Algebra & Number Theory

Volume 18
2024
No. 9

Prime values of $f(a,b^2)$ and $f(a,p^2)$, f quadratic

Stanley Yao Xiao

msp

ALGEBRA AND NUMBER THEORY 18:9 (2024)
https://doi.org/10.2140/ant.2024.18.1619

msp

Prime values of $f(a,b^2)$ and $f(a,p^2)$, f quadratic

Stanley Yao Xiao

Dedicated to the occasion of John Friedlander's 80th birthday

We prove an asymptotic formula for primes of the shape $f(a, b^2)$ with a, b integers and of the shape $f(a, p^2)$ with p prime. Here f is a binary quadratic form with integer coefficients, irreducible over $\mathbb{Q}$ and has no local obstructions. This refines the seminal work of Friedlander and Iwaniec on primes of the form $x^2 + y^4$ and of Heath-Brown and Li on primes of the form $a^2 + p^4$, as well as earlier work of the author with Lam and Schindler on primes of the form $f(a, p)$ with f a positive definite form.

1. Introduction

Two of the most stunning results in prime number theory in the last thirty years are the seminal works of Friedlander and Iwaniec [5] and Heath-Brown [9], demonstrating that the polynomials $x^2 + y^4$ and $x^3 + 2y^3$, respectively, take on infinitely many prime values. In particular, Friedlander and Iwaniec obtained the beautiful asymptotic formula

$$\sum_{a^2+b^4\le X}\sum \Lambda(a^2+b^4) = \frac{2\Gamma(1/4)^2}{3\pi\sqrt{2\pi}}X^{3/4}\left(1+O\left(\frac{\log\log X}{\log X}\right)\right), \tag{1-1}$$

where $\Lambda(\cdot)$ is the von Mangoldt function and Γ is the Gamma function.

Heath-Brown's result on $x^3 + 2y^3$ was quickly generalized by Heath-Brown and Moroz in [11], which demonstrated that any admissible binary cubic form takes on infinitely many prime values. More recently, X. Li has proved that the cubic form $x^3 + 2y^3$ takes on infinitely many prime values with y restricted to a short interval [14]. One also notes the stunning work of J. Maynard on representation of primes by incomplete norm forms, a substantial generalization of Heath-Brown's work [15].

Despite the passage of more than two decades, a generalization akin to that of Heath-Brown and Moroz [11] has yet to materialize for the main result of [5], despite the authors of that paper claiming that such a result should be readily obtainable from their arguments.[1] That is, there has yet to be a proof that $f(x, y^2)$ takes on infinitely many prime values for any binary quadratic form f other than $f(x, y) = x^2 + y^2$.

More precisely, it can be seen from [9; 11] that the work needed to go from prime values of $x^3 + 2y^3$ to prime values of $F(x, y)$ for arbitrary admissible cubic forms F is purely algebraic. In particular, the

MSC2020: primary 11N32; secondary 11N35, 11N36, 11R45.
Keywords: primes, prime values of quadratic forms, Friedlander–Iwaniec theorem.

[1] "We expect, but did not check, that the methods carry over to the prime values of $\phi(a, b^2)$ for ϕ a quite general binary quadratic form." [5, p. 947].

analytic machinery established by Heath-Brown in [9] essentially only depends on the $\mathbb{Z}$-module structure of $\mathbb{Z}\big[\sqrt[3]{2}\,\big]$, which means it can be easily adapted to work with sets of ideal numbers.

Such is not the case with a^2+b^4. In fact the analytic machinery in [5] is far more delicate, as they needed to work around the lack of homogeneity of the polynomial. Much of this machinery is quite subtle. Therefore, in addition to establishing an appropriate algebraic framework akin to the work of Heath-Brown and Moroz, it is necessary to generalize some of the analytic machinery in [5] as well. For the algebraic framework, we use language established by Heath-Brown and Moroz, but in principle we can use the same type of explicit language used by Lam, Schindler, and the author in [12].

Fortuitously, we are able to salvage a significant portion of the analytic machinery established by Friedlander and Iwaniec [5] and Heath-Brown and Li [10]. There is one notable exception, which can be viewed as the most novel contribution of this paper: the so-called Jacobi–Kubota symbol. In fact this symbol, introduced in [5], works well *only* in the ring of Gaussian integers $\mathbb{Z}[i]$. This symbol is subtle because unlike much of the other pieces of analytic machinery, it relies also on the *arithmetic structure* of the sets of ideal numbers of the quadratic field associated to f. Obtaining a generalization of the multiplicativity of the Jacobi–Kubota symbol workable in the general setting is crucial to our arguments.

In another direction, one might ask whether *reducible polynomials* take on infinitely many *semiprime values*, with the order of the semiprime being equal to the number of irreducible factors. A first example of this type of result is due to Fouvry and Iwaniec [3], who showed that the binary cubic form $y(x^2+y^2)$ takes on infinitely many values with exactly two prime factors. This work paved the way for the later work of Friedlander and Iwaniec [5]. Heath-Brown and Li then combined the result of Fouvry and Iwaniec and Friedlander and Iwaniec in [10], showing that the polynomial $y(x^2+y^4)$ takes on infinitely many values with exactly two prime factors. In particular they obtained the asymptotic formula

$$\sum_{a^2+b^4\le X}\sum \lambda(b)\lambda(a^2+b^4) = \frac{2\Gamma(1/4)^2}{3\pi\sqrt{2\pi}}\frac{X^{3/4}}{(\log X)^2}\left(1+O_\varepsilon\left(\frac{1}{(\log X)^{1-\varepsilon}}\right)\right), \tag{1-2}$$

where λ is the prime indicator function.

Lam, Schindler and the author generalized the work of Fouvry and Iwaniec in another direction, proving that for any admissible positive definite binary quadratic form f the cubic form $yf(x,y)$ takes on infinitely many values with exactly two prime factors. Our main result implies

$$\sum_{f(m,\ell)\le X}\sum \Lambda(\ell)\Lambda(f(m,\ell)) = \nu_f \mathfrak{S}'_f X + O_A(X(\log X))^{-A}), \tag{1-3}$$

where ν_f is a product of local densities given by

$$\nu_f = \prod_{p\nmid\Delta(f)}\left(1-\frac{\rho_f(p)}{p}\right)\left(1-\frac{1}{p}\right)^{-1}\prod_{p\mid\Delta(f)}\left(1-\frac{1}{p}\right)^{-1}, \tag{1-4}$$

$\mathfrak{S}'_f$ is given by (1-9), and $\rho_f(m) = \#\{x \pmod m : f(x,1)\equiv 0 \pmod m\}$.

We simultaneously generalize the main results of Friedlander and Iwaniec [5] and Heath-Brown and Li [10]. If f is definite put

$$\mathfrak{S}_f = \text{Area}\{(x, y) \in \mathbb{R}^2 : f(x, y^2) \le 1\}$$

and for f indefinite we define

$$\mathfrak{S}_f = \lim_{X\to\infty} \frac{\text{Area}\{(x, y) \in \mathbb{R}^2 : 0 < f(x, y^2) < X,\ 0 < y \le X^{1/4}\}}{X^{3/4}}.$$

Our first main result is:

Theorem 1.1. *Let $f(x, y) = f_2x^2 + f_1xy + f_0y^2 \in \mathbb{Z}[x, y]$ be an irreducible and primitive binary quadratic form, with the property that $f(x, 1) \not\equiv x(x + 1) \pmod 2$. Then for f positive definite we have*

$$\sum_{\substack{m,\ell\in\mathbb{Z}\\ f(m,\ell^2)\le X}} \lambda(f(m, \ell^2)) = \frac{\nu_f\mathfrak{S}_f X^{3/4}}{\log X}\left(1 + O\left(\frac{\log\log X}{\log X}\right)\right) \tag{1-5}$$

and for f indefinite we have

$$\sum_{\substack{m,\ell\in\mathbb{Z}\\ 0<f(m,\ell^2)\le X\\ 0<\ell\le X^{1/4}}} \lambda(f(m, \ell^2)) = \frac{\nu_f\mathfrak{S}_f X^{3/4}}{\log X}\left(1 + O\left(\frac{\log\log X}{\log X}\right)\right). \tag{1-6}$$

The condition that $f(x, 1) \not\equiv x(x + 1) \pmod 2$ is necessary, as otherwise $f(x, k)$ is divisible by 2 whenever k is odd, precluding the possibility that it could be a prime square unless $k = 4$. Theorem 1.1 recovers Theorem 1 of [5] upon setting $f(x, y)=x^2+y^2$. It also implies, for example, that the polynomials $x^2 + xy^2 + y^4$ and $x^2 - 2y^4$ represent infinitely many primes.

The choice of cutting off the y-variable at $X^{1/4}$ is somewhat arbitrary, and is mostly done for aesthetic reasons. Of course, in the indefinite case some such cut-off is necessary. In particular such a choice guarantees that we do not need to worry about long cusps if we insist only on the condition $|f(x, y^2)| \le X$.

Both Theorems 1.1 and 1.2 apply to *indefinite* as well as definite forms. Although, for economy, we state the two cases together, there are some differences in the proof and, so far as we are aware, these give the first examples of asymptotic formulae for the number of prime values of indefinite nonhomogeneous polynomials of degree exceeding two.

Our proof, which further develops ideas in [10], yields the following general version of Theorem 1 of [10] or (1-2):

Theorem 1.2. *Let $f(x, y) = f_2x^2 + f_1xy + f_0y^2 \in \mathbb{Z}[x, y]$ be an irreducible and primitive binary quadratic form, with the property that $f(x, 1) \not\equiv x(x + 1) \pmod 2$. Then for f positive definite we have*

$$\sum_{\substack{m,\ell\in\mathbb{Z}\\ 0<f(m,\ell^2)\le X}} \lambda(\ell)\lambda(f(m, \ell^2)) = \frac{\nu_f\mathfrak{S}_f X^{3/4}}{(\log X)^2}\left(1 + O\left(\frac{\log\log X}{\log X}\right)\right) \tag{1-7}$$

and for f indefinite we have

$$\sum_{\substack{m,\ell\in\mathbb{Z}\\ 0<f(m,\ell^2)\le X\\ 0<\ell\le X^{1/4}}}\lambda(\ell)\lambda(f(m,\ell^2))=\frac{\nu_f\mathfrak{S}_f X^{3/4}}{(\log X)^2}\left(1+O\left(\frac{\log\log X}{\log X}\right)\right). \tag{1-8}$$

Theorem 1.2 implies that there are infinitely many integers x and primes p for which $f(x, p^2)$ is prime. We further note that the error term in Theorem 1.2 is slightly better than in (1-2), due to choosing a slightly different sieving parameter.

In [11], the key new insight is that the arithmetic of *ideal numbers* allows one to connect the multiplicative structure on the set of ideals of a ring of integers, which has unique factorization, to the arithmetic of the elements in a ring of integers which need not have unique factorization. This breaks a key barrier in [9] where the fact that $\mathbb{Z}\big[\sqrt[3]{2}\,\big]$ is a unique factorization domain is used in a crucial manner. The analytic estimates obtained by Heath-Brown in [9] can be applied with relatively few changes in the general setting [11].

In [12] we essentially pursued the same approach, although we did not state things in terms of ideal numbers but rather worked out an explicit composition law for binary quadratic forms, in the spirit of Gauss and Dirichlet. We have decided to adopt the approach of Heath-Brown and Moroz and use ideal numbers, as this is a more elegant and general approach.

In order to prove Theorems 1.1 and 1.2 we adopt an approach introduced by Heath-Brown in [9], which we call Heath-Brown's *comparison sieve*. This involves applying the same sieve procedure to two comparable sequences $\mathcal{A}=(a_n)$ and $\mathcal{B}=(b_n)$, producing cancellation at appropriate junctures. This was used again by Heath-Brown and Li in [10] for the proof of their result.

In order to prove Theorem 1.1 we choose our sequence $\mathcal{B}$ to simply be the set of prime ideals of the ring of integers $\mathcal{O}_K$, where $K=\mathbb{Q}\big(\sqrt{\Delta(f)}\,\big)$ is the splitting field of our form f. The sequence $\mathcal{B}$ used by Heath-Brown and Li is exactly the sequence studied by Fouvry and Iwaniec in [3]. For positive definite forms f we may then apply the result in [12], and for indefinite forms we will need to prove an extension of our main result with Lam and Schindler in [12], which gives an asymptotic formula for the number of representation of primes by $f(x, p)$, with p prime.

For f positive definite put

$$\mathfrak{S}'_f=\operatorname{Area}\{(x,y)\in\mathbb{R}^2: f(x,y)\le 1\} \tag{1-9}$$

and for f indefinite put

$$\mathfrak{S}'_f=\lim_{X\to\infty}\frac{\operatorname{Area}\{(x,y)\in\mathbb{R}^2: 0<f(x,y)<X,\ 0<y<X^{1/2}\}}{X}.$$

Theorem 1.3. *Let $f(x, y)=f_2x^2+f_1xy+f_0y^2\in\mathbb{Z}[x, y]$ be an irreducible and primitive binary quadratic form, with the property that $f(x, 1)\not\equiv x(x+1)\pmod 2$. Then for f positive definite we have*

$$\sum_{\substack{m,\ell\in\mathbb{Z}\\ 0<f(m,\ell)\le X}}\Lambda(\ell)\Lambda(f(m,\ell))=\nu_f\mathfrak{S}'_f X+O_A\left(\frac{X}{(\log X)^A}\right) \tag{1-10}$$

and for f indefinite we have

$$\sum_{\substack{m,\,\ell\in\mathbb{Z}\\ 0<f(m,\,\ell)\le X\\ 0<\ell\le X^{1/2}}}\Lambda(\ell)\Lambda(f(m,\,\ell))=\nu_f\mathfrak{S}'_f X+O_A\bigg(\frac{X}{(\log X)^A}\bigg). \tag{1-11}$$

Here ν_f is as in Theorem 1.2 and $\mathfrak{S}'_f$ is as in (1-9).

Theorem 1.3 is stated with the von Mangoldt function rather than λ to emphasize that a substantially better error term, giving an arbitrary log-power saving, is possible.

Theorem 1.3 implies the following, which completely settles *Schinzel's hypothesis* for binary cubic forms:

Corollary 1.4. *Let $F(x,\,y)$ be a reducible binary cubic form of the shape $F(x,\,y)=L(x,\,y)Q(x,\,y)$, where Q is an irreducible binary quadratic form. Then there are infinitely many pairs of integers $x,\,y$ such that $F(x,\,y)$ is divisible by exactly two primes.*

Corollary 1.4 is the final case of *Schinzel's hypothesis* in the setting of binary cubic forms. The hardest case, that of irreducible binary cubic forms, is settled by the work of Heath-Brown [9] and Heath-Brown and Moroz in [11]. The case with F reducible with a positive definite quadratic factor is settled by the author's joint work with Lam and Schindler in [12]. The special cases when the irreducible quadratic factor is x^2+y^2 was settled by Fouvry and Iwaniec in [3] and the special case when the quadratic factor is x^2+xy+y^2 was settled by M. Pandey [17]. The totally reducible case was settled by van der Corput [2], and was of course famously generalized by B. J. Green and T. Tao to cover arbitrarily long arithmetic progressions [8]; see also B. J. Green's work on Roth's theorem in the primes [7].

As in [11], our basic objective is to invoke composition laws involving ideal numbers of a fixed quadratic field in order to reduce the problem to one that is amenable to the analytic methods developed by Friedlander and Iwaniec in [5] and Heath-Brown and Li in [10]. In [11] the relevant analytic methods developed by Heath-Brown in [9] can be applied with only minor modifications once the relevant algebraic framework is established, since these estimates depend only on the $\mathbb{Z}$-module structure. However, the analytic estimates employed by Friedlander and Iwaniec in [5] are far more delicate and depend subtly on the fine arithmetic properties of the ring $\mathbb{Z}[i]$ rather than simply its structure as a rank-two $\mathbb{Z}$-module. Indeed, the obvious analogue of the so-called Jacobi–Kubota symbol introduced by Friedlander and Iwaniec in [5] does not seem to behave nicely and special care must be taken to define and work with the twisting factor $\xi_w(z)$ needed to recover multiplicativity. Again we emphasize that the definition and application of these generalized Jacobi–Kubota symbols and their twisting factors may be viewed as the most novel contribution of this paper. We give a rough explanation of this in the following section.

Organization of the paper. In Section 2 we give a brief overview of the ideas in this paper, emphasizing key new ingredients. In Section 3 we discuss our approach to implementing the asymptotic sieve for primes, in the manner introduced by Heath-Brown in [9] which we dub *Heath-Brown's comparison sieve*,

also used by Heath-Brown and Moroz in [11] and Heath-Brown and Li in [10]. In Section 4 we introduce the necessary algebraic number theory involving the arithmetic ideal numbers, necessary to establish the framework needed to apply the analytic estimates in [5; 10]. In Section 5 we establish the needed level of distribution or type-I estimates. In Section 6 we will prove the necessary bilinear sum estimates to obtain the analogue of the main theorem of [12] in the indefinite case, which for us is needed to apply Heath-Brown's comparison sieve in the indefinite case. In Section 7 we establish the preliminary steps to proving our two key technical propositions, being Propositions 7.5 and 7.6, which are analogues of Heath-Brown and Li's Propositions 6 and 7 in [10]. In Section 8 we prove Proposition 7.5, the proof being identical to that of [10] except we avoid the language of Gaussian integers. In Sections 9 and 10 we modify Heath-Brown and Li's proof of their Proposition 7 in the setting of a general quadratic field K, thereby proving our Proposition 7.6, which then completes the proof of Theorem 1.2, conditioned on certain character sum estimates that they imported from [5]. Finally, in Section 11 we introduce the analogues of Friedlander and Iwaniec's notion of *Jacobi–Kubota symbols* in the setting of a general quadratic field, as well as the analogue of their symbol $[\cdot]$ which in some sense measures the "spin" of an ideal in $\mathbb{Z}[i]$, allowing us to prove versions of their Proposition 23.1 and Theorem ψ which are needed by Heath-Brown and Li. This may be of independent interest.

Notation. Throughout, we fix our binary quadratic form

$$f(x, y) = f_2x^2 + f_1xy + f_0y^2 \in \mathbb{Z}[x, y],$$

which satisfies the hypothesis that for all primes p there exist integers x_p, y_p such that $p \nmid f(x_p, y_p)$, and $f(x, 1) \not\equiv x(x+1) \pmod 2$. We will use both the Landau and Vinogradov notation $\ll$ and $O(\cdot)$.

2. Sketch of the main ideas

To sketch our ideas it is necessary to give a quick summary of the works of Friedlander–Iwaniec [5], Heath-Brown and Li [10], as well as the works of Heath-Brown [9] and Heath-Brown and Moroz [11]. In particular, we will see in this section that the paths to Theorems 1.1 and 1.2 are not as straightforward as going from [9] to [11].

We divide our arguments into ideas that are essentially algebraic, ideas which are essentially analytic in nature, and the final subsection is devoted to the new ingredient needed to tie these two bags of tools together to give the proof.

Analytic (sieve theoretic) ideas. In [5] the principal strategy is to verify the hypotheses of the asymptotic sieve for primes, introduced by Friedlander and Iwaniec in [4], hold for the sequence $\mathcal{A} = (a_n)$ defined by

$$a_n = \sum_{x^2+y^4=n} 1. \tag{2-1}$$

In order to do so they obtain an optimal level of distribution (or type-I estimates) for the sequence $\mathcal{A}$, which is extended and refined in their subsequent work [6]; see also [1]. The strength of their result relies

on the remarkable property that roots of quadratic congruences are extraordinarily well-spaced modulo 1. However, the main obstacle they overcome in [5] is to obtain acceptable estimates for *bilinear sums* of the shape

$$\sum_m \alpha(m) \sum_{\substack{N\le n<2N\\ mn\le X}} \beta(n) a_{mn}$$

for quite general complex sequences $(\alpha(m))$ and $(\beta(n))$.

To do so, Friedlander and Iwaniec converted the problem to one about estimating solutions to a family of quadratic congruences via Fourier analysis. They then succeeded in obtaining satisfactory estimates for the number of solutions after a herculean effort in [5]. This estimation constitutes the bulk of the work done in [5].

They partitioned their argument into estimating solutions with small, medium, or large moduli. As usual, the contribution from the small moduli is expected to be relatively straightforward since explicit asymptotic formulae are expected to exist. In [5] this was done from the ground up and in [10] this was obtained by applying the general Siegel–Walfisz type theorems of Mitsui [16].

One pillar of [5], which treats the lion's share of possible moduli in the middle, is their Proposition 14.1. There they cleverly used quadratic reciprocity to achieve *moduli flipping*, which allows one to swap a large modulus with a small one via complementary divisors as long as the modulus is not too large (so that its complementary divisor is not too small). This aspect of Friedlander–Iwaniec is imported without change in [10].

We state Proposition 14.1 in [5] here for convenience:

Proposition 2.1 [5, Proposition 14.1]. *Let $D, R, S \ge 1$. For any complex numbers α_{rs} with* $\gcd(r, 2s) = 1$ *supported in the box $R < r < 2R$, $S < s < 2S$ we have*

$$\sum_{D<d\le 2D} \sum_{a \,(\mathrm{mod}\, d)} \left| \sum_{\bar{r}s\equiv a \,(\mathrm{mod}\, d)} \sum_{} \alpha_{rs} \left(\frac{r}{d}\right) \right|^2 \le \mathcal{N}(D, R, S) \sum_r \sum_s \tau(r) |\alpha_{rs}|^2,$$

where $\mathcal{N}(D, R, S)$ satisfies the bound

$$\mathcal{N}(D, R, S) \ll_\varepsilon D + D^{-1/2}RS + D^{1/3}(RS)^{2/3}(\log 2RS)^4 + (R+S)^{1/12}(RS)^{1/12+\varepsilon}$$

for any $\varepsilon > 0$. Here $\left(\frac{\cdot}{\cdot}\right)$ is the Jacobi symbol if d is odd and is extended for d even via the Hilbert symbol.

It is important to emphasize that Proposition 14.1 in [5] only involves rational integers, and therefore only depends on the structure of $\mathbb{Z}^2$ as a $\mathbb{Z}$-module. The relevance to the present setting is that our particular quadratic field is of no concern when invoking this proposition.

The treatment of large moduli constitutes the bulk of the hard work in [5; 10]. Here the main obstacle to overcome is an acceptable estimate of a sum of the shape

$$\sum_{z_1, z_2} \beta'_{z_1} \beta'_{z_2} \tag{2-2}$$

where β' takes the form

$$\beta'_z = \beta_z i^{(x-1)/2}\left(\frac{y}{x}\right)$$

with $z = (x, y)$ and i is the imaginary unit. The function β_z is supported on a small region in $\mathbb{R}^2$. The key property needed here is that the sum (2-2) can be split as a product

$$\sum_{z_1, z_2} \beta'_{z_1}\beta'_{z_2} = \Big(\sum_{z_1} \beta'_{z_1}\Big)\Big(\sum_{z_2} \beta'_{z_2}\Big) \tag{2-3}$$

using properties of the Jacobi symbol and the arithmetic of $\mathbb{Z}[i]$. This will be discussed in detail later in Section 10.

We remark that two key results in [10], namely Corollaries 1 and 2 which are a refinement of the Barban–Davenport–Halberstam theorem and a Siegel–Walfisz type estimate, respectively, are not explicitly invoked here. This is because these two results are used in [10] to prove their Proposition 6 which, surprisingly, can be applied more or less without change in our case.

Algebraic ideas. The ideas in this subsection are introduced by Heath-Brown and Moroz [11], and are also related to the work of the author with Lam and Schindler [12].

The main role played by algebraic number theory in the present work is to obtain an analogue of equation (5.2) in [5], which we state here for convenience:

$$a_{mn} = \frac{1}{4}\sum_{|w|^2=m}\sum_{|z|^2=n} \mathfrak{Z}(\mathrm{Re}(\overline{w}z)),$$

where $\mathfrak{Z}$ is the indicator function for square integers. This crucial formula allows one to decompose the terms a_n given in (2-1) multiplicatively, which is the principal reason why such strong bilinear sum estimates can be obtained. In [5] they used the fact that the Gaussian integers $\mathbb{Z}[i]$ is a principal ideal domain, and more crucially has a *canonical basis*, in order to obtain their equation (5.2).

In general the quadratic order $\mathcal{O}_K$ we are working in is not a PID. Indeed if we are interested in general binary quadratic forms it is not enough to only work with $\mathcal{O}_K$ but rather *all* sets of ideal numbers simultaneously. If we denote by h the class number of $\mathcal{O}_K$ and $A_1, \dots, A_h$ the corresponding sets of ideal numbers, what we require is a choice of a basis of the A_j's as $\mathbb{Z}$-modules and a *composition law* connecting them, expressed in terms of the given bases.

These ideas are already expressed fully by Heath-Brown and Moroz in [11]. In [12] we adopted a more down-to-earth but ultimately more explicit approach. In the present paper we adapt the ideas in [11] instead.

The key algebraic result we will need is Proposition 4.1 which gives the analogue of equation (5.2) in [5]. This means that we again have sums which can be decomposed multiplicatively which enables us to obtain strong bilinear sum estimates.

The new input: a generalized Jacobi–Kubota symbol and its twist factor. So far we have discussed how Friedlander and Iwaniec relied on the fact that $\mathbb{Z}[i]$ is a PID in order to obtain their decomposition

formula. Likewise, Heath-Brown had relied on the fact that $\mathbb{Z}\big[\sqrt[3]{2}\,\big]$ is a PID in order to obtain the decomposition formula he needed in [9]. The key new idea in [11] was to use the algebraic structure provided by Hecke's ideal numbers in order to reduce the problem of finding prime values of a binary cubic form F to estimating certain sums on $\mathbb{Z}$-modules. The latter constitutes the analytic portion of the argument. In essence, in [11] they successfully separated the algebraic and analytic arguments.

We are doing much the same, but there is one component of the arguments in [5] (and this is inherited by [10]) that is solidly wedged in between the algebraic and analytic worlds and requires separate treatment: the *Jacobi–Kubota symbol*.

The Jacobi–Kubota symbol as defined in [5] is essential in obtaining the decomposition property of (2-2) given as (2-3). Indeed, this is the lynchpin that holds together the arguments needed to obtain suitable estimates for the largest moduli in [5; 10].

The Jacobi–Kubota symbol is extremely specific to the Gaussian integers $\mathbb{Z}[i]$. In particular to obtain the nice properties derived in [5] one must use in an essential way the following properties of $\mathbb{Z}[i]$:

- the class number of $\mathbb{Z}[i]$ is 1;
- the norm of $\mathbb{Z}[i]$ is the same as the Euclidean norm on $\mathbb{C}$; and
- the odd rational primes that split in $\mathbb{Z}[i]$ are precisely those that are congruent to $1 \pmod 4$.

