

# DIRICHLET $L$ -FUNCTIONS OF QUADRATIC CHARACTERS OF PRIME CONDUCTOR AT THE CENTRAL POINT

SIEGFRED BALUYOT AND KYLE PRATT

ABSTRACT. We prove that more than nine percent of the central values  $L(\frac{1}{2}, \chi_p)$  are non-zero, where  $p \equiv 1 \pmod{8}$  ranges over primes and  $\chi_p$  is the real primitive Dirichlet character of conductor  $p$ . Previously, it was not known whether a positive proportion of these central values are non-zero. As a by-product, we obtain the order of magnitude of the second moment of  $L(\frac{1}{2}, \chi_p)$ , and conditionally we obtain the order of magnitude of the third moment. Assuming the Generalized Riemann Hypothesis, we show that our lower bound for the second moment is asymptotically sharp.

## TABLE OF CONTENTS

1. Introduction and results	2
2. Notation and conventions	5
3. Outline of the proof of Theorem 1.1	6
4. Lemmata	9
5. Sieve estimates	13
6. The mollified first moment	18
6.1. Main term	19
6.2. Preparation of the off-diagonal	21
6.3. Regime I	24
6.4. Regime II	27
6.5. Regime III	30
6.6. Dénouement	32
7. The mollified second moment	32
7.1. The contribution of $S_R^+$	33
7.2. Poisson summation	35
7.3. The contribution from $k = 0$	36
7.4. The contribution from $k \neq 0$ : splitting into cases	40
7.5. Evaluation of the sum with $\mathcal{Q}_1$	41
7.6. Bounding the contribution of $\mathcal{R}_1$	47
7.7. Conditions for the parameters	52
7.8. Evaluating the sums of the other terms with $k \neq 0$	53
7.9. Putting together the estimates	56
8. Choosing the mollifier: finishing the proof of Theorem 1.1	65
9. The second moment of $L(\frac{1}{2}, \chi_p)$	68
9.1. The upper bound in Theorem 1.2	68

---

2010 *Mathematics Subject Classification.* 11M20, 11N36, 11R42.

*Keywords and phrases:* central point, mollifier, moments, non-vanishing, primes, quadratic Dirichlet character, Selberg sieve.

9.2. The lower bound in Theorem 1.2	69
9.3. Proof of Theorem 1.3	72
10. Proof of Theorem 1.4	78
10.1. The lower bound	78
10.2. The upper bound	78
Acknowledgements	79
References	80

## 1. INTRODUCTION AND RESULTS

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 non-vanishing is particularly important. For instance, the well-known Birch and Swinnerton-Dyer conjecture [44] 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 [20] show that the non-vanishing of  $L$ -functions associated with holomorphic cusp forms is closely related to the Landau-Siegel zero problem. Thus the question of non-vanishing 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  $\chi$  a primitive quadratic Dirichlet character [8, 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  $\chi$ . One family that has attracted a lot of attention is the family of  $L(s, \chi)$  with  $\chi$  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 [4] 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, [5, 10, 39]). Iwaniec and Sarnak [19] 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 non-vanishing of  $L$ -functions at the central point. Bui [6] and Khan and Ngo [27] introduced new ideas and further improved the lower bound  $\frac{1}{3}$ . The second author [36] has shown that more than fifty percent of the central values are non-vanishing when one additionally averages over the conductors. For further interesting research on this and other families of  $L$ -functions, see [7, 11, 24, 25, 26, 28, 29, 30, 31, 32].

The family of  $L(s, \chi)$  with  $\chi$  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

$\chi_d$  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 [18, p. 514]. Jutila [22] initiated the study of non-vanishing 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  $\chi_d$  with  $|d| \leq X$  have  $L(\frac{1}{2}, \chi_d) \neq 0$ . Özlük and Snyder [33] examined the low-lying zeros of this family, and found the first evidence of its symplectic behavior. Assuming the Generalized Riemann Hypothesis (GRH), they showed that at least  $\frac{15}{16}$  of the central values  $L(\frac{1}{2}, \chi_d)$  are non-zero [34]. Katz and Sarnak independently obtained the same result in unpublished work (see [23, 40]).

Soundararajan [40] 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 biggest difficulty lies in analyzing the contribution of the “off-diagonal” terms in the evaluation of a mollified second moment. Soundararajan discovered that there is, in fact, a main contribution arising from these off-diagonal terms. (See Section 3 for more discussion.)

The case of real primitive characters with prime conductor is more difficult still. Jutila [22] 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 the authors are aware, no asymptotic evaluation of the second moment has appeared in the literature. However, Andrade and Keating [2] asymptotically evaluated the second moment of an analogous family over function fields. Andrade and the first author [1] 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 an 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 1.1.** *There exists an absolute, effective constant  $X_0$  such that if  $X \geq X_0$  then*

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

The proof of Theorem 1.1 proceeds via the mollification method, which we discuss briefly in Section 3 below. Our methods build on those of Jutila [22] and Soundararajan [40]. As in the work of Soundararajan, the main difficulty lies in evaluating the contribution of certain off-diagonal terms. The difference now is that we are summing over primes instead of over square-free integers, and so we cannot directly use his approach. A key idea in the proof of Theorem 1.1 is the use of upper bound sieves to turn intractable sums over primes into manageable sums over integers. The use of sieves in studying central values of  $L$ -functions has also appeared in some other contexts (see [17], also [37, p. 1035]).

The tools developed for the proof of Theorem 1.1 allow us to obtain the order of magnitude of the second moment of  $L(\frac{1}{2}, \chi_p)$ .

**Theorem 1.2.** *Let  $\mathfrak{c}$  be the positive constant*

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

For large  $X$  we have

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

One would rather have an upper bound in Theorem 1.2 that asymptotically matches the lower bound, but this seems difficult to prove unconditionally. By adapting a method of Soundararajan and Young [42] we are able, however, to prove such an asymptotic formula on GRH.

**Theorem 1.3.** *Let  $\mathfrak{c}$  be as in Theorem 1.2. Assume the Riemann Hypothesis for  $\zeta(s)$  and for all Dirichlet  $L$ -functions  $L(s, \chi_p)$  with  $p \equiv 1 \pmod{8}$ . Then*

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

After we completed this paper, Maksym Radziwiłł informed us about work in progress with Julio Andrade, Roger Heath-Brown, Xiannan Li, and K. Soundararajan in which they derive an unconditional asymptotic formula for the second moment of  $L(\frac{1}{2}, \chi_p)$ . Their approach similarly introduces sieve weights, and they also observed that this idea could lead to a non-vanishing result.

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 1.4.** *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  $X$*

$$\sum_{\substack{p \leq X \\ p \equiv 1 \pmod{8}}} (\log p) L\left(\frac{1}{2}, \chi_p\right)^3 \asymp X(\log X)^6.$$

Throughout this paper, we work exclusively with  $p \equiv 1 \pmod{8}$  for convenience, but our methods are not specific to this residue class. With some modifications one could state similar results for other residue classes modulo 8. See the end of Section 3 for more details.

Our work indicates that Soundararajan's lower bound [40] for the proportion of non-vanishing for fundamental discriminants  $d \equiv 0 \pmod{8}$  also holds for the case of fundamental discriminants  $d \equiv 1 \pmod{8}$ . Proving this involves re-doing the calculations in Section 7, but without applying an upper bound sieve. To complete the proof, one would also need a first moment calculation. We omit the details and instead refer the reader to [40, Section 4].

It is natural to ask about the limitations of our method, and how much we can increase the lower bound in Theorem 1.1. If we assume that we can use arbitrarily long mollifiers [13], then we obtain a higher percentage of non-vanishing. However, in view of the parity problem of sieve theory [14], we could not reach a proportion greater than  $\frac{1}{2}$  via our method. On the other hand, by a different method [1], the Density Conjecture of Katz and Sarnak would imply that 100% of the central values  $L(\frac{1}{2}, \chi_p)$  are nonzero.

The outline of the rest of the paper is as follows. In Section 2 we establish some notation and conventions that hold throughout this work. Section 3 outlines the basic strategy for

the proof of Theorem 1.1. In Sections 4 and 5 we state a number of important technical results which are used in the proofs of our theorems. The proof of Theorem 1.1 is spread across Sections 6, 7, and 8. In Section 6 and its subsections we study the mollified first moment problem. The very long Section 7 and its subsections handle the mollified second moment. We choose our mollifier and finish the proof of Theorem 1.1 in Section 8. We prove Theorems 1.2 and 1.3 in Section 9, and we prove Theorem 1.4 in Section 10.

## 2. NOTATION AND CONVENTIONS

We define  $\chi_n(\cdot) = \left(\frac{\cdot}{n}\right)$ , the Kronecker symbol, for all nonzero integers  $n$ , even if  $n$  is not a fundamental discriminant. Note that this means  $\chi_n$  has conductor  $|n|$  only when  $n$  is a fundamental discriminant. We write  $S(Q)$  for the set of all real primitive characters  $\chi$  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  $\varepsilon$  is present, in some fashion, in an inequality or error term, we allow implied constants to depend on  $\varepsilon$  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,  $Y \ll_A Z$ . When we write an error term of the form  $O_A((\log X)^{-A})$ , we always mean that  $A$  is some large but unspecified constant.

Throughout this paper, we denote by  $\Phi(x)$  a smooth function, compactly supported on  $[\frac{1}{2}, 1]$ , which satisfies  $0 \leq \Phi(x) \leq 1$  and  $\Phi(x) = 1$  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  $j \geq 0$ . In Sections 9 and 10 we allow for the possibility of  $\Phi(x)$  being a smooth majorant for the indicator function of  $[\frac{1}{2}, 1]$  with similar properties; we leave it to the reader to determine in any given subsection whether a majorant or minorant is required. 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  $e(x) = e^{2\pi i x}$ . 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 5.2).

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

$$(2.1) \quad g_s(t) = g(t) t^{s/2}.$$

Note that

$$(2.2) \quad \hat{\Phi}(0) = \Phi^\dagger(1) = \check{\Phi}(0) = \frac{1}{2} + O\left(\frac{1}{\log X}\right).$$

The letter  $p$  always denotes a prime number. We write  $\varphi$  for the Euler phi function, and  $d_k(n)$  for the  $k$ -fold divisor function, so that  $d_k(n) = 1$  for  $k = 1$ . 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  $\bar{a} \pmod{q}$  for the multiplicative inverse of  $a$  modulo  $q$ .

### 3. OUTLINE OF THE PROOF OF THEOREM 1.1

The proof of Theorem 1.1 proceeds through the mollification method. The method was introduced by Bohr and Landau [5], but later greatly refined in the hands of Selberg [39]. 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

$$\begin{aligned} S_1 &:= \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) L\left(\frac{1}{2}, \chi_p\right) M(p), \\ S_2 &:= \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) L\left(\frac{1}{2}, \chi_p\right)^2 M(p)^2. \end{aligned} \tag{3.1}$$

If the mollifier is chosen well then  $S_1 \gg X$  and  $S_2 \ll X$ . By the Cauchy-Schwarz inequality we have

$$\sum_{\substack{p \equiv 1 \pmod{8} \\ L(\frac{1}{2}, \chi_p) \neq 0}} (\log p) \Phi\left(\frac{p}{X}\right) \geq \frac{S_1^2}{S_2}, \tag{3.2}$$

and this implies that a positive proportion of  $L(\frac{1}{2}, \chi_p)$  are non-zero.

Our mollifier takes the form

$$M(p) := \sum_{\substack{m \leq M \\ m \text{ odd}}} \frac{b_m}{\sqrt{m}} \chi_p(m), \tag{3.3}$$

for some coefficients  $b_m$  we describe shortly. Here we set

$$M = X^\theta, \quad \theta \in (0, \tfrac{1}{2}) \quad \text{fixed.} \tag{3.4}$$

The larger one can take  $\theta$ , the better proportion of non-vanishing one can achieve.

The coefficients  $b_m$  are a smoothed version of the Möbius function  $\mu(m)$ . Specifically, we choose

$$b_m = \mu(m) H\left(\frac{\log m}{\log M}\right), \tag{3.5}$$

where  $H(t)$  is a smooth function compactly supported in  $[-1, 1]$  that we choose in Section 8. 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 6). Using an approximate functional equation for the central value  $L(\frac{1}{2}, \chi_p)$  (Lemma 4.2), 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 8 to handle the sum over  $p$ . 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 6.1).

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  $\chi_q$  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 [22].

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  $\chi_{q^*}$  is exceptional: that is, the associated  $L$ -function  $L(s, \chi_{q^*})$  has a real zero  $\beta_*$  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  $\chi_{q^*}$  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  $X^\delta$  for some small, fixed  $\delta > 0$ . Rather than treating the sum over primes analytically, as we did when  $Q$  was small or medium-sized, we treat the sum over 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 4.4).

We now describe our plan of attack for  $S_2$  (see Section 7). Recall the definition of  $S_2$  in (3.1). As we see from Theorem 1.3, we only barely obtain an asymptotic formula for the second moment

$$\sum_{\substack{p \leq X \\ p \equiv 1 \pmod{8}}} (\log p) L\left(\frac{1}{2}, \chi_p\right)^2$$

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 5). By positivity we have

$$S_2 \leq (\log X) \sum_{n \equiv 1 \pmod{8}} \mu^2(n) \Phi\left(\frac{n}{X}\right) \left( \sum_{d|n} \lambda_d \right) L\left(\frac{1}{2}, \chi_n\right)^2 M(n)^2,$$

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 [40] (see [40, Section 5]).

We begin by writing

$$(3.6) \quad \mu^2(n) = N_Y(n) + R_Y(n),$$

where

$$(3.7) \quad N_Y(n) := \sum_{\substack{\ell^2 | n \\ \ell \leq Y}} \mu(\ell), \quad R_Y(n) := \sum_{\substack{\ell^2 | n \\ \ell > Y}} \mu(\ell),$$

and  $Y$  is a small power of  $X$ . The sum

$$\sum_{n \equiv 1 \pmod{8}} \Phi\left(\frac{n}{X}\right) R_Y(n) \left( \sum_{d|n} \lambda_d \right) L\left(\frac{1}{2}, \chi_n\right)^2 M(n)^2$$

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 4.5).

The main task is therefore to asymptotically evaluate the sum

$$\sum_{n \equiv 1 \pmod{8}} \Phi\left(\frac{n}{X}\right) N_Y(n) \left( \sum_{d|n} \lambda_d \right) L\left(\frac{1}{2}, \chi_n\right)^2 M(n)^2.$$

We use an approximate functional equation to represent the central values  $L(\frac{1}{2}, \chi_n)^2$  and arrive at expressions of the form

$$\sum_{\ell \leq Y} \mu(\ell) \sum_{d \leq D} \lambda_d \sum_{m_1, m_2 \leq M} \frac{b_{m_1} b_{m_2}}{\sqrt{m_1 m_2}} \sum_{\nu=1}^{\infty} \frac{d_2(\nu)}{\sqrt{\nu}} \sum_{\substack{n \equiv 1 \pmod{8} \\ d|n \\ \ell^2 | n}} \left( \frac{m_1 m_2 \nu}{n} \right) \Phi\left(\frac{n}{X}\right) \omega\left(\frac{\nu}{n}\right),$$

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 resulting sum over  $m$  into a sum basically of the form

$$\sum_{k \in \mathbb{Z}} \left( \frac{[d, \ell^2]k}{m_1 m_2 \nu} \right) e \left( \frac{k[d, \ell^2]m_1 m_2 \nu}{8} \right) \hat{F}_\nu \left( \frac{kX}{[d, \ell^2]m_1 m_2 \nu} \right),$$

for some smooth function  $F_\nu$ . (The quadratic character here is really a kind of quadratic Gauss sum, but we simplify the situation for the sake of this outline.) 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 (a convention we utilize only in this outline section), the  $k = 0$  contribution represents the expected “diagonal” contribution from  $m_1 m_2 \nu = \square$ . 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 [40]. The situation is complicated by the presence of the additive character  $e(\cdot)$ , which is not present in [40]. 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  $\lambda_d$ , 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. [18, 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 8 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 1.1.

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  $\chi_p$  is still a primitive character of conductor  $p$ . Second, we use a variant of the approximate functional equation (Lemma 4.2) with  $\omega_j$ , defined in (4.1), 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.

#### 4. 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

$$(4.1) \quad \omega_j(\xi) = \frac{1}{2\pi i} \int_{(c)} \frac{\Gamma\left(\frac{s}{2} + \frac{1}{4}\right)^j}{\Gamma\left(\frac{1}{4}\right)^j} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)^j \xi^{-s} \frac{ds}{s}.$$

**Lemma 4.1.** *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 [40, Lemma 2.1], but we give details for completeness. The function  $\omega_j(s)$  is real-valued because the change of variable  $\text{Im}(s) \rightarrow -\text{Im}(s)$  shows that  $\omega_j$  is equal to its complex conjugate. Moreover, uniform convergence for  $\xi$  in compact subintervals of  $(0, \infty)$  implies that  $\omega_j$  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  $|\Gamma(x+iy)| \leq \Gamma(x)$  for  $x > 0$  and  $z\Gamma(z) = \Gamma(z+1)$ . Thus, for  $c \geq 2$  we obtain

$$\begin{aligned} |\omega_j^{(\nu)}(\xi)| &\ll \Gamma\left(\frac{c}{2} + \frac{5}{4} + \nu\right)^j \left(1 + \frac{2^c}{\sqrt{2}}\right)^j \xi^{-c-\nu} \int_{(c)} \frac{1}{|s||\frac{s}{2} + \frac{1}{4} + \nu|} \prod_{k=0}^{\nu-1} \frac{|s+k|}{|\frac{s}{2} + \frac{1}{4} + k|} |ds| \\ &\ll \Gamma\left(\frac{c}{2} + \frac{5}{4} + \nu\right)^j \left(\frac{2^j}{\xi}\right)^c \left(\frac{2}{\xi}\right)^\nu c^{-1}, \end{aligned}$$

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 4.2.** *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*

$$L\left(\frac{1}{2}, \chi_n\right)^j = \frac{2}{\left(1 - \frac{1}{\sqrt{2}}\right)^{2j}} \sum_{\substack{\nu=1 \\ \nu \text{ odd}}}^{\infty} \frac{\chi_n(\nu) d_j(\nu)}{\sqrt{\nu}} \omega_j \left( \nu \left(\frac{\pi}{n}\right)^{j/2} \right) =: \mathcal{D}_j(n).$$

*Proof.* The proof follows along standard lines (e.g. [18, 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  $n$  is positive we have  $\chi_n(-1) = 1$ , and therefore we have the functional equation (see [9, Proposition 2.2.24], [12, Chapter 9])

$$\Lambda(z, \chi_n) = \Lambda(1 - z, \chi_n).$$

Recall also that  $\Lambda(z, \chi_n)$  is entire because  $\chi_n$  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  $\omega_j$  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  $s = 0$ :

$$\begin{aligned} I &= \frac{\left(\frac{n}{\pi}\right)^{-j/4}}{\Gamma\left(\frac{1}{4}\right)^j} \left(1 - \frac{1}{\sqrt{2}}\right)^{2j} \Lambda\left(\frac{1}{2}, \chi_n\right)^j \\ &\quad + \frac{1}{2\pi i} \int_{(-c)} \frac{\left(\frac{n}{\pi}\right)^{-j/4}}{\Gamma\left(\frac{1}{4}\right)^j} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)^j \left(1 - \frac{1}{2^{\frac{1}{2}+s}}\right)^j \Lambda\left(\frac{1}{2} + s, \chi_n\right)^j \frac{ds}{s}. \end{aligned}$$

In this latter integral we change variables  $s \rightarrow -s$  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.  $\square$

We frequently encounter exponential sums which are analogous to Gauss sums. Given an odd integer  $n$ , we define for all integers  $k$

$$(4.2) \quad G_k(n) = \left(\frac{1-i}{2} + \left(\frac{-1}{n}\right) \frac{1+i}{2}\right) \sum_{a \pmod{n}} \left(\frac{a}{n}\right) e\left(\frac{ak}{n}\right)$$

and

$$(4.3) \quad \tau_k(n) = \sum_{a \pmod n} \left( \frac{a}{n} \right) e \left( \frac{ak}{n} \right) = \left( \frac{1+i}{2} + \left( \frac{-1}{n} \right) \frac{1-i}{2} \right) G_k(n).$$

We require knowledge of  $G_k(n)$  for all odd  $n$ .

**Lemma 4.3.** (i) (Multiplicativity) Suppose  $m$  and  $n$  are coprime odd integers. Then  $G_k(mn) = G_k(m)G_k(n)$ .

(ii) Suppose  $p^\alpha$  is the largest power of  $p$  dividing  $k$ . (If  $k = 0$  set  $\alpha = \infty$ .) Then for  $\beta \geq 1$

$$G_k(p^\beta) = \begin{cases} 0 & \text{if } \beta \leq \alpha \text{ is odd,} \\ \varphi(p^\beta) & \text{if } \beta \leq \alpha \text{ is even,} \\ -p^\alpha & \text{if } \beta = \alpha + 1 \text{ is even,} \\ \left( \frac{kp^{-\alpha}}{p} \right) p^\alpha \sqrt{p} & \text{if } \beta = \alpha + 1 \text{ is odd,} \\ 0 & \text{if } \beta \geq \alpha + 2. \end{cases}$$

*Proof.* This is [40, 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 [16].

**Lemma 4.4.** Let  $N$  and  $Q$  be positive integers, and let  $a_1, \dots, a_N$  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

$$\sum_{|m| \leq M} \frac{1}{m_2} \left| \sum_{n \leq N} a_n \left( \frac{m}{n} \right) \right|^2 \ll (MN)^\varepsilon N(M+N).$$

*Proof.* This is [40, Lemma 2.4]. □

**Lemma 4.5.** Suppose  $\sigma + it$  is a complex number with  $\sigma \geq \frac{1}{2}$ . 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 [40, Lemma 2.5]. □

## 5. SIEVE ESTIMATES

Our main sieve will be a variant of the Selberg sieve (see [15, Chapter 7]). To lessen the volume of calculations, we also use Brun's pure sieve [15, Chapter 6] as a preliminary sieve to handle small prime factors. We set

$$(5.1) \quad z_0 := \exp((\log X)^{1/3})$$

and

$$(5.2) \quad R := X^\vartheta, \quad \vartheta \in (0, \tfrac{1}{2}) \quad \text{fixed.}$$

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

$$P(y) = \prod_{p \leq y} p.$$

Then, for  $n \asymp X$ , our basic sieve inequality is

$$(5.3) \quad \mathbf{1}_{\{n: n \text{ prime}\}} \leq \mathbf{1}_{\{n: (n, P(z_0))=1\}} \mathbf{1}_{\{n: (n, P(R)/P(z_0))=1\}},$$

We write  $\omega(n)$  for the number of distinct prime factors of  $n$ . To bound the first factor on the right-hand side of (5.3), we use Brun's upper bound sieve condition (see [15, (6.1)])

$$(5.4) \quad \mathbf{1}_{\{n: (n, P(z_0))=1\}}(n) \leq \sum_{\substack{b|(n, P(z_0)) \\ \omega(b) \leq 2r_0}} \mu(b),$$

where

$$r_0 := \lfloor (\log X)^{1/3} \rfloor.$$

We use an “analytic” Selberg sieve (e.g. [35]) for the second factor of (5.3). 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

$$(5.5) \quad \begin{aligned} \mathbf{1}_{\{n: (n, P(R)/P(z_0))=1\}}(n) &\leq \left( \sum_{\substack{d|n \\ (d, P(z_0))=1}} \mu(d) G\left(\frac{\log d}{\log R}\right) \right)^2 \\ &= \sum_{\substack{j, k \leq R \\ [j, k] | n \\ (jk, P(z_0))=1}} \mu(j) \mu(k) G\left(\frac{\log j}{\log R}\right) G\left(\frac{\log k}{\log R}\right). \end{aligned}$$

We mention also that the properties of  $G$  imply

$$(5.6) \quad \int_0^\infty G'(t)^2 dt = 1 + O\left(\frac{1}{\log \log X}\right) = 1 + o(1).$$

Note that the fundamental theorem of calculus and Cauchy-Schwarz yield the lower bound

$$\int_0^\infty G'(t)^2 dt \geq 1.$$

From (5.3), (5.4), and (5.5), we arrive at the upper bound sieve condition

$$(5.7) \quad \mathbf{1}_{\{n:n \text{ prime}\}}(n) \leq \sum_{d|n} \lambda_d,$$

where the coefficients  $\lambda_d$  are defined by

$$(5.8) \quad \lambda_d = \sum_{\substack{b|P(z_0) \\ \omega(b) \leq 2r_0}} \sum_{\substack{m,n \leq R \\ b[m,n]=d \\ (mn, P(z_0))=1}} \mu(b)\mu(m)\mu(n) G\left(\frac{\log m}{\log R}\right) G\left(\frac{\log n}{\log R}\right).$$

If  $b|P(z_0)$  and  $\omega(b) \leq 2r_0$ , then  $b \leq z_0^{2r_0} \leq \exp(2(\log X)^{2/3})$ . Hence  $\lambda_d \neq 0$  only for  $d \leq D$ , where

$$(5.9) \quad D = R^2 \exp(2(\log X)^{2/3}) \ll_{\varepsilon} R^2 X^{\varepsilon}.$$

In our evaluation of sums involving the sieve coefficients (5.8) we use the following version of the fundamental lemma of sieve theory (see also [15, Section 6.5]).

**Lemma 5.1.** *Let  $0 < \delta < 1$  be a fixed constant,  $r$  a positive integer with  $r \asymp (\log X)^{\delta}$ , and  $z_0$  as in (5.1). Suppose that  $g$  is a multiplicative function such that  $|g(p)| \ll 1$  uniformly for all primes  $p$ . Then*

$$\sum_{\substack{b|P(z_0) \\ \omega(b) \leq r \\ (b,\ell)=1}} \frac{\mu(b)}{b} g(b) = \prod_{\substack{p \leq z_0 \\ p \nmid \ell}} \left(1 - \frac{g(p)}{p}\right) + O\left(\exp(-r \log \log r)\right)$$

uniformly for all positive integers  $\ell$ .

*Proof.* The proof is standard. Complete the sum on the left-hand side by adding to it all the terms with  $\omega(b) > r$ . Bound the sum of those terms by taking absolute values, dropping by positivity the condition  $(b, \ell) = 1$ . The resulting bound is  $\ll \exp(-(1+o(1))r \log r) \ll \exp(-r \log \log r)$  (e.g. [18, §6.3]). The completed sum is equal to the Euler product on the right-hand side.  $\square$

The basic tool in our application of the Selberg sieve is the following lemma.

**Lemma 5.2.** *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  $\vartheta$  (see (5.2)) with  $E_0(X) \rightarrow 0$  as  $X \rightarrow \infty$ , such that*

$$(5.10) \quad \begin{aligned} \sum_{\substack{m,n \leq R \\ (mn, \ell P(z_0))=1}} \frac{\mu(m)\mu(n)}{[m,n]} G\left(\frac{\log m}{\log R}\right) G\left(\frac{\log n}{\log R}\right) \prod_{p|mn} (1 + h(p)) \\ = \frac{1 + E_0(X)}{\log R} \prod_{p \leq z_0} \left(1 - \frac{1}{p}\right)^{-1} + O_{\varepsilon, A} \left(\frac{1}{(\log R)^A}\right), \end{aligned}$$

uniformly for  $\ell \ll X^{O(1)}$ .

*Proof.* Let  $\mathcal{S}$  denote the left-hand side of (5.10). If  $m, n \leq R$  and  $(mn, P(z_0)) = 1$ , then  $\omega(mn) \ll \log R$ , and each prime dividing  $mn$  is larger than  $z_0$ . Thus

$$\prod_{p|mn} (1 + h(p)) = 1 + O_\varepsilon \left( \frac{\log R}{z_0^\varepsilon} \right),$$

and so

$$(5.11) \quad \mathcal{S} = \sum_{\substack{m, n \leq R \\ (mn, \ell P(z_0))=1}} \frac{\mu(m)\mu(n)}{[m, n]} G\left(\frac{\log m}{\log R}\right) G\left(\frac{\log n}{\log R}\right) + O\left(\frac{(\log R)^4}{z_0^\varepsilon}\right).$$

We may ignore the condition  $(mn, \ell) = 1$  in (5.11) because

$$\sum_{\substack{m, n \leq R \\ (mn, P(z_0))=1 \\ (mn, \ell) > 1}} \frac{1}{[m, n]} \leq \sum_{\substack{m, n \leq R \\ (mn, P(z_0))=1}} \frac{1}{[m, n]} \sum_{\substack{p|\ell \\ p|mn}} 1 \ll (\log R)^3 \sum_{\substack{p|\ell \\ p > z_0}} \frac{1}{p} \ll \frac{(\log \ell)(\log R)^3}{z_0}.$$

We next insert the Fourier inversion formula

$$(5.12) \quad G(t) = \int_{-\infty}^{\infty} g(z) e^{-t(1+iz)} dz$$

into (5.11), where

$$(5.13) \quad g(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^t G(t) e^{izt} dt.$$

As  $G(t)$  is supported in  $[-1, 1]$  we may let  $m$  and  $n$  range over all positive integers in (5.11). We then interchange the order of summation and integration and write the sum as an Euler product to deduce that

$$(5.14) \quad \mathcal{S} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(z_1) g(z_2) \prod_{p > z_0} \left( 1 - \frac{1}{p^{1+\frac{1+iz_1}{\log R}}} - \frac{1}{p^{1+\frac{1+iz_2}{\log R}}} + \frac{1}{p^{1+\frac{2+iz_1+iz_2}{\log R}}} \right) dz_1 dz_2 + O\left(\frac{(\log R)^4}{z_0^\varepsilon}\right).$$

By integrating (5.13) 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 (5.14) 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

(5.15)

$$\begin{aligned} \mathcal{S} = & \iint_{|z_i| \leq \sqrt{\log R}} g(z_1)g(z_2) \frac{\zeta\left(1 + \frac{2+iz_1+iz_2}{\log R}\right)}{\zeta\left(1 + \frac{1+iz_1}{\log R}\right)\zeta\left(1 + \frac{1+iz_2}{\log R}\right)} \\ & \times \prod_{p \leq z_0} \frac{\left(1 - \frac{1}{p^{1+\frac{2+iz_1+iz_2}{\log R}}}\right)}{\left(1 - \frac{1}{p^{1+\frac{1+iz_1}{\log R}}}\right)\left(1 - \frac{1}{p^{1+\frac{1+iz_2}{\log R}}}\right)} \prod_{p > z_0} \left(1 + O\left(\frac{1}{p^2}\right)\right) dz_1 dz_2 + O\left(\frac{1}{(\log R)^A}\right). \end{aligned}$$

The product over primes  $p > z_0$  in (5.15) is  $1 + O(1/z_0)$ . To estimate the product over  $p \leq z_0$ , observe that if  $|s| \ll (\log R)^{-1/2}$ , then

$$\sum_{p \leq z_0} \frac{1}{p-1} (1 - p^{-s}) \ll \sum_{p \leq z_0} \frac{|s| \log p}{p} \ll |s| \log z_0 \ll \frac{(\log X)^{1/3}}{(\log R)^{1/2}},$$

which implies that

$$\begin{aligned} \prod_{p \leq z_0} \left(1 - \frac{1}{p^{1+s}}\right) &= \exp\left(\sum_{p \leq z_0} \log\left(1 + \frac{1}{p-1} (1 - p^{-s})\right)\right) \prod_{p \leq z_0} \left(1 - \frac{1}{p}\right) \\ &= \left(1 + O\left(\frac{(\log X)^{1/3}}{(\log R)^{1/2}}\right)\right) \prod_{p \leq z_0} \left(1 - \frac{1}{p}\right). \end{aligned}$$

We may also expand each zeta-function in (5.15) into its Laurent series. With these approximations, we deduce from (5.15) that

$$\begin{aligned} \mathcal{S} = & \frac{1}{\log R} \prod_{p \leq z_0} \left(1 - \frac{1}{p}\right)^{-1} \iint_{|z_i| \leq \sqrt{\log R}} g(z_1)g(z_2) \frac{(1+iz_1)(1+iz_2)}{2+iz_1+iz_2} (1 + E(X, \vartheta, z_1, z_2)) dz_1 dz_2 \\ & + O\left((\log R)^{-A}\right), \end{aligned}$$

uniformly for  $\log \ell \ll \log X$ . Here  $E(X, \vartheta, z_1, z_2)$  tends to zero as  $X \rightarrow \infty$ . 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 (5.12) under the integral sign and Fubini's theorem, we find

$$\iint_{\mathbb{R}^2} g(z_1)g(z_2) \frac{(1+iz_1)(1+iz_2)}{2+iz_1+iz_2} dz_2 dz_1 = \int_0^\infty G'(t)^2 dt.$$

The lemma now follows from this and (5.6).  $\square$

**Lemma 5.3.** *Let  $\lambda_d$  and  $D$  be as defined in (5.8) and (5.9), 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 5.2 we have*

$$\sum_{\substack{d \leq D \\ (d, \ell)=1}} \frac{\lambda_d}{d} g(d) = \frac{1 + E_0(X)}{\log R} \prod_{\substack{p \leq z_0 \\ p \nmid \ell}} \left(1 - \frac{g(p)}{p}\right) \prod_{p \leq z_0} \left(1 - \frac{1}{p}\right)^{-1} + O_{A, \varepsilon} \left(\frac{1}{(\log R)^A}\right),$$



uniformly for  $\ell \ll X^{O(1)}$ .

