TOPICS IN ANALYTIC NUMBER THEORY

BY

KYLE PRATT

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mathematics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Doctoral Committee:

Professor Scott Ahlgren, Chair
Professor Kevin Ford, Director of Research
Professor Bruce Berndt
Professor Alexandru Zaharescu

ABSTRACT

We investigate properties of prime numbers and L-functions, and interactions between these two topics. First, we discuss the problem of primes in thin sequences, expanding on work of Maynard and Friedlander-Iwaniec. Next, motivated by work of Iwaniec and Sarnak, we study the question of average nonvanishing of Dirichlet L-functions at the central point. Finally, in joint work with Siegfred Baluyot, we build on work of Soundararajan and synthesize our studies of primes and L-functions by examining Dirichlet L-functions of quadratic characters of prime conductor.

For Dana.

ACKNOWLEDGMENTS

I express my thanks to the mathematics department of the University of Illinois at Urbana-Champaign for support, and for providing a pleasant and enjoyable work environment. I am grateful to the number theory group for awarding me the Bateman Prize.

I express my thanks to George Shakan, a friend and academic brother, for countless insightful conversations.

I express my thanks to Siegfred Baluyot, with whom the work of Chapter 3 came about in a wonderful and enjoyable collaboration.

I express my thanks to my advisor, Kevin Ford, for spending much of his valuable time on meeting with me, encouraging me, and sharing his ideas with me. I also thank him for his very generous financial support, which enabled me to attend many more conferences than would otherwise have been possible.

I express my thanks and love to my dear wife, Dana, who gave so much of herself to assist me in bringing to pass this thesis. I am a better mathematician for having her in my life. Thanks also to my son, Andrew, who let me leave behind the toy trucks and trains so I could go to work, and who was always eager to play again when I came back home.

Some of the work in this thesis was supported by the National Science Foundation under grant number DMS-1501982, and by the National Science Foundation Graduate Research Fellowship under grant number DGE-1144245. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

TABLE OF CONTENTS

INTRODUCTION

The prime numbers are the fundamental building blocks of arithmetic, and consequently the primes are of considerable interest to number theorists. In fact, primes should be of interest to everyone, mathematician or not, since the cryptographic protocols upon which modern internet commerce are based rely crucially on the indivisible, primordial nature of prime numbers.

It is easy to ask questions about prime numbers, but much harder to get the primes to reveal their secrets. Mathematicians have invented many different tools in attempting to uncover these secrets. Some of the most powerful tools for studying primes are objects known as L-functions. An L-function is a kind of mathematical function that encodes a vast amount of symmetry and arithmetic information.

In this thesis we will investigate three problems relating to prime numbers and L-functions. First, we discuss the problem of primes in thin sequences, expanding on work of Maynard and Friedlander-Iwaniec. Next, motivated by work of Iwaniec and Sarnak, we study the question of average nonvanishing of Dirichlet L-functions at the central point. Finally, in joint work with Siegfred Baluyot, we build on work of Soundararajan and synthesize our studies of primes and L-functions by examining Dirichlet L-functions of quadratic characters of prime conductor.

0.1 Primes in thin sequences

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 [1, 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 proof[1].

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, for $x^2 + y^2$ 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 $|\mathcal{S}| \leq x^{1-\epsilon}$ for some $\epsilon > 0$). 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 [3] 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 $\approx x(\log x)^{-2}$, 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 [4] 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 [5] on primes of the form $p = x^3 + 2y^3$ and Heath-Brown and Moroz [6] on primes represented by cubic forms, and Maynard [7] on primes represented by incomplete norm forms. Heath-Brown and Li [8] 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 $a_0 \in \{0, 1, 2, \ldots, 9\}$ be

[1] Euler finally found a proof more than a century later [2].

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

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

and note that

$$\sum_{\ell \leq y} \mathbf{1}_{\mathcal{A}}(\ell) \asymp y^{\gamma_0}, \quad y \geq 2. \tag{0.1.1}$$

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.

James Maynard showed in a beautiful paper [9] that 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, [10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21] 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 [3] and the work of Friedlander and Iwaniec [4]. We prove there are infinitely many primes of the form $p = m^2 + \ell^2$, where ℓ is missing a fixed digit in its decimal expansion. The following theorem gives a precise statement.

Theorem 0.1.1. *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

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

where γ *denotes Euler's constant,*

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

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.*

It is technically convenient for the proof of Theorem 0.1.1 to have ℓ not divisible by small prime factors. We also remark that it is potential exceptional zeros for certain Hecke L-functions that make the implied constant in Theorem 0.1.1 ineffective.

0.2 Nonvanishing of L-functions at the central point

The values of L-functions at special points on the complex plane are of great interest. At the fixed point of the functional equation, called the central point, the question of nonvanishing is particularly important. For instance, the well-known Birch and Swinnerton-Dyer conjecture [22] relates the order of vanishing of certain L-functions at the central point to the arithmetic of elliptic curves. Katz and Sarnak [23] discuss several examples of families of L-functions and describe how the zeros close to $s=\frac{1}{2}$ give evidence of some underlying symmetry group for each of these families. They suggest that understanding these symmetries may in turn lead to finding a natural spectral interpretation of the zeros of the L-functions. The analysis of each family they discuss leads to a *Density Conjecture* that, if true, would imply that almost all L-functions in the family do not vanish at the central point. Iwaniec and Sarnak [24] show that the nonvanishing of L-functions associated with holomorphic cusp forms is closely related to the Landau-Siegel zero problem. Thus the question of nonvanishing at the central point is connected to many deep arithmetical problems.

A considerable amount of research has been done towards answering this question for families of Dirichlet L-functions. Chowla conjectured that $L(\frac{1}{2}, \chi) \neq 0$ for χ a primitive quadratic Dirichlet character [25, p. 82, problem 3]. It has since become a sort of folklore conjecture that $L(\frac{1}{2}, \chi) \neq 0$ for all primitive Dirichlet characters χ. One family that has attracted much attention is the family of $L(s, \chi)$ with χ varying over primitive characters modulo a fixed conductor. This family is widely believed to have a unitary symmetry type, as in the philosophy of Katz and Sarnak. Balasubramanian and Murty [26] were the first to prove that a (small) positive proportion of this family does not vanish at the central point. They used the celebrated technique of mollified moments, a method that has been highly useful in other contexts (see, for example, [27, 28, 29]). Iwaniec and Sarnak [30] developed a simpler, stronger version of the method and improved this proportion to $\frac{1}{3}$. The approach of Iwaniec and Sarnak has since become standard in the study of nonvanishing of L-functions at the central point. Bui [31] and Khan and Ngo [32] introduced new ideas and further improved the lower bound $\frac{1}{3}$.

If one assumes the Generalized Riemann Hypothesis, one can show that at least half of the primitive characters $\chi \pmod q$ satisfy $L\left(\frac{1}{2}, \chi\right) \neq 0$ [33, 34],[35, Exercise 18.2.8]. One uses the explicit formula, rather than mollification, and the proportion $\frac{1}{2}$ arises from the choice of a test function with certain positivity properties.

It seems plausible that one may obtain a larger proportion of nonvanishing by also averaging over moduli q. Indeed, Iwaniec and Sarnak [30] already claimed that by averaging over moduli one can prove at least half of the central values are nonzero. This is striking, in that it is as strong, on average, as the proportion obtained via GRH.

A natural question is whether, by averaging over moduli, one can breach the 50% barrier, thereby going beyond the immediate reach of GRH. We answer this question in the affirmative[2].

Let $\sum^*_{\chi(q)}$ denote a sum over the primitive characters modulo q, and define $\varphi^*(q)$ to be the number of primitive characters modulo q.

Theorem 0.2.1. *Let Ψ be a fixed, nonnegative smooth function, compactly supported*

[2]This work, which appears in Chapter 2, was published as [36].

in $[\frac{1}{2}, 2]$, *which satisfies*

$$\int_{\mathbb{R}} \Psi(x)dx > 0.$$

Then for Q *sufficiently large we have*

$$\sum_q \Psi\left(\frac{q}{Q}\right) \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\substack{\chi(q) \\ L\left(\frac{1}{2},\chi\right)\neq 0}} 1 \geq 0.50073 \sum_q \Psi\left(\frac{q}{Q}\right) \frac{q}{\varphi(q)} \varphi^*(q).$$

Thus, roughly speaking, a randomly chosen central value $L\left(\frac{1}{2}, \chi\right)$ is more likely nonzero than zero. We remark also that the appearance of the arithmetic weight $\frac{q}{\varphi(q)}$ is technically convenient, but not essential.

For further interesting research on this and other families of L-functions, see [37, 38, 39, 40, 41, 42, 43, 44, 45, 46].

The family of $L(s, \chi)$ with χ varying over all real primitive characters has also been extensively studied. This family is of particular significance because it seems to be of symplectic rather than unitary symmetry. Thus we encounter new phenomena not seen in the unitary case. For d a fundamental discriminant, set $\chi_d(\cdot) = \left(\frac{d}{\cdot}\right)$, the Kronecker symbol. Then χ_d is a real primitive character with conductor $|d|$. The hypothetical positivity of central values $L(\frac{1}{2}, \chi_d)$ has implications for the class number of imaginary quadratic fields [47, p. 514]. Jutila [48] initiated the study of nonvanishing at the central point for this family and proved that $L(\frac{1}{2}, \chi_d) \neq 0$ for infinitely many fundamental discriminants d. His methods show that $\gg X/\log X$ of the quadratic characters χ_d with $|d| \leq X$ have $L(\frac{1}{2}, \chi_d) \neq 0$. Özlük and Snyder [49] examined the low-lying zeros of this family, and found the first evidence of its symplectic behavior. Assuming the Generalized Riemann Hypothesis, they showed that at least $\frac{15}{16}$ of the central values $L(\frac{1}{2}, \chi_d)$ are non-zero [50]. Katz and Sarnak independently obtained the same result in unpublished work (see [23, 51]). Soundararajan [51] made a breakthrough when he proved unconditionally that at least $\frac{7}{8}$ of the central values $L(\frac{1}{2}, \chi_d)$ with $d \equiv 0 \pmod 8$ are non-zero.

The case of real primitive characters with prime conductor is more difficult still. Jutila [48] initiated the study of $L(\frac{1}{2}, \chi_p)$, where p is a prime. His methods yield that $\gg X/(\log X)^3$ of the primes $p \leq X$ satisfy $L(\frac{1}{2}, \chi_p) \neq 0$. The difficulty in studying this family is that its moments involve sums over primes, and thus are more complicated to investigate. In fact, Jutila only evaluated the first moment of this family. As far as we are aware, no asymptotic evaluation of the second moment has appeared in the literature. However, Andrade and Keating [52] asymptotically evaluated the second moment of an analogous family over function fields. Andrade and Baluyot [53] have continued the study of the family of $L(\frac{1}{2}, \chi_p)$, showing that it is likely governed by a symplectic law. Conditionally on GRH, they prove that at least 75% of primes $p \leq X$ satisfy $L(\frac{1}{2}, \chi_p) \neq 0$.

We prove the first unconditional positive proportion result for the central values $L(\frac{1}{2}, \chi_p)$. In fact, we prove that more than nine percent of these central values are non-zero.

Theorem 0.2.2. *There exists an absolute, effective constant X_0 such that if $X \geq X_0$ then*

$$\sum_{\substack{p \leq X \\ p \equiv 1 \ (\mathrm{mod}\ 8) \\ L(\frac{1}{2}, \chi_p) \neq 0}} 1 \geq .0964 \sum_{\substack{p \leq X \\ p \equiv 1 \ (\mathrm{mod}\ 8)}} 1.$$

The tools developed for the proof of Theorem 0.2.2 allow one to obtain the order of magnitude of the second moment of $L(\frac{1}{2}, \chi_p)$.

Theorem 0.2.3. *Let $\mathfrak{c}$ be the positive constant*

$$\mathfrak{c} := \left(144\zeta(2)\left(1 - \frac{1}{\sqrt{2}}\right)^2\right)^{-1} = .0492\ldots.$$

For large X we have

$$(\mathfrak{c} - o(1))\frac{X}{4}(\log X)^3 \leq \sum_{\substack{p \leq X \\ p \equiv 1 \ (\mathrm{mod}\ 8)}} (\log p) L\left(\tfrac{1}{2}, \chi_p\right)^2 \leq (4\mathfrak{c} + o(1))\frac{X}{4}(\log X)^3.$$

One would rather have an upper bound in Theorem 0.2.3 that asymptotically matches the lower bound, but this seems difficult to prove unconditionally. By adapting a method of Soundararajan and Young [54] we are able, however, to prove such an asymptotic formula on GRH.

Theorem 0.2.4. *Let $\mathfrak{c}$ be as in Theorem 0.2.3. Assume the Riemann Hypothesis for*

$\zeta(s)$ and for all Dirichlet L-functions $L(s,\chi_p)$ with $p \equiv 1 \pmod 8$. Then

$$\sum_{\substack{p \le X \\ p \equiv 1 \ (\mathrm{mod}\ 8)}} (\log p) L\left(\tfrac{1}{2}, \chi_p\right)^2 = \mathfrak{c}\frac{X}{4}(\log X)^3 + O(X(\log X)^{11/4}).$$

Our methods further yield the order of magnitude of the third moment of $L(\frac{1}{2}, \chi_p)$, assuming that the central values $L(\frac{1}{2}, \chi_n)$ are non-negative for certain fundamental discriminants n. This non-negativity hypothesis follows, of course, from GRH.

Theorem 0.2.5. *Assume that for all positive square-free integers n with $n \equiv 1 \pmod 8$ it holds that $L(\frac{1}{2}, \chi_n) \geq 0$. Then for large X*

$$\sum_{\substack{p \le X \\ p \equiv 1 \ (\mathrm{mod}\ 8)}} (\log p) L\left(\tfrac{1}{2}, \chi_p\right)^3 \asymp X(\log X)^6.$$

CHAPTER 1

PRIMES FROM SUMS OF TWO SQUARES AND MISSING DIGITS

1.1 Introductory remarks

The goal of this chapter[1] is to present the proof of Theorems 0.1.1. We are able to avoid more sophisticated sieves like Harman's sieve [55], and instead we require only Vaughan's identity (see (1.2.1)). The application of Vaughan's identity reduces the problem to the estimation of "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 1.3.1). 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 1.1.1. *Let $\mathcal{B} \subset \{0, 1, \ldots, 9\}$ 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 $\{0, 1, \ldots, 9\} \backslash \mathcal{B}$. Let*

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

[1]The work in this chapter has been submitted for publication.

Then, with the notation as above, we have

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

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.

When $|\mathcal{B}|=3$, Theorem 1.1.1 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.

In this chapter we shall give a complete proof of Theorem 0.1.1, and at the end of the chapter we shall describe the modifications which are necessary for the proof of Theorem 1.1.1.

Throughout this chapter 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 1.6 onward we allow the implied constants to depend on L (see (1.6.7)) without indicating this in the notation. A subscript such as $f \ll_\epsilon g$ means the implied constant depends on ϵ.

We use the convention that ϵ denotes an arbitrarily small positive quantity that may vary from one occurrence to the next. Thus, 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, and $P^+(n), P^-(n)$ for the largest and smallest prime factors of n, respectively.

1.2 Initial manipulations and outline

We begin the proof of Theorem 0.1.1 by setting out to estimate

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

where

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

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

Let $U, V > 2$ be real parameters to be chosen later (see (1.6.3)). For an arithmetic function $f : \mathbb{N} \to \mathbb{C}$ and $W \ge 1$, define

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

and write $f_{>W} = f - f_{\le W}$. Then Vaughan's identity is

$$\Lambda = \Lambda_{\le U} + \mu_{\le V} \star \log - \Lambda_{\le U} \star \mu_{\le V} \star 1 + \Lambda_{>U} \star \mu_{>V} \star 1. \tag{1.2.1}$$

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

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

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

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 (1.2.1) into the definition of $S(x)$ gives

$$\begin{aligned} S(x) = \sum_{n\leq x} a(n)\Lambda(n) = \sum_{n\leq U} a(n)\Lambda(n) + \sum_{n\leq x} a(n)(\mu_{\leq V} \star \log)(n) \\ - \sum_{n\leq x} a(n)(\mu_{\leq V} \star \Lambda_{\leq U} \star 1)(n) + \sum_{n\leq x} a(n)(\mu_{>V} \star \Lambda_{>U} \star 1)(n). \end{aligned} \tag{1.2.2}$$

By trivial estimation

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

the last inequality following by (0.1.1). In what follows we shall have many occasions to use the bound

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

and we do so without further comment.

For the second sum in (1.2.2) we interchange the order of summation and separate the logarithmic factors to obtain

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

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 (1.2.2), the Type II sum, we interchange the order of summation and change variables to obtain

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

In short,

$$S(x) = A(x;U,V) + B(x;U,V) + O((\log U)U^{\frac{1}{2}+\frac{\gamma_0}{2}}), \tag{1.2.3}$$

where

$$A(x;U,V) := \sum_{d\leq V}\mu(d)\left(A_d'(x) - A_d(x)\log d - \sum_{m\leq U}\Lambda(m)A_{dm}(x)\right) \tag{1.2.4}$$

and

$$B(x;U,V) := \sum_{U<m\leq x/V}(\Lambda_{>U}\star 1)(m)\sum_{V<n\leq x/m}\mu(n)a(mn). \tag{1.2.5}$$

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

$$A_d'(x) = A_d(x)\log x - \int_1^x A_d(t)\frac{dt}{t}. \tag{1.2.6}$$

Define

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

where

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

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 terms

$$R_d(x) := A_d(x) - M_d(x), \qquad R(x, D) := \sum_{d \leq D} |R_d(x)|. \tag{1.2.7}$$

Inserting (1.2.6) into (1.2.4) and writing $A_d(x) = M_d(x) + R_d(x)$, we obtain

$$A(x; U, V) = M(x; U, V) + R(x; U, V), \tag{1.2.8}$$

where

$$M(x; U, V) = \sum_{n \leq x} \sum_{d \leq V} \frac{\mu(d)}{d} \left(a_d(n) \log(n/d) - \sum_{m \leq U} \frac{\Lambda(m)}{m} a_{dm}(n) \right) \tag{1.2.9}$$

and

$$R(x; U, V) = \sum_{d \leq V} \mu(d) \left(R_d(x) \log(x/d) - \int_1^x R_d(t) \frac{dt}{t} - \sum_{m \leq U} \Lambda(m) R_{md}(x) \right). \tag{1.2.10}$$

This completes our preliminary manipulations of $S(x)$.

The outline of the rest of the chapter is as follows. In Section 1.3 we show that $R(x; U, V)$ contributes only to the error term in Theorem 0.1.1. The analysis in Section 1.4 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 0.1.1. We use the fundamental lemma of sieve theory to remove this condition in Section 1.5, and this yields the desired main term.

We estimate the bilinear form $B(x; U, V)$ in Sections 1.6 through 1.10. In Section 1.6 we perform some technical reductions like separating variables. These reductions allow us to enter the Gaussian domain $\mathbb{Z}[i]$ in Section 1.7. A congruence modulo Δ arises, and this introduces further complications. We address many of these in Section 1.8. 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 1.9 we are mostly able to remove the congruence modulo Δ, which simplifies our working considerably. With the congruence removed we devote Section 1.10 to extracting cancellation from the sign changes of the Möbius function using the theory of Hecke L-functions. Theorem 0.1.1 follows from (1.3.1), (1.4.17), (1.5.8), (1.6.1), and Proposition 1.6.1.

In the last section of the chapter, Section 1.11, we show how to modify the proof of

Theorem 0.1.1 to prove Theorem 1.1.1.

1.3 The sieve remainder term

Our goal in this section is to show that

$$R(x;U,V) \ll x^{\frac{1}{2}+\frac{\gamma_0}{2}-\epsilon}, \tag{1.3.1}$$

provided $U,V > 2$ and $UV \leq x^{\gamma_0-\epsilon}$.

Applying the triangle inequality to (1.2.10), we get

$$|R(x;U,V)| \ll (\log x)R(x,UV) + \int_1^x R(t,V)\frac{dt}{t}. \tag{1.3.2}$$

The following is the key result we use to estimate remainder terms.

Lemma 1.3.1. *For $1 \leq D \leq x$ and $\epsilon > 0$ we have*

$$R(x,D) = \sum_{d \leq D} |R_d(x)| \ll D^{\frac{1}{4}} x^{\frac{1}{2}+\frac{\gamma_0}{4}+\epsilon},$$

the implied constant depending only on ϵ.

Proof. This is a specialization of [3, Lemma 4]. In the notation of [3], we take $\lambda_\ell = \mathbf{1}_{\mathcal{A}}(\ell)$ for $\ell \leq x^{1/2}$. We then observe that

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

the last inequality following by (0.1.1). □

With Lemma 1.3.1 in hand we can show the contribution from (1.3.2) is sufficiently small. The contribution from $R(x,UV)$ is negligible provided

$$UV \leq x^{\gamma_0-\delta}, \tag{1.3.3}$$

where $\delta > 0$ is any small fixed quantity. We henceforth assume (1.3.3). We can also immediately estimate the part of the integral with $t \geq V$:

$$\int_V^x R(t,V)\frac{dt}{t} \ll \int_V^x V^{\frac{1}{4}} t^{\frac{1}{2}+\frac{\gamma_0}{4}+\epsilon}\frac{dt}{t} \ll V^{\frac{1}{4}} x^{\frac{1}{2}+\frac{\gamma_0}{4}+\epsilon}. \tag{1.3.4}$$

This is sufficiently small provided $V \leq x^{\gamma_0-\delta}$, which already follows from (1.3.3) since $U > 2$. To show (1.3.1) it therefore suffices to prove

$$\int_1^V R(t,V)\frac{dt}{t} \ll V^{\frac{1}{2}+\frac{\gamma_0}{2}+\epsilon}. \tag{1.3.5}$$

We 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

$$\begin{aligned}\sum_{d\leq V} A_d(t) &\leq \sum_{d\leq V} \sum_{\substack{m^2+\ell^2\leq t \\ m^2+\ell^2\equiv 0(d)}}\sum \mathbf{1}_{\mathcal{A}}(\ell) \leq \sum_{m^2+\ell^2\leq t}\sum \mathbf{1}_{\mathcal{A}}(\ell)\tau(m^2+\ell^2) \\ &\ll t^{\epsilon} \sum_{m^2+\ell^2\leq t}\sum \mathbf{1}_{\mathcal{A}}(\ell) \ll t^{\frac{1}{2}+\frac{\gamma_0}{2}+\epsilon}.\end{aligned} \tag{1.3.6}$$

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 1.3.2. *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 [3, (3.4)]. □

Observe that Lemma 1.3.2 implies

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

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

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

The following lemma illustrates how we estimate sums involving ρ_ℓ.

Lemma 1.3.3. *Let $y \ge 2$, and let ℓ be an integer. Then*

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

Proof. We factor n as $n = em$, where $e \mid \ell^\infty$ and m is coprime to ℓ. By multiplicativity and Lemma 1.3.2 we obtain

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

We use multiplicativity and Lemma 1.3.2 again to obtain

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

□

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

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

We apply Lemma 1.3.3 and obtain

$$\sum_{d\le V} M_d(t) \ll \sum\sum_{m^2+\ell^2\le t}\mathbf{1}_{\mathcal{A}}(\ell)(\log V)^2\tau(\ell) \ll t^{\frac{1}{2}+\frac{\gamma_0}{2}}(tV)^\epsilon, \tag{1.3.7}$$

and combining this with our bound (1.3.6) yields (1.3.5).

1.4 The sieve main term

In this section we begin to show how $M(x;U,V)$ yields the main term for Theorem 0.1.1: we show that $M(x;U,V)$ is equal to

$$\frac{4}{\pi}C \sum\sum_{\substack{m^2+\ell^2\le x\\(\ell,\Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell),$$

up to negligible error. The estimates involved are standard, but we give details for completeness.

From (1.2.9) we derive

$$M(x;U,V) = \sum\sum_{\substack{g^2+\ell^2\le x\\(\ell,\Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell)\Bigg(\log(g^2+\ell^2)\sum_{d\le V}\frac{\mu(d)\rho_\ell(d)}{d} - \sum_{d\le V}\frac{\mu(d)\rho_\ell(d)\log d}{d} - \sum_{m\le U}\frac{\Lambda(m)}{m}\sum_{d\le V}\frac{\mu(d)\rho_\ell(dm)}{d}\Bigg). \tag{1.4.1}$$

The main term arises from the second term on the right side of (1.4.1), and the other two terms contribute only to the error.

We begin by estimating

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

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

$$\sum_{p\le z}\frac{\rho_\ell(p)}{p}=\log\log z+c_\ell+O_\ell(\exp(-c\sqrt{\log z})),$$

for some constant c_ℓ depending on ℓ. By [56, (2.4)], this implies

$$\sum_{d=1}^{\infty}\frac{\mu(d)\rho_\ell(d)}{d}=0. \tag{1.4.2}$$

From (1.4.2) and partial summation it follows that

$$\begin{aligned}\sum_{d\le V}\frac{\mu(d)\rho_\ell(d)}{d} &= -\sum_{d>V}\frac{\mu(d)\rho_\ell(d)}{d}=\lim_{H\to\infty}\left(-\sum_{V<d\le H}\frac{\mu(d)\rho_\ell(d)}{d}\right) \qquad (1.4.3)\\ &= \lim_{H\to\infty}\left(-H^{-1}\sum_{d\le H}\mu(d)\rho_\ell(d)V^{-1}\sum_{d\le V}\mu(d)\rho_\ell(d)+\int_V^H\frac{1}{t^2}\left(\sum_{d\le t}\mu(d)\rho_\ell(d)\right)dt\right).\end{aligned}$$

We will show

$$\sum_{d\le z}\mu(d)\rho_\ell(d)\ll\prod_{p\mid\ell}\left(1+\frac{26}{p^{2/3}}\right)\ z\exp(-c\sqrt{\log z}), \tag{1.4.4}$$

uniformly in ℓ and $z\ge 1$. The bound 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 satisfies $g^{(j)}\ll_j y^{-j}$. Estimating trivially,

$$\sum_{d\le z}\mu(d)\rho_\ell(d)=O(y\log z)+\sum_{d}\mu(d)\rho_\ell(d)g(d). \tag{1.4.5}$$

Mellin inversion yields

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

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

$$\widehat{g}(s) \ll z^\sigma \left(1 + (y/z)^2 t^2\right)^{-1}, \tag{1.4.6}$$

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 (1.4.6) to $\sigma = 1 + \frac{1}{\log z}$ and estimate trivially the contribution from $|t| \geq T$, with T a parameter to be chosen. This gives

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

For $|t| \leq T$ we move the line of integration to $\sigma = 1 - \frac{c}{\log T}$, where c is chosen small enough that $\zeta(s)^{-1}L(s,\chi)^{-1}$ has no zeros in $\sigma \geq 1 - \frac{c}{\log T}$, $|t| \leq T$, and add in horizontal connecting lines. We estimate everything trivially to find this error is

$$\ll z(\log zT)^{O(1)} \exp(2b\sqrt{\log z}) \prod_{p|\ell}\left(1 + \frac{\chi^2(p)}{p^{2/3} - 1 - \chi^2(p)}\right)\left(\frac{1}{T} + \exp\left(-c\frac{\log z}{\log T}\right)\right).$$

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

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

and this completes the proof of (1.4.4).

The fact that ℓ is coprime to Π implies

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

From (1.4.3) we see that (1.4.4) and $(\ell, \Pi) = 1$ yield

$$\sum_{d\le V}\frac{\mu(d)\rho_\ell(d)}{d} \ll \exp(-c\sqrt{\log V}). \tag{1.4.7}$$

This shows that the first term of (1.4.1) satisfies the bound

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

provided

$$V \ge x^{\delta}$$

for some absolute constant $\delta > 0$.

We turn to estimating

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

We add and subtract the quantity

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

which yields

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

by (1.4.7). From Perron's formula we obtain

$$\sum_{d\le V}\frac{\mu(d)\rho_\ell(d)}{d}\log(V/d)=\frac{1}{2\pi i}\int_{(1)}\frac{x^s}{s^2}\sum_{d=1}^{\infty}\frac{\mu(d)\rho_\ell(d)}{d^{1+s}}ds. \tag{1.4.8}$$

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 (1.4.4), but here there is a main term coming from the simple pole of the integrand in (1.4.8) at $s=0$. Since $L(1,\chi)=\frac{\pi}{4}$, we deduce

$$\begin{aligned}-\sum_{d\le V}\frac{\mu(d)\rho_\ell(d)\log d}{d}&=\frac{4}{\pi}\prod_{p|\ell}\left(1+\frac{\chi(p)}{p-1-\chi(p)}\right)\\&\quad\times\prod_{p}\left(1-\frac{\chi(p)}{(p-1)(p-\chi(p))}\right)+O(\exp(-c\sqrt{\log V})).\end{aligned} \tag{1.4.9}$$

The expression in (1.4.9) gives rise to the main term in Theorem 0.1.1.

The last term of $M(x;U,V)$ we estimate similarly to the first. We aim to show that

$$\sum_{m\le U}\frac{\Lambda(m)}{m}\sum_{d\le V}\frac{\mu(d)\rho_\ell(dm)}{d}\ll(\log \ell V)^3P^{-1/2}. \tag{1.4.10}$$

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

$$\begin{aligned}\sum_{m\le U}\frac{\Lambda(m)}{m}\sum_{d\le V}\frac{\mu(d)\rho_\ell(dm)}{d}&=\sum_{m\le U}\frac{\Lambda(m)\rho_\ell(m)}{m}\sum_{\substack{d\le V\\(d,m)=1}}\frac{\mu(d)\rho_\ell(d)}{d}\\&\quad-\sum_{p^k\le U}\frac{(\log p)\rho_\ell(p^{k+1})}{p^{k+1}}\sum_{\substack{e\le V/p\\(e,p)=1}}\frac{\mu(e)\rho_\ell(e)}{e}.\end{aligned} \tag{1.4.11}$$

It is not difficult to deal with the sum over d in the first term of (1.4.11) using an argument analogous to that which gave (1.4.7), as the condition $(d,m)=1$ causes no great complications. To bound the sum over m we use Lemma 1.3.3, obtaining

$$\sum_{m \le U} \frac{\Lambda(m)\rho_\ell(m)}{m} \le \sum_{\substack{m \le U \\ (m,\ell)=1}} \frac{\Lambda(m)\rho_\ell(m)}{m} + (\log U) \sum_{\substack{p^k \\ p \mid \ell}} \frac{\rho_\ell(p^k)}{p^k}$$
$$\ll \log U + (\log U) \sum_{\substack{p^k \\ p \mid \ell}} \frac{p^{k/2}}{p^k} \ll \log U.$$

The last inequality follows since $p \mid \ell$ implies $p > P$. Therefore

$$\sum_{m \le U} \frac{\Lambda(m)\rho_\ell(m)}{m} \sum_{\substack{d \le V \\ (d,m)=1}} \frac{\mu(d)\rho_\ell(d)}{d} \ll (\log U) \exp(-c\sqrt{\log V}). \tag{1.4.12}$$

We turn our attention to the second term of (1.4.11). We first remove those p that are not coprime to ℓ. By trivial estimation

$$\begin{aligned} &\sum_{\substack{p^k \le U \\ p \mid \ell}} \frac{(\log p)\rho_\ell(p^{k+1})}{p^{k+1}} \sum_{\substack{e \le V/p \\ (e,p)=1}} \frac{\mu(e)\rho_\ell(e)}{e} \\ &\ll (\log V)^2 \sum_{p \mid \ell} (\log p) \sum_{k=1}^{\infty} \frac{1}{p^{k/2}} \ll (\log \ell V)^3 P^{-1/2} \end{aligned} \tag{1.4.13}$$

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

To handle those p that are coprime to ℓ, we assume that

$$U \ge x^\delta$$

for some absolute constant $\delta > 0$. We then estimate trivially the contribution from $p > R = \exp(\sqrt{\log V})$. Observe that $R < U$. Then

$$\begin{aligned} &\sum_{\substack{p^k \le U \\ p > R \\ (p,\ell)=1}} \frac{(\log p)\rho_\ell(p^{k+1})}{p^{k+1}} \sum_{\substack{e \le V/p \\ (e,p)=1}} \frac{\mu(e)\rho_\ell(e)}{e} \ll (\log V)^2 \sum_{p > R} \log p \sum_{k=2}^{\infty} \frac{k}{p^k} \\ &\ll (\log V)^2 \sum_{p > R} \frac{\log p}{p^2} \ll \frac{(\log V)^2}{R}, \end{aligned} \tag{1.4.14}$$

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

$$\sum_{\substack{p^k \leq U \\ p \leq R \\ (p,\ell)=1}} \frac{(\log p)\rho_\ell(p^{k+1})}{p^{k+1}} \sum_{\substack{e \leq V/p \\ (e,p)=1}} \frac{\mu(e)\rho_\ell(e)}{e} \ll \exp(-c\sqrt{\log V}) \tag{1.4.15}$$

by arguing as before, since V/p is close to V in the logarithmic scale. Taking (1.4.12), (1.4.13), (1.4.14), and (1.4.15) together gives (1.4.10). We combine (1.4.7), (1.4.9), and (1.4.12) to derive

$$\begin{aligned} M(x;U,V) = &\frac{4}{\pi} C \sum\sum_{\substack{m^2+\ell^2 \leq x \\ (\ell,\Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \prod_{p \mid \ell} \left(1 + \frac{\chi(p)}{p-1-\chi(p)}\right) \\ &+ O\left((\log x)^3 x^{\frac{1}{2}+\frac{\gamma_0}{2}} P^{-1/2}\right), \end{aligned} \tag{1.4.16}$$

provided $U, V \geq x^\delta$ for some absolute constant $\delta > 0$. Here

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

is the constant in Theorem 0.1.1. Since $P^-(\ell) > P$ we have

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

and so (1.4.16) becomes

$$M(x;U,V) = \frac{4}{\pi} C \sum\sum_{\substack{m^2+\ell^2 \leq x \\ (\ell,\Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) + O\left((\log x)^3 x^{\frac{1}{2}+\frac{\gamma_0}{2}} P^{-1/2}\right). \tag{1.4.17}$$

1.5 The sieve main term: fundamental lemma

We wish to simplify the main term of (1.4.17) by removing the condition $(\ell, \Pi) = 1$, which we accomplish with the fundamental lemma of sieve theory.

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

$$F_Y(\theta) := Y^{-\log 9/\log 10} \left| \sum_{0 \le n < Y} \mathbf{1}_{\mathcal{A}}(n) e(n\theta) \right|, \tag{1.5.1}$$

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.

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 (1.2.1)). From the definition (1.5.1) it is not difficult to derive (see [9, p. 6]) the identity

$$F_{UV}(\theta) = F_U(\theta) F_V(U\theta). \tag{1.5.2}$$

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 1.5.1. *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 [9, Lemma 10.1]. □

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

Lemma 1.5.2. *For $q \ge 1$ we have*

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

Proof. This is a slight weakening of the first part of [9, Lemma 10.5]. □

Lemma 1.5.3. *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 [9, Lemma 10.5]. □

Now that the necessary results are in place, we proceed with the estimation of the main term in (1.4.17). We write

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

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 [57, Chapter 5] for terminology). Therefore

$$\sum_{\substack{d\leq D\\d|\Pi\\(d,10)=1}} \lambda_d^- \sum_{\substack{\ell\leq z\\\ell\equiv 0(d)\\(\ell,10)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \leq \sum_{\substack{\ell\leq z\\(\ell,\Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \leq \sum_{\substack{d\leq D\\d|\Pi\\(d,10)=1}} \lambda_d^+ \sum_{\substack{\ell\leq z\\\ell\equiv 0(d)\\(\ell,10)=1}} \mathbf{1}_{\mathcal{A}}(\ell). \tag{1.5.3}$$

The upper and lower bounds turn out to be asymptotically equal, and we 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 [12]). 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

$$\sum_{\substack{\ell \le z \\ \ell \equiv 0 (d) \\ (\ell,10)=1}} \mathbf{1}_{\mathcal{A}}(\ell) = \sum_{n \in S(z)} \sum_{\substack{nY \le \ell < (n+1)Y \\ \ell \equiv 0(d) \\ (\ell,10)=1}} \mathbf{1}_{\mathcal{A}}(\ell) + O\left(\sum_{\substack{z-Y \le \ell \le z+Y \\ \ell \equiv 0(d)}} \mathbf{1}_{\mathcal{A}}(\ell) \right). \tag{1.5.4}$$

Here $S(z)$ is some set of size $\ll P$. We remark that we will repeatedly see this technique of breaking an interval into shorter subintervals, with each subinterval having length a power of 10, in the estimation of the bilinear sum $B(x; U, V)$.

We first illustrate how to use Lemma 1.5.3 to handle the error term in (1.5.4). On summing over d, we must estimate

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

Because the estimation of $\mathcal{E}$ introduces a number of important ideas that we use throughout the proof of Theorem 0.1.1, we encapsulate the estimation in a lemma.

Lemma 1.5.4. *With the notation as above,*

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

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

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

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 $0 \le t < X$. 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 $0 \le t < X$. If $a_0 = 0$ then $\mathbf{1}_{\mathcal{A}}(kX+t) = 0$ for $0 \le t < X/10$ and $\mathbf{1}_{\mathcal{A}}(kX+t) = \mathbf{1}_{\mathcal{A}}(t)$ for $X/10 \le t < X$. We can unite the two subcases by writing

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

where

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

By inclusion-exclusion and the triangle inequality we find

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

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

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

By the triangle inequality,

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

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

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

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

$$\begin{aligned} \sum_{1 < d \le D} \frac{1}{d} \sum_{r=1}^{d-1} F_X\left(\frac{r}{d}\right) &= \sum_{1 < d \le D} \frac{1}{d} \sum_{\substack{q \mid d \\ q > 1}} \sum_{\substack{1 \le b \le q \\ (b,q)=1}} F_X\left(\frac{b}{q}\right) \\ &\ll (\log D) \sum_{1 < q \le D} \frac{1}{q} \sum_{\substack{1 \le b \le q \\ (b,q)=1}} F_X\left(\frac{b}{q}\right). \end{aligned}$$

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

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

By Lemma 1.5.3,

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

and this completes the proof. □

From Lemma 1.5.4 it follows that

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

We detect the congruence with additive characters and obtain

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

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

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

For $1 \le r \le d-1$ we handle the condition $(t,10)=1$ with Möbius inversion. We then reverse our short interval decomposition to get

$$\sum_{\substack{d\le D\\ d|\Pi\\ (d,10)=1}} \lambda_d \sum_{\substack{\ell\le z\\ \ell\equiv 0(d)\\ (\ell,10)=1}} \mathbf{1}_{\mathcal{A}}(\ell) = \kappa \sum_{\substack{d\le D\\ d|\Pi\\ (d,10)=1}} \frac{\lambda_d}{d}\sum_{\ell\le z}\mathbf{1}_{\mathcal{A}}(\ell) + O\left(x^{\gamma_0/2}\left(P^{-1/2}+EP^{\gamma_0}\right)\right), \quad (1.5.5)$$

where

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

and X is a power of 10 with $X \asymp Y$. Similarly to the estimation of $\mathcal{E}$ in Lemma 1.5.4 above, we put, for pedagogical reasons, the estimation of E into a lemma.

Lemma 1.5.5. *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

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

We apply (1.5.2) 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 $10a \to a$ to obtain

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

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

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

Set $Q_1 = \exp(\varepsilon\sqrt{\log z})$, where $\varepsilon > 0$ is a small positive constant. If $Q > Q_1$ we use Lemma 1.5.3, and if $Q \le Q_1$ we use Lemma 1.5.1. Provided ε in the definition of Q_1 is taken sufficiently small in terms of c_0 in Lemma 1.5.1, we obtain

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

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

We take (1.5.5) with Lemma 1.5.5, along with the fact that $\log P = o(\sqrt{\log z})$, to get

$$\sum_{\substack{d \le D \\ d \mid \Pi \\ (d,10)=1}} \lambda_d \sum_{\substack{\ell \le z \\ \ell \equiv 0(d) \\ (\ell,10)=1}} \mathbf{1}_{\mathcal{A}}(\ell) = \kappa \sum_{\substack{d \le D \\ d \mid \Pi \\ (d,10)=1}} \frac{\lambda_d}{d} \sum_{\ell \le z} \mathbf{1}_{\mathcal{A}}(\ell) + O\left(x^{\gamma_0/2} P^{-1/2}\right). \tag{1.5.6}$$

By the fundamental lemma of the linear sieve (see [57, Lemma 6.11])

$$\sum_{\substack{d \le D \\ d \mid \Pi \\ (d,10)=1}} \frac{\lambda_d}{d} = \left(1 + O\left(\exp\left(-\frac{1}{2} s \log s\right)\right)\right) \prod_{\substack{p \le P \\ p \nmid 10}} \left(1 - \frac{1}{p}\right), \tag{1.5.7}$$

where

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

It follows that

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

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 $c>0$. Observe that our lower bound for $\log P$ implies

$$\exp\left(-c\sqrt{\log P}\right)\leq \exp\left(-c(\log\log x)^2\right),$$

so this error term is acceptable for Theorem 0.1.1. Therefore,

$$\sum_{\substack{m^2+\ell^2\leq x\\ (\ell,\Pi)=1}}\sum \mathbf{1}_{\mathcal{A}}(\ell)=\frac{10}{\varphi(10)}\kappa\frac{e^{-\gamma}}{\log P}\sum_{m^2+\ell^2\leq x}\sum \mathbf{1}_{\mathcal{A}}(\ell)+O\left(x^{\frac{1}{2}+\frac{\gamma_0}{2}}\exp\left(-c\sqrt{\log P}\right)\right). \quad (1.5.8)$$

Combining (1.5.8) with (1.4.17) yields the main term of Theorem 0.1.1.

1.6 Bilinear form in the sieve: first steps

Let us summarize what we have accomplished thus far. We take (1.2.3), (1.2.8), (1.3.1), (1.4.17), and (1.5.8) to obtain

$$\begin{aligned} S(x)=\frac{4C\kappa_1}{\pi}\frac{e^{-\gamma}}{\log P}\sum_{m^2+\ell^2\leq x}\sum \mathbf{1}_{\mathcal{A}}(\ell)+B(x;U,V)\\ +O\left(x^{\frac{1}{2}+\frac{\gamma_0}{2}}\exp\left(-c\sqrt{\log P}\right)\right). \end{aligned} \quad (1.6.1)$$

This holds provided our parameters U,V in Vaughan's identity (1.2.1) satisfy

$$UV\leq x^{\gamma_0-\delta}, \qquad U,V\geq x^{\delta}, \quad (1.6.2)$$

for some absolute $\delta>0$. The task now is to show that the Type II sum $B(x;U,V)$ contributes only to the error term of (1.6.1). (We note that the implied constant in the error term of (1.6.1) is effectively computable.)

In the course of our estimations we encounter more severe restrictions on U,V than those in (1.6.2), and we note these as we go along. It transpires that there is some flexibility in choosing U and V. For those unwilling to wait in suspense, we mention that the choices

$$U = x^{7/10}, \quad V = x^{1/5}, \tag{1.6.3}$$

say, are acceptable for Theorem 0.1.1 (see (1.6.11) and (1.8.4)).

Recall that

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

We shall prove the following proposition.

Proposition 1.6.1. *Let x be large, and let $L > 0$ be fixed. There exist absolute constants $C, \omega > 0$ such that*

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

The implied constant is ineffective.

Observe that Proposition 1.6.1 implies Theorem 0.1.1 by taking $L = \omega^{-1}(A + C)$. To prove Theorem 0.1.1 it therefore suffices to prove Proposition 1.6.1, and this task occupies the remainder of the chapter. In what follows we allow implied constants to depend on L without indicating it in the notation.

In this section we perform technical reductions that reduce the estimation of $B(x; U, V)$ to the estimation of sums of the form

$$\sum_m \left| \sum_{(n, m\Pi) = 1} \mu(n) a(mn) \right|. \tag{1.6.4}$$

Here the intervals of summation of m and n are independent of one another. This separation of variables is accomplished by a short interval decomposition. Once m and n are separated, we remove the small prime factors of n and transfer them to m. This has the immediate benefit of insuring that m and n are almost always coprime, but also confers substantial technical advantages in later calculations. The contribution from those m and n satisfying $(m, n) > 1$ is then trivially estimated and shown to be negligible, which gives the reduction to (1.6.4) (see (1.6.21)).

We mention that the arguments in this section have some similarity to those in [4, Section 10] and [56, Section 4].

Since $(\Lambda_{>U} \star 1)(m) \leq (\Lambda \star 1)(m) = \log m$, we see

$$B(x;U,V) \leq (\log x) \sum_{U<m\leq x/V} \left| \sum_{V<n\leq x/m} \mu(n)a(mn) \right|. \tag{1.6.5}$$

It is easy to obtain a trivial bound for $B(x;U,V)$ that is not far from the bound of Proposition 1.6.1.

Lemma 1.6.2. *Let $x \geq 2$. Then*

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

Proof. We change variables in (1.6.5) 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 1.3.1 to bound the sum of $R_d(x)$, giving

$$B(x;U,V) \ll (\log x) \sum\sum_{\substack{m^2+\ell^2\leq x \\ (\ell,\Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{d\leq x^{1/2}} \frac{\rho_\ell(d)}{d}. \tag{1.6.6}$$

By Lemma 1.3.3 we have

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

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

In the proof of Lemma 1.6.2 we used Lemma 1.3.3 to control averages of ρ_ℓ. We shall need more elaborate versions of this argument in several of our reductions of $B(x;U,V)$.

Our first step is to separate the variables m and n so that they run independently over intervals of summation. We accomplish this with a short interval decomposition.

Set

$$\theta = (\log x)^{-L}, \tag{1.6.7}$$

where $L > 0$ is fixed as in Proposition 1.6.1. We break the summation over n into subintervals $N < n \le (1+\theta)N$. By the triangle inequality,

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

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

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

and so for fixed N we have

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

In order to apply Lemma 1.3.1 we require $N \le x/U \ll x^{\gamma_0-\epsilon}$, which holds provided

$$U \ge x^{1-\gamma_0+\epsilon}. \tag{1.6.8}$$

This supersedes the lower bound for U in (1.6.2). Assuming (1.6.8) we obtain by Lemma 1.3.1

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

We have

$$\sum\sum_{x<m^2+\ell^2\le(1+\theta)x} \mathbf{1}_{\mathcal{A}}(\ell) = \sum_{\ell\le\sqrt{(1+\theta)x}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{\sqrt{x-\ell^2}<|m|\le\sqrt{(1+\theta)x-\ell^2}} 1$$
$$\ll \theta^{1/2}x^{1/2} \sum_{\ell\ll x^{1/2}} \mathbf{1}_{\mathcal{A}}(\ell) \ll \theta^{1/2}x^{\frac{1}{2}+\frac{\gamma_0}{2}}.$$

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

$$B(x;U,V) \ll (\log x) \sum_{\substack{N=(1+\theta)^j V\\ j\ge 0\\ N\le x/U}} \sum_{U<m\le x/N} \left| \sum_{N<n\le(1+\theta)N} \mu(n)a(mn) \right| + O((\log x)^3\theta^{1/2}x^{\frac{1}{2}+\frac{\gamma_0}{2}}). \tag{1.6.9}$$

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

$$\sum_{U<m\le x/N} \left| \sum_{N<n\le(1+\theta)N} \mu(n)a(mn) \right| \le \sum_{\substack{M=2^jU\\ j\ge 0\\ MN\le x}} \sum_{M<m\le 2M} \left| \sum_{N<n\le(1+\theta)N} \mu(n)a(mn) \right|.$$

We define

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

and note that the variables m and n are separated in $B_1(M,N)$.

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

$$B_1(M,N) \ll \sum_{N<n\le(1+\theta)N} \mu^2(n) \sum_{\substack{k\ll\theta x\\ n|k}} a(k) \ll (\log N)\theta^{1+\gamma_0}x^{\frac{1}{2}+\frac{\gamma_0}{2}}, \tag{1.6.10}$$

the latter inequality following essentially by the argument that gave (1.6.9). In order to prove Proposition 1.6.1 it suffices by virtue of (1.6.9) and (1.6.10) to prove the following result.

Proposition 1.6.3. *Let x be large, and for $L > 0$ fixed set $\theta = (\log x)^{-L}$. We then* *have*

$$B_1(M,N) \ll_{\epsilon,L} (\log MN)^{O(1)}\theta^{5/2}(MN)^{\frac{1}{2}+\frac{\gamma_0}{2}}$$

uniformly in

$$x^{1/2-\gamma_0/2+\epsilon} \le N \le x^{25/77-\epsilon}, \qquad \theta x < MN \le x. \tag{1.6.11}$$

The implied constant is ineffective.

It is not yet apparent why N must be of the size given in (1.6.11). We gradually introduce stronger conditions on N as the proof requires, and find in the last instance that (1.6.11) is sufficient.

Now that the variables m and n are separated from one another, we wish to remove the small prime factors from n. We write $n = n_0n_1$, where $(n_0, \Pi) = 1$ and $n_1 \mid \Pi$, then set

$$C = \exp((\log P)^2)$$

(there should be no cause to confuse the C given here with the absolute constant C in Proposition 1.6.1). Observe that $C > P$, and that $C = x^{o(1)}$ by our upper bound for $\log P$ in Theorem 0.1.1.

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

$$\begin{aligned}\sum_{M<m\le 2M} \left| \sum_{\substack{N<n_0n_1\le(1+\theta)N \\ (n_0,\Pi)=1 \\ n_1\mid\Pi \\ n_1>C}}\sum \mu(n_0)\mu(n_1)a(mn_0n_1) \right| &\le \sum_{\substack{C<d\le CP \\ d\mid\Pi}} \sum_{\substack{n\ll MN \\ d\mid n}} a(n)\tau_3(n) \\ &\le \sum_{\substack{C<d\le CP \\ d\mid\Pi}} \sum_{\substack{n\ll MN \\ d\mid n}} a(n)\tau(n)^2 =: B_1',\end{aligned}$$

say. We utilize the following lemma to handle the divisor function.

Lemma 1.6.4. *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 [8, Lemma 4]. □

Applying Lemma 1.6.4 with $k = 2$ yields

$$B_1' \ll \sum_{\substack{C<d\leq CP \\ d|\Pi}} \sum_{e \ll (MN)^{1/4}} \tau(e)^8 \sum_{\substack{n \ll MN \\ [d,e]|n}} a(n),$$

where $[d, e]$ is the least common multiple of d and e. By trivial estimation (i.e. no need for recourse to Lemma 1.3.1 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

$$\begin{aligned} B_1' &\ll (MN)^{1/2} \sum_{\substack{\ell \ll (MN)^{1/2} \\ (\ell,\Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{\substack{C<d\leq CP \\ d|\Pi}} \sum_{e \ll (MN)^{1/4}} \frac{\tau(e)^8 \rho_\ell([d,e])}{[d,e]} \\ &\ll (MN)^{1/2} \sum_{\substack{\ell \ll (MN)^{1/2} \\ (\ell,\Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{\substack{C<d\leq CP \\ d|\Pi}} \sum_{\substack{r \ll (MN)^{1/4} \\ P^+(r)\leq P}} \frac{\tau(r)^8 \rho_\ell([d,r])}{[d,r]} \sum_{\substack{s \ll (MN)^{1/4} \\ P^-(s)>P}} \frac{\tau(s)^8 \rho_\ell(s)}{s}. \end{aligned}$$

We bound the sum over s by working as in Lemma 1.3.3. We have

$$\begin{aligned} \sum_{\substack{s \ll (MN)^{1/4} \\ P^-(s)>P}} \frac{\tau(s)^8 \rho_\ell(s)}{s} &\leq \sum_{s \ll (MN)^{1/4}} \frac{\tau(s)^8 \rho_\ell(s)}{s} \\ &\leq \sum_{d|\ell^\infty} \frac{\tau(d)^8 \rho_\ell(d)}{d} \sum_{\substack{t \ll (MN)^{1/4} \\ (t,\ell)=1}} \frac{\tau(t)^9}{t} \qquad (1.6.12) \\ &\ll (\log MN)^{2^9} \prod_{p|\ell} \left(1 + \frac{2^9}{p^{1/2}}\right) \ll (\log MN)^{2^9}. \end{aligned}$$

By (1.6.12) and the change of variables $n = [d, r]$, we obtain

$$B_1' \ll (\log MN)^{2^9} \sum_{\substack{\ell \ll (MN)^{1/2} \\ (\ell,\Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{\substack{n>C \\ P^+(n)\le P}} \frac{\tau(n)^8 \tau_3(n) \rho_\ell(n)}{n}.$$

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

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

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

$$C^{-\varepsilon} \prod_{p\le P} \left(1 + \frac{2^{14}}{p^{1+\varepsilon}}\right) \le C^{-\varepsilon} \zeta(1+\varepsilon)^{2^{14}} \ll C^{-\varepsilon} \varepsilon^{-2^{14}} \le (\log MN)^{2^{13}} P^{-1}.$$

We deduce

$$B_1' \ll (\log MN)^{2^{14}} (MN)^{\frac{1}{2} + \frac{\gamma_0}{2}} P^{-1}. \tag{1.6.13}$$

By (1.6.13) and our lower bound for P the contribution from $n_1 > C$ is acceptable for Proposition 1.6.3. It follows that

$$B_1(M, N) \ll \theta^{5/2} (MN)^{\frac{1}{2} + \frac{\gamma_0}{2}} + \sum_{M < m \le 2M} \sum_{\substack{n_1 \le C \\ n_1 | \Pi}} \left| \sum_{\substack{N < n_0 n_1 \le (1+\theta)N \\ (n_0,\Pi)=1}} \mu(n_0) a(m n_0 n_1) \right|. \tag{1.6.14}$$

We wish 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

$$\sum_{\substack{G=(1+\theta^{5/2})^j\\ j\geq -1\\ G\leq C}} \sum_{M<m\leq 2M} \sum_{\substack{G<n_1\leq(1+\theta^{5/2})G\\ n_1\mid\Pi}} \left| \sum_{\substack{N<n_0n_1\leq(1+\theta)N\\ (n_0,\Pi)=1}} \mu(n_0)a(mn_0n_1) \right|.$$

In the sum over n_0 we wish to replace the conditions $N < n_0n_1$ and $n_0n_1 \leq (1+\theta)N$ by the conditions $N < n_0G$ and $n_0G \leq (1+\theta)N$, respectively. If $n_0n_1 > N$ but $n_0G \leq N$, then

$$N < n_0n_1 \leq (1+\theta^{5/2})n_0G \leq (1+\theta^{5/2})N,$$

and the error in replacing the condition $n_0n_1 > N$ by $n_0G > N$ is

$$\ll (\log C)\theta^{-5/2} \sup_{G\leq C} \sum_{G<n_1\leq(1+\theta^{5/2})G} \mu^2(n_1)$$
$$\sum_{(1+\theta^{5/2})^{-1}N/G<n_0\leq(1+\theta^{5/2})N/G} \mu^2(n_0) \sum_{\substack{n\leq 3MN\\ n_0n_1\mid n}} a(n).$$

We write these three sums as

$$\sum_{n_1}\sum_{n_0} A_{n_0n_1}(3MN) = \sum_{n_1}\sum_{n_0} \left(M_{n_0n_1}(3MN) + R_{n_0n_1}(3MN)\right).$$

To estimate the remainder term we change variables

$$\sum_{n_1}\sum_{n_0} |R_{n_0n_1}(3MN)| \leq \sum_{d} \tau(d)|R_d(3MN)|,$$

then apply the divisor bound and Lemma 1.3.1. We estimate the main term as we have before, and find that

$$\sum_{n_1}\sum_{n_0} A_{n_0n_1}(3MN) \ll (\log MN)^{O(1)}\theta^{5/2}(MN)^{\frac{1}{2}+\frac{\gamma_0}{2}}.$$

We similarly acquire the condition $n_0G \leq (1+\theta)N$. We then change variables $mn_1 \to$

$m, n_0 \to n$ to obtain

$$B_1(M,N) \ll (\log MN)^{O(1)}\theta^{5/2}(MN)^{\frac{1}{2}+\frac{\gamma_0}{2}} + (\log MN)\theta^{-5/2} \times \sup_{G\le C} \sum_{MG<m\le 2(1+\theta^{5/2})MG} \tau(m) \left| \sum_{\substack{N/G<n_0\le(1+\theta)N/G \\ (n,\Pi)=1}} \mu(n)a(mn) \right|. \tag{1.6.15}$$

In order to prove Proposition 1.6.3 it therefore suffices to show that

$$\begin{aligned} B_2(M,N) &:= \sum_{M<m\le 2M} \tau(m) \left| \sum_{\substack{N<n\le(1+\theta)N \\ (n,\Pi)=1}} \mu(n)a(mn) \right| \\ &\ll \theta^5 (\log MN)^{O(1)}(MN)^{\frac{1}{2}+\frac{\gamma_0}{2}} \end{aligned} \tag{1.6.16}$$

uniformly in

$$x^{1/2-\gamma_0/2+\epsilon} \le N \le x^{25/77-\epsilon}, \qquad \theta x \ll MN \ll x. \tag{1.6.17}$$

Note the slight (inconsequential) difference between (1.6.17) and (1.6.11).

We have removed the small prime factors from n. This will aid us in making m and n coprime, which in turn 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 is more difficult to deal with later. By the Cauchy-Schwarz inequality

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

where

$$B_3(M,N) := \sum_{M<m\le 2M} \left| \sum_{\substack{N<n\le(1+\theta)N \\ (n,\Pi)=1}} \mu(n)a(mn) \right|,$$

$$B_2'(M,N) := \sum_{M<m\le 2M} \tau(m)^2 \left| \sum_{\substack{N<n\le(1+\theta)N \\ (n,\Pi)=1}} \mu(n)a(mn) \right|.$$

We bound $B_2'(M,N)$ trivially.

Lemma 1.6.5. *We have* $B_2'(M,N) \ll \theta(\log MN)^{2^{23}}(MN)^{\frac{1}{2}+\frac{\gamma_0}{2}}$.

Proof. We have the trivial bound

$$B_2'(M,N) \ll \sum_{N<n\leq(1+\theta)N} \mu^2(n) \sum_{\substack{k\leq 3MN \\ d|k}} a(k)\tau(k)^2.$$

We impose here a more severe condition on U, and therefore N, than (1.6.8). We require

$$U \geq x^{1/2}, \tag{1.6.18}$$

which implies $N \ll x^{1/2}$. Stricter conditions than (1.6.18) are imposed later, so there is no loss in imposing this condition now. We apply Lemma 1.6.4 with $k=2$ to arrive at

$$\begin{aligned} B_2'(M,N) &\ll \sum_{N<n\leq(1+\theta)N} \sum_{e\ll(MN)^{1/4}} \tau(e)^8 \sum_{\substack{k\leq 3MN \\ [n,e]|k}} a(k) \\ &= \sum_{N<n\leq(1+\theta)N} \sum_{e\ll(MN)^{1/4}} \tau(e)^8 \left(M_{[n,e]}(3MN) + R_{[n,e]}(3MN)\right). \end{aligned}$$

The contribution from the remainder terms is

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

and since $Nx^{1/4} \leq x^{\gamma_0-\epsilon}$ we may bound the remainder terms with Lemma 1.3.1.

