CONSEQUENCES OF THE GROSS/ZAGIER FORMULAE: STABILITY OF AVERAGE L-VALUES, SUBCONVEXITY, AND NON-VANISHING MOD p

PHILIPPE MICHEL AND DINAKAR RAMAKRISHNAN

In memory of Serge Lang

Introduction

In this paper we investigate some consequences of the Gross/Zagier type formulae which were introduced by Gross and Zagier and then generalized in various directions by Hatcher, Zhang, Kudla and others [GZ, Gro, Hat2, Zha, DZ]. Let us now recall these formulae in the classical context. Denote by K an imaginary quadratic field of discriminant $-D$ say, with associated quadratic character $\chi_{-D} = (\frac{-D}{\cdot})$, Ψ a character of the ideal class group $\mathrm{Pic}(\mathcal{O}_K)$ of K, $\mathcal{H}$ the upper half plane, and g_Ψ the weight one theta series associated with Ψ:

$$g_\Psi(z) = \sum_{m\geq 0} r_\Psi(m)q^m,\ \ q = \exp(2\pi\iota z), z \in \mathcal{H},$$

where for $m \geq 1$

$$r_\Psi(m) = \sum_{N(\mathfrak{a})=m} \Psi(\mathfrak{a})$$

and $\mathfrak{a} \subset \mathcal{O}_K$ ranging over the $\mathcal{O}_K$-ideal of norm m. We will denote the trivial character of $\mathrm{Pic}(\mathcal{O}_K)$ by 1_K.

Now let f be an holomorphic new cuspform of level N coprime with D, trivial nebentypus and weight $2k$:

$$f(z) = \sum_{m\geq 1} a_m(f)q^m.$$

Depending on how the primes dividing N split in K, the Gross/Zagier formula expresses the central value at $s = k$ (or the derivative of that value) of the Rankin-Selberg L-function

$$L(s, f, \Psi) := L(2s, \chi_{-D}) \sum_{m\geq 1} a_m(f)r_\Psi(m)m^{-s}$$

in term of an intersection/height pairing of the f-isotypic component $e_{\Psi,f}$ of a cycle e_Ψ living on some Hecke module $M = M_{k,N}$: Denoting this pairing by $\langle \cdot, \cdot \rangle_M$ and the Petersson inner product on $S_{2k}(N)$ by

$$\langle f, g\rangle = \int_{Y_0(N)} = f(z)\overline{g(z)}y^{2k-2}dxdy,$$

where $Y_0(N)$ denotes the open modular curve $\Gamma_0(N)\backslash\mathcal{H}$, one has

$$c_{k,K}\frac{L^{(i)}(k, f, \Psi)}{\langle f, f\rangle} = \langle e_{\Psi,f}, e_{\Psi,f}\rangle_M \tag{1}$$

for some constant $c_{k,K} > 0$ and the order of derivative $i = i_{K,N}$ is 0 or 1 (depending on the sign of the functional equation). Originally the formula was proven as follows (for $i = 0$): let $M_{2k}(N)$ (resp. $S_{2k}(N)$) denote the space of holomorphic forms (resp. cusp forms) of weight $2k$ level N and trivial nebentypus. The map

$$f \mapsto L(s, f, \Psi)$$

being linear on $S_{2k}(N)$, can be represented by a kernel $f \mapsto \langle f, G_\Psi\rangle$ for some $G_\Psi \in M_{2k}(N)$ (same for the first derivative). By the Rankin-Selberg theory

$$L(k, f, \Psi) = \int_{Y_0(N)} f(z)g_\Psi(z)E_{2k-1}(z)y^{(2k+1)/2-2}dxdy$$

for a suitable holomorphic Eisenstein series E_{2k-1} of weight $2k-1$. The determination of G_Ψ amounts to first taking the trace from level $N' = lcm(4, N)$ to N, and then computing the projection of $g_\Psi(z)E_{2k-1}(z)$ on $M_{2k}(N)$. This can be done and one infers from the computation of the Fourier expansion of $g_\Psi(z)E_{2k-1}(z)$, that the Fourier coefficients $a_m(G_\Psi)$ of G_Ψ are relatively elementary expressions involving the arithmetical functions r_Ψ and variants thereof: see below for an example. One the other hand, using the theory of complex multiplication, Gross and Zagier, and subsequently other people, showed by an auxiliary computation that

$$G_\Psi(z) = a_0(G_\Psi) + \sum_{m\geq 1}\langle T_m e_\Psi, e_\Psi\rangle_M q^m$$

where T_m denote the m-th Hecke operator acting on the module M. The final result follows then from a formal argument involving the multiplicity one theorem. The main observation underlying this paper is that the above computation provides formally an expression for the *average* of the central values $L(k, f, \Psi)$. Namely, if $S_{2k}^{new}(N)$ denote the set of arithmetically normalized new forms, then $\{f/\langle f, f\rangle^{1/2}\}_{f\in S_{2k}^{new}(N)}$

may be completed to an orthonormal basis of $S_{2k}(N)$. Then decomposing G_Ψ along such an orthonormal basis, and taking the m-th Fourier coefficient in the above decomposition, one deduces, for any $m \geq 1$,

$$\sum_{f\in S_{2k}^{new}(N)} \frac{L(k,f,\Psi)}{\langle f,f\rangle} a_m(f) \;=\; a_m(G_\Psi) + +\mathcal{A}_{\mathrm{old}}(m) + \mathcal{A}_{\mathrm{Eis}}(m),$$

where $\mathcal{A}_{\mathrm{old}}(m)$, resp. $\mathcal{A}_{\mathrm{Eis}}(m)$, is the contribution from the old forms, resp. the Eisenstein series, of weight $2k$ and level N. In principle, the Eisenstein series contribution could be evaluated explicitly, while the old forms contribution could be computed by induction on N by following the same scheme, though there is an added complication of finding a suitable orthonormal basis. We shall consider here the nicest possible situation for which these additional contributions have a particularly simple expression, in fact where the old part vanishes! Therefore we obtain, by the first step of the proof of the Gross/Zagier type formulae, a simple expression for the first moment

$$\sum_{f\in S_{2k}^{new}(N)} \frac{L(k,f,\Psi)}{\langle f,f\rangle} a_m(f).$$

Let us turn to a more specific example. Set $h = h_K = |\mathrm{Pic}(\mathcal{O}_K)|$, the class number of K, $u = |\mathcal{O}_K^\times/\{\pm 1\}|$, and

$$R(m) := \begin{cases} h/2u, & m = 0 \\ \sum\limits_{\substack{\mathfrak{a}\subset\mathcal{O}_K \\ N(\mathfrak{a})=m}} 1, & m \geq 1 \end{cases},$$

