

PRIMES FROM SUMS OF TWO SQUARES AND MISSING DIGITS

KYLE PRATT

ABSTRACT. Let \mathcal{A}' be the set of integers missing any three fixed digits from their decimal expansion. We produce primes in a thin sequence by proving an asymptotic formula for counting primes of the form $p = m^2 + \ell^2$, with $\ell \in \mathcal{A}'$.

The proof draws on ideas from the work of Friedlander-Iwaniec on primes of the form $p = x^2 + y^4$, as well as ideas from the work of Maynard on primes with restricted digits.

TABLE OF CONTENTS

1. Introduction	1
2. Initial manipulations and outline	4
3. The sieve remainder term	7
4. The sieve main term	10
5. The sieve main term: fundamental lemma	15
6. Bilinear form in the sieve: first steps	20
7. Bilinear form in the sieve: transformations	30
8. Coprimality conditions and further reductions	37
9. Polar boxes and congruence exercises	44
10. Simplifications and endgame	53
11. Modifications for Theorem 1.2	59
Acknowledgements	62
References	62

1. INTRODUCTION

Some of the most interesting questions in analytic prime number theory arise from interactions with themes and ideas from other areas of mathematics. The famous twin prime conjecture, for example, arises from placing the multiplicative notion of a prime number in an additive context.

An early instance of this phenomenon is in Fermat’s 1640 “Christmas letter” to Marin Mersenne [26, pp. 212-217], wherein he describes which numbers may be written as a sum of two integral squares (Fermat phrased his observations in terms of integers appearing as hypotenuses of right triangles). Along the way he noted that every prime $p \equiv 1 \pmod{4}$ may be written as $p = x^2 + y^2$, but in true Fermat fashion he supplied no proof¹.

2010 *Mathematics Subject Classification*. 11N05, 11N32, 11N36, 11A63.

Keywords and phrases: prime numbers, sum of two squares, thin sequence, missing digit.

¹Euler finally found a proof more than a century later [12].

At first glance the equation $p = x^2 + y^2$ looks like an additive equation involving primes, but with the benefit of substantial hindsight we see this is in fact a multiplicative problem, since $x^2 + y^2$ is the norm form of the algebraic number field $\mathbb{Q}(i)$.

Other famous problems in prime number theory concern primes in “thin” sequences, such as primes in short intervals, or primes of the form $p = n^2 + 1$. A set of integers $\mathcal{S} \subset [1, x]$ is *thin* if there are few elements of \mathcal{S} relative to x (think $|\mathcal{S}| \leq x^{1-\delta}$ for some $\delta > 0$). It is natural to ask under what conditions \mathcal{S} contains prime numbers, but often these questions are very hard. Most often one needs the set \mathcal{S} to have some nice multiplicative structure to exploit.

Several authors have proved the existence of infinitely many primes within different thin sequences. Fouvry and Iwaniec [13] proved there are infinitely many primes of the form $p = m^2 + q^2$, where q is a prime number. The set $\{m^2 + q^2 \leq x : q \text{ prime}\}$ has size $\ll x(\log x)^{-2}$, and so is thin in the sense that it has zero density inside of the primes. This is a nice example of additively-structured primes in a thin sequence.

Friedlander and Iwaniec [15] built on the foundation laid by Fouvry and Iwaniec, and proved there are infinitely many primes of the form $p = x^2 + y^4$. This is a much thinner sequence of primes than those considered by Fouvry and Iwaniec, and consequently the proof is much more difficult. It is crucial for the work of Friedlander and Iwaniec that $x^2 + y^4 = x^2 + (y^2)^2$.

Other striking examples are the works of Heath-Brown [18] on primes of the form $p = x^3 + 2y^3$ and Heath-Brown and Moroz [20] on primes represented by cubic forms, and Maynard [23] on primes represented by incomplete norm forms. Heath-Brown and Li [19] refined the theorem of Friedlander and Iwaniec by showing there are infinitely many primes of the form $p = x^2 + q^4$, where q is a prime. Each of these results relies heavily on the fact that the underlying polynomial is related to the norm form of an algebraic number field.

Polynomials offer one source of thin sequences, but they are not the only source. Particularly attractive are other, more exotic, thin sequences, like the set of integers missing a fixed digit from their decimal expansion. To be precise, let $a_0 \in \{0, 1, 2, \dots, 9\}$ be fixed, and let \mathcal{A} be the set of nonnegative integers without the digit a_0 in their decimal expansion. We write $\mathbf{1}_{\mathcal{A}}$ for the indicator function of this set. We define

$$\gamma_0 := \frac{\log 9}{\log 10} = 0.954\dots,$$

and note that

$$(1.1) \quad \sum_{\ell \leq y} \mathbf{1}_{\mathcal{A}}(\ell) \asymp y^{\gamma_0}, \quad y \geq 2.$$

Our goal is to tie together several different mathematical strands by proving there are infinitely many primes p of the form $p = m^2 + \ell^2$, where $\ell \in \mathcal{A}$. Note that $\frac{1}{2} + \frac{\gamma_0}{2} < 1$, so this sequence of primes is indeed thin.

The present work was inspired by Maynard’s beautiful paper [24], wherein he showed there are infinitely many primes in the thin sequence \mathcal{A} . The key to the whole enterprise is that the Fourier transform of \mathcal{A} has remarkable properties. Exploiting this Fourier structure has been vital in works on digit-related problems (see, for example, [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 22] and the works cited therein). We also rely on this Fourier structure.

It turns out that we ultimately use few of the tools Maynard developed. Rather, our work is closer in spirit to the work of Fouvry and Iwaniec [13] and the work of Friedlander and Iwaniec [15].

Our basic strategy is to use a sieve to count the primes $p = m^2 + \ell^2$, where $\ell \in \mathcal{A}$. It is tempting to try and estimate the sum

$$(1.2) \quad \sum_{\substack{m^2 + \ell^2 \leq x \\ (\ell, 10) = 1}} \mathbf{1}_{\mathcal{A}}(\ell) \Lambda(m^2 + \ell^2),$$

but for technical reasons it is convenient to prove a stronger theorem in which ℓ has no small prime factors. This ensures that ℓ is almost always coprime to other variables.

Theorem 1.1. *Let x be large, and let $A > 0$ be fixed. Let P be a parameter which satisfies*

$$(\log \log x)^4 \leq \log P \leq \frac{\sqrt{\log x}}{\log \log x},$$

and define

$$\Pi := \prod_{p \leq P} p.$$

We then have

$$\sum_{\substack{m^2 + \ell^2 \leq x \\ (\ell, \Pi) = 1}} \mathbf{1}_{\mathcal{A}}(\ell) \Lambda(m^2 + \ell^2) = \frac{4\mathfrak{C}\kappa_1}{\pi} \frac{e^{-\gamma}}{\log P} \sum_{m^2 + \ell^2 \leq x} \mathbf{1}_{\mathcal{A}}(\ell) + O\left(x^{1/2 + \gamma_0/2} (\log x)^{-A}\right),$$

where γ denotes Euler's constant,

$$\mathfrak{C} := \prod_p \left(1 - \frac{\chi(p)}{(p-1)(p-\chi(p))}\right),$$

χ is the nonprincipal character modulo 4, and

$$\kappa_1 := \begin{cases} \frac{10}{9}, & (a_0, 10) \neq 1, \\ \frac{10(\varphi(10)-1)}{9\varphi(10)}, & (a_0, 10) = 1. \end{cases}$$

The implied constant depends on A and is ineffective.

Remark. It is potential exceptional zeros for certain Hecke L -functions that make the implied constant in Theorem 1.1 ineffective.

We are able to avoid more sophisticated sieves like Harman's sieve [17]; instead we require only Vaughan's identity (see (2.1)). The application of Vaughan's identity reduces the problem to the estimation of certain arithmetic "Type I" and "Type II" sums. The Type I information, which is quite strong, comes from a general result of Fouvry and Iwaniec (see Lemma 3.2). The strength of the Type I bound relies on the homogeneous nature of the polynomial $x^2 + y^2$. For the Type II sums we follow the outlines of the argument of Friedlander-Iwaniec. Our argument is less complicated in some places and more complicated in others. The desired cancellation eventually comes from an excursion into a zero-free region for Hecke L -functions.

We obtain Type II information in a wide interval, much wider than that which is required given our amount of Type I information. This suggests the possibility of finding primes of the form $p = m^2 + \ell^2$, where ℓ is missing more than one digit in its decimal expansion.

Theorem 1.2. *Let $\mathcal{B} \subset \{0, 1, \dots, 9\}$ satisfy $1 \leq |\mathcal{B}| \leq 3$, and let \mathcal{A}' denote the set of nonnegative integers whose decimal expansions consist only of the digits in $\{0, 1, \dots, 9\} \setminus \mathcal{B}$. Let*

$$\gamma_{\mathcal{B}} := \frac{\log(10 - |\mathcal{B}|)}{\log 10}.$$

Then, with the notation as above, we have

$$\sum_{\substack{m^2 + \ell^2 \leq x \\ (\ell, \Pi) = 1}} \sum \mathbf{1}_{\mathcal{A}'}(\ell) \Lambda(m^2 + \ell^2) = \frac{4\mathfrak{C}_{\mathcal{B}}}{\pi} \frac{e^{-\gamma}}{\log P} \sum_{m^2 + \ell^2 \leq x} \mathbf{1}_{\mathcal{A}'}(\ell) + O\left(x^{1/2 + \gamma_{\mathcal{B}}/2} (\log x)^{-A}\right),$$

where

$$\kappa_{\mathcal{B}} := \frac{10}{\varphi(10)} \frac{\varphi(10) + |\{a \in \mathcal{B} : (a, 10) \neq 1\}| - |\mathcal{B}|}{10 - |\mathcal{B}|}.$$

The implied constant depends on A and is ineffective.

Remark. When $|\mathcal{B}| = 3$, Theorem 1.2 shows the existence of primes in a set of integers of size $\ll x^{\frac{1}{2} + \frac{1}{2} \frac{\log 7}{\log 10}} \approx x^{0.9225}$. One may take $|\mathcal{B}|$ to be larger by using a more complicated sieve argument and imposing extra conditions on the elements of \mathcal{B} , but we do not pursue this here.

Throughout the paper we make use of asymptotic notation $\ll, \gg, O(\cdot)$, and $o(\cdot)$. We write $f \asymp g$ if $f \ll g$ and $f \gg g$. Usually the implied constants are absolute, but from Section 6 onward we allow the implied constants to depend on L (see Proposition 6.2) without indicating this in the notation. A subscript such as $f \ll_{\epsilon} g$ means the implied constant depends on ϵ .

The quantity x is always assumed to be large. In particular, x is large compared to any fixed constant. Thus, for example, we have $(\log x)^B \leq x^{\delta}$ for any fixed positive constants B, δ .

We use the convention that ϵ denotes an arbitrarily small (and sufficiently small) positive quantity that may vary from one occurrence to the next. We may write $x^{\epsilon + o(1)} \leq x^{\epsilon}$, for example, with no difficulties.

In order to economize on space, we often write the congruence $n \equiv v \pmod{d}$ as $n \equiv v(d)$. The notation $n \mid m^{\infty}$ means there is some positive integer N such that n divides m^N . We use the symbol \star to denote Dirichlet convolution.

We write φ for the Euler totient function, τ for the divisor function, and $P^+(n), P^-(n)$ for the largest and smallest prime factors of n , respectively.

2. INITIAL MANIPULATIONS AND OUTLINE

We begin the proof of Theorem 1.1 by setting out to estimate

$$S(x) := \sum_{n \leq x} a(n) \Lambda(n),$$

where

$$a(n) := \sum_{\substack{m^2 + \ell^2 = n \\ (\ell, \Pi) = 1}} \mathbf{1}_{\mathcal{A}'}(\ell).$$

In the definition of $a(n)$ we allow m to range over both positive and negative integers.

In this section we perform some preliminary maneuvers that reduce the proof of Theorem 1.1 to proving certain “Type I” and “Type II” estimates (see Proposition 2.1 near the end of this section).

Let $U, V > 2$ be real parameters to be chosen later (see (6.3)). For an arithmetic function $f : \mathbb{N} \rightarrow \mathbb{C}$ and real $W \geq 1$, define

$$f_{\leq W}(n) := \begin{cases} f(n), & n \leq W, \\ 0, & n > W, \end{cases}$$

and write $f_{>W} := f - f_{\leq W}$. Then Vaughan’s identity [21, Proposition 13.4] is

$$(2.1) \quad \Lambda = \Lambda_{\leq U} + \mu_{\leq V} \star \log - \Lambda_{\leq U} \star \mu_{\leq V} \star 1 + \Lambda_{>U} \star \mu_{>V} \star 1.$$

The different pieces of Vaughan’s identity decompose $S(x)$ into several sums, which we handle with different techniques. The first term $\Lambda_{\leq U}$ we treat trivially, since we may choose U to be small compared to x . The terms $\mu_{\leq V} \star \log$ and $\Lambda_{\leq U} \star \mu_{\leq V} \star 1$ are Type I sums, and require estimation of the congruence sums

$$A_d(x) := \sum_{\substack{n \leq x \\ n \equiv 0(d)}} a(n),$$

$$A'_d(x) := \sum_{\substack{n \leq x \\ n \equiv 0(d)}} a(n) \log n.$$

The last term $\Lambda_{>U} \star \mu_{>V} \star 1$ gives rise to a Type II or “bilinear” sum, and the estimation of this sum requires much more effort than estimating the Type I sums.

Let us carry out this decomposition explicitly. Inserting (2.1) into the definition of $S(x)$ gives

$$(2.2) \quad \begin{aligned} S(x) &= \sum_{n \leq x} a(n) \Lambda(n) = \sum_{n \leq U} a(n) \Lambda(n) + \sum_{n \leq x} a(n) (\mu_{\leq V} \star \log)(n) \\ &\quad - \sum_{n \leq x} a(n) (\mu_{\leq V} \star \Lambda_{\leq U} \star 1)(n) + \sum_{n \leq x} a(n) (\mu_{>V} \star \Lambda_{>U} \star 1)(n). \end{aligned}$$

By trivial estimation

$$\begin{aligned} \sum_{n \leq U} a(n) \Lambda(n) &\leq (\log U) \sum_{n \leq U} a(n) = (\log U) \sum_{\substack{m^2 + \ell^2 \leq U \\ (\ell, \Pi) = 1}} \mathbf{1}_{\mathcal{A}}(\ell) \\ &\leq (\log U) \left(\sum_{|m| \leq U^{1/2}} 1 \right) \left(\sum_{\ell \leq U^{1/2}} \mathbf{1}_{\mathcal{A}}(\ell) \right) \ll (\log U) U^{1/2 + \gamma_0/2}, \end{aligned}$$

the last inequality following by (1.1). In what follows we shall have many occasions to use the bound

$$\sum_{n \leq z} \left(\sum_{m^2 + \ell^2 = n} \mathbf{1}_{\mathcal{A}}(\ell) \right) \ll z^{1/2 + \gamma_0/2},$$

and we do so without further comment.

For the second sum in (2.2) we interchange the order of summation and separate the logarithmic factors to obtain

$$\begin{aligned} \sum_{n \leq x} a(n)(\mu_{\leq V} \star \log)(n) &= \sum_{d \leq V} \mu(d) \sum_{\substack{n \leq x \\ d|n}} a(n) \log(n/d) \\ &= \sum_{d \leq V} \mu(d) A'_d(x) - \sum_{d \leq V} \mu(d)(\log d) A_d(x). \end{aligned}$$

We similarly show that the third sum is

$$-\sum_{n \leq x} a(n)(\mu_{\leq V} \star \Lambda_{\leq U} \star 1)(n) = -\sum_{d \leq V} \sum_{m \leq U} \mu(d) \Lambda(m) A_{dm}(x).$$

For the last sum in (2.2), the Type II sum, we interchange the order of summation and change variables to obtain

$$\begin{aligned} \sum_{n \leq x} a(n)(\mu_{>V} \star \Lambda_{>U} \star 1)(n) &= \sum_{\substack{mn \leq x \\ n > V \\ m > U}} \mu(n)(\Lambda_{>U} \star 1)(m) a(mn) \\ &= \sum_{U < m \leq x/V} (\Lambda_{>U} \star 1)(m) \sum_{V < n \leq x/m} \mu(n) a(mn). \end{aligned}$$

In short,

$$S(x) = A(x; U, V) + B(x; U, V) + O((\log U)U^{1/2+\gamma_0/2}),$$

where

$$(2.3) \quad A(x; U, V) := \sum_{d \leq V} \mu(d) \left(A'_d(x) - A_d(x) \log d - \sum_{m \leq U} \Lambda(m) A_{dm}(x) \right)$$

and

$$(2.4) \quad B(x; U, V) := \sum_{U < m \leq x/V} (\Lambda_{>U} \star 1)(m) \sum_{V < n \leq x/m} \mu(n) a(mn).$$

We can exchange $A'_d(x)$ in $A(x; U, V)$ for quantities involving $A_d(t)$ using partial summation:

$$(2.5) \quad A'_d(x) = A_d(x) \log x - \int_1^x A_d(t) \frac{dt}{t}.$$

Define

$$M_d(x) := \frac{1}{d} \sum_{n \leq x} a_d(n),$$

where

$$a_d(n) := \sum_{\substack{m^2 + \ell^2 = n \\ (\ell, \Pi) = 1}} \mathbf{1}_A(\ell) \rho_\ell(d)$$

and $\rho_\ell(d)$ denotes the number of solutions ν to $\nu^2 + \ell^2 \equiv 0 \pmod{d}$. We expect that $M_d(x)$ is a good approximation to $A_d(x)$, at least on average. We therefore define the remainder term

$$(2.6) \quad R_d(x) := A_d(x) - M_d(x).$$

Inserting (2.5) into (2.3) and writing $A_d(x) = M_d(x) + R_d(x)$, we obtain

$$A(x; U, V) = M(x; U, V) + R(x; U, V),$$

where

$$(2.7) \quad M(x; U, V) := \sum_{n \leq x} \sum_{d \leq V} \frac{\mu(d)}{d} \left(a_d(n) \log(n/d) - \sum_{m \leq U} \frac{\Lambda(m)}{m} a_{dm}(n) \right)$$

and

$$(2.8) \quad R(x; U, V) := \sum_{d \leq V} \mu(d) \left(R_d(x) \log(x/d) - \int_1^x R_d(t) \frac{dt}{t} - \sum_{m \leq U} \Lambda(m) R_{md}(x) \right).$$

We have therefore proved the following result.

Proposition 2.1. *Let $U, V > 2$ be real parameters. With the notation given in this section, we have*

$$S(x) = M(x; U, V) + R(x; U, V) + B(x; U, V) + O((\log U)U^{1/2+\gamma_0/2}).$$

The outline of the rest of the paper is as follows. In Section 3 we show that $R(x; U, V)$ contributes only to the error term in Theorem 1.1. The analysis in Section 4 gives a partial analysis of $M(x; U, V)$, showing that, up to the condition $(\ell, \Pi) = 1$, the term $M(x; U, V)$ yields the main term of Theorem 1.1. We use the fundamental lemma of sieve theory to remove this condition in Section 5, and this yields the desired main term.

We estimate the bilinear form $B(x; U, V)$ in Sections 6 through 10. In Section 6 we perform some technical reductions like separating variables. These reductions allow us to enter the Gaussian domain $\mathbb{Z}[i]$ in Section 7. A congruence to a large modulus Δ arises, and this introduces further complications. We address many of these in Section 8. A particularly delicate issue is that \mathcal{A} is not well-distributed in arithmetic progressions modulo Δ when Δ shares a factor with 10. At the end of Section 9 we are mostly able to remove the congruence modulo Δ , which simplifies our work considerably. With the congruence removed we devote Section 10 to extracting cancellation from the sign changes of the Möbius function using the theory of Hecke L -functions. Theorem 1.1 follows from Propositions 6.1 and 6.2.

In the last section, Section 11, we show how to modify the proof of Theorem 1.1 to prove Theorem 1.2.

3. THE SIEVE REMAINDER TERM

In this section we show that the remainder term $R(x; U, V)$ in Proposition 2.1 is acceptably small.

Proposition 3.1. *Let $0 < \delta \leq \frac{1}{100}$ be fixed. Let $U, V > 2$ be real parameters such that*

$$(3.1) \quad UV \leq x^{\gamma_0 - \delta}.$$

Then

$$|R(x; U, V)| \ll_{\delta} x^{1/2+\gamma_0/2-\delta/5}.$$

The following lemma is the key input for Proposition 3.1.

Lemma 3.2. *For $1 \leq D \leq x$ and $\epsilon > 0$ we have*

$$R(x, D) := \sum_{d \leq D} |R_d(x)| \ll_{\epsilon} D^{1/4} x^{1/2 + \gamma_0/4 + \epsilon}.$$

Proof. This is a specialization of [13, Lemma 4]. In the notation of [13] we take $\lambda_{\ell} = \mathbf{1}_{\mathcal{A}}(\ell)$ for $\ell \leq x^{1/2}$. We then observe that

$$\|\lambda\| \leq \left(\sum_{\ell \leq x^{1/2}} \mathbf{1}_{\mathcal{A}}(\ell) \right)^{1/2} \ll x^{\gamma_0/4},$$

the last inequality following by (1.1). \square

We devote the rest of this section to the proof of Proposition 3.1. Applying the triangle inequality to (2.8), we get

$$(3.2) \quad |R(x; U, V)| \ll (\log x) R(x, UV) + \int_1^x R(t, V) \frac{dt}{t}.$$

By Lemma 3.2 and (3.1) we have

$$(\log x) R(x, UV) \ll_{\epsilon} (\log x) (x^{\gamma_0 - \delta})^{1/4} x^{1/2 + \gamma_0/4 + \epsilon} \ll x^{1/2 + \gamma_0/2 + \epsilon - \delta/4} \ll_{\delta} x^{1/2 + \gamma_0/2 - \delta/5}.$$

We can also immediately use Lemma 3.2 estimate the part of the integral in (3.2) with $t \geq V$:

$$\int_V^x R(t, V) \frac{dt}{t} \ll_{\epsilon} \int_V^x V^{1/4} t^{1/2 + \gamma_0/4 + \epsilon} \frac{dt}{t} \ll V^{1/4} x^{1/2 + \gamma_0/4 + \epsilon}.$$

This is of size $O_{\delta}(x^{1/2 + \gamma_0/2 - \delta/5})$ provided $V \leq x^{\gamma_0 - \delta}$, which already follows from (3.1) since $U > 2$. To prove Proposition 3.1 it therefore suffices to show

$$(3.3) \quad \int_1^V R(t, V) \frac{dt}{t} \ll_{\epsilon} V^{1/2 + \gamma_0/2 + \epsilon}.$$

We use (2.6) to write

$$R(t, V) = \sum_{d \leq V} |R_d(t)| \leq \sum_{d \leq V} (A_d(t) + M_d(t))$$

and estimate the sums involving A_d and M_d separately.

For the term involving A_d we use the divisor bound to obtain

$$(3.4) \quad \begin{aligned} \sum_{d \leq V} A_d(t) &\leq \sum_{d \leq V} \sum_{\substack{m^2 + \ell^2 \leq t \\ m^2 + \ell^2 \equiv 0(d)}} \mathbf{1}_{\mathcal{A}}(\ell) \leq \sum_{m^2 + \ell^2 \leq t} \mathbf{1}_{\mathcal{A}}(\ell) \tau(m^2 + \ell^2) \\ &\ll_{\epsilon} t^{\epsilon} \sum_{m^2 + \ell^2 \leq t} \mathbf{1}_{\mathcal{A}}(\ell) \ll t^{1/2 + \gamma_0/2 + \epsilon}. \end{aligned}$$

The estimation of the term involving M_d is slightly more complicated due to the presence of the function $\rho_{\ell}(d)$. Recall that $\rho_{\ell}(d)$ counts the number of residue classes $\nu \pmod{d}$ such that $\nu^2 + \ell^2 \equiv 0 \pmod{d}$. If ℓ is coprime to d , then we can divide both sides of the congruence by ℓ^2 and we find that $\rho_{\ell}(d) = \rho(d)$, where $\rho(d)$ counts the number of solutions to $\nu^2 + 1 \equiv 0 \pmod{d}$. In general, a slightly more complicated relationship holds.

Lemma 3.3. *Let ℓ, d be positive integers. Let $r(d)$ denote the largest integer r such that $r^2 \mid d$. Then*

$$\rho_\ell(d) = (r(d), \ell) \rho(d/(d, \ell^2)).$$

Proof. See [13, (3.4)]. □

Observe that Lemma 3.3 implies

$$\rho_\ell(d) \leq \rho(d) \leq \tau(d)$$

whenever d is squarefree or coprime to ℓ . If p divides ℓ , then

$$\rho_\ell(p^e) \leq 2p^{e/2}.$$

The following lemma illustrates how we estimate sums involving ρ_ℓ .

Lemma 3.4. *Let $y \geq 2$, and let ℓ be an integer. Then*

$$\sum_{n \leq y} \frac{\rho_\ell(n)}{n} \ll (\log y)^2 \prod_{p \mid \ell} \left(1 + \frac{7}{p^{1/2}}\right).$$

Proof. We factor n as $n = em$, where $e \mid \ell^\infty$ and m is coprime to ℓ . By multiplicativity and Lemma 3.3 we obtain

$$\sum_{n \leq y} \frac{\rho_\ell(n)}{n} \leq \sum_{e \mid \ell^\infty} \frac{\rho_\ell(e)}{e} \sum_{\substack{m \leq y \\ (m, \ell) = 1}} \frac{\rho_\ell(m)}{m} \leq \sum_{e \mid \ell^\infty} \frac{\rho_\ell(e)}{e} \sum_{\substack{m \leq y \\ (m, \ell) = 1}} \frac{\tau(m)}{m} \ll (\log y)^2 \sum_{e \mid \ell^\infty} \frac{\rho_\ell(e)}{e}.$$

We use multiplicativity and Lemma 3.3 again to obtain

$$\begin{aligned} \sum_{e \mid \ell^\infty} \frac{\rho_\ell(e)}{e} &= \prod_{p \mid \ell} \left(\sum_{j=0}^{\infty} \frac{\rho_\ell(p^j)}{p^j} \right) \leq \prod_{p \mid \ell} \left(1 + 2 \sum_{j=1}^{\infty} \frac{1}{p^{j/2}} \right) = \prod_{p \mid \ell} \left(1 + \frac{2}{p^{1/2} - 1} \right) \\ &\leq \prod_{p \mid \ell} \left(1 + \frac{7}{p^{1/2}} \right). \end{aligned} \quad \square$$

By the definition of $M_d(t)$ we find

$$\sum_{d \leq V} M_d(t) \leq \sum_{m^2 + \ell^2 \leq t} \sum_{d \leq V} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{d \leq V} \frac{\rho_\ell(d)}{d}.$$

We apply Lemma 3.4 and obtain

$$\sum_{d \leq V} M_d(t) \ll \sum_{m^2 + \ell^2 \leq t} \sum_{d \leq V} \mathbf{1}_{\mathcal{A}}(\ell) (\log V)^2 \tau(\ell) \ll_\epsilon t^{1/2 + \gamma_0/2} (tV)^\epsilon,$$

and combining this with our bound (3.4) yields (3.3). This completes the proof of Proposition 3.1.

4. THE SIEVE MAIN TERM

In this section we begin to show how the term $M(x; U, V)$ in Proposition 2.1 yields the main term for Theorem 1.1. This section is devoted to the proof of the following proposition.

Proposition 4.1. *Let $0 < \delta < 1$ be fixed. Let U and V satisfy $x^\delta \leq U, V \leq x$. Then*

$$(4.1) \quad M(x; U, V) = \frac{4}{\pi} \mathfrak{C} \sum_{\substack{m^2 + \ell^2 \leq x \\ (\ell, \Pi) = 1}} \mathbf{1}_{\mathcal{A}}(\ell) + O_\delta((\log x)^3 x^{1/2 + \gamma_0/2} P^{-1/2}),$$

where \mathfrak{C}, Π , and P are as in Theorem 1.1.

