ON QUADRATIC FAMILIES OF CM ELLIPTIC CURVES

RITABRATA MUNSHI

Abstract. Given a CM elliptic curve with Weierstrass equation $y^2 = f(x)$, and a positive definite binary quadratic form $Q(u,v)$, we show that there are infinitely many reduced integer pairs $(u,v)$ such that the twisted elliptic curve $Q(u,v)y^2 = f(x)$ has analytic rank (and consequently Mordell-Weil rank) one. In fact it follows that the number of such pairs with $|u|, |v| \leq X$ is at least $X^2 - \varepsilon$ for any $\varepsilon > 0$.

1. Introduction

Let $E : y^2 = f(x)$ be an elliptic curve defined over $\mathbb{Q}$ and let $Q(u,v)$ be a positive definite binary quadratic form over $\mathbb{Q}$. We define a family of elliptic curves given by

$$E_{u,v} : Q(u,v)y^2 = f(x), \quad u, v \in \mathbb{Z}, \quad \gcd(u,v) = 1.$$  \hfill (1)

In this paper we will prove the following theorem.

Theorem 1. Let $E$ be a CM elliptic curve of conductor $q$. Let $Q(u,v)$ be a positive definite binary quadratic form of discriminant $D$. Suppose $(D,q) = 1$. Furthermore, if $q$ is a square, then suppose that there is a pair $(u_0,v_0)$ such that the elliptic curve $E_{u_0,v_0}$ has root number $-1$. Then there are infinitely many coprime integer pairs $(u,v)$ such that the elliptic curve $E_{u,v}$ has analytic rank one.

A similar result for linear families, i.e. $E_{u,v} : l(u,v)y^2 = f(x)$ with $l$ a linear form, has been proved by Goldfeld-Hoffstein-Patterson [4] for CM elliptic curves, and by Murty-Murty [11] and Bump-Friedberg-Hoffstein [2] for general elliptic curves. The situation in the present paper is much more delicate owing to the fact that the quadratic family is too sparse for any direct application of harmonic analysis. This will be explained in some detail below.

In light of the well-known results of Gross-Zagier, Kolyvagin and Rubin, linking analytic rank with Mordell-Weil rank, we have the following corollary.

Corollary 1. Under the same hypothesis there are infinitely many coprime integer pairs $(u,v)$ such that the elliptic curve $E_{u,v}$ has Mordell-Weil rank one.

For an application of our theorem consider the (congruent number) elliptic curve

$$E : y^2 = x^3 - x,$$
which has complex multiplication by the ring of integers \( \mathbb{Z}[\sqrt{-1}] \). The prime \( p = 2 \) is the only place of bad reduction. Let \( Q(u, v) \) be a quadratic form with odd discriminant \( D \). Notice that the discriminant is coprime with the conductor of the elliptic curve. We consider the family of curves given by

\[
E_{u, v} : y^2 = x^3 - Q(u, v)^2x.
\]

From the theory of quadratic forms and the root number formula [1] it follows that there is a coprime integer pair \((u_0, v_0)\) such that the corresponding curve \( E_{u_0, v_0} \) has root number \(-1\). Consequently, from our theorem it follows that the quadratic form \( Q(u, v) \) represents infinitely many congruent numbers.

**Corollary 2.** Any positive definite binary quadratic form of odd discriminant properly represents infinitely many congruent numbers.

For our next application recall that in [8] we established that the level of distribution of the sequence of central derivatives \( \Lambda'(1/2, E^d) \), where \( d \) runs through the set of fundamental discriminants coprime with the conductor of the elliptic curve \( E \), is at least one-half. The main motivation of [8] was to show that for any quadratic field \( K \) there are infinitely many square-free integers \( d \) containing no inert primes and such that the twisted elliptic curve \( E^d \) has analytic rank one. This result would follow from the fundamental lemma of the half dimensional sieve if we had a level of distribution slightly larger than one-half. In [8] we just fell short of beating the threshold one-half and hence the half dimensional sieve was not effective. However now we have the following corollary.

**Corollary 3.** Let \( K \) be an imaginary quadratic field with discriminant coprime with the conductor of the CM elliptic curve \( E \). Then there are infinitely many square-free integers \( d \) such that all the prime factors of \( d \) split in \( K \) and such that the twisted elliptic curve \( E^d \) has analytic rank (and consequently Mordell-Weil rank) one.

We will prove the above theorem by computing the averages of the central derivatives of the \( L \)-functions over the family. Let \( r_Q(d) \) denote the number of representations of the integer \( d \) by the quadratic forms in the genus of \( Q \). So if \( r_Q(d) > 0 \), then we have a coprime integer pair \((u, v)\), and an integer \( w \) such that \( Q(u, v) = dw^2 \). Consequently, the fiber \( E_{u, v} \) is isogenous (over \( \mathbb{Q} \)) to the quadratic twist \( E^d : dy^2 = f(x) \). To prove Theorem [1] it is enough to prove that there are infinitely many square-free integers \( d \) such that \( r_Q(d) > 0 \), and \( E^d \) has analytic rank one. The following weighted mean value theorem serves the purpose.

**Theorem 2.** We assume as before that \( (q, D) = 1 \). Let \( r_Q(d) \) denote the number of representations of \( d \) by the quadratic forms in the genus of \( Q \). Then we have

\[
\sum_{\substack{d \text{ \&-free} \\ (d, 2qD) = 1}} r_Q(d) \Lambda'(1/2, E^d) F(d/Y) = \alpha Y \log Y + O(Y(\log Y)^\delta)
\]

for some \( \delta < 1 \). The leading coefficient \( \alpha \neq 0 \) when either of the following conditions are satisfied: (i) \( q \) is not a square, (ii) the root number of \( E \) is \(-1\).

In general we are interested in obtaining a similar mean value result for quadratic twists of modular forms. Recently some progress was made in this direction (see
and it was shown that
\[
\sum_d r_Q(d)\Lambda^{(l)}(1/2, f \otimes \chi_d)F(d/Y) = YP_l(\log Y) + O(Y(\log Y)^{7+\frac{1}{2}})
\]
for any newform \(f\). Here \(\chi_d\) denotes the quadratic character modulo \(d\), and \(P_l\) is a polynomial of degree \(l\). Consequently, if the order of the derivative \(l\) is at least 8, one can show that the main term dominates the error term by explicitly computing the leading coefficients of \(P_l\). It seems that the method applied in [9] has a natural barrier and it may not be possible, in general, to get a better bound than \(O(Y(\log Y)^{1+\varepsilon})\) for the error term, without any new input. Of course for the Diophantine problem, which is our primary motivation, we require an error term which is \(o(Y\log Y)\). The purpose of this paper is to show that when the modular form \(f\) is associated with a CM elliptic curve then we may do better by carefully exploiting the distribution of the sizes of the Fourier coefficients of \(f\).

Let us briefly recall the main issues in evaluating the weighted mean value (3). Let \(\lambda_f(n)\) denote the normalized Fourier coefficient of \(f\). Using the approximate functional equation the left hand side of (3) is reduced to a double sum which is roughly of the form
\[
\sum_{d \sim Y} r_Q(d) \sum_{m \leq Y} \lambda_f(n)\chi_d(m)\sqrt{m} \log \left(\frac{Y}{m}\right).
\]

The above sum can be compared with
\[
\sum_{n \sim Y} \sum_{m \leq Y} \lambda_g(n)\lambda_f(m)\left(\frac{m}{n}\right),
\]
where \(f\) and \(g\) are \(GL_2\) automorphic forms. This is a well-known ‘deadlock situation’ as the Fourier transform (Voronoi summation formula), in either variable, is of no help. Indeed the summation formula acts like an identity. In [9] we abandoned the idea of getting cancellation in both the sums, and just obtained a reasonable upper bound by breaking the deadlock using some inequalities of multiplicative functions. It worked due to the presence of the weights \(\log(Y/m)\) which elevate the main term, but do not hamper the error term. In this paper we will follow the same line as [9], but with finer analysis of the error term.

To get a quantitative result we use the well-known bound for the second moment of the central derivatives (see [5]),
\[
\sum_d |\Lambda'(1/2, f \otimes \chi_d)|^2F(d/Y) = O(Y^{1+\varepsilon}).
\]

A direct application of Cauchy’s inequality then gives us the following corollary.

**Corollary 4.** In the setup of the preceding theorem, we have
\[
\#\{d : |d| < D, r_Q(d)\Lambda'(1/2, E_d) \neq 0\} \gg D^{1-\varepsilon}
\]
for any \(\varepsilon > 0\).

From this we can derive the following quantitative version of Theorem [11].