Moreover extend, for any ideal class group character Ψ, the definition of $r_\Psi(m)$ to $m = 0$ by setting

$$r_\Psi(0) = \begin{cases} 0, & \text{if } \Psi \neq 1_K \\ h/2u, & \text{if } \Psi = 1_K. \end{cases}$$

We also set

$$\sigma_N(m) = \sum_{\substack{d|m \\ (d,N)=1}} d$$

Specializing to a generalization by Hatcher [Hat2, Hat1] of a formula of Gross [Gro], we obtain

Theorem 1. *Let $-D < 0$ be an odd fundamental discriminant; let N be a prime which inert in $K = \mathbb{Q}(\sqrt{-D})$ and let $k \geq 2$ be an even* *integer. For Ψ a character of $\mathrm{Pic}(\mathcal{O}_K)$, then for any positive integer m, we have the following exact identity:*

$$(2)\qquad \frac{(2k-2)!D^{1/2}u^2}{2\pi(4\pi)^{2k-1}}\sum_{f\in\mathcal{F}_{2k}(N)}\frac{L(f,\Psi,k)}{\langle f,f\rangle}a_m(f) = -\delta\frac{12h^2}{N-1}\sigma_N(m)+um^{k-1}r_\Psi(m)h+u^2m^{k-1}\sum_{n=1}^{\frac{mD}{N}}\Phi_k(n,\Psi,N)$$

Here

$$\Phi_k(n,\Psi,N) = d((n,D))\delta_1(\Psi)R(n)r_\Psi(mD-nN)P_{k-1}(1-\frac{2nN}{mD}),$$

with P_{k-1} denoting the $(k-1)$-th Legendre polynomial; $\delta\in\{0,1\}$ is 1 iff $(k,\Psi)=(1,1_K)$; $\delta_1(\Psi)\in\{0,1\}$ is 1 if D is prime, and when D is composite, it is 1 iff $\Psi^2=1_K$ and there exist ideals $\mathfrak{a},\mathfrak{b}$, of respective norms $mD-nN$ and n, such that, for a prime ideal Q congruent to $-N$ mod D, the class of $\mathfrak{a}\mathfrak{b}Q$ is a square in $\mathrm{Pic}(\mathcal{O}_K)$.

An asymptotic formula involving the average on the left was first established for $k=1,\Psi=1_K$ by W. Duke ([Duk]), which spurred a lot of other work, including that of Iwaniec and Sarnak ([IS1]) relating it to the problem of Siegel zeros for $L(s,\chi_{-D})$. In the work of the second named author with J. Rogawski ([RR]), a different proof of Duke's result was given (for all weights), using Jacquet's relative trace formula involving the integration of the kernel over the square of the split torus, and in addition, the intervening measure was identified.

It is important to note that one obtains a *stability theorem* when N is sufficiently large compared with D and m, and this could perhaps be considered the most unexpected consequence of our approach. Indeed, when $N>mD$, the sum on the far right of the identity furnished by Theorem 1 becomes zero, and our exact average simplifies as follows:

Corollary 1. *(*Stability*) With the above notations and assumptions, suppose moreover $N>mD$, then one has*

$$\frac{(2k-2)!D^{1/2}u^2}{2\pi(4\pi)^{2k-1}}\sum_{f\in\mathcal{F}_{2k}(N)}\frac{L(f,\Psi,k)}{\langle f,f\rangle}a_m(f) = -\delta\frac{12h^2}{N-1}\sigma_N(m)+um^{k-1}r_\Psi(m)h$$

We call the range $N>mD$, the *stable range.* As one can check with other instances of the Gross/Zagier formulas, such as for the derivative in the case of odd order of vanishing, this phenomenon appears to be quite general. It has been recently generalized to Hilbert modular forms of square-free level by B. Feigon and D. Whitehouse ([FW]), using the relative trace formula, now by integrating the kernel over a non-split torus.

When $\Psi = 1_K$, we have the factorization

$$L(s, f, 1_K) = L(s, f_K) = L(s, f)L(s, f \otimes \chi_{-D}),$$

where f_K denotes the base change of f to K, $L(s, f)$ the Hecke L-function of f, and $f \otimes \chi_{-D}$ the twist of f by χ_{-D}. Thus for $m = 1$ and $N > D$, we get the following explicit identity involving the class number of K:

$$\frac{(2k-2)!D^{1/2}u}{2\pi(4\pi)^{2k-1}} \sum_{f \in \mathcal{F}_{2k}(N)} \frac{L(k, f)L(k, f \otimes \chi_{-D})}{\langle f, f \rangle} = h\big(1 - \delta\frac{12h}{u(N-1)}\big)$$

In the weight 2 case, as N is taken to be a prime here, the cardinality of $\mathcal{F}_2(N)$ is just the genus $g_0(N)$ of the compactification $X_0(N)$ of $Y_0(N)$. It is amusing to note that when $g_0(N)$ is zero, one finds that

$$h = \frac{(N-1)u}{12},$$

implying that $h = 1$ when $(-D, N)$ is $(-3, 5)$, $(-7, 13)$, $(-8, 13)$ or $(-11, 13)$, agreeing with known data. Similarly, $X_0(11)$ is an elliptic curve $E/\mathbb{Q}$, and if we denote by E_{-D} the $-D$-twist of E, we see, for $D = 3$, that the algebraic special value $A(1, E)A(1, E_{-3})$ is just $1/5$. In general one gets more complicated identities, involving average central values, which are all compatible with the Birch and Swinnerton-Dyer conjecture for E, E_{-D}, and the Shafarevich-Tate groups $\mathrm{Sh}(E)$, $\mathrm{Sh}(E_{-D})$.