The elements of the proof of Proposition 4.1 are standard, but we give details for completeness.

From (2.7) we derive

$$(4.2) \quad M(x; U, V) = \sum_{\substack{g^2 + \ell^2 \leq x \\ (\ell, \Pi) = 1}} \mathbf{1}_{\mathcal{A}}(\ell) \left(\log(g^2 + \ell^2) \sum_{d \leq V} \frac{\mu(d) \rho_\ell(d)}{d} - \sum_{d \leq V} \frac{\mu(d) \rho_\ell(d) \log d}{d} \right. \\ \left. - \sum_{m \leq U} \frac{\Lambda(m)}{m} \sum_{d \leq V} \frac{\mu(d) \rho_\ell(dm)}{d} \right).$$

The main term in Proposition 4.1 arises from the second term on the right side of (4.2), and the other two terms contribute only to the error.

We begin by estimating

$$\sum_{d \leq V} \frac{\mu(d) \rho_\ell(d)}{d}$$

uniformly in ℓ . We note that

$$\rho_\ell(p) = \begin{cases} 1 + \chi(p), & p \nmid \ell, \\ 1, & p \mid \ell, \end{cases}$$

(Recall that χ is the nonprincipal character modulo 4.) The prime number theorem in arithmetic progressions then gives

$$(4.3) \quad \sum_{p \leq z} \frac{\rho_\ell(p)}{p} = \log \log z + \eta_\ell + O_\ell(\exp(-c\sqrt{\log z})),$$

for some constant η_ℓ depending on ℓ (to which we need not refer again). Since ρ_ℓ is multiplicative, (4.3) and [14, (2.4)] imply

$$(4.4) \quad \sum_{d=1}^{\infty} \frac{\mu(d) \rho_\ell(d)}{d} = 0.$$

From (4.4) and partial summation it follows that

$$(4.5) \quad \begin{aligned} \sum_{d \leq V} \frac{\mu(d)\rho_\ell(d)}{d} &= - \sum_{d > V} \frac{\mu(d)\rho_\ell(d)}{d} = \lim_{Z \rightarrow \infty} \left(- \sum_{V < d \leq Z} \frac{\mu(d)\rho_\ell(d)}{d} \right) \\ &= \lim_{Z \rightarrow \infty} \left(-Z^{-1} \sum_{d \leq Z} \mu(d)\rho_\ell(d) + V^{-1} \sum_{d \leq V} \mu(d)\rho_\ell(d) + \int_V^Z \frac{1}{t^2} \left(\sum_{d \leq t} \mu(d)\rho_\ell(d) \right) dt \right). \end{aligned}$$

We will show, uniformly in ℓ and $z \geq 1$, that

$$(4.6) \quad \sum_{d \leq z} \mu(d)\rho_\ell(d) \ll \prod_{p|\ell} \left(1 + \frac{26}{p^{2/3}} \right) z \exp(-c\sqrt{\log z}),$$

where c is some sufficiently small positive constant. For the rest of the section we shall let $c_j, j \geq 1$, denote sufficiently small positive constants, the precise values of which may vary from line to line.

The bound (4.6) is trivial if z is bounded, so we may suppose that z is large. Let $y = z \exp(-b\sqrt{\log z})$, where $b > 0$ is a parameter to be chosen later. Let g be a smooth function, supported in $[1/2, z]$, which is identically equal to one on $[y, z - y]$ and which satisfies $g^{(j)} \ll_j y^{-j}$. As $|\rho_\ell(d)| \leq \tau(d)$ for d squarefree and $\mu(d) = 0$ for d not squarefree, trivial estimation yields

$$\sum_{d \leq z} \mu(d)\rho_\ell(d) = O(y \log z) + \sum_d \mu(d)\rho_\ell(d)g(d).$$

By Mellin inversion we derive

$$(4.7) \quad \sum_d \mu(d)\rho_\ell(d)g(d) = \frac{1}{2\pi i} \int_{(2)} \widehat{g}(s) \sum_{d=1}^{\infty} \frac{\mu(d)\rho_\ell(d)}{d^s} ds.$$

From the derivative bounds on g we find that the Mellin transform $\widehat{g}(s)$ satisfies

$$\widehat{g}(s) \ll z^\sigma (1 + (y/z)^2 t^2)^{-1},$$

where $s = \sigma + it$ and $\sigma \geq \frac{2}{3}$, say.

An Euler product computation yields

$$\sum_{d=1}^{\infty} \frac{\mu(d)\rho_\ell(d)}{d^s} = \zeta(s)^{-1} L(s, \chi)^{-1} H(s) f_s(\ell),$$

where

$$H(s) := \prod_p \frac{1 - \frac{1+\chi(p)}{p^s}}{\left(1 - \frac{1}{p^s}\right) \left(1 - \frac{\chi(p)}{p^s}\right)}$$

is analytic in $\sigma \geq \frac{2}{3}$, say, and

$$f_s(\ell) := \prod_{p|\ell} \frac{1 - \frac{1}{p^s}}{1 - \frac{1+\chi(p)}{p^s}} = \prod_{p|\ell} \left(1 + \frac{\chi(p)}{p^s - 1 - \chi(p)} \right).$$

We move the line of integration in (4.7) to $\sigma = 1 + \frac{1}{\log z}$ and estimate trivially the contribution from $|t| \geq T$, with T a parameter to be chosen. In this region one has the easy bounds $|\zeta(s)|^{-1}, |L(s, \chi)|^{-1} \ll \log(2 + |t|)$, $|\widehat{g}(s)| \ll z^3(yt)^{-2}$, and $|H(s)| \ll 1$, and these give

$$\int_{|t| \geq T} \ll (\log z)^{O(1)} \frac{z^3}{y^2 T} \prod_{p|\ell} \left(1 + \frac{\chi^2(p)}{p-1-\chi^2(p)}\right).$$

For $|t| \leq T$ we move the line of integration to $\sigma = 1 - \frac{c_1}{\log T}$, where c_1 is chosen small enough that $\zeta(s)L(s, \chi)$ has no zeros in $\sigma \geq 1 - \frac{c_1}{\log T}$, $|t| \leq T$, and add in horizontal connecting lines. In this zero-free region we have the bounds $|\zeta(s)|^{-1}, |L(s, \chi)|^{-1} \ll \log(2 + |t|)$ (e.g. [25, Theorem 11.4], [27, (3.11.8)]). We estimate everything trivially to arrive at

$$\int_{|t| \leq T} \ll (\log zT)^{O(1)} \frac{z^3}{y^2} \prod_{p|\ell} \left(1 + \frac{\chi^2(p)}{p^{2/3}-1-\chi^2(p)}\right) \left(\frac{1}{T} + \exp\left(-c_1 \frac{\log z}{\log T}\right)\right).$$

We set $T = \exp(\sqrt{\log z})$, and take $b = c_1/3$. With a small amount of calculation we see that

$$\frac{\chi^2(p)}{p^{2/3}-1-\chi^2(p)} < \frac{26}{p^{2/3}}.$$

Substituting in these bounds and our choice for T , we find that the integral in (4.7) is bounded by

$$\begin{aligned} &\ll (\log z)^{O(1)} \frac{z^3}{y^2} \prod_{p|\ell} \left(1 + \frac{26}{p^{2/3}}\right) \left(\exp(-\sqrt{\log z}) + \exp(-c_1 \sqrt{\log z})\right) \\ &\ll \prod_{p|\ell} \left(1 + \frac{26}{p^{2/3}}\right) z \exp(-c_2 \sqrt{\log z}). \end{aligned}$$

This completes the proof of (4.6).

The fact that ℓ is coprime to Π implies

$$\prod_{p|\ell} \left(1 + \frac{26}{p^{2/3}}\right) \ll 1.$$

From (4.5) we see that (4.6) and $(\ell, \Pi) = 1$ yield

$$(4.8) \quad \sum_{d \leq V} \frac{\mu(d)\rho_\ell(d)}{d} \ll \exp(-c_3 \sqrt{\log V}).$$

This shows that the first term of (4.2) satisfies the bound

$$\sum_{\substack{g^2 + \ell^2 \leq x \\ (\ell, \Pi) = 1}} \mathbf{1}_{\mathcal{A}}(\ell) \log(g^2 + \ell^2) \sum_{d \leq V} \frac{\mu(d)\rho_\ell(d)}{d} \ll x^{1/2+\gamma_0/2} \exp(-c_4 \delta \sqrt{\log x}),$$

since $V \geq x^\delta$.

We turn to estimating

$$-\sum_{d \leq V} \frac{\mu(d)\rho_\ell(d) \log d}{d}.$$

We add and subtract the quantity

$$\log V \sum_{d \leq V} \frac{\mu(d)\rho_\ell(d)}{d},$$

which yields

$$-\sum_{d \leq V} \frac{\mu(d)\rho_\ell(d) \log d}{d} = \sum_{d \leq V} \frac{\mu(d)\rho_\ell(d)}{d} \log(V/d) + O(\exp(-c_5 \sqrt{\log V}))$$

by (4.8). From Perron's formula we obtain

$$(4.9) \quad \sum_{d \leq V} \frac{\mu(d)\rho_\ell(d)}{d} \log(V/d) = \frac{1}{2\pi i} \int_{(1)} \frac{x^s}{s^2} \sum_{d=1}^{\infty} \frac{\mu(d)\rho_\ell(d)}{d^{1+s}} ds.$$

An Euler product computation reveals

$$\sum_{d=1}^{\infty} \frac{\mu(d)\rho_\ell(d)}{d^{1+s}} = \zeta(1+s)^{-1} L(1+s, \chi)^{-1} H(1+s) \prod_{p|\ell} \left(1 + \frac{\chi(p)}{p^{1+s} - 1 - \chi(p)} \right).$$

We proceed in nearly identical fashion to the proof of (4.6), but here there is a main term coming from the simple pole of the integrand in (4.9) at $s = 0$. Since $L(1, \chi) = \frac{\pi}{4}$, we deduce

$$(4.10) \quad -\sum_{d \leq V} \frac{\mu(d)\rho_\ell(d) \log d}{d} = \frac{4}{\pi} \prod_{p|\ell} \left(1 + \frac{\chi(p)}{p - 1 - \chi(p)} \right) \prod_p \left(1 - \frac{\chi(p)}{(p-1)(p-\chi(p))} \right) + O(\exp(-c_6 \sqrt{\log V})).$$

The expression in (4.10) gives rise to the main term in Proposition 4.1.

The last term of (4.2) we estimate similarly to the first. We aim to show that

$$(4.11) \quad \sum_{m \leq U} \frac{\Lambda(m)}{m} \sum_{d \leq V} \frac{\mu(d)\rho_\ell(dm)}{d} \ll (\log \ell V)^3 P^{-1/2}.$$

It is convenient to distinguish two cases for d : those d that are coprime to m , and those that are not. If d is not coprime to $m = p^k$, then the presence of the Möbius function implies $d = ep$ with $(e, p) = 1$. Therefore

$$(4.12) \quad \begin{aligned} \sum_{m \leq U} \frac{\Lambda(m)}{m} \sum_{d \leq V} \frac{\mu(d)\rho_\ell(dm)}{d} &= \sum_{m \leq U} \frac{\Lambda(m)\rho_\ell(m)}{m} \sum_{\substack{d \leq V \\ (d, m)=1}} \frac{\mu(d)\rho_\ell(d)}{d} \\ &\quad - \sum_{\substack{p^k \leq U \\ p|\ell}} \frac{(\log p)\rho_\ell(p^{k+1})}{p^{k+1}} \sum_{\substack{e \leq V/p \\ (e, p)=1}} \frac{\mu(e)\rho_\ell(e)}{e}. \end{aligned}$$

It is not difficult to deal with the sum over d in the first term of (4.12) using an argument analogous to that which gave (4.8), as the condition $(d, m) = 1$ causes no great complications.

We derive

$$\sum_{\substack{d \leq V \\ (d, m)=1}} \frac{\mu(d)\rho_\ell(d)}{d} \ll \prod_{p|m} \left(1 + \frac{1}{p^{1/2}} \right) \prod_{p|\ell} \left(1 + \frac{26}{p^{2/3}} \right) \exp(-c_7 \sqrt{\log V}) \ll \exp(-c_7 \sqrt{\log V}).$$

To bound the sum over m we use Lemma 3.3, obtaining

$$\begin{aligned} \sum_{m \leq U} \frac{\Lambda(m) \rho_\ell(m)}{m} &\leq \sum_{\substack{m \leq U \\ (m, \ell)=1}} \frac{\Lambda(m) \rho_\ell(m)}{m} + (\log U) \sum_{\substack{p^k \\ p|\ell}} \frac{\rho_\ell(p^k)}{p^k} \\ &\ll \log U + (\log U) \sum_{\substack{p^k \\ p|\ell}} \frac{p^{k/2}}{p^k} \ll \log U. \end{aligned}$$

The last inequality follows since $p \mid \ell$ implies $p > P$, and $P > (\log x)^2$. Therefore

$$(4.13) \quad \sum_{m \leq U} \frac{\Lambda(m) \rho_\ell(m)}{m} \sum_{\substack{d \leq V \\ (d, m)=1}} \frac{\mu(d) \rho_\ell(d)}{d} \ll \exp(-c_8 \sqrt{\log V}).$$

We turn our attention to the second term of (4.12). We first remove those p that are not coprime to ℓ . By trivial estimation

$$(4.14) \quad \sum_{\substack{p^k \leq U \\ p|\ell}} \frac{(\log p) \rho_\ell(p^{k+1})}{p^{k+1}} \sum_{\substack{e \leq V/p \\ (e, p)=1}} \frac{\mu(e) \rho_\ell(e)}{e} \ll (\log V)^2 \sum_{p|\ell} (\log p) \sum_{k=1}^{\infty} \frac{1}{p^{k/2}} \ll (\log \ell V)^3 P^{-1/2}.$$

Here we have again used the fact that $P^-(\ell) > P$.

To handle those p that are coprime to ℓ , we first estimate trivially the contribution from $p > R = \exp(\sqrt{\log V})$. Observe that $R < U$. Then

$$(4.15) \quad \begin{aligned} \sum_{\substack{p^k \leq U \\ p > R \\ (p, \ell)=1}} \frac{(\log p) \rho_\ell(p^{k+1})}{p^{k+1}} \sum_{\substack{e \leq V/p \\ (e, p)=1}} \frac{\mu(e) \rho_\ell(e)}{e} &\ll (\log V)^2 \sum_{p > R} \log p \sum_{k=2}^{\infty} \frac{k}{p^k} \\ &\ll (\log V)^2 \sum_{p > R} \frac{\log p}{p^2} \ll \frac{(\log V)^2}{R}, \end{aligned}$$

and this is an acceptably small error. We may then show

$$(4.16) \quad \sum_{\substack{p^k \leq U \\ p \leq R \\ (p, \ell)=1}} \frac{(\log p) \rho_\ell(p^{k+1})}{p^{k+1}} \sum_{\substack{e \leq V/p \\ (e, p)=1}} \frac{\mu(e) \rho_\ell(e)}{e} \ll \exp(-c_8 \sqrt{\log V})$$

by arguing as before, since V/p is close to V in the logarithmic scale. Taking (4.13), (4.14), (4.15), and (4.16) together gives (4.11). We combine (4.8), (4.10), and (4.13) to derive

$$M(x; U, V) = \frac{4}{\pi} \mathfrak{C} \sum_{\substack{m^2 + \ell^2 \leq x \\ (\ell, \Pi)=1}} \sum_{p|\ell} \mathbf{1}_{\mathcal{A}}(\ell) \prod \left(1 + \frac{\chi(p)}{p-1-\chi(p)} \right) + O((\log x)^3 x^{1/2+\gamma_0/2} P^{-1/2}).$$

Here

$$\mathfrak{C} = \prod_p \left(1 - \frac{\chi(p)}{(p-1)(p-\chi(p))} \right)$$

is the constant in Theorem 1.1. Since $P^-(\ell) > P$ we have

$$\prod_{p|\ell} \left(1 + \frac{\chi(p)}{p-1-\chi(p)}\right) = 1 + O\left(\frac{\log \ell}{P}\right),$$

and this yields Proposition 4.1.

5. THE SIEVE MAIN TERM: FUNDAMENTAL LEMMA

We wish to simplify the main term of Proposition 4.1 by removing the condition $(\ell, \Pi) = 1$, which we accomplish with the fundamental lemma of sieve theory. In this section we prove the following proposition.

Proposition 5.1. *Assume the notation of Theorem 1.1. Then*

$$\sum_{\substack{m^2+\ell^2 \leq x \\ (\ell, \Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) = \kappa_1 \frac{e^{-\gamma}}{\log P} \sum_{m^2+\ell^2 \leq x} \mathbf{1}_{\mathcal{A}}(\ell) + O\left(x^{1/2+\gamma_0/2} \exp\left(-c\sqrt{\log P}\right)\right),$$

where $c > 0$ is some absolute constant.

In order to apply the sieve we require information about the elements of \mathcal{A} in arithmetic progressions. We invariably detect congruence conditions on elements of \mathcal{A} via additive characters, so we require information on exponential sums over \mathcal{A} . It is convenient to normalize these exponential sums so that we may study them at different scales. For Y an integral power of 10, we define

$$(5.1) \quad F_Y(\theta) := Y^{-\log 9 / \log 10} \left| \sum_{0 \leq n < Y} \mathbf{1}_{\mathcal{A}}(n) e(n\theta) \right|,$$

so $F_Y(\theta) \ll 1$ for all Y and real numbers θ . Observe that F_Y is a periodic function with period 1. We emphasize that Y is always a power of 10 when it appears in a subscript like this.

Let U and V be two integral powers of ten (here U and V have nothing to do with the U and V from Vaughan's identity (2.1)). From the definition (5.1) it is not difficult to derive (see [24, (10.3)]) the identity

$$(5.2) \quad F_{UV}(\theta) = F_U(\theta) F_V(U\theta).$$

We take the opportunity to collect in one place the lemmas we need to estimate F_Y and various averages of F_Y .

The first result is a sort of Siegel-Walfisz result for F_Y .

Lemma 5.2. *Let $q < Y^{1/3}$ be of the form $q = q_1 q_2$ with $(q_1, 10) = 1$ and $q_1 > 1$. Then for any integer a coprime to q we have*

$$F_Y\left(\frac{a}{q}\right) \ll \exp\left(-c_0 \frac{\log Y}{\log q}\right)$$

for some absolute constant $c_0 > 0$.

Proof. This is a slight weakening of [24, Lemma 10.1]. □

The next two lemmas are results of large sieve type for F_Y .

Lemma 5.3. *For $q \geq 1$ we have*

$$\sup_{\beta \in \mathbb{R}} \sum_{a \leq q} F_X \left(\frac{a}{q} + \beta \right) \ll q^{27/77} + \frac{q}{X^{50/77}}.$$

Proof. This is a slight weakening of the first part of [24, Lemma 10.5]. \square

Lemma 5.4. *For $Q \geq 1$ we have*

$$\sup_{\beta \in \mathbb{R}} \sum_{q \leq Q} \sum_{\substack{1 \leq a \leq q \\ (a,q)=1}} F_Y \left(\frac{a}{q} + \beta \right) \ll Q^{54/77} + \frac{Q^2}{Y^{50/77}}.$$

Proof. This is a slight weakening of the second part of [24, Lemma 10.5]. \square

Now that the necessary results are in place, we proceed with the proof of Proposition 5.1. We write

$$\begin{aligned} \sum_{\substack{m^2 + \ell^2 \leq x \\ (\ell, \Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) &= \sum_{|m| \leq x^{1/2}} \sum_{\substack{\ell \leq \sqrt{x-m^2} \\ (\ell, \Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \\ &= \sum_{|m| \leq \sqrt{1-P^{-2}}x^{1/2}} \sum_{\substack{\ell \leq \sqrt{x-m^2} \\ (\ell, \Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) + O(x^{1/2+\gamma_0/2}P^{-1}), \end{aligned}$$

the second equality following by trivial estimation.

With the restriction $|m| \leq \sqrt{1-P^{-2}}x^{1/2}$ on m we estimate each sum over ℓ individually. Set $z = z(m) = \sqrt{x-m^2}$. We apply upper- and lower-bound linear sieves of level

$$D = z^{1/5}$$

(see [16, Chapter 5] for terminology). Therefore

$$\sum_{\substack{d \leq D \\ d \mid \Pi \\ (d,10)=1}} \lambda_d^- \sum_{\substack{\ell \leq z \\ \ell \equiv 0(d) \\ (\ell,10)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \leq \sum_{\substack{\ell \leq z \\ (\ell, \Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \leq \sum_{\substack{d \leq D \\ d \mid \Pi \\ (d,10)=1}} \lambda_d^+ \sum_{\substack{\ell \leq z \\ \ell \equiv 0(d) \\ (\ell,10)=1}} \mathbf{1}_{\mathcal{A}}(\ell).$$

The sequences $(\lambda_d^\pm)_{d \leq D}$ are real, and the terms satisfy $|\lambda_d^\pm| \leq 1$. The upper and lower bounds turn out to be asymptotically equal, and we simply write λ_d for λ_d^+ or λ_d^- .

It is difficult to work with elements of \mathcal{A} over intervals whose lengths are not a power of 10. We put ourselves in this situation with a short interval decomposition (a similar technique is applied in [3]). Let Y be the largest power of 10 that satisfies $Y \leq zP^{-1}$. We break the summation over ℓ into intervals of the form $[nY, (n+1)Y)$, where n is a nonnegative integer. This gives

$$(5.3) \quad \sum_{\substack{\ell \leq z \\ \ell \equiv 0(d) \\ (\ell,10)=1}} \mathbf{1}_{\mathcal{A}}(\ell) = \sum_{n \in S(z)} \sum_{\substack{nY \leq \ell < (n+1)Y \\ \ell \equiv 0(d) \\ (\ell,10)=1}} \mathbf{1}_{\mathcal{A}}(\ell) + O\left(\sum_{\substack{z-Y \leq \ell < z+Y \\ \ell \equiv 0(d)}} \mathbf{1}_{\mathcal{A}}(\ell) \right).$$

Here $S(z)$ is some set of size $\ll P$.

Remark. We will repeatedly see this technique of breaking an interval into shorter subintervals, with each subinterval having length a power of 10, in the later estimation of the bilinear sum $B(x; U, V)$.

We first illustrate how to use Lemma 5.4 to handle the error term in (5.3). On summing over d , we must estimate

$$\mathcal{E} := \sum_{d \leq D} \sum_{\substack{z-Y \leq \ell \leq z+Y \\ \ell \equiv 0(d)}} \mathbf{1}_{\mathcal{A}}(\ell).$$

Because the estimation of \mathcal{E} introduces a number of important ideas that we use throughout the proof of Theorem 1.1, we encapsulate the estimation in a lemma.

Lemma 5.5. *With the notation as above,*

$$\mathcal{E} \ll (\log D)^2 Y^{\gamma_0}.$$

Proof. For X some power of 10 with $Y \leq X \ll Y$ and some integer k depending only on z, Y , and X , we have

$$\mathcal{E} \ll \sum_{d \leq D} \sum_{\substack{kX \leq \ell < (k+1)X \\ \ell \equiv 0(d)}} \mathbf{1}_{\mathcal{A}}(\ell).$$

If $\mathbf{1}_{\mathcal{A}}(k) = 0$ then the sum over ℓ is empty and $\mathcal{E} = 0$. Suppose then that $\mathbf{1}_{\mathcal{A}}(k) = 1$. We write $\ell = kX + t$, where $0 \leq t < X$. There are now two subcases to consider, depending on whether the missing a_0 is equal to 0 or not. If $a_0 \neq 0$ then $\mathbf{1}_{\mathcal{A}}(kX + t) = \mathbf{1}_{\mathcal{A}}(t)$ for $0 \leq t < X$. If $a_0 = 0$ then $\mathbf{1}_{\mathcal{A}}(kX + t) = 0$ for $0 \leq t < X/10$ and $\mathbf{1}_{\mathcal{A}}(kX + t) = \mathbf{1}_{\mathcal{A}}(t)$ for $X/10 \leq t < X$. We can unite the two subcases by writing

$$\mathcal{E} \leq \sum_{d \leq D} \sum_{\substack{\delta(a_0)X/10 \leq t < X \\ t \equiv -kX(d)}} \mathbf{1}_{\mathcal{A}}(t),$$

where

$$\delta(n) = \begin{cases} 1, & n = 0, \\ 0, & n \neq 0. \end{cases}$$

By inclusion-exclusion and the triangle inequality we find

$$\mathcal{E} \ll \sum_{d \leq D} \sum_{\substack{t < X \\ t \equiv -kX(d)}} \mathbf{1}_{\mathcal{A}}(t).$$

We detect the congruence via the orthogonality of additive characters, which yields

$$\mathcal{E} \ll \sum_{d \leq D} \frac{1}{d} \sum_{r=1}^d e\left(\frac{rkX}{d}\right) \sum_{t < X} \mathbf{1}_{\mathcal{A}}(t) e\left(\frac{rt}{d}\right).$$

By the triangle inequality,

$$\mathcal{E} \ll X^{\gamma_0} \sum_{d \leq D} \frac{1}{d} \sum_{r=1}^d F_X\left(\frac{r}{d}\right).$$

We remove the terms with $r = d$ (the “zero” frequency), which gives

$$\mathcal{E} \ll (\log D)X^{\gamma_0} + X^{\gamma_0} \sum_{1 < d \leq D} \frac{1}{d} \sum_{r=1}^{d-1} F_X\left(\frac{r}{d}\right).$$

For the “non-zero” frequencies we reduce to primitive fractions and obtain

$$\sum_{1 < d \leq D} \frac{1}{d} \sum_{r=1}^{d-1} F_X \left(\frac{r}{d} \right) = \sum_{1 < d \leq D} \frac{1}{d} \sum_{\substack{q|d \\ q > 1}} \sum_{\substack{1 \leq b \leq q \\ (b,q)=1}} F_X \left(\frac{b}{q} \right) \ll (\log D) \sum_{1 < q \leq D} \frac{1}{q} \sum_{\substack{1 \leq b \leq q \\ (b,q)=1}} F_X \left(\frac{b}{q} \right).$$

We perform a dyadic decomposition on the range of q to get

$$\mathcal{E} \ll (\log D)^2 X^{\gamma_0} \sup_{Q \leq D} \frac{1}{Q} \sum_{q \leq Q} \sum_{\substack{1 \leq b \leq q \\ (b,q)=1}} F_X \left(\frac{b}{q} \right).$$

By Lemma 5.4,

$$\begin{aligned} \mathcal{E} &\ll (\log D)^2 X^{\gamma_0} \sup_{Q \leq D} \left(\frac{1}{Q^{23/77}} + \frac{Q}{X^{50/77}} \right) \ll (\log D)^2 X^{\gamma_0} \left(1 + \frac{D}{X^{50/77}} \right) \\ &\ll (\log D)^2 X^{\gamma_0}, \end{aligned}$$

and this completes the proof. \square

From (5.3) and Lemma 5.5 it follows that

$$\sum_{\substack{d \leq D \\ d|10 \\ (d,10)=1}} \lambda_d \sum_{\substack{\ell \leq z \\ \ell \equiv 0(d) \\ (\ell,10)=1}} \mathbf{1}_{\mathcal{A}}(\ell) = \sum_{n \in S(z)} \mathbf{1}_{\mathcal{A}}(n) \sum_{\substack{d \leq D \\ d|10 \\ (d,10)=1}} \lambda_d \sum_{\substack{\delta(a_0)Y/10 \leq t < Y \\ t \equiv -nY(d) \\ (t,10)=1}} \mathbf{1}_{\mathcal{A}}(t) + O(x^{\gamma_0/2} P^{-1/2}).$$

We detect the congruence with additive characters and obtain

$$\sum_{\substack{\delta(a_0)Y/10 \leq t < Y \\ t \equiv -nY(d) \\ (t,10)=1}} \mathbf{1}_{\mathcal{A}}(t) = \frac{1}{d} \sum_{r=1}^d e \left(\frac{rnY}{d} \right) \sum_{\substack{\delta(a_0)Y/10 \leq t < Y \\ (t,10)=1}} \mathbf{1}_{\mathcal{A}}(t) e \left(\frac{rt}{d} \right).$$

Naturally we extract the main term from $r = d$.