*Proof.* The definitions (5.8) and (5.9) of  $\lambda_d$  and  $D$  imply

$$\sum_{\substack{d \leq D \\ (d, \ell)=1}} \frac{\lambda_d}{d} g(d) = \sum_{\substack{b|P(z_0) \\ \omega(b) \leq 2r_0 \\ (b, \ell)=1}} \sum_{\substack{m, n \leq R \\ (mn, \ell P(z_0))=1}} \frac{\mu(b)\mu(m)\mu(n)}{b[m, n]} G\left(\frac{\log m}{\log R}\right) G\left(\frac{\log n}{\log R}\right) g(b[m, n]).$$

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 5.2 and then Lemma 5.1 to arrive at Lemma 5.3.  $\square$

**Lemma 5.4.** *Let  $\lambda_d, D, g$  be as in Lemma 5.3. 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 5.2 we have*

$$\begin{aligned} \sum_{\substack{d \leq D \\ (d, \ell)=1}} \frac{\lambda_d}{d} g(d) \sum_{p|d} h(p) &= -\frac{1 + E_0(X)}{\log R} \prod_{p \leq z_0} \left(1 - \frac{1}{p}\right)^{-1} \\ &\quad \times \sum_{\substack{p \leq z_0 \\ p \nmid \ell}} \frac{g(p)h(p)}{p} \prod_{\substack{q \leq z_0 \\ q \nmid p\ell}} \left(1 - \frac{g(q)}{q}\right) + O_{A, \varepsilon} \left(\frac{1}{(\log R)^A}\right), \end{aligned}$$

uniformly for all integers  $\ell$  such that  $\log \ell \ll \log X$ . (Here, the index  $q$  runs over primes  $q$ .)

*Proof.* The definitions (5.8) and (5.9) of  $\lambda_d$  and  $D$  imply

$$\begin{aligned} \sum_{\substack{d \leq D \\ (d, \ell)=1}} \frac{\lambda_d}{d} g(d) \sum_{p|d} h(p) &= \sum_{\substack{b|P(z_0) \\ \omega(b) \leq 2r_0 \\ (b, \ell)=1}} \sum_{\substack{m, n \leq R \\ (mn, \ell P(z_0))=1}} \frac{\mu(b)\mu(m)\mu(n)}{b[m, n]} G\left(\frac{\log m}{\log R}\right) G\left(\frac{\log n}{\log R}\right) \\ &\quad \times g(b[m, n]) \sum_{p|bmn} h(p). \end{aligned}$$

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  $mn \leq R^2$  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 5.2 to deduce that

$$\begin{aligned} \sum_{\substack{d \leq D \\ (d, \ell)=1}} \frac{\lambda_d}{d} g(d) \sum_{p|d} h(p) &= \frac{1 + E_0(X)}{\log R} \prod_{p \leq z_0} \left(1 - \frac{1}{p}\right)^{-1} \\ (5.16) \quad &\quad \times \sum_{\substack{b|P(z_0) \\ \omega(b) \leq 2r_0 \\ (b, \ell)=1}} \frac{\mu(b)}{b} g(b) \sum_{p|b} h(p) + O_A \left(\frac{1}{(\log R)^A}\right). \end{aligned}$$

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 5.4 now follows from Lemma 5.1 and (5.16).  $\square$

## 6. THE MOLLIFIED FIRST MOMENT

Our goal in this section is to asymptotically evaluate  $S_1$ . Recall from (3.1) that

$$S_1 = \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) L\left(\frac{1}{2}, \chi_p\right) M(p).$$

Recall the definition of  $M(p)$  from (3.3), and the choice (3.5) we made for the mollifier coefficients  $b_m$ . We shall prove the following result.

**Proposition 6.1.** *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. We apply Lemma 4.2 to write  $L(\frac{1}{2}, \chi_p)$  as a Dirichlet series. We insert the definition of  $M(p)$  and obtain

$$S_1 = \frac{2}{\left(1 - \frac{1}{\sqrt{2}}\right)^2} \sum_{\substack{m \leq M \\ m \text{ odd}}} \frac{b_m}{\sqrt{m}} \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{1}{\sqrt{n}} \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) \omega_1\left(n \sqrt{\frac{\pi}{p}}\right) \left(\frac{mn}{p}\right).$$

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

$$(6.1) \quad \begin{aligned} S_1^\square &= \frac{2}{\left(1 - \frac{1}{\sqrt{2}}\right)^2} \sum_{\substack{m \leq M \\ m \text{ odd}}} \sum_{\substack{n=1 \\ n \text{ odd} \\ mn=\square}}^{\infty} \frac{b_m}{\sqrt{m}} \frac{1}{\sqrt{n}} \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) \omega_1\left(n \sqrt{\frac{\pi}{p}}\right) \left(\frac{mn}{p}\right), \\ S_1^\neq &= \frac{2}{\left(1 - \frac{1}{\sqrt{2}}\right)^2} \sum_{\substack{m \leq M \\ m \text{ odd}}} \sum_{\substack{n=1 \\ n \text{ odd} \\ mn \neq \square}}^{\infty} \frac{b_m}{\sqrt{m}} \frac{1}{\sqrt{n}} \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) \omega_1\left(n \sqrt{\frac{\pi}{p}}\right) \left(\frac{mn}{p}\right). \end{aligned}$$

We treat first the main term  $S_1^\square$ , and later we will bound the error term  $S_1^\neq$ .

**6.1. 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

$$S_1^\square = \frac{2}{\left(1 - \frac{1}{\sqrt{2}}\right)^2} \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) \sum_{\substack{m \leq M \\ (m, 2p)=1}} \frac{b_m}{m} \sum_{\substack{k=1 \\ (k, 2p)=1}}^{\infty} \frac{1}{k} \omega_1\left(mk^2 \sqrt{\frac{\pi}{p}}\right).$$

By the rapid decay of  $\omega_1$  (Lemma 4.1) 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  $m \leq M < p$ . We insert the definition (4.1) of  $\omega_1(\xi)$  and interchange to deduce that for any  $c > 0$  we have

$$\begin{aligned} & \sum_{\substack{k=1 \\ (k, 2)=1}}^{\infty} \frac{1}{k} \omega_1\left(mk^2 \sqrt{\frac{\pi}{p}}\right) \\ &= \frac{1}{2\pi i} \int_{(c)} \frac{\Gamma(\frac{s}{2} + \frac{1}{4})}{\Gamma(\frac{1}{4})} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right) \left(1 - \frac{1}{2^{1+2s}}\right) \zeta(1+2s) \left(\frac{p}{\pi}\right)^{s/2} m^{-s} \frac{ds}{s}. \end{aligned}$$

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 (3.4). Writing the residue at  $s = 0$  as an integral along a small circle around 0, we deduce that

$$\begin{aligned} (6.1.1) \quad S_1^\square &= O(X^{1-\varepsilon}) + \frac{2}{\left(1 - \frac{1}{\sqrt{2}}\right)^2} \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) \sum_{\substack{m \leq M \\ (m, 2)=1}} \frac{b_m}{m} \\ &\quad \times \frac{1}{2\pi i} \oint_{|s|=\frac{1}{2\log X}} \frac{\Gamma(\frac{s}{2} + \frac{1}{4})}{\Gamma(\frac{1}{4})} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right) \left(1 - \frac{1}{2^{1+2s}}\right) \zeta(1+2s) \left(\frac{p}{\pi}\right)^{s/2} m^{-s} \frac{ds}{s}. \end{aligned}$$

We next use the definition  $b_m = \mu(m)H\left(\frac{\log m}{\log M}\right)$  and the Fourier inversion formula (compare with (5.12), (5.13))

$$(6.1.2) \quad H(t) = \int_{-\infty}^{\infty} h(z) e^{-t(1+iz)} dz,$$

where

$$(6.1.3) \quad h(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^t H(t) e^{izt} dt,$$

to write

$$\begin{aligned} \sum_{\substack{m \leq M \\ (m, 2)=1}} \frac{b_m}{m} m^{-s} &= \int_{-\infty}^{\infty} h(z) \sum_{\substack{m=1 \\ (m, 2)=1}}^{\infty} \frac{\mu(m)}{m^{1+s+\frac{1+iz}{\log M}}} dz \\ &= \int_{-\infty}^{\infty} h(z) \left(1 - \frac{1}{2^{1+s+\frac{1+iz}{\log M}}}\right)^{-1} \zeta^{-1}\left(1+s+\frac{1+iz}{\log M}\right) dz. \end{aligned}$$

Here we have used the fact that  $b_m = 0$  for  $m > M$  due to the support of  $H(t)$ . From repeated integration by parts we obtain

$$(6.1.4) \quad h(z) \ll_j \frac{1}{(1 + |z|)^j},$$

and therefore we may truncate this integral to the range  $|z| \leq \sqrt{\log M}$ . Thus,

$$\begin{aligned} \sum_{\substack{m \leq M \\ (m,2)=1}} \frac{b_m}{m} m^{-s} &= \int_{|z| \leq \sqrt{\log M}} h(z) \left(1 - \frac{1}{2^{1+s+\frac{1+iz}{\log M}}}\right)^{-1} \zeta^{-1} \left(1 + s + \frac{1+iz}{\log M}\right) dz \\ &\quad + O_A \left( \frac{1}{(\log X)^A} \right). \end{aligned}$$

For  $|s| = \frac{1}{2\log X}$  and  $|z| \leq \sqrt{\log M}$ , 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

$$\begin{aligned} \sum_{\substack{m \leq M \\ (m,2)=1}} \frac{b_m}{m} m^{-s} &= 2 \int_{|z| \leq \sqrt{\log M}} h(z) \left(s + \frac{1+iz}{\log M}\right) dz \\ &\quad + O \left( \frac{1}{(\log X)^2} \int_{|z| \leq \sqrt{\log M}} |h(z)|(1 + |z|^2) dz \right) \\ &= 2 \int_{|z| \leq \sqrt{\log M}} h(z) \left(s + \frac{1+iz}{\log M}\right) dz + O \left( \frac{1}{(\log X)^2} \right). \end{aligned}$$

We may extend the range of integration to the entire real line, with negligible error, because of (6.1.4). Differentiating (6.1.2) leads to

$$H'(t) = - \int_{-\infty}^{\infty} (1+iz)h(z)e^{-t(1+iz)} dz.$$

This and (6.1.2) thus imply

$$\int_{-\infty}^{\infty} h(z) \left(s + \frac{1+iz}{\log M}\right) dz = sH(0) - \frac{1}{\log M} H'(0),$$

and hence

$$(6.1.5) \quad \sum_{\substack{m \leq M \\ (m,2)=1}} \frac{b_m}{m} m^{-s} = 2sH(0) - \frac{2}{\log M} H'(0) + O \left( \frac{1}{(\log X)^2} \right).$$

We insert (6.1.5) into (6.1.1) to obtain

$$\begin{aligned} S_1^\square &= \frac{4}{\left(1 - \frac{1}{\sqrt{2}}\right)^2} \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi \left( \frac{p}{X} \right) \frac{1}{2\pi i} \oint_{|s|=\frac{1}{2\log X}} \frac{\Gamma(\frac{s}{2} + \frac{1}{4})}{\Gamma(\frac{1}{4})} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right) \\ &\quad \times \left(1 - \frac{1}{2^{1+2s}}\right) \zeta(1+2s) \left(\frac{p}{\pi}\right)^{s/2} \left(sH(0) - \frac{1}{\log M} H'(0)\right) \frac{ds}{s} + O \left( \frac{X}{\log X} \right). \end{aligned}$$

We evaluate the integral using the formula

$$(6.1.6) \quad \operatorname{Res}_{s=0} g(s) = \frac{1}{(n-1)!} \frac{d^{n-1}}{ds^{n-1}} s^n g(s) \Big|_{s=0}$$

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  $\Phi$  we have  $\log p = \log X + O(1)$ . We then use the prime number theorem in arithmetic progressions and partial summation to obtain

$$S_1^\square = \frac{1}{\left(1 - \frac{1}{\sqrt{2}}\right)} \left(H(0) - \frac{\log X}{2 \log M} H'(0)\right) \frac{X}{4} \widehat{\Phi}(0) + O\left(\frac{X}{\log X}\right).$$

This gives the main term for Proposition 6.1, by (2.2).

**6.2. Preparation of the off-diagonal.** We turn to bounding  $S_1^\neq$ . In order to complete the proof of Proposition 6.1, we shall prove

$$(6.2.1) \quad S_1^\neq \ll \frac{X}{(\log X)^{1-\varepsilon}}.$$

We need to perform some technical massaging before  $S_1^\neq$  is in a suitable form. Recall from (6.1) that

$$S_1^\neq = \frac{2}{\left(1 - \frac{1}{\sqrt{2}}\right)^2} \sum_{\substack{m \leq M \\ m \text{ odd}}} \sum_{\substack{n=1 \\ n \text{ odd} \\ mn \neq \square}}^\infty \frac{b_m}{\sqrt{mn}} \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) \omega_1\left(n \sqrt{\frac{\pi}{p}}\right) \left(\frac{mn}{p}\right).$$

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

$$S_1^\neq = \frac{2}{\left(1 - \frac{1}{\sqrt{2}}\right)^2} \sum_{\substack{m \leq M \\ m \text{ odd}}} \frac{b_m}{\sqrt{m}} \sum_{\substack{r=1 \\ r \text{ odd} \\ r \neq m}}^\infty \sum_{\substack{k=1 \\ k \text{ odd}}}^\infty \frac{\mu^2(r)}{k \sqrt{r}} \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) \omega_1\left(rk^2 \sqrt{\frac{\pi}{p}}\right) \left(\frac{mrk^2}{p}\right)$$

We next factor out the greatest common divisor, say  $g$ , of  $m$  and  $r$ . We change variables  $m \rightarrow gm, r \rightarrow gr$  and obtain

$$\begin{aligned} S_1^\neq &= \frac{2}{\left(1 - \frac{1}{\sqrt{2}}\right)^2} \sum_{g \text{ odd}} \frac{\mu^2(g)}{g} \sum_{\substack{m \leq M/g \\ (m, 2g)=1}} \frac{b_{mg}}{\sqrt{m}} \sum_{\substack{r=1 \\ (r, 2g)=1 \\ (m, r)=1 \\ mr > 1}}^\infty \frac{\mu^2(r)}{\sqrt{r}} \sum_{\substack{k=1 \\ k \text{ odd}}}^\infty \frac{1}{k} \\ &\quad \times \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) \omega_1\left(grk^2 \sqrt{\frac{\pi}{p}}\right) \left(\frac{mr g^2 k^2}{p}\right). \end{aligned}$$

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  $\left(\frac{g^2 k^2}{p}\right) = 1$  for  $p \nmid gk$  and  $= 0$  otherwise. Since  $g \leq M < p$ , the condition  $p \nmid g$  is automatically satisfied. By Lemma 4.1 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 over  $r$  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  $\omega_1$  again to extend the sum over  $k$  to infinity. It follows that

$$(6.2.2) \quad S_1^\neq = \frac{2}{\left(1 - \frac{1}{\sqrt{2}}\right)^2} \sum_{g \text{ odd}} \frac{\mu^2(g)}{g} \sum_{\substack{m \leq M/g \\ (m, 2g)=1}} \frac{b_{mg}}{\sqrt{m}} \sum_{\substack{r \leq X^{1/2+\varepsilon} \\ (r, 2g)=1 \\ (m, r)=1 \\ mr > 1}} \frac{\mu^2(r)}{\sqrt{r}} \sum_{\substack{k=1 \\ k \text{ odd}}}^{\infty} \frac{1}{k} \\ \times \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) \omega_1\left(grk^2 \sqrt{\frac{\pi}{p}}\right) \left(\frac{mr}{p}\right) + O(X^{-1}).$$

We next detect the congruence condition  $p \equiv 1 \pmod{8}$  with multiplicative characters modulo 8 and write

$$(6.2.3) \quad \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) \omega_1\left(grk^2 \sqrt{\frac{\pi}{p}}\right) \left(\frac{mr}{p}\right) \\ = \frac{1}{4} \sum_{\gamma \in \{\pm 1, \pm 2\}} \sum_p (\log p) \Phi\left(\frac{p}{X}\right) \omega_1\left(grk^2 \sqrt{\frac{\pi}{p}}\right) \left(\frac{\gamma mr}{p}\right).$$

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  $\gamma mr$  is square-free. Therefore  $\gamma mr \equiv 1, 2$ , or  $3 \pmod{4}$ . 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  $3 \pmod{4}$ , then  $\left(\frac{4\gamma mr}{\cdot}\right)$  is a real primitive character modulo  $|4\gamma mr|$  (see [9, 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 (6.2.3) is equal to

$$(6.2.4) \quad \frac{1}{4} \sum_{\gamma \in \{\pm 1, \pm 2\}} \sum_p (\log p) \Phi\left(\frac{p}{X}\right) \omega_1\left(grk^2 \sqrt{\frac{\pi}{p}}\right) \chi_{\gamma mr}(p),$$

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  $3 \pmod{4}$ , so that  $\chi_{\gamma mr}(\cdot)$  is a real primitive character for all the relevant  $\gamma, m, r$ . Also, since  $mr > 1$ , we see that  $\gamma mr$  is never 1, so each  $\chi_{\gamma mr}$  is nonprincipal.

We insert the definition of  $\omega_1$  into (6.2.4) in order to facilitate a separation of variables. Recalling (6.2.2) and (6.2.3), we interchange the order of summation and integration to obtain

$$S_1^\neq = O(1) + \frac{2}{\left(1 - \frac{1}{\sqrt{2}}\right)^2} \sum_{g \text{ odd}} \frac{\mu^2(g)}{g} \sum_{\substack{m \leq M/g \\ (m, 2g)=1}} \frac{b_{mg}}{\sqrt{m}} \sum_{\substack{r \leq X^{1/2+\varepsilon} \\ (r, 2g)=1 \\ (m, r)=1 \\ mr > 1}} \frac{\mu^2(r)}{\sqrt{r}} \frac{1}{4} \sum_{\gamma \in \{\pm 1, \pm 2\}} \sum_{\substack{k=1 \\ k \text{ odd}}}^{\infty} \frac{1}{k} \\ \times \frac{1}{2\pi i} \int_{(c)} \frac{\Gamma\left(\frac{s}{2} + \frac{1}{4}\right)}{\Gamma\left(\frac{1}{4}\right)} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right) \pi^{-s/2} (grk^2)^{-s} \sum_p (\log p) \Phi\left(\frac{p}{X}\right) \chi_{\gamma mr}(p) p^{s/2} \frac{ds}{s}.$$

We choose  $c = \frac{1}{\log X}$ , so that  $p^{s/2}$  is bounded in absolute value. We can put the summation over  $k$  inside of the integral, where it becomes a zeta factor, and we obtain

$$S_1^\# = O(1) + \frac{2}{\left(1 - \frac{1}{\sqrt{2}}\right)^2} \sum_{g \text{ odd}} \frac{\mu^2(g)}{g} \sum_{\substack{m \leq M/g \\ (m, 2g)=1}} \frac{b_{mg}}{\sqrt{m}} \sum_{\substack{r \leq X^{1/2+\varepsilon} \\ (r, 2g)=1 \\ (m, r)=1 \\ mr > 1}} \frac{\mu^2(r)}{\sqrt{r}} \frac{1}{4} \sum_{\gamma \in \{\pm 1, \pm 2\}} \frac{1}{2\pi i} \int_{(c)} \frac{\Gamma(\frac{s}{2} + \frac{1}{4})}{\Gamma(\frac{1}{4})} \\ \times \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right) \left(1 - \frac{1}{2^{1+2s}}\right) \zeta(1+2s) \pi^{-s/2} (gr)^{-s} \sum_p (\log p) \Phi\left(\frac{p}{X}\right) \chi_{\gamma mr}(p) p^{s/2} \frac{ds}{s}.$$

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  $\Gamma$  function in vertical strips we can truncate the integral to  $|\operatorname{Im}(s)| \leq (\log X)^2$ , at the cost of a negligible error. We therefore obtain

$$(6.2.5) \quad S_1^\# = O(X^{1-\varepsilon}) + \frac{2}{\left(1 - \frac{1}{\sqrt{2}}\right)^2} \sum_{g \text{ odd}} \frac{\mu^2(g)}{g} \sum_{\substack{m \leq M/g \\ (m, 2g)=1}} \frac{b_{mg}}{\sqrt{m}} \sum_{\substack{r \leq X^{1/2+\varepsilon} \\ (r, 2g)=1 \\ (m, r)=1 \\ mr > 1}} \frac{\mu^2(r)}{\sqrt{r}} \frac{1}{4} \sum_{\gamma \in \{\pm 1, \pm 2\}} \\ \times \frac{1}{2\pi i} \int_{\frac{1}{\log X} - i(\log X)^2}^{\frac{1}{\log X} + i(\log X)^2} \frac{\Gamma(\frac{s}{2} + \frac{1}{4})}{\Gamma(\frac{1}{4})} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right) \left(1 - \frac{1}{2^{1+2s}}\right) \zeta(1+2s) \\ \times \left(\frac{X}{\pi}\right)^{s/2} (gr)^{-s} \sum_n \Lambda(n) \Phi_s\left(\frac{n}{X}\right) \chi_{\gamma mr}(n) \frac{ds}{s},$$

where we recall (see (2.1)) that  $\Phi_s(x) = \Phi(x)x^{s/2}$ .

Having arrived at (6.2.5), we are finished with the preparatory technical manipulations. We proceed to show that  $S_1^\#$  is small. As discussed in Section 3, 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

$$(6.2.6) \quad S_1^\# = E_1 + E_2 + O(X^{1-\varepsilon}),$$

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

$$(6.2.7) \quad E_1 \ll \frac{X}{(\log X)^{1-\varepsilon}}, \\ E_2 \ll X \exp(-c\varpi\sqrt{\log X}),$$

where  $c > 0$  is some absolute constant. Taking together (6.2.6) and (6.2.7) gives (6.2.1), and this yields Proposition 6.1. It therefore suffices to show (6.2.7).

We remark that there is some flexibility in choosing the “boundary” between Regime I and Regime II. If one wishes to use Siegel’s theorem, then one may choose the boundary to be  $mr \leq (\log X)^B$ , and then the treatment of Regime I is essentially a simple application of the Siegel-Walfisz theorem. The error bound one obtains for Regime II would then be a weaker  $O_A(X/(\log X)^A)$ , compared to the stronger bound of (6.2.7).

**6.3. Regime I.** We first bound  $E_1$ , which is precisely the contribution of Regime I. Apart from an unimportant factor of  $\frac{1}{4} \cdot 2(1 - 2^{-1/2})^{-2}$ , we have

$$\begin{aligned}
 (6.3.1) \quad E_1 &:= \sum_{g \text{ odd}} \frac{\mu^2(g)}{g} \sum_{\substack{m \leq M/g \\ (m, 2g)=1}} \frac{b_{mg}}{\sqrt{m}} \sum_{\substack{r \leq X^{1/2+\varepsilon} \\ (r, 2g)=1 \\ (m, r)=1 \\ 1 < mr \ll \exp(\varpi \sqrt{\log X})}} \frac{\mu^2(r)}{\sqrt{r}} \sum_{\gamma \in \{\pm 1, \pm 2\}} \\
 &\times \frac{1}{2\pi i} \int_{\frac{1}{\log X} - i(\log X)^2}^{\frac{1}{\log X} + i(\log X)^2} \frac{\Gamma(\frac{s}{2} + \frac{1}{4})}{\Gamma(\frac{1}{4})} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right) \left(1 - \frac{1}{2^{1+2s}}\right) \zeta(1+2s) \\
 &\times \left(\frac{X}{\pi}\right)^{s/2} (gr)^{-s} \sum_n \Lambda(n) \Phi_s\left(\frac{n}{X}\right) \chi_{\gamma mr}(n) \frac{ds}{s}.
 \end{aligned}$$

We transform the sum over  $n$  with partial summation to obtain

$$(6.3.2) \quad \sum_n \Lambda(n) \Phi_s\left(\frac{n}{X}\right) \chi_{\gamma mr}(n) = - \int_0^\infty \frac{1}{X} \Phi'_s\left(\frac{w}{X}\right) \left(\sum_{n \leq w} \Lambda(n) \chi_{\gamma mr}(n)\right) dw.$$

By [12, equation (8) of Chapter 20], we have

$$(6.3.3) \quad \sum_{n \leq w} \Lambda(n) \chi_{\gamma mr}(n) = -\frac{w^{\beta_1}}{\beta_1} + O\left(w \exp(-c_1 \sqrt{\log w})\right),$$

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  $\beta_1$  which satisfies  $\beta_1 > 1 - \frac{c_2}{\log |\gamma mr|}$  for some sufficiently small constant  $c_2 > 0$ . (It is important here that  $mr \ll \exp(\varpi \sqrt{\log X})$ .) All the constants in (6.3.3), implied or otherwise, are effective.

The contribution from the error term in (6.3.3) is easy to control. Observe that

$$(6.3.4) \quad \int_0^\infty \frac{1}{X} \left| \Phi'_s\left(\frac{w}{X}\right) \right| dw = \int_0^\infty |\Phi'_s(u)| du \ll |s| + \log X,$$

uniformly in  $s$  with  $\text{Re}(s)$  bounded. Taking (6.3.1), (6.3.2) and (6.3.4) together, we see the error term of (6.3.3) contributes

$$(6.3.5) \quad \ll X \exp((c_3 \varpi - c_1) \sqrt{\log X})$$

to  $E_1$ , where  $c_3 > 0$  is some absolute constant; we have used here the fact that  $mr \ll \exp(\varpi \sqrt{\log X})$ . The bound (6.3.5) is more than adequate for (6.2.7) provided we choose  $\varpi > 0$  sufficiently small in terms of  $c_1$ .

Since  $mr \ll \exp(\varpi \sqrt{\log X})$ , the conductor of the primitive character  $\chi_{\gamma mr}$  is  $\ll \exp(\varpi \sqrt{\log X}) \leq \exp(2\varpi \sqrt{\log X})$ . We apply Page’s theorem [12, 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

$$(6.3.6) \quad \beta_1 > 1 - \frac{c_4}{2\varpi\sqrt{\log X}}.$$

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 (6.3.2) and (6.3.3). 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 (6.3.6) does exist, for otherwise we already have the acceptable bound (6.3.5) for  $E_1$ . Let  $q^*$  be such that  $\chi_{q^*}$  is the exceptional character with a real zero  $\beta_1$  satisfying (6.3.6). Then we have

$$(6.3.7) \quad \begin{aligned} E_1 = & -\frac{1}{2\pi i} \frac{X^{\beta_1}}{\sqrt{|q^*|}} \sqrt{\gamma^*} \int_{\frac{1}{\log X} - i(\log X)^2}^{\frac{1}{\log X} + i(\log X)^2} \frac{\Gamma(\frac{s}{2} + \frac{1}{4})}{\Gamma(\frac{1}{4})} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right) \left(1 - \frac{1}{2^{1+2s}}\right) \\ & \times \left(\frac{X}{\pi}\right)^{s/2} \zeta(1+2s) \Phi^\dagger\left(\frac{s}{2} + \beta_1\right) \sum_{\substack{1 < mr \ll \exp(\varpi\sqrt{\log X}) \\ (mr, 2)=1 \\ (m, r)=1 \\ \gamma mr = q^*}} \mu^2(r) r^{-s} \sum_{(g, 2mr)=1} \frac{\mu^2(g) b_{gm}}{g^{1+s}} \frac{ds}{s} \\ & + O\left(X \exp(-c_5 \sqrt{\log X})\right), \end{aligned}$$

where  $c_5 > 0$  is some absolute constant,  $\gamma \in \{\pm 1, \pm 2\}$  is fixed, and  $\gamma^*$  is some bounded power of two. There is at most one choice of  $\gamma$  that can give rise to the exceptional character since  $mr$  is odd and positive.

We next write  $b_{gm} = \mu(gm) H(\frac{\log gm}{\log M})$  and apply Fourier inversion as in (6.1.2), (6.1.3) to obtain

$$(6.3.8) \quad \sum_{(g, 2mr)=1} \frac{\mu^2(g) b_{gm}}{g^{1+s}} = \mu(m) \int_{-\infty}^\infty \frac{1}{m^{\frac{1+iz}{\log M}}} h(z) \prod_{p|2mr} \left(1 - \frac{1}{p^{1+s+\frac{1+iz}{\log M}}}\right)^{-1} \zeta^{-1}\left(1+s+\frac{1+iz}{\log M}\right) dz.$$

By (6.1.4) we can truncate the integral in (6.3.8) 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 (6.3.7)

$$\ll_B \frac{X}{(\log X)^{B-O(1)}},$$

which is acceptable. We therefore have

(6.3.9)

$$\begin{aligned}
E_1 = & -X^{\beta_1} \frac{\sqrt{\gamma^*}}{\sqrt{|q^*|}} \sum_{\substack{1 < mr \ll \exp(\varpi \sqrt{\log X}) \\ (mr, 2)=1 \\ (m, r)=1 \\ \gamma mr = q^*}} \mu(m) \mu^2(r) \frac{1}{2\pi i} \int_{\frac{1}{\log X} - i(\log X)^2}^{\frac{1}{\log X} + i(\log X)^2} \frac{\Gamma(\frac{s}{2} + \frac{1}{4})}{\Gamma(\frac{1}{4})} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right) r^{-s} \\
& \times \left(1 - \frac{1}{2^{1+2s}}\right) \zeta(1+2s) \left(\frac{X}{\pi}\right)^{s/2} \Phi^\dagger\left(\frac{s}{2} + \beta_1\right) \int_{|z| \leq \sqrt{\log M}} \frac{1}{m^{\frac{1+iz}{\log M}}} h(z) \\
& \times \prod_{p|2mr} \left(1 - \frac{1}{p^{1+s+\frac{1+iz}{\log M}}}\right)^{-1} \zeta^{-1}\left(1 + s + \frac{1+iz}{\log M}\right) dz \frac{ds}{s} + O_A\left(\frac{X}{(\log X)^A}\right).
\end{aligned}$$

We handle the  $s$ -integral in (6.3.9) by moving the line of integration to  $\operatorname{Re}(s) = -\frac{c_6}{\log \log X}$ , where  $c_6 > 0$  is small enough that

$$(6.3.10) \quad \zeta(1+z) \ll \log |\operatorname{Im}(z)| \quad \text{and} \quad \frac{1}{\zeta(1+z)} \ll \log |\operatorname{Im}(z)|$$

for  $\operatorname{Re}(z) \geq -c_6/\log |\operatorname{Im}(z)|$  and  $|\operatorname{Im}(z)| \geq 1$  (see, for example, Theorem 3.5 and (3.11.8) of Titchmarsh [43]). We estimate the integral on the line  $\operatorname{Re}(s) = -\frac{c_6}{\log \log X}$  with trivial estimates, along with the bounds (6.3.10) and

$$|X^{s/2}| = \exp\left(-\frac{c_6}{2} \frac{\log X}{\log \log X}\right).$$

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

(6.3.11)

$$\begin{aligned}
E_1 = & -X^{\beta_1} \frac{\sqrt{\gamma^*}}{\sqrt{|q^*|}} \sum_{\substack{1 < mr \ll \exp(\varpi \sqrt{\log X}) \\ (mr, 2)=1 \\ (m, r)=1 \\ \gamma mr = q^*}} \mu(m) \mu^2(r) \frac{1}{2\pi i} \oint_{|s|=\frac{1}{\log X}} \frac{\Gamma(\frac{s}{2} + \frac{1}{4})}{\Gamma(\frac{1}{4})} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right) r^{-s} \\
& \times \left(1 - \frac{1}{2^{1+2s}}\right) \zeta(1+2s) \left(\frac{X}{\pi}\right)^{s/2} \Phi^\dagger\left(\frac{s}{2} + \beta_1\right) \int_{|z| \leq \sqrt{\log M}} \frac{1}{m^{\frac{1+iz}{\log M}}} h(z) \\
& \times \prod_{p|2mr} \left(1 - \frac{1}{p^{1+s+\frac{1+iz}{\log M}}}\right)^{-1} \zeta^{-1}\left(1 + s + \frac{1+iz}{\log M}\right) dz \frac{ds}{s} + O_A\left(\frac{X}{(\log X)^A}\right).
\end{aligned}$$

We have the bound

$$(6.3.12) \quad \beta_1 < 1 - \frac{c_7}{\sqrt{|q^*|}(\log |q^*|)^2},$$

where  $c_7 > 0$  is a fixed absolute constant (see [12, equation (12) of Chapter 14]). If  $q^*$  satisfies  $|q^*| \leq (\log X)^{2-\varepsilon}$  then by (6.3.12) we derive

$$X^{\beta_1} \ll X \exp(-c_7(\log X)^{\varepsilon/3}).$$

By estimating (6.3.11) trivially we then obtain

$$E_1 \ll_A \frac{X}{(\log X)^A} + X \exp(-c_7(\log X)^{\varepsilon/4}),$$

which is an acceptable bound. We may therefore assume that  $q^*$  satisfies

$$(6.3.13) \quad |q^*| > (\log X)^{2-\varepsilon}.$$

For  $|s| = \frac{1}{\log X}$  we have the bounds

$$\zeta(1+2s) \ll \log X, \quad \zeta^{-1}\left(1+s+\frac{1+iz}{\log M}\right) \ll \frac{1+|z|}{\log X}.$$

Using these bounds and (6.3.13) we deduce by trivial estimation that

$$(6.3.11) \ll \frac{X}{|q^*|^{1/2-o(1)}} \ll \frac{X}{(\log X)^{1-\varepsilon}}.$$

This completes the proof of the bound for  $E_1$  in (6.2.7).

**6.4. Regime II.** It remains to prove the bound for  $E_2$  in (6.2.7). From (6.2.5) and (6.2.6) 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.