Clearly, no other ring of quadratic integers except for possibly suborders of $\mathbb{Z}[i]$ possess all three of these properties.

To see how these properties come into play, note that Friedlander and Iwaniec defined the Jacobi–Kubota symbol in [5, equation (20.1)] as

$$[z] = [r + is] = i^{(r-1)/2}\left(\frac{s}{|r|}\right)$$

where $\left(\frac{\cdot}{\cdot}\right)$ is the Jacobi symbol. One sees right away that the choice of basis is relevant: the components s, r in the definition cannot make sense without a choice of basis. Right now it appears that the Jacobi–Kubota symbol depends only on the $\mathbb{Z}$-module structure of $\mathbb{Z}[i]$. This is indeed the case: once we have fixed a basis for our relevant $\mathbb{Z}$-module we can define the Jacobi–Kubota symbol analogously.

The trouble is that in order to have the desired multiplicative property, namely

$$[z][w] = \varepsilon[zw]\xi_w(z), \tag{2-4}$$

where $\varepsilon = \pm 1$ depending only on the quadrants containing z, w respectively, the "twist factor" $\xi_w(z)$ must satisfy nice properties that critically depend on the arithmetic structure of $\mathbb{Z}[i]$ as well as the niceness of the canonical basis. In particular, the twist factor $\xi_w(z)$ satisfies:

(1) It is multiplicative for each $w \in \mathbb{Z}[i]$: one has $\xi_w(z_1)\xi_w(z_2) = \xi_w(z_1 z_2)$.

(2) It is symmetric: $\xi_w(z) = \xi_z(w)$ for $w, z \in \mathbb{Z}[i]$.

(3) (Lemma 21.1 in [6]) For $q = |w_1 w_2|^2$ and $d = |\gcd(w_1, \overline{w_2})|^2$ one has

$$\sum_{\zeta \,(\mathrm{mod}\, q)} \xi_{w_1}(\zeta)\xi_{w_2}(\zeta) = \begin{cases} q\varphi(d)\varphi(q/d) & \text{if } q, d \text{ are squares,} \\ 0 & \text{otherwise.} \end{cases}$$

(4) For $w = u + iv$ and $\omega \equiv -v\bar{u} \pmod{q}$ with $q = |w|^2$, one has

$$\xi_w(z) = \left(\frac{ur - vs}{q}\right) \quad \text{and} \quad \xi_w(z) = \left(\frac{r + \omega s}{q}\right),$$

where $z = r + is$.

In order for the Jacobi–Kubota symbol, defined analogously to [5], to have nice properties we must introduce an analogous twist factor to $\xi_w(z)$. One sees right away that this is a tall order. Simply writing down the definition will require much of the setup which will take place throughout the paper, so we defer this until Section 11.

3. Heath-Brown's comparison sieve

We describe the ideas given by Heath-Brown in [9] and expanded upon and refined in [11] and [10]. Heath-Brown's great insight is that quite often it is possible to establish the infinitude of primes in a sequence $\mathcal{A}$ by comparing it to a suitable sequence $\mathcal{B}$ known to contain infinitely many primes, suitably weighted. For example in [9] Heath-Brown compared the sequence of values of the binary cubic form $x^3 + 2y^3$ (weighted by multiplicity) and the sequence of values taken by the norm form of the cubic field $K = \mathbb{Q}(\sqrt[3]{2})$.

We shall consider two nonnegative sequences $\mathcal{A} = (a_n)$, $\mathcal{B} = (b_n)$ supported on positive integers $n \leq X$, and put

$$\pi(\mathcal{A}) = \sum_p a_p \quad \text{and} \quad \pi(\mathcal{B}) = \sum_p b_p, \tag{3-1}$$

where the summations run over primes. If one establishes an asymptotic relation of the form

$$\pi(\mathcal{A}) = \varkappa \pi(\mathcal{B})(1 + o(1))$$

say, then an asymptotic formula for $\pi(\mathcal{B})$ implies an asymptotic formula for $\pi(\mathcal{A})$. In particular, this allows us to avoid working through the difficult harmonic analysis in [5], and allows one to work with estimates that apply to general complex sequences rather than relying on properties of the Möbius function.

To simplify matters, we will restrict the variable of interest, namely ℓ, to a short interval of the shape $I(X) = (X^*, (1+\eta)X^*]$ where $\eta \asymp (\log X)^{-1}$ and $X^{1/2}(\log X)^{-4} \leq X^* \leq c_f X^{1/2}$ where

$$c_f = \begin{cases} \sup_{f(x,y)\leq 1} y & \text{if } f \text{ is definite,} \\ 1 & \text{if } f \text{ is indefinite.} \end{cases}$$

We then define

$$a_n = \sum_{\substack{f(m,\ell)=n \\ \ell \in I(X)}} \mathfrak{Z}(\ell) \tag{3-2}$$

and

$$b_n = \sum_{\substack{f(m,\ell)=n \\ \ell \in I(X)}} \Lambda(\ell). \tag{3-3}$$

Here

$$\mathfrak{Z}(\ell) = \begin{cases} 2p \log p & \text{if } \ell = p^2, \\ 0 & \text{otherwise,} \end{cases} \tag{3-4}$$

and Λ is the von Mangoldt function. In the definite case Lam, Schindler, and the author proved that $\pi(\mathcal{B})$ satisfies an asymptotic formula. We will extend this to the indefinite, irreducible case.

One notes that the sequences (a_n), (b_n) introduced in (3-2) and (3-3) are analogous to the sequences introduced in [10]. The analogous sequences $\mathcal{A}^\spadesuit$, $\mathcal{B}^\spadesuit$ for the purpose of Theorem 1.1 are

$$a_n^\spadesuit = \sum_{\substack{f(m,\ell)=n \\ \ell \in I(X)}} \mathfrak{Z}^\spadesuit(\ell) \text{ and } b_n^\spadesuit = \sum_{\substack{f(m,\ell)=n \\ \ell \in I(X)}} 1, \tag{3-5}$$

respectively, where

$$\mathfrak{Z}^\spadesuit(\ell) = \begin{cases} 2k & \text{if } \ell = k^2, \\ 0 & \text{otherwise.} \end{cases} \tag{3-6}$$

We emphasize that the integer k appearing in (3-6) is not required to be prime, unlike in (3-4).

Having established the asymptotic formula for $\pi(\mathcal{B})$, $\pi(\mathcal{B}^\spadesuit)$, we will then prove an analogue of Proposition 1 in [10]. In [10] they introduced the quantity

$$\mu(I) = \int_I \sqrt{X - t^2}\, dt = \int_I \int_0^{\sqrt{X-t^2}} ds\, dt.$$

In other words, $\mu(I)$ is the area of the subset of the positive half-disk with y-coordinate restricted to I. We generalize this definition to

$$\mu_f(I) = \text{Area}\{(x, y) \in \mathbb{R}^2 : 0 < f(x, y) < X,\ y \in I(X)\} = \int_I \int_{0 < f(s,t) < X} ds\, dt. \tag{3-7}$$

Observe that $\mu_f(I) \ll_f \sqrt{X} \cdot |I|$, where $|I|$ is the length of I. This brings us to the following statement:

Proposition 3.1. *Let $\mathcal{A} = (a_n)$, $\mathcal{B} = (b_n)$ be given as in* (3-2) *and* (3-3)*. Then we have the asymptotic relation*

$$|\pi(\mathcal{A}) - \pi(\mathcal{B})| \ll_\varepsilon \frac{\mu_f(I) \log\log X}{(\log X)^2}.$$

Similarly, for $\mathcal{A}^\spadesuit$, $\mathcal{B}^\spadesuit$ given by (3-5) *one has*

$$|\pi(\mathcal{A}^\spadesuit) - \pi(\mathcal{B}^\spadesuit)| \ll_\varepsilon \frac{\mu_f(I) \log\log X}{(\log X)^2}.$$

We will see that this is enough to prove Theorems 1.1 and 1.2 as in the proof of Theorem 1 from Proposition 1 in [10]. First we will prove that

$$\pi(\mathcal{B}) = \frac{\nu_f \mu_f(I)}{\log X}\left(1 + O\left(\frac{1}{\log X}\right)\right), \tag{3-8}$$

this following from Theorem 1.3 via partial summation. In the case of $\mathcal{B}$ and f is definite we start with the asymptotic formula (1-10) and write it as

$$\sum_{q \le X} \Lambda(q) \sum_{f(m,\ell)=q} \Lambda(\ell) = \nu_f \mathfrak{S}'_f X + O_A(X(\log X)^{-A}).$$

Writing $\Psi(q) = \sum_{f(m,\ell)=q} \Lambda(\ell)$ and replacing $\Lambda(q)$ with $\log q$ (supported on primes), we have by partial summation

$$\log X \sum_{q \le X} \Psi(q) - \int_1^X \frac{1}{t}\left(\sum_{q \le t} \Psi(q)\right) dt = \nu_f \mathfrak{S}'_f X + O_A(X(\log X)^{-A}).$$

An upper bound sieve gives that

$$\sum_{q \le X} \Psi(q) = O\left(\frac{X}{\log X}\right),$$

hence

$$\log X \sum_{q \le X} \Psi(q) = \nu_f \mathfrak{S}'_f X + O_A(X(\log X)^{-A}) + O\left(\int_1^X \frac{dt}{\log t}\right),$$

and thus

$$\sum_{q \le X} \Psi(q) = \frac{\nu_f \mathfrak{S}'_f X}{\log X}\left(1 + O\left(\frac{1}{\log X}\right)\right).$$

By replacing $\Psi(q)$ with

$$\Psi'(q) = \sum_{\substack{f(m,\ell)=q \\ \ell \in I(X)}} \Lambda(\ell),$$

we see from the same argument that

$$\sum_{q \le X} \Psi'(q) = \frac{\nu_f \mu_f(I)}{\log X}\left(1 + O\left(\frac{1}{\log X}\right)\right),$$

as desired. The same argument applies to the indefinite case, following (1-11).

Thus Proposition 3.1 gives

$$\pi(\mathcal{A}) = \frac{\nu_f \mu_f(I)}{\log X}\left(1 + O\left(\frac{\log\log X}{\log X}\right)\right). \tag{3-9}$$

We then proceed by partial summation as in [10]. We consider intervals $I_j = (X_j, X_j(1+\eta)]$ that form a partition of $(X^{1/2}(\log X)^{-4}, c_f X^{1/2}]$. Here $\eta \asymp (\log X)^{-1}$ is chosen so we have an exact partition. We

let $\mathcal{A}_j$ be defined as in (3-2) with $I(X) = I_j$. The number of pairs (a, p) with $0 < f(a, p^2) \le X$ and $p|a$, $p^2 \in I$ is bounded by

$$\sum_{p^2 \in I} \frac{\sqrt{X}}{p} \ll_\varepsilon X^{1/2+\varepsilon}.$$

It follows that

$$\begin{aligned}
&\#\{(a, p) : 0 < f(a, p^2) \le X \text{ is prime }, \ p \text{ is prime}, \ p \le X^{1/4}\} \\
&\quad = \sum_j \frac{1}{\sqrt{X_j}\log X_j}\pi(\mathcal{A}_j)\left(1 + O\left(\frac{1}{\log X}\right)\right) + O\left(\frac{X^{3/4}}{(\log X)^3}\right) \\
&\quad = \frac{\nu_f + O((\log X)^{-1}\log\log X)}{(\log X)^2}\sum_j \frac{\mu_f(I_j)}{\sqrt{X_j}} + O\left(\frac{X^{3/4}}{(\log X)^3}\right) \\
&\quad = \frac{\nu_f + O((\log X)^{-1}\log\log X)}{(\log X)^2}\int_{\sqrt{X}/(\log X)^4}^{\sqrt{X}} \frac{1}{\sqrt{t}}\int_{0<f(s,t)<X} ds\, dt + O\left(\frac{X^{3/4}}{(\log X)^3}\right) \\
&\quad = \frac{\nu_f \mathfrak{S}_f X^{3/4}}{(\log X)^2}\left(1 + O\left(\frac{\log\log X}{\log X}\right)\right).
\end{aligned}$$

Thus Theorem 1.2 follows from Proposition 3.1. Next we do something similar to deduce Theorem 1.1. In this case it is trivial that

$$\pi(\mathcal{B}^\spadesuit) = \frac{\nu_f \mu_f(I)}{\log X}\left(1 + O\left(\frac{1}{\log X}\right)\right),$$

since this is a direct consequence of Landau's prime ideal theorem. Therefore Proposition 3.1 gives

$$\pi(\mathcal{A}^\spadesuit) = \frac{\nu_f \mu_f(I)}{\log X}\left(1 + O\left(\frac{1}{\log X}\right)\right).$$

We then proceed by partial summation as above, but noting that the weight is $2k$ rather than $2p\log p$. The same calculation then gives

$$\#\{(a, b) : 0 < f(a, b^2) \le X \text{ is prime}, \ b \le X^{1/4}\} = \frac{\nu_f \mathfrak{S}_f X^{3/4}}{\log X}\left(1 + O\left(\frac{1}{\log X}\right)\right),$$

which suffices to prove Theorem 1.1.

In order to establish Proposition 3.1 we apply the same sieve procedure to the pairs $(\mathcal{A}, \mathcal{B})$ and $(\mathcal{A}^\spadesuit, \mathcal{B}^\spadesuit)$, producing cancellation at key junctures and upper bounding the rest. For any complex sequence $\mathcal{C} = (c_n)$ supported on the positive integers put

$$S(\mathcal{C}, Z) = \sum_{\substack{n \in \mathbb{N} \\ p|n \Rightarrow p > Z}} c_n$$

and for each $d \in \mathbb{N}$ put

$$\mathcal{C}_d = \{c_{dn} : n \in \mathbb{N}\}.$$

We fix

$$\delta_1 = \delta_1(X) = (\log X)^{\varpi - 1} \quad \text{and} \quad \delta_2 = \delta_2(X) = \frac{A_1 \log\log X}{\log X} \tag{3-10}$$

for some large positive number A_1 and small number $0 < \varpi < 1$ which we specify later. In [10] they have a single parameter δ which is equal to our δ_1. The reason why we are having two separate parameters is to obtain the superior error term in Theorem 1.2 and the error term in Theorem 1.1.

We also fix $Y > X^{1/3}$, where the specific choice of Y will be made when it is relevant. Now put

$$S_1(\mathcal{C}) = S(\mathcal{C}, X^{\delta_1}), \quad S_2(\mathcal{C}) = \sum_{X^{\delta_1} \le p < Y} S(\mathcal{C}_p, p), \quad S_3(\mathcal{C}) = \sum_{Y \le p < X^{1/2-\delta_2}} S(\mathcal{C}_p, p). \tag{3-11}$$

The astute reader will note that $S_1(\mathcal{C})$ is readily handled by the fundamental lemma of sieve theory, giving an asymptotic formula; see, for example, Corollary 6.10 in [6]. By Buchstab's identity, we have

$$\pi(\mathcal{C}) = S(\mathcal{C}, X^{1/2}) = S_1(\mathcal{C}) - S_2(\mathcal{C}) - S_3(\mathcal{C}) - \sum_{X^{1/2-\delta_2} \le p \le X^{1/2}} S(\mathcal{C}_p, p).$$

The last sum can be handled by Selberg's upper bound sieve, and we conclude:

Lemma 3.2. *For $Y = X^{17/48}$ and $\mathcal{C} = \mathcal{A}, \mathcal{B}, \mathcal{A}^\spadesuit, \mathcal{B}^\spadesuit$ we have*

$$\pi(\mathcal{C}) = S_1(\mathcal{C}) - S_2(\mathcal{C}) - S_3(\mathcal{C}) + O\left(\frac{\delta_2 \mu_f(I)}{\log X}\right).$$

We will see that $S_3(\mathcal{C})$ can be written in terms of appropriate bilinear forms, but $S_2(\mathcal{C})$ will require further treatment. Let us put

$$T^{(n)}(\mathcal{C}) = \sum_{\substack{X^{\delta_1} \le p_n < \cdots < p_1 < Y \\ p_1 \cdots p_n < Y}} S(\mathcal{C}_{p_1 \cdots p_n}, X^{\delta_1})$$

and

$$U^{(n)}(\mathcal{C}) = \sum_{\substack{X^{\delta_1} \le p_{n+1} < \cdots < p_1 < Y \\ p_1 \cdots p_n < Y \le p_1 \cdots p_{n+1}}} S(\mathcal{C}_{p_1 \cdots p_{n+1}}, p_{n+1}).$$

We then have:

Lemma 3.3. *For $n_0 = \left\lfloor \frac{\log Y}{\delta_1 \log X} \right\rfloor$ we have*

$$S_2(\mathcal{C}) = \sum_{1 \le n \le n_0} (-1)^{n-1} (T^{(n)}(\mathcal{C}) - U^{(n)}(\mathcal{C})).$$

The sums

$$|S_1(\mathcal{A}) - S_1(\mathcal{B})|, \quad |S_1(\mathcal{A}^\spadesuit) - S_1(\mathcal{B}^\spadesuit)| \tag{3-12}$$

and

$$\sum_{1 \le n \le n_0} |T^{(n)}(\mathcal{A}) - T^{(n)}(\mathcal{B})|, \quad \sum_{1 \le n \le n_0} |T^{(n)}(\mathcal{A}^\spadesuit) - T^{(n)}(\mathcal{B}^\spadesuit)| \tag{3-13}$$

can be handled by our type-I estimate Proposition 5.1 and the fundamental lemma; see Lemma 2 in [10]. To control these sums it suffices to prove:

Proposition 3.4. *Let $\mathfrak{Q}$ be a set of square-free numbers not exceeding $Y = X^{17/48}$. Then for any $A > 0$ we have*

$$\left|\sum_{q\in\mathfrak{Q}} S(\mathcal{A}_q, X^{\delta_1}) - \sum_{q\in\mathfrak{Q}} S(\mathcal{B}_q, X^{\delta_1})\right| \ll_A \frac{X}{(\log X)^A}$$

and

$$\left|\sum_{q\in\mathfrak{Q}} S(\mathcal{A}_q^{\spadesuit}, X^{\delta_1}) - \sum_{q\in\mathfrak{Q}} S(\mathcal{B}_q^{\spadesuit}, X^{\delta_1})\right| \ll_A \frac{X}{(\log X)^A}$$

By the definitions of $S_1(\mathcal{C})$ and $T^{(n)}(\mathcal{C})$, clearly Proposition 3.4 gives the bound of $O_A(X(\log X)^{-A})$ for both (3-12) and (3-13).

We now give a proof for Proposition 3.4, which depends on Proposition 5.1.

Proof of Proposition 3.4. The fundamental lemma allows us to give an asymptotic formula for the sum

$$\sum_{q\in\mathfrak{Q}} S(\mathcal{C}_q, X^{\delta_1})$$

for $\mathcal{C} = \mathcal{A}, \mathcal{B}, \mathcal{A}^{\spadesuit}, \mathcal{B}^{\spadesuit}$. Recall that $\delta_1 = (\log X)^{\varpi-1}$. Proposition 5.1 gives us a level of distribution of $X^{3/4}(\log X)^{-B}$ for some large B. We then apply an upper and lower bound sieve of level of distribution $X^{1/4}$, so that the sifting variable

$$s = \frac{\log D}{\log z} = \frac{\log X^{1/4}}{\log X^{\delta_1}} = \frac{1}{4\delta_1}.$$

We use the usual notation

$$V(z) = \prod_{p<z}(1-g(p)) = \prod_{p<z}\left(1-\frac{\rho_f(p)}{p}\right),$$

where $\rho_f(p)$ counts the number of linear factors of f modulo p; see (5-2). We then have

$$R_d(\mathcal{C}) = |A_d(\mathcal{C}) - M_d(\mathcal{C})|$$

with $M_d(\mathcal{C})$ as in Proposition 5.1. By Corollary 6.10 in [6] and applying Proposition 5.1 we obtain

$$\begin{aligned}\sum_{q\in\mathfrak{Q}} S(\mathcal{C}_q, X^{\delta_1}) &= V(X^{\delta_1})\sum_{q\in\mathfrak{Q}}\frac{\rho_f(q)}{q}\mu_f(I)\big(1+O\big(\exp(-(4\delta)^{-1})\big)\big) + O\bigg(\sum_{q\in\mathfrak{Q}}\sum_{d<X^{1/4}} R_{dq}(\mathcal{C})\bigg)\\ &= V(X^{\delta_1})\sum_{q\in\mathfrak{Q}}\frac{\rho_f(q)}{q}\mu_f(I)\bigg(1+O\bigg(\frac{1}{(\log X)^A}\bigg)\bigg) + O\bigg(\sum_{d<X^{3/4-1/8}}\tau(d)R_q(\mathcal{C})\bigg)\\ &= V(X^{\delta_1})\sum_{q\in\mathfrak{Q}}\frac{\rho_f(q)}{q}\mu_f(I)\bigg(1+O\bigg(\frac{1}{(\log X)^A}\bigg)\bigg) + O_A(X(\log X)^{-A})\end{aligned}$$

for any $A > 0$. The last line is independent of whether $\mathcal{C} = \mathcal{A}, \mathcal{B}, \mathcal{A}^\spadesuit$ or $\mathcal{C} = \mathcal{B}^\spadesuit$. Since $V(X^{\delta_1}) \le 1$ it follows that

$$\sum_{q\in\mathfrak{Q}}(S(\mathcal{A}_q, X^{\delta_1}) - S(\mathcal{B}_q, X^{\delta_1})) \ll_A \frac{1}{(\log X)^A}\mu_f(I)\sum_{q\in\mathfrak{Q}}\frac{\rho_f(q)}{q} + X(\log X)^{-A}$$
$$\ll_A X(\log X)^{-A+2},$$

since $\rho_f(q) \ll \tau(q)$. Likewise,

$$\sum_{q\in\mathfrak{Q}}(S(\mathcal{A}_q^\spadesuit, X^{\delta_1}) - S(\mathcal{B}_q^\spadesuit, X^{\delta_1})) \ll_A X(\log X)^{-A+2}. \qquad \square$$

Thus it remains to show that

$$|S_3(\mathcal{A}) - S_3(\mathcal{B})| \ll_A \frac{X}{(\log X)^A} \quad \text{and} \quad |U^{(n)}(\mathcal{A}) - U^{(n)}(\mathcal{B})| \ll_A \frac{X}{(\log X)^A} \quad \text{for } n \ge 3 \tag{3-14}$$

and

$$|U^{(n)}(\mathcal{A}) - U^{(n)}(\mathcal{B})| \ll \frac{\delta_2\mu_f(I)}{\log X} \quad \text{for } n = 1, 2, \tag{3-15}$$

with analogous statements for $\mathcal{A}^\spadesuit, \mathcal{B}^\spadesuit$.

We proceed to reduce the verification of (3-14) and (3-15) to a bilinear sum estimate.

Reduction to a bilinear sum bound. Let us write $U^{(1)}$ and $U^{(2)}$ into a more convenient form, as in [10]. To do so let us put

$$U_1^{(1)}(\mathcal{C}) = \sum_{\substack{X^{\delta_1}\le p_2<p_1<Y\\ Y\le p_1p_2<X^{1/2-\delta_2}}} S(\mathcal{C}_{p_1p_2}, p_2),$$
$$U_2^{(1)}(\mathcal{C}) = \sum_{\substack{X^{\delta_1}\le p_2<p_1<Y\\ p_1p_2\ge X^{1/2+\delta_2}}} S(\mathcal{C}_{p_1p_2}, p_2),$$
$$U_1^{(2)}(\mathcal{C}) = \sum_{\substack{X^{\delta_1}\le p_3<\cdots<p_1<Y\\ p_1p_2<Y\le p_1p_2p_3<X^{1/2-\delta_2}}} S(\mathcal{C}_{p_1p_2p_3}, p_3),$$
$$U_2^{(2)}(\mathcal{C}) = \sum_{\substack{X^{\delta_1}\le p_3<\cdots<p_1<Y\\ p_1p_2<Y\le p_1p_2p_3\\ p_1p_2p_3\ge X^{1/2+\delta_2}}} S(\mathcal{C}_{p_1p_2p_3}, p_3).$$

We now state Lemmas 6 and 7 from [10]. Their proofs apply equally well, but since for us δ_1, δ_2 are different we write out the proofs.