**Corollary 5.** In the setup of Theorem [11] we have
\[
\#\{(u, v) : |u|, |v| < D, \Lambda'(1/2, E_{u, v}) \neq 0\} \gg D^{2-\varepsilon}
\]
for any \(\varepsilon > 0\).
In this context let us remark that in [7] we have shown that the cubic surface given by the affine equation \( Q(u, v) = f(x) \) (where \( f \) is an irreducible cubic) has at least \( Y/(\log Y)^{3/2} \) many rational points of height at most \( Y \). It is expected that most of these points come from elliptic curves of Mordell-Weil rank one. The above corollary supports this expectation.

Finally we note that our method works (with minor changes) for negative definite forms and indefinite forms. Also it works for any dihedral form associated with a grossencharacter of a quadratic field.

**Theorem 3.** Let \( f_\psi \) be a dihedral form associated with the grossencharacter \( \psi \) of a quadratic field. Let \( r_Q(d) \) denote the number of representations of \( d \) by the quadratic forms in the genus of \( Q \). Then we have

\[
\sum_d r_Q(d)\Lambda(l)(1/2, f_\psi \otimes \chi_d)F(d/Y) = YP_l(\log Y) + O(Y(\log Y)^\delta)
\]

for some \( \delta < 1 \). Here \( P_l \) is a polynomial of degree \( l \).

2. Preliminaries

Let us begin by recalling some basic facts about representation of integers by positive definite binary quadratic forms. Let \( Q(x, y) \) be such a form of discriminant \( D \), and let us assume that \( D \) is fundamental. A square-free integer \( d \) is properly represented by a form of discriminant \( D \) if and only if

\[
\chi_D(p) = 1 \quad \forall \ p|d,
\]

where \( \chi_D \) is the quadratic character associated with the imaginary quadratic field \( \mathbb{Q}(\sqrt{D}) \). (Recall that by proper representation we mean \( Q(u, v) = d \) with \( (u, v) = 1 \).) To determine whether \( d \) is represented by the form \( Q \) we introduce the genus characters. These are given by the quadratic residue symbols \( (\frac{\cdot}{p}) \), one for each odd prime \( p \) dividing \( D \). Moreover for discriminants with 4 \( | D \) we may have either one or two more characters of modulus 4 or 8. For convenience we will label the characters as

\[ \{\psi_1, \ldots, \psi_h\}, \]

where \( h = \omega(D) \) or \( \omega(D) \pm 1 \), depending on the nature of \( D \). (Here \( \omega(D) \) denotes the number of prime factors of \( D \).) Now associated with the given form \( Q \) we have a string \( (\varepsilon_1, \ldots, \varepsilon_h) \), with \( \varepsilon_i = \pm 1 \) and \( \prod_i \varepsilon_i = 1 \), such that the given integer \( d \) is represented by a form in the genus of \( Q \) if and only if \( d \) satisfies (9) and \( \psi_i(d) = \varepsilon_i \) for all \( i \). Note that for the principal genus the associated string is given by \( \varepsilon_i = 1 \) for all \( i \). Finally for any given form \( Q(x, y) \) there is a choice for a fixed set of squares \( \eta_j^2 \) for \( j = 1, \ldots, h^*(D) \) such that one of \( d\eta_j^2 \) is represented by \( Q(x, y) \) whenever \( d \) is represented by some form in the genus of \( Q \). Here \( h^*(D) \) denotes the number of forms in any genus.

We want to study the set of coprime integer pairs \( (u, v) \) such that the elliptic curve

\[
E_{u,v} : Q(u, v)y^2 = f(x)
\]

has analytic rank one. We consider this as a problem of understanding the set of square-free integers \( d \) such that \( d \) is represented by the quadratic form \( Q \), and the twisted elliptic curve \( E^d : dy^2 = f(x) \) has analytic rank one. Now since \( E^d \) and \( E^{d\eta_j^2} \) are isogenous over \( \mathbb{Q} \), we can replace the condition that \( d \) is represented by the form \( Q \) by the weaker condition that \( d \) is represented by some form in the genus
of $Q$. The last condition can be written in terms of characters. Indeed using genus theory, as explained above, a square-free positive integer $d$ is represented by a form in the genus of $Q$ if any only if

$$r_Q(d) = \prod_i \left( \psi_i(d) + \varepsilon_i \right) \prod_{p|d} \left( 1 + \chi_D(p) \right)$$

is nonzero. We set

$$R_D(d) = \prod_{p|d} \left( 1 + \chi_D(p) \right),$$

and note that it is a multiplicative function on the set of square-free integers.

We will use the theory of quadratic forms to simplify our job. Indeed we will pick a curve in the family, say $E_{u,v}$, and then study the twists of it by the principal quadratic form $Q_0$, i.e.

$$E_{u,v}^*: Q_0(s,t)Q(u,v)y^2 = f(x).$$

By Gauss composition the integers $Q_0(s,t)Q(u,v)$ are represented by the form $Q$, and hence the family of curves $E_{u,v}^*$ is in fact a subfamily of the original family.

Next we will recall some well-known facts about elliptic curves. Let $E$ be a CM elliptic curve of conductor $q$ defined over $\mathbb{Q}$ with complex multiplication by an order in the imaginary quadratic field $K$. To this we can associate a newform $f$ of weight 2 and level $q$, such that the Hasse-Weil $L$-function of $E$ coincides with the $L$-function of the form $f$

$$L(s, E) = L(s, f) = \sum_{n=1}^{\infty} \lambda(n)n^{-s}. \tag{12}$$

Recall that we normalize the $L$-functions such that $1/2$ is the central point. So the Fourier expansion of $f$ is given by

$$f(z) = \sum_{n=1}^{\infty} n^{1/2} \lambda(n)e(nz),$$

and $|\lambda(n)| \leq \tau(n)$ where $\tau$ is the divisor function. Moreover for primes $p$ which are inert in $K$ we have $\lambda(p) = 0$, and for the remaining primes the angles $\theta_p$, where $\lambda(p) = 2\cos\theta_p$, are uniformly distributed over $[0, \pi]$ (see [3]). We also define the two set of primes

$$Q = \{ p : p \text{ inert in } \mathcal{O}_K \}, \quad \mathcal{P} = \{ p : \lambda(p) > 1 \}. \tag{13}$$

Note that the first set of primes has density 1/2, and the second set of primes has density 1/3. This special property of CM elliptic curves will play a crucial role in this paper.

For integers $d \equiv 0$, or $1$ (mod 4) we put $\chi_d(n) = (\frac{d}{n})$. So $\chi_d$ is a real character with conductor $|d|$. If $d$ is a fundamental discriminant, then $\chi_d$ is a primitive character of conductor $|d|$ associated with the quadratic field $\mathbb{Q}(\sqrt{d})$. We take a square-free $d$ such that $(d, 2q) = 1$, then the Hasse-Weil $L$-function of the quadratic twist elliptic curve $E^d$: $dy^2 = f(x)$ is given by

$$L(s, E^d) = L(s, f, \chi_d) = \sum_{n=1}^{\infty} \lambda(n)\chi_d(n)n^{-s}.$$
This satisfies the functional equation \( \Lambda(s, f, \chi_d) = \varepsilon_f \chi_d(q^*) \Lambda(1 - s, f, \chi_d) \), where \( q^* \) is the square-free part of \( q \), \( \varepsilon_f \) is the sign of the functional equation of \( L(s, f) \) (or the root number of the elliptic curve \( E \)), and the completed \( L \)-function is given by

\[
\Lambda(s, f, \chi_d) = \left( \frac{\sqrt{|d|}}{2\pi} \right)^{s - \frac{1}{2}} \Gamma \left( s + \frac{1}{2} \right) L(s, f, \chi_d).
\]

(13)

To evaluate the \( L \)-function inside the critical strip we can use the functional equation to get a rapidly decaying series expansion called the approximate functional equation. For the first derivative at the center we have

\[
\Lambda'(1/2, f, \chi_d) = (1 - \varepsilon_f \chi_d(q^*)) \sum_{n=1}^{\infty} \frac{\lambda(n) \chi_d(n)}{\sqrt{n}} \sqrt{n} V \left( \frac{2\pi n \sqrt{|d|}}{q} \right),
\]

(14)

where the smooth function \( V \) is given by

\[
V(y) = \frac{1}{2\pi i} \int_{(3)} \frac{d}{ds} \left( \frac{\Gamma(s + 1)}{y^s} \right) ds.
\]

This function decays rapidly as \( y \) gets larger, in fact we have

\[
V(y) \ll_N y^{-N}
\]

for all \( N \geq 1 \). This implies that in (14), the terms with \( n \gg d \) (the tail) make a very small contribution, or in other words the first \( \asymp d \) terms of the sum give a very good approximation for the central value. In addition, when \( y \) is small

\[
V(y) = (\log y) + v + O(y),
\]

for some constant \( v \). This fact will play an important role in our analysis. Observe that this implies that the beginning terms in the sum in (14) are weighted by \( \log |d| \), whereas near the end, i.e. when \( n \sim |d| \) the terms are unweighted.