We estimate the main term using the same types of arguments that gave (1.6.13). We factor $n=bd$ and $e=rs$ to bound the main term by

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

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

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

If $b\leq N^{1/2}$ we use Lemma 1.6.4 with $k=1$ to deduce

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

the last inequality following from the lower bound (1.6.2). 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

$$\sum_{\substack{n>N^{1/2}\\ n\mid\ell^\infty}} \frac{\tau(n)^8\tau_3(n)\rho_\ell(n)}{n} \ll N^{-1/4}\prod_{p\mid\ell}\left(1+\frac{2^{12}}{p^{1/4}}\right).$$

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

$$B_2'(M,N) \ll (\log MN)^{2^{23}}\theta(MN)^{\frac{1}{2}+\frac{\gamma_0}{2}},$$

as desired. □

By Lemma 1.6.5 we see that in order to prove (1.6.16) it suffices to show that

$$B_3(M,N) \ll \theta^9(\log MN)^{O(1)}(MN)^{\frac{1}{2}+\frac{\gamma_0}{2}}. \tag{1.6.19}$$

We are finally 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

$$B_3'(M,N) := \sum_{M<m\le 2M}\left|\sum_{\substack{N<n\le(1+\theta)N\\ (n,\Pi)=1\\ (n,m)\neq 1}}\mu(n)a(mn)\right| \ll \sum_{P<p\ll(MN)^{1/2}}\ \sum_{\substack{k\ll MN\\ p^2\mid k}} a(k)\tau(k).$$

We trivially estimate the contribution from $p>(MN)^{1/10}$ using the bound $a(k)\tau(k)\ll_\epsilon (MN)^\epsilon$. Thus

$$\begin{aligned} B_3'(M,N) &\ll (MN)^{9/10+\epsilon} + \sum_{P<p\le(MN)^{1/10}}\ \sum_{\substack{k\ll MN\\ p^2\mid k}} a(k)\tau(k)\\ &\ll \sum_{P<p\le(MN)^{1/10}}\ \sum_{d\ll(MN)^{1/2}}\ \sum_{\substack{k\ll MN\\ [d,p^2]\mid k}} a(k). \end{aligned}$$

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

p^2 divides d), we find that

$$B_3'(M,N) \ll \sum_{P<p\le (MN)^{1/10}} \sum_{d\ll (MN)^{1/2}} \sum_{\substack{k\ll MN \\ dp^2 | k}} a(k).$$

We apply Lemma 1.3.1 to deduce

$$\begin{aligned}
B_3'(M,N) &\ll (MN)^{1/2} \sum_{\substack{\ell \ll (MN)^{1/2} \\ (\ell,\Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{P<p\le (MN)^{1/10}} \frac{1}{p^2} \sum_{d\ll (MN)^{1/2}} \frac{\rho_\ell(dp^2)}{d} \\
&\ll (MN)^{1/2} \sum_{\substack{\ell \ll (MN)^{1/2} \\ (\ell,\Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{p>P} \frac{1}{p^2} \sum_{k=0}^{\infty} \sum_{\substack{d\ll (MN)^{1/2}/p^k \\ (d,p)=1}} \frac{\rho_\ell(dp^{k+2})}{dp^k} \\
&\ll (\log MN)^2 (MN)^{1/2} \sum_{\substack{\ell \ll (MN)^{1/2} \\ (\ell,\Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{p>P} \frac{1}{p^2} \sum_{k=0}^{\infty} \frac{\rho_\ell(p^{k+2})}{p^k}.
\end{aligned}$$

In going from the second line to the third line we have used Lemma 1.3.3 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

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

where we have used Lemma 1.3.2 to control the behavior of $\rho_\ell(p^{k+2})$. It follows that

$$\begin{aligned}
B_3(M,N) &\ll (\log MN)^{O(1)} (MN)^{\frac{1}{2}+\frac{\gamma_0}{2}} P^{-1} \\
&\quad + \sum_{M<m\le 2M} \left| \sum_{\substack{N<n\le (1+\theta)N \\ (n,m\Pi)=1}} \mu(n) a(mn) \right|.
\end{aligned} \tag{1.6.20}$$

In order to prove (1.6.19) it therefore suffices to show that

$$B_4(M,N) := \sum_{M<m\le 2M} \left| \sum_{\substack{N<n\le(1+\theta)N \\ (n,m\Pi)=1}} \mu(n)a(mn) \right| \ll \theta^9 (\log MN)^{O(1)} (MN)^{\frac{1}{2}+\frac{\gamma_0}{2}} \tag{1.6.21}$$

for M and N satisfying (1.6.17).

1.7 Bilinear form in the sieve: transformations

Now that m and n are coprime we are able to enter the realm of the Gaussian integers. This is the key step that allows us to estimate successfully the bilinear form $B_4(M,N)$ (see the discussion in [4, Section 5] for more insight on this). Since m and n are coprime the unique factorization in $\mathbb{Z}[i]$ gives

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

Since $(n,\Pi)=1$ we have $(z\overline{z},\Pi)=1$, 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

$$r \equiv 1 \pmod 2, \qquad s \equiv r-1 \pmod 4, \tag{1.7.1}$$

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

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

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 $(w\overline{w}, z\overline{z}) = 1$ which is a nuisance since we wish for w and z to range independently of one another. Because $z\overline{z}$ has no small prime factors, it suffices to estimate trivially the complimentary sum in which $(w\overline{w}, z\overline{z}) \neq 1$.

The arguments of this section bear some semblance to those in [4, Section 5] and [57, Section 20.4]. The plan of this section is as follows. We remove the condition $(w\overline{w}, z\overline{z}) = 1$ 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

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

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

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

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 $\overline{z_1} z_2$. 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.

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

Lemma 1.7.1. *With the notation as above, we have*

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

Proof. We find

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

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

Given fixed 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 $u \equiv \overline{r}\ell \pmod s$. If $p \mid \ell$ this gives the claim, so assume $p \nmid \ell$. Then $v \equiv \overline{s}(\ell - ru) \pmod p$, which yields

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

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 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

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

We also note that

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

$$\ll \frac{N\log N}{p}.$$

Therefore

$$\mathcal{B}_1'(M,N) \ll (\log N)(MN)^{1/2} \sum_{\ell\ll (MN)^{1/2}} \mathbf{1}_{\mathcal{A}}(\ell) \sum_{P<p\ll N} \frac{(\ell,p)}{p^2} + (\log N)^2 (MN)^{\gamma_0/2} N. \tag{1.7.2}$$

The second term is sufficiently small if $N \le x^{1/2-\epsilon}$, which is satisfied if

$$U \ge x^{1/2+\epsilon}. \tag{1.7.3}$$

This lower bound supersedes (1.6.18), and implies $M > N$ since $MN \gg \theta x$.

To bound the first term we note that

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

and this gives the bound

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

□

Lemma 1.7.1 proves that (1.6.21) follows from the bound

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

We now apply the Cauchy-Schwarz inequality, obtaining

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

where

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

Note that we have used positivity to extend the sum over w. It therefore suffices to show that

$$\mathcal{D}_1(M,N) \ll \theta^{18}(\log MN)^{O(1)}(MN)^{\gamma_0}N. \tag{1.7.4}$$

Expanding the square in $\mathcal{D}_2(M,N)$ 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

$$\mathcal{D}_1(M,N) \le D_1'(M,N) := \sum_{|w|^2\le 2M} \sum\sum_{N<|z_1|^2,|z_2|^2\le 2N} \mathbf{1}_{\mathcal{A}}(\mathrm{Re}(z_1\overline{w}))\mathbf{1}_{\mathcal{A}}(\mathrm{Re}(z_2\overline{w})).$$

Lemma 1.7.2. *For $M \ge N \ge 2$ we have*

$$D_1'(M,N) \ll \left((MN)^{\frac{1}{2}+\frac{\gamma_0}{2}} + (MN)^{\gamma_0}N\right)(\log MN)^{36}.$$

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

We can bound the diagonal terms by

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

By an argument similar to that which yielded (1.7.2), 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

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

and we deduce that

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

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

$$\Delta = \Delta(z_1,z_2) = \frac{1}{2i}\left(\overline{z_1}z_2 - z_1\overline{z_2}\right) \ne 0,$$

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

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

We note that the division takes place in the Gaussian integers, and that $\ell_1 z_2 - \ell_2 z_1 \ne 0$ (see (1.7.7) 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 $\ell_1 s_2 - \ell_2 s_1 \ne 0$. Given ℓ_1, ℓ_2, s_1, s_2, and $\Delta \ne 0$, 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 $\delta = (s_1, s_2)$ and $s_1 = \delta s_1^*, s_2 = \delta s_2^*$ (so that $(s_1^*, s_2^*) = 1$), we see that

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

Observe that $|\ell_1 s_2^* - \ell_2 s_1^*| \ll \sqrt{M}N$. We apply Lemma 1.6.4 with $k = 2$ to get

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

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

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

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

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

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

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

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 1.5.4). 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, a dyadic division, and Lemma 1.5.3, we find this last quantity is

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

Observe that

$$F = (\sqrt{M}N)^{1/4} \leq (\sqrt{MN})^{1/2} \ll (\sqrt{MN})^{50/77-\epsilon} \asymp X^{50/77-\epsilon},$$

which yields

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

□

With Lemma 1.7.2 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

$$\sum_{P<|\pi|^2\ll N} \sum_{|w|^2 \ll M} \sum\sum_{\frac{N}{|\pi|^2}<|z_1|^2, |z_2|^2 \leq \frac{2N}{|\pi|^2}} \mathbf{1}_{\mathcal{A}}(\mathrm{Re}(z_1\pi\overline{w}))\mathbf{1}_{\mathcal{A}}(\mathrm{Re}(z_2\pi\overline{w})).$$

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

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 $P < P_1 \ll N$. By Lemma 1.7.2 this bound becomes

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

The second term is satisfactorily small, and the first is sufficiently small provided

$$V \geq x^{1/2-\gamma_0/2+\epsilon}. \tag{1.7.5}$$

This lower bound for V supersedes the one in (1.6.2). In order to show (1.7.4) it then suffices to show that

$$\begin{aligned} \mathcal{D}_2(M,N) := \sum_{|w|^2 \leq 2M} \sum_{\substack{N < |z_1|^2, |z_2|^2 \leq (1+\theta)N \\ (|z_1|^2 |z_2|^2, \Pi) = 1 \\ (z_1, z_2) = 1 \\ (\mathrm{Re}(z_1 \overline{w}) \mathrm{Re}(z_2 \overline{w}), \Pi) = 1}} \sum \mu(|z_1|^2) \mu(|z_2|^2) \mathbf{1}_{\mathcal{A}}(\mathrm{Re}(z_1 \overline{w})) \mathbf{1}_{\mathcal{A}}(\mathrm{Re}(z_2 \overline{w})) \\ \ll \theta^{18} (\log MN)^{O(1)} (MN)^{\gamma_0} N \end{aligned} \tag{1.7.6}$$

for M and N satisfying (1.6.17). Since $Vx^{-o(1)} \ll N \ll x/U$ we see that (1.7.5) yields the lower bound on N in (1.6.17).

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

$$z_1 \overline{w} + \overline{z_1} w = 2\ell_1,$$
$$z_2 \overline{w} + \overline{z_2} w = 2\ell_2.$$

We set $\Delta = \Delta(z_1, z_2) = \mathrm{Im}(\overline{z_1} z_2) = \frac{1}{2i}(\overline{z_1} z_2 - z_1 \overline{z_2})$, and note that

$$iw\Delta = \ell_1 z_2 - \ell_2 z_1. \tag{1.7.7}$$

It follows that

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

Observe that the congruence is a congruence of Gaussian integers.

The contribution from $\Delta = 0$ is bounded by

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

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

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

Thus

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

and this is acceptable for (1.7.6) since $MN \gg \theta x$. It therefore suffices to show that

$$\mathcal{D}_3(M,N) \ll \theta^{18}(\log MN)^{O(1)}(MN)^{\gamma_0} N, \tag{1.7.8}$$

where $\mathcal{D}_3$ is $\mathcal{D}_2$ with the additional condition that $\Delta \neq 0$.

1.8 Congruence exercises

Our next major task, which requires much preparatory work, is to simplify $\mathcal{D}_3$ by removing the congruence condition entangling z_1, z_2, ℓ_1, and ℓ_2. 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 $\ell_2 \equiv b\ell_1 \pmod{|\Delta|}$.

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

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

To see this, note that we only require $(z_1, |\Delta|) = 1$. Now, if π is a Gaussian prime dividing z_1 and $|\Delta|$, then the congruence condition on b implies $\pi \mid z_2$, which contradicts the fact that z_1 and z_2 are coprime Gaussian integers.

One problem we face is that the congruence $\ell_2 \equiv b\ell_1 \pmod{|\Delta|}$ is not a congruence of rational 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

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

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 Δ 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

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

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 $|\Delta|$.

One 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.

We begin by removing those $|\Delta|$ that are unusually small (see [8, (17)] and the following discussion for a similar computation). Since $\Delta = \mathrm{Im}(\overline{z_1}z_2)$ 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 1.8.1. *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

$$D_3'(M,N) := \sum\sum_{\substack{N<|z_1|^2,|z_2|^2\leq(1+\theta)N \\ (z_1,z_2)=1 \\ 0<|\Delta|\leq\theta^{18}N}} \sum_{\substack{b(|\Delta|) \\ bz_1\equiv z_2(|\Delta|)}} \sum_{\ell_1\ll(MN)^{1/2}} \mathbf{1}_{\mathcal{A}}(\ell_1) \sum_{\substack{\ell_2\equiv\mathrm{Re}(b)\ell_1(|\Delta|) \\ (1.8.1)}} \mathbf{1}_{\mathcal{A}}(\ell_2),$$

where (1.8.1) denotes the condition

$$\left|\ell_2 - \ell_1\frac{z_2}{z_1}\right| \ll \theta^{18}(MN)^{1/2}. \tag{1.8.1}$$

Observe that (1.8.1) forces ℓ_2 to lie in an interval $I = I(\ell_1, z_2, z_2)$ of length $\leq c\theta^{18}(MN)^{1/2}$, for some positive constant c.

We use the "intervals of length a power of ten" technique we deployed in analyzing (1.4.17) (see (1.5.4)). 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 non-negative integer (observe that we require only $O(1)$ subintervals to cover I). Recalling from the proof of Lemma 1.5.4 that we argue slightly differently depending on whether a_0 is zero or not, we have

$$\sum_{\substack{\ell_2\equiv\mathrm{Re}(b)\ell_1(|\Delta|) \\ (1.8.1)}} \mathbf{1}_{\mathcal{A}}(\ell_2) \leq \sum_{n\in S(I)} \mathbf{1}_{\mathcal{A}}(n) \sum_{\substack{\delta(a_0)Y/10\leq t<Y \\ t+nY\equiv\mathrm{Re}(b)\ell_1(|\Delta|)}} \mathbf{1}_{\mathcal{A}}(t),$$

where $S(I)$ is some set of integers depending on 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

$$\sum_{\substack{\ell_2 \equiv \mathrm{Re}(b)\ell_1 (|\Delta|) \\ (1.8.1)}} \mathbf{1}_{\mathcal{A}}(\ell_2) \ll \frac{1}{|\Delta|}(\theta^{18}\sqrt{MN})^{\gamma_0} + \frac{1}{|\Delta|}(\theta^{18}\sqrt{MN})^{\gamma_0} \sum_{r=1}^{|\Delta|-1} F_X\left(\frac{r}{|\Delta|}\right),$$

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

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

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 $\mathrm{Im}(\overline{z_1}z_2) = rv - su$. 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 1.5.3,

$$\frac{1}{Q}\sum_{q}\sum_{(b,q)=1} F_X\left(\frac{b}{q}\right) \ll \frac{1}{Q^{23/77}} + \frac{Q}{X^{50/77}} \ll 1 + \frac{\theta^{18}N}{X^{50/77}}. \tag{1.8.2}$$

We wish for the quantity in (1.8.2) to be $\ll 1$, so it suffices to have $N \ll x^{25/77-\epsilon}$, and this in turn requires

$$U \geq x^{52/77+\epsilon}. \tag{1.8.3}$$

The constraint (1.8.3) replaces (1.7.3), and is the last lower bound condition we need to put on U. 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. $\square$

We make now a brief detour to discuss our restrictions on N, U, and V. With the upper bound on UV from (1.6.2), our lower bound for V (1.7.5), and our lower bound for U (1.8.3) in hand, there are no more conditions to put on U or V, and the range of N in (1.6.17) is now clear. For these constraints to be consistent with one another it suffices to have

$$\begin{aligned} &U = x^{\alpha}, \qquad V = x^{\beta}, \\ &\frac{52}{77} = 0.675\ldots < \alpha < \gamma_0 - \left(\frac{1}{2} - \frac{\gamma_0}{2}\right) = 0.931\ldots, \\ &\frac{1}{2} - \frac{\gamma_0}{2} = 0.0228\ldots < \beta < \gamma_0 - \alpha. \end{aligned} \tag{1.8.4}$$

Note that (1.8.4) is consistent with the specific choice we made in (1.6.3).

Let us return to estimations. With Lemma 1.8.1 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 0.1.1.

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

$$D_3''(M,N) := \sum_{p>P} \sum\sum_{\substack{N<|z_1|^2,|z_2|^2\leq(1+\theta)N \\ (z_1,z_2)=1 \\ p||\Delta|}} \sum_{bz_1\equiv z_2(|\Delta|)} \sum_{\substack{\ell_1<X \\ p|\ell_1}} \mathbf{1}_{\mathcal{A}}(\ell_1) \sum_{\substack{\ell_2<X \\ \ell_2\equiv \mathrm{Re}(b)\ell_1(|\Delta|)}} \mathbf{1}_{\mathcal{A}}(\ell_2),$$

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

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

By Lemma 1.5.2 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

$$D_3''(M,N) \ll (MN)^{\gamma_0} N \sum_{P<p\ll N} p^{-50/77} \sum_{\substack{d\ll N \\ p|d}} \frac{1}{d} \sum_{r=1}^{d} F_X\left(\frac{r}{d}\right).$$

We separate the contribution of the zero frequency $r = d$, and find that it contributes

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

For the nonzero frequencies we reduce to primitive fractions, obtaining

$$\begin{aligned}
\sum_{p>P} p^{-50/77} \sum_{\substack{d\ll N\\ p|d}} \frac{1}{d} \sum_{r=1}^{d} F_X\left(\frac{r}{d}\right) &\ll \sum_{p>P} p^{-50/77} \sum_{\substack{d\ll N\\ p|d}} \frac{1}{d} \sum_{\substack{q|d\\ q>1}} \sum_{(a,q)=1} F_X\left(\frac{a}{q}\right)\\
&\ll \sum_{q\ll N} \frac{1}{q} \sum_{(a,q)=1} F_X\left(\frac{a}{q}\right) \sum_{d'\ll N/q} \frac{1}{d'} \sum_{\substack{p>P\\ p|q}} p^{-50/77}\\
&+ \sum_{q\ll N} \frac{1}{q} \sum_{(a,q)=1} F_X\left(\frac{a}{q}\right) \sum_{d'\ll N/q} \frac{1}{d'} \sum_{\substack{p>P\\ p|d'}} p^{-50/77}.
\end{aligned}$$

Here we have written $d = qd'$ and used the fact that $p \mid qd'$ implies $p \mid q$ or $p \mid d'$. We change variables $d' = pd''$, say, and use the bound

$$\sum_{\substack{p>P\\ p|k}} \frac{1}{p^{50/77}} \ll (\log k) P^{-50/77},$$

to obtain

$$\begin{aligned}
D_3''(M,N) &\ll (\log N)^2 (MN)^{\gamma_0} N P^{-50/77} \sum_{q\ll N} \frac{1}{q} \sum_{(a,q)=1} F_X\left(\frac{a}{q}\right)\\
&\ll (\log N)^3 (MN)^{\gamma_0} N P^{-50/77}.
\end{aligned}$$

The second inequality follows by a dyadic decomposition and Lemma 1.5.3. Thus the contribution from those ℓ_1 not coprime to $|\Delta|$ is negligible.

Now that b is a rational residue class, it remains only 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

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

In order to prove our desired bound (1.7.8), it therefore suffices to study

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

and show

$$\mathcal{D}_4(M,N) \ll \theta^{18}(\log MN)^{O(1)}(MN)^{\gamma_0}N. \tag{1.8.5}$$

Before we proceed further, there is another technical issue to resolve. As mentioned above, the sequence $\mathcal{A}$ is nicely distributed in residue classes to moduli that are coprime to 10, but things become more complicated if the modulus is not coprime to 10. In an effort to isolate this poor behavior at the primes 2 and 5 we write

$$\Delta = \Delta_{10}|\Delta'|,$$

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 (1.7.1)). 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 to $\mathcal{D}_4(M,N)$ from large Δ_{10} to be negligible. We finish this section with the following result.

Lemma 1.8.2. *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

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

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

$$\sum_{\substack{\ell_2 < X \\ \ell_2 \equiv b\ell_1(\Delta_{10}) \\ \ell_2 \equiv b\ell_1(|\Delta'|)}} \mathbf{1}_{\mathcal{A}}(\ell_2) = \frac{1}{\Delta_{10}|\Delta'|} \sum_{s=1}^{\Delta_{10}} e\left(-\frac{sb\ell_1}{\Delta_{10}}\right) \sum_{k=1}^{|\Delta'|} e\left(-\frac{kb\ell_1}{|\Delta'|}\right)$$
$$\times \sum_{\ell_2 < X} \mathbf{1}_{\mathcal{A}}(\ell_2) e\left(\frac{k\ell_2}{|\Delta'|} + \frac{s\ell_2}{\Delta_{10}}\right).$$

We first consider the contribution from $k = |\Delta'|$. By the triangle inequality, the contribution from $k = |\Delta'|$ to $\mathcal{D}_4'$ is

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

the last inequality following by Lemma 1.5.2. 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 1.5.5. Arguing as in the case $k = |\Delta'|$ and changing variables, this contribution is bounded by

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

Reducing from fractions with denominator e to primitive fractions gives that this last quantity 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 $q \le Q$ and $q > Q$, where $Q = \exp(\varepsilon\sqrt{\log N})$ with $\varepsilon > 0$ sufficiently small. 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 1.5.3, which gives that the contribution from $q > Q$ is

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

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

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

the last inequality following from Lemma 1.5.2. By Rankin's trick, we obtain the bound

$$Q \sum_{\substack{t>T \\ t|10^\infty}} \frac{1}{t^{50/77}} \ll \frac{Q}{T^{1/2}} = \frac{1}{Q}.$$

It therefore suffices to bound

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

At this point we avail ourselves of the product formula (1.5.2) for F. We take U to be a power of 10 such that t divides U for every $t \mid 10^\infty$ with $t \le Q^4$, and set $V = X/U$. Any such $t \le Q^4$ may be written as $t = 2^a 5^b$ with $2^a 5^b \le Q^4$. Clearly $a, b \le \frac{4}{\log 2} \log Q$. We take $U = 10^c$ with $c = \frac{4}{\log 2} \log Q + O(1)$, so that

$$U \asymp Q^{\frac{4}{\log 2} \log 10} \ll Q^{14},$$

say. Since F is 1-periodic we obtain

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

Observe that V and X are asymptotically equal in the logarithmic scale since $Q = N^{o(1)}$. We then apply Lemma 1.5.1 to bound each F_V individually, and find

$$\mathcal{D}_4' \ll (\log N)^4 (MN)^{\gamma_0} N \left(\theta^{18} + \exp(-c'\sqrt{\log MN})\right). \tag{1.8.6}$$

□

In light of Lemma 1.8.2 it suffices to show that

$$\mathcal{D}_5(M,N) \ll \theta^{18} (\log MN)^{O(1)} (MN)^{\gamma_0} N, \tag{1.8.7}$$

where $\mathcal{D}_5(M,N)$ is the same as $\mathcal{D}_4(M,N)$, but with the additional condition that $\Delta_{10} \le \theta^{-28}$.

1.9 Polar boxes and the fundamental lemma

In this section we remove the congruence condition modulo $|\Delta'|$, which will simplify the situation considerably. Not surprisingly, there are several 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. 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 task is then to get cancellation from the Möbius function in the main term, which we do in Section 1.10.

We begin with a preparatory lemma.

Lemma 1.9.1. *Let* $\exp(-(\log MN)^{1/3}) < \delta < \frac{1}{2}$*, and let* $C' > 0$ *be an absolute constant. Then*

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

Proof. We begin by handling the congruence condition as we did in imposing the condition $\Delta_{10} \leq \theta^{-28}$ in Lemma 1.8.2. We detect the congruences modulo Δ_{10} and $|\Delta'|$ with additive characters. The nonzero frequencies modulo $|\Delta'|$ contribute an acceptably small error term, by the argument that led to (1.8.6). For the zero frequency modulo $|\Delta'|$ we apply orthogonality of additive characters to reintroduce the congruence modulo Δ_{10}. We find

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

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

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

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

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

Summing over r and s then completes the proof. □

We now introduce a parameter $\lambda = k^{-1}$, for some $k \in \mathbb{N}$ to be chosen. We break the sums over z_1, z_2 into polar boxes, so that

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

Note that $N < R_i \le (1+\theta)N$ and $\delta_j = 2\pi j\lambda$ for $0 \le j \le \lambda^{-1}-1$ an integer. For such a polar box, let $z(\mathfrak{B}) := R_i e^{i\delta_j}$. The number of polar boxes for z_1, z_2 is $O(\lambda^{-4})$, and we have the trivial bound

$$\sum_{z\in\mathfrak{B}} 1 \ll \lambda^2 N.$$

Set $\Delta(\mathfrak{B}_1,\mathfrak{B}_2) := \Delta(z(\mathfrak{B}_1), z(\mathfrak{B}_2))$. 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 $\theta_i = \delta_i + O(\lambda)$. If we assume that $\lambda \leq \varepsilon\theta^{18}$ for some sufficiently small $\varepsilon > 0$, then 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

$$\Delta(z_1,z_2) = (1+O(\lambda'))\Delta(\mathfrak{B}_1,\mathfrak{B}_2), \tag{1.9.1}$$

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 sum over polar boxes not satisfying this condition is bounded by

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

for some absolute constant $C' > 0$, and by Lemma 1.9.1 this quantity is $\ll \theta^{-28}\lambda'(MN)^{\gamma_0}N$. This bound is acceptable for (1.8.7) provided

$$\lambda \leq \theta^{64},$$

which we now assume.

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 1.9.1, we find the error made by this approximation is

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

and this error is acceptable for (1.8.7) since we have already imposed the condition $\lambda \leq \theta^{64}$. We therefore have

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

where

$$\mathcal{D}_1(\mathfrak{B}_1,\mathfrak{B}_2) := \sum_{\substack{z_1\in\mathfrak{B}_1, z_2\in\mathfrak{B}_2 \\ (z_1\overline{z_1}z_2\overline{z_2},\Pi)=1 \\ (z_1|\Delta|,z_2)=1 \\ \Delta_{10}\leq\theta^{-28}}}\sum \mu(|z_1|^2)\mu(|z_2|^2) \sum_{bz_1\equiv z_2(|\Delta|)} \times \sum_{\substack{\ell_1\leq\sqrt{2(1+\theta)}(MN)^{1/2} \\ (\ell_1,\Pi|\Delta|)=1}} \mathbf{1}_{\mathcal{A}}(\ell_1) \sum_{\substack{\ell_2\leq\sqrt{2(1+\theta)}(MN)^{1/2} \\ \ell_2\equiv b\ell_1(|\Delta|) \\ |\ell_1z_2-\ell_2z_1|^2\leq 2\Delta^2M \\ (\ell_2,\Pi)=1}} \mathbf{1}_{\mathcal{A}}(\ell_2).$$

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

$$\mathcal{D}_1(\mathfrak{B}_1,\mathfrak{B}_2) \ll \theta^{18}(\log MN)^{O(1)}\lambda^4(MN)^{\gamma_0}N \tag{1.9.2}$$

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

We now work to make the condition $|\ell_1z_2-\ell_2z_1|^2 \leq 2\Delta^2M$ less dependent on z_1 and z_2. We can rearrange to get the condition

$$\left|\ell_2 - \frac{\ell_1z_2}{z_1}\right| \leq \sqrt{2}\frac{|\Delta(z_1,z_2)|M^{1/2}}{|z_1|}. \tag{1.9.3}$$

We wish to replace (1.9.3) by

$$\left|\ell_2 - \frac{\ell_1z_2}{z_1}\right| \leq \sqrt{2}\frac{|\Delta(\mathfrak{B}_1,\mathfrak{B}_2)|M^{1/2}}{|z(\mathfrak{B}_1)|}. \tag{1.9.4}$$

Since

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

we see it suffices to bound the contribution from those ℓ_2 that satisfy

$$(1-C\lambda')K \leq \left|\ell_2 - \frac{\ell_1 z_2}{z_1}\right| \leq (1+C\lambda')K, \tag{1.9.5}$$

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 (1.9.5) places ℓ_2 in a bounded number of intervals (depending on ℓ_1, z_1, z_2) of length $\ll (\lambda')^{1/2}J$. For notational simplicity, write $A = (1-C\lambda')K$ and $B = (1+C\lambda')K$. Then (1.9.5) 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

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

There are two cases now to consider: $v \geq A$ and $v < A$. If $v \geq A$ then $A^2 - v^2 \leq 0$, and the lower bound is therefore 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 thus 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 (1.9.5). At this point we should have enough experience to see how we should proceed. We let Y be the largest power of 10 satisfying $Y \leq (\lambda')^{1/2}K$, and cover the intervals (1.9.5) with subintervals of the form $[nY, nY + Y)$, $n \geq 0$ an integer. The number of subintervals is $O(1)$. We can reduce to summing the indicator function $\mathbf{1}_{\mathcal{A}}(t)$ over $0 \leq t < Y$, and then deal with the congruence modulo $|\Delta|$ by considering it as a congruence modulo $|\Delta'|$ and Δ_{10}. We obtain a bound of

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

and this is acceptable for (1.9.2) provided $\lambda \ll \theta^{153}$. We set

$$\lambda \asymp \theta^{153}.$$

We now have the conditions

$$\begin{aligned} \ell_2 &\leq \sqrt{2(1+\theta)}(MN)^{1/2}, \\ \left|\ell_2 - \frac{\ell_1 z_2}{z_1}\right| &\leq \sqrt{2}\frac{|\Delta(\mathfrak{B}_1, \mathfrak{B}_2)|M^{1/2}}{|z(\mathfrak{B}_1)|}, \\ \ell_2 &\equiv b\ell_1(|\Delta|) \end{aligned} \tag{1.9.6}$$

on ℓ_2. Recall that the congruence 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 (1.9.6) we obtain acceptable contributions. 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 $\mathcal{D}(\mathfrak{B}_1, \mathfrak{B}_2)$ with the condition (1.9.3) replaced by (1.9.4)). We partition Σ as

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

where in Σ_+ we sum over those z_1, z_2 such that $\mu(|z_1|^2)\mu(|z_2|^2) > 0$, and in Σ_- we sum over those z_1, z_2 such that $\mu(|z_1|^2)\mu(|z_2|^2) < 0$. We get an upper bound on Σ_+ using an upper-bound linear sieve of level D

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

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

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

where D is chosen shortly (see (1.9.8)). This yields an upper bound on Σ. Reversing λ^+ and λ^- we get a lower bound on Σ, and we show that these bounds are the same asymptotically. Thus, for some sign ϵ, it suffices to study

$$\Sigma_\epsilon^\pm := \sum_{\substack{z_i \in \mathfrak{B}_i \\ \mu(|z_1|^2)\mu(|z_2|^2) = \epsilon}} \sum_{\substack{d \leq D \\ d \mid \Pi \\ (d, 10|\Delta|)=1}} \lambda_d^\pm \sum_{b(|\Delta|)} \sum_{\ell_1} \sum_{n} \sum_{\substack{nY < \ell_2 \leq nY+Y \\ \ell_2 \equiv b\ell_1 (|\Delta|) \\ \ell_2 \equiv 0 (d) \\ (\ell_2, 10)=1}} \mathbf{1}_{\mathcal{A}}(\ell_2). \tag{1.9.7}$$

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 (1.9.7) becomes a linear combination of a bounded number of quantities of the form

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

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 $\Sigma^{\pm}_{\epsilon,E}$, which comes from $1 \leq f \leq d|\Delta'| - 1$. The argument is similar to that which gave (1.8.6) in Lemma 1.8.2.

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

$$\begin{aligned}\Sigma^{\pm}_{\epsilon,E} &\ll \lambda^{-1}(MN)^{\gamma_0} \sum_{\substack{d\leq D \\ (d,10)=1}} \sum_{\substack{|\Delta'|\ll N \\ (|\Delta'|,10)=1 \\ d|\Delta'|>1}} \frac{1}{d|\Delta'|} \sum_{f=1}^{d|\Delta'|-1} \sum_{\substack{\Delta_{10}\leq\theta^{-22} \\ \Delta_{10}|10^\infty}} \frac{1}{\Delta_{10}} \sum_{g=1}^{\Delta_{10}} \sum_{h|10} \sum_{k=1}^{h} \\ &\quad \times F_{Y'}\left(\frac{f}{d|\Delta'|} + \frac{g}{\Delta_{10}} + \frac{k}{h}\right) \sum_{z_1\in\mathfrak{B}_1} \sum_{\substack{z_2\in\mathfrak{B}_2 \\ \mathrm{Im}(\overline{z_1}z_2)=\Delta_{10}\Delta'}} 1 \\ &\ll \lambda(MN)^{\gamma_0} N \sum_{\substack{1<d\ll DN \\ (d,10)=1}} \frac{\tau(d)}{d} \sum_{f=1}^{d-1} \sum_{\substack{\Delta_{10}\leq\theta^{-22} \\ \Delta_{10}|10^\infty}} \frac{1}{\Delta_{10}} \sum_{g=1}^{\Delta_{10}} \sum_{h|10} \sum_{k=1}^{h} F_{Y'}\left(\frac{f}{d} + \frac{g}{\Delta_{10}} + \frac{k}{h}\right).\end{aligned}$$

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

$$\begin{aligned}\Sigma^{\pm}_{\epsilon,E} \ll (\log N)^2 \lambda(MN)^{\gamma_0} N &\sum_{\substack{1<q\ll DN \\ (q,10)=1}} \frac{\tau(q)}{q} \sum_{\substack{r=1 \\ (r,q)=1}}^{q} \sum_{\substack{\Delta_{10}\leq\theta^{-22} \\ \Delta_{10}|10^\infty}} \\ &\times \frac{1}{\Delta_{10}} \sum_{g=1}^{\Delta_{10}} \sum_{h|10} \sum_{k=1}^{h} F_{Y'}\left(\frac{r}{q} + \frac{g}{\Delta_{10}} + \frac{k}{h}\right).\end{aligned}$$

We choose

$$D := x^{1/\log\log x}, \tag{1.9.8}$$

so that $DN \ll x^{25/77-\epsilon}$. We estimate the contribution from $q > Q = \exp((\log MN)^{1/3})$ using the divisor bound, dyadic decomposition, and Lemma 1.5.3. For $q \leq Q$ we first use the product formula for $F_{Y'}$ to eliminate $\frac{g}{\Delta_{10}} + \frac{k}{h}$, and then use Lemma 1.5.1.

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 acceptable error terms, the main term is then given by

$$\Sigma_{\epsilon,0}^{\pm} := \sum_{\substack{z_i \in \mathfrak{B}_i \\ \mu(|z_1|^2)\mu(|z_2|^2)=\epsilon}}\sum \frac{1}{|\Delta'|} \sum_{\substack{d \leq D \\ d|\Pi \\ (d,10|\Delta|)=1}} \frac{\lambda_d^{\pm}}{d} \sum_{\substack{\ell_1 \leq \sqrt{2(1+\theta)}(MN)^{1/2} \\ (\ell,\Pi|\Delta|)=1}} \mathbf{1}_{\mathcal{A}}(\ell_1) \sum_{\substack{\ell_2 \leq \sqrt{2(1+\theta)}(MN)^{1/2} \\ \ell_2 z_1 \equiv \ell_1 z_2 (\Delta_{10}) \\ (1.9.4) \\ (\ell_2,10)=1}} \mathbf{1}_{\mathcal{A}}(\ell_2).$$

From the fundamental lemma of sieve theory (see (1.5.7), for example) we have

$$\sum_{\substack{d \leq D \\ d|\Pi \\ (d,10|\Delta|)=1}} \frac{\lambda_d^{\pm}}{d} = \left(1 + \exp\left(-\frac{1}{2}s\log s\right)\right) \prod_{\substack{p \leq P \\ p \nmid 10|\Delta|}} \left(1 - \frac{1}{p}\right), \tag{1.9.9}$$

where

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

The error term of (1.9.9) is therefore acceptably small for (1.9.2). We write

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

and observe that the error term is again acceptable by our lower bound for P. Thus $\Sigma^{+}_{\epsilon,0}$ and $\Sigma^{-}_{\epsilon,0}$ are asymptotically equal, and up to acceptable error terms we have

$$\begin{aligned}\Sigma = &\sum\sum_{\substack{z_1\in\mathfrak{B}_1, z_2\in\mathfrak{B}_2\\ (z_1\overline{z_1}z_2\overline{z_2},\Pi)=1\\ (z_1|\Delta|,z_2)=1\\ \Delta_{10}\le\theta^{-28}}}\mu(|z_1|^2)\mu(|z_2|^2)\frac{1}{|\Delta'|}\frac{10|\Delta|}{\varphi(10|\Delta|)}\\ &\times\sum_{\substack{\ell_1\le\sqrt{2(1+\theta)}(MN)^{1/2}\\ (\ell,\Pi|\Delta|)=1}}\mathbf{1}_{\mathcal{A}}(\ell_1)\sum_{\substack{\ell_2\le\sqrt{2(1+\theta)}(MN)^{1/2}\\ \ell_2 z_1\equiv\ell_1 z_2(\Delta_{10})\\ (1.9.4)\\ (\ell_2,10)=1}}\mathbf{1}_{\mathcal{A}}(\ell_2).\end{aligned}$$

We may use trivial estimations to replace condition (1.9.4) by

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

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 acceptable error. 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

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

where

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

Recall the lower bound $|\Delta(\mathfrak{B}_1,\mathfrak{B}_2)| \gg \theta^{18}N$. In order to prove (1.9.2) it therefore suffices to show that

$$\mathcal{D}_2(\mathfrak{B}_1,\mathfrak{B}_2) \ll \theta^{36}\lambda^4(\log MN)^{O(1)}(MN)^{\gamma_0}N^2. \tag{1.9.10}$$

1.10 Simplifications and endgame

We have removed the congruence condition to modulus $|\Delta'|$. From this point onwards our estimates are more straightforward, since we do not have to work with congruence conditions on elements of $\mathcal{A}$ to large moduli.

Recall that our goal is to use the cancellation induced by the Möbius function to show that $\mathcal{D}_2$ is small. 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. 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.

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

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

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 show (1.9.10) it therefore suffices to prove

$$\mathcal{C} := \sum_{\substack{z_i \in \mathfrak{B}_i \\ (|z_i|^2,\Pi)=1 \\ \Delta_{10}=f \\ \ell_1 z_2 \equiv \ell_2 z_2 (f) \\ |\ell_1 z_2 - \ell_2 z_1|^2 \leq 2\Delta(\mathfrak{B}_1,\mathfrak{B}_2)^2 M}} \sum \mu(|z_1|^2)\mu(|z_2|^2) \frac{|\Delta|}{\varphi(|\Delta|)} \ll \theta^{92} (\log MN)^{O(1)} \lambda^4 N^2 \quad (1.10.1)$$

uniformly in $f \leq \theta^{-28}$ with $f \mid 10^\infty$, 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.

If n is a positive integer, then

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

and therefore

$$\mathcal{C} = \sum_{d \ll N} \frac{\mu^2(d)}{\varphi(d)} \sum_{\substack{z_i \in \mathfrak{B}_i \\ (|z_i|^2,\Pi)=1 \\ \Delta_{10}=f \\ \ell_1 z_2 \equiv \ell_2 z_2 (f) \\ \Delta \equiv 0 (d) \\ |\ell_1 z_2 - \ell_2 z_1|^2 \leq 2\Delta(\mathfrak{B}_1,\mathfrak{B}_2)^2 M}} \sum \mu(|z_1|^2)\mu(|z_2|^2). \quad (1.10.2)$$

We introduce a parameter W, and estimate trivially the contribution from $d > W$ in (1.10.2). Writing z_1 and z_2 in rectangular coordinates, we see the contribution from $d > W$ is bounded by

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

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

$$W = \theta^{-92}\lambda^{-4} \asymp \theta^{-704}$$

then gives an acceptable contribution for (1.10.1).

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 $f = 2^a5^b$. Then $\Delta_{10} = f$ if and only if

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

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

$$\sum_{\substack{z_1,z_2 \\ \Delta_{10}=f}}\!\!\sum = \sum_{\substack{\mathfrak{m},\mathfrak{n}(10f) \\ \mathrm{Im}(\overline{\mathfrak{m}}\mathfrak{n})\equiv\nu_i(10f)}}\!\!\sum \;\sum_{z_1\equiv\mathfrak{m}(10f)} \;\sum_{z_2\equiv\mathfrak{n}(10f)}.$$

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

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

so to prove (1.10.1) it suffices to show that

$$\sum_{\substack{z_i \in \mathfrak{B}_i \\ (|z_i|^2, \Pi)=1 \\ |\ell_1 z_2 - \ell_2 z_1|^2 \leq 2\Delta(\mathfrak{B}_1, \mathfrak{B}_2)^2 M}}\sum \mu(|z_1|^2)\psi'(z_1)\mu(|z_2|^2)\psi(z_2) \ll \theta^{204}\lambda^4(\log MN)^{O(1)}N^2 \quad (1.10.3)$$

uniformly in characters ψ' and ψ. Here

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

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 moduli at most $O(d^2 f^2) = O(\theta^{-1464})$. Taking the supremum over z_1, it suffices to show that

$$\mathcal{S} := \sum_{\substack{z_2 \in \mathfrak{B}_2 \\ (z_2\overline{z_2}, \Pi)=1 \\ (1.10.5)}} \mu(|z_2|^2)\psi(z_2) \ll \theta^{204}\lambda^2(\log MN)^{O(1)}N, \quad (1.10.4)$$

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

$$\left| z_2 - z_1 \frac{\ell_2}{\ell_1} \right| \leq \sqrt{2}\frac{|\Delta(\mathfrak{B}_1, \mathfrak{B}_2)| M^{1/2}}{\ell_1}. \quad (1.10.5)$$

We see that (1.10.5) 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 introduce a parameter ϖ. We cover $\mathfrak{B}_2$ in polar boxes, which we call ϖ-polar boxes, of the form

$$R \leq |z_2|^2 \leq (1+\varpi)R,$$
$$\vartheta \leq \arg(z_2) \leq \vartheta + 2\pi\varpi.$$

For technical convenience we use smooth partitions of unity to accomplish this. 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

$$R < n \leq (1 + O(\varpi))R, \qquad R \asymp N, \tag{1.10.6}$$

and which satisfies

$$g^{(j)}(n) \ll_j (\varpi N)^{-j}, \qquad j \geq 0. \tag{1.10.7}$$

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

$$q^{(j)}(\alpha) \ll_j \varpi^{-j}, \qquad j \geq 0. \tag{1.10.8}$$

We observe that the boundary of the intersection between $\mathfrak{B}_2$ and the disc (1.10.5) 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

$$\varpi = \theta^{204}\lambda \asymp \theta^{357}.$$

It follows that

$$\mathcal{S} = O(\theta^{204}\lambda^2 N) + \sum\sum_{(g,q)\in S(z_1,\ell_1,\ell_2)} \sum_{(z_2\overline{z_2},\Pi)=1} \mu(|z_2|^2)\psi(z_2)g(|z_2|^2)q(\arg(z_2)).$$

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 (1.10.4) it suffices to show that

$$\mathcal{S}_{g,q} := \sum_{(z_2\overline{z_2},\Pi)=1} \mu(|z_2|^2)\psi(z_2)g(|z_2|^2)q(\arg(z_2)) \ll \theta^{204}\varpi^2(\log N)^{O(1)}N \tag{1.10.9}$$

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 [4, (16.14)]). Our treatment of $\mathcal{S}_{g,q}$ follows their treatment of $S_\chi^k(\beta)$ quite closely, and we quote the relevant statements and results of [4, 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 (1.10.8) we see the Fourier coefficients satisfy

$$\widehat{q}(h) \ll \frac{\varpi}{1+\varpi^2 h^2}. \tag{1.10.10}$$

By means of (1.10.10) we obtain the truncated Fourier series

$$q(\alpha) = \sum_{|h|\leq H} \widehat{q}(h) e^{ih\alpha} + O(\varpi^{-1}H^{-1}). \tag{1.10.11}$$

The contribution of the error term in (1.10.11) to $\mathcal{S}_{g,q}$ is $O(NH^{-1})$.

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

$$g(n) = \frac{1}{2\pi i}\int_{(\sigma)} \widehat{g}(s) n^{-s} ds, \qquad s = \sigma + it. \tag{1.10.12}$$

As g is supported in the interval (1.10.6) and satisfies (1.10.7), we find that the Mellin transform $\widehat{g}(s)$ is entire and satisfies

$$\widehat{g}(s) \ll \frac{\varpi N^\sigma}{1+\varpi^2 t^2}. \tag{1.10.13}$$

Applying (1.10.11) and (1.10.12) we obtain

$$\mathcal{S}_{g,q} = \sum_{|h|\leq H} \widehat{q}(h) \frac{1}{2\pi i}\int_{(\sigma)} \widehat{g}(s) Z_\psi^h(s) ds + O(NH^{-1}), \tag{1.10.14}$$

where

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

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. Since z is odd and primary, there is a one-to-one correspondence between elements z and odd ideals $\mathfrak{a}$, given by $\mathfrak{a} = (z)$. Omitting subscripts and superscripts for simplicity, we then have

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

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

$$P(s) := \prod_{\substack{p\leq P\\ p\equiv 1(4)}} \left(1-\frac{\xi(\mathfrak{p})}{p^s}\right)^{-1}\left(1-\frac{\xi(\overline{\mathfrak{p}})}{p^s}\right)^{-1}$$

where $\mathfrak{p}\overline{\mathfrak{p}} = (p)$, and $G(s)$ is given by an Euler product that converges absolutely and uniformly in $\sigma \geq \frac{1}{2}+\epsilon$. In the region $\sigma \geq 1-\frac{1}{\log P}$ the inequality $p^{-\sigma} < 3p^{-1}$ holds, and this gives the bound

$$P(s) \ll (\log P)^3, \qquad \sigma \geq 1-\frac{1}{\log P}. \tag{1.10.15}$$

Let k be the modulus of ξ (recall that $k \ll \theta^{-1464}$). Then $L(s,\xi)$ is nonzero (see [4, (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 ([4, Lemma 16.1]) one may show that when ξ is real, $L(s,\xi)$ has no zeros in the region

$$\sigma \geq 1 - \frac{c(\varepsilon)}{k^{\varepsilon}}, \qquad 0 < \varepsilon \leq \frac{1}{4}. \tag{1.10.16}$$

The constant $c(\varepsilon)$ is ineffective, and for this reason the implied constants in Proposition 1.6.3 and Theorem 0.1.1 are ineffective.

The inequality (1.10.16) allows one to establish ([4, (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

$$Z(s) \ll k(\log(|h| + |t| + N))^5. \tag{1.10.17}$$

We now estimate the integral

$$I := \frac{1}{2\pi i} \int_{(\sigma)} \widehat{g}(s) Z(s) ds. \tag{1.10.18}$$

We move the contour of integration to

$$\begin{aligned} s &= 1 + it, \qquad |t| \geq T, \\ s &= 1 - \beta + it, \qquad |t| \leq T, \end{aligned}$$

and add in horizontal connecting segments

$$s = \sigma + \pm iT, \qquad 1 - \beta \leq \sigma \leq 1.$$

Estimating trivially we find by (1.10.13) and (1.10.17) that

$$I \ll \varpi^{-1}\theta^{-1464}\left(T^{-1}+N^{-\beta}\right)N(\log(N+T))^5.$$

We set $T := 3\exp(\sqrt{\log N})$. Recalling that $\log P \leq \frac{\sqrt{\log x}}{\log\log x} \ll \frac{\sqrt{\log N}}{\log\log N}$, we see that

$$I \ll (\log N)^5\varpi^{-1}\theta^{-1464}N\left(\exp\left(-\frac{c(\varepsilon)\sqrt{\log N}}{k^{\varepsilon}}\right)+\exp\left(-c\sqrt{\log N}\right)\right) \tag{1.10.19}$$

uniformly in $|h| \leq 2\exp(\sqrt{\log N})$. We choose $H := \exp(\sqrt{\log N})$, then take (1.10.14) together with (1.10.19) and sum over $|h| \leq H$ by means of (1.10.10). Provided $\varepsilon > 0$ is sufficiently small in terms of θ (take $\varepsilon = \varepsilon(L)$, compare (1.6.7)), we obtain the bound

$$\mathcal{S}_{g,q} \ll_{\epsilon} N\exp\left(-c(\epsilon)(\log N)^{\frac{1}{2}-\epsilon}\right). \tag{1.10.20}$$

The bound (1.10.20) implies (1.10.9), and this completes the proof of Proposition 1.6.3.

1.11 Modifications for Theorem 1.1.1

The proof of Theorem 1.1.1 follows the same lines as the proof of Theorem 0.1.1. 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 $\{a_1,\ldots,a_d\} \subset \{0,1,2,\ldots,9\}$ be a fixed set. Denote by $\mathcal{A}_d$ the set of nonnegative integers missing the digits $a_1,\ldots,a_d$ in their decimal expansions. Let $\gamma_d := \frac{\log(10-d)}{\log 10}$. For Y a power of 10 we define

$$F_{Y,d}(\theta) := Y^{-\gamma_d}\left|\sum_{n<Y}\mathbf{1}_{\mathcal{B}_d}(n)e(n\theta)\right|.$$

We note that if $Y = 10^k$ then

$$\begin{aligned} F_{Y,d}(\theta) &= \prod_{i=0}^{k-1}\frac{1}{10-d}\left|\sum_{n_i<10}\mathbf{1}_{\mathcal{B}_d}(n_i)e(n_i10^i\theta)\right| \\ &= \prod_{i=1}^{k}\frac{1}{10-d}\left|\frac{e(10^i\theta-1)}{e(10^{i-1}\theta)-1}-\sum_{r=1}^{d}e(a_r10^{i-1}\theta)\right|. \end{aligned}$$

We therefore have the product formula

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

The most important task is to obtain analogues of Lemmas 1.5.1, 1.5.2, and 1.5.3 for the functions $F_{Y,d}$. By arguing as in the proof of [9, Lemma 10.1] it is not difficult to prove the analogue of Lemma 1.5.1.

Lemma 1.11.1. *Let $q < Y^{\frac{1}{3}}$ be of the form $q = q_1q_2$ 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 1.5.2 and 1.5.3. 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 1.11.2. *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

$$G_d(t_0, \dots, t_J) := \sup_{|\gamma|\le 10^{-J-1}} \frac{1}{10-d}\left|\frac{e\left(\sum_{j=0}^J t_j 10^{-j} + 10\gamma\right) - 1}{e\left(\sum_{j=0}^J t_j 10^{-j-1} + \gamma\right) - 1}\right.$$
$$\left. - \sum_{r=1}^{d} e\left(\sum_{j=0}^{J} a_r t_j 10^{-j-1} + a_r\gamma\right)\right|.$$

Then

$$\sum_{0\le a<10^k} F_{10^k,d}\left(\frac{a}{10^k}\right)^t \ll_{t,J,d} \lambda_{t,J,d}^k.$$

Proof. Following the proof of [9, Lemma 10.2], we find that

$$F_{Y,d}\left(\sum_{i=1}^k \frac{t_i}{10^i}\right) \le \prod_{i=1}^k G_d(t_i,\ldots,t_{i+J}),$$

where $t_j = 0$ for $j > k$. Maynard proceeds at this point using a Markov chain argument, but we give here a different argument due to Kevin Ford (private communication).

Write $M_{t,d} = (m_{ij})_{i,j}$ (we suppress the dependence on d for notational convenience), where m_{ij} is zero unless

$$\begin{aligned} i-1 &= a_2 + a_3 10 + \cdots + 10^{J-1}a_{J+1},\\ j-1 &= a_1 + 10a_2 + \cdots + 10^{J-1}a_J \end{aligned}$$

for some digits $a_1,\ldots,a_J$. Thus

$$(M_{t,d}^k)_{i,j} = \sum_{i_1,\ldots,i_{k-1}} m_{i,i_1} m_{i_1,i_2}\cdots m_{i_{k-1},j},$$

where the product is nonzero only if

$$\begin{aligned} j-1 &= a_1 + 10a_2 + \cdots + 10^{J-1}a_J,\\ i_{k-1}-1 &= a_2 + 10a_3 + \cdots + 10^{J-1}a_{J+1},\\ &\vdots\\ i_1 - 1 &= a_k + 10a_{k+1} + \cdots + 10^{J-1}a_{k+J-1},\\ i-1 &= a_{k+1} + 10a_{k+2} + \cdots + 10^{J-1}a_{k+J}. \end{aligned}$$

Fixing $i=1$, so that $a_{k+1} = \cdots = a_{J+k} = 0$, and summing over j we obtain

$$\sum_j (M_{t,d}^k)_{1,j} = \sum_{a_1,\ldots,a_k} G_d(a_1,\ldots,a_{J+1})^t \cdots G(a_k,\ldots,a_{k+J})^t \ge \sum_{0\le a<10^k} F_{10^k,d}\left(\frac{a}{10^k}\right)^t.$$

One may then use the Perron-Frobenius theorem to obtain the conclusion of the lemma.

□

The following is a consequence of Lemma 1.11.2 and some numerical computation.

Lemma 1.11.3. *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

$$\alpha_2 = \frac{54}{125}, \qquad \alpha_3 = \frac{99}{200}.$$

Proof. We use bounds on $\lambda_{1,2,2}$ and $\lambda_{1,2,3}$. By numerical calculation[2] we find

$$\lambda_{1,2,2} < 10^{\frac{54}{125}}$$

for all choices of $\{a_1, a_2\} \subset \{0, \dots, 9\}$, and

$$\lambda_{1,2,3} < 10^{\frac{99}{200}}$$

for all choices of $\{a_1, a_2, a_3\} \subset \{0, \dots, 9\}$. By the argument of [9, 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

$$\begin{aligned}\int_0^1 F_{Y,d}(t)dt &= \sum_{0\le a<Y} \int_{\frac{a}{Y}}^{\frac{a}{Y}+\frac{1}{Y}} F_{Y,d}(t)dt = \int_0^{\frac{1}{Y}} \sum_{0\le a<Y} F_{Y,d}\left(\frac{a}{Y} + t\right) dt \\ &\le Y^{-1} \sup_{\beta\in\mathbb{R}} \sum_{a<Y} F_{Y,d}\left(\frac{a}{Y} + \beta\right) \ll Y^{-1+\alpha_d}.\end{aligned}$$

□

[2]Mathematica® files with these computations can be found at https://arxiv.org/abs/1806.02699.

We note it is crucial for the proof of Theorem 1.1.1 that $\alpha_d < \frac{1}{2}$. For $d \geq 4$ there exist choices of excluded digits which force $\alpha_d > \frac{1}{2}$.

Lemma 1.11.4. *We have*

$$\sup_{\beta\in\mathbb{R}} \sum_{a\leq q} \left| F_{Y,d}\left(\frac{a}{q}+\beta\right)\right| \ll q^{\alpha_d} + \frac{q}{Y^{1-\alpha_d}},$$
$$\sup_{\beta\in\mathbb{R}} \sum_{q\leq Q} \sum_{\substack{1\leq a\leq q\\(a,q)=1}} \left| F_{Y,d}\left(\frac{a}{q}+\beta\right)\right| \ll Q^{2\alpha_d} + \frac{Q^2}{Y^{1-\alpha_d}}$$

Proof. We use the large sieve argument of [9, Lemma 10.5] with Lemma 1.11.3. □

Let us now give a broad sketch of the proof of Theorem 1.1.1. We proceed as in the proof of Theorem 0.1.1, only we use Lemma 1.11.4 instead of Lemma 1.5.2 or Lemma 1.5.3.

Our sequence

$$\sum_{\substack{m^2+\ell^2=n\\(\ell,\Pi)=1}}\sum \mathbf{1}_{\mathcal{A}_d}(\ell)$$

has level of distribution

$$D \leq x^{\gamma_d-\epsilon},$$

and we have an acceptable Type II bound provided

$$x^{\frac{1}{2}-\frac{\gamma_d}{2}+\epsilon} \ll N \ll x^{\frac{1}{2}(1-\alpha_d)-\epsilon}. \tag{1.11.1}$$

(Compare (1.11.1) with (1.6.11).) 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 (1.2.1) (compare with (1.8.4)).

At various points in the proof of Theorem 0.1.1 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 1.5 leading up to Lemmas 1.5.4 and 1.5.5). The short interval decomposition depended on whether or not the missing digit was the zero digit. In the case of Theorem 1.1.1 one argues similarly, and finds that the short interval decomposition depends only on whether $0 \in \mathcal{A}_d$.

CHAPTER 2

AVERAGE NONVANISHING OF DIRICHLET L-FUNCTIONS AT THE CENTRAL POINT

2.1 Mollification, and a sketch for Theorem 0.2.1

In this chapter[1] we prove Theorem 0.2.1. We give an overview of the proof technique before getting into the actual details.

The proof of Theorem 0.2.1 relies on the powerful technique of mollification. For each character χ we associate a function $\psi(\chi)$, called a mollifier, that serves to dampen the large values of $L(\frac{1}{2},\chi)$. By the Cauchy-Schwarz inequality we have

$$\frac{\left|\sum_{q\asymp Q}\sum^{*}_{\chi(q)} L\left(\frac{1}{2},\chi\right)\psi(\chi)\right|^2}{\sum_{q\asymp Q}\sum^{*}_{\chi(q)} \left|L\left(\frac{1}{2},\chi\right)\psi(\chi)\right|^2} \leq \sum_{q\asymp Q}\sideset{}{^*}\sum_{\substack{\chi(q)\\ L\left(\frac{1}{2},\chi\right)\neq 0}} 1. \tag{2.1.1}$$

The better the mollification by ψ, the larger proportion of nonvanishing one can deduce.

It is natural to choose $\psi(\chi)$ such that

$$\psi(\chi) \approx \frac{1}{L\left(\frac{1}{2},\chi\right)}.$$

Since $L\left(\frac{1}{2},\chi\right)$ can be written as a Dirichlet series

$$L\left(\frac{1}{2},\chi\right) = \sum_{n=1}^{\infty}\frac{\chi(n)}{n^{\frac{1}{2}}}, \tag{2.1.2}$$

[1]First published in *Algebra & Number Theory* in volume 13 (2019), no. 1, published by Mathematical Sciences Publishers.

this suggests the choice

$$\psi(\chi) \approx \sum_{\ell \le y} \frac{\mu(\ell)\chi(\ell)}{\ell^{\frac{1}{2}}}. \tag{2.1.3}$$

We have introduced a truncation y in anticipation of the need to control various error terms that will arise. We write $y = Q^{\theta}$, where $\theta > 0$ is a real number. At least heuristically, larger values of θ yield better mollification by (2.1.3). Iwaniec and Sarnak [30] made this choice (2.1.3) (up to some smoothing), and found that the proportion of nonvanishing attained was

$$\frac{\theta}{1+\theta}. \tag{2.1.4}$$

When $\theta = 1$ we see (2.1.4) is exactly $\frac{1}{2}$, so we need $\theta > 1$ in order to conclude Theorem 0.2.1. This seems beyond the range of present technology. Without averaging over moduli we may take $\theta = \frac{1}{2} - \varepsilon$, and the asymptotic large sieve of Conrey, Iwaniec, and Soundararajan [58] allows one to take $\theta = 1 - \varepsilon$ if one averages over moduli. This just falls short of our goal.

Thus, a better mollifier than (2.1.3) is required. Part of the problem is that (2.1.2) is an inefficient representation of $L\left(\frac{1}{2}, \chi\right)$. A better representation of $L\left(\frac{1}{2}, \chi\right)$ may be obtained through the approximate functional equation, which states

$$L\left(\frac{1}{2}, \chi\right) \approx \sum_{n \le q^{1/2}} \frac{\chi(n)}{n^{\frac{1}{2}}} + \epsilon(\chi) \sum_{n \le q^{1/2}} \frac{\overline{\chi}(n)}{n^{\frac{1}{2}}}. \tag{2.1.5}$$

Here $\epsilon(\chi)$ is the root number, which is a complex number of modulus 1 defined by

$$\epsilon(\chi) = \frac{1}{q^{\frac{1}{2}}} \sum_{h (\mathrm{mod}\ q)} \chi(h) e\left(\frac{h}{q}\right). \tag{2.1.6}$$

Inspired by (2.1.5), Michel and VanderKam [45] chose a mollifier

$$\psi(\chi) \approx \sum_{\ell \le y} \frac{\mu(\ell)\chi(\ell)}{\ell^{\frac{1}{2}}} + \overline{\epsilon}(\chi) \sum_{\ell \le y} \frac{\mu(\ell)\overline{\chi}(\ell)}{\ell^{\frac{1}{2}}}. \tag{2.1.7}$$

We note that Soundararajan [59] earlier used a mollifier of this shape in the context of the Riemann zeta function.

For $y = Q^\theta$, Michel and VanderKam found that (2.1.7) gives a nonvanishing proportion of

$$\frac{2\theta}{1+2\theta}. \tag{2.1.8}$$

Thus, we need $\theta = \frac{1}{2} + \varepsilon$ in order for (2.1.8) to imply a proportion of nonvanishing greater than $\frac{1}{2}$. However, the more complicated nature of the mollifier (2.1.7) means that, without averaging over moduli, only the choice $\theta = \frac{3}{10} - \varepsilon$ is acceptable [32].

As we allow ourselves to average over moduli, however, one might hope to obtain (2.1.8) for $\theta = \frac{1}{2} + \varepsilon$. Again we fall just short of our goal. Using a powerful result of Deshouillers and Iwaniec on cancellation in sums of Kloosterman sums (see Lemma 2.4.1 below) we shall show that $\theta = \frac{1}{2} - \varepsilon$ is acceptable, but increasing θ any further seems very difficult. It follows that we need any extra amount of mollification in order to obtain a proportion of nonvanishing strictly greater than $\frac{1}{2}$.

The solution is to attach yet another piece to the mollifier $\psi(\chi)$, but here we wish for the mollifier to have a very different shape from (2.1.7). Such a mollifier was utilized by Bui [31], who showed that

$$\psi_{\mathrm{B}}(\chi) \approx \frac{1}{\log q} \sum\sum_{bc \le y} \frac{\Lambda(b)\mu(c)\overline{\chi}(b)\chi(c)}{(bc)^{\frac{1}{2}}} \tag{2.1.9}$$

is a mollifier for $L\left(\frac{1}{2}, \chi\right)$. It turns out that adding (2.1.9) to (2.1.7) gives a sufficient mollifier to conclude Theorem 0.2.1.