Lemma 3.5 [10, Lemma 6]. *For $\mathcal{C} = \mathcal{A}, \mathcal{B}$ we have $U^{(j)}(\mathcal{C})$ satisfies*

$$U^{(1)}(\mathcal{C}) = U_1^{(1)}(\mathcal{C}) + U_2^{(1)}(\mathcal{C}) + O\left(\frac{\delta_2\mu_f(I)}{\log X}\right) \tag{3-16}$$

and

$$U^{(2)}(\mathcal{C}) = U_1^{(2)}(\mathcal{C}) + U_2^{(2)}(\mathcal{C}) + O\left(\frac{\delta_2 \mu_f(I)}{\log X}\right). \tag{3-17}$$

Proof. To prove (3-16) it suffices to show

$$\sum_{\substack{X^{\delta_1} \le p_2 < p_1 < Y \\ X^{1/2-\delta_2} < p_1 p_2 \le X^{1/2+\delta_2}}} S(\mathcal{C}_{p_1 p_2}, p_2) \ll \frac{\delta_2 \mu_f(I)}{\log X}.$$

In the sum above we have

$$p_2 \ge \frac{X^{1/2-\delta_2}}{p_1} > \frac{X^{1/2-\delta_2}}{Y} > X^{1/10}$$

so we may apply Selberg's upper bound sieve and our level of distribution to obtain

$$\begin{aligned}\sum_{\substack{X^{\delta_1} \le p_2 < p_1 < Y \\ X^{1/2-\delta_2} < p_1 p_2 \le X^{1/2+\delta_2}}} S(\mathcal{C}_{p_1 p_2}, p_2) &\ll \sum_{\substack{X^{\delta_1} \le p_2 < p_1 < Y \\ X^{1/2-\delta_2} < p_1 p_2 \le X^{1/2+\delta_2}}} S(\mathcal{C}_{p_1 p_2}, X^{1/10}) \\ &\ll \frac{\mu_f(I)}{\log X} \sum_{\substack{X^{1/10} < p_2 < p_1 < Y \\ X^{1/2-\delta_2} < p_1 p_2 < X^{1/2+\delta_2}}} \frac{1}{p_1 p_2} \\ &\ll \frac{\delta_2 \mu_f(I)}{\log X}.\end{aligned}$$

The proof for (3-17) follows similarly. □

Lemma 3.6 [10, Lemma 7]. *Let κ be a positive number satisfying $X^{-\delta_1} \le \kappa \le 1$. Let N_1, N_2 be positive numbers in the interval $[X^{\delta}, X^{1/3}]$. We then have, for any $A > 0$,*

$$\sum_{N_1 \le p_1 \le (1+\kappa)N_1} \sum_{N_2 \le p_2 \le (1+\kappa)N_2} \sum_{n \equiv 0 \,(\mathrm{mod}\, p_1 p_2)} c_n \tau(n) \ll_A \kappa^2 X (\log X)^{2^{17}} + \frac{X}{(\log X)^A}.$$

For $k \ge 3$, the condition of summation in $U^{(k)}(\mathcal{C})$ is

$$X^{17/48} = Y \le p_1 \cdots p_{k+1} < (p_1 \cdots p_k)^{(k+1)/k} \le Y^{4/3} < X^{1/2-\delta_2}.$$

Therefore, upon defining

$$U_*^{(k)}(\mathcal{C}) = \sum_{X^{\delta_1} \le p_{k+1} < \cdots < p_1 \cdots p_k < Y \le p_1 \cdots p_{k+1} < X^{1/2-\delta}} S(\mathcal{C}_{p_1 \cdots p_{k+1}}, p_{k+1})$$

we have

$$S_3(\mathcal{C}) = U_*^{(0)}(\mathcal{C}),\ U_1^{(1)}(\mathcal{C}) = U_*^{(1)}(\mathcal{C}),\ U_1^{(2)}(\mathcal{C}) = U_*^{(2)}(\mathcal{C})$$

and

$$U^{(k)}(\mathcal{C}) = U_*^{(k)}(\mathcal{C}) \quad \text{for } k \ge 3.$$

If $p \in \mathcal{J} = [V, (1+\kappa)V)$ and n is an integer is counted by $S(\mathcal{C}_{pq}, V)$ but not by $S(\mathcal{C}_{pq}, p)$, then n has at least two prime factors p_1, p_2 in $\mathcal{J}$. In our application we will have $V \le X^{1/2-\delta_2}$ and $n \ge X(\log X)^{-8}$. Note that n has a divisor exceeding one and coprime to $p_1 p_2$. It follows that

$$V^3 \le n \le X.$$

A given integer n may be counted multiple times by $U_*^{(k)}(\mathcal{C})$ but the multiplicity is bounded by the number of choices for $p_{k+1} < \cdots < p_1$ all dividing n, and therefore the multiplicity is at most $\tau(n)$. Applying Lemma 3.6 and setting

$$\mathcal{J}(r) = [V_r, V_{r+1}) = [X^{\delta_1}(1+\kappa)^r, X^{\delta_1}(1+\kappa)^{r+1}), \quad r \ge 0$$

and $R \ll \kappa^{-1} \log X$ satisfying $X^{\delta_1}(1+\kappa)^R > X$, we obtain

$$U_*^{(k)}(\mathcal{C}) = \sum_{0 \le r \le R} \sum_{p \in \mathcal{J}(r)} \sum_{p < p_k < \cdots < p_1 \cdots p_k < Y \le p_1 \cdots p_k p < X^{1/2-\delta_2}} S(\mathcal{C}_{p_1 \cdots p_k p}, V_r) + O_A\left(\kappa X (\log X)^{1+2^{17}} + \kappa^{-1} \frac{X}{(\log X)^{A-1}}\right). \quad (3\text{-}18)$$

We need to make sure that both

$$\kappa X (\log X)^{1+2^{17}}, \quad \kappa^{-1} \frac{X}{(\log X)^{A-1}}$$

are $O(X(\log X)^{-A'})$ for some $A' > 1$. This compels us to choose

$$\kappa = (\log X)^{-A/2}.$$

This gives

$$\kappa X (\log X)^{1+2^{17}} = X(\log X)^{1+2^{17}-A/2} \quad \text{and} \quad \kappa^{-1} \frac{X}{(\log X)^{A-1}} = \frac{X}{(\log X)^{A/2-1}}. \quad (3\text{-}19)$$

This procedure allows us to reduce our proof to estimations of certain bilinear sums since

$$\sum_{p \in \mathcal{J}(r)} \sum_{p < p_k < \cdots < p_1 \cdots p_k < Y \le p_1 \cdots p_k p < X^{1/2-\delta_2}} S(\mathcal{C}_{p_1 \cdots p_k p}, V_r) = \sum_{m,n} \alpha_m^{(r)} \beta_n^{(r)} c_{mn}, \quad (3\text{-}20)$$

where $\alpha_m^{(r)}$ is the characteristic function for the integers m all of whose prime factors are at least V_r and $\beta_n^{(r)}$ is the characteristic function for integers $n = p_1 \cdots p_k p$ satisfying

$$p \in \mathcal{J}(r), \quad p < p_k < \cdots < p_1 < Y \quad \text{and} \quad p_1 \cdots p_k < Y \le p_1 \cdots p_k p < X^{1/2-\delta_2}.$$

Observe that $\beta_n^{(r)}$ is supported on integers $n \in [Y, X^{1/2-\delta_2})$.

The procedure for $U_2^{(1)}(\mathcal{C})$ and $U_2^{(2)}(\mathcal{C})$ will be somewhat different. We may use Lemma 3.6 to replace $S(\mathcal{C}_{p_1p_2}, p_2)$ in $U_2^{(1)}(\mathcal{C})$ by $S(\mathcal{C}_{p_1p_2}, V_r)$ when $p_2 \in J(r)$. This yields

$$U_2^{(1)}(\mathcal{C}) = \sum_{0\le r\le R} \sum_{p_2\in J(r)} \sum_{\substack{p_1\ge X^{1/2-\delta_2}/p_2\\ p_2<p_1<Y}} S(\mathcal{C}_{p_1p_2}, V_r) + O(\kappa X(\log X)^{1+2^{17}}) + O\left(\kappa^{-1}\frac{X}{(\log X)^{A-1}}\right).$$

The sum on the right can be expressed as

$$\sum_{0\le r\le R} \sum_{m,n} \alpha_m^{(r)}\beta_n^{(r)} c(mn),$$

where we now take $\alpha_m^{(r)}$ to be the characteristic function for numbers $m = p_1p_2$ with $p_2 \in J(r)$, $p_2 < p_1 < Y$ and $p_1p_2 \ge X^{1/2+\delta_2}$, and $\beta_n^{(r)}$ to be the characteristic function for those numbers n all of whose prime factors are at least V_r. Since $c(n)$ is supported in

$$((X^*)^2, c_f X] \subseteq (X(\log X)^{-8}, c_f X]$$

we may assume that $\beta_n^{(r)}$ is supported in

$$(X(\log X)^{-8}Y^{-2}, X^{1/2-\delta_2}] \subseteq (X^{1/4+1/48}, X^{1/2-\delta_2}].$$

This is sufficient for our purposes. We may handle $U_2^{(2)}(\mathcal{C})$ in an analogous fashion.

On setting $\kappa = (\log X)^{-A/2}$ we find that each of

$$S_3(\mathcal{C}), \quad U_1^{(1)}(\mathcal{C}), \quad U_2^{(1)}(\mathcal{C}), \quad U_1^{(2)}(\mathcal{C}), \quad U_2^{(2)}(\mathcal{C}), \quad \text{and} \quad U^{(k)}(\mathcal{C})$$

for $k \ge 3$ can be expressed as a sum of $O(R)$ bilinear sums as in (3-20), together with an error term of $O_A(X(\log X)^{1+2^{17}-A/2})$. Thus it will be sufficient to prove:

Proposition 3.7 (main bilinear sum estimate). *Let $\xi > 0$ and suppose $X^{1/4+\xi} \le N < X^{1/2-\delta_2}$. Suppose (α_m), (β_n) are two complex sequences having sup-norm at most 1 supported on natural numbers with no prime factors less than X^δ. Then for any $A > 0$ we have*

$$\sum_{N<n\le 2N} \sum_{m<X/N} \alpha_m\beta_n(a_{mn} - b_{mn}) \ll_{A,\xi} \frac{X}{(\log X)^A} \tag{3-21}$$

and

$$\sum_{N<n\le 2N} \sum_{m<X/N} \alpha_m\beta_n(a^\spadesuit_{mn} - b^\spadesuit_{mn}) \ll_{A,\xi} \frac{X}{(\log X)^A} \tag{3-22}$$

It will be important that the sequences $\{\alpha_m\}$, $\{\beta_n\}$ are supported on those numbers whose prime factors all exceed X^{δ_1}, and in particular, they are supported on odd numbers.

The remainder of the paper is devoted to proving Proposition 3.7. In particular, Propositions 7.5 and 7.6 will imply Proposition 3.7. In order to get there, we need to decompose the terms c_{mn} for any positive integers m, n into components that resemble c_m, c_n. This turns out to be somewhat delicate and we will

require the composition laws of the ideals of $\mathcal{O}_K$, expressed in terms of ideal numbers. This will be the primary focus of the next section.

4. Algebraic characterization of the multiplicative structure in terms of ideal numbers

The main purpose of this section is obtain an analogue of Proposition 2.3 in [12]. However, instead of using an explicit Dirichlet composition law as in [12] we will instead adopt the language of Hecke's *ideal numbers* as in [11].

Choose ideals $\mathfrak{a}_1, \ldots, \mathfrak{a}_t$ whose classes generate the ideal class group of $\mathcal{O}_K$ so that every fractional ideal $\mathfrak{a} \subseteq \mathcal{O}_K$ has a unique decomposition

$$\mathfrak{a} = (\alpha)\mathfrak{a}_1^{\ell_1} \cdots \mathfrak{a}_t^{\ell_t}$$

where $\alpha \in K^*$ and $\ell_j \in \mathbb{Z}$ with $0 \leq \ell_j < h_j$, with h_j the smallest positive integer such that $\mathfrak{a}_j^{h_j} = (\alpha_j)$ is principal. Then the class number $h(K)$ of $\mathcal{O}_K$ is equal to

$$h(K) = \prod_{j=1}^{t} h_j. \tag{4-1}$$

Let us choose complex numbers $b_1, \ldots, b_t$ so that

$$b_j^{h_j} = \alpha_j \quad \text{for } j = 1, \ldots, t,$$

and $b_j^{(i)}$ are complex numbers such that

$$(b_j^{(i)})^{h_j} = \alpha_j^{(i)} \quad \text{for } i = 1, 2.$$

Now put $L = K(b_1, \ldots, b_t)$ and $\mathfrak{J}(K)^*$ the subgroup of L^* generated by K^* and $\{b_j : 1 \leq j \leq t\}$. Then $\mathfrak{J}(K) = \{0\} \cup \mathfrak{J}(K)^*$ is the domain of *ideal numbers* of K. The quotient group $\mathfrak{J}(K)^*/\mathcal{O}_K^*$ is then isomorphic to the group of fractional ideals of K. Each $\gamma \in \mathfrak{J}(K)$ corresponds a unique fractional ideal $J(\gamma)$; the norm of the ideal $J(\gamma)$ is given by the product

$$N(\gamma) = N(J(\gamma)) = \gamma^{(1)}\gamma^{(2)}.$$

Note that $\gamma^{(2)}$ is the algebraic conjugate of $\gamma^{(1)} = \gamma$.

Further, we have $J(\gamma)$ is an integral ideal of $\mathcal{O}_K$ if and only if $\gamma \in \mathcal{O}_L$.

We thus have a correspondence between the ideal classes of $\mathcal{O}_K$ and a subset of integers in $\mathcal{O}_L$. Indeed, we can say that $\gamma, \gamma' \in \mathfrak{J}(K)$ belong to the same class if and only if the corresponding ideals $J(\gamma), J(\gamma') \subseteq \mathcal{O}_K$ are in the same ideal class. It follows that we may partition $\mathfrak{J}(K)$ into $h(K)$ classes, corresponding to the ideal classes of $\mathcal{O}_K$. Such a class of ideal numbers, say A, has an integral basis $\{w_1, w_2\}$ such that

$$A = \{a_1 w_1 + a_2 w_2 : (a_1, a_2) \in \mathbb{Q}^2\}$$

and

$$A \cap \mathcal{O}_L = \{a_1 w_1 + a_2 w_2 : (a_1, a_2) \in \mathbb{Z}^2\}.$$

Indeed, this follows by noting that

$$A = \{\gamma \beta_1^{\ell_1} \cdots \beta_t^{\ell_t} : \gamma \in K\},$$

where $(\ell_1, \ldots, \ell_t)$ is a fixed tuple of nonnegative integers. If (v_1, v_2) is an integral basis of K, then we may take

$$w_1 = v_1 \beta_1^{\ell_1} \cdots \beta_t^{\ell_t}, \quad w_2 = v_2 \beta_1^{\ell_1} \cdots \beta_t^{\ell_t}.$$

Further, the discriminant of A, viewed as a $\mathbb{Z}$-lattice, is equal to $\Delta(K)$. Moreover for any basis $\{w_1, w_2\}$ of A and $\alpha \in A \setminus \{0\}$ we have that $\{\alpha^{-1} w_1, \alpha^{-1} w_2\}$ is a basis of $K/\mathbb{Q}$. This implies that there is a unique *dual basis* $\{\widetilde{w_1}, \widetilde{w_2}\}$ of A^{-1} defined by the condition

$$\mathrm{Tr}(w_i \widetilde{w_j}) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{4-2}$$

We use the notation $\mathrm{Cl}\,\mathfrak{a}$, $\mathrm{Cl}\,\alpha$ for the ideal class of the integral ideal $\mathfrak{a} \subset \mathcal{O}_K$ and the class of ideal numbers of the ideal number α.

Next we show that there is a correspondence between rank-two submodules of $\mathcal{O}_K$ and $\mathrm{SL}_2(\mathbb{Z})$-equivalence classes of irreducible integral binary quadratic forms having splitting field K. To establish this correspondence, first start with a rank-two submodule

$$\Lambda = \{a_1 \omega_1 + a_2 \omega_2 : a_1, a_2 \in \mathbb{Z}\}$$

with $\omega_1, \omega_2 \in \mathcal{O}_K$. Put

$$\mathfrak{S} = \gcd\{N_{K/\mathbb{Q}}(\omega_1 x + \omega_2) : x, y \in \mathbb{Z}\}.$$

Then the form

$$g(x, y) = N_{K/\mathbb{Q}}(\omega_1 x + \omega_2 y)\mathfrak{S}^{-1} \tag{4-3}$$

is an irreducible integral binary quadratic form with splitting field K.

Conversely, take an arbitrary irreducible integral binary quadratic form g which splits over K. Then there exists an integral nonsingular matrix M such that

$$g(x, y) = g^*((x, y)M),$$

where g^* is a primitive integral binary quadratic form with discriminant equal to $\Delta(K)$. Gauss's composition law then implies that g^* corresponds to an ideal class α, and in particular, can be expressed in the form

$$g^*(x, y) = N_{K/\mathbb{Q}}(\alpha_1 x + \alpha_2 y)N(\alpha^{-1})$$

with $\alpha = (\alpha_1, \alpha_2)$. Viewing α as a $\mathbb{Z}$-module and applying the transformation induced by M then gives the form g.

Now let $\mathfrak{f}$ be the $\mathbb{Z}$-module associated to f with basis $\{\nu_1, \nu_2\}$ so that

$$f(x, y) = N_{K/\mathbb{Q}}(\nu_1 x + \nu_2 y) N(\mathfrak{d}(f)^{-1}), \tag{4-4}$$

where $\mathfrak{d}(f) = (\nu_1, \nu_2)$ is the ideal generated by ν_1, ν_2. Let ψ_f be the ideal number of the ideal $\mathfrak{d}(f)$. Having identified $\mathfrak{f}$ we define the set of ideals

$$\mathfrak{A}(f) = \{(\nu_1 a_1 + \nu_2 a_2)\mathfrak{d}(f)^{-1} : a_1, a_2 \in \mathbb{Z},\ \gcd(a_1, a_2) = 1\}.$$

We now put $\mathcal{L}$ for the set of ideals in $\mathcal{O}_K$ which are not divisible by a rational prime. An integral ideal number $\gamma \in \mathfrak{I}(K)$ is said to be primitive if $J(\gamma) \in \mathcal{L}$. If K is a real quadratic field, put $\mathcal{L}_0$ the set of primitive ideal numbers γ satisfying the condition

$$\gamma = (N_{L/\mathbb{Q}}(\gamma))^{1/2} \varepsilon_0^z, \quad -\tfrac{1}{2} < z \leq \tfrac{1}{2},\ \gamma > 0,$$

where $\varepsilon_0 > 1$ is a fundamental unit of $\mathcal{O}_K$. If K is an imaginary quadratic field we may simply take $\mathcal{L}_0$ to be the set of primitive ideal numbers.

We now want to use the above discussion to obtain a meaningful decomposition for

$$c_n = \sum_{\substack{f(m,\ell)=n \\ \ell \in I(X)}} \Upsilon(\ell). \tag{4-5}$$

We follow the setup in [11] and introduce, for a given primitive vector $\boldsymbol{u} = (u_1, u_2)$ let $\mathfrak{F}(\boldsymbol{u})$ be the ideal in $\mathfrak{A}(f)$ given by $(\nu_1 u_1 + \nu_2 u_2) N(\mathfrak{d}(f)^{-1})$. We now put

$$\mathfrak{R}(X; n) = \{(u_1, u_2) \in \mathbb{Z}^2 : u_2 \in I(X),\ f(u_1, u_2) = n\}.$$

Note that $\mathfrak{R}(X; n)$ is finite for all $X > 0$ and $n \in \mathbb{Z}$. We then have

$$c_n = \sum_{\boldsymbol{u} \in \mathfrak{R}(X; n)} \Upsilon(u_2).$$

Via the correspondence

$$(u_1, u_2) \mapsto (\nu_1 u_1 + \nu_2 u_2) N(\mathfrak{d}(f)^{-1}) = \mathfrak{F}(u_1, u_2)$$

$\mathfrak{R}(X; n)$ corresponds to a set of ideals. For a given integer mn we then see that each element (u_1, u_2) of $\mathfrak{R}(X; mn)$ corresponds to a set of ideal factorizations of the form

$$\mathfrak{m}\mathfrak{n} = \mathfrak{F}(u_1, u_2) \tag{4-6}$$

with $N(\mathfrak{m}) = m$, $N(\mathfrak{n}) = n$. Now associate to $\mathfrak{m}$, $\mathfrak{n}$ ideal numbers $m^*, n^* \in \mathcal{L}_0$. Then (4-6) can be interpreted as multiplication in the set of ideal numbers. To make this concrete, first choose $\{w_1, w_2\}$ to be a basis for the ideal class $\mathrm{Cl}\,\mathfrak{d}(f)^{-1}$ such that $w_1 \psi_f^{-1} = z\nu_1$ and $w_2 \psi_f^{-1} = \nu_2$ for some integer z.

For each pair of ideal classes A, $B = A^{-1}\operatorname{Cl} f$ and any bases $\{a_1, a_2\}$, $\{b_1, b_2\}$ of A, B, respectively, we have a composition law

$$(a_1x_1 + a_2x_2)(b_1y_1 + b_2y_2) = \psi_f^{-1}(w_1R_{A,B}(\boldsymbol{x};\boldsymbol{y}) + w_2Q_{A,B}(\boldsymbol{x};\boldsymbol{y})).$$

By our choice of $\{w_1, w_2\}$ this is equivalent to

$$(a_1x_1 + a_2x_2)(b_1y_1 + b_2y_2) = z\nu_1R_{A,B}(\boldsymbol{x};\boldsymbol{y}) + \nu_2Q_{A,B}(\boldsymbol{x};\boldsymbol{y}).$$

This gives a bilinear mapping

$$\Phi_{A,B} : (\mathcal{L}_0 \cap A) \times (\mathcal{L}_0 \cap B) \to \{(x, y) \in \mathbb{R}^2 : y \in I(X)\},$$
$$\Phi_{A,B}(m_1, m_2; n_1, n_2) = (R_{A,B}(\boldsymbol{m};\boldsymbol{n}),\ Q_{A,B}(\boldsymbol{m};\boldsymbol{n}))$$

say. Let us write $A_0 = A \cap \mathcal{L}_0$ and $B_0 = B \cap \mathcal{L}_0$ for convenience. We then have

$$c_{mn} = \sum_{A\cdot B = \operatorname{Cl} f} \sum_{\substack{\boldsymbol{m}\in A,\ \boldsymbol{n}\in B\\ N(\boldsymbol{m})=m,\ N(\boldsymbol{n})=n\\ Q_{A,B}(\boldsymbol{m};\boldsymbol{n})\in I(X)}}\sum \Upsilon(Q_{A,B}(\boldsymbol{m};\boldsymbol{n})). \tag{4-7}$$

This is the desired analogue to equation (5.2) in [5]. We summarize this below:

Proposition 4.1. *For $\mathcal{C} = \mathcal{A}, \mathcal{B}, \mathcal{A}^{\spadesuit}, \mathcal{B}^{\spadesuit}$, (4-7) holds.*

5. Type-I estimates

We will establish the necessary type-I estimate we need, following the work of Friedlander and Iwaniec in [6]. For this section, we shall put $\lambda(\ell)$ to be any function bounded by one supported on r-th powers of integers, and put

$$a_n = \sum_{\substack{f(\ell,m)=n\\ \ell\in I(X)}} \lambda(\ell). \tag{5-1}$$

We recall that

$$A_d(X) = \sum_{\substack{n\le X\\ n\equiv 0 \,(\mathrm{mod}\, d)}} a_n.$$

For a given positive integer ℓ put

$$\mathcal{I}(\ell; X) = \{x \in \mathbb{R}^2 : 0 < f(x, \ell) < X\}$$

and $\iota(\ell; X)$ to be the length of $\mathcal{I}(\ell; X)$. We then expect $A_d(X)$ to be well-approximated by

$$M_d(X) = \frac{\rho_f(d)}{d} \sum_{\substack{\ell\in I(X)\\ \gcd(\ell,d)=1}} \lambda(\ell)\frac{\varphi(\ell)}{\ell}\iota(\ell; X),$$

where φ is the Euler totient function and $\rho_f(d)$ is the number of solutions to the congruence

$$f(x, 1) \equiv 0 \ (\mathrm{mod}\, d). \tag{5-2}$$

Our goal is to establish:

Proposition 5.1. *Suppose that λ is supported on r-th powers. Then uniformly for $X^{1/2} \le D \le X^{(r+1)/(2r)}$ we have*

$$\sum_{d \le D} |A_d(X) - M_d(X)| \ll D^{1/4} X^{3(r+1)/(8r)} (\log X)^{24}.$$

As usual, our starting point is the following result from [1], which states that the roots of quadratic congruences are separated as much as possible:

Proposition 5.2 [1, Proposition 3]. *Let $F(x, y) = \alpha x^2 + \beta xy + \gamma y^2 \in \mathbb{Z}[x, y]$ be an arbitrary binary quadratic form whose discriminant is not a perfect square. For any sequence (α_n) of complex numbers and positive real numbers D, N we have*

$$\sum_{D \le d \le 2D} \sum_{F(1,\nu) \equiv 0 \,(\mathrm{mod}\, d)} \left| \sum_{n \le N} \alpha_n e\left(\frac{\nu n}{d}\right) \right|^2 \ll_F (D + N) \sum_n |\alpha_n|^2.$$

It is the fact that such a strong large sieve inequality exists for roots of quadratic congruences that enables such powerful results to be proved about thin variables as in [3; 5; 10]. We show how to derive the type-I estimates we need by following the same steps carried out in [6; 12]. We first replace $A_d(X)$, $M_d(X)$ with their smooth counterparts. Consider an auxiliary smooth function $\phi : \mathbb{R} \to \mathbb{R}$ satisfying

(1) $\phi(u) = 1$ if $0 < u \le X - Y$;

(2) $\phi^{(j)}(u) \ll Y^{-j}$ for $j \ge 0$; and

(3) $\phi(u) = 0$ if $u \ge X$.