Comparing with (6.2.5), we see that we may write  $E_2$  as

$$\begin{aligned} & \frac{1}{2\pi i} \int \frac{1}{\log X} \frac{e^{i(\log X)^2 s}}{e^{-i(\log X)^2 s}} K(s) \sum_{\substack{g \leq M \\ g \text{ odd}}} \frac{\mu^2(g)}{g^{1+s}} \sum_{\substack{m \leq M \\ (m, 2g)=1}} \frac{b_{mg}}{m^{1/2}} \\ & \times \sum_{\substack{r \leq X^{1/2+\varepsilon} \\ (r, 2g)=1 \\ (m, r)=1 \\ mr \gg \exp(\varpi \sqrt{\log X})}} \frac{\mu^2(r)}{r^{1/2+s}} \sum_{\gamma \in \{\pm 1, \pm 2\}} \sum_n \Lambda(n) \Phi_s\left(\frac{n}{X}\right) \chi_{\gamma mr}(n) ds, \end{aligned}$$

where  $K(s)$  satisfies  $|K(s)| \ll (\log X)^{O(1)}$ . We apply the triangle inequality and take a supremum in  $s$  to see that, for some complex number  $s_0$  satisfying  $\operatorname{Re}(s_0) = \frac{1}{\log X}$ ,  $|\operatorname{Im}(s_0)| \leq (\log X)^2$  we have

$$\begin{aligned} E_2 & \ll (\log X)^{O(1)} \sum_{\substack{g \leq M \\ g \text{ odd}}} \frac{\mu^2(g)}{g} \sum_{\substack{m \leq M \\ (m, 2g)=1}} \frac{\mu^2(m)}{m^{1/2}} \\ & \times \sum_{\substack{r \leq X^{1/2+\varepsilon} \\ (r, 2g)=1 \\ (m, r)=1 \\ mr \gg \exp(\varpi \sqrt{\log X})}} \frac{\mu^2(r)}{r^{1/2}} \sum_{\gamma \in \{\pm 1, \pm 2\}} \left| \sum_n \Lambda(n) \Phi_{s_0}\left(\frac{n}{X}\right) \chi_{\gamma mr}(n) \right|. \end{aligned}$$

Summing over  $g$  then contributes an additional factor of  $\log X$ . We write  $q = \gamma mr$  and use the divisor bound to control the multiplicity of representations of  $q$ . After breaking the

range of  $q$  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

$$\mathcal{E}(Q) := Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \left| \sum_n \Lambda(n) \Phi_{s_0} \left( \frac{n}{X} \right) \chi(n) \right|.$$

In order to prove (6.2.7) it therefore suffices to show, for some absolute constant  $c_8 > 0$ , that

$$(6.4.1) \quad \mathcal{E}(Q) \ll X \exp(-c_8 \varpi \sqrt{\log X})$$

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  $\Phi_{s_0}$  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

$$(6.4.2) \quad \left| \Phi^\dagger(\sigma + it + \frac{s_0}{2}) \right| \ll_{\sigma,j} (\log X)^j \left( 1 + \left| t - \frac{\text{Im}(s_0)}{2} \right| \right)^{-j}$$

for every non-negative integer  $j$ .

We shift the line of integration to  $\text{Re}(w) = -\frac{1}{2}$ , leaving residues from all of the zeros of  $L(w, \chi)$  in the critical strip. We bound the new integral by applying the estimate

$$\left| \frac{L'}{L}(w, \chi) \right| \ll \log(q|w|),$$

valid for  $\text{Re}(w) = -\frac{1}{2}$ , and deduce that

$$\sum_n \Lambda(n) \Phi_{s_0} \left( \frac{n}{X} \right) \chi(n) = - \sum_{\substack{L(\rho, \chi)=0 \\ 0 \leq \beta \leq 1}} X^\rho \Phi^\dagger \left( \rho + \frac{s_0}{2} \right) + O \left( \frac{(\log X)^{O(1)}}{X^{1/2}} \right).$$

We have written here  $\rho = \beta + i\gamma$ . The error term is, of course, completely acceptable for (6.4.1) when summed over  $\chi \in S(Q)$ .

By (6.4.2), the contribution to  $\mathcal{E}(Q)$  from those  $\rho$  with  $|\gamma| > Q^{1/2}$  is

$$\ll XQ^{-100},$$

say, and this gives an acceptable bound. We have therefore obtained

$$(6.4.3) \quad \mathcal{E}(Q) \ll X \exp(-\varpi \sqrt{\log X}) + Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \sum_{\substack{L(\rho, \chi)=0 \\ 0 \leq \beta \leq 1 \\ |\gamma| \leq Q^{1/2}}} X^\beta.$$

In order to bound the right side of (6.4.3), we first need to introduce some notation. For a primitive Dirichlet character  $\chi$  modulo  $q$ , let  $N(T, \chi)$  denote the number of zeros of  $L(s, \chi)$  in the rectangle

$$0 \leq \beta \leq 1, \quad |\gamma| \leq T.$$

For  $T \geq 2$ , say, we have [12, Chapter 16]

$$(6.4.4) \quad N(T, \chi) \ll T \log(qT).$$

For  $\frac{1}{2} \leq \alpha \leq 1$ , define  $N(\alpha, T, \chi)$  to be the number of zeros  $\rho = \beta + i\gamma$  of  $L(s, \chi)$  in the rectangle

$$\alpha \leq \beta \leq 1, \quad |\gamma| \leq T,$$

and define

$$N(\alpha, Q, T) = \sum_{q \leq Q} \sum_{\chi \pmod{q}}^* N(\alpha, T, \chi).$$

In  $N(\alpha, Q, T)$  the summation over  $\chi$  is over primitive characters. We shall employ Jutila's zero-density estimate [21, (1.7)]

$$(6.4.5) \quad N(\alpha, Q, T) \ll (QT)^{4(1-\alpha)+\varepsilon},$$

which holds for  $\alpha \geq \frac{4}{5}$ .

In (6.4.3), we separate the zeros  $\rho$  according to whether  $\beta < \frac{4}{5}$  or  $\beta \geq \frac{4}{5}$ . Using (6.4.4) we deduce

$$(6.4.6) \quad Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \sum_{\substack{L(\rho, \chi)=0 \\ 0 \leq \beta < 4/5 \\ |\gamma| \leq Q^{1/2}}} X^\beta \ll X^{\frac{4}{5}} Q^{1+\varepsilon}.$$

For those zeros with  $\beta \geq \frac{4}{5}$  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  $\leq Q$ . Applying (6.4.4) and (6.4.5), we obtain

$$\begin{aligned} \sum_{\chi \in S(Q)} \sum_{\substack{L(\rho, \chi_q)=0 \\ 4/5 \leq \beta \leq 1 \\ |\gamma| \leq Q^{\frac{1}{2}}}} X^\beta &\ll X^{\frac{4}{5}} Q^{\frac{3}{2}+\varepsilon} + (\log X) \int_{4/5}^1 X^\alpha N(\alpha, Q, Q^{\frac{1}{2}}) d\alpha \\ &\ll X^{\frac{4}{5}} Q^{\frac{3}{2}+\varepsilon} + Q^\varepsilon \int_{4/5}^1 X^\alpha Q^{6(1-\alpha)} d\alpha. \end{aligned}$$

Since  $Q \ll X^{\frac{1}{10}}$ , the integrand of this latter integral is maximized when  $\alpha = 1$ . It follows that

$$(6.4.7) \quad Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \sum_{\substack{L(\rho, \chi_q)=0 \\ 4/5 \leq \beta \leq 1 \\ |\gamma| \leq Q^{\frac{1}{2}}}} X^\beta \ll X^{\frac{4}{5}} Q^{1+\varepsilon} + X Q^{-\frac{1}{2}+\varepsilon} \ll X Q^{-\frac{1}{2}+\varepsilon}.$$

Combining (6.4.7), (6.4.6), and (6.4.3) yields

$$\mathcal{E}(Q) \ll XQ^{-\frac{1}{2}+\varepsilon} + X \exp(-\varpi \sqrt{\log X}),$$

and this suffices for (6.4.1).

**6.5. 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 (3.4)). 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 4.4.

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 [16, 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  $f : \mathbb{N} \rightarrow \mathbb{C}$  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  $f_{>W}(n) = f(n) - f_{\leq W}(n)$ .

We write  $\star$  for Dirichlet convolution. Our starting place is Vaughan's identity [18, Proposition 13.4]. Given a parameter  $V > 1$ , we have

$$(6.5.1) \quad \Lambda(n) = \Lambda_{\leq V}(n) + (\mu_{\leq V} \star \log)(n) - (\mu_{\leq V} \star \Lambda_{\leq V} \star 1)(n) + (\mu_{>V} \star \Lambda_{>V} \star 1)(n).$$

We apply (6.5.1) 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 (6.5.1), but  $\Lambda_{\leq V}(n)$  is identically zero for  $n \asymp X$ .

We have

$$\begin{aligned} \mathcal{E}_1(Q) &:= Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \left| \sum_n (\mu_{\leq V} \star \log)(n) \Phi_{s_0} \left( \frac{n}{X} \right) \chi(n) \right| \\ &\ll Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \sum_{v \leq V} \mu^2(v) \left| \sum_m (\log m) \Phi_{s_0} \left( \frac{mv}{X} \right) \chi(m) \right|. \end{aligned}$$

Let us temporarily define  $f(t) = (\log t) \Phi_{s_0} \left( \frac{tv}{X} \right)$ . Observe that  $f$  is supported on real numbers  $t \asymp X/v$ , and that  $|f'(t)| \ll t^{-1}(\log X)^{O(1)}$ . By partial summation we derive

$$\sum_m (\log m) \Phi_{s_0} \left( \frac{mv}{X} \right) \chi(m) = \int_{t \asymp X/v} f'(t) \sum_{m \leq t} \chi(m) dt \ll (\log X)^{O(1)} \int_{t \asymp X/v} \frac{1}{t} \left| \sum_{m \leq t} \chi(m) \right| dt.$$

The Pólya-Vinogradov inequality [12, Chapter 23] implies that

$$\left| \sum_{m \leq t} \chi(m) \right| \ll Q^{1/2} \log Q,$$

and therefore

$$(6.5.2) \quad \mathcal{E}_1(Q) \ll QVX^\varepsilon \ll X^{\frac{1}{2}+\theta+\frac{1}{3}(\frac{1}{2}-\theta)+\varepsilon} \ll X^{1-\varepsilon};$$

this last inequality holds for  $\varepsilon = \varepsilon(\theta) > 0$  sufficiently small.

We use the same argument to estimate  $\mathcal{E}_2(Q)$ :

$$\begin{aligned}
 \mathcal{E}_2(Q) &:= Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \left| \sum_n (\mu_{\leq V} \star \Lambda_{\leq V} \star 1)(n) \Phi_{s_0} \left( \frac{n}{X} \right) \chi(n) \right| \\
 (6.5.3) \quad &\leq Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \sum_{v \leq V^2} (\mu^2 \star \Lambda)(v) \left| \sum_m \Phi_{s_0} \left( \frac{mv}{X} \right) \chi(m) \right| \\
 &\ll Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \sum_{v \leq V^2} (\mu^2 \star \Lambda)(v) Q^{1/2+\varepsilon} \\
 &\ll Q^{1+\varepsilon} V^2 \ll X^{\frac{1}{2}+\theta+\frac{2}{3}(\frac{1}{2}-\theta)+\varepsilon} \ll X^{1-\varepsilon}.
 \end{aligned}$$

The last sum to estimate is  $\mathcal{E}_3(Q)$ :

$$\begin{aligned}
 \mathcal{E}_3(Q) &:= Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \left| \sum_n (\mu_{>V} \star \Lambda_{>V} \star 1)(n) \Phi_{s_0} \left( \frac{n}{X} \right) \chi(n) \right| \\
 &= Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \left| \sum_{k,\ell} \alpha(k) \beta(\ell) \Phi_{s_0} \left( \frac{k\ell}{X} \right) \chi(k\ell) \right|,
 \end{aligned}$$

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  $\ell$ , so that  $k \asymp K, \ell \asymp L$ , with

$$(6.5.4) \quad V \ll K \ll XV^{-1}, \quad V \ll L \ll XV^{-1}, \quad KL \asymp X.$$

We next separate the variables by Mellin inversion on  $\Phi_{s_0}$ :

$$\mathcal{E}_1(Q) \ll (\log X)^{O(1)} \sup_{K,L} \int_{(0)} \left| \Phi^\dagger \left( w + \frac{s_0}{2} \right) \right| Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \left| \sum_{\substack{k \asymp K \\ \ell \asymp L}} \alpha(k) \beta(\ell) (k\ell)^{-w} \chi(k\ell) \right| |dw|.$$

The integral of  $|\Phi^\dagger|$  has size  $\ll (\log X)^{O(1)}$ , so we obtain

$$\mathcal{E}_3(Q) \ll \sup_{\substack{K,L \\ t \in \mathbb{R}}} Q^{-\frac{1}{2}+\varepsilon} \sum_{\chi \in S(Q)} \left| \sum_{\substack{k \asymp K \\ \ell \asymp L}} \tilde{\alpha}(k) \tilde{\beta}(\ell) \chi(k\ell) \right|,$$

where  $\tilde{\alpha}(k) = \alpha(k)k^{-it}$  and  $\tilde{\beta}(\ell) = \beta(\ell)\ell^{-it}$ . For notational convenience we suppress the dependence of  $\tilde{\alpha}$  and  $\tilde{\beta}$  on  $t$ .

By multiplicativity and Cauchy-Schwarz we obtain

$$\mathcal{E}_3(Q) \ll \sup_{\substack{K,L \\ t \in \mathbb{R}}} Q^{-\frac{1}{2}+\varepsilon} \left( \sum_{\chi \in S(Q)} \left| \sum_{k \asymp K} \tilde{\alpha}(k) \chi(k) \right|^2 \right)^{\frac{1}{2}} \left( \sum_{\chi \in S(Q)} \left| \sum_{\ell \asymp L} \tilde{\beta}(\ell) \chi(\ell) \right|^2 \right)^{\frac{1}{2}}.$$

Applying Lemma 4.4 and recalling (6.5.4) yields

$$\begin{aligned}
 \mathcal{E}_3(Q) &\ll \sup_{K,L} \frac{X^\varepsilon}{Q^{\frac{1}{2}}} ((Q+K)K)^{\frac{1}{2}} ((Q+L)L)^{\frac{1}{2}} \\
 (6.5.5) \quad &\ll \sup_{K,L} X^\varepsilon \left( (KL)^{\frac{1}{2}} Q^{\frac{1}{2}} + \frac{KL}{K^{\frac{1}{2}}} + \frac{KL}{L^{\frac{1}{2}}} + \frac{KL}{Q^{\frac{1}{2}}} \right) \\
 &\ll X^\varepsilon \left( X^{\frac{3}{4}+\frac{\theta}{2}} + \frac{X}{V^{\frac{1}{2}}} + \frac{X}{Q^{\frac{1}{2}}} \right) \ll X^{1-\varepsilon}.
 \end{aligned}$$

The last inequality follows since  $V = X^{\frac{1}{3}(\frac{1}{2}-\theta)}$  and  $Q \gg X^{\frac{1}{10}}$ . Then (6.5.2), (6.5.3), and (6.5.5) imply

$$\mathcal{E}(Q) \ll X^{1-\varepsilon},$$

and this suffices for (6.4.1).

**6.6. Dénouement.** We can extract from our proof of Proposition 6.1 the following result on character sums over primes, which we shall have occasion to use later.

**Lemma 6.1.** *Let  $X$  be a large real number, and let  $\delta > 0$  be small and fixed. Let  $s_0$  be a complex number with  $|\operatorname{Re}(s_0)| \leq \frac{A_1}{\log X}$  and  $|\operatorname{Im}(s_0)| \leq (\log X)^{A_2}$ , for some positive real numbers  $A_1$  and  $A_2$ . Given any positive real numbers  $A_3, A_4$ , and  $B$ , we have*

$$\sum_{\substack{q \leq X^{1-\delta} \\ q \text{ odd} \\ q \neq \square}} \frac{d_2(q)^{A_3} (\log q)^{A_4}}{\sqrt{q}} \left| \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi_{s_0} \left( \frac{p}{X} \right) \left( \frac{q}{p} \right) \right| \ll_{A_1, A_2, A_3, A_4, B, \delta} \frac{X}{(\log X)^B}.$$

*The implied constant is ineffective.*

*Proof.* Follow the proof of (6.2.7), but with a few modifications. Instead of the inequality (6.3.12) use Siegel's theorem [12, Chapter 21]

$$\beta_1 \leq 1 - \frac{c(\varepsilon)}{|q^*|^\varepsilon};$$

the constant  $c(\varepsilon)$  is ineffective if  $\varepsilon \leq \frac{1}{2}$ . Therefore, instead of using the lower bound (6.3.13) we use the lower bound  $|q^*| > (\log X)^D$ , where  $D$  is chosen sufficiently large in terms of  $B$ , and  $\varepsilon$  in Siegel's theorem is chosen sufficiently small in terms of  $D$ .  $\square$

Lemma 6.1 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 6.1 with the upper bound for  $q$  replaced by  $X^{1+\varepsilon}$  for any  $\varepsilon > 0$ .

## 7. 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 (3.1). Our main result for this section is the following (recall (3.4) and (5.2)).

**Proposition 7.1.** *Let  $\delta > 0$  be small and fixed, and let  $\theta, \vartheta$  satisfy  $\theta + 2\vartheta < \frac{1}{2}$ . If  $X \geq X_0(\delta, \theta, \vartheta)$ , then*

$$S_2 \leq \frac{1 + \delta}{2(1 - \frac{1}{\sqrt{2}})^2} \frac{\Im X}{\vartheta^4},$$



where

$$\begin{aligned} \mathfrak{I} = & -2 \int_0^1 H(x)H'(x)dx + \frac{1}{\theta} \int_0^1 H(x)H''(x)dx + \frac{1}{\theta} \int_0^1 H'(x)^2 dx \\ & - \frac{1}{2\theta^2} \int_0^1 H'(x)H''(x)dx + \frac{1}{24\theta^3} \int_0^1 H''(x)^2 dx. \end{aligned}$$

The proof of Proposition 7.1 follows the ideas outlined in Section 3. First, we note that  $\log p \leq \log X$  in (3.1) because  $\Phi$  is supported on  $[\frac{1}{2}, 1]$ . By positivity we may apply the upper bound sieve condition (5.7) to write

$$S_2 \leq (\log X)S^+,$$

where  $S^+$  is defined by

$$(7.1) \quad S^+ = \sum_{n \equiv 1 \pmod{8}} \mu^2(n) \left( \sum_{\substack{d|n \\ d \leq D}} \lambda_d \right) \Phi\left(\frac{n}{X}\right) L\left(\frac{1}{2}, \chi_n\right)^2 M(n)^2.$$

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 (5.8), and so  $\lambda_d = \mu^2(d)\lambda_d$ . We use Lemma 4.2 to write  $L(\frac{1}{2}, \chi_n)^2 = \mathcal{D}_2(n)$ , then insert (3.6) into (7.1) to write

$$(7.2) \quad S^+ = S_N^+ + S_R^+,$$

where

$$(7.3) \quad S_N^+ = \sum_{n \equiv 1 \pmod{8}} N_Y(n) \left( \sum_{\substack{d|n \\ d \leq D}} \mu^2(d)\lambda_d \right) \Phi\left(\frac{n}{X}\right) \mathcal{D}_2(n) M(n)^2$$

and

$$S_R^+ = \sum_{n \equiv 1 \pmod{8}} R_Y(n) \left( \sum_{\substack{d|n \\ d \leq D}} \mu^2(d)\lambda_d \right) \Phi\left(\frac{n}{X}\right) \mathcal{D}_2(n) M(n)^2$$

We first obtain a bound on  $S_R^+$ . The remainder of this section will then be devoted to an analysis of  $S_N^+$ .

**7.1. The contribution of  $S_R^+$ .** In this subsection we show

$$(7.1.1) \quad S_R^+ \ll X^\varepsilon \left( \frac{X}{Y} + X^{1/2} M \right).$$

The arguments here are almost identical to those in [40, 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  $\ell$  and  $h$  are odd and  $h \equiv 1 \pmod{8}$ . By the divisor bound we have

$$|R_Y(n)| \ll n^\varepsilon, \quad \left| \sum_{\substack{d|n \\ d \leq D}} \mu^2(d)\lambda_d \right| \ll n^\varepsilon,$$

and therefore

$$S_R^+ \ll X^\varepsilon \sum_{\substack{Y < \ell \leq \sqrt{X} \\ 2 \nmid \ell}} \sum_{\substack{X/2\ell^2 < h \leq X/\ell^2 \\ h \equiv 1 \pmod{8}}} \mu^2(h) |M(\ell^2 h)^2 \mathcal{D}_2(\ell^2 h)|.$$

There is a mild complication compared to [40] in that it is possible to have  $h = 1$ , in which case the character  $\chi_h$  is principal.

We apply Cauchy-Schwarz and obtain

(7.1.2)

$$S_R^+ \ll X^\varepsilon \sum_{\substack{Y < \ell \leq \sqrt{X} \\ 2 \nmid \ell}} \left( \sum_{\substack{X/2\ell^2 < h \leq X/\ell^2 \\ h \equiv 1 \pmod{8}}} \mu^2(h) |M(\ell^2 h)^2|^2 \right)^{1/2} \left( \sum_{\substack{X/2\ell^2 < h \leq X/\ell^2 \\ h \equiv 1 \pmod{8}}} \mu^2(h) |\mathcal{D}_2(\ell^2 h)|^2 \right)^{1/2}.$$

We have

$$M(\ell^2 h)^2 = \sum_{\substack{m \leq M^2 \\ (m, 2\ell) = 1}} \frac{\alpha(m)}{\sqrt{m}} \left( \frac{h}{m} \right)$$

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 4.4 as in [40, p. 460]. We therefore have

$$(7.1.3) \quad \sum_{\substack{X/2\ell^2 < h \leq X/\ell^2 \\ h \equiv 1 \pmod{8}}} \mu^2(h) |M(\ell^2 h)^2|^2 \ll X^\varepsilon \left( \frac{X}{\ell^2} + M^2 \right).$$

Now, by the definition (4.1) of  $\omega_2(\xi)$ , and the definition of  $\mathcal{D}_2$  in Lemma 4.2, for any  $c > \frac{1}{2}$ ,

$$\mathcal{D}_2(\ell^2 h) = \frac{2}{(1 - \frac{1}{\sqrt{2}})^4} \frac{1}{2\pi i} \int_{(c)} \frac{\Gamma^2(\frac{s}{2} + \frac{1}{4})}{\Gamma^2(\frac{1}{4})} \left( 1 - \frac{1}{2^{1/2-s}} \right)^2 \left( \frac{\ell^2 h}{\pi} \right)^s L^2\left(\frac{1}{2} + s, \chi_h\right) \mathcal{E}(s, 2\ell) \frac{ds}{s},$$

where

$$\mathcal{E}(s, k) = \prod_{p|k} \left( 1 - \frac{\chi_h(p)}{p^{1/2+s}} \right)^2.$$

If  $h = 1$  then  $L^2(\frac{1}{2} + s, \chi_h) = \zeta^2(\frac{1}{2} + s)$ . 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 4.5, we obtain an analogue of [40, (3.5)],

$$(7.1.4) \quad \sum_{\substack{X/2\ell^2 < h \leq X/\ell^2 \\ h \equiv 1 \pmod{8}}} \mu^2(h) |\mathcal{D}_2(\ell^2 h)|^2 \ll \frac{X^{1+\varepsilon}}{\ell^2}.$$

Combining (7.1.2), (7.1.3), and (7.1.4) yields (7.1.1).

**7.2. Poisson summation.** We begin our evaluation of  $S_N^+$  by inserting into (7.3) the definition (3.3) of the mollifier  $M(n)$ . We then use the definition of  $\mathcal{D}_2$  (see Lemma 4.2) to write

$$(7.2.1) \quad S_N^+ = \frac{8}{(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ m_1, m_2 \text{ odd}}} \frac{b_{m_1} b_{m_2}}{\sqrt{m_1 m_2}} \sum_{\substack{n \equiv 1 \pmod{8} \\ d|n}} N_Y(n) \Phi\left(\frac{n}{X}\right) \\ \times \sum_{\substack{\nu=1 \\ \nu \text{ odd}}}^{\infty} \frac{d_2(\nu)}{\sqrt{\nu}} \omega_2\left(\frac{\nu\pi}{n}\right) \left(\frac{n}{m_1 m_2 \nu}\right).$$

We next apply Poisson summation to evaluate the  $n$ -sum. Denote the  $n$ -sum in (7.2.1) by  $Z$ , i.e. define  $Z$  by

$$(7.2.2) \quad Z = Z(d, \nu, m_1 m_2; X, Y) = \sum_{\substack{n \equiv 1 \pmod{8} \\ d|n}} N_Y(n) \Phi\left(\frac{n}{X}\right) \omega_2\left(\frac{\nu\pi}{n}\right) \left(\frac{n}{m_1 m_2 \nu}\right).$$

We insert the definition (3.7) of  $N_Y(n)$  and interchange the order of summation to write  $Z$  as

$$(7.2.3) \quad Z = \sum_{\substack{\alpha \leq Y \\ \alpha \text{ odd}}} \mu(\alpha) \sum_{\substack{n \equiv 1 \pmod{8} \\ [\alpha^2, d]|n}} F_{\nu}\left(\frac{n}{X}\right) \left(\frac{n}{m_1 m_2 \nu}\right),$$

where  $F_{\nu}(t)$  is defined by

$$(7.2.4) \quad F_{\nu}(t) = \Phi(t) \omega_2\left(\frac{\nu\pi}{tX}\right).$$

If  $\alpha$  and  $d$  are square-free, then  $[\alpha^2, d] = \alpha^2 d_1$ , where

$$(7.2.5) \quad d_1 = \frac{d}{(d, \alpha)}.$$

We may thus relabel  $n$  as  $\alpha^2 d_1 m$  in (7.2.3), and then split the resulting  $m$ -sum according to the congruence class of  $m \pmod{m_1 m_2 \nu}$ . We deduce from (7.2.3) that

$$Z = \sum_{\substack{\alpha \leq Y \\ (\alpha, 2m_1 m_2 \nu)=1}} \mu(\alpha) \left(\frac{d_1}{m_1 m_2 \nu}\right) \sum_{b \pmod{m_1 m_2 \nu}} \left(\frac{b}{m_1 m_2 \nu}\right) \sum_{\substack{m \equiv \alpha^2 d_1 \pmod{8} \\ m \equiv b \pmod{m_1 m_2 \nu}}} F_{\nu}\left(\frac{\alpha^2 d_1 m}{X}\right).$$

By the Chinese Remainder Theorem, we may write the congruence conditions on  $m$  as a single condition  $m \equiv \gamma \pmod{8m_1 m_2 \nu}$  for some integer  $\gamma$  depending on  $\alpha, d, b$ . Thus, we may relabel  $m$  as  $8jm_1 m_2 \nu + \gamma$ , where  $j$  ranges over all integers, and arrive at

$$(7.2.6) \quad Z = \sum_{\substack{\alpha \leq Y \\ (\alpha, 2m_1 m_2 \nu)=1}} \mu(\alpha) \left(\frac{d_1}{m_1 m_2 \nu}\right) \sum_{b \pmod{m_1 m_2 \nu}} \left(\frac{b}{m_1 m_2 \nu}\right) \sum_{j \in \mathbb{Z}} F_{\nu}\left(\frac{\alpha^2 d_1 (8jm_1 m_2 \nu + \gamma)}{X}\right).$$

We apply Poisson summation to the  $j$ -sum to write

$$\sum_{j \in \mathbb{Z}} F_{\nu}\left(\frac{\alpha^2 d_1 (8jm_1 m_2 \nu + \gamma)}{X}\right) = \frac{X}{8\alpha^2 d_1 m_1 m_2 \nu} \sum_{k \in \mathbb{Z}} e\left(\frac{k\gamma}{8m_1 m_2 \nu}\right) \hat{F}_{\nu}\left(\frac{kX}{8\alpha^2 d_1 m_1 m_2 \nu}\right).$$

We insert this into (7.2.6) and apply the reciprocity relation

$$e\left(\frac{k\gamma}{8m_1m_2\nu}\right) = e\left(\frac{k\bar{8}b}{m_1m_2\nu}\right) e\left(\frac{k\overline{\alpha^2d_1m_1m_2\nu}}{8}\right),$$

where  $\bar{8}$  is the inverse of 8 modulo  $m_1m_2\nu$  and  $\overline{\alpha^2d_1m_1m_2\nu}$  is the inverse of  $\alpha^2d_1m_1m_2\nu$  modulo 8. We then evaluate the  $b$ -sum using the definition (4.3) of the Gauss sum. The result is

$$\begin{aligned} Z &= \frac{X}{8m_1m_2\nu} \sum_{\substack{\alpha \leq Y \\ (\alpha, 2m_1m_2\nu)=1}} \frac{\mu(\alpha)}{\alpha^2d_1} \left(\frac{2d_1}{m_1m_2\nu}\right) \\ &\quad \times \sum_{k \in \mathbb{Z}} e\left(\frac{k\overline{\alpha^2d_1m_1m_2\nu}}{8}\right) \hat{F}_\nu\left(\frac{kX}{8\alpha^2d_1m_1m_2\nu}\right) \tau_k(m_1m_2\nu). \end{aligned}$$

Recalling (7.2.1) and (7.2.2), we arrive at

$$\begin{aligned} (7.2.7) \quad S_N^+ &= \frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1m_2, 2d)=1}} \frac{b_{m_1}b_{m_2}}{(m_1m_2)^{3/2}} \sum_{\substack{\nu=1 \\ (\nu, 2d)=1}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2}} \sum_{\substack{\alpha \leq Y \\ (\alpha, 2m_1m_2\nu)=1}} \frac{\mu(\alpha)}{\alpha^2d_1} \\ &\quad \times \left(\frac{2d_1}{m_1m_2\nu}\right) \sum_{k \in \mathbb{Z}} e\left(\frac{k\overline{\alpha^2d_1m_1m_2\nu}}{8}\right) \hat{F}_\nu\left(\frac{kX}{8\alpha^2d_1m_1m_2\nu}\right) \tau_k(m_1m_2\nu). \end{aligned}$$

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 (7.2.7) as

$$(7.2.8) \quad S_N^+ = \mathcal{T}_0 + \mathcal{B},$$

where  $\mathcal{T}_0$  is the contribution from  $k = 0$  in (7.2.7), while  $\mathcal{B}$  is the contribution from  $k \neq 0$  in (7.2.7). We evaluate  $\mathcal{T}_0$  in the next subsection, and  $\mathcal{B}$  in later subsections.

**7.3. The contribution from  $k = 0$ .** By (4.3),  $\tau_0(n) = \varphi(n)$  if  $n$  is a perfect square, and  $\tau_0(n) = 0$  otherwise. Hence the term  $\mathcal{T}_0$  in (7.2.7) is

$$\begin{aligned} (7.3.1) \quad \mathcal{T}_0 &= \frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1m_2, 2d)=1}} \frac{b_{m_1}b_{m_2}}{(m_1m_2)^{3/2}} \sum_{\substack{\nu=1 \\ (\nu, 2d)=1 \\ m_1m_2\nu=\square}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2}} \sum_{\substack{\alpha \leq Y \\ (\alpha, 2m_1m_2\nu)=1}} \frac{\mu(\alpha)}{\alpha^2d_1} \\ &\quad \times \hat{F}_\nu(0) \varphi(m_1m_2\nu). \end{aligned}$$

We first extend the sum over  $\alpha$  to infinity. Since  $\varphi(n) \leq n$ , the error introduced in doing so is

$$(7.3.2) \quad \ll X \sum_{d \leq D} |\lambda_d| \sum_{m_1, m_2 \leq M} \frac{|b_{m_1}b_{m_2}|}{\sqrt{m_1m_2}} \sum_{\substack{\nu=1 \\ m_1m_2\nu=\square}}^{\infty} \frac{d_2(\nu)}{\sqrt{\nu}} \sum_{\alpha > Y} \frac{1}{\alpha^2d_1} |\hat{F}_\nu(0)|.$$

By Lemma 4.1,  $\hat{F}_\nu(0) \ll 1$  uniformly for all  $\nu > 0$ , and  $\hat{F}_\nu(0) \ll \exp(-\frac{\pi\nu}{8X})$  for  $\nu > X^{1+\varepsilon}$ . Moreover, (5.8) implies that  $|\lambda_d| \ll d^\varepsilon$ , while  $|b_m| \ll 1$  by (3.5). It follows from these bounds

that (7.3.2) is

$$(7.3.3) \quad \ll X^{1+\varepsilon} \sum_{d \leq D} \sum_{m_1, m_2 \leq M} \frac{1}{\sqrt{m_1 m_2}} \sum_{\substack{\nu \leq X^{1+\varepsilon} \\ m_1 m_2 \nu = \square}} \frac{1}{\sqrt{\nu}} \sum_{\alpha > Y} \frac{1}{\alpha^2 d_1} + \exp(-X^\varepsilon).$$

Since  $m_1 m_2 \nu$  is a perfect square, the sum over  $m_1, m_2, \nu$  in (7.3.3) is  $\ll X^\varepsilon$ . Also, the definition (7.2.5) 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 (7.3.3) is  $O(X^{1+\varepsilon}/Y)$ . This bounds the error in extending the sum over  $\alpha$  in (7.3.1) to infinity, and we arrive at

$$\begin{aligned} \mathcal{T}_0 = \frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} \sum_{\substack{\nu=1 \\ (\nu, 2d)=1 \\ m_1 m_2 \nu = \square}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2}} \sum_{\substack{\alpha=1 \\ (\alpha, 2m_1 m_2 \nu)=1}}^{\infty} \frac{\mu(\alpha)}{\alpha^2 d_1} \\ \times \hat{F}_\nu(0) \varphi(m_1 m_2 \nu) + O\left(\frac{X^{1+\varepsilon}}{Y}\right). \end{aligned}$$

Writing the  $\alpha$ -sum as an Euler product, we deduce that

$$(7.3.4) \quad \begin{aligned} \mathcal{T}_0 = \frac{4X}{3(\sqrt{2}-1)^4 \zeta(2)} \sum_{\substack{d \leq D \\ d \text{ odd}}} \frac{\mu^2(d) \lambda_d}{d} \prod_{p|d} \left(\frac{p}{p+1}\right) \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{\sqrt{m_1 m_2}} \\ \times \sum_{\substack{\nu=1 \\ (\nu, 2d)=1 \\ m_1 m_2 \nu = \square}}^{\infty} \frac{d_2(\nu)}{\sqrt{\nu}} \hat{F}_\nu(0) \prod_{p|m_1 m_2 \nu} \left(\frac{p}{p+1}\right) + O\left(\frac{X^{1+\varepsilon}}{Y}\right). \end{aligned}$$

We next evaluate the sum over  $d$ . Lemma 5.3 implies

$$(7.3.5) \quad \begin{aligned} \sum_{\substack{d \leq D \\ (d, 2m_1 m_2 \nu)=1}} \frac{\mu^2(d) \lambda_d}{d} \prod_{p|d} \left(\frac{p}{p+1}\right) = \frac{1 + E_0(X)}{\log R} \prod_{\substack{p|2m_1 m_2 \nu \\ p \leq z_0}} \left(1 + \frac{1}{p}\right) \prod_{p \leq z_0} \left(\frac{p^2}{p^2-1}\right) \\ + O_A((\log R)^{-A}). \end{aligned}$$

Recall that  $E_0(X) \rightarrow 0$ , and depends only on  $X, G$ , and  $\vartheta$ . Heretofore we just write  $o(1)$  instead of  $E_0(X)$ .