One may roughly motivate a mollifier of the shape (2.1.9) as follows. Working formally,

$$\begin{aligned}
\frac{1}{L(\frac{1}{2},\chi)} = \frac{L(\frac{1}{2},\overline{\chi})}{L(\frac{1}{2},\chi)L(\frac{1}{2},\overline{\chi})} &= \sum\sum\sum_{r,s,v} \frac{\overline{\chi}(r)\mu(s)\overline{\chi}(s)\mu(v)\chi(v)}{(rsv)^{\frac{1}{2}}} \\
&\approx \sum\sum\sum_{r,s,v} \frac{\log r}{\log q} \frac{\overline{\chi}(r)\mu(s)\overline{\chi}(s)\mu(v)\chi(v)}{(rsv)^{\frac{1}{2}}} \\
&= \frac{1}{\log q} \sum\sum_{u,v} \frac{(\mu \star \log)(u)\overline{\chi}(u)\mu(v)\chi(v)}{(uv)^{\frac{1}{2}}}.
\end{aligned}$$

One might wonder what percentage of nonvanishing one can obtain using only a mollifier of the shape (2.1.9). The analysis for Bui's mollifier is more complicated, and it does not seem possible to write down simple expressions like (2.1.4) or (2.1.8) that give a percentage of nonvanishing for (2.1.9) in terms of θ. If one assumes, perhaps optimistically, that averaging over moduli allows one to take any $\theta < 1$ in (2.1.9), then some numerical computation indicates that the nonvanishing percentage does not exceed 27%, say.

We remark that, in the course of the proof, the main terms are easily extracted and we have no need here for the averaging over moduli. We require the averaging over moduli in order to estimate some of the error terms.

The structure of the remainder of the chapter is as follows. In Section 2.2 we reduce the proof of Theorem 0.2.1 to two technical results, Lemma 2.2.3 and Lemma 2.2.4, which give asymptotic evaluations of certain mollified sums. In Section 2.3 we extract the main term of Lemma 2.2.3, and in Section 2.4 we use estimates on sums of Kloosterman sums to complete the proof of this lemma. Section 2.5 similarly proves the main term of Lemma 2.2.4, but this derivation is longer than that given in Section 2.3 because the main terms are more complicated. In the final section, Section 2.6, we bound the error term in Lemma 2.2.4, again using results on sums of Kloosterman sums.

2.2 Proof of Theorem 0.2.1: first steps

Let us fix some notation and conventions that shall hold for the remainder of the chapter.

The notation $a \equiv b(q)$ means $a \equiv b \pmod q$, and when $a(q)$ occurs beneath a sum it indicates a summation over residue classes modulo q.

We denote by ϵ an arbitrarily small positive quantity that may vary from one line to the next, or even within the same line. Thus, we may write $X^{2\epsilon} \leq X^{\epsilon}$ with no reservations.

We need to treat separately the even primitive characters and odd primitive characters. We focus exclusively on the even primitive characters, since the case of odd characters is nearly identical. We write $\sum^{+}_{\chi(q)}$ for a sum over even primitive characters modulo q, and we write $\varphi^+(q)$ for the number of such characters. Observe that $\varphi^+(q) = \frac{1}{2}\varphi^*(q) + O(1)$.

We shall encounter the Ramanujan sum $c_q(n)$ (see the proof of Proposition 2.4.2), defined by

$$c_q(n) = \sum_{\substack{a(q)\\(a,q)=1}} e\left(\frac{an}{q}\right).$$

We shall only need to know that $c_q(1) = \mu(q)$ and $|c_q(n)| \leq (q,n)$, where (q,n) is the greatest common divisor of q and n.

We now fix a smooth function Ψ as in the statement of Theorem 0.2.1, and allow all implied constants to depend on Ψ. We let Q be a large real number, and set $y_i = Q^{\theta_i}$ for $i \in \{1,2,3\}$, where $0 < \theta_i < \frac{1}{2}$ are fixed real numbers. We further define $L = \log Q$. The notation $o(1)$ denotes a quantity that goes to zero as Q goes to infinity.

Let us now begin the proof of Theorem 0.2.1 in earnest. As discussed in Section 2.1, we choose our mollifier $\psi(\chi)$ to have the form

$$\psi(\chi) = \psi_{\mathrm{IS}}(\chi) + \psi_{\mathrm{B}}(\chi) + \psi_{\mathrm{MV}}(\chi), \tag{2.2.1}$$

where

$$\begin{aligned} \psi_{\mathrm{IS}}(\chi) &= \sum_{\ell \leq y_1} \frac{\mu(\ell)}{\ell^{\frac{1}{2}}} P_1\left(\frac{\log(y_1/\ell)}{\log y_1}\right), \\ \psi_{\mathrm{B}}(\chi) &= \frac{1}{L} \sum\sum_{bc \leq y_2} \frac{\Lambda(b)\mu(c)\overline{\chi}(b)\chi(c)}{(bc)^{\frac{1}{2}}} P_2\left(\frac{\log(y_2/bc)}{\log y_2}\right), \\ \psi_{\mathrm{MV}}(\chi) &= \epsilon(\overline{\chi}) \sum_{\ell \leq y_3} \frac{\mu(\ell)\overline{\chi}(\ell)}{\ell^{\frac{1}{2}}} P_3\left(\frac{\log(y_3/\ell)}{\log y_3}\right). \end{aligned} \tag{2.2.2}$$

The smoothing polynomials P_i are real and satisfy $P_i(0) = 0$. For notational convenience we write

$$P_i\left(\frac{\log(y_i/x)}{\log y_i}\right) = P_i[x].$$

There is some ambiguity in this notation because of the y_i-dependence in the polynomials, and this needs to be remembered in calculations.

Now define sums S_1 and S_2 by

$$S_1 = \sum_q \Psi\left(\frac{q}{Q}\right) \frac{q}{\varphi(q)} \sum_{\chi(q)}{}^{+} L\left(\frac{1}{2}, \chi\right) \psi(\chi),$$
$$S_2 = \sum_q \Psi\left(\frac{q}{Q}\right) \frac{q}{\varphi(q)} \sum_{\chi(q)}{}^{+} \left|L\left(\frac{1}{2}, \chi\right) \psi(\chi)\right|^2. \tag{2.2.3}$$

We apply Cauchy-Schwarz as in (2.1.1) and get

$$\sum_q \Psi\left(\frac{q}{Q}\right) \frac{q}{\varphi(q)} \sum_{\substack{\chi(q) \\ L\left(\frac{1}{2},\chi\right)\neq 0}}{}^{+} 1 \geq \frac{S_1^2}{S_2}. \tag{2.2.4}$$

The proof of Theorem 0.2.1 therefore reduces to estimating S_1 and S_2. We obtain asymptotic formulas for these two sums.

Lemma 2.2.1. *Suppose* $0 < \theta_1, \theta_2 < 1$ *and* $0 < \theta_3 < \frac{1}{2}$. *Then*

$$S_1 = \left(P_1(1) + P_3(1) + \frac{\theta_2}{2}\widetilde{P_2}(1) + o(1)\right) \sum_q \Psi\left(\frac{q}{Q}\right) \frac{q}{\varphi(q)} \varphi^+(q),$$

where

$$\widetilde{P_2}(x) = \int_0^x P_2(u)du.$$

Lemma 2.2.2. *Let* $0 < \theta_1, \theta_2, \theta_3 < \frac{1}{2}$ *with* $\theta_2 < \theta_1, \theta_3$. *Then*

$$S_2 = \left(2P_1(1)P_3(1) + P_3(1)^2 + \frac{1}{\theta_3}\int_0^1 P_3'(x)^2 dx + \kappa + \lambda + o(1)\right)$$
$$\times \sum_q \Psi\left(\frac{q}{Q}\right) \frac{q}{\varphi(q)} \varphi^+(q),$$

where

$$\kappa = 3\theta_2 P_3(1)\widetilde{P_2}(1) - 2\theta_2 \int_0^1 P_2(x)P_3(x)dx$$

and

$$\lambda = P_1(1)^2 + \frac{1}{\theta_1}\int_0^1 P_1'(x)^2dx - \theta_2 P_1(1)\widetilde{P_2}(1) + 2\theta_2\int_0^1 P_1\left(1 - \frac{\theta_2(1-x)}{\theta_1}\right)P_2(x)dx$$
$$+ \frac{\theta_2}{\theta_1}\int_0^1 P_1'\left(1 - \frac{\theta_2(1-x)}{\theta_1}\right)P_2(x)dx + \theta_2^2\int_0^1 (1-x)P_2(x)^2dx$$
$$+ \frac{\theta_2}{2}\int_0^1 (1-x)^2P_2'(x)^2dx - \frac{\theta_2^2}{4}\widetilde{P_2}(1)^2 + \frac{\theta_2}{4}\int_0^1 P_2(x)^2dx.$$

Proof of Theorem 0.2.1. Lemmas 2.2.1 and 2.2.2 give the evaluations of S_1 and S_2 for even characters. The identical formulas hold for odd characters. Theorem 0.2.1 then follows from (2.2.4) upon choosing $\theta_1 = \theta_3 = \frac{1}{2}$, $\theta_2 = 0.163$, and

$$P_1(x) = 4.86x + 0.29x^2 - 0.96x^3 + 0.974x^4 - 0.17x^5,$$
$$P_2(x) = -3.11x - 0.3x^2 + 0.87x^3 - 0.18x^4 - 0.53x^5,$$
$$P_3(x) = 4.86x + 0.06x^2.$$

These choices actually yield a proportion[2]

$$\geq 0.50073004\ldots,$$

which allows us to state Theorem 0.2.1 with a clean inequality. □

We note without further comment the curiosity in the proof of Theorem 0.2.1 that the largest permissible value of θ_2 is not optimal.

We can dispense with S_1 quickly.

Proof of Lemma 2.2.1. Apply [31, Theorem 2.1] and the argument of [45, Section 3], using the facts $L = \log q + O(1)$ and $y_i = q^{\theta_i + o(1)}$. □

The analysis of S_2 is much more involved, and we devote the remainder of the paper to this task. We first observe that (2.2.1) yields

$$|\psi(\chi)|^2 = |\psi_{\rm IS}(\chi) + \psi_{\rm B}(\chi)|^2 + 2{\rm Re}\left\{\psi_{\rm IS}(\chi)\psi_{\rm MV}(\overline{\chi}) + \psi_{\rm B}(\chi)\psi_{\rm MV}(\overline{\chi})\right\} + |\psi_{\rm MV}(\chi)|^2.$$

[2] A Mathematica® file with this computation can be found at https://arxiv.org/abs/1804.01445.

By [31, Theorem 2.2] we have

$$\sum_{\chi(q)}^{+} \left| L\left(\frac{1}{2}, \chi\right) \right|^2 |\psi_{\mathrm{IS}}(\chi) + \psi_{\mathrm{B}}(\chi)|^2 = \lambda \varphi^+(q) + O\left(qL^{-1+\epsilon}\right),$$

where λ is as in Lemma 2.2.2. We also have

$$\begin{aligned}\frac{1}{\varphi^+(q)} \sum_{\chi(q)}^{+} \left| L\left(\frac{1}{2}, \chi\right) \right|^2 |\psi_{\mathrm{MV}}(\chi)|^2 &= \frac{1}{\varphi^+(q)} \sum_{\chi(q)}^{+} \left| L\left(\frac{1}{2}, \chi\right) \right|^2 \left| \sum_{\ell \leq y_3} \frac{\mu(\ell)\chi(\ell)P_3[\ell]}{\ell^{\frac{1}{2}}} \right|^2 \\ &= P_3(1)^2 + \frac{1}{\theta_3} \int_0^1 P_3'(x)^2 dx + O(L^{-1+\epsilon}),\end{aligned}$$

by the analysis of the Iwaniec-Sarnak mollifier (see [31, Section 2.3]).

Therefore, in order to prove Lemma 2.2.2 it suffices to prove the following two results.

Lemma 2.2.3. *For $0 < \theta_1, \theta_3 < \frac{1}{2}$ we have*

$$\begin{aligned}&\sum_q \Psi\left(\frac{q}{Q}\right) \frac{q}{\varphi(q)} \sum_{\chi(q)}^{+} \left| L\left(\frac{1}{2}, \chi\right) \right|^2 \psi_{IS}(\chi)\psi_{MV}(\overline{\chi}) \\ &\quad = (P_1(1)P_3(1) + o(1)) \sum_q \Psi\left(\frac{q}{Q}\right) \frac{q}{\varphi(q)} \varphi^+(q).\end{aligned}$$

Lemma 2.2.4. *Let $0 < \theta_2 < \theta_3 < \frac{1}{2}$. Then*

$$\begin{aligned}&\sum_q \Psi\left(\frac{q}{Q}\right) \frac{q}{\varphi(q)} \sum_{\chi(q)}^{+} \left| L\left(\frac{1}{2}, \chi\right) \right|^2 \psi_B(\chi)\psi_{MV}(\overline{\chi}) \\ &\quad = \left(\frac{3\theta_2}{2} P_3(1)\widetilde{P_2}(1) - \theta_2 \int_0^1 P_2(x)P_3(x)dx + o(1) \right) \sum_q \Psi\left(\frac{q}{Q}\right) \frac{q}{\varphi(q)} \varphi^+(q).\end{aligned}$$

2.3 Lemma 2.2.3: main term

The goal of this section is to extract the main term in Lemma 2.2.3. The main term analysis is given in [45, Section 6], but as the ideas also appear in the proof of Lemma 2.2.4 we give details here.

We begin with two lemmas.

Lemma 2.3.1. *Let χ be a primitive even character modulo q. Let $G(s)$ be an even polynomial satisfying $G(0) = 1$, and which vanishes to second order at $\frac{1}{2}$. Then we have*

$$\left|L\left(\frac{1}{2}, \chi\right)\right|^2 = 2\sum_{m,n}\sum \frac{\chi(m)\overline{\chi}(n)}{(mn)^{\frac{1}{2}}} V\left(\frac{mn}{q}\right),$$

where

$$V(x) = \frac{1}{2\pi i}\int_{(1)} \frac{\Gamma^2\left(\frac{s}{2}+\frac{1}{4}\right)}{\Gamma^2\left(\frac{1}{4}\right)} \frac{G(s)}{s}\pi^{-s}x^{-s}ds. \tag{2.3.1}$$

Proof. See [30, (2.5)]. The result follows along the lines of [47, Theorem 5.3]. □

We remark that V satisfies $V(x) \ll_A (1+x)^{-A}$, as can be seen by moving the contour of integration to the right. We also note that the choice of $G(s)$ in Lemma 3.3.2 is almost completely free. In particular, we may choose G to vanish at whichever finite set of points is convenient for us (see (2.3.6) below for an application).

Lemma 2.3.2. *Let $(mn, q) = 1$. Then*

$$\sum_{\chi(q)}^{+} \chi(m)\overline{\chi}(n) = \frac{1}{2}\sum_{\substack{vw=q \\ w|m\pm n}}\sum \mu(v)\varphi(w).$$

Proof. See [37, Lemma 4.1], for instance. □

We do not need the averaging over q in order to extract the main term of Lemma 2.2.3. We insert the definitions of the mollifiers $\psi_{\mathrm{IS}}(\chi)$ and $\psi_{\mathrm{MV}}(\overline{\chi})$, then apply Lemma 3.3.2 and interchange orders of summation. We obtain

$$\begin{aligned}
&\sum_{\chi(q)}^{+} \left|L\left(\frac{1}{2}, \chi\right)\right|^2 \psi_{\mathrm{IS}}(\chi)\psi_{\mathrm{MV}}(\overline{\chi}) \\
&= 2\sum_{\substack{\ell_1 \leq y_1 \\ \ell_3 \leq y_3 \\ (\ell_1\ell_3, q)=1}}\sum \frac{\mu(\ell_1)\mu(\ell_3)P_1[\ell_1]P_3[\ell_3]}{(\ell_1\ell_3)^{\frac{1}{2}}} \sum_{(mn,q)=1}\sum \frac{1}{(mn)^{\frac{1}{2}}} V\left(\frac{mn}{q}\right) \\
&\times \sum_{\chi(q)}^{+} \epsilon(\chi)\chi(m\ell_1\ell_3)\overline{\chi}(n).
\end{aligned} \tag{2.3.2}$$

Opening $\epsilon(\chi)$ using (2.1.6) and applying Lemma 2.3.2, we find after some work (see

[30, (3.4) and (3.7)]) that

$$\sum_{\chi(q)}^{+} \epsilon(\chi)\chi(m\ell_1\ell_3)\overline{\chi}(n) = \frac{1}{q^{1/2}} \sum_{\substack{vw=q \\ (v,w)=1}}\sum \mu^2(v)\varphi(w)\cos\left(\frac{2\pi n\overline{m\ell_1\ell_3 v}}{w}\right). \qquad (2.3.3)$$

The main term comes from $m\ell_1\ell_3 = 1$. With this constraint in place we apply character orthogonality in reverse, obtaining that the main term $M_{\mathrm{IS,MV}}$ of Lemma 2.2.3 is

$$M_{\mathrm{IS,MV}} = 2P_1(1)P_3(1)\sum_{\chi(q)}^{+} \epsilon(\chi)\sum_n \frac{\overline{\chi}(n)}{n^{1/2}} V\left(\frac{n}{q}\right).$$

We have the following proposition.

Proposition 2.3.3. *Let χ be a primitive even character modulo q, and let $T > 0$ be a real number. Let V be defined as in* (2.3.1)*. Then*

$$\sum_n \frac{\overline{\chi}(n)}{n^{\frac{1}{2}}} V\left(\frac{Tn}{q}\right) = L\left(\frac{1}{2}, \overline{\chi}\right) - \epsilon(\overline{\chi})\sum_n \frac{\chi(n)}{n^{\frac{1}{2}}} F\left(\frac{n}{T}\right),$$

where

$$F(x) = \frac{1}{2\pi i}\int_{(1)} \frac{\Gamma\left(\frac{s}{2}+\frac{1}{4}\right)\Gamma\left(-\frac{s}{2}+\frac{1}{4}\right)}{\Gamma^2\left(\frac{1}{4}\right)} \frac{G(s)}{s} x^{-s} ds. \qquad (2.3.4)$$

Before proving Proposition 2.3.3, let us see how to use it to finish the evaluation of $M_{\mathrm{IS,MV}}$. Proposition 2.3.3 gives

$$M_{\mathrm{IS,MV}} = 2P_1(1)P_3(1)\sum_{\chi(q)}^{+} \epsilon(\chi) L\left(\frac{1}{2}, \overline{\chi}\right) - 2P_1(1)P_3(1)\sum_{\chi(q)}^{+}\sum_n \frac{\chi(n)}{n^{\frac{1}{2}}} F(n),$$

and by the first moment analysis (see [45, Section 3], also Section 2.5 below) we have

$$2P_1(1)P_3(1)\sum_{\chi(q)}^{+} \epsilon(\chi) L\left(\frac{1}{2}, \overline{\chi}\right) = (1+o(1))2P_1(1)P_3(1)\varphi^+(q). \qquad (2.3.5)$$

For the other piece, we apply Lemma 2.3.2 to obtain

$$-2P_1(1)P_3(1)\sum_{\chi(q)}^{+}\sum_n \frac{\chi(n)}{n^{\frac{1}{2}}} F(n) = -P_1(1)P_3(1)\sum_{w|q}\varphi(w)\mu(q/w)\sum_{\substack{n\equiv\pm1(w) \\ (n,q)=1}} \frac{1}{n^{\frac{1}{2}}} F(n).$$

We choose G to vanish at all the poles of

$$\Gamma\left(\frac{s}{2}+\frac{1}{4}\right)\Gamma\left(-\frac{s}{2}+\frac{1}{4}\right)$$

in the disc $|s| \leq A$, where $A > 0$ is large but fixed. By moving the contour of integration to the right we see

$$F(x) \ll \frac{1}{(1+x)^{100}}, \tag{2.3.6}$$

say, and therefore the contribution from $n > q^{\frac{1}{10}}$ is negligible. By trivial estimation the contribution from $w \leq q^{\frac{1}{4}}$ is also negligible. For $w > q^{\frac{1}{4}}$ and $n \leq q^{\frac{1}{10}}$, we can only have $n \equiv \pm 1 \pmod{w}$ if $n = 1$. Adding back in the terms with $n \leq q^{\frac{1}{4}}$, the contribution from these terms is therefore

$$-(1+o(1))2P_1(1)P_3(1)F(1)\varphi^+(q). \tag{2.3.7}$$

Since the integrand in $F(1)$ is odd, we may evaluate $F(1)$ through a residue at $s = 0$. We shift the line of integration in (2.3.4) to $\mathrm{Re}(s) = -1$, picking up a contribution from the simple pole at $s = 0$. In the integral on the line $\mathrm{Re}(s) = -1$ we change variables $s \to -s$. This yields the relation $F(1) = 1 - F(1)$, whence $F(1) = \frac{1}{2}$. Combining (2.3.5) and (2.3.7), we obtain

$$M_{\mathrm{IS,MV}} = (1+o(1))P_1(1)P_3(1)\varphi^+(q),$$

as desired. This yields the main term of Lemma 2.2.3.

Proof of Proposition 2.3.3. We write V using its definition and interchange orders of summation and integration to get

$$\sum_n \frac{\overline{\chi}(n)}{n^{1/2}} V\left(\frac{Tn}{q}\right) = \frac{1}{2\pi i}\int_{(1)} \frac{\Gamma^2\left(\frac{s}{2}+\frac{1}{4}\right)}{\Gamma^2\left(\frac{1}{4}\right)} \frac{G(s)}{s}\left(\frac{q}{\pi}\right)^s T^{-s} L\left(\frac{1}{2}+s,\overline{\chi}\right) ds.$$

We move the line of integration to $\mathrm{Re}(s) = -1$, picking up a contribution of $L\left(\frac{1}{2},\overline{\chi}\right)$ from the pole at $s = 0$. Observe that we do not get any contribution from the double pole of $\Gamma^2\left(\frac{s}{2}+\frac{1}{4}\right)$ at $s = -\frac{1}{2}$ because of our assumption that G vanishes at $s = \pm\frac{1}{2}$ to second order.

Now, for the integral on the line $\mathrm{Re}(s) = -1$, we apply the functional equation for

$L\left(\frac{1}{2}+s,\overline{\chi}\right)$ and then change variables $s \to -s$ to obtain

$$-\epsilon(\overline{\chi})\frac{1}{2\pi i}\int_{(1)}\frac{\Gamma\left(\frac{s}{2}+\frac{1}{4}\right)\Gamma\left(-\frac{s}{2}+\frac{1}{4}\right)}{\Gamma^2\left(\frac{1}{4}\right)}\frac{G(s)}{s}T^s L\left(\frac{1}{2}+s,\chi\right)ds.$$

The desired result follows by expanding $L\left(\frac{1}{2}+s,\chi\right)$ in its Dirichlet series and interchanging the order of summation and integration. □

2.4 Lemma 2.2.3: error term

Here we show that the remainder of the terms in (2.3.2) (those with $m\ell_1\ell_3 \neq 1$) contribute only to the error term of Lemma 2.2.3. Here we must avail ourselves of the averaging over q.

Inserting (2.3.3) into (2.3.2) and averaging over moduli, we wish to show that

$$\begin{aligned}\mathcal{E}_1 = &\sum\sum_{(v,w)=1}\mu^2(v)\frac{v}{\varphi(v)}\frac{w^{\frac{1}{2}}}{v^{\frac{1}{2}}}\Psi\left(\frac{vw}{Q}\right)\sum\sum_{\substack{\ell_1\le y_1\\ \ell_3\le y_3\\ (\ell_1\ell_3,vw)=1}}\frac{\mu(\ell_1)\mu(\ell_3)P_1[\ell_1]P_3[\ell_3]}{(\ell_1\ell_3)^{\frac{1}{2}}}\\ &\times\sum\sum_{(mn,vw)=1}\frac{1}{(mn)^{\frac{1}{2}}}Z\left(\frac{mn}{vw}\right)\cos\left(\frac{2\pi n\overline{m\ell_1\ell_3 v}}{w}\right)\ll Q^{2-\epsilon+o(1)},\end{aligned}\tag{2.4.1}$$

where $m\ell_1\ell_3 \neq 1$, but we do not indicate this in the notation. The function Z is actually just V in (2.3.1), but we do not wish to confuse the function V with the scale V that shall appear shortly.

Observe that the arithmetic weight $\frac{q}{\varphi(q)}$ has become $\frac{v}{\varphi(v)}\frac{w}{\varphi(w)}$ by multiplicativity, and that this factor of $\varphi(w)$ has canceled with $\varphi(w)$ in (2.3.3), making the sum on w smooth.

The main tool we use to bound $\mathcal{E}_1'$ is the following result, due to Deshouillers and Iwaniec, on cancellation in sums of Kloosterman sums.

Lemma 2.4.1. *Let C, D, N, R, S be positive numbers, and let $b_{n,r,s}$ be a complex sequence supported in $(0,N]\times(R,2R]\times(S,2S]\cap\mathbb{N}^3$. Let $g_0(\xi,\eta)$ be a smooth function* *having compact support in* $\mathbb{R}^+ \times \mathbb{R}^+$*, and let* $g(c,d) = g_0(c/C, d/D)$*. Then*

$$\sum_c \sum_d \sum_{\substack{n \\ (rd,sc)=1}} \sum_r \sum_s b_{n,r,s} g(c,d) e\left(n \frac{\overline{rd}}{sc}\right)$$
$$\ll_{\epsilon, g_0} (CDNRS)^\epsilon K(C,D,N,R,S) \|b_{N,R,S}\|_2,$$

where

$$\|b_{N,R,S}\|_2 = \left(\sum_n \sum_r \sum_s |b_{n,r,s}|^2 \right)^{\frac{1}{2}}$$

and

$$K(C,D,N,R,S)^2 = CS(RS+N)(C+RD) + C^2 DS\sqrt{(RS+N)R} + D^2 NR.$$

Proof. This is the corrected version [60, Lemma 2.1] of [61, Lemma 1], which is an easy consequence of a corrected version of [62, Theorem 12]. □

We need to massage (2.4.1) before it is in a form where an application of Lemma 2.4.1 is appropriate. Let us briefly describe our plan of attack. We apply partitions of unity to localize the variables and then separate variables with integral transforms. By using the orthogonality of multiplicative characters we will be able to assume that v is quite small, which is advantageous when it comes time to remove coprimality conditions involving v. We next reduce to the case in which n is somewhat small. This is due to the fact that the sum on n is essentially a Ramanujan sum, and Ramanujan sums experience better than squareroot cancellation on average. We next use Möbius inversion to remove the coprimality condition between n and w. This application of Möbius inversion introduces a new variable, call if f, and another application of character orthogonality allows us to assume f is small. We then remove the coprimality conditions on m. We finally apply Lemma 2.4.1 to get the desired cancellation, and it is crucial here that f and v are no larger than Q^ϵ.

Let us turn to the details in earnest. We apply smooth partitions of unity (see [63, Lemma 1.6], for instance) in all variables, so that $\mathcal{E}_1$ can be written

$$\sum_{M,N,L_1,L_3,V,W} \cdots \sum \mathcal{E}_1(M,N,L_1,L_3,V,W), \tag{2.4.2}$$

where

$$\mathcal{E}_1(M,N,L_1,L_3,V,W) = \sum_{(v,w)=1}\sum \mu^2(v)\frac{v}{\varphi(v)}\frac{w^{\frac{1}{2}}}{v^{\frac{1}{2}}}\Psi\left(\frac{vw}{Q}\right)G\left(\frac{v}{V}\right)G\left(\frac{w}{W}\right)$$
$$\times \sum_{\substack{\ell_1\leq y_1\\ \ell_3\leq y_3\\ (\ell_1\ell_3,vw)=1}}\sum \frac{\mu(\ell_1)\mu(\ell_3)P_1[\ell_1]P_3[\ell_3]}{(\ell_1\ell_3)^{\frac{1}{2}}}G\left(\frac{\ell_1}{L_1}\right)G\left(\frac{\ell_3}{L_3}\right)$$
$$\times \sum_{(mn,vw)=1}\sum \frac{1}{(mn)^{\frac{1}{2}}}Z\left(\frac{mn}{vw}\right)G\left(\frac{m}{M}\right)G\left(\frac{n}{N}\right)\cos\left(\frac{2\pi n\overline{m\ell_1\ell_3 v}}{w}\right).$$

Here G is a smooth, nonnegative function supported in $[\frac{1}{2},2]$, and the numbers M,N,L_i,V,W in (2.4.2) range over powers of two. We may assume

$$M,N,L_1,L_3,V,W \gg 1, \qquad VW \asymp Q, \qquad L_i \ll y.$$

Furthermore, by the rapid decay of Z we may assume $MN \leq Q^{1+\epsilon}$. Thus, the number of summands $\mathcal{E}_1(M,\ldots,W)$ in (2.4.2) is $\ll Q^{o(1)}$.

Up to changing the definition of G, we may rewrite $\mathcal{E}_1(M,\cdots,W)$ as

$$\mathcal{E}_1(M,N,L_1,L_3,V,W) = \frac{W^{\frac{1}{2}}}{(MNL_1L_3V)^{\frac{1}{2}}}\sum_{(v,w)=1}\sum \alpha(v)G\left(\frac{v}{V}\right)G\left(\frac{w}{W}\right)\Psi\left(\frac{vw}{Q}\right)$$
$$\times \sum_{\substack{\ell_i\leq y_i\\ (\ell_i,vw)=1}}\sum \beta(\ell_1)\gamma(\ell_3)G\left(\frac{\ell_1}{L_1}\right)G\left(\frac{\ell_3}{L_3}\right)$$
$$\times \sum_{(mn,vw)=1}\sum Z\left(\frac{mn}{vw}\right)G\left(\frac{m}{M}\right)G\left(\frac{n}{N}\right)\cos\left(\frac{2\pi n\overline{m\ell_1\ell_3 v}}{w}\right),$$

where α,β,γ are sequences satisfying $|\alpha(v)|,|\beta(\ell_1)|,|\gamma(\ell_3)| \ll Q^{o(1)}$.

We separate the variables in Z by writing Z using its definition as an integral (2.3.1) and moving the line of integration to $\mathrm{Re}(s) = L^{-1}$. By the rapid decay of the Γ function in vertical strips we may restrict to $|\mathrm{Im}(s)| \leq Q^\epsilon$. We similarly separate the variables in Ψ using the inverse Mellin transform. Therefore, up to changing the definition of some of the functions G, it suffices to prove that

$$\begin{aligned}\mathcal{E}_1'(M,N,L_1,L_3,V,W) = \frac{W^{\frac{1}{2}}}{(MNL_1L_3V)^{\frac{1}{2}}} &\sum\sum_{(v,w)=1} \alpha(v) G\left(\frac{v}{V}\right) G\left(\frac{w}{W}\right) \\ &\times \sum\sum_{\substack{\ell_i \le y_i \\ (\ell_i, vw)=1}} \beta(\ell_1)\gamma(\ell_3) G\left(\frac{\ell_3}{L_3}\right) G\left(\frac{\ell_3}{L_3}\right) \\ &\times \sum\sum_{(mn,vw)=1} G\left(\frac{m}{M}\right) G\left(\frac{n}{N}\right) e\left(\frac{n\overline{m\ell_1\ell_3 v}}{w}\right) \ll Q^{2-\epsilon+o(1)}.\end{aligned} \tag{2.4.3}$$

Our smooth functions G all satisfy $G^{(j)}(x) \ll_j Q^{j\epsilon}$ for $j \geq 0$. To save on space we write the left side of (2.4.3) as simply $\mathcal{E}_1'$.

Observe that the trivial bound for $\mathcal{E}_1'$ is

$$\mathcal{E}_1' \ll V^{\frac{1}{2}} W^{\frac{3}{2}} (MN)^{\frac{1}{2}} (L_1L_3)^{\frac{1}{2}} Q^{o(1)} \ll \frac{Q^{2+\epsilon}(y_1y_3)^{\frac{1}{2}}}{V}. \tag{2.4.4}$$

This bound is worst when V is small. Since y_i will be taken close to $Q^{\frac{1}{2}}$, we therefore need to save $\approx Q^{\frac{1}{2}}$ in order to obtain (2.4.1). The trivial bound does show, however, that the contribution from $V > Q^{\frac{1}{2}+2\epsilon}$ is acceptably small, and we may therefore assume that $V \leq Q^{\frac{1}{2}+2\epsilon}$. Note this implies $W \gg Q^{\frac{1}{2}-\epsilon}$.

We now reduce to the case $V \ll Q^{\epsilon}$. We accomplish this by re-introducing multiplicative characters. The orthogonality of multiplicative characters yields

$$e\left(\frac{n\overline{m\ell_1\ell_3 v}}{w}\right) = \frac{1}{\varphi(w)} \sum_{\chi(w)} \tau(\overline{\chi})\chi(n)\overline{\chi}(m\ell_1\ell_3 v). \tag{2.4.5}$$

Using the Gauss sum bound $|\tau(\overline{\chi})| \ll w^{\frac{1}{2}}$ we then arrange $\mathcal{E}_1'$ as

$$\mathcal{E}_1' \ll \frac{W}{(MNL_1L_3V)^{\frac{1}{2}}} \sum_{w \asymp W} \frac{1}{\varphi(w)} \sum_{v \asymp V} \left| \sum\sum_{(mn,v)=1} \chi(n)\overline{\chi}(m) \right| \left| \sum\sum_{(\ell_1\ell_3,v)=1} \overline{\chi}(\ell_1\ell_3) \right|,$$

where we have suppressed some things in the notation for brevity. By Cauchy-Schwarz and character orthogonality we obtain

$$\sum_{\chi(w)} \left| \sum\sum_{m,n} \right| \left| \sum\sum_{\ell_1,\ell_3} \right| \ll Q^{o(1)} (MNL_1L_3)^{\frac{1}{2}} (MN+W)^{\frac{1}{2}} (L_1L_3+W)^{\frac{1}{2}},$$

which yields a bound of

$$Q^{-o(1)}\mathcal{E}_1' \ll \frac{Q(MN)^{\frac{1}{2}}(y_1y_3)^{\frac{1}{2}}}{V^{\frac{1}{2}}} + \frac{Q^{\frac{3}{2}}(MN)^{1/2}}{V} + \frac{Q^{\frac{3}{2}}(y_1y_3)^{\frac{1}{2}}}{V} + \frac{Q^2}{V^{\frac{3}{2}}}. \qquad (2.4.6)$$

We observe that (2.4.6) is acceptable for $V \geq Q^{3\epsilon}$, say. We may therefore assume $V \leq Q^{\epsilon}$.

We next show that $\mathcal{E}_1'$ is small provided N is somewhat large.

Proposition 2.4.2. *Assume the hypotheses of Lemma 2.2.3. If* $N \geq MQ^{-2\epsilon}$ *and* $m\ell_1\ell_3 \neq 1$, *then* $\mathcal{E}_1' \ll Q^{2-\epsilon+o(1)}$.

Proof. We make use only of cancellation in the sum on n, say

$$\Sigma_N = \sum_{(n,vw)=1} G\left(\frac{n}{N}\right) e\left(\frac{n\overline{m\ell_1\ell_3 v}}{w}\right).$$

We use Möbius inversion to detect the condition $(n,v)=1$, and then break n into primitive residue classes modulo w. Thus

$$\Sigma_N = \sum_{d|v} \mu(d) \sum_{(a,w)=1} e\left(\frac{ad\overline{m\ell_1\ell_3 v}}{w}\right) \sum_{n\equiv a(w)} G\left(\frac{dn}{N}\right).$$

We apply Poisson summation to each sum on n, and obtain

$$\Sigma_N = \sum_{d|v} \mu(d) \sum_{(a,w)=1} e\left(\frac{ad\overline{m\ell_1\ell_3 v}}{w}\right) \frac{N}{dw} \sum_{|h|\leq W^{1+\epsilon}d/N} e\left(\frac{ah}{w}\right) \widehat{G}\left(\frac{hN}{dw}\right) + O_\epsilon\left(Q^{-100}\right),$$

say. The contribution of the error term is, of course, negligible. The contribution of the zero frequency $h=0$ to Σ_N is

$$\widehat{G}(0)\frac{N}{w}\sum_{d|v}\frac{\mu(d)}{d}\sum_{(a,w)=1} e\left(\frac{ad\overline{m\ell_1\ell_3 v}}{w}\right) = \widehat{G}(0)\mu(w)\frac{N}{w}\frac{\varphi(v)}{v},$$

and upon summing this contribution over the remaining variables, the zero frequency contributes

$$\ll V^{\frac{1}{2}}W^{\frac{1}{2}}(MN)^{\frac{1}{2}}(y_1y_3)^{\frac{1}{2}}Q^{o(1)} \ll Q^{\frac{3}{2}}$$

to $\mathcal{E}_1'$, and this contribution is sufficiently small.

It takes just a bit more work to bound the contribution of the nonzero frequencies $|h| > 0$. We rearrange the sum as

$$\sum_{d|v} \mu(d) \frac{N}{dw} \sum_{|h| \leq W^{1+\epsilon} d/N} \widehat{G}\left(\frac{hN}{dw}\right) \sum_{(a,w)=1} e\left(\frac{ad\overline{m\ell_1\ell_3 v}}{w} + \frac{ah}{w}\right).$$

By a change of variables the inner sum is equal to the Ramanujan sum $c_w(hm\ell_1\ell_3 v + d)$. Note that $hm\ell_1\ell_3 v + d \neq 0$ because $m\ell_1\ell_3 \neq 1$. The nonzero frequencies therefore contribute to $\mathcal{E}_1'$ an amount

$$\ll Q^\epsilon \frac{(VWL_1L_3M)^{\frac{1}{2}}}{N^{\frac{1}{2}}} \sup_{0<|k|\ll Q^{O(1)}} \sum_{w \asymp W} |c_w(k)|.$$

Since $|c_w(k)| \leq (k, w)$ the sum on w is $\ll W^{1+o(1)}$. It follows that

$$\mathcal{E}_1' \ll Q^{\frac{3}{2}} + Q^{\frac{3}{2}+\epsilon}(y_1y_3)^{\frac{1}{2}} \frac{M^{1/2}}{N^{\frac{1}{2}}}.$$

Since $y_i = Q^{\theta_i}$ and $\theta_i < \frac{1}{2} - 3\epsilon$, say, this bound for $\mathcal{E}_1'$ is acceptable provided $N \geq MQ^{-2\epsilon}$. □

By Proposition 2.4.2 we may assume $N \leq MQ^{-2\epsilon}$. Since $MN \leq Q^{1+\epsilon}$, the condition $N \leq MQ^{-2\epsilon}$ implies $N \leq Q^{\frac{1}{2}}$.

We now pause to make a comment on the condition $m\ell_1\ell_3 \neq 1$, which we have assumed throughout this section but not indicated in the notation for $\mathcal{E}_1'$. Observe that this condition is automatic if $ML_1L_3 > 2019$, say. If $ML_1L_3 \ll 1$, then we may use the trivial bound (2.4.4) along with the bound $N \leq Q^{\frac{1}{2}} \leq Q^{1-\epsilon}$ to obtain

$$\mathcal{E}_1' \ll Q^{2-\epsilon}.$$

We may therefore assume $ML_1L_2 \gg 1$, so that the condition $m\ell_1\ell_3 \neq 1$ is satisfied.

We now remove the coprimality condition $(n, w) = 1$. By Möbius inversion we have

$$\mathbf{1}_{(n,w)=1} = \sum_{\substack{f|n \\ f|w}} \mu(f).$$

We move the sum on f to be the outermost sum, and note $f \ll N$. We then change variables $n \to nf, w \to wf$. If a_*, say, is any lift of the multiplicative inverse of $m\ell_1\ell_3 v$

modulo wf, then $a_* \equiv \overline{m\ell_1\ell_3 v} \pmod{w}$, and therefore

$$\frac{nf\overline{m\ell_1\ell_3 v}}{wf} \equiv \frac{n\overline{m\ell_1\ell_3 v}}{w} \pmod{1}.$$

It follows that

$$\mathcal{E}_1' = \frac{W^{\frac{1}{2}}}{(MNL_1L_3V)^{\frac{1}{2}}} \sum_{f \ll N} \mu(f) \sum\sum_{(v,wf)=1} \alpha(v) G\left(\frac{v}{V}\right) G\left(\frac{wf}{W}\right)$$
$$\times \sum_{\substack{\ell_i \le y_i \\ (\ell_i, fvw)=1}}\sum \beta(\ell_1)\gamma(\ell_3) G\left(\frac{\ell_1}{L_1}\right) G\left(\frac{\ell_3}{L_3}\right) \sum_{\substack{(m,fvw)=1 \\ (n,v)=1}}\sum G\left(\frac{m}{M}\right) G\left(\frac{nf}{N}\right) e\left(\frac{n\overline{m\ell_1\ell_3 v}}{w}\right).$$

We next reduce the size of f by a similar argument to the one that let us impose the condition $V \le Q^\epsilon$. We obtain by transitioning to multiplicative characters (recall (2.4.5)) that the sum over $v, w, m, n, \ell_1, \ell_3$ is bounded by

$$\ll \frac{W^{\frac{1}{2}+o(1)} V^{\frac{1}{2}}}{f^{\frac{1}{2}}} \sum_{w \asymp W/f} \frac{1}{w^{\frac{1}{2}}} \left(\frac{(MN)^{\frac{1}{2}}}{f^{\frac{1}{2}}} + w^{\frac{1}{2}}\right) \left((L_1L_3)^{\frac{1}{2}} + w^{\frac{1}{2}}\right) \ll \frac{Q^{2+\epsilon}}{f^{\frac{3}{2}}},$$

and therefore the contribution from $f > Q^{4\epsilon}$ is negligible.

Now the only barrier to applying Lemma 2.4.1 is the conditions $(m, f) = 1$ and $(m, v) = 1$. We remove both of these conditions with Möbius inversion, obtaining

$$\sum_{f \ll \min(N, Q^\epsilon)} \mu(f) \sum_{h|f} \mu(h) \sum_{t \ll V} \mu(t) \frac{W^{\frac{1}{2}}}{(MNL_1L_3V)^{\frac{1}{2}}} \sum_{\substack{(v,wf)=1 \\ (w,ht)=1}}\sum \alpha(v) G\left(\frac{vt}{V}\right) G\left(\frac{wf}{W}\right)$$
$$\times \sum_{\substack{\ell_i \le y_i \\ (\ell_i, fvw)=1}}\sum \beta(\ell_1)\gamma(\ell_3) G\left(\frac{\ell_1}{L_1}\right) G\left(\frac{\ell_3}{L_3}\right) \sum_{\substack{(m,w)=1 \\ (n,v)=1}}\sum G\left(\frac{mht}{M}\right) G\left(\frac{nf}{N}\right) e\left(\frac{n\overline{mht^2\ell_1\ell_3 v}}{w}\right).$$

We set

$$b_{n,ht^2k} = \mathbf{1}_{(n,v)=1} G\left(\frac{nf}{N}\right) \sum_{\substack{\ell_1 \quad \ell_3 \quad v \\ \ell_1\ell_3 v = k \\ (\ell_1\ell_3, v)=1}}\sum\sum \beta(\ell_1)\gamma(\ell_3)\alpha(v) G\left(\frac{vt}{V}\right) G\left(\frac{\ell_1}{L_1}\right) G\left(\frac{\ell_3}{L_3}\right)$$

if $(k, f) = 1$, and for integers r not divisible by ht^2 we set $b_{n,r} = 0$. It follows that if $b_{n,r} \neq 0$, then $n \asymp N/f$ and $r \asymp htL_1L_3V$ with $r \equiv 0(ht^2)$. The sum over n, r, m, w is therefore a sum of the form to which Lemma 2.4.1 may be applied. We note that

$$\|b_{N,R}\|_2 \ll \frac{Q^{o(1)}}{(ft)^{\frac{1}{2}}}(NL_1L_3V)^{\frac{1}{2}},$$

and therefore by Lemma 2.4.1 we have

$$\begin{aligned}
\mathcal{E}_1' &\ll Q^\epsilon \sum_{f \ll Q^\epsilon} \frac{1}{f^{\frac{1}{2}}} \sum_{h|f} \sum_{t \ll Q^\epsilon} \frac{1}{t^{\frac{1}{2}}} \frac{W^{\frac{1}{2}}}{M^{\frac{1}{2}}} \\
&\times \left\{ \frac{W^{\frac{1}{2}}}{f^{\frac{1}{2}}} \left((htL_1L_3V)^{\frac{1}{2}} + \frac{N^{\frac{1}{2}}}{f^{\frac{1}{2}}} \right) \left(\frac{W^{\frac{1}{2}}}{f^{\frac{1}{2}}} + (ML_1L_3V)^{\frac{1}{2}} \right) \right. \\
&\left. + \frac{W}{f} \frac{M^{\frac{1}{2}}}{(ht)^{\frac{1}{2}}} \left((htL_1L_3V)^{\frac{1}{2}} + (htL_1L_3NV)^{\frac{1}{4}} \right) + \frac{M}{ht}(htL_1L_3NV)^{\frac{1}{2}} \right\} \\
&\ll Q^\epsilon \left(\frac{W^{\frac{3}{2}}(y_1y_3)^{\frac{1}{2}}}{M^{\frac{1}{2}}} + Wy_1y_3 + W^{\frac{3}{2}} \frac{N^{\frac{1}{2}}}{M^{\frac{1}{2}}} + W(y_1y_3)^{\frac{1}{2}} N^{\frac{1}{2}} + W^{\frac{3}{2}}(y_1y_3)^{\frac{1}{2}} \right. \\
&\left. + W^{\frac{3}{2}}(y_1y_3)^{\frac{1}{4}} N^{\frac{1}{4}} + W^{\frac{1}{2}} Q^{\frac{1}{2}} (y_1y_3)^{\frac{1}{2}} \right) \ll Q^{2-\epsilon},
\end{aligned}$$

upon recalling the bounds $W \ll Q$, $y_i \leq Q^{\theta_i}$ with $\theta_i < \frac{1}{2}$, and $N \leq Q^{1-\epsilon}$. This completes the proof of Lemma 2.2.3.

2.5 Lemma 2.2.4: main term

In this section we obtain the main term of Lemma 2.2.4. We allow ourselves to recycle some notation from Sections 2.3 and 2.4.

Recall that we wish to asymptotically evaluate

$$\sum_{\chi(q)}^{+} \left| L\left(\frac{1}{2}, \chi\right) \right|^2 \psi_{\mathrm{B}}(\chi)\psi_{\mathrm{MV}}(\overline{\chi}).$$

We begin precisely as in Section 2.3. Inserting the definitions of $\psi_{\mathrm{B}}(\chi)$ and $\psi_{\mathrm{MV}}(\overline{\chi})$, we must asympotically evaluate

$$\frac{2}{L}\sum\sum_{\substack{bc\leq y_2\\(bc,q)=1}}\frac{\Lambda(b)\mu(c)P_2[bc]}{(bc)^{\frac12}}\sum_{\substack{\ell\leq y_3\\(\ell,q)=1}}\frac{\mu(\ell)P_3[\ell]}{\ell^{\frac12}}$$
$$\times\sum\sum_{(mn,q)=1}\frac{1}{(mn)^{\frac12}}V\left(\frac{mn}{q}\right)\sum_{\chi(q)}{}^{+}\epsilon(\chi)\chi(c\ell m)\overline{\chi}(bn). \tag{2.5.1}$$

The main term of Lemma 2.2.3 arose from $m\ell_1\ell_3 = 1$. In the present case, the main term contains more than just $c\ell m = 1$; the main term arises from those $c\ell m$ which divide b. The support of the von Mangoldt function constrains b to be a prime power, so the condition $c\ell m \mid b$ is straightforward, but tedious, to handle.

There are three different cases to consider. The first case is $c\ell m = 1$. In the second case we have $c\ell m = p$ and $b = p$. Both of these cases contribute to the main term. The third case is everything else ($b = p^j$ with $j \geq 2$ and $c\ell m \mid b$ with $c\ell m \geq p$), and this case contributes only to the error term.

First case: $c\ell m = 1$

If $c\ell m$ is equal to 1 then certainly $c\ell m$ divides b for every b. The contribution from $c\ell m = 1$ is equal to

$$M = \frac{2P_3(1)}{L}\sum_{\chi(q)}{}^{+}\epsilon(\chi)\sum_{b\leq y_2}\frac{\Lambda(b)\overline{\chi}(b)P_2[b]}{b^{\frac12}}\sum_{n}\frac{\overline{\chi}(n)}{n^{\frac12}}V\left(\frac{n}{q}\right).$$

By an application of Proposition 2.3.3,

$$M = M_1 + M_2, \tag{2.5.2}$$

where

$$M_1 = \frac{2P_3(1)}{L} \sum_{\substack{b \le y_2 \\ (b,q)=1}} \frac{\Lambda(b)P_2[b]}{b^{\frac{1}{2}}} \sideset{}{^+}\sum_{\chi(q)} \epsilon(\chi)\overline{\chi}(b) L\left(\frac{1}{2}, \overline{\chi}\right),$$

$$M_2 = -\frac{2P_3(1)}{L} \sum_{\substack{b \le y_2 \\ (bn,q)=1}}\sum \frac{\Lambda(b)P_2[b]}{(bn)^{\frac{1}{2}}} F(n) \sideset{}{^+}\sum_{\chi(q)} \chi(n)\overline{\chi}(b),$$

and F is the rapidly decaying function given by (2.3.4). A main term arises from M_1, and M_2 contributes only to the error term.

Let us first investigate M_2. By Lemma 2.3.2 we have

$$M_2 = -\frac{P_3(1)}{L} \sum_{w|q} \varphi(w)\mu(q/w) \sum_{\substack{b \le y_2 \\ b \equiv \pm n(w) \\ (bn,q)=1}}\sum \frac{\Lambda(b)P_2[b]}{(bn)^{\frac{1}{2}}} F(n).$$

By the rapid decay of F (recall (2.3.6)) we may restrict n to $n \le q^{\frac{1}{10}}$. The contribution from $w \le q^{\frac{1}{2}+\epsilon}$ is then trivially $\ll q^{1-\epsilon}$, since $y_2 \ll q^{\frac{1}{2}-\epsilon}$. For the remaining terms, the congruence condition $b \equiv \pm n(w)$ becomes $b = n$, and thus

$$M_2 \ll q^{1-\epsilon} + \frac{1}{L} \sum_{\substack{w|q \\ w > q^{\frac{1}{2}+\epsilon}}} \varphi(w) \sum_{b \le q^{\frac{1}{10}}} \frac{\Lambda(b)P_2[b]}{b} F(b) \ll qL^{-1}.$$

Let us turn to M_1. We use the following lemma to represent the central value $L\left(\frac{1}{2}, \overline{\chi}\right)$.

Lemma 2.5.1. *Let $\overline{\chi}$ be a primitive even character modulo q. Then*

$$L\left(\frac{1}{2}, \overline{\chi}\right) = \sum_n \frac{\overline{\chi}(n)}{n^{\frac{1}{2}}} V_1\left(\frac{n}{q^{\frac{1}{2}}}\right) + \epsilon(\overline{\chi}) \sum_n \frac{\chi(n)}{n^{\frac{1}{2}}} V_1\left(\frac{n}{q^{\frac{1}{2}}}\right),$$

where

$$V_1(x) = \frac{1}{2\pi i} \int_{(1)} \frac{\Gamma\left(\frac{s}{2}+\frac{1}{4}\right)}{\Gamma\left(\frac{1}{4}\right)} \frac{G_1(s)}{s} \pi^{-s/2} x^{-s} ds$$

and $G_1(s)$ is an even polynomial satisfying $G_1(0) = 1$.

Proof. See [30, (2.2)]. □

Applying Lemma 2.5.1, the main term M_1 naturally splits as $M_1 = M_{1,1} + M_{1,2}$, where

$$M_{1,1} = \frac{2P_3(1)}{L} \sum_{\substack{b \le y_2 \\ (bn,q)=1}}\sum \frac{\Lambda(b)P_2[b]}{(bn)^{\frac{1}{2}}} V_1\left(\frac{n}{q^{\frac{1}{2}}}\right) \sum_{\chi(q)}{}^{+} \epsilon(\chi)\overline{\chi}(bn),$$
$$M_{1,2} = \frac{2P_3(1)}{L} \sum_{\substack{b \le y_2 \\ (bn,q)=1}}\sum \frac{\Lambda(b)P_2[b]}{(bn)^{\frac{1}{2}}} V_1\left(\frac{n}{q^{\frac{1}{2}}}\right) \sum_{\chi(q)}{}^{+} \chi(n)\overline{\chi}(b).$$

Applying character orthogonality to $M_{1,1}$ we arrive at

$$M_{1,1} = \frac{2P_3(1)}{Lq^{1/2}} \sum_{\substack{vw=q \\ (v,w)=1}}\sum \mu^2(v)\varphi(w) \sum_{\substack{b \le y_2 \\ (bn,q)=1}}\sum \frac{\Lambda(b)P_2[b]}{(bn)^{\frac{1}{2}}} V_1\left(\frac{n}{q^{\frac{1}{2}}}\right) \cos\left(\frac{2\pi bn\overline{v}}{w}\right),$$

and a trivial estimation shows

$$M_{1,1} \ll q^{1-\epsilon}.$$

Let us lastly examine $M_{1,2}$, from which a main term arises. By character orthogonality we have

$$M_{1,2} = \frac{P_3(1)}{L} \sum_{w|q} \varphi(w)\mu(q/w) \sum_{\substack{b \le y_2 \\ b \equiv \pm n(w) \\ (b,q)=1}}\sum \frac{\Lambda(b)P_2[b]}{(bn)^{\frac{1}{2}}} V_1\left(\frac{n}{q^{\frac{1}{2}}}\right).$$

By trivial estimation, the contribution from $w \le q^{\frac{1}{2}+\epsilon}$ is

$$\ll \sum_{\substack{w|q \\ w \le q^{\frac{1}{2}+\epsilon}}} \varphi(w) \sum_{b \le y_2} \frac{1}{b^{1/2}} \sum_{\substack{n \le q^{\frac{1}{2}+\epsilon} \\ n \equiv \pm b(w)}} \frac{1}{n^{\frac{1}{2}}} \ll y_2^{\frac{1}{2}} \sum_{\substack{w|q \\ w \le q^{\frac{1}{2}+\epsilon}}} \varphi(w) \left(\frac{q^{\frac{1}{4}+\epsilon}}{w} + O(1)\right) \ll q^{\frac{3}{4}+\epsilon}.$$

By the rapid decay of V_1, for $w > q^{\frac{1}{2}+\epsilon}$ the congruence $b \equiv \pm n(w)$ becomes $b = n$. Adding back in the terms $w \le q^{\frac{1}{2}+\epsilon}$, we have

$$M_{1,2} = \frac{2P_3(1)}{L}\varphi^{+}(q) \sum_{\substack{b \le y_2 \\ (b,q)=1}} \frac{\Lambda(b)P_2[b]}{b} V_1\left(\frac{b}{q^{\frac{1}{2}}}\right) + O(q^{1-\epsilon}).$$

For $x \ll 1$ we see by a contour shift that

$$V_1(x) = 1 + O(x^{\frac{1}{3}}),$$

and we have $bq^{-1/2} \ll q^{-\epsilon}$. It follows that

$$M_{1,2} = O(q^{1-\epsilon}) + \frac{2P_3(1)}{L}\varphi^+(q) \sum_{\substack{b \leq y_2 \\ (b,q)=1}} \frac{\Lambda(b)P_2[b]}{b}.$$

We have

$$\sum_{(b,q)>1} \frac{\Lambda(b)}{b} \ll 1 + \sum_{p|q} \frac{\log p}{p} \ll \log\log q,$$

and therefore we may remove the condition $(b,q) = 1$ at the cost of an error $O\left(qL^{-1+\epsilon}\right)$. From the estimate

$$\sum_{n \leq x} \frac{\Lambda(n)}{n} = \log x + O(1),$$

summation by parts, and elementary manipulations, we obtain

$$\sum_{b \leq y_2} \frac{\Lambda(b)P_2[b]}{b} = (\log y_2) \int_0^1 P_2(u)du + O(1).$$

Therefore, the contribution to the main term of Lemma 2.2.4 from $c\ell m = 1$ is

$$(2\theta_2 P_3(1)\widetilde{P_2}(1) + o(1))\varphi^+(q). \tag{2.5.3}$$

Second case: $c\ell m = p$, $b = p$

Another main term which contributes to Lemma 2.2.4 comes from $c\ell m = p$ and $b = p$. There are three subcases: $(c, \ell, m) = (p, 1, 1), (1, p, 1)$, or $(1, 1, p)$. These three cases give (compare with (2.5.1))

$$N_1 = -\frac{2P_3(1)}{L} \sum_{\substack{p \le y_2^{1/2} \\ (p,q)=1}} \frac{(\log p) P_2\left(\frac{\log(y_2^{1/2}/p)}{\log(y_2^{1/2})}\right)}{p} \sum_{\chi(q)}^{+} \epsilon(\chi) \sum_n \frac{\overline{\chi}(n)}{n^{\frac{1}{2}}} V\left(\frac{n}{q}\right),$$

$$N_2 = -\frac{2}{L} \sum_{\substack{p \le y_2 \\ (p,q)=1}} \frac{(\log p) P_2[p] P_3[p]}{p} \sum_{\chi(q)}^{+} \epsilon(\chi) \sum_n \frac{\overline{\chi}(n)}{n^{\frac{1}{2}}} V\left(\frac{n}{q}\right),$$

$$N_3 = \frac{2P_3(1)}{L} \sum_{\substack{p \le y_2 \\ (p,q)=1}} \frac{(\log p) P_2[p]}{p} \sum_{\chi(q)}^{+} \epsilon(\chi) \sum_n \frac{\overline{\chi}(n)}{n^{\frac{1}{2}}} V\left(\frac{pn}{q}\right).$$

The first two are somewhat easier to handle than the last one. We apply Proposition 2.3.3 then argue as in Section 2.3 and the $c\ell m = 1$ case to obtain

$$\sum_{\chi(q)}^{+} \epsilon(\chi) \sum_n \frac{\overline{\chi}(n)}{n^{1/2}} V\left(\frac{n}{q}\right) = \frac{1}{2}\varphi^+(q) + O(q^{1-\epsilon}).$$

It follows that

$$N_1 = -\left(\frac{\theta_2}{2} P_3(1)\widetilde{P_2}(1) + o(1)\right)\varphi^+(q), \tag{2.5.4}$$

$$N_2 = -\left(\theta_2 \int_0^1 P_2(u)P_3(u)du + o(1)\right)\varphi^+(q).$$

Combining (2.5.3) and (2.5.4) gives the main term of Lemma 2.2.4.

The final term N_3 is more difficult because the inner sum now depends on p. However, M_3 contributes only to the error term. By Proposition 2.3.3 with $T = p$,

$$\sum_n \frac{\overline{\chi}(n)}{n^{\frac{1}{2}}} V\left(\frac{pn}{q}\right) = L\left(\frac{1}{2}, \overline{\chi}\right) - \sum_n \frac{\chi(n)}{n^{\frac{1}{2}}} F\left(\frac{n}{p}\right). \tag{2.5.5}$$

The first term on the right side of (2.5.5) contributes to N_3 an amount

$$\left(2\theta_2 P_3(1)\widetilde{P_2}(1) + o(1)\right)\varphi^+(q). \tag{2.5.6}$$

For the second term on the right side of (2.5.5) we use character orthogonality and get

$$-\frac{2P_3(1)}{L}\sum_{\substack{p\leq y_2\\(p,q)=1}}\frac{(\log p)P_2[p]}{p}\frac{1}{2}\sum_{w|q}\varphi(w)\mu(q/w)\sum_{n\equiv\pm1(w)}\frac{1}{n^{\frac{1}{2}}}F\left(\frac{n}{p}\right).$$

By the rapid decay of F the contribution from $n > p^{\frac{11}{10}}$, say, is $O(qL^{-1})$. We next estimate trivially the contribution from $w \leq q^{\frac{3}{5}}$, say. We have the bound

$$\sum_{\substack{n\equiv\pm1(w)\\n\leq p^{\frac{11}{10}}}}\frac{1}{n^{\frac{1}{2}}}F\left(\frac{n}{p}\right)\ll q^{\epsilon}\left(\frac{p^{\frac{11}{20}}}{w}+1\right),$$

and this contributes to N_3 an amount

$$\ll q^{\frac{3}{5}+\epsilon}+q^{\epsilon}\sum_{p\leq y_2}p^{-\frac{9}{20}}\ll q^{\frac{3}{5}+\epsilon},$$

since $y_2 \ll q^{\frac{1}{2}}$. For $w > q^{\frac{3}{5}}$ and $n \leq p^{\frac{11}{10}}$ the congruence $n \equiv \pm1(w)$ becomes $n = 1$. By a contour shift we have

$$F\left(\frac{1}{p}\right)=1+O\left(p^{-\frac{1}{2}}\right).$$

Thus, the second term on the right side of (2.5.5) contributes to N_3 an amount

$$-\left(2\theta_2P_3(1)\widetilde{P_2}(1)+o(1)\right)\varphi^{+}(q), \tag{2.5.7}$$

and (2.5.6) and (2.5.7) together imply N_3 is negligible.

Third case: everything else

This case is the contribution from $b = p^j$ with $j \geq 2$ and $c\ell m \mid b$ with $c\ell m \geq p$. This case contributes an error of size $O(qL^{-1+\epsilon})$, essentially because the sum

$$\sum_{\substack{p^k \\ k \geq 2}} \frac{\log(p^k)}{p^k}$$

converges. There are four different subcases to consider, since the Möbius functions attached to c and ℓ imply $c, \ell \in \{1, p\}$. The same techniques we have already employed allow one to bound the resulting sums, so we leave the details for the interested reader. This completes the proof of Lemma 2.2.4.