We conclude this section by recalling the fundamental result of Kolyvagin. First observe that if root number is \(+1\), then the derivative of the completed \( L \)-function vanishes at the center \( \Lambda'(1/2, f, \chi_d) = 0 \). So the nonvanishing of \( \Lambda'(1/2, f, \chi_d) \) implies that root number is \(-1\) and \( \text{ord}_{s=1/2} L(s, E^d) = 1 \). The result of Kolyvagin now says that the last condition implies that the Mordell-Weil rank of the elliptic curve \( E^d \) is also one, i.e.

\[
\Lambda'(1/2, f, \chi_d) \neq 0 \implies \text{rank}_2 E^d(Q) = 1.
\]

3. Initial reductions

For the sake of simplicity we will assume that \((D, q) = 1\). However our treatment is more general, and can be worked out without this coprimality assumption. In that case we have to be careful about the primes dividing \((D, q)\) in the computation of the leading term in the next section.

Let \( Q \) be the principal quadratic form of discriminant \( D \). We are interested in the sum

\[
S = \sum_{(d, 2qD) = 1}^{b} (1 + \chi_{-4}(d)) r(d) \Lambda'(1/2, f, \chi_D \chi_d) F(d/Y),
\]

(15)
where $F$ is a nonnegative smooth function with compact support. (The notation $\flat$ on the top-right of the summation sign indicates that the sum is over square-free integers.) The weight $r(d)$ is given by

$$r(d) = \prod_{p \mid D} \left( 1 + \left( \frac{p}{D} \right) \right) \times R_D(d),$$

which is related (via reciprocity) with $r_Q(d)$ as defined in (10) for the principal form $Q$. In fact they differ only slightly when $D$ is even. But in any case $(1 + \chi_{-4}(d))r(d) \neq 0$ implies that $r_Q(d) \neq 0$, i.e. $d$ is represented by the principal form. Notice that we are imposing a condition so that only square-free odd integers $d \equiv 1 \pmod{4}$ contribute to the sum $S$. This is necessary for using reciprocity. We are also throwing in the character $\chi_D$, which will simplify the computation of the main term. Observe that this does not affect our goal as $-D$ is represented by the principal form of discriminant $D$, and we may take for $f$ the newform associated with the CM elliptic curve $(-1)y^2 = f(x)$.

**Lemma 1.** We have

$$S = C \left( 1 - \varepsilon_f \prod_{p \mid q^*} \frac{\lambda(p)}{\sqrt{p}} \left( 1 + \frac{1}{p} \right)^{-1} \right) Y \log Y + O(Y(\log Y)^{10}),$$

for a nonzero constant $C$ which depends on the form $f$, the discriminant $D$ and the cut-off function $F$.

Indeed this is the mean value theorem that we will prove in the rest of the paper. Compare the above lemma with Theorem 2.

Using the approximate functional equation (14), and (10) we write

$$S = \sum \chi S_{\chi},$$

where $\chi$ ranges over the set of quadratic characters which we get by opening up the product

$$\left( 1 + \left( \frac{-1}{d} \right) \right) \left( 1 - \left( \frac{q^*}{d} \right) \varepsilon_f \right) \prod_{p \mid D} \left( 1 + \left( \frac{p}{d} \right) \right).$$

Notice that we are using the positivity of $d$. The coefficients $c_\chi$ are obvious. Finally

$$S_{\chi} = \sum_{(d,2qD)=1}^{b} R_D(d) \chi(d) \sum_{n=1}^{\infty} \frac{\lambda(n)\chi_D(n)}{\sqrt{n}} \left( \frac{n}{d} \right) G_n(d/Y),$$

where $R_D$ is as given in (11), and

$$G_n(y) = V \left( \frac{2\pi n}{\sqrt{qY}} \right) F(y).$$

Since $F$ is compactly supported and smooth, the analytic nature of $G_n$ is inherited from $V$, in particular we have

$$G_n(y) \ll (n/Y)^{-N}$$

for any $N \geq 1$. So the computation of $S$ boils down to computing sums of the form

$$T = \sum_{(d,2qD)=1}^{b} R_D(d) \left( \frac{q^*}{d} \right) \sum_{n=1}^{\infty} \frac{\lambda(n)\chi_D(n)}{\sqrt{n}} \left( \frac{n}{d} \right) G_n(d/Y),$$

where $\lambda(n)\chi_D(n)$ is a square-free integer.
for a given square-free integer \(a\) (which in our case is \(D\) and \(\theta = \pm 1, \pm q^*\)). Now we will interchange the order of summation and use inverse Mellin transform. To this end we define the Mellin transform

\[
H_n(s) = \int_0^\infty G_n(y) y^s dy.
\]

Since \(G_n(y)\) is compactly supported, the function \(H_n(s)\) is entire and in any vertical strip the function decays faster than \(\Im(s)^{-N}\) for any \(N \geq 1\). In fact we have the bound

\[
H_n(\sigma + it) \ll_{\sigma,N,M} (n/Y)^{-N}(1 + |t|)^{-M}.
\]

Now the inverse Mellin transform yields

\[
G_n(y) = \frac{1}{2\pi i} \int_{(\sigma)} H_n(s) y^{-s} ds,
\]

for any \(\sigma\). Taking \(\sigma > 1\) and replacing in the expression \((17)\) and interchanging the order of summations and integration we get

\[
T = \frac{1}{2\pi i} \int_{(\sigma)} Y^s \left\{ \sum_{n=1}^{\infty} \lambda(n) \chi_D(n) \frac{H_n(s)}{\sqrt{n}} \sum_{(d,2qD)=1} R_D(d) \left( \frac{an}{d} \right)^s \right\} ds
\]

\[
= \frac{1}{2\pi i} \int_{(\sigma)} Y^s \left\{ \sum_{n=1}^{\infty} \lambda(n) \chi_D(n) \frac{H_n(s) L(s, \chi_{4an}) L(s, \chi_{4Dan}) \zeta(2s)^{-3} D_n(s)}{\sqrt{n}} \right\} ds.
\]

To derive the last equality we used \((11)\). The Dirichlet series \(D_n(s)\) converges absolutely for \(\sigma > 1/3\), and in the domain \(\sigma \geq 1/2\) it satisfies the bound \(|D_n(s)| \ll \frac{n}{\phi(n)}\), where the implied constant depends only on \(q, D\). Observe that the inner sum over \(n\) in \((18)\) converges absolutely and uniformly in any given compact domain for \(s\), away from the possible poles coming from the \(L\)-function (i.e. away from \(s = 1\)). Hence the sum defines an analytic function with a possible pole at \(s = 1\). We shift the contour of integration to the critical line \(\sigma = 1/2\) and in the process we pick up residue at the only possible pole at \(s = 1\). This gives

\[
T = M + E
\]

where \(M\) is coming from the residue at \(s = 1\) and it, as we will see in the next section, contributes to the main term. The term \(E\) is given by the integral over the central line \(\sigma = 1/2\) and it contributes to the error term. We observe that in our analysis of \(E\) the terms with \(n > Y^{1+\varepsilon}\) for any \(\varepsilon > 0\) contribute a negligible quantity compared to the other terms, and hence can be ignored. So it follows that

\[
E \ll \int_{(1/2)} \sqrt{Y} \left\{ \sum_{n \leq Y^{1+\varepsilon}} \frac{\lambda^*(n) |H_n(s)|}{\sqrt{n}} \right\} \left| \frac{L(s, \chi_{4an}) |L(s, \chi_{4Dan})|}{\zeta(2s)^{-3} D_n(s)} \right| ds + O(Y^{-2009}),
\]

where \(\lambda^*(n) = \lambda(n)n/\phi(n)\).
4. The main term

In this section we analyze the main term

\[ M = \sum_{n=1}^{\infty} \frac{\lambda(n)\chi_D(n)}{\sqrt{n}} \text{Res}_{s=1} \left\{ H_n(s) L(s, \chi_{4an}) L(s, \chi_{4Dan}) \zeta(2s) - 3 D_n(s) Y^s \right\}. \]

Observe that since \( D \) is not a square, the expression inside the braces may have at most a simple pole at \( s = 1 \). In fact there is a simple pole when either \( an = \varnothing \) or \( Dan = \varnothing \). Recall that \( a = \delta \theta \) where \( \delta | D \) and \( \theta = \pm 1, \pm q^* \). Now observe that due to the presence of the character \( \chi_D \) we have \( (n, D) = 1 \). So \( an = \varnothing \) can only hold if \( a = 1, q^* \), and similarly \( Dan = \varnothing \) can only hold if \( a = D, Dq^* \). So in the computation of the main term we are only concerned with these four cases. This is due to the presence of the character \( \chi_D \) in (15).