Define

$$\kappa := \begin{cases} \frac{\varphi(10)}{9}, & (a_0, 10) \neq 1, \\ \frac{\varphi(10)-1}{9}, & (a_0, 10) = 1. \end{cases}$$

It is easy to check that

$$\sum_{\substack{t < 10^k \\ (t,10)=1}} \mathbf{1}_{\mathcal{A}}(t) = \kappa \sum_{t < 10^k} \mathbf{1}_{\mathcal{A}}(t),$$

which implies

$$\sum_{\substack{\delta(a_0)Y/10 \leq t < Y \\ (t,10)=1}} \mathbf{1}_{\mathcal{A}}(t) = \kappa \sum_{\delta(a_0)Y/10 \leq t < Y} \mathbf{1}_{\mathcal{A}}(t).$$

We can then reverse our short interval decomposition.

For the “non-zero frequencies” $1 \leq r \leq d-1$, we observe that only those $n \in S(z)$ with $\mathbf{1}_{\mathcal{A}}(n) \neq 0$ contribute. There are $\ll P^{\gamma_0}$ such n . We then use inclusion-exclusion to replace

the condition $\delta(a_0)Y/10 \leq t < Y$ by $t < X$, for $X \asymp Y$ a power of 10. We then use the triangle inequality to find that the contribution from the non-zero frequencies is

$$\ll P^{\gamma_0} \sum_{\substack{d \leq D \\ (d,10)=1}} \frac{1}{d} \sum_{r=1}^{d-1} \left| \sum_{\substack{t < X \\ (\ell,10)=1}} \mathbf{1}_{\mathcal{A}}(t) e\left(\frac{rt}{d}\right) \right|.$$

We handle the condition $(t, 10) = 1$ with Möbius inversion, and detect the resulting congruences with additive characters. Therefore

$$\sum_{\substack{t < X \\ (\ell,10)=1}} \mathbf{1}_{\mathcal{A}}(t) e\left(\frac{rt}{d}\right) = \sum_{b|10} \frac{\mu(b)}{b} \sum_{s=1}^b \sum_{t < X} \mathbf{1}_{\mathcal{A}}(t) e\left(\frac{rt}{d} + \frac{st}{b}\right),$$

and putting the contribution from $r = d$ and $1 \leq r \leq d - 1$ together it follows that

$$(5.4) \quad \sum_{\substack{d \leq D \\ d|10}} \lambda_d \sum_{\substack{\ell \leq z \\ \ell \equiv 0(d) \\ (\ell,10)=1}} \mathbf{1}_{\mathcal{A}}(\ell) = \kappa \sum_{\substack{d \leq D \\ d|10}} \frac{\lambda_d}{d} \sum_{\ell \leq z} \mathbf{1}_{\mathcal{A}}(\ell) + O\left(x^{\gamma_0/2} (P^{-1/2} + EP^{\gamma_0})\right),$$

where

$$E := \sum_{\substack{1 < d \leq D \\ (d,10)=1}} \frac{1}{d} \sum_{e=1}^{10} \sum_{r=1}^{d-1} F_X\left(\frac{r}{d} + \frac{e}{10}\right).$$

Similarly to the estimation of \mathcal{E} in Lemma 5.5 above, we put, for pedagogical reasons, the estimation of E into a lemma.

Lemma 5.6. *With the notation given above,*

$$E \ll \exp(-c\sqrt{\log z})$$

for some absolute constant $c > 0$.

Proof. We reduce to primitive fractions to derive

$$E = \sum_{\substack{1 < d \leq D \\ (d,10)=1}} \frac{1}{d} \sum_{e=1}^{10} \sum_{\substack{q|d \\ q>1}} \sum_{\substack{a=1 \\ (a,q)=1}}^q F_X\left(\frac{a}{q} + \frac{e}{10}\right).$$

We apply (5.2) with $U = 10, V = X/10$ to obtain

$$\begin{aligned} F_X\left(\frac{a}{q} + \frac{e}{10}\right) &= F_{10}\left(\frac{a}{q} + \frac{e}{10}\right) F_V\left(\frac{10a}{q} + \frac{10e}{10}\right) \\ &= F_{10}\left(\frac{a}{q} + \frac{e}{10}\right) F_V\left(\frac{10a}{q}\right) \ll F_V\left(\frac{10a}{q}\right). \end{aligned}$$

Since $(10, q) = 1$, we may change variables $10a \rightarrow a$ to obtain

$$E \ll \sum_{\substack{1 < d \leq D \\ (d,10)=1}} \frac{1}{d} \sum_{\substack{q|d \\ q>1}} \sum_{\substack{a=1 \\ (a,q)=1}}^q F_V\left(\frac{a}{q}\right) \ll (\log D) \sum_{\substack{1 < q \leq D \\ (q,10)=1}} \frac{1}{q} \sum_{\substack{a=1 \\ (a,q)=1}}^q F_V\left(\frac{a}{q}\right).$$

We perform a dyadic decomposition on the range of q to obtain

$$E \ll (\log D)^2 \sup_{Q \leq D} \frac{1}{Q} \sum_{\substack{1 < q \leq Q \\ (q, 10)=1}} \sum_{\substack{a=1 \\ (a, q)=1}}^q F_V \left(\frac{a}{q} \right).$$

Set $Q_1 := \exp(\varepsilon \sqrt{\log z})$, where $\varepsilon > 0$ is a small positive constant. If $Q > Q_1$ we use Lemma 5.4, and if $Q \leq Q_1$ we use Lemma 5.2. It is helpful here to recall that $D = z^{1/5} \leq x^{1/10}$. Provided ε in the definition of Q_1 is taken sufficiently small in terms of c_0 in Lemma 5.2, we obtain

$$E \ll \exp(-c\sqrt{\log z}),$$

where $c > 0$ is some absolute constant. □

We take (5.4) with Lemma 5.6, along with the fact that $\log P = o(\sqrt{\log z})$, to get

$$\sum_{\substack{d \leq D \\ d|\Pi \\ (d, 10)=1}} \lambda_d \sum_{\substack{\ell \leq z \\ \ell \equiv 0(d) \\ (\ell, 10)=1}} \mathbf{1}_{\mathcal{A}}(\ell) = \kappa \sum_{\substack{d \leq D \\ d|\Pi \\ (d, 10)=1}} \frac{\lambda_d}{d} \sum_{\ell \leq z} \mathbf{1}_{\mathcal{A}}(\ell) + O(x^{\gamma_0/2} P^{-1/2}).$$

By the fundamental lemma of the linear sieve (see [16, Lemma 6.11])

$$(5.5) \quad \sum_{\substack{d \leq D \\ d|\Pi \\ (d, 10)=1}} \frac{\lambda_d}{d} = (1 + O(\exp(-\tfrac{1}{2}s \log s))) \prod_{\substack{p \leq P \\ p \nmid 10}} \left(1 - \frac{1}{p}\right),$$

where

$$s = \frac{\log D}{\log P} \gg (\log \log x) \sqrt{\log x}.$$

It follows that

$$\sum_{\substack{d \leq D \\ d|\Pi \\ (d, 10)=1}} \lambda_d \sum_{\substack{\ell \leq z \\ \ell \equiv 0(d) \\ (\ell, 10)=1}} \mathbf{1}_{\mathcal{A}}(\ell) = \frac{10}{\varphi(10)} \kappa \prod_{p \leq P} \left(1 - \frac{1}{p}\right) \sum_{\ell \leq z} \mathbf{1}_{\mathcal{A}}(\ell) + O(x^{\gamma_0/2} P^{-1/2}).$$

We use Mertens' theorem with prime number theorem error term to get

$$\prod_{p \leq P} \left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}}{\log P} \left(1 + O\left(\exp\left(-c\sqrt{\log P}\right)\right)\right),$$

for some constant $c > 0$. This completes the proof of Proposition 5.1.

6. BILINEAR FORM IN THE SIEVE: FIRST STEPS

By combining the work of previous sections, we may prove the following proposition. Recall from (2.4) that

$$B(x; U, V) = \sum_{U < m \leq x/V} (\Lambda_{>U} \star 1)(m) \sum_{V < n \leq x/m} \mu(n) a(mn).$$

Proposition 6.1. *Let $0 < \delta \leq \frac{1}{100}$ be fixed. Let U and V be real parameters satisfying*

$$UV \leq x^{\gamma_0 - \delta}, \quad U, V \geq x^\delta.$$

Then

$$(6.1) \quad S(x) = \frac{4\mathfrak{C}\kappa_1}{\pi} \frac{e^{-\gamma}}{\log P} \sum_{m^2 + \ell^2 \leq x} \mathbf{1}_{\mathcal{A}}(\ell) + B(x; U, V) + O_\delta \left(x^{1/2 + \gamma_0/2} \exp \left(-c\sqrt{\log P} \right) \right).$$

Proof. Proposition 2.1 states

$$S(x) = M(x; U, V) + R(x; U, V) + B(x; U, V) + O((\log U)U^{1/2 + \gamma_0/2}).$$

The error term is sufficiently small since $U \leq x^{\gamma_0}$. By Proposition 3.1 we have

$$|R(x; U, V)| \ll_\delta x^{1/2 + \gamma_0/2 - \delta/5}.$$

Proposition 4.1 gives

$$M(x; U, V) = \frac{4}{\pi} \mathfrak{C} \sum_{\substack{m^2 + \ell^2 \leq x \\ (\ell, \Pi) = 1}} \mathbf{1}_{\mathcal{A}}(\ell) + O_\delta((\log x)^3 x^{1/2 + \gamma_0/2} P^{-1/2}),$$

and by Proposition 5.1

$$\sum_{\substack{m^2 + \ell^2 \leq x \\ (\ell, \Pi) = 1}} \mathbf{1}_{\mathcal{A}}(\ell) = \kappa_1 \frac{e^{-\gamma}}{\log P} \sum_{m^2 + \ell^2 \leq x} \mathbf{1}_{\mathcal{A}}(\ell) + O \left(x^{1/2 + \gamma_0/2} \exp \left(-c\sqrt{\log P} \right) \right). \quad \square$$

The task now is to show that the Type II sum $B(x; U, V)$ contributes only to the error term of (6.1). (We note that the implied constant in the error term of (6.1) is effectively computable.) We shall prove the following proposition.

Proposition 6.2. *Let $L > 0$ be large but fixed. Assume that U and V satisfy*

$$(6.2) \quad \begin{aligned} U &= x^\alpha, & V &= x^\beta, \\ \frac{52}{77} &= 0.675\dots < \alpha < \gamma_0 - \left(\frac{1}{2} - \frac{\gamma_0}{2} \right) = 0.931\dots, \\ \frac{1}{2} - \frac{\gamma_0}{2} &= 0.0228\dots < \beta < \gamma_0 - \alpha, \end{aligned}$$

with α and β being fixed real numbers. Then there exist absolute constants $C, \Omega > 0$ (i.e. independent of L) such that

$$B(x; U, V) \ll_L (\log x)^{-\Omega L + C} x^{1/2 + \gamma_0/2}.$$

The implied constant is ineffective.

Proof of Theorem 1.1 assuming Proposition 6.2. We choose U and V in Vaughan's identity (2.1) to be

$$(6.3) \quad U = x^{7/10}, \quad V = x^{1/5}.$$

Then U and V satisfy the hypotheses of Propositions 6.1 and 6.2. Theorem 1.1 follows upon taking $L = \Omega^{-1}(A + C)$. \square

To prove Theorem 1.1 it therefore suffices to prove Proposition 6.2, and this task commands our attention through Section 10. In what follows we allow implied constants to depend on L without indicating it in the notation.

We mention that the arguments in this section have some similarity to those in [14, Section 10] and [15, Section 4].

Since $(\Lambda_{>U} \star 1)(m) \leq (\Lambda \star 1)(m) = \log m$, we have the upper bound

$$(6.4) \quad B(x; U, V) \leq (\log x) \sum_{U < m \leq x/V} \left| \sum_{V < n \leq x/m} \mu(n) a(mn) \right|.$$

It is easy to obtain a trivial bound for $B(x; U, V)$ that is not too far from the bound of Proposition 6.2.

Lemma 6.3. *Let $x \geq 2$. Then*

$$B(x; U, V) \ll (\log x)^3 x^{1/2+\gamma_0/2}.$$

Proof. We change variables in (6.4) and deduce

$$B(x; U, V) \ll (\log x) \sum_{k \leq x} a(k) \tau(k) \ll (\log x) \sum_{d \leq x^{1/2}} \sum_{\substack{k \leq x \\ d|k}} a(k) = (\log x) \sum_{d \leq x^{1/2}} A_d(x).$$

We write $A_d(x) = M_d(x) + R_d(x)$ and use Lemma 3.2 to bound the sum of $R_d(x)$, giving

$$B(x; U, V) \ll (\log x) \sum_{\substack{m^2 + \ell^2 \leq x \\ (\ell, \Pi) = 1}} \sum_{\ell} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{d \leq x^{1/2}} \frac{\rho_\ell(d)}{d}.$$

By Lemma 3.4 we have

$$B(x; U, V) \ll (\log x)^3 \sum_{\substack{m^2 + \ell^2 \leq x \\ (\ell, \Pi) = 1}} \sum_{\ell} \mathbf{1}_{\mathcal{A}}(\ell) \prod_{p|\ell} \left(1 + \frac{7}{p^{1/2}} \right) \ll (\log x)^3 x^{1/2+\gamma_0/2},$$

the last inequality following since $P^-(\ell) > P$. \square

In the proof of Lemma 6.3 we used Lemma 3.4 to control averages of ρ_ℓ . We shall need more elaborate versions of this argument in several of our reductions of $B(x; U, V)$.

We wish to reduce the estimation of $B(x; U, V)$ to a bilinear sum in which the variables m and n are separated from one another, and in which n has no prime factors $\leq P$; equivalently, $(n, \Pi) = 1$. We will see that the condition $(n, \Pi) = 1$ confers substantial technical advantages.

We spend the rest of this section reducing Proposition 6.2 to the following proposition.

Proposition 6.4. *Let $0 < \delta \leq \frac{1}{100}$ be fixed, and for fixed $L > 0$ set $\theta = (\log x)^{-L}$. Then for real numbers M, N with*

$$(6.5) \quad x^{1/2-\gamma_0/2+\delta} \leq N \leq x^{25/77-\delta}, \quad \theta x \ll MN \ll x$$

we have

$$B_4(M, N) := \sum_{M < m \leq 2M} \left| \sum_{\substack{N < n \leq (1+\theta)N \\ (n, m\Pi) = 1}} \mu(n) a(mn) \right| \ll_\delta \theta^9 (\log MN)^{O(1)} (MN)^{1/2+\gamma_0/2}.$$

The implied constant is ineffective.

We first focus on the task of separating the variables m and n . We break the summation over n into subintervals $N < n \leq (1 + \theta)N$. By the triangle inequality,

$$B(x; U, V) \leq (\log x) \sum_{\substack{N=(1+\theta)^j V \\ j \geq 0 \\ N \leq x/U}} \sum_{U < m \leq x/V} \left| \sum_{\substack{N < n \leq (1+\theta)N \\ mn \leq x}} \mu(n) a(mn) \right|.$$

We wish to replace the condition $mn \leq x$ by $mN \leq x$. Clearly $mn \leq x$ implies $mN \leq x$ since $n > N$. Thus, suppose $mN \leq x$ but $mn > x$. Then

$$x < mn \leq (1 + \theta)mN \leq (1 + \theta)x,$$

and so for fixed N we have

$$\sum_{U < m \leq x/V} \left| \sum_{\substack{N < n \leq (1+\theta)N \\ mN \leq x \\ mn > x}} \mu(n) a(mn) \right| \leq \sum_{N < n \leq (1+\theta)N} \mu^2(n) \sum_{\substack{x < k \leq (1+\theta)x \\ n|k}} a(k).$$

Since by (6.2) we have $N \leq x/U \leq x^{1/2}$, say, we obtain by Lemma 3.2

$$\begin{aligned} & \sum_{N < n \leq (1+\theta)N} \mu^2(n) \sum_{\substack{x < k \leq (1+\theta)x \\ n|k}} a(k) \\ & \ll \sum_{x < m^2 + \ell^2 \leq (1+\theta)x} \sum_{N < n \leq (1+\theta)N} \mathbf{1}_{\mathcal{A}}(\ell) \frac{\mu^2(n) \rho_{\ell}(n)}{n} + R(x, N) + R((1 + \theta)x, N) \\ & \ll \sum_{x < m^2 + \ell^2 \leq (1+\theta)x} \sum_{N < n \leq (1+\theta)N} \mathbf{1}_{\mathcal{A}}(\ell) \frac{\tau(n)}{n} + x^{1/2 + \gamma_0/2 - 1/1000} \\ & \ll \theta(\log N) \sum_{x < m^2 + \ell^2 \leq (1+\theta)x} \mathbf{1}_{\mathcal{A}}(\ell) + x^{1/2 + \gamma_0/2 - 1/1000}. \end{aligned}$$

In making these estimates we have used the fact that $\mu^2(n) \rho_{\ell}(n) \leq \tau(n)$, which is a consequence of Lemma 3.3. We have

$$\begin{aligned} \sum_{x < m^2 + \ell^2 \leq (1+\theta)x} \mathbf{1}_{\mathcal{A}}(\ell) &= \sum_{\ell \leq \sqrt{(1+\theta)x}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{\sqrt{x - \ell^2} < |m| \leq \sqrt{(1+\theta)x - \ell^2}} 1 \\ &\ll \theta^{1/2} x^{1/2} \sum_{\ell \ll x^{1/2}} \mathbf{1}_{\mathcal{A}}(\ell) \ll \theta^{1/2} x^{1/2 + \gamma_0/2}. \end{aligned}$$

We have therefore shown

$$\sum_{U < m \leq x/V} \left| \sum_{\substack{N < n \leq (1+\theta)N \\ mN \leq x \\ mn > x}} \mu(n) a(mn) \right| \ll (\log x) \theta^{3/2} x^{1/2 + \gamma_0/2}.$$

Since the number of intervals $N < n \leq (1 + \theta)N$ is $\ll (\log x)\theta^{-1}$ we see

$$(6.6) \quad \begin{aligned} B(x; U, V) &\ll (\log x) \sum_{\substack{N=(1+\theta)^j V \\ j \geq 0 \\ N \leq x/U}} \sum_{U < m \leq x/N} \left| \sum_{N < n \leq (1+\theta)N} \mu(n)a(mn) \right| \\ &\quad + (\log x)^3 \theta^{1/2} x^{1/2+\gamma_0/2}, \end{aligned}$$

and the error term here is acceptable for Proposition 6.2.

We now fix one such N with $V \leq N \leq x/U$, and perform a dyadic decomposition on the range of m , which yields

$$\begin{aligned} \sum_{U < m \leq x/N} \left| \sum_{N < n \leq (1+\theta)N} \mu(n)a(mn) \right| &\leq \sum_{\substack{M=2^j U \\ j \geq 0 \\ MN \leq x}} \sum_{M < m \leq 2M} \left| \sum_{N < n \leq (1+\theta)N} \mu(n)a(mn) \right| \\ &= \sum_{\substack{M=2^j U \\ j \geq 0 \\ MN \leq x}} B_1(M, N), \end{aligned}$$

where

$$B_1(M, N) := \sum_{M < m \leq 2M} \left| \sum_{N < n \leq (1+\theta)N} \mu(n)a(mn) \right|.$$

Observe that if $MN \ll \theta x$ then

$$(6.7) \quad B_1(M, N) \ll \sum_{N < n \leq (1+\theta)N} \mu^2(n) \sum_{\substack{k \leq \theta x \\ n|k}} a(k) \ll (\log N) \theta^{1+\gamma_0} x^{1/2+\gamma_0/2},$$

the latter inequality following essentially by the argument that gave (6.6).

If U and V satisfy (6.2), then N satisfies (6.5) for some sufficiently small $\delta > 0$ (which depends on α and β in (6.2)) since $V \ll N \ll x/U$. In order to prove Proposition 6.2 it therefore follows from (6.6) and (6.7) that we need to show

$$(6.8) \quad B_1(M, N) \ll \theta^{5/2} (\log MN)^{O(1)} (MN)^{1/2+\gamma_0/2}$$

for M and N satisfying (6.5).

Note that the bilinear form $B_4(M, N)$ in Proposition 6.4 is $B_1(M, N)$ with the additional condition $(n, m\Pi) = 1$ imposed. We write the variable n in $B_1(M, N)$ as $n = n_0 n_1$, where $(n_0, \Pi) = 1$ and $n_1 \mid \Pi$. We then set

$$\mathcal{T} := \exp((\log P)^2).$$

Observe that $\mathcal{T} > P$ and $\mathcal{T} \ll x^{o(1)}$ by our upper bound for $\log P$ in Theorem 1.1.

We first show that the contribution from $n_1 > \mathcal{T}$ to $B_1(M, N)$ is negligible. If $n_1 \mid \Pi$ and $n_1 > \mathcal{T}$, then there is a divisor d of n_1 that satisfies $\mathcal{T} < d \leq P\mathcal{T}$. Indeed, writing $n_1 = p_1 \cdots p_r$ where $p_1 < \cdots < p_r$, we see there is a minimal j such that $p_1 \cdots p_j \leq \mathcal{T}$ but $p_1 \cdots p_{j+1} > \mathcal{T}$. The desired divisor is $d = p_1 \cdots p_{j+1}$. The contribution to $B_1(M, N)$ from

$n_1 > \mathcal{T}$ is

$$\begin{aligned} \sum_{M < m \leq 2M} \left| \sum_{\substack{N < n_0 n_1 \leq (1+\theta)N \\ (n_0, \Pi)=1 \\ n_1 | \Pi \\ n_1 > \mathcal{T}}} \mu(n_0) \mu(n_1) a(m n_0 n_1) \right| &\leq \sum_{\substack{\mathcal{T} < d \leq P\mathcal{T} \\ d | \Pi}} \sum_{\substack{n \ll MN \\ d | n}} a(n) \tau_3(n) \\ &\leq \sum_{\substack{\mathcal{T} < d \leq P\mathcal{T} \\ d | \Pi}} \sum_{\substack{n \ll MN \\ d | n}} a(n) \tau(n)^2 =: B'_1, \end{aligned}$$

say.

We utilize the following lemma to gain control over the divisor function.

Lemma 6.5. *For any $n, k \geq 1$ there exists a divisor $d \mid n$ such that $d \leq n^{1/2^k}$ and*

$$\tau(n) \leq 2^{2^k-1} \tau(d)^{2^k}.$$

Proof. This is [19, Lemma 4]. □

Applying Lemma 6.5 with $k = 2$ yields

$$B'_1 \ll \sum_{\substack{\mathcal{T} < d \leq P\mathcal{T} \\ d | \Pi}} \sum_{e \ll (MN)^{1/4}} \tau(e)^8 \sum_{\substack{n \ll MN \\ [d,e] | n}} a(n),$$

where $[d, e]$ is the least common multiple of d and e . By trivial estimation (i.e. no need for recourse to Lemma 3.2 since $[d, e] \leq (MN)^{1/4+o(1)}$ is so small) we find that

$$\sum_{\substack{n \ll MN \\ [d,e] | n}} a(n) \ll (MN)^{1/2} \sum_{\substack{\ell \leq (MN)^{1/2} \\ (\ell, \Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \frac{\rho_{\ell}([d, e])}{[d, e]}.$$

Recall that $P^+(n)$ and $P^-(n)$ denote the greatest and least prime factors of n , respectively. We factor e uniquely as $e = rs$, where $P^+(r) \leq P$ and $P^-(s) > P$. Thus

$$\begin{aligned} B'_1 &\ll (MN)^{1/2} \sum_{\substack{\ell \ll (MN)^{1/2} \\ (\ell, \Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{\substack{\mathcal{T} < d \leq P\mathcal{T} \\ d | \Pi}} \sum_{e \ll (MN)^{1/4}} \frac{\tau(e)^8 \rho_{\ell}([d, e])}{[d, e]} \\ &\ll (MN)^{1/2} \sum_{\substack{\ell \ll (MN)^{1/2} \\ (\ell, \Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{\substack{\mathcal{T} < d \leq P\mathcal{T} \\ d | \Pi}} \sum_{\substack{r \ll (MN)^{1/4} \\ P^+(r) \leq P}} \frac{\tau(r)^8 \rho_{\ell}([d, r])}{[d, r]} \sum_{\substack{s \ll (MN)^{1/4} \\ P^-(s) > P}} \frac{\tau(s)^8 \rho_{\ell}(s)}{s}. \end{aligned}$$

We bound the sum over s by working as in Lemma 3.4. We have

$$\begin{aligned} \sum_{\substack{s \ll (MN)^{1/4} \\ P^-(s) > P}} \frac{\tau(s)^8 \rho_{\ell}(s)}{s} &\leq \sum_{s \ll (MN)^{1/4}} \frac{\tau(s)^8 \rho_{\ell}(s)}{s} \leq \sum_{d | \ell^{\infty}} \frac{\tau(d)^8 \rho_{\ell}(d)}{d} \sum_{\substack{t \ll (MN)^{1/4} \\ (t, \ell)=1}} \frac{\tau(t)^9}{t} \\ (6.9) \quad &\ll (\log MN)^{2^9} \prod_{p | \ell} \left(1 + \frac{2^9}{p^{1/2}} \right) \ll (\log MN)^{2^9}. \end{aligned}$$

By (6.9) and the change of variables $n = [d, r]$, we obtain

$$B'_1 \ll (\log MN)^{2^9} \sum_{\substack{\ell \ll (MN)^{1/2} \\ (\ell, \Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{\substack{n > \mathcal{T} \\ P^+(n) \leq P}} \frac{\tau(n)^8 \tau_3(n) \rho_\ell(n)}{n}.$$

Since $P^-(\ell) > P$ we see that $(n, \ell) = 1$, and therefore $\rho_\ell(n) \leq \tau(n)$. Set $\varepsilon := (\log P)^{-1}$. By Rankin's trick and the inequality $\tau_3(n) \leq \tau(n)^2$ we obtain

$$\begin{aligned} \sum_{\substack{n > \mathcal{T} \\ P^+(n) \leq P}} \frac{\tau(n)^8 \tau_3(n) \rho_\ell(n)}{n} &\leq \mathcal{T}^{-\varepsilon} \sum_{P^+(n) \leq P} \frac{\tau(n)^{11}}{n^{1-\varepsilon}} = \mathcal{T}^{-\varepsilon} \prod_{p \leq P} \left(1 + \sum_{k=1}^{\infty} \frac{\tau(p^k)^{11}}{p^{k(1-\varepsilon)}} \right) \\ &\ll \mathcal{T}^{-\varepsilon} \prod_{p \leq P} \left(1 + \frac{2^{11}}{p^{1-\varepsilon}} \right) \leq \mathcal{T}^{-\varepsilon} \prod_{p \leq P} \left(1 + \frac{2^{14}}{p^{1+\varepsilon}} \right). \end{aligned}$$

The last inequality follows since $p^{2\varepsilon} \leq e^2 < 8$. We finish by observing that

$$\mathcal{T}^{-\varepsilon} \prod_{p \leq P} \left(1 + \frac{2^{14}}{p^{1+\varepsilon}} \right) \leq \mathcal{T}^{-\varepsilon} \zeta(1+\varepsilon)^{2^{14}} \ll \mathcal{T}^{-\varepsilon} \varepsilon^{-2^{14}} \leq (\log MN)^{2^{13}} P^{-1},$$

and therefore

$$(6.10) \quad B'_1 \ll (\log MN)^{2^{14}} (MN)^{1/2+\gamma_0/2} P^{-1}.$$

By (6.10) and our lower bound for P the contribution from $n_1 > \mathcal{T}$ is acceptably small for (6.8). It follows that

$$B_1(M, N) \ll \theta^{5/2} (MN)^{1/2+\gamma_0/2} + \sum_{M < m \leq 2M} \sum_{\substack{n_1 \leq \mathcal{T} \\ n_1 | \Pi}} \left| \sum_{\substack{N < n_0 n_1 \leq (1+\theta)N \\ (n_0, \Pi)=1}} \mu(n_0) a(mn_0 n_1) \right|.$$