2.6 Lemma 2.2.4: error term

After the results of the previous section, it remains to finish the proof of Lemma 2.2.4 by showing the error term of (2.5.1) is negligible. The argument is very similar to that given in Section 2.4, and, indeed, the arguments are identical after a point.

The error term has the form

$$\mathcal{E}_2 = \sum\sum_{(v,w)=1} \mu^2(v) \frac{v}{\varphi(v)} \frac{w^{\frac{1}{2}}}{v^{\frac{1}{2}}} \Psi\left(\frac{vw}{Q}\right) \sum_{\substack{\ell \leq y_3 \\ (\ell, vw)=1}} \frac{\mu(\ell) P_3[\ell]}{\ell^{\frac{1}{2}}}$$
$$\times \sum\sum_{\substack{bc \leq y_2 \\ (bc,vw)=1}} \frac{\Lambda(b)\mu(c)P_2[bc]}{(bc)^{\frac{1}{2}}} \sum\sum_{(mn,vw)=1} \frac{1}{(mn)^{\frac{1}{2}}} V\left(\frac{mn}{vw}\right) \cos\left(\frac{2\pi bn\overline{c\ell m v}}{w}\right),$$

where we also have the condition $c\ell m \nmid b$, which we do not indicate in the notation. This condition is awkward, but turns out to be harmless.

We note that we may separate the variables b and c from one another in $P_2[bc]$ by linearity, the additivity of the logarithm, and the binomial theorem. Thus, it suffices to study $\mathcal{E}_1$ with $P_2[bc]$ replaced by $(\log b)^{j_1}(\log c)^{j_2}$, for j_i some fixed nonnegative integers. Arguing as in the reduction to (2.4.3), we may bound $\mathcal{E}_2$ by $\ll Q^{o(1)}$ instances of

$\mathcal{E}_2' = \mathcal{E}_2'(B,C,L,M,N,V,W)$, where

$$\begin{aligned}\mathcal{E}_2' = \frac{W^{\frac{1}{2}}}{(BCLMNV)^{\frac{1}{2}}} &\sum\sum_{(v,w)=1} \alpha(v) G\left(\frac{v}{V}\right) G\left(\frac{w}{W}\right) \\ \times &\sum_{\substack{\ell \leq y_3 \\ (\ell, vw)=1}} \beta(\ell) G\left(\frac{\ell}{L}\right) \sum\sum_{\substack{bc \leq y_2 \\ (bc,vw)=1}} \gamma(b)\delta(c) G\left(\frac{b}{B}\right) G\left(\frac{c}{C}\right) \\ \times &\sum\sum_{(mn,vw)=1} G\left(\frac{m}{M}\right) G\left(\frac{n}{N}\right) e\left(\frac{bn\overline{c\ell m v}}{w}\right),\end{aligned} \tag{2.6.1}$$

the function G is smooth as before, and $\alpha, \beta, \gamma, \delta$ are sequences f satisfying $|f(z)| \ll Q^{o(1)}$. We also have the conditions

$$VW \asymp Q, \quad MN \leq Q^{1+\epsilon}, \quad BC \ll y_2, \quad L \ll y_3, \quad B,C,L,M,N,V,W \gg 1.$$

By the argument that gave (2.4.6) we may also assume $V \leq Q^\epsilon$. Lastly, we may remove the condition $bc \leq y_2$ by Mellin inversion, at the cost of changing γ and δ by b^{it_0}, c^{it_0}, respectively, where $t_0 \in \mathbb{R}$ is arbitrary (see [64, Lemma 9], for instance).

Recall the condition $c\ell m \nmid b$. This condition is unnecessary if $CLM > 2019B$, say, so it is only in the case $CLM \ll B$ where we need to deal with it. However, the case $CLM \ll B$ is exceptional, since B is bounded by $y_2 \ll Q^{\frac{1}{2}}$ but generically we would expect CLM to be much larger than $Q^{\frac{1}{2}}$.

Indeed, we now show that when $CLM \ll B$ it suffices to get cancellation from the n variable alone. The proof is essentially Proposition 2.4.2, so we just remark upon the differences. By Möbius inversion and Poisson summation we have

$$\begin{aligned}&\sum_{(n,vw)=1} G\left(\frac{n}{N}\right) e\left(\frac{bn\overline{c\ell m v}}{w}\right) = \mu(w) \frac{N}{w}\frac{\varphi(v)}{v} \\ &+ \sum_{d|v} \mu(d) \frac{N}{dw} \sum_{|h| \leq W^{1+\epsilon} d/N} \widehat{G}\left(\frac{hN}{dw}\right) \sum_{(a,w)=1} e\left(\frac{abd\overline{c\ell m v}}{w} + \frac{ah}{w}\right) \\ &+ O(Q^{-100}).\end{aligned}$$

The first and third terms contribute acceptable amounts, so consider the second term. The sum over a is the Ramanujan sum $c_w(hc\ell mv + bd)$, and since $c\ell m$ does not divide b the argument of the Ramanujan sum is non-zero. Following the proof of Proposition

2.4.2, we therefore obtain a bound of

$$\mathcal{E}_2' \ll \frac{Q^{\frac{3}{2}+\epsilon}(BCLM)^{\frac{1}{2}}}{N^{\frac{1}{2}}}. \tag{2.6.2}$$

By the reasoning immediately after Proposition 2.4.2, the bound (2.6.2) allows us to assume $N \leq MQ^{-2\epsilon}$, so that $N \leq Q^{\frac{1}{2}}$, regardless of whether $CLM \ll B$. In the case $CLM \ll B$, the bound (2.6.2) becomes

$$\mathcal{E}_2' \ll \frac{Q^{\frac{3}{2}+\epsilon}B}{N^{\frac{1}{2}}} \ll Q^{\frac{3}{2}+\epsilon}B \ll Q^{\frac{3}{2}+\theta_2+\epsilon} \ll Q^{2-\epsilon},$$

which of course is acceptable.

At this point we can follow the rest of the proof in Section 2.4. We change variables $bn \to n$, and the rest follows *mutatis mutandis* (it is important that with $N \ll Q^{\frac{1}{2}}$ we have $BN \ll Q^{1-\epsilon}$). This completes the proof of Lemma 2.2.4.

CHAPTER 3

DIRICHLET L-FUNCTIONS OF QUADRATIC CHARACTERS OF PRIME CONDUCTOR AT THE CENTRAL POINT

In this chapter[1] we prove Theorems 0.2.2, 0.2.3, 0.2.4, and 0.2.5. The outline of this chapter is as follows. In Section 3.1 we establish some notation and conventions that hold throughout this chapter. Section 3.2 outlines the basic strategy for the proof of Theorem 0.2.2. In Sections 3.3 and 3.4 we state a number of important technical results which are used in the proofs of our theorems. The proof of Theorem 0.2.2 is spread across Sections 3.5, 3.6, and 3.7. In Section 3.5 and its subsections we study the mollified first moment problem. The very long Section 3.6 and its subsections handle the mollified second moment. We choose our mollifier and finish the proof of Theorem 0.2.2 in Section 3.7. We prove Theorems 0.2.3 and 0.2.4 in Section 3.8, and we prove Theorem 0.2.5 in Section 3.9.

3.1 Notation and conventions

We define $\chi_n(\cdot) = \left(\frac{n}{\cdot}\right)$, the Kronecker symbol, for all nonzero integers n, even if n is not a fundamental discriminant. Note that this means χ_n has conductor $|n|$ only when n is a fundamental discriminant. We write $S(Q)$ for the set of all real primitive characters χ with conductor $\leq Q$. For an integer n, we write $n = \square$ or $n \neq \square$ according to whether or not n is a perfect square.

We let $\varepsilon > 0$ denote an arbitrarily small constant whose value may vary from one line to the next. When ε is present, in some fashion, in an inequality or error term, we allow implied constants to depend on ε without necessarily indicating this in the notation. At times we indicate the dependence of implied constants on other quantities by use of subscripts: for example, $Y \ll_A Z$.

Throughout this chapter, we denote by $\Phi(x)$ a smooth function, compactly supported

[1]All work in this chapter is joint with Siegfred Baluyot. This work has been submitted for publication.

in $[\frac{1}{2}, 1]$, which satisfies $\Phi(x) = 1$ for $x \in [\frac{1}{2} + \frac{1}{\log X}, 1 - \frac{1}{\log X}]$ and $\Phi^{(j)}(x) \ll_j (\log X)^j$ for all $j \geq 0$. We could state our results for arbitrary smooth functions supported in $[\frac{1}{2}, 1]$, but we avoid this in an attempt to achieve some simplicity.

We write $e(x) = e^{2\pi ix}$. For g a compactly supported smooth function, we define the Fourier transform $\hat{g}(y)$ of g by

$$\hat{g}(y) = \int_{\mathbb{R}} g(x)e(-xy)dx.$$

At times, however, we find it convenient to use a slightly different normalization of the Fourier transform (see Lemma 3.4.2).

We define the Mellin transform $g^\dagger(s)$ of g by

$$g^\dagger(s) = \int_0^\infty g(x)x^{s-1}dx.$$

It is also helpful to define a modified Mellin transform $\check{g}(w)$ by

$$\check{g}(w) = \int_0^\infty g(x)x^w dx.$$

Observe that $\check{g}(w) = g^\dagger(1 + w)$. Lastly, for a complex number s, we define

$$g_s(t) = g(t)t^{s/2}.$$

Note that

$$\hat{\Phi}(0) = \Phi^\dagger(1) = \check{\Phi}(0) = \frac{1}{2} + O\left(\frac{1}{\log X}\right).$$

The letter p always denotes a prime number. We write φ for the Euler phi function, and d_k for the k-fold divisor function. If a and b are integers we write $[a, b]$ for their least common multiple and (a, b) for their greatest common divisor. It will always be clear from context whether $[a, b]$, say, denotes a least common multiple or a real interval.

Given coprime integers a and q, we write $\overline{a} \pmod q$ for the multiplicative inverse of a modulo q.

3.2 Outline of the proof of Theorem 0.2.2

The proof of Theorem 0.2.2 proceeds through the mollification method. The method was introduced by Bohr and Landau [27], but later greatly refined in the hands of Selberg [29]. The idea is to introduce a Dirichlet polynomial $M(p)$, known as a mollifier, which dampens the occasional wild behavior of the central values $L(\frac{1}{2},\chi_p)$. We study the first and second moments

$$\begin{aligned} S_1 &:= \sum_{p\equiv 1 \pmod 8} (\log p)\Phi\left(\frac{p}{X}\right) L\left(\tfrac{1}{2},\chi_p\right) M(p), \\ S_2 &:= \sum_{p\equiv 1 \pmod 8} (\log p)\Phi\left(\frac{p}{X}\right) L\left(\tfrac{1}{2},\chi_p\right)^2 M(p)^2. \end{aligned} \tag{3.2.1}$$

If the mollifier is chosen well then $S_1 \gg X$ and $S_2 \ll X$. By the Cauchy-Schwarz inequality we have

$$\sum_{\substack{p\equiv 1 \pmod 8 \\ L(\frac{1}{2},\chi_p)\neq 0}} (\log p)\Phi\left(\frac{p}{X}\right) \geq \frac{S_1^2}{S_2}, \tag{3.2.2}$$

and this implies that a positive proportion of $L(\frac{1}{2},\chi_p)$ are non-zero.

Our mollifier takes the form

$$M(p) := \sum_{\substack{m\leq M \\ m \text{ odd}}} \frac{b_m}{\sqrt{m}}\chi_p(m), \tag{3.2.3}$$

for some coefficients b_m we describe shortly. Here we set

$$M = X^\theta, \qquad \theta \in \left(0,\tfrac{1}{2}\right) \quad \text{fixed.} \tag{3.2.4}$$

The larger one can take θ, the better proportion of nonvanishing one can achieve.

The coefficients b_m are a smoothed version of the Möbius function $\mu(m)$. Specifically, we choose

$$b_m = \mu(m)H\left(\frac{\log m}{\log M}\right), \tag{3.2.5}$$

where $H(t)$ is smooth function compactly supported in $[-1,1]$ which we choose in Section 3.7. It will be convenient in a number of places that b_m is supported on square-free integers.

We outline our strategy for estimating S_1 and S_2. We simplify the presentation here in comparison to the actual proofs. The sum S_1 is by far the simpler of the two, so we start here (see Section 3.5). Using an approximate functional equation for the central value $L(\frac{1}{2}, \chi_p)$ (Lemma 3.3.2), we write S_1 as

$$S_1 \approx \sum_{m \leq M} \frac{b_m}{\sqrt{m}} \sum_{k \leq X^{1/2+\varepsilon}} \frac{1}{\sqrt{k}} \sum_{p \equiv 1 \pmod 8} (\log p) \Phi\left(\frac{p}{X}\right) \chi_p(mk).$$

The main term arises from the "diagonal" terms $mk = \square$. The character values $\chi_p(mk)$ are then all equal to one, and we simply use the prime number theorem for arithmetic progressions modulo eight to handle the sum on p. The sum over k contributes a logarithmic factor, but this logarithmic loss is canceled out by a logarithmic gain coming from a cancellation in the mollifier coefficients. This yields the main term for S_1, which is of size $\asymp X$ (Proposition 3.5.1).

The "off-diagonal" terms $mk \neq \square$ contribute only to the error term. After some manipulations the off-diagonal terms are essentially of the form

$$\mathcal{E} := \sum_{\substack{q \leq MX^{1/2+\varepsilon} \\ q \neq \square}} \frac{\alpha(q)}{q^{\frac{1}{2}}} \sum_{p} (\log p) \Phi\left(\frac{p}{X}\right) \chi_q(p),$$

where $\alpha(q)$ is some function satisfying $|\alpha(q)| \ll_\varepsilon q^\varepsilon$. We assume here for simplicity that all of the characters χ_q are primitive characters. We bound the character sum over primes in $\mathcal{E}$ in three different ways, depending on the size of q. These three regimes correspond to small, medium, and large values of q. Some of the arguments are similar to those of Jutila [48].

In the regime of small q we appeal to the prime number theorem for arithmetic progressions with error term. The sum over primes p is small, except in the case where one of the characters χ_{q^*} is exceptional: that is, the associated L-function $L(s, \chi_{q^*})$ has a real zero β_* very close to $s = 1$. Siegel's theorem gives $q^* \geq c(B)(\log X)^B$ with $B > 0$ arbitrarily large. This would immediately dispatch any exceptional characters, but unfortunately the constant $c(B)$ is not effectively computable. To get an effective estimate we use Page's theorem, which states that at most one such exceptional character χ_{q^*} exists. We then study carefully the contribution of this one exceptional character and show it is acceptably small.

In regimes of medium and large q, we take advantage of the averaging over q present in $\mathcal{E}$. We bound $\mathcal{E}$ in terms of instances of

$$\mathcal{E}(Q) := Q^{-\frac{1}{2}+\varepsilon} \sum_{\substack{Q/2<q\leq Q \\ q\neq\square}} \left| \sum_p (\log p)\Phi\left(\frac{p}{X}\right)\chi_q(p) \right|,$$

where Q is of moderate size, or is large.

When Q is medium-sized, we use the explicit formula to bound $\mathcal{E}(Q)$ by sums over zeros of the L-functions $L(s,\chi_q)$. We then use zero-density estimates.

We are left with the task of bounding $\mathcal{E}(Q)$ when Q is large, which means Q is larger than X^δ for some small, fixed $\delta > 0$. Rather than treating the sum on primes analytically, as we did when Q was small or medium-sized, we treat the sum on primes combinatorially. We use Vaughan's identity to write the character sum over the primes as a linear combination of linear and bilinear sums. The linear sums are handled easily with the Pólya-Vinogradov inequality. We bound the bilinear sums by appealing to a large sieve inequality for real characters due to Heath-Brown (Lemma 3.3.4).

We now describe our plan of attack for S_2 (see Section 3.6). Recall that

$$S_2 = \sum_{p\equiv 1 \pmod 8} (\log p)\Phi\left(\frac{p}{X}\right) L\left(\tfrac{1}{2},\chi_p\right)^2 M(p)^2.$$

As we see from Theorem 0.2.4, we only barely obtain an asymptotic formula for the second moment

$$\sum_{\substack{p\leq X \\ p\equiv 1 \pmod 8}} (\log p) L\left(\tfrac{1}{2},\chi_p\right)^2$$

under the assumption of the Generalized Riemann Hypothesis. Thus, it might seem doubtful that one can say anything useful about S_2, since the central value $L(\frac{1}{2},\chi_p)^2$ is further twisted by the square of a Dirichlet polynomial. The key idea is that we do not need an asymptotic formula for S_2, but only an upper bound of the right order of magnitude (with a good constant). We therefore avail ourselves of sieve methods (see Section 3.4). By positivity we have

$$S_2 \leq (\log X) \sum_{n\equiv 1 \pmod 8} \mu^2(n)\Phi\left(\frac{n}{X}\right)\left(\sum_{d|n}\lambda_d\right) L\left(\tfrac{1}{2},\chi_n\right)^2 M(n)^2,$$

where

$$\sum_{d|n} \lambda_d$$

is an upper bound sieve supported on coefficients with $d \leq D$. Since we are now working with ordinary integers instead of prime numbers, the analysis for S_2 becomes similar to the second moment problem considered in [51] (see [51, Section 5]).

We begin by writing

$$\mu^2(n) = N_Y(n) + R_Y(n), \tag{3.2.6}$$

where

$$N_Y(n) := \sum_{\substack{\ell^2 | n \\ \ell \leq Y}} \mu(\ell), \qquad R_Y(n) := \sum_{\substack{\ell^2 | n \\ \ell > Y}} \mu(\ell), \tag{3.2.7}$$

and Y is a small power of X. The sum

$$\sum_{n \equiv 1 \pmod 8} \Phi\left(\frac{n}{X}\right) R_Y(n) \left(\sum_{d|n} \lambda_d\right) L\left(\tfrac{1}{2}, \chi_n\right)^2 M(n)^2$$

is an error term, and is shown to be small in a straightforward fashion by applying moment estimates for $L(\frac{1}{2}, \chi_n)$ due to Heath-Brown (Lemma 3.3.5).

The main task is therefore to asymptotically evaluate the sum

$$\sum_{n \equiv 1 \pmod 8} \Phi\left(\frac{n}{X}\right) N_Y(n) \left(\sum_{d|n} \lambda_d\right) L\left(\tfrac{1}{2}, \chi_n\right)^2 M(n)^2.$$

We use an approximate functional equation to represent the central values $L\left(\frac{1}{2}, \chi_n\right)^2$ and arrive at expressions of the form

$$\sum_{\ell \leq Y} \mu(\ell) \sum_{d \leq D} \lambda_d \sum\sum_{m_1, m_2 \leq M} \frac{b_{m_1} b_{m_2}}{\sqrt{m_1 m_2}} \sum_{\nu=1}^{\infty} \frac{d(\nu)}{\sqrt{\nu}} \sum_{\substack{n \equiv 1 \pmod 8 \\ d|n \\ \ell^2 | n}} \left(\frac{m_1 m_2 \nu}{n}\right) \Phi\left(\frac{n}{X}\right) \omega\left(\frac{\nu}{n}\right),$$

where $\omega(x)$ is some rapidly decaying smooth function that satisfies $\omega(x) \approx 1$ for small x. We then make the change of variables $n = m[d, \ell^2]$.

We use Poisson summation to transform the sum over m into a sum basically of the form

$$\sum_{k\in\mathbb{Z}} \left(\frac{[d,\ell^2]k}{m_1m_2\nu}\right) e\left(\frac{k\overline{[d,\ell^2]m_1m_2\nu}}{8}\right) \hat{F}_\nu\left(\frac{kX}{[d,\ell^2]m_1m_2\nu}\right),$$

for some smooth function F_ν. The zero frequency $k=0$ gives rise to a main term. Since $(\frac{0}{h}) = 1$ or 0 depending on whether h is a square, the $k=0$ contribution represents the expected "diagonal" contribution from $m_1m_2\nu = \square$. There is an additional, off-diagonal, main term which arises, essentially, from the terms with $[d,\ell^2]k = \square$. We adapt here the delicate off-diagonal analysis of [51]. The situation is complicated by the presence of the additive character $e(\cdot)$, which is not present in [51]. The additive character necessitates a division of the integers k into residue classes modulo 8. We then use Fourier expansion to write the additive character as a linear combination of multiplicative characters. After many calculations the off-diagonal main term arises as a sum of complex line integrals. When we combine the various pieces the integrand becomes an even function, exhibiting a symmetry which none of the pieces separately possessed. This fact proves to be very convenient in the final steps of the main term analysis.

One intriguing feature of the main term in S_2 is a kind of "double mollification". We must account for the savings coming from the mollifier $M(n)$, but must also account for the savings coming from the sieve weights λ_d, which act as a sort of mollifier on the natural numbers. It is crucial that we get savings in both places, and therefore our sieve process must be very precise. We find that a variation on the ideas of Selberg (see e.g. [47, Section 6.5]) is sufficient.

At length we arrive at an upper bound $S_{2,U}$, say, for S_2 of size $S_{2,U} \ll X$. We make an optimal choice of the function $H(x)$ in Section 3.7 to maximize the ratio $S_1^2/S_{2,U}$. The resulting mollifier is not the optimal mollifier, but it gives results that are asymptotically equivalent to those attained with the optimal mollifier. This yields Theorem 0.2.2.

To treat other residue classes of $p \pmod 8$, we make the following changes. First, we change the definition of $\chi_p(\cdot)$ to $\left(\frac{(-1)^{\mathfrak{a}}p}{\cdot}\right)$, where $\mathfrak{a}=0$ if $p\equiv 1 \pmod 4$ and $\mathfrak{a}=1$ if $p\equiv 3 \pmod 4$. Thus χ_p is still a primitive character of conductor p. Second, we use a variant of the approximate functional equation (Lemma 3.3.2) with ω_j, defined in (3.3.1), replaced by

$$\frac{1}{2\pi i}\int_{(c)} \frac{\Gamma\left(\frac{s}{2}+\frac{1+2\mathfrak{a}}{4}\right)^j}{\Gamma\left(\frac{1+2\mathfrak{a}}{4}\right)^j}\left(1-\frac{\chi_p(2)}{2^{\frac{1}{2}-s}}\right)^j \xi^{-s} W(s)\frac{ds}{s}.$$

The function $W(s)$ here is $16\left(s^2-\frac{1}{4}\right)^2$. Its purpose is to cancel potential poles at $s=\frac{1}{2}$ in the analysis.

3.3 Lemmata

We represent the central values of L-functions by using an approximate functional equation. We first investigate some properties of the smooth functions which appear in our approximate functional equations. For $j=1,2$ and $c>0$, define

$$\omega_j(\xi)=\frac{1}{2\pi i}\int_{(c)} \frac{\Gamma\left(\frac{s}{2}+\frac{1}{4}\right)^j}{\Gamma\left(\frac{1}{4}\right)^j}\left(1-\frac{1}{2^{\frac{1}{2}-s}}\right)^j \xi^{-s}\frac{ds}{s}. \tag{3.3.1}$$

Lemma 3.3.1. *Let $j=1,2$. The function $\omega_j(\xi)$ is real-valued and smooth on $(0,\infty)$. If $\xi>0$ we have*

$$\omega_j(\xi)=\left(1-\frac{1}{\sqrt{2}}\right)^j+O_\varepsilon(\xi^{\frac{1}{2}-\varepsilon}).$$

For any fixed integer $\nu\geq 0$ and $\xi\geq 4\nu+10$, we have

$$\omega_j^{(\nu)}(\xi)\ll(\xi/2)^{\nu+3}\exp\left(-\frac{1}{4}\xi^{\frac{2}{j}}\right)\ll_\nu \exp\left(-\frac{1}{8}\xi^{\frac{2}{j}}\right).$$

Proof. The proof is similar to [51, Lemma 2.1], but we give details for completeness. The function $\omega_j(s)$ is real-valued because the change of variable $\mathrm{Im}(s)\to-\mathrm{Im}(s)$ shows that ω_j is equal to its complex conjugate. Moreover, uniform convergence for ξ in compact subintervals of $(0,\infty)$ shows that ω_j is smooth.

To prove the first estimate of the lemma, move the line of integration in the definition of $\omega_j(\xi)$ to $c=-\frac{1}{2}+\varepsilon$. The pole at $s=0$ contributes $\left(1-\frac{1}{\sqrt{2}}\right)^j$, and the new integral is $O_\varepsilon(\xi^{\frac{1}{2}-\varepsilon})$.

Let us turn to the last estimate of the lemma. We may suppose $\xi^{\frac{2}{j}}\geq 4\nu+10$. By differentiation under the integral sign we find

$$\omega_j^{(\nu)}(\xi) = \frac{(-1)^\nu}{2\pi i} \int_{(c)} \frac{\Gamma\left(\frac{s}{2}+\frac{1}{4}\right)^j}{\Gamma\left(\frac{1}{4}\right)^j} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)^j s(s+1)\cdots(s+\nu-1)\xi^{-s-\nu}\frac{ds}{s}.$$

Recall that $|\Gamma(x+iy)| \le \Gamma(x)$ for $x \ge 1$ and $z\Gamma(z) = \Gamma(z+1)$. Thus, for $c \ge 2$ we obtain

$$\begin{aligned}|\omega_j^{(\nu)}(\xi)| &\ll \Gamma\left(\frac{c}{2}+\frac{5}{4}+\nu\right)^j \left(1+\frac{2^c}{\sqrt{2}}\right)^j \xi^{-c-v} \int_{(c)} \frac{1}{|s||\frac{s}{2}+\frac{1}{4}+\nu|} \prod_{k=0}^{\nu-1} \frac{|s+k|}{|\frac{s}{2}+\frac{1}{4}+k|}|ds| \\ &\ll \Gamma\left(\frac{c}{2}+\frac{5}{4}+\nu\right)^j \left(\frac{2^j}{\xi}\right)^c \left(\frac{2}{\xi}\right)^\nu c^{-1},\end{aligned}$$

where the implied constants are absolute. By Stirling's formula this is

$$\ll \left(\frac{c+2\nu+3}{2e}\right)^{\frac{j}{2}(c+2\nu+3)} \left(\frac{2^j}{\xi}\right)^c \left(\frac{2}{\xi}\right)^\nu.$$

We choose $c = \frac{1}{2}\xi^{\frac{2}{j}} - 2\nu - 3$, which we note is > 2. Thus, the quantity in question is

$$\ll \left(\frac{\xi}{2}\right)^{\nu+3} \exp\left(-\frac{1}{4}\xi^{\frac{2}{j}}\right),$$

as desired. □

We will find it technically convenient to use an approximate functional equation in which the variable of summation is restricted to odd integers.

Lemma 3.3.2. *Let $n \equiv 1 \pmod 8$ be square-free and satisfy $n > 1$. Let $\chi_n(\cdot) = \left(\frac{n}{\cdot}\right)$ denote the real primitive character of conductor n. Then for $j = 1, 2$ we have*

$$L\left(\tfrac{1}{2}, \chi_n\right)^j = \frac{2}{\left(1-\frac{1}{\sqrt{2}}\right)^{2j}} \sum_{\substack{\nu=1 \\ \nu \text{ odd}}}^{\infty} \frac{\chi_n(\nu)d_j(\nu)}{\sqrt{\nu}} \omega_j\left(\nu\left(\frac{\pi}{n}\right)^{j/2}\right) =: \mathcal{D}_j(n).$$

Proof. The proof follows along standard lines (e.g. [47, Theorem 5.3]), but we give a proof since our situation is slightly different.

Let $\Lambda(z,\chi_n) = \left(\frac{n}{\pi}\right)^{z/2}\Gamma\left(\frac{z}{2}\right)L(z,\chi_n)$. Since $n \equiv 1 \pmod 4$ we have $\chi_n(-1) = 1$, and therefore we have the functional equation (see [65, Proposition 2.2.24], [66, Chapter 9])

$$\Lambda(z,\chi_n) = \Lambda(1-z,\chi_n).$$

Recall also that $\Lambda(z,\chi_n)$ is entire because χ_n is primitive.

Now consider the sum

$$I := \sum_{\nu \text{ odd}} \frac{\chi_n(\nu) d_j(\nu)}{\sqrt{\nu}} \omega_j \left(\nu \left(\frac{\pi}{n} \right)^{j/2} \right).$$

We use the definition of ω_j and interchange the order of summation and integration. Since $\chi_n(2) = 1$ we have

$$\begin{aligned} I &= \frac{1}{2\pi i} \int_{(c)} \frac{\Gamma\left(\frac{s}{2}+\frac{1}{4}\right)^j}{\Gamma\left(\frac{1}{4}\right)^j} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)^j \left(1 - \frac{1}{2^{\frac{1}{2}+s}}\right)^j \left(\frac{n}{\pi}\right)^{js/2} L\left(\frac{1}{2}+s,\chi_n\right)^j \frac{ds}{s} \\ &= \frac{1}{2\pi i} \int_{(c)} \frac{\left(\frac{n}{\pi}\right)^{-j/4}}{\Gamma\left(\frac{1}{4}\right)^j} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)^j \left(1 - \frac{1}{2^{\frac{1}{2}+s}}\right)^j \Lambda\left(\frac{1}{2}+s,\chi_n\right)^j \frac{ds}{s}. \end{aligned}$$

We move the line of integration to $\text{Re}(s) = -c$, picking up a contribution from the simple pole at $s = 0$:

$$\begin{aligned} I = &\frac{\left(\frac{n}{\pi}\right)^{-j/4}}{\Gamma\left(\frac{1}{4}\right)^j} \left(1 - \frac{1}{\sqrt{2}}\right)^{2j} \Lambda\left(\frac{1}{2},\chi_n\right)^j \\ &+ \frac{1}{2\pi i} \int_{(-c)} \frac{\left(\frac{n}{\pi}\right)^{-j/4}}{\Gamma\left(\frac{1}{4}\right)^j} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)^j \left(1 - \frac{1}{2^{\frac{1}{2}+s}}\right)^j \Lambda\left(\frac{1}{2}+s,\chi_n\right)^j \frac{ds}{s}. \end{aligned}$$

In this latter integral we change variables $s \to -s$ and then apply the functional equation $\Lambda\left(\frac{1}{2}-s,\chi_n\right) = \Lambda\left(\frac{1}{2}+s,\chi_n\right)$ to obtain

$$\frac{\left(\frac{n}{\pi}\right)^{-j/4}}{\Gamma\left(\frac{1}{4}\right)^j} \left(1 - \frac{1}{\sqrt{2}}\right)^{2j} \Lambda\left(\frac{1}{2},\chi_n\right)^j = 2I = 2 \sum_{\nu \text{ odd}} \frac{\chi_n(\nu) d_j(\nu)}{\sqrt{\nu}} \omega_j \left(\nu \left(\frac{\pi}{n} \right)^{j/2} \right).$$

We then rearrange to obtain the desired conclusion. □

We frequently encounter exponential sums which are analogous to Gauss sums.

Given an odd integer n, we define for all integers k

$$G_k(n) = \left(\frac{1-i}{2} + \left(\frac{-1}{n}\right)\frac{1+i}{2}\right) \sum_{a \pmod n} \left(\frac{a}{n}\right) e\left(\frac{ak}{n}\right) \tag{3.3.2}$$

and

$$\tau_k(n) = \sum_{a \pmod n} \left(\frac{a}{n}\right) e\left(\frac{ak}{n}\right) = \left(\frac{1+i}{2} + \left(\frac{-1}{n}\right)\frac{1-i}{2}\right) G_k(n). \tag{3.3.3}$$

We require knowledge of $G_k(n)$ for all n.

Lemma 3.3.3. *(i) (Multiplicativity) Suppose m and n are coprime odd integers. Then $G_k(mn) = G_k(m)G_k(n)$.*
(ii) Suppose p^α is the largest power of p dividing k. (If $k = 0$ set $\alpha = \infty$.) Then for $\beta \geq 1$

$$G_k(p^\beta) = \begin{cases} 0 & \text{if } \beta \leq \alpha \text{ is odd,} \\ \varphi(p^\beta) & \text{if } \beta \leq \alpha \text{ is even,} \\ -p^\alpha & \text{if } \beta = \alpha + 1 \text{ is even,} \\ (\frac{kp^{-\alpha}}{p})p^\alpha\sqrt{p} & \text{if } \beta = \alpha + 1 \text{ is odd,} \\ 0 & \text{if } \beta \geq \alpha + 2. \end{cases}$$

Proof. This is [51, Lemma 2.3]. □

The following two results are useful for bounding various character sums that arise. Both results are corollaries of a large sieve inequality for quadratic characters developed by Heath-Brown [67].

Lemma 3.3.4. *Let N and Q be positive integers, and let $a_1, \ldots, a_N$ be arbitrary complex numbers. Then*

$$\sum_{\chi \in S(Q)} \left| \sum_{n \leq N} a_n \chi(n) \right|^2 \ll_\varepsilon (QN)^\varepsilon (Q+N) \sum_{n_1 n_2 = \square} |a_{n_1} a_{n_2}|,$$

for any $\varepsilon > 0$. Let M be a positive integer, and for each $|m| \leq M$ write $4m = m_1 m_2^2$, where m_1 is a fundamental discriminant, and m_2 is positive. Suppose the sequence a_n

satisfies $|a_n| \ll n^{\varepsilon}$*. Then*

$$\sum_{|m|\le M} \frac{1}{m_2} \left| \sum_{n\le N} a_n \left(\frac{m}{n}\right) \right|^2 \ll (MN)^{\varepsilon} N(M+N).$$

Proof. This is [51, Lemma 2.4]. □

Lemma 3.3.5. *Suppose* $\sigma + it$ *is a complex number with* $\sigma \ge \frac{1}{2}$*. Then*

$$\sum_{\chi \in S(Q)} |L(\sigma + it, \chi)|^4 \ll Q^{1+\varepsilon}(1+|t|)^{1+\varepsilon}$$

and

$$\sum_{\chi \in S(Q)} |L(\sigma + it, \chi)|^2 \ll Q^{1+\varepsilon}(1+|t|)^{\frac{1}{2}+\varepsilon}.$$

Proof. This is [51, Lemma 2.5]. □

3.4 Sieve estimates

Our main sieve will be a variant of the Selberg sieve (see [57, Chapter 7]). To lessen the volume of calculations, we also use Brun's pure sieve [57, Chapter 6] as a preliminary sieve to handle small prime factors. We set

$$z_0 := \exp((\log X)^{1/3}) \tag{3.4.1}$$

and

$$R := X^{\vartheta}, \qquad \vartheta \in \left(0, \tfrac{1}{2}\right) \quad \text{fixed.} \tag{3.4.2}$$

Given a set $\mathcal{A}$ of integers we write $\mathbf{1}_{\mathcal{A}}(n)$ for the indicator function of this set. For $y > 2$ we define

$$P(y) = \prod_{p \le y} p.$$

Then, for $n \asymp X$, our basic sieve inequality is

$$\mathbf{1}_{\{n:n \text{ prime}\}} \leq \mathbf{1}_{\{n:(n,P(z_0))=1\}}\mathbf{1}_{\{n:(n,P(R)/P(z_0))=1\}}, \tag{3.4.3}$$

We write $\omega(n)$ for the number of distinct prime factors of n. To bound the first factor on the right-hand side of (3.4.3), we use Brun's upper bound sieve condition (see [57, (6.1)])

$$\mathbf{1}_{\{n:(n,P(z_0))=1\}}(n) \leq \sum_{\substack{b|(n,P(z_0)) \\ \omega(b)\leq 2r_0}} \mu(b), \tag{3.4.4}$$

where

$$r_0 := \lfloor (\log X)^{1/3} \rfloor.$$

We use an "analytic" Selberg sieve (e.g. [68]) for the second factor of (3.4.3). We introduce a smooth, non-negative function $G(t)$ which is supported on the interval $[-1,1]$. We further require $G(t)$ to satisfy $|G(t)| \ll 1, |G^{(j)}(t)| \ll_j (\log\log X)^{j-1}$ for j a positive integer, and on the interval $[0,1]$ we require $G(t) = 1-t$ for $t \leq 1-(\log\log X)^{-1}$. Then

$$\begin{aligned}\mathbf{1}_{\{n:(n,P(R)/P(z_0))=1\}}(n) &\leq \left(\sum_{\substack{d|n \\ (d,P(z_0))=1}} \mu(d) G\left(\frac{\log d}{\log R}\right)\right)^2 \\ &= \sum\sum_{\substack{j,k\leq R \\ [j,k]|n \\ (jk,P(z_0))=1}} \mu(j)\mu(k) G\left(\frac{\log j}{\log R}\right) G\left(\frac{\log k}{\log R}\right).\end{aligned} \tag{3.4.5}$$

We mention also that the properties of G imply

$$\int_0^\infty G'(t)^2 dt = 1 + O\left(\frac{1}{\log\log X}\right) = 1 + o(1). \tag{3.4.6}$$

Note that the fundamental theorem of calculus and Cauchy-Schwarz yield the lower bound

$$\int_0^\infty G'(t)^2 dt \geq 1.$$

From (3.4.3), (3.4.4), and (3.4.5), we arrive at the upper bound sieve condition

$$\mathbf{1}_{\{n:n \text{ prime}\}}(n) \le \sum_{d|n} \lambda_d, \tag{3.4.7}$$

where the coefficients λ_d are defined by

$$\lambda_d = \sum_{\substack{b|P(z_0)\\ \omega(b)\le 2r_0}} \sum\sum_{\substack{m,n\le R\\ b[m,n]=d\\ (mn,P(z_0))=1}} \mu(b)\mu(m)\mu(n) G\left(\frac{\log m}{\log R}\right) G\left(\frac{\log n}{\log R}\right). \tag{3.4.8}$$

If $b|P(z_0)$ and $\omega(b) \le 2r_0$, then $b \le z_0^{2r_0} = \exp(2(\log X)^{2/3})$. Hence $\lambda_d \ne 0$ only for $d \le D$, where

$$D = R^2 \exp(2(\log X)^{2/3}) \ll_\varepsilon R^2 X^\varepsilon. \tag{3.4.9}$$

In our evaluation of sums involving the sieve coefficients (3.4.8) we use the following version of the fundamental lemma of sieve theory (see also [57, Section 6.5]).

Lemma 3.4.1. *Let $0 < \delta < 1$ be a fixed constant, r a positive integer with $r \asymp (\log X)^\delta$, and z_0 as in* (3.4.1)*. Suppose that g is a multiplicative function such that $|g(p)| \ll 1$ uniformly for all primes p. Then*

$$\sum_{\substack{b|P(z_0)\\ \omega(b)\le r\\ (b,\ell)=1}} \frac{\mu(b)}{b} g(b) = \prod_{\substack{p\le z_0\\ p\nmid \ell}} \left(1 - \frac{g(p)}{p}\right) + O\Big(\exp(-r\log\log r)\Big)$$

uniformly for all positive integers ℓ.

Proof. The proof is standard. Complete the sum on the left-hand side by adding to it all the terms with $\omega(b) > r$, dropping by positivity the condition $(b,\ell) = 1$. The error introduced in doing so is $\ll \exp(-(1+o(1))r\log r) \ll \exp(-r\log\log r)$ (e.g. [47, §6.3]). The completed sum is equal to the Euler product on the right-hand side. □

The basic tool in our application of the Selberg sieve is the following lemma.

Lemma 3.4.2. *Let $z_0 = \exp((\log X)^{1/3})$. Let G be as above. Suppose h is a function such that $|h(p)| \ll_\varepsilon p^{-\varepsilon}$ uniformly for all primes p. Let $A > 0$ be a fixed real number. Then there exists a function $E_0(X)$, which depends only on X, G, and ϑ (see* (3.4.2)*)* *with* $E_0(X) \to 0$ *as* $X \to \infty$*, such that*

$$\sum\sum_{\substack{m,n\le R\\(mn,\ell P(z_0))=1}} \frac{\mu(m)\mu(n)}{[m,n]}\, G\left(\frac{\log m}{\log R}\right) G\left(\frac{\log n}{\log R}\right) \prod_{p|mn}\Big(1+h(p)\Big) \\ = \frac{1+E_0(X)}{\log R}\prod_{p\le z_0}\left(1-\frac{1}{p}\right)^{-1} + \; O_{\varepsilon,A}\left(\frac{1}{(\log R)^A}\right), \tag{3.4.10}$$

uniformly for $\ell \ll X^{O(1)}$.

Proof. Let $\mathcal{S}$ denote the left-hand side of (3.4.10). If $m,n\le R$ and $(mn,P(z_0))=1$, then $\omega(mn) \ll \log R$, and each prime dividing mn is larger than z_0. Thus

$$\prod_{p|mn}\Big(1+h(p)\Big) = 1 + O_\varepsilon\left(\frac{\log R}{z_0^\varepsilon}\right),$$

and so

$$\mathcal{S} = \sum\sum_{\substack{m,n\le R\\(mn,\ell P(z_0))=1}} \frac{\mu(m)\mu(n)}{[m,n]}\, G\left(\frac{\log m}{\log R}\right) G\left(\frac{\log n}{\log R}\right) + O\left(\frac{(\log R)^4}{z_0^\varepsilon}\right). \tag{3.4.11}$$

We may ignore the condition $(mn,\ell)=1$ in (3.4.11) because

$$\sum\sum_{\substack{m,n\le R\\(mn,P(z_0))=1\\(mn,\ell)>1}} \frac{1}{[m,n]} \le \sum\sum_{\substack{m,n\le R\\(mn,P(z_0))=1}} \frac{1}{[m,n]} \sum_{\substack{p|\ell\\p|mn}} 1 \ll (\log R)^3 \sum_{\substack{p|\ell\\p>z_0}} \frac{1}{p} \ll \frac{(\log \ell)(\log R)^3}{z_0}.$$

We next insert the Fourier inversion formula

$$G(t) \;=\; \int_{-\infty}^{\infty} g(z)e^{-t(1+iz)}\,dz \tag{3.4.12}$$

into (3.4.11), where

$$g(z) \;=\; \int_{-\infty}^{\infty} e^t G(t) e^{izt}\,dt. \tag{3.4.13}$$

We then interchange the order of summation and integration and write the sum as an

Euler product to deduce that

$$\mathcal{S} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(z_1)g(z_2)\prod_{p>z_0}\left(1-\frac{1}{p^{1+\frac{1+iz_1}{\log R}}}-\frac{1}{p^{1+\frac{1+iz_2}{\log R}}}+\frac{1}{p^{1+\frac{2+iz_1+iz_2}{\log R}}}\right)dz_1dz_2 + O\!\left(\frac{(\log R)^4}{z_0^{\varepsilon}}\right). \tag{3.4.14}$$

By integrating (3.4.13) by parts repeatedly we see

$$g(z) \ll_A \left(\frac{\log\log X}{1+|z|}\right)^A,$$

and we have the trivial bound

$$\prod_{p>z_0}\left(1-\frac{1}{p^{1+\frac{1+iz_1}{\log R}}}-\frac{1}{p^{1+\frac{1+iz_2}{\log R}}}+\frac{1}{p^{1+\frac{2+iz_1+iz_2}{\log R}}}\right) \ll (\log R)^{O(1)}.$$

Therefore, we may truncate the double integral in (3.4.14) to the region $|z_1|,|z_2| \leq \sqrt{\log R}$, with an error of size $O_A((\log R)^{-A})$. After doing so, we multiply and divide the integrand by Euler products of zeta-functions to arrive at

$$\mathcal{S} = \iint_{|z_i|\leq\sqrt{\log R}} g(z_1)g(z_2)\frac{\zeta\left(1+\frac{2+iz_1+iz_2}{\log R}\right)}{\zeta\left(1+\frac{1+iz_1}{\log R}\right)\zeta\left(1+\frac{1+iz_2}{\log R}\right)} \times \prod_{p\leq z_0}\frac{\left(1-\frac{1}{p^{1+\frac{2+iz_1+iz_2}{\log R}}}\right)}{\left(1-\frac{1}{p^{1+\frac{1+iz_1}{\log R}}}\right)\left(1-\frac{1}{p^{1+\frac{1+iz_2}{\log R}}}\right)}\prod_{p>z_0}\left(1+O\!\left(\frac{1}{p^2}\right)\right)dz_1dz_2 + O\!\left(\frac{1}{(\log R)^A}\right). \tag{3.4.15}$$

The product over primes $p > z_0$ in (3.4.15) is $1 + O(1/z_0)$. To estimate the product over $p \leq z_0$, observe that if $|s| \ll \sqrt{\log R}$, then

$$\sum_{p\leq z_0}\frac{1}{p-1}\left(1-p^{-s}\right) \ll \sum_{p\leq z_0}\frac{|s|\log p}{p} \ll |s|\log z_0 \ll \frac{(\log X)^{1/3}}{(\log R)^{1/2}},$$

which implies that

$$\prod_{p\le z_0}\left(1-\frac{1}{p^{1+s}}\right)=\exp\left(\sum_{p\le z_0}\log\left(1+\frac{1}{p-1}\left(1-p^{-s}\right)\right)\right)\prod_{p\le z_0}\left(1-\frac{1}{p}\right)$$
$$=\left(1+O\left(\frac{(\log X)^{1/3}}{(\log R)^{1/2}}\right)\right)\prod_{p\le z_0}\left(1-\frac{1}{p}\right).$$

We may also expand each zeta-function in (3.4.15) into its Laurent series. With these approximations, we deduce from (3.4.15) that

$$\mathcal{S}=\frac{1}{\log R}\prod_{p\le z_0}\left(1-\frac{1}{p}\right)^{-1}\iint\limits_{|z_i|\le\sqrt{\log R}} g(z_1)g(z_2)$$
$$\times\frac{(1+iz_1)(1+iz_2)}{2+iz_1+iz_2}\left(1+E(X,\vartheta,z_1,z_2)\right)dz_1dz_2$$
$$+O\Big((\log R)^{-A}\Big),$$

uniformly for $\log\ell\ll\log X$. Here $E(X,\vartheta,z_1,z_2)$ tends to zero as $X\to\infty$. By the rapid decay of $g(z)$, we may extend the range of integration to $\mathbb{R}^2$ without affecting our bound for the error term. By differentiating (3.4.12) under the integral sign and Fubini's theorem, we find

$$\iint\limits_{\mathbb{R}^2} g(z_1)g(z_2)\frac{(1+iz_1)(1+iz_2)}{2+iz_1+iz_2}dz_2dz_1=\int_0^\infty G'(t)^2dt. \tag{3.4.16}$$

The lemma now follows from (3.4.16) and (3.4.6). □

Lemma 3.4.3. *Let λ_d and D be as defined in* (3.4.8) *and* (3.4.9)*, respectively. Suppose that g is a multiplicative function such that $g(p)=1+O(p^{-\varepsilon})$ for all primes p. Then with $E_0(X)$ as in Lemma 3.4.2 we have*

$$\sum_{\substack{d\le D\\(d,\ell)=1}}\frac{\lambda_d}{d}g(d)=\frac{1+E_0(X)}{\log R}\prod_{\substack{p\le z_0\\p\nmid\ell}}\left(1-\frac{g(p)}{p}\right)\prod_{p\le z_0}\left(1-\frac{1}{p}\right)^{-1}+O_\varepsilon\left(\frac{1}{(\log R)^{2019}}\right),$$

uniformly in $\ell\ll X^{O(1)}$.

Proof. The definitions (3.4.8) and (3.4.9) of λ_d and D imply

$$\sum_{\substack{d\le D\\(d,\ell)=1}} \frac{\lambda_d}{d} g(d) = \sum_{\substack{b|P(z_0)\\ \omega(b)\le 2r_0\\ (b,\ell)=1}} \sum\sum_{\substack{m,n\le R\\ (mn,\ell P(z_0))=1}} \frac{\mu(b)\mu(m)\mu(n)}{b[m,n]} G\left(\frac{\log m}{\log R}\right) G\left(\frac{\log n}{\log R}\right) g(b[m,n]).$$

In the sum on the right-hand side, $g(b[m,n]) = g(b)g([m,n])$ because b and mn are coprime. Thus we may apply Lemma 3.4.2 and then Lemma 3.4.1 to arrive at Lemma 3.4.3. □

Lemma 3.4.4. *Let λ_d, D, g be as in Lemma 3.4.3. Suppose that h is a function such that $|h(p)| \ll_\varepsilon p^{-1+\varepsilon}$ for all primes p. Then with $E_0(X)$ as in Lemma 3.4.2 we have*

$$\sum_{\substack{d\le D\\(d,\ell)=1}} \frac{\lambda_d}{d} g(d) \sum_{p|d} h(p) = -\frac{1+E_0(X)}{\log R} \prod_{p\le z_0}\left(1-\frac{1}{p}\right)^{-1}$$
$$\times \sum_{\substack{p\le z_0\\ p\nmid \ell}} \frac{g(p)h(p)}{p} \prod_{\substack{q\le z_0\\ q\nmid p\ell}} \left(1-\frac{g(q)}{q}\right) + O_\varepsilon\left(\frac{1}{(\log R)^{2019}}\right),$$

uniformly for all integers ℓ such that $\log \ell \ll \log X$. (Here, the index q runs over primes q.)

Proof. The definitions (3.4.8) and (3.4.9) of λ_d and D imply

$$\sum_{\substack{d\le D\\(d,\ell)=1}} \frac{\lambda_d}{d} g(d) \sum_{p|d} h(p) = \sum_{\substack{b|P(z_0)\\ \omega(b)\le 2r_0\\ (b,\ell)=1}} \sum\sum_{\substack{m,n\le R\\ (mn,\ell P(z_0))=1}} \frac{\mu(b)\mu(m)\mu(n)}{b[m,n]} G\left(\frac{\log m}{\log R}\right) G\left(\frac{\log n}{\log R}\right)$$
$$\times g(b[m,n]) \sum_{p|bmn} h(p).$$

Since b and mn are coprime, $g(b[m,n]) = g(b)g([m,n])$ and

$$\sum_{p|bmn} h(p) = \sum_{p|b} h(p) + \sum_{p|mn} h(p).$$

We may ignore the sum over the $p|mn$ because the conditions $(mn, P(z_0)) = 1$ and $mn \le R^2$ imply

$$\sum_{p|mn} h(p) \ll \sum_{p|mn} p^{-1+\varepsilon} \ll \frac{\log R}{z_0^{1-\varepsilon}}.$$

We factor out $g(b)$ and $\sum_{p|b} h(p)$ from the sum over m, n and then apply Lemma 3.4.2 to deduce that

$$\sum_{\substack{d\leq D\\(d,\ell)=1}} \frac{\lambda_d}{d} g(d) \sum_{p|d} h(p) = \frac{1+E_0(X)}{\log R} \prod_{p\leq z_0} \left(1-\frac{1}{p}\right)^{-1} \times \sum_{\substack{b|P(z_0)\\ \omega(b)\leq 2r_0\\ (b,\ell)=1}} \frac{\mu(b)}{b} g(b) \sum_{p|b} h(p) + O\left(\frac{1}{(\log R)^{2019}}\right). \tag{3.4.17}$$

To estimate the b-sum, we interchange the order of summation and then relabel b as bp to write

$$\sum_{\substack{b|P(z_0)\\ \omega(b)\leq 2r_0\\ (b,\ell)=1}} \frac{\mu(b)}{b} g(b) \sum_{p|b} h(p) = \sum_{\substack{p\leq z_0\\ p\nmid \ell}} h(p) \sum_{\substack{b|P(z_0)\\ \omega(b)\leq 2r_0\\ (b,\ell)=1\\ p|b}} \frac{\mu(b)}{b} g(b)$$

$$= -\sum_{\substack{p\leq z_0\\ p\nmid \ell}} \frac{g(p)h(p)}{p} \sum_{\substack{b|P(z_0)\\ \omega(b)\leq 2r_0-1\\ (b,p\ell)=1}} \frac{\mu(b)}{b} g(b).$$

Lemma 3.4.4 now follows from Lemma 3.4.1 and (3.4.17). □

3.5 The mollified first moment

Our goal in this section is to asymptotically evaluate S_1. Recall from (3.2.1) that

$$S_1 = \sum_{p\equiv 1 \ (\mathrm{mod}\ 8)} (\log p) \Phi\left(\frac{p}{X}\right) L\left(\tfrac{1}{2}, \chi_p\right) M(p).$$

Recall the definition of $M(p)$ from (3.2.3), and the choice (3.2.5) we made for the mollifier coefficients b_m. We shall prove the following result.

Proposition 3.5.1. *Let $0 < \theta < \frac{1}{2}$ be fixed. If $X \geq X_0(\theta)$, then*

$$S_1 = \frac{1}{2(1-\frac{1}{\sqrt{2}})} \left(H(0) - \frac{1}{2\theta} H'(0)\right) \frac{X}{4} + O\left(\frac{X}{(\log X)^{1-\varepsilon}}\right).$$

The implied constant in the error term is effectively computable.

Let us begin in earnest, following the outline in Section 3.2. We apply Lemma 3.3.2 to write $L(\frac{1}{2},\chi_p)$ as a Dirichlet series. We insert the definition of $M(p)$ and obtain

$$S_1 = \frac{2}{\left(1-\frac{1}{\sqrt{2}}\right)^2} \sum_{\substack{m\le M\\ m \text{ odd}}} \frac{b_m}{\sqrt{m}} \sum_{\substack{n=1\\ n \text{ odd}}}^{\infty} \frac{1}{\sqrt{n}} \sum_{p\equiv 1 \,(\mathrm{mod}\ 8)} (\log p)\Phi\left(\frac{p}{X}\right)\omega_1\left(n\sqrt{\frac{\pi}{p}}\right)\left(\frac{mn}{p}\right).$$

The main term arises from the terms with $mn = \square$. Let us denote this portion of S_1 by $S_1^{\square}$. We denote the complementary portion with $mn \neq \square$ by $S_1^{\neq}$. Therefore

$$S_1 = S_1^{\square} + S_1^{\neq},$$

where

$$\begin{aligned}
S_1^{\square} &= \frac{2}{\left(1-\frac{1}{\sqrt{2}}\right)^2} \sum_{\substack{m\le M\\ m \text{ odd}\\ mn=\square}} \sum_{\substack{n=1\\ n \text{ odd}}}^{\infty} \frac{b_m}{\sqrt{m}}\frac{1}{\sqrt{n}} \sum_{p\equiv 1 \,(\mathrm{mod}\ 8)} (\log p)\Phi\left(\frac{p}{X}\right)\omega_1\left(n\sqrt{\frac{\pi}{p}}\right)\left(\frac{mn}{p}\right),\\
S_1^{\neq} &= \frac{2}{\left(1-\frac{1}{\sqrt{2}}\right)^2} \sum_{\substack{m\le M\\ m \text{ odd}\\ mn\neq\square}} \sum_{\substack{n=1\\ n \text{ odd}}}^{\infty} \frac{b_m}{\sqrt{m}}\frac{1}{\sqrt{n}} \sum_{p\equiv 1 \,(\mathrm{mod}\ 8)} (\log p)\Phi\left(\frac{p}{X}\right)\omega_1\left(n\sqrt{\frac{\pi}{p}}\right)\left(\frac{mn}{p}\right).
\end{aligned} \tag{3.5.1}$$

We treat first the main term $S_1^{\square}$, and later we will bound the error term $S_1^{\neq}$.

3.5.1 Main term

Recall that b_m is supported on square-free integers m. Therefore, $mn = \square$ if and only if $n = mk^2$, where k is a positive integer. We make this change of variables and then interchange orders of summation to obtain

$$S_1^{\square} = \frac{2}{\left(1-\frac{1}{\sqrt{2}}\right)^2} \sum_{p\equiv 1 \,(\mathrm{mod}\ 8)} (\log p)\Phi\left(\frac{p}{X}\right) \sum_{\substack{m\le M\\ (m,2p)=1}} \frac{b_m}{m} \sum_{\substack{k=1\\ (k,2p)=1}}^{\infty} \frac{1}{k}\omega_1\left(mk^2\sqrt{\frac{\pi}{p}}\right).$$

By the rapid decay of ω_1 (Lemma 3.3.1) we see that the contribution from those k with $(k,p) > 1$ is $O_A(X^{-A})$, so we may safely ignore this condition. We may also ignore the condition $(m,p)=1$, since $m \le M < p$. We insert the definition (3.3.1) of $\omega_1(\xi)$ and interchange to deduce that for any $c>0$ we have

$$\sum_{\substack{k=1\\(k,2)=1}}^{\infty} \frac{1}{k}\omega_1\left(mk^2\sqrt{\frac{\pi}{p}}\right)$$
$$= \frac{1}{2\pi i}\int_{(c)} \frac{\Gamma(\frac{s}{2}+\frac{1}{4})}{\Gamma(\frac{1}{4})}\left(1-\frac{1}{2^{\frac{1}{2}-s}}\right)\left(1-\frac{1}{2^{1+2s}}\right)\zeta(1+2s)\left(\frac{p}{\pi}\right)^{s/2} m^{-s}\,\frac{ds}{s}.$$

We move the line of integration to $\operatorname{Re} s = -\frac{1}{2}+\varepsilon$, leaving a residue at $s=0$. The new integral is $O_\varepsilon\left(p^{-\frac{1}{4}+\varepsilon}m^{\frac{1}{2}-\varepsilon}\right)$. Using $b_m \ll 1$, we see that the total contribution of this error term is $\ll X^{\frac{3}{4}+\varepsilon}M^{\frac{1}{2}}$. This is $O(X^{1-\varepsilon})$ by (3.2.4). Writing the residue at $s=0$ as an integral along a small circle around 0, we deduce that

$$\begin{aligned} S_1^{\square} = O(X^{1-\varepsilon}) + \frac{2}{\left(1-\frac{1}{\sqrt{2}}\right)^2} &\sum_{p\equiv 1 \,(\mathrm{mod}\ 8)} (\log p)\Phi\left(\frac{p}{X}\right)\sum_{\substack{m\le M\\(m,2)=1}}\frac{b_m}{m} \\ &\times \frac{1}{2\pi i}\oint_{|s|=\frac{1}{2\log X}} \frac{\Gamma(\frac{s}{2}+\frac{1}{4})}{\Gamma(\frac{1}{4})}\left(1-\frac{1}{2^{\frac{1}{2}-s}}\right)\left(1-\frac{1}{2^{1+2s}}\right) \\ &\times \zeta(1+2s)\left(\frac{p}{\pi}\right)^{s/2} m^{-s}\,\frac{ds}{s}. \end{aligned} \tag{3.5.1.1}$$

We next use the definition $b_m = \mu(m)H\left(\frac{\log m}{\log M}\right)$ and the Fourier inversion formula (compare with (3.4.12),(3.4.13))

$$H(t) \;=\; \int_{-\infty}^{\infty} h(z)e^{-t(1+iz)}\,dz, \tag{3.5.1.2}$$

where

$$h(z) \;=\; \int_{-\infty}^{\infty} e^t H(t)e^{izt}\,dt, \tag{3.5.1.3}$$

to write

$$\begin{aligned} \sum_{\substack{m\le M\\(m,2)=1}} \frac{b_m}{m}m^{-s} &= \int_{-\infty}^{\infty} h(z)\sum_{\substack{m=1\\(m,2)=1}}^{\infty}\frac{\mu(m)}{m^{1+s+\frac{1+iz}{\log M}}}\,dz \\ &= \int_{-\infty}^{\infty} h(z)\left(1-\frac{1}{2^{1+s+\frac{1+iz}{\log M}}}\right)^{-1}\zeta^{-1}\left(1+s+\frac{1+iz}{\log M}\right)dz. \end{aligned}$$

From repeated integration by parts we obtain

$$h(z) \ll_j \frac{1}{(1+|z|)^j}, \tag{3.5.1.4}$$

and therefore we may truncate this integral to the range $|z| \le \sqrt{\log M}$. Thus,

$$\sum_{\substack{m \le M \\ (m,2)=1}} \frac{b_m}{m} m^{-s} = \int_{|z| \le \sqrt{\log M}} h(z) \left(1 - \frac{1}{2^{1+s+\frac{1+iz}{\log M}}}\right)^{-1} \zeta^{-1}\left(1 + s + \frac{1+iz}{\log M}\right) dz$$
$$+ O_A\left(\frac{1}{(\log X)^A}\right).$$

For $|s| = \frac{1}{2\log X}$ and $|z| \le \sqrt{\log M}$, we may write

$$\left(1 - \frac{1}{2^{1+s+\frac{1+iz}{\log M}}}\right)^{-1} \zeta^{-1}\left(1 + s + \frac{1+iz}{\log M}\right)$$

as a power series and arrive at

$$\sum_{\substack{m \le M \\ (m,2)=1}} \frac{b_m}{m} m^{-s} = 2\int_{|z| \le \sqrt{\log M}} h(z) \left(s + \frac{1+iz}{\log M}\right) dz + O\left(\frac{1}{(\log X)^2}\right).$$

We may extend the range of integration to the entire real line, with negligible error, because of (3.5.1.4). The definition of $H(t)$ implies that

$$H'(t) = -(1+iz) \int_{-\infty}^{\infty} h(z) e^{-t(1+iz)}\, dz.$$

Therefore

$$\int_{-\infty}^{\infty} h(z) \left(s + \frac{1+iz}{\log M}\right) dz = sH(0) - \frac{1}{\log M} H'(0),$$

and hence

$$\sum_{\substack{m \le M \\ (m,2)=1}} \frac{b_m}{m} m^{-s} = 2sH(0) - \frac{2}{\log M} H'(0) + O\left(\frac{1}{(\log X)^2}\right). \tag{3.5.1.5}$$

We insert (3.5.1.5) into (3.5.1.1) to obtain

$$
\begin{aligned}
S_1^{\square} =& \frac{4}{\left(1-\frac{1}{\sqrt{2}}\right)^2} \sum_{p\equiv 1 \pmod 8} (\log p)\Phi\left(\frac{p}{X}\right) \frac{1}{2\pi i} \oint_{|s|=\frac{1}{2\log X}} \frac{\Gamma(\frac{s}{2}+\frac{1}{4})}{\Gamma(\frac{1}{4})} \left(1-\frac{1}{2^{\frac{1}{2}-s}}\right) \\
&\times \left(1-\frac{1}{2^{1+2s}}\right)\zeta(1+2s)\left(\frac{p}{\pi}\right)^{s/2}\left(sH(0)-\frac{1}{\log M}H'(0)\right)\frac{ds}{s} \\
&+ O\left(\frac{X}{\log X}\right).
\end{aligned}
$$

We evaluate the integral using the formula

$$
\operatorname*{Res}_{s=0} g(s) = \frac{1}{(n-1)!}\frac{d^{n-1}}{ds^{n-1}} s^n g(s)\bigg|_{s=0} \tag{3.5.1.6}
$$

for a pole of a function $g(s)$ at $s=0$ of order at most n. This yields

$$
S_1^{\square} = \frac{1}{\left(1-\frac{1}{\sqrt{2}}\right)} \sum_{p\equiv 1 \pmod 8} (\log p)\Phi\left(\frac{p}{X}\right) \left(H(0)-\frac{\log p}{2\log M}H'(0)\right) + O\left(\frac{X}{\log X}\right).
$$

By the support of Φ we have $\log p = \log X + O(1)$. We then use the prime number theorem in arithmetic progressions and partial summation to obtain

$$
S_1^{\square} = \frac{1}{\left(1-\frac{1}{\sqrt{2}}\right)} \left(H(0)-\frac{\log X}{2\log M}H'(0)\right)\frac{X}{4}\widehat{\Phi}(0) + O\left(\frac{X}{\log X}\right). \tag{3.5.1.7}
$$

Now (3.5.1.7) gives the main term for Proposition 3.5.1.