We may omit the condition  $p \leq z_0$  by trivial estimation and (5.1). It follows from (7.3.5) and (7.3.4) that

$$(7.3.6) \quad \begin{aligned} \mathcal{T}_0 = \frac{2X}{(\sqrt{2}-1)^4} \frac{1 + o(1)}{\log R} \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2)=1}} \frac{b_{m_1} b_{m_2}}{\sqrt{m_1 m_2}} \sum_{\substack{\nu=1 \\ (\nu, 2)=1 \\ m_1 m_2 \nu = \square}}^{\infty} \frac{d_2(\nu)}{\sqrt{\nu}} \hat{F}_\nu(0) \\ + O_A\left(\frac{X}{(\log R)^A}\right) + O\left(\frac{X^{1+\varepsilon}}{Y}\right). \end{aligned}$$

The next task is to carry out the summation over  $m_1, m_2$ , and  $\nu$ . Let  $\Upsilon_0$  be defined by

$$(7.3.7) \quad \Upsilon_0 = \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2)=1}} \sum_{\substack{\nu=1 \\ (\nu, 2)=1 \\ m_1 m_2 \nu = \square}} \frac{b_{m_1} b_{m_2}}{\sqrt{m_1 m_2}} \sum_{\nu=1}^{\infty} \frac{d_2(\nu)}{\sqrt{\nu}} \hat{F}_\nu(0).$$

We insert into (7.3.7) the definition (3.5) of  $b_m$  and the definitions (7.2.4) and (4.1) of  $F_\nu$  and  $\omega_2$ , and then apply the Fourier inversion formula (6.1.2). After interchanging the order of summation, we arrive at

$$(7.3.8) \quad \Upsilon_0 = \frac{1}{2\pi i} \int_{(c)} \frac{\Gamma(\frac{s}{2} + \frac{1}{4})^2}{\Gamma(\frac{1}{4})^2} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)^2 \left(\frac{X}{\pi}\right)^s \check{\Phi}(s) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(z_1) h(z_2) \\ \times \sum_{\substack{(m_1 m_2 \nu, 2)=1 \\ m_1 m_2 \nu = \square}} \frac{\mu(m_1) \mu(m_2) d_2(\nu)}{(m_1 m_2 \nu)^{\frac{1}{2}} m_1^{\frac{1+iz_1}{\log M}} m_2^{\frac{1+iz_2}{\log M}} \nu^s} dz_1 dz_2 \frac{ds}{s},$$

where we take  $c = \frac{1}{\log X}$  to facilitate later estimations. We may write the sum over  $m_1, m_2, \nu$  as an Euler product

$$(7.3.9) \quad \sum_{\substack{(m_1 m_2 \nu, 2)=1 \\ m_1 m_2 \nu = \square}} \frac{\mu(m_1) \mu(m_2) d_2(\nu)}{(m_1 m_2 \nu)^{\frac{1}{2}} m_1^{\frac{1+iz_1}{\log M}} m_2^{\frac{1+iz_2}{\log M}} \nu^s} = \prod_{p>2} \sum_{\substack{m_1=0 \\ m_1+m_2+\nu \text{ even}}}^1 \sum_{\substack{m_2=0 \\ m_1+m_2+\nu \text{ even}}}^1 \sum_{\nu=0}^{\infty} \frac{(-1)^{m_1+m_2} (\nu+1)}{p^{\frac{m_1+m_2+\nu}{2} + m_1 \left(\frac{1+iz_1}{\log M}\right) + m_2 \left(\frac{1+iz_2}{\log M}\right) + \nu s}}.$$

In the latter sum, the terms with  $(m_1, m_2, \nu) \neq (0, 0, 0), (0, 0, 2), (0, 1, 1), (1, 0, 1),$  or  $(1, 1, 0)$  add up to  $O(p^{-2+\varepsilon})$  whenever  $z_1, z_2, s$  are complex numbers such that  $\frac{1+iz_1}{\log M}, \frac{1+iz_2}{\log M}$ , and  $s$  each have real part  $\geq -\varepsilon$ . Hence, for such  $z_1, z_2, s$ , (7.3.9) equals

$$\prod_{p>2} \left(1 + \frac{3}{p^{1+2s}} - \frac{2}{p^{1+\frac{1+iz_1}{\log M}+s}} - \frac{2}{p^{1+\frac{1+iz_2}{\log M}+s}} + \frac{1}{p^{1+\frac{1+iz_1}{\log M}+\frac{1+iz_2}{\log M}}} + O\left(\frac{1}{p^{2-\varepsilon}}\right)\right).$$

Thus (7.3.9) equals

$$(7.3.10) \quad \zeta^3(1+2s) \zeta\left(1 + \frac{2+iz_1+iz_2}{\log M}\right) \zeta^{-2}\left(1 + \frac{1+iz_1}{\log M} + s\right) \zeta^{-2}\left(1 + \frac{1+iz_2}{\log M} + s\right) Q\left(\frac{1+iz_1}{\log M}, \frac{1+iz_2}{\log M}, s\right),$$

where  $Q(w_1, w_2, s)$  is an Euler product that is uniformly bounded and holomorphic when each of  $\text{Re}(w_1), \text{Re}(w_2)$ , and  $\text{Re}(s)$  is  $\geq -\varepsilon$ . From this definition of  $Q$  and a calculation, we see that

$$(7.3.11) \quad Q(0, 0, 0) = 1,$$

a fact we use shortly. We insert the expression (7.3.10) for the  $m_1, m_2, \nu$ -sum into (7.3.8) and arrive at

$$\Upsilon_0 = \frac{1}{2\pi i} \int_{(c)} \frac{\Gamma(\frac{s}{2} + \frac{1}{4})^2}{\Gamma(\frac{1}{4})^2} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)^2 \left(\frac{X}{\pi}\right)^s \check{\Phi}(s) \zeta^3(1+2s) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(z_1) h(z_2) \\ \times \zeta\left(1 + \frac{2+iz_1+iz_2}{\log M}\right) \zeta^{-2}\left(1 + \frac{1+iz_1}{\log M} + s\right) \zeta^{-2}\left(1 + \frac{1+iz_2}{\log M} + s\right) Q\left(\frac{1+iz_1}{\log M}, \frac{1+iz_2}{\log M}, s\right) dz_1 dz_2 \frac{ds}{s}.$$

By (6.1.4) and the rapid decay of the gamma function, we may truncate the integrals to the region  $|z_1|, |z_2| \leq \sqrt{\log M}$  and  $|\text{Im}(s)| \leq (\log X)^2$ , 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 (6.3.10) holds on  $L_1, L_2$ , and  $L_3$ . 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  $\Gamma$  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

(7.3.12)

$$\begin{aligned} \Upsilon_0 = & \frac{1}{2\pi i} \oint_{|s|=\frac{1}{\log X}} \frac{\Gamma(\frac{s}{2} + \frac{1}{4})^2}{\Gamma(\frac{1}{4})^2} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)^2 \left(\frac{X}{\pi}\right)^s \check{\Phi}(s) \zeta^3(1+2s) \\ & \times \int \int_{|z_i| \leq \sqrt{\log M}} h(z_1) h(z_2) \zeta\left(1 + \frac{2+iz_1+iz_2}{\log M}\right) \zeta^{-2}\left(1 + \frac{1+iz_1}{\log M} + s\right) \zeta^{-2}\left(1 + \frac{1+iz_2}{\log M} + s\right) \\ & \times Q\left(\frac{1+iz_1}{\log M}, \frac{1+iz_2}{\log M}, s\right) dz_1 dz_2 \frac{ds}{s} + O_A\left(\frac{1}{(\log X)^A}\right). \end{aligned}$$

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 (7.3.11) that

$$\begin{aligned} \Upsilon_0 = & \frac{1}{16\pi i} \oint_{|s|=\frac{1}{\log X}} \frac{\Gamma(\frac{s}{2} + \frac{1}{4})^2}{\Gamma(\frac{1}{4})^2} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)^2 \left(\frac{X}{\pi}\right)^s \check{\Phi}(s) \int \int_{|z_i| \leq \sqrt{\log M}} h(z_1) h(z_2) \\ & \times \left(\frac{\log M}{2 + iz_1 + iz_2}\right) \left(\frac{1 + iz_1}{\log M} + s\right)^2 \left(\frac{1 + iz_2}{\log M} + s\right)^2 dz_1 dz_2 \frac{ds}{s^4} + O\left(\frac{1}{\log X}\right). \end{aligned}$$

By (6.1.4), we may extend the integrals over  $z_1, z_2$  to  $\mathbb{R}^2$ , introducing a negligible error. We then apply the formula

$$\begin{aligned} (7.3.13) \quad & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(z_1) h(z_2) \frac{(1 + iz_1)^j (1 + iz_2)^k}{2 + iz_1 + iz_2} dz_1 dz_2 \\ & = \int_0^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(z_1) h(z_2) (1 + iz_1)^j (1 + iz_2)^k e^{-t(1+iz_1)-t(1+iz_2)} dz_1 dz_2 dt \\ & = (-1)^{j+k} \int_0^{\infty} H^{(j)}(t) H^{(k)}(t) dt, \end{aligned}$$

which follows from repeated differentiation of (6.1.2), to obtain

$$\begin{aligned} \Upsilon_0 = & \frac{1}{16\pi i} \oint_{|s|=\frac{1}{\log X}} \frac{\Gamma(\frac{s}{2} + \frac{1}{4})^2}{\Gamma(\frac{1}{4})^2} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)^2 \left(\frac{X}{\pi}\right)^s \check{\Phi}(s) \left\{ \frac{1}{(\log M)^3} \int_0^1 H''(t)^2 dt \right. \\ & - \frac{4s}{(\log M)^2} \int_0^1 H'(t) H''(t) dt + \frac{2s^2}{\log M} \int_0^1 H(t) H''(t) dt + \frac{4s^2}{\log M} \int_0^1 H'(t)^2 dt \\ & \left. - 4s^3 \int_0^1 H(t) H'(t) dt + s^4 \log M \int_0^1 H(t)^2 dt \right\} \frac{ds}{s^4} + O\left(\frac{1}{\log X}\right). \end{aligned}$$

We evaluate the  $s$ -integral as a residue using (6.1.6). The result is

$$\begin{aligned} \Upsilon_0 = & \frac{\check{\Phi}(0)}{8} \left(1 - \frac{1}{\sqrt{2}}\right)^2 \left\{ \frac{1}{6} \left(\frac{\log X}{\log M}\right)^3 \int_0^1 H''(t)^2 dt - 2 \left(\frac{\log X}{\log M}\right)^2 \int_0^1 H'(t)H''(t) dt \right. \\ & \left. + 2 \frac{\log X}{\log M} \int_0^1 H(t)H''(t) dt + 4 \frac{\log X}{\log M} \int_0^1 H'(t)^2 dt - 4 \int_0^1 H(t)H'(t) dt \right\} + O\left(\frac{1}{\log X}\right). \end{aligned}$$

From this, (2.2), (7.3.6), and the definition (7.3.7) of  $\Upsilon_0$ , we arrive at

(7.3.14)

$$\begin{aligned} \mathcal{T}_0 = & \frac{X}{8 \left(1 - \frac{1}{\sqrt{2}}\right)^2} \frac{1 + o(1)}{\log R} \left\{ \frac{1}{24} \left(\frac{\log X}{\log M}\right)^3 \int_0^1 H''(t)^2 dt \right. \\ & - \frac{1}{2} \left(\frac{\log X}{\log M}\right)^2 \int_0^1 H'(t)H''(t) dt + \frac{\log X}{2 \log M} \int_0^1 H(t)H''(t) dt + \frac{\log X}{\log M} \int_0^1 H'(t)^2 dt \\ & \left. - \int_0^1 H(t)H'(t) dt \right\} + O\left(\frac{X}{(\log X)^2} + \frac{X^{1+\varepsilon}}{Y}\right). \end{aligned}$$

**7.4. The contribution from  $k \neq 0$ : splitting into cases.** Having estimated the term  $\mathcal{T}_0$  in (7.2.8), 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  $e(k\alpha^2 d_1 m_1 m_2 \nu / 8)$  in (7.2.7) 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 (7.2.8) according to these four cases. For the terms with odd  $k$ , we use the identity

$$e\left(\frac{h}{8}\right) = \frac{\sqrt{2}}{2} \left(\frac{2}{h}\right) + \frac{\sqrt{2}}{2} \left(\frac{-2}{h}\right) i, \quad h \text{ odd},$$

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 (4.3) 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  $e(k\alpha^2 d_1 m_1 m_2 \nu / 8)$  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, (7.2.7), and (7.2.8), we write

$$\begin{aligned} (7.4.1) \quad \mathcal{B} = & \frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} \sum_{\substack{\nu=1 \\ (\nu, 2d)=1}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2}} \sum_{\substack{\alpha \leq Y \\ (\alpha, 2m_1 m_2 \nu)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \\ & \times (\mathcal{Q}_1 + \mathcal{Q}_2 + \mathcal{Q}_3 + \mathcal{Q}_4 + \mathcal{U}_1 + \mathcal{U}_2 + \mathcal{V} + \mathcal{W}), \end{aligned}$$

where

$$(7.4.2) \quad \mathcal{Q}_1 = \left(\frac{1+i}{2}\right) \frac{\sqrt{2}}{2} \left(\frac{2d_1}{m_1 m_2 \nu}\right) \sum_{\substack{k \in \mathbb{Z} \\ k \text{ odd}}} \left(\frac{2}{k d_1 m_1 m_2 \nu}\right) \hat{F}_\nu \left(\frac{kX}{8\alpha^2 d_1 m_1 m_2 \nu}\right) G_k(m_1 m_2 \nu),$$



$$(7.4.3) \quad \mathcal{Q}_2 = \left( \frac{1-i}{2} \right) \frac{\sqrt{2}}{2} \left( \frac{-2d_1}{m_1 m_2 \nu} \right) \sum_{\substack{k \in \mathbb{Z} \\ k \text{ odd}}} \left( \frac{2}{k d_1 m_1 m_2 \nu} \right) \hat{F}_\nu \left( \frac{kX}{8\alpha^2 d_1 m_1 m_2 \nu} \right) G_k(m_1 m_2 \nu),$$

$$(7.4.4) \quad \mathcal{Q}_3 = \left( \frac{1+i}{2} \right) i \frac{\sqrt{2}}{2} \left( \frac{2d_1}{m_1 m_2 \nu} \right) \sum_{\substack{k \in \mathbb{Z} \\ k \text{ odd}}} \left( \frac{-2}{k d_1 m_1 m_2 \nu} \right) \hat{F}_\nu \left( \frac{kX}{8\alpha^2 d_1 m_1 m_2 \nu} \right) G_k(m_1 m_2 \nu),$$

$$(7.4.5) \quad \mathcal{Q}_4 = \left( \frac{1-i}{2} \right) i \frac{\sqrt{2}}{2} \left( \frac{-2d_1}{m_1 m_2 \nu} \right) \sum_{\substack{k \in \mathbb{Z} \\ k \text{ odd}}} \left( \frac{-2}{k d_1 m_1 m_2 \nu} \right) \hat{F}_\nu \left( \frac{kX}{8\alpha^2 d_1 m_1 m_2 \nu} \right) G_k(m_1 m_2 \nu),$$

$$(7.4.6) \quad \mathcal{U}_1 = \left( \frac{1+i}{2} \right) \left( \frac{2d_1}{m_1 m_2 \nu} \right) \sum_{\substack{k \in \mathbb{Z} \\ k \equiv 2 \pmod{4}}} e \left( \frac{k\alpha^2 d_1 m_1 m_2 \nu}{8} \right) \hat{F}_\nu \left( \frac{kX}{8\alpha^2 d_1 m_1 m_2 \nu} \right) G_k(m_1 m_2 \nu),$$

$$(7.4.7) \quad \mathcal{U}_2 = \left( \frac{1-i}{2} \right) \left( \frac{-2d_1}{m_1 m_2 \nu} \right) \sum_{\substack{k \in \mathbb{Z} \\ k \equiv 2 \pmod{4}}} e \left( \frac{k\alpha^2 d_1 m_1 m_2 \nu}{8} \right) \hat{F}_\nu \left( \frac{kX}{8\alpha^2 d_1 m_1 m_2 \nu} \right) G_k(m_1 m_2 \nu),$$

$$(7.4.8) \quad \mathcal{V} = \left( \frac{2d_1}{m_1 m_2 \nu} \right) \sum_{\substack{k \in \mathbb{Z} \\ k \equiv 4 \pmod{8}}} e \left( \frac{k\alpha^2 d_1 m_1 m_2 \nu}{8} \right) \hat{F}_\nu \left( \frac{kX}{8\alpha^2 d_1 m_1 m_2 \nu} \right) \tau_k(m_1 m_2 \nu),$$

and

$$(7.4.9) \quad \mathcal{W} = \left( \frac{2d_1}{m_1 m_2 \nu} \right) \sum_{\substack{k \in \mathbb{Z} \\ k \equiv 0 \pmod{8} \\ k \neq 0}} e \left( \frac{k\alpha^2 d_1 m_1 m_2 \nu}{8} \right) \hat{F}_\nu \left( \frac{kX}{8\alpha^2 d_1 m_1 m_2 \nu} \right) \tau_k(m_1 m_2 \nu).$$

**7.5. Evaluation of the sum with  $\mathcal{Q}_1$ .** In this subsection, we evaluate the sum

$$(7.5.1) \quad \mathcal{Q}_1^* := \sum_{\substack{\nu=1 \\ (\nu, 2d)=1}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2}} \sum_{\substack{\alpha \leq Y \\ (\alpha, 2m_1 m_2 \nu)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \mathcal{Q}_1,$$

with  $\mathcal{Q}_1$  defined by (7.4.2). We may cancel the two Jacobi symbols  $\left( \frac{2}{m_1 m_2 \nu} \right)$  in (7.4.2), insert the resulting expression into (7.5.1), and then apply the Mellin inversion formula to the  $\nu$ -sum to deduce that

$$(7.5.2) \quad \begin{aligned} & \mathcal{Q}_1^* = \left( \frac{1+i}{2} \right) \frac{\sqrt{2}}{2} \sum_{\substack{\alpha \leq Y \\ (\alpha, 2m_1 m_2)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \left( \frac{d_1}{m_1 m_2} \right) \sum_{\substack{k \in \mathbb{Z} \\ k \text{ odd}}} \left( \frac{2}{k d_1} \right) \\ & \times \frac{1}{2\pi i} \int_{(c)} \int_0^\infty \hat{F}_t \left( \frac{kX}{8\alpha^2 d_1 m_1 m_2 t} \right) t^{w-1} dt \sum_{\substack{\nu=1 \\ (\nu, 2\alpha d)=1}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2+w}} \left( \frac{d_1}{\nu} \right) G_k(m_1 m_2 \nu) d\nu \end{aligned}$$

for any  $c > 1$ . The interchange in the order of summation is justified by absolute convergence. The next step is to write the  $\nu$ -sum as an Euler product, as follows.

**Lemma 7.1.** *Let  $d_1$  be as defined by (7.2.5). For each nonzero integer  $k$ , define  $k_1$  and  $k_2$  uniquely by the equation*

$$(7.5.3) \quad 4kd_1 = k_1 k_2^2,$$

where  $k_1$  is a fundamental discriminant and  $k_2$  is a positive integer. If  $\ell$  is a positive integer and  $\operatorname{Re}(s) > 1$ , 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

$$\begin{aligned} \mathcal{G}_{0,p}(s; k, \ell, \alpha, d) &= \left( 1 - \frac{1}{p^s} \left( \frac{k_1}{p} \right) \right)^2 \quad \text{if } p|2\alpha d, \quad \text{and} \\ \mathcal{G}_{0,p}(s; k, \ell, \alpha, d) &= \left( 1 - \frac{1}{p^s} \left( \frac{k_1}{p} \right) \right)^2 \sum_{r=0}^{\infty} \frac{r+1}{p^{rs}} \left( \frac{d_1}{p^r} \right) \frac{G_k(p^{r+\operatorname{ord}_p(\ell)})}{p^{r/2}} \quad \text{if } p \nmid 2\alpha d. \end{aligned}$$

The function  $\mathcal{G}_0(s; k, \ell, \alpha, d)$  is holomorphic for  $\operatorname{Re}(s) > \frac{1}{2}$ . Moreover, if  $k_3$  and  $k_4$  are defined by the equation

$$(7.5.4) \quad k = k_3 k_4^2,$$

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  $\operatorname{Re}(s) \geq \frac{1}{2} + \varepsilon$ .

*Proof.* It follows from the definition of  $\mathcal{G}_{0,p}(s; k, \ell, \alpha, d)$  and Lemma 4.3 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  $p \nmid 2\alpha d k \ell$ , since  $\left( \frac{d_1 k}{p} \right) = \left( \frac{k_1}{p} \right)$  for odd primes  $p$ , by (7.5.3). The rest of the proof is similar to that of [40, Lemma 5.3].  $\square$

We also need some analytic properties of the function  $h(\xi, w)$  defined for  $\operatorname{Re}(w) > 0$  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 7.2.** *Let  $F_t$  be defined by (7.2.4). If  $\xi \neq 0$  is a real number, 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

$$(7.5.5) \quad \frac{1}{2\pi i} \int_{(c)} \frac{\Gamma\left(\frac{s}{2} + \frac{1}{4}\right)^2}{\Gamma\left(\frac{1}{4}\right)^2} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)^2 \frac{X^s}{(\pi|\xi|)^s} (2\pi)^{-s+w} \Gamma(s-w) \\ \times \left\{ \cos\left(\frac{\pi}{2}(s-w)\right) - i \operatorname{sgn}(\xi) \sin\left(\frac{\pi}{2}(s-w)\right) \right\} \frac{ds}{s}$$

for any  $c$  with  $\max\{0, \operatorname{Re}(w)\} < c < \operatorname{Re}(w) + 1$ . If  $\xi \neq 0$  is real, then  $h(\xi, w)$  is an entire function of  $w$ . In the region  $-1 < \operatorname{Re}(w) \leq -\delta_0$ , where  $\delta_0 > 0$  is fixed and small, it satisfies the bound

$$h(\xi, w) \ll_{\delta_0} (1 + |w|)^{-\operatorname{Re}(w) - \frac{1}{2}} \exp\left(-\frac{1}{10} \frac{\sqrt{|\xi|}}{\sqrt{X}(|w|+1)}\right) |\xi|^{\operatorname{Re}(w)} |\check{\Phi}(w)|.$$

*Proof.* The proof is similar to that of [40, Lemma 5.2].  $\square$

Now, by these lemmas and the rapid decay of  $\check{\Phi}(w)$  as  $|\operatorname{Im}(w)| \rightarrow \infty$  in a fixed vertical strip, we may move the line of integration of the  $w$ -integral in (7.5.2) to  $\operatorname{Re}(w) = -\frac{1}{2} + \varepsilon$ . This leaves a residue from a pole at  $w = 0$  only when  $\chi_{k_1}$  is a principal character, which holds if and only if  $k_1 = 1$ . By (7.5.3),  $k_1 = 1$  if and only if  $kd_1$  is a perfect square. Hence

$$(7.5.6) \quad \mathcal{Q}_1^* = \mathcal{P}_1 + \mathcal{R}_1,$$

where  $\mathcal{P}_1$  is defined by

$$(7.5.7) \quad \mathcal{P}_1 = \operatorname{Res}_{w=0} \left( \frac{1+i}{2} \right) \frac{\sqrt{2}}{2} \sum_{\substack{\alpha \leq Y \\ (\alpha, 2m_1m_2)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \left( \frac{d_1}{m_1 m_2} \right) \sum_{\substack{k \in \mathbb{Z} \\ k \text{ odd} \\ kd_1 = \square}} h\left(\frac{kX}{8\alpha^2 d_1 m_1 m_2}, w\right) \\ \times \zeta(1+w)^2 \mathcal{G}_0(1+w; k, m_1 m_2, \alpha, d)$$

and  $\mathcal{R}_1$  is defined by

$$(7.5.8) \quad \mathcal{R}_1 = \left( \frac{1+i}{2} \right) \frac{\sqrt{2}}{2} \sum_{\substack{\alpha \leq Y \\ (\alpha, 2m_1m_2)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \left( \frac{d_1}{m_1 m_2} \right) \sum_{\substack{k \in \mathbb{Z} \\ k \text{ odd}}} \left( \frac{2}{kd_1} \right) \\ \times \frac{1}{2\pi i} \int_{(-\frac{1}{2}+\varepsilon)} h\left(\frac{kX}{8\alpha^2 d_1 m_1 m_2}, w\right) L(1+w, \chi_{k_1})^2 \mathcal{G}_0(1+w; k, m_1 m_2, \alpha, d) dw.$$

We bound  $\mathcal{R}_1$  in Subsection 7.6. To estimate  $\mathcal{P}_1$ , observe that  $d_1$  is square-free by its definition (7.2.5) 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 (7.5.7), we may relabel  $k$  as  $d_1 j^2$ , where  $j$  runs through all the odd positive integers. With this and Lemma 7.2, we deduce

from (7.5.7) that

(7.5.9)

$$\begin{aligned} \mathcal{P}_1 = \operatorname{Res}_{w=0} \left( \frac{1+i}{2} \right) \frac{\sqrt{2}}{2} \sum_{\substack{\alpha \leq Y \\ (\alpha, 2m_1m_2)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \zeta(1+w)^2 \check{\Phi}(w) X^w \frac{1}{2\pi i} \int_{(c)} \frac{\Gamma\left(\frac{s}{2} + \frac{1}{4}\right)^2}{\Gamma\left(\frac{1}{4}\right)^2} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)^2 \\ \times \pi^{-s} \Gamma_2(s-w) (8\alpha^2 m_1 m_2)^{s-w} \sum_{\substack{j=1 \\ j \text{ odd}}}^{\infty} j^{-2s+2w} \left( \frac{d_1}{m_1 m_2} \right) \mathcal{G}_0(1+w; d_1 j^2, m_1 m_2, \alpha, d) \frac{ds}{s}, \end{aligned}$$

where  $\Gamma_2(u)$  is defined by

$$(7.5.10) \quad \Gamma_2(u) = (2\pi)^{-u} \Gamma(u) \left( \cos\left(\frac{\pi}{2}u\right) - i \sin\left(\frac{\pi}{2}u\right) \right),$$

and where we take  $c > \frac{1}{2}$  to guarantee the absolute convergence of the  $j$ -sum.

We next write the  $j$ -sum in (7.5.9) as an Euler product. First observe that, by (ii) of Lemma 4.3, if  $j$  is a positive integer then

$$(7.5.11) \quad \left( \frac{d_1}{p^\beta} \right) G_{d_1 j^2}(p^\beta) = G_{j^2}(p^\beta)$$

for all  $p \nmid 2\alpha d$  and  $\beta \geq 1$ . To see this, suppose that  $p^{2n}$  is the largest power of  $p$  dividing  $j^2$ , so that it is also the largest power of  $p$  dividing  $d_1 j^2$ , by (7.2.5). If  $\beta \leq 2n$  is even, then  $\left( \frac{d_1}{p^\beta} \right) = 1$  and both sides of (7.5.11) equal  $\varphi(p^\beta)$ . If  $\beta = 2n + 1$ , then the left-hand side of (7.5.11) equals

$$\left( \frac{d_1}{p^{2n+1}} \right) \left( \frac{d_1 j^2 p^{-2n}}{p} \right) p^{2n} \sqrt{p} = \left( \frac{d_1}{p^{2n}} \right) \left( \frac{d_1}{p} \right)^2 \left( \frac{j^2 p^{-2n}}{p} \right) p^{2n} \sqrt{p} = \left( \frac{j^2 p^{-2n}}{p} \right) p^{2n} \sqrt{p},$$

which equals the right-hand side of (7.5.11). In all other possible cases for  $\beta$ , both sides of (7.5.11) are zero. This proves (7.5.11). From it and the multiplicativity stated in (i) of Lemma 4.3, we deduce that

$$\left( \frac{d_1}{m_1 m_2 \nu} \right) G_{d_1 j^2}(m_1 m_2 \nu) = G_{j^2}(m_1 m_2 \nu)$$

for all positive integers  $\nu$  with  $(\nu, 2\alpha d) = 1$ . It follows from this and Lemma 7.1 that

$$(7.5.12) \quad \left( \frac{d_1}{m_1 m_2} \right) \zeta^2(1+w) \mathcal{G}_0(1+w; d_1 j^2, m_1 m_2, \alpha, d) = \sum_{\substack{\nu=1 \\ (\nu, 2\alpha d)=1}}^{\infty} \frac{d_2(\nu) G_{j^2}(m_1 m_2 \nu)}{\nu^{1+w} \sqrt{\nu}} \\ = \zeta^2(1+w) \mathcal{G}_0(1+w; j^2, m_1 m_2, \alpha d, 1)$$

for  $\operatorname{Re}(w) > 0$ . This and analytic continuation implies

$$(7.5.13) \quad \left( \frac{d_1}{m_1 m_2} \right) \mathcal{G}_0(1+w; d_1 j^2, m_1 m_2, \alpha, d) = \mathcal{G}_0(1+w; j^2, m_1 m_2, \alpha d, 1)$$

for  $\operatorname{Re}(w) > -\frac{1}{2}$ . Now if  $f(j, p)$  is a function such that  $f(j, p) = f(p^{\operatorname{ord}_p(j)}, p)$  for all positive integers  $j$  and all primes  $p$ , then

$$(7.5.14) \quad \sum_{j=1}^{\infty} \prod_p f(j, p) = \prod_p \sum_{b=0}^{\infty} f(p^b, p)$$

if absolute convergence holds for both sides. This can be proved using a standard argument (see, for example, Theorem 11.7 of [3]) together with the fact that  $\prod_{p>y} f(1, p) \rightarrow 1$  as  $y \rightarrow \infty$ . From (7.5.13), the Euler product definition of  $\mathcal{G}_0$  in Lemma 7.1, and (7.5.14) with  $f(1, 2) = 1$  and  $f(2^b, 2) = 0$  for  $b > 0$ , we arrive at

$$\begin{aligned}
 (7.5.15) \quad \sum_{\substack{j=1 \\ j \text{ odd}}}^{\infty} j^{-2s+2w} \left( \frac{d_1}{m_1 m_2} \right) \mathcal{G}_0(1+w; d_1 j^2, m_1 m_2, \alpha, d) &= \sum_{\substack{j=1 \\ j \text{ odd}}}^{\infty} j^{-2s+2w} \mathcal{G}_0(1+w; j^2, m_1 m_2, \alpha d, 1) \\
 &= \left( 1 - \frac{1}{2^{1+w}} \right)^2 \prod_{p>2} \sum_{b=0}^{\infty} p^{2b(w-s)} \mathcal{G}_{0,p}(1+w; p^{2b}, m_1 m_2, \alpha d, 1) \\
 &= \left( 1 - \frac{1}{4^{s-w}} \right) \prod_p \sum_{b=0}^{\infty} p^{2b(w-s)} \mathcal{G}_{0,p}(1+w; p^{2b}, m_1 m_2, \alpha d, 1).
 \end{aligned}$$

Since  $G_k(n) = G_{4k}(n)$  for all odd  $n$ , Lemma 5.3 of [40] and Lemma 7.1 imply that

$$\mathcal{G}_{0,p}(1+w; j^2, m_1 m_2, \alpha d, 1) = \mathcal{G}_p(1+w; j^2, m_1 m_2, \alpha d),$$

where  $\mathcal{G}_p$  is defined in Lemma 5.3 of [40]. Therefore the latter Euler product in (7.5.15) is the same one that has been evaluated in [40, p. 471]. From its evaluation and (7.5.15), we deduce that

$$\begin{aligned}
 (7.5.16) \quad \sum_{\substack{j=1 \\ j \text{ odd}}}^{\infty} j^{-2s+2w} \left( \frac{d_1}{m_1 m_2} \right) \mathcal{G}_0(1+w; d_1 j^2, m_1 m_2, \alpha, d) \\
 = \left( 1 - \frac{1}{4^{s-w}} \right) (m_1 m_2)^{1-s+w} \ell_1^{s-w-\frac{1}{2}} \zeta(2s-2w) \zeta(2s+1) \mathcal{H}_1(s-w, 1+w; m_1 m_2, \alpha d),
 \end{aligned}$$

where  $\ell_1$  is the square-free integer defined by the equation

$$(7.5.17) \quad m_1 m_2 = \ell_1 \ell_2^2, \quad \mu^2(\ell_1) = 1, \quad \ell_2 \in \mathbb{Z},$$