We want to make mn_1 into a single variable, but before we can do this we need to separate the variables n_0 and n_1 . We achieve this with another short interval decomposition. We are reduced to studying

$$\sum_{\substack{G=(1+\theta^{5/2})^j \\ j \geq -1 \\ G \leq \mathcal{T}}} \sum_{M < m \leq 2M} \sum_{\substack{G < n_1 \leq (1+\theta^{5/2})G \\ n_1 | \Pi}} \left| \sum_{\substack{N < n_0 n_1 \leq (1+\theta)N \\ (n_0, \Pi)=1}} \mu(n_0) a(mn_0 n_1) \right|.$$

In the sum over n_0 we wish to replace the conditions $N < n_0 n_1$ and $n_0 n_1 \leq (1+\theta)N$ by the conditions $N < n_0 G$ and $n_0 G \leq (1+\theta)N$, respectively. If $n_0 n_1 > N$ but $n_0 G \leq N$, then

$$N < n_0 n_1 \leq (1+\theta^{5/2})n_0 G \leq (1+\theta^{5/2})N,$$

and the error in replacing the condition $n_0 n_1 > N$ by $n_0 G > N$ is

$$\ll (\log \mathcal{T}) \theta^{-5/2} \sup_{G \leq \mathcal{T}} \sum_{G < n_1 \leq (1+\theta^{5/2})G} \mu^2(n_1) \sum_{(1+\theta^{5/2})^{-1}N/G < n_0 \leq (1+\theta^{5/2})N/G} \mu^2(n_0) \sum_{\substack{n \leq 3MN \\ n_0 n_1 | n}} a(n).$$

We write these three sums as

$$\sum_{n_1} \sum_{n_0} A_{n_0 n_1}(3MN) = \sum_{n_1} \sum_{n_0} (M_{n_0 n_1}(3MN) + R_{n_0 n_1}(3MN)).$$

To estimate the remainder term we change variables

$$\sum_{n_1} \sum_{n_0} |R_{n_0 n_1}(3MN)| \leq \sum_d \tau(d) |R_d(3MN)|,$$

then apply the divisor bound $\tau(d) \ll d^{o(1)}$ and Lemma 3.2. We estimate the main term as we have before, and find that

$$\sum_{n_1} \sum_{n_0} A_{n_0 n_1}(3MN) \ll (\log MN)^{O(1)} \theta^{5/2} (MN)^{1/2+\gamma_0/2}.$$

We similarly acquire the condition $n_0 G \leq (1+\theta)N$. We then change variables $mn_1 \rightarrow m, n_0 \rightarrow n$ to obtain

$$B_1(M, N) \ll (\log MN)^{O(1)} \theta^{5/2} (MN)^{1/2+\gamma_0/2} \\ + (\log MN) \theta^{-5/2} \sup_{G \leq T} \sum_{MG < m \leq 2(1+\theta^{5/2})MG} \tau(m) \left| \sum_{\substack{N/G < n \leq (1+\theta)N/G \\ (n, \Pi)=1}} \mu(n) a(mn) \right|.$$

Observe that since $G \ll x^{o(1)}$ the size of N/G is still appropriate for Proposition 6.4 (with a smaller value of δ). In order to prove Proposition 6.2 it therefore suffices to show that

$$(6.11) \quad B_2(M, N) := \sum_{M < m \leq 2M} \tau(m) \left| \sum_{\substack{N < n \leq (1+\theta)N \\ (n, \Pi)=1}} \mu(n) a(mn) \right| \\ \ll \theta^5 (\log MN)^{O(1)} (MN)^{1/2+\gamma_0/2}.$$

We have removed the small prime factors from n . This will aid us in making m and n coprime, which in the next section will allow us to perform a factorization of our bilinear form over $\mathbb{Z}[i]$. Before estimating the contribution of those m and n which are not coprime, however, it is useful to remove the divisor function on m , as it would cause complications later on. By the Cauchy-Schwarz inequality

$$B_2(M, N) \leq B_3(M, N)^{1/2} B'_2(M, N)^{1/2},$$

where

$$B_3(M, N) := \sum_{M < m \leq 2M} \left| \sum_{\substack{N < n \leq (1+\theta)N \\ (n, \Pi)=1}} \mu(n) a(mn) \right|, \\ B'_2(M, N) := \sum_{M < m \leq 2M} \tau(m)^2 \left| \sum_{\substack{N < n \leq (1+\theta)N \\ (n, \Pi)=1}} \mu(n) a(mn) \right|.$$

We bound $B'_2(M, N)$ trivially.

Lemma 6.6. *We have $B'_2(M, N) \ll \theta (\log MN)^{2^{23}} (MN)^{1/2+\gamma_0/2}$.*

Proof. We have the trivial bound

$$B'_2(M, N) \ll \sum_{N < n \leq (1+\theta)N} \mu^2(n) \sum_{\substack{k \leq 3MN \\ d|k}} a(k) \tau(k)^2.$$

As $N \leq x^{1/2}$, say, we may apply Lemma 6.5 with $k = 2$ to arrive at

$$\begin{aligned} B'_2(M, N) &\ll \sum_{N < n \leq (1+\theta)N} \sum_{e \ll (MN)^{1/4}} \tau(e)^8 \sum_{\substack{k \leq 3MN \\ [n,e]|k}} a(k) \\ &= \sum_{N < n \leq (1+\theta)N} \sum_{e \ll (MN)^{1/4}} \tau(e)^8 (M_{[n,e]}(3MN) + R_{[n,e]}(3MN)). \end{aligned}$$

The contribution from the remainder terms is

$$\ll \sum_{d \ll N(MN)^{1/4}} \left(\sum_{\substack{n_1, n_2 \\ [n_1, n_2] = d}} 1 \right) |R_d(3MN)| \ll_\epsilon (MN)^\epsilon \sum_{d \ll N(MN)^{1/4}} |R_d(3MN)|,$$

and since $Nx^{1/4} \leq x^{3/4}$ we may bound the remainder terms with Lemma 3.2.

We estimate the main term using the same types of arguments that gave (6.10). We factor $n = bd$ and $e = rs$ to bound the main term by

$$\begin{aligned} &\ll (MN)^{1/2} \sum_{\substack{\ell \ll (MN)^{1/2} \\ (\ell, \Pi) = 1}} \mathbf{1}_A(\ell) \sum_{\substack{b \leq (1+\theta)N \\ b|\ell^\infty}} \sum_{\substack{N/b < d \leq (1+\theta)N/b \\ (d, \ell) = 1}} \\ &\times \sum_{\substack{r \ll (MN)^{1/4} \\ r|\ell^\infty}} \frac{\tau(r)^8 \rho_\ell([b, r])}{[b, r]} \sum_{\substack{s \ll (MN)^{1/4} \\ (s, \ell) = 1}} \frac{\tau(s)^8 \rho_\ell([d, s])}{[d, s]}. \end{aligned}$$

Since $(ds, \ell) = 1$ we have $\rho_\ell([d, s]) \leq \tau([d, s]) \leq \tau(ds) \leq \tau(d)\tau(s)$. We write

$$\frac{1}{[d, s]} = \frac{1}{ds} (d, s) \leq \frac{1}{ds} \sum_{\substack{f|d \\ f|s}} f,$$

which yields that the main term is

$$\begin{aligned} &\ll (MN)^{1/2} \sum_{\substack{\ell \ll (MN)^{1/2} \\ (\ell, \Pi) = 1}} \mathbf{1}_A(\ell) \sum_{\substack{b \leq (1+\theta)N \\ b|\ell^\infty}} \sum_{\substack{r \ll (MN)^{1/4} \\ r|\ell^\infty}} \frac{\tau(r)^8 \rho_\ell([b, r])}{[b, r]} \\ &\times \sum_{N/b < d \leq (1+\theta)N/b} \frac{\tau(d)}{d} \sum_{f|d} f \sum_{\substack{s \ll (MN)^{1/4} \\ f|s}} \frac{\tau(s)^9}{s} \\ &\ll (\log MN)^{2^9} (MN)^{1/2} \sum_{\substack{\ell \ll (MN)^{1/2} \\ (\ell, \Pi) = 1}} \mathbf{1}_A(\ell) \sum_{\substack{b \leq (1+\theta)N \\ b|\ell^\infty}} \sum_{\substack{r \ll (MN)^{1/4} \\ r|\ell^\infty}} \frac{\tau(r)^8 \rho_\ell([b, r])}{[b, r]} \sum_{N/b < d \leq (1+\theta)N/b} \frac{\tau(d)^{11}}{d}. \end{aligned}$$

If $b \leq N^{1/2}$ we use Lemma 6.5 with $k = 1$ to deduce

$$\begin{aligned} \sum_{N/b < d \leq (1+\theta)N/b} \frac{\tau(d)^{11}}{d} &\ll \frac{b}{N} \sum_{k \ll (N/b)^{1/2}} \tau(k)^{22} \sum_{\substack{N/b < d \leq (1+\theta)N/b \\ k|d}} 1 \ll (\log N)^{2^{22}} (\theta + (b/N)^{-1/2}) \\ &\ll (\log N)^{2^{22}} (\theta + N^{-1/4}) \ll (\log N)^{2^{22}} \theta, \end{aligned}$$

the last inequality following from the lower bound for N in (6.5). For $b > N^{1/2}$ we estimate the sum over d trivially and change variables $n = [b, r]$ to get

$$\ll (\log MN)^{2^{12}} (MN)^{1/2} \sum_{\substack{\ell \ll (MN)^{1/2} \\ (\ell, \Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{\substack{n > N^{1/2} \\ n|\ell^\infty}} \frac{\tau(n)^8 \tau_3(n) \rho_\ell(n)}{n}.$$

By Rankin's trick

$$\sum_{\substack{n > N^{1/2} \\ n|\ell^\infty}} \frac{\tau(n)^8 \tau_3(n) \rho_\ell(n)}{n} \ll N^{-1/4} \prod_{p|\ell} \left(1 + \sum_{k=1}^{\infty} \frac{2(k+1)^{10} p^{k/2}}{p^{3k/4}} \right) \ll N^{-1/4} \prod_{p|\ell} \left(1 + \frac{2^{12}}{p^{1/4}} \right),$$

Since ℓ has no small prime factors this last quantity is $\ll N^{-1/4}$. We deduce that

$$B'_2(M, N) \ll (\log MN)^{2^{23}} \theta(MN)^{1/2+\gamma_0/2},$$

as desired. \square

By Lemma 6.6 we see that in order to prove (6.11) it suffices to show that

$$(6.12) \quad B_3(M, N) \ll \theta^9(\log MN)^{O(1)} (MN)^{1/2+\gamma_0/2}.$$

We are now in a position where we can make our variables m and n coprime to one another. Since n is only divisible by primes $p > P$, if $(m, n) \neq 1$ it follows that there exists a prime $p > P$ with $p \mid m$ and $p \mid n$. Therefore the contribution from those m and n that are not coprime is bounded by

$$B'_3(M, N) := \sum_{M < m \leq 2M} \left| \sum_{\substack{N < n \leq (1+\theta)N \\ (n, \Pi)=1 \\ (n, m) \neq 1}} \mu(n) a(mn) \right| \ll \sum_{P < p \leq (MN)^{1/2}} \sum_{\substack{k \leq MN \\ p^2 \mid k}} a(k) \tau(k).$$

We estimate the contribution from $p > (MN)^{1/10}$ using the trivial bound

$$a(k) \tau(k) \ll (MN)^{1/100},$$

say. Thus

$$\begin{aligned} B'_3(M, N) &\ll (MN)^{9/10+1/100} + \sum_{P < p \leq (MN)^{1/10}} \sum_{\substack{k \leq MN \\ p^2 \mid k}} a(k) \tau(k) \\ &\ll \sum_{P < p \leq (MN)^{1/10}} \sum_{d \leq (MN)^{1/2}} \sum_{\substack{k \leq MN \\ [d, p^2] \mid k}} a(k). \end{aligned}$$

Considering separately three cases (d and p are coprime, p divides d but p^2 does not, p^2 divides d), we find that

$$B'_3(M, N) \ll \sum_{P < p \leq (MN)^{1/10}} \sum_{d \leq (MN)^{1/2}} \sum_{\substack{k \leq MN \\ dp^2 \mid k}} a(k).$$

We apply Lemma 3.2 to deduce

$$\begin{aligned}
B'_3(M, N) &\ll (MN)^{1/2} \sum_{\substack{\ell \ll (MN)^{1/2} \\ (\ell, \Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{P < p \leq (MN)^{1/10}} \frac{1}{p^2} \sum_{d \ll (MN)^{1/2}} \frac{\rho_\ell(dp^2)}{d} \\
&\ll (MN)^{1/2} \sum_{\substack{\ell \ll (MN)^{1/2} \\ (\ell, \Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{p > P} \frac{1}{p^2} \sum_{k=0}^{\infty} \sum_{\substack{d \ll (MN)^{1/2}/p^k \\ (d, p)=1}} \frac{\rho_\ell(dp^{k+2})}{dp^k} \\
&\ll (\log MN)^2 (MN)^{1/2} \sum_{\substack{\ell \ll (MN)^{1/2} \\ (\ell, \Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{p > P} \frac{1}{p^2} \sum_{k=0}^{\infty} \frac{\rho_\ell(p^{k+2})}{p^k}.
\end{aligned}$$

In going from the second line to the third line we have used Lemma 3.4 to bound the sum over d .

We consider separately the cases $(p, \ell) = 1$ and $p \mid \ell$:

$$\sum_{\substack{p > P \\ (p, \ell)=1}} \frac{1}{p^2} \sum_{k=0}^{\infty} \frac{\rho_\ell(p^{k+2})}{p^k} \ll \sum_{p > P} \frac{1}{p^2} \ll P^{-1}$$

and

$$\sum_{\substack{p > P \\ p \mid \ell}} \frac{1}{p^2} \sum_{k=0}^{\infty} \frac{\rho_\ell(p^{k+2})}{p^k} \ll \sum_{\substack{p > P \\ p \mid \ell}} \frac{1}{p} \ll (\log \ell) P^{-1},$$

where we have used Lemma 3.3 to control the behavior of $\rho_\ell(p^{k+2})$. It follows that

$$B_3(M, N) \ll (\log MN)^{O(1)} (MN)^{1/2+\gamma_0/2} P^{-1} + \sum_{M < m \leq 2M} \left| \sum_{\substack{N < n \leq (1+\theta)N \\ (n, m\Pi)=1}} \mu(n) a(mn) \right|.$$

In order to prove (6.12) it therefore suffices to show that

$$B_4(M, N) = \sum_{M < m \leq 2M} \left| \sum_{\substack{N < n \leq (1+\theta)N \\ (n, m\Pi)=1}} \mu(n) a(mn) \right| \ll \theta^9 (\log MN)^{O(1)} (MN)^{1/2+\gamma_0/2},$$

and this is precisely the content of Proposition 6.4. Hence Proposition 6.2 follows from Proposition 6.4.

7. BILINEAR FORM IN THE SIEVE: TRANSFORMATIONS

In this section we perform further manipulations on the bilinear form $B_4(M, N)$ appearing in the statement of Proposition 6.4. By hypothesis we have that m and n are coprime, and this will allow us to enter the realm of the Gaussian integers. This is the key step that allows us to prove Proposition 6.4 (see the discussion in [15, Section 5] for more insight on this).

Since m and n are coprime the unique factorization in $\mathbb{Z}[i]$ gives

$$a(mn) = \frac{1}{4} \sum_{\substack{|w|^2=m, |z|^2=n \\ (\text{Re}(z\bar{w}), \Pi)=1}} \mathbf{1}_{\mathcal{A}}(\text{Re}(z\bar{w})).$$

(It shall be highly convenient in a number of places, though not essential, that the ring of integers of the number field $\mathbb{Q}(i)$ has class number one.) Since $(n, \Pi) = 1$ we have $(z\bar{z}, \Pi) = 1$, so in particular z is odd. Multiplying w and z by a unit we can rotate z to a number satisfying

$$z \equiv 1 \pmod{2(1+i)}.$$

Such a number is called primary, and is determined uniquely by its ideal. In rectangular coordinates $z = r + is$ being primary is equivalent to

$$(7.1) \quad r \equiv 1 \pmod{2}, \quad s \equiv r - 1 \pmod{4},$$

so that r is odd and s is even. We therefore obtain

$$B_4(M, N) \leq \mathcal{B}_1(M, N) := \sum_{M < |w|^2 \leq 2M} \left| \sum_{\substack{N < |z|^2 \leq (1+\theta)N \\ (z\bar{z}, w\bar{w}\Pi)=1 \\ (\text{Re}(z\bar{w}), \Pi)=1}} \mu(|z|^2) \mathbf{1}_{\mathcal{A}}(\text{Re}(z\bar{w})) \right|.$$

Here we assume that z runs over primary numbers, so that the factor of $\frac{1}{4}$ does not occur. Further, the presence of the Möbius function implies we may take z to be primitive, that is, $z = r + is$ with $(r, s) = 1$. Henceforth a summation over Gaussian integers z is always assumed to be over primary, primitive Gaussian integers.

The condition $(m, n) = 1$ was crucial in obtaining a factorization of our bilinear form over $\mathbb{Z}[i]$, but now this condition has become $(w\bar{w}, z\bar{z}) = 1$ which is a nuisance since we wish for w and z to run independently of one another. Because $z\bar{z}$ has no small prime factors, it suffices to estimate trivially the complimentary sum in which $(w\bar{w}, z\bar{z}) \neq 1$.

The arguments of this section bear some semblance to those in [15, Section 5] and [16, Section 20.4]. The plan of this section is as follows. We remove the condition $(w\bar{w}, z\bar{z}) = 1$ in order to make w and z more independent. With this condition gone we apply the Cauchy-Schwarz inequality to arrive at sums of the form

$$\sum_w \left| \sum_z \mu(|z|^2) \mathbf{1}_{\mathcal{A}}(\text{Re}(z\bar{w})) \right|^2 = \sum_w \sum_{z_1, z_2} \mu(|z_1|^2) \mu(|z_2|^2) \mathbf{1}_{\mathcal{A}}(\text{Re}(z_1\bar{w})) \mathbf{1}_{\mathcal{A}}(\text{Re}(z_2\bar{w})).$$

For technical reasons it is convenient to impose the condition that z_1 and z_2 are coprime to each other. The key is again the fact that $|z_i|^2$ has no small prime factors. Once this is accomplished, we change variables to arrive at sums of the form

$$\sum_{z_1, z_2} \mu(|z_1|^2) \mu(|z_2|^2) \sum_{\ell_1, \ell_2} \mathbf{1}_{\mathcal{A}}(\ell_1) \mathbf{1}_{\mathcal{A}}(\ell_2),$$

where ℓ_1, ℓ_2 are rational integers. The variable w has disappeared, but now there are numerous conditions entangling z_1, z_2 and the ℓ_i . Foremost among these conditions is a congruence to modulus Δ , which is the imaginary part of $\bar{z}_1 z_2$. The contribution from $\Delta = 0$ is easily dispatched, but the estimation of the terms with $\Delta \neq 0$ is much more involved and is handled in future sections.

In particular, in this section we reduce the proof of Proposition 6.4 to the proof of the following result.

Proposition 7.1. *Let $0 < \delta \leq \frac{1}{100}$ be fixed. For fixed $L > 0$ set $\theta = (\log x)^{-L}$. Assume the notation above. Then*

$$\mathcal{D}_3(M, N) := \sum_{\substack{N < |z_1|^2, |z_2|^2 \leq (1+\theta)N \\ (|z_1|^2 |z_2|^2, \Pi) = 1 \\ (z_1, z_2) = 1 \\ \Delta \neq 0}} \mu(|z_1|^2) \mu(|z_2|^2) \sum_{\substack{\ell_1, \ell_2 \leq \sqrt{2(1+\theta)}(MN)^{1/2} \\ \ell_1 z_2 \equiv \ell_2 z_1 \pmod{|\Delta|} \\ |\ell_1 z_2 - \ell_2 z_1|^2 \leq 2\Delta^2 M \\ (\ell_1 \ell_2, \Pi) = 1}} \mathbf{1}_{\mathcal{A}}(\ell_1) \mathbf{1}_{\mathcal{A}}(\ell_2)$$

satisfies

$$\mathcal{D}_3(M, N) \ll_{\delta} \theta^{18} (\log MN)^{O(1)} (MN)^{\gamma_0} N$$

uniformly for

$$x^{\delta} \leq N \leq x^{25/77-\delta}, \quad \theta x \ll MN \ll x.$$

Let $\mathcal{B}'_1(M, N)$ denote the contribution to $\mathcal{B}_1(M, N)$ from those w and z with $(w\bar{w}, z\bar{z}) \neq 1$. We estimate $\mathcal{B}'_1(M, N)$ trivially and show that it is sufficiently small.

Lemma 7.2. *With the notation as above, we have*

$$\mathcal{B}'_1(M, N) \ll (\log MN)^2 (MN)^{1/2+\gamma_0/2} P^{-1}.$$

Proof. We find

$$\mathcal{B}'_1(M, N) \ll \sum_{\ell \ll (MN)^{1/2}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{p > P} \sum_{\substack{N < r^2 + s^2 \leq 2N \\ r^2 + s^2 \equiv 0(p) \\ (r, s) = 1}} \sum_{\substack{M < u^2 + v^2 \leq 2M \\ u^2 + v^2 \equiv 0(p) \\ ru + sv = \ell}} 1.$$

Observe that $p \nmid rs$ since $r^2 + s^2 \equiv 0 \pmod{p}$ and $(r, s) = 1$.

Given fixed p, ℓ, r , and s , we claim that the residue class of u is fixed modulo $ps/(\ell, p)$. Indeed, we see that u is in a fixed residue class modulo s , since $ru + sv = \ell$ implies $u \equiv \bar{r}\ell \pmod{s}$. If $p \mid \ell$ this gives the claim, so assume $p \nmid \ell$. Then $v \equiv \bar{s}(\ell - ru) \pmod{p}$, which yields

$$0 \equiv u^2 + v^2 \equiv u^2 + (\bar{s})^2(\ell - ru)^2 \pmod{p}.$$

We multiply both sides of the congruence by s^2 , expand out $(\ell - ru)^2$, and use the fact that $r^2 + s^2 \equiv 0(p)$. This gives

$$2\ell ru \equiv \ell^2 \pmod{p}.$$

Since ℓ is coprime to p we can divide both sides by ℓ , and we can divide by $2r$ since $p \nmid 2r$. Thus the class of u is fixed modulo p . Since the class of u is fixed modulo p and modulo s , and since $(p, s) = 1$, the Chinese remainder theorem gives that the class of u is fixed modulo ps . This completes the proof of the claim.

If p, ℓ, r, s , and u are given, then v is determined. The sum over u, v is then bounded by

$$\ll \frac{M^{1/2}(\ell, p)}{ps} + 1.$$

By the symmetry of u and v we also have that the sum over u, v is bounded by

$$\ll \frac{M^{1/2}(\ell, p)}{pr} + 1.$$

Since $r^2 + s^2 > N$, either $r \gg N^{1/2}$ or $s \gg N^{1/2}$, so we may bound the sum over u, v by

$$\ll \frac{M^{1/2}(\ell, p)}{N^{1/2} p} + 1.$$

We also note that

$$\sum_{\substack{N < r^2 + s^2 \leq 2N \\ r^2 + s^2 \equiv 0(p) \\ (r, s) = 1}} \sum_{\substack{n \leq N \\ p|n}} 1 \ll \sum_{\substack{n \leq N \\ p|n}} \tau(n) \ll \frac{N \log N}{p}.$$

Therefore

$$(7.2) \quad \mathcal{B}'_1(M, N) \ll (\log N)(MN)^{1/2} \sum_{\ell \leq (MN)^{1/2}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{P < p \leq N} \frac{(\ell, p)}{p^2} + (\log N)^2 (MN)^{\gamma_0/2} N.$$

The second term in (7.2) is sufficiently small for Proposition 6.4 since $N \leq x^{1/2-1/100}$, say.

To bound the first term we note that

$$\sum_{P < p \leq N} \frac{(\ell, p)}{p^2} \leq \sum_{\substack{p > P \\ p \nmid \ell}} \frac{1}{p^2} + \sum_{\substack{p > P \\ p \mid \ell}} \frac{1}{p} \ll (\log \ell) P^{-1},$$

and this gives the bound

$$\mathcal{B}'_1(M, N) \ll (\log MN)^2 (MN)^{1/2+\gamma_0/2} P^{-1}. \quad \square$$

Lemma 7.2 proves that Proposition 6.4 follows from the bound

$$\mathcal{B}_2(M, N) := \sum_{M < |w|^2 \leq 2M} \left| \sum_{\substack{N < |z|^2 \leq (1+\theta)N \\ (z\bar{z}, \Pi) = 1 \\ (\text{Re}(z\bar{w}), \Pi) = 1}} \mu(|z|^2) \mathbf{1}_{\mathcal{A}}(\text{Re}(z\bar{w})) \right| \ll \theta^9 (\log MN)^{O(1)} (MN)^{1/2+\gamma_0/2}.$$

We now apply the Cauchy-Schwarz inequality, obtaining

$$\mathcal{B}_2(M, N)^2 \ll M \mathcal{D}_1(M, N),$$

where

$$\mathcal{D}_1(M, N) := \sum_{|w|^2 \leq 2M} \left| \sum_{\substack{N < |z|^2 \leq (1+\theta)N \\ (z\bar{z}, \Pi) = 1 \\ (\text{Re}(z\bar{w}), \Pi) = 1}} \mu(|z|^2) \mathbf{1}_{\mathcal{A}}(\text{Re}(z\bar{w})) \right|^2.$$

Note that we have used positivity to extend the sum over w . Proposition 6.4 then follows from the bound

$$(7.3) \quad \mathcal{D}_1(M, N) \ll \theta^{18} (\log MN)^{O(1)} (MN)^{\gamma_0} N.$$

Expanding the square in $\mathcal{D}_1(M, N)$ gives a sum over w, z_1 , and z_2 , say. As mentioned above, we wish to impose the condition that z_1 and z_2 are coprime. To do so we first require a trivial bound. Observe that

$$\mathcal{D}_1(M, N) \leq D'_1(M, N) := \sum_{|w|^2 \leq 2M} \sum_{N < |z_1|^2, |z_2|^2 \leq 2N} \mathbf{1}_{\mathcal{A}}(\operatorname{Re}(z_1 \bar{w})) \mathbf{1}_{\mathcal{A}}(\operatorname{Re}(z_2 \bar{w})).$$

Lemma 7.3. *We have*

$$D'_1(M, N) \ll ((MN)^{1/2+\gamma_0/2} + (MN)^{\gamma_0} N) (\log MN)^{36}.$$

Proof. We consider separately the diagonal $|z_1| = |z_2|$ and the off-diagonal $|z_1| \neq |z_2|$ cases.