3.5.2 Preparation of the off-diagonal

We turn to bounding $S_1^{\neq}$. In order to complete the proof of Proposition 3.5.1, we prove

$$
S_1^{\neq} \ll \frac{X}{(\log X)^{1-\varepsilon}}. \tag{3.5.2.1}
$$

We need to perform some technical massaging before $S_1^{\neq}$ is in a suitable form. Recall from (3.5.1) that

$$S_1^{\neq} = \frac{2}{\left(1-\frac{1}{\sqrt{2}}\right)^2} \sum_{\substack{m\leq M \\ m \text{ odd}\\ mn\neq\square}} \sum_{\substack{n=1\\ n \text{ odd}}}^{\infty} \frac{b_m}{\sqrt{mn}} \sum_{p\equiv 1 \,(\mathrm{mod}\, 8)} (\log p)\Phi\left(\frac{p}{X}\right)\omega_1\left(n\sqrt{\frac{\pi}{p}}\right)\left(\frac{mn}{p}\right).$$

We begin by uniquely writing $n = rk^2$, where r is square-free and k is an integer (this variable k is unrelated to the variable k appearing in the analysis for $S_1^{\square}$). The condition $mn \neq \square$ is equivalent to $m \neq r$, since both m and r are square-free. It follows that

$$S_1^{\neq} = \frac{2}{\left(1-\frac{1}{\sqrt{2}}\right)^2} \sum_{\substack{m\leq M \\ m \text{ odd}}} \frac{b_m}{\sqrt{m}} \sum_{\substack{r=1\\ r \text{ odd}\\ r\neq m}}^{\infty} \sum_{\substack{k=1\\ k \text{ odd}}}^{\infty} \frac{\mu^2(r)}{k\sqrt{r}}$$
$$\times \sum_{p\equiv 1 \,(\mathrm{mod}\, 8)} (\log p)\Phi\left(\frac{p}{X}\right)\omega_1\left(rk^2\sqrt{\frac{\pi}{p}}\right)\left(\frac{mrk^2}{p}\right)$$

We next factor out the greatest common divisor, say g, of m and r. We change variables $m \to gm, r \to gr$ and obtain

$$S_1^{\neq} = \frac{2}{\left(1-\frac{1}{\sqrt{2}}\right)^2} \sum_{g \text{ odd}} \frac{\mu^2(g)}{g} \sum_{\substack{m\leq M/g \\ (m,2g)=1}} \frac{b_{mg}}{\sqrt{m}} \sum_{\substack{r=1\\ (r,2g)=1\\ (m,r)=1\\ mr>1}}^{\infty} \frac{\mu^2(r)}{\sqrt{r}} \sum_{\substack{k=1\\ k \text{ odd}}}^{\infty} \frac{1}{k}$$
$$\times \sum_{p\equiv 1 \,(\mathrm{mod}\, 8)} (\log p)\Phi\left(\frac{p}{X}\right)\omega_1\left(grk^2\sqrt{\frac{\pi}{p}}\right)\left(\frac{mrg^2k^2}{p}\right).$$

Observe that the support of b_{gm} forces $g \leq M < X^{\frac{1}{2}}$, but we prefer not to indicate this explicitly.

Clearly we have $\left(\frac{g^2k^2}{p}\right) = 1$ for $p \nmid gk$ and $= 0$ otherwise. Since $g \leq M < p$ the condition $p \nmid g$ is automatically satisfied. By Lemma 3.3.1 we may truncate the sum over k to $k \leq X^{\frac{1}{4}+\varepsilon}$ at the cost of an error $O(X^{-1})$, say. We may similarly truncate the sum on r to $r \leq X^{\frac{1}{2}+\varepsilon}$. With k suitably reduced we may drop the condition $p \nmid k$, and then we use the rapid decay of ω_1 again to extend the sum on k to infinity. It follows that

$$S_1^{\neq} = \frac{2}{\left(1-\frac{1}{\sqrt{2}}\right)^2} \sum_{g \text{ odd}} \frac{\mu^2(g)}{g} \sum_{\substack{m \le M/g \\ (m,2g)=1}} \frac{b_{mg}}{\sqrt{m}} \sum_{\substack{r \le X^{1/2+\varepsilon} \\ (r,2g)=1 \\ (m,r)=1 \\ mr>1}} \frac{\mu^2(r)}{\sqrt{r}} \sum_{\substack{k=1 \\ k \text{ odd}}}^{\infty} \frac{1}{k}$$

$$\times \sum_{p \equiv 1 \pmod 8} (\log p) \Phi\left(\frac{p}{X}\right) \omega_1\left(grk^2\sqrt{\frac{\pi}{p}}\right)\left(\frac{mr}{p}\right) + O(X^{-1}). \tag{3.5.2.2}$$

We next detect the congruence condition $p \equiv 1 \pmod 8$ with multiplicative characters modulo 8. Therefore

$$\begin{aligned}\sum_{p \equiv 1 \pmod 8} &(\log p) \Phi\left(\frac{p}{X}\right) \omega_1\left(grk^2\sqrt{\frac{\pi}{p}}\right)\left(\frac{mr}{p}\right) \\ &= \frac{1}{4} \sum_{\gamma \in \{\pm 1, \pm 2\}} \sum_p (\log p) \Phi\left(\frac{p}{X}\right) \omega_1\left(grk^2\sqrt{\frac{\pi}{p}}\right)\left(\frac{\gamma mr}{p}\right).\end{aligned} \tag{3.5.2.3}$$

Since m and r are odd and square-free and $(m,r)=1$, it follows that mr is odd and square-free. Hence, for each $\gamma \in \{1,-1,2,-2\}$, the integer γmr is square-free. Therefore $\gamma mr \equiv 1$, 2, or 3 (mod 4). If $\gamma mr \equiv 1 \pmod 4$, then $\left(\frac{\gamma mr}{\cdot}\right)$ is a real primitive character modulo $|\gamma mr|$, while if $\gamma mr \equiv 2$ or 3 (mod 4), then $\left(\frac{4\gamma mr}{\cdot}\right)$ is a real primitive character modulo $|4\gamma mr|$ (see [65, Theorem 2.2.15]). Moreover, for p odd, $\left(\frac{4\gamma mr}{p}\right) = \left(\frac{\gamma mr}{p}\right)$. Therefore the sum in (3.5.2.3) is equal to

$$\frac{1}{4} \sum_{\gamma \in \{\pm 1, \pm 2\}} \sum_p (\log p) \Phi\left(\frac{p}{X}\right) \omega_1\left(grk^2\sqrt{\frac{\pi}{p}}\right) \chi_{\gamma mr}(p), \tag{3.5.2.4}$$

where $\chi_{\gamma mr}(\cdot) = \left(\frac{\gamma mr}{\cdot}\right)$ if $\gamma mr \equiv 1 \pmod 4$, and $\chi_{\gamma mr}(\cdot) = \left(\frac{4\gamma mr}{\cdot}\right)$ if $\gamma mr \equiv 2$ or 3 (mod 4), so that $\chi_{\gamma mr}(\cdot)$ is a real primitive character for all the relevant γ, m, r. Also, since $mr > 1$, we see that γmr is never 1, so each $\chi_{\gamma mr}$ is nonprincipal.

We insert the definition of ω_1 into (3.5.2.4) in order to facilitate a separation of variables. Recalling (3.5.2.2) and (3.5.2.3), we interchange the order of summation and integration to obtain

$$S_1^{\neq} = O(1) + \frac{2}{\left(1-\frac{1}{\sqrt{2}}\right)^2} \sum_{g \text{ odd}} \frac{\mu^2(g)}{g} \sum_{\substack{m \le M/g \\ (m,2g)=1}} \frac{b_{mg}}{\sqrt{m}} \sum_{\substack{r \le X^{1/2+\varepsilon} \\ (r,2g)=1 \\ (m,r)=1 \\ mr>1}} \frac{\mu^2(r)}{\sqrt{r}} \frac{1}{4} \sum_{\gamma \in \{\pm 1, \pm 2\}} \sum_{\substack{k=1 \\ k \text{ odd}}}^{\infty} \frac{1}{k}$$
$$\times \frac{1}{2\pi i} \int_{(c)} \frac{\Gamma(\frac{s}{2}+\frac{1}{4})}{\Gamma(\frac{1}{4})} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right) \pi^{-s/2} \left(grk^2\right)^{-s} \sum_p (\log p) \Phi\left(\frac{p}{X}\right) \chi_{\gamma mr}(p) p^{s/2} \frac{ds}{s}.$$

We choose $c = \frac{1}{\log X}$, so that $p^{s/2}$ is bounded in absolute value. We can put the summation on k inside of the integral, where it becomes a zeta factor, and we obtain

$$S_1^{\neq} = O(1) + \frac{2}{\left(1-\frac{1}{\sqrt{2}}\right)^2} \sum_{g \text{ odd}} \frac{\mu^2(g)}{g} \sum_{\substack{m \le M/g \\ (m,2g)=1}} \frac{b_{mg}}{\sqrt{m}} \sum_{\substack{r \le X^{1/2+\varepsilon} \\ (r,2g)=1 \\ (m,r)=1 \\ mr>1}} \frac{\mu^2(r)}{\sqrt{r}}$$
$$\times \frac{1}{4} \sum_{\gamma \in \{\pm 1, \pm 2\}} \frac{1}{2\pi i} \int_{(c)} \frac{\Gamma(\frac{s}{2}+\frac{1}{4})}{\Gamma(\frac{1}{4})} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right) \left(1 - \frac{1}{2^{1+2s}}\right) \zeta(1+2s)$$
$$\times \pi^{-s/2} (gr)^{-s} \sum_p (\log p) \Phi\left(\frac{p}{X}\right) \chi_{\gamma mr}(p) p^{s/2} \frac{ds}{s}.$$

It is more convenient to replace the $\log p$ factor with the von Mangoldt function $\Lambda(n)$. By trivial estimation we have

$$\sum_p (\log p) \Phi\left(\frac{p}{X}\right) \chi_{\gamma mr}(p) p^{s/2} = \sum_n \Lambda(n) \Phi\left(\frac{n}{X}\right) \chi_{\gamma mr}(n) n^{s/2} + O(X^{1/2}).$$

When we sum the error term over m, g, r and integrate over s, the total contribution is $O(X^{1-\varepsilon})$, provided $\varepsilon = \varepsilon(\theta) > 0$ is sufficiently small. By the rapid decay of the Γ function in vertical strips we can truncate the integral to $|\text{Im}(s)| \le (\log X)^2$, at the cost of a negligible error. We therefore obtain

$$S_1^{\neq} = O(X^{1-\varepsilon}) + \frac{2}{\left(1-\frac{1}{\sqrt{2}}\right)^2} \sum_{g \text{ odd}} \frac{\mu^2(g)}{g} \sum_{\substack{m \le M/g \\ (m,2g)=1}} \frac{b_{mg}}{\sqrt{m}} \sum_{\substack{r \le X^{1/2+\varepsilon} \\ (r,2g)=1 \\ (m,r)=1 \\ mr>1}} \frac{\mu^2(r)}{\sqrt{r}} \frac{1}{4} \sum_{\gamma \in \{\pm 1, \pm 2\}}$$
$$\times \frac{1}{2\pi i} \int_{\frac{1}{\log X} - i(\log X)^2}^{\frac{1}{\log X} + i(\log X)^2} \frac{\Gamma(\frac{s}{2}+\frac{1}{4})}{\Gamma(\frac{1}{4})} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)\left(1 - \frac{1}{2^{1+2s}}\right) \zeta(1+2s)$$
$$\times \left(\frac{X}{\pi}\right)^{s/2} (gr)^{-s} \sum_n \Lambda(n) \Phi_s\left(\frac{n}{X}\right) \chi_{\gamma mr}(n) \, \frac{ds}{s}. \tag{3.5.2.5}$$

Having arrived at (3.5.2.5), we are finished with the preparatory technical manipulations. We proceed to show that $S_1^{\neq}$ is small. As discussed in Section 3.2, we apply three different arguments, depending on the size of mr. We call these ranges Regimes I, II, and III, which correspond to small, medium, and large values of mr. In Regime I we have $1 < mr \ll \exp(\varpi\sqrt{\log x})$, where $\varpi > 0$ is a sufficiently small, fixed constant. Regime II corresponds to $\exp(\varpi\sqrt{\log x}) \ll mr \ll X^{\frac{1}{10}}$, and Regime III corresponds to $X^{\frac{1}{10}} \ll mr \ll MX^{\frac{1}{2}+\varepsilon}$. We then write

$$S_1^{\neq} = E_1 + E_2, \tag{3.5.2.6}$$

where E_1 contains those terms with $mr \ll \exp(\varpi\sqrt{\log x})$, and E_2 contains those terms with $mr \gg \exp(\varpi\sqrt{\log x})$. We claim the bounds

$$\begin{aligned} E_1 &\ll \frac{X}{(\log X)^{1-\varepsilon}}, \\ E_2 &\ll X \exp(-c\varpi\sqrt{\log x}), \end{aligned} \tag{3.5.2.7}$$

where $c > 0$ is some absolute constant. Taking together (3.5.2.6) and (3.5.2.7) clearly gives (3.5.2.1), and this yields Proposition 3.5.1. It therefore suffices to show (3.5.2.7).

3.5.3 Regime I

We first bound E_1, which is precisely the contribution of Regime I. By definition, we have

$$E_1 := \frac{2}{\left(1-\frac{1}{\sqrt{2}}\right)^2} \sum_{g \text{ odd}} \frac{\mu^2(g)}{g} \sum_{\substack{m \le M/g \\ (m,2g)=1}} \frac{b_{mg}}{\sqrt{m}} \sum_{\substack{r \le X^{1/2+\varepsilon} \\ (r,2g)=1 \\ (m,r)=1 \\ 1 < mr \ll \exp(\varpi\sqrt{\log x})}} \frac{\mu^2(r)}{\sqrt{r}} \frac{1}{4} \sum_{\gamma \in \{\pm 1, \pm 2\}}$$
$$\times \frac{1}{2\pi i} \int_{\frac{1}{\log X} - i(\log X)^2}^{\frac{1}{\log X} + i(\log X)^2} \frac{\Gamma(\frac{s}{2}+\frac{1}{4})}{\Gamma(\frac{1}{4})} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)\left(1 - \frac{1}{2^{1+2s}}\right) \zeta(1+2s)$$
$$\times \left(\frac{X}{\pi}\right)^{s/2} (gr)^{-s} \sum_n \Lambda(n) \Phi_s\left(\frac{n}{X}\right) \chi_{\gamma mr}(n) \frac{ds}{s}. \tag{3.5.3.1}$$

We transform the sum on n with partial summation to obtain

$$\sum_n \Lambda(n) \Phi_s\left(\frac{n}{X}\right) \chi_{\gamma mr}(n) = -\int_0^\infty \frac{1}{X} \Phi_s'\left(\frac{w}{X}\right) \left(\sum_{n \le w} \Lambda(n) \chi_{bmr}(n)\right) dw. \tag{3.5.3.2}$$

By [66, equation (8) of Chapter 20], we have

$$\sum_{n \le w} \Lambda(n) \chi_{\gamma mr}(n) = -\frac{w^{\beta_1}}{\beta_1} + O\left(w \exp(-c_1\sqrt{\log w})\right), \tag{3.5.3.3}$$

where $c_1 > 0$ is some absolute constant, and the term $-w^{\beta_1}/\beta_1$ only appears if $L(s, \chi_{\gamma mr})$ has a real zero β_1 which satisfies $\beta_1 > 1 - \frac{c_2}{\log|\gamma mr|}$ for some sufficiently small constant $c_2 > 0$. All the constants in (3.5.3.3), implied or otherwise, are effective.

The contribution from the error term in (3.5.3.3) is easy to control. Observe that

$$\int_0^\infty \frac{1}{X} \left|\Phi_s'\left(\frac{w}{X}\right)\right| dw = \int_0^\infty |\Phi_s'(u)|\, du \ll |s| + 1, \tag{3.5.3.4}$$

uniformly in s with $\mathrm{Re}(s)$ bounded. Taking (3.5.3.1),(3.5.3.2) and (3.5.3.4) together, we see the error term of (3.5.3.3) contributes

$$\ll X \exp(c_3(\varpi - c_1)\sqrt{\log X}) \tag{3.5.3.5}$$

to E_1, where $c_3 > 0$ is some absolute constant. The bound (3.5.3.5) is adequate for (3.5.2.7) provided we choose $\varpi > 0$ sufficiently small in terms of c_1.

The conductor of the primitive character $\chi_{\gamma mr}$ is $\leq \exp(2\varpi\sqrt{\log X})$. We apply Page's theorem [66, equation (9) of Chapter 14], which implies that, for some fixed absolute constant $c_4 > 0$, there is at most one real primitive character $\chi_{\gamma mr}$ with modulus $\leq \exp(2\varpi\sqrt{\log X})$ for which the L-function $L(s, \chi_{\gamma mr})$ has a real zero satisfying

$$\beta_1 > 1 - \frac{c_4}{2\varpi\sqrt{\log X}}. \tag{3.5.3.6}$$

To estimate the contribution of the possible term $-\frac{w^{\beta_1}}{\beta_1}$, we evaluate the integral

$$\int_0^\infty \frac{w^{\beta_1}}{\beta_1}\frac{1}{X}\Phi_s'\left(\frac{w}{X}\right)\,dw$$

arising from (3.5.3.2) and (3.5.3.3). We make the change of variable $\frac{w}{X} \mapsto u$ and integrate by parts to see that this integral equals

$$X^{\beta_1}\int_0^\infty \frac{u^{\beta_1}}{\beta_1}\Phi_s'(u)\,du \;=\; -X^{\beta_1}\int_0^\infty \Phi_s(u)u^{\beta_1-1}\,du \;=\; -X^{\beta_1}\Phi^\dagger\left(\frac{s}{2}+\beta_1\right).$$

We assume that a real zero satisfying (3.5.3.6) does exist, for otherwise we already have an acceptable bound for E_1. Let q^* denote the conductor of the exceptional character $\chi_{\gamma mr}$ for which the real zero β_1 satisfying (3.5.3.6) exists. Then we have

$$\begin{aligned} E_1 = {} & -\frac{1}{2\pi i}\frac{\sqrt{\gamma^*}}{2\left(1-\frac{1}{\sqrt{2}}\right)^2}\frac{X^{\beta_1}}{\sqrt{|q^*|}}\int_{\frac{1}{\log X}-i(\log X)^2}^{\frac{1}{\log X}+i(\log X)^2}\frac{\Gamma(\frac{s}{2}+\frac{1}{4})}{\Gamma(\frac{1}{4})} \\ & \times\left(1-\frac{1}{2^{\frac{1}{2}-s}}\right)\left(1-\frac{1}{2^{1+2s}}\right)\left(\frac{X}{\pi}\right)^{s/2}\zeta(1+2s)\Phi^\dagger\left(\frac{s}{2}+\beta_1\right) \\ & \times\sum\sum_{\substack{1<mr\ll\exp(\varpi\sqrt{\log X}) \\ (mr,2)=1 \\ (m,r)=1 \\ \gamma mr=q^*}}\frac{\mu^2(r)}{r^s}\sum_{(g,2mr)=1}\frac{\mu^2(g)b_{gm}}{g^{1+s}}\frac{ds}{s} \\ & + O\left(X\exp(-c_5\sqrt{\log X})\right), \end{aligned} \tag{3.5.3.7}$$

where $c_5 > 0$ is some constant, and γ^* is some bounded power of two.

We next write $b_{gm} = \mu(gm)H(\frac{\log gm}{\log M})$ and apply Fourier inversion as in (3.5.1.2) and

(3.5.1.3) to obtain

$$\begin{aligned}\sum_{(g,2mr)=1} \frac{b_{mg}}{g^{1+s}} &= \mu(m) \int_{-\infty}^{\infty} \frac{1}{m^{\frac{1+iz}{\log M}}} h(z) \\ &\times \prod_{p|2mr} \left(1 - \frac{1}{p^{1+s+\frac{1+iz}{\log M}}}\right)^{-1} \zeta^{-1}\left(1+s+\frac{1+iz}{\log M}\right) dz.\end{aligned} \tag{3.5.3.8}$$

By (3.5.1.4) we can truncate the integral in (3.5.3.8) to $|z| \leq \sqrt{\log M}$ at the cost of an error of size $O_B(d_2(mr)(\log X)^{-B})$. This error contributes to (3.5.3.7)

$$\ll_B \frac{X}{(\log X)^{B+O(1)}},$$

which is acceptable. We therefore have

$$\begin{aligned} E_1 = &-\frac{X^{\beta_1}}{2\left(1-\frac{1}{\sqrt{2}}\right)^2} \frac{\sqrt{\gamma^*}}{\sqrt{|q^*|}} \sum_{\substack{m \leq M, r \leq X^{\frac{1}{2}+\varepsilon} \\ (mr,2)=1 \\ (m,r)=1 \\ \gamma mr = q^*}} \sum \mu(m)\mu^2(r) \\ &\times \frac{1}{2\pi i} \int_{\frac{1}{\log X} - i(\log X)^2}^{\frac{1}{\log X} + i(\log X)^2} \frac{\Gamma(\frac{s}{2}+\frac{1}{4})}{\Gamma(\frac{1}{4})} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right) r^{-s} \\ &\times \left(1 - \frac{1}{2^{1+2s}}\right) \zeta(1+2s) \left(\frac{X}{\pi}\right)^{s/2} \Phi^\dagger\left(\frac{s}{2}+\beta_1\right) \int_{|z| \leq \sqrt{\log M}} \frac{1}{m^{\frac{1+iz}{\log M}}} h(z) \\ &\times \prod_{p|2mr} \left(1 - \frac{1}{p^{1+s+\frac{1+iz}{\log M}}}\right)^{-1} \zeta^{-1}\left(1+s+\frac{1+iz}{\log M}\right) dz \frac{ds}{s} + O\left(\frac{X}{\log X}\right). \end{aligned} \tag{3.5.3.9}$$

We handle the s-integral in (3.5.3.9) by moving the line of integration to $\mathrm{Re}(s) = -\frac{c_6}{\log \log X}$, where $c_6 > 0$ is small enough that $\zeta(1+s+\frac{1+iz}{\log M})$ has no zeros in the region $\mathrm{Re}(s) \geq -\frac{c_6}{\log\log X}, \mathrm{Im}(s) \leq (\log X)^2$. By moving the line of integration we pick up a contribution from the pole at $s = 0$. We write this residue as an integral around a circle of small radius centered at the origin, and thereby deduce

$$
\begin{aligned}
E_1 = &-\frac{X^{\beta_1}}{2\left(1-\frac{1}{\sqrt{2}}\right)^2}\frac{\sqrt{\gamma^*}}{\sqrt{|q^*|}}\sum_{\substack{m\le M, r\le X^{1/2+\varepsilon}\\ (mr,2)=1\\ (m,r)=1\\ \gamma mr=q^*}}\sum \mu(m)\mu^2(r)\\
&\times\frac{1}{2\pi i}\oint_{|s|=\frac{1}{\log X}}\frac{\Gamma(\frac{s}{2}+\frac14)}{\Gamma(\frac14)}\left(1-\frac{1}{2^{\frac12-s}}\right)r^{-s}\\
&\times\left(1-\frac{1}{2^{1+2s}}\right)\zeta(1+2s)\left(\frac{X}{\pi}\right)^{s/2}\Phi^\dagger\left(\frac{s}{2}+\beta_1\right)\int_{|z|\le\sqrt{\log M}}\frac{1}{m^{\frac{1+iz}{\log M}}}h(z)\\
&\times\prod_{p|2mr}\left(1-\frac{1}{p^{1+s+\frac{1+iz}{\log M}}}\right)^{-1}\zeta^{-1}\left(1+s+\frac{1+iz}{\log M}\right)dz\frac{ds}{s}+O\left(\frac{X}{\log X}\right).
\end{aligned}
\tag{3.5.3.10}
$$

We have the bound

$$\beta_1 < 1-\frac{c_7}{\sqrt{|q^*|}(\log|q^*|)^2}, \tag{3.5.3.11}$$

where $c_7>0$ is a fixed absolute constant (see [66, equation (12) of Chapter 14]). If q^* satisfies $|q^*|\le(\log X)^{2-\varepsilon}$ then by (3.5.3.11) we derive

$$X^{\beta_1}\ll X\exp(-c_7(\log X)^{\varepsilon/3}).$$

By estimating (3.5.3.10) trivially we then obtain

$$E_1\ll X\exp(-c_7(\log X)^{\varepsilon/4}),$$

which is an acceptable bound. We may therefore assume that q^* satisfies

$$|q^*|>(\log X)^{2-\varepsilon}. \tag{3.5.3.12}$$

For $|s|=\frac{1}{\log X}$ we have the bounds

$$\zeta(1+2s)\ll\log X,\qquad \zeta^{-1}\left(1+s+\frac{1+iz}{\log M}\right)\ll\frac{1+|z|}{\log X}.$$

Using these bounds and (3.5.3.12) we deduce by trivial estimation that

$$(3.5.3.10) \ll \frac{X}{|q^*|^{1/2-o(1)}} \ll \frac{X}{(\log X)^{1-\varepsilon}}.$$

This completes the proof of the bound for E_1 in (3.5.2.7).

3.5.4 Regime II

It remains to prove the bound for E_2 in (3.5.2.7). From (3.5.2.5) and (3.5.2.6) we see that E_2 is the contribution from those m and r in Regimes II and III. The estimates in regimes II and III are less delicate than those in regime I, and consequently the arguments are easier.

In (3.5.2.5) we write $q = \gamma m r$. After breaking q into dyadic segments we find

$$E_2 \ll (\log X)^{O(1)} \sum_{\substack{Q=2^j \\ Q \gg \exp(\varpi\sqrt{\log X}) \\ Q \ll MX^{1/2+\varepsilon}}} \mathcal{E}(Q),$$

where

$$\mathcal{E}(Q) := Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \left| \sum_n \Lambda(n) \Phi_{s_0}\left(\frac{n}{X}\right) \chi(n) \right|.$$

Here s_0 is some complex number with $\mathrm{Re}(s_0) = \frac{1}{\log X}$ and $|\mathrm{Im}(s_0)| \leq (\log X)^2$. In order to prove (3.5.2.7) it therefore suffices to show that

$$\mathcal{E}(Q) \ll X \exp(-c_8 \varpi \sqrt{\log X}) \tag{3.5.4.1}$$

for each Q satisfying $\exp(\varpi\sqrt{\log X}) \ll Q \ll MX^{\frac{1}{2}+\varepsilon}$. In this subsection we treat the Q belonging to Regime II, that is, those Q which satisfy $Q \ll X^{\frac{1}{10}}$. In the next subsection we treat the Q in Regime III, which satisfy $Q \gg X^{\frac{1}{10}}$.

In Regime II we employ zero-density estimates. We begin by writing Φ_{s_0} as the integral of its Mellin transform, yielding

$$\sum_n \Lambda(n)\, \Phi_{s_0}\left(\frac{n}{X}\right) \chi(n) = \frac{1}{2\pi i} \int_{(2)} X^w \Phi^\dagger\left(w + \frac{s_0}{2}\right) \left(-\frac{L'}{L}(w,\chi)\right) dw.$$

Observe that from repeated integration by parts we have

$$\left|\Phi^{\dagger}(\sigma+it+\frac{s_0}{2})\right| \ll_{\sigma,j} (\log X)^j \left(1+\left|t-\frac{\mathrm{Im}(s_0)}{2}\right|\right)^{-j} \tag{3.5.4.2}$$

for every non-negative integer j.

We shift the line of integration to $\mathrm{Re}(w) = -\frac{1}{2}$, picking up residues from all of the zeros in the critical strip. On the line $\mathrm{Re}(w) = -\frac{1}{2}$ we have the bound

$$\left|\frac{L'}{L}(w,\chi)\right| \ll \log(q|w|),$$

and this yields

$$\sum_n \Lambda(n)\,\Phi_{s_0}\left(\frac{n}{X}\right)\chi(n) = \sum_{\substack{L(\rho,\chi)=0\\ 0\le\beta\le1}} X^{\rho}\Phi^{\dagger}\left(\rho+\frac{s_0}{2}\right) + O\left(\frac{(\log X)^{O(1)}}{X^{1/2}}\right).$$

We have written here $\rho = \beta + i\gamma$. The error term is, of course, completely acceptable for (3.5.4.1) when summed over $q \ll Q$.

By (3.5.4.2), the contribution to $\mathcal{E}(Q)$ from those ρ with $|\gamma| > Q^{1/2}$ is

$$\ll XQ^{-100},$$

say, and this gives an acceptable bound. We have therefore obtained

$$\mathcal{E}(Q) \ll X\exp(-\varpi\sqrt{\log X}) + Q^{-\frac{1}{2}+\varepsilon}\sum_{\chi\in S(Q)}\ \sum_{\substack{L(\rho,\chi)=0\\ 0\le\beta\le1\\ |\gamma|\le Q^{1/2}}} X^{\beta}. \tag{3.5.4.3}$$

In order to bound the right side of (3.5.4.3), we first need to introduce some notation. For a primitive Dirichlet character χ modulo q, let $N(T,\chi)$ denote the number of zeros of $L(s,\chi)$ in the rectangle

$$0 \le \beta \le 1, \qquad |\gamma| \le T.$$

For $T \ge 2$, say, we have [66, Chapter 16]

$$N(T,\chi) \ll T\log(qT). \tag{3.5.4.4}$$

For $\frac{1}{2} \le \alpha \le 1$, define $N(\alpha, T, \chi)$ to be the number of zeros $\rho = \beta + i\gamma$ of $L(s,\chi)$ in the rectangle

$$\alpha \le \beta \le 1, \qquad |\gamma| \le T,$$

and define

$$N(\alpha, Q, T) = \sum_{q \le Q} \sideset{}{^*}\sum_{\chi (\mathrm{mod}\ q)} N(\alpha, T, \chi).$$

In $N(\alpha, Q, T)$ the summation on χ is over primitive characters. We employ Jutila's zero-density estimate [69, (1.7)]

$$N(\alpha, Q, T) \ll (QT)^{4(1-\alpha)+\varepsilon}, \tag{3.5.4.5}$$

which holds for $\alpha \ge \frac{4}{5}$.

In (3.5.4.3), we separate the zeros ρ according to whether $\beta < \frac{4}{5}$ or $\beta \ge \frac{4}{5}$. Using (3.5.4.4) we deduce

$$Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \sum_{\substack{L(\rho,\chi)=0 \\ 0 \le \beta < 4/5 \\ |\gamma| \le Q^{1/2}}} X^{\beta} \ll X^{\frac{4}{5}} Q^{1+\varepsilon}. \tag{3.5.4.6}$$

For those zeros with $\beta \ge \frac{4}{5}$ we write

$$X^{\beta} = X^{4/5} + (\log X) \int_{4/5}^{\beta} X^{\alpha} d\alpha.$$

We then embed $S(Q)$ into the set of all primitive characters with conductors $\le Q$. Applying (3.5.4.6) and (3.5.4.5), we obtain

$$\sum_{\chi \in S(Q)} \sum_{\substack{L(\rho,\chi_q)=0 \\ 4/5 \le \beta \le 1 \\ |\gamma| \le Q^{\frac{1}{2}}}} X^{\beta} \ll X^{\frac{4}{5}} Q^{1+\varepsilon} + (\log X) \int_{4/5}^{1} X^{\alpha} N(\alpha, Q, Q^{\frac{1}{2}}) d\alpha$$

$$\ll X^{\frac{4}{5}} Q^{1+\varepsilon} + Q^{\varepsilon} \int_{4/5}^{1} X^{\alpha} Q^{6(1-\alpha)} d\alpha.$$

Since $Q \ll X^{\frac{1}{10}}$ the integrand of this latter integral is maximized when $\alpha = 1$. It follows that

$$Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \sum_{\substack{L(\rho,\chi_q)=0 \\ 4/5 \leq \beta \leq 1 \\ |\gamma| \leq Q^{\frac{1}{2}}}} X^{\beta} \ll X^{\frac{4}{5}} Q^{1+\varepsilon} + XQ^{-\frac{1}{2}+\varepsilon} \ll XQ^{-\frac{1}{2}+\varepsilon}. \tag{3.5.4.7}$$

Combining (3.5.4.7) and (3.5.4.6) yields

$$\mathcal{E}(Q) \ll XQ^{-\frac{1}{2}+\varepsilon},$$

and this suffices for (3.5.4.1).

3.5.5 Regime III

In Regime III we have $X^{\frac{1}{10}} \ll Q \ll MX^{\frac{1}{2}+\varepsilon} = X^{\frac{1}{2}+\theta+\varepsilon}$ (recall (3.2.4)). Here we depart from the philosophy of the previous two regimes, in that we do not bound $\mathcal{E}(Q)$ by considerations of zeros of L-functions. Rather, we exploit the combinatorial structure of the von Mangoldt function and Lemma 3.3.4.

We observe that in Regime III one may still proceed with zero-density estimates by appealing to Heath-Brown's zero-density estimate for L-functions of quadratic characters [67, Theorem 3]. We present our method for the sake of variety, and because it might prove useful in other contexts.

Let us move to our treatment of $\mathcal{E}(Q)$ for these large Q. Given an arithmetic function $f : \mathbb{N} \to \mathbb{C}$ and a real number $W > 1$, let $f_{\leq W}(n)$ denote the arithmetic function

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

We write $f_{>W}(n) = f(n) - f_{\leq W}(n)$.

Our starting place is Vaughan's identity [47, Proposition 13.4]. Given a parameter $V > 1$, we have

$$\begin{aligned} \Lambda(n) &= \Lambda_{\leq V}(n) + (\mu_{\leq V} \star \log)(n) \\ &\quad - (\mu_{\leq V} \star \Lambda_{\leq V} \star 1)(n) + (\mu_{>V} \star \Lambda_{>V} \star 1)(n). \end{aligned} \tag{3.5.5.1}$$

We apply (3.5.5.1) for $n \asymp X$, and we set $V := X^{\frac{1}{3}(\frac{1}{2}-\theta)}$. This reduces the estimation of $\mathcal{E}(Q)$ to the estimation of three different sums, say $\mathcal{E}_i(Q)$, for $i \in \{1,2,3\}$. Observe that there are four terms on the right side of (3.5.5.1), but $\Lambda_{\leq V}(n)$ is identically zero for $n \asymp X$.

We have

$$\begin{aligned}\mathcal{E}_1(Q) &:= Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \left| \sum_n (\mu_{\leq V} \star \log)(n) \Phi_{s_0}\left(\frac{n}{X}\right) \chi(n) \right| \\ &\ll Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \sum_{v \leq V} \mu^2(v) \left| \sum_m (\log m) \Phi_{s_0}\left(\frac{mv}{X}\right) \chi(m) \right|.\end{aligned}$$

By partial summation and the Pólya-Vinogradov inequality, we find that

$$\mathcal{E}_1(Q) \ll Q^{1+\varepsilon} V \ll X^{\frac{1}{2}+\theta+\frac{1}{3}(\frac{1}{2}-\theta)+\varepsilon} \ll X^{1-\varepsilon}, \tag{3.5.5.2}$$

the last inequality holding for $\varepsilon = \varepsilon(\theta) > 0$ sufficiently small.

The estimation of $\mathcal{E}_2(Q)$ is entirely similar, and we obtain

$$\begin{aligned}\mathcal{E}_2(Q) &:= Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \left| \sum_n (\mu_{\leq V} \star \Lambda_{\leq V} \star 1)(n) \Phi_{s_0}\left(\frac{n}{X}\right) \chi(n) \right| \\ &\ll Q^{1+\varepsilon} V^2 \ll X^{\frac{1}{2}+\theta+\frac{2}{3}(\frac{1}{2}-\theta)+\varepsilon} \ll X^{1-\varepsilon}.\end{aligned} \tag{3.5.5.3}$$

The last sum to estimate is $\mathcal{E}_3(Q)$:

$$\begin{aligned}\mathcal{E}_3(Q) &:= Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \left| \sum_n (\mu_{>V} \star \Lambda_{>V} \star 1)(n) \Phi_{s_0}\left(\frac{n}{X}\right) \chi(n) \right| \\ &= Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \left| \sum\sum_{k,\ell} \alpha(k)\beta(\ell) \Phi_{s_0}\left(\frac{k\ell}{X}\right) \chi(k\ell) \right|,\end{aligned}$$

where $\alpha(k) = \mu_{>V}(k)$ and $\beta(\ell) = (\Lambda_{>V} \star 1)(\ell)$. Observe that both $\alpha(\cdot)$ and $\beta(\cdot)$ are supported on integers m satisfying

$$V \ll m \ll XV^{-1}.$$

We further observe that $|\alpha(k)| \leq 1, |\beta(\ell)| \leq \log(\ell)$. We perform dyadic decompositions on the ranges of k and ℓ, so that $k \asymp K, \ell \asymp L$, with

$$V \ll K \ll XV^{-1}, \qquad V \ll L \ll XV^{-1},$$

and $KL \asymp X$.

We next separate the variables by Mellin inversion on Φ_{s_0}:

$$\mathcal{E}_1(Q) \ll (\log X)^{O(1)} \sup_{K,L} \int_{(0)} \left| \Phi^{\dagger}_{s_0}\left(w + \frac{s_0}{2}\right) \right| \\ \times Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \left| \sum_{\substack{k \asymp K \\ \ell \asymp L}}\sum \alpha(k)\beta(\ell)(k\ell)^{-w}\chi(k\ell) \right| dw.$$

The integral of $|\Phi^{\dagger}_{s_0}|$ has size $\ll (\log X)^{O(1)}$, so we obtain

$$\mathcal{E}_3(Q) \ll \sup_{K,L}\ Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \left| \sum_{\substack{k \asymp K \\ \ell \asymp L}}\sum \tilde{\alpha}(k)\tilde{\beta}(\ell)\chi(k\ell) \right|,$$

where $\tilde{\alpha}, \tilde{\beta}$ are complex sequences with $|\tilde{\alpha}(k)| = |\alpha(k)|, |\tilde{\beta}(\ell)| = |\beta(\ell)|$ for all k, ℓ.

By multiplicativity and Cauchy-Schwarz we obtain

$$\mathcal{E}_3(Q) \ll \sup_{K,L}\ Q^{-\frac{1}{2}+\varepsilon} \left(\sum_{\chi \in S(Q)} \left| \sum_{k \asymp K} \tilde{\alpha}(k)\chi(k) \right|^2 \right)^{\frac{1}{2}} \left(\sum_{\chi \in S(Q)} \left| \sum_{\ell \asymp L} \tilde{\beta}(\ell)\chi(\ell) \right|^2 \right)^{\frac{1}{2}}.$$

Applying Lemma 3.3.4 yields

$$\begin{aligned} \mathcal{E}_3(Q) &\ll \sup_{K,L}\ \frac{X^{\varepsilon}}{Q^{\frac{1}{2}}} \left((Q+K)K\right)^{\frac{1}{2}} \left((Q+L)L\right)^{\frac{1}{2}} \\ &\ll \sup_{K,L}\ X^{\varepsilon} \left(X^{\frac{1}{2}}Q^{\frac{1}{2}} + \frac{KL}{K^{\frac{1}{2}}} + \frac{KL}{L^{\frac{1}{2}}} + \frac{KL}{Q^{\frac{1}{2}}} \right) \qquad (3.5.5.4) \\ &\ll X^{\varepsilon} \left(X^{\frac{3}{4}+\frac{\theta}{2}} + \frac{X}{V^{\frac{1}{2}}} + \frac{X}{Q^{\frac{1}{2}}} \right) \ll X^{1-\varepsilon}. \end{aligned}$$

The last inequality follows since $V = X^{\frac{1}{3}(\frac{1}{2}-\theta)}$ and $Q \gg X^{\frac{1}{10}}$. Then (3.5.5.2), (3.5.5.3), and (3.5.5.4) imply

$$\mathcal{E}(Q) \ll X^{1-\varepsilon},$$

and this suffices for (3.5.4.1).

3.5.6 Dénouement

We can extract from our proof of Proposition 3.5.1 the following result on character sums over primes, which we shall have occasion to use later.

Lemma 3.5.2. *Let X be a large real number, and let $\delta > 0$ be small and fixed. Let s_0 be a complex number with $|\mathrm{Re}(s_0)| \leq \frac{A_1}{\log X}$ and $|\mathrm{Im}(s_0)| \leq (\log X)^{A_2}$, for some positive real numbers A_1 and A_2. Given any positive real numbers A_3, A_4, and B, we have*

$$\sum_{\substack{q \leq X^{1-\delta} \\ q \text{ odd} \\ q \neq \square}} \frac{\tau(q)^{A_1}(\log q)^{A_2}}{\sqrt{q}} \left| \sum_{p \equiv 1 \,(\mathrm{mod}\ 8)} (\log p) \Phi_{s_0}\left(\frac{p}{X}\right)\left(\frac{q}{p}\right) \right| \ll_{A_1,A_2,A_3,A_4,B,\delta} \frac{X}{(\log X)^B}.$$

The implied constant is ineffective.

Proof. Follow the proof of (3.5.2.7), but instead use the lower bound $q^* > c(D)(\log X)^D$, which holds for arbitrary $D > 0$. The constant $c(D)$ is ineffective if $D \geq 2$. □

Lemma 3.5.2 is quite strong since it corresponds, roughly, to square root cancellation on average in the sums over p. Thus, one would not expect to be able to prove an analogue of Lemma 3.5.2 with the upper bound for q replaced by $X^{1+\varepsilon}$ for any $\varepsilon > 0$.

3.6 The mollified second moment

In this section we derive an upper bound of the correct order of magnitude for the sum S_2 defined in (3.2.1). Our main result for this section is the following (recall (3.2.4) and (3.4.2)).

Proposition 3.6.1. *Let $\delta > 0$ be small and fixed, and let θ, ϑ satisfy $\theta + 2\vartheta < \frac{1}{2}$. If $X \geq X_0(\delta, \theta, \vartheta)$, then*

$$S_2 \leq \frac{1+\delta}{2(1-\frac{1}{\sqrt{2}})^2} \frac{\mathfrak{I}}{\vartheta} \frac{X}{4},$$

where

$$\mathfrak{I} = -2\int_0^1 H(x)H'(x)dx + \frac{1}{\theta}\int_0^1 H(x)H''(x)dx + \frac{1}{\theta}\int_0^1 H'(x)^2dx \\ - \frac{1}{2\theta^2}\int_0^1 H'(x)H''(x)dx + \frac{1}{24\theta^3}\int_0^1 H''(x)^2dx.$$

The proof of Proposition 3.6.1 follows the ideas outlined in Section 3.2. First, we note that $\log p \leq \log X$ in (3.2.1) because Φ is supported on $[\frac{1}{2}, 1]$. By positivity we may apply the upper bound sieve condition (3.4.7) to write

$$S_2 \leq (\log X)S^+,$$

where S^+ is defined by

$$S^+ = \sum_{n\equiv 1 \,(\mathrm{mod}\ 8)} \mu^2(n)\left(\sum_{\substack{d|n \\ d\leq D}} \lambda_d\right)\Phi\left(\frac{n}{X}\right)L(\tfrac{1}{2}, \chi_n)^2 M(n)^2. \tag{3.6.1}$$

Note that d is odd since $d \mid n$ and $n \equiv 1 \pmod 8$. Also, $\lambda_d \neq 0$ only for square-free d by the definition (3.4.8), and so $\lambda_d = \mu^2(d)\lambda_d$. We use Lemma 3.3.2 to write $L(\frac{1}{2}, \chi_n)^2 = \mathcal{D}_2(n)$, then insert (3.2.6) into (3.6.1) to write

$$S^+ = S_N^+ + S_R^+, \tag{3.6.2}$$

where

$$S_N^+ = \sum_{n\equiv 1 \,(\mathrm{mod}\ 8)} N_Y(n)\left(\sum_{\substack{d|n \\ d\leq D}} \mu^2(d)\lambda_d\right)\Phi\left(\frac{n}{X}\right)\mathcal{D}_2(n)M(n)^2 \tag{3.6.3}$$

and

$$S_R^+ = \sum_{n\equiv 1 \,(\mathrm{mod}\ 8)} R_Y(n)\left(\sum_{\substack{d|n \\ d\leq D}} \mu^2(d)\lambda_d\right)\Phi\left(\frac{n}{X}\right)\mathcal{D}_2(n)M(n)^2$$

We first obtain a bound on S_R^+. The remainder of this section will then be devoted to an analysis of S_N^+.

3.6.1 The contribution of S_R^+

In this subsection we show

$$S_R^+ \ll X^{\varepsilon}\left(\frac{X}{Y} + X^{1/2}M\right). \tag{3.6.1.1}$$

The arguments here are almost identical to those in [51, Section 3]. Observe that $R_Y(n) = 0$ unless $n = \ell^2 h$ with $\ell > Y$ and h square-free. If $n \equiv 1 \pmod 8$ then ℓ and h are odd and $h \equiv 1 \pmod 8$. By the divisor bound we have

$$|R_Y(n)| \ll n^{\varepsilon}, \qquad \left|\sum_{\substack{d|n \\ d \le D}} \mu^2(d)\lambda_d\right| \ll n^{\varepsilon},$$

and therefore

$$S_R^+ \ll X^{\varepsilon} \sum_{\substack{Y < \ell \le \sqrt{X} \\ 2 \nmid \ell}} \sum_{\substack{X/2\ell^2 < h \le X/\ell^2 \\ h \equiv 1 \,(\mathrm{mod}\ 8)}} \mu^2(h)|M(\ell^2 h)^2 \mathcal{D}_2(\ell^2 h)|.$$

There is a mild complication compared to [51] in that it is possible to have $h = 1$, in which case the character χ_h is principal.

We apply Cauchy-Schwarz and obtain

$$S_R^+ \ll X^{\varepsilon} \sum_{\substack{Y < \ell \le \sqrt{X} \\ 2 \nmid \ell}} \left(\sum_{\substack{X/2\ell^2 < h \le X/\ell^2 \\ h \equiv 1 \,(\mathrm{mod}\ 8)}} \mu^2(h)|M(\ell^2 h)^2|^2\right)^{1/2} \tag{3.6.1.2}$$

$$\times \left(\sum_{\substack{X/2\ell^2 < h \le X/\ell^2 \\ h \equiv 1 \,(\mathrm{mod}\ 8)}} \mu^2(h)|\mathcal{D}_2(\ell^2 h)|^2\right)^{1/2}. \tag{3.6.1.3}$$

We have

$$M(\ell^2 h)^2 = \sum_{\substack{m \le M^2 \\ (m, 2\ell) = 1}} \frac{\alpha(m)}{\sqrt{m}} \left(\frac{h}{m}\right)$$

for some coefficients $\alpha(m)$ satisfying $|\alpha(m)| \ll m^{\varepsilon}$. For $h = 1$ we use the trivial bound

$M(\ell^2)^4 \ll M^2 X^\varepsilon$. For $h > 1$ we use Lemma 3.3.4. We therefore have

$$\sum_{\substack{X/2\ell^2 < h \leq X/\ell^2 \\ h \equiv 1 \ (\mathrm{mod}\ 8)}} \mu^2(h) |M(\ell^2 h)^2|^2 \ll X^\varepsilon \left(\frac{X}{\ell^2} + M^2 \right). \tag{3.6.1.4}$$

Now observe that, for any $c > \frac{1}{2}$,

$$\mathcal{D}_2(\ell^2 h) = \frac{2}{(1 - \frac{1}{\sqrt{2}})^4} \frac{1}{2\pi i} \int_{(c)} \frac{\Gamma^2 \left(\frac{s}{2} + \frac{1}{4} \right)}{\Gamma^2 \left(\frac{1}{4} \right)} \left(1 - \frac{1}{2^{1/2-s}} \right)^2 \times \left(\frac{\ell^2 h}{\pi} \right)^s L^2 \left(\frac{1}{2} + s, \chi_h \right) \mathcal{E}(s, 2\ell) \frac{ds}{s},$$

where

$$\mathcal{E}(s,k) = \prod_{p|k} \left(1 - \frac{\chi_h(p)}{p^{1/2+s}} \right)^s.$$

If $h = 1$ then $L^2 \left(\frac{1}{2} + s, \chi_h \right) = \zeta^2(\frac{1}{2} + s)$. In any case, we move the line of integration to $c = \frac{1}{\log X}$, and we do not pick up contributions from any poles. When $h > 1$ this is obvious, and when $h = 1$ the double pole of $\zeta^2(\frac{1}{2} + s)$ is canceled out by the double zero of $(1 - 2^{-(1/2-s)})^2$. By trivial estimation we have then $|\mathcal{D}_2(\ell^2)| \ll X^\varepsilon$. For $h > 1$ we apply Cauchy-Schwarz to obtain

$$|\mathcal{D}_2(\ell^2 h)|^2 \ll X^\varepsilon \int_{(\frac{1}{\log X})} \left| \Gamma \left(\frac{s}{2} + \frac{1}{4} \right) \right|^2 \left| L \left(\frac{1}{2} + s, \chi_h \right) \right|^4 |ds|.$$

Summing over h and using Lemma 3.3.5, we obtain

$$\sum_{\substack{X/2\ell^2 < h \leq X/\ell^2 \\ h \equiv 1 \ (\mathrm{mod}\ 8)}} \mu^2(h) |\mathcal{D}_2(\ell^2 h)|^2 \ll \frac{X^{1+\varepsilon}}{\ell^2}. \tag{3.6.1.5}$$

Combining (3.6.1.2), (3.6.1.4), and (3.6.1.5) yields (3.6.1.1).

3.6.2 Poisson summation

We begin our evaluation of S_N^+ by inserting into (3.6.3) the definition (3.2.3) of the mollifier $M(n)$. We then use the definition of $\mathcal{D}_2$ (see Lemma 3.3.2) to write

$$S_N^+ = \frac{8}{(\sqrt{2}-1)^4} \sum_{\substack{d\le D \\ d \text{ odd}}} \mu^2(d)\lambda_d \sum\sum_{\substack{m_1,m_2\le M \\ m_1,m_2 \text{ odd}}} \frac{b_{m_1}b_{m_2}}{\sqrt{m_1m_2}} \sum_{\substack{n\equiv 1 \ (\mathrm{mod}\ 8) \\ d|n}} N_Y(n)\Phi\left(\frac{n}{X}\right) \times \sum_{\substack{\nu=1 \\ \nu \text{ odd}}}^{\infty} \frac{d_2(\nu)}{\sqrt{\nu}}\omega_2\left(\frac{\nu\pi}{n}\right)\left(\frac{n}{m_1m_2\nu}\right). \tag{3.6.2.1}$$

We next apply Poisson summation to evaluate the n-sum. Denote the n-sum in (3.6.2.1) by Z, i.e. define Z by

$$Z = Z(d,\nu,m_1m_2;X,Y) = \sum_{\substack{n\equiv 1 \ (\mathrm{mod}\ 8) \\ d|n}} N_Y(n)\Phi\left(\frac{n}{X}\right)\omega_2\left(\frac{\nu\pi}{n}\right)\left(\frac{n}{m_1m_2\nu}\right). \tag{3.6.2.2}$$

We insert the definition (3.2.7) of $N_Y(n)$ and interchange the order of summation to write Z as

$$Z = \sum_{\substack{\alpha\le Y \\ \alpha \text{ odd}}} \mu(\alpha) \sum_{\substack{n\equiv 1 \ (\mathrm{mod}\ 8) \\ [\alpha^2,d]|n}} F_\nu\left(\frac{n}{X}\right)\left(\frac{n}{m_1m_2\nu}\right), \tag{3.6.2.3}$$

where $F_\nu(t)$ is defined by

$$F_\nu(t) = \Phi(t)\omega_2\left(\frac{\nu\pi}{tX}\right). \tag{3.6.2.4}$$

If α and d are square-free, then $[\alpha^2,d] = \alpha^2 d_1$, where

$$d_1 = \frac{d}{(d,\alpha)}. \tag{3.6.2.5}$$

We may thus relabel n as $\alpha^2 d_1 m$ in (3.6.2.3), and then split the resulting sum on m according to the congruence class of $m \pmod{m_1m_2\nu}$. We deduce from (3.6.2.3) that

$$Z = \sum_{\substack{\alpha\le Y \\ (\alpha,2m_1m_2\nu)=1}} \mu(\alpha)\left(\frac{d_1}{m_1m_2\nu}\right) \sum_{b \pmod{m_1m_2\nu}} \left(\frac{b}{m_1m_2\nu}\right) \times \sum_{\substack{m\equiv \overline{\alpha^2 d_1} \ (\mathrm{mod}\ 8) \\ m\equiv b \ (\mathrm{mod}\ m_1m_2\nu)}} F_\nu\left(\frac{\alpha^2 d_1 m}{X}\right).$$

By the Chinese Remainder Theorem, we may write the congruence conditions on m as a single condition $m \equiv \gamma \pmod{8m_1m_2\nu}$ for some integer γ depending on α, d, b. Thus, we may relabel m as $8jm_1m_2\nu + \gamma$, where j ranges over all integers, and arrive at

$$\begin{aligned} Z = & \sum_{\substack{\alpha \le Y \\ (\alpha, 2m_1m_2\nu)=1}} \mu(\alpha) \left(\frac{d_1}{m_1m_2\nu} \right) \sum_{b \pmod{m_1m_2\nu}} \left(\frac{b}{m_1m_2\nu} \right) \\ & \times \sum_{j\in\mathbb{Z}} F_\nu \left(\frac{\alpha^2 d_1 (8jm_1m_2\nu + \gamma)}{X} \right). \end{aligned} \tag{3.6.2.6}$$

We apply Poisson summation to the j-sum to write

$$\sum_{j\in\mathbb{Z}} F_\nu \left(\frac{\alpha^2 d_1 (8jm_1m_2\nu + \gamma)}{X} \right) = \frac{X}{8\alpha^2 d_1 m_1m_2\nu} \sum_{k\in\mathbb{Z}} e\left(\frac{k\gamma}{8m_1m_2\nu} \right) \hat{F}_\nu \left(\frac{kX}{8\alpha^2 d_1 m_1m_2\nu} \right).$$

We insert this into (3.6.2.6), apply the reciprocity relation

$$e\left(\frac{k\gamma}{8m_1m_2\nu} \right) = e\left(\frac{k\overline{8}b}{m_1m_2\nu} \right) e\left(\frac{k\overline{\alpha^2 d_1 m_1m_2\nu}}{8} \right),$$

and then evaluate the b-sum using the definition (3.3.3) of the Gauss sum. Therefore

$$\begin{aligned} Z = & \frac{X}{8m_1m_2\nu} \sum_{\substack{\alpha \le Y \\ (\alpha, 2m_1m_2\nu)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \left(\frac{2d_1}{m_1m_2\nu} \right) \\ & \times \sum_{k\in\mathbb{Z}} e\left(\frac{k\overline{\alpha^2 d_1 m_1m_2\nu}}{8} \right) \hat{F}_\nu \left(\frac{kX}{8\alpha^2 d_1 m_1m_2\nu} \right) \tau_k(m_1m_2\nu). \end{aligned}$$

Recalling (3.6.2.1) and (3.6.2.2), we arrive at

$$\begin{aligned} S_N^+ = & \frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \le D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \le M \\ (m_1m_2, 2d)=1}}\sum \frac{b_{m_1} b_{m_2}}{(m_1m_2)^{3/2}} \sum_{\substack{\nu=1 \\ (\nu, 2d)=1}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2}} \sum_{\substack{\alpha \le Y \\ (\alpha, 2m_1m_2\nu)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \\ & \times \left(\frac{2d_1}{m_1m_2\nu} \right) \sum_{k\in\mathbb{Z}} e\left(\frac{k\overline{\alpha^2 d_1 m_1m_2\nu}}{8} \right) \hat{F}_\nu \left(\frac{kX}{8\alpha^2 d_1 m_1m_2\nu} \right) \tau_k(m_1m_2\nu). \end{aligned} \tag{3.6.2.7}$$

Note that we may impose the condition $(m_1m_2\nu, d) = 1$ because otherwise $(\frac{2d_1}{m_1m_2\nu}) = 0$.

We write (3.6.2.7) as

$$S_N^+ = \mathcal{T}_0 + \mathcal{B}, \tag{3.6.2.8}$$

where $\mathcal{T}_0$ is the contribution from $k = 0$ in (3.6.2.7), while $\mathcal{B}$ is the contribution from $k \neq 0$ in (3.6.2.7). We evaluate $\mathcal{T}_0$ in the next subsection, and $\mathcal{B}$ in later subsections.

3.6.3 The contribution from $k = 0$

By (3.3.3), $\tau_0(n) = \varphi(n)$ if n is a perfect square, and $\tau_0(n) = 0$ otherwise. Hence the term $\mathcal{T}_0$ in (3.6.2.7) is

$$\mathcal{T}_0 = \frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \le D \\ d \text{ odd}}} \mu^2(d)\lambda_d \sum\sum_{\substack{m_1,m_2 \le M \\ (m_1m_2,2d)=1}} \frac{b_{m_1}b_{m_2}}{(m_1m_2)^{3/2}} \sum_{\substack{\nu=1 \\ (\nu,2d)=1 \\ m_1m_2\nu=\square}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2}} \sum_{\substack{\alpha \le Y \\ (\alpha,2m_1m_2\nu)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \times \hat{F}_\nu(0)\varphi(m_1m_2\nu). \tag{3.6.3.1}$$

We first extend the sum over α to infinity. Since $\varphi(n) \le n$, the error introduced in doing so is

$$\ll X \sum_{d \le D} |\lambda_d| \sum\sum_{m_1,m_2 \le M} \frac{|b_{m_1}b_{m_2}|}{\sqrt{m_1m_2}} \sum_{\substack{\nu=1 \\ m_1m_2\nu=\square}}^{\infty} \frac{d_2(\nu)}{\sqrt{\nu}} \sum_{\alpha > Y} \frac{1}{\alpha^2 d_1} |\hat{F}_\nu(0)|. \tag{3.6.3.2}$$

By Lemma 3.3.1, $\hat{F}_\nu(0) \ll 1$ uniformly for all $\nu > 0$, and $\hat{F}_\nu(0) \ll \exp(-\frac{\pi\nu}{8X})$ for $\nu > X^{1+\varepsilon}$. Moreover, (3.4.8) implies that $|\lambda_d| \ll d^\varepsilon$, while $|b_m| \ll 1$ by (3.2.5). It follows from these bounds that (3.6.3.2) is

$$\ll X^{1+\varepsilon} \sum_{d \le D} \sum\sum_{m_1,m_2 \le M} \frac{1}{\sqrt{m_1m_2}} \sum_{\substack{\nu \le X^{1+\varepsilon} \\ m_1m_2\nu=\square}} \frac{1}{\sqrt{\nu}} \sum_{\alpha > Y} \frac{1}{\alpha^2 d_1} + \exp\left(-X^\varepsilon\right). \tag{3.6.3.3}$$

Since $m_1m_2\nu$ is a perfect square, the sum over m_1, m_2, ν in (3.6.3.3) is $\ll X^\varepsilon$. Also, the definition (3.6.2.5) of d_1 implies that

$$\sum_{\alpha > Y} \frac{1}{\alpha^2 d_1} = \frac{1}{d} \sum_{j|d} \varphi(j) \sum_{\substack{\alpha > Y \\ j|\alpha}} \frac{1}{\alpha^2} \ll \frac{1}{d^{1-\varepsilon}Y}.$$

Therefore (3.6.3.3) is $O(X^{1+\varepsilon}/Y)$. This bounds the error in extending the sum over α in (3.6.3.1) to infinity, and we arrive at

$$\mathcal{T}_0 = \frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \le D \\ d \text{ odd}}} \mu^2(d)\lambda_d \sum\sum_{\substack{m_1, m_2 \le M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} \sum_{\substack{\nu=1 \\ (\nu,2d)=1 \\ m_1 m_2 \nu = \square}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2}} \sum_{\substack{\alpha=1 \\ (\alpha, 2m_1 m_2 \nu)=1}}^{\infty} \frac{\mu(\alpha)}{\alpha^2 d_1} \times \hat{F}_\nu(0) \varphi(m_1 m_2 \nu) + O\left(\frac{X^{1+\varepsilon}}{Y}\right).$$

Writing the sum on α as an Euler product, we deduce that

$$\mathcal{T}_0 = \frac{4X}{3(\sqrt{2}-1)^4 \zeta(2)} \sum_{\substack{d \le D \\ d \text{ odd}}} \frac{\mu^2(d)\lambda_d}{d} \prod_{p \mid d} \left(\frac{p}{p+1}\right) \sum\sum_{\substack{m_1, m_2 \le M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{\sqrt{m_1 m_2}} \times \sum_{\substack{\nu=1 \\ (\nu,2d)=1 \\ m_1 m_2 \nu = \square}}^{\infty} \frac{d_2(\nu)}{\sqrt{\nu}} \hat{F}_\nu(0) \prod_{p \mid m_1 m_2 \nu} \left(\frac{p}{p+1}\right) + O\left(\frac{X^{1+\varepsilon}}{Y}\right). \tag{3.6.3.4}$$

We next evaluate the sum over d. Lemma 3.4.3 implies

$$\sum_{\substack{d \le D \\ (d, 2m_1 m_2 \nu)=1}} \frac{\mu^2(d)\lambda_d}{d} \prod_{p \mid d} \left(\frac{p}{p+1}\right) = \frac{1+E_0(X)}{\log R} \prod_{\substack{p \mid 2m_1 m_2 \nu \\ p \le z_0}} \left(1 + \frac{1}{p}\right) \prod_{p \le z_0} \left(\frac{p^2}{p^2-1}\right) + O\left((\log R)^{-2019}\right). \tag{3.6.3.5}$$

Recall that $E_0(X) \to 0$, and depends only on X, G, and ϑ. Heretofore we just write $o(1)$ instead of $E_0(X)$.