Here $X^{7/8} \le Y \le X$ will be chosen later. We then introduce (by abuse of notation)

$$A_d(\phi) = \sum_{n \equiv 0 \,(\mathrm{mod}\, d)} a_n \phi(n) \tag{5-3}$$

and

$$M_d(\phi) = \frac{\rho_f(d)}{d} \sum_{\gcd(\ell, d) = 1} \lambda(\ell) \frac{\varphi(d)}{d} \int_0^\infty \phi(f(\ell, t))\, dt. \tag{5-4}$$

We estimate the differences by elementary means as follows. Note that

$$\sum_{d \le D} |A_d(X) - A_d(\phi)| \le \sum_{\substack{X - Y < f(m,\ell) \le X \\ \gcd(\ell, m) = 1}}{}' \lambda(\ell) \tau(f(m, \ell)) + O\big(\sqrt{X} \log X\big),$$

where $\sum'$ means that the terms with a value of ℓ closest to $\sqrt{X}$ are omitted. We then have the following consequence of Landreau's inequality [13], resulting in the bound

$$\sum_{\ell \ll \sqrt{X}}{}' \sum_{\substack{d \le X^{1/4} \\ \gcd(d, \ell) = 1}} \tau(d)^8 \sum_{\substack{X - Y < f(m,\ell) \le X \\ f(m,\ell) \equiv 0 \,(\mathrm{mod}\, d)}} 1.$$

The conditions

$$X - Y < f(m,\ell) \le X \quad \text{and} \quad \ell \in I(X)$$

imply that m is restricted to an interval of length $O_f\big(Y/\sqrt{X+\ell^2}\big)$. Splitting into residue classes $m \equiv \alpha\ell \pmod d$ with α running over the roots of (5-2) we see that the above sum is bounded by

$$O\bigg(Y\bigg(\sum_{d\le X^{1/4}} \tau(d)^8 \frac{\rho_f(d)}{d}\bigg)\bigg(\sum_{\ell\ll\sqrt{X}}{}' |\lambda(\ell)|\,(X+\ell^2)^{-1/2}\bigg) + X^{1/4+1/(2r)}(\log X)^{256}\bigg).$$

We have the bounds

$$\sum_{d\le X^{1/4}} \tau(d)^8 \frac{\rho_f(d)}{d} \ll (\log X)^{256}$$

and

$$\begin{aligned}\sum_{\ell\ll\sqrt{X}}{}' |\lambda(\ell)|(X+\ell^2)^{-1/2} &\le \sum_{k\ll X^{1/2r}}{}' (X+k^{2r})^{-1/2}\\ &\ll X^{(1-2r)/(4r)} \sum_{k\ll X^{1/2r}}{}' (X^{1/(2r)}+k)^{-1/2}\\ &\ll X^{(1-2r+1)/(4r)} = X^{(1-r)/(2r)}.\end{aligned}$$

It follows that

$$\sum_{d\le D} |A_d(X) - A_d(\phi)| \ll Y X^{(1-r)/(2r)} (\log X)^{256}. \tag{5-5}$$

Similarly, we obtain

$$\sum_{d\le D} |M_d(X) - M_d(\phi)| \ll Y X^{(1-r)/(2r)} (\log X)^{256}. \tag{5-6}$$

We then proceed to decompose $A_d(\phi)$ as

$$\begin{aligned}A_d(\phi) &= \sum_{\substack{f(m,\ell)\equiv 0 \,(\mathrm{mod}\, d)\\ \gcd(\ell,m)=1}}\sum \lambda(\ell)\phi(f(m,\ell))\\ &= \sum_{f(\alpha,1)\equiv 0\,(\mathrm{mod}\,d)} \sum_{\ell} \lambda(\ell) \sum_{\substack{m\equiv\alpha\ell\,(\mathrm{mod}\,d)\\ \gcd(\ell,m)=1}} \phi(f(\ell,m))\\ &= \sum_{f(\alpha,1)\equiv 0\,(\mathrm{mod}\,d)} \sum_{a} \mu(a) \sum_{\ell} \lambda(a\ell) \sum_{m\equiv\alpha\ell\,(\mathrm{mod}\, d/\gcd(a,d))} \phi(a^2 f(m,\ell)),\end{aligned} \tag{5-7}$$

where we applied Möbius inversion to the inner sum to remove the awkward coprimality condition. We then apply Poisson's formula to the inner sum to obtain

$$\sum_{m\equiv\alpha\ell\,(\mathrm{mod}\, d/\gcd(a,d))} \phi(a^2 f(m,\ell)) = \frac{\gcd(a,d)}{d} \sum_{h\in\mathbb{Z}} e\bigg(\alpha h\ell \frac{\gcd(a,d)}{d}\bigg) \Phi_{a\ell}\bigg(\frac{h\gcd(a,d)}{d}\bigg),$$

where $\Phi_{a\ell}(v)$ is the Fourier integral

$$\Phi_{a\ell}(v) = \int_{-\infty}^{\infty} \phi(a^2 f(\ell, t)) e(-vt)\, dt. \tag{5-8}$$

The zero-frequency $h = 0$ gives exactly $M_d(\phi)$. Integration by parts yields

$$\Phi_{a\ell}(v) = (2\pi i v)^{-j} \int_{-\sqrt{X}/a}^{\sqrt{X}/a} e(-vt) \frac{\partial^j}{\partial t^j} \phi(a^2 f(\ell, t))\, dt.$$

Using our hypotheses on ϕ we estimate

$$\frac{\partial^j}{\partial t^j} \phi(a^2 f(\ell, t)) \ll \left(\frac{\sqrt{X}}{aY}\right)^j.$$

It follows that

$$\Phi_{a\ell}(v) \ll \frac{\sqrt{X}}{a} \left(\frac{\sqrt{X}}{aYv}\right)^j. \tag{5-9}$$

Now, for $R_d(\phi) = A_d(\phi) - M_d(\phi)$ we obtain from (5-7) that

$$\begin{aligned} R_d(\phi) &= \sum_{f(\alpha,1)\equiv 0 \,(\mathrm{mod}\, d)} \sum_a \mu(a) \sum_\ell \lambda(a\ell) \frac{\gcd(a,d)}{d} \sum_{h\neq 0} e\left(\alpha h \ell \frac{\gcd(a,d)}{d}\right) \Phi_{a\ell}\left(\frac{h \gcd(a,d)}{d}\right) \\ &= \frac{2}{d} \sum_a \mu(a) \sum_{\substack{bc=d \\ b\mid a}} b \sum_{\substack{\alpha \,(\mathrm{mod}\, bc) \\ f(\alpha,1)\equiv 0 \,(\mathrm{mod}\, bc)}} \sum_{h>0} \sum_\ell \lambda(a\ell) e\left(\frac{\alpha h \ell}{c}\right) \Phi_{a\ell}\left(\frac{h}{c}\right) \\ &= \frac{2}{d} \sum_a \mu(a) \sum_{\substack{bc=d \\ b\mid a}} b\rho_f(b) \sum_{\substack{\alpha \,(\mathrm{mod}\, c) \\ f(\alpha,1)\equiv 0 \,(\mathrm{mod}\, c)}} \sum_{h>0} \sum_\ell \lambda(a\ell) e\left(\frac{\alpha h \ell}{c}\right) \Phi_{a\ell}\left(\frac{h}{c}\right). \end{aligned} \tag{5-10}$$

Applying (5-9) for $h \geq a^{-1}Y^{-1}DX^{1/2+\psi(r)} = H$ for some small $\psi(r) > 0$ and choosing $j = j(r)$ sufficiently large, we may assume that $\Phi_{a\ell}(h/c) \ll h^{-2}D^{-1}$. Bounding absolutely we then conclude that the tail is bounded by

$$O(\rho_f(d) d^{-1} \|\lambda\|_1).$$

Since $|a\ell| \ll \sqrt{X}$, we thus conclude that

$$\|\lambda\|_1 \ll X^{1/(2r)} \quad \text{and} \quad \sum_{D\leq d<2D} \frac{\rho_f(d)}{d} \|\lambda\|_1 \ll X^{1/(2r)} \log D, \tag{5-11}$$

which is sufficiently small. To handle the remaining range, we apply a change of variables to obtain

$$\Phi_{a\ell}\left(\frac{h}{c}\right) = \frac{2\sqrt{X}}{ah} \int_0^\infty \phi\left(a^2 f\left(\ell, \frac{s\sqrt{X}}{ah}\right)\right) e\left(-\frac{s\sqrt{X}}{ac}\right) ds.$$

The integrand vanishes unless

$$\ell \ll \frac{\sqrt{X}}{a} \quad \text{and} \quad h \gg s.$$

It follows that

$$d\,|R_d(\phi)| \ll \sum_a^\flat \frac{\sqrt{X}}{a} \sum_{\substack{bc=d\\ b|a}} b\rho_f(b) \int_0^H \sum_{f(\alpha,1)\equiv 0 \,(\mathrm{mod}\, c)} \Bigg| \sum_{\substack{s<h<H\\ |\ell|\ll \sqrt{X}/a}} h^{-1}\lambda(a\ell)\phi\bigg(a^2 f\bigg(\ell, \frac{s\sqrt{X}}{ah}\bigg)\bigg) e\bigg(\frac{\alpha h\ell}{c}\bigg)\Bigg| ds.$$

We reorganize the inner sum as

$$\sum_{\substack{s<h<H\\ |\ell|\ll \sqrt{X}/a}} h^{-1}\lambda(a\ell)\phi\bigg(a^2 f\bigg(\ell, \frac{s\sqrt{X}}{ah}\bigg)\bigg) e\bigg(\frac{\alpha h\ell}{c}\bigg)$$
$$= \sum_{n\ll H\sqrt{X}/a} \Bigg(\sum_{\substack{h\ell=n\\ s<h<H}} \frac{1}{h}\lambda(a\ell)\phi\bigg(a^2 f\bigg(\ell, \frac{s\sqrt{X}}{ah}\bigg)\bigg)\Bigg) e\bigg(\frac{\alpha n}{c}\bigg) = \sum_{n\ll H\sqrt{X}/a} \xi_n(s) e\bigg(\frac{\alpha n}{c}\bigg),$$

say. Next we write

$$\sum_{D\le d<2D} |R_d(\phi)| \ll O(X^{1/(4r)}\log D)$$
$$+ \frac{1}{D}\int_0^H \Bigg(\sum_a^\flat \frac{\sqrt{X}}{a} \sum_{b|a} \rho_f(b) b \sum_{D/b\le c<2D/b} \sum_{f(\alpha,1)\equiv 0\,(\mathrm{mod}\, c)} \Bigg| \sum_{n\ll H\sqrt{X}/a} \xi_n(s) e\bigg(\frac{\alpha n}{c}\bigg)\Bigg|\Bigg) ds.$$

Applying Cauchy–Schwarz we obtain

$$\sum_{C\le c<2C} \sum_{f(\alpha,1)\equiv 0\,(\mathrm{mod}\, c)} \Bigg| \sum_{n\ll H\sqrt{X}/a} \xi_n(s) e\bigg(\frac{\alpha n}{c}\bigg)\Bigg|$$
$$\le C^{1/2}\Bigg(\sum_{C\le c<2C} \sum_{f(\alpha,1)\equiv 0\,(\mathrm{mod}\, c)} \Bigg| \sum_{n\ll H\sqrt{X}/a} \xi_n(s) e\bigg(\frac{\alpha n}{c}\bigg)\Bigg|^2\Bigg)^{1/2}$$
$$\ll C^{1/2}\big(C + H\sqrt{X}/a\big)^{1/2}\|\xi\|_2 \tag{5-12}$$

by Proposition 5.2. Next we note that

$$\|\xi(s)\|_2^2 \le \frac{1}{s^2} \sum_{n\ll H\sqrt{X}/a} \Bigg(\sum_{\substack{h\ell=n\\ s\le h<H}} \lambda(a\ell)\Bigg)^2$$

It follows that

$$\sum_{D\le d<2D} |R_d(\phi)| \ll$$
$$\frac{\sqrt{X}}{D} \sum_a^\flat \frac{1}{a} \sum_{b|a} \rho_f(b) b \bigg(\frac{D}{b}\bigg)^{1/2} \bigg(\frac{D}{b} + \frac{H\sqrt{X}}{a}\bigg)^{1/2} \Bigg| \sum_{n\ll H\sqrt{X}/a} \Bigg(\sum_{\substack{h\ell=n\\ H\le h<2H}} \lambda(a\ell)\Bigg)^2\Bigg|^{1/2} \int_0^H \frac{1}{s}\, ds. \tag{5-13}$$

Next we evaluate

$$\sum_{n \ll H\sqrt{X}/a} \Bigg(\sum_{\substack{h\ell = n \\ 0 < h < H}} \lambda(a\ell) \Bigg)^2.$$

Since a is square-free and $a\ell$ is an r-th power, it follows that $\ell = a^{r-1}m^r$ with $m \le a^{-1}X^{1/(2r)} = M$, say. Therefore we see that the sum above is bounded by the number of solutions to

$$h_1 m_1^r = h_2 m_2^r$$

with $H \le h_1, h_2 < 2H$ and $m_1, m_2 \le M$. The solutions are parametrized by $m_1 = st_1, m_2 = st_2$ with $\gcd(t_1, t_2) = 1$, $st_1, st_2 \le M$. Observe that

$$s \le \frac{M}{\max(t_1, t_2)} \quad \text{and} \quad k \le \frac{2H}{\max(t_1^r, t_2^r)}.$$

This gives

$$\sum_n \Bigg(\sum_{\substack{h\ell = n \\ 0 < h < H}} \lambda(a\ell) \Bigg)^2 \ll HM \sum_{t_1 \le t_2 \le M} \frac{1}{t_2^{r+1}} \ll HM \log M.$$

We thus obtain the upper bound of $O(Ha^{-1}X^{1/(2r)}(\log X))$. Inserting this into (5-13) and summing gives

$$\sum_{D \le d < 2D} |R_d(\phi)| \ll D^{-1/2}\big(D + H\sqrt{X}\big)^{1/2} H^{1/2} X^{1/2+1/(4r)} (\log X)^2. \tag{5-14}$$

Inserting

$$H \le DY^{-1}X^{1/2+\psi(r)}$$

into (5-14) gives the bound

$$\begin{aligned} \sum_{D \le d < 2D} |R_d(\phi)| &\ll X^{1/2} D^{-1} D^{1/2} H^{1/2} X^{1/4} H^{1/2} X^{1/(4r)} (\log X)^2 \\ &\ll_\varepsilon D^{-1/2} (Y^{-1} D X^{1/2+\psi(r)}) X^{3/4+1/(4r)} (\log X)^2 \\ &= D^{1/2} Y^{-1} X^{(5r+1)/(4r)+\psi(r)} (\log X)^2. \end{aligned} \tag{5-15}$$

This bound holds uniformly for $d \le D$. We may thus choose

$$Y = D^{1/4} X^{(7r-1)/(8r)-\psi(r)} (\log X)^{-26}.$$

This in turn gives the estimate

$$\sum_{d \le D} |A_d(\phi) - M_d(\phi)| \ll D^{1/4} X^{3(r+1)/(8r)} (\log X)^{24},$$

which is enough to prove Proposition 5.1.

6. Estimating $\pi(\mathcal{B})$: bilinear sum bounds

We will deal with the sum

$$P(X) = \sum_{n \le X} b_n \Lambda(n)$$

in the case of $\pi(\mathcal{B})$ via Vaughan's identity [18], which is an elegant combinatorial identity which decomposes the von Mangoldt function. The ideas recorded here are from [3]. Suppose $Y, Z \ge 1$ and suppose $n > Z$. Then

$$\Lambda(n) = \sum_{\substack{m|n \\ m \le Y}} \mu(m) \log \frac{n}{m} - \sum_{\substack{mc|n \\ m \le Y \\ c \le Z}} \mu(m)\Lambda(c) + \sum_{\substack{mc|n \\ b > Y \\ c > Z}} \mu(m)\Lambda(c) \tag{6-1}$$

and if $n \le Z$, the right hand side is zero. For $X > YZ$ then Vaughan's identity implies that

$$\begin{aligned}
P(X) &= P(Z) + \sum_{n \le X} b_n \Bigg(\sum_{\substack{m|n \\ m \le Y}} \mu(m) \log \frac{n}{m} - \sum_{\substack{mc|n \\ b \le Y \\ c \le Z}} \mu(m)\Lambda(c) + \sum_{\substack{mc|n \\ m > Y \\ c > Z}} \mu(m)\Lambda(c) \Bigg) \\
&= P(Z) + \sum_{m \le Y} \mu(m) \Bigg(\sum_{\substack{n \le X \\ m|n}} b_n \log n - \sum_{\substack{n \le X \\ m|n}} b_n \log m - \sum_{c \le Z} \Lambda(c) \sum_{\substack{n \le X \\ mc|n}} b_n \Bigg) + \sum_{m > Y} \mu(m) \sum_{c > Z} \Lambda(c) \sum_{\substack{n \le X \\ mc|n}} b_n \\
&= P(Z) + A(X; Y, Z) + \sum_{\substack{md \le X \\ m > Y}} \mu(m) \Bigg(\sum_{\substack{c|d \\ c > Z}} \Lambda(c) \Bigg) b_{md} \\
&= P(Z) + A(X; Y, Z) + B(X; Y, Z),
\end{aligned}$$

say. We can treat $P(Z)$ by applying trivial bounds provided that Z is sufficiently small with respect to X. The term $A(X; Y, Z)$ can be dealt with using the appropriate type-I estimates; see Proposition 5.1. The term $B(X; Y, Z)$, as expected, will require some type-II estimates. Given our treatment of the algebraic aspects of bilinear sums in Section 4, the treatment below is very similar to that given in [3; 12] so we will be fairly terse on the details.

Our target is the estimate

$$B(X; Y, Z) \ll \Delta X (\log X)^5,$$

with $\Delta = (\log X)^{-A}$ for any large, fixed $A > 5$. Recall that

$$B(X; Y, Z) = \sum_{Z < d < X/Y} \Bigg(\sum_{\substack{c|d \\ c > Z}} \Lambda(c) \Bigg) \sum_{Y < m \le X/d} \mu(m) b_{md}.$$

Using the trivial estimate

$$\sum_{\substack{c|d \\ c > Z}} \Lambda(c) \le \log X$$

we then find that

$$|B(X; Y, Z)| \le (\log X) \sum_{d>Z} \Big| \sum_{Y<m\le X/d} \mu(m) b_{md} \Big|.$$

We wish to break the sum into short sums of the shape

$$\mathcal{B}(M, N) = \sum_{M<m\le 2M} \Big| \sum_{N<n\le N'} \mu(n) b_{mn} \Big| \tag{6-2}$$

with $N' = e^{k\Delta} N$. Considering $M = 2^j Z$ and $N = e^{\Delta k} y$ for various j, k, we then see that

$$|\mathcal{B}(X; Y, Z)| \le (\log X) \sum_{\substack{\Delta X<MN<X \\ M\ge Z \\ N\ge Y}} \sum \mathcal{B}(M, N) + O(\Delta X (\log X)^2), \tag{6-3}$$

where the error term $O(\Delta X(\log X)^2)$ represents a trivial bound for the contribution of $\mu(m)b_{md}$ with $md \le 2\Delta X$ or $e^{-2\Delta} X < md \le X$, where the terms are not covered exactly. There are at most $2\Delta^{-1}(\log X)^2$ short sums $\mathcal{B}(M, N)$ in (6-3) so it suffices to show that

$$\mathcal{B}(M, N) \ll \Delta^2 X (\log X)^2 \tag{6-4}$$

for all M, N in the relevant range. We have a trivial bound

$$\mathcal{B}(M, N) \le \sum_{M<m\le 2M} \rho_f(m) \sum_{N<n\le N'} \rho_f(n) \ll \Delta MN,$$

and we can use this bound to obtain

$$\mathcal{B}(M, N) \le \sum_{d\le \Delta^{-1}} \mathcal{B}_d(M, N) + O(\Delta^2 X),$$

where $\mathcal{B}_d(M, N)$ consists of the subsum of $\mathcal{B}(M, N)$ where $\gcd(m, n) = d$. The error term $O(\Delta^2 X)$ comes from the trivial bound and the condition $d > \Delta^{-1}$. Next observe that

$$\mathcal{B}_d(M, N) \le \mathcal{B}_1(dM, N/d),$$

and so it suffices to show

$$\mathcal{B}_1(M, N) \ll \Delta^3 X (\log X)^2 \tag{6-5}$$

for M, N satisfying $M \ge Z$, $N \ge \Delta Y$ and $\Delta X < MN < X$.

Applying (4-7) to (6-2) we then obtain

$$\mathcal{B}_1(M, N) \le \sum_{A\cdot B=\mathrm{Cl}\, f} \sum_{\substack{\boldsymbol{m}\in A \\ M<N(J(\boldsymbol{m}))\le 2M}} \Big| \sum_{\substack{\boldsymbol{n}\in B \\ N<N(J(\boldsymbol{n}))\le N' \\ \gcd(N(J(\boldsymbol{n})), N(J(\boldsymbol{m})))=1}} \mu(N(J(\boldsymbol{n})))\Lambda(Q_{A,B}(\boldsymbol{m}; \boldsymbol{n})) \Big|.$$

Removing the coprimality condition via Möbius inversion as in [3; 12], as well as partitioning the sum $\mathcal{B}_1(M, N)$ based on the classes A, B, it suffices to show that the sums

$$\mathcal{C}_r(M, N) = \sum_{\substack{M<g_1(x_1,x_2)\le 2M\\(x_1,x_2)\in\mathcal{K}_1}} \left| \sum_{\substack{N<g_2(y_1,y_2)\le N'\\(y_1,y_2)\in\mathcal{K}_2}} \mu(rg_2(y_1, y_2))\Lambda(Q(x_1, x_2; y_1, y_2)) \right| \tag{6-6}$$

are bounded by $O(\Delta^5 X(\log X)^2)$ for every r, M, N satisfying

$$r < \Delta^{-2}, \quad M \ge Z, \quad N \ge \Delta^3 Y \quad \text{and} \quad \Delta X < MN < X$$

and $\mathcal{K}_1, \mathcal{K}_2$ domains which are contained in $[-CX, CX]^2$ for some absolute constant C depending only on our choices of fundamental domains.

If we write

$$Q(x_1, x_2; y_1, y_2) = x_1\ell_1(y_1, y_2) + x_2\ell_2(y_1, y_2)$$

for linear forms $\ell_1, \ell_2 \in \mathbb{Z}[x, y]$ then the condition that $Q(\boldsymbol{x}; \boldsymbol{y}) = 0$ implies that $(\ell_1(y_1, y_2), \ell_2(y_1, y_2))$ is proportional to $(-x_2, x_1)$. We then make a change of variables in the inner sum, obtaining

$$\mathcal{C}_r(M, N) = \sum_{\substack{M<g_1(x_1,x_2)\le 2M\\(x_1,x_2)\in\mathcal{K}_1}} \left| \sum_{\substack{N<g_2^*(z_1,z_2)\le N'\\(y_1,y_2)\in\mathcal{K}_2}} \mu(rg_2^*(z_1, z_2))\Lambda(x_1z_1 + x_2z_2) \right|,$$

where $z_i = \ell_i(y_1, y_2)$ and g_2^* is such that $g_2^*(z_1, z_2) = g_2(y_1, y_2)$. We are then left with the bilinear sum

$$C(\alpha, \beta; \lambda) = \sum_{\boldsymbol{z}} {\sum_{\boldsymbol{w}}}^{*} \alpha(\boldsymbol{z})\beta(\boldsymbol{w})\lambda(\mathcal{Q}(\boldsymbol{z}; \boldsymbol{w})), \tag{6-7}$$

where α is supported in a disk of radius R_1 and β supported on an annulus $\mathbb{A}(R_2, 2R_2)$ having inner radius R_2 and outer radius $2R_2$, say. Further, we assume that λ is supported on $|\ell| \le CAB$ for some absolute constant C depending only on f, so in particular the ℓ^2-norm of λ is finite. Applying the Cauchy–Schwarz inequality we obtain

$$|\mathcal{C}(\alpha, \beta; \lambda)| \le \sum_{\ell} |\lambda(\ell)| {\sum_{\boldsymbol{y}}}^{*} |\beta(\boldsymbol{y})| \left| \sum_{\mathcal{Q}(\boldsymbol{x};\boldsymbol{y})=\ell} \alpha(\boldsymbol{x}) \right| \le \|\lambda\|_2 \cdot \|\beta\|_2 \mathcal{D}(\alpha)^{1/2}, \tag{6-8}$$

where $\|\cdot\|_2$ denotes the ℓ^2-norm and

$$\mathcal{D}(\alpha) = {\sum_{\boldsymbol{y}}}^{*} \mathcal{G}(\boldsymbol{y}) \sum_{\ell} \left| \sum_{\mathcal{Q}(\boldsymbol{x};\boldsymbol{y})=\ell} \alpha(\boldsymbol{x}) \right|^2,$$

where $\mathcal{G}$ is any nonnegative function with $\mathcal{G}(\boldsymbol{y}) \ge 1$ on the annulus $\mathbb{A}(R_2, 2R_2)$. As in [3; 12] it will be convenient to suppose that $\mathcal{G}$ is a radial, compactly supported, and smooth function. Squaring out we obtain

$$\mathcal{D}(\alpha) = {\sum_{\boldsymbol{y}}}^{*} \mathcal{G}(\boldsymbol{y}) \sum_{\mathcal{Q}(\boldsymbol{x};\boldsymbol{y})=0} (\alpha * \alpha)(\boldsymbol{x}), \tag{6-9}$$

with

$$(\alpha * \alpha)(\boldsymbol{x}) = \sum_{\boldsymbol{u}-\boldsymbol{v}=\boldsymbol{x}} \alpha(\boldsymbol{u})\bar{\alpha}(\boldsymbol{v}).$$

Note that

$$(\alpha * \alpha)(0) = \|\alpha\|_2^2.$$

The orthogonality relation $\boldsymbol{x} \cdot \boldsymbol{y} = 0$ for a primitive $\boldsymbol{x}$ in (6-9) is equivalent to the statement that $\boldsymbol{y}$ is a rational integer multiple of $\boldsymbol{x}' = (-x_2, x_1)$. It follows that

$$\mathcal{D}(\alpha) = \sum_{c\in\mathbb{Z}} \sideset{}{^*}\sum_{\boldsymbol{y}} \mathcal{G}(\boldsymbol{y})(\alpha * \alpha)(c\boldsymbol{y}) = \mathcal{D}_0(\alpha) + 2\mathcal{D}^*(\alpha), \tag{6-10}$$

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

$$\mathcal{D}_0(\alpha) = \|\alpha\|_2^2 \sideset{}{^*}\sum_{\boldsymbol{y}} \mathcal{G}(\boldsymbol{y}) \ll \|\alpha\|_2^2 B^2$$

and

$$\mathcal{D}^*(\alpha) = \sum_{\boldsymbol{x}\neq\boldsymbol{0}} \mathcal{G}(\boldsymbol{x}^*)(\alpha * \alpha)(z),$$

where $\boldsymbol{x}^*$ is a primitive vector proportional to $\boldsymbol{x}$. Again, we may apply Möbius inversion to remove the primitivity conditions, and obtain

$$\mathcal{D}^*(\alpha) = \sum_{b,c>0} \sum \mu(b)\mathcal{D}(\alpha; bc)$$

where

$$\mathcal{D}(\alpha; bc) = \sum_{\boldsymbol{x}\equiv 0 \,(\mathrm{mod}\, bc)} \mathcal{G}(c^{-1}\boldsymbol{x})(\alpha * \alpha)(\boldsymbol{x}).$$

From here, the treatment is identical to the one given in [3; 12] as no structure of the Gaussian integers or even an imaginary quadratic field is necessary. This completes our treatment for $\pi(\mathcal{B})$.

7. Type-II estimates for $\pi(\mathcal{A}) - \pi(\mathcal{B})$: preliminary steps

We discuss the proof of Proposition 3.7. We note that Proposition 3.7 is exactly analogous to Proposition 5 in [10], though our sequences $\mathcal{A}, \mathcal{B}$ are different. We have largely divorced the arithmetic of our field K with the analysis of bilinear sums in Section 4, and so we are in good shape to import results from [10] directly. We will make clear which components of [10] can be used without change, and where we need to make suitable modifications.

We substitute (4-7) into (3-21) to obtain

$$\sum_{N<n\leq 2N} \sum_{m<X/N} \alpha_m \beta_n (a_{mn} - b_{mn})$$
$$= \sum_{A\cdot B = \mathrm{Cl}\, f} \sum_{\substack{w\in A_0 \\ N<N(J(w))\leq 2N}} \beta_w \sum_{\substack{v\in B_0 \\ N(J(v))<X/N}} \alpha_v (\mathfrak{Z}(Q_{A,B}(v,w)) - \Lambda(Q_{A,B}(v,w)), \tag{7-1}$$

where $\alpha_v = \alpha_{N(J(v))}$, $\beta_w = \beta_{N(J(w))}$. Writing each bilinear form Q above as $w_1\ell_1(v_1, v_2) + w_2\ell_2(v_1, v_2)$ say and applying a linear change of variables to the inner sum, we transform the inner sum

$$\sum_{\substack{v\in B_0\\ N(J(v))<X/N}} \alpha_v(\mathfrak{Z}(Q(v, w)) - \Lambda(Q(v, w)) = \sum_z \alpha_z(\mathfrak{Z}(w_1z_1 + w_2z_2) - \Lambda(w_1z_1 + w_2z_2))$$

say, with the support of z being the image of the support of the sum on the left under the linear transformation. The linear transformation depends only on Q and not X.