We can bound the diagonal terms by

$$D'_=(M, N) := \sum_{\ell \ll (MN)^{1/2}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{\substack{N < r^2 + s^2 \leq 2N \\ (r,s)=1}} \tau(r^2 + s^2) \sum_{\substack{u^2 + v^2 \leq 2M \\ ru + sv = \ell}} 1.$$

By an argument similar to that which yielded (7.2), we bound the sum over u, v by

$$\ll \min \left(\frac{M^{1/2}}{r} + 1, \frac{M^{1/2}}{s} + 1 \right) \ll \frac{M^{1/2}}{N^{1/2}} + 1 \ll \frac{M^{1/2}}{N^{1/2}}.$$

The sum over r and s is bounded by

$$\sum_{\substack{N < r^2 + s^2 \leq 2N \\ (r,s)=1}} \tau(r^2 + s^2) \ll \sum_{n \leq 2N} \tau(n)^2 \ll N(\log N)^3,$$

and we deduce that

$$D'_=(M, N) \ll (\log N)^3 (MN)^{1/2+\gamma_0/2}.$$

We turn now to bounding the off-diagonal terms with $|z_1| \neq |z_2|$. Observe that

$$\Delta = \Delta(z_1, z_2) = \frac{1}{2i} (\bar{z}_1 z_2 - z_1 \bar{z}_2) \neq 0,$$

since $(z_1, \bar{z}_1) = (z_2, \bar{z}_2) = 1$. The off-diagonal terms therefore contribute

$$D'_{\neq}(M, N) \ll \sum_{\ell_1, \ell_2 \ll (MN)^{1/2}} \mathbf{1}_{\mathcal{A}}(\ell_1) \mathbf{1}_{\mathcal{A}}(\ell_2) \sum_{\substack{N < |z_1|^2, |z_2|^2 \leq 2N \\ |z_1| \neq |z_2| \\ \Delta |(\ell_1 z_2 - \ell_2 z_1)}} 1.$$

We note that the division takes place in the Gaussian integers, and that $\ell_1 z_2 - \ell_2 z_1 \neq 0$ (see (7.5) below). Using rectangular coordinates $z_1 = r_1 + is_1, z_2 = r_2 + is_2$, we see that $\Delta = r_1 s_2 - r_2 s_1$ and

$$\begin{aligned} \ell_1 r_2 &\equiv \ell_2 r_1 \pmod{\Delta}, \\ \ell_1 s_2 &\equiv \ell_2 s_1 \pmod{\Delta}, \end{aligned}$$

where now the congruences are congruences of rational integers. By symmetry we may assume that $\ell_1 s_2 - \ell_2 s_1 \neq 0$. Given ℓ_1, ℓ_2, s_1, s_2 , and $\Delta \neq 0$, we see that the residue class of r_1 modulo $s_1/(s_1, s_2)$ is fixed, and then r_2 is determined by the relation $\Delta = r_1 s_2 - r_2 s_1$. The number of pairs r_1, r_2 is then bounded by

$$\ll \sqrt{N} \frac{(s_1, s_2)}{s_1}.$$

Letting $\omega = (s_1, s_2)$ and $s_1 = \omega s_1^*, s_2 = \omega s_2^*$ (so that $(s_1^*, s_2^*) = 1$), we see that

$$D'_{\neq}(M, N) \ll \sqrt{N} \sum_{\omega \ll N^{1/2}} \tau(\omega) \sum_{\substack{s_1^*, s_2^* \ll N^{1/2}/\omega \\ (s_1^*, s_2^*)=1}} \sum_{\substack{\ell_1, \ell_2 \ll (MN)^{1/2} \\ \ell_1 s_2^* - \ell_2 s_1^* \neq 0}} \frac{1}{s_1^*} \mathbf{1}_{\mathcal{A}}(\ell_1) \mathbf{1}_{\mathcal{A}}(\ell_2) \tau(\ell_1 s_2^* - \ell_2 s_1^*).$$

We have used the fact that, given $\ell_1, s_1^*, s_2^*, \ell_2$, and ω , the number of choices of Δ is at most $\tau(\omega) \tau(\ell_1 s_2^* - \ell_2 s_1^*)$, since $\Delta \mid \ell_1 s_2 - \ell_2 s_1 = \omega(\ell_1 s_2^* - \ell_2 s_1^*)$.

Observe now that $|\ell_1 s_2^* - \ell_2 s_1^*| \ll \sqrt{MN}$. We apply Lemma 6.5 with $k = 2$ to get

$$D'_{\neq}(M, N) \ll \sqrt{N} \sum_{\omega \ll N^{1/2}} \tau(\omega) \sum_{\substack{s_1^*, s_2^* \ll N^{1/2}/\omega \\ (s_1^*, s_2^*)=1}} \sum_{f \ll F} \frac{1}{s_1^*} \tau(f)^4 \sum_{\substack{\ell_1, \ell_2 \ll (MN)^{1/2} \\ \ell_1 s_2^* \equiv \ell_2 s_1^* (f)}} \mathbf{1}_{\mathcal{A}}(\ell_1) \mathbf{1}_{\mathcal{A}}(\ell_2),$$

where $F = (\sqrt{MN})^{1/4}$. Taking the supremum over s_2^* and ω gives

$$D'_{\neq}(M, N) \ll (\log N)^2 N \sum_{\substack{s_1^* \leq N' \\ (s_1^*, s_2^*)=1}} \sum_{f \ll F} \frac{1}{s_1^*} \tau(f)^4 \sum_{\substack{\ell_1, \ell_2 \ll (MN)^{1/2} \\ \ell_1 s_2^* \equiv \ell_2 s_1^* (f)}} \mathbf{1}_{\mathcal{A}}(\ell_1) \mathbf{1}_{\mathcal{A}}(\ell_2)$$

for some $N', s_2^* \ll N^{1/2}$. We now write $f = gh, s_1^* = hs$ with $(g, s) = 1$. Observe that $(h, s_2^*) = 1$. Then the congruence $\ell_1 s_2^* \equiv \ell_2 s_1^* (f)$ yields the congruences

$$\begin{aligned} \ell_1 &\equiv 0 \pmod{h}, \\ \ell_2 &\equiv \bar{s} s_2^* (\ell_1/h) \pmod{g}, \end{aligned}$$

where \bar{s} is the inverse of s modulo g . We deduce that

$$D'_{\neq}(M, N) \ll (\log N)^2 N \sum_{gh \leq F} \frac{\tau(g)^4 \tau(h)^4}{h} \sum_{s \leq N'/h} \frac{1}{s} \sum_{\substack{\ell_1 < X \\ h \mid \ell_1}} \mathbf{1}_{\mathcal{A}}(\ell_1) \sum_{\substack{\ell_2 < X \\ \ell_2 \equiv \nu(g)}} \mathbf{1}_{\mathcal{A}}(\ell_2),$$

where X is a power of 10 with $X \asymp (MN)^{1/2}$ and $\nu = \nu(h, \ell_1, s, s_2^*)$ is a residue class. We detect the congruence on ℓ_2 with additive characters, and then apply the triangle inequality to eliminate ν (we have already seen this technique in the proof of Lemma 5.5). We then drop the divisibility condition on ℓ_1 , obtaining

$$D'_{\neq}(M, N) \ll (\log MN)^{19} (MN)^{\gamma_0} N \sum_{g \leq F} \frac{\tau(g)^4}{g} \sum_{r=1}^g F_X \left(\frac{r}{g} \right).$$

Reducing to primitive fractions gives

$$\sum_{g \leq F} \frac{\tau(g)^4}{g} \sum_{r=1}^g F_X \left(\frac{r}{g} \right) \ll (\log MN)^{16} \sum_{1 \leq q \leq F} \frac{\tau(q)^4}{q} \sum_{\substack{1 \leq a \leq q \\ (a, q)=1}} F_X \left(\frac{a}{q} \right).$$

By the divisor bound $\tau(q) \ll q^{o(1)}$, a dyadic division, and Lemma 5.4, we find this last quantity is

$$\ll (\log MN)^{17} \sup_{Q \leq F} \left(\frac{1}{Q^{23/77-o(1)}} + \frac{Q^{1+o(1)}}{X^{50/77}} \right) \ll (\log MN)^{17} \left(1 + \frac{F^{1+o(1)}}{X^{50/77}} \right).$$

Observe that

$$F = (\sqrt{MN})^{1/4} \leq (\sqrt{MN})^{1/2} \ll X^{1/2},$$

which yields

$$D'_{\neq}(M, N) \ll (\log MN)^{36} (MN)^{\gamma_0} N. \quad \square$$

With Lemma 7.3 in hand, we can show that the contribution from $(z_1, z_2) \neq 1$ in $\mathcal{D}_1(M, N)$ is negligible. This is due to the fact that $(|z_i|^2, \Pi) = 1$. Denoting by π a Gaussian prime, the contribution from $(z_1, z_2) \neq 1$ is bounded by

$$\sum_{P < |\pi|^2 \ll N} \sum_{|w|^2 \ll M} \sum_{\substack{N < |z_1|^2, |z_2|^2 \leq \frac{2N}{|\pi|^2} \\ (|z_i|^2, \Pi) = 1}} \sum_{\substack{(z_1, z_2) = 1 \\ (\text{Re}(z_1 \bar{w}) \text{Re}(z_2 \bar{w}), \Pi) = 1}} \mathbf{1}_{\mathcal{A}}(\text{Re}(z_1 \pi \bar{w})) \mathbf{1}_{\mathcal{A}}(\text{Re}(z_2 \pi \bar{w})).$$

We break the range of $|\pi|^2$ into dyadic intervals $P_1 < |\pi|^2 \leq 2P_1$, and put $w' = w\bar{\pi}$. We observe that

$$\sum_{\pi | \mathfrak{z}} 1 \ll \log |\mathfrak{z}|$$

for any Gaussian integer \mathfrak{z} , so the contribution from the pairs z_1, z_2 that are not coprime is bounded by

$$\ll (\log MN)^2 D'_1(MP_1, NP_1^{-1}),$$

for some $P \ll P_1 \ll N$. By Lemma 7.3 this bound becomes

$$\ll (\log MN)^{38} ((MN)^{1/2+\gamma_0/2} + (MN)^{\gamma_0} NP^{-1}).$$

The second term is satisfactorily small, and the first is sufficiently small since $N \geq x^{1/2-\gamma_0/2+\delta}$ by the hypotheses of Proposition 6.4. (The lower bound on N in Proposition 6.4 is required to control this particular diagonal contribution. After this point we require only the weaker lower bound on N in Proposition 7.1.)

In order to show (7.3) it then suffices to show that

$$\begin{aligned} \mathcal{D}_2(M, N) &:= \sum_{|w|^2 \leq 2M} \sum_{\substack{N < |z_1|^2, |z_2|^2 \leq (1+\theta)N \\ (|z_1|^2 |z_2|^2, \Pi) = 1 \\ (z_1, z_2) = 1 \\ (\text{Re}(z_1 \bar{w}) \text{Re}(z_2 \bar{w}), \Pi) = 1}} \mu(|z_1|^2) \mu(|z_2|^2) \mathbf{1}_{\mathcal{A}}(\text{Re}(z_1 \bar{w})) \mathbf{1}_{\mathcal{A}}(\text{Re}(z_2 \bar{w})) \\ (7.4) \quad &\ll \theta^{18} (\log MN)^{O(1)} (MN)^{\gamma_0} N. \end{aligned}$$

Now that z_1 and z_2 are coprime we change variables in order to rid ourselves of the variable w . We put $\ell_1 = \text{Re}(z_1 \bar{w})$ and $\ell_2 = \text{Re}(z_2 \bar{w})$, that is,

$$\begin{aligned} z_1 \bar{w} + \bar{z}_1 w &= 2\ell_1, \\ z_2 \bar{w} + \bar{z}_2 w &= 2\ell_2. \end{aligned}$$

We set $\Delta = \Delta(z_1, z_2) = \text{Im}(\bar{z}_1 z_2) = \frac{1}{2i}(\bar{z}_1 z_2 - z_1 \bar{z}_2)$, and note that

$$(7.5) \quad iw\Delta = \ell_1 z_2 - \ell_2 z_1.$$

It follows that

$$\mathcal{D}_2(M, N) = \sum_{\substack{N < |z_1|^2, |z_2|^2 \leq (1+\theta)N \\ (|z_1|^2 |z_2|^2, \Pi) = 1 \\ (z_1, z_2) = 1}} \mu(|z_1|^2) \mu(|z_2|^2) \sum_{\substack{\ell_1, \ell_2 \leq \sqrt{2(1+\theta)}(MN)^{1/2} \\ \ell_1 z_2 \equiv \ell_2 z_1 \pmod{|\Delta|} \\ |\ell_1 z_2 - \ell_2 z_1|^2 \leq 2\Delta^2 M \\ (\ell_1 \ell_2, \Pi) = 1}} \mathbf{1}_{\mathcal{A}}(\ell_1) \mathbf{1}_{\mathcal{A}}(\ell_2).$$

Note that the congruence $\ell_1 z_2 \equiv \ell_2 z_1 \pmod{|\Delta|}$ is a congruence of Gaussian integers.

The contribution from $\Delta = 0$ to \mathcal{D}_2 is bounded by

$$\mathcal{D}'_2 := \sum_{\substack{|z_1|^2, |z_2|^2 \ll N \\ \text{Im}(\bar{z}_1 z_2) = 0}} \sum_{\ell \ll (MN)^{1/2}} \mathbf{1}_{\mathcal{A}}(\ell),$$

since if $\Delta = 0$ the triple (z_1, z_2, ℓ_1) determines ℓ_2 . The summation over ℓ is bounded by $O((MN)^{\gamma_0/2})$. Writing $z_1 = r + is$ and $z_2 = u + iv$, we may bound the sum over z_1, z_2 by

$$\sum_{r, v \ll N^{1/2}} \sum_{\substack{s, u \ll N^{1/2} \\ su = rv}} 1 \leq \sum_{r, v \ll N^{1/2}} \tau(rv) \ll N \log^2 N.$$

Thus

$$\mathcal{D}'_2 \ll (\log N)^2 (MN)^{\gamma_0/2} N,$$

and this is acceptable for (7.4) since $MN \gg \theta x$. It therefore suffices to show that

$$\mathcal{D}_3(M, N) \ll \theta^{18} (\log MN)^{O(1)} (MN)^{\gamma_0} N,$$

where \mathcal{D}_3 is \mathcal{D}_2 with the additional condition that $\Delta \neq 0$. Note that \mathcal{D}_3 is the sum appearing in Proposition 7.1, and therefore Proposition 6.4 follows from Proposition 7.1.

8. COPRIMALITY CONDITIONS AND FURTHER REDUCTIONS

In the previous section we performed a factorization of our bilinear form and entered the Gaussian domain. This factorization is one of the key points in the proof of Theorem 1.1. The next key point is handling the congruence $\ell_1 z_2 \equiv \ell_2 z_1 \pmod{|\Delta|}$ appearing in the bilinear form $\mathcal{D}_3(M, N)$ of Proposition 7.1. The modulus $|\Delta|$ is usually large, and this creates many difficulties. In this section we perform further reductions and manipulations, paving the way for (mostly) removing this congruence modulo $|\Delta|$ in the next section.

To handle the condition $\ell_1 z_2 \equiv \ell_2 z_1 \pmod{|\Delta|}$, we sum over all residue classes b modulo $|\Delta|$ such that $bz_1 \equiv z_2 \pmod{|\Delta|}$. Then, with ℓ_1 fixed, we sum over $\ell_2 \equiv b\ell_1 \pmod{|\Delta|}$.

Recall that the congruence $\ell_1 z_2 \equiv \ell_2 z_1 \pmod{|\Delta|}$ is a congruence of Gaussian integers, so that $b \in \mathbb{Z}[i]/|\Delta|\mathbb{Z}[i]$. We shall show shortly, however, that we may take $b \in \mathbb{Z}/|\Delta|\mathbb{Z}$.

A helpful observation is that, since z_1 and z_2 are coprime, b is uniquely determined modulo $|\Delta|$. That is,

$$\sum_{\substack{b \pmod{|\Delta|} \\ bz_1 \equiv z_2 \pmod{|\Delta|}}} 1 = 1.$$

Note that this follows at once if $(z_1, |\Delta|) = 1$. But if π is a Gaussian prime dividing both z_1 and $|\Delta|$, then the congruence $bz_1 \equiv z_2 \pmod{|\Delta|}$ implies $\pi \mid z_2$, which contradicts the fact that z_1 and z_2 are coprime Gaussian integers. Hence $(z_1, |\Delta|) = 1$ and b is uniquely determined modulo $|\Delta|$.

One problem we face is that, *a priori*, the congruence $\ell_2 \equiv b\ell_1 \pmod{|\Delta|}$ is not a congruence of rational integers, but a congruence of Gaussian integers. If we write $b = r + is$, then we see that the Gaussian congruence $\ell_2 \equiv b\ell_1 \pmod{|\Delta|}$ is equivalent to the rational congruences

$$\begin{aligned} \ell_2 &\equiv r\ell_1 \pmod{|\Delta|}, \\ s\ell_1 &\equiv 0 \pmod{|\Delta|}. \end{aligned}$$

If we can take ℓ_1 to be coprime to $|\Delta|$, then this implies $s \equiv 0 \pmod{|\Delta|}$. As s is only defined modulo $|\Delta|$ we may then take s to be zero, which implies b is rational.

Lastly, with a view towards using the fundamental lemma to control the condition $(\ell_2, \Pi) = 1$, we anticipate sums of the form

$$\sum_{\substack{\ell_2 \equiv b\ell_1(|\Delta|) \\ \ell_2 \equiv 0(d)}} \mathbf{1}_{\mathcal{A}}(\ell_2).$$

If we can ensure that b is coprime to $|\Delta|$, then the first congruence implies ℓ_2 is coprime to $|\Delta|$ (recall we are assuming for the moment that $(\ell_1, |\Delta|) = 1$). Taking the first and second congruences together we see that $(d, |\Delta|) = 1$, so that the set of congruences may be combined by the Chinese remainder theorem into a single congruence modulo $d|\Delta|$. We can take b to be coprime to $|\Delta|$ by imposing the condition $(z_1 z_2, |\Delta|) = 1$. Actually, we saw above that z_1 is already coprime to $|\Delta|$, so we only need to make z_2 coprime to $|\Delta|$.

Another technical obstacle to overcome is that the set \mathcal{A} is not well-distributed in residue classes to moduli that are not coprime to 10. Since we have essentially no control over the 2- or 5-adic valuation of $|\Delta|$, we need to work around the “10-adic” part of $|\Delta|$ somehow. In an effort to isolate this poor behavior at the primes 2 and 5 we write

$$(8.1) \quad \Delta = \Delta_{10} |\Delta'|,$$

where Δ_{10} is a positive divisor of 10^∞ and $(|\Delta'|, 10) = 1$. Note that $2 \mid \Delta_{10}$ since z_1 and z_2 are primary (see (7.1)). By the Chinese remainder theorem we can think about the congruence $\ell_2 \equiv b\ell_1(|\Delta|)$ as two separate rational congruences, one to modulus Δ_{10} and one to modulus $|\Delta'|$. Because integers divisible only by the primes 2 and 5 form a very sparse subset of all the integers, we expect the contribution from large Δ_{10} to be negligible.

Proposition 8.1. *Let $0 < \delta \leq \frac{1}{100}$ be fixed. For fixed $L > 0$ set $\theta = (\log x)^{-L}$. Assume the notation above. Then*

$$\begin{aligned} \mathcal{D}_5(M, N) := & \sum_{\substack{N < |z_1|^2, |z_2|^2 \leq (1+\theta)N \\ (z_1 \bar{z}_1 z_2 \bar{z}_2, \Pi) = 1 \\ (z_1 | \Delta, z_2) = 1 \\ |\Delta| > \theta^{18} N \\ \Delta_{10} \leq \theta^{-28}}} \mu(|z_1|^2) \mu(|z_2|^2) \sum_{bz_1 \equiv z_2(|\Delta|)} \\ & \times \sum_{\substack{\ell_1 \leq \sqrt{2(1+\theta)(MN)^{1/2}} \\ (\ell_1, \Pi|\Delta|) = 1}} \mathbf{1}_{\mathcal{A}}(\ell_1) \sum_{\substack{\ell_2 \leq \sqrt{2(1+\theta)(MN)^{1/2}} \\ \ell_2 \equiv b\ell_1(|\Delta|) \\ |\ell_1 z_2 - \ell_2 z_1|^2 \leq 2\Delta^2 M \\ (\ell_2, \Pi) = 1}} \mathbf{1}_{\mathcal{A}}(\ell_2) \end{aligned}$$

satisfies

$$\mathcal{D}_5(M, N) \ll_\delta \theta^{18} (\log MN)^{O(1)} (MN)^{\gamma_0} N$$

uniformly in

$$x^\delta \leq N \leq x^{25/77-\delta}, \quad \theta x \ll MN \ll x.$$

The implied constant is ineffective.

We use the remainder of this section to show that Proposition 7.1 follows from Proposition 8.1.

We begin by removing from \mathcal{D}_3 those $|\Delta|$ that are unusually small (compare [19, (17)] and the following discussion for a similar computation). Since $\Delta = \text{Im}(\bar{z}_1 z_2)$ and $|z_i| \asymp N^{1/2}$, we expect that typically $|\Delta| \approx N$, and perhaps that those $|\Delta|$ that are much smaller than N should have a negligible contribution.

Lemma 8.2. *The contribution to \mathcal{D}_3 from $|\Delta| \leq \theta^{18}N$ is*

$$\ll \theta^{18}(\log N)^2(MN)^{\gamma_0}N.$$

Proof. We estimate trivially the contribution from $|\Delta| \leq \theta^{18}N$. By the triangle inequality, this contribution is bounded by

$$D'_3(M, N) := \sum_{\substack{N < |z_1|^2, |z_2|^2 \leq (1+\theta)N \\ (z_1, z_2) = 1 \\ 0 < |\Delta| \leq \theta^{18}N}} \sum_{\substack{b(|\Delta|) \\ bz_1 \equiv z_2(|\Delta|)}} \sum_{\ell_1 \ll (MN)^{1/2}} \mathbf{1}_{\mathcal{A}}(\ell_1) \sum_{\substack{\ell_2 \equiv \text{Re}(b)\ell_1(|\Delta|) \\ (8.2)}} \mathbf{1}_{\mathcal{A}}(\ell_2),$$

where (8.2) denotes the condition

$$(8.2) \quad \left| \ell_2 - \ell_1 \frac{z_2}{z_1} \right| \ll \theta^{18}(MN)^{1/2}.$$

Observe that (8.2) forces ℓ_2 to lie in an interval $I = I(\ell_1, z_2, z_1)$ of length $\leq c\theta^{18}(MN)^{1/2}$, for some positive absolute constant c .

We use the “intervals of length a power of ten” technique we deployed in analyzing (4.1) (see (5.3)). Let Y be the largest power of 10 satisfying $Y \leq \theta^{18}(MN)^{1/2}$, and cover the interval I with subintervals of the form $[nY, (n+1)Y)$, where n is a nonnegative integer. Recalling from the proof of Lemma 5.5 that we argue slightly differently depending on whether a_0 is zero or not, we have

$$\sum_{\substack{\ell_2 \equiv \text{Re}(b)\ell_1(|\Delta|) \\ (8.2)}} \mathbf{1}_{\mathcal{A}}(\ell_2) \leq \sum_{n \in S(I)} \mathbf{1}_{\mathcal{A}}(n) \sum_{\substack{\delta(a_0)Y/10 \leq t < Y \\ t+nY \equiv \text{Re}(b)\ell_1(|\Delta|)}} \mathbf{1}_{\mathcal{A}}(t),$$

where $S(I)$ is some set of integers depending on I . The size of $S(I)$ is $O(1)$, and this bound is independent of I . We detect the congruence condition via additive characters, and separate the zero frequency from the nonzero frequencies. On the nonzero frequencies we apply inclusion-exclusion and then the triangle inequality so that t runs over an interval of the form $t < Y/10$ or $t < Y$. This application of the triangle inequality also removes the dependence on n, b , and ℓ_1 . It follows that

$$\sum_{\substack{\ell_2 \equiv \text{Re}(b)\ell_1(|\Delta|) \\ (8.2)}} \mathbf{1}_{\mathcal{A}}(\ell_2) \ll \frac{1}{|\Delta|} (\theta^{18}\sqrt{MN})^{\gamma_0} + \frac{1}{|\Delta|} (\theta^{18}\sqrt{MN})^{\gamma_0} \sum_{r=1}^{|\Delta|-1} F_X \left(\frac{r}{|\Delta|} \right),$$

where X is a power of 10 with $X \asymp Y$.

The contribution $D'_{3,0}(M, N)$ to $D'_3(M, N)$ coming from the first term here is

$$D'_{3,0}(M, N) \ll \theta^{18\gamma_0}(MN)^{\gamma_0} \sum_{0 < |\Delta| \leq \theta^{18}N} \frac{1}{|\Delta|} \sum_{N < |z_1|^2 \leq (1+\theta)N} \sum_{\substack{N < |z_2|^2 \leq (1+\theta)N \\ \text{Im}(\bar{z}_1 z_2) = \Delta}} 1.$$

Let $z_1 = r + is$ with $(r, s) = 1$. Since $r^2 + s^2 > N$ this implies $rs \neq 0$. Let $z_2 = u + iv$, and note that $\text{Im}(\bar{z}_1 z_2) = rv - su$. Let (u_0, v_0) be a pair such that $rv_0 - su_0 = \Delta$. Then for any other pair (u_1, v_1) such that $rv_1 - su_1 = \Delta$, we have

$$r(v_1 - v_0) - s(u_1 - u_0) = 0.$$

Since r and s are coprime, we see that $v_1 - v_0 = ks$ for some integer k , and $u_1 - u_0 = \ell r$ for some integer ℓ . As $rs \neq 0$ we find that $k = \ell$, and thus $u_1 + iv_1 = u_0 + iv_0 + kz_1$. Since $|z_1| \asymp |z_2| \asymp N^{1/2}$, it follows that the number of choices for z_2 , given Δ and z_1 , is $O(1)$, and therefore

$$D'_{3,0}(M, N) \ll \theta^{18\gamma_0+1}(\log N)(MN)^{\gamma_0}N \ll \theta^{18}(\log N)(MN)^{\gamma_0}N.$$

We now turn to bounding the contribution of the nonzero frequencies $D'_{3,*}(M, N)$. Arguing as with $D'_{3,0}(M, N)$, we deduce that

$$D'_{3,*}(M, N) \ll \theta^{18\gamma_0+1}(MN)^{\gamma_0}N \sum_{d \leq \theta^{18}N} \frac{1}{d} \sum_{r=1}^{d-1} F_X\left(\frac{r}{d}\right).$$

We reduce to primitive fractions and perform dyadic decompositions to obtain

$$D'_{3,*}(M, N) \ll \theta^{18\gamma_0+1}(\log N)^2(MN)^{\gamma_0}N \sup_{Q \ll \theta^{18}N} \frac{1}{Q} \sum_{q \ll Q} \sum_{(b,q)=1} F_X\left(\frac{b}{q}\right).$$

By Lemma 5.4,

$$(8.3) \quad \frac{1}{Q} \sum_q \sum_{(b,q)=1} F_X\left(\frac{b}{q}\right) \ll \frac{1}{Q^{23/77}} + \frac{Q}{X^{50/77}} \ll 1 + \frac{\theta^{18}N}{X^{50/77}}.$$

We wish for the quantity in (8.3) to be $\ll 1$ so it suffices to have $N \ll x^{25/77-\delta}$, which is precisely the case. We deduce that the total contribution from $|\Delta| \leq \theta^{18}N$ is

$$\ll D'_3(M, N) \ll \theta^{18}(\log N)^2(MN)^{\gamma_0}N,$$

as desired. \square

With Lemma 8.2 we have removed those moduli $|\Delta|$ that are substantially smaller than expected, and we now proceed with our task of making b a rational residue class. We saw above that it suffices to impose the condition $(\ell_1, |\Delta|) = 1$. We expect to be able to impose this condition with the cost of only a small error since $(\ell_1, \Pi) = 1$. Indeed, it is for this step alone that we introduced the condition $(\ell, \Pi) = 1$ at the beginning of the proof of Theorem 1.1 (see the comments after (1.2)).