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

$$(7.5.18) \quad \mathcal{H}_{1,p} = \begin{cases} \left( 1 - \frac{1}{p^{1+w}} \right)^2 \left( 1 - \frac{1}{p^{1+2s}} \right) & \text{if } p|2\alpha d \\ \frac{\left( 1 - \frac{1}{p^{1+w}} \right)^2}{\left( 1 - \frac{1}{p^{1+2s}} \right)} \left( 1 + \frac{2}{p^{1+w}} - \frac{2}{p^{1+2s-w}} + \frac{1}{p^{1+2s}} - \frac{3}{p^{2+2s}} + \frac{1}{p^{3+4s}} \right) & \text{if } p \nmid 2\alpha d m_1 m_2 \\ \frac{\left( 1 - \frac{1}{p^{1+w}} \right)^2}{\left( 1 - \frac{1}{p^{1+2s}} \right)} \left( 1 - \frac{1}{p^{2s-2w}} + \frac{2}{p^{2s-w}} - \frac{2}{p^{1+2s-w}} + \frac{1}{p^{1+2s}} - \frac{1}{p^{1+4s-2w}} \right) & \text{if } p|\ell_1 \\ \frac{\left( 1 - \frac{1}{p^{1+w}} \right)^2}{\left( 1 - \frac{1}{p^{1+2s}} \right)} \left( 1 - \frac{1}{p} + \frac{2}{p^{1+w}} - \frac{2}{p^{1+2s-w}} + \frac{1}{p^{1+2s}} - \frac{1}{p^{2+2s}} \right) & \text{if } p|m_1 m_2, \quad p \nmid \ell_1. \end{cases}$$

Inserting (7.5.16) into (7.5.9), we find that

$$(7.5.19) \quad \mathcal{P}_1 = \left( \frac{1+i}{2} \right) \frac{\sqrt{2}}{2} \sum_{\substack{\alpha \leq Y \\ (\alpha, 2m_1 m_2) = 1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \mathcal{I},$$

where

$$(7.5.20) \quad \begin{aligned} \mathcal{I} = & \operatorname{Res}_{w=0} \zeta(1+w)^2 \check{\Phi}(w) X^w \frac{1}{2\pi i} \int_{(c)} \frac{\Gamma\left(\frac{s}{2} + \frac{1}{4}\right)^2}{\Gamma\left(\frac{1}{4}\right)^2} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)^2 \pi^{-s} \Gamma_2(s-w) (8\alpha^2)^{s-w} \\ & \times \left(1 - \frac{1}{4^{s-w}}\right) m_1 m_2 \ell_1^{s-w-\frac{1}{2}} \zeta(2s-2w) \zeta(2s+1) \mathcal{H}_1(s-w, 1+w; m_1 m_2, \alpha d) \frac{ds}{s}. \end{aligned}$$

The next step is to extend the  $\alpha$ -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 (7.5.20) closer to 0 to guarantee the absolute convergence of the  $\alpha$ -sum. We first evaluate the residue using (6.1.6) to see that (7.5.20) is the same as

$$(7.5.21) \quad \begin{aligned} \mathcal{I} = & \frac{\check{\Phi}(0)}{2\pi i} \int_{(c)} \frac{\Gamma\left(\frac{s}{2} + \frac{1}{4}\right)^2}{\Gamma\left(\frac{1}{4}\right)^2} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)^2 \pi^{-s} \Gamma_2(s) (8\alpha^2)^s \left(1 - \frac{1}{4^s}\right) m_1 m_2 \\ & \times \ell_1^{s-\frac{1}{2}} \zeta(2s) \zeta(2s+1) \mathcal{H}_1(s, 1; m_1 m_2, \alpha d) \left\{ 2\gamma + \frac{(\check{\Phi})'(0)}{\check{\Phi}(0)} + \log \left( \frac{X}{8\alpha^2 \ell_1} \right) \right. \\ & \left. - \frac{\Gamma'_2(s)}{\Gamma_2(s)} + \frac{\log 4}{(1-4^s)} - 2 \frac{\zeta'(2s)}{\zeta(2s)} + \frac{\frac{\partial}{\partial w} \mathcal{H}_1(s-w, 1+w; m_1 m_2, \alpha d)}{\mathcal{H}_1(s-w, 1+w; m_1 m_2, \alpha d)} \right\} \bigg|_{w=0} \frac{ds}{s}. \end{aligned}$$

Here  $\gamma$  denotes the Euler-Mascheroni constant. The definition (7.5.18) of  $\mathcal{H}_1(s-w, 1+w; m_1 m_2, \alpha d)$  implies that it is holomorphic for  $\operatorname{Re}(s) > 0$  and  $|\operatorname{Re}(w)| < \min\{\frac{1}{2}, 2\operatorname{Re}(s)\}$ , and that it and its first partial derivatives at  $w = 0$  are bounded by  $\ll (\alpha X)^\varepsilon$  for  $\operatorname{Re}(s) \geq \frac{1}{\log X}$ . Thus, by the rapid decay of the gamma function, we may move the line of integration in (7.5.21) to  $\operatorname{Re}(s) = \frac{1}{\log X}$ . 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

$$\ll m_1 m_2 \ell_1^{-\frac{1}{2}+\varepsilon} \alpha^\varepsilon X^\varepsilon \int_{\left(\frac{1}{\log X}\right)} |\Gamma\left(\frac{s}{2} + \frac{1}{4}\right)|^2 \max\{|\Gamma_2(s)|, |\Gamma'_2(s)|\} (1 + |s|)^{\frac{1}{2}+\varepsilon} |ds|,$$

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

$$(7.5.22) \quad \sum_{\substack{\alpha > Y \\ (\alpha, 2m_1 m_2) = 1}} \frac{\mu^2(\alpha)}{\alpha^2 d_1} |\mathcal{I}| \ll \frac{m_1 m_2 \ell_1^{-\frac{1}{2}+\varepsilon} X^\varepsilon}{d^{1-\varepsilon} Y^{1-\varepsilon}}$$

because, by (7.2.5), 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 (7.5.19), (7.5.22), and (7.5.20) now with  $c = \frac{1}{\log X}$ , we arrive at

$$(7.5.23) \quad \begin{aligned} \mathcal{P}_1 = \operatorname{Res}_{w=0} \zeta(1+w)^2 \left( \frac{1+i}{2} \right) \frac{\sqrt{2}}{2} X^w \frac{1}{2\pi i} \int_{(\frac{1}{\log X})} \Gamma_2(s-w) 8^{s-w} \left( 1 - \frac{1}{4^{s-w}} \right) \\ \times \mathcal{K}(s, w; m_1 m_2, d) \frac{ds}{s} + O\left( \frac{m_1 m_2 \ell_1^{-\frac{1}{2}+\varepsilon} X^\varepsilon}{d^{1-\varepsilon} Y^{1-\varepsilon}} \right), \end{aligned}$$

with  $\mathcal{K}(s, w; m_1 m_2, d)$  defined by

$$(7.5.24) \quad \begin{aligned} \mathcal{K}(s, w; m_1 m_2, d) = \check{\Phi}(w) \frac{\Gamma\left(\frac{s}{2} + \frac{1}{4}\right)^2}{\Gamma\left(\frac{1}{4}\right)^2} \left( 1 - \frac{1}{2^{\frac{1}{2}-s}} \right)^2 \pi^{-s} m_1 m_2 \ell_1^{s-w-\frac{1}{2}} \zeta(2s-2w) \\ \times \zeta(2s+1) \sum_{\substack{\alpha=1 \\ (\alpha, 2m_1 m_2)=1}}^{\infty} \frac{\mu(\alpha)}{\alpha^{2-2s+2w} d_1} \mathcal{H}_1(s-w, 1+w; m_1 m_2, \alpha d), \end{aligned}$$

where, as before,  $\ell_1$  is defined by (7.5.17),  $d_1$  is defined by (7.2.5), and  $\mathcal{H}_1$  is defined as the product of (7.5.18) over all primes.

It is convenient for later calculations to write  $\mathcal{P}_1$  in terms of a residue, as in (7.5.23), rather than in terms of derivatives as in (7.5.21).

**7.6. Bounding the contribution of  $\mathcal{R}_1$ .** Having handled  $\mathcal{P}_1$  in (7.5.6), we next turn to  $\mathcal{R}_1$ , defined by (7.5.8). It will be convenient to denote

$$(7.6.1) \quad \begin{aligned} \mathcal{R}(\ell, d) = \frac{1}{\ell} \left( \frac{1+i}{2} \right) \frac{\sqrt{2}}{2} \sum_{\substack{\alpha \leq Y \\ (\alpha, 2\ell)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \left( \frac{d_1}{\ell} \right) \sum_{\substack{k \in \mathbb{Z} \\ k \text{ odd}}} \left( \frac{2}{kd_1} \right) \\ \times \frac{1}{2\pi i} \int_{(-\frac{1}{2}+\varepsilon)} h\left( \frac{kX}{8\alpha^2 d_1 \ell}, w \right) L(1+w, \chi_{k_1})^2 \mathcal{G}_0(1+w; k, \ell, \alpha, d) dw, \end{aligned}$$

so that  $\mathcal{R}_1 = m_1 m_2 \mathcal{R}(m_1 m_2, d)$ . We will bound  $|\mathcal{R}(\ell, d)|$  on average as  $\ell$  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  $\ell, d$  with  $J \leq \ell < 2J$  and  $V \leq d < 2V$ , where  $J, V \geq 1$ . We then insert the definition (7.6.1) and bring the  $d, \ell$ -sum inside the integral to deduce that

$$(7.6.2) \quad \sum_{\substack{d=V \\ (d, 2)=1}}^{2V-1} \sum_{\substack{\ell=J \\ (\ell, 2d)=1}}^{2J-1} |\mathcal{R}(\ell, d)| = \sum_{\substack{d=V \\ (d, 2)=1}}^{2V-1} \sum_{\substack{\ell=J \\ (\ell, 2d)=1}}^{2J-1} \beta_{\ell, d} \mathcal{R}(\ell, d) \ll \sum_{\substack{\alpha \leq Y \\ (\alpha, 2)=1}} \frac{1}{\alpha^2} \sum_{\substack{k \in \mathbb{Z} \\ k \text{ odd}}} \int_{(-\frac{1}{2}+\varepsilon)} U(\alpha, k, w) |dw|,$$

where for brevity we denote

$$U(\alpha, k, w) = \sum_{\substack{d=V \\ (d,2)=1}}^{2V-1} \frac{1}{d_1} |L(1+w, \chi_{k_1})|^2 \left| \sum_{\substack{\ell=J \\ (\ell, 2\alpha d)=1}}^{2J-1} \frac{\beta_{\ell,d}}{\ell} \left( \frac{d_1}{\ell} \right) \mathcal{G}_0(1+w; k, \ell, \alpha, d) h \left( \frac{kX}{8\alpha^2 d_1 \ell}, w \right) \right|.$$

We split the  $k$ -sum into dyadic blocks  $K \leq |k| < 2K$ , with  $K \geq 1$ , and apply Cauchy's inequality to write

(7.6.3)

$$\begin{aligned} \sum_{\substack{K \leq |k| < 2K \\ k \text{ odd}}} U(\alpha, k, w) &\ll \left( \sum_{\substack{d=V \\ (d,2)=1}}^{2V-1} \frac{1}{d_1} \sum_{\substack{K \leq |k| < 2K \\ k \text{ odd}}} k_2 |L(1+w, \chi_{k_1})|^4 \right)^{\frac{1}{2}} \\ &\times \left( \sum_{\substack{d=V \\ (d,2)=1}}^{2V-1} \frac{1}{d_1} \sum_{\substack{K \leq |k| < 2K \\ k \text{ odd}}} \frac{1}{k_2} \left| \sum_{\substack{\ell=J \\ (\ell, 2\alpha d)=1}}^{2J-1} \frac{\beta_{\ell,d}}{\ell} \left( \frac{d_1}{\ell} \right) \mathcal{G}_0(1+w; k, \ell, \alpha, d) h \left( \frac{kX}{8\alpha^2 d_1 \ell}, w \right) \right|^2 \right)^{\frac{1}{2}}, \end{aligned}$$

where  $k_2$  is defined by (7.5.3). To bound the first factor on the right-hand side of (7.6.3), 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 (7.2.5) to deduce that

$$\sum_{\substack{d=V \\ (d,2)=1}}^{2V-1} \frac{1}{d_1} \sum_{\substack{K \leq |k| < 2K \\ k \text{ odd}}} k_2 |L(1+w, \chi_{k_1})|^4 \leq \frac{\alpha}{V} \sum_{0 < |k_1| \ll KV} |L(1+w, \chi_{k_1})|^4 \sum_{k_2 \ll \sqrt{\frac{KV}{|k_1|}}} k_2 \sum_{\substack{d=V \\ (d,2)=1 \\ d_1 |k_1 k_2^2}}^{2V-1} 1.$$

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 4.5. It follows from this and (7.6.3) that

(7.6.4)

$$\begin{aligned} \sum_{\substack{K \leq |k| < 2K \\ k \text{ odd}}} U(\alpha, k, w) &\ll \left( \alpha K^{1+\varepsilon} V^\varepsilon (1+|w|)^{1+\varepsilon} \right)^{\frac{1}{2}} \\ &\times \left( \sum_{\substack{d=V \\ (d,2)=1}}^{2V-1} \frac{1}{d_1} \sum_{\substack{K \leq |k| < 2K \\ k \text{ odd}}} \frac{1}{k_2} \left| \sum_{\substack{\ell=J \\ (\ell, 2\alpha d)=1}}^{2J-1} \frac{\beta_{\ell,d}}{\ell} \left( \frac{d_1}{\ell} \right) \mathcal{G}_0(1+w; k, \ell, \alpha, d) h \left( \frac{kX}{8\alpha^2 d_1 \ell}, w \right) \right|^2 \right)^{\frac{1}{2}}. \end{aligned}$$

The next task is to bound the second factor on the right-hand side. To this end we prove the following two lemmas.

**Lemma 7.3.** *Let  $\alpha \leq Y$ ,  $d$ ,  $K$ , 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  $\gamma_\ell$  with  $|\gamma_\ell| \leq 1$ ,*

$$\sum_{\substack{K \leq |k| < 2K \\ k \text{ odd}}} \frac{1}{k_2} \left| \sum_{\substack{\ell=J \\ (\ell, 2\alpha d)=1}}^{2J-1} \frac{\gamma_\ell}{\ell} \mathcal{G}_0(1+w; k, \ell, \alpha, d) h \left( \frac{kX}{8\alpha^2 d_1 \ell}, w \right) \right|^2$$



is bounded by

$$\ll_{\varepsilon} |\check{\Phi}(w)|^2 \frac{d_1 \alpha^{2+\varepsilon} J^{2+\varepsilon} K^{\varepsilon} d^{\varepsilon}}{X^{1-\varepsilon}} \exp \left( -\frac{1}{20} \frac{\sqrt{K}}{\alpha \sqrt{d_1 J(1+|w|)}} \right).$$

and also by

$$\ll_{\varepsilon} ((1+|w|)\alpha d J K X)^{\varepsilon} |\check{\Phi}(w)|^2 \frac{\alpha^2 d_1 (J K + J^2)}{K X}.$$

**Lemma 7.4.** Let  $\delta_{\ell} \ll \ell^{\varepsilon}$  be any sequence of complex numbers and let  $\operatorname{Re}(w) = -\frac{1}{2} + \varepsilon$ . Then

$$\sum_{K \leq |k| < 2K} \frac{1}{k_2} \left| \sum_{\substack{\ell=J \\ (\ell, 2\alpha d)=1}}^{2J-1} \frac{\delta_{\ell}}{\sqrt{\ell}} \mathcal{G}_0(1+w; k, \ell, \alpha, d) \right|^2 \ll_{\varepsilon} (\alpha d J K)^{\varepsilon} J(J+K).$$

*Proof of Lemma 7.3 assuming Lemma 7.4.* To prove the first bound, we use the triangle inequality and apply the bounds for  $\mathcal{G}_0$  from Lemma 7.1 and  $h(\xi, w)$  from Lemma 7.2 to deduce that the sum in question is

$$\ll |\check{\Phi}(w)|^2 \frac{d_1 \alpha^{2+\varepsilon} J^{\varepsilon} K^{\varepsilon} d^{\varepsilon}}{X^{1-\varepsilon}} \exp \left( -\frac{1}{20} \frac{\sqrt{K}}{\alpha \sqrt{d_1 J(1+|w|)}} \right) \sum_{\substack{K \leq |k| < 2K \\ k \text{ odd}}} \frac{1}{|k| k_2} \left( \sum_{\substack{\ell=J \\ (\ell, 2\alpha d)=1}}^{2J-1} (\ell, k_4^2)^{\frac{1}{2}} \right)^2.$$

We then estimate the  $k$ -sum by splitting it according to the values of  $k_3$  and  $k_4$  and using  $(\ell, k_4^2) \leq k_4^2$  and  $\frac{1}{k_2} \leq \frac{1}{k_4}$ , which follows from (7.5.3) and (7.5.4). This leads to the first bound of the lemma.

To prove the second bound, we apply Lemma 7.2 and write the integral (7.5.5) 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  $\ell$ -sum inside the integral and use the triangle inequality to deduce that

$$\begin{aligned} & \left| \sum_{\substack{\ell=J \\ (\ell, 2\alpha d)=1}}^{2J-1} \frac{\gamma_{\ell}}{\ell} \mathcal{G}_0(1+w; k, \ell, \alpha, d) h \left( \frac{kX}{8\alpha^2 d_1 \ell}, w \right) \right| \\ & \ll |\check{\Phi}(w)| \left( \frac{\alpha^{1+\varepsilon} d_1^{\frac{1}{2}+\varepsilon}}{|k|^{\frac{1}{2}-\varepsilon} X^{\frac{1}{2}-\varepsilon}} \right) \int_{(\varepsilon)} \left| g(s, w; \operatorname{sgn}(k)) \sum_{\substack{\ell=J \\ (\ell, 2\alpha d)=1}}^{2J-1} \frac{\gamma_{\ell}}{\ell^{1+w-s}} \mathcal{G}_0(1+w; k, \ell, \alpha, d) \right| |ds|. \end{aligned}$$

Thus, since  $g(s, w; \operatorname{sgn}(k)) \ll_{\varepsilon} (1+|w|)^{\varepsilon} \exp(-(\frac{\pi}{2}-\varepsilon)|\operatorname{Im}(s)|)$  by Stirling's formula, it follows from Cauchy's inequality that

$$\begin{aligned} & \left| \sum_{\substack{\ell=J \\ (\ell, 2\alpha d)=1}}^{2J-1} \frac{\gamma_{\ell}}{\ell} \mathcal{G}_0(1+w; k, \ell, \alpha, d) h \left( \frac{kX}{8\alpha^2 d_1 \ell}, w \right) \right|^2 \ll (1+|w|)^{\varepsilon} |\check{\Phi}(w)|^2 \left( \frac{\alpha^{2+\varepsilon} d_1^{1+\varepsilon}}{|k|^{1-\varepsilon} X^{1-\varepsilon}} \right) \\ & \times \int_{(\varepsilon)} \exp(-(\frac{\pi}{2}-\varepsilon)|\operatorname{Im}(s)|) \left| \sum_{\substack{\ell=J \\ (\ell, 2\alpha d)=1}}^{2J-1} \frac{\gamma_{\ell}}{\ell^{1+w-s}} \mathcal{G}_0(1+w; k, \ell, \alpha, d) \right|^2 |ds|. \end{aligned}$$

The second bound of the lemma follows from this and Lemma 7.4.  $\square$

*Proof of Lemma 7.4.* For any integer  $k = \pm \prod_{i, a_i \geq 1} p_i^{a_i}$ , let  $a(k)$  and  $b(k)$  be defined by

$$(7.6.5) \quad a(k) = \prod_i p_i^{a_i+1} \quad \text{and} \quad b(k) = \prod_{a_i=1} p_i \prod_{a_i \geq 2} p_i^{a_i-1}.$$

From the definition of  $\mathcal{G}_0$  in Lemma 7.1, we see for  $(\ell, 2\alpha d) = 1$  that  $\mathcal{G}_0(1+w; k, \ell, \alpha, d) = 0$  unless  $\ell$  can be written as  $gm$  with  $g|a(k)$  and  $m$  square-free and relatively prime to  $k$ . With this expression for  $\ell$ , it follows from Lemma 4.3 that if  $(\ell, 2\alpha d) = 1$  then

$$(7.6.6) \quad \mathcal{G}_0(1+w; k, \ell, \alpha, d) = \sqrt{m} \left( \frac{k}{m} \right) \prod_{p|m} \left( 1 + \frac{2}{p^{1+w}} \left( \frac{k_1}{p} \right) \right)^{-1} \mathcal{G}_0(1+w; k, g, \alpha, d).$$

From this and Cauchy's inequality, we arrive at

$$(7.6.7) \quad \sum_{K \leq |k| < 2K} \frac{1}{k_2} \left| \sum_{\substack{\ell=J \\ (\ell, 2\alpha d)=1}}^{2J-1} \frac{\delta_\ell}{\sqrt{\ell}} \mathcal{G}_0(1+w; k, \ell, \alpha, d) \right|^2 \ll_\varepsilon K^\varepsilon \sum_{K \leq |k| < 2K} \frac{1}{k_2} \sum_{\substack{g|a(k) \\ g < 2J}} (\Psi_1(k, g) + \Psi_2(k, g)),$$

where

$$\Psi_1(k, g) = \left| \sum_{\substack{\frac{J}{g} \leq m < \frac{2J}{g} \\ (m, 2\alpha d)=1 \\ 3 \nmid m}} \frac{\mu^2(m) \delta_{gm}}{\sqrt{g}} \mathcal{G}_0(1+w; k, g, \alpha, d) \left( \frac{k}{m} \right) \prod_{p|m} \left( 1 + \frac{2}{p^{1+w}} \left( \frac{k_1}{p} \right) \right)^{-1} \right|^2$$

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  $\Psi_1$ . We factor out  $g^{-1/2} \mathcal{G}_0(1+w; k, g, \alpha, d)$  and apply the bound from Lemma 7.1 to deduce that

$$(7.6.8) \quad \Psi_1(k, g) \ll_\varepsilon (\alpha d K)^\varepsilon g^{1+\varepsilon} \left| \sum_{\substack{\frac{J}{g} \leq m < \frac{2J}{g} \\ (m, 6\alpha d)=1}} \mu^2(m) \delta_{gm} \left( \frac{k}{m} \right) \prod_{p|m} \left( 1 + \frac{2}{p^{1+w}} \left( \frac{k_1}{p} \right) \right)^{-1} \right|^2.$$

In the sum, if  $\left( \frac{k}{m} \right) \neq 0$ , then

$$\begin{aligned} \prod_{p|m} \left( 1 + \frac{2}{p^{1+w}} \left( \frac{k_1}{p} \right) \right)^{-1} &= \prod_{p|m} \left( 1 - \frac{4}{p^{2+2w}} \right)^{-1} \prod_{p|m} \left( 1 - \frac{2}{p^{1+w}} \left( \frac{k_1}{p} \right) \right) \\ &= \prod_{p|m} \left( 1 - \frac{4}{p^{2+2w}} \right)^{-1} \sum_{j|m} \frac{\mu(j) d_2(j)}{j^{1+w}} \left( \frac{k_1}{j} \right). \end{aligned}$$

We insert this into (7.6.8), interchange the order of summation, and apply Cauchy's inequality to see that

$$\Psi_1(k, g) \ll_\varepsilon (\alpha d K)^\varepsilon g^{1+\varepsilon} \sum_{j < \frac{2J}{g}} \left| \sum_{\substack{\frac{J}{g} \leq m < \frac{2J}{g} \\ (m, 6\alpha d)=1 \\ j|m}} \mu^2(m) \delta_{gm} \left( \frac{k}{m} \right) \prod_{p|m} \left( 1 - \frac{4}{p^{2+2w}} \right)^{-1} \right|^2.$$

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  $\operatorname{Re}(w) \geq -\frac{1}{2} + \varepsilon$  and  $p > 3$  for all  $p|m$ . The result is

$$(7.6.9) \quad \Psi_1(k, g) \ll_\varepsilon (\alpha dJK)^\varepsilon g^{1+\varepsilon} \sum_{j < \frac{2J}{g}} \left| \sum_{\substack{\frac{J}{gj} \leq m < \frac{2J}{gj} \\ (m, 6\alpha dj)=1}} \mu^2(m) \delta_{gjm} \left(\frac{k}{m}\right) \prod_{p|m} \left(1 - \frac{4}{p^{2+2w}}\right)^{-1} \right|^2.$$

Now, by (7.6.5),  $g|a(k)$  implies  $b(g)|k$ . Thus we may interchange the order of summation to write

$$\sum_{K \leq |k| < 2K} \frac{1}{k_2} \sum_{\substack{g|a(k) \\ g < 2J}} \Psi_1(k, g) \leq \sum_{g < 2J} \sum_{\substack{K \leq |k| < 2K \\ b(g)|k}} \frac{1}{k_2} \Psi_1(k, g) = \sum_{g < 2J} \sum_{\substack{\frac{K}{b(g)} \leq |f| < \frac{2K}{b(g)}}} \frac{1}{k_2} \Psi_1(fb(g), g),$$

where we have relabeled  $k$  in the last sum as  $fb(g)$ , so that, by (7.5.3),  $k_2 > 0$  satisfies  $4fb(g)d_1 = k_1 k_2^2$ , with  $k_1$  a fundamental discriminant. From this and (7.6.9), we arrive at

$$(7.6.10) \quad \sum_{K \leq |k| < 2K} \frac{1}{k_2} \sum_{\substack{g|a(k) \\ g < 2J}} \Psi_1(k, g) \ll_\varepsilon (\alpha dJK)^\varepsilon \sum_{g < 2J} g \sum_{\substack{\frac{K}{b(g)} \leq |f| < \frac{2K}{b(g)}}} \frac{1}{k_2} \\ \times \sum_{j < \frac{2J}{g}} \left| \sum_{\substack{\frac{J}{gj} \leq m < \frac{2J}{gj} \\ (m, 6\alpha dj)=1}} \mu^2(m) \delta_{gjm} \left(\frac{fb(g)}{m}\right) \prod_{p|m} \left(1 - \frac{4}{p^{2+2w}}\right)^{-1} \right|^2.$$

If  $4f = f_1 f_2^2$ , with  $f_1$  a fundamental discriminant and  $f_2$  a positive integer, then the equation  $4fb(g)d_1 = k_1 k_2^2$  implies that  $f_2 | 2k_2$ , and thus  $k_2^{-1} \ll f_2^{-1}$ . This and (7.6.10) imply

$$\sum_{K \leq |k| < 2K} \frac{1}{k_2} \sum_{\substack{g|a(k) \\ g < 2J}} \Psi_1(k, g) \ll_\varepsilon (\alpha dJK)^\varepsilon \sum_{g < 2J} g \sum_{j < \frac{2J}{g}} \sum_{\substack{\frac{K}{b(g)} \leq |f| < \frac{2K}{b(g)}}} \frac{1}{f_2} \left| \sum_{\substack{\frac{J}{gj} \leq m < \frac{2J}{gj} \\ (m, 6\alpha dj)=1}} a_m \left(\frac{f}{m}\right) \right|^2,$$

where

$$a_m = \mu^2(m) \delta_{gjm} \left(\frac{b(g)}{m}\right) \prod_{p|m} \left(1 - \frac{4}{p^{2+2w}}\right)^{-1}.$$

It follows from this and Lemma 4.4 that

$$\begin{aligned} \sum_{K \leq |k| < 2K} \frac{1}{k_2} \sum_{\substack{g|a(k) \\ g < 2J}} \Psi_1(k, g) &\ll_\varepsilon (\alpha dJK)^\varepsilon \sum_{g < 2J} g \sum_{j < \frac{2J}{g}} \frac{2J}{gj} \left(\frac{2K}{b(g)} + \frac{2J}{gj}\right) \\ &\ll_\varepsilon (\alpha dJK)^\varepsilon J(K+J) \sum_{g < 2J} \frac{1}{b(g)} \sum_{j < \frac{2J}{g}} \frac{1}{j} \\ &\ll_\varepsilon (\alpha dJK)^\varepsilon J(K+J) \end{aligned}$$

since  $\sum_{g < 2J} b(g)^{-1} \ll (\log J)^2$ . This proves the desired bound for the sum of  $\Psi_1(k, g)$  in (7.6.7). To bound the sum of  $\Psi_2(k, g)$ , we argue in the same way, but instead of (7.6.6) 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

$$\mathcal{G}_0^*(1+w; k, g, \alpha, d) = \left( 1 - \frac{1}{3^{1+w}} \left( \frac{k_1}{3} \right) \right)^2 \prod_{p \neq 3} \mathcal{G}_{0,p}(1+w; k, g, \alpha, d),$$

with  $\mathcal{G}_{0,p}$  as defined in Lemma 7.1.  $\square$

We now estimate the contribution of  $\mathcal{R}_1$ . From the first bound of Lemma 7.3, we see that the sum of the right-hand side of (7.6.4) over all  $K = 2^j > \alpha^2 V J (1 + |w|) (\log X)^4$  is negligible. On the other hand, if  $K \leq \alpha^2 V J (1 + |w|) (\log X)^4$  then it follows from (7.6.4) and the second bound in Lemma 7.3 that

$$\begin{aligned} \sum_{\substack{K \leq |k| < 2K \\ k \text{ odd}}} U(\alpha, k, w) &\ll_{\varepsilon} (1 + |w|)^{\frac{1}{2} + \varepsilon} |\check{\Phi}(w)| (\alpha J K V X)^{\varepsilon} \left( \frac{\alpha^3 V (J K + J^2)}{X} \right)^{\frac{1}{2}} \\ &\ll_{\varepsilon} (1 + |w|)^{1 + \varepsilon} |\check{\Phi}(w)| (\alpha J K V X)^{\varepsilon} \frac{\alpha^{\frac{5}{2}} V J}{X^{\frac{1}{2}}}. \end{aligned}$$

We sum this over all  $K = 2^j$ ,  $j$  a nonnegative integer, with  $K \leq \alpha^2 V J (1 + |w|) (\log X)^4$ , and then multiply the resulting sum by  $\alpha^{-2}$ . We then integrate over all  $w$  with  $\operatorname{Re}(w) = -\frac{1}{2} + \varepsilon$  and sum over all integers  $\alpha \leq Y$  to deduce from (7.6.2) that

$$(7.6.11) \quad \sum_{\substack{d=V \\ (d,2)=1}}^{2V-1} \sum_{\substack{\ell=J \\ (\ell,2d)=1}}^{2J-1} |\mathcal{R}(\ell, d)| \ll \frac{V^{1+\varepsilon} J^{1+\varepsilon} Y^{\frac{3}{2}+\varepsilon}}{X^{\frac{1}{2}-\varepsilon}}.$$

Recall from (7.5.8) and (7.6.1) that  $\mathcal{R}_1 = m_1 m_2 \mathcal{R}(m_1 m_2, d)$ . Since  $\lambda_d \ll d^{\varepsilon}$  by (5.8) and  $b_m \ll 1$  by (3.5), it thus follows from (7.6.11) that

$$(7.6.12) \quad \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} |\mathcal{R}_1| \ll \frac{D^{1+\varepsilon} M^{1+\varepsilon} Y^{\frac{3}{2}+\varepsilon}}{X^{\frac{1}{2}-\varepsilon}}.$$

**7.7. Conditions for the parameters.** From (7.5.1), (7.5.6), (7.5.23), and (7.6.12), we see that the total contribution of the sum with  $\mathcal{Q}_1$  to  $\mathcal{B}$  in (7.4.1) is

$$\begin{aligned} (7.7.1) \quad & \frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} \sum_{\substack{\nu=1 \\ (\nu, 2d)=1}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2}} \sum_{\substack{\alpha \leq Y \\ (\alpha, 2m_1 m_2 \nu)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \mathcal{Q}_1 \\ &= \left( \frac{1+i}{2} \right) \frac{\sqrt{2} X}{2(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} \operatorname{Res}_{w=0} \zeta^2(1+w) \frac{X^w}{2\pi i} \int_{(\frac{1}{\log X})} \Gamma_2(s-w) \\ &\quad \times 8^{s-w} \left( 1 - \frac{1}{4^{s-w}} \right) \mathcal{K}(s, w; m_1 m_2, d) \frac{ds}{s} + O \left( \frac{X^{1+\varepsilon} D^{\varepsilon} M^{\varepsilon}}{Y^{1-\varepsilon}} + X^{\frac{1}{2}+\varepsilon} D^{1+\varepsilon} M^{1+\varepsilon} Y^{\frac{3}{2}+\varepsilon} \right). \end{aligned}$$

Recall the definition (3.4) of  $M$ . Also, recall the definitions (5.9) and (5.2) of  $D$  and  $R$ , respectively. So that the error terms in (7.7.1) are  $O(X^{1-\varepsilon})$ , we assume the parameters  $\theta$  and  $\vartheta$  satisfy

$$\theta + 2\vartheta < \frac{1}{2},$$

and we take the parameter  $Y$  in (3.6) to be

$$Y = X^\delta$$

with  $\delta = \delta(\theta, \vartheta)$  sufficiently small.

