COMPOSITIO MATHEMATICA

On mean values and non-vanishing of derivatives of L-functions in a nonlinear family

Ritabrata Munshi

Compositio Math. **147** (2011), 19–34.

doi:10.1112/S0010437X10004732

FOUNDATION COMPOSITIO MATHEMATICA

The London Mathematical Society

Compositio Math. 147 (2011) 19–34
doi:10.1112/S0010437X10004732

On mean values and non-vanishing of derivatives of L-functions in a nonlinear family

Ritabrata Munshi

Abstract

We prove a mean-value result for derivatives of L-functions at the center of the critical strip for a family of forms obtained by twisting a fixed form by quadratic characters with modulus which can be represented as sum of two squares. Such a family of forms is related to elliptic fibrations given by the equation $q(t)y^2 = f(x)$ where $q(t) = t^2 + 1$ and $f(x)$ is a cubic polynomial. The aim of the paper is to establish a prototype result for such quadratic families. Though our method can be generalized to prove similar results for any positive definite quadratic form in place of sum of two squares, we refrain from doing so to keep the presentation as clear as possible.

1. Introduction

Let f be a primitive form with trivial central character, and let $\Lambda(s, f)$ denote the completed L-function associated with it. The function $\Lambda(s, f)$ has analytic continuation to the whole of the complex plane and satisfies a functional equation with $s \mapsto 1 - s$. (Here and elsewhere we normalize all the functional equations so that $s = 1/2$ is the center of the critical strip.) The non-vanishing of the L-function at the center $s = 1/2$ is of arithmetic importance, and has been extensively studied in the literature. The prototype problem in this field is the following: given a family of forms $\mathcal{F}$ show that there exists a form $f \in \mathcal{F}$ such that the lth derivative $\Lambda^{(l)}(1/2, f) \neq 0$. Often one is interested in proving a more quantitative result regarding the density of such forms in the family (like infinitely many non-vanishing or, better, a positive proportion non-vanishing).

One way of proving such a non-vanishing result is by forming a filtration $\mathcal{F} = \bigcup_Y \mathcal{F}(Y)$, with $|\mathcal{F}(Y)| < \infty$, and asymptotically computing the mean value (or the moment)

$$\sum_{f \in \mathcal{F}(Y)} \Lambda^{(l)}(1/2, f), \tag{1}$$

as $Y \to \infty$. Sometimes it is also useful to introduce certain weights, e.g. spectral weights or smooth cut-off functions, in (1). To obtain an asymptotic formula for (1) we usually require some sort of 'spectral completeness' for the family $\mathcal{F}$, so that we have a precise estimate for the sum

$$\sum_{f \in \mathcal{F}(Y)} q_f^{-1} \lambda_f(n), \tag{2}$$

where $\lambda_f(n)$ stands for the nth normalized Fourier coefficient of f and q_f denotes the level. Examples of such families include

$$\mathcal{F}_{\text{twist}} = \{g \otimes \chi : \chi\},$$

Received 29 May 2009, accepted in final form 14 November 2009, published online 20 April 2010.
2000 Mathematics Subject Classification 11F67 (primary), 11M41, 11G40 (secondary).
Keywords: derivatives of L-functions, non-vanishing of central values.

the family of twists of a fixed form g by all characters χ, and also the subfamily $\mathcal{F}_{\text{real twist}}$ consisting of all twists of g by real (quadratic) characters. For $\mathcal{F}_{\text{twist}}$ one can utilize the orthogonality of the characters to compute (2). However, when one restricts to the subfamily of quadratic twists, this crucial orthogonality relation is lost. Nevertheless, we still have enough harmonics in the family, and the Poisson summation formula can be used instead to compute (2). This yields the following mean value result

$$\sum_{f\in\mathcal{F}_{\text{real twist}}} \Lambda^{(l)}(1/2, f)F\left(\frac{q_f}{Y}\right) = YP_l(\log Y) + E(Y),$$

where F is a smooth compactly supported function, P_l is a polynomial of degree l and $E(Y)$ stands for the error term. It is not difficult to show that $E(Y) = o(Y)$, and with some work one may prove the stronger bound (see [Mun09, MM91]),

$$E(Y) = O(Y^{3/4+\varepsilon}).$$

The polynomial P_l depends on both the form g and the smooth cut-off function F. In particular, the leading coefficient is related to the value of the symmetric square L-function of g at $s = 1$, which is non-zero. (Some other interesting subfamilies of $\mathcal{F}_{\text{twist}}$ are given by restricting to twists by special characters like cyclotomic characters, cubic characters, quadratic characters with prime modulus etc.)

Other examples of interesting families which have been in the focus of research [IS00, KMV00] include: (i) the collection of all primitive forms of a given level and varying weight; and (ii) primitive forms of a given weight and varying level. In such cases we can evaluate the sum (2) using the spectral theorem of automorphic forms (e.g. the Petersson formula). The results that we get in these cases are even better, compared with $\mathcal{F}_{\text{real twist}}$, as the summation formula is much stronger. For example, in these families we can show that a positive proportion of the central values do not vanish (see [IS00]).

The purpose of this paper is to study the family of forms given by

$$\mathcal{F} = \{f\otimes\chi : \chi = \chi_d \text{ primitive real}, p|d \implies p\equiv 1 \pmod 4\}, \tag{3}$$

i.e. the family of twists of a primitive form f by real primitive characters χ_d such that d is odd and is a sum of two squares. This is a nonlinear algebraic subfamily of $\mathcal{F}_{\text{real twist}}$. For example, if f is a modular form associated with an elliptic curve $E : y^2 = f(x)$, then this family is associated with the quadratic elliptic fibration $(t^2 + 1)y^2 = f(x)$ with $t \in \mathbb{Q}$. In general we can also consider fibrations given by $q(t)y^2 = f(x)$ where $q(t)$ is a quadratic polynomial. In this context the family $\mathcal{F}_{\text{real twist}}$ corresponds to linear elliptic fibrations of the form $l(t)y^2 = f(x)$ where $l(t)$ is a linear polynomial.