To compute the residue we go back to the expression (15) and observe that

\[ L(s, \chi_{4an}) L(s, \chi_{4Dan}) \zeta(2s) - 3 D_n(s) = \prod_{p \mid 4D} \left\{ 1 + \frac{1}{p^s} \left( \frac{an}{p} \right) + \frac{1}{p^s} \left( \frac{Dan}{p} \right) \right\}. \]

Hence it follows that when \( an = \varnothing \) or \( Dan = \varnothing \), we have

\[ \text{Res}_{s=1} L(s, \chi_{4an}) L(s, \chi_{4Dan}) \zeta(2s) - 3 D_n(s) = C_{q,D} \prod_{p \mid n, p \mid 2qD} \left\{ 1 + \frac{1}{p} + \frac{(D_p)}{p} \right\}^{-1}, \]

for some constant \( C_{q,D} = C \neq 0 \). So we obtain that

\[ M = CY \sum_{n=1}^{\infty} \frac{\lambda(n)\chi_D(n)}{\sqrt{n}} H_n(1) \prod_{p \mid n, p \mid 2qD} \left\{ 1 + \frac{1}{p} + \frac{(D_p)}{p} \right\}^{-1}, \]

where the * on the top-right of the summation sign indicates that either \( an = \varnothing \) or \( Dan = \varnothing \). We will write \( l_p \) for the local \( L \)-factor in the above expression. We seek to evaluate the above sum using contour integration. To this end recall that \( H_n(1) \) is given by

\[ H_n(1) = \int_{0}^{\infty} V \left( \frac{2\pi n}{\sqrt{qY}} \right) F(y) dy = \frac{1}{2\pi i} \int_{0}^{\infty} \int_{(3)} \frac{d}{ds} \left( \frac{1}{(\sqrt{qY})^s} \right) \frac{\Gamma(s+1)}{(2\pi n)^s} \frac{dy}{ds} F(y) dy. \]

Both the integrals appearing above are absolutely convergent. This together with the absolute convergence of the Dirichlet series

\[ D(s) = \sum_{n=1}^{\infty} \frac{\lambda(n)\chi_D(n)}{n^{s+1/2}} \prod_{p \mid n, p \mid 2qD} l_p \]

in the half-plane \( \sigma > 0 \), justify the interchange in the order of summation and integration, and also the term-by-term differentiation which yields

\[ M = CY \int_{0}^{\infty} \int_{(3)} \frac{d}{ds} \left( \frac{1}{(\sqrt{qY})^s} \frac{\Gamma(s+1)}{(2\pi n)^s} \right) \frac{dy}{ds} F(y) dy. \]

Now we will shift the contour of integration to the left and in the process pick up the residue at the pole \( s = 0 \). The residue will of course give us the leading term,
and the integral over the vertical line to the left of \( s = 0 \) will contribute to the error term. To make this precise we need to analyze the Dirichlet series \( D(s) \). In general it is related to the symmetric square \( L \)-function of \( f \). Using the multiplicativity of the Fourier coefficient we get that

\[
D(s) = L(s) \times \prod_{p \mid 2q} \left( 1 + \frac{\lambda(p^2)}{(p^2)^{s+1/2}} + \ldots \right) \prod_{p \mid 2qD} \left( 1 + \frac{\lambda(p^2)l_p}{(p^2)^{s+1/2}} + \ldots \right)
\]

where

\[
L(s) = \begin{cases} 
1 & \text{if } a = 1, D; \\
\prod_{p \mid q^*} \frac{\lambda(p)}{p^{s}} \left( 1 + \frac{1}{p^{s+1}} \right)^{-1} & \text{if } a = q^*, Dq^*.
\end{cases}
\]

There is a Dirichlet series \( D_2(s) \) which is absolutely convergent and has a Euler product expression in the half-plane \( \sigma > -1/4 \), such that

\[
D(s) = L(2s + 1, \text{Sym}^2 f) D_2(s).
\]

Notice that \( D_2(0) \neq 0 \) as it is given by a convergent Euler product. It is a well-known fact that the symmetric square \( L \)-function \( L(s, \text{Sym}^2 f) \) has analytic continuation to the whole of the complex plane and it does not vanish at \( s = 1 \). Hence \( D(0) \neq 0 \).

By contour integration we get

\[
M = CY \int_0^\infty \frac{d}{ds} \left( \Gamma(s + 1)(4\gamma^2Y)^s \right) \bigg|_{s=0} D(0)F(y)dy + R
\]

where the remainder term \( R \) is given by the same integral as in (19), but now over the vertical line \( \sigma = -1/4 + \varepsilon \). Hence it follows from standard estimates that

\[
R \ll Y^{3/4+\varepsilon}.
\]

The leading term is given by

\[
L = CY \log Y D(0) \int_0^\infty F(y)dy + C' Y,
\]

for some constant \( C' \) (which will not play any role in our analysis) depending on \( a \).

Now we can derive the leading term in the statement of Lemma 1. Let \( L_x \) denote the leading term corresponding to \( S_x \) in (16). We have \( L_x = 0 \) except in the following four cases

\[
L_x = \begin{cases} 
C^* Y \log Y & \text{if } \chi = 1 \text{ or } \chi_D; \\
C^* Y \log Y \times \prod_{p \mid q^*} \frac{\lambda(p)}{\sqrt{p}} \left( 1 + \frac{1}{p} \right)^{-1} & \text{if } \chi = (q^*) \text{ or } \chi_D(q^*),
\end{cases}
\]

where \( C^* \) is a nonzero constant depending on the form \( f \), the discriminant \( D \) and the smooth cut-off function \( F \). The corresponding coefficients \( c_x \) can be easily computed. Putting them together we get the following lemma.

**Lemma 2.** The leading term of \( S \) as defined in (15) is given by

\[
\sum_{\chi} c_x L_x = 2C^* \left( 1 - \varepsilon f \prod_{p \mid q^*} \frac{\lambda(p)}{\sqrt{p}} \left( 1 + \frac{1}{p} \right)^{-1} \right) Y \log Y.
\]
We conclude this section with some remarks concerning the leading term. First we observe that if \( q^\ast = 1 \), i.e. \( q = \square \), then the leading term vanishes if \( \varepsilon_f = 1 \). Indeed in this case all the curves in the family that we are taking into account in (15) have root number +1. We may investigate what happens if we introduce twists \( \chi_d \) in (15) such that \((d, 2q) \neq 1\). To bypass this issue we rather prefer to introduce the condition that there exists a pair of integers \((u, v)\) such that the root number of the curve \( E_{u,v} \) is \(-1\). Then we study the twists of this curve by the principal quadratic form of discriminant \( D \). As we have already noted, by the theory of quadratic forms these curves are also members of the original family.

5. The error term

We now turn to the analysis of the error term \( E \) which we have already seen to be bounded above by

\[
E \ll \int_{(1/2)} \frac{\sqrt{Y}}{|\zeta(2s)|^3} \left\{ \sum_{n \leq Y^{1+\varepsilon}} \frac{|\lambda^\ast(n)||H_n(s)|}{\sqrt{n}} |L(s, \chi_{4n})| |L(s, \chi_{4Dn})| \right\} |ds| + O_\varepsilon (Y^{-2009}),
\]

where \( \lambda^\ast(n) = \lambda(n)n/\phi(n) \). Notice that we have already sacrificed the hope of getting any cancellation from the sign changes of the Fourier coefficient \( \lambda(n) \). This seems to be beyond the reach of the current technology. (For more discussion regarding this issue see [9].) However we will be able to get a nice bound for the right hand side of (24).