0.1. **Application to the subconvexity problem.** We now discuss some simple applications of the above exact average formula, the first one being a subconvex estimate for the central values $L(k, f, \Psi)$. We refer to [IS2] for a general discussion on the subconvexity problem. In the present case the convexity bound is given by

$$L(k, f, \Psi) \ll_\varepsilon (kND)^\varepsilon kN^{1/2}D^{1/2}, \tag{3}$$

for any $\varepsilon > 0$. We prove here

Corollary 2. *Preserve the notations of Theorem 1. Then for any $\varepsilon > 0$, we have*

$$L(k, f, \Psi) \ll_\varepsilon (kDN)^\varepsilon kN^{1/2}D^{1/2}\big(\frac{1}{N^{1/2}} + \frac{N^{1/2}}{D^{1/2}}\big).$$

In particular this improves on convexity as long as

$$(kD)^{\delta} \leq N \leq D(kD)^{-\delta}$$

for some fixed $\delta > 0$.

Note that this breaks convexity for any fixed k, as long as N is between D^{δ} and $D^{1-\delta}$. The beauty is that we can also vary k in an appropriate region, obtaining a *hybrid subconvexity*.

At this point we do not know of any application to these subconvex estimates, but we are intrigued by them because they come for free and seem to be hard to prove with the current methods of analytic number theory (eg. see [DFI, KMV]). Note also that such bounds are fundamentally limited to the critical center $s = k$. For a generalization to the Hilbert modular case, where Ψ is allowed to be any ray class character, see [FW].

0.2. **Application to non-vanishing problems.** Another line of application addresses the existence of f for which $L(k, f, \Psi)$ does not vanish. Indeed several variants of such problems have been considered in the past by various methods [Duk, IS1, KM, OS, Vat2]. Here we obtain non-vanishing results which are valid with a fairly large uniformity in the parameters, and again such uniformity seems hard to achieve by purely analytic methods.

Theorem 2. *Assumptions being as for Theorem A. Suppose that*

$$N \gg_{\delta} D^{1/2+\delta}$$

for some $\delta > 0$*, then there exists* $f \in S_{2k}^{new}(N)$ *such that*

$$L(k, f, \Psi) \neq 0.$$

The same conclusion holds as long as $N > D$ *and either* $k \neq 1$ *or* $\Psi \neq 1_K$.

When $\Psi = 1_K$, we also obtain non-vanishing result in a somewhat greater range:

Theorem 3. *Suppose* $\Psi = 1_K$, $k = 1$ *and*

$$h < \frac{N-1}{12}.$$

Then there exist f *such that*

$$L(k, f)L(k, f \otimes \chi_{-D}) \neq 0.$$

Non-vanishing theorems of this kind, with an *explicit* dependence between N and D (like $N > D$ or $N - 1 > 12h$), are of some interest. For instance, in the paper [Mer], Merel needs to consider the following problem: Given a prime p and a character χ of conductor p which is not even and quadratic, does there exist an $f \in \mathcal{F}_2(p)$ such that $L(1, f \otimes \chi) \neq 0$? In the appendix of that paper, the first named author and E. Kowalski prove that this is the case when p is greater than an explicit but very large number. In particular, it has so far not been possible to answer the problem numerically in the finitely many remaining cases; this has been answered however for $p < 1000$ [MS]. Closer to the main concern of the present paper, Ellenberg [Ell1, Ell2] uses analytic methods to prove the non-vanishing of the twisted L-function $L(1, f \otimes \chi_{-4})$ for some f in $\mathcal{F}_2(N)$ for N of the form p^2 or $2p^2$ (p an odd prime) and with prescribed eigenvalues at the Atkin/Lehner operators w_2, w_p, subject to an *explicit* lower bound on p. Ellenberg concludes from this the non-existence of primitive integral solutions to the generalized Fermat equation $A^4 + B^2 = C^p$ as long as $p > 211$; that this equation has only a finite number of primitive solutions is a theorem of Darmon and Granville. Another related set of examples is in the work of Dieulefait and Urroz ([DU]). In a sequel to this paper under preparation ([MR]), we will develop a suitable generalization of the exact average formula to a class of composite levels N, and investigate similar questions by modifying the method. This extension is subtle for three reasons: N is not square-free, D is not odd, and N, D are not relatively prime.

0.3. **Nonvanishing modulo p.** The exactness of the Gross/Zagier formulae even enable us to obtain *average non-vanishing results* for the *algebraic part* of the $L(k, f, \Psi)$ modulo suitable primes p. Again, such a question has been considered in the past, see for example [BJK$^+$, Vat2]. However, these earlier works addressed the question of the existence of the non-vanishing of $L(k, f, \Psi)$ mod p when the form f is *fixed* and when the character Ψ varies. Here our results go in the other direction as we fix p and let N and f vary. Given $f \in \mathcal{F}_{2k}(N)$ and g_Ψ as above, we denote by $L^{\mathrm{alg}}(k, f, \Psi)$ the algebraic part of $L(k, f, \Psi)$ (see section 5, (11), for a precise definition). It follows from the work of Shimura that $L^{\mathrm{alg}}(k, f, \Psi)$ is an algebraic number satisfying the reciprocity law

$$L^{\mathrm{alg}}(k, f, \Psi)^\sigma = L^{\mathrm{alg}}(k, f^\sigma, \Psi^\sigma)$$

for any σ automorphism of $\mathbb{C}$ [Shi].

Theorem 4. *Let $p > 2k + 1$ be a prime, $\mathcal{P}$ be a chosen place in $\overline{\mathbb{Q}}$ above p and let N, D be as in Theorem 1. Suppose moreover that p*

does not divide $h = h_{-D}$, that $N > D$, and that N is greater that some absolute constant. Then there exists $f \in \mathcal{F}_{2k}(N)$ such that

$$L^{\mathrm{alg}}(k, f, \Psi) \not\equiv 0 \ (\mathrm{mod}\ \mathcal{P}).$$

Naturally, the question of integrality of $L^{\mathrm{alg}}(k, f, \Psi)$, which is subtle, and our result only concerns the numerator of the L-value. When $\Psi = 1_K$, we also prove the following variant:

Theorem 5. *Notations and assumptions as in Theorem 4. Suppose moreover that $\Psi = 1$ and $N > pD$. Then there exists $f \in \mathcal{F}_{2k}(N)$ such that*

$$\sqrt{D}(2\pi)^{-2k}\frac{L(k, f)L(k, f \otimes \chi_{-D})}{\langle f, f\rangle}a_p(f) \not\equiv 0 \ (\mathrm{mod}\ \mathcal{P}^{2k-1}).$$

The assertion makes sense because the left hand side is (see section 5.1) a p-unit times $a_p(f)$ times $L^{\mathrm{alg}}(k, f, 1_K)$.