It turns out that the usual (as in the case of $\mathcal{F}_{\text{real twist}}$) direct application of the Poisson summation formula is not sufficient to prove a mean value result for such a family. Let me briefly explain the problem. Let $r(d)$ denote the number of representations of d as a sum of two squares. Then we are interested in evaluating the following sum:

$$\sum_d r(d)L(1/2, f, \chi_d)F(d/Y),$$

where $L(1/2, f, \chi_d)$ is the L-function associated with the form $f\otimes\chi_d$. Using the approximate functional equation and interchanging the order of summation we reduce the problem to evaluating the sum

$$\sum_d r(d)\chi_d(n)F(d/Y)$$

https://doi.org/10.1112/S0010437X10004732 Published online by Cambridge University Press

where, in the worst case scenario, n is as large as Y. Notice that without the arithmetic weights $r(d)$ this sum can be evaluated using the Poisson summation formula, as the length of the sum (i.e. Y) is much larger than the square root of the modulus of the character (i.e. $\sqrt{n}$). However, if we use the Poisson (or Voronoi) summation formula with the weights $r(d)$, then the dual sum is roughly the same as the original sum (in length and complexity), and as such we do not gain anything from such a summation formula. Recall that it is useful to apply the twisted Voronoi summation formula for GL_2 only if the length of the sum is longer than the size of the modulus of the character, which is not the case here. In fact, the new input that we need is an upper bound for certain bilinear forms involving real characters. We now state the main result of this paper. (The notation $\square$ stands for a non-zero square integer.)

THEOREM 1. *Let f be a primitive form of level q, and let $r(d)$ denote the number of representations of d as a sum of two squares. Then we have*

$$\sum_{\substack{(d,2q)=1 \\ d\,\square\text{-free}}} r(d)\Lambda^{(l)}(1/2, f, \chi_d)F(d/Y) = YQ_l(\log Y) + E(Y),$$

where Q_l is a polynomial of degree l and the error term $E(Y)$ is bounded by

$$E(Y) \ll Y(\log Y)^{7+\frac{1}{2}}.$$

The error term is smaller than the leading term if the order of the derivative l is larger than seven. In such a case we may explicitly calculate the leading coefficient of the polynomial Q_l and hope to show that it is non-zero. This will imply that there are infinitely many twists $f \otimes \chi_d$ with $d = \square + \square$ such that the lth derivative of the completed L-function does not vanish at the central point. We also remark that $r(d)$ can be replaced by the more general $r_Q(d)$ which counts the number of representation of d by the positive definite binary quadratic form Q. (Of course there will be extra complications related to class number. But such problems can be tackled (see [IM10]).)

It will be apparent that our method can be carefully tuned to obtain a better bound for the error term. Such an endeavor, though possible, would increase the length of the paper to twice its present size. What is more discouraging is the fact that it might not be possible to beat the bound $E(Y) \ll Y(\log Y)^{1+\varepsilon}$ without any new ideas. So the most interesting cases of the central value and the first derivative would still remain out of reach. (We will briefly describe the possible improvements in the last section.)

Recently, double Dirichlet series have been used to prove various mean value results about L-functions and their derivatives. In particular, the mean value for the quadratic twist family $\mathcal{F}_{\text{real twist}}$ can be proved this way. The main advantage of this method is its ready adaptability with higher number fields and twists by higher order residue symbols. For the family $\mathcal{F}$ defined in (3), the 'model' double Dirichlet series is given by

$$\sum_d \frac{r(d)L(s, f \otimes \chi_d)}{d^w}. \tag{4}$$

Using the standard heuristics (see [CFH06]) we may expect to obtain a modified double Dirichlet series with analytic continuation if the 'expected' functional equations of (4) generate a finite group. One of the expected functional equations for (4) is given by $s \mapsto 1 - s$ and $w \mapsto w + 2s - 1$. To get another functional equation we expand the L-function as a series and interchange the order of summation. Then we get a double Dirichlet series of the form

$$\sum_n \frac{\lambda_f(n)L(w,\chi_n)\zeta(w)}{n^s}.$$

This gives the expected functional equation given by $w \mapsto 1-w$ and $s \mapsto s+2w-1$. Observe that these two functional equations generate an infinite group. (This is exactly the same deadlock that we encountered when we tried to use the Voronoi summation formula.) So we do not expect to get a modified double Dirichlet series with analytic continuation. Hence it seems that results of the form given in Theorem 1 are not well adapted in the double Dirichlet series scenario.

Finally, let us point out that we choose to work with the derivatives of the completed L-function $\Lambda(s,f)$ not just for the sake of convenience. Indeed our result cannot be translated to yield a similar mean-value theorem for the central derivatives $L^{(l)}(1/2,f)$, for any l. The usual yoga of expressing $L^{(l)}(1/2,f)$ as a linear combination of $\Lambda^{(j)}(1/2,f)$ with $j=1,\ldots,l$, does not work, as the coefficients in this linear combination involve powers of $\log q$ and this blows up the error term in Theorem 1. This is an extreme example where the symmetry of the functional equation is so crucial that a slight perturbation ruins the result completely. The lth derivative of the completed L-function $\Lambda^{(l)}(s,f)$ satisfies a symmetric functional equation for any l. This is not the case with the derivatives of the non-completed L-function $L(s,f)$.

2. Approximate functional equation

Let f be a primitive form of weight k and level q (which for convenience we assume to be odd), given by a Fourier expansion

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

with normalized Fourier coefficients $\lambda(n)$. Deligne has proved that $|\lambda(n)| \leqslant \tau(n)$ where τ is the divisor function. (Note that Deligne's bound will play a crucial role in our analysis and so our result does not hold for a Maass form.) For a Dirichlet character χ we define the twisted L-function

$$L(s,f,\chi)=\sum_{n=1}^{\infty}\lambda(n)\chi(n)n^{-s},$$

which is absolutely convergent in the region $\sigma := \Re(s) > 1$. It is well-known that $L(s,f,\chi)$ can be extended to an entire function. In this paper we will focus on twists by quadratic characters. For integers $d \equiv 0$, or $1 \pmod 4$ we put $\chi_d(n)=(d/n)$. So χ_d is a real character with conductor less than or equal to $|d|$. If d is a fundamental discriminant then χ_d is a primitive character of conductor $|d|$ associated with the quadratic field $\mathbb{Q}(\sqrt{d})$. If in addition $(d,2q)=1$ then the twisted L-function $L(s,f,\chi_d)$ satisfies a functional equation. More precisely we define the completed L-function by

$$\Lambda(s,f,\chi_d)=(\hat{q}|d|)^{s-1/2}\Gamma\left(s+\frac{k-1}{2}\right)L(s,f,\chi_d) \quad \text{where } \hat{q}=\frac{\sqrt{q}}{2\pi}, \tag{5}$$

then we have

$$\Lambda(s,f,\chi_d)=\varepsilon_f\chi_d(-q)\Lambda(1-s,f,\chi_d)$$

where ε_f is the sign of the functional equation of $L(s,f)$. To study 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 lth derivative at the center we have

$$\Lambda^{(l)}(1/2, f, \chi_d) = (1 + (-1)^l \varepsilon_f \chi_d(-q)) \sum_{n=1}^{\infty} \frac{\lambda(n)\chi_d(n)}{\sqrt{n}} V\left(\frac{n}{\hat{q}|d|}\right), \tag{6}$$

where the smooth function V is given by

$$V(y) = \frac{1}{2\pi i} \int_{(3)} \left(\frac{\Gamma(s+k/2)}{y^s}\right)^{(l)} \frac{ds}{s}.$$