Due to the fast decay of the function \( H_n(s) \) in the vertical line the outer integral is convergent. The main task is to give a sharp upper bound for the weighted average of the \( L \)-functions. To this end we will break up the inner sum into dyadic blocks and in the block \( N \leq n < 2N \) we will use the bound \( H_n(s) \ll H(N)(1 + |t|)^{-7} \), where

\[
H(N) = \begin{cases} 
1 + \log(Y/N) & \text{if } N \leq Y, \\
(Y/N) & \text{if } Y < N \ll Y^{1+\varepsilon}.
\end{cases}
\]

Then using Cauchy-Schwarz inequality, and using the trivial bound for the zeta function on the vertical line, we get that

\[
E \ll \sqrt{Y} (\log \log Y) \sum_{\text{dyadic blocks}} H(N) N^{-1/2} \int_{\mathbb{R}} U(N, t)(1 + |t|)^{-7}|t|^3 dt,
\]

where the sum is over \( \log Y \) many dyadic blocks, and

\[
U(N, t) = \sum_{N \leq n < 2N} |\lambda(n)||L\left(\frac{1}{2} + it, \chi_{4n}\right)|^2.
\]

Observe that in (26) we have a decay of order three near \( t = 0 \) coming from the pole of the Riemann zeta function at 1. Also we are using the relatively mild bound \( |\zeta(1 + it)|^{-1} \ll (1 + |t|) \). The above deduction may need more explanations. First we are replacing \( an \) by \( n \) and in the process increasing the \( n \)-sum by the fixed factor \( a \) (thanks to positivity). Also \( U(N, t) \) should in fact be defined as the maximum of the expression in the right of (27) and a similar expression with \( \chi_{4Dn} \) in place of \( \chi_{4n} \). However this does not alter the magnitude of \( U(N, t) \). Now to expand
the \( L \)-function in the critical line we first reduce the modulus \( n \) to a square-free number. Writing \( 4n \) as \( km^2 \) with \( k \) a fundamental discriminant we observe that

\[
U(N, t) \ll \sum_{m \leq \sqrt{N}} \tau(m^2) \tau(m) \sum_{k \sim N/m^2}^b |\lambda(k)| \left| L \left( \frac{1}{2} + it, \chi_k \right) \right|^2.
\]

(28)

Here we have used the Hasse bound. (Note the slight abuse of notation. The sum over \( k \) is over fundamental discriminants which, once we use the approximate functional equation, for all practical purposes can be thought of as a sum over square-free integers. So the \( \sqcup \) on the top-right of the summation formula is quite justified.) The inner sum being over fundamental discriminants \( k \), we can use the approximate functional equation to replace the \( L \)-function by a sum of two rapidly decaying series. We will continue our analysis with one such series

\[
L_k(t) = \sum_d \frac{\chi_k(d)}{d^{\frac{1}{2} + it}} W_t \left( \frac{d}{\sqrt{k}} \right).
\]

Here the smooth function is given by

\[
W_t(y) = \frac{1}{2\pi i} \int_{(3)} G(u) \frac{\Gamma \left( \frac{1}{4} + \frac{it}{2} + \frac{y}{2} \right)}{\Gamma \left( \frac{1}{4} + \frac{y}{2} \right)} (\sqrt{\pi y})^{-u} d \frac{u}{u}
\]

where \( G(u) \) is any function which is holomorphic in the strip \(-4 < \Re(u) < 4\), even and normalized by \( G(0) = 1 \) (see Theorem 5.3. in [6]).

Let

\[
U^*(N, t) = \sum_{N \leq n < 2N} \sum_d \frac{\chi_n(d)}{d^{\frac{1}{2} + it}} W_t \left( \frac{d}{\sqrt{n}} \right)
\]

\[
\left( \sum_{d \leq n} \chi_n(d) \right)^2.
\]

Notice that in the inner sum terms up to \( d \ll \sqrt{N(1 + |t|)} \) contribute to the sum; the other terms of the sum make a negligible contribution. (This follows from the exponential decay of the function \( W_t \) (see [6]).) Our job is now reduced to producing a decent upper bound for \( U^*(N, t) \).

The next step is the most crucial one where we break the deadlock as explained in the introduction by an inequality of multiplicative functions. We will first replace the weight \( |\lambda(n)| \) by an upper bound which will have the advantage of breaking the \( n \)-sum into two parts—a short sum and a long sum. The shorter sum is more complicated and we do not hope to gain anything from this part, i.e. we will evaluate this sum trivially. However the longer sum is without any arithmetic weights and we can execute this sum using the Poisson summation formula.

6. An Inequality

In [9] we used an inequality of the type

\[
|\lambda(n)| \ll \sum_{d \mid n} \beta_k
\]

for square-free \( n \). For each positive integer \( k \) there is an allowable range for \( \beta_k \) (see [7] or [10]). In light of the result in the next section, regarding the level of distribution of the sequence of absolute value squares of Dirichlet \( L \)-function at points on the critical line, we need an inequality of type (30) with \( k > 2 \). Fortunately it turns out that for our purpose \( k = 3 \) is quite sufficient. However in anticipation
Lemma 3. Let \( \tau_k(n) = \sum_{d_1 \ldots d_k = n} 1 \). Let \( 1 \leq a < b \) be two integers. Then for any square-free \( n \) we have

\[
\frac{\tau_b(n)}{\tau_{b-a}(n)} \leq b! \sum_{d|n \atop d \leq n^{a/b}} \frac{\tau_a(d)}{\tau_{b-a}(d)}.
\]

Proof. We start with the simple combinatorial inequality

\[
\tau_b(n) = \sum_{d_1 \ldots d_k = n} 1 \leq b! \sum_{d_1 \leq d_2 \leq \ldots \leq d_b} 1.
\]

For \( j > a \) it follows that \( d_j \geq d_a \geq \cdots \geq d_1 \), and hence \( d_j \geq (d_1 \ldots d_a)^{1/a} \). Consequently

\[
n = (d_1 \ldots d_a)d_{a+1} \ldots d_b \geq (d_1 \ldots d_a)^{1+(b-a)/a} = (d_1 \ldots d_a)^{b/a}.
\]

Hence we get that

\[
\tau_b(n) \leq b! \sum_{d_1 \ldots d_a|n \atop (d_1 \ldots d_a) \leq n^{a/b}} \sum_{d_{a+1} \ldots d_b|n/d_1 \ldots d_a} 1 \leq b! \sum_{d|n \atop d \leq n^{a/b}} \tau_a(d)\tau_{b-a}(\frac{n}{d}).
\]

Now we obtain the lemma by using the multiplicativity of \( \tau_{b-a} \) and the assumption that \( n \) is square-free. \( \square \)

For any fraction \( \delta \) we are interested in proving an inequality of the form

\[
|\sigma(n)| \ll \sum_{d|n \atop d \leq n^\delta} |\sigma(d)|^{\beta(\delta)}
\]

for the divisor function \( \sigma = \tau_2 \). We want to find the smallest possible value for the exponent \( \beta(\delta) \). We set \( V(d) = |\sigma(d)|^\beta \), and apply Hölder’s inequality to obtain

\[
\sum_{d|n \atop d \leq n^\delta} \tau_a(d)V(d)^{\frac{1}{t}} \leq \left[ \sum_{d|n \atop d \leq n^\delta} V(d) \right]^{\frac{1}{t}} \left[ \sum_{d|n \atop d \leq n^\delta} \frac{\tau_a(d)^{\frac{1}{t}}}{\tau_{b-a}(d)^\frac{1}{t} V(d)^{\frac{1}{t}}} \right]^{\frac{1}{t}}
\]

for any \( t > 0 \). Combining this inequality with Lemma 3 we get

\[
\left[ \frac{\tau_b(n)}{\tau_{b-a}(n)} \right]^{t+1} \left[ \sum_{d|n \atop d \leq n^\delta} \frac{\tau_a(d)^{\frac{1}{t+1}}}{\tau_{b-a}(d)^{\frac{1}{t+1}} V(d)^{\frac{1}{t}}} \right]^{-t} \leq (b!)^{t+1} \left[ \sum_{d|n \atop d \leq n^\delta} V(d) \right]^{t+1}.
\]

Observe that the left hand side is multiplicative. So to establish (31) we have to show that there is a \( t > 0 \) such that

\[
2 \leq \left[ \frac{b}{b-a} \right]^{t+1} \left[ 1 + \frac{a^{t+1}}{(b-a)^{t+1} 2^t} \right]^{-t}.
\]

The right hand side of the above inequality is given by

\[
\frac{b}{b-a} \left[ 1 - \frac{a}{b} \left( 1 - \left( \frac{a}{2^t(b-a)} \right)^{\frac{1}{t}} \right) \right]^{-t}.
\]
As $t \to \infty$ this converges to
\[
\frac{b}{b-a} \left( \frac{a}{2^\beta (b-a)} \right)^{-a/b},
\]
which is larger than 2 if $\beta$ satisfies the inequality
\[
\beta \log_2 \delta + \delta^{-1} \{ 1 + (1 - \delta) \log_2 (1 - \delta) \}.
\]
So for such $\beta$ we have a choice of $t$ and hence the inequality (31) follows. We will denote the right hand side of the above inequality by $\beta(\delta)$.