There are two fundamental periods $c^+(f)$ and $c^-(f)$ associated to f such that for any Dirichlet character ν, the special value $L^{\mathrm{alg}}(k, f \otimes \nu)$, defined as $L(k, f \otimes \nu)/c^{\mathrm{sgn}(\nu(-1))}(f)$ times a simple factor (see section 5, (12)) is an algebraic number. One gets the near-factorization

$$\eta_f L^{\mathrm{alg}}(k, f, 1_K) \ = \ L^{\mathrm{alg}}(k, f)L^{\mathrm{alg}}(k, f \otimes \chi_{-D}),$$

where η_f is essentially the order of the congruence module considered by Hida, Wiles, Taylor, Flach, Diamond, and others, which measures the congruences f has with other modular forms modulo p. The needed non-divisibility properties of η_f (for suitable p) are understood (at least) if f is ordinary or $k = 1$. Now finally, let us suppose we are in the classical weight 2 situation, i.e., with $\Psi = 1_K$ and $k = 1$.

Theorem 6. *Let p an odd prime not dividing Dh_{-D}, with D odd. Then there exist infinitely many newforms of f of prime level N and weight 2 such that*

$$\mathrm{num}\left(\frac{L^{\mathrm{alg}}(1, f \otimes \chi_{-D})}{\eta_f}\right) \not\equiv 0 \pmod{p},$$

where η_f is the order of the congruence module of f.

See section 5 for a discussion of η_f, which measures the congruences which f may have with other modular forms of the same weight and level. An analogue of Theorem 6 should also hold, in a suitable range of p, for forms of higher weight, and this question will be taken up elsewhere.

0.4. **Acknowledgement.** Serge Lang always conveyed infectious excitement about Mathematics to anyone he came into contact with, and he will be missed. He was quite interested in the values of L-functions and in the *divisibility properties* of arithmetic invariants, and it is a pleasure to dedicate this article to him. The first author would like to thank Caltech for its hospitality during the preparation of this work. The second author would like to thank Flach, Hida, Prasanna and Vatsal for helpful conversations concerning the last part, and the National Science Foundation for support through the grant DMS0402044.

1. The weight 2 case

It may be instructive to explain why the exact average formula holds in the weight 2 case when $\Psi = 1$. Let B be a quaternion division algebra over $\mathbb{Q}$, ramified only at N and ∞, with maximal order R. Put Y is the associated rational curve such that $\mathrm{Aut}(Y) = B^*/\mathbb{Q}^*$. Put

$$X = B^*\backslash Y \times \hat{B}^*/\hat{R}^* = \cup_{j=1}^n \Gamma_j\backslash Y,$$

where $\hat{B}^* = \prod_p{}' B_p^*$ and $\hat{R}^* = \prod_p R_p^*$, with each Γ_j being a finite group. Then $\mathrm{Pic}(X)$ identifies with $\{e_1, e_2, \ldots, e_n\}$, where each e_j is the class of $\Gamma_j\backslash Y$. Since N is inert in $K = \mathbb{Q}[\sqrt{-D}]$, there is an embedding $f \in \mathrm{Hom}(K, B) = Y(K)$. It results in certain *Heegner points* $x = (f, b)$ of discriminant $-D$ in X, with $b \in \hat{B}^*/\hat{R}^*$. For any eigenform f, let c_f denote the f-component of $c = \sum_A x_A$, where A runs over ideal classes of K. Then by a beautiful theorem of B. Gross ([G]), providing an analogue for the L-value of the Gross-Zagier theorem for the first derivative, one has

$$\langle c_f, c_f\rangle \;=\; u^2\sqrt{D}\frac{L(1,f)L(1,f\otimes\chi_{-D})}{(f,f)},$$

where $\langle\cdot,\cdot\rangle$ is a natural *height pairing* on $\mathrm{Pic}(x)$. We have by orthogonality,

$$\langle c, T_m c\rangle = \langle c_E, T_m c_E\rangle + \sum_f \langle c_f, T_m c_f\rangle,$$

where T_m is the operator corresponding to the m-the Hecke operator on $M_2(N)$, f runs over newforms in $M_2(N)$, and E denotes the unique (holomorphic) Eisenstein series (of weight 2 and level N). Using the fact that f and E are Hecke eigenforms, and that $\langle c_E, c_E\rangle = \frac{12h^2}{N-1}$, we get by averaging Gross's formula,

$$u^2\sqrt{|D|}\sum_f \frac{L(1,f)L(1,f\otimes\chi_{-D})}{(f,f)} = -\sigma_N(m)\frac{12h^2}{N-1} + \langle c, T_m c\rangle.$$

One has

$$\langle c, T_m c\rangle = \sum_A \sum_B \langle x_B, T_m x_{AB}\rangle.$$

If we pick $q \equiv -N (\mathrm{mod}\ D)$, with $q\mathcal{O}_K = Q\overline{Q}$ in K, one sees that

$$\sum_B \langle x_B, T_m x_{AB}\rangle = uhR_A(m) + \sum_{n=1}^{mD/N} R_A(mD - nN) d((n, D)) R_{\{QA\}}(n),$$

with

$$R_{\{QA\}}(n) = |\{I : N(I) = n, QAI \in \mathrm{Pic}(\mathcal{O}_K)^2\}|.$$

Note that $R_{\{QA\}}(n)$ is just $R_A(n)$ when D is prime. The assertion of Theorem 1 now follows by summing over A. Moreover, when mD is less than N, $\sum\limits_B \langle x_B, T_m x_{AB}\rangle$ simply equals $uhR_A(m)$, and this furnishes Corollary 1 (stability) in the weight 2 case.