We estimate trivially the contribution from $(\ell_1, |\Delta|) \neq 1$. By the triangle inequality, it suffices to estimate

$$D''_3(M, N) := \sum_{p > P} \sum_{N < |z_1|^2, |z_2|^2 \leq (1+\theta)N} \sum_{\substack{(z_1, z_2)=1 \\ p \nmid |\Delta|}} \sum_{b z_1 \equiv z_2(|\Delta|)} \sum_{\substack{\ell_1 < X \\ p \nmid \ell_1}} \mathbf{1}_{\mathcal{A}}(\ell_1) \sum_{\substack{\ell_2 < X \\ \ell_2 \equiv \text{Re}(b)\ell_1(|\Delta|)}} \mathbf{1}_{\mathcal{A}}(\ell_2),$$

where X is a power of 10 with $X \asymp (MN)^{1/2}$. As has become typical, we introduce characters to detect the congruence on ℓ_2 and then apply the triangle inequality to eliminate the

dependence on b and ℓ_1 . We also apply additive characters to detect the congruence on ℓ_1 , obtaining

$$D_3''(M, N) \ll (MN)^{\gamma_0} \sum_{P < p \ll N} \sum_{\substack{N < |z_1|^2, |z_2|^2 \leq (1+\theta)N \\ (z_1, z_2)=1 \\ p \parallel |\Delta|}} \frac{1}{p|\Delta|} \sum_{k=1}^p F_X\left(\frac{k}{p}\right) \sum_{r=1}^{|\Delta|} F_X\left(\frac{r}{|\Delta|}\right).$$

By Lemma 5.3 we find

$$\frac{1}{p} \sum_{k=1}^p F_X\left(\frac{k}{p}\right) \ll p^{-50/77} + X^{-50/77} \ll p^{-50/77},$$

the last inequality following since $N < M$. Thus

$$\begin{aligned} D_3''(M, N) &\ll (MN)^{\gamma_0} N \sum_{P < p \ll N} p^{-50/77} \sum_{\substack{d \ll N \\ p \mid d}} \frac{1}{d} \sum_{r=1}^d F_X\left(\frac{r}{d}\right) \\ &\ll (MN)^{\gamma_0} N \sum_{d \ll N} \sum_{r=1}^d \frac{1}{d} F_X\left(\frac{r}{d}\right) \sum_{\substack{p \mid d \\ p > P}} p^{-50/77} \\ &\ll (\log N)(MN)^{\gamma_0} NP^{-50/77} \sum_{d \ll N} \frac{1}{d} \sum_{r=1}^d F_X\left(\frac{r}{d}\right). \end{aligned}$$

The contribution of the zero frequency $r = d$ is

$$\ll (\log N)^2 (MN)^{\gamma_0} NP^{-50/77}.$$

For the nonzero frequencies we reduce to primitive fractions, obtaining

$$\begin{aligned} &\ll (\log N)(MN)^{\gamma_0} NP^{-50/77} \sum_{d \ll N} \frac{1}{d} \sum_{\substack{q \mid d \\ q > 1}} \sum_{(a, q)=1} F_X\left(\frac{a}{q}\right) \\ &\ll (\log N)^2 (MN)^{\gamma_0} NP^{-50/77} \sum_{q \ll N} \frac{1}{q} \sum_{(a, q)=1} F_X\left(\frac{a}{q}\right) \\ &\ll (\log N)^3 (MN)^{\gamma_0} NP^{-50/77}. \end{aligned}$$

The last inequality follows by a dyadic decomposition and Lemma 5.4. Therefore

$$D_3''(M, N) \ll (\log N)^3 (MN)^{\gamma_0} NP^{-50/77}.$$

Thus the contribution from those ℓ_1 not coprime to $|\Delta|$ is negligible.

Now that b is a rational residue class, we wish to make b coprime with $|\Delta|$. It suffices to make z_2 coprime to $|\Delta|$ (recall that z_1 is already coprime to $|\Delta|$ since z_1, z_2 are coprime). Since z_2 has no small prime factors this condition is easy to impose. The details are by now familiar so we omit them. The error terms involved are of size

$$\ll (\log N)^3 (MN)^{\gamma_0} NP^{-1}.$$

In order to prove Proposition 7.1, it therefore suffices to study

$$\begin{aligned} \mathcal{D}_4(M, N) := & \sum_{\substack{N < |z_1|^2, |z_2|^2 \leq (1+\theta)N \\ (z_1 \bar{z}_1 z_2 \bar{z}_2, \Pi)=1 \\ (z_1 | \Delta, z_2)=1 \\ |\Delta| > \theta^{18} N}} \sum \mu(|z_1|^2) \mu(|z_2|^2) \sum_{bz_1 \equiv z_2(|\Delta|)} \\ & \times \sum_{\substack{\ell_1 \leq \sqrt{2(1+\theta)}(MN)^{1/2} \\ (\ell_1, \Pi|\Delta)=1}} \mathbf{1}_{\mathcal{A}}(\ell_1) \sum_{\substack{\ell_2 \leq \sqrt{2(1+\theta)}(MN)^{1/2} \\ \ell_2 \equiv b\ell_1(|\Delta|) \\ |\ell_1 z_2 - \ell_2 z_1|^2 \leq 2\Delta^2 M \\ (\ell_2, \Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell_2), \end{aligned}$$

and show

$$\mathcal{D}_4(M, N) \ll \theta^{18} (\log MN)^{O(1)} (MN)^{\gamma_0} N.$$

We observe that \mathcal{D}_4 differs from the sum \mathcal{D}_5 in Proposition 8.1 only in its absence of the condition $\Delta_{10} > \theta^{-28}$ (recall the factorization (8.1)). We show that \mathcal{D}_4 and \mathcal{D}_5 differ from one another by an amount acceptably small for Proposition 7.1. The following result completes the reduction of Proposition 7.1 to Proposition 8.1.

Lemma 8.3. *The contribution to $\mathcal{D}_4(M, N)$ from $\Delta_{10} > \theta^{-28}$ is*

$$\ll \theta^{18} (\log MN)^{O(1)} (MN)^{\gamma_0} N.$$

Proof. The contribution to $\mathcal{D}_4(M, N)$ from $\Delta_{10} > \theta^{-28}$ is bounded above by a constant multiple of

$$\mathcal{D}'_4 := \sum_{\substack{|z_1|^2, |z_2|^2 \ll N \\ (z_1, z_2)=1 \\ \Delta_{10} > \theta^{-28}}} \sum_{\substack{b(|\Delta|) \\ bz_1 \equiv z_2(|\Delta|)}} \sum_{\ell_1 < X} \mathbf{1}_{\mathcal{A}}(\ell_1) \sum_{\substack{\ell_2 < X \\ \ell_2 \equiv b\ell_1(\Delta_{10}) \\ \ell_2 \equiv b\ell_1(|\Delta'|)}} \mathbf{1}_{\mathcal{A}}(\ell_2),$$

where $X \asymp (MN)^{1/2}$ is a power of 10. We apply additive characters to detect the congruences and then apply the triangle inequality, obtaining

$$\sum_{\substack{\ell_2 < X \\ \ell_2 \equiv b\ell_1(\Delta_{10}) \\ \ell_2 \equiv b\ell_1(|\Delta'|)}} \mathbf{1}_{\mathcal{A}}(\ell_2) \ll \frac{(MN)^{\gamma_0/2}}{\Delta_{10} |\Delta'|} \sum_{s=1}^{\Delta_{10}} \sum_{k=1}^{|\Delta'|} F_X \left(\frac{k\ell_2}{|\Delta'|} + \frac{s\ell_2}{\Delta_{10}} \right).$$

The contribution from $k = |\Delta'|$ to \mathcal{D}'_4 is

$$\begin{aligned} & \ll (MN)^{\gamma_0} \sum_{\substack{|z_1|^2, |z_2|^2 \ll N \\ (z_1, z_2)=1 \\ \Delta_{10} > \theta^{-28}}} \sum \frac{1}{\Delta_{10} |\Delta'|} \sum_{s=1}^{\Delta_{10}} F_X \left(\frac{s}{\Delta_{10}} \right) \\ & \ll (MN)^{\gamma_0} N \sum_{\substack{\theta^{-28} < d \ll N \\ d|10^\infty}} \frac{1}{d} \sum_{s=1}^d F_X \left(\frac{s}{d} \right) \sum_{|\Delta'| \ll N} \frac{1}{|\Delta'|} \\ & \ll (\log N) (MN)^{\gamma_0} N \sum_{\substack{d > \theta^{-28} \\ d|10^\infty}} \frac{1}{d^{50/77}}, \end{aligned}$$

the last inequality following by Lemma 5.3. By Rankin's trick and an Euler product computation,

$$\sum_{\substack{d > \theta^{-28} \\ d|10^\infty}} \frac{1}{d^{50/77}} \ll \theta^{28(50/77-1/\log \log N)} \sum_{d|10^\infty} d^{-1/\log \log N} \ll \theta^{18} (\log \log N)^2.$$

Let us now turn to the case in which $1 \leq k \leq |\Delta'| - 1$. The argument is a more elaborate version of the proof of Lemma 5.6. Arguing as in the case $k = |\Delta'|$ and changing variables, this contribution is bounded by

$$\ll (MN)^{\gamma_0} N \sum_{\substack{t > \theta^{-28} \\ t|10^\infty}} \frac{1}{t} \sum_{s=1}^t F_X \left(\frac{s}{t} \right) + (MN)^{\gamma_0} N \sum_{\substack{t \ll N \\ t|10^\infty}} \frac{1}{t} \sum_{s=1}^t \sum_{\substack{1 < e \ll N \\ (e,10)=1}} \frac{1}{e} \sum_{k=1}^{e-1} F_X \left(\frac{k}{e} + \frac{s}{t} \right).$$

The first term here accounts for the possibility that $\Delta_{10} = \Delta$, and contributes at most $O(\theta^{18}(\log N)(MN)^{\gamma_0} N)$. It therefore suffices to bound the second term. Reducing from fractions with denominator e to primitive fractions gives that this second term is bounded by

$$\ll (\log N)(MN)^{\gamma_0} N \sum_{\substack{t \ll N \\ t|10^\infty}} \frac{1}{t} \sum_{s=1}^t \sum_{\substack{1 < q \ll N \\ (q,10)=1}} \frac{1}{q} \sum_{(r,q)=1} F_X \left(\frac{r}{q} + \frac{s}{t} \right).$$

We break the sum over q into $q \leq Q$ and $q > Q$, where $Q := \exp(\varepsilon \sqrt{\log N})$ with $\varepsilon > 0$ sufficiently small constant. We first handle $q > Q$. Taking the supremum over s and t , the contribution from $q > Q$ is

$$\ll (\log N)^3 \sup_{\beta \in \mathbb{R}} \sum_{\substack{Q < q \ll N \\ (q,10)=1}} \frac{1}{q} \sum_{(r,q)=1} F_X \left(\frac{r}{q} + \beta \right).$$

We break the range of q into dyadic segments and apply Lemma 5.4, which gives that the contribution from $q > Q$ is

$$\ll \frac{(\log N)^4}{Q^{23/77}}.$$

Now we turn to $q \leq Q$. We first show that the contribution from $t > Q^4$, say, is negligible. Interchanging the order of summation,

$$\begin{aligned} & \sum_{\substack{1 < q \leq Q \\ (q,10)=1}} \frac{1}{q} \sum_{(r,q)=1} \sum_{\substack{t > Q^4 \\ t|10^\infty}} \frac{1}{t} \sum_{r=1}^t F_X \left(\frac{s}{t} + \frac{r}{q} \right) \ll Q \sup_{\beta \in \mathbb{R}} \sum_{\substack{t > Q^4 \\ t|10^\infty}} \frac{1}{t} \sum_{r=1}^t F_X \left(\frac{s}{t} + \beta \right) \\ & \ll Q \sum_{\substack{t > Q^4 \\ t|10^\infty}} \frac{1}{t^{50/77}}, \end{aligned}$$

the last inequality following from Lemma 5.3. By Rankin's trick, we obtain the bound

$$Q \sum_{\substack{t > Q^4 \\ t|10^\infty}} \frac{1}{t^{50/77}} \ll \frac{Q}{(Q^4)^{1/2}} \sum_{t|10^\infty} \frac{1}{t^{50/77-1/2}} \ll \frac{1}{Q}.$$

It therefore suffices to bound

$$\sum_{\substack{t \leq Q^4 \\ t|10^\infty}} \frac{1}{t} \sum_{s=1}^t \sum_{\substack{1 < q \leq Q \\ (q,10)=1}} \frac{1}{q} \sum_{(r,q)=1} F_X \left(\frac{r}{q} + \frac{s}{t} \right).$$

At this point we avail ourselves of the product formula (5.2) for F . We take U to be a power of 10 such that t divides U for every $t | 10^\infty$ with $t \leq Q^4$, and set $V = X/U$. We may take $U \asymp Q^{14}$, since any $t | 10^\infty$ with $t \leq Q^4$ may be written as $t = 2^a 5^b$ where $a, b \leq \frac{4}{\log 2} \log Q$, and $\frac{4 \log 10}{\log 2} < 14$. Since F is 1-periodic we obtain

$$F_X \left(\frac{r}{q} + \frac{s}{t} \right) = F_U \left(\frac{r}{q} + \frac{s}{t} \right) F_V \left(\frac{Ur}{q} + \frac{Us}{t} \right) = F_U \left(\frac{r}{q} + \frac{s}{t} \right) F_V \left(\frac{Ur}{q} \right) \ll F_V \left(\frac{Ur}{q} \right).$$

Observe that V and X are asymptotically equal in the logarithmic scale since $Q = N^{o(1)}$. We then apply Lemma 5.2 to bound each F_V individually, and find

$$(8.4) \quad \mathcal{D}'_4 \ll (\log N)^4 (MN)^{\gamma_0} N \left(\theta^{18} + \exp(-c\sqrt{\log MN}) \right),$$

for some positive constant c . This completes the proof. \square

9. POLAR BOXES AND CONGRUENCE EXERCISES

In this section we remove the congruence condition modulo $|\Delta'|$ in \mathcal{D}_5 , the sum appearing in the statement of Proposition 8.1. Not surprisingly, there are several further technical barriers to overcome before this can be accomplished. For instance, the condition

$$|\ell_1 z_2 - \ell_2 z_1|^2 \leq 2\Delta^2 M$$

entangles the four variables z_1, z_2, ℓ_1 , and ℓ_2 in an undesirable way. We put z_1 and z_2 into polar boxes in order to reduce some of this dependence. After restricting to “generic” boxes and removing as much z_1 and z_2 dependence as we can, we break the sum over ℓ_2 into short intervals in preparation for applying additive characters. We employ the fundamental lemma to handle the condition $(\ell_2, \Pi) = 1$. The error term is estimated as we have done before, using distribution results for F_Y . With the congruence condition modulo $|\Delta'|$ removed, we can make some simplifications and adjustments in the main term. The last major task is then to get cancellation from the Möbius function in the main term, which we do in Section 10.

Before we state our next proposition, we need to introduce some notation. We break the sums over z_1 and z_2 in \mathcal{D}_5 into polar boxes. We now introduce a new quantity $\lambda := k_0^{-1} \asymp \theta^{153}$, for some $k_0 \in \mathbb{N}$ of size θ^{-153} . We split up the summation over z_1 and z_2 into polar boxes of the form

$$\mathfrak{B}_{i,j} = \left\{ w \in \mathbb{C} : R_i < |w|^2 \leq (1 + \lambda)R_i, \delta_j \leq \arg(w) \leq \delta_j + 2\pi\lambda \right\},$$

where $R_i = (1 + \lambda)^i N$. Note that $N < R_i \leq (1 + \theta)N$ and $\delta_j = 2\pi j\lambda$ for $0 \leq j \leq \lambda^{-1} - 1$ an integer. For such a polar box, let $z(\mathfrak{B}_{i,j}) := R_i e^{i\delta_j}$ (there should be no confusing the index i and the imaginary unit i). Observe that

$$(9.1) \quad z = z(\mathfrak{B}_{i,j})(1 + O(\lambda)).$$

The number of such polar boxes is $O(\lambda^{-4})$, and for each polar box $\mathcal{B}_{i,j}$ we have the trivial bound

$$\sum_{z \in \mathcal{B}_{i,j}} 1 \ll \lambda^2 N.$$

To ease the burden on the subscripts, we abuse notation and simply write $z_1 \in \mathfrak{B}_1, z_2 \in \mathfrak{B}_2$. We also set $\Delta(\mathfrak{B}_1, \mathfrak{B}_2) := \Delta(z(\mathfrak{B}_1), z(\mathfrak{B}_2))$.

Proposition 9.1. *Let $\delta > 0$ be sufficiently small and fixed. For fixed $L > 0$ set $\theta = (\log x)^{-L}$. Assume the notation above. For*

$$\begin{aligned} \mathcal{D}_2(\mathfrak{B}_1, \mathfrak{B}_2) := & \sum_{\substack{z_1 \in \mathfrak{B}_1, z_2 \in \mathfrak{B}_2 \\ (z_1 \bar{z}_1 z_2 \bar{z}_2, \Pi)=1 \\ \Delta_{10} \leq \theta^{-28}}} \mu(|z_1|^2) \mu(|z_2|^2) \Delta_{10} \frac{10|\Delta|}{\varphi(10|\Delta|)} \\ & \times \sum_{\substack{\ell_1, \ell_2 \leq \sqrt{2(1+\theta)}(MN)^{1/2} \\ \ell_2 z_1 \equiv \ell_1 z_2 \pmod{\Delta_{10}} \\ | \ell_1 z_2 - \ell_2 z_1 |^2 \leq 2\Delta(\mathfrak{B}_1, \mathfrak{B}_2)^2 M \\ (\ell_1, \Pi)=1, (\ell_2, 10)=1}} \mathbf{1}_{\mathcal{A}}(\ell_1) \mathbf{1}_{\mathcal{A}}(\ell_2) \end{aligned}$$

we have

$$\mathcal{D}_2(\mathfrak{B}_1, \mathfrak{B}_2) \ll_{\delta} \theta^{36} \lambda^4 (\log MN)^{O(1)} (MN)^{\gamma_0} N^2$$

uniformly in

$$N \gg x^{\delta}, \quad MN \ll x.$$

The implied constant is ineffective.

Observe that the congruence modulo $|\Delta|$ in Proposition 8.1 has been replaced by a congruence modulo Δ_{10} in Proposition 9.1. This congruence condition is much more mild since $\Delta_{10} \ll (\log MN)^{O(1)}$.

We use the remainder of this section to show that Proposition 8.1 follows from Proposition 9.1.

From the lower bound for $|\Delta|$, we see the polar boxes $\mathfrak{B}_1, \mathfrak{B}_2$ cannot be too close to one another, in a sense. Writing $z_i = r_i e^{i\theta_i}$, we see

$$|\Delta(z_1, z_2)| = r_1 r_2 |\sin(\theta_2 - \theta_1)| > \theta^{18} N,$$

after using the fact that $e^{i\theta} = \cos \theta + i \sin \theta$. Since $r_1, r_2 \asymp N^{1/2}$, we have

$$|\sin(\theta_2 - \theta_1)| \gg \theta^{18}.$$

Recall that $\theta_i = \delta_i + O(\lambda)$. Since $\lambda \leq \theta^{19}$, the sine angle addition formula and the triangle inequality imply

$$|\sin(\delta_2 - \delta_1)| \gg \theta^{18}.$$

Thus the angles δ_1, δ_2 cannot be too close to each other. Given this fact, we may show in the same manner that

$$(9.2) \quad \Delta(z_1, z_2) = (1 + O(\lambda')) \Delta(\mathfrak{B}_1, \mathfrak{B}_2),$$

where $\lambda' := \theta^{-18} \lambda$.

We claim it suffices to sum over polar boxes $\mathfrak{B}_1, \mathfrak{B}_2$ such that $|\Delta(\mathfrak{B}_1, \mathfrak{B}_2)| > (1 + \lambda')\theta^{18}N$. Indeed, the contribution from all polar boxes not satisfying this condition is bounded by

$$(9.3) \quad \sum_{\substack{N < |z_1|^2, |z_2|^2 \leq 2N \\ (z_1, z_2) = 1 \\ (1 - C\lambda')\theta^{18}N < |\Delta| \leq (1 + C\lambda')\theta^{18}N \\ \Delta_{10} \leq \theta^{-28}}} \sum_{\substack{\ell_1, \ell_2 \ll (MN)^{1/2} \\ \ell_1 z_2 \equiv \ell_2 z_1 \pmod{|\Delta|}}} \mathbf{1}_{\mathcal{A}}(\ell_1) \mathbf{1}_{\mathcal{A}}(\ell_2)$$

for some absolute constant $C > 0$. We bound this contribution using the following lemma.

Lemma 9.2. *Let $\exp(-(\log MN)^{1/3}) < \eta < \frac{1}{2}$, and let $C > 0$ be an absolute constant. Then*

$$\mathcal{D}'_5(M, N) := \sum_{\substack{N < |z_1|^2, |z_2|^2 \leq 2N \\ (z_1, z_2) = 1 \\ (1 - C\eta)\theta^{18}N < |\Delta| \leq (1 + C\eta)\theta^{18}N \\ \Delta_{10} \leq \theta^{-28}}} \sum_{\substack{\ell_1, \ell_2 \ll (MN)^{1/2} \\ \ell_1 z_2 \equiv \ell_2 z_1 \pmod{|\Delta|}}} \mathbf{1}_{\mathcal{A}}(\ell_1) \mathbf{1}_{\mathcal{A}}(\ell_2) \ll \eta \theta^{-28} (MN)^{\gamma_0} N.$$

Proof. We begin by handling the congruence condition as we did in imposing the condition $\Delta_{10} \leq \theta^{-28}$ in Lemma 8.3. We detect the congruences modulo Δ_{10} and $|\Delta'|$ with additive characters. The nonzero frequencies modulo $|\Delta'|$ contribute an error term of size $O((MN)^{\gamma_0} N \exp(-c\sqrt{\log MN}))$, by the argument that led to (8.4). For the zero frequency modulo $|\Delta'|$ we apply orthogonality of additive characters to reintroduce the congruence modulo Δ_{10} . We find

$$\begin{aligned} \mathcal{D}'_5(M, N) &= \sum_{\substack{N < |z_1|^2, |z_2|^2 \leq 2N \\ (z_1, z_2) = 1 \\ (1 - C\delta)\theta^{18}N < |\Delta| \leq (1 + C\delta)\theta^{18}N \\ \Delta_{10} \leq \theta^{-28}}} \frac{1}{|\Delta'|} \sum_{\substack{\ell_1, \ell_2 \ll (MN)^{1/2} \\ \ell_1 z_2 \equiv \ell_2 z_1 \pmod{\Delta_{10}}}} \mathbf{1}_{\mathcal{A}}(\ell_1) \mathbf{1}_{\mathcal{A}}(\ell_2) \\ &\quad + O((MN)^{\gamma_0} N \exp(-c\sqrt{\log MN})). \end{aligned}$$

Since

$$\theta^{18}N \ll |\Delta| = \Delta_{10}|\Delta'|,$$

we see that

$$|\Delta'|^{-1} \ll \theta^{-28} J^{-1},$$

where $J = \theta^{18}N$. Dropping the congruence condition modulo Δ_{10} , it follows that

$$\begin{aligned} \mathcal{D}'_5(M, N) &\ll (MN)^{\gamma_0} N \exp(-c\sqrt{\log MN}) + \frac{\theta^{-28}}{J} (MN)^{\gamma_0} \sum_{\substack{N < |z_1|^2, |z_2|^2 \leq 2N \\ (1 - C\eta)J < |\Delta| \leq (1 + C\eta)J}} 1 \\ &\ll (MN)^{\gamma_0} N \exp(-c\sqrt{\log MN}) \\ &\quad + \frac{\theta^{-28}}{J} (MN)^{\gamma_0} \sum_{\substack{N < r^2 + s^2 \leq 2N \\ (r, s) = 1}} \sum_{\substack{u, v \ll N^{1/2} \\ (1 - C\eta)J < |rv - su| \leq (1 + C\eta)J}} 1. \end{aligned}$$

Observe that the conditions on r and s imply $rs \neq 0$. Given r, s , and u the number of v is $\ll \eta J/|r|$, and given r, s , and v the number of u is $\ll \eta J/|s|$. Since $\max(|r|, |s|) \gg N^{1/2}$, we

see that

$$\sum_{\substack{u,v \ll N^{1/2} \\ (1-C\delta)J < |rv-su| \leq (1+C\delta)J}} 1 \ll \eta J.$$

Summing over r and s then completes the proof. \square

By Lemma 9.2 the contribution of (9.3) is $\ll \theta^{-28} \lambda'(MN)^{\gamma_0} N$. This bound is acceptably small for Proposition 8.1 since $\lambda \leq \theta^{64}$.

The number of boxes intersecting the boundary of $\{z : N < |z|^2 \leq (1+\theta)N\}$ is $O(\lambda^{-2})$. Handling the congruences modulo $|\Delta'|$ and Δ_{10} as in Lemma 9.2, we find the error made by this approximation is

$$\ll \theta^{-46} \lambda^2(MN)^{\gamma_0} N,$$

and this error is acceptable for Proposition 8.1. We therefore have

$$\mathcal{D}_5(M, N) = O(\theta^{18}(MN)^{\gamma_0} N) + \sum_{\substack{\mathfrak{B}_1, \mathfrak{B}_2 \\ |\Delta(\mathfrak{B}_1, \mathfrak{B}_2)| > (1+\lambda')\theta^{18}N}} \mathcal{D}_1(\mathfrak{B}_1, \mathfrak{B}_2),$$

where

$$\begin{aligned} \mathcal{D}_1(\mathfrak{B}_1, \mathfrak{B}_2) := & \sum_{\substack{z_1 \in \mathfrak{B}_1, z_2 \in \mathfrak{B}_2 \\ (z_1 \overline{z_1} z_2 \overline{z_2}, \Pi) = 1 \\ (z_1 | \Delta, z_2) = 1 \\ \Delta_{10} \leq \theta^{-28}}} \mu(|z_1|^2) \mu(|z_2|^2) \sum_{bz_1 \equiv z_2(|\Delta|)} \\ & \times \sum_{\substack{\ell_1 \leq \sqrt{2(1+\theta)}(MN)^{1/2} \\ (\ell_1, \Pi|\Delta|) = 1}} \mathbf{1}_{\mathcal{A}}(\ell_1) \sum_{\substack{\ell_2 \leq \sqrt{2(1+\theta)}(MN)^{1/2} \\ \ell_2 \equiv b\ell_1(|\Delta|) \\ |\ell_1 z_2 - \ell_2 z_1|^2 \leq 2\Delta^2 M \\ (\ell_2, \Pi) = 1}} \mathbf{1}_{\mathcal{A}}(\ell_2). \end{aligned}$$

Observe that $\mathcal{D}_1(\mathfrak{B}_1, \mathfrak{B}_2)$ depends on M and N , but we have suppressed this in the notation. In order to prove Proposition 8.1 it therefore suffices to show that

$$(9.4) \quad \mathcal{D}_1(\mathfrak{B}_1, \mathfrak{B}_2) \ll \theta^{18} (\log MN)^{O(1)} \lambda^4 (MN)^{\gamma_0} N$$

uniformly in \mathfrak{B}_1 and \mathfrak{B}_2 .

We now work to make the condition $|\ell_1 z_2 - \ell_2 z_1|^2 \leq 2\Delta^2 M$ in $\mathcal{D}_1(\mathfrak{B}_1, \mathfrak{B}_2)$ less dependent on z_1 and z_2 . We can rearrange to get the condition

$$(9.5) \quad \left| \ell_2 - \frac{\ell_1 z_2}{z_1} \right| \leq \sqrt{2} \frac{|\Delta(z_1, z_2)| M^{1/2}}{|z_1|}.$$

We wish to replace (9.5) by

$$(9.6) \quad \left| \ell_2 - \frac{\ell_1 z_2}{z_1} \right| \leq \sqrt{2} \frac{|\Delta(\mathfrak{B}_1, \mathfrak{B}_2)| M^{1/2}}{|z(\mathfrak{B}_1)|}.$$

Since

$$\frac{|\Delta(z_1, z_2)| M^{1/2}}{|z_1|} = (1 + O(\lambda')) \frac{|\Delta(\mathfrak{B}_1, \mathfrak{B}_2)| M^{1/2}}{|z(\mathfrak{B}_1)|}$$

(recall (9.1) and (9.2)), we see it suffices to bound the contribution from those ℓ_2 in $\mathcal{D}_1(\mathfrak{B}_1, \mathfrak{B}_2)$ that satisfy

$$(9.7) \quad (1 - C\lambda')K \leq \left| \ell_2 - \frac{\ell_1 z_2}{z_1} \right| \leq (1 + C\lambda')K,$$

where

$$K := \sqrt{2} \frac{|\Delta(\mathfrak{B}_1, \mathfrak{B}_2)| M^{1/2}}{|z(\mathfrak{B}_1)|}$$

and $C > 0$ is a sufficiently large absolute constant.