**Lemma 4.** Let $\tau(n) = \tau_2(n)$ be the divisor function. Given $1/2 \geq \delta > 0$, suppose $\beta > \beta(\delta)$. Then for any square-free integer $n$ we have
\[
|\tau(n)| \ll \sum_{d|n, d \leq n^{1/3}} |\tau(d)|^\beta,
\]
where the implied constant depends only on $\delta$ and $\beta$.

Recall that in the derivation of inequalities like (30) we take $a = 1$ and $b = k$. We note that $\beta(1/3) = 0.24511 \ldots$, which is such that $2^{\beta(1/3)} = 1.185 \ldots < 3/2$. Though we will not be using any other $\delta$, we note the following interesting exponents
\[
\beta(1/2) = 0, \beta(1/4) = 3 \log_2 3 - 4 = 0.75488 \ldots, \text{and} \beta(2/9) = 1.06179 \ldots.
\]
Given any square-free integer $n$, it can be uniquely written as $n = ab$ where all the prime factors of $a$ are in the set $\mathcal{P}$ (see (12)), and the prime factors of $b$ lie outside $\mathcal{P}$. Using the definition of the set $\mathcal{P}$, and the Hasse bound it follows that
\[
|\lambda(n)| \leq |\lambda(a)| \leq \tau(a).
\]
Combining this with Lemma 4 we obtain the following corollary.

**Corollary 6.** We have pairs $(\delta, \beta)$ with $2^\beta < 3/2$ and $\delta < 1/2$, such that for any square-free $n$ we have
\[
|\lambda(n)| \ll_{\delta, \beta} \sum_{d|n, d \leq n^{1/3}, p|d \Rightarrow p \in \mathcal{P}} |\tau(d)|^\beta.
\]
This inequality is of outmost importance in this paper. A concrete pair is given by $\delta = 1/3$ and $\beta = 1/4$
\[
|\lambda(n)| \ll \sum_{d|n, d \leq n^{1/3}, p|d \Rightarrow p \in \mathcal{P}} |\tau(d)|^{1/2}.
\]
After using this inequality we remove the square-free condition from $n$, and also smooth out the sharp cut-off using an appropriate smooth function $K$ and obtain
\[
U^*(N, t) \ll \sum_{a \leq N^{1/4}, p|a \Rightarrow p \in \mathcal{P}} |\tau(a)|^{1/2} U^*(N, t, a)
\]
where
\[
U^*(N, t, a) = \sum_{p|n} \left| \sum_d \frac{\chi_{an}(d)}{d^{1/2} + it} W_i \left( \frac{d}{\sqrt{an}} \right) \right|^2 K \left( \frac{an}{N} \right).
\]
Of course before removing the square-free condition on $n$, we utilize the fact that for square-free $n$ we have $\lambda(n) = 0$ unless all the prime factors of $n$ are outside $\mathcal{Q}$ (see [12]).

7. The Level of Distribution

In this section we want to address the question about the level of distribution of the sequence

\[
A(t) = \left\{ \alpha_n = \left| \sum_d \frac{1}{d^{\frac{3}{2} + it}} \left( \frac{n}{d} \right) W_t \left( \frac{d}{\sqrt{n}} \right) \right|^2 K \left( \frac{n}{N} \right) : n = 1, 2, \ldots \right\}.
\]

Our goal is to show, following the arguments given in [8] (the techniques are already present in [12]), that the sequence in question has level of distribution at least one-half. Any improvement of this will be of substantial interest.

Let

\[
A_a := \sum_{n \equiv 0 (\mod a)} \alpha_n.
\]

**Lemma 5.** Let $a$ be a square-free integer and $t \neq 0$. Then we have

\[
A_a = \left[ \alpha_1(t) \prod_{p | a} L_p(0, 0) + \alpha_2(t) \prod_{p | a} L_p(-it, it) \right] \frac{N}{a} \log N + \beta(a, N; t) \frac{N}{a} + E_a(N, t),
\]

where $L_p$ is defined in (45). We have a function $\beta(t)$ such that $|\beta(a, N; t)| \leq \prod_{p | a} (1 + \frac{1}{p}) \beta(t)$. The functions $\alpha_i$ and $\beta$ decay at infinity, and near $t = 0$ we have $|\alpha_i(t)| \ll |t|^{-2}$ and $|\beta(t)| \ll |t|^{-3}$. Moreover the error term satisfies the following:

\[
\sum_{a < A} |E_a(N, t)| \ll (1 + |t|)^2 AN^{1/2}(AN)^{\varepsilon}.
\]

Since the proof of this lemma is computational intensive, and basically follows the same line of arguments as given in [8], [12], we will take the liberty to be rather sketchy in some parts.

**Proof.** We will evaluate

\[
A_a = \sum_b \left| \sum_d \frac{1}{d^{\frac{3}{2} + it}} \left( \frac{ab}{d} \right) W_t \left( \frac{d}{\sqrt{ab}} \right) \right|^2 K \left( \frac{ab}{N} \right),
\]

using the Poisson summation formula. Opening the absolute square of the inner sum and interchanging the order of summation we get

\[
A_a = \sum_{d_1, d_2} \frac{1}{d_1^{\frac{1}{2} + it} d_2^{\frac{3}{2} - it}} \left( \frac{a}{d_1 d_2} \right) \sum_b \left( \frac{b}{d_1 d_2} \right) J_{d_1, d_2, t} \left( \frac{ab}{N} \right)
\]

where

\[
J_{d_1, d_2, t}(y) = W_t \left( \frac{d_1}{\sqrt{yN}} \right) W_t \left( \frac{d_2}{\sqrt{yN}} \right) K(y).
\]

Using the Poisson summation formula for the $b$-sum and then interchanging the order of summation we obtain

\[
A_a = \frac{N}{a} \sum_k \sum_{d_1, d_2} \frac{g(k, d_1 d_2)}{d_1^{\frac{1}{2} + it} d_2^{\frac{3}{2} - it}} \left( \frac{a}{d_1 d_2} \right) J_{d_1, d_2, t} \left( \frac{kN}{ad_1 d_2} \right).
\]
where $\tilde{J}_{d_1,d_2,t}(y) = \int_{\mathbb{R}} J_{d_1,d_2,t}(x)e(-xy)dx$, and

$$g(k, d_1d_2) = \sum_{\alpha \pmod{d_1d_2}} \left(\frac{\alpha}{d_1d_2}\right)e\left(\frac{k\alpha}{d_1d_2}\right)$$

is the Gauss sum. We will show that the principal contribution to the above sum comes from the zero frequency $k = 0$ and the frequencies corresponding to $ak = \square$. The other terms contribute to the error term.

Let us begin by evaluating the principal term

$$P(N,t,a) = \frac{N}{a} \sum_{d_1, d_2} g(0, d_1d_2) \left(\frac{a}{d_1d_2}\right) \tilde{J}_{d_1,d_2,t}(0).$$

Recall that $g(0, d_1d_2) = 0$ if $d_1d_2 \neq \square$, and $g(0, d_1d_2) = \phi(d_1d_2)$ if $d_1d_2 = \square$. Hence it follows that

$$P(N,t,a) = \frac{N}{a} \sum_{\substack{d_1, d_2 \in \square \quad \text{where } (d_1,d_2,a) = 1}} \phi(d_1d_2) \left(\frac{\alpha}{d_1d_2}\right) \tilde{J}_{d_1,d_2,t}(0).$$

This can be evaluated using the integral representation of $W_t(y)$ as given in (29). It follows that the principal term is given by

$$P(N,t,a)$$

where $\tilde{K}$ is the Mellin transform of $K$,

$$H_t(u_1, u_2) = G(u_1)G(u_2)\frac{\Gamma\left(\frac{1}{4} + \frac{it}{2} + \frac{u_1}{2}\right) \Gamma\left(\frac{1}{4} - \frac{it}{2} + \frac{u_1}{2}\right) \Gamma\left(\frac{1}{4} + \frac{it}{2} + \frac{u_2}{2}\right) \Gamma\left(\frac{1}{4} - \frac{it}{2} + \frac{u_2}{2}\right)}{\Gamma\left(\frac{1}{4} + \frac{it}{2}\right) \Gamma\left(\frac{1}{4} - \frac{it}{2}\right)},$$

and $L(u_1, u_2)$ is a double Dirichlet series

$$L(u_1, u_2) = \sum_{\substack{d_1, d_2 \in \square \quad \text{where } (d_1,d_2,a) = 1}} \frac{\phi(d_1d_2)}{d_1^{\frac{1}{4}+u_1+it}d_2^{\frac{1}{4}+u_2-it}}.$$