2. **Proof of the main identity for all** $k \geq 1$

2.1. **Preliminaries.** For $N \geq 1$, let $M_{2k}(N)$ (resp $S_{2k}(N)$) denote, as usual, the space of holomorphic modular forms (resp. cusp forms) of weight $2k$ level N and trivial character. For $f \in M_{2k}(N)$, we write the Fourier expansion at the infinite cusp as

$$f(z) = \sum_{m \geq 0} a_m(f) q^m, q = \exp(2\pi \imath z).$$

We denote by $\mathcal{F}_{2k}(N)$, the set of cuspidal new forms f (normalized in the usual way, so that the first Fourier coefficient $a_1(f)$ is 1. Whenever it converges, we denote the Petersson inner product on $M_{2k}(N)$ by

$$\langle f, g\rangle = \int_{Y_0(N)} f(z)\overline{g(z)} y^{2k-2} dx dy.$$

Let $-D < 0$ be an odd fundamental discriminant, $K = \mathbb{Q}(\sqrt{-D}), \mathcal{O}_k$ the maximal order of $K, \mathrm{Pic}(\mathcal{O}_K)$ the ideal class group, and $u = u_k = |\mathcal{O}_K{}^\times|/2$. For any ideal class $A \in \mathrm{Pic}(\mathcal{O}_k)$, define

$$r_A(m) = \begin{cases} |\{\mathfrak{a} \subset \mathcal{O}_K, N(\mathfrak{a}) = m, \mathfrak{a} \in A\}| & \text{if } m \geq 1 \\ \frac{1}{2u} & \text{if } m = 0 \end{cases}$$

The theta series

$$\theta_A(z) = \sum_{m \geq 0} r_A(m) q^m, q = \exp(2\pi \imath z)$$

is a modular form of weight 1, level D and central character χ_{-D}. Moreover, for any $\Psi \in \widehat{\mathrm{Pic}(\mathcal{O}_K)}$, put

$$\theta_\Psi(z) = \sum_A \overline{\Psi}(A)\theta_A(z),$$

whose Fourier coefficients are then given by

$$a_m(\theta_\Psi) = \sum_A \overline{\Psi}(A)a_m(\theta_A).$$

In particular, the constant term $a_0(\theta_\Psi)$ equals $\frac{1}{2u}\sum_A \overline{\Psi}(A)$, which is, by orthogonality, zero iff $\Psi \neq 1_K$, when θ_Ψ is a cusp form. Setting

$$L(s,f,A) := \sum_{\substack{n>1\\(n,N)=1}} \frac{\chi_{-D}(n)}{n^{1+1(s-k)}} \sum_{m\geq 1} \frac{a_m(f)r_A(m)}{m^s},$$

one has

$$L(s,f,\Psi) = \sum_{A\in \mathrm{Pic}(\mathcal{O}_K)} \Psi(A)L(s,f,A).$$

Define a holomorphic function G_A on the upper half plane $\mathcal{H}$, invariant under $z \longrightarrow z+1$, by means of its Fourier expansion at infinity:

$$G_A(z) := \sum_{m=0}^{\infty} b_{m,A}q^m, \tag{2}$$

where

$$\begin{aligned} b_{m,A} &= m^{k-1}\frac{h}{u}r_A(Dm) \\ &+ m^{k-1}\sum_{n=1}^{mD/N} \delta(n)r_A(mD-nN)R_{(-NA)}(n)P_{k-1}\left(1-\frac{2nN}{mD}\right) \end{aligned} \tag{3}$$

In this definition, u and $R(n) = \sum_A r_A(n)$ are as in the Introduction, $\delta(n)$ is 1 (resp. 2) if (m,D) is 1 (resp. $\neq 1$), and for $r \geq 0, P_r$ is the r-th Legendre polynomial defined by

$$P_r(x) := \frac{1}{2^r}\sum_{m=1}^{[r/2]}(-1)^m \binom{r}{m}\binom{2r-2m}{r} x^{r-2m}.$$

The following result, due to B. Gross, D. Zagier and R. Hatcher, is crucial to us:

Theorem 7. *G_A is a modular form of weight $2k$, level N, and trivial character; it is cuspidal if $k > 1$, and for every newform f of weight $2k$ and level N, we have*

$$L(k, f, A) = \frac{(4\pi)^{2k}}{2(2k-2)!D^{1/2}}(f, G_A).$$

For $k = 1$, see [11], Prop. 9.1, and for general k, this is in [12], Theorem 5.6 and [14], Theorem 5.2. (See also [13], where the case D prime is treated.)

2.2. **The exact average formula.** Let

$$E \;=\; E_{2,N} \;=\; \sum_{n=0}^{\infty} a_n(E)q^n$$

denote a holomorphic Eisenstein series for $\Gamma_0(N)$ of weight 2. Since N is prime, the modular curve $Y_0(N)$ has only two cusps, namely ∞ and 0. It then follows that E is unique up to scalar multiple, and so $E(z)/a_0(E)$ is well defined with constant term 1 at ∞. To be specific, we will take

$$E(z) = \frac{N-1}{12} + \sum_{m=1}^{\infty} \sigma_N(m)q^m,$$

where $\sigma_N(m) = \sum_{d|m,(d,N)=1} d$.

For $A \in \mathrm{Pic}(\mathcal{O}_K)$, with G_A being as in the previous section, put

$$G_A^{\mathrm{cusp}}(z) : G_A(z) - \delta_{k=1}\frac{b_{0,A}}{a_0(E)}E(z), \tag{4}$$

with $\delta_{k,1}$ being 1 (resp. 0) if $k = 1$ (resp. $k \neq 1$). Then G_A^{cusp} is a holomorphic cusp form of level N, weight $2k$, and trivial character, with coefficients $a_m(G_A^{\mathrm{cusp}})$.

Lemma 2.1. *For $-D$ an odd fundamental discriminant and N a prime inert in K, we have, for any $m \geq 1$,*

$$\frac{2(2k-2)!D^{1/2}}{(4\pi)^{2k}} \sum_{f \in \mathcal{F}_{2k}(N)} \frac{L(k, f, A)}{\langle f, f\rangle} a_m(f)$$
$$= a_m(G_A^{\mathrm{cusp}}) = b_{m,A} - \delta_{k=1}\frac{b_{0,A}}{a_0(E)}a_m(E)$$

In order to prove this, we first need the following

Lemma 2.2. *Assume that N is a prime which is inert in $K = \mathbb{Q}[\sqrt{-D}]$. Let φ be any old form in $S_{2k}(N)$. Then we have, for every $A \in \mathrm{Pic}(\mathcal{O}_K)$,*

$$(\varphi, G_A^{\mathrm{cusp}}) = 0.$$

There is nothing to prove when $k < 6$, since $S_{2k}(1)$ is zero in that case (cf. [Lan], for example.) Such a Lemma will not in general hold for composite N.