This function decays rapidly as y gets larger; in fact we have

$$V(y) \ll_N y^{-N}$$

for all $N \geqslant 1$. This implies that in (6) 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

$$V(y) = \Gamma(k/2)(-\log y)^l + P(-\log y) + O(y),$$

where P is a polynomial of degree at most $l-1$. This fact will play an important role in our analysis. Observe that this implies that the beginning terms in the sum in (6) are weighted by $(\log |d|)^l$, whereas near the end, i.e. when $n \sim |d|$, the terms are unweighted. This gives us the bound (14) which is crucial in the analysis of the error term in § 4.

We are interested in the sum

$$S = \sideset{}{^\flat}\sum_{(d,2q)=1} r(d)\Lambda^{(l)}(1/2, f, \chi_d)F(d/Y), \tag{7}$$

where $r(d)$ denotes the number of representation of d as a sum of two squares, and F is a non-negative smooth function with compact support. (The notation $\sum^\flat$ indicates that the sum is over square-free integers.) Notice that only square-free positive odd integers $d \equiv 1 \pmod 4$ contribute to the sum S. Consequently we have $\chi_d(-q) = \chi_d(q)$. Using the approximate functional equation we can write

$$S = S_1 + (-1)^l \varepsilon_f S_2$$

where

$$S_1 = \sideset{}{^\flat}\sum_{(d,2q)=1} r(d) \sum_{n=1}^{\infty} \frac{\lambda(n)\chi_d(n)}{\sqrt{n}} G_n(d/Y),$$

$$S_2 = \sideset{}{^\flat}\sum_{(d,2q)=1} r(d)\chi_d(q) \sum_{n=1}^{\infty} \frac{\lambda(n)\chi_d(n)}{\sqrt{n}} G_n(d/Y),$$

and

$$G_n(y) = V\left(\frac{n}{\hat{q}yY}\right)F(y).$$

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

$$\sideset{}{^\flat}\sum_{(d,2q)=1} r(d)\chi_d(a) \sum_{n=1}^{\infty} \frac{\lambda(n)\chi_d(n)}{\sqrt{n}} G_n(d/Y), \tag{8}$$

https://doi.org/10.1112/S0010437X10004732 Published online by Cambridge University Press

for a given odd positive integer a (which in our case is either 1 or q). For simplicity we will assume that a is square-free. Observe that since $d \equiv 1 \pmod 4$ we can use reciprocity to write $\chi_d(a) = (a/d)$ and $\chi_d(n) = (n/d)$ in terms of the Jacobi symbol. Since the weight $r(d)$ is given by the convolution $r(d) = 4\sum_{\delta|d}(-1/\delta)$, we get that (8) is given by

$$4 \sum_{\substack{d_1,d_2\\(d_1d_2,2q)=1}} \mu(d_1d_2)^2 \left(\frac{-a}{d_1}\right)\left(\frac{a}{d_2}\right)\sum_n \frac{\lambda(n)}{\sqrt{n}}\left(\frac{n}{d_1d_2}\right)G_n(d_1d_2/Y). \tag{9}$$

To free the sum over d_1 and d_2 from the coprimality and square-free conditions we use the standard method involving the Möbius function. This leads us to consider sums of the form

$$T = \sum_{d_1 \text{ odd}}\sum_{d_2 \text{ odd}} \left(\frac{-a}{d_1}\right)\left(\frac{a}{d_2}\right)\sum_n \frac{\lambda(n)}{\sqrt{n}}\left(\frac{n}{rd_1d_2}\right)G_n(rd_1d_2/Y),$$

where r is a small positive odd integer; in fact we will have $r \ll Y^\delta$ for any small $\delta > 0$. The larger values of r contribute a negligible amount. To see why this is so, observe that $r = r_1r_2^2$ where r_1 is essentially a divisor of q^2 and r_2 is the variable that is unbounded. Then one uses the Cauchy–Schwarz inequality and the upper bound

$$\sum_d |L(1/2, f, \chi_d)|^2 F(d/Y) \ll Y^{1+\varepsilon}$$

for the second moment (where the sum is over square-free odd integers d) and the fact that $\sum_{r>A} r^{-2} \ll A^{-1}$. This upper bound for the second moment follows easily from the mean value estimate given in [Hea95] for real characters.

Next we separate the variables d_1 and d_2 from each other using an inverse Mellin transform. To this end we define the Mellin transform

$$H_n(s) = \int_0^\infty G_n(y)y^s \frac{dy}{y}.$$

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 $N \geqslant 1$. In fact we have the bound

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

Now the inverse Mellin transform yields

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

for any σ. Taking $\sigma > 1$ and replacing in the expression for T, and interchanging the order of summations and integration we get

$$T = \frac{1}{2\pi i}\int_{(\sigma)} \left(\frac{Y}{r}\right)^s \left\{\sum_n \frac{\lambda(n)H_n(s)}{\sqrt{n}}\left(\frac{n}{r}\right)L(s, \chi_{4an})L(s, \chi_{-4an})\right\} ds. \tag{10}$$

Observe that the inner sum over n 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

$$T = M + E$$

https://doi.org/10.1112/S0010437X10004732 Published online by Cambridge University Press

where M is coming from the residue at $s=1$ and, as we will see in the next section, it 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

$$E \ll \int_{(1/2)} \sqrt{\frac{Y}{r}} \left\{ \sum_{n \ll Y^{1+\varepsilon}} \frac{|\lambda(n)||H_n(s)|}{\sqrt{n}} |L(s,\chi_{4an})||L(s,\chi_{-4an})| \right\} |ds|.$$

3. The main term

In this section we analyze the main term,

$$M = \sum_n \frac{\lambda(n)}{\sqrt{n}} \left(\frac{n}{r}\right) \mathrm{Res}_{s=1} \left\{ H_n(s) L(s,\chi_{4an}) L(s,\chi_{-4an}) \left(\frac{Y}{r}\right)^s \right\}.$$