We claim that (9.7) places ℓ_2 in a bounded number of intervals of length $\ll (\lambda')^{1/2}K$. It is important to note that the number of intervals will be bounded independently of ℓ_1, z_1, z_2 , but the intervals themselves still depend on ℓ_1, z_1 , and z_2 . For notational simplicity, write $A = (1 - C\lambda')K$ and $B = (1 + C\lambda')K$. Then (9.7) gives

$$A \leq |\ell_2 - (u + iv)| \leq B$$

for some real numbers u, v . Since ℓ_2 is real, we obtain by squaring and rearranging

$$A^2 - v^2 \leq (\ell_2 - u)^2 \leq B^2 - v^2.$$

Clearly, if $|v| \geq B$ then there is at most one choice for ℓ_2 ; the claim then follows immediately. We may therefore assume $|v| < B$. There are then two cases to consider: $|v| \geq A$ and $|v| < A$. If $|v| \geq A$ then $A^2 - v^2 \leq 0$, and the lower bound on $(\ell_2 - u)^2$ is automatically satisfied. We therefore obtain

$$|\ell_2 - u| \leq \sqrt{B^2 - v^2} \leq \sqrt{B^2 - A^2} \ll (\lambda')^{1/2}K.$$

Now suppose that $|v| < A$. Then

$$\sqrt{A^2 - v^2} \leq |\ell_2 - u| \leq \sqrt{B^2 - v^2},$$

and hence ℓ_2 is in two intervals of length $\leq \sqrt{B^2 - v^2} - \sqrt{A^2 - v^2} + 2$, say. We then have

$$\sqrt{B^2 - v^2} - \sqrt{A^2 - v^2} = \frac{B^2 - A^2}{\sqrt{B^2 - v^2} + \sqrt{A^2 - v^2}} \leq \frac{B^2 - A^2}{\sqrt{B^2 - v^2}} \leq \sqrt{B^2 - A^2},$$

and this completes the proof of the claim.

We now bound the contribution of those ℓ_2 satisfying (9.7). At this point we should have enough experience to see how we proceed. We let Y be the largest power of 10 satisfying $Y \leq (\lambda')^{1/2}K$, and cover the intervals (9.7) with subintervals of the form $[nY, nY + Y)$, $n \geq 0$ an integer. Observe that Y is independent of ℓ_1, z_1, z_2 , but the integers n are not. The number of integers n is $O(1)$. We can reduce to summing the indicator function $\mathbf{1}_{\mathcal{A}}(t)$ over $0 \leq t < Y'$, $Y' \asymp Y$, and then deal with the congruence modulo $|\Delta|$ by considering it as a congruence modulo $|\Delta'|$ and Δ_{10} . As usual, we use the triangle inequality to eliminate any dependence on n . We obtain a bound of

$$\ll \theta^{-46} (\lambda')^{\gamma_0/2} \lambda^4 (MN)^{\gamma_0} N,$$

and this is acceptable for (9.4) since $\lambda \asymp \theta^{153}$.

We now have that ℓ_2 in $\mathcal{D}_1(\mathfrak{B}_1, \mathfrak{B}_2)$ must satisfy the conditions

$$(9.8) \quad \begin{aligned} \ell_2 &\leq \sqrt{2(1+\theta)}(MN)^{1/2}, \\ \left| \ell_2 - \frac{\ell_1 z_2}{z_1} \right| &\leq \sqrt{2} \frac{|\Delta(\mathfrak{B}_1, \mathfrak{B}_2)| M^{1/2}}{|z(\mathfrak{B}_1)|}, \\ \ell_2 &\equiv b\ell_1(|\Delta|). \end{aligned}$$

Recall that the congruence in (9.8) is a congruence of rational integers. To handle the first two conditions we perform a short interval decomposition. Let Y be the largest power of 10 which satisfies

$$Y \leq \lambda \frac{|\Delta(\mathfrak{B}_1, \mathfrak{B}_2)| M^{1/2}}{|z(\mathfrak{B}_1)|}.$$

We cover the interval $\ell_2 \leq \sqrt{2(1+\theta)}(MN)^{1/2}$ with subintervals of the form $[nY, nY + Y)$, as we have done many times before. For the subintervals that intersect the boundary of the second condition of (9.8) we obtain a contribution of $\ll (\log MN)^{O(1)} \lambda^{4+\gamma_0/2} \theta^{-46} (MN)^{\gamma_0} N$, which is acceptable. The sum over ℓ_2 has therefore become

$$\sum_{\substack{n \in \mathbb{Z} \\ n \in S(\ell_1, z_1, z_2)}} \sum_{\substack{nY < \ell_2 \leq nY + Y \\ \ell_2 \equiv b\ell_1(|\Delta|) \\ (\ell_2, \Pi) = 1}} \mathbf{1}_{\mathcal{A}}(\ell_2),$$

for some set $S(\ell_1, z_1, z_2)$ of size $O(\lambda^{-1})$.

We handle the condition $(\ell_2, \Pi) = 1$ using the fundamental lemma. Let

$$\Sigma := \sum_{z_i \in \mathfrak{B}_i} \sum_{b(|\Delta|)} \sum_{\ell_1} \sum_n \sum_{\ell_2}$$

be the sum we wish to bound (up to acceptable errors, Σ is $\mathcal{D}_1(\mathfrak{B}_1, \mathfrak{B}_2)$ with the condition (9.5) replaced by (9.6)). We partition Σ as

$$\Sigma = \Sigma_+ + \Sigma_-,$$

where in Σ_+ we sum over those z_1, z_2 such that $\mu(|z_1|^2)\mu(|z_2|^2) = +1$, and in Σ_- we sum over those z_1, z_2 such that $\mu(|z_1|^2)\mu(|z_2|^2) = -1$. We get an upper bound on Σ_+ using an upper-bound linear sieve of level D

$$\mathbf{1}_{(\ell_2, \Pi)=1} \leq \mathbf{1}_{(\ell_2, 10)=1} \sum_{\substack{d \leq D \\ d|\Pi/10 \\ d|\ell_2}} \lambda_d^+,$$

and a lower bound on Σ_- using a lower-bound linear sieve of level D

$$\mathbf{1}_{(\ell_2, \Pi)=1} \geq \mathbf{1}_{(\ell_2, 10)=1} \sum_{\substack{d \leq D \\ d|\Pi/10 \\ d|\ell_2}} \lambda_d^-,$$

where D is chosen shortly (see (9.10)). This yields an upper bound on Σ . Reversing λ^+ and λ^- we get a lower bound on Σ , and we show that these bounds for Σ are asymptotically the

same. Thus, for some sign $\varepsilon \in \{\pm 1\}$, it suffices to study

$$(9.9) \quad \Sigma_\varepsilon^\pm := \sum_{\substack{z_i \in \mathfrak{B}_i \\ \mu(|z_1|^2)\mu(|z_2|^2)=\varepsilon}} \sum_{\substack{d \leq D \\ d|\Pi \\ (d,10|\Delta|=1)}} \lambda_d^\pm \sum_{b(|\Delta|)} \sum_{\ell_1} \mathbf{1}_A(\ell_1) \sum_n \sum_{\substack{nY < \ell_2 \leq nY+Y \\ \ell_2 \equiv b\ell_1(|\Delta|) \\ \ell_2 \equiv 0(d) \\ (\ell_2,10)=1}} \mathbf{1}_A(\ell_2).$$

Observe that we have suppressed several conditions in the notation, but these conditions are not to be forgotten.

We write the congruence modulo $|\Delta|$ as two congruences modulo Δ_{10} and $|\Delta'|$, and then use the Chinese remainder theorem to combine the congruences modulo d and $|\Delta'|$ into a congruence modulo $d|\Delta'|$. Considering separately the cases $a_0 \neq 0$ and $a_0 = 0$ and then applying inclusion-exclusion if necessary, we can reduce to having the sum over ℓ_2 be a sum over $0 \leq t < Y'$, where $Y' = Y$ or $Y' = Y/10$. Applying additive characters, the sum over ℓ_2 in (9.9) becomes a linear combination of a uniformly bounded number of quantities of the form

$$\frac{1}{d|\Delta'|} \sum_{f=1}^{d|\Delta'|} e\left(\frac{-f\nu}{d|\Delta'|}\right) \sum_{\substack{t < Y' \\ t+nY \equiv b\ell_1(\Delta_{10}) \\ (t,10)=1}} \mathbf{1}_A(t) e\left(\frac{ft}{d|\Delta'|}\right),$$

where $\nu = \nu(z_1, z_2, \ell_1, n)$ is some residue class. The term $f = d|\Delta'|$ supplies the main term, which we discuss later. For now we turn our attention to the error term $\Sigma_{\varepsilon,E}^\pm$, which is the contribution of $1 \leq f \leq d|\Delta'| - 1$ to Σ_ε^\pm . The argument is essentially that which gave (8.4) in Lemma 8.3.

We apply additive characters to detect the congruence modulo Δ_{10} , apply Möbius inversion to trade the condition $(t, 10) = 1$ for congruence conditions, and then apply additive characters again to detect these latter congruence conditions. We then apply the triangle inequality to eliminate the dependencies on ℓ_1, b , and n . We obtain

$$\begin{aligned} \Sigma_{\varepsilon,E}^\pm &\ll \lambda^{-1}(MN)^{\gamma_0} \sum_{\substack{d \leq D \\ (d,10)=1}} \sum_{\substack{|\Delta'| \leq N \\ (|\Delta'|,10)=1 \\ d|\Delta'| > 1}} \frac{1}{d|\Delta'|} \sum_{f=1}^{d|\Delta'|-1} \sum_{\substack{\Delta_{10} \leq \theta^{-28} \\ \Delta_{10} | 10^\infty}} \frac{1}{\Delta_{10}} \sum_{g=1}^{\Delta_{10}} \sum_{h|10}^h \sum_{k=1}^h \\ &\times F_{Y'} \left(\frac{f}{d|\Delta'|} + \frac{g}{\Delta_{10}} + \frac{k}{h} \right) \sum_{z_1 \in \mathfrak{B}_1} \sum_{\substack{z_2 \in \mathfrak{B}_2 \\ \text{Im}(z_1 z_2) = \Delta_{10} \Delta'}} 1 \\ &\ll \lambda(MN)^{\gamma_0} N \sum_{\substack{1 < d \leq DN \\ (d,10)=1}} \frac{\tau(d)}{d} \sum_{f=1}^{d-1} \sum_{\substack{\Delta_{10} \leq \theta^{-28} \\ \Delta_{10} | 10^\infty}} \frac{1}{\Delta_{10}} \sum_{g=1}^{\Delta_{10}} \sum_{h|10}^h \sum_{k=1}^h F_{Y'} \left(\frac{f}{d} + \frac{g}{\Delta_{10}} + \frac{k}{h} \right). \end{aligned}$$

The second inequality follows, among other things, by changing variables $d|\Delta'| \rightarrow d$. We reduce to primitive fractions to obtain

$$\Sigma_{\varepsilon,E}^\pm \ll (\log N)^2 \lambda(MN)^{\gamma_0} N \sum_{\substack{1 < q \leq DN \\ (q,10)=1}} \frac{\tau(q)}{q} \sum_{\substack{r=1 \\ (r,q)=1}}^q \sum_{\substack{\Delta_{10} \leq \theta^{-28} \\ \Delta_{10} | 10^\infty}} \frac{1}{\Delta_{10}} \sum_{g=1}^{\Delta_{10}} \sum_{h|10}^h \sum_{k=1}^h F_{Y'} \left(\frac{r}{q} + \frac{g}{\Delta_{10}} + \frac{k}{h} \right).$$

We now choose

$$(9.10) \quad D := x^{1/\log \log x},$$

so that $DN \leq x^{25/77-\delta+o(1)}$.

We introduce a parameter $\mathcal{Q} := \exp((\log MN)^{1/3})$. Let us first bound the contribution from $q > \mathcal{Q}$. Taking the supremum over $\frac{g}{\Delta_{10}} + \frac{k}{h}$, we see the contribution from $q > \mathcal{Q}$ is bounded by

$$\ll (\log N)^3 \lambda(MN)^{\gamma_0} N \sup_{\beta \in \mathbb{R}} \sum_{\substack{\mathcal{Q} < q \leq DN \\ (q,10)=1}} \frac{\tau(q)}{q} \sum_{(r,q)=1} F_{Y'} \left(\frac{r}{q} + \beta \right).$$

We split the summation over q into dyadic intervals, apply the divisor bound $\tau(q) \ll q^{o(1)}$, and then apply Lemma 5.4 to obtain

$$\begin{aligned} & \sup_{\beta \in \mathbb{R}} \sum_{\substack{\mathcal{Q} < q \leq DN \\ (q,10)=1}} \frac{\tau(q)}{q} \sum_{(r,q)=1} F_{Y'} \left(\frac{r}{q} + \beta \right) \\ & \ll (\log N) \sup_{\mathcal{Q} < Q \leq DN} \frac{1}{Q^{1-o(1)}} \sup_{\beta \in \mathbb{R}} \sum_{q \leq Q} \sum_{(r,q)=1} F_{Y'} \left(\frac{r}{q} + \beta \right) \\ & \ll (\log N) \sup_{\mathcal{Q} < Q \leq DN} \frac{1}{Q^{1-o(1)}} \left(Q^{54/77} + \frac{Q^2}{(Y')^{50/77}} \right) \\ & \ll (\log N) (\mathcal{Q}^{-1/4} + x^{-2\delta+o(1)}) \\ & \ll \mathcal{Q}^{-1/5}. \end{aligned}$$

The contribution from $q > \mathcal{Q}$ is therefore $O((MN)^{\gamma_0} N \mathcal{Q}^{-1/6})$, and this is acceptably small for Proposition 8.1.

Let us turn to bounding the contribution from $q \leq \mathcal{Q}$. We shall use the product formula (5.2) to eliminate the presence of $\frac{g}{\Delta_{10}} + \frac{k}{h}$. Let U and V be integral powers of 10 with $UV = Y'$ and $U \asymp \theta^{-94}$, and note that since $\Delta_{10} \leq \theta^{-28}$ we have $\Delta_{10} \mid U$. Therefore

$$\begin{aligned} F_{Y'} \left(\frac{r}{q} + \frac{g}{\Delta_{10}} + \frac{k}{h} \right) &= F_U \left(\frac{r}{q} + \frac{g}{\Delta_{10}} + \frac{k}{h} \right) F_V \left(\frac{Ur}{q} + \frac{Ug}{\Delta_{10}} + \frac{Uk}{h} \right) \\ &= F_U \left(\frac{r}{q} + \frac{g}{\Delta_{10}} + \frac{k}{h} \right) F_V \left(\frac{Ur}{q} \right) \ll F_V \left(\frac{Ur}{q} \right). \end{aligned}$$

The contribution from $q \leq \mathcal{Q}$ is then

$$\ll (\log N)^3 (MN)^{\gamma_0} N \sum_{\substack{1 < q \leq \mathcal{Q} \\ (q,10)=1}} \frac{\tau(q)}{q} \sum_{(r,q)=1} F_V \left(\frac{Ur}{q} \right).$$

By Lemma 5.2 we have $F_V \left(\frac{Ur}{q} \right) \ll \exp(-\sqrt{\log MN})$, and therefore the contribution from $q \leq \mathcal{Q}$ is

$$\ll (MN)^{\gamma_0} N \exp(-\tfrac{1}{2}\sqrt{\log MN}),$$

which is acceptably small for Proposition 8.1.

(The upper bound $N \leq x^{25/77-\delta}$ has been required to bound the error terms arising from studying $\mathbf{1}_{\mathcal{A}}(\ell)$ in arithmetic progressions. With the congruence modulo $|\Delta'|$ removed we no longer have the need for this condition, hence the less restrictive conditions on N in Proposition 9.1 compared to Proposition 8.1.)

Let us now turn to the main term we alluded to above. We reverse the transition from ℓ_2 to t , and then undo our short interval decomposition. Up to error terms of size $\ll (\log MN)^{O(1)} \theta^{18} \lambda^4 (MN)^{\gamma_0} N$, the main term of $\Sigma_{\varepsilon}^{\pm}$ is then given by

$$\Sigma_{\varepsilon,0}^{\pm} := \sum_{\substack{z_i \in \mathfrak{B}_i \\ \mu(|z_1|^2) \mu(|z_2|^2) = \varepsilon}} \frac{1}{|\Delta'|} \sum_{\substack{d \leq D \\ d|\Pi \\ (d, 10|\Delta|)=1}} \frac{\lambda_d^{\pm}}{d} \sum_{\substack{\ell_1 \leq \sqrt{2(1+\theta)}(MN)^{1/2} \\ (\ell, \Pi|\Delta|)=1}} \mathbf{1}_{\mathcal{A}}(\ell_1) \sum_{\substack{\ell_2 \leq \sqrt{2(1+\theta)}(MN)^{1/2} \\ \ell_2 z_1 \equiv \ell_1 z_2 \pmod{\Delta_{10}} \\ (9.6) \\ (\ell_2, 10)=1}} \mathbf{1}_{\mathcal{A}}(\ell_2).$$

From the fundamental lemma of sieve theory (see (5.5), for example) we have

$$(9.11) \quad \sum_{\substack{d \leq D \\ d|\Pi \\ (d, 10|\Delta|)=1}} \frac{\lambda_d^{\pm}}{d} = (1 + O(\exp(-\tfrac{1}{2}s \log s))) \prod_{\substack{p \leq P \\ p \nmid 10|\Delta|}} \left(1 - \frac{1}{p}\right),$$

where

$$s = \frac{\log D}{\log P} \geq \sqrt{\log x} \gg \sqrt{\log MN}.$$

The error term of (9.11) contributes

$$\ll \theta^{-46} (MN)^{\gamma_0} N \exp(-c\sqrt{\log MN}),$$

where $c > 0$ is some fixed constant, and this is acceptably small for (9.4). We write

$$\begin{aligned} \prod_{\substack{p \leq P \\ p \nmid 10|\Delta|}} \left(1 - \frac{1}{p}\right) &= \prod_{p \leq P} \left(1 - \frac{1}{p}\right) \prod_{\substack{p \leq P \\ p \nmid 10|\Delta|}} \left(1 - \frac{1}{p}\right)^{-1} \\ &= \left(1 + O\left(\frac{\log N}{P}\right)\right) \prod_{p \leq P} \left(1 - \frac{1}{p}\right) \prod_{p \nmid 10|\Delta|} \left(1 - \frac{1}{p}\right)^{-1} \\ &= \left(1 + O\left(\frac{\log N}{P}\right)\right) \prod_{p \leq P} \left(1 - \frac{1}{p}\right) \frac{10|\Delta|}{\varphi(10|\Delta|)}, \end{aligned}$$

and observe that the error term contributes at most

$$\ll (\log N) \theta^{-46} \lambda^4 (MN)^{\gamma_0} N P^{-1},$$

which is again acceptable for (9.4) by our lower bound for P . Thus $\Sigma_{\varepsilon,0}^+$ and $\Sigma_{\varepsilon,0}^-$ are asymptotically equal, and up to acceptable error terms we have

$$\begin{aligned} \Sigma = & \sum_{\substack{z_1 \in \mathfrak{B}_1, z_2 \in \mathfrak{B}_2 \\ (z_1 \bar{z}_1 z_2 \bar{z}_2, \Pi)=1 \\ (z_1|\Delta|, z_2)=1 \\ \Delta_{10} \leq \theta^{-28}}} \mu(|z_1|^2) \mu(|z_2|^2) \frac{1}{|\Delta'|} \frac{10|\Delta|}{\varphi(10|\Delta|)} \\ & \times \sum_{\substack{\ell_1 \leq \sqrt{2(1+\theta)}(MN)^{1/2} \\ (\ell, \Pi|\Delta)=1}} \mathbf{1}_{\mathcal{A}}(\ell_1) \sum_{\substack{\ell_2 \leq \sqrt{2(1+\theta)}(MN)^{1/2} \\ \ell_2 z_1 \equiv \ell_1 z_2 (\Delta_{10}) \\ (9.6) \\ (\ell_2, 10)=1}} \mathbf{1}_{\mathcal{A}}(\ell_2). \end{aligned}$$

We may use trivial estimations to replace the condition (9.6) by

$$|\ell_1 z_2 - \ell_2 z_1|^2 \leq 2\Delta(\mathfrak{B}_1, \mathfrak{B}_2)^2 M$$

at the cost of an error of size $O(\theta^{-46} \lambda^{4+\gamma_0/2} (MN)^{\gamma_0} N)$. Further, by trivial estimation we may also remove the conditions $(z_2, |\Delta|) = 1$, $(z_1, z_2) = 1$, and $(\ell_1, |\Delta|) = 1$ at the cost of an error of size $O((\log N) \theta^{-46} (MN)^{\gamma_0} N P^{-50/77})$. Having removed these conditions, we then write

$$\frac{1}{|\Delta'|} = \frac{\Delta_{10}}{|\Delta|} = (1 + O(\lambda')) \frac{\Delta_{10}}{|\Delta(\mathfrak{B}_1, \mathfrak{B}_2)|}$$

It follows that

$$\mathcal{D}_1(\mathfrak{B}_1, \mathfrak{B}_2) = |\Delta(\mathfrak{B}_1, \mathfrak{B}_2)|^{-1} \prod_{p \leq P} \left(1 - \frac{1}{P}\right) \mathcal{D}_2(\mathfrak{B}_1, \mathfrak{B}_2) + O\left(\theta^{18} (\log MN)^{O(1)} \lambda^4 (MN)^{\gamma_0} N\right),$$

where we recall that

$$\begin{aligned} \mathcal{D}_2(\mathfrak{B}_1, \mathfrak{B}_2) = & \sum_{\substack{z_1 \in \mathfrak{B}_1, z_2 \in \mathfrak{B}_2 \\ (z_1 \bar{z}_1 z_2 \bar{z}_2, \Pi)=1 \\ \Delta_{10} \leq \theta^{-28}}} \mu(|z_1|^2) \mu(|z_2|^2) \Delta_{10} \frac{10|\Delta|}{\varphi(10|\Delta|)} \\ & \times \sum_{\substack{\ell_1, \ell_2 \leq \sqrt{2(1+\theta)}(MN)^{1/2} \\ \ell_2 z_1 \equiv \ell_1 z_2 (\Delta_{10}) \\ |\ell_1 z_2 - \ell_2 z_1|^2 \leq 2\Delta(\mathfrak{B}_1, \mathfrak{B}_2)^2 M \\ (\ell_1, \Pi)=1, (\ell_2, 10)=1}} \mathbf{1}_{\mathcal{A}}(\ell_1) \mathbf{1}_{\mathcal{A}}(\ell_2). \end{aligned}$$

Upon using the lower bound $|\Delta(\mathfrak{B}_1, \mathfrak{B}_2)| \gg \theta^{18} N$ we see Proposition 8.1 follows from Proposition 9.1.

10. SIMPLIFICATIONS AND ENDGAME

We are left with the task of proving Proposition 9.1. Our goal is to use the cancellation induced by the Möbius function to show that $\mathcal{D}_2(\mathfrak{B}_1, \mathfrak{B}_2)$ is small. From this point on we do not need to perform any averaging over ℓ_1 and ℓ_2 , so we reduce to considering a sum over z_1 and z_2 ; there are no further complications coming from the fact that ℓ_1 and ℓ_2 belong to a sparse set.

After some manipulations, including splitting into more polar boxes to separate z_1 and z_2 , we reduce to finding cancellation when z_1 and z_2 are summed over arithmetic progressions

whose moduli are bounded by a fixed (but large) power of a logarithm. We detect these congruences with multiplicative characters. We can then get cancellation from the zero-free region for Hecke L -functions, even in the presence of an exceptional zero.

Proof of Proposition 9.1. We interchange the order of summation in $\mathcal{D}_2(\mathfrak{B}_1, \mathfrak{B}_2)$, putting the sums over ℓ_1 and ℓ_2 on the outside and the sums over z_1 and z_2 on the inside. With ℓ_1 and ℓ_2 fixed, we then write

$$\sum_{\substack{z_1 \in \mathfrak{B}_1, z_2 \in \mathfrak{B}_2 \\ \Delta_{10} \leq \theta^{-28}}} \Delta_{10} \leq \sum_{\substack{f \leq \theta^{-28} \\ f | 10^\infty}} f \left| \sum_{\substack{z_1 \in \mathfrak{B}_1, z_2 \in \mathfrak{B}_2 \\ \Delta_{10} = f}} \right|.$$

We can exchange $10|\Delta|/\varphi(10|\Delta|)$ for $|\Delta|/\varphi(|\Delta|)$ by considering separately those f divisible by 5 and those f not divisible by 5, and pulling out potential factors of $5/\varphi(5)$ (recall that $|\Delta|$ is always divisible by 2). To prove Proposition 9.1 it therefore suffices to show

$$(10.1) \quad \mathcal{C} := \sum_{\substack{z_i \in \mathfrak{B}_i \\ (|z_i|^2, \Pi) = 1 \\ \Delta_{10} = f \\ \ell_1 z_2 \equiv \ell_2 z_1 (f) \\ |\ell_1 z_2 - \ell_2 z_1|^2 \leq 2\Delta(\mathfrak{B}_1, \mathfrak{B}_2)^2 M}} \mu(|z_1|^2) \mu(|z_2|^2) \frac{|\Delta|}{\varphi(|\Delta|)} \ll \theta^{64} (\log MN)^{O(1)} \lambda^4 N^2$$

uniformly in $f \leq \theta^{-28}$ with $f | 10^\infty$, and $\ell_1, \ell_2 \ll (MN)^{1/2}$ with $(\ell_1 \ell_2, 10) = 1$. Note that \mathcal{C} depends on $\mathfrak{B}_1, \mathfrak{B}_2, \ell_1, \ell_2$, and f , but we have suppressed this dependence for notational convenience.

For n a positive integer

$$\frac{n}{\varphi(n)} = \sum_{d|n} \frac{\mu^2(d)}{\varphi(d)},$$

and therefore

$$(10.2) \quad \mathcal{C} = \sum_{d \ll N} \frac{\mu^2(d)}{\varphi(d)} \sum_{\substack{z_i \in \mathfrak{B}_i \\ (|z_i|^2, \Pi) = 1 \\ \Delta_{10} = f \\ \ell_1 z_2 \equiv \ell_2 z_1 (f) \\ \Delta \equiv 0(d) \\ |\ell_1 z_2 - \ell_2 z_1|^2 \leq 2\Delta(\mathfrak{B}_1, \mathfrak{B}_2)^2 M}} \mu(|z_1|^2) \mu(|z_2|^2).$$

We introduce a parameter W , and estimate trivially the contribution from $d > W$ in (10.2). Writing z_1 and z_2 in rectangular coordinates, we see the contribution from $d > W$ is bounded by

$$E_W := \sum_{W < d \ll N} \frac{\mu^2(d)}{\varphi(d)} \sum_{\substack{r, s, u, v \ll N^{1/2} \\ rv \equiv su(d)}} \sum \sum \sum \sum 1.$$

If d, r, s , and u are fixed, then v is fixed modulo $d/(d, r)$, which yields

$$\begin{aligned} E_W &\ll N^{3/2} \sum_{W < d \ll N} \frac{\mu^2(d)}{\varphi(d)d} \sum_{r \ll N^{1/2}} (r, d) + (\log N) N^{3/2} \\ &\ll N^2 \sum_{d > W} \frac{\mu^2(d)\tau(d)}{\varphi(d)d} + (\log N)^2 N^{3/2} \\ &\ll (\log W) W^{-1} N^2 + (\log N)^2 N^{3/2}. \end{aligned}$$

Setting

$$W = \theta^{-64} \lambda^{-4} \asymp \theta^{-676}$$

then gives an acceptable contribution for (10.1).