**7.8. Evaluating the sums of the other terms with  $k \neq 0$ .** The procedure for evaluating the sum with  $\mathcal{Q}_2$  in (7.4.1) 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 (7.4.3). This causes the residues in the versions of (7.5.6) and (7.5.7) for  $\mathcal{Q}_2$  to have each  $-kd_1$  equal to a perfect square instead of  $kd_1 = \square$ . This means  $\text{sgn}(k) = -1$ . Hence, because of the factor  $\text{sgn}(\xi)$  in (7.5.5), the version of (7.5.9) for  $\mathcal{Q}_2$  has the function

$$(2\pi)^{-u}\Gamma(u)(\cos(\tfrac{\pi}{2}u) + i\sin(\tfrac{\pi}{2}u))$$

in place of the function  $\Gamma_2(u)$  defined by (7.5.10). These lead to a version of (7.7.1) for  $\mathcal{Q}_2$  that we may combine with (7.7.1) using the identity

$$(7.8.1) \quad \left(\frac{1+i}{2}\right)(\cos u - i\sin u) + \left(\frac{1-i}{2}\right)(\cos u + i\sin u) = \cos u + \sin u.$$

The result is

$$(7.8.2) \quad \begin{aligned} & \frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} \sum_{\substack{\nu=1 \\ (\nu, 2d)=1}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2}} \sum_{\substack{\alpha \leq Y \\ (\alpha, 2m_1 m_2 \nu)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} (\mathcal{Q}_1 + \mathcal{Q}_2) \\ &= \frac{\sqrt{2}X}{2(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} \text{Res}_{w=0} \zeta^2(1+w) \frac{X^w}{2\pi i} \int_{(\frac{1}{\log X})} \Gamma_1(s-w) \\ & \quad \times 8^{s-w} \left(1 - \frac{1}{4^{s-w}}\right) \mathcal{K}(s, w; m_1 m_2, d) \frac{ds}{s} + O(X^{1-\varepsilon}), \end{aligned}$$

where

$$(7.8.3) \quad \Gamma_1(u) = (2\pi)^{-u}\Gamma(u)(\cos(\tfrac{\pi}{2}u) + \sin(\tfrac{\pi}{2}u))$$

and the bound  $O(X^{1-\varepsilon})$  for the error term is guaranteed by the conditions in Subsection 7.7.

The evaluation of the sums in (7.4.1) with  $\mathcal{Q}_3$  and  $\mathcal{Q}_4$  defined by (7.4.4) and (7.4.5) is similar. The version of (7.5.7) 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 (7.4.1) with  $\mathcal{Q}_3$  and  $\mathcal{Q}_4$  is exactly the same as the right-hand side of (7.8.2). Therefore

$$\begin{aligned}
& \frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} \sum_{\substack{\nu=1 \\ (\nu, 2d)=1}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2}} \sum_{\substack{\alpha \leq Y \\ (\alpha, 2m_1 m_2 \nu)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \sum_{j=1}^4 \mathcal{Q}_j \\
&= \frac{\sqrt{2}X}{(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} \operatorname{Res}_{w=0} \zeta^2(1+w) \frac{X^w}{2\pi i} \int_{(\frac{1}{\log X})} \Gamma_1(s-w) \\
&\quad \times 8^{s-w} \left(1 - \frac{1}{4^{s-w}}\right) \mathcal{K}(s, w; m_1 m_2, d) \frac{ds}{s} + O(X^{1-\varepsilon}).
\end{aligned}$$

To estimate the sum with  $\mathcal{U}_1$  in (7.4.1), we first relabel  $k$  in (7.4.6) as  $2k$ , now with  $k$  odd, to write

$$(7.8.5) \quad \mathcal{U}_1 = \left(\frac{1+i}{2}\right) \left(\frac{2d_1}{m_1 m_2 \nu}\right) \sum_{\substack{k \in \mathbb{Z} \\ k \text{ odd}}} e\left(\frac{k \alpha^2 d_1 m_1 m_2 \nu}{4}\right) \hat{F}_\nu \left(\frac{kX}{4\alpha^2 d_1 m_1 m_2 \nu}\right) G_{2k}(m_1 m_2 \nu).$$

From the definition (4.2) 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 (7.8.5) that

$$\mathcal{U}_1 = i \left(\frac{1+i}{2}\right) \left(\frac{-d_1}{m_1 m_2 \nu}\right) \sum_{\substack{k \in \mathbb{Z} \\ k \text{ odd}}} \left(\frac{-1}{kd_1}\right) \hat{F}_\nu \left(\frac{kX}{4\alpha^2 d_1 m_1 m_2 \nu}\right) G_k(m_1 m_2 \nu).$$

We then proceed as we did for  $\mathcal{Q}_1$ . We treat the sum with  $\mathcal{U}_2$ , defined by (7.4.7), in a similar way. We combine the resulting expressions using the identity (7.8.1), and we arrive at

$$\begin{aligned}
& \frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} \sum_{\substack{\nu=1 \\ (\nu, 2d)=1}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2}} \sum_{\substack{\alpha \leq Y \\ (\alpha, 2m_1 m_2 \nu)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} (\mathcal{U}_1 + \mathcal{U}_2) \\
&= \frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} \operatorname{Res}_{w=0} \zeta^2(1+w) \frac{X^w}{2\pi i} \int_{(\frac{1}{\log X})} \Gamma_1(s-w) \\
&\quad \times 4^{s-w} \left(1 - \frac{1}{4^{s-w}}\right) \mathcal{K}(s, w; m_1 m_2, d) \frac{ds}{s} + O(X^{1-\varepsilon}).
\end{aligned}$$

Next, to evaluate the sum with  $\mathcal{V}$  in (7.4.1), we relabel  $k$  in (7.4.8) as  $4k$ , now with  $k$  odd, to see that

$$\mathcal{V} = - \left(\frac{2d_1}{m_1 m_2 \nu}\right) \sum_{\substack{k \in \mathbb{Z} \\ k \text{ odd}}} \hat{F}_\nu \left(\frac{kX}{2\alpha^2 d_1 m_1 m_2 \nu}\right) \tau_k(m_1 m_2 \nu)$$

since  $e(h/2) = -1$  for odd  $h$  and  $\tau_{4k}(n) = \tau_k(n)$  for odd  $n$  by (4.3). Into this we insert the second expression for  $\tau_k(n)$  in (4.3). Since  $\left(\frac{-1}{n}\right) G_k(n) = G_{-k}(n)$  by (4.2), 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

$$(7.8.7) \quad \mathcal{V} = - \left( \frac{2d_1}{m_1 m_2 \nu} \right) \sum_{\substack{k \in \mathbb{Z} \\ k \text{ odd}}} \tilde{F}_\nu \left( \frac{kX}{2\alpha^2 d_1 m_1 m_2 \nu} \right) G_k(m_1 m_2 \nu),$$

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 [40, Lemma 5.2] instead of Lemma 7.2. We arrive at versions of (7.5.6), (7.5.7), and (7.5.8) which show that the residue at  $w = 0$  equals zero because  $2kd_1 \neq \square$  when  $kd_1$  is odd. This leads to

$$(7.8.8) \quad \frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} \sum_{\substack{\nu=1 \\ (\nu, 2d)=1}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2}} \sum_{\substack{\alpha \leq Y \\ (\alpha, 2m_1 m_2 \nu)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \mathcal{V} = O(X^{1-\varepsilon})$$

under the conditions in Subsection 7.7.

Lastly, to estimate the sum with  $\mathcal{W}$  in (7.4.1), we relabel  $k$  in (7.4.9) as  $8k$  to write

$$\mathcal{W} = \left( \frac{d_1}{m_1 m_2 \nu} \right) \sum_{\substack{k \in \mathbb{Z} \\ k \neq 0}} \hat{F}_\nu \left( \frac{kX}{\alpha^2 d_1 m_1 m_2 \nu} \right) \tau_k(m_1 m_2 \nu)$$

using the fact that  $e(h) = 1$  for any integer  $h$  and  $\tau_{8k}(n) = \left(\frac{2}{n}\right) \tau_k(n)$  for odd  $n$  by (4.3). Into this we insert the second expression for  $\tau_k(n)$  in (4.3), 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 (7.8.7) to deduce that

$$\mathcal{W} = \left( \frac{d_1}{m_1 m_2 \nu} \right) \sum_{\substack{k \in \mathbb{Z} \\ k \neq 0}} \tilde{F}_\nu \left( \frac{kX}{\alpha^2 d_1 m_1 m_2 \nu} \right) G_k(m_1 m_2 \nu).$$

We then proceed as we did for  $\mathcal{Q}_1$ , using [40, Lemma 5.2] instead of Lemma 7.2. Since we are now summing over all nonzero integers  $k$  and not just the odd ones, instead of (7.5.15) and (7.5.16) we use

$$\begin{aligned} \sum_{j=1}^{\infty} j^{-2s+2w} \mathcal{G}_0(1+w; j^2, m_1 m_2, \alpha d, 1) &= \prod_p \sum_{b=0}^{\infty} p^{2b(w-s)} \mathcal{G}_{0,p}(1+w; p^{2b}, m_1 m_2, \alpha d, 1) \\ &= (m_1 m_2)^{1-s+w} \ell_1^{s-w-\frac{1}{2}} \zeta(2s-2w) \zeta(2s+1) \mathcal{H}_1(s-w, 1+w; m_1 m_2, \alpha d). \end{aligned}$$

We arrive at

$$\begin{aligned}
(7.8.9) \quad & \frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} \sum_{\substack{\nu=1 \\ (\nu, 2d)=1}}^{\infty} \frac{d_2(\nu)}{\nu^{3/2}} \sum_{\substack{\alpha \leq Y \\ (\alpha, 2m_1 m_2 \nu)=1}} \frac{\mu(\alpha)}{\alpha^2 d_1} \mathcal{W} \\
& = \frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} \\
& \quad \times \operatorname{Res}_{w=0} \zeta^2(1+w) \frac{X^w}{2\pi i} \int_{\left(\frac{1}{\log X}\right)} \Gamma_1(s-w) \mathcal{K}(s, w; m_1 m_2, d) \frac{ds}{s} + O(X^{1-\varepsilon}).
\end{aligned}$$

**7.9. Putting together the estimates.** From (7.4.1), (7.8.4), (7.8.6), (7.8.8), and (7.8.9), we deduce that

$$\begin{aligned}
\mathcal{B} &= \frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} \operatorname{Res}_{w=0} \zeta^2(1+w) \frac{X^w}{2\pi i} \int_{\left(\frac{1}{\log X}\right)} \Gamma_1(s-w) \\
& \quad \times \left( 8^{s-w} \sqrt{2} + 4^{s-w} - 2^{s-w} \sqrt{2} \right) \mathcal{K}(s, w; m_1 m_2, d) \frac{ds}{s} + O(X^{1-\varepsilon}),
\end{aligned}$$

where we recall that  $\Gamma_1$  is defined in (7.8.3) and  $\mathcal{K}$  is defined in (7.5.24). 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 (6.1.6) with  $n = 2$  to write

$$\begin{aligned}
(7.9.1) \quad \mathcal{B} &= \frac{X}{(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} \frac{1}{2\pi i} \int_{\left(\frac{1}{\log X}\right)} \Gamma_1(s) \left( 8^s \sqrt{2} + 4^s - 2^s \sqrt{2} \right) \\
& \quad \times \mathcal{K}(s, 0; m_1 m_2, d) \left\{ 2\gamma + \log X - \frac{\Gamma'_1(s)}{\Gamma_1(s)} - (\log 2) \frac{3 \cdot 8^s \sqrt{2} + 2 \cdot 4^s - 2^s \sqrt{2}}{8^s \sqrt{2} + 4^s - 2^s \sqrt{2}} \right. \\
& \quad \left. + \frac{\partial}{\partial w} \mathcal{K}(s, w; m_1 m_2, d) \bigg|_{w=0} \right\} \frac{ds}{s} + O(X^{1-\varepsilon}).
\end{aligned}$$

To simplify this, we evaluate the expressions involving  $\mathcal{K}$  by using Euler products, as follows. By (7.2.5), the definition (7.5.18) of  $\mathcal{H}_1$ , and the general formula (7.5.14), we may write the  $\alpha$ -sum in the definition (7.5.24) of  $\mathcal{K}$  as

$$\begin{aligned}
(7.9.2) \quad & \sum_{\substack{\alpha=1 \\ (\alpha, 2m_1 m_2)=1}}^{\infty} \frac{\mu(\alpha)}{\alpha^{2-2s+2w} d_1} \mathcal{H}_1(s-w, 1+w; m_1 m_2, \alpha d) = \frac{1}{d} \prod_{p|2m_1 m_2} \mathcal{H}_{1,p} \\
& \quad \times \prod_{p|d} \left( 1 - \frac{1}{p^{1+w}} \right)^2 \left( 1 - \frac{1}{p^{1+2s}} \right) \left( 1 - \frac{1}{p^{1-2s+2w}} \right) \prod_{p \nmid 2m_1 m_2 d} \frac{\left( 1 - \frac{1}{p^{1+w}} \right)^2}{\left( 1 - \frac{1}{p^{1+2s}} \right)} \left( 1 + \frac{2}{p^{1+w}} \right. \\
& \quad \left. - \frac{2}{p^{1+2s-w}} + \frac{1}{p^{1+2s}} - \frac{3}{p^{2+2s}} + \frac{1}{p^{3+4s}} - \frac{1}{p^{2-2s+2w}} + \frac{2}{p^{3+2w}} - \frac{1}{p^{4+2s+2w}} \right).
\end{aligned}$$



We insert this into (7.5.24), put  $w = 0$ , and use the identity

$$\left(8^s\sqrt{2} + 4^s - 2^s\sqrt{2}\right) \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right) \left(1 - \frac{1}{2^{1+2s}}\right) = 4^s \left(1 - \frac{1}{2^{\frac{1}{2}+s}}\right) \left(\frac{5}{2} - 4^s - 4^{-s}\right)$$

to deduce that

(7.9.3)

$$\begin{aligned} & \Gamma_1(s) \left(8^s\sqrt{2} + 4^s - 2^s\sqrt{2}\right) \mathcal{K}(s, 0; m_1 m_2, d) \\ &= \frac{\check{\Phi}(0)}{4} \frac{\Gamma\left(\frac{s}{2} + \frac{1}{4}\right)^2}{\Gamma\left(\frac{1}{4}\right)^2} \left(\frac{4}{\pi}\right)^s \Gamma_1(s) \zeta(2s) \zeta(2s+1) \left(1 - \frac{1}{2^{\frac{1}{2}+s}}\right) \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right) \left(\frac{5}{2} - 4^s - 4^{-s}\right) \\ & \times \frac{\varphi(dm_1 m_2)^2}{d^3 m_1 m_2 \sqrt{\ell_1}} \sum_{ab=\ell_1} \left(\frac{a}{b}\right)^s \prod_{\substack{p|m_1 m_2 \\ p \nmid \ell_1}} \left(1 + \frac{1}{p}\right) \prod_{p|d} \left(1 - \frac{1}{p^{1+2s}}\right) \left(1 - \frac{1}{p^{1-2s}}\right) \\ & \times \prod_{p \nmid 2m_1 m_2 d} \left\{ \left(1 - \frac{1}{p}\right)^2 \left(1 + \frac{2}{p} + \frac{1}{p^3} - \frac{1}{p^{2-2s}} - \frac{1}{p^{2+2s}}\right) \right\}, \end{aligned}$$

where  $\ell_1$  is defined by (7.5.17). Furthermore, we insert (7.9.2) into (7.5.24), take the logarithmic derivative with respect to  $w$  at  $w = 0$ , and use the identities

$$-(\log 2) \frac{3 \cdot 8^s \sqrt{2} + 2 \cdot 4^s - 2^s \sqrt{2}}{8^s \sqrt{2} + 4^s - 2^s \sqrt{2}} + \frac{2 \log 2}{2^{1+2s} - 1} = -3 \log 2 + \frac{\log 2}{(\sqrt{2} + 2^s)(\sqrt{2} + 2^{-s})}$$

and

$$-\sum_p \frac{2 \log p}{p^{1+2s} - 1} = 2 \frac{\zeta'}{\zeta}(2s+1), \quad \operatorname{Re}(s) > 0,$$

to arrive at

$$\begin{aligned} & -\frac{\Gamma'_1(s)}{\Gamma_1(s)} - (\log 2) \frac{3 \cdot 8^s \sqrt{2} + 2 \cdot 4^s - 2^s \sqrt{2}}{8^s \sqrt{2} + 4^s - 2^s \sqrt{2}} + \frac{\frac{\partial}{\partial w} \mathcal{K}(s, w; m_1 m_2, d)}{\mathcal{K}(s, w; m_1 m_2, d)} \Big|_{w=0} \\ &= \frac{(\check{\Phi})'(0)}{\check{\Phi}(0)} - \log(2\ell_1) - \frac{\Gamma'_1(s)}{\Gamma_1(s)} - 2 \frac{\zeta'}{\zeta}(2s) + 2 \frac{\zeta'}{\zeta}(2s+1) + \frac{\log 2}{(\sqrt{2} + 2^s)(\sqrt{2} + 2^{-s})} \\ (7.9.4) \quad & + \sum_{p|d} \left( \frac{2 \log p}{p-1} + \frac{2 \log p}{p^{1+2s}-1} + \frac{2 \log p}{p^{1-2s}-1} \right) + \sum_{p|m_1 m_2} \frac{2 \log p}{p-1} - \sum_{\substack{p|m_1 m_2 \\ p \nmid \ell_1}} \frac{2 \log p}{p+1} \\ & + \sum_{p \nmid 2m_1 m_2 d} \left( \frac{2 \log p}{p-1} - \left( \frac{2 \log p}{p} \right) \frac{1 + \frac{2}{p^2} - \frac{1}{p}(p^{2s} + p^{-2s})}{1 + \frac{2}{p} + \frac{1}{p^3} - \frac{1}{p^2}(p^{2s} + p^{-2s})} \right). \end{aligned}$$

Now the definition (7.8.3) 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 (7.9.3) and (7.9.4) are even functions of  $s$ . It follows that the integrand in (7.9.1) is an odd function of  $s$ . We move the line of integration in (7.9.1) to  $\text{Re}(s) = -\frac{1}{\log X}$ , 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 (7.9.1). 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

$$(7.9.5) \quad \mathcal{B} = \frac{X}{2(\sqrt{2}-1)^4} \sum_{\substack{d \leq D \\ d \text{ odd}}} \mu^2(d) \lambda_d \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2d)=1}} \frac{b_{m_1} b_{m_2}}{(m_1 m_2)^{3/2}} \frac{1}{2\pi i} \oint_{|s|=\frac{1}{\log X}} \Gamma_1(s) \left( 8^s \sqrt{2} + 4^s - 2^s \sqrt{2} \right) \\ \times \mathcal{K}(s, 0; m_1 m_2, d) \left\{ 2\gamma + \log X - \frac{\Gamma'_1(s)}{\Gamma_1(s)} - (\log 2) \frac{3 \cdot 8^s \sqrt{2} + 2 \cdot 4^s - 2^s \sqrt{2}}{8^s \sqrt{2} + 4^s - 2^s \sqrt{2}} \right. \\ \left. + \frac{\frac{\partial}{\partial w} \mathcal{K}(s, w; m_1 m_2, d)}{\mathcal{K}(s, w; m_1 m_2, d)} \Big|_{w=0} \right\} \frac{ds}{s} + O(X^{1-\varepsilon}).$$

The next step is to carry out the summation over  $d$ . From (7.9.3) and (7.9.4), we see that we need to evaluate the sums  $\Sigma_1$  and  $\Sigma_2$  defined by

$$(7.9.6) \quad \Sigma_1 = \sum_{\substack{d \leq D \\ (d, 2m_1 m_2)=1}} \mu^2(d) \lambda_d \frac{\varphi(d)^2}{d^3} \prod_{p|d} \left( 1 - \frac{1}{p^{1+2s}} \right) \left( 1 - \frac{1}{p^{1-2s}} \right) \\ \times \prod_{p|2m_1 m_2 d} \left\{ \left( 1 - \frac{1}{p} \right)^2 \left( 1 + \frac{2}{p} + \frac{1}{p^3} - \frac{1}{p^{2-2s}} - \frac{1}{p^{2+2s}} \right) \right\}$$

and

$$(7.9.7) \quad \Sigma_2 = \sum_{\substack{d \leq D \\ (d, 2m_1 m_2)=1}} \mu^2(d) \lambda_d \frac{\varphi(d)^2}{d^3} \prod_{p|d} \left( 1 - \frac{1}{p^{1+2s}} \right) \left( 1 - \frac{1}{p^{1-2s}} \right) \\ \times \prod_{p|2m_1 m_2 d} \left\{ \left( 1 - \frac{1}{p} \right)^2 \left( 1 + \frac{2}{p} + \frac{1}{p^3} - \frac{1}{p^{2-2s}} - \frac{1}{p^{2+2s}} \right) \right\} \sum_{p|d} J(p, s),$$

where

$$(7.9.8) \quad J(p, s) = \frac{2 \log p}{p^{1+2s} - 1} + \frac{2 \log p}{p^{1-2s} - 1} + \left( \frac{2 \log p}{p} \right) \frac{1 + \frac{2}{p^2} - \frac{1}{p} (p^{2s} + p^{-2s})}{1 + \frac{2}{p} + \frac{1}{p^3} - \frac{1}{p^2} (p^{2s} + p^{-2s})}$$

and  $|s| = \frac{1}{\log X}$ . We only estimate  $\Sigma_1$  since  $\Sigma_2$  may be treated in the same way, except using Lemma 5.4 instead of Lemma 5.3. We rearrange the factors in (7.9.6) to write  $\Sigma_1$  as

$$(7.9.9) \quad \begin{aligned} \Sigma_1 = & \prod_{p \nmid 2m_1m_2} \left\{ \left(1 - \frac{1}{p}\right)^2 \left(1 + \frac{2}{p} + \frac{1}{p^3} - \frac{1}{p^{2-2s}} - \frac{1}{p^{2+2s}}\right) \right\} \sum_{\substack{d \leq D \\ (d, 2m_1m_2)=1}} \frac{\mu^2(d)\lambda_d}{d} \\ & \times \prod_{p|d} \left(1 - \frac{1}{p^{1+2s}}\right) \left(1 - \frac{1}{p^{1-2s}}\right) \left(1 + \frac{2}{p} + \frac{1}{p^3} - \frac{1}{p^{2-2s}} - \frac{1}{p^{2+2s}}\right)^{-1}. \end{aligned}$$

Now recall the definition (5.1) of  $z_0$  and the definition (5.8) of  $\lambda_d$ . Factoring out the product over primes  $p > z_0$ , we see that

$$\begin{aligned} & \prod_{p \nmid 2m_1m_2} \left\{ \left(1 - \frac{1}{p}\right)^2 \left(1 + \frac{2}{p} + \frac{1}{p^3} - \frac{1}{p^{2-2s}} - \frac{1}{p^{2+2s}}\right) \right\} \\ & = \left(1 + O\left(\frac{1}{z_0}\right)\right) \prod_{\substack{p \nmid 2m_1m_2 \\ p \leq z_0}} \left\{ \left(1 - \frac{1}{p}\right)^2 \left(1 + \frac{2}{p} + \frac{1}{p^3} - \frac{1}{p^{2-2s}} - \frac{1}{p^{2+2s}}\right) \right\}. \end{aligned}$$

From this, (7.9.9), Lemma 5.3, and some simplification, we deduce that

$$(7.9.10) \quad \Sigma_1 = \left(1 + O\left(\frac{1}{z_0}\right)\right) \frac{1 + o(1)}{\log R} \prod_{\substack{p \nmid 2m_1m_2 \\ p \leq z_0}} \left(1 - \frac{1}{p^2}\right) \prod_{\substack{p \nmid 2m_1m_2 \\ p \leq z_0}} \left(1 - \frac{1}{p}\right)^{-1} + O_A\left(\frac{1}{(\log R)^A}\right).$$

The condition  $p \leq z_0$  may be omitted because  $\prod_{p > z_0} (1 + O(\frac{1}{p^2})) = 1 + O(\frac{1}{z_0})$  and

$$\prod_{\substack{p \nmid 2m_1m_2 \\ p > z_0}} \left(1 - \frac{1}{p}\right)^{-1} = \left(1 + O\left(\frac{1}{z_0}\right)\right)^{O(\log X)} = 1 + O\left(\frac{\log X}{z_0}\right).$$

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 (7.9.10), we arrive at

$$(7.9.11) \quad \Sigma_1 = \frac{2m_1m_2}{\varphi(m_1m_2)} \prod_{p \nmid 2m_1m_2} \left(1 - \frac{1}{p^2}\right) \frac{1 + o(1)}{\log R} + O_A\left(\frac{1}{(\log R)^A}\right).$$

In a similar way, but using Lemma 5.4 instead of Lemma 5.3, we deduce from (7.9.7) that

$$(7.9.12) \quad \begin{aligned} \Sigma_2 = & -\frac{2m_1m_2}{\varphi(m_1m_2)} \prod_{p \nmid 2m_1m_2} \left(1 - \frac{1}{p^2}\right) \frac{1 + o(1)}{\log R} \\ & \times \sum_{p \nmid 2m_1m_2} \frac{J(p, s)}{p+1} \left(1 - \frac{1}{p^{1+2s}}\right) \left(1 - \frac{1}{p^{1-2s}}\right) + O_A\left(\frac{1}{(\log R)^A}\right). \end{aligned}$$

We now combine the above pieces to get a new expression for  $\mathcal{B}$ . Take the expression for  $\mathcal{B}$  in (7.9.5) and substitute in (7.9.3) and (7.9.4). Next bring the sum over  $d$  inside to see

that we have  $\Sigma_1$  as in (7.9.6) and  $\Sigma_2$  as in (7.9.7). Then insert (7.9.11) for  $\Sigma_1$  and (7.9.12) for  $\Sigma_2$ . It now follows that

(7.9.13)

$$\begin{aligned} \mathcal{B} = & \frac{X\check{\Phi}(0)}{3\zeta(2)(\sqrt{2}-1)^4} \frac{1+o(1)}{\log R} \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2)=1}} \sum_{p|\ell_1} \frac{b_{m_1} b_{m_2}}{\sqrt{m_1 m_2 \ell_1}} \prod_{p|\ell_1} \left( \frac{p}{p+1} \right) \\ & \times \frac{1}{2\pi i} \oint_{|s|=\frac{1}{\log X}} \sum_{ab=\ell_1} \left( \frac{a}{b} \right)^s \Gamma_1(s) \frac{\Gamma\left(\frac{s}{2} + \frac{1}{4}\right)^2}{\Gamma\left(\frac{1}{4}\right)^2} \left( \frac{4}{\pi} \right)^s \zeta(2s) \zeta(2s+1) \left( 1 - \frac{1}{2^{\frac{1}{2}+s}} \right) \left( 1 - \frac{1}{2^{\frac{1}{2}-s}} \right) \\ & \times \left( \frac{5}{2} - 4^s - 4^{-s} \right) \left\{ \log \left( \frac{X}{2\ell_1} \right) + 2\gamma + \frac{(\check{\Phi})'(0)}{\check{\Phi}(0)} - \frac{\Gamma_1'(s)}{\Gamma_1(s)} - 2\frac{\zeta'(2s)}{\zeta(2s)} + 2\frac{\zeta'(2s+1)}{\zeta(2s+1)} \right. \\ & \left. + \frac{\log 2}{(\sqrt{2}+2^s)(\sqrt{2}+2^{-s})} + \sum_{p \neq 2} \eta_1(p, s) + \sum_{p|\ell_1} \eta_2(p, s) + \sum_{\substack{p|m_1 m_2 \\ p \nmid \ell_1}} \eta_3(p, s) \right\} \frac{ds}{s} + O_A \left( \frac{X}{(\log R)^A} \right), \end{aligned}$$

where

$$\eta_1(p, s) = \frac{2 \log p}{p-1} - \left( \frac{2 \log p}{p} \right) \frac{1 + \frac{2}{p^2} - \frac{1}{p} (p^{2s} + p^{-2s})}{1 + \frac{2}{p} + \frac{1}{p^3} - \frac{1}{p^2} (p^{2s} + p^{-2s})} - \frac{J(p, s)}{p+1} \left( 1 - \frac{1}{p^{1+2s}} \right) \left( 1 - \frac{1}{p^{1-2s}} \right),$$

$$(7.9.14) \quad \eta_2(p, s) = \frac{2 \log p}{p-1} - \eta_1(p, s),$$

and

$$\eta_3(p, s) = \frac{2 \log p}{p-1} - \frac{2 \log p}{p+1} - \eta_1(p, s),$$

with  $J(p, s)$  defined by (7.9.8).

Next, we carry out the summation over  $m_1, m_2$ . Define  $\Upsilon_1, \Upsilon_2, \Upsilon_3$ , and  $\Upsilon_4$  by

$$(7.9.15) \quad \Upsilon_1 = \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2)=1}} \sum_{p|\ell_1} \frac{b_{m_1} b_{m_2}}{\sqrt{m_1 m_2 \ell_1}} \prod_{p|\ell_1} \left( \frac{p}{p+1} \right) \sum_{ab=\ell_1} \left( \frac{a}{b} \right)^s,$$

$$(7.9.16) \quad \Upsilon_2 = - \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2)=1}} \sum_{p|\ell_1} \frac{b_{m_1} b_{m_2}}{\sqrt{m_1 m_2 \ell_1}} \prod_{p|\ell_1} \left( \frac{p}{p+1} \right) \sum_{ab=\ell_1} \left( \frac{a}{b} \right)^s \log \ell_1,$$

$$(7.9.17) \quad \Upsilon_3 = \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2)=1}} \sum_{p|\ell_1} \frac{b_{m_1} b_{m_2}}{\sqrt{m_1 m_2 \ell_1}} \prod_{p|\ell_1} \left( \frac{p}{p+1} \right) \sum_{ab=\ell_1} \left( \frac{a}{b} \right)^s \sum_{p|\ell_1} \eta_2(p, s),$$

and

$$(7.9.18) \quad \Upsilon_4 = \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2)=1}} \sum_{p|\ell_1} \frac{b_{m_1} b_{m_2}}{\sqrt{m_1 m_2 \ell_1}} \prod_{p|\ell_1} \left( \frac{p}{p+1} \right) \sum_{ab=\ell_1} \left( \frac{a}{b} \right)^s \sum_{\substack{p|m_1 m_2 \\ p \nmid \ell_1}} \eta_3(p, s),$$

so that (7.9.13) can be written as

$$\begin{aligned}
 \mathcal{B} = & \frac{X\check{\Phi}(0)}{3\zeta(2)(\sqrt{2}-1)^4} \frac{1+o(1)}{\log R} \frac{1}{2\pi i} \oint_{|s|=\frac{1}{\log X}} \Gamma_1(s) \frac{\Gamma\left(\frac{s}{2}+\frac{1}{4}\right)^2}{\Gamma\left(\frac{1}{4}\right)^2} \left(\frac{4}{\pi}\right)^s \zeta(2s)\zeta(2s+1) \\
 & \times \left(1 - \frac{1}{2^{\frac{1}{2}+s}}\right) \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right) \left(\frac{5}{2} - 4^s - 4^{-s}\right) \left\{ \Upsilon_1 \left( \log\left(\frac{X}{2}\right) + 2\gamma + \frac{(\check{\Phi})'(0)}{\check{\Phi}(0)} \right. \right. \\
 (7.9.19) \quad & \left. \left. - \frac{\Gamma_1'(s)}{\Gamma_1(s)} - 2\frac{\zeta'(2s)}{\zeta(2s)} + 2\frac{\zeta'(2s+1)}{\zeta(2s+1)} + \frac{\log 2}{(\sqrt{2}+2^s)(\sqrt{2}+2^{-s})} + \sum_{p \neq 2} \eta_1(p, s) \right) \right. \\
 & \left. + \Upsilon_2 + \Upsilon_3 + \Upsilon_4 \right\} \frac{ds}{s} + O_A\left(\frac{X}{(\log R)^A}\right).
 \end{aligned}$$

We need to estimate the sums  $\Upsilon_1$ ,  $\Upsilon_2$ ,  $\Upsilon_3$ , and  $\Upsilon_4$  for  $|s| = \frac{1}{\log X}$ .