After applying these linear transformations, we have now changed all of our bilinear forms Q to

$$Q_0(x_1, x_2; y_1, y_2) = x_1y_1 + x_2y_2.$$

We write $\mathcal{S}_1(X) \times \mathcal{S}_2(X)$ for the union of the images of the supports of w, v in (7-1), so that (7-1) becomes

$$h(K) \sum_{\boldsymbol{w}\in\mathcal{S}_1(X)} \sum_{\boldsymbol{v}\in\mathcal{S}_2(X)} \alpha_{\boldsymbol{w}}\beta_{\boldsymbol{v}}(\mathfrak{Z}(w_1v_1 + w_2v_2) - \Lambda(w_1v_1 + w_2v_2)). \tag{7-2}$$

Remark 7.1. Since the linear transformations depend only on the class $1 \le j \le h(K)$ and the corresponding choice of fundamental domain, the image of the set $\mathcal{F}_j(X)$ with $N < N(J(w)) \le 2N$ is contained in the annulus $\mathbb{A}(c_1N, c_2N)$ for some positive numbers c_1, c_2 independent of N. Similarly, the image of $\mathcal{F}'_j(X)$ with $N(J(v)) \le X/N$ is contained in the disk $\mathbb{D}(c_3X/N)$ for some $c_3 > 0$ depending at most on f. This observation is crucial because we will use the Euclidean norm and the corresponding geometry to treat our sums when we wish to import estimates from [5; 10], and switch to using the norm on $\mathcal{O}_K$ and the corresponding induced norm on ideal numbers when the arithmetic of K is relevant.

Since we are looking to save an arbitrary power of log, it suffices to further subdivide the support of (7-2), and consider sums of the shape

$$\sum_{N<\|\boldsymbol{w}\|_2\le 2N} \alpha_{\boldsymbol{w}} \sum_{\|z\|_2\le X/N} \beta_z(\mathfrak{Z}(w_1z_1 + w_2z_2) - \Lambda(w_1z_1 + w_2z_2)).$$

Remark 7.2. We abuse notation and refer to the terms β_n for some positive integer n as well as β_z for some vector $z \in \mathbb{Z}^2$. In the former case we interpret the support of β_n to be a set of ideal numbers of $\mathcal{O}_K$ in a fixed class having norm equal to n, and in the latter we simply interpret the set of ideal numbers as a $\mathbb{Z}$-module.

Put

$$S_1(z, \boldsymbol{w}) = \sum_{\substack{p^2\in I\\ w_1z_1+w_2z_2=p^2}} 2p\log p \quad \text{and} \quad S_2(z, \boldsymbol{w}) = \sum_{\substack{p\in I\\ w_1z_1+w_2z_2=p}} \log p \tag{7-3}$$

and

$$S_1^{\spadesuit}(z, \boldsymbol{w}) = \sum_{\substack{k^2\in I(X)\\ w_1z_1+w_2z_2=k^2}} 2k \quad \text{and} \quad S_2^{\spadesuit}(z, \boldsymbol{w}) = \sum_{\substack{k\in I\\ w_1z_1+w_2z_2=k}} 1.$$

Our aim is to obtain the estimates

$$\sum_{\substack{\boldsymbol{w}\\ N<\|\boldsymbol{w}\|_2\le 2N}} \sum_{\substack{z\\ \|z\|_2\le X/N}} \alpha_{\boldsymbol{w}}\beta_z(S_1(z,\boldsymbol{w})-S_2(z,\boldsymbol{w})) \ll_A \frac{X}{(\log X)^A} \tag{7-4}$$

and

$$\sum_{\substack{\boldsymbol{w}\\ N<\|\boldsymbol{w}\|_2\le 2N}} \sum_{\substack{z\\ \|z\|_2\le X/N}} \alpha_{\boldsymbol{w}}\beta_z(S_1^{\spadesuit}(z,\boldsymbol{w})-S_2^{\spadesuit}(z,\boldsymbol{w})) \ll_A \frac{X}{(\log X)^A}.$$

We are almost ready to import the remaining argument from [10]. Let us put

$$\mathcal{R}(N;X)=\left\{z\in\mathbb{Z}^2 : N\le\|z\|_2<2N,\ \left|\arg(z)-\frac{k\pi}{2}\right|\le(\log X)^{-A} \text{ for some } k\in\mathbb{Z}\right\}.$$

We note that, as we will use repeatedly later (and we will remind the reader of this again when this becomes relevant), that once we subdivide the regions into small dyadic ranges that the conditions $\|z\|_2\sim N$ and $N(z)\sim N$ are nearly identical. Here $\boldsymbol{z}=\hat{z}$ is the vector associated to z, viewed as an ideal number of K.

The following results from [10] can now be imported without change:

Lemma 7.3 [10, Lemma 9]. *Suppose that both $\boldsymbol{z}$ and q are fixed. Then the number of possible $\boldsymbol{w}$ with $q=w_1z_1+w_2z_2$ is $O((M/N)^{1/2})$, where $M=X/N$.*

Lemma 7.4 [10, Lemma 10]. *We have*

$$\sum_{z\in\mathcal{R}(N;X)} \sum_{\boldsymbol{w}} \beta_z\alpha_{\boldsymbol{w}}S_j(z,\boldsymbol{w}) \ll_A X(\log X)^{-A}$$

for $j=1,2$.

We remark that Lemma 7.4 apply equally well with $S_j(z,\boldsymbol{w})$ replaced with $S_j^{\spadesuit}(z,\boldsymbol{w})$.

As is standard at this juncture (see [3; 5; 10]), we apply Cauchy–Schwarz to obtain

$$\left(\sum_{\boldsymbol{w}}\alpha_{\boldsymbol{w}}\sum_{z}\beta_z(S_1(z,\boldsymbol{w})-S_2(z,\boldsymbol{w}))\right)^2 \le \sum_{\boldsymbol{w}}\alpha_{\boldsymbol{w}}^2\sum_{\boldsymbol{w}}\left(\sum_{z}\beta_z(S_1(z,\boldsymbol{w})-S_2(z,\boldsymbol{w}))\right)^2.$$

It is then sufficient to show that

$$\sum_{\boldsymbol{y},z}\beta_{\boldsymbol{y}}\beta_z\sum_{\boldsymbol{w}}(S_1^{\spadesuit}(\boldsymbol{y},\boldsymbol{w})-S_2^{\spadesuit}(\boldsymbol{y},\boldsymbol{w}))(S_1^{\spadesuit}(z,\boldsymbol{w})-S_2^{\spadesuit}(z,\boldsymbol{w})) \ll_A \frac{XN}{(\log X)^A} \tag{7-5}$$

and

$$\sum_{\boldsymbol{y},z}\beta_{\boldsymbol{y}}\beta_z\sum_{\boldsymbol{w}}(S_1(\boldsymbol{y},\boldsymbol{w})-S_2(\boldsymbol{y},\boldsymbol{w}))(S_1(z,\boldsymbol{w})-S_2(z,\boldsymbol{w})) \ll_A \frac{XN}{(\log X)^A} \tag{7-6}$$

for any $A>0$.

We emphasize, as this will be relevant later, that the vectors $\boldsymbol{z}$, $\boldsymbol{y}$ represent elements in the same ideal class.

Next we consider the diagonal contribution coming from $\boldsymbol{y} = \boldsymbol{z}$. This gives the sums

$$\sum_{\boldsymbol{z}} \beta_{\boldsymbol{z}} \sum_{\boldsymbol{w}} \alpha_{\boldsymbol{w}} (S_1^{\spadesuit}(\boldsymbol{z}, \boldsymbol{w}) - S_2^{\spadesuit}(\boldsymbol{z}, \boldsymbol{w}))^2 = \sum_{\boldsymbol{z}} \beta_{\boldsymbol{z}} \sum_{\boldsymbol{w}} \alpha_{\boldsymbol{w}} (S_1^{\spadesuit}(\boldsymbol{z}, \boldsymbol{w})^2 - 2S_1^{\spadesuit}(\boldsymbol{z}, \boldsymbol{w}) S_2^{\spadesuit}(\boldsymbol{z}, \boldsymbol{w}) + S_2^{\spadesuit}(\boldsymbol{z}, \boldsymbol{w})^2)$$

and

$$\sum_{\boldsymbol{z}} \beta_{\boldsymbol{z}} \sum_{\boldsymbol{w}} \alpha_{\boldsymbol{w}} (S_1(\boldsymbol{z}, \boldsymbol{w}) - S_2(\boldsymbol{z}, \boldsymbol{w}))^2 = \sum_{\boldsymbol{z}} \beta_{\boldsymbol{z}} \sum_{\boldsymbol{w}} \alpha_{\boldsymbol{w}} (S_1(\boldsymbol{z}, \boldsymbol{w})^2 - 2S_1(\boldsymbol{z}, \boldsymbol{w}) S_2(\boldsymbol{z}, \boldsymbol{w}) + S_2(\boldsymbol{z}, \boldsymbol{w})^2).$$

Clearly,

$$S_1(\boldsymbol{z}, \boldsymbol{w}) S_2(\boldsymbol{z}, \boldsymbol{w}) = S_1^{\spadesuit}(\boldsymbol{z}, \boldsymbol{w}) S_2^{\spadesuit}(\boldsymbol{z}, \boldsymbol{w}) = 0$$

since their supports are incompatible. Next we have the trivial estimate

$$\begin{aligned}
\sum_{\boldsymbol{z}} \sum_{\boldsymbol{w}} S_1(\boldsymbol{z}, \boldsymbol{w}) &\ll \sum_{N < \|\boldsymbol{z}\|_2 \le 2N} \sum_{p^2 \in I} p \log p \sum_{\substack{\boldsymbol{w} \\ w_1 z_1 + w_2 z_2 = p^2}} 1 \\
&\ll \sqrt{\frac{M}{N}} \sum_{p^2 \in I} p \log p \sum_{N \le \|\boldsymbol{z}\|_2 < 2N} 1 \\
&\ll \sqrt{MN} \sum_{p^2 \in I} p \log p \\
&\ll_{\varepsilon} \sqrt{MN} X^{1/2+\varepsilon} \ll_{\varepsilon} X^{1+\varepsilon}.
\end{aligned}$$

Similarly, we conclude

$$\begin{aligned}
\sum_{\boldsymbol{z}} \sum_{\boldsymbol{w}} S_2(\boldsymbol{z}, \boldsymbol{w}) &\ll_{\varepsilon} X^{1+\varepsilon}, \\
\sum_{\boldsymbol{z}} \sum_{\boldsymbol{w}} S_1^{\spadesuit}(\boldsymbol{z}, \boldsymbol{w}) &\ll_{\varepsilon} X^{1+\varepsilon}, \\
\sum_{\boldsymbol{z}} \sum_{\boldsymbol{w}} S_2^{\spadesuit}(\boldsymbol{z}, \boldsymbol{w}) &\ll_{\varepsilon} X^{1+\varepsilon}.
\end{aligned}$$

From here we obtain

$$\begin{aligned}
\sum_{\boldsymbol{z}} \sum_{\boldsymbol{w}} S_1(\boldsymbol{z}, \boldsymbol{w})^2 + S_2(\boldsymbol{z}, \boldsymbol{w})^2 &\ll X^{1/4} \log X \sum_{\boldsymbol{z}} \sum_{\boldsymbol{w}} S_1(\boldsymbol{z}, \boldsymbol{w}) + \log X \sum_{\boldsymbol{z}} \sum_{\boldsymbol{w}} S_2(\boldsymbol{z}, \boldsymbol{w}) \\
&\ll_{\varepsilon} X^{5/4+\varepsilon}.
\end{aligned}$$

and

$$\sum_{\boldsymbol{z}} \sum_{\boldsymbol{w}} S_1^{\spadesuit}(\boldsymbol{z}, \boldsymbol{w})^2 + S_2^{\spadesuit}(\boldsymbol{z}, \boldsymbol{w})^2 \ll_{\varepsilon} X^{5/4+\varepsilon}.$$

At this stage, we expunge the references to the Gaussian domain $\mathbb{Z}[i]$ in [10] to make it clear that much of their treatment of bilinear sums apply equally well in our situation, despite the fact that our number field is different from $\mathbb{Q}(i)$. For $\boldsymbol{y}, \boldsymbol{z} \in \mathbb{Z}^2$ put $\Delta(\boldsymbol{y}, \boldsymbol{z}) = y_1 z_2 - y_2 z_1$. Given $\boldsymbol{w}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{Z}^2$ such that

$$w_1 y_1 + w_2 y_2 = q_1 \quad \text{and} \quad w_1 z_1 + w_2 z_2 = q_2,$$

we have

$$\begin{bmatrix} y_1 & y_2 \\ z_1 & z_2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}.$$

Inverting the matrix on the left we see that

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \frac{1}{\Delta(\boldsymbol{z}, \boldsymbol{y})} \begin{bmatrix} z_2 & -y_2 \\ -z_1 & y_1 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}.$$

Since $\boldsymbol{w} = (w_1, w_2) \in \mathbb{Z}^2$, it follows that

$$q_1 z_2 - q_2 y_2 \equiv q_1 z_1 - q_2 y_1 \equiv 0 \pmod{\Delta(\boldsymbol{z}, \boldsymbol{y})}. \tag{7-7}$$

Let $C(q_1, q_2, \boldsymbol{z}, \boldsymbol{y})$ be the statement that $q_1, q_2, \boldsymbol{z}, \boldsymbol{y}$ satisfy (7-7). Next we have

$$\begin{aligned} \|q_1(z_1, z_2) - q_2(y_1, y_2)\|_2 &= \sqrt{(q_1 z_1 - q_2 y_1)^2 + (q_1 z_2 - q_2 y_1)^2} \\ &= \sqrt{(w_1 \Delta(\boldsymbol{z}, \boldsymbol{y}))^2 + (w_2 \Delta(\boldsymbol{z}, \boldsymbol{y}))^2} \\ &= \Delta(\boldsymbol{z}, \boldsymbol{y}) \sqrt{w_1^2 + w_2^2} \leq \Delta(\boldsymbol{z}, \boldsymbol{y}) M. \end{aligned} \tag{7-8}$$

We also wish to impose the condition that $\Delta(\boldsymbol{z}, \boldsymbol{y})$ is small. In particular, we wish to only consider those $\boldsymbol{z}, \boldsymbol{y}$ with

$$\Delta(\boldsymbol{z}, \boldsymbol{y}) > \mathfrak{D}_0 = N(\log X)^{-A-6}. \tag{7-9}$$

For brevity, let us write

$$h^{\dagger}(q) = \begin{cases} 2p \log p & \text{if } q = p^2 \in I(X), \\ 0 & \text{otherwise,} \end{cases} \qquad h^{\ddagger}(q) = \begin{cases} \log p & \text{if } q = p \in I(X), \\ 0 & \text{otherwise,} \end{cases}$$

and

$$h(q) = h^{\dagger}(q) - h^{\ddagger}(q).$$

Similarly, let us write

$$h^{\spadesuit,\dagger}(q) = \begin{cases} 2p \log p & \text{if } q = p^2 \in I(X), \\ 0 & \text{otherwise,} \end{cases} \qquad h^{\spadesuit,\ddagger}(q) = \begin{cases} \log p & \text{if } q = p \in I(X), \\ 0 & \text{otherwise,} \end{cases}$$

and

$$h^{\spadesuit}(q) = h^{\spadesuit,\dagger}(q) - h^{\spadesuit,\ddagger}(q).$$

For any subinterval $J \subset I(X)$ we have

$$\sum_{q \in J} h(q) = O_C\left(\frac{X^{1/4}}{(\log X)^C}\right)$$

for any $C > 0$. This is a consequence of our choice of weights.

As in [10], we want to carve up the support of $\boldsymbol{z}, \boldsymbol{y}$ into regions of the form

$$\mathcal{U} = \mathcal{U}(c, \theta_0) = \left\{\boldsymbol{z} : c\sqrt{N} < \|\boldsymbol{z}\|_2 \leq c(1+\omega_1)\sqrt{N},\ \theta_0 < \arg(\boldsymbol{z}) \leq \theta_0 + \omega_2\right\}, \tag{7-10}$$

for fixed $1 \le c \le \sqrt{2}$ and θ_0. We may choose ω_1 and ω_2 so that the regions $\mathcal{U}$ form a partition of the region

$$\{\boldsymbol{z} : N \le \|\boldsymbol{z}\|_2 < 2N,\ z_1 > 0\} \setminus \mathcal{R}.$$

The number of regions needed for the sum over $\boldsymbol{z}$, $\boldsymbol{y}$ is $O((\log X)^{4L})$. Here, as in [10], we allow the parameters ω_1 and ω_2, both of order $(\log X)^{-A}$, to be different in order to perfectly cover our region.

As in [10] let us write $\mathfrak{C}_1(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)$ as the condition that all $(\boldsymbol{z}, \boldsymbol{y}, q_1, q_2) \in \mathcal{U}_1 \times \mathcal{U}_2 \times J_1 \times J_2$ satisfy (7-8) and (7-9). We remark that such tuples are the most intricate to estimate; in fact it is only in the treatment of these tuples where we must diverge from the argument given in [10].

Similarly, let $\mathfrak{C}_2(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)$ denote the condition that there exists some tuple $(\boldsymbol{z}, \boldsymbol{y}, q_1, q_2) \in \mathcal{U}_1 \times \mathcal{U}_2 \times J_1 \times J_2$ which satisfies (7-8) and there exists some tuple $(\boldsymbol{z}', \boldsymbol{y}', q_1', q_2') \in \mathcal{U}_1 \times \mathcal{U}_2 \times J_1 \times J_2$ which does not satisfy (7-8). Finally, let $\mathfrak{C}_3(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)$ be the condition that all tuples $(\boldsymbol{z}, \boldsymbol{y}, q_1, q_2) \in \mathcal{U}_1 \times \mathcal{U}_2 \times J_1 \times J_2$ satisfy (7-8) but there exists some tuple $(\boldsymbol{z}, \boldsymbol{y}, q_1, q_2) \in \mathcal{U}_1 \times \mathcal{U}_2 \times J_1 \times J_2$ which does not satisfy (7-9).

Recall that $C(q_1, q_2, \boldsymbol{z}, \boldsymbol{y})$ is the condition that $\boldsymbol{z}, \boldsymbol{y}, q_1, q_2$ satisfy (7-7). We also introduce the condition $\sum^{\flat}$ to indicate a summation over primitive $\boldsymbol{z} \notin \mathcal{R}(N; X)$. Observe that we do not insist that $\boldsymbol{z} \equiv (1, 0) \pmod 2$ as in [10]. For $\mathcal{U}_1, \mathcal{U}_2, J_1, J_2$ satisfying $\mathfrak{C}_1(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)$ put

$$T(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) = \sum_{\substack{\boldsymbol{z} \in \mathcal{U}_1 \\ \boldsymbol{y} \in \mathcal{U}_2}}^{\flat} \beta_{\boldsymbol{z}} \beta_{\boldsymbol{y}} \sum_{\substack{q_1 \in J_1 \\ q_2 \in J_2 \\ C(q_1, q_2, \boldsymbol{z}, \boldsymbol{y})}} h(q_1) h(q_2), \tag{7-11}$$

and otherwise set $T(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) = 0$. Further, let

$$T'(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) = \sum_{\substack{\boldsymbol{z} \in \mathcal{U}_1 \\ \boldsymbol{y} \in \mathcal{U}_2}}^{\flat} \sum_{\substack{q_1 \in J_1 \\ q_2 \in J_2 \\ C(q_1, q_2, \boldsymbol{z}, \boldsymbol{y})}} |h(q_1) h(q_2)|. \tag{7-12}$$

Similarly, define

$$T_{\spadesuit}(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) \text{ and } T'_{\spadesuit}(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)$$

analogously with h replaced with $h^{\spadesuit}$. Then to obtain (7-5) and (7-6) it suffices to show that

$$\sum_{\substack{\mathcal{U}_1, \mathcal{U}_2, J_1, J_2 \\ \mathfrak{C}_1(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)}} T_{\spadesuit}(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) + \sum_{\substack{\mathcal{U}_1, \mathcal{U}_2, J_1, J_2 \\ \mathfrak{C}_2(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) \text{ or } \mathfrak{C}_3(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)}} T'_{\spadesuit}(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) \ll_A \frac{XN}{(\log X)^A} \tag{7-13}$$

and

$$\sum_{\substack{\mathcal{U}_1, \mathcal{U}_2, J_1, J_2 \\ \mathfrak{C}_1(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)}} T(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) + \sum_{\substack{\mathcal{U}_1, \mathcal{U}_2, J_1, J_2 \\ \mathfrak{C}_2(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) \text{ or } \mathfrak{C}_3(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)}} T'(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) \ll_A \frac{XN}{(\log X)^A}. \tag{7-14}$$

As in [10], we will show that the contribution from $\mathfrak{C}_i(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)$ is negligible for $i = 2, 3$. Indeed, we shall obtain:

Proposition 7.5. *We have*

$$\sum_{\substack{\mathcal{U}_1,\mathcal{U}_2,J_1,J_2\\ \mathfrak{C}_2(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2)\text{ or }\mathfrak{C}_3(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2)}} T'_{\spadesuit}(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2) \ll_A \frac{XN}{(\log X)^A}$$

and

$$\sum_{\substack{\mathcal{U}_1,\mathcal{U}_2,J_1,J_2\\ \mathfrak{C}_2(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2)\text{ or }\mathfrak{C}_3(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2)}} T'(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2) \ll_A \frac{XN}{(\log X)^A}.$$

In fact, Proposition 7.5 is exactly analogous to Proposition 6 in [10]. More strikingly, the proof does not need to be modified and we can simply apply Proposition 6 of [10]. However, given that our setups are not identical we will explain why our situations are indeed interchangeable.

We will also need the following analogue of Proposition 7 in [10]:

Proposition 7.6. *For fixed J_1, J_2 and $L = 6A + 52$ we have*

$$\sum_{\substack{\mathcal{U}_1,\mathcal{U}_2\\ \mathfrak{C}_1(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2)}} T_{\spadesuit}(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2) \ll_A \frac{XN}{(\log X)^{A+2L}}.$$

and

$$\sum_{\substack{\mathcal{U}_1,\mathcal{U}_2\\ \mathfrak{C}_1(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2)}} T(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2) \ll_A \frac{XN}{(\log X)^{A+2L}}.$$

Unlike Proposition 7.5 we cannot simply import Proposition 7 from [10]. This is because Proposition 7.5, by the definition of $T'(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2)$, is insensitive to the nature of the coefficients β_z and so the treatment in [10] is directly applicable to our situation. However in order to prove Proposition 7 in [10] the specific shape of β_z was needed. That said, the modifications needed to adapt their proof to our case are minor, and we will still be able to follow their argument for the most part.

In the next few sections we will give proofs for Propositions 7.5 and 7.6. We will largely follow the structure of the argument given in [10].

8. Proof of Proposition 7.5

First we have the following lemma, which is Lemma 12 from [10]:

Lemma 8.1. *We have the bounds*

$$\sum_{\boldsymbol{z},\boldsymbol{y}}^{\flat} \sum_{\substack{q_1\in J_1,q_2\in J_2\\ C(q_1,q_2,\boldsymbol{z},\boldsymbol{y})\\ \gcd(q_1q_2,\Delta(\boldsymbol{z},\boldsymbol{y}))>1}} |h^{\spadesuit}(q_1)h^{\spadesuit}(q_2)| \ll N^2\sqrt{X}(\log X)^3$$

and

$$\sum_{\boldsymbol{z},\boldsymbol{y}}^{\flat} \sum_{\substack{q_1\in J_1,q_2\in J_2\\ C(q_1,q_2,\boldsymbol{z},\boldsymbol{y})\\ \gcd(q_1q_2,\Delta(\boldsymbol{z},\boldsymbol{y}))>1}} |h(q_1)h(q_2)| \ll N^2\sqrt{X}(\log X)^3.$$

Proof. See Section 7 in [10]. □

Lemma 8.1 allows us, as in [10], to write

$$T(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) = \sum_{D \le 2N} \sideset{}{^*}\sum_{a \,(\mathrm{mod}\, D)} \mathcal{Y}(a, D; h, h)\mathcal{Z}(a, D) + O\big(N^2\sqrt{X}(\log X)^3\big) \tag{8-1}$$

where

$$\mathcal{Z}(a, D) = \sideset{}{^\flat}\sum_{\substack{(z,y)\in\mathcal{U}_1\times\mathcal{U}_2\\ \Delta(z,y)=D\\ ay\equiv z\,(\mathrm{mod}\, D)}} \beta_z\beta_y$$

and

$$\mathcal{Y}(a, D; h_1, h_2) = \sum_{\substack{q_1\in J_1,\, q_2\in J_2\\ q_1\equiv aq_2\,(\mathrm{mod}\, D)\\ \gcd(q_1q_2, D)=1}} h_1(q_1)h_2(q_2).$$

Similarly, we have

$$T_\spadesuit(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2) = \sum_{D \le 2N} \sideset{}{^*}\sum_{a \,(\mathrm{mod}\, D)} \mathcal{Y}^\spadesuit(a, D; h^\spadesuit, h^\spadesuit)\mathcal{Z}(a, D) + O\big(N^2\sqrt{X}(\log X)^3\big) \tag{8-2}$$

where

$$\mathcal{Y}^\spadesuit(a, D; h_1^\spadesuit, h_2^\spadesuit) = \sum_{\substack{q_1\in J_1,\, q_2\in J_2\\ q_1\equiv aq_2\,(\mathrm{mod}\, D)\\ \gcd(q_1q_2, D)=1}} h_1^\spadesuit(q_1)h_2^\spadesuit(q_2)$$

This crucial decomposition allows us to separate $T(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)$ and $T_\spadesuit(\mathcal{U}_1, \mathcal{U}_2, J_1, J_2)$ into components $\mathcal{Z}(a, D)$ containing the coefficients β_z, β_y and a congruence sum which no longer has anything to do with the coefficients β. To treat (7-14) requires a treatment of $\mathcal{Y}(a, D)$ involving primes. For this purpose they needed a refinement of the Barban–Davenport–Heilbronn theorem, which we will not go into more detail here as we can use their Proposition 6 directly.