Proof of Lemma 2.2. Since φ is cuspidal, it suffices to prove that $(\varphi, G_A) = 0$. Put

$$G_\Psi := \sum_{A \in \mathrm{Pic}(\mathcal{O}_K)} \Psi(A) G_A$$

which is modular form of weight 1 and character χ_{-D}. It is sufficient to show that $(\varphi, G_\Psi) = 0$ for all ideal class characters Ψ of K. If $\varphi = \sum_{n=1}^{\infty} a_n(\varphi) q^n$, put

$$D(s, \varphi \times \theta_\Psi) = \sum_{n=1}^{\infty} \frac{a_n(\varphi)\overline{a}_n(\theta_\Psi)}{n^s} \tag{5}$$

Then the Rankin-Selberg method give the identity

$$(2\pi)^{-k}\Gamma(k) D(k, \varphi \times \theta_\Phi) = \langle f, \mathrm{Tr}_{ND/N}(\theta_\Phi \mathcal{E}_{2k-1,N})\rangle \tag{6}$$

where $\mathcal{E}_{2k-1,N}$ is the result of slashing a holomorphic Eisentein series of weight $2k-1$ (and character χ_{-D}) with the Atkin involution u_N, and $\mathrm{Tr}_{ND/D}$ denotes the trace from $S_{2k}(ND)$ to $S_{2k}(N)$. In fact, the calculations of Gross and Zagier ([GZ]) show that

$$G_\Psi = \mathrm{Tr}_{ND/N}(\theta_\Psi \mathcal{E}_{2k-1,N}). \tag{7}$$

Now let φ be a newform of level 1 (and weight $2k$). Then since N is prime, it defines two old forms of level N, namely $\varphi_1(z) = \varphi(z)$ and $\varphi_2(z) = \varphi(Nz)$, so that $a_m(\varphi_2)$ is zero unless $N|m$, and $a_m N(\varphi_2) = a_m(\varphi)$. Since the new and old forms are orthogonal to each other under $(\cdot,\cdot)$, and since the space of old forms of level N are spanned by $\{\varphi_d, d = 1, N\}$ with φ running overl all the cusp forms of level 1, it suffices to prove that each $D(k, \varphi_d \times \theta_\Psi) = 0$. Let $d = 1$. Then one obtains (by section 3, Lemma 1, of [Sh]):

$$L(2k, \chi_{-D}) D(k, \varphi_d \times \theta_\Psi) = L(k, \varphi \times \theta_\Psi). \tag{8}$$

Since $L(x, \chi_{-D})$ is non-zero at $s = 2k$ (which is in the region of absolute convergence), it reduces to checking the vanishing of the right hand side. Since φ has level 1, the root number of $L(k, \varphi \times \theta_\Psi)$ is -1, yielding the requisite vanishing. When $d = N, D(k, \varphi_d \times \theta_\Psi)$ is still a non-zero multiple of $L(k, \varphi \times \theta_\Psi)$, which is zero.

□

Proof of Lemma 2.1 We may choose an orthogonal basis $\mathcal{B}$ of $S_{2k}(N)$ to be of the form $\mathcal{F}_{2k}(N) \cup \mathcal{B}'$, where $\mathcal{B}'$ consists of old forms. Clearly we have

$$\sum_{f \in \mathcal{B}} \frac{(f, G_A^{\text{cusp}})}{\langle f, f \rangle} = G_A^{\text{cusp}}. \tag{9}$$

In view of the Lemma, the sum on the left hand side needs to run only over newforms f. Applying Theorem 6, and using (8), we obtain

$$\frac{2(2k-2)!D^{1/2}}{(4\pi)^{2k}} \sum_{f \in \mathcal{F}_{2k}(N)} \frac{L(f, \Psi, k)}{\langle f, f \rangle} = G_A^{\text{cusp}}.$$

The lemma now follows by taking the m-coefficient of the above identity.

□

Proof of Theorem 1 The exact average formula follows by performing the averaging $\sum_{A \in \text{Pic}(\mathcal{O}_K)} \Psi(A) \dots$ on both sides of the formula in Lemma 2.1 using the formula (5) for the coefficients $b_{m.A}$, and by noting that

$$\frac{a_m(E)}{a_0(E)} = \frac{12}{N-1} \sigma_N(m)$$

and that $b_{0,A} = \frac{h}{2u^2}$.

□

3. Subconvex Bounds

In this section, we prove Corollary 2. By the work of Waldspurger, Guo and Jacquet ([Guo, Wal]; also [KZ] for $\Psi = 1_K$),

$$L(k, f, \Psi) \geq 0.$$

Thus from formula (2) for $m = 1$, we have

$$\frac{(2k-2)!D^{1/2}}{2(4\pi)^{2k}} \frac{L(f, \Psi, k)}{\langle f, f \rangle} \leq \frac{h}{u} + \sum_{n=1}^{\frac{D}{N}} |\Psi_k(n, \Psi, N)|$$

Since $|P_{k-1}(x)| \leq 1$ for $|x| \leq 1$ and $R(n), |r_\Psi(n)| \leq d(n)$ (where $d(n)$ is the number of divisors of n), so that

$$R(n)|r_\Psi(D - nN)| \leq d(n)^2 + d(D - nN)^2,$$

we see that the n-sum on the right side is bounded by $\frac{D}{N}(\log D)^3$. From the class number formula, we have

$$h \ll D^{1/2} \log D$$

and

$$\langle f, f\rangle \ll (4\pi)^{-2k}(2k-1)!N(\log kN)^3$$

as follows from [ILS], (2.3), (unlike the corresponding bound for Maass forms ([HL]) this upper bound is elementary since f holomorphic so its Fourier coefficients satisfy the Ramanujan—Petersson bound). Thus we see that

$$L(f, \Psi, k) \ll (\log kN)^3(\log D)^3 k(N + D^{1/2}).$$

□

4. **Application to non-vanishing**

We prove here Theorem 2. Arguing exactly as above we have

$$\begin{aligned}\frac{(2k-2)!D^{1/2}}{2(4\pi)^{2k}} \sum_{f\in\mathcal{F}_{2k}(N)} \frac{L(f,\Psi,k)}{\langle f,f\rangle} &= \frac{h}{u} - \delta\frac{6(h/u)^2}{N-1} + O\left(\frac{D}{N}(\log D)^3\right)\\ &= \frac{h}{u} + O\left(\frac{D}{N}(\log D)^3\right)\end{aligned}$$