To estimate  $\Upsilon_1$ , observe that if  $m_1$  and  $m_2$  are square-free then (7.5.17) implies

$$(7.9.20) \quad \ell_1 = \frac{m_1 m_2}{(m_1, m_2)^2}$$

and

$$(7.9.21) \quad \sum_{ab=\ell_1} \left(\frac{a}{b}\right)^s = \prod_{p|\ell_1} (p^s + p^{-s}).$$

From these, the definition (3.5) of  $b_m$ , and the Fourier inversion formula (6.1.2), we deduce from (7.9.15) that

$$\Upsilon_1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(z_1)h(z_2) \sum_{(m_1 m_2, 2)=1} \sum \frac{\mu(m_1)\mu(m_2)(m_1, m_2)}{m_1^{1+\frac{1+iz_1}{\log M}} m_2^{1+\frac{1+iz_2}{\log M}}} \prod_{\substack{p|m_1 m_2 \\ p \nmid (m_1, m_2)}} (p^s + p^{-s}) \left(\frac{p}{p+1}\right) dz_1 dz_2.$$

Thus, writing the sum as an Euler product, we see that

$$\begin{aligned}
 \Upsilon_1 = & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(z_1)h(z_2) \prod_{p>2} \left( 1 - \frac{1}{p^{1+\frac{1+iz_1}{\log M}}} (p^s + p^{-s}) \left(\frac{p}{p+1}\right) \right. \\
 & \left. - \frac{1}{p^{1+\frac{1+iz_2}{\log M}}} (p^s + p^{-s}) \left(\frac{p}{p+1}\right) + \frac{1}{p^{1+\frac{2+iz_1+iz_2}{\log M}}} \right) dz_1 dz_2.
 \end{aligned}$$

We write this as

$$(7.9.22) \quad \Upsilon_1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{h(z_1)h(z_2)\zeta\left(1 + \frac{2+iz_1+iz_2}{\log M}\right) W(s, z_1, z_2, \frac{1}{\log M}) dz_1 dz_2}{\zeta\left(1 + \frac{1+iz_1}{\log M} + s\right) \zeta\left(1 + \frac{1+iz_1}{\log M} - s\right) \zeta\left(1 + \frac{1+iz_2}{\log M} + s\right) \zeta\left(1 + \frac{1+iz_2}{\log M} - s\right)},$$

where  $W(s, z_1, z_2, \frac{1}{\log M})$  is an Euler product that is bounded and holomorphic for  $|s| \leq \varepsilon$  and complex  $z_1, z_2$  with  $|\operatorname{Im}(z_1)|, |\operatorname{Im}(z_2)| \leq \varepsilon \log M$ . Note that this definition of  $W$  implies

$$(7.9.23) \quad W(0, 0, 0, 0) = 8 \prod_{p>2} \left(1 - \frac{4}{p+1} + \frac{1}{p}\right) \left(1 - \frac{1}{p}\right)^{-3} = 6\zeta(2),$$

a fact we use shortly. By (6.1.4), we may truncate the integrals in (7.9.22) to the range  $|z_1|, |z_2| \leq \sqrt{\log M}$ , introducing a negligible error. On this range of  $z_1$  and  $z_2$ , the function  $W$  and the zeta-functions in (7.9.22) may be written as Laurent series. The contributions of the terms other than the first terms of these Laurent expansions are smaller than that of the first terms of the expansion by a factor of  $\log X$ . The first term of the Laurent expansion of  $W$  is given by (7.9.23). We thus arrive at

$$\begin{aligned} \Upsilon_1 = 6\zeta(2) \int \int_{|z_i| \leq \sqrt{\log M}} h(z_1)h(z_2) & \left( \frac{\log M}{2 + iz_1 + iz_2} \right) \left( \frac{1 + iz_1}{\log M} - s \right) \left( \frac{1 + iz_1}{\log M} + s \right) \\ & \times \left( \frac{1 + iz_2}{\log M} + s \right) \left( \frac{1 + iz_2}{\log M} - s \right) dz_1 dz_2 + O\left( \frac{1}{(\log X)^4} \right). \end{aligned}$$

By (6.1.4), we may extend the range of integration to  $\mathbb{R}^2$ , introducing a negligible error. We then apply (7.3.13) to deduce that

$$\begin{aligned} \Upsilon_1 = 6\zeta(2) & \left( \frac{1}{\log^3 M} \int_0^1 H''(t)^2 dt - \frac{2s^2}{\log M} \int_0^1 H(t)H''(t) dt \right. \\ (7.9.24) \quad & \left. + s^4 \log M \int_0^1 H(t)^2 dt \right) + O\left( \frac{1}{(\log X)^4} \right). \end{aligned}$$

Having evaluated  $\Upsilon_1$ , we next estimate  $\Upsilon_2$ . 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, (7.9.16), (7.9.20), (7.9.21), the definition (3.5) of  $b_m$ , and the Fourier inversion formula (6.1.2), it follows that

$$\begin{aligned} \Upsilon_2 = \frac{1}{2\pi i} \oint_{|y|=\frac{1}{2\log X}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(z_1)h(z_2) & \sum_{(m_1 m_2, 2)=1} \sum \frac{\mu(m_1)\mu(m_2)(m_1, m_2)^{1+2y}}{m_1^{1+\frac{1+iz_1}{\log M}+y} m_2^{1+\frac{1+iz_2}{\log M}+y}} \\ & \times \prod_{\substack{p|m_1 m_2 \\ p \nmid (m_1, m_2)}} (p^s + p^{-s}) \left( \frac{p}{p+1} \right) dz_1 dz_2 \frac{dy}{y^2}. \end{aligned}$$

We express the sum as an Euler product to see that

$$\begin{aligned} \Upsilon_2 = \frac{1}{2\pi i} \oint_{|y|=\frac{1}{2\log X}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(z_1)h(z_2) & \prod_{p>2} \left( 1 - \frac{1}{p^{1+\frac{1+iz_1}{\log M}+y}} (p^s + p^{-s}) \left( \frac{p}{p+1} \right) \right. \\ & \left. - \frac{1}{p^{1+\frac{1+iz_2}{\log M}+y}} (p^s + p^{-s}) \left( \frac{p}{p+1} \right) + \frac{1}{p^{1+\frac{2+iz_1+iz_2}{\log M}}} \right) dz_1 dz_2 \frac{dy}{y^2}. \end{aligned}$$

Write this as

$$\begin{aligned} \Upsilon_2 = \frac{1}{2\pi i} \oint_{|y|=\frac{1}{2\log X}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(z_1)h(z_2) \zeta\left(1 + \frac{2+iz_1+iz_2}{\log M}\right) V(s, z_1, z_2, \frac{1}{\log M}, y) \\ \times \zeta^{-1}\left(1 + \frac{1+iz_1}{\log M} + y + s\right) \zeta^{-1}\left(1 + \frac{1+iz_1}{\log M} + y - s\right) \\ \times \zeta^{-1}\left(1 + \frac{1+iz_2}{\log M} + y + s\right) \zeta^{-1}\left(1 + \frac{1+iz_2}{\log M} + y - s\right) dz_1 dz_2 \frac{dy}{y^2}, \end{aligned}$$

where  $V(s, z_1, z_2, \frac{1}{\log M}, y)$  is an Euler product that is bounded and holomorphic for  $|s|, |y| \leq \varepsilon$  and complex  $z_1, z_2$  with  $|\operatorname{Im}(z_1)|, |\operatorname{Im}(z_2)| \leq \varepsilon \log M$ . This definition of  $V$  implies that  $V(0, 0, 0, 0, 0) = 6\zeta(2)$ . As in our treatment of  $\Upsilon_1$ , we use (6.1.4) 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

$$\begin{aligned} \Upsilon_2 = \frac{6\zeta(2)}{2\pi i} \oint_{|y|=\frac{1}{2\log X}} \int \int_{|z_i| \leq \sqrt{\log M}} h(z_1)h(z_2) \left(\frac{\log M}{2 + iz_1 + iz_2}\right) \left(\frac{1 + iz_1}{\log M} + y - s\right) \\ \times \left(\frac{1 + iz_1}{\log M} + y + s\right) \left(\frac{1 + iz_2}{\log M} + y + s\right) \left(\frac{1 + iz_2}{\log M} + y - s\right) dz_1 dz_2 \frac{dy}{y^2} + O\left(\frac{1}{(\log X)^3}\right). \end{aligned}$$

We carry out the integration over  $y$  by applying the formula (6.1.6) with  $n = 2$  and deduce that

$$\begin{aligned} \Upsilon_2 = 6\zeta(2) \int \int_{|z_i| \leq \sqrt{\log M}} h(z_1)h(z_2) \left(\frac{\log M}{2 + iz_1 + iz_2}\right) \\ \times \left\{ \left(\frac{1 + iz_1}{\log M} + s\right) \left(\frac{(1 + iz_2)^2}{(\log M)^2} - s^2\right) + \left(\frac{1 + iz_1}{\log M} - s\right) \left(\frac{(1 + iz_2)^2}{(\log M)^2} - s^2\right) \right. \\ \left. + \left(\frac{1 + iz_2}{\log M} + s\right) \left(\frac{(1 + iz_1)^2}{(\log M)^2} - s^2\right) + \left(\frac{1 + iz_2}{\log M} - s\right) \left(\frac{(1 + iz_1)^2}{(\log M)^2} - s^2\right) \right\} dz_1 dz_2 \\ + O\left(\frac{1}{(\log X)^3}\right). \end{aligned}$$

We extend the integral and apply (7.3.13). After simplifying, we arrive at

$$(7.9.25) \quad \Upsilon_2 = 6\zeta(2) \left( -\frac{4}{\log^2 M} \int_0^1 H'(t)H''(t) dt + 4s^2 \int_0^1 H(t)H'(t) dt \right) + O\left(\frac{1}{(\log X)^3}\right).$$

We next estimate  $\Upsilon_3$  defined by (7.9.17). We interchange the order of summation over  $m_1, m_2$  and over  $p$ . From (7.9.20), 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

$$\Upsilon_3 = 2 \sum_{2 < q \leq M} \eta_2(q, s) \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2) = 1 \\ q \nmid m_1, q \nmid m_2}} \frac{b_{m_1} b_{m_2}}{\sqrt{m_1 m_2} \ell_1} \prod_{p|\ell_1} \left(\frac{p}{p+1}\right) \sum_{ab=\ell_1} \left(\frac{a}{b}\right)^s.$$

From this, the definition (3.5) of  $b_m$ , (7.9.20), and (7.9.21), it follows that

$$\begin{aligned} \Upsilon_3 = 2 \sum_{2 < q \leq M} \eta_2(q, s) \sum_{\substack{m_1, m_2 \leq M \\ (m_1 m_2, 2) = 1 \\ q | m_1, q \nmid m_2}} \frac{\mu(m_1) \mu(m_2)}{[m_1, m_2]} \prod_{\substack{p | m_1 m_2 \\ p \nmid (m_1, m_2)}} \left( \frac{p}{p+1} \right) (p^s + p^{-s}) \\ \times H \left( \frac{\log m_1}{\log M} \right) H \left( \frac{\log m_2}{\log M} \right). \end{aligned}$$

We relabel  $m_1$  as  $qm_1$  to write this as

$$\begin{aligned} \Upsilon_3 = -2 \sum_{2 < q \leq M} \frac{\eta_2(q, s)}{q+1} (q^s + q^{-s}) \sum_{m_1 \leq \frac{M}{q}} \sum_{\substack{m_2 \leq M \\ (m_1 m_2, 2q) = 1}} \frac{\mu(m_1) \mu(m_2)}{[m_1, m_2]} \\ \times \prod_{\substack{p | m_1 m_2 \\ p \nmid (m_1, m_2)}} \left( \frac{p}{p+1} \right) (p^s + p^{-s}) H \left( \frac{\log qm_1}{\log M} \right) H \left( \frac{\log m_2}{\log M} \right). \end{aligned}$$

We insert the Fourier inversion formula (6.1.2), interchange the order of summation, and then write the  $m_1, m_2$ -sum as an Euler product to deduce that

$$\begin{aligned} \Upsilon_3 = -2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{2 < q \leq M} \frac{\eta_2(q, s) (q^s + q^{-s})}{(q+1) q^{\frac{1+iz_1}{\log M}}} h(z_1) h(z_2) \prod_{p \nmid 2q} \left( 1 - \frac{1}{p^{1+\frac{1+iz_1}{\log M}}} (p^s + p^{-s}) \left( \frac{p}{p+1} \right) \right. \\ \left. - \frac{1}{p^{1+\frac{1+iz_2}{\log M}}} (p^s + p^{-s}) \left( \frac{p}{p+1} \right) + \frac{1}{p^{1+\frac{2+iz_1+iz_2}{\log M}}} \right) dz_1 dz_2. \end{aligned}$$

We may express the Euler product in terms of zeta-functions to write

$$\begin{aligned} (7.9.26) \quad \Upsilon_3 = -2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{2 < q \leq M} \frac{\eta_2(q, s) (q^s + q^{-s})}{(q+1) q^{\frac{1+iz_1}{\log M}}} h(z_1) h(z_2) \zeta \left( 1 + \frac{2+iz_1+iz_2}{\log M} \right) \zeta^{-1} \left( 1 + \frac{1+iz_1}{\log M} + s \right) \\ \times \zeta^{-1} \left( 1 + \frac{1+iz_1}{\log M} - s \right) \zeta^{-1} \left( 1 + \frac{1+iz_2}{\log M} + s \right) \zeta^{-1} \left( 1 + \frac{1+iz_2}{\log M} - s \right) U_q(s, z_1, z_2, \frac{1}{\log M}) dz_1 dz_2, \end{aligned}$$

where  $U_q(s, z_1, z_2, \frac{1}{\log M})$  is an Euler product that is uniformly bounded for  $2 < q \leq M$  prime,  $|s| \leq \varepsilon$ , and real  $z_1, z_2$ . Using (6.1.4), we may truncate the integrals to the range  $|z_1|, |z_2| \leq \sqrt{\log M}$  and introduce only a negligible error. In this range, and for  $|s| = \frac{1}{\log X}$ , the quotient of zeta-functions in (7.9.26) is  $\ll (1 + |z_1|^2)(1 + |z_2|^2)(\log X)^{-3}$ . Moreover, (7.9.14) implies  $\eta_2(q, s) \ll q^{-1+\varepsilon}$  for  $2 < q \leq M$  and  $|s| = \frac{1}{\log X}$ . It thus follows that

$$(7.9.27) \quad \Upsilon_3 \ll \frac{1}{(\log X)^3}.$$

A similar argument applies to  $\Upsilon_4$  defined by (7.9.18), except we use the fact that, for a prime  $q$ ,  $q | m_1 m_2$  and  $q \nmid \ell_1$  both hold if and only if  $q$  divides both  $m_1$  and  $m_2$ , by (7.9.20). This leads to

$$(7.9.28) \quad \Upsilon_4 \ll \frac{1}{(\log X)^3}.$$



It now follows from (7.9.19), (7.9.24), (7.9.25), (7.9.27), and (7.9.28) that

$$\begin{aligned} \mathcal{B} = & \frac{2X\check{\Phi}(0)}{(\sqrt{2}-1)^4} \frac{1+o(1)}{\log R} \frac{1}{2\pi i} \oint_{|s|=\frac{1}{\log X}} \Gamma_1(s) \frac{\Gamma\left(\frac{s}{2} + \frac{1}{4}\right)^2}{\Gamma\left(\frac{1}{4}\right)^2} \left(\frac{4}{\pi}\right)^s \zeta(2s)\zeta(2s+1) \\ & \times \left(1 - \frac{1}{2^{\frac{1}{2}+s}}\right) \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right) \left(\frac{5}{2} - 4^s - 4^{-s}\right) \left\{ \left(\log\left(\frac{X}{2}\right) + 2\gamma + \frac{(\check{\Phi})'(0)}{\check{\Phi}(0)}\right. \right. \\ & \left. \left. - \frac{\Gamma_1'(s)}{\Gamma_1(s)} - 2\frac{\zeta'(2s)}{\zeta(2s)} + 2\frac{\zeta'(2s+1)}{\zeta(2s+1)} + \frac{\log 2}{(\sqrt{2}+2^s)(\sqrt{2}+2^{-s})} + \sum_{p \neq 2} \eta_1(p, s) \right) \right. \\ & \times \left( \frac{1}{\log^3 M} \int_0^1 H''(t)^2 dt - \frac{2s^2}{\log M} \int_0^1 H(t)H''(t) dt + s^4 \log M \int_0^1 H(t)^2 dt \right) \\ & \left. - \frac{4}{\log^2 M} \int_0^1 H'(t)H''(t) dt + 4s^2 \int_0^1 H(t)H'(t) dt \right\} \frac{ds}{s} + O\left(\frac{X}{(\log X)^2}\right). \end{aligned}$$

Evaluating the  $s$ -integral as a residue, we deduce that

$$\begin{aligned} \mathcal{B} = & \frac{X\check{\Phi}(0)}{4\left(1 - \frac{1}{\sqrt{2}}\right)^2} \frac{1+o(1)}{\log R} \left\{ \frac{\log X}{2\log M} \int_0^1 H(t)H''(t) dt - \int_0^1 H(t)H'(t) dt \right\} \\ & + O\left(X(\log X)^{-2}\right). \end{aligned}$$

From this, (2.2), (7.3.14), (7.2.8), (7.2), and (7.1.1), it now follows that

$$\begin{aligned} S^+ = & \frac{X}{8\left(1 - \frac{1}{\sqrt{2}}\right)^2} \frac{1+o(1)}{\log R} \left\{ \frac{1}{24} \left(\frac{\log X}{\log M}\right)^3 \int_0^1 H''(t)^2 dt \right. \\ & - \frac{1}{2} \left(\frac{\log X}{\log M}\right)^2 \int_0^1 H'(t)H''(t) dt + \frac{\log X}{\log M} \int_0^1 H(t)H''(t) dt + \frac{\log X}{\log M} \int_0^1 H'(t)^2 dt \\ & \left. - 2 \int_0^1 H(t)H'(t) dt \right\} + O\left(\frac{X}{(\log X)^2} + \frac{X^{1+\varepsilon}}{Y} + X^{\frac{1}{2}+\varepsilon}M\right). \end{aligned}$$

The error terms are acceptable by the choices in Subsection 7.7, and this yields Proposition 7.1.

## 8. CHOOSING THE MOLLIFIER: FINISHING THE PROOF OF THEOREM 1.1

In this section we complete the proof of Theorem 1.1 by making an optimal choice for the smooth function  $H(x)$  (see (3.3), (3.5)).

By (3.2), Proposition 6.1, and Proposition 7.1, one derives the inequality

$$(8.1) \quad \sum_{\substack{p \equiv 1 \pmod{8} \\ L(\frac{1}{2}, \chi_p) \neq 0}} (\log p) \Phi\left(\frac{p}{X}\right) \geq \frac{X}{(1+\delta_0)8} \cdot \vartheta \frac{(H(0) - \frac{1}{2\theta}H'(0))^2}{\mathfrak{I}}$$

for  $\theta + 2\vartheta < \frac{1}{2}$  and  $X \geq X_0(\delta_0, \theta, \vartheta)$ , 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 (8.1) is an increasing function of  $\vartheta$ , and so  $\vartheta$  should be as large as possible. The hypotheses of Proposition 7.1 allow  $\vartheta = \frac{1}{2}(\frac{1}{2} - \theta) - \varepsilon$ , and therefore

$$(8.2) \quad \sum_{\substack{X/2 < p \leq X \\ p \equiv 1 \pmod{8} \\ L(\frac{1}{2}, \chi_p) \neq 0}} 1 \geq \frac{X}{(1 + 2\delta_0)8 \log X} \cdot \varrho,$$

where

$$\varrho := \frac{1}{2} \left( \frac{1}{2} - \theta \right) \frac{(H(0) - \frac{1}{2\theta} H'(0))^2}{\mathfrak{J}}.$$

We seek a choice of  $H(x)$  that maximizes  $\varrho$ .

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

$$\begin{aligned} \int_0^1 H(x) H'(x) dx &= -\frac{1}{2} A^2, \\ \int_0^1 H(x) H''(x) dx &= AB - \int_0^1 H'(x)^2 dx, \\ \int_0^1 H'(x) H''(x) dx &= -\frac{1}{2} B^2, \end{aligned}$$

we have, by the definition of  $\mathfrak{J}$  in Proposition 7.1, that

$$\mathfrak{J} = \left( A + \frac{1}{2\theta} B \right)^2 + \frac{1}{24\theta^3} \int_0^1 H''(x)^2 dx.$$

We choose  $H(x)$  such that on  $[0, 1]$  it is a smooth approximation to the optimal function  $H_*(x)$  that minimizes the integral

$$(8.3) \quad \int_0^1 H''_*(x)^2 dx$$

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

$$(1 + \delta_0) \int_0^1 H''_*(x)^2 dx \geq \int_0^1 H''(x)^2 dx.$$

By the Euler-Lagrange equation, we find that an  $H_*(x)$  which minimizes (8.3) 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

$$\int_0^1 H_*''(x)^2 dx = 3A^2 + (2B - 3A)^2,$$

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

$$(8.4) \quad \varrho \geq \frac{1 - O(\delta_0)}{2} \left( \frac{1}{2} - \theta \right) \frac{2\theta(3 + 6\theta + 4\theta^2)}{(1 + 2\theta)^3}.$$

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  $\varrho$  is invariant under multiplication of  $H$  by scalars, we arrive at the convenient expression

$$(8.5) \quad H_*(x) = (1 - x)^2 \left( 2 + \frac{3}{2\theta} + x \right).$$

If we set  $x = \frac{\log m}{\log M}$  in (8.5), 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 [40, 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 (8.2) and (8.4), we obtain

$$(8.6) \quad \sum_{\substack{X/2 < p \leq X \\ p \equiv 1 \pmod{8} \\ L(\frac{1}{2}, \chi_p) \neq 0}} 1 \geq \frac{X}{(1 + O(\delta_0))8 \log X} \cdot \rho(\theta).$$

The maximum of  $\rho(\theta)$  on  $(0, \frac{1}{2})$  occurs at the unique positive root  $\theta_0$  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

$$(8.7) \quad \rho(\theta_0) = 0.09645 \dots$$

We then choose  $\theta = \theta_0$ . Since

$$\sum_{\substack{X/2 < p \leq X \\ p \equiv 1 \pmod{8}}} 1 = (1 + o(1)) \frac{X}{8(\log X)},$$

we deduce Theorem 1.1 from (8.6) and (8.7) upon summing over dyadic intervals.

## 9. THE SECOND MOMENT OF $L(\frac{1}{2}, \chi_p)$

In this section we prove Theorems 1.2 and 1.3. We first consider separately the upper and lower bounds for Theorem 1.2.

**9.1. The upper bound in Theorem 1.2.** We define

$$(9.1.1) \quad M_2 := \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) L\left(\frac{1}{2}, \chi_p\right)^2.$$

In this subsection we prove

$$(9.1.2) \quad M_2 \leq (4\mathfrak{c} + o(1)) \frac{X}{8} (\log X)^3.$$

The upper bound of Theorem 1.2 then follows from (9.1.2) upon summation over dyadic intervals.

The proof of (9.1.2) follows the lines of the proof of Proposition 7.1, 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 (7.1.1).

We follow the argument of Section 7 up to (7.2.8), obtaining

$$S_N^+ = \mathcal{T}_0 + \mathcal{B}.$$

Since we have no mollifier here, we find that (7.3.6) in this case is

$$\mathcal{T}_0 = \frac{2X}{(\sqrt{2}-1)^4} \frac{1+o(1)}{\log R} \sum_{\substack{\nu=1 \\ (\nu,2)=1 \\ \nu=\square}}^{\infty} \frac{d_2(\nu)}{\sqrt{\nu}} \hat{F}_\nu(0) + O_A\left(\frac{X}{(\log R)^A}\right) + O\left(\frac{X^{1+\varepsilon}}{Y}\right).$$

We insert into this the definitions (7.2.4) and (4.1) of  $F_\nu$  and  $\omega_2$ , interchange the order of summation, and then write the sum over  $\nu$  as an Euler product. The result is

$$\begin{aligned} \mathcal{T}_0 &= \frac{2X}{(\sqrt{2}-1)^4} \frac{1+o(1)}{\log R} \frac{1}{2\pi i} \int_{(c)} \frac{\Gamma(\frac{s}{2} + \frac{1}{4})^2}{\Gamma(\frac{1}{4})^2} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right)^2 \left(\frac{X}{\pi}\right)^s \check{\Phi}(s) \left(1 - \frac{1}{2^{1+2s}}\right)^3 \\ &\quad \times \zeta(1+2s)^3 \left(1 - \frac{1}{2^{2+4s}}\right)^{-1} \zeta(2+4s)^{-1} \frac{ds}{s} + O_A\left(\frac{X}{(\log R)^A}\right) + O\left(\frac{X^{1+\varepsilon}}{Y}\right). \end{aligned}$$

As before, we truncate the integral to the range  $|\operatorname{Im}(s)| \leq (\log X)^2$ , and then deform the path of integration to the path made up of the line segments  $L_1, L_2, L_3$  defined above (7.3.12) to

see that the main contribution arises from the residue of the integrand at  $s = 0$ . We evaluate the residue using (6.1.6) 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 (2.2) and the definition of  $\mathfrak{c}$ , we deduce that

$$(9.1.3) \quad \mathcal{T}_0 \leq (\mathfrak{c} + \varepsilon) \frac{X (\log X)^3}{8 \log R} + O \left( X \log X + \frac{X^{1+\varepsilon}}{Y} \right).$$

Moreover, we see from (7.9.13) that if  $M = 1$  and  $b_1 = 1$ , then

$$(9.1.4) \quad \mathcal{B} \ll X \frac{\log X}{\log R} \ll X$$

since we may deform the path of integration in (7.9.13) to a circle  $|s| = \varepsilon$ . The condition  $\theta + 2\vartheta < \frac{1}{2}$  in Subsection 7.7 with  $\theta = 0$  allows us to take  $\vartheta = \frac{1}{4} - \varepsilon$  in (9.1.3). We then set  $Y = X^\delta$ , for some small, fixed  $\delta > 0$ . We see that the upper bound (9.1.2) then follows from (9.1.3) and (9.1.4) after sending  $\varepsilon$  to zero sufficiently slowly.

**9.2. The lower bound in Theorem 1.2.** Recall the definition (9.1.1) of  $M_2$ . Our goal is to prove the following result.

**Proposition 9.1.** *As  $X \rightarrow \infty$ , 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 1.2.

The lower bound for Theorem 1.2 easily follows from Proposition 9.1 by summing over dyadic intervals.

The main idea in the proof of Proposition 9.1 is a standard one. For any Dirichlet polynomial  $A(p)$ , the Cauchy-Schwarz inequality implies

$$(9.2.1) \quad M_2 \geq \frac{\left( \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi \left( \frac{p}{X} \right) L \left( \frac{1}{2}, \chi_p \right) A(p) \right)^2}{\sum_{p \equiv 1 \pmod{8}} (\log p) \Phi \left( \frac{p}{X} \right) A(p)^2}.$$

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 4.2. For a positive real number  $\alpha$ , we define

$$(9.2.2) \quad A_\alpha(p) := \frac{2}{\left( 1 - \frac{1}{\sqrt{2}} \right)^2} \sum_{n \text{ odd}} \frac{\chi_p(n)}{\sqrt{n}} \omega_1 \left( n \sqrt{\frac{\pi}{p^\alpha}} \right).$$

With  $\varepsilon_0 > 0$  small and fixed, we then choose  $A(p)$  in (9.2.1) to be

$$(9.2.3) \quad A(p) := A_{1-\varepsilon_0}(p).$$

Observe that taking  $\alpha = 1$  in (9.2.2) yields

$$(9.2.4) \quad A_1(p) = L \left( \frac{1}{2}, \chi_p \right).$$

**Proposition 9.2.** *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 9.1 assuming Proposition 9.2.* By (9.2.1), (9.2.3), and (9.2.4), we have

$$M_2 \geq \frac{M_{1-\varepsilon_0, 1}^2}{M_{1-\varepsilon_0, 1-\varepsilon_0}}.$$

We apply Proposition 9.2 to obtain

$$M_2 \geq \frac{1}{2}(\mathfrak{c} + O(\varepsilon_0)) \frac{X}{4} (\log X)^3.$$

Proposition 9.1 follows upon letting  $\varepsilon_0 = \varepsilon_0(X)$  go to zero sufficiently slowly as  $X \rightarrow \infty$ .  $\square$

We devote the rest of this subsection to the proof of Proposition 9.2.

*Proof of Proposition 9.2.* By definition,

$$M_{\alpha_1, \alpha_2} = \frac{4}{(1 - \frac{1}{\sqrt{2}})^4} \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) \sum_{m, n \text{ odd}} \sum_{\substack{\chi_p(mn) \\ \sqrt{mn}}} \omega_1\left(m \sqrt{\frac{\pi}{p^{\alpha_1}}}\right) \omega_1\left(n \sqrt{\frac{\pi}{p^{\alpha_2}}}\right).$$

Let  $M_{\neq}$  denote the contribution to  $M_{\alpha_1, \alpha_2}$  from  $mn \neq \square$ . An application of Lemma 6.1 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

$$M_{\alpha_1, \alpha_2} = \frac{4}{(1 - \frac{1}{\sqrt{2}})^4} \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) \sum_{\substack{(mn, 2p)=1 \\ mn=\square}} \frac{1}{\sqrt{mn}} \omega_1\left(m \sqrt{\frac{\pi}{p^{\alpha_1}}}\right) \omega_1\left(n \sqrt{\frac{\pi}{p^{\alpha_2}}}\right) + O(X).$$

We use Lemma 4.1 to remove the condition  $(mn, p) = 1$  at the cost of a negligible error. We then open  $\omega_1$  using its definition as an integral, and interchange the order of summation and integration. After some simplification, we arrive at

$$M_{\alpha_1, \alpha_2} = \frac{4}{(1 - \frac{1}{\sqrt{2}})^4} \frac{1}{(2\pi i)^2} \int_{(c_1)} \int_{(c_2)} K(s_1, s_2) \left(\frac{X^{\alpha_1}}{\pi}\right)^{\frac{s_1}{2}} \left(\frac{X^{\alpha_2}}{\pi}\right)^{\frac{s_2}{2}} \zeta(1 + 2s_1) \zeta(1 + 2s_2) \\ \times \zeta(1 + s_1 + s_2) \left( \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) \left(\frac{p}{X}\right)^{\frac{\alpha_1 s_1 + \alpha_2 s_2}{2}} \right) \frac{ds_1 ds_2}{s_1 s_2} + O(X),$$

where  $c_\ell = \text{Re}(s_\ell)$  is a positive real number, and

$$K(s_1, s_2) = \zeta^{-1}(2 + 2s_1 + 2s_2) \left(1 + \frac{1}{2^{1+s_1+s_2}}\right)^{-1} \\ \times \prod_{\ell=1}^2 \frac{\Gamma(\frac{s_\ell}{2} + \frac{1}{4})}{\Gamma(\frac{1}{4})} \left(1 - \frac{1}{2^{\frac{1}{2}-s_\ell}}\right) \left(1 - \frac{1}{2^{1+2s_\ell}}\right).$$

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  $|\text{Im}(s_\ell)| \leq (\log X)^2$  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 over  $p$  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  $M_{\alpha_1, \alpha_2}$ . We then remove the condition on  $\text{Im}(s_\ell)$  by the same means we installed it and obtain

$$\begin{aligned} M_{\alpha_1, \alpha_2} &= \frac{4}{(1 - \frac{1}{\sqrt{2}})^4} \frac{X}{4} \int_0^\infty \Phi(x) \frac{1}{(2\pi i)^2} \int_{(c_1)} \int_{(c_2)} K(s_1, s_2) \left(\frac{(xX)^{\alpha_1}}{\pi}\right)^{\frac{s_1}{2}} \left(\frac{(xX)^{\alpha_2}}{\pi}\right)^{\frac{s_2}{2}} \\ &\quad \times \zeta(1 + 2s_1) \zeta(1 + 2s_2) \zeta(1 + s_2 + s_2) \frac{ds_1 ds_2}{s_1 s_2} dx + O(X). \end{aligned}$$

We wish to separate the variables  $s_1$  and  $s_2$ . Since  $c_\ell > 0$ , we may expand  $\zeta(1 + s_1 + s_2)$  as an absolutely convergent Dirichlet series. Interchanging the order of summation and integration, we obtain

$$\begin{aligned} M_{\alpha_1, \alpha_2} &= \frac{4}{(1 - \frac{1}{\sqrt{2}})^4} \frac{X}{4} \int_0^\infty \Phi(x) \sum_{n=1}^\infty \frac{1}{n} \frac{1}{(2\pi i)^2} \int_{(c_1)} \int_{(c_2)} K(s_1, s_2) \left(\frac{(xX)^{\alpha_1}}{\pi n^2}\right)^{\frac{s_1}{2}} \left(\frac{(xX)^{\alpha_2}}{\pi n^2}\right)^{\frac{s_2}{2}} \\ &\quad \times \zeta(1 + 2s_1) \zeta(1 + 2s_2) \frac{ds_1 ds_2}{s_1 s_2} dx + O(X). \end{aligned}$$

To truncate the summation over  $n$ , first we move the lines 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  $\text{Re}(s_2)$  to  $c_2 = \frac{1}{\log X}$  and estimate trivially, getting an error term of  $O(X(\log X)^2)$ . If  $n \ll X^{\frac{\alpha_1}{2}}$ , we then move  $c_1$  to  $\frac{1}{\log X}$  as well, obtaining

$$\begin{aligned} M_{\alpha_1, \alpha_2} &= \frac{4}{(1 - \frac{1}{\sqrt{2}})^4} \frac{X}{4} \int_0^\infty \Phi(x) \sum_{n \leq \sqrt{(xX)^{\alpha_1}/\pi}} \frac{1}{n} \frac{1}{(2\pi i)^2} \int_{(\frac{1}{\log X})} \int_{(\frac{1}{\log X})} K(s_1, s_2) \\ &\quad \times \left(\frac{(xX)^{\alpha_1}}{\pi n^2}\right)^{\frac{s_1}{2}} \left(\frac{(xX)^{\alpha_2}}{\pi n^2}\right)^{\frac{s_2}{2}} \zeta(1 + 2s_1) \zeta(1 + 2s_2) \frac{ds_1 ds_2}{s_1 s_2} dx + O(X(\log X)^2). \end{aligned}$$

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  $\text{Re}(s_1) = \text{Re}(s_2) = -\delta$ , 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  $\text{Re}(s_\ell) = -\delta$  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

$$\begin{aligned} M_{\alpha_1, \alpha_2} &= \frac{4}{(1 - \frac{1}{\sqrt{2}})^4} \frac{X}{4} \int_0^\infty \Phi(x) \sum_{n \leq \sqrt{(xX)^{\alpha_1}/\pi}} \frac{1}{n} \frac{1}{(2\pi i)^2} \oint_{|s_\ell|=(\log X)^{-1}} \oint K(s_1, s_2) \\ &\quad \times \left(\frac{(xX)^{\alpha_1}}{\pi n^2}\right)^{\frac{s_1}{2}} \left(\frac{(xX)^{\alpha_2}}{\pi n^2}\right)^{\frac{s_2}{2}} \zeta(1 + 2s_1) \zeta(1 + 2s_2) \frac{ds_1 ds_2}{s_1 s_2} dx + O(X(\log X)^2). \end{aligned}$$

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

$$\begin{aligned} M_{\alpha_1, \alpha_2} &= \frac{2}{3\zeta(2)(1 - \frac{1}{\sqrt{2}})^2} \frac{X}{4} \int_0^\infty \Phi(x) \sum_{n \leq \sqrt{(xX)^{\alpha_1}/\pi}} \frac{1}{n} \frac{1}{(2\pi i)^2} \oint_{|s_\ell|=(\log X)^{-1}} \oint \left( \frac{(xX)^{\alpha_1}}{\pi n^2} \right)^{\frac{s_1}{2}} \\ &\quad \times \left( \frac{(xX)^{\alpha_2}}{\pi n^2} \right)^{\frac{s_2}{2}} \zeta(1 + 2s_1) \zeta(1 + 2s_2) \frac{ds_1 ds_2}{s_1 s_2} dx + O(X(\log X)^2). \end{aligned}$$

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

$$\begin{aligned} M_{\alpha_1, \alpha_2} &= \frac{1}{6\zeta(2)(1 - \frac{1}{\sqrt{2}})^2} \frac{X}{4} \int_0^\infty \Phi(x) \sum_{n \leq \sqrt{(xX)^{\alpha_1}/\pi}} \frac{1}{n} \log \left( \frac{1}{n} \sqrt{\frac{(xX)^{\alpha_1}}{\pi}} \right) \log \left( \frac{1}{n} \sqrt{\frac{(xX)^{\alpha_2}}{\pi}} \right) \\ &\quad + O(X(\log X)^2). \end{aligned}$$

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 using (2.2) we arrive at

$$M_{\alpha_1, \alpha_2} = \frac{1}{2} (\mathfrak{c} + O(\varepsilon_0)) \frac{X}{4} (\log X)^3.$$

□

**9.3. Proof of Theorem 1.3.** We turn now to the proof of Theorem 1.3. Throughout this subsection we set  $\eta := 100 \log \log X / \log X$ . Recalling the definition (9.2.2) of  $A_\alpha(p)$ , we then have

$$L\left(\frac{1}{2}, \chi_p\right) = A_{1-\eta}(p) + B(p),$$

say. Thus

$$(9.3.1) \quad M_2 = \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) \{A_{1-\eta}(p)^2 + O(|A_{1-\eta}(p)B(p)| + |B(p)|^2)\}.$$

We shall prove, on GRH, that

$$(9.3.2) \quad \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) A_{1-\eta}(p)^2 = \mathfrak{c} \frac{X}{8} (\log X)^3 + O(X(\log X)^{2+\varepsilon})$$



and

$$(9.3.3) \quad \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) |B(p)|^2 \ll X(\log X)^{5/2}.$$

Theorem 1.3 then follows from (9.3.1), (9.3.2), and (9.3.3) after applying Cauchy-Schwarz and summing over dyadic ranges.