We may omit the condition $p \le z_0$ by trivial estimation and (3.4.1). It follows from (3.6.3.5) and (3.6.3.4) that

$$\mathcal{T}_0 = \frac{2X}{(\sqrt{2}-1)^4} \frac{1+o(1)}{\log R} \sum\sum_{\substack{m_1, m_2 \le M \\ (m_1 m_2, 2)=1}} \frac{b_{m_1} b_{m_2}}{\sqrt{m_1 m_2}} \sum_{\substack{\nu=1 \\ (\nu,2)=1 \\ m_1 m_2 \nu = \square}}^{\infty} \frac{d_2(\nu)}{\sqrt{\nu}} \hat{F}_\nu(0) + O\left(\frac{X}{(\log R)^{2019}} + \frac{X^{1+\varepsilon}}{Y}\right). \tag{3.6.3.6}$$

The next task is to carry out the summation over m_1, m_2, and ν. Let Υ_0 be defined

by

$$\Upsilon_0 = \sum\sum_{\substack{m_1,m_2\le M\\(m_1m_2,2)=1}} \frac{b_{m_1}b_{m_2}}{\sqrt{m_1m_2}} \sum_{\substack{\nu=1\\(\nu,2)=1\\m_1m_2\nu=\square}}^{\infty} \frac{d_2(\nu)}{\sqrt{\nu}}\hat{F}_\nu(0). \tag{3.6.3.7}$$

We insert into (3.6.3.7) the definition (3.2.5) of b_m and the definitions (3.6.2.4) and (3.3.1) of F_ν and ω_2, and then apply the Fourier inversion formula (3.5.1.2). After interchanging the order of summation, we arrive at

$$\Upsilon_0 = \frac{1}{2\pi i}\int_{(c)} \frac{\Gamma(\frac{s}{2}+\frac14)^2}{\Gamma(\frac14)^2}\left(1-\frac{1}{2^{\frac12-s}}\right)^2\left(\frac{X}{\pi}\right)^s \check{\Phi}(s)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(z_1)h(z_2)$$
$$\times \sum\sum\sum_{\substack{(m_1m_2\nu,2)=1\\m_1m_2\nu=\square}} \frac{\mu(m_1)\mu(m_2)d_2(\nu)}{(m_1m_2\nu)^{\frac12} m_1^{\frac{1+iz_1}{\log M}} m_2^{\frac{1+iz_2}{\log M}}\nu^s}\, dz_1dz_2\frac{ds}{s}, \tag{3.6.3.8}$$

where we take $c = \frac{1}{\log X}$ to facilitate later estimations. We may write the sum on m_1, m_2, ν as an Euler product

$$\sum\sum\sum_{\substack{(m_1m_2\nu,2)=1\\m_1m_2\nu=\square}} \frac{\mu(m_1)\mu(m_2)d_2(\nu)}{(m_1m_2\nu)^{\frac12} m_1^{\frac{1+iz_1}{\log M}} m_2^{\frac{1+iz_2}{\log M}}\nu^s}$$
$$\times = \prod_{p>2}\sum_{\substack{m_1=0\\m_1+m_2+\nu\text{ even}}}^{1}\sum_{m_2=0}^{1}\sum_{\nu=0}^{\infty} \frac{(-1)^{m_1+m_2}(\nu+1)}{p^{\frac{m_1+m_2+\nu}{2}+m_1\left(\frac{1+iz_1}{\log M}\right)+m_2\left(\frac{1+iz_2}{\log M}\right)+\nu s}}.$$

This can also be written as

$$\zeta^3(1+2s)\zeta\left(1+\tfrac{2+iz_1+iz_2}{\log M}\right)\zeta^{-2}\left(1+\tfrac{1+iz_1}{\log M}+s\right)\zeta^{-2}\left(1+\tfrac{1+iz_2}{\log M}+s\right)Q\left(\tfrac{1+iz_1}{\log M},\tfrac{1+iz_2}{\log M},s\right), \tag{3.6.3.9}$$

where $Q(w_1, w_2, s)$ is an Euler product that is uniformly bounded and holomorphic when each of $\mathrm{Re}(w_1)$, $\mathrm{Re}(w_2)$, and $\mathrm{Re}(s)$ is $\ge -\varepsilon$. From this definition of Q and a calculation, we see that

$$Q(0,0,0) = 1, \tag{3.6.3.10}$$

a fact we use shortly. We insert the expression (3.6.3.9) for the m_1, m_2, ν-sum into

(3.6.3.8) and arrive at

$$\Upsilon_0 = \frac{1}{2\pi i}\int_{(c)} \frac{\Gamma(\frac{s}{2}+\frac{1}{4})^2}{\Gamma(\frac{1}{4})^2}\left(1-\frac{1}{2^{\frac{1}{2}-s}}\right)^2\left(\frac{X}{\pi}\right)^s \check{\Phi}(s)\zeta^3(1+2s)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(z_1)h(z_2)$$
$$\times\, \zeta\Big(1+\tfrac{2+iz_1+iz_2}{\log M}\Big)\zeta^{-2}\Big(1+\tfrac{1+iz_1}{\log M}+s\Big)\zeta^{-2}\Big(1+\tfrac{1+iz_2}{\log M}+s\Big)Q\left(\tfrac{1+iz_1}{\log M},\tfrac{1+iz_2}{\log M},s\right)\, dz_1dz_2\frac{ds}{s}.$$

By (3.5.1.4) and the rapid decay of the gamma function, we may truncate the integrals to the region $|z_1|, |z_2| \le \sqrt{\log M}$ and $|\mathrm{Im}(s)| \le (\log X)^2$, introducing a negligible error. We then deform the path of integration of the s-integral to the path made up of the line segment L_1 from $\frac{1}{\log X} - i(\log X)^2$ to $-\frac{c'}{\log\log X} - i(\log X)^2$, followed by the line segment L_2 from $-\frac{c'}{\log\log X} - i(\log X)^2$ to $-\frac{c'}{\log\log X} + i(\log X)^2$, and then by the line segment L_3 from $-\frac{c'}{\log\log X} + i(\log X)^2$ to $\frac{1}{\log X} + i(\log X)^2$, where c' is a constant chosen so that

$$\zeta(1+z) \ll \log|\mathrm{Im}(z)| \quad \text{and} \quad \frac{1}{\zeta(1+z)} \ll \log|\mathrm{Im}(z)| \tag{3.6.3.11}$$

for $\mathrm{Re}(z) \ge -c'/\log|\mathrm{Im}(z)|$ and $|\mathrm{Im}(z)| \ge 1$ (see, for example, Theorem 3.5 and (3.11.8) of Titchmarsh [70]). This leaves a residue from the pole at $s = 0$. The contributions of the integrals over L_1 and L_3 are negligible because of the rapid decay of the Γ function, while the contribution of the integral over L_2 is negligible because $X^s \ll \exp\left(-c'\frac{\log X}{\log\log X}\right)$ for s on L_2. Hence the main contribution arises from the residue of the pole at $s = 0$. Writing this residue as an integral along a circle centered at 0, we arrive at

$$\Upsilon_0 = \frac{1}{2\pi i}\oint_{|s|=\frac{1}{\log X}} \frac{\Gamma(\frac{s}{2}+\frac{1}{4})^2}{\Gamma(\frac{1}{4})^2}\left(1-\frac{1}{2^{\frac{1}{2}-s}}\right)^2\left(\frac{X}{\pi}\right)^s \check{\Phi}(s)\zeta^3(1+2s)$$
$$\times \iint\limits_{|z_i|\le\sqrt{\log M}} h(z_1)h(z_2)\zeta\Big(1+\tfrac{2+iz_1+iz_2}{\log M}\Big)\zeta^{-2}\Big(1+\tfrac{1+iz_1}{\log M}+s\Big)\zeta^{-2}\Big(1+\tfrac{1+iz_2}{\log M}+s\Big)$$
$$\times Q\left(\tfrac{1+iz_1}{\log M},\tfrac{1+iz_2}{\log M},s\right)\, dz_1dz_2\frac{ds}{s} + O\left(\frac{1}{(\log X)^{2019}}\right).$$

We may expand the zeta-functions and the function Q into Laurent series. The main contribution arises from the first terms of the Laurent expansions, and so we deduce using (3.6.3.10) that

$$\Upsilon_0 = \frac{1}{16\pi i}\oint_{|s|=\frac{1}{\log X}} \frac{\Gamma(\frac{s}{2}+\frac14)^2}{\Gamma(\frac14)^2}\left(1-\frac{1}{2^{\frac12 - s}}\right)^2 \left(\frac{X}{\pi}\right)^s \check{\Phi}(s) \int\int_{|z_i|\le\sqrt{\log M}} h(z_1)h(z_2)$$
$$\times \left(\frac{\log M}{2+iz_1+iz_2}\right)\left(\frac{1+iz_1}{\log M}+s\right)^2\left(\frac{1+iz_2}{\log M}+s\right)^2 dz_1dz_2\frac{ds}{s^4} + O\left(\frac{1}{(\log X)^{1-\varepsilon}}\right).$$

By (3.5.1.4), we may extend the integrals over z_1, z_2 to $\mathbb{R}^2$, introducing a negligible error. We then apply the formula

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(z_1)h(z_2)\frac{(1+iz_1)^j(1+iz_2)^k}{2+iz_1+iz_2}\,dz_1dz_2 \\
&= \int_0^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(z_1)h(z_2)(1+iz_1)^j(1+iz_2)^k e^{-t(1+iz_1)-t(1+iz_2)}\,dz_1dz_2dt \qquad (3.6.3.12)\\
&= (-1)^{j+k}\int_0^{\infty} H^{(j)}(t)H^{(k)}(t)\,dt
\end{aligned}$$

to obtain

$$\begin{aligned}
\Upsilon_0 = \frac{1}{16\pi i}\oint_{|s|=\frac{1}{\log X}} &\frac{\Gamma(\frac{s}{2}+\frac14)^2}{\Gamma(\frac14)^2}\left(1-\frac{1}{2^{\frac12 - s}}\right)^2 \left(\frac{X}{\pi}\right)^s \check{\Phi}(s)\Bigg\{\frac{1}{(\log M)^3}\int_0^1 H''(t)^2\,dt \\
&-\frac{4s}{(\log M)^2}\int_0^1 H'(t)H''(t)\,dt + \frac{2s^2}{\log M}\int_0^1 H(t)H''(t)\,dt + \frac{4s^2}{\log M}\int_0^1 H'(t)^2\,dt \\
&-4s^3\int_0^1 H(t)H'(t)\,dt + s^4\log M\int_0^1 H(t)^2\,dt\Bigg\}\frac{ds}{s^4} + O\left(\frac{1}{(\log X)^{1-\varepsilon}}\right).
\end{aligned}$$

We evaluate the s-integral as a residue using (3.5.1.6). The result is

$$\begin{aligned}
\Upsilon_0 = \frac{\check{\Phi}(0)}{8}&\left(1-\frac{1}{\sqrt{2}}\right)^2\Bigg\{\frac16\left(\frac{\log X}{\log M}\right)^3\int_0^1 H''(t)^2\,dt \\
&\times -2\left(\frac{\log X}{\log M}\right)^2\int_0^1 H'(t)H''(t)\,dt + 2\frac{\log X}{\log M}\int_0^1 H(t)H''(t)\,dt \\
&\times +4\frac{\log X}{\log M}\int_0^1 H'(t)^2\,dt - 4\int_0^1 H(t)H'(t)\,dt\Bigg\} + O\left(\frac{1}{(\log X)^{1-\varepsilon}}\right).
\end{aligned}$$

From this, (3.6.3.6), and the definition (3.6.3.7) of Υ_0, we arrive at

$$\begin{aligned}\mathcal{T}_0 = \frac{X}{8\left(1-\frac{1}{\sqrt{2}}\right)^2}\frac{1+o(1)}{\log R}\Bigg\{&\frac{1}{24}\left(\frac{\log X}{\log M}\right)^3\int_0^1 H''(t)^2\,dt \\ &-\frac{1}{2}\left(\frac{\log X}{\log M}\right)^2\int_0^1 H'(t)H''(t)\,dt+\frac{\log X}{2\log M}\int_0^1 H(t)H''(t)\,dt \\ &+\frac{\log X}{\log M}\int_0^1 H'(t)^2\,dt-\int_0^1 H(t)H'(t)\,dt\Bigg\}+O\left(\frac{X}{(\log X)^{1-\varepsilon}}+\frac{X^{1+\varepsilon}}{Y}\right).\end{aligned} \tag{3.6.3.13}$$

3.6.4 The contribution from $k \neq 0$: splitting into cases

Having estimated the term $\mathcal{T}_0$ in (3.6.2.8), we now begin our analysis of $\mathcal{B}$. The analysis of $\mathcal{B}$ is much more complicated than the analysis for $\mathcal{T}_0$.

The behavior of the additive character $e(k\overline{\alpha^2 d_1 m_1 m_2 \nu}/8)$ in (3.6.2.7) depends upon the residue class of k modulo 8. We therefore distinguish the following cases for k: k is odd, $k \equiv 2 \pmod 4$, $k \equiv 4 \pmod 8$, or $k \equiv 0 \pmod 8$. We split our analysis of the sum $\mathcal{B}$ in (3.6.2.8) according to these four cases. For the terms with odd k, we use the identity

$$e\left(\frac{h}{8}\right) = \frac{\sqrt{2}}{2}\left(\frac{2}{h}\right) + \frac{\sqrt{2}}{2}\left(\frac{-2}{h}\right)i, \qquad h \text{ odd},$$

and treat separately the contributions of each term on the right-hand side. Moreover, for the terms with odd k or $k \equiv 2 \pmod 4$, we use the second expression in (3.3.3) for $\tau_k(n)$ and treat separately the contributions of the terms $\left(\frac{1+i}{2}\right)G_k(n)$ and $\left(\frac{-1}{n}\right)\left(\frac{1-i}{2}\right)G_k(n)$. We can treat these two contributions together as one combined sum for the terms with $k \equiv 0, 4 \pmod 8$, because, for those k, the additive character $e(k\overline{\alpha^2 d_1 m_1 m_2 \nu}/8)$ is constant and the conditions $k \equiv 0, 4 \pmod 8$ are invariant with respect to the substitution $k \mapsto -k$. Hence, in view of these considerations, (3.6.2.7), and (3.6.2.8), we write

$$\mathcal{B} = \frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \le D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum\sum_{\substack{m_1, m_2 \le M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} \sum_{\substack{\nu=1 \\ (\nu, 2d)=1}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2}} \sum_{\substack{\alpha \le Y \\ (\alpha, 2m_1 m_2 \nu)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \times (\mathcal{Q}_1 + \mathcal{Q}_2 + \mathcal{Q}_3 + \mathcal{Q}_4 + \mathcal{U}_1 + \mathcal{U}_2 + \mathcal{V} + \mathcal{W}), \tag{3.6.4.1}$$

where

$$\mathcal{Q}_1 = \left(\frac{1+i}{2}\right) \frac{\sqrt{2}}{2} \left(\frac{2d_1}{m_1 m_2 \nu}\right) \sum_{\substack{k \in \mathbb{Z} \\ k \text{ odd}}} \left(\frac{2}{k d_1 m_1 m_2 \nu}\right) \hat{F}_\nu \left(\frac{kX}{8\alpha^2 d_1 m_1 m_2 \nu}\right) G_k(m_1 m_2 \nu), \tag{3.6.4.2}$$

$$\mathcal{Q}_2 = \left(\frac{1-i}{2}\right) \frac{\sqrt{2}}{2} \left(\frac{-2d_1}{m_1 m_2 \nu}\right) \sum_{\substack{k \in \mathbb{Z} \\ k \text{ odd}}} \left(\frac{2}{k d_1 m_1 m_2 \nu}\right) \hat{F}_\nu \left(\frac{kX}{8\alpha^2 d_1 m_1 m_2 \nu}\right) G_k(m_1 m_2 \nu), \tag{3.6.4.3}$$

$$\mathcal{Q}_3 = \left(\frac{1+i}{2}\right) i\frac{\sqrt{2}}{2} \left(\frac{2d_1}{m_1 m_2 \nu}\right) \sum_{\substack{k \in \mathbb{Z} \\ k \text{ odd}}} \left(\frac{-2}{k d_1 m_1 m_2 \nu}\right) \hat{F}_\nu \left(\frac{kX}{8\alpha^2 d_1 m_1 m_2 \nu}\right) G_k(m_1 m_2 \nu), \tag{3.6.4.4}$$

$$\mathcal{Q}_4 = \left(\frac{1-i}{2}\right) i\frac{\sqrt{2}}{2} \left(\frac{-2d_1}{m_1 m_2 \nu}\right) \sum_{\substack{k \in \mathbb{Z} \\ k \text{ odd}}} \left(\frac{-2}{k d_1 m_1 m_2 \nu}\right) \hat{F}_\nu \left(\frac{kX}{8\alpha^2 d_1 m_1 m_2 \nu}\right) G_k(m_1 m_2 \nu), \tag{3.6.4.5}$$

$$\mathcal{U}_1 = \left(\frac{1+i}{2}\right) \left(\frac{2d_1}{m_1 m_2 \nu}\right) \sum_{\substack{k \in \mathbb{Z} \\ k \equiv 2 \pmod 4}} e\left(\frac{k\overline{\alpha^2 d_1 m_1 m_2 \nu}}{8}\right) \hat{F}_\nu \left(\frac{kX}{8\alpha^2 d_1 m_1 m_2 \nu}\right) G_k(m_1 m_2 \nu), \tag{3.6.4.6}$$

$$\mathcal{U}_2 = \left(\frac{1-i}{2}\right) \left(\frac{-2d_1}{m_1 m_2 \nu}\right) \sum_{\substack{k \in \mathbb{Z} \\ k \equiv 2 \pmod 4}} e\left(\frac{k\overline{\alpha^2 d_1 m_1 m_2 \nu}}{8}\right) \hat{F}_\nu \left(\frac{kX}{8\alpha^2 d_1 m_1 m_2 \nu}\right) G_k(m_1 m_2 \nu), \tag{3.6.4.7}$$

$$\mathcal{V} = \left(\frac{2d_1}{m_1 m_2 \nu}\right) \sum_{\substack{k \in \mathbb{Z} \\ k \equiv 4 \pmod 8}} e\left(\frac{k\overline{\alpha^2 d_1 m_1 m_2 \nu}}{8}\right) \hat{F}_\nu \left(\frac{kX}{8\alpha^2 d_1 m_1 m_2 \nu}\right) \tau_k(m_1 m_2 \nu), \tag{3.6.4.8}$$

and

$$\mathcal{W} = \left(\frac{2d_1}{m_1 m_2 \nu}\right) \sum_{\substack{k\in\mathbb{Z} \\ k\equiv 0 \,(\mathrm{mod}\ 8) \\ k\neq 0}} e\left(\frac{k\overline{\alpha^2 d_1 m_1 m_2 \nu}}{8}\right) \hat{F}_\nu\left(\frac{kX}{8\alpha^2 d_1 m_1 m_2 \nu}\right) \tau_k(m_1 m_2 \nu). \tag{3.6.4.9}$$

3.6.5 Evaluation of the sum with $\mathcal{Q}_1$

In this subsection, we evaluate the sum

$$\mathcal{Q}_1^* := \sum_{\substack{\nu=1 \\ (\nu,2d)=1}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2}} \sum_{\substack{\alpha\leq Y \\ (\alpha,2m_1m_2\nu)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \mathcal{Q}_1, \tag{3.6.5.1}$$

with $\mathcal{Q}_1$ defined by (3.6.4.2). We may cancel the two Jacobi symbols $(\frac{2}{m_1 m_2 \nu})$ in (3.6.4.2), insert the resulting expression into (3.6.5.1), and then apply the Mellin inversion formula to the ν-sum to deduce that

$$\mathcal{Q}_1^* = \left(\frac{1+i}{2}\right)\frac{\sqrt{2}}{2} \sum_{\substack{\alpha\leq Y \\ (\alpha,2m_1m_2\nu)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1}\left(\frac{d_1}{m_1 m_2}\right) \sum_{\substack{k\in\mathbb{Z} \\ k \text{ odd}}} \left(\frac{2}{kd_1}\right)$$
$$\times \frac{1}{2\pi i}\int_{(c)}\int_0^\infty \hat{F}_t\left(\frac{kX}{8\alpha^2 d_1 m_1 m_2 t}\right) t^{w-1}\,dt \sum_{\substack{\nu=1 \\ (\nu,2d)=1}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2+w}}\left(\frac{d_1}{\nu}\right) G_k(m_1 m_2 \nu)\,dw \tag{3.6.5.2}$$

for any $c > 1$. The interchange in the order of summation is justified by absolute convergence. The next step is to write the ν-sum as an Euler product, as follows.

Lemma 3.6.2. *Let d_1 be as defined by* (3.6.2.5)*. For each nonzero integer k, define k_1 and k_2 uniquely by the equation*

$$4kd_1 = k_1 k_2^2, \tag{3.6.5.3}$$

where k_1 is a fundamental discriminant and k_2 is a positive integer. If ℓ is a positive *integer and Re(s) > 1, then*

$$\sum_{\substack{\nu=1 \\ (\nu,2\alpha d)=1}}^{\infty} \frac{d_2(\nu)}{\nu^s} \left(\frac{d_1}{\nu}\right) \frac{G_k(\ell\nu)}{\sqrt{\nu}} = L(s,\chi_{k_1})^2 \prod_p \mathcal{G}_{0,p}(s;k,\ell,\alpha,d)$$

$$=: L(s,\chi_{k_1})^2 \mathcal{G}_0(s;k,\ell,\alpha,d),$$

where $\chi_{k_1}(\cdot) = \left(\frac{k_1}{\cdot}\right)$ *and* $\mathcal{G}_{0,p}(s;k,\ell,\alpha,d)$ *is defined by*

$$\mathcal{G}_{0,p}(s;k,\ell,\alpha,d) = \left(1 - \frac{1}{p^s}\left(\frac{k_1}{p}\right)\right)^2 \quad \text{if } p|2\alpha d, \quad \text{and}$$

$$\mathcal{G}_{0,p}(s;k,\ell,\alpha,d) = \left(1 - \frac{1}{p^s}\left(\frac{k_1}{p}\right)\right)^2 \sum_{r=0}^{\infty} \frac{r+1}{p^{rs}} \left(\frac{d_1}{p^r}\right) \frac{G_k(p^{r+\mathrm{ord}_p(\ell)})}{p^{r/2}} \quad \text{if } p \nmid 2\alpha d.$$

The function $\mathcal{G}_0(s;k,\ell,\alpha,d)$ *is holomorphic for Re(s)* $> \frac{1}{2}$. *Moreover, if* k_3 *and* k_4 *are defined by the equation*

$$k = k_3 k_4^2, \tag{3.6.5.4}$$

with k_3 *square-free and* k_4 *a positive integer, then*

$$\mathcal{G}_0(s;k,\ell,\alpha,d) \ll_\varepsilon (\alpha d|k|\ell)^\varepsilon \ell^{1/2} (\ell, k_4^2)^{1/2}$$

uniformly for Re(s) $\geq \frac{1}{2} + \varepsilon$.

Proof. It follows from the definition of $\mathcal{G}_{0,p}(s;k,\ell,\alpha,d)$ and Lemma 3.3.3 that

$$\mathcal{G}_{0,p}(s;k,\ell,\alpha,d) = \left(1 - \frac{1}{p^s}\left(\frac{k_1}{p}\right)\right)^2 \left(1 + \frac{2}{p^s}\left(\frac{d_1 k}{p}\right)\right) = 1 - \frac{3}{p^{2s}} + \frac{2}{p^{3s}}\left(\frac{k_1}{p}\right)$$

for each $p \nmid 2\alpha dk\ell$, since $\left(\frac{d_1 k}{p}\right) = \left(\frac{k_1}{p}\right)$ for odd primes p, by (3.6.5.3). The rest of the proof is similar to that of [51, Lemma 5.3]. □

We also need some analytic properties of the function $h(\xi, w)$ defined for $\mathrm{Re}(w) > 0$ by

$$h(\xi,w) = \int_0^\infty \hat{F}_t\left(\frac{\xi}{t}\right) t^{w-1}\, dt.$$

These are embodied in the following lemma. As a bit of notation, for a real number x

we define

$$\operatorname{sgn}(x) = \begin{cases} 1, & x \geq 0, \\ -1, & x < 0. \end{cases}$$

Lemma 3.6.3. *Let F_t be defined by* (3.6.2.4). *If $\xi \neq 0$ then*

$$h(\xi, w) = |\xi|^w \check{\Phi}(w) \int_0^\infty \omega_2\left(\frac{|\xi|\pi}{Xz}\right)(\cos(2\pi z) - i\operatorname{sgn}(\xi)\sin(2\pi z))\,\frac{dz}{z^{w+1}}.$$

The integral above may be expressed as

$$\begin{aligned}&\frac{1}{2\pi i}\int_{(c)} \frac{\Gamma\left(\frac{s}{2}+\frac{1}{4}\right)^2}{\Gamma\left(\frac{1}{4}\right)^2}\left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)^2 \frac{X^s}{(\pi|\xi|)^s}(2\pi)^{-s+w}\Gamma(s-w) \\ &\times \left\{\cos\left(\tfrac{\pi}{2}(s-w)\right) - i\operatorname{sgn}(\xi)\sin\left(\tfrac{\pi}{2}(s-w)\right)\right\}\,\frac{ds}{s}\end{aligned} \tag{3.6.5.5}$$

for any c with $\operatorname{Re}(w)+1 > c > \max\{0, \operatorname{Re}(w)\}$. If $\xi \neq 0$, then $h(\xi, w)$ is an entire function of w. In the region $1 \geq \operatorname{Re}(w) > -1$, it satisfies the bound

$$h(\xi, w) \ll (1+|w|)^{-\operatorname{Re}(w)-\frac{1}{2}} \exp\left(-\frac{1}{10}\frac{\sqrt{|\xi|}}{\sqrt{X(|w|+1)}}\right)|\xi|^w|\check{\Phi}(w)|.$$

Proof. The proof is similar to that of [51, Lemma 5.2]. □

Now, by these lemmas and the rapid decay of $\check{\Phi}(w)$ as $|\operatorname{Im}(w)| \to \infty$ in a fixed vertical strip, we may move the line of integration of the w-integral in (3.6.5.2) to $\operatorname{Re}(w) = -\frac{1}{2} + \varepsilon$. This leaves a residue from a pole at $w = 0$ only when χ_{k_1} is a principal character, which holds if and only if $k_1 = 1$. By (3.6.5.3), $k_1 = 1$ if and only if kd_1 is a perfect square. Hence

$$\mathcal{Q}_1^* = \mathcal{P}_1 + \mathcal{R}_1, \tag{3.6.5.6}$$

where $\mathcal{P}_1$ is defined by

$$\mathcal{P}_1 = \operatorname*{Res}_{w=0} \left(\frac{1+i}{2}\right)\frac{\sqrt{2}}{2} \sum_{\substack{\alpha \le Y \\ (\alpha, 2m_1m_2)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \left(\frac{d_1}{m_1m_2}\right) \sum_{\substack{k\in\mathbb{Z} \\ k \text{ odd} \\ kd_1=\square}} h\left(\frac{kX}{8\alpha^2 d_1 m_1 m_2}, w\right) \times \zeta(1+w)^2 \mathcal{G}_0(1+w; k, m_1m_2, \alpha, d) \tag{3.6.5.7}$$

and $\mathcal{R}_1$ is defined by

$$\mathcal{R}_1 = \left(\frac{1+i}{2}\right)\frac{\sqrt{2}}{2} \sum_{\substack{\alpha \le Y \\ (\alpha, 2m_1m_2)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \left(\frac{d_1}{m_1m_2}\right) \sum_{\substack{k\in\mathbb{Z} \\ k \text{ odd}}} \left(\frac{2}{kd_1}\right) \times \frac{1}{2\pi i}\int_{(-\frac{1}{2}+\varepsilon)} h\left(\frac{kX}{8\alpha^2 d_1 m_1 m_2}, w\right) L(1+w, \chi_{k_1})^2 \mathcal{G}_0(1+w; k, m_1m_2, \alpha, d)\, dw. \tag{3.6.5.8}$$

We bound $\mathcal{R}_1$ in Subsection 3.6.6. To estimate $\mathcal{P}_1$, observe that d_1 is square-free by its definition (3.6.2.5) and the fact that d is square-free. This implies that kd_1 is a perfect square if and only if k equals d_1 times a perfect square. Hence, in (3.6.5.7), we may relabel k as $d_1 j^2$, where j runs through all the odd positive integers. With this and Lemma 3.6.3, we deduce from (3.6.5.7) that

$$\begin{aligned}\mathcal{P}_1 = \operatorname*{Res}_{w=0} &\left(\frac{1+i}{2}\right)\frac{\sqrt{2}}{2} \sum_{\substack{\alpha \le Y \\ (\alpha, 2m_1m_2)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1}\zeta(1+w)^2 \check{\Phi}(w) X^w \\ &\times \frac{1}{2\pi i}\int_{(c)} \frac{\Gamma\left(\frac{s}{2}+\frac{1}{4}\right)^2}{\Gamma\left(\frac{1}{4}\right)^2}\left(1-\frac{1}{2^{\frac{1}{2}-s}}\right)^2 \pi^{-s}\Gamma_2(s-w)(8\alpha^2 m_1m_2)^{s-w} \\ &\times \sum_{\substack{j=1 \\ j \text{ odd}}}^{\infty} j^{-2s+2w}\left(\frac{d_1}{m_1m_2}\right)\mathcal{G}_0(1+w; d_1j^2, m_1m_2, \alpha, d)\,\frac{ds}{s},\end{aligned} \tag{3.6.5.9}$$

where $\Gamma_2(u)$ is defined by

$$\Gamma_2(u) = (2\pi)^{-u}\Gamma(u)\left(\cos\left(\tfrac{\pi}{2}u\right) - i\sin\left(\tfrac{\pi}{2}u\right)\right), \tag{3.6.5.10}$$

and where we take $c > \frac{1}{2}$ to guarantee the absolute convergence of the j-sum.

We next write the j-sum in (3.6.5.9) as an Euler product. By (ii) of Lemma 3.3.3, if

j is a positive integer then

$$\left(\frac{d_1}{p^\beta}\right) G_{d_1 j^2}(p^\beta) = G_{j^2}(p^\beta)$$

for all $p \nmid 2\alpha d$ and $\beta \geq 1$. From this and the definition of $\mathcal{G}_0$ in Lemma 3.6.2, we see that

$$\left(\frac{d_1}{m_1 m_2}\right) \mathcal{G}_0(1+w; d_1 j^2, m_1 m_2, \alpha, d) = \mathcal{G}(1+w; j^2, m_1 m_2, \alpha d),$$

where $\mathcal{G}$ is defined by [51, (5.8)]. Hence we may write the inner j-sum in (3.6.5.9) as an Euler product

$$\begin{aligned} &\sum_{\substack{j=1 \\ j \text{ odd}}}^{\infty} j^{-2s+2w} \mathcal{G}(1+w; j^2, m_1 m_2, \alpha d) \\ &= \left(1 - \frac{1}{2^{1+w}}\right)^2 \prod_{p>2} \sum_{b=0}^{\infty} p^{2b(w-s)} \mathcal{G}_p(1+w; p^{2b}, m_1 m_2, \alpha d) \\ &= \left(1 - \frac{1}{4^{s-w}}\right) \prod_{p} \sum_{b=0}^{\infty} p^{2b(w-s)} \mathcal{G}_p(1+w; p^{2b}, m_1 m_2, \alpha d). \end{aligned} \tag{3.6.5.11}$$

This latter expression is [51, p. 471]

$$\left(1 - \frac{1}{4^{s-w}}\right) (m_1 m_2)^{1-s+w} \ell_1^{s-w-\frac{1}{2}} \zeta(2s-2w) \zeta(2s+1) \mathcal{H}_1(s-w, 1+w; m_1 m_2, \alpha d),$$

where ℓ_1 is the square-free integer defined by the equation

$$m_1 m_2 = \ell_1 \ell_2^2, \quad \mu^2(\ell_1) = 1, \ \ell_2 \in \mathbb{Z}, \tag{3.6.5.12}$$

and $\mathcal{H}_1$ is defined by an Euler product

$$\mathcal{H}_1(s-w, 1+w; m_1 m_2, \alpha d) = \prod_p \mathcal{H}_{1,p}.$$

The local factors $\mathcal{H}_{1,p}$ are

$$\mathcal{H}_{1,p} = \begin{cases} \left(1-\frac{1}{p^{1+w}}\right)^2\left(1-\frac{1}{p^{1+2s}}\right) & \text{if } p|2\alpha d \\ \frac{\left(1-\frac{1}{p^{1+w}}\right)^2}{\left(1-\frac{1}{p^{1+2s}}\right)}\left(1+\frac{2}{p^{1+w}}-\frac{2}{p^{1+2s-w}}+\frac{1}{p^{1+2s}}-\frac{3}{p^{2+2s}}+\frac{1}{p^{3+4s}}\right) & \text{if } p\nmid 2\alpha d m_1 m_2 \\ \frac{\left(1-\frac{1}{p^{1+w}}\right)^2}{\left(1-\frac{1}{p^{1+2s}}\right)}\left(1-\frac{1}{p^{2s-2w}}+\frac{2}{p^{2s-w}}-\frac{2}{p^{1+2s-w}}+\frac{1}{p^{1+2s}}-\frac{1}{p^{1+4s-2w}}\right) & \text{if } p|\ell_1 \\ \frac{\left(1-\frac{1}{p^{1+w}}\right)^2}{\left(1-\frac{1}{p^{1+2s}}\right)}\left(1-\frac{1}{p}+\frac{2}{p^{1+w}}-\frac{2}{p^{1+2s-w}}+\frac{1}{p^{1+2s}}-\frac{1}{p^{2+2s}}\right) & \text{if } p|m_1m_2,\ p\nmid \ell_1. \end{cases} \tag{3.6.5.13}$$

Inserting this expression for the j-sum in (3.6.5.9) into (3.6.5.9), we find that

$$\mathcal{P}_1 = \left(\frac{1+i}{2}\right)\frac{\sqrt{2}}{2}\sum_{\substack{\alpha\le Y\\(\alpha,2m_1m_2)=1}}\frac{\mu(\alpha)}{\alpha^2 d_1}\,\mathcal{I}, \tag{3.6.5.14}$$

where

$$\mathcal{I} = \operatorname*{Res}_{w=0}\zeta(1+w)^2\check{\Phi}(w)X^w\frac{1}{2\pi i}\int_{(c)}\frac{\Gamma\left(\frac{s}{2}+\frac{1}{4}\right)^2}{\Gamma\left(\frac{1}{4}\right)^2}\left(1-\frac{1}{2^{\frac{1}{2}-s}}\right)^2\pi^{-s}\Gamma_2(s-w)(8\alpha^2)^{s-w} \\ \times\left(1-\frac{1}{4^{s-w}}\right)m_1m_2\ell_1^{s-w-\frac{1}{2}}\zeta(2s-2w)\zeta(2s+1)\mathcal{H}_1(s-w,1+w;m_1m_2,\alpha d)\,\frac{ds}{s}. \tag{3.6.5.15}$$

The next step is to extend the α-sum to infinity and show that the error introduced in doing so is small. To do this, we need to move the line of integration in (3.6.5.15) closer to 0 to guarantee the absolute convergence of the α-sum. We first evaluate the residue to see that (3.6.5.15) is the same as

$$\mathcal{I} = \frac{\check{\Phi}(0)}{2\pi i}\int_{(c)}\frac{\Gamma\left(\frac{s}{2}+\frac{1}{4}\right)^2}{\Gamma\left(\frac{1}{4}\right)^2}\left(1-\frac{1}{2^{\frac{1}{2}-s}}\right)^2\pi^{-s}\Gamma_2(s)(8\alpha^2)^s\left(1-\frac{1}{4^s}\right)m_1m_2 \\ \times\ell_1^{s-\frac{1}{2}}\zeta(2s)\zeta(2s+1)\mathcal{H}_1(s,1;m_1m_2,\alpha d)\left\{2\gamma+\frac{(\check{\Phi})'(0)}{\check{\Phi}(0)}+\log\left(\frac{X}{8\alpha^2\ell_1}\right)\right. \\ \left.-\frac{\Gamma_2'}{\Gamma_2}(s)+\frac{\log 4}{(1-4^s)}-2\frac{\zeta'}{\zeta}(2s)+\frac{\frac{\partial}{\partial w}\mathcal{H}_1(s-w,1+w;m_1m_2,\alpha d)}{\mathcal{H}_1(s-w,1+w;m_1m_2,\alpha d)}\Bigg|_{w=0}\right\}\frac{ds}{s}. \tag{3.6.5.16}$$

Here γ denotes the Euler-Mascheroni constant. The definition (3.6.5.13) of $\mathcal{H}_1(s-w,1+$

$w; m_1m_2, \alpha d)$ implies that it is holomorphic for $\mathrm{Re}(s) > 0$ and $|w| < \max\{\frac{1}{2}, 2|s|\}$, and that it and its first partial derivatives at $w = 0$ are bounded by $\ll (\alpha X)^\varepsilon$ for $\mathrm{Re}(s) \geq \frac{1}{\log X}$. Thus, by the rapid decay of the gamma function, we may move the line of integration in (3.6.5.16) to $\mathrm{Re}(s) = \frac{1}{\log X}$. There is no residue because the poles of $\zeta(2s)$ and $\frac{\zeta'}{\zeta}(2s)$ at $s = \frac{1}{2}$ are canceled by the zero of the factor $(1 - 2^{s-\frac{1}{2}})^2$. Using well-known bounds for $\zeta(2s)$ and $\zeta'(2s)$ implied by the Phragmén-Lindelöf principle, we see that the new integral is now bounded by

$$\ll m_1 m_2 \ell_1^{-\frac{1}{2}+\varepsilon} \alpha^\varepsilon X^\varepsilon \int_{\left(\frac{1}{\log X}\right)} \left|\Gamma\left(\frac{s}{2} + \frac{1}{4}\right)\right|^2 \max\{|\Gamma_2(s)|, |\Gamma_2'(s)|\} (1 + |s|)^{\frac{1}{2}+\varepsilon} \, |ds|,$$

which is $\ll m_1 m_2 \ell_1^{-\frac{1}{2}+\varepsilon} \alpha^\varepsilon X^\varepsilon$ by the rapid decay of the gamma function. Dividing this bound by $\alpha^2 d_1$ and summing the result over all $\alpha > Y$, we deduce that

$$\sum_{\substack{\alpha > Y \\ (\alpha, 2m_1m_2)=1}} \frac{\mu^2(\alpha)}{\alpha^2 d_1} \, |\mathcal{I}| \ll \frac{m_1 m_2 \ell_1^{-\frac{1}{2}+\varepsilon} X^\varepsilon}{d^{1-\varepsilon} Y^{1-\varepsilon}} \tag{3.6.5.17}$$

because, by (3.6.2.5), if $\varphi(j)$ is the Euler totient function, then

$$\sum_{\alpha > Y} \frac{1}{\alpha^{2-\varepsilon} d_1} = \frac{1}{d} \sum_{j|d} \varphi(j) \sum_{\substack{\alpha > Y \\ j|\alpha}} \frac{1}{\alpha^{2-\varepsilon}} \ll \frac{1}{d^{1-\varepsilon} Y^{1-\varepsilon}}.$$

From (3.6.5.14), (3.6.5.17), and (3.6.5.15) now with $c = \frac{1}{\log X}$, we arrive at

$$\begin{aligned} \mathcal{P}_1 = \operatorname*{Res}_{w=0} \left(\frac{1+i}{2}\right) \frac{\sqrt{2}}{2} X^w \frac{1}{2\pi i} \int_{\left(\frac{1}{\log X}\right)} \Gamma_2(s-w) 8^{s-w} \left(1 - \frac{1}{4^{s-w}}\right) \\ \times \mathcal{K}(s, w; m_1m_2, d) \, \frac{ds}{s} + O\left(\frac{m_1 m_2 \ell_1^{-\frac{1}{2}+\varepsilon} X^\varepsilon}{d^{1-\varepsilon} Y^{1-\varepsilon}}\right), \end{aligned} \tag{3.6.5.18}$$

with $\mathcal{K}(s,w;m_1m_2,d)$ defined by

$$\mathcal{K}(s,w;m_1m_2,d) = \zeta(1+w)^2\check{\Phi}(w)\frac{\Gamma\left(\frac{s}{2}+\frac{1}{4}\right)^2}{\Gamma\left(\frac{1}{4}\right)^2}\left(1-\frac{1}{2^{\frac{1}{2}-s}}\right)^2\pi^{-s}m_1m_2\ell_1^{s-w-\frac{1}{2}}$$
$$\times\zeta(2s-2w)\zeta(2s+1)\sum_{\substack{\alpha=1\\(\alpha,2m_1m_2)=1}}^{\infty}\frac{\mu(\alpha)}{\alpha^{2-2s+2w}d_1}\mathcal{H}_1(s-w,1+w;m_1m_2,\alpha d), \tag{3.6.5.19}$$

where, as before, ℓ_1 is defined by (3.6.5.12), d_1 is defined by (3.6.2.5), and $\mathcal{H}_1$ is defined as the product of (3.6.5.13) over all primes.

It is convenient for later calculations to write $\mathcal{P}_1$ in terms of a residue, as in (3.6.5.18), rather than in terms of logarithmic derivatives as in (3.6.5.16).

3.6.6 Bounding the contribution of $\mathcal{R}_1$

Having handled $\mathcal{P}_1$ in (3.6.5.6), we next turn to $\mathcal{R}_1$, defined by (3.6.5.8). It will be convenient to denote

$$\mathcal{R}(\ell,d) = \frac{1}{\ell}\left(\frac{1+i}{2}\right)\frac{\sqrt{2}}{2}\sum_{\substack{\alpha\leq Y\\(\alpha,2\ell)=1}}\frac{\mu(\alpha)}{\alpha^2d_1}\left(\frac{d_1}{\ell}\right)\sum_{\substack{k\in\mathbb{Z}\\k\text{ odd}}}\left(\frac{2}{kd_1}\right)$$
$$\times\frac{1}{2\pi i}\int_{(-\frac{1}{2}+\varepsilon)}h\left(\frac{kX}{8\alpha^2d_1\ell},w\right)L(1+w,\chi_{k_1})^2\mathcal{G}_0(1+w;k,\ell,\alpha,d)\,dw, \tag{3.6.6.1}$$

so that $\mathcal{R}_1 = m_1m_2\mathcal{R}(m_1m_2,d)$. We will bound $|\mathcal{R}(\ell,d)|$ on average as ℓ and d each range over a dyadic interval.

Let $\beta_{\ell,d} = \overline{\mathcal{R}(\ell,d)}/|\mathcal{R}(\ell,d)|$ if $\mathcal{R}(\ell,d)\neq 0$, and $\beta_{\ell,d}=1$ otherwise. Then $|\beta_{\ell,d}|=1$ and $|\mathcal{R}(\ell,d)| = \beta_{\ell,d}\mathcal{R}(\ell,d)$. We sum this over all ℓ,d with $J\leq\ell<2J$ and $V\leq d<2V$, where $J,V\geq 1$. We then insert the definition (3.6.6.1) and bring the d,ℓ-sum inside the integral to deduce that

$$\sum_{\substack{d=V\\(d,2)=1}}^{2V-1}\sum_{\substack{\ell=J\\(\ell,2d)=1}}^{2J-1}|\mathcal{R}(\ell,d)| = \sum_{\substack{d=V\\(d,2)=1}}^{2V-1}\sum_{\substack{\ell=J\\(\ell,2d)=1}}^{2J-1}\beta_{\ell,d}\mathcal{R}(\ell,d)$$
$$\ll\sum_{\substack{\alpha\leq Y\\(\alpha,2)=1}}\frac{1}{\alpha^2}\sum_{\substack{k\in\mathbb{Z}\\k\text{ odd}}}\int_{(-\frac{1}{2}+\varepsilon)}U(\alpha,k,w)\,|dw|, \tag{3.6.6.2}$$

where for brevity we denote

$$U(\alpha,k,w)=\sum_{\substack{d=V\\(d,2)=1}}^{2V-1}\frac{1}{d_1}|L(1+w,\chi_{k_1})|^2\Bigg|\sum_{\substack{\ell=J\\(\ell,2\alpha d)=1}}^{2J-1}\frac{\beta_{\ell,d}}{\ell}\left(\frac{d_1}{\ell}\right)$$
$$\times\,\mathcal{G}_0(1+w;k,\ell,\alpha,d)h\left(\frac{kX}{8\alpha^2 d_1\ell},w\right)\Bigg|.$$

We split the k-sum into dyadic blocks $K\leq|k|<2K$, with $K\geq 1$, and apply Cauchy's inequality to write

$$\sum_{\substack{K\leq|k|<2K\\k\text{ odd}}}U(\alpha,k,w)\ll\left(\sum_{\substack{d=V\\(d,2)=1}}^{2V-1}\frac{1}{d_1}\sum_{\substack{K\leq|k|<2K\\k\text{ odd}}}k_2|L(1+w,\chi_{k_1})|^4\right)^{\frac{1}{2}}$$
$$\times\left(\sum_{\substack{d=V\\(d,2)=1}}^{2V-1}\frac{1}{d_1}\sum_{\substack{K\leq|k|<2K\\k\text{ odd}}}\frac{1}{k_2}\Bigg|\sum_{\substack{\ell=J\\(\ell,2\alpha d)=1}}^{2J-1}\frac{\beta_{\ell,d}}{\ell}\left(\frac{d_1}{\ell}\right)\right. \tag{3.6.6.3}$$
$$\left.\times\,\mathcal{G}_0(1+w;k,\ell,\alpha,d)h\left(\frac{kX}{8\alpha^2 d_1\ell},w\right)\Bigg|^2\right)^{\frac{1}{2}},$$

where k_2 is defined by (3.6.5.3). To bound the first factor on the right-hand side of (3.6.6.3), we split the k-sum according to the values of k_1 and k_2 and interchange the order of summation. Then we use the fact that $d_1\geq d/\alpha$ by (3.6.2.5) to deduce that

$$\sum_{\substack{d=V\\(d,2)=1}}^{2V-1}\frac{1}{d_1}\sum_{\substack{K\leq|k|<2K\\k\text{ odd}}}k_2|L(1+w,\chi_{k_1})|^4$$
$$\leq\frac{\alpha}{V}\sum_{0<|k_1|\ll KV}|L(1+w,\chi_{k_1})|^4\sum_{k_2\ll\sqrt{\frac{KV}{k_1}}}k_2\sum_{\substack{d=V\\(d,2)=1\\d_1|k_1k_2^2}}^{2V-1}1.$$

We estimate the inner sum using the divisor bound, and find that the above is

$$\ll\alpha K^{1+\varepsilon}V^{\varepsilon}\sum_{0<|k_1|\ll KV}\frac{1}{k_1}|L(1+w,\chi_{k_1})|^4\ll\alpha K^{1+\varepsilon}V^{\varepsilon}(1+|w|)^{1+\varepsilon}$$

by Lemma 3.3.5. It follows from this and (3.6.6.3) that

$$\begin{aligned}\sum_{\substack{K\le |k|<2K\\ k \text{ odd}}} U(\alpha,k,w) \ll &\Big(\alpha K^{1+\varepsilon}V^{\varepsilon}(1+|w|)^{1+\varepsilon}\Big)^{\frac{1}{2}}\\ &\times\Bigg(\sum_{\substack{d=V\\(d,2)=1}}^{2V-1}\frac{1}{d_1}\sum_{\substack{K\le|k|<2K\\ k \text{ odd}}}\frac{1}{k_2}\Bigg|\sum_{\substack{\ell=J\\(\ell,2\alpha d)=1}}^{2J-1}\frac{\beta_{\ell,d}}{\ell}\left(\frac{d_1}{\ell}\right)\\ &\times\mathcal{G}_0(1+w;k,\ell,\alpha,d)h\left(\frac{kX}{8\alpha^2 d_1\ell},w\right)\Bigg|^2\Bigg)^{\frac{1}{2}}.\end{aligned} \tag{3.6.6.4}$$

The next task is to bound the second factor on the right-hand side. To this end we prove the following two lemmas.

Lemma 3.6.4. *Let $\alpha \le Y$, d, K, and J be positive integers, and suppose w is a complex number with real part $-\frac{1}{2}+\varepsilon$. Then for any choice of complex numbers γ_ℓ with $|\gamma_\ell|\le 1$,*

$$\sum_{\substack{K\le|k|<2K\\ k \text{ odd}}}\frac{1}{k_2}\Bigg|\sum_{\substack{\ell=J\\(\ell,2\alpha d)=1}}^{2J-1}\frac{\gamma_\ell}{\ell}\mathcal{G}_0(1+w;k,\ell,\alpha,d)h\left(\frac{kX}{8\alpha^2 d_1\ell},w\right)\Bigg|^2$$

is bounded by

$$\ll_\varepsilon |\check{\Phi}(w)|^2\frac{d_1\alpha^{2+\varepsilon}J^{2+\varepsilon}K^{\varepsilon}d^{\varepsilon}}{X^{1-\varepsilon}}\exp\left(-\frac{1}{20}\frac{\sqrt{K}}{\alpha\sqrt{d_1J(1+|w|)}}\right).$$

and also by

$$\ll_\varepsilon ((1+|w|)\alpha dJKX)^{\varepsilon}|\check{\Phi}(w)|^2\frac{\alpha^2 d_1(JK+J^2)}{KX}.$$

Lemma 3.6.5. *Let $\delta_\ell \ll \ell^\varepsilon$ be any sequence of complex numbers and let* $\mathrm{Re}(w) = -\frac{1}{2}+\varepsilon$. Then

$$\sum_{K\le|k|<2K}\frac{1}{k_2}\left|\sum_{\substack{\ell=J\\(\ell,2\alpha d)=1}}^{2J-1}\frac{\delta_\ell}{\sqrt{\ell}}\mathcal{G}_0(1+w;k,\ell,\alpha,d)\right|^2 \ll_\varepsilon (\alpha dJK)^{\varepsilon}J(J+K).$$

Proof of Lemma 3.6.4 assuming Lemma 3.6.5. To prove the first bound, we use the triangle inequality and apply the bounds for $\mathcal{G}_0$ from Lemma 3.6.2 and $h(\xi,w)$ from

Lemma 3.6.3 to deduce that the sum in question is

$$\ll |\check{\Phi}(w)|^2 \frac{d_1 \alpha^{2+\varepsilon} J^\varepsilon K^\varepsilon d^\varepsilon}{X^{1-\varepsilon}} \exp\left(-\frac{1}{20} \frac{\sqrt{K}}{\alpha \sqrt{d_1 J(1+|w|)}} \right)$$
$$\times \sum_{\substack{K \le |k| < 2K \\ k \text{ odd}}} \frac{1}{|k| k_2} \left(\sum_{\substack{\ell=J \\ (\ell, 2\alpha d)=1}}^{2J-1} (\ell, k_4^2)^{\frac{1}{2}} \right)^2 .$$

We then estimate the k-sum by splitting it according to the values of k_1 and k_2 and using $(\ell, k_4^2) \le k_4^2 \le k_2^2$, which follows from (3.6.5.3) and (3.6.5.4). This leads to the first bound of the lemma.