The following lemma is critical to the proof of Proposition 7.5:

Lemma 8.2. *Let*

$$\widetilde{\mathcal{Z}}(a, D) = \sideset{}{^\flat}\sum_{\substack{(z,y)\in\mathcal{U}_1\times\mathcal{U}_2\\ \Delta(z,y)=D\\ ay\equiv z\,(\mathrm{mod}\, D)}} 1.$$

We then have the bounds

$$\sum_D \tau(D) \sideset{}{^*}\sum_{a \,(\mathrm{mod}\, D)} \widetilde{\mathcal{Z}}(a, D) \ll \omega^4 N^2(\log X)^{16}, \tag{8-3}$$

$$\sum_{\mathcal{U}_1,\mathcal{U}_2} \sideset{}{^*}\sum_{a \,(\mathrm{mod}\, D)} \widetilde{\mathcal{Z}}(a, D) \ll N, \tag{8-4}$$

and

$$\sum_{a \,(\mathrm{mod}\, D)}^{*} \widetilde{\mathcal{Z}}(a, D) \ll (\log X)^3 \frac{N^2}{D} \tau(D)^6. \tag{8-5}$$

Proof. See Lemma 13 in [10]. □

For an interval J and a function $\mathfrak{h}$, put

$$\mathcal{Y}(J, \mathfrak{h}; D) = \sum_{\substack{q \in J \\ \gcd(q, D)=1}} \mathfrak{h}(q)$$

and

$$\mathcal{Y}_{\mathfrak{h}_1, \mathfrak{h}_2}(D) = \mathcal{Y}(D) = \frac{1}{\varphi(D)} \mathcal{Y}(J_1, \mathfrak{h}_1; D) \mathcal{Y}(J_2, \mathfrak{h}_2; D).$$

Recall that q_1, q_2 appearing in $\mathcal{Y}(a, D; \mathfrak{h}_1, \mathfrak{h}_2)$ satisfy $\gcd(q_1 q_2, D) = 1$. If h_1 or h_2 is equal to $h^\ddagger$, then $\mathcal{Y}(D)$ is the expected value of $\mathcal{Y}(a, D; h_1, h_2)$. If $h_1 = h_2 = h^\dagger$, note that $p_1^2 \equiv a p_2^2 \pmod{D}$ implies that $p_1 \equiv b p_2 \pmod{D}$ for some b such that $a \equiv b^2 \pmod{D}$. Here, $\mathcal{Y}(a, D; h_1, h_2) = 0$ if a is not a square modulo D. Therefore

$$\sum_{a \,(\mathrm{mod}\, D)}^{*} \mathcal{Y}(a, D; h_1, h_2) \mathcal{Z}(a, D) = \sum_{b \,(\mathrm{mod}\, D)}^{*} \mathcal{Y}_{h^\dagger}(b, D) \mathcal{Z}(b^2, D)$$

where

$$\mathcal{Y}_{h^\dagger}(b, D) = \sum_{\substack{p_1^2 \in J_1,\, p_2^2 \in J_2 \\ p_1 \equiv b p_2 \,(\mathrm{mod}\, D) \\ \gcd(p_1 p_2, D)=1}} h^\dagger(p_1^2) h^\dagger(p_2^2).$$

When $h_1 = h_2 = h^\dagger$, then $\mathcal{Y}(D)$ is the expected value of $Y_{h^\dagger}(b, D)$. Now put

$$\mathcal{E}(N) = \sum_{D \le 2N} \sum_{a \,(\mathrm{mod}\, D)}^{*} |\mathcal{Y}(a, D; h_1, h_2) - \mathcal{Y}_{h_1, h_2}(D)| \, \widetilde{\mathcal{Z}}(a, D)$$

if either $h_1 = h^\dagger$ or $h_2 = h^\dagger$, and

$$\mathcal{E}_{h^\dagger}(N) = \sum_{D \le 2N} \sum_{b \,(\mathrm{mod}\, D)}^{*} |\mathcal{Y}_{h^\dagger}(b, D) - \mathcal{Y}_{h^\dagger, h^\dagger}(D)| \, \widetilde{\mathcal{Z}}(b^2, D)$$

if $h_1 = h_2 = h^\dagger$. We then have the following proposition, which is Proposition 8 from [10]:

Proposition 8.3. *For any $C > 0$ we have*

$$\mathcal{E}(N) \ll_C \frac{XN}{(\log X)^C} \quad \textit{and} \quad \mathcal{E}_{h^\dagger}(N) \ll_C \frac{XN}{(\log X)^C}.$$

With this proposition in hand, we may proceed to prove Proposition 7.5 in the exact same way as Proposition 6 in [10]. We will not repeat the details.

We now move to the proof of Proposition 7.6. Most of the arguments can be adapted from the proof of Proposition 7 in [10], but since we rely on some properties of the coefficients β_z in this argument we

cannot follow all of the arguments in [10] verbatim. We will especially emphasize those points where modifications are required.

9. Proof of Proposition 7.6: some maneuvers

Supposing that one of the functions h_1, h_2 is $h^\ddagger$, we have according to Proposition 8.3 that

$$\sum_{D\le 2N}\ \sideset{}{^*}\sum_{a\,(\mathrm{mod}\,D)} \mathcal{Y}(a,D;h_1,h_2)\mathcal{Z}(a,D) = \sum_{D\le 2N}\ \sideset{}{^*}\sum_{a\,(\mathrm{mod}\,D)} \mathcal{Y}_{h_1,h_2}(D)\mathcal{Z}(a,D) + O_C\Big(\frac{XN}{(\log X)^C}\Big)$$

for any $C>0$. In the remaining case with $h_1=h_2=h^\dagger$, we have

$$\sum_{D\le 2N}\ \sideset{}{^*}\sum_{a\,(\mathrm{mod}\,D)} \mathcal{Y}(a,D;h^\dagger,h^\dagger)\mathcal{Z}(a,D) = \sum_{D\le 2N}\ \sideset{}{^*}\sum_{b\,(\mathrm{mod}\,D)} \mathcal{Y}_{h^\dagger,h^\dagger}(D)\mathcal{Z}(b^2,D) + O_C\Big(\frac{XN}{(\log X)^C}\Big).$$

As in [10] we may replace $\mathcal{Y}_{h^\dagger,h^\dagger}(D)$ by $|J_1|\,|J_2|/\varphi(D)$ in each case, with a total error of

$$O\big(X\exp\big(\sqrt{-\log X}\,\big)N(\log X)^2\big).$$

Our remaining task is the inequality

$$|J_1|\,|J_2| \sum_{\substack{\mathcal{U}_1,\mathcal{U}_2\\ \mathfrak{C}_1(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2)}} \sum_{D\le 2N}\frac{1}{\varphi(D)}\Big(\sideset{}{^*}\sum_{b\,(\mathrm{mod}\,D))}\mathcal{Z}(b^2,D) - \sideset{}{^*}\sum_{a\,(\mathrm{mod}\,D)}\mathcal{Z}(a,D)\Big) \ll \frac{XN}{(\log X)^{A+2L}},$$

or

$$\mathcal{E}' = \sum_{\substack{\mathcal{U}_1,\mathcal{U}_2\\ \mathfrak{C}_1(\mathcal{U}_1,\mathcal{U}_2,J_1,J_2)}} \sum_{D}\frac{1}{\varphi(D)}\Big(\sideset{}{^*}\sum_{b\,(\mathrm{mod}\,D)}\mathcal{Z}(b^2,D) - \sideset{}{^*}\sum_{a\,(\mathrm{mod}\,D)}\mathcal{Z}(a,D)\Big) \ll \frac{N}{(\log X)^A}.$$

Here we dropped the condition $D\le 2N$, which follows automatically since β_z is supported on $\|z\|_2\le 2N$.

We may follow Heath-Brown and Li's arguments in [10] to conclude that it suffices to obtain the estimate

$$\mathcal{E}_1(\mathcal{U}_1,\mathcal{U}_2) = \sum_{D}\frac{D}{\varphi(D)}\Big(\sideset{}{^*}\sum_{b\,(\mathrm{mod}\,D)}\mathcal{Z}(b^2,D) - \sideset{}{^*}\sum_{a\,(\mathrm{mod}\,D)}\mathcal{Z}(a,D)\Big) \ll \frac{N^2}{(\log X)^{C_1}} \tag{9-1}$$

for any $C_1>0$ and for fixed $\mathcal{U}_1,\mathcal{U}_2$. By Möbius inversion we deduce that

$$\mathcal{E}_1(\mathcal{U}_1,\mathcal{U}_2) = \sum_{D=1}^{\infty}\sum_{k=1}^{\infty}\frac{D\mu(k)}{\varphi(D)}\Big(\sideset{}{^*}\sum_{b\,(\mathrm{mod}\,D)}W(b^2,k,D) - \sideset{}{^*}\sum_{a\,(\mathrm{mod}\,D)}W(a,k,D)\Big),$$

where

$$W(a,k,D) = \sum_{\substack{(z,y)\in\mathcal{U}_1\times\mathcal{U}_2\\ kD\mid\Delta(z,y)\\ ay\equiv z\,(\mathrm{mod}\,D)}}{}^{\flat}\ \beta_z\beta_y.$$

Here the condition $\Delta(z,y)=D$ which appears in the definitions of $\mathcal{Z}(a,D)$, $\widetilde{\mathcal{Z}}(a,D)$ is replaced with a divisibility condition via Möbius inversion.

If kD divides $\Delta(z, y)$, there is a unique integer $c = c(z, y; kD)$ modulo kD such that $cy \equiv z \pmod{kD}$, and conversely this congruence implies that kD divides $\Delta(z, y)$. We have $\gcd(c, kD) = 1$ and

$$\begin{aligned}\#\{b \pmod D : b^2 y \equiv z \pmod D\} &= \#\{b \pmod D : b^2 \equiv c \pmod D\} \\ &= \sum_{\substack{\chi \pmod D \\ \chi^2 = \chi_0}} \chi(c).\end{aligned}$$

It now follows that

$$\sum_{b \pmod D}^{*} W(b^2, k, D) - \sum_{a \pmod D}^{*} W(a, k, D) = \sum_{\substack{\chi \pmod D \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \sum_{c \pmod{kD}}^{*} \sum_{\substack{(z,y) \in \mathcal{U}_1 \times \mathcal{U}_2 \\ cy \equiv z \pmod{kD}}}^{\flat} \beta_z \beta_y \chi(c),$$

and hence

$$\mathcal{E}_1(\mathcal{U}_1, \mathcal{U}_2) = \sum_{D=1}^{\infty} \sum_{k=1}^{\infty} \frac{D\mu(k)}{\varphi(D)} \sum_{\substack{\chi \pmod D \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}} \sum_{c \pmod{kD}}^{*} \sum_{\substack{(z,y) \in \mathcal{U}_1 \times \mathcal{U}_2 \\ cy \equiv z \pmod{kD}}}^{\flat} \beta_z \beta_y \chi(c).$$

Let $d = d(\chi)$ be the conductor of χ and write $D = de$ and $ek = \mathfrak{k}$, giving

$$\mathcal{E}(\mathcal{U}_1, \mathcal{U}_2) = \sum_{d>1} \sum_{\mathfrak{k}} C(d, \mathfrak{k}) \sum_{\substack{\chi \pmod d \\ \chi^2 = \chi_0}}^{*} \sum_{c \pmod{d\mathfrak{k}}}^{*} \sum_{\substack{(z,y) \in \mathcal{U}_1 \times \mathcal{U}_2 \\ cy \equiv z \pmod{d\mathfrak{k}}}}^{\flat} \beta_z \beta_y \chi(c),$$

where

$$C(d, \mathfrak{k}) = \sum_{d_1 k = d_2} \frac{de\mu(k)}{\varphi(de)} = \frac{d}{\varphi(d)} \sum_{ek = \mathfrak{k}} \frac{\varphi(d) e \mu(k)}{\phi(de)}.$$

The sum for $\chi \pmod d$ is empty unless $d = d_1, 4d_1, 8d_1$ with d_1 odd and square-free, in which cases there are at most two possible characters χ. For fixed d the function

$$\varphi_d(e) = \frac{\varphi(d)e}{\varphi(de)}$$

is multiplicative in e. Further, for $v \geq 1$ we have

$$(\varphi_e * \mu)(p^v) = \begin{cases} (p-1)^{-1} & \text{if } v = 1 \text{ and } p \nmid d, \\ 0 & \text{otherwise.} \end{cases}$$

We then see that

$$C(d, \mathfrak{k}) = \frac{d\mu^2(\mathfrak{k})}{\varphi(d\mathfrak{k})}$$

if $\gcd(d, \mathfrak{k}) = 1$ and $C(d, \mathfrak{k}) = 0$ otherwise. This gives the expression

$$\mathcal{E}_1(\mathcal{U}_1, \mathcal{U}_2) = \sum_{\substack{\mathfrak{k}, d \\ \gcd(d, \mathfrak{k}) = 1}} \frac{d\mu^2(\mathfrak{k})}{\varphi(d\mathfrak{k})} \sum_{\substack{\chi \pmod d \\ \chi^2 = \chi_0 \\ \chi \neq \chi_0}}^{*} \left(\sum_{c \pmod{d\mathfrak{k}}}^{*} \sum_{\substack{(z,y) \in \mathcal{U}_1 \times \mathcal{U}_2 \\ cy \equiv z \pmod{d\mathfrak{k}}}}^{\flat} \beta_z \beta_y \chi(c) \right). \tag{9-2}$$

We proceed to show that large values of $\mathfrak{k}$ make a negligible contribution. Since $d\mathfrak{k} \mid \Delta(z, y)$ we have $d\mathfrak{k} \le 2N$. Since $0 \le \beta_z \le 1$ we find that

$$\begin{aligned}\sum_{\mathfrak{k}>\mathfrak{K}} \sum_{\substack{\mathfrak{k}\\ \gcd(d,\mathfrak{k})=1}} \frac{d\mu^2(\mathfrak{k})}{\varphi(d\mathfrak{k})} \sideset{}{^*}\sum_{\substack{\chi \,(\mathrm{mod}\, d)\\ \chi^2=\chi_0}} \Bigg| \sideset{}{^*}\sum_{c \,(\mathrm{mod}\, d\mathfrak{k})} \sideset{}{^\flat}\sum_{\substack{(z,y)\in\mathcal{U}_1\times\mathcal{U}_2\\ cy\equiv z \,(\mathrm{mod}\, d\mathfrak{k})}} \beta_z\beta_y\chi(c)\Bigg| &\ll (\log X)\sum_{\mathfrak{k}>\mathfrak{K}} \mathfrak{k}^{-1} \sum_{d\le 2N/\mathfrak{k}} \sum_{\substack{d\mathfrak{k}\mid D\\ D\le 2N}} \sideset{}{^*}\sum_{a \,(\mathrm{mod}\, D)} \widetilde{\mathcal{Z}}(a, D)\\ &\ll (\log X)\sum_{\mathfrak{k}>\mathfrak{K}} \mathfrak{k}^{-1} \sum_{d\le 2N/\mathfrak{k}} \sum_{\substack{d\mathfrak{k}\mid D\\ D\le 2N}} N\\ &\ll \frac{N^2(\log X)^2}{\mathfrak{K}}.\end{aligned}$$

Choosing

$$\mathfrak{K} = (\log X)^{C_1+2}$$

and applying Lemma 8.2 then gives a satisfactory bound.

Observe that the argument above only depends on the property that $0 \le \beta_z \le 1$, and so no modification is necessary from the argument given by Heath-Brown and Li in [10]. As in [10] we divide into three ranges for d, namely

$$d \le D_1, \quad D_1 < d \le D_2, \quad \text{and} \quad d > D_2,$$

where

$$D_1 = \mathfrak{K}^{10}(\log X)^{2C_1+14} \quad \text{and} \quad D_2 = \frac{N}{\mathfrak{K}^{15}(\log X)^{3C_1+21}}.$$

Next we handle the middle range of d. The treatment given here is identical to that in [10], since again the specific shape of β_z is of no consequence in this part. Set

$$\mathcal{E}_1(D) = \sum_{\mathfrak{k}\le\mathfrak{K}} \mathfrak{k}^{-1}\mu^2(\mathfrak{k}) \sum_{\substack{D<d\le 2D\\ \gcd(d,\mathfrak{k})=1}} \sideset{}{^*}\sum_{\substack{\chi \,(\mathrm{mod}\, d)\\ \chi^2=\chi_0}} \Bigg| \sideset{}{^*}\sum_{c \,(\mathrm{mod}\, d\mathfrak{k})} \sideset{}{^\flat}\sum_{\substack{(z,y)\in\mathcal{U}_1\times\mathcal{U}_2\\ cy\equiv z \,(\mathrm{mod}\, d\mathfrak{k})}} \beta_z\beta_y\chi(c)\Bigg|.$$

We remark on the significance that the sum is over primitive characters in the definition of $\mathcal{E}_1(D)$. Indeed, as seen in [10] this property is necessary to decompose the characters into Jacobi symbols.

Heath-Brown and Li obtains the following bound, which we summarize in the following lemma:

Lemma 9.1. *For any $\varepsilon > 0$ we have*

$$\mathcal{E}_1(D) \ll_\varepsilon \mathfrak{K}^5(\log X)^6(D + D^{-1/2}N + D^{1/3}N^{2/3} + N^{23/24+\varepsilon})N.$$

Summing over dyadic ranges of D, we see that the values of d in the range $D_1 \le d \le D_2$ make a satisfactory contribution given our choices of D_1, D_2.

It then remains to give estimates for the small and large ranges of d, where we must depart somewhat from Heath-Brown and Li's treatment due to the dependence on the specific shapes of the coefficients β_z.

10. Proof of Proposition 7.6: remaining ranges

Large d. We will obtain the bound

$$\sum_{\substack{d>D_2\\ \gcd(d,\mathfrak{k})=1}} \sideset{}{^*}\sum_{\substack{\chi \,(\mathrm{mod}\, d)\\ \chi^2=\chi_0}} \Bigg(\sideset{}{^*}\sum_{c \,(\mathrm{mod}\, d\mathfrak{k})} \sideset{}{^\flat}\sum_{\substack{(z,y)\in \mathcal{U}_1\times\mathcal{U}_2\\ cy\equiv z \,(\mathrm{mod}\, d\mathfrak{k})}} \beta_z\beta_y \chi(c)\Bigg) \ll_C \frac{N^2}{(\log X)^C}$$

for any $C>0$ and $\mathfrak{k}\le \mathfrak{K}$. There is still more mileage we can get from the argument given in [10]. In particular, we follow their argument in Section 11 of [10] and decompose d as d_1d_2, as well as $\chi=\chi_1\chi_2$. We have $d\mathfrak{k}\mid \Delta(z,y)$ and thus we may set $\Delta(z,y)=d_1et$ where e is odd and t is a power of 2. Our conditions on $\mathcal{U}_1,\mathcal{U}_2$ guarantee that $0<\Delta(z,y)\le 2N$, hence $1\le et\le 16N/D_2\ll (\log X)^{18C_1+51}$. We split the sums over z,y into congruence classes $z\equiv u \pmod{8et}$, $y\equiv v \pmod{8et}$ and fix the parameters

$$\mathfrak{k},\quad d_2,\quad \chi_2,\quad e,\quad u,\quad v,\quad \text{and}\quad t. \tag{10-1}$$

Each admissible pair u,v corresponds to a unique integer $k \pmod{\Delta(z,y)}$ with the property that $ky\equiv z \pmod{\Delta(z,y)}$, and then

$$\chi(c)=\chi(k)=\chi_2(k)\left(\frac{k}{d_1}\right),$$

where $\chi_2(k)$ is determined by the parameters (10-1). The number of choices for the parameters (10-1) is bounded by a fixed power of $\log X$ and so it suffices to show that

$$\sum_{\substack{d_1>D_2/d_2\\ \gcd(d_2,2\mathfrak{k})=1}} \frac{d_1\mu^2(d_1)}{\varphi(d_1)}\Bigg(\sideset{}{^*}\sum_{k\,(\mathrm{mod}\, d_1et)} \sideset{}{^\flat}\sum_{z,y} \beta_z\beta_y\left(\frac{k}{d_1}\right)\Bigg) \ll_C \frac{N^2}{(\log X)^C}$$

for every $C>0$, where the sum over z,y satisfies the conditions

$$(z,y)\in\mathcal{U}_1\times\mathcal{U}_2,\quad ky\equiv z \pmod{\Delta(z,y)},\quad z\equiv u \pmod{8et},\quad y\equiv v \pmod{8et},\quad \text{and}\quad \Delta(z,y)=d_1et.$$

Following the same analysis in Section 11.1 of [10], we conclude that it is sufficient to obtain the bound

$$\sum_{\substack{(z,y)\in\mathcal{U}_1\times\mathcal{U}_2\\ z\equiv u,\, y\equiv v \,(\mathrm{mod}\, 8etn)\\ \Delta(z,y)>etD_2/d_2}} \beta'_z\beta'_y \ll_C \frac{N^2}{(\log X)^C}$$

where

$$\beta'_z=\beta_z(-1)^{(z_1-1)/2}\left(\frac{z_2}{z_1}\right).$$

for every fixed $C>0$, for each choice of parameters $e,t,n\le(\log X)^C$, and for each u,v. Further subdividing into congruence classes it suffices to handle

$$\sum_{\substack{(z,y)\in\mathcal{U}_1\times\mathcal{U}_2\\ z\equiv u,\, y\equiv v \,(\mathrm{mod}\, 8etn)}} \beta'_z\beta'_y = \Bigg(\sum_{\substack{z\in\mathcal{U}_1\\ z\equiv u \,(\mathrm{mod}\, 8etn)}} \beta'_z\Bigg)\Bigg(\sum_{\substack{z\in\mathcal{U}_2\\ z\equiv v \,(\mathrm{mod}\, 8etn)}} \beta'_z\Bigg). \tag{10-2}$$

At this stage, we must diverge from Heath-Brown and Li's treatment. We briefly discuss why this is necessary. In order to proceed, Heath-Brown and Li rely on the crucial property that their β_z are supported on Gaussian integers z such that $N(z)$ has no small prime factors. The analogous condition for us is that the ideal number $\gamma(\boldsymbol{z})$ has norm (equal to the norm of the ideal $J(\gamma(\boldsymbol{z}))$ in $\mathcal{O}_K$) without small prime factors. Thus, now going to the perspective that $\boldsymbol{z}$ represents an ideal number γ, we see that $N(\gamma) = N(J(\gamma))$ is automatically coprime to $8etn$ and therefore we may assume that υ, ν (the ideal numbers corresponding to $\boldsymbol{u}$, $\boldsymbol{v}$, respectively) are coprime to $8etn$. This allows us to pick out the congruence condition $\gamma \equiv \upsilon, \nu \pmod{8etn}$ using multiplicative characters. In order to make this precise, we borrow from the algebraic treatment given in [11] and put

$$\mathfrak{J}(q) = \{\alpha \in \mathfrak{J} : \gcd(\alpha, q) = 1\}$$

and $\mathfrak{J}_1(q) = \mathfrak{J}(q) \cap K$. Further, put

$$\mathfrak{J}_0(q) = \{\alpha \in K : \alpha \equiv 1 \pmod{q}\}.$$

Then our congruence conditions can be picked out using characters of the quotient group $\mathfrak{J}_1(q)/\mathfrak{J}_0(q)$, and we conclude that

$$\sum_{\substack{\boldsymbol{a} \in \mathcal{U}_j \\ \alpha \equiv \nu \,(\mathrm{mod}\, 8etn)}} = \frac{1}{\varphi_K(8etn)} \sum_{\chi \,(\mathrm{mod}\, 8etn)} \bar{\chi}(\nu) \mathcal{S}(\chi, \mathcal{U}_j),$$

where φ_K is the Euler-φ function for $\mathcal{O}_K$ and

$$\mathcal{S}(\chi, \mathcal{U}) = \sum_{\boldsymbol{a} \in \mathcal{U}} \beta'_\alpha \chi(\alpha).$$

In order to obtain acceptable estimates for $\mathcal{S}(\chi, \mathcal{U})$, we will need to generalize certain results from [5] to apply to general quadratic fields. This work may be of independent interest and is recorded in the next section; see Propositions 11.7 and 11.9 in particular. We emphasize that these results rely on the setup in (10-2): in particular, we need $\boldsymbol{z}$, $\boldsymbol{w}$ to come from the same ideal class and that they satisfy a congruence condition modulo $8etn$.

We now proceed to pick out the condition that we are constrained in a narrow sector using a twice-differentiable periodic function $\upsilon(\theta)$, where

$$\upsilon(\theta) = \begin{cases} 1 & \text{if } \theta \in (\theta_0, \theta_0 + \varpi_2) \pmod{2\pi}, \\ 0 & \text{if } \theta \notin [\theta_0 - (\log X)^{-C}, \theta_0 + \varpi_2 + (\log X)^{-C}] \pmod{2\pi}, \end{cases}$$

and where $|\upsilon''(\theta)| \ll (\log X)^{-2C}$. Then

$$\mathcal{S}(\chi, \mathcal{U}) = \sum_{N' < N(z) \le N'(1+\varpi)} \beta'_z \chi(z) \upsilon(\arg z) + O\left(\frac{N}{(\log X)^C}\right).$$

The Fourier coefficients of υ satisfy $c_k \ll k^{-2}(\log X)^{2C}$ for $k \neq 0$, and so

$$\upsilon(\arg z) = \sum_k c_k \left(\frac{z}{|z|}\right)^k = \sum_{|k| \leq (\log X)^{3C}} c_k \left(\frac{z}{|z|}\right)^k + O((\log X)^{-C}).$$

It then suffices to show that

$$\mathcal{S}(\chi, N', k) = \sum_{N' < N(z) \leq N'(1+\varpi)} \beta'_z \chi(z) \left(\frac{z}{|z|}\right)^k \ll_C N(\log X)^{-4C}$$

for any $C > 0$, and for $|k| \leq (\log X)^{3C}$. As in [10] we can obtain in fact a small power-saving in N. We recall that $\beta_z = \beta_{N(z)}$ is the indicator function of a set of one of the shapes

$$Q_j = \{p_1 \cdots p_{j+1} \in (N', N'(1+\varpi)] : p_{j+1} \in J,\ p_{j+1} < \cdots < p_1,\ p_1 \cdots p_j < Y \leq p_1 \cdots p_{j+1} < X^{1/20\delta}\}$$

or

$$R = \{n \in (N', N'(1+\varpi)] : \gcd(n, P(V)) = 1\}.$$

Here we will have $0 \leq j \leq n_0 = \lfloor \log Y/(\delta \log X) \rfloor$, and $J = [V, V(1+\kappa)) \subseteq [X^\delta, X^{1/2-\delta})$. In particular we interpret Q_0 to be $\{p : p \in J \cap (N', N'(1+\varpi))\}$.