We may easily prove (9.3.2), since the treatment is substantially similar to the proof of Proposition 9.2. By the definition (9.2.2) of  $A_{1-\eta}(p)$ , the left-hand side of (9.3.2) is

$$\frac{4}{(1 - \frac{1}{\sqrt{2}})^4} \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) \sum_{m,n} \sum_{\text{odd}} \frac{\chi_p(mn)}{\sqrt{mn}} \omega_1\left(m\sqrt{\frac{\pi}{p^{1-\eta}}}\right) \omega_1\left(n\sqrt{\frac{\pi}{p^{1-\eta}}}\right).$$

We argue as in Proposition 9.2 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 (see Chapter 20 of [12]) implies that the contribution to (9.3.2) from  $mn \neq \square$  is  $O(X/\log X)$ , say.

**Lemma 9.1.** *Let  $\chi$  be a non-principal Dirichlet character modulo  $q$ . Let  $\chi^*$  be the primitive character inducing  $\chi$ , 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 (9.3.3) is more subtle. Here the method of proof is that of Soundararajan and Young [42]. As the arguments are very similar, our exposition will be sparse, and we refer the reader to [42] for more details. We perform some initial manipulations, and then we state the main proposition that will yield (9.3.3).

By the definition of  $B(p)$ , Lemma 4.2, and the definition (4.1) of  $\omega_1$ , we have

$$(9.3.4) \quad B(p) = \frac{1}{2\pi i} \int_{(c)} g(s) L\left(\frac{1}{2} + s, \chi_p\right) \frac{p^{s/2} - p^{(1-\eta)s/2}}{s} ds,$$

where  $c > 0$  and

$$g(s) = \frac{2}{(1 - \frac{1}{\sqrt{2}})^2} \frac{\Gamma(\frac{s}{2} + \frac{1}{4})}{\Gamma(\frac{1}{4})} \left(1 - \frac{1}{2^{\frac{1}{2}-s}}\right) \left(1 - \frac{1}{2^{\frac{1}{2}+s}}\right) \pi^{-s/2}.$$

The function  $(p^{s/2} - p^{(1-\eta)s/2})/s$  is entire, so we may move the line of integration in (9.3.4) to  $\text{Re}(s) = 0$ . On the line  $\text{Re}(s) = 0$  we have the bound  $|(p^{s/2} - p^{(1-\eta)s/2})/s| \ll \log \log X$ , and hence the left side of (9.3.3) is

$$(9.3.5) \quad \ll (\log X)^{1+\varepsilon} \int_{\mathbb{R}} \int_{\mathbb{R}} |g(it_1)g(it_2)| \sum_{\substack{p \leq X \\ p \equiv 1 \pmod{8}}} |L(\tfrac{1}{2} + it_1, \chi_p) L(\tfrac{1}{2} + it_2, \chi_p)| dt_1 dt_2.$$

To state the proposition we need, we first establish some notation, following [42, Section 6]. Given  $x \geq 10$ , say, and a complex number  $z$ , we define

$$\mathcal{L}(z, x) = \begin{cases} \log \log x, & |z| \leq (\log x)^{-1}, \\ -\log |z|, & (\log x)^{-1} \leq |z| \leq 1, \\ 0, & |z| \geq 1. \end{cases}$$

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

$$\begin{aligned} \mathcal{V}(z_1, z_2, x) = \frac{1}{2} \Big\{ & \mathcal{L}(2z_1, x) + \mathcal{L}(2z_2, x) + \mathcal{L}(2\operatorname{Re}(z_1), x) + \mathcal{L}(2\operatorname{Re}(z_2), x) \\ & + 2\mathcal{L}(z_1 + z_2, x) + 2\mathcal{L}(z_1 + \overline{z_2}, x) \Big\}. \end{aligned}$$

It is helpful to know that for the values of  $z_1$  and  $z_2$  we consider, we have  $\log \log X \leq \mathcal{V}(z_1, z_2, X) \leq 4 \log \log X$ .

The following result, an analogue of [42, Theorem 6.1], is the key input we need.

**Proposition 9.3.** *Let  $X$  be large, and let  $z_1$  and  $z_2$  be complex numbers with  $0 \leq \operatorname{Re}(z_i) \leq \frac{1}{\log X}$  and  $|z_i| \leq X$ . 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 real  $r > 0$  and any  $\varepsilon > 0$  we have*

$$\begin{aligned} & \sum_{\substack{p \leq X \\ p \equiv 1 \pmod{8}}} \left| L\left(\frac{1}{2} + z_1, \chi_p\right) L\left(\frac{1}{2} + z_2, \chi_p\right) \right|^r \\ & \ll_{r, \varepsilon} \frac{X}{(\log X)^{1-\varepsilon}} \exp \left( r \mathcal{M}(z_1, z_2, X) + \frac{r^2}{2} \mathcal{V}(z_1, z_2, X) \right). \end{aligned}$$

*Proof of (9.3.3) assuming Proposition 9.3.* Recall (9.3.5). If  $t_1$  or  $t_2$  satisfies  $|t_i| > X$  we use Cauchy-Schwarz, Lemma 4.5, and the rapid decay of  $g$  to get a negligible error.

We may therefore assume that  $|t_i| \leq X$ . We then consider, for a parameter  $0 < \alpha < 1$  at our disposal, two cases: (1) both  $t_1$  and  $t_2$  satisfy  $|t_i| \leq (\log X)^{-\alpha}$ , or (2) at least one of  $t_1, t_2$  satisfies  $|t_i| \geq (\log X)^{-\alpha}$ . In case (1) we use the trivial bounds

$$\begin{aligned} \mathcal{M}(it_1, it_2, X) & \leq \log \log X, \\ \mathcal{V}(it_1, it_2, X) & \leq 4 \log \log X, \end{aligned}$$

while in case (2) we use the bounds

$$\begin{aligned} \mathcal{M}(it_1, it_2, X) & \leq \frac{1+\alpha}{2} \log \log X, \\ \mathcal{V}(it_1, it_2, X) & \leq \frac{7+\alpha}{2} \log \log X + O(1). \end{aligned}$$

Since  $|g(it)| \ll (1+t^2)^{-1}$  we obtain by Proposition 9.3 that the quantity in (9.3.5) is

$$\ll X(\log X)^\varepsilon \left( (\log X)^{3-2\alpha} + (\log X)^{9/4+3\alpha/4} \right) = X(\log X)^{27/11+\varepsilon} \leq X(\log X)^{5/2}$$

upon choosing  $\alpha = 3/11$ . □

To prove Proposition 9.3 we establish estimates for how often  $|L(\frac{1}{2} + z_1, \chi_p) L(\frac{1}{2} + z_2, \chi_p)|$  can be large. The following is very similar to [42, Proposition 6.2].

**Proposition 9.4.** *Assume the hypotheses of Proposition 9.3. Let  $\mathcal{N}(V; z_1, z_2, X)$  denote the number of primes  $p \leq X$ ,  $p \equiv 1 \pmod{8}$ , such that  $\log |L(\frac{1}{2} + z_1, \chi_p) L(\frac{1}{2} + z_2, \chi_p)| \geq V + \mathcal{M}(z_1, z_2, X)$ . In the range  $3 \leq V \leq 4r\mathcal{V}(z_1, z_2, X)$  we have*

$$\mathcal{N}(V; z_1, z_2, X) \ll \frac{X}{(\log X)^{1-or(1)}} \exp\left(-\frac{V^2}{2\mathcal{V}(z_1, z_2, X)}\right),$$

and for larger  $V$  we have

$$\mathcal{N}(V; z_1, z_2, X) \ll \frac{X}{(\log X)^{1-or(1)}} \exp(-2rV).$$

*Proof of Proposition 9.3.* We have

$$\begin{aligned} & \sum_{\substack{p \leq X \\ p \equiv 1 \pmod{8}}} |L(\tfrac{1}{2} + z_1, \chi_p) L(\tfrac{1}{2} + z_2, \chi_p)|^r \\ &= r \int_{-\infty}^{\infty} \exp(rV + r\mathcal{M}(z_1, z_2, X)) \mathcal{N}(V; z_1, z_2, X) dV. \end{aligned}$$

Then use Proposition 9.4. □

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 9.2.** *Let  $X$  and  $y$  be real numbers and  $k$  a natural number with  $y^k \leq X^{\frac{1}{2} - \frac{1}{\log_2 X}}$ . For any complex numbers  $a(q)$  we have*

$$\sum_{\substack{p \leq X \\ p \equiv 1 \pmod{8}}} \left| \sum_{2 < q \leq y} \frac{a(q) \chi_p(q)}{q^{\frac{1}{2}}} \right|^{2k} \ll \frac{X \log_2 X}{\log X} \frac{(2k)!}{2^k k!} \left( \sum_{q \leq y} \frac{|a(q)|^2}{q} \right)^k,$$

where the summation variable  $q$  runs through the primes. The implied constant is absolute.

*Proof.* This result is similar to [42, Lemma 6.3], so we give only a sketch. Since we are assuming GRH we could use Lemma 9.1, but we get an unconditional result that is almost as good by appealing to sieve theory.

Since  $p \equiv 1 \pmod{8}$ , we have  $\chi_p(q) = \chi_q(p)$ , by the quadratic reciprocity law. Observe that  $\chi_q$  is a primitive character with conductor  $\leq 4q$ . We then introduce an upper bound sieve with weights  $\lambda_d$  supported on  $d \leq D = X^{\frac{1}{\log_2 X}}$ . (For example, we may take the weights  $\lambda_d$  in (5.8) with  $R \approx X^{\frac{1}{\log_2 X}}$ ). 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, we see that the sum in question is then

$$\begin{aligned} &\ll \sum_{n \leq X} \left( \sum_{d|n} \lambda_d \right) \left| \sum_{2 < q \leq y} \frac{a(q) \chi_q(n)}{q^{\frac{1}{2}}} \right|^{2k} \\ &\ll \sum_{\substack{q_i \leq y \\ q_1 \cdots q_{2k} = \square}} \frac{|a(q_1) \cdots a(q_{2k})|}{\sqrt{q_1 \cdots q_{2k}}} \sum_{n \leq X} \left( \sum_{d|n} \lambda_d \right) + D \log(y^{2k}) \sum_{q_1, \dots, q_{2k} \leq y} |a(q_1) \cdots a(q_{2k})|. \end{aligned}$$

For the first term we obtain

$$\begin{aligned} \sum_{\substack{q_i \leq y \\ q_1 \cdots q_{2k} = \square}} \frac{|a(q_1) \cdots a(q_{2k})|}{\sqrt{q_1 \cdots q_{2k}}} \sum_{n \leq X} \left( \sum_{d|n} \lambda_d \right) &\ll \frac{X}{\log D} \sum_{\substack{q_i \leq y \\ q_1 \cdots q_{2k} = \square}} \frac{|a(q_1) \cdots a(q_{2k})|}{\sqrt{q_1 \cdots q_{2k}}} \\ &\ll \frac{X \log_2 X (2k!)}{\log X} \left( \sum_{q \leq y} \frac{|a(q)|^2}{q} \right)^k, \end{aligned}$$

where the sieved sum on  $n$  contributes the factor  $\frac{X}{\log D} = \frac{X}{\log X} \log_2 X$  (this follows, for example, from Lemma 5.3), and the last inequality follows as in [42, Lemma 6.3]. For the second term we use Cauchy-Schwarz and the assumption  $y^k \leq X^{\frac{1}{2} - \frac{1}{\log_2 X}}$  to obtain

$$D \log(y^{2k}) \sum_{q_1, \dots, q_{2k} \leq y} |a(q_1) \cdots a(q_{2k})| \ll X \frac{k \log X}{D} \left( \sum_{q \leq y} \frac{|a(q)|^2}{q} \right)^k.$$

□

*Proof of Proposition 9.4.* Assume GRH for  $L(s, \chi_p)$ . A modification of the proof of the Proposition in [41] then yields, for  $2 \leq x \leq X$ ,

$$\begin{aligned} \log |L(\tfrac{1}{2} + z_1, \chi_p) L(\tfrac{1}{2} + z_2, \chi_p)| &\leq \operatorname{Re} \left( \sum_{q^\ell \leq x} \frac{\chi_p(q^\ell)}{\ell q^{\ell(\frac{1}{2} + \frac{1}{\log x})}} (q^{-\ell z_1} + q^{-\ell z_2}) \frac{\log(x/q^\ell)}{\log x} \right) \\ &\quad + 2 \frac{\log X}{\log x} + O\left(\frac{1}{\log x}\right), \end{aligned}$$

where the sum here is a double sum over primes  $q$  and positive integers  $\ell$ . The terms with  $\ell \geq 3$  contribute  $O(1)$ . For  $\ell = 2$  we use the Riemann hypothesis for  $\zeta(s)$  (see [42, (6.4)]) and obtain

$$\frac{1}{2} \sum_{q \leq \sqrt{x}} \frac{1}{q^{1 + \frac{2}{\log x}}} (q^{-2z_1} + q^{-2z_2}) \frac{\log(x/q^2)}{\log x} = \mathcal{M}(z_1, z_2, x) + O(\log \log \log X).$$

Since  $\mathcal{M}(z_1, z_2, x) \leq \mathcal{M}(z_1, z_2, X)$ , we obtain

$$\begin{aligned} (9.3.6) \quad \log |L(\tfrac{1}{2} + z_1, \chi_p) L(\tfrac{1}{2} + z_2, \chi_p)| &\leq \operatorname{Re} \sum_{2 < q \leq x} \frac{\chi_p(q)}{q^{\frac{1}{2} + \frac{1}{\log x}}} (q^{-z_1} + q^{-z_2}) \frac{\log(x/q)}{\log x} \\ &\quad + \mathcal{M}(z_1, z_2, X) + 2 \frac{\log X}{\log x} + O(\log \log \log X). \end{aligned}$$

Now recall that we wish to bound the number  $\mathcal{N}(V; z_1, z_2, X)$ . By taking  $x = \log X$  in (9.3.6) and estimating trivially, we see that we may assume  $V \leq \frac{3 \log X}{\log \log X}$ , for otherwise  $\mathcal{N}(V; z_1, z_2, X) = 0$ . We may also assume that  $V^2 \geq \mathcal{V}(z_1, z_2, X)$ , since otherwise the conclusion of Proposition 9.4 would follow immediately from the prime number theorem. We put  $\mathcal{V} = \mathcal{V}(z_1, z_2, X)$ , and define

$$T = \begin{cases} \frac{1}{2} \log \log \log X, & V \leq \mathcal{V}, \\ \frac{\mathcal{V}}{2V} \log \log \log X, & \mathcal{V} < V \leq \frac{1}{16} \mathcal{V} \log \log \log X, \\ 8, & V > \frac{1}{16} \mathcal{V} \log \log \log X. \end{cases}$$

We take  $x = X^{T/V}$  in (9.3.6), and define  $z = x^{1/\log \log X}$ . We then obtain

$$(9.3.7) \quad \log \left| L\left(\frac{1}{2} + z_1, \chi_p\right) L\left(\frac{1}{2} + z_2, \chi_p\right) \right| \leq S_1 + S_2 + \mathcal{M}(z_1, z_2, X) + 3 \frac{V}{T},$$

where  $S_1$  is the sum over  $q$  truncated to  $q \leq z$ , and  $S_2$  is the remainder of the sum. Observe that if  $\log \left| L\left(\frac{1}{2} + z_1, \chi_p\right) L\left(\frac{1}{2} + z_2, \chi_p\right) \right| \geq V + \mathcal{M}(z_1, z_2, X)$ , then (9.3.7) implies that either

$$S_2 \geq \frac{V}{T} \quad \text{or} \quad S_1 \geq V \left(1 - \frac{6}{T}\right) =: V_1.$$

Therefore  $\mathcal{N}(V; z_1, z_2, X) \leq \mathcal{N}_1 + \mathcal{N}_2$ , where  $\mathcal{N}_2$  is the number of primes  $p \leq X$  with  $p \equiv 1 \pmod{8}$  such that  $S_2 \geq \frac{V}{T}$ , and  $\mathcal{N}_1$  is the number of  $p \leq X$  with  $p \equiv 1 \pmod{8}$  such that  $S_1 \geq V_1$ . Hence, to bound  $\mathcal{N}(V; z_1, z_2, X)$ , it suffices to bound  $\mathcal{N}_2$  and  $\mathcal{N}_1$ .

To bound  $\mathcal{N}_2$ , we take  $k = \lfloor (\frac{1}{2} - \frac{1}{\log_4 X}) \frac{V}{T} \rfloor - 1$  in Lemma 9.2 and apply the usual Chebyshev-type maneuver

$$\#\{p \leq X, p \equiv 1 \pmod{8} : S_2 \geq V/T\} \leq \sum_{\substack{p \leq X \\ p \equiv 1 \pmod{8}}} \frac{|S_2|^{2k}}{(V/T)^{2k}}$$

to deduce that (see also [42, p. 1115])

$$\mathcal{N}_2 \ll \frac{X \log_2 X}{\log X} \exp \left( -\frac{V}{4T} \log V \right).$$

By a similar argument via Lemma 9.2, this time using the fact that  $\operatorname{Re}(q^{-z_1} + q^{-z_2}) = \frac{1}{2}(q^{-z_1} + q^{-\bar{z}_1} + q^{-z_2} + q^{-\bar{z}_2})$ , we see for any  $k \leq (\frac{1}{2} - \frac{1}{\log_2 X}) \frac{V \log \log X}{T}$  that

$$\mathcal{N}_1 \ll \frac{X \log_2 X}{\log X} \left( \frac{2k(\mathcal{V}(z_1, z_2, X) + O(\log \log \log X))}{eV_1^2} \right)^k$$

(see also [42, p. 1115]). For  $V \leq (\log \log X)^2$  we take  $k = \lfloor V_1^2/2\mathcal{V} \rfloor$ , and for  $V > (\log \log X)^2$  we take  $k = \lfloor 10V \rfloor$ . We arrive at

$$\mathcal{N}_1 \ll \frac{X \log_2 X}{\log X} \exp \left( -\frac{V_1^2}{2\mathcal{V}} \left( 1 + O \left( \frac{\log \log \log X}{\log \log X} \right) \right) \right) + \frac{X \log_2 X}{\log X} \exp(-V \log V).$$

□

## 10. PROOF OF THEOREM 1.4

The proof of Theorem 1.4 breaks naturally into two parts: the lower bound, and the upper bound. The argument for the lower bound is very similar to that in [38], and we therefore give only a sketch. The argument for the upper bound is similar to that in Section 7. In either case, we crucially use the assumption that the central values are non-negative.

**10.1. The lower bound.** Let  $d_{1/2}(n)$  be the multiplicative function with Dirichlet convolution  $(d_{1/2} \star d_{1/2})(n) = 1$ . 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

$$\begin{aligned} T_1 &:= \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) L\left(\frac{1}{2}, \chi_p\right) R(p)^4, \\ T_2 &:= \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) R(p)^6. \end{aligned}$$

In  $T_2$  we open up  $R(p)^6$ , and obtain a sum over  $n_1, \dots, n_6$ , 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 6.1.

For  $T_1$ , we write  $L(\frac{1}{2}, \chi_p)$  using Lemma 4.2. After opening  $R(p)^4$ , we have a sum over  $n_1, \dots, n_4, m$ , 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 6.1. This gives the lower bound.

**10.2. 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

$$\begin{aligned} M_3 &:= \sum_{p \equiv 1 \pmod{8}} (\log p) \Phi\left(\frac{p}{X}\right) L\left(\frac{1}{2}, \chi_p\right)^3 \\ &\leq (\log X) \sum_{n \equiv 1 \pmod{8}} \mu^2(n) \left( \sum_{\substack{d|n \\ d \leq D}} \lambda_d \right) \Phi\left(\frac{n}{X}\right) L\left(\frac{1}{2}, \chi_n\right)^3 \end{aligned}$$

The coefficients  $\lambda_d$  of the sieve are given, as before, by (5.8). We take  $R$  to be a sufficiently small power of  $X$ .

We use the approximate functional equation

$$L\left(\frac{1}{2}, \chi_n\right)^3 = \frac{16}{(\sqrt{2}-1)^6} \sum_{\substack{\nu=1 \\ \nu \text{ odd}}}^{\infty} \frac{d_3(\nu) \left(\frac{\nu}{n}\right)}{\sqrt{\nu}} \omega_3\left(\nu \left(\frac{\pi}{n}\right)^{3/2}\right),$$

where  $\omega_3(\xi)$  is defined by taking  $j = 3$  in (4.1). 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 7.1. For  $N_Y(n)$  we use Poisson summation as before. Up to negligible error, we therefore have the upper bound

$$M_3 \leq (\log X) \frac{16}{(\sqrt{2} - 1)^6} \sum_{\substack{d \leq D \\ d \text{ odd}}} \lambda_d \sum_{\substack{\nu=1 \\ (\nu, 2d)=1}}^{\infty} \frac{d_3(\nu)}{\sqrt{\nu}} \sum_{\substack{\alpha \leq Y \\ (\alpha, 2\nu)=1}} \mu(\alpha) \\ \times \left( \frac{2[\alpha^2, d]}{\nu} \right) \frac{X}{[\alpha^2, d]8\nu} \sum_{k \in \mathbb{Z}} e \left( \frac{k[\alpha^2, d]\nu}{8} \right) \hat{F}_\nu \left( \frac{kX}{[\alpha^2, d]8\nu} \right) \tau_k(\nu),$$

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 7.3, 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 second expression in (4.3) for  $\tau_k(n)$  and treat separately the contributions of the terms  $(\frac{1+i}{2}) G_k(n)$  and  $(\frac{-1}{n}) (\frac{1-i}{2}) G_k(n)$ . We then follow the method of Section 7 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 7.2. 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 7. We therefore do not need to concern ourselves with any symmetry properties of the integrand (cf. the symmetry argument yielding (7.9.5)).

#### ACKNOWLEDGEMENTS

The authors would like to thank the anonymous referee for a thorough examination of the paper, and for several suggestions that have improved the clarity of our exposition. We would also like to thank Steve Gonek and Matthew Young for helpful comments that have improved our exposition.

The second author was supported by NSF grant DMS-1501982, and by the National Science Foundation Graduate Research Program under grant number DGE-1144245.

## REFERENCES

- [1] J. C. Andrade, S. Baluyot. *Small zeros of Dirichlet  $L$ -functions of quadratic characters of prime modulus*. Res. Number Theory 6, 18 (2020).
- [2] J. C. Andrade, J. P. Keating. *Mean value theorems for  $L$ -functions over prime polynomials for the rational function field*. Acta Arith. 161 (2013), no. 4, 371–385.
- [3] T. M. Apostol. *Introduction to Analytic Number Theory*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1976.
- [4] R. Balasubramanian, V. K. Murty. *Zeros of Dirichlet  $L$ -functions*. Ann. Sci. École Norm. Sup. (4) 25 (1992), no. 5, 567–615.
- [5] H. Bohr, E. Landau. *Sur les zeros de la fonction  $\zeta(s)$  de Riemann*. C. R. Acad. des Sciences Paris 158 (1914), 158–162.
- [6] H. M. Bui. *Non-vanishing of Dirichlet  $L$ -functions at the central point*. Int. J. Number Theory 8 (2012), no. 8, 1855–1881.
- [7] H. M. Bui, M. Milinovich. *Central values of derivatives of Dirichlet  $L$ -functions*. Int. J. Number Theory 7 (2011), no. 2, 371–388.
- [8] S. Chowla. *The Riemann hypothesis and Hilbert’s tenth problem*. Mathematics and Its Applications, Vol. 4 Gordon and Breach Science Publishers, New York-London-Paris 1965
- [9] Henri Cohen. *Number theory. Vol. I. Tools and Diophantine equations*. Graduate Texts in Mathematics, 239. Springer, New York, 2007.
- [10] J. B. Conrey, A. Ghosh, S. M. Gonek. *Simple zeros of the zeta-function of a quadratic number field. II*. Analytic number theory and Diophantine problems (Stillwater, OK, 1984), 87–114, Progr. Math., 70, Birkhäuser Boston, Boston, MA, 1987.
- [11] S. Das, R. Khan. *The third moment of symmetric square  $L$ -functions*. Q. J. Math. 69 (2018), no. 3, 1063–1087.
- [12] Harold Davenport. *Multiplicative number theory*. Third edition. Revised and with a preface by Hugh L. Montgomery. Graduate Texts in Mathematics, 74. Springer-Verlag, New York, 2000.
- [13] David W. Farmer. *Long mollifiers of the Riemann zeta-function*. Mathematika 40 (1993), no. 1, 71–87.
- [14] J. Friedlander, H. Iwaniec. *What is ... the parity phenomenon?* Notices Amer. Math. Soc. 56 (2009), no. 7, 817–818
- [15] J. Friedlander, H. Iwaniec. *Opera de Cribro*. American Mathematical Society Colloquium Publications, 57. American Mathematical Society, Providence, RI, 2010.
- [16] D. R. Heath-Brown. *A mean value estimate for real character sums*. Acta Arith. 72 (1995), no. 3, 235–275.
- [17] J. Hoffstein, W. Luo. *Nonvanishing of  $L$ -series and the combinatorial sieve*. With an appendix by David E. Rohrlich. Math. Res. Lett. 4 (1997), no. 2-3, 435–444.
- [18] H. Iwaniec, E. Kowalski. *Analytic number theory*. American Mathematical Society Colloquium Publications, 53. American Mathematical Society, Providence, RI, 2004.
- [19] H. Iwaniec, P. Sarnak. *Dirichlet  $L$ -functions at the central point*. Number theory in progress, Vol. 2 (Zakopane-Kościelisko, 1997), 941–952, de Gruyter, Berlin, 1999.
- [20] H. Iwaniec, P. Sarnak. *The non-vanishing of central values of automorphic  $L$ -functions and Landau-Siegel zeros*. Israel J. Math. 120 (2000), part A, 155–177.
- [21] M. Jutila. *Zero-density estimates for  $L$ -functions*. Acta Arith. 32 (1977), no. 1, 55–62.
- [22] M. Jutila. *On the mean value of  $L(\frac{1}{2}, \chi)$  for real characters*. Analysis 1 (1981), no. 2, 149–161.
- [23] N. Katz, P. Sarnak. *Zeros of zeta functions and symmetry*. Bull. Amer. Math. Soc. 36 (1999), no. 1, 1–26.
- [24] Rizwanur Khan. *Non-vanishing of the symmetric square  $L$ -function at the central point*. Proc. Lond. Math. Soc. (3) 100 (2010), no. 3, 736–762.
- [25] Rizwanur Khan. *Simultaneous non-vanishing of  $GL(3) \times GL(2)$  and  $GL(2)$   $L$ -functions*. Math. Proc. Cambridge Philos. Soc. 152 (2012), no. 3, 535–553.
- [26] R. Khan, D. Milićević, H. T. Ngo. *Non-vanishing of Dirichlet  $L$ -functions in Galois orbits*. Int. Math. Res. Not. IMRN 2016, no. 22, 6955–6978.
- [27] R. Khan, H. T. Ngo. *Nonvanishing of Dirichlet  $L$ -functions*. Algebra Number Theory 10 (2016), no. 10, 2081–2091.



- [28] E. Kowalski, P. Michel. *A lower bound for the rank of  $J_0(q)$* . Acta Arith. 94 (2000), no. 4, 303–343.
- [29] E. Kowalski, P. Michel, J. VanderKam. *Non-vanishing of high derivatives of automorphic  $L$ -functions at the center of the critical strip*. J. Reine Angew. Math. 526 (2000), 1–34.
- [30] E. Kowalski, P. Michel, J. VanderKam. *Mollification of the fourth moment of automorphic  $L$ -functions and arithmetic applications*. Invent. Math. 142 (2000), no. 1, 95–151.
- [31] P. Michel, J. VanderKam. *Non-vanishing of high derivatives of Dirichlet  $L$ -functions at the central point*. J. Number Theory 81 (2000), no. 1, 130–148.
- [32] K. Ono, C. Skinner. *Non-vanishing of quadratic twists of modular  $L$ -functions*. Invent. Math. 134 (1998), no. 3, 651–660.
- [33] A. E. Özlük, C. Snyder. *Small zeros of quadratic  $L$ -functions*. Bull. Austral. Math. Soc. 47 (1993), no. 2, 307–319.
- [34] A. E. Özlük, C. Snyder. *On the distribution of the nontrivial zeros of quadratic  $L$ -functions close to the real axis*. Acta Arith. 91 (1999), no. 3, 209–228.
- [35] D. H. J. Polymath. *Variants of the Selberg sieve, and bounded intervals containing many primes*. Res. Math. Sci. 1 (2014), Art. 12, 83 pp.
- [36] Kyle Pratt. *Average non-vanishing of Dirichlet  $L$ -functions at the central point*. Algebra Number Theory 13 (2019), no. 1, 227–249.
- [37] M. Radziwiłł, K. Soundararajan. *Moments and distribution of central  $L$ -values of quadratic twists of elliptic curves*. Invent. Math. 202 (2015), no. 3, 1029–1068.
- [38] Z. Rudnick, K. Soundararajan. *Lower bounds for moments of  $L$ -functions: symplectic and orthogonal examples. Multiple Dirichlet series, automorphic forms, and analytic number theory*, 293–303, Proc. Sympos. Pure Math., 75, Amer. Math. Soc., Providence, RI, 2006.
- [39] Atle Selberg. *On the zeros of Riemann’s zeta-function*. Skr. Norske Vid. Akad. Oslo I. 1942, (1942). no. 10, 59 pp.
- [40] K. Soundararajan. *Nonvanishing of quadratic Dirichlet  $L$ -functions at  $s = \frac{1}{2}$* . Ann. of Math. (2) 152 (2000), no. 2, 447–488.
- [41] K. Soundararajan. *Moments of the Riemann zeta function*. Ann. of Math. (2) 170 (2009), no. 2, 981–993.
- [42] K. Soundararajan, M. Young. *The second moment of quadratic twists of modular  $L$ -functions*. J. Eur. Math. Soc. (JEMS) 12 (2010), no. 5, 1097–1116.
- [43] E. C. Titchmarsh, *The Theory of the Riemann Zeta-function*. Second edition. Revised by D. R. Heath-Brown. The Clarendon Press, Oxford University Press, New York, 1986.
- [44] Andrew Wiles. *The Birch and Swinnerton-Dyer Conjecture. The millennium prize problems*, 31–41, Clay Math. Inst., Cambridge, MA, 2006.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, 1409 WEST GREEN STREET, URBANA, IL 61801, USA

*Current address:* American Institute of Mathematics, 600 East Brokaw Road, San Jose, CA 95112, USA

*Email address:* sieg.baluyot@gmail.com

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, 1409 WEST GREEN STREET, URBANA, IL 61801, USA

*Current address:* All Souls College, Oxford OX1 4AL, United Kingdom

*Email address:* kyle.pratt@all-souls.ox.ac.uk, kvpratt@gmail.com