To prove the second bound, we apply Lemma 3.6.3 and write the integral (3.6.5.5) as

$$\frac{1}{2\pi i} \int_{(c)} g(s, w; \operatorname{sgn}(\xi)) \left(\frac{X}{\pi |\xi|} \right)^s ds$$

with $c = \varepsilon$. We then bring the ℓ-sum inside the integral and use the triangle inequality to deduce that

$$\left| \sum_{\substack{\ell=J \\ (\ell, 2\alpha d)=1}}^{2J-1} \frac{\gamma_\ell}{\ell} \mathcal{G}_0(1+w; k, \ell, \alpha, d) h\left(\frac{kX}{8\alpha^2 d_1 \ell}, w \right) \right| \ll |\check{\Phi}(w)| \left(\frac{\alpha^{1+\varepsilon} d_1^{\frac{1}{2}+\varepsilon}}{|k|^{\frac{1}{2}-\varepsilon} X^{\frac{1}{2}-\varepsilon}} \right)$$
$$\times \int_{(\varepsilon)} \left| g(s, w; \operatorname{sgn}(k)) \sum_{\substack{\ell=J \\ (\ell, 2\alpha d)=1}}^{2J-1} \frac{\gamma_\ell}{\ell^{1+w-s}} \mathcal{G}_0(1+w; k, \ell, \alpha, d) \right| |ds|.$$

Thus, since $g(s, w; \operatorname{sgn}(k)) \ll_\varepsilon (1+|w|)^\varepsilon \exp(-(\frac{\pi}{2} - \varepsilon)|\mathrm{Im}(s)|)$ by Stirling's formula, it follows from Cauchy's inequality that

$$\left| \sum_{\substack{\ell=J \\ (\ell, 2\alpha d)=1}}^{2J-1} \frac{\gamma_\ell}{\ell} \mathcal{G}_0(1+w; k, \ell, \alpha, d) h\left(\frac{kX}{8\alpha^2 d_1 \ell}, w \right) \right|^2 \ll (1+|w|)^\varepsilon |\check{\Phi}(w)|^2 \left(\frac{\alpha^{2+\varepsilon} d_1^{1+\varepsilon}}{|k|^{1-\varepsilon} X^{1-\varepsilon}} \right)$$
$$\times \int_{(\varepsilon)} \exp(-(\tfrac{\pi}{2} - \varepsilon)|\mathrm{Im}(s)|) \left| \sum_{\substack{\ell=J \\ (\ell, 2\alpha d)=1}}^{2J-1} \frac{\gamma_\ell}{\ell^{1+w-s}} \mathcal{G}_0(1+w; k, \ell, \alpha, d) \right|^2 |ds|.$$

The second bound of the lemma follows from this and Lemma 3.6.5. □

Proof of Lemma 3.6.5. For any integer $k = \pm \prod_{i,\ a_i \ge 1} p_i^{a_i}$, let $a(k)$ and $b(k)$ be defined

by

$$a(k) = \prod_i p_i^{a_i+1} \quad \text{and} \quad b(k) = \prod_{a_i=1} p_i \prod_{a_i \ge 2} p_i^{a_i-1}. \tag{3.6.6.5}$$

From the definition of $\mathcal{G}_0$ in Lemma 3.6.2, we see that $\mathcal{G}_0(1+w;k,\ell,\alpha,d) = 0$ unless ℓ can be written as gm with $g|a(k)$ and m square-free and relatively prime to k. With this expression for ℓ, it follows from Lemma 3.3.3 that if $(\ell, 2\alpha d) = 1$ then

$$\mathcal{G}_0(1+w;k,\ell,\alpha,d) = \sqrt{m}\left(\frac{k}{m}\right)\prod_{p|m}\left(1+\frac{2}{p^{1+w}}\left(\frac{k_1}{p}\right)\right)^{-1}\mathcal{G}_0(1+w;k,g,\alpha,d). \tag{3.6.6.6}$$

From this and Cauchy's inequality, we arrive at

$$\begin{aligned} &\sum_{K\le|k|<2K}\frac{1}{k_2}\left|\sum_{\substack{\ell=J\\(\ell,2\alpha d)=1}}^{2J-1}\frac{\delta_\ell}{\sqrt{\ell}}\mathcal{G}_0(1+w;k,\ell,\alpha,d)\right|^2 \\ &\ll_\varepsilon K^\varepsilon \sum_{K\le|k|<2K}\frac{1}{k_2}\sum_{\substack{g|a(k)\\g<2J}}(\Psi_1(k,g)+\Psi_2(k,g)), \end{aligned} \tag{3.6.6.7}$$

where

$$\Psi_1(k,g) = \left|\sum_{\substack{\frac{J}{g}\le m<\frac{2J}{g}\\(m,2\alpha d)=1\\3\nmid m}}\frac{\mu^2(m)\delta_{gm}}{\sqrt{g}}\mathcal{G}_0(1+w;k,g,\alpha,d)\left(\frac{k}{m}\right)\prod_{p|m}\left(1+\frac{2}{p^{1+w}}\left(\frac{k_1}{p}\right)\right)^{-1}\right|^2$$

and $\Psi_2(k,g)$ is the same, but with the condition $3|m$ instead of $3\nmid m$. We first bound the contribution of Ψ_1. We factor out $g^{-1/2}\mathcal{G}_0(1+w;k,g,\alpha,d)$ and apply the bound from Lemma 3.6.2 to deduce that

$$\Psi_1(k,g) \ll_\varepsilon (\alpha d K)^\varepsilon g^{1+\varepsilon}\left|\sum_{\substack{\frac{J}{g}\le m<\frac{2J}{g}\\(m,6\alpha d)=1}}\mu^2(m)\delta_{gm}\left(\frac{k}{m}\right)\prod_{p|m}\left(1+\frac{2}{p^{1+w}}\left(\frac{k_1}{p}\right)\right)^{-1}\right|^2. \tag{3.6.6.8}$$

If $\left(\frac{k}{m}\right) \neq 0$, then

$$\prod_{p|m}\left(1+\frac{2}{p^{1+w}}\left(\frac{k_1}{p}\right)\right)^{-1} = \prod_{p|m}\left(1-\frac{4}{p^{2+2w}}\right)^{-1}\prod_{p|m}\left(1-\frac{2}{p^{1+w}}\left(\frac{k_1}{p}\right)\right)$$
$$= \prod_{p|m}\left(1-\frac{4}{p^{2+2w}}\right)^{-1}\sum_{j|m}\frac{\mu(j)d_2(j)}{j^{1+w}}\left(\frac{k_1}{j}\right).$$

We insert this into (3.6.6.8), interchange the order of summation, and apply Cauchy's inequality to see that

$$\Psi_1(k,g) \ll_\varepsilon (\alpha d K)^\varepsilon g^{1+\varepsilon}\sum_{j<\frac{2J}{g}}\left|\sum_{\substack{\frac{J}{g}\le m<\frac{2J}{g}\\(m,6\alpha d)=1\\ j|m}}\mu^2(m)\delta_{gm}\left(\frac{k}{m}\right)\prod_{p|m}\left(1-\frac{4}{p^{2+2w}}\right)^{-1}\right|^2.$$

We next relabel m as jm, factor out $\mu^2(j)\left(\frac{k}{j}\right)\prod_{p|j}\left(1-\frac{4}{p^{2+2w}}\right)^{-1}$ from the m-sum, and observe that $\prod_{p|j}\left(1-\frac{4}{p^{2+2w}}\right)^{-1} \ll_\varepsilon j^\varepsilon$ because $\mathrm{Re}(w) \geq -\frac{1}{2}+\varepsilon$ and $p>3$ for all $p|m$. The result is

$$\Psi_1(k,g) \ll_\varepsilon (\alpha d J K)^\varepsilon g^{1+\varepsilon}\sum_{j<\frac{2J}{g}}\left|\sum_{\substack{\frac{J}{gj}\le m<\frac{2J}{gj}\\(m,6\alpha dj)=1}}\mu^2(m)\delta_{gjm}\left(\frac{k}{m}\right)\prod_{p|m}\left(1-\frac{4}{p^{2+2w}}\right)^{-1}\right|^2. \tag{3.6.6.9}$$

Now, by (3.6.6.5), $g|a(k)$ implies $b(g)|k$. Thus we may interchange the order of summation to write

$$\sum_{K\le|k|<2K}\frac{1}{k_2}\sum_{\substack{g|a(k)\\ g<2J}}\Psi_1(k,g) \le \sum_{g<2J}\sum_{\substack{K\le|k|<2K\\ b(g)|k}}\frac{1}{k_2}\Psi_1(k,g) = \sum_{g<2J}\sum_{\frac{K}{b(g)}\le|f|<\frac{2K}{b(g)}}\frac{1}{k_2}\Psi_1(fb(g),g),$$

where we have relabeled k in the last sum as $fb(g)$, so that, by (3.6.5.3), $k_2>0$ satisfies $4fb(g)d_1 = k_1k_2^2$, with k_1 a fundamental discriminant. From this and (3.6.6.9),

we arrive at

$$\sum_{K\le|k|<2K}\frac{1}{k_2}\sum_{\substack{g|a(k)\\g<2J}}\Psi_1(k,g)\ll_\varepsilon(\alpha dJK)^\varepsilon\sum_{g<2J}g\sum_{\frac{K}{b(g)}\le|f|<\frac{2K}{b(g)}}\frac{1}{k_2}$$
$$\times\sum_{j<\frac{2J}{g}}\left|\sum_{\substack{\frac{J}{gj}\le m<\frac{2J}{gj}\\(m,6\alpha dj)=1}}\mu^2(m)\delta_{gjm}\left(\frac{fb(g)}{m}\right)\prod_{p|m}\left(1-\frac{4}{p^{2+2w}}\right)^{-1}\right|^2. \tag{3.6.6.10}$$

If $4f=f_1f_2^2$, with f_1 a fundamental discriminant and f_2 a positive integer, then the equation $4fb(g)d_1=k_1k_2^2$ implies that $f_2|2k_2$, and thus $k_2^{-1}\ll f_2^{-1}$. Hence it follows from (3.6.6.10) and Lemma 3.3.4 that

$$\sum_{K\le|k|<2K}\frac{1}{k_2}\sum_{\substack{g|a(k)\\g<2J}}\Psi_1(k,g)\ll_\varepsilon(\alpha dJK)^\varepsilon J(J+K).$$

This proves the desired bound for the sum of $\Psi_1(k,g)$ in (3.6.6.7). To bound the sum of $\Psi_2(k,g)$, we argue in the same way, but instead of (3.6.6.6) we use

$$\mathcal{G}_0(1+w;k,\ell,\alpha,d)=\sqrt{m}\left(\frac{k}{m}\right)\prod_{\substack{p|m\\p>3}}\left(1+\frac{2}{p^{1+w}}\left(\frac{k_1}{p}\right)\right)^{-1}\mathcal{G}_0^*(1+w;k,g,\alpha,d),$$

where

$$\mathcal{G}_0^*(1+w;k,g,\alpha,d)\;=\;\left(1-\frac{1}{3^{1+w}}\left(\frac{k_1}{3}\right)\right)^2\prod_{p\ne3}\mathcal{G}_{0,p}(1+w;k,g,\alpha,d),$$

with $\mathcal{G}_{0,p}$ as defined in Lemma 3.6.2. □

We now estimate the contribution of $\mathcal{R}_1$. From the first bound of Lemma 3.6.4, we see that the sum of the right-hand side of (3.6.6.4) over all $K=2^j>\alpha^2VJ(1+|w|)(\log X)^4$ is negligible. On the other hand, if $K\le\alpha^2VJ(1+|w|)(\log X)^4$ then it follows from

(3.6.6.4) and the second bound in Lemma 3.6.4 that

$$\sum_{\substack{K\leq|k|<2K\\ k\text{ odd}}} U(\alpha,k,w) \ll_\varepsilon (1+|w|)^{\frac{1}{2}+\varepsilon}|\check{\Phi}(w)|(\alpha JKVX)^\varepsilon\left(\frac{\alpha^3V(JK+J^2)}{X}\right)^{\frac{1}{2}}$$

$$\ll_\varepsilon (1+|w|)^{1+\varepsilon}|\check{\Phi}(w)|(\alpha JKVX)^\varepsilon\frac{\alpha^{\frac{5}{2}}VJ}{X^{\frac{1}{2}}}.$$

We sum this over all $K=2^j$, j a positive integer, with $K\leq\alpha^2VJ(1+|w|)(\log X)^4$, and then multiply the resulting sum by α^{-2}. We then integrate over all w with $\mathrm{Re}(w)=-\frac{1}{2}+\varepsilon$ and sum over all integers $\alpha\leq Y$ to deduce from (3.6.6.2) that

$$\sum_{\substack{d=V\\ (d,2)=1}}^{2V-1}\sum_{\substack{\ell=J\\ (\ell,2d)=1}}^{2J-1}|\mathcal{R}(\ell,d)| \ll \frac{V^{1+\varepsilon}J^{1+\varepsilon}Y^{\frac{3}{2}+\varepsilon}}{X^{\frac{1}{2}-\varepsilon}}. \tag{3.6.6.11}$$

Recall from (3.6.5.8) and (3.6.6.1) that $\mathcal{R}_1=m_1m_2\mathcal{R}(m_1m_2,d)$. Since $\lambda_d\ll d^\varepsilon$ by (3.4.8) and $b_m\ll 1$ by (3.2.5), it thus follows from (3.6.6.11) that

$$\sum_{\substack{d\leq D\\ d\text{ odd}}}\mu^2(d)\lambda_d\sum\sum_{\substack{m_1,m_2\leq M\\ (m_1m_2,2d)=1}}\frac{b_{m_1}b_{m_2}}{(m_1m_2)^{3/2}}|\mathcal{R}_1| \ll \frac{D^{1+\varepsilon}M^{1+\varepsilon}Y^{\frac{3}{2}+\varepsilon}}{X^{\frac{1}{2}-\varepsilon}}. \tag{3.6.6.12}$$

3.6.7 Conditions for the parameters

From (3.6.5.1), (3.6.5.6), (3.6.5.18), and (3.6.6.12), we see that the total contribution of the sum with $\mathcal{Q}_1$ to $\mathcal{B}$ in (3.6.4.1) is

$$\begin{aligned}
&\frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d\le D \\ d \text{ odd}}} \mu^2(d)\lambda_d \sum\sum_{\substack{m_1,m_2\le M \\ (m_1m_2,2d)=1}} \frac{b_{m_1}b_{m_2}}{(m_1m_2)^{3/2}} \sum_{\substack{\nu=1 \\ (\nu,2d)=1}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2}} \sum_{\substack{\alpha\le Y \\ (\alpha,2m_1m_2\nu)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \mathcal{Q}_1 \\
&= \left(\frac{1+i}{2}\right) \frac{\sqrt{2}X}{2(\sqrt{2}-1)^4} \sum_{\substack{d\le D \\ d \text{ odd}}} \mu^2(d)\lambda_d \sum\sum_{\substack{m_1,m_2\le M \\ (m_1m_2,2d)=1}} \frac{b_{m_1}\overline{b_{m_2}}}{(m_1m_2)^{3/2}} \operatorname*{Res}_{w=0} X^w \\
&\times \frac{1}{2\pi i} \int_{\left(\frac{1}{\log X}\right)} \Gamma_2(s-w) 8^{s-w} \left(1-\frac{1}{4^{s-w}}\right) \mathcal{K}(s,w;m_1m_2,d) \frac{ds}{s} \\
&+ O\left(\frac{X^{1+\varepsilon}D^{\varepsilon}M^{\varepsilon}}{Y^{1-\varepsilon}} + X^{\frac{1}{2}+\varepsilon}D^{1+\varepsilon}M^{1+\varepsilon}Y^{\frac{3}{2}+\varepsilon}\right).
\end{aligned} \tag{3.6.7.1}$$

Recall the definition (3.2.4) of M. Also, recall the definitions (3.4.9) and (3.4.2) of D and R of D, respectively. So that the error terms in (3.6.7.1) are $O(X^{1-\varepsilon})$, we assume the parameters θ and ϑ satisfy

$$\theta + 2\vartheta < \frac{1}{2},$$

and we take the parameter Y in (3.2.6) to be

$$Y = X^{\delta}$$

with $\delta = \delta(\theta,\vartheta)$ sufficiently small.

3.6.8 Evaluating the sums of the other terms with $k \neq 0$

The procedure for evaluating the sum with $\mathcal{Q}_2$ in (3.6.4.1) is largely similar to the above process for $\mathcal{Q}_1$, with only a few differences. The main difference arises from the negative sign in the character $\left(\frac{-2d_1}{m_1m_2\nu}\right)$ in (3.6.4.3). This causes the residues in the versions of (3.6.5.6) and (3.6.5.7) for $\mathcal{Q}_2$ to have each $-kd_1$ equal to a perfect square instead of $kd_1 = \square$. This means $\operatorname{sgn}(k) = -1$. Hence, because of the factor $\operatorname{sgn}(\xi)$ in

(3.6.5.5), the version of (3.6.5.9) for $\mathcal{Q}_2$ has the function

$$(2\pi)^{-u}\Gamma(u)(\cos\left(\tfrac{\pi}{2}u\right) + i\sin\left(\tfrac{\pi}{2}u\right))$$

in place of the function $\Gamma_2(u)$ defined by (3.6.5.10). These lead to a version of (3.6.7.1) for $\mathcal{Q}_2$ that we may combine with (3.6.7.1) using the identity

$$\left(\frac{1+i}{2}\right)(\cos u - i\sin u) + \left(\frac{1-i}{2}\right)(\cos u + i\sin u) = \cos u + \sin u. \tag{3.6.8.1}$$

The result is

$$\begin{aligned}
&\frac{X}{(\sqrt{2}-1)^4}\sum_{\substack{d\le D\\ d \text{ odd}}}\mu^2(d)\lambda_d \sum\sum_{\substack{m_1,m_2\le M\\ (m_1m_2,2d)=1}}\frac{b_{m_1}b_{m_2}}{(m_1m_2)^{3/2}}\\
&\times\sum_{\substack{\nu=1\\ (\nu,2d)=1}}^{\infty}\frac{d_2(\nu)}{\nu^{3/2}}\sum_{\substack{\alpha\le Y\\ (\alpha,2m_1m_2\nu)=1}}\frac{\mu(\alpha)}{\alpha^2 d_1}(\mathcal{Q}_1+\mathcal{Q}_2)\\
&=\frac{\sqrt{2}X}{2(\sqrt{2}-1)^4}\sum_{\substack{d\le D\\ d \text{ odd}}}\mu^2(d)\lambda_d \sum\sum_{\substack{m_1,m_2\le M\\ (m_1m_2,2d)=1}}\frac{b_{m_1}b_{m_2}}{(m_1m_2)^{3/2}}\operatorname*{Res}_{w=0} X^w\frac{1}{2\pi i}\int_{\left(\frac{1}{\log X}\right)}\Gamma_1(s-w)\\
&\times 8^{s-w}\left(1-\frac{1}{4^{s-w}}\right)\mathcal{K}(s,w;m_1m_2,d)\,\frac{ds}{s} + O(X^{1-\varepsilon}),
\end{aligned} \tag{3.6.8.2}$$

where

$$\Gamma_1(u) = (2\pi)^{-u}\Gamma(u)(\cos\left(\tfrac{\pi}{2}u\right) + \sin\left(\tfrac{\pi}{2}u\right)) \tag{3.6.8.3}$$

and the bound $O(X^{1-\varepsilon})$ for the error term is guaranteed by the conditions in Subsection 3.6.7.

The evaluation of the sums in (3.6.4.1) with $\mathcal{Q}_3$ and $\mathcal{Q}_4$ defined by (3.6.4.4) and (3.6.4.5) is similar. The version of (3.6.5.7) for $\mathcal{Q}_3$ has an extra -1 factor because the Kronecker symbol $\left(\frac{-2}{kd_1}\right)$ equals -1 when $-kd_1$ is an odd perfect square. The resulting expression for the sums in (3.6.4.1) with $\mathcal{Q}_3$ and $\mathcal{Q}_4$ is exactly the same as the right-hand side of (3.6.8.2). Therefore

$$\frac{X}{(\sqrt{2}-1)^4}\sum_{\substack{d\le D\\ d \text{ odd}}}\mu^2(d)\lambda_d \sum\sum_{\substack{m_1,m_2\le M\\ (m_1m_2,2d)=1}}\frac{b_{m_1}b_{m_2}}{(m_1m_2)^{3/2}}\sum_{\substack{\nu=1\\ (\nu,2d)=1}}^{\infty}\frac{d_2(\nu)}{\nu^{3/2}}\sum_{\substack{\alpha\le Y\\ (\alpha,2m_1m_2\nu)=1}}\frac{\mu(\alpha)}{\alpha^2 d_1}\sum_{j=1}^{4}\mathcal{Q}_j$$

$$= \frac{\sqrt{2}X}{(\sqrt{2}-1)^4}\sum_{\substack{d\le D\\ d \text{ odd}}}\mu^2(d)\lambda_d \sum\sum_{\substack{m_1,m_2\le M\\ (m_1m_2,2d)=1}}\frac{b_{m_1}b_{m_2}}{(m_1m_2)^{3/2}}\operatorname*{Res}_{w=0} X^w\frac{1}{2\pi i}\int_{\left(\frac{1}{\log X}\right)}\Gamma_1(s-w)$$

$$\times 8^{s-w}\left(1-\frac{1}{4^{s-w}}\right)\mathcal{K}(s,w;m_1m_2,d)\,\frac{ds}{s}+O(X^{1-\varepsilon}). \tag{3.6.8.4}$$

To estimate the sum with $\mathcal{U}_1$ in (3.6.4.1), we first relabel k in (3.6.4.6) as $2k$, now with k odd, to write

$$\mathcal{U}_1=\left(\frac{1+i}{2}\right)\left(\frac{2d_1}{m_1m_2\nu}\right)\sum_{\substack{k\in\mathbb{Z}\\ k \text{ odd}}} e\left(\frac{k\overline{\alpha^2d_1m_1m_2\nu}}{4}\right)\hat{F}_\nu\left(\frac{kX}{4\alpha^2d_1m_1m_2\nu}\right)G_{2k}(m_1m_2\nu). \tag{3.6.8.5}$$

From the definition (3.3.2) of $G_k(n)$, we see that $G_{2k}(n)=\left(\frac{2}{n}\right)G_k(n)$ for all odd integers n. Also, the orthogonality of Dirichlet characters modulo 4 implies that $e(\frac{h}{4})=i(\frac{-1}{h})$ for odd h. It follows from these and (3.6.8.5) that

$$\mathcal{U}_1=i\left(\frac{1+i}{2}\right)\left(\frac{-d_1}{m_1m_2\nu}\right)\sum_{\substack{k\in\mathbb{Z}\\ k \text{ odd}}}\left(\frac{-1}{kd_1}\right)\hat{F}_\nu\left(\frac{kX}{4\alpha^2d_1m_1m_2\nu}\right)G_k(m_1m_2\nu).$$

We then proceed as we did for $\mathcal{Q}_1$. We treat the sum with $\mathcal{U}_2$, defined by (3.6.4.7), in a similar way. We combine the resulting expressions using the identity (3.6.8.1), and

we arrive at

$$\frac{X}{(\sqrt{2}-1)^4}\sum_{\substack{d\le D\\ d\text{ odd}}}\mu^2(d)\lambda_d\sum\sum_{\substack{m_1,m_2\le M\\ (m_1m_2,2d)=1}}\frac{b_{m_1}b_{m_2}}{(m_1m_2)^{3/2}}$$
$$\times\sum_{\substack{\nu=1\\ (\nu,2d)=1}}^{\infty}\frac{d_2(\nu)}{\nu^{3/2}}\sum_{\substack{\alpha\le Y\\ (\alpha,2m_1m_2\nu)=1}}\frac{\mu(\alpha)}{\alpha^2 d_1}(\mathcal{U}_1+\mathcal{U}_2)$$
$$=\frac{X}{(\sqrt{2}-1)^4}\sum_{\substack{d\le D\\ d\text{ odd}}}\mu^2(d)\lambda_d\sum\sum_{\substack{m_1,m_2\le M\\ (m_1m_2,2d)=1}}\frac{b_{m_1}b_{m_2}}{(m_1m_2)^{3/2}}\operatorname*{Res}_{w=0}X^w\frac{1}{2\pi i}\int_{\left(\frac{1}{\log X}\right)}\Gamma_1(s-w)$$
$$\times 4^{s-w}\left(1-\frac{1}{4^{s-w}}\right)\mathcal{K}(s,w;m_1m_2,d)\,\frac{ds}{s}+O(X^{1-\varepsilon}). \tag{3.6.8.6}$$

Next, to evaluate the sum with $\mathcal{V}$ in (3.6.4.1), we relabel k in (3.6.4.8) as $4k$, now with k odd, to see that

$$\mathcal{V}=-\left(\frac{2d_1}{m_1m_2\nu}\right)\sum_{\substack{k\in\mathbb{Z}\\ k\text{ odd}}}\hat{F}_\nu\left(\frac{kX}{2\alpha^2d_1m_1m_2\nu}\right)\tau_k(m_1m_2\nu)$$

since $e(h/2)=-1$ for odd h and $\tau_{4k}(n)=\tau_k(n)$ for odd n by (3.3.3). Into this we insert the second expression for $\tau_k(n)$ in (3.3.3). Since $\left(\frac{-1}{n}\right)G_k(n)=G_{-k}(n)$ by (3.3.2), we may split our sum expression for $\mathcal{V}$ into two, one with $G_k(n)$ and the other with $G_{-k}(n)$. We relabel k as $-k$ in the latter and combine the result with the former to arrive at

$$\mathcal{V}=-\left(\frac{2d_1}{m_1m_2\nu}\right)\sum_{\substack{k\in\mathbb{Z}\\ k\text{ odd}}}\tilde{F}_\nu\left(\frac{kX}{2\alpha^2d_1m_1m_2\nu}\right)G_k(m_1m_2\nu), \tag{3.6.8.7}$$

where $\tilde{F}(\xi)$ is defined by

$$\tilde{F}(\xi)=\frac{1+i}{2}\hat{F}(\xi)+\frac{1-i}{2}\hat{F}(-\xi)=\int_{-\infty}^{\infty}(\cos(2\pi\xi x)+\sin(2\pi\xi x))F(x)\,dx.$$

We then proceed as we did for $\mathcal{Q}_1$, using [51, Lemma 5.2] instead of Lemma 3.6.3. We arrive at versions of (3.6.5.6), (3.6.5.7), and (3.6.5.8) which show that the residue at

$w = 0$ equals zero because $2kd_1 \neq \square$ when kd_1 is odd. This leads to

$$\frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \le D \\ d \text{ odd}}} \mu^2(d)\lambda_d \sum\sum_{\substack{m_1,m_2 \le M \\ (m_1m_2,2d)=1}} \frac{b_{m_1}b_{m_2}}{(m_1m_2)^{3/2}} \times \sum_{\substack{\nu=1 \\ (\nu,2d)=1}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2}} \sum_{\substack{\alpha \le Y \\ (\alpha,2m_1m_2\nu)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \mathcal{V} \ll X^{1-\varepsilon} \tag{3.6.8.8}$$

under the conditions in Subsection 3.6.7.

Lastly, to estimate the sum with $\mathcal{W}$ in (3.6.4.1), we relabel k in (3.6.4.9) as $8k$ to write

$$\mathcal{W} = \left(\frac{d_1}{m_1m_2\nu}\right) \sum_{\substack{k \in \mathbb{Z} \\ k \neq 0}} \hat{F}_\nu\left(\frac{kX}{\alpha^2 d_1 m_1 m_2 \nu}\right) \tau_k(m_1m_2\nu)$$

using the fact that $e(h) = 1$ for any integer h and $\tau_{8k}(n) = (\frac{2}{n})\tau_k(n)$ for odd n by (3.3.3). Into this we insert the second expression for $\tau_k(n)$ in (3.3.3), apply $\left(\frac{-1}{n}\right) G_k(n) = G_{-k}(n)$, and recombine the k and $-k$ terms as we did for $\mathcal{V}$ in (3.6.8.7) to deduce that

$$\mathcal{W} = \left(\frac{d_1}{m_1m_2\nu}\right) \sum_{\substack{k \in \mathbb{Z} \\ k \neq 0}} \tilde{F}_\nu\left(\frac{kX}{\alpha^2 d_1 m_1 m_2 \nu}\right) G_k(m_1m_2\nu).$$

We then proceed as we did for $\mathcal{Q}_1$, using [51, Lemma 5.2] instead of Lemma 3.6.3. Since we are now summing over all nonzero integers k and not just the odd ones, instead of (3.6.5.11) we use

$$\begin{aligned}\sum_{j=1}^{\infty} j^{-2s+2w} \mathcal{G}(1+w; j^2, m_1m_2, \alpha d) &= \prod_p \sum_{b=0}^{\infty} p^{2b(w-s)} \mathcal{G}_p(1+w; p^{2b}, m_1m_2, \alpha d) \\ &= (m_1m_2)^{1-s+w} \ell_1^{s-w-\frac{1}{2}} \zeta(2s-2w)\zeta(2s+1)\mathcal{H}_1(s-w, 1+w; m_1m_2, \alpha d).\end{aligned}$$

We arrive at

$$\frac{X}{(\sqrt{2}-1)^4}\sum_{\substack{d\le D\\ d\text{ odd}}}\mu^2(d)\lambda_d\sum\sum_{\substack{m_1,m_2\le M\\ (m_1m_2,2d)=1}}\frac{b_{m_1}b_{m_2}}{(m_1m_2)^{3/2}}$$

$$\times\sum_{\substack{\nu=1\\ (\nu,2d)=1}}^{\infty}\frac{d_2(\nu)}{\nu^{3/2}}\sum_{\substack{\alpha\le Y\\ (\alpha,2m_1m_2\nu)=1}}$$

$$=\frac{X}{(\sqrt{2}-1)^4}\sum_{\substack{d\le D\\ d\text{ odd}}}\mu^2(d)\lambda_d\sum\sum_{\substack{m_1,m_2\le M\\ (m_1m_2,2d)=1}}\frac{b_{m_1}b_{m_2}}{(m_1m_2)^{3/2}}\operatorname*{Res}_{w=0}X^w\frac{1}{2\pi i}\int_{\left(\frac{1}{\log X}\right)}\Gamma_1(s-w)$$
$$\times\mathcal{K}(s,w;m_1m_2,d)\,\frac{ds}{s}+O(X^{1-\varepsilon}).\qquad(3.6.8.9)$$

3.6.9 Putting together the estimates

From (3.6.4.1), (3.6.8.4), (3.6.8.6), (3.6.8.8), and (3.6.8.9), we deduce that

$$\mathcal{B}=\frac{X}{(\sqrt{2}-1)^4}\sum_{\substack{d\le D\\ d\text{ odd}}}\mu^2(d)\lambda_d\sum\sum_{\substack{m_1,m_2\le M\\ (m_1m_2,2d)=1}}\frac{b_{m_1}b_{m_2}}{(m_1m_2)^{3/2}}\operatorname*{Res}_{w=0}X^w\frac{1}{2\pi i}\int_{\left(\frac{1}{\log X}\right)}\Gamma_1(s-w)$$
$$\times\left(8^{s-w}\sqrt{2}+4^{s-w}-2^{s-w}\sqrt{2}\right)\mathcal{K}(s,w;m_1m_2,d)\,\frac{ds}{s}+O(X^{1-\varepsilon}).$$

We next evaluate the residue at $w=0$. Note that, for fixed s, the integrand has a pole of order at most 2 at $w=0$. We use (3.5.1.6) with $n=2$ to write

$$\mathcal{B}=\frac{X}{(\sqrt{2}-1)^4}\sum_{\substack{d\le D\\ d\text{ odd}}}\mu^2(d)\lambda_d\sum\sum_{\substack{m_1,m_2\le M\\ (m_1m_2,2d)=1}}\frac{b_{m_1}b_{m_2}}{(m_1m_2)^{3/2}}$$
$$\times\frac{1}{2\pi i}\int_{\left(\frac{1}{\log X}\right)}\Gamma_1(s)\left(8^s\sqrt{2}+4^s-2^s\sqrt{2}\right)$$
$$\times\mathcal{K}(s,0;m_1m_2,d)\left\{\log X-\frac{\Gamma_1'(s)}{\Gamma_1(s)}-(\log 2)\frac{3\cdot 8^s\sqrt{2}+2\cdot 4^s-2^s\sqrt{2}}{8^s\sqrt{2}+4^s-2^s\sqrt{2}}\right.\qquad(3.6.9.1)$$
$$\left.+\left.\frac{\frac{\partial}{\partial w}\mathcal{K}(s,w;m_1m_2,d)}{\mathcal{K}(s,w;m_1m_2,d)}\right|_{w=0}\right\}\frac{ds}{s}+O(X^{1-\varepsilon}).$$

From the definitions (3.6.5.19) and (3.6.5.13) of $\mathcal{K}$ and $\mathcal{H}$, we see that, after some simplification,

$$\begin{aligned}
&\left(8^s\sqrt{2}+4^s-2^s\sqrt{2}\right)\mathcal{K}(s,0;m_1m_2,d)\\
&=\frac{\check{\Phi}(0)}{4}\frac{\Gamma\left(\frac{s}{2}+\frac{1}{4}\right)^2}{\Gamma\left(\frac{1}{4}\right)^2}\left(\frac{4}{\pi}\right)^s\zeta(2s)\zeta(2s+1)\left(1-\frac{1}{2^{\frac{1}{2}+s}}\right)\left(1-\frac{1}{2^{\frac{1}{2}-s}}\right)\left(\frac{5}{2}-4^s-4^{-s}\right)\\
&\quad\times\frac{\varphi(dm_1m_2)^2}{d^3m_1m_2\sqrt{\ell_1}}\sum_{ab=\ell_1}\left(\frac{a}{b}\right)^s\prod_{\substack{p|m_1m_2\\ p\nmid\ell_1}}\left(1+\frac{1}{p}\right)\prod_{p|d}\left(1-\frac{1}{p^{1+2s}}\right)\left(1-\frac{1}{p^{1-2s}}\right)\\
&\quad\times\prod_{p\nmid 2m_1m_2d}\left\{\left(1-\frac{1}{p}\right)^2\left(1+\frac{2}{p}+\frac{1}{p^3}-\frac{1}{p^{2-2s}}-\frac{1}{p^{2+2s}}\right)\right\},
\end{aligned}\tag{3.6.9.2}$$

where ℓ_1 is defined by (3.6.5.12), and

$$\begin{aligned}
&-(\log 2)\frac{3\cdot 8^s\sqrt{2}+2\cdot 4^s-2^s\sqrt{2}}{8^s\sqrt{2}+4^s-2^s\sqrt{2}}+\left.\frac{\frac{\partial}{\partial w}\mathcal{K}(s,w;m_1m_2,d)}{\mathcal{K}(s,w;m_1m_2,d)}\right|_{w=0}\\
&=2\gamma+\frac{(\check{\Phi})'(0)}{\check{\Phi}(0)}-\log(2\ell_1)-2\frac{\zeta'}{\zeta}(2s)+2\frac{\zeta'}{\zeta}(2s+1)+\frac{\log 2}{\left(\sqrt{2}+2^s\right)\left(\sqrt{2}+2^{-s}\right)}\\
&\quad+\sum_{p|d}\left(\frac{2\log p}{p-1}+\frac{2\log p}{p^{1+2s}-1}+\frac{2\log p}{p^{1-2s}-1}\right)+\sum_{p|m_1m_2}\frac{2\log p}{p-1}-\sum_{\substack{p|m_1m_2\\ p\nmid\ell_1}}\frac{2\log p}{p+1}\\
&\quad+\sum_{p\nmid 2m_1m_2d}\left(\frac{2\log p}{p-1}-\left(\frac{2\log p}{p}\right)\frac{1+\frac{2}{p^2}-\frac{1}{p}\left(p^{2s}+p^{-2s}\right)}{1+\frac{2}{p}+\frac{1}{p^3}-\frac{1}{p^2}\left(p^{2s}+p^{-2s}\right)}\right).
\end{aligned}\tag{3.6.9.3}$$

Now the definition (3.6.8.3) of $\Gamma_1(u)$, the Legendre duplication formula, the functional equation of $\zeta(s)$, and the identity $\Gamma(z)\Gamma(1-z)=\pi\csc(\pi z)$ imply that the functions

$$\frac{\Gamma^2\left(\frac{s}{2}+\frac{1}{4}\right)}{\Gamma^2\left(\frac{1}{4}\right)}\left(\frac{4}{\pi}\right)^s\Gamma_1(s)\zeta(2s)\zeta(2s+1)$$

and

$$-\frac{\Gamma_1'(s)}{\Gamma_1(s)}-2\frac{\zeta'}{\zeta}(2s)+2\frac{\zeta'}{\zeta}(2s+1)$$

are even functions of s. Hence (3.6.9.2) and (3.6.9.3) are even functions of s. It follows that the integrand in (3.6.9.1) is an odd function of s. We move the line of integration in (3.6.9.1) to $\mathrm{Re}(s) = -\frac{1}{\log X}$, leaving a residue at $s = 0$. In the new integral, we make a change of variables $s \mapsto -s$ to see that, since its integrand is odd, it equals the negative of the original integral in (3.6.9.1). Therefore twice the original integral equals the residue at $s = 0$. We write this residue as an integral along the circle $|s| = \frac{1}{\log X}$, taken in the positive direction, and arrive at

$$\begin{aligned}\mathcal{B} = &\frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \le D \\ d \text{ odd}}} \mu^2(d)\lambda_d \sum\sum_{\substack{m_1,m_2 \le M \\ (m_1m_2,2d)=1}} \frac{b_{m_1}b_{m_2}}{(m_1m_2)^{3/2}} \\ &\times \frac{1}{4\pi i} \oint_{|s|=\frac{1}{\log X}} \Gamma_1(s)\Big(8^s\sqrt{2} + 4^s - 2^s\sqrt{2}\Big) \\ &\times \mathcal{K}(s,0;m_1m_2,d)\left\{ \log X - \frac{\Gamma_1'(s)}{\Gamma_1(s)} - (\log 2)\frac{3\cdot 8^s\sqrt{2} + 2\cdot 4^s - 2^s\sqrt{2}}{8^s\sqrt{2}+4^s-2^s\sqrt{2}} \right. \\ &\left. + \left.\frac{\frac{\partial}{\partial w}\mathcal{K}(s,w;m_1m_2,d)}{\mathcal{K}(s,w;m_1m_2,d)}\right|_{w=0} \right\} \frac{ds}{s} + O(X^{1-\varepsilon}).\end{aligned} \tag{3.6.9.4}$$

The next step is to carry out the summation over d. From (3.6.9.2) and (3.6.9.3), we see that we need to evaluate the sums Σ_1 and Σ_2 defined by

$$\begin{aligned}\Sigma_1 = &\sum_{\substack{d \le D \\ (d,2m_1m_2)=1}} \mu^2(d)\lambda_d \frac{\varphi(d)^2}{d^3} \prod_{p|d}\left(1 - \frac{1}{p^{1+2s}}\right)\left(1 - \frac{1}{p^{1-2s}}\right) \\ &\times \prod_{p \nmid 2m_1m_2d} \left\{ \left(1-\frac{1}{p}\right)^2 \left(1 + \frac{2}{p} + \frac{1}{p^3} - \frac{1}{p^{2-2s}} - \frac{1}{p^{2+2s}}\right)\right\}\end{aligned} \tag{3.6.9.5}$$

and

$$\begin{aligned}\Sigma_2 = &\sum_{\substack{d \le D \\ (d,2m_1m_2)=1}} \mu^2(d)\lambda_d \frac{\varphi(d)^2}{d^3} \prod_{p|d}\left(1 - \frac{1}{p^{1+2s}}\right)\left(1 - \frac{1}{p^{1-2s}}\right) \\ &\times \prod_{p \nmid 2m_1m_2d} \left\{ \left(1-\frac{1}{p}\right)^2 \left(1 + \frac{2}{p} + \frac{1}{p^3} - \frac{1}{p^{2-2s}} - \frac{1}{p^{2+2s}}\right)\right\} \sum_{p|d} J(p,s),\end{aligned} \tag{3.6.9.6}$$

where

$$J(p,s) = \frac{2\log p}{p^{1+2s}-1} + \frac{2\log p}{p^{1-2s}-1} + \left(\frac{2\log p}{p}\right) \frac{1+\frac{2}{p^2}-\frac{1}{p}\left(p^{2s}+p^{-2s}\right)}{1+\frac{2}{p}+\frac{1}{p^3}-\frac{1}{p^2}\left(p^{2s}+p^{-2s}\right)} \tag{3.6.9.7}$$

and $|s| = \frac{1}{\log X}$. We only estimate Σ_1 since Σ_2 may be treated in the same way, except using Lemma 3.4.4 instead of Lemma 3.4.3. We rearrange the factors in (3.6.9.5) to write Σ_1 as

$$\begin{aligned}\Sigma_1 = \prod_{p \nmid 2m_1m_2} &\left\{\left(1-\frac{1}{p}\right)^2\left(1+\frac{2}{p}+\frac{1}{p^3}-\frac{1}{p^{2-2s}}-\frac{1}{p^{2+2s}}\right)\right\} \sum_{\substack{d\le D\\(d,2m_1m_2)=1}} \frac{\mu^2(d)\lambda_d}{d} \\ &\times \prod_{p|d}\left(1-\frac{1}{p^{1+2s}}\right)\left(1-\frac{1}{p^{1-2s}}\right)\left(1+\frac{2}{p}+\frac{1}{p^3}-\frac{1}{p^{2-2s}}-\frac{1}{p^{2+2s}}\right)^{-1}.\end{aligned} \tag{3.6.9.8}$$

Now recall the definition (3.4.1) of z_0 and the definition (3.4.8) of λ_d. Factoring out the product over primes $p > z_0$, we see that

$$\begin{aligned}&\prod_{p\nmid 2m_1m_2}\left\{\left(1-\frac{1}{p}\right)^2\left(1+\frac{2}{p}+\frac{1}{p^3}-\frac{1}{p^{2-2s}}-\frac{1}{p^{2+2s}}\right)\right\} \\ &\quad = \left(1+O\left(\frac{1}{z_0}\right)\right)\prod_{\substack{p\nmid 2m_1m_2\\p\le z_0}}\left\{\left(1-\frac{1}{p}\right)^2\left(1+\frac{2}{p}+\frac{1}{p^3}-\frac{1}{p^{2-2s}}-\frac{1}{p^{2+2s}}\right)\right\}.\end{aligned}$$

From this, (3.6.9.8), Lemma 3.4.3, and some simplification, we deduce that

$$\begin{aligned}\Sigma_1 = &\left(1+O\left(\frac{1}{z_0}\right)\right)\frac{1+o(1)}{\log R}\prod_{\substack{p\nmid 2m_1m_2\\p\le z_0}}\left(1-\frac{1}{p^2}\right)\prod_{\substack{p|2m_1m_2\\p\le z_0}}\left(1-\frac{1}{p}\right)^{-1} \\ &+O\left(\frac{1}{(\log R)^{2019}}\right).\end{aligned} \tag{3.6.9.9}$$

The condition $p \le z_0$ may be omitted because $\prod_{p>z_0}(1+O(\frac{1}{p^2})) = 1+O(\frac{1}{z_0})$ and

$$\prod_{\substack{p|2m_1m_2\\p>z_0}}\left(1-\frac{1}{p}\right)^{-1} = \left(1+O\left(\frac{1}{z_0}\right)\right)^{O(\log X)} = 1+O\left(\frac{\log X}{z_0}\right).$$

The contributions of the error terms $O\left(\frac{1}{z_0}\right)$ and $O\left(\frac{\log X}{z_0}\right)$ are negligible. From these and (3.6.9.9), we arrive at

$$\Sigma_1 = \frac{2m_1m_2}{\varphi(m_1m_2)} \prod_{p\nmid 2m_1m_2} \left(1-\frac{1}{p^2}\right) \frac{1+o(1)}{\log R} + O\left(\frac{1}{(\log R)^{2019}}\right). \tag{3.6.9.10}$$

In a similar way, but using Lemma 3.4.4 instead of Lemma 3.4.3, we deduce from (3.6.9.6) that

$$\begin{aligned}\Sigma_2 = &-\frac{2m_1m_2}{\varphi(m_1m_2)} \prod_{p\nmid 2m_1m_2} \left(1-\frac{1}{p^2}\right) \frac{1+o(1)}{\log R} \\ &\times \sum_{p\nmid 2m_1m_2} \frac{J(p,s)}{p+1}\left(1-\frac{1}{p^{1+2s}}\right)\left(1-\frac{1}{p^{1-2s}}\right) + O\left(\frac{1}{(\log R)^{2019}}\right).\end{aligned} \tag{3.6.9.11}$$

In view of the expressions (3.6.9.2) and (3.6.9.3) and the definitions (3.6.9.5) and (3.6.9.6), it now follows from (3.6.9.4), (3.6.9.10) and (3.6.9.11) that

$$\begin{aligned}\mathcal{B} = &\frac{X\check{\Phi}(0)}{3\zeta(2)(\sqrt{2}-1)^4}\frac{1+o(1)}{\log R} \sum\sum_{\substack{m_1,m_2\le M\\(m_1m_2,2)=1}} \frac{b_{m_1}b_{m_2}}{\sqrt{m_1m_2\ell_1}} \prod_{p\mid \ell_1}\left(\frac{p}{p+1}\right) \\ &\times \frac{1}{2\pi i}\oint_{|s|=\frac{1}{\log X}} \sum_{ab=\ell_1}\left(\frac{a}{b}\right)^s \Gamma_1(s)\frac{\Gamma\left(\frac{s}{2}+\frac14\right)^2}{\Gamma\left(\frac14\right)^2}\left(\frac{4}{\pi}\right)^s \zeta(2s)\zeta(2s+1) \\ &\times \left(1-\frac{1}{2^{\frac12+s}}\right)\left(1-\frac{1}{2^{\frac12-s}}\right)\left(\frac52-4^s-4^{-s}\right)\Bigg\{\log\left(\frac{X}{2\ell_1}\right)+2\gamma+\frac{(\check{\Phi})'(0)}{\check{\Phi}(0)}-\frac{\Gamma_1'(s)}{\Gamma_1(s)} \\ &-2\frac{\zeta'}{\zeta}(2s)+2\frac{\zeta'}{\zeta}(2s+1)+\frac{\log 2}{\left(\sqrt{2}+2^s\right)\left(\sqrt{2}+2^{-s}\right)}+\sum_{p\neq 2}\eta_1(p,s) \\ &+\sum_{\substack{p\mid m_1m_2\\p\nmid \ell_1}}\eta_2(p,s)+\sum_{p\mid \ell_1}\eta_3(p,s)\Bigg\}\frac{ds}{s}+O\left(\frac{X}{(\log R)^{2019}}\right),\end{aligned} \tag{3.6.9.12}$$

where

$$\begin{aligned}\eta_1(p,s) = &\frac{2\log p}{p-1} - \left(\frac{2\log p}{p}\right)\frac{1+\frac{2}{p^2}-\frac1p\left(p^{2s}+p^{-2s}\right)}{1+\frac2p+\frac{1}{p^3}-\frac{1}{p^2}\left(p^{2s}+p^{-2s}\right)} \\ &-\frac{J(p,s)}{p+1}\left(1-\frac{1}{p^{1+2s}}\right)\left(1-\frac{1}{p^{1-2s}}\right),\end{aligned}$$

$$\eta_2(p,s) = \frac{2\log p}{p-1} - \frac{2\log p}{p+1} - \eta_1(p,s), \tag{3.6.9.13}$$

and

$$\eta_3(p,s) = \frac{2\log p}{p-1} - \eta_1(p,s),$$

with $J(p,s)$ defined by (3.6.9.7).

Next, we carry out the summation over m_1, m_2. We see from (3.6.9.12) that we need to evaluate the sums Υ_1, Υ_2, Υ_3, and Υ_4 defined by

$$\Upsilon_1 = \sum_{\substack{m_1,m_2\le M\\(m_1m_2,2)=1}}\sum \frac{b_{m_1}b_{m_2}}{\sqrt{m_1m_2\ell_1}} \prod_{p|\ell_1}\left(\frac{p}{p+1}\right) \sum_{ab=\ell_1}\left(\frac{a}{b}\right)^s, \tag{3.6.9.14}$$

$$\Upsilon_2 = -\sum_{\substack{m_1,m_2\le M\\(m_1m_2,2)=1}}\sum \frac{b_{m_1}b_{m_2}}{\sqrt{m_1m_2\ell_1}} \prod_{p|\ell_1}\left(\frac{p}{p+1}\right) \sum_{ab=\ell_1}\left(\frac{a}{b}\right)^s \log\ell_1, \tag{3.6.9.15}$$

$$\Upsilon_3 = \sum_{\substack{m_1,m_2\le M\\(m_1m_2,2)=1}}\sum \frac{b_{m_1}b_{m_2}}{\sqrt{m_1m_2\ell_1}} \prod_{p|\ell_1}\left(\frac{p}{p+1}\right) \sum_{ab=\ell_1}\left(\frac{a}{b}\right)^s \sum_{p|\ell_1}\eta_2(p,s), \tag{3.6.9.16}$$

and

$$\Upsilon_4 = \sum_{\substack{m_1,m_2\le M\\(m_1m_2,2)=1}}\sum \frac{b_{m_1}b_{m_2}}{\sqrt{m_1m_2\ell_1}} \prod_{p|\ell_1}\left(\frac{p}{p+1}\right) \sum_{ab=\ell_1}\left(\frac{a}{b}\right)^s \sum_{\substack{p|m_1m_2\\p\nmid\ell_1}}\eta_3(p,s), \tag{3.6.9.17}$$

with $|s| = \frac{1}{\log X}$.

To estimate Υ_1, observe that if m_1 and m_2 are square-free then (3.6.5.12) implies

$$\ell_1 = \frac{m_1m_2}{(m_1,m_2)^2} \tag{3.6.9.18}$$

and

$$\sum_{ab=\ell_1}\left(\frac{a}{b}\right)^s = \prod_{p|\ell_1}(p^s+p^{-s}). \tag{3.6.9.19}$$

From these, the definition (3.2.5) of b_m, and the Fourier inversion formula (3.5.1.2), we deduce from (3.6.9.14) that

$$\Upsilon_1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(z_1)h(z_2) \sum\sum_{(m_1m_2,2)=1} \frac{\mu(m_1)\mu(m_2)(m_1,m_2)}{m_1^{1+\frac{1+iz_1}{\log M}} m_2^{1+\frac{1+iz_2}{\log M}}} \times \prod_{\substack{p|m_1m_2\\ p\nmid(m_1,m_2)}} (p^s+p^{-s}) \left(\frac{p}{p+1}\right) dz_1dz_2.$$

Thus, writing the sum as an Euler product, we see that

$$\Upsilon_1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(z_1)h(z_2) \prod_{p>2} \left(1 - \frac{1}{p^{1+\frac{1+iz_1}{\log M}}}(p^s+p^{-s})\left(\frac{p}{p+1}\right) - \frac{1}{p^{1+\frac{1+iz_2}{\log M}}}(p^s+p^{-s})\left(\frac{p}{p+1}\right) + \frac{1}{p^{1+\frac{2+iz_1+iz_2}{\log M}}}\right) dz_1dz_2.$$

We write this as

$$\Upsilon_1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{h(z_1)h(z_2)\zeta\left(1+\frac{2+iz_1+iz_2}{\log M}\right) W(s,z_1,z_2,\frac{1}{\log M})\,dz_1dz_2}{\zeta\left(1+\frac{1+iz_1}{\log M}+s\right)\zeta\left(1+\frac{1+iz_1}{\log M}-s\right)\zeta\left(1+\frac{1+iz_2}{\log M}+s\right)\zeta\left(1+\frac{1+iz_2}{\log M}-s\right)}, \tag{3.6.9.20}$$

where $W(s,z_1,z_2,\frac{1}{\log M})$ is an Euler product that is bounded and holomorphic for $|s|\le\varepsilon$ and complex z_1,z_2 with $|\mathrm{Im}(z_1)|,|\mathrm{Im}(z_2)|\le\varepsilon\log M$. Note that this definition of W implies

$$W(0,0,0,0) = 8\prod_{p>2}\left(1-\frac{4}{p+1}+\frac{1}{p}\right)\left(1-\frac{1}{p}\right)^{-3} = 6\zeta(2), \tag{3.6.9.21}$$

a fact we use shortly. By (3.5.1.4), we may truncate the integrals in (3.6.9.20) to the range $|z_1|,|z_2|\le\sqrt{\log M}$, introducing a negligible error. On this range of z_1 and z_2, the function W and the zeta-functions in (3.6.9.20) may be written as Laurent series. The contributions of the terms other than the first terms of these Laurent expansions are a factor of $(\log X)^{1-\varepsilon}$ smaller than the contribution of the first terms. The first term of the Laurent expansion of W is given by (3.6.9.21). We thus arrive at

$$\Upsilon_1 = 6\zeta(2) \iint\limits_{|z_i|\le\sqrt{\log M}} h(z_1)h(z_2)\left(\frac{\log M}{2+iz_1+iz_2}\right)\left(\frac{1+iz_1}{\log M}-s\right)\left(\frac{1+iz_1}{\log M}+s\right)$$
$$\times\left(\frac{1+iz_2}{\log M}+s\right)\left(\frac{1+iz_2}{\log M}-s\right)dz_1dz_2+O\left(\frac{1}{(\log X)^{4-\varepsilon}}\right).$$

By (3.5.1.4), we may extend the range of integration to $\mathbb{R}^2$, introducing a negligible error. We then apply (3.6.3.12) to deduce that

$$\Upsilon_1 = 6\zeta(2)\Bigg(\frac{1}{\log^3 M}\int_0^1 H''(t)^2\,dt - \frac{2s^2}{\log M}\int_0^1 H(t)H''(t)\,dt + s^4\log M\int_0^1 H(t)^2\,dt\Bigg)+O\left(\frac{1}{(\log X)^{4-\varepsilon}}\right). \tag{3.6.9.22}$$

Having evaluated Υ_1, we next estimate Υ_2. Using the residue theorem, we write

$$-\log \ell_1 = \frac{1}{2\pi i}\oint_{|y|=\frac{1}{2\log X}} \ell_1^{-y}\,\frac{dy}{y^2}.$$

From this, (3.6.9.15), (3.6.9.18), (3.6.9.19), the definition (3.2.5) of b_m, and the Fourier inversion formula (3.5.1.2), it follows that

$$\Upsilon_2 = \frac{1}{2\pi i}\oint_{|y|=\frac{1}{2\log X}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(z_1)h(z_2)\sum_{(m_1m_2,2)=1}\sum \frac{\mu(m_1)\mu(m_2)(m_1,m_2)^{1+2y}}{m_1^{1+\frac{1+iz_1}{\log M}+y}m_2^{1+\frac{1+iz_2}{\log M}+y}}$$
$$\times\prod_{\substack{p\mid m_1m_2\\ p\nmid(m_1,m_2)}}(p^s+p^{-s})\left(\frac{p}{p+1}\right)dz_1dz_2\frac{dy}{y^2}.$$

We express the sum as an Euler product to see that

$$\Upsilon_2 = \frac{1}{2\pi i}\oint_{|y|=\frac{1}{2\log X}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(z_1)h(z_2)\prod_{p>2}\Bigg(1-\frac{1}{p^{1+\frac{1+iz_1}{\log M}+y}}(p^s+p^{-s})\left(\frac{p}{p+1}\right)$$
$$-\frac{1}{p^{1+\frac{1+iz_2}{\log M}+y}}(p^s+p^{-s})\left(\frac{p}{p+1}\right)+\frac{1}{p^{1+\frac{2+iz_1+iz_2}{\log M}}}\Bigg)dz_1dz_2\,\frac{dy}{y^2}.$$

Write this as

$$\Upsilon_2 = \frac{1}{2\pi i}\oint_{|y|=\frac{1}{2\log X}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(z_1)h(z_2)\zeta\left(1+\tfrac{2+iz_1+iz_2}{\log M}\right) V(s,z_1,z_2,\tfrac{1}{\log M},y)$$
$$\times \zeta^{-1}\left(1+\tfrac{1+iz_1}{\log M}+y+s\right)\zeta^{-1}\left(1+\tfrac{1+iz_1}{\log M}+y-s\right)$$
$$\times \zeta^{-1}\left(1+\tfrac{1+iz_2}{\log M}+y+s\right)\zeta^{-1}\left(1+\tfrac{1+iz_2}{\log M}+y-s\right)\, dz_1 dz_2 \frac{dy}{y^2},$$

where $V(s,z_1,z_2,\frac{1}{\log M},y)$ is an Euler product that is bounded and holomorphic for $|s|,|y| \le \varepsilon$ and complex z_1,z_2 with $|\mathrm{Im}(z_1)|,|\mathrm{Im}(z_2)| \le \varepsilon \log M$. This definition of V implies that $V(0,0,0,0,0) = 6\zeta(2)$. As in our treatment of Υ_1, we use (3.5.1.4) to truncate the integrals. Then we write the function V and the zeta-functions as Laurent series. The main contribution arises from the first terms of the Laurent expansions, and we arrive at

$$\Upsilon_2 = \frac{6\zeta(2)}{2\pi i}\oint_{|y|=\frac{1}{2\log X}} \iint\limits_{|z_i|\le\sqrt{\log M}} h(z_1)h(z_2)\left(\frac{\log M}{2+iz_1+iz_2}\right)\left(\frac{1+iz_1}{\log M}+y-s\right)$$
$$\times\left(\frac{1+iz_1}{\log M}+y+s\right)\left(\frac{1+iz_2}{\log M}+y+s\right)\left(\frac{1+iz_2}{\log M}+y-s\right) dz_1 dz_2 \frac{dy}{y^2}$$
$$+O\left(\frac{1}{(\log X)^{3-\varepsilon}}\right).$$

We carry out the integration over y by applying the formula (3.5.1.6) with $n=2$ and deduce that

$$\Upsilon_2 = 6\zeta(2) \iint\limits_{|z_i|\le\sqrt{\log M}} h(z_1)h(z_2)\left(\frac{\log M}{2+iz_1+iz_2}\right)$$
$$\times\Bigg\{\left(\frac{1+iz_1}{\log M}+s\right)\left(\frac{(1+iz_2)^2}{(\log M)^2}-s^2\right)+\left(\frac{1+iz_1}{\log M}-s\right)\left(\frac{(1+iz_2)^2}{(\log M)^2}-s^2\right)$$
$$+\left(\frac{1+iz_2}{\log M}+s\right)\left(\frac{(1+iz_1)^2}{(\log M)^2}-s^2\right)$$
$$+\left(\frac{1+iz_2}{\log M}-s\right)\left(\frac{(1+iz_1)^2}{(\log M)^2}-s^2\right)\Bigg\} dz_1 dz_2 + O\left(\frac{1}{(\log X)^{3-\varepsilon}}\right).$$

We extend the integral and apply (3.6.3.12). After simplifying, we arrive at

$$\begin{aligned}\Upsilon_2 = 6\zeta(2)\left(-\frac{4}{\log^2 M}\int_0^1 H'(t)H''(t)\,dt + 4s^2\int_0^1 H(t)H'(t)\,dt\right)\\ + O\left(\frac{1}{(\log X)^{3-\varepsilon}}\right).\end{aligned} \tag{3.6.9.23}$$

We next estimate Υ_3 defined by (3.6.9.16). We interchange the order of summation over m_1, m_2 and over p. From (3.6.9.18), we see for a prime q and square-free m_1 and m_2 that $q|\ell_1$ if and only if q divides exactly one of m_1 or m_2. If q divides m_2 and not m_1, then we may relabel m_1 as m_2 and vice versa. Hence

$$\Upsilon_3 = 2\sum_{2<q\le M}\eta_2(q,s)\sum\sum_{\substack{m_1,m_2\le M\\ (m_1m_2,2)=1\\ q|m_1,\ q\nmid m_2}}\frac{b_{m_1}b_{m_2}}{\sqrt{m_1m_2\ell_1}}\prod_{p|\ell_1}\left(\frac{p}{p+1}\right)\sum_{ab=\ell_1}\left(\frac{a}{b}\right)^s.$$

From this, the definition (3.2.5) of b_m, (3.6.9.18), and (3.6.9.19), it follows that

$$\begin{aligned}\Upsilon_3 = 2\sum_{2<q\le M}\eta_2(q,s)\sum\sum_{\substack{m_1,m_2\le M\\ (m_1m_2,2)=1\\ q|m_1,\ q\nmid m_2}}\frac{\mu(m_1)\mu(m_2)}{[m_1,m_2]}\prod_{\substack{p|m_1m_2\\ p\nmid(m_1,m_2)}}\left(\frac{p}{p+1}\right)(p^s+p^{-s})\\ \times H\left(\frac{\log m_1}{\log M}\right)H\left(\frac{\log m_2}{\log M}\right).\end{aligned}$$

We relabel m_1 as qm_1 to write this as

$$\begin{aligned}\Upsilon_3 = -2\sum_{2<q\le M}\frac{\eta_2(q,s)}{q+1}(q^s+q^{-s})\sum_{m_1\le\frac{M}{q}}\sum_{\substack{m_2\le M\\ (m_1m_2,2q)=1}}\frac{\mu(m_1)\mu(m_2)}{[m_1,m_2]}\\ \times\prod_{\substack{p|m_1m_2\\ p\nmid(m_1,m_2)}}\left(\frac{p}{p+1}\right)(p^s+p^{-s})H\left(\frac{\log qm_1}{\log M}\right)H\left(\frac{\log m_2}{\log M}\right).\end{aligned}$$

We insert the Fourier inversion formula (3.5.1.2), interchange the order of summation, and then write the m_1, m_2-sum as an Euler product to deduce that

$$\Upsilon_3 = -2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sum_{2<q\le M}\frac{\eta_2(q,s)(q^s+q^{-s})}{(q+1)q^{\frac{1+iz_1}{\log M}}}h(z_1)h(z_2)$$
$$\times\prod_{p\nmid 2q}\left(1-\frac{1}{p^{1+\frac{1+iz_1}{\log M}}}(p^s+p^{-s})\left(\frac{p}{p+1}\right)\right.$$
$$\left.-\frac{1}{p^{1+\frac{1+iz_2}{\log M}}}(p^s+p^{-s})\left(\frac{p}{p+1}\right)+\frac{1}{p^{1+\frac{2+iz_1+iz_2}{\log M}}}\right)dz_1dz_2.$$

We may express the Euler product in terms of zeta-functions to write

$$\Upsilon_3 = -2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sum_{2<q\le M}\frac{\eta_2(q,s)(q^s+q^{-s})}{(q+1)q^{\frac{1+iz_1}{\log M}}}h(z_1)h(z_2)$$
$$\times\zeta\left(1+\tfrac{2+iz_1+iz_2}{\log M}\right)\zeta^{-1}\left(1+\tfrac{1+iz_1}{\log M}+s\right)\zeta^{-1}\left(1+\tfrac{1+iz_1}{\log M}-s\right)\zeta^{-1}\left(1+\tfrac{1+iz_2}{\log M}+s\right)$$
$$\times\zeta^{-1}\left(1+\tfrac{1+iz_2}{\log M}-s\right)U_q(s,z_1,z_2,\tfrac{1}{\log M})dz_1dz_2, \tag{3.6.9.24}$$

where $U_q(s,z_1,z_2,\frac{1}{\log M})$ is an Euler product that is uniformly bounded for $2<q\le M$ prime, $|s|\le\varepsilon$, and real z_1,z_2. Using (3.5.1.4), we may truncate the integrals to the range $|z_1|,|z_2|\le\sqrt{\log M}$ and introduce only a negligible error. In this range, and for $|s|=\frac{1}{\log X}$, the quotient of zeta-functions in (3.6.9.24) is $(\log M)^{-3+\varepsilon}$. Moreover, (3.6.9.13) implies $\eta_2(q,s)\ll byq^{-1+\varepsilon}$ for $2<q\le M$ and $|s|=\frac{1}{\log X}$. It thus follows that

$$\Upsilon_3 \ll \frac{1}{(\log X)^{3-\varepsilon}}. \tag{3.6.9.25}$$

A similar argument applies to Υ_4 defined by (3.6.9.17), except we use the fact that, for a prime q, $q|m_1m_2$ and $q\nmid\ell_1$ both hold if and only if q divides both m_1 and m_2, by (3.6.9.18). This leads to

$$\Upsilon_4 \ll \frac{1}{(\log X)^{3-\varepsilon}}. \tag{3.6.9.26}$$

It now follows from (3.6.9.12); the definitions (3.6.9.14) through (3.6.9.17) of Υ_1, Υ_2,

Υ_3, and Υ_4; and the estimates (3.6.9.22), (3.6.9.23), (3.6.9.25), and (3.6.9.26) that

$$\mathcal{B} = \frac{2X\check{\Phi}(0)}{(\sqrt{2}-1)^4}\frac{1+o(1)}{\log R}\,\frac{1}{2\pi i}\oint_{|s|=\frac{1}{\log X}}\Gamma_1(s)\frac{\Gamma\left(\frac{s}{2}+\frac{1}{4}\right)^2}{\Gamma\left(\frac{1}{4}\right)^2}\left(\frac{4}{\pi}\right)^s\zeta(2s)\zeta(2s+1)$$
$$\times\left(1-\frac{1}{2^{\frac{1}{2}+s}}\right)\left(1-\frac{1}{2^{\frac{1}{2}-s}}\right)\left(\frac{5}{2}-4^s-4^{-s}\right)$$
$$\times\left\{\log\left(\frac{X}{2}\right)+2\gamma+\frac{(\check{\Phi})'(0)}{\check{\Phi}(0)}-\frac{\Gamma_1'(s)}{\Gamma_1(s)}-2\frac{\zeta'}{\zeta}(2s)\right.$$
$$\left.+2\frac{\zeta'}{\zeta}(2s+1)+\frac{\log 2}{\left(\sqrt{2}+2^s\right)\left(\sqrt{2}+2^{-s}\right)}+\sum_{p\neq 2}\eta_1(p,s)\right\}\left\{\frac{1}{\log^3 M}\int_0^1 H''(t)^2\,dt\right.$$
$$-\frac{2s^2}{\log M}\int_0^1 H(t)H''(t)\,dt+s^4\log M\int_0^1 H(t)^2\,dt-\frac{4}{\log^2 M}\int_0^1 H'(t)H''(t)\,dt$$
$$\left.+4s^2\int_0^1 H(t)H'(t)\,dt\right\}\frac{ds}{s}+O\left(\frac{X}{(\log X)^{2-\varepsilon}}\right).$$

Evaluating the s-integral as a residue, we deduce that

$$\mathcal{B} = \frac{X\check{\Phi}(0)}{4\left(1-\frac{1}{\sqrt{2}}\right)^2}\frac{1+o(1)}{\log R}\left\{\frac{\log X}{2\log M}\int_0^1 H(t)H''(t)\,dt-\int_0^1 H(t)H'(t)\,dt\right\}$$
$$+O\left(X(\log X)^{-2+\varepsilon}\right).$$

From this, (3.6.3.13), (3.6.2), (3.6.1.1), and (3.6.2.8), it now follows that

$$S^+ = \frac{X}{8\left(1-\frac{1}{\sqrt{2}}\right)^2}\frac{1+o(1)}{\log R}\left\{\frac{1}{24}\left(\frac{\log X}{\log M}\right)^3\int_0^1 H''(t)^2\,dt\right.$$
$$-\frac{1}{2}\left(\frac{\log X}{\log M}\right)^2\int_0^1 H'(t)H''(t)\,dt+\frac{\log X}{\log M}\int_0^1 H(t)H''(t)\,dt$$
$$\left.+\frac{\log X}{\log M}\int_0^1 H'(t)^2\,dt-\ 2\int_0^1 H(t)H'(t)\,dt\right\}$$
$$+O\left(\frac{X}{(\log X)^{2-\varepsilon}}+\frac{X^{1+\varepsilon}}{Y}+X^{\frac{1}{2}+\varepsilon}M\right).$$

The error terms are acceptable by the choices in Subsection 3.6.7, and this yields Proposition 3.6.1.

3.7 Choosing the mollifier: finishing the proof of Theorem 0.2.2

In this section we complete the proof of Theorem 0.2.2 by making an optimal choice for the smooth function $H(x)$ (see (3.2.3),(3.2.5)).

By (3.2.2), Proposition 3.5.1, and Proposition 3.6.1, one derives the inequality

$$\sum_{\substack{p\equiv 1 \pmod 8\\ L(\frac{1}{2},\chi_p)\neq 0}} (\log p)\Phi\left(\frac{p}{X}\right) \geq \frac{X}{(1+\delta_0)8}\cdot\vartheta\frac{\left(H(0)-\frac{1}{2\theta}H'(0)\right)^2}{\mathfrak{I}}, \tag{3.7.1}$$

where $\delta_0 > 0$ is sufficiently small and fixed. We also have the upper bound

$$\sum_{\substack{p\equiv 1 \pmod 8\\ L(\frac{1}{2},\chi_p)\neq 0}} (\log p)\Phi\left(\frac{p}{X}\right) \leq (\log X)\sum_{\substack{X/2<p\leq X\\ p\equiv 1 \pmod 8\\ L(\frac{1}{2},\chi_p)\neq 0}} 1.$$

The right side of (3.7.1) is an increasing function of ϑ, and so ϑ should be as large as possible. The hypotheses of Proposition 3.6.1 allow $\vartheta = \frac{1}{2}(\frac{1}{2}-\theta)-\varepsilon$, and therefore

$$\sum_{\substack{X/2<p\leq X\\ p\equiv 1 \pmod 8\\ L(\frac{1}{2},\chi_p)\neq 0}} 1 \ \geq \frac{X}{(1+2\delta_0)8\log X}\cdot\varrho, \tag{3.7.2}$$

where

$$\varrho := \frac{1}{2}\left(\frac{1}{2}-\theta\right)\frac{\left(H(0)-\frac{1}{2\theta}H'(0)\right)^2}{\mathfrak{I}}.$$

We seek a choice of $H(x)$ which maximizes ϱ.

