helped me grow as a mathematician. Moreover, I would like to thank my office mates Giulio, Jinbi, Raymond, Stefan and Steven for providing a great atmosphere during work.

My time in Leiden would not have been the same without the great friends I made. I had a lot of fun going to Einstein with Abtien, Djordjo, Giulio, Janusz and Neha. I will also not forget the many great nights spent in Pelibar with Amine, Djordjo, Francesco, Giulio, Ilaria, Jared, Margherita, Martina, Matteo, Neha, Stefan and Thibault. It was also a great pleasure for me to play boardgames with Amine, David, Djordjo, Erik, Garnet, Giulio, Ilaria, Jared, Matteo, Margherita, Martina, Raymond, Stefan and Thibault.

Thank you to Alessandro, Andrea, Garnet, Guido, Julian, Marco, Marta, Martin, Mima, Peter, Robin, Rosa, Rosa, Wouter, all the other people at the Snellius, the chess players from Schaakclub Oegstgeest and the people from the ISN board game nights.

In Eindhoven I had a lot of fun going out and playing various board and card games with Arthur, Bas, Guus, Ingrid, Iris, Jarich, Jeroen, Jessica, Jochem, Martijn, Tim and Ton. I will not forget the many Friday nights we spent together. Finally I would like to thank my girlfriend Ilaria and my family, in particular my grandmother, my father Ron, my mother Letty and my sister Karin.

Curriculum vitae

Peter Hubrecht Koymans was born on the 24th of July 1992 in Eindhoven, where he also received his pre-university education from 2004 to 2010. During this time he participated in the Dutch Mathematical Olympiad with a top 10 placement. After graduating cum laude from high school, he went to Eindhoven University of Technology to study applied mathematics and computer science from 2010 to 2013 graduating cum laude in both. He also won the “Jong Talent Prijs” from the KNAW.

After his bachelor, he went to Universiteit Leiden from 2013-2015 to obtain his master degree in “Algebra, Geometry and Number Theory”. His master thesis was written under the supervision of dr. Jan-Hendrik Evertse, and Peter graduated cum laude. After obtaining his master degree, he did a Ph.D. in mathematics at Universiteit Leiden, supervised by prof. dr. Peter Stevenhagen and dr. Jan-Hendrik Evertse. The Ph.D. resulted in several published articles, this booklet and the KWG Ph.D. prize.

After his Ph.D., Peter will join the Max Planck Institute in Bonn from September 2019 to August 2020 for a postdoctoral position followed by a three year postdoctoral position at the University of Michigan.