The L-function $L(s,\chi_{-4an})$ is entire for any n, whereas the other L-function $L(s,\chi_{4an})$ has a simple pole at $s=1$ for $an=\square$. Hence it follows that

$$M = \frac{Y}{r} \sum_{\substack{n \\ an=\square}} \frac{\lambda(n) H_n(1)}{\sqrt{n}} \left(\frac{n}{r}\right) L(1,\chi_{-4an})\ \mathrm{Res}_{s=1} L(s,\chi_{4an}).$$

The residue is given by

$$\mathrm{Res}_{s=1} L(s,\chi_{4an}) = \prod_{p|2an} \left(1-\frac{1}{p}\right).$$

If a is square-free then the condition $an=\square$ is equivalent to $n=am^2$ for some m, from which we get

$$M = \frac{Y}{r} \frac{L(1,\chi_{-4})}{\sqrt{a}} \left(\frac{a}{r}\right) \sum_{(m,r)=1} \frac{\lambda(am^2) H_{am^2}(1)}{m} \prod_{p|2am} \left(1-\frac{1}{p}\right)\left(1-\frac{\chi_{-4}(p)}{p}\right).$$

We will write l_p for the local L-factor $(1-p^{-1})(1-\chi_{-4}(p)p^{-1})$. We seek to evaluate the above sum using contour integration. To this end recall that $H_n(1)$ is given by

$$\begin{aligned} H_n(1) &= \int_0^\infty V\left(\frac{n}{\hat{q}yY}\right) F(y)\, dy \\ &= \frac{1}{2\pi i} \int_0^\infty \int_{(3)} \left(\frac{\Gamma(s+k/2)(\hat{q}yY)^s}{n^s}\right)^{(l)} \frac{ds}{s} F(y)\, dy. \end{aligned}$$

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

$$D(s) = \sum_{(m,r)=1} \frac{\lambda(am^2) \prod_{p|2am} l_p}{a^{s+1/2} m^{2s+1}}$$

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

$$M = \frac{YL(1,\chi_{-4})}{r} \left(\frac{a}{r}\right) \frac{1}{2\pi i} \int_0^\infty \int_{(3)} (\Gamma(s+k/2)(\hat{q}yY)^s D(s))^{(l)} \frac{ds}{s} F(y)\, dy. \qquad (11)$$

https://doi.org/10.1112/S0010437X10004732 Published online by Cambridge University Press

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 $D(s)$. First using the multiplicativity of the Fourier coefficient we get

$$D_{\text{odd}}(s) = \prod_{p|a}\left\{\frac{\lambda(p^3)l_p}{(p^3)^{s+1/2}} + \frac{\lambda(p^5)l_p}{(p^5)^{s+1/2}} + \cdots\right\} \prod_{p\nmid 2ar}\left\{1 + \frac{\lambda(p^2)l_p}{(p^2)^{s+1/2}} + \frac{\lambda(p^4)l_p}{(p^4)^{s+1/2}} + \cdots\right\},$$

where $D_{\text{odd}}(s)$ is the Dirichlet series obtained from $D(s)$ by removing the Euler factor at $p=2$. Hence the Dirichlet series $D(s)$ is related to the symmetric square L-function of the form f. 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

$$D(s) = L(2s+1, \mathrm{Sym}^2 f)D_2(s).$$

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 $L(s, \mathrm{Sym}^2 f)$ has analytic continuation to the whole of the complex plane and it does not vanish at $s=1$. Hence $D(0) \neq 0$.

We use the expansion

$$Y^s = 1 + s\log Y + \frac{(s\log Y)^2}{2!} + \cdots$$

to conclude that the lth coefficient in the power series expansion of the holomorphic function $\Gamma(s+k/2)(4\hat{q}yY)^s D(s)$ around the origin is given by a polynomial in $\log Y$ of degree l,

$$c_l = (l!)^{-1}\Gamma(k/2)(\log Y)^l D(0) + \text{lower degree terms in} \log Y.$$

Hence the residue of the integrand in (11) at $s=0$ is also a polynomial in $\log Y$ of degree l, given by

$$\mathrm{Res}_{s=0} s^{-1}(\Gamma(s+k/2)(\hat{q}yY)^s D(s))^{(l)} = l!c_l.$$

By contour integration we get

$$M = \frac{D(0)}{r}\left(\frac{a}{r}\right)\Gamma(k/2)L(1,\chi_{-4})Y(\log Y)^l \int_0^\infty F(y)\,dy + YQ(\log Y) + R, \tag{12}$$

where Q is a polynomial of degree $l-1$, and the remainder term R is given by the same integral as in (11), but now over the vertical line $\sigma = -1/4+\varepsilon$. Hence it follows from standard estimates that

$$R \ll \frac{Y^{3/4}}{r}(Yr)^\varepsilon.$$

The first two terms in the right of the expression (12) give the the leading term in Theorem 1. Also observe that the leading term in M is of size $Y(\log Y)^l$ and the coefficient is given by a nice formula. In particular the coefficient is a multiplicative function in r. So we can easily execute the sum over the variable r. Recall that the variable r appears when we use the Möbius function to deal with the coprimality and the square-free conditions in the sum (8).

4. The error term

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

$$E \ll \int_{(1/2)}\sqrt{\frac{Y}{r}}\left\{\sum_{n\ll Y^{1+\varepsilon}} \frac{|\lambda(n)||H_n(s)|}{\sqrt{n}}|L(s,\chi_{4an})||L(s,\chi_{-4an})|\right\}|ds|. \tag{13}$$

https://doi.org/10.1112/S0010437X10004732 Published online by Cambridge University Press

Notice that we have already sacrificed the hope of getting any cancellation from the sign changes of the Fourier coefficient $\lambda(n)$. Experience shows that such a desire is insatiable because we do not have any information about how the L-functions behave inside the critical strip. The only way to get cancellation would be to use an approximate functional equation to open the L-functions and in the process lose $Y^{1/4}$ each for the two L-functions. Then of course the cancellation in the n-sum gives us back this loss, but nothing more. So we are again back to where we started. In short, in a double sum of equal lengths with joint oscillation one may not hope to get more than one-fourth cancellation (i.e. one-half cancellation from one sum of one's choice).