By Siegel's Theorem, which gives $h = D^{1/2+o(1)}$, we see that the right side is positive as soon as $N > D^{1/2+\delta}$ for some $\delta > 0$. If $N > D$, then we are in the stable range and we have

$$\frac{(2k-2)!D^{1/2}}{2(4\pi)^{2k}} \sum_{f\in\mathcal{F}_{2k}(N)} \frac{L(f,\Psi,k)}{\langle f,f\rangle} = \frac{h}{u}\left(1 - \delta\frac{6(h/u)}{N-1}\right). \tag{10}$$

When $\delta = 0$, this concludes the proof of Theorem 2 since $h \geq 1$. □

Suppose now that $\delta = 1$ (ie. $k = 1, \Psi = 1_K$). Then we remark that

$$\sum_{n=1}^{\frac{D}{N}} \Psi_1(n, 1, N) \geq 0$$

so that

$$\frac{(2k-2)!D^{1/2}}{2(4\pi)^{2k}} \sum_{f\in\mathcal{F}_{2k}(N)} \frac{L(f,\Psi,k)}{\langle f,f\rangle} \geq \frac{h}{u}\left(1 - \frac{6(h/u)}{N-1}\right)$$

combining the proof of Theorem 3. □

5. **Non-vanishing mod** p

5.1. **Algebraic Parts of L-values.** Let us put

$$L^{\mathrm{alg}}(k,f,\Psi) \;=\; (-1)^k(2\pi)^{-2k}(k-1)!^2 g(\chi_D)\frac{L(k,f,\Psi)}{\langle f,f\rangle}, \tag{11}$$

where $g(\overline{\Psi})$ is the Gauss sum. Then it is known, by Shimura ([Shi], see also [Hid1]), that $L^{\mathrm{alg}}(k,f,\psi)$ is an algebraic number obeying the reciprocity law:

$$L^{\mathrm{alg}}(k,f^\sigma,\Psi^\sigma) = L^{\mathrm{alg}}(k,f,\Psi)^\sigma,$$

for every automorphism σ of $\mathbb{C}$.

Next recall that for $\Psi = 1_K$, $L(k,f,\Psi)$ factors as $L(k,f)L(k,f\otimes\chi_{-D})$. For any Dirichlet character ν, the algebraic part of $L(k,f\otimes\nu)$ is given by

$$L^{\mathrm{alg}}(k,f\otimes\nu) \;=\; g(\overline{\nu})(k-1)!\frac{L(k,f,\nu)}{(-2\pi i)^k c_\pm(f)}, \tag{12}$$

where $c_\pm(f)$ is a fundamental period of f, with $\pm = \nu(-1)$. Again, one has for any automorphism σ of $\mathbb{C}$, $L^{\mathrm{alg}}(k,f^\sigma\otimes\nu^\sigma)$ is $L^{\mathrm{alg}}(k,f\otimes\nu)^\sigma$.

This leads to the near-factorization

$$\eta_f L^{\mathrm{alg}}(k,f,1_K) \;=\; L^{\mathrm{alg}}(k,f)L^{\mathrm{alg}}(k,f\otimes\chi_{-D}), \tag{13}$$

where η_f equals, thanks to a series of papers of Hida (cf. [Hid1], [Hid2]), Wiles ([Wil]), Taylor-Wiles ([TW]), and Diamond-Flach-Guo ([DFG]), the order of the congruence module of f, i.,e the number which counts the congruences of f with other modular forms of the same weight and level.

5.2. **Proof of Theorems 4 and 5.** From the definition of the algebraic part, the hypothesis of Theorem 4 and the formula (9), used in conjunction with $\delta = 0$, we have (up to multiplication by a p-unit)

$$\sum_{f\in\mathcal{F}_{2k}(N)} L^{\mathrm{alg}}(k,f,\Psi) = \frac{h}{u}.$$

The conclusion of Theorem 4 is immediate.

For the proof of Theorem 5, we have, assuming that $N > pD$,

$$\sum_{f\in\mathcal{F}_{2k}(N)} L^{\mathrm{alg}}(k,f,1_K) = \frac{h}{u}\left(1-\frac{12(h/u)}{N-1}\right).$$

Therefore the conclusion holds except possibly if $p|(1-\frac{6(h/u)}{N-1})$. Suppose we are in that latter case. Then we apply the exact formula of Corollary

1 with $m = p$ and get

$$\sum_{f\in\mathcal{F}_{2k}(N)} L^{\mathrm{alg}}(k,f,1_K)a_p(f) = \frac{h}{u}\left(R(p) - \frac{6(h/u)}{N-1}(p+1)\right)$$

$R(p)$ is either 0 or 2, if it is zero, then the left hand side of the previous formula is not divisible by p. If $R(p) = 2$, then $2 - \frac{6(h/u)}{N-1}$ is not divisible by p since by assumption $p|(1 - \frac{6(h/u)}{N-1})$. So we are done in all cases. □

5.3. **Proof of Theorem 6.** Here are restricting to the weight 2 case, and by the theory of modular symbols, cf. Stevens [Ste] and Vatsal [Vat1] - see also Prasanna [Pra] - we know that for any Dirichlet character ν, the special value $L^{\mathrm{alg}}(1, f \otimes \nu)$ is integral except possibly at the Eisenstein primes; these are the primes dividing

$$\tilde{N} := \prod_{q|N} q(q^2 - 1),$$

which is related to the order of the cuspidal divisor class group, studied for modular curves, among others, by Kubert and Lang.

We may, and we will, choose N to lie in the infinite family of primes which are inert in K and are such that $p \nmid \tilde{N}$.

Now Theorem 6 follows by the near-factorization (13) of $L^{\mathrm{alg}}(1, f, 1_K)$. It may be useful to note that when f has $\mathbb{Q}$-coefficients, with associated elliptic curve E over $\mathbb{Q}$, one knows (cf. Flach [Fla]) that any prime dividing η_F also divides the degree of the modular parametrization $X_0(N) \to E$.

References

[BJK+] J. H. Bruinier, K. James, W. Kohnen, K. Ono, C. Skinner, and V. Vatsal. Congruence properties of values of L-functions and applications. In *Topics in number theory (University Park, PA, 1997)*, volume 467 of *Math. Appl.*, pages 115–125. Kluwer Acad. Publ., Dordrecht, 1999.

[DZ] H. Darmon and S.-W. Zhang, editors. *Heegner points and Rankin L-series*, volume 49 of *Mathematical Sciences Research Institute Publications*. Cambridge University Press, Cambridge, 2004. Papers from the Workshop on Special Values of Rankin L-Series held in Berkeley, CA, December 2001.