As $H(x)$ is a smooth function supported in $[-1,1]$, we have $H(1) = H'(1) = 0$. For notational simplicity we set $H(0) = A, -H'(0) = B$. Since

$$\int_0^1 H(x)H'(x)dx = -\frac{1}{2}A^2,$$
$$\int_0^1 H(x)H''(x)dx = AB - \int_0^1 H'(x)^2dx,$$
$$\int_0^1 H'(x)H''(x)dx = -\frac{1}{2}B^2,$$

we have

$$\mathfrak{I} = \left(A + \frac{1}{2\theta}B\right)^2 + \frac{1}{24\theta^3}\int_0^1 H''(x)^2dx.$$

We choose $H(x)$ such that on $[0,1]$ it is a smooth approximation to the optimal function $H_*(x)$ which minimizes the integral

$$\int_0^1 H_*''(x)^2dx \tag{3.7.3}$$

among all $H_1 \in \mathcal{C}^3([0,1])$ satisfying the boundary conditions $H_1(0) = A, -H_1'(0) = B, H_1(1) = H_1'(1) = 0$. We may choose $H(x)$ such that

$$(1+\delta_0)\int_0^1 H_*''(x)^2dx \geq \int_0^1 H''(x)^2dx.$$

By the Euler-Lagrange equation, we find that an $H_*(x)$ which minimizes (3.7.3) must satisfy

$$H_*^{(4)}(x) = 0.$$

Thus, $H_*(x)$ is a polynomial of degree at most three. Recalling the boundary conditions, we find

$$H_*(x) = (2A-B)x^3 + (2B-3A)x^2 - Bx + A.$$

By direct computation we obtain

$$\int_0^1 H_*''(x)^2dx = 3A^2 + (2B-3A)^2,$$

and therefore

$$\varrho \geq \frac{1-O(\delta_0)}{2}\left(\frac{1}{2}-\theta\right)\left(1+\frac{3A^2+(2B-3A)^2}{24\theta^3(A+\frac{1}{2\theta}B)^2}\right)^{-1}.$$

It is now a straightforward, but tedious, calculus exercise to find that

$$A = \frac{B(4\theta+3)}{6(\theta+1)}$$

is an optimal choice. Thus

$$\varrho \geq \frac{1-O(\delta_0)}{2}\left(\frac{1}{2}-\theta\right)\frac{2\theta(3+6\theta+4\theta^2)}{(1+2\theta)^3}. \tag{3.7.4}$$

With this choice of A we have

$$H_*(x) = \frac{2B\theta}{6(\theta+1)}(1-x)^2\left(2+\frac{3}{2\theta}+x\right).$$

Since ϱ is invariant under multiplication of H by scalars, we arrive at the convenient expression

$$H_*(x) = (1-x)^2\left(2+\frac{3}{2\theta}+x\right). \tag{3.7.5}$$

If we set $x=\frac{\log m}{\log M}$ in (3.7.5), we obtain that the mollifier coefficients b_m satisfy

$$b_m \approx \mu(m)\frac{\log^2(M/m)}{\log^2 M}\frac{\log(X^{3/2}M^2m)}{\log M}.$$

One might wish to compare this with the description of $\lambda(\ell)$ in [51, p. 449].

Define

$$\rho(\theta) := \frac{1}{2}\left(\frac{1}{2}-\theta\right)\frac{2\theta(3+6\theta+4\theta^2)}{(1+2\theta)^3} = \frac{1}{2}\left(\frac{1}{2}-\theta\right)\left(1-\frac{1}{(1+2\theta)^3}\right).$$

By (3.7.2) and (3.7.4), we obtain

$$\sum_{\substack{X/2<p\leq X \\ p\equiv 1 \,(\mathrm{mod}\ 8) \\ L(\frac{1}{2},\chi_p)\neq 0}} 1 \;\geq \frac{X}{(1+O(\delta_0))8\log X}\cdot\rho(\theta). \tag{3.7.6}$$

The maximum of $\rho(\theta)$ on $(0, \frac{1}{2})$ occurs at the unique positive root θ_0 of the polynomial $16\theta^4 + 32\theta^3 + 24\theta^2 + 12\theta - 3$. By numerical calculation we find

$$\theta_0 = 0.17409\dots$$

and

$$\rho(\theta_0) = 0.09645\dots. \tag{3.7.7}$$

We then choose $\theta = \theta_0$. Since

$$\sum_{\substack{X/2<p\leq X\\ p\equiv 1 \,(\mathrm{mod}\ 8)}} 1 = (1+o(1))\frac{X}{8(\log X)},$$

we deduce Theorem 0.2.2 from (3.7.6) and (3.7.7) upon summing over dyadic intervals.

3.8 The second moment of $L(\frac{1}{2}, \chi_p)$

In this section we prove Theorems 0.2.3 and 0.2.4. We first consider separately the upper and lower bounds for Theorem 0.2.3.

3.8.1 The upper bound in Theorem 0.2.3

We define

$$M_2 := \sum_{p\equiv 1 \,(\mathrm{mod}\ 8)} (\log p)\Phi\left(\frac{p}{X}\right) L\left(\tfrac{1}{2}, \chi_p\right)^2. \tag{3.8.1.1}$$

In this subsection we prove

$$M_2 \leq (4\mathfrak{c} + o(1))\frac{X}{8}(\log X)^3. \tag{3.8.1.2}$$

The upper bound of Theorem 0.2.3 then follows from (3.8.1.2) upon summation over dyadic intervals.

The proof of (3.8.1.2) follows the lines of the proof of Proposition 3.6.1, taking $M(p) = 1$. We employ positivity to replace $\log p$ by $\log X$ and then introduce an upper bound sieve. After applying the approximate functional equation we split $\mu^2(n) = N_Y(n) + R_Y(n)$, and employ the bound (3.6.1.1).

We follow the argument of Section 3.6 down to (3.6.2.8), obtaining

$$S_N^+ = \mathcal{T}_0 + \mathcal{B}.$$

Since we have no mollifier here, we find

$$\mathcal{T}_0 = \frac{2X}{(\sqrt{2}-1)^4}\frac{1+o(1)}{\log R}\sum_{\substack{\nu=1\\(\nu,2)=1\\\nu=\square}}^{\infty}\frac{d_2(\nu)}{\sqrt{\nu}}\hat{F}_\nu(0) + O\left(\frac{X}{(\log R)^{2019}}\right) + O\left(\frac{X^{1+\varepsilon}}{Y}\right).$$

We insert into this the definitions (3.6.2.4) and (3.3.1) of F_ν and ω_2, interchange the order of summation, and then write the sum on ν as an Euler product. The result is

$$\mathcal{T}_0 = \frac{2X}{(\sqrt{2}-1)^4}\frac{1+o(1)}{\log R}\frac{1}{2\pi i}\int_{(c)}\frac{\Gamma(\frac{s}{2}+\frac{1}{4})^2}{\Gamma(\frac{1}{4})^2}\left(1-\frac{1}{2^{\frac{1}{2}-s}}\right)^2\left(\frac{X}{\pi}\right)^s\check{\Phi}(s)\left(1-\frac{1}{2^{1+2s}}\right)^3$$
$$\times\,\zeta(1+2s)^3\left(1-\frac{1}{2^{2+4s}}\right)^{-1}\zeta(2+4s)^{-1}\frac{ds}{s} + O\left(\frac{X}{(\log R)^{2019}} + \frac{X^{1+\varepsilon}}{Y}\right).$$

As before, we truncate the integral to the range $|\mathrm{Im}(s)| \le (\log X)^2$, and then deform the path of integration to the path made up of the line segments L_1, L_2, L_3 defined above (3.6.3.11) to see that the main contribution arises from the residue of the integrand at $s = 0$. We evaluate the residue using (3.5.1.6) and arrive at

$$\mathcal{T}_0 = \left(144\zeta(2)\left(1-\frac{1}{\sqrt{2}}\right)^2\right)^{-1}\frac{X\check{\Phi}(0)}{4}\frac{1+o(1)}{\log R}(\log X)^3 + O\left(X\log X + \frac{X^{1+\varepsilon}}{Y}\right).$$

Recalling the definition of $\mathfrak{c}$, we have

$$\mathcal{T}_0 \le (\mathfrak{c}+\varepsilon)\frac{X}{8}\frac{(\log X)^3}{\log R} + O\left(X\log X + \frac{X^{1+\varepsilon}}{Y}\right). \tag{3.8.1.3}$$

Moreover, we see from (3.6.9.12) that if $M = 1$ and $b_1 = 1$, then

$$\mathcal{B} \ll X\frac{\log X}{\log R} \ll X \tag{3.8.1.4}$$

since we may deform the path of integration in (3.6.9.12) to a circle $|s| = \varepsilon$. The condition $\theta + 2\vartheta < \frac{1}{2}$ in Subsection 3.6.7 with $\theta = 0$ allows us to take $\vartheta = \frac{1}{4} - \varepsilon$ in (3.8.1.3). We then set $Y = X^{\delta}$, for some small, fixed $\delta > 0$. We see that the upper bound (3.8.1.2) then follows from (3.8.1.3) and (3.8.1.4) after sending ε to zero sufficiently slowly.

3.8.2 The lower bound in Theorem 0.2.3

Recall the definition (3.8.1.1) of M_2. Our goal is to prove the following result.

Proposition 3.8.1. *For large X we have*

$$M_2 \geq \frac{1}{2}(\mathfrak{c} - o(1))\frac{X}{4}(\log X)^3,$$

where $\mathfrak{c}$ is the positive constant defined in Theorem 0.2.3, and $o(1)$ is some quantity that goes to zero as $X \to \infty$.

The lower bound for Theorem 0.2.3 easily follows from Proposition 3.8.1 by summing over dyadic intervals.

The main idea in the proof of Proposition 3.8.1 is a standard one. For any Dirichlet polynomial $A(p)$, the Cauchy-Schwarz inequality implies

$$M_2 \geq \frac{\left(\sum_{p \equiv 1 \pmod 8} (\log p) \Phi\left(\frac{p}{X}\right) L\left(\frac{1}{2}, \chi_p\right) A(p)\right)^2}{\sum_{p \equiv 1 \pmod 8} (\log p) \Phi\left(\frac{p}{X}\right) A(p)^2}. \tag{3.8.2.1}$$

Clearly, we should choose $A(p)$ to be an approximation to $L(\frac{1}{2}, \chi_p)$. Our choice is inspired by the approximate functional equation in Lemma 3.3.2. For a positive real number α, we define

$$A_\alpha(p) := \frac{2}{\left(1 - \frac{1}{\sqrt{2}}\right)^2} \sum_{n \text{ odd}} \frac{\chi_p(n)}{\sqrt{n}} \omega_1\left(n\sqrt{\frac{\pi}{p^\alpha}}\right). \tag{3.8.2.2}$$

With $\varepsilon_0 > 0$ small and fixed, we then choose $A(p)$ in (3.8.2.1) to be

$$A(p) := A_{1-\varepsilon_0}(p). \tag{3.8.2.3}$$

Observe that taking $\alpha = 1$ in (3.8.2.2) yields

$$A_1(p) = L\left(\tfrac{1}{2}, \chi_p\right). \tag{3.8.2.4}$$

Proposition 3.8.2. *Let $\varepsilon_0 > 0$ be small. Let $\alpha_1 \leq \alpha_2$ be real numbers with $\alpha_1, \alpha_2 \in \{1-\varepsilon_0, 1\}$, and $(\alpha_1, \alpha_2) \neq (1,1)$. Then*

$$M_{\alpha_1,\alpha_2} := \sum_{p \equiv 1 \pmod 8} (\log p)\Phi\left(\frac{p}{X}\right) A_{\alpha_1}(p) A_{\alpha_2}(p) = \frac{1}{2}(\mathfrak{c} + O(\varepsilon_0))\frac{X}{4}(\log X)^3.$$

Proof of Proposition 3.8.1 assuming Proposition 3.8.2. By (3.8.2.1), (3.8.2.3), and (3.8.2.4), we have

$$M_2 \geq \frac{M_{1-\varepsilon_0,1}^2}{M_{1-\varepsilon_0,1-\varepsilon_0}}.$$

We apply Proposition 3.8.2 to obtain

$$M_2 \geq \frac{1}{2}(\mathfrak{c} + O(\varepsilon_0))\frac{X}{4}(\log X)^3.$$

The proposition follows upon letting $\varepsilon_0 = \varepsilon_0(X)$ go to zero sufficiently slowly as $X \to \infty$. □

We devote the rest of this subsection to the proof of Proposition 3.8.2.

Proof of Proposition 3.8.2. By definition,

$$M_{\alpha_1,\alpha_2} = \frac{4}{(1-\frac{1}{\sqrt{2}})^4} \sum_{p \equiv 1 \pmod 8} (\log p)\Phi\left(\frac{p}{X}\right)$$
$$\times \sum\sum_{m,n \text{ odd}} \frac{\chi_p(mn)}{\sqrt{mn}} \omega_1\left(m\sqrt{\frac{\pi}{p^{\alpha_1}}}\right)\omega_1\left(n\sqrt{\frac{\pi}{p^{\alpha_2}}}\right).$$

Let $M_{\neq}$ denote the contribution to M_{α_1,α_2} from $mn \neq \square$. An application of Lemma 3.5.2 shows that $M_{\neq} \ll X$, say. We note that for bounding $M_{\neq}$ it is crucial that $\alpha_1 = 1 - \varepsilon_0$.

We therefore have

$$M_{\alpha_1,\alpha_2} = \frac{4}{(1-\frac{1}{\sqrt{2}})^4} \sum_{p\equiv 1 \,(\mathrm{mod}\ 8)} (\log p)\Phi\left(\frac{p}{X}\right)$$
$$\sum\sum_{\substack{(mn,2p)=1\\ mn=\square}} \frac{1}{\sqrt{mn}}\omega_1\left(m\sqrt{\frac{\pi}{p^{\alpha_1}}}\right)\omega_1\left(n\sqrt{\frac{\pi}{p^{\alpha_2}}}\right) + O(X).$$

We use Lemma 3.3.1 to remove the condition $(mn,p)=1$ at the cost of a negligible error. We then open ω_1 using its definition as an integral, and interchange the order of summation and integration. After some simplification we arrive at

$$\begin{aligned} M_{\alpha_1,\alpha_2} = &\frac{4}{(1-\frac{1}{\sqrt{2}})^4}\frac{1}{(2\pi i)^2}\int_{(c_1)}\int_{(c_2)} K(s_1,s_2)\left(\frac{X^{\alpha_1}}{\pi}\right)^{\frac{s_1}{2}}\left(\frac{X^{\alpha_2}}{\pi}\right)^{\frac{s_2}{2}} \\ &\times \zeta(1+2s_1)\zeta(1+2s_2)\zeta(1+s_1+s_2) \\ &\times\left(\sum_{p\equiv 1\,(\mathrm{mod}\ 8)}(\log p)\Phi\left(\frac{p}{X}\right)\left(\frac{p}{X}\right)^{\frac{\alpha_1 s_1+\alpha_2 s_2}{2}}\right)\frac{ds_1 ds_2}{s_1 s_2} + O(X), \end{aligned}$$

where $c_\ell = \mathrm{Re}(s_\ell)$ is a positive real number, and

$$\begin{aligned} K(s_1,s_2) = &\zeta^{-1}(2+2s_1+2s_2)\left(1+\frac{1}{2^{1+s_1+s_2}}\right)^{-1} \\ &\times\prod_{\ell=1}^{2}\frac{\Gamma\left(\frac{s_\ell}{2}+\frac{1}{4}\right)}{\Gamma\left(\frac{1}{4}\right)}\left(1-\frac{1}{2^{\frac{1}{2}-s_\ell}}\right)\left(1-\frac{1}{2^{1+2s_\ell}}\right). \end{aligned}$$

For the moment we choose $c_1 = c_2 = \frac{1}{\log X}$. By the rapid decay of $K(s_1,s_2)$ in vertical strips, we may truncate to $|\mathrm{Im}(s_\ell)| \leq (\log X)^2$ at the cost of a negligible error. With this condition in place, we use the prime number theorem in arithmetic progressions to obtain that the sum on p is

$$\frac{X}{4}\int_0^\infty \Phi(x)x^{\frac{\alpha_1 s_1+\alpha_2 s_2}{2}}dx + O\left(X\exp(-c\sqrt{\log X})\right).$$

The error term clearly makes an acceptable contribution to M_{α_1,α_2}. We then remove the condition on $\mathrm{Im}(s_\ell)$ by the same means we installed it and obtain

$$M_{\alpha_1,\alpha_2} = \frac{4}{(1-\frac{1}{\sqrt{2}})^4}\frac{X}{4}\int_0^\infty \Phi(x)\frac{1}{(2\pi i)^2}\int_{(c_1)}\int_{(c_2)} K(s_1,s_2)\left(\frac{(xX)^{\alpha_1}}{\pi}\right)^{\frac{s_1}{2}} \times \left(\frac{(xX)^{\alpha_2}}{\pi}\right)^{\frac{s_2}{2}} \times \zeta(1+2s_1)\zeta(1+2s_2)\zeta(1+s_2+s_2)\frac{ds_1 ds_2}{s_1 s_2}dx + O(X).$$

We wish to separate the variables s_1 and s_2. Since $c_\ell > 0$ we expand $\zeta(1+s_1+s_2)$ as an absolutely convergent Dirichlet series. Interchanging the order of summation and integration, we obtain

$$M_{\alpha_1,\alpha_2} = \frac{4}{(1-\frac{1}{\sqrt{2}})^4}\frac{X}{4}\int_0^\infty \Phi(x)\sum_{n=1}^\infty \frac{1}{n}\frac{1}{(2\pi i)^2}\int_{(c_1)}\int_{(c_2)} K(s_1,s_2) \times \left(\frac{(xX)^{\alpha_1}}{\pi n^2}\right)^{\frac{s_1}{2}}\left(\frac{(xX)^{\alpha_2}}{\pi n^2}\right)^{\frac{s_2}{2}} \times \zeta(1+2s_1)\zeta(1+2s_2)\frac{ds_1 ds_2}{s_1 s_2}dx + O(X).$$

To truncate the summation over n, first we move the contours of integration to the right to $c_1 = c_2 = 1$. By trivial estimation we deduce that the contribution from $n \gg X^{\frac{\alpha_1+\alpha_2}{4}}$ is $O(X)$. For n in the range $X^{\frac{\alpha_1}{2}} \ll n \ll X^{\frac{\alpha_1+\alpha_2}{4}}$, we move $\mathrm{Re}(s_2)$ to $c_2 = \frac{1}{\log X}$ and estimate trivially, getting an error term of $O(X(\log X)^2)$. With $n \ll X^{\frac{\alpha_1}{2}}$ we then move c_1 to $\frac{1}{\log X}$, obtaining

$$M_{\alpha_1,\alpha_2} = \frac{4}{(1-\frac{1}{\sqrt{2}})^4}\frac{X}{4}\int_0^\infty \Phi(x)\sum_{n\le\sqrt{(xX)^{\alpha_1}/\pi}} \frac{1}{n} \times \frac{1}{(2\pi i)^2}\int_{(\frac{1}{\log X})}\int_{(\frac{1}{\log X})} K(s_1,s_2)\left(\frac{(xX)^{\alpha_1}}{\pi n^2}\right)^{\frac{s_1}{2}}\left(\frac{(xX)^{\alpha_2}}{\pi n^2}\right)^{\frac{s_2}{2}} \times \zeta(1+2s_1)\zeta(1+2s_2)\frac{ds_1 ds_2}{s_1 s_2}dx + O(X(\log X)^2).$$

The variables s_1 and s_2 are almost separated, except they are entangled inside of $K(s_1,s_2)$. We move the lines of integration to $\mathrm{Re}(s_1) = \mathrm{Re}(s_2) = -\delta$, for some small, fixed $\delta > 0$. In doing so we pick up contributions from the poles at $s_1, s_2 = 0$. The contribution from the integrals on $\mathrm{Re}(s_\ell) = -\delta$ is trivially bounded by $O(X\log X)$. We write the contributions from the poles at $s_\ell = 0$ as contour integrals around small circles, thereby obtaining

$$M_{\alpha_1,\alpha_2} = \frac{4}{(1-\frac{1}{\sqrt{2}})^4}\frac{X}{4}\int_0^\infty \Phi(x) \sum_{n\le\sqrt{(xX)^{\alpha_1}/\pi}} \frac{1}{n}$$
$$\times \frac{1}{(2\pi i)^2} \oint\oint_{|s_\ell|=(\log X)^{-1}} K(s_1,s_2)\left(\frac{(xX)^{\alpha_1}}{\pi n^2}\right)^{\frac{s_1}{2}}\left(\frac{(xX)^{\alpha_2}}{\pi n^2}\right)^{\frac{s_2}{2}}$$
$$\times \zeta(1+2s_1)\zeta(1+2s_2)\frac{ds_1 ds_2}{s_1 s_2}dx + O(X(\log X)^2).$$

Since $|s_\ell| = (\log X)^{-1}$ we have

$$K(s_1,s_2) = K(0,0) + O\left(\frac{1}{\log X}\right) = \frac{1}{6\zeta(2)}\left(1-\frac{1}{\sqrt{2}}\right)^2 + O\left(\frac{1}{\log X}\right),$$

and therefore

$$M_{\alpha_1,\alpha_2} = \frac{2}{3\zeta(2)(1-\frac{1}{\sqrt{2}})^2}\frac{X}{4}\int_0^\infty \Phi(x) \sum_{n\le\sqrt{(xX)^{\alpha_1}/\pi}} \frac{1}{n}$$
$$\times \frac{1}{(2\pi i)^2} \oint\oint_{|s_\ell|=(\log X)^{-1}} \left(\frac{(xX)^{\alpha_1}}{\pi n^2}\right)^{\frac{s_1}{2}}\left(\frac{(xX)^{\alpha_2}}{\pi n^2}\right)^{\frac{s_2}{2}}$$
$$\times \zeta(1+2s_1)\zeta(1+2s_2)\frac{ds_1 ds_2}{s_1 s_2}dx + O(X(\log X)^2).$$

Expanding in Laurent and power series yields

$$\frac{1}{2\pi i}\oint_{|s_\ell|=(\log X)^{-1}} \left(\frac{(xX)^{\alpha_\ell}}{\pi n^2}\right)^{\frac{s_\ell}{2}} \zeta(1+2s_\ell)\frac{ds_\ell}{s_\ell} = \frac{1}{2}\log\left(\frac{1}{n}\sqrt{\frac{(xX)^{\alpha_\ell}}{\pi}}\right) + O(1),$$

and hence

$$M_{\alpha_1,\alpha_2} = \frac{1}{6\zeta(2)(1-\frac{1}{\sqrt{2}})^2}\frac{X}{4}\int_0^\infty \Phi(x) \sum_{n\le\sqrt{(xX)^{\alpha_1}/\pi}} \frac{1}{n}\log\left(\frac{1}{n}\sqrt{\frac{(xX)^{\alpha_1}}{\pi}}\right)$$
$$\times \log\left(\frac{1}{n}\sqrt{\frac{(xX)^{\alpha_2}}{\pi}}\right) + O(X(\log X)^2).$$

Partial summation yields

$$\sum_{n\leq\sqrt{(xX)^{\alpha_1}/\pi}} \frac{1}{n}\log\left(\frac{1}{n}\sqrt{\frac{(xX)^{\alpha_1}}{\pi}}\right)\log\left(\frac{1}{n}\sqrt{\frac{(xX)^{\alpha_2}}{\pi}}\right) = \frac{1+O(\varepsilon_0)}{24}(\log X)^3,$$

and therefore

$$M_{\alpha_1,\alpha_2} = \frac{1}{2}(\mathfrak{c}+O(\varepsilon_0))\frac{X}{4}(\log X)^3.$$

□

3.8.3 Proof of Theorem 0.2.4

We turn now to the proof of Theorem 0.2.4. Throughout this subsection we set $\eta := 100\log\log X/\log X$. Recalling the definition (3.8.2.2) of $A_\alpha(p)$, we then have

$$L\left(\tfrac{1}{2},\chi_p\right) = A_{1-\eta}(p) + B(p),$$

say. Thus

$$M_2 = \sum_{p\equiv 1\ (\mathrm{mod}\ 8)} (\log p)\Phi\left(\frac{p}{X}\right)\left\{A_{1-\eta}(p)^2 + O(|A_{1-\eta}(p)B(p)| + |B(p)|^2)\right\}. \qquad (3.8.3.1)$$

We prove, on GRH, that

$$\sum_{p\equiv 1\ (\mathrm{mod}\ 8)} (\log p)\Phi\left(\frac{p}{X}\right)A_{1-\eta}(p)^2 = \mathfrak{c}\frac{X}{8}(\log X)^3 + O(X(\log X)^{2+\varepsilon}) \qquad (3.8.3.2)$$

and

$$\sum_{p\equiv 1\ (\mathrm{mod}\ 8)} (\log p)\Phi\left(\frac{p}{X}\right)|B(p)|^2 \ll X(\log X)^{5/2}. \qquad (3.8.3.3)$$

Theorem 0.2.4 then follows from (3.8.3.1), (3.8.3.2), and (3.8.3.3) after applying Cauchy-Schwarz and summing over dyadic ranges.

We may easily prove (3.8.3.2), since the treatment is substantially similar to the proof of Proposition 3.8.2. Applying the approximate functional equation, the main term of (3.8.3.2) is

$$\frac{4}{(1-\frac{1}{\sqrt{2}})^4}\sum_{p\equiv 1\,(\mathrm{mod}\ 8)}(\log p)\Phi\left(\frac{p}{X}\right)\sum_{m,n\ \mathrm{odd}}\sum\frac{\chi_p(mn)}{\sqrt{mn}}\omega_1\left(m\sqrt{\frac{\pi}{p^{1-\eta}}}\right)\omega_1\left(n\sqrt{\frac{\pi}{p^{1-\eta}}}\right).$$

We argue as in Proposition 3.8.2 and obtain that the contribution from $mn=\square$ is

$$\mathfrak{c}\frac{X}{8}(\log X)^3+O(X(\log X)^{2+\varepsilon}).$$

The following standard result implies that the contribution to (3.8.3.2) from $mn\neq\square$ is $O(X/\log X)$, say.

Lemma 3.8.3. *Let χ be a non-principal Dirichlet character modulo q. Let χ^* be the primitive character inducing χ, and assume that GRH holds for $L(s,\chi^*)$. If $q\leq X^M$ for some fixed positive constant M, then*

$$\sum_{p\leq X}\chi(p)(\log p)\ll_M X^{1/2}(\log X)^2.$$

The proof of (3.8.3.3) is more subtle. Here the method of proof is that of Soundararajan and Young [54]. As the arguments are very similar, our exposition will be sparse, and we refer the reader to [54] for more details. We perform some initial manipulations, and then we state the main proposition which will yield (3.8.3.3).

By definition, we have

$$B(p)=\frac{1}{2\pi i}\int_{(c)}g(s)L\left(\frac{1}{2}+s,\chi_p\right)\frac{p^{s/2}-p^{(1-\eta)s/2}}{s}ds, \tag{3.8.3.4}$$

where $c>0$ and

$$g(s)=\frac{2}{(1-\frac{1}{\sqrt{2}})^2}\frac{\Gamma\left(\frac{s}{2}+\frac{1}{4}\right)}{\Gamma\left(\frac{1}{4}\right)}\left(1-\frac{1}{2^{\frac{1}{2}-s}}\right)\left(1-\frac{1}{2^{\frac{1}{2}+s}}\right)\pi^{-s/2}.$$

The function $(p^{s/2}-p^{(1-\eta)s/2})/s$ is entire, so we may move the line of integration in (3.8.3.4) to $\mathrm{Re}(s)=0$. On the line $\mathrm{Re}(s)=0$ we have the bound $|(p^{s/2}-p^{(1-\eta)s/2})/s|\ll$

$\log\log X$, and hence the left side of (3.8.3.3) is

$$\ll (\log X)^{1+\varepsilon} \int_{\mathbb{R}} \int_{\mathbb{R}} |g(it_1)g(it_2)| \sum_{\substack{p\le X \\ p\equiv 1 \pmod 8}} \left| L\left(\tfrac{1}{2}+it_1,\chi_p\right) L\left(\tfrac{1}{2}+it_2,\chi_p\right)\right| dt_1\, dt_2. \tag{3.8.3.5}$$

To state the proposition we need, we first establish some notation, following [54, Section 6]. Given $x \ge 10$, say, and a complex number z, we define

$$\mathcal{L}(z,x) = \begin{cases} \log\log x, & |z| \le (\log x)^{-1}, \\ -\log|z|, & (\log x)^{-1} \le |z| \le 1, \\ 0, & |z| \ge 1. \end{cases}$$

For complex numbers z_1 and z_2 we define

$$\mathcal{M}(z_1,z_2,x) = \frac{1}{2}(\mathcal{L}(z_1,x)+\mathcal{L}(z_2,x)),$$

and

$$\begin{aligned}\mathcal{V}(z_1,z_2,x) = \frac{1}{2}[&\mathcal{L}(2z_1,x)+\mathcal{L}(2z_2,x)+\mathcal{L}(2\mathrm{Re}(z_1),x)+\mathcal{L}(2\mathrm{Re}(z_2),x)) \\ &+2\mathcal{L}(z_1+z_2,x)+2\mathcal{L}(z_1+\overline{z_2},x)].\end{aligned}$$

It is helpful to know that for the values of z_1 and z_2 we consider, we have $\log\log X \le \mathcal{V}(z_1,z_2,X) \le 4\log\log X$.

The following result, an analogue of [54, Theorem 6.1], is the key input we need.

Proposition 3.8.4. *Let X be large, and let z_1 and z_2 be complex numbers with $0 \le \mathrm{Re}(z_i) \le \frac{1}{\log X}$ and $|z_i| \le X$. Assume the Riemann Hypothesis for the Riemann zeta function $\zeta(s)$ and for all Dirichlet L-functions $L(s,\chi_p)$ with $p \equiv 1 \pmod 8$. Then for any $r > 0$ in $\mathbb{R}$ and any $\varepsilon > 0$ we have*

$$\sum_{\substack{p\le X \\ p\equiv 1 \pmod 8}} \left| L\left(\tfrac{1}{2}+z_1,\chi_p\right) L\left(\tfrac{1}{2}+z_2,\chi_p\right)\right|^r \ll_{r,\varepsilon} \frac{X}{(\log X)^{1-\varepsilon}} \exp\left(r\mathcal{M}(z_1,z_2,X)+\frac{r^2}{2}\mathcal{V}(z_1,z_2,X)\right).$$

Proof of (3.8.3.3) *assuming Proposition 3.8.4.* Recall (3.8.3.5). If t_1 or t_2 satisfies $|t_i| > X$ we use Cauchy-Schwarz, Lemma 3.3.5, and the rapid decay of g to get a negligible error.

We may therefore assume that $|t_i| \le X$. We then consider, for a parameter $0 < \alpha < 1$ at our disposal, two cases: (1) both t_1 and t_2 satisfy $|t_i| \le (\log X)^{-\alpha}$, or (2) one of t_1, t_2 satisfies $|t_i| \ge (\log X)^{-\alpha}$. In case (1) we use the trivial bounds

$$\mathcal{M}(it_1, it_2, X) \le \log\log X,$$
$$\mathcal{V}(it_1, it_2, X) \le 4\log\log X,$$

while in case (2) we use the bounds

$$\mathcal{M}(it_1, it_2, X) \le \frac{1+\alpha}{2}\log\log X,$$
$$\mathcal{V}(it_1, it_2, X) \le \frac{7+\alpha}{2}\log\log X + O(1).$$

Since $|g(it)| \ll (1+t^2)^{-1}$ we obtain by Proposition 3.8.4 that the quantity in (3.8.3.5) is

$$\ll X(\log X)^{\varepsilon}\left((\log X)^{3-2\alpha} + (\log X)^{9/4+3\alpha/4}\right) = X(\log X)^{27/11+\varepsilon} \le X(\log X)^{5/2}$$

upon choosing $\alpha = 3/11$. □

To prove Proposition 3.8.4 we establish estimates for how often

$$\left| L\left(\frac{1}{2} + z_1, \chi_p\right) L\left(\frac{1}{2} + z_2, \chi_p\right) \right|$$

can be large. The following is very similar to [54, Proposition 6.2].

Proposition 3.8.5. *Assume the hypotheses of Proposition 3.8.4. Let $\mathcal{N}(V; z_1, z_2, X)$ denote the number of primes $p \le X$, $p \equiv 1 \pmod 8$, such that*

$$\log\left| L\left(\frac{1}{2} + z_1, \chi_p\right) L\left(\frac{1}{2} + z_2, \chi_p\right) \right| \ge V + \mathcal{M}(z_1, z_2, X).$$

In the range $3 \le V \le 4r\mathcal{V}(z_1, z_2, X)$ we have

$$\mathcal{N}(V; z_1, z_2, X) \ll \frac{X}{(\log X)^{1-o_r(1)}} \exp\left(-\frac{V^2}{2\mathcal{V}(z_1, z_2, X)}\right),$$

and for larger V we have

$$\mathcal{N}(V;z_1,z_2,X) \ll \frac{X}{(\log X)^{1-o_r(1)}}\exp(-4rV).$$

Proof of Proposition 3.8.4. We have

$$\sum_{\substack{p\le X\\ p\equiv 1 \ (\mathrm{mod}\ 8)}} \left|L\left(\tfrac{1}{2}+z_1,\chi_p\right)L\left(\tfrac{1}{2}+z_2,\chi_p\right)\right|^r$$
$$= r\int_{-\infty}^{\infty}\exp(rV+r\mathcal{M}(z_1,z_2,X))\mathcal{N}(V;z_1,z_2,X)dV.$$

Then use Proposition 3.8.5. □

We use the following lemma to determine how frequently a Dirichlet polynomial can be large. We write $\log_2 X$ for $\log\log X$.

Lemma 3.8.6. *Let X and y be real numbers and k a natural number with $y^k \le X^{\frac{1}{2}-\frac{1}{\log_2 X}}$. For any complex numbers $a(q)$ we have*

$$\sum_{\substack{p\le X\\ p\equiv 1 \ (\mathrm{mod}\ 8)}} \left|\sum_{2<q\le y}\frac{a(q)\chi_p(q)}{q^{\frac{1}{2}}}\right|^{2k} \ll \frac{X\log_2 X}{\log X}\frac{(2k!)}{2^k k!}\left(\sum_{q\le y}\frac{|a(q)|^2}{q}\right)^k,$$

where the implied constant is absolute.

Proof. This result is similar to [54, Lemma 6.3], so we give only a sketch. Since we are assuming GRH we could use Lemma 3.8.3, but we get an unconditional result that is almost as good by appealing to sieve theory.

Since $p\equiv 1 \pmod 8$, we have $\chi_p(q)=\chi_{q^*}(p)$, where for an odd integer n we define $n^*=(-1)^{\frac{n-1}{2}}n$. Observe that χ_{q^*} is a primitive character with conductor $\le 4q$. We then introduce an upper bound sieve supported on $d\le D = X^{\frac{1}{\log_2 X}}$. With the upper bound sieve in place we drop the congruence condition modulo 8 and the condition that p is a prime. Opening the square and using the Pólya-Vinogradov inequality, the sum in question is then

$$
\begin{aligned}
&\ll \sum_{n\le X}\left(\sum_{d|n}\lambda_d\right)\left|\sum_{2<q\le p}\frac{a(q)\chi_{q^*}(n)}{q^{\frac{1}{2}}}\right|^{2k} \\
&\ll \sum_{\substack{q_i\le y\\ q_1\cdots q_{2k}=\square}}\frac{|a(q_1)\cdots a(q_{2k})|}{\sqrt{q_1\cdots q_{2k}}}\sum_{n\le X}\left(\sum_{d|n}\lambda_d\right) \\
&+ D\log(y^{2k})\sum_{q_1,\dots,q_{2k}\le y}|a(q_1)\cdots a(q_{2k})|.
\end{aligned}
$$

For the first term we obtain

$$
\sum_{\substack{q_i\le y\\ q_1\cdots q_{2k}=\square}}\frac{|a(q_1)\cdots a(q_{2k})|}{\sqrt{q_1\cdots q_{2k}}}\sum_{n\le X}\left(\sum_{d|n}\lambda_d\right)\ll \frac{X\log_2 X}{\log X}\frac{(2k!)}{2^k k!}\left(\sum_{q\le y}\frac{|a(q)|^2}{q}\right)^k,
$$

and for the second term we use Cauchy-Schwarz to obtain

$$
D\log(y^{2k})\sum_{q_1,\dots,q_{2k}\le y}|a(q_1)\cdots a(q_{2k})|\ll X\frac{k\log X}{D}\left(\sum_{q\le y}\frac{|a(q)|^2}{q}\right)^k.
$$

□

Proof of Proposition 3.8.5. Assume GRH for $L(s,\chi_p)$. A modification of the proof of the Proposition in [71] then yields

$$
\begin{aligned}
\log\left|L\left(\tfrac{1}{2}+z_1,\chi_p\right)L\left(\tfrac{1}{2}+z_2,\chi_p\right)\right| \le \operatorname{Re}&\left(\sum_{q^\ell\le x}\frac{\chi_p(q^\ell)}{\ell q^{\ell(\frac{1}{2}+\frac{1}{\log x})}}(p^{-\ell z_1}+p^{-\ell z_2})\frac{\log(x/q^\ell)}{\log x}\right) \\
&+2\frac{\log X}{\log x}+O\left(\frac{1}{\log x}\right).
\end{aligned}
$$

The terms with $\ell\ge 3$ contribute $O(1)$. For $\ell=2$ we use the Riemann hypothesis for $\zeta(s)$ (see [54, (6.4)]) and obtain

$$
\frac{1}{2}\sum_{q\le\sqrt{x}}\frac{1}{q^{1+\frac{2}{\log x}}}(q^{-2z_1}+q^{-2z_2})\frac{\log(x/q^2)}{\log x}=\mathcal{M}(z_1,z_2,x)+O(\log\log\log X).
$$

Since $\mathcal{M}(z_1, z_2, x) \le \mathcal{M}(z_1, z_2, X) + 2\frac{\log X}{\log x}$, we obtain

$$\log\left|L\left(\tfrac{1}{2}+z_1,\chi_p\right)L\left(\tfrac{1}{2}+z_2,\chi_p\right)\right| \le \operatorname{Re}\sum_{2<q\le x}\frac{\chi_p(q)}{q^{\frac{1}{2}+\frac{1}{\log x}}}(q^{-z_1}+q^{-z_2}) \tag{3.8.3.6}$$
$$+\mathcal{M}(z_1, z_2, X) + 4\frac{\log X}{\log x} + O(\log\log\log X).$$

We put $\mathcal{V} = \mathcal{V}(z_1, z_2, X)$, and define

$$T = \begin{cases} \frac{1}{2}\log\log\log X, & V \le \mathcal{V}, \\ \frac{\mathcal{V}}{2V}\log\log\log X, & \mathcal{V} < V \le \frac{1}{16}\mathcal{V}\log\log\log X, \\ 8, & V > \frac{1}{16}\mathcal{V}\log\log\log X. \end{cases}$$

We take $x = X^{T/V}$, and $z = x^{1/\log\log X}$.

Taking $x = \log X$ in (3.8.3.6) and estimating trivially, we may assume $V \le \frac{5\log X}{\log\log X}$. In (3.8.3.6) we then have

$$\log\left|L\left(\tfrac{1}{2}+z_1,\chi_p\right)L\left(\tfrac{1}{2}+z_2,\chi_p\right)\right| \le S_1 + S_2 + \mathcal{M}(z_1, z_2, X) + 5\frac{V}{T},$$

where S_1 is the sum on q truncated to $q \le z$, and S_2 is the remainder of the sum. Since $\log\left|L\left(\tfrac{1}{2}+z_1,\chi_p\right)L\left(\tfrac{1}{2}+z_2,\chi_p\right)\right| \ge V + \mathcal{M}(z_1, z_2, X)$ we have

$$S_2 \ge \frac{V}{T} \qquad \text{or} \qquad S_1 \ge V\left(1 - \frac{6}{T}\right) =: V_1.$$

We take $k = \lfloor(\frac{1}{2} - \frac{1}{\log_4 X})\frac{V}{T}\rfloor - 1$ in Lemma 3.8.6 and apply the usual Chebyshev-type maneuver to deduce that the number of $p \le X$ with $S_2 \ge V/T$ is

$$\ll \frac{X\log_2 X}{\log X}\exp\left(-\frac{V}{4T}\log V\right).$$

It remains to bound the number of p for which S_1 is large. By Lemma 3.8.6, for any $k \le (\frac{1}{2} - \frac{1}{\log_2 X})\frac{V\log\log X}{T}$ the number of $p \le X$ with $S_1 \ge V_1$ is

$$\ll \frac{X\log_2 X}{\log X}\left(\frac{2k\mathcal{V}(z_1, z_2, X) + O(\log\log\log X)}{eV_1^2}\right)^k.$$

For $V \le (\log\log X)^2$ we take $k = \lfloor V_1^2/2\mathcal{V}\rfloor$, and for $V > (\log\log X)^2$ we take $k =$

$\lfloor 10V \rfloor$. It follows that the number of p for which $S_1 \geq V_1$ is

$$\ll \frac{X \log_2 X}{\log X} \exp\left(-\frac{V_1^2}{2\mathcal{V}}\left(1 + O\left(\frac{\log\log\log X}{\log\log X}\right)\right)\right) + \frac{X \log_2 X}{\log X} \exp(-V \log V).$$

□

3.9 Proof of Theorem 0.2.5

The proof of Theorem 0.2.5 breaks naturally into two parts: the lower bound, and the upper bound. The argument for the lower bound is very similar to that in [72], and we therefore give only a sketch. The argument for the upper bound is similar to that in Section 3.6. In either case, we crucially use the assumption that the central values are non-negative.

3.9.1 The lower bound

Let $d_{1/2}(n)$ be the multiplicative function with $(d_{1/2} \star d_{1/2})(n) = 1$. For a prime $p \equiv 1 \pmod 4$ and large X define

$$R(p) := \sum_{n \leq X^{1/500}} \frac{d_{1/2}(n)\chi_p(n)}{\sqrt{n}}.$$

By Hölder's inequality and the assumption $L(\frac{1}{2}, \chi_p) \geq 0$ we have

$$\sum_{p \equiv 1 \pmod 8} (\log p)\Phi\left(\frac{p}{X}\right) L\left(\frac{1}{2}, \chi_p\right)^3 \geq \frac{T_1^3}{T_2^2},$$

where

$$T_1 := \sum_{p \equiv 1 \pmod 8} (\log p)\Phi\left(\frac{p}{X}\right) L\left(\frac{1}{2}, \chi_p\right) R(p)^4,$$
$$T_2 := \sum_{p \equiv 1 \pmod 8} (\log p)\Phi\left(\frac{p}{X}\right) R(p)^6.$$

In T_2 we open up $R(p)^6$, and obtain a sum over $n_1, \ldots, n_6$, and p. The terms with $n_1 \cdots n_6 = \square$ yield a main term of size $\ll X(\log X)^6$, and the terms with $n_1 \cdots n_6 \neq \square$

are shown to be an error term by using Lemma 3.5.2.

For T_1, we write $L(\frac{1}{2},\chi_p)$ using Lemma 3.3.2. After opening $R(p)^4$, we have a sum over $n_1,\ldots,n_4,m$, and p, where m is the variable of summation in the approximate functional equation. The main term $mn_1\cdots n_4=\square$ is of size $\gg X(\log X)^6$, and the error term $mn_1\cdots n_4\neq\square$ is small by Lemma 3.5.2. This gives the lower bound.

3.9.2 The upper bound

Assuming that $L(\frac{1}{2},\chi_n)\geq 0$ for all square-free $n\equiv 1 \pmod 8$, we can use an upper bound sieve and positivity to write

$$\begin{aligned} M_3 &:= \sum_{p\equiv 1 \,(\mathrm{mod}\ 8)} (\log p)\Phi\left(\frac{p}{X}\right) L\left(\tfrac{1}{2},\chi_p\right)^3 \\ &\leq (\log X) \sum_{n\equiv 1\,(\mathrm{mod}\ 8)} \mu^2(n)\left(\sum_{\substack{d|n\\ d\leq D}}\lambda_d\right)\Phi\left(\frac{n}{X}\right) L\left(\tfrac{1}{2},\chi_n\right)^3 \end{aligned}$$

The coefficients λ_d of the sieve are given, as before, by (3.4.8). We take R to be a sufficiently small power of X.

We use the approximate functional equation

$$L(\tfrac{1}{2},\chi_n)^3 \;=\; \frac{16}{(\sqrt{2}-1)^6}\sum_{\substack{\nu=1\\ \nu\text{ odd}}}^{\infty}\frac{d_3(\nu)\left(\frac{\nu}{n}\right)}{\sqrt{\nu}}\omega_3\left(\nu\left(\frac{\pi}{n}\right)^{3/2}\right),$$

where $\omega_3(\xi)$ is defined by taking $j=3$ in (3.3.1). Our function $\omega_3(\xi)$ is not the same as $\omega_3(\xi)$ in [51]. After using the approximate functional equation to represent $L(\frac{1}{2},\chi_n)^3$, we write $\mu^2(n)=N_Y(n)+R_Y(n)$. The contribution from $R_Y(n)$ is bounded using arguments similar to those in Subsection 3.6.1. For $N_Y(n)$ we use Poisson summation as before. Up to negligible error, we therefore have the upper bound

$$\begin{aligned} M_3 \leq (\log X)&\frac{16}{(\sqrt{2}-1)^6}\sum_{\substack{d\leq D\\ d|P(z)\\ d\text{ odd}}}\lambda_d\sum_{\substack{\nu=1\\ \nu\text{ odd}}}^{\infty}\frac{d_3(\nu)}{\sqrt{\nu}}\sum_{\substack{\alpha\leq Y\\ \alpha\text{ odd}}}\mu(\alpha) \\ &\times\left(\frac{2[\alpha^2,d]}{\nu}\right)\frac{X}{[\alpha^2,d]8\nu}\sum_{k\in\mathbb{Z}}e\left(\frac{k\overline{[\alpha^2,d]\nu}}{8}\right)\hat{F}_\nu\left(\frac{kX}{[\alpha^2,d]8\nu}\right)\tau_k(\nu), \end{aligned}$$

where

$$F_\nu(t) \;=\; \Phi(t)\omega_3\left(\nu\left(\frac{\pi}{tX}\right)^{3/2}\right).$$

We treat separately the contributions from $k = 0$ and $k \neq 0$. The calculations are somewhat easier in that ultimately we seek only upper bounds, not asymptotic formulas.

The contribution from $k = 0$ is treated as in Subsection 3.6.3, and is

$$\ll X\frac{\log X}{\log R}(\log X)^6 \ll X(\log X)^6.$$

For $k \neq 0$ the presence of the additive character necessitates a splitting of k into residue classes modulo 8. When necessary, we write the additive character as a linear combination of multiplicative characters. We use the identity

$$\tau_k(n) \;=\; \left(\frac{1+i}{2} + \left(\frac{-1}{n}\right)\left(\frac{1-i}{2}\right)\right)G_k(n)$$

and treat the two terms separately. We then follow the method of Section 3.6 to obtain that the contribution from $k \neq 0$ is

$$\ll X\frac{\log X}{\log R}(\log X)^6 \ll X(\log X)^6.$$

One difference that arises is in proving analogues of Lemma 3.6.3. Here we have $\check{\Phi}(w+\frac{s}{2})$ inside of an integral, instead of just $\check{\Phi}(w)$ outside of an integral. It is helpful to use the bound

$$\check{\Phi}(y) \ll_j \left(\frac{\log X}{|y|}\right)^j.$$

Another difference is that we have a factor of $X^{s/2}$ in the integrals, whereas this factor disappeared for the $k \neq 0$ terms in Section 3.6. We therefore do not need to concern ourselves with any symmetry properties of the integrand (cf. the symmetry argument yielding (3.6.9.4)).

REFERENCES

[1] P. Tannery and C. H. (eds.), *Oeuvres de Fermat, 4 volumes.* Paris, 1891-1912.

[2] L. Euler, "Letter to Christian Goldbach, dated April 12, 1749." http://eulerarchive.maa.org//correspondence/letters/OO0852.pdf.

[3] E. Fouvry and H. Iwaniec, "Gaussian primes," *Acta Arith.*, vol. 79, no. 3, pp. 249–287, 1997.

[4] J. B. Friedlander and H. Iwaniec, "The polynomial $X^2 + Y^4$ captures its primes," *Ann. of Math. (2)*, vol. 148, no. 3, pp. 945–1040, 1999.

[5] D. R. Heath-Brown, "Primes represented by $x^3 + 2y^3$," *Acta Math.*, vol. 186, no. 1, pp. 1–84, 2001.

[6] D. R. Heath-Brown and B. Z. Moroz, "On the representation of primes by cubic polynomials in two variables," *Proc. London Math. Soc. (3)*, vol. 88, no. 2, pp. 289–312, 2004.

[7] J. Maynard, "Primes represented by incomplete norm forms," 2015, preprint, arXiv: 1507.05080.

[8] D. R. Heath-Brown and X. Li, "Prime values of $a^2 + p^4$," *Invent. Math.*, vol. 208, no. 2, pp. 441–499, 2017.

[9] J. Maynard, "Primes with restricted digits," *Invent. Math.*, vol. to appear, 2019.

[10] W. Banks, A. Conflitti, and I. Shparlinski, "Character sums over integers with restricted g-ary digits," *Illinois J. Math.*, vol. 46, no. 3, pp. 819–836, 2002.

[11] W. Banks and I. Shparlinski, "Arithmetic properties of numbers with restricted digits," *Acta Arith.*, vol. 112, no. 4, pp. 313–332, 2004.

[12] J. Bourgain, "Prescribing the binary digits of primes, ii," *Israel J. Math.*, vol. 206 (1), pp. 165–182, 2015.

[13] S. Col, "Diviseurs des nombres ellipséphiques," *Period. Math. Hungar.*, vol. 58, no. 1, pp. 1–23, 2009.

[14] J. Coquetl, "On the uniform distribution modulo one of some subsequences of polynomial sequences," *J. Number Theory*, vol. 10, no. 3, pp. 291–296, 1978.

[15] J. Coquetl, "On the uniform distribution modulo one of some subsequences of polynomial sequences, ii," *J. Number Theory*, vol. 12, no. 2, pp. 244–250, 1980.

[16] C. Dartyge and C. Mauduit, "Nombres presque premiers dont l'écriture en base r ne comporte pas certains chiffres," *J. Number Theory*, vol. 81, no. 2, pp. 270–291, 2000.

[17] C. Dartyge and C. Mauduit, "Ensembles de densité nulle contenant des entiers possédant au plus deux facteurs premiers," *J. Number Theory*, vol. 91, no. 2, pp. 230–255, 2001.

[18] M. Drmota and C. Mauduit, "Weyl sums over integers with affine digit restrictions," *J. Number Theory*, vol. 130, no. 11, pp. 2404–2427, 2010.

[19] P. Erdős, C. Mauduit, and A. Sárközy, "On arithmetic properties of integers with missing digits. I. distribution in residue classes," *J. Number Theory*, vol. 70, no. 2, pp. 99–120, 1998.

[20] P. Erdős, C. Mauduit, and A. Sárközy, "On arithmetic properties of integers with missing digits. II. prime factors," *Paul Erdős memorial collection. Discrete Math.*, vol. 200, no. 1-3, pp. 149–164, 1999.

[21] S. Konyagin, "Arithmetic properties of integers with missing digits: distribution in residue classes," *Period. Math. Hungar.*, vol. 42, no. 1-2, pp. 145–162, 2001.

[22] A. Wiles, "The Birch and Swinnerton-Dyer conjecture." *The millennium prize problems*, vol. Clay Math. Inst., pp. 31–41, 2006.

[23] N. Katz and P. Sarnak, "Zeroes of zeta functions and symmetry," *Bull. Amer. Math. Soc.*, vol. 36, no. 1, pp. 1–26, 1999.

[24] H. Iwaniec and P. Sarnak, "The non-vanishing of central values of automorphic L-functions and Landau-Siegel zeros," *Israel J. Math.*, vol. 120, part A, pp. 155–177, 2000.

[25] S. Chowla, *The Riemann hypothesis and Hilbert's tenth problem*, ser. Mathematics and Its Applications, Vol. 4. Gordon and Breach Science Publishers, 1965.

[26] R. Balasubramanian and V. K. Murty, "Zeros of Dirichlet L-functions," *Ann. Sci. École Norm. Sup. (4)*, vol. 25, no. 5, pp. 567–615, 1992.

[27] H. Bohr and E. Landau, "Sur les zeros de la fonction $\zeta(s)$ de Riemann," *C. R. Acad. des Sciences Paris*, vol. 158, pp. 158–162, 1914.

[28] J. B. Conrey, A. Ghosh, and S. M. Gonek, "Simple zeros of the zeta-function of a quadratic number field. ii," *Analytic number theory and Diophantine problems (Stillwater, OK, 1984)*, vol. Progr. Math., 70, pp. 87–114, 1987.

[29] A. Selberg, "On the zeros of Riemann's zeta-function," *Skr. Norske Vid. Akad. Oslo I.*, vol. no. 10, p. 59 pp., 1942.

[30] H. Iwaniec and P. Sarnak, "Dirichlet L-functions at the central point," pp. 941–952, 2004.

[31] H. M. Bui, "Non-vanishing of Dirichlet L-functions at the central point," *Int. J. Number Theory*, vol. 8, no. 8, pp. 1855–1881, 2012.

[32] R. Khan and H. T. Ngo., "Nonvanishing of Dirichlet L-functions," *Algebra Number Theory*, vol. 10, no. 10, pp. 2081–2091, 2016.

[33] M. R. Murty, "On simple zeros of certain L-series," pp. 427–439, 1990.

[34] F. Sica, *The order of vanishing of L-functions at the center of the critical strip.* McGill University, PhD thesis, 1998.

[35] S. J. Miller and R. Takloo-Bighash, *An invitation to modern number theory.* Princeton University Press, 2006.

[36] K. Pratt, "Average nonvanishing of Dirichlet L-functions at the central point," *Algebra Number Theory*, vol. 13, pp. 227–249, 2019.

[37] H. M. Bui and M. B. Milinovich, "Central values of derivatives of Dirichlet L-functions," *Int. J. Number Theory*, vol. 7, no. 2, pp. 371–388, 2011.

[38] S. Das and R. Khan, "The third moment of symmetric square L-functions," *Q. J. Math.*, vol. 69, no. 3, pp. 1063–1087, 2018.

[39] R. Khan, "Non-vanishing of the symmetric square L-function at the central point." *Proc. Lond. Math. Soc. (3)*, vol. 100, no. 3, pp. 736–762, 2010.

[40] R. Khan, "Simultaneous non-vanishing of $GL(3)\times GL(2)$ and $GL(2)$ L-functions," *Math. Proc. Cambridge Philos. Soc.*, vol. 152, no. 3, pp. 535–553, 2012.

[41] R. Khan, D. Milićević, and H. T. Ngo., "Non-vanishing of Dirichlet L-functions in galois orbits," *Int. Math. Res. Not.*, vol. 22, pp. 6955–6978, 2016.

[42] E. Kowalski and P. Michel, "A lower bound for the rank of $J_0(q)$," *Acta Arith.*, vol. 94, no. 4, pp. 303–343, 2000.

[43] E. Kowalski, P. Michel, and J. VanderKam, "Non-vanishing of high derivatives of automorphic L-functions at the center of the critical strip," *J. Reine Angew. Math.*, vol. 526, pp. 1–34, 2000.

[44] E. Kowalski, P. Michel, and J. VanderKam, "Mollification of the fourth moment of automorphic L-functions and arithmetic applications," *Invent. Math.*, vol. 142, no. 1, pp. 95–151, 2000.

[45] P. Michel and J. VanderKam., "Non-vanishing of high derivatives of Dirichlet L-functions at the central point," *J. Number Theory*, vol. 81, no. 1, pp. 130–148, 2000.

[46] K. Ono and C. Skinner, "Non-vanishing of quadratic twists of modular L-functions," *Invent. Math.*, vol. 134, no. 3, pp. 651–660, 1998.

[47] H. Iwaniec and E. Kowalski, *Analytic number theory*, ser. American Mathematical Society Colloquium Publications, 53. American Mathematical Society, 2004.

[48] M. Jutila, "On the mean value of $L(\frac{1}{2}, \chi)$ for real characters." *Analysis*, vol. 1, no. 2, pp. 149–161, 1981.

[49] A. E. Özlük and C. Snyder, "Small zeros of quadratic L-functions," *Bull. Austral. Math. Soc.*, vol. 47, no. 2, pp. 307–319, 1993.

[50] A. E. Özlük and C. Snyder, "On the distribution of the nontrivial zeros of quadratic L-functions close to the real axis." *Acta Arith.*, vol. 91, no. 3, pp. 209–228, 1999.

[51] K. Soundararajan, "Nonvanishing of quadratic Dirichlet L-functions at $s = \frac{1}{2}$," *Ann. of Math. (2)*, vol. 152, no. 2, pp. 447–488, 2000.

[52] J. C. Andrade and J. P. Keating, "Mean value theorems for L-functions over prime polynomials for the rational function field," *Acta Arith.*, vol. 161, no. 4, pp. 371–385, 2013.

[53] J. C. Andrade and S. Baluyot, "Small zeros of Dirichlet L-functions of quadratic characters of prime modulus," 2018, preprint, arXiv: 1802.03413.

[54] K. Soundararajan and M. Young, "The second moment of quadratic twists of modular L-functions," *J. Eur. Math. Soc. (JEMS)*, vol. 12, no. 5, pp. 1097–1116, 2010.

[55] G. Harman, *Prime-detecting sieves*, ser. London Mathematical Society Monographs Series, 33. Princeton University Press, 2007.

[56] J. B. Friedlander and H. Iwaniec, "Asymptotic sieve for primes," *Ann. of Math. (2)*, vol. 148, no. 3, pp. 1041–1065, 1999.

[57] J. B. Friedlander and H. Iwaniec, *Opera de cribro*, ser. American Mathematical Society Colloquium Publications, 57. American Mathematical Society, 2010.

[58] J. B. Conrey, H. Iwaniec, and K. Soundararajan, "The mean square of the product of a Dirichlet L-function and a Dirichlet polynomial," 2018, preprint, arXiv:1808.02879.

[59] K. Soundararajan, "Mean-values of the Riemann zeta-function," *Mathematika*, vol. 42, no. 1, pp. 158–174, 1995.

[60] J. B. F. E. Bombieri and H. Iwaniec., "Some corrections to an old paper," 2019, preprint, arXiv:1903.01371.

[61] J. B. F. E. Bombieri and H. Iwaniec., "Primes in arithmetic progressions to large moduli," *Acta Math.*, vol. 156, no. 3-4, pp. 203–251, 1986.

[62] J. M. Deshouillers and H. Iwaniec, "Kloosterman sums and fourier coefficients of cusp forms," *Invent. Math.*, vol. 70, no. 2, pp. 219–288, 1982/83.

[63] V. B. et al., "On moments of twisted L-functions," *Amer. J. Math.*, vol. 139, no. 3, pp. 707–768, 2017.

[64] W. Duke, J. B. Friedlander, and I. Iwaniec, "Bilinear forms with kloosterman fractions," *Invent. Math.*, vol. 128, no. 1, pp. 23–43, 1997.

[65] H. Cohen, *Number theory. Vol. I. Tools and Diophantine equations*, ser. Graduate Texts in Mathematics, 239. Springer, 2007.

[66] H. Davenport, *Multiplicative number theory.*, ser. Graduate Texts in Mathematics, 74. Springer-Verlag, 2000.

[67] D. R. Heath-Brown, "A mean value estimate for real character sums." *Acta Arith.*, vol. 72, no. 3, pp. 235–275, 1995.

[68] D. H. J. Polymath, "Variants of the Selberg sieve, and bounded intervals containing many primes." *Res. Math. Sci.*, vol. 1, Art. 12, p. 83 pp., 2014.

[69] M. Jutila, "Zero-density estimates for L-functions." *Acta Arith.*, vol. 32, no. 1, pp. 55–62, 1977.

[70] E. C. Titchmarsh, *The Theory of the Riemann Zeta-function.* Oxford University Press, 1986, second edition. Revised by D. R. Heath-Brown.

[71] K. Soundararajan, "Moments of the Riemann zeta function," *Ann. of Math. (2)*, vol. 170, no. 2, pp. 981–993, 2009.

[72] Z. Rudnick and K. Soundararajan, "Lower bounds for moments of L-functions: symplectic and orthogonal examples," *Multiple Dirichlet series, automorphic forms, and analytic number theory.*, vol. Proc. Sympos. Pure Math., 75, pp. 293–303, 2006.