Owing 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 $N \leqslant n < 2N$ we will use the bound $H_n(s) \ll H(N)(1+|t|)^{-4}$, where

$$H(N) = \begin{cases} 1+\log^l(Y/N) & \text{if } N \leqslant Y, \\ (Y/N) & \text{if } Y < N \ll Y^{1+\varepsilon}. \end{cases} \tag{14}$$

Then using the Cauchy–Schwarz inequality we get that

$$E \ll \sqrt{Y/r} \sum_{\text{dyadic blocks}} H(N) N^{-1/2} \int_{\mathbb{R}} U(N,t)(1+|t|)^{-4}\, dt, \tag{15}$$

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

$$U(N,t) = \sum_{N \leqslant n < 2N} |\lambda(n)| \left| L\left(\frac{1}{2}+it, \chi_{4n}\right)\right|^2. \tag{16}$$

The above deduction may need some 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 (16) and a similar expression with χ_{-4n} in place of χ_{4n}. A moments reflection convinces us that 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 n as km^2 with k square-free we observe that

$$U(N,t) \ll \sum_{m \leqslant \sqrt{N}} \tau(m^2)\tau(m) \sideset{}{^\flat}\sum_{k \sim N/m^2} |\lambda(k)| \left| L\left(\frac{1}{2}+it, \chi_k\right)\right|^2. \tag{17}$$

(Note the slight abuse of notation. Indeed χ_k stands for the character associated with the field $\mathbb{Q}(\sqrt{k})$.) Here we have used the Deligne bound but this is not the crucial usage that we mentioned above. The inner sum being over square-free numbers k, we can use 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

$$W_t(y) = \frac{1}{2\pi i} \int_{(3)} G(u) \frac{\Gamma((1/4)+(it/2)+(u/2))}{\Gamma((1/4)+(it/2))} (\sqrt{\pi} y)^{-u} \frac{du}{u} \tag{18}$$

where $G(u)$ is any function which is holomorphic in the strip $-4 < \Re(u) < 4$, even and normalized by $G(0)=1$ (see [IK04, Theorem 5.3]).

Let

$$U^*(N,t) = \sideset{}{^\flat}\sum_{N\leqslant n<2N} |\lambda(n)| \left| \sum_d \frac{\chi_n(d)}{d^{1/2+it}} W_t\left(\frac{d}{\sqrt{n}}\right)\right|^2.$$

Notice that in the inner sum, terms up to $d \ll \sqrt{N(1+|t|)}$ contribute to the sum, while the other terms of the sum make a negligible contribution. (This follows from the exponential decay of the function W_t (see [IK04]).) Our job is now reduced to producing a decent upper bound for $U^*(N,t)$. An application of Deligne bound at this point yields

$$U^*(N,t) \ll \sideset{}{^\flat}\sum_{N\leqslant n<2N} \tau(n) \left| \sum_d \frac{\chi_n(d)}{d^{1/2+it}} W_t\left(\frac{d}{\sqrt{n}}\right)\right|^2.$$

Observe that the expression in the right-hand side of the above inequality is actually comparable with the original sum (7); the only gain that we have accomplished is that we no more have the lth derivative, i.e. we have saved a power of $\log Y$. If we now open the absolute square, interchange the order of summation and try to evaluate the sum over n using a summation formula, we will face the same problem that we faced before, and the dual sum is roughly the same as the original sum. This happens as the modulus of the character $(\frac{\cdot}{d_1 d_2})$ in the worst case scenario is as large as $N(1+|t|)$, which is roughly the length of the n-sum. This is the deadlock case in the twisted Voronoi summation formula as we have mentioned in the introduction. To break this deadlock 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.

We use the following inequality (see [IM10]) for the multiplicative function $|\lambda(n)|$ for square-free n

$$|\lambda(n)| \ll \sideset{}{^\flat}\sum_{\substack{a\mid n\\ a\leqslant n^{1/7}}} |\lambda(a)|^3. \tag{19}$$

For the sake of completeness we include a proof of the above inequality. The starting point is the following simple combinatorial inequality:

$$\tau_7(n) = \sum_{d_1\ldots d_7=n} 1 \leqslant 7 \sum_{\substack{d_1\ldots d_7=n\\ d_1\leqslant n^{1/7}}} 1 = 7 \sum_{\substack{d_1\mid n\\ d_1\leqslant n^{1/7}}} \tau_6\left(\frac{n}{d_1}\right) = 7\tau_6(n) \sum_{\substack{d_1\mid n\\ d_1\leqslant n^{1/7}}} \frac{1}{\tau_6(d_1)}.$$

(Recall that $\tau_k(n) = \sum_{d_1\ldots d_k=n} 1$.) We can assume that $|\lambda(p)|>1$ for all primes $p|n$, otherwise $|\lambda(n)| \leqslant |\lambda(n/p)|$ and it is enough to prove (19) for n/p. Now we set $V(d)=|\lambda(d)|^3$ for any $d|n$, and apply Hölder's inequality to obtain

$$\sum_{\substack{d\mid n\\ d\leqslant n^{1/7}}} \frac{V(d)^{1/(t+1)}}{\tau_6(d)V(d)^{1/(t+1)}} \leqslant \left[\sum_{\substack{d\mid n\\ d\leqslant n^{1/7}}} V(d)\right]^{1/(t+1)} \left[\sum_{d\mid n} \frac{1}{\tau_6(d)^{(t+1)/t} V(d)^{1/t}}\right]^{t/(t+1)}$$

for any $t>0$. Combining the above two inequalities we get

$$\left[\frac{\tau_7(n)}{\tau_6(n)}\right]^{t+1} \left[\sum_{d\mid n} \frac{1}{\tau_6(d)^{(t+1)/t} V(d)^{1/t}}\right]^{-t} \leqslant 7^{t+1} \left[\sum_{\substack{d\mid n\\ d\leqslant n^{1/7}}} V(d)\right].$$

Observe that the left-hand side is multiplicative. Therefore, to establish (19) we have to show that there is a $t>0$ (not depending on p) such that

$$|\lambda(p)| \leqslant \left[\frac{7}{6}\right]^{t+1}\left[1+\frac{1}{6^{(t+1)/t}V(p)^{1/t}}\right]^{-t}.$$

Since $1<|\lambda(p)|\leqslant 2$ (using Deligne's bound), we are only required to show that there is a $t>0$ such that

$$\frac{1}{x} \leqslant \left[\frac{7}{6}\right]^{t+1}\left[1+\frac{x^{3/t}}{6^{(t+1)/t}}\right]^{-t}$$

for all $x\in[1/2,1]$. The above inequality is equivalent to

$$1 \leqslant \frac{7x}{6}\left[1+\frac{1}{6}\left(1-\frac{x^{2/t}}{7^{1/t}}\right)\right]^{t}.$$

As $t\to\infty$ the right-hand side of the above expression converges uniformly to $(7x/6)(7/x^2)^{1/6}$, which is larger than 1 for $x\geqslant 1/2$. This concludes the proof of (19).