We now write

$$\lambda(n) = \sum_{N(z)=n}^{\wedge} \chi(z) \left(\frac{z}{|z|}\right)^k u^{(x-1)/2} \left(\frac{z_2}{z_1}\right),$$

where $\sum^\wedge$ denotes a sum over primitive ideal numbers z in a fixed class of ideal numbers, with $\hat{z} = (z_1, z_2)$. We then have

$$\mathcal{S}(\chi, N', k) = \sum_n \lambda(n),$$

where n runs over R or Q_j for some j. As in [10], the treatment for R and Q_j are similar. To begin, we first handle the contribution from those n whose largest prime factor, say $\mathcal{P}(n)$, exceeds $N^{99/100}$. The contribution from such integers is

$$\sum_{m \leq 2N^{1/100}} \sum_{\substack{p > \max\{\mathcal{P}(m),\, N^{99/100}\} \\ mp \in Q_j}} \lambda(mp).$$

Since p is the largest prime factor of mp one sees from the definition of the set Q_j that one may rewrite the conditions $p > \mathcal{P}(m)$ and $mp \in Q_j$ to say that p runs over an interval $I_j(m) \subseteq [N/m, 2N/m)$. We may then apply Proposition 11.9 to conclude that

$$\begin{aligned}\sum_{m \leq 2N^{1/100}} \sum_{\substack{p > \max\{\mathcal{P}(m),\, N^{99/100}\} \\ mp \in Q_j}} \lambda(mp) &\ll q_0(|k|+1) \sum_{m \leq 2N^{1/100}} m(N/m)^{76/77} \\ &\ll q_0(|k|+1) N^{76/77 + (78/77)/100}.\end{aligned}$$

Since $76/77 + (78/77)/100 < 1$, this gives the required power-saving bound.

Next we deal with the terms where every prime factor is at most $N^{99/100}$. To do so we rewrite our sum in terms of bilinear sums. Suppose $n = p_1 \cdots p_{j+1}$ as in the description of the set Q_j, and divide the range of each prime p_i into intervals of the shape $(B_i, 2B_i]$. This will give us at most $(2\log N)^{1+n_0}$ sets of dyadic ranges, and since $n_0 \ll \delta^{-1} = (\log X)^{1-\varpi}$ there will be at most $O_\varepsilon(N^\varepsilon)$ such ranges. Moreover we may suppose

$$\prod_{i=1}^{j+1} B_i \ll N \ll 2^{j+1} \prod_{i=1}^{j+1} B_i.$$

Since we may now assume that $B_1 \le N^{99/100}$ there will be an index u such that

$$N^{1/100} \le \prod_{i=1}^{u} B_i \le N^{99/100}.$$

Fixing such an index u we split $n = n_1 n_2$ with

$$n_1 = \prod_{i=1}^{u} p_i \quad \text{and} \quad n_2 = \prod_{i=u+1}^{j+1} p_i,$$

so that $n_1 \le N_1$ and $n_2 \le N_2$ with

$$N_1 = 2^{1+n_0} \prod_{i=1}^{u} B_i \quad \text{and} \quad N_2 = 2^{1+n_0} \prod_{i=u+1}^{j+1} B_i.$$

It follows that

$$N_1 N_2 \ll_\varepsilon N^{1+\varepsilon} \quad \text{and} \quad N_1, N_2 \ll_\varepsilon N^{99/100+\varepsilon}.$$

This implies that

$$N_1 N^{-\varepsilon} \ll n_1 \le N_1 \quad \text{and} \quad N_2 N^{-\varepsilon} \ll n_2 \le N_2.$$

We may thus reinterpret our description of Q_j by requiring that $n_1 \in Q_{j,u}$ and $n_2 \in Q'_{j,u}$ for appropriate sets $Q_{j,u}$, $Q'_{j,u}$, together with the conditions that

$$n_1 n_2 \in I = (N', N'(1+\varpi)] \cap [Y, X^{1/2-\delta}),\ p_{j+1}^{-1} n_1 n_2 < Y, \text{ and } p_{u+1} < p_u). \tag{10-3}$$

In other words, we put

$$Q_{j,u} = \{n_1 = p_1 \cdots p_u : p_i \in (B_i, 2B_i],\ p_u < \cdots < p_1\}$$

and

$$Q'_{j,u} = \{n_2 = p_{u+1} \cdots p_{j+1} : p_i \in (B_i, 2B_i],\ p_{j+1} \in J,\ p_{j+1} < \cdots < p_{u+1} < Y\}.$$

In order to separate the variables n_1, n_2 completely we subdivide the available ranges for n_1, n_2, p_{j+1}, p_u, and p_{u+1} into intervals of the shape $(A, A + A/L)$, $(A', A' + A'/L]$, $(B'_{j+1}, B'_{j+1} + B'_{j+1}/L]$, $(B'_u, B'_u + B'_u/L]$ and $(B'_{u+1}, B'_{u+1} + B'_{u+1}/L]$. Here the parameter L will be chosen to be a small power of N. These intervals may have length less than one. Indeed such an interval may contain no integers at all.

There will be $O(L^5(\log X)^2)$ such intervals and there will be some for which $n_1n_2 \in I$, $p_{j+1}^{-1}n_1n_2 < Y$ and $p_{u+1} < p_u$ for every choice of $p_1, \dots, p_{j+1}$ satisfying

$$n_1 \in (A, A + A/L], \quad n_2 \in (A', A' + A'/L],$$
$$p_{j+1} \in (B'_{j+1}, B'_{j+1} + B'_{j+1}/L], \quad p_u \in (B'_u, B'_u + B'_u/L], \quad p_{u+1} \in (B'_{u+1}, B'_{u+1} + B'_{u+1}/L],$$

and

$$p_i \in I_i \quad \text{with } i \neq 1, u, u+1.$$

This case gives the subsum

$$\sum_{\substack{n_1 \in Q_{j,u} \cap (A, A+N_1/L] \\ p_u \in (B'_u, B'_u + B'_u/K]}} \sum_{\substack{n_2 \in Q'_{j,u} \cap (A', A'+A'/L] \\ p_{j+1} \in (B'_{j+1}, B'_{j+1}+B'_{j+1}/L] \\ p_{u+1} \in (B'_{u+1}, B'_{u+1}+B'_{u+1}/L]}} \lambda(n_1n_2),$$

so that we have separated the variables n_1, n_2. For such sums we can apply Proposition 11.7 which gives the bound

$$O_\varepsilon((N_1 + N_2)^{1/12}(N_1N_2)^{1/12+\varepsilon}) = O_\varepsilon(N^{(9/100)\cdot(1/12)} \cdot N^{1/12+\varepsilon}) = O_\varepsilon(N^{1-1/1200+\varepsilon}).$$

Since there are $O_\varepsilon(L^5N^\varepsilon)$ such subsums the overall contribution will be $O(L^5N^{1-1/200+\varepsilon})$.

It remains to consider the contribution from the remaining "bad" sets of ranges which are not exclusively contained in the region given by (10-3). First suppose that the interval I is given by $[e_1, e_2]$ say, and that there are integers $n_1, n_1' \in (A, A + A/L]$ and $n_2, n_2' \in (A', A' + A'/L]$ such that $n_1n_2 \in I$ but $n_1'n_2' \notin I$. Then we must have $n_1n_2 = (1 + O(L^{-1})e_1$ or $n_1n_2 = (1 + O(L^{-1}))e_2$. We now consider the total contribution from integers $n \in Q_j$ for all such "bad" choices of intervals $(A, A + A/L)$, $(A', A' + A'/L]$, $(B'_{j+1}, B'_{j+1} + B'_{j+1}/L]$, $(B'_u, B'_u + B'_u/L]$ and $(B'_{u+1}, B'_{u+1} + B'_{u+1}/L]$. Since each integer n occurs at most once, and $\lambda(n) = O(\tau(n))$, the contribution will be

$$O_\varepsilon\Bigg(\sum_{n=(1+O(L^{-1}))e_1} \tau(n)\Bigg) = O_\varepsilon(N^{1+\varepsilon}L^{-1}).$$

Similarly, if we have $p_{j+1}^{-1}n_1n_2 < Y$ but $(p'_{j+1})^{-1}n_1'n_2' \geq Y$, then $p_{j+1}^{-1}n_1n_2 = (1 + O(L^{-1}))Y$. This gives

$$B_{j+1}Y \asymp AA' \leq N_1N_2 \ll_\varepsilon N^{1+\varepsilon},$$

so any n which is counted in this case will have a prime factor $p \ll N^{1+\varepsilon}/Y$ and such that $p^{-1}n = (1 + O(L^{-1}))Y$. Thus, on writing $n = pm$, we see that the total contribution in this case is

$$O\Bigg(\sum_{p \ll N^{1+\varepsilon}/Y} \sum_{m=(1+O(L^{-1}))Y} \tau(pm)\Bigg) = O_\varepsilon(N^{1+\varepsilon}Y^{-1}(1 + L^{-1}Y)) = O_\varepsilon(N^{1+\varepsilon}L^{-1}),$$

for $L \leq Y$.

Finally, if $B_u = B_{u+1}$, then it may happen that the condition $p_{u+1} < p_u$ is satisfied by some, but not all, pairs of primes (p_u, p_{u+1}) from the intervals $(B'_u, B'_u + B'_u/L]$ and $(B'_{u+1}, B'_{u+1} + B'_{u+1}/L]$. Clearly this problem cannot arise when $L \geq 2P_u$ since then the intervals $(B'_u, B'_u + B'_u/L]$ and $(B'_{u+1}, B'_{u+1} + B'_{u+1}/L]$ contain at most one prime each. It follows that any such n to be counted in this case must have two prime factors $p' > p \geq P_u \geq L/2$ with $p' = (1 + O(L^{-1}))p$. Hence the corresponding contribution is

$$O\Bigg(\sum_{\substack{p'>p\geq L/2 \\ p'=(1+O(L^{-1}))p}} \sum_{\substack{n\ll N \\ p'p\mid n}} \tau(n) \Bigg) = O_\varepsilon\Bigg(\sum_{\substack{p'>p\geq L/2 \\ p'=(1+O(L^{-1}))p}} \frac{N^{1+\varepsilon}}{p'p} \Bigg) = O_\varepsilon(N^{1+\varepsilon}L^{-1}).$$

We therefore find that our sum is bounded by

$$O_\varepsilon(L^5 N^{1-1/1200+\varepsilon} + N^{1+\varepsilon}L^{-1}),$$

whenever $L \leq Y$. We may then choose $L = N^{10^{-5}}$ say, to achieve the claimed power saving in the case of large d.

Small d. To handle small d it suffices to show that for any $\mathfrak{k} \leq C$, $d \leq D_1$, and any nonprincipal $\chi \pmod d$,

$$\sideset{}{^*}\sum_{c \,(\mathrm{mod}\, d\mathfrak{k})} \sum_{\substack{(z,y)\in \mathcal{U}_1\times\mathcal{U}_2 \\ cz\equiv y \,(\mathrm{mod}\, d\mathfrak{k})}}{}^{\flat} \beta_z\beta_y\chi(c) \ll_C \frac{N^2}{(\log X)^C}$$

for every $C > 0$. As is usually the case in prime number theory, the case of small moduli can be handled using some type of Siegel–Walfsiz theorem; we shall use the results of Mitsui [16] following the argument in [10].

Since

$$\sideset{}{^*}\sum_{c \,(\mathrm{mod}\, d\mathfrak{k})} \chi(c) = 0,$$

it suffices to prove that if $\mathcal{U} = \mathcal{U}_1$ or $\mathcal{U}_2$ then there is a number $\mathfrak{M} = \mathfrak{M}(\mathcal{U}, d\mathfrak{k})$ such that

$$\sum_{\substack{z\in\mathcal{U} \\ z\equiv\alpha \,(\mathrm{mod}\, 2d\mathfrak{k})}} \beta_z = \mathfrak{M} + O_C\Bigg(\frac{N}{(\log X)^C}\Bigg)$$

for any $\gcd(\alpha, 2d\mathfrak{k}) = 1$ and $C > 0$, since β_z is supported on those z free of small prime factors, and $2d\mathfrak{k}$ is small. As before we may drop the summation condition $\flat$. For notational convenience, we set $q = 2d\mathfrak{k}$ and note that $q \leq (\log X)^{C_0}$ for some $C_0 > 0$.

As in the previous subsection we may assume that $\beta_z = \beta_{N(z)}$, where β_n is the indicator function of either Q_j or R. We describe the procedure for Q_j, the method for R being similar. We decompose z as s_1s_2 with $N(s_1)$ being the largest prime factor of $N(s_1s_2)$. The requirement that $n \in Q_j$ is then equivalent to a condition of the form $N(s_2) \in Q'_j$ together with a restriction of the type $N(s_1) \in I(s_2)$ for some real interval $I(s_2)$. Specifically, we have

$$Q'_{j+1} = \{p_2 \cdots p_{j+1} : p_{j=1} \in J,\ p_{j+1} < \cdots < p_2\}$$

and

$$I(s_2) = (p_2, \infty) \cap \left(\frac{N'}{N(s_2)}, \frac{N'(1+\varpi)}{N(s_2)}\right] \cap \left[\frac{Y}{N(s_2)}, \frac{X^{1/2-\delta}}{N(s_2)}\right),$$

where p_2 is the largest prime factor of $N(s_2)$. When $\mathcal{U}$ is given by (7-10) the condition on the size of $N(s_1 s_2)$ is exactly the condition

$$N(s_1) \in \left(\frac{N'}{N(s_2)}, \frac{N'(1+\omega)}{N(s_2)}\right],$$

and we have $\theta_0 < \arg z \le \theta_0 + \omega_2$ exactly when

$$\theta_1(s_2) < \arg s_1 \le \theta_1(s_2) + \omega_2,$$

with $\theta_1(s_2) = \theta_1 - \arg s_2$. It follows that

$$\sum_{\substack{z \in \mathcal{U} \\ z \equiv \alpha \,(\mathrm{mod}\, q)}} \beta_z = \sum_{\substack{N(s_2) \in Q_j' \\ \gcd(s_2, q) = 1}} \mathcal{N}(s_2, \alpha), \tag{10-4}$$

where $\mathcal{N}(s_2, \alpha)$ is the number of ideal numbers s_1 satisfying

$$s_1 s_2 \equiv \alpha \pmod q, \quad N(s_1) \in I(s_2), \quad \text{and} \quad \theta_1(s_2) < \arg s_1 \le \theta_1(s_2) + \omega_2$$

and for which $N(s_1)$ is prime. We can estimate $\mathcal{N}(s_2, \alpha)$ using a form of the prime number theorem for arithmetic progressions over number fields, due to Mitsui [16]. As we remarked earlier, we can easily redivide our sectors in accordance with the condition $N(z) \sim N$ as opposed to $\|z\|_2 \sim N$, so we may apply Mitsui's theorem without worry in each of our sectors. If we put $\pi(X; q, \alpha, \theta)$ for the number of prime ideal numbers $\mathfrak{p}$ in a fixed ideal class satisfying $\mathfrak{p} \equiv \alpha \pmod q$ and having norm at most X with $0 \le \arg \mathfrak{p} \le \theta$, then Mitsui's theorem gives the estimate

$$\pi(X; q, \alpha, \theta) = \frac{w\theta R_K}{2^{r_1} h_K \varphi_K(\mathfrak{a})} \mathrm{Li}(X) + O_K\left(X \exp\left(-c\sqrt{\log X}\right)\right), \tag{10-5}$$

where r_1 is the number of real embeddings of K, w the number of roots of unity in K, R_K the regulator of K, and h_K the class number of K. Here c is an absolute constant. Since we do not care about dependence on K, we may take the implied constant in (10-5) as an absolute constant. We emphasize that (10-5) holds uniformly for $\theta \in [0, 2\pi]$ and for all $q \le (\log X)^A$.

Applying (10-5) with $q = 2d\mathfrak{k}$ to estimate $\mathcal{N}(s_2, \alpha)$, we have $I(s_2) \subseteq (0, 2N/N(s_2)]$ and so we will need to know that $q = 2d\mathfrak{k} \le (\log 2N/N(s_2))^A$ for some constant A. This holds whenever p divides an element of Q_j then one has $p \ge X^{\delta_1}$ with $\delta = (A \log\log X)/\log X$. Thus we will have $2N/N(s_2) \ge X^{\delta_1}$ and so

$$\delta_1 \log X \le \log\left(\frac{N}{N(s_2)}\right),$$

which implies that

$$\log X \le \left(\log\left(\frac{N}{N(s_2)}\right)\right)^{1/\varpi}.$$

Therefore whenever $2d\mathfrak{k} \le (\log X)^{C_0}$ we have

$$2d\mathfrak{k} \le (\log X)^{C_0} \le \left(\log\left(\frac{2N}{N(s_2)}\right)\right)^{C_0/\varpi}.$$

The required condition therefore holds when $\mathfrak{k} \le \mathfrak{K}$ and $d \le D_1$.

We may then conclude, as in [10], that

$$\mathcal{N}(s_2, \alpha) = \mathfrak{M}(s_2, d\mathfrak{k}, j, \mathcal{U}) + O\left(\frac{N}{N(s_2)} \exp(-c(\log X)^{\varpi/2})\right),$$

where the main term crucially is independent of α. Feeding this into (10-4) then completes our treatment of small d, and hence the proof of Proposition 3.7.

11. A generalization of the Jacobi–Kubota symbol and consequences

We will introduce and prove analogues of Proposition 23.1 and Theorem ψ in [5]. We introduce, for an ideal number α in a fixed class A, the vector

$$\widehat{\alpha} = (a_1, a_2) \in \mathbb{Z}^2$$

corresponding to the class A with basis produced as in Section 4. We then introduce the *Jacobi–Kubota symbol*

$$[\alpha] = i^{(a_1-1)/2}\left(\frac{a_2}{|a_1|}\right), \tag{11-1}$$

where $\left(\frac{\cdot}{\cdot}\right)$ is the Jacobi symbol. Our generalized Jacobi–Kubota symbol $[\,\cdot\,]$ depends on the class A *and the choice of basis*, which we have suppressed.

Our goal is to obtain an analogue of Lemma 20.1 in [5], which shows that while $[\,\cdot\,]$ is not multiplicative, a suitable result exists to separate $[zw]$ into $[z][w]\kappa(zw)$, where $|\kappa(zw)| = 1$ and κ can be described explicitly. To do so we need to introduce an analogue of the "twist factor" $\xi_w(z)$ in [5]. Defining the analogue of $\xi_w(z)$ in the present setting is tricky, due to the fact that in general $\mathcal{O}_K$ need not be a unique factorization domain. In fact the situation is even more delicate than that; in order for our $\xi_w(z)$ to have nice properties, we must restrict the ideal classes of w, z as well as requiring w, z to satisfy a congruence condition like in (10-2).

To prepare for our definition, we first gather several of the key properties satisfied by Friedlander and Iwaniec's $\xi_w(z)$ in [5]. In particular, it satisfies the following:

(1) It satisfies an equation of the form

$$[z][w] = \varepsilon[zw]\xi_w(z),$$

where $\varepsilon = \pm 1$ depending only on the quadrants containing z, w, respectively.

(2) It is multiplicative for each $w \in \mathbb{Z}[i]$: one has $\xi_w(z_1)\xi_w(z_2) = \xi_w(z_1z_2)$.

(3) It is symmetric: $\xi_w(z) = \xi_z(w)$ for $w, z \in \mathbb{Z}[i]$.

(4) (Lemma 21.1 in [5]) For $q = |w_1 w_2|^2$ and $d = |\gcd(w_1, \overline{w_2})|^2$ one has

$$\sum_{\zeta \,(\mathrm{mod}\, q)} \xi_{w_1}(\zeta)\xi_{w_2}(\zeta) = \begin{cases} q\varphi(d)\varphi(q/d) & \text{if } q, d \text{ are squares,} \\ 0 & \text{otherwise.} \end{cases}$$

(5) For $w = u + iv$ and $\omega \equiv -v\bar{u} \pmod{q}$ with $q = |w|^2$, one has

$$\xi_w(z) = \left(\frac{ur - vs}{q}\right) \quad \text{and} \quad \xi_w(z) = \left(\frac{r + \omega s}{q}\right),$$

where $z = r + is$.

We would like to define our function $\xi_\alpha(z)$ to have the same properties. Unfortunately, it seems that at least some of these properties require special structures of the Gaussian integers $\mathbb{Z}[i]$. Thus, some more preparatory work is needed before we can define our stand-in for the symbol $\xi_w(z)$. We then check that our analogous symbol has the necessary properties to carry out the proofs of analogous statements in [5].

First we note that our symbol $\xi_\alpha(z)$ depends on α, and in particular, depends on the class A of α. This of course is a trivial point when $K = \mathbb{Q}(i)$, since $\mathbb{Z}[i]$ has unique factorization. Next we will also need to restrict the class of the *inputs* z, in order for our symbol to be well-behaved. This is far from ideal and is likely too restrictive, but it suffices for our purposes in this paper. Indeed, later we will see that it is necessary to define a separate symbol ξ for each class of ideal numbers *along with a basis* of said ideal numbers.

The most important property turns out to be (1), so we define our symbol with this in mind. To simplify matters we will assume that in our composition law the bilinear form $Q_{A,B}(w, z)$ is given by $w_1 z_1 + w_2 z_2$. In particular, we fix bases $\{\alpha_1, \alpha_2\} \subset A$, $\{\beta_1, \beta_2\} \subset B$, $\{\gamma_1, \gamma_2\} \subset C = A \cdot B$ so that

$$(\alpha_1 x_1 + \alpha_2 x_2)(\beta_1 y_1 + \beta_2 y_2) = (x_1 \ell_1(y_1, y_2) + x_2 \ell_2(y_1, y_2))\gamma_1 + (x_1 y_1 + x_2 y_2)\gamma_2.$$

We begin with the Jacobi symbol

$$\left(\frac{w_1 z_1 + w_2 z_2}{|w_1 \ell_1(z_1, z_2) + w_2 \ell_2(z_1, z_2)|}\right),$$

where $R_{A,B}(w, z) = w_1 \ell_1 + w_2 \ell_2$. We can extend the definition of the Jacobi symbol by setting

$$\left(\frac{a}{b}\right) = \left(\frac{a}{|b|}\right)(a, b)_\infty,$$

where

$$(a, b)_\infty = \begin{cases} -1 & \text{if } a, b < 0, \\ 1 & \text{otherwise,} \end{cases}$$

is the Hilbert symbol. Next we note quadratic reciprocity, which states for a, b odd and coprime that

$$\left(\frac{a}{|b|}\right)\left(\frac{b}{|a|}\right) = (-1)^{((a-1)/2)\cdot((b-1)/2)}(a, b)_\infty \tag{11-2}$$

and for $d > 0$ odd we have

$$\left(\frac{2}{d}\right) = (-1)^{(d^2-1)/4}.$$

Clearly, not both $Q_{A,B}$, $R_{A,B}$ can be even otherwise the corresponding ideal number is not primitive. Without loss of generality, let us suppose that $w_1z_1 + w_2z_2$ is odd. Let 2^k be the highest power of 2 dividing $w_1\ell_1(z_1, z_2) + w_2\ell_2(z_1, z_2)$. Then

$$\left(\frac{w_1z_1 + w_2z_2}{|w_1\ell_1(z_1, z_2) + w_2\ell_2(z_1, z_2)|}\right) = \left(\frac{w_1z_1 + w_2z_2}{2^{-k}|w_1\ell_1(z_1, z_2) + w_2\ell_2(z_1, z_2)|}\right).$$

We put

$$u = w_1z_1 + w_2z_2, \quad v = w_1\ell_1(z_1, z_2) + w_2\ell_2(z_1, z_2)$$

for simplicity. Applying quadratic reciprocity (11-2) then gives

$$\begin{aligned}\left(\frac{w_1z_1 + w_2z_2}{2^{-k}(w_1\ell_1(z_1, z_2) + w_2\ell_2(z_1, z_2))}\right)(u, v)_\infty &= \left(\frac{2^{-k}v}{u}\right)(-1)^{((u-1)/2)\cdot((2^{-k}v-1)/2)} \\ &= \left(\frac{2^k}{u}\right)(-1)^{((u-1)/2)\cdot((2^{-k}v-1)/2)}\left(\frac{v}{u}\right).\end{aligned}$$

Let us write $u_1 = \gcd(u, z_2)$ and $u_2 = u/u_1$. From the definition we see that $u_1 = \gcd(w_1, z_2)$. Put $w_1 = u_1w_1^*$, $z_2 = u_1z_2^*$ with $\gcd(w_1^*, z_2^*) = 1$. We now make use of the fact

$$w_1z_1 + w_2z_2 \equiv 0 \pmod{u}. \tag{11-3}$$

We treat the congruence modulo u_2 first. Plainly, $\gcd(z_2, u_2) = 1$. (11-3) then implies

$$w_2 \equiv -z_2^{-1}w_1z_1 \pmod{u_2}.$$

Substituting this into $R_{A,B}(w, z)$ gives

$$\begin{aligned}w_1\ell_1(z_1, z_2) + w_2\ell_2(z_1, z_2) &\equiv w_1\ell_1(z_1, z_2) - z_2^{-1}w_1z_1\ell_2(z_1, z_2) \pmod{u_2} \\ &\equiv z_2^{-1}w_1(z_2\ell_1(z_1, z_2) - z_1\ell_2(z_1, z_2)) \pmod{u_2}.\end{aligned}$$

Modulo u_1 we must have

$$\begin{aligned}w_1\ell_1(z_1, z_2) + w_2\ell_2(z_1, z_2) &\equiv w_2\ell_2(z_1, z_2) \pmod{u_1} \\ &\equiv -z_1^{-1}w_2(z_2\ell_1(z_1, z_2) - z_1\ell_2(z_1, z_2)) \pmod{u_1}.\end{aligned}$$

In both cases, we see that $R_{A,B}(w, z)$ is congruent to a multiple of the quadratic form

$$g(z_1, z_2) = z_2\ell_1(z_1, z_2) - z_1\ell_2(z_1, z_2),$$

which we now interpret. By definition, our composition law gives the relation

$$\begin{aligned}(z_2\alpha_1 - z_1\alpha_2)(z_1\beta_1 + z_2\beta_2) &= R_{A,B}(z_2, -z_1; z_1, z_2)\gamma_1 + Q_{A,B}(z_2, -z_1; z_1, z_2)\gamma_2 \\ &= (z_2\ell_1(z_1, z_2)) - z_1\ell_2(z_1, z_2)\gamma_1 + (z_2z_1 - z_1z_2)\gamma_2 \\ &= g(z_1, z_2)\gamma_1. \end{aligned}\tag{11-4}$$

Dividing both sides by γ_1 we then see that $g(z_1, z_2)$ must be equivalent to the norm form of $\mathcal{O}_K$.