[DFG] F. Diamond, M. Flach, and L. Guo. The Bloch-Kato conjecture for adjoint motives of modular forms. *Math. Res. Lett.* **8** (2001), 437–442.

[DU] L. Dieulefait and J. Urroz. Solving Fermat-type equations $x^4 + dy^2 = z^p$ via modular Q-curves over polyquadratic fields, Preprint, 2006.

[Duk] W. Duke. The critical order of vanishing of automorphic L-functions with large level. *Invent. Math.* **119** (1995), 165–174.

[DFI] W. Duke, J. Friedlander, and H. Iwaniec. Bounds for automorphic L-functions. II. *Invent. Math.* **115** (1994), 219– 239.

[Ell1] J. S. Ellenberg. Galois representations attached to $\mathbb{Q}$-curves and the generalized Fermat equation $A^4 + B^2 = C^p$. *Amer. J. Math.* **126** (2004), 763–787.

[Ell2] J. S. Ellenberg. On the error term in Duke's estimate for the average special value of L-functions. *Canad. Math. Bull.* **48** (2005), 535–546.

[FW] B. Feigon and D. Whitehouse. Averages of central L-values of Hilbert modular forms with an application to subconvexity, Preprint. 2007.

[Fla] M. Flach. On the degree of modular parametrizations. In *Séminaire de Théorie des Nombres, Paris, 1991–92*, volume 116 of *Progr. Math.*, pages 23–36. Birkhäuser Boston, Boston, MA, 1993.

[GZ] B. Gross and D. Zagier. Heegner points and derivatives of L-series. *Invent. Math.* **84** (1986), 225–320.

[Gro] B. H. Gross. Heights and the special values of L-series. (1987), 115–187.

[Guo] J. Guo. On the positivity of the central critical values of automorphic L-functions for GL(2). *Duke Math. J.* **83** (1996), 157–190.

[Hat1] R. L. Hatcher. Special values of L-series. *Proc. Amer. Math. Soc.* **114** (1992), 337–343.

[Hat2] R. L. Hatcher. Heights and L-series. *Canad. J. Math.* **42** (1990), 533–560.

[Hid1] H. Hida. A p-adic measure attached to the zeta functions associated with two elliptic modular forms. I. *Invent. Math.* **79** (1985), 159–195.

[Hid2] H. Hida. On the search of genuine p-adic modular L-functions for $\mathrm{GL}(n)$. *Mém. Soc. Math. Fr. (N.S.)* (1996), vi+110. With a correction to: "On p-adic L-functions of $\mathrm{GL}(2) \times \mathrm{GL}(2)$ over totally real fields" [Ann. Inst. Fourier (Grenoble) **41** (1991), no. 2, 311–391; MR1137290 (93b:11052)].

[HL] J. Hoffstein and P. Lockhart. Coefficients of Maass forms and the Siegel zero. *Ann. of Math. (2)* **140** (1994), 161–181. With an appendix by D. Goldfeld, J. Hoffstein and D. Lieman.

[IS1] 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.

[IS2] H. Iwaniec and P. Sarnak. Perspectives on the analytic theory of L-functions. *Geom. Funct. Anal.* (2000), 705–741. GAFA 2000 (Tel Aviv, 1999).

[ILS] H. Iwaniec, W. Luo, and P. Sarnak. Low lying zeros of families of L-functions. *Inst. Hautes Études Sci. Publ. Math.* (2000), 55–131 (2001).

[KZ] W. Kohnen and D. Zagier. Values of L-series of modular forms at the center of the critical strip. *Invent. Math.* **64** (1981), 175–198.

[KM] E. Kowalski and P. Michel. The analytic rank of $J_0(q)$ and zeros of automorphic L-functions. *Duke Math. J.* **100** (1999), 503–542.

[KMV] E. Kowalski, P. Michel, and J. VanderKam. Rankin-Selberg L-functions in the level aspect. *Duke Math. J.* **114** (2002), 123–191.

[Lan] S. Lang. *Introduction to modular forms*, volume 222 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1995. With appendixes by D. Zagier and Walter Feit, Corrected reprint of the 1976 original.

[MS] L. Merel and W. Stein. The field generated by the points of small prime order on an elliptic curve. *Internat. Math. Res. Notices* (2001), 1075–1082.

[Mer] L. Merel. Sur la nature non-cyclotomique des points d'ordre fini des courbes elliptiques. *Duke Math. J.* **110** (2001), 81–119. With an appendix by E. Kowalski and P. Michel.

[MR] P. Michel and D. Ramakrishnan. Exact averages of central L-values, class numbers, and a diophantine application, In preparation. 2007.

[OS] K. Ono and C. Skinner. Fourier coefficients of half-integral weight modular forms modulo l. *Ann. of Math. (2)* **147** (1998), 453–470.

[Pra] K. Prasanna. Arithmetic properties of the theta correspondence and periods of modular forms, preprint, 2006.

[RR] D. Ramakrishnan and J. Rogawski. Average values of modular L-series via the relative trace formula. *Pure Appl. Math. Q.* **1** (2005), 701–735.

[Shi] G. Shimura. The special values of the zeta functions associated with cusp forms. *Comm. Pure Appl. Math.* **29** (1976), 783–804.

[Ste] G. Stevens. The cuspidal group and special values of L-functions. *Trans. Amer. Math. Soc.* **291** (1985), 519–550.

[TW] R. Taylor and A. Wiles. Ring-theoretic properties of certain Hecke algebras. *Ann. of Math. (2)* **141** (1995), 553–572.

[Vat1] V. Vatsal. Canonical periods and congruence formulae. *Duke Math. J.* **98** (1999), 397–419.

[Vat2] V. Vatsal. Special values of anticyclotomic L-functions. *Duke Math. J.* **116** (2003), 219–261.

[Wal] J.-L. Waldspurger. Sur les valeurs de certaines fonctions L automorphes en leur centre de symétrie. *Compositio Math.* **54** (1985), 173–242.

[Wil] A. Wiles. Modular elliptic curves and Fermat's last theorem. *Ann. of Math. (2)* **141** (1995), 443–551.

[Zha] S. Zhang. Heights of Heegner cycles and derivatives of L-series. *Invent. Math.* **130** (1997), 99–152.