The inequality (19) is of outmost importance in this paper and so Deligne's bound is absolutely necessary. (Also note that the exponent 3 in the right-hand side of (19) is not optimum, and there is a slight room of improvement (see [IM10]).) 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

$$U^*(N,t) \ll \sideset{}{^\flat}\sum_{a\leqslant N^{1/7}} |\lambda(a)|^3 U^*(N,t,a) \tag{20}$$

where

$$U^*(N,t,a)=\sum_b\left|\sum_d \frac{1}{d^{1/2+it}}\left(\frac{ab}{d}\right)W_t\left(\frac{d}{\sqrt{ab}}\right)\right|^2 K\left(\frac{ab}{N}\right).$$

We will evaluate the above sum using the Poisson summation formula.

Now opening the absolute square of the inner sum and interchanging the order of summation we get

$$U^*(N,t,a)=\sum_{d_1,d_2}\frac{1}{d_1^{1/2+it}d_2^{1/2-it}}\left(\frac{a}{d_1d_2}\right)\sum_b\left(\frac{b}{d_1d_2}\right)J_{d_1,d_2,t}\left(\frac{ab}{N}\right),$$

where

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

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

$$U^*(N,t,a)=\frac{N}{a}\sum_k\sum_{d_1,d_2}\frac{g(k,d_1d_2)}{d_1^{3/2+it}d_2^{3/2-it}}\left(\frac{a}{d_1d_2}\right)\check{J}_{d_1,d_2,t}\left(\frac{kN}{ad_1d_2}\right),$$

where $\check{J}_{d_1,d_2,t}$ stands for the Fourier transform

$$\check{J}_{d_1,d_2,t}(y)=\int_{\mathbb{R}} J_{d_1,d_2,t}(x)e(-xy)\,dx,$$

and

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

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 giving an upper bound for the principal term,

$$P(N, t, a) = \frac{N}{a} \sum_{d_1, d_2} \frac{g(0, d_1d_2)}{d_1^{3/2+it} d_2^{3/2-it}} \left(\frac{a}{d_1d_2}\right) \breve{J}_{d_1,d_2,t}(0).$$

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

$$P(N, t, a) \ll \frac{N}{a} \sum_{\substack{d_1, d_2 \\ d_1d_2 = \square \\ (d_1d_2, a)=1}} \frac{\phi(d_1d_2)}{d_1^{3/2} d_2^{3/2}} |\breve{J}_{d_1,d_2,t}(0)|. \tag{21}$$

Now using an appropriate test function $G(u)$ in (18) we can show that for any $A > 0$ we have

$$\breve{J}_{d_1,d_2,t}(0) \ll \left(1 + \frac{d_1}{\sqrt{N(1+|t|)}}\right)^{-A} \left(1 + \frac{d_2}{\sqrt{N(1+|t|)}}\right)^{-A}. \tag{22}$$

This implies that essentially the terms in (21) with $d_1, d_2 \ll \sqrt{N(1+|t|)}$ contribute to the sum and the other terms make a negligible contribution. Replacing the above bound for the Fourier transform in (21) and executing the sum over d_1 and d_2 we obtain

$$P(N, t, a) \ll \frac{N(\log N)^3}{a} (1+|t|)^2.$$

Next we compute the contribution of the terms with $ak = \square$ to $U^*(N, t, a)$. This is given by

$$P_\square(N, t, a) = \frac{N}{a} \sum_{\substack{ak=\square \\ k \neq 0}} \sum_{d_1, d_2} \frac{g(k, d_1d_2)}{d_1^{3/2+it} d_2^{3/2-it}} \left(\frac{a}{d_1d_2}\right) \breve{J}_{d_1,d_2,t}\left(\frac{kN}{ad_1d_2}\right).$$

As we do not expect to get any cancellation in this case, we will evaluate the sum trivially. Writing each d_i as l_id_i with $l_i | k^\infty$ and $(d_i, k) = 1$, we see that the above sum is dominated by

$$\frac{N}{a} \sum_{\substack{ak=\square \\ k \neq 0}} \sum_{\substack{l_1, l_2 | k^\infty \\ (l_1l_2, a)=1}} \frac{|g(k, l_1l_2)|}{(l_1l_2)^{3/2}} \sum_{\substack{d_1, d_2 \\ (d_1d_2, ak)=1}} \frac{|g(1, d_1d_2)|}{(d_1d_2)^{3/2}} \left| \breve{J}_{l_1d_1, l_2d_2, t}\left(\frac{kN}{al_1d_1l_2d_2}\right) \right|. \tag{23}$$

(The notation $a | b^\infty$ means that every prime divisor of a is a prime divisor of b, or equivalently a divides some power of b.) We need an estimate for the Fourier transform. The bound given in (22) is not enough, as we need to have a bound which restricts the contribution from the k-sum. We derive from the analytic nature of the Fourier transform $\breve{J}_{d_1,d_2,t}$ that frequencies k of large size, namely $|k| \gg aN^\varepsilon$, can be ignored from analysis as they make a negligible contribution. For smaller non-zero frequencies we use integration by parts (as $k \neq 0$) to obtain

$$\breve{J}_{d_1,d_2,t}\left(\frac{kN}{ad_1d_2}\right) \ll \frac{ad_1d_2}{kN} \left(1 + \frac{d_1}{\sqrt{N(1+|t|)}}\right)^{-A} \left(1 + \frac{d_2}{\sqrt{N(1+|t|)}}\right)^{-A}.$$

Also we have square-root cancellation in the Gauss sum, i.e.,

$$g(1, d_1d_2) \ll \sqrt{d_1d_2}.$$

https://doi.org/10.1112/S0010437X10004732 Published online by Cambridge University Press