The rational congruence $\Delta \equiv 0(d)$ is equivalent to the Gaussian congruence $\overline{z_1} z_2 \equiv z_1 \overline{z_2} (2d)$. Since $(z_i \overline{z_i}, \Pi) = 1$ and $2d \ll W$, we see that $(z_1 \overline{z_1} z_2 \overline{z_2}, 2d) = 1$. We detect this congruence with multiplicative characters modulo $2d$. Since $(\ell_1 \ell_2, 10) = 1$, we may also detect the congruence $z_1 \ell_2 \equiv z_2 \ell_1 (f)$ with multiplicative characters.

We handle the condition $\Delta_{10} = f$ as follows. Write $f = 2^a 5^b$. Then $\Delta_{10} = f$ if and only if

$$\begin{aligned} \Delta &\equiv 0 \pmod{2^a}, & \Delta &\not\equiv 0 \pmod{2^{a+1}}, \\ \Delta &\equiv 0 \pmod{5^b}, & \Delta &\not\equiv 0 \pmod{5^{b+1}}. \end{aligned}$$

These congruences are equivalent to

$$\begin{aligned} \Delta &\equiv 2^a \pmod{2^{a+1}}, \\ \Delta &\equiv 5^b, 2 \cdot 5^b, 3 \cdot 5^b, \text{ or } 4 \cdot 5^b \pmod{5^{b+1}}, \end{aligned}$$

and by the Chinese remainder theorem these are equivalent to

$$\Delta \equiv \nu_1, \nu_2, \nu_3, \text{ or } \nu_4 \pmod{10f},$$

for some residue classes ν_i . We therefore write our sum over z_1 and z_2 as

$$\sum_{\substack{z_1, z_2 \\ \Delta_{10}=f}} = \sum_{\substack{\mathbf{m}, \mathbf{n}(10f) \\ \text{Im}(\overline{\mathbf{m}}\mathbf{n}) \equiv \nu_i(10f)}} \sum_{z_1 \equiv \mathbf{m}(10f)} \sum_{z_2 \equiv \mathbf{n}(10f)}.$$

Observe that the residue classes \mathbf{m}, \mathbf{n} are primitive since $(z_i, \Pi) = 1$. We trivially have

$$\sum_{\substack{\mathbf{m}, \mathbf{n}(10f) \\ \text{Im}(\overline{\mathbf{m}}\mathbf{n}) \equiv \nu_i(10f)}} 1 \ll f^4,$$

so to prove (10.1) it suffices to show that

$$(10.3) \quad \sum_{\substack{z_i \in \mathfrak{B}_i \\ (|z_i|^2, \Pi)=1 \\ |\ell_1 z_2 - \ell_2 z_1|^2 \leq 2\Delta(\mathfrak{B}_1, \mathfrak{B}_2)^2 M}} \mu(|z_1|^2) \psi'(z_1) \mu(|z_2|^2) \psi(z_2) \ll \theta^{176} \lambda^4 (\log MN)^{O(1)} N^2$$

uniformly in characters ψ' and ψ . Here

$$\psi(\mathbf{m}) = \chi(\mathbf{m}) \overline{\chi}(\overline{\mathbf{m}}) \zeta(\mathbf{m}) \phi(\mathbf{m}),$$

where χ is a character modulo $2d$, ζ is a character modulo f , and ϕ is a character modulo $10f$. The character ψ' is given similarly. The bar denotes complex conjugation and not

multiplicative inversion. Observe that ψ, ψ' are characters with modulus at most $O(d^2 f^2) = O(\theta^{-1408})$. Taking the supremum over z_1 , we see that to prove (10.3) it suffices to show that

$$(10.4) \quad \mathcal{S} := \sum_{\substack{z_2 \in \mathfrak{B}_2 \\ (z_2 \overline{z_2}, \Pi) = 1 \\ (10.5)}} \mu(|z_2|^2) \psi(z_2) \ll \theta^{176} \lambda^2 (\log MN)^{O(1)} N,$$

uniformly in ψ, z_1, ℓ_1 , and ℓ_2 . The last condition in the summation conditions for \mathcal{S} is

$$(10.5) \quad \left| z_2 - z_1 \frac{\ell_2}{\ell_1} \right| \leq \sqrt{2} \frac{|\Delta(\mathfrak{B}_1, \mathfrak{B}_2)| M^{1/2}}{\ell_1}.$$

We see that (10.5) forces z_2 to lie in some disc in the Gaussian integers. Since z_2 already lies in a polar box, we need to understand the intersection of a polar box with a disc.

We now introduce a parameter ϖ , which will be small compared to λ . We cover \mathfrak{B}_2 in polar boxes, which we call ϖ -polar boxes, of the form

$$\begin{aligned} R &< |z_2|^2 \leq (1 + \varpi)R, \\ \vartheta &< \arg(z_2) \leq \vartheta + 2\pi\varpi. \end{aligned}$$

There are $O(\lambda^2/\varpi^2)$ such choices for R and ϑ . For technical convenience we use smooth partitions of unity to accomplish this splitting. This amounts to attaching smooth functions $g(|z_2|^2)$ and $q(\arg(z_2))$, where $g(n)$ is a smooth function supported on an interval

$$(10.6) \quad R < n \leq (1 + O(\varpi))R, \quad R \asymp N,$$

and which satisfies

$$(10.7) \quad g^{(j)}(n) \ll_j (\varpi N)^{-j}, \quad j \geq 0.$$

Further, q is a smooth, 2π -periodic function supported on an interval of length $O(\varpi)$ which satisfies

$$(10.8) \quad q^{(j)}(\alpha) \ll_j \varpi^{-j}, \quad j \geq 0.$$

We observe that the boundary of the intersection between \mathfrak{B}_2 and the disc (10.5) is a finite union of circular arcs and line segments. It is straightforward to check that the boundary has length $\ll \lambda N^{1/2}$. Any ϖ -polar box that intersects the boundary is contained in a $O(\varpi N^{1/2})$ -neighborhood of the boundary. We deduce that the total contribution from those boxes not strictly contained in the intersection is

$$\ll \varpi \lambda N,$$

and this is acceptable if we set

$$\varpi = \theta^{176} \lambda \asymp \theta^{329}.$$

(Observe that we have not required ϖ to be the reciprocal of an integer, as we did with λ . Thus, the covering is not perfect, but we have just shown that the residual contribution is negligible.) It follows that

$$\mathcal{S} = O(\theta^{176} \lambda^2 N) + \sum_{(g,q) \in S(z_1, \ell_1, \ell_2)} \sum_{(z_2 \overline{z_2}, \Pi) = 1} \mu(|z_2|^2) \psi(z_2) g(|z_2|^2) q(\arg(z_2)).$$

The number of pairs $(g, q) \in S(z_1, \ell_1, \ell_2)$ is $\ll (\log N)^2 \varpi^{-2} \lambda^2$, so to prove (10.4) it suffices to show that

$$(10.9) \quad \mathcal{S}_{g,q} := \sum_{(z_2 \bar{z}_2, \Pi)=1} \mu(|z_2|^2) \psi(z_2) g(|z_2|^2) q(\arg(z_2)) \ll \theta^{176} \varpi^2 (\log N)^{O(1)} N$$

uniformly in g and q .

Our sum $\mathcal{S}_{g,q}$ is very similar to the sum $S_\chi^k(\beta)$ treated by Friedlander and Iwaniec (see [15, (16.14)]). Our treatment of $\mathcal{S}_{g,q}$ follows their treatment of $S_\chi^k(\beta)$, and we quote the relevant statements and results of [15, Section 16] as necessary. Friedlander and Iwaniec work with characters having moduli divisible by 4, but this is a distinction without material consequence.

We expand $q(\alpha)$ in its Fourier series. From the derivative bounds (10.8) we see the Fourier coefficients satisfy

$$(10.10) \quad \widehat{q}(h) \ll \frac{\varpi}{1 + \varpi^2 h^2}.$$

By means of (10.10) we obtain the truncated Fourier series

$$(10.11) \quad q(\alpha) = \sum_{|h| \leq H} \widehat{q}(h) e^{ih\alpha} + O(\varpi^{-1} H^{-1}).$$

The contribution of the error term in (10.11) to $\mathcal{S}_{g,q}$ is $O(NH^{-1})$.

We next use Mellin inversion to write $g(n)$ as

$$(10.12) \quad g(n) = \frac{1}{2\pi i} \int_{(\sigma)} \widehat{g}(s) n^{-s} ds, \quad s = \sigma + it.$$

As g is supported in the interval (10.6) and satisfies (10.7), we find that the Mellin transform $\widehat{g}(s)$ is entire and satisfies

$$(10.13) \quad \widehat{g}(s) \ll \frac{\varpi N^\sigma}{1 + \varpi^2 t^2}.$$

Applying (10.11) and (10.12) we obtain

$$(10.14) \quad \mathcal{S}_{g,q} = \sum_{|h| \leq H} \widehat{q}(h) \frac{1}{2\pi i} \int_{(\sigma)} \widehat{g}(s) Z_\psi^h(s) ds + O(NH^{-1}),$$

where

$$Z_\psi^h(s) := \sum_{(z \bar{z}, \Pi)=1} \mu(Nz) \psi(z) \left(\frac{z}{|z|} \right)^h (Nz)^{-s}$$

and Nz denotes the norm $|z|^2$ of z . Call an ideal *odd* if it contains no primes over 2 in its factorization into prime ideals. We may equate primary $z \in \mathbb{Z}[i]$ such that $(z \bar{z}, \Pi) = 1$ with ideals \mathfrak{a} satisfying $\gcd(\mathfrak{a} \bar{\mathfrak{a}}, \Pi) = 1$ under the correspondence $\mathfrak{a} = (z)$. Note that all such ideals are odd. Omitting subscripts and superscripts for simplicity, we then have

$$Z(s) = \sum_{(\mathfrak{a} \bar{\mathfrak{a}}, \Pi)=1} \xi(\mathfrak{a}) \mu(N\mathfrak{a}) (N\mathfrak{a})^{-s},$$

where

$$\xi(\mathfrak{a}) := \psi(z) \left(\frac{z}{|z|} \right)^h$$

and z is the unique primary generator of \mathfrak{a} . From the Euler product it follows that

$$Z(s) = L(s, \xi)^{-1} P(s) G(s),$$

where $L(s, \xi)$ is the Hecke L -function

$$L(s, \xi) := \sum_{\mathfrak{a}} \frac{\xi(\mathfrak{a})}{(N\mathfrak{a})^s},$$

$P(s)$ is the Dirichlet polynomial given by

$$P(s) := \prod_{\substack{p \leq P \\ p \equiv 1(4)}} \left(1 - \frac{\xi(\mathfrak{p})}{p^s} \right)^{-1} \left(1 - \frac{\xi(\bar{\mathfrak{p}})}{p^s} \right)^{-1}$$

where $\mathfrak{p}\bar{\mathfrak{p}} = (p)$, and $G(s)$ is given by an Euler product that converges absolutely and uniformly in $\sigma \geq \frac{2}{3}$, say. In the region $\sigma \geq 1 - \frac{1}{\log P}$ the inequality $p^{-\sigma} < 3p^{-1}$ holds, and this gives the bound

$$P(s) \ll (\log P)^3, \quad \sigma \geq 1 - \frac{1}{\log P}.$$

Let k be the modulus of ξ (recall that $k \ll \theta^{-1408}$). Then $L(s, \xi)$ is nonzero (see [15, (16.20)]) in the region

$$\sigma \geq 1 - \frac{c}{\log(k + |h| + |t|)},$$

except for possibly an exceptional real zero when ξ is real. By applying the method of Siegel ([15, Lemma 16.1]) one may show that when ξ is real $L(s, \xi)$ has no zeros in the region

$$(10.15) \quad \sigma \geq 1 - \frac{c(\varepsilon)}{k^\varepsilon}, \quad 0 < \varepsilon \leq \frac{1}{4}.$$

The constant $c(\varepsilon)$ is ineffective, and for this reason the implied constant in Proposition 9.1 is ineffective.

The inequality (10.15) allows one to establish ([15, (16.23) and (16.24)]) the upper bound

$$L(s, \xi)^{-1} \ll k(\log(|h| + |t| + 3))^2$$

in the region

$$\sigma \geq 1 - \frac{c(\varepsilon)}{k^\varepsilon \log(|h| + |t| + 3)}.$$

For $T \geq |h| + 3$, we set

$$\beta := \min \left(\frac{c(\varepsilon)}{k^\varepsilon \log T}, \frac{1}{\log P} \right),$$

so that in the region $\sigma \geq 1 - \beta$ we have the bound

$$(10.16) \quad Z(s) \ll k(\log(|h| + |t| + N))^5.$$

We now estimate the integral

$$(10.17) \quad I := \frac{1}{2\pi i} \int_{(\sigma)} \widehat{g}(s) Z(s) ds.$$

We move the contour of integration to

$$\begin{aligned} s &= 1 + it, & |t| &\geq T, \\ s &= 1 - \beta + it, & |t| &\leq T, \end{aligned}$$

and add in horizontal connecting segments

$$s = \sigma + \pm iT, \quad 1 - \beta \leq \sigma \leq 1.$$

Estimating trivially we find by (10.13) and (10.16) that

$$I \ll \varpi^{-1} \theta^{-1408} (T^{-1} + N^{-\beta}) N(\log(N + T))^5.$$

We set $T := 3 \exp(\sqrt{\log N})$. Recalling that $\log P \leq \frac{\sqrt{\log x}}{\log \log x} \ll_{\delta} \frac{\sqrt{\log N}}{\log \log N}$, we see that

$$(10.18) \quad I \ll (\log N)^5 \varpi^{-1} \theta^{-1408} N \left(\exp \left(-\frac{c(\varepsilon) \sqrt{\log N}}{k^{\varepsilon}} \right) + \exp \left(-c \sqrt{\log N} \right) \right)$$

uniformly in $|h| \leq 2 \exp(\sqrt{\log N})$. We choose $H := \exp(\sqrt{\log N})$, then take (10.14) together with (10.18) and sum over $|h| \leq H$ by means of (10.10). Provided $\varepsilon > 0$ is sufficiently small in terms of θ (take $\varepsilon = \varepsilon(L)$, with L as in the statement of Proposition 9.1), we obtain the bound

$$(10.19) \quad \mathcal{S}_{g,q} \ll N \exp \left(-c_L (\log N)^{1/3} \right),$$

say, where the positive constant c_L is ineffective. By hypothesis we have $N \geq x^{\delta}$, so the bound (10.19) implies (10.9). This completes the proof of Proposition 9.1. \square

11. MODIFICATIONS FOR THEOREM 1.2

The proof of Theorem 1.2 follows the same lines as the proof of Theorem 1.1. We provide a sketch of the modified argument, and leave the task of fleshing out complete details to the interested reader.

We let $d \in \{2, 3\}$, and let $\{a_1, \dots, a_d\} \subset \{0, 1, 2, \dots, 9\}$ be a fixed set. Denote by \mathcal{A}_d the set of nonnegative integers missing the digits a_1, \dots, a_d in their decimal expansions. Let $\gamma_d := \frac{\log(10-d)}{\log 10}$. For Y a power of 10 we define

$$F_{Y,d}(\theta) := Y^{-\gamma_d} \left| \sum_{n < Y} \mathbf{1}_{\mathcal{A}_d}(n) e(n\theta) \right|.$$

We note that if $Y = 10^k$ then

$$\begin{aligned} F_{Y,d}(\theta) &= \prod_{i=0}^{k-1} \frac{1}{10-d} \left| \sum_{n_i < 10} \mathbf{1}_{\mathcal{A}_d}(n_i) e(n_i 10^i \theta) \right| \\ &= \prod_{i=1}^k \frac{1}{10-d} \left| \frac{e(10^i \theta) - 1}{e(10^{i-1} \theta) - 1} - \sum_{r=1}^d e(a_r 10^{i-1} \theta) \right|. \end{aligned}$$

We therefore have the product formula

$$F_{UV,d}(\theta) = F_{U,d}(\theta)F_{V,d}(U\theta).$$

The most important task is to obtain analogues of Lemmas 5.2, 5.3, and 5.4 for the functions $F_{Y,d}$. By arguing as in the proof of [24, Lemma 10.1] it is not difficult to prove the analogue of Lemma 5.2.

Lemma 11.1. *Let $q < Y^{\frac{1}{3}}$ be of the form $q = q_1 q_2$ with $(q_1, 10) = 1$ and $q_1 > 1$. Then for any integer a coprime with q we have*

$$F_{Y,d}\left(\frac{a}{q}\right) \ll \exp\left(-c \frac{\log Y}{\log q}\right)$$

for some absolute constant $c > 0$.

It is a little more difficult to obtain the analogues of Lemmas 5.3 and 5.4. They will follow from a good upper bound for

$$\sup_{\beta \in \mathbb{R}} \sum_{a < Y} F_{Y,d}\left(\frac{a}{Y} + \beta\right).$$

The key is that we can estimate moments of $F_{Y,d}$ by numerically computing the largest eigenvalue of a certain matrix.

Lemma 11.2. *Let J be a positive integer. Let $\lambda_{t,J,d}$ be the largest eigenvalue of the $10^J \times 10^J$ matrix $M_{t,d}$, given by*

$$(M_{t,d})_{i,j} := \begin{cases} G_d(a_1, \dots, a_{J+1})^t, & \text{if } i-1 = \sum_{\ell=1}^J a_{\ell+1} 10^{\ell-1}, j-1 = \sum_{\ell=1}^J a_{\ell} 10^{\ell-1} \\ & \text{for some } a_1, \dots, a_{J+1} \in \{0, \dots, 9\}, \\ 0, & \text{otherwise,} \end{cases}$$

where

$$G_d(t_0, \dots, t_J) := \sup_{|\gamma| \leq 10^{-J-1}} \frac{1}{10-d} \left| \frac{e\left(\sum_{j=0}^J t_j 10^{-j} + 10\gamma\right) - 1}{e\left(\sum_{j=0}^J t_j 10^{-j-1} + \gamma\right) - 1} - \sum_{r=1}^d e\left(\sum_{j=0}^J a_r t_j 10^{-j-1} + a_r \gamma\right) \right|.$$

Then

$$\sum_{0 \leq a < 10^k} F_{10^k,d}\left(\frac{a}{10^k}\right)^t \ll_{t,J,d} \lambda_{t,J,d}^k.$$

Proof. This is an easy adaptation of the proof of [24, Lemma 10.2]. □

The following is a consequence of Lemma 11.2 and some numerical computation.

Lemma 11.3. *We have*

$$\sup_{\beta \in \mathbb{R}} \sum_{a < Y} F_{Y,d}\left(\beta + \frac{a}{Y}\right) \ll Y^{\alpha_d}$$

and

$$\int_0^1 F_{Y,d}(t) dt \ll Y^{-1+\alpha_d},$$

where

$$\alpha_2 = \frac{54}{125}, \quad \alpha_3 = \frac{99}{200}.$$

Proof. We use bounds on $\lambda_{1,2,2}$ and $\lambda_{1,2,3}$. By numerical calculation² we find

$$\lambda_{1,2,2} < 10^{54/125}$$

for all choices of $\{a_1, a_2\} \subset \{0, \dots, 9\}$, and

$$\lambda_{1,2,3} < 10^{99/200}$$

for all choices of $\{a_1, a_2, a_3\} \subset \{0, \dots, 9\}$. By the argument of [24, Lemma 10.3] this then yields

$$\sup_{\beta \in \mathbb{R}} \sum_{a < Y} F_{Y,d} \left(\beta + \frac{a}{Y} \right) \ll Y^{\alpha_d}.$$

To complete the proof we observe

$$\begin{aligned} \int_0^1 F_{Y,d}(t) dt &= \sum_{0 \leq a < Y} \int_{\frac{a}{Y}}^{\frac{a}{Y} + \frac{1}{Y}} F_{Y,d}(t) dt = \int_0^{\frac{1}{Y}} \sum_{0 \leq a < Y} F_{Y,d} \left(\frac{a}{Y} + t \right) dt \\ &\leq Y^{-1} \sup_{\beta \in \mathbb{R}} \sum_{a < Y} F_{Y,d} \left(\frac{a}{Y} + \beta \right) \ll Y^{-1+\alpha_d}. \end{aligned} \quad \square$$

Remark. It is crucial for the proof of Theorem 1.2 that $\alpha_d < \frac{1}{2}$. For $d \geq 4$ there exist choices of excluded digits which force $\alpha_d > \frac{1}{2}$.

Lemma 11.4. *We have*

$$\begin{aligned} \sup_{\beta \in \mathbb{R}} \sum_{a \leq q} \left| F_{Y,d} \left(\frac{a}{q} + \beta \right) \right| &\ll q^{\alpha_d} + \frac{q}{Y^{1-\alpha_d}}, \\ \sup_{\beta \in \mathbb{R}} \sum_{q \leq Q} \sum_{\substack{1 \leq a \leq q \\ (a,q)=1}} \left| F_{Y,d} \left(\frac{a}{q} + \beta \right) \right| &\ll Q^{2\alpha_d} + \frac{Q^2}{Y^{1-\alpha_d}}. \end{aligned}$$

Proof. We use the large sieve argument of [24, Lemma 10.5] with Lemma 11.3. \square

Let us now give a broad sketch of the proof of Theorem 1.2. We proceed as in the proof of Theorem 1.1, only we use Lemma 11.4 instead of Lemma 5.3 or Lemma 5.4.

The sequence

$$\sum_{\substack{m^2 + \ell^2 = n \\ (\ell, \Pi) = 1}} \sum \mathbf{1}_{\mathcal{A}_d}(\ell)$$

has level of distribution

$$D \leq x^{\gamma_d - \delta},$$

and we have an acceptable Type II bound provided

$$(11.1) \quad x^{1/2 - \gamma_d/2 + \delta} \leq N \leq x^{\frac{1}{2}(1 - \alpha_d) - \delta}.$$

²Mathematica® files with these computations are included with this work on arxiv.org.

(Compare (11.1) with (6.5).) Since

$$\frac{1}{2}(1 - \alpha_d) - \left(\frac{1}{2} - \frac{\gamma_d}{2}\right) > 1 - \gamma_d$$

there exists an appropriate choice of U and V in Vaughan's identity (2.1) (compare with (6.2)).

At various points in the proof of Theorem 1.1 we had to perform a short interval decomposition in order to gain control on elements of \mathcal{A} in arithmetic progressions (see the arguments in Section 5 leading up to Lemmas 5.5 and 5.6). The short interval decomposition depended on whether or not the missing digit was the zero digit. In the case of Theorem 1.2 one argues similarly, and finds that the short interval decomposition depends only on whether $0 \in \mathcal{A}_d$.

ACKNOWLEDGEMENTS

The author wishes to thank his adviser Kevin Ford for suggesting the problem which resulted in this paper, and for many helpful conversations. The author expresses gratitude to Junxian Li for some assistance with Mathematica®. Many thanks to the anonymous referee, whose careful reading of this paper and thoughtful comments have improved the quality of this work.

Part of this research was conducted while the author was supported by NSF grant DMS-1501982.

REFERENCES

- [1] W. Banks, A. Conflitti, I. Shparlinski. *Character sums over integers with restricted g -ary digits*. Illinois J. Math. 46 (2002), no. 3, 819-836
- [2] W. Banks, I. Shparlinski. *Arithmetic properties of numbers with restricted digits*. Acta Arith. 112 (2004), no. 4, 313-332
- [3] Jean Bourgain. *Prescribing the binary digits of primes, II*. Israel J. Math. 206 (2015), no. 1, 165-182
- [4] Sylvain Col. *Diviseurs des nombres ellipsépiques*. Period. Math. Hungar. 58 (2009), no. 1, 1-23
- [5] Jean Coquet. *On the uniform distribution modulo one of some subsequences of polynomial sequences*. J. Number Theory 10 (1978), no. 3, 291-296
- [6] Jean Coquet. *On the uniform distribution modulo one of some subsequences of polynomial sequences. II*. J. Number Theory **12** (1980), no. 2, 244-250
- [7] C. Dartyge, C. Mauduit. *Nombres presque premiers dont l'écriture en base r ne comporte pas certains chiffres*. J. Number Theory 81 (2000), no. 2, 270-291
- [8] C. Dartyge, C. Mauduit. *Ensembles de densité nulle contenant des entiers possédant au plus deux facteurs premiers*. J. Number Theory 91 (2001), no. 2, 230-255
- [9] M. Drmota, C. Mauduit. *Weyl sums over integers with affine digit restrictions*. J. Number Theory 130 (2010), no. 11, 2404-2427
- [10] P. Erdős, C. Mauduit, A. Sárközy. *On arithmetic properties of integers with missing digits. I. Distribution in residue classes*. J. Number Theory 70 (1998), no. 2, 99-120
- [11] P. Erdős, C. Mauduit, A. Sárközy. *On arithmetic properties of integers with missing digits. II. Prime factors*. Paul Erdős memorial collection. Discrete Math. 200 (1999), no. 1-3, 149-164
- [12] Leonhard Euler, letter to Christian Goldbach, dated April 12, 1749. <http://eulerarchive.maa.org/correspondence/letters/000852.pdf>
- [13] E. Fouvry, H. Iwaniec. *Gaussian primes*. Acta Arith. **79** (1997), no. 3, 249-287
- [14] J. Friedlander, H. Iwaniec. *Asymptotic sieve for primes*. Ann. of Math. (2) **148** (1998), no. 3, 1041-1065
- [15] J. Friedlander, H. Iwaniec. *The polynomial $x^2 + y^4$ captures its primes*. Ann. of Math. (2) **148** (1998), no. 3, 945-1040
- [16] J. Friedlander, H. Iwaniec. *Opera de cribro*. American Mathematical Society Colloquium Publications, 57. American Mathematical Society, Providence, RI, 2010

- [17] Glyn Harman. *Prime-detecting sieves*. London Mathematical Society Monographs Series, 33. Princeton University Press, Princeton, NJ, 2007
- [18] D. R. Heath-Brown. *Primes represented by $x^3 + 2y^3$* . Acta Math. 186 (2001), no. 1, 1-84
- [19] D. R. Heath-Brown, X. Li. *Prime values of $a^2 + p^4$* . Invent. Math. **208** (2017), no. 2, 441-499
- [20] D. R. Heath-Brown, B. Z. Moroz. *On the representation of primes by cubic polynomials in two variables*. Proc. London Math. Soc. (3) 88 (2004), no. 2, 289-312
- [21] E. Kowalski, H. Iwaniec. *Analytic number theory*. American Mathematical Society Colloquium Publications, 53. American Mathematical Society, Providence, RI, 2004
- [22] Sergei Konyagin. *Arithmetic properties of integers with missing digits: distribution in residue classes*. Period. Math. Hungar. 42 (2001), no. 1-2, 145-162
- [23] James Maynard. *Primes represented by incomplete norm forms*. Preprint. <https://arxiv.org/abs/1507.05080>
- [24] James Maynard. *Primes with restricted digits*. Invent. Math. 217 (2019), no. 1, 127-218
- [25] H. L. Montgomery, R. C. Vaughan. *Multiplicative Number Theory. I. Classical theory*. Cambridge Studies in Advanced Mathematics, 97. Cambridge University Press, Cambridge, 2007
- [26] *Oeuvres de Fermat*, ed. Paul Tannery and Charles Henry, 4 vols (Paris, 1891-1912), ii
- [27] E. C. Titchmarsh. *The theory of the Riemann zeta function*. Second edition. Edited and with a preface by D. R. Heath-Brown. The Clarendon Press, Oxford University Press, New York, 1986

ALL SOULS COLLEGE, UNIVERSITY OF OXFORD, OXFORD OX1 4AL

E-mail address: kvpratt@gmail.com