We must now relate $g(z_1, z_2)$ to $N(z) = N(J(z_1\beta_1 + z_2\beta_2))$. Note that

$$g(z_1, z_2) = \gamma_1^{-1}(\alpha_1 z_2 - \alpha_2 z_1)(\beta_1 z_1 + \beta_2 z_2)$$

is divisible by $z = \beta_1 z_1 + \beta_2 z_2$, which implies that $g(z_1, z_2)$ is a rational integer divisible by $N(z)$. By primitivity we then see that $g(z_1, z_2)$ must be a constant multiple of $N(z)$, the constant depending only on the classes A, B. We summarize this as a lemma:

Lemma 11.1. *Let $g(x, y)$ be the integral binary quadratic form which arises from the composition law* (11-4)*. Then $g(z_1, z_2)$ is a constant multiple of $N(J(\beta_1 z_1 + \beta_2 z_2))$, with the constant depending only on the classes A, B and choices of bases of A, B, $A \cdot B$.*

Similarly, since $v = w_1\ell_1 + w_2\ell_2$ is divisible by 2^k, we may assume without loss of generality that ℓ_1 is odd to obtain

$$w_1 \equiv -\ell_1^{-1} w_2 \ell_2 \pmod{2^k}$$

and this implies that

$$\begin{aligned} w_1 z_1 + w_2 z_2 &\equiv -\ell_1^{-1} w_2 \ell_2 z_1 + w_2 z_2 \pmod{2^k} \\ &\equiv -\ell_1^{-1} w_2 (z_2 \ell_1 - z_1 \ell_2) \pmod{2^k} \\ &\equiv -\ell_1^{-1} w_2 g(z_1, z_2) \pmod{2^k}. \end{aligned}$$

Since $u = w_1 z_1 + w_2 z_2$ is odd by assumption, it follows that $g(z_1, z_2)$ must be odd as well.

Continuing on, with $u = u_1 u_2$ as before, we have

$$\begin{aligned} z_2^{-1} w_1 (z_2 \ell_1(z_1, z_2) - z_1 \ell_2(z_1, z_2)) &\equiv z_2^{-1} w_1 g(z_1, z_2) \pmod{u_2}, \\ -z_1^{-1} w_2 (z_2 \ell_1(z_1, z_2) - z_1 \ell_2(z_1, z_2)) &\equiv -z_1^{-1} w_2 g(z_1, z_2) \pmod{u_1} \end{aligned}$$

which implies that

$$\begin{aligned} \left(\frac{v}{u}\right) &= \left(\frac{-z_1^{-1} w_2 g(z_1, z_2)}{u_1}\right)\left(\frac{z_2^{-1} w_1 g(z_1, z_2)}{u_2}\right) \\ &= \left(\frac{-z_1 w_2}{u_1}\right)\left(\frac{z_2^* w_1^*}{u_2}\right)\left(\frac{g(z_1, z_2)}{u}\right). \end{aligned}$$

Observe that, by definition, we have

$$u_2 = w_1^* z_1 + w_2 z_2^*.$$

Since u is odd, it follows that exactly one of the pairs $\{w_1, z_1\}$, $\{w_2, z_2\}$ consists of two odd numbers. Without loss of generality, we assume that w_1, z_1 are both odd. We write $z_2^* = 2^{k_2}\nu_2$ with ν_2 odd. Applying

(11-2) we find that

$$\begin{aligned}\left(\frac{w_1^* z_2^*}{u_2}\right) &= \left(\frac{2^{k_2}}{u_2}\right)\left(\frac{w_1^* \nu_2}{w_1^* z_1 + w_2 z_2^*}\right)\\ &= \left(\frac{2^{k_2}}{u_2}\right)(-1)^{(u_2-1)/2}(-1)^{(w_1^*\nu_2-1)/2}\left(\frac{w_1^* z_1 + w_2 z_2^*}{|w_1^*\nu_2|}\right)\\ &= \left(\frac{2^{k_2}}{u_2}\right)(-1)^{(u_2-1)/2}(-1)^{(w_1^*\nu_2-1)/2}\left(\frac{w_2 z_2^*}{|w_1^*|}\right)\left(\frac{w_1^* z_1}{|\nu_2|}\right)\\ &= \left(\frac{2^{k_2}}{u_2}\right)(-1)^{(u_2-1)/2}(-1)^{(w_1^*\nu_2-1)/2}\left(\frac{w_2}{|w_1^*|}\right)\left(\frac{z_2^*}{|w_1^*|}\right)\left(\frac{w_1^*}{|\nu_2|}\right)\left(\frac{z_1}{|\nu_2|}\right).\end{aligned}$$

Applying (11-2) repeatedly we obtain

$$\left(\frac{w_1^* z_2^*}{u_2}\right) = \varepsilon_1\left(\frac{w_2}{|w_1^*|}\right)\left(\frac{\nu_2}{|z_1|}\right) \tag{11-5}$$

where

$$\varepsilon_1 = \left(\frac{2^{k_2}}{u_2}\right)\left(\frac{2^{k_2}}{|w_1^*|}\right)(-1)^{(u_2-1)/2}(-1)^{(w_1^*\nu_2-1)/2}(-1)^{(w_1^*-1)/2}(-1)^{(z_1-1)/2}(w_1^*, \nu_2)_\infty(\nu_2, z_1)_\infty. \tag{11-6}$$

Next we note that

$$\left(\frac{-z_1 w_2}{u_1}\right) = \left(\frac{-1}{u_1}\right)\left(\frac{z_1}{u_1}\right)\left(\frac{w_2}{u_1}\right).$$

It follows that

$$\begin{aligned}\left(\frac{-z_1 w_2}{u_1}\right)\left(\frac{z_2^* w_1^*}{u_2}\right) &= \varepsilon_1\left(\frac{-1}{u_1}\right)\left(\frac{u_1}{|z_1|}\right)(-1)^{(w_2-1)/2}(-1)^{(z_1-1)/2}(w_1, u_1)_\infty\left(\frac{w_2}{|u_1|}\right)\left(\frac{w_2}{|w_1^*|}\right)\left(\frac{\nu_2}{|z_1|}\right)\\ &= \varepsilon_2\left(\frac{z_2}{|z_1|}\right)\left(\frac{w_2}{|w_1|}\right).\end{aligned}$$

Here we have

$$\varepsilon_2 = \varepsilon_1(-1)^{(w_2-1)/2}(-1)^{(z_1-1)/2}(w_1, u_1)_\infty\left(\frac{-1}{u_1}\right)\left(\frac{2^{k_2}}{|z_1|}\right). \tag{11-7}$$

Finally, by (11-2) we have

$$\left(\frac{u}{g(z_1, z_2)}\right) = (-1)^{(g(z_1,z_2)-1)/2}(-1)^{(u-1)/2}\left(\frac{g(z_1, z_2)}{u}\right).$$

Collecting these calculations we conclude that

$$\left(\frac{w_1\ell_1(z_1, z_2) + w_2\ell_2(z_1, z_2)}{|w_1 z_1 + w_2 z_2|}\right) = \left(\frac{w_2}{|w_1|}\right)\left(\frac{z_2}{|z_1|}\right)\left(\frac{w_1 z_1 + w_2 z_2}{g(z_1, z_2)}\right)\varepsilon(w, z), \tag{11-8}$$

where

$$\varepsilon(w, z) = \varepsilon_2(-1)^{(g(z_1,z_2)-1)/2}(-1)^{(u-1)/2}. \tag{11-9}$$

From (11-9) we make the following conclusion:

Lemma 11.2. *Let $\varepsilon(w, z)$ be given as in* (11-9)*. Then $\varepsilon(w, z) \in \{\pm 1\}$ and its value is determined by the quadrants of (z_1, z_2), (w_1, w_2), the congruence classes of w_1, w_2, z_1, z_2 modulo* 8*, and whether* 2 *divides z_2 an odd or an even number of times.*

Since we have insisted that w, z belong to fixed congruence classes modulo $8etn$ as in (10-2) it follows that $\varepsilon(w, z)$ can be determined as a function of the congruence class alone, except for the condition on whether z_2 is divisible by an even or odd power of 2. We can treat the two cases separately, and in each case assume that $\varepsilon(w, z)$ is constant.

These calculations compel us to define our analogue of the twist factor in the multiplication law for the Jacobi–Kubota symbol as

$$\xi_w(z) = \left(\frac{w_1 z_1 + w_2 z_2}{g(w_1, w_2)}\right). \tag{11-10}$$

Note that $\xi_w(z)$ depends on the ideal classes of w, z and a choice of basis for the ideal classes.

Next we observe for w, z satisfying (10-2), w, z are in the same class and therefore $R_{A,B}(w, z) = R_{A,A}(w, z)$ must be symmetric in w, z. From here it follows that

$$\begin{aligned} z_2^{-1} w_1 g(z_1, z_2) &\equiv R_{A,A}(w, z) \pmod{u} \\ &\equiv R_{A,A}(z, w) \pmod{u} \\ &\equiv w_1^{-1} z_2 g(w_1, w_2) \pmod{u}. \end{aligned}$$

This implies that

$$\left(\frac{g(z_1, z_2)}{u}\right)\left(\frac{g(w_1, w_2)}{u}\right) = 1.$$

Thus, up to a factor ε depending at most on congruence classes and signs of w, z, we have

$$\xi_w(z) = \varepsilon \xi_z(w). \tag{11-11}$$

Summarizing, we obtain the following analogue of Lemma 20.1 in [6]:

Lemma 11.3. *Let w, z satisfy the hypothesis given in* (10-2)*. Then there exist numbers $\theta(w, z) \in \{-1, 1\}$ depending only on the signs and congruence classes of w, z modulo $8etn$ such that*

$$\left(\frac{Q_{A,B}(w, z)}{|R_{A,B}(w, z)|}\right) = \theta(w, z)\left(\frac{w_2}{|w_1|}\right)\left(\frac{z_2}{|z_1|}\right)\xi_z(w). \tag{11-12}$$

Next we show that the analogue of Lemma 21.1 in [5] holds:

Lemma 11.4. *For fixed elements w, v in the class A and*

$$q = g(w_1, w_2)g(v_1, v_2) \quad \text{and} \quad d = \gcd(g(w_1, w_2), g(v_1, v_2)),$$

we have

$$\sum_{z \,(\mathrm{mod}\, q)} \xi_{w_1}(z)\xi_{w_2}(z) = \begin{cases} q\varphi(d)\varphi(q/d) & \text{if } q, d \text{ are squares,} \\ 0 & \text{otherwise.} \end{cases}$$

Proof. We have

$$\sum_{z \,(\mathrm{mod}\, q)} \xi_w(z)\xi_v(z) = \sum_{z \,(\mathrm{mod}\, q)} \left(\frac{w_1 z_1 + w_2 z_2}{g(w_1, w_2)}\right)\left(\frac{v_1 z_1 + v_2 z_2}{g(v_1, v_2)}\right)$$
$$= \sum_{z \,(\mathrm{mod}\, q)} \left(\frac{(w_1 z_1 + w_2 z_2)(v_1 z_1 + v_2 z_2)}{d}\right)\left(\frac{w_1 z_1 + w_2 z_2}{g(w_1, w_2)/d}\right)\left(\frac{v_1 z_1 + v_2 z_2}{g(v_1, v_2)/d}\right).$$

From here we see that the final sum is zero unless each of the summands is equal to 1 or 0 identically. This is only the case when d, $g(w_1, w_2)/d$, $g(w_1, w_2)/d$ are all squares. Since $d \mid \gcd(g(w_1, w_2), g(v_1, v_2))$ and $d \nmid \Delta(f)$ it follows that $w_1 x + w_2 y$, $v_1 x + v_2 y$ are not proportional modulo d. From here we see that, modulo d, the number of solutions to $\gcd(w_1 x + w_2 y, d) = \gcd(v_1 x + v_2 y, d) = 1$ is equal to $\varphi(d)^2$. Similarly, modulo $g(w_1, w_2)/d$ and $g(v_1, v_2)/d$ there are $g(w_1, w_2)\varphi(g(w_1, w_2)/d)/d$ solutions to $\gcd(w_1 x + w_2 y, g(w_1, w_2)/d) = 1$ and $\gcd(v_1 x + v_2 y, g(v_1, v_2)/d) = 1$, respectively. Lifting to the modulus q yields

$$\frac{q^2}{d^2} \cdot \varphi(d)^2 \cdot \frac{g(w_1, w_2)g(v_1, v_2)}{d^2}\varphi(g(w_1, w_2)/d)\varphi(g(v_1, v_2)/d) = q\varphi(d)\varphi(q/d),$$

since $\gcd(q/d^2, d) = 1$. This completes the proof. □

Lemma 11.4 is analogous to Lemma 21.1 in [5].

We now prove the following analogue of Lemma 21.2 in [5]:

Proposition 11.5. *Let A, B be classes of ideal numbers. Put*

$$\mathcal{Q}(M, N) = \sum_w^* \sum_z \alpha_w \beta_z \xi_w(z), \tag{11-13}$$

where α_w, β_z are bounded real coefficients supported in appropriate fundamental domains for A, B having norm bounded by M, N, respectively. Then for all $\varepsilon > 0$ we have

$$\mathcal{Q}(M, N) \ll_\varepsilon (M + N)^{1/12}(MN)^{1/12+\varepsilon}. \tag{11-14}$$

Proof. Applying Cauchy's inequality we obtain

$$|\mathcal{Q}(M, N)|^2 \le \|\beta\|_2^2 \sum_z \left|\sum_w^* \alpha_w \xi_w(z)\right|^2$$
$$= \|\beta\|_2^2 \sum_{w_1}^* \sum_{w_2}^* \alpha_{w_1}\alpha_{w_2} \sum_z \xi_{w_1}(z)\xi_{w_2}(z).$$

We then find that splitting z into congruence classes modulo $q = g(w_1)g(w_2)$ that

$$\sum_z \xi_{w_1}(z)\xi_{w_2}(z) = \sum_{\zeta \,(\mathrm{mod}\, q)} \xi_{w_1}(\zeta)\xi_{w_2}(\zeta) \cdot \left(\frac{c_f N}{q^2} + O_f\left(\frac{\sqrt{N}}{q} + 1\right)\right)$$

where

$$c_f = \lim_{s \to 1}(s - 1)\zeta_K(s).$$

We obtain, by Lemma 11.1 and using (11-11) if necessary,

$$\mathcal{Q}(M, N)^2 \ll N^2 \sum_{\substack{m_1, m_2 \leq M \\ m_1 m_2 = \square}} \sum \tau(m_1 m_2) + NM^4\big(\sqrt{N} + M^2\big), \tag{11-15}$$

which gives the bound

$$\mathcal{Q}(M, N) \ll_\varepsilon (M^3 N^{1/2} + M^2 N^{3/4} + M^{1/2} N)(MN)^\varepsilon.$$

In the next step we shall apply Hölder's inequality to obtain

$$\mathcal{Q}(M, N)^k \ll M^{k-1} \sum_w{}^* \left| \sum_z \beta_z \xi_w(z) \right|^k = M^{k-1} \widetilde{\mathcal{Q}}(M, N^k),$$

say. In [5] the next step is to argue that $\widetilde{\mathcal{Q}}(M, N^k)$ can be written as a bilinear form of the shape (11-13), using the fact that in the case $K = \mathbb{Q}(i)$ that $\xi_w(z)$ is multiplicative in z. In general this is not the case. However, we are free to choose a basis for the class B^k for each positive integer k, which allows one to write

$$\xi_w(z_1) \cdots \xi_w(z_k) = \xi_w^{(k)}(z_1 \cdots z_k) \tag{11-16}$$

in a consistent way. Recalling (11-10), we note that

$$\xi_w(z_1)\xi_w(z_2) = \left(\frac{Q_{B,B}(z_1) Q_{B,B}(z_2)}{g(w_1, w_2)} \right).$$

The numerator is a bilinear form in z_1, z_2. Using composition laws to write

$$z_1 z_2 = R_{B^2}(z_1, z_2)\gamma_1^{(2)} + Q_{B^2}(z_1, z_2)\gamma_2^{(2)}$$

as ideal numbers, we see that we can apply a change of variables, depending only on w, the class B, and the choice of bases, so that the numerator $Q_{B,B}(z_1)Q_{B,B}(z_2)$ is a linear form in $R_{B^2}(z_1, z_2)$, $Q_{B^2}(z_1, z_2)$. Inductively, we then find that

$$\xi_w(z_1) \cdots \xi_w(z_k) = \left(\frac{L_w(z_1 \cdots z_k)}{g(w_1, w_2)} \right),$$

where L_w is a linear form in two variables with coefficients depending at most on w and evaluates $z_1 \cdots z_k$ in terms of its representation as an element in the lattice of the corresponding ideal numbers. Defining the right-hand side as $\xi_w^{(k)}(z_1 \cdots z_k)$ we obtain (11-16). Replacing $\xi_w(\cdot)$ with $\xi_w^{(k)}(\cdot)$ in (11-13) shows that (11-15) holds, and therefore we may proceed as in [5] after applying Hölder's inequality to conclude

$$\mathcal{Q}(M, N)^k \ll_\varepsilon M^{k-1}(M^3 N^{k/2} + M^2 N^{3k/4} + M^{1/2} N^k)(MN)^\varepsilon,$$

which upon taking k-th roots gives us the bound

$$\mathcal{Q}(M, N) \ll_\varepsilon (M^{1+2/k} N^{1/2} + M^{1+1/k} N^{3/4} + M^{1-1/2k} N)(MN)^\varepsilon$$

for all positive $k \in \mathbb{N}$. Switching the roles of M, N and applying Lemma 11.4, we obtain as in [5] that

$$\mathcal{Q}(M, N) \ll_\varepsilon (M+N)^{1/12}(MN)^{1/12+\varepsilon}$$

upon setting $k = 6$. □

Next we move on to proving the analogue of Proposition 22.1 in [5]. We define, for any ideal number z, a rational integer k, and a character χ modulo $4d$ the Hecke character

$$\psi(z) = \chi(z)\left(\frac{z}{|z|}\right)^k. \tag{11-17}$$

Consider the sum

$$\mathcal{K}(N) = \sideset{}{^\wedge}\sum_{z\in\mathfrak{B}} \psi(z)\,[wz]$$

and

$$\mathcal{K}^*(N) = \sideset{}{^\wedge}\sum_{\substack{z\in\mathfrak{B}\\ \gcd(z,\,w)=1}} \psi(z)\,[wz],$$

where $\mathfrak{B}$ is a narrow sector contained in the intersection of a fundamental domain for the ideal class numbers containing z having norm bounded N. We treat w as a fixed primitive ideal number. Our analogue of Proposition 22.1 in [5] is thus:

Proposition 11.6. *Given ψ and w as above we have*

$$\mathcal{K}(N) \ll d(|k|+1)\,|w|\,N^{3/4}\log(|w|\,N) \tag{11-18}$$

and

$$\mathcal{K}^*(N) \ll d(|k|+1)\,|w|\,\tau(N(w))N^{3/4}\log(|w|\,N). \tag{11-19}$$

Proof. Just like the proof of Proposition 22.1 in [5], the key result needed to obtain the necessary cancellation is the Polya–Vinogradov theorem, which asserts that

$$\sum_{n\le N}\chi(n) \ll \sqrt{q}\log q$$

for every nontrivial Dirichlet character $\chi \pmod q$ with an absolute implied constant. To estimate $\mathcal{K}(N)$ we apply Lemma 11.3 to obtain

$$\mathcal{K}(N) = [w]\sideset{}{^\wedge}\sum_{z\in\mathfrak{B}} \varepsilon(w,z)\,\psi(z)\,[z]\,\xi_w(z),$$

and by breaking the sum up to finitely many congruence classes if necessary, we may factor the ε-factor out (because it will be constant) to obtain

$$\mathcal{K}(N) = [w]\varepsilon\sideset{}{^\wedge}\sum_{z\in\mathfrak{B}} \psi(z)\,[z]\,\xi_w(z).$$

Breaking the sum up into a double sum over rational integers forming vectors running over $\mathfrak{B}$ as in [5] and applying Polya–Vinogradov we obtain (11-18) and (11-19) as required. □

Put

$$\lambda(n) = \sideset{}{^\wedge}\sum_{N(z)=n} \psi(z)[z],$$

the sum restricted to a fundamental domain of ideal numbers so each ideal is represented at most once. Consider the sum

$$\mathcal{L}(M, N) = \sum_m \sum_n \alpha(m)\beta(n)\lambda(cmn), \tag{11-20}$$

where α, β are complex coefficients having norm at most 1 and supported on $1 \le m \le M$ and $n \le N$. Likewise, let $\mathcal{L}^*(M, N)$ be the subsum of (11-20) restricted to $\gcd(m, n) = 1$. Combining Proposition 11.5 and Lemma 11.3 then gives the following analogue of Proposition 23.1 in [5]:

Proposition 11.7. *For any complex coefficients $\alpha(m)$, $\beta(n)$ as above and for any positive integer c,*

$$\mathcal{L}(M, N) \ll \tau(c)(M + N)^{1/12}(MN)^{1/12+\varepsilon}. \tag{11-21}$$

We also introduce the analogues of $\mathcal{K}(N)$, $\mathcal{K}^*(N)$:

$$\mathcal{L}(N) = \sum_{n \le N} \lambda(mn), \quad \mathcal{L}^*(N) = \sum_{\substack{n \le N \\ \gcd(m,n)=1}} \lambda(mn) \tag{11-22}$$

and obtain the following analogue of Proposition 23.2 in [5] by applying Proposition 11.6:

Proposition 11.8. *For ψ as defined by* (11-17) *and positive integer m we have the bounds*

$$\mathcal{L}(N) \ll d(|k| + 1)\tau(m)^4\sqrt{m}N^{3/4}\log(mN) \tag{11-23}$$

and

$$\mathcal{L}^*(N) \ll d(|k| + 1)\tau(m)^2\sqrt{m}N^{3/4}\log(mN). \tag{11-24}$$

These estimates then imply the following analogue of Theorem ψ in [5]:

Proposition 11.9. *For any $c \ge 1$ we have*

$$\sum_{n \le X} \Lambda(n)\lambda(cn) \ll cd(|k| + 1)X^{6/77} \tag{11-25}$$

with the absolute constant dependent only on f.

Proof. This is the same as the proof of Theorem ψ in [5] with Propositions 23.1 and 23.2 replaced by Propositions 11.7 and 11.8, respectively. □

Acknowledgements

The author owes an incalculable debt of gratitude to John Friedlander, whose encouragement and guidance made this paper possible. The author also thanks D. R. Heath-Brown whose work on prime number theory is an inspiration for the present work, D. Schindler and J. Maynard for helpful discussions, to C. L. Stewart for a careful reading of an earlier version of this paper and for providing instrumental advice, and to S. Yamagishi whose collaboration and friendship was instrumental in the author's pursuit of prime number theory. The author would like to express his deepest appreciation for the hard work done by the referee and for the helpful suggestions made.

References

[1] A. Balog, V. Blomer, C. Dartyge, and G. Tenenbaum, "Friable values of binary forms", *Comment. Math. Helv.* **87**:3 (2012), 639–667. MR Zbl

[2] J. G. van der Corput, "Über Summen von Primzahlen und Primzahlquadraten", *Math. Ann.* **116**:1 (1939), 1–50. MR Zbl

[3] E. Fouvry and H. Iwaniec, "Gaussian primes", *Acta Arith.* **79**:3 (1997), 249–287. MR Zbl

[4] J. Friedlander and H. Iwaniec, "Asymptotic sieve for primes", *Ann. of Math.* (2) **148**:3 (1998), 1041–1065. MR Zbl

[5] J. Friedlander and H. Iwaniec, "The polynomial $X^2 + Y^4$ captures its primes", *Ann. of Math.* (2) **148**:3 (1998), 945–1040. MR Zbl

[6] J. B. Friedlander and H. Iwaniec, "Gaussian sequences in arithmetic progressions", *Funct. Approx. Comment. Math.* **37** (2007), 149–157. MR Zbl

[7] B. Green, "Roth's theorem in the primes", *Ann. of Math.* (2) **161**:3 (2005), 1609–1636. MR Zbl

[8] B. Green and T. Tao, "The primes contain arbitrarily long arithmetic progressions", *Ann. of Math.* (2) **167**:2 (2008), 481–547. MR Zbl

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

[10] D. R. Heath-Brown and X. Li, "Prime values of $a^2 + p^4$", *Invent. Math.* **208**:2 (2017), 441–499. MR Zbl

[11] D. R. Heath-Brown and B. Z. Moroz, "Primes represented by binary cubic forms", *Proc. Lond. Math. Soc.* (3) **84**:2 (2002), 257–288. MR Zbl

[12] P. C.-H. Lam, D. Schindler, and S. Y. Xiao, "On prime values of binary quadratic forms with a thin variable", *J. Lond. Math. Soc.* (2) **102**:2 (2020), 749–772. MR Zbl

[13] B. Landreau, "Majorations de fonctions arithmétiques en moyenne sur des ensembles de faible densité", exposé 13 in *Séminaire de Théorie des Nombres* (Talence, France, 1987-1988), Univ. Bordeaux I, Talence, France, 1988. MR Zbl

[14] X. Li, "Prime values of a sparse polynomial sequence", *Duke Math. J.* **171**:1 (2022), 101–208. MR Zbl

[15] J. Maynard, "Primes represented by incomplete norm forms", *Forum Math. Pi* **8** (2020), art. id. e3. MR Zbl

[16] T. Mitsui, "Generalized prime number theorem", *Jpn. J. Math.* **26** (1956), 1–42. MR Zbl

[17] M. Pandey, "On Eisenstein primes", *Integers* **18** (2018), art. id. A59. MR Zbl

[18] R. C. Vaughan, "Mean value theorems in prime number theory", *J. Lond. Math. Soc.* (2) **10** (1975), 153–162. MR Zbl

Communicated by Andrew Granville
Received 2022-07-17 Revised 2023-08-30 Accepted 2023-10-12

stanleyyao.xiao@unbc.ca *Department of Mathematics and Statistics, University of Northern British Columbia, Prince George, Canada*

Algebra & Number Theory

msp.org/ant

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2024 is US $525/year for the electronic version, and $770/year (+$65, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online.

ANT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

Algebra & Number Theory

Volume 18 No. 9 2024