Using the above bounds we can now evaluate the inner sum in (23). The sum over l_i can be evaluated trivially. Finally to execute the sum over k we use the observation that the condition $ak = \square$ is equivalent to having $k = am^2$ for some $m \neq 0$ (as a is square-free). It follows that the contribution from the non-zero square frequencies is smaller in magnitude than the contribution from the zero frequency, and we get

$$P(N,t,a) + P_{\square}(N,t,a) \ll \frac{N(\log N)^3}{a}(1+|t|)^2.$$

5. The remainder terms

It remains for us to evaluate the contribution of the remaining frequencies. It turns out that we cannot estimate this trivially as we have done for the square frequencies. Of course, if we do so then we are off by an amount aN^{ε}. This is the amount we have to save utilizing the cancellation in the sum over d_i, as in the present case we have oscillation in these sums. Also the length of the sum over any of the d_i is roughly $\sqrt{N}$ which is satisfactory compared to the modulus of oscillation. Now since the frequencies k with $|k| > \kappa := a(1+|t|)^2 N^{\varepsilon}$ make a negligible contribution, our job reduces to estimating

$$R(N,t,a) = \frac{N}{a} \sum_{\substack{ak \neq \square \\ 0<|k|<\kappa}} \sum_{d_1,d_2} \frac{g(k,d_1d_2)}{d_1^{3/2+it} d_2^{3/2-it}} \left(\frac{a}{d_1d_2}\right) \breve{J}_{d_1,d_2,t}\left(\frac{kN}{ad_1d_2}\right). \tag{24}$$

We write $d_1 = ln$ with $l|(2kd_2)^{\infty}$ and $(n, 2kd_2) = 1$. We want to estimate

$$\sum_{(n,2kd_2)=1} \frac{g(1,n)}{n^{3/2+it}} \left(\frac{ak}{n}\right) \breve{J}_{nl,d_2,t}\left(\frac{kN}{anld_2}\right) \tag{25}$$

using partial summation. Recall that $g(1,n) = 0$ if n is not square-free and $g(1,n) = \sqrt{n}$ if n is square-free. Now we will use the Polya–Vinogradov inequality or the convexity bound for the Dirichlet L-function to evaluate the character sum over square-free integers. This gives us the inequality

$$\sideset{}{^\flat}\sum_{n \leqslant y} \left(\frac{akd_2^2}{n}\right) \ll (ak)^{1/4} y^{1/2} (akd_2)^{\varepsilon}.$$

(Of course one may use Burgess' bound instead of the Polya–Vinogradov inequality. This will lead to a longer allowable range for the a-sum in (19). In fact, instead of $a \leqslant n^{1/7}$ we will have $a \leqslant n^{1/6}$, the optimum value of the exponent for this range is given by $\eta = 2.099\ldots$ and hence we will be able to save a fractional power of $\log Y$ when we use the Rankin bound below.) Also, for the smooth function

$$f(y) = y^{-1-it} \breve{J}_{yl,d_2,t}\left(\frac{kN}{yald_2}\right)$$

we have the bound for the derivative,

$$f'(y) \ll \frac{ald_2}{ykN}\left(1 + \frac{ly}{\sqrt{N(1+|t|)}}\right)^{-A}\left(1 + \frac{d_2}{\sqrt{N(1+|t|)}}\right)^{-A}(1+|t|),$$

so it follows using partial summation that the sum in (25) is dominated by

$$\frac{a^{5/4} l^{1/2} d_2}{k^{3/4} N^{3/4}} (1+|t|)^{5/4} \left(1 + \frac{d_2}{\sqrt{N(1+|t|)}}\right)^{-A} (akd_2)^{\varepsilon}.$$

https://doi.org/10.1112/S0010437X10004732 Published online by Cambridge University Press

Plugging this into expression (24) we get

$$R(N,t,a) \ll a^{1/4}N^{1/4}(1+|t|)^{5/4} \sum_{0<|k|<\kappa} \sum_{d_2} \sum_{l|(2kd_2)^\infty} \frac{|g(k,ld_2)|}{ld_2^{1/2}k^{3/4}} \left(1+\frac{d_2}{\sqrt{N(1+|t|)}}\right)^{-A} (akd_2)^\varepsilon.$$

We now use the standard explicit bound for the Gauss sum to execute the summation over l, d_2 and k. We obtain

$$R(N,t,a) \ll a^{1/2}N^{3/4}(1+|t|)^3(aN)^\varepsilon.$$

We have therefore established that

$$\begin{aligned} U^*(N,t,a) &\ll P(N,t,a) + P_\square(N,t,a) + R(N,t,a) \\ &\ll \frac{N}{a}(1+|t|)^3(\log N)^3 + a^{1/2}N^{3/4}(1+|t|)^3(aN)^\varepsilon. \end{aligned}$$

We now substitute this bound into expression (20). The second term after the sum over a is dominated by $N^{27/28+\varepsilon}(1+|t|)^3$, which is negligible. Now we will use the following bound due to Rankin [Ran95]:

$$\sum_{a\leqslant y} |\lambda(a)|^3 \ll y(\log y)^{\sqrt{2}-1}.$$

To have a neat expression we will replace the upper bound by $y\sqrt{\log y}$. Then using partial summation it follows that

$$U^*(N,t) \ll N(1+|t|)^3(\log N)^{9/2}.$$

Now from the expression given in (17) we get that

$$\begin{aligned} U(N,t) &\ll \sum_{m\leqslant\sqrt{N}} \tau(m^2)\tau(m)U^*(N/m^2,t) \\ &\ll N(1+|t|)^3(\log N)^{9/2} \sum_{m\leqslant\sqrt{N}} \frac{\tau(m^2)\tau(m)}{m^2} \\ &\ll N(1+|t|)^3(\log N)^{9/2}. \end{aligned}$$

Replacing this bound in (15) and computing the integral over t, we get that the error term

$$E \ll \sqrt{Y/r} \sum_{\text{dyadic blocks}} H(N)\sqrt{N}(\log N)^{9/2} \ll \frac{Y(\log Y)^{9/2}}{\sqrt{r}}.$$

Observe that we do not lose an extra $\log Y$ for breaking the sum over n into dyadic segments.

From the above bound we can derive the claimed bound for $E(Y)$ in Theorem 1. Recall that $r=r_1r_2^2$ with $r_1|q^2$ and hence bounded so the sum over r_1 does not alter the size of the error term. However, the variable r_2 is a product of three independent variables, $r_2=l_1l_2l_3$, where l_1 and l_2 are the variables used to free the sum (9) from the square-free condition on d_1 and d_2 respectively. The third variable l_2 was introduced to free (9) from the coprimality condition $(d_1,d_2)=1$. Therefore, when we do the sum over r_2 we will lose $(\log Y)^3$. Hence

$$E(Y) \ll Y(\log Y)^{7+\frac{1}{2}}.$$

Notice that if we had used Burgess' bound instead of the Polya–Vinogradov inequality we would have got $E(Y) \ll Y(\log Y)^{7.07\cdots}$, and so the integral part of the exponent would remain unchanged. Of course there are many possibilities of improvements as we indicate in the next section.