We can write the double Dirichlet series as a Euler product over all primes $p \mid a$. A careful study of the local factors gives that $L(u_1, u_2) = \prod_{p \mid a} L_p(u_1, u_2)^{-1}$, where

$$L_p(u_1, u_2)^{-1} = \frac{1}{p} + \left(1 - \frac{1}{p}\right)\left(1 + \frac{1}{p^{u_1+u_2}}\right)\left(1 - \frac{1}{p^{u_1+2u_2+2it}}\right)^{-1}\left(1 - \frac{1}{p^{1+2u_2-2it}}\right)^{-1}.$$
where the Euler product $D(u_1, u_2)$ converges absolutely in the tube domain $\Re(u_1), \Re(u_2) \geq -1/4$. Now the idea is to shift both the contours one by one to the left. In the process we will encounter poles and the residue there will contribute to the leading term and the integral on the left will contribute to the error term. To this end we first fix $u_1$ and move the contour over $u_2$ to the line $\Re(u_2) = -1/4$. There are simple poles at $u_2 = 0, it$. Now we shift the contour over $u_1$ to the line $\Re(u_1) = -1/4$. This time we encounter both simple and double poles, and they are located at the points $u_1 = 0, -it$. The residue computation now yields

$$P(N, t, a) = \left[ \alpha_1(t) \prod_{p \mid a} L_p(0, 0) + \alpha_2(t) \prod_{p \mid a} L_p(-it, it) \right] \frac{N}{a} \log N + \beta_1(a, N; t) \frac{N}{a} + O\left( \frac{N^{7/8}}{a} \right),$$

where $|\alpha_i(t)|$ decays as $|t| \to \infty$ (this is achieved by choosing $G(u) = (\cos \frac{\pi u}{4A})^{-4dA}$ with a positive integer $A$ large enough (see Chapter 5 of [1]). However $|\alpha_i(t)|$ have a singularity of type $t^{-2}$ at $t = 0$. The function $\beta_1(a, N; t)$ is oscillatory and includes $N^{\pm it}$. A closer analysis reveals that we have $|\beta_1(a, N; t)| \leq \beta(t)$ where the function $\beta(t)$ decays as $|t| \to \infty$, and it has a singularity of type $t^{-3}$ at $t = 0$. Finally we observe that the error term, which comes from the remaining integral, has an absolute implied constant.

We remark that we can recover the principal term for the case $t = 0$ from $P(N, t, a)$ by taking limit $t \to 0$. In the limit the oscillatory part of $\beta_1(a, N; t)$ yields a term containing $(\log N)^3$, and thus it becomes the leading term.

The secondary principal contribution comes from the terms with $ak = \square$ to $U^*(N, t, a)$. This is given by

$$P_{\square}(N, t, a) = \frac{N}{a} \sum_{ak = \square} \sum_{d_1, d_2 \neq 0} \frac{g(k, d_1d_2)}{d_1^{2+it}d_2^{-it}} \left( \frac{a}{d_1d_2} \right) J_{d_1, d_2, t} \left( \frac{kN}{ad_1d_2} \right).$$

Since for $(d_1, d_2, k) = 1$ (the generic case) the Gauss sum $g(k, d_1d_2) = (\frac{k}{d_1d_2}) \sqrt{d_1d_2}$, we see that its oscillatory part is cancelled by the character $(\frac{k}{d_1d_2})$, as $ak = \square$. So there is a main term present in this expression. We can again take recourse to Mellin inversion and Cauchy’s theorem to extract the main term. This yields

$$P_{\square}(N, t, a) = \beta_2(a, N; t) \frac{N}{a} + O\left( \frac{N^{7/8}}{a} \right).$$

We also have $|\beta_2(a, N; t)| \leq \beta(t)$, with $\beta(t)$ as above.

It remains to evaluate the contribution of the remaining frequencies. Since the frequencies $k$ with $|k| > \kappa := a(1 + |t|)^2N^\varepsilon$ make a negligible contribution, our job reduces to estimating

$$R(N, t, a) = \frac{N}{a} \sum_{ak \neq \square} \sum_{0 < |k| < \kappa} \frac{g(k, d_1d_2)}{d_1^{2+it}d_2^{-it}} \left( \frac{a}{d_1d_2} \right) J_{d_1, d_2, t} \left( \frac{kN}{ad_1d_2} \right).$$
The analysis of this remainder term is quite delicate. We again employ inverse Mellin transform to see that $R(N, t, a)$ is basically given by the integral

$$
\frac{N}{a} \sum_{\substack{ak \neq \square \\
0 < |k| < \kappa}} \iint H_t(u_1, u_2) \left( \frac{N}{\pi} \right)^{u_3} \sum_{d_1, d_2} \frac{g(k, d_1 d_2)(\frac{a}{d_1 d_2})}{d_1^{\frac{1}{2}+u_1+it} d_2^{\frac{1}{2}+u_2-it}} \int K(y)e\left(\frac{-kNy}{ad_1 d_2}\right)y^{u_3} dy du_1 du_2,
$$

where $u_3 = (u_1 + u_2)/2$ and the first two integrals are over $\Re(u_1) = 1/4$ and $\Re(u_2) = 1/4$, respectively. Notice that there is no problem with convergence. Now in the innermost integral we again employ inverse Mellin transform and interchange the order of integration to get

$$
\int K(y)e\left(\frac{-kNy}{ad_1 d_2}\right)y^{u_3} dy = \frac{1}{2\pi i} \int \mathcal{K}(w) \int e\left(\frac{-kNy}{ad_1 d_2}\right)y^{u_3-w} dy dw.
$$

The last integral is conditionally convergent if $-1 < \frac{1}{4} - c < 0$. Then changing variables and executing the inner integral we obtain

$$
\frac{1}{2\pi i} \int \left( \frac{ad_1 d_2}{|k|N} \right)^{u_3-w} \mathcal{K}(w) e\left(-\text{sgn}(k)(u_3-w+1)\right) \frac{\Gamma(u_3-w+1)}{2^{u_3-w+1}} dw.
$$

We will denote the product of the last three terms in the integrand by $I_k(u_3, w)$. Substituting this in the expression (48) we obtain

$$
\frac{N}{a} \sum_{\substack{ak \neq \square \\
0 < |k| < \kappa}} \iint H_t(u_1, u_2) \frac{N^w}{\pi u_3} \left( \frac{a}{|k|} \right)^{u_3-w} \mathcal{L}_k(u_4, w; t) I_k(u_3, w) dw du_1 du_2,
$$

where $u_4 = (u_1 - u_2)/2$ and

$$
\mathcal{L}_k(u_4, w; t) = \sum_{d_1, d_2} \frac{g(k, d_1 d_2)(\frac{a}{d_1 d_2})}{d_1^{\frac{1}{2}+w+u_4+it} d_2^{\frac{1}{2}+w-u_4-it}}.
$$

Explicitly computing the local Euler factors of the above Dirichlet series it follows that there exists a Dirichlet series $D_k(u_4, w; t)$ which is absolutely convergent for $\Re(w) > -1/2$ when $\Re(u_4) \geq 0$, and such that $\mathcal{L}_k(u_4, w; t)$ is given by the product

$$
L\left(1 + w + u_4 + it, \chi_{ak}\right) L\left(1 + w - u_4 - it, \chi_{ak}\right) D_k(u_4, w; t).
$$

We will shift the contour over $w$ to $\Re(w) = -1/2 + \varepsilon$. Since $ak \neq \square$ we do not encounter any poles of the Dirichlet series. Also there are poles of the gamma factors in the region. So we obtain that (48) is dominated above by

$$
\iint |H_t(u_1, u_2)| I_+(u_3, w) N^{\frac{1}{2}+\varepsilon} a \left[ \sum_{\substack{ak \neq \square \\
0 < |k| < \kappa}} \left( \frac{a}{|k|} \right) \frac{1}{2} \varepsilon |L(s_1, \chi_{ak})||L(s_2, \chi_{ak})|\right] |dw| du_1 |du_2| |u_1| |u_2|,
$$

plus a similar expression for the negative frequencies $-\kappa < k < 0$. Here $s_1 = 1 + w + u_4 + it$ and $s_2 = 1 + w - u_4 - it$. To estimate the sum inside the braces, we
break it into dyadic blocks, and in each block we use Cauchy’s inequality to obtain

\[
\ldots \ll \sum_{\text{dyadic blocks}} \left( \frac{a}{K} \right)^{3+\varepsilon} \left\{ \sum_{ak \not\in \mathbb{D}, k \sim K} |L(s_1, \chi_{ak})|^2 \right\}^{1/2} \left\{ \sum_{ak \not\in \mathbb{D}, k \sim K} |L(s_2, \chi_{ak})|^2 \right\}^{1/2}.
\]

Now using Heath-Brown’s large sieve inequality for quadratic characters [5] (see also [12]), we obtain that the second moments are individually bounded by \((aK + \sqrt{aK(1 + |\Im(s_i)|)})(aK)^\varepsilon\). Hence

\[
\ldots \ll a(1 + |t|)^{2(1 + |\Im(s_i)|)}(1 + |\Im(s_i)|)(aN)^\varepsilon.
\]

Recall that \(|k| < \kappa := a(1 + |t|)^2N^\varepsilon\). After integrating over \(u_i\) and \(w\) trivially we get

\[
R(N, t, a) \ll (1 + |t|)^{2} (aN)^{\varepsilon}.
\]

8. A SMALL SIEVE

We are left with the task of estimating the sum

\[
U^*(N, t, a) = \sum_n \left| \sum_d \frac{\chi_{an}(d)}{d^{\frac{3}{2} + it}} W_i \left( \frac{d}{\sqrt{an}} \right) \right|^2 K\left( \frac{an}{N} \right),
\]

where \(a \leq N^{\frac{1}{3}}\). We just want to get an upper bound of the right order of magnitude. To this end we can use a half-dimensional upper bound sieve \(\xi_+\) supported on the set of primes \(Q\) and having very small level, say \(\Delta = N^{\frac{1}{10}}\). So \(\xi_+ = 0\) for \(\nu > \Delta\). Then we have

\[
U^*(N, t, a) \leq \sum_{\nu \leq \Delta} \xi_+ \left\{ \sum_{n | \nu} \sum_d \frac{\chi_{an}(d)}{d^{\frac{3}{2} + it}} W_i \left( \frac{d}{\sqrt{an}} \right) \right\} K\left( \frac{an}{N} \right).
\]

Recall that (see (37)) all the prime factors of \(a\) belong to the set \(P\), whereas all the prime factors of \(\nu\) belong to \(Q\). Hence \(av\) is square-free. Now we apply Lemma 5 to the inner sum and obtain that \(U^*(N, t, a)\) is dominated by

(49) \[
\sum_{\nu \leq \Delta} \xi_+ \left\{ \alpha_1(t) \prod_{p | av} L_p(0, 0) + \alpha_2(t) \prod_{p | av} L_p(-it, it) \right\} \frac{N}{av} \log N

+ \beta(av, N; t) \frac{N}{av} + E_{av}(N, t) \right\}.
\]

To estimate the leading term of (49) we use the standard results (fundamental lemma) from sieve theory. We obtain, using the multiplicativity of the coefficients, that

\[
\sum_{\nu \leq \Delta} \xi_+ \prod_{p | \nu} L_p(0, 0) \ll \prod_{\substack{p \in Q \atop p \leq \Delta}} \left( 1 - \frac{L_p(0, 0)}{p} \right) \ll \frac{1}{\sqrt{\log N}}.
\]
In the last inequality we have used the explicit form of $L_p$ as given in (15). In a similar vein we obtain

$$\sum_{\nu \leq \Delta} \xi^+_\nu \prod_{p \mid \nu} L_p(-it, it) \ll \frac{1}{\sqrt{\log N}}.$$ 

Then it follows that

$$\sum_{\nu \leq \Delta} \xi^+_\nu \left[ \alpha_1(t) \prod_{p \mid \nu} L_p(0, 0) + \alpha_2(t) \prod_{p \mid \nu} L_p(-it, it) \right] \frac{N}{\nu \log N} \ll \alpha(t) \prod_{p \mid \nu} \left( 1 + \frac{1}{p} \right) \frac{N}{\nu} \sqrt{\log N},$$

where the function $\alpha(t)$ behaves like the functions $\alpha_i(t)$.

We estimate the contribution of the other terms trivially, not seeking any cancellation from the sum over $\nu$, and obtain that

$$\sum_{\nu \leq \Delta} \xi^+_\nu \beta_{\nu, N; t} \ll \sum_{\nu \leq \Delta} |E_{av}(N, t)|.$$ 

Comparing the terms we conclude the following upper bound.

**Lemma 6.** We have

$$U^*(N, t, a) \ll \gamma(t) \prod_{p \mid \nu} \left( 1 + \frac{1}{p} \right) N \sqrt{\log N} + \sum_{\nu \leq \Delta} |E_{av}(N, t)|,$$

where $\gamma(t) = \alpha(t) + \beta(t)$.

The above lemma gives us an upper bound for $U^*(N, t, a)$ of right order of magnitude.

9. Conclusion

Substituting the upper bound of $U^*(N, t, a)$ from Lemma 6 into (37), we obtain

$$U^*(N, t) \ll \gamma(t) \left[ \sum_{a \leq N^{\frac{1}{2}}} \frac{\tau(a)}{a} \prod_{p \mid a} \left( 1 + \frac{1}{p} \right) \right] N \sqrt{\log N} + \sum_{a \leq N^{\frac{1}{2}}} \sum_{\nu \leq \Delta} |E_{av}(N, t)| N^\varepsilon.$$

Now using Lemma 5 we obtain that the error term is bounded by

$$\sum_{a \leq N^{\frac{1}{2}}} \sum_{\nu \leq \Delta} |E_{av}(N, t)| N^\varepsilon \ll (1 + |t|)^2 \sqrt{N \Delta} N^{\frac{1}{2} + \varepsilon} \ll (1 + |t|)^2 N^{\frac{1}{2} + \frac{1}{10} + \varepsilon}.$$ 

The last inequality follows from our choice of $\Delta$.

Now we are left with the task of estimating

$$\sum_{a \leq N^{\frac{1}{2}}} \frac{\tau(a)}{a} \prod_{p \mid a} \left( 1 + \frac{1}{p} \right).$$
The above expression is trivially bounded above by the product
\[ \prod_{p \leq N, p \in \mathcal{P}} \left( 1 + \frac{2^k}{p} \left( 1 + \frac{1}{p^3} \right) \right). \]
Using the fact that the density of the set \( \mathcal{P} \) is one-third we get
\[ \prod_{p \leq N, p \in \mathcal{P}} \left( 1 + \frac{2^k}{p} \left( 1 + \frac{1}{p^3} \right) \right) \ll (\log N)^{1/4} \ll (\log N)^{1/2} - \theta \]
where \( \theta = \frac{1}{2} - \frac{2^{1/4}}{3} = 0.1035 \cdots > 0. \) Putting all together it follows that
\[ U^*(N, t) \ll \gamma(t)N(\log N)^{1-\theta} + (1 + |t|)^2N. \]
Substituting this bound in (28) we finally get
\[ U(N, t) \ll \sum_{m \leq \sqrt{N}} \tau(m^2)\tau(m) \left\{ \gamma(t) \frac{N}{m^2} (\log N)^{1-\theta} + (1 + |t|)^2 \frac{N}{m^2} \right\}. \]
Executing the sum trivially we arrive at the bound
\[ U(N, t) \ll \delta(t)N(\log N)^{1-\theta} \]
where the function \( \delta(t) \) grows like \( |t|^2 \) as \( |t| \to \infty \) and near \( t = 0 \) it blows up like \( |t|^{-3} \). Now we come to the final step of substituting the above bound in (26) and integrating over \( t \). This gives us
\[ E \ll \sqrt{Y}(\log \log Y) \sum_{\text{dyadic blocks}} H(N)N^{1/2}(\log N)^{1-\theta} \int_{\mathbb{R}} \delta(t)(1 + |t|)^{-7}|t|^3 dt. \]
The integral converges. Observe that we have a vanishing of order three at \( t = 0 \) coming from the pole of the zeta function [21]. Hence we obtain
\[ E \ll \sqrt{Y}(\log \log Y)(\log Y)^{1-\theta} \sum_{\text{dyadic blocks}} H(N)N^{1/2}. \]
Then we add up the contribution of all the dyadic blocks using the bound in (25), and obtain
\[ E \ll Y(\log Y)^{1-\theta}(\log \log Y)^3. \]
This completes the proof of Lemma 1.

Acknowledgement

The author thanks Kannan Soundararajan for his interest in this work and the referee for helpful comments.

References


Institute for Advanced Study, Einstein Drive, Princeton New Jersey 08540

E-mail address: rmunshi@math.ias.edu

Current address: School of Mathematics, Tata Institute of Fundamental Research, 1 Homi Bhabha Road, Colaba, Mumbai 400005, India

E-mail address: rmunshi@math.tifr.res.in