6. Concluding remarks

In this section we briefly indicate how we may obtain a better bound for the error term $E(Y)$ using a more delicate analysis. There is lots of room for improvement and in several points. First when we use the Möbius function to free the sum in (8) from the coprimality and the square-free conditions, and obtain the expression in (9), we are forfeiting one $\log Y$ for each of the Möbius functions. However, this is not essential and we can continue our analysis in § 2 with

$$T = \frac{1}{2\pi i}\int_{(\sigma)} Y^s\bigg\{\sum_n \frac{\lambda(n)H_n(s)}{\sqrt{n}} L(s,\chi_{4an})L(s,\chi_{-4an})\zeta(2s)^{-3}D_n(s)\bigg\}\,ds,$$

instead of the expression (10). Here $D_n(s)$ is a Dirichlet series which has an Euler product expression and converges absolutely for $\sigma > 1/3$. Observe that this minor change does not hamper the shift of the contour integral to $\sigma = 1/2$. Of course the analysis of the main term in § 3 goes through with some minor changes, $D_n(s)$ being quite amiable.

In our analysis of the error term, already in the expression on the right of (13) we have lost a power of Y. Indeed we are not trying to get a cancellation in the sum over n in E. However, a closer analysis shows that we really do not need the b-sum (in (20)) to be that long (the level of distribution for a GL_2 L-function is at least 1/2 (see [Mun09])), so we can use the following inequality:

$$|\lambda(n)||D_n(1/2+it)| \ll \sideset{}{^\flat}\sum_{\substack{a\mid n\\ a\leqslant n^{1/4}}} |\lambda(a)|\prod_{p\mid a}\bigg(1+\frac{1}{p}\bigg)^3.$$

This can be proved in the same spirit as (19) coupled with the observation that $|D_n(1/2+it)| \ll \prod_{p\mid n}(1+1/p)^3$. In this case Hölder's inequality together with Rankin's mean-value result [Ran95] gives

$$\sum_{a\leqslant y}|\lambda(a)|\prod_{p\mid a}\bigg(1+\frac{1}{p}\bigg)^3 \ll \bigg[\sum_{a\leqslant y}|\lambda(a)|^{3/2}\bigg]^{2/3}\bigg[\sum_{a\leqslant y}\prod_{p\mid a}\bigg(1+\frac{1}{p}\bigg)^9\bigg]^{1/3} \ll y(\log y)^{2F(3/4)/3},$$

where

$$F(3/4) = \frac{2^{-1/4}}{5}(2^{3/4}+3^{5/4}) - 1 = -0.053\ldots.$$

Hence, instead of losing a power of $\log Y$ in the sum over a, we will in fact gain a fractional power of $\log Y$. This will of course come as a boon when we are right at the threshold and need to save a fractional power of $\log Y$.

Also, one can show that for t away from zero, we have

$$\sum_{N\leqslant n<2N}\bigg|L\bigg(\frac{1}{2}+it,\chi_{4n}\bigg)\bigg|^2 \ll N(\log N).$$

(Of course the uniformity of the bound in the t-aspect is an issue here. Indeed for any fixed $t\neq 0$ we can take a smooth version of the sum appearing on the left-hand side and evaluate it asymptotically. The leading term is of the form $a(N,t)N\log N + b(N,t)N$ where the functions $a(N,t)$ and $b(N,t)$ contain oscillatory terms, comparable with $(N^{it}-1)/t$, which as $t\to 0$ yields extra powers of $\log N$. It follows that for $t \ll (\log N)^{-1}$ the true order of magnitude is $N(\log N)^3$. We take up this analysis in detail in the forthcoming work [Mun] where we establish Theorem 1 with an error term $O(Y(\log Y)^{9/10})$ in the special case of dihedral form f.) Hence it will follow if we choose to be very careful that

$$U(N,t) \ll N(1+|t|)^3(\log N)^{2-\delta},$$

for some δ in the range $(0,1)$ and $t \neq 0$, and the final bound will turn out to be

$$E(Y) \ll Y(\log Y)^{2-\delta}. \tag{26}$$

Therefore, we will have a non-vanishing result from the second derivative onwards. Of course the most interesting case is the first derivative, as it is related to the existence of fibers of rank one in the quadratic elliptic fibration $q(t)y^2 = f(x)$.

It seems that (26) is the limit of the method and any improvement over this will require new ideas. One may conjecture that the right-hand side of (13) is bounded by $O(Y(\log Y)^{1-\delta})$ for some $\delta \in (0,1)$. Indeed, this is what we get if we replace each of the L-functions and the Fourier coefficients appearing in (13) by their respective average sizes, but unfortunately the inequalities of the type (19), which we employ to break the deadlock, cost us an extra $\log Y$.

Acknowledgements

I thank Professor Henryk Iwaniec for many enlightening conversations. I also thank the referee for many helpful suggestions.

References

CFH06 G. Chinta, S. Friedberg and J. Hoffstein, *Multiple Dirichlet series and automorphic forms*, in *Multiple Dirichlet series, automorphic forms, and analytic number theory*, Proceedings of Symposia in Pure Mathematics, vol. 75 (American Mathematical Society, Providence, RI, 2006), 3–41.

Hea95 D. R. Heath-Brown, *A mean value estimate for real character sums*, Acta Arith. **72** (1995), 235–275.

IK04 H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53 (American Mathematical Society, Providence, RI, 2004), xii+615.

IM10 H. Iwaniec and R. Munshi, *Cubic polynomials and quadratic forms*, J. Lond. Math. Soc. (2) **81** (2010), 45–64.

IS00 H. Iwaniec and P. Sarnak, *The non-vanishing of central values of automorphic L-functions and Landau–Siegel zeros*, Israel J. Math. **120** (2000), 155–177, part A.

KMV00 E. Kowalski, P. Michel and J. VanderKam, *Non-vanishing of high derivatives of automorphic L-functions at the center of the critical strip*, J. Reine Angew. Math. **526** (2000), 1–34.

Mun09 R. Munshi, *The level of distribution of the special values of L-functions*, Acta Arith. **138** (2009), 239–257.

Mun R. Munshi, *Quadratic families of CM elliptic curves*, Preprint.

MM91 M. R. Murty and V. K. Murty, *Mean values of derivatives of modular L-series*, Ann. of Math. (2) **133** (1991), 447–475.

Ran95 R. A. Rankin, *Sums of powers of cusp form coefficients II*, Math. Ann. **272** (1985), 593–600.

Ritabrata Munshi rmunshi@math.ias.edu
Department of Mathematics, Rutgers University, Hill Center, Piscataway, NJ 08854, USA

